Author manuscript, published in "Journal of Climate 22, 5 (2009) 1313–1324"
DOI : 10.1175/2008JCLI2383.1

NOTES AND CORRESPONDENCE

Spatial coherence of monsoon onset over Western and Central Sahel (1950-2000)

Romain MARTEAU [1], Vincent MORON [2, 3, 4, 5] and Nathalie PHILIPPON [1].

[1] UMR 5210 CNRS, Centre de Recherches en Climatologie, Université de Bourgogne, Dijon, France.

[2] UMR 6635, CEREGE, Europôle méditerranéen de l'Arbois, France.

[3] UFR des Sciences Géographiques et de l'Aménagement, Université d'Aix-Marseille I, Aix en Provence, France.

[4] International Research Institute of Climate and Society, Columbia University, New York, USA.

[5] Institut Universitaire de France

Journal of Climate, submitted.

December, 10 2007

Revised version

July, 07 2008

Second revision

September, 02 2008

Corresponding author address: Romain Marteau, Centre de Recherches de Climatologie, UMR 5210 CNRS, Université de Bourgogne, Faculté des Sciences Gabriel, 6 Bd Gabriel, 21000 Dijon, France.

E-mail: romain.marteau@u-bourgogne.fr

Abstract:

The spatial coherence of boreal monsoon onset over Western and Central Sahel (Senegal, Mali, Burkina-Faso) is studied through the analysis of daily rainfall data for 103 stations from 1950 to 2000. Onset date is defined using a local agronomic definition, i.e. the first wet day (> 1 mm) of 1 or 2 consecutive days receiving at least 20 mm without a 7-day dry spell receiving less than 5 mm in the following 20 days. Changing either the length and/or the amplitude of the initial wet spell, or the length of the following dry spell modify the long-term mean local-scale onset date but has only a weak impact either on its interannual variability or its spatial coherence. Onset date exhibits a seasonal progression from southern Burkina-Faso (mid May) to northwestern Senegal and Saharian edges (early August). Interannual variability of the local-scale onset date does not seem to be strongly spatially coherent. The amount of common or covariant signal across the stations is far weaker than the inter-station noise at the interannual time scale. In particular, a systematic spatially-consistent, advance or delay of the onset is hardly observed across the whole Western and Central Sahel. In consequence, the seasonal predictability of local-scale onset over the Western and Central Sahel associated for example with large-scale sea surface temperatures, is, at best, weak.

1. Introduction

The rainy season over the Sahelian belt spans only few months, mainly from July to September (Lamb 1978; Nicholson 1979, 1980), and is associated with the latitudinal shift of the whole monsoon system. Previous work has shown that the northward shift, which determines the large-scale onset of the rainy season across the Sudanian and Sahelian belts, usually occurs at the end of June, as an abrupt jump between Guinean (i.e. 5°-8°N) to Sahelian (i.e. 10°-14°N) latitudes (Sultan and Janicot 2000, 2003; Sultan et al. 2003; Fontaine and Louvet 2006). The monsoon withdrawal in late September over the Sahelian belt is much smoother (Sultan and Janicot 2000). "Pre-onset" scattered rains could occur over the Sahelian belt before the large-scale S-N jump, while most of the Sahelian seasonal rainfall are associated with westward-moving squall-lines (SLs) of mean horizontal extent of ~ 500 km and propagation speed of 10-

15 m/s (i.e. Martin and Schreiner 1981; D'Amato and Lebel 1998; Laurent et al. 1998; Mathon et al. 2002).

The timing of the start of the rainy season is crucial to decide when to plant crops (Sivakumar 1992). This is of particular importance for the Sahelian belt where economy is mostly based on rainfed agriculture, with weak capacity to mitigate adverse effects of interannual variability of rainfall (Ingram et al. 2002). In that context, any reliable prediction of the local-scale onset date would be of a great value to assist on time preparation of farmlands, mobilisation of seed, manpower and equipment, and would also reduce the risks of planting at an unfavorable time (Omotosho et al. 2000). Sultan et al. (2005) clearly demonstrated that the "ideal" sowing date, i.e. the one leading to the highest yields, is close to the large-scale onset date, defined as the northward jump of the Intertropical Convergence Zone (ITCZ).

The definition of the onset of the rainy season over West Africa has generally followed two approaches: (i) a local-scale one and (ii) a large-scale one. Methods for detecting the onset of the rainy season at local-scale have been proposed as early as the 70's for specific Sahelian settings (in Niger mainly, Davey et al. 1976; Benoit 1977). Most methods rely on *in-situ* daily measurements of rainfall meeting subjective thresholds (Davey et al. 1976; Stern et al. 1981; Sivakumar 1988; Jolliffe and Sarria-Dodd 1994; Balme et al. 2005). These empirical thresholds are chosen according to the climatological properties of the rainy season as well as agronomic criteria (these points will be detailed in section 3). In this approach, the onset is primarily assumed as a change in the properties, frequency and/or intensity of rainfall events associated with the arrival of the monsoon air mass. Other parameters have also been taken into consideration to detect local onsets, such as evapotranspiration (Benoit 1977; Ati et al. 2002).

A more recent large-scale approach is based on sub-continental indexes of observed rainfall or rainfall proxies such as outgoing long wave radiation (OLR) (Sultan and Janicot 2000). The seasonal cycles of pentad rainfall averaged over West African longitudes (e.g. 10°W-10°E) show an alternation of steep increases and pauses (Louvet et al. 2003) during its northward progression. One of these pulses appears to be associated with the abrupt northward shift of the ITCZ, the so-called "ITCZ jump" (Sultan and Janicot 2000). The onset is thus only defined from the planetary-scale northward shift of the monsoon, and a single date is given for a given band of longitudes.

The approach of Camberlin and Diop (2003) can be viewed as an intermediate approach. In their study, an Empirical Orthogonal Function (EOF) analysis is applied to local-scale, daily rainfall anomalies in Senegal for a multi-year period. A single onset date was then computed for the whole country as the minimum of the cumulative sum of the leading principal component (PC) time series on a yearly basis. This date is representative of the local-scale onset dates for the stations highly correlated with the leading PC, and part of local-scale interstation noise is filtered out through the spatial weighting provided by the EOF.

This brief review demonstrates that at least two different strategies are used to define the onset. The local-scale definition is clearly the most suitable for end-users so that they can use the information as a valuable decision-making parameter. In that context, three important related questions are explored in this paper:

1. How pertinent is the local-scale definition, proposed in previous studies for specific settings across the Sahel, for a larger longitudinal belt? The previous definitions are based on subjective thresholds chosen according to climatological and/or agronomic features. The sensitivity of the onset date (in terms of mean and interannual variability) to this parametrization has never been evaluated.
2. To what extent do the average and local-scale onset date match with the regional-scale (i.e. as defined by Sultan and Janicot, 2000) ones? Is the abrupt jump of the ITCZ noticeable across the network? These issues are related to the shape and intensity of potential modes of variability of onset date. They are also intrinsically related to the intensity of the spatial coherence and potential predictability of onset date.
3. The spatial coherence of a phenomenon is partly indicative of its potential predictability: It is indeed expected that global or regional-scale sea surface temperatures anomalies should induce a rather smooth, roughly spatially-uniform, anomalous signal at regional-scale, at least for a flat area such as the Sahel (Moron et al. 2006, 2007). As forecasts of the rainy season onset is desired by user communities, an estimate of its potential predicatbility is a first step toward this goal.

These three issues are analyzed from a 103 rain-gauge network covering Western and Central Sahel (Senegal, Mali and Burkina-Faso) from 1950 to 2000. Section 2 describes the different daily rainfall datasets. Section 3 provides the main results in terms of (i) definition of the local-scale onset date, (ii) climatological mean and characteristics of the rainfall field before and after the onset, (iii) interannual variability and spatial coherence of onset date. Conclusions are given in section 4.

2. Daily rainfall data

The daily rainfall analysed in this study come from two different databases: (i) the FRIEND-AOC (Flow Regimes from International Experimental and Network Data - Afrique de l'Ouest et Centrale) group (http://armspark.msem.univ-montp2.fr/FRIENDAOC/), and (ii) the Senegalese DMN (Direction de la Météorologie Nationale) that provided us with *in situ* daily data of 103 stations for the period 1950 to 2000 (Fig. 1).

a. Rainfall data

Data for 33 stations in Mali and 41 stations in Burkina-Faso were compiled by the FRIEND-AOC group from various sources (IRD – Institut de Recherche pour le Développement –, ASECNA – Agence pour la SECurité de la Navigation Aérienne en Afrique –, DMN – Direction Météorologique Nationale – and CIEH – Comité Inter-Africain d'Etudes Hydrauliques –).

Daily rainfall amounts recorded at 29 stations come from the Senegal DMN. Most of the stations are located in the western part of the country while the eastern part is less sampled (Fig. 1).

b. Filling the missing values

Missing daily rainfall observations (less than 3%) were replaced with a local scaling (Widmann et al. 2003; Ines and Hansen 2005; Schmidli et al. 2006) of daily ERA-40 rainfall. For each station, the closest ERA-40 grid point is chosen, and its daily rainfall is scaled so that the frequency of occurrence of wet day > 1 mm and the mean intensity of daily rainfall during wet days match the long-term monthly mean of the available data. Missing entries are then replaced by the calibrated ERA-40 rainfall for these particular days. Note that replacing the missing entries from May to October with a simple stochastic weather generator (Wilks 1999) leads to very similar results.

3. Results

a. Agronomic definition of onset

The local-scale agronomic definition of the onset is the most relevant to agriculture management. This definition is based on local-scale daily rainfall using empirical thresholds of rainfall that ensure enough soil moisture during planting and growing periods to avoid crop failure (Walter 1967; Omotosho 1990 1992; Omotosho et al. 2000). Table 1 lists the different thresholds used in selected previous studies dedicated to the Sahelian domain. Most of the studies agree on two criteria; (i) 1 or 2 consecutive days receiving at least 20-25 mm of rainfall, and (ii) no 7-day dry spell in the following 30 days from the onset. The threshold of 20 mm amount was adopted following a study by ICRISAT dedicated to millet in Niger (Davey et al. 1976; Sivakumar 1992) and corresponds to the minimum water requirement for crop survival. Considering at least 2 consecutive days is a reasonable choice to take into account a rain event that occurs across the recording time (usually between 7 and 9 a.m local time). In fact, SLs develop themselves mainly in late afternoon and persist sometimes during the night (Martin and Schreiner 1981). The 7-day dry spell after the onset is known as the "control period" and prevents against "false" onsets that have disastrous consequences on germination and crop development (Sivakumar 1992), and require farmers to sow again while the seed’s stock is usually small (Diop 1996). Moreover, anything less than 50 percent of the weekly crop water requirement (CWR) will likely lead to crop failure (i.e. Omotosho et al. 2000). The CWR equals at least 10 mm across the Sahel, leading to a threshold of 5 mm in 7 days to define a dangerous limit for crop survival after the first rains.

In the current study, a wet (dry) day is defined as a day receiving more (less) than 1 mm (Diop 1999) because (i) rainfall amounts between 0 and 1 mm are not equally reported across countries and stations (i.e. synoptic or not), and (ii) considering daily amounts less than 1 mm usually increases the interstation noise (Moron et al. 2007) when determining onset. The onset date (OD hereafter) is defined here as the first wet day of 1 or 2 consecutive days receiving at least 20 mm without any 7-day dry spell receiving less than 5 mm during the following 20 days counted from the onset. OD is computed from May, 1st. These criteria are obviously too restrictive for the driest stations and years. While less than 0.4% of station-year do not experience any 2-day receiving at least 20 mm during May-November season, 8.6% of onset are undefined when the post-onset 7-day dry spell receiving less than 5 mm in the following 20 days is included in the definition. This percentage reaches 19% when the control period length is extended to 30 days from the initial wet spell. A window of 20-day is thus retained as a compromise between the need to consider "false" starts and a too restrictive threshold that excludes many stations and years. Undefined onset dates still account for 40% of years for the 14 stations north of 15°N. In most of these cases, no rainy season is actually experienced with scattered rainy events spread across the season. In the following analyses (except for EOF in section 3d), the undefined onset dates (8.6%) are replaced by the latest available onset date observed across the network for the given year, weakly increasing the spatial coherence over the northern edge of the domain.

Figure 2a displays the mean onset dates averaged over the 51 years. The mean OD exhibits a rather regular northward shift , mixed with a secondary northwestward one over Senegal, from mid May (south Burkina Faso) to early August (NW Senegal and central-northern Mali). Respectively, 17.9%, 31.4%, 36.1% and 12.7% of onset occur in May, June, July and August (Fig. 2b).

b. Mean characteristics of rainfall field before and after the onset

Figure 3 shows the mean characteristics of the rainfall field averaged over the 15 days before and 15 days after the local-scale OD. These periods are extracted for each station on a yearly basis, and the frequencies of occurrence of wet days receiving respectively between 1 and 20 mm and > 20 mm are computed (Fig. 3a-d) as well as the mean length of the wet and dry spells (Fig. 3e-h).

Before the onset, the frequency of occurrence of wet days receiving between 1 and 20 mm is far from zero (usually 10-20%). By contrast, the frequency of occurrence of wet days receiving > 20 mm is very low (1-4%), but this is almost fully explained by the criteria used to define OD (see section 3a). The mean dry spell length is rather homogeneous, usually between 4 and 6 days, without any clear northward increase (Fig. 3g). Similarly, the mean wet-spell length is rather spatially uniform, between 1 and 2 days, still without any clear northward decrease (Fig. 3e).

From the onset, the frequency of occurrence of wet day receiving between 1 and 20 mm increases smoothly, by a factor of 2, while the frequency of occurrence of wet day > 20 mm is roughly multiplied by 5. The multiplication of wet days breaks up the dry spells, that are consistently shorter than before the onset (less than 3 days except for several northern stations, Fig. 3g,h) while the mean wet spell length remains shorter than 2 days (Fig. 3i,j). In other terms, at local-scale, the onset is usually associated with more and wetter rainy days afterwards, but those wet days remain rather isolated or clustered into 2-day wet sequences. This could be related to the dominant influence of SLs and the usual scarcity of long-lasting wet spells at these latitudes.

Figure 4 further illustrates the relative rainfall field near the onset date, in particular its spatial pattern. For each station and each year, the frequency of occurrence of daily rainfall > 1 mm for 2 days before (Fig. 4a), and after (Fig. 4b) the local-scale onset (and their difference, Fig. 4c) is computed for the whole rainfall field relocated relative to the reference station (i.e. (0,0) location on Fig. 4). In other words, Figures 4b,c illustrate the spatial coherence of the "first kind", i.e. the frequency of occurrence of a wet day around a station when its onset occurs. Figure 4a is roughly similar to the mean climatological pattern of frequency of occurrence, i.e. a northward gradient, also visible in the mean onset date (Fig. 2). Considering the long-term mean of the frequency of occurrence of any other sample of 2 days leads roughly to the same pattern, except that the spatial mean varies according to the seasonal cycle of rainfall (not shown). The occurrence of local-scale OD slightly alters this pattern (Fig. 4b), mostly through an increase of the frequency of occurrence of an elongated WSW-ENE shape around the reference station (Fig. 4c). A small asymmetric surface of 3°-4° (~ 1°) in longitude by 2° (~ 1°) in latitude around the reference station exhibits significant (at the two-sided 90% level according to a Student's T test) increase of the frequency of occurrence for 75% (90%) of the time between pre- and post-onset 2 days (Fig. 4c). Thus, figure 4 suggests that local-scale OD seems hardly related to coherent space-time propagating features at meso- or synoptic-scale and/or even the planetary-scale because neither regular northward (possibly associated with an abrupt jump of the ITCZ), nor westward (possibly associated with a SL) progression appears between panel (a) and (b) of Figure 4. Increasing the length of the time period averaged before and after the onset to 5, 10, 20 and 30 days is associated with a wider ~ W-E belt of increased frequency of occurrence of rainfall, probably in relation with the mean northward shift of the overall rain belt, especially for periods lasting at least 10 days (not shown).

c. Spatial coherence of the onset date at interannual time scale

The spatial coherence of the "second kind" is the possible systematic modulation of onset dates (or any other seasonal characteristic) at interannual time scale. Note that a weak spatial coherence of the "first kind" (section 3b) does not necessarily forbid this effect because a large-scale forcing could synchronize the onset dates in time, by systematically delaying or bringing forward the onset of the rainy season at local-scale (Moron et al. 2008). The spatial coherence of the OD at interannual time scale has been estimated through degrees of freedom (DOF) (Moron 1994; Fraedrich et al. 1995; Bretherton et al. 1999; Moron et al. 2006) and interannual variance of the standardized anomaly index (SAI) (var[SAI]) (Katz and Glantz 1986). The DOF are computed following the equation of Fraedrich et al. (1995):

$$DOF = \frac{M^2}{\sum_{i=1}^{M} e_i^2}$$

where M is the number of stations and e, the eigenvalues of the correlation matrix. The DOF varies between 1 and the rank of the correlation matrix. In the limiting case of $e_i = 1$ for all orders, DOF = *M*, that is, each station conveys independent information and the common "signal" is zero. On the other hand, if the first eigenvalue accounts for all variance of the field, then DOF = 1, that is, each station conveys the same information equal to the signal and the noise is zero. Note that due to finite length, the highest DOF expected for a 51 x 103 matrix of independent white noise is not 50, but rather ~ 34. The var[SAI] ranges between 0 (correlation of -1 between two equal-sized and perfectly out-of-phase samples), 1/M (~ 0.01 in our case, for spatially independent variations), and 1 (perfect correlation between stations) (Moron et al. 2006, 2007).

Here the DOF of the local-scale OD is high (23.7) indicating a weak spatial coherence at the regional scale. The var[SAI] equals 0.07 and is also close to the value associated with a white noise time series. For comparison, the DOF and var[SAI] equal 6.1 (respectively 4.8) and 0.36 (respectively 0.41) for the May-November seasonal rainfall amounts (respectively frequency of occurrence of wet days > 1 mm). Therefore, according to DOF and var[SAI], the agronomic OD is characterized by a considerable amount of inter-station noise, and in consequence, is assumed to be poorly predictable from large-scale predictors at the interannual time scale (Moron et al. 2006, 2007). To test the sensitivity of the spatial coherence to parametrization, DOF and var[SAI] have been computed again, by changing (i) the length of initial wet spell (L_w = 2 to 10 days) and (ii) the amount of rainfall received during the initial wet spell (A_w = 1, 5, 10, 15, 20, 25, 30, 35 and 40 mm) while the other criterion (i.e. a 7-day dry spell receiving less than 5 mm in the following 20 days) is kept fixed. Note that the percentage of undefined onset dates is always < 10% ~~in mean~~, except when $A_w > 25$ mm and $L_w < 4$ days (maximum of undefined onset = 19.1% for $A_w = 40$ and $L_w = 2$). As before, the missing entries are replaced by the latest interannual onset date, slightly increasing the spatial coherence on the northern edges.

DOF and var[SAI] are indeed very stable amongst the 81 different estimates; DOF varies between ~23 and ~28 with 88% of values between 23 and 25 and var[SAI] varies between 0.065 and 0.086. In other words, increasing the length and/or amplitude of the initial wet spell has a very weak impact on the spatial coherence. Even the climatological long-term mean and the interannual variability at local-scale are only weakly sensitive to the parametrization of A_w and L_w.. 52% (respectively 76%) of the local-scale mean bias between two different estimates of OD are less than 5 days (respectively 10 days) in absolute value and 77% (respectively 88%) of the local-scale correlations between two different estimates of OD are equal or greater than 0.7 (respectively 0.6). The impact of the length of the control dry spell (L_d) was also tested with A_w and L_w set to 20 mm and 2 days respectively, while L_d was set equal to 5, 7, 9, 11, 13, and 15 days in the 20 days from the onset. Note that undefined onsets decrease almost exponentially from 25% to 1% between the extreme values of L_d. DOF and var[SAI] are almost stable between 25 and 26 and < 0.10 respectively. In summary, the weak spatial coherence of "second kind" is not related to the parametrization of the OD.

d. Interannual to interdecadal variability of the onset date

To further examine the interannual variability of OD, an Empirical Orthogonal Function (EOF) analysis based on its correlation matrix is computed. Note that undefined onsets are left untouched here and that correlation matrix is computed only from pairs of available ODs. The leading EOF mode of OD accounts for 8.7% of the total variance consistently with the weak spatial coherence depicted above. By contrast, the leading EOF of the May-September seasonal amount (the frequency of occurrence of wet days > 1 mm) accounts for 38.1% (respectively 43.4%) of the total variance. The second unrotated EOF mode of onset date explains less than 6.4% of total variance. The first unrotated mode (Fig. 5a) describes a relatively coherent structure of OD variability mostly over Western and Central Senegal. This behaviour could be at least partly related to the relatively high spatial sampling (Fig. 1), but could be also related to a slightly stronger signal (Camberlin and Diop 2003). The corresponding leading EOFs of seasonal amount (Fig. 5b) and frequency of occurrence of wet days > 1 mm (Fig. 5c) exhibit larger scale patterns with loadings > 0.6 spread all across the domain without a clear spatial gradient. The leading PC of onset date (respectively seasonal amount and frequency of occurrence) (Fig. 5d) exhibits a positive (negative) and statistically significant (at one-sided 95% confidence level) trend, i.e. a delayed onset (decreased seasonal rainfall and frequency of occurrence) from 1950-2000. The median of local-scale correlations between onset on one hand and seasonal amount and frequency of occurrence on the other hand equals only -0.35.

In order to analyse the relationships between the local-scale and the regional-scale onset date as defined by Sultan and Janicot (2000) through the S-N jump of the ITCZ the local-scale OD are extracted from 1968 (available online at http://www.lodyc.jussieu.fr/~bslod/monsoon.htm). Note that DOF of local-scale OD for the 1968-2000 period equals ~19, but this apparent higher spatial coherence could be an artefact of the decrease of the matrix rank from 50 to 32. The correlations between the large-scale and the 103 local-scale ODs are between -0.54 and 0.41, with only 7 values being significant at the two-sided 95% level according to a random phase test (Janicot et al. 1996). Moreover, the correlation between regional-scale OD as defined by Sultan and Janicot (2000) and a regional index (SAI) or the leading PC of onset date equals only 0.06 and 0.24 (both values are not significant at the two-sided 90% level). The large-scale onset is divided in 3 terciles of early (1969, 1982, 1985, 1973, 1978, 1994, 1996, 1999, 1993, 1974, 1977), near-normal (1990, 1992, 1975, 1980, 1970, 1983, 1986, 1968, 1976, 1971, 1979) and late (1981, 1972, 2000, 1985, 1988, 1997, 1984, 1989, 1987, 1998, 1991) jumps of the ITCZ. The local-scale OD anomalies are then averaged for the three sets of years. There is a considerable amount of inter-station noise for each of the terciles and the number of stations where the OD anomaly is significantly different from zero (based on a one-sided t-test at the 90% confidence level) is always less than 15% (not shown).

Lastly, the teleconnections between local scale OD and four sea surface temperature (SST) indices, capturing ocean sectors known for exerting an influence over West African Monsoon were considered (i.e. Nicholson 1986; Ropelewski and Halpert 1987; Rowell et al. 1995;

Nicholson and Kim 1997; Janicot et al. 1996, 2001). These indices include the east equatorial Pacific (NINO3.4, 170°W-120°W; 5°S-5°N), the North Tropical Atlantic (NATL: 60°W-30°W; 5°N-20°N), the South Tropical Atlantic (SATL: 30°W-10°E; 0°-20°S) and the Guinean Gulf (GGUI: 10°W-10°E; 5°S-5°N). Figure 6 displays distribution of correlation coefficients between local-scale OD and each SST indices averaged over May-July (i.e. including ~85% of the local-scale onset, Fig. 1) for 1950-2000, 1950-1969 and 1970-2000. At least 50% of correlation coefficients are between -0.2 and +0.2 and the significant value are usually scattered in space (not shown). In summary, the teleconnections between the SST related to the ENSO phenomenon, the Tropical Atlantic dipole, or Guinea Gulf and local-scale OD seem weak.

4. Conclusion

The main goal of this paper was to analyze the spatial coherence, robustness and the interannual variability of the local-scale onset date (OD) of the boreal summer monsoon over Western and Central Sahel (Mali, Burkina-Faso, Senegal) from a network of 103 rain gauges (Fig. 1) with almost complete records from 1950 to 2000. The local-scale OD is first defined as the first wet day ~~> 1 mm~~ of one or two consecutive day receiving at least 20 mm without any 7-day dry spell receiving less than 5 mm in the following 20 days from the onset. This local-scale definition follows previous studies (Stern et al. 1981; Sivakumar 1988; Omotosho 1990; Jolliffe and Sarria-Dodd 1994; Omotosho et al. 2000; Dodd and Jolliffe 2001) and is best suited for end-user purposes since it explicitly takes into account the rainfall demand for crop seeding and survival during the first stages of growth. This analysis complements previous similar studies (summarized in Table 1) which typically considered a smaller subset of stations and did not deeply investigate the sensitivity of the mean OD and its variability related to the subjective parameters used in its definition such as the length and amplitude of the initial wet spell.

The spatial coherence of OD is analyzed from two points of view: (i) a quasi-instantaneous pattern of rainfall when onset date occurs at a particular station (Fig. 4) and (ii) the synchronization of onset dates at interannual time scale, i.e. the possible systematic modulation of onset dates across the Sahelian band (Fig. 5). The spatial coherence of the "first kind" exhibits a small, significant increase of frequency of occurrence of rainfall relative to pre-onset 2-day (Fig. 4c). This could be associated with a range of factors, from convective cells to meso-convective clusters that probably trigger the local-scale onset. The spatial coherence of the "second kind" is analyzed through empirical estimates of the degrees of freedom (DOF) or interannual variance of the Standardized Anomaly Index (var[SAI]). Both estimates suggest a weak spatial coherence, i.e. the onset date is hardly systematically synchronized at interannual time scales. This is also demonstrated by the weak amount of variance accounted for by the leading EOF (Fig. 5a). These results contrast with the seasonal amount and the frequency of occurrence which exhibit a large-scale pattern across the whole belt analyzed (Fig. 5b,c). The weak spatial coherence of the "second kind" is almost independent of the parametrizations used in the definition of the onset date.

These suggest that large-scale potential forcing seems unable to systematically synchronize the occurrence of the first rains above a fixed threshold across the western and central Sahel. The occurrence of the local-scale onset could be hardly viewed as a systematic event across year and station, but merely a complex product of the multi-scale phenomena that produce rainfall over this area. The discrepancy between the local-scale OD defined here and the regional-scale onset defined by the northward jump of the monsoon is especially puzzling and needs further investigation. It is possible that local-scale onset as defined here could be associated either with a localized convective event, even before the S-N jump of the ITCZ, or with a meso-scale squall line. But even in this latter case, the timing of occurrence, size, location and track of the first squall line of the rainy season are hardly reproducible from one year to the next. In other words, there are too many different phenomena, or different scales of motion, that could lead to 5-40 mm rainy event in 2-10 days at local-scale. If the northward jump of the monsoon enhances the probability of such events across the Sahel, the fact that ITCZ has moved to its northernmost seasonal location, does not necessarily induce local-scale onset everywhere at the same moment or even on a short time step.

An important consequence is that local-scale OD appears to show little potential predictability based on large and regional-scale boundary conditions such as sea surface temperatures and/or soil moisture (Folland et al. 1986; Philippon and Fontaine 2002; Douville et al. 2007). This is because the regional-scale seasonal potential predictability implies that the variable of interest shares a variable, but significant, amount of common or covariant information at this scale, at least for a flat area as Sahel. The fact that seasonal rainfall amount and frequency of occurrence are far more spatially-coherent than the onset date (i.e. DOF equals respectively 6.1, 4.8, and 23-28) suggests that most of the seasonal predictability of rainfall over the Sahelian belt is actually not associated with a systematic delay or advance of the local-scale onset as defined here, but stands out for example as a seasonally-varying or constant systematic modulation of the frequency of occurrence of rainfall across the season (Moron et al. 2007). This is not a trivial result, because local-scale onset sometimes conveys a spatially-consistent signal and potential predictability, as in Indonesia (Moron et al. 2008). This analysis emphasizes also the need to carefully examine the context of any local-scale analysis of the onset, because the weak spatial coherence implies a potential large uncertainty due to random sampling. More work is also needed to look at other definitions of the onset, for example integrating the daily rainfall in time and/or in space. The spatial noise is not necessarily evenly distributed across the scales and it would be interesting to look at intermediate scales between the local-scale and large-scale onsets. In particular, any definition that filters out some of the local-scale noise as the cumulative daily rainfall anomalies (Camberlin and Diop 2003; Liebmann et al. 2007) is especially appealing in this context. The apparent spatial increase of the signal over Western and Central Senegal (Fig. 5a) warrants also further studies.

Acknowledgements: We acknowledge the FRIEND-AOC group and Senegal DNM for providing us the rainfall data. R. Marteau is partly supported by the AMMA program. Based on a French initiative, AMMA was built by an international scientific group and is currently funded by a large number of agencies, especially from France, UK, US and Africa. It has been the beneficiary of a large financial contribution from the European Community's sixth Framework Research Program. Detailed information on scientific coordination and funding is available on the AMMA International web site: http://www.amma-international.org. ECMWF ERA-40 data used in this study have been freely obtained from the ECMWF Data Server (http://www.ecmwf.int). We thank Gil Mahé and Benjamin Sultan for fruitful discussions. Lastly, we thank two anonymous reviewers and the editor of this paper (D. Straus) for their constructive comments which help us on clarifying the goals and results of our paper, and lastly Bradfield Lyon (IRI, Columbia University) for his careful reading of the revised version of this paper

References

Ati, O.F., Stigter, C.J., and Oladipo, E.O., 2002: A comparison of methods to determine the onset of the growing season in Northern Nigeria. *Int. J. Climatol.*, **22,** 731-742.

Balme, M., Galle, S. and Lebel, T., 2005: Démarrage de la saison des pluies au Sahel : variabilité aux échelles hydrologique et agronomique, analysée à partir des données EPSAT-Niger. *Sécheresse*, **16**, 15-22.

Benoit, P., 1977: The start of the growing season in Northern Nigeria. *Agric. For. Meteorol.*, **18,** 91-99.

Bretherton, C.S., Widmann, M., Dymnikov, V.P., Wallace, J.M., and Bladé, I., 1999: The effective number of degrees of freedom of a time varying field. *J. Clim.*, **12,** 1990-2009.

Camberlin, P. and Diop, M., 2003: Application of daily rainfall principal component analysis to the assessment of rainy season characteristics in Senegal. *Clim. Res.*, **23,** 159-169.

D'Amato, N. and Lebel, T., 1998: on the characteristics of the rainfall events in the Sahel with a view to the analysis of climatic variability. *Int. J. Climatol.*, **18,** 955-974.

Davey, E.G., Mattei, F., and Soloman, S.I., 1976: An evaluation of climate and water resources for the development of agriculture in the Soudano-Saelian zone of West Africa. *WMO Special Environment Report No. 9.*

Diop, M., 1996: A propos de la durée de la saison des pluies au Sénégal. *Sécheresse*, 7**,** 7-15.

Diop, M., 1999: Caractérisation du facteur hydrique en agriculture pluviale au Sénégal: le cas de l'arachide et du mil. *Thèse de doctorat***,** Université de Bourgogne, Dijon**,** 185 p.

Dodd, D.O.S. and Jolliffe, I.T., 2001: Early detection of the start of the wet season in semiarid tropical climates of western Africa. *Int. J. Climatol.*, **21,** 1251-1262.

Douville, H., Conil, S., Tyteca, A., and Voldoire, A., 2007: Soil moisture memory and West African monsoon predictability: artefact or reality? *Clim. Dyn.***,** DOI 10.1007/s00382-006-0207-8.

Folland, C.K., Palmer, T.N., and Parker, D.E., 1986: Sahel rainfall and worldwide sea temperatures. *Nature*, **320,** 602-607.

Fontaine, B. and Louvet, S., 2006: Sudan-Sahel rainfall onset : definition of an objective index, types of years, and experimental hindcast. *J. Geophys. Res.*, **111,** doi:10.1029/2005JD007019.

Fraedrich, K., Ziehmann, C., and Sielmann, F., 1995: Estimates of spatial degrees of freedom.

J. Clim., **8,** 361-369.

Ines, A.V.M. and Hansen, J.W., 2005: Bias correction of daily GCM outputs for crop simulation studies. *Agric. For. Meteorol.*, **138,** 44-53.

Ingram, K.T., Roncoli, M.C., and Kirshen, P.H., 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.

Janicot, S., Moron, V., and Fontaine, B., 1996: Sahel drought and ENSO dynamics. *Geophys. Res. Lett.*, **23,** 515-518.

Janicot, S., Trzaska, S., and Poccard, I., 2001: Summer Sahel-ENSO teleconnection and decadal time scale SST variations. *Clim. Dyn.*, **18,** 303-320.

Jolliffe, I.T. and Sarria-Dodd, D.E., 1994: Early detection of the start of the wet season in tropical climates. *Int. J. Climatol.*, **14,** 71-76.

Katz, R.W. and Glantz, M.H., 1986: Anatomy of a rainfall index. *Mon. Wea. Rev.*, **114,** 764-771.

Lamb, J.P., 1978: Case studies of tropical Atlantic surface circulation patterns associated with subsaharan weather anomalies: 1967 and 1968. *Mon. Wea. Rev.*, **106,** 1052-1055.

Laurent, H., D'Amato, N., and Lebel, T., 1998: How important is the contribution of the Mesoscale Convective Complexes to the Sahelian rainfall? *Phys. Chem. Earth*, **23,** 629-633.

Liebman, B., Camargo, S., Seth, A., Marengo, J.A., Carvalho, L.M.V., Allured, D., Fu, R., and Vera, C.S., 2007: Onset and end of the rainy season in South America in observations and the ECHAM 4.5 Atmospheric General Circulation Model. *J. Clim.*, **20,** 2037-2050.

Louvet, S., Fontaine, B., and Roucou, P., 2003: Active phases and pauses during the installation of the West African monsoon through 5-day CMAP rainfall data (1979-2001). *Geophys. Res. Lett.*, **30,** doi:10.1029/2003GL018058.

Martin, D.W., and Schreiner, A.J., 1981: Characteristics of West African and east Atlantic cloud clusters: a survey from GATE. *Mon. Wea. Rev.*, **109,** 1671-1688.

Mathon, V., Laurent, H., and Lebel, T., 2002: Mesoscale convective system rainfall in the Sahel. *J. Appl. Meteo.*, **41,** 1081-1092.

Moron, V., 1994: Guinean and Sahelien rainfall anomaly indices at annual and monthly scales (1933-1990). *Int. J. Climatol.*, **14,** 325-341.

Moron, V., Robertson, A.W., and Ward, M.N., 2006: Seasonal predictability and spatial coherence of rainfall characteristics in the tropical setting of Senegal. *Mon.Wea. Rev.*, **134,** 3248-3262.

Moron, V., Robertson, A.W., Ward, M.N., and Camberlin, P., 2007: Spatial coherence of tropical rainfall at the regional scale. *J. Clim.*, **20,** 5244-5263.

Moron, V., Robertson, A.W., and Boer, R., 2008: Spatial coherence and seasonal predictability of monsoon onset over Indonesia. *J. Clim.***,** in press.

Nicholson, S., 1979: Revised rainfall series for the West-African subtropics. *Mon. Wea.*

Rev., **107,** 620-623.

Nicholson, S., 1980: The nature of rainfall fluctuations in sub-tropical Africa. *Mon. Wea. Rev.*, **108,** 473-484.

Nicholson, S., 1986: The quasi-periodic behavior of rainfall variability in Africa and its relationship to the Southern Oscillation. *Arch. Meteo. Geophys. Bio.*, **A 34,** 311-348.

Nicholson, S. and Kim, J., 1997: The relationship of the El Niño-Southern Oscillation to African rainfall. *Int. J. Climatol.*, **17,** 177-135.

Omotosho, J.B., 1990: Onset of thunderstorms and precipitation over northern Nigeria. *Int. J. Climatol.*, **10,** 849-860.

Omotosho, J.B., 1992: Long-range prediction of the onset and end of the rainy season in the West African Sahel. *Int. J. Climatol.*, **12,** 369-382.

Omotosho, J.B., Balogun, A.A., and Ogunjobi, K., 2000: Predicting monthly and seasonal rainfall, onset and cessation of the rainy season in West Africa using only surface data. *Int. J. Climatol.*, **20,** 865-880.

Philippon, N., and Fontaine, B., 2002: The relationship between the Sahelian and previous 2nd Guinean rainy seasons: a monsoon regulation by soil wetness ? *Ann. Geophysicae*, **20,** 575-582.

Ropelewski, C.F. and Halpert, M.S., 1987: Global and regional scale precipitations associated with El Niño/Southern Oscillation. *Mon.Wea. Rev.*, **115,** 985-996.

Rowell, D.P., Folland, C.K., Maskell, K., and Ward, M.N., 1995: Variability of summer rainfall over tropical North Africa (1906-1992): Observations and modelling. *Quart. J. Meteo. Soc.*, **113,** 669-674.

Schmidli, J., Frei, C., and Vidale, P.L., 2006: Downscalling from GCM precipitation : a benchmark for dynamical and statistical downscalling methods. *Int. J. Climatol.*, **26,** 679-689.

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. Meteo.*, **42,** 295-305.

Sivakumar, M.V.K., 1992: Empirical analysis of dry spells for agricultural applications in West Africa. *J. Clim.*, **5,** 532-539.

Stern, R.D., Dennett, M.D., and Garbutt, D.J., 1981: The start of the rains in West Africa. *Int. J. Climatol.*, **1,** 59-68.

Sultan, B., and Janicot, S., 2000: Abrupt shift of the ITCZ over West Africa and intra-seasonal variability. *Geophys. Res. Lett.*, **27,** 3353-3356.

Sultan, B., Janicot, S., and Diedhiou, A., 2003: The West African monsoon dynamics. Part I : Documentation of intraseasonal variability. *J. Clim.*, **16,** 3389-3406.

Sultan, B., and Janicot, S., 2003: The West African monsoon dynamics. Part II : The "pre-onset" and the "onset" of the summer monsoon. *J. Clim.*, **16,** 3407-3427.

Sultan, B., Baron, C., Dingkuhn, M., Sarr, B., and Janicot, S., 2005: Agricultural impacts of the large-scale variability of the West African monsoon. *Agric. For. Meteo.*, **128,** 93-110.

Walter, M.W., 1967: Length of rainy season in Nigeria. *Niger. Geogr. J.*, **10,** 123-128.

Widmann, M., Bretherton, C.S., and Salaté, P., 2003: Statistical precipitation downscaling over the Northwestern United States using numerically simulated precipitation as a predictor. *J. Clim.*, **16,** 799-816.

Wilks, D.S., 1999: Interannual variability and extreme-value characteristics of several stochastic daily precipitation models. *Agric. For. Meteorol.*, **93,** 153-169.

Figures captions

Figure 1: Location of the 103 Sahelian rain-gauge used in this study.

Figure 2: (a) Climatological (1950-2000) local-scale onset date (b) distribution of cumulative frequency of the local onset date (25, 50 and 75 percentiles are indicated by vertical dashed line)

Figure 3: Rainfall field statistics averaged over the 15 days before (left column) and from (right column) the local-scale onset date. Mean frequency of occurrence (x 100) of wet days receiving between 1 and 20 mm (a, c) and more than 20 mm (b, d). Mean wet spell length (e, f) and dry spell length (g, h) in tenth of day.

Figure 4: Mean frequency of occurrence of daily rainfall > 1 mm for 2-day before (a) and from (b) the local-scale onset, and the "(b) minus (a)" difference (c). All fields are relocated relatively to the reference station (coordinates = 0,0). In the latter panel, the simple (double) circles indicate a relative location where 75% (90%) of the differences (b)-(a) are significant at the two-sided 90% level according to a Student's T test.

Figure 5: (a) Leading unrotated empirical orthogonal function (EOF) modes of the local-scale onset date, expressed as loadings, i.e. correlation between the corresponding Principal Component (PC) and the raw data. (b) Leading unrotated EOF of seasonal (MJJAS) amount of rainfall. (c) Leading unrotated EOF of seasonal (MJJAS) frequency of occurrence of rainfall > 1 mm. Bold squares displaying correlations statistically significant at the two-sided 95% confidence level according to a Bravais-Pearson test. (d) leading principal component

timeseries of onset date (bars), seasonal amount (circle) and frequency of occurrence of rainfall (upper triangle).

Figure 6: Boxplot of correlation coefficients between local-scale onset date and NINO3.4, Tropical Atlantic dipole and Guinea gulf SST index for periods 1950-2000, 1950-1969 and 1970-2000. ~~Boxes are bounded by the first quartile, and third quartile of the distribution and the internal horizontal line is the median. Whiskers extend from the box out to the most extreme data value within 1.5 by the interquartile range.~~ The boxes have lines at the lower quartile, median, and upper quartile values. The whiskers are lines extending from each end of the boxes to show the extent of the range of the data within 1.5 by the interquartile range from the upper and lower quartiles. ~~The outliers are displayed by a cross~~. Outliers, displayed by a cross, are data with values beyond the ends of the whiskers. Italic values indicate the number of local-scale correlations significant at two-sided 95% confidence level according to a random phase test (Janicot et al. 1996).

Tables captions

Table 1. Criteria of the agronomic definitions of the local-scale onset date of the rainy season used in selected previous studies. All of these definitions start from daily rainfall observed at rain-gauge stations.

Authors	Data	Details of definitions	Study area	Number of stations	Study period
Stern et al. (1981)	Daily rainfall	- 2 successive days - receiving at least 20 mm	Transect S/N Benin, Nigeria, Niger, Mali	11 stations	1934-1965
Sivakumar (1988)	Daily rainfall	- 3 consecutive days - accumulating at least 20 mm - no dry period of 7 or more consecutive days in the following 30 days	Niger Burkina-Faso	58 stations	25 Years (not specified)
Omotosho (1990, 1992)	Daily rainfall	- first 4 falls receiving at least 10 mm - no dry spell of 7-day between any 2 such rains	Northern Nigeria & West Africa	3 stations	1973-1988
Dodd and Jolliffe (1994)	Daily rainfall	- 5 consecutive days (with at least two other wet days in the period) - accumulating at least 25 mm - no dry period of 7 or more consecutive days in the following 30 days	Burkina-Faso	22 stations	1902-1989

Table 1. Criteria of the agronomic definitions of the local-scale onset date of the rainy season used in selected previous studies. All of these definitions start from daily rainfall observed at rain-gauge stations.

<u>FIGURES:</u>

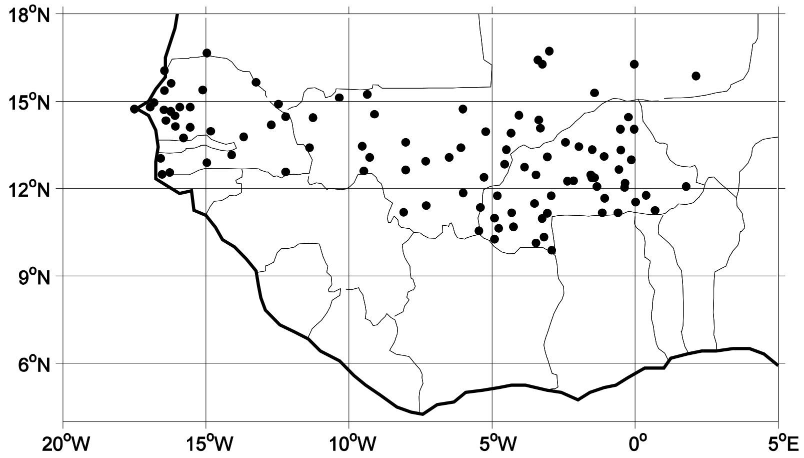


Figure 1. Location of the 103 Sahelian rain-gauge stations used in this study.

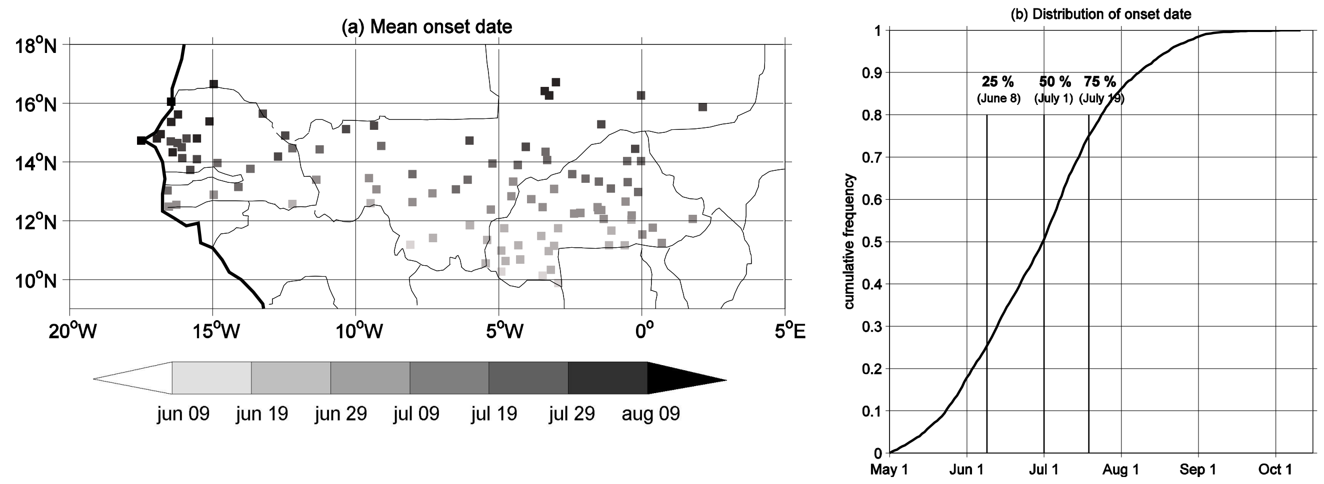
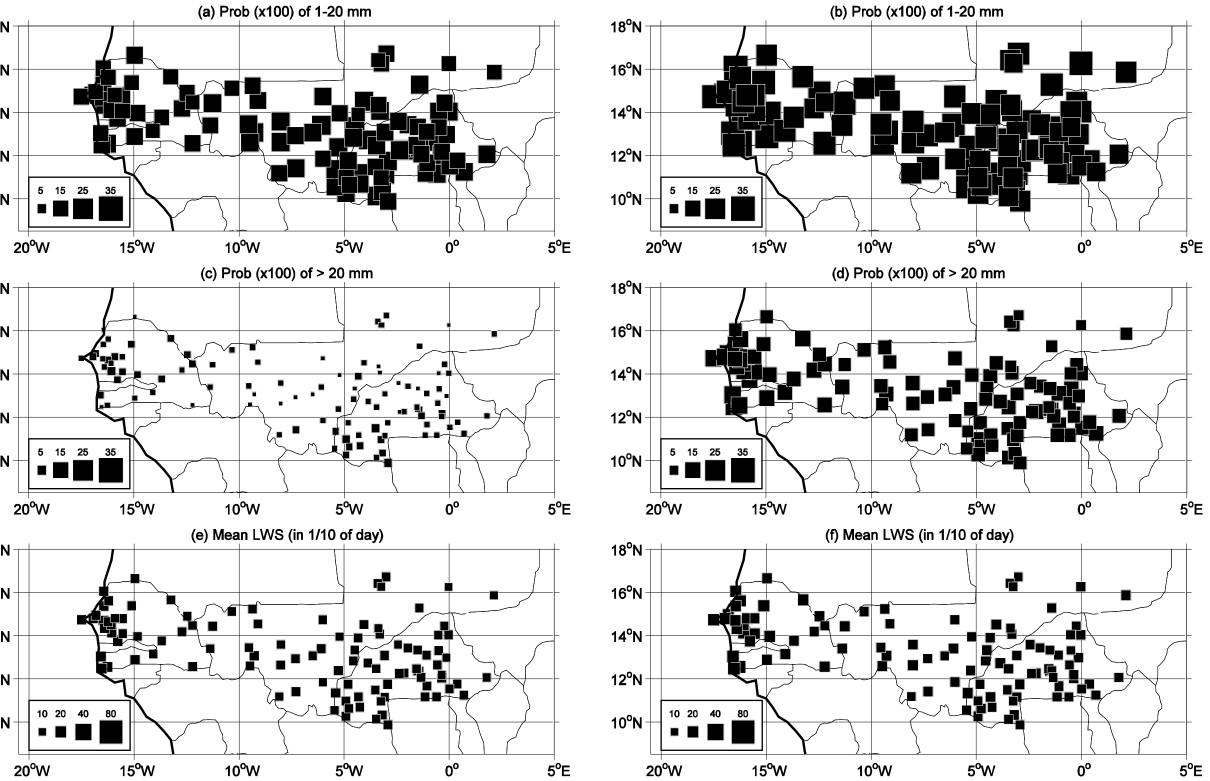


Figure 2 : (a) Climatological (1950-2000) local-scale onset date (b) distribution of cumulative frequency of the local onset date (25, 50 and 75 percentiles are indicated by vertical dashed line).



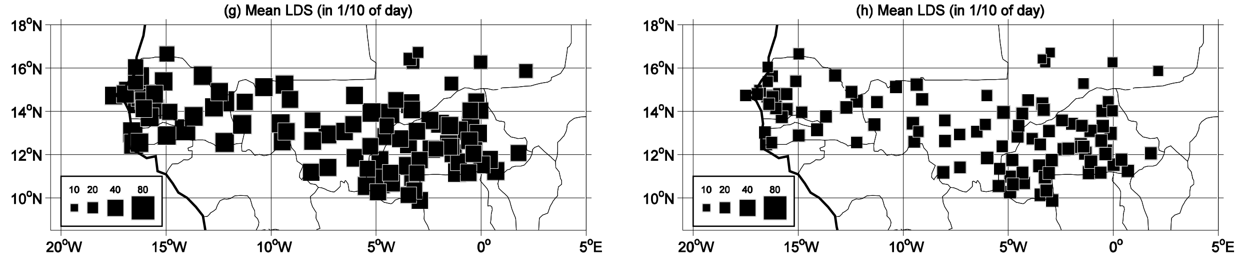


Figure 3. Rainfall field statistics averaged over the 15 days before (left column) and from (right column) the local-scale onset date. Mean frequency of occurrence (× 100) of wet days receiving between 1 and 20 mm (a, c) and > 20 mm (b, d). Mean wet spell length (e, f) and dry spell length (g, h) in tenth of day.

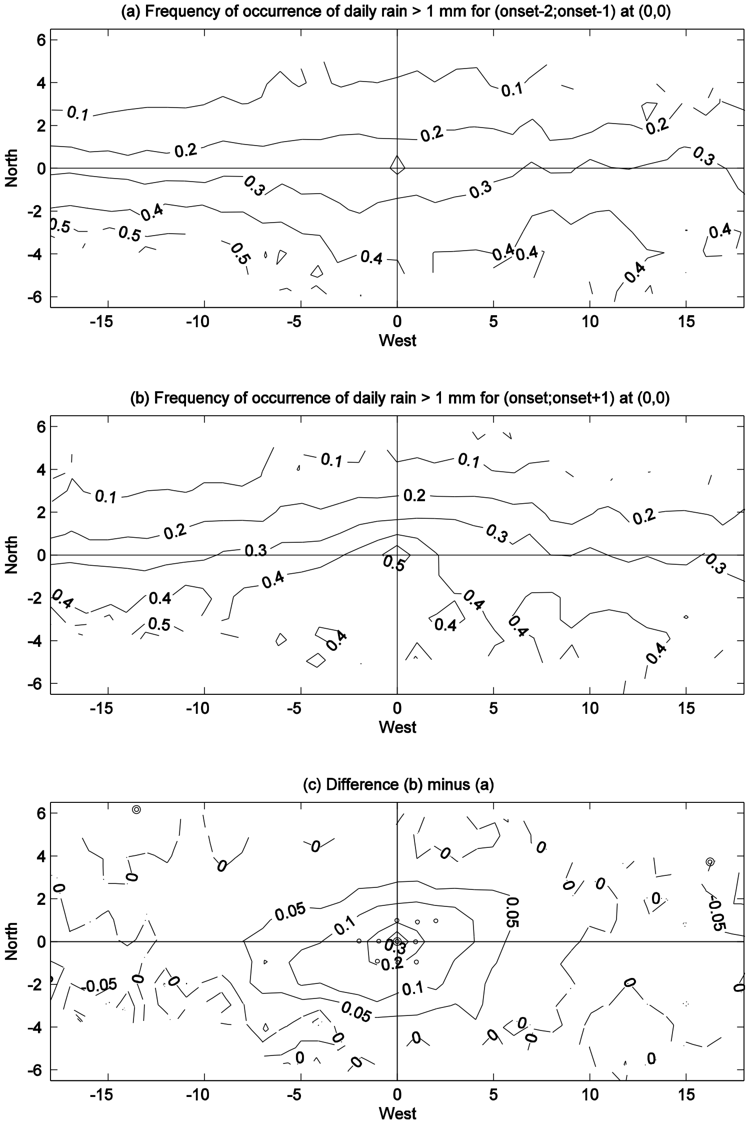
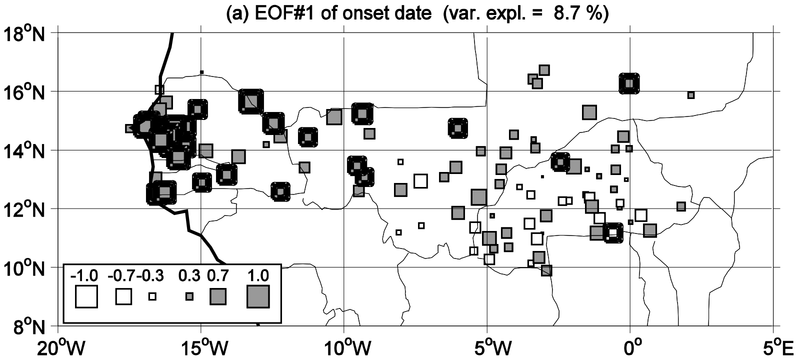
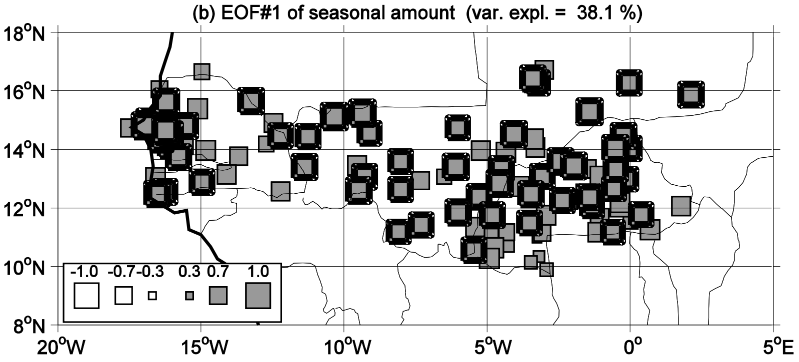


Figure 4. Mean frequency of occurrence of daily rainfall > 1 mm for 2-day before (a) and from (b) the local-scale onset, and the "(b) minus (a)" difference (c). All fields are relocated relatively to the base station (coordinates = 0,0). In the latter panel, the simple (double) circles indicate a relative location where 75% (90%) of the difference (b)-(a) are significant at the two-sided 90% level according to a Student's T test.

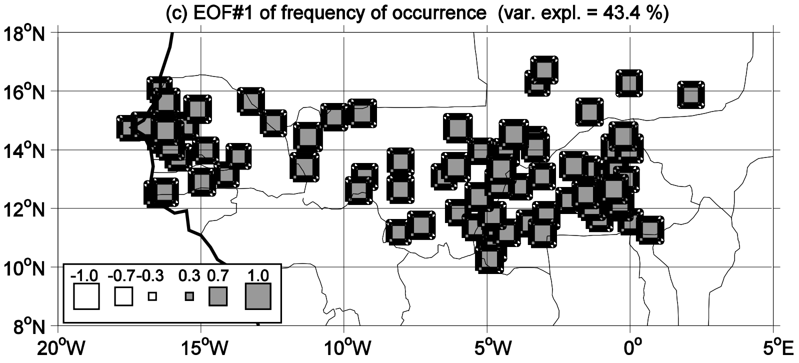
(a) EOF#1 of onset date (var. expl. = 8.7 %)



(b) EOF#1 of seasonal amount (var. expl. = 38.1 %)



(c) EOF#1 of frequency of occurrence (var. expl. = 43.4 %)



(d) Leading principal component

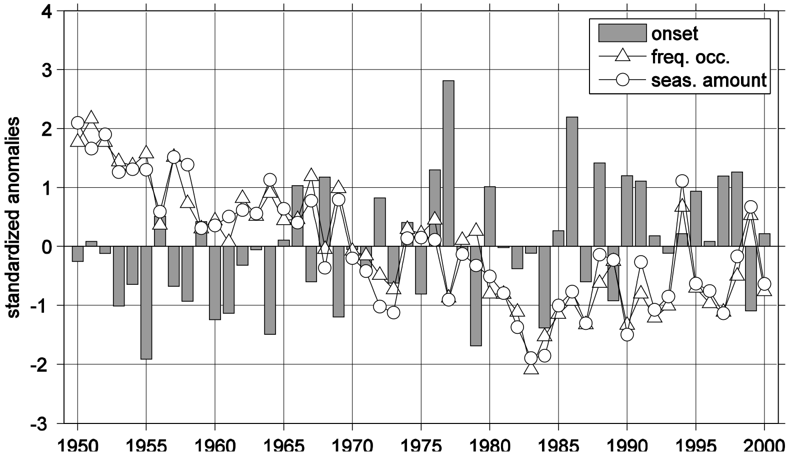


Figure 5. (a) Leading unrotated empirical orthogonal function (EOF) modes of the local-scale onset date, expressed as loadings, i.e. Pearson correlation between the corresponding Principal Component (PC) and the raw data. (b) Leading unrotated EOF of seasonal amount of rainfall. (c) Leading unrotated EOF of frequency of occurrence of rainfall > 1 mm. (d) Leading principal component timeseries of onset date (bars), seasonal amount (circle) and frequency of

occurrence of rainfall (upper triangle). Bold squares displaying correlations statistically significant at the two-sided 95% confidence level according to Bravais-Pearson test.

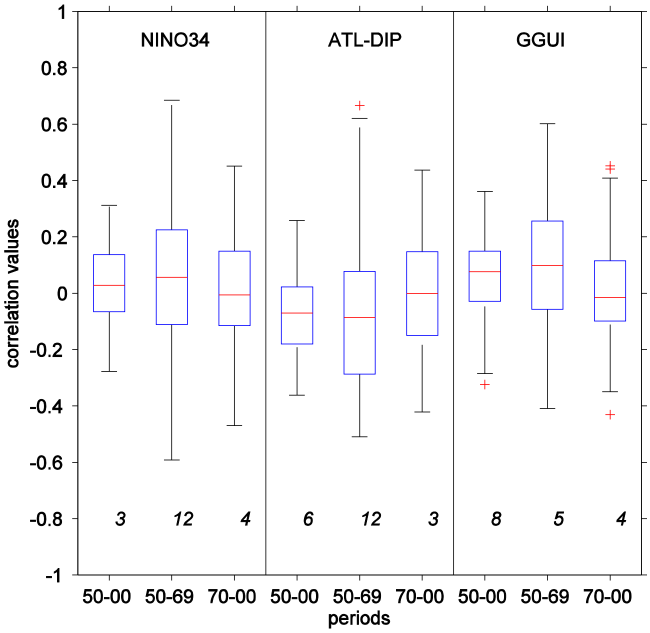


Figure 6 : Boxplot of correlation coefficients between local-scale onset date and NINO3.4, Tropical Atlantic dipole and Guinea gulf SST index for periods 1950-2000, 1950-1969 and 1970-2000. ~~Boxes are bounded by the first quartile, and third quartile of the distribution and the internal horizontal line is the median. Whiskers extend from the box out to the most extreme data value within 1.5 by the interquartile range~~. The boxes have lines at the lower quartile, median, and upper quartile values. The whiskers are lines extending from each end of the boxes to show the extent of the range of the data. ~~The outliers are displayed by a cross.~~ Outliers, displayed by a cross, are data with values beyond the ends of the whiskers. Italic values indicate the number of local-scale correlations significant at two-sided 95% confidence level according to a random phase test (Janicot et al., 1996).