



## Assessment of machine learning-based approaches to improve sub-seasonal to seasonal forecasting of precipitation in Senegal.

Dioumacor Faye<sup>1</sup>, Felipe M. de Andrade<sup>2</sup>, Roberto Suárez-Moreno<sup>3</sup>, Dahirou Wane<sup>1</sup>, Michaela I. Hegglin<sup>4,5</sup>, Abdou L. Dieng<sup>1</sup>, François Kaly<sup>6</sup>, Redouane Lguensat<sup>7</sup>, and Amadou T. Gaye<sup>1</sup>

<sup>1</sup>Laboratoire Physique de l'Atmosphère et de l'Océan-Siméon Fongang (LPAO-SF), Ecole Supérieure Polytechnique (ESP), Université Cheikh Anta Diop de Dakar, Dakar-Fann 5085, Dakar, Senegal

<sup>2</sup>National Institute for Space Research, Cachoeira Paulista, São Paulo, Brazil

<sup>3</sup>Geophysical Institute and Bjerknes Centre for Climate Research University of Bergen 5020, Bergen, Norway

<sup>4</sup>Institute of Climate and Energy Systems - Stratosphere (ICE-4), Forschungszentrum Jülich, Wilhelm-Johnen-Straße, 52428, Jülich, Germany

<sup>5</sup>Department of Meteorology, University of Reading, RG6 6BB, Reading, United Kingdom

<sup>6</sup>Department of Computer Science, UFR of Sciences and Technologies, Université Iba Der THIAM de, Thiès, 21000, Thiès, Senegal

<sup>7</sup>Institut Pierre-Simon Laplace, IRD, 4 place de Jussieu, Paris, France

**Correspondence:** Dioumacor FAYE (dioumacor.faye@ucad.edu.sn)

**Abstract.** In Senegal, the West African monsoon (WAM) season is characterized by pronounced subseasonal to seasonal (S2S) rainfall fluctuations in response to complex interactions between large-scale atmospheric and oceanic variability patterns and mesoscale convective systems. Indeed, the general circulation models (GCMs) used in the development of S2S forecasting systems often struggle to represent the mechanisms yielding WAM predictability. This study explores the potential of machine learning (ML) approaches to improve S2S precipitation forecasting in Senegal. We evaluate a set of ML models, including ridge regression, linear regression, random forest, support vector machine, Adaboost, and multilayer perceptron for S2S forecasting of precipitation during the monsoon season. To this aim, we use a combination of high-resolution global precipitation estimates from ground and satellite observations, along with atmospheric and oceanic reanalysis products. Our methodology relies on a non-filtering approach to extract significant S2S signals as predictors, enabling real-time application. We demonstrate that integrating different predictor variables from a range of atmospheric and oceanic fields significantly enhances prediction skill. Notably, the ridge regression model outperforms state-of-the-art GCM-derived S2S predictions. The study highlights the potential for developing operational S2S forecasting systems for West African precipitation using ML techniques to complement GCM-based forecast systems, offering valuable tools for climate risk anticipation and water resource management. Such ML-based systems not only provide skillful predictions but are also computationally more efficient compared to GCMs, and can be extended to diverse climatic zones.

**Keywords.** Subseasonal to Seasonal forecasting, Senegal, Machine learning, Precipitation



## 1 Introduction

Extreme floods and droughts are becoming more widespread in the current context of climate change, causing considerable economic damage and threatening the livelihood of the population, according to the 6<sup>th</sup> Assessment Report of the Intergovernmental Panel on Climate Change (IPCC). For West Africa in general, and Senegal in particular, reliable subseasonal-to-seasonal (S2S) forecasts of rainfall during the West African monsoon (WAM) season can provide actionable data for informed decision-making to mitigate the harmful impact of extreme events on the vulnerable population, whose economy and subsistence are mainly based on rain-fed agriculture. Indeed, S2S predictability of rainfall in this region is crucial for crop management, disaster risk reduction and food security. This S2S forecasting time scale, which typically ranges from two weeks to two months, helps bridge the gap between weather forecasting and climate prediction (Hung et al., 2009; Brunet et al., 2010; Vitart et al., 2012). However, rainfall variability associated with the WAM is highly complex due to the combined impact of local mesoscale convective systems and remote sea surface temperature (SST) forcing (Suárez-Moreno et al., 2018; Biasutti, 2019), which contributes to preventing skillful S2S predictability (Vigaud and Giannini, 2019).

From a societal perspective, decision-making in the framework of agriculture, food security, water resources, risk management and health care would greatly benefit from improved S2S forecasts. However, this time scale has long been considered a "predictability desert", being much less studied compared to medium-range and seasonal forecasting (Robertson et al., 2020). Recent studies have shown that predictability at the S2S time scale could be enhanced through various factors, including more complete and reliable observational networks, a better understanding and representation of intra-seasonal atmospheric processes, and improved initialization and assimilation of land, ocean, cryosphere and stratosphere components in general circulation models (GCMs)-based prediction systems (Pegion et al., 2019). Accordingly, several initiatives have been carried out, such as the S2S prediction project and the Subseasonal eXperiment (SubX; now the Subseasonal Consortium, SubC), to provide GCM-derived S2S rainfall forecasts up to 60 days in advance (Vitart et al., 2017; Pegion et al., 2019). However, such forecasts are still far from skillful (de Andrade et al., 2019), as GCMs struggle to represent key atmospheric processes, mainly those on a small scale (Vitart and Robertson, 2019). As a result, post-processing is often required to improve the accuracy and reliability of forecasts (Li et al., 2022). Some studies used, for instance, Bayesian joint probability to post-process S2S rainfall forecasts over different regions, resulting in improved skill and reliability compared to the raw forecasts (Schepen et al., 2012; Li et al., 2021). Moreover, a new spatial correction method proposed by (Vigaud et al., 2020) showed improved intra-seasonal rainfall estimates derived from multimodel ensembles. Nevertheless, the accuracy of these post-processed estimates degrades sharply for lead times (i.e., the difference in time between the onset of an observed event and the issuance of the forecast of such event) beyond 10-14 days (Li et al., 2022).

Alternative approaches for S2S rainfall forecasting include machine learning (ML) models that operate from mathematical relationships between rainfall data and preceding indices of atmospheric and/or oceanic variability. Although GCMs have traditionally excelled in short- and medium-term forecasting, and ML models have historically been reserved for longer timescales (Abbot and Marohasy, 2014; Tuel and Eltahir, 2018), recent advancements in ML, such as with models like GraphCast, have demonstrated unprecedented accuracy in medium-term weather forecasting. Many data-driven algorithms, such as multiple lin-



ear regression, maximum covariance analysis or canonical correlation analysis, among others, have been devised for seasonal rainfall expectations based on the assumption that seasonal anomalies are caused by slow variations in SST, snow cover and other boundary conditions (Barnston and Smith, 1996; Hwang et al., 2001; Eden et al., 2015; Suárez-Moreno and Rodríguez-Fonseca, 2015). Additional examples include an empirical cluster-based method to predict winter precipitation anomalies over European and Mediterranean regions using SST, geopotential height, sea level pressure, snow cover extent, and sea ice concentration as predictors (Totz et al., 2017). A random forest-based model, whose predictors were extracted from SST data, was developed to predict seasonal precipitation in Central and Southern Asia (Gerlitz et al., 2016). However, there is a lack of studies focusing on intra-seasonal (or S2S) forecasting in West Africa.

This study takes a step forward to develop a ML-based intra-seasonal rainfall forecasting system for Senegal. Through the exploitation of the links between rainfall variability and atmospheric and oceanic fields, we evaluate the performance of several ML approaches to predict intra-seasonal weekly precipitation anomalies during the WAM season. Our results reveal how these models outperform GCM-based prediction systems. The work is organized as follows: Data and methods are described in Section 2. The forecasting capabilities of the models are evaluated in Section 3. Section 4 is dedicated to discussion and conclusions, including comments on the potential application of the ML approaches for disaster risk reduction, water resources management, and other climate-related risks.

## 2 Data and methodology

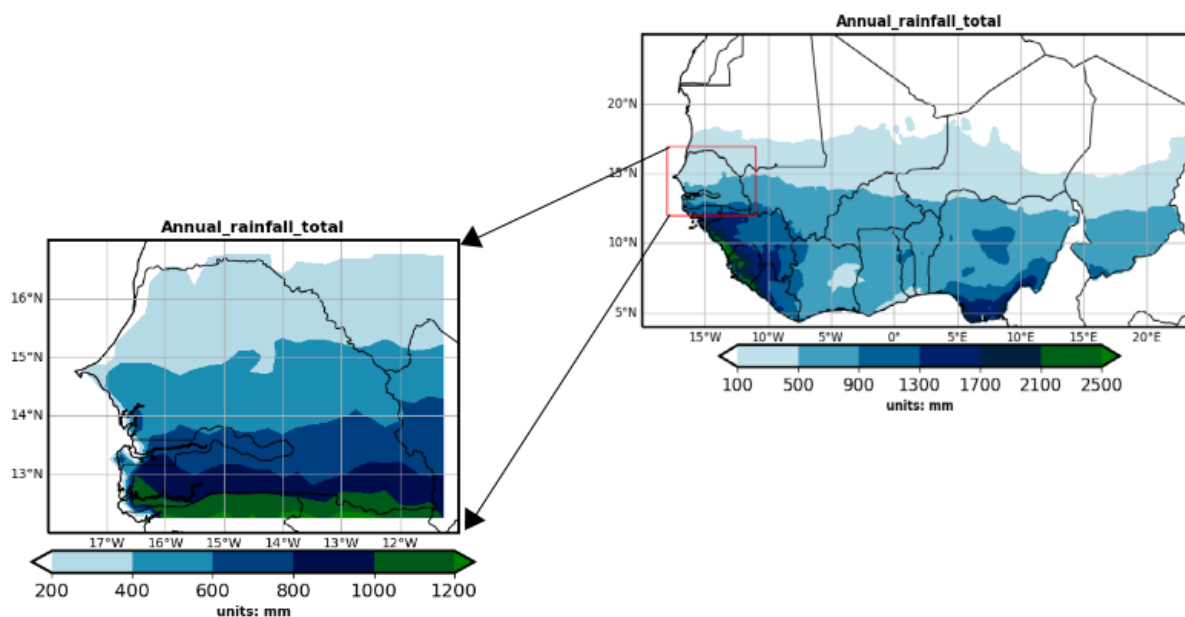
### 2.1 Data

#### 2.1.1 Geographical Information

Senegal is located in the westernmost part of West Africa, between latitudes  $12^{\circ}8$  and  $16^{\circ}41$  north and longitudes  $11^{\circ}21$  and  $17^{\circ}32$  west (red box on Fig. 1). It is characterized by two main seasons: a dry season (from November to May) marked by the predominance of maritime trade winds from the west and continental trade winds from the northeast in the interior; and a rainy season (from May - June to October), dominated by the southwesterly flow that characterizes the WAM (Fall et al., 2020). The maximum rainfall is received in August and September, coinciding with the period during which the intertropical convergence zone (ITCZ) reaches its northernmost position above Senegal (Sane et al., 2018). Fig. 1 illustrates the spatial distribution of the annual cumulative precipitation in West Africa, along with the area of interest. A meridional gradient of precipitation is observed, with higher values in the south and a northward rainfall decreasing (Faye et al., 2024).

#### 2.1.2 Target variable: Precipitation Data

Daily precipitation data are derived from the Climate Hazards group InfraRed Precipitation with Stations (CHIRPS) dataset (Funk et al., 2015). CHIRPS combines ground observation data and satellite observations to produce daily and monthly gridded rainfall estimates worldwide from 1981 to the present. CHIRPS has a high spatial resolution (5 Km), exhibiting more realistic spatial patterns and greater accuracy over land than other observational datasets. A study by (Tarnavsky et al., 2014)



**Figure 1.** Spatial distribution of annual cumulative precipitation in West Africa and Senegal. The red outlined rectangle denotes Senegal as the study region (Faye et al., 2024).

compared CHIRPS with Tropical Rainfall Measuring Mission (TRMM) data (Simpson et al., 1988) for the West African region. The results showed that the two datasets are well correlated, but CHIRPS tends to better capture short-duration rainfall events. Likewise, comparing with the European Centre for Medium-Range Weather Forecasts (ECMWF) dataset, Funk et al. (2015) showed that CHIRPS is more realistic in reproducing lower rainfall intensities. Recently, Faye et al. (2024) showed that CHIRPS data, compared to those from the National Hydrological and Meteorological Services (NHMS), provide a more accurate representation of the total annual rainfall distribution, as well as the onset and cessation of the rainy season in Senegal.

### 2.1.3 Atmospheric and oceanic variables

To represent the intraseasonal atmospheric oscillation, outgoing longwave radiation (OLR) and zonal winds (U) in the upper (U200 hPa) and lower (U850 hPa) troposphere are used. Although several indices, including the real-time multivariate Madden-Julian Oscillation (MJO) index (RMM) (Wheeler and Hendon, 2004) and the boreal summer intraseasonal oscillation (BSISO) index (Lee et al., 2013), have been proposed to track the sub-seasonal oscillation propagation, these may not cover the patterns that could be important for subseasonal rainfall in some regions (Li et al., 2022). In addition, correlations with geopotential height (H) in the lower (H850), middle (H500), and upper (H200) troposphere are also analyzed. Leung and Qian (2017) showed that H at 850, 500, and 200 hPa are able to reflect the MJO structure as well as the zonal wind. Daily OLR data used in this study are provided by the National Oceanic and Atmospheric Administration (NOAA) on a 2.5° resolution global grid (Liebmann and Smith, 1996). The OLR data is derived from high-resolution infrared sounders and is valuable for a wide



range of applications. The daily averaged U850, U200, H850, H500 and H200 data are derived from the fifth generation of the  
100 ECMWF reanalysis (ERA5) dataset. Reanalysis products combine model data with observations from across the world into  
a globally complete and consistent dataset using the laws of physics. This principle, called data assimilation, is based on the  
method used by numerical weather prediction centres, in which given a certain time interval a previous forecast is combined  
with newly available observations in an optimal way to produce a new best estimate of the state of the atmosphere, called  
analysis, from which an updated, improved forecast is issued. The ERA5 data used in this study have a horizontal resolution of  
105 0.25 degrees. To focus on large-scale features and increase computational efficiency, the ERA5 reanalysis data are interpolated  
to 2.5° X 2.5° spatial resolution.

Regarding SST, we use the NOAA Daily Optimum Interpolation Sea Surface Temperature (OISST) (Huang et al., 2021)  
which is a long-term climate record integrating observations from different platforms (satellites, ships, buoys and Argo floats)  
on a regular global grid with a latitude-longitude resolution of 0.25°. The dataset is interpolated to fill gaps on the grid and  
110 create a spatially complete SST map. To be consistent with the temporal coverage of the OISST, we focus on the period  
1982-2019.

#### 2.1.4 S2S hindcasts

Hindcasts, often referred to as reforecasts, are retrospective climate forecasts. The precipitation hindcasts from the S2S database  
were assessed for the models of the ECMWF, the Met Office (UKMO), and the National Centers for Environmental Predic-  
115 tion (NCEP). These forecasts exhibit various configurations, including forecast range, spatial resolution, frequency, period,  
ensemble size, and coupling effects (Table 1), (Vitart et al., 2017; de Andrade et al., 2021). Furthermore, ECMWF and UKMO  
reforecasts are generated gradually by updating their model versions according to near real-time forecasts, whereas in the  
NCEP model, reforecasts are based on a fixed date for a given model version. We analyzed ECMWF and UKMO reforecasts  
corresponding to model version dates of the year 2018. Four start dates per month were chosen based on UKMO initializations  
120 (1st, 9th, 17th, and 25th of each month), as indicated by (de Andrade et al., 2021). We select the closest start date for some  
non-matching ECMWF initializations. This discrepancy regarding models initialization restricted a multimodel evaluation. To  
ensure a fair comparison between models, we select six perturbed members, in addition to the control, to obtain the ensemble  
average for each model. For the NCEP model, which has a total of three perturbations, we have used three additional perturba-  
tions by modifying the initial states. This procedure allows all models to have at least seven members in their ensemble. Since  
125 the subseasonal time scale exceeds the limit of weather prediction, a weekly time frame was used to represent the subseasonal  
forecast range more adequately. Weekly precipitation was obtained considering four accumulation lead times: days 5–11 (week  
1), 12–18 (week 2), 19–25 (week 3), and 26–32 (week 4). The results of our ML-based models will be compared to those from  
these S2S dynamical prediction systems (see Table 1) to assess their performance. The study period selected for comparison  
extends from 1999 to 2010, determined by the availability of NCEP model reforecasts.

130

**Table 1.** The main features of the three S2S operational models and their hindcasts (de Andrade et al., 2021).



Model	Forecast length	Spatial resolution	Hindcast frequency	Hindcast period	Ensemble size	Ocean coupled	Sea ice coupled
ECMWF	46 days	Tco639 /319L91	Two per week	Past 20 years	11	Yes	No
UKMO	60 days	N216 L85	Four per month	1993-2016	7	Yes	Yes
NCEP	44 days	T126 L64	Daily	1999-2010	4 + 3 <sup>a</sup>	Yes	Yes

<sup>a</sup>Three more perturbed members, extracted from 1-day lag after initializations, were added to the NCEP ensemble size.

## 2.2 Methodology

### 2.2.1 Extraction of intraseasonal signals

135 In this part, we discuss the extraction of significant intraseasonal signals, important for S2S precipitation (Li et al., 2022). The raw daily data of atmospheric variables (U850, U200, OLR, H850, H500 and H200), daily precipitation, and SST contain high-frequency noise. Bandpass filtering methods, such as the fast Fourier transform, are commonly used to isolate the intraseasonal scale (10- to 60-day signals) (Zhang, 2005). However, these traditional methods are not practical for real-time applications, as they will introduce future information beyond the current date.

140 As proposed by (Li et al., 2022), we use a non-filtering method to extract 10- to 60-day signals from atmospheric and oceanic variables and CHIRPS precipitation. Compared to traditional methods of extracting intraseasonal signal, this approach could be used for real-time applications. The climatological annual cycle of the raw daily data is first removed by subtracting a 90-day low-pass filtered climatological component, as follows:

$$X' = X - \bar{X} \quad (1)$$

145 where  $X$  is the daily data.  $\bar{X}$  is the corresponding climatological 90-day low-pass filtered component derived by the Lanczos filtering method (Duchon, 1979). The period during which the low-pass filter is applied spans from 1982 to 2019.

In the second step, low-frequency signals of more than 60 days are removed by subtracting the 30-day moving average as follows:

$$X'' = X' - \overline{X'}^{30d} \quad (2)$$

where  $\overline{X'}^{30d}$  is the running mean of the 30 days of  $X'$ .

150 A weekly mean is performed on  $X''$  (for atmospheric and oceanic signals.) to remove higher frequency signals, as follows:

$$X^* = \overline{X''}^{7d} \quad (3)$$



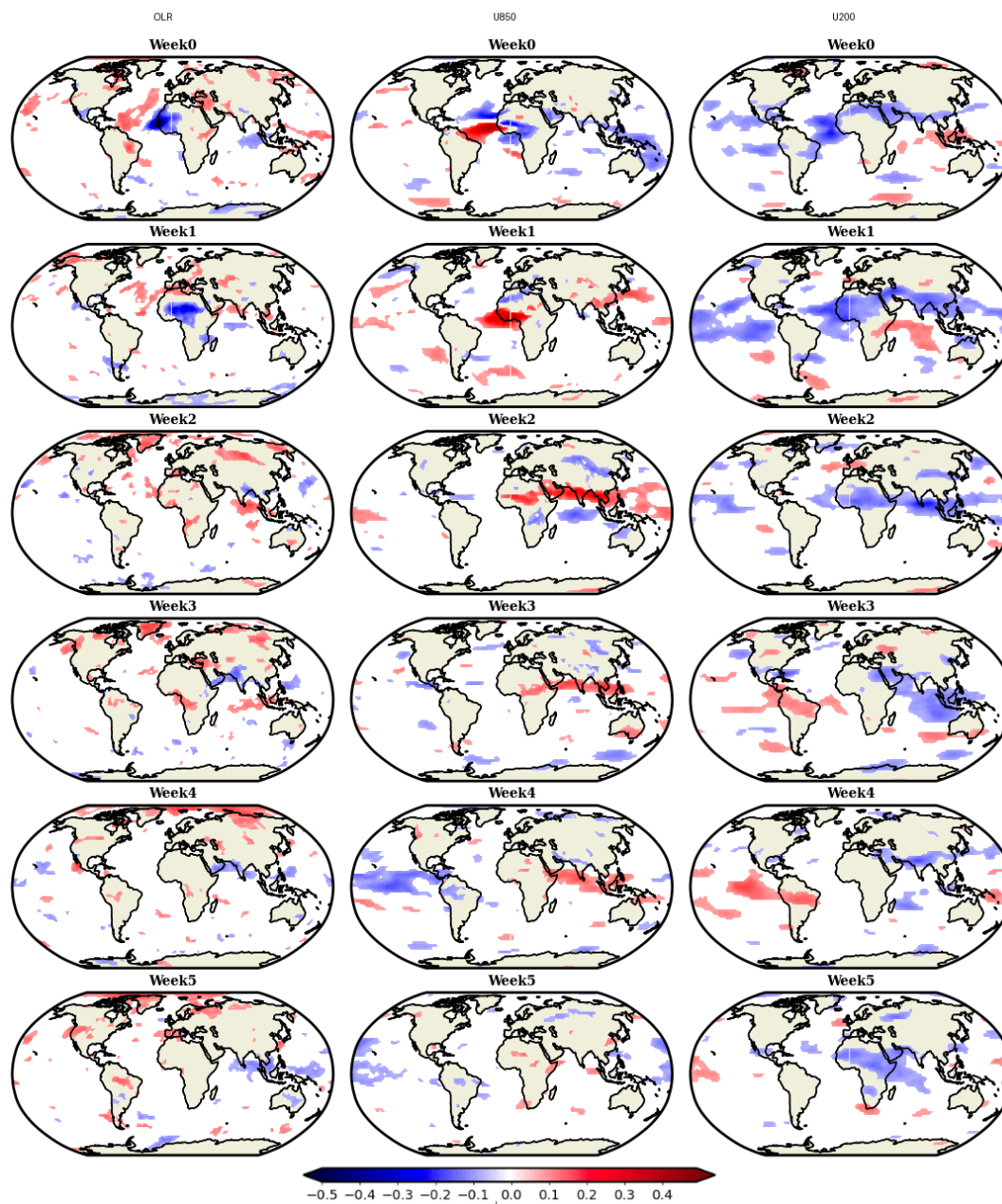
As a result, the signal derived represents the 10- to 60-day signal of  $X$ . The daily intraseasonal signals are averaged into weekly data to further reduce noise and improve predictability (Li et al., 2022). Hereinafter, the 10- to 60-day accumulated weekly precipitation signal will be referred to as the weekly precipitation anomaly.

## 155 2.2.2 Definition of predictors

In order to identify the geographical areas for which the atmospheric and oceanic fields show a potential influence on the subseasonal variability of rainfall in Senegal, we analyze the correlation between weekly precipitation anomaly (i.e., WPA) during the period of 1982–2019, from June to September in Senegal and atmospheric parameters, with weeks 0, 1, 2, 3, 4, and 5. We also did the same with SST oceanic signals. Consider, for example, predicting the accumulated weekly precipitation for  
160 the period from June 1 to 7, 2010. In this case, the mean weekly intraseasonal oscillation (ISO) signals of the atmospheric field for the periods from June 1 to 7 (week 0), May 25 to 31 (week 1), May 18 to 24 (week 2), May 11 to 17 (week 3), May 4 to 10 (week 4), and April 27 to May 3 (week 5) are used as predictors to generate precipitation forecasts at different lead times.

Fig. 2 depicts the correlation between the previous 10 to 60 days mean weekly signals of OLR, U850, and U200 and WPA over Senegal at different temporal lags. At weeks 5 and 4, significantly correlated OLR signals are primarily located above the  
165 Philippine Sea, the Bay of Bengal. These signals seem to propagate towards East Africa at week 4. OLR anomalies remain near the above Africa, and the Indian Ocean at week 3 to week 2. At week 1 and week 0, OLR signals are more pronounced above Africa. The spatial distribution of OLR signals is more concentrated over Central and West Africa at weeks 1 and 0, respectively, indicating more robust statistical relationships in this area. At week 1, signals that were located above Central Africa appear further west (particularly towards Senegal). The high negative correlation over Senegal at week 0 confirms the  
170 results found in the region during the boreal summer. Accordingly, studies have shown that negative (positive) OLR anomalies are indicative of more (less) cloud coverage and hence enhanced (suppressed) convective precipitation in the region (Mohino et al., 2011; Janicot et al., 2008). Thus, when precipitation occurs, a decrease in outgoing infrared radiation in the atmosphere is observed. This phenomenon is attributed to cloud formation and the onset of atmospheric convection, two processes closely linked to precipitation in West Africa. Convective clouds reflect a portion of the infrared radiation back into space, thereby  
175 reducing the amount of infrared radiation detected by satellites (Janicot et al., 2008). Significantly correlated U850 signals are mainly located above the Indian Ocean, the Sea of Oman, Eastern Africa, and the West Pacific at weeks 4 and 5. These U850 hPa signals are between the Philippine Sea and the Horn of Africa at week 3 to week 2 and then seem to propagate toward Central Africa at week 2. A more westward progression is noted for week 1 to week 0, with a spatial distribution of U850 signals more concentrated between subtropical latitudes of both hemispheres at week 0, indicating more robust statistical  
180 relationships in this area.

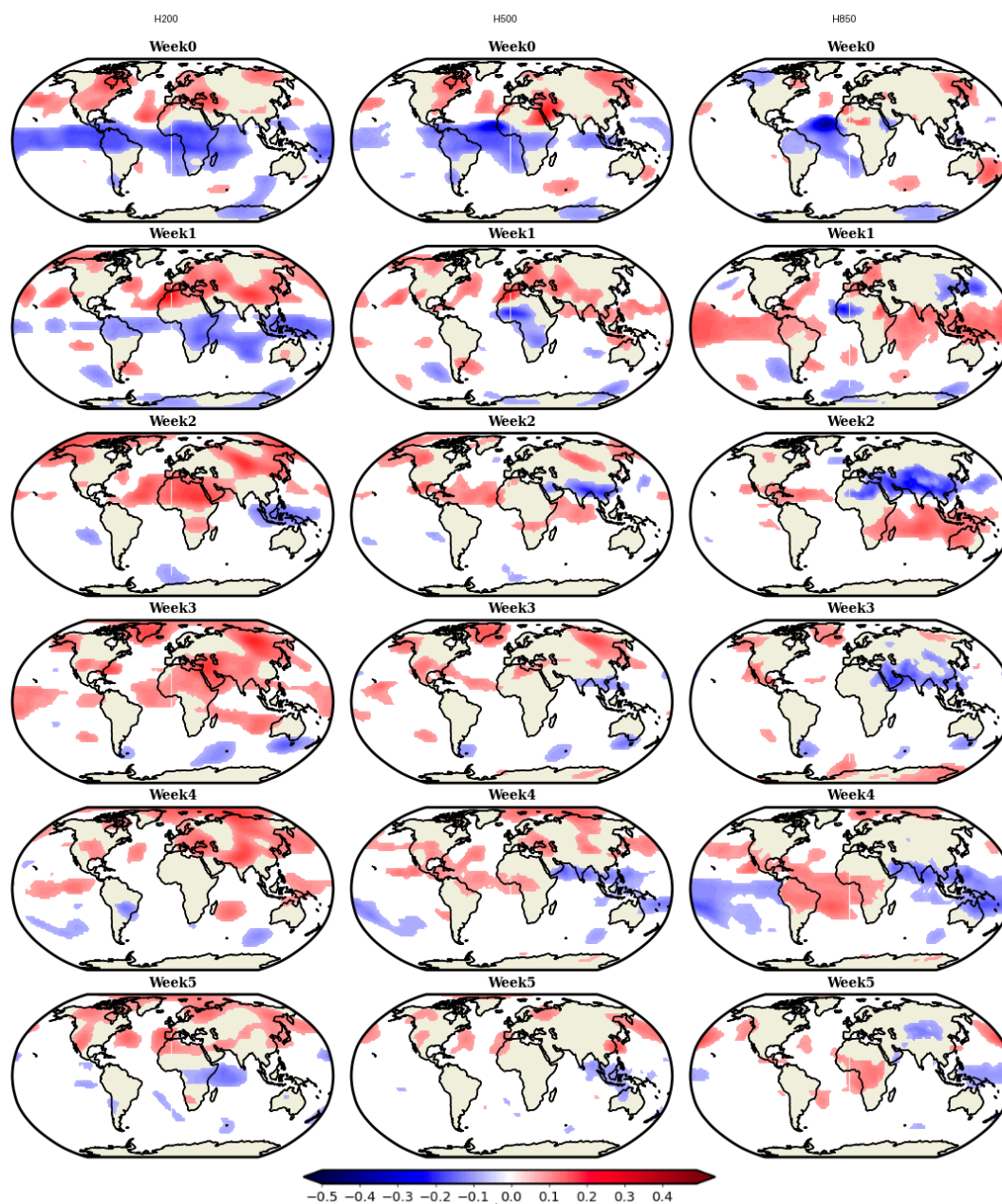
This dipole is also present between the Atlantic Ocean and the African continent, which is potentially related to the onset of the summer monsoon system in West Africa or low-level trade winds carrying moisture from the ocean to the continent. In this context, (Sultan et al., 2003) show that the boreal summer circulation dipole between ocean and continent is often associated with the pressure difference between these two areas. Statistically significant U200 anomaly correlations are found over the  
185 Indian and Pacific Oceans during weeks 5-3. At week 2, the signals appear above Africa and Indian Ocean, whereas at weeks



**Figure 2.** Correlation coefficient between the preceding weekly mean 10–60d signals of (left column) OLR, (mid column) U850, and (right column) U200 anomalies, and weekly precipitation anomalies over Senegal at different lead times during the period of 1982–2019 from June to September. Correlation coefficients that are statistically significant at the 5% level are shaded.



1-0 are located westward over the Atlantic Ocean. These results in U200 correlations are related with the so-called Tropical Easterly Jet (TEJ) (Nicholson and Grist, 2003; Wu et al., 2009). The intensity and latitudinal position of this easterly jet influence precipitation in West Africa during the monsoon season (boreal summer). Specifically, a strengthening (intensification) of this easterly jet at 200 hPa is linked to increased rainfall over the Sahel and countries like Senegal. A more intense easterly jet promotes the ascent of humid air from the Gulf of Guinea, reinforcing convection and precipitation over the Sahel. The more northerly position of an intense jet (as in the case of weeks 2, 1, and 0) allows for a deeper penetration of the monsoon towards Sahelian latitudes, thus favoring precipitation. It appears that U850 correlations exhibit similar characteristics to U200 correlations in many regions, but with opposite signs, which is consistent with upper- and lower-level circulations associated with monsoon systems (tropical baroclinic circulations). Fig. 3 depicts the correlation between the previous 10 to 60 days mean weekly H850, H500, and H200 signals and Senegal precipitation anomalies at different temporal lags. At weeks 5-4, significantly correlated H850 signals are primarily seen over Africa and progressively spread over the Atlantic. The signals are scattered across the tropics at weekly lead times 2 and 1. A very intense signal appears on the Senegalese coast at week 0. For H500 anomalies, the distribution is somewhat similar to that of H850 correlation signals, with the intense H850 signal on the Senegalese coast at week 0 also found in H500 correlations. In contrast, H200 anomaly correlations seem much more scattered compared to H850 and H500 throughout the weeks. At weeks 1-0, significantly correlated H200 signals are mainly seen over the tropical zone.

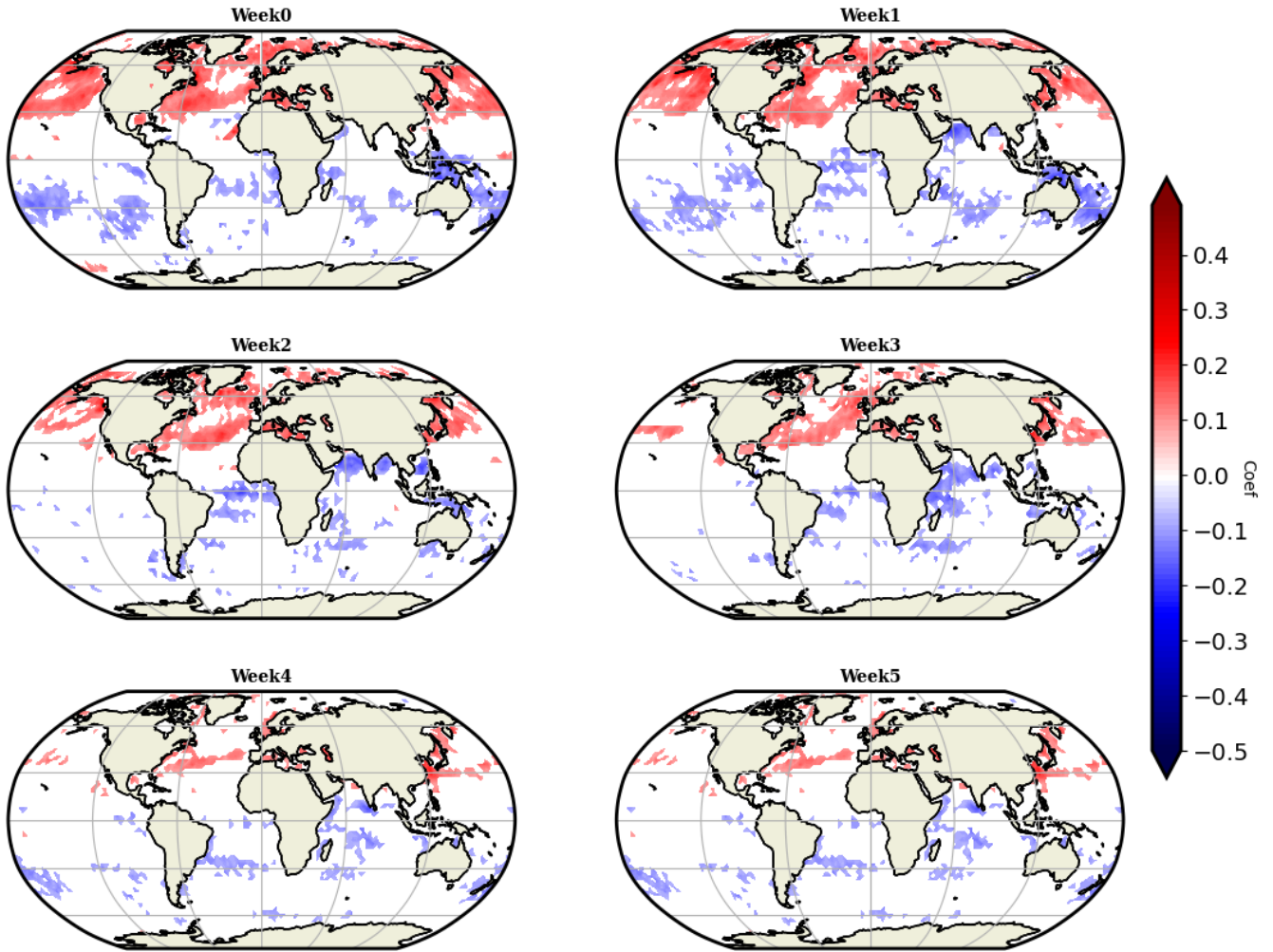


**Figure 3.** Correlation coefficient between the preceding weekly mean 10–60d signals of (left column) H850, (mid column) H500, and (right column) H200 anomalies, and weekly precipitation anomalies over Senegal at different lead times during the period of 1982–2019 from June to September. Correlation coefficients that are statistically significant at the 5% level are shaded.

Fig. 4 displays the correlation between the previous 10 to 60 days mean weekly SST signals and Senegal precipitation anomalies at different temporal lags. Positive signals are seen throughout the North Atlantic, the Mediterranean Sea, and the



North Pacific during various weekly lead times. These positive signals in the North Atlantic are much more widespread and  
205 intense as the lag decreases. In other words, more robust signals with a greater spatial distribution in the North Atlantic are  
found at weeks 3-0. Significant negative correlations are also identified in certain regions, such as the South Atlantic, the  
eastern Indian Ocean, and some areas of the South Pacific, during different time lags. These results are consistent with studies  
by (Gaetani et al., 2010; Mohino et al., 2011; Fontaine et al., 2011; Diakhaté et al., 2019; Thiam et al., 2024), showing that  
strong and significant positive SST anomalies in the Mediterranean precede a wetter than average summer in the Sahel. Jung  
210 (2006) identified a significant increase in precipitation in the Sahel following the Mediterranean heatwave of 2003. Strong  
associations are also found in the North Atlantic, with correlations exceeding 0.4 in the Gulf Stream region north of 30°N  
(Wang et al., 2012; Monerie et al., 2023; Liu et al., 2014) or in the northwest Pacific: extratropical warming of the northern  
hemisphere indeed induces a significant increase in precipitation in the Sahel through the modification of large-scale meridional  
heat distribution, according to (Park et al., 2015; Suárez-Moreno et al., 2018). Thus, all these results are in perfect agreement  
215 with our findings.



**Figure 4.** Correlation coefficient between the preceding weekly mean 10–60 day SST anomaly signals and weekly precipitation anomalies over Senegal at different lead times during the period from 1982 to 2019, from June to September. Correlation coefficients that are statistically significant at the 5% level are shaded.

The spatiotemporal coupled covariance patterns are then constructed for a grid point where the correlation is statistically significant at the 5% level. The predictor is defined by calculating the sum of the products of the covariance patterns and the 10- to 60-day signals of atmospheric and oceanic fields for each preceding weekly interval following (Li et al. (2022)):

$$\text{cov}(X_{i,k}, Y) = \frac{1}{T} \sum_{t=1}^T (Y_t - E(Y))(X_{i,k,t} - E(X_{i,k})) \quad (4)$$



$$220 \quad X_k = \sum_{i=1}^N \text{cov}(X_{i,k}, Y) \cdot X_{i,k} \quad (5)$$

$X_{i,k}$  denotes the weekly mean 10-60 day filtered signal of the  $k^{\text{th}}$  atmospheric and oceanic field at grid point  $i$ , where the correlation between  $X_{i,k}$  and  $Y$  is statistically significant at the 5% level. Here,  $k$  ranges from 1 to 7, representing 6 different atmospheric fields and the SST field.  $Y$  denotes the weekly mean precipitation anomalies.  $T$  is the total number of weeks, and  $N$  is the total number of grid points at which the correlation between  $X_{i,k}$  and  $Y$  is statistically significant at the 5% level.

225 Thus, for each atmospheric or oceanic field and each preceding week, there is only one predictor  $X_k$ .

### 2.2.3 Machine Learning Modeling

In the previous steps, predictors were defined by analyzing the relationship between global intraseasonal (weekly) signals, from atmospheric and oceanic fields, and precipitation anomalies in Senegal. The derived predictors can be used to forecast weekly precipitation anomalies in Senegal. Once the predictors were defined, ML models were constructed. Various modeling  
230 techniques were utilized, including Ridge regression, linear regression, random forests, multi-layer perceptron, and support vector machines.

**Linear regression (LR)** is a statistical method used to assess the linear relationship between a dependent variable  $y$  and one or more independent variables  $X$  (Draper, 1998). The linear regression model is written as:

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \varepsilon \quad (6)$$

235 where  $\beta_i$  are the regression coefficients and  $\varepsilon$  is the error term.

The coefficients  $\beta_i$  are estimated using the least squares method, which minimizes the sum of the squared residuals between the observed  $y$  and modeled  $y$  values.

In general, linear regression is used when the relationship between the predictors  $X$  and outcome  $y$  can be reasonably approximated as linear (Weisberg, 2005). It works well when the predictors are not too highly correlated. The model can be  
240 used for prediction, inference and interpretation of the impact of each predictor on the outcome.

Linear regression is widely used in climate science for tasks like modeling temperature changes over time or predicting precipitation levels based on atmospheric variables. The interpretability of the model makes it a popular choice.

**Ridge regression (also known as regularized regression)** is a method that adds a regularization term to the cost function of linear regression in order to reduce the variance of the model (Hoerl and Kennard, 1970). The regularization term introduces  
245 bias into the estimates of the regression coefficients, but reduces variance by pulling them towards zero. The ridge model cost function is written as:

$$\min \|X\beta - y\|^2 + \lambda \|\beta\|^2 \quad (7)$$

where  $\lambda$  is a hyperparameter that controls the intensity of regularization. Several techniques can be used to estimate its optimal value, such as cross-validation (Golub et al., 1979).

250 In general, ridge regression outperforms simple linear regression in the case of multicollinearity between predictors, as it reduces the instability of coefficient estimates (Marquardt and Snee, 1975). Additionally, it is well-suited for high-dimensional problems by limiting overfitting risks (Zou and Hastie, 2005). This is why it is often used in climate science with a large number of variables. Ridge regression was used to model the relationship between CO<sub>2</sub> concentrations and temperature (Gregory et al., 2004).

255 **Support vector machine (SVM)** is a supervised ML algorithm and can be used for both classification and regression (Vapnik et al., 1998). SVM uses kernels function which can be linear or polynomial in order to obtain non-linear function (Steve, 1998). SVM minimizes the error by adding the hyperplane and maximizing the margin between the prediction and the actual values (Karatzoglou et al., 2006). This model presents many advantages. It is very effective even with high dimensional data. It also works very well if the classes in the data are well separated points. SVM can also work with image data. Nevertheless, SVM  
260 model has some disadvantages. In fact, it is not easy to choose a good kernels function. Moreover, the training period has to be very long for large datasets. It is also difficult to understand and interpret the final model, the weights of the variables and the individual impact and to fine-tune these hyperparameters (Kirchner and Signorino, 2018).

**Random forests (RF)** are an ensemble method based on decision trees. The principle is to construct a collection of decision trees, each trained on a bootstrap sample of the original data (resampling with replacement) (Breiman, 2001). In addition, for  
265 each split of a tree, only a random subset of the predictors is considered. Three hyperparameters need to be tuned in the RF algorithm, namely the number of trees in the "forest", the number of features to consider when looking for the best split, and the maximum depth of the tree. In general, RF generates a higher quality global model than single decision tree models (Mutanga et al., 2012), because RF compensates for the bias introduced by the single decision tree due to its random character. Moreover, RF also shows efficiency in handling large dimensional datasets (Vincenzi et al., 2011), which helps to analyze the dataset of  
270 this study.

**AdaBoost, or "Adaptive Boosting,"** is an ensemble learning method that enhances the accuracy of prediction models by combining multiple weak models (often called "weak learners") to create a strong model. This technique is particularly effective for classification tasks, although it can also be adapted for regression. AdaBoost uses weak learners, typically shallow decision trees (stumps), as base models. The models are trained sequentially, with each model attempting to correct the errors  
275 of previous models. The model predictions are combined with weighting to produce the final prediction.

**Multilayer Perceptron (MLP)** is a type of artificial neural network capable of learning nonlinear relationships between variables (Rosenblatt, 1958). The MLP consists of an input layer, one or more hidden layers, and an output layer that are fully connected. Each neuron computes a linear combination of the inputs with associated weights, and applies a nonlinear activation function like the tanh or relu. The weights are adjusted through backpropagation of the error gradient (Rumelhart  
280 et al., 1986). The MLP can approximate any continuous function thanks to its hidden layers (Cybenko, 1989). It is able to handle complex classification and regression problems. In climatology, the MLP is used for reconstructing missing data



(Suzuki, 2011), precipitation forecasting (Hung et al., 2009), and solar radiation modeling (Li et al., 2020). Its ability to capture nonlinear relationships makes it a powerful tool.

The choice of these models was based on the performance they demonstrated in the studies conducted by (Gerlitz et al., 2016; Cai et al., 2019; Toure et al., 2023; Sarr and Sultan, 2023). These references demonstrated the higher predictive effectiveness of these models compared to other alternatives. Therefore, we selected these particular ML models for our own analysis. In a recent study, PMM (Predictive Mean Matching), RF, and NORM (Bayesian Linear Regression) were used for multiple imputation to improve climate databases in Senegal (Toure et al., 2023). The results highlight the superior performance of the RF model in terms of accuracy and explained variance compared to other ML models. Sarr and Sultan (2023) used these ML techniques to predict crop yields in Senegal. Their results showed that combining climate and vegetation data with ML methods yields the best performance.

To determine optimal hyperparameters for selected ML models, we used a grid search method. This approach consists of exhaustively testing all possible combinations of hyperparameters from a predefined grid of values. Although computationally expensive, grid search has the advantage of being simple to implement. It allowed us to systematically identify the best hyperparameter configuration to maximize the performance of models. We used a one-year cross-validation (or k-fold cross-validation) to evaluate the machine learning models and assess the predictive performance of individual and combined predictors. Cross-validation is a procedure used to estimate the performance of a machine learning algorithm when making predictions on data not used during the training of the model. The cross-validation has a single hyperparameter “k” (here k = one-year) that controls the number of subsets that a dataset is split into. Once split, each subset is given the opportunity to be used as a test set while all other subsets together are used as a training dataset. This means that k-fold cross-validation involves fitting and evaluating k models. This, in turn, provides k estimates of a model’s performance on the dataset, which can be reported using summary statistics such as the mean and standard deviation. Then, to compare with the S2S models, we refined our approach by selecting the two best-performing models, Ridge and linear regression (LR). In addition to achieving the best results, these models require fewer computational resources, allowing us to apply the Leave-One-Year-Out (LOYO) validation without encountering limitations related to computing capacity. Although LOYO validation is a computationally demanding method, it provides a reliable and unbiased estimate of model performance. Leave-one-out cross-validation is a specific configuration of k-fold cross-validation, where k is equal to the number of examples in the dataset. LOYO represents an extreme version of this approach, involving the highest computational cost. Indeed, it requires training and evaluating a model for each example in the training set. The advantage of such a large number of evaluations is that it provides a more robust estimate of model performance, as each observation has the opportunity to represent the entire test dataset. Thus, with a dataset covering 38 years, thirty-eight training and validation processes were conducted using the LOYO method.

#### 2.2.4 Prediction skill-scores

We conducted leave-one-year-out cross-validation to assess the practicality of the models, i.e., using all years’ data from 1982 to 2019 except the target year to train the model and then make a prediction for the target year. This approach is an extensively used cross-validation method because of its simplicity, universality, and superiority in avoiding the issue of over-fitting. In



this study, we utilize the mean absolute error (MAE) to provide an overall assessment of the forecast accuracy for weekly precipitation anomalies. The MAE is calculated as follows:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (8)$$

where  $n$  is the total number of data (sample size),  $y_i$  is the actual value of the  $i$ -th data point, and  $\hat{y}_i$  is the corresponding predicted value. The mean absolute error, which quantifies the absolute difference between the values predicted by the model and the observations, is a positive metric. Thus, the closer the value of this average error tends towards zero, the more the model's predictive performance is judged to be excellent. A low MAE therefore demonstrates high accuracy of the forecasts generated by the model compared to actual measurements.

Additionally, we used the Pearson correlation coefficient to evaluate the linear relationship between the predicted and observed values. The Pearson correlation coefficient is calculated as follows:

$$r = \frac{\sum_{i=1}^n (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2} \sqrt{\sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})^2}} \quad (9)$$

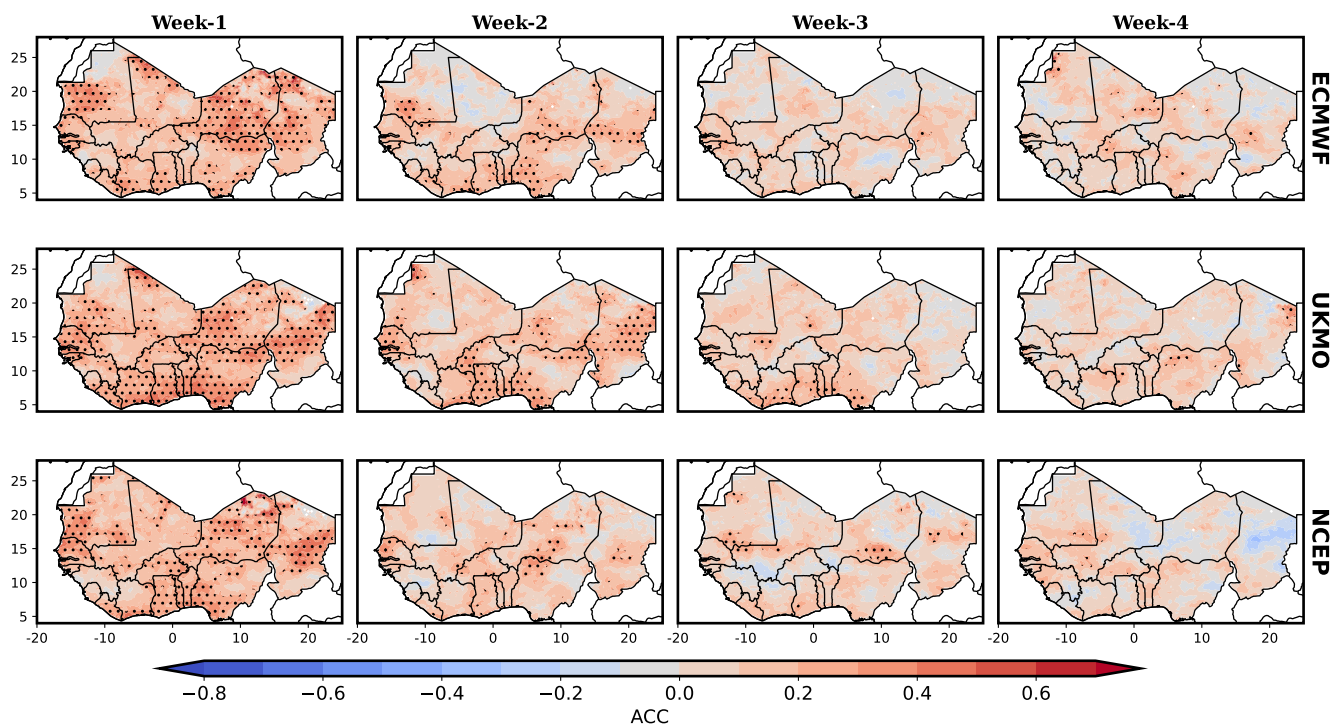
where  $\bar{y}$  and  $\bar{\hat{y}}$  are the means of the actual and predicted values, respectively. The Pearson correlation coefficient,  $r$ , ranges from -1 to 1, where values closer to 1 indicate a strong positive linear relationship, values closer to -1 indicate a strong negative linear relationship, and values around 0 indicate no linear relationship.

By computing both the MAE and the Pearson correlation coefficient (or anomaly correlation coefficient - ACC), we can provide a comprehensive evaluation of the model's performance in predicting weekly precipitation anomalies.

### 3 Results

#### 3.1 Assessment of S2S Model Performance for Precipitation Forecasting

In this section, we evaluate the quality of weekly precipitation hindcasts over West Africa using deterministic forecast verification metrics. The assessment covers the period from May to September during 1999-2010, focusing on three selected S2S models, namely, ECMWF, UKMO, and NCEP (for details on these models, see Section 2.1.4). Figure 5 illustrates the correlation between the hindcast ensemble mean and observed anomalies of accumulated precipitation for each S2S model at different weekly lead times. The highest associations are observed at week 1, with correlations decreasing as the lead time increases. Significant correlations are primarily found in countries such as Senegal, Mauritania, Niger, Nigeria, and Burkina Faso, notably during the first and second weeks, with positive correlations persisting across all lead times in Senegal. Weak associations are seen for lead times from 3 to 4 weeks ahead, particularly over the Sahel. A comparison between models reveals that UKMO outperforms ECMWF and NCEP in several Sahelian areas, including Senegal, Chad, and the coastal countries of the Gulf of Guinea. This finding aligns with the results of de Andrade et al. (2021), who demonstrated significant correlations up to week 4 over West Africa near the Gulf of Guinea for almost all models.

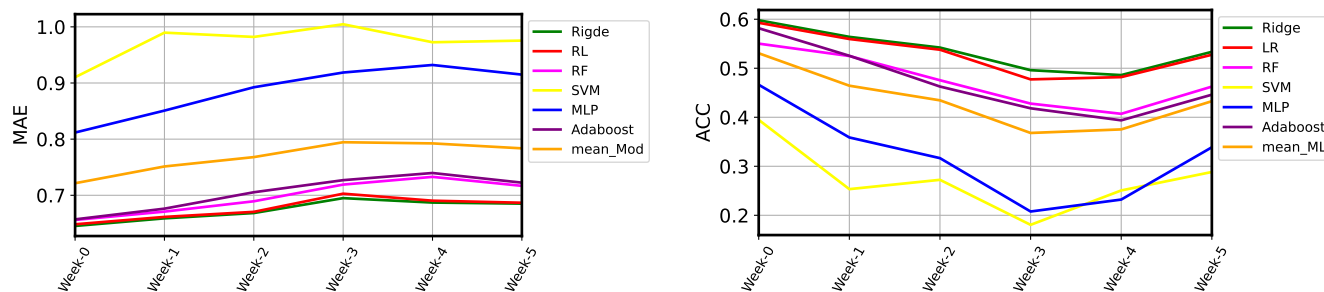


**Figure 5.** Anomaly correlation coefficient (ACC) between hindcast ensemble mean and observed accumulated precipitation anomalies (CHIRPS) for each S2S model (rows) during weeks 1 to 4 (columns). Models are initialized from May to September over the 1999-2010 period in West African countries. Stippling indicates statistically significant values at the 95% confidence level.

### 345 3.2 Comparative Analysis of Machine Learning Models for Precipitation Forecasting

We use six ML algorithms, detailed in the methods section: Ridge, LR, RF, SVM, Adaboost, and MLP (see section 2.2.3 for details). We then calculate the ensemble average of these models (hereinafter *mean\_ML*). Figure 6 shows the MAE and ACC of weekly precipitation anomalies predicted by the different ML algorithms, as well as their ensemble average. The results show higher predictive ability for Ridge Regression in terms of ACC and MAE for all forecast intervals, followed by the LR model. This suggests that the Ridge method further enhances prediction skill across all intervals, yielding this technique for our subsequent analysis.

350



**Figure 6.** Cross-validated scores for forecasting weekly precipitation anomalies at different lead times during JJAS (1982-2019) in Senegal. Left: MAE. Right: ACC. Colors represent different ML models and their ensemble mean (*mean\_ML*).

### 3.3 Predictive Performance of Individual and Combined Predictors

To better understand the primary sources of predictability for intra-seasonal precipitation, we employ the Ridge regression model for each predictor (i.e., atmospheric and/or oceanic fields) individually. Figure 7 presents a comparison of MAE and ACC for forecasts of weekly precipitation anomalies obtained from different predictors. Overall, predictors such as OLR, U200, and U850 demonstrate high predictive ability (low MAE values and high ACC values) compared to predictors such as H200, H500, and H850 for nearly all lead times. This finding aligns with the ACC values shown in Fig. 3. The skillful predictors (OLR, U200, and U850) exhibit stronger linear relationships with precipitation, with OLR showing the strongest association. These results suggest that ISO signals from OLR, U200, and U850 contribute significantly to the prediction skill of sub-seasonal precipitation. Notably, as we move to longer timescales, ACC scores are generally influenced not only by OLR (as seen for shorter leads, such as week 0) but also increasingly by other variables, particularly SST (e.g., highest ACC for week 5), which is physically consistent, as the ocean operates on longer timescales due to its higher heat capacity and lower thermal inertia. For lead times of 2-4 weeks, U200 and U850 also exert some influence, which aligns with the fact that atmospheric circulation (i.e., winds) impact weather patterns on the ground. Comparing the Ridge regression model built with a single predictor to the one built with all atmospheric predictors (denoted as *All.atm*), we find that the latter enhances forecasting capability. This improvement is even more evident when all atmospheric and oceanic predictors (denoted as *All* in Figure 7) are incorporated, thereby enhancing the predictive ability of the model. These results indicate enhanced forecast accuracy of intraseasonal precipitation with the simultaneous inclusion of different predictors, suggesting the relative relevance of oceanic and atmospheric variables under longer and shorter lead times, respectively.



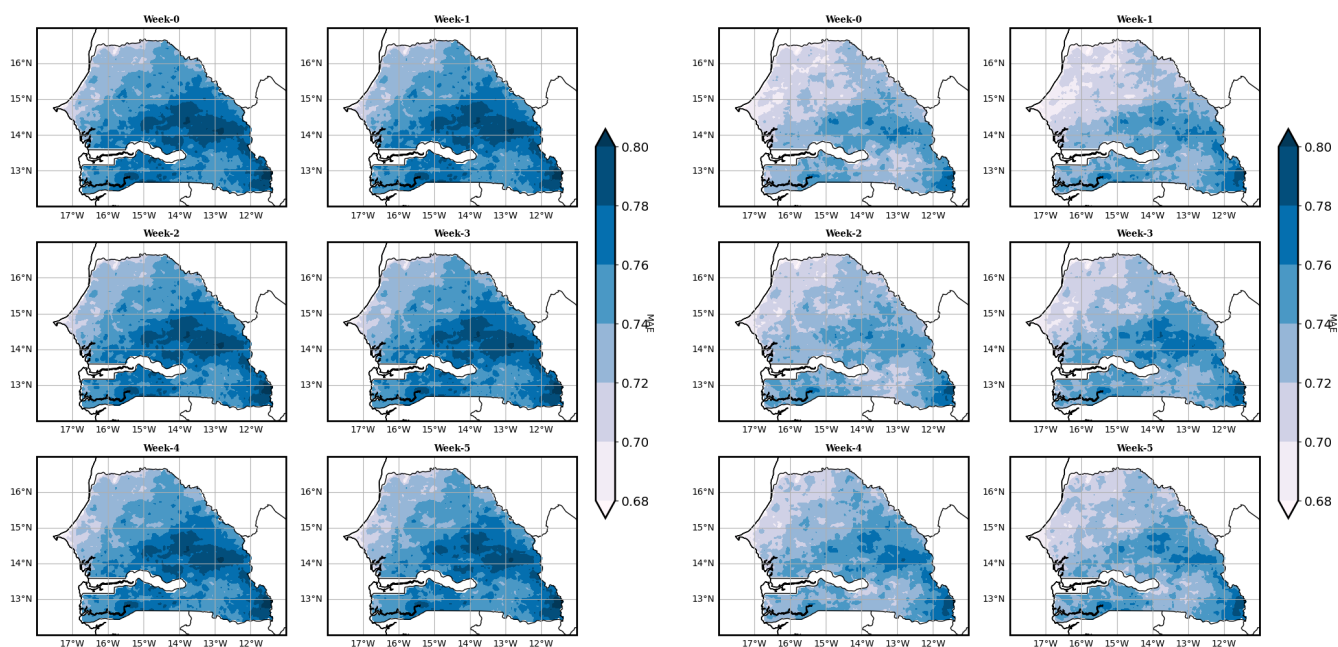
**Figure 7.** Cross-validated scores for weekly precipitation anomaly forecasts during JJAS 1982-2019 in Senegal. Left: MAE. Right: ACC. Rows represent different lead times. Columns show individual predictors, "All\_atm" (all atmospheric signals), and "All" (all atmospheric and oceanic signals).

### 370 3.4 Spatial Variability in Forecast Accuracy for Weekly Precipitation Anomalies

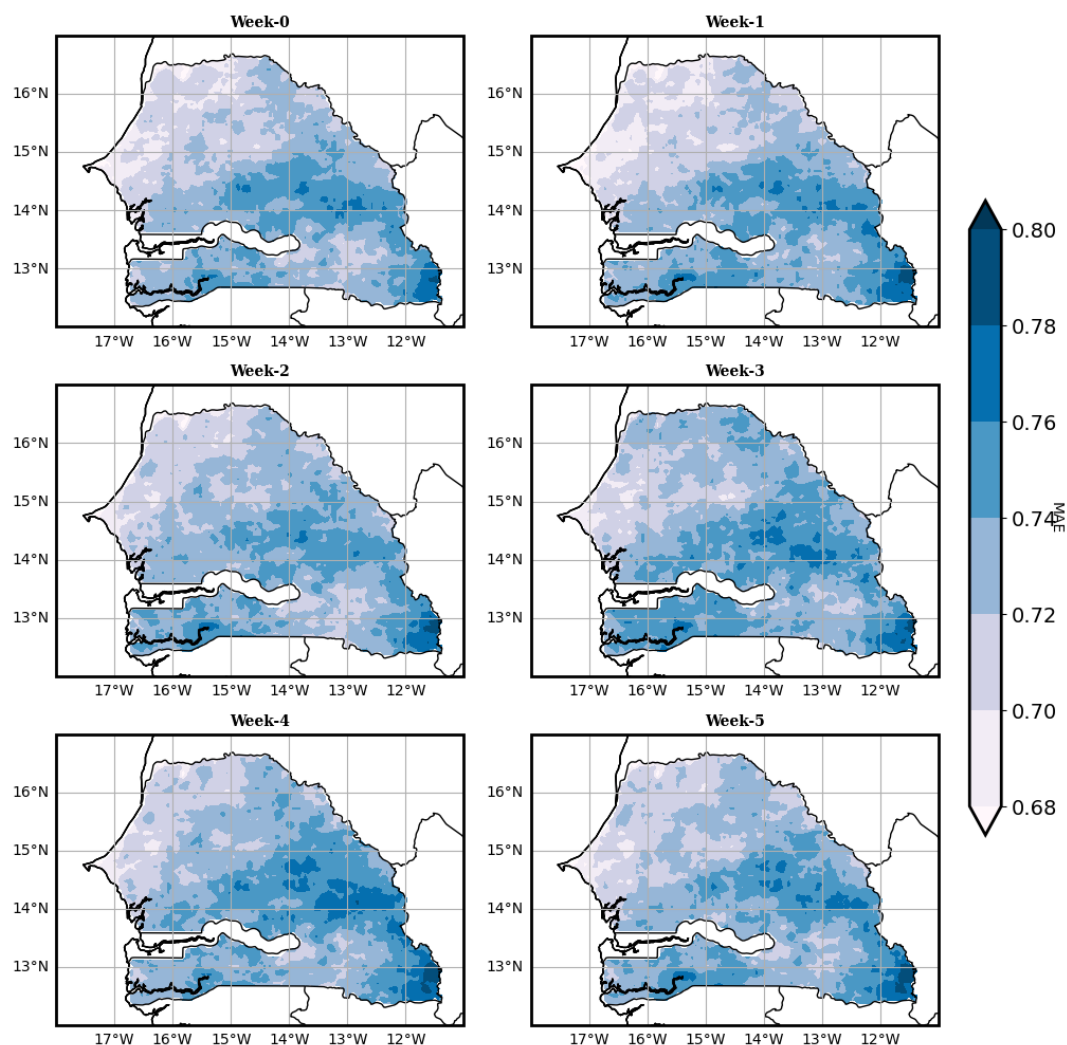
In this section, we focus on the spatial aspects of prediction skill for weekly precipitation anomalies in Senegal, the core region of interest. Figure 8 illustrates the cross-validated MAE obtained from a Ridge regression model for subseasonal forecasts of weekly precipitation anomalies. The analysis covers the period from June to September, 1982 to 2019, across Senegal under different weekly lead times. The 6-panel block on the left depicts the results using intraseasonal SST predictors. A northwest-southeast gradient of MAE values is observed, consistent across lead times. Relatively lower MAE values are consistent in the northwestern area of Senegal across the different lead times, indicating higher model performance. In contrast, the southeastern part of the country exhibits higher MAE values, showing moderate prediction skill. The apparent improvement in model performance as lead time increases is attributed to the longer-term predictive capability of ocean-atmosphere interactions compared to local atmospheric predictors. The block to the right in Figure 8 presents the cross-validated MAE using solely intraseasonal atmospheric variables as predictors. The spatial pattern follows a similar northwest-southeast gradient as the previous case for the oceanic predictors. Model performance improves under shorter lead times and in the northwest region with a gradual decline towards the southeast. Figure 9 shows the results of evaluating the performance of the Ridge regression model using both intraseasonal atmospheric and oceanic variables as predictors. As in the previous cases, cross-validated MAE scores reveal a persistent northwest-southeast gradient for all lead times. Decreased predictive ability (increased MAE scores) is observed as lead time increases. These spatial variations can be attributed to the differential contributions between atmospheric and oceanic variables according to lead times and impact regions. Given the similarity between Figure 9 and the atmospheric predictors-only block in Figure 8 (right panels), atmospheric signals appear to play a more prominent role in precipitation forecasting compared to SST variations when it comes to the subseasonal time scale (up to week 5 lead time). This is attributed to longer lead times in the significant response of precipitation to SST forcing, particularly in the



390 southeastern area of Senegal. When assessing the weekly precipitation forecast in Senegal using the ACC metric with all predictors (supplementary material), we observe a significant decline in ACC as the forecast lead time increases, which further underscores the challenges in maintaining high prediction skill at longer lead times.



**Figure 8.** Spatial variability of cross-validated MAE using the Ridge regression model for subseasonal forecasts of weekly precipitation anomalies during JJAS (1982-2019) in Senegal under different lead times (indicated above the panels). Results are shown in two 6-panel blocks: (Left) using SST predictors; (right) using all atmospheric signals (*All.atm*) as predictors.



**Figure 9.** Spatial variability of cross-validated MAE using the Ridge regression model for subseasonal forecasts of weekly precipitation anomalies during JJAS (1982-2019) in Senegal under different lead times (indicated above the panels). Predictors include both intraseasonal atmospheric and oceanic variables.

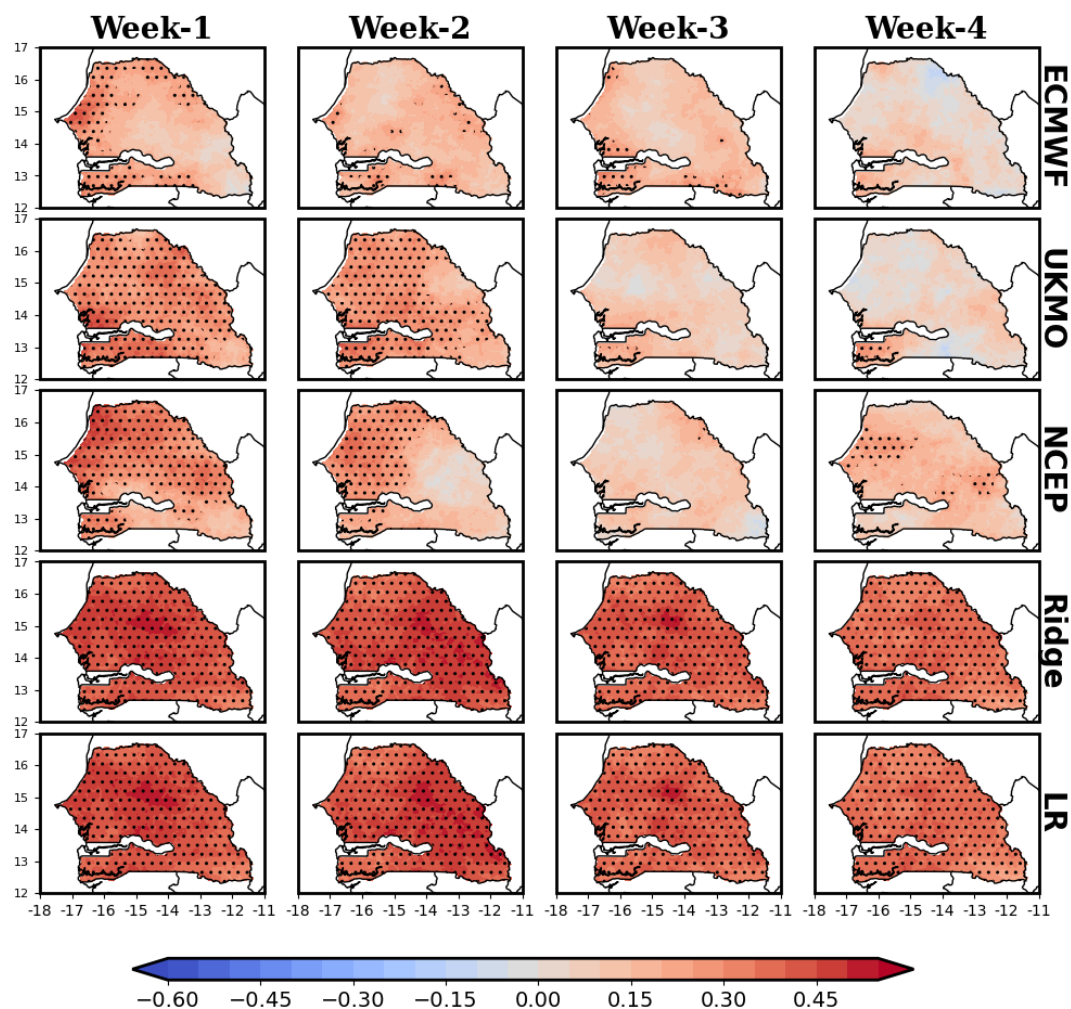
395 However, the model performance evaluation using LOYO validation (see supplementary material) reveals a similar MAE gradient but with significantly lower values than those obtained with standard cross-validation (Figure 9). This indicates that models trained with LOYO validation are considerably more effective. This improvement can be attributed to the fact that LOYO validation leverages a more diverse training set by excluding an entire year at each iteration, allowing the model to be less biased by year-specific variations and better generalize to unseen data. In contrast, standard cross-validation, by randomly mixing training and test data, can overestimate model performance by incorporating temporally close information between training and validation. The same result is observed when considering other metrics such as ACC and RMSE (see



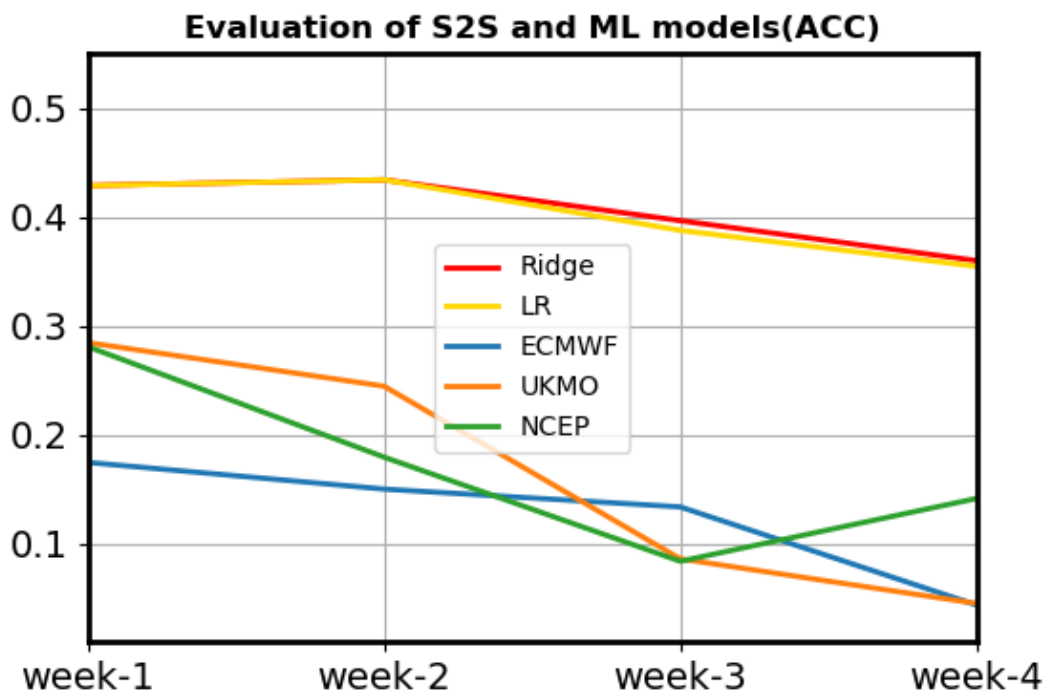
400 supplementary material). This is why we used the outputs of models trained with LOYO validation to compare them with our three models from the S2S database.

### 3.5 Comparative Analysis of ML and S2S Models for Precipitation Forecasting

This section presents a comprehensive comparison between the top-performing ML models and S2S models in forecasting precipitation in Senegal. Figure 10 illustrates the ACC between the ML approaches that relatively show the best predictive ability (Ridge, LR) and GCM-based models from the S2S database (ECMWF, UKMO, and NCEP). We noticed that UKMO seems more skillful than ECMWF and NCEP over Senegal for weeks 1 and 2. For all S2S models (GCM-based models), ACC scores show a consistent degradation as forecast lead time increases. Statistically significant ACC values are observed during weeks 1 and 2, spreading over much of Senegal, particularly for the UKMO and NCEP models, while ECMWF's predictive skill is limited even under short lead times, being confined mainly to the coastal band. The remarkable improvement in skill is evident for ML-based models (Ridge, LR), which exhibit significant positive ACC values for all lead times, extending to the entire country. The improvement with respect to S2S models is outstanding as the prediction horizon increases, placing ML models as an efficient complement or even alternative to GCM-based subseasonal forecast systems. These performance differences are further explored in Figure 11. It depicts the regionally averaged ACC scores, as a function of forecast lead times, for both S2S and ML models. The result unequivocally shows that ML models outperform S2S.



**Figure 10.** Comparison of spatial ACC scores between S2S and top-performing ML models calculated between precipitation hindcasts and observations for the June-September 1999-2010 study period. From top to bottom row, S2S models are the ECMWF, UKMO and NCEP, respectively. ML models are Ridge and RL, respectively. From left to right, each column represents weekly forecast lead time from 1 to 4. Stippling indicates statistically significant ACC at the 95% confidence level.



**Figure 11.** Regionally averaged ACC scores as a function of forecast lead time from 1 to 4 weeks. Ridge and LR (ML models), ECMWF, UKMO and NCEP (S2S models). Different models are represented by colors (see labels within the panels).

415 In the southeastern region of Senegal, all models, particularly the S2S models, exhibit a notable skill degradation. This diminished forecast accuracy stems from the unique geographical characteristics of the region, including the presence of vegetation associated with land-atmosphere interactions that plays a role in precipitation formation. This emphasizes the significance of considering these geographical factors when analyzing and predicting weather phenomena in Senegal, and highlights potential avenues for enhancing the accuracy of prediction models in the future. Our comparative analysis underscores the remarkable  
420 potential of ML approaches in improving subseasonal precipitation forecasting in Senegal. The remarkable performance of these approaches, both spatially and for longer lead times, indicate the added value of ML to provide supplementary insights alongside existing dynamical forecasting systems. In the immediate future, while further work is needed to improve GCMs to achieve a better understanding of the physical mechanisms underlying climate predictability, ML techniques are highly valuable to improve the quality of subseasonal forecast, leading to more reliable predictions of precipitation, crucial for water  
425 resource management, agricultural planning, and risk assessment.

#### 4 Conclusion and Discussion

Through this study, we advance on the field of sub-seasonal to seasonal precipitation forecasting in Senegal by leveraging state-of-the-art ML techniques, namely, Ridge regression, LR, RF, SVM, Adaboost, and MLP. As a starting point, our ap-



proach introduced a comprehensive analysis of the links between various intra-seasonal atmospheric and oceanic signals, and  
430 observed precipitation patterns in Senegal. In this framework, we conducted an in-depth examination of spatiotemporal cor-  
relations between atmospheric (U850, U200, OLR, H850, H500, and H200) and oceanic (SST) variables, and precipitation.  
With this analysis, we identified regions of statistically significant ACC scores, serving to construct spatiotemporal covariance  
models to define our set of predictors through the summation of the product of the covariance fields and intra-seasonal signals.  
To evaluate the performance of ML models, we used MAE and ACC metrics. The results were rigorously compared over  
435 the period 1982-2019 using a leave-one-year-out cross-validation method, showing that ML models provide skillful weekly  
precipitation forecasts for Senegal. Notably, the Ridge regression model consistently outperformed all other models and the  
ensemble mean across almost all lead times. Our analysis of individual sets of predictors, separating between atmospheric  
and oceanic variables, revealed that OLR, U200, and U850 contributed most significantly to improving the prediction skill of  
sub-seasonal precipitation under shorter lead times, particularly OLR. As the prediction horizon increases, oceanic variables  
440 improve their forecast ability, while atmospheric signals alone showed progressive skill degradation. As a third case study, the  
combination of both atmospheric and oceanic predictors enhanced forecasting skill-scores. This shift in the primary drivers of  
predictability from atmospheric to oceanic factors as forecast lead time increases occurs because oceans have a much greater  
heat capacity and thermal inertia than the atmosphere, allowing SST to maintain and gradually release energy over extended  
periods. As a result, oceanic conditions become increasingly important for longer lead times, providing a "memory" to the  
445 climate system that extends predictability beyond the chaotic limits of atmospheric processes alone e.g.,(Mariotti et al., 2018;  
Bach et al., 2019). On the side of the GCM-based models used in our study (NCEP, ECMWF and UKMO), we found a pro-  
nounced degradation in prediction skill as forecast lead time increases, with UKMO showing enhanced performance. Our  
comparative analysis against ML approaches revealed the potential of ML for subseasonal-to-seasonal forecasting, particularly  
for longer lead times, enhancing sub-seasonal forecasting performance compared to GCM-based systems. It is worth noting  
450 that GCMs often struggle to accurately represent the complex climate variability in West Africa, particularly the WAM dy-  
namics (Rodrigues et al., 2014). Specific challenges are in simulating key features like the African Easterly Jet, mesoscale  
convective systems, and the interactions between land surface processes and atmospheric circulation (Roehrig et al., 2013;  
Kniffka et al., 2020). Consequently, GCM-based forecasts for West Africa often exhibit moderate skill scores, particularly for  
precipitation, supporting the results found in this work. While ML models used in this work do not inherently yield physical  
455 causality, they offer valuable insights into the underlying climate dynamics. The careful selection of predictors, combined with  
our existing knowledge of the mechanisms associated with oceanic and atmospheric climate drivers, allows these ML models  
to provide a certain degree of physical interpretability. This approach bridges the gap between purely statistical forecasting  
and process-based understanding, potentially offering new perspectives on the complex interactions governing West African  
climate variability. We identify certain limitations in our approach regarding model robustness, which could be improved by  
460 incorporating predictors to consider land-atmosphere interactions. Thus, future research directions include:

- Integrating ground observations from Senegal to enhance model robustness by considering spatial features. Additional explanatory variables include, but are not limited to topography, vegetation cover and soil moisture.



- Applying localized ML models based on homogeneous climatic zones.
- Using physics informed ML (PIML) models.

465 Our application in Senegal demonstrates the substantial potential of ML for improved sub-seasonal to seasonal forecasting  
in West African countries, bridging the gap between weather and climate predictions. Despite limitations, the operational  
deployment of these models could significantly advance hydrometeorological risk assessment and water resource management,  
with far-reaching implications for the predominant rain-fed agriculture sector. Furthermore, the methods we have evaluated  
can be extended to other regions with different climatic features, allowing for broader application and impact of ML-based  
470 forecasting approaches.

*Code availability.* A software repository leading to the results of this study is freely available at: <https://doi.org/10.5281/zenodo.14894884>  
If you need assistance in reproducing this work, you can refer to the ReadMeS2S and readmeML files, which can be found in the same folder  
as the code.

*Data availability.* CHIRPS data: [https://data.ceda.ac.uk/badc/chirps/data/CHIRPS-2.0/global\\_daily/netcdf/p05](https://data.ceda.ac.uk/badc/chirps/data/CHIRPS-2.0/global_daily/netcdf/p05)

475 OLR data: <https://psl.noaa.gov/data/gridded/data.olrcdr.interp.html>

S2S data: <https://apps.ecmwf.int/datasets/data/s2s-reforecasts-instantaneous-accum-ecmf/levtype=sfc/type=cf/>

SST data: <https://www.ncei.noaa.gov/products/climate-data-records/sea-surface-temperature-optimum-interpolation>

ERA5 data: <https://cds.climate.copernicus.eu/datasets/reanalysis-era5-pressure-levels?tab=overview>

*Author contributions.* Evaluation of S2S models and Grads Software (DF and FMA); original draft preparation, writing guidance and editing  
480 (RSM and DF); machine learning modeling and software development (DF, RL); manuscript review (RSM, MIH, FMA, DW); conceptual-  
ization (DF, RSM, FK, ALD); supervision (ATG)

*Competing interests.* The results of this study are interpreted in an appropriate context. There are no conflicts of interest to declare.

ML Machine Learning WAM West African monsoon S2S subseasonal-to-seasonal GCMs General Circulation Models MAE  
Mean Absolute Error NCEP National Centers for Environmental Prediction ECMWF European Centre for Medium-Range  
485 Weather Forecasts UKMO United Kingdom Met Office U200 Wind at 200 hPa U850 Wind at 850 hPa H850 Geopotential at  
850 hPa H500 Geopotential at 500 hPa H200 Geopotential at 200 hPa ACC Anomaly Correlation Coefficient MCA Maximum  
Covariance Analysis CCA Canonical Correlation Analysis MJO Madden-Julian Oscillation ISO Intra-seasonal Oscillation  
RMM Real-time Multivariate MJO TRMM Tropical Rainfall Measuring Mission NHMS National Hydrological and Meteorological  
Services NOAA National Oceanic and Atmospheric Administration BSISO Boreal Summer Intraseasonal Oscillation



490 SST Sea Surface Temperature OISST Optimum Interpolation Sea Surface Temperature RF Random Forests LR Linear regression AdaBoost Adaptive Boosting, SVM Support vector machine MLP Multilayer Perceptron JJAS June July August September SubX Subseasonal Experiment ISO Intraseasonal Oscillation IPCC Intergovernmental Panel on Climate Change ITCZ Intertropical Convergence Zone AEJ African Easterly Jet CHIRPS Climate Hazards Group Infrared Precipitation with Stations BJP Bayesian joint probability OLR outgoing longwave radiation PIML Physics Informed ML

495 *Acknowledgements.* I would like to express my sincere gratitude to the entire team who contributed to the success of this paper. Your hard work, dedication, and collaboration were invaluable throughout the research process. Thank you for your support and commitment to excellence.



## References

- Abbot, J. and Marohasy, J.: Input selection and optimisation for monthly rainfall forecasting in Queensland, Australia, using artificial neural networks, *Atmospheric Research*, 138, 166–178, 2014.
- Bach, E., Motesharrei, S., Kalnay, E., and Ruiz-Barradas, A.: Local atmosphere–ocean predictability: Dynamical origins, lead times, and seasonality, *Journal of Climate*, 32, 7507–7519, 2019.
- Barnston, A. G. and Smith, T. M.: Specification and prediction of global surface temperature and precipitation from global SST using CCA, *Journal of Climate*, 9, 2660–2697, 1996.
- 505 Biasutti, M.: Rainfall trends in the African Sahel: Characteristics, processes, and causes, *Wiley Interdisciplinary Reviews: Climate Change*, 10, e591, 2019.
- Breiman, L.: Random forests, *Machine learning*, 45, 5–32, 2001.
- Brunet, G., Shapiro, M., Hoskins, B., Moncrieff, M., Dole, R., Kiladis, G. N., Kirtman, B., Lorenc, A., Mills, B., Morss, R., et al.: Collaboration of the weather and climate communities to advance subseasonal-to-seasonal prediction, *Bulletin of the American Meteorological Society*, 91, 1397–1406, 2010.
- 510 Cai, Y., Guan, K., Lobell, D., Potgieter, A. B., Wang, S., Peng, J., Xu, T., Asseng, S., Zhang, Y., You, L., et al.: Integrating satellite and climate data to predict wheat yield in Australia using machine learning approaches, *Agricultural and forest meteorology*, 274, 144–159, 2019.
- Cybenko, G.: Approximation by superpositions of a sigmoidal function, *Mathematics of control, signals and systems*, 2, 303–314, 1989.
- 515 de Andrade, F. M., Coelho, C. A., and Cavalcanti, I. F.: Global precipitation hindcast quality assessment of the Subseasonal to Seasonal (S2S) prediction project models, *Climate Dynamics*, 52, 5451–5475, 2019.
- de Andrade, F. M., Young, M. P., MacLeod, D., Hirons, L. C., Woolnough, S. J., and Black, E.: Subseasonal precipitation prediction for Africa: Forecast evaluation and sources of predictability, *Weather and Forecasting*, 36, 265–284, 2021.
- Diakhaté, M., Rodriguez-Fonseca, B., Gómara, I., Mohino, E., Dieng, A. L., and Gaye, A. T.: Oceanic forcing on interannual variability of Sahel heavy and moderate daily rainfall, *Journal of Hydrometeorology*, 20, 397–410, 2019.
- 520 Draper, N. R.: Applied regression analysis bibliography update 1994-97, *Communications in Statistics-Theory and Methods*, 27, 2581–2623, 1998.
- Duchon, C. E.: Lanczos filtering in one and two dimensions, *Journal of Applied Meteorology and Climatology*, 18, 1016–1022, 1979.
- Eden, J., van Oldenborgh, G. J., Hawkins, E., and Suckling, E. B.: A global empirical system for probabilistic seasonal climate prediction, *Geoscientific Model Development*, 8, 3947–3973, 2015.
- 525 Fall, M., Dieng, A. L., Sall, S. M., Sane, Y., and Diakhaté, M.: Synoptic analysis of extreme rainfall event in West Africa: the case of Linguère, *American Journal of Environmental Protection*, 8, 1–9, 2020.
- Faye, D., Kaly, F., Dieng, A. L., Wane, D., Fall, C. M. N., Mignot, J., and Gaye, A. T.: Regionalization of the Onset and Offset of the Rainy Season in Senegal Using Kohonen Self-Organizing Maps, *Atmosphere*, 15, 378, 2024.
- 530 Fontaine, B., Gaetani, M., Ullmann, A., and Roucou, P.: Time evolution of observed July–September sea surface temperature-Sahel climate teleconnection with removed quasi-global effect (1900–2008), *Journal of Geophysical Research: Atmospheres*, 116, 2011.
- Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., et al.: The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes, *Scientific data*, 2, 1–21, 2015.



- 535 Gaetani, M., Fontaine, B., Roucou, P., and Baldi, M.: Influence of the Mediterranean Sea on the West African monsoon: Intraseasonal variability in numerical simulations, *Journal of Geophysical Research: Atmospheres*, 115, 2010.
- Gerlitz, L., Vorogushyn, S., Apel, H., Gafurov, A., Unger-Shayesteh, K., and Merz, B.: A statistically based seasonal precipitation forecast model with automatic predictor selection and its application to central and south Asia, *Hydrology and Earth system sciences*, 20, 4605–4623, 2016.
- 540 Golub, G. H., Heath, M., and Wahba, G.: Generalized cross-validation as a method for choosing a good ridge parameter, *Technometrics*, 21, 215–223, 1979.
- Gregory, J. M., Ingram, W. J., Palmer, M., Jones, G. S., Stott, P., Thorpe, R., Lowe, J. A., Johns, T., and Williams, K.: A new method for diagnosing radiative forcing and climate sensitivity, *Geophysical research letters*, 31, 2004.
- Hoerl, A. E. and Kennard, R. W.: Ridge regression: Biased estimation for nonorthogonal problems, *Technometrics*, 12, 55–67, 1970.
- Huang, B., Liu, C., Banzon, V., Freeman, E., Graham, G., Hankins, B., Smith, T., and Zhang, H.-M.: Improvements of the daily optimum interpolation sea surface temperature (DOISST) version 2.1, *Journal of Climate*, 34, 2923–2939, 2021.
- 545 Hung, N. Q., Babel, M. S., Weesakul, S., and Tripathi, N.: An artificial neural network model for rainfall forecasting in Bangkok, Thailand, *Hydrology and Earth System Sciences*, 13, 1413–1425, 2009.
- Hwang, S.-O., Schemm, J.-K. E., Barnston, A. G., and Kwon, W.-T.: Long-lead seasonal forecast skill in far eastern Asia using canonical correlation analysis, *Journal of Climate*, 14, 3005–3016, 2001.
- 550 Janicot, S., Thorncroft, C. D., Ali, A., Asencio, N., Berry, G., Bock, O., Bourlès, B., Caniaux, G., Chauvin, F., Deme, A., et al.: Large-scale overview of the summer monsoon over West Africa during the AMMA field experiment in 2006, in: *Annales Geophysicae*, vol. 26, pp. 2569–2595, Copernicus Publications Göttingen, Germany, 2008.
- Jung, G.: Regional climate change and the impact on hydrology in the Volta Basin of West Africa, Publish/report a document, 2006.
- Karatzoglou, A., Meyer, D., and Hornik, K.: Support vector machines in R, *Journal of statistical software*, 15, 1–28, 2006.
- 555 Kirchner, A. and Signorino, C. S.: Using support vector machines for survey research, *Survey Practice*, 11, 2018.
- Kniffka, A., Knippertz, P., Fink, A. H., Benedetti, A., Brooks, M. E., Hill, P. G., Maranan, M., Pante, G., and Vogel, B.: An evaluation of operational and research weather forecasts for southern West Africa using observations from the DACCIWA field campaign in June–July 2016, *Quarterly Journal of the Royal Meteorological Society*, 146, 1121–1148, 2020.
- Lee, J.-Y., Wang, B., Wheeler, M. C., Fu, X., Waliser, D. E., and Kang, I.-S.: Real-time multivariate indices for the boreal summer intraseasonal oscillation over the Asian summer monsoon region, *Climate Dynamics*, 40, 493–509, 2013.
- 560 Leung, J. C.-H. and Qian, W.: Monitoring the Madden–Julian oscillation with geopotential height, *Climate Dynamics*, 49, 1981–2006, 2017.
- Li, P., Bessafi, M., Morel, B., Chabriat, J.-P., Delsaut, M., and Li, Q.: Daily surface solar radiation prediction mapping using artificial neural network: the case study of Reunion Island, *Journal of Solar Energy Engineering*, 142, 021 009, 2020.
- Li, Y., Wu, Z., He, H., Wang, Q. J., Xu, H., and Lu, G.: Post-processing sub-seasonal precipitation forecasts at various spatiotemporal scales across China during boreal summer monsoon, *Journal of Hydrology*, 598, 125 742, 2021.
- 565 Li, Y., Wu, Z., He, H., and Yin, H.: Probabilistic subseasonal precipitation forecasts using preceding atmospheric intraseasonal signals in a Bayesian perspective, *Hydrology and Earth System Sciences*, 26, 4975–4994, 2022.
- Liebmann, B. and Smith, C. A.: Description of a complete (interpolated) outgoing longwave radiation dataset, *Bulletin of the American Meteorological Society*, 77, 1275–1277, 1996.
- 570 Liu, Y., Chiang, J. C., Chou, C., and Patricola, C. M.: Atmospheric teleconnection mechanisms of extratropical North Atlantic SST influence on Sahel rainfall, *Climate dynamics*, 43, 2797–2811, 2014.



- Mariotti, A., Ruti, P. M., and Rixen, M.: Progress in subseasonal to seasonal prediction through a joint weather and climate community effort, *Npj Climate and Atmospheric Science*, 1, 4, 2018.
- Marquardt, D. W. and Snee, R. D.: Ridge regression in practice, *The American Statistician*, 29, 3–20, 1975.
- 575 Mohino, E., Janicot, S., and Bader, J.: Sahel rainfall and decadal to multi-decadal sea surface temperature variability, *Climate dynamics*, 37, 419–440, 2011.
- Monerie, P.-a., Biasutti, M., Mignot, J., Mohino, E., Pohl, B., and Zappa, G.: Storylines of Sahel Precipitation Change: Roles of the North Atlantic and Euro-Mediterranean Temperature, *Journal of Geophysical Research: Atmospheres*, 128, e2023JD038 712, 2023.
- Mutanga, O., Adam, E., and Cho, M. A.: High density biomass estimation for wetland vegetation using WorldView-2 imagery and random  
580 forest regression algorithm, *International Journal of Applied Earth Observation and Geoinformation*, 18, 399–406, 2012.
- Nicholson, S. E. and Grist, J. P.: The seasonal evolution of the atmospheric circulation over West Africa and equatorial Africa, *Journal of climate*, 16, 1013–1030, 2003.
- Park, J.-Y., Bader, J., and Matei, D.: Northern-hemispheric differential warming is the key to understanding the discrepancies in the projected Sahel rainfall, *Nature communications*, 6, 5985, 2015.
- 585 Pegion, K., Kirtman, B. P., Becker, E., Collins, D. C., LaJoie, E., Burgman, R., Bell, R., DelSole, T., Min, D., Zhu, Y., et al.: The Subseasonal Experiment (SubX): A multimodel subseasonal prediction experiment, *Bulletin of the American Meteorological Society*, 100, 2043–2060, 2019.
- Robertson, A. W., Vitart, F., and Camargo, S. J.: Subseasonal to seasonal prediction of weather to climate with application to tropical cyclones, *Journal of Geophysical Research: Atmospheres*, 125, e2018JD029 375, 2020.
- 590 Rodrigues, L. R. L., García-Serrano, J., and Doblas-Reyes, F.: Seasonal forecast quality of the West African monsoon rainfall regimes by multiple forecast systems, *Journal of Geophysical Research: Atmospheres*, 119, 7908–7930, 2014.
- Roehrig, R., Bouniol, D., Guichard, F., Hourdin, F., and Redelsperger, J.-L.: The present and future of the West African monsoon: A process-oriented assessment of CMIP5 simulations along the AMMA transect, *Journal of Climate*, 26, 6471–6505, 2013.
- Rosenblatt, F.: The perceptron: a probabilistic model for information storage and organization in the brain., *Psychological review*, 65, 386,  
595 1958.
- Rumelhart, D. E., Hinton, G. E., and Williams, R. J.: Learning representations by back-propagating errors, *nature*, 323, 533–536, 1986.
- Sane, Y., Panthou, G., Bodian, A., Vischel, T., Lebel, T., Dacosta, H., Quantin, G., Wilcox, C., Ndiaye, O., Diongue-Niang, A., et al.: Intensity–duration–frequency (IDF) rainfall curves in Senegal, *Natural Hazards and Earth System Sciences*, 18, 1849–1866, 2018.
- Sarr, A. B. and Sultan, B.: Predicting crop yields in Senegal using machine learning methods, *International Journal of Climatology*, 43,  
600 1817–1838, 2023.
- Schepen, A., Wang, Q., and Robertson, D.: Evidence for using lagged climate indices to forecast Australian seasonal rainfall, *Journal of Climate*, 25, 1230–1246, 2012.
- Simpson, J., Adler, R. F., and North, G. R.: A proposed tropical rainfall measuring mission (TRMM) satellite, *Bulletin of the American meteorological Society*, 69, 278–295, 1988.
- 605 Steve, G.: Support vector machines for classification and regression, University of Southampton ISIS (Image Speech and Intelligent Systems group) Technical Report, pp. 1–52, 1998.
- Suárez-Moreno, R. and Rodríguez-Fonseca, B.: S 4 CAST v2. 0: sea surface temperature based statistical seasonal forecast model, *Geoscientific model development*, 8, 3639–3658, 2015.



- Suárez-Moreno, R., Rodríguez-Fonseca, B., Barroso, J. A., and Fink, A. H.: Interdecadal changes in the leading ocean forcing of Sahelian rainfall interannual variability: atmospheric dynamics and role of multidecadal SST background, *Journal of Climate*, 31, 6687–6710, 2018.
- Sultan, B., Janicot, S., and Diedhiou, A.: The West African monsoon dynamics. Part I: Documentation of intraseasonal variability, *Journal of Climate*, 16, 3389–3406, 2003.
- Suzuki, K.: Artificial neural networks: methodological advances and biomedical applications, BoD–Books on Demand, 2011.
- 615 Tarnavsky, E., Grimes, D., Maidment, R., Black, E., Allan, R. P., Stringer, M., Chadwick, R., and Kayitakire, F.: Extension of the TAMSAT satellite-based rainfall monitoring over Africa and from 1983 to present, *Journal of Applied Meteorology and Climatology*, 53, 2805–2822, 2014.
- Thiam, M., Oruba, L., De Coetlogon, G., Wade, M., Diop, B., and Farota, A. K.: Impact of the Sea Surface Temperature in the North-Eastern Tropical Atlantic on Precipitation Over Senegal, *Journal of Geophysical Research: Atmospheres*, 129, e2023JD040 513, 2024.
- 620 Totz, S., Tziperman, E., Coumou, D., Pfeiffer, K., and Cohen, J.: Winter precipitation forecast in the European and Mediterranean regions using cluster analysis, *Geophysical Research Letters*, 44, 12–418, 2017.
- Toure, M., Klutse, N. A. B., Sarr, M. A., Kenne, A. D., Bhuiyanr, M. A. E., Ndiaye, O., Badiane, D., Thiaw, W. M., Sy, I., Mbow, C., et al.: A New Multiple Imputation Approach Using Machine Learning to Enhance Climate Databases in Senegal, Research Square Company, 2023.
- 625 Tuel, A. and Eltahir, E. A.: Seasonal precipitation forecast over Morocco, *Water Resources Research*, 54, 9118–9130, 2018.
- Vapnik, V. N., Vapnik, V., et al.: Statistical learning theory, *The Nature of Statistical Learning theory*, 1998.
- Vigaud, N. and Giannini, A.: West African convection regimes and their predictability from submonthly forecasts, *Climate dynamics*, 52, 7029–7048, 2019.
- Vigaud, N., Tippett, M. K., Yuan, J., Robertson, A. W., and Acharya, N.: Spatial correction of multimodel ensemble subseasonal precipitation forecasts over North America using local Laplacian eigenfunctions, *Monthly Weather Review*, 148, 523–539, 2020.
- 630 Vincenzi, S., Zucchetto, M., Franzoi, P., Pellizzato, M., Pranovi, F., De Leo, G. A., and Torricelli, P.: Application of a Random Forest algorithm to predict spatial distribution of the potential yield of *Ruditapes philippinarum* in the Venice lagoon, Italy, *Ecological Modelling*, 222, 1471–1478, 2011.
- Vitart, F. and Robertson, A. W.: Introduction: Why sub-seasonal to seasonal prediction (S2S)?, in: *Sub-seasonal to seasonal prediction*, pp. 3–15, Elsevier, 2019.
- 635 Vitart, F., Robertson, A., Kumar, A., Hendon, H., Takaya, Y., Lin, H., Arribas, A., Lee, J., Waliser, D., Kirtman, B., et al.: Subseasonal to seasonal prediction: Research implementation plan, Tech. rep., WWRP/THORPEX-WCRP Report, 2012.
- Vitart, F., Ardilouze, C., Bonet, A., Brookshaw, A., Chen, M., Codorean, C., Déqué, M., Ferranti, L., Fucile, E., Fuentes, M., et al.: The subseasonal to seasonal (S2S) prediction project database, *Bulletin of the American Meteorological Society*, 98, 163–173, 2017.
- 640 Wang, X. L., Feng, Y., and Swail, V.: North Atlantic wave height trends as reconstructed from the 20th century reanalysis, *Geophysical Research Letters*, 39, 2012.
- Weisberg, S.: *Applied linear regression*, vol. 528, John Wiley & Sons, 2005.
- Wheeler, M. C. and Hendon, H. H.: An all-season real-time multivariate MJO index: Development of an index for monitoring and prediction, *Monthly weather review*, 132, 1917–1932, 2004.
- 645 Wu, M.-L. C., Reale, O., Schubert, S. D., Suarez, M. J., Koster, R. D., and Pegion, P. J.: African easterly jet: Structure and maintenance, *Journal of Climate*, 22, 4459–4480, 2009.

<https://doi.org/10.5194/egusphere-2024-4040>

Preprint. Discussion started: 28 February 2025

© Author(s) 2025. CC BY 4.0 License.



Zhang, C.: Madden-Julian oscillation, *Reviews of Geophysics*, 43, 2005.

Zou, H. and Hastie, T.: Regularization and variable selection via the elastic net, *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 67, 301–320, 2005.