HAL
open science

Definition and predictability of an OLR based West African monsoon onset

Bernard Fontaine, Samuel Louvet, Pascal Roucou

To cite this version:

Bernard Fontaine, Samuel Louvet, Pascal Roucou. Definition and predictability of an OLR based West African monsoon onset. International Journal of Climatology, 2008, 28 (13), pp.1097-1798. 10.1002/joc.1674 . hal-00320782

HAL Id: hal-00320782

https://hal.science/hal-00320782

Submitted on 12 Jan 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Definition and predictability of an OLR based West African monsoon onset

by Bernard Fontaine, Samuel Louvet and Pascal Roucou (Centre de Recherches de Climatologie, CNRS / University of Burgundy, Dijon, France)

Abstract

The monsoon onset is documented in terms of latitudinal shift of deep convection areas within the ITCZ using an interpolated version of the NOAA OLR at a 5-day time-step over West Africa for the period 1979-2004. Signals in moist convection derived from OLR values lower than 180 W/m2 allow better determination of onset dates (ODs) than the use of other thresholds or of the raw values of OLR. Such ODs are defined without any time filtering or spatial averaging along the meridional plane. They are also significantly correlated with ODs based on other datasets such as the CMAP and GPCP rainfall estimates and seem more realistic, especially during years when the latitudinal shift is unclear and delayed. However the respective means (30 June for ODs based on OLR vs 25 June for ODs based on CMAP and GPCP) and standard deviations (15.6 days vs 8.1) are slightly different. These differences illustrate the fact that the monsoon onset corresponds to a 10-15 day transition in the West African monsoon seasonal evolution which does not affect concomitantly all variables describing the system.

The onset is clearly linked to the time evolution of a few key-descriptors of the monsoon system at regional scale such as the installation of northward meridional gradients of moist static energy in low levels and of the monsoon cell over the continent. It is also associated with specific variations of the relative vorticity in low troposphere and of the velocity potential in upper levels over the Sahel region. Moreover OD time series exhibits good predictability. Multivariate linear regressions have been performed in a leave-one-out cross-validation way using OLR and selected atmospheric predictors in May. The best results, explain 60% of the OD variance and are obtained with 2 types of predictors: the OLR gradients between the Gulf of Guinea and the African continent, and the northward migration of the West African Monsoon cell.

I. Introduction:

The West African Monsoon consists of a low-level moist flow originating from the equatorial and southern tropical Atlantic basin and converging onto the continent. Following by about 4-6 weeks the Solar zenithal position, this inter tropical convergence zone (ITCZ), moves

therefore northward in spring then southward in autumn and concentrates most of moist convergence in low troposphere along with the major part of organized convective systems and precipitation. As a result, the annual course of rainy seasons is closely associated with the meridional excursion of this monsoon 'system'. So, the Guinean regions record two rainy seasons centered on May-June and September-October whereas areas northward to 8°N register a single rainy season in July-September when ITCZ and its associated rain band reach their northernmost location (Amma, 2004; Ward et al., 2004). During the northward migration of the system the starting of the Guinean then of the Sahelian rainy seasons is usually associated with clear increases in moist convergence, convection and precipitation at regional scale (Louvet *et al.* 2003; Sultan and Janicot 2003). These phenomena called monsoon onsets are nevertheless different: the Guinean onset is mainly due to local increase in convection just southward to the line coast by 5°-6°N; the Sahelian onset results from an abrupt northward shift of the ITCZ from 5°N to 10°N and is preceded in May-June by a preonset stage. The preonset corresponds to the northward shift up to 15°N of the confluence line between moist southwesterly monsoon winds and dry northeasterly Harmattan allowing isolated convective systems to develop ITCZ (Janicot and Sultan 2001, Sultan *et al.* 2003).

The problem of definition and predictability of onsets has been investigated using several approaches. For example, Sultan and Janicot (2003) utilized the daily NCEP/NCAR reanalyses and the daily OLR low-filtered data to remove variability lower than 10 days and applied a semi-empirical criterion based on the synchronous evolution of OLR series averaged over the Sahelian and Guinean belts. Louvet et al. (2003) used Principal Components Analysis (PCA) with a Varimax orthogonal rotation performed on the 5-day CMAP data after application of low pass filter (< 1 month) for detecting the main active and inactive phases of the monsoon through an automatic criterion based on the time evolution of series. More recently, Fontaine and Louvet (2006) compared the sensitivity of onset dates (ODs) to the 5-day GPCP and CMAP rainfall estimates. They define a Northern index (NI) averaging precipitation between 7.5°N and 20°N and a Southern Index (SI) for the region between 7.5°N and the equator after elimination of time variability < 15 days. In such a context the onset date (OD) is the first pentad registering a positive Onset Indice (OI) where OI is the normalized difference between the NI and SI standardized indexes, in CMAP and/or GPCP: *OI= stan [stan(SSI) - stan(GI)]*.

Although conditioned by different purposes and choices, the above approaches give similar ODs in terms of temporal means (24-30 June) and standard deviations (8-10 days). This partly results from use of time filtering and latitudinal averaging along the meridional plane which artificially reduce the real interannual variability of the dates of onset and tend also to artificially reinforce the potential skill of OD statistical prediction. Moreover the 3 approaches allow description of only a small part of variability within the ITCZ core. For example, they do not take directly into account the observed abrupt latitudinal shift in deep convection at the moment of onset since they integrate external information issued from its northern (the Sahel) and southern (the Gulf of Guinean) margins.

The purpose of this article is first to compare different new approaches of monsoon onset detection through selected variables and datasets describing precipitation, convection and atmospheric dynamics. Its first aim is to propose a more direct and precise depiction of the onset at a 5-day time scale without any time filtering and spatial averaging along the meridional plane. The second one is to define a few regional atmospheric descriptors describing the basic key-features of the West African monsoon linked to the onset. The third aim is to analyze the potential predictability of onset dates. The work is based on the empirical evidence that the latitudinal shift in deep convection within ITCZ averaged between 10W and 10E can be considered as a robust and basic signal of the monsoon onset at regional scale. In the following section the data and methodological approaches will be briefly described while sections 3 and 4 will present selected diagnostics and hindcasts. Section 5 will summarize the results.

II. Data and methods

The Sudan-Sahel onset is here defined at the daily and 5-day time-steps using a few basic variables describing the mean processes associated with the West African monsoon, i.e., convection, precipitation, energy and dynamics.

The Outgoing Longwave Radiation (OLR) data are first selected for depicting the onset through convection within the ITCZ. This is estimated from the daily NOAA Interpolated OLR in a 2.5°*2.5° resolution. In this version all gaps have been filled with sophisticated temporal and spatial interpolation as described in Liebmann and Smith (Bulletin of the American Meteorological Society, June 1996). See http://www.cdc.noaa.gov/cdc/data.interp_OLR.html for more details on satellites and day

time crossing. Figure 1 displays OLR values along the meridional plane, for different longitudinal windows in West Africa for the period 1979-2005. If a West African onset can be observed between 20°W and 10°E it is mainly attested eastward to 10°W where the zonal extension of the Guinean coast enhances all meridional sea-land contrasts (Figure 1b,c). Westward to this window i.e., over the Atlantic, or eastward to 10°E, the northward shift is less observed (Fig. 1a,d,e). So all following results will refer to the 10°W-10°E window.

It is worth mentioning that OLR daily values are not direct measurements of organized convective systems because it can exist many individual cloud-types embedded within a common cirrostratus canopy with cirrus and cumulonimbus. More important in clear-sky situations, OLR measures the direct emission from the warmer land or oceanic surfaces. Thus moist convective signals at regional scale cannot be directly derived from spatial averages of OLR values in W/m2 since these arithmetic means will merge both atmospheric and surface emissions. The definition of specific OLR thresholds is hence essential to retain at best the only influence of highly reflective cloud systems. Over West Africa theoretical and empirical evidence shows that OLR daily values lower than 240 W/m2 allow elimination of the signal from warmer surfaces whereas values < 180 W/m2 capture only areas with organized convective systems and many cold and high cloud-tops.
For depicting moderate (strong) deep convective systems within the ITCZ, we adopt the following method: for any day of the period 1979-2005 and for any grid-box of the domain [Eq.-20°N; 10°W-10°E], we allocate the value of 1 if OLR < 240 W/m2 (< 180 W/m2) and 0 if not. Then, we calculate the corresponding sums over each 5-day periods and average by latitude to portray the mean position of the ITCZ core between 10°W and 10°E. In that context, the 5-day values range from 0, meaning no day with OLR < 240 (or < 180) for all longitudes, to 5 meaning 5 days with OLR < 240 (or < 180) at every longitude. The results are then converted in percentages of occurrence for convenience. These percentages will be used for computing OLR based dates of onset.

Rainfall data have been selected from satellite products like the CPC Merged Analysis of Precipitation (CMAP, Xie and Arkin, 1997) and the Global Precipitation Climatology Project (GPCP, Adler et al, 2003 and Xie et al, 2003). These Satellite estimates are available on a 2.5 x 2.5 degree latitude/longitude grid at a 5-day time-scale over the period 1979-2005. It is obviously of paramount importance to estimate the validity of these data versus the real world. Fontaine and Louvet (2006) showed that for onset diagnosis and predictability

purposes the two datasets fit rather well over the West African domain: the CMAP-GPCP relative differences expressed in percentages are negative over land and positive over the Gulf of Guinea but always low (< 10%); temporal correlations are also very high (> 0.96). More details are available in Gruber et al. (2000) and Yin et al (2004) who compare the two precipitation datasets at larger scale.

It is more difficult to compare directly satellite estimates with land-based rainfall data at a daily or 5-day scale. The main reason is that a large number of West African stations suffer from many gaps over the last decades and especially after 2004. For example rainfall stations located between 10W and 10E referenced in the Global Surface Summary of Day (GSOD), one of the most complete land-based data base, show that for many stations, 50% or more of the daily precipitation amounts can be considered as doubtful (incomplete or underestimated) over recent periods. However reliable regional indexes can be defined with selected stations over both the Guinean (05N-7.5N: 37 stations) and Sudano-Sahelian (8.5N-12N: 38 stations) regions for the period 1979-2004. When synoptic variability (< 10 days) is eliminated these rainfall indexes are positively correlated with the respective GPCP and CMAP indices although the results are very sensitive to the time filter retained.

The CMAP and GPCP satellite estimates show also good agreement with land-based rainfall data provided by the CILSS and IRD for the Sudan-Sahel belt (see Louvet et al., 2007 for more details). Over these regions GPCP is generally closer to observations than CMAP in terms of correlation (intra-seasonal rhythms) whereas CMAP is better in terms of bias (pentad anomalies) and of statistical distribution. Obviously such comparisons are challenging over areas documented by a small number of rainfall stations as in Western Sudan (Ali et al., 2004 and 2005). In any case, the gauge data dominates the final GPCP and CMAP analysis over land areas (Dr. Xie's personal communication).

So we choose to define rainfall-based dates of onset mainly through the CPCP and CMAP data. Additionally to obtain a better benchmark of onset database closer to the 'real world' we will compare with ODs based on selected GSOD stations over the Guinean (37 stations) and Sudan-Sahel regions (38 stations).

Atmospheric data have been provided by the NCEP/DOE AMIP-II Reanalysis, the 2nd second version (R-2) of NCEP/NCAR reanalysis (Kanamitsu et al., 2002) over the period 1979-2005. They are available at the same spatial resolution than the OLR, CMAP and GPCP datasets. In

this version of reanalysis, model precipitation has been replaced with observed (satellite + gauge) 5-day precipitation in a similar manner to CMAP (Kanamitsu et al., 2002). Moreover R-2 improves upon the NCEP/NCAR Reanalysis(R-1) by fixing the errors and by updating the parameterizations of the physical processes. Use of SSM/I data and assimilation of rainfall data have been also incorporated. The main improvements regarding the water cycle, as well as the effect of satellite and aircraft, on analyses are extensively discussed in Kanamitsu et al. (2002), in particular the correction method equivalent to using observed 5-day precipitation amounts in the hydrological calculations. Maurer et al. (2001) found that R-2 provides more accurate pictures of soil wetness, near surface temperature and surface hydrology budget over land, and radiation fluxes over Ocean. By contrast the new boundary layer and convection schemes have modified the water vapor profile: R-2 has more moisture in low levels than R-1. For convenience atmospheric variables and indexes used in the present article will be described in section 4.

III. An OLR index for the West African monsoon onset

Hovmöller diagrams in Figure 2 (left panels a,c,e) display mean annual evolutions of OLR values versus percentages of occurrences of moderate and strong moist convection using the OLR thresholds defined above. The superimposed black curves mark OLR minimums and maximums of convection. Clearly, a 180 W/m2 threshold allows better determination of the meridional shift of deep convection within ITCZ (Fig. 2e) than the 240 W/m2 limit (Fig. 2c) or than the raw values (Fig.2a). The northward displacement of moist convective areas in Fig. 2e is more obvious and more precise in time and latitude: it occurs clearly at the turn of June/July between 5°N and 12°N.

Right panels in Fig. 2 b,d,f detail the latitudinal evolutions of OLR and convection: the onset follows a period of OLR increase from the end of May to mid-June, as shown by the values higher than 220 W/m2 in Fig. 2b; this is followed by a clear decrease in both moderate and strong convection within ITCZ, i.e., values < 70% and 15 % in Fig. 2d,f respectively. In particular, Fig. 2f shows that when the ITCZ reaches 10°N, the geographic extent of the strongest convective signals (OLR < 180 W/m2) is smaller than what is observed at 5°N before the onset -25-30% in April-May against less than 20% after-: this change is not attested with moderate convection (Fig. 2d). Only the 180 W/m2 threshold correctly monitor both the abrupt meridional shift and the subsequent decrease in convective maximums within ITCZ. The advantage of using low OLR thresholds such as 180 W/m2 or 200 W/m2 for capturing

better the meridional shift of ITCZ has also been confirmed through composite and clustering analyses (not shown here).

An OLR-based onset date (OD) can then be defined when the 3 following conditions occur successively: i) convective maxima are already located at -or northward to- 5°N during 2-3 pentads; ii) part of the strongest maxima shifts abruptly to 10°N where percentages of OLR < 180 W/m2 are lower than 25%; iii) they remain further at that latitude or move northward.

Figure 3a presents these OLR based onset dates against the CMAP and GPCP OD estimates in bold, solid and dash, respectively. ODcmap and ODgpcp time series are highly correlated (r=+ 0.85): the curves are largely superimposed on the figure because their yearly differences equal zero in 92% of years. These 2 series can therefore be averaged for defining new ODs through precipitation: ODpre= (ODcmap+ODgpcp) / 2, where the subscript ‘ pre’ means precipitation. For comparison dates of onset based on the selected GSOD rain gauges mentioned in the ‘data and methods’ section are also defined. This additional descriptor, called ODgsod, is computed by making each year the rainfall difference between the Sudano-Sahelian and Guinean areas documented by 38 and 37 selected stations respectively, and after elimination of the synoptic variability (< 12 days) with a butterworth filter: the onset date is then the first pentad of any 20-day or longer period registering a positive difference.

Tables 1 and 2 display the basic statistics and ODs obtained with the 3 indexes: they differ regarding their means and variances but are positively correlated: r=+0.76 between ODpre and ODolr; r=+0.39 between ODpre and ODgsod. The lower correlation and higher standard deviation registered with ODgsod are easy to explain: in some cases daily values at station are not reported or underestimated. In Mali and Benin for instance only 50% or less of the GSOD daily values correspond really to the summation of 4 reports of 6-hour precipitation amount, or of 2 reports of 12-hour precipitation amount or of 1 report of 24-hour precipitation amount. As a result the correct geographic coverage can change with days, months and years, which limits the precision of ODgsod.

Figure 3b displays the mean OLR field during the last pentad of June a time of onset. The major convective areas are signaled by values < 240 W/m2. In complement the composite map in figure 3c describes OLR evolution around the onset date by making the OLR difference after and before each ODolr: here positive/negative values stand for local increases/decreases in OLR, and the shadings point out the significant differences at the p= 0.05 level. Significant signals are encountered in the northern subtropics marked by high OLR

values, hence no deep convection, as shown in fig 3b. Notice also that the ODolr and ODpre composites differ in the vicinity of ITCZ: ODolr exhibits negative values in the 7°N-16°N zone just northward to the ITCZ mean position indicated by values < 240 W/m2; by contrast positive OLR differences occur in the Gulf of Guinea. Such diagnostics are coherent with a dipolar pattern evolution at time of onset: decreasing/increasing OLR values, and hence increasing/decreasing moist convection, over the Sahelian/Guinean latitudes. This evolution is more significant eastward to 5°E.

The composite map in figure 3d describes OLR evolution after and before each ODpre. Here also the differences exhibit a dipolar pattern: they are clearly negative southward to ITCZ over the eastern Atlantic and the Gulf of Guinea and slightly positive over central Sahel. Such signals are consistent with decreasing/increasing convection south/north of the ITCZ at the onset date and in harmony with figure 2a,c,e. Since ODolr tends to follow ODpre, both in mean and median, the above observations suggest that convection decreases first over the eastern Atlantic north of the equator and westward to 5°E then, 5-10 days after, increases on Central Sahel eastward to 5°E. The onset is hence well captured by the abrupt 5°N-12°N meridional jump inside the 10°W-10°E window, although at larger scale it is also associated with a SW to NW migration (Fig. 3c,d).

The greatest differences between ODolr and ODpre reach about 3 pentads or 15 days. They tend to happen in years registering the latest ODs, those occurring in July, as observed in Fig. 3a during the years 1988, 1995 and 1998. To focus on this particularity, figure 4 displays together the April-October latitudinal OLR variations and associated maximums in convection. Here ODpre and ODolr series are marked by plus and asters, respectively. It is noteworthy that during these years the onset is neither sudden in time nor abrupt in its latitudinal shift: it corresponds more closely to a succession of short phases of strong convection extended along the meridional plane at least between 5°N and 10°N. Additionally no clear northward displacement of convective maximums within the ITCZ can be attested. In these cases ODolr depicts better the latitudinal shift than ODpre.

IV. Associated atmospheric signals and ODolr predictability

This section analyses time evolution of dynamics and energy in West Africa at time of onset using complementary atmospheric spatial indexes all defined over the 10°W-10°E window:

- Raw OLR in Watt/m2 at 10°N, 5°N and Equator: low (high) values depict more (less) moist convection
- Two indexes based on moist static energy (MSE) in low levels. Let us recall that MSE is linked to the transformation of enthalpy and latent energy available in the lower troposphere into geopotential energy in the upper levels: MSE = gz + CpT + Lq, where gz is the gravitational-potential energy (with g the gravitational acceleration and z the geopotential height of isobaric surfaces); CpT the enthalpy (with Cp the specific heat of dry air at constant pressure (Cp=1000 m2 s-2 deg-1), T the temperature (Kelvin) of isobaric surfaces), and Lq the latent energy associated with evaporation and condensation of water (with L the latent heat of evaporation (L=2501 J/g) and q the specific humidity of isobaric surfaces.
    1. an index describing the MSE flux convergence in the low monsoon layer, vertically integrated from the surface to 850 hPa. This index will be called conv(MSE),
    2. an index depicting the continental meridional gradient of MSE at 1000 hPa between the Sudano-Sahelian and Guinean belts, i.e., 10N-15N and 5N-10N, respectively.
- Three descriptors of atmospheric dynamics in the monsoon region:
    1. A West African Monsoon index (WAMI, Fontaine and Janicot, 1992) to diagnose the West African monsoon cell circulation. WAMI is the difference between the standardized values of the zonal wind at 200 hPa and the wind modulus at 925 hPa: WAMI = stan(M925) – stan(U200). A high/low WAMI represents an enhanced/reduced southwesterly monsoon flux associated with a stronger/weaker Tropical Easterly Jet.
    2. The relative vorticity at 925 hPa averaged over the Western, Central and Eastern Sahel for depicting local motion in the lower branch of the monsoon cell captured by WAMI
    3. The velocity potential at 200 hPa to capture larger-scale dynamics in the upper branch of the monsoon cell.

More details on these descriptors can be found on the website http://www.u-bourgogne.fr/climatologie/AMMA_D1.1.3/ along with comments, maps and time series. This material is downloadable.

Mean annual OLR cycles are displayed as a function of latitude in Fig. 5a for illustrating the northward penetration of moist convection areas -where OLR is low-: in March-April these minima are located by the equator, in May over the Guinean coast by 5°N and in August at 10°N. Since 0LR varies inversely in Sudan-Sahel and along the Guinean coast, the intersection between the 5°N and 10°N curves around the end of June in Fig. 5a marks the mean onset date.

This inverse evolution is not perceptible with WAMI because this index depicts more the extent and coherence of the monsoon cell at a sub-continental scale: evolutions of the WAMI index over the Sahelian, Sudanian and Guinean belts do not clearly differ; they all peak during the JAS rainy season. In Fig. 5c the period of onset corresponds to the 3 WAMI maximums by the end of June, at the very beginning of the tropical rainy season.

By contrast MSE convergence there is more evidence for a latitude dependence. In Fig. 5e the values peak 2-3 months before the different rainy seasons: over the Guinean belt the highest values occur in February and August before the May-June and September-October rains, respectively; over Sudan and Sahel they take place in April and May before the July-September season. The mean date of onset occurs when the Sahelian curve crosses the two others in figure 5e. The Sahelian onset is then concomitant of i) a clear decrease in MSE convergence over the Sahel, and ii) an increase at the Guinean latitudes while a minimum happens between the two regions.

The right panels in Fig. 5 display the major changes observed in OLR, WAMI and MSE evolutions for a period of sixty days centered each year on the date of monsoon onset to verify whether coherent leading signals exist in the 5-day tendencies before the monsoon jump. Tendencies are defined as the differences between 2 consecutive pentads. Each 5-day tendency is judged significant if it exceeds the mean June-July tendency +/- 2 standard deviations. This is indicated by markers in Fig. 5b,d,e.

At time of onset (OD on the x-axis) OLR diminishes at 10°N whereas the values slightly augment southward mainly between 5°N and the equator (Fig 5b). This signal is preceded by a modification of monsoon dynamics: ten days before OD, all WAMI indexes increases strongly, especially over the Sahelian belt and reach a first maximum at the onset (Fig. 5d). Similarly, the MSE flux convergence increases over the Sahel whereas divergence occurs southward, as shown in Fig. 5f with the Guinean minimum.

In order to see if such signals could be useful for prediction purposes with sufficient leading-time, we calculate correlation coefficients between ODolr and some selected atmospheric and OLR indexes over a 60-day period preceding each OD. Markers are superimposed when the values are judged significant at P=0.05, taking into account autocorrelation in the series. Fig. 6a displays the results obtained with meridional OLR gradient over the continent, defined as the differences between the 10°N-5°N and 5°N-Eq: a northward gradient is defined positive. Clear changes in correlation appear 30 days before the onset date. A late onset tends to be preceded by 1) a southward OLR gradient anomaly over land between 10N and 5N, and 2) a northward gradient over the Gulf of Guinea from equator to 5N. An early onset is associated with the reverse signals: a northward/southward OLR gradient over land/ocean.

Significant correlations are also obtained with the WAMI and Conv(MSE) meridional gradients, defined as the differences between the Sahelian and Guinean belts (Fig. 6b). These gradients capture respectively the northward migrations of the monsoon cell and of the MSE flux convergence. All curves exhibit significant signals 5-20-day before onsets: a late/early onset is preceded by positive/negative WAMI gradients and by negative/positive Conv(MSE) gradient.

Monsoon onsets are also directly linked to the evolution of dynamical atmospheric variables over the Sahelian belt. This can be illustrated through the relative vorticity in low tropospheric levels (RV at 925 hPa) and the velocity potential in upper levels (VP at 200 hPa) averaged over the Western, Central and Eastern Sahel regions. Here RV features atmospheric dynamics in low levels in terms of wind curl and is positive with cyclonic motion. By contrast VP refers to larger-scale atmospheric motion in high troposphere, since the VP gradients are proportional to the irrotational flux: VP is positive when large-scale divergence occurs in this flux. The diagrams 7a,b present the mean evolution of these quantities over a 300-day time window centered each year on ODolr.

In figure 7a, one can notice that onsets are concomitant with RV minimum in Central Sahel. This indicates a northward migration of cyclonic curl within the Saharan heat low associated with RV maximums which strongly develops at solstice. Southward to that region the irrotational component of the low-level flow can reinforce. Figure 5b shows a VP maximum at onset which expresses the upsurge of the divergent component of the flow in high levels: this tends to reinforce the upper branch of the monsoon cell which is directly linked to the upper and lower flows via latent heat release over regions of deep convection. Thus most onsets are captured both by variations in moist static energy which measures the quantity of

available energy in the atmosphere and by the WAMI index directly linked to the lower and upper branch of the monsoon cell.

Figure 8a displays the mean annual evolution of the monsoon system as a function of MSE gradients and WAMI. Additionally, the period during which conv(MSE) is greater over the Sahel than over Sudan is marked with circles whereas the beginning and end of this season are indicated by pentads #33 and #53, respectively (between mid-June and the end of September). Notice that on the average pentad # 36 i.e., end of June, is the first one registering the 3 following conditions: grad(MSE) > 0, WAMI > 0.1 in standardized units and conv(MSE) stronger over Sahel than southward. So it is assumed that the onset can be also captured by the sole MSE and WAMI descriptors. It is hence possible to define an atmospheric onset date: ODatm occurs when, for a period of at least 20 days, the MSE gradient between Sudan-Sahel and Guinea is oriented northward while the WAMI index is clearly positive (> 0.1) according to the installation of the monsoon cell over the continent, as a positive WAMI depicts by construction SW winds in low levels and easterlies at 200 hPa.

Figure 8b presents the respective ODatm versus ODpre and ODolr series in Julian days over the period 1979-2004: ODatm is marked with squares while ODolr and ODpre are displayed with asters and circles, respectively: clearly the onset dates depend on the used criterion. The figure illustrates also the internal consistency of all descriptors: ODatm and ODolr series register about the same mean (day#181) and standard deviation (13 days) and are positively correlated (r=+0.52). This confirms that the onset cannot be reduced to a sudden change affecting all variables describing the monsoon circulation: it is a 10-15 day transition period separating 2 quasi-equilibrium periods centered in May-June and July-September which does not affect simultaneously all variables at a daily or 5-day time scale.

The question of ODolr predictability is now examined at a 5-day timestep through multivariate linear regressions using evolution in May, i.e. pentads # 25-30 of a few key atmospheric variables. Two indexes linked to the northward displacement of the monsoon system appeared efficient: 1) the OLR differences between 5°N and the equator for estimating displacements in deep moist convection over the Gulf of Guinea; 2) the WAMI difference between Sahel and Guinea for measuring the northward migration of the monsoon cell. Table 3 shows that when the regression is performed in a leave-one-out cross-validation way (Michaelsen, 1987), 60% of the OD variance is explained by the northward evolution of these

2 descriptors. Notice that the selected predictors are not collinear since the variance inflation factor is low, largely < 5 in Table 3. Signs of coefficients show that late ODs are associated with 1) OLR gradient anomalies oriented southward in May which depicts a northward gradient in moist convection anomalies; 2) northward WAMI gradient anomalies in May, meaning a northernmost location of the monsoon cell over West Africa. Early ODs are obviously linked to the reverse signals.

Additional analyses have been performed with the same predictors by dividing the series into different training and validation periods. Despite some decrease in the variance explained the results show generally good agreement. For example when the validation period refers to 1992-2004 (training period =1979-1991) the correlation between observations and hindcasted series is +0.65; a very similar value is obtained when the training and validation periods are permuted. If the training refers to 20 consecutive years, correlations over the validation periods (each of 6 years) can show stronger variations due to the small length of the series. When the training is defined over 1979-1998, the correlation on the validation period 1999-2004 is +0.77.

V. Conclusion.

The purpose of this article was to propose a direct approach of the latitudinal shift in deep convection within ITCZ, without any time filtering and spatial averaging along the meridional plane. It used mainly an interpolated version of the NOAA OLR at a 5-day time-step over the domain Eq.-20°N; 10°W-10°E (1979-2004). This choice is based on the evidence that moist convective signals of regional scale cannot be directly derived from spatial indexes computed with raw values of OLR in W/m2, because any average will tend to merge both atmospheric and surface emissions and not the only influence of highly reflective cloud systems. At time of onset, OLR values increase (i.e., convection decreases) first on the eastern Atlantic northward to the equator and westward to 5°E, then they decreases (i.e., convection increases) on Central Sahel eastward to 5°E.

Efficiencies of specific OLR thresholds, 240 W/m2 and 180 W/m2, have then been discussed. Theoretical and empirical evidence exists that indexes based on OLR < 180 W/m2 capture better deep convective areas inside the ITCZ than OLR < 240 W/m2. The strongest convective signals with OLR < 180 W/m2 allowed us to correctly portray several basic

features of the Sahelian onset in particular its meridional shift and the subsequent decrease in convection maximums within ITCZ.

So, Onset Dates (ODs) have been defined with OLR signals < 180 W/m2 without any time filtering or spatial averaging in latitude as follows: i) convective maxima are located at -or northward to- 5°N during 2-3 pentads; ii) part of these strongest maxima shifts abruptly to 10°N where percentages of areas with OLR < 180 W/m2 are lower than 25%; iii) they remain further at that latitude or move northward.

.

They show high correlation (r=+0.76) with ODs based on CMAP and GPCP rainfall estimates but differ in means (30 June with OLR versus 25 June with precipitation) and standard deviations (15.6 days with OLR against 8.1 with precipitation). This larger variability is more realistic because the monsoon onset does not affect simultaneously all variables of the system at a 5-day time step and cannot be reduced to an abrupt change in monsoon circulation: it is essentially a 10-15 day transition which separates 2 quasi-equilibrium periods of the monsoon system centered on May-June and July-August-September. It has been more complex to define accurate ODs with land-based rain gauges using due to the scarcity of the station network and to the quality of many reported amounts in the GSOD dataset.

OD time series based on the northward migration of deep convective signals inside the ITCZ depict better onsets than ODs based on large spatial averages or defined with other variables such as satellite estimates or land-based precipitation. This appears well when the onset is unclear or delayed, as in years 1988, 1995 and 1998. ODs occur when convection over the Gulf of Guinea decreases while monsoon dynamics changes over the continent. This relationship has been portrayed with a few selected descriptors describing atmospheric dynamics at regional scale, such as the monsoon cell depicted by the WAMI index, moist static energy gradients, flux convergence and relative vorticity in low levels, velocity potential at 200 hPa. In fact most onsets can be captured by the sole variations of moist static energy gradients and of the monsoon cell.

Onset dates based on deep convective signals exhibit also good predictability. Multivariate linear regressions have been performed through both a leave-one-out cross-validation way and different training and validation periods using 2 basic atmospheric indexes in May: OLR differences between 5°N and the equator and WAMI differences between Sahel and Guinea.

The results show that 60% of the OD variance is explained when the regression is performed in a leave-one-out cross-validation way: late ODs are preceded by a northward displacement of moist convection over the equatorial Atlantic associated and a northward displacement of the monsoon cell over West Africa. Early ODs are obviously linked to the reverse signals.

Acknowledgements:

The authors are very grateful *to the* NOAA-CIRES ESRL/PSD Climate Diagnostics branch, Boulder, Colorado, USA, for providing OLR data (at http://www.cdc.noaa.gov/). The study was supported by the Global change and Ecosystems programme (EU Integrated project: African Monsoon Multidisciplinary Analysis (AMMA) and the French component of AMMA. Based on an 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 major financial contribution from the European Community's Sixth Framework Research Programme. Detailed information on scientific coordination and funding is available on the AMMA International web site http://www.amma-international.org

References

Adler, R. F., Huffman , G. J., Chang, R., Ferraro, A., Xie, P., Janowiak, J. E., Rudolf, B., Schneider, U., Curtis, S., Bolvin, D. T., Gruber, Susskind, J. and Arkin, P. 2003. The version 2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979-present), *J. Hydrometeor*, *4* (6), 1147-1167.

Ali A., Amani A. and Lebel T. 2004. Estimation des pluies au Sahel: utilisation d'un modèle d'erreur pour évaluer les réseaux sol et produits satellitaires. Sécheresse 3 (15):271-278.

Ali A., Amani A., Diedhiou A. and Lebel T. Rainfall Estimation in the Sahel 2005. Part II: Evaluation of Rain Gauge Networks in the CILSS Countries and Objective Intercomparison of Rainfall Products. Journal of Applied Meteorology 44 (11):1707–1722.

Amma 2004. African Monsoon Multidisciplinary Analysis : description of work, Paris, 350 pp.

Barnston, A.G. 1992. Correspondence among the correlation, RMSE and Heidke score, *Wea Forecast*, 7, 699-709.

Chatterjee, S. and Price, B. 1977 Regression analysis by example, John Wiley & Sons.

Folland, C.K., J. Owen, M.N. Ward, and A.W. Colman 1991. Investigation of seasonal rainfall in the Sahel region using empirical and dynamical methods, *Jour. of Forecasting*, **10**, 21-56.

Fontaine B. and Janicot S. 1992. Wind field coherence and its variations over West Africa**.** J. of Climate., 5, 512-524.

Fontaine B. and Louvet S. 2006. Sudan-Sahel rainfall onset: Definition of an objective index, types of years, and experimental hindcasts, J. Geophys. Res., 111, D20103, doi:10.1029/2005JD007019.

Gruber, A., Su, X., Kanamitsu M. and Schemm , J. 2000. The comparison of two merger rain gauge – Satellite precipitation datasets, *Bull. American, Met. Soc., 81* (11), 2631-2643,.

Janicot, S., B., Sultan 2001. Intra-seasonal modulation of convection in the West African monsoon, *Geophysical Research Letters, 28*, 523-526.

Kanamitsu, M., W. Ebisuzaki, J. Woolen, S.K. Yarg, J.J. Hnilo, M. Fiorino and G.L. Potter 2002. NCEP-DOE AMIP-II reanalysis (R-2) *Bull. Amer. Meteor. Soc., 83*, 1631-1643.

Louvet, S., B., Fontaine and P. Roucou, 2003. Active phases and pauses during the installation of the West African monsoon through 5-day CMAP rainfall data (1979-2001), *Geophysical Research Letters, 30, 24*, 2271-2275.

Louvet, S., B., Fontaine and P. Roucou, 2006. Which rainfall dataset can be used to study African monsoon at intra-seasonal timescale?, submitted to the Journal of Applied Meteorology and Climatology, JAMC-1859.

Maurer EP, O'Donnell GM and Lettenmaier DP 2001. Evaluation of the Land Surface Water Budget in NCEP/NCAR and NCEP/DOE Reanalyses using an Off-line Hydrologic Model. Journal of Geophysical Research 106 (D16): 17.841-17.862.

Michaelsen, J., 1987. Cross-validation in statistical climate forecast models, *J. Climate Appl. Meteor., 26*, 1589-1600.

Sultan, B. and S. Janicot 2003. The West African Monsoon Dynamics. Part II : The "Preonset" and the "Onset" of the summer Monsoon, *Journal of Climate, 16*, 3407-3427.

Sultan, B. and S., Janicot 2000. Abrupt shift of the ITCZ over West Africa and intra-seasonal variability, *Geophysical Research Letters, 27*, 3353-3356.

Sultan, B., Janicot, S., and A. Diedhiou 2003. The West African Monsoon Dynamics. Part I : Documentation of Intra-seasonal variability, *Journal of Climate, 16, 21*, 3389-3406.

Ward, N., K. Cook, A. Diedhiou, B. Fontaine, A. Giannini, A. Kamga, P.J. Lamb, A. Ben Mohamed, A. Nassor and C. Thorncroft *2004.* Seasonal-to-Decadal Predictability and Prediction of West African Climate, *CLIVAR Exchanges Volume 9 No.3, 14-20.*

Xie P. and P.A. Arkin 1997. Global precipitation: a 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs, *Bull. Amer. Meteor. Soc., 78*, 2539-2558.

Xie, P., Janowiak, J. E., Arkin, P. A., Adler, R. F., Gruber, A., Ferraro, R., Huffman, G. J. and Curtis, S. 2003. GPCP pentad precipitation analyses: An experimental dataset based on gauge observations and satellite estimates, *Journal of Climate, 16*, 2197–2214.

Yin, X., Gruber, A. and Arkin, P. A. 2004. Comparison of the GPCP and CMAP merged gauge satellite monthly precipitation products for the period 1979-2001, *Journal of Hydrometeorology, 5*, 1207-1222.

Tables

1979-2004	Mean	Median	Std	Range
ODolr	181 (June 30)	183 (July 2)	15.6	153-208
ODpre	176 (June 25)	174 (June 23)	8.1	165-193
ODgsod	190 (July, 10)	187 (July, 7)	20.5	142-247

Table 1: Basic statistics for ODolr and ODpre in julian days over the period 1979-2004.

	1979	**1980**	**1981**	**1982**	**1983**	**1984**	**1985**	**1986**	**1987**	**1988**	**1989**	**1990**	**1991**
ODcmap	167	172	182	172	172	192	167	172	182	177	182	177	192
ODgpcp	167	167	182	172	172	192	172	172	177	177	182	177	192
ODgsod	202	182	192	187	182	217	167	187	182	187	202	192	217
	1992	**1993**	**1994**	**1995**	**1996**	**1997**	**1998**	**1999**	**2000**	**2001**	**2002**	**2003**	**2004**
ODcmap	172	177	172	167	167	172	182	192	177	187	167	167	187
ODgpcp	172	172	172	167	167	157	182	192	177	182	172	167	172
ODgsod	177	177	182	177	202	197	142	247	187	192	217	177	167

Table 2 : Dates of onset derived from the 3 rainfall datasets in julian days: ODcmap, ODgpcp, ODgsod.

	P#26 5N-Eq. OLR grad.	P#29 5N-Eq. OLR grad	P#27 SAH-GUI WAMI diff	P#30 SAH-GUI WAMI diff
Reg.Coeff R2=60%	-0.41	-0.47	+0.25	+0.18
Corr. Coeff.	-0.45	-0.60	+0.48	+0.46
VIF=1.1	1.05	1.08	1.12	1.17

Table 3: Cross-validated hindcasts of ODolr using pentad (P) values of 4 predictors in May (pentads# 26-30) based on OLR and WAMI meridional gradients (line 1), along with explained variance (R2) and the regression coefficients obtained with standardized series (line 2). The variance inflation factor (VIF) is displayed in line 4 following Chatterjee and Price (1977).

Figure captions

Figure 1: April-October latitudinal mean variations of OLR values in W/m2 averaged over five longitudinal windows: (a) 30°W-20°W, (b) 20°W-10°W, (c) 10°W-10°E, (d) 10°E-20°E, (e) 20°E-30°E windows. Period 1979-2005.

Figure 2 : Comparison between selected April-October signals in OLR averaged in the 10°W-10°E window as a function of latitude: (a) OLR mean, (b) OLR minimum; (c,e) percentages of occurrences with OLR < 240 and < 180 W/m2, respectively; (d,f) percentages of occurrences for convection maximums. The superimposed bold curve on left panels marks the core of ITCZ using OLR minimums (a) and convection maximums (c,e). Period 1979-2005.

Figure 3: (a) Time series of the different onset dates: ODcmap, ODgpcp and ODolr; (b) mean OLR field in W/m2 around the onset date; (c) OLR tendency = mean field difference between OLR observed after and before ODolr each year (see Table 1); (d) as (c) but for ODpre. In c,d the shadings point out the significant differences at p= 0.05 regarding a Student t-test. Period 1979-2005.

Figure 4: Hovmöller diagrams averaged in the 10°W-10°E window for the years 1988, 1995 and 1998: OLR values (left) and maximums of convection (right). The superimposed '+' and '*' mark ODpre and ODolr, respectively.

Figure 5: Mean composite evolution and tendencies before and after ODs for selected indexes averaged in the 10°W-10°E window : (a,b), OLR at 10°N, 5°N and equator; (c,d) WAMI and Conv(MSE) (e,f) over central Sahel, central Sudan and central Guinea. Markers are superimposed when a local (5-day) tendency is greater than mean tendency plus 2 standard deviations in June-July.

Figure 6: Time evolution of correlation coefficients between ODolr and selected indexes averaged in the 10°W-10°E window:

a) OLR difference between 10°N and 5°N (dashed line) and between 5°N and equator (solid); b) WAMI and Conv(MSE) differences between central Sahel and Guinea (dashed and solid lines respectively).

Figure 7: Time evolution of atmospheric dynamics stratified against ODolr (lag time in days) for the Western (dashed line), Central (red bolded line) and Eastern Sahel (dash-dotted line). (a) Relative vorticity at 925 hPa in s^{-1} ; (b) Velocity potential at 200 hPa in s^{-1}. The onset date (OD on x-axis) is also indicated by markers on the curves.

Figure 8: a) Mean annual evolution of the meridional MSE gradient between [10N-15N] and [05N-10N] and of the WAMI index averaged over the Sahelian and Guinean zones (standardized units). Letters for months; the circles mark the period during which conv(MSE) is greater over Sahel than over Sudan between pentads #33 and #53. Pentad # 36 is the mean date for ODatm, i.e., when grad(MSE) > 0, WAMI > 0.1 and conv(MSE) occurs mainly over Sahel.

b) Comparison between ODolr in Julian days, precipitation (circles) and atmospheric descriptors (squares) relative to the longitudes 10W-10E and the period 1979-2004.

Figure 1: April-October latitudinal mean variations of OLR values in W/m2 averaged over five longitudinal windows: (a) 30°W-20°W, (b) 20°W-10°W, (c) 10°W-10°E, (d) 10°E-20°E, (e) 20°E-30°E windows. Period 1979-2005.

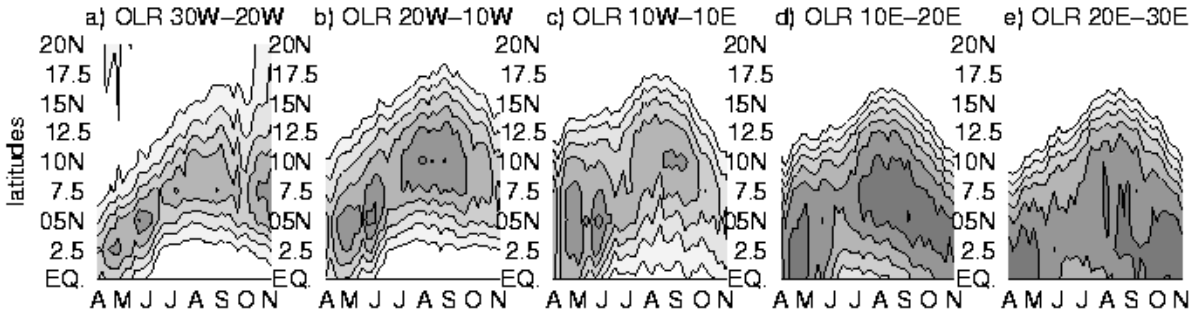

Figure 2 : Comparison between selected April-October signals in OLR averaged in the 10°W-10°E window as a function of latitude: (a) OLR mean, (b) OLR minimum; (c,e) percentages of occurrences with OLR < 240 and < 180 W/m2, respectively; (d,f) percentages of occurrences for convection maximums. The superimposed red curve on the left panels marks the heart of ITCZ using OLR minimums (a) and convection maximums (c,e). Period 1979-2005.

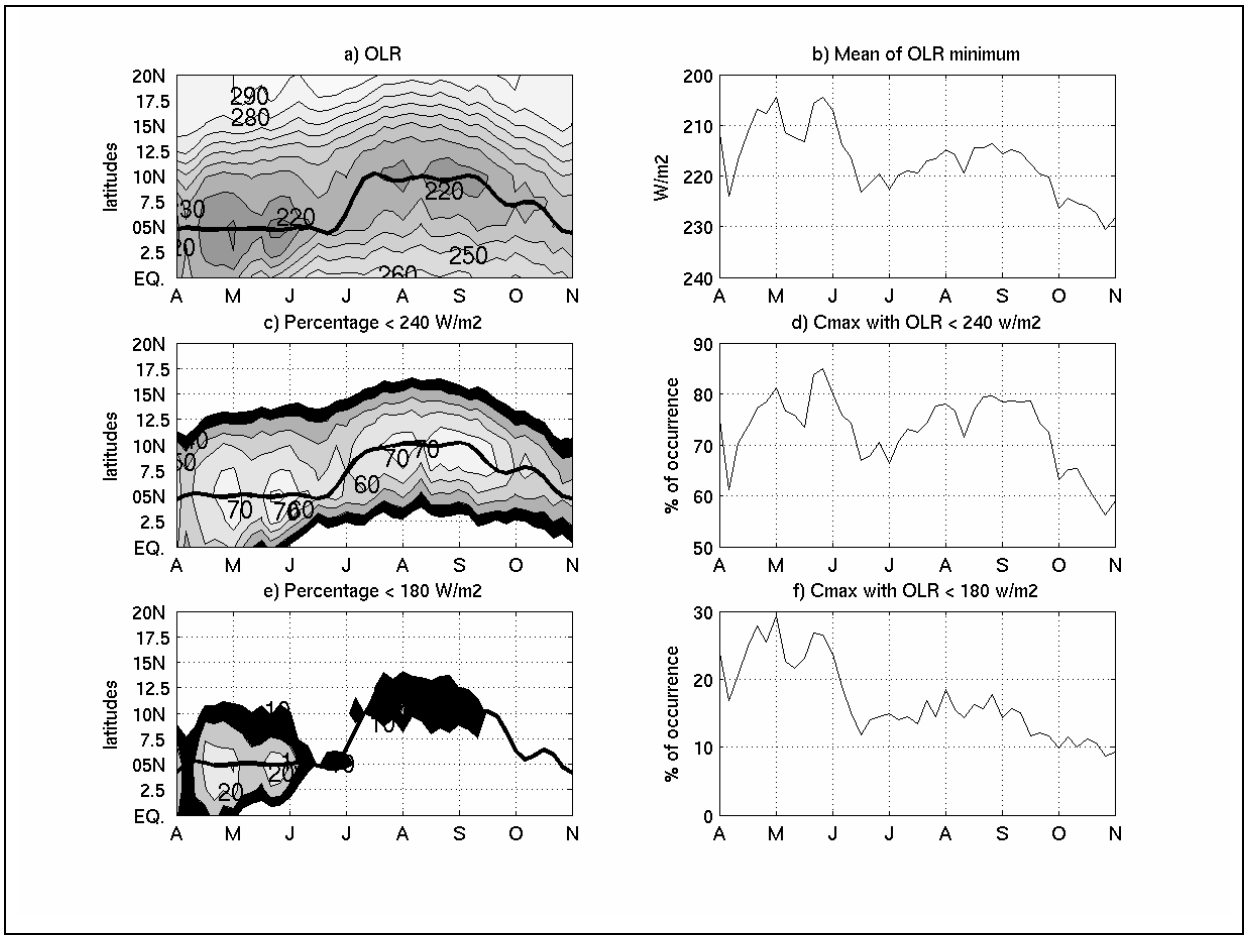

Figure 3: (a) Time series of the different onset dates: OD(cmap), OD(gpcp) and OD(olr); (b) mean field of OLR around OD(olr) in W/m2; (c) OLR tendency (mean field difference between OLR observed after OD(olr) minus OLR observed before each year (see Table 1); (d) as (c) but for OD(pre). In c,d solid (dashed) lines point out negative (positive) OLR differences associated with increasing (decreasing) convection and shadings the significant differences at p= 0.05 regarding a Student t-test. Period 1979-2005.

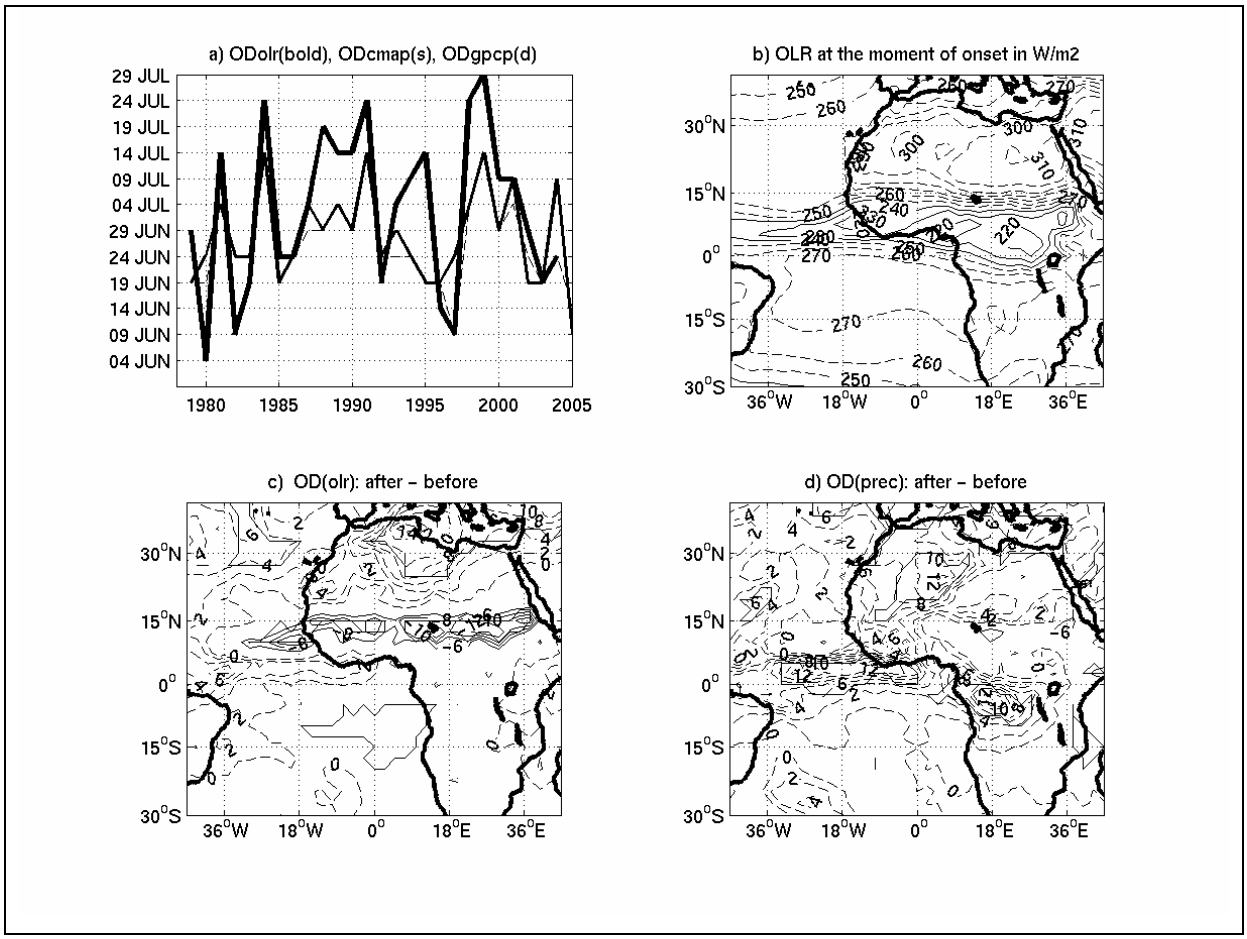

Figure 4: Hovmöller diagrams averaged in the 10°W-10°E window for the years 1988, 1995 and 1998: OLR values (left) and maximums of convection (right). The superimposed '+' and '*' mark OD(pre) and OD(olr), respectively.

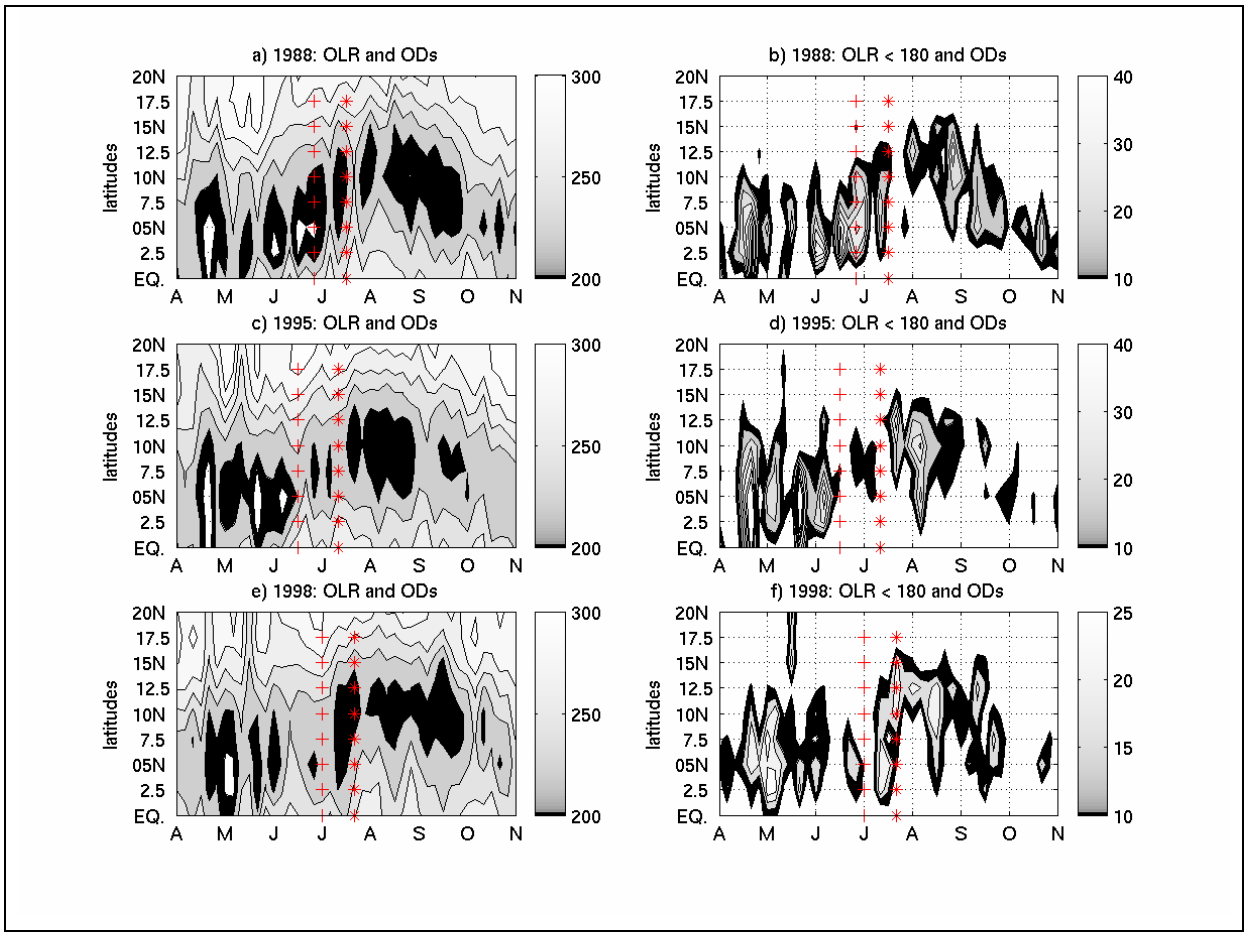

Figure 5: Mean annual and tendencies (composite evolution before and after ODs) of selected indexes averaged in the 10°W-10°E window : (a,b), OLR at 10°N, 5°N and equator; (c,d) WAMI and (e,f) Conv(MSE) over central Sahel, central Sudan and central Guinea. Markers are superimposed when a local (5-day) tendency is greater than mean tendency plus 2 standard deviations in June-July.

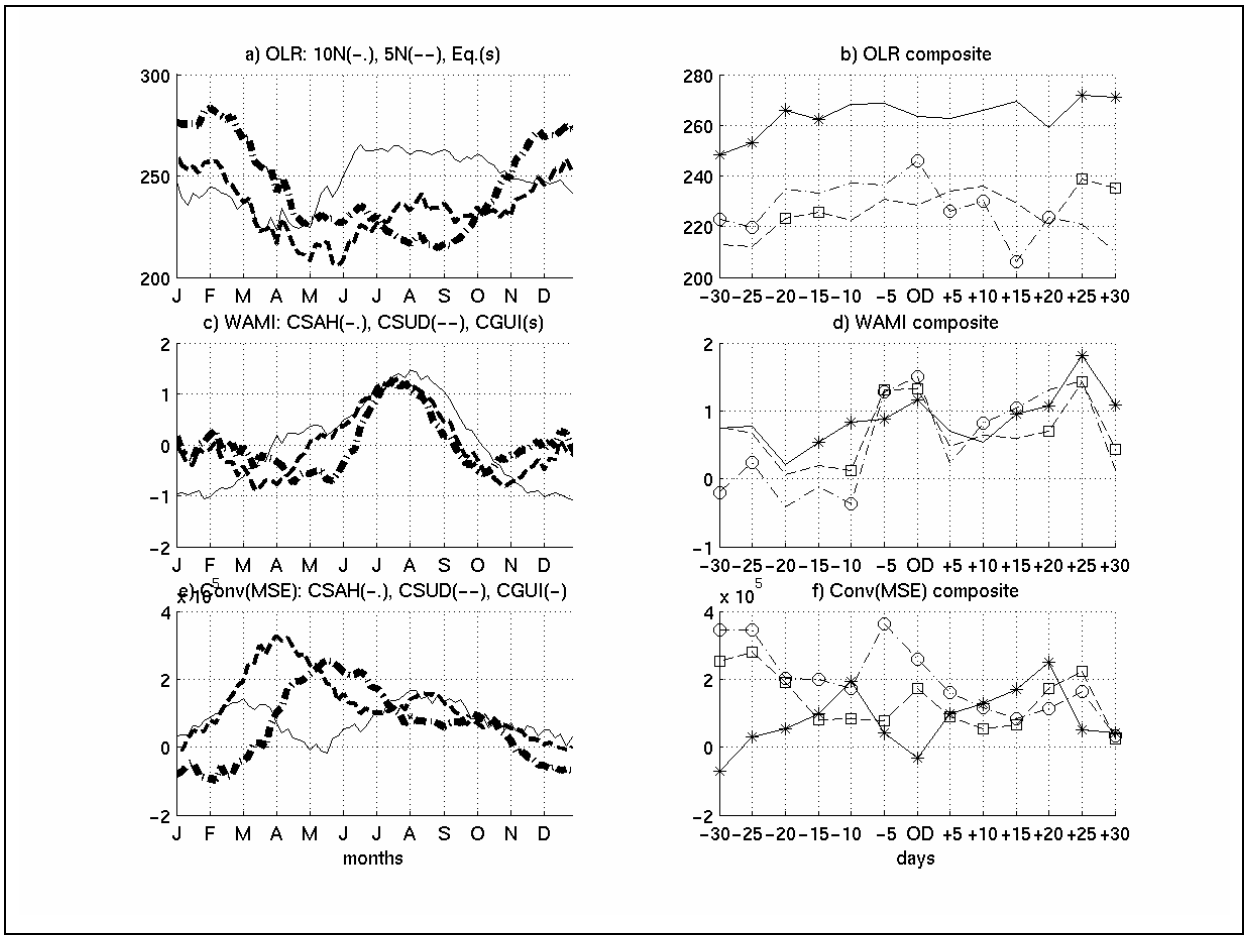

Figure 6: Time evolution of correlation coefficients between OD(olr) and selected indexes averaged in the 10°W-10°E window:

a) OLR difference between 10°N and 5°N (dashed line) and between 5°N and equator (solid);

b) WAMI differences between central Sahel and Guinea (dashed line) and Conv(MSE) difference between central Sudan and Guinea (solid line).

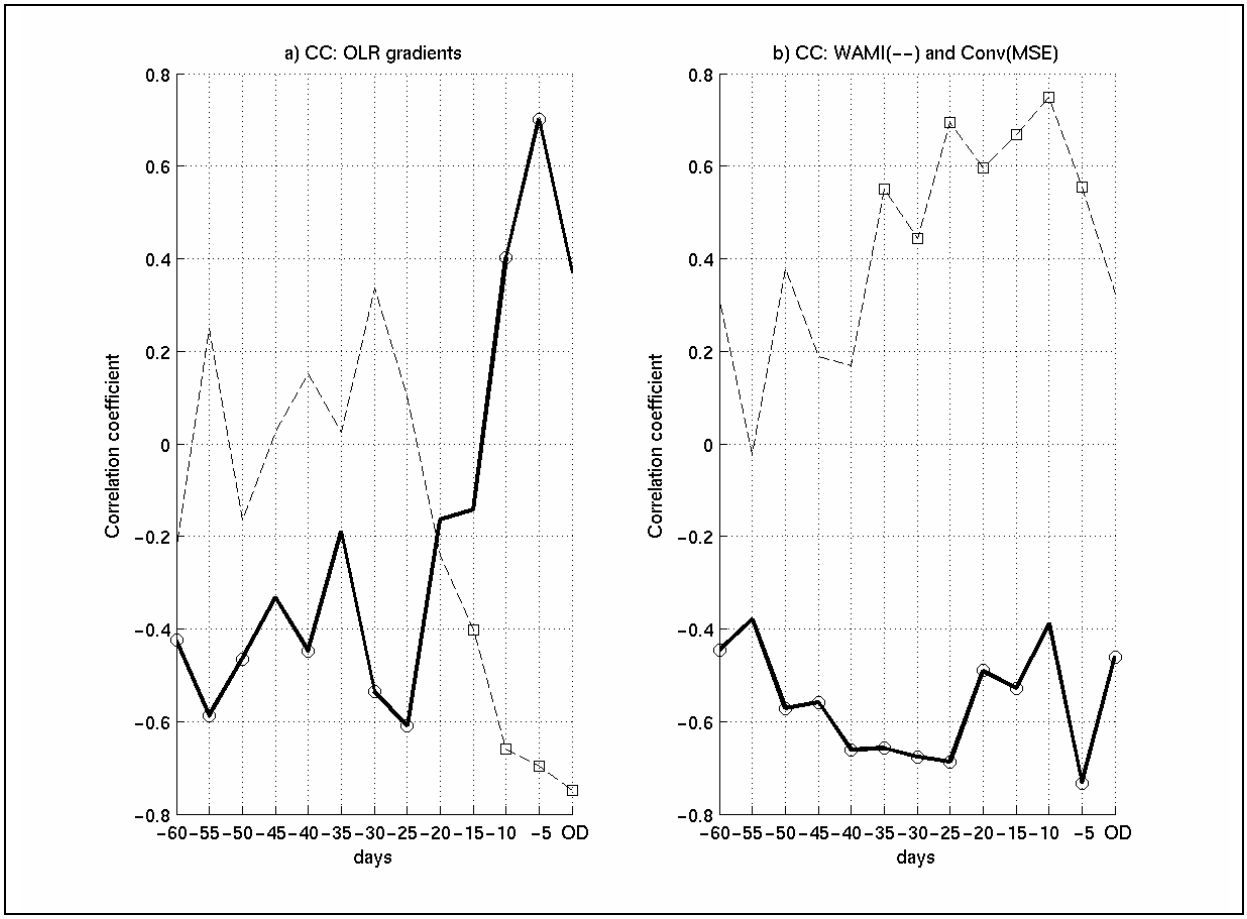

Figure 7: Time evolution of atmospheric dynamics stratified against OD(olr) (lag time in days) for the Western (dashed line with a square), Central (bolded line) and Eastern Sahel (dash-dotted line with a triangle). The onset date (OD on x-axis) is also indicated by markers on the curves. (a) Relative vorticity at 925 hPa in s^{-1} (b) Velocity potential at 200 hPa in s^{-1}.

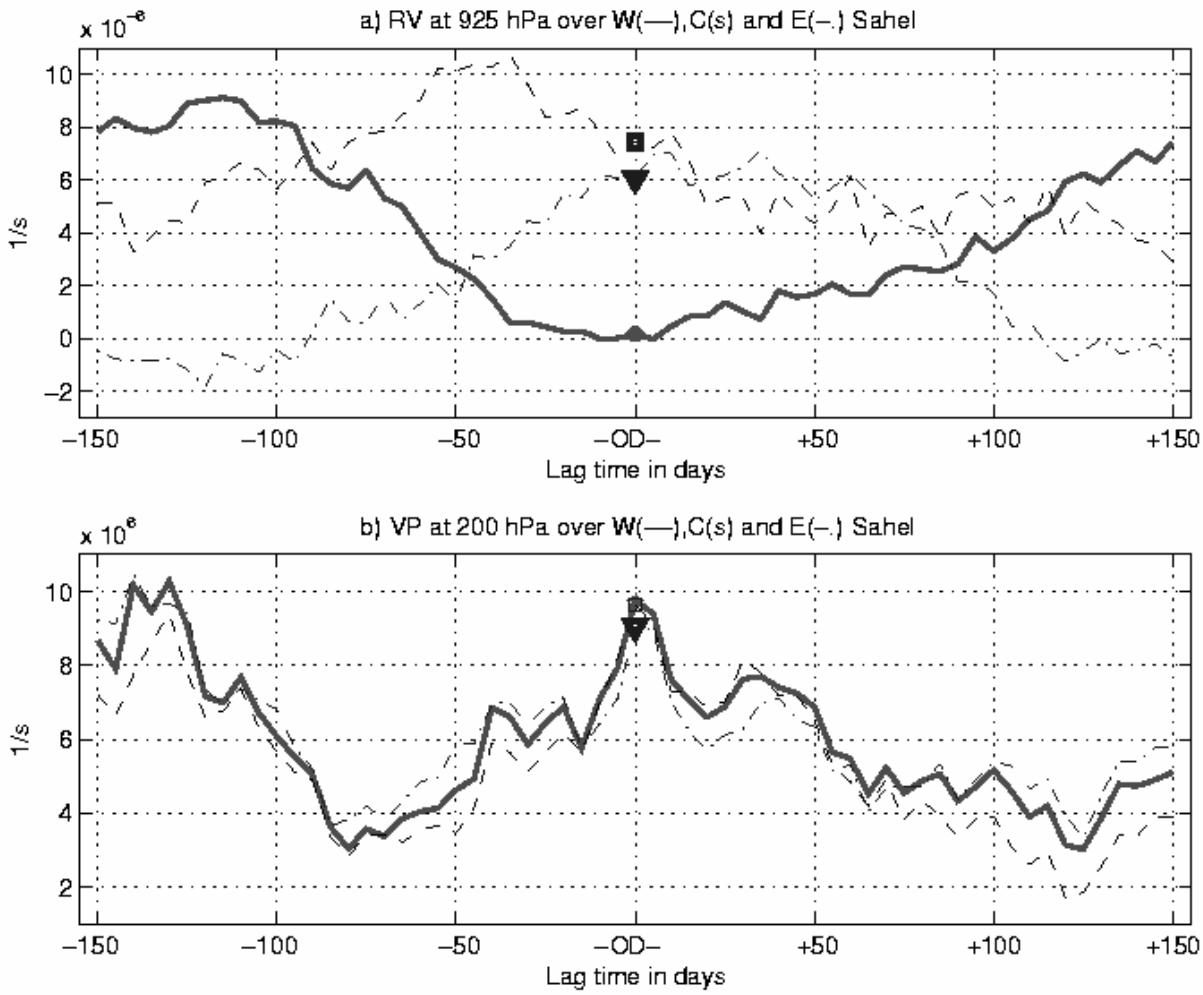

Figure 8: a) Mean annual cycle of selected atmospheric parameters: the meridional MSE gradient between [10N-15N] and [05N-10N] vs WAMI over the Sahelian and Guinean zones (standardized units). Letters for months; the circles mark pentads when conv(MSE) is greater over Sahel than over Sudan between pentads #33 and #53. Pentad # 36 is the mean date for OD(atm), i.e., when grad(MSE) > 0, WAMI > 0.1 and conv(MSE) occurs mainly over Sahel. b) Comparison between ODs(olr) (in Julian days), precipitation (circles) and atmospheric descriptors (squares). Longitudes [10W-10E], period 1979-2004.