

Spatial Coherence of tropical rainfall at Regional Scale

Vincent Moron (*,**)¹, Andrew W. Robertson (*), M. Neil Ward (*), Pierre Camberlin (@)

(*) International Research Institute for Climate and Society, The Earth Institute of Columbia
University, Palisades, New York

(**) CEREGE, Université d'Aix-Marseille I, Aix-en-Provence, France

(@) Centre de Recherches de Climatologie, Université de Bourgogne, Dijon, France

submitted to *Journal of Climate*

Manuscript JCLI-1623

Submitted: August 4, 2006

Revised: March 5, 2007

¹ Corresponding author address: Vincent Moron, Université d'Aix-Marseille I and Cerege, UMR 6635 CNRS, Europôle Méditerranéen de l'Arbois, BP 80, 13545 Aix en Provence, France (moron@cerege.fr).

Abstract

This study examines the spatial coherence characteristics of daily station observations of rainfall in 5 tropical regions during the principal rainfall season(s): the Brazilian Nordeste, Senegal, Kenya, Northwestern India and Northern Queensland. The rainfall networks include between 9 and 81 stations, and 29-70 seasons of observations. Seasonal-mean rainfall totals are decomposed in terms of daily rainfall frequency (i.e the number of wet days) and mean intensity (i.e. the mean rainfall amount on wet days).

Despite the diverse spatio-temporal sampling, orography and land-cover between regions, three general results emerge. (1) Interannual anomalies of rainfall frequency is usually the most spatially coherent variable, generally followed closely by the seasonal amount, with the daily mean intensity in distant third place. In some cases, such as Northwestern India which is characterized by large daily rainfall amounts, the frequency of occurrence is much more coherent than the seasonal amount. (2) On daily time scales, the inter-station correlations between amounts on wet days always fall to insignificant values beyond a distance of about 100 km. The spatial scale of daily rainfall occurrence is larger and is more variable amongst the networks. (3) The regional-scale signal of the seasonal amount is primarily related to a systematic spatially-coherent modulation of the frequency of occurrence.

1. Introduction

Tropical rainfall is mainly convective, with rainfall events characterized by short durations, high rain rates and small spatial scales (e.g. Houze and Cheng, 1977; Leary and Houze, 1979; Chen *et al.*, 1996; Rickenbach and Rutledge, 1998). The spatial correlation length, that is the approximate distance at which the correlation between a grid-point and the neighboring ones becomes lower than $1/e \sim 0.37$, of tropical deep convection at 3-hourly time scale has been estimated from satellite infra-red radiance data to be 95-155 km over the continents (Ricciardulli and Sardeshmukh, 2002). Smith *et al.* (2005) found similar length-scales for tropical rainfall from Tropical Rainfall Measuring Mission (TRMM) data, with large variations between regions (their Fig. 6 for example). A rainy season is comprised of a large number of individual rainfall events and this summation smoothes the rainfall field to a certain extent (e.g. Bacchi and Kottigoda, 1995; Abdou *et al.*, 2003). At the annual time scale, Dai *et al.* (1997) estimated that the spatial correlation of rainfall amounts falls to an insignificant level beyond a separation distance of about 200 km for the northern Tropics and about 550 km for the southern Tropics. New *et al.* (2000) found similar estimates using monthly rainfall.

Current seasonal prediction schemes concentrate on larger spatio-temporal scales by issuing three-month average predictions of rainfall amounts across homogeneous areas or at GCM grid-points (e.g. Goddard *et al.*, 2001; Gong *et al.*, 2003). However, potential users of seasonal predictions of tropical rainfall often need estimates at smaller spatio-temporal scale, such as onset date and length of the rainy season, or the amplitude and frequency of dry or wet spells at local scale (e.g. Ingram *et al.*, 2002; Baron *et al.*, 2005; Hansen *et al.*, 2006). These characteristics depend ultimately on the occurrence of rainfall and the distribution of rainfall amounts on wet days (e.g. Stern and Coe, 1984; Wilks, 1999). The extent to which the frequency of occurrence, and the shape and scale of the probability density function of rainfall at local scale are potentially predictable remains to be estimated.

An upper bound of the potential predictability may be inferred from the spatial coherence of regional-scale anomalies based on the hypothesis that any large-scale climate forcing, such as that provided by sea surface temperature (SST) anomalies, would, to first order, tend to give rise to a rather spatially-uniform signal in precipitation (Haylock and McBride, 2001), in the absence of strong lower-boundary gradients such as orography. In a recent study of daily rainfall on a network of 13 stations over Senegal during the West African monsoon season, Moron *et al.* (2006, hereafter MRW06) found seasonal anomalies of the frequency of occurrence of daily rainfall (i.e. the number of wet days) to be much more spatially coherent between the 13 stations, than seasonal anomalies in the average rainfall amount on wet days, i.e. the mean rainfall intensity. Consistent with this result, MRW06 used a set of atmospheric general circulation model (GCM) simulations made with prescribed historical SSTs, to demonstrate much higher skill in the GCM's simulation of seasonal rainfall frequency than mean intensity over the region of Senegal.

Many factors impact the spatio-temporal properties of rainfall, and their relative importance depends strongly on the scales analyzed. Tropical weather phenomena have been decomposed into at least four hierarchical spatial scales: the cumulus scale (~ 1-10 km), meso-scale (~ 10-300 km), cloud-cluster scale (~ 300-1000 km), and synoptic scale (~ 1000 km) (Houze and Chen, 1977), in addition to the tropical convergence zones at the planetary scale. The spatial extent of an individual rainfall event is ultimately related to the processes that generate rainfall, but is much smaller than the regional scale analyzed here. Very high rain rates are associated with small-scale individual convective cells (i.e. cumulus scale) embedded within meso-scale convective complexes (MCCs) and cloud clusters (e.g. Leary and Houze, 1979; Chen *et al.*, 1996; Rickenbach and Rutledge, 1998). The MCCs and cloud clusters contain both regions of convective cells and stratiform rain, more extended in space and lengthier in time (e.g. Lopez, 1978; Leary and Houze, 1979; Rickenbach and Rutledge, 1998). The MCCs and cloud clusters

are usually propagating features, so that there is a mismatch between the Lagrangian frame in which it is most natural to view them, and the fixed-in-space Eulerian frame of the irregular and widely-spaced networks of raingauges. The MCCs and cloud clusters themselves could be in turn be embedded within larger-scale convergence zones such as synoptic-scale lows and depressions, the most spectacular ones being tropical cyclones, and ultimately within the planetary-scale inter-tropical convergence zone. These aspects are illustrated schematically in Fig. 1.

Clearly, a large amount of statistical sampling variability is inevitable, and the daily, monthly and seasonal integration is crucial to isolate the impact of larger-scale organization and its potentially-predictable slow temporal modulation on station-scale daily rainfall characteristics. Considering daily rainfall occurrence, rather than daily totals, can be expected to filter out small-scale details of amount-variability associated with individual convective cells, thus tending to emphasize larger-scale organization. In the long-term climatological mean, differences between the stations are mostly related to fixed factors including geographical location, orography etc. Beyond these long-term differences, the “fast” variations (i.e. $<$ season), primarily associated with regional and local-scale atmospheric internal variability, and the “slow” variations (i.e. $>$ season), more related to SST and soil-moisture anomalies, contribute differently to the spatial coherence of rainfall on different time scales.

The purpose of this paper is to determine the generality of the spatial-coherence findings of MRW06 by considering four additional tropical regions, and to include analysis of daily as well as seasonal-scale rainfall coherence, bridging the gap between largely intermittent and chaotic daily rainfall fields (from a seasonal perspective) and spatially coherent and partly potentially predictable seasonal ones. The regions are NE Brazil (hereafter referred as the Nordeste), Kenya, Northwestern India, and Northern Queensland in Australia, where in each case we consider the principal monsoon season (two in the case of Kenya). These regions are considered to be fairly

spatially homogenous at least as far as the interannual variability of seasonal rainfall amounts is concerned (Camberlin and Diop, 1999 for Senegal; Parthasarathy *et al.*, 1993, 1994 for Northwestern India; Ogallo, 1989 and Indeje *et al.*, 2000 for Kenya; Moron *et al.*, 1995 for Nordeste; Lough, 1993, 1996 for Queensland). While Senegal is rather flat, the other regions are more heterogeneous as a result of contrasting orography, especially in Kenya. The five regions (including Senegal) span a range of monsoon climates affected by various meso to synoptic-scale meteorological features, such as squall lines in Senegal (e.g. Laurent *et al.*, 1998; Mathon and Laurent, 2001) and tropical depressions in India (e.g. Mooley, 1973; Mittra *et al.*, 1997), allowing us to determine how generalizable the specific results obtained over Senegal (MRW06) are in different tropical settings.

The paper proceeds as follows. Section 2 describes the data used. The analyses of spatial coherence of the interannual variability of seasonal seasonal amount, frequency of occurrence and mean intensity of rainfall during wet days in the observed station datasets are reported in section 3. The spatial scales of daily and seasonal rainfall are analyzed in section 4, with the summary and discussion in section 5.

2. Daily rainfall data

A summary of the 5 data networks is provided in Table 1. The networks range in area between ~120 000 (Nordeste) and 450 000 (Northwestern India) km² (Fig. 2). The analysis is restricted in each case to the rainy season excluding the dryer months on either side where the spatial coherence would be artificially inflated, at least at daily and short time scales.

a. Senegal

The West African monsoon season, peaking in August over Senegal, is associated with the northernmost migration of the African inter-tropical convergence zone (ITCZ). As for the entire

Sahelian belt, rainfall is mostly associated with westward moving meso-scale squall-lines (e.g. Le Barbé and Lebel, 1997; Laurent *et al.*, 1998; Mathon and Laurent, 2001). The 13 stations used in MRW06, provided by the “Direction de la Météorologie Nationale” are used here. We use daily rainfall amounts during the 92-day season, July 1 – September 30, for the 38-year period 1961-1998. Senegal is a flat country and the main heterogeneity is associated with land cover. The rain gauges are spaced fairly regularly over the country (Fig. 2a). This dataset contains no missing data though some data are doubtful. There are three non-synoptic stations, such as Diouloulou in SW Senegal that record far fewer small rainfall amounts (< 1 mm) than neighboring synoptic stations, such as Ziguinchor. There are also several unexplained very long continuous dry spells (e.g. August-September in 1996 in Kolda).

b. Northern Queensland

The summer monsoon season over Northern Queensland is centered on January and corresponds to the southernmost location of the ITCZ over the Australian continent (Lough, 1993). Rainfall is produced by various phenomena including synoptic-scale lows and depressions (e.g. Hopkins and Holland, 1997). We use daily rainfall amounts at 11 stations during the 121-day season, December 1 – March 31, for the 40-year period 1958-1998. These data were obtained from the Patched Point Dataset (PPD) (Jeffrey *et al.* 2001), and used previously in the study of Robertson *et al.* (2006). The PPD combines observed Australian Bureau of Meteorology (BoM) daily rainfall records with high quality and rigorously tested data infilling and deaccumulation of missing or accumulated rainfall. Four of the 11 stations have more than 10% of missing days infilled: station 2 (18.0%), station 4 (41.6%), station 8 (20.9%), and station 9 (10.3%). Of these, station 8 might be viewed with some caution, since it has considerable infilling and is situated in a region of orography. The rain gauges are not homogenously distributed, with 6 along or near the coast, and 5 stations in the interior (Fig. 2b). A small mountain range along the coast enhances

rainfall in the coastal stations (> 1300 mm and 60 wet days) while the interior is drier and the westernmost stations receive less than 400 mm of rain over 22-31 wet days.

c. Northwestern India

We use daily rainfall amounts at 28 stations over the state of Northwestern India largely within the states of Rajasthan and Gujarat for the 122-day summer monsoon season, June 1–September 30, for the 71-year period 1900-1970. The chosen network is located over the northwestern part of the so-called “homogeneous monsoon zone”, or continental tropical convergence zone (e.g. Sikka and Gadgil, 1980; Webster *et al.*, 1998; Gadgil, 2003). This area is affected by synoptic-scale lows and tropical depressions which form over the northern Bay of Bengal and move westward across India (e.g. Mooley, 1973; Mitra *et al.*, 1997; Goswami *et al.*, 2003). These data were taken from the Global Daily Climatology Network (GDCN), archived at the National Climatological Data Center (NCDC). The stations are fairly regularly spaced (Fig. 2c) and the terrain is rather flat. There are nominally 0.07% of missing daily values. Seasonal rainfall totals range from 200 mm in the north to 1200 mm in the southeast (not shown). The number of wet day (receiving more than 0 mm of rainfall) varies accordingly from 10-20 days at the driest stations, to 40-65 days at the wettest ones.

d. The Nordeste

We use daily rainfall amounts recorded at 81 stations over Ceara state, that is the northernmost part of Brazilian Nordeste, for the 89-day (90-day for leap years) February 1 – April 29 (FMA) rainfall season, over the 29-year period 1974-2002. This region receives its highest rainfall in March-April when the ITCZ is at its most southerly position (Ratisbona, 1976; Hastenrath and Heller, 1977). Rainfall is produced by various phenomena, such as slow westward-propagating depressions (Ramos, 1975), and northward-moving cold fronts which enhance convective activity across Nordeste (Kousky, 1979, 1985). This data comes from a larger dataset of 700 stations

provided by FUNCEME (Fundacao Cearesense de Meteorologia). The 81 stations having more than 90% of daily values present are retained in this study. There are 4.49% of missing days. This is the densest network (Fig. 2d and Table 1) analyzed. The orography in this area is relatively smooth with highest peaks below 600 m. Seasonal rainfall totals range from less than 350 mm in the extreme southwest to more than 1000 mm in the northwest. The driest area, on a SW-NE diagonal, receives less than 500 mm during FMA. The number of wet days receiving more than 0 mm of rainfall is typically between 23 and 35, and reaches 50 at the wettest locations (not shown).

e. Kenya

We use daily rainfall amounts at 9 stations (Fig. 2e) for the 92-day March 1 – May 31 and October 1 – December 31 rainfall seasons, over the 39-year period 1960-1998. There are two rainy seasons associated with the passage of the ITCZ over the country (Ogallo, 1985; Nicholson, 1996). However, the ITCZ is very diffuse in the region. Partly due to the complex topography, organized weather systems are not common, and local convection, combined with orographical effects, plays a major role in the space-time distribution of rainfall (Ogallo, 1985; Nicholson, 1996). The dataset was assembled from two sets provided by the Kenyan Meteorological Department. The daily data were checked for errors against “monthly weather summaries” since spurious, extremely high daily amounts are found for Lodwar, Narok and Moyale (with more than 10 times the monthly mean). These errors, were found to correspond to a multiplication by a factor of 10 or 100, and were corrected to match the monthly amounts. A few missing values remain (respectively 1.7% and 2.8% of missing values in MAM and OND) but are concentrated during a few years and stations. Kenya is mountainous and this network is, by far, the least homogeneous of those analyzed in this study. Given the complex orography, this network (Fig. 2e) is certainly unable to sample the whole range of variability, but the Kenyan case is particularly interesting since it enables to compare two contrasted rainy seasons in a given region.

During the MAM “long rains”, rainfall is low in the northwest (< 100 mm in Lodwar) and varies between 280 and 560 mm elsewhere, reaching a maximum in the southwest and along the Indian Ocean coast. The occurrence frequency of wet days is 12 in Lodwar and 31-55 elsewhere. During the OND “short rains”, rainfall amounts are lower but less variable, between 140 mm in Narok and 300 mm in Kisumu (except < 50 mm for Lodwar) and the number of wet days is between 6 (Lodwar) and 40 (Kisumu) as typically between 20 and 30.

f. Treatment of missing values

Missing entries over the Nordeste, Kenya, Northwestern India were filled using a simple stochastic generator applied independently to each station (Wilks, 1999) before the following analyses. If a whole month is missing, this method creates a daily time series where the amplitude of daily amount as well as the persistence of dry and wet days are consistent with the long-term mean. This simple scheme will underestimate any spatial coherence at the daily time scale, but the small fraction of missing data (at most 4.5% for the networks where the missing entries are not filled *a priori*) limits this bias. Moreover, the missing values tend to be scattered randomly in time and thus their impact on seasonal quantities is small.

3. Estimates of spatial coherence

a. Degrees of freedom and relationships between seasonal amount, frequency of occurrence and mean intensity

In their study of the Senegal network, MRW06 proposed that, while interannual variability of seasonal amount (S) at each station is accounted for rather equally by year-to-year changes in the frequency of occurrence (O) and mean intensity (I) (see Fig. 2 of MRW06), the predictable component of the interannual variability of S is largely restricted to O , with I being essentially unpredictable. Here O is simply estimated as the relative frequency of daily rainfall > 1 mm, and

$I=S/O$ is the mean rainfall amount on wet days. Sensitivity to the wet-day threshold is investigated in Sect. 3c.

We now explore this hypothesis across the 6 seasonal networks at interannual scale. First, the component of interannual variability of S that is linearly related to that of O is extracted using linear regression, computed independently at each station;

$$(1) \quad S_{occ}=\alpha O+\beta$$

We then compute the correlation between the residuals of the regression ($S_{occ}-S$) and the interannual variability of I at each station. The network average of the squared correlation (i.e. the common variance) is shown for each region in Table 2, together with an estimate of the number of spatial degrees of freedom (DOF) for each variable. The DOF (Der Mégrédichtian, 1979; Moron, 1994; Fraedrich *et al.*, 1995; Bretherton *et al.*, 1999; MRW06) gives an empirical estimate of the spatial coherence in terms of empirical orthogonal functions (EOFs), with higher values denoting lower spatial coherence. Further details are given in the appendix.

As in the case of Senegal in MRW06, the correlations between S and O on one hand and between S and I on the other hand are both positive and usually high. In consequence, the network averages of the common variance between S and I are usually comparable to those between S and O across all the networks. In contrast, the common variance between O and I is usually close to zero, except for Kenya in MAM (Table 2) and the DOF of I and $S_{occ}-S$ are also consistently much larger than the DOF of O , again supporting the hypothesis of MRW06 that the spatially coherent part of the interannual variability of S is mainly derived from that of O .

The DOF of interannual variability of S , O and I over Kenya are plotted as a function of the number (M) of stations considered in Fig. 3, with the two seasons (MAM and OND) considered

separately. Every combination of M stations is considered, with M varying from 2 to 9. Similar computations have been made for every network (not shown). Again the results support those reported by MWR06 for Senegal: spatial coherence at interannual scale is moderate-to-strong for the seasonal amount (Fig. 3a) and rainfall occurrence (Fig. 3b), but weak for the mean intensity of rainfall (Fig. 3c). Over Kenya, the spatial coherence is stronger during OND than MAM, which is consistent with the well-known larger potential predictability during OND (e.g. Rowell *et al.*, 1994; Camberlin and Philippon, 2002; Philippon *et al.*, 2003).

b. Robustness to methodological considerations

The comparison between the DOF estimates of different networks is not bias free, particularly when the spatial coherence is low (that is when DOF is high). This bias is related to the fact that DOF is bounded by the rank of the correlation matrix, which is its smallest dimension (see appendix). An alternative measure of the spatial coherence of interannual anomalies is provided the interannual variance of station-averaged (standardized) anomalies – the standardized anomaly index (SAI , see appendix)

The DOF and $var[SAI]$ estimates of the interannual variability of S , O and I are plotted for each network in Fig. 4. There is a near-linear relationship between both estimates for moderate and higher values of spatial coherence [$DOF < \sim 6$ and $var[SAI] > \sim 0.2$]. Thus, both measures of spatial coherence are consistent with each other, except when spatial coherence is low, as is the case for I where the estimate becomes bounded by the dimensions of the network. Figure 4 also shows that the spatial coherence of I is always far weaker than the one of O and S .

Note that the variation of the spatial coherence of O across the networks does not depend significantly on the climatological frequency of occurrence itself. The percentage of wet days (i.e. receiving > 1 mm) varies between 14.3% (Kenya in OND) and 39.4% (Nordeste). The Spearman

rank correlation between these percentages and *DOF* of the frequency of occurrence is 0.14. Similarly, the rank correlation between the mean seasonal amount of each network and the *DOF* of the seasonal amount is -0.37 . Both correlations are not significant at the one-sided 90% significance level.

c. Robustness to rainfall thresholds

It is instructive to test the dependence of *DOF* on the wet-day threshold, and the amplitude of daily rainfall amounts. Measurements errors, which are almost innately spatially-independent, as well as the threshold considered for recording rain (e.g. 0.1, 0.5 or 1 mm) could both induce noise in the different data sets, especially if these thresholds are variable in time and/or in space. It is also interesting to analyze the impact of large daily amounts on the spatial coherence of *S*.

The robustness of *DOF* values is studied by (i) changing the threshold for defining wet days between 0 and 10 mm; and (ii) analyzing the seasonal sums of truncated daily rainfall values. The seasonal sum $S(z)$ of truncated daily rainfall for a given cut-off t is defined as (Snijders, 1986)

$$(2) \quad S(z) = \sum_{i=1}^n \min\{r_i, z\},$$

where r_i is the rainfall recorded on day i . The seasonal sum $S(z)$ is thus insensitive to the magnitudes of rainfall amounts above t . If $t = \infty$, $S(z)$ is simply the seasonal amount (Snijders, 1986).

The impact of wet-day threshold amount on the *DOF* for the interannual variability of *O* is plotted in Fig. 5a, with the impact of rainfall amount truncation on the *DOF* of the interannual variability of *S* plotted in Fig. 5b. In general, both impacts are rather weak. This is especially true for Queensland, Nordeste and Kenya-OND, which are the three networks with the strongest spatial coherence values (Table 2). For Northwestern India and Senegal, the *DOF* of *O* is fairly

constant for wet-day thresholds below 5 mm, but then increases for higher thresholds (Fig. 5a). A 0 mm threshold clearly adds noise in the Senegal and Northwestern India cases. Very large daily amounts of rainfall (i.e. > 50 mm) also markedly increase the noise in S for Northwestern India (Fig. 5b). The fraction of wet days (> 0 mm) receiving 50 mm or more is 9.2% for Queensland, 7.9% for Northwestern India, and less than 5% for the other networks. For Queensland and Northwestern India, these high-rainfall days account for 46.2% and 40.0% of seasonal rainfall respectively. However, the DOF of the frequency of occurrence of high-rainfall days are very different : 4.2 for Queensland and 12.1 for Northwestern India. respectively. Although the sample sizes of these high-rainfall events are relatively small, this suggests that they occur much more randomly in space and time in Northwestern India than they do in Queensland, and that seasonal anomalies of the latter are much more potentially predictable than the former.

4. Spatial coherence and spatial scales of rainfall

In order to understand better the spatial coherence of seasonally-averaged rainfall, we examine next the spatial scales of rainfall variability at daily-to-seasonal time scales. The spatial scales of daily rainfall are estimated using spatial autocorrelation of rainfall intensity (i.e. wet-day amount, Bacchi and Kottegoda, 1995) and the probability of rainfall occurrence. The spatial correlation function of seasonal anomalies and the effect of the temporal integration across the season are then analyzed. Very similar results are obtained when the spatial correlation calculations are repeated using only pairs of stations having both data values present in the original data set (not shown).

a. Spatial scales of daily rainfall intensity

The spatial correlation of daily rainfall intensity is computed for each pair of stations, using only wet days (> 1 mm) at both stations. The spatial scale is usually defined as the de-correlation distance, τ_D , at which the correlation falls to $1/e = 0.37$ (Dai *et al.*, 1997; New *et al.*, 2000;

Ricciardulli and Sardeshmukh, 2002; Smith *et al.*, 2005). However, the low station densities of our networks do not allow a precise definition of τ_D , except perhaps for the Nordeste. The analysis is therefore restricted to a general assessment of the relationship between inter-station correlations and distances.

The Pearson correlation between each station pair is plotted against distance in Fig. 6 for each network, with average values denoted by circles, and the $1/e$ value dashed. All the networks exhibit a general exponential decay of correlation with increasing distance (Ricciardulli and Sardeshmukh, 2002; Smith *et al.*, 2005). Correlations lie well below $1/e$ for distances > 100 km for all networks, with considerable differences between them for shorter distances. Thus, 100 km can be considered as an upper bound on τ_D for daily rainfall amounts, with smaller τ_D for the dense Nordeste network (Fig. 6d). This rough estimate is consistent with the general results of Ricciardulli and Sardeshmukh (2002) and Smith *et al.* (2005) using satellite measurements of deep convection and rainfall at the 3-hourly time scale. Note that considering the square root of daily amount to reduce skewness leads to very similar results (Fig. 6h).

b. Spatial scales of daily rainfall occurrence

The spatial autocorrelation is computed for each pair of stations using the frequency of occurrence coded as 0 for dry and 1 for wet day. The linear correlation between two binary variables is known as the phi correlation (ϕ) and can be computed in terms of simple frequency counts of $[0,0]=A$, $[1,1]=B$, $[0,1]=C$ and $[1,0]=D$ (Garson, 1982):

$$(3) \quad \phi = \frac{AB - CD}{\sqrt{(A+C)(A+D)(D+B)(C+B)}}.$$

Figure 7 shows plots of ϕ for each network. The correlations are larger than for rainfall intensity (Fig. 6) for every network, and are closer to a linear decay of correlation with increased distance. Differences between the networks are somewhat larger than for intensities. The ϕ -values are

relatively small in the Nordeste (Fig. 7d) while they are larger for Senegal (Fig. 7a) and Northwestern India (Fig. 7c). Organized MCCs, cloud-clusters or synoptic-scale meteorological phenomena, for example in the form of squall-lines over Senegal and tropical depressions over Northwestern India, may be associated with the larger occurrence correlations there, as compared to the Nordeste. Note that the ϕ correlation of rainfall occurrences is not necessarily larger than the Pearson correlation between rainfall amounts. The former depends ultimately on the probability of similar characteristics, wet or dry, at both stations; for example, two stations having 50% (respectively 75%) of common dry or wet days will have a ϕ correlation near 0 (respectively near 0.5) if both stations have roughly the same probability of wet days. For a pair of stations having a ϕ correlation of 0.5, the correlation between intensities will be closer to 1 if rainfall intensities are broadly ranked in the same order for both stations (not shown).

c. Impact of temporal integration

The previous two sub-sections show that daily intensities are largely spatially-incoherent except at short distances, while daily occurrence tends to exhibit larger spatial scales. To tie these findings to our previous seasonal-scale results, we next investigate the impact of temporal integration on the spatial correlation function of daily rainy events from daily to seasonal time scale.

Figure 8 displays the spatial correlation functions for the frequency of occurrence (left), mean intensity (center) and amount (right), across the networks. The results are shown for averaging periods of 2, 5, 10, 30 days, in addition to the daily and seasonal-average values. By our definition, the spatial correlation function is identical for amount and mean intensity at the daily time scale, while the amount calculations include dry days for averaging periods of 2 days or more.

Frequency-of-occurrence spatial correlation values (Fig. 8, left) generally increase quite regularly with temporal averaging period. These increases are almost independent of inter-station distance, and are fairly similar for all the networks. Thus, temporal integration acts to increase spatial coherence across the entire network, and this is true for all the networks; it strongly suggests a common network-scale climate forcing on occurrence frequency. On the other hand, the near-linear shape of the spatial correlation function of daily rainfall clearly survives the temporal integration, so that the seasonal-scale autocorrelation function resembles a superposition of daily and seasonal effects. Note that the largest autocorrelations are not necessarily achieved at the seasonal scale, but at monthly time scales in the case of Northwestern India, and less clearly in Kenya (MAM) and Senegal. Monthly values are likely to be inflated by spatial coherence associated with a strong seasonal cycle. This has been evaluated by removing an estimated seasonal cycle from the daily occurrence and amount data. The mean seasonal cycle was computed on a daily basis by averaging across years and low-pass filtering to remove periods shorter than 30 days. The resulting spatial correlation functions now increase quite regularly from daily to seasonal scale with the largest spatial correlations occurring on the seasonal scale (not shown). Over Northwestern India and Kenya in MAM, where the monthly and seasonal values are almost identical (not shown).

In stark contrast to rainfall frequency, there is very little impact of temporal averaging on mean intensity (Fig. 8, center), except for Kenya in OND, and to a small extent, Queensland. A spatially coherent signal in the mean intensity cannot be excluded in these cases, though it is much weaker than for frequency. The picture for rainfall amount (Fig. 8, right) is closer to that of occurrence frequency, with a rather regular increase of correlation with time scale.

The seasonal scale spatial correlations are compared among networks in Fig. 9, for amount and occurrence. There are large differences between the networks, with the spread larger for amount (Fig. 9a). The ranking of the curves in Fig. 9 is highly consistent with the *DOF* values in Table 2, with amounts being generally less spatially coherent.

In terms of temporal integration, the spatial correlations of amount at the daily scale are lower than for occurrence, and this difference tends to persist to the seasonal scale (Fig. 9). However, the Nordeste network exhibits the smallest spatial correlations between daily amounts (Fig. 6d) and frequency of occurrences (Fig. 7d), yet is among the most coherent at the seasonal scale (Fig. 9a,b). Another example is that the spatial scales of Kenyan daily rainfall are not obviously different in the MAM and OND seasons (Fig. 6-7e,f), while a larger spatial coherence is clearly seen at seasonal scale in OND (Fig. 9a,b). This suggests that spatial coherence of seasonal anomalies emerges mostly through a persistent modulation of the quasi-random daily rainfall rather than an organized pattern appearing repeatedly across the season (Krishnamurthy and Shukla, 2000). The ‘randomness’ of daily rainfall should be understood not necessarily as random size or patterns, but rather as random tracks of rainy systems across the network.

d. Signal-to-noise interpretation

We return now to the simple statistical model of MRW06 in which interannual variability of rainfall anomalies at regional-scale was decomposed into a spatially-uniform signal, related to external forcing(s), plus a spatially-independent noise, resulting from small-scale processes and statistical sampling variations. Details of the model are provided for completeness in the appendix. This model leads to a constant spatial correlation function that is independent of distance, given by the contribution of signal to the total variance, while the dispersion of the correlations between pairs of stations is solely due to random sampling of the noise component. As it stands, this model cannot account for the decrease in spatial correlation with distance shown

in Fig. 9. We now demonstrate that a straightforward extension of the model to include a spatial modulation of the signal-to-noise ratio is able to account for the quasi-linear decrease of spatial correlation of occurrence frequency with distance.

We have created 100 random simulations of seasonal rainfall occurrence O for each network according to equation A.5, with the proportion of the signal being set to match the external variance ratio estimated from the observed network (MRW06). The noise component is given by station independent random white noise. The amplitude of the noise at a given station is then weighted by the mean distance (w_i) between this station and all others. The weights w_i , are standardized so that their mean equals one. The complementary weight ($1 - w_i$) is applied to the signal. Thus isolated stations get more noise, and the signal-to-noise ratio is largest for stations that are bunched together. By construction, stations that are closer together will feel the common signal more strongly and thus be more highly correlated, while widely spaced station pairs will experience a larger noise fraction and thus be correlated more weakly.

Results from the Queensland and Northwestern India networks are shown in Fig. 10 for seasonal frequency of occurrence of rainfall. In both cases the statistical model can qualitatively account for the observed autocorrelation behavior, with a correlation decay distance near 700-800 km. Thus, the decrease in inter-station correlation with distance is not inconsistent with the simple conceptual model of the addition of a spatially-uniform signal plus a spatially-independent noise where the amplitude of signal and noise is spatially modulated. The difference between Queensland and Northwestern India is mainly due to the amount of signal, computed using the external variance ratio (MRW06), of 67% for Queensland and 58% for Northwestern India. Note that the random variability associated with stochastic sampling is large, even with the long records available in Queensland and Northwestern India. Similar results were obtained for the other networks (not shown).

5. Summary and discussion

a. Summary

In this paper we have extended the analysis of spatial coherence of daily station rainfall characteristics, carried out by MRW06 over Senegal, to consider the principal rainy seasons of six different station networks – Senegal, the Brazilian Nordeste, Northern Queensland, Northwestern India, and Kenya (with two rainy seasons). They share similar geographical extents (120 000 - 450 000 km²), but the station density (Fig. 2), as well as topography and the seasonal evolution of rainfall differ among them. The primary rainfall statistics analyzed were spatial coherence of interannual anomalies in seasonal amount (S), wet-day occurrence frequency (O), and the mean intensity of rainfall on wet days ($I=S/O$).

Our principal results generalize the previous findings of MRW06 for Senegal.

- Over all the six tropical rainy seasons considered, seasonal anomalies of O and S were found to be much more spatially coherent than mean daily intensity (Figs. 3-4 and Table 2).
- The spatial coherence of O was found to be quite similar across the 6 rainy seasons, while S showed more variation, being higher for Queensland, the Nordeste and Kenya-OND, and lower for Northwestern India, Kenya-MAM, and Senegal (Fig. 4 and Table 2).
- While the spatial coherence of I was found to be considerably lower than that of O in all the networks, the Kenya and Queensland networks exhibited larger spatial coherence of I , particularly over Kenya during the OND season (Table 2).
- Although interannual variability in S was found to be almost equally accounted for by I and O at station scale across all networks, the part of variance of S that is linearly

unrelated to O was found to be spatially incoherent and almost fully attributable to I in all the networks (Table 2).

A second set of results addresses the relationship between the spatial coherence and the scales of rainfall, together with their behavior on daily-to-seasonal time scales. These were expressed in terms of the spatial autocorrelation function between station rainfall occurrences and intensity.

- The spatial autocorrelation functions of daily rainfall occurrence probability and daily rainfall intensity were both found to be consistent from network to network (Figs. 6-7). Daily intensities were found to exhibit an exponential decay with distance and to be virtually uncorrelated beyond about 100 km; differences amongst the networks below this distance could not be assessed accurately because of insufficient station density of our networks.
- The spatial scales of daily rainfall occurrence probability were found to be larger than for intensity with a more linear spatial decay; they showed greater differences between the networks (Fig. 7). Average spatial scales were found to be larger for Northwestern India than for the Nordeste, possibly because of the impact of synoptic-scale lows and depressions over Northwestern India.
- At the seasonal scale, larger differences were found in rainfall scales between networks, particularly for seasonal amount (Fig. 9). These differences were found to be highly consistent with estimates of spatial coherence in terms of degrees of freedom.
- Temporal integration was found to act very consistently on O and amount, while having almost no impact on I (Fig. 8).
- The shape of the spatial autocorrelation function of seasonal rainfall frequency can be recovered from a model of a spatially-uniform signal added to a spatially-independent

noise, provided that a realistic variation of the amplitude of signal and noise with geographical distances is included (Fig. 10).

A third set of general results is methodological: the spatial coherence was not found to be severely biased by the exact formulation of the statistical estimate, the definition of wet days, and the *a priori* geographical definition of the regions.

- Estimates of spatial coherence in terms of spatial degrees of freedom (*DOF*) and variance of the standardized anomaly index ($\text{var}[SAI]$) were found to be consistent with each other across all regions (Fig. 4).
- Estimates of spatial coherence are least robust when it is weak, where they become sensitive to the number of stations on the network. This is only the case for *I*, for which a denser network would be expected to yield a more precise and reliable estimate of the spatial coherence.
- The spatial coherence of *S* and *O* is usually robust with respect to the threshold used to define wet days, as well as truncation of large daily rainfall amounts (Fig. 5). Over Northwestern India, very large daily amounts tend to decrease the spatial coherence of *S* (Table 2 and Fig. 5). In most cases, very small daily rainfall amounts (between 0 and 1 mm) are an additional source of noise and thus lower the spatial coherence (Fig. 5a).

b. Discussion

The main message of this paper is that daily rainfall properties are remarkably similar in different continental monsoonal regions. Interannual station anomalies in rainfall frequency and seasonal amount are always much more spatially coherent than those of mean rainfall intensity. The spatial coherence of seasonal rainfall totals largely reflects that of rainfall frequency. Seasonal anomalies in intensity are found to be relatively incoherent in all of the 6 networks and their contribution to spatial coherence of seasonal amounts is secondary because the *DOF* and $\text{var}[SAI]$ estimates of

the latter are relatively insensitive to additive spatial noise; they effectively isolate the spatially coherent part.

At the daily scale, rainfall probability is found to be always more spatially coherent than rainfall intensity. The difference is especially large in regions of large-scale rainfall organization, i.e. the intraseasonal oscillation and monsoon depressions over Northwestern India, and African easterly waves and/or squall lines over Senegal. However, despite geographical and meteorological differences between the networks, temporal integration acts in a remarkably similar way across all of them; its impact is generally more pronounced on amount than occurrence, while being very small on intensity in most cases.

The simplest interpretation is that large-scale climate forcing acts across the season on the occurrence, while intensity can largely be regarded as a random process. The general lack of spatial coherence of intensity does not exclude any modulation at interannual scale, but simply means that this modulation is not spatially coherent across the network. Our results are consistent with previous findings over the Sahel (Le Barbé and Lebel, 1997; Lebel *et al.*, 2003) and India (Gadgil, 2003), and help to interpret the low predictability of rainfall intensity over North Queensland documented by Robertson *et al.* (2006) and over Senegal documented in MRW06. Although small, some spatial coherence of intensity over Kenya in OND, and perhaps Queensland was found which could be of potential significance.

Station daily rainfall intensity (i.e. the amount recorded on wet days) tends to be strongly influenced by the smallest spatial scales of the precipitation process, i.e. the cumulus scale (e.g. Houze and Chen, 1977; Lopez 1978; Chen *et al.*, 1996; Rickenbach and Rutledge, 1998). In contrast, daily rainfall probability tends to reflect larger-scale organization, from cloud clusters, synoptic-scale disturbances to large-scale tropical convergence zones. The respective spatial

correlation functions at the daily scale are thus found to be distinctly different, with a much more rapid fall-off for intensity than for occurrence across all networks (Fig. 6 vs. Fig. 7). Temporal integration of the highly-skewed daily intensity distribution is sensitive to outliers and thus produces seasonal-average intensities that are still largely spatially incoherent (Fig. 8). This is consistent with Mooley (1973) who showed that the amount of daily rainfall associated with synoptic-scale depressions over India displayed a large spatial variability, even if averaged rainfall over a large number of events is considered (see also Mitra *et al.*, 1997). Similar findings have been reported for West-African MCCs (e.g. Abdou *et al.*, 2003; Lebel *et al.*, 2003; Balme *et al.*, 2007). Previous studies have not analyzed rainfall occurrence in the same way, but our results suggest that temporal integration of daily rainfall probability yields to a substantial spatial coherence of station-scale rainfall frequencies and amounts at the seasonal scale. This occurs for all the networks, even in the case of Kenya in MAM where the association with ENSO is weak.

Each tropical region is subject to its own complex mixture of various meteorological phenomena. Studies point to the larger impact of MCCs and synoptic scales over Northwestern India and Senegal, compared to Kenya and the Nordeste, where radar echoes are especially small (Da Silva Aragao *et al.*, 2000). Propagation speed and areal extent of disturbances will also play a role. It is well-known that the intra-seasonal oscillation over India organizes the occurrence and path of monsoon depressions, and influences seasonal monsoon rainfall totals (e.g. Hartmann and Michelsen, 1989; Gadgil and Joseph, 2003; Goswami *et al.*, 2003; Greene *et al.*, 2007).

Each of the station networks considered was chosen a priori, largely from data availability considerations. This could potentially lead to an underestimation of the true spatial coherence, if the networks encompass homogeneous sub-regions that behave independently. We have examined this issue in terms of the degree to which coherent sub regions can be identified from EOF analysis. However, the results (not shown) suggest that this is not a serious issue for any of

the networks considered, except perhaps for Kenya where both stations of Malindi are rather independent from the interior.

One striking result is the different spatial coherence observed over Kenya where the OND season shows a distinctly greater coherence than during MAM (Table 2). Organised disturbances are seldom encountered, and there is little evidence of a significant change in the weather systems and/or in size of the area affected simultaneously by rain, between the two seasons. In fact, the spatial scales at daily time scale are hardly different between OND and MAM (Fig. 6e,f and 7e,f). The greater OND spatial coherence is associated with a robust seasonal teleconnection with ENSO (Ogallo, 1989; Hastenrath, 2000; Mutai and Ward, 2000; Camberlin and Philippon, 2002). An important implication of our study is that the seasonal predictability of station-scale rainfall is likely to be enhanced quite generally by considering rainfall frequency in place of seasonal rainfall total. This may be of particular relevance to agriculture, and underlines the need for daily station data.

Our results suggest that the analysis of the spatial coherence of observed fields is a valuable tool to infer properties of the underlying processes that control the spatio-temporal variability. Previous work suggests that potential predictability may be large at specific time scales, such as for the intraseasonal oscillation over India in summer (e.g. Webster *et al.*, 1998; Gadgil and Joseph, 2003) and the analysis of spatial coherence at these time scales deserves further investigation. The five networks considered here are continental, and it would be instructive to consider maritime regions as well.

Appendix A: Definitions and empirical estimates of spatial coherence

Two scores are used to provide empirical estimates of the spatial coherence of seasonal anomalies between stations: the interannual variance of the standardized anomaly index (Katz and Glantz, 1986), and the number of spatial degrees of freedom (Der M gr dichtian, 1979; Fraedrich *et al.*, 1993 ; Bretherton *et al.*, 1999). We write the individual station time series of rainfall amount (S), occurrence (O) and mean intensity (I) as x_{ij} , where $i = 1 \dots N$ denotes the year and $j = 1 \dots M$ denotes the station, and the $N \times M$ matrices of S , O and I as \mathbf{X} . These are firstly normalized to zero mean and unit variance

$$(A1) \quad x'_{ij} = \frac{(x_{ij} - \bar{x}_j)}{\sigma_j}$$

where \bar{x}_j is the long-term time mean and σ_j is the interannual standard deviation for station j .

The standardized anomaly index (SAI) is defined as the average of the normalized station time series of seasonal averages over the M stations (Katz and Glantz, 1986) ;

$$(A2) \quad SAI_i = \frac{1}{M} \sum_{j=1}^M x'_{ij}$$

The interannual variance of the SAI ($var[SAI]$) is a measure of the spatial coherence since it depends on the inter-station correlations (ρ_{ij}). Note that $var[SAI]$ reaches a value of 1 when all M stations are perfectly correlated. If all stations are assumed independent of each other (i.e. their mean correlation equals zero), its value ($= 1/M$) ranges from 0.0123 (Nordeste) to 0.11 (Kenya).

The number of degrees of freedom (DOF) can be estimated through an eigen-analysis, i.e. empirical orthogonal function (EOF) analysis, of the correlation matrix formed from the station seasonal-mean time series:

$$(A3) \quad DOF = \frac{M^2}{\sum_{i=1}^M e_i^2},$$

where e_i are the eigenvalues of the correlation matrix. In the limiting case of $e_i = 1$ for all stations, $DOF = M$ (or $DOF = N-1$ if $N < M$) i.e. each station conveys independent information and the common “signal” is zero. On the other hand, if a single eigenvalue accounts for all variance of the field (= trace of the correlation matrix, that is M), then $DOF = 1$; i.e. each station conveys the same information equal to the “signal” and the noise is zero. In the latter case, the station network can be described by a single EOF.

Both scores are linear and thus sensitive on large deviations from a Gaussian distribution. This is not expected to be a serious issue for the rainy seasons considered in this paper. However, a Box-Cox transform (Box and Cox, 1964) can be used to obtain a more nearly Gaussian distribution; the transformation from x to $x(\lambda)$ is defined only for positive values and is given by

$$(A4) \quad x(\lambda) = \frac{x^\lambda - 1}{\lambda} \text{ if } \lambda \neq 0 \text{ and } x(\lambda) = \ln(x) \text{ if } \lambda = 0.$$

The estimates of spatial coherence of seasonal amount, frequency of occurrence and daily mean intensity of rainfall shown in the paper have been re-computed after having applied the Box-Cox transform (Box and Cox, 1964) to the raw data. This leads to very similar results (not shown).

The spatio-temporal variability of the rainfall variable \mathbf{X} , recorded over a regional (i.e. roughly 10^5 - 10^6 km²) window can be considered to result from the complex interactions between multi-scale fixed (e.g. topography) and variable (e.g. SST) boundary forcings, together with the internal dynamics of the atmosphere. A simple conceptual model of the regional-scale seasonal anomalies relative to long-term mean can be constructed following MRW06 by adding a spatially uniform “signal” (\mathbf{C}) to a spatially independent “noise” (\mathbf{N})

$$(A5) \quad \mathbf{X} = \mathbf{C} + \mathbf{N}.$$

Where the M columns of \mathbf{C} each contain the single signal time series and \mathbf{N} contains a set of M independent white noise random time series. Here \mathbf{C} can be interpreted as the integration of large-

scale variable forcings while N represents all small-scale variable forcings as well as uncertainties associated with sampling and recording.

In this context, it is possible for $var[SAI]$ to be zero if half the stations are perfectly anti-correlated with the others. Similarly, it is theoretically possible to have $DOF=1$ in the pathological case where neighboring stations are anti-correlated with each other, in a “noodled” pattern that is anything but spatially coherent in the usual sense. Thus, an inverse quasi-linear relationship is generally expected between DOF and $var[SAI]$, except in the case of a “noodled” pattern where $DOF=1$ and $var[SAI]=0$.

Acknowledgements

We acknowledge Matayo Indeje (IRI and the Kenyan Meteorological Department) and the Kenyan Meteorological Department for providing us the Kenyan rainfall, Ousmane Ndiaye (IRI and Direction de la Météorologie Nationale) for the Senegal rainfall, Joshua Qian (IRI) for compiling Nordeste data, and Arthur Greene (IRI) for compiling the Indian data. The Queensland Centre for Climate Applications supplied the Queensland rainfall data. We thank James Hansen and Arthur Greene for discussions at various stages of this work. We thank also the editor (David Straus) and the three anonymous reviewers whose constructive comments lead to a substantially improved manuscript. This research was supported by NOAA through a block grant to the International Research Institute for Climate and Society (VM, AWR, MNW), and by Department of Energy grant DE-FG02-02ER63413 (AWR).

References

Abdou A., T. Lebel, and A. Amani, 2003: Invariance in the spatial structure of Sahelian rain fields at climatological scales. *J. Hydrometeorology*, **4**, 996-1011.

Bacchi B., and N.T. Kottegoda, 1995: Identification and calibration of spatial correlation patterns of rainfall. *J. Hydrology*, **165**, 311-346.

Balme M., T. Vischel, T. Lebel, C. Peugeot, and S. Galle, 2007: Assessing the water balance in the Sahel: Impact of small scale rainfall variability on runoff. Part 1 : rainfall variability analysis. *J. Hydrology*, **331**, 336-348.

Baron C., B. Sultan, M. Balme, B. Sarr, S. Traore, T. Lebel, S. Janicot, and M. Dingkuhn, 2005. From GCM grid cell to agricultural plot: scale issues affecting modelling of climate impact. *Philosophical Trans. Roy Society B*, **360**, 2095-2108.

Box G.E.P., and D.R. Cox, 1964: An analysis of transformations, *J. Royal Stat. Society*, **127**, 211-252.

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

Camberlin P., and M. Diop, 1999: Inter-relationships between groundnut yield in Senegal, interannual rainfall variability and sea surface temperature. *Theor. App. Climatol.*, **63**, 163-181.

Camberlin P., and N. Philippon, 2002: The East African March-May rainy season: Associated atmospheric dynamics and predictability over the 1968-97 period. *J. Climate*, **15**, 1002-1019.

Camberlin P., and O. Planchon, 1997: Coastal precipitation regimes in Kenya. *Geografiska Annaler*, **79A**, 1-2, 109-119.

Chen S.S., Houze Jr R.A., and B.E. Mapes, 1996: Multiscale variability of deep convection in relation to large-scale circulation in TOGA COARE. *J. Atmos. Sci.*, **23**, 1380-1409.

Dai A., I.Y. Fung, and A.D. del Genio, 1997: Surface Global observed land precipitations variations: 1900-1988. *J. Climate*, **10**, 2943-2962.

Da Silva Aragão M.R., Correia M.D.F., and H.A. de Araujo, 2000: Characteristics of C-Band meteorological radar echoes at Petrolina, Northeast Brazil. *Int. J. Climatol.*, **20**, 279-298.

Der Mégrédichtian G., 1979: L'optimisation du réseau d'observation des champs météorologiques. *La Météorologie*, **VI^{ème} serie**, 51-66.

Fraedrich K., C. Ziehmann, and F. Sielmann, 1995: Estimates of spatial degrees of freedom. *J. Climate*, **8**, 361-369.

Gadgil S., 2003: The Indian monsoon and its variability. *Annu. Rev. Earth. Planet. Sci.*, **31**, 429-467.

Gadgil S., and P.V. Joseph, 2003: On breaks of Indian monsoon. *Proc. Indian Acad. Sci. (Earth Planet Sci.)*, **112**, 529-558.

Garson G.D., 1982: *Political Science Method*, Hollbrook, Boston, 288 p.

Goddard L., S.J. Mason, S.E. Zebiak, C.F. Ropelewski, R. Basher, and M.A. Cane, 2001: Current approaches to seasonal to interannual climate predictions, *Int. J. Climatol.*, **21**, 1111-1152.

Gong X., A.G. Barnston, and M.N. Ward, 2003: The effect of spatial aggregation on the skill of seasonal rainfall forecasts. *J. Climate*, **16**, 3059-3071.

Goswami B.N., Ajayamohan R.S., Xavier P.K., and D. Sengupta, 2003: Clustering of synoptic activity by Indian summer monsoon intraseasonal oscillations. *Geophys. Res. Letters*, **30**, 1431.

Greene A., A.W. Robertson, and S. Kirshner, 2007: Analysis of Indian monsoon daily rainfall on subseasonal to multidecadal time scales using a hidden Markov model. *Quart. J. Meteo. Soc.*, **128**, in press.

Haylock M., and J. McBride, 2001: Spatial coherence and predictability of Indonesian wet season rainfall. *J. Climate*, **14**, 3882-3887.

Hansen J., A. Challinor, A. Ines, T. Wheeler, and V. Moron, 2006: Translating climate forecasts into agricultural terms: advances and challenges. *Climate Research*, **33**, 27-41.

Hartmann D.L., and M.L. Michelsen, 1989: Intraseasonal periodicities in Indian rainfall. *J. Atmos. Sci.*, **46**, 2838-2862.

Hastenrath S., and L. Heller, 1977: Dynamics of climatic hazards in NorthEast Brazil, *Quart. J. Meteo. Soc.*, **103**, 77-92.

Hastenrath S., 2000: Zonal circulations over the Equatorial Indian Ocean. *J. Climate*, **13**, 2746-2756.

Houze Jr R.A., and C.P. Cheng, 1977: Radar characteristics of tropical convection observed during GATE: mean properties and trends over the summer season. *Mon. Wea. Rev.*, **105**, 964-980.

Hopkins L.C., and G.J. Holland, 1997: Australian heavy-rain days and associated East Coast cyclones: 1958-92. *J. Climate*, **10**, 621-635.

Indeje M., F.H.M. Semazzi, and L.J. Ogallo, 2000: ENSO signal in East African rainy seasons. *Int. J. Climatol.*, **20**, 19-46.

Ingram K.T., M.C. Roncoli, and P.H. Kirshen, 2002: Opportunities and constraints for farmers of West Africa to use seasonal precipitation forecasts with Burkina-Faso as a case study. *Agricultural Systems*, **74**, 331-349.

Jeffrey S.J., J.O Carter, K.B. Moodie, and A.R. Beswick, 2001: Using spatial interpolation to construct a comprehensive archive of Australian climate data. *Environmental Modeling and Software*, **16**, 309-330.

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

Krishnamurthy V., and J. Shukla, 2000: Intraseasonal and interannual variability of rainfall over India. *J. Climate*, **13**, 4366-4377.

Kousky V.E., 1979: Frontal influences on Northeast Brazil, *Mon. Wea. Rev.*, **107**, 1140-1153.

Kousky V.E., 1985: Atmospheric circulation changes associated with rainfall anomalies over Tropical Brazil. *Mon. Wea. Rev.*, **113**, 1951-1957.

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

Leary C.A., and R.A. Houze Jr, 1979: The structure and evolution of convection in a tropical cloud cluster. *J. Atmos. Sci.*, **36**, 437-457.

Lebel T., A. Diedhiou, and H. Laurent, 2003: Seasonal cycle and interannual variability of the Sahelian rainfall at hydrological scales. *J. Geophys. Res.*, **108**, doi:10.1029/2001JD001580.

Le Barbe L., and T. Lebel, 1997: Rainfall climatology of the HAPEX-Sahel region during the years 1950-1990. *J. Hydrology*, **188-189**, 43-73.

Lopez R.E., 1978: Internal structure and development processes of C-scale aggregates of cumulus clouds. *Mon. Wea. Rev.*, **106**, 1488-1494.

Lough J.M., 1993: Variations of Some Seasonal Rainfall Characteristics in Queensland, Australia: 1921-1987. *Int. J. Climatol.*, **13**, 391-409.

Lough J.M., 1996: Regional indices of climate variation: temperature and rainfall in Queensland, Australia. *Int. J. Climatol.*, **16**, 55-66.

Mathon V., and H. Laurent, 2001: Life cycle of sahelian mesoscale convective cloud systems. *Quart. J. Meteo. Soc.*, **127**, 377-406.

Mitra A.K., A.K. Bohra, and D. Rajan, 1997: Daily rainfall analysis for Indian summer monsoon region, *Int. J. Climatol.*, **17**, 1083-1092.

Mooley D.A., 1973: Some aspects of Indian monsoon depressions and the associated rainfall. *Mon. Wea. Rev.*, **101**, 271-280.

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

Moron V., S. Bigot, and P. Roucou, 1995: Rainfall variability in subequatorial Africa and South America (1951-1990) and relationships with the main SST modes. *Int. J. Climatol.*, **15**, 1297-1322.

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

Mutai C.C., and M.N. Ward, 2000: East African rainfall and the tropical circulation / convection on intraseasonal to interannual timescales. *J. of Climate*, **13**, 3915-3939

New M., M. Hulme, and P. Jones, 2000: Representing twentieth-century space-time climate variability. Part II: development of 1901-96 monthly grids of terrestrial surface climate. *J. Climate*, **13**, 2217-2238.

Nicholson S.E., 1996: A review of climate dynamics and climate variability in Eastern Africa. In "The limnology, climatology and paleoclimatology of the east african lakes", Johnson T.C. and Odada E.O., Eds. Gordon and Breach, 25-56.

Ogallo L.J., 1985: Climatology of rainfall in East Africa. WMO Conference on GATE, WAMEX, Trop. met., 10-14 Dec. 1985, Dakar. WMO T.M.P. n°16, 96-102.

Ogallo L.J., 1989: The spatial and temporal patterns of the East African seasonal rainfall derived from principal component analysis. *J. Climatol.*, **9**, 145-167.

Parthasarhy B., A.R. Kumar, and A.A. Munot, 1993, Homogeneous Indian monsoon rainfall – variability and prediction. *Proc. Indian. Acad. Sci. (Earth and Planetary Sciences)*, **102**, 121-155.

Parthasarhy B., A.A. Munot, and D.R. Kothawale 1994: All India monthly and seasonal time series 1871-1993. *Theor. Appl. Clim.*, **49**, 217-224.

Philippon N., P. Camberlin and N. Fauchereau, 2003: Empirical predictability study of October-December East African rainfall. *Quart. J. Meteo. Soc.*, **128**, 2239-2256.

Ramos R.P.L., 1975: Precipitation characteristics in the Northeast Brazil dry region, *J. Geophys. Res.*, **80**, 1665-1678.

Ratisbona C.R., 1976: The climate of Brazil. *Climates of Central and South America*, W. Schwerdtfeger and H.E. Landsberg Eds, *World Survey of Climatology*, vol. 12, Elsevier, 219-293.

Ricciardulli L., and P.D. Sardeshmukh, 2002: Local time- and space scales of organized tropical deep convection. *J. Climate*, **15**, 2775-2790.

Rickenbach T.M., and S.A. Rutledge, 1998: Convection in TOGA COARE: horizontal scale, morphology and rainfall production. *J. Atmos. Sci.*, **55**, 2715-2729.

Robertson A.W., S. Kirshner, P. Smyth, B. Bates, and S. Charles, 2006: Subseasonal-to-interdecadal variability of the Australian monsoon over North Queensland. *Quarterly J. Meteorol. Soc.*, **132**, 519-542.

Rowell D.P., Ininda J.M., Ward M.N., 1994: The impact of global sea surface temperature patterns on seasonal rainfall in East Africa. *Proc. Int. Conf. Monsoon Variability and Prediction*, Trieste, Italie, 9-13 Mai 1994. WMO/TD n°619. 666-672.

Sikka D.R., and S. Gadgil, 1980: On the maximum cloud zone and the ITCZ over India longitude during the SouthWest monsoon, *Mon. Wea. Rev.*, **108**, 1840-1853.

Smith D.F., A.J. Gasiewski, D.L. Jackson, and G.A. Wick, 2005: Spatial scales of tropical precipitation inferred from TRMM Microwave imager data, *IEEE Trans. Geos. Rem. Sensing*, **43**, 1542-1551.

Snijders T.A.B., 1986 : Interstation correlations and nonstationnarity of Burkina Faso rainfall. *J. Clim App. Meteo.*, **25**, 524-531.

Stern R.D., and R. Coe, 1984: A model fitting analysis of daily rainfall data. *J. R. Statistic. Soc. A*, **147**, 1-34.

Webster P.J., V.O. Magana, T.N. Palmer, J. Shukla, R.T. Tomas, M. Yanai, and T. Yasunari, 1998: Monsoons, predictability and the prospects of prediction. *J. Geophys. Res.*, **103**, 14451-14510.

Wilks D.S., 1999: Interannual variability and extreme-value characteristics of several daily precipitation models. *Agricultural and Forest Meteorology*, **93**, 153-169.

Figures captions

Figure 1: Scheme of convective cells (shaded small circles) embedded in a meso-Scale Convective Complex (MCC) (large white circle), across a network of stations (denoted as crosses). The location of the MCC is indicated at t_0 (simple line), $t+6$ hours (double line) and $t+12$ hours (triple line). The speed and direction (here westward) of moving of MCC, as well as the shape and the relative location of convective cells inside the MCC are varying across the tropics.

Figure 2: Geographical location of the stations in (a) Senegal, (b) Queensland, (c) Northwestern India, (d) Nordeste and (e) Kenya. The stations are displayed as dots on a regional window of the same size (i.e. 10 deg x 9 deg) approximately centered on each network to ease the comparison. In Kenya, two stations are located at Malindi on the coast of Indian Ocean.

Figure 3: Degree of freedom –DOF– for (a) seasonal amount, (b) occurrence of rainfall (= number of wet days > 1 mm) and (c) mean intensity of rain during rainy days (= seasonal amount / number of wet days) of the 9-stations Kenyan network. The full line with white circle (black circle) indicates *DOF* for March-May (October-December) and the dashed lines are the maximum and minimum values of the combinations considering M (between 2 and 8) stations amongst 9. The dotted lines indicate the lowest *DOF* reached in 10, 5 and 1% of a 39 x 9 white noise time series.

Figure 4: Scatter-plot of degree of freedom –DOF– (in ordinate) and interannual variance of the standardized anomaly index $-var[SAI]$ – (in abscissa) for the 6 networks ('K1'; Kenya in March-May, 'K2'; Kenya in October-December, 'NO'; Nordeste in February-April; 'SE'; Senegal in July-September; 'IN'; Northwestern India in June-September; 'QU'; Queensland in December-March) and the 3 seasonal variables (black circle: seasonal amount; white circle: occurrence of rainfall (wet day = day receiving measurable rain > 1 mm); cross; mean intensity of rainfall during wet days). Ellipses have been added to emphasize the range of both estimates of spatial

coherence for seasonal amount (full line), frequency of occurrence (dashed line) and mean intensity of rainfall (dotted line).

Figure 5: (a) Degree of freedom $-DOF-$ (ordinate) of the interannual variability of frequency of occurrence with different thresholds to define a day as wet (in abscissa), i.e. frequency of occurrence is computed for each season and station using a threshold of $> 0, 1, 2, \dots, 10$ mm to define a day as wet; (b) DOF (ordinate) of the interannual variability of seasonal amounts with different levels of truncation of daily amounts (in abscissa) i.e. seasonal amounts are computed using raw daily amounts (Inf), and truncated daily amounts to respectively 100, 50, 20, 10, 5, 1 mm when they are above this amount.

Figure 6: Spatial correlation (in ordinate) function of daily rainfall amounts vs distance in km (in abscissa) for (a) Senegal in July-September, (b) Queensland in December-March, (c) Northwestern India in June-September, (d) Nordeste in February-April, (e) Kenya in March-May and (f) Kenya in October-December. The correlations are computed for each pair, represented as dot, using only wet day (> 1 mm) at both stations. The bold circles give the averaged correlations for mean distance in range 0-99, 100-199 km, etc. The last two panels shows the average spatial correlation functions for the daily rainfall amounts (g) and the square root of daily rainfall amounts (h) for the 6 networks. The $1/e$ correlation (~ 0.37) is shown for convenience as a dashed line.

Figure 7: Spatial correlation (in ordinate) function of daily frequency of occurrence $-\phi$ correlation $-\phi$ of rainfall vs distance in km (in abscissa) for (a) Senegal in July-September, (b) Queensland in December-March, (c) Northwestern India in June-September, (d) Nordeste in February-April, (e) Kenya in March-May and (f) Kenya in October-December. ϕ correlation is computed for each pair, represented as dot, using frequency of occurrence of wet day > 1 mm. The bold circle gives the averaged ϕ correlation for mean distance in range 0-99, 100-199 km,

etc.. The panel (g) shows the spatial correlation functions for the daily rainfall occurrence for the 6 networks. The $1/e$ correlation (~ 0.37) is shown for convenience as a dashed line.

Figure 8: Mean spatial correlation (in ordinate) function versus the mean distance in km (in abscissa) for the frequency of occurrence (first column; (a) Senegal in July-September, (d) Queensland in December-March, (g) Northwestern India in June-September (j) Nordeste in February-April (m) Kenya in March-May and (p) Kenya in October-December), mean intensity of rainfall (second column; (b) Senegal in July-September, (e) Queensland in December-March, (h) Northwestern India in June-September (k) Nordeste in February-April (n) Kenya in March-May and (q) Kenya in October-December) and amount (third column; (c) Senegal in July-September, (f) Queensland in December-March, (i) Northwestern India in June-September (l) Nordeste in February-April (o) Kenya in March-May and (r) Kenya in October-December) for the daily (full bold line), 2-day (line + lower triangle), 5-day (line + upper triangle), 10-day (line + circle), 30-day (line + square) and seasonal time scales. The spatial correlation functions are displayed on the same scale to ease the comparison.

Figure 9: Spatial correlation (in ordinate) function of seasonal amount (a) and frequency of occurrence (b) vs distance (in abscissa). The mean spatial correlation is averaged over each network and shown for mean distance in range 0-99, 100-199 km, etc. The $1/e$ value (~ 0.37) is shown for convenience as a dashed line.

Figure 10: Observed (large gray dots) and simulated (small black dots) spatial correlation (in ordinate) function vs distance in km (in abscissa) for seasonal frequency occurrence of rainfall for Queensland (a) and Northwestern India (b). The simulations come from 100 random fields having the same external variance ratio (= amount of signal) as the observed fields and where the amplitude of signal (respectively noise) is spatially modulated as the 1 minus the mean distance (respectively mean distance) of each station with the remaining ones (see text). The mean spatial correlation for 0-99, 100-199 km is shown for observation (circle + full line) and simulation (square + dashed line).

Table captions

Table 1 : Summary of the 6 seasonal networks analyzed in this study.

Table 2 : Squared correlations between seasonal-average quantities (*S*: seasonal amount; *O*: frequency of wet day and *I*: mean intensity of rainfall, all computed using a threshold of 1 mm for defining wet days), averaged over each network (columns 2-5), together with network-averaged estimates of the number of degrees of freedom (*DOF*) of each station quantity (columns 6-9).

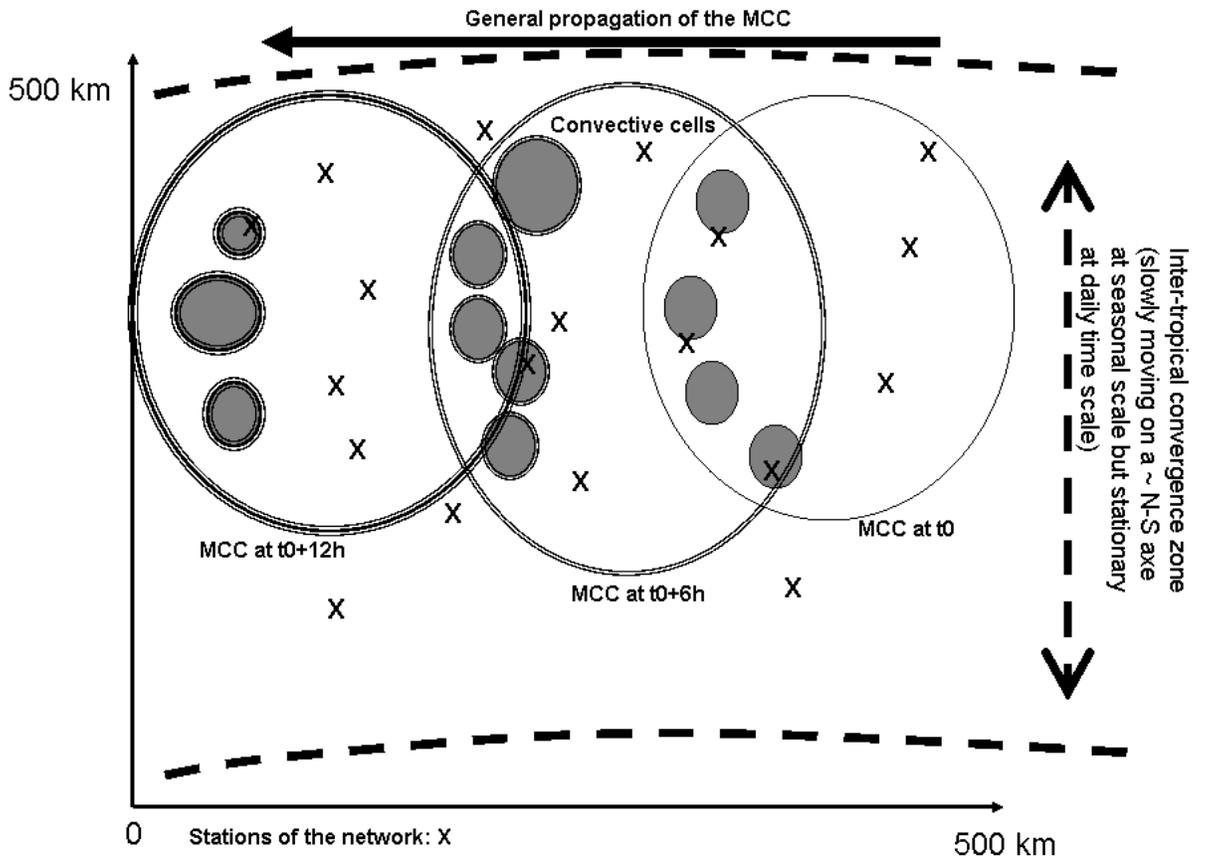


Figure 1: Scheme of convective cells (shaded small circles) embedded in a Meso-Scale Convective Complex (MCC) (large white circle), across a network of stations (denoted as crosses). The location of the MCC is indicated at t_0 (simple line), $t+6$ hours (double line) and $t+12$ hours (triple line). The speed and direction (here westward) of moving of MCC, as well as the shape and the relative location of convective cells inside the MCC are varying across the tropics.

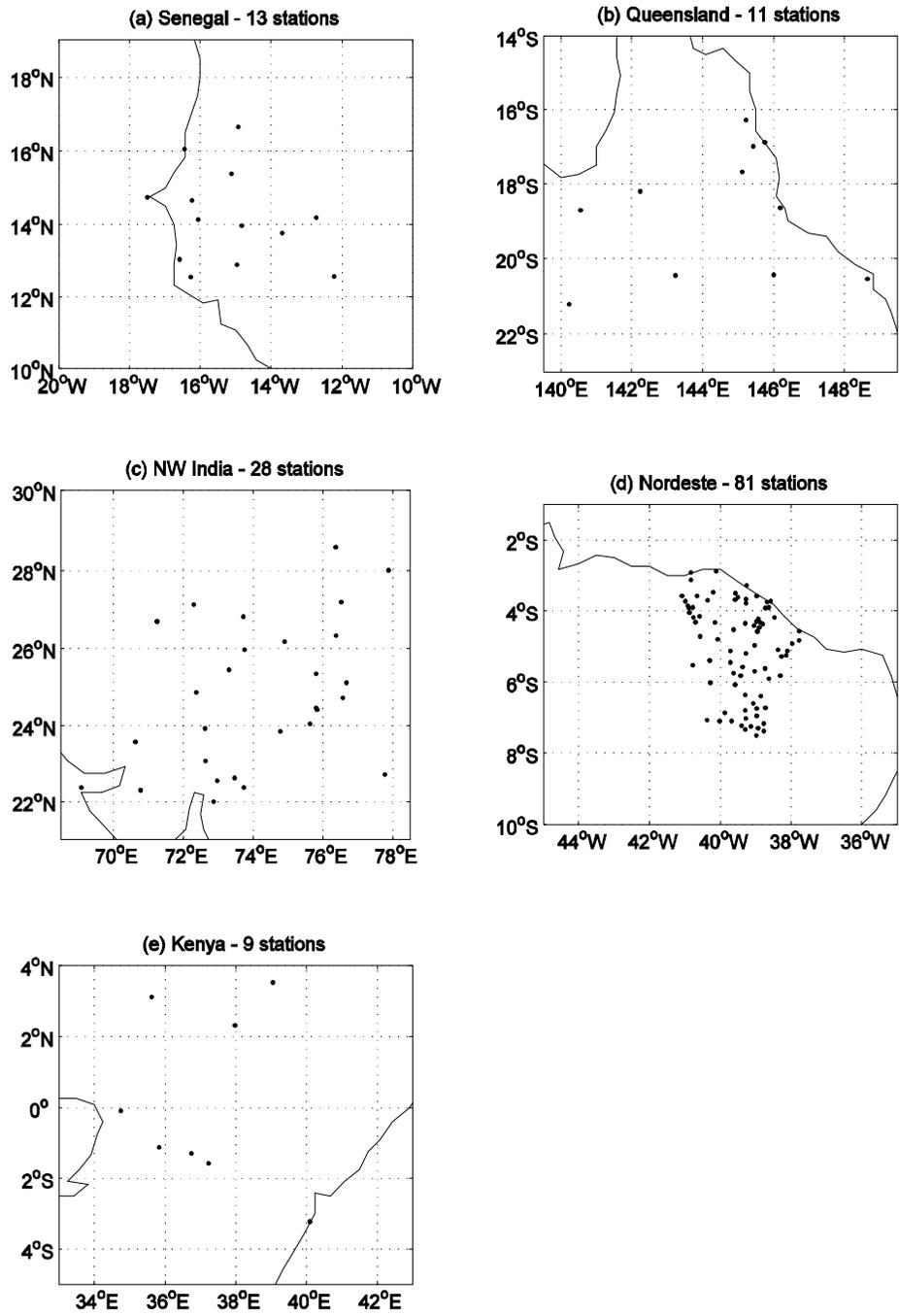


Figure 2: Geographical location of the stations in (a) Senegal, (b) Queensland, (c) Northwestern India, (d) Nordeste and (e) Kenya. The stations are displayed as dots on a regional window of the same size (i.e. 10 deg x 9 deg) approximately centered on each network to ease the comparison. In Kenya, two stations are located at Malindi on the coast of Indian Ocean.

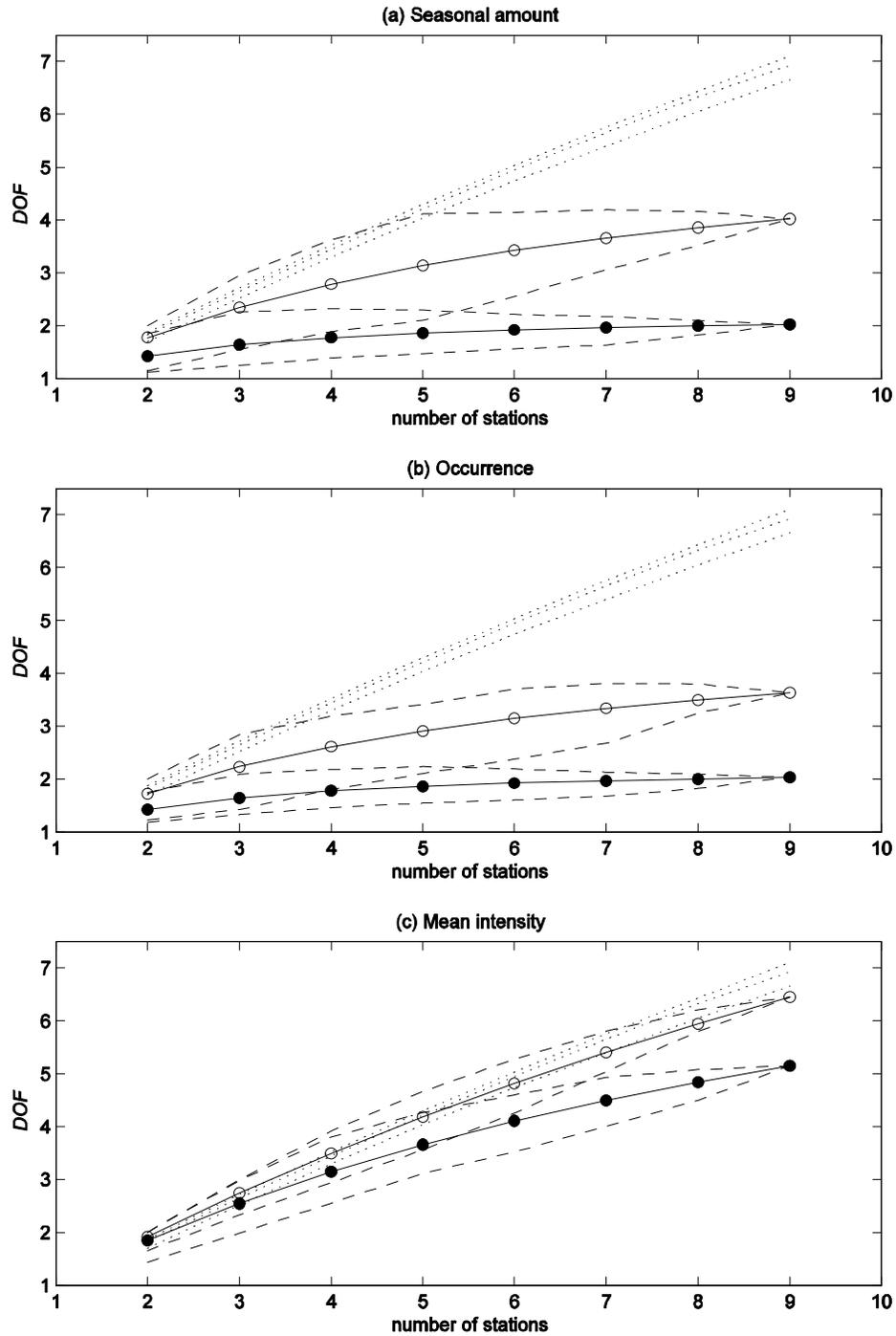


Figure 3: Degree of freedom –DOF– for (a) seasonal amount, (b) occurrence of rainfall (= number of wet days > 1 mm) and (c) mean intensity of rain during rainy days (= seasonal amount / number of wet days) of the 9-stations Kenyan network. The full line with white circle (black circle) indicates *DOF* for March-May (October-December) and the dashed lines are the maximum and minimum values of the combinations considering M (between 2 and 8) stations amongst 9. The dotted lines indicate the lowest *DOF* reached in 10, 5 and 1% of a 39 x 9 white noise time series.

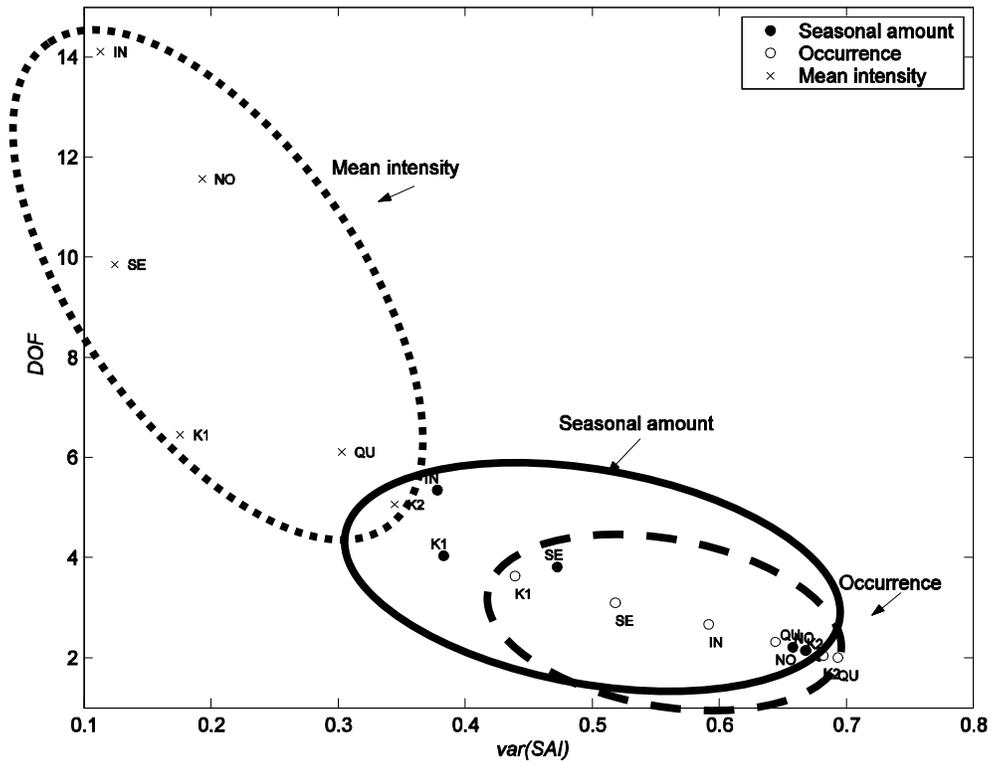


Figure 4: Scatter-plot of degree of freedom $-DOF-$ (in ordinate) and interannual variance of the standardized anomaly index $-var[SAI]-$ (in abscissa) for the 6 networks ('K1'; Kenya in March-May, 'K2'; Kenya in October-December, 'NO'; Nordeste in February-April; 'SE'; Senegal in July-September; 'IN'; Northwestern India in June-September; 'QU'; Queensland in December-March) and the 3 seasonal variables (black circle: seasonal amount; occurrence of rainfall (wet day = day receiving measurable rain $>$ 1 mm); cross; mean intensity of rainfall during wet days). Ellipses have been added to emphasize the range of both estimates of spatial coherence for seasonal amount (full line), frequency of occurrence (dashed line) and mean intensity of rainfall (dotted line).

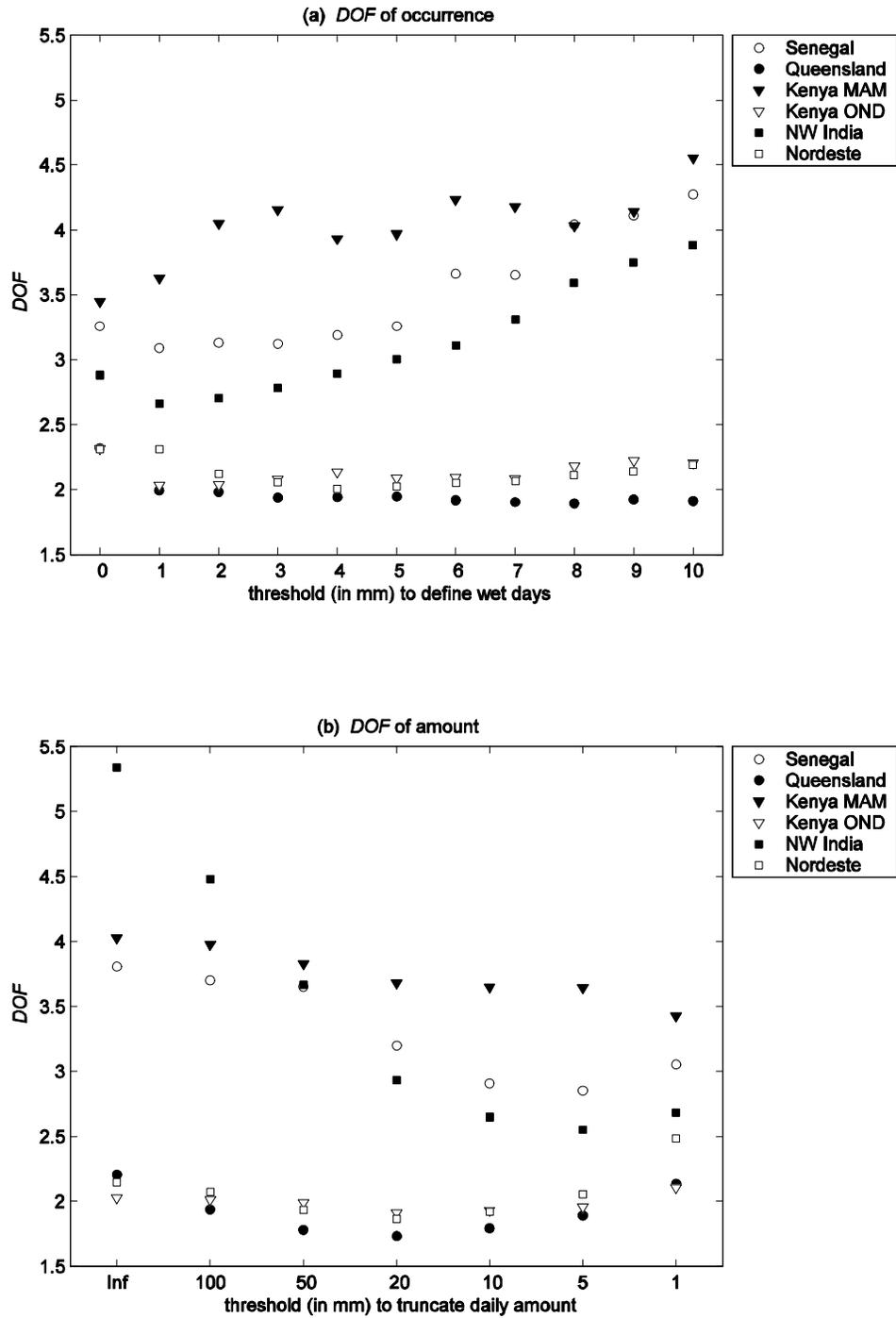


Figure 5: (a) Degree of freedom *-DOF-* (ordinate) of the interannual variability of frequency of occurrence with different thresholds to define a day as wet (in abscissa) i.e. frequency of occurrence is computed for each season and station using a threshold of $> 0, 1, 2, \dots, 10$ mm to define a day as wet; (b) *DOF* (ordinate) of the interannual variability of seasonal amounts with different levels of truncation of daily amounts (in abscissa) i.e. seasonal amounts are computed using raw daily amounts (Inf), and truncated daily amounts to respectively 100, 50, 20, 10, 5, 1 mm when they are above this amount.

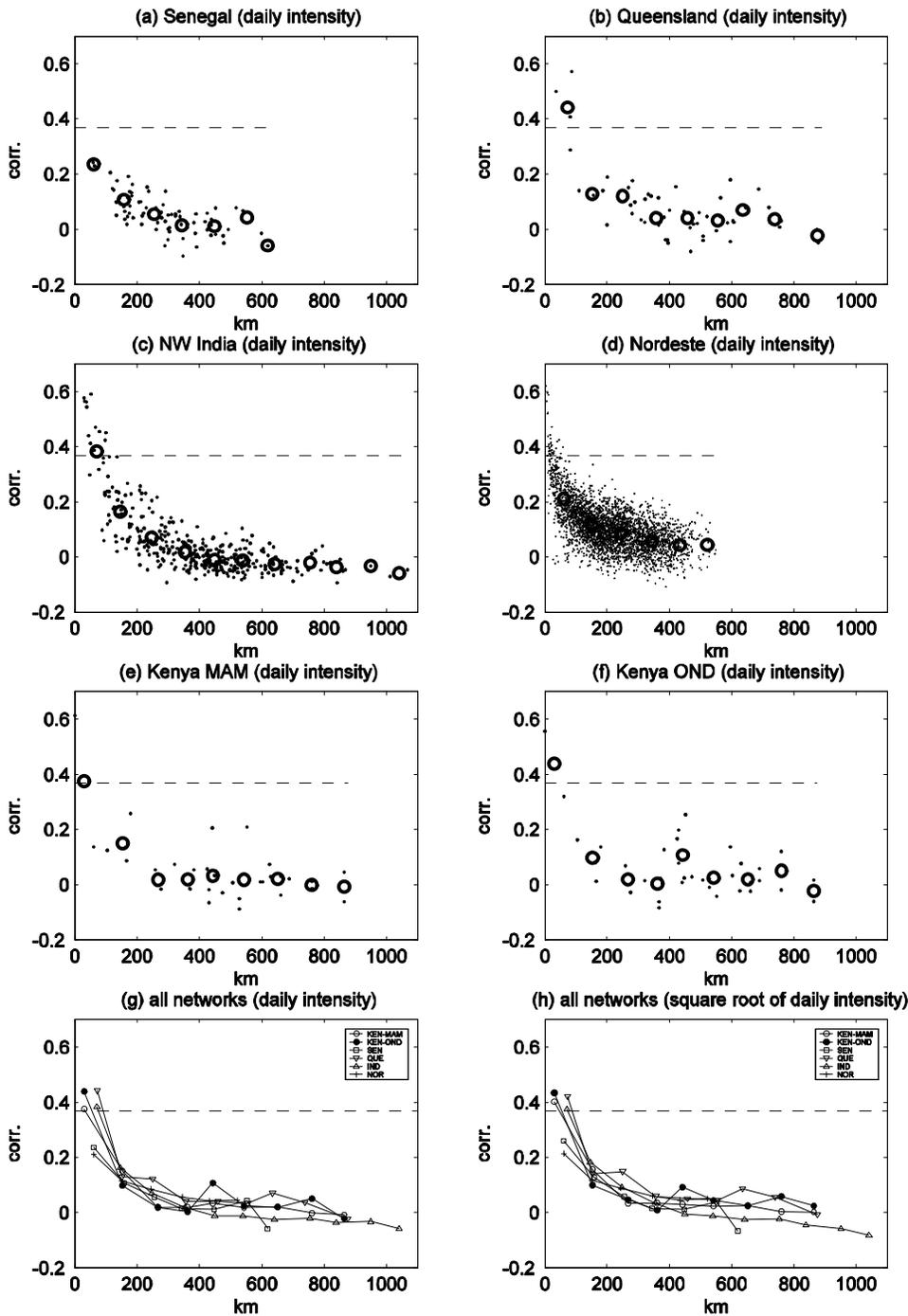


Figure 6: Spatial correlation (in ordinate) function of daily rainfall amounts vs distance in km (in abscissa) for (a) Senegal in July-September, (b) Queensland in December-March, (c) Northwestern India in June-September, (d) Nordeste in February-April, (e) Kenya in March-May and (f) Kenya in October-December. The correlations are computed for each pair, represented as dot, using only wet day (> 1 mm) at both stations. The bold circles give the averaged correlations for mean distance in range 0-99, 100-199 km, etc. The last two panels shows the average spatial correlation functions for the daily rainfall amounts (g) and the square root of daily rainfall amounts (h) for the 6 networks. The $1/e$ correlation (~ 0.37) is shown for convenience as a dashed line.

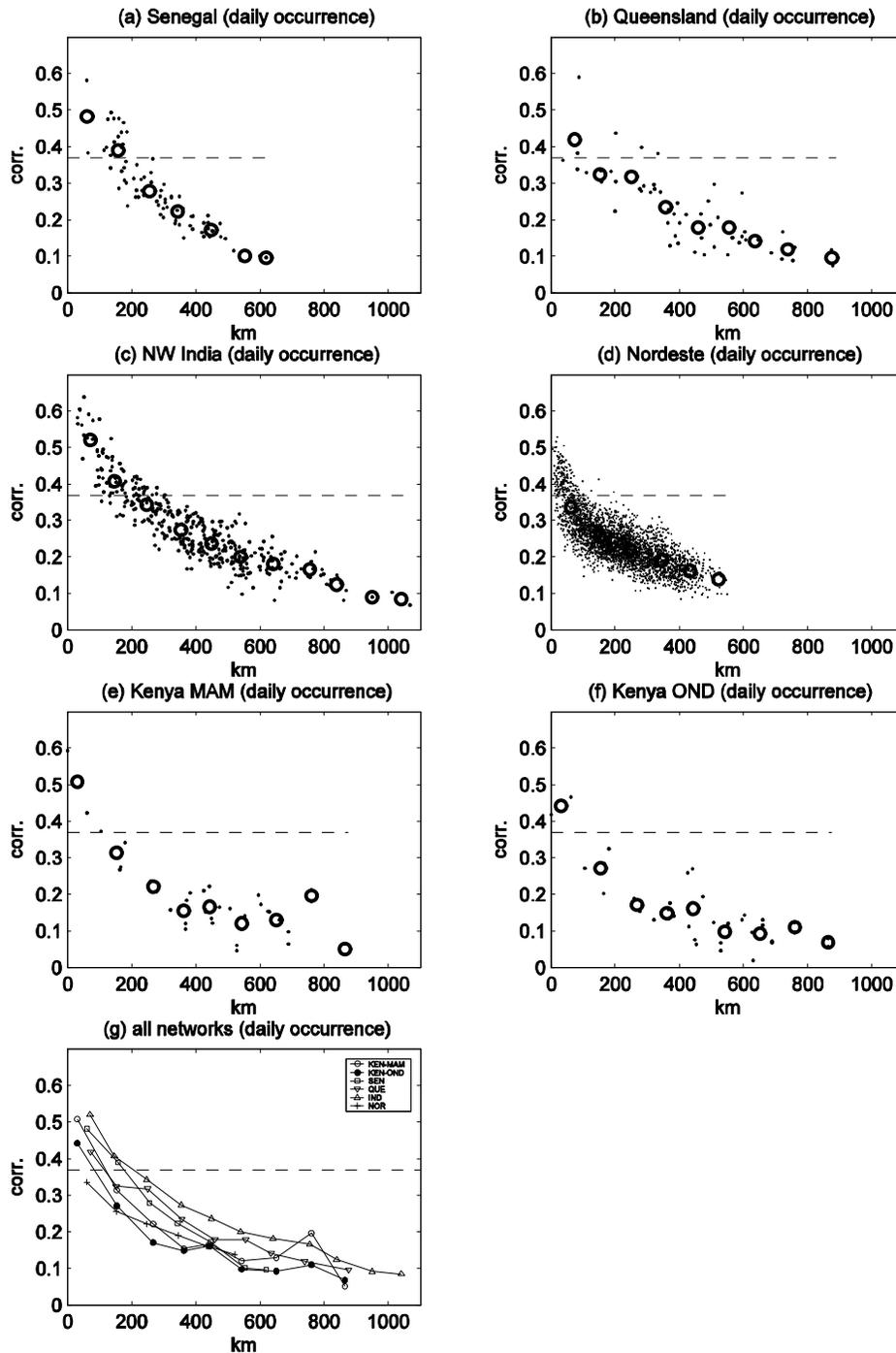


Figure 7: Spatial correlation (in ordinate) function of daily frequency of occurrence – ϕ correlation – of rainfall vs distance in km (in abscissa) for (a) Senegal in July-September, (b) Queensland in December-March, (c) Northwestern India in June-September, (d) Nordeste in February-April, (e) Kenya in March-May and (f) Kenya in October-December. ϕ correlation is computed for each pair, represented as dot, using frequency of occurrence of wet day > 1 mm. The bold circle gives the averaged ϕ correlation for mean distance in range 0-99, 100-199 km, etc.. The panel (g) shows the spatial correlation functions for the daily rainfall occurrence for the 6 networks. The $1/e$ correlation (~ 0.37) is shown for convenience as a dashed line.

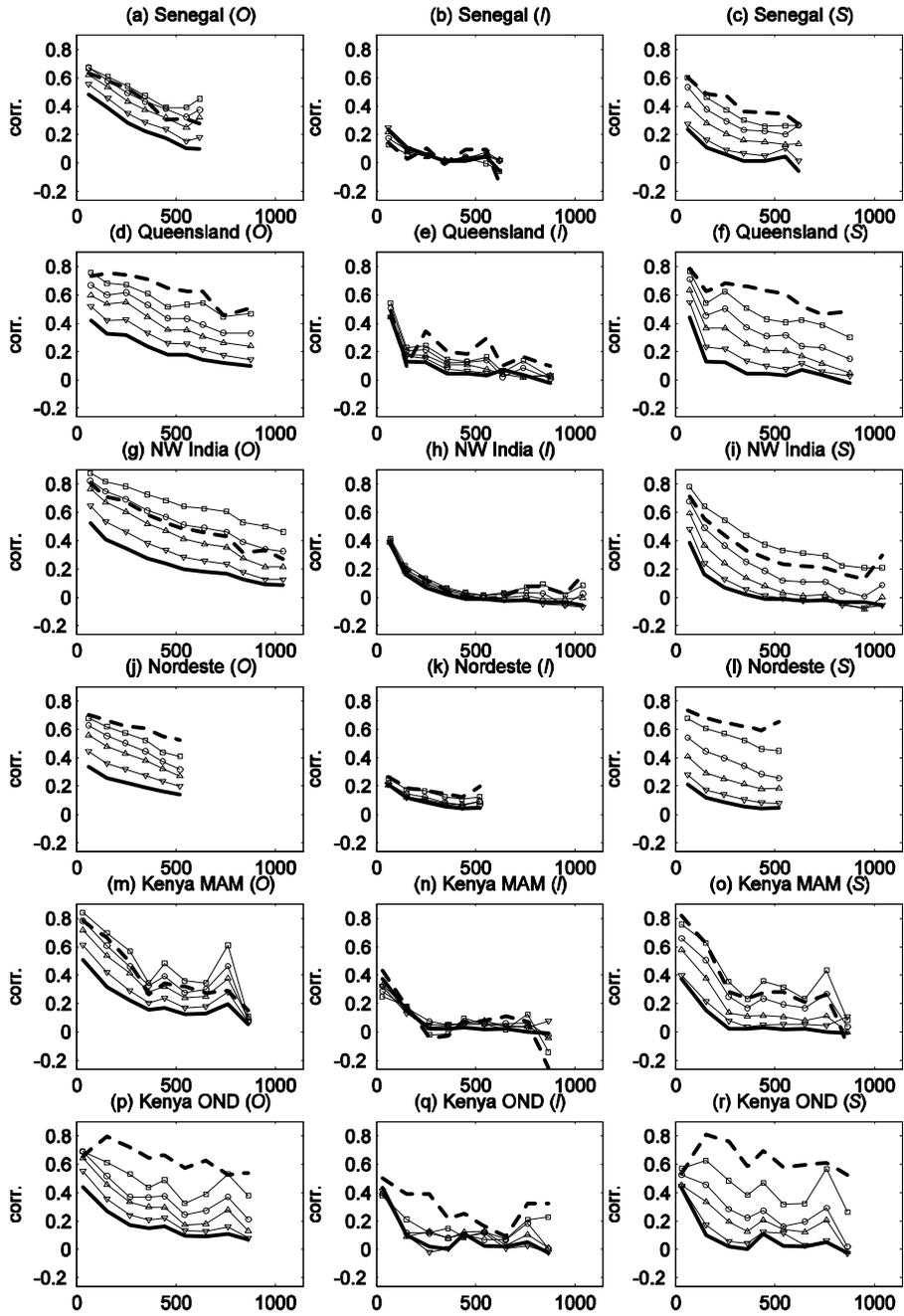


Figure 8: Mean spatial correlation (in ordinate) function versus the mean distance in km (in abscissa) for the frequency of occurrence (first column; (a) Senegal in July-September, (d) Queensland in December-March, (g) Northwestern India in June-September (j) Nordeste in February-April (m) Kenya in March-May and (p) Kenya in October-December), mean intensity of rainfall (second column; (b) Senegal in July-September, (e) Queensland in December-March, (h) Northwestern India in June-September (k) Nordeste in February-April (n) Kenya in March-May and (q) Kenya in October-December) and amount (third column; (c) Senegal in July-September, (f) Queensland in December-March, (i) Northwestern India in June-September (l) Nordeste in February-April (o) Kenya in March-May and (r) Kenya in October-December) for the daily (full bold line), 2-day (line + lower triangle), 5-day (line + upper triangle), 10-day (line + circle), 30-day (line + square) and seasonal time scales. The spatial correlation functions are displayed on the same scale to ease the comparison.

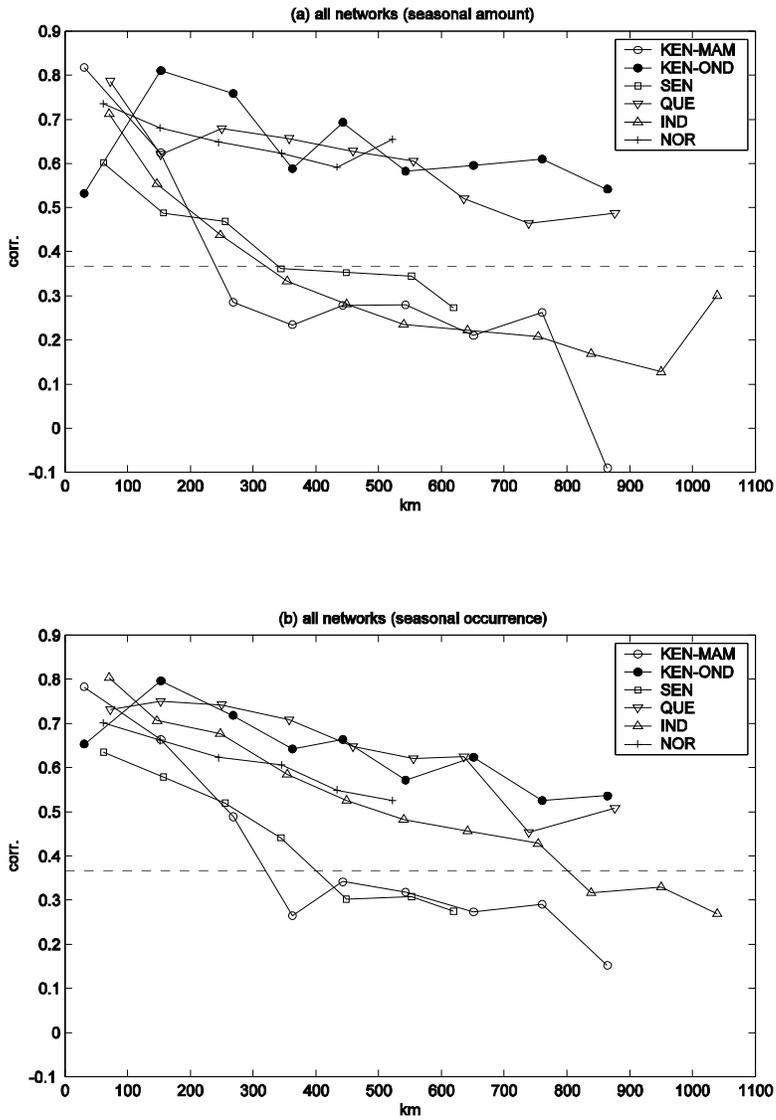


Figure 9: Spatial correlation (in ordinate) function of seasonal amount (a) and frequency of occurrence (b) vs distance (in abscissa). The mean spatial correlation is averaged over each network and shown for mean distance in range 0-99, 100-199 km, etc. The 1/e value (~ 0.37) is shown for convenience as a dashed line.

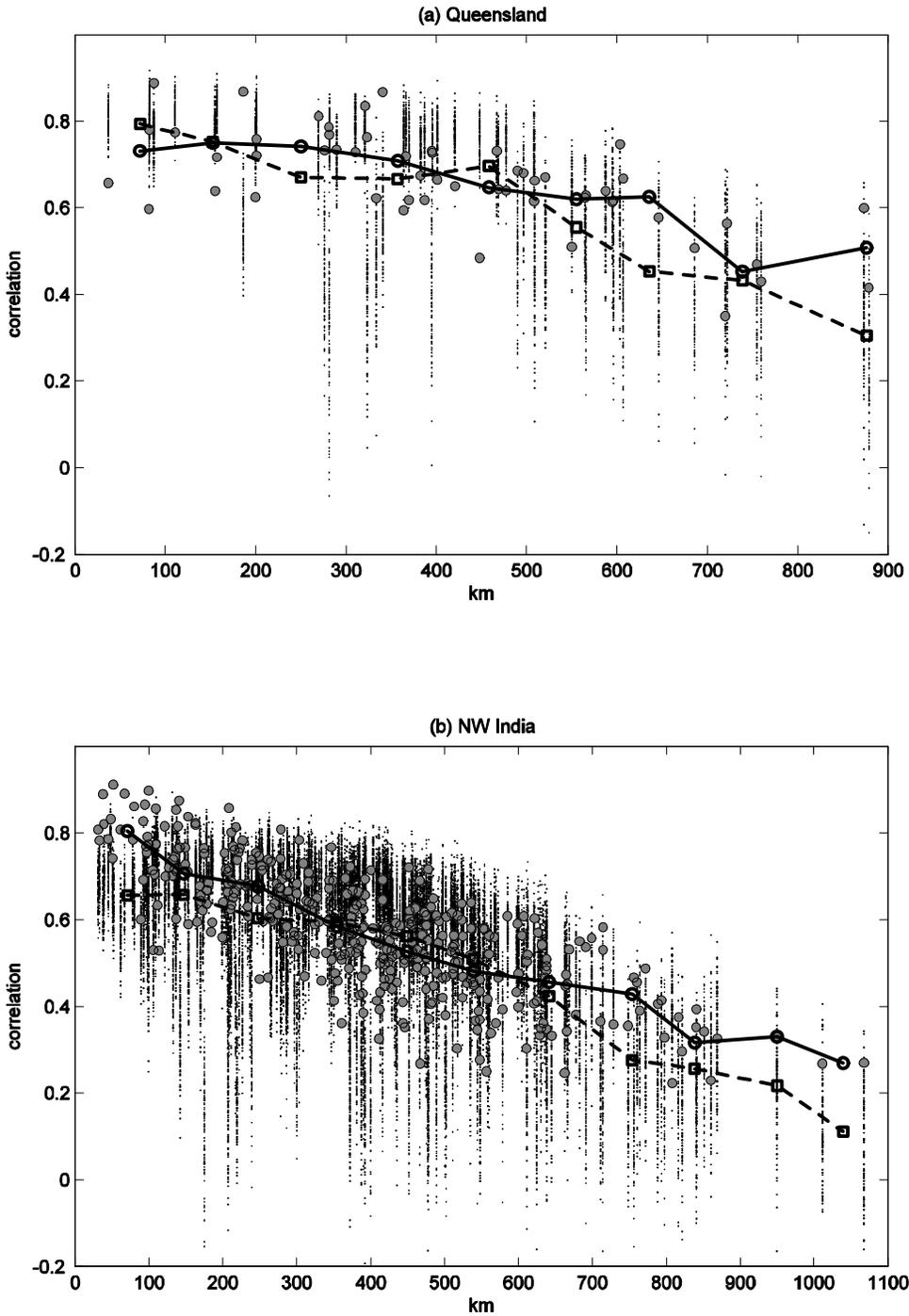


Figure 10: Observed (large gray dots) and simulated (small black dots) spatial correlation (in ordinate) function vs distance in km (in abscissa) for seasonal frequency occurrence of rainfall for Queensland (a) and Northwestern India (b). The simulations come from 100 random fields having the same external variance ratio (= amount of signal) as the observed fields and where the amplitude of signal (respectively noise) is spatially modulated as the 1 minus mean distance (respectively mean distance) of each station with the remaining ones (see text). The mean spatial correlation for 0-99, 100-199 km is shown for observation (circle + full line) and simulation (square + dashed line).

Table 1: Summary of the 6 seasonal networks analyzed in this study.

	Number of Stations	Months	Years	Density (stations / km ²)
Senegal	13	JAS	1961-98	1/13000
Queensland	11	DJFM	1959-98	1/26000
NW India	28	JJAS	1901-70	1/15000
Nordeste	81	FMA	1974-02	1/1600
Kenya	9	MAM	1960-98	1/29000
Kenya	9	OND	1960-98	1/29000

Table 2 : Squared correlations between seasonal-average quantities (*S*: seasonal amount; *O*: frequency of wet day and *I*: mean intensity of rainfall, all computed using a threshold of 1 mm for defining wet days), averaged over each network (columns 2-5), together with network-averaged estimates of the number of degrees of freedom (*DOF*) of each station quantity (columns 6-9).

	<i>S</i> vs <i>O</i>	<i>S</i> vs <i>I</i>	<i>O</i> vs <i>I</i>	<i>S_{occ}</i> - <i>S</i> vs <i>I</i>	<i>DOF O</i>	<i>DOF S</i>	<i>DOF</i> (<i>S</i> - <i>S_{occ}</i>)	<i>DOF I</i>
Senegal	.49	.41	.05	.90	3.1	3.8	9.2	9.9
Queensland	.61	.51	.04	.88	2.0	2.2	6.5	6.1
NW India	.65	.35	.05	.86	2.7	5.3	14.7	14.1
Nordeste	.55	.55	.07	.85	2.3	2.1	11.4	11.6
Kenya MAM	.76	.58	.20	.62	3.6	4.0	6.8	6.4
Kenya OND	.49	.49	.02	.88	2.0	2.0	6.0	5.2