

Drivers and Subseasonal Predictability of Heavy Rainfall in Equatorial East Africa and Relationship with Flood Risk

DAVID A MACLEOD,^{a,b} RUTGER DANKERS,^{c,d} RICHARD GRAHAM,^d KISWENDSIDA GUIGMA,^e LUKE JENKINS,^e MARTIN C. TODD,^e AUGUSTINE KIPTUM,^{e,f} MARY KILAVI,^f ANDREW NJOGU,^f AND EMMAH MWANGI^g

^a *Atmospheric, Oceanic and Planetary Physics, University of Oxford, Oxford, United Kingdom*

^b *School of Geographical Sciences, University of Bristol, Bristol, United Kingdom*

^c *Wageningen Environmental Research, Wageningen University and Research, Wageningen, Netherlands*

^d *Met Office, Exeter, United Kingdom*

^e *Department of Geography, University of Sussex, Brighton, United Kingdom*

^f *Kenya Meteorological Department, Nairobi, Kenya*

^g *Kenya Red Cross Society, Nairobi, Kenya*

(Manuscript received 28 August 2020, in final form 18 January 2021)

ABSTRACT: Equatorial East Africa (EEA) suffers from significant flood risks. These can be mitigated with preemptive action; however, currently available early warnings are limited to a few days' lead time. Extending warnings using subseasonal climate forecasts could open a window for more extensive preparedness activity. However, before these forecasts can be used, the basis of their skill and relevance for flood risk must be established. Here we demonstrate that subseasonal forecasts are particularly skillful over EEA. Forecasts can skillfully anticipate weekly upper-quintile rainfall within a season, at lead times of 2 weeks and beyond. We demonstrate the link between the Madden–Julian oscillation (MJO) and extreme rainfall events in the region, and confirm that leading forecast models accurately represent the EEA teleconnection to the MJO. The relevance of weekly rainfall totals for fluvial flood risk in the region is investigated using a long record of streamflow from the Nzoia River in western Kenya. Both heavy rainfall and high antecedent rainfall conditions are identified as key drivers of flood risk, with upper-quintile weekly rainfall shown to skillfully discriminate flood events. We additionally evaluate GloFAS global flood forecasts for the Nzoia basin. Though these are able to anticipate some flooding events with several weeks lead time, analysis suggests action based on these would result in a false alarm more than 50% of the time. Overall, these results build on the scientific evidence base that supports the use of subseasonal forecasts in EEA, and activities to advance their use are discussed.

KEYWORDS: Africa; Madden-Julian oscillation; Forecast verification/skill; Flood events

1. Introduction

Heavy rainfall can have very serious societal impacts. These arise indirectly through its contribution to high streamflow and fluvial flooding and, more directly, pluvial flash flooding. Heavy rainfall can also trigger landslides. Estimates suggest that flooding affects more people than any other geophysical hazard globally with associated economic losses second only to storms (UNIDRR 2015).

Recent events in 2018–20 have highlighted flood risk across East Africa in particular with extensive flood loss and damage from exceptionally wet seasons. During the “long rains” season of 2018 flooding across Kenya caused the displacement of 300 000 people (OCHA 2018), shortly followed by the “short rains” October–December season flooding of 2019, associated

with a strong Indian Ocean dipole event (Doi et al. 2020; Wainwright et al. 2020). Anomalous wet conditions persisted across East Africa through to May 2020 resulting in hundreds of deaths from flood and landslides and hundreds of thousands of people displaced across the region (Floodlist 2020).

A key pathway to mitigate risks is taking preventative action based on early warnings from weather and climate forecasts. Early warning systems (EWS) have long been recognized as a key component of flood risk management and international initiatives have supported their development for many years. The Sendai Framework for Disaster Risk Reduction established a global target: “Substantially increase the availability of and access to multi-hazard early warning systems (MHEWS) and disaster risk information and assessments to the people by 2030” (UNIDRR 2015). More broadly, movement toward anticipatory action is growing, through initiatives such as Forecast-based Finance pioneered by the International Federation of Red Cross and Red Crescent Societies, among many others (Wilkinson et al. 2018).

Early warnings of extreme weather and flooding are provided with short lead times, as they are based on traditional weather forecasts, generally targeting no further than a week ahead. However, longer lead times in the subseasonal forecast time scale (with lead times of 10–30 days) would allow for more effective preparation and implementation of a wider range of preventative actions (White et al. 2017). Further, subseasonal forecasts can bridge the gap between weather forecasts and

Denotes content that is immediately available upon publication as open access.

Supplemental information related to this paper is available at the Journals Online website: <https://doi.org/10.1175/JHM-D-20-0211.s1>.

Corresponding author: David MacLeod, david.macleod@bristol.ac.uk

DOI: 10.1175/JHM-D-20-0211.1

© 2021 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBSReuseLicenses).

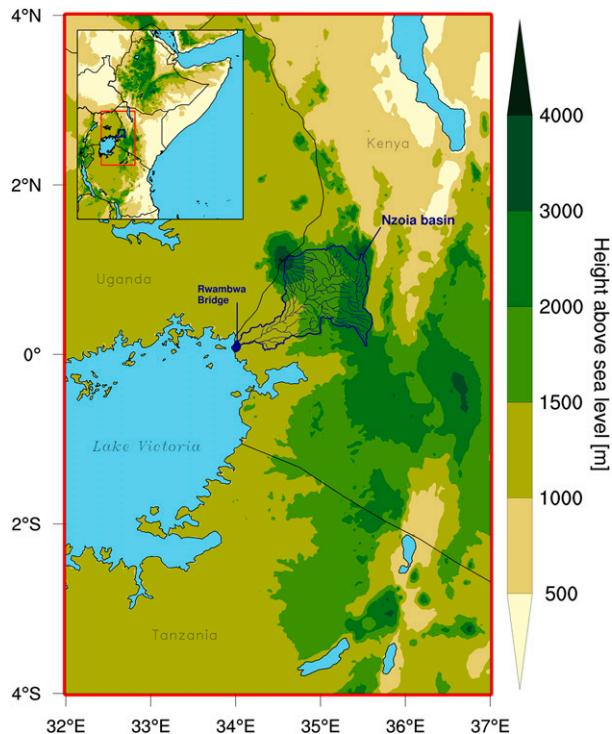


FIG. 1. Topography of the study area, including the location of the Nzoia basin and associated river network (dark blue) within GHA (inset). The dark blue marker shows the location of Rwambwa Bridge, the gauge data station used in the study. The magnified region (outlined in red) corresponds to the domain used to evaluate drivers of large-scale modes of variability (see section 2c).

seasonal outlooks, which can be linked through, for example a “Ready. Set. Go!” approach to seamless lead-time risk management (Bazo et al. 2019).

As recent events illustrate, the wider equatorial region of East Africa, hereafter EEA, centered approximately on Lake Victoria (Fig. 1) in particular suffers impacts of heavy rainfall. There is a clear need for improved localized EWS with longer lead time for anticipatory action. Existing early warnings of extreme rainfall and/or flood risk in EEA are generally limited to a few days lead time. For example, in Kenya, the Kenya Meteorological Department (KMD) issues heavy rainfall advisories with a few days’ lead time (MacLeod et al. 2020) and a 3-day streamflow forecast is produced for the Nzoia River basin of western Kenya. In Uganda, the Uganda Red Cross in collaboration with other stakeholders are piloting the use of the Global Flood Awareness System (GloFAS) to initiate early actions up to a week in advance. These examples are representative of the lead time of currently available basin-scale flood forecasts in the region.

Despite a demand for reliable longer-range forecasts over East Africa (e.g., Kilavi et al. 2018) there remains little uptake of subseasonal forecasting by national meteorological centers in the East Africa region. Contributing factors include a lack of awareness of the potential, a lack of capacity and limited access to

raw forecast model outputs from leading subseasonal forecast models. To counter this, the ongoing WMO/WCRP Subseasonal to Seasonal (S2S) real-time pilot seeks to extend the dissemination and use of subseasonal forecasts (Vitart 2017).

However, in order to build robust early action protocols based on subseasonal forecasts, the forecast information needs to be reliable. Generally, for much of the global tropics the basis of subseasonal predictability is relatively well established and is associated with large-scale modes of variability, notably the Madden–Julian oscillation (MJO; Waliser 2011; Zaitchik 2017). For the EEA region there is evidence of the MJO as a primary driver of the westerly low-level wind anomalies linked to wet days/spells (e.g., Pohl et al. 2007; Berhane and Zaitchik 2014; Finney et al. 2020; Kilavi et al. 2018; Nicholson 2017). In general our understanding of weather and climate is much weaker for EEA than for the wider Greater Horn of Africa (GHA) region, although recent dedicated research projects have been undertaken to address this, e.g., the U.K.-funded HyCristal (<https://futureclimateafrica.org/project/hycristal>) and Highway (<https://www.metoffice.gov.uk/about-us/what/working-with-other-organisations/international/projects/wiser/highway>). Our current understanding of the nature, drivers and predictability of subseasonal climate variability over EEA, required to support risk management is insufficient at present.

In this context this paper contributes to the scientific evidence base for predictability of extreme rainfall and flood risk in the EEA region and addresses the following research questions:

- What are the large-scale drivers of extreme rainfall, relevant for both fluvial and pluvial flood?
- How skillful are rainfall predictions from subseasonal forecast systems and how well do they represent teleconnections with these main sources of predictability?
- What is the link between extreme rainfall and fluvial flooding and what are the implications for flood predictability?

We also present and discuss the trialing of real-time subseasonal forecast information for early warning in Kenya, supported by the ForPac project (Toward Forecast-based Preparedness Action, www.forpac.org), which has been exploring the use of forecasts for drought and flood preparedness across Kenya.

Our paper is organized as follows. Data and methodology are described in section 2. The large-scale drivers of extreme rainfall over EEA are evaluated in section 3. Then in section 4 we present a skill evaluation of state-of-the-art subseasonal prediction systems up to several weeks ahead over EEA and the wider surrounding GHA in all rainy seasons. In section 5 we focus on a specific flood risk case study of the Nzoia River of western Kenya (Fig. 1) in order to evaluate the potential for extending the lead time of fluvial flood EWS. The Nzoia River is one of the most flood-prone basins in Kenya, with the only established operational flood EWS in the region backed by a long and high-quality time series of discharge observations. For this reason it is an ideal candidate for evaluation in the current study. The paper then concludes with a discussion of the findings along with recommendations for further development of EWS in the region.

2. Data and methodologies

a. Observational and reanalysis data

The Climate Hazards Group Infrared Precipitation with Station Data (CHIRPS), are used to assess drivers of extreme rainfall, verify subseasonal forecasts and to explore the drivers of fluvial flooding. CHIRPS merges microwave and infrared satellite data with in situ gauge observations to produce daily rainfall estimates at 0.05° from 1981 to the present (Funk et al. 2015). To evaluate the driving tropical modes of variability and the associated atmospheric circulation we use, respectively, satellite-derived outgoing longwave radiation (OLR) data from the National Oceanic and Atmospheric Administration (NOAA) on a 2.5° grid (Liebmann and Smith 1996) and ERA5 reanalysis data, produced by the European Centre for Medium-Range Weather Forecasts (ECMWF). ERA5 provides a spatially and temporally complete estimate of hourly atmospheric fields at 30-km resolution (Copernicus Climate Change Service 2017). Real-time multivariate MJO (RMM) indices (RMM1 and RMM2) are used to evaluate MJO forecasts and teleconnections, which are derived from OLR and upper-level winds from NCEP reanalysis and provided by the Bureau of Meteorology (<http://www.bom.gov.au/climate/mjo/graphics/rmm.74toRealttime.txt>). Daily discharge data for the Nzoia River at Rwambwa Bridge (0.0944°N, 33.9919°E) in the basin floodplain in Siaya County from 1950 to 2018 are provided by KMD.

b. Forecast model data

We evaluate skill and representation of teleconnections in leading subseasonal forecast models, using reforecasts produced by both ECMWF and the U.K. Met Office.

Twice a week ECMWF issues a 51-member ensemble forecast out to 46 days (hereafter, ENS-extended), along with a parallel 11-member reforecast of the previous 20 years. Here we use the reforecast generated for 1997–2016, based on the 2017 model versions to evaluate forecast skill. The resolution of this particular reforecast is around 18km for the first 15 days and changes to around 36 km for days 16–46. A full description of ENS-extended can be found at <https://www.ecmwf.int/en/forecasts/documentation-and-support/extended-range-forecasts>.

The U.K. Met Office uses the GloSea5 system to produce forecasts of the following 60 days. GloSea5 forecasts are produced on the 1st, 9th, 17th, and 25th of each month, using seven perturbed members per start date and an atmospheric resolution of approximately 50 km. A reforecast is run for the period 1993–2016; here we use a period coincident with the ENS-extended hindcast (1997–2016) following the same approach as the ENS-extended data and treating each start date as an independent forecast (the alternative being a pooling of multiple start dates to increase the ensemble size, as is done in operational GloSea5 forecast production). A full description of GloSea5 can be found in MacLachlan et al. (2015).

From both the ENS-extended and GloSea5 reforecasts we extract forecasts of daily precipitation accumulation. We also use forecasted daily RMM data to evaluate MJO skill and teleconnections, which are calculated following Vitart (2017) and provided through the S2S project database.

c. Methodologies

1) EVALUATING LARGE-SCALE DRIVERS OF EXTREME RAINFALL OVER EEA

To evaluate the large-scale drivers of rainfall, CHIRPS rainfall data are first high-pass filtered using a Lanczos filter (Duchon 1979) to retain high-frequency variability (less than 90 days). Extreme rainfall days are then defined at both gridcell and regional levels. At a given grid cell, a day is considered extreme if the daily rainfall exceeds the 95th percentile of the distribution across all days at that grid cell. Then for the EEA regional level, an extreme rainfall day corresponds to a day where at least 10% of grid cells within the EEA domain exceed their local 95th percentile. This domain (32°–37°E, 4°S–4°N, Fig. 1) is centered on the Nzoia catchment and is chosen to be in accordance with previous studies (e.g., Fink et al. 2020; Finney et al. 2020).

To detect tropical modes including the MJO and equatorial waves (Rossby, Kelvin, mixed Rossby–gravity, and African easterly waves), a wavenumber–frequency decomposition is performed on the OLR data as in Wheeler and Kiladis (1999). The characteristics of each mode (Kiladis et al. 2009) are used to retain the corresponding harmonics. The convective activity of each tropical mode is further assessed by filtering OLR data at a reference longitude of 35°E (averaged between 4°S and 4°N) and normalizing using the standard deviation. The resulting time series and its associated time derivatives are further combined to determine the amplitude and the local phases of the mode as in Riley et al. (2011). The mode is considered active only when its amplitude is greater than one.

The impact of the activity of tropical modes on the occurrence of extreme rainfall is then assessed by first calculating the fraction of extreme rainfall falling on active mode phases and then comparing the observed frequency of rainfall during active versus inactive phases, similar to Xavier et al. (2014). The modulation of extreme rainfall occurrence by a given mode on phase x is thus calculated using

$$M = \frac{(P_x - P_N)}{P_N}, \quad (1)$$

where P_x is extreme rainfall frequency associated with an active phase x of a given mode, i.e., the number of days when extreme rain occurs during phase x of the mode over the total number of days of the season, and P_N is the frequency of extreme rainfall when the mode is inactive. This formula is applied to the gridcell data of extreme rainfall events using the RMM MJO indices. The analysis is done separately for relevant rainfall seasons detailed later and results are shown in section 3.

2) EVALUATING EXTENDED RANGE FORECASTS OF RAINFALL AND THE DRIVERS OF EXTREME RAINFALL OVER EEA AND WIDER GHA

We use a range of verification metrics to evaluate subseasonal forecasts; descriptions of all may be found in Jolliffe and Stephenson (2012). For evaluation of MJO forecasts we use reliability and sharpness diagrams, which indicate biases in

model-generated probabilities and the boldness of issued probabilities. MJO teleconnections are evaluated by calculating composites of weekly rainfall based on an active MJO in each phase (activity defined as the magnitude of $RMM1^2 + RMM2^2$ exceeding unity). Model composites are derived by calculating average rainfall totals forecast across all start dates and ensemble members when the model itself (not the observations) indicates a specific MJO state. To calculate the statistical significance of each MJO teleconnection map the number of active weeks, N , in each phase is used as a basis for bootstrap resampling. A total of 1000 composites of N weeks chosen at random are used to generate 2.5th and 97.5th percentiles of the composite anomaly expected by chance. When the true composite lies outside this interval it is then classified as statistically significant at the 95% level. Lower significance thresholds are found for model data as member as an independent realization of model behavior and so the sample size is higher.

Forecasts of anomalous weekly rainfall totals are evaluated according to the season in which the target week falls. We use four seasons defined by Yang et al. (2015), corresponding to the long rains [March–May (MAM)]; the short rains [October–December (OND)]; the boreal summer season [June–September (JJAS)], which is relatively wet over parts of EEA, including the Nzoia River basin, but dry over most of East Africa (also mostly corresponding to the Kiremt rainy season over Ethiopia); and the boreal winter dry season [January–February (JF)]. Forecasts for 1–6 weeks ahead are evaluated and anomalies are calculated by removing the mean of the rainfall total for all reforecast weeks falling in that season. Anomalies are calculated independently for each lead time in order to account for potential bias in lead time.

Maps of ensemble mean correlation provide a broad indication of skill, with 95th percentile significance levels calculated with a t test. Verification of probabilities focuses on rainfall total above the 80th percentile (upper quintile) as indicative of a wet week. Such an event occurs once in 5 weeks on average; for a 3-month rainy season this corresponds to the 2 or 3 wettest weeks during the season. Analysis of flooding over Kenya demonstrates that wet seasons can be dominated by pronounced intraseasonal events falling in a few distinct weeks (Kilavi et al. 2018). Discrimination ability of wet weeks is calculated using the Relative Operating Characteristic Area Under Curve (ROC AUC), with 95th percentile statistical significance values calculated using the Mann–Whitney U transform described in Mason and Graham (2002).

Maps of reliability and sharpness metrics are also presented. Reliability is quantified for each grid point by categorizing the slope of the weighted regression line through the reliability diagram. Notional categories are used: reliable (above 0.9), slightly overconfident (0.7–0.9), overconfident (0.3–0.7), highly overconfident (0.1–0.3), and unreliable (below 0.1). These boundaries are intended as broad indications of different classes of forecast reliability. Such a categorization of reliability diagrams was first carried out by Weisheimer and Palmer (2014) for seasonal forecasts, using both the slope and its uncertainties to define categories. However, the relatively large sample size of subseasonal forecasts here gives much smaller uncertainty in the slope, hence we define categories

based on the slope alone. A metric to quantify sharpness is defined here: the percentage of forecasts where the forecasted probabilities of an event are more than twice the climatological frequency of that event. For the quintile events studied here this corresponds to the percentage of forecasts showing a probability of an upper-quintile event to be greater than 40%.

3) EVALUATING THE LINK BETWEEN HEAVY RAINFALL AND FLUVIAL FLOOD RISK

We focus our analysis of basin scale fluvial flood on the Nzoia River basin, a medium-sized river basin (area 12 700 km²) in western Kenya which rises in the highlands of Mount Elgon flowing to the southwest into Lake Victoria (Fig. 1). This basin is selected as it is one of the most flood-prone basins in Kenya, as well as benefiting from a long and high-quality dataset of daily river flow.

To explore the role of precursor conditions in conditioning flood risk we use the antecedent precipitation index (API), which is defined on a daily time step t ,

$$API_t = kAPI_{t-1} + p_t, \quad (2)$$

where API is calculated by adding the day's precipitation p to the previous day's API, modified by a parameter k . API is monitored at KMD as a flood indicator, using a value of $k = 0.97$ (chosen to maximize the correlation between API and streamflow). API is essentially the sum of previous days' precipitation, weighted toward recent days, where the weight of the n th previous day is given by k^n . This corresponds to a weighting of 0.4, 0.16, and 0.06 for rainfall falling 1, 2, and 3 months previously.

To assess the predictability of Nzoia River basin flooding we assess reforecasts of daily river discharge from GloFAS version 2.0 (Alfieri et al. 2013), at the grid cell nearest to the Rwambwa Bridge monitoring station. GloFAS (<http://www.globalfloods.eu/>) is the global flood service of the European Commission's Copernicus Emergency Management Service (CEMS), an operational system for monitoring and forecasting floods across the world [see Alfieri et al. (2013) for more detail]. GloFAS is based on rainfall runoff generated by the ENS-extended hindcast described previously, and we assess skill using the reforecast which has two start dates per week and 11 ensemble members for 1997–2017. While the ENS-extended input to GloFAS is uncalibrated, GloFAS is calibrated using streamflow data where they are available. For the Nzoia basin GloFAS has recently been calibrated using the flow record mentioned above; however, the reforecasts used here were created with a model version which was not calibrated for the basin.

3. Large-scale drivers of extreme rainfall

Low-level westerly anomalies bringing moisture from the Congo basin into East Africa have been highlighted by previous studies as a driver of rainfall (e.g., Nicholson 2017; Kilavi et al. 2018; Finney et al. 2020). Analysis of the EEA domain shown in Fig. 1 is consistent and shows extreme precipitation days over this region to be accompanied by increased moisture flux and westerly wind anomalies (not shown). The westerlies

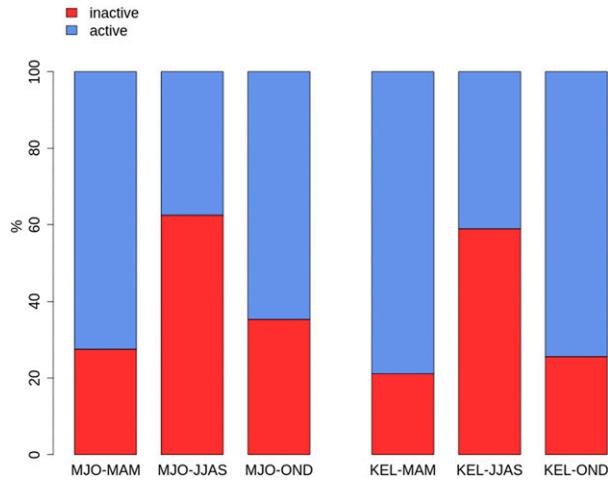


FIG. 2. Distribution of extreme rainfall days over the East Africa box between the local (red shades) inactive and (blue shades) active phases of the MJO and Kelvin waves during the MAM, JJAS, and OND seasons. The number of extreme rainfall days for MAM is 156, 56 for JAS, and 82 for OND.

themselves, at the intraseasonal scale, may appear as a footprint of the passage of large eastward tropical modes over the region (e.g., Fink et al. 2020). As a result, it is relevant to analyze the direct statistical link between them and extreme precipitation over the region as this may serve to establish potential mechanisms of predictability.

Analysis of tropical modes shows that activity of eastward modes, namely, the MJO and the equatorial Kelvin waves, have the strongest relationship with extreme precipitation over

EEA. Figure 2 shows the coincidence of extreme rainfall days with either MJO or Kelvin wave activity, broken down by season. During OND and MAM, 75% of extreme rainfall days (above the 95th percentile) are coincident with Kelvin wave activity, while 70% (60%) of MAM (OND) extreme days occur alongside MJO activity. It should be noted that these modes are often simultaneously active (e.g., over 50% of extreme rainfall in MAM occurs when both modes are active). In JJAS, the activity of the MJO and Kelvin waves weakens significantly over Africa (Guigma et al. 2020) and they become less associated with extreme rainfall events whose number also decreases.

The MJO is a source of subseasonal predictability (Vitart 2017) and suggests potential predictability of heavy rainfall. In particular Fig. 3 indicates that the MJO in its phases 2–4 (when convection is located over East Africa) significantly increases the likelihood of extreme rainfall in the region. This is consistent with previous work linking MJO activity in these phases to rainfall variability over the Greater Horn of Africa (Berhane and Zaitchik 2014; Zaitchik 2017). In MAM, MJO phases 3 and 4 increase the probability of extreme rainfall by up to 150%, mostly over central Tanzania and western Kenya. In OND, there is an even more important role of the MJO than in MAM with excess risk of 200% (i.e., a threefold increase) on phases 2 (Uganda and western Kenya), 3 (Uganda, western Kenya and the whole of Tanzania), and 4 (central Tanzania and western Kenya). In JJAS wide parts of the region do not record extreme rainfall and the MJO is less active compared to other seasons.

4. Skill of subseasonal forecasts

The relationship between MJO activity and extreme rainfall suggests subseasonal forecasting potential. We explore this

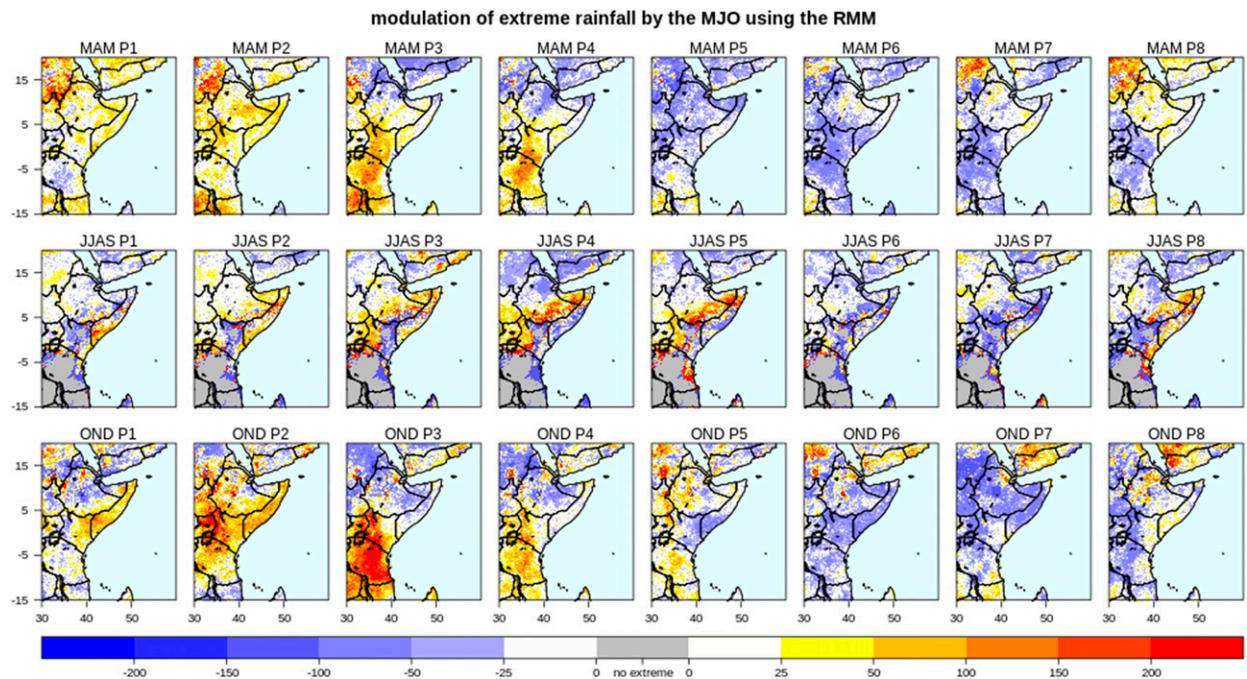


FIG. 3. Modulation (%) of extreme rainfall probability of occurrence by the MJO based on the RMM index.

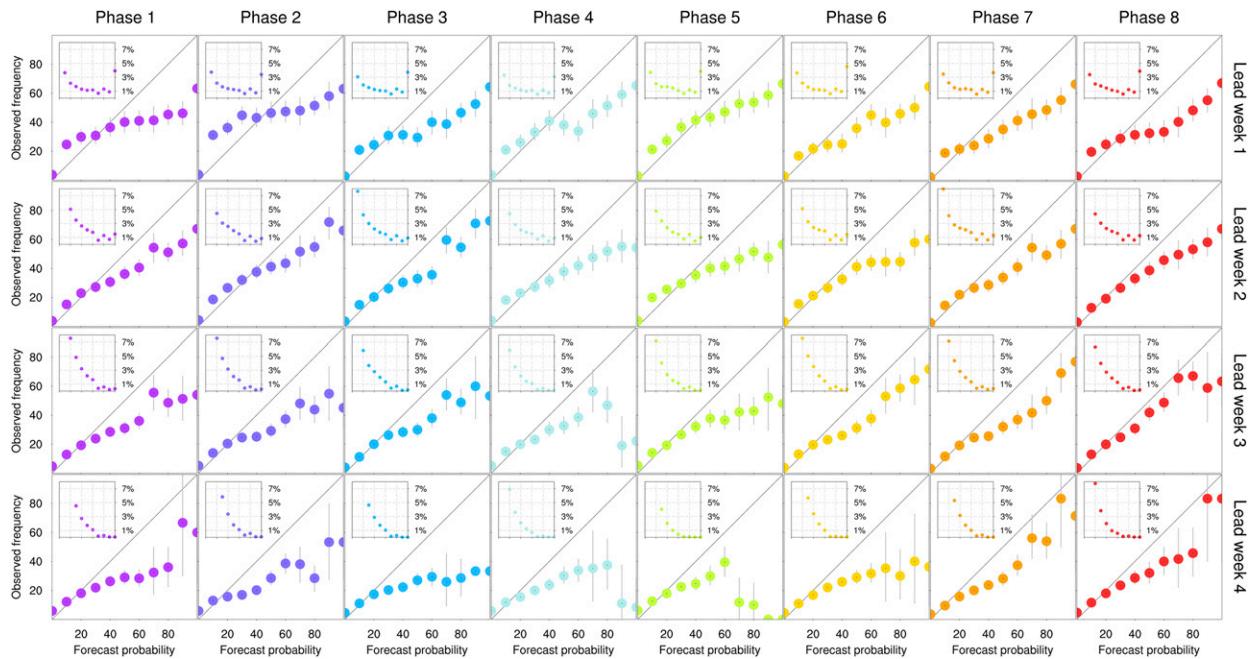


FIG. 4. Reliability diagrams for ENS-extended for daily MJO activity in any zone 1–8 (from left to right) falling in lead week 1–4 (from top to bottom). The inset plot indicates log axis frequency distribution of forecast probabilities, where the x axis follows the x axis of the larger plots.

through analysis of the two leading subseasonal forecast systems: the ENS-extended and GloSea5.

Reliability diagrams for MJO activity forecasts from ENS-extended are shown in Fig. 4 (corresponding GloSea5 results are shown in Fig. S1 in the online supplemental material). Probabilities show a positive relationship with observed frequencies out to 4 weeks ahead, although are slightly overconfident overall, consistent with the underdispersion reported by Vitart (2017). Issued forecast probabilities are relatively bold; for example, during the second week at least 5% of all forecasts are issued with probabilities greater than 50%. For context, the climatological frequency of MJO in each phase in the reforecast period ranges from 5% to 9% (not shown) and so a forecast of 50% represents an enhancement of 5–10 times the normal probability of occurrence. At 4-week lead the sharpness drops significantly across all phases. In general, probabilities appear to be quite well calibrated in the lower probability range, becoming significantly overconfident at higher probabilities. The slope of the reliability diagram is similar across most phases; however, for phases two and three in particular we note a flattening of the curve at longer lead times, indicating an increase in overconfidence at long leads.

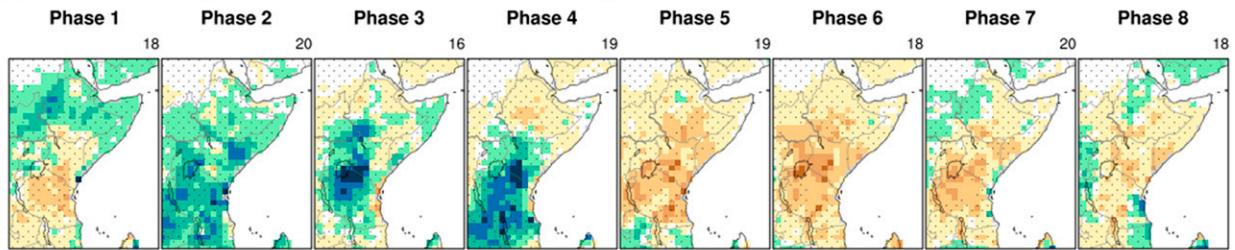
We next assess the model representation of the MJO teleconnection to rainfall over the GHA region. Weekly rainfall composited on MJO activity is shown in Fig. 5 in observations and reforecasts. Results are shown for forecasts at 1-, 2-, and 6-week lead time for ENS-extended forecast weeks falling in the long rains, while results for the short rains and other seasons and for GloSea5 are shown in Figs. S2–S8 (we omit weeks 3–5 as they are highly similar to week 2). Observations show a wet (dry) signal with MJO activity in phases 2–4 (5–8), consistent

with composites of extreme daily rainfall in Fig. 3 and previous work (Berhane and Zaitchik 2014; Zaitchik 2017). This is present across all seasons and is largely captured by both forecasting systems. On average the strength of the signal in models is lower than the observed signal, although after a slight drop in the strength of the signal between lead week 1 and 2, the strength does not degrade further.

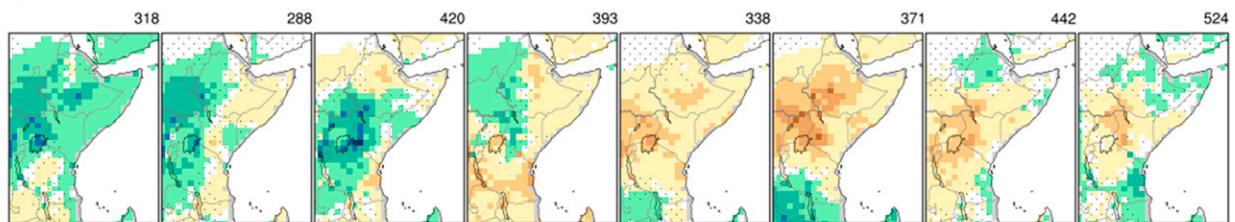
We note a few instances where simulated teleconnections diverge significantly from observations. Most notably during the MAM long rains both models incorrectly simulate a wet signal over Kenya during phase 1 when a dry signal is expected. ENS-extended also fails to capture the strong phase 4 wet signal over Kenya in this season, although this is simulated correctly by GloSea5 (Fig. S5). In addition during the short rains the phase 4 wet signal in the south of the region is too weak. However, in general, the teleconnection is largely well captured by both models. Along with skillful forecasts of MJO activity this establishes the foundation for subseasonal skill of rainfall in the region.

Figure 6 shows verification of ENS-extended forecasts for heavy rainfall weeks falling in the long rains (results for other seasons and GloSea5 are shown in Figs. S9–S15). We note correlations over 0.6 over a large region for the first week and up to 0.6 for the second week. The highest skill is seen for a region extending from Lake Victoria to southern Ethiopia, but notably over western Kenya and including the Nzoia River basin, with statistically significant correlations at this region even at a 6-week lead time. Discrimination of wet weeks over the Nzoia basin is high for week 2 forecasts (ROC AUC up to 0.8) and is even statistically significant up to forecast week 6. Model probabilities are slightly overconfident for week 1 and 2

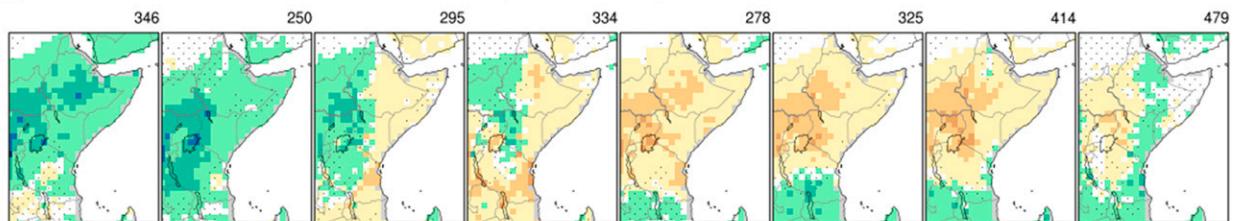
(a) Observed weekly rainfall teleconnection to MJO activity



(b) Teleconnection in ENS-extended reforecast (one week lead)



(c) Teleconnection in ENS-extended reforecast (two week lead)



(d) Teleconnection in ENS-extended reforecast (six week lead)

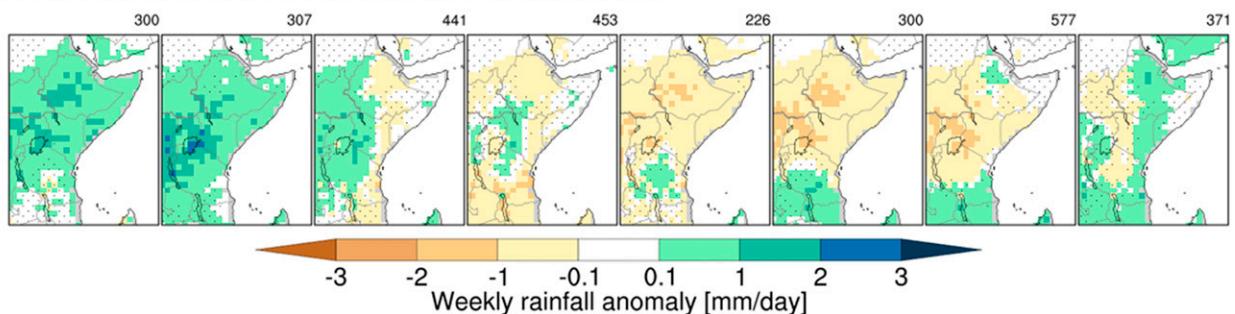


FIG. 5. Anomalous weekly rainfall (mm day^{-1}) during MAM associated with MJO activity in zone 1–8 (from left to right) in (a) observations (using CHIRPS rainfall) and (b)–(d) the ENS-extended (the first, second, and sixth week of the forecast). Numbers above each panel show sample size, and stippling indicates where anomalies are not significant at the 95% level.

forecasts and become increasingly overconfident with lead time. However, the slope of the reliability diagram line remains positive over the Nzoia basin at all lead times, indicating that probabilities calculated directly from the model ensemble carry information about increased risk of heavy rainfall, although they tend to overestimate this risk. Probabilities are reasonably sharp, with at least 1 in 5 lead one forecasts indicating a doubling of risk, dropping to less than 1 in 10 forecasts in week 6. GloSea5 shows a very similar pattern of skill, including highest values seen over western Kenya (Fig. S13).

Verification results are largely similar for forecasts targeting weeks during the short rains (Figs. S11 and S15). Forecasts for January–February (Figs. S9 and S12) and June–September

(Figs. S10 and S14) are generally less skillful, with scores largely dropping below statistical significance beyond week 3. However, both ENS-extended and GloSea5 still show relatively good discrimination (ROC AUC of 0.7) wet weeks at lead week 2 over western Kenya for both seasons, and for JJAS models show skill at lead week 2 in regions experiencing significant rainfall during JJAS, including Uganda, western Kenya, South Sudan, and Ethiopia, along with coastal regions.

The evaluation presented here supports the use of sub-seasonal forecasts in the region. To this end a trial of real-time forecast information has been implemented through the ForPac project since 2018, where the U.K. Met Office has provided probabilistic weekly rainfall tercile forecasts from

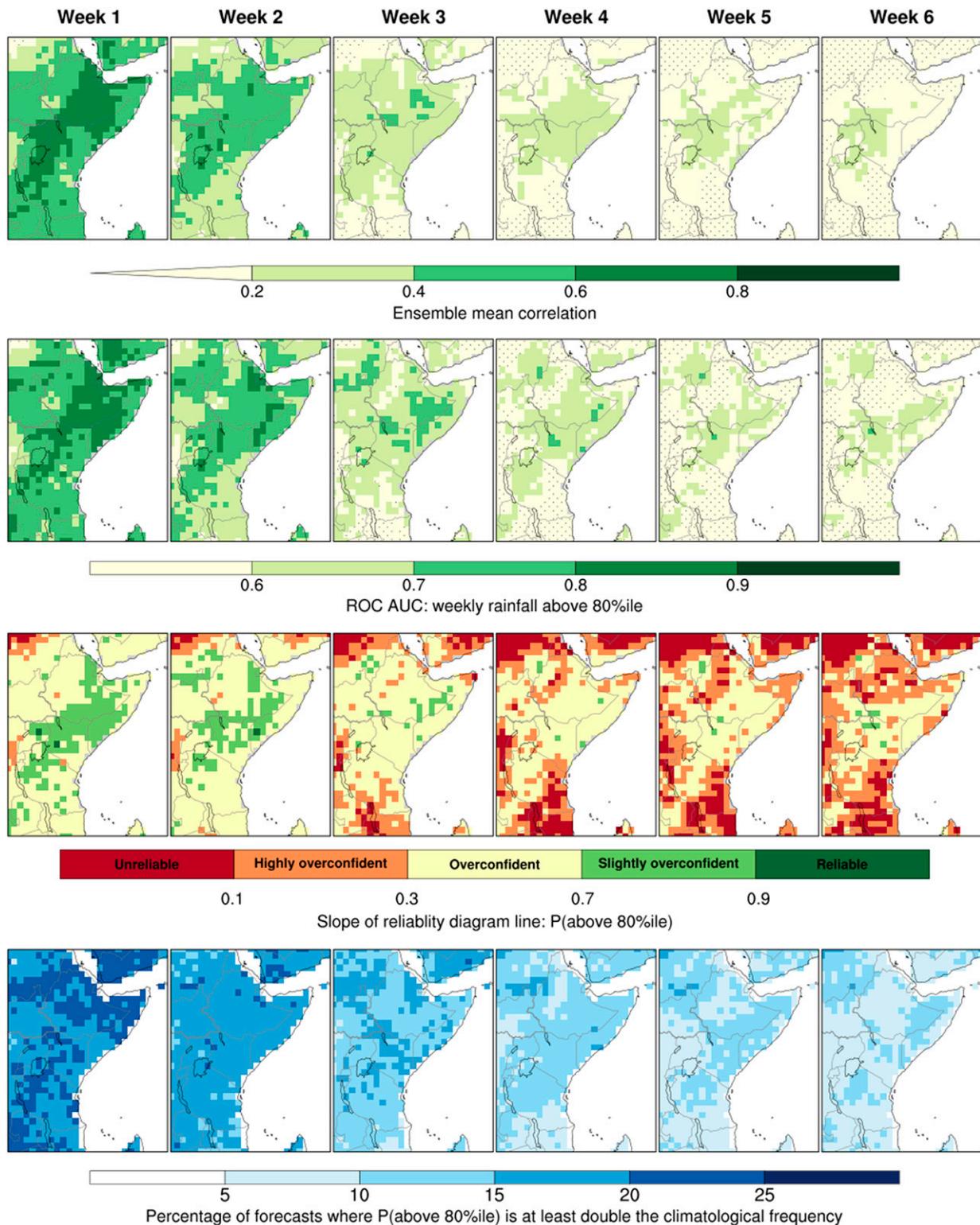


FIG. 6. Performance of ENS-extended forecasts of rainfall totals 1–6 weeks ahead (from left to right) during MAM. (first row) Ensemble mean correlation, with values below the 95th significance level indicated with stippling. Rows 2–4 show verification of forecast probabilities for a wet week (defined here as rainfall total falling above the 80th percentile). (second row) ROC AUC (values below significance are stippled). (third row) Slope of the reliability diagram, where a value of one indicates perfectly reliable probabilities and a value of zero indicates that probabilities have no relationship with outcomes. (fourth row) Sharpness, defined here as the number of times the system indicates a probability of an 80th percentile event greater than double the climatological frequency (i.e., greater than 40%). Results here are calculated for weeks falling in the long rains; results for other seasons are provided in the supplemental material.

GloSea5 up to 4 weeks ahead to KMD and the Intergovernmental Authority on Development (IGAD) Climate Predictions and Applications Centre (ICPAC). This trial was carried out ahead of the S2S project real-time pilot phase (which commenced in November 2019) and in East Africa served as a valuable first exploration of potential. Here, we note that consistent with observed outcomes, the forecasts warned of enhanced chance of late (early) onset of the long (short) rains 2019 (not shown). In addition, flooding during late April 2020 which led to over 200 deaths and affecting 800 000 people (Floodlist 2020) was well anticipated at least 3 weeks ahead, as witnessed by the tercile category probability forecasts from GloSea5 provided by the U.K. Met Office for the week of 20–26 April 2020 from 30 March (4 weeks lead) and updated weekly which consistently indicated strongly raised probability of rainfall in the upper tercile (Fig. S17). The active MJO during this time, a likely driver of the prolonged heavy rains, was also well anticipated by GloSea5 (not shown).

Overall subseasonal forecasts show clear potential value for the warning of the direct impacts of heavy rainfall. However, the relationship between rainfall and fluvial flooding is complicated by the role of hydrology; this is considered in the following section.

5. Links between heavy rainfall and streamflow in the Nzoia River basin and potential for long lead early warnings of flood risk

Accumulated precipitation and soil moisture are often key drivers of flood risk (Berghuijs et al. 2019), such that rainfall forecasts alone are unlikely to be sufficient for triggering fluvial flood preparedness. Here we evaluate the relationship between extreme rainfall and fluvial flood in western Kenya through the use of a record of daily discharge observations for the Nzoia basin. API is used to characterize long-term antecedent rainfall conditions (see section 2c for details).

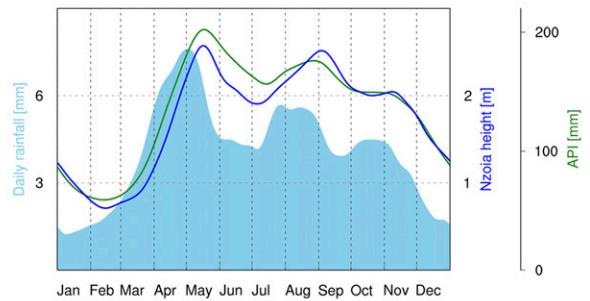
Nzoia is one of the most flood-prone basins in the region and in 2008 the KMD Flood Diagnostic and Forecasting Centre established a flood EWS. Predetermined river flow thresholds are used to release alerts or warnings are released (Hoedjes et al. 2014): a forecasted height of 4 m triggers a flood alert, while 5 m approximates bankfull discharge. This information is disseminated through community radio, emails, and updated website information (Shilenje and Ogwang 2015).

We proceed first with a characterization of the Nzoia River seasonal cycle and interannual variability, along with an assessment of precursors to recent flooding events (section 5a). We then determine the importance of long-term rainfall accumulation in conditioning the relationship between heavy rainfall events and flood risk, and outline the potential for subseasonal flood forecasting over the Nzoia with a brief evaluation of the GloFAS system (section 5b).

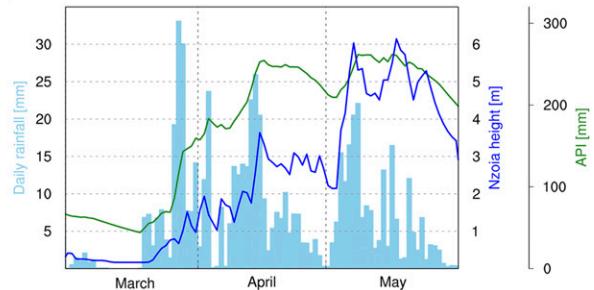
a. Drivers of flooding in the Nzoia basin

The seasonal cycle of the river level, catchment rainfall, and API is shown in Fig. 7a. There are not two distinct seasons like large parts of Kenya (Nicholson 2017). Annual rainfall minimum occurs at the start of the year, with increases through

(a) Nzoia daily climatology



(b) 1985



(c) 2013

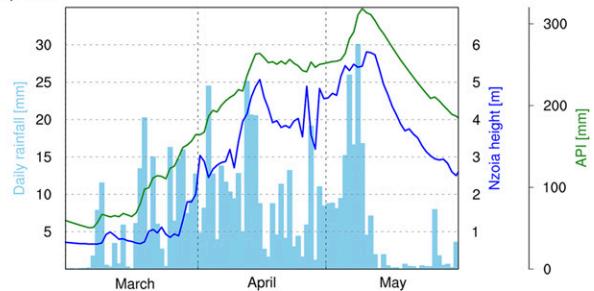


FIG. 7. (a) Seasonal cycle of rainfall, Nzoia height, and API over Nzoia basin estimated based on 1981–2018 data. (b),(c) Evolution of hydrology during the two floodiest periods since 1981: MAM 1985 and 2013.

March and April and an annual maximum in early May. Rainfall declines slightly through May before increasing again through July to a second maxima in August, before declining slightly in September, plateauing in October and continuing to decline through November and December. The climatology of API lags rainfall by around 15 days from January through to June, then remains relatively high throughout the rest of the year before declining at the end of the year. The climatology of river height is well matched by API, which it lags by a few days.

An overview of high-flow events is shown in Fig. 8, indicating all periods with river height above 4 and 5 m. We estimate the return period for a 5-m event to be around 3 years (based on the approach of fitting a generalized extreme value distribution to annual maximum discharge, analysis not shown). Nzoia heights exceeding 5 m have been seen in all months bar January and February, although most flood events occur in May, August, and September (and to a lesser extent November). The concentration of flood events in these months

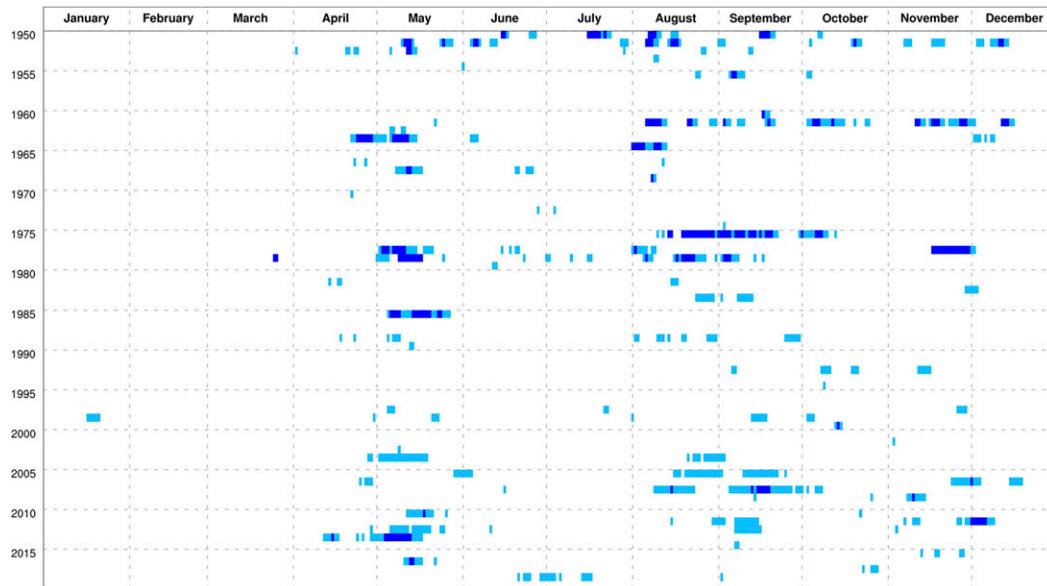


FIG. 8. Showing days on record where Nzoia level at Rwambwa Bridge was over 4 m (light blue) and 5 m (dark blue).

rather than April and May when mean river height and rainfall are highest gives a first-order indication that soil-moisture-dependent precipitation excess rather than extreme rainfall is the primary flood generating mechanism, which we explore below. In this case it may be forecasting moderate to heavy (rather than extreme) rainfall may be most important for flood forecasting.

There is pronounced interannual variability in flood events; the river remains below 4 m in some years while others experience extended flood periods. A few years contain the majority of 5-m exceedance days, notably 1950, 1951, 1961, 1975, 1977, and 1985. After 1985 no 5-m exceedance was observed until a brief event in October 1999, with more following in 2007, 2011, and 2013. Some of these extreme years are associated with major climate events affecting the larger GHA region, notably the Indian Ocean dipole event of 1961 (Hameed 2018). However, other major events such as the El Niños of 1982, 1997, or 2015 (MacLeod and Caminade 2019) do not trigger high-flow periods. This suggests a key role of subseasonal variability. Overall, no strong and systematic association of flood occurrence with multiday or monthly precipitation was found (not shown), consistent with the work of Stephens et al. (2015), who demonstrate that monthly total rainfall is not a good indication of “floodiness” for many regions worldwide.

To better understand the drivers of flooding events we look in detail at individual flooding events. Figures 7b and 7c shows the evolution of river height, daily rainfall totals and API for two of the floodiest periods in the record, occurring during the long rains of 1985 and 2013. Here API tracks river height well, demonstrating that the association does not simply arise when averaging over multiple years. The events also indicate that flood peaks are not necessarily immediately preceded by extreme rainfall. For 1985 the days in mid-May where the river rose beyond 5 m were preceded by a single day of 10 mm,

among a series with relatively moderate rainfall of around 5 mm day^{-1} . On the other hand, the first 5-m event in 2013 occurred after 3 days of 20 mm, a relatively significant total of more than 60 mm. This indicates two Nzoia flooding pathways; one triggered by moderate rainfall falling on highly saturated soil and another triggered by heavy rainfall. To skillfully predict flood events, both pathways must be represented.

Figure 9 shows the immediate precursors of all flood events during the period of daily precipitation coverage (1981 onward). In this period there were 43 days where river levels were over 5 m; in this figure we focus just on those where the level had not already exceeded 5 m in the preceding week (indicating relatively fast onset and unexpected flooding), resulting in 11 events. The precursor meteorological states were assessed in order to identify potential common drivers; however, no consistent drivers were found (analysis not shown). This indicates that other factors beyond basin-average rainfall and API (e.g., the subbasin distribution of rainfall) may control the exact timing of threshold crossing events. On the other hand, all eleven 5-m crossing events occurred following both weekly rainfall and API above the 80th percentile. This indicates that a weekly rainfall total and API falling in the upper quintile may be relevant predictors of flooding.

The effectiveness of weekly rainfall and API as flood predictors is assessed, by plotting ROC curves. These are formed by plotting the hit rate (proportion of events successfully identified) as a function of false alarm rate (proportion of nonevents incorrectly classified). Hit rate and false alarm rate are calculated in this case by assuming weekly rainfall above a series of increasing thresholds would result in a 5-m event the following day (note this analysis considers all 43 days of above 5-m exceedance a flood event).

Figure 10a shows that the curve is significantly above the diagonal, indicating that weekly rainfall thresholds alone are

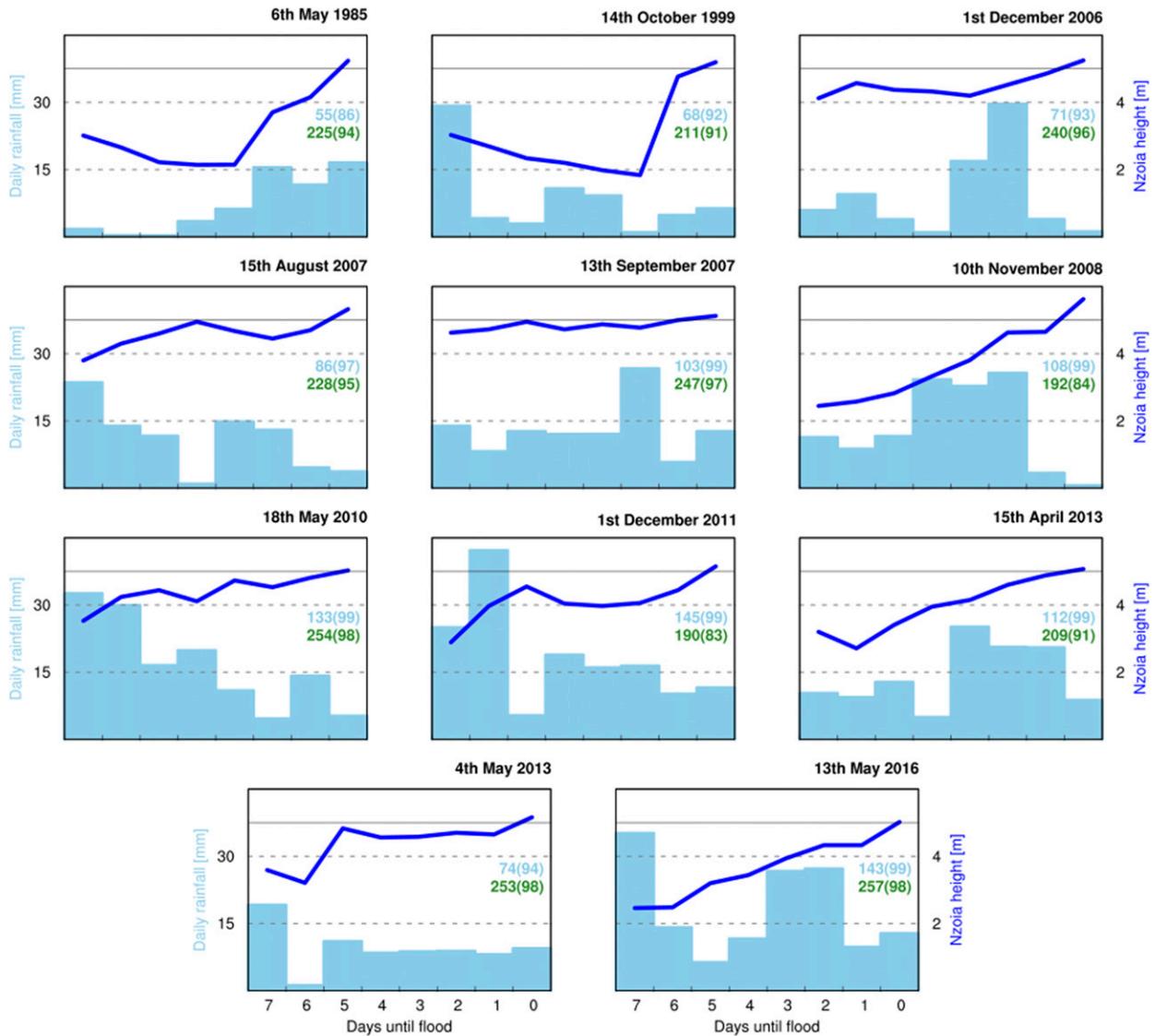


FIG. 9. Daily rainfall (blue bars, left axis) and river height in the week preceding the eleven 5-m crossing events 1981–2019. The light blue and green numbers inset on the right of each plot report the weekly rainfall total and initial API value, with the number in parenthesis indicating the corresponding percentile.

able to discriminate flood events. A threshold of 80th percentile rainfall identifies around 90% of all events, with a false alarm rate of 20%. Using higher thresholds of 90th or 95th percentile reduces the hit rate to 60% and 40%, respectively, and the false alarm rate to around 10% and 5%. Figure 10a also show the discrimination achieved by requiring both precipitation and API to be above a threshold in order to “predict” a 5-m event. Adding 80th percentile API as a predictor significantly improves the discrimination ability: the false alarm rate is reduced from 20% to around 5% without degrading the hit rate. Similar results are found when using 90th percentile API, although use of 95th percentile API degrades the hit rate, as more events are missed.

While encouraging, caution must be taken in interpreting these as necessarily useful predictors of flooding in isolation. A

hit rate of 90% with false alarm rate of 20% certainly indicates skillful discrimination ability, with around 38/43 days correctly identified. However, there are around 14 200 nonflood days in the sample, so a false alarm rate corresponds to around 3000 incorrectly identified days. Given the rarity of the event, it is instructive to show the modified ROC curve (described in Richardson et al. 2011), which plots hit rate as a function of false alarm ratio, rather than false alarm rate. The false alarm ratio indicates the number of false alarms as a percentage of the total number of forecasted events (rather than as a percentage of the total number of observed nonevents) and so is insensitive to the large number of nonevents. Results are shown in Fig. 10b and show that while upper-quintile precipitation and API anticipate over 90% of flood events, 90% of the time it is not followed by flooding. This high false alarm ratio is

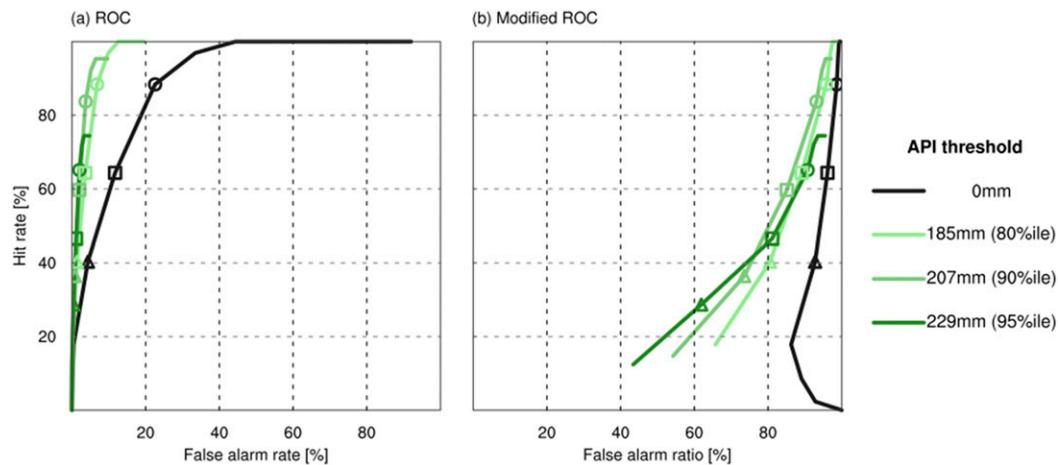


FIG. 10. Showing (a) ROC and (b) modified ROC curves obtained if weekly rainfall accumulation exceeding a range of percentile thresholds (1st–99th percentile) is used to predict if the Nzoia will be above 5 m at the end of that week. Where the curve is marked with a circle, square or triangle indicates the score obtained by using 80th, 90th, and 95th percentile weekly rainfall exceedance (corresponding to 48, 62, and 76 mm). Curves for unstratified data are shown in black, while light, medium, and dark green curves show results using only data where the week begins with API above the 80th, 90th, or 95th percentiles (corresponding to 185, 207, or 229 mm).

reduced to 60% with the use of 95th percentile precipitation and API; however, this combination only successfully identifies 33% of events.

Figure 10b shows that while a combination of high weekly rainfall and API has skill in discriminating flood days, it would result in frequent false positives relative to the total number of flood days. Inclusion of other factors such as subbasin distribution of rainfall and soil drainage characteristics may reduce the number of false positives, while temperature-driven increases in evaporative demand may dampen the possibility of flooding after some heavy rainfall events. The changing pattern of land use over time may drive nonstationarity in flood-relevant thresholds, given increases in runoff over agricultural land.

b. Evaluation of GloFAS flood forecasts for the Nzoia basin

Upper-quintile weekly rainfall over EEA is predictable (section 4) and is also associated with all flooding events (section 5a). A flood model over Nzoia using skillful sub-seasonal rainfall forecasts while representing antecedent rainfall conditions and hydrological dynamics has the potential to accurately forecast flood events. The GloFAS system is one such system and here we briefly evaluate it over Nzoia (a comprehensive evaluation is forthcoming; A. Kiptum 2021, personal communication).

First, we confirm that the ENS-extended probabilities for upper quintile averaged over the Nzoia basin are highly reliable and sharp at 2 weeks' lead time (Fig. 11a). Next we show the GloFAS forecasts at Rwambwa Bridge during the flooding occurring during the long rains of 2013, Fig. 11b. Probabilities of 99.2 percentile discharge exceedance are indicated (99.2 percentile corresponds with a 5-m exceedance in the river height observations). For the first brief 5-m event on 15 April, forecasts on 20 March show a slightly raised probability of 5% of such an event on this day over 4 weeks ahead. This

probability increases to 20%, 30%, and 50% for the forecasts initialized on 27 March, 3 April, and 10 April, respectively. For the later extended period of flooding beginning 4 May probabilities are even greater; the forecast initialized on 10 April indicates probabilities of 20%, with the following forecasts on 17 April, 24 April, and 1 May indicating probabilities increasing to 30%, 50%, and 100%, respectively.

This case study demonstrates the potential of the GloFAS system. However, a full evaluation is essential in the basin before operational use; we take a first step toward this in Fig. 11c. This shows ensemble median GloFAS discharge forecasts at 15-day lead time plotted against river height, for all GloFAS start dates. Ensemble median forecasts above a threshold indicate instances where action would be triggered, if a 50% probability trigger were used. Verification using other lead times and probability thresholds is forthcoming; however, with this concise analysis we a single hit, five false alarms, and six misses for these specific criteria. This corresponds to a hit rate of 13% and a false alarm ratio of 85%. Clearly GloFAS does not perfectly discriminate all flood events at this lead time. Forecasting a 4-m threshold exceedance event gives 27 hits, 48 false alarms, and 56 misses, corresponding to a hit rate of 32% and a false alarm ratio of 64%, this indicates that GloFAS may be more appropriate for indication of less extreme, but still flood-relevant events.

Finally we note significant forecast busts in GloFAS. For example the model does not appear to give any advance warning of recent flooding in April 2020, shown in Fig. 12 (we note that this event is outside the period of both the Nzoia daily flow record we have access to and the GloFAS reforecast: we therefore examine the operational GloFAS forecast). Five-meter exceedance was seen in late April 2020 after significant rainfall in the week beginning 13 April (Fig. 12a). Verification of the ENS-extended rainfall used to drive GloFAS shows large-scale rainfall to be well forecast several weeks ahead

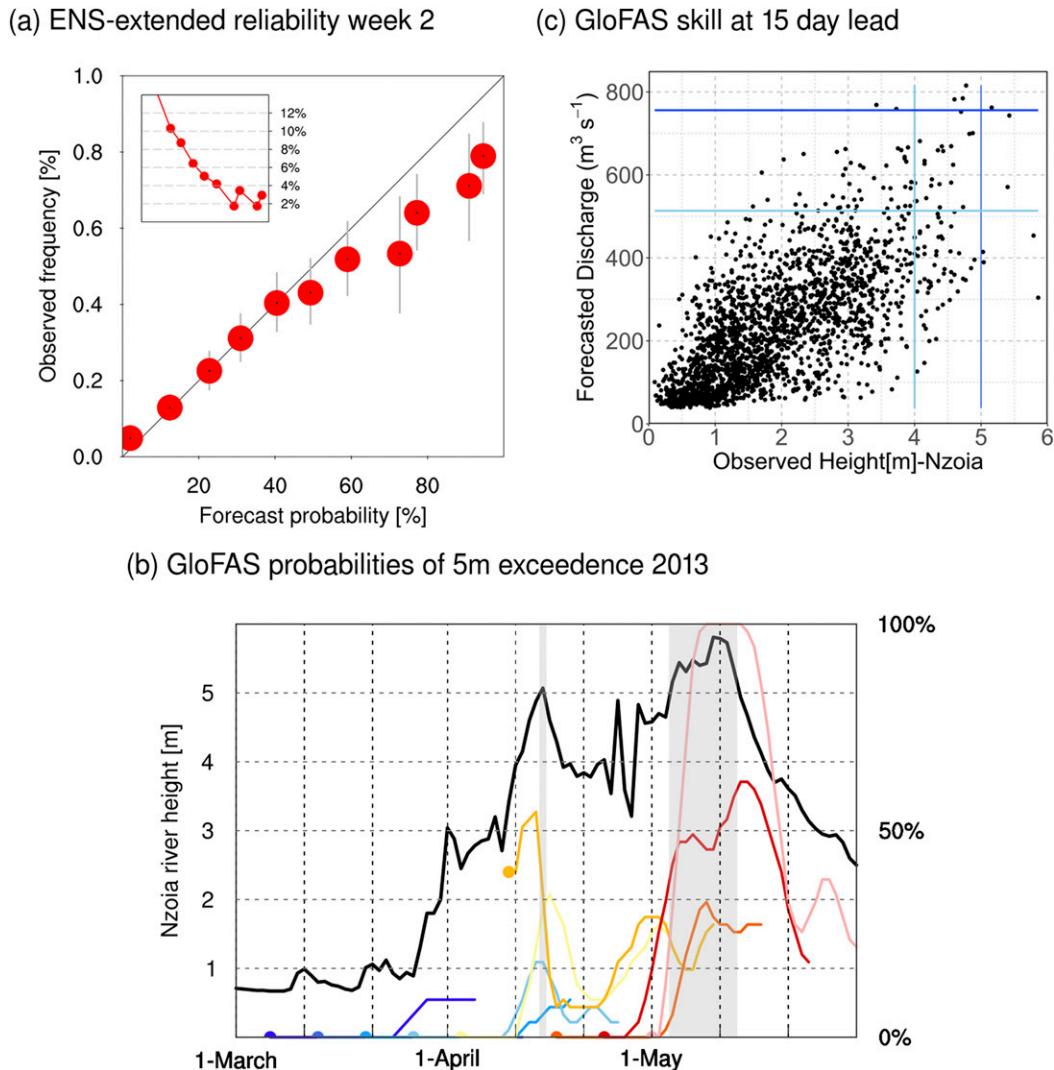


FIG. 11. (a) Reliability of ENS-extended probabilities of weekly Nzoia basin rainfall accumulation exceeding the 80th percentile at lead time 8–14 ahead (bars indicate sampling uncertainty; the size of dots and inset plot show sample size of bins). (b) GloFAS probability of exceedance of a 99.2th percentile event during 2013 (corresponding to 5-m level); a selection of forecasts is shown, with the initial date of each indicated by a circle. Observations are indicated by the black line and gray shading indicates periods where the level exceeded 5 m. Probability forecasts correspond to the right axis; colors are only used to help visually identify individual forecasts. (c) All GloFAS ensemble median values at 15-day lead time plotted against observed height; vertical and horizontal lines show percentile values associated with 4- and 5-m exceedance values shown in light and dark blue.

(Fig. 12b). ENS-extended forecasts initialized on 6 April (for example), show clear indication of the wet event, however the GloFAS run initialized on the same day (Fig. 12c) does not show any indication of enhanced probabilities of high flow, nor do later forecasts (not shown). It should be noted that unlike the reforecast used in Fig. 12, the latest GloFAS operational forecasts have been calibrated for the Nzoia basin based on the gauge data presented here. Yet in this case at least, flooding is not simulated despite a strong indication of heavy rainfall in the driving meteorological data. The reason for this bust is unknown and has not been investigated further; however, given the apparent skill in the rainfall forecast and the analysis

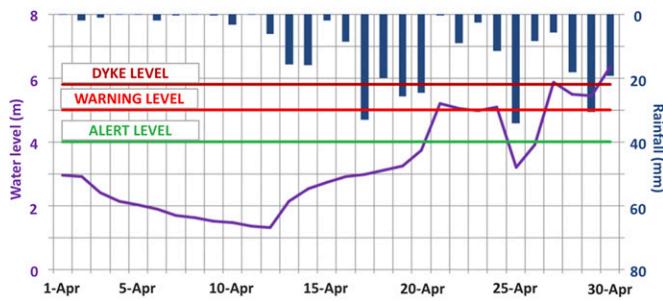
presented in section 5a, it does suggest that a simpler flood prediction using API and weekly rainfall from ENS-extended may have provided a better warning of this event.

6. Conclusions and recommendations

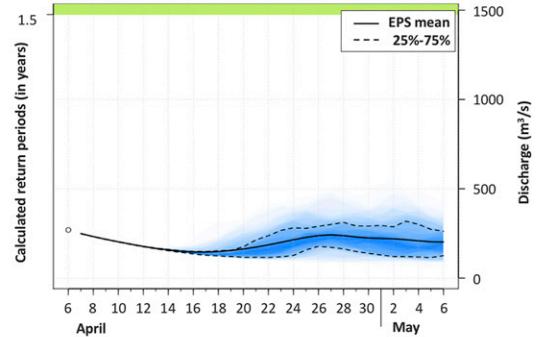
To establish the scientific basis for extending flood related early warnings in EEA with subseasonal forecasts, we set out three research questions, which are reiterated here and addressed in turn:

- What are the large-scale drivers of extreme rainfall?

(a) KMD observations of Nzoia April 2020



(c) GloFAS forecast initialized 6th April 2020



(b) ENS-extended forecasts for 13-19th April 2020

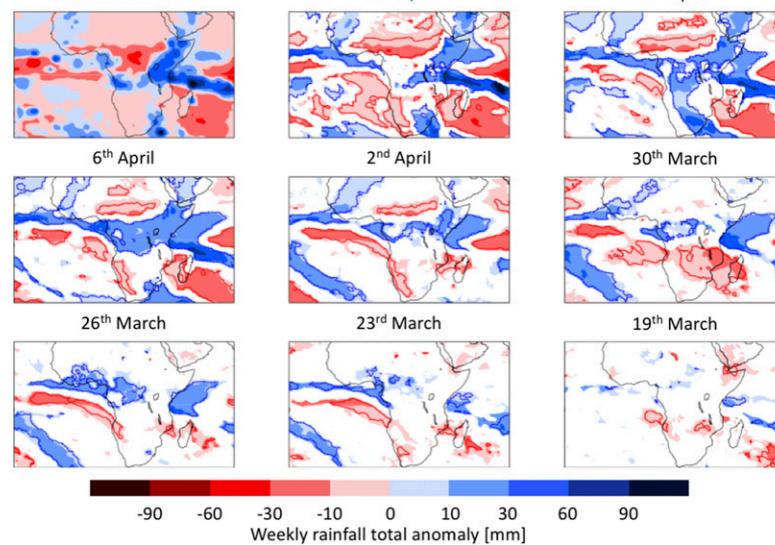


FIG. 12. (a) Courtesy KMD; Nzoia River exceeded 5 m in late April 2020 after significant rainfall in the week beginning 13 Apr. (b) ECMWF ENS-extended ensemble mean anomalies for the week beginning 13 Apr at several weeks' lead (initialization date shown in subtitles), indicate that even forecasts as issued early as 6 Apr indicated heavy rainfall (reproduced from www.ecmwf.int). (c) However, the GloFAS hydrograph initialized on 6 Apr 2020 did not indicate enhanced probabilities of flooding during April (reproduced from www.globalfloods.eu).

- How skillful are rainfall predictions from subseasonal forecast systems and how well do they represent teleconnections with these main sources of predictability?
- What is the link between extreme rainfall and fluvial flooding and what are the implications for flood predictability?

We first demonstrate that the MJO and equatorial Kelvin waves are the main large-scale drivers of extreme precipitation in the region, consistent with previous work on MJO teleconnections (Berhane and Zaitchik 2014; Zaitchik 2017). Consistent with previous work we find high skill for subseasonal predictability of the MJO in leading subseasonal forecasts from ECMWF and the U.K. Met Office (Vitart 2017). This subseasonal predictability is realized in rainfall forecasts, which show probabilities of heavy rainfall weeks which are both reliable and sharp at lead week 2 and beyond.

These findings are consistent with global analysis of subseasonal models (de Andrade et al. 2019) in which EEA can be seen as one of a few locations worldwide with high subseasonal predictability. This high predictability is likely driven by good representation of the MJO to rainfall which we demonstrate here. However, we note some errors which may be limiting forecast skill. In particular, models do not capture the magnitude of the observed strong wet signal associated with MJO activity in phase 4. Future work might attempt to diagnose the reasons for and fix such errors, while the development of mixed dynamical–statistical methods may improve forecast skill by combining observed MJO–rainfall relationships with dynamically predicted MJO forecasts (Glahn and Lowry 1972). On the other hand, given that models do mostly correctly simulate MJO teleconnections, it might reasonably be expected that any improvements in MJO forecasts will translate directly into improved rainfall forecasts for the region. Overall, the results

presented here support the use of subseasonal forecasts to extend early warnings of heavy rainfall over EEA.

Analysis of the Nzoia River identifies both a heavy rainfall trigger and high antecedent rainfall as key drivers of flooding. Both upper-quintile weekly rainfall totals and upper-quintile API anticipate all recent flooding events, although triggering early action on these factors alone would lead to false alarms in the majority of cases. Consideration of subbasin rainfall variability may be necessary to explain the occurrence (and non-occurrence) of flooding. In addition spatial heterogeneity in soil hydrology and land cover may have an important role to play in determining the likelihood of heavy rainfall triggering flooding, by modulating patterns of runoff. High temperatures may also be relevant, reducing the chance for triggering by increasing evapotranspiration fluxes from the surface. Inclusion of all such factors in a well-calibrated hydrological model is likely to be a prerequisite to accurately identifying all historical flood events, while avoiding false alarms.

Finally we assess the GloFAS flood forecasts for Nzoia. Since GloFAS uses the ENS-extended rainfall forecasts (shown to be particularly skillful over the Nzoia basin) it has high potential for skill. It shows promise, by indicating that 2013 flooding was anticipated several weeks in advance. On the other hand, evaluation of the uncalibrated reforecasts shows that use of 2-week lead time forecasts is likely to lead to high numbers of false alarms with many missed events. Operational forecasts for Nzoia have now been calibrated to the observed discharge record through collaboration between KMD and ECMWF. However, we note that the calibrated model does not appear to give any advance warning of recent flooding in April 2020, despite a well-forecast wet event in the driving ENS-extended. We therefore advise caution with the use of GloFAS without a verification of forecast skill, focusing specifically on those extreme flooding forecasts which will be used to trigger actions. This is particularly pertinent as GloFAS is currently being proposed by Kenya Red Cross as a forecast trigger for advanced humanitarian preparedness actions, and has been proposed as a trigger for Red Cross Early Action Plans in many countries. Our results reinforce the need for comprehensive verification. Diagnosis of the reason for GloFAS forecast busts is essential, and such analysis is in progress (A. Kiptum 2021, personal communication). However, proposed actions would have to be very low regret to tolerate the false alarm ratios found here, which are quite high once the event under consideration is extreme enough to correspond to actual flooding.

While GloFAS has significant advantages as a flood forecasting system, a significant drawback is the lack of assimilation of monitoring of soil moisture in real time [the reasons for which are outlined in [Lavers et al. \(2019\)](#)]. Analysis presented here indicates that the precursor hydrological state driven by antecedent rainfall is a key factor in flood generation, and the lack of assimilated soil moisture observations in GloFAS may be significantly limiting forecast skill. By comparison the existing Nzoia forecast system at KMD does utilize soil moisture observations so optimal forecasts may be obtained by running this system with subseasonal rainfall forecasts. However, significant technical barriers are faced with using ensemble data

with the current system. Notably, ForPac has worked with KMD to develop capacity in probabilistic flood through the Met Office providing 5-day forecast of daily Nzoia rainfall with a three-member ensemble, using the 10th, 50th, and 90th percentile of the ensemble, allowing simulation of best, median, worst-case scenario forecasts, respectively.

Another simple-to-implement option for extending lead times is to link observations to subseasonal forecasts to create a subseasonal API forecast. Although using high API and weekly rainfall alone would lead to a relatively large number of false alarms, a simple forecasting tool based on this would have the potential to identify all significant flooding events as long as the rainfall forecast is relatively good, such as the April 2020 flood event, which was missed by GloFAS, despite a good indication of heavy rain in ENS-extended. An API forecast would be quite simple to implement and we recommend the evaluation and development of such a tool during the S2S real-time trial, in which KMD is involved through ForPac.

The analysis presented here supports the development of the existing Nzoia flood forecasting system to run with longer lead subseasonal ensemble forecasts. In addition, moving beyond a basin-average model to represent subbasin distribution of rainfall and spatial heterogeneity in hydrological characteristics may improve forecasts, both by simulating hydrological processes more faithfully as well as better exploiting any predictable information contained in the spatial variability revealed by the subbasin network of hydrometeorological stations, recently extended and upgraded. Flood management in Kenya is being advanced substantially under the ongoing World Bank-funded Kenya Water Security and Climate Resilience Project (<https://projects.worldbank.org/en/projects-operations/document-detail/P117635>); for the Nzoia River the existing flood forecasting system has been upgraded using the lumped catchment GR4J model, but for other flood prone basins initiatives to calibrate and test the suite of MIKE distributed hydrological models are underway. The potential for ensemble forecasting in these cases is evident from our work.

In conclusion there is a clear need for extending flood EWS, and here we show the potential of subseasonal forecasts, which provide reliable and sharp forecasts of extreme rainfall and the drivers of extreme rainfall out to several weeks ahead. EEA is a particular global “sweet spot” for subseasonal predictability and the science supports the development of long-lead forecast-based early action protocols. Moving forward, a key barrier to uptake of the information is that it is not always freely available to National Meteorological and Hydrological Services (NMHS) (for instance, the ECMWF noncommercial annual license to nonmember state NMHS costs over 40 000 euros). The S2S pilot studies are a promising step forward and are currently ongoing through the ForPac and SWIFT (Science for Weather Information and Forecasting, www.africanswift.org), where subseasonal forecasts are being provided to KMD for use in flood and drought early warning across Kenya. This optimism is tempered by the fact that temporary access to real-time forecast data has been provided in the past without continuation beyond the lifetime of the project. However, we hope that the S2S pilot is a bold step toward a long-term sustainable model in which ability to pay is not a barrier to access to the forecasts that can save lives.

Acknowledgments. We thank KMD for provision of the Nzoia streamflow observations, and the S2S project for working toward improving access to subseasonal forecasts and making freely available the forecast data used here. This research was supported by the Science for Humanitarian Emergencies and Resilience (SHEAR) consortium project “Towards Forecast-based Preparedness Action” (ForPac, www.forpac.org), Grants NE/P000673/1, NE/P000568/1, and NE/P000428/1. The SHEAR programme is funded by the U.K. Natural Environment Research Council, the Economic and Social Research Council, and the U.K. Department for International Development. AK is supported by the SHEAR Doctoral Training Cohort. KG is very generously supported through the Peter Carpenter African Climate Scholarship programme. LJ is supported by the U.K. International Development Challenges Fund.

Data availability statement. Subseasonal forecast data used here are freely available through the S2S project database (www.s2sprediction.net). Nzoia flow data are held by KMD and are not publicly available. Other observational databases used are publicly available; details can be found in the relevant citations provided.

REFERENCES

- Alfieri, L., P. Burek, E. Dutra, B. Krzeminski, D. Muraro, J. Thielen, and F. Pappenberger, 2013: GloFAS-global ensemble streamflow forecasting and flood early warning. *Hydrol. Earth Syst. Sci.*, **17**, 1161–1175, <https://doi.org/10.5194/hess-17-1161-2013>.
- Bazo, J., R. Singh, M. Destrooper, and E. C. de Perez, 2019: Pilot experiences in using seamless forecasts for early action: The “Ready-Set-Go!” approach in the red cross. *Sub-Seasonal to Seasonal Prediction*, Elsevier, 387–398.
- Berghuijs, W. R., S. Harrigan, P. Molnar, L. J. Slater, and J. W. Kirchner, 2019: The relative importance of different flood-generating mechanisms across Europe. *Water Resour. Res.*, **55**, 4582–4593, <https://doi.org/10.1029/2019WR024841>.
- Berhane, F., and B. Zaitchik, 2014: Modulation of daily precipitation over East Africa by the Madden–Julian Oscillation. *J. Climate*, **27**, 6016–6034, <https://doi.org/10.1175/JCLI-D-13-00693.1>.
- Copernicus Climate Change Service, 2017: ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate. Accessed 24 August 2020, <https://cds.climate.copernicus.eu/cdsapp#!/home>.
- de Andrade, F. M., C. A. Coelho, and I. F. Cavalcanti, 2019: Global precipitation hindcast quality assessment of the Subseasonal to Seasonal (S2S) prediction project models. *Climate Dyn.*, **52**, 5451–5475, <https://doi.org/10.1007/s00382-018-4457-z>.
- Doi, T., S. K. Behera, and T. Yamagata, 2020: Predictability of the super IOD event in 2019 and its link with El Niño Modoki. *Geophys. Res. Lett.*, **47**, e2019GL086713, <https://doi.org/10.1029/2019GL086713>.
- Duchon, C. E., 1979: Lanczos filtering in one and two dimensions. *J. Appl. Meteor.*, **18**, 1016–1022, [https://doi.org/10.1175/1520-0450\(1979\)018<1016:LFIOAT>2.0.CO;2](https://doi.org/10.1175/1520-0450(1979)018<1016:LFIOAT>2.0.CO;2).
- Fink, A. H., A. Schlueter, R. van der Linden, and J. G. Pinto, 2020: A systematic comparison of tropical waves over Western and Eastern Equatorial Africa. *33rd Conf. on Climate Variability and Change and 100th American Meteorological Society Annual Meeting*, Boston, MA, Amer. Meteor. Soc., 2A.1, <https://ams.confex.com/ams/2020Annual/webprogram/Paper364112.html>.
- Finney, D. L., J. H. Marsham, D. P. Walker, C. E. Birch, B. J. Woodhams, L. S. Jackson, and S. Hardy, 2020: The effect of westerlies on East African rainfall and the associated role of tropical cyclones and the Madden–Julian oscillation. *Quart. J. Roy. Meteor. Soc.*, **146**, 647–664, <https://doi.org/10.1002/qj.3698>.
- Floodlist, 2020: Kenya – Floods hit north and central regions as death toll rises to 237. 13 May, accessed 24 August 2020, <https://perma.cc/3TA2-A6SM>.
- Funk, C., and Coauthors, 2015: The climate hazards infrared precipitation with stations—A new environmental record for monitoring extremes. *Sci. Data*, **2**, 150066, <https://doi.org/10.1038/sdata.2015.66>.
- Glahn, H. R., and D. A. Lowry, 1972: The use of model output statistics (MOS) in objective weather forecasting. *J. Appl. Meteor.*, **11**, 1203–1211, [https://doi.org/10.1175/1520-0450\(1972\)011<1203:TUOMOS>2.0.CO;2](https://doi.org/10.1175/1520-0450(1972)011<1203:TUOMOS>2.0.CO;2).
- Guigma, K. H., F. Guichard, M. Todd, P. Peyrille, and Y. Wang, 2020: Atmospheric tropical modes are important drivers of Sahelian springtime heatwaves. *Climate Dyn.*, <https://doi.org/10.1007/S00382-020-05569-9>.
- Hameed, S. N., 2018: The Indian Ocean Dipole. Oxford Research Encyclopedia of Climate Science, <https://doi.org/10.1093/acrefore/9780190228620.013.619>.
- Hoedjes, J. C., and Coauthors, 2014: A conceptual flash flood early warning system for Africa, based on terrestrial microwave links and flash flood guidance. *ISPRS Int. J. Geoinf.*, **3**, 584–598, <https://doi.org/10.3390/ijgi3020584>.
- Jolliffe, I. T., and D. B. Stephenson, 2012: *Forecast Verification: A Practitioner’s Guide in Atmospheric Science*. John Wiley & Sons, 288 pp.
- Kiladis, G. N., M. C. Wheeler, P. T. Haertel, K. H. Straub, and P. E. Roundy, 2009: Convectively coupled equatorial waves. *Rev. Geophys.*, **47**, RG2003, <https://doi.org/10.1029/2008RG000266>.
- Kilavi, M., and Coauthors, 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**, 472, <https://doi.org/10.3390/atmos9120472>.
- Lavers, D. A., S. Harrigan, E. Andersson, D. S. Richardson, C. Prudhomme, and F. Pappenberger, 2019: A vision for improving global flood forecasting. *Environ. Res. Lett.*, **14**, 121002, <https://doi.org/10.1088/1748-9326/AB52b2>.
- Liebmann, B., and C. A. Smith, 1996: Description of a complete (interpolated) outgoing longwave radiation dataset. *Bull. Amer. Meteor. Soc.*, **77**, 1275–1277, <https://doi.org/10.1175/1520-0477-77.6.1274>.
- MacLachlan, C., and Coauthors, 2015: Global Seasonal Forecast System version 5 (GloSea5): A high-resolution seasonal forecast system. *Quart. J. Roy. Meteor. Soc.*, **141**, 1072–1084, <https://doi.org/10.1002/qj.2396>.
- MacLeod, D., and C. Caminade, 2019: The moderate impact of the 2015 El Niño over East Africa and its representation in seasonal reforecasts. *J. Climate*, **32**, 7989–8001, <https://doi.org/10.1175/JCLI-D-19-0201.1>.
- , M. Kilavi, E. Mwangi, M. Ambani, J. Robbins, R. Graham, P. Rowhani, and M. C. Todd, 2020: Are Kenya Meteorological Department heavy rainfall advisories useful for forecast-based early action and early preparedness for flooding? *Nat. Hazards Earth Syst. Sci.*, **21**, 261–277, <https://doi.org/10.5194/NHESS-21-261-2021>.
- Mason, S. J., and N. E. Graham, 2002: Areas beneath the relative operating characteristics (ROC) and relative operating levels

- (ROL) curves: Statistical significance and interpretation. *Quart. J. Roy. Meteor. Soc.*, **128**, 2145–2166, <https://doi.org/10.1256/003590002320603584>.
- Nicholson, S. E., 2017: Climate and climatic variability of rainfall over eastern Africa. *Rev. Geophys.*, **55**, 590–635, <https://doi.org/10.1002/2016RG000544>.
- OCHA, 2018: OCHA flash update #6: Floods in Kenya 7 June 2018. UN Office for the Coordination of Humanitarian Affairs, accessed 24 August 2020, <https://perma.cc/B47A-HSYF>.
- Pohl, B., Y. Richard, and N. Fauchereau, 2007: Influence of the Madden–Julian oscillation on southern African summer rainfall. *J. Climate*, **20**, 4227–4242, <https://doi.org/10.1175/JCLI4231.1>.
- Richardson, D., and Coauthors, 2011: Verification statistics and evaluations of ECMWF forecasts in 2010–2011. ECMWF Tech. Memo. 654, 53 pp., <https://www.ecmwf.int/sites/default/files/elibrary/2011/11916-verification-statistics-and-evaluations-ecmwf-forecasts-2010-2011.pdf>.
- Riley, E. M., B. E. Mapes, and S. N. Tulich, 2011: Clouds associated with the Madden–Julian oscillation: A new perspective from CloudSat. *J. Atmos. Sci.*, **68**, 3032–3051, <https://doi.org/10.1175/JAS-D-11-030.1>.
- Shilenje, Z. W., and B. A. Ogwang, 2015: The role of Kenya meteorological service in weather early warning in Kenya. *Int. J. Atmos. Sci.*, **2015**, 302076, <https://doi.org/10.1155/2015/302076>.
- Stephens, E., J. J. Day, F. Pappenberger, and H. Cloke, 2015: Precipitation and floodiness. *Geophys. Res. Lett.*, **42**, 10 316–10 323, <https://doi.org/10.1002/2015GL066779>.
- UNIDRR, 2015: Sendai framework for disaster risk reduction 2015–2030. *Proc. Third United Nations World Conf. on DRR*, Sendai, Japan, United Nations Office for Disaster Risk Reduction, 14–18.
- Vitart, F., 2017: Madden–Julian oscillation prediction and teleconnections in the S2S database. *Quart. J. Roy. Meteor. Soc.*, **143**, 2210–2220, <https://doi.org/10.1002/qj.3079>.
- Wainwright, C. M., D. L. Finney, M. Kilavi, E. Black, and J. H. Marsham, 2020: Extreme rainfall in East Africa, October 2019–January 2020 and context under future climate change. *Weather*, **76**, 26–31, <https://doi.org/10.1002/WEA.3824>.
- Waliser, D. E., 2011: Predictability and Forecasting. *Intraseasonal Variability in the Atmosphere–Ocean Climate System*, W. K.-M. Lau and D. E. Waliser, Eds., Springer, 613 pp.
- Weisheimer, A., and T. Palmer, 2014: On the reliability of seasonal climate forecasts. *J. Roy. Soc. Interface*, **11**, 20131162, <https://doi.org/10.1098/RSIF.2013.1162>.
- Wheeler, M., and G. N. Kiladis, 1999: Convectively coupled equatorial waves: Analysis of clouds and temperature in the wavenumber–frequency domain. *J. Atmos. Sci.*, **56**, 374–399, [https://doi.org/10.1175/1520-0469\(1999\)056<0374:CCEWAO>2.0.CO;2](https://doi.org/10.1175/1520-0469(1999)056<0374:CCEWAO>2.0.CO;2).
- White, C. J., and Coauthors, 2017: Potential applications of subseasonal-to-seasonal (S2S) predictions. *Meteor. Appl.*, **24**, 315–325, <https://doi.org/10.1002/met.1654>.
- Wilkinson, E., L. Weingärtner, R. Choularton, M. Bailey, M. Todd, D. Kniveton, and C. Cabot Venton, 2018: Forecasting hazards, averting disasters: Implementing forecast-based early action at scale. Tech. Rep., Overseas Development Institute, 38 pp., <https://www.odi.org/sites/odi.org.uk/files/resource-documents/12104.pdf>.
- Xavier, P., R. Rahmat, W. K. Cheong, and E. Wallace, 2014: Influence of Madden–Julian oscillation on Southeast Asia rainfall extremes: observations and predictability. *Geophys. Res. Lett.*, **41**, 4406–4412, <https://doi.org/10.1002/2014GL060241>.
- Yang, W., R. Seager, M. A. Cane, and B. Lyon, 2015: The annual cycle of East African precipitation. *J. Climate*, **28**, 2385–2404, <https://doi.org/10.1175/JCLI-D-14-00484.1>.
- Zaitchik, B. F., 2017: Madden–Julian oscillation impacts on tropical African precipitation. *Atmos. Res.*, **184**, 88–102, <https://doi.org/10.1016/j.atmosres.2016.10.002>.