1
2 **Decadal prediction of Sahel rainfall: where does the skill (or lack thereof)**
3 **come from?**
4
5
6

7 Elsa Mohino (*)
8   Meteorology and Geophysical Department, Universidad Complutense de
9   Madrid
10
11   Address:
12   Dpto. Física de la Tierra, Astronomía y Astrofísica I
13   Facultad Ciencias Físicas
14   Universidad Complutense de Madrid
15   Ciudad Universitaria, Plaza Ciencias, 1, 28040 Madrid, SPAIN
16
17
18 Noel Keenlyside
19
20   Geophysical Institute, University of Bergen
21   Bjerknes Centre for Climate Research
22
23   Address:
24   Geophysical Institute, University of Bergen, Allegaten 70, 5007, Bergen,
25   Norway
26
27
28 Holger Pohlmann
29
30   Max-Planck-Institut für Meteorologie, Hamburg
31
32   Address:
33   Max Planck Institute for Meteorology, Bundesstrasse 53, 20146 Hamburg,
34   Germany
35
36
37 (*) Corresponding author: Elsa Mohino
38 email: emohino@ucm.es
39 Telephone: +34 91 394 4390
40 Fax: +34 91 394 4398

41
42 ABSTRACT
43
44 Previous works suggest decadal predictions of Sahel rainfall could be skillful. However, the sources
45 of such skill are still under debate. In addition, previous results are based on short validation periods
46 (i.e. less than 50 years). In this work we propose a framework based on multi-linear regression
47 analysis to study the potential sources of skill for predicting Sahel trends several years ahead. We
48 apply it to an extended decadal hindcast performed with the MPI-ESM-LR model that span from 1901
49 to 2010 with one year sampling interval. Our results show that the skill mainly depends on how well
50 we can predict the timing of the global warming (GW), the Atlantic multidecadal variability (AMV)
51 and, to a lesser extent, the inter-decadal Pacific oscillation (IPO) signals, and on how well the system
52 simulates the associated SST and West African rainfall response patterns. In the case of the MPI-
53 ESM-LR decadal extended hindcast, the observed timing is well reproduced only for the GW and
54 AMV signals. However, only the West African rainfall response to the AMV is correctly reproduced.
55 Thus, for most of the lead times the main source of skill in the decadal hindcast of West African
56 rainfall is from the AMV. The GW signal degrades skill because the response of West African rainfall
57 to GW is incorrectly captured. Our results also suggest that initialized decadal predictions of West
58 African rainfall can be further improved by better simulating the response of global SST to GW and
59 AMV. Furthermore, our approach may be applied to understand and attribute prediction skill for other
60 variables and regions.
61
62
63 KEYWORDS: decadal climate predictions, Sahel, Atlantic Multidecadal variability, Global Warming,
64 Climate Variability

2

65 1. INTRODUCTION
66
67 The Sahel is an African semiarid region located between the Sahara desert to the north and the
68 Savanna to the south. The economy of the region is mostly based on rain-fed agriculture and pastures
69 for livestock, which makes the Sahel highly vulnerable to rainfall variability (Kandji et al. 2006;
70 Ickowicz et al. 2012). Rainfall over the Sahel is strongly tied to the West African Monsoon and shows
71 variability at many different time scales (Rodríguez-Fonseca et al. 2015). However, its signal at
72 decadal timescales is outstanding (Dai et al. 2004): it showed a wet period in the 1950s-1960s
73 followed by a devastating drought in the 1970s and 1980s and a recent recovery since the 1990s.
74 Predicting the trends in Sahel rainfall several years ahead would be highly beneficial for decision
75 making and planning in the region.
76
77 The decadal fluctuations in Sahel rainfall have been associated with worldwide changes in the sea
78 surface temperatures (SST) in different ocean basins (Rodriguez-Fonseca et al. 2015, and references
79 therein). The warming of the Indian (e.g. Giannini et al. 2003; Bader and Latif 2003), the tropical
80 Pacific (e.g. Lu and Delworth 2005; Caminade and Terray 2010) and the tropical Atlantic south of
81 10ºN (Hagos and Cook 2008) all lead to drought conditions over the Sahel. On the other hand, the
82 warming of the Mediterranean Sea (Park et al. 2016) or the differential warming of the northern and
83 southern hemispheres, particularly over the Atlantic basin (e.g. Hoerling et al. 2006; Ting et al. 2009),
84 has also been suggested as a driver of Sahel rainfall variability. In turn, Mohino et al. (2011)
85 suggested that the evolution of Sahel rainfall at decadal timescales could be explained by the
86 competing effects of the Atlantic Multidecadal Variability (AMV), the Interdecadal Pacific
87 Oscillation (IPO) and the global warming trend (GW), especially in the tropics. Decadal prediction of
88 Sahel rainfall could be possible if these factors could be predicted.
89
90 There have been several attempts to dynamically predict climate changes at decadal time scales
91 (Smith et al. 2007; Pohlmann et al. 2009; Keenlyside et al. 2008; Mochizuki et al. 2010; Taylor et al.
92 2012). These predictions lie between an initial value problem and a forced boundary condition
93 problem (Meehl et al. 2009a; Keenlyside and Ba 2010): On the one hand, at these time scales the
94 changes in the external forcing (CO_2 or aerosols, for instance) can affect the evolution of climate
95 variables and, on the other, the effect of internal variability can be of comparable magnitude (Meehl et
96 al. 2009a). The decadal climate predictions are based on coupled atmosphere-ocean models run for 10
97 years forced by the observed and projected changes in the external forcing and initialized from an
98 observed state at the beginning of the simulation (Taylor et al. 2012). These have shown skill on
99 decadal timescales in predicting SST in the extra-tropical North Atlantic and the Indian Ocean and
100 Western Pacific, but very little skill in other regions (Doblas-Reyes et al. 2013).
101
102 Several works have analysed the prediction skill of Sahelian rainfall using decadal hindcasts. The
103 work by van Oldenborgh et al. (2012) suggested no significant skill in the ENSEMBLES multi-model
104 system, and little or no impact of the initialization. However, García-Serrano et al. (2013) suggested
105 that, though not skillful, there was an encouraging tendency to produce positive correlation scores
106 between the observed and modelled Sahelian mode of variability in the multi-model. Bellucci et al.
107 (2015) reported statistically significant skill for the multi-model mean hindcasts run in the framework
108 of EU COMBINE project, though a large spread for individual models. In addition, the decadal
109 climate hindcasts performed by some models in the framework of the Coupled Model
110 Intercomparison Project – Phase 5 (CMIP5) show skill in reproducing the decadal variations of the
111 Sahelian rainfall index (Gaetani and Mohino 2013; Martin and Thorncroft 2014; Otero et al. 2015).
112 The sources for such skill are under debate. Martin and Thorncroft (2014) suggest that the skill comes
113 from the correct reproduction of the link between Sahel rainfall and the relative SST index (RSI),
114 defined in Giannini et al. (2013) as the difference of SSTs in the subtropical North Atlantic minus the

3

115   mean in the Tropics. In turn, Gaetani and Mohino (2013) suggest that skillful predictions are provided
116   by models that properly reproduce the observed SST-precipitation relationship at multi-decadal time
117   scales, which they relate to the AMV and IPO modes. This is in agreement with García-Serrano et al.
118   (2015), who suggested that skillful predictions of Sahel rainfall could be obtained with the
119   initialization of AMV and a proper representation of the Sahel-AMV teleconnection in the models.
120
121   However, all the previous studies are based on a validation period typically spanning from 1960 to
122   2010, which is too short when taking into account that the Sahel rainfall is related to the AMV, a
123   phenomenon with a timescale of 60-80 years (Kerr 2000). In particular, Müller et al. (2014) showed
124   that the variability in the North Atlantic is dominated by the trend from the 1960s, while the
125   oscillating components become more important for longer periods. In this work we aim to investigate
126   the sources of skill (or lack thereof) for the decadal prediction of Sahel rainfall using the decadal
127   climate predictions performed with the MPI-ESM-LR model for the period 1901-2010 (Müller et al.
128   2014). In section 2 we describe the experiments and methods used and in section 3.1 we show the
129   skill of the model in predicting Sahel rainfall. To better understand the sources of skill and errors in
130   the prediction of Sahel rainfall we propose a new methodology based on the analysis of the fidelity of
131   the simulations in reproducing the observed decadal modes of SST variability (namely the GW, AMV
132   and IPO) and their impact on West African rainfall (section 3.2). In section 3.3 we subsequently use
133   multi-linear regression analysis to partition the variance of the Sahel rainfall indices into different
134   terms related to the GW, AMV and IPO signals. The same analysis is used to decompose the
135   correlation of the simulated indices with respect to the observed ones in terms of the same SST signals.
136   In section 4 we discuss our results and in section 5 we draw the main conclusions from our research.
137   This research will shed light into the feasibility of decadal predictions in an area highly vulnerable to
138   climate variability and change.
139
140   2. DATA AND METHODS
141
142   2.1 Observed data:
143
144   We use CRUTS3.1 (Harris et al. 2014) for rainfall estimates. It is a global gridded dataset based on
145   observed precipitation amounts with data only over land regions. It spans the period 1901 to 2009
146   with a horizontal resolution of 0.5º.
147
148   To account for the uncertainty in sea surface temperature, especially in the trends (Vecchi et al. 2008;
149   Falvey and Garreaud 2009), two different datasets are used: HadISST1 (Rayner et al. 2003) and
150   ERSSTv3b (Smith et al. 2008). Both are reconstructed datasets based on ship measurements.
151   HadISST1 also uses satellite estimates from the mid 1980s onwards. The datasets also differ on the
152   sources of ship data (ICOADS for ERSSTv3b, and Met Office Marine Data Bank for HadISST1) and
153   on the reconstruction methodology. They are both gridded data sets with a horizontal resolution of 1°
154   x 1° and 2° x 2° for the HadISST1 and the ERSSTv3b, respectively, and global coverage. They span
155   the period from 1870 to present for HadISST1 and from 1854 onwards for the ERSSTv3b data set.
156
157   Observed rainfall and SST data are re-gridded to the coarser resolution mesh of the model
158   (approximately 1.9º in longitude and latitude).
159
160
161   2.2 Simulations:
162
163   Two different types of simulations from the low resolution version of the Max Plank Institute climate

164 model (MPI-ESM-LR, Müller et al. 2014) are used, namely the historical simulation and the decadal
165 hindcasts. The former is a long forced run performed in the framework of the Coupled Model
166 Intercomparison Project Phase 5 (CMIP5, Taylor et al. 2012) that spans from 1850 to 2005 and is
167 initialized from a preindustrial control run. It takes into account the observed changes in
168 anthropogenic (greenhouse-gas composition and aerosols) and natural (solar and volcanic) forcing.
169 The decadal hindcasts consist in a set of 10-year long runs that use the same forcing as the historical
170 run, but are anomaly initialized every year from 1901 to 2010 (Müller et al. 2014). The initial
171 conditions are taken from an assimilation experiment in which the model is nudged towards the ocean
172 salinity and temperature anomalies added to the model's climatology. These anomalies are obtained
173 from runs of the ocean component of the model forced with individual members of the 20CR (Compo
174 et al., 2011) reanalysis (Müller et al. 2015). The results shown in this work are based on averaging the
175 three ensemble members available for each experiment. By such averaging, we attempt to remove
176 climate variability in the model not coming from the external forcing nor from the initialization.
177
178 2.3 Methods:
179
180 To characterize the temporal evolution of Sahel rainfall, we define the Sahel rainfall index (SRI) as
181 the average of rainfall in the region 18ºW-10ºE, 10ºN-20ºN during the summer months (July to
182 September).
183
184 In this work we define the prediction skill as a measurement of the skill one can expect when trying to
185 predict some quantity, like the Sahel rainfall, using only information available before the actual
186 prediction issue date: If one wants to predict an index in year X with a lead time of Y, we can only use
187 simulations initialized previous to year X-Y. Two measures of the prediction skill are used: Anomaly
188 correlation coefficients (ACC) and root mean square error (RMSE). To avoid noise from interannual
189 phenomena, these scores are applied to averages of 4 years (similar results are obtained when using 5
190 years, not shown). For the observations and the historical simulations, such averaging is equivalent to
191 applying a 4-year running mean. For the decadal hindcasts the averages are applied separately in each
192 10-yr long simulation, with the first possible forecast verification being the average of 1 to 4 year lead
193 time. Following the World Climate Research Programme recommendations, anomalies in the decadal
194 hindcasts are defined with respect to the lead time (ICPO 2011). In this way, the climatology to
195 subtract at each lead time τ is calculated using only the year τ of all the decadal hindcast runs. This
196 procedure removes the mean forecast drift. For direct comparison, the historical simulation is treated
197 in the same way. The period to evaluate the skill scores is between 1914 and 2004 for all simulations
198 and lead times.
199
200 It is important to attribute the skill to physical mechanisms. These are assessed by estimating the
201 ability of the model to simulate the observed variations and this is analysed in a different framework
202 than the prediction one: We are no longer constrained by the issue date of the predictions and we can
203 also use information that would be available later than the date we are analysing. This allows the use
204 of low-pass filters with sharper response functions at the cut-off frequency than the running mean.
205 This is a standard approach for analysing historical simulations. Here we extended it to the decadal
206 hindcasts by repeating the analysis on the separate time series constructed by concatenating the
207 hindcast data for each lead time (i.e., constructing a hindcast series for each lead time). For the sake of
208 brevity, in the following we will refer to the degree of reproduction of an observed phenomenon as
209 the simulation fidelity. Thus we will assess how the simulation fidelity varies with hindcast lead time.
210
211 Mohino et al. (2011) found that most of the decadal and multi-decadal evolution of Sahel rainfall
212 could be explained by the combined influence of three modes of variability of SSTs at these time
213 scales: the GW, AMV and IPO. In this work we follow the same methodology to analyse the fidelity

5

in simulating the SRI and its sources. The GW signal index is defined as the SST averaged between 45ºS and 60ºN and low-pass filtered with a cut-off period of 40 years (using a 4th order Butterworth filter). The other two signals, AMV and IPO, are estimated as the principal component associated with the first empirical orthogonal function (EOF) over the Atlantic and the Indo-Pacific basins, respectively. Prior to the EOF calculation, the GW signal is removed at each grid point by means of a linear regression and the field is low-pass filtered with a cut-off period of 13 years (using a 4th order Butterworth filter).

To estimate the rainfall and SST patterns related with each component, the raw rainfall data are regressed onto the appropriate standardized index. The statistical significance of correlations for filtered time series is estimated with the following Monte Carlo approach, also used in Mohino et al. (2011): 1,000 pairs of synthetic time series of the same length as the original ones are produced using a Gaussian distribution and low-pass filtered using the same filter. The correlation between each pair is calculated to build a probability density distribution of correlations against which the original correlation is compared. The use of an AR-1 process to generate the synthetic time series does not alter the conclusions (not shown). To evaluate the fidelity in simulating Sahel rainfall decadal variability we have low-pass filtered the SRI with a cut-off period of 13 years (again with a 4th order Butterworth filter), which we name hereafter as decadal-SRI. Unlike the observations and the historical simulations, in which there is only one temporal coordinate, in the decadal predictions the analysis of the relation of the indices (GW, AMV, IPO) with rainfall has been applied for each lead time separately. In this way we can analyse the ability of the model to simulate the SST modes and associated indices as a function of the lead time. We can see whether the fidelity in simulating the temporal and spatial characteristics of a particular signal changes as the model evolves from the initialised state close to the observed.

We use multilinear regression analysis to evaluate the contribution of each signal to the decadal variability of the SRI:

$$SRI_{t,k} = \alpha_k GW_{t,k} + \beta_k AMV_{t,k} + \gamma_k IPO_{t,k} + \epsilon_{t,k} \quad (1)$$

where $SRI_{t,k}$ denotes the decadal low pass filtered SRI; $GW_{t,k}$, $AMV_{t,k}$ and $IPO_{t,k}$ the GW, AMV and IPO indices, respectively; α_k, β_k and γ_k are the coefficients of the multilinear regression with GW, AMV and IPO indices, respectively; $\epsilon_{t,k}$ is the residual of the fitting; the subscripts t and k refer respectively to the time and the data subset (observations, historical run, and the decadal hindcasts for each lead time).

The variance of the SRI ($Var[SRI_k]$) can then be partitioned into the following components:

$$Var[SRI_k] = {\alpha_k}^2 + {\beta_k}^2 + {\gamma_k}^2 + 2\beta_k \gamma_k Cov[AMV_k, IPO_k] + Var[\epsilon_k] \quad (2)$$

where we have taken into account that the GW, AMV and IPO indices are standardized (to have unit variance) and that the GW component is orthogonal (by construction) to the other two SST indices. The term $Cov[AMV_k, IPO_k]$ stands for the covariance between the AMV and the IPO indices.

In addition, the correlations of any simulated SRI index (SRI_k) with the observed one can be explained in terms of the multilinear regression coefficients and the correlations between the observed SRI index (SRI_o) and the simulated SST indices GW_k, AMV_k and IPO_k:

$$262 \quad \rho(SRI_o, SRI_k) = \frac{\alpha_k}{\sqrt{Var[SRI_k]}} \rho(SRI_o, GW_k) + \frac{\beta_k}{\sqrt{Var[SRI_k]}} \rho(SRI_o, AMV_k)$$

$$263 \quad + \frac{\gamma_k}{\sqrt{Var[SRI_k]}} \rho(SRI_o, IPO_k) + \frac{\sqrt{Var[\epsilon_k]}}{\sqrt{Var[SRI_k]}} \rho(SRI_o, \epsilon_k) \qquad (3)$$

264

265 where, $\rho(X,Y)$ is the correlation between time series X and Y and $Var(X)$ the variance of time series
266 X.

267

268 The simulation fidelity is analysed for the 1910-2005 period when there is data for the historical
269 simulation and for all lead times of the decadal hindcasts.

270

271 3. RESULTS

272

273 3.1 SRI Prediction Skill

274

275 The decadal hindcasts of Sahel rainfall show statistically significant ACC for all lead times, although
276 less skilful than persistence; the ACC skill of historical runs is not significant (Fig. 1). In accordance
277 with previous works (Gaetani and Mohino 2013; Bellucci et al. 2015; Martin and Thorncroft 2014;
278 García-Serrano et al. 2013; Otero et al. 2015) the skill in predicting Sahel rainfall changes with lead
279 time: the biggest ACC scores are obtained at lead times 2-5 and 3-4 years. ACC scores are
280 subsequently reduced until lead time 7-10 years, when they increase. Roughly the opposite can be
281 observed for RMSE scores. There are several possible reasons for such changes in skill: initial shock
282 and non-linear drifts could be responsible for lower skill values at short lead times, while long-term
283 trends could enhance skill at long lead times, as Bellucci et al. (2015) found. However, when the ACC
284 scores are recomputed using the same time series with the trend previously removed (labeled as
285 "detrended" in Fig. 1), they are not reduced at middle lead times and are even enhanced at long leads
286 (7-9 years), which suggests that the skill at these time scales is not coming from the long-term trend.

287

288 3.2 Simulation fidelity

289

290 To understand the reasons for the prediction skill scores obtained in Fig. 1, in this section we examine
291 the model fidelity in reproducing the characteristics of the GW, AMV and IPO decadal modes of SST
292 variability: their phase and their associated SST and West African rainfall patterns. We show that the
293 timing is only well captured by the decadal hindcasts for the first two. We also show that the SST and
294 rainfall patterns are adequately reproduced only for the AMV case, though the magnitude of rainfall
295 anomalies over West Africa is too weak.

296

297 *Indices*

298

299   We begin by considering the ability of the historical simulation and the decadal hindcasts in
300 capturing the phase of the three SST signals (GW, AMV and IPO) and of the SRI (Fig. 2). It is worth
301 noting that for the decadal hindcasts the indices presented in Fig. 2 have been obtained independently
302 for each lead time τ by concatenating in time the simulated outputs (i.e., we create a hindcast-series
303 for each lead time). The Sahel rainfall index presented in Fig. 2d is the low-pass filtered SRI time
304 series, using a cut-off frequency of 13 years, which we name decadal-SRI.

305

306 The observed GW index shows a warming trend modulated by multidecadal variability, with two
307 strong warming periods (1920-1940 and 1970-2000), a hiatus from the 1950s to 1960s, and a
308 suggestion of a hiatus at the end of the record (Fig. 2a). There is still a debate on the causes for such

309   periods of acceleration and deceleration in the warming trend (Keenlyside and Ba 2010; Booth et al.
310   2012; Tung and Zhou 2013; Kosaka and Xie 2013; Chen and Tung 2014; Mann et al. 2014; England
311   et al. 2015). The historical simulation slightly overestimates the total amplitude of the warming and
312   the 1970-2000 trend and underestimates the 1920-1940 trend and the mid-20th Century warming
313   slowdown. The decadal hindcasts overestimate the amplitude of the 20th Century warming at all lead
314   times, especially at the short ones (maximum overestimation is 50% of the observed amplitude at lead
315   time 3 years). In addition, the mid-20th Century slowdown is only shown for long lead times. The
316   correlations between the GW indices in observations and in the decadal hindcasts at all lead times are
317   positive and statistically significant (Fig. 3a). This is also the case for the historical simulation,
318   suggesting that this component is strongly related to the external forcing.
319
320   The observed AMV time series are consistent between both data sets and with previous works (Fig. 2;
321   e. g. Cai and Whetton 2001; Baines and Folland 2007; Ting et al. 2009; Trenberth and Shea 2006;
322   Zhang and Delworth 2006; Tung and Zhou 2013; Svendsen et al. 2014). The historical simulation
323   ensemble is not able to capture the time series of the observed AMV signal (Fig. 2b), showing
324   negative correlations (Fig. 3b). Conversely, the decadal hindcasts capture the phasing of the observed
325   signal (Fig. 2b), showing positive correlations for all lead times (Fig. 3b). Such results are in
326   agreement with previous studies that show prediction skill for the AMV component for a range of
327   lead times (García-Serrano et al. 2015). However, the decadal hindcasts fail to capture the observed
328   second maxima in the 1950s and tend to delay the AMV changes of phase in the mid-1960s and late
329   1990s at long lead times. The decadal hindcasts also show an unrealistic cold period in the 1910s,
330   which is related to too cold anomalies over the central and western North Atlantic in the initial
331   conditions provided (not shown).
332
333   The observed IPO index shows regime shifts from a positive phase to a negative one in the mid 1940s
334   and late 1990s and a change from a negative phase to a positive one in the mid-1970s (Fig. 2c),
335   coherent with previous works (Power et al. 1999; Mantua and Hare 2002; Meehl et al. 2009b; Mohino
336   et al. 2011; Villamayor and Mohino 2015). However, neither the historical simulation nor the decadal
337   hindcasts capture such shifts, showing statistically non-significant or even negative correlations (Fig.
338   3c). Though disappointing, such results are consistent with the lack of skill for decadal prediction of
339   SSTs that models tend to show in the Pacific basin (Kim et al. 2012; Doblas-Reyes et al. 2013;
340   Bellucci et al. 2015). Despite this lack of ability to capture the observed IPO, there is a certain
341   memory in the IPO index from the decadal hindcasts: except for the longest one (10 years), the
342   different lead times show some consistency among themselves (Fig. 2d), with always positive
343   correlations and mostly statistically significant (at the 5% level), especially at lead time 6 years (table
344   1).
345
346   Regarding the decadal-SRI, the observation shows two main regimes, a wet one up to 1970, with
347   maximum rainfall anomalies occurring in the 1950s, and a dry one since 1970, peaking in the 1980s
348   (Fig. 2d). The change from the 1950s to the 1980s represents nearly a 30% of the total mean average
349   rainfall of the region. The variance of the decadal-SRI represents just over 50% of the total variability
350   of the unfiltered SRI time series. There is also a suggestion of a recent recovery in the 1990s, though
351   it does not reach the levels previous to 1970. The weak recovery shown in our analysis is probably
352   due to the area chosen to define the SRI, which covers the western Sahel (18ºW-10ºW), where rainfall
353   did not recover, and its central part (10ºW-10ºE), where the recovery was clear (Lebel and Ali 2009).
354
355   All the simulated and predicted decadal-SRI show smaller amplitude than the observed one (Fig. 2d).
356   The decadal-SRI from the historical simulation does neither capture the strong positive rainfall regime
357   in the 1950s nor the strong negative one in the 1980s. Compared to GW and AMV indices, the SRI
358   indices in the decadal hindcasts are less coherent with observations or among different lead times (Fig.

8

2). Nevertheless, most lead times show positive trends between 1910-1940, and between 1970-2000, consistent with observations and previous analysis (Müller et al. 2014). There is also a negative trend in the 1970s that is delayed for longer lead times, consistently with the behaviour in the hindcast AMV indices. The 1950s observed rainfall maximum is not well simulated for most lead times, while they tend to capture the 1930s one, which is again coherent with the simulation of the AMV signal.

The decadal-SRI index from the hindcasts show correlations above 0.4 for most lead times (Fig. 3d). Note these correlations are evaluated by comparing observed and hindcast decadal-SRI indices (Fig. 2d) and are different from the ACC scores presented in Fig. 1 (see methods for details on the calculation of the ACC scores to evaluate prediction skill). The correlations show the biggest values at lead time of 3 and 8 years, while the lowest are for lead times of 7 and 9 years. This is roughly consistent with the prediction ACC skill scores (Fig. 1, undetrended): ACC prediction scores are biggest at short lead times (2 to 5 and 3 to 6 years), which take into account the simulation of year 3. Consistently with the prediction skill (Fig. 1), the historical experiment shows negative correlations for the simulation of the decadal-SRI.

Regression patterns
To further understand the skill in simulating the decadal-SRI we show the SST and rainfall patterns associated with the three SST indices. The regression of the observed SST onto the GW index show remarkable differences between both observational data sets, especially over the Pacific Ocean (Fig. 4, Vecchi et al. 2008; Falvey and Garreaud 2009; Mohino et al. 2011). Thus, despite the strong consistency in the time evolution of the GW signal, there is high uncertainty in the spatial pattern associated with it. However, both datasets tend to show higher loads over the tropical and southern Atlantic and Indian basins and reduced loads over the north Atlantic and north-west Pacific ocean. Such an SST pattern is related to a reduction of rainfall over the Sahel, especially in the western part (Fig. 4a,b). Unlike the observations, the historical simulation shows a SST pattern with higher loads in the Northern Hemisphere than in the tropics, especially over the Pacific basin (Fig. 4m). The GW signal is associated with increased rainfall over the entire West Africa, except at the westernmost side of the Sahel. This agrees with Park et al. (2015), who show that models with a northward global SST inter-hemispheric gradient tend to show increased rainfall over the Sahel. Rainfall increases are not restricted to the Sahel, but they are also shown in other regions north of the equator, like the Asian monsoon or north-west South America, while there are rainfall deficits south of the equator (Fig. 4m). This suggest a northward shift of the intertropical convergence zone, which could be related to the global northward SST inter-hemispheric gradient (Hwang and Frierson 2013).

Regarding the decadal hindcasts, the SST patterns associated with the GW signal show changes with lead time (Fig. 4): at short lead times, there are strong loads in the tropics, which peak in year 3, coinciding with the strongest amplitude of the GW time series (Fig. 2a). Such strong tropical SST gradients are neither consistent with the observations nor with the historical simulation. As the decadal hindcast evolves in lead time, the trends in the tropics are gradually reduced, while those in the north Pacific get stronger and resemble more the GW SST pattern in the historical simulation (Fig. 4m). This suggests that the strong tropical trends shown by the decadal hindcasts in the first four years of the simulation are related to the initialization and could be connected to the 20CR (Compo et al. 2011) reanalysis used to force the ocean component of the model (Müller et al. 2014), though further research is needed to clarify this point. The regression of the GW index on rainfall also changes with hindcast lead time: there are mainly negative loads over the Sahel at short lead times, especially at lead time 3 years, while positive loads dominate for long lead times. This is consistent with the evolution of the GW SST pattern: the warm tropical SSTs have been related to a decrease of rainfall over the Sahel (e.g. Giannini et al. 2003, 2013; Bader and Latif 2003; Lu and Delworth 2005;

9

409 Caminade and Terray 2010; Mohino et al. 2011), while the northward SST gradient shown at long
410 lead times leads to increased Sahel rainfall, as discussed before.
411
412 The observed AMV SST pattern shows positive anomalies over the whole of the North Atlantic, with
413 strongest loads over the Atlantic subpolar gyre region (Fig. 5). Positive loads are also shown over the
414 North Pacific, while negative anomalies are present in the Indian and the South Atlantic oceans. The
415 observed AMV signal is strongly related to positive anomalies of Sahel rainfall (Fig. 5). The
416 historical simulation exhibits an AMV SST pattern with positive anomalies over the north Atlantic,
417 especially over its eastern border, and is associated with positive Sahel rainfall anomalies, though
418 very weak. The AMV is not significantly related to SST anomalies outside the Atlantic in the
419 historical simulations. Conversely, the AMV SST patterns in the decadal hindcasts show much higher
420 consistency with observations: positive anomalies over the north Atlantic, mainly over the subpolar
421 and midlatitude regions, and the north Pacific. It should be also noted that there is a great persistence
422 in the AMV SST regression patterns with lead time, higher than for the GW component (Fig. 4). In
423 particular, the positive anomalies in the northern extratropics are present throughout the 10 years of
424 the simulations. In addition, the AMV related rainfall anomalies, though weaker than the observed,
425 are positive over the Sahel for all lead times. Their amplitude varies with lead time, being maximum
426 for lead time 3 years (approximately 60 % of the observed anomaly) and minimum for 1 and 2 year
427 lead times (approximately 15 and 20 % of the observed anomaly, respectively). There are, however,
428 some mismatches between the simulated and observed patterns, the most striking is the eastward shift
429 of the strong positive SST anomalies in the north Atlantic in the simulation. In addition, the SST
430 anomalies over the tropical north Atlantic are weaker than observed. This could be related to a
431 reduced atmospheric feedback mechanism (Zhang 2007), since the rainfall anomalies are weaker (Fig.
432 5), or to an under-representation of the external forcing fingerprint on AMV, as has been found to be
433 more important in the tropics than in the extratropics (Terray 2012).
434
435 Regarding the SST patterns associated with the IPO index, the observations show positive SST
436 anomalies over the eastern tropical Pacific, while in the subtropics there are negative SST anomalies,
437 which are more intense in the Northern Hemisphere (Fig. 6; e.g., Meehl et al. 2009b). The observed
438 IPO is linked to reduced Sahel rainfall (Mohino et al. 2011; Villamayor and Mohino 2015). Despite
439 the model's inability to capture the timing of the IPO signal, its SST pattern is well simulated in the
440 historical run. The anomalies of Sahel rainfall, though they agree in sign with the observed ones, are
441 weak and non-significant. In the decadal hindcasts, the IPO pattern shows less consistency with
442 observations, and also varies greatly with lead time: there are unrealistically high positive loads in the
443 tropical Pacific at short lead times that weaken at long lead times (except for lead time 10 years).
444 Sahel rainfall anomalies related to the IPO in the decadal hindcasts are negative for all lead times, but
445 are only statistically significant for some of them.
446
447
448
449 3.3. Multi-linear regression analysis
450
451 In this section we apply multi-linear regression analysis between the decadal-SRI and the time series
452 of GW, AMV and IPO. Using this fit, we decompose the variance of the decadal-SRI and the
453 correlations shown in Fig. 3d into terms related to the three SST signals (see section 2.3 for details in
454 the calculations). This decomposition allows the quantification of the effect of each SST signal on the
455 reproduction of the observed decadal-SRI. We show that the coefficients of the multi-linear regression
456 analysis are highly related to the regression patterns shown in the previous section, and we assess
457 causes for their discrepancies to observations. With them, we further understand the large RMSE
458 errors in Fig.1 as due to the weak response of rainfall to the SST patterns and we show that the main

459 positive contributor to the correlations is the AMV component.
460

461 *Relationship between the multi-linear regression coefficients and the regression patterns*
462 We begin by evaluating the relationship between the coefficients of the multi-linear regression
463 analysis ($\alpha_k, \beta_k, \gamma_k$) with the SST and rainfall regression patterns shown in the previous section (Fig.
464 4-6). Note that, unlike for simple linear regression analysis, the multi-linear regression coefficients
465 can differ from the average rainfall regression over the Sahel. In Fig. 7a we show that for the GW case,
466 the α_k coefficient is approximately equal to the regression of the GW index over Sahel region. These
467 coefficients are always negative, as in the observations, except for the historical experiment and for
468 the decadal hindcasts at lead times 7 and 9. In addition, we find that α_k is related to the averaged
469 tropical SSTs patterns of the GW signal (between 15ºS and 15ºN, Fig. 7d). This is consistent with the
470 AGCM experiments in Mohino et al. (2011) that showed most of the rainfall response to the GW SST
471 pattern over West Africa could be simulated using only the tropical part of the pattern. The overall
472 warming in the tropical basins leads to drought over the Sahel: the Indo-Pacific warming leads to
473 subsidence over West Africa and the warming in the Atlantic leads to a southward shift of the West
474 African monsoon (Mohino et al. 2011). However, as discussed before, such positive tropical warming,
475 especially over the Tropical Pacific, is not consistent with observations (Fig. 4 and 7d) and seems to
476 be related to an artifact of the initialization (not shown).
477

478 For the AMV case, there is a strong and linear relationship between β_k and the corresponding average
479 regression of rainfall over the Sahel in Fig. 5 (Fig. 7b). These values are always positive. In addition,
480 β_k shows sensitivity to the meridional SST gradient near the west coast of the Sahel (Fig. 7d). The
481 link between β_k and SST pattern is less clear than for the α_k term (Fig. 7d,e). A weaker link is
482 expected as the GW SST regression patterns show much greater diversity than the AMV ones, which
483 are much more consistent among themselves (Fig. 4, 5). The strength of the precipitation response is
484 nonetheless found to be sensitive to the coastal gradient of SSTs in the eastern tropical North Atlantic.
485

486 For the IPO case, the magnitudes of the γ_k coefficients are smaller than those obtained from
487 averaging rainfall regression values over the Sahel in Fig. 6 (Fig. 7c). A linear relationship between
488 both is still present, though weaker than for the other two SST components. Conversely, the γ_k
489 coefficients do not show a clear relationship with the SST regression patterns shown in Fig. 6: unlike
490 what would be expected from previous works (Villamayor and Mohino, 2015), the γ_k coefficient is
491 not negatively related to the magnitude of SST anomalies over the tropical Pacific (Fig. 7f).
492

493 *Variance of the decadal-SRI*
494 We now partition the total variance of the decadal-SRI into components related to GW, AMV, IPO,
495 an AMV-IPO interaction term, and an unexplained residual (equation 2, Fig 8a). The observational
496 decadal-SRI is mostly explained by the AMV and GW components (between 43 % and 57 % for the
497 former and nearly 20 % for the latter, amounting to between 63 % and 75 % of the total variance
498 when using ERSSTv3b and HadISST1 datasets, respectively), with a smaller contribution of the IPO
499 and the AMV-IPO covariance terms. The total variance accounted for in the observations is 86 %
500 (80 %) when using HadISST1 (ERSSTv3b) data set. In the decadal hindcasts the AMV is the main
501 signal contributing to the total variance, while the residual dominates in the historical simulation (Fig.
502 8a).
503

504 The simulated decadal-SRI variance is roughly 10% of the observed (between 3% and 27%,
505 depending on the lead time). A fraction of this underestimation is due to the averaging over the 3
506 ensemble members. However, accounting for this by instead averaging the variance of the individual
507 ensemble members only increases variance of the simulated decadal-SRI to 20% of the observed. The
508 underestimation of the decadal variability of Sahel rainfall is consistent with the results from most

509 coupled models (Villamayor and Mohino 2015), and could be related to the inability of the
510 atmospheric component to simulate such decadal variability (Rodriguez-Fonseca et al. 2015). Poor
511 representation of land surface-atmosphere interactions, including vegetation, is the possible cause for
512 the underestimation (Giannini et al. 2003; Wang et al. 2004). However, recently Vellinga et al. (2016)
513 suggested that the main problem is coarse model resolution that hinders the simulation of strong
514 rainfall events, which conveys the decadal signal. Further evidence to support the hypothesis of the
515 underestimation being primary related to the atmospheric component of the model comes from the
516 regression patterns in Figs. 4-6: for all three signals, the strength of the SST anomalies simulated in
517 both experiments is similar to the observed one, while over West Africa, the model underestimates the
518 observed rainfall anomalies.
519
520 From this analysis we can conclude that the low total variance of the SRI in the simulations is mainly
521 due to a weak response of Sahel rainfall to the simulated SST signals, especially the AMV. The
522 underestimation of variance contributes to the large RMSE of the historical simulations and decadal
523 hindcasts of SRI (Fig 1).
524
525 *Correlations of simulated and observed decadal-SRI*
526 We further investigate the correlations of the decadal-SRI from the historical simulation and the
527 decadal hindcasts with the observations (Fig. 3d), partitioning them into four terms related to the GW,
528 AMV, IPO and the residual in the multi-linear regression analysis (equation 3, fig 8b). The total
529 correlation is given by the blue dashed line (corresponding to Fig 3d). The contribution from the
530 AMV term is the largest one (except in the historical simulation, Fig. 8b). This term contributes
531 positively to the correlations in the decadal hindcasts and negatively in the historical simulation. The
532 magnitude of the contribution is relatively stable. The minimum values are found for the historical
533 experiment and for lead times 1 and 2 years in the decadal hindcasts. To further explain these
534 contributions we turn to the $\frac{\beta_k}{\sqrt{Var[SRI_k]}}\rho(SRI_o, AMV_k)$ term in equation (3), which accounts for the
535 AMV contribution to the correlations. The β_k coefficients are smallest for the historical simulation
536 and the first two lead times (Fig. 7b), which explains the lower magnitude of their contributions at
537 those lead times and experiment. In addition, the correlations between the observed decadal-SRI and
538 the hindcast AMV indices ($\rho(SRI_o, AMV_k)$) are strong and positive for all lead times (table 2). This
539 together with the positive β_k coefficients explains the systematic positive contribution of AMV to the
540 correlation between observed and hindcast decadal-SRI. Conversely, as the $\rho(SRI_o, AMV_k)$ is
541 negative for the historical simulation (table 2), the AMV contributes negatively to the correlation
542 between observed and simulated decadal-SRI. Therefore, the positive AMV contribution to the
543 fidelity of the hindcast decadal-SRI results from skill in both capturing the AMV timing and
544 reproducing the associated SST and precipitation patterns, while for the historical simulation the
545 incorrect timing of the AMV causes its reduced decadal-SRI fidelity.
546
547 On average the GW signal is the second largest contributor to the correlations between simulated and
548 observed decadal-SRI. Interestingly, the GW signal clearly enhances the decadal hindcasts
549 correlations only at short lead times (1 to 4 years), while its contribution is weak at long ones. The
550 contribution is even negative for lead times 7 and 9 years in the decadal hindcasts and for the
551 historical simulation. To further explain these contributions we turn to the $\frac{\alpha_k}{\sqrt{Var[SRI_k]}}\rho(SRI_o, GW_k)$
552 term in equation (3), which accounts for the GW contribution to the correlations between simulated
553 and observed decadal-SRI. The correlations between the observed decadal-SRI and the modelled GW
554 indices ($\rho(SRI_o, GW_k)$) are strong and negative for the decadal hindcasts for all lead times and the
555 historical simulation (Table 2). Therefore, the different contribution of GW to the decadal-SRI
556 correlation is explained by the $\frac{\alpha_k}{\sqrt{Var[SRI_k]}}$ term, which is only strong and negative (as in the

observations) for lead times 1 to 4 years. Therefore, the main problems with the GW contribution are not due to the phasing of the index, which is well captured, but with the simulated SST and rainfall patterns: the response of Sahel rainfall is consistent with the observed one only in the decadal hindcasts and mainly at short lead times. This explains why detrending the SRI time series enhances the prediction skill at long lead times (Fig 1). Note, however, that the consistent Sahel response at short lead times is due to an unrealistic SST pattern, as commented in the previous section.

The IPO contribution to the simulation skill of the decadal-SRI is typically smaller than for the GW and AMV signals (Fig. 8b). In the historical experiment and for most lead times in the decadal hindcasts its contribution is positive. The exceptions can be understood analysing the $\frac{\gamma_k}{\sqrt{Var[SRI_k]}} \rho(SRI_o, IPO_k)$ term: the negative contribution at lead time 1 year is due to the temporal correlations between the hindcasts IPO index and the observed decadal-SRI ($\rho(SRI_o, IPO_k)$), which is positive (table 2) while the regression coefficient γ_k is negative (as observed); and at lead time year 5 the γ_k regression coefficient is positive (opposite to observed, Fig. 7c) while the temporal correlation between the hindcast IPO index and the observed decadal-SRI is negative (table 2). For the historical experiment and the decadal hindcasts at other lead times the temporal correlation between the simulated IPO index and the observed decadal-SRI is negative and γ_k is negative (as observed). The maximum contribution to the decadal-SRI fidelity is at lead time 9 years. The very weak contribution of IPO signal to the skill for the historical simulations and for the lead time 10 years of the decadal hindcasts is due to the temporal correlation between such indices and the observed decadal-SRI (table 2). The large variations in the IPO contribution among lead times suggest the contribution of the IPO to fidelity results from chance, rather than any underlying mechanism.

4. DISCUSSION

We have used multiple linear regression to decompose the decadal-SRI into GW, AMV, and IPO components. The regression coefficients associated with GW (α_k) and AMV (β_k), explain most of the variance of the decadal-SRI in both the observations and model. However, the model underestimates strongly the variance associated with these components contributing to the large RMSE of the hindcast and historical experiments. Furthermore, these two regression coefficients, rather than the phase of the GW and AMV indices, explain most of the correlations between hindcast and observed decadal-SRI (Fig. 8b). These coefficients are related to the rainfall response to the GW and AMV SST patterns (Fig. 4 to 5), which show a mismatch between model and observations: For both, α_k and β_k, the model coefficients are much weaker than the observed ones for similar levels of SST averages (Fig. 7). This is consistent with other studies, discussed above, that show models (probably their atmospheric component) tend to underestimate the response of Sahel rainfall to decadal SST signals. However, this alone would not explain the incorrect positive α_k values found in some simulations, which are associated with a stronger warming of the tropics than the observed. Instead the warmer conditions in the northern hemisphere (between 15ºN and 60ºS) shown in the historical-simulation GW patterns (0.27 ºC mean anomaly) with respect to the observed patterns (0.14 ºC mean anomaly) would help increase rainfall over the Sahel, despite the warm tropical anomalies (Park et al. 2015). In addition, the MPI-ESM-LR atmospheric component is also very sensitive to the direct Greenhouse Gas effect (Janicot et al. 2015), which has been shown to increase Sahel rainfall (Haarsma et al. 2005; Held et al. 2005; Hoerling et al. 2006; Giannini 2010; Dong and Sutton 2015). As the model only simulates the direct aerosol effect, it may also underestimate the tropospheric cooling effect of anthropogenic aerosols, leading to an overestimation of Sahel rainfall long-term trends (Kawase et al. 2010).

There is currently a debate on the origin of the AMV component, some works point at atmospheric

13

aerosols accounting for much of the 20th Century AMV evolution (Rotstayn and Lohmann 2001; Ottera et al. 2010; Booth et al. 2012), while other works suggest an internal origin related to the strength of the Atlantic Meridional Overturning Circulation (AMOC; Zhang 2007; Knight et al. 2005; Knight 2009; Ting et al. 2009, 2011; Knudsen et al. 2011; Zhang et al. 2013; Hodson et al. 2014, Keenlyside et al. 2015). Recent works also suggest that the indirect anthropogenic aerosol influence in the last part of the 20th Century might be irrelevant due to a saturation effect of cloud albedo sensitivity to aerosol emission (Carslaw et al. 2013; Stevens 2013). The historical simulation captures both the tropical part of the SST pattern related to the AMV as well as the rainfall response over the Sahel. However, the AMV index shows statistically significant negative correlation with the observed one. Thus while in this model the external forcing projects onto a tropical AMV pattern, it induces a signal with incorrect timing, decreasing the fidelity in simulating the decadal-SRI (Fig. 8). The weaker response of Sahel rainfall to this component in the historical simulation (Fig. 4) also reduces the relative contribution to the total variance of the decadal-SRI (Fig. 8a). In contrast, in the decadal hindcasts the AMV signal is captured skilfully (Fig. 3b), and it contributes to the greater total decadal-SRI variance, except for lead time 1 year, and to the positive correlation with the observed decadal-SRI (Fig. 8). Thus, in our experiments external forcing does not seem to account for the hindcast fidelity in reproducing the AMV signal. It should be noted, however, that the model is not simulating aerosol indirect effects, which are found by Booth et al. (2012) to explain most of the aerosol impact on North Atlantic SSTs. Nevertheless, the decadal hindcasts show a negative impact of the simulated AMV signal on subsurface temperatures in the Tropical Atlantic south of 20ºN-25ºN (not shown), consistent with AMOC changes (Zhang 2007; Zhang et al. 2013), whereas the historical simulation shows no such change (not shown). This is in agreement with other studies which show an important role for ocean dynamics in climate prediction of the North Atlantic region (Yeager et al. 2012).

The lack of skill in hindcasting the IPO may be due to inherent less predictability (Latif et al. 2006), which could in part arise from the multiple factors composing its north Pacific component (Schneider et al. 2002), or because of model and initialization errors. Such lack of skill negatively affects our expectations for predicting Sahel rainfall at decadal time scales. However, as the observed variance of Sahel rainfall at these time scales is mainly related to the AMV and GW components (Fig. 8a), decadal predictions systems could still attain better skill over the Sahel by focusing on these two components rather than the IPO.

5. CONCLUSIONS

Unlike previous works that evaluate the skill of decadal predictions of rainfall over West Africa, in this study we have focused on analysing the potential sources for such skill. For this aim, we proposed a new framework based on the multi-linear regression analysis of the decadal Sahel rainfall index (SRI), decomposing it into the contributions of GW, AMV and IPO (Mohino et al. 2011). We have tested such approach using the decadal hindcasts performed by Müller et al. (2014), which provide 10-year long predictions issued every year from 1901 to 2010, and by comparing with the historical simulation with the same version of the MPI-ESM-LR model.

The overall skill in predicting the decadal-SRI can be understood as a combination of how well we can predict the timing of the three different SST signals (GW, AMV and IPO) and how well we can simulate the SST and West African rainfall patterns in response to such signals.

In the case of the GW signal, both experiments show high skill in capturing the timing of the signal, as it is mostly related with the 20th Century trend. However, the decadal hindcasts tend to overestimate the observed amplitude of the signal, especially at lead times 2 to 4 years over the

14

tropical Pacific. Despite a reasonable timing, the SST and rainfall response patterns are not as well captured. The average rainfall response to the GW over the Sahel is weak in the simulations and even positive in some cases (contrary to observations). Such positive rainfall loads lead to a negative contribution of the GW signal to the skill in simulating the decadal-SRI. The most extreme case is the historical simulation in which the timing is the best captured (Fig. 3a) and the averaged northward SST gradient is maximum (Fig. 4m), leading to the most negative contribution to the skill. Interestingly, the biggest positive contribution of the GW to reproducing the decadal-SRI is in the decadal hindcasts at lead times 1 to 4 years. However in this case we have a good response for the wrong reasons: the too strong spurious tropical warming trends correct other model deficiencies in simulating the GW impact on Sahel rainfall and lead to an average negative Sahel rainfall response.

For the AMV signal, the timing and the SST and rainfall response patterns are adequately reproduced in the decadal hindcasts. Such reproduction leads to a consistent positive contribution of this component to the skill in reproducing the decadal-SRI. The main differences among the different lead times are related to the strength of the rainfall response rather than the rainfall pattern. In turn, such strength seems to be related to the local gradients of SST in the east tropical North Atlantic, off the coast of West Africa. Conversely, for the historical simulations, the timing is wrongly reproduced, which leads to a negative contribution to the skill in simulating the decadal-SRI.

For the IPO signal, the observed timing is not captured by any of the simulations. The observed global SST patterns are better captured in terms of spatial correlation and root mean square error (not shown), but not properly translated to the precipitation patterns over West Africa. This component shows, however, marginal contributions to enhance the skill in simulating the decadal-SRI at some lead times, though part of such enhancement could be related to a contamination by the AMV signal.

The analysis of the different contributions to the correlation between observed and simulated (hindcast and historical) decadal-SRI (i.e., low-pass filtered time series that incorporate future information into the forecast) indices explain some of the characteristics of the actual prediction skill (Fig. 1): at long lead times the skill obtained for the total predicted SRI is similar or even smaller than the one obtained after detrending the data because the impact of the GW signal on rainfall is wrongly simulated. Moreover, the biggest ACC skill scores for the total predicted SRI is obtained for lead times 2-5 and 3-6 for the wrong reasons: the initialization introduces a spurious trend, mainly in the tropical Pacific, which overcomes the tendency of the model to enhance Sahel rainfall with the GW (contrary to the observed behaviour). Most of the skill in the predicted SRI comes from the AMV signal, because the initialization adequately captures the timing of the signal as well as a reasonablly realistic SST pattern and associated rainfall response over the Sahel. Thus, conversely to the suggestions from other works (Bellucci et al. 2015), the skill at long lead times does not stem only from the external forcing: the AMV obtained through the initialization was the only source of skill for Sahel rainfall in our decadal hindcasts at long lead times.

Our multi-linear regression analysis framework can also explain the total variance obtained in the decadal-SRI. The main contribution in both observations and simulations comes from the AMV component. Those simulations for which the AMV signal is weakly translated to the Sahel rainfall, the total variance is very low.

It should be noted that our results are based on only one model and, as shown by previous work, the predictive skill of Sahel rainfall is model dependent (Gaetani and Mohino 2013; García-Serrano et al. 2013; Martin and Thorncroft 2014; Bellucci et al. 2015). In the framework of CMIP5, not many models show skill. When analysing rainfall outputs, approximately 30% (4 out of the 14 analysed) of models show significant skill and the MPI-ESM-LR is one of them (Otero et al. 2015). Gaetani and

Mohino (2013) showed that for this particular model, the initialization was playing the key role by incorporating a multi-decadal SST-precipitation relationship similar to the observed one and that this was the case only in 37% of the models analysed.

The analysis presented in this work was possible due to the long length of the validation period (the whole 20th Century), for which a separation of the GW, AMV and IPO components following our methodology was possible. Though issuing predictions for such a long period presents big challenges and some problems may arise (like the too cold AMV signal in the 1910s or the too strong SST trends at short lead times in the GW component), it offers a unique opportunity to evaluate the skill of predicting phenomena with very long periods like the AMV (period of 65 years) for which the standard CMIP5 approach of simulating from 1960 to 2010 does not even cover 1 cycle of the phenomenon.

Finally, leaving aside the IPO influence, we have shown that part of the reduced skill shown by the simulated decadal-SRI index is due to failures in the model to represent properly some characteristics of the SST patterns and, especially, translate the GW and AMV signals correctly to the precipitation anomalies. If atmospheric response to these two signals was correctly simulated then skill could be greatly increased. To quantify this we construct a new multiple regression model using the GW and AMV model indices but replacing the corresponding regression coefficients from the model with the observed ones; we neglect the IPO and residual contributions. This significantly enhances the correlations obtained (Fig 9). By improving the rainfall response to the two SST components, the attainable fidelity in simulating the decadal-SRI skill is much higher in the initialized simulations than in the historical one. These results suggest that there is still room for skill improvement. In particular, for the MPI-ESM-LR model, there are two key factors to address: On the one hand, the SST pattern associated with the GW component should be improved to resemble better the observations (i.e. show less warming in the Northern Hemisphere extratropics and more in the Southern Ocean in response to the slowly varying external forcing). This would likely lead to a negative response of the simulated Sahel rainfall (as observed) and enhance the ACC scores in the decadal hindcasts. On the other hand, the magnitude of the response of the simulated Sahel rainfall to decadal SST modes should be increased to match the observed one. This would likely reduce the RMSE of the hindcast SRI.

6. ACKNOWLEDGEMENTS

REFERENCES

Bader J, Latif M (2003) The impact of decadal scale Indian Ocean SST anomalies on Sahelian rainfall and the North Atlantic Oscillation. Geophys Res Lett 30:2169. doi:10.1029/2003 GL018426

16

755 Baines PG, Folland CK (2007) Evidence for a rapid global climate shift across the late 1960s. J Clim
756   20:2721–2744. doi:10.1175/ JCLI4177.1

757 Bellucci A, Haarsma R, Gualdi S et al. (2015) An assessment of a multi-model ensemble of decadal
758   climate predictions. Clim Dyn 44:2787-2806

759 Booth BBB, Dunstone NJ, Halloran PR, Andrews T, Bellouin N (2012) Aerosols implicated as a
760   prime driver of twentieth-century North Atlantic climate variability. Nature 484:229-232

761 Cai W, Whetton PH (2001) Modes of SST variability and the fluctuation of global mean temperature.
762   Clim Dyn 17:889–901

763 Caminade C, Terray L (2010) Twentieth century Sahel rainfall variability as simulated by the
764   ARPEGE AGCM, and future changes. Clim Dyn. 35:75-94. doi:10.1007/s00382-009-0545-4

765 Carslaw KS, Lee LA, Reddington CL et al. (2013) Large contribution of natural aerosols to
766   uncertainty in indirect forcing. Nature 503:67-17

767 Chen X, Tung KK (2014) Varying planetary heat sink led to global-warming slowdown and
768   acceleration. Science 345:897-903

769 Compo GP, Whitaker JS, Sardeshmukh PD, Matsui N, Allan RJ, Yin X, Gleason BE, Vose RS,
770   Rutledge G, Bessemoulin P, Brönnimann S, Brunet M, Crouthamel RI, Grant AN, Groisman PY,
771   Jones PD, Kruk MC, Kruger AC, Marshall GJ, Mauger M, Mok HY, Nordli Ø, Ross TF, Trigo RM,
772   Wang XL, Woodruff SD, Worley SJ (2011) The twentieth century reanalysis project. Q J R
773   Meteorol Soc 137:1–28

774 Dai A, Trenberth KE, Qian T (2004) A global data set of Palmer Drought Severity Index for 1870–
775   2002: relationship with soil moisture and effects of surface warming. J Hydrometeorol 5:1117–
776   1130

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

780 Dong B, Sutton R (2015) Dominant role of greenhouse-gas forcing in the recovery of Sahel rainfall.
781   Nature Climate Change 5:757-760

782 England MH, Kajtar JB, Maher N (2015) Robust warming projections despite the recent hiatus. Nat
783   Clim Change 5: 394-396

784 Falvey M, Garreaud RD (2009) Regional cooling in a warming world: recent temperature trends in
785   the southeast Pacific and along the west coast of subtropical South America (1979–2006). J
786   Geophys Res 114. doi:10.1029/2008JD010519

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

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

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

17

794 Giannini A (2010) Mechanisms of Climate Change in the Semiarid African Sahel: The Local View. J
795   Clim 23:743-756

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

798 Giannini A, Salack S, Lodoun T, Ali A, Gaye AT, Ndiaye O (2013) A unifying view of climate
799   change in the Sahel linking intra-seasonal, interannual and longer time scales. Environ Res Lett
800   8:024010, doi:10.1088/1748-9326/8/2/024010

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

803 Hagos S, Cook K (2008) Ocean warming and late-twentieth-century Sahel drought and recovery. J
804   Clim 21:3797–3814. doi: 10.1175/2008JCLI2055.1

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

807 Held IM, Delworth TL, Lu J, Findell KL, Knutson TR (2005) Simulation of Sahel drought in the 20th
808   and 21st centuries. PNAS 102:17891–17896

809 Hoerling M, Hurrell JW, Eischeid J, Phillips AS (2006) Detection and attribution of twentieth-century
810   northern and southern African rainfall change. J Clim 19:3989–4008

811 Hodson DLR, Robson JI, Sutton RT (2014) An anatomy of the cooling of the North Atlantic Ocean in
812   the 1960s and 1970s. J Clim 27:8229-8243

813 Hwang YT, Frierson DMW (2013) Link between the double-intertropical convergence zone problem
814   and cloud biases over the Souther Ocean. PNAS 110:4935-4940

815 Ickowicz A, Ancey V, Corniaux C, Duteurtre G, Poccard-Chappuis R, Toure I, Vall E and Wane A
816   (2012) Crop-livestock production systems in the Sahel—increasing resilience for adaptation to
817   climate change and preserving food security. Building resilience for adaptation to climate change in
818   the agriculture sector FAO/ OECD Rome 243–276

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

821 Janicot S, Gaetani M, Hourdin F et al. (2015) The recent partial recovery in Sahel rainfall: a
822   fingerprint of Greenhouse Gases Forcing? GEWEX 27:11-15

823 Kawase H, Abe M, Yamada Y, Takemura T, Yokohata T, Nozawa T (2010) Physical mechanism of
824   long-term drying trend over tropical North Africa. Geophys Res Lett 37:L09706

825 Kandji ST, Verchot S, Mackensen J (2006) Climate Change and variability in the Sahel region:
826   impacts and adaptation strategies in the Agricultural sector. World Agroforestry Centre (ICRAF)
827   and United Nations Environment Programme (UNEP). UNEP 2006:1–48

828 Keenlyside NS and Ba J (2010) Prospects for decadal climate prediction. Wiley In terdisciplinary
829   Reviews: Climate Change, 1, 627-635

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

832 Keenlyside NS, Ba J, Mecking J, Omrani NO, Latif M, Zhang R, Msadek R (2015) North Atlantic

multi-decadal variability - mechanisms and predictability. Climate Change: Multidecadal and Beyond, Chang C-P, Ghil M, Latif M, Wallace M, Eds, World Scientific Publishing Company, Singapore, n/a. ISBN 978-9814579926

Kerr RA (2000) A North Atlantic climate pacemaker for the centuries. Science 288:1984–1985

Kim HM, Websetr PJ, Curry JA. (2012) Evaluation of short-term climate change prediction in mutli-model CMIP5 decadal hindcasts. Geophys Res Lett 39:L10701

Knight JR (2009) The Atlantic multidecadal oscillation inferred from the forced climate response in coupled general circulation models. J Clim 22:1610–1625. doi:10.1175/2008JCLI2628.1

Knight JR, Allan RJ, Folland CK, Vellinga M, Mann ME (2005) A signature of persistent natural thermohaline circulation cycles in observed climate. Geophys Res Lett 32. doi:10.1029/2005 GL024233

Knudsen MF, Seidenkrantz MS, Jacobsen BH, Kuijpers A (2011) Tracking the Atlantic Multidecadal Oscillation through the last 8,000 years. Nat Comm 2:178

Kosaka Y, SP Xie (2013) Recent global-warming hiatus tied to equatorial Pacific surface cooling. Nature 501:403-407

Latif M, Collins M, Pohlmann J, Keenlyside M (2006) A review of predictability studies of Atlantic sector climate on decadal time scales. J Clim 19:5971-5987

Lebel T, Ali A (2009) Recent trends in the Central and Western Sahel rainfall regime (1990-2007). J Hydrol 375:52-64

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

Mann ME, Steinman BA, Miller SK (2014) On forced temperautre changes, internal variability and the AMO. Geophys Res Lett 41:3211-3219

Mantua NJ, Hare SR (2002) The Pacific decadal oscillation. J Oceanogr 58:35–44

Martin ER, Thorncroft C (2014) Sahel rainfall in multimodel CMIP5 decadal hindcasts. Geophys Res Lett: 41. doi:10.1002/201 4GL059338

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

Meehl GA, Hu A, Santer BD (2009b) The mod-1970s climate shift in the Pacific and the relative roles of forced versus inherent decadal variability. J Clim 22:780–792

Mochizuki T, and Coauthors (2010): Pacific decadal oscillation hindcasts relevant to near-term climateMeehl GA et al (2009a) Decadal prediction: can it be skillful? Bull Amer Meteor Soc 90:1467–1485 prediction. Proc. Natl. Acad. Sci. USA, 107, 1833–1837

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

Müller WA, Pohlmann H, Sienz, F, Smith D (2014) Decadal climate predictions for the period 1901-2010 with a coupled climate model. Geophys Res Lett 41(6):2100–2107, doi:10.1002/2014GL059259

Müller WA, Matei D, Bersch M, Jungclaus JH, Haak H, Lohmann K, Compo GP, Sardeshmukh PD,

871   Marotzke J (2015) A twentieth century reanalysis forced ocean model to reconstruct the North
872   Atlantic climate variation during the 1920s . Clim Dyn 44:1935-1955, doi:10.1007/s00382-014-
873   2267-5

874 Otero N, Mohino E, Gaetani M (2015) Decadal prediction of Sahel rainfall using dynamics- based
875   indices . Clim Dyn, doi:10.1007/s00382-015-2738-3

876 Ottera OH, Bentsen M, Drange H, Suo L (2010) External forcing as a metronome for Atlantic
877   multidecadal variability. Nat Geosci 3:688-694

878 Park JY, Bader J, Matei D (2015) Northern-hemispheric differential warming is the key to
879   understanding the discrepancies in the projected Sahel rainfall. Nat Comm 6,
880   doi:10.1038/ncomms6985

881 Park JY, Bader J, Matei D (2016) Anthropogenic Mediterranean warming essential driver for present
882   and future Sahel rainfall. Nat Comm, doi:10.1038/nclimate3065

883 Pohlmann H,Jungclaus JH, Köhl A, Stammer D, Marotzke J (2009) Initializing decadal climate
884   predictions with the GECCO oceanic synthesis: Effect on the North Atlantic. J. Climate, 22, 3926–
885   3938

886 Power S, Casey T, Folland C, Colman A, Mehta V (1999) Inter- decadal modulation of the impact of
887   ENSO on Australia. Clim Dyn 15:319–324

888 Rayner NA, Parker DE, Horton EB, Folland CK, Alexander LV, Rowell DP (2003) Global analyses
889   of sea surface temperature, sea ice, and night marine air temperature since the nineteenth century. J
890   Geophys Res 108. doi:10.1029/2002JD002670

891 Rodriguez-Fonseca B, Mohino E, Mechoso CR et al (2015) Variability and predictability of West
892   African drougthss: a review on the role of sea surface temperature anomalies. J Clim 28:4034-4060

893 Rotstayn LD, Lohman U (2002) Tropical rainfall trends and the indirect aerosolo effect. J Clim
894   15:2103-2116

895 Schneider N, Miller AJ, Pierce DW (2002) Anatomy of North Pacific decadal variability. J Clim
896   15:586-605

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

899 Smith TM, Reynolds RW, Peterson TC, Lawrimore J (2008) Improvements to NOAA's historical
900   merged land-ocean surface temperature analysis (1880–2006). J Clim 21:2283–2296. doi:
901   10.1175/2007JCLI2100.1

902 Stevens B (2013) Uncertain then, irrelevant now. Nature 503: 47-48

903 Svendsen L, Kvamsto NG, Keenlyside N (2014) Weakening AMOC connects Equatorial Atlantic and
904   Pacific interannual variability. Clim Dyn 43:2931-2941

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

907 Terray L (2012) Evidence for multiple drivers of North Atlantic multi-decadal climate variability.
908   Geophys Res Lett 29:L19712

909 Ting M, Kushnir Y, Seager R, Li C (2009) Forced and internal 20th century SST trends in the North

910   Atlantic. J Clim 22:1469– 1481

911   Ting M, Kushnir Y, Seager R, Li C (2011) Robust features of Atlantic multi-decadal variability and
912   its climate impacts. Geophys Res Lett 38:L17705

913   Trenberth KE, Shea DJ (2006) Atlantic hurricanes and natural variability in 2005. Geophys Res Lett
914   33. doi:10.1029/2006G L026894

915   Tung KK, Zhou J (2013) Using data to attribute episodes of warming and cooling in instrumental
916   records. PNAS 110:2058-2063

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

919   Vecchi GA, Clement A, Solden BJ (2008) Examining the tropical Pacific's response to global
920   warming. EOS 89(9):81–83

921   Vellinga M, Roberts M, Vidale PL, Mizielinski MS, Demory ME, Schiemann R, Strachan J, Bain C
922   (2016) Sahel decadal rainfall variability and the role of model horizontal resolution. Geoph Res
923   Lett 43:326-333. doi:10.1002/2015GL066690

924   Villamayor J , Mohino E (2015) Robust Sahel drought due to the Interdecadal Pacific Oscillation in
925   CMIP5 simulations. Geophys Res Lett 42:1214-1222

926   Wang GL, Eltahir EAB, Foley JA, Pollard D, Levis S (2004) Decadal variability of rainfall in the
927   Sahel: Results from the coupled GENESIS-IBIS atmosphere–biosphere model. Clim Dyn 22:625–
928   637. doi:10.1007/s00382-004-0411-3

929   Yeager S, Karspeck A, Danabasoglu G, Tribbia J, Teng HY (2012) A Decadal Prediction Case Study:
930   Late Twentieth-Century North Atlantic Ocean Heat Content. J Clim 25:5173-5189

931   Zhang R (2007) Anticorrelated multidecadal variations between surface and subsurface tropical North
932   Atlantic. Geophys Res Lett 34: L12713

933   Zhang R, Delworth TL (2006) Impact of Atlantic multidecadal oscillations on India/Sahel rainfall and
934   Atlantic hurricanes. Geophys Res Lett 33. doi:10.1029/2006GL026267

935   Zhang R, Delworth TL, Sutton R et al. (2013) Have aerosolos caused the observed Atlantic Multidecadal
936   variability? J Atmos Sci 70:1135-1144

937
938 TABLES

939

940 Table 1: Anomaly correlation coefficient among the different simulated IPO indices. Statistically
941 significant correlations (at the 5% level) are shown in bold.

	LT1	LT2	LT3	LT4	LT5	LT6	LT7	LT8	LT9	LT10
LT2	**0.84**									
LT3	**0.56**	**0.84**								
LT4	0.28	**0.61**	**0.88**							
LT5	0.45	**0.66**	**0.75**	**0.83**						
LT6	**0.57**	**0.73**	**0.77**	**0.75**	**0.95**					
LT7	0.43	**0.59**	**0.65**	**0.60**	**0.78**	**0.79**				
LT8	0.28	0.39	0.48	0.43	**0.68**	**0.67**	**0.81**			
LT9	0.12	0.35	**0.51**	0.45	**0.63**	**0.65**	**0.61**	**0.83**		
LT10	0.08	0.19	0.40	0.28	0.13	0.29	0.21	0.01	0.23	
HIST	0.01	0.18	0.18	0.25	0.29	0.24	-0.03	-0.17	0.13	0.02

942

943

944 Table 2: Correlation between the observed decadal-SRI and the simulated GW, AMV and IPO time
945 series. The time series were not detrended before the correlation calculation.
946

	SRIo-GWs	SRIo-AMVs	SRIo-IPOs	IPOs-AMVs
LT1	-0.51	0.47	0.25	-0.08
LT2	-0.55	0.48	-0.03	-0.22
LT3	-0.58	0.53	-0.23	-0.46
LT4	-0.60	0.54	-0.37	-0.44
LT5	-0.61	0.52	-0.33	-0.29
LT6	-0.60	0.52	-0.26	-0.27
LT7	-0.60	0.52	-0.31	-0.07
LT8	-0.60	0.53	-0.35	-0.36
LT9	-0.59	0.52	-0.44	-0.67
LT10	-0.56	0.50	-0.03	-0.01
HIST	-0.61	-0.32	-0.01	-0.31

947
948
949

950

952 **Figure 1** Prediction skill: Anomaly correlation coefficients (ACC, adimensional, bars) and root mean
953 square errors (RMSE in mm/day, stems) scores for Sahel rainfall at different lead times (1 to 4, 2 to 5,
954 3 to 6, 4 to 7, 5 to 9, 6 to 9 and 7 to 10: seven first columns, respectively) in the decadal hindcast and
955 in the historical uninitialized simulation (last column) for raw (dark blue) and de-trended data (light
956 blue) in the 1914-2004 period. The detrended scores are calculated over the same time series as the
957 raw ones except that the linear trends are previously removed. Note that the detrended time series
958 cannot be calculated in real-time prediction. Solid lines indicate the ACC skill of persistence for raw
959 (dark blue) and detrended data (light blue). Persistence is calculated as the average over the 4 years
960 preceding the model initialization. The dot-dashed line show the threshold to reject the null hypothesis
961 that the correlations come from chance (at the 5% level) for 4-year running mean filtered data.
962 Reference data for both metrics is CRUTS3.1 rainfall estimates.
963
964 **Figure 2** Time series of the anomalous indices with respect to the 1910-2005 period calculated for the
965 observations (HadISST1 and ERSST3vb are used for SST-based indices, labeled HAD and ERS
966 respectively and CRUTS3.1 for the rainfall index in black), historical simulation (dashed blue) and
967 separately for each 10 lead times of the decadal hindcast of: (a) GW (in K); (b) AMV (standardized);
968 (c) PDO (standardized); (d) decadal-SRI (in mm/day). For the definition of each index see details in
969 the text.
970
971 **Figure 3** Simulation skill: Anomaly correlation coefficients calculated between the observed
972 (HadISST1 dataset is used for reference) and the simulated indices of : (a) GW; (b) AMV; (c) PDO;
973 (d) decadal-SRI (in mm/day) shown in Fig. 2. Blue (red) bars mark positive (negative) correlations.
974 The black horizontal line shows the threshold above which the null hypothesis that the correlations
975 come from chance (at the 5% level) for the same type of filtered data should be rejected.
976
977 **Figure 4** GW regression patterns: Regression of the SST (K per standard deviation of the GW index,
978 shown over ocean areas) and precipitation (mm/day per standard deviation of the GW index, shown
979 only over land areas) fields onto the GW indices shown in Fig. 2 for: a) and b) the observations
980 (HadISST1 and ERSST3vb datasets are used for SSTs, labeled HAD and ERS respectively, and
981 CRUTS3.1 is used for rainfall in both plots); c) to l) the 10 lead times of the decadal hincast; and (m)
982 the historical experiment. Areas statistically non-significant are marked with dots (at the 5% level).
983
984 **Figure 5** AMV regression patterns: same as fig. 4 but using the AMV indices.
985
986 **Figure 6** IPO regression patterns: same as fig. 4 but using the IPO indices.
987
988
989 **Figure 7** Relation between the multi-linear regression coefficients and the rainfall and SST patterns:
990 Scatter plot between (a) α coefficient and the average rainfall regression of the GW pattern over the
991 Sahel in Fig. 4; (b) β coefficient and the average rainfall regression of the AMV pattern over the Sahel
992 in Fig. 5; (c) γ coefficient and the average rainfall regression of the IPO pattern over the Sahel in Fig.
993 6; (d) α coefficient and the tropical average SSTs (between 15ºS and 15ºN) in the GW patterns in Fig.
994 4; (e) β coefficient and the tropical North Atlantic SST gradient in the eastern part of the basin
995 (calculated as the difference between the SST average in the region 30ºW-18ºW, 16ºN-30ºN and in
996 the region 30ºW-18ºW, Eq-14ºN) in the AMV patterns in Fig. 5 ; (f) γ coefficient and the tropical
997 Pacific average SSTs (in the region 180ºW-95ºW 15ºS-15ºN) in the GW patterns in Fig. 6. In blue we
998 show the regression line and correlation coefficient using the 10 lead times of the decadal hindcast

999 and the historical experiments. In gray we show the regression line and correlation coefficient
1000 obtained without taking into account the historical simulation. Units for rainfall regression averages
1001 and multi-linear coefficients (SST regression averages) are mm/day (K) per standard deviation of the
1002 corresponding index

1003
1004 **Figure 8** Explained Variance AND SRI ACC scores: (a) Variance of the decadal-SRI (in mm^2day^{-2})
1005 and its partition into five components following the multi-linear regression analysis, which correspond
1006 to: GW, AMV, IPO, AMV-IPO covariance and residual of the fit (labeled as EPS). (b) Anomaly
1007 correlation coefficient (ACC) of the simulated decadal-SRI for each of the 10 lead times in the
1008 decadal hindcast and the historical experiment (blue dashed bars) and its decomposition into four
1009 terms following the multi-linear regression analysis, which are due to: GW, AMV, IPO and the
1010 residual of the fit (labeled as EPS). Positive and negative contributions are shown separately as
1011 stacked bars. The subtraction of the total stack positive bar minus the stack negative bar provides the
1012 ACC scores.

1013
1014
1015
1016 **Figure 9** Potential simulation skill: Anomaly correlation coefficient (ACC) between the decadal-SRI
1017 from observations and a synthetic one obtained using the GW and AMV from simulations and the α_k
1018 and β_k coefficients from the multi-linear regression analysis applied to observations (blue dashed
1019 bars). Decomposition of such ACC into the two terms (GW and AMV). Positive and negative
1020 contributions are shown separately as stacked bars. The subtraction of the total stack positive bar
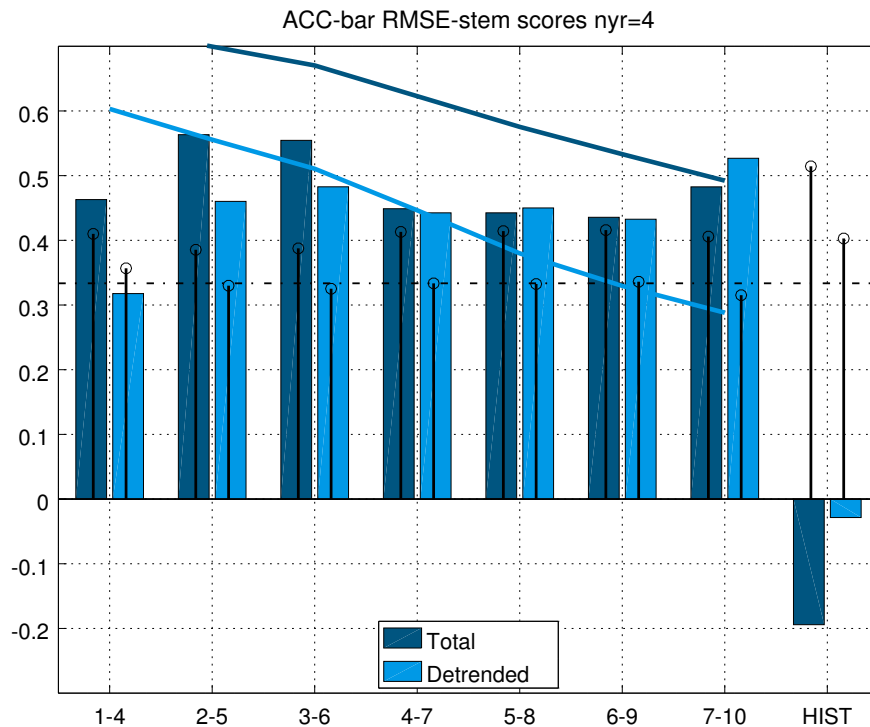1021 minus the stack negative bar provides the ACC scores.

Fig. 1: Prediction skill: Anomaly correlation coefficients (ACC, adimensional, bars) and root mean square errors (RMSE in mm/day, stems) scores for Sahel rainfall at different lead times (1 to 4, 2 to 5, 3 to 6, 4 to 7, 5 to 9, 6 to 9 and 7 to 10: seven first columns, respectively) in the decadal hindcast and in the historical uninitialized simulation (last column) for raw (dark blue) and de-trended data (light blue) in the 1914-2004 period. The detrended scores are calculated over the same time series as the raw ones except that the linear trends are previously removed. Note that the detrended time series cannot be calculated in real-time prediction. Solid lines indicate the ACC skill of persistence for raw (dark blue) and detrended data (light blue). Persistence is calculated as the average over the 4 years preceding the model initialization. The dot-dashed line show the threshold to reject the null hypothesis that the correlations come from chance (at the 5% level) for 4-year running mean filtered data. Reference data for both metrics is CRUTS3.1 rainfall estimates.
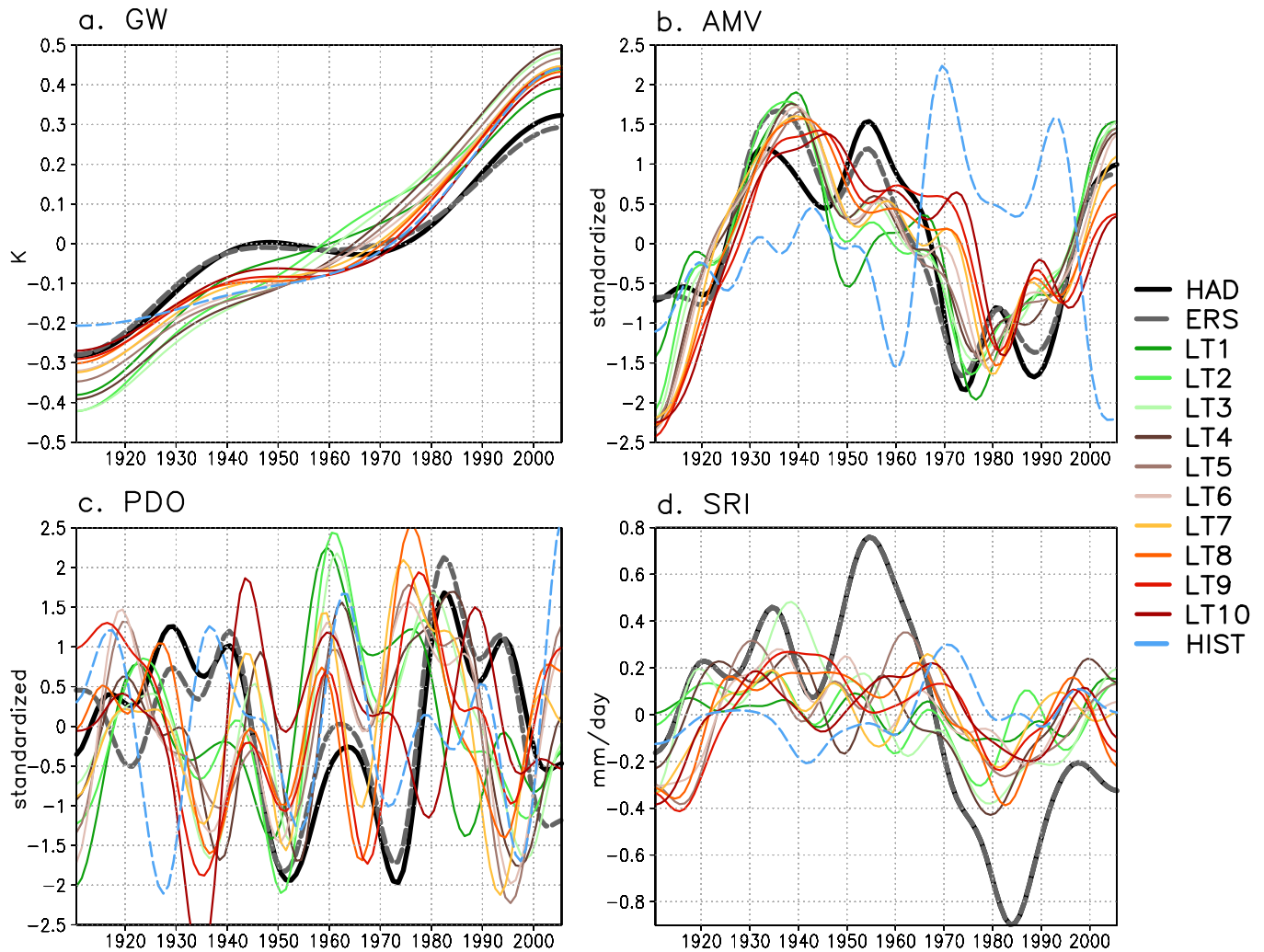
Fig. 2: Time series of the anomalous indices with respect to the 1910-2005 period calculated for the observations (HadISST1 and ERSST3vb are used for SST-based indices, labeled HAD and ERS respectively and CRUTS3.1 for the rainfall index in black), historical simulation (dashed blue) and separately for each 10 lead times of the decadal hindcast of: (a) GW (in K); (b) AMV (standardized); (c) PDO (standardized); (d) decadal-SRI (in mm/day). For the definition of each index see details in the text.
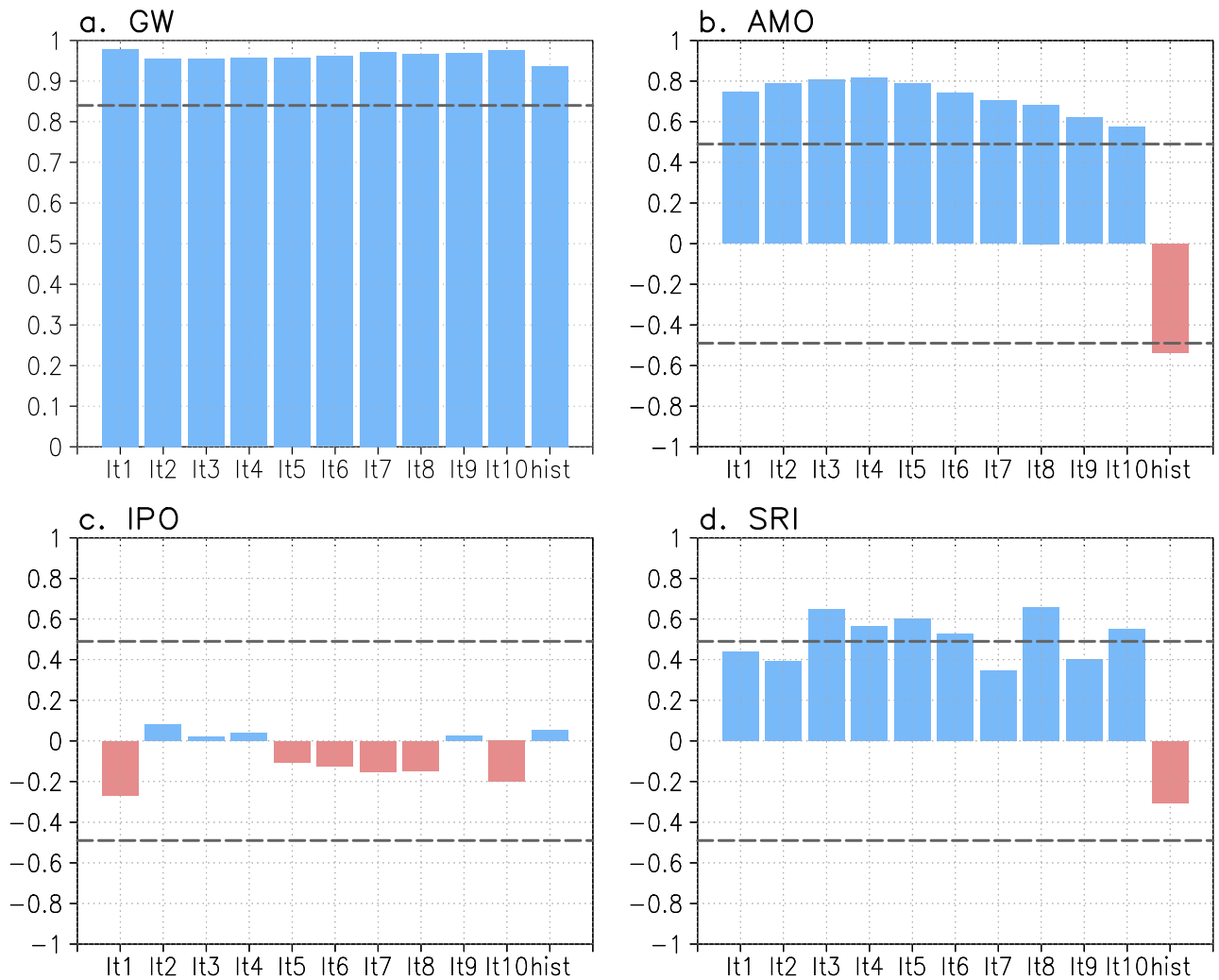
Fig. 3: Simulation skill: Anomaly correlation coefficients calculated between the observed (HadISST1 dataset is used for reference) and the simulated indices of : (a) GW; (b) AMV; (c) PDO; (d) decadal-SRI (in mm/day) shown in Fig. 2. Blue (red) bars mark positive (negative) correlations. The black horizontal line shows the threshold above which the null hypothesis that the correlations come from chance (at the 5% level) for the same type of filtered data should be rejected.
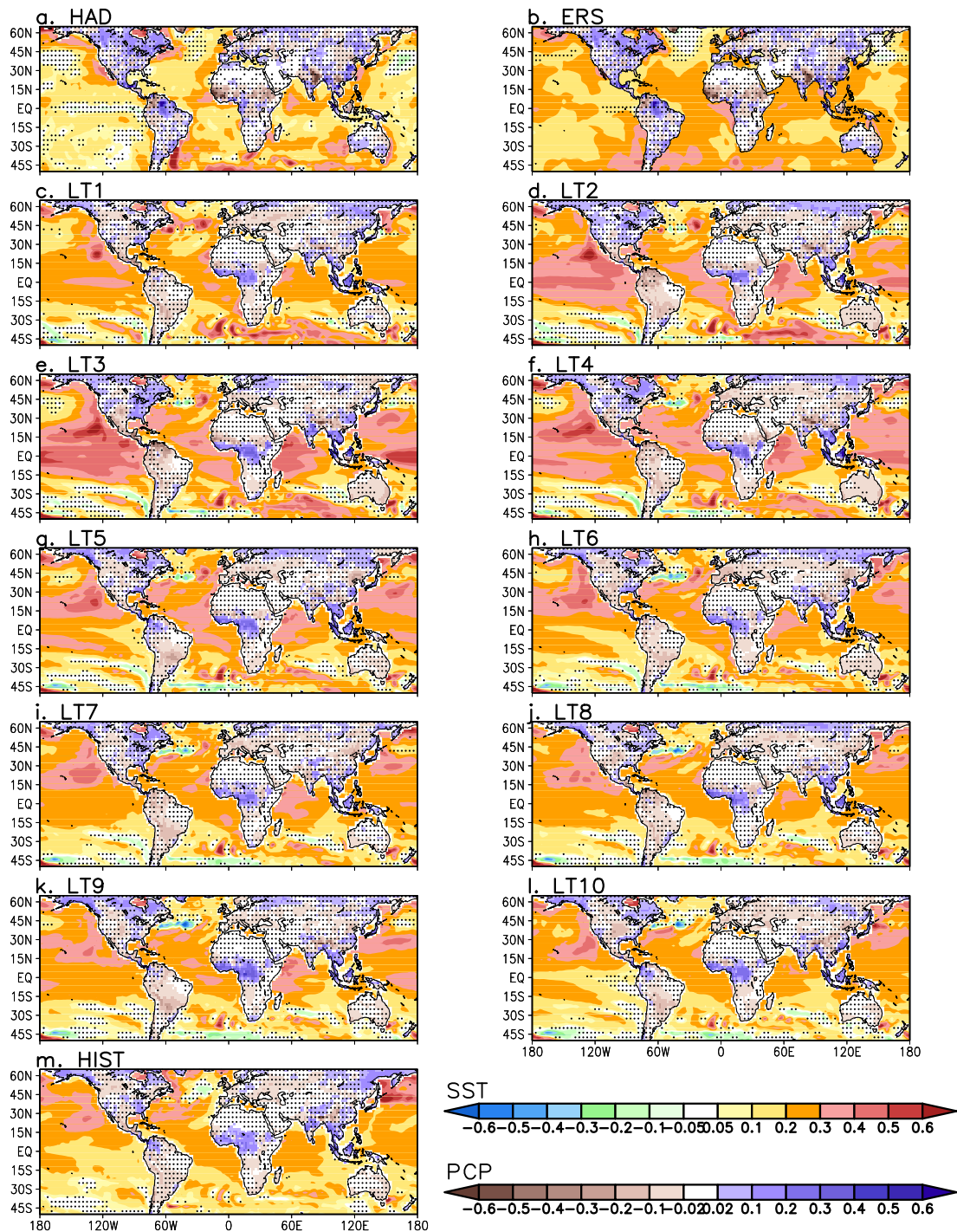
Fig. 4: GW regression patterns: Regression of the SST (K per standard deviation of the GW index, shown over ocean areas) and precipitation (mm/day per standard deviation of the GW index, shown only over land areas) fields onto the GW indices shown in Fig. 2 for: a) and b) the observations (HadISST1 and ERSST3vb datasets are used for SSTs, labeled HAD and ERS respectively, and CRUTS3.1 is used for rainfall in both plots); c) to l) the 10 lead times of the decadal hincasts; and (m) the historical experiment. Areas statistically non-significant are marked with dots (at the 5% level).
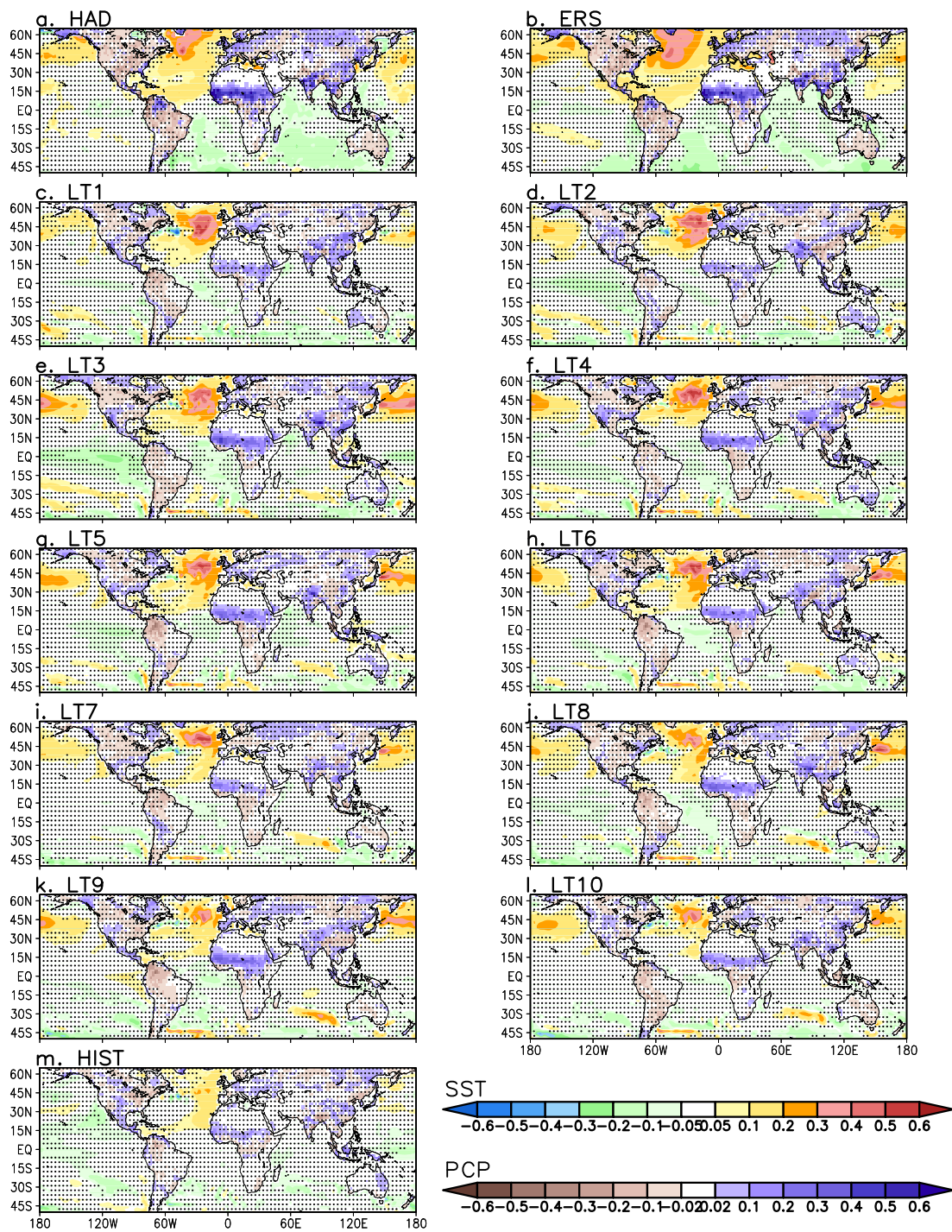
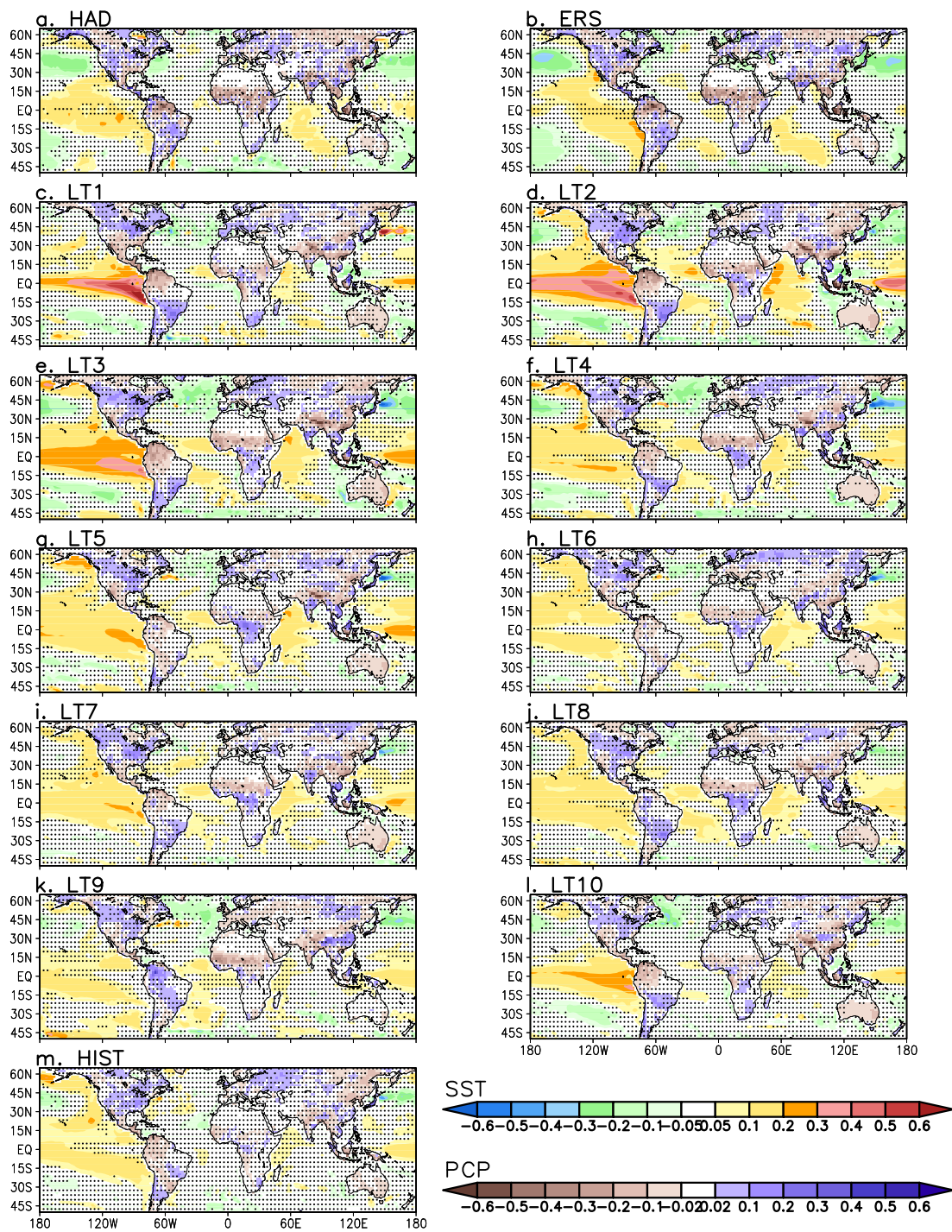Fig. 5: AMV regression patterns: same as fig. 4 but using the AMV indices

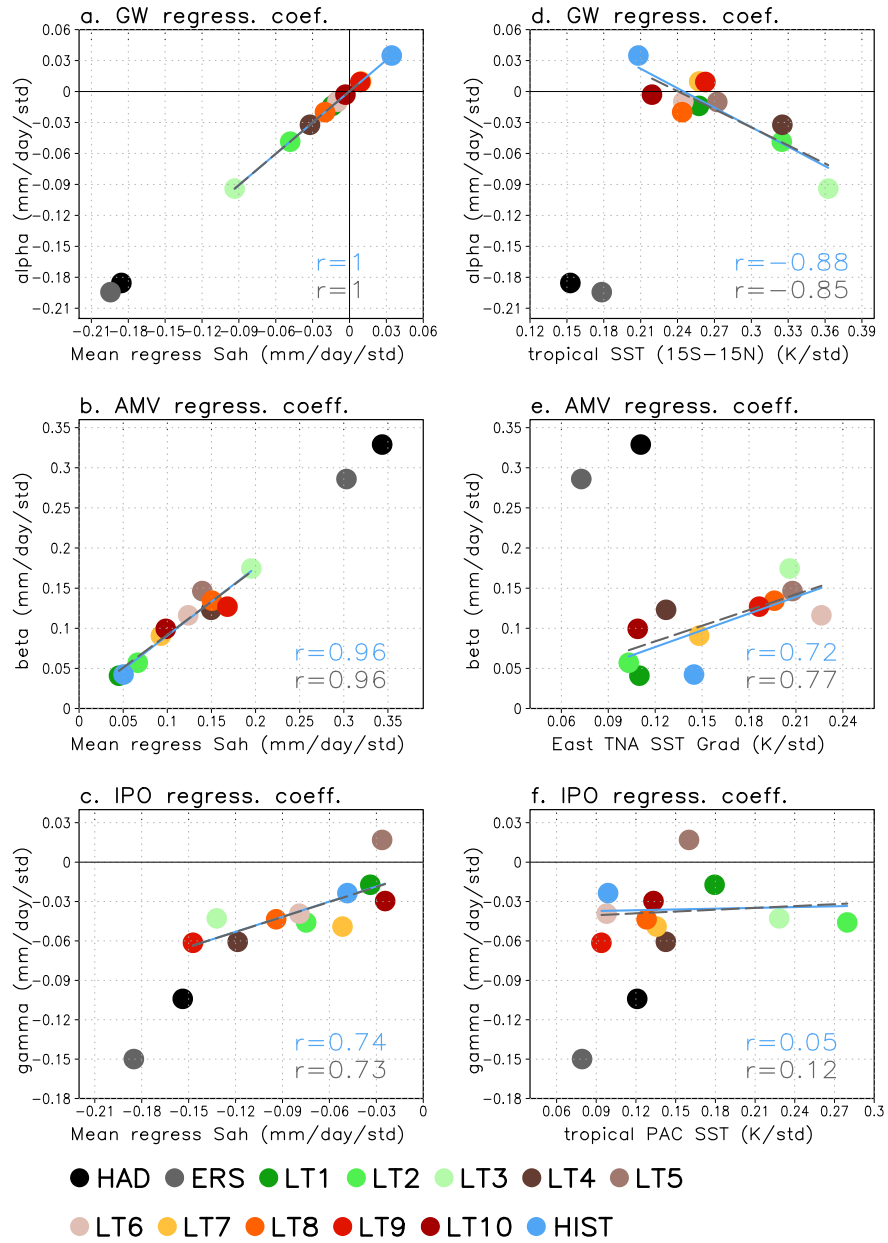Fig. 6: IPO regression patterns: same as fig. 4 but using the IPO indices

Fig. 7: Relation between the multi-linear regression coefficients and the rainfall and SST patterns: Scatter plot between (a) α coefficient and the average rainfall regression of the GW pattern over the Sahel in Fig. 4; (b) β coefficient and the average rainfall regression of the AMV pattern over the Sahel in Fig. 5; (c) γ coefficient and the average rainfall regression of the IPO pattern over the Sahel in Fig. 6; (d) α coefficient and the tropical average SSTs (between 15°S and 15°N) in the GW patterns in Fig. 4; (e) β coefficient and the tropical North Atlantic SST gradient in the eastern part of the basin (calculated as the difference between the SST average in the region 30°W-18°W, 16°N-30°N and in the region 30°W-18°W, Eq-14°N) in the AMV patterns in Fig. 5; (f) γ coefficient and the tropical Pacific average SSTs (in the region 180°W-95°W 15°S-15°N) in the GW patterns in Fig. 6. In blue we show the regression line and correlation coefficient using the 10 lead times of the decadal hindcast and the historical experiments. In gray we show the regression line and correlation coefficient obtained without taking into account the historical simulation. Units for rainfall regression averages and multi-linear coefficients (SST regression averages) are mm/day (K) per standard deviation of the corresponding index
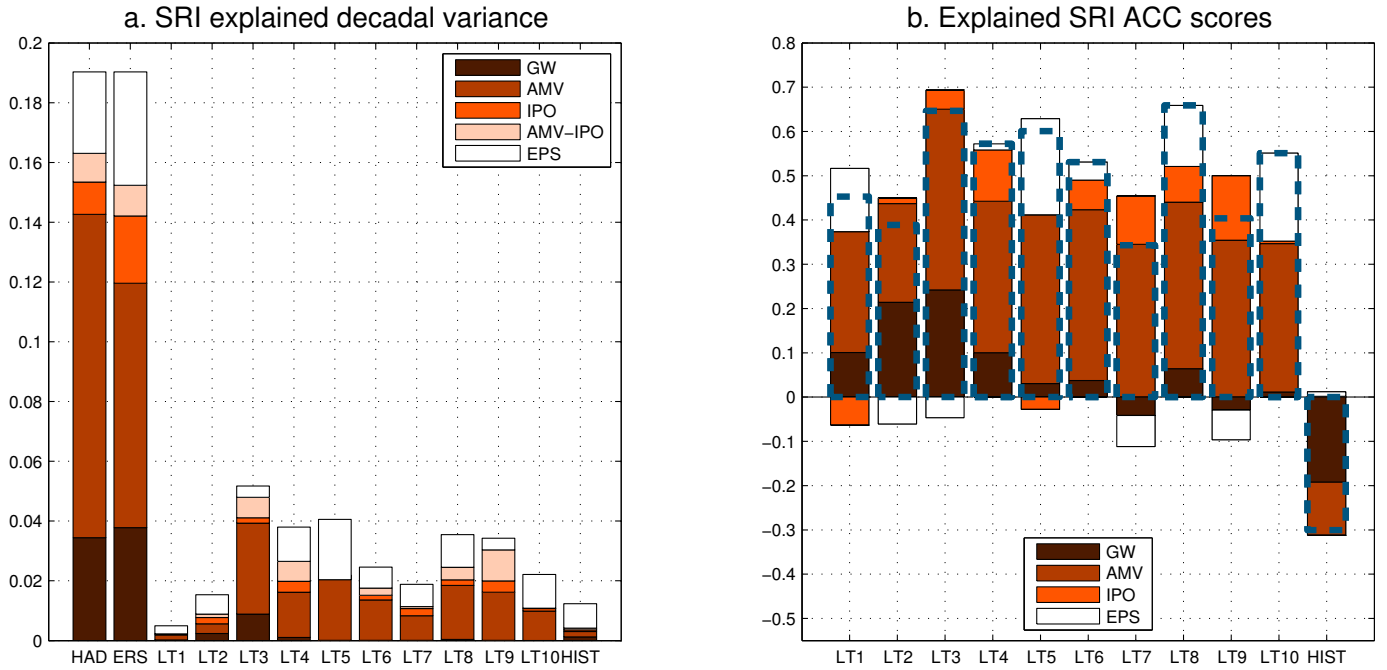
Fig. 8: Explained Variance AND SRI ACC scores (a) Variance of the decadal-SRI (in mm²day⁻²) and its partition into five components following the multi-linear regression analysis, which correspond to: GW, AMV, IPO, AMV-IPO covariance and residual of the fit (labeled as EPS). (b) Anomaly correlation coefficient (ACC) of the simulated decadal-SRI for each of the 10 lead times in the decadal hindcast and the historical experiment (blue dashed bars) and its decomposition into four terms following the multi-linear regression analysis, which are due to: GW, AMV, IPO and the residual of the fit (labeled as EPS). Positive and negative contributions are shown separately as stacked bars. The subtraction of the total stack positive bar minus the stack negative bar provides the ACC scores.
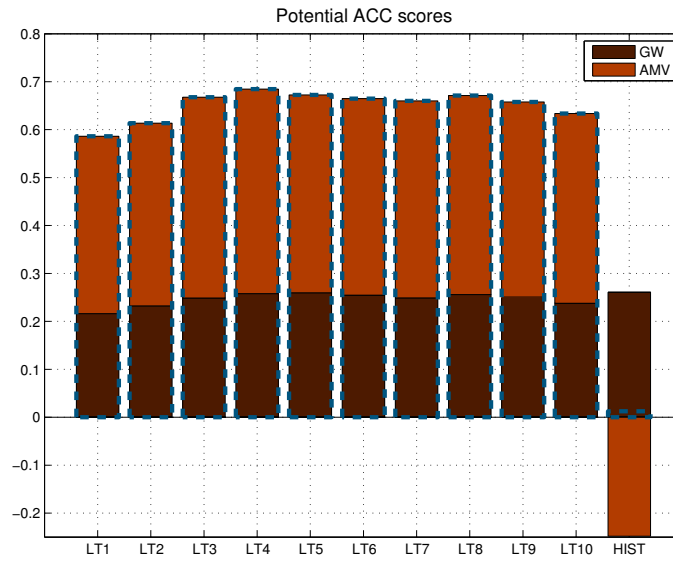
Fig. 9: Potential simulation skill: Anomaly correlation coefficient (ACC) between the decadal-SRI from observations and a synthetic one obtained using the GW and AMV from simulations and the α and β coefficients from the multi-linear regression analysis applied to observations (blue dashed bars). Decomposition of such ACC into the two terms (GW and AMV). Positive and negative contributions are shown separately as stacked bars. The subtraction of the total stack positive bar minus the stack negative bar provides the ACC scores.