

Components of rainy seasons' variability in Equatorial East Africa: onset, cessation, rainfall frequency and intensity

Pierre Camberlin · Vincent Moron · Raphael Okoola ·
Nathalie Philippon · Wilson Gitau

Received: 28 August 2008 / Accepted: 30 January 2009 / Published online: 18 February 2009
© Springer-Verlag 2009

Abstract The inter-annual and spatial variability of different rainfall variables is analysed over Equatorial East Africa (Kenya and northeastern Tanzania). At the station level, three variables are considered: the total precipitation amount (P), the number of rain days (NRD) and the daily rainfall intensity (INT). Using a network of 34 stations, inter-station correlations (1958–1987) are computed for each of these variables. The spatial coherence of monthly or seasonal P and NRD is always much higher than that of rainfall intensity. However, large variations in spatial coherence are found in the course of the seasonal cycle. Coherence is highest at the peak of the short rains (October–December)

and low during the long rains (March–May), except at its beginning. The inter-annual variability of the onset and cessation of the rains is next considered, at the regional scale, and the study extended to 2001. In the long rains, the onset and cessation dates are independent of NRD and INT during the rainy season. Hence, the long rains seasonal rainfall total depends on a combination of virtually unrelated factors, which may account for the difficulty in its prediction. However, the onset, which exhibits a large inter-annual variability and a strong spatial coherence, has a prime role. Conversely, in the short rains, though the onset is again more decisive than the cessation, the different intra-seasonal descriptors of the rains are more strongly inter-related.

P. Camberlin (✉) · N. Philippon · W. Gitau
Centre de Recherches de Climatologie (CRC),
Université de Bourgogne/CNRS,
Dijon, France
e-mail: camber@u-bourgogne.fr

V. Moron
CEREGE, Université d'Aix-Marseille I,
Aix-en-Provence, France

V. Moron
International Research Institute for Climate and Society (IRI),
Columbia University,
Palisades, NY, USA

V. Moron
Institut Universitaire de France (IUP),
Paris, France

R. Okoola · W. Gitau
Department of Meteorology, University of Nairobi,
Nairobi, Kenya

W. Gitau
IGAD Climate Prediction and Applications Centre (ICPAC),
Nairobi, Kenya

1 Introduction

The performance of a given rainy season, for optimal crop growth, does not only lie in the overall total amount, but requires an adequate distribution of the rains during the year. This is particularly so in the semi-arid and sub-humid zones of tropical Africa where irrigation is underdeveloped and the rains fall within a limited period of time. The causes of the rains' failure range from a delayed onset of the rains, an early withdrawal, or short but intense rainfall events separated by long dry spells. Though in the last three decades much research has been devoted to the inter-annual variations of seasonal rainfall in the Tropics, it is still not well-known what actually makes a high seasonal rainfall total, in terms of the distribution of the rains. Is it simply more rainfall events? A greater efficiency of each event in terms of rain amount per rain day? A longer rainy season? This knowledge is of notable concern when elaborating seasonal rainfall predictions. In fact, the processes which result into a smaller number of rainfall events are unlikely

to be the same as those which drive a delayed onset of the rains or smaller amounts per rain day.

In the West African Sahel, rainfall variability was found to be linked to the variations in the number of events rather than in the magnitude of these events (D'Amato and Lebel 1998). In West Africa as a whole, only the reduction in the number of rain events shows a pattern consistent with that of the drought initiated in the 1970s (Le Barbé et al. 2002). This drought was also found to result in a shorter rainy season, as shown in Senegal (Camberlin and Diop 2003) and in northern Nigeria (Olaniran and Sumner 1990; Oladipo and Kyari 1993). However, the evidence is less clear in some countries like Mali (Traoré et al. 2000). Such discrepancies could partly be due to the methods used to define onset and cessation, since at station level, there is inherent “noise” in the distribution of daily rainfall. The parts played by the rainy season onset and retreat variations are also not necessarily symmetrical. In Nigeria, Oladipo and Kyari (1993) found the inter-annual variations of the growing season to be more dependent on the starting date than on the cessation date of the rains. Studies on India summer monsoon rainfall indicate that there is little relationship between the seasonal rainfall accumulation and the date of monsoon onset in southern India (Dhar et al. 1980; Nayagam et al. 2008). Similarly, Marengo et al. (2001) found an unexpectedly weak relationship between the length of the rainy season and seasonal precipitation amount for the Brazilian Amazon Basin. Liebmann and Marengo (2001), based on a slightly different method, found higher correlations and demonstrated that the sea surface temperature (SST) influence on Amazon rainfall was through the onset rather than the within-season rain rate. A study of five tropical regions (Moron et al. 2007) showed that inter-annual variations in the seasonal number of rain days (NRD) tend to be slightly more coherent in space than those of the seasonal rainfall amounts. Mean intensities per rain day are not consistent from station to station at inter-annual time scale. This means that, at the regional scale, the inter-annual variability of seasonal rainfall totals depends mostly on the frequency of rainfall events and, with a few exceptions, not at all on their “productivity”.

The case of East Africa is of particular interest since much of the region exhibits two rainy seasons, whose characteristics are partly dissimilar. The one occurring during the boreal autumn, the short rains, shows large inter-annual fluctuations, strongly related to large-scale circulation anomalies in the Indian and Pacific oceans (Hastenrath et al. 1993; Mutai and Ward 2000; Philippon et al. 2002; Black et al. 2003). The other one, the long rains, in boreal spring, exhibits much more inconsistent variations from station to station, which are also harder to relate to anomalies of the climate system (hence, are more

difficult to predict; Ogallo 1989; Nicholson 1996; Camberlin and Philippon 2002). As a set-off, the long rains tend to be more reliable and constitute the main agricultural season in most parts of Kenya (Woodhead 1970; Akonga et al. 1988). However, their onset displays substantial year-to-year variations, partly related to pressure gradients between the Indian and the Atlantic oceans (Camberlin and Okoola 2003). Therefore, the long rains seasonal rainfall depends more strongly on the onset than on the cessation dates. Such information is not known for the short rains. Moron et al. (2007) included Kenya in their comparative study. They found that, as in the four other tropical regions they considered, inter-annual rainfall variations were much more related to the frequency of the rain events than to daily rainfall intensity (INT). However, the short rains, for which some spatial coherence was found in rainfall intensity, stood out as an exception. However, the study was based on a small number of stations (nine only) and it considered a fixed definition of the rainy seasons (March–April–May [MAM] for the long rains, October–November–December [OND] for the short rains). Therefore, no attempt was made to analyse the variations in the onset/end of the rains.

On the whole, too little is still known about the inter-annual variations of the components of both rainy seasons (onset, cessation, frequency of rain events, etc) and how these components individually contribute to higher or lower rainfall amounts for the season as a whole. Also, beyond the above-noted differences in the spatial coherence of rainfall between the long and the short rains, it is not clear whether this spatial coherence remains steady all along each rainy season (e.g. during the onset, core and withdrawal phases) and how strong it is in the other seasons (June–September and January–February), which are not always totally dry across Kenya. Addressing these issues will be done in four successive steps:

1. Identify key descriptors of rainfall behaviour (timing of the seasonal rains and within-season distribution).
2. Investigate the spatial coherence of these descriptors (how well their inter-annual variability matches between all the stations of a network), as a complement to the study by Moron et al. (2007) based on a small number of stations.
3. For those descriptors not attached to the timing of the rains, assess how the spatial coherence varies all along the seasonal cycle.
4. Analyse how these descriptors are related to each other. In particular, for the region as a whole, determine which descriptors best explain the inter-annual variability of the seasonal rainfall amounts.

The climatic characteristics of the region under study and the data to be used are presented in Section 2. The components of the rainy seasons and the methods used for

their analysis are discussed under Section 3. The results, first regarding the spatial coherence of inter-annual variations, then the relationships between the various components of the rainy season variability, are shown in Section 4.

2 Data

Data consist of daily precipitation amounts obtained from the Kenya Meteorological Department (KMD) and completed by data from the Tanzania Meteorological Agency and from the National Climate Data Center (NCDC/NOAA, United States). For consistency with prior work on the timing of the rains in East Africa, the same network of stations as in Camberlin and Okoola (2003), covering the period 1958–1987, is used in this study. It comprises 34 stations (Fig. 1a), 24 of which are located in Kenya. The others are in northeastern Tanzania, which experiences a bimodal rainfall regime similar to the one found in most of Kenya. There are a few missing values (3%) scattered in both space and time. Their handling is discussed below.

A second, incomplete data set of 24 stations covering the period 1988–2001 has been collated. It is only used in the regional assessment (Section 4.2), not in the station analyses. The aim is to document a possible evolution in the various descriptors of the rainy season for Kenya as a whole. There are many missing values, and the number of stations available on a given day ranges from 15 to 24. Their location is shown on Fig. 1 (filled squares).

The region under study lies across the equator. The climate is arid to sub-humid, with annual rainfall mostly in the range 300–1,000 mm (Fig. 1a). Higher values are found

in some highland areas. The rainfall regimes are mainly bimodal. Two rainy seasons occur in March–May (MAM, known as the long rains) and October–December (OND, the short rains). These rainy seasons accompany the meridional shift of the intertropical convergence zone (ITCZ), moving from the southern to the northern hemisphere during the long rains and back to the southern hemisphere during the short rains. The rainy seasons are separated by two dry seasons in June–September and January–February. The June–September season, however, remains relatively wet around the western Kenya highlands and on the Indian Ocean coast, whilst in January–February, the dryness is often tempered by unseasonable rains. Therefore, some analyses will consider the whole annual cycle and not only the two main rainy seasons.

3 Materials and methods

A set of variables is defined as descriptors of the rainy season. On a local basis (station), as in Moron et al. (2007), three variables are computed for each rainy season:

1. P_T is the total precipitation amount (in millimetres) over the period of time T ;
2. NRD_T is the number of rain days (1 mm or above) over the same period;
3. INT_T is the intensity of the rains, in millimetres per rain day.

No attempt is made to estimate each daily missing value. Instead, it is the descriptor itself, for period T , which is reconstructed, whenever some days are missing during this

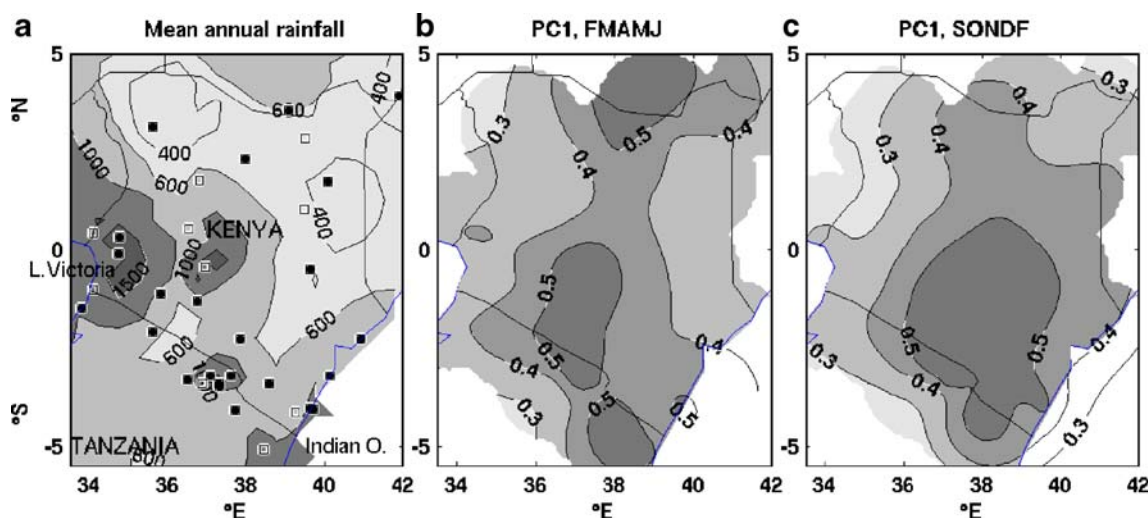


Fig. 1 Mean annual rainfall map (a) and loading patterns of the first principal component of daily rainfall for February–June (b) and September–February (c) 1958–1987. On a, the squares show the stations used in the study (empty squares data on 1958–1987 only,

filled squares stations with some additional data available on 1988–2001). On b and c, the shadings show the correlation values between the PC scores and the station time series

period. To obtain the reconstructed value of the descriptor for year Y , we first compute the actual statistic of the descriptor (e.g. INT_7) using all the available days during period T in year Y . This statistic is then complemented by the long-term average of the descriptor for the missing days. The final estimate is the weighted average of the two values (weights are the number of non-missing and missing days, respectively). This procedure enables us to make maximum use of the available data (when a seasonal value is missing, it is often because of only a few days or a single month), whilst not requiring questionable reconstructions of the daily values. It has been verified that the method of gap filling has very little effect on the results.

Using the inter-annual time series (1958–1987) of each descriptor, inter-station correlations are computed in order to assess the spatial coherence of the inter-annual variations. This is first done on two fixed periods of time: MAM, as representative of the long rains, and OND, of the short rains. The 30-day means, using a mobile window starting on January 1 and ending on December 31, are next considered, in order to show how the inter-station correlations vary in the course of the annual cycle.

The three descriptors are also computed for the actual rainy season in each year, allowing for variations in the onset and the cessation of rains. The onset (ONS) and the cessation (CES) of the rainy season are defined based on the method of cumulative anomalies. Accumulated rainfall is often considered in order to detect onset dates (e.g. Ilesanmi 1972; Nicholls 1984). The use of rainfall anomalies (i.e. departures from the long-term mean) enables us to better separate the dry and the rainy seasons with respect to the mean climate at each station, hence avoiding the difficult choice of absolute thresholds (Liebmann and Marengo 2001; Camberlin and Diop 2003; Liebmann et al. 2007). This method has been used in Kenya to detect the onset and cessation of the “long rains”, on a regional basis (Camberlin and Okoola 2003), and in this study will again be retained to get “whole Kenya” onset and cessation dates for the long and short rains. It is based on a principal component analysis (PCA) of all stations’ time series, over a period longer than the rainy season (about twice its usual length), including parts of both the antecedent and posterior dry seasons. The daily rainfall amounts are first square-rooted in order to reduce skewness in the distribution. For each year separately, accumulated scores of the leading principal component (PC1) then serve to identify regional ONS and CES dates. The onset date corresponds to the time at which the minimum accumulated rainfall score is reached for that year. Reciprocally, the time on which the maximum accumulated rainfall score is attained locates the cessation date. Discussions on the robustness and representativeness of the results are provided in Camberlin and Diop (2003) and Camberlin and Okoola (2003).

For the long rains, the computation of onset and cessation dates is based on the analysis of daily rainfall from February to June, and for the short rains, from September to February. The short rains generally end in December. However, they occasionally continue into the (usually) dry months of January–February, as in 1979 for instance (Wairoto and Nyenzi 1987). Therefore, the period used for the identification of the short rains onset and cessation had to be extended to late February. The possibility theoretically exists that no real discontinuity is found between the short and the long rains, but in practise, it is almost never encountered. PC1 accounts for 21.2% and 20.8% of the variance for FMAMJ and SONDJF, respectively. For both periods, PC1 depicts a uniform signal over most of Kenya and northeastern Tanzania (Fig. 1b, c). It materialises partly in-phase rainfall variability across the region, incorporating seasonal, inter-annual and intra-seasonal signals. Of course, all of them are to be considered together since the aim is to analyse inter-annual variations in the onset and cessation dates of the rainy season. The accumulated PC1 scores are used to detect the onset and cessation for each year and each season, following Camberlin and Okoola (2003). Tests were carried out to verify the robustness of the results. It was found that the addition of 1 month at the beginning or the end of the period used for PCA did not change appreciably the onset and cessation dates of the rainy seasons.

Regional indices of P , NRD, and INT are also computed. They are weighted averages of all the stations of the network. For consistency with the onset and cessation dates, which are based on the PC1 of daily rainfall, the PC1 loading coefficients are used as weights applied on the standardised station time series. Since the NRD is of course very much dependant on the length of the rainy season, the frequency of rain days (FRD) is defined as the percentage number of days, within the rainy season, when precipitation was recorded. FRD is used in place of NRD in the regional analysis. The five resulting descriptors, for the Kenya–northern Tanzania region as a whole, are then noted as P_R , FRD_R , INT_R , ONS_R and CES_R .

The regional indices are initially computed on the period 1958–1987. Their extension to 2001 is permitted by the use of the additional data set of incomplete rainfall records. For these extra years, the weighted mean of all the available stations is computed, as explained above. It is important to make sure that the standardisation of the time series for each descriptor is based on the reference period 1958–1987 for all stations, including those whose records extend to 1988–2001. The onset and cessation dates for these additional years are obtained by reconstructing the PC1 scores based on the available stations, then determining the dates from the scores as above.

The resulting regional indices, as well as the station time series, serve to assess how the variations in the

seasonal precipitation amount P depend on the various aspects of rainfall distribution (frequency of the rains—FRD, intensity—INT and timing—ONS and CES). The relationships between all these variables are first studied, and then a descriptive multi-linear regression model relating P to the other variables is proposed.

4 Results

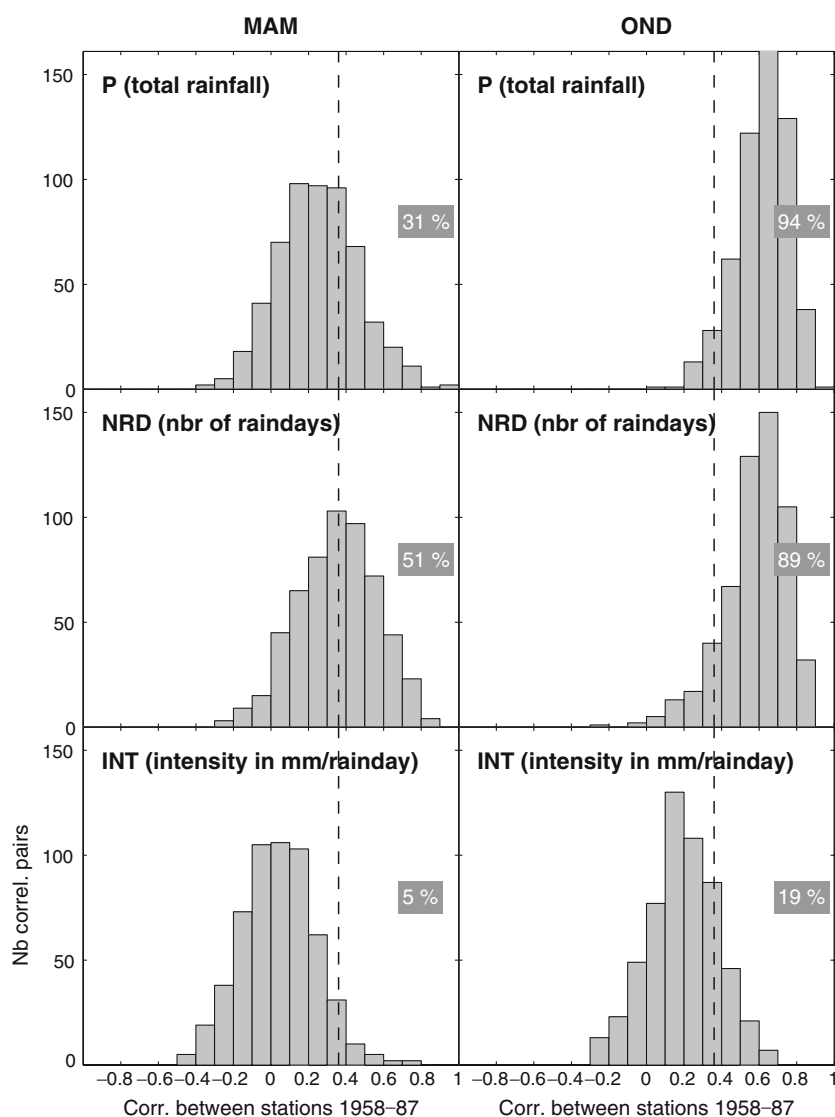
4.1 Spatial coherence of inter-annual variability

In this section, we first consider each rainy season as a fixed period of time, i.e. MAM for the long rains and OND for the short rains. As a way of assessing the spatial coherence of rainfall variability, all the inter-station correlations in the inter-annual variations of P , NRD and INT are plotted on Fig. 2, first for MAM (left panels). It is evident

that the spatial coherence is weak (though not completely lacking) during this rainy season. For P , only 31% of the correlations are significant at the one-sided 95% confidence level (c.l.). The coherence in the year-to-year variations of the rain day occurrence (NRD) is higher, with 51% of significant correlations. There is less coherence in the mean seasonal intensity of the rains (INT). About 5% only of the correlations are significant, which means that the inter-annual variations in the precipitation intensity at the various stations are almost totally independent from each other.

The patterns for OND are radically different (Fig. 2, right panels). As much as 94% of the correlations for P are significant (95% c.l.), with most of them comprised between 0.6 and 0.8. The same applies to NRD, though the percentage of significant correlations is marginally lower (89%), which contrasts with MAM during which rainfall occurrence was spatially more coherent than total rainfall amount. Coherence in daily rainfall INT is again

Fig. 2 Distribution of the inter-station correlation coefficients between the inter-annual variations (1958–1987) of each rainfall descriptor for the MAM season (*left panels*) and the OND season (*right panels*). The vertical line indicates the 95% c.l. for the correlation between two series displaying no autocorrelation. The percent values show the percentage of correlations which are significant



much lower than that found for the two other variables, but still greater than MAM rainfall intensity, with 19% of significant correlations, as against 5%. This (slightly) increased coherence in INT contributes to the very high coherence of P (on a station basis, P is the product of NRD and INT).

The above results point to a weaker inter-station coherence in MAM. In order to confirm the difference between the two rainy seasons, spatial numbers of degrees of freedom (DOF), as discussed in Moron et al. (2007), were computed. For the NRD, the DOF value found in MAM (5.1) is twice that of OND (2.6). Such a small value for OND, which is actually the same as the one found for the smaller network analysed in Moron et al. (2007), is indicative of a very strong spatial coherence in inter-annual variability. For MAM, the higher DOF value means that the coherence is reduced, though not nil (in such a case, the DOF would be close to 34). Two hypotheses are proposed. The first one is that there actually exists a common regional signal, but weaker than in OND, all across the region. The second one is that two (or even more) distinct climatic signals superimpose. To test these options, we adopt an approach similar to that followed in Moron et al. (2006). The results are presented for NRD, but similar ones are obtained for the other variables. Two ensemble data sets comprising the same number of years (30) and stations (34) as the observed one are created. In the first one, a uniform regional-scale signal is added to random white noise, in such a way that the signal represents 36% of the total variance (as in the observation for NRD). In the second ensemble data set, two independent climatic signals (together explaining 36% of the total variance as well) are added to random white noise. The inter-station correlations obtained for each ensemble data set are compared with those obtained for the observed data set, by showing their cumulative distribution functions (Fig. 3). Note that the information plotted, for the observed data set, is the same as that in Fig. 2 (central left panel) except that it is the cumulative frequency which is now displayed. It is noteworthy that the observed pattern is much closer to that of a uniform signal, than to that of two distinct climatic signals. Two distinct signals in MAM would actually result into two populations of correlations (Fig. 3), one made up of much lower (mostly around -0.2 to $+0.2$), the other one of much higher (around $+0.7$) correlations than those actually observed (median $+0.35$). In summary, this suggests that there is some form of spatially coherent and uniform climate forcing in MAM in the region although it is much weaker than in OND.

These differences between the two rainy seasons have been pointed to by Hastenrath (2000, 2007) and Hastenrath et al. (2007). They attributed the greater spatial coherence in boreal autumn to the manifestation of a powerful zonal-vertical circulation cell along the Equatorial Indian Ocean

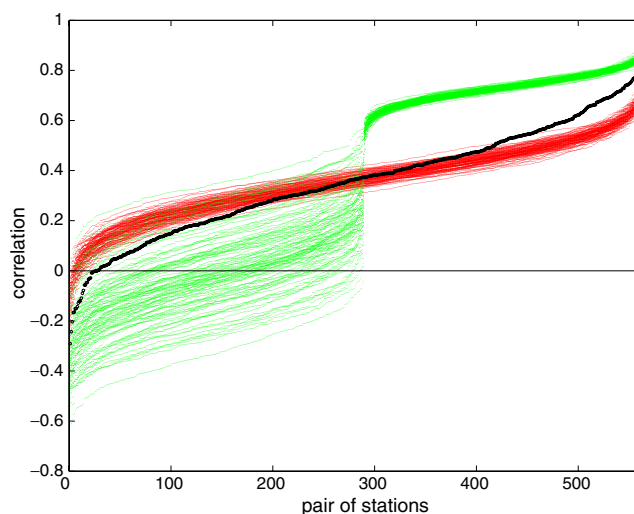


Fig. 3 Correlations between all possible pairs of stations for the inter-annual variations in the NRD during MAM. The plots show the cumulative distribution functions of the correlations in the observation and under two theoretical hypotheses. *Black curves* frequency of observed correlations (1958–1987), *red curves* frequency of correlations for an ensemble of random data superimposed on a single (uniform) regional-scale climate signal, *green curves* same as the red curves but for an ensemble of random data superimposed on two independent regional-scale climate signals

during this season. This cell controls the vertical motion over most of East Africa, and thus the rainfall activity.

A possible reason for the lower spatial coherence in MAM could also be that this season is temporally less coherent than OND (Beltrando 1990; Mutai and Ward 2000). To test this hypothesis, inter-station correlations have again been computed, now based on 30-day moving windows (Fig. 4). The analysis has been extended to the whole year, since the January–February and June–September periods are not totally dry across Kenya and northern Tanzania (see on Fig. 4 the squares which display the mean rainfall amounts). For each variable, Fig. 4 shows the detailed seasonal cycle of spatial coherence, in the form of the median value of the inter-station correlations. For instance, the value of 0.33 found on 1 January for P (thick solid line) is the median (across Kenya) of all the inter-station correlations between the inter-annual variations (1958–1987) of the 17 December to 15 January accumulated rainfall. The computation is repeated for all 30-day windows from January to December. Another analysis (not shown), based on the number of significant inter-station correlations (using Monte Carlo testing and taking into account autocorrelation in the time series), yielded seasonal patterns which were very close to those displayed on Fig. 4.

Two rainfall descriptors, P and NRD, display quite similar patterns, with strong seasonal variations in their spatial coherence (Fig. 4). The coherence in precipitation occurrence (NRD) is often slightly higher than that of

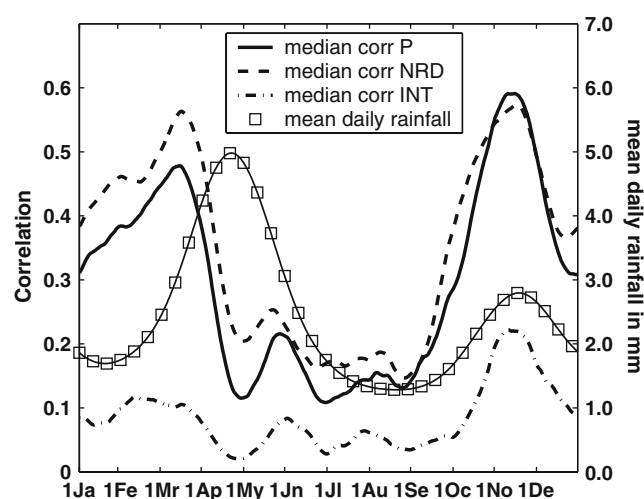


Fig. 4 Seasonal evolution of the median inter-station correlation, computed on shifting 30-day periods, in the inter-annual variations of total rainfall amount (P , solid line), NRD (dashed line) and daily rainfall INT (dash-dotted line). Approximate 95% c.l.=0.36. The 30-year mean daily rainfall for the region as a whole (smoothed using a 120-day low-pass filter) is shown by the solid line with squares

precipitation amount (P), especially from January to May. Both variables show two peaks. Maximum coherence is found in November, in the middle of the short rains. Another maximum occurs in March. It is noteworthy that, in this case, it does not coincide with the peak rainfall amount which, for the long rains, is found in late April for Kenya as a whole (Fig. 4, thin line with squares). The spatial coherence actually drops very fast to insignificant values at the peak of the long rains. These important results tend to indicate that large-scale mechanisms govern the short rains as a whole and only the onset of the long rains. This is in accordance with the spatial patterns of rainfall variability as deduced from PCA, which show that the first eigenvector depicts a coherent signal throughout East Africa in the short rains, whilst the long rains display a much more fragmented pattern (Ogallo 1989; Beltrando 1990; Camberlin and Philippon 2002). It is likely that the (weak) spatially uniform signal in MAM, as evidenced above (Fig. 3), is mainly related to the onset of the rainy season. This hypothesis will be further tested below. The lowest spatial coherence, for P and NRD, is found from April to August (Fig. 4). Coherence strongly increases in September, whilst the mean rainfall amount is still low, also suggesting that some large-scale mechanisms govern the onset of the short rains. Finally, a relatively high coherence bridges the gap between the short and the long rains in January–February. Occasional heavy rains during this period, like in 1979, 1993 or 1998 tend to occur simultaneously at most stations.

In contrast with the other two variables, the inter-station coherence in daily rainfall INT is very low. The median correlation value does not reach the 95% significance level

at any time during the year, though it shows a peak in the middle of the short rains (November; Fig. 4). Moron et al. (2007) similarly found that the spatial coherence in the inter-annual variations of mean rain per rain day was very low in all the five tropical regions they analysed, with the notable exception of the short rains in Kenya. Even in this case, its coherence, however, remains much smaller than those of the total precipitation amount and NRD.

The use of the median in Fig. 4 may mask out regional contrasts in spatial coherence. To deal with this issue, all correlations with stations located within in a 400-km radius around each station are retained. Different radii were tested, giving more or less the same results. Around each station, the correlations are then regressed against distance, and the predicted correlation value (as from the regression model) reached at an arbitrary distance of 200 km is plotted (Fig. 5). A high value means that there is coherence in inter-annual rainfall variations around the station which is considered. The analysis is carried out for each month and each variable (P , NRD and INT). For both P and NRD, the greater coherence of the half-year between October and March is clear (Fig. 5). The peaks in November and March are also conspicuous. Coherence strongly decreases in April, though secondary coherence is found in the south, especially along the coast, from May to September. These months are mostly dry, except in both western and coastal Kenya. The coast is known to be affected by organised disturbances, likely to explain the in-phase relationships in rainfall variability. Conversely, the daily rainfall INT maps rarely display significant correlations, with the notable exception of November, and to some extent, the southern/coastal areas from June to October. Whatever the variable, it is noteworthy that no particular shift (following the ITCZ movement for instance) in the area of maximum correlations can be observed. For instance, in January–February, when the ITCZ is located away from the region (further south), the correlations are high, whereas during the other dry season (June–August), when the ITCZ is again away from the region (further north), the correlations are low.

4.2 Relationships between the different properties of each rainy season

The fact that the coherence in the short rains peaks in the middle of the season, whilst that of the long rains peaks in its early part, suggests that the overall inter-annual rainfall variability of each season may be associated with different mechanisms. In particular, we can assume that the length of the rainy season may play a different part in the two rainy seasons.

This aspect is explored by computing correlations between the various descriptors of the rainy season variability. The approach is regional. To depict the inter-annual variations in the onset, end and length of the rainy season, the detection

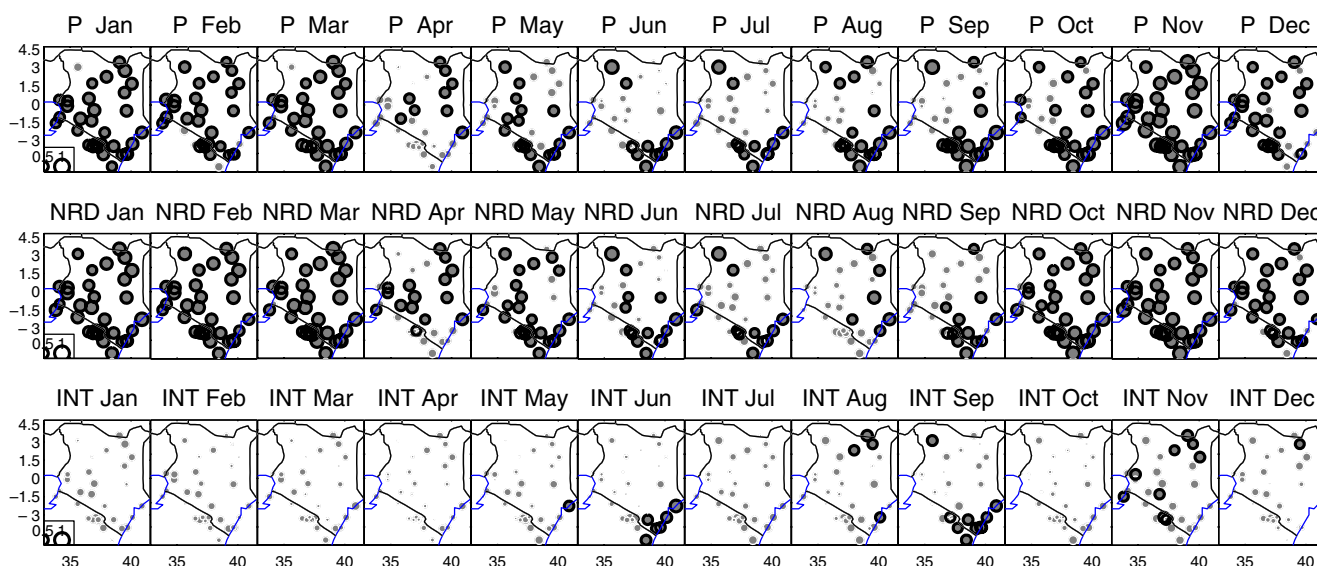


Fig. 5 Month-to-month changes in the average inter-station correlation, at a distance of 200 km around each station, between the inter-annual variations of P , NRD and INT (see text). A black circle indicates a correlation at 200 km from the station exceeding the 95% significance level

method presented above and in Camberlin and Okoola (2003) is used. Two descriptors of the rainy season characteristics as used in Section 4.1, P and INT, are also retained and computed on the period delimited by the onset and cessation dates; therefore, varying every year. Instead of the NRD, which is inherently related to the duration of the rainy season, we use the FRD during each rainy season, expressed as the ratio between NRD and the length of the season. Weighted means of the stations' time series constitute the regional indices (see Section 3 above). It must be remembered that the regional indices are less representative in the long rains than in the short rains. However, even in MAM, the seasonal totals of 31 stations out of 34 are significantly correlated with the regional index, which justifies the consideration of a single index for the long rains.

Statistics for the two rainy seasons are first compared (Table 1). The two rainy seasons last on average about 2 months, but the variability of the short rains is larger (by 10 days for both the standard deviations of the onset and cessation) than that of the long rains. The onset is more variable than the cessation, this result being an extension to the short rains of the one obtained by Camberlin and

Okoola (2003) for the long rains. The greater inter-annual variability of the short rains is also found in the other three seasonal descriptors, P , FRD and INT. The higher standard deviations of the short rains are all the more remarkable when it is realised that both the mean frequency of the rains and their intensity are much lower during the short rains. The large inter-annual variability of the short rains has been widely reported before (e.g. Ogallo 1989; Hastenrath et al. 1993; Black et al. 2003), but the present results do show that these large variations occur in conjunction with those of the various rainy season attributes (onset, cessation, rainfall occurrence and rainfall intensity).

Table 2 shows the correlations between the descriptors of the long rains in the period 1958–1987. The inter-annual variations of P are weakly related to the rainfall occurrence (FRD), the intensity of the rains (INT) and the CES date. The confidence levels are virtually the same, since correlations range between 0.43 and 0.48. A higher rainfall amount (P) during the long rains is partly explained by more rainy days, more intense rains and a late end of the season. However, a much higher correlation with P is found for the onset date (-0.83), meaning that much of the long

Table 1 Statistics on onset, cessation, total rainfall, frequency of rain days and daily rainfall intensity for the long rains and the short rains (regional values for Kenya and northern Tanzania), 1958–1987

		ONS	CES	P (mm)	FRD (rain days/day)	INT (mm/day)
Mean	Long rains	25 March	21 May	357	0.43	14.3
	Short rains	23 October	26 December	226	0.30	10.9
Standard deviation of regional index	Long rains	14.5 days	10.3 days	137 (38%)	0.05 (12%)	1.2 (8%)
	Short rains	24.7 days	20.4 days	174 (77%)	0.06 (20%)	2.7 (25%)

The percentage of the mean for the corresponding season is shown in parentheses

Table 2 Correlations between the regional descriptors of the long rains (1958–1987)

	<i>P</i>	FRD	INT	ONS	CES
FRD (rain days/day)	<i>0.43</i>				
INT (mm/day)	<i>0.48</i>	<i>0.44</i>			
ONS	<i>-0.83</i>	<i>-0.21</i>	<i>-0.32</i>		
CES	<i>0.45</i>	<i>-0.18</i>	<i>-0.02</i>	<i>-0.08</i>	
Length	<i>0.91</i>	<i>0.07</i>	<i>0.24</i>	<i>-0.83</i>	<i>0.62</i>

Significant correlations (95% c.l.) are shown in italics

rains variability is controlled by inter-annual variations in the onset, with higher amounts in seasons starting early. The collinearity between all the descriptors is found to be low (Table 2). For instance, earlier onsets (or late retreats) of the rainy season are no clues that there will be more wet days, or greater daily rainfall amounts, during the rainy season. A slight exception in this absence of collinearity is the correlation between FRD and INT. The positive value (+0.44, significant at the 95% c.l.) denotes that more intense rainfall tends to be recorded in years when it rains more often, though the relationship remains weak. This is not a trivial result, since in most tropical regions investigated by Moron et al. (2007), with a slight exception for Kenya, the common variance between the rainfall occurrence (“O”) and the rainfall intensity (“I”) is close to zero at station scale.

A multiple linear regression is then defined to model the inter-annual variations of *P*, as a function of the other four descriptors of the rainy season. For the long rains, the regression equation (with a null intercept, since all the independent variables are first standardised) is as follows:

$$P_{R \text{ long rains}} = -0.69 \text{ ONS}_R + 0.46 \text{ CES}_R + 0.30 \text{ FRD}_R + 0.14 \text{ INT}_R.$$

Overall, 97.7% of the variance of *P* is explained by the combination of the four descriptors. A 100% value may not be reached since the INT and FRD descriptors are spatial averages of the values computed at individual stations. An analysis of variance (not shown) indicates that each descriptor separately contributes to explain a significant part of the inter-annual variations of *P*. It is confirmed that the main factor accounting for the fluctuation of *P* is the date of onset of the rains, which has the greatest coefficient in the equation. This date is highly variable from year to year, with a standard deviation close to 15 days (Table 1). The date of cessation ranks second. The length of the rainy season is, therefore, decisive and explains 83% of the variance of *P*. The contribution of FRD is smaller, whilst INT ranks last. Heavy (and more frequent) rains do not

usually compensate the rainfall deficit induced by a late start of the rainy season.

Figure 6 shows for all years between 1958 and 2001 the anomalies in the various descriptors of the long rains. Circles indicate years with values below -0.5 standard deviation. Most years with a significant deficit during the long rains (circles in the left column of Fig. 6) coincide with a late starting season. A similar coincidence can be found between excess rainfall and early onsets (plus signs on Fig. 6). However, a close inspection reveals that, in individual years, there are many different combinations to explain deficits. For instance, the year 2000 drought (DMC 2000) is related to a short rainy season, an insufficient number of rain events and low rainfall yields per event. By contrast, the 1984 drought only exhibited a shortened rainy season whilst other variables were normal or above average. The 1965 season started on 25 March, which is the long-term average onset date, but an early end, low intensities and, moreover, a small NRD all combined to result in a very poor year. In some cases, late onsets may be partly compensated by late cessation dates, like in 1972–1973, 1983 or 1987.

The same analysis is performed on the short rains. Table 3 shows that the correlations between the descriptors are generally higher than for the long rains. In particular, the total rainfall amount during the rainy season is now strongly tied to the FRD ($r=0.67$) and daily intensity (INT, $r=0.72$). The correlations obtained between *P* and the timing of the rains are more or less the same as those found during the long rains. The onset ($r=-0.78$) is much more decisive than the cessation ($r=+0.49$) in fixing the total seasonal amount.

The multiple linear equation is as follows:

$$P_{R \text{ short rains}} = -0.52 \text{ ONS}_R + 0.36 \text{ CES}_R + 0.28 \text{ FRD}_R + 0.27 \text{ INT}_R.$$

The percentage of variance of *P* explained by the regression is marginally smaller than in the long rains (95.9%). As in the long rains, the leading contributor to *P* variability is the onset date, though the coefficient is now much smaller. The cessation date ranks second. The main difference with the long rains is that a larger contribution is accounted for by the intensity of the rains, with a regression coefficient almost as large as that of FRD. Hence, the timing of the rains, despite their very large inter-annual variations (Table 1), has a relatively smaller share in explaining the season variability than during the long rains. However, Table 3 shows that there is more collinearity between all the descriptors in the short rains. For instance, the onset is partly related to both FRD and INT whilst it was independent from them in the long rains: late starting seasons tend to have relatively less frequent rain days, each

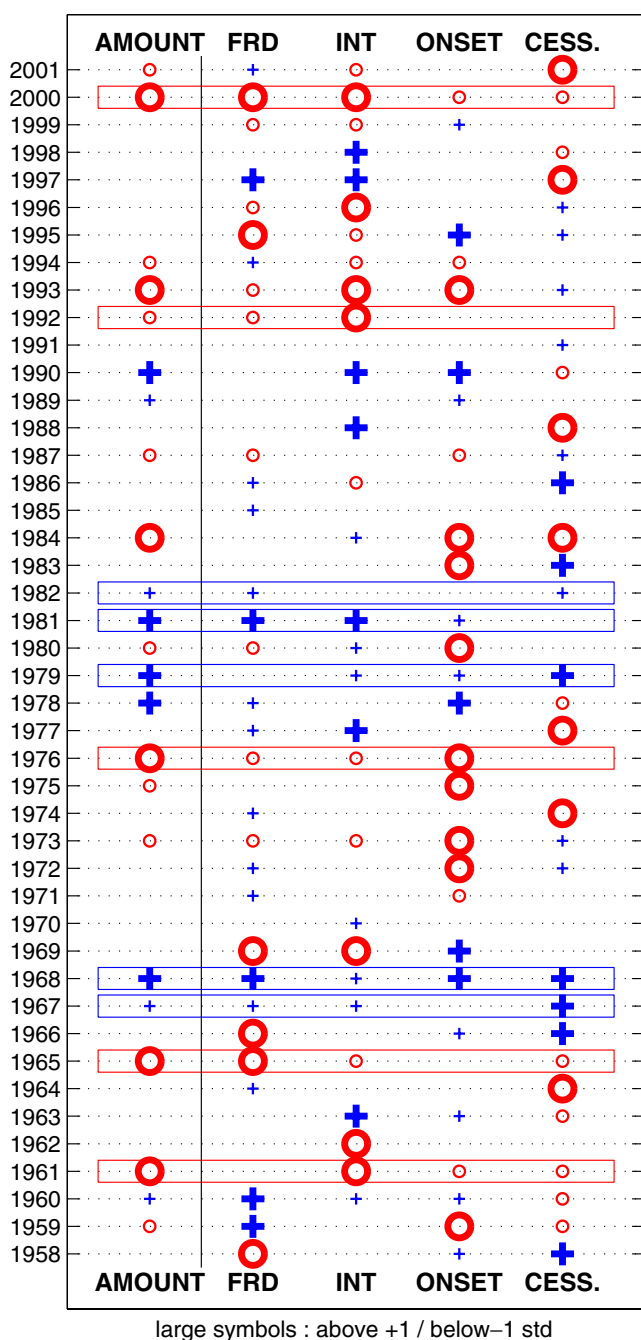


Fig. 6 Anomalies of the various components of the long rains for each year between 1958 and 2001. “Bad years” are denoted as *circles* (i.e. anomalies below -0.5 standard deviation, with *large symbols* below -1 standard deviation). “Good years” are denoted as *pluses* (i.e. anomalies above $+0.5$ standard deviation, with *large symbols* above $+1$ standard deviation). The sign of the onset has been switched so that positive anomalies now denote early onsets and negative anomalies late onsets. Data for 1988–2001 is based on a reduced network of stations. *Amount* is the total rainfall recorded from the onset to the cessation dates. *Rectangles* indicate years for which there is consistency between the majority of the variables

one recording lower precipitation. Except for the cessation date which is more independent, the interpretation should not separate the different descriptors. It can be, therefore, concluded that good short rains are related to a general enhancement of the rainy season, including its lengthening (especially an earlier start), more frequent rainfall events and a greater intensity of each one of them, all these features being partly associated.

The covariations of several descriptors of the short rains are well-shown in Fig. 7. Dry years are often found to have a below normal FRD, low rainfall intensities, a late onset and early cessation, or at least two of these features whilst the others remain close to normal. It is also evident that the wettest years (1961, 1972, 1977, 1978, 1982 and 1997) result from a conjunction of favourable conditions for most descriptors. A few years show contrasted anomalies (e.g. 1969 had a long season, but a low FRD), but on the whole there is reasonable agreement between the different descriptors, except for the cessation date, which, as shown in Table 3, behaves more independently.

5 Discussion

The analysis of the space–time characteristics of precipitation in Kenya and northern Tanzania revealed three main issues which deserve further discussion. First, inter-station correlations showed that the spatial coherence of monthly or seasonal rainfall occurrence, throughout the year, was much higher than that of rainfall intensity. There is growing evidence that, in the Tropics, most large-scale or regional-scale signals of inter-annual (Robertson et al. 2006; Moron et al. 2007) as well as decadal variations (Le Barbé et al. 2002) involve a modulation of rainfall occurrence, rather than a modulation of the intensity of individual rainfall events. These observations are confirmed, for East Africa, by the above findings. It is suggested that higher mean monthly or seasonal intensities in a given year at a given location may only be due to one or two heavy rainfall events, which are unlikely to be recorded at the same time in the other stations of the network. It should be noted that,

Table 3 Correlations between the regional descriptors of the short rains (1958–1987)

	<i>P</i>	FRD	INT	ONS	CES
FRD (rain days/day)	<i>0.67</i>				
INT (mm/day)	<i>0.72</i>	<i>0.48</i>			
ONS	-0.78	-0.45	-0.43		
CES	<i>0.49</i>	0.10	0.27	-0.06	
Length	<i>0.89</i>	<i>0.40</i>	<i>0.49</i>	-0.79	<i>0.66</i>

Significant correlations (95% c.l.) are shown in *italics*

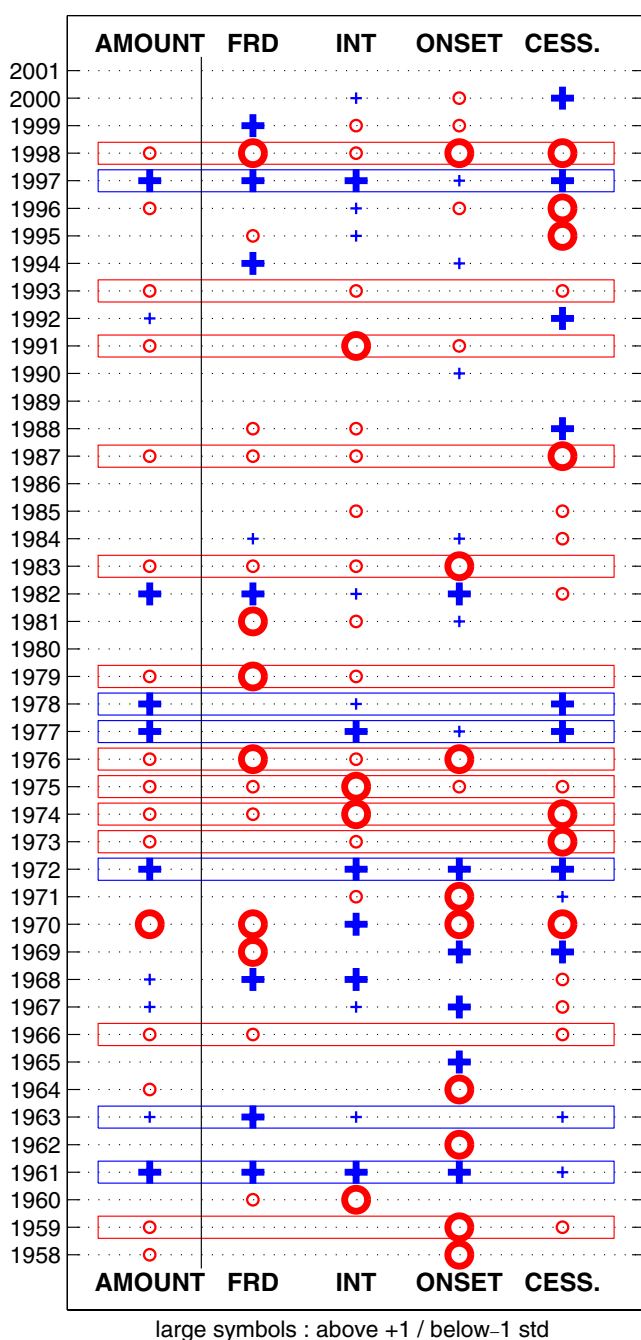


Fig. 7 Same as Fig. 6 but for the short rains

independently from inter-annual variations, mean patterns of rainfall intensity (per rain day) in the Tropics, in both time and space, often markedly differ from those of rainfall amounts. The seasonal cycle sometimes display higher intensities in the early (and drier) part of the rainy season (Jackson 1986). Over West Africa (Le Barbé et al. 2002) and Southwest USA (Wang et al. 2007), whilst rainfall occurrence replicates the rainfall regimes, rainfall intensity generally shows little variations between the beginning and the end of the rainy season. Spatial patterns also display

such discrepancies. In Tanzania, the monthly maps of mean daily rainfall INT, with the exception of April, bear only loose similarities with those of monthly rainfall (Jackson 1972).

A slight exception to the absence of spatial coherence for mean intensity was noticed in Moron et al. (2007) for the East African short rains and confirmed in the present study using a denser station network. Inter-station correlations in the inter-annual variations of rainfall intensity, though lower than those of rainfall occurrence, markedly increase at this time of the year. It remains difficult to interpret this specificity of the short rains. It is clear that the factors which trigger rainfall occurrence also act to modulate rainfall intensity at this time of the year. Excess short rains are related to large-scale Walker circulation anomalies along the equatorial Indian Ocean (e.g. Hastenrath et al. 1993, 2007), resulting in increased moisture convergence, instability and convection over all parts of Eastern Africa.

A second set of results is that, besides rainfall intensity, there exist marked seasonal changes in the spatial coherence of rainfall occurrence and rainfall amounts over East Africa. The inter-annual variations of these two variables are spatially much less consistent in MAM than in OND. More generally, it transpires that the year may be divided into two contrasted periods. The first one from September to March (boreal winter half of the year) is one of strong spatial coherence; the second one from April to August (boreal summer half of the year) shows little spatial coherence. The weakening of the coherence between the early (March) and late part (May) of the long rains is outstanding and coincides with the gradual replacement of the north-easterlies by thickening south-easterlies (Fig. 8). It is suggested that, in the early part of the long rains season, rainfall occurrence is contingent upon the northward shift of the ITCZ, especially

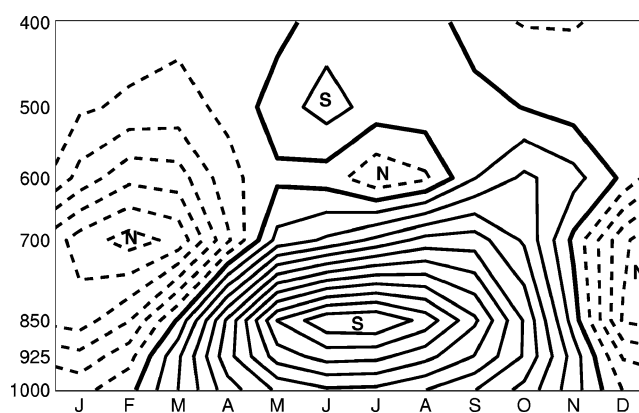


Fig. 8 Mean meridional wind flow over Eastern Equatorial Africa (5° S–5° N, 35–40° E), as a function of month and pressure level, between 1,000 and 400 hPa (source: Oettli 2008, based on NCEP/NCAR reanalysis data; solid lines southerly flow, dashed lines northerly flow. The contours are spaced at a 0.5-m s⁻¹ interval, with the zero isotach shown as bold line)

the development of a low-level moisture influx associated with south-easterly winds. This shift being related to large-scale circulation, we expect years on which it is early (late) to result into wet (dry) conditions at most stations at the same time, hence the strong spatial coherence. Later in the season, more random convection is expected whilst the air mass is more uniform, hence the weak spatial coherence. In the short rains, the ITCZ does not have the structure shown in the long rains. There is widespread convergence between the north-easterlies and the south-easterlies, but its intensity depends on large-scale Walker circulation anomalies across the Indian Ocean as noted above. These anomalies are persistent all along the boreal fall season, which is the reason why spatial coherence is very strong at this time of the year, and is maintained for several consecutive months. All aspects of the temporal organisation of the rains are, therefore, concerned, in a consistent way (onset, rainfall occurrence and, to a lesser extent, mean intensity and cessation).

The third set of results, which is tightly related to the last one, is on the dependence of the seasonal rains to the timing of the rainy season and within-season distribution of the rain events. The onset is the feature which has the greatest effect on the overall seasonal rainfall total. Variations in the cessation date have a much smaller impact. The greater dependence of both the long and the short rains to the onset than to the cessation may be compared with results for other regions of the world. For the Indian monsoon, using a hydrological definition of its onset and withdrawal, Fasullo and Webster (2003) indicate that the seasonal precipitation amount (all-India July–September rainfall) is more strongly correlated with the withdrawal than the onset dates ($r=0.58$ and -0.33 over the period 1948–2000, respectively). Other Indian monsoon onset indices are actually very poorly correlated with all-India precipitation amounts (Dhar et al. 1980; Nayagam et al. 2008). By contrast, Liebmann and Marengo (2001) found that it was through a modulation of the onset that SST were impacting seasonal rainfall variability in the Amazon Basin. This result is close to what is found for the long rains in Kenya and Tanzania, where, at the regional scale, most of the inter-annual variability of the seasonal rains is related to that of the onset. In the short rains, the more consistent variations in all the rainy season characteristics, hence the dependence on the onset is less obvious. These slightly different results obtained for the two rainy seasons, added to the above-noted findings for India, indicate that there is no unique relationship between seasonal rainfall and the temporal characteristics of rainfall distribution.

6 Conclusions

This study provides new information on both the spatial coherence of rainfall in East Africa, and the contribution of

the various intra-seasonal characteristics of the rains to the seasonal precipitation totals. Based on a network of 34 stations covering Kenya and northeastern Tanzania, four main variables depicting the elements which contribute to the inter-annual variability of the rains were retained: the number of rain days (NRD), the mean daily rainfall intensity (INT), the onset (ONS) and the cessation (CES) dates of the rainy season.

The main results are summarised as follows:

1. The spatial coherence of the rains strongly varies in the course of the seasonal cycle. These variations do not replicate those of the rainfall regimes: whilst maximum coherence is found at the peak of the October–November short rains, the March–May long rains show a strong spatial coherence at its beginning only. These differences are suggested to be due to contrasts in the ITCZ structure and dependence on large-scale circulation.
2. Most of the spatial coherence is attached to the occurrence of the rainfall events, rather than on their intensity. In agreement with results obtained elsewhere in the Tropics, daily rainfall INT shows almost no coherence, except weakly at the peak of the short rains.
3. In both rainy seasons, the onset is more variable than the cessation. A possible reason is that the wetter land surface at the end of each rainy season has a positive feedback on the atmosphere, which dampens remote forced variations, whereas, for the onset, external forcings are stronger.
4. At the regional scale, inter-annual variations in the onset date of the rainy season have the biggest impact on the seasonal rains, which are less dependent on within-season variations of both the NRD and daily rainfall INT.
5. However, there exist some differences between the two rainy seasons. A greater inter-annual variability of the short rains is found for all the four seasonal descriptors (NRD, INT, ONS and CES). The large inter-annual variability of the short rains has been widely reported before (e.g. Ogallo 1989; Hastenrath et al. 1993; Black et al. 2003), but the present results do show that these large variations occur as a conjunction of those of the various rainy season attributes. Whilst in the short rains these attributes are all related, in the long rains they are more independent. An earlier onset (or late retreat) of the long rains is no clue that there will be more wet days, or greater daily rainfall amounts, during the rainy season. However, setting aside the impact of the season length, in wet long rains, it tends to rain more often rather than more heavily.

These results have several implications in terms of seasonal rainfall forecasting. The seasonal prediction of the long rains has remained a very difficult task to date, which is likely to result from the fact that the total rainfall

amount for this season is a combination of many independent features (variations in onset, rain day occurrence, rainfall intensity and cessation date). For the short rains, the different descriptors of the rains are more strongly inter-related. Additionally, they show a much stronger spatial coherence than in the long rains. This enhanced spatial coherence reflects the fact that rainfall variability at this time of the year is governed by large-scale circulation changes. In the long rains, rainfall tends to be more localised, except during the onset phase.

Acknowledgements This study is a contribution to the PICREVAT project, funded by the "Agence Nationale de la Recherche" under its VMCS program. The authors wish to thank the Kenya Meteorological Department, Nairobi, the Tanzania Meteorological Agency, Dar-Es-Salaam, as well as Patrick Valimba (Department of Water Resources Engineering, University of Dar-Es-Salaam), for providing a large part of the data used in this study.

References

- Akong A J, Downing TE, Konijn NT, Mungai DN, Muturi HR, Potter HL (1988) The effects of climatic variations on agriculture in central and eastern Kenya. In: Parry ML, Carter TR, Konijn NT (eds) *The impact of climatic variations on agriculture*. Kluwer, Dordrecht, The Netherlands, pp 121–270
- Beltrando G (1990) Space-time variability of rainfall in April and October–November over East Africa during the period 1932–83. *J Climatol* 10:691–702
- Black E, Slingo J, Sperber KR (2003) An observational study of the relationship between excessively strong short rains in Coastal East Africa and Indian Ocean SST. *Mon Weather Rev* 131:74–94
- Camberlin P, Diop M (2003) Application of daily rainfall principal component analysis to the assessment of the rainy season characteristics in Senegal. *Clim Res* 23:159–169
- Camberlin P, Okoola RE (2003) The onset and cessation of the 'long rains' in eastern Africa and their interannual variability. *Theor Appl Climatol* 75:43–54
- Camberlin P, Philippon N (2002) The east African March–May rainy season: associated atmospheric dynamics and predictability over the 1968–97 period. *J Climate* 15:1002–1019
- D'Amato N, Lebel T (1998) On the Characteristics of the rainfall events in the Sahel with a view to the analysis of climatic variability. *Int J Climatol* 18:955–974
- Dhar ON, Rakhecha PR, Mandal BN (1980) Does the early or late onset of monsoon provide any clue to subsequent rainfall during the monsoon season? *Mon Weather Rev* 108:1069–1072
- DMC (2000) DEKAD 19 report, 1–10 July, 2000. Ten day bulletin no. DMCN/01/337/19/07/2000. Drought Monitoring Centre, Nairobi, Kenya
- Fasullo J, Webster PJ (2003) A hydrological definition of the Indian summer monsoon onset and withdrawal. *J Climate* 16:3200–3211
- Hastenrath S (2000) Zonal circulations over the equatorial Indian Ocean. *J Climate* 13:2746–2756
- Hastenrath S (2007) Circulation mechanisms of climate anomalies in East Africa and the equatorial Indian Ocean. *Dyn Atmos Ocean* 43:25–35
- Hastenrath S, Nicklis A, Greischar L (1993) Atmospheric–hydrospheric mechanisms of climate anomalies in the western equatorial Indian Ocean. *J Geophys Res* 98:20219–20235
- Hastenrath S, Polzin D, Mutai C (2007) Diagnosing the 2005 drought in equatorial East Africa. *J Climate* 20:4628–4637
- Ilesanmi OO (1972) An empirical formulation of the onset, advance and retreat of rainfall in Nigeria. *J Trop Geogr* 34:15–33
- Jackson IJ (1972) Mean daily rainfall intensity and number of rain days over Tanzania. *Geogr Ann* 54A(3–4):36–375
- Jackson IJ (1986) Relationships between raindays, mean daily intensity and monthly rainfall in the Tropics. *J Climatol* 6:117–134
- Le Barbé L, Lebel T, Tapsoba D (2002) Rainfall variability in West Africa during the years 1950–1990. *J Climate* 15:187–202
- Liebmann B, Marengo JA (2001) Interannual variability of the rainy season and rainfall in the Brazilian Amazon Basin. *J Climate* 14:4308–4318
- Liebmann B, Camargo S, Marengo J, Vera C, Seth A, Carvalho L, Allured D, Fu R (2007) Onset and end of the rainy season in South America in observations and an atmospheric general circulation model. *J Climate* 20:2037–2050
- Marengo JA, Liebmann B, Kousky VE, Filizola N, Wainer IC (2001) Onset and end of the rainy season in the Brazilian Amazon Basin. *J Climate* 14:833–852
- Moron V, Robertson AW, Ward MN (2006) Seasonal predictability and spatial coherence of rainfall characteristics in the tropical setting of Senegal. *Mon Weather Rev* 134:3248–3262
- Moron V, Robertson AW, Ward MN, Camberlin P (2007) Spatial coherence of tropical rainfall at regional scale. *J Climate* 20:5244–5263
- Mutai CC, Ward MN (2000) East African rainfall and the tropical circulation/convection on intraseasonal to interannual timescales. *J Climate* 13:3915–3939
- Nayagam LR, Janardanan R, Ram Mohan HS (2008) An empirical model for the seasonal prediction of southwest monsoon rainfall over Kerala, a meteorological subdivision of India. *Int J Climatol* 28:823–831
- Nicholls N (1984) A system for predicting the onset of the north Australian wet-season. *Int J Climatol* 4:425–435
- Nicholson SE (1996) A review of climate dynamics and climate variability in Eastern Africa. In: Johnson TC, Odada EO (eds) *The limnology, climatology and paleoclimatology of the east African lakes*. Gordon and Breach, Toronto, pp 25–56
- Oettli P (2008) Précipitations et relief en Afrique orientale et australe: modélisations statistiques et géostatistiques. Ph.D. thesis, Université de Bourgogne, p 287
- Ogallal LJ (1989) The spatial and temporal patterns of the East African seasonal rainfall derived from principal component analysis. *J Climatol* 9:145–167
- Oladipo EO, Kyari JO (1993) Fluctuations in the onset, termination and length of the growing season in Northern Nigeria. *Theor Appl Climatol* 47:241–250
- Olaniran OJ, Summer GN (1990) Long-term variations of annual and growing season rainfalls in Nigeria. *Theor Appl Climatol* 41:41–53
- Philippon N, Camberlin P, Fauchereau N (2002) Empirical predictability study of October–December East African rainfall. *Q J R Meteorol Soc* 128:2239–2256
- Robertson AW, Kirshner S, Smyth P, Charles SP, Bates BC (2006) Subseasonal-to-interdecadal variability of the Australian Monsoon over North Queensland. *Q J R Meteorol Soc* 132:519–542
- Traoré SB, Reyniers FN, Vaksman M, Kouressy M, Yattara K, Yorote A, Sidibe A, Kone B (2000) Adaptation à la sécheresse des écotypes locaux de sorgho du Mali. *Sécheresse* 11:227–237
- Wairoto JG, Nyenzi BS (1987) The unseasonal heavy rains in Kenya during January 1979. In *Proceedings of the First Technical Conference on Meteorological Research in Eastern and Southern Africa*, KMD, Nairobi, 6–9 January, pp 173–178
- Wang J, Anderson BT, Salvucci GD (2007) Stochastic modeling of daily summertime rainfall over the southwestern U.S. Part II: intraseasonal variability. *J Hydrometeorol* 8:938–951
- Woodhead TW (1970) Confidence limits for seasonal rainfall and their value in Kenya agriculture. *Exp Agric* 6:81–86