



## Sudan-Sahel rainfall onset: Definition of an objective index, types of years, and experimental hindcasts

Bernard Fontaine<sup>1</sup> and Samuel Louvet<sup>1</sup>

Received 21 December 2005; revised 22 May 2006; accepted 26 June 2006; published 21 October 2006.

[1] Using rainfall estimates from the 5-day version of CPC Merged Analysis of Precipitation (CMAP) and Global Precipitation Climatology Project data along with multiscale spatial key descriptors of atmospheric dynamics from National Center for Environmental Prediction/Department of Energy reanalysis 2 (NCEP/DOE), we first define a West African monsoon onset index to determine its successive dates of occurrence over the period 1979–2004 (28–29 June in mean with a standard deviation of 8.5 days). Then we focus on the three main types of time evolutions of that index at the moment of the monsoon onset in terms of precipitation, pressure, temperature, and winds at different levels to detect the most robust associated signals and select on the mid-May, mid-June period different sets of potential predictors for onset dates. Basically, late (early) onsets are preceded by more (less) rainfall southward to the equator by the end of May, then by a clear decrease (increase) of the normal northward cross-equatorial gradient in mid-June. Finally, we present experimental cross-validated hindcasts of the dates of onset always based on four predictors using both reanalyzed atmospheric data and observed rainfall estimates. Accurate forecasting schemes are obtained, especially with CMAP rainfall (time tendencies and cross-equatorial gradients) over the Gulf of Guinea. In this context the hindcasted series are highly significant when compared to observations (75% of explained common variance with the sole rainfall predictors) and can reach 82% of variance when a zonal wind descriptor of regional scale is added. Basically, a late (early) onset is preceded by more (less) rainfall southward to the equator by the end of May, then by a clear decrease (increase) of the normal northward cross-equatorial gradient by mid-June.

**Citation:** Fontaine, B., and S. Louvet (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.

### 1. Introduction

[2] At first order the West African monsoon (WAM) circulation, mainly results from sea-land contrasts and consists of a southwesterly flow in the lower levels, which transports water vapor from the Atlantic Ocean. Because of the horizontal component of the Coriolis acceleration, this flow turns eastward north of the equator and converges onto the continent in the ITCZ and its associated rainbelt, which together migrate between the Guinean coast to the southern margins of the Sahara. This movement, toward north in spring, and south in autumn, explains the mean annual evolution of precipitation: the Guinean regions record two clear rainy seasons peaking in May–June and September–October whereas the Sudanian and Sahelian belts register a single rainy season which peaks in August after the monsoon onset. At regional scale this generates strong meridional gradients of specific humidity, precipitation,

vegetation, and soil moisture: see, among others, *Moron et al.* [1995], *Polcher* [2004], and *Ward et al.* [2004] for more details.

[3] Monsoon onsets observed over the Guinean and Sahelian regions have been recently studied as intraseasonal components of rainfall and atmospheric variability in the WAM system by *Grodsky and Carton* [2001], *Sultan and Janicot* [2000], *Janicot and Sultan* [2001], *Sultan et al.* [2003], and *Louvet et al.* [2003], among others. The purpose of this study is not to suggest a new line of attack of intraseasonal variability over WAM but to evaluate the potential forecast skill linked to rainfall and atmospheric conditions occurring a few weeks before WAM onsets. *Omotsho et al.* [2000] have prediction schemes for some West African stations (Bamako, Niamey, and Kano) using a reanalysis of in situ rainfall data. We propose hereinafter a preliminary statistical approach of onset diagnostics and predictability as well as some experimental hindcasts using basic descriptors issued from a few selected rainfall and atmospheric data sets currently available. Diagnostic of onset takes into account the abruptness and amplitude of the latitudinal shift of the ITCZ as well as the concomitant increase/decrease of Sahelian/Guinean precipitation. The

<sup>1</sup>Centre de Recherches de Climatologie, CNRS/University of Burgundy, Dijon, France.

**Table 1.** Comparison Between the CMAP and GPCP 5-Day Precipitation Data Sets Over the Period 1979–2004 After Removing the Mean Annual Cycles<sup>a</sup>

	JFM		AMJ		JAS		OND	
	RD, %	TCC						
Land	+0.4	+0.99	-3.1	+0.97	-3.9	+0.98	-2.3	+0.97
Ocean	+4.5	+0.96	+4.4	+0.96	+8.8	+0.97	+5.3	+0.98

<sup>a</sup>Relative differences (RD) in percentages  $50(\text{CMAP} - \text{GPCP})/(\text{CMAP} + \text{GPCP})$  and temporal correlation coefficients (TCC) are averaged by season on the continental ( $20^{\circ}\text{N}-5^{\circ}\text{N}$ ) and oceanic ( $5^{\circ}\text{N}-5^{\circ}\text{S}$ ) West African monsoon (WAM) areas. CMAP, CPC Merged Analysis of Precipitation; GPCP, Global Precipitation Climatology Project; JFM, January, February, March; AMJ, April, May, June; JAS, July, August, September; OND, October, November, December.

ultimate objective (i.e., in near future) will be to provide more accurate and comprehensive hindcasts and, if possible, real-time predictions of these events using both historical data and numerical simulations such as the ECMWF forecasts for the time range 10–30 days, operational since October 2004.

[4] In the following section the data and methodological approach will be briefly described while section 3 will present a few selected results. We will provide first some statistics on the dates of onset and on the relevant aspects of rainfall and atmospheric field evolutions around these dates, then we will focus on onset predictability and prediction by presenting and discussing several experimental hindcasts of the dates of monsoon onset. Section 4 will summarize the results and give the future perspectives.

## 2. Data and Methods

[5] Rainfall data have been selected from the CPC Merged Analysis of Precipitation (CMAP) [Xie and Arkin, 1997] and from the Global Precipitation Climatology Project (GPCP) [Adler et al., 2003; Xie et al., 2003] estimates on a  $2.5 \times 2.5$  degree latitude/longitude grid at a 5-day timescale over the period 1979–2004. The two precipitation data sets were compared by Gruber et al. [2000] and Yin et al. [2004]. The first authors show good spatial and temporal correlations between the two sets. In particular, the temporal correlation over West Africa and the tropical Atlantic Ocean exceed +0.90 for the monthly field and +0.70 with the annual cycle removed, the best fit occurring over the continent south of the Sahara. In northern spring and summer, the mean precipitation differences (CMAP–GPCP) are positive over land and negative over ocean with a typical difference of  $\pm 0.1$  mm/day [Gruber et al., 2000, Figure 7, Table 3, p. 2638]. This could be because, in contrast with GPCP, CMAP uses both SSM/I microwave scattering and Microwave sounding unit over the ocean, uncorrected rain gauge over land, and an OLR precipitation index over land and ocean [Gruber et al., 2000]. Yin et al. [2004] precise that (1) the GPCP–CMAP spatial correlation is higher over land than over ocean; (2) in general, over-ocean precipitation represented by the GPCP is more reasonable while over land the two products are close; and (3) discrepancies can however exist in sensitive areas such as equatorial West Africa. For our purpose (onset diagnosis and predictability) and over the West African monsoon (WAM) domain here considered, i.e., a squared

area of  $20^{\circ}$  extension in latitude and longitude (equ- $20^{\circ}\text{N}$ ;  $10^{\circ}\text{W}-10^{\circ}\text{E}$ ), the two data sets fit rather well as shown in Table 1: the CMAP–GPCP relative differences expressed in percentages are negative over land and positive over the Gulf of Guinea but always rather low ( $<10\%$ ) and the temporal correlations very high ( $>+0.96$ ). So to avoid redundancy, the results displayed in other tables and figures will refer to the CMAP data set; however, the text will indicate the differences when necessary. This choice is justified by the results of Ali et al. [2004, 2005]. These authors have recently compared on a monthly basis and a  $2.5^{\circ} \times 2.5^{\circ}$  latitude/longitude resolution the in situ precipitation estimates using the Comité Inter-étatique de Lutte contre la sécheresse au Sahel (CILSS) rain gauge network (managed by AGRHYMET) in the  $10^{\circ}\text{N}-20^{\circ}\text{N}$  West African belt during the rainy season, versus a lot of rainfall estimates: Global Precipitation Climatology Centre (GPCC), Global Precipitation Index (GPI), GPCP, and CMAP. Ali et al. [2004, 2005] clearly demonstrate that the CMAP data set is the best one in terms of bias, root mean square, and distribution at this timescale.

[6] The selected CMAP version depends mainly on observations: rainfall forecasts from the model are not assimilated and observations from rain gauges have been merged with estimates from infrared and microwave data from satellites. CMAP analysis comprises two phases. In the first one, all estimates except the gauge-based analysis are combined using a maximum likelihood estimate with weights derived by comparison to the gauge analysis (over land) or in a simplified form (over oceans). This reduces the overall random error, but leaves the systematic error. The gauge analysis is then used over land, when sufficient gauge observations are available, or atoll-based gauge observations over the tropical oceans, to determine the absolute value of the analysis, while preserving the gradients derived earlier.

[7] Atmospheric data provide from NCEP/DOE AMIP-II reanalysis (R-2) [Kanamitsu et al., 2002] over the period 1979–2004 available at the same spatial resolution than CMAP data. We have primarily considered the fields known to impact WAM energetics and dynamics: sea level pressure (SLP), skin temperature, moist static energy (MSE), and wind speed at 925, 600, and 200 hPa for documenting the monsoon flow, the African easterly jet and the tropical easterly jet, respectively.

[8] The statistical approach comprises four steps:

[9] 1. To depict the monsoon onset, we first define two rainfall indexes over a squared area of  $20^{\circ}$  extension in latitude and longitude (equ- $20^{\circ}\text{N}$ ;  $10^{\circ}\text{W}-10^{\circ}\text{E}$ ) using both the 5-day CMAP and GPCP data: a northern index (NI) averaging precipitation north of  $7.5^{\circ}\text{N}$  and a southern index (SI) for the region extending between  $7.5^{\circ}\text{N}$  and the equator. Then a WAM onset index (WAMOI) is defined as the difference between the NI and SI standardized indexes, after elimination of time variability  $<15$  days by a butterworth time filter [Murakami, 1979]: When the rainbelt migrates northward (southward), WAMOI values increase (decrease). In that context the onset date is defined as the first pentad of a 20-day (or longer) period registering positive WAMOI in both CMAP and GPCP. WAMOI refers to the  $10^{\circ}\text{W}-10^{\circ}\text{E}$  window where a clear meridional sea-land contrast exists.

**Table 2.** Dates of Onset in Pentad Numbers Using CMAP and GPCP, if Different, Over the Period 1979–2004<sup>a</sup>

	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991
CMAP	34	35	37	35	35	39	34	35	37	36	37	36	39
GPCP		34					35		36				
	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
CMAP	35	36	35	34	34	35	37	39	36	38	34	34	38
GPCP		35				32				37	35		35

<sup>a</sup>Means are 35.9 versus 35.6, standard deviations are 1.7 versus 1.7, and correlation coefficient between the two series is +0.85\*\*\*.

[10] 2. A clustering algorithm (K-means) is then applied on selected subsamples of WAMOI series (the April–July seasons) to detail the main types of spring to summer evolutions on the period 1979–2004. K-means clustering is an algorithm to classify objects (here years) based into K number of groups, by minimizing the sum of squares of distances between data and the corresponding cluster centroid. Here the clustering algorithm is repeated 1000 times with a new set of initial centroids to assess robustness. This allows us to separate the robust signals from those occurring possibly by chance or too noisy. Several K numbers (2 to 5 classes) have been tried and we report only those with K = 3 which can be estimated as the most efficient choice with respect to the length of the series (26 years).

[11] 3. Rainfall and atmospheric patterns describing time tendencies at the moment of the onset are analyzed using CMAP and selected National Center for Environmental Prediction/Department of Energy reanalysis 2 (NCEP/DOE) (NCEP2/DOE) fields like sea level pressure, skin temperature and horizontal wind field at 925 hPa (monsoon level), 600 hPa (midtroposphere), and 200 hPa (upper troposphere). The most robust signals are then considered as potential predictors and interpreted in terms of pressure and moist static energy gradients along the meridional and zonal planes.

[12] 4. Finally, experimental hindcasts of the dates of onset are performed using multivariate linear regression analyses with objective (stepwise) selection of different sets of potential predictors (rainfall, pressure, temperature, moist static energy, and winds at different levels) after application of a cross-validation procedure. The reported results will concern always cross-validated hindcasts with four predictors. To ensure robust skills we applied a leave-one-out cross-validation, meaning that the predictive information of the year to predict is not used; that is, coefficients of regression for a given year are therefore calculated on the other (25) years. As a result, each yearly hindcast derives from a model on all available data excluding one year to minimize potential skill inflation, which may arise from the multiannual persistence of predictor and/or predictand conditions. However, one important point must be kept in mind: if cross-validation is used to obtain, to first order, uninflated skill estimates and provides a good estimate of the true forecast skill in large samples [Michaelsen, 1987; Barnston, 1994], they are not similar to a real operational context; in particular, the skill obtained is generally slightly better.

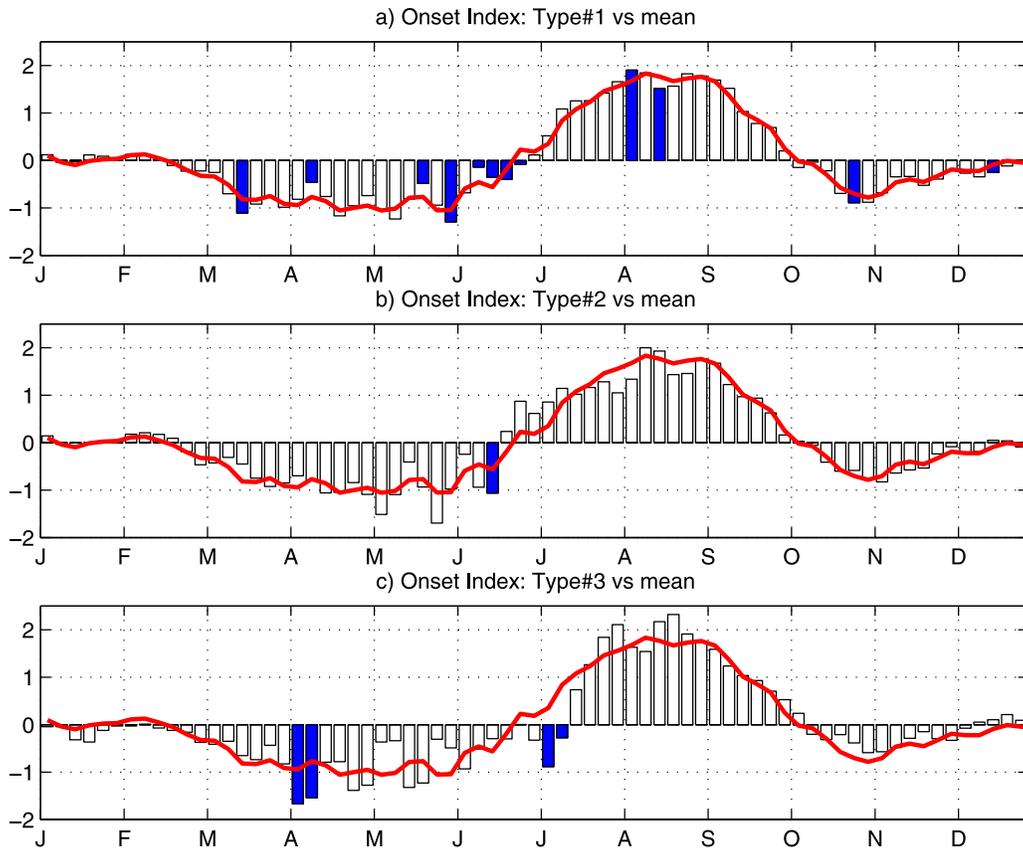
### 3. Dates of Onset and Dominant Types of Spring to Summer Rainfall Evolutions

[13] Several onset definitions for the WAM domain have been already published. *Omotosho et al.* [2000, p. 869] use

a few in situ rainfall stations with the following onset definition: "... the first three or four rainfalls of at least 10 mm with not more than 7 days between them." By contrast, *Sultan and Janicot* [2003] defined the Sahelian's onset as the moment of the latitudinal shift of the ITCZ from 5°N to 10°N through daily gridded rainfall data during the period 1968–1990, and OLR gridded data over the 10°W–10°E longitudinal window and the period 1979–1990. *Louvet et al.* [2003] diagnose the most coherent rainfall evolutions over homogeneous regions using principal components analysis (PCA) to document the main modes of 5-day CMAP variability (main patterns and associated time components over the period 1979–2001) after elimination of fluctuations (<1 month). In this last approach the Sahelian onset is viewed as an important intraseasonal modulation of precipitations among others, which does not only concern Sahelian latitudes. All these approaches give similar mean dates, although some discrepancies appear in specific years.

[14] Here the dates of onset have been defined through the WAM onset index already defined (see point 1 above) using the 5-day CMAP and GPCP data sets over the period 1979–2004 and the longitudinal window 10°W–10°E: This window is the region where the abrupt shift of the ITCZ is the most obvious [Sultan and Janicot, 2003; Louvet et al., 2003].

[15] The results can be compared in Table 2 both in terms of means and interannual variability. They are very similar: the CMAP and GPCP mean dates occur in the last days of June (35.9 versus 35.6 pentad numbers, respectively, meaning June 28–29) with a standard deviations of 1.7 pentads (8.5 days); the correlation coefficient between CMAP and GPCP time series is high and significant (+0.85\*\*\*). Indeed, the CMAP-GPCP differences are equal to zero in 69% of occurrences (18 years out of 26) and do not exceed one pentad in 92% of years. The slight difference in the mean dates (0.3 pentad or 1–2 days) is coherent with the CMAP underestimation/overestimation over land/ocean already shown in Table 1. Given this global accordance we will focus hereafter on the sole CMAP results to avoid redundancy. It is noteworthy, however, that the monsoon onset is not exactly similar to the beginning of the tropical rainy season, which usually refers to smaller spatial scale and is motivated by an agro-meteorological approach. For example, our onset dates differ for several years from the growing season onset dates in northern Nigeria published by *Ati et al.* [2002], who use in situ observations over six stations northward to 12°N. By contrast, our onset dates are in good coherence with the results of *Louvet et al.* [2003] and *Sultan and Janicot* [2003] and are directly linked to monsoon dynamics. Thus it is possible to determine the mean onset date with an index of the northward shift of the



**Figure 1.** (a–c) Three main types of mean annual evolutions for West African monsoon onset index (WAMOI, bars) along with the 1979–2004 mean, after elimination of time variability <15 days by a butterworth time filter. Shadings point out the significant signals at  $p = 0.05$  regarding a Student  $t$  test. Years composing each type and the respective dates of onset are listed in Table 3.

African Easterly Jet in midtroposphere at the moment of the monsoon onset: the AEJ core located by  $10^{\circ}\text{N}$ – $12^{\circ}\text{N}$  decreases abruptly because (1) the meridional thermal gradient weakens at this moment (not shown) and (2) intense African easterly waves (AEWs) appear as shown by *Gu and Adler* [2004]. The spring to summer transition over West Africa is also marked by changes in synoptic-scale waves, that is, from the eastward-propagating Kelvin-type modes during spring to the westward-propagating AEWs during summer.

[16] Years entering into each type of years are listed in Table 3 along with their mean dates of onset. The respective mean annual rainfall evolutions are shown in Figure 1. Type 1 is clearly the dominant type and represents 50% of years (13 years out of 26, Table 3). Its mean onset date, marked in Figure 1a by the first positive bar at the end of June (the 36th pentad), is very close to both the mean date and annual evolution (curve on Figure 1) and is coherent with *Sultan and Janicot* [2000] and *Louvet et al.* [2003]. For these

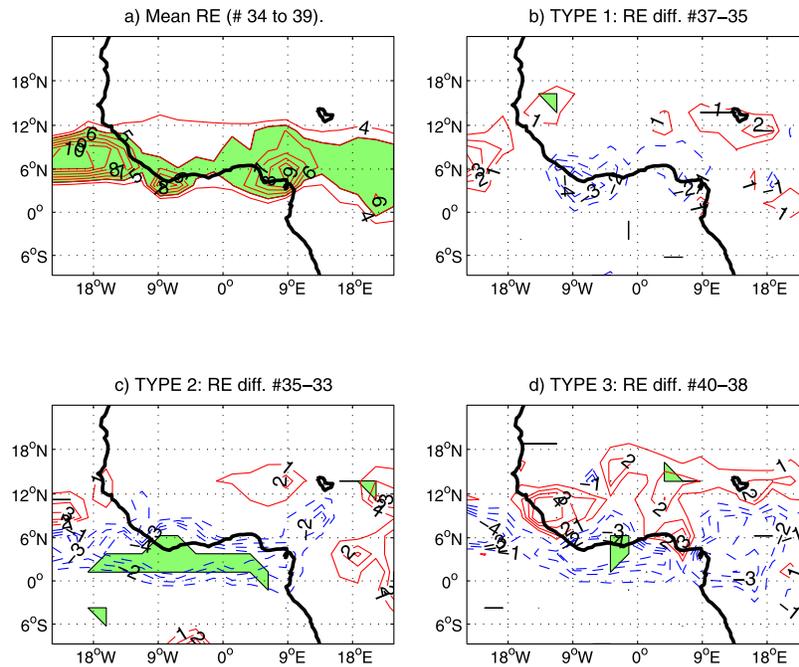
reasons, type 1 will be called hereinafter “normal” onset. By contrast, years classified into type 2 (type 3) tend to register onsets in advance (in late) with a mean date of occurrence by mid-June (before mid-July). So, types 2 (31% of years) and 3 (19%) will be called “early” and “late” onsets, respectively.

[17] For each type, Figure 1 points out the most significant pentads ( $p = 0.05$ , bars shaded) all along the year regarding a Student  $t$  test calculated on the differences between the years composing the type and all the remaining years: these signals tend to occur one or several pentads before the mean dates of onset (Figures 1a–1c).

[18] CMAP estimates (RE) at the time of the mean onset date (pentad number 36) is displayed in Figure 2a after elimination of synoptic variability (maxima greater than 5 mm/day are shaded). The rainband, which marks the preferential ITCZ locations, extends inside the  $3^{\circ}\text{N}$ – $12^{\circ}\text{N}$  latitudinal belt. However, within the  $10^{\circ}\text{W}$ – $10^{\circ}\text{E}$  window this band is thinner and displaced southward, toward the

**Table 3.** Mean Dates of Onset in Pentad Numbers and Years by Types and for All the Available Period 1979–2004 Using CMAP

Type	Mean Date of Onset	Years
Mean	35.9	all years
Type 1	35.8	1981, 1982, 1986, 1987, 1988, 1989, 1990, 1992, 1993, 1996, 1997, 1998, 2000
Type 2	34.4	1979, 1980, 1983, 1985, 1994, 1995, 2002, 2003
Type 3	38.6	1984, 1991, 1999, 2001, 2004



**Figure 2.** (a) Rainfall estimates (RE) around the onset in mm/day and (b–d) composite evolutions before and after the yearly onsets for the three dominant types. For Figures 2b–2d the values in mm/day are relative to the dates of onset listed in Table 2 and expressed as simple differences between 5-day rainfall amounts registered after and before the onset date: If  $RE(i)$  denotes the rainfall estimates registered during one pentad ( $i$ ) during which the onset occurs, the values refer to  $RE(i + 1) - RE(i - 1)$ . Shadings point out values greater than 5 mm/day in Figure 2a and significant differences at  $p = 0.05$  regarding a Student  $t$  test in Figures 2b–2d.

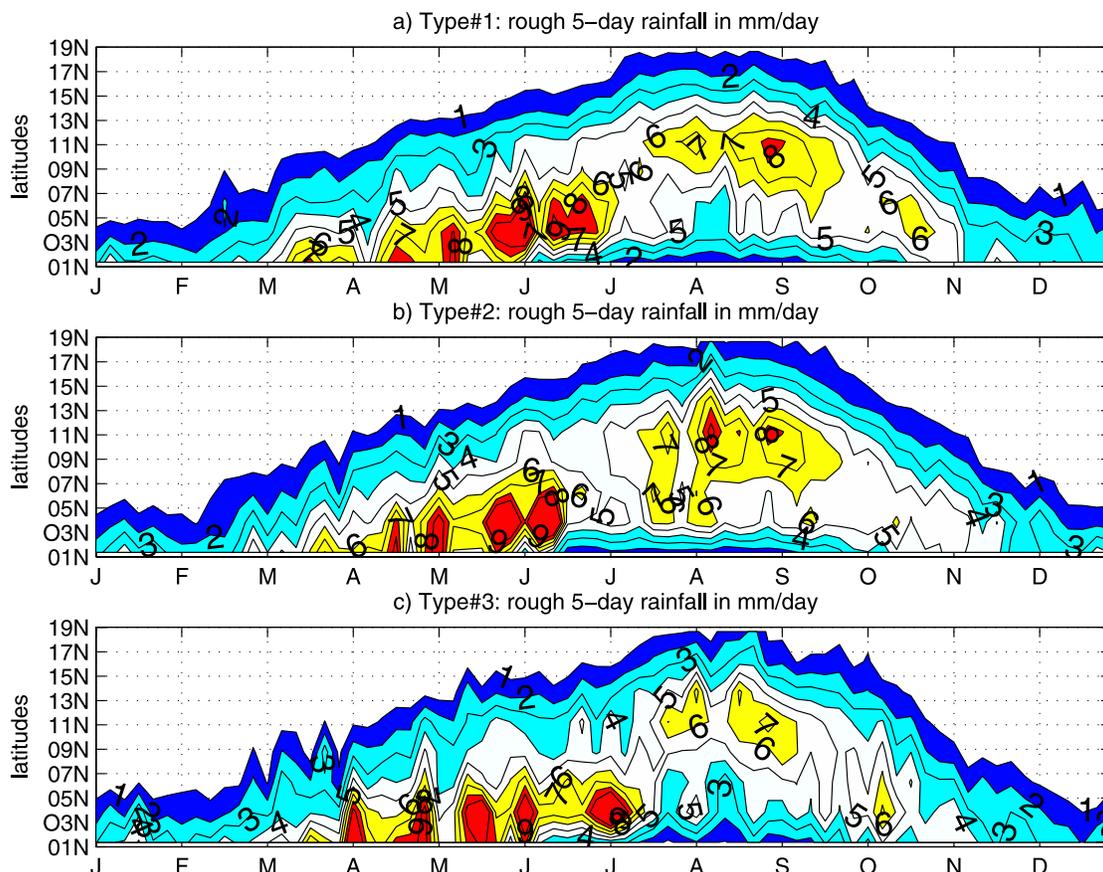
Guinean coast. The composite fields of the differences between the 5-day rainfall amounts registered after and before the “normal,” “late,” and “early” onset dates are displayed in Figures 2b–2d, respectively: Here any positive (negative) value means an increase (a decrease) in precipitation at the onset period whereas the shadings point out the significant signals regarding a Student  $t$  test ( $p = 0.05$ ). Basically, the onset corresponds to a sudden northward excursion of the rainband: the largest rainfall increases (northward to  $10^{\circ}\text{N}$ ) are associated with decreasing precipitation south of the Guinea coast. However, this feature characterizes more the “normal” onsets than the “early” and “late” situations, which are somewhat different. For example, “early” onsets (Figure 2c) appear linked to decreasing rainfall over the Atlantic Ocean south of  $6^{\circ}\text{N}$  but do not exhibit any important Sahelian increase in the  $10^{\circ}\text{W}–10^{\circ}\text{E}$  window. If we consider the total period, and hence the three types, the most robust signals are therefore limited to the equatorial Atlantic, south of the Guinean coast.

[19] In order to better detail such rainfall signals over the  $10^{\circ}\text{W}–10^{\circ}\text{E}$  longitudinal window in years with “normal,” “early,” and “late,” onsets, Figure 3 displays the three time latitude diagrams of rough (unfiltered) 5-day CMAP data. It is first noteworthy that the onset tends to occur just after the vanishing of the last physically powerful rain events ( $>6$  mm/day) south of  $10^{\circ}\text{N}$  at the end of the first Guinean rainy season and leads by 2–4 pentads rainfall maxima over Sudan-Sahel ( $10^{\circ}\text{N}–13^{\circ}\text{N}$ ). As a result, the onset date is not accurate at a daily scale. Compared to “normal” years (Figure 3a), “early” years (Figure 3b) register wetter con-

ditions and a longer rainy season north of  $10^{\circ}\text{N}$  after the onset date, whereas the rainy seasons following “late” onsets are drier and shorter.

#### 4. Associated Atmospheric Fields

[20] Similar analyses have been conducted using the reanalyzed atmospheric fields to give us an idea about the robustness of these signals and to determine their spatial scale. Figures 4a and 5a display the mean fields of skin temperature (SKT) and sea level pressure (SLP) conditions by the mean date of onset (the 36th pentad of the year, by the end of June). In these figures, SKT (SLP) fields registering values lower than  $30^{\circ}\text{C}$  ( $1015$  hPa) are marked by dashed lines. Notice that, northward to the Sahelian belt ( $\sim 15^{\circ}\text{N}$ ) and especially over western Sahara and Mauritania, the SKT maxima are close to the SLP minima (Figures 4–5): this points out the continental heat low region bordered by the two Azores and Santa-Helena high-pressure cells over the tropical Atlantic. Figures 4b–4d and 5b–5d present the mean SKT and SLP tendencies at the dates of “normal,” “early,” and “late” onsets, respectively. Here the positive (negative) tendencies are displayed with solid (dashed) lines and with shadings if significant at  $p = 0.05$ . Basically, the onset period corresponds to a large northward translation of SKT maxima over the western Sahara as shown by the positive values extending north of  $15^{\circ}\text{N}$  and the negative values southward in Figures 4b–4d. However, some distinctions must be made. Years of type 1 (“normal onset”) exhibit the largest signals over the WAM region, especially a clear dipolar pattern (warming/cooling



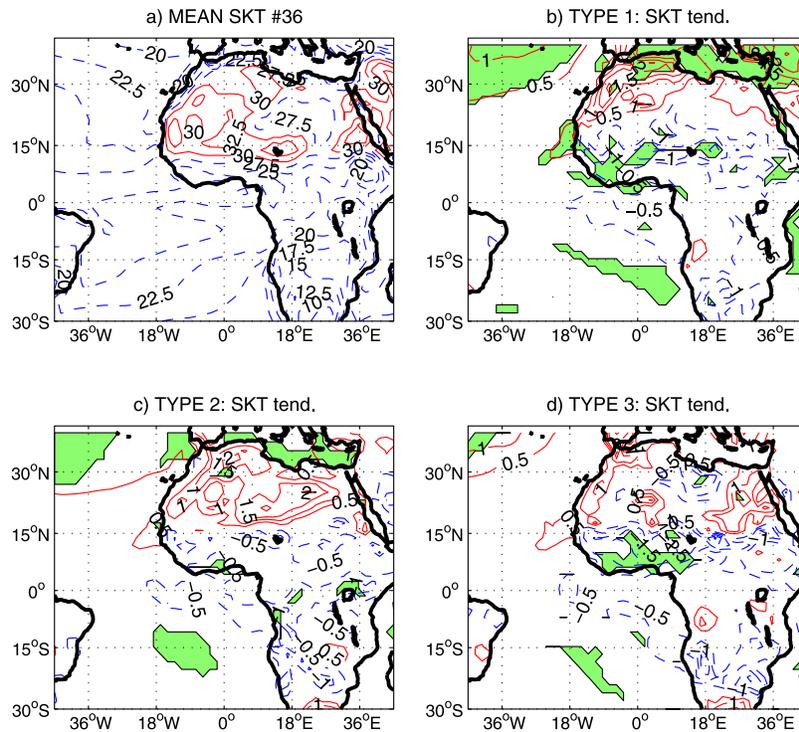
**Figure 3.** (a–c) Three WAMOI types in terms of time-latitude diagrams of rough (not filtered) 5-day CPC Merged Analysis of Precipitation (CMAP) data (period: 1979–2004), averaged over the  $10^{\circ}\text{W}$ – $10^{\circ}\text{E}$  longitudinal window. The units are expressed in mm/day.

north/south of  $\sim 20^{\circ}\text{N}$ ) with significant values south of  $15^{\circ}\text{N}$  (Figure 4b). This is coherent with (1) the SLP increase in the vicinity of the two Azores and Santa-Helena anticyclonic cells (positive values in Figure 5b) and (2) the SLP decrease (negative values) over the heat low region in western Sahara. Such a SLP anomaly pattern reinforces the generated pressure force anomaly region toward the WAM region, and hence wind convergence. By contrast, type 2 (“early onset”) does not show any significant signal located over the continent (Figures 4c and 5c) but (1) a reinforcement of the Azores cell associated with a warming over the northern tropical Atlantic and the Mediterranean Sea and (2) a cooling over the eastern Atlantic by  $\sim 15^{\circ}\text{S}$ . Type 3 (“late onset”) is mainly associated with a significant cooling on the continental Guinean belt and with a large pressure increase in the equatorial Atlantic between  $10^{\circ}\text{N}$  and  $10^{\circ}\text{S}$ .

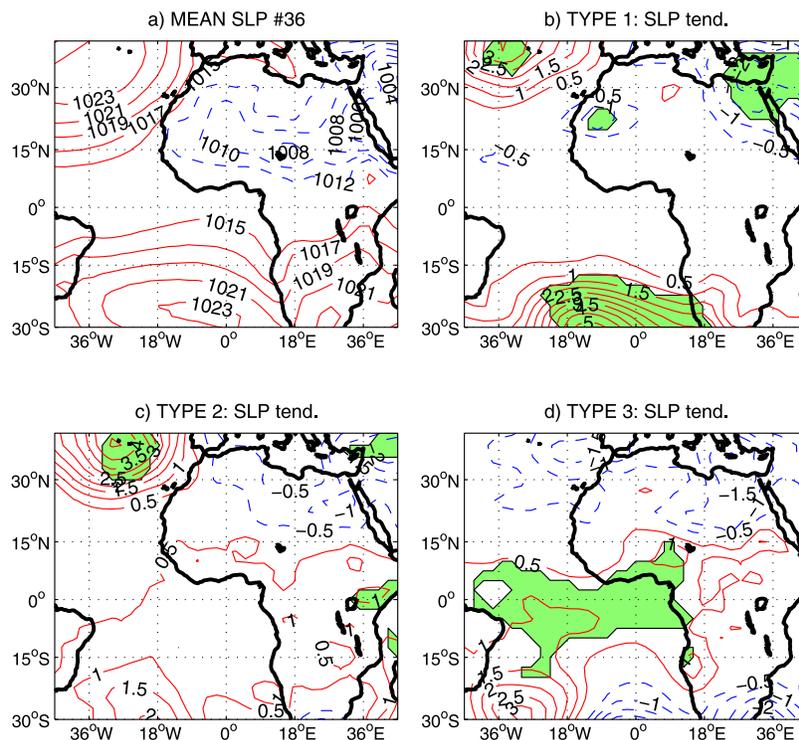
[21] These results show that at the moment of onset, robust signals of regional extent can be captured in the SKT and SLP meridional and zonal gradients. Such gradients are well known to drive monsoon energetics and dynamics, either directly via changes in low-level energy contents and pressure gradient force (and hence horizontal wind field) or indirectly via their new vertical arrangement.

[22] To focus on the last point, Figures 6–8 present similar diagrams but for the wind fields at 925 hPa (the monsoon flow), 600 hPa (the African Easterly Jet level,

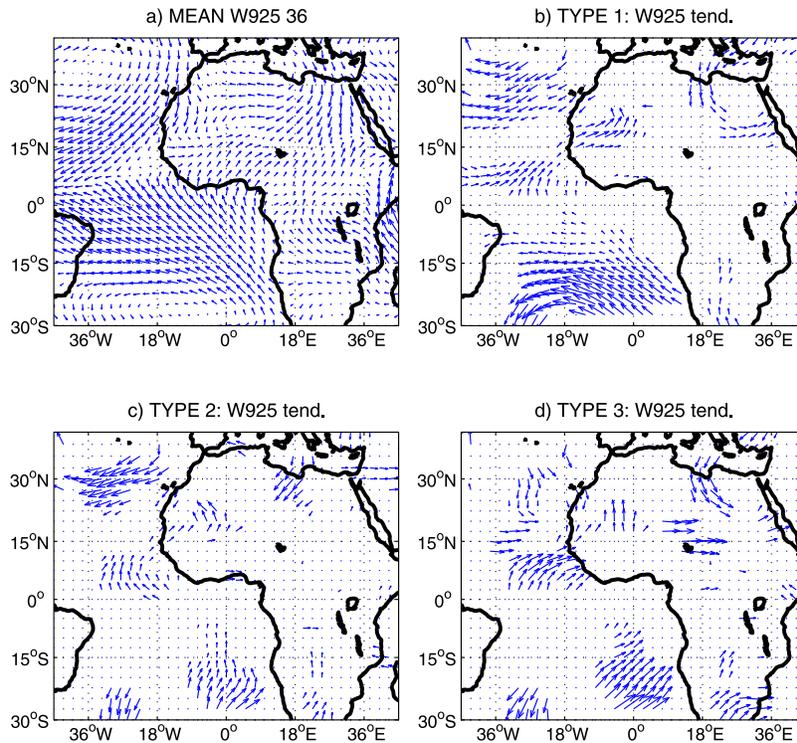
AEJ), and 200 hPa (the Tropical Easterly Jet, TEJ), respectively. Figures 6a, 7a, and 8a refer to the mean conditions before (after) the mean onset date, whereas Figures 6b–6d, Figures 7b–7d, and Figures 8b–8d display the significant wind field tendencies at the moment of the “normal,” “early,” and “late” onsets (i.e., the “after” minus “before” differences regarding the dates in Table 2). For the three types, the onset period shows a large enhancement of the mean circulation in low levels (Figures 6b–6d). This can be observed in the vicinity of both the Santa Helena cell (northwestward vectors in coherency with SLP increasing already pointed out), and of the Azores cell (southerlies and southwesterlies). Between these anticyclonic gyres the monsoon flux reinforces as noticed by the westerly and southerly anomalies over the tropical and the western part of the continent, in coherency with the SLP decrease already noticed over the Saharan heat low region (see Figure 4). The reinforcement of the zonal component (westerly winds) in low levels within the latitudinal band  $5^{\circ}\text{N}$ – $15^{\circ}\text{N}$  and near the African coast ( $30^{\circ}\text{W}$ – $10^{\circ}\text{W}$ ) is probably linked to a reinforcement of the westerly wind jet at the moment of the monsoon onset. The westerly jet installs over this region from May to September and brings humid maritime air and thus rainfall inland [Grodsky *et al.*, 2003]. In midlevels, years of type 1 (“normal” onset) show a large enhancement of easterlies and southeasterlies from the eastern Mediterranean basin to the Sahel and the western Sahara along with



**Figure 4.** (a) Mean skin temperature (SKT in Celsius) by the end of June (pentad #36); dashed lines for values lower than 30°C. (b–d) Mean tendencies at the dates of “normal,” “early,” and “late” onsets, respectively. Values are the mean 5-day SKT differences between the pentad before and after the date of onset observed each year, where positive/negative values mean warming/cooling. Shading for the significant differences is at  $p = 0.05$  regarding a Student  $t$  test.



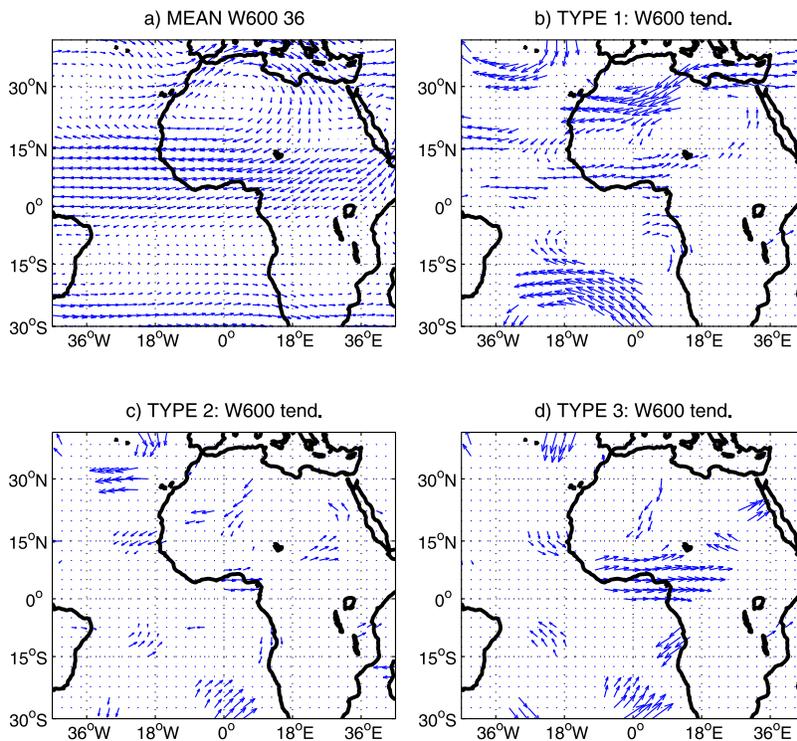
**Figure 5.** Same as Figure 4 except for sea level pressure (SLP). Dashed lines are for values lower than 1015 hPa.



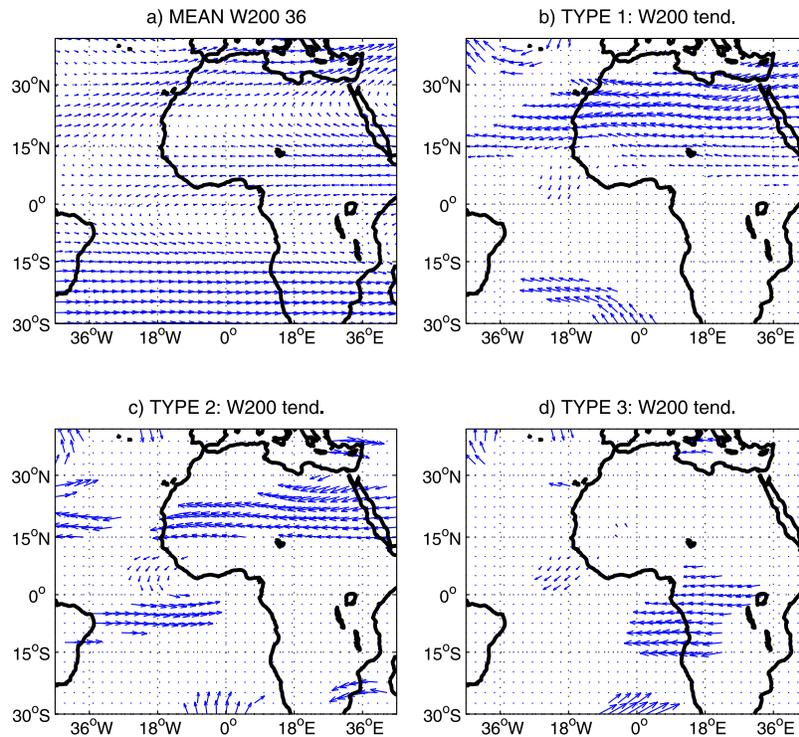
**Figure 6.** (a–d) Same as Figure 4 except for the wind field at 925 hPa (W925).

a decrease in AEJ attested by a westerly tendency northward to the Guinean coast (Figure 7b). Type 2 (“early”) and especially type 3 (“late”) exhibit also this tendency but displaced somewhat southward as shown northward to the equator in Figures 7c and 7d. This is coherent with a

northward migration of the AEJ axis. In the upper troposphere a large area is concerned by a westward tendency extending north of 15°N in “normal” and “early” years (types 1 and 2 in Figures 8b and 8c) or southward to the equator for the “late” years (type 3 in Figure 8d). Such



**Figure 7.** (a–d) Same as Figure 4 except for the wind field at 600 hPa (W600).



**Figure 8.** (a–d) Same as Figure 4 except for the wind field at 200 hPa (W200).

patterns reinforce the mean easterly circulation between the tropics (Figure 8a). In all composites northward anomalies are also always present in the southern Atlantic. In Figure 8c, notice the southward anomaly winds over the equatorial Atlantic offshore of Senegal linked to large westward anomalies over the continent.

[23] All these results (Figures 2–8) show that robust signals are linked to the dates of onset. They are either of subregional (i.e., rainfall evolution at the Guinean latitudes), or of regional scale (i.e., pressure and energy gradients, temperature, and wind fields). A statistical prediction of onset dates based on such signals is hence feasible. The following section briefly illustrates this issue.

## 5. Experimental Hindcasts of the Dates of Monsoon Onset

### 5.1. Definition of Experiments

[24] Four types of short-term cross-validated experimental hindcasts of onset dates have been performed using multivariate linear regression analyses with objective selection of different sets of potential predictors taken from mid-May to mid-June (i.e., six pentads from the 28th to the 33rd pentads).

[25] 1. In a first experiment the potential predictors are defined in a physical sense to describe monsoon energetics and dynamics. They take directly into account the rough (unfiltered) meridional and zonal gradients of sea level pressure (SLP) and moist static energy (MSE) at 1000 hPa, because this information is physically linked to the monsoon and has been successfully used for prediction of the Sahel rainy season [Fontaine et al., 1999; Philippon and Fontaine, 2002]. These gradients refer to different spatial scales and areas over the domain (30°N–Eq.; 20°W–10°E).

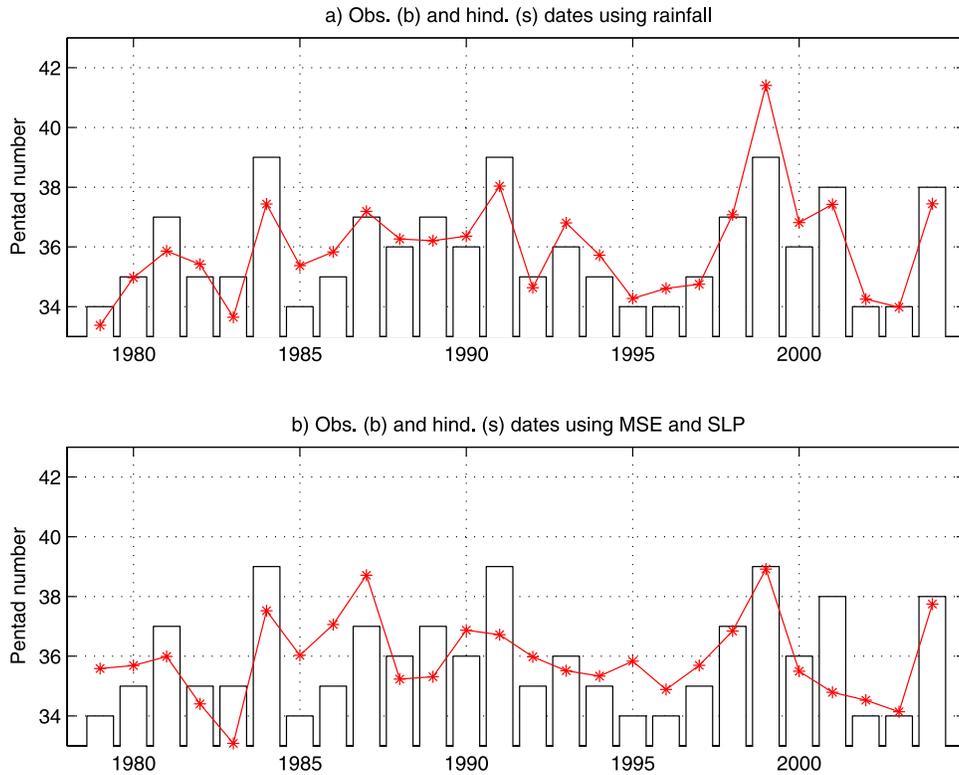
[26] 2. In a second experiment we use field descriptors through the rough (unfiltered) NCEP2/DOE reanalyzed atmospheric data (SLP, temperature, winds at 925, 600, and 200 hPa) since atmospheric variability appeared to be significantly linked to the onset in Figures 4–8. These modes, defined in a pure statistical sense (i.e., the 10 first leading EOF of each field with and without Varimax rotation), are hence of regional extent and refer to the same domain.

[27] 1. In experiment 3 we focus on the sole rainfall signals of smaller scale discussed in the above section (Figures 2 and 3), using the rough (unfiltered) 5-day CMAP rainfall information southward to the Sahara (between 20°N and 5°S). Three adjacent latitudinal belts of 5° large are defined over land between 20°N and 5°N in the 10°W–10°E window for documenting the Sahelian, Sudanian and Guinean belts, plus two others in the Gulf of Guinea (GoG) for describing its northern (N.GoG: equator–5°N in the 10°W–10°E window) and southern (S.GoG: 5°N–5°S in the 15°W–10°E window) parts. So, we only consider these five rainfall zonal indexes (three over land and two over the ocean), along with the four meridional differences between the five adjacent areas to detail rainfall gradients.

[28] 2. The fourth experiment consists of merging the potential predictors used in all experiments (rainfall and atmospheric descriptors).

### 5.2. Results

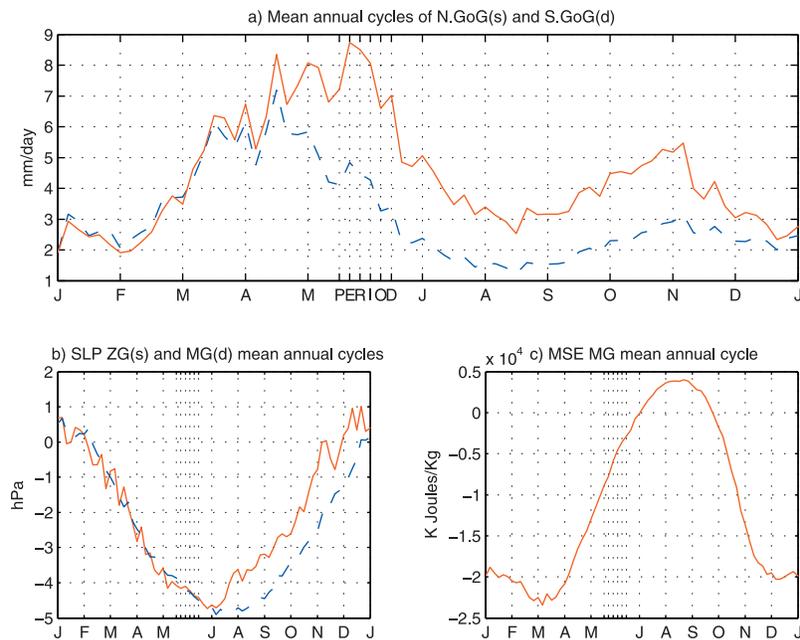
[29] If hindcasts based on atmospheric predictors (experiments 1 and 2) from NCEP/DOE2 are encouraging, those using “observed” rainfall estimates over a more limited domain (experiments 3 and 4) are significantly better. The following list summarizes briefly the main results obtained in terms of common variance between observed and hind-



**Figure 9.** Experimental hindcasts of the dates of monsoon onset through multivariate linear regression analyses using four predictors in cross-validation mode: (a) with rainfall predictors and (b) with SLP and MSE predictors.

casted time series ( $R^2$  coefficient) and of statistical significance (one, two, and three asterisks denote the significant levels at  $p = 0.1, 0.05,$  and  $0.001,$  respectively, after 1000 Monte Carlo simulations). More details will be given after.

[30] 1. In experiment 1, the SLP and MSE gradient predictors explain 36%\* of the common variance:  $r = +0.60$ ) in good coherency with the physical approach of the monsoon and the statistical schemes used by *Fontaine et*



**Figure 10.** Mean annual cycles of selected predictors: (a) rainfall, (b) SLP zonal gradients, and (c) MSE meridional gradients. The period mid-May to mid-June and the mean onset date are marked “PERIOD” and “J,” respectively.

**Table 4.** Results From Experiment 3<sup>a</sup>

With Rainfall EV: 75% ***	S.GoG #28	S.GoG #30	S.GoG #32	GoG_MG #33
Reg. C	+0.41	+0.40	-0.43	-0.64
C(l)Tot CC	+0.53***	-0.34*	-0.44*	-0.58***

<sup>a</sup>S.GoG #28, S.GoG #30, S.GoG #32, and GoG\_MG #33 are selected predictors. Reg. C values are the regression coefficients for normalized series. C(l)Tot CC values are the total correlation coefficients. Basic statistics are as follows (see Figure 10a): AFA, 0.66\*\*\*; AA, 0.05\*\*\*; RMSE, 0.87\*\*\*; RMSE-A, 1.01\*\*\*; ABSE, 0.68\*\*\*; BIAS, -0.04. EV, explained variance; AFA, actual forecast ability; AA, artificial ability; RMSE and RMSE-A, root mean square (normal and adjusted); ABSE, absolute error and BIAS following *Ward et al.* [1990], *Ward and Folland* [1991], and *Knaff and Landsea* [1997]. Levels of significance at  $p = 0.01$ , 0.05, and 0.01 are marked by one, two, and three asterisks, respectively. They are estimated by taking into account the number of independent values in the series for CCs and, for the basic statistics, by performing 1000 Monte Carlo simulations.

*al.* [1999] and *Philippon and Fontaine* [2002] for predicting the Sahelian rainy season. More details will be given hereinafter.

[31] 2. In experiment 2, the 10 leading EOF modes of the basic atmospheric fields shown in Figures 4–8 (SLP, temperature and wind components at 925, 600, and 200 hPa) enable one to obtain better skills with four predictors:  $R^2 = 31\%^*$  with SLP (i.e.,  $r = +0.57$ ),  $46\%^{**}$  with winds at 600 hPa,  $50\%^{**}$  with skin temperature,  $51\%^{**}$  with winds at 200hPa, and  $R^2 = 65\%^{***}$  with the 925 hPa wind field ( $r = +0.81$ ). When all atmospheric field descriptors (SLP, temperature and winds) are considered, the skill can reach  $73\%^{***}$ , and  $74\%^{***}$  when onset dates are defined through GPCP. Better scores are obtained when EOF modes are defined over larger windows. This demonstrates the good predictability of the dates of onset with reanalyzed atmospheric products.

[32] 3. In experiment 3, the potential predictors are the sole CMAP rainfall indexes averaged over a few small areas of interest (i.e., the subregional indexes describing rainfall amounts and meridional gradients southward to the Sahara already presented). Predictors and results will be detailed hereafter. In this experiment the cross-validated predictive skill explains  $75\%^{***}$  of common variance ( $r = +0.87$ ). More details will be given hereinafter.

[33] 4. This last score can be slightly improved ( $R^2 = 82\%^{***}$ ) by considering both rainfall and atmospheric signals (experiment 4). Here the procedure automatically selects 3 rainfall predictors over the Gulf of Guinea explaining 77% of the relationship. The fourth predictor (23% of the relationship) is a low energetic zonal wind mode of variability in low levels having stronger weights on the WAM domain and on the equatorial and southern tropical Atlantic.

[34] Figures 9 and 10 and Tables 4 and 5 describe the cross-validated hindcasted results (basic statistics, selected predictors) issued from experiments 3 and 1. These two experiments are chosen because they allow us to better summarize and discuss the overall results. Experiment 3 (Figure 9a and Table 4) is not the best one in a pure statistical sense since it explains only 36% of. However, it is surely the most potentially useful since it uses only one “observed” data set over a limited area (CMAP rainfall estimates in the Gulf of Guinea). By contrast, experiment 1

(Figure 9b and Table 5) is less powerful in a statistical sense but its predictors (SLP and MSE gradients) are directly linked to monsoon energy and dynamics and therefore more helpful to discuss the results in physical/dynamical terms.

[35] Table 4 confirms the accurate predictive skill obtained with the sole observed rainfall variability over the Gulf of Guinea, as no continental rainfall predictor is retained by the selection procedure: the retained predictors (Table 4) are the 5-day rainfall amounts during the 28th, 30th, and 32nd pentads over southern GoG (S.GoG:  $15^\circ\text{W} - 10^\circ\text{E}$ ;  $5^\circ\text{S} - 5^\circ\text{N}$ ) and the transequatorial meridional gradient (GoG\_MG) during the 33rd pentad. This could be owing to the fact that, CMAP uses both SSM/I microwave scattering and microwave sounding unit over the ocean [*Gruber et al.*, 2000].

[36] The selected rainfall predictors are not significantly cross-correlated, the coefficients ranging from  $-0.15$  to  $+0.27$  with a variance inflation factor (VIF) lower than the maximum threshold of 5 recommended by *Chatterjee and Price* [1977]:  $VIF = 1.79$ , meaning that the four predictors are not collinear. Lines 2 and 3 in Table 4 list the respective regression coefficients in conjunction with the total correlations between the dates of onset and the four predictors. Sign reverses between pentads point to the role of rainfall time tendencies. For example, the total correlations show that an onset in late (advance) is significantly associated with decreasing rainfall over the equator (positive/negative values with S.GoG for pentad numbers 28/30). Coefficients of regression demonstrate that a too late onset is typically associated with (1) more rainfall by mid-May over the equatorial Atlantic (positive coefficient in pentad #28 with S.GoG), (2) a rapid and persistent decrease over the same region (negative coefficients after), and (3) a clear southward cross-equatorial gradient anomaly by mid-June over the Gulf of Guinea (negative value with GoG\_MG during the 33rd pentad). As shown by statistics (last line of Table 4), the cross-validated forecast skill is good and significant (i.e.,  $75\%^{***}$  of explained variance:  $r = +0.87$ ). Figure 9a enables comparison between the observed and predicted dates of onset as bars and curve, respectively. One can notice the close relationship between the two time series, except perhaps in 1999 when the predicted date lags by two pentads the observed onset (pentad 41 versus 39).

[37] It is noteworthy that the above skill is largely higher than the score ( $R^2 = 45\%^{**}$ ) obtained with rainfall field predictors of regional scale resulting from EOF analyses without rotation over the domain  $30^\circ\text{S} - 40^\circ\text{N}$ ;  $45^\circ\text{W} - 45^\circ\text{E}$ . By contrast, when rainfall field predictors are defined over a

**Table 5.** Same as Table 4 but for Experiment 1<sup>a</sup>

With MSE + SLP EV: 36% ***	SLP_MG #31	SLP_ZG #31	SLP_ZG #33	MSE_MG #33
Reg. C	+0.34	+0.35	-0.29	-0.33
Tot. CC	+0.40**	+0.37*	-0.21	-0.53***

<sup>a</sup>SLP\_MG stands for the SLP meridional gradients across the equator between ( $0 - 20^\circ\text{N}$ ) and ( $0 - 20^\circ\text{S}$ ) in the  $10^\circ\text{W} - 10^\circ\text{E}$  window. SLP\_ZG stands for the SLP zonal gradients in the northern tropics ( $20^\circ\text{N} - 30^\circ\text{N}$ ) between the eastern Atlantic ( $20^\circ\text{W} - 10^\circ\text{W}$ ) and the western Sahara ( $10^\circ\text{W} - 0^\circ$ ). MSE\_MG stands for the MSE meridional gradients between the ( $10^\circ - 15^\circ\text{N}$ ) and ( $5^\circ\text{N} - 10^\circ\text{N}$ ) latitudinal belts in the  $10^\circ\text{W} - 10^\circ\text{E}$  window. Basic statistics are as follows: AFA, 0.13\*; AA, 0.12\*; RMSE, 1.36; RMSE-A, 1.58; ABSE, 1.11; BIAS,  $\sim 0.00$ .

more limited WAM region ( $0\text{--}20^\circ\text{N}$ ;  $10^\circ\text{W}\text{--}10^\circ\text{E}$ ) the hindcasted results are improved ( $R^2 = 65\%^{***}$ ) and  $73\%^{***}$  if a Varimax rotation is performed to give more weight to signals of lower scale. This proves that special attention must be paid not only to the type of variable but also to the spatial scale of descriptors and to the way they are defined. Unfortunately, the predictive capability of CMAP rainfall decreases rapidly when the lead time increases:  $56\%^{**}$  of explained variance with predictors defined from April 25 to May 25, for example.

[38] Let us now focus on the mean annual cycles of the selected rainfall predictors in experiment 3 (i.e., the northern and southern parts of the Gulf of Guinea, Figure 10a). The period mid-May to mid-June and the mean onset date are marked “PERIOD” and “J,” respectively. Clearly, the mean date of onset (end of June, just before the letter J on the  $x$  axis) follows a rainfall maximum over N.GoG ( $8.5\text{ mm/day}$  on the solid line in Figure 10a) and a concomitant rainfall decreasing over the northern and southern (dashed line) parts of the Gulf of Guinea. This generates a persistent northward gradient (precipitation is higher over N.GoG than over S.GoG). So, any southward gradient anomaly (when rainfall amounts over S.GoG exceed those registered over N.GoG) will tend to reduce the mean northward gradient displayed in Figure 10a. Such a situation will favor an abnormal southward position of the rain belt and therefore will tend to delay the date of onset (see the negative coefficients between the date and GoG\_MG in Table 4).

[39] By comparison hindcasts based on SLP and MSE gradients known to directly impact monsoon dynamics show clearly lower skill (Figures 9b and Table 5). However, scores are significant and the selected predictors (Figures 10b and 10c) are not colinear ( $\text{VIF} = 1.16$ ) and enable physical interpretation. The onset date is concomitant of an extremum ( $\sim -4.5\text{ hPa}$  in Figure 10b) in SLP differences along the meridional and zonal directions at regional scale: between  $[0\text{--}20^\circ\text{N}]$  and  $[0\text{--}20^\circ\text{S}]$  in the  $10^\circ\text{W}\text{--}10^\circ\text{E}$  window; between the eastern Atlantic  $[20^\circ\text{W}\text{--}10^\circ\text{W}]$  and the western Sahara  $[10^\circ\text{W}\text{--}0^\circ]$  in the northern tropics  $[20^\circ\text{N}\text{--}30^\circ\text{N}]$ . These negative gradients directed toward south and east increase therefore the pressure gradient force in the opposite directions, driving horizontal wind toward the heat low region and hence enhancing horizontal convergence: the monsoon onset occurs when the northward and westward pressure gradient forces register their annual maximum at regional scale. Since SLP gradients are defined positive toward the north or east, a too late onset is linked to (1) northward anomaly gradients across the equator (see the positive coefficients with SLP\_MG in Table 5) and (2) increasing eastward anomaly gradients in the northern tropics between the western Sahara and the eastern Atlantic (see the reverse in sign with SLP\_ZG in Table 5). This tends to slow down the seasonal increase in SLP gradients (negative values shown in Figure 4b) and hence to delay their annual extremums.

[40] In terms of monsoon energetics, the onset date is related to a sign reverse in MSE meridional gradients between the  $(10^\circ\text{--}15^\circ\text{N})$  and  $(5^\circ\text{N}\text{--}10^\circ\text{N})$  latitudinal belts, as shown by the zero value in Figure 10c: the gradients are negative before the onset (directed southward) and positive after (directed northward). So, any negative MSE anomaly

gradients on the continent will tend to delay the reverse in MSE meridional gradients and hence the date of onset.

## 6. Discussion and Conclusion

[41] The objective of this article was to propose a preliminary statistical study of the West African monsoon onset in terms of diagnostics, predictability and short-term prediction. A four-step approach has been proposed: (1) definition of a WAM onset index (WAMOI) using both the 5-day CMAP and GPCP data; (2) application of clustering algorithm (K-means) on April–July WAMOI series to detail the main types of spring to summer evolutions; (3) analyses of the most robust leading signals in the CMAP estimates and NCEP2/DOE fields (sea level pressure, skin temperature, wind at 925, 600, and 200 hPa, moist static energy) to select potential rainfall and atmospheric predictors; and (4) production of experimental cross-validated hindcasted onset dates using objective selections of these predictors.

[42] The CMAP and GPCP rainfall estimates allowed us to define a mean date of onset over the period 1979–2004 by June 28–29 with a standard deviation of 8.5 days and to show that onset dates are not accurate at a daily scale. Classifications (K-mean clusterings) well separate years with a “normal” onset date (36th pentad, 50% of cases) from those registering an “early” (mid-June, 31%) or a “late” (before mid-July, 19%) onset. These three types of situation are also different regarding precipitation: basically the “normal” and “late” onsets correspond to a large northward excursion of the rainband associated with significant increase over Sudan-Sahel and decrease in the Gulf of Guinea; an “early” onset shows does not show any important Sahelian increase in the  $10^\circ\text{W}\text{--}10^\circ\text{E}$  window. So, over the 26-year total period the most accurate signals in precipitation linked to the onset period are located just southward to the Guinea coast. They are hence of lower extent than the regional scale (i.e., subregional).

[43] Other associated and significant signals can be defined using the basic atmospheric fields (sea level pressure, skin temperature, winds at 925, 600, and 200 hPa, moist static energy). The onset period is associated with a large enhancement of the mean circulation in low levels, mainly in the vicinity of both the Santa-Helena cell and the Azores cell, and between these gyres a reinforcement of the low-level westerly wind jet and thus of the monsoon flux over the tropical Atlantic and the western part of West Africa. In midlevels, easterlies from the eastern Mediterranean basin to the Sahel reinforce, but the African Easterly Jet (AEJ) lessens over the Guinean coast, which is coherent with a northward migration of the AEJ axis. In the upper troposphere, the mean easterly circulation between the tropics is reinforced. All these signals are in general agreement with an abrupt intensification and northward shift of the monsoon rains at time of onset. For example, any stronger westerly jet in low levels will favor the development of easterly waves by enhancing zonal wind shear between low levels and midlevels [Grodky *et al.*, 2003] and will cause enhanced marine air transport into the continent [Grist and Nicholson, 2001].

[44] The monsoon onset occurs when the northward cross-equatorial rainfall gradient is at maximum (i.e., in-

creasing rainfall differences between the oceanic regions north and south of equator). It is also concomitant to (1) a maximum in sea level pressure differences along both the meridional and zonal directions: the onset takes place when the northward and westward pressure gradient forces at regional scale exceed a certain threshold; (2) a sign reversing in moist static energy gradient between the (10°–15°N) and (5°N–10°N) latitudinal belts in the 10°W–10°E window. These gradients are directed southward before the onset but northward after.

[45] These results (and others not reported here) showed that statistical predictions of onset dates based on some rainfall and/or atmospheric signals (i.e., SLP, temperature, MSE, and wind field) selected on the period mid-May/mid-June are feasible or possible. All of them demonstrate the potential ability of reanalyzed atmospheric data (period mid-May to mid-June) to predict the onset dates. In this study the CMAP rainfall gradients and time tendencies defined over the equatorial Atlantic in the Gulf of Guinea appeared also very efficient. In particular, skills issued from cross-validated hindcasts with four rainfall predictors are highly significant (75% of explained common variance with observations). Basically, late (early) onsets are preceded by more (less) rainfall southward to the equator by the end of May, then by a clear decrease (increase) of the normal northward cross-equatorial gradient by mid-June. This skill can be slightly improved ( $R^2 = 82\%$ ) by merging this rainfall information with low-level zonal wind variability. However, in this context rainfall descriptors over the Gulf of Guinea remain largely dominant in the regression (three predictors out of four, explaining 77% of the relationship), certainly because these indicators have been defined from mid-May to mid-June and over small adjacent areas of subregional scale. Other experiments not reported here have been made for verifying the importance of CMAP rainfall versus NCEP/DOE2 atmospheric descriptors: in general, they explain more than 2/3 of the total variance of the prediction. In particular, the best schemes have been obtained with spatial rainfall indexes averaged over a few small areas of interest or through rotated EOFs defined over the sole WAM domain [0°–20°N; 10°W–10°E]. This proves that special attention must be paid to the spatial scale of descriptors and to the way they are defined. Unfortunately, the predictive capability of CMAP rainfall decreases rapidly when the lead time increases: 56% of explained variance with predictors defined from April 25 to May 25, for example.

[46] The first objective is now to select operational precipitation data for better describing the “real world” and detailing anomalies associated with the three types of onsets. The second one is to increase the lead time of prediction for the AMMA program by improving our statistical schemes with more specific atmospheric WAM and surface multiscale descriptors, like those already available on the link AMMA-WP1.1.3 at <http://www.u-bourgogne.fr/climatologie/>. The third objective is to assess the importance of spatial scales for the predictors and to evaluate the predictability of rainfall variability at intraseasonal time-scale, including breaks and surges of the monsoon as defined by Louvet *et al.* [2003]. In this study the WAM onset index refers to the 10°W–10°E window since this is the region where the abrupt shift of the ITCZ is the most obvious [Sultan and Janicot, 2003]: an interesting research

issue will be to investigate other WAM longitudes (i.e., 15°W–5°W, 5°W–5°E, and 5°E–15°E). The last purpose will be to investigate the capability of atmospheric predictors defined in GCMs to evaluate the new ECMWF numerical forecasts for the time range 10–30 days in terms of monsoon onset, in order to propose new schemes based on both observed (or reanalyzed) historical data and numerical simulations.

[47] **Acknowledgments.** The authors are very grateful to the National Center for Atmospheric Research (USA) for providing the NCEP/DOE2 data and to the anonymous reviewers for their interesting comments. 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). On the basis of a French initiative, AMMA was built by an international scientific group and is currently funded by a large number of agencies, especially from France, the United Kingdom, the United States, 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., *et al.* (2003), The version 2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–present), *J. Hydro-meteorol.*, *4*, 1147–1167.
- Ali, A., A. Amani, and T. Lebel (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., A. Amani, A. Diedhiou, and T. Lebel (2005), Rainfall estimation in the Sahel, part II, Evaluation of rain gauge networks in the CILSS countries and objective intercomparison of rainfall products, *J. Appl. Meteorol.*, *44*(11), 1707–1722.
- Ati, O. F., C. J. Stigter, and E. O. Oladipo (2002), A comparison of methods to determine the onset of the growing season in northern Nigeria, *Int. J. Climatol.*, *22*(6), 731–742.
- Barnston, A. G. (1994), Linear statistical short-term climate predictive skill in the Northern Hemisphere, *J. Clim.*, *7*, 1513–1564.
- Chatterjee, S., and B. Price (1977), *Regression Analysis by Example*, John Wiley, Hoboken, N. J.
- Fontaine, B., N. Philippon, and P. Camberlin (1999), An improvement of June–September rainfall forecasting in the Sahel based upon region April–May moist static energy content (1968–1997), *Geophys. Res. Lett.*, *26*, 2041–2044.
- Grist, J. P., and S. E. Nicholson (2001), A study of the dynamics factors influencing the rainfall variability in the West African Sahel, *J. Clim.*, *14*, 1337–1359.
- Grodsky, S. A., and J. A. Carton (2001), Coupled land/atmosphere interactions in the West African Monsoon, *Geophys. Res. Lett.*, *28*, 1503–1506.
- Grodsky, S. A., J. A. Carton, and S. Nigam (2003), Near surface westerly wind jet in the Atlantic ITCZ, *Geophys. Res. Lett.*, *30*(19), 2009, doi:10.1029/2003GL017867.
- Gruber, A., X. Su, M. Kanamitsu, and J. Schemm (2000), The comparison of two merger rain gauge–satellite precipitation datasets, *Bull. Am. Meteorol. Soc.*, *81*(11), 2631–2643.
- Gu, G., and R. F. Adler (2004), Seasonal evolution and variability associated with the West African monsoon system, *J. Clim.*, *17*, 3364–3377.
- Janicot, S., and B. Sultan (2001), Intra-seasonal modulation of convection in the West African monsoon, *Geophys. Res. Lett.*, *28*, 523–526.
- Kanamitsu, M., W. Ebisuzaki, J. Woollen, S.-K. Yarg, J. J. Hnilo, M. Fiorino, and G. L. Potter (2002), NCEP-DOE AMIP-II reanalysis (R-2), *Bull. Am. Meteorol. Soc.*, *83*, 1631–1643.
- Knaff, J. A., and C. W. Landsea (1997), An El Niño–Southern Oscillation climatology and persistence (CLIPER) forecasting scheme, *Weather Forecasting*, *12*, 633–652.
- 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), *Geophys. Res. Lett.*, *30*(24), 2271, doi:10.1029/2003GL018058.
- Michaelsen, J. (1987), Cross-validation in statistical climate forecast models, *J. Clim. Appl. Meteorol.*, *26*, 1589–1600.
- Moron, V., S. Bigot, and P. Roucou (1995), Rainfall variability in subequatorial America and Africa and relationships with the main Sea-Surface Temperature modes (1951–1990), *Int. J. Climatol.*, *15*, 1297–1322.

- Murakami, M. (1979), Large scale aspects of deep convective activity over the Gate area, *Mon. Weather Rev.*, 107, 994–1013.
- Omotosho, J. B., A. A. Balogun, and K. Ogunjobi (2000), Predicting monthly and seasonal rainfall, onset and cessation of the rainy season in the West Africa using only surface data, *Int. J. Climatol.*, 20, 865–880.
- Philippon, N., and B. Fontaine (2002), The relationship between the Sahelian and previous 2nd Guinean rainy seasons: A monsoon regulation by soil wetness?, *Ann. Geophys.*, 20(4), 575–582.
- Polcher, J., (Ed.) (2004), African Monsoon Multidisciplinary Analysis, 350 pp., Cent. Natl. de la Rech. Sci., Paris.
- Sultan, B., and S. Janicot (2000), Abrupt shift of the ITCZ over West Africa and intra-seasonal variability, *Geophys. Res. Lett.*, 27, 3353–3356.
- Sultan, B., and S. Janicot (2003), The West African Monsoon dynamics, part II, The “Preonset” and the “Onset” of the summer Monsoon, *J. Clim.*, 16, 3407–3427.
- Sultan, B., S. Janicot, and A. Diedhiou (2003), The West African Monsoon dynamics, part I, Documentation of intra-seasonal variability, *J. Clim.*, 16(21), 3389–3406.
- Ward, M. N., and C. K. Folland (1991), Prediction of seasonal rainfall in the north Nordeste of Brazil using eigenvectors of sea-surface temperature, *Int. J. Climatol.*, 11, 711–743.
- Ward, M. N., C. K. Folland, and K. Maskell (1990), Understanding and predicting seasonal rainfall in subSaharan Africa, *Trop. Meteorol. Res. Program Rep. Ser.*, 36, 157–161.
- Ward, N., et al. (2004), Seasonal-to-Decadal Predictability and Prediction of West African Climate, *CLIVAR Exchanges*, 9(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. Am. Meteorol. Soc.*, 78, 2539–2558.
- Xie, P., J. E. Janowiak, P. A. Arkin, R. F. Adler, A. Gruber, R. Ferraro, G. J. Huffman, and S. Curtis (2003), GPCP pentad precipitation analyses: An experimental dataset based on gauge observations and satellite estimates, *J. Clim.*, 16, 2197–2214.
- Yin, X., A. Gruber, and P. A. Arkin (2004), Comparison of the GPCP and CMAP merged gauge satellite monthly precipitation products for the period 1979–2001, *J. Hydrometeorol.*, 5, 1207–1222.

---

B. Fontaine and S. Louvet, Centre de Recherches de Climatologie, CNRS, Faculte de Sciences Gabriel, University of Burgundy, 6 Boulevard Gabriel, F21004 Dijon, France. (fontaine@u-bourgogne.fr)