



# Lagged oceanic effects on the East African short rains

Erik W. Kolstad<sup>1</sup> · David MacLeod<sup>2</sup>

Received: 16 September 2021 / Accepted: 25 January 2022 / Published online: 15 February 2022  
© The Author(s) 2022

## Abstract

The East African ‘short rains’ in October–December (OND) exhibit large interannual variability. Drought and flooding are not unusual, and long-range rainfall forecasts can guide planning and preparedness for such events. Although seasonal forecasts based on dynamical models are making inroads, statistical models based on sea surface temperature (SST) precursors are still widely used, making it important to better understand the strengths and weaknesses of such models. Here we define a simple statistical forecast model, which is used as a tool to shed light on the dynamics that link SSTs and rainfall across time and space, as well as on why such models sometimes fail. Our model is a linear regression, where the August states of El Niño–Southern Oscillation (ENSO) and the Indian Ocean Dipole (IOD) predict about 40% of the short rains variability in 1950–2020. The forecast errors are traced back to the initial SSTs: too-wet (too-dry) forecasts are linked linearly to positive (negative) initial ENSO and IOD states in August. The link to the initial IOD state is mediated by changes in the IOD between August and OND, highlighting a physical mechanism for prediction busts. We also identify asymmetry and nonlinearity: when ENSO and/or the IOD are positive in August, the range and variance of OND forecast errors are larger than when the SST indices are negative. Upfront adjustments of predictions conditional on initial SSTs would have helped in some years with large forecast busts, such as the dry 1987 season during a major El Niño, for which the model erroneously predicts copious rainfall, but it would have exacerbated the forecast in the wet 2019 season, when the IOD was strongly positive and the model predicts too-dry conditions.

**Keywords** East Africa · Short rains · Climate prediction · ENSO · Indian Ocean Dipole

## 1 Introduction

Climate variability in East Africa is high, with floods and droughts significantly impacting lives and livelihoods (e.g., Little et al. 2001; Conway et al. 2005; Haile et al. 2019). The ‘short rains’ in the boreal autumn (Dunning et al. 2016; Nicholson 2017) in particular show substantial interannual variability, with extreme impacts from both flooding (e.g., in 2019; Chang’a et al. 2020; Nicholson et al. 2021; Wainwright et al. 2021) and drought (e.g., in 2010; Dutra et al. 2013).

A strong association between El Niño–Southern Oscillation (ENSO) and the short rains has been known for decades

(e.g., Ropelewski and Halpert 1987), and the key importance of the Indian Ocean for the short rains also has a long history (e.g., Nyenzi 1988; Beltrando and Camberlin 1993; Hastenrath et al. 1993). The Indian Ocean Dipole (IOD; Saji et al. 1999), which peaks around September–November (e.g., Zheng et al. 2021), is an oceanic driver of the short rains (Latif et al. 1999; Black et al. 2003; Black 2005; Bahaga et al. 2019), specifically through its role in a local east–west-oriented Walker cell (e.g., Goddard and Graham 1999; Hastenrath et al. 2004; Behera et al. 2005; Ummenhofer et al. 2009; Nicholson 2015; Zhao and Cook 2021). Numerous studies have shown that ENSO in boreal summer leads the IOD in boreal autumn (e.g., Stuecker et al. 2017; McKenna et al. 2020), suggesting a mainly indirect role of ENSO in the context of predicting the short rains. Some studies have estimated the spatial patterns of rainfall anomalies associated with events where both ENSO and the IOD are active, compared to ENSO-only events; the latter showing a weaker and less homogenous rainfall response (MacLeod and Caminade 2019; MacLeod et al. 2021).

✉ Erik W. Kolstad  
ekol@norceresearch.no

<sup>1</sup> NORCE Norwegian Research Centre, Bjerknes Centre for Climate Research, Bergen, Norway

<sup>2</sup> School of Geographical Sciences, University of Bristol, Bristol, UK

Seasonal forecasts can play a central role in planning and preparedness (e.g., Lemos et al. 2012; Bruno Soares et al. 2018; Tall et al. 2018; Bazo et al. 2019; Nyamekye et al. 2021). Forecast skill is underpinned by seasonal predictability arising from teleconnections with large-scale drivers, and with a strong link to ENSO and the IOD, short rains forecasts generally show high levels of skill at long lead times (e.g., Mwangi et al. 2014; Bahaga et al. 2015; MacLeod 2019). The season is then a prime candidate for forecast-based early actions, and seasonal forecasts have been provided at the Greater Horn of Africa Regional Climate Outlook Forum (GHACOF) meetings for over a decade (Hansen et al. 2011; Walker et al. 2019), with the outlook for October–December (OND) usually produced in late August.

The GHACOF forecasts are issued by the Intergovernmental Authority on Development (IGAD) Climate Prediction and Applications Centre (ICPAC) and produced in collaboration with the National Meteorological and Hydrological Services (NMHS) of 11 countries, from Sudan in the north to Tanzania in the south. Previously, the GHACOF forecasts were the outcome of an opaque ‘consensus’ process, which included the use of dynamical forecast models, statistical regression models, and ‘analogue years’ based on the current tropical SST anomalies (mainly ENSO and the IOD). The consensus process has had a ‘tendency to over-forecast the near normal category of rainfall’ (Walker et al. 2019), which is unfortunate because it leads to under-prediction of lower- and upper-tercile events. In an effort to mitigate this problem, the production of the forecasts has gradually evolved towards being based on ‘objective’ dynamical model forecasts in recent years, but there is an understandable reluctance to fully embrace dynamical models.

In fact, conversations that we have had with NMHS representatives in East Africa reveal that methods based on SST regressions and analogue years are still widely used in the region. It is therefore important to study the performance of statistical forecast models in predicting the short rains, to learn more about the dynamical pathways linking August SST anomalies and OND rainfall, as well as assessing the years when statistical forecasts fail. This is the purpose of the study described here, which we hope will contribute to a more informed use of statistical models.

We note here that many previous studies have described regression-based short rains prediction models based on SST indices and atmospheric precursors (e.g., Mutai et al. 1998; Philippon et al. 2002; Ntale et al. 2003; Hastenrath et al. 2004; Nicholson 2014). The skill of several of these models is higher than the skill of the model that we present here. Our purpose is not to create the best possible statistical prediction model, but rather to study a simple SST-based model to try to understand the strengths and limitations of that model.

Specifically, we seek to understand the ability of climate conditions (ENSO and IOD) which are present at the time of the late August GHACOF to anticipate OND rainfall, in order to diagnose those climate states which are (and are not) a useful indication of upcoming climate variability. We do this through the use of a linear regression model based on reanalysis data from 1950 to 2020, relating ENSO and IOD to rainfall and variables that describe the Walker circulation over the Indian Ocean. We then harness the errors of the reanalysis-based linear model as a diagnostic tool, to provide insight into the relationship between August SSTs, the OND Walker circulation, and the short rains. Changes in the SST forcing between the prediction time in August and the OND itself is investigated as a possible source of errors. Furthermore, we study the relationship between OND forecast errors and precursor SST conditions in August to determine if there exist certain conditions which can indicate a priori (upfront) the level of uncertainty in a seasonal forecast. We recently used a similar approach to show that the ECMWF monthly forecasting system has a systematic conditional short rains bias contingent on the initial IOD state (Kolstad et al. 2021).

Studies of conditional forecast model errors and biases such as the one described herein are relevant for both dynamical forecasts initialised with SST fields which have had notable forecast busts in recent years (Kilavi et al. 2018; MacLeod and Caminade 2019), and for statistical forecasts based on precursor SST conditions. They may also guide the use of analogue forecasts which produce an outlook based on diagnoses of historical seasons with similar SST conditions.

## 2 Data and methods

### 2.1 Data and detrending

ERA5 reanalysis (Hersbach et al. 2020c) monthly mean data from 1950 to 2020 are used throughout. The following variables are analysed: precipitation, SST, vertical velocity at 500 hPa ( $w_{500}$ ; positive upwards), and zonal wind at 850 hPa ( $u_{850}$ ). The SSTs in ERA5 are prescribed from independent datasets. To compensate for trends during the study period, the SST and precipitation data from ERA5 are linearly detrended. As the data for the other variables have less consistent trends, these are not detrended.

For precipitation, datasets which are based on direct and satellite-derived observations exist, such as the Climate hazards infrared precipitation with stations (CHIRPS) dataset (Funk et al. 2015). A comparison between ERA5 and CHIRPS in East Africa has indicated that ERA5 is slightly wetter than CHIRPS in October and November, particularly near the Equator (Gleixner et al. 2020). ERA5 also exhibits a stronger drying trend from the 1980s and onwards than

CHIRPS, but this appears to be a bigger problem in Central Africa than in our focus region. Nevertheless, any spurious ERA5 precipitation trends are removed from our analysis by first detrending the data. The main reason for using ERA5 precipitation data is that we gain 31 years for the analysis, as the CHIRPS period starts in 1981. To assess the differences between ERA5 and CHIRPS, we calculated the Empirical Orthogonal Functions (EOFs; see Sect. 2.3) of East African precipitation for both datasets, and the interannual correlation between the first principal components was 0.97 in the overlapping period between 1981 and 2020. (Note that we included oceanic grid points in the calculation of the ERA5-based EOFs, while CHIRPS is only defined over land.) Although there are doubtless differences between ERA5 and CHIRPS on the local scale and daily time scales, the two data sources give highly similar results on the aggregated temporal and spatial scales studied here.

### 2.2 SST indices

Two SST indices are used, and these are based on standard regions (e.g., Stuecker et al. 2017): the NINO3.4 index (referred to as N34 from now), calculated as area-averaged SST anomalies from 170° W to 120° W and between 5° S and 5° N; and the IOD dipole mode index (DMI hereafter), computed as the difference between area-averaged SST anomalies in the western (50° E to 70° E and 10° S to 10° N) and eastern (90° E to 110° E and 10° S to 0°) Indian Ocean. The detrending was performed after the indices had been computed, and then the N34 and DMI time series were standardized.

### 2.3 EOF analysis

The rainfall analysis is based on EOFs, which reduce vast amounts of data to a limited number of interannual time series. We first calculate detrended ERA5 OND precipitation anomalies for each grid point (irrespective of the land mask) from 10° S to 12° N and 30° E to 52° E—the same region used by Vigaud et al. (2017) and Kolstad et al. (2021). This gives us 71 precipitation maps, which are used as input to the *eofs* software package for Python (Dawson 2016).

### 2.4 Linear prediction

We predict several variables during OND based on SST indices in August. The choice of August as the initial month is motivated by the timing of the GHACOFs addressing the OND season, which are usually held in late August. In addition to predicting the rainfall, we also predict certain variables related to the Walker circulation, namely SST,  $w_{500}$ , and  $u_{850}$ , all in OND. Naming the SST indices  $X_1$  (N34) and

$X_2$  (DMI), we define a regression that describes their lagged relationship with any variable  $Y$ :

$$Y = c_1 X_1 + c_2 X_2 + \rho, \tag{1}$$

where  $\rho$  is the residual. The non-calibrated prediction of  $Y$  is written as:

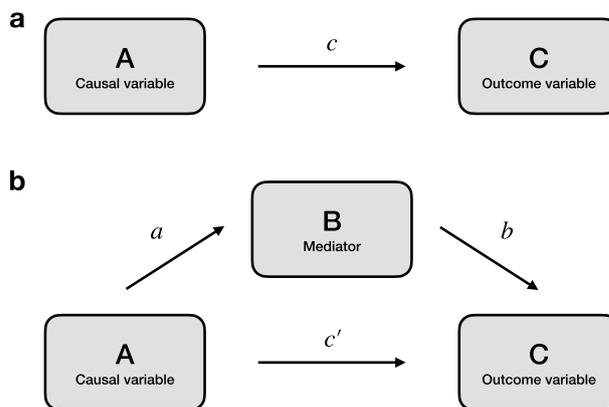
$$\hat{y} = \hat{c}_1 X_1 + \hat{c}_2 X_2. \tag{2}$$

We estimate  $\hat{c}_i$  separately for each year by using the time series of  $Y$ ,  $X_1$  and  $X_2$  for all the other years; this method is usually referred to as ‘out-of-sample’ or ‘leave-one-out’ cross-validation. (Note that when predicting the principal components, we do not compute out-of-sample EOFs for each year, as this might for some years create different spatial patterns and principal components with opposite signs; we just leave the in-sample elements of the principal components out in the prediction.) To scale the prediction, we divide  $\hat{y}$  by its out-of-sample standard deviation and multiply by the out-of-sample standard deviation of  $Y$  to obtain a calibrated prediction:

$$\hat{Y} = \sigma(Y)\hat{y}/\sigma(\hat{y}). \tag{3}$$

### 2.5 Mediation

If a ‘causal variable’  $A$  is significantly correlated with an ‘outcome variable’  $C$ , we can express this relationship as  $A \rightarrow C$ . However, the direct effect of  $A$  on  $C$  may be mediated by a variable  $B$ . This can be expressed as  $A \rightarrow B \rightarrow C$ , and the validity of the pathway is checked by going through four ‘steps’ (e.g., Baron and Kenny 1986), using regressions illustrated in Fig. 1.



**Fig. 1** **a** A diagram illustrating the direct effect  $c$  of the causal variable  $A$  on the outcome variable  $C$ . **b** A diagram illustrating the direct effect  $a$  of  $A$  on the mediator  $B$ , as well as the regression coefficients of Eq. 4,  $b$  and  $c'$

First, the coefficient  $c$  in Fig. 1a, which describes the direct effect of  $A$  on  $C$ , must be significant. Otherwise, there is no effect to be mediated. Second, the coefficient  $a$  in Fig. 1b, the direct effect of  $A$  on  $B$ , must also be significant to ensure that the mediator is influenced by the causal variable. To check the remaining steps, we define a linear regression:

$$C = c'A + bB, \quad (4)$$

The third step is that  $b$  must be significant, which means that  $B$  has an effect on  $C$  when accounting for  $A$ . The fourth requirement for mediation is that the coefficient  $c'$  must be non-significant, as this signifies that the outcome  $C$  is conditionally independent of  $A$  when accounting for  $B$ . If all the four steps are satisfied, the mediator  $B$  fully mediates the effect of  $A$  on  $C$ . We emphasize here that full mediation by  $B$  does not necessarily mean that  $B$  is the only mediator (e.g., VanderWeele and Vansteelandt 2014).

## 2.6 Significance testing

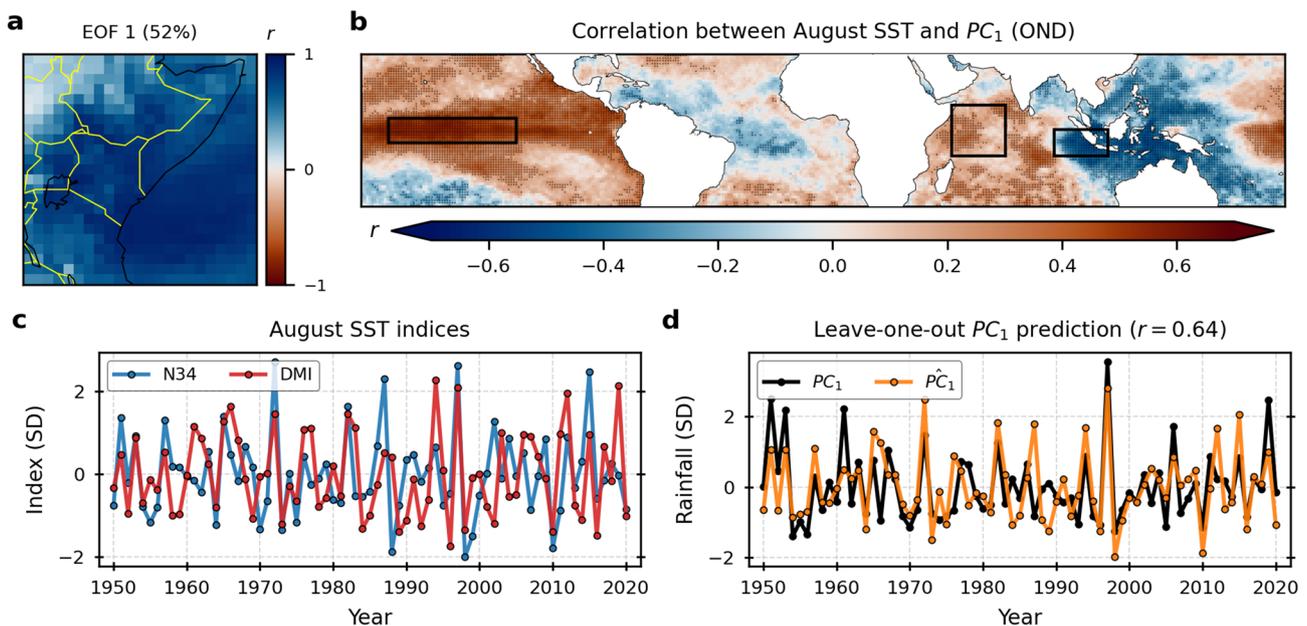
Bootstrapping is used to calculate statistical significance. In each case, a set of 1,000 artificial time series is produced by replacing each datum with the datum from a random year between 1950 and 2020 (with replacement). Correlations or regression coefficients are deemed significant at the 5% level (which is used throughout) if they are either less than

the 2.5th percentile of the bootstrapped set, or greater than its 97.5th percentile.

## 3 Results

### 3.1 Lagged relationships between SST and rainfall variability

In Fig. 2a, the loading pattern of the leading EOF of East African rainfall in OND is shown, expressed as the interannual correlation between the detrended precipitation and the leading principal component ( $PC_1$  hereafter, defined uniquely for OND) in each grid point. EOF 1 describes 52% of the rainfall variance and represents large-scale precipitation anomalies affecting most of the region. This agrees well with the leading EOF in both Schreck and Semazzi (2004) and Wenhaji Ndomeni et al. (2018), both of whom used other rainfall data sets and reference periods, and slightly different region boundaries. The highest correlations are also in good qualitative agreement with the region with a biannual rainfall regime and October/November onset in Dunning et al. (2016; their Fig. 5d). On the interannual scale, the time series of  $PC_1$  is highly correlated with a time series of area-averaged detrended precipitation anomalies inside the study region, using land points only ( $r = 0.95$ ), confirming that EOF 1 is a suitable proxy for the large-scale rainfall in



**Fig. 2** **a** The loading pattern of the leading EOF of East African rainfall, with country borders in yellow. The whole region used to compute the EOFs is shown. **b** Interannual correlation between SSTs in August and  $PC_1$ , with significant correlations indicated with dots and

the boundaries of the N34 and IOD regions shown as rectangles. **c** Detrended interannual time series of August N34 and DMI; **d** Time series of  $PC_1$  and its prediction  $\hat{PC}_1$

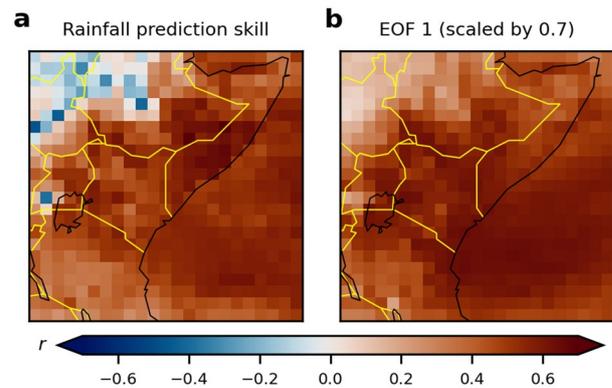
East Africa.  $PC_1$  is also significantly correlated with both N34 ( $r = 0.58$ ) and DMI ( $r = 0.73$ ) in OND.

In Fig. 2b, we show the spatial structure of the lagged correlation between tropical SSTs in August and  $PC_1$ . Inside the NINO3.4 region, the correlations are positive, and the oppositely signed correlations in the two IOD regions are consistent with a positive lagged correlation with DMI. Indeed, in index form  $PC_1$  is significantly lag-correlated with both N34 ( $r = 0.56$ ) and DMI ( $r = 0.55$ ) in August. There are indications of significant correlations outside the N34 and DMI regions, such as in the Atlantic, but the two indices appear to represent a large share of the interannual correlations.

Figure 2c shows standardized interannual time series of the detrended August SST indices. The correlation between the two indices is 0.45. In some years both indices have large values with the same sign, such as during the major El Niño episodes in 1972 and 1997 and the La Niña events in 1998 and 2010. Other years, including 2019, have large DMI absolute values but weak ENSO conditions.

The time series for  $PC_1$  is shown in Fig. 2d, along with its (leave-one-out) prediction  $\widehat{PC}_1$ , where Eqs. 1–3 are used. The Eq. 2 regression coefficients are both significant and similar in magnitude ( $\widehat{c}_1 = 0.40$  and  $\widehat{c}_2 = 0.37$ ). The correlation between  $\widehat{PC}_1$  and  $PC_1$  is 0.64, and although this may be a slight overestimation of the actual skill (since the EOFs were not computed out-of-sample; see Sect. 2.4), the high and significant correlation demonstrates substantial predictive power on the seasonal time scale: about 40% of the interannual variability of the first principal component of OND rainfall is accounted for by the two SST indices in August, 1–2 months before the rainy season onset.

Having demonstrated a temporal link between August SST and East African rainfall (Fig. 2d), we now show the spatial dependence of this link across East Africa in Fig. 3. First, we predict rainfall for each grid point based on the two SST predictors, using the same leave-one-out cross-validation methodology that we used to predict  $PC_1$  (i.e., Eqs. 1–3). Figure 3a shows the correlation between this predicted field and ERA5 rainfall for each grid point. Second, in Fig. 3b we show the loading pattern of the leading EOF (this is the same as Fig. 2a except that we here inverted the colours and scaled by 0.7 for easy comparison with Fig. 3a). The similarity between the two panels of Fig. 3 shows that the parts of East Africa that are most affected by the leading mode of variability (EOF 1) are also the ones that are most skilfully predicted by the August SST indices. It is not obvious that the patterns should be similar, as the SST indices do not enter in the calculation of the EOFs. However, this result is consistent with physical understanding, since previous work has shown the spatial expression of the N34 and DMI influences on East Africa to be strongest along the coast of Somalia, Kenya and southeast Ethiopia (MacLeod



**Fig. 3** **a** The interannual correlation between East African rainfall and the rainfall prediction (using leave-one-out cross-validation) based on August N34 and DMI, with country borders in yellow. **b** The same EOF pattern as the one shown in Fig. 2a, but scaled by 0.7 and with inverted colours

**Table 1** Eq. 2 regression coefficients for initial SST indices from June to October

Predictor month	N34 coefficient ( $\widehat{c}_1$ )	DMI coefficient ( $\widehat{c}_2$ )	Correlation between $\widehat{PC}_1$ and $PC_1$
June	0.29	0.29	0.48
July	0.43	0.30	0.60
August	0.40	0.36	0.64
September	0.30	0.48	0.68
October	(0.16)	0.61	0.71

The number in parentheses indicates a non-significant coefficient, and the remaining correlations are significant. The last column shows the correlation between the predicted and actual rainfall index

et al. 2021). We expect predictions based on SST indices alone to be most skilful in this region. These drivers are also known to dominate the variance of OND rainfall, and thus we expect the first EOF to also reflect their influence.

As mentioned, the zero-lag correlation with OND rainfall is higher for DMI than for N34, whereas for the August SST indices, the N34 coefficient  $\widehat{c}_1$  is slightly larger than the DMI coefficient  $\widehat{c}_2$ . We now investigate how the roles of the indices change when the initial SST state is taken from different months. Table 1 shows that the relative importance of N34 with respect to DMI is highest for July initial states, and thereafter it declines (recall that the IOD peaks in September–November). For October initial states, the N34 coefficient is non-significant, which demonstrates that it is the IOD that directly influences the East African rainfall at short lags.

Furthermore, a mediation analysis of the effect of August N34 on  $PC_1$  shows that this effect is fully mediated by DMI in OND. Using the notation in Sect. 2.5, we let the causal

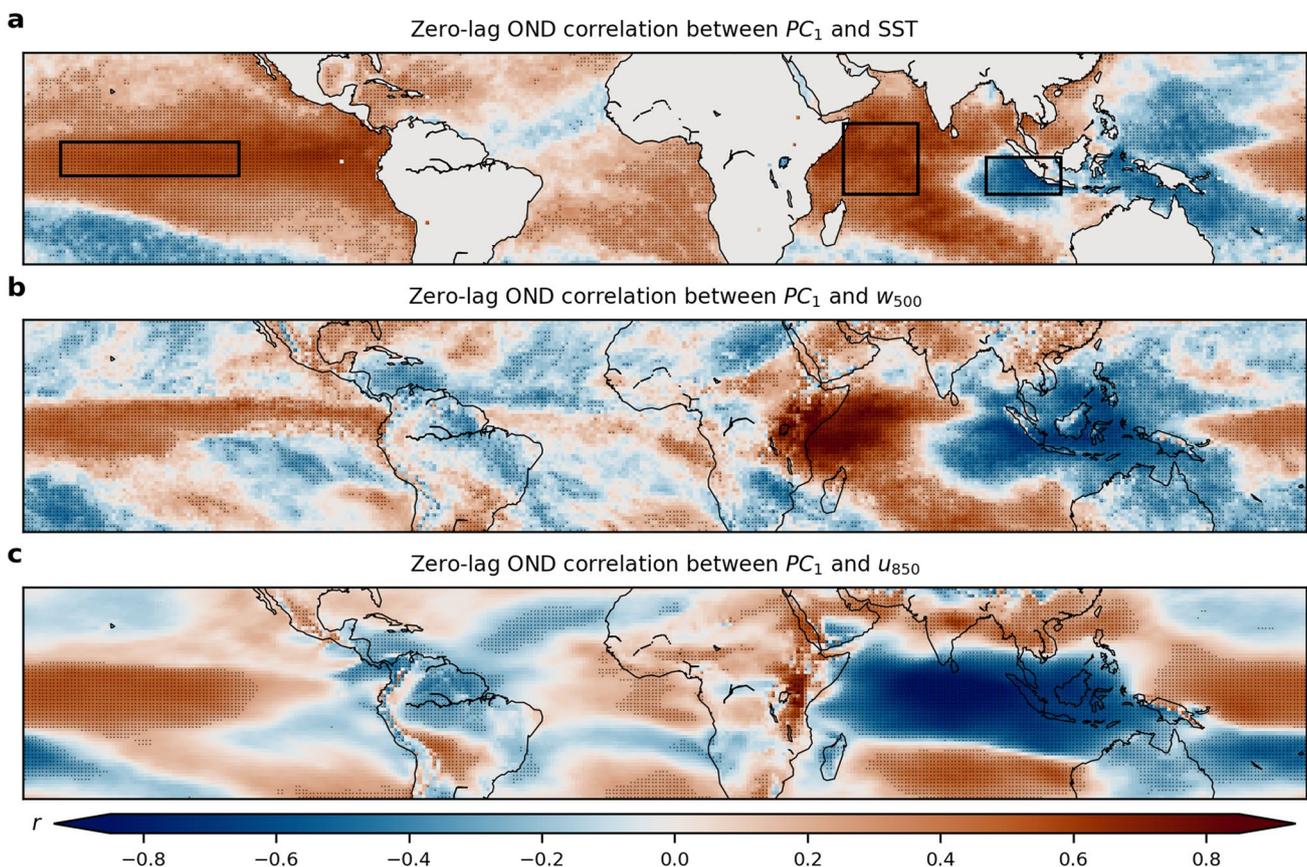
variable  $A$  be N34 in August, the mediator  $B$  is DMI in OND, and the outcome  $C$  is  $PC_1$ , and then we go through the four steps required for mediation sequentially. First, as mentioned earlier,  $c$  is significant (the correlation between August N34 and  $PC_1$  is 0.56). Second, the coefficient  $a$  is also significant with a value of 0.65, representing the ENSO-leading-IOD relationship (Stuecker et al. 2017). Third, the coefficient  $b$  in Eq. 4 is significant with a value of 0.63. Fourth, the coefficient  $c'$  is non-significant, which shows that  $PC_1$  is independent of August N34 given DMI in OND.

### 3.2 Dynamical pathways

We now investigate  $PC_1$  and its prediction  $\widehat{PC}_1$  in more detail. Our main purpose is not to validate the prediction model, but rather to use that model as a tool to better understand the dynamical linkages between August SSTs and large-scale OND rainfall. Our approach is to investigate when and why the linear model fails: when  $\widehat{PC}_1$  deviates from  $PC_1$ . There are linear aspects to these failures, but we are also interested in nonlinear aspects, and especially large

deviations between  $\widehat{PC}_1$  and  $PC_1$ . For brevity, we refer to the forecast error as  $\varepsilon \stackrel{\text{def}}{=} \widehat{PC}_1 - PC_1$ .

Before we study the forecast errors, we look at how  $PC_1$  relates to concurrent anomalies of SST,  $w_{500}$ , and  $u_{850}$ . Figure 4a shows that rainfall and SST anomalies are positively correlated in the western part of the Indian Ocean, and negatively correlated in the eastern part. This is consistent with the positive correlation between DMI and East African rainfall. The positive correlations in the eastern Pacific, including in the NINO3.4 region, are probably mainly due to teleconnections with the IOD region (see Sect. 3.1). For the vertical velocity, a large area over the western Indian Ocean and East Africa has high positive correlations with  $PC_1$ , with negative correlations in evidence over the eastern Indian Ocean (Fig. 4b). In Fig. 4c, the low-level zonal wind is shown to be negatively correlated with  $PC_1$  over the whole central Indian Ocean. In sum, the correlations over the Indian Ocean in all three panels of Fig. 4 are consistent with a positive IOD state and an enhanced Walker circulation during periods with higher-than-normal rainfall, and



**Fig. 4** Interannual correlation between  $PC_1$  and the OND anomalies of the following variables: SST (a); 500 hPa vertical velocity (b); and 850 hPa zonal wind (c). Significant correlations are indicated with

dots. In (a), the boundaries of the IOD and ENSO regions are shown as rectangles

a negative IOD state and a suppressed Walker circulation during drier-than-normal periods.

We now use Eqs. 1–3 to predict SST, vertical velocity, and zonal wind in OND, based on the two SST indices in August, just as we predicted  $PC_1$ . This is done for all grid points individually, and then we compute the correlation between the forecast errors  $\hat{Y} - Y$  and  $\varepsilon$ , where  $Y$  is SST,  $w_{500}$ , and  $u_{850}$ , in turn. These correlations are shown in Fig. 5, along with yellow contours representing the correlations that were shown in Fig. 4. It is immediately clear that the overall patterns in Fig. 5 are highly similar to the patterns in Fig. 4, and especially over the Indian Ocean. When too much rainfall is predicted, too-positive SST anomalies are predicted in the western Indian Ocean (Fig. 5a), with too-positive overlying  $w_{500}$  anomalies (Fig. 5b), and too-negative  $u_{850}$  anomalies are predicted over the Indian Ocean (Fig. 5c). In the eastern part of the Indian Ocean, too-negative SST anomalies and too-negative  $w_{500}$  anomalies complete the Walker circulation signature.

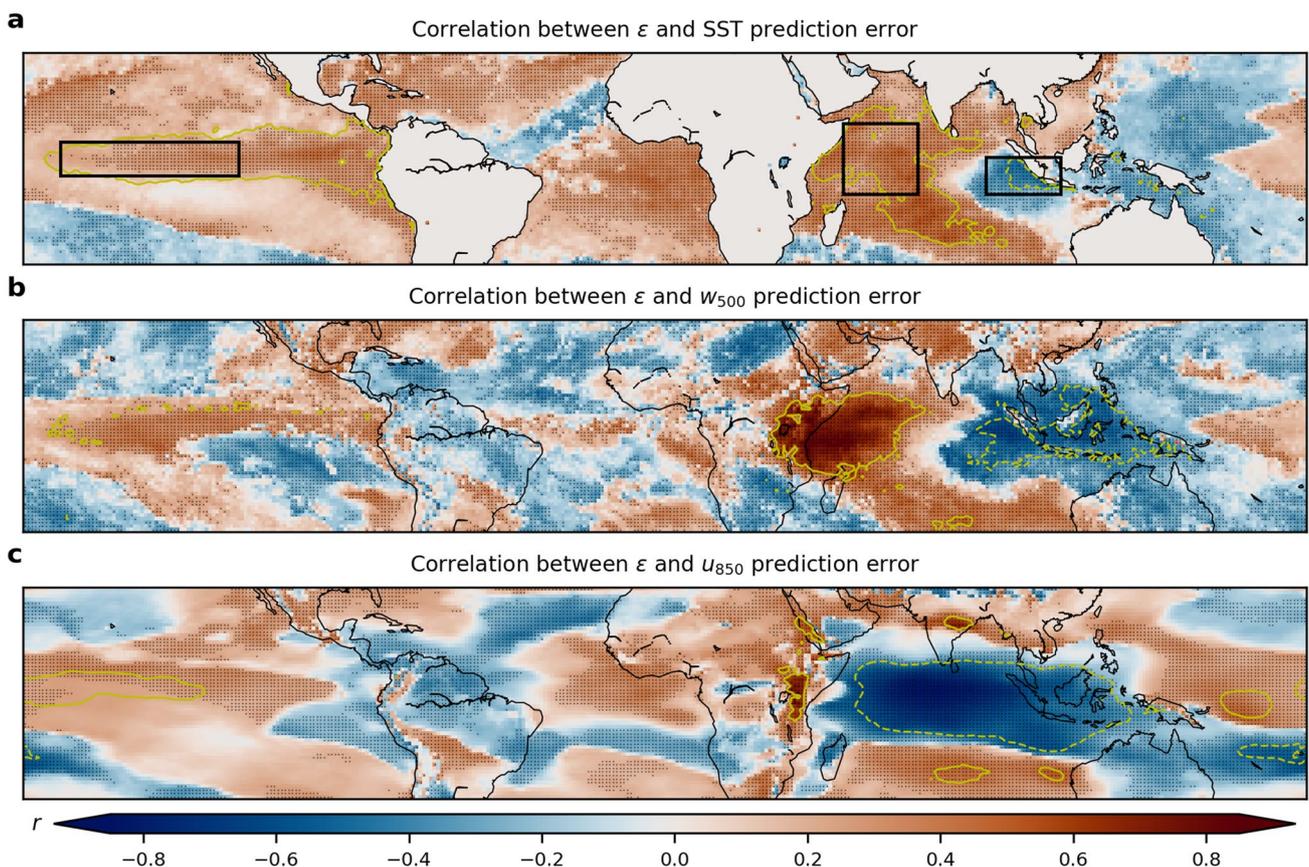
In other words, when the linear East African rainfall prediction model fails, similar linear models fail to predict

SST, vertical velocity, and low-level zonal wind anomalies in the areas where said variables are correlated with East African rainfall. One interpretation of this finding is that the SST indices in August usually predict the Walker circulation in OND quite well, and hence also East African rainfall, but when the Walker circulation deviates from the expected ‘response’ to the August SSTs, the OND rainfall prediction fails.

What could drive an unexpected OND Walker circulation response to August SSTs and hence an error in the rainfall prediction based on August SST indices? One possibility, which we investigate in the next section, is that the SSTs in the Indian Ocean evolve substantially between August and OND, which again might lead to different SST forcing to the one predicted by the linear model.

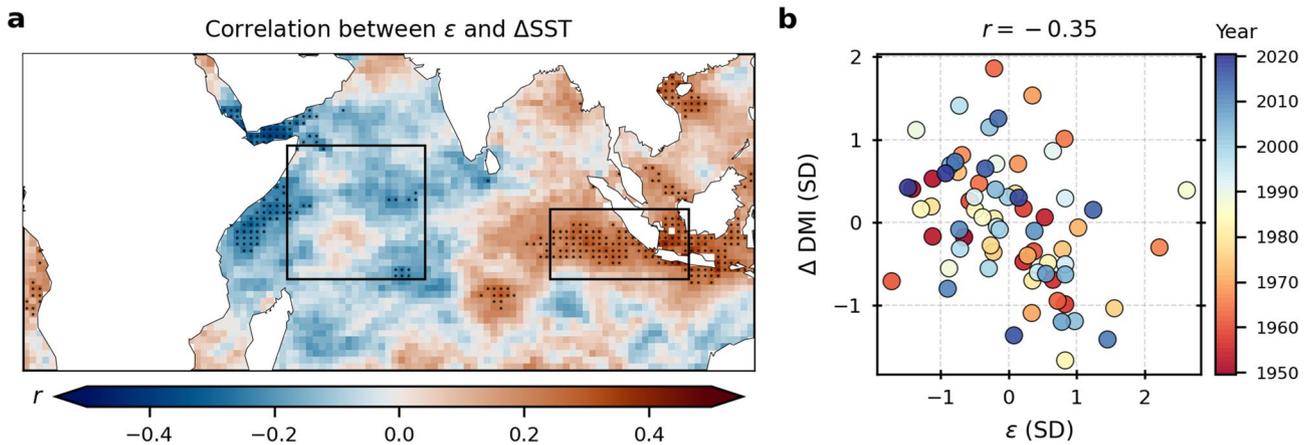
### 3.3 Role of SST changes

The map in Fig. 6a shows the correlation between  $\varepsilon$  and  $\Delta$ SST (defined as the change in SST from August to OND) in the Indian Ocean. As  $\Delta$ SST in the N34 region is not



**Fig. 5** Interannual correlation between  $\varepsilon$  and the OND prediction error of the following variables: SST (a); 500 hPa vertical velocity (b); and 850 hPa zonal wind (c). Significant correlations are indicated with dots, and the yellow contours show the +0.5 (solid) and -0.5

(dashed) isolines for the correlation between  $PC_1$  and OND anomalies of each variable. In (a), the boundaries of the IOD and ENSO regions are shown as rectangles



**Fig. 6** **a** Interannual correlation between  $\Delta$ SST (change in SST between August and OND) and  $\epsilon$ . Significant correlations are indicated with dots, and the rectangles indicate the outlines of the IOD

regions. **b** Scatterplot of  $\epsilon$  on the  $x$ -axis versus  $\Delta$ DMI on the  $y$ -axis (both in standard deviations). The colours of the circles correspond to the years

significantly correlated with  $\epsilon$ , the Pacific region is not shown. The general pattern is that when too much rainfall is predicted, it is because the western part of the Indian Ocean cools between August and OND, while the eastern part warms. This pattern is compatible with a de-intensification of the Walker circulation, which again is consistent with decreased East African rainfall. In other words, if DMI decreases, the actual rainfall tends to be less than what the model predicts. Conversely, when the prediction is too dry, it is linked to a Walker circulation intensification due to SST warming in the west and cooling in the east, which is linked to actual rainfall greater than the predicted rainfall.

Although the significant blue areas in Fig. 6a in the western Indian Ocean are concentrated to the west of the western IOD region, there is clearly a negative correlation between changes in DMI and the rainfall forecast error. In Fig. 6b,  $\Delta$ DMI (the change in DMI from August to OND) is plotted against  $\epsilon$  for each year. The correlation is  $-0.35$ , which is statistically significant. In all the years where DMI decreases by more than one standard deviation, the rainfall forecast error is positive. However, there are years with large positive forecast errors where DMI does not decrease notably. In the El Niño year of 1987 (see Fig. 2c), the largest positive  $\epsilon$  value is found, but DMI in fact increased slightly between August and OND. We also note that large positive  $\Delta$ DMI values are not uniformly linked to negative rainfall forecast errors. For instance, the year with the second-largest DMI increase (1969),  $\epsilon$  is positive.

In sum, the correlation of  $-0.35$  (Fig. 6b) shows that about 10% of the rainfall forecast error can be linked linearly to DMI changes. However, it is not generally possible to know in advance how SSTs will change between August and OND, although dynamical climate models may have some skill, and a hybrid statistical–dynamical forecasting

approach appears to be promising (Colman et al. 2020). It would be more useful if we could identify a relationship between the initial state in August and the subsequent prediction errors in OND, as these could potentially make it possible to foresee a priori whether the linear model is more likely to fail.

### 3.4 Role of initial SSTs

First, to get a qualitative overview of the most severe prediction busts, we study the seven years (roughly 10% of the 71 years in the study period) with the largest negative errors, as well as the seven years with the largest positive errors, i.e., when the model is too dry and too wet, respectively. For these years,  $PC_1$ , its prediction  $\widehat{PC}_1$ , the error  $\epsilon$ , and the August SST indices are listed in Table 2.

For some of the years, the errors are large because  $\widehat{PC}_1$  has the wrong sign, such as in 1952, for which the model predicts drier-than-average conditions ( $\widehat{PC}_1 = -0.7$ ) but the actual rainfall was wetter than average ( $PC_1 = 0.5$ ). In the El Niño year of 1987, the model predicts heavy rain ( $\widehat{PC}_1 = 1.8$ ), yet the actual conditions were drier than normal ( $PC_1 = -0.8$ ). In other cases,  $\widehat{PC}_1$  has the right sign, but the magnitude is incorrect, such as in 2019, when the IOD was strongly positive, for which  $\widehat{PC}_1 = 1.0$ , but the actual rainfall was extremely high (Wainwright et al. 2021):  $P = 2.5$ .

The too-dry cases shown in Table 2 have no consistent August N34 and DMI sign; across the seven years the average state is neutral for both SST indices. Correspondingly the rainfall prediction  $\widehat{PC}_1$  is also near-zero on average, whilst observed rainfall is positive in all years, with a strongly positive average  $PC_1$  value (1.2). The situation is different for the seven too-wet cases. Here there is a

**Table 2** The seven years in each category (too-dry and too-wet), ranked by the rainfall forecast error magnitude

Too-dry forecasts						Too-wet forecasts					
Year	$\epsilon$	$PC_1$	$\widehat{PC}_1$	August N34	August DMI	Year	$\epsilon$	$PC_1$	$\widehat{PC}_1$	August N34	August DMI
1961	-1.7	2.2	0.5	-0.1	1.2	1987	2.6	-0.8	1.8	2.3	0.5
2019	-1.5	2.5	1.0	0.0	2.1	1966	2.2	-1.0	1.3	0.5	1.6
1951	-1.4	2.5	1.1	1.4	0.5	1976	1.6	-0.7	0.9	0.4	1.1
1989	-1.4	0.1	-1.3	-0.8	-1.4	2012	1.4	0.2	1.7	0.9	2.0
1984	-1.3	0.2	-1.1	-0.5	-1.3	2015	1.2	0.8	2.1	2.5	1.0
1978	-1.1	0.6	-0.5	-0.1	-0.8	1972	1.0	1.5	2.5	2.7	1.5
1952	-1.1	0.5	-0.7	-0.2	-1.0	2003	1.0	-0.4	0.5	-0.1	1.0
<b>Average</b>	<b>-1.4</b>	<b>1.2</b>	<b>-0.1</b>	<b>-0.1</b>	<b>-0.1</b>	<b>Average</b>	<b>1.6</b>	<b>-0.1</b>	<b>1.5</b>	<b>1.3</b>	<b>1.2</b>

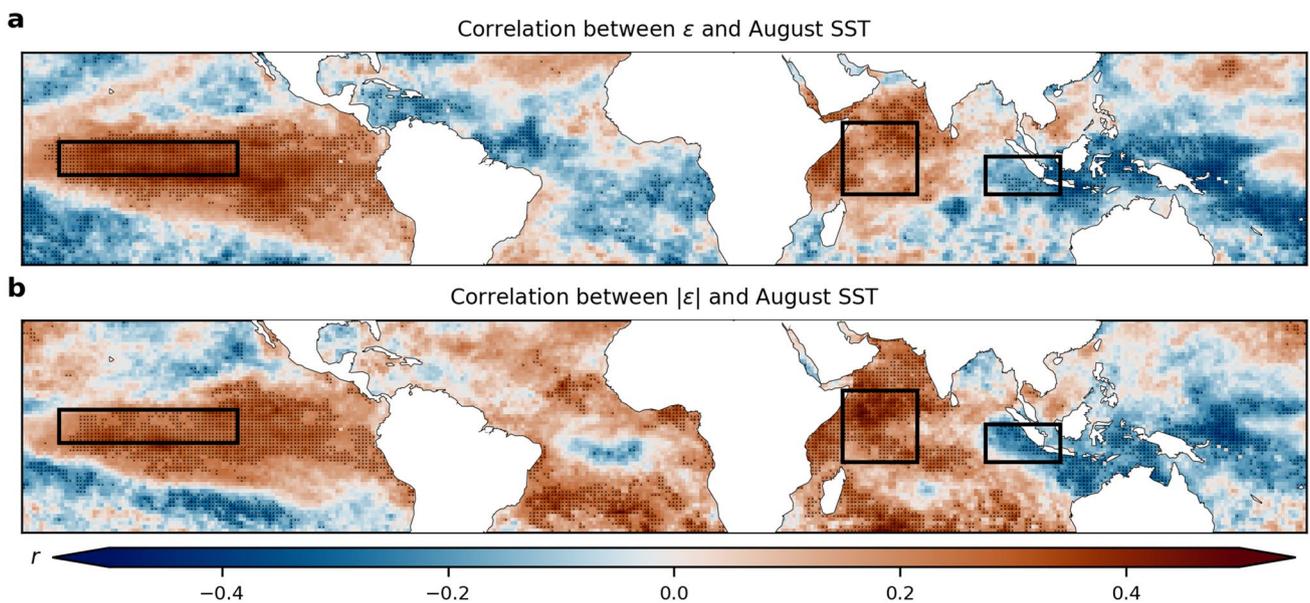
The unit is standard deviations. Note that rounding errors leads to some apparently inconsistent  $\epsilon$  values

consistent sign in August N34 and DMI values: both indices are positive in nearly all cases and strongly positive on average. The average rainfall prediction  $\widehat{PC}_1$  in these too-wet years is strongly positive (1.5), as expected from the SST indices, but there is no consistent sign in the observed rainfall indices wet forecasts, showing a near-zero average  $PC_1$  value of -0.1.

We now return to all 71 years to more comprehensively evaluate the link between the August SST state and the rainfall prediction error. In Fig. 7a, the correlation between August SST anomalies and  $\epsilon$  is shown. This shows a clear linear relationship between  $\epsilon$  and both ENSO and the IOD, with positive phases of the SST indices in August tending to result in too-wet rainfall predictions, and negative phases

tending to result in too-dry predictions. We quantify this with a correlation between  $\epsilon$  and the indices, resulting in 0.36 and 0.34 for N34 and DMI, respectively. This is an important result, but the last row of Table 2 shows that, for the most extreme cases, there is an asymmetry in that the too-wet predictions are linked to positive initial ENSO and IOD states, but the too-dry cases are not linked to negative initial ENSO and IOD conditions.

This asymmetry points to a nonlinear aspect to the relationship between the initial oceanic state and  $\epsilon$ , and we examine this potential link in Fig. 7b by showing the correlation between  $|\epsilon|$  and August SST anomalies for all the 71 years. Clearly, both negative and positive rainfall prediction errors are correlated with both ENSO and the IOD in



**Fig. 7** Maps of the interannual correlation between SST anomalies in August and  $\epsilon$  (a) and its magnitude,  $|\epsilon|$  (b). Significant correlations are indicated with dots, and the boxes indicate the outlines of the N34 and IOD regions

August. In particular, we note that the SSTs in both IOD regions show a stronger correlation with  $|\epsilon|$  than with  $\epsilon$  (Fig. 7a). We now investigate these asymmetric relationships further by dividing the 71 years into three bins, organized by ascending initial N34 and DMI index values, in turn.

The top row of Fig. 8 shows boxplots (e.g., Krzywinski and Altman 2014) representing the distribution of the forecast error  $\epsilon$  in each bin. We first note that the median of each bin increases from left to right, reflecting the positive linear relationship between  $\epsilon$  and the August SST indices seen in Fig. 7a. For the N34 index in Fig. 8a, we see that the standard deviation of the errors is lowest in the first bin, where the N34 index is negative. In the third bin, where N34 is positive, the range of the errors is largest, and the lowermost whisker of the rightmost bin extends as far into negative territory as the lowermost whisker of the leftmost bin. Similar behaviour is seen for DMI in Fig. 8b, but the standard deviation of the third bin is more than twice as large as the standard deviation of the first bin.

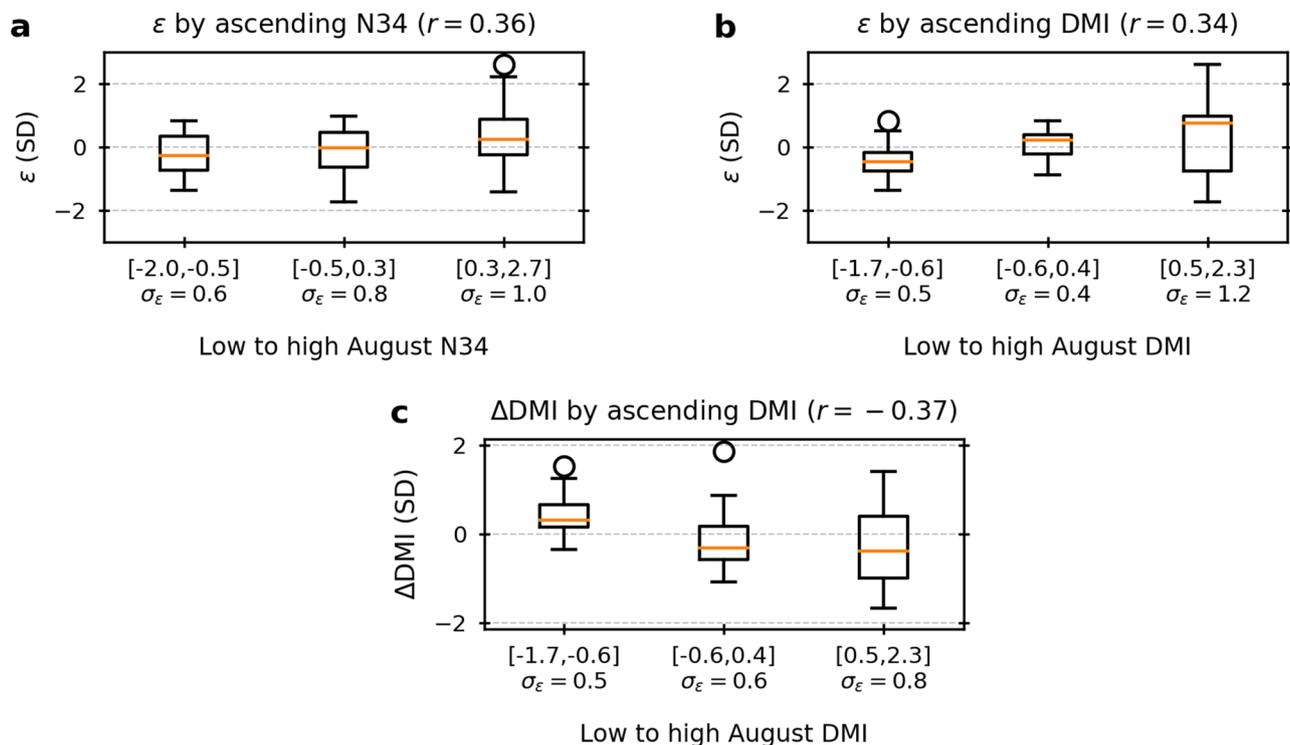
In sum, the boxplots in the top row of Fig. 8 illustrate the nonlinear relationship between the initial SST indices and the rainfall forecast error  $\epsilon$ . When the SST indices are positive in August, the error of the linear prediction model is less

predictable and spans a larger interval than when the SST indices are neutral or negative (see also Fig. 7b), making  $\epsilon$  difficult to predict a priori.

Figure 8c shows that  $\Delta\text{DMI}$  is also both linearly and non-linearly associated with August DMI values. The correlation between August DMI and  $\Delta\text{DMI}$  is significant and negative ( $-0.37$ ), which implies that, on average, DMI regresses towards the mean between August and OND. The nonlinearity of the relationship is clear: the standard deviation of  $\Delta\text{DMI}$  in the bin with the highest August DMI values is substantially higher than the standard deviation in the bin with the lowest initial values.

### 3.5 Mediating role of DMI changes

In the preceding sections, we have shown that both the rainfall forecast error  $\epsilon$  and  $\Delta\text{DMI}$  are correlated with August DMI values. We now use the mediation analysis framework in Sect. 2.5 to show that  $\Delta\text{DMI}$  fully mediates the lagged effect of August DMI on  $\epsilon$ . We note, for completeness, that  $\Delta\text{DMI}$  is significantly correlated with  $\Delta\text{N34}$ , with a correlation coefficient of 0.28, but as there is no significant effect of  $\Delta\text{N34}$  on  $\epsilon$ ,  $\Delta\text{N34}$  is not a potential mediator. We also note



**Fig. 8** Top row: Box plots of the distribution of  $\epsilon$ , sorted into bins according to August N34 (a) and DMI (b) values (increasing from left to right). Each rectangle extends from the lower (Q1) to the upper quartile (Q3) of  $\epsilon$ , and the horizontal lines show the median. The upper ‘whiskers’ extend to the highest data points lower than  $Q3 + 1.5$  IQR, where the interquartile range  $\text{IQR} = Q3 - Q1$ , and the lower

whiskers extend to the lowest data points greater than  $Q1 - 1.5$  IQR. Outliers are shown with circles. The bracket under each bar indicates the range of the SST index values, and under that the standard deviation of  $\epsilon$  in each bin is shown. **c**  $\Delta\text{DMI}$  by ascending August DMI values, with conventions as in the two preceding panels

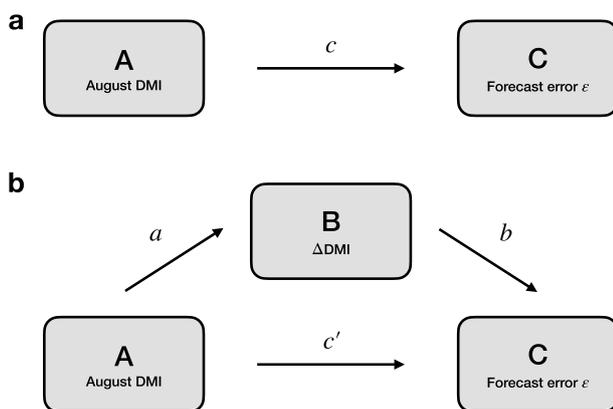
that we checked whether  $\Delta\text{DMI}$  is a mediator of the lagged effect of August N34 on  $\epsilon$ , but it is not.

In Fig. 9, we have updated Fig. 1 with the names of the causal variable  $A$  (DMI in August), the potential mediator  $B$  ( $\Delta\text{DMI}$ ), and the outcome variable  $C$  ( $\epsilon$ ). We now go through the four steps required for mediation. First, we already know from the previous section (see Fig. 8b) that the positive correlation coefficient  $c$  in Fig. 9a is significant. Second, the previous section also showed that negative correlation coefficient  $a$  in Fig. 9b is significant. Using Eq. 4, we can calculate the coefficients  $b$  and  $c'$ , which are  $-0.30$  and  $0.21$ , respectively. Bootstrapping shows that  $b$  is significant, but  $c'$  is non-significant (although its value is close to the 2.5th percentile of the bootstrapped set of artificial  $c'$  coefficients). This means that all the four steps in the mediation check are satisfied, and we have shown that  $\Delta\text{DMI}$  mediates the lagged effect of August DMI on  $\epsilon$ . Note that even if  $c'$  had been significant, its value is considerably smaller than the coefficient  $c$  ( $c'/c = 0.61$ ). Cases for which the first three steps are satisfied, but where  $c'$  is significant but  $|c'/c| < 1$ , are often referred to as ‘partial mediation’.

## 4 Summary and discussion

We have studied the lagged relationship between tropical SSTs in August and the East African short rains from October to December. By working with detrended ERA5 reanalysis data from 1950 to 2020, we obtained the following main results, which are discussed further here:

1. A linear prediction model based on the ENSO and IOD states in August accounts for about 40% of the interannual variance of an East African short rains index, which



**Fig. 9** **a**, A diagram illustrating the direct effect  $c$  of the causal variable  $A$  (August DMI) on the outcome variable  $C$  ( $\epsilon$ ). **b**, A diagram illustrating the direct effect  $a$  of  $A$  on the mediator  $B$  ( $\Delta\text{DMI}$ ), as well as the regression coefficients of Eq. 4,  $b$  and  $c'$

again accounts for more than half of the spatial and temporal OND rainfall variance in the region. Although the N34 index in August is a significant predictor of OND rainfall, this effect is fully mediated by the IOD index DMI in OND. The high prediction skill of the linear model is consistent with previous research and here we confirm the result for a long 70-year record. Since linear regression models are still widely used in East Africa, it is important to know why and when such models fail. For this reason, the remaining analysis uses this linear model, and particularly its errors, as a tool to study the dynamical linkages across time and space between August SSTs and OND rainfall, as well as to assess whether it is possible to predict the forecast errors a priori, based on the initial SSTs.

2. Tropical SSTs in August are linked to the East African short rains in OND through a Walker circulation across the Indian Ocean. We investigated the correlations between the linear rainfall prediction errors and errors of similarly defined linear prediction models for vertical velocity, low-level zonal wind, and tropical SST. The rationale for this approach is that areas where the error correlations are positive pinpoint areas where the other variables ‘communicate’ the lagged effects of the SSTs on the rainfall. When too much rainfall is predicted, a too-strong Walker circulation is simultaneously predicted, and when the linear model predicts too little rain, the predicted Walker circulation is too weak.
3. OND rainfall forecast errors are also linked to changes in Indian Ocean SSTs between August and OND. On average, negative DMI changes—cooling in the western and warming in the eastern parts of the basin—are associated with positive rainfall prediction errors. In other words, the model predicts too much rain because DMI and the Walker circulation are weaker in OND than they are predicted to be based on the state in August. Conversely, too little rain is predicted in the opposite case when the Walker circulation intensifies between August and OND. However, there are several exceptions to this ‘rule’, and the DMI changes only account for just over 10% of the prediction errors.
4. We identified both linear and nonlinear linkages across time between rainfall forecast errors in OND with the oceanic state in August. The linear part of this relationship is that the sign of the ENSO and IOD conditions in August is correlated with the sign of the subsequent prediction error (the nonlinear part is discussed below). This can potentially be exploited by correcting rainfall predictions upfront. For example, when DMI and N34 are positive in August and the linear model predicts higher-than-average rainfall, the prediction might be adjusted downwards a priori. Referring to the list of too-wet forecasts in Table 2, such an approach would have

worked in 1987, when the linear model predicts abundant rainfall, but in reality the conditions were drier-than-normal. It would also have worked in the high-DMI years 1966, 1976, and 2012, and it would have worked in 1972 and 2015, when both N34 and DMI were strongly positive. However, the approach would only have been successful for four (1972, 1984, 1989 and 2003) of the seven years with the largest negative rainfall prediction errors, when the predicted rainfall might have been adjusted upwards if either N34 or DMI (or both) had been negative. Notably, the approach would not have worked in 2019, when the region experienced massive flooding (Nicholson et al. 2021; Wainwright et al. 2021). The 2019 case is a telling example of the nonlinear relationship between initial SST conditions and subsequent rainfall errors, which we discuss next.

5. The relationship between OND rainfall forecast errors and the August SST conditions also have nonlinear aspects. By separating the 71 years into three bins according to August N34 and DMI, we showed that the bin with the highest initial index values has substantially larger error variability than the bin with the lowest initial values. A practical implication is that an a priori adjustment of the forecast is less reliable when the ENSO and/or the IOD are in the positive phase in August. The 2019 season is a prime example. In August, DMI was 2.1 standard deviations above normal, and as the model prediction for 2019 is already too dry, a further downward adjustment of the prediction based on the positive IOD state would have exacerbated the forecast error. Examples such as this are an argument for using coupled dynamical models, which one can hope are or will be able to capture non-linear evolutions of SSTs and Walker circulations in and over the Indian Ocean.
6. Despite the nonlinear nature of the lagged relationships between August SSTs and short rains prediction errors, we identified a mechanism which mediates the lagged linear effect of the IOD state in August on these errors, namely changes in DMI between August and OND. The causal pathway is as follows. The DMI in August is negatively correlated with the change in DMI between August and OND. This means that both positive and negative August values tend to regress toward the mean (which is a neutral index). The DMI in August is positively correlated with the rainfall in OND, but it is also positively correlated with the forecast *error*. This means that when DMI is positive in August, the linear model predicts higher-than-normal rainfall, but it often predicts too much rainfall, as it does for all the seven years in the right part of Table 2. The mediation by the change in DMI from August to OND shows that one reason for this is that DMI tends to become more neutral between August and OND, which again means that the Walker

circulation in OND is less vigorous than expected by the prediction model in August. This often means that it rains less than predicted. It is important to emphasize that although the DMI change fully mediates the direct effect of August DMI on the OND forecast error, there may be other mechanisms that also mediate this effect. The simple mediation framework used here does not imply exclusivity. Finding more mediators of the effect investigated here would further enhance our understanding of the lagged relationships between tropical SSTs and East African rainfall.

In summary, this study has shown that statistical short rains forecasts based on the IOD and ENSO states in August are quite skilful. However, the linear prediction model has a systematic bias which is conditional, in both a linear and a nonlinear sense, on the initial state. In some cases, an a priori adjustment of the forecast can lead to improved forecasts, but the nonlinear characteristic of the forecast bias makes such adjustments perilous, especially when the initial IOD state is positive. These results highlight knowledge gaps around the relationships between the East African short rains and the large-scale drivers which underlie the high predictability of this season.

**Author contributions** Both authors conceived and wrote the paper, and EWK performed the data analysis and produced the figures and tables.

**Funding** Open Access funding provided by NORCE Norwegian Research Centre AS. The research was supported by funding from the European Union's Horizon 2020 program through the CONFER (grant 869730) and DOWN2EARTH (grant 869550) projects, and from the Research Council of Norway through the Seasonal Forecasting Engine project (grant 270733) and the Climate Futures research centre (grant 309562).

**Availability of data and material** The ERA5 data are available from the Copernicus Climate Data Store (Hersbach et al. 2018a, b, 2020a, b).

**Code availability** N/A.

## Declarations

**Conflict of interest** The authors declare that they have no conflict of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will

need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Bahaga TK, Mengistu Tsidu G, Kucharski F, Diro GT (2015) Potential predictability of the sea-surface temperature forced equatorial East African short rains interannual variability in the 20th century. *Quart J Roy Meteorol Soc* 141(686):16–26. <https://doi.org/10.1002/qj.2338>
- Bahaga TK, Fink AH, Knippertz P (2019) Revisiting interannual to decadal teleconnections influencing seasonal rainfall in the Greater Horn of Africa during the 20th century. *Int J Climatol* 39(5):2765–2785. <https://doi.org/10.1002/joc.5986>
- Baron RM, Kenny DA (1986) The moderator–mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J Pers Soc Psychol* 51(6):1173–1182. <https://doi.org/10.1037/0022-3514.51.6.1173>
- Bazo J, Singh R, Destrooper M, Coughlan de Perez E (2019) Chapter 18—pilot experiences in using seamless forecasts for early action: the “Ready-Set-Go!” approach in the red cross. In: Robertson AW, Vitart F (eds) *Sub-seasonal to seasonal prediction*. Elsevier, Amsterdam
- Behera SK, Luo J-J, Masson S, Delecluse P, Gualdi S, Navarra A, Yamagata T (2005) Paramout impact of the Indian Ocean dipole on the east african short rains: a CGCM study. *J Clim* 18(21):4514–4530. <https://doi.org/10.1175/jcli3541.1>
- Beltrando G, Camberlin P (1993) Interannual variability of rainfall in the eastern horn of Africa and indicators of atmospheric circulation. *Int J Climatol* 13(5):533–546. <https://doi.org/10.1002/joc.3370130505>
- Black E (2005) The relationship between Indian Ocean sea-surface temperature and East African rainfall. *Phil Trans R Soc A* 363(1826):43–47. <https://doi.org/10.1098/rsta.2004.1474>
- Black E, Slingo J, Sperber KR (2003) An observational study of the relationship between excessively strong short rains in Coastal East Africa and Indian Ocean SST. *Mon Weather Rev* 131(1):74–94. [https://doi.org/10.1175/1520-0493\(2003\)131%3c0074:Aosotr%3e2.0.Co;2](https://doi.org/10.1175/1520-0493(2003)131%3c0074:Aosotr%3e2.0.Co;2)
- Bruno Soares M, Daly M, Dessai S (2018) Assessing the value of seasonal climate forecasts for decision-making. *Wires Clim Chang* 9(4):e523. <https://doi.org/10.1002/wcc.523>
- Chang’a LB, Kijazi AL, Mafuru KB, Kondowe AL, Osima SE, Mtongori HI, Ng’ongolo HK, Juma OH, Michael E (2020) Assessment of the evolution and socio-economic impacts of extreme rainfall events in October 2019 over the east Africa. *Atmos and Clim Sci* 10(3):319–338. <https://doi.org/10.4236/acs.2020.103018>
- Colman AW, Graham RJ, Davey MK (2020) Direct and indirect seasonal rainfall forecasts for East Africa using global dynamical models. *Int J Climatol* 40(2):1132–1148. <https://doi.org/10.1002/joc.6260>
- Conway D, Allison E, Felstead R, Goulden M (2005) Rainfall variability in East Africa: implications for natural resources management and livelihoods. *Phil Trans R Soc A* 363(1826):49–54. <https://doi.org/10.1098/rsta.2004.1475>
- Dawson A (2016) eofs: a library for eof analysis of meteorological, oceanographic, and climate data. *J Open Res Softw*. <https://doi.org/10.5334/jors.122>
- Dunning CM, Black ECL, Allan RP (2016) The onset and cessation of seasonal rainfall over Africa. *J Geophys Res Atmos* 121(19):11405–411424. <https://doi.org/10.1002/2016JD025428>
- Dutra E, Magnusson L, Wetterhall F, Cloke HL, Balsamo G, Bousssetta S, Pappenberger F (2013) The 2010–2011 drought in the Horn of Africa in ECMWF reanalysis and seasonal forecast products. *Int J Climatol* 33(7):1720–1729. <https://doi.org/10.1002/joc.3545>
- Funk C, Peterson P, Landsfeld M, Pedreros D, Verdin J, Shukla S, Husak G, Rowland J, Harrison L, Hoell A, Michaelsen J (2015) The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Sci Data* 2(1):150066. <https://doi.org/10.1038/sdata.2015.66>
- Gleixner S, Demissie T, Diro GT (2020) Did ERA5 improve temperature and precipitation reanalysis over East Africa? *Atmosphere* 11(9):996. <https://doi.org/10.3390/atmos11090996>
- Goddard L, Graham NE (1999) Importance of the Indian Ocean for simulating rainfall anomalies over eastern and southern Africa. *J Geophys Res Atmos* 104(D16):19099–19116. <https://doi.org/10.1029/1999JD900326>
- Haile GG, Tang Q, Sun S, Huang Z, Zhang X, Liu X (2019) Droughts in East Africa: causes, impacts and resilience. *Earth-Sci Rev* 193:146–161. <https://doi.org/10.1016/j.earscirev.2019.04.015>
- Hansen JW, Mason SJ, Sun L, Tall A (2011) Review of seasonal climate forecasting for agriculture in Sub-Saharan Africa. *Exp Agric* 47(2):205–240. <https://doi.org/10.1017/S0014479710000876>
- Hastenrath S, Nicklis A, Greischar L (1993) Atmospheric-hydrospheric mechanisms of climate anomalies in the western equatorial Indian Ocean. *J Geophys Res: Oceans* 98(6111):20219–20235. <https://doi.org/10.1029/93JC02330>
- Hastenrath S, Polzin D, Camberlin P (2004) Exploring the predictability of the ‘Short Rains’ at the coast of East Africa. *Int J Climatol* 24(11):1333–1343. <https://doi.org/10.1002/joc.1070>
- Hersbach H, Bell B, Berrisford P, Biavati G, Horányi A, Muñoz Sabater J, Nicolas J, Peubey C, Radu R, Rozum I, Schepers D, Simmons A, Soci C, Dee D, Thépaut J-N (2018a) ERA5 monthly averaged data on pressure levels from 1979 to present
- Hersbach H, Bell B, Berrisford P, Biavati G, Horányi A, Muñoz Sabater J, Nicolas J, Peubey C, Radu R, Rozum I, Schepers D, Simmons A, Soci C, Dee D, Thépaut J-N (2018b) ERA5 monthly averaged data on single levels from 1979 to present
- Hersbach H, Bell B, Berrisford P, Biavati G, Horányi A, Muñoz Sabater J, Nicolas J, Peubey C, Radu R, Rozum I, Schepers D, Simmons A, Soci C, Dee D, Thépaut J-N (2020a) ERA5 monthly averaged data on single levels from 1950 to 1978 (preliminary version)
- Hersbach H, Bell B, Berrisford P, Biavati G, Horányi A, Muñoz Sabater J, Nicolas J, Peubey C, Radu R, Rozum I, Schepers D, Simmons A, Soci C, Dee D, Thépaut J-N (2020b) ERA5 monthly averaged data on pressure levels from 1950 to 1978 (preliminary version)
- Hersbach H, Bell B, Berrisford P, Hirahara S, Horányi A, Muñoz-Sabater J, Nicolas J, Peubey C, Radu R, Schepers D, Simmons A, Soci C, Abdalla S, Abellan X, Balsamo G, Bechtold P, Biavati G, Bidlot J, Bonavita M, De Chiara G, Dahlgren P, Dee D, Diamantakis M, Dragani R, Flemming J, Forbes R, Fuentes M, Geer A, Haimberger L, Healy S, Hogan RJ, Hólm E, Janisková M, Keeley S, Laloyaux P, Lopez P, Lupu C, Radnoti G, de Rosnay P, Rozum I, Vamborg F, Villaume S, Thépaut J-N (2020c) The ERA5 global reanalysis. *Quart J Roy Meteorol Soc* 146(730):1999–2049. <https://doi.org/10.1002/qj.3803>
- Kilavi M, MacLeod D, Ambani M, Robbins J, Dankers R, Graham R, Tittle H, Salih AAM, Todd MC (2018) Extreme rainfall and flooding over central kenya including Nairobi City during the long-rains season 2018: causes, predictability, and potential for early warning and actions. *Atmosphere* 9(12):472. <https://doi.org/10.3390/atmos9120472>
- Kolstad EW, MacLeod D, Demissie TD (2021) Drivers of subseasonal rainfall forecast errors over East Africa. *Geophys Res Lett*. <https://doi.org/10.1029/2021GL093292>
- Krzywinski M, Altman N (2014) Visualizing samples with box plots. *Nat Methods* 11(2):119–120. <https://doi.org/10.1038/nmeth.2813>

- Latif M, Dommenges D, Dima M, Grötzner A (1999) The role of Indian Ocean sea surface temperature in forcing East African rainfall anomalies during December–January 1997/98. *J Clim* 12(12):3497–3504. [https://doi.org/10.1175/1520-0442\(1999\)012%3c3497:TROIOS%3e2.0.CO;2](https://doi.org/10.1175/1520-0442(1999)012%3c3497:TROIOS%3e2.0.CO;2)
- Lemos MC, Kirchhoff CJ, Ramprasad V (2012) Narrowing the climate information usability gap. *Nature Clim Change* 2(11):789–794. <https://doi.org/10.1038/nclimate1614>
- Little PD, Mahmoud H, Coppock DL (2001) When deserts flood: risk management and climatic processes among East African pastoralists. *Clim Res* 19(2):149–159. <https://doi.org/10.3354/cr019149>
- MacLeod D (2019) Seasonal forecast skill over the Greater Horn of Africa: a verification atlas of System 4 and SEAS5—part 1/2. ECMWF report. <https://www.ecmwf.int/node/18906>
- MacLeod D, Caminade C (2019) The moderate impact of the 2015 El Niño over East Africa and its representation in seasonal reforecasts. *J Clim* 32(22):7989–8001. <https://doi.org/10.1175/JCLI-D-19-0201.1>
- MacLeod D, Graham R, O'Reilly C, Otieno G, Todd M (2021) Causal pathways linking different flavours of ENSO with the Greater Horn of Africa short rains. *Atmos Sci Lett* 22(2):e1015. <https://doi.org/10.1002/joc.1015>
- McKenna S, Santoso A, Gupta AS, Taschetto AS, Cai W (2020) Indian Ocean Dipole in CMIP5 and CMIP6: characteristics, biases, and links to ENSO. *Sci Rep* 10(1):11500. <https://doi.org/10.1038/s41598-020-68268-9>
- Mutai CC, Ward MN, Colman AW (1998) Towards the prediction of the East Africa short rains based on sea-surface temperature–atmosphere coupling. *Int J Climatol* 18(9):975–997. [https://doi.org/10.1002/\(SICI\)1097-0088\(199807\)18:9%3c975::AID-JOC259%3e3.0.CO;2-U](https://doi.org/10.1002/(SICI)1097-0088(199807)18:9%3c975::AID-JOC259%3e3.0.CO;2-U)
- Mwangi E, Wetterhall F, Dutra E, Di Giuseppe F, Pappenberger F (2014) Forecasting droughts in East Africa. *Hydrol Earth Syst Sci* 18(2):611–620. <https://doi.org/10.5194/hess-18-611-2014>
- Nicholson SE (2014) The predictability of rainfall over the Greater Horn of Africa. Part I: prediction of seasonal rainfall. *J Hydrometeorol* 15(3):1011–1027. <https://doi.org/10.1175/JHM-D-13-062.1>
- Nicholson SE (2015) Long-term variability of the East African ‘short rains’ and its links to large-scale factors. *Int J Climatol* 35(13):3979–3990. <https://doi.org/10.1002/joc.4259>
- Nicholson SE (2017) Climate and climatic variability of rainfall over eastern Africa. *Rev Geophys* 55:590–635. <https://doi.org/10.1002/2016RG000544>
- Nicholson SE, Fink AH, Funk C, Klotter DA, Sathesh AR (2021) Meteorological causes of the catastrophic rains of October/November 2019 in equatorial Africa. *Global Planet Change*. <https://doi.org/10.1016/j.gloplacha.2021.103687>
- Ntale HK, Gan TY, Mwale D (2003) Prediction of East African seasonal rainfall using simplex canonical correlation analysis. *J Climate* 16:2105–2112. [https://doi.org/10.1175/1520-0442\(2003\)016%3c2105:POEASR%3e2.0.CO;2](https://doi.org/10.1175/1520-0442(2003)016%3c2105:POEASR%3e2.0.CO;2)
- Nyamekye AB, Nyadzzi E, Dewulf A, Werners S, Van Slobbe E, Biesbroek RG, Termmer CJAM, Ludwig F (2021) Forecast probability, lead time and farmer decision-making in rice farming systems in Northern Ghana. *Clim Risk Manag* 31:100258. <https://doi.org/10.1016/j.crm.2020.100258>
- Nyenzi BS (1988) Mechanisms of East African rainfall variability. Ph.D. thesis, The Florida State University.
- Philippon N, Camberlin P, Fauchereau N (2002) Empirical predictability study of October–December East African rainfall. *Quart J Roy Meteorol Soc* 128(585):2239–2256. <https://doi.org/10.1256/qj.01.190>
- Ropelewski CF, Halpert MS (1987) Global and regional scale precipitation patterns associated with the El Niño/Southern oscillation. *Mon Weather Rev* 115(8):1606–1626. [https://doi.org/10.1175/1520-0493\(1987\)115%3c1606:GARSPP%3e2.0.CO;2](https://doi.org/10.1175/1520-0493(1987)115%3c1606:GARSPP%3e2.0.CO;2)
- Saji NH, Goswami BN, Vinayachandran PN, Yamagata T (1999) A dipole mode in the tropical Indian Ocean. *Nature* 401(6751):360–363. <https://doi.org/10.1038/43854>
- Schreck CJI, Semazzi FHM (2004) Variability of the recent climate of eastern Africa. *Int J Climatol* 24(6):681–701. <https://doi.org/10.1002/joc.1019>
- Stuecker MF, Timmermann A, Jin F-F, Chikamoto Y, Zhang W, Wittenberg AT, Widiasih E, Zhao S (2017) Revisiting ENSO/Indian Ocean Dipole phase relationships. *Geophys Res Lett* 44(5):2481–2492. <https://doi.org/10.1002/2016GL072308>
- Tall A, Coulibaly JY, Diop M (2018) Do climate services make a difference? A review of evaluation methodologies and practices to assess the value of climate information services for farmers: Implications for Africa. *Clim Serv* 11:1–12. <https://doi.org/10.1016/j.cliser.2018.06.001>
- Ummenhofer CC, Sen Gupta A, England MH, Reason CJ (2009) Contributions of Indian Ocean sea surface temperatures to enhanced East African rainfall. *J Clim* 22(4):993–1013. <https://doi.org/10.1175/2008JCLI2493.1>
- VanderWeele TJ, Vansteelandt S (2014) Mediation analysis with multiple mediators. *Epidemiol Methods* 2(1):95–115. <https://doi.org/10.1515/em-2012-0010>
- Vigaud N, Lyon B, Giannini A (2017) Sub-seasonal teleconnections between convection over the Indian Ocean, the East African long rains and tropical Pacific surface temperatures. *Int J Climatol* 37(3):1167–1180. <https://doi.org/10.1002/joc.4765>
- Wainwright CM, Finney DL, Kilavi M, Black E, Marsham JH (2021) Extreme rainfall in East Africa, October 2019–January 2020 and context under future climate change. *Weather* 76(1):26–31. <https://doi.org/10.1002/wea.3824>
- Walker DP, Birch CE, Marsham JH, Scaife AA, Graham RJ, Segele ZTJCD (2019) Skill of dynamical and GHACOF consensus seasonal forecasts of East African rainfall. *Clim Dyn*. <https://doi.org/10.1007/s00382-019-04835-9>
- Wenhaji Ndomeni C, Cattani E, Merino A, Levizzani V (2018) An observational study of the variability of East African rainfall with respect to sea surface temperature and soil moisture. *Quart J Roy Meteorol Soc* 144(S1):384–404. <https://doi.org/10.1002/qj.3255>
- Zhao S, Cook KH (2021) Influence of walker circulations on East African rainfall. *Clim Dyn* 56:2127–2147. <https://doi.org/10.1007/s00382-020-05579-7>
- Zheng X-T, Lu J, Hui C (2021) Response of seasonal phase locking of Indian Ocean Dipole to global warming. *Clim Dyn*. <https://doi.org/10.1007/s00382-021-05834-5>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.