Understanding Rainfall Prediction Skill Over the Sahel in NMME Seasonal Forecast

Verónica Martín-Gómez (vero.martin.gomez@gmail.com)
Universidad Complutense de Madrid https://orcid.org/0000-0002-7855-4893
Elsa Mohino
Complutense University of Madrid: Universidad Complutense de Madrid
Belén Rodriguez - Fonseca
Complutense University of Madrid: Universidad Complutense de Madrid
Emilia Sanchez - Gomez
CERFACS

Research Article

Keywords: teleconnections, (SSTa), (NMME), Sahel teleconnection, Mediterranean, Sahelian rainfall anomalies

Posted Date: October 22nd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-995460/v1

Version of Record: A version of this preprint was published at Climate Dynamics on April 5th, 2022. See the published version at https://doi.org/10.1007/s00382-022-06263-8.

Understanding rainfall prediction skill over the Sahel in NMME seasonal forecast

Verónica Martín – Gómez[1,2], Elsa Mohino[1], Belén Rodriguez – Fonseca[1,2], Emilia Sanchez – Gómez[3]

(1) Departamento de Física de la Tierra y Astrofísica, Universidad Complutense de Madrid, Spain.

(2) Instituto de Geociencias IGEO – CSIC, Madrid, Spain.

(3) CERFACS, CECI, Toulouse, France.

[1]Corresponding author: vero.martin.gomez@gmail.com

Abstract

Sahelian rainfall presents large interannual variability which is partly controlled by the sea surface temperature anomalies (SSTa) over the eastern Mediterranean, equatorial Pacific and Atlantic oceans, making seasonal prediction of rainfall changes in Sahel potentially possible. However, it is not clear whether seasonal forecast models present skill to predict the Sahelian rainfall anomalies. Here, we consider the set of models from the North American Multi-model ensemble (NMME) and analyze their skill in predicting the Sahelian precipitation and address the sources of this skill.

Results show that though the skill in predicting the Sahelian rainfall is generally low, it and can be mostly explained by a combination of how well models predict the SSTa in the Mediterranean and in the equatorial Pacific regions, and how well they simulate the teleconnections of these SSTa with Sahelian rainfall. Our results suggest that Sahelian rainfall skill is improved for those models in which the Pacific SST - Sahel rainfall teleconnection is correctly simulated. On the other hand, models present a good ability to reproduce the sign of the Mediterranean SSTa – Sahel teleconnection, albeit with underestimated amplitude due to an underestimation of the variance of the SSTa over this oceanic region. However, they fail to correctly predict the SSTa over this basin, which is the main reason for the poor Sahel rainfall skill in models. Therefore, results suggest models need to improve their ability to reproduce the variability of the SSTa over the Mediterranean as well as the teleconnections of Sahelian rainfall with Pacific and Mediterranean SSTa.

1 Introduction

The Sahel is a semiarid region located in the westernmost part of the tropical African continent, between the south of the Sahara desert and the humid savanna (e.g., Nicholson 2013). Most of the population in this region reside in rural areas and their sustenance is mainly based on the development of agriculture and pastoralism activities, sectors vulnerable to rainfall variability (Mortimore et al. 2001; Kanji et al.

2006). Therefore, the understanding of the changes in rainfall as well as having good predictions thereof are crucial for this region.

Rainfall over the Sahel presents a strong meridional gradient, with annual rainfall mean values of roughly 550 mm over its southern part and those on the order of 150 mm over its northern part (Nicholson 2013). Throughout the year, rainfall mainly occurs during the summer months (July - August - September, JAS), in association with the West African Monsoon (WAM), with a maximum developed in August. Occasional rainfall might also be observed during winter time, though related to extratropical systems (Nicholson 2013).

Boreal summer seasonal precipitation over the Sahel presents large variability from interannual to interdecadal timescales (Kitoh et al. 2020 and references therein). Previous studies have shown that the sea surface temperature anomalies (SSTa) over different basins can impact rainfall interannual variability over the Sahel (Rodriguez-Fonseca et al. 2011, 2015). Analysis of observations and model experiments have shown that warm SST anomalies in the eastern equatorial Atlantic reduce the land-atmosphere temperature and surface pressure gradients and tend to be associated with an anomalous dipole of rainfall with positive values over the Guinea Gulf and negative ones over the Sahel (e.g. Vizy and Cook 2002; Losada et al. 2010; Polo et al. 2008). Warm anomalies in the equatorial Pacific tend to produce a stabilization of the air column and subsidence over West Africa, weakening the monsoon and reducing precipitation seasonal amounts and the occurrence of heavy precipitation events (Rowell 2001; Janicot 2000; Mohino et al. 2011; Parhi et al. 2016; Joly and Voldoire 2009; Diakhaté et al. 2019). Cold SST anomalies over the equatorial Pacific tend to promote the opposite effect. Out of the tropics, warm Mediterranean SST anomalies enhance local evaporation leading, through southerly moisture advection, to an increase of low-level moisture convergence and destabilization over the Sahel. This strengthens the monsoon and increases seasonal rainfall averages there. The contrary occurs for cold SST anomalies over the Mediterranean (Rowell 2003; Fontaine et al. 2010, 2011; Gaetani et al. 2010).

Note that the above rainfall anomalies associated with tropical Atlantic, Pacific and extratropical Mediterranean SSTs represent the direct response to each isolated SST forcing. Nevertheless, oceans are interconnected and pantropical interactions have been detected during certain decades (Cai et al. 2019; Wang 2019; Kitoh et al. 2020). In particular, it is know that, from the 1970's, the Atlantic and Pacific Niños appear in opposition of phases in summer (Rodriguez-Fonseca et al. 2009). As warm SSTs over both, the equatorial Atlantic and Pacific oceans, decrease rainfall, a counteracting effect over the Sahel appears under that "opposition of phase" configuration (Polo et al. 2008, Losada et al. 2012; Suarez-Moreno et al. 2018).

In addition, non stationarities are found in the impact of SSTs on Sahelian rainfall (Rodriguez-Fonseca et al. 2011, 2015, 2016, Suarez-Moreno et al. 2018). During the period between the 1950 and 1980s no impact from the Mediterranean is detected in the observations, being significant in the previous and later decades (Suárez-Moreno et al. 2018). In turn, in the last decades, the relation with the tropical Pacific is strong, while that with the Atlantic appears absent. This lack of connection may be the result of the above mentioned counteracting effect with the Pacific (Losada et al. 2012). Thus, in recent decades, Pacific and Mediterranean seem to dominate the interannual variability of Sahelian rainfall.

The slow varying SST and their influence over continental areas at seasonal timescales constitutes the physical basis of the seasonal predictions. Similarly to other regions in the world, seasonal forecasts over the Sahel were initially developed from statistical methods based on empirical teleconnections between SST anomaly patterns and continental anomalies (Folland et al. 1991). In a second phase, seasonal predictions were performed by using atmospheric general circulation models (AGCM) forced by observed SSTs (Lau et al. 2009, Ndiaye et al. 2009). Nevertheless, the dynamical predictions presented limited skill in the Sahel and often, complementary approaches relating both dynamical and statistical predictions were developed (Garric et al. 2002, Ndiaye et al. 2009). In a third phase, and together with the development of coupled models, seasonal forecasts were carried out with coupled systems (atmosphere model coupled to an ocean model). Today, seasonal predictions based on coupled systems are made operational and delivered by the main centers of operational forecasting around the world. Moreover, in the late 1990s, several international initiatives proposed the development of multi-model seasonal climate predictions systems (i.e. DEMETER, ENSEMBLES, SINTEX), with the aim of joining efforts and comparing seasonal predictions performed with different coupled models. Currently, multi-model ensemble forecasting is a mainstream method used for seasonal predictions. Several studies based on these multi-model ensembles have revealed that the multi-model predictions present higher skill than individual systems (Palmer et al. 2004, Doblas-Reyes et al. 2009), as they also account for model uncertainty. Focusing on the Sahel, Batté and Déqué (2011) pointed out that the multi-model ENSEMBLES seasonal predictions enhance the skill of the western Africa precipitation, by reducing the skill-spread ratio. Rodrigues et al. (2014) conclude that the state-of-the-art EUROSIP (European Seasonal to Interannual Prediction) and NMME (North American Multi-model Ensemble) seasonal predictions are reliable in predicting the interannual variations of the Sahel precipitation regimes. Recently, Giannini et al. (2020) analyzed the precipitation skill of five seasonal forecast models from the NMME and, based on the multi-model mean, showed that precipitation anomalies during the monsoon season can be predicted even with lead times as far as 3-4 months. They also found that such skill comes mainly from ENSO and the North Atlantic sea surface temperatures. Nevertheless, the rest of the models in the NMME were not analyzed, nor were other SSTa regions which could contribute to skill addressed.

Here, we extend the number of NMME models and analyze the precipitation prediction skill seeking to understand where the skill or lack thereof comes from. Our interest is on the interannual rainfall variability associated with oceanic forcing. The starting hypothesis is that much of the precipitation prediction skill in models should come from the teleconnections with the SSTa in different regions. Therefore, we first analyze the main sources of predictability for the Sahel rainfall in observations, and then we evaluate whether these potential predictors obtained from observations are also sources of predictability in models. The final aim is to evaluate the skill of the seasonal forecast models to predict rainfall over the Sahel, analyzing whether the models are able to reproduce the SST over the observed potential predictor regions and the sign and amplitude of the SST-Sahel teleconnections.

Data and Methodology

2.1 Data

2.1.1 Observational data

Observational data is considered to contrast results from models and analyze their seasonal forecasting skill. We employ the monthly precipitation values from GPCPv2.3 with a spatial resolution of 2.5º x 2.5º. They are provided by the *NOAA/OAR/ESRL PSD, Boulder, Colorado, USA,* and are available at the website*: https://www.esrl.noaa.gov/psd/* (Adler et al. 2003). Additionally, we also consider the monthly sea surface temperatures (SST) data from HadISSTv1.1 with 1º x 1º of resolution (Rayner et al., 2003). The study is focused on July - August – September (JAS), which is the season when the monsoon takes place (Rodriguez – Fonseca et al. 2015; Nicholson 2013; Thorncroft et al. 2011). In order to check the sensitivity of the results to the choice of the observational dataset, the analysis is also performed using SST from ERSSTv5 (with a resolution of 2º x 2º, Huang et al. 2017) and precipitation from CRU TS 4.03 (Harris et al. 2020). CRU TS is derived by the interpolation of monthly climate anomalies from extensive networks of weather station observations. However, GPCP considers data from rain gauge stations, satellites, and sounding observations from 1979 to present. Regarding SST, HadISST uses reduced space optimal interpolation applied to SSTs from the Marine Data Bank (mainly ship tracks) and the International Comprehensive Ocean–Atmosphere Dataset (ICOADS) through 1981 and a blend of in-situ and adjusted satellite-derived SSTs for 1982-onwards. ERSSTv5 uses new data sets from ICOADS Release 3.0 SST, Argo floats above 5 meters and Hadley Centre Ice-SST version 2 (HadISST2) ice concentration. Given that conclusions are not altered, in this work we only show the results obtained by contrasting model simulations with observations from GPCPv2.3 and HadISSTv1.1. Additionally, they are not dependent on the actual choice of the peak season, as results for August-September are consistent with the ones presented for JAS in the paper (not shown).

Finally, we also consider the monthly mean sea level pressure (MSLP) and horizontal winds at 850hPa and 200hPa from ERA5 reanalysis (Hersbach et al., 2020) in order to analyze the spatial teleconnection patterns of eMED and Niño3 with the Sahel.

2.1.2 NMME models

We consider the monthly hindcast of SST and precipitation from a set of fifteen seasonal forecasting models belonging to the North American Multi-model ensemble (NMME). Table 1 shows a summary of the models used. Though models have different native spatial resolution, their output are monthly forecasts with a similar resolution of 1º x 1º. They are available at the web page http://iridl.ldeo.columbia.edu/SOURCES/.Models/.NMME/ (Kirtman et al. 2014). The common period for which models have complete predictions is 1982 - 2010 and, therefore, this will be our period of study. Finally, in order to analyze the prediction skill, six were considered for all the models, from 1st July to 1st February.

The analysis of the skill for individual models is performed by considering the ensemble mean, constructed by averaging 4 different simulations of each model. We also analyze the skill of the multi-model mean, constructed following the methodology of pooling of first n-members of each ensemble (Hemri et al. 2020) to construct the multi-model mean of all the NMME models. The parameter n is selected as 4, since this is the maximum number of ensemble members in one of the models (NASA-GEOSS2S). The results are qualitatively in agreement with a different definition of the multi-model mean based on an average of each model mean in which all the available members (not just 4) are taken into account (not shown).

Finally, we also analyze the monthly hindcasts of mean sea level pressure and horizontal winds at 200mb and 850mb. These fields are considered in order to evaluate the atmospheric patterns of the teleconnections in models. Given that only the models CanCM4i, CMC1-CanCM3, CMC2-CanCM4 and GEM-NEMO have available atmospheric data, the analysis of the teleconnection mechanisms will be evaluated for these 4 models and, for the sake of briefness, it will only be shown for the average of them.

2.2 Methodology

We begin by revisiting the main SST sources of predictability for Sahel precipitation in observations for the period of study. To do so, we obtain summer (JAS) seasonal anomalies of SSTs and precipitation by subtracting the seasonal mean. Seasonal forecasts aim at providing predictions of climate anomalies for the forthcoming season (Balmaseda et al. 2009), which should be taken into account when assessing their skill. In this work, we use a 21-yr window for the analysis and furthermore remove the trends in all variables prior to any other calculation. This is particularly relevant in the case of the Sahel rainfall,

which shows strong climate variability at decadal timescales (Kitoh et al. 2020). We then calculate the regression map of SST anomalies worldwide onto the Sahel precipitation index (Fig. 1). The precipitation index is defined as the seasonal average of the precipitation anomalies over the Sahel region (see domain on Table 2).

Secondly, we define the indices that represent the temporal evolution of the SST anomalies over the regions associated with the main observed predictors. As it will be shown in section 3.1, the potential sources of interannual predictability for rainfall over Sahel in JAS during the period of study are the eastern Mediterranean (eMED) and the eastern equatorial Pacific (Nino3) (in agreement with Suárez - Moreno et al. (2018)). These indices are computed as the average of JAS seasonal SST anomalies in the appropriate regions (Table 2). Note that the indices so defined do not show a linear trend, as this was removed in the definition of the SST anomalies to avoid introducing long-term variability in our analysis. Additionally, in order to analyze the spatial atmospheric teleconnection patterns, we also compute the (1) correlation maps of the eMED index onto the mean sea level pressure anomalies and winds anomalies at 850hPa, and (2) the correlation maps of Niño3 index and the anomalous velocity potential difference between 200hPa and 850hPa levels ($Vpot_{200} - Vpot_{850}$) as a way of illustrating the baroclinic atmospheric response associated with ENSO forcing. Velocity potential ($Vpot$) is obtained by solving the Poisson's equation: $\nabla \cdot V = \Delta Vpot$, where V represent the horizontal winds.

Thirdly, to evaluate the contribution of the eMED and Niño3 signals to precipitation variability over Sahel, the precipitation index is fitted with a multilinear regression model as follows:

$$PCP_{reg} = \alpha \cdot eMED_{index} + \beta \cdot Niño3_{index} + \varepsilon \quad \text{(eq. 1)}$$

where PCP_{reg} is the precipitation index obtained from multiple regression analysis, α *and* β represent the coefficients of multilinear regression for eMED and Niño3, respectively, $eMED_{index}$ and $Nino3_{index}$ are the standardized indices associated with the eastern Mediterranean and equatorial Pacific El Niño, respectively. Finally, ϵ represents the residual fitting. The statistical significance of the multiple linear regression is assessed by considering a F-test with a 95% confidence level.

Using this fit, the total variance of precipitation can be decomposed into the following components (2):

$$Var(PCP_{reg}) = \alpha^2 + \beta^2 + 2 \cdot \alpha \cdot \beta \cdot cov(eMED, Niño3) + var(\varepsilon) \quad \text{(eq. 2)}$$

where α^2, β^2 represent the part of the total precipitation variance which is explained by eMED and Niño3, respectively, and the term $2 \cdot \alpha \cdot \beta \cdot cov(eMED, Niño3)$ stands for the covariance between eMED and Niño3.

Fourthly, we analyze the precipitation prediction skill by means of the anomaly correlation coefficient (ACC), the root mean squared error (RMSE) and the mean squared error skill score (MSESS). ACC is computed as the correlation between the observed and the modeled indices. We consider that a model presents skill in terms of ACC when the ACC value is statistically significant at 95% confidence level in two - tailed t – test with an effective number of degrees of freedom (Mitchell et al. 1966; Bretherton et al. 1999). ACC values are computed considering the ensemble mean of each model.

As it was aforementioned, our starting hypothesis is that much of the skill for predicting Sahel precipitation should come from it teleconnections with the SST anomalies in different regions worldwide. Therefore, to investigate where the precipitation ACC skill scores comes from, we also evaluate the models' skill in predicting the SST anomalies over the main sources of predictability and their ability in reproducing the teleconnections with precipitation. To evaluate the teleconnections skill, we compare the regression and correlation coefficients between each predictor and Sahel precipitation in the models with the one obtained from observations. We also compare the physical mechanisms for the teleconnections in models and observations by analyzing the atmospheric patterns that each one of the potential predictors induce. Given the availability of atmospheric data in models, this can be only done considering 4 of the 15 models.

To evaluate the contribution of the different SSTa signals on precipitation ACC skill score in models, we use the same multilinear regression analysis than in observations. Considering this analysis, the contribution of each predictor to the precipitation ACC skill score in models can be estimated in terms of the multilinear regression coefficients and the correlations between the observed precipitation index (PCP_{obs}) and the simulated oceanic indices (as in Mohino et al., 2016):

$$\text{ACC}=\rho(PCP_{obs}, PCP_{nmme}) = \frac{\alpha}{\sqrt{var(PCP_{nmme})}}\rho(PCP_{obs}, eMED_{nmme}) +$$

$$\frac{\beta}{\sqrt{var(PCP_{nmme})}}\rho(PCP_{obs}, Niño3_{nmme}) + \frac{\sqrt{var(\varepsilon_{nmme})}}{\sqrt{var(PCP_{nmme})}}\rho(PCP_{obs}, \varepsilon_{nmme}) \qquad (eq.5)$$

where ρ represents the correlation coefficient, PCPobs the precipitation index from observations, $PCPnmme$ the precipitation index from NMME, $eMEDnmme, Niño3nmme$, the indices of eMED and Niño3 from NMME models, and ϵ the residual fitting from the multiple regression in NMME models. With this decomposition, the first and second terms in the right-hand side of the equation can be understood as the part of the precipitation ACC skill score explained by the eMED and Niño3 indices, respectively. The third term corresponds to the unexplained ACC skill score, which could be related to unaccounted sources.

Finally, as mentioned previously, forecast skill is also assessed considering the RMSE and MSESS. RMSE is defined as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_{i,nmme} - y_{i,obs}\right)^2} \qquad (eq.3)$$

where $y_{i,obs}$ is the observed value ($PCP_{i,obs}$, $eMED_{i,obs}$ or $Niño3_{i,obs}$) and $y_{i,nmme}$ is the forecasted one ($PCP_{i,nmme}$, $eMED_{i,nmme}$ or $Niño3_{i,nmme}$) (Déqué 2011). On the other hand, MSESS is defined as:

$$MSESS = 1 - \frac{MSE}{MSE_{clim}} \qquad (eq.4)$$

where $MSE = \frac{1}{n}\sum_{i=1}^{n}\left(y_{i,nmme} - y_{i,obs}\right)^2$ and $MSE_{clim} = \frac{1}{n}\sum_{i=1}^{n}\left(\overline{y_{obs}} - y_{i,obs}\right)^2 = var(Y_{obs})$, being $Y_{obs} = y_i,\ldots,y_n$ is the time series of the observed values and $\overline{y_{obs}}$ its mean. The maximum value of MSESS is 1 and occurs when the MSE=0, that is, when model gives a perfect forecast. MSESS = 0 takes place when $MSE=MSE_{clim}$, which implies that the model forecast skill in terms of MSE is equal to that one provided by a climatological forecast. Finally, negative MSESS values implies that model forecast skill in terms of MSE is worse than consider a climatological forecast (Murphy 1998; Déqué 2011).

3 Results

3.1 Observational analysis

In this section we briefly analyze the teleconnection of Sahel precipitation and SST during the period of study and estimate the percentage of the total precipitation variance explained by each predictor in observations. Figure 1 shows the regression map between the precipitation index over Sahel and worldwide SST anomalies. In the period under analysis, Sahel precipitation shows statistically significant positive regression coefficients with positive SST anomalies over the eastern Mediterranean (eMED) and negative ones over equatorial Pacific (Niño3), in accordance with previous studies (e.g. Janicot 2001; Rowell 2001, 2003; Fontaine et al. 2010, Fontaine et al. 2011; Gaetani et al. 2010; Diahaté et al. 2019). It also shows positive connections with North Atlantic and Pacific SSTs. However, these signals are centered in the northern hemisphere subtropical gyres, which suggests they could be due to Ekman induced transport in response to changes in the subtropical high pressure systems (i.e., Barrier et al. 2014), as a consequence of atmospheric teleconnections from El Niño and the Mediterranean. Particularly, the North Atlantic positive significant area (defined as the average in the appropriate box in Fig. 1) shows a statistically significant correlation of 0.6 with eMED index. Conversely, the Niño 3 and eMED indices show no statistically significant correlation among themselves (correlation value is -0.15) and are, therefore, selected as the two independent predictors for rainfall over Sahel in the rest of the study.

The teleconnection mechanisms through which eMED and Niño3 influence rainfall variability over Sahel are shown on Fig. 2(a) and (b), respectively. Focusing on the eMED teleconnection (Fig. 2(a)), an anomalous warming over the eastern Mediterranean induces an anomalous low level pressure over the eastern Sahara (in agreement with Gomara et al., 2018). Additionally, Fig. 2(a) shows an intensification of the southwesterly monsoonal flow due to the strengthening of the meridional MSLP gradient between the Gulf of Guinea and the Sahara. As a consequence, a warming over the eMED is related to an increase of precipitation over Sahel. These results are consistent with previous studies (Diakhaté et al. 2019 ; Gómara et al. 2018 ; Fontaine et al. 2010, 2011; Gaetani et al. 2010 ; Jung et al. 2006).

Regarding the influence of Niño3 over the Sahel (Fig. 2(b)), the anomalous warming over the equatorial Pacific generates an anomalous increase of the difference between the anomalous velocity potential at 200hPa and 850hPa (VPOT200/850), suggesting a weakening the monsoon and reducing convection and precipitation over Sahel (in agreement with Diakhaté et al. 2019; Gomara et al. 2018; Suárez-Moreno et al 2018; Mohino et al. 2011; Joly and Voldoire 2009).

The multiple regression analysis performed for Sahel rainfall considering only the eMED and Niño3 indices suggests that these two predictors can explain in total 58% of the total precipitation variance at interannual timescales and that the eMED is the most dominant influence (see Table 3). If the analysis is repeated with the eMED and Niño3 indices leading the JAS Sahel rainfall with lags from 0 to 6 months, the percentage of total precipitation variance explained by these two predictors decreases with the lag (see section 1 in additional material). On the other hand, the percentage explained by the residue (42%) represents the part of the total precipitation variability not explained by changes in the eMED and nino3 indices. We speculate that this could be partly driven by land-atmosphere and aerosol-radiative processes (e.g. Nicholson et al. 2013; Rodriguez-Fonseca et al. 2015).

3.2 Prediction skill for precipitation

Figure 3 shows the precipitation prediction ACC skill scores in JAS for each FST and model. Results show that models in general lack skill to predict rainfall over Sahel, although there are some of them which present statistically significant ACC values for specific FST. For example, GFDL-CM2p5-FLOR-A06 and GFDL-CM2p5-FLOR-B01 present skill for FST 1st July and 1st June, GFDL-CM2p1 for FST from 1st May to 1st March, GFDL-CM2p1-aer04 for FST= 1st June, 1st April and 1st February, NCEP-CFSv2 for FST= 1st June, 1st April and 1st March, and COLA-RSMAS-CCSM4 for FST=1st Feb. Note that even in these cases, ACC skill scores are low. Comparing results from the multi-model mean and the models, the former presents larger precipitation ACC values than most of the models for all the FST, suggesting that the pooling of models leads to a forecast skill greater than the majority of single model systems. Nonetheless, the multimodel mean only presents statistically significant ACC values

for FST=1st June and 1st Feb and is outperformed by some single models (e.g.GFDL-CM2p5-FLOR-A06, GFDL-CM2p5-FLOR-B01 for FST=1st June; and GFDL-CM2p1-aer04 and COLA-RSMAS-CCSM4 for FST= 1st Feb).

We also computed the RMSE of precipitation (Fig. 4(a)). Results show that most models present a RMSE smaller than 3 mm/day, with the best representation provided by GFDL-CM2p5-FLOR-A06 and NCEP-CFSv2 models for FST 1st July, GFDL-CM2p1 and NCEP-CFSv2 for FST=1st June, COLA-RSMAS-CCSM3 and Multimodel mean for FST 1st May, and COLA-RSMAS-CCSM3 for FST 1st April to 1st May, and GFDL-CM2p1 for FST=1st Feb. On the other hand, the model with highest RMSE is NASA-GEOSS2S for all the FST. Regarding the MSESS, Fig. 4(b) shows that most of models present negative MSESS, which suggest that models forecast skill in terms of MSE is, in general, worse than consider the climatological forecast. The model with worst MSESS values is NASA-GEOSS2S, in agreement with results from RMSE. Nonetheless, some models present positive MSESS for specific FST, such as GFDL-CM2p5-FLOR-A06 for FST=1st July, CMC2-CanCM4 for FST= 1st Feb, NCAR-CESM1 for FST=1st Feb and NCEP-CFSv2 for FST=1st June, suggesting that precipitation forecast given by these models at those FST is better, in terms of MSE, than the climatological forecast. Finally, results from the multimodel-mean of the precipitation index are one of the best in terms of RMSE and MSESS in comparison with the rest of the models (see Fig. 4(a) and (b)), although it MSESS remains negative.

On the basis of these results, we present 2 questions: Does the lack of precipitation skill in NMME arise from a wrong prediction of the SST over the eMED and equatorial Pacific? Or does the lack of precipitation skill in NMME arise from an incorrect prediction of the SST-Sahel rainfall teleconnection? In the next section we try to give answers to both questions and quantify the relative importance of both factors on the prediction precipitation skill over the Sahel.

3.3 Prediction skill for SSTs

3.3.1 Eastern Mediterranean (eMED)

Figure 5 shows the ACC values for eMED. Most models present statistically significant ACC values for FST= 1st July. However, only GFDL-CM2p5-FLOR-A06 and NCAR-CESM1 have statistical significant ACC values for earlier FST (FST= 1st June). Results are similar for the case of the multi-model mean, which loses its ability for FSTs before than 1st July.

Figure 6(a) shows that most of the models present RMSE smaller than 12ºC, with the best representation provided by GFDL-CM2p1-aer04. The model with higher RMSE is IRI-ECHAM4p5-DirectCoupled followed by IRI-ECHAM4p5-AnomalyCoupled, COLA-RSMAS-CCSM3 and COLA-RSMAS-

CCSM4. Regarding MSESS, Fig. 6(b) shows that all models present negative MSESS, suggesting that the eMED model forecast in terms of MSE is worse than the climatological forecast. In this case, the multimodel-mean does not present one of the best MSESS values. This is related to the existence of some models (IRI-ECHAM4p5-DirectCoupled, IRI-ECHAM4p5-AnomalyCoupled, COLA-RSMAS-CCSM3 and COLA-RSMAS-CCSM4) with very high RMSE values in comparison with the rest of the models (see Fig. 6(b)).

The observed positive connection between eMED and Sahel precipitation is reproduced by most models, although with an underestimation of the magnitude in both correlation value and multiple linear regression coefficient (Fig. 7, first and second columns respectively). The underestimation of regression values (α from eq. (1)) suggests a weaker sensitivity of Sahel rainfall to eMED anomalies in models. In addition to this lower sensitivity, models tend to show too low variance of the SST anomalies over the eMED in models, compared to observations (see section 2 in additional material), which could allow other sources of precipitation variability to become more dominant in models and further explain the underestimation in correlation values. The underestimation of the eMED - precipitation teleconnection is also evident in the case of the multi-model mean (Fig. 7(b)). Note that the multi-model mean is the average of the 15 models which show very poor ACC skill scores to predict eMED (Fig. 5). The lack of phasing of the eMED among the models reduces the eMED signal in the multi-model and therefore, weakens the eMED - precipitation teleconnection signal in this case. All these results suggest that the generalized lack of skill to predict the precipitation anomalies in models could be related not only to the lack of skill (in terms of ACC) to reproduce the SST anomalies in eMED (Fig. 5) but also to the underestimation of the of the eMED - precipitation teleconnection amplitude (Fig. 7(b)).

In order to assess how an improvement on eMED SST skill impacts the skill of its teleconnection with the Sahel, Fig. 8(a) and (b) show the scatter plots of precipitation prediction ACC skill scores vs. eMED prediction ACC skill scores considering all the models and the first two FST separately, 1st July and 1st June, respectively. Information regarding the rest of the FST can be found in Table 4 (or section 3 in additional material). The correlation between these two variables is statistically significant at 95% confidence level in a two tailed t - test only for the FST= 1st July and 1st May (see Table 4), suggesting that for these FST an increase of the eMED ACC skill scores is related to an increase of the PCP prediction ACC skill scores in models (Fig. 8(a) and (b) and Table 4). Additionally, Table 4 shows that the largest correlation between PCP ACC skill score and eMED ACC skill score is found at FST=1st July, when a large part of the models present skill (in terms of ACC) to predict eMED (see Fig. 5). Figure 9(a) shows the correlation maps between the eMED index and the MSLP and horizontal winds at 850hPa for the FST: 1st July. Though this map is obtained by averaging the 4 NMME models for which we have available atmospheric data, it is representative of all models as the spatial structure of wind pattern is similar in the 4 cases (see section 4 in additional material). Figure 9(a) shows that, in

agreement with observations (compare Figs. 9(a) and 2(a)), an anomalous warming over the eMED generates an anomalous low level pressure at the east of Sahara desert which favors the southward advection of moisture toward Sahel (in agreement with Gomara et al., 2018). Additionally, there is a strengthening of the meridional MSLP across the Sahel that intensifies the southeasterly monsoonal flow, and therefore, it favors moisture supply and positive precipitation anomalies (in agreement with Rowell 2003; Fontaine et al. 2010; Gomara et al., 2018).

3.3.2 Equatorial Pacific (Niño3)

In the case of the tropical Pacific, all models present good ability for reproducing the variability of the SST anomalies over the equatorial Pacific at all considered FST, although their ability is gradually reduced as the FST is moved backwards (Fig. 10). The multimodel-mean presents larger ACC scores than most individual models for all the FST (Fig. 10), suggesting that the pooling of models can improve the ability to predict Niño3 variability for most of them. Nonetheless, the multi-model mean is outperformed by some particular models (e.g. NASA-GEOSS2S and GEM-NEMO for FST=1st July, and COLA-RSMAS-CCSM3 for FST=1st Feb).

Figure 11(a) shows that most models present an RMSE smaller than 2ºC, with the best representation provided by the multimodel-mean for most of the FST. The model with highest RMSE is GDFL-CM2p1. Regarding MSESS, Fig. 11(b) shows that models CMC2-CanCM4, CanCM4i, GEM.NEMO, COLA-RSMAS-CCSM4 (GFDL-CM2p5-FLOR-A06, GFDL-CM2p5-FLOR-B01 and NASA-GEOSS2S) present positive MSESS for FST from 1st July to 1st May (from 1st July to 1st June), suggesting that for these FST the Niño3 prediction provided by these models is better, in terms of the MSE, than the climatological forecast. On the opposite side, models GFDL-CM2p1, GFDL-CM2p5-FLOR-A06, NCEP-CFSv2 and CMC1-CanCM3 show negative MSESS values all the FST. In general terms, it is found that the earlier the FST, the lower the MSESS (more negative), and that most models show negative MSESS for FST before 1st May, indicating that the Niño3 index forecasted by models is worse than considering the climatological forecast for that FSTs. The worst Niño3 predictions are given by GDFL-CM2p1for most of the FST (Fig. 6(b)), consistent with results of RMSE (see Fig. 11 (a)). Comparing results from the multimodel-mean and models (Fig. 11(b)), the former present larger MSESS values than the rest of the models for all the FST, suggesting that the pooling of models can improve the ability to predict Niño3.

Figure 12 (first column) shows the histogram of the Niño3 – precipitation correlation in NMME models. Most models correctly reproduce the negative sign of the observed Niño3 – precipitation teleconnection, although with a generally reduced regression value (Fig. 12, second column), suggesting Sahel precipitation in models tends to be less sensitive to Niño3 anomalies than observed. Correlation values are also underestimated, although less than in the eMED-Sahel precipitation case, which could

be explained by a proper representation of the Niño3 index variance (see section 2 in additional material). Conversely, the correlation value is overestimated in the case of the multimodel mean. Note that in this case we are comparing the correlation/regression of individual simulations with the correlation of the mean of all simulations. A reasonable explanation for such overestimation is that the averaging of all the simulations and models filters out the internal atmospheric variability and those signals that the model cannot predict, making Niño3 the main signal explaining Sahel precipitation variability, unlike in observations.

For the case of Niño3-Sahel teleconnection, it is found that the better the skill in reproducing the Niño3-Sahel precipitation teleconnection, the better the ACC skill score in predicting Sahel precipitation (Fig. 13(a) and (b), Table 4) and the larger Niño3 contribution to precipitation prediction ACC skill score in models (Fig. 13(c) and (d), Table 4). This suggests that enhanced precipitation prediction skill can be obtained by improving the simulation of the precipitation - Niño3 teleconnection, in accordance with Giannini et al., (2020). Figure 9(b) shows the correlation map between the Niño3 index and the anomalous velocity potential difference between the levels 200hPa and 850hPa for the FST: 1st July (see section 5 in the additional material for the rest of the FST). Results show that, in agreement with results from reanalysis (compare Figs. 9(b) and 2(b)), an anomalous warming over the equatorial Pacific reduces the vertical ascent motions, leading to a weakening of the monsoon and to negative rainfall anomalies over Sahel (in agreement with Gomara et al. 2018; Joly and Voldoire 2009). Although these spatial patterns are obtained considering the average of the atmospheric fields from 4 different models, the ones for which atmospheric data was available, the atmospheric teleconnection patterns in the individual models are largely similar to the mean (not shown).

In summary, we are able now to answer question 1 and question 2. There is skill for Pacific but not for Mediterranean SSTs. Additionally, it is found that the better the skill in simulating the Pacific SST-rainfall teleconnection, the better the skill in predicting Sahelian rainfall, however, in the case of the eMED, models with a better ACC skill scores in simulating Mediterranean SSTs tend to better reproduce the observed Sahel rainfall when there is skill for predicting eMED (FST=1st July).

3.4 Explained model variance and ACC scores.

Variance of the Sahel precipitation index in models is partitioned following the same multiple regression analysis already applied to the observed index (see eqs (1) and (2)). Figure 14 shows the variance of the precipitation index partitioned into the components explained by the eMED, Niño3, the cov(eMED,Niño3) and the residue. The latest (residual term) represents the part of the Sahelian precipitation variance which is not explained by the multiple linear regression, or in other words, the part of the variance not explained by the considered oceanic signals (eMED and Niño3). Given that the variance of precipitation in models is much lower than in observations (see section 2 in additional

material), precipitation indices in Fig. 14 are standardized for an easy comparison with observations. The standardization of the precipitation index in each model is done by dividing the index by the standard deviation of the index in the model. Figure 14 only shows results for the first two FST (1st July and 1st June), whereas the rest of the FST can be found in the additional material (see section 6). Additionally, Table 5 shows the percentage of total variance of precipitation that is explained by the eMED and Niño3 indices for each model and FST.

In observations, the precipitation variance explained by the eMED and Niño SSTs is approximately 58%. Focusing on Table 5 and Fig. 14, this value is lower in models, suggesting that models underestimate the influence of these basins (eMED and Niño3) on rainfall over Sahel. The larger precipitation variance explained by the residue in models could be related to the underestimation of the sensitivity of Sahel precipitation to the eMED and Niño3 signals (Figs. 7 and 12, second columns), and to the smaller variance of the eMED SST anomalies in models (see section 2 in additional material). Additionally, the residual shows no statistical significant ACC scores when compared with the observed index for all models and FSTs (except for the NCEP-CFSv2 model at FST=1st May), suggesting that no further sources of predictability aside from the eMED and Niño are present in the models (not shown). Note that other factors as those related to the interbasin interactions, which could introduce counteracting effects, are not considered in this analysis (Polo et al. 2008, Losada et al. 2012, Suarez-Moreno et al. 2018).

Focusing on FST = 1st July, Fig. 14(a) shows different cases in models' behavior. Whereas in observations the precipitation variance explained by eMED is much larger than the one explained by Niño3, in models there are cases in which (1) the eMED explain more percentage than the Niño3 (e.g., CMC2-CanCM4, GEN-NEMO, COLA-RSMAS-CCSM3, NCEP-CFSv2), (2) others in which Niño3 explain a larger percentage of precipitation variability than eMED (e.g., CMC1-CanCM3, CanCM4i, COLA-RSMAS-CCSM4, the 4 GFDLs models and NASA-GEOSS2S), and (3) others in which the SSTs (eMED and Niño3) explain a very small percentage (both IRI-ECHAM models).

Finally, in Fig. 15 we decompose rainfall ACC skill score provided by each model in the contributions coming from each predictor and the residue. To do so, we follow eq. (5). The eMED/Niño3/residue contribution to precipitation prediction ACC skill score must be understood as the part of the precipitation ACC skill score which is explained by eMED/Niño3/residue, respectively. In Fig. 15 we only show results for the first two FST (1st July and 1st June). The rest of the FST can be found in the supplementary material (see section 7). Even though the eMED and Niño3 indices did not account, in general, for a great part of the modeled rainfall variability, using only the eMED and Niño3 SST indices, a large fraction of the ACC skill scores obtained for Sahel precipitation in Fig. 3 can be explained for

most models and FST. Among them, the Niño3 is the main contributor to these precipitation ACC skill scores (see Figs.15(a) and (b) and section 7 in the additional material).

The contribution of eMED to precipitation skill is weak and its role is mainly restricted to FST 1st July to 1st May (see Figs. 8(a) and (b) and section 3 in supplementary material). For these FST, it is found that the greater eMED ACC skill score in models, the larger contribution of eMED to precipitation skill (Fig. 8(c) and (d), Table 4). Focusing on FST = 1st July, when most models present skill to predict eMED, it is found the larger eMED skill, the larger precipitation prediction ACC skill scores (see Fig. 8(a)). These results suggest that increased skill in Sahel precipitation can be expected with improved predictions of eMED SSTs and improved simulation of the Niño3-Sahel precipitation teleconnection.

4 Discussion

Results of this study suggest that, although most of the NMME models show poor skill for predicting precipitation over the Sahel, the better they represent SST signals and SST-rainfall teleconnections, the better the skill in predicting rainfall. These results highlight the importance of ocean variability for the predictability of Sahelian rainfall and that the simulation of teleconnections is a key element to consider for a correct forecast.

As previously shown, models in general lack skill to predict rainfall over Sahel, although there are some of them which present statistically significant ACC skill scores for specific FSTs (GFDL-CM2p5-FLOR-A06 and GFDL-CM2p5-FLOR-B01 present skill for FST= 1st July and 1st June, GFDL-CM2p1 for FST from 1st May to 1st March, GFDL-CM2p1-aer04 for FST= 1st June, 1st April and 1st February, NCEP-CFSv2 for FST= 1st June, 1st April and 1st March, and COLA-RSMAS-CCSM4 for FST=1st Feb). Results are similar for the multi-model mean, which presents statistically significant skill scores only for FST=1st June and 1st Feb (see Fig. 3). These low skill scores are in striking contrast to the statistically significant ACC skill scores obtained by Giannini et al. (2020), especially for the multi-model ensemble, which presented statistically significant skill scores for JAS Sahel precipitation well above the significant threshold for FSTs up to 5 months before. Several factors could contribute to these contrasting results, like the number of models considered (5 vs 15), the zonal domain used to define the Sahel (20ºW-40ºE vs 15ºW-º15ºE), the period of study (1982-2016 vs 1982-2010) and the type of observational dataset employed to assess skill scores (CHIRPS (Funk et al., 2015) vs GPCP and CRU). However, we find that the main reason for such differences is related to the methodology used to preprocess the forecasts. While Giannini et al. 2020 did not remove the long-term trend in the data, we remove it because we focus on the interannual precipitation variability over the Sahel. In Fig. 16(a) and (b) we show the skill of 5 NMME models when the trend is removed (Fig. 16(a)) and when the trend is not removed (Fig. 16(b)). In this figure and, as in Giannini et al. (2020), we only consider one model per modelling group. From the comparison of these two figures we can see that when the trend is not

503 removed (as in Giannini et al. 2020) precipitation skill for most models and FST increases. Results in
504 Fig. 16(b) are more consistent with the ones obtained by Gianini et al. 2020 and suggest that a part of
505 the significant skill scores obtained in that work are related to the long-term trend. Nevertheless,
506 regardless of the actual values of the ACC skill scores, both studies agree on the key role of the ocean
507 and, in particular, ENSO, in the seasonal predictability of Sahel rainfall.

508 There is another factor which merits further discussion: the role of the equatorial Atlantic. As stated in
509 the introduction, studies show that an anomalous warming over the equatorial Atlantic (Atl3) reduces
510 precipitation over the Sahel (see references in the introduction), and viceversa for a cooling. The sign
511 of the correlation between the Atl3 and rainfall coincides with the one from Niño3 - precipitation
512 teleconnection. Rodríguez - Fonseca et al. 2009 showed that, after 70s, the Atl3 and Niño3 started to be
513 connected in such a way that anomalous warming over the equatorial Pacific would be concomitant
514 with anomalous cooling over the eastern equatorial Atlantic. Taking into account that our period of
515 study is after 70s, warmer Niño3 events could be coexisting with cooler Atl3 events, counteracting their
516 effects (Losada et al. 2012). This could be a reason for the nonsignificant correlation found in Fig. 1
517 over the equatorial Atlantic.

518 **5 Conclusions**

519 The objective of this study is to analyze the PCP prediction skill over the Sahel in 15 seasonal
520 forecasting models from NMME, understanding where the skill or lack thereof comes from and what
521 models need to improve to get better precipitation predictions.

522 The forecast skill is analyzed considering the ACC skill score, RMSE and MSESS. Results show that
523 the precipitation ACC skill scores over Sahel are low and that most of models present negative MSESS
524 (except GFDL-CM2p5-FLOR-A06 for FST=1st July, CMC2-CanCM4 for FST= 1st Feb, NCAR-
525 CESM1 for FST=1st Feb and NCEP-CFSv2 for FST=1st June), indicating that models forecast skill in
526 terms of MSE is, in general, worse than considering the climatological forecast. The multimodel-mean
527 shows one of the best results in terms MSESS and ACC, suggesting that, although its precipitation
528 prediction is still worse than the climatological forecast (because the negative MSESS values observed
529 in Fig. 4(b)), the pooling of models leads to a forecast skill greater than the majority of single model
530 systems (see Figs. 3 and 4(b)). On the other hand, most models present an RMSE lower than 3 mm/day.
531 The model which provides a better representation of the precipitation (lower RMSE) depends on the
532 FST and NASA-GEOSS2S shows the highest RMSE values for all the FST.

533 In general, results of this paper highlight the importance of El Niño and the Mediterranean Sea surface
534 temperature in explaining rainfall predictability. Although the precipitation ACC skill scores over Sahel
535 are low, the better the SST variability and SST-rainfall teleconnections is represented by models, the

higher the precipitation ACC skill scores. The starting hypothesis of this study is that in observations roughly half of the total variance of Sahel precipitation at interannual timescales can be explained by its teleconnections with SSTs in the tropical Pacific and Mediterranean regions. Given that SSTs can provide long-term memory to the climate system, these areas could also be sources of predictability for Sahel rainfall. In this work we show that for the 1982-2010 period, the main sources of interannual variability for Sahel rainfall are the SST anomalies over the eMED and Niño3 (see Fig. 1), explaining together up to 58% of the total precipitation variance. However, this percentage is reduced in models (see Fig. 14 and Table 5). Two reasons could be behind of this result. The most important is the lack of skill of the NMME models to correctly reproduce the SST anomalies over eMED (see section 8 in additional material), which is the main source of precipitation variability in observations (Fig. 1 or the bar from observations on Fig. 12). In general, models presents low eMED ACC skill scores and negative MSESS values, indicating that models do not reproduce the observed variability of the SST over eMED and that the eMED SST forecast is, in terms of MSE, worse than the climatological forecast. Additionally, when most of the models present statistically significant ACC skill scores to predict eMED (FST=1st July), it is also found that the larger eMED ACC skill score in NMME models, the larger precipitation prediction ACC skill score (Fig. 8(a)). However, although models correctly reproduce the sign of the eMED - precipitation teleconnection (see Fig. 7, first column), they strongly underestimate the amplitude thereof (see Fig. 7, second column). This could be another reason for the lack of precipitation skill and the reduced percentage of total variance of precipitation explained by eMED and Niño3 indices in models. The underestimation of the amplitude of this teleconnection could be related to the lower variance of the SST anomalies over this region in models (see section 2 in additional material), making the teleconnection weaker.

On the other hand, all the models present good skill for predicting the variability of the SST anomalies over Niño3 (Fig. 10) and also for reproducing the sign of the Niño3 – PCP teleconnection (Fig. 12, first column), although most of them underestimate the amplitude (Fig.12, second column). These results make the Niño3, unlike observations, to be the main contributor to the predictability of precipitation (and precipitation skill) in models (see Figs. 14 and 15), a result that agrees with a recent study of Giannini et al. (2020). Although models present a good skill for predicting the variability of the SST anomalies in the equatorial Pacific (high Niño3 ACC skill scores), most of them only show positive MSESS values for the first two FST, suggesting that the Niño3 SST forecast given by them is, in terms of MSE, better than the climatological forecast only for these two first FST.

The variance of the SST anomalies over Niño3 region in models is similar to the one in observations (see section 2 in additional material). Thus, the underestimation of the teleconnection Niño3 - precipitation amplitude could be mainly related to difficulties in models to correctly reproduce the intensity of the teleconnection mechanism. On the other hand, it is also found the larger correlation

Niño3 - precipitation in models, the larger contribution to precipitation ACC skill score in models (see Fig. 13(c) and (d), Table 4) and precipitation ACC skill score (see Fig. 13(a) and (b), Table 4). Additionally, results from Fig. 15 suggest that the election of the predictors is appropriate, given that most of the models' skill is explained considering eMED and Niño3 as predictors.

Therefore, results from this study suggest the models need two requirements for having a better precipitation prediction skill: an improvement of the models' ability to reproduce the SST anomalies variability over the eMED region and a better simulation of the amplitude of the teleconnections: Niño3 – precipitation and eMED – precipitation.

Finally, is it important to mention that other sources of predictability should be considered in the future, as this study is just valid for the assessment period used. SST - rainfall telconnections are not stationary on time and Mediterranean, Atlantic, Indian and Pacific oceans could exert different impact on Sahel rainfall depending on the time period considered. Also, interbasin teleconnections could be counteracting or adding their effects on the Sahel rainfall depending on the period considered, a feature that should be checked with observations when analyzing the seasonal prediction system of study.

Acknowledgement

We acknowledge the agencies that support the NMME-Phase II system, and we thank the climate modeling groups (Environment Canada, NASA, NCAR, NOAA/GFDL, NOAA/NCEP, and University of Miami) for producing and making available their model output. NOAA/NCEP, NOAA/CTB, and NOAA/CPO jointly provided coordinating support and led development of the NMME-Phase II system. Authors also thank the support of the Spanish Ministry of Science and Innovation for the support through project PRE4CAST (CGL2017-86415-R) and the EU H2020 project TRIATLAS (No. 817578).

References

Adler RF, Huffman GJ, Chang A, Ferraro R, Xie P, Janowiak J, Rudolf B, Schneider U, Curtis S, Bolvin D, Gruber A, Susskind J and Arkin P (2003) The Version 2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979-Present). J. Hydrometeor., 4,1147-1167.

Arizmendi F, Marti A, Barreiro M (2014) Evolution of atmospheric connectivity in the 20th century Nonlinear Proc. Geophys., 21,825-839.

Balmaseda MA, Aives OJ, Arribas A, Awaji T, Behringer DW, Ferry N, FUJI y, Lee T, Rienecker M, Rosati T and Stammer D (2009). Ocean initialization for seasonal forecasts. Oceanography, 22(3), 154-159.

Barreiro M (2010) Influence of ENSO and South Atlantic Ocean on climate predictability over southeastern South America. Clim. Dyn. 35: 1493–1508, doi: 10.1007/s00382-0666-9.

Barreiro M, Tippmann A (2008) Atlantic modulation of El Nino influence on summertime rainfall over Southeastern South America. Geophys. Res. Lett. 3: L16704, doi: 10.1029/2008GL035019.

Batté L, DeÉQueÉ M (2011). Seasonal predictions of precipitation over Africa using coupled ocean-atmosphere general circulation models: Skill of the ENSEMBLES project multimodel ensemble forecasts. Tellus A: Dynamic Meteorology and Oceanography, 63(2), 283-299.

Barrier N, Cassou C, Deshayes J, Treguier AM (2014) Response of North Atlantic Ocean circulation to atmospheric weather regimes. Journal of physical oceanography, 44(1), 179-201.

Batté L, Ardilouze C, Déqué M (2018) Forecasting West African heat waves at subseasonal and seasonal time scales. Monthly Weather Review, 146(3), 889-907.

Bretherton CS, Widmann M, Dymnikov VP, Wallace JM and Bladé, I (1999) The effective number of spatial degrees of freedom of a time-varying field. *Journal of climate*, *12*(7), 1990-2009.

Cai W, Wu L et al (2019) Pantropical climate interactions. Science, 363(6430)

Déqué, M. (2011). Deterministic forecasts of continuous variables. *Forecast Verification: A Practitioner's Guide in Atmospheric Science*, 77-94.

Diakhate M, Rodriguez-Fonseca B, Gómara I, Mohino E, Dieng AL, Gaye AT (2019) Oceanic Forcing on Interannual Variability of Sahel Heavy and Moderate Daily Rainfall. Journal of Hydrometeorology, 20(3), 397-410.

Doblas-Reyes FJ, Weisheimer A, Déqué M, Keenlyside N, McVean M, Murphy M, Rogel P, Smith D and Palmer N (2009) Addressing model uncertainty in seasonal and annual dynamical ensemble forecasts. Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography, 135(643), 1538-1559.

Folland C, Owen J, Ward MN, Colman A (1991) Prediction of seasonal rainfall in the sahel region using empirical and dynamical methods. J. Forecast., 10: 21-56. doi:10.1002/for.3980100104

Fontaine B, Garcia-Serrano J, Roucou P, Rodriguez-Fonseca B, Losada T, Chauvin F, Gervois S, Sijikumar S, Ruti P, Janicot, S (2010) Impacts of warm and cold situations in the Mediterranean basins on the West African monsoon: observed connection patterns (1979–2006) and climate simulations. Climate Dynamics, 35(1), 95-114.

Fontaine B, Monerie PA, Gaetani M, Roucou P (2011) Climate adjustments over the African-Indian monsoon regions accompanying Mediterranean Sea thermal variability. Journal of Geophysical Research: Atmospheres, 116(D23).

Funk C, Nicholson SE, Landsfeld M, Klotter D, Peterson P, Harrison L (2015) The centennial trends Greater Horn of Africa precipitation dataset. Scientific data, 2(1), 1-17.

Gaetani M, Fontaine B, Roucou P, Baldi M (2010) Influence of the Mediterranean Sea on the West African monsoon: Intraseasonal variability in numerical simulations. Journal of Geophysical Research: Atmospheres, 115(D24).

639 Garric G, Douville H, Déqué M (2002) Prospects for improved seasonal predictions of monsoon
640 precipitation over Sahel. Int. J. Climatol., 22: 331-345. doi:10.1002/joc.736

641 Giannini A, Ali A, Kelley CP, Lamptey BL, Minoungou B, Ndiaye O (2020) Advances in the lead time
642 of Sahel rainfall prediction with the North American Multi-Model Ensemble. Geophysical Research
643 Letters, e2020GL087341.

644 Harris I, Osborn TJ, Jones P, Lister D (2020) Version 4 of the CRU TS monthly high-resolution gridded
645 multivariate climate dataset. Scientific data, 7(1), 1-18.

646 Hemri S, Bhend J, Liniger MA, Manzanas R, Siegert S, Stepheson DB, Guitérrez JM, Brookshaw A
647 and Doblas-Reyes FJ (2020) How to create an operational multi-model of seasonal forecasts?. Clim
648 Dyn **55,** 1141–1157. https://doi.org/10.1007/s00382-020-05314-2

649 Hersbach H, Bell B, Berrisford P, Hirahara S, Horányi A, Muñoz-Sabater J, ... and Thépaut JN (2020)
650 The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999-
651 2049.

652 Huang B, Thorne PW et al (2017). NOAA Extended Reconstructed Sea Surface Temperature (ERSST),
653 Version 5. NOAA National Centers for Environmental Information.

654 Janicot S, Trzaska S, Poccard I (2001) Summer Sahel-ENSO teleconnection and decadal time scale SST
655 variations. Climate Dynamics, 18(3-4), 303-320.

656 Joly M, Voldoire A (2009) Influence of ENSO on the West African monsoon: temporal aspects and
657 atmospheric processes. Journal of Climate, 22(12), 3193-3210.

658 Kandji ST, Verchot L, Mackensen J (2006) Climate change and variability in the Sahel region: impacts
659 and adaptation strategies in the agricultural sector.

660 Kitoh A, Mohino E, Ding Y, Rajendran K, Ambrizzi T, Jose M, and Magaña V (2020). Combined
661 oceanic influences on continental climates. In Mechoso, C. R., editor, Interacting climates of ocean
662 basins, chapter 7, pages 216–257. Cambridge University Press, Cambridge, United Kingdom

663 Kirtman Ben P, Min D et al (2014) The North American Multimodel Ensemble: Phase-1 seasonal-to-
664 interannual prediction; Phase-2 toward developing intraseasonal prediction. Bull. Amer. Meteor. Soc.,
665 95, 585–601. doi: http://dx.doi.org/10.1175/BAMS-D-12-00050.1

666 Li LZ (2006) Atmospheric GCM response to an idealized anomaly of the Mediterranean sea surface
667 temperature. Climate dynamics, 27(5), 543-552.

668 Losada T, Rodríguez-Fonseca B, Janicot S, Gervois S, Chauvin F, Ruti P (2010) A multi-model
669 approach to the Atlantic Equatorial mode: impact on the West African monsoon. Climate dynamics,
670 35(1), 29-43.

671 Losada T, Rodriguez-Fonseca B, Mohino E, Bader J, Janicot S, Mechoso CR (2012) Tropical SST and
672 Sahel rainfall: A non-stationary relationship. Geophysical Research Letters, 39(12).

673 Martín-Gómez V and Barreiro M (2016) Analysis of oceans' influence on spring time rainfall variability
674 over Southeastern South America during the 20th century. International Journal of Climatology, 36(3),
675 1344-1358.

676 Mitchell JM, Dzerdzeevskii B, Flohn H, Hofmeyr WL, Lamb HH, Rao KN, and Wallen CC (1966)
677 Climatic change. WMO, Technical Note, No. 79. World Meteorological Organization.

678 Mohino E, Rodríguez-Fonseca B, Mechoso CR, Gervois S, Ruti P, Chauvin F (2011) Impacts of the
679 tropical Pacific/Indian Oceans on the seasonal cycle of the West African monsoon. Journal of climate,
680 24(15), 3878-3891.

681 Mohino E, Keenlyside N, Pohlmann H (2016) Decadal prediction of Sahel rainfall: where does the skill
682 (or lack thereof) come from?. Climate Dynamics, 47(11), 3593-3612.

683 Mortimore MJ, Adams WM (2001) Farmer adaptation, change and 'crisis' in the Sahel. Glob. Environ.
684 Chang. 2001, 11, 49–57.

685 Murphy, A. H. (1998). The early history of probability forecasts: Some extensions and
686 clarifications. *Weather and forecasting*, *13*(1), 5-15.

687 Ndiaye O, Goddard L, Ward MN (2009) Using regional wind fields to improve general circulation
688 model forecasts of July–September Sahel rainfall. International Journal of Climatology: A Journal of
689 the Royal Meteorological Society, 29(9), 1262-1275.

690 Ndiaye O, Ward MN, Thiaw WM (2011) Predictability of seasonal Sahel rainfall using GCMs and lead-
691 time improvements through the use of a coupled model. Journal Climate, 24, 1931–1949.

692 Nicholson SE (2013) The West African Sahel: A review of recent studies on the rainfall regime and its
693 interannual variability. ISRN Meteorology, 2013.

694 Palmer TN, Alessandri A et al (2004) Development of a European Multi- model Ensemble System for
695 Seasonal-to-Interannual Prediction (DEMETER). Bull. Amer. Meteor. Soc., 85, 853–872.

696 Parhi P, Giannini A, Gentine P, Lall U (2016) Resolving contrasting regional rainfall responses to El
697 Niño over tropical Africa. Journal of Climate, 29(4), 1461-1476.

698 Polo I, Rodríguez-Fonseca B, Losada T, García-Serrano J (2008) Tropical Atlantic variability modes
699 (1979–2002). Part I: Time-evolving SST modes related to West African rainfall. Journal of
700 Climate, 21(24), 6457-6475.

701 Rayner NAA, Parker DE, Horton EB, Folland CK, Alexander LV, Rowell DP, Kaplan A (2003) Global
702 analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth
703 century. Journal of Geophysical Research: Atmospheres, 108(D14).

704 Richter I, Doi T, Behera SK, Keenlyside N (2018) On the link between mean state biases and prediction
705 skill in the tropics: an atmospheric perspective. Climate Dynamics, 50(9-10), 3355-3374.

706 Rodríguez-Fonseca, B., Polo, I., García-Serrano, J., Losada, T., Mohino, E., Mechoso, C. R., &
707 Kucharski, F. (2009). Are Atlantic Niños enhancing Pacific ENSO events in recent decades?.
708 Geophysical Research Letters, 36(20).

709 Rodrigues LRL, García-Serrano J, Doblas-Reyes F (2014) Seasonal forecast quality of the West African
710 monsoon rainfall regimes by multiple forecast systems. Journal of Geophysical Research: Atmospheres,
711 119(13), 7908-7930.

712 Rodríguez-Fonseca B, Polo I, García-Serrano J, Losada T, Mohino E, Mechoso CR, Kucharski F (2009)
713 Are Atlantic Niños enhancing Pacific ENSO events in recent decades?. Geophysical Research
714 Letters, 36(20).

715 Rodríguez-Fonseca, B., Janicot, S., Mohino, E., Losada, T., Bader, J., Caminade, C., ... & Voldoire, A.
716 (2011). Interannual and decadal SST-forced responses of the West African monsoon. *Atmospheric*
717 *Science Letters*, *12*(1), 67-74.

718 Rodríguez-Fonseca B, Suárez-Moreno R, Ayarzagüena B, López-Parages J, Gómara I, Villamayor J,
719 Mohino E, Losada T, Castaño-Tierno, A (2016) A review of ENSO influence on the North Atlantic. A
720 non-stationary signal. Atmosphere, 7(7), 87.

721 Rodríguez-Fonseca B, Mohino E, Mechoso CR, Caminade C, Biasutti M, Gaetani M et al (2015).
722 Variability and predictability of West African droughts: a review on the role of sea surface temperature
723 anomalies. Journal of Climate, 28(10), 4034-4060.

724 Rowell DP (2003) The impact of Mediterranean SSTs on the Sahelian rainfall season. Journal of
725 climate, 16(5), 849-862.

726 Rowell DP (2001). Teleconnections between the tropical Pacific and the Sahel. Quarterly Journal of the
727 Royal Meteorological Society, 127(575), 1683-1706.

728 Saravanan R and Chang P (2000) Interaction between tropical Atlantic variability and El Niño-Southern
729 Oscillation. J. Clim. 13: 2177–2194.
730
731 Suárez-Moreno R, Rodríguez-Fonseca B, Barroso JA, Fink AH (2018) Interdecadal changes in the
732 leading ocean forcing of sahelian rainfall interannual variability: atmospheric dynamics and role of
733 multidecadal SST background. Journal of Climate, 31(17), 6687-6710.
734
735 Thorncroft, C. D., Nguyen, H., Zhang, C., & Peyrillé, P. (2011). Annual cycle of the West African monsoon:
736 regional circulations and associated water vapour transport. *Quarterly Journal of the Royal Meteorological*
737 *Society*, *137*(654), 129-147.

738 Vizy EK, Cook KH (2002) Development and application of a mesoscale climate model for the tropics:
739 Influence of sea surface temperature anomalies on the West African monsoon. Journal of Geophysical
740 Research: Atmospheres, 107(D3), ACL-2.

741 Wang, C. Three-ocean interactions and climate variability: a review and perspective. *Clim*
742 *Dyn* **53,** 5119–5136 (2019). https://doi.org/10.1007/s00382-019-04930-x

743

744

745

746

747

Model name	Institute
CMC1-Can3	Canadian Meteorological Center
CMC2-Can4	Canadian Meteorological Center
CanCM4i	Canadian Centre for Climate Modelling and Analysis
GEM-NEMO	Recherche en Prévision Numérique (Canada)
COLA-RSMAS-CCSM3	National Center for atmospheric Research (NCAR)
COLA-RSMAS-CCSM4	National Center for atmospheric Research (NCAR)
GFDL-CM2p1	Geophysical Fluid Dynamics Laboratory NOAA
GFDL-CM2p1-aer04	Geophysical Fluid Dynamics Laboratory NOAA
GFDL-CM2p5-A06	Geophysical Fluid Dynamics Laboratory NOAA

GFDL-CM2p5-B01	Geophysical Fluid Dynamics Laboratory NOAA
IRI-ECHAM4p5-AnomalyCoupled	International Research Institute for climate and Society (IRI)
IRI-ECHAM4p5-DirectCoupled	International Research Institute for climate and Society (IRI)
NASA-GEOSS2S	National Aeronautics and Space Administration (NASA)
NCAR-CESM1	National Center for Atmospheric Research
NCEP-CFSv2	National Center for Environmental Prediction

748 **Table 1.** NMME models considered in this study. See Kirtmann et al., (2014) for further details. The
749 ensemble mean of 4 simulations for each model is considered to analyze the skill.

750

751

752

753

754

755

756

757

758

759

Index	Spatial Domain
Eastern Mediterranean (eMED)	(30ºN-40ºN, 16ºE-38ºE)
Equatorial Pacific (Niño3)	(5ºN-5ºS, 150ºW-90ºW)
Precipitation over Sahel (PCPs)	(10ºN-20ºN, 15ºW-15ºE)

760 **Table 2.** Spatial domain where indices were computed.

761

Predictor	multiple regression coefficient		Percentage of Precipitation variance explained
eMED	α	(0.35±0.16)	46%
Niño3	β	(-0.18± 0.16)	12%

762 **Table 3.** Results from multiple regression analysis considering eMED and Niño3 as predictors. The
763 part of the precipitation variance explained by each predictor is computing following equation 2. Values
764 on table are obtained considering SST from HadISSTv1.1 and PCP from GPCPv2.3.

765

766

767

768

769

Forecast Start Time	PCP ACC vs eMED ACC	eMED contribution to PCP ACC vs. eMED ACC	PCP ACC vs corr(Niño3-PCP)	Niño3 contribution to PCP ACC vs corr(Niño3-PCP)
1sy July	**0.84**	**0.71**	**-0.86**	**-0.94**
1st June	0.31	**0.59**	**-0.85**	**-0.94**
1st May	**0.52**	**0.70**	**-0.51**	**-0.93**
1st April	-0.43	-0.02	**-0.69**	**-0.95**
1st March	-0.45	0.06	**-0.66**	**-0.93**
1st February	0.20	0.10	**-0.77**	**-0.93**

Table 4. Correlation between PCP skill and eMED skill (second column), eMED skill and the eMED contribution to PCP skill (third column), PCP skill and Niño3-PCP teleconnection skill (fourth column), and between the Niño3 contribution to PCP skill and Niño3-PCP teleconnection skill (fifth column). Correlations were computed for each forecast start time and considering all the models from Table 1. Values in bold are statistically significant.

	Forecast Start Time					
Model	**1st July**	**1st June**	**1st May**	**1st Apr**	**1st Mar**	**1st Feb**
CMC1-CanCM3	23	37	33	39	28	13
CMC2-CanCM4	40	18	25	29*	19	13
CanCM4i	29	39	16	17	13	21
GEM-NEMO	14	23	25	3	10	0
COLA-RSMAS-CCSM3	18	6	12*	6	12	15
COLA-RSMAS-CCSM4	18	22	27	61*	43*	46*
GFDL-CM2p1	50*	26	45	48	47	64
GFDL-CM2p1-aer04	42	28	30	45	37	51
GFDL-CM2p5-FLOR-A06	28	46	2	2	12	7
GFDL-CM2p5-FLOR-B01	23	32	13*	19	14	2
IR-AnomalyCoupled	1	2	9*	14	3	16
IR-DirectCoupled	1	4	9*	1	20	14
NASA-GEOSS2S	35*	63*	45	38*	40*	58*
NCAR-CESM1	4	10	24	31	35	15
NCEP-CFSv2	25*	42*	25*	6	29*	36*
Multi – model mean	41	40	36	49	35	41

789 **Table 5**. Percentage of total PCP variance explained by oceans in models. Values with * present
790 negative covariance values between eMED and Niño3, so the percentage of total PCP variance
791 explained by the eMED and Niño indices would be lower.

792

Figures

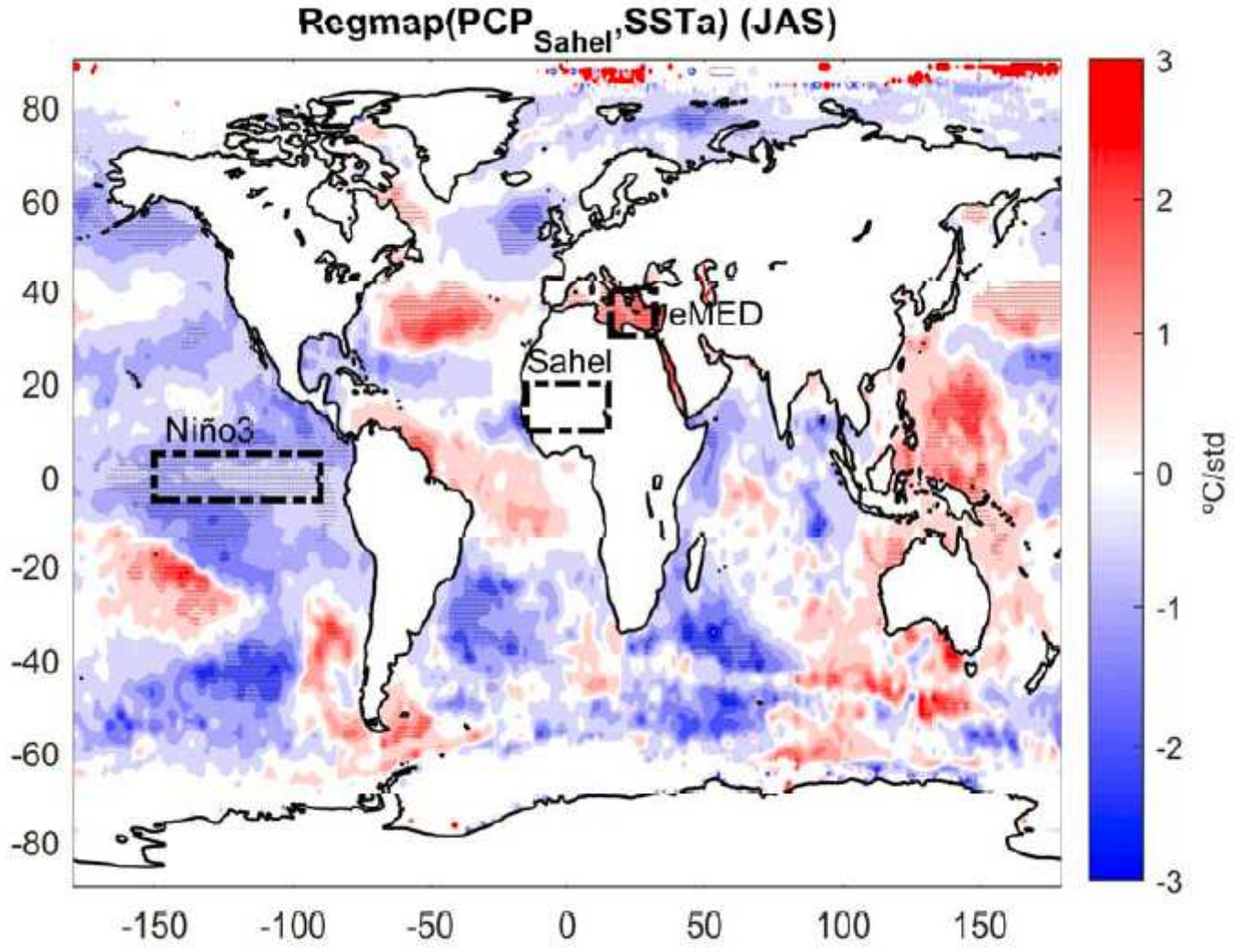
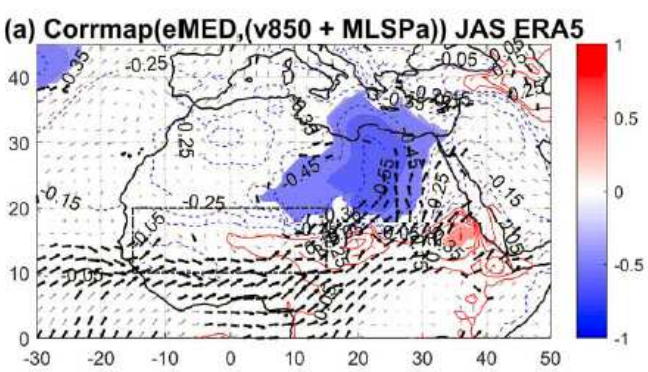


Figure 1

Regression map between the SST anomaly field and the precipitation index over Sahel. Dotted regions are significant at 95% confidence level in two tailed t-test with an effective number of degrees of freedom. SST from HadISST v1.1 and PCP from GPCPv2.3. Regression is computed considering the average of the anomalies in JAS during the period (1982-2010).



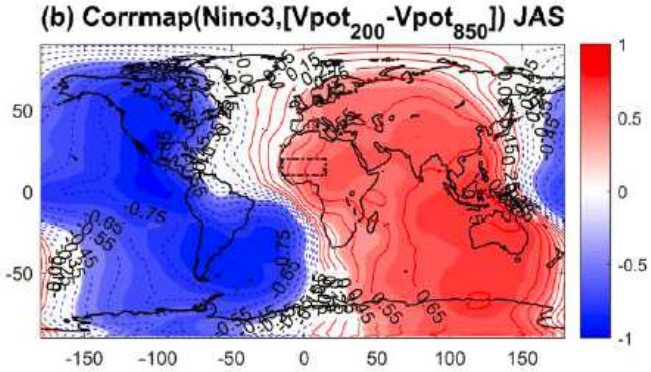


Figure 2

(a) Correlation map between the fields wind anomalies at 850hPa and mean sea level pressure anomalies and the eMED index. (b) Correlation map between the velocity potential field anomaly and the Niño3 index. The velocity potential fields used here is the difference between the velocity potential at 200hPa and 850hPa (VPOT200/850). Shaded regions are significant at 95% confidence level in two tailed t-test with an effective number of degrees of freedom. Mean sea level pressure, horizontal winds are from ERA5 reanalysis. Correlations are computed considering the average of the anomalies in JAS during the period (1982-2010).

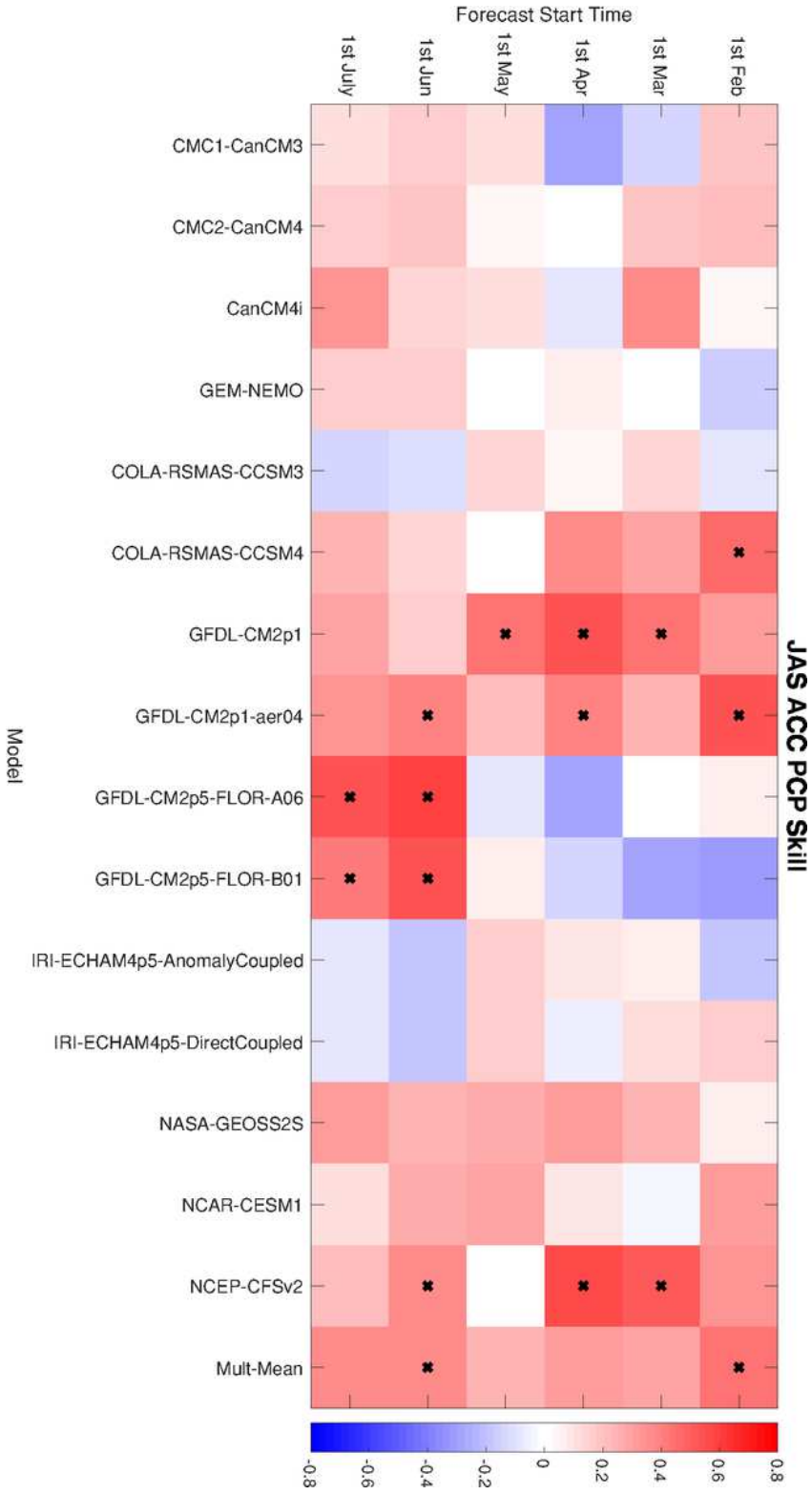


Figure 3

Precipitation prediction ACC skill scores for the different models and FST. Colors represent the ACC values and the boxes marked with “x” present the ACC values statistically significant at 95% confidence level from two tailed t-test with an effective number of degrees of freedom. Note that FST = 1st July corresponds to lead time 0. ACC values are computed considering precipitation index from GPCPv2.3. Results are similar considering precipitation from CRU TS 4.03 (not shown).

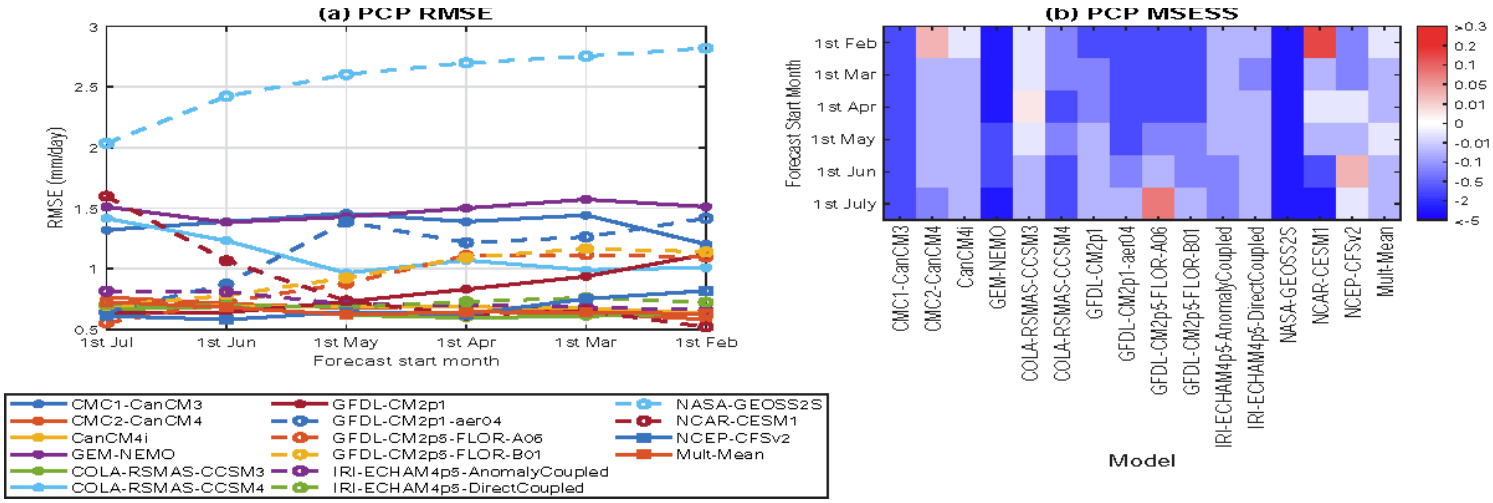


Figure 4

(a) Root mean squared error (RMSE) of precipitation in models (Units: mm/day). (b) Mean squared error skill score for precipitation in models

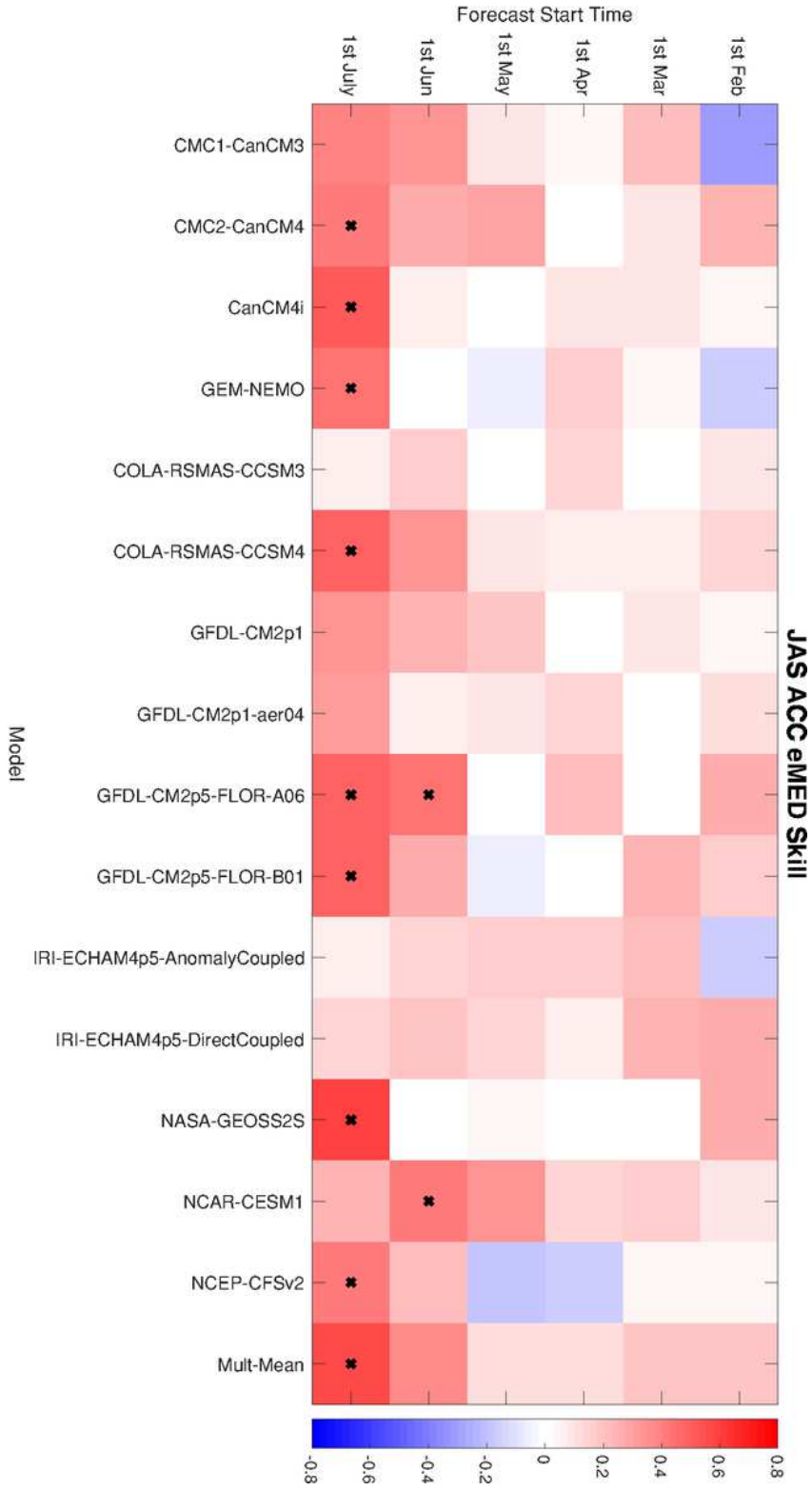


Figure 5

eMED prediction skill in terms of ACC for the different models and FST. Colors represent the ACC values and the boxes marked with “x” present the ACC values statistically significant at 95% confidence level from two tailed t-test with an effective number of degrees of freedom. Note that FST = 1st July corresponds to lead time 0. ACC values are computed considering SST from HadISSTv1.1. Results are similar considering precipitation from ERSSTv5 (not shown).

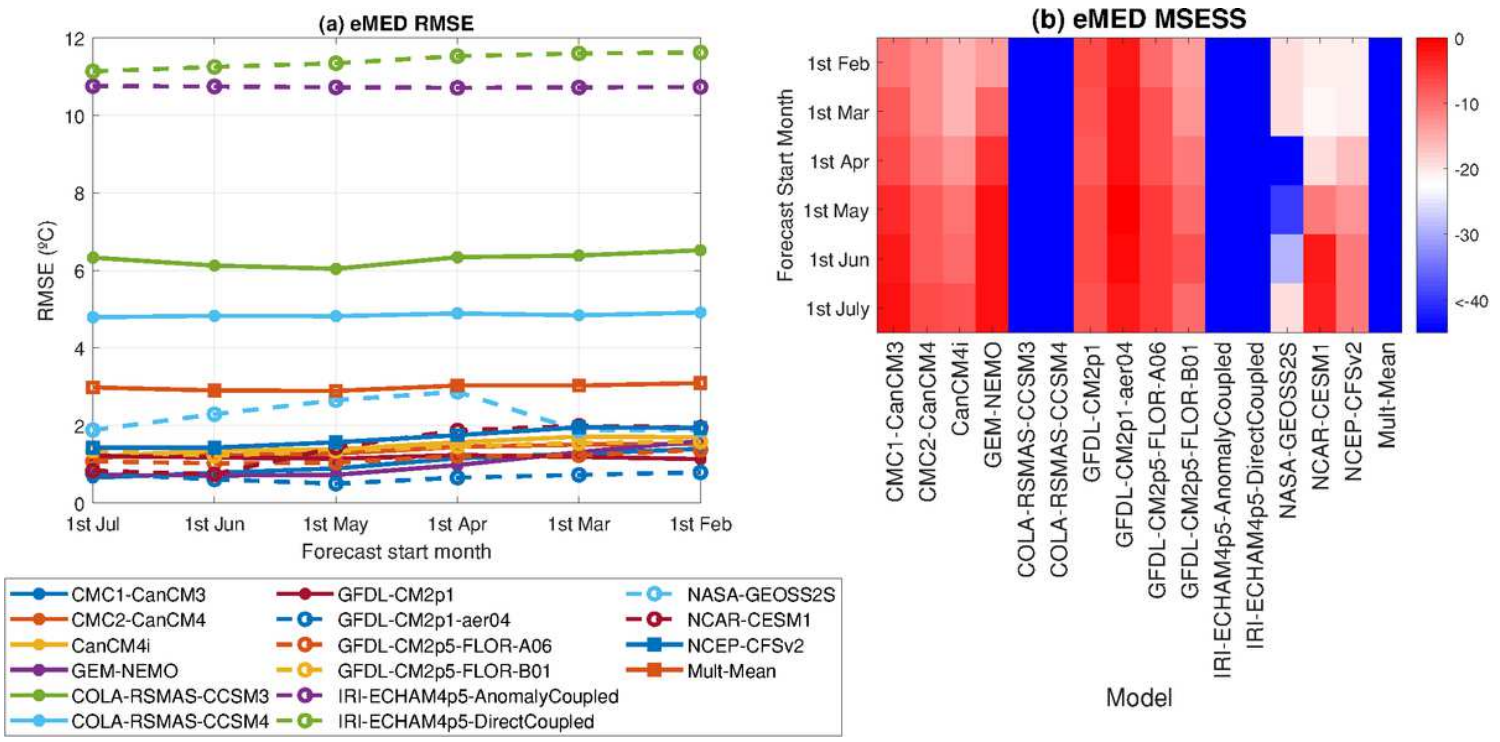


Figure 6

(a) Root mean squared error (RMSE) of eMEDin models (Units: ºC). (b) Mean squared error skill score for eMED in models.

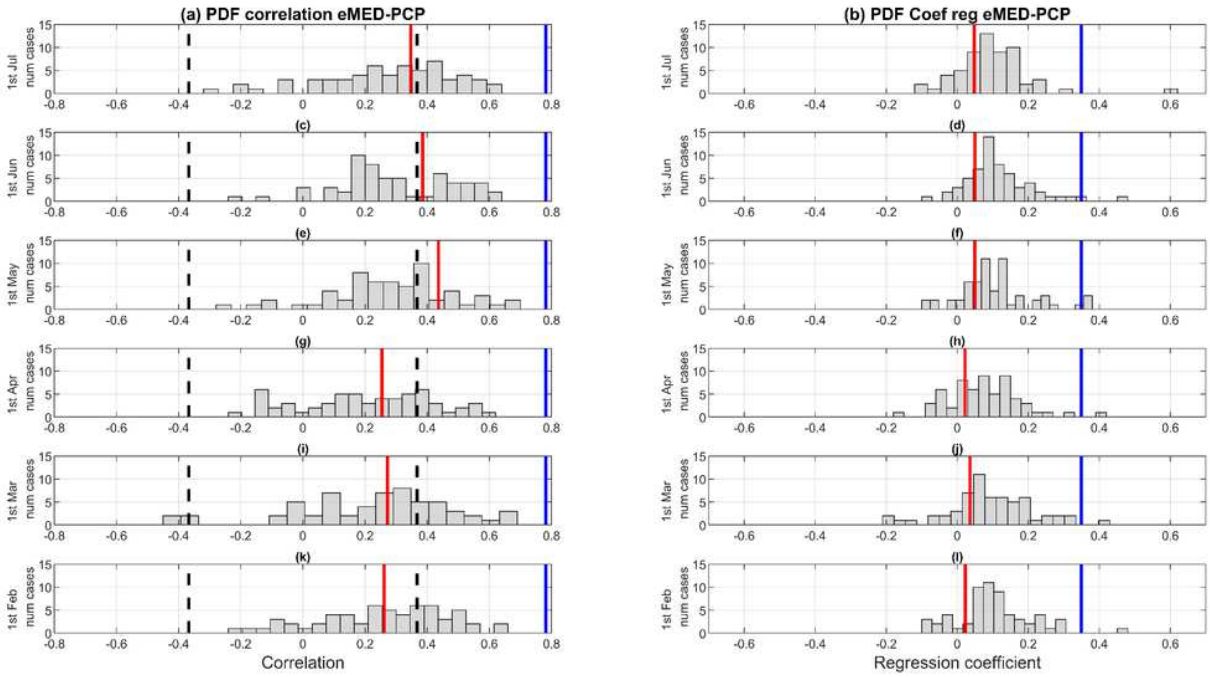


Figure 7

Histograms of the correlation (first column) and regression (second column) coefficients between the precipitation index over Sahel and eMED. The regression coefficient plotted on the second column is α from eq. (1). Each one of the rows makes reference to a lead time, from 0 in the first row to 5 in the sixth. 1st July corresponds to lead time 0 and 1st Feb to lead time 5. The vertical black dot lines, the blue and red ones represent the significance threshold levels, the observed correlation value and the correlation for multimodel mean, respectively. In the case of the multi-model mean, the value is computed as the correlation between the precipitation index of the multi-model mean and the eMED index of the multi.model mean. The threshold level (vertical black dot line) was established considering values exceeding the 95% confidence level from one tailed t – test. Correlations were computed for each one of the 4 simulations of the 15 NMME models. Observed Precipitation is from GPCPv2.3 and observed SST is from HadISSTv1.1.

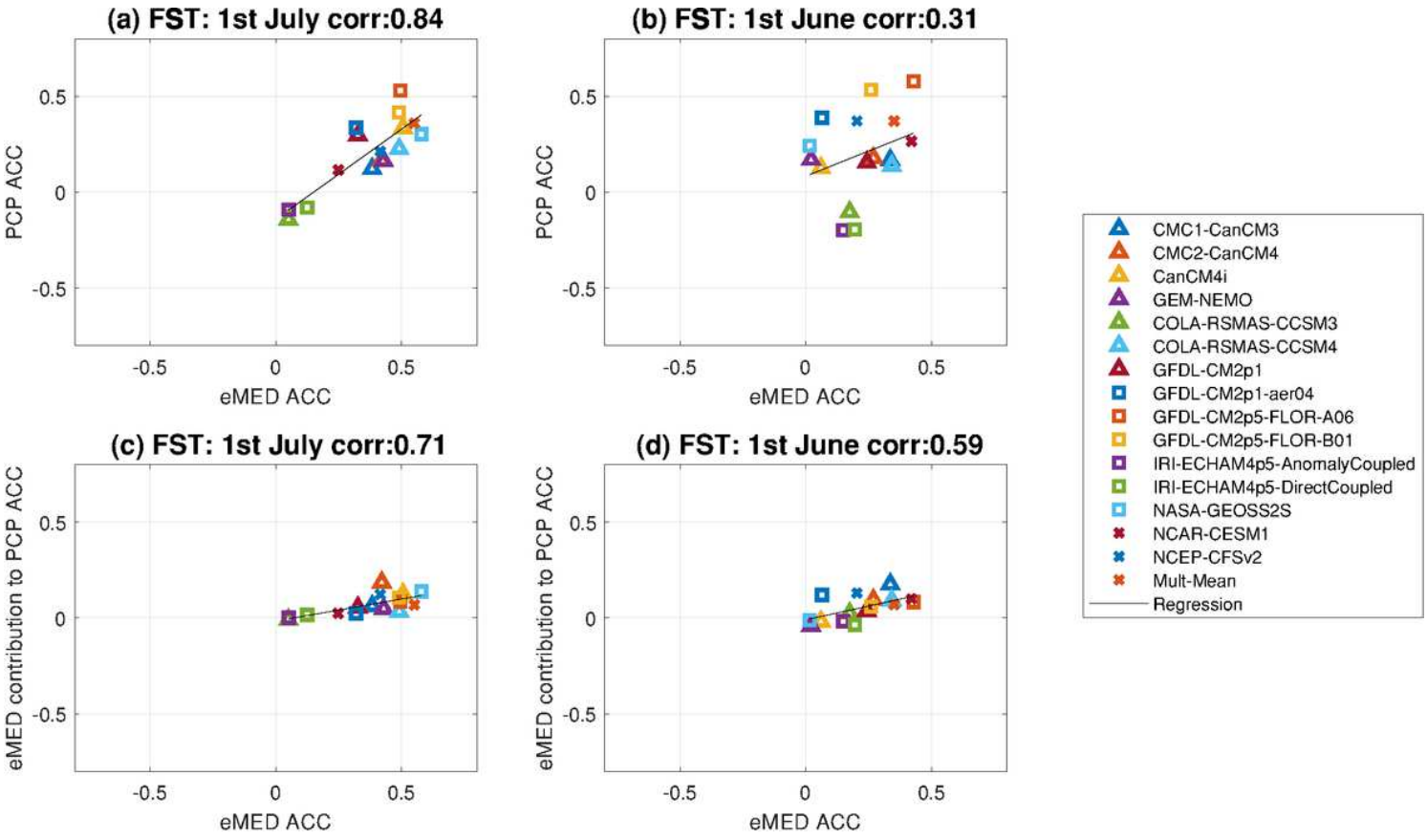


Figure 8

Scatter plots PCP skill in terms of ACC vs eMED skill in terms of ACC for FST= 1st July (a), and FST=1st June (b). Scatter plots eMED contribution to PCP skill in terms of ACC vs eMED skill in terms of ACC for FST= 1st July (c), and FST=1st June (d). Black line represents the linear regression between variables and the correlation values are shown on the title. Precipitation from GPCPv2.3 and SST from HadISSTv1.1. Threshold correlation value is 0.50 considering a two-tailed t -test with a 95% confidence level. Correlations for the rest of the FST appear on Table 4.

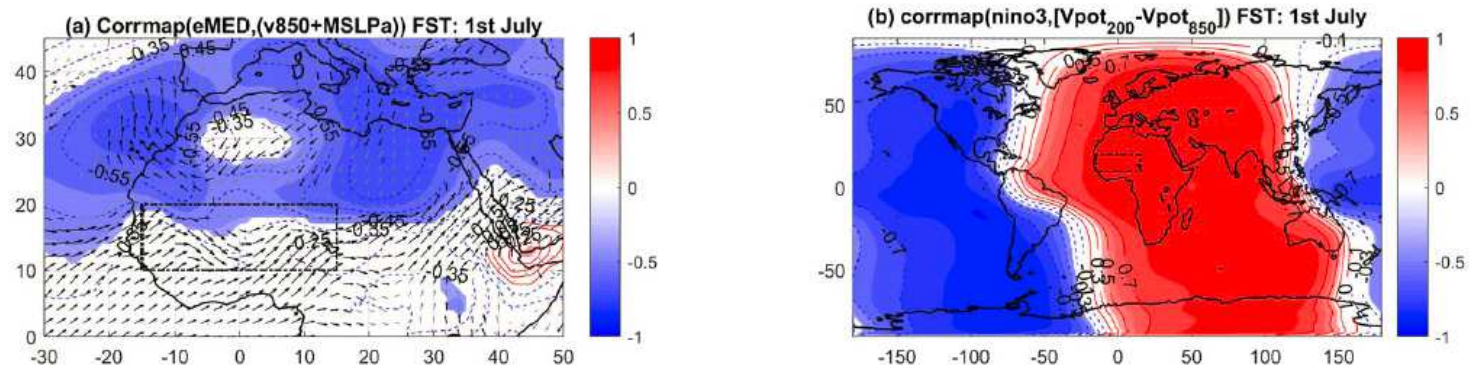


Figure 9

(a) Correlation map between the fields wind anomalies at 850hPa and mean sea level pressure anomalies and the eMED index. (b) Correlation map between the velocity potential field anomaly (VPOT200/850) and the Niño3 index. The velocity potential fields used here is the difference between the velocity potential at 200hPa and 850hPa. Shaded regions are significant at 95% confidence level in two tailed t-test with an effective number of degrees of freedom. Mean sea level pressure, horizontal winds are from the Multi-model Mean constructed considering the models: CanCM4i, CMC1-CanCM3, CMC2-CanCM4, GEM-NEMO. Correlations are computed considering the average of the anomalies in JAS during the period (1982-2010). Figures represent results for FST= 1st July. The rest of the FST can be found in the additional material.

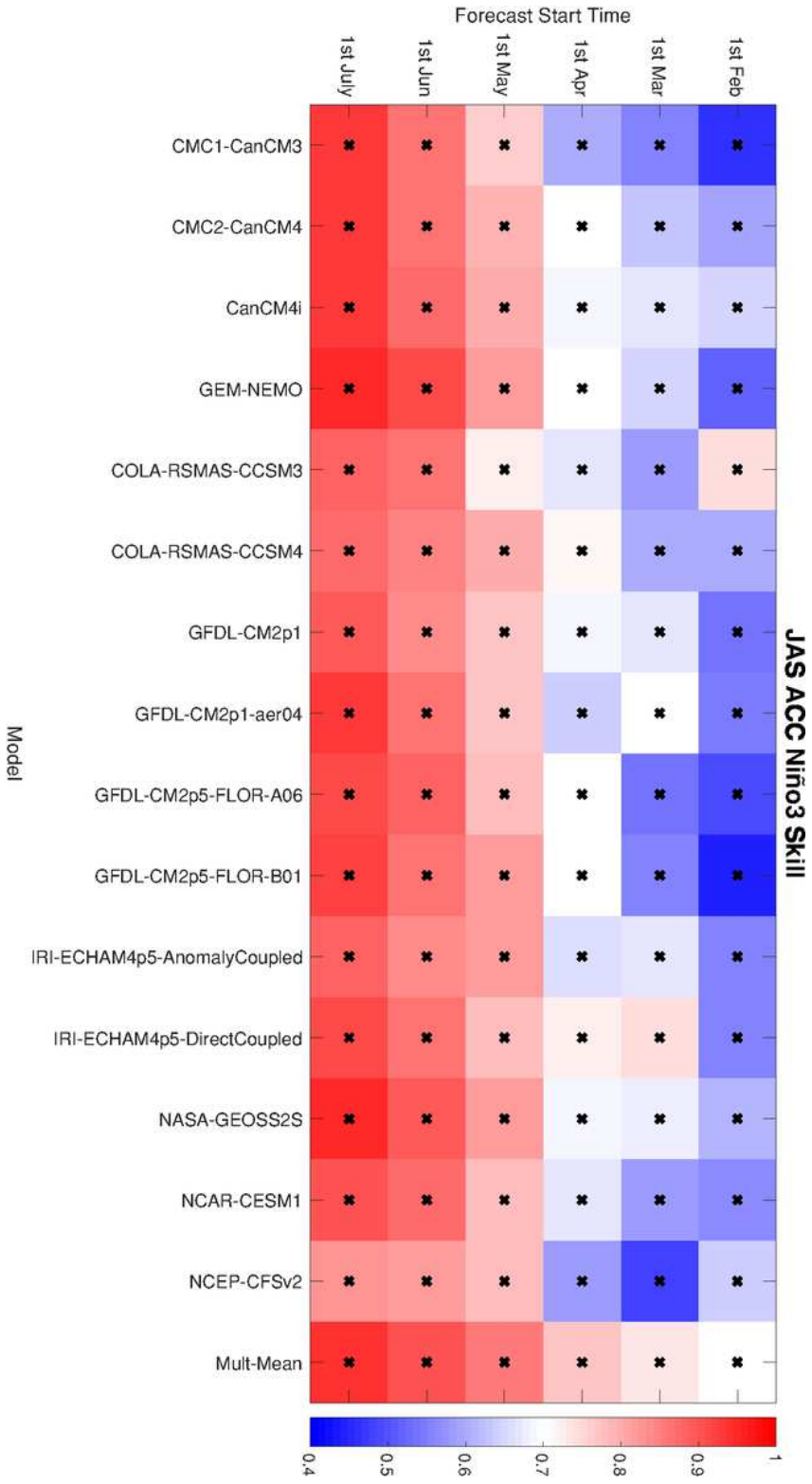


Figure 10

Niño3 prediction skill in terms of ACC for the different models and FST. Colors represent the ACC values and the boxes marked with “x” present the ACC values statistically significant at 95% confidence level from two tailed t-test with an effective number of degrees of freedom. Note that FST = 1st July corresponds to lead time 0. ACC values are computed considering SST from HadISSTv1.1. Results are similar considering precipitation from ERSSTv5 (not shown).

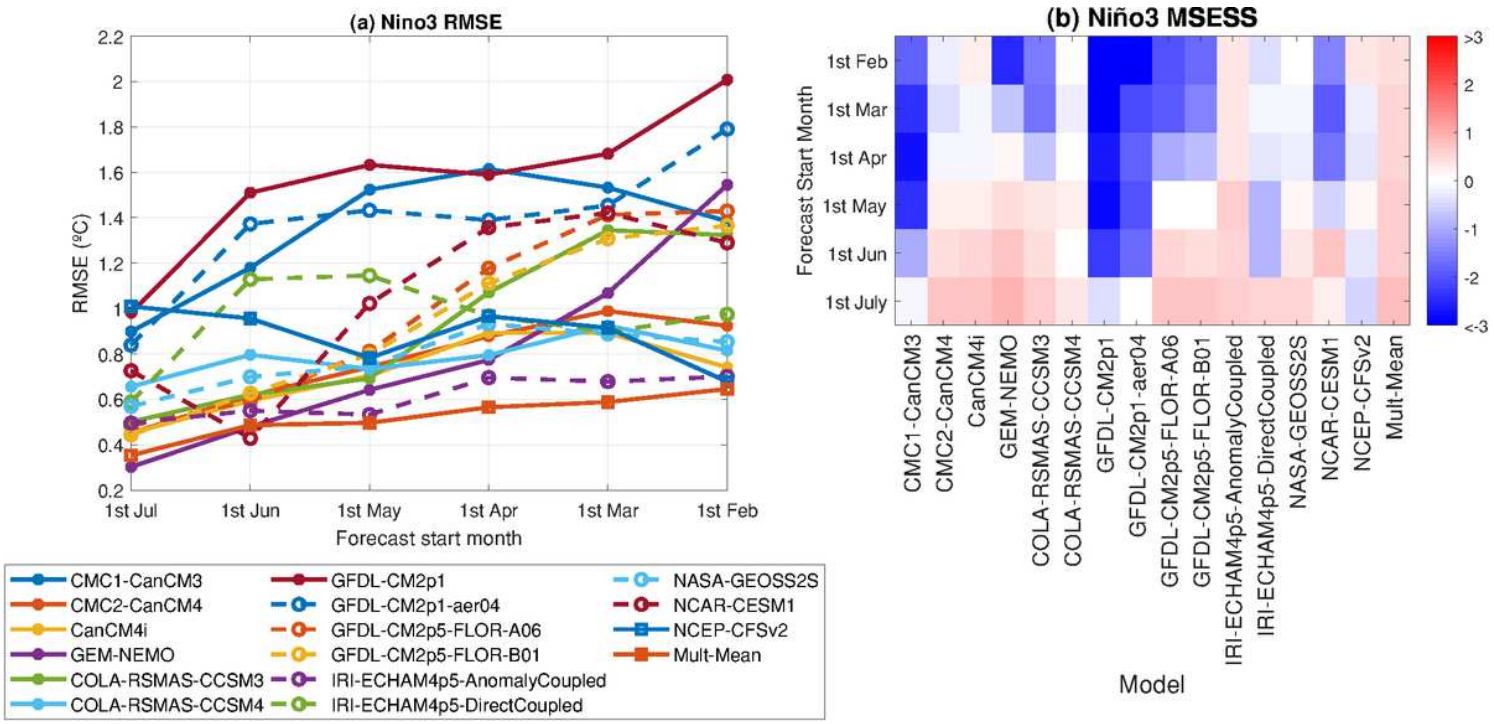


Figure 11

(a) Root mean squared error (RMSE) of Niño3 in models (Units: ºC). (b) Mean squared error skill score for Niño3 in models.

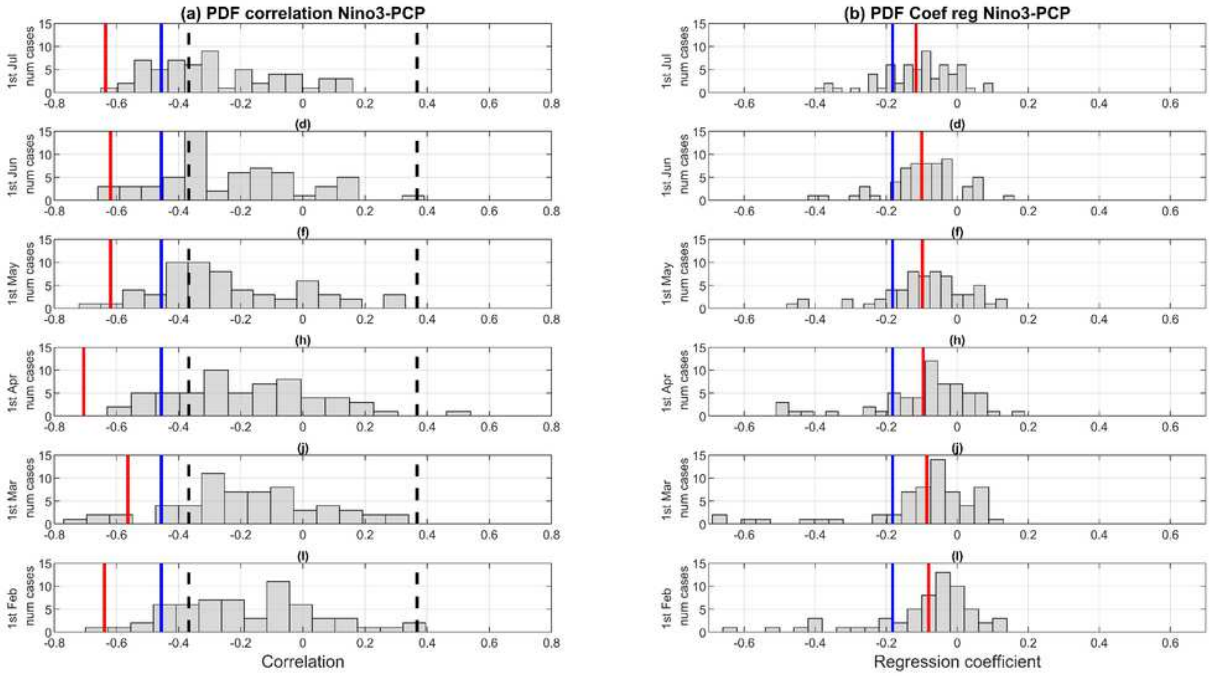


Figure 12

Histograms of the correlation (first column) and regression (second column) coefficients between the precipitation index over Sahel and Niño3. The regression coefficient plotted on the second column is β

from eq. (1). Each one of the rows makes reference to a lead time, from 0 in the first row to 5 in the sixth. 1st July corresponds to lead time 0 and 1st Feb to lead time 5. The vertical black dot lines, the blue and red ones represent the significance threshold levels, the observed correlation value and the correlation for multimodel mean, respectively. The threshold level was established considering values exceeding the 95% confidence level from one tailed t – test. Correlations were computed for each ensemble member of each NMME model. Precipitation from GPCPv2.3 and SST from HadISSTv1.1.

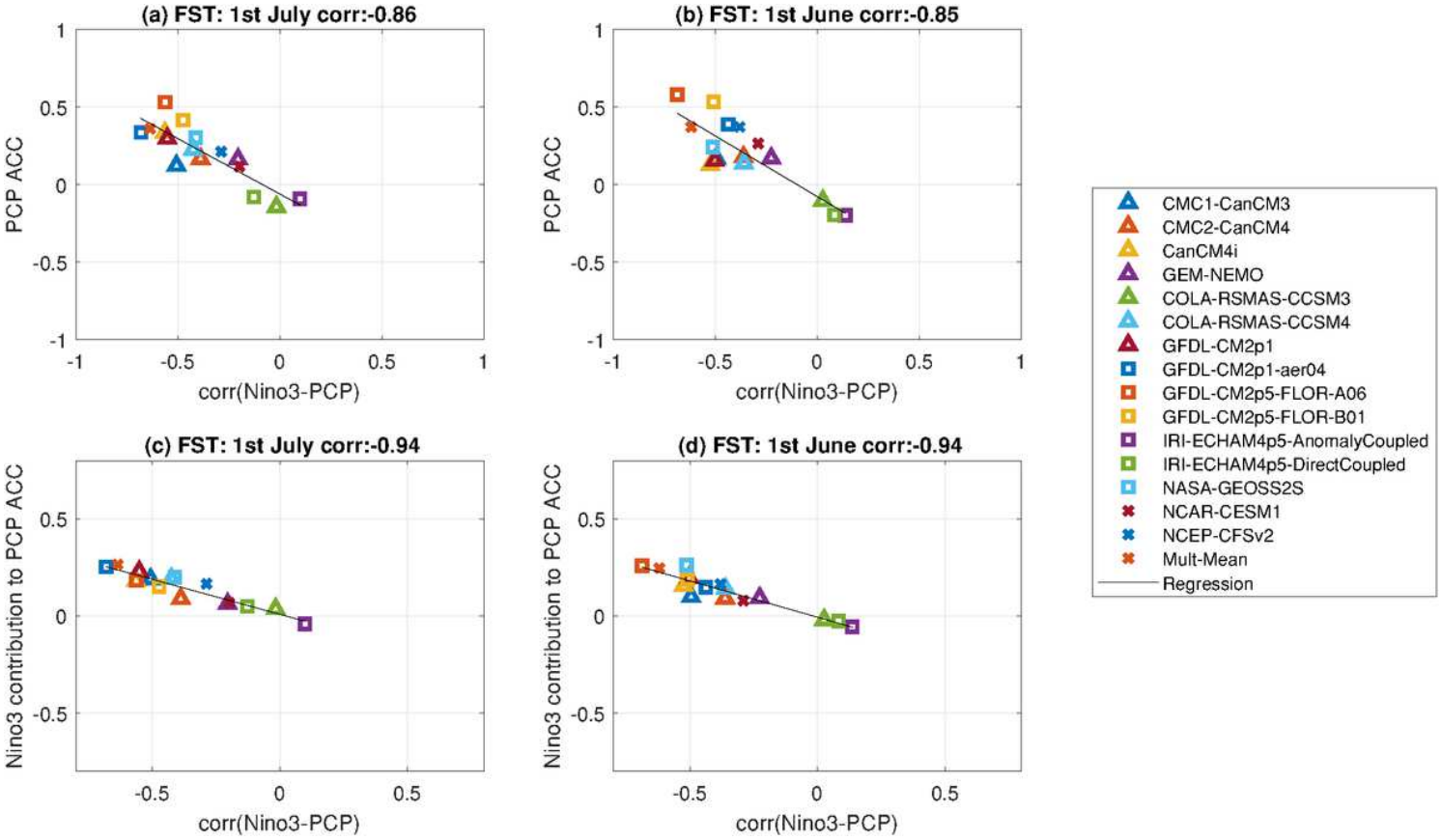


Figure 13

Scatter plots PCP ACC vs Niño3 - PCP teleconnection skill for FST= 1st July (a), and FST=1st June (b). Scatter plots Niño3 contribution to PCP ACC vs Niño3 - PCP teleconnection skill for FST= 1st July (c), and FST=1st June (d). Black line represents the linear regression between variables and the correlation values are shown on the title. Precipitation from GPCPv2.3 and SST from HadISSTv1.1. Threshold correlation value is 0.50 considering a two-tailed t -test with a 95% confidence level. Correlations for the rest of the FST appear on Table 4.

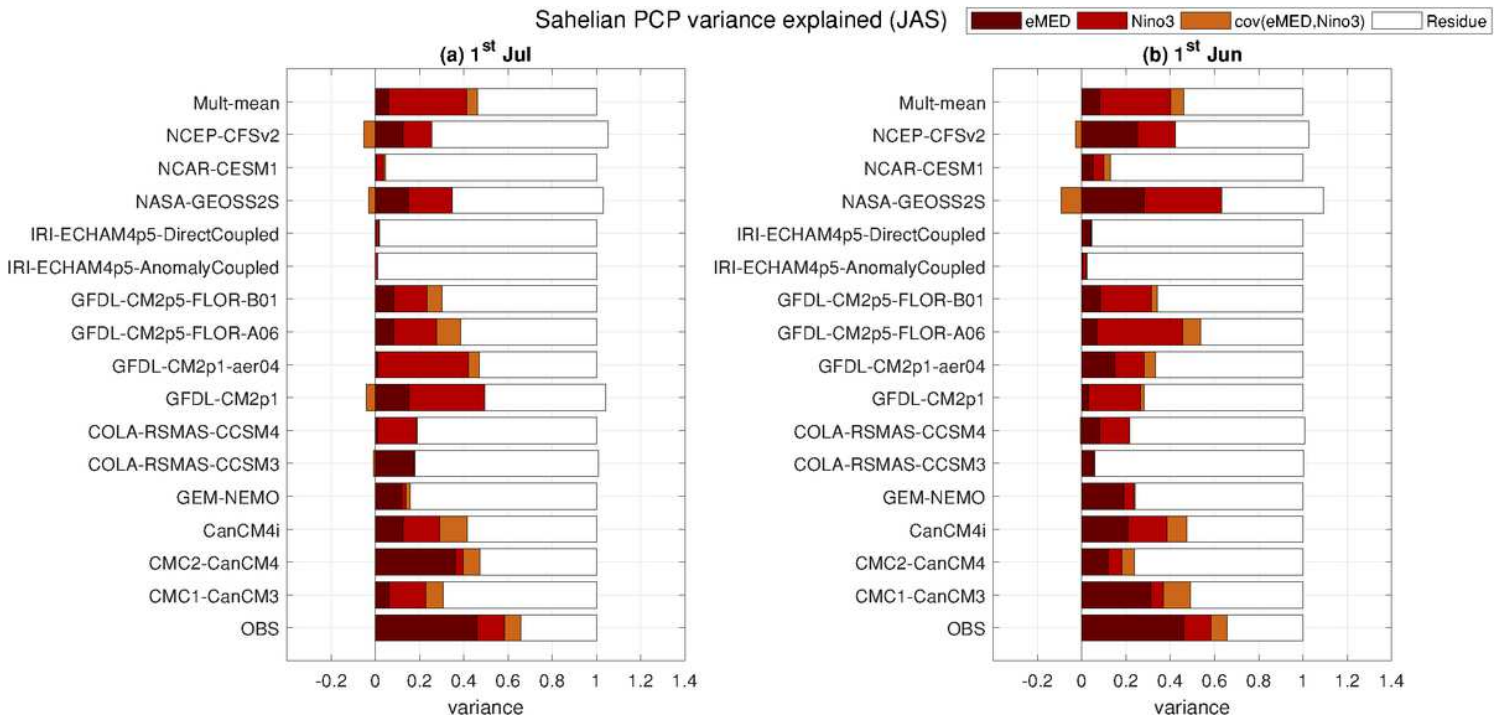


Figure 14

Total precipitation variance decomposition. Given that the variance of precipitation index in observations is larger than in models, precipitation indices are standardized in order to better compare results from observations and models. OBS means observations.

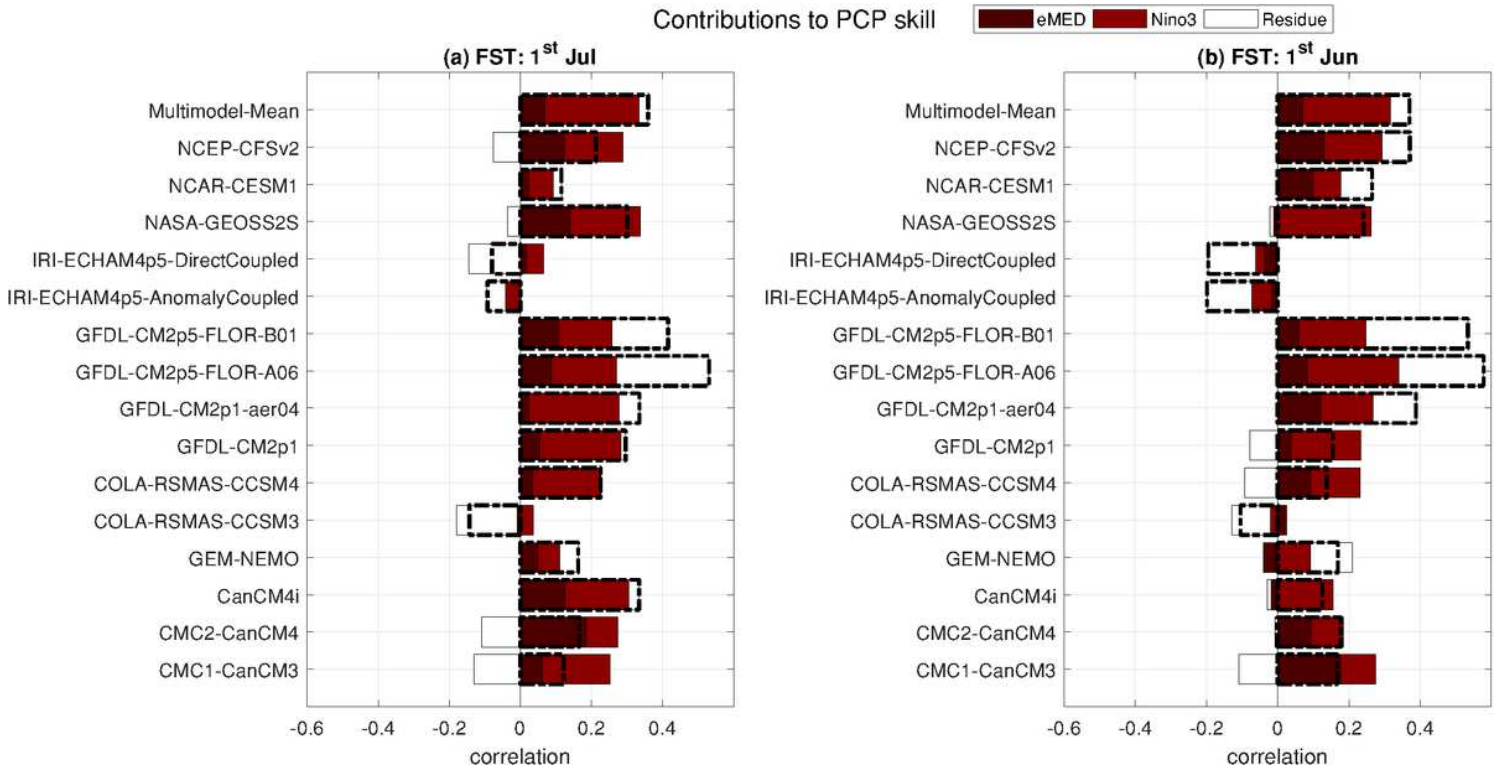


Figure 15

Contribution of eMED , Niño3 and residue to JAS ACC precipitation skill score. Boxes marked with discontinuous black lines represent the value of the correlation between PCPnmme and PCPobs (PCP skill). Boxes in colors represent the part of the PCP ACC skill score which is due to each one of the predictors considered in the multiple linear regression model. Results considering oceanic indices from HadISSTv.1 and precipitation index from GPCPv2.3.

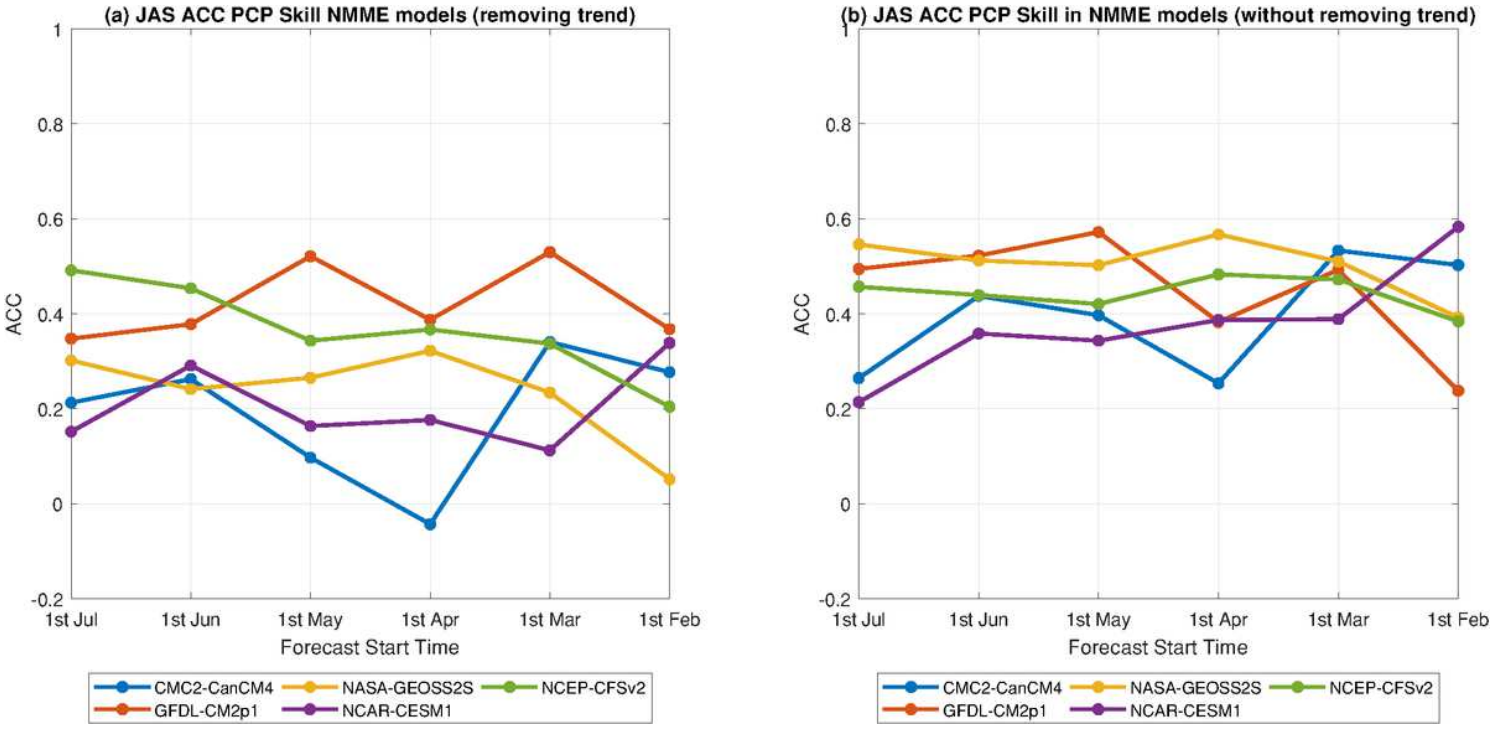


Figure 16

(a) Precipitation prediction skill in NMME models removing trend in data and (b) without removing the trend. Here, we consider 5 models, one per modeling group. The horizontal black dot line represents the threshold level considering a two tailed t-test with a 95% confidence level.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalmaterial.docx