

Forecasting the vegetation photosynthetic activity over the Sahel: a Model Output Statistics approach

Philippon N.^{°*}

Martiny N[§]. and Camberlin P[°].

[°]Centre de Recherches de Climatologie

UMR5210 CNRS / Université de Bourgogne

6 blvd Gabriel, BP21877, 21078 Dijon CEDEX, FRANCE

UMR5805 EPOC-OASU

Site de Talence, Université Bordeaux 1

Avenue des Facultés

33405 Talence CEDEX - FRANCE

* corresponding author: nphilipp@u-bourgogne.fr

submitted to International Journal Of Climatology

4 April 2007

Revised 26 September 2008

Abstract

The predictability of the mean August-September photosynthetic activity of vegetation over the Sahel for the period 1982-2002 is explored through a Model Output Statistics approach using ECHAM4.5 retrospective forecasts. Given the poor ability of Atmospheric General Circulation Models (AGCMs) to correctly simulate rainfall over the Sahel, the stress is put on using atmospheric dynamics alone. The mean July-September predicted fields of zonal wind at 600hPa and humidity flux at 850hPa are selected because of their key-role in the West African Monsoon system and their consistency in AGCMs. Coupled modes of NDVI / atmospheric dynamics are extracted using Canonical Correlation Analyses performed in leave-one-out cross-validation. The most relevant modes (using NCEP/DOE 2 or ECHAM4.5 atmospheric dynamics) associate enhanced greenness to a weakened African Easterly Jet displaced northward and strengthened moisture fluxes from the Guinean Gulf. They are linked with increased rainfall over the Sahel and positive (negative) SST anomalies over the Mediterranean (the Eastern equatorial Pacific).

Used as predictors in a Multiple Linear Regression model, the 3rd cross-validated canonical coefficient derived from ECHAM4.5 simulations added to the August-September NDVI value of the previous year enable to explain 30% of the variance of a Sahelian regional index with a 2-month lead-time. Applied at an 8-km spatial resolution, the statistical model possesses a usable skill (>.5) for 24% of the pixels analysed. The NDVI of pixels covered by open grassland appears as the most predictable.

Keywords:

Normalised Difference Vegetation Index – Sahel – Model Output Statistics – hindcasts

1. Introduction

The monitoring and understanding of the interannual evolution of the vegetation over the Sahel by use of the Normalised Difference Vegetation Index (NDVI) satellite product have been the subject of numerous studies in the last two decades.

The earlier studies by Justice and Hiernaux (1986), Tucker *et al.* (1986), Dregne and Tucker (1988) were amongst the first to attempt the monitoring of vegetation over the Sahel. They pointed out the ability of the NDVI product to detect spatial variations in the primary production from one year to another for the very dry years 1983-1985. Tucker *et al.* (1991) used NDVI to estimate annual precipitation and analysed the latitudinal shifts of the Saharan-Sahelian boundary. They observed that the boundary mean position in 1990 was ~130km south of its position in 1980.

The extension of the NDVI data set temporal coverage (more than 20-year records by now) have enabled to further investigate the factors and characteristics of the NDVI interannual variability over the Sahel. The prime driver of NDVI variability is rainfall. A strong linear and positive relationship is observed between annual (and seasonal) NDVI and rainfall amounts (Malo and Nicholson, 1990, Nicholson *et al.*, 1990, Herrmann *et al.*, 2005, Camberlin *et al.*, 2007). However, for a given rainfall amount, the strength of the relationship significantly varies with the vegetation type : open grassland and cropland areas exhibit the highest NDVI/rainfall correlations (Camberlin *et al.* 2007). Memory effects from one year to another have also been recently evoked by Martiny *et al.* (2005) and Philippon *et al.* (2007) to explain the persistence of marked NDVI anomalies after very wet or very dry years (e.g. 1984).

In view of the characteristics of the NDVI variability, two main points recently emerged. At first, the NDVI variability is spatially coherent over the region. Yearly maps of NDVI anomalies by Anyamba and Tucker (2005) clearly highlight Sahel-wide patterns. Jarlan *et al.* (2005) and Philippon *et al.* (2007) extracted regional-scale modes of NDVI variability which partly match those of rainfall (such as the global and West Sahelian modes, Moron 1994) by applying Principal or Independent Component Analyses to the NDVI data set. Secondly, seasonal NDVI presents a positive trend since 1982 over most of the Sahel (Eklundh and Olsson, 2003, Herrmann *et al.*, 2005, Anyamba and Tucker, 2005, Heumann *et al.*, 2007). This greenness increase has proved to be related at first with

rainfall recovery from the early 80's drought conditions (Hickler *et al.*, 2005) but the synchronous changes observed in land use and land management are also assumed to play a role (Olsson *et al.*, 2005).

Thus, the studies performed until now have offered to the scientific community a clearer picture of the factors and characteristics of the spatial and temporal variability of the vegetation photosynthetic activity over the Sahel. The next step must be toward the NDVI forecasting. Several early warning systems (for famine, epidemics) use satellite remote sensing data to supplement ground-based observations (Gulaid, 1986). Indeed, remote sensing data have an advantage in timeliness and for the identification of hot-spots areas where to take action (Hutchinson, 1991).

Numerous studies have demonstrated that Net Primary Production and crop yield could be estimated from NDVI (Tucker *et al.*, 1986, Groten, 1993, Maselli *et al.*, 1993, 2000). Forecasts of NDVI such as the ones proposed in the current study have been demonstrated to be a valuable tool for forecasting crop yields (Rasmussen, 1997) and grain potential prices movement (Brown *et al.* 2008), planning resources for livestock for the dry season (Wylie *et al.*, 1991). Used as an indicator for wetness/dryness conditions in semi-arid Africa, NDVI forecasts can also potentially inform on epidemic and epizootic risks (Linthicum *et al.*, 1999, Indeje *et al.*, 2006).

The predictability of NDVI has recently been evaluated by Funk and Brown (2006), Mangiarotti *et al.* (2006), Indeje *et al.* (2006), and Martiny *et al.* (2008) for the whole or specific areas of semi-arid Africa. The two former studies focused on projections of the intra-seasonal NDVI changes and used in particular the NDVI itself as a predictor. The two latter studies worked on seasonal NDVI to be related with lagged observed or modelled oceano-atmospheric conditions. In addition, Jarlan *et al.* (2005) and Philippon *et al.* (2007) highlighted significant lagged relationships between the NDVI over the Sahel and Sea Surface Temperatures anomalies over the Pacific, the Indian or the Atlantic oceans. On the whole, these studies point out the high potential predictability for the NDVI over semi-arid areas of Africa. Thus, the objective of this study is to evaluate the predictability of the vegetation photosynthetic activity for the Sahelian region specifically. The approach adopted follows the one developed by Indeje *et al.* (2006) for forecasting NDVI in East Africa: we first look for the coupled modes of variability between the NDVI and selected observed or modelled atmospheric

fields. We secondly apply a Model Output Statistics (MOS) approach to hindcast the NDVI from the main modes of atmospheric variability detected. However, our study is distinct from the one of Indeje *et al.* (2006) regarding (i) the fields used as predictors: in particular, we did not use rainfall as a potential predictor of the NDVI, (ii) the way the CCAs are computed: we applied leave-one-out cross-validations which means that skill is lowered but more realistic in particular in view of operational forecasting, and (iii) the pixel resolution considered: we tested the MOS for the NDVI field at an 8-km spatial resolution, and examined predictability according to the land-cover type. The study is organised into 6 sections. Section 2 presents the NDVI and observed / modelled atmospheric data used. Section 3 gives a brief overview of the statistical tools and MOS approach employed to relate NDVI to the atmospheric fields. In section 4, the main coupled modes of NDVI and observed / modelled atmospheric fields are presented, compared, and discussed in terms of linkages with the rainfall and sea surface temperature fields. The atmospheric components of the coupled modes detected in section 4 are then used as predictors in forecasting models presented in section 5. Section 6 closes the paper.

2. Data sets:

a. Normalized Difference Vegetation Index:

The NDVI is a combination of measures in the red and infra-red spectral regions which allows to estimate the photosynthetic capacity of vegetated areas. The NDVI data set used in this study is from the new process by the Global Inventory Mapping and Monitoring Studies (GIMMS) group at the University of Maryland (available at <http://glcf.umiacs.umd.edu/data/gimms>). This data set from the Advanced Very High Resolution Radiometer instrument on-board the NOAA meteorological satellites provides NDVI global records at an 8-km spatial resolution for the period July 1981 – December 2003. In the current study, we worked on the Africa subset, sampled at a 10-day time step. This data set differs from the previous data sets in several ways: the use of an empirical mode decomposition/reconstruction to minimise effects of orbital drift, NOAA-9 descending node data from September 1994 to January 1995, a volcanic stratospheric aerosol correction for 1982-1984 and 1991-1994 (Tucker *et al.*, 2005). As a consequence, NDVI variations arising from calibration, view

geometry, volcanic aerosols, and other effects not related to actual vegetation change are significantly reduced.

For the purposes of our study, we extracted data documenting the domain 10-20°N/20°W-20°E over the period 1982-2002, and we averaged the 10-day periods of August-September. August-September is the peak of the photosynthetic activity over the Sahel, and the interannual variability of NDVI for these months and domain is strongly coherent (Philippon *et al.*, 2007). Pixels with at least one 10-day value below 0.11 were flagged. Below this value, NDVI is no longer representative of vegetated areas (Martiny *et al.* 2005). In addition, we computed a regional index ('Sahelian NDVI index') which is the spatial average of the 8km pixels correlated at 0.4 with the Sahelian index computed in a previous study by Philippon *et al.* (2007, using an older version of the GIMMS NDVI dataset). This index will be used as predictand in the multiple linear regression model developed in section 5.

b. Atmospheric dynamics:

'Observed' data

The NCEP/DOE 2 reanalyses (Kanamitsu *et al.*, 2002) were used as a proxy for the observed atmospheric dynamics. We selected 4 parameters: the zonal component of the wind at 600hPa (U600 hereafter), the zonal and meridional components of the wind and the specific humidity at 850hPa (U,V and SH850). The 3 latter fields were combined to compute the humidity fluxes at 850hPa (HF850 hereafter). We extracted them over the window 0-40°N/40°W-40°E to capture the enlarged domain of the West African Monsoon (WAM). We selected the months of July to September (JAS) to take into account the one-month time-lag that exists between rainfall (as well as the atmospheric dynamics) and vegetation photosynthetic activity over the region (Malo and Nicholson, 1990, Justice *et al.*, 1991, Martiny *et al.*, 2006), .

The choice of the U600 and HF850 fields was first motivated by their importance for the WAM. Figure 1 top panels provide their climatology for JAS and the period 1982-2002, over the above mentioned domain. U600 documents the African Easterly Jet (AEJ) in JAS, located around 10°N from 20°E to 20°W (Fig. 1 top left panel, -10m/s contour, see also Afiesimama 2007). The AEJ intensity has proved to be negatively correlated with the JAS rainfall amounts (Fontaine and Janicot,

1992, Fontaine *et al.*, 1995, Adedoyin 2000): indeed, a strong AEJ is a limiting factor for deep convection. It is to be noticed that the AEJ is also sensitive to the land surface state, and the soil moisture meridional gradients (Cook 1999) in particular. More generally, from the decadal to the intra-seasonal time-scales, the WAM variability has been proved to be strongly sensitive and related to the land surface conditions (Douville 2002). The HF850 field enables to capture variations in the monsoon layer thickness, and in moisture advections from the Gulf of Guinea and the Mediterranean (Fig. 1 top right panel). Recently, Rowell (2003), Raichich *et al.* (2003) and Jung *et al.* (2006) highlighted a strong teleconnection between the Mediterranean SST and the WAM through moisture advections. A second motivation for the selection of the U600 and HF850 fields is that they are usually correctly simulated by the atmospheric general circulation models (AGCMs). Moron *et al.* (2004) analysed an ensemble of SST forced AGCMs simulations. They showed that interannual variations of the AEJ and the West African Monsoon Index (WAMI) are well reproduced by models in comparison with rainfall. Garric *et al.* (2002) compared the skill of rainfall forecasts based on ARPEGE AGCM atmospheric dynamics (namely indices depicting the AEJ, the WAM or the Tropical Easterly Jet) outputs to the direct rainfall outputs. They noticed that direct rainfall outputs are less skillful than forecasts based on the atmospheric dynamics outputs.

Model outputs

The simulated atmospheric dynamics data set was obtained from the International Research Institute for Climate and Society (IRI). Outputs are from retrospective forecasts with the ECHAM4.5 AGCM (available from http://iridl.ldeo.columbia.edu/SOURCES/.IRI/.FD/.ECHAM4p5/.Forecast/.ca_sst/ and released in early 2006). The retrospective ensemble selected for this study has been produced by forcing ECHAM4.5 with constructed analogue SST ('ca_sst' hereafter) over tropical oceans (30S-30N) and persisted SST over the extra-tropical oceans as boundary conditions. Analogue SST were computed based on the methodology proposed by van den Dool (1994), lying on a linear combination of past observed anomaly patterns as close as desired of the initial state, here the first 5 PCs of the global SST field at consecutive 3-month periods prior to forecast time. Persisted SST were computed by adding the dampened SST anomalies of the preceding months (with a dampened coefficient having an e-folding time of three months) to the climatological annual cycle. The skill comparison

performed by Li and Goddard (2005) demonstrates that retrospective forecasts under analogues significantly improve upon retrospective forecasts under persisted SST anomalies.

The *ca_sst* forecasts consist in 24-member ensemble forecasts covering the period 1957-2002, with lead-times of 0 to 6 months. We selected the simulations starting and initialised in early June and we extracted the same variables as for NCEP/DOE2 (i.e., U600, and U, V, SH850 combined into HF850 over the domain 40°N/40°W-40°E) over the same domain for JAS and the 1982-2002 period. Given the fact that the NDVI is considered from August to September, the statistical models built in this study by use of dynamical outputs have potentially a 2-month lead-time. Notice that all the analyses were performed using the 24-member ensemble mean. Figure 1 bottom panels present the JAS climatological fields of U600 and HF850 as obtained from ECHAM4.5 *ca_sst* simulations. As compared to the reanalyses (Fig. 1 top left), the AEJ is weakened (-8m/s against -10m/s); it is also located northward and extends westward (Fig. 1 bottom left). The 850hPa humidity fluxes (Fig. 1 bottom right) display an enhanced zonal component at 10°N. Conversely, moisture advections from the Mediterranean are reduced as well as the North-East fluxes associated with the Azores high pressure. Despite these differences, ECHAM4.5 *ca_sst* simulates reasonably well the climatological July-September U600 and HF850 fields. ECHAM4.5 skill in simulating the interannual variability is presented and discussed in the 4th section.

c. Rainfall and Sea Surface Temperatures:

The rainfall and sea surface temperature (SST) fields were used to interpret and understand the teleconnections underlying the coupled modes of NDVI / atmospheric dynamics variability detected. Rainfall data were extracted from the Climate Research Unit database which provides data for the global land areas at a monthly time resolution and a 0.5° lat/lon spatial resolution (New *et al.*, 2000). We selected data for the months of July-September (we computed the JAS seasonal amount) and the domain 10-18°N/20°W-20°E. The choice of this data set that documents the period 1900-2000 only, was motivated by the balance between a high spatial resolution (0.5°) and a good quality. For example the CMAP dataset that documents the period 1979-2006 has a coarse resolution of 2.5°. SST data originate from the Hadley Centre HadISST database (Rayner *et al.*, 2003). This global

coverage data set provides monthly mean fields of SST and sea-ice concentration from 1870 to present on a $1^\circ \times 1^\circ$ grid. It improves upon the older GISST2.3b (Global Sea Ice and Sea Surface Temperature version 2.3b) data set. We extracted data documenting the months of July-September and the domain $40^\circ\text{S}-40^\circ\text{N}/180^\circ\text{W}-180^\circ\text{E}$.

3. Methods

The methodology used in this study follows the one proposed by Indeje *et al.* (2006) and comprises two steps: (i) the detection of coupled modes of NDVI / U600-HF850 (from NCEP/DOE2 and ECHAM4.5 ca_sst) variability and (ii) the hindcast of NDVI using a Multiple Linear Regression Model in which the predictors are the U600-HF850 part of the coupled modes detected. This latter approach is commonly referred to as the Model Output Statistics (MOS) approach (Wilks, 2006). It enables to link large-scale simulated fields of atmospheric dynamics to fine-scale observed fields such as rainfall (Landman and Goddard, 2002, Paeth and Hense, 2003) or NDVI (Indeje *et al.* 2006). Note that in this study predictands are (1) a regional index of NDVI and (2) NDVI for each 8-km pixel over the Sahelian domain. An additional predictor has also been considered: the August-September NDVI of the previous year in order to take into account inter-season memory effects highlighted by Martiny *et al.* (2005).

a. The Canonical Correlation Analysis

The detection of the coupled modes is through a Canonical Correlation Analysis (CCA). This method is designed to identify the linear combinations of variables in one field (here a combination of U600 and HF850) that are the most strongly correlated with linear combinations of variables in another field (here NDVI) (Barnston and Ropelewski, 1992, Bretherton *et al.*, 1992, Moron *et al.*, 2001 among others). In this study, each CCA mode is depicted by:

- (i) two time series called 'canonical coefficients' (CC hereafter), one for U600/HF850 and one for NDVI. The U600/HF850 canonical coefficients will be used as potential predictors of NDVI.
- (ii) three 'correlation patterns': the first two maps display the correlations between the U600/HF850 canonical coefficient and the U600 and HF850 fields; they are called

'homogeneous' maps ('*g-map*' hereafter) and localise the covarying parts of U600 and HF850. The third map, called 'heterogeneous' map ('*hn-map*' hereafter), displays the correlations with the NDVI field. The *hn-map* is of importance since it demonstrates the extent to which NDVI is predictable from the atmospheric dynamics. Attached to each map is the percentage of total variance explained (table 2).

As suggested by Barnett and Preisendorfer (1987), it is advisable to remove unnecessary noise in the raw data by applying Principal Component Analyses (PCA) then reconstructing each field as a combination of the first few principal components. We thus submitted the standardised NDVI and U600-HF850 fields to separate PCA and used the PCs as inputs to the CCA instead of the original grid-point values (note that the PCA for U600 and HF850 is a combined PCA). Table 1 displays the number of PCs retained as well as the percentage of variance explained for the NDVI (line 2) and the observed and simulated combined U600-FH850 fields (lines 3-4). The number of PCs retained is chosen according to the scree-slope criterion (Wilks, 2006). The CCA results appear robust: in particular the leading CCA modes are not very sensitive to variations in the number of PCs retained. These leading modes are thought to carry potentially predictable signals. Lastly, note that the CCA has been developed in a leave-one-out cross-validation way (Elsner and Schmertmann 1994). This procedure consists in computing training and testing sub-samples by leaving one year in turn. It usually provides a more realistic measure of skill than considering the whole dataset. Moreover, this procedure is particularly appropriate for small sample size (Barnston and van den Dool, 1993).

b. Building of the Model Output Statistics

The hindcasts of NDVI are based on a Multiple Linear Regression (MLR) model where the cross-validated U600-HF850 canonical coefficients (from NCEP/DOE2 or ECHAM4.5 *ca_sst*) and previous year NDVI value are the predictors. The use of the canonical coefficients as predictors rather than atmospheric dynamics indices (representative of the AEJ or the Mediterranean advections for example) preserves against (*i*) the biases frequently observed in AGCM simulations in the location of the main components of the atmospheric dynamics and (*ii*) colinear predictors. We first run the MLR model with the Sahelian index of NDVI as the predictand, and selected the most

powerful predictors (among the 8 ones available) through a stepwise procedure with a significance level for a predictor to enter or leave the model fixed at 5%. Then all the 8-km resolution pixels analysed were hindcasted using the same predictors. In the case where ECHAM4.5 ca_sst forecasts would be perfect, the skill obtained with the MLR model fed with these forecasts should be the same as the skill obtained with the MLR model fed with NCEP/DOE2 data. Two skill-scores measures are used to evaluate the performance of our models: the correlation coefficient (r_0 hereafter) and the linear error in probability space (LEPS) revised by Potts *et al.* (1996). The adjusted R-square ('R².adj') which is a modification of R-square has also been used. It takes into account the number of explanatory terms in a model, and enables to compare the skill of models using different numbers of predictors. It is defined as

$$R_{adj} = \frac{1 - R^2}{1 - \frac{p}{n}} \quad \text{where } n \text{ is the sample size and } p \text{ the total number}$$

of predictors.

LEPS measures forecasts errors according to their errors in the climatological probability distribution, giving greater relative penalty to forecasts errors near the centre than near the extreme of the cumulative probability distribution (Ward and Folland, 1991). Ranging between -1 and 2, it is defined as

$$LEPS = 3 \left(1 - \frac{|P_v - P_f|}{P_v^2 - P_f^2} \right) - 1 \quad \text{where } P_v \text{ and } P_f \text{ are the cumulative}$$

distribution function of the observation and the forecast respectively. LEPS appears as an alternative unbiased skill measure to the common root mean-squared error and anomaly correlation (Potts *et al.*, 1996).

4. Coupled modes of NDVI / atmospheric dynamics variability and their links with rainfall and SST

The search for coupled modes of NDVI / atmospheric dynamics has been done with NDVI field upscaled to a 48-km spatial resolution using a simple pixel average.

a. Results with NCEP/DOE2 U600-HF850 fields

Figure 2 presents the first CCA mode of NDVI and U600-HF850 from the NCEP/DOE2 reanalyses. Table 2 provides the percentage of variance explained for each of the three fields (column 2), and the correlation between the canonical coefficients (last line, 'R.CC').

The first CCA mode associates positive anomalies of NDVI over the central and eastern Sahel mainly (*hn-map* Fig. 2a, 18.6% of variance explained) with a weakened northward shift of the AEJ (negative correlations along 20°N, *g-map* Fig. 2b), and strengthened moisture advections from the Gulf of Guinea (*g-map* Fig. 2c). Such atmospheric dynamics patterns are known to be favourable to rainfall. This is confirmed by the correlation map of the U600-HF850 mode 1 canonical coefficient (CC1 hereafter) with the JAS rainfall field (Fig. 3b): it features positive values over most of the Sahel, the central and East Sahel recording values above 0.7. Similarly, Figure 3a provides CC1 correlation map with the JAS SST field. Time series were detrended in order to highlight teleconnections at the interannual time-scale. The main signal appears over the Mediterranean Sea with positive correlations. This is not contradictory with Fig. 2c where no significant signal can be observed in the humidity fluxes from the Mediterranean. Indeed, the correlation pattern computed between CC1 and the 850hPa humidity field alone (without combining it with the wind, not shown) depicts a South-West / North-East broad band of significant positive correlations (up to 0.7) spreading from the Mediterranean to the Chad. Positive correlations are also observed with SST over the northern extra-tropical basins. Two additional signals take place in the equatorial Pacific and in the tropical South Atlantic to a lesser extent. Cold SST anomalies in these basins are favourable to high photosynthetic activity over the Sahel (Mennis, 2006, Philippon *et al.* 2007). These results are in agreement with previous studies relating the Sahelian rainfall variability with ENSO (Janicot and Fontaine, 1996, Janicot *et al.*, 2001, Rowell, 2001) and to the equatorial and South tropical Atlantic (Lough, 1986, Druryan, 1991, Fontaine and Janicot, 1996, Janicot *et al.*, 1998, Camberlin *et al.*, 2001 among others). Lastly, negative correlations are also found off the Senegal coast and could explain the low correlations values observed in the *hn-map* (Fig. 2a) and the rainfall field (Fig. 3b) over the West Sahel. In this case, cold SST there inhibit rainfall. A noteworthy point is that when considering the canonical coefficient relative to NDVI (Fig. 2d, line), the significant correlations with the equatorial Pacific are no longer observed while the ones with the equatorial and south tropical Atlantic are

stronger (not shown). This could suggest that NDVI is more sensitive to rainfall variations related to the availability of moisture (advected from the Guinean Gulf) than those due to convection activity level within the ITCZ (partly modulated by ENSO).

The standardised cross-validated canonical coefficients are correlated at 0.81 (Fig. 2d and table 1) and oppose the 'green' years 1988, 1994, 1999 to the 'yellow' years 1983, 1984, 1987. They also feature a positive trend which has already been observed over the Sahel at the pixel scale (Eklundh and Olsson, 2003, Olsson *et al.*, 2005, Heumann *et al.*, 2007) and in the rainfall field (Nicholson, 2005), the greening of the Sahel being primarily attributed to the rainfall fluctuations (Hickler *et al.*, 2005). It is worth mentioning there that the NDVI CC is in good agreement with the Sahelian index built by Philippon *et al.* (2007) as a correlation of 0.76 is found between the two.

Among the next modes (not shown), the 2nd mode depicts more particularly the NDVI variability in the western Sahel (agreeing with results obtained by Jarlan *et al.*, 2005) but explains a low fraction of variance (Table 2, 3rd column). Positive NDVI anomalies there are associated with a cyclonic circulation in the 850hPa humidity flux over the western Sahara, and an AEJ located eastward. The CC associated with this mode feature negative trends as opposed to the positive ones observed for CC1. Correlation maps with rainfall and SST fields highlight above normal rainfall over the Senegal and a La Niña event in the equatorial eastern Pacific in years of enhanced greenness in the Western Sahel. The signal in this oceanic basin is much more marked than the one observed for the 1st CCA mode: correlation values are up to 0.6 against 0.4 respectively. It may reflect a spatial shift in the Sahelian rainfall / ENSO teleconnection over the two last decades. All the next modes explain strongly decreasing variance and become awkward to interpret.

b. Results with ECHAM4.5 ca_sst U600-HF850 fields

The 1st and 2nd modes resulting from the CCA computed between NDVI and U600/HF850 fields from ECHAM4.5 ca_sst simulations, are not shown. Indeed, the variance explained is very low (table 2, columns 4 and 5) and the associated atmospheric dynamics and NDVI patterns are noisy and awkward to interpret. In addition, the canonical coefficients do not show any significant relationship with the observed Sahelian rainfall field nor clear signals with the observed SST field. Thus they are

assumed to be strongly model-dependent. Figure 4 presents the 3rd CCA mode using ECHAM4.5 U600/HF850 fields. This mode roughly matches the 1st mode of the CCA performed with the reanalyses fields (Fig. 2). The *hn-map* (Fig. 4a) depicts an in-phase NDVI variability over the whole Sahelian band, while the *g-maps* show a weaker shifted northward AEJ (Fig. 4b), and reinforced moisture advections from the Guinean Gulf and the Mediterranean (Fig. 4c). However, as compared to reanalyses, correlations values (particularly in the *hn-map*, Fig. 4a) and explained variance (table 2) are less significant, and too much weight is given to advections from the Mediterranean and to those associated with the eastern part of the monsoon. This latter point is consistent with the mean HF850 field (Fig. 1d) for which this bias was already noticed. Correlations of the U600-HF850 canonical coefficient with the JAS SST and rainfall fields (Fig. 5ab) are also in agreement with results obtained with the reanalyses (Fig. 3ab). For rainfall, positive significant correlations are observed from the east of Senegal to the west of Burkina Faso and at the east of Niger (Fig. 5b). For SST, positive (negative) correlations appear over the Mediterranean and the boreal extratropical basins (the equatorial eastern Pacific, Fig. 5a) in agreement with observations (Fig.3a) but once again correlations are less significant. Thus, ECHAM4.5 *ca_sst* simulations capture in a fairly good way the atmospheric dynamics associated with the rainfall and NDVI anomalies over the Sahel and its teleconnection with the SST anomalies. The standardised cross-validated canonical coefficients (Fig. 4d) are correlated at 0.45 ('R.CC' table 2). Among the years of below (above) normal NDVI, notice 1984, 1987, 2002 (1988, 1994) which are also extreme years in the canonical coefficients associated with the NCEP/DOE2 reanalyses (Fig. 2d). CC3 from ECHAM4.5 simulation and CC1 from NCEP/DOE2 are correlated at 0.76 (0.74 when the trends are removed), indicating again the good consistency between the observed and simulated atmospheric dynamics components of the coupled modes.

5. NDVI hindcasts using a Model Output Statistics approach

Analyses carried out in the previous section have highlighted the potential predictability of the NDVI at a 48-km spatial resolution from key components of the large-scale atmospheric dynamics

associated with the West African Monsoon. Moreover, the ECHAM4.5 retrospective forecasts reproduce in a relatively accurate way the main observed coupled modes of variability despite percentages of variance that remain low (table 2, 6th column). The Model Output Statistics approach is now used to hindcast the NDVI at its highest resolution, i.e. 8km, from NCEP/DOE2 reanalyses and ECHAM4.5 forecasts ('ND2 MOS' and 'EC4.5 MOS' hereafter). The former informs about the skill that could be potentially reached with perfect forecasts.

Table 3 presents the canonical coefficients selected as predictors in the ND2 and EC4.5 MOS. It is reminded that selection of predictors is based on MOS developed with a Sahelian NDVI index as predictand. The skill of these MOS is displayed in Table 3 (column 3-4) while the observed and hindcasted time-series of the Sahelian NDVI index are shown in Figure 6. Among the 7 canonical coefficients and the previous year NDVI value available as predictors, only the 1st CC is selected in the ND2 MOS. It enables to explain 77% of the Sahelian NDVI index variance. The EC4.5 MOS uses the 3rd CC and the previous year NDVI value but only explains 30% of variance. This agrees with the lower percentage of variance explained by the EC4.5 modes. For the two MOS, the LEPS scores are above 0.3. The fit between observations and hindcasts appears in Figure 6. It is particularly good for the ND2 MOS (Fig. 6, full circles) : the most noticeable error concerns the key year 1984 for which NDVI is over-estimated. 1984 is the driest year of the 2nd half of the 20th century and is associated with a warm equatorial Atlantic (Druryan and Hastenrath, 1992, Hisard *et al.*, 1986). The over-estimation is thought to be due to the weak teleconnection of CC1 with this basin (Fig. 3a) but also to vegetation memory effects observed from one year to another as proposed by Martiny *et al.* (2005). Indeed, adding the NDVI value of the previous year in the ND2 MOS improves the hindcast for that year. Lastly, the ND2 MOS tends to accentuate the trend. The error in 1984 also concerns the EC4.5 MOS which even more largely over-estimates NDVI (again the over-estimation increases if the NDVI of the previous year is removed from the EC4.5 MOS). The EC4.5 MOS (Fig. 6, stars) appears also less skillfull for the key years 1994 and 1999 that are under-estimated. 1994 was a particularly green year over the Central Sahel (see Philippon *et al.*, 2007, their 3rd mode from an independent component analysis). However, this part of the Sahel is not well depicted in the MOS (NDVI there is depicted by the 2nd CCA mode which has been shown to be unrelated to observed rainfall and SST).

Skill of NDVI hindcasts at 8km spatial resolution are provided in Figure 7. Panels ab display the correlation maps and allow to locate the regions where NDVI is the best hindcasted. Figure 7c summarises the distribution of the correlation values and allows a better comparison of the skill of the two MOS. As expected, the regions recording the highest skill logically match those highlighted in the 'hm-maps' in Fig. 2a and Fig. 4a. On the whole, pixels ranging from 13-18°N and 18°W-20°E have the highest potential predictability, the Sahelian domain being clearly discerned. Most of the pixels located south of this domain, i.e. the Sudanian domain, record r_0 values below 0.3, suggesting that either NDVI there is not predictable from the atmospheric dynamics or is in opposite phase with the northern region. The former alternative is supported by Camberlin *et al.* (2007) study which shows non significant or nil correlation values between annual NDVI and the synchronous rainfall amount over the Guinean and Sudanian regions (see their figure 2). In this case, the low sensitivity of NDVI to rainfall variations logically makes it less prone to be teleconnected to the atmospheric dynamics. The second alternative is to be considered too since Philippon *et al.* (2007) detected a dipole-like mode of NDVI variability over West Africa, opposing the Sahelian domain to the Guinean-Sudanian one in August particularly (this mode prevailed in 1984, 1987, 1998 and 2001). At finer scale, skill variations are also observable within the Sahelian domain. High skill is observable over South-Eastern Mauritania, Central Niger, Western Tchad and Northern Burkina Faso especially for the ND2 MOS (Fig. 7a). As for the Sahelian NDVI index, a loss of skill is observed for the NDVI field with EC4.5 MOS (Fig. 7bc): 24% of the grid-points of the domain considered record a r_0 above 0.5, against 37% with ND2 MOS. A similar behaviour is observed in Indeje *et al.* (2006) results (see their fig. 11). Pixels recording the highest skill in the EC4.5 MOS are located over Mali, Burkina Faso and Niger (Fig. 7b).

According to Barnston and Ropelewski (1992), models with a r_0 below 0.5 are though to be unusable for operational forecasting. However, despite the low skills obtained for a large part of pixels, results are nonetheless encouraging given the fact that (i) they concern 8-km spatial resolution pixels, (ii) they are issued from MOS that do not use any information relative to rainfall.

Lastly, a deeper insight on the skill of the MOS over the Sahelian domain is proposed in Figure 8. Based on the land-cover map for Africa (Mayaux *et al.*, 2004) prepared in the framework of the

Global Land Cover 2000 project (Bartholomé *et al.*, 2005), skill is provided as a function of the cover type. Note that the GLC2000 Africa map provides data at an initial 1km resolution that have been up-scaled at the 8km resolution for the purposes of our study. Pixels are shared among 14 different classes (table 4). Looking at figure 8, it appears that the predictability level varies according to the cover-type. In particular pixels covered by grassland (classes 12 to 15) and cropland (class 17) are the most predictable. Focusing on results obtained with ND2 MOS, two cover types in particular emerge as having a NDVI level noticeably predictable from the atmospheric dynamics: the open grasslands (class 13) and croplands (class 17), the two having a median r_0 above 0.5, i.e. the usable level for operational forecasting. With EC4.5 MOS, the median r_0 reaches only 0.4 for these two classes. This loss of skill agrees with the one observed in figure 7c considering all the grid-points. Note also that differences in predictability among the classes is less pronounced with the EC4.5 (Fig. 8b) than ND2 (Fig. 8a) MOS. These results are concordant with the study by Camberlin *et al.* (2007). These authors noted, for the whole tropical Africa, the highest median correlation between annual rainfall and annual NDVI for pixels covered by the open grasslands (with 80% of the pixels displaying significant correlations). Thus and logically, vegetation for which the photosynthetic activity is the most predictable from AGCMs simulations is the one that is the most strongly sensitive to rainfall variability. A last point concerns pixels classified as 'bare soil' whose predictability is on average higher than the one of woody cover types (7, 9 and 18) for ND2. This could be explained by the mix of bare soil and open or sparse grassland covers in the 1km and 8km spatial resolution pixels.

6. Discussion and Conclusion

The purpose of this study was to assess the predictability of the mean photosynthetic activity of the Sahelian vegetation in August-September. To that aim we looked for relevant potential predictors and developed and tested forecasting models based on a Model Output Statistics approach.

Whereas most of the predictability studies use rainfall as predictors, we worked with predictors depicting the atmospheric dynamics and simulated by an AGCM run in a forecast mode (i.e. forced by constructed analogue SSTs). There are two main reasons for such a choice. Firstly, the interannual

variability of rainfall over West Africa is usually inaccurately reproduced or predicted by atmospheric (Moron et al. 1994) and coupled (Bouali et al. 2008) GCMs that better simulate the decadal variability (Giannini et al. 2003, Hoerling et al 2006). Secondly, the one-month time-lag that exists between rainfall and NDVI is not sufficient for using observed rainfall (or atmospheric dynamics) in operational forecasting, given the frequency of the data updates. By use of AGCM outputs initialised in early June, we take advantage of a 2-month lead-time.

The atmospheric dynamics fields selected are the zonal wind at 600hPa and the humidity fluxes at 850hPa over the domain of the West African Monsoon. They enable to capture the African Easterly Jet and moisture advections from the Gulf of Guinea, which are two key components of the WAM. The selection of these fields is of particular relevance. Indeed, enhanced August-September greenness over the Sahel is associated with a weakened AEJ located northward, and reinforced moisture advections from the Gulf of Guinea and the Mediterranean, in July-September. These atmospheric dynamics patterns of anomalies appear in turn to be linked to a synchronous abnormally warm Mediterranean and cool eastern equatorial Pacific (La Niña event), and lead to above normal rainfall amounts over the Sahel. Then, these signals are reasonably simulated by ECHAM4.5 and along with the NDVI of the previous year (to take into account vegetation memory effects) explain a non negligible part of the NDVI variability: 30% of a Sahelian index variance is explained. At a 8-km spatial resolution, 24% of the pixels analysed have a useful skill (i.e., r_0 above 0.5). If these skills can seem low as compared to the study by Indeje *et al.* (2006) for East Africa, it must be reminded that our models (1) do not use rainfall as predictor, (2) are fully developed in a cross-validation way (applied to the CCA), and (3) tested at the 8km resolution over a 21-yr period. Moreover, these skills can also be partly model dependent. Further investigations and predictability assessment would worth be conducted using outputs from AGCMs forced with persisted SSTs and following the ensemble mean approach (single or multi-model ensemble mean). These latter allow first to reduce uncertainty due to model parametrisation and second to express forecasts in terms of probabilities which is much relevant for societal application purposes. Lastly, the attempt to integrate in our models the 'persistence effect' as observed by Martiny *et al.* (2005), i.e. the NDVI recorded the previous year, improves the skill mainly for the ECHAM4.5 based MOS. It also reduces errors in NCEP/DOE2

based MOS for specific years, such as 1984.

Three additional noticeable points emerged from this study. At first, the degree to which NDVI is predictable significantly varies according to the land-cover type. Pixels covered by grasslands have the most predictable photosynthetic activity. This is coherent with findings by Camberlin *et al.* (2007) showing that the strongest NDVI / rainfall relationship is observed for this land-cover type. Secondly, looking at the correlation map between the rainfall field and the atmospheric dynamics modes, it seems that NDVI is as predictable as the rainfall field (if not more since it is considered at 8km resolution). Following the study by Grist *et al.* (1997), estimations of the rainfall field for areas where NDVI is highly predictable (which are also those where the NDVI / rainfall relationship is the strongest) could then be produced and used e.g., as inputs in hydrological, crop yields (Rasmussen, 1997) and crop prices (Brown *et al.* 2008) models. Lastly, the fact that the teleconnexion patterns with SST of the dynamical and NDVI parts of the NCEP/DOE2 based CCA are different suggest that the intra-seasonal and spatial distribution of rainfall rank number two after the rainfall amount for explaining NDVI variability. This calls for further analyses on NDVI and its links with intra-seasonal components of the rainy season.

Acknowledgements

The authors are grateful to (i) IRI for providing them with ECHAM4.5 ca_sst integrations, made possible by a computing grant from the National Center for Atmospheric Research (NCAR) Climate Simulation Laboratory (CSL) to the IRI (Principal Investigator Steve Zebiak); (ii) the Global Land Cover Facility (GLCF) which is the source for NDVI data; (iii) the AMMA program. Based on a French initiative, AMMA was built by an international scientific group and is currently funded by a large number of agencies, especially from France, UK, US and Africa. It has been the beneficiary of a large financial contribution from the European Community's sixth Framework Research Program. Detailed information on scientific coordination and funding is available on the AMMA International web site : <http://www.amma-international.org>. The authors also thank an anonymous reviewer for his comments as well as Pr V Moron for the fruitful discussion.

References

- Adedoyin JA 2000. SST-induced climate change in tropical North Africa: the intermediary role of the lower tropospheric oscillations. *Meteorol. Atmos. Phys.* 75: 135-147.
- Afiesimama EA 2007. Annual cycle of the mid-tropospheric easterly jet over West Africa. *Theor. Appl. Climatol.* Doi: 10.1007/s00704-006-0284y.
- Anyamba A, Tucker C 2005. Analysis of Sahelian vegetation dynamics using NOAA-AVHRR NDVI data from 1981-2003. *Journal of Arid Environment* 63: 596-614.
- Barnett CW, Preisendorfer R 1987. Origins and levels of monthly and seasonal forecasts skill for United States surface temperatures determined by canonical correlation analysis. *Mon. Wea. Rev.* 115: 1825-1850.
- Barnston AG, Ropelewski CF 1992. Prediction of ENSO episodes using Canonical Correlation Analysis. *J. Climate* 5: 1316-1345.
- Barnston AG, van den Dool HM 1993. A degeneracy in cross-validated skill in regression-based forecasts. *J. Climate* 6: 963-977.
- Bartholomé E, Belward AS 2005. GCL2000: a new approach to global land cover mapping from Earth observation data. *Int. J. Remote Sens.* 26: 1959-1977.
- Bouali L, Philippon N, Fontaine B, Lemond J 2008. Performance of DEMETER calibration for rainfall forecasting purposes: Application to the July–August Sahelian rainfall. *J. Geophys. Res.* 113: D15111, doi:10.1029/2007JD009403.
- Bretherton CS, Smith C, Wallace JM 1992. An intercomparison of methods for finding coupled patterns in climate data. *J. Climate* 5: 541-560.
- Brown ME, Pinzon JE, Prince SD 2008. Using Satellite Remote Sensing Data in a Spatially Explicit Price Model: Vegetation Dynamics and Millet Prices. *Land Economics*, 84: 340-357 doi:10.3368/le.84.2.340
- Camberlin P, Janicot S, Pocard I 2001. Seasonality and atmospheric dynamics of the teleconnection between african rainfall and tropical sea-surface temperature: Atlantic vs. ENSO. *Int. J. Climatol.* 21: 973-1005.
- Camberlin P, Martiny N, Philippon N, Richard Y 2007. Determinants of the interannual relationships between remote sensed photosynthetic activity and rainfall in tropical Africa. *Remote Sens. Environ.* 106: 199-216.
- Cook KH 1999. Generation of the African Easterly Jet and its role in determining West African precipitation. *J. Climate* 12: 1165-1184.
- Douville, H., 2002: Influence of Soil Moisture on the Asian and African Monsoons. Part II: Interannual Variability. *J. Climate*, **15**, 701–720.
- Dregne HE, Tucker CJ 1988. Green biomass and rainfall in semi-arid sub-Saharan Africa. *Journal of*

Arid Environment 15: 245-252.

- Druyan LM 1991. The sensitivity of sub-Saharan precipitation to Atlantic SST. *Climatic Change* 18: 17-36.
- Druyan LM, Hastenrath S 1992. GCM simulation of the Sahel 1984 drought with alternative specifications of observed SST. *Int. J. Climatol.* 12: 521-526.
- Eklundh L, Olsson L 2003. Vegetation index trends for the African Sahel 1982-1999. *Geophys. Res. Lett.* 30: doi:10.1029/2002GL016772.
- Elsner JB, Schmertmann CP 1994. Assessing forecast skill through cross validation. *Weather and Forecasting* 9: 619-624.
- Fontaine B, Janicot S 1992. Wind-Field Coherence and Its Variations over West Africa. *J. Climate* 5: 512-524.
- Fontaine B, Janicot S 1996. Near-global sea surface temperature variability associated with West African anomaly types. *J. Climate* 9: 2935-2940.
- Fontaine B, Janicot S, Moron V 1995. Rainfall anomaly patterns and wind field signals over West Africa in August (1958-1989). *J. Climate* 8: 1503-1510.
- Funk CC, Brown ME 2006. Intra-seasonal NDVI change projections in semi-arid Africa. *Remote Sens. Environ.* 101: 249-256.
- Garric G, Douville H, Déqué M 2002. Prospects for improved seasonal predictions of monsoon precipitation over West Africa. *Int. J. Climatol.* 22: 331-345.
- Giannini A, Saravanan R., P. Chang 2003. Oceanic Forcing of Sahel Rainfall on Interannual to Interdecadal Time Scales. *Science.* 302: 1027-1030.
- Grist J, Nicholson S, Mpolokang A 1997. On the use of NDVI for estimating rainfall fields in the Kalahari of Bostwana. *Journal of Arid Environment* 35: 195-214.
- Groten SME 1993. NDVI-crop monitoring and early yield assessment of Burkina Faso. *Int. J. Remote Sens.* 14: 1495-1515.
- Gulaid AA 1986. Contribution of remote sensing to food security and early warning systems in drought affected countries in Africa. *7th Remote sensing for ressources development and environmental management International symposium.* Rotterdam 25th August. 457-460.
- Herrmann SM, Anyamba A, Tucker CJ 2005. Recent trends in vegetation dynamics in the African Sahel and their relationship to climate. *Global Environmental Change* 15: 394-404.
- Heumann BW, Seaquist JW, Eklundh L, Jönsson P 2007. AVHRR derived phenological change in the Sahel and Soudan, Africa, 1982-2005. *Remote Sens. Environ.* 108: 385-392..
- Hickler T, Eklundh L, Seaquist JW, Smith B, Ardö J, Olsson L, Sykes MT, Sjöström M 2005. Precipitation controls Sahel greening trend. *Geophys. Res. Lett.* 32: L21415, doi:10.1029/2005GL024370.

- Hisard P, Henin C, Houghton R, Piton B, Rual P 1986. Oceanic conditions in the tropical Atlantic during 1983 and 1984. *Nature* 322: 243-245.
- Hoerling MP, Hurrell JW, Eischeid J, Phillips AS 2006: Detection and attribution of 20th century northern and southern African monsoon change. *J. Climate* 19: 3989–4008.
- Hutchinson CF, 1991. Uses of satellite data for famine early warning in sub-Saharan Africa. *Int. J. Remote Sensing*. 12: 1405-1421. DOI: 10.1080/01431169108929733
- Indeje M, Ward NM, Ogallo LJ, Davies G, Dilley M, Anyamba A 2006. Predictability of the Normalized Difference Vegetation Index in Kenya and potential applications as an indicator of Rift Valley fever outbreaks in the Greater Horn of Africa. *J. Climate* 19: 1673-1687.
- Janicot S, Moron V, Fontaine B 1996. Sahel droughts and ENSO dynamics. *Geophys. Res. Lett.* 23: 515-518.
- Janicot S, Trzaska S, Pocard I 2001. Summer Sahel-ENSO teleconnection and decadal time scale SST variations. *Climate Dyn.* 18: 303-320.
- Janicot S, Harzallah A, Fontaine B, Moron V 1998. West African monsoon dynamics and the Eastern equatorial Atlantic and Pacific SST anomalies (1970-88). *J. Climate* 11: 1874-1882.
- Jarlan L, Tourre YM, Mougou E, Philippon N, Mazzega P 2005. Dominant patterns of AVHRR NDVI interannual variability over the Sahel and linkages with key climate signals (1982-2003). *Geophys. Res. Lett.* 32: doi:10.1029/2004GL021841.
- Jung T, Ferranti L, Tompkins AM 2006. Response to the summer of 2003 Mediterranean SST anomalies over Europe and Africa. *J. Climate* 19: 5439-5454.
- Justice CO, Hiernaux P 1986. Monitoring the grasslands of the Sahel using NOAA/AVHRR data. *Int. J. Remote Sens.* 7: 1475-1498.
- Justice CO, Dugdale G, Townshend JR, Narracott AS, Kumar M 1991. Synergism between NOAA-AVHRR and Meteosat data for studying vegetation development in semi-arid West Africa. *Int. J. Remote Sens.* 12: 1349-1368.
- Kanamitsu M, Ebisuzaki W, Woollen J, Yang SK, Hnilo JJ, Fiorino M, Potter GL 2002. NCEP-DEO AMIP-II Reanalysis (R-2). *Bull. Amer. Meteor. Soc.* 83: 1631-1643.
- Landman WA, Goddard L 2002. Statistical recalibration of GCM forecasts over Southern Africa using model output statistics. *J. Climate* 15: 2038-2055.
- Li S, Goddard L 2005. Retrospective forecasts with the ECHAM4.5 AGCM. The International Research Institute for Climate and Society. IRI Technical Report, IRI-TR /05/2: 16.
- Linthicum KJ, Anyamba A, Tucker CJ, Kelley PW, Myers MF, Peters CJ 1999. Climate and satellite indicators to forecast Rift Valley fever epidemics in Kenya. *Science* 285: 397-400.
- Lough JM 1986. Tropical Atlantic sea surface temperature and rainfall variations in subsaharan Africa. *Mon. Wea. Rev.* 114: 561-570.
- Malo AR, Nicholson S 1990. A study of the rainfall and vegetation dynamics in the African Sahel

using the normalised difference vegetation index. *Journal of Arid Environment* 19: 1-24.

Martiny N, Camberlin P, Richard Y 2005. Interannual persistence effects in vegetation dynamics of semi-arid Africa. *Geophys. Res. Lett.* 32: L24403, doi:10.1029/2005GL024634.

Martiny N, Camberlin P, Richard Y, Philippon N 2006. Compared regimes of NDVI and rainfall in semi-arid regions of Africa. *Int. J. Remote Sens.* 27: 5201-5223.

Martiny N, Philippon N, Camberlin P, Richard Y 2008. Predictability of NDVI in semi-arid regions of Africa from ocean-atmospheric fields. *J. Climate*. revised.

Maselli F, Romanelli S, Bottai L, Maracchi G 2000. Processing of GAC NDVI data for yield forecasting in the Sahelian region. *Int. J. Remote Sens.* 21: 3509 - 3523.

Maselli F, Conese C, Petkov L, Gilabert MA 1993. Environmental monitoring and crop forecasting in the Sahel through the use of NOAA NDVI data. A case study: Niger 1986-89. *Int. J. Remote Sens.* 14: 3471 – 3487.

Mayaux P, Bartholomé E, Fritz S, Beward A 2004. A new land-cover map of Africa for the year 2000. *Journal of Biogeography* 31: 861-877.

Mennis J 2006. Exploring the influence of ENSO on African vegetation variability using Multidimensional Map Algebra. *GIScience and Remote Sensing*. 43: 352-376.

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

Moron V, Ward MN, Navarra A 2001. Observed and SST-forced seasonal rainfall variability across tropical America. *Int. J. Climatol.* 21: 1467-1501.

Moron V, Philippon N, Fontaine B 2004. Simulation of a West-African monsoon index in four Atmospheric General Circulation Models forced by prescribed SST. *J. Geophys. Res.* 109: doi:10.1029/2004JD004760.

New M, Hulme M, Jones P 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 2005. On the question of the "recovery" of the rains in the West African Sahel. *Journal of Arid Environment* 63: 615-641.

Nicholson S, Davenport M, Malo AR 1990. A comparison of the vegetation response to rainfall in the Sahel and East Africa using Normalized Difference Vegetation Index from NOAA AVHRR. *Climatic Change* 17: 209-241.

Olsson L, Eklundh L, Ardö J 2005. A recent greening of the Sahel - trends, patterns and potential causes. *Journal of Arid Environment* 63: 556-566.

Paeth H, Hense A 2003. Seasonal forecast of sub-sahelian rainfall using cross validated model output statistics. *Meteorologische zeitschrift* 12: 157-173.

Philippon N, Jarlan L, Martiny N, Camberlin P, Mougou E 2007. Characterization of the interannual and intraseasonal variability of West African vegetation between 1982 and 2002 by means of

NOAA AVHRR NDVI data. *J. Climate* 20: 1202-1218.

Potts JM, Folland CK, Jolliffe IT, Sexton D 1996. Revised LEPS scores for assessing climate model simulations and long-range forecasts. *J. Climate* 9: 34-53.

Raich F, Pinardi N, Navarra A 2003. Teleconnections between Indian monsoon and Sahel rainfall and the Mediterranean. *Int. J. Climatol.* 23: 173-186.

Rasmussen MS 1997. Operational yield forecast using AVHRR NDVI data: Reduction of environmental and inter-annual variability. *Int. J. Remote Sens.* 18: 1059– 1077.

Rowell DP 2001. Teleconnections between the tropical Pacific and the Sahel. *Quart. J. Roy. Meteor. Soc.* 127: 1683-1706.

Rowell DP 2003. The impact of Mediterranean SSTs on the Sahelian rainfall season. *J. Climate* 16: 849-862.

Tucker CJ 2005. An Extended AVHRR 8-km NDVI Data Set compatible with MODIS and SPOT Vegetation NDVI Data. *Int. J. Remote Sens.* 26: 4485-4498.

Tucker CJ, Justice CO, Prince SD 1986. Monitoring the grasslands of the Sahel 1984-1985. *Int. J. Remote Sens.* 7: 1571-1581.

Tucker CJ, Dregne HE, Newcomb WW 1991. Expansion and contraction of the Sahara Desert between 1980 and 1990. *Science* 253: 299-301.

van den Dool HM 1994. Searching for analogues, how long must we wait? *Tellus* 46A: 314-324.

Ward MN, Folland CK 1991. Prediction of seasonal rainfall in North Nordeste Brasil using eigenvectors of sea surface temperature. *Int. J. Climatol.* 11: 711-743.

Wylie BK, Harrington Jr. JA, Prince SD, Denda I 1991. Satellite and ground-based pasture production assessment in Niger: 1986-1988. *Int. J. Remote Sensing.* 12: 1281-1300. DOI: 10.1080/01431169108929726

Wilks DS 2006. Statistical methods in the atmospheric sciences, Dmowska R, Hartmann D and Rossby HT. Elsevier: San Diego; 627.

Table 1: percentage of variance and number of modes kept after the PCA pre-filtering of NDVI and U600/HF850 (from NCEP/DOE2 and ECHAM4.5 ca_sst) fields.

		% of variance	Number of modes
<i>NDVI</i>		.64	7
<i>U600/HF850</i>	NCEP/DOE2	.81	10
	ECHAM4.5 ca_sst	.93	11

Table 2: percentage of variance explained by the main CCA modes for the NDVI field, and the U 600hPa and HF 850hPa fields from NCEP/DOE2 and ECHAM4.5 ca_sst. 'R.CC' is the correlation between the canonical coefficients. the CCA has been developed in a leave-one-out cross-validation mode.

	NCEP/DOE2		ECHAM4.5 ca_sst		
	CCA 1	CCA 2	CCA 1	CCA 2	CCA 3
<i>NDVI</i>	18.6	6.4	5.4	4.4	7.3
<i>U600</i>	20.8	12.1	8.5	8.9	13.4
<i>HF850 (U part)</i>	20.3	10.6	6.5	8.2	13.3
<i>HF850 (V part)</i>	10.7	10.1	6.5	6.9	9.8
<i>R.CC</i>	.81	.62	.16	.12	.45

Table 3: number of the canonical coefficients selected as predictors, r_0 , adjusted R^2 and LEPS skill-scores of the 1-yr cross-validated MOS models ('ND2' and 'EC4.5' in text) based on a Sahelian NDVI index as predictand.

	CC retained	r_0 (R^2 .adj)	LEPS
<i>ND2 (NCEP/DOE2)</i>	N° 1	.77 (56%)	.47
<i>EC4.5 (ECHAM4.5 Ca_sst + NDVI y-1)</i>	N° 3	.58 (30%)	.32

Table 4: number and denomination of the cover type associated with the land-cover classes defined in GLC2000 for pixels considered in the study.

Class number	Cover type
<i>1</i>	Closed evergreen lowland forest
<i>5</i>	Mangrove
<i>7</i>	Mosaic forest / savanna
<i>9</i>	Deciduous woodland
<i>10</i>	Deciduous shrubland
<i>12</i>	Closed grassland
<i>13</i>	Open grassland with sparse shrubs
<i>14</i>	Open grassland
<i>15</i>	Sparse grassland
<i>16</i>	Swamp bushland and grassland
<i>17</i>	croplands
<i>18</i>	Croplands with open woody vegetation
<i>19</i>	Irrigated croplands
<i>21</i>	Bare soils

Figure 1: mean July-September fields of the zonal wind at 600hPa (left panels) and the humidity fluxes at 850hPa (right panels) from NCEP/DOE2 reanalyses (top panels) and ECHAM4.5 ca_sst (bottom) for the period 1982-2002.

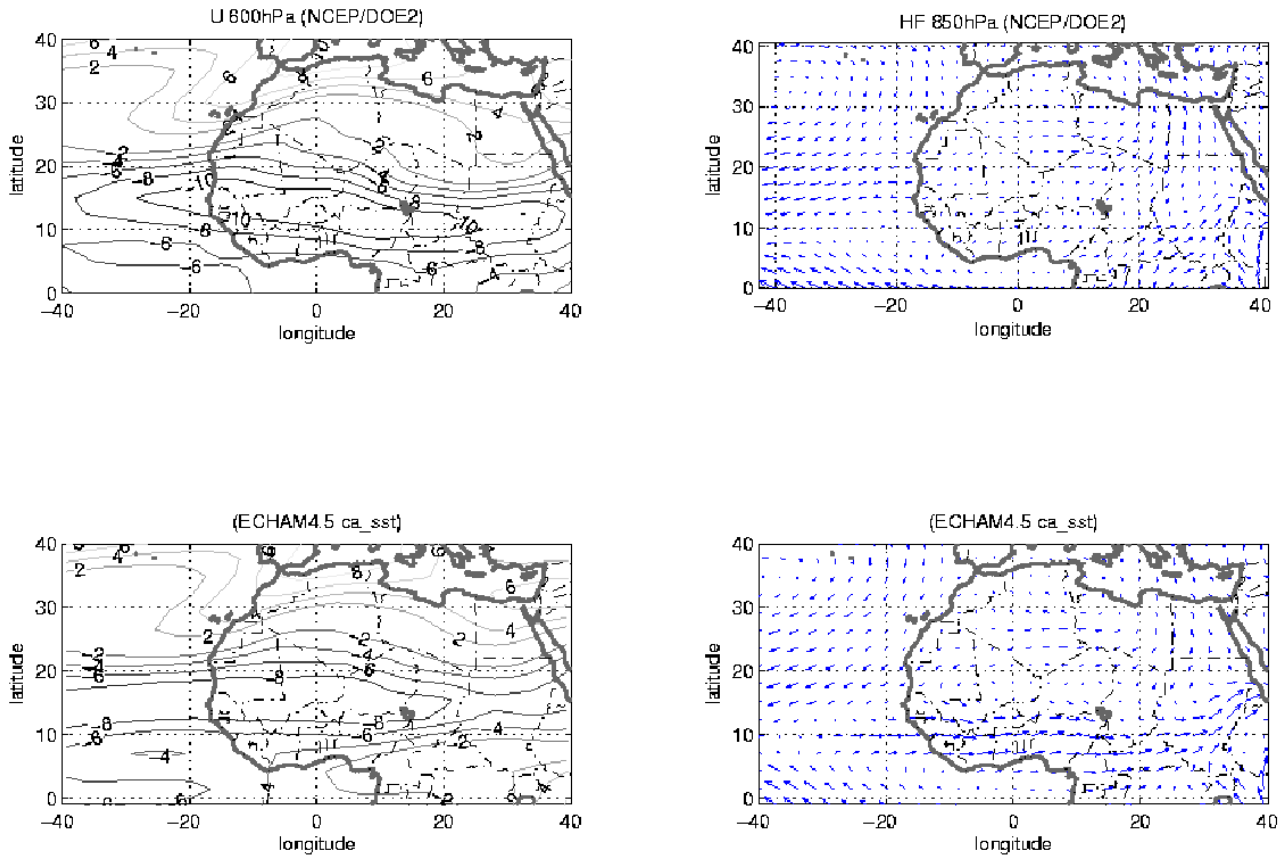


Figure 2: 1st CCA mode between NDVI and NCEP/DOE2 U600/HF850. (a) NDVI heterogeneous pattern (*hn-map*). (b and c) U600 and HF850 homogeneous patterns (*g-map*). (d) canonical coefficients. Shading in panels bc denotes correlations significant at the 95% level according to Student's T-test. CCA has been developed in a leave-one-out cross-validation.

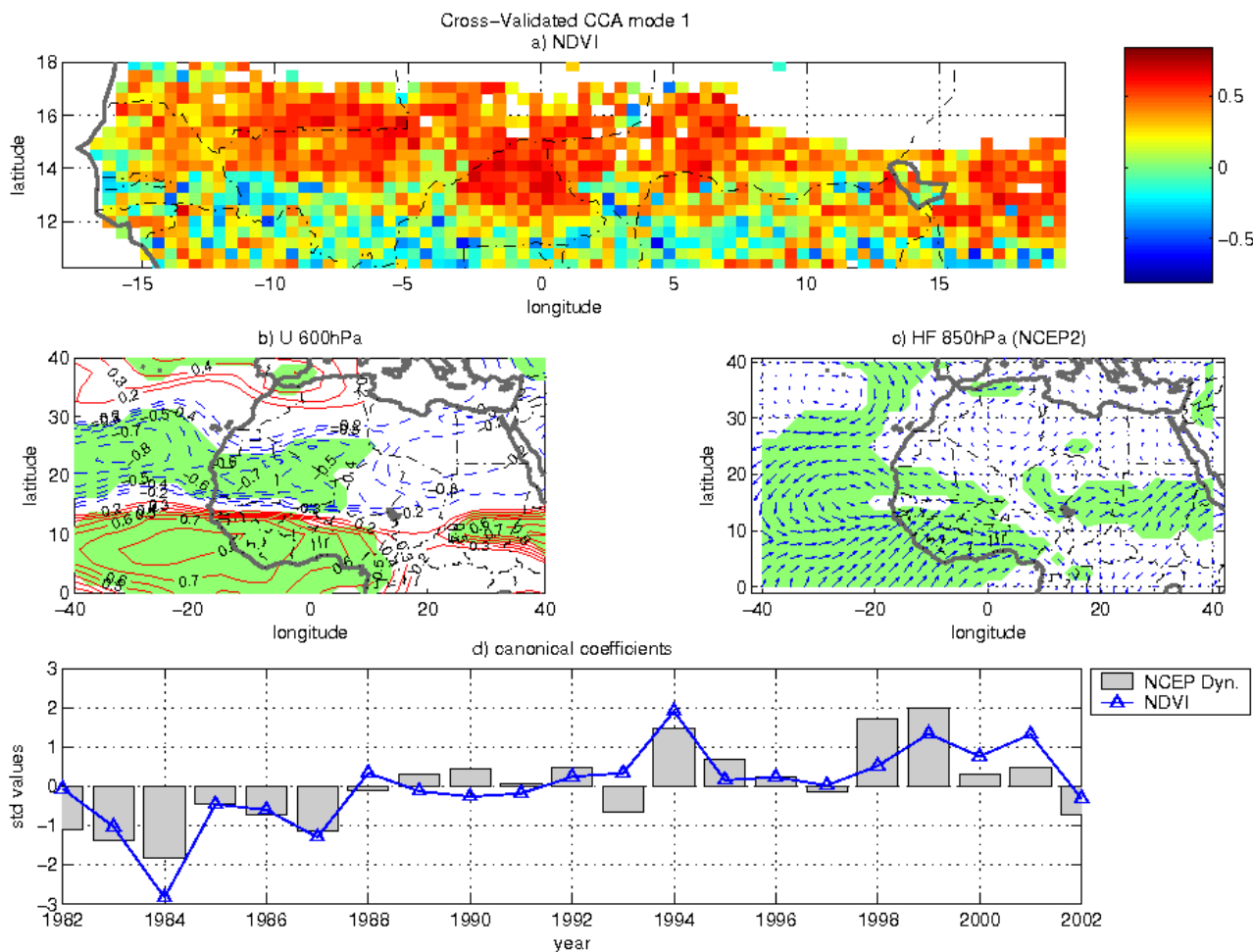


Figure 3: Correlation maps between the cross-validated 1st canonical coefficient from the CCA between NDVI and NCEP/DOE2 atmospheric dynamics and (a) July-September SST field (detrended), and (b) rainfall field (from CRU). Thick (thin) contours for positive (negative) correlations. Shading denotes correlations significant at the 95% level according to Student's T-test.

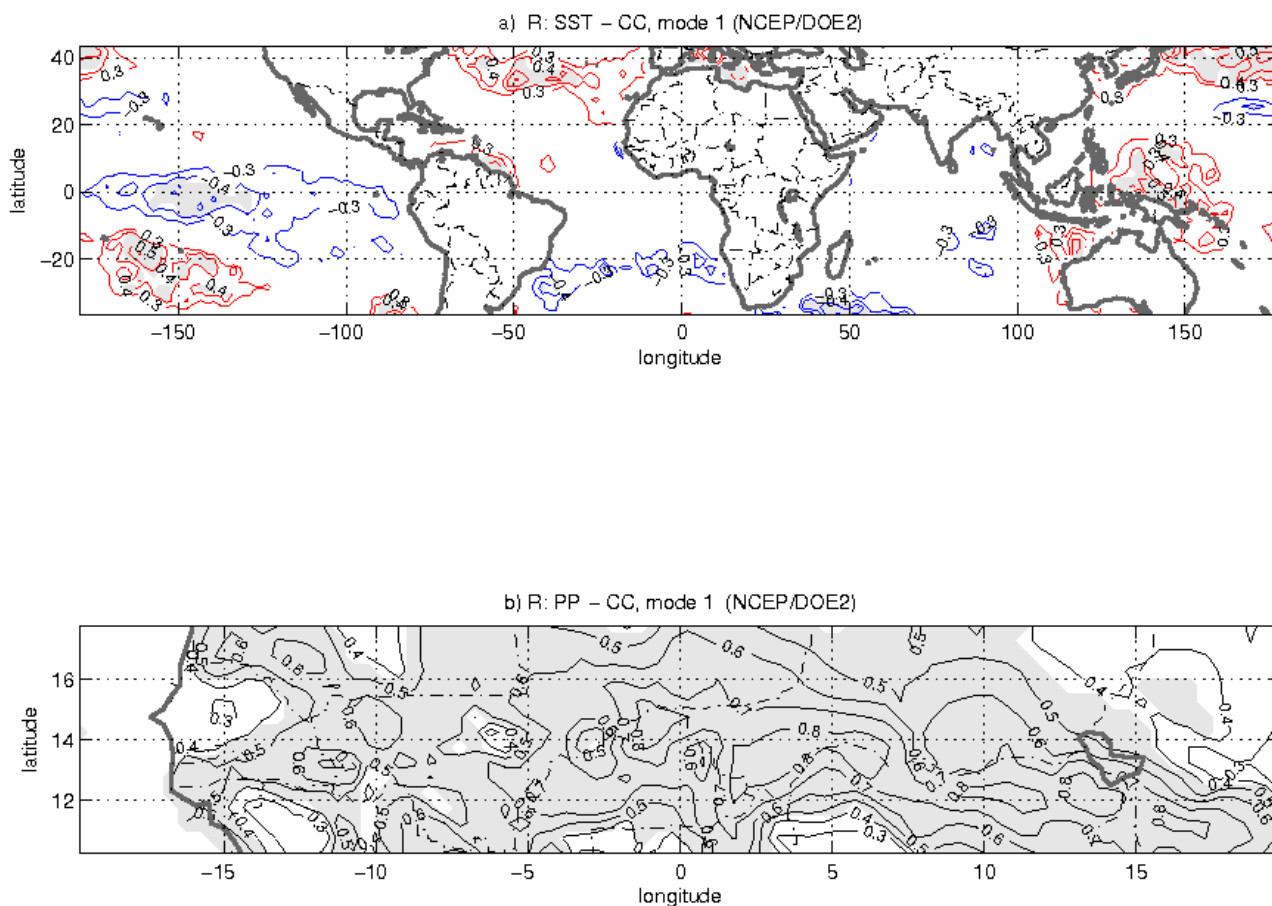


Figure 4: as Figure 2 but between NDVI and ECHAM4.5 ca_sst U600/HF850 and for the 3rd CCA mode.

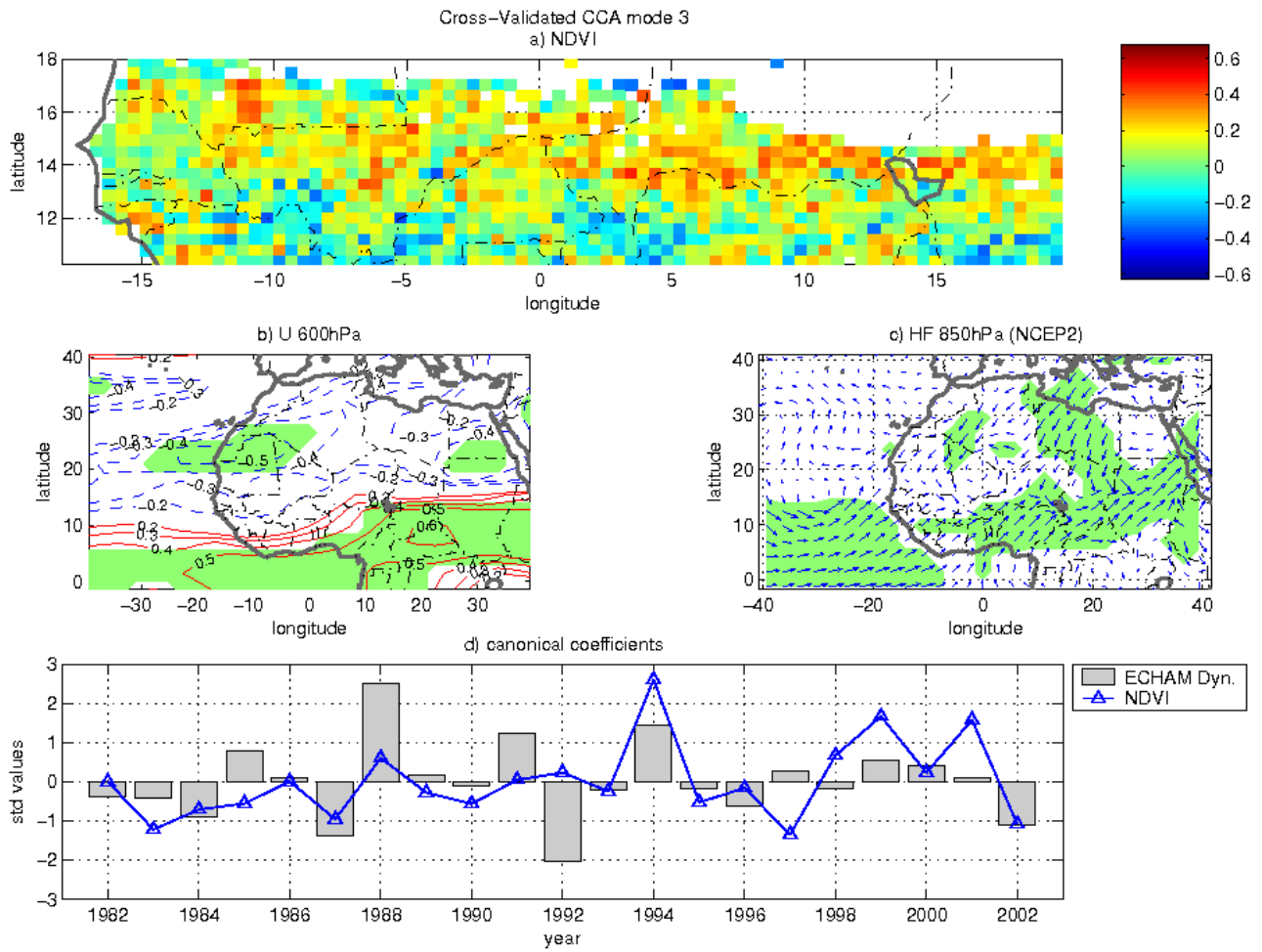


Figure 5 : as Figure 3 but between the 3rd canonical coefficient from the CCA between NDVI and ECHAM4.5 ca_sst atmospheric dynamics.

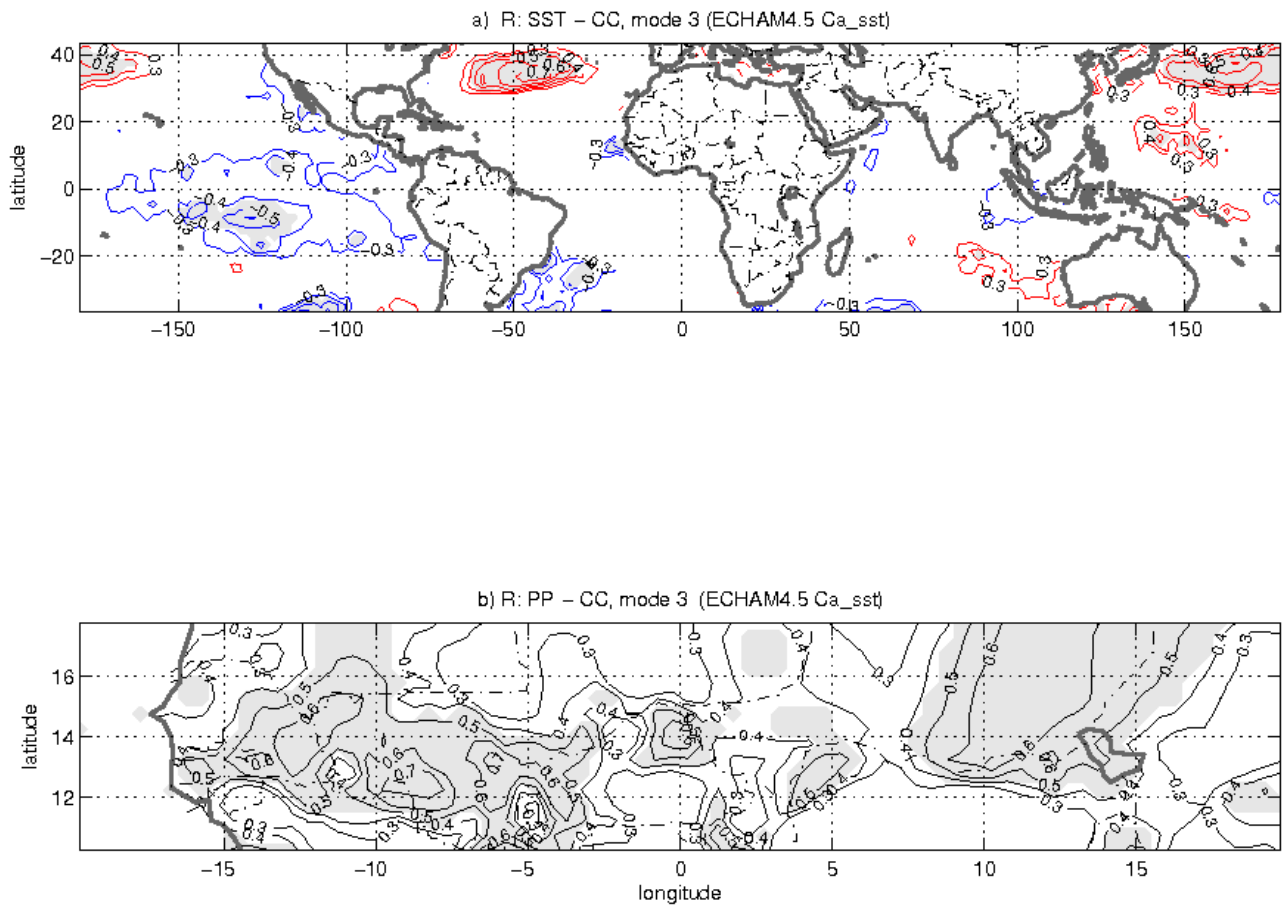


Figure 6: 1-yr cross-validated hindcasts from ND2 (circle) and EC4.5(*) MOS (see text), and observations (bar) of a standardised Sahelian NDVI index.

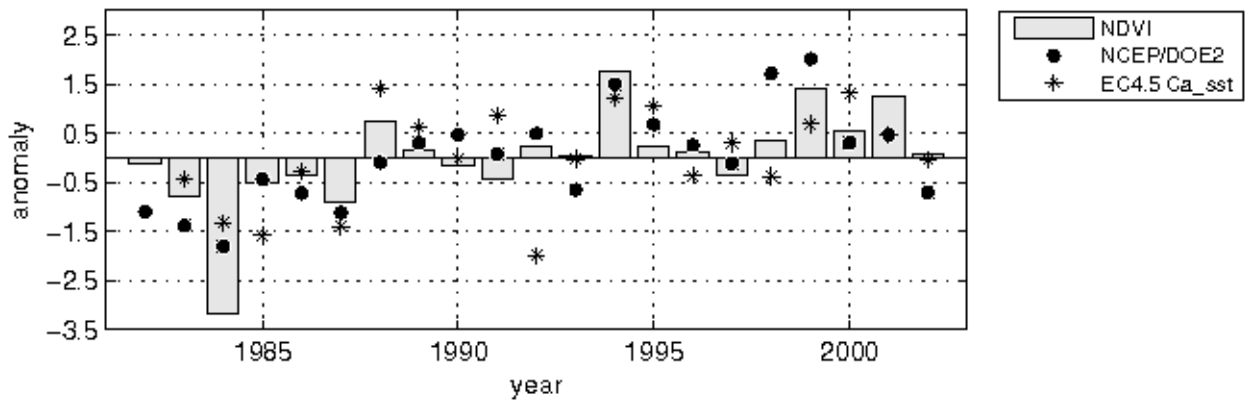


Figure 7: Correlation (r_0) maps between 1-yr cross-validated observed and hindcasted NDVI using (a) ND2 and (b) EC4.5 MOS. (c) cumulative percentage of grid points as a function of the correlation value threshold for hindcasts from ND2 (black line) and EC4.5 (red line) MOS. The dash line indicates the $r_0=0.5$ threshold above which modes are usable for operational forecasting. The full line indicates the 95% significance threshold for r_0 .

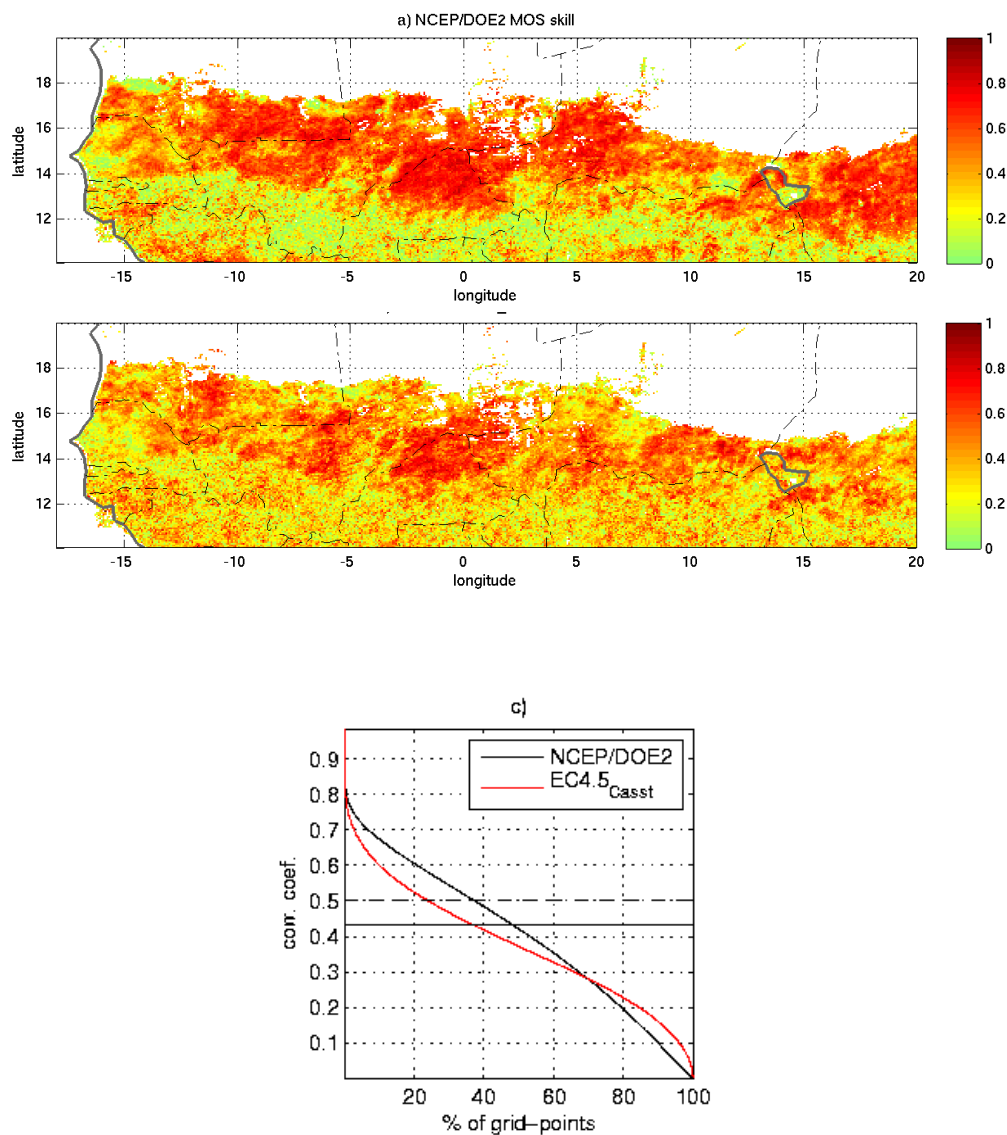


Figure 8 : Box-plot of correlation coefficients (r_0) between 1-yr cross-validated observed and hindcasted NDVI from (a) ND2 and (b) EC4.5 MOS, as a function of the land-cover type (as defined by GLC2000; see full classes denominations under table 4). Only those classes which are represented by at least 100 grid-points are shown. Top numbers: number of grid-points. Horizontal dashed line: 95% confidence level. Box-plots are represented in the usual way (i.e., lower and upper limits of the box: lower and upper quartiles; central line: median; whiskers: lowest and highest values, with '+' showing outliers beyond 1.5 times the inter-quartile range).

