Wavelet Analysis of Variability, Teleconnectivity, and Predictability of the September–November East African Rainfall

DAVISON MWALE AND THIAN YEW GAN

Department of Civil and Environmental Engineering, University of Alberta, Edmonton, Alberta, Canada

(Manuscript received 2 October 2003, in final form 10 July 2004)

ABSTRACT

By applying wavelet analysis and wavelet principal component analysis (WPCA) to individual wavelet-scale power and scale-averaged wavelet power, homogeneous zones of rainfall variability and predictability were objectively identified for September–November (SON) rainfall in East Africa (EA). Teleconnections between the SON rainfall and the Indian Ocean and South Atlantic Ocean sea surface temperatures (SST) were also established for the period 1950–97. Excluding the Great Rift Valley, located along the western boundaries of Tanzania and Uganda, and Mount Kilimanjaro in northeastern Tanzania, EA was found to exhibit a single leading mode of spatial and temporal variability. WPCA revealed that EA suffered a consistent decrease in the SON rainfall from 1962 to 1997, resulting in 12 droughts between 1965 and 1997. Using SST predictors identified in the April–June season from the Indian and South Atlantic Oceans, the prediction skill achieved for the SON (one-season lead time) season by the nonlinear model known as artificial neural network calibrated by a genetic algorithm (ANN-GA) was high [Pearson correlation $r$ ranged between 0.65 and 0.9, Hansen–Kuipers (HK) scores ranged between 0.2 and 0.8, and root-mean-square errors (rmse) ranged between 0.4 and 0.75 of the standardized precipitation], but that achieved by the linear canonical correlation analysis model was relatively modest ($r$ between 0.25 and 0.55, HK score between $-0.05$ and 0.3, and rmse between 0.4 and 1.2 of the standardized precipitation).

1. Introduction

East Africa (EA), which consists of Kenya, Uganda, and Tanzania (Fig. 1), has two main rainy seasons: September–November (SON) and March–May (MAM). Several investigators have studied the spatial and temporal characteristics of EA seasonal rainfall and its associations with sea surface temperature (SST) variations in the surrounding oceans (e.g., Ogallo 1989; Basalirwa 1995; Mutai et al. 1998; Basalirwa et al. 1999; Ntale et al. 2003). There is a general consensus that the spatial and temporal variabilities of EA rainfall are complex, and, as such, EA has previously been divided into 6–26 homogenous zones of rainfall variability (e.g., Ogallo 1989; Ntale 2001). The spatial inhomogeneities of rainfall variability are attributed to the complex topography, the effects of inland lakes, the seasonal migration of the intertropical convergence zone (ITCZ), and the influence of Indian Ocean SST variations (Ntale 2001). In the last four decades, it has been found that the SON rainfall exhibited greater temporal variability than the MAM rainfall, with one flood in 1961 and another in 1997 (Philippon et al. 2002) and 12 droughts between 1965 and 1997 (Ntale 2001), whereas there was only one recorded failure in the MAM rainfall, which occurred in 1984 (Ntale 2001).

Even though the spatial variability and temporal variability of the SON rainfall have been documented and linked to external forcings such as SST and sea level pressure, it is still a challenge to try to predict the SON rainfall accurately (e.g., Ntale et al. 2003). The deficiency in prediction skill has been observed in other parts of the world, especially in the last two decades (Hastenrath 1995; Webster et al. 2002; Zebiak 2003) and has been attributed to nonstationarity in climate (Landman et al. 2001), nonlinearity in the interaction of climate elements (Barnston et al. 1994), failure to identify robust predictor data (Singh et al. 1995), and the inadequacy of prediction models (Webster et al. 2002). The frequent drought episodes in the latter half of the twentieth century and two floods (in 1961 and 1997) are good examples of the nonstationary character of the SON rainfall of EA. In addition, previous work by Potts (1971), Rodhe and Virji (1976), and Ropelewski and Halpert (1987) found oscillatory peaks in EA rainfall of 2–8 yr. Because of nonstationarity, the temporal resolution of these oscillations is unknown. Therefore, to investigate the SON rainfall variability and its relation-
ship to other climatic elements such as SST, it has become imperative to analyze the events irregularly distributed in time that depict nonstationary power. This paper attempts to solve the problem of nonstationarity by using wavelet analysis and wavelet-based empirical orthogonal function analysis (WEOF), also known in this paper as wavelet principal component analysis (WPCA). Furthermore, because the relationship between rainfall and SST is nonlinear (e.g., Landman and Mason 1999), this study also proposes to teleconnect the SON rainfall to SST using a nonlinear prediction model, the artificial neural network (ANN). The results from the nonlinear prediction model are compared with the linear canonical correlation analysis (CCA) model.

Efforts to improve climate prediction will have economic significance, because agriculture (both rain fed and irrigated) and water resources are inextricably linked to the timing and amount of rainfall. Because there is a clear need to explore longer lead times for predicting seasonal rainfall in EA and to continuously monitor prediction relationships, the objectives of this study are as follows:

1) to identify and analyze the spatial, temporal, and frequency variability of the dominant oscillations of the rainfall fields of EA using wavelet analysis and WEOF,
2) to use the information obtained from objective 1 to explore teleconnectivity between rainfall fields in EA and SST in the Indian and South Atlantic Oceans,
3) and, on the basis of SST predictor fields identified from objective 2, to use a linear statistical model (CCA) and a nonlinear model [artificial neural network calibrated by a genetic algorithm (ANN-GA)] to predict SON rainfall of EA at one-season lead time.

The rest of the paper is organized as follows: Section 2 presents the data and methods used in the analysis. The dominant modes and variability of the SON rainfall are identified and analyzed in section 3. Relationships between EA rainfall variability and the Indian and South Atlantic Ocean SST variability are identified and discussed in section 4. The models used to predict the SON rainfall are presented in section 5. The SON rainfall season is predicted in section 6 using SST data selected from the Indian and Atlantic Oceans. The summary, conclusions, and future work are presented in section 7.

2. Data and methods

a. Data

Monthly rainfall data (1950–97) from 21 grid locations at a resolution of 2.5° × 3.75° latitude–longitude were extracted for EA over 2°N–12°S and 30°–43°E (Fig. 1). The rainfall data are part of a monthly precipitation dataset for land areas worldwide from 1900 to 1998 provided by the Met Office (UKMO). These data were constructed from station datasets averaged onto 2.5° × 3.75° grids using Thiessen polygon weights. The quality control of the data is described in Hulme (1994).

Monthly SST anomaly grid data at 5° × 5° latitude–longitude resolution were extracted from the Indian Ocean (20°N–40°S, 40°–105°E; Fig. 1). The SST datasets were 48 yr long (1950–97) and were transformed into seasonal and annual data by computing 3-month averages [January–March (JFM), April–June (AMJ), July–September (JAS), and October–December (OND)] and 12-month averages, respectively. The SST dataset is part of the UKMO historical SST6 data (MOHSST6), a historical global dataset of mean monthly global SST anomalies with respect to the 1961–90 normals.

b. Methods

1) Wavelet analysis

The wavelet transform of a real signal $X(t)$ with respect to a mother wavelet $\Psi$ is a convolution integral given as

$$W(b, a) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} X(t)\Psi^\ast\left(\frac{t - b}{a}\right) dt,$$

where $W(b, a)$ is the wavelet function, $b$ is the time shift, $a$ is the scale parameter, and $\Psi^\ast$ is the complex conjugate of the mother wavelet $\Psi$.
where $\Psi^*$ is the complex conjugate of $\Psi$, $t$ is time, $T$ is the total length of the time series, and $W(b, a)$ is a wavelet function, a matrix of energy coefficients of the decomposed time series at each scale $a$ and time $b$. The magnitude of the wavelet spectrum coefficients shows how well the wavelet matches with the time series. At each scale, the wavelet spectrum coefficients also depict the amplitude of a time series. In a rainfall or SST time series, power at each scale will therefore be a good measure of the magnitude of the rainfall or SST anomalies. In addition to power at individual scales, power over a range of scales [the scale-averaged wavelet power (SAWP), which represents the mean variance of wavelet coefficients over a range of scales] might also be used. The SAWP of the wavelet spectrum is computed as follows (Torrence and Compo 1998):

$$W^2(b) = \frac{\delta_i \delta_j \sum_{j_1}^{j_2} |W(b, a_j)|^2}{a_j}, \quad (2)$$

where $C_\delta$ is 0.776 for the Morlet wavelet (used in this paper), $\delta$ is a factor for scale averaging, $\delta$ is the sampling period, and $j_1, \ldots, j_2$ represent scales over which the SAWP is computed. There are many different candidate functions to be used as the mother wavelet, and the Morlet wavelet is just one of those (albeit, one of the most commonly used). Because the SAWP is a time series of average variance in a certain band, it can be used to examine the modulation of one time series by another or the modulation of one frequency by another within the same time series (Torrence and Compo 1998). Having obtained the SAWP and individual-scale power, the EOF analysis was then used to extract the joint modes of spatial and temporal variability of the SAWP and individual-scale power.

2) **Wavelet Empirical Orthogonal Function Analysis**

Empirical orthogonal functions analysis, also known as principal component analysis, is among the most popular methods for extracting information from raw data fields to identify objectively the spatially uncorrelated modes of variability (e.g., Mason 1995; Venegas et al. 1997). An EOF transforms an “$n \times K$” data matrix $X$ into another “$n \times K$” matrix of principal components (PC) $U$ that accounts for all the variability of the matrix $X$. Here, $n$ and $K$ represent the length of the time series and the number of data stations, respectively. However, just a few PCs account for the majority of the joint variation of the matrix $X$. Here, they are designated as $u_m$:

$$u_m = e_m^T X' = \sum_{k=1}^{K} e_{km} x_k', \quad m = 1, \ldots, M, \quad (3)$$

where $e_{km}$ are the eigenvectors, $x_k'$ are the anomalies of $X$, and $M$ represents the number of PCs retained. To distinguish EOF of raw data from EOF performed on the wavelet spectra SAWP, the latter is called wavelet empirical orthogonal function analysis or wavelet principal components analysis and their corresponding PCs are referred to as wavelet principal components (WPCs). The WPCs obtained from SAWP are interpreted as frequency-compacted energy variability, and they also indicate the magnitude of events of time series averaged from various scales.

An important decision in EOF is to select an appropriate number of PCs that most strongly capture the joint variability of the original data, without discarding important information carried in the original data. For an “$n \times K$” matrix $X$, there are $K$ eigenvectors. However, atmospheric data typically contain substantial covariances among the original variables, and this attribute implies that the first few eigenvectors will locate directions in which the joint variability of the data is largely accounted for (Wilks 1995). Several techniques are available for determining the optimal number of PCs; however, no universal consensus exists for a single clear criterion, because all criteria are subjective. Note that the number of PCs to be retained for individual scales is no different for the wavelet decomposition than for the original time series, because wavelet analysis does not affect the underlying relationship between the variables at any scale (Bakshi 1998). The scree plot is often used to determine the number of WPCs. In this paper, the numbers of WPCs extracted for analysis were, however, not based on the scree plot. Instead, because WEOF was applied to EA rainfall for the first time, the plotted WPCs were visually inspected for physical explanations; in most cases, WPCs that explained less than 10% of the variance did not have extensive spatial patterns and were excluded from the analysis. The first and second components normally explained a large proportion of the total energy variance of SON rainfall. The third and higher components explained very little variability, and their spatial patterns generally covered small areas.

3. **Dominant Modes and Variability of the SON Rainfall**

Local and global wavelet spectra were computed for the Indian and South Atlantic Oceans and for EA. Figure 2 shows an example of a local and corresponding global wavelet spectrum of SON rainfall. The shaded areas in the wavelet spectrum delineate the areas where the power is significant at the 95% confidence level of a red-noise process. The solid dashed line drawn through the wavelet spectrum delineates the cone of influence (COI). Because the length of the data used (1950–97) is finite, the ends of the time series were padded with zeros to bring the total length of the time series to the next-higher power of 2: 512, 1024, and so on. This padding facilitated the computation of wavelet
power at larger scales and also speeded up the transformation. However, padding the ends of time series with zeros always introduces discontinuities at the end points of the time series, and, as one goes toward larger scales, the amplitude near the edges decreases as more zeros enter the analysis. Therefore, at the end and beginning of the SAWP time series, the variation of power is suppressed and is not easy to interpret.

In Fig. 2 and other figures (not shown), appreciable power was seen at high and low periods. However, the variation of rainfall and SST was dominated by peaks whose periods were between 2 and 8 yr, indicating that ENSO might be within the identified scales. The concentration of energy within these periods was clearly visible as statistically significant peaks in the global wavelet spectra. Hence, for this study, the periods from 2 to 8 yr were chosen for the analysis of rainfall and SST variability and also for the computation of SAWP.

We applied the WEOF analysis to the SAWP of the 21 grid stations and retained the two leading WPCs (WPC1 and WPC2), which explained 63% and 17% of the total SAWP variance, respectively. The third WPC explained only 8% of the total variance, covered small areas, and was discarded. Instead of having to contend with 21 separate SAWP vectors, WEOF thus allows us to express the total SON energy variability by using only two vectors.

The WEOF analysis was based on the correlation matrix, and the spatial distributions of the WPCs are shown in the form of the correlation coefficient between WPCs and the 21 SAWP time series (see Fig. 3). Although the selection of the correlation matrix, as opposed to the covariance matrix, is much more suitable for resolving spatial oscillations (Overland and Preisendorfer 1982), we did not notice any difference between the results obtained by either approach.

WPC1 describes the joint variation of the SAWP of the 21 grid stations within the 2–8-yr period band. Statistically significant variations of WPC1 occur over southwestern Kenya, but strong variations are also found in eastern Uganda and eastern Tanzania, that is, regions of correlation between 0.6 and 0.9. The correlations between WPC1 and SAWP decrease rapidly from 34° to 30°E. This stretch of the region contains the Great Rift Valley (GRV), a geological fault that borders the equatorial Congo basin. The variation not accounted for by WPC1 in this region and northeast Tanzania is explained by WPC2. WPC2 is positively correlated with the SAWP of northeast Tanzania (where Mount Kilimanjaro of 5895 m MSL is located) and southwest Tanzania and is negatively correlated with Uganda. The positive correlation of WPC2 with southwest Tanzania and negative correlation to northern and western Uganda (both areas along the GRV) likely suggest that the mode describes the variation of the SAWP of extreme altitudes (GRV and Mount Kilimanjaro) and some effects from the equatorial Congo basin that are unaccounted for by WPC1.

To determine which periods (also referred to in this paper as scales) explain more variance of the SAWP, and also to extract any interesting scale-based features, WEOF was also applied to the time series of wavelet power of individual scales. The scales used are 2.3, 3.5, and 5–6 yr (Nicholson 1996), others between 2 and 5 yr (Rodhe and Virji 1976), and the scales between 5 and 8 yr shown by wavelet spectrum (Fig. 2). Here, the variance explained by individual scales of the SAWP was computed as a square of the correlation between WPCs of the individual scales and the SAWP for each of the 21 grid locations (Houghton and Tourre 1992). Most of the variance in the SAWP is explained by periods between 2 and 3 yr with a spatial distribution similar to Fig. 3a. The explained variance gradually decreases from 3.5- to 8-yr cycles. The spatial distribution patterns of the explained variances for periods of 2- and 8-yr cycles are shown in Fig. 4.

The dominance of the 2–3-yr cycles in EA and the strong spatial and temporal patterns throughout the
1950–97 period suggest that strong interannual cycles are important in the SON rainfall variability. The internal variability is homogenous over much of EA and is only modulated by areas of extreme altitude, that is, along the GRV and Mount Kilimanjaro. This finding contrasts with those of Ntale (2001), Basalirwa (1995), and Ogallo (1989), who found numerous zones of homogenous rainfall variability. Based on the six zones, Ntale et al. (2003) used CCA in an attempt to predict the SON rainfall. The prediction skill they achieved was
low over most of EA. Given the strong and spatially coherent interannual variations detected, we believe that, outside regions of extreme altitude, SON rainfall is generally predictable with relatively high skill in EA.

The temporal variability of the WPCs is shown in Fig. 5. The variance of WPC1 was high between 1954 and 1974 but decreased considerably between 1976 and 1997. The energy peak in 1961 and the decrease of energy between 1962 and 1997 are consistent with the flood and drought episodes recorded during this period (Ntale 2001). The low variance between 1950 and 1954 and in 1997 (flood year) is most likely due to the zero padding. Consistent with the strong correlation between WPC1 and SAWP of rainfall in Kenya, eastern Uganda, and northern parts of Tanzania (Fig. 3), rainfall in these regions peaked in 1961 and has been on the decline for over three decades (1962–97). Thus, these energy changes are directly related to changes in rainfall and accurately represent the temporal variability of EA SON rainfall.

In comparison with WPC1, WPC2 is less variable between 1954 and 1974. WPC2 suggests that northern and western Uganda (being negatively correlated) generally experienced an increase in SON rainfall between 1982 and 1997, while Mount Kilimanjaro in the northeast of Tanzania and southwest Tanzania experienced decreased rainfall. Because the correlation between WPC1 and SAWP in the GRV area decreases rapidly, this area might have experienced a relatively lower reduction in rainfall as compared with areas with higher correlations. Because the SAWP of southwest Tanzania and much of western and northern Uganda are out of phase, the reduction was probably greater for the former than for the latter.

4. Relationships between SON rainfall WPCs and the Indian and Atlantic Ocean SST SAWP

The objective of this section is to establish relationships between SON rainfall variability in EA and SST variability in the Indian and Atlantic Oceans between 1950 and 1997 by linearly correlating the WPCs of EA rainfall and the SAWP of Indian and Atlantic Ocean SST. Instead of using untransformed (raw) rainfall or SST time series, which contain a wide range of frequencies and noise, we used the SAWP, which predominantly contains frequencies for which the energy was significant at the 95% level, to establish the presence of teleconnections. The relationships were first determined between annual Indian and South Atlantic Ocean SST SAWP and WPC1 of SON rainfall. Next, the relationship between the SST SAWP of the AMJ season and the SON rainfall WPC1 was similarly determined. Among the AMJ, OND, JFM, and SON seasons, AMJ SST SAWP was chosen to exploit the lag relationship between SST and SON rainfall. Furthermore, as shown in Fig. 6c, for the AMJ season, the variabilities of the northern and southern parts of the Indian Ocean are out of phase with each other. Figure 7b also shows a similar out-of-phase relationship between the WPC1 of the SON rainfall of EA and the AMJ SAWP of the Indian Ocean. This result demonstrates that the SON rainfall of EA has been modulated by the Indian Ocean’s SST.

Because of strong persistence in the SAWP and WPC time series, the effective length of the SAWP and WPCs, $N_{eff}$, was estimated using the approximation (Wilks 1995)

$$N_{eff} = \frac{N}{1 + \rho_1},$$  \hspace{1cm} (4)

where $\rho_1$ is the lag-1 autocorrelation. Then $N_{eff}$, which was much shorter than $N$, was used to determine the 95% significance level of the correlation between WPCs and the SAWP.

a. SON rainfall WPCs’ association with Indian Ocean SST SAWP

The spatial correlation pattern between WPCs of the SON rainfall of EA and individual 5° × 5° SST SAWP time series of the Indian Ocean is shown in Fig. 7. For both the annual period and the AMJ season, rainfall WPC1 is positively correlated with the SST SAWP of the southwest Indian Ocean but negatively correlated with the northwest portion of the Indian Ocean. Because SON WPC1 is out of phase with the northern Indian Ocean and is in phase with the southern Indian Ocean SAWP, warming of the northern Indian Ocean SST is associated with decreased SON rainfall for EA. In converse, a cooling of the southern Indian Ocean SST is associated with decreased SON rainfall. The temporal variability of WPC1 and WPC2 of the annual Indian Ocean SST (see Fig. 7c) shows that the northern Indian Ocean began warming in 1960 and continued warming until the early 1990s, whereas the southern Indian Ocean began cooling from 1970 until 1997. A comparison of Figs. 5 and 8c shows that EA responded
immediately to changes in the northern Indian Ocean SST.

b. SON rainfall WPCs’ association with Atlantic Ocean SST SAWP

Figures 7c and 7d also show the correlations between SON rainfall WPC1 and individual 5° × 5° SST SAWP of the South Atlantic Ocean. At the annual scale, rainfall WPC1 is positively correlated with the Brazil Current and the Sierra Leone basin and is negatively but weakly correlated with the Benguela Current and the South Atlantic Ocean between 30° and 40°S. At the seasonal scale, rainfall WPC1 is only positively correlated with the Brazil Current and negatively but weakly correlated with the Guinea Current and the South Atlantic Ocean between 30° and 40°S. The positive relationship between rainfall WPC1 and the Brazil Current SST at both seasonal and annual time scales shows that variabilities in these ocean currents have important climatological implications for EA rainfall.

The Brazil and Guinea Current SSTs appear to be located too far away from the Indian Ocean SST to affect EA rainfall. Using the OND season, that is, SON shifted by one month, Philippon et al. (2002) showed that teleconnections between OND rainfall and various atmospheric and oceanic parameters in the Atlantic Ocean SST exist. We found that the spatial and temporal patterns of the SON (see Fig. 3) and OND WPC1 (figure not shown) are very similar to each other. In addition, we also found that the correlation between SON and OND rainfall of EA is significant at the 99% confidence level, which implies that, on the whole, the results of Philippon et al. (2002) are complementary to ours. We note that the OND rainfall is higher than SON rainfall in southern Tanzania and that from the middle of Tanzania (5°S, 33.75°E) to Uganda and Kenya (2.5°N, 41.25°E) the SON and OND rainfall totals are almost equal, showing that the December rainfall is greater for southern Tanzania but it decreases farther north. Because the SON rainfall variability in EA is positively correlated with the SST in the Brazil Current, warming in these ocean basins during AMJ is associated with increased SON rainfall in EA, whereas the
relationship is out of phase with the Guinea Current SST. The comparison between SON and OND rainfall was intended simply for comparing our results to Philippon et al. (2002).

5. Models for predicting rainfall

a. Artificial neural network calibrated by the genetic algorithm

The neural network used in this study has three layers (input, hidden, and outer layer) and three hidden neurons. A genetic algorithm (GA) was used to calibrate the parameters of the ANN-GA model. GAs are global optimization techniques that create new solutions to a given problem by exploiting the past performance of previous solutions based on a model of natural, biological evolution (Goldberg 1989). The solution space (or population) from which the individual solutions are drawn is represented in the form of finite lengths of strings called chromosomes. The adaptability of chromosomes is improved through a process of systematic modifications made up of crossover and mutation. For the ANN-GA model, the chromosomes were formed from weights and biases assigned to each node, which have traditionally been found using the back-propagation algorithm (Rumelhart et al. 1986). In recent years, the back-propagation algorithm is less preferred than more efficient optimization algorithms such as the conjugate gradient method, simulated annealing, and genetic algorithms (Hsieh and Tang 1998).

The ANN-GA creates an initial set of weights $W^1$ and $W^2$ and biases $B^1$ and $B^2$ (superscripts indicate the first and second layers of the ANNs) for a large number of neural networks. At the start of training, each output of the neural network in the population is initially evaluated against a known predictand. The objective function (fitness) used is based on the Pearson correlation. Next, all the neural networks are ranked according to the computed fitness: the best network at the top and the worst network at the bottom.
The best 85% of the ranked population are randomly selected to compose offspring of the next generation. Selecting from the best 85% of the ranked population ensures that, on average, the new generation has comparatively better fitness than the original population. Selection thus shifts the search space toward better solution spaces of the problem. When selection is made from less than 85% of the ranked population, the search prematurely converges. On the other hand, when selection is made from more than 85% of the ranked population, on average there is no systematic improvement in the average fitness of the subsequent populations and the search results in suboptimum solutions. Because the population is usually maintained as constant, some members of the old population are selected more than once.

Next, pairs of neural networks are selected in the same order in which they were selected from the previous population, and their weights and biases are exchanged. This procedure is called crossover. To effect exchange of weights and biases, a location is randomly chosen in the hidden layer and weights on either side of the location are exchanged between the two neural networks (shown as thick lines in Fig. 9). This procedure is repeated between all other pairs of neural networks in the population.

Next, a small percentage (typically between 0.1\% and 1\%; in this study, 1\% was used) of the population is randomly selected for which some of the weights and biases are randomly replaced (see thick lines in Fig. 9). The process is called mutation and is designed to restore good weights and biases eliminated during selection. Because mutation is a purely random process, it is always kept to a minimum to prevent the search from degenerating into a random process. If mutation results in a better neural network, that network will likely survive in the next selection; if mutation results in a more inferior individual, that individual will likely perish in the next selection. The neural networks are once again evaluated against the same known predictand.

This procedure is repeated through several generations. At each generation, the best network is kept until a better solution is found in successive generations. Convergence is reached when no better solution can be found and when at least 95\% of the networks have the same weights and biases. At the end of the run, the weights and biases of the best surviving network are kept to be used for making predictions with new input data.

To evaluate each neural network, the predictand \( y \) is obtained as a nonlinear translation of the weighted average of the predictor data \( x \), which have been normalized [that is \( x = (x - \bar{x})/\sigma_x \), where \( \bar{x} \) and \( \sigma_x \) are the mean and standard deviation of \( x \), respectively].

The first step is to compute the weighted input to the \( j \)th hidden unit, \( \text{hidunit}_{pj} \):

\[
\text{hidunit}_{pj} = \sum_{i=1}^{N} W_{ji}x_{pi} + B_{pj},
\]

where \( W_{ji} \) is the weight connecting the \( i \)th predictor to the \( j \)th hidden unit, \( B_{pj} \) is the bias for the \( j \)th hidden unit, and \( x_{pi} \) is the input to the \( i \)th predictor. The output of the \( j \)th hidden unit is then computed as a nonlinear function of the weighted input. The final output of the neural network is obtained by aggregating the outputs of the hidden units.
where $N$ is the total number of input nodes, $W_{ji}$ are the weights from input unit $i$ to the hidden unit $j$, $B_{jj}$ are the biases for hidden neuron $j$, and $x_{pi}$ is the $i$th input of pattern $p$ (in this case SST PCs are used). The hidden layer undergoes a nonlinear translation:

$$f \left( \text{hidunit}_{pj} \right)$$

where $f$ is the $j$th neuron nonlinear activation function. The output, $y_{pk}$ is computed as a weighted average of the hidden units:

$$y_{pk} = \sum_{j=1}^{M} W_{kj} f \left( \text{hidunit}_{pj} \right) + B_{ko},$$

where $M$ is the number of hidden units, $W_{kj}$ represents
the weight connecting the hidden node $j$ to the output $k$, and $B_{jk}$ is the bias for output neuron $k$.

b. Canonical correlation analysis

CCA is an established statistical forecasting scheme and so its details are not discussed here. Readers interested in CCA can refer to Barnett and Preisendorfer (1987) and Glahn (1968). Because CCA is a widely used linear model, it is used here as a benchmark for comparison with the nonlinear ANN-GA. The raw predictor fields from the Indian Ocean and the South Atlantic Ocean identified through WEOF were first reduced to five dominant EOF or PCA modes as input in CCA. In applying CCA, 38 yr of prior data were used in the model to predict rainfall of the following season, such as using the 1950–87 AMJ SST data to predict the 1987 SON rainfall, 1951–88 AMJ SST data to predict the 1988 SON rainfall, and so on.

6. Predicting the SON rainfall of East Africa (1987–97)

The strong interannual cycles of the SON rainfall and the strong seasonal relationships between the SST SAWP of the South Atlantic and Indian Oceans and the EA rainfall WPCs found in section 3 show that the prediction of SON rainfall in EA based on the preceding AMJ SST of the two oceans is possible. Raw SST predictor data were extracted for April, May, and June from the southwest and northwest Indian Ocean and the Brazil Current region in locations where the correlation between WPC1 of SON and the SAWP SST exceeded 0.5 (Figs. 7b,d). Because the autocorrelation for the SAWP and the WPCs was above 0.9, the effective length of the SAWP and the WPCs was reduced from 48 to about 3 yr [from Eq. (4)] and, at 95% significance level, the correlation between SON WPCs and SST SAWP (about 3-yr in length) was above 0.997. However, the ocean areas covered by this level of correlation of 0.997 were small, and very few data could be collected. Because the spatial correlation pattern between SON WPC1 and SON SAWP of Fig. 3 covering most of the region varied between 0.5 and 1.0, it was decided that SST data would be collected from areas of the ocean where the correlation between SON WPC1 and SST SAWP exceeded 0.5, which was a correlation level above which sufficient data points could be collected. The raw data were standardized and averaged over the 3 months for each grid station in the two oceans to give one AMJ dataset. Rainfall data for the SON season were also extracted for each of the 21 grid stations. The rainfall data were also standardized.

To evaluate the prediction skill, the Pearson correlation, the Hansen–Kuipers (HK) skill score, and rmse were used. The Pearson correlation coefficient $\rho$ is computed as

$$\rho = \frac{\sum_{k=1}^{n} (\text{obs}_k - \bar{\text{obs}})(\text{pred}_k - \bar{\text{pred}})}{\left[ \sum_{k=1}^{n} (\text{obs}_k - \bar{\text{obs}})^2 \sum_{k=1}^{n} (\text{pred}_k - \bar{\text{pred}})^2 \right]^{1/2}},$$

where $\text{obs}_k$ and $\text{pred}_k$ are the observed and predicted values for sample $k$, $\bar{\text{obs}}$ and $\bar{\text{pred}}$ are their respective means, and $n$ is the sample size. This $\rho$ varies between 1 and −1, with the maximum and minimum values indicating perfect positive and negative linear relationships, respectively. To compute the HK skill score, the predicted and observed rainfall data are grouped into categories: “dry,” “near normal,” and “wet.” Tercile percentages of below 33%, 33%–66%, and above 66% of the range of standardized SON rainfall were used to define the categories in a square contingency table. Then,

$$\text{HK} = \frac{H - E_c}{T - E_m},$$

where $H$ is the total number of correct forecasts, $T$ is the total number of correct forecasts obtainable with a perfect forecast model, $E_c$ is the number of correct hits expected by chance, and $E_m$ is the marginal number of correct (observation) hits expected by chance. For an $L \times L$ contingency table, the HK score may be expressed in terms of probabilities as

$$\text{HK} = \frac{\sum_{j=1}^{L} p(\text{obs}_j, \text{pred}_j) - \sum_{j=1}^{L} p(\text{obs}_j) \times p(\text{pred}_j)}{1 - \sum_{j=1}^{L} [p(\text{obs}_j)]^2},$$

where $\text{obs}_j$ and $\text{pred}_j$ are the $j$th observed and predicted values. The HK score values range from −1 to +1. Perfect forecasts receive a score of 1, random forecasts receive a score of 0, and forecasts inferior to random forecasts receive negative scores. In this paper $L = 3$. The root-mean-square error is computed as

$$\text{rmse} = \left[ \frac{1}{n} \sum_{k=1}^{n} (\text{obs}_k - \text{pred}_k)^2 \right]^{1/2}.$$

The spatial display of the correlation, HK scores, and rmse between predicted and observed SON rainfall is shown in Fig. 10 for both ANN-GA and CCA. Correlations between 0.7 and 0.9, rmse of 0.4 and 0.75, and HK scores between 0.2 and 0.8 were achieved using ANN-GA. For CCA, the correlations ranged between 0.3 and 0.6, the rmse ranged between 0.5 and 1.2, and HK scores ranged between −0.05 and 0.3. For the ANN-GA model, high prediction skill was achieved in central and eastern Tanzania, all of Kenya, and almost all of Uganda, with correlations greater than 0.8, rmse less than 0.65, and HK scores greater than 0.4. These regions were also found to have the largest positive correlations between SON rainfall WPC1 and the SON rainfall SAWP (see Fig. 3a). The prediction skill decreases westward from about 33° to 30°E in a fashion similar to the decrease in strength of the correlation between SON rainfall WPC1 and SAWP. In Fig. 3, this area was identified as the GRV, where SON rainfall is probably modulated by the equatorial Congo basin or by the sudden change in altitude when entering into the GRV. The prediction skill of this region was low in comparison with the rest of the region. In the GRV, the model predicted the wrong category of rainfall 6 out of 11 yr; of those 6 yr, 2 were overpredicted and 4 were underpredicted (see Table 1). On the other hand, Table 2 shows that for an area in Kenya 10 out of 11 yr were accurately predicted, with only one prediction falling into the wrong category; instead of near normal, the year was predicted as wet.

Similar to the example shown in Table 2, the majority of the contingency tables for Kenya, eastern Uganda, and northern Tanzania had more years in the dry category than did the stations in the GRV area, showing that the reduction in rainfall was experienced more by the former than by the latter in the 1987–97 period, as was found in section 3a.

The skill scores for the CCA are shown in Figs. 10d–f. In contrast to ANN-GA, the skill scores were lower. Correlations between observed and predicted SON rainfall were lowest around Mount Kilimanjaro. Elsewhere, they ranged between 0.30 and 0.50. The rmse generally increased from east to west, with the largest errors recorded in southwest Uganda. The HK scores were generally low (between −0.05 and 0.15) throughout much of EA. The increase in error from east to west most likely reflects the weakening variations of the SON signal toward the GRV. The weak correlations (0.3–0.5) likely suggest that CCA could not quite capture the nonlinear co-variations between the SST and SON rainfall, a result that is expected because it is a linear model. With HK scores of between

![Fig. 10. Contour plots at 0.05 or 0.1 intervals (to avoid crowding) showing the (a), (d) correlation coefficient, (b), (e) rmse, and (c), (f) HK scores between SON predicted and observed standardized rainfall for (a)–(c) ANN-GA and (d)–(f) CCA driven by SST predictor fields from the Indian and Atlantic Oceans shown in Figs. 7b,d where correlations were greater than 0.5.](image-url)
0.05 and 0.25, the prediction skill was poor for most of EA.

The above results show that the SON rainfall variability is influenced by SST variations in the southwest and northwest Indian Ocean and the Atlantic Ocean. Overall, using sectors of predictor fields identified in the oceans, the nonlinear ANN-GA system could predict between 49% and 81% of the rainfall variability at 2-month lead time, whereas the linear CCA system could only predict between 6% and 36% of the SON rainfall variability. The fact that the prediction skill of ANN-GA was consistently higher than CCA shows that the nonlinear relationships between EA rainfall and SST variations of the Indian and Atlantic Ocean basins are captured better by the former than by the latter.

7. Summary, conclusions, and future work

We have used wavelet spectra information 1) to identify and to analyze the spatial, temporal, and frequency variability of dominant oscillations in the rainfall of EA and SST in the South Atlantic and Indian Oceans, 2) to explore teleconnection patterns between SST fields from the Indian and South Atlantic Oceans and rainfall in EA to identify relevant SST predictor fields from the two oceans to drive the teleconnection models, and 3) to assess the prediction skill of linear and nonlinear statistical teleconnection models (CCA and ANN-GA, respectively) for predicting the SON rainfall of EA at a 2-month lead time.

The results show that SON rainfall has a strong signal in much of EA, except along the Great Rift Valley and Mount Kilimanjaro. We found that the signal is linked to the South Atlantic and the Indian Oceans. Cooling of the Brazil Current and the southwest Indian Ocean coupled with cooling of the southern Indian Ocean—in particular, the southwest Indian Ocean—and warming of the northern Indian Ocean are associated with below-normal rainfall and sometimes droughts in some parts of EA. We also confirmed that failure of the SON rainfall causes droughts mainly in Kenya, Uganda, and northern Tanzania, where the SON rainfall accounts for over 30% of the annual rainfall and where the signal is strongest.

Using energy extracted at individual frequencies within the 2–8-yr ENSO band, we found that the dominant frequencies predominantly lie between 2 and 3 yr, followed by those between 3.5 and 8 yr. We also found that the WPC1 of SAWP accurately represents the declining SON rainfall and the history of floods and droughts between 1950 and 1997. The decline in SON rainfall resulted in the 12 droughts experienced in EA during this period.

Using a nonlinear ANN-GA model, high skill was achieved in predicting the SON rainfall over most of EA, whereas the linear CCA model performed poorly in all of EA. The overall summary statistics for the ANN-GA model are as follows: ρ between 0.70 and 0.9, HK score between 0.2 and 0.8, and rmse between 0.4 and 0.75; those for CCA are as follows: ρ between 0.25 and 0.55, HK score between −0.05 and 0.3, and rmse between 0.4 and 1.2. It is apparent that nonstationarity and nonlinearity were prominent characteristics of climate data (rainfall and SST) in EA and the two oceans, which means that an application of a nonlinear teleconnection model, such as the ANN-GA presented in this study, is more likely to be successful than an application of the standard linear CCA.

Further analysis of the EA rainfall variability (both SON and MAM) is being performed using the empirical mode decomposition and the Hilbert spectrum method (Huang et al. 1998).

Acknowledgments. This study is partly funded by the Natural Science and Engineering Research Council (NSERC) of Canada, and the first author is supported by the Commonwealth scholarship of CIDA, Canada. UKMO provided both the SST grid data of the Indian and Atlantic Ocean (part of MOHSST6) and the rainfall data for East Africa. The wavelet analysis was done using the software of Torrence and Compo (1998) downloaded from http://paos.colorado.edu/research/wavelets.

REFERENCES


Glahn, H. R., 1968: Canonical correlation and its relationship to

Table 1. Contingency table for a grid in the GRV with HK = 0.18.

<table>
<thead>
<tr>
<th></th>
<th>Dry</th>
<th>Near normal</th>
<th>Wet</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dry</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Near normal</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Wet</td>
<td>0</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 2. Contingency table for a grid in Kenya with HK = 0.85.

<table>
<thead>
<tr>
<th></th>
<th>Dry</th>
<th>Near normal</th>
<th>Wet</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dry</td>
<td>7</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Near normal</td>
<td>0</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Wet</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>


