Teleconnections between Ethiopian summer rainfall and sea surface temperature: part II. Seasonal forecasting

G. T. Diro · D. I. F. Grimes · E. Black

Abstract A seasonal forecasting system that is capable of skilfully predicting rainfall totals on a regional scale would be of great value to Ethiopia. Here, we describe how a statistical model can exploit the teleconnections described in part 1 of this pair of papers to develop such a system. We show that, in most cases, the predictors selected objectively by the statistical model can be interpreted in the light of physical teleconnections with Ethiopian rainfall, and discuss why, in some cases, unexpected regions are chosen as predictors. We show that the forecast has skill in all parts of Ethiopia, and argue that this method could provide the basis of an operational seasonal forecasting system for Ethiopia.

Keywords Statistical model · Seasonal forecasting · Teleconnection · Ethiopia · Rainfall · Sea surface temperature

1 Introduction

This paper forms part of an investigation into the relationship between global sea surface temperatures (SSTs) and rainfall in Ethiopia. A preliminary investigation (Gissila et al. 2004) showed that a statistical model making use of the correlation between SST in the Pacific and Indian Ocean and rainfall in various part of Ethiopia could

be used to forecast Ethiopian seasonal rainfall total. In Diro et al. (2008) (hereafter, DBG08), the approach was further developed making use of more data and more sophisticated statistical methods to identify global SST regions which give the optimum forecast of Belg (February–May) rainfall. In part I of this paper, Diro et al. (2010) (hereafter DGB10), the teleconnection mechanisms linking global SSTs and Kiremt (June–September) rainfall were analysed and described. In this work (part II), the statistically optimal SST regions for forecasting Kiremt rains are identified and we discuss the extent to which they can be explained by the teleconnections described in part I.

Accurate and reliable seasonal forecasts of rainfall have enormous potential benefit in many areas of Ethiopian society. According to the latest UN report (http://data.un.org/CountryProfile.aspx?crName=Ethiopia) close to 85% of the Ethiopian population are living in rural areas where they are engaged in rain fed agriculture. Zhou et al. (2004) investigated the relationship between climate variability and malaria epidemics in the East African highlands including Ethiopia and concluded that reliable malaria early warning is feasible only if seasonal rainfall climate forecasts are incorporated in the process. In the energy sector, according to the Ethiopian electric and power corporation (EEPCO) over 98% of the energy production in 2005, 2006 and 2007 comes from hydropower although that percentage has slightly decreased temporarily over the last couple of years (http://www.gibe3.com.et/EEPCo2001.pdf).

Seasonal forecasting systems are usually categorised as dynamical, statistical or hybrid. Dynamical methods are based on forecasts made by coupled atmosphere–ocean General Circulation Models (GCMs). Operational dynamical seasonal forecasts are provided by for example, the European Centre for Medium-Range Weather Forecasts (ECMWF). Statistical methods of seasonal forecasting use

G. T. Diro (✉)
Earth System Physics Section, The Abdus Salam International
Centre for Theoretical Physics, Strada Costiera 11, Trieste, Italy
e-mail: gtefera@ictp.it

G. T. Diro · D. I. F. Grimes · E. Black
Department of Meteorology, University of Reading,
Reading, UK

the historical relationship between rainfall and other weather/climate related parameters. SST is a key indicator because of its relatively slow rate of change and the strong ocean atmosphere coupling. Hybrid methods usually employ a combination of statistical and dynamical techniques, for example by using statistical techniques to forecast SSTs which are then input into a dynamical atmosphere-only GCM to generate seasonal rainfall forecasts.

Currently, dynamical forecasts do not show as much skill in the East African region as statistical forecasts (Jan van Oldenborgh et al. 2005), probably due to the complexities of the East African climate (DGB10, Segele et al. (2009)). There is now a growing body of work on statistical methods for East African seasonal forecasting using oceanic and atmospheric variables as predictors (Mutai et al. 1998; Hastenrath et al. 2004; Philippon et al. 2002; Gissila et al. 2004; Korecha and Barnston 2007; Block and Rajagopalan 2007; Yeshanew and Jury 2007; Camberlin and Philippon 2002; Ntale et al. 2003; Mwale and Gan 2005). Among these only few of them (e.g. Gissila et al. 2004; Korecha and Barnston 2007; Block and Rajagopalan 2007) have focused on the main rainy season of Ethiopia. The methods used to develop a forecast for East Africa include regression-based deterministic algorithms e.g. Gissila et al. (2004); Korecha and Barnston (2007); Hastenrath et al. (2004); Yeshanew and Jury (2007) or probabilistic methods (e.g. Block and Rajagopalan 2007) or linear discriminant analysis based methods (e.g. Mutai et al. 1998; Philippon et al. 2002).

Undoubtedly, a skilful seasonal forecasting system would be valuable for Ethiopia. However, developing and implementing such a system has proved difficult—in part because of the topographic complexity of the region and the spatial variability of the meteorology. The following section briefly describes the spatial variability in Ethiopian climate.

In Ethiopia, altitude ranges from 200 m below sea level in the northeast to over 4,000 m above sea level in the northern highlands. Ethiopia's topography, together with its geographical location lead to high spatial variability in precipitation, with the seasonal cycle, the factors governing interannual variability and the annual total all exhibiting large variability within the country.

Broadly speaking, Ethiopia can be divided into three regions, based on the seasonal cycle in rainfall. The northern and central western part of the country has a single rainy season with a peak in July/August; central and eastern Ethiopia has two rainy periods February–May (Belg) and June–September (Kiremt). Southern Ethiopia has two rainy seasons, the long rains (March–May) and the short rains (September–November). The shape of the seasonal cycle is controlled by the passage of the Inter-tropical convergence zone (ITCZ). During the course of the year the main ITCZ moves from 15N to 15S and back again, passing over Ethiopia in the boreal summer and spring. Additionally, a meridional arm of the ITCZ, induced by the difference in the heat capacity of the land surface and Indian Ocean, produces rainfall over the southwestern Ethiopia in February and March, when the main ITCZ is in the southern hemisphere (Kassahun 1987).

Ethiopia experiences high inter-annual variability in rainfall in all seasons. Previous observational studies have suggested that in summer this interannual variability is ultimately controlled by large scale phenomena, such as the El Niño Southern Oscillation (for example Gissila et al. 2004; Bekele 1997); Segele and Lamb 2005; Segele et al. 2009). Modelling studies (e.g. Folland et al. 1986) also suggest that Pacific and Indian Ocean SST influence Ethiopian rainfall. Specifically, Ethiopian rainfall is negatively associated with both Indian Ocean and Eastern Pacific SST. This means that excess rainfall tends to occur when there is a La Nina, and deficit rainfall when there is an El-Nino. Consistent with this high summer rainfall in Ethiopia is associated with a strong Indian monsoon. As is shown later, these statements referring to Ethiopian rainfall teleconnections are generalisations. There is considerable variability at a more local level.

Large-scale, remote phenomena, such as ENSO, affect Ethiopian rainfall via their influence on the positions and strengths of the nearby subtropical high pressure systems, upper level jet streams and low level winds (for example Camberlin and Philippon 2002 for spring and DGB10 for summer). The Tropical Easterly Jet (TEJ) is the most important upper level feature for the summer rains in Ethiopia, with a stronger (weaker) TEJ associated with shorter (longer) dry spells (Segele and Lamb 2005). Circulation at lower levels also plays a significant role. Specifically, moisture flux from the Indian and Atlantic Oceans, and hence precipitation, is affected by the intensities and positions of the Mascarene and St. Helena highs (Kassahun 1987).

As has been described previously, the seasonal cycle in rainfall varies substantially within Ethiopia. The first part of this study (DGB10) demonstrated that teleconnections with remote SSTs also exhibit large spatial variations within Ethiopia. The remote effect of equatorial Atlantic SST was found to be particularly complex, with a warm Gulf of Guinea both weakening the TEJ and strengthening the ITCZ. The net effect is that a warm Gulf of Guinea is associated with higher rainfall in the north and lower rainfall in the south.

The variability in teleconnections within Ethiopia described in DGB10 motivates the development of seasonal forecasts at a regional rather than national scale. In this paper, we describe the development of such a system for the Kiremt rains. The study described in this paper

primarily aims to extend our understanding of how tele-connections with remote SSTs can be utilised to forecast Ethiopian rainfall. With reference to the first part of the study, it will also comment on the physical mechanisms underlying forecast skill. In order to relate the results of this study to the idealised experiments described in DGB10, the model developed here only considers SST anomalies as potential predictors. Additional factors that may improve the skill, such as circulation and land–surface anomalies, are not considered.

In summary, this study aims to address the following questions:

- What is the relationship between objectively selected predictors and the physical teleconnections described in DGB10?
- How skillful are regional statistical forecasts of rainfall using only SSTs in advance of the season? How much does skill improve if contemporaneous predictors are used?

2 Data and data processing

Monthly rainfall totals from 45 raingauges in Ethiopia covering the period 1969–2003 were provided by the Ethiopian National Meteorological Agency (NMA). For the purpose of analysis, Ethiopia was sub-divided into six climatologically homogeneous rainfall zones as shown in Fig. 1. Locations of the gauges and quality control methods were described in DBG08 and DGB10 and the method used in determining the zone boundaries is described in detail in DBG08. The nomenclature of the zones is for consistency with previous work.

Sea surface temperature (SSTs) for the same time period were obtained from the UK Met Office Hadley Centre Global Sea Ice and Sea Surface Temperature (HadISST)
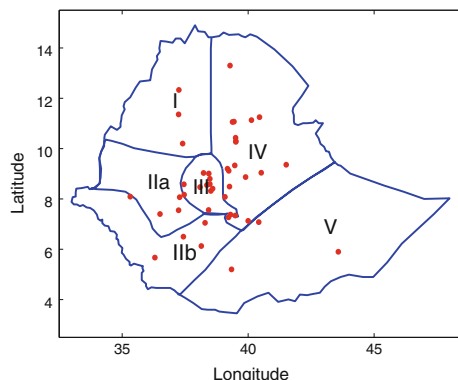

Fig. 1 Homogeneous rainfall zones for Kiremt season. After Diro et al. (2010)

version 1.1 (Rayner et al. 2003). This product gives monthly mean temperatures with a resolution of $1° \times 1°$. In situ sea surface observations and satellite derived esti-mates of SST are included in the data.

3 Seasonal forecasting

The seasonal forecasting system has three stages:

- selection of predictors,
- implementation of a statistical forecasting model for each homogeneous rainfall zone
- evaluating the forecast skill.

The three stages are described in the following subsections.

3.1 Selection of predictors

Candidate predictors were selected as in DBG08 by identifying the months and oceanic areas for which SST was most highly correlated with Kiremt total rainfall in a given rainfall zone. Correlations were determined for individual calendar months with lead times up to 8 months before the Kiremt Season for all zones. Examples of the contemporaneous correlations are shown in DGB10. Where the same SST region shows a significant correlation with rainfall for different lead times per season, only the one with the highest correlation over the largest area was included. For example in Fig. 2, it can be seen that Feb-ruary, March, April and May SST of the tropical Pacific are significantly correlated with Kiremt Zone I rainfall. In this case the May SST was used because it showed the highest correlation over the largest area. Using this method, up to 20 potential candidate predictors were obtained from various oceanic regions with different lag times for each zone.

As in DBG08, so as to avoid problems with collinearity and overfitting, stepwise regression, analysis of variance inflation factor and stepwise discriminant analysis were used to eliminate redundant predictors. The details of these procedures are described in DBG08. Combining these two selection methods with inclusion or exclusion of contem-poraneous predictors, four sets of predictors plus an equatorial Pacific only set were generated. i.e.

- Set A: selected by stepwise regression, excluding predictors contemporaneous with Kiremt rainfall
- Set A′: as Set A except contemporaneous season predictors were included
- Set B: selected by stepwise discriminant analysis excluding predictors contemporaneous with Kiremt rainfall

Fig. 2 Spatial pattern of correlation between Kiremt Zone I rainfall and SST in months of previous Belg season. Contour lines represent significant correlation at 0.1 level
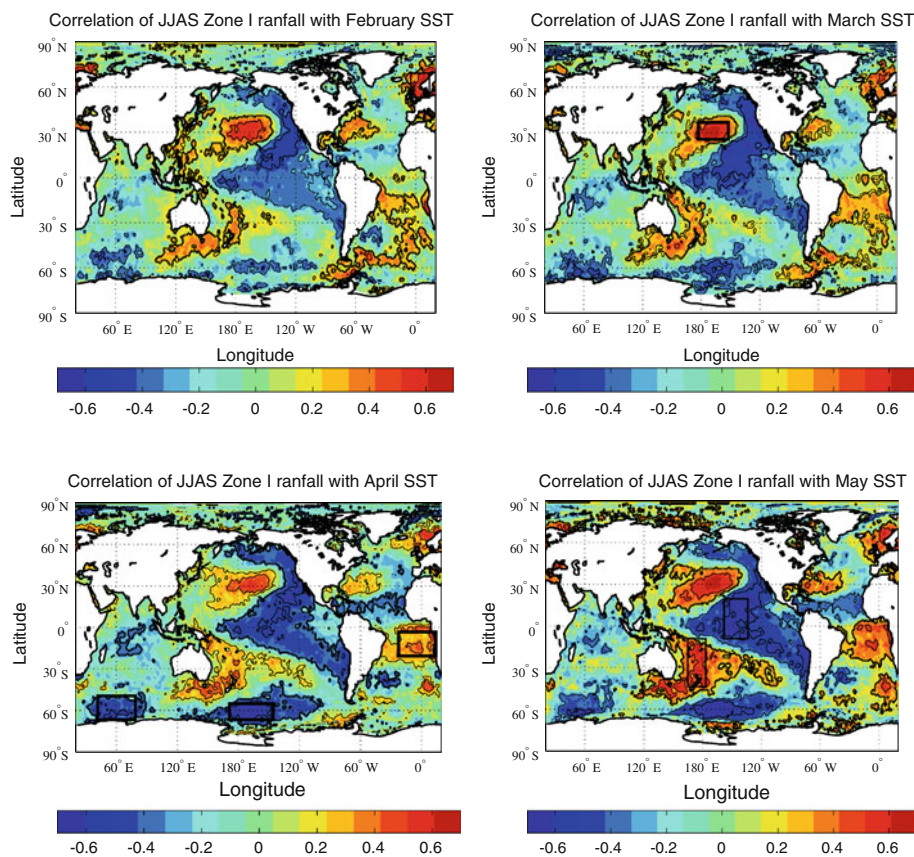
- Set B′: as Set B but contemporaneous predictors were included
- Set C: Equatorial Pacific only

Table 1 shows that the number of selected predictors ranges from two to five when excluding contemporaneous season predictors and up to seven when contemporaneous predictors are included. The geographical location of the predictors is shown in Fig. 3. It can be seen that the predictors chosen vary from zone to zone and that different methods of selecting predictors may result in different predictor sets. This sensitivity to method probably arises from the strong cross correlation between candidate predictors.

The coefficient of determination (R^2), which is defined as the proportion of variance explained by the regression model is shown in Table 2. R^2 values adjusted to take

Table 1 The number of selected predictors for each predictor set for Kiremt

Predictors set	Zone I	Zone IIa	Zone IIb	Zone III	Zone IV
Set A′	7	6	3	4	7
Set B′	3	5	3	4	6
Set A	5	5	4	4	3
Set B	4	5	3	3	4

account of the varying number of degrees of freedom are also shown. It can be seen that (R^2) varies both with zone and predictor set. Set A′ (including contemporaneous SSTA) gives higher values except in the case of Zone IIb as shown in Table 2. The $\lambda_{\text{Wilk's}}$ representing the ratio of within category variance to the total variance is shown in Table 3. The smaller the value of $\lambda_{\text{Wilk's}}$, the better the discrimination. From the values of $\lambda_{\text{Wilk's}}$ for predictor set B (selected using stepwise discriminant analysis by excluding the contemporaneous season predictors) it can be seen that the best separation between categories is observed for Zone I ($\lambda_{\text{Wilk's}} = 0.12$).

As with R^2 the best discrimination is observed when contemporaneous predictors are included for all zones except Zone IIb. The fact that including contemporaneous SST predictors does not always improve the forecast skill was discussed by DBG08 in the context of the short (Belg) rainy season.

DGB10 showed that the strength, and in some cases even polarity, of teleconnections between precipitation and remote SST varies within Ethiopia. It is therefore not surprising that the predictors vary from one region to another. The relationship between the teleconnections described in DGB10 and the predictors is not, however, obvious. For example DGB10 found that in much of Ethiopia, SST anomalies in the equatorial Pacific affect
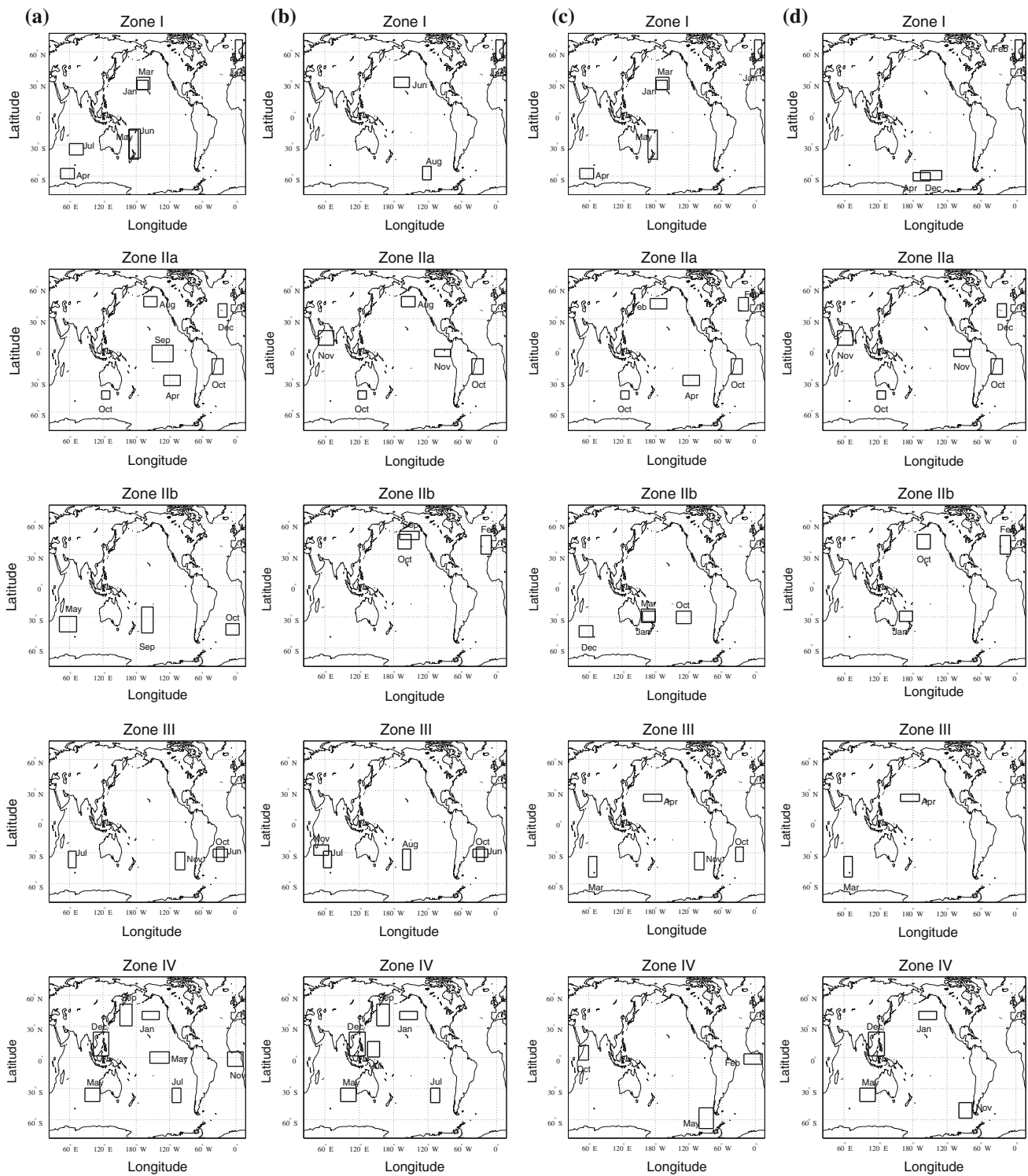
Fig. 3 Location and lag time of predictors used in the forecast model for each zone for Kiremt. Predictors in **a** are chosen by including the contemporaneous season and using stepwise regression **b** the same as in **a** but using stepwise discriminant analysis **c** by excluding those from the contemporaneous season and using stepwise regression **d** the same as **c** but using stepwise discriminant analysis

Kiremt precipitation, with warm anomalies associated with low rainfall. Despite this, Fig. 3 (c and d) shows that SST in the equatorial Pacific is not selected as a predictors for

Zones I, IIb, III and IV. Instead, the Pacific predictors are located in the central and northwest and south Pacific. Both the observed and modelling studies described in DGB10

Table 2 Summary of the coefficient of determination (R^2) of the regression model for Kiremt. The figures in bracket are the adjusted R^2

Predictors	Zone I	Zone IIa	Zone IIb	Zone III	Zone IV
Set A′	0.84 (0.80)	0.83 (0.79)	0.57 (0.53)	0.76 (0.73)	0.79 (0.74)
Set A	0.79 (0.76)	0.72 (0.67)	0.60 (0.54)	0.58 (0.52)	0.62 (0.58)

Table 3 Summary of Wilk's λ of the discriminant analysis model for Kiremt

Predictors	Zone I	Zone IIa	Zone IIb	Zone III	Zone IV
Set B′	0.12	0.11	0.25	0.16	0.10
Set B	0.12	0.18	0.31	0.38	0.25

suggest that anomalies in the northwest Pacific are causally linked to precipitation only in Zone IV (northeast Ethiopia).

There are two physical reasons for an apparently non-causal predictor (in this case extra tropical Pacific SST anomalies) to be selected over a causal predictor (in this case equatorial pacific SST anomalies). Firstly, the idealised modelling studies and composites presented in DGB10 are based on contemporaneous links. The prediction model, on the other hand, considers SST anomalies several months in advance of the season. It is unlikely that SST anomalies anywhere on the globe this far in advance of the rainy season have a direct causal link with precipitation amounts. We are thus seeking SST anomalies that are pre-cursors to those that directly influence rainfall. In the case of the equatorial Pacific, January, February and March SST anomalies in the northwest Pacific have higher correlation with JJAS SST anomalies than do the SST anomalies in the region itself ($R = -0.41$ as compared with $R = -0.34$ for March for example), probably because of the character of the ENSO cycle.

Secondly, the statistical prediction model does not directly account for non-linearities in climate dynamics, and hence the nature of the causal relationships between SST anomalies and rainfall. However, candidate predictors that reflect these non-linearities will be better correlated with rainfall than those that do not. Consider again the equatorial Pacific. DGB10 reported that the relationship between Ethiopian rainfall and equatorial Pacific is non-linear, with warm SST having a stronger link than cold SST. The relationship between equatorial Pacific SST anomalies and those in the northwest Pacific exhibit a similar non-linearity, with warm anomalies in the equatorial Pacific affecting the northwest Pacific more strongly than cold anomalies (because of the differences in the patterns of anomalies associated with La Nina and El Nino). This is reflected by a significantly stronger correlation between Kiremt rainfall and March SST in the northwest Pacific (0.67 as compared to -0.42 for equatorial Pacific).

The idealised studies reported in DGB10 suggested that anomalies in the midlatitude northwest Pacific directly cause rainfall anomalies in rainfall Zone IV (northeast Ethiopia) towards the end of the rainy season in August, and particularly in September. Specifically, cool SSTs cause a southward shift of the ITCZ and low rainfall, while warm SSTs enhance the ITCZ and cause high rainfall (DGB10). Consistent with this, September SST anomalies in the midlatitude northwest Pacific are selected as a predictor for Zone IV.

DGB10 has shown that, in addition to the equatorial Pacific, SST anomalies in the Atlantic are causally linked with precipitation in central and southern Ethiopia. The observational studies described in DGB10 and Diro (2008) suggested that, SST anomalies in the Gulf of Guinea in advance of the rainy season may affect Ethiopian rainfall via the low level circulation of the Atlantic and humidity anomalies over the eastern Atlantic. Consistent with this, Gulf of Guinea SST anomalies well in advance of the rainy season (November and February) are selected as predictors for Zone IV. However, for Zones I and II, Gulf of Guinea SST anomalies were not selected because the linear correlations between Gulf of Guinea SST in advance of the season, and rainfall are low (for example for Zone I, ranging from near zero in February to 0.26 in June). The low linear correlations reflect the non-linear nature of the link between Gulf of Guinea SST and rainfall in western Ethiopia (cold SST has a weaker effect than warm SST). The correlations are also weakened by the variability in the strength of the teleconnection within rainfall zones. Specifically, northern Ethiopia is less affected by Gulf of Guinea SST than southern and central Ethiopia. The observational and modelling studies reported in DGB10 suggest that this variability arises from the enhancement of the ITCZ (caused by warming of the Gulf of Guinea) having a stronger effect on northern Ethiopia. The strongest pre-cursors of Gulf of Guinea SST anomalies are found in the southern mid- and high latitudes of the Pacific and Atlantic, with cooling in February and March negatively correlated with June SST in the Gulf of Guinea ($R < -0.6$ in much of the region). Consistent with this, high/ higher mid-latitude southern SST anomalies are selected as predictors for all of the regions except IIa (central western Ethiopia).

In most cases, as has been shown, it is possible to explain the selection of predictors physically. However, in

a few cases, this is not true. For example, it is difficult to explain why the SST anomalies in the northern Atlantic (around Great Britain) are well correlated with rainfall in northern Ethiopia. Whether these teleconnections are statistical artefacts or demonstrate a real but obscure teleconnection can only be ascertained by careful analysis of more years of data and further modelling studies.

3.2 Development of a forecast model

Two empirical statistical techniques were used to formulate a statistical model: multiple linear regression (MLR) and linear discriminant analysis (LDA). The primary difference between MLR and LDA is that in LDA the output is a categorical probabilistic forecast whereas in MLR the output is a deterministic value together with the uncertainty associated with it. For the MLR, a probablistic form of regression (similar to the form of the LDA) was used by converting the prediction errors of the regression to a probablity density function. These methods are described in depth in DBG08 and in "Appendix" to this paper.

3.3 Results of the seasonal forecasting

To estimate the true predictive skill, it is crucial that the set of forecasts are independent of the data used in the calibration of the model. In other words, the verification data should be completely different from the calibration data. The calibration error statistic cannot be used as an indication of the performance on future data, because the new data will not be exactly the same as the calibration data. Any leakage of information from the calibration sample to the verification sample will bias the predictive skill estimate.

One way of generating forecasts that are independent of the calibration dataset is to split the whole period in to two parts, a calibration period and a verification period. An alternative method is 'leave k years out' cross-validation. In this study, cross validation is used because the 35-year data span was too short to divide into a reasonable training and verification periods. It can also be argued that cross-validation is closer to the operational situation, where all past data are utilised for each forecast.

To avoid any bias to the estimation of forecast skill arising from persistence, 3 years were withheld for validation: the target year, the year before and the year after. Both LDA and MLR were applied using each of the four sets of predictors (A, A′, B, B′), which meant that eight sets of forecasts were carried out for each homogeneous rainfall zone. In addition, LDA and MLR forecasts were carried out for Zone IV using predictor set C (equatorial Pacific only).

The models were assessed on their skill in making categorical forecasts. This forecast was carried out for both quintile and tercile categories. The observed rainfall totals were ranked and assigned to tercile (1–3) and quintile (1–5) categories with 1 being driest in both cases.

The output from the LDA forecasts was the probability of the total rainfall being in each category. The deterministic forecast from the MLR model was transformed to a probabilistic categorical forecast (similar to the form of the LDA) by converting errors in the regression model to a probability density function (pdf). In the pdf the threshold values for each categories (i.e. very dry to very wet) are computed from the forecast climatology. Then the probability for that particular category is obtained by computing the definite integral bounded by the threshold values for that category. The probablistic tercile forecast is converted to a categorical tercile forecast by selecting the highest probability as the forecast category.

The skill of the forecast models was assessed using two standard measures: relative operational characteristics (ROC) and ranked probability skill score (RPSS). The ROC score compares the skill of the forecasts against a random forecast whereas the RPSS compares the skill of the forecasts against a reference forecast (climatology in this case). Descriptions of these skill scores follow below.

The ROC method of assessing skill compares successful forecasts of an event (hits) to forecasts of an event that did not happen (false alarms). For a skill-less forecast, the hit rate (proportion of events successfully forecast) is equal to the false alarm rate (proportion of forecasts of an event that did not occur). Conversely, if the hit rate exceeds the false alarm rate then the forecast has some skill. In probabilistic forecasts, the ROC curve is a plot of a set of hit rates against the corresponding false alarm rates for different warning thresholds (Masson and Graham 1999). The ROC score (ROCS) is the area under the ROC curve. The skill of a forecast can be summarised by the ROCS. A zero skill forecast would have a score of 0.5 or less, and for a perfect forecast the ROCS will be 1.

The other method used in this study is the ranked probability skill score (RPSS). The ranked probability score (RPS) measures the squared difference between a categorical forecast and an observed category in probability space.

The RPSS is the RPS of the forecast compared to the RPS of the reference forecast. In this case the reference forecast is climatology, which assigns probabilities of 0.33 for terciles and 0.2 for quintiles to each of the categories. The mathematical derivation of RPS and RPSS are described in DBG08.

The ROC scores for tercile and quintile forecasts (averaged for all rainfall zones) are given in Tables 4 and 5, respectively, for Kiremt. It can be seen that the ROCS for Kiremt ranged from 0.58 to 0.93 for terciles and 0.51 to 0.91 for quintile forecasts. Thus, both tercile and quintile

Table 4 Summary of ROC scores for tercile forecasts for Kiremt

Predictors	Dry		Normal		Wet	
	MLR	LDA	MLR	LDA	MLR	LDA
Set A′	**0.91**	**0.87**	*0.76*	0.65	**0.93**	**0.88**
Set B′	**0.89**	**0.87**	0.72	0.64	**0.93**	**0.89**
Set A	**0.84**	**0.81**	**0.71**	0.61	**0.87**	**0.85**
Set B	**0.87**	**0.85**	0.63	0.58	**0.85**	**0.83**

The values are averages over all homogeneous rainfall zones except Zone V. Bold figures are significant at the 95% confidence level. Figures in italics are significant at the 90% confidence level

forecasts are better than a random forecast in all categories. Figure 4 reveals that both models show highest skill for northwestern Ethiopia (Zone I) and lowest for the southwestern part (Zone IIb). The low skill for Zone IIb (the southwest) reflects the fact that the correlation between rainfall and global SST is weaker compared to the other regions and the other dominant control on precipitation for this region is probably the Quasi Biennial Oscillation (QBO) (Diro 2008).

For Zone I, the skill remains stable when predictors are excluded from the contemporaneous season, whereas for

Table 5 Summary of ROC scores for quintile forecasts for Kiremt

Predictors	Very Dry		Dry		Normal		Wet		Very Wet	
	MLR	LDA	MLR	LDA	MLR	LDA	MLR	LDA	MLR	LDA
Set A′	**0.87**	**0.80**	**0.79**	0.72	0.71	0.55	*0.72*	0.62	**0.91**	**0.85**
Set B′	**0.85**	*0.84*	0.68	0.76	0.69	*0.72*	*0.70*	*0.72*	**0.90**	**0.84**
Set A	**0.85**	**0.83**	0.71	0.57	0.67	0.51	*0.73*	0.62	**0.82**	*0.76*
Set B	**0.81**	*0.80*	*0.70*	0.67	0.60	0.64	0.68	**0.71**	**0.84**	**0.81**

The values are averages over all homogeneous rainfall zones except Zone V. Bold figures are significant at the 95% confidence level. Figures in italics are significant at the 90% confidence level
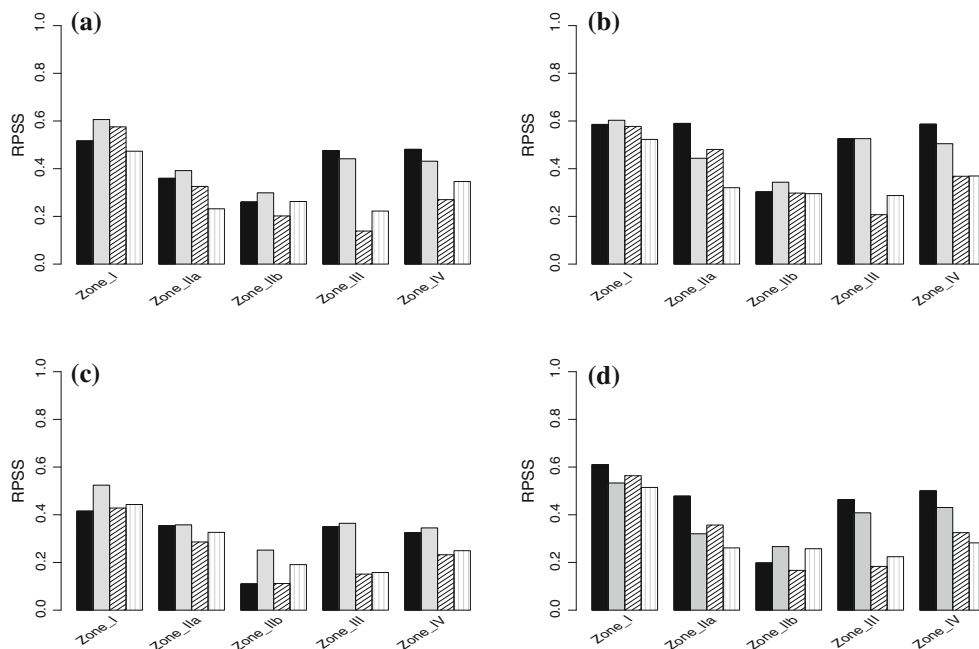

Fig. 4 RPSS for the Kiremt season: *top row* is for tercile forecasts, *bottom row* is for quintile forecasts, *lefthand* column is for LDA, *righthand* column is for MLR. The four *bars* represent the four sets of predictors. The *black shaded bar* is for the set of predictors selected by a stepwise regression and including contemporaneous season predictors. The *gray shaded bar* is for the predictor set selected by a stepwise discriminant analysis and including contemporaneous season predictors. The *diagonal hatched bar* is for the predictor set selected by a stepwise regression and excluding contemporaneous season predictors. The *vertical hatched bar* is for the predictor set selected by a stepwise discriminant analysis and excluding contemporaneous season predictors

central Ethiopia (Zone III), the skill dropped almost by a factor of two when contemporaneous season predictors were excluded from the model. This information is useful when considering operational forecasting methods. For zones in which the contemporaneous SSTAs make a significant improvement using a hybrid model, in which contemporaneous SSTAs are predicted by a GCM, has the potential to improve the skill.

It is noteworthy that forecasts with the highest skill were observed for the outer categories (very wet and very dry for quintile and wet and dry for tercile forecasts), whereas the inner (normal or near normal) categories had low skill. In fact, only the outer categories for both tercile and quintile forecasts had significant skill at the 95% level. This is a common feature in seasonal forecasting, probably because the predictive signal (for example La Niña/El Niño conditions) tends to be stronger for larger anomalies of the predictand (for example very dry or very wet conditions). Also the skill of the near normal category is likely to be smaller than the outer categories because the normal and near normal categories are bounded from both sides. A detailed explanation for the low skill of these categories is found in Kharin and Zwiers (2003).

A theme of this pair of papers is the relevance of physical mechanism for forecast skill. The observational studies and idealised modelling carried out in DGB10 showed that the equatorial Pacific is causally related to rainfall in all regions except for Zone I (where only the observational studies indicate a link). However, previous sections of this paper showed that, in most cases, the equatorial Pacific is not selected as a predictor by objective statistical methods. Section 3.1 proposed some explanations for this. As an example, and a test for the arguments given in Sect. 3.1, a statistical forecast using a fifth predictor set (set C), which comprised only the equatorial Pacific was implemented for Zone IV only.

The results are shown in Table 6 and Fig. 5 which compares regions in the full predictor set and equatorial Pacific only model. It is clear that the equatorial Pacific only forecasts have less skill than those utilising a full set
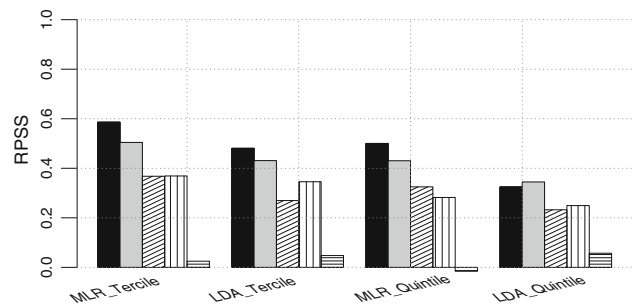

Fig. 5 RPSS for both LDA and MLR comparing different predictor sets for Zone IV. The five *bars* represent the various sets of predictors. The *black*, the *gray*, the *vertical* and *diagonal hatched bar* represent the same as in Fig 4. The *horizontal hatched bar* is when equatorial pacific is the only predictor

of objectively chosen predictors. This is consistent with Sect. 3.1, which argued that the non linear nature of the ENSO teleconnection together with the dynamics of the ENSO cycle would make the SSTAs in the equatorial Pacific a poor predictor-despite their causal link with the Ethiopian rainfall.

4 Summary and conclusions

Rainfall over Ethiopia has considerable spatial variability on both interannual and seasonal time scales. For this reason forecasts have been developed and evaluated for five homogeneous rainfall zones. Statistical forecasts were devised using two techniques [multiple linear regression (MLR) and linear discriminant analysis (LDA)], were applied to five sets of predictors (the first four selected by either stepwise regression or discriminant analysis and either including or excluding the contemporaneous season, and the last one just using the equatorial Pacific). Only SST anomalies were considered as predictors; linkages to other phenomena such as the Quasi-Biennial Oscillation (QBO) were not included in this analysis. The results show that forecasts that utilized a full objectively selected set of predictors had better skill than those based only on the equatorial Pacific.

When the full predictor sets were used, forecast skill was shown to be greater for extreme years (very rainy and very dry) than for normal and near normal years, for all forecast models and rainfall zones. This suggests that quintile forecasts would be more useful than the current tercile forecasts as they more clearly highlight extreme situations which are the most important to predict correctly.

The forecasts had greatest skill over northwest Ethiopia. Including contemporaneous predictors improved the skill in mosts zones, and by a up to a factor of two in central Ethiopia. This implies that more reliable forecasts could be

Table 6 ROCS for LDA and MLR models when different predictors set is used for Zone IV

ROC score for Zone IV		Dry		Normal		Wet	
		MLR	LDA	MLR	LDA	MLR	LDA
Predictors	Set A′	**0.92**	**0.89**	**0.81**	*0.76*	**0.96**	**0.89**
	Set B′	**0.88**	**0.87**	*0.78*	0.68	**0.95**	**0.89**
	Set A	*0.79*	*0.74*	*0.73*	0.68	**0.92**	**0.90**
	Set B	**0.89**	**0.88**	0.61	*0.72*	**0.85**	**0.81**
	ENSO only	*0.71*	*0.72*	0.26	0.22	*0.71*	*0.72*

Bold figures are significant at 0.05 level and figures with italics are significant at 0.1 level

made with a hybrid model which predicts the contemporaneous SST values as a first step. These can be then fed into the statistical model along with the observed SST of previous months. The skill of the forecast deteriorated when the model used equatorial Pacific SSTs as the only input. Therefore it is recommended that operational methods should include other oceanic areas rather than concentrating entirely on those associated with El Niño.

The results of this study suggest that the limitations of statistical prediction based on linear relationships can potentially be mitigated by blending objective selection of predictors with a more subjective approach. For instance, objective selection of predictors can circumvent the difficulties of applying a linear model to the non-linear climate system by identifying predictors that reflect non-linearity. There is, however, a possibility that a completely objective model will select predictors that have only a coincidental link with Ethiopian rainfall. However, it should become obvious over time that such spurious predictors do not provide useful information and they may be removed from the model. This underlines the importance of expert knowledge to monitor the performance of statistical models being used in an operational situation.

Gissila et al. (2004) commented on the feasibility of using a simple regression model to forecast Ethiopian rainfall operationally at a regional scale, concluding that so long as the appropriate data were available, in principle this type of model could be used by meteorological agencies. While this is no doubt true, this study, DGB10 and DBG08 have demonstrated that developing a statistical model requires a thorough understanding of the regional variability in Ethiopian climate and teleconnections. In particular, continuous monitoring of changes in the Ethiopian climate and its teleconnections would be necessary, to ensure that such a statistical model remains robust in the long term.

Appendix

Linear discriminant analysis

LDA is a method of classifying a response into discrete groups or categories by obtaining linear combination of predictors. A detailed mathematical derivation about linear discriminant analysis can be found in Afifi and Azen (1979). This method is implemented for a seasonal forecasting method by Mutai et al. (1998), Ward and Folland

(1991), Philippon et al. (2002) and DBG08. For a categorical forecast, LDA uses a Bayesian probability approach to estimate the posterior probability of a category given the predictor values. For a single predictor X, and prior probability $\text{Prob}(W_i)$ of a tercile or quintile category W_i, the posterior probability can be computed:

$$\text{Prob}(W_i|X) = \frac{\text{Prob}(W_i)\text{Prob}(X|W_i)}{\sum_j \text{Prob}(W_j)\text{Prob}(X|W_j)} \quad (1)$$

where, $\text{Prob}(W_i) \Rightarrow$ prior probability of category i in original population which is 0.333 for tercile and 0.2 for quintile by definition $\text{Prob}(X|W_i) \Rightarrow$ probability of observing the predictor value of X given the individual comes from category i (i.e. W_i)

$\text{Prob}(W_i|X) \Rightarrow$ posterior probability.

If there is more than one predictor, the multivariate distribution of predictors is used.

Assuming the predictors obey a multivariate normal distribution and the covariance of all categories are identical (i.e. $\sigma_1^2 = \sigma_2^2 = \ldots = \sigma_n^2 = \Sigma$), then the probability $\text{Prob}(X|W_i)$ can be computed from the density function $f(X|W_i)$ i.e.

$$\text{Prob}(X|W_i) \approx f(X|W_i) = \frac{1}{\sqrt{(2\pi)^2|\Sigma|}} e^{(X-\mu_i)^T \Sigma^{-1}(X-\mu_i)} \quad (2)$$

where, Σ is the covariance matrix of predictors and μ_i is the mean value of each predictor in those years when $\text{Prob}(W_i)$ is observed.

The linear discriminant analysis equation for multivariate situation can be derived by substituting Eq. 2 in to Eq. 1 to obtain

$$\text{Prob}(W_i|X) = \frac{e^{d_i(X)}}{\sum_j e^{d_j(X)}} \quad (3)$$

where, $d_i(X)$ is the discriminant score for the category W_i

$$d_i(X) = (\mu_i^T \Sigma^{-1})X - \frac{1}{2}\mu_i^T \Sigma^{-1}\mu_i + \ln(\text{Prob}(W_i))$$

Multiple linear regression

In a regression model, if standardised rainfall ($\mathbf{Y}$) is a response and SSTAs ($\mathbf{X}$) are predictors, the forecasting model can be expressed in matrix form as:

$$\mathbf{Y} = \boldsymbol{\beta}\mathbf{X} + \epsilon \quad (4)$$

where, $\mathbf{Y}$ is an ($n \times 1$) vector of predictand (rainfall), $\mathbf{X}$ is an ($n \times p$) matrix of predictors, $\boldsymbol{\beta}$ is a ($p \times 1$) vector of parameters, ϵ is an ($n \times 1$) vector of errors which are assumed to be independent and distributed normally with zero mean and variance of σ_o^2 i.e. $\epsilon \sim N(0, \mathbf{I}\sigma_\mathbf{0}^2)$. The number of years is denoted by n, and p is the number of predictors $+1$.

The deterministic forecast from the MLR model can be converted to a probabilistic categorical forecast (similar to the form of the LDA) by converting the prediction errors of the regression to a probability density function. The Probabilistic form of Regression can be written as:

$$\mathbf{Prob}(\mathbf{Y}|\mathbf{X}) \sim N(\hat{\mathbf{Y}}, \sigma^2) \tag{5}$$

where,

$$\hat{\mathbf{Y}} = \mathbf{X}\hat{\beta} \quad \text{and} \quad \sigma^2 = \sigma_o^2[1 + X_o^T(X^TX)^{-1}X_o] \quad \text{and}$$
$$\sigma_o^2 = \frac{\epsilon^T\epsilon}{n-p}$$

where ˆ means estimated value, σ^2 is the predicted variance, $\mathbf{X_o}$ is the vector of predictor variable and X is the matrix of training period predictor data.

References

Afifi A, Azen S (1979) Statistical analysis: a computer oriented approach, 2nd edn. Academic press, New York

Bekele F (1997) Ethiopian use of ENSO information in its seasonal forecasts. Internet J Afr Stud 2

Block P, Rajagopalan B (2007) Interannual variability and ensemble forecast of Upper Blue Nile Basin Kiremt season precipitation. J Hydrometeorol 8:327–343

Camberlin P, Philippon N (2002) The East African March–May rainy season: associated atmospheric dynamics and predictability over the 1968–1997 period. J Clim 15:1002–1019

Diro G (2008) Seasonal forecasting of ethiopian rainfall. PhD thesis, Department of Meteorology, Reading University, Reading

Diro G, Black E, Grimes D (2008) Seasonal forecasting of Ethiopian spring rains. Meteorol Applications 15:73–83

Diro G, Grimes D, Black (2010) Teleconnection between Ethiopian summer rainfall and sea surface temperature: part I. Observation and modelling. Clim Dyn. doi:10.1007/s00382-010-0837-8

Folland C, Palmer T, Parker D (1986) Sahel rainfall and worldwide sea temperatures 1901–1985. Nature 320: 602–607

Gissila T, Black E, Grimes D, Slingo J (2004) Seasonal forecasting of the Ethiopian summer rains. Int J Clim 24:1345–1358

Hastenrath S, Polzin D, Camberlin P (2004) Exploring the predictability of the 'short rains' at the coast of East Africa. Int J Climatol 24:1333–1343

Jan van Oldenborgh G, Balmaseda M, Ferranti L, Stockdale T, Anderson D (2005) Did the ECMWF seasonal forecast model outperform statistical ENSO forecast models over the last 15 years?. J Clim 18:3240–3249

Kassahun B (1987) Weather systems over ethiopia. In: Proceedings of First Tech Conf on Meteorological Research in Eastern and Southern Africa. Kenya Meteorological Department, Nairobi, Kenya, pp 53–57

Kharin V, Zwiers F (2003) On the ROC score of probability forecasts. J Clim 16:4145–4150

Korecha D, Barnston A (2007) Predictability of June–September rainfall in Ethiopia. Mon Weather Rev 135:628–650

Masson S, Graham N (1999) Conditional probability, relative operating characteristics and relative operating levels. Weather Forecast 14:713–725

Mutai C, Ward M, Colman A (1998) Towards the prediction of the East Africa short rains based on sea surface temperature-atmosphere coupling. Int J Climatol 18:975–997

Mwale D, Gan T (2005) Wavelet analysis of variability, teleconnectivity, and predictability of September–November East Africa rainfall. J Appl Meteorol 44:256–269

Ntale H, Gan T, Mwale D (2003) Prediction of East African seasonal rainfall using simplex canonical correlation analysis. J Clim 16:2105–2112

Philippon N, Camberlin P, Fauchereau N (2002) Empirical predictability study of October-December east African rainfall. Q J Royal Meteorol Soc 128:2239–2256

Rayner N, Parker D, Horton E, Folland C, Alexander L, Rowell D, Kent E, Kaplan A (2003) Global analysis of Sea surface temperature, sea ice and night marine air temperature since the late nineteenth century. J Geophys Res 108(D14)

Segele ZT, Lamb P (2005) Characterization and variability of Kiremt rainy season over Ethiopia. Meteorol Atmos Phys 89:153–180

Segele ZT, Lamb P, Leslie L (2009) Large-scale atmospheric circulation and global sea surface temperature associations with Horn of Africa June–September rainfall. Int J Climatol 29:1075–1100

Ward M, Folland C (1991) Prediction of seasonal rainfall in the north nordeste of Brazil using eigen vectors of sea surface temperature. Int J Climatol 11:711–743

Yeshanew A, Jury M (2007) North African climate variability. Part 3: resource prediction. Theoret Appl Climatol 89:51–62

Zhou G, Minakawa N, Githeko A, Yan G (2004) Association between climate variability and malaria epidemics in the East African highlands. Proc Nat Acad Sci USA 101(8):2375–2380