Seasonal Outlook on the West African Rainy Season: Monitoring Onset Date Variations using GPM IMERG Satellite-Based Precipitation Data

C. B. JAYASANKAR[a,c] AND VASUBANDHU MISRA[a,b,c]

[a] Center for Ocean-Atmospheric Prediction Studies, Florida State University, Tallahassee, Florida
[b] Department of Earth, Ocean and Atmospheric Science, Florida State University, Tallahassee, Florida
[c] Florida Climate Institute, Florida State University, Tallahassee, Florida

(Manuscript received 9 May 2024, in final form 18 December 2024, accepted 10 January 2025)

ABSTRACT: In this study, we diagnose the onset and demise of the rainy season from the daily rainfall over West Africa. We then produce a probabilistic seasonal outlook of the rainy season over this region based on the observed variations of the onset date of the season, which verifies well with observations. We generated 101 ensemble members at every grid point by randomly perturbing the observed series of daily rainfall data obtained from the Integrated Multi-satellitE Retrievals for Global Precipitation Measurement (IMERG) mission, version 6, rainfall analysis. The generated ensemble of time series of daily rainfall accounts for uncertainties at meso- to synoptic scales that could arise in the generation of the observed rainfall analysis. The ensemble members provide a robust estimate of the onset and demise dates of the rainy season. We find that the interannual variations in the onset and demise dates of the rainy season in West Africa significantly influence the corresponding anomalies of the seasonal length and rainfall. Additionally, the interannual variability of the onset dates dominates over the demise dates of the rainy season across West Africa. In contrast, their association with remote, large-scale forcing is not found to be as significant. In addition, we found that the African easterly jet (AEJ) is displaced southward or northward in early or late onset seasons, respectively. This study highlights the effectiveness of utilizing the intrinsic relationships between onset date, seasonal length, and rainfall anomaly to produce useful seasonal outlooks of the rainy season solely by monitoring the onset date of the rainy season.

SIGNIFICANCE STATEMENT: The West African region comprising of Liberia, Sierra Leone, Guinea, and Guinea-Bissau is critically dependent on the seasonal evolution of the rainfall for their agricultural activities amid significant seasonal-to-interannual variations of their rainy season. In this study, we offer a simple but reliable methodology to provide a seasonal outlook of the rainy season over West Africa by simply monitoring the evolution of the onset date of the rainy season. Given the availability of gridded rainfall analysis at 10-km grid resolution from the Integrated Multi-satellitE Retrievals for Global Precipitation Measurement (IMERG) mission, version 6, at 12-h latency, it is now possible to adapt the proposed methodology for real-time monitoring and seasonal outlook of the rainy season. The veracity of the seasonal outlook of the rainy season from the proposed methodology is analyzed to show that it has much higher skill than a random forecast and could serve as a complementary approach to current practices of seasonal forecast for the region.

KEYWORDS: Atmosphere; Africa; Monsoons; Precipitation; Satellite observations; Interannual variability

1. Introduction

West Africa receives rainfall mainly during the months of June–September as part of the West African summer monsoon (WASM) season. WASM rainfall exhibits significant spatiotemporal variability (Sylla et al. 2010). This holds a profound impact given that a majority of the people in West Africa are dependent on agriculture or agricultural products (Sultan et al. 2005). Sivakumar (1988) demonstrated that there is a substantial association between the length of the growing season and the onset of the rainy season over West Africa. In another related study, Ingram et al. (2002) mentioned that farmers in West Africa frequently identify interannual variations of the monsoon season by variations in the length of the season. For example, the later onset of the rainy season and an earlier demise of the rainy season with a corresponding deficit in the seasonal rainfall are known as a risky season for their crop yield. Hence, reliable forecast estimates of the onset date (OD), demise date (DD), seasonal length (SL), and seasonal rainfall of the WASM season are crucial for planning rain-fed agriculture over the region.

The WASM is characterized by two distinct systems: (i) a strong seasonal wind shift caused by thermodynamic differences between the land (the Sahara) and the ocean (the equatorial Atlantic) and (ii) a moisture-laden southwesterly flow established between the Atlantic cold tongue and the Saharan heat low (Peyrillé et al. 2007; Thorncroft et al. 2011). These two systems have a strong relationship with the annual evolution of moisture fluxes, convergence, and precipitation over West Africa (Nicholson 2013). The lower-tropospheric equatorial westerlies associated with the southwest monsoon flow,

Supplemental information related to this paper is available at the Journals Online website: https://doi.org/10.1175/JHM-D-24-0067.s1.

Corresponding author: C. B. Jayasankar, cbjayasankar@gmail.com

DOI: 10.1175/JHM-D-24-0067.1

the midtropospheric African easterly jet (AEJ), and the upper-level tropical easterly jet (TEJ) are the major tropical circulation features linked to the WASM. The amount of WASM rainfall and its variabilities in each season are determined by the strength and location of these circulation patterns (Redelsperger et al. 2002). There have been many previous studies that have discussed the variability of the WASM and its teleconnections at the interannual scales (e.g., Rowell et al. 1992; Janicot et al. 2001; Giannini et al. 2005; Joly et al. 2007; Joly and Voldoire 2009; Li et al. 2012). Many of these studies are focused on the El Niño–Southern Oscillation (ENSO)–WASM teleconnection. In contrast, the teleconnection of WASM with other external forcing (like the tropical Atlantic or the Indian Ocean variability) is comparatively fewer (Losada et al. 2010; Gu and Adler 2004; Flaounas et al. 2012).

Several studies have been undertaken to determine the onset of the WASM by using several meteorological variables such as precipitation, moist static energy, outgoing longwave radiation, mean sea level pressure, and zonal winds (Sultan and Janicot 2003; Fontaine et al. 2008; Nguyen et al. 2014; Fitzpatrick et al. 2015; Dunning et al. 2016). In a comprehensive review, Fitzpatrick et al. (2015) listed 18 definitions for the WASM onset dates which included both the regional and local onset definitions. According to Sultan and Janicot (2003), the abrupt northward shift of the intertropical convergence zone (ITCZ) from 5° to 10°N is marked as the onset of the WASM. In addition to this northward shift of the ITCZ, the AEJ also shows latitudinal migration at the time of the onset of the WASM (Redelsperger et al. 2002). Drobinski et al. (2009) found that the anomalous northeasterly direction of the AEJ and its interaction with the orography can cause delayed onset in West African countries such as Niger and Mali. Sylla et al. (2010) revealed that the meridional displacement of the core of the AEJ during the WASM has a significant impact on the intraseasonal changes in the location of the rain belt. For instance, AEJ advances in an equatorial direction favor dry conditions over the Sahel. However, some studies use a simple objective method (similar to the method adopted in this study) by utilizing only the daily rainfall time series to define the onset and demise of West Africa (e.g., Dunning et al. 2016).

Many studies report that Africa has a lack of in situ precipitation measurements (e.g., Maidment et al. 2013; Dezfuli et al. 2017; Atiah et al. 2020). Due to the sparse rain gauge observation network, an alternative is to use gridded merged rain gauge–satellite precipitation datasets for the analysis (Diaconescu et al. 2015; Satgé et al. 2020). One of the most often used precipitation datasets in this area is the Tropical Rainfall Measuring Mission (TRMM) Multisatellite Precipitation Analysis (TMPA; Huffman et al. 2010). In a related study, Dunning et al. (2016) used multiple observations and reanalysis of precipitation datasets to identify the onset and demise of the WASM. It is important to note that Dunning et al. (2016) found that onset/demise dates of the WASM from other observations and TMPA observations did not differ significantly, despite the satellite measurements being known to overestimate precipitation over West Africa (Macharia et al. 2022). The recently released Integrated Multi-satellitE Retrievals for Global Precipitation Measurement (GPM) (IMERG; Huffman et al. 2019) provides better quality and spatiotemporal resolution of precipitation data than the TMPA (Dezfuli et al. 2017). Dezfuli et al. (2017) find that IMERG improves upon TMPA at diurnal scales. However, at other time scales including the seasonal cycle, the evaluation metrics yield results that vary by season, region, and the metric to compare IMERG and TMPA across Africa.

Since the WASM exhibits a strong seasonal cycle and year-to-year fluctuations, it is important to identify the features of these variations of the WASM. There are very few studies that have used IMERG observations to study the summer monsoon over West Africa (e.g., Becker et al. 2021). Hence, the primary objective of this study is to investigate the seasonal variability of the rainy season over the WASM region using the IMERG rainfall product. Since only rainfall is used in this study, we make a subtle distinction of the rainy season from the WASM season with the former primarily concerned with the seasonal evolution of rainfall and the latter would include other upper-air variables like circulation, thermal gradients. This study differs from previous studies that employed rainfall-based definitions for onset/demise, such as Dunning et al. (2016), by using a perturbed ensemble of daily rainfall data to define onset/demise (Misra et al. 2023). Unlike previous studies, we employ a perturbation technique to construct 101 ensemble members, allowing for a more comprehensive account of both observational and analytical uncertainties, thus providing a robust estimate of the onset and demise dates of the WASM season. In addition, the choice of using the 12-h latency product of IMERG offers flexibility for use in real-time monitoring of the rainy season as demonstrated over Florida in Misra et al. (2022). Hence, in this study, we also intend to explore the possibility of real-time monitoring of the rainy season onset date at a local scale, which could be useful for anticipating anomalous rainy seasons. In addition, this study explores the interannual variation of the evolution of the rainy season and its link to the AEJ and large-scale forcing. The paper is organized as follows: section 2 explains the datasets and the methods used in the study. The results are presented in section 3 with concluding remarks in section 4.

2. Data and methodology

a. Datasets

This study used the IMERG, version 6 (Huffman et al. 2019), precipitation dataset, which comes from the GPM mission launched in 2014 and operated jointly by the National Aeronautics and Space Administration (NASA) and the Japan Aerospace Exploration Agency (JAXA). The GPM mission follows the TRMM, which was launched in 1997, as its successor. When compared to TRMM, GPM provides wider spatial coverage and a higher temporal resolution and can also identify light rain events (<0.5 mm h^{-1}). The GPM Core Observatory carries two most advanced precipitation-measuring instruments in space: the dual-frequency precipitation radar (Ka band and Ku band) and a conically scanning multifrequency passive microwave radiometer called the GPM Microwave Imager (10–183 GHz).

IMERG data are accessible at 0.1° grid spacing at half-hourly intervals from June 2000 to the present. It has early,

late, and final run products that have ~4, ~12 h, and ~3.5 months latency. In this study, we used the IMERG late daily precipitation dataset with 12-h latency. The IMERG late run uses a variety of data sources, including satellite microwave and infrared estimations, precipitation gauge analysis, and potentially other precipitation estimators, covering fine time and space scales for the TRMM and GPM eras globally. The incorporation of both forward and backward propagation improves its interpolation process. In addition to IMERG, we have used the Climate Prediction Center (CPC) unified gauge-based analysis of global daily precipitation product from the National Oceanic and Atmospheric Administration (NOAA; Xie et al. 2007). This rain gauge–based product is accessible from 1979 to the present at 0.5° grid spacing. This dataset makes use of gauge reports from 30 000 stations around the global land areas. We also used the high-resolution daily Climate Hazards Group Infrared Precipitation with Station (CHIRPS) data precipitation data (Funk et al. 2015), which was created by blending rain gauge observations with infrared satellite data. We also used Niño-3.4, North Atlantic Oscillation (NAO), tropical southern Atlantic (TSA) index, and tropical northern Atlantic (TNA) index provided by NOAA's Physical Sciences Laboratory. The daily rainfall at six stations over West Africa was obtained from Global Surface Summary of the Day (GSOD), which is a by-product of the integrated surface hourly (ISH) dataset (NOAA NCEI 1999). We also used the zonal wind at 600 hPa, which is obtained from the fifth major global reanalysis produced by ECMWF (ERA5; Hersbach et al. 2023).

b. Methodology

1) Onset and demise dates

We used a rainfall-based objective definition identical to that developed by Liebmann and Marengo (2001) and Liebmann et al. (2007), for defining the onset date and demise date of the WASM. For a year n, the cumulative anomaly of the daily precipitation at a day m [i.e., $D_n(m)$] at each grid point is computed as

$$D_n(m) = \sum_{m=1}^{M} [P_n(m) - \overline{P}], \tag{1}$$

where $P_n(m)$ is the precipitation for the day m, M is the last day of the year (31 December), and $\overline{P}$ is the annual mean precipitation climatology. The $D_n(m)$ for 365/366 days in a year is termed as the cumulative precipitation anomaly curve. The onset and demise dates are determined by finding the minima and maxima of the cumulative anomaly curve of the precipitation, respectively. The seasonal length is estimated as the number of days between the demise and the onset dates. The daily precipitation is added up from the onset dates to the demise date for estimating the seasonal accumulated precipitation of the rainy season.

The goal of perturbing the time series is to account for the uncertainty introduced by random synoptic or mesoscale events, which may be unrelated to the seasonal cycle and may have an impact on the determination of the onset or demise date of the rainy season. Because the proposed methodology is at the granularity of a relatively high spatial resolution rainfall analysis, there is a high risk of incorrectly diagnosing the onset or end dates of the rainy season from random rain events that are unconnected to the seasonal cycle. As a result, it is critical to create an ensemble of rainy season onset/demise dates that considers this uncertainty.

The original time series at each grid point is perturbed by shuffling the series and picking the rainfall for each day randomly from the range of ±3 days of the original time series. The ±3-day range used to generate the perturbed time series accounts for uncertainties from the occurrence of synoptic (1–7 days) and mesoscale (1–3 days) rain features. This process is repeated 100 times to generate an ensemble of 101 time series (100 perturbed + 1 original). By doing so, extremely anomalous onset/demise dates diagnosed from random rain events unconnected to the seasonal cycle are likely to appear far less frequently in the ensemble. It should be noted that a larger ensemble did not affect the results presented in this paper (not shown).

2) Signal-to-noise ratio

Since an ensemble of 101 time series is generated, we can compute the signal-to-noise ratio (SNR) of the four quantities as

$$\mathrm{SNR} = \frac{\sigma^2_{\mathrm{signal}}}{\sigma^2_{\mathrm{noise}}}, \tag{2}$$

where

$$\sigma^2_{\mathrm{noise}} = \frac{1}{M(N-1)} \sum_{i=1}^{M} \sum_{j=1}^{N} (A_{ij} - \overline{A}_i)^2, \tag{3}$$

$$\sigma^2_{\mathrm{signal}} = \sigma^2_{\mathrm{em}} - \frac{1}{N} \sigma^2_{\mathrm{noise}}. \tag{4}$$

The variance of the noise is represented as $\sigma^2_{\mathrm{noise}}$, the variance of the signal is represented as $\sigma^2_{\mathrm{signal}}$, A could be any one of the onset date, demise date, seasonal length, and seasonal rainfall anomalies, i and j are the indices for individual years and individual ensemble members, respectively, M is the total number of years, and N is the total number of ensemble members. The σ^2_{em} represents the variance of the ensemble mean.

$$\sigma^2_{\mathrm{em}} = \frac{1}{(M-1)} \sum_{i=1}^{M} (\overline{A}_i - \overline{\overline{A}})^2, \tag{5}$$

where $\overline{A}_i = [1/(N-1)] \sum_{j=1}^{N} A_{ij}$ is the ensemble mean for year i and $\overline{\overline{A}} = [1/(M-1)] \sum_{i=1}^{M} \overline{A}_i$ is the ensemble mean climatology of the variable.

When SNR < 1, it signifies that noise is dominant, and when SNR $\gg 1$, then the signal dominates. A dominating signal suggests that the perturbations of the time series will not significantly modulate the "A" diagnostic of the time series.

3) AUC

In this study, we use the area under the receiver operating characteristic (ROC) curve (AUC) method to analyze the probabilistic skill of predicting anomalous seasons based on the diagnosis of the ensemble of onset dates of the rainy season across the Roberts flight information region (FIR; includes Liberia, Sierra Leone, and Guinea), and the Guinea-Bissau region (hereafter LSGB). The AUC has been commonly used in several earlier studies to assess prediction skill (Mason and Graham 2002; Misra 2004; Narotsky and Misra 2022). The steps to estimate the AUC are provided in section S1 in the online supplemental material. The ROC curve is obtained from plotting the hit rate to the false alarm rate for probability forecasts of three equiprobable categories of below normal, normal, and above normal. The onset date, seasonal length, and seasonal rainfall are divided into terciles based on their rank order, where the lower tercile indicates an early onset, shorter season, and drier season, while the upper tercile indicates a delayed onset, lengthier season, and wetter season. The hit and the false alarm rates are prepared from a contingency table (supplemental material Table 1) to evaluate the probability of categorical forecasts, such as an early or late onset leading to a shorter or longer, and wetter or drier seasons, respectively. The AUC is determined by using the trapezoid rule. An AUC value of 0.5 or less indicates that the model's predictions are no better than random chance. However, an AUC value over 0.5 is considered skillful. An AUC value of 1.0 implies a perfect forecast with no false alarms or misses, and all forecasts are precisely verified.

In addition to using AUC, we also used Heidke skill score (HSS; Heidke 1926) to support our assessment on seasonal forecast verification. Like the AUC, HSS is also used to measure the skill of a categorical forecast (or binary outcomes, e.g., yes/no events) compared to a reference, such as random chance or climatology.

c. Study area

The WASM is a component of the global monsoon system and is responsible for receiving nearly 80% of West Africa's annual rainfall during the period between June and September (Akinsanola and Zhou 2020). West Africa has two major rainfall-receiving areas: one is located over the Roberts FIR, and the other is located over the land region adjacent to the Gulf of Guinea. The annual total rainfall in these regions ranges between 3000 and 4500 mm yr^{-1}. These two regions have a distinct wet season and are comparatively elevated than the rest of West Africa with elevation varying between ~1500 and 3300 m. In this study, we focus on the Roberts FIR region and the Guinea-Bissau region (i.e., LSGB region outlined in Fig. 1a). The LSGB region has a wet tropical climate with a distinct wet season. The location of the FIR within LSGB is significant due to its role in facilitating major flight operations to multiple West African countries. Additionally, LSGB encompasses the drainage basins of significant rivers like the Niger, Senegal, and Gambia. Changes in rainfall patterns within this region directly influence water availability, river flow, and ecosystem dynamics, consequently impacting human populations and biodiversity.

3. Results

Primarily, we looked at the spatial distribution of the 23-yr annual mean rainfall climatology from IMERG shown in Fig. 1a. It shows that the coastal areas west of the LSGB region receive high rainfall annually, with precipitation gradually decreasing toward the east. It is important to verify the spatiotemporal distribution of IMERG 12-h latency rainfall across the LSGB region in comparison to other well-known observations. In Fig. 1b, we depict the scatter between the area-averaged 23-yr (2001–23) daily IMERG rainfall data and the CHIRPS/CPC rainfall over the LSGB region. The CHIRPS data are commonly employed in studies over the African region (Dunning et al. 2016, 2018; Satgé et al. 2020; Nicholson and Klotter 2021; Palmer et al. 2023) due to its high spatial resolution (0.05° × 0.05°), extensive temporal coverage (1981 to the present), and superior performance in this region. The IMERG aligns closely with the CHIRPS, with a Pearson correlation coefficient (CC) of 0.87, a Spearman rank-order correlation of 0.93, and a root-mean-square error (RMSE) of 3.3 mm day^{-1}. In addition, we utilized the CPC dataset to compare IMERG, which is a gauge-based precipitation analysis. Satgé et al. (2020) compared 23 gridded rainfall datasets and observed that CPC tends to consistently underestimate rainfall over West Africa. However, when compared to CPC rainfall data, IMERG tends to overestimate the rainfall, with a Pearson correlation coefficient of 0.51, Spearman rank-order correlation of 0.71, and an RMSE of 5.86 mm day^{-1}. The disparity between IMERG and CPC could be attributed to either a smaller number of gauges available for CPC (e.g., Satgé et al. 2020) or to the inherent tendency of satellite-based products to overestimate rainfall (Dezfuli et al. 2017).

Figures 1c and 1d show the quantile–quantile (Q–Q) plot which compares the area-averaged daily rainfall quantiles of IMERG analysis with CHIRPS and CPC over the LSGB region. Likewise, Fig. 1e shows the Q–Q plot between CHIRPS and CPC. Figure 1c shows that the rainfall data from IMERG and CHIRPS in the lower quartile range (from 0 to ~15 mm) are nearly comparable to each other. However, at larger quantiles (over ~15 mm), the distributions diverge from the 1:1 line, with CHIRPS displaying higher rainfall values than IMERG. This shows that CHIRPS exhibit some fraction of the rainy days with more intensity than IMERG. Similarly, when we look at the Q–Q plot for IMERG and CPC rainfall (Fig. 1d), the lower quantiles (from 0 to ~25 mm) are below the 1:1 line which indicates an overestimation of rainfall from IMERG relative to CPC. On the other hand, at higher quantiles (over ~25 mm), it is above the 1:1 line which indicates an underestimation of rainfall by IMERG in comparison to CPC. Similarly, Fig. 1e shows the comparison between CHIRPS and CPC, which indicates an underestimation of rainfall by CPC relative to CHIRPS at most quantiles. The point of these comparisons is to suggest that there is uncertainty in the gridded rainfall analysis that has been acknowledged in earlier studies as well (Satgé et al. 2020; Dos Santos et al. 2022). The paucity of rain gauge data in the region affects rainfall analysis directly dependent on them (e.g., CPC rainfall analysis), while also affecting the calibration of the remotely sensed rainfall products

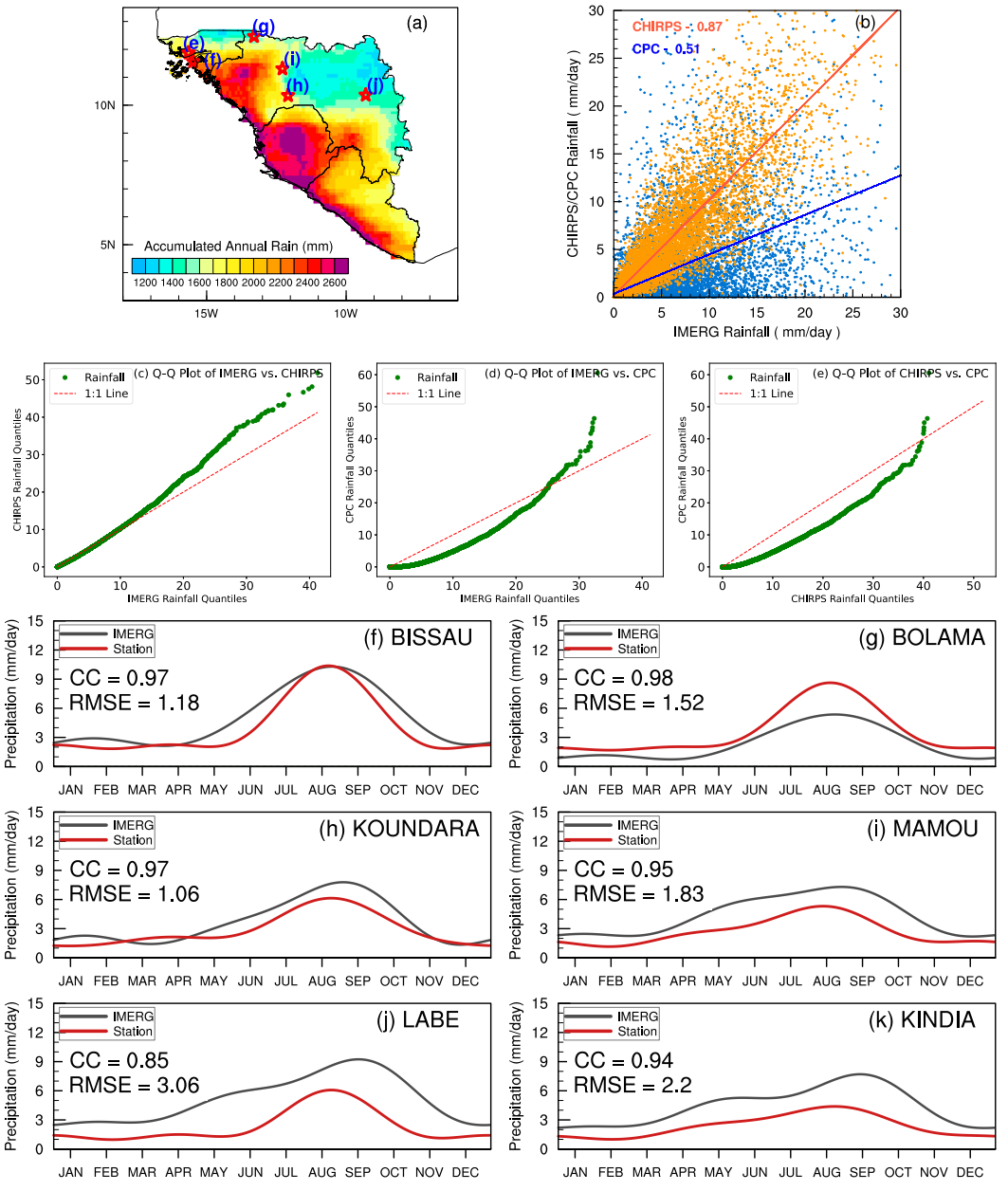

FIG. 1. (a) The spatial distribution of annual accumulated rainfall climatology (mm) over the Guinea coast of West Africa (LSGB) using IMERG data. (b) The scatter between the area-averaged 23-yr (2001–23) daily IMERG rainfall (mm day^{-1}) and the CHIRPS (orange dots) and CPC (blue dots) rainfall over the LSGB region. The corresponding Pearson CC between IMERG and CHIRPS and IMERG and CPC are indicated in the legend in orange and blue, respectively. The orange and blue lines are the best-fit lines for the scatter of IMERG rainfall data with corresponding CHIRPS- and CPC-based rainfall, respectively. The Q–Q plot for (c) IMERG vs CHIRPS, (d) IMERG vs CPC, and (e) CHIRPS vs CPC. The climatological seasonal cycle of IMERG rainfall corresponds to six stations such as (f) Bissau, (g) Bolama, (h) Koundara, (i) Mamou, (j) Labe, and (k) Kindia. The CC and RMSE (mm day^{-1}) for (f)–(k) are marked inset.

that are then used in multisourced rainfall analysis (e.g., CHIRPS, IMERG). However, despite these uncertainties, earlier studies find that the phase of the seasonal cycle is comparable among many of these gridded rainfall products (Dunning et al. 2016; Satgé et al. 2020; Dos Santos et al. 2022), which is critical for this study as we are interested to diagnose the timing of the onset and demise of the rainy season.

We have validated the IMERG rainfall against six GSOD station observations across LSGB (these six stations marked in Fig. 1a were selected based on the availability of more than

10 years of data between 2001 and 2023). The climatological seasonal cycle comparison is shown in Figs. 1f–k, and the seasonal cycle is calculated by adding the first four harmonics of daily precipitation climatology (Murakami et al. 1986; Jayasankar et al. 2015). The seasonal cycle of IMERG with the corresponding GSOD observations at all six stations compares favorably with regard to the seasonal phase, but the magnitude of the seasonal cycle is overestimated at five stations and underestimated at Bolama (Fig. 1g). The range of the average difference in the seasonal cycle between GSOD and IMERG is as follows: from −0.11 to 2.28 mm for Bissau, from −3.28 to −0.49 mm for Bolama, from −0.64 to 2.00 mm for Koundara, from 0.47 to 2.97 mm for Mamou, from −1.05 to 5.15 mm for Labe, and from 0.79 to 3.88 mm for Kindia. Some of the differences in these figures could be attributed to the inherent issues of calibration of IMERG data with in situ data over the region (Dezfuli et al. 2017; Satgé et al. 2020). Four stations, including Bissau Osvaldo Vieira International Airport (Guinea-Bissau), Bolama (Guinea-Bissau), Koundara (Bissau), and Mamou (Bissau), had Pearson correlation coefficients of 0.95 or higher, with RMSE values ranging from 1.06 to 1.83. The stations Labe (Bissau; 0.85, 3.06 mm day^{-1}) and Kindia (Bissau; 0.94, 2.2 mm day^{-1}) had slightly lower correlation coefficients and higher RMSE, respectively. In addition, the Q–Q plot for IMERG and the six GSOD stations (Fig. S1) indicates significant inconsistency across all six stations. It is important to note that, while the IMERG dataset is temporally continuous, the GSOD dataset contains many missing days, which could also contribute to these discrepancies.

Based on the area-averaged rainfall over the LSGB region, we estimated the onset and demise dates of the rainy season for all years which is shown in Fig. 2. The wet season typically starts between April (earliest onset recorded on 21 April 2010) and June (latest onset recorded on 10 June 2019). The peak of the wet season usually occurs from August to September, and the season typically ends around October (earliest demise recorded on 16 October 2017) and November (latest demise recorded on 13 November 2021). The standard deviation for the onset date is 13.6 days, while for the demise dates, it is 6.7 days. We estimated the seasonal length (i.e., the number of days between the onset and demise) for all the years which varies from 136 days (in 2019) to 198 days (in 2010). Similarly, the seasonal rainfall accumulated from onset day to demise day of the rainy season varies from 1262.5 mm (in 2020) to 1912.6 mm (in 2023).

The 23-yr climatological local onset, demise, seasonal length, and seasonal accumulated rainfall (SR) of the LSGB region are presented in Figs. 3a–d. In general, the southern part of the LSGB region experiences an early onset (Fig. 3a), while delayed onset is typical in the northern regions. However, early demise (Fig. 3b) tends to occur in the eastern regions, whereas late demise is more common in the western regions. The spatial pattern of the length of the seasons (Fig. 3c) closely follows those of the onset date and the demise dates. Additionally, the spatial distribution of seasonal accumulated rainfall (Fig. 3d) indicates higher rainfall over the western parts of the LSGB, particularly in Sierra Leone, the northwest of Guinea, and the south of Guinea-Bissau, compared to the eastern parts of LSGB. Figures 3e and 3f illustrate the spatial distribution of

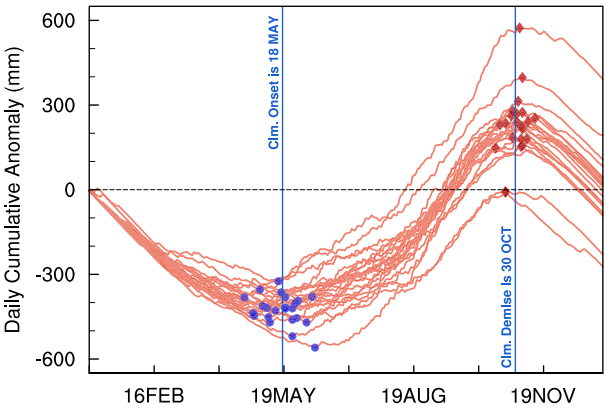

FIG. 2. Individual year (2001–23) cumulative anomaly curves for the area-averaged rainfall over the LSGB region. Each year's OD is marked as filled circles, the DD is marked as rhombus, and the climatological onset and demise are marked as vertical lines.

standard deviations for the onset, demise, seasonal length, and seasonal accumulated rainfall. Figure 3e highlights strong variability (10%–15% of the mean) in the onset, primarily in the southeast regions of the LSGB region. The variability of the demise date is considerably low (Fig. 3f) relative to the onset date (Fig. 3e), while the season's length exhibits high variability (Fig. 3g). Furthermore, there is notable high standard deviation in seasonal total rainfall found over the southwest LSGB region.

First, we estimated the correlation coefficient between onset, demise, seasonal length, and seasonal accumulated rainfall by using area-averaged rainfall over the entire LSGB region. These correlation coefficients are useful for identifying the possible relationships between these four quantities. We noticed a strong, significant negative correlation between the onset date and seasonal rain, which was calculated to be −0.72. Additionally, a strong significant negative correlation was found between the onset and seasonal length, with a coefficient of −0.89. Similarly, we estimated the correlation coefficient between these four variables at each grid point to analyze their spatial variability (Figs. 4a–e). The correlation between onset dates and seasonal lengths (Fig. 4a) reveals a significant negative correlation (<-0.7) over the majority of the grid points. Similarly, the correlation between onset dates and seasonal accumulated rainfall (Fig. 4b) also indicates a significant negative correlation in many grid points. These significant linear correlations suggest that seasons with earlier or later onset dates in the LSGB region are likely to have lengthier and wetter or shorter and drier seasons, respectively. These findings are consistent with a model study by Dunning et al. (2018), who found that a later onset of the rainy season resulted in a shorter length and a decrease in total seasonal rainfall, with more than 75% of the models agreeing. Furthermore, Cook and Vizy (2012) discovered that a delayed onset date is associated with a reduction in the number of growing season days west of 0°.

The correlation between demise dates and seasonal lengths (Fig. 4c) exhibits a significant positive correlation over the northern LSGB region. However, the association between

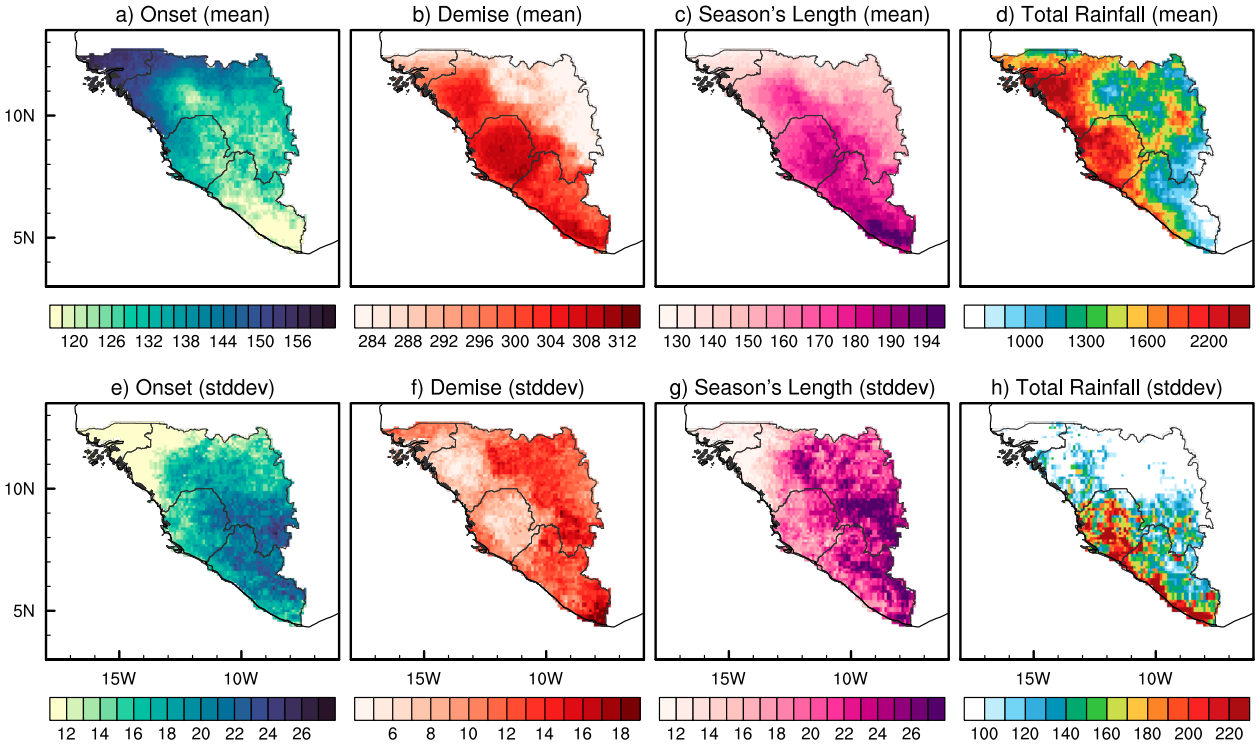

FIG. 3. The 23-yr (2001–23) (a)–(d) climatological mean and (e)–(h) standard deviation of (a),(e) OD (Julian day), (b),(f) DD (Julian day), (c),(g) SL (days), and (d),(h) SR (mm) from the IMERG dataset.

demise dates and seasonal accumulated rainfall (Fig. 4d) is comparatively weaker. Furthermore, the onset date does not show a significant correlation with variations in the demise date (Fig. 4e). The significant correlations observed with the onset date variations (Figs. 4a,b), together with the prominent interannual variations (Figs. 3e–h), emphasize the importance and potential for forecasting these variables reliably from monitoring the onset date of the wet season. Our findings suggest it is possible to provide an outlook of the seasonal length and seasonal rainfall anomaly of the wet season of LSGB by solely monitoring the onset date. Using a similar strategy, Misra et al. (2022) carried out a real-time seasonal outlook for Florida's rainy season. This method has also proven to be effective for real-time seasonal outlooks for the rainy season of northern Australia (Misra et al. 2023).

We utilized the 101-member ensemble of time series data to assess the signal-to-noise ratio of four quantities (onset, demise, seasonal length, and seasonal accumulated rainfall) at each grid point. This metric helps measure the spread within the ensemble members relative to the externally forced or persistent signal (Shukla et al. 2000). For the onset dates and demise dates of the LSGB, the signal-to-noise ratios are mostly below 1 across many grid points (Figs. 5a,b). This suggests that there is more noise than signal in the variability of these dates, indicating a high level of ensemble spread. In contrast, a noticeable north–south gradient is observed in the signal-to-noise ratios of the seasonal length (Fig. 5b). The values are slightly higher than 1 over the northern LSGB, indicating comparatively less spread, whereas in the southern region, the ratio falls below 1. In comparison to the other quantities, the seasonal accumulated rainfall generally exhibits higher signal-to-noise ratios (>1) across most grid points (Fig. 5d). This indicates that the signal dominates the noise, resulting in the likelihood of better predictability compared to the other quantities.

Figure 6 shows the probabilistic skill scores derived from AUC for the six distinct categories. The anomalous seasons (Figs. 6a–d) exhibit higher skills (AUC ≥ 0.5) across LSGB than the normal seasons (Figs. 6e,f). This is because the prediction of seasonal length and rainfall is based on their linear relationships with the onset date variations. Therefore, the larger the anomalies, the higher the likelihood of their verification, which implies higher AUC. In Fig. 6a, the season with an early start and lengthier season shows the most skill (>0.9) over Sierra Leone and northwest Guinea. Similarly, the early start of the season with the wetter season (Fig. 6b) and the late start with the drier season also show high skill scores across most of the areas. Whereas, the late start and associated shorter season (Fig. 6c) exhibit the highest skill score over Sierra Leone, east Guinea, and some parts of Liberia. In contrast, the seasons with normal start show AUC skill scores of less than 0.85 across the majority of the grid points (Figs. 6e,f). However, the skill scores of a normal start of the season with normal rain (Fig. 6f) are slightly better than those for a normal start season with a normal seasonal length (Fig. 6e). This is because the demise dates of the rainy seasons also have a significant

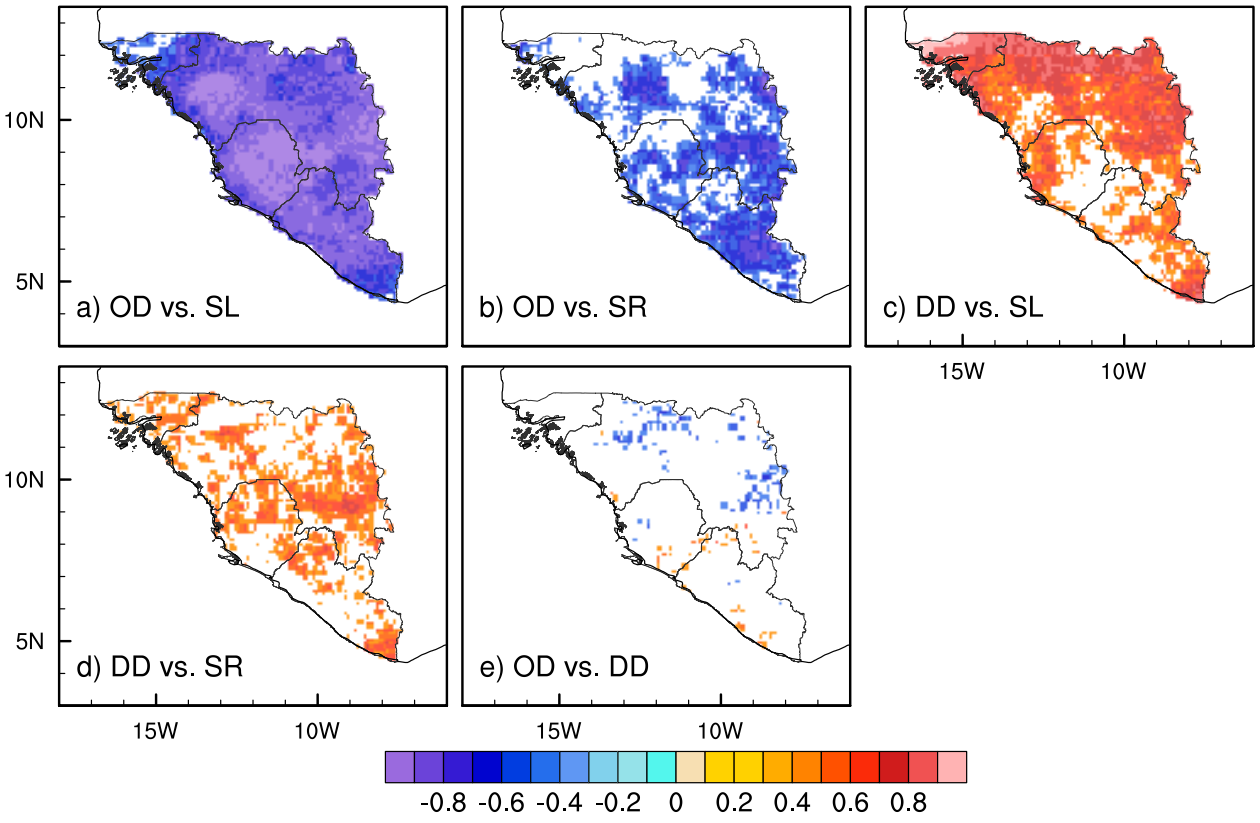

FIG. 4. The CC of (a) the ODs with SLs, (b) the ODs with SR, (c) the DDs with SLs, (d) the DDs with SR, and (e) the ODs with DDs. Shading indicates significant correlations at a 95% confidence interval as determined by the *t* test.

bearing on the length of the seasons, and the former is insignificantly related to onset date variations.

To support our assessment, we have also estimated the HSS for the six distinct categories (Fig. S2), and the results are in line with the AUC. The HSS measures how well the prediction performs relative to a reference, with values ranging from −1 (worse than the reference) to 1 (perfect prediction). The anomalous seasons (Figs. S2a–d) exhibit larger positive HSS across LSGB compared to the normal seasons (Figs. S2e,f). The figures show that early start with longer seasons (Fig. S2a) and wetter seasons (Fig. S2b), as well as late start with shorter (Fig. S2c) and drier seasons (Fig. S2d), have high skill scores (>0.3), especially over Sierra Leone and Guinea. In contrast, normal-start seasons show lower skill scores (<0.35).

We further investigated the interannual variability of the onset dates over the LSGB region and its association with the AEJ. The AEJ appears as the midtropospheric (~600 hPa) easterly jet, which flows across West Africa, often between latitudes of 10° and 15°N, and lasts around 6 months, peaking in July with the strongest easterly winds of 10–11 m s^{-1} (Wu et al. 2009). Drobinski et al. (2009) suggests that the onset of the WASM is modulated by the variations of the AEJ and its interaction with the orography. Figure 7 exhibits the correlation between the 23-yr onset dates over the LSGB region and the seasonal average of zonal winds at 600 hPa obtained from ERA5. The seasonal mean is based on the average of the zonal wind from the onset date to the demise date of the rainy season. It is found that the onset date has a significant negative correlation to the north of the jet (between 15° and 25°N) and a significant positive correlation to the south of the jet (between 2° and 8°N). This indicates that the AEJ shifts slightly southward or northward during early or late onset seasons, respectively.

Further, we investigated the interannual variability of these four quantities with the other potential external forcings such as ENSO, NAO, TNA, and TSA. Previous studies show that the WASM exhibits high sensitivity to SSTs in the global tropics (Giannini et al. 2005; Joly et al. 2007). Several studies suggest teleconnection between ENSO and the WASM rainfall anomalies at interannual time scales (Rowell et al. 1992; Janicot et al. 2001; Joly and Voldoire 2009; Emmanuel 2022). Joly et al. (2007) found that El Niño events exhibit significant correlation with drought conditions in the Sahel and northern parts of LSGB. Emmanuel (2022) studied the impact of ENSO and showed El Niño (La Niña) years correspond to positive (negative) rainfall anomalies of the LSGB region with a correlation between ENSO and precipitation ranging from 0.28 to 0.46. Similarly, Li et al. (2012) found that the WASM rainfall is negatively correlated with boreal spring NAO.

Figures 8a–d show the correlation between the onset, demise, seasonal length, and seasonal rainfall with the Niño-3.4

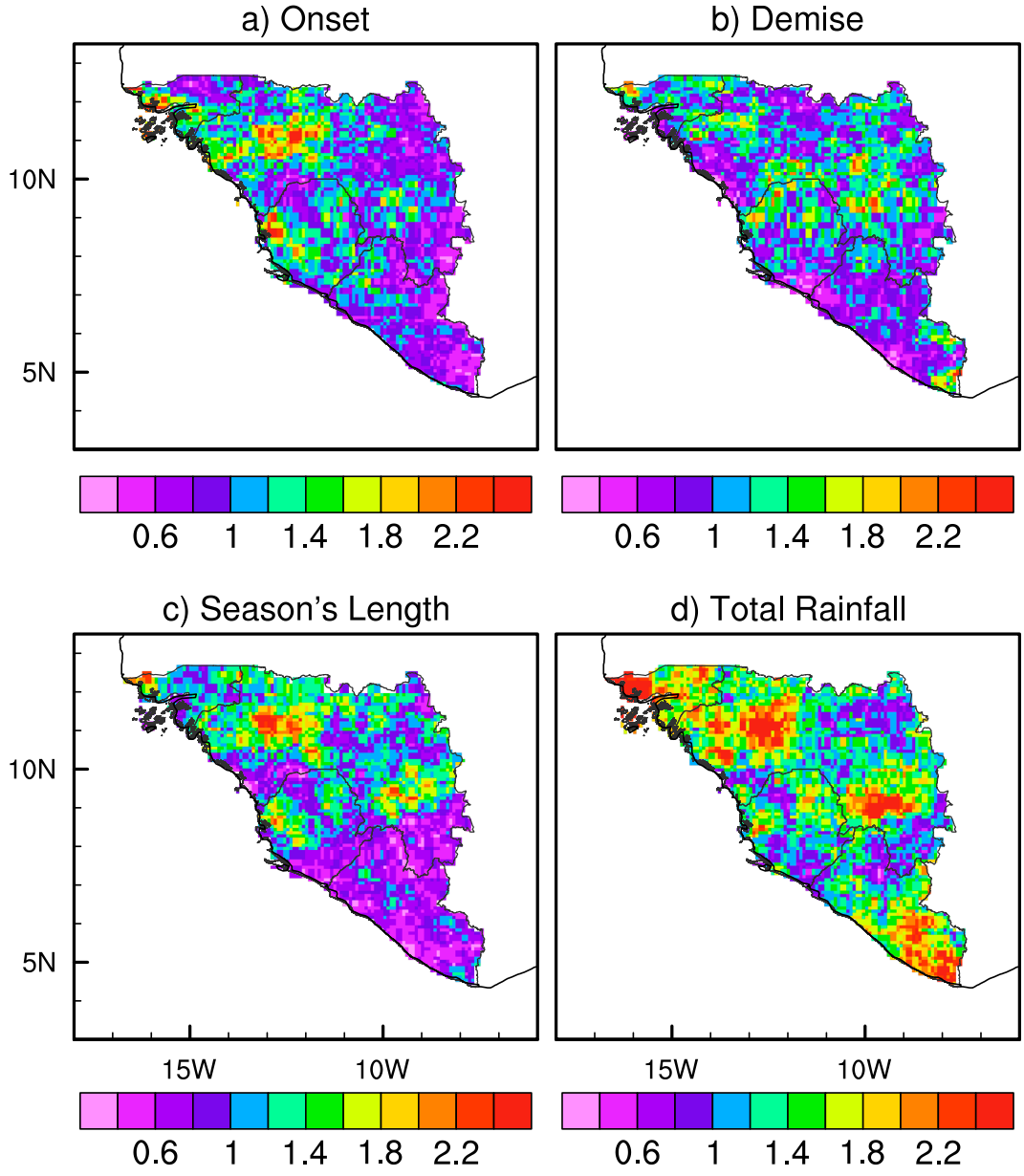

FIG. 5. The 23-yr (2001–23) climatological SNR for (a) ODs, (b) DDs, (c) SLs, and (d) SR from the IMERG dataset.

indices. There is a positive correlation between the onset dates and Niño-3.4 indices during MAM in the majority of the grid points. However, only a few grid points in southern LSGB regions show significant correlations. The demise date shows an insignificant correlation across the region with the Niño-3.4 index during the JJA season. Similarly, the seasonal length and seasonal accumulated rainfall show insignificant correlations in the majority of the grid points. There exists a significant negative correlation between seasonal rainfall and JJA Niño-3.4 indices over southern Liberia and a significant positive correlation over east Sierra Leone. Figures 8e–h show the correlation between the onset, demise, seasonal length, and seasonal rainfall with the NAO indices. The correlation between the onset dates and NAO indices during MAM in the majority of the grid points is insignificant except over northwest Guinea. The demise date, seasonal length, and seasonal rain also show insignificant correlations across the region with the NAO index during the JJA season with the exception of a significant negative correlation over many grid points in east Guinea. Figures 8i–l (Figs. 8m–p) show the correlation between the onset, demise, seasonal length, and seasonal rainfall with the TNA (TSA) indices. The significant correlations with the start date, demise date, seasonal length, and seasonal rainfall are yet again sparse. However, in many grid points over east Guinea, the correlation between the onset date and the TNA indices during MAM shows a significant

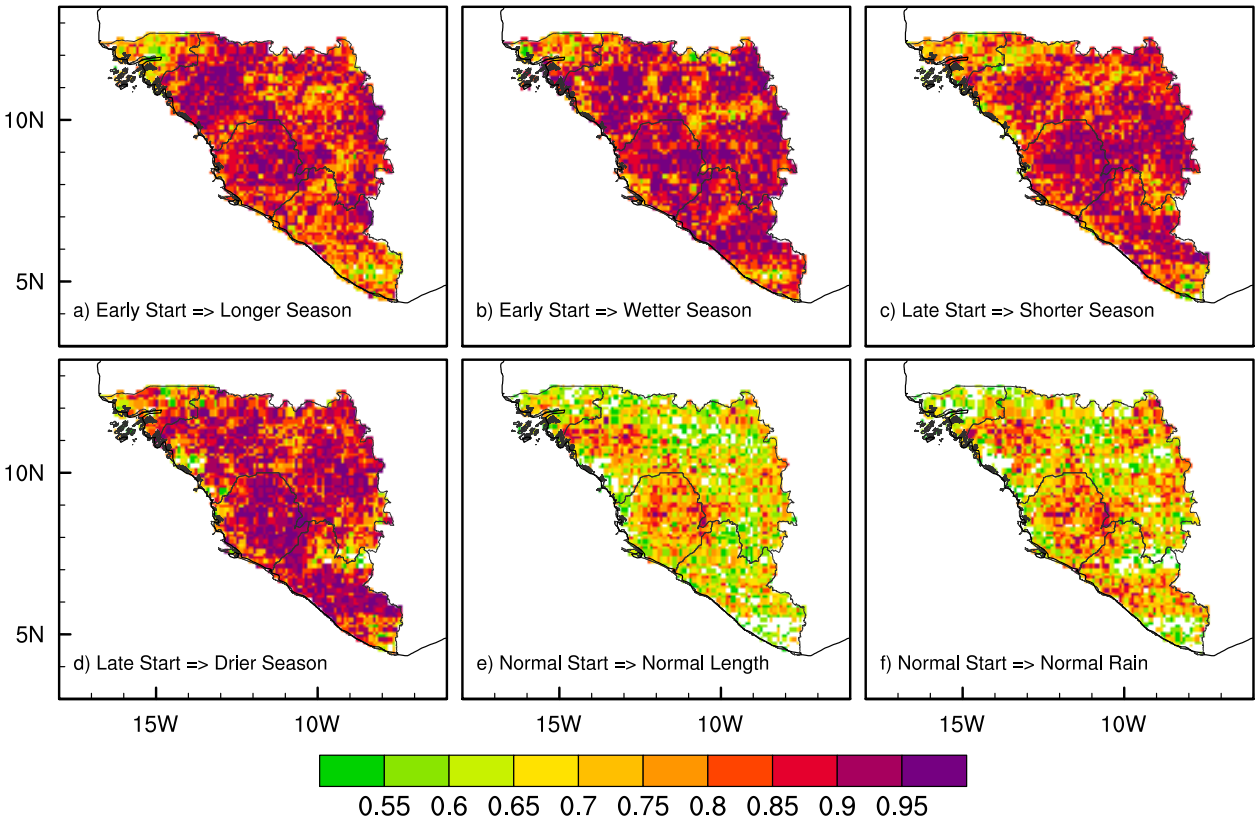

FIG. 6. The AUC for (a),(b) early onset seasons resulting in (a) longer and (b) wetter than normal wet seasons, (c),(d) later onset seasons resulting in (c) shorter and (d) drier than normal wet seasons, and (e),(f) near-normal onset seasons resulting in (e) normal length and (f) normal seasonal rainfall of the rainy season.

negative correlation. The analysis in Fig. 8 clearly shows that variability of the wet season over LSGB is not strongly dictated by variations of the large scale as represented by these indices (ENSO, NAO, TNA, TSA) relative to that shown in Figs. 4a and 4b. This suggests that the seasonal predictability of WASM rainfall is a huge challenge as these indices are not robust predictors for the wet season variation over the LSGB region. These results are also evident from the relatively low signal-to-noise ratio shown earlier (Fig. 5).

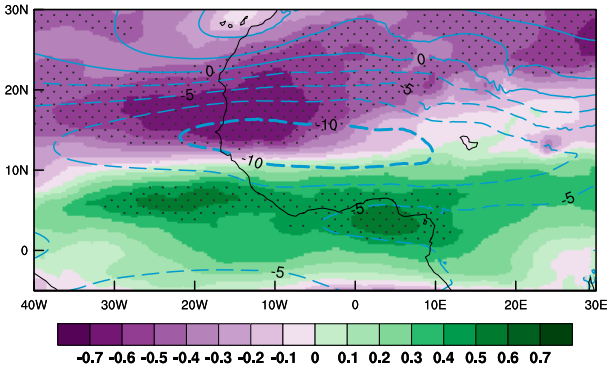

FIG. 7. The correlation of 23-yr ODs (averaged over the LSGB region) with the seasonal mean zonal winds at 600 hPa obtained from ERA5 (seasonal mean is the average zonal wind from OD to DD) is shaded. The contours (sky blue color; dashed is for negative and solid is for positive) are the climatology of seasonal mean zonal wind (m s^{-1}). The significant correlations at a 99% confidence interval according to the *t* test are stippled.

4. Concluding remarks

West Africa stands among the world's most susceptible regions in terms of water resources and food production, owing to demographic and socioeconomic issues compounded by high seasonal climate variability. The variabilities in the WASM rainfall anomalies at the interannual time scale play a decisive role in the agrarian economy of West Africa. We assessed the fidelity of the IMERG intermediate 12-h latency precipitation product with other independent rainfall analysis products and station data. Our analysis yielded a correlation of 0.87 between the IMERG and CHIRPS. CHIRPS is one of the datasets that has been more commonly adopted for the WASM region (e.g., Dunning et al. 2016, 2018; Satgé et al. 2020; Nicholson and Klotter 2021; Palmer et al. 2023). We also compared IMERG with the CPC rainfall dataset, which is primarily based on rain gauge data. The correlation between IMERG and CPC is comparatively lower at 0.51, which is likely due to the inadequate number of gauges over

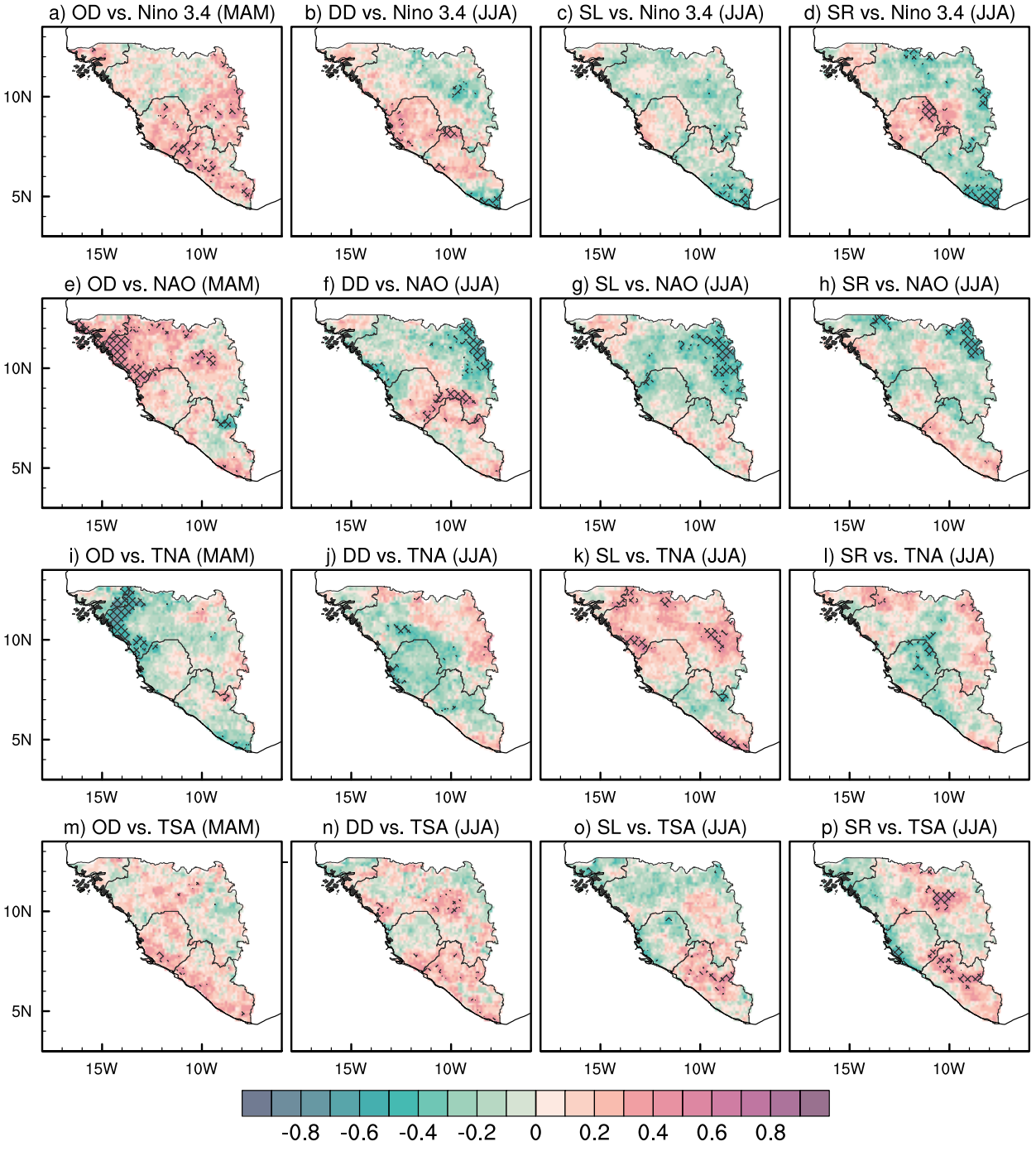

FIG. 8. The 23-yr (2001–23) CC of MAM (a) Niño-3.4 (based on SST from ERSSTv5), (e) NAO, (i) TNA, and (m) TSA with anomalies of the OD. Similarly, the CC of JJA (b) Niño-3.4, (f) NAO, (j) TNA, and (n) TSA with DD; (c) Niño-3.4, (g) NAO, (k) TNA, and (o) TSA with SL; and (d) Niño-3.4, (h) NAO, (l) TNA, and (p) TSA with SR of the rainy season. Stippling indicates the significant correlations at a 95% confidence interval as determined by the *t* test.

the LSGB region (Satgé et al. 2020). From the Q–Q plot, we further highlighted the uncertainty in the gridded rainfall analyses of IMERG, CHIRPS, and CPC over the region. The scarcity of rain gauges over West Africa affects all gridded rainfall analyses over the region. However, verification of the climatological seasonal cycle rainfall obtained from the GSOD station dataset shows that the IMERG is a reliable dataset. This is a critical point because this study is hinged on diagnosing the evolution of the rainy season from its onset to demise.

This study defined the onset and demise dates of the rainy season using an objective method based solely on daily rainfall over the LSGB region, at the granularity of the IMERG rainfall data (i.e., 10-km grid resolution). The analysis reveals that the onset dates exhibit a greater variability across the region compared to the demise dates. The interannual variations in the onset dates of the rainy season are significantly correlated with the length and seasonal accumulated rainfall variations of the rainy season. The generated 101 ensemble members at each grid point by randomly perturbing daily rainfall data provide a robust estimate of start and demise dates, seasonal length, and seasonal accumulated rainfall of the rainy season.

The medians from 101 ensemble members across the LSGB region indicate a strong correlation between the timing of the rainy season's onset and its duration and rainfall anomalies, suggesting that an early or late start date corresponds to lengthier and wetter or shorter and drier conditions, respectively. The estimated climatological mean signal-to-noise ratios for seasonal accumulated rainfall show values greater than one at several grid points throughout the region, indicating the dominance of signal over noise and, thus, potentially implying prospects for better seasonal predictability. Moreover, we employed the area under the receiver operating characteristic (ROC) curve (AUC) and Heidke skill score (HSS) method to evaluate the probabilistic skill of the seasonal outlooks of the seasonal length and seasonal rainfall based on the variations of the onset date of the wet season. Our analysis suggests useful seasonal outlooks could be attained from monitoring of the onset date of the rainy season. Furthermore, anomalous seasons demonstrate higher skill across the LSGB region compared to normal seasons.

We found that the onset date exhibits a significant negative correlation north of the AEJ (15°–25°N) and a significant positive correlation south of the AEJ (2°–8°N), indicating stronger easterlies to the south and weaker easterlies to the north during early onset years, and vice versa during late onset years. In other words, the AEJ shifts southward or northward in early or late onset seasons, respectively. The analysis of interannual variability of the WASM rainfall season with external forcings such as ENSO, NAO, TNA, and TSA reveals that the teleconnections are comparatively weaker than the intrinsic relationships with the onset date variations. Consequently, these large-scale forcings may not serve as reliable potential predictors for variations in the wet season over the LSGB region. Hence, the proposed monitoring of the onset date by utilizing the 12-h latency product of the IMERG precipitation in near real time allows us a feasible alternative to understand and anticipate the evolution of the rainy season at the local scale, which could be usefully adapted for applications in various sectors over the West African region.

Acknowledgments. We acknowledge the funding provided by NASA Grant 80NSSC22K0595. We thank NASA's Global Precipitation Measurement (GPM) program for providing the daily 12-h latency IMERG dataset, which is archived at the NASA GES DISC. We also thank the Climate Hazards Center at the University of California, Santa Barbara (UCSB), for providing the CHIRPS data, and the National Oceanic and Atmospheric Administration (NOAA) Climate Prediction Center (CPC) for their rainfall datasets. We thank NOAA NCEI for the GSOD dataset. We also acknowledge NOAA's Physical Sciences Laboratory (PSL) for providing the Nino-3.4, NAO, TSA, and TAN indices and the fifth major global reanalysis produced by ECMWF (ERA5), for their wind data.

Data availability statement. The IMERG rainfall from NASA was obtained from https://gpm1.gesdisc.eosdis.nasa.gov/data/GPM_L3/GPM_3IMERGDL.07/. The CHIRPS precipitation data are obtained from https://data.chc.ucsb.edu/products/CHIRPS-2.0, and the CPC precipitation data are obtained from https://downloads.psl.noaa.gov/Datasets/cpc_global_precip. The ERA5 wind data are obtained from https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=form. The Nino-3.4 index is from https://psl.noaa.gov/data/correlation/nina34.anom.data, the NAO index is from https://psl.noaa.gov/data/correlation/nao.data, the TNA index is from https://psl.noaa.gov/data/correlation/tna.data, and the TSA index is from https://psl.noaa.gov/data/correlation/tsa.data. The GSOD dataset was from https://www.ncei.noaa.gov/access/search/data-search/global-summary-of-the-day.

REFERENCES

Akinsanola, A. A., and W. Zhou, 2020: Understanding the variability of West African summer monsoon rainfall: Contrasting tropospheric features and monsoon index. *Atmosphere*, **11**, 309, https://doi.org/10.3390/atmos11030309.

Atiah, W. A., L. K. Amekudzi, J. N. A. Aryee, K. Preko, and S. K. Danuor, 2020: Validation of satellite and merged rainfall data over Ghana, West Africa. *Atmosphere*, **11**, 859, https://doi.org/10.3390/atmos11080859.

Becker, T., P. Bechtold, and I. Sandu, 2021: Characteristics of convective precipitation over tropical Africa in storm-resolving global simulations. *Quart. J. Roy. Meteor. Soc.*, **147**, 4388–4407, https://doi.org/10.1002/qj.4185.

Cook, K. H., and E. K. Vizy, 2012: Impact of climate change on mid-twenty-first century growing seasons in Africa. *Climate Dyn.*, **39**, 2937–2955, https://doi.org/10.1007/s00382-012-1324-1.

Dezfuli, A. K., C. M. Ichoku, G. J. Huffman, K. I. Mohr, J. S. Selker, N. van de Giesen, and F. O. Annor, 2017: Validation of IMERG precipitation in Africa. *J. Hydrometeor.*, **18**, 2817–2825, https://doi.org/10.1175/JHM-D-17-0139.1.

Diaconescu, E. P., P. Gachon, J. Scinocca, and R. Laprise, 2015: Evaluation of daily precipitation statistics and monsoon onset/retreat over western Sahel in multiple data sets. *Climate Dyn.*, **45**, 1325–1354, https://doi.org/10.1007/s00382-014-2383-2.

Dos Santos, V., R. A. Jucá Oliveira, P. Datok, S. Sauvage, A. Paris, M. Gosset, and J. M. Sánchez-Pérez, 2022: Evaluating the performance of multiple satellite-based precipitation products in the Congo River Basin using the SWAT model. *J. Hydrol.*, **42**, 101 168, https://doi.org/10.1016/j.ejrh.2022.101168.

Drobinski, P., S. Bastin, S. Janicot, O. Bock, A. Dabas, P. Delville, O. Reitebuch, and B. Sultan, 2009: On the late northward propagation of the West African monsoon in summer 2006 in

the region of Niger/Mali. *J. Geophys. Res.*, **114**, D09108, https://doi.org/10.1029/2008JD011159.

Dunning, C. M., E. C. L. Black, and R. P. Allan, 2016: The onset and cessation of seasonal rainfall over Africa. *J. Geophys. Res. Atmos.*, **121**, 11 405–11 424, https://doi.org/10.1002/2016JD025428.

——, E. Black, and R. P. Allan, 2018: Later wet seasons with more intense rainfall over Africa under future climate change. *J. Climate*, **31**, 9719–9738, https://doi.org/10.1175/JCLI-D-18-0102.1.

Emmanuel, I., 2022: Linkages between El Niño-Southern Oscillation (ENSO) and precipitation in West Africa regions. *Arabian J. Geosci.*, **15**, 675, https://doi.org/10.1007/s12517-022-09942-2.

Fitzpatrick, R. G. J., C. L. Bain, P. Knippertz, J. H. Marsham, and D. J. Parker, 2015: The West African monsoon onset: A concise comparison of definitions. *J. Climate*, **28**, 8673–8694, https://doi.org/10.1175/JCLI-D-15-0265.1.

Flaounas, E., S. Janicot, S. Bastin, R. Roca, and E. Mohino, 2012: The role of the Indian monsoon onset in the West African monsoon onset: Observations and AGCM nudged simulations. *Climate Dyn.*, **38**, 965–983, https://doi.org/10.1007/s00382-011-1045-x.

Fontaine, B., S. Janicot, and P. Roucou, 2008: Definition and predictability of an OLR-based West African monsoon onset. *Int. J. Climatol.*, **28**, 1787–1798, https://doi.org/10.1002/joc.1674.

Funk, C., P. Peterson, M. Landsfeld, D. Pedreros, J. Verdin, S. Shukla, and J. Michaelsen, 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.

Giannini, A., R. Saravanan, and P. Chang, 2005: Dynamics of the boreal summer African monsoon in the NSIPP1 atmospheric model. *Climate Dyn.*, **25**, 517–535, https://doi.org/10.1007/s00382-005-0056-x.

Gu, G., and R. F. Adler, 2004: Seasonal evolution and variability associated with the West African monsoon system. *J. Climate*, **17**, 3364–3377, https://doi.org/10.1175/1520-0442(2004)017<3364:SEAVAW>2.0.CO;2.

Heidke, P., 1926: Berechnung des Erfolges und der Gute der Windstarkevorhersagen im Sturmwarnungsdienst (Measures of success and goodness of wind force forecasts by the gale-warning service). *Geogr. Ann.*, **8**, 301–349.

Hersbach, H., and Coauthors, 2023: ERA5 hourly data on single levels from 1940 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS), accessed 25 January 2023, https://doi.org/10.24381/cds.adbb2d47.

Huffman, G. J., R. F. Adler, D. T. Bolvin, and E. J. Nelkin, 2010: The TRMM Multi-satellite Precipitation Analysis (TMPA). *Satellite Rainfall Applications for Surface Hydrology*, M. Gebremichael and F. Hossain, Eds., Springer, 3–22.

——, E. F. Stocker, D. T. Bolvin, E. J. Nelkin, and J. Tan, 2019: GPM IMERG late precipitation L3 1 day 0.1 degree × 0.1 degree V06. Goddard Earth Sciences Data and Information Services Center (GES DISC), accessed 5 June 2023, https://doi.org/10.5067/GPM/IMERGDL/DAY/06.

Ingram, K. T., M. C. Roncoli, and P. H. Kirshen, 2002: Opportunities and constraints for farmers of West Africa to use seasonal precipitation forecasts with Burkina Faso as a case study. *Agric. Syst.*, **74**, 331–349, https://doi.org/10.1016/S0308-521X(02)00044-6.

Janicot, S., S. Trzaska, and I. Poccard, 2001: Summer Sahel-ENSO teleconnection and decadal time scale SST variations. *Climate Dyn.*, **18**, 303–320, https://doi.org/10.1007/s003820100172.

Jayasankar, C. B., S. Surendran, and K. Rajendran, 2015: Robust signals of future projections of Indian summer monsoon rainfall by IPCC AR5 climate models: Role of seasonal cycle and interannual variability. *Geophys. Res. Lett.*, **42**, 3513–3520, https://doi.org/10.1002/2015GL063659.

Joly, M., and A. Voldoire, 2009: Influence of ENSO on the West African monsoon: Temporal aspects and atmospheric processes. *J. Climate*, **22**, 3193–3210, https://doi.org/10.1175/2008JCLI2450.1.

——, ——, H. Douville, P. Terray, and J.-F. Royer, 2007: African monsoon teleconnections with tropical SSTs: Validation and evolution in a set of IPCC4 simulations. *Climate Dyn.*, **29** (1), 1–20, https://doi.org/10.1007/s00382-006-0215-8.

Li, H., H. Wang, and Y. Yin, 2012: Interdecadal variation of the West African summer monsoon during 1979–2010 and associated variability. *Climate Dyn.*, **39**, 2883–2894, https://doi.org/10.1007/s00382-012-1426-9.

Liebmann, B., and J. Marengo, 2001: Interannual variability of the rainy season and rainfall in the Brazilian Amazon basin. *J. Climate*, **14**, 4308–4318, https://doi.org/10.1175/1520-0442(2001)014<4308:IVOTRS>2.0.CO;2.

——, S. J. Camargo, A. Seth, J. A. Marengo, L. M. V. Carvalho, D. Allured, R. Fu, and C. S. Vera, 2007: Onset and end of the rainy season in South America in observations and the ECHAM 4.5 atmospheric general circulation model. *J. Climate*, **20**, 2037–2050, https://doi.org/10.1175/JCLI4122.1.

Losada, T., B. Rodríguez-Fonseca, S. Janicot, S. Gervois, F. Chauvin, and P. Ruti, 2010: A multi-model approach to the Atlantic Equatorial mode: Impact on the West African monsoon. *Climate Dyn.*, **35**, 29–43, https://doi.org/10.1007/s00382-009-0625-5.

Macharia, D., K. Fankhauser, J. S. Selker, J. C. Neff, and E. A. Thomas, 2022: Validation and intercomparison of satellite-based rainfall products over Africa with TAHMO in situ rainfall observations. *J. Hydrometeor.*, **23**, 1131–1154, https://doi.org/10.1175/JHM-D-21-0161.1.

Maidment, R. I., D. I. F. Grimes, R. P. Allan, H. Greatrex, C. Rojas, and O. Leo, 2013: Evaluation of satellite-based and model re-analysis rainfall estimates for Uganda. *Meteor. Appl.*, **20**, 308–317, https://doi.org/10.1002/met.1283.

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.

Misra, V., 2004: An evaluation of the predictability of austral summer season precipitation over South America. *J. Climate*, **17**, 1161–1175, https://doi.org/10.1175/1520-0442(2004)017<1161:AEOTPO>2.0.CO;2.

——, C. B. Jayasankar, P. Beasley, and A. Bhardwaj, 2022: Operational monitoring of the evolution of rainy season over Florida. *Front. Climate*, **4**, 793959, https://doi.org/10.3389/fclim.2022.793959.

——, S. Dixit, and C. B. Jayasankar, 2023: The regional diagnosis of onset and demise of the rainy season over tropical and subtropical Australia. *Earth Interact.*, **27**, https://doi.org/10.1175/EI-D-22-0026.1.

Murakami, T., L.-X. Chen, A. Xie, and M. L. Shrestha, 1986: Eastward propagation of 30–60 day perturbations as revealed from outgoing longwave radiation data. *J. Atmos. Sci.*, **43**,

961–971, https://doi.org/10.1175/1520-0469(1986)043<0961:EPODPA>2.0.CO;2.

Narotsky, C. D., and V. Misra, 2022: The seasonal predictability of the wet season over Peninsular Florida. *Int. J. Climatol.*, **42**, 3408–3417, https://doi.org/10.1002/joc.7423.

Nguyen, D.-Q., J. Renwick, and J. McGregor, 2014: Variations of monsoon rainfall: A simple unified index. *Geophys. Res. Lett.*, **41**, 575–581, https://doi.org/10.1002/2013GL058155.

Nicholson, S. E., 2013: The West African Sahel: A review of recent studies on the rainfall regime and its interannual variability. *Int. Scholarly Res. Not.*, **2013**, 453521, https://doi.org/10.1155/2013/453521.

——, and D. A. Klotter, 2021: Assessing the reliability of satellite and reanalysis estimates of rainfall in equatorial Africa. *Remote Sens.*, **13**, 3609, https://doi.org/10.3390/rs13183609.

NOAA/NCEI, 1999: Global Surface Summary of the Day – GSOD, version 1.0. NOAA National Centers for Environmental Information, accessed 10 July 2024, https://www.ncei.noaa.gov/access/metadata/landing-page/bin/iso?id=gov.noaa.ncdc:C00516.

Palmer, P. I., and Coauthors, 2023: Drivers and impacts of eastern African rainfall variability. *Nat. Rev. Earth Environ.*, **4**, 254–270, https://doi.org/10.1038/s43017-023-00397-x.

Peyrillé, P., J.-P. Lafore, and J.-L. Redelsperger, 2007: An idealized two-dimensional framework to study the West African monsoon. Part I: Validation and key controlling factors. *J. Atmos. Sci.*, **64**, 2765–2782, https://doi.org/10.1175/JAS3919.1.

Redelsperger, J.-L., A. Diongue, A. Diedhiou, J.-P. Ceron, M. Diop, J.-F. Gueremy, and J.-P. Lafore, 2002: Multi-scale description of a Sahelian synoptic weather system representative of the West African monsoon. *Quart. J. Roy. Meteor. Soc.*, **128**, 1229–1257, https://doi.org/10.1256/003590002320373274.

Rowell, D. P., C. K. Folland, K. Maskell, J. A. Owen, and M. N. Ward, 1992: Modelling the influence of global sea surface temperatures on the variability and predictability of seasonal Sahel rainfall. *Geophys. Res. Lett.*, **19**, 905–908, https://doi.org/10.1029/92GL00939.

Satgé, F., D. Defrance, B. Sultan, M.-P. Bonnet, F. Seyler, N. Rouché, F. Pierron, and J.-E. Paturel, 2020: Evaluation of 23 gridded precipitation datasets across West Africa. *J. Hydrol.*, **581**, 124412, https://doi.org/10.1016/j.jhydrol.2019.124412.

Shukla, J., and Coauthors, 2000: Dynamical seasonal prediction. *Bull. Amer. Meteor. Soc.*, **81**, 2593–2606, https://doi.org/10.1175/1520-0477(2000)081<2593:DSP>2.3.CO;2.

Sivakumar, M. V. K., 1988: Predicting rainy season potential from the onset of rains in southern Sahelian and Sudanian climatic zones of West Africa. *Agric. For. Meteor.*, **42**, 295–305, https://doi.org/10.1016/0168-1923(88)90039-1.

Sultan, B., and S. Janicot, 2003: The West African monsoon dynamics. Part II: The "preonset" and "onset" of the summer monsoon. *J. Climate*, **16**, 3407–3427, https://doi.org/10.1175/1520-0442(2003)016<3407:TWAMDP>2.0.CO;2.

——, C. Baron, M. Dingkuhn, B. Sarr, and S. Janicot, 2005: Agricultural impacts of large-scale variability of the West African monsoon. *Agric. For. Meteor.*, **128**, 93–110, https://doi.org/10.1016/j.agrformet.2004.08.005.

Sylla, M. B., E. Coppola, L. Mariotti, F. Giorgi, P. M. Ruti, A. Dell'Aquila, and X. Bi, 2010: Multiyear simulation of the African climate using a regional climate model (RegCM3) with the high resolution ERA-interim reanalysis. *Climate Dyn.*, **35**, 231–247, https://doi.org/10.1007/s00382-009-0613-9.

Thorncroft, C. D., H. Nguyen, C. Zhang, and P. Peyrillé, 2011: Annual cycle of the West African monsoon: Regional circulations and associated water vapour transport. *Quart. J. Roy. Meteor. Soc.*, **137**, 129–147, https://doi.org/10.1002/qj.728.

Wu, M.-L. C., O. Reale, S. D. Schubert, M. J. Suarez, R. D. Koster, and P. J. Pegion, 2009: African easterly jet: Structure and maintenance. *J. Climate*, **22**, 4459–4480, https://doi.org/10.1175/2009JCLI2584.1.

Xie, P., M. Chen, S. Yang, A. Yatagai, T. Hayasaka, Y. Fukushima, and C. Liu, 2007: A gauge-based analysis of daily precipitation over East Asia. *J. Hydrometeor.*, **8**, 607–626, https://doi.org/10.1175/JHM583.1.

Supplemental Material

Journal of Hydrometeorology
Seasonal Outlook on the West African Rainy Season: Monitoring Onset Date Variations using GPM IMERG Satellite-Based Precipitation Data
https://doi.org/10.1175/JHM-D-24-0067.1

Supplemental Material

Seasonal Outlook on the West African Rainy Season: Monitoring Onset Date Variations using GPM IMERG Satellite-based Precipitation Data

Jayasankar C B[1, 3, #], Vasubandhu Misra[1, 2, 3]

[1]Center for Ocean-Atmospheric Prediction Studies, Florida State University, Tallahassee, Florida, USA.

[2]Department of Earth, Ocean and Atmospheric Science, Florida State University, Tallahassee, Florida, USA.

[3]Florida Climate Institute, Florida State University, Tallahassee, Florida, USA.

#cbjayasankar@gmail.com

Section S1: *Steps to estimate area under the relative operating characteristic curve (AROC) for early onset with longer season*

1. Using 100 perturbed onset dates at a single grid point, calculate the median onset date (Reference) for each 23-year period from 2001 to 2023, then rank the reference onset dates from 1 to 23.
2. Separate the ranks into terciles to determine the reference thresholds for early onset (lower tercile) and late onset (upper tercile).
3. Rank the onset dates (1 to 2323) generated by the perturbation method over all 23 years and 101 ensemble members (100 perturbed + 1 original). Divide the ranked onset dates into terciles and identify the prediction thresholds for early onset (lower tercile) and late onset (upper tercile).
4. For year 2001, estimate the number of ensemble members predict early onset date. Examine the seasonal length, is it longer season (occurs in the upper tercile)?
    - If so, the answer is "Yes" in the contingency table (Table S1). Go to (a).
    - If not, the answer is "No" in the contingency table (Table S1). Go to (b).

    a. Using the count of ensemble members predict early onset date, ask:
    Do 10% (or 10 out of 101) ensemble members predict early onset in 2001? If yes, assign Hit=1. If no, assign Miss=1.

    Do 20% (or 20 out of 101) ensemble members predict early onset in 2001? If yes, assign Hit=1. If no, assign Miss=1.

    Continue like this up to 90% of ensemble members.

    b. Using the count of ensemble members predict early onset date, ask:
    Do 10% (or 10 out of 101) ensemble members predict early onset in 2001? If yes, assign False Alarm=1. If no, assign Correct Rejection=1.

    Do 20% (or 20 out of 101) ensemble members predict early onset in 2001? If yes, assign False Alarm=1. If no, assign Correct Rejection=1.

    Continue like this up to 90% of ensemble members.

Repeat step 4 through all 23 years.

5. For each percentage, sum up the values of Hit, Miss, False Alarm, and Correct Rejection across all years.
6. For each percentage, compute the Hit Rate (HR) and the False Alarm Rate (FAR).

    HR = Hit / (Hit + Miss)
    FAR = False Alarm / (False Alarm + Correct Rejection)

7. Plot the HR against the FAR. This is the Relative Operating Characteristic Curve (ROC). Compute the area under the ROC.

8. Repeat this for other grid points.

Repeat steps 2-8 for other event types such as “early onset with wetter season”, “late onset with drier season”, “late onset with shorter season”, and “normal onset with normal length” and “normal onset with normal rainfall”.

Table 1*: A contingency table is used to create relative operating characteristic curves and measure prediction (perturbation) uncertainty.*

Probability of Event	**Exceeds Threshold**	**Does Not Exceed Threshold**
Event Observed (Yes)	Hit	Miss
Event Not Observed (No)	False Alarm	Correct Rejection

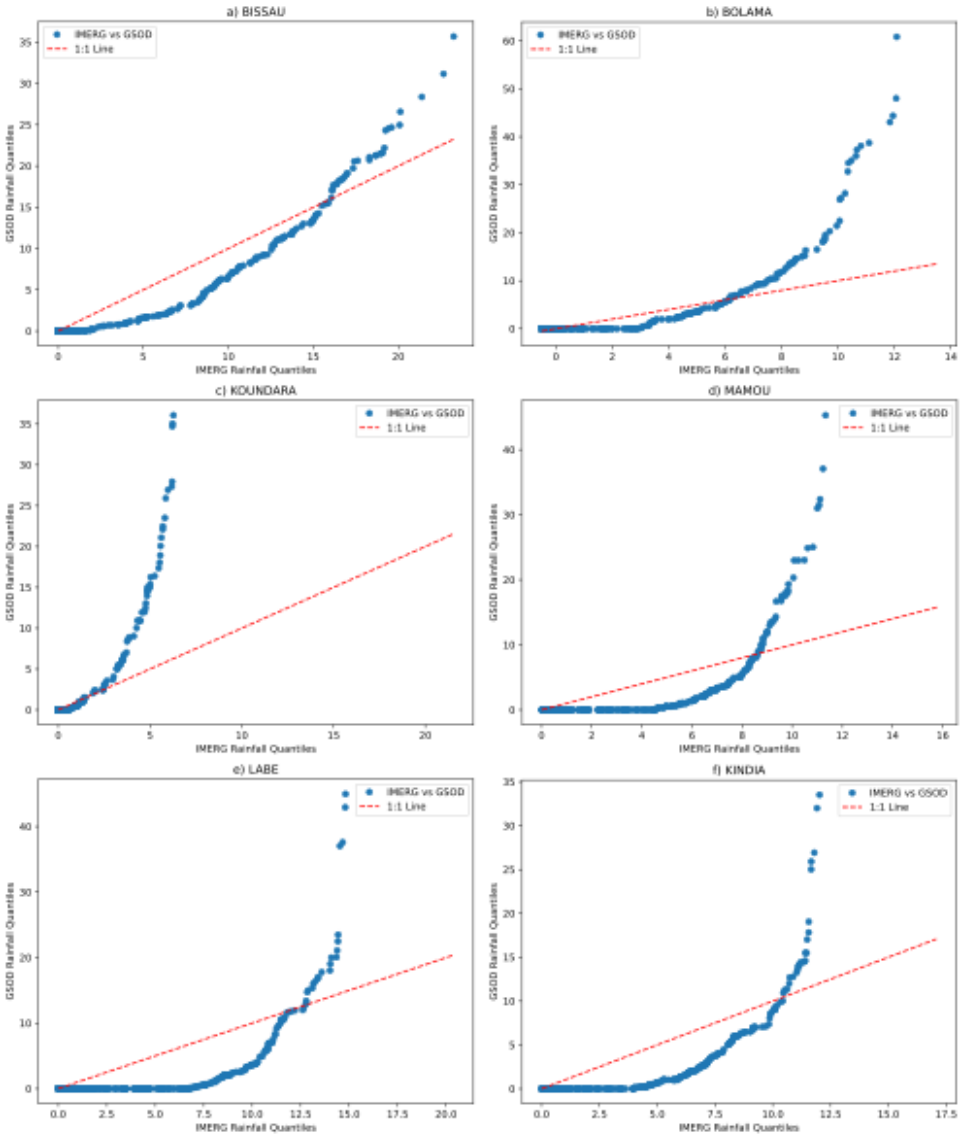

Figure S1*: The Q-Q plot for c) IMERG vs. GSOD for six stations such as a) Bissau, b) Bolama, c) Koundara, d) Mmamou, e) Labe, and f) Kindia.*

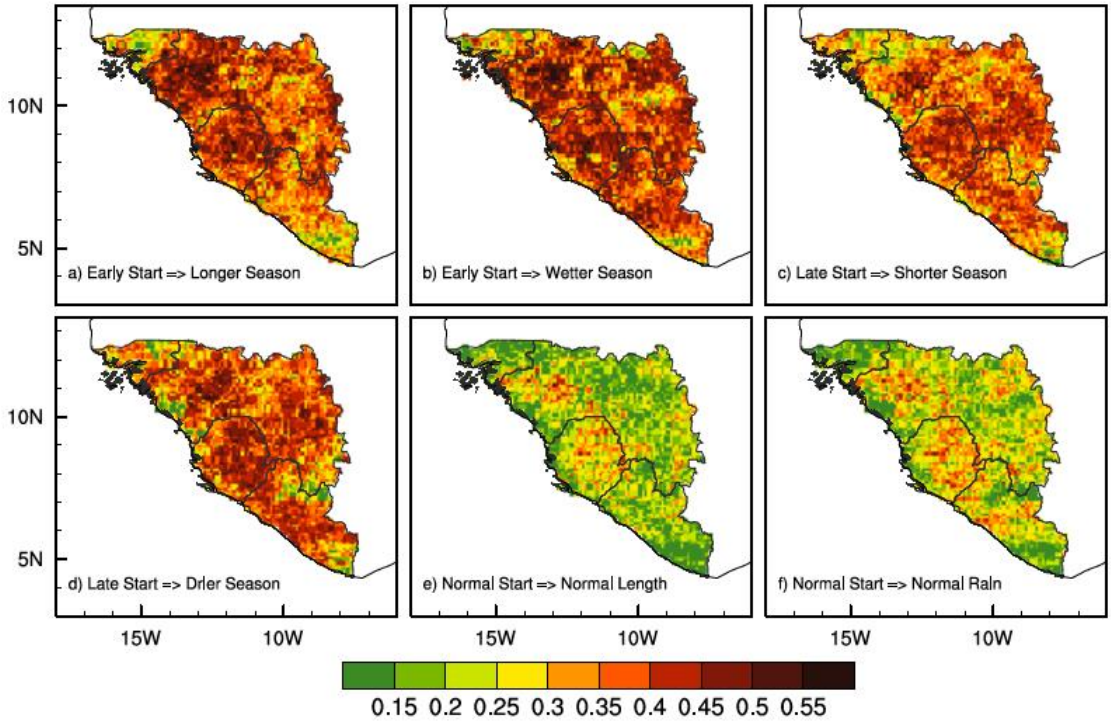

Figure S2*: The Heidke Skill Score (HSS) for (a, b) early onset season resulting in (a) longer and (b) wetter than normal wet season, and (c, d) later onset season resulting (c) shorter and (d) drier than normal wet season, and (e, f) near normal onset season resulting (e) normal length and (f) normal seasonal rainfall of the rainy season.*