CrossMark

Decadal prediction of Sahel rainfall using dynamics-based indices

Noelia Otero[1] · Elsa Mohino[2] · Marco Gaetani[3]

Abstract At decadal time scales, the capability of state-of-the-art atmosphere-ocean coupled climate models in predicting the precipitation in Sahel is assessed. A set of 14 models participating in the Coupled Model Intercomparison Project Phase 5 (CMIP5) is selected and two experiments are analysed, namely initialized decadal hindcasts and forced historical simulations. Considering the strong linkage of the atmospheric circulation signatures over West Africa with the rainfall variability, this study aims to investigate the potential of using wind fields for decadal predictions. Namely, a West African monsoon index (WAMI) is defined, based on the coherence of low (925 hPa) and high (200 hPa) troposphere wind fields, which accounts for the intensity of the monsoonal circulation. A combined empirical orthogonal functions analysis is applied to explore the wind fields' covariance modes, and a set of indices is defined on the basis of the identified patterns. The WAMI predictive skill is assessed by comparing WAMI from coupled models with WAMI from reanalysis products and with a standardized precipitation index (SPI) from observations. Results suggest that the predictive skill is highly model dependent and it is strongly related to the WAMI definition. In addition, hindcasts are more skilful than historical simulations in both deterministic and probability forecasts, which suggests an added value of initialization for decadal predictability. Moreover, coupled models are more skilful in predicting the observed SPI than the WAMI obtained from reanalysis. WAMI performance is also compared with decadal predictions from CMIP5 models based on a Sahelian precipitation index, and an improvement in predictive skill is observed in some models when WAMI is used. Therefore, we conclude that dynamics-based indices are potentially more effective for decadal prediction of precipitation in Sahel than precipitation-based indices for those models in which Sahel rainfall variability is not well simulated. We thus recommend a two-fold approach when testing the performance of models in predicting Sahel rainfall, based not only on rainfall but also on the dynamics of the West African monsoon.

✉ Noelia Otero
  Noelia.OteroFelipe@iass-potsdam.de

[1] Institute for Advanced Sustainability Studies (IASS), Postdam, Germany

[2] Dpto. Física de la Tierra, Astronomía y Astrofísica I, Universidad Complutense de Madrid, Madrid, Spain

[3] LATMOS-IPSL, CNRS, Sorbonne Universites, UPMC and UVSQ, Paris, France

1 Introduction

The Sahel is a semiarid region in West Africa between 10°N and 20°N. The economy of most of the Sahelian countries is based on rain-fed agriculture and pastures for livestock, which makes them highly vulnerable to rainfall variability (Kandji et al. 2006; Ickowicz et al. 2012). The Sahel suffered an intense and prolonged drought during the 1970s and 1980s, which is recognized as one of the main recent climate change signals (Dai et al. 2004; Trenberth

et al. 2007), with dramatic socio-economic consequences, noticeably on food security (McIntire 1981; Kandji et al. 2006). Rainfall variability is also highly relevant in other sectors as public health, because it is a precursor for vector-borne diseases like Malaria (Diouf et al. 2013). Therefore, reliable predictions of Sahel rainfall variability and trends could be highly beneficial to manage climate risk and vulnerability at seasonal-to-decadal time scales.

Rainfall in the Sahel is driven by the West African monsoon (WAM) system, which is the dominant circulation regime in the region (Nicholson 2013). The Sahel rainy season starts in late June and lasts throughout September, and its variability is strongly related to sea surface temperature (SST) anomalies (Folland et al. 1986; Rodríguez-Fonseca et al. 2015). At decadal time scales, many studies have shown a connection between Sahel drought and a warming of the tropical Indian and Pacific oceans (e.g. Bader and Latif 2003; Giannini et al. 2003; Lu and Delworth 2005; Caminade and Terray 2010; Mohino et al. 2011c). The Atlantic multidecadal variability (AMV) and the Pacific decadal oscillation (PDO) have also been shown to drive Sahel rainfall (e.g. Lu and Delworth 2005; Hoerling et al. 2006; Mohino et al. 2011a; Villamayor and Mohino 2015). In addition, some works point to a significant contribution of radiative external forcings (Haarsma et al. 2005; Biasutti et al. 2008). In this context, it has been highlighted that the 1970s–1980s drought in the Sahel was mainly driven by SST variability and amplified by land-surface processes (e.g. Zeng et al. 1999; Giannini et al. 2003). Such a strong linkage with SSTs makes climate predictions of Sahel rainfall at decadal time scales potentially feasible, which is of prominent importance for the management of climate risk and vulnerability. Moreover, some authors point out the relevance of feedback mechanisms that operate between the vegetation and the atmosphere, which may be responsible for the severity and persistence of the drought (Wang and Eltahir 2000).

Decadal climate predictions aim at providing climate information from a few years to a few decades into the future. At such a time horizon, predictability is influenced by both internal variability of the climate system and external forcings, combining initial and boundary condition problems (Meehl et al. 2009). Decadal forecast systems are based on coupled atmosphere-ocean climate models running for 10-to-30 years. The models are initialized with the observed climate state at the beginning of the simulation and forced by the observed and/or projected changes in the external forcings, specifically the anthropogenic emissions (Kirtman et al. 2013). Specific efforts are devoted to the assessment of predictive performances and the improvement of the forecast systems for decadal climate predictions. Early studies show that climate models are able to produce skilful decadal

predictions of global SSTs (Smith et al. 2007) and also at regional scale over the North Atlantic and Pacific oceans (Keenlyside et al. 2008; Mochizuki et al. 2010). In the framework of the ENSEMBLES FP6 project, a small set of forecast systems were used for running coordinated decadal simulations (Doblas-Reyes et al. 2010), producing encouraging results for sea surface temperatures over the North Atlantic (García-Serrano and Doblas-Reyes 2012). More recently, in the context of the Coupled Model Intercomparison Project Phase 5 (CMIP5; Taylor et al. 2012), results from a broader set of models confirm the potential decadal predictability of global and large-scale surface temperatures (Kim and Webster 2012; Doblas-Reyes et al. 2013).

Despite the promising outcomes for temperature, decadal predictability of precipitation at global scale is still problematic (Kirtman et al. 2013), and studies specific to Sahel rainfall also show ambiguous results (van Oldenborgh et al. 2012). Through the evaluation of decadal hindcasts in a set of ENSEMBLES simulations, García-Serrano et al. (2013) suggested no skill for Sahel rainfall predictions. However, Gaetani and Mohino (2013) showed that significant skill can be obtained from some CMIP5 models, a result later on confirmed by Martin and Thorncroft (2014). The limitations in the model simulation of precipitation are related to inadequate description of the convective processes (Kohler et al. 2010), as well as to the underestimation of the interannual variability (Ineson and Scaife 2009). As a consequence, confidence in decadal predictability and future projections of Sahelian precipitation is still low (Kirtman et al. 2013).

So far, the studies on decadal prediction of Sahel rainfall have analysed the precipitation fields simulated by models. However, several works have shown that, even forced with prescribed SSTs, global circulation models (GCM) have problems reproducing the variability of Sahel rainfall (Moron et al. 2004; Mohino et al. 2011b). On the other hand, seasonal predictions of parameters relative to the atmospheric dynamics are generally more skilful (Moron et al. 2004; Philippon et al. 2010; Ndiaye et al. 2011), because models perform better in capturing large-scale dynamics fluctuations than regional-scale rainfall anomalies (Garric et al. 2002).

The wind field coherence over West Africa can provide a physical basis for diagnosing Sahel rainfall variability (Fontaine and Janicot 1992). Indeed, Sahel droughts are strongly related with weaker than average upper level easterlies and lower level south-westerlies (e.g. Kidson 1977; Newell and Kidson 1984; Grist and Nicholson 2001; Fontaine et al. 1995; Nicholson 2013). The present study aims at assessing the skill of CMIP5 decadal prediction systems in forecasting Sahel rainfall at decadal time scales using dynamics-based indices.

The paper is structured as follows: in Sect. 2 data and methodology used are described, in Sect. 3 the decadal hindcasts and historical simulations are analysed and the predictive skills are assessed, the main results are discussed in Sect. 4 and finally summary and conclusions are drawn in Sect. 5.

2 Data and methods

In this study we assess the predictive skill of CMIP5 dynamics-based indices against observed precipitation and dynamics-based indices. For this purpose, observational data, reanalysis products, and model outputs are analysed. In order to compare datasets with different resolutions, all the precipitation and dynamics fields are interpolated to a T42 Gaussian grid (~2.8° in latitude and longitude), which is consistent with the resolution of the models (Table 1).

2.1 Climate models

We use monthly wind and precipitation data simulated by 14 atmosphere-ocean coupled GCM participating in the CMIP5 Project (Taylor et al. 2012; Table 1). Two sets of experiments have been analysed, namely, decadal prediction experiments (also referred as decadal hindcasts) and historical simulations.

The decadal hindcasts are 10-year long coupled simulations initialized every 5 years over the period 1960–2005 (i.e., the start dates of the simulations are the end of 1960, 1965, … 2005, and the hindcasts cover the decades 1961–1970, 1966–1975, … 2006–2015). These decadal prediction experiments respond to the time-varying atmospheric composition and are initialized with observed climate states, which include atmosphere, ocean and sea ice conditions. Therefore, decadal hindcasts provide information on both internal variability and externally-forced climate change (Meehl et al. 2009). Ocean and atmosphere initialization methods (e.g. full-field or anomaly initialization) are chosen at the discretion of each modelling group (Meehl et al. 2014).

The historical coupled simulations span the period 1850–2005. They are typically based on multicentury control integrations and respond to time-varying atmospheric composition (Taylor et al. 2012). We focus on the 1960–2005 period and use these simulations as a "non-initialized" baseline for the assessment of hindcasts' skill (Meehl et al. 2009). A brief description of the 14 models and simulations used in this study is presented in Table 1.

Table 1 Details on models used in the analysis: short name, modelling centre, spatial resolution, size of the ensemble for decadal and historical simulations, and initialization procedure

Model name	Modelling center	Resolution lon-lat	Decadal	Historical	Initialization
BCC-CSM-1	Beijing Climate Center, China Meteorological Administration	T42 (~2.8°)	4	3	Full field
CanCM4	Canadian Centre for Climate Modelling and Analysis	T42 (~2.8°)	10	10	Full field
CCSM4	National Center for Atmospheric Research	1.25° × 0.9°	10	6	Full field
CNRM-CM5	Centre National de Recherches Météorologiques/Centre Européen de Recherche et Formation Avancée en Calcul Scientifique	T127 (~1.4°)	10	10	Full field
EC-EARTH	EC-EARTH consortium	T159 (~1.125°)	5	4	Full field
FGOALS-g2	LASG, Institute of Atmospheric Physics, Chinese Academy of Sciences and CESS, Tsinghua University	2.8° × 3.0°	3	5	Full field
GFDL-CM2.1	NOAA Geophysical Fluid Dynamics Laboratory	2.5° × 2.0°	10	10	Full field
HadCM3	Met Office Hadley Centre (additional HadGEM2-ES realizations contributed by Instituto Nacional de Pesquisas Espaciais)	3.75° × 2.5°	10	10	Full field
IPSL-CM5A-LR	Institut Pierre-Simon Laplace	3.75° × 1.875°	6	6	Anomaly
MIROC4 h	Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology	T213 (~0.6°)	3	3	Anomaly
MIROC5		T127 (~1.4°)	6	5	Anomaly
MPI-ESM-LR	Max-Planck-Institut für Meteorologie (Max Planck Institute for Meteorology)	T63 (~1.9°)	10	3	Anomaly
MPI-ESM-MR		T63 (~1.9°)	3	3	Anomaly
MRI-CGCM3	Meteorological Research Institute	T159 (~1.125°)	9	3	Anomaly

2.2 Observations and reanalysis data

To assess the skill of Sahel rainfall decadal predictions, we use the observed monthly precipitation from the Climate Research Unit (CRU), version 3.10.01 (Harris et al. 2014). The CRU database is based on observations from meteorological stations, and provides monthly mean data at $0.5° \times 0.5°$ horizontal resolution for global land areas (excluding Antarctica). It covers the period 1901–2009 and has been shown to compare well with precipitation measured from meteorological stations over the Sahel (Fink et al. 2010). In addition, monthly mean precipitation extracted from the Global Precipitation Climatology Project dataset (GPCP; Adler et al. 2003), which is available on a $2.5° \times 2.5°$ horizontal grid from 1979 to present, is used to extend the comparison with decadal predictions up to 2014.

Dynamics-based indices have also been computed by using wind data from NCEP-NCAR and ERA-40 reanalyses (Kalnay et al. 1996; Uppala et al. 2005). The former dataset is available from January 1948 until present with a spectral horizontal resolution of T62, resulting in a $2.5° \times 2.5°$ grid with 17 levels in the vertical. The latter is the second generation reanalysis of the European Centre for Medium Range Weather Forecasting and covers the period from September 1957 to August 2002, with a $2.5° \times 2.5°$ spatial resolution and 23 vertical levels available.

2.3 West African Monsoon Index

Strong links between WAM dynamics and rainfall anomalies in the Sahel have been long recognized (Kidson 1977;

Newell and Kidson 1984). Fontaine and Janicot (1992) showed the existence of regional meridional (Hadley-type) and zonal (Walker-type) cells on a monthly time scale, which is a significant feature of the atmospheric circulation over West Africa. Based on the coherent relationship found between lower and higher troposphere variability, Fontaine et al. (1995) defined the West African monsoon index (WAMI) as the standardized time series of the difference between standardized anomalies of wind modulus at 925 hPa and zonal wind at 200 hPa at the location 7.5°N, 0°E. The index is intended to capture the most prominent seasonal signals of the monsoonal circulation over West Africa: the low level monsoonal flow and the tropical easterly jet (TEJ). A high (low) WAMI value indicates an enhanced (reduced) monsoonal circulation as well as a stronger (weaker) TEJ (Fontaine et al. 1995). The WAMI has been also defined in different areas over the Sahel (Fontaine et al. 2011; Garric et al. 2002). Such index is considered as a proxy of the monsoon cell intensity at regional scale (Fontaine et al. 2011) and it shows a strong positive correlation with Sahel rainfall at interannual and decadal time scales (Fontaine et al. 1995).

In this study we follow a definition for the WAMI similar to the one proposed by Fontaine et al. (2011). We first calculate the standardized anomalies of July to September (JAS) seasonal means of wind modulus at 925 hPa and zonal wind at 200 hPa averaged over certain regions. We then calculate WAMI as the difference between such standardized anomalies. Finally, WAMI index is also standardized in time. Due to the time coverage of model simulations and observational data, we focus our study on the

Table 2 Domains selected for the computation of the West African monsoon index, and correlations coefficients between WAMI and the CEOF1 PCs time series

Model	MOD925		UA200		PC-WAMI
	Latitude	Longitude	Latitude	Longitude	
BCC-CSM1-1	7–14N	15W–20E	4–10N	5W–20E	0.92
CanCM4	2–12N	20W–20E	5–15N	10W–20E	0.99
CCSM4	5–15N	20W–10E	10–15N	0–20E	0.99
CNRM-CM5	6–14N	10–25E	3–13N	8W–15E	0.94
EC-EARTH	5–15N	10W–20E	5–16N	10W–10E	0.98
FGOALS-g2	2–10N	10W–15E	3–15N	0–20E	0.97
GFDL-CM2p1	5–10N	10W–20E	0–20N	10W–20E	0.99
HadCM3	2–12N	20W–12E	5–12N	10W–15E	0.93
IPSL-CM5A	0–7N	15W–5E	4–14N	8–22E	0.96
MIROC4 h	2–13N	20W–5E	5–15N	6W–18E	0.94
MIROC5	3–12N	20W–10E	4–14N	5E–18E	0.97
MPI-ESM-LR	4–12N	20W–15 W	5–12N	7W–10E	0.96
MPI-ESM-MR	2–12N	20W–15E	5–15N	10W–15E	0.95
MRI-CGCM3	2–8N	15W–5E	10–20N	0–20E	0.88
MME	5–10N	20W–5E	0–10N	15W–10E	0.98
ERA40	2–10N	10W–5E	3–16N	10W–15E	0.92
NCEP	4–12N	15W–10 N	2–12N	15W–20E	0.98

1961–2009 period. Nevertheless, the particular regions used to calculate the JAS averages differ among models and with observations. Such regions are carefully chosen, for each dataset and model, on the basis of the particular dominant dynamics features of the monsoonal circulation cell, which is, in turn, characterized using a combined empirical orthogonal functions (CEOF) analysis.

The CEOF method is one of the simplest variants of the standard empirical orthogonal functions (EOF) analysis (Venegas 2001) and it is used to explore the common modes of variability of two or more variables. Specifically, two or more fields are concatenated one after another along the spatial dimension in the covariance matrix, and the eigenvalues problem is solved as in a standard EOF analysis. For each mode of variability, the CEOF method yields an anomaly pattern for each variable, and a principal component (PC) time series common to all the variables used in the analysis. All fields are previously normalized to avoid the dominance of one field over another (Venegas 2001). In our study we apply the CEOF analysis to the standardized JAS seasonal means of the zonal wind at 200 hPa and the wind modulus at 925 hPa to identify their common modes of variability. For models, we use all ensemble members of the historical simulation concatenated in time to compute the CEOF analysis. Based on the wind patterns of the dominant mode of covariance, we determine the domains used for computing the WAMI (Table 2).

Since our main interest is to define a dynamic proxy that stands for the monsoonal circulation in terms of both vertical and horizontal coherence, a sensitivity test has been performed to evaluate the robustness of the CEOF with regard to the domain size. Indeed, the variability of the lower troposphere flow over West Africa may be also influenced by the northerlies from the Mediterranean basin, which have been shown to modulate the monsoonal precipitation over Sahel (Gaetani et al. 2010; Gaetani and Fontaine 2013). Therefore, we checked whether the CEOF patterns are affected and the monsoonal cell coherence is preserved when modifying the domain limits, in order to recognize possible mixtures of circulation regimes with nearby regions. Thus, the CEOF are examined in two different areas, a first one that extends northwards to 40°N, and a second smaller to 20°N.

2.4 Detection of forced vs internal variability

To highlight the predictive skill that could come from the initialization, the forced component has been isolated by means of a signal-to-noise EOF analysis. Such method is designed to distinguish between externally forced climate responses and natural internal variability (Solomon et al. 2011) and can be applied when there is a weak forced signal and strong internal variability (Davies et al. 1997;

Venzke et al. 1999; Chang et al. 2000; Ting et al. 2009). A detailed description of the method is presented in Venzke et al. (1999). In this study the forced signal is detected for each model in the historical simulation and the forcing patterns in wind field and SST are subtracted from the decadal hindcasts to define the residual fields. The predictive skill obtained in such residual fields is attributed to the initialization (Mochizuki et al. 2010; Gaetani and Mohino 2013).

2.5 Drift correction

When models are initialized with an observed climate state, errors appear as a consequence of the models' drift toward their own attractor (Kirtman et al. 2013). Such errors depend on the forecast time. To correct them, we follow the method recommended by the World Climate Research Programme (IPCO 2011): we calculate anomalies ($Y'_{j\tau}$) by subtracting the climatology at each forecast year:

$$Y'_{j\tau} = Y_{j\tau} - \bar{Y}_\tau,$$

where j is the start date, τ is the forecast year, $Y_{j\tau}$ the seasonal average at forecast year τ from simulation started at year j, and $\bar{Y}_\tau$ the climatology for forecast year τ. In the decadal simulations analysed in this work, the initialization year j spans the period 1960–2005 with an increment of 5 years (i.e., j = 1960, 1965,…, 2005) and, as the simulations are 10-year long: $\tau = 1, 2,…,10$. García Serrano et al. (2015) suggest that such 5-year interval between start dates is enough to yield the main features over the Atlantic at decadal time scales. For each forecast year τ, the climatology $\bar{Y}_\tau$ is calculated using the years $1960 + \tau$ in the first simulation, $1965 + \tau$ in the second one,… and $2005 + \tau$ in the last one. For direct comparison, historical simulations are treated in the same way as the decadal ones, except that the former are analysed for the 1961–2005 period and the latter in the 1961–2009 period. The procedure is similar when the reanalysis datasets are employed, using the period 1961–2009 for the NCEP and 1961–2002 for the ERA40. In this study, anomalies are also standardized at the forecast time.

2.6 Predictive skill metrics

The predictive hindcast skill is evaluated through the anomaly correlation coefficient (ACC), widely used to quantify the correlation between forecast and observed anomalies (Miyakoda et al. 1972), and the root-mean-square error (RMSE), which measures the distance between the models output and observations. Before computing these verification measures, a 4-year average is applied to focus on decadal time scales. A 4-year average has been shown to be a compromise between the capability of partially removing

the unpredictable interannual variability in near-term dynamical forecasting and the ability to show some of the skill evolution along the forecast time (García-Serrano and Doblas-Reyes 2012). This is a common approach widely used in decadal prediction (van Oldenborgh et al. 2012; Kim and Webster 2012; Gaetani and Mohino 2013). Given the small size of the samples, a Monte Carlo test with 200 permutations is applied to estimate the statistical significance of the correlations. Hence, the role of the initialization process is assessed by comparing the ACC and RMSE calculated for the initialized hindcast and the no initialized experiments. The robustness of the estimation of the added-value coming from the initialization alone is then assessed by applying the skill measures to the decadal residuals, for which the forced component is removed (Mochizuki et al. 2010; Gaetani and Mohino 2013).

We also use the 4-year averaged WAMIs from the 14 models jointly to test the performance of a probabilistic forecast. We base such forecast on three mutually exclusive events with equal probability, which are "above normal", "near normal" and "below normal" conditions. We apply an unadjusted Gaussian probabilistic approach to estimate the probabilities of each category and follow Kharin and Zwiers (2003a) to estimate them as:

$$\hat{P}(A) = F_G\left(\frac{\hat{\beta} - x_a}{\hat{\sigma}}\right)$$

$$\hat{P}(B) = F_G\left(\frac{\hat{\beta} - x_a}{\hat{\sigma}}\right)$$

$$\hat{P}(N) = 1 - \hat{P}(A) - \hat{P}(B)$$

where $\hat{P}(A)$ is the estimated probability of having above normal conditions, $\hat{P}(B)$ of below normal and $\hat{P}(N)$ the probability of normal conditions; $F_G()$ is the Gaussian cumulative distribution for zero mean and a standard deviation of one; $\hat{\beta}$ is the estimation of the potentially predictable signal; $\hat{\sigma}$ is the estimation of the standard deviation of the non-predictable signal; and x_a is the lower boundary for the above normal category (which is equal to the minus lower boundary as we are working with anomalies). The $\hat{\beta}$ and $\hat{\sigma}$ parameters are estimated as the average of all ensemble predictions and the square root of the mean over all hindcasts of the standard deviation across all ensemble members, respectively. For parameter x_a, we use

$$x_a = \hat{\sigma}_X F_G^{-1}\left(2/3\right)$$

where $F_G^{-1}\left(2/3\right)$ is the inverse of the cumulative Gaussian distribution for probability 2/3, which is approximately 0.43; and $\hat{\sigma}_X$ is the estimated standard deviation of the observation taking all events in the hindcast period.

To test the skill of such probabilistic scheme, we use the Relative Operating Characteristic (ROC) curves and ROC scores. The ROC curve is a plot of the hit rate (HR, the fraction of the events for which an alarm was rightly issued) versus the false-alarm rate (FAR, the fraction of non-events for which an alarm was wrongly issued), varying the probability threshold to launch the alarm. The ROC score is twice the area under the ROC curve minus one. Perfect forecast show a ROC score of 1, while a ROC score of 0 means no skill. Interested readers are referred to Kharin and Zwiers (2003b) for further details.

3 Results

3.1 CEOF analysis and WAMI definition

Figures 1 and 2 show the patterns of wind modulus at 925 hPa and zonal wind at 200 hPa associated to the first CEOF mode (CEOF1) calculated in the 20°W–30°E, 0–20°N domain for the historical simulations and for NCEP and ERA40 reanalysis. As it was mentioned in the previous section, in order to avoid possible mixtures of regimes, the CEOF analysis was also applied in a larger area (30°W–50°E, 10ºS–40°N) (not shown). The correlations between the PCs obtained for both regions are high (all above 0.83) and statistically significant (95 % level of confidence), which suggests that the covariance analysis is not strongly affected by the selected area. We, therefore, restrict our subsequent analysis to the results obtained with the CEOF method applied in the smaller area.

When CEOF is applied to reanalysis, the resulting patterns for the leading mode account for 57 % of the variance in NCEP and 39 % in ERA40. The CEOF1 of the wind modulus at 925 hPa for NCEP data shows a pattern of anomalies over West Africa, with a maximum that extends approximately between 15°W and 10°E and 5–10°N (Fig. 1). The CEOF1 pattern for the ERA40 shows two maxima, though in this case they are significantly less pronounced. One of the maxima is located north of 10°N on the northern side of the Atlantic Intertropical Convergence Zone, while the second one is near to the Guinean coast (10°W–5°E, 2°–10°N). Differences between NCEP and ERA40 CEOF1 patterns are also reflected in the zonal wind at 200 hPa (Fig. 2). The strongest anomalies are shown by the NCEP, data which expand across the West African domain, with the highest amplitude around 6°N along 0–5°E, while maximum anomalies in ERA40 show a larger latitudinal extension from 2°N to 15°N and are placed eastwards with respect to the NCEP ones.

The CEOF1 patterns explain above 40 % of the variance in most models, which suggests a high degree of covariance between the high and low tropospheric levels. BCC-CSM1

Fig. 1 Correlation of JAS wind modulus at 925 hPa with the PC time series associated to CEOF1 computed in the 20°W–30°E, 0–20°N domain. For the models, the CEOF analysis was applied to the historical simulations in the 1960–2005 period. *Black boxes* show the regions used to compute the regional averages of wind modulus at 925 hPa used in the definition of WAMI (see details in the text). Percentages of explained variability by the CEOF mode are shown in the *top right corner of each plot*. *Grey shades* correspond to grid-cells with missing values
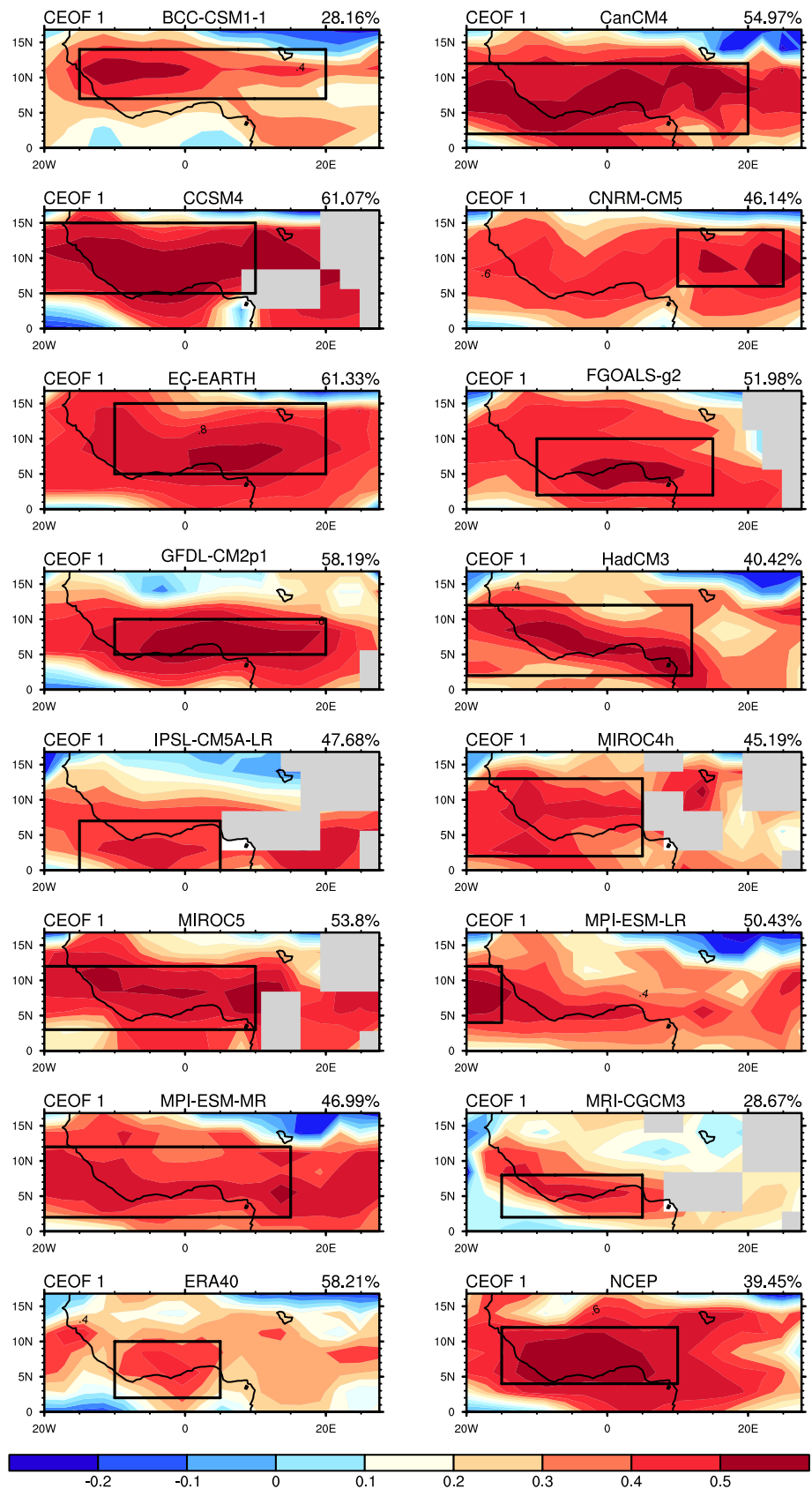
Fig. 2 Same as Fig. 1, but for zonal wind at 200 hPa
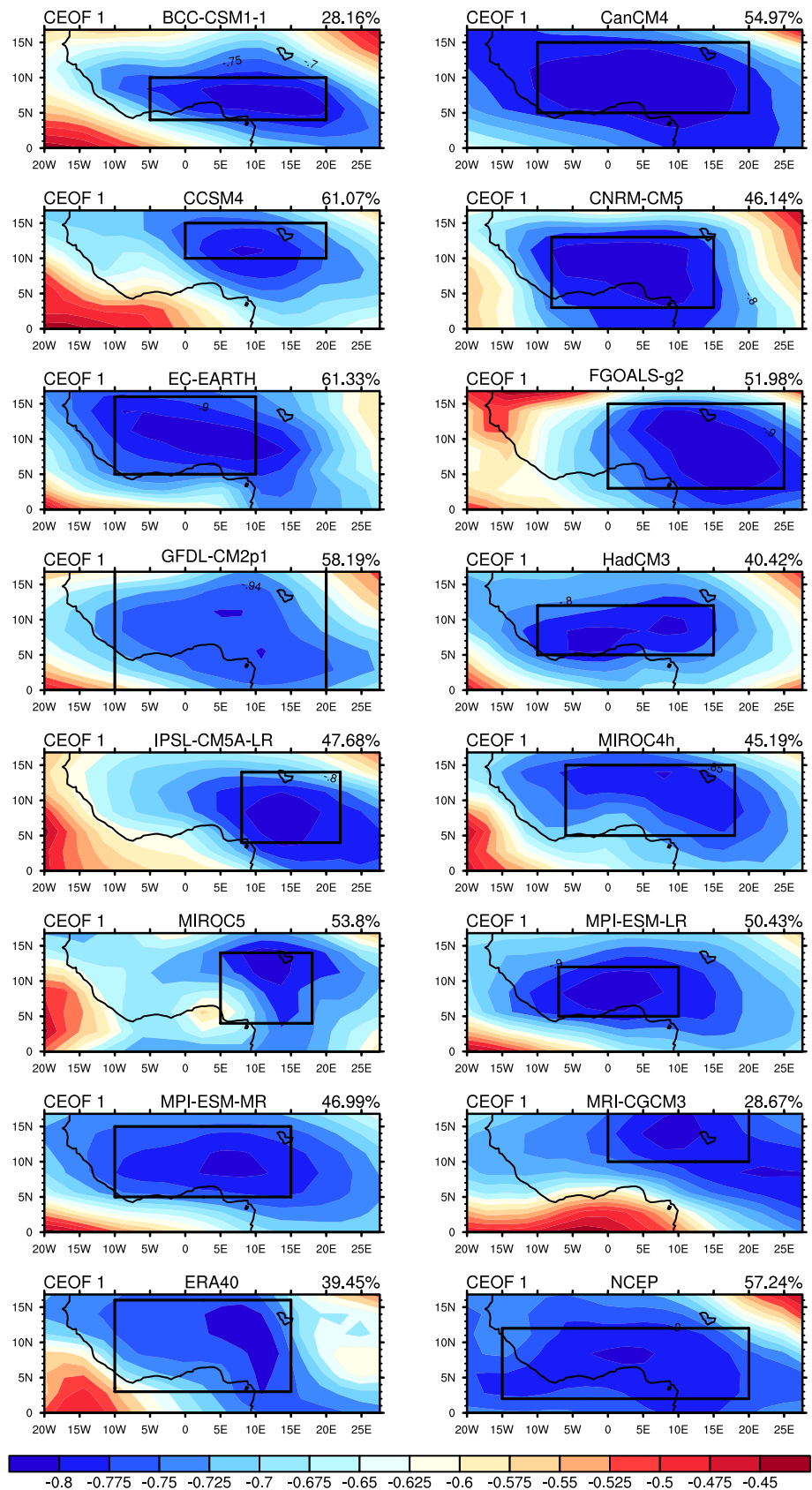
and MRI-CGCM3 models are the exception, with just below 30 % of explained variance. Models generally capture the main features of CEOF1 pattern of wind modulus at 925 hPa shown by reanalysis, especially regarding the latitudinal position of the maximum correlations, which are situated within the band 5–15°N, with only IPSL-CM5A-LR showing them south of 5° (Fig. 1). However, there are substantial differences on the longitudinal position of the maximum correlations, ranging from 20°W to 20°E. Models such as CNRM-CM5, MPI-ESM-LR or MPI-ESM-MR, exhibit the maximum correlation areas far from the observed monsoonal cell, which is displaced along 0–20°N and 10°W–10°E (Fontaine and Janicot 1992). Differences in the CEOF1 patterns of the zonal wind at 200 hPa are also evident, with some models showing wider areas of maximum correlation (CanCM4, CNRM-CM5, FGOALS-g2), and others characterized by a narrower peak (e.g. CCSM4, MIROC5, see Fig. 2). Moreover, most of the models show better agreement with the NCEP pattern, with strong correlations along the Guinean coast (BCC-CSM1-1, CanCM4, CNRM-CM5, EC-EARTH, GFDL-CM2p1, HadCM3, MIROC4 h, MPI-ESM-LR, MPI-ESM-MR), while the others show an eastward maximum more similar to the one obtained with ERA40 data.

To evaluate the covariance of the precipitation and wind fields, the CEOF method was also applied including the precipitation fields (not shown). The correlation coefficients between the PC time series from the CEOF analysis applied to wind and precipitation fields and the one applied only to wind fields are above 0.97, which suggests that the wind patterns do not suffer significant changes when the precipitation is added. This analysis supports the WAMI definition without prior precipitation consideration (Fontaine et al. 1995).

Our CEOF analysis suggests that there are regional differences in the monsoon dynamics simulated by the different models. To allow for these differences, the areas where the wind components of WAMI are averaged are defined separately for each model. We carefully select such areas to capture the high loads in the CEOF1 patterns (Figs. 1, 2; Table 2). The same procedure is also applied to the reanalysis data.

In order to verify that this procedure captures the main mode of covariance between high and low tropospheric levels, we correlate the WAMI obtained for the historical experiments with the PC associated with CEOF1 (Table 2). In general, strong correlations are obtained (Table 2), which suggest that the calculated WAMI is able to capture the co-variability of wind fields in the upper and lower troposphere and hence, it is able to diagnose the West African monsoon cell circulation (Fontaine et al. 2011).

To verify whether the WAMI-precipitation relationship is well simulated, we calculate the correlations between the WAMI and the precipitation fields for each individual model in the historical experiments, and between WAMI in reanalysis and CRUTS3.1 rainfall field (Fig. 3). The ERA40 and NCEP maps show significant correlations over the Sahel region with maximum values at 12°–15°N, 14 W–14E. Most of the models show robust WAMI-precipitation correlation patterns, but with substantial differences. This relationship is robust over the Sahel for several models (i.e. CCSM4, EC-EARTH, GFDL-CM2p1, MIROC5, MPI-ESM-LR, MPI-ESM-MR), whereas other models exhibit significant correlations placed eastward (CanCM4, FGOALS-g2, MIRCO4 h) or southward (HadCM3, IPSL-ESM-LR, MRI-CGCM3). The behaviour of CNRM-CM5 is less clear, with significant correlations over the Guinea coast and also over Sahara. Correlations are weaker and less spatially coherent only in the case of BCC-CSM1.

3.2 Skill of modelled WAMI vs observed precipitation

A Standardized Precipitation Index (SPI) is defined over the 10°–20°N, 15°W–15°E domain, and the predictive skill metrics (ACC and RMSE) are computed between the observed SPI and the simulated WAMI. For both SPI and WAMI, the standardization of anomalies is calculated at the forecast years. The standardization of precipitation and WAMI allows the computation of RMSE between dimensionally different variables (Barnston 1992). Both skill scores have been applied to the initialized hindcast, historical experiment and resulting decadal-forcing residuals (Fig. 4).

Decadal hindcasts and historical simulations show a similar behaviour, with a reduction of RMSE values at long lead times (5–10 years). In general, models perform better in decadal hindcasts than in historical simulations, whereas decadal-forcing residuals show a larger spread in RMSE, with some models further improving their skill (e.g. MIROC5 and CanCM4) and some others downgrading their performance (e.g. MPI-ESM-MR). The RMSE for the residuals of MIROC5 decreases for short lead-times (1–5) compared to the scores in the decadal hindcasts, showing more skill when the forced component is removed. Unlike other models, as the CanCM4 or HadCM3, which lose their skill when residuals are considered.

The ACC analysis shows that there is forecast skill for some models in the decadal hindcasts. The CanCM4 is one of the most skilful models and shows significant correlations for both short (1–5 years) and long lead time (5–10 years). HadCM3 is skilful at the 2–6 years and 5–8 years lead times, MIROC5 is skilful at short lead times (1–5 years) and BCC-CSM1-1 at long lead times (6–9 years). These three models, unlike CanCM4, did not show any skill when ACC was applied to observed and simulated precipitation. The forecast skill at long lead

Fig. 3 Correlations between WAMI, computed according to the CEOF analysis, and precipitation field in the period 1961–2005, for the historical experiments, and for reanalysis and CRU data. *Dots* show the areas with significance at 95 % defined using a two-tailed Student's *t* test. *Grey shades* correspond to grid-cells with missing values
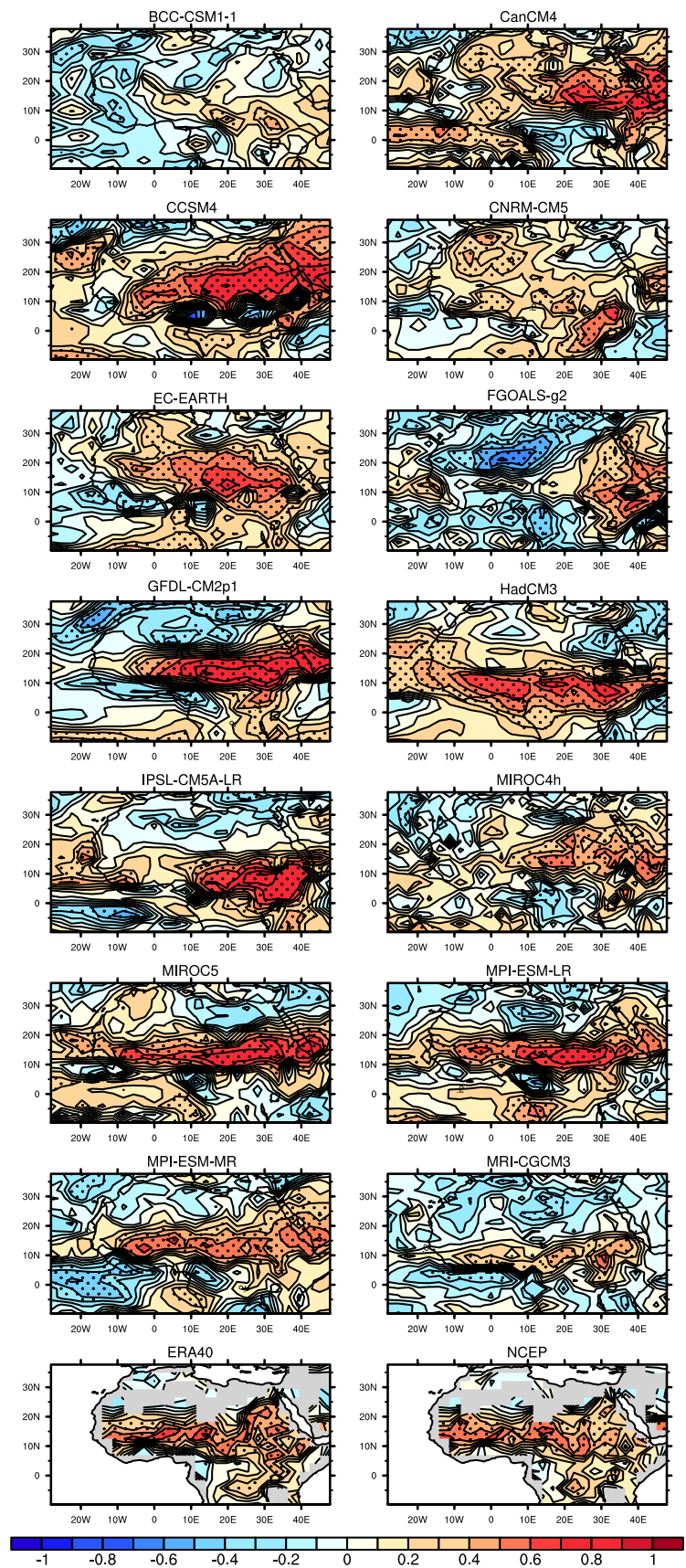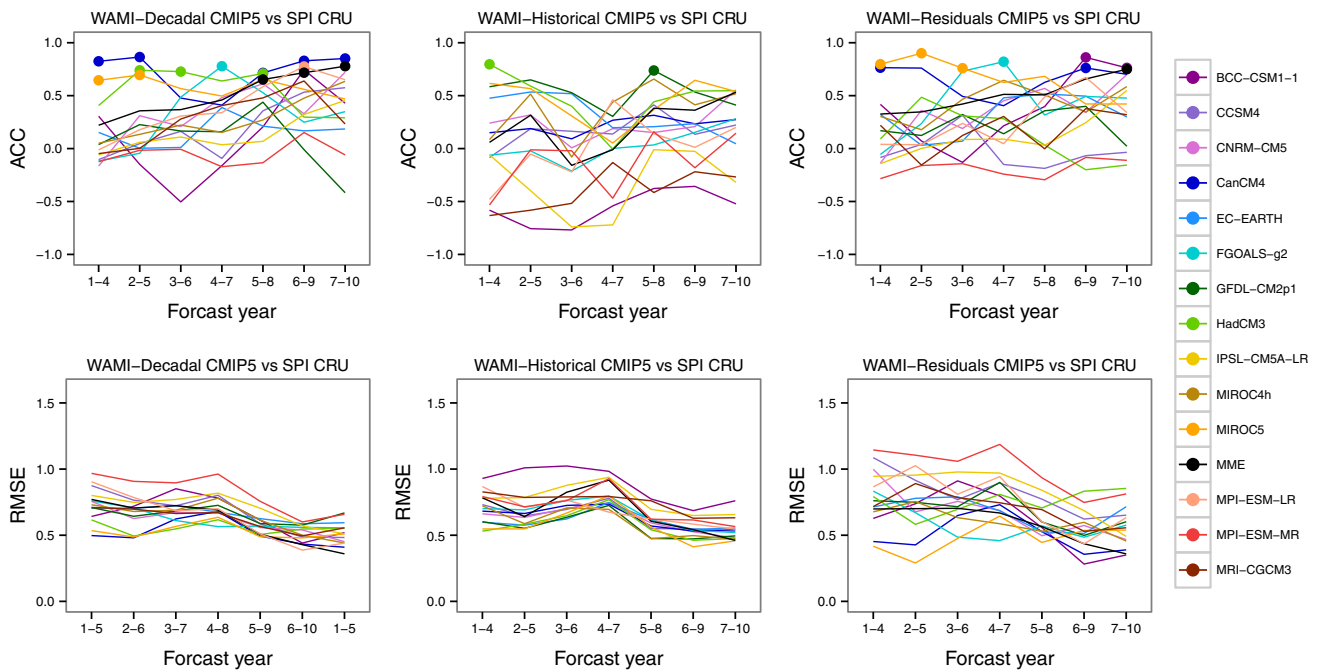
Fig. 4 ACC (*top*) and RMSE (*bottom*) scores for modelled WAMI in decadal hindcasts (*left*), historical experiment (*middle*) and decadal-forcing residuals (*right*), tested against the Sahel precipitation index (SPI) from CRU data. *Dots* indicate significant ACC at 95 % level of confidence

times (6–9 years) obtained for rainfall in the MPI-ESM-LR is also obtained when using WAMI (Fig. 4). Multi-model ensemble (MME) indices are defined by averaging the indices computed for each model, and they show significant correlations at long lead times (5–10 years). Similar results are found when analysing MMEs independently computed from those models with "full-field" and "anomaly" initialization (see Table 1), although it is observed a slight better performance of the "anomaly" ensemble (not shown). When assessing the skill in the historical experiments, ACC values show a large spread among the models and only a couple of significant correlations. Namely, only HadCM3 and GFDL-g2 are skilful at the 1–4 and 5–8 years lead time, respectively. When the ACC is computed for the decadal-forcing residuals the results for the CanCM4, HadCM4 and the MME suggest that in these cases the forced component is also relevant to obtain skill for decadal predictions. Nevertheless, the forced component is decreasing the forecast skill in the BCC-CSM1-1 model and the MIROC5, which show significant ACC values in the decadal-forcing residuals at long (6–10 years) and short (1–5 years) lead times, respectively. The results for the FGOALS-g2 in the residuals remain as in the decadal hindcasts. Similar results were found when using longer windows filtering, specifically 5 years, although with a slight reduction in the skill (not shown). Better performances in the decadal simulations and partially for the decadal-forcing residuals, compared to the skills in the

historical simulations, suggest the contribution of the initialization to the prediction skills.

We also tested a probabilistic forecast for rainfall. Given the reduced number of members for some models (Table 1), we decided to test only the probabilistic hindcasts issued from the 14 models jointly, using the 4-year averaged WAMI from each. In addition, we only have between 8 to 10 hindcasts, depending on the experiment and lead time, which is a
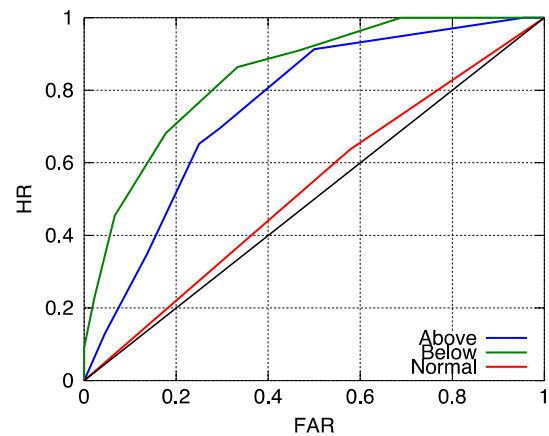

Fig. 5 ROC curves (hit rate -HR- as function of the false alarm rate—FAR) for the above normal (A), below normal (B) and normal (N) categories estimated for the 6–9 lead time using an ensemble of 14 models WAMI obtained in the decadal experiments. The probabilistic forecasts are tested against Sahel rainfall index from CRU

Table 3 ROC scores for the three tercile categories (A: above normal, B: below normal, N: normal) for the probabilistic hindcasts obtained from the decadal, historical and decadal residual experiments

	Decadal	Hitorical	Decadal_residuals
Above (A)	0.51	0.14	0.45
Below (B)	0.69	0.33	0.58
Normal (N)	0.06	0.07	−0.04

The ROC scores were computed using the 14 models (one WAMI per model) and tested against Standard Precipitation Index (SPI) defined over the Sahel (10°–20°N, 15°W–15°E)

very small sample to test the probability forecast approach. For this reason, we use all seven lead times at once for each experiment, which gives us an average estimation of the skill of such probabilistic predictions. This approach allows us to assess the prediction capability of the forecast system as a whole, focusing on the differences between initialized and non-initialized experiments. Figure 5 shows the ROC curve for the decadal simulations taking into account all lead times. The ROC score for the curves in Fig. 5 are shown in Table 3 for the "above", "below" and "normal" categories using the WAMIs estimated from the decadal hindcasts. We also add for comparison the scores obtained for the historical and decadal residual simulations.

The below normal category is more skilfully forecasted than the above normal one. Our probabilistic forecasts

show no skill for the normal category, which is common to other seasonal probabilistic forecasts (Kharin and Zwiers 2003b). Regarding the different simulations, for those categories in which we might expect to have skill (above and below normal conditions), the decadal experiments outperform the historical experiments (Table 3). The decadal residual hindcasts retain most of the skill for those two categories, which suggests that the skill of probabilistic decadal simulations comes entirely from the initialization of the simulations. Such result is in agreement with our deterministic analysis using ACC scores.

3.3 Skill of modelled WAMI vs reanalysis WAMI

The WAMI, computed in both NCEP and ERA40 reanalysis, is also employed for investigating the capability of models in predicting the monsoonal circulation over West Africa. As in the previous analysis, ACC and RSME are used to assess prediction skills.

Models' RMSE scores tested against NCEP WAMI show a similar behaviour to the one obtained when comparing with observed rainfall (Fig. 6), with a decrease at long lead times. However, RMSE scores are considerable higher for the three experiments, decadal, historical and especially when the forced component is removed. When evaluating ACC, there is a larger spread of score values in comparison to those computed against observed rainfall (Fig. 6). For the decadal hindcasts, CNRM-CM5 and CanCM4 are
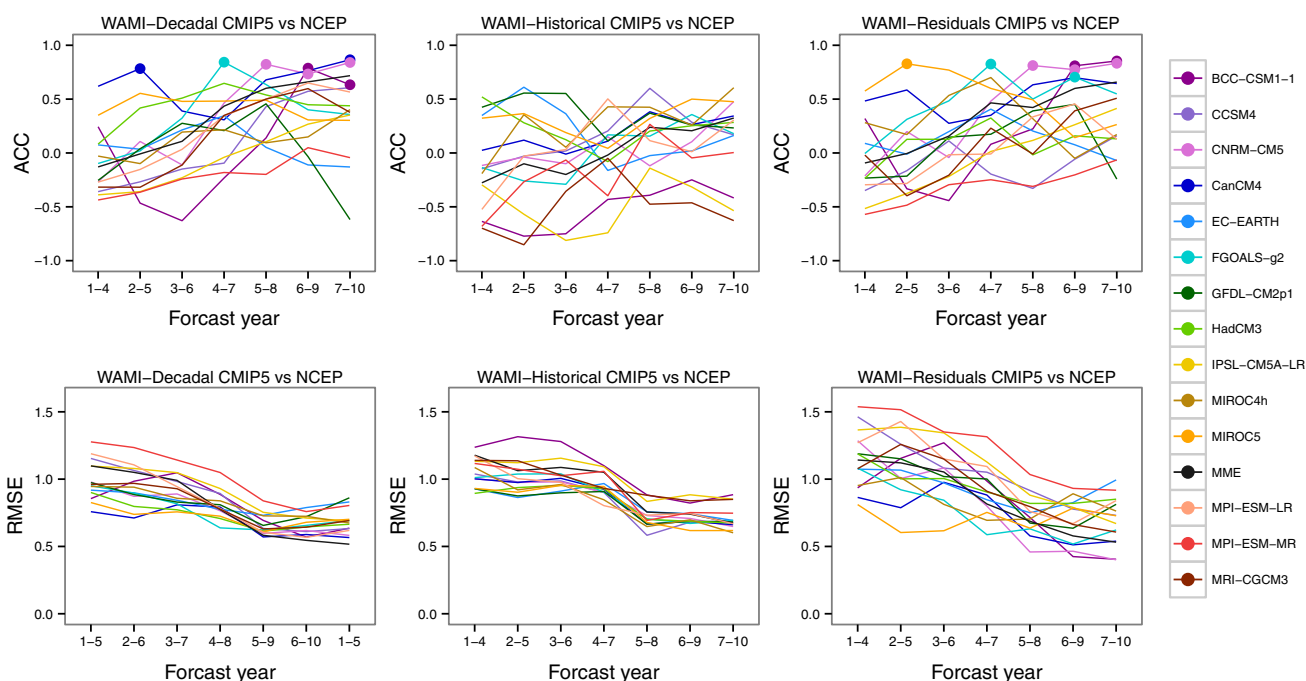

Fig. 6 ACC (*top*) and RMSE (*bottom*) scores for modelled WAMI in decadal hindcasts (*left*), historical experiment (*middle*) and decadal-forcing residuals (*right*), tested against the WAMI from NCEP reanalysis. *Dots* indicate significant ACC at 95 % level of confidence
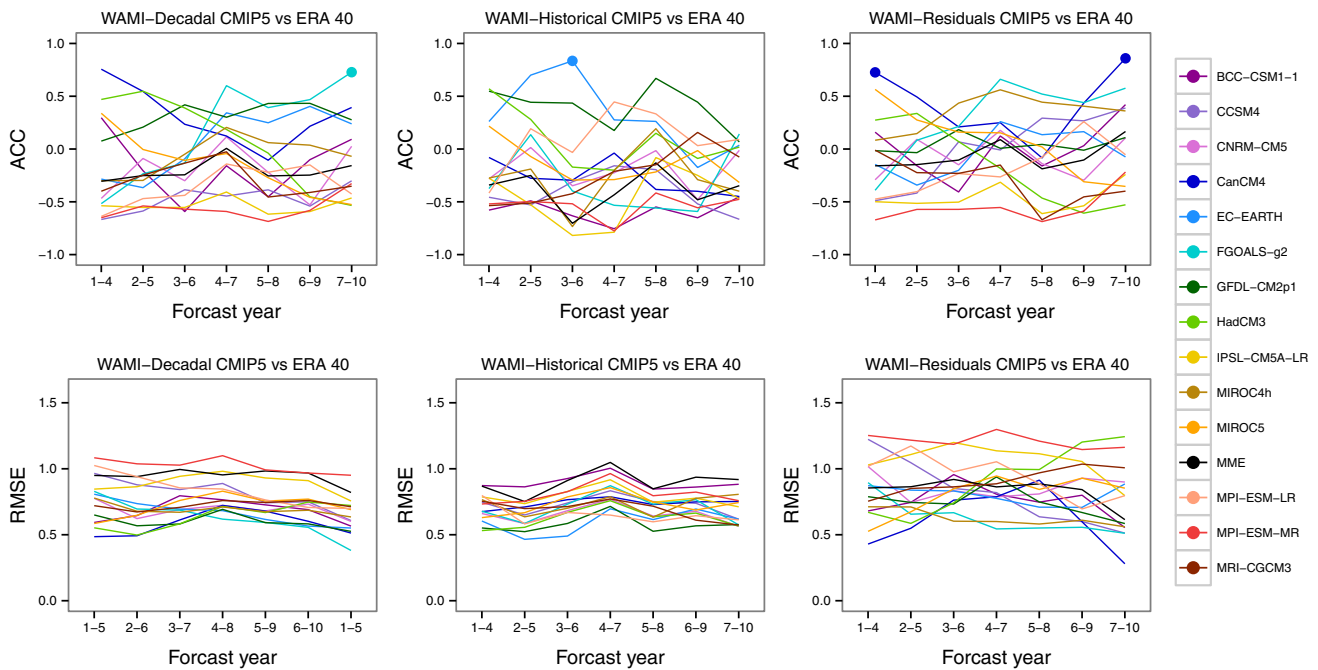
Fig. 7 Same as Fig. 5 but for ERA40 reanalysis

the most skilful models with significant correlations at long lead times (5–10 years). Similar skill was found for this last model in the precipitation analysis (Figure S2). The CanCM4 model shows also significant correlations at the 2–5 years lead time, while BCC-CSM1-1 is skilful for long lead times (6–10 years). The FGOALS-g2 model shows the same skill when tested against reanalysis WAMI and observed SPI (Figs. 4, 6). When the correlations are computed by using the historical experiments, a large spread of the values is again observed and significant skills are completely lost. The results for the decadal-forcing residuals look similar to those obtained for the decadal hindcasts, though the CanCM4 model performance is worse when the forced component is removed. BCC-CSM1-1 and CNRM-CM5 retain the skill, while FGOALS-g2 improves its ability at the 6–9 years lead time, as the MIROC5 for the 2–5 years.

There are considerable differences when verification measures are applied to WAMI computed for ERA40 reanalysis (Fig. 7). Indeed, skills are now drastically reduced for each model in all the experiments. The RMSE is in the 0.4–1.2 range and, unlike the previous analysis where the RMSE values decreased for longer lead times, the RMSE values do not show improvement at long lead times. Overall, the ACC shows a wide spread among the models, even for the decadal hindcasts, with only one significant correlation: FGOALS-g2 for the decadal and GFDL-CM2p1 for the historical experiments. The CanCM4 model shows significant skill at the first and the last lead time for the

decadal-forcing residuals. To test whether the disagreement could come from the different boxes used in the definition of WAMI in both reanalysis, we re-calculated WAMI for ERA40 using the same regions as for the NCEP one (Table 2). However, the models still showed no significant skill when tested against this new ERA40 WAMI (not shown).

Summarizing, the different performances of the decadal and historical experiments are evident when comparing the metrics between the simulated WAMI index and the observed SPI or the observed WAMI from NCEP reanalysis, pointing out both the robustness of WAMI as predictor of the decadal variability of precipitation, and the added value coming from initialization. Nevertheless, there is no agreement when analysing the models skill in predicting WAMI from ERA40 reanalysis. In that case, models show a wider spread in the verification measures and, in general, a poorer performance, irrespective of whether the initialization is included or not in the simulations. The distinctive behaviour shown by ERA40 reanalysis data has been further investigated and discussed in Section S.4, and a possible cause is found in the low-frequency component of the wind modulus at 925 hPa.

4 Discussion

Our results indicate that dynamics-based indices can be used to forecast Sahel precipitation at decadal time scales
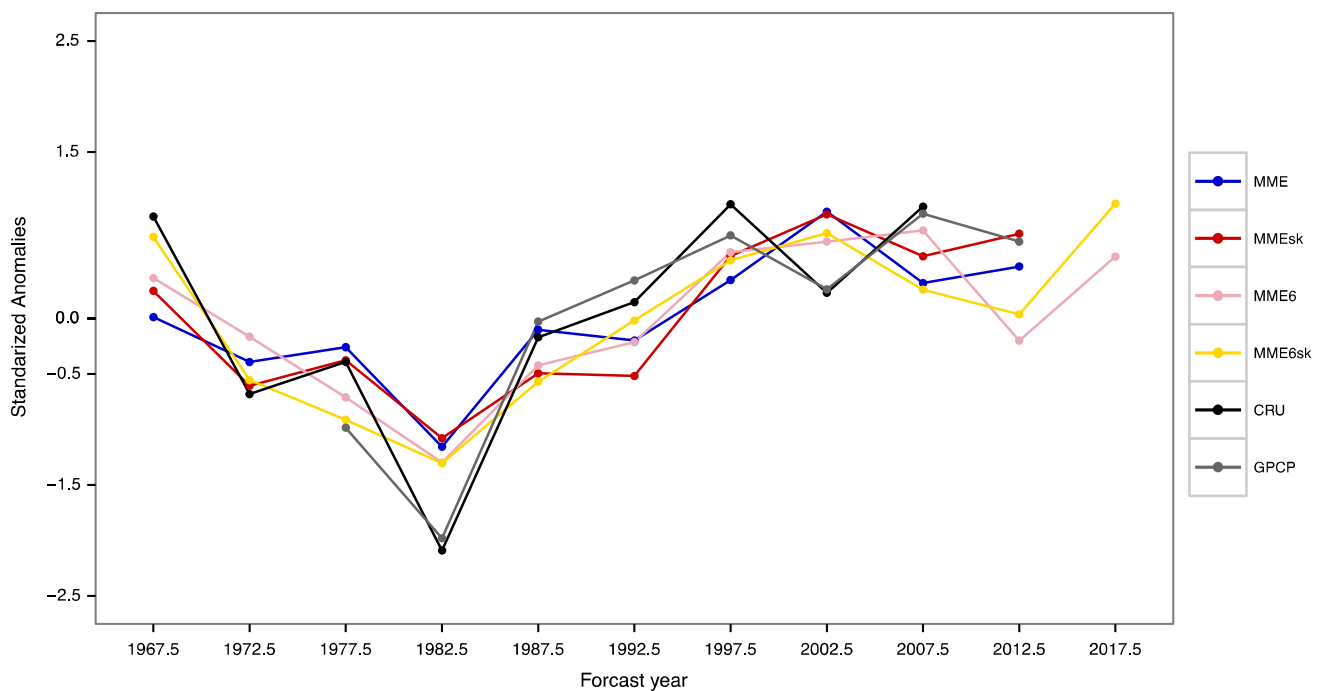
Fig. 8 Standardised anomalies of the observed SPI from CRU and GPCP datasets, and multi-model ensemble mean WAMI computed for all (MME) and skilful (MMEsk) models available up to the 2006–2015 decade, and for all (MME6) and skilful (MME6sk) models available up to the 2011–2020 decade

due to the strong linkage between the WAM dynamical fields and the precipitation over the Sahel (Fontaine et al. 1995). Indeed, the WAMI used in this work, which represents the main features of the WAM circulation (Fontaine et al. 1995), shows skill in predicting decadal variability of Sahelian precipitation for some models. Such result suggests that we could use the decadal simulations to give an outlook into the evolution of Sahel rainfall in the near future (decade 2011–2020).

In Fig. 8 we show the multidecadal variability of observed precipitation along with the WAMI simulated in the decadal hindcasts. Standardised anomalies of SPI are computed for the CRU dataset until 2009 and GPCP dataset up to 2014. In this plot, we show the results for the 6–9 years lead-time, since the skill of prediction for the models has been detected mostly at long lead times (Fig. 4). Unfortunately, at the time of analysis, not all models were providing the decadal simulation initialized in 2010 (i.e., forecasts for the 2011–2020 decade). Therefore, a subset of only 6 models is considered for the future outlook (CanCM4, GFDL-CM2p1, MIROC5, MPI-ESM-LR, MPI-ESM-MR, MRI-CGCM3). Standardized anomalies of the indices are computed averaging the WAMI for several multi-model ensembles: a multimodel mean for all models from the decadal hindcasts (MME); a mean for the most skilful models in our analysis (BCC-CSM1-C1, CanCM4, HadCM3, FGOALS-g2, MIROC5, MPI-ESM-LR)

(MMEsk); a mean for the models simulating the 2011–2020 decade (MME6); and the last one computed using only the skilful models within the MME6 subset (CanCM4, MIROC5, MPI-ESM-LR) (MME6sk).

The observed precipitation for both CRU and GPCP datasets shows decadal variability with a remarkable dry period starting in the middle of the 60 s until early 80 s (Fig. 8). The rainfall recovery is appreciated in the following years with a partial reduction at the beginning of the twentieth Century and an increase in the recent years. In general, the decadal variability shown by the observed precipitation is partially captured by the dynamic indices for the four multimodel ensembles represented in Fig. 8. In spite of their underestimation of rainfall in the 60 s, they are able to reproduce the decrease of the precipitation during the 70 s and subsequent recovery in the 80 s. Discrepancies with the observed rainfall are more evident for the MME and MMEsk (which only comprise decadal runs initialized from 1960 to 2005), especially in the 2000 decade. Nevertheless, the multimodels that extend one decade further (MME6 and MM6sk) are more consistent with the observed recent trend, which is represented by the GPCP. The results for the near future (2016–2019) indicate a positive tendency with respect to the previous four years period (2011–2014) (MME6 and MME6sk).

We also tested the WAMI simulated by the models against the one obtained with two reanalysis data set. According to

our analysis (Sect. 3.3), several models show skill when the WAMI from the decadal runs is tested against the NCEP one, with the main contribution coming from the initialization (Fig. 6). Nevertheless, there are evident differences in the results when using the ERA40 reanalysis, for which models do not show any prediction ability (Fig. 7), regardless of the actual regions used to define the WAMI for this data set (not shown). In Sect. 3.1 we also highlighted some remarkable differences between the main CEOF modes obtained from NCEP and ERA40 reanalyses, with the former explaining a much higher variability than the latter (57 and 39 %, respectively). Such differences are also reflected in the principal component time series associated with the CEOF1 (Figure S1). Moreover, the wind patterns associated with this first mode show stronger and more uniform loads over West Africa for NCEP than for ERA40 (Figs. 1, 2). These results are also supported by those obtained when the precipitation is included in the CEOF analysis. Also in this case, the patterns of the leading mode CEOF1 show substantial differences (not shown), which are reflected in the variance explained (46 % for NCEP and 32 % for ERA40). Therefore, our results suggest that the description of decadal variability of the West African monsoon using wind modulus at 925 hPa and zonal wind at 200 hPa is better captured in NCEP reanalysis than in ERA40.

Gaetani and Mohino (2013), by using simulated and observed Sahelian precipitation indices, assessed the ability of several CMIP5 models in predicting the decadal variability of rainfall over Sahel, and they showed that the predictive skills are highly model dependent, in agreement with our results. In their study, they highlighted the ability of the CanCM4 and the MPI-ESM-LR models in predicting Sahel rainfall at long lead times. Consistently with such results, both models show skill when testing the performance of WAMI against observed rainfall (Fig. 4). In addition, we show that the unskilful results of BCC-CSM1-1, FGOALS-g2, MIROC5 and HadCM3 models using precipitation indices (Figure S2) can be improved when WAMI is used as a proxy of rainfall (Tables S1, S2 and S3 summarize the predictive skills, for precipitation indices and WAMI, of the 14 models analysed, reporting the results from decadal hindcasts, historical simulations and residuals, respectively). Such result suggests that these models are capable of simulating the dynamics associated with the West African monsoon and its variability at decadal time scales, but unable to translate it into rainfall, which shows a weak response (BCC-CSM1-1 and FGOALS-g2 models) or it is badly located to the south (HadCM3 and MIROC5 models) (Fig. 4).

5 Conclusions

The capability to predict the decadal variability of Sahel rainfall has been evaluated in 14 state-of-the-art coupled ocean-atmosphere global climate models using dynamics-based indices. Our approach is based on the strong link between the West African monsoon winds at low and high tropospheric levels and rainfall over the Sahel, which led Fontaine et al. (1995) to propose the WAMI as a proxy of Sahel rainfall. To select the most convenient regions for the definition of the indices, we have performed a CEOF analysis of wind modulus at 925 hPa and zonal wind at 200 hPa. All models show stronger monsoon flow at low levels in connection to a stronger TEJ at high levels (Figs. 1, 2). However, the anomaly patterns associated to CEOF1 can show substantial differences among the analysed models. The correlation coefficients between the CEOF1 PCs time series and the WAMI defined of the basis of the CEOF analysis confirm the robustness of the CEOF method for defining the WAMI.

When assessing the predictive skills by using the observed SPI, models' performance for the decadal hindcasts is higher than for the historical experiments. Similar results are obtained when the observed WAMI is defined for the NCEP reanalysis. However, substantial differences are observed when using ERA40 reanalysis: first, CEOF1 shows smaller loads and explains less variability than for NCEP reanalysis; second, most of the models show no skill when WAMI is tested against the one defined from the ERA40 reanalysis. These results suggest that NCEP reanalysis should be preferred when assessing decadal variability of the WAM dynamics, while ERA40 description of the WAM dynamics would deserve further investigation to understand the observed discrepancies.

Our results show there is significant skill when predicting Sahel rainfall at decadal time scales in some models and that the initialization plays a major role, in agreement with Gaetani and Mohino (2013), and Martin and Thorncroft (2014). As in their work, we find significant skill in the decadal hindcasts of the CanCM4 and MPI-ESM-LR models using the WAMI. Moreover, our analysis reveals significant skill in the decadal hindcasts for several models, which did not show skill when a simulated Sahelian precipitation index was used (BCC-CSCM-1, FGOALS-g2, HadCM3 and MIROC5). However, there is no predictive skill for the WAMI in other models that showed significant skills for their rainfall outputs (CNRM-CM5, GFDL-CM2p1 and MME). Therefore, our results suggest that it is potentially possible to improve the skill for predicting rainfall over Sahel in some models by using dynamics-based indices. We thus recommend a two-fold approach when testing the performance of models in predicting Sahel rainfall, based not only on rainfall but also on the dynamics of the WAM.

Our results also suggest that the initialized decadal experiments can be used for skilful decadal probabilistic forecasts of above, below and near normal rainfall seasons over the Sahel. In addition, issuing decadal hindcasts

every year (instead of every 5 years, which is the standard approach in CMIP5) would also allow the analysis of such probabilistic forecast separately for each lead time.

Finally, based on the WAMI analysis, we show that models predict a positive tendency in Sahel rainfall over the next four year period (2016–2019) with respect to the last one (2011–2014).

References

Adler RF, Huffman GJ, Chang A et al (2003) The version 2 global precipitation climatology project (GPCP) monthly precipitation analysis (1979-Present). J Hydrometeor 4:1147–1167

Bader J, Latif M (2003) The impact of decadal-scale Indian Ocean sea surface temperature anomalies on Sahelian rainfall and the North Atlantic Oscillation. Geophys Res Lett 30:2169–2173

Barnston AG (1992) Correspondence among the correlation, RMSE, and heidke forecast verification measures; refinement of the heidke score. Weather Forecast 7:699–709

Biasutti M, Held IM, Sobe AH, Giannini A (2008) SST forcings and Sahel rainfall variability in simulations of the twentieth and twenty-first centuries. J Clim 21:3471–3486

Caminade C, Terray L (2010) Twenty century Sahel rainfall variability as simulated by the ARPEGE AGCM and future changes. Clim Dyn 35:75–94

Chang P, Saravanan R, Ji L, Hegerl GC (2000) The effect of local sea surface temperatures on atmospheric circulation over the tropical Atlantic sector. J Clim 13:2195–2216

Dai A, Trenberth KE, Qian T (2004) A global data set of Palmer Drought Severity Index for 1870–2002: relationship with soil moisture and effects of surface warming. J Hydrometeorol 5:1117–1130

Davies JR, Rowell DP, Folland CK (1997) North Atlantic and European seasonal predictability using an ensemble of multi- decadal AGCM simulations. Int J Climatol 17:1263–1284

Diouf I, Deme A, Ndione JA, Gaye AT, Rodríguez-Fonseca B, Cissé M (2013) Climate and health: observation and modeling of malaria in the Ferlo (Senegal). C R Biologies 336:253–260

Doblas-Reyes FJ, Weisheimer A, Palmer TN, Murphy JM, Smith D (2010) Forecast quality assessment of the ENSEMBLES seasonal-to-decadal stream 2 hindcasts. ECMWF Technical Memoranda 621

Doblas-Reyes FJ, Andreu-Burillo I, Chikamoto Y, Garcia-Serrano J, Guemas V, Kimoto M, Mochizuki T, Rodrigues LRL, van Oldenborgh GJ (2013) Initialized near-term regional climate change prediction. Nat Commun 4:1715. doi:10.1038/ncomms2704

Fink AH, Schrage JM, Kotthaus S (2010) On the potential causes of the nonstationary correlations between West African precipitation and Atlantic hurricane activity. J Clim 23:5437–5456

Folland C, Palmer T, Parker D (1986) Sahel rainfall and worldwide sea temperatures, 1901–1985. Nature 320:602–607

Fontaine B, Janicot S (1992) Wind field coherence and its variations over West Africa. J Clim 5:512–524

Fontaine B, Janicot S, Moron V (1995) Rainfall anomaly patterns and wind field signals over West Africa in August (1958–1989). J Clim 8:1503–1510

Fontaine B, Gaetani M, Ullmann A, Roucou P (2011) Time evolution of observed July–September sea surface temperature-Sahel climate teleconnection with removed quasi-global effect (1900–2008). J Geophys Res 116:D04105. doi:10.1029/2010JD014843

Gaetani M, Fontaine B (2013) Interaction between the West African Monsoon and the summer Mediterranean climate: an overview. Física de la Tierra 25:41–55

Gaetani M, Mohino E (2013) Decadal prediction of the Sahelian precipitation in CMIP5 simulations. J Clim 26:7708–7719. doi:10.1175/JCLI-D-12-00635.1

Gaetani M, Fontaine B, Roucou P, Baldi M (2010) Influence of the Mediterranean Sea on the West African monsoon: intraseasonal variability in numerical simulations. J Geophys Res 115:D24115

García Serrano J, Guemas V, Doblas-Reyes FJ (2015) Added-value from initialization in predictions of Atlantic multi-decadal variability. Clim Dyn 44:2539–2555. doi:10.1007/s00382-014-2370-7

García-Serrano J, Doblas-Reyes FJ (2012) On the assessment of near-surface global temperature and North Atlantic multi-decadal variability in the ENSEMBLES decadal hindcast. Clim Dyn 39:2025–2040

García-Serrano J, Doblas-Reyes FJ, Haarsma RJ, Polo I (2013) Decadal prediction of the dominant West African monsoon rainfall modes. J Geophys Res Atmos 118:5260–5279. doi:10.1002/jgrd.50465

Garric G, Douville H, Déqué M (2002) Prospects for improved seasonal predictions of monsoon precipitation over Sahel. Int J Climatol 22:331–341. doi:10.1002/joc.736

Giannini A, Saravanan R, Chang P (2003) Oceanic forcing of Sahel rainfall on interannual to interdecadal time scale. Science 302:1027–1030

Grist JP, Nicholson SE (2001) A study of the dynamic factors influencing the rainfall variability in the West African Sahel. J Clim 14:1337–1359

Haarsma RJ, Selten FM, Weber SL, Kliphuis M (2005) Sahel rainfall variability and response to greenhouse warming. Geophys Res Lett 32:L17702. doi:10.1029/2005GL023232

Harris I, Jones PD, Osborn TJ, Lister HD (2014) Updated high-resolution grids of monthly climatic observations—the CRU TS3.10 Dataset. Int J Climatology 34:623–642

Hoerling MP, Hurrell J, Eischeid JK (2006) Detection and attribution of 20th Century northern and southern African rainfall change. J Clim 19:3989–4008. doi:10.1175/JCLI3842.1

Ickowicz A, Ancey V, Corniaux C, Duteurtre G, Poccard-Chappuis R, Toure I, Vall E and Wane A (2012) Crop-livestock production systems in the Sahel—increasing resilience for adaptation to climate change and preserving food security. Building resilience for adaptation to climate change in the agriculture sector FAO/OECD Rome 243–276

Ineson S, Scaife AA (2009) The role of the stratosphere in the European climate response to El Nino. Nature Geosci 2:32–36

International CLIVAR Project Office (ICPO) (2011) Data and bias correction for decadal climate predictions. International CLIVAR Project Office CLIVAR Publication Series 150:6

Kalnay E, Kanamitsu M, Kistler R et al (1996) The NCEP/NCAR 40-year reanalysis project. Bull Am Meteorol Soc 77:437–471. doi:10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2

Kandji ST, Verchot S, Mackensen J (2006) Climate Change and variability in the Sahel region: impacts and adaptation strategies in the Agricultural sector. World Agroforestry Centre (ICRAF)

and United Nations Environment Programme (UNEP). UNEP 2006:1–48

Keenlyside NS, Latif M, Jungclaus J et al (2008) Advancing decadal scale climate prediction in the North Atlantic sector. Nature 453:84–88. doi:10.1038/nature06921

Kharin VV, Zwiers FW (2003a) Improved seasonal probability forecasts. J Clim 16:1684–1701

Kharin VV, Zwiers FW (2003b) On the ROC score of probability forecasts. J Clim 16:4145–4150

Kidson JW (1977) African rainfall and its relation to the upper air circulation. Quart J Roy Meteor Soc 103:441–456

Kim HM, Webster PJ (2012) Curry JA (2012) Evaluation of short-term climate change prediction in multi-model CMIP5 decadal hindcasts. Geophys Res Lett 39:L10701. doi:10.1029/2012GL051644

Kirtman B, Power SB, Adedoyin JA et al (2013) Near-term climate change: projections and predictability. In: Stocker TF, Qin D, Plattner G-K, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM (eds) Climate change 2013: the physical science basis. Contribution of Working Group I to the fifth assessment report of the intergovernmental panel on climate change. Cambridge University Press, Cambridge, UK

Kohler M, Kalthoff N, Kottmeier Ch (2010) The impact of soil moisture modifications on CBL characteristics in West Africa: a case-study from the AMMA campaign. Quart J Roy Meteor Soc 136:442–455. doi:10.1002/qj.430

Lu J, Delworth T (2005) Oceanic forcing of the late 20th century Sahel drought. Geophys Res Lett 32:L22706. doi:10.1029/2005GL023316

Martin ER, Thorncroft C (2014) Sahel rainfall in multimodel CMIP5 decadal hindcasts. Geophys Res Lett: 41. doi:10.1002/2014GL059338

McIntire J (1981) Food security in the Sahel: variable import levy, grain reserves and foreign exchange assistance. Research report 26 International Food Policy Research Institute Washington USA

Meehl GA et al (2009) Decadal prediction: can it be skillful? Bull Amer Meteor Soc 90:1467–1485

Meehl GA, Goddard L, Boer G et al (2014) Decadal climate prediction: an update from the trenches. Bull Am Meteorol Soc 95:243–267. doi:10.1175/BAMS-D-12-00241.1

Miyakoda K, Hembree GD, Strickler RF, Shulman I (1972) Cumulative results of extended forecast experiments I. Model performance for winter cases. Mon Wea Rev 100:836–855. doi:10.1175/1520-0493(1972)100<0836:CROEFE>2.3.CO;2

Mochizuki T, Ishii M, Kimoto M et al (2010) Pacific decadal oscillation hindcasts relevant to near-term climate prediction. Proc Natl Acad Sci USA 107:1833–1837

Mohino E, Janicot S, Bader J (2011a) Sahel rainfall and decadal to multi-decadal sea surface temperature variability. Clim Dyn 37:419–440

Mohino E, Rodríguez-Fonseca B, Losada T, Gervois S, Janicot S, Bader J, Ruti P, Chauvin F (2011b) SST-forced signals on West African rainfall from AGCM simulations-Part I: changes in the interannual modes and model intercomparison. Clim Dyn 37:1707–1725. doi:10.1007/s00382-011-1093-2

Mohino E, Rodríguez-Fonseca B, Mechoso CR, Gervois S, Ruti P, Chauvin F (2011c) Impacts of the tropical Pacific/Indian Oceans on the seasonal cycle of the West African monsoon. J Clim 24:3878–3891. doi:10.1175/2011JCLI3988.1

Moron V, Philippon N, Fontaine B (2004) Simulation of West African monsoon circulation in four atmospheric general circulation models forced by prescribed sea surface temperature. J Geophys Res 109:D24105. doi:10.1029/2004JD004760

Ndiaye O, Ward MN, Thiaw WM (2011) Predictability of seasonal Sahel rainfall using GCMs and lead-time improvements through the use of a coupled model. J Clim 24:1931–1949. doi:10.1175/2010JCLI3557.1

Newell RE, Kidson JE (1984) African mean wind changes between Sahelian wet and dry periods. J Climatol 4:24–33

Nicholson SE (2013) The West African Sahel: a review of recent studies on the rainfall regime and its interannual variability. ISRN Meteorology. doi:10.1155/2013/453521

Philippon N, Doblas-Reyes FJ, Ruti P (2010) Skill, reproducibility and potential predictability of the West African monsoon in coupled GCMs. Clim Dyn 35:53–74. doi:10.1007/s00382-010-0856-5

Rodríguez-Fonseca B, Mohino E, Mechoso CR et al (2015) Variability and predictability of West African droughts: a review in the role of sea surface temperature anomalies. J Clim 28:4034–4060. doi:10.1175/JCLI-D-14-00130.1

Smith DM, Cusack S, Colman AW, Folland CK, Harris GR, Murphy JM (2007) Improved surface temperature prediction for the coming decade from a global climate model. Science 317:796–799

Solomon A, Goddard L, Kumar A, Carton J, Deser C, Fukumori I, Stockdale T (2011) Distinguishing the role of natural and anthropogenically forced decadal climate variability: implications for prediction. Bull Am Meteorol Soc 2:141–156

Taylor KE, Stouffer RJ, Meehl GA (2012) An overview of CMIP5 and the experiment design. Bull Amer Meteor Soc 93:485–498

Ting M, Kushnir Y, Seager R, Li C (2009) Forced and internal twentieth-century SST trends in the North Atlantic. J Clim 22:1469–1481

Trenberth KE et al (2007) Observations: Surface and atmospheric climate change. In: Climate change 2007: the physical science basis. Contribution of working group I to the fourth assessment report of the intergovernmental panel on climate change [Solomon S et al (eds)]. Cambridge University Press, Cambridge and New York, NY

Uppala SM, Ållberg PW, Simmons AJ et al (2005) The ERA-40 reanalysis. Q J R Meteorol Soc 131:2961–3012

van Oldenborgh GJ, Doblas-Reyes FJ, Wouters B, Hazeleger W (2012) Decadal prediction skill in a multi- model ensemble. Clim Dyn 38:1263–1280. doi:10.1007/s00382-012-1313-4

Venegas SA (2001) Statistical methods for signal detection in climate. Danish Center for Earth System Science Rep 2:46

Venzke S, Allen MR, Sutton RT, Rowell DP (1999) The atmospheric response over the North Atlantic to decadal changes in sea surface temperature. J Clim 12:2562–2584

Villamayor J, Mohino E (2015) Robust Sahel drought due to the Interdecadal Pacific Oscillation in CMIP5 simulations. Geophys Res Lett Res Lett 42:1214–1222. doi:10.1002/2014GL062473

Wang G, Eltahir EAB (2000) Role of vegetation dynamics in enhancing the low-frequency variability of the Sahel rainfall. Water Resour Res 36:1013–1021. doi:10.1029/1999WR900361

Zeng N, Neelin JD, Lau KM, Tucker CJ (1999) Enhancement of interdecadal climate variability in the Sahel by vegetation interaction. Science 286:1537–1540