S^4CAST v2.0: Sea Surface Temperature based Statistical Seasonal Forecast Model

R. Suárez-Moreno [1,2] and B. Rodríguez-Fonseca [1,2]

[1] {Departamento de Geofísica y Meteorología, Facultad de Físicas, Universidad Complutense de Madrid, Plaza de las Ciencias 1, 28040 Madrid, Spain}

[2] {Instituto de Geociencias (IGEO), Facultad de Ciencias Geológicas, Universidad Complutense de Madrid - CSIC, C/ José Antonio Novais 12, 28040 Madrid, Spain}

Correspondence to: R. Suárez-Moreno (roberto.suarez@fis.ucm.es)

Abstract

Sea Surface Temperature is the key variable when tackling seasonal to decadal climate forecast. Dynamical models are unable to properly reproduce tropical climate variability, introducing biases that prevent a skillful predictability. Statistical methodologies emerge as an alternative to improve the predictability and reduce these biases. In addition, recent studies have put forward the non-stationary behavior of the teleconnections between tropical oceans, showing how the same tropical mode has different impacts depending on the considered sequence of decades. To improve the predictability and investigate possible teleconnections, the Sea Surface Temperature based Statistical Seasonal foreCAST model (S^4CAST) introduces the novelty of considering the non-stationary links between the predictor and predictand fields. This paper describes the development of S^4CAST model whose operation is focused on studying the impacts of sea surface temperature on any climate-related variable. Two applications focused on analyzing the predictability of different climatic events have been implemented as benchmark examples.

1. Introduction

Global oceans have the capacity to store and release heat as energy that is transferred to the atmosphere altering global atmospheric circulation. Therefore, fluctuations in monthly sea surface temperature (SST) may be considered as an important source of energy affecting seasonal predictability and improving the ability to forecast climate-related variables. Many research works have been conducted to study the impacts of worldwide sea surface temperature anomalies (SSTA) by means of dynamical models, observational studies and statistical methods. In this way, tropical oceans purchase greater relevance (Rasmusson and Carpenter, 1982; Harrison and Larkin, 1998; Klein et al., 1999; Saravanan and Chang, 2000; Trenberth et al., 2002; Chang et al., 2006; Ding et al., 2012; Wang et al., 2012; Ham, 2013a; 2013b; Keenlyside et al., 2013). Because of the persistence shown by SSTA, alterations that occur in the oceans are slower than changes occurring in the atmosphere. Once the thermal equilibrium between the ocean and the atmosphere is broken, oceans are able to release their energy, changing the atmospheric circulation for some time before dissipating, leading in turn to an influence on other variables. This fact explains why the SSTA can be used as potential predictor of the anomalous associated impacts.

The S^4CAST model presented in this work is focused on the study of the predictability and teleconnections of climate-related variables based on the remote influence of the SSTA. It has been shown that such variables can be SST (Rasmusson and Carpenter, 1982; Latif and Barnett, 1995; Harrison and Larkin, 1998; Klein et al., 1999; Trenberth et al., 2002), rainfall (Janicot et al., 2001; Drosdowsky and Chambers, 2001; Giannini et al., 2001, 2003; Rowell, 2001, 2003; Chung and Ramathan, 2006; Haylock et al., 2006; Polo et al., 2008; Joly and Voldoire, 2009; Lu, 2009; Gaetani et al., 2010; Shin and Sardeshmukh, 2010; Fontaine et al., 2011; Nnamchi and Li, 2011; Bulic and Kucharski, 2012; López-Parages and Rodríguez-Fonseca, 2012), and other climate-related variables. In this way, there are studies that have focused on the role of the tropical Pacific on vegetation, crop yields and the economic consequences resulting from these impacts (Hansen et al., 1998, 2001; Adams et al., 1999; Legler et al., 1999; Li and Kafatos, 2000; Naylor et al., 2001; Tao et al., 2004; Deng et al., 2010; Phillips et al., 1998; Verdin et al., 1999; Podestá et al., 1999; Travasso et al., 2009). Regarding human health, tropical SST patterns have been widely linked to the development and propagation of diseases (Linthicum et al., 2010), where El Niño-southern Oscillation (ENSO) related variability plays a crucial role mainly affecting tropical and subtropical regions around the world (Kovats, 2000; Patz, 2002; Kovats et al., 2003; Patz et al., 2005; McMichael et al., 2006).

The study of the impacts of tropical global SST on climate has become increasingly important during the last decades. Thus, there are dynamical and statistical prediction models that attempt to define and predict seasonal averages from interannual to multidecadal time scales. In this way, General Circulation Models (GCMs) emerged from the need to reproduce the ocean-atmosphere interactions, responsible for much of climate variability whose major component is attributed to ENSO phenomenon (Bjerknes, 1969; Gill, 1980). Numerous research centers have done a hard work to create their own prediction systems in which coupled ocean-atmosphere GCMs are used in conjunction with statistical methods to achieve reliable ENSO variability predictions and analyze the skill of these models (Cane et al., 1986; Barnett and Preisendorfer, 1987; Zebiak and Cane, 1987; Barnston and Ropelewski, 1992; Barnett et al., 1993; Barnston et al., 1994, 1999; Ji et al., 1994a, 1994b; Van den Dool, 1994; Mason et al., 1999). Predictability of rainfall has become a scope for these models, finding works that have focused on this issue by means of dynamical and statistical models (Garric et al., 2002; Coelho et al., 2006). However, the difficulty of GCMs to adequately reproduce the tropical climate variability remains a real problem, so that in recent years the number of studies focusing on specific aspects of the biases of these models has increased exponentially (Biasutti et al., 2006; Richter and Xie, 2008; Wahl et al., 2011; Doi et al., 2012; Li and Xie, 2012; Richter et al., 2012; Bellenguer et al., 2013; Brown et al., 2013; Toniazzo and Woolnough, 2013; Vanniere et al., 2013; Xue et al., 2013; Li and Xie, 2014).

Statistical models have been widely used as an alternative way of climate forecasting, including several techniques in their development. Model Output Statistics (MOS) determine a statistical relationship between the predictand and the variables obtained from dynamic models (Glahn and Lowry, 1972; Klein and Glahn, 1974; Vislocky and Fritsch, 1995). Stochastic climate models were defined in the 1970s to be first applied to predict SSTA and thermocline variability (Hasselmann, 1976; Frankignoul and Hasselmann, 1977) and later addressing non-linearity problems (Majda et al., 1999). Moreover, Linear Inverse Modeling (Penland and Sardeshmukh, 1995) has been used in predicting variables such as tropical Atlantic SSTA (Penland and Matrosova, 1998) and the study of Atlantic Meridional Mode (Vimont, 2012). Statistical modeling with neural networks is also applied in climate prediction (Gardner and Dorling, 1998; Hsieh and Tang, 1998; Tang et al., 2000; Hsieh, 2001; Knutti et al., 2003; Baboo and Shereef, 2010; Shukla et al., 2011) with the potential to be a nonlinear method capable of addressing the problems in atmospheric processes that are overlooked in other statistical methodologies (Tang et al., 2000; Hsieh, 2001).

A special mention goes to two linear statistical methods: Maximum Covariance Analysis

(MCA) used in the S^4CAST model and Canonical Correlation Analysis (CCA). These methods have been widely used in seasonal climate forecasting, either to complement dynamical models or to be applied independently. In this way, Climate Predictability Tool (CPT) developed at International Research Institute for Climate and Society (IRI) allows user to apply multivariate linear regression techniques (e.g., CCA) to get their own predictions (Korecha and Barnston, 2007; Recalde-Coronel et al., 2014; Barnston and Tippet, 2014). In essence, these techniques serve to isolate co-variability coupled patterns between two variables that act as predictor and predictand respectively (Bretherton et al., 1992). Based on the ability of the SSTA as predictor field, these methods were originally applied to analyze the predictability of phenomenon like ENSO (Barnston and Ropelewski, 1992), 500-mb height anomalies (Wallace et al., 1992) or global surface temperature and rainfall (Barnston and Smith, 1996). Nevertheless, there are works discussing the use of these methods, focusing on the differences between the two techniques (Cherry, 1996, 1997) and on the limitations in their applications (Newman and Sardeshmukh, 1995).

The co-variability patterns between SSTA themselves might fluctuate from one given study period to another, determining non-stationary behavior along time. In this way, teleconnections associated with El Niño or with the Tropical Atlantic are effective in some periods but not in others. In this way, Rodríguez-Fonseca et al. (2009) suggested how the interanual variability in the equatorial Atlantic could be used as predictor of Pacific ENSO after the 1970's, a theory that has been subsequently reinforced by further analysis (Martín-Rey et al., 2012; 2014; 2015; Polo et al., 2015). The non-stationarity in terms of predictability of rainfall has also been found for West African rainfall (Janicot et al., 1996; Fontaine et al., 1998; Mohino et al., 2011; Losada et al., 2012; Rodriguez-Fonseca et al., 2011; 2015); and Europe (López-Parages and Rodriguez-Fonseca, 2012; López-Parages et al., 2014). Thus, the existence of non-stationarities is a key factor in the development of the statistical model.

The present paper describes a statistical model based on the predictive nature of SSTA treating the stationarity in the relationships between the predictor and predictand fields. Section 2 describes the theoretical framework including the statistical methodology and the significance of the statistical analysis. Section 3 is dedicated to S^4CAST model description including the determination of stationary periods, hindcast and forecast calculations and validation. Section 4 describes two case studies concerning the predictability of Sahelian rainfall and tropical Pacific SSTA.

2. Theoretical framework

2.1. Statistical methodology

Maximum Covariance Analysis (MCA) is a broadly used statistical discriminant analysis methodology based on calculating principal directions of maximum covariance between two variables. This statistical analysis considers two fields, Y (predictor) and Z (predictand) (Bretherton et al., 1992; Cherry, 1997; Widmann, 2005) for applying the Singular Value Decomposition (SVD) to the cross-covariance matrix (C) in order to be maximized. SVD is an algebraical technique that diagonalizes non-squared matrices, as it can be the case of the matrices of the two fields to be maximized.

In the meteorological context, C is dimensioned in time (n_t) and space domains (n_Y and n_Z for Y and Z respectively), although the spatial domain can be more complex depending on the user requirements. SVD calculates linear combinations of the time series of Y and Z, named as expansion coefficients (hereinafter U and V for Y and Z respectively) that maximize C. The expansion coefficients are computed by diagonalization of C. As C is non-squared, diagonalization is first done to $A = CC^T$ and then to $B = C^T C$. The singular vectors R and Q are the resultant eigenvectors from each diagonalization, which are the spatial configurations of the co-variability modes. The associated loadings on time domain are the expansion coefficients U and V. The eigenvalues are a measure of the percentage of variance explained by each mode.

Mathematically, the time anomalies of both, Z and Y fields are calculated by removing the climatological seasonal cycle to the seasonal means.

$$Z' = Z - \overline{Z} \quad (1)$$

$$Y' = Y - \overline{Y} \quad (2)$$

Then, the cross-covariance matrix is calculated as:

$$C_{Y'Z'} = \frac{Y'Z'^T}{(n_t - 1)} \quad (3)$$

MCA diagonalizes (3) by SVD methodology, obtaining the singular vectors R and Q from which the expansion coefficients are obtained according to the following expression:

$$U = R^T Y \tag{4}$$

$$V = Q^T Z \tag{5}$$

Using the eigenvectors, the percentage of explained covariance is calculated as

$$scf_k = \frac{\lambda_k^2}{\sum_i^r \lambda_i^2}; \lambda_k = [\lambda_1, \lambda_2, ..., \lambda_n] \tag{6}$$

Where k is the eigenvalue for each k mode and r represents the number of modes taken into account for the analysis.

The expression from which an estimation of the predictand is obtained is a linear model as:

$$\hat{Z} = \Phi Y \tag{7}$$

Where Φ is the so-called regression coefficient and $\hat{Z}$ denotes an estimation of the data to be predicted (hindcast).

Taking into account that S is the regression map of the field Z onto the direction of U

$$S = UZ^T \tag{8}$$

And assuming good prediction $\hat{Z}$, it follows that

$$S = U\hat{Z}^T \tag{9}$$

Introducing the equality $(UU^T)(UU^T)^{-1} = I$ and multiplying in (9) the following expression is obtained:

$$(UU^T)(UU^T)^{-1} S = U\hat{Z}^T \tag{10}$$

Removing U from both terms

$$\hat{Z} = \left[U^T (UU^T)^{-1} S\right]^T \tag{11}$$

Considering now the expression $U = Y^T R$ it follows that

$$\hat{Z} = YR(UU^T)^{-1} S \tag{12}$$

Comparing this expression with (7) and introducing (8) it can be concluded that

$$\Phi = R\left(UU^T\right)^{-1} UZ^T \tag{13}$$

Which is the regression coefficient to be calculated when defining the linear model from which the predictions and hindcasts will be obtained.

2.2. Statistical field significance

There are many statistical tests to assess the robustness of a result. The S^{4}CAST uses a non-parametric test because, a priori, the model doesn't know the distribution of the predictand field. Thus, applying Monte Carlo testing assesses the robustness of the results and is used to validate the S^{4}CAST model skill. This method involves performing a large number ($N > 500$) of permutations from the original time series. Each permuted time series is used to repeat the calculation and compare the obtained results with the real values. Once this is done, the values obtained with the N permutations are taken to create a random distribution to finally determine the position of the real value within the distribution, which will indicate the statistical significance of the obtained value. This method has been described and used in many previous works (Livezey and Chen, 1987; Barnett, 1995; Maia et al., 2007). The user inputs the level of statistical significance at which the test is applied, being the most used 90% (0.10), 95% (0.05) and 99% (0.01).

3. S^{4}CAST model

S^{4}CAST v2.0 model is conceived as a statistical tool to study the predictability and teleconnections of variables that strongly covary with SSTA variability in remote and nearby locations to a particular region of study. The code has been developed as a MATLAB® toolbox. The software requirements are variable and depend on user needs. The spatial resolution and size of data files used as inputs are directly proportional to memory requirements. The software generates an 'out of memory' message whenever it requests a segment of memory from the operating system that is larger than what is currently available. The model software consists of three main modules (figure 1), each composed of a set of sub-modules whose operation is described below.

3.1. Model Inputs

S^{4}CAST v2.0 has a direct execution mode. By simply typing *'S4cast'* in the command window, the user is prompted to enter a series of input parameters in a simple and intuitive way.

3.1.1. Loading databases

The model is ready to work with Network Common Data Form (NetCDF) data files. There are different conventions to set the attributes of the variables contained in NetCDF files. In this way, the data structure must conform as far as possible to the Cooperative Ocean/Atmosphere Research Service (COARDS) convention. Execution errors that may occur due to the selection of data files are easily corrected by minor modifications of data assimilation scripts. Data files can be easily introduced at the request of the user. Once downloaded from the website of a determined center of climate and environmental research, the user inserts data files into the directory set by default (*S4CAST_v2.0/data_files)*.

3.1.2. Input parameters

In order to correctly introduce the input parameters, it is convenient to present some terms commonly used in seasonal forecasting. In this way, the forecast period corresponds to the n-month seasonal period concerning the predictand for which the forecast and hindcasts are performed. Moreover, the lead-time refers to time expressed in months between the last month comprising the predictor monthly period and the first month comprising the forecast period. Thus, medium-range forecast refers to a lead-time set to zero, while long-range forecast refers to a lead-time equal or larger than one month. Strictly, we can't speak about lead-time when the predictor monthly period partially or totally overlaps the forecast period. In this case we refer to lag-time expressed in moths between the last month comprising the forecast period and the last month for predictor period. The relationship between lead-time and lag-time depends on the number of months comprising the forecast period. Finally, the forecast-time is commonly used to describe the time gap expressed in months between the predictor and predictand monthly periods, assuming the same concept represented by the lead-time.

In the first step, predictand and predictor data files are selected. In this way, the predictand field can be precipitation, SST, or any variable susceptible to be predicted from SSTA. The predictor is restricted to SST.

Once predictor and predictand fields are selected, the available common time period between them is analyzed and displayed so that the user is prompted to select the whole common period for analysis or other within it. The same temporal dimension in both fields is required in the statistical analysis to construct the cross-covariance matrix (see section 2.1.).

The next step is for selecting the n-month forecast period in which the predictand is considered. The model allows a selection from one (n = 1) to four (n = 4) months. From the forecast period, the user determines a specific lead-time, relative to the predictor, from which medium-range (lead-time 0) or long-range (lead-time > 0) forecast can be performed. In order to study and evaluate possible teleconnections, the temporal overlapping between the forecast period and the predictor is also available by defining the monthly lags between both fields from monthly lag 0 (synchronous) referred to the case in which the predictor and the predictand fields are taken at the same n-month period, through partial overlapping to eliminate the overlapping (medium-range forecast). Note that synchronous and partially overlapping seasons between predictor and predictand fields are not useful when referring to predictability, although this option is available in order to perform simulations focused on the study of physical mechanisms (teleconnections) between the predictor and predictand fields. Thus, it is worth noting that the model may be focused in the study of the predictability but it can be also used to detect teleconnections between SST (predictor) and a predictand field.

Monthly lags indicating forecast times (lead-times) are user selectable. To illustrate the above, taking a hypothetical case in which the forecast period corresponds to the months from February to April (FMA) whatever the region, the synchronous option will consider the predictor in FMA, while partially overlapping occurs when the predictor is taken for January-to-March (JFM) and December-to-February (DJF). Avoiding overlapping, lead time 0 will be NDJ (November-to-January), lead time 1 will be OND (October-to-December), lead time 2 will be SON (September-to-November) and so on, without overlapping FMA season of the previous year. Thus, the user can select any 3-month isolated period from FMA (synchronous) to MJJ (May-to-July).

Next, the spatial domains of both predictor and predictand fields are easily selected from its latitudinal and longitudinal values. Considering the above options, the user can select a sequence of successive monthly lags or only one so that the predictor is taken for the total amount of selected information (e.g., NDJ+OND+SON).

Later, there is the possibility of applying a filter to the time series of predictor and predictand fields. The current version uses a Butterworth filter, either as high-pass or low-pass filter frequently used in climate-related studies (e.g., Roe and Steig, 2004; Enfield and Cid-Serrano, 2006; Mokhov and Smirnov, 2006; Ault and George, 2012; Schurer and Hegerl, 2013), although the selection of low pass filter is not suitable for seasonal forecast and subsequently is not useful in the current version. Anyway, the possibility of selecting a low pass filter is maintained in order to include decadal predictability in a future version of the model. The application of a filter allows the user to isolate the frequencies at which the variability operates, which can have different sources of predictability. In this way, the user selects the cutoff frequency, following the expression *2dt/T*, being *dt* the sampling interval and *T* the period to be filtered both in the same units of time. If no filter is applied, the raw data is used. There are plenty of filters that could be applied and future versions of the model will include different possibilities.

In case of multiple time selection for predictor, the statistical methodology is firstly applied considering the largest lead-time and successively adding information for other lead-times up to the present. So, continuing with the example above in which the forecast period corresponds to FMA if selected lead-times from 0 to 3, the first predictor selection is made considering the 3-months lead-time period (SON). After, the 2-months lead-time period is added (ASO+SON). Next, up to the period 1-month delayed (ASO+SON+OND), and finally the case up to the period with a lead-time equal to zero (ASO+SON+OND+NDJ). Previous example is illustrated in figure 2.

Once the matrices are determined for each predictor time selection, the statistical methodology is applied. Up to now, the model applies the MCA discriminant analysis technique, although other statistical methodologies will be included in future releases, including CCA or non-linear methods as neural network and Bayesian methodologies. As indicated in the previous section, MCA determines a new vector base in which the relations between the variables are maximized. Thus, it is important to choose a number of modes (principal directions) to be considered in the computations, selecting either a single mode or a set of them, always consecutive. The analysis of stationarity is performed for a single mode selection. For multi-mode selection, the whole time series will be considered.

The statistical field significance level is set for the first time to assess the analysis of stationarity. Thus, the model runs for the entire period and for those periods for which the relationships are considered stationary within it. This is internally established by applying the method explained later in the section 3.2.1.

3.1.3. Data preprocessing

From selected data files and input parameters previously defined, preprocessing of data is performed so that the data are prepared for implementing statistical methodology.

3.2. Statistical Tools

At this point the statistical procedure described in the methodology is applied considering different periods based on the previously described stationary analysis.

3.2.1. Analysis of stationarity

Stationarity refers to changes along time in the co-variability pattern between two variables. Thus, we speak about stationarity when such a pattern of co-variability keeps invariant within a time period and therefore will be non-stationary when showing changes. To evaluate how much the predictor (Y) and the predictand (Z) fields are related to each other, the model calculates running mean correlations between the expansion coefficients indicated in (4) and (5) for the selected k^{th} mode along the record. This technique has been widely used to determine the stationarity of the relationships between the time series of climate indices (e.g., Camberlin et al., 2001; Rimbu et al., 2003; Van Oldenborgh and Burgers, 2005). Next, the significance level of correlation coefficients is calculated according to the method explained in section 2.2. In this way, stationary relationships between the predictor (Y) and the predictand (Z) fields are established by applying a 21 years moving correlation windows analysis between the leading expansion coefficients of both fields obtained from the discriminant analysis method (section 2.1.) using the whole record in accordance with the evolution of the correlation coefficient. To do this, three types of 21 years moving correlation windows are user selectable: 'delayed' to correlate one year and the 20 previous years; 'centered' to correlate one year, the 10 previous years and the 10 next years; or 'advanced' to correlate one year and the 20 next years. Note that delayed correlation coefficients are the most suitable in a forecast context when referring to future prediction. Nevertheless, centered and advanced correlation coefficients are also available for application no matter the aim of the user.

From previous analysis, three different periods are analyzed depending on the stationarity of the predictability: use the significant correlation period (hereinafter SC) for which the expansion coefficients are significantly correlated; use no significant correlation period (hereinafter NSC), and work with the entire period (hereinafter EP). The model performs all calculations for each period separately and, from them, the simulated maps (hindcasts) of the predictand for each year are calculated by applying cross-validation.**3.2.2. Model**

validation

Cross-validation is used in climate forecasting as part of statistical models when assessing forecast skill (Michaelsen, 1987; Barnston and Van den Dool, 1993; Elsner and Schmertmann, 1994). This method is intended as a model validation technique in which the data for the predictor and the predictand for a given time step is removed from the analysis to make an estimate of it with the rest of data, comparing the simulated value with the removed one. In this way, a cross-validated hindcast is obtained. In the S^4CAST model, the leave-one-out method is applied as described by (Dayan et al., 2013). From the comparison between the predicted value and the original one, the skill of the model can be inferred using different skill-scores. S^4CAST considers the Pearson correlation coefficients and the root mean square error (RMSE) although other scores will be introduced in future versions.

3.3. Model Outputs

Modes of co-variability are related to spatial patterns of different variables that co-vary over time, and thus, are linked to each other. In the case of MCA, the covariance matrix is computed and the SVD method is applied to provide a new basis of eigenvectors for the predictor and predictand fields which covariance is maximized. The obtained singular vectors describe spatial patterns of anomalies in each of the variables that tend to be related to each other. Regression and correlation maps and corresponding expansion coefficients determine each mode of co-variability for the predictor and predictand fields. The expansion coefficients indicate the weight of these patterns in each of the time steps. Thus, regression and correlation co-variability maps can be represented. This is done with the original anomalous matrix, highlighting those grid points whose time series are highly correlated with the obtained expansion coefficients, showing large co-variability and determining the key regions of prediction. To represent it, regression and correlation maps are calculated to analyze the coupling between variables and to understand the physical mechanisms involved in the link.

On the other hand, the time series of the expansion coefficients determine the scores of the regression and correlation maps at each time along the study period. The model represent the expansion coefficients used to calculate the regression coefficients. Thus, those years in which the expansion coefficients for the predictor and the predictand are highly correlated will coincide with years in which we can expect a better estimation.

In the current version of the model, the root mean square error (RMSE) and the Pearson correlation coefficients skill scores have been included. These techniques are applied to compare the observed and simulated maps (hindcasts) of the predictand field obtaining correlation and RMSE maps and time series. On the one hand, maps are obtained calculating for each grid point the skill scores between the hindcast and the observed maps. On the other hand, time series are obtained for each time by applying correlation and RMSE between the area average of the observed and estimated maps. Some comments on these techniques are addressed by Barnston (1992). The S^4CAST model generates the hindcast within the EP, SC and NSC periods separately from applying the one-leave-out method (Dayan et al., 2013) and then the statistical methodology.

4. Application of the model: case studies

Two different case studies have been simulated as benchmark examples. Both cases are focused on the predictive ability of the tropical Atlantic SSTA. In a first simulation, the predictand field corresponds to Sahelian rainfall. In a second simulation, winter tropical Pacific SSTA have been used as predictand field. The links between tropical Atlantic Ocean and the two variables selected as predictand fields have been widely studied exhibiting non-stationary relationships. The results obtained by applying the model have been contrasted in the following sections. Tables 1 and 2 list the entries for both case studies to be easily reproduced by the user.

4.1 Tropical Atlantic – Sahelian rainfall

In this first case study the model has been applied to validate its use in the study of seasonal rainfall predictability in the Sahel taking tropical Atlantic as predictor field. The West African Monsoon (WAM) is characterized by a strong seasonal rainfall regime that occurs from July to September related to the semi-annual shift of the Intertropical Convergence Zone (ITCZ) together with the presence of a strong thermal gradient between the Sahara and the ocean in the Gulf of Guinea. The interannual fluctuations in seasonal rainfall are due to various causes, being the changes in global SST the main driver of WAM variability (Folland, 1986; Palmer, 1986; Fontaine et al., 1998; Rodríguez-Fonseca et al., 2015). Particularly, several observational studies suggest the influence of tropical Atlantic SSTA on the WAM at interannual time scales (Giannini et al., 2003; Polo et al., 2008; Joly and Voldoire, 2009; Nnamchi and Li, 2011).

Regarding the input parameters (table 1), the predictand field corresponds to precipitation from GPCC Full Data Reanalysis monthly means of precipitation appended with GPCC monitoring dataset from 2011 onwards with a resolution of 1.0º x 1.0º covering the period from January 1901 to March 2015 (Rudolf et al., 2010; Becker et al., 2013; Schneider et al., 2014; *http://gpcc.dwd.de*). The forecast period consists of July to September (JAS), computing seasonal anomalous rainfall in the Sahelian domain (18W-10E; 12N-18N). No frequency filter is applied for predictand. The predictor field corresponds to NOAA Extended Reconstructed SST (ERSST) V3b monthly means of SST with a resolution of 2.0º x 2.0º spanning the period from January 1854 to May 2015 (Smith and Reynolds 2003; 2004; Smith et al., 2008; *http://www.ncdc.noaa.gov/oa/climate/research/sst/ersstv3.php*). The spatial domain corresponds to southern subtropical and equatorial Atlantic band (60W-20E; 20S-4N). A high pass filter with cutoff frequency set to 7 years has been applied to the predictor time series in order to analyze the influence of SSTA interannual variability, which includes leading oceanic interannual variability modes such as the Atlantic equatorial mode (AEM) (Polo et al., 2008) or the South Atlantic Ocean dipole (SAOD) (Nnamchi et al., 2011). Medium-range forecast has been taken into account setting the lead-time to zero (equivalent to monthly lag 3). In this way, April-to-June (AMJ) is the selected season for predictor.

For applying the methodology, the leading mode of co-variability ($k = 1$) has been selected. The correlation curve (figure 3) reflects the stationary periods (SC and NSC) within EP period as stated in section 3.2.1. The SC period is almost restricted to years from 1932 to 1971 with some exceptions. The remaining years are taken to analyze the predictability for the NSC period.

Figure 4 show regression maps associated with the leading mode for the periods SC, EP and NSC explaining 50%, 32% and 41% of co-variability respectively. For the SC period (figure 4, top panels), the co-variability pattern exhibits a quasi-isolated cooling in the tropical Atlantic associated with a rainfall dipole over West Africa with negative anomalies in the region of the Gulf of Guinea and opposite in the Sahel. The opposite co-variability pattern takes place under negative scores of the expansion coefficient. These results are in agreement with those found in the last decades of the 20th century by several authors who have discussed the role of the tropical Atlantic SST as a dominant factor in the WAM variability at interannual and seasonal time scales (Janowiak, 1988; Janicot, 1992; Fontaine and Janicot, 1996). Losada et al. (2010b) found how the response to an isolated positive equatorial Atlantic Niño event is a dipolar rainfall pattern in which the decrease of rainfall in Sahel is related to the increase of rainfall in Guinea (as in figure 4) due to changes in the sea-land pressure gradient between Gulf of Guinea SSTs and the Sahel. Mohino et al (2011) and Rodríguez-Fonseca et al. (2011) have found in the observations how this dipolar behavior takes place for some particular decades coinciding with the SC periods, confirming in this way the correct determination of the leading co-variability mode by the model. When considering the EP period (figure 4, middle panels), a co-variability pattern similar to that observed for the SC period is appreciated with small differences. Regarding the predictand field, the anomalous rainfall signal is less intense when compared to SC. For the predictor, the cooling in the tropical Atlantic is accompanied by opposite weak anomalies in the north subtropical and tropical Pacific. Regarding the NSC period (figure 4, bottom panels), as for the previous periods (SC, EP) a cooling in the tropical Atlantic is observed concerning the predictor associated with negative rainfall anomalies in the Gulf of Guinea and a weak positive signal in the eastern Sahel, virtually disappearing the rainfall dipole. The global SSTA regression map shows a significant warming in the tropical Pacific. The opposite pattern should be considered under negative scores of the expansion coefficient.

The results presented above support the existence of a non-stationary behavior of the teleconnections between SSTA variability and rainfall associated with WAM. Several authors have addressed the dipolar anomalous rainfall pattern as a response of an isolated tropical Atlantic warming (cooling) (Rodríguez-Fonseca et al., 2011, Losada et al., 2010a; 2010b; Mohino et al., 2011) restricted to the period 1957-78 in the observations. The uniform rainfall signal over the whole West Africa, with negative anomalies related to a cooling over tropical Atlantic and an opposite sign pattern over tropical Pacific is only observed for the period from 1979 in advance. These results agree with Losada et al. (2012), who focused on non-stationary influences of tropical global SST in WAM variability, explaining how the disappearance of the dipole was due to the counteracting effect of the anomalous responses of the Pacific and Atlantic on the Sahel. Recently, Diatta and Fink (2014) have documented similar non-stationary relationships.

The associated skill of the model to reproduce the rainfall is shown in figure 5 in terms of correlation maps and time series for SC and EP periods. A qualitative improvement is observed when considering the SC periods instead of the whole period (EP). This result points to a better spatial distribution of the significant values for particular decades in which the signal extends to a larger spatial domain. In order to analyze the performance of the simulation for each particular year, the correlation between observed and predicted maps at each time step is calculated and shown in figure 5. Since it has only been considered the leading mode of co-variability, the time series of validation between observed and simulated rainfall should evolve following the absolute values of the expansion coefficients. Thus, when the expansion coefficient (U) of the predictor (SST) shows high scores in the leading mode, good hindcasts are obtained.

4.2. Tropical Atlantic – Tropical Pacific

A non-stationary behavior in the association between tropical Atlantic and tropical Pacific SSTA has been recently documented in some works suggesting that the tropical Atlantic SSTA during the boreal summer could be a potential predictor of winter tropical Pacific SSTA variability after the 1970s (Rodríguez-Fonseca et al., 2009; Ding et al., 2012). In this section, the S4CAST model has been applied to corroborate the non-stationarity in the teleconnection between tropical Atlantic considered as predictor field and tropical Pacific variability, a feature that has been also demonstrated in Martin del Rey et al (2015).

The input parameters are listed in table 2. Both predictor and predictand fields corresponds to NOAA ERSST introduced in the previous section (4.1) covering the period from January 1854 to May 2015. The forecast period consists of December-to-March (DJFM). The selected region for predictand corresponds to SSTA in the tropical Pacific domain (120E-60W; 30S-20N), while the predictor corresponds to tropical Atlantic SSTA (60W-20E; 20S-4N) and has been considered for the period July-to-October (JASO), which means long-range forecast setting the lead-time to one month. A high pass filter with cutoff frequency set to 7 years has been applied to both predictor and predictand time series in order to analyze the predictability considering interannual variability. For applying the methodology and assess the stationary periods (SC and NSC) within EP, the leading mode of co-variability ($k = 1$) has been selected.

The correlation curve (figure 7) presents the SC period clearly divided into two intervals: from 1889 to 1939 and from 1985 up to the present (2015). Consequently, the NSC period corresponds to the remaining years within the study period (1854 – 2015).

The leading mode (figure 8) for the periods SC, NSC and EP explains 52%, 28% and 43% of co-variability respectively. Regarding the SC (figure 8; top panels) and EP (figure 8; middle panels) periods it is observed how a cooling (warming) in the tropical Atlantic is related to a warming (cooling). Thus the co-varibility pattern is defined by opposite sign anomalies between predictor and predictand fields, although the magnitude of the anomalies is greater concerning the SC period. Considering the NSC period (figure 8; bottom panels), a signal in tropical Pacific is not observed in response to the tropical Atlantic cooling (warming).

Previous results are in agreement with former studies in which a similar tropical SSTA pattern with opposite temperature anomalies in the equatorial Atlantic and Pacific in summer has been documented to occur in the decades within the SC period (Rodríguez-Fonseca et al., 2009; Martin-Rey et al., 2012). Thus, Martín-Rey et al. (2014, 2015) point to a non-stationary relationship that seems to take place in the early 20th century and after the 1970s, confirming the correct determination of the leading co-variability mode by the model.

The mechanism from which the teleconnection takes place, has been explained by Polo et al. (2015), who suggest that a cooling in the equatorial Atlantic results in enhanced equatorial convection, altering the Walker circulation and consequently enhancing subsidence and surface wind divergence over the equatorial Pacific during the period July-to-August (JASO). The anomalous wind piles up water in the western tropical Pacific, triggering a Kelvin wave eastward from autumn to winter, setting up the conditions for a cold event in the equatorial east Pacific during the period December-to-March (DJFM). Considering a cooling in the tropical Atlantic, the opposite sequence takes place.

The skill of the model in reproducing tropical Pacific SSTA (figure 9) is also restricted to stationary conditions. Thus, depending on the considered sequence of decades within the period EP (figure 9; middle panels), the model provides better results for period SC (figure 9; top panels), while it is not able to produce reliable estimations when period NSC (figure 9; bottom panels) is taken into account. These results highlight the need to consider different periods and possible modulations when tackling seasonal predictability of tropical Pacific SSTA, in agreement with recent results of Martin del Rey et al. (2015).

5. Discussion and conclusions

It is well known how dynamical models are far to produce very accurate seasonal climate forecast for non-ENSO events, partly due to the presence of strong biases in some regions, as the tropical Atlantic (Barnston et al., 2015). In contrast, statistical models, despite being an useful and effective supplement, they are mostly unable to reproduce the non-linearity in the ocean-atmosphere system, exceptions include neural networks and Bayesian methods. Attempts to implement new statistical models constitute a fundamental contribution aimed to enhance and complement the dynamical models. Anyway, statistical models have evolved linked to dynamical models, either as an alternative or within them as a hybrid model.

Following this reasoning, this paper introduces the S^4CAST v2.0 model. The model was created from the first version (S^4CAST v1.0) developed as the main part of a cooperation project between the *Laboratoire de Physique de l'Atmosphère et de l'Océan Siméon Fongang* of the University Cheik Anta Diop (UCAD) in Dakar (Senegal) and the *Complutense* University of Madrid (UCM) within the VIII UCM Call for Cooperation and Development projects (VR: 101/11) and was named "Creation and Donation of a statistical seasonal forecast model for West African rainfall". Thereby, the authors wanted to respect the number of the donation version despite not having a publication. As a brief explanation on the history, the original model was restricted to study the predictability of West African rainfall from tropical global SSTA under some input parameters much more limited respect version 2.0. Thus, the reason for developing and improve the model for publication is the motivation arising from colleagues in different institutions along Africa and Europe to expand the model and use it as an alternative tool to look for SST-related predictability due to the strong SST bias that coupled dynamical models exhibit nowadays.

The model is based on the predictive power of the SST. Concerning the association along time between SSTA and any climate-related variable susceptible of being predicted from it, the concept of stationarity is raised as one of the motivating factors in creating the S^4CAST model. The stationarity refers to changes in the co-variability patterns between the predictor and the predictand fields along a given sequence of decades, so that it can be kept invariant (stationary) or changing (non-stationary). This concept has been addressed by different authors (Janicot et al., 1996; Fontaine et al., 1998; Rodríguez-Fonseca et al., 2009, 2011; Mohino et al., 2011; Martín-Rey et al., 2012; Losada et al., 2012) and becomes the main novelty and contribution introduced by S^4CAST as a key factor to consider in seasonal forecasting provided by current prediction models, either dynamical or statistical. Thus, S^4CAST model is an alternative to enhance and complement the estimates made by dynamical models, which have a number of systematic errors to adequately reproduce the tropical climate variability (Biasutti et al., 2006; Richter and Xie, 2008; Wahl et al., 2011; Doi et al., 2012; Richter et al., 2012; Bellenguer et al., 2013; Brown et al., 2013; Li and Xie, 2013; Toniazzo and Woolnough, 2013; Vanniere et al., 2013; Xue et al., 2013). For the time being, the S^4CAST model cannot be applied for strict operational forecasts, although its application in determining stationary relationships between two fields and their co-variability patterns can be crucial for improving the estimates provided by the operating prediction models currently used.

The model is proposed for use in two areas: the study of seasonal predictability and the study of teleconnections, both based on the influence of SST. On the one hand, we refer to predictability when predictor is considered from a lead-time equal to 0 months (medium-range forecast) in advance (long-range forecast). On the other hand, we speak about the study of teleconnections when predictor seasonal selection partially or totally overlaps (synchronous) the forecast period, meaning that one can not speak about lead-time, instead we speak about a monthly lag between the last month in the forecast period and the last month comprising the predictor monthly period.

In addition to previous considerations, the model always provides the predictions in hindcast mode for the different periods of stationarity (SC, NSC and EP), while the forecast mode depends on input parameters and data files used for predictor and predictand fields. For instance, considering from September to November (SON) as forecast period concerning the predictand and selecting a lead- time of two months for the prediction, which means taking the predictor two months before September (from April to June; AMJ), the prediction for SON 2015 will be performed if predictand field is available at least until November 2014 and predictor is available at least until June 2015. Thus, the model constructs the regression coeficient by using the common period until November 2014. Regression coefficients along with predictor data (AMJ 2015) will provide the forecast for SON 2015. In this way, the model firstly checks data availability related to the input parameters and shows by screen if future forecast is enabled. If enabled, the model performs three types of forecast by computing the regression coefficient respectively for each period (SC, NSC, EP). Finally, the user should determine the better forecast by a study of the modulations of each stationary period and the sequence of hindcasts immediately preceding the present.

In the applications shown in this paper we have focused in the results from MCA. This statistical methodology, along with Canonical Correlation Analysis (CCA), have been widely used in studies of predictability during the last decades (Barnston and Ropelewski, 1992; Bretherton et al., 1992; Wallace et al., 1992; Barnston and Smith, 1996; Fontaine et al., 1999; Korecha and Barnston, 2007; Barnston and Tippet, 2014; Recalde-Coronel et al., 2014). Integration of the methodology and intuitive use through a user interface are some of the main advantages of the S^4CAST model, allowing the selection of a big number of inputs. Future releases of the model will include other methodologies that are currently being introduced and tested.

Originally, the model was created to tackle the study of the predictability of anomalous rainfall associated with WAM, which co-varies in a different way with the tropical band of Atlantic and Pacific ocean basins, being an indicator of non-stationarity (Losada et al., 2012). The transition between SC and NSC periods, around the 1970s, has served as the starting point of many studies focusing on the influence of global SSTA before and after that period (Mohino et al., 2011; Rodríguez-Fonseca et al., 2011; 2015; Losada et al., 2012) while being one of the motivations to create S^4CAST.

The choice of the case study related to Sahelian rainfall predictability is motivated by two main reasons: on the one hand, SST in the tropical Atlantic is well known to strongly influence the dynamics of the ITCZ (Fontaine et al., 1998) which in turn determines the subsequent WAM. Nevertheless, dynamical models do not reproduce the influence of SST on the ITCZ (Lin, 2007; Richter and Xie, 2008; Doi et al., 2012; Tonniazzo and Woolnough, 2013) becoming the statistical prediction an alternative way to predict WAM variability. The second reason is related to the non-stationary influence of the tropical Atlantic on Sahelian rainfall reported in some studies (Janicot et al., 1996, 1998; Ward, 1998; Rodríguez-Fonseca et al., 2011; Mohino et al., 2011; Losada et al., 2012).

The second case study has served as a benchmark to certify the ability of the S^4CAST model in the study of SSTA predictability by the corroboration of the Equatorial Atlatnic variability as preditor of ENSO· This is a recently discovered relationship (Rodríguez-Fonseca et al., 2009; Ding et al., 2011; Polo et al., 2015) that has been found to be non-stationary (Martín del Rey et al., 2014, 2015).

The application of moving correlation windows between expansion coefficients obtained from MCA analysis results in three periods of stationarity depending on the statistically significant correlation: entire period (EP), significant correlation period (SC) and no-significant correlation period (NSC). For the case in which non-stationarity is considered we refer to EP period, assuming changes in co-variability patterns. Stationarity is referred to SC and NSC periods. These periods may slightly vary depending on the type of moving correlation windows: advanced, centered or delayed. Stationary analysis to determine the three different work periods (SC, NSC, EP) is limited to the selection of a single mode of co-variability.

When selecting a set of modes, the stationarity analysis is not applied so that simulations are only developed for EP period, whereby the whole time series is considered for both the predictor and predictand fields.

Three conditions may enhance the degree of confidence in a given predictor. The first has to do with the selection of moving correlation windows (see section 3.1.2.) used to determine the working scenarios (SC, NSC, EP). Delayed moving correlation windows can help in this task. Thus, if correlation coefficients between the expansion coefficients (U and V) exhibit significant values for the present year and the previous 21 study years, greater confidence is assumed for the predictor. The second condition is determined by the value of the expansion coefficient (U) for the current year so that the higher its value, the better the forecast. The last condition has to do with the percentage of variance explained by the selected co-variability mode, the higher its value, the better the forecast. Nevertheless, despite previous conditions, the influence of other remote and nearby oceanic predictors must be considered in order to provide a full and reliable predictability study.

So far, the data files used as predictor and predictand fields correspond to observations and reanalysis from several institutions. The use of new data files is simple and can be performed according to user needs. The upgrade of data files from respective websites must be checked periodically to strengthen the results. In addition, it is also advisable to launch the same simulations using different data files in order to compare the results and assess the robustness of the forecast. The results shown in this work for different selections have been verified by following these criteria.

The results obtained by using the S^4CAST model put forward the consideration of non-stationarity in the co-variability patterns and therefore in climatic teleconnections. Thus, it is important to determine the multidecadal modulator of the interannual variability in order to know which predictor is the one affecting in particular periods and regions (Rodríguez-Fonseca et al., 2015).

6. Code availability

The model consists of a software package organized in folders containing libraries, functions and scripts developed as a MATLAB® toolbox from version R2010b onwards. Two of the folders, named as *mexcdf* and *netcdf_toolbox*, corresponds to libraries needed for working with NetCDF files and have been downloaded from www.mexcdf.sourceforge.net and built-in into the model. The file containing the model core with the executable code is named *S4core*. Once the toolbox has been added to the MATLAB® path and by simply typing '*S4cast*' in the command window, the user is prompted to enter a number of input parameters required to launch a simulation. The software package *S4plot* dedicated to plot figures has been added so that the user can use this software by typing *'figures'* in the command window. Note that figures presented in this work have been further improved manually. The code is Open Access and can be downloaded from the Zenodo repository (DOI 10.5281/zenodo.15985) in the URL *https://zenodo.org/record/15985*. To facilitate the execution of the model leading to the results shown in this paper, used data files that have been previously defined in Section 4, are included in the directories */S4CAST_v2.0/data_files/predictand* and */S4CAST_v2.0/data_files/predictor*. The second case study requires NOAA ERSST as predictor and predictand. The code has been thoroughly analyzed by using several data files and input parameters. However, the emergence of software bugs is not ruled out, being mostly associated with problems to adapt and use NetCDF files. To solve these hypothetical code bugs, please do not hesitate to contact authors.

Acknowledgements

The research leading to these results received funding from the PREFACE-EU project (EU FP7/2007-2013) under grant agreement no. 603521, Spanish national project MINECO (CGL2012-38923-C02-01) and the VR: 101/11 project from the VIII UCM Call for Cooperation and Development projects. We also appreciate the work done by SOURCEFORGE.NET® staff in creating NetCDF libraries for MATLAB®. And of course, thanks also to the reviewers, editors and their advice and/or criticism.

References

Adams, R. M., Chen, C. C., McCarl, B. A., Weiher, R. F. (1999). The economic consequences of ENSO events for agriculture. Climate Research, 13(3), 165-172.

Ault, T. R., Cole, J. E., & St George, S. (2012). The amplitude of decadal to multidecadal variability in precipitation simulated by state-of-the-art climate models. Geophysical Research Letters, 39(21).

Baboo, S. S., & Shereef, I. K. (2010). An efficient weather forecasting system using artificial neural network. International journal of environmental science and development, 1(4), 2010-0264.

Barnett, T. P., Preisendorfer, R. (1987). Origins and levels of monthly and seasonal forecast skill for United States surface air temperatures determined by canonical correlation analysis. Monthly Weather Review, 115(9), 1825-1850.

Barnett, T. P., Graham, N., Pazan, S., White, W., Latif, M., Flügel, M. (1993). ENSO and ENSO-related predictability. Part I: Prediction of equatorial Pacific sea surface temperature with a hybrid coupled ocean-atmosphere model. Journal of Climate, 6(8), 1545-1566.

Barnett, T. P. (1995). Monte Carlo climate forecasting. Journal of climate, 8(5), 1005-1022.

Barnston, A. G. (1992). Correspondence among the correlation, RMSE, and Heidke forecast verification measures; refinement of the Heidke score. Weather and Forecasting, 7(4), 699-709.

Barnston, A. G., Ropelewski, C. F. (1992). Prediction of ENSO episodes using canonical correlation analysis. Journal of Climate, 5(11), 1316-1345.

Barnston, A. G., van den Dool, H. M. (1993). A degeneracy in cross-validated skill in regression-based forecasts. Journal of Climate, 6(5), 963-977.

Barnston, A. G., van den Dool, H. M., Rodenhuis, D. R., Ropelewski, C. R., Kousky, V. E., O'Lenic, E. A., Leetmaa, A. (1994). Long-lead seasonal forecasts-Where do we stand?. Bulletin of the American Meteorological Society, 75(11), 2097-2114.

Barnston, A. G., Smith, T. M. (1996). Specification and prediction of global surface temperature and precipitation from global SST using CCA. Journal of Climate, 9(11), 2660-2697.

Barnston, A. G., He, Y., Glantz, M. H. (1999). Predictive skill of statistical and dynamical climate models in SST forecasts during the 1997-98 El Niño episode and the 1998 La Niña onset. Bulletin of the American Meteorological Society, 80(2), 217-243.

Barnston, A. G., Tippett, M. K. (2014). Climate information, outlooks, and understanding–where does the IRI stand?. Earth Perspectives, 1(1), 1-17.

Barnston, A. G., Tippet, M. K., van den Dool, H. M., Unger, D. A. (2015). Toward an Improved Multi-model ENSO Prediction. Journal of Applied Meteorology and Climatology, (2015) Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B., Schamm, K., Schneider, U., Ziese, M. (2013). A description of the global land-surface precipitation data products of the Global Precipitation Climatology Center with sample applications including centennial (trend) analysis from 1901-present. Earth System Science Data, 5, 71-99.

Bellenger, H., Guilyardi, E., Leloup, J., Lengaigne, M., Vialard, J. (2013). ENSO representation in climate models: from CMIP3 to CMIP5. Climate Dynamics, 42(7-8), 1999-2018.

Biasutti, M., Sobel, A. H., Kushnir, Y. (2006). AGCM precipitation biases in the tropical Atlantic. Journal of climate, 19(6), 935-958.

Bjerknes, J. (1969). Atmospheric teleconnections from the equatorial pacific 1. Monthly Weather Review, 97(3), 163-172.

Bretherton, C. S., Smith, C., Wallace, J. M. (1992). An intercomparison of methods for finding coupled patterns in climate data. Journal of climate, 5(6), 541-560.

Brown, J. N., Gupta, A. S., Brown, J. R., Muir, L. C., Risbey, J. S., Whetton, P., Wijffels, S. E. (2013). Implications of CMIP3 model biases and uncertainties for climate projections in the western tropical Pacific. Climatic Change, 119(1), 147-161.

Bulić, I. H., Kucharski, F. (2012). Delayed ENSO impact on spring precipitation over North/Atlantic European region. Climate dynamics, 38(11-12), 2593-2612.

Camberlin, P., Janicot, S., & Poccard, I. (2001). Seasonality and atmospheric dynamics of the teleconnection between African rainfall and tropical sea-surface temperature: Atlantic vs. ENSO. International Journal of Climatology, 21(8), 973-1005.

Cane, M. A., Zebiak, S. E., Dolan, S. C. (1986). Experimental forecasts of EL Nino. Nature, 321(6073), 827-832.

Chang, P., Fang, Y., Saravanan, R., Ji, L., & Seidel, H. (2006). The cause of the fragile relationship between the Pacific El Nino and the Atlantic Nino. Nature, 443(7109), 324-328.

Cherry, S. (1996). Singular value decomposition analysis and canonical correlation analysis. Journal of Climate, 9(9), 2003-2009.

Cherry, S. (1997). Some comments on singular value decomposition analysis. Journal of Climate, 10(7), 1759-1761.

Chung, C. E., Ramanathan, V. (2006). Weakening of North Indian SST gradients and the monsoon rainfall in India and the Sahel. Journal of Climate, 19(10), 2036-2045.

Coelho, C. A. S., Stephenson, D. B., Balmaseda, M., Doblas-Reyes, F. J., van Oldenborgh, G. J. (2006). Toward an integrated seasonal forecasting system for South America. Journal of Climate, 19(15), 3704-3721.

Dayan, H., Vialard, J., Izumo, T., & Lengaigne, M. (2014). Does sea surface temperature outside the tropical Pacific contribute to enhanced ENSO predictability?. Climate Dynamics, 43(5-6), 1311-1325.

Deng, X., Huang, J., Qiao, F., Naylor, R. L., Falcon, W. P., Burke, M., Battisti, D. (2010). Impacts of El Nino-Southern Oscillation events on China's rice production. Journal of Geographical Sciences, 20(1), 3-16.

Diatta, S., Fink, A. H. (2014). Statistical relationship between remote climate indices and West African monsoon variability. International Journal of Climatology.

Ding, H., Keenlyside, N. S., & Latif, M. (2012). Impact of the equatorial Atlantic on the El Nino southern oscillation. Climate dynamics, 38(9-10), 1965-1972.

Doi, T., Vecchi, G. A., Rosati, A. J., Delworth, T. L. (2012). Biases in the Atlantic ITCZ in seasonal-interannual variations for a coarse-and a high-resolution coupled climate model. Journal of Climate, 25(16), 5494-5511.

Drosdowsky, W., Chambers, L. E. (2001). Near-global sea surface temperature anomalies as predictors of Australian seasonal rainfall. Journal of Climate, 14(7), 1677-1687.

Elsner, J. B., Schmertmann, C. P. (1994). Assessing forecast skill through cross validation. Weather and Forecasting, 9(4), 619-624.

Enfield, D. B., & Cid‐Serrano, L. (2006). Projecting the risk of future climate shifts. International Journal of Climatology, 26(7), 885-895.

Folland, C. K., Palmer, T. N., Parker, D. E. (1986). Sahel rainfall and worldwide sea temperatures, 1901–85. Nature, 320(6063), 602-607.

Fontaine, B., Janicot, S. (1996). Sea surface temperature fields associated with West African rainfall anomaly types. Journal of climate, 9(11), 2935-2940.

Fontaine, B., Trzaska, S., Janicot, S. (1998). Evolution of the relationship between near global and Atlantic SST modes and the rainy season in West Africa: statistical analyses and sensitivity experiments. Climate Dynamics, 14(5), 353-368.

Fontaine, B., Philippon, N., & Camberlin, P. (1999). An improvement of June–September rainfall forecasting in the Sahel based upon region April–May moist static energy content (1968–1997). Geophysical Research Letters, 26(14), 2041-2044.

Fontaine, B., Monerie, P. A., Gaetani, M., Roucou, P. (2011). Climate adjustments over the African‐Indian monsoon regions accompanying Mediterranean Sea thermal variability. Journal of Geophysical Research: Atmospheres (1984–2012), 116(D23).

Frankignoul, C., Hasselmann, K. (1977). Stochastic climate models, part II application to sea‐surface temperature anomalies and thermocline variability. Tellus, 29(4), 289-305.

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 (1984–2012), 115(D24).

Gardner, M. W., Dorling, S. R. (1998). Artificial neural networks (the multilayer perceptron)--a review of applications in the atmospheric sciences. Atmospheric environment, 32(14-15), 2627-2636.

Garric, G., Douville, H., Déqué, M. (2002). Prospects for improved seasonal predictions of monsoon precipitation over Sahel. International journal of climatology, 22(3), 331-345.

Giannini, A., Chiang, J. C., Cane, M. A., Kushnir, Y., Seager, R. (2001). The ENSO teleconnection to the tropical Atlantic Ocean: contributions of the remote and local SSTs to rainfall variability in the tropical Americas*. Journal of Climate, 14(24), 4530-4544.

Giannini, A., Saravanan, R., Chang, P. (2003). Oceanic forcing of Sahel rainfall on interannual to interdecadal time scales. Science, 302(5647), 1027-1030.

Gill, A. (1980). Some simple solutions for heat‐induced tropical circulation. Quarterly Journal of the Royal Meteorological Society, 106(449), 447-462.

Glahn, H. R., Lowry, D. A. (1972). The use of model output statistics (MOS) in objective weather forecasting. Journal of applied meteorology, 11(8), 1203-1211.

Hansen, J. W., Hodges, A. W., Jones, J. W. (1998). ENSO Influences on Agriculture in the Southeastern United States*. Journal of Climate, 11(3), 404-411.

Ham, Y. G., Kug, J. S., Park, J. Y., & Jin, F. F. (2013a). Sea surface temperature in the north tropical Atlantic as a trigger for El Niño/Southern Oscillation events. Nature Geoscience, 6(2), 112-116.

Ham, Y. G., Sung, M. K., An, S. I., Schubert, S. D., & Kug, J. S. (2013b). Role of tropical Atlantic SST variability as a modulator of El Niño teleconnections. Asia-Pacific Journal of Atmospheric Sciences, 1-15.

Harris, I., Jones, P. D., Osborn, T. J., & Lister, D. H. (2014). Updated high‐resolution grids of monthly climatic observations–the CRU TS3. 10 Dataset. International Journal of Climatology, 34(3), 623-642.

Harrison, D. E., Larkin, N. K. (1998). El Niño‐Southern Oscillation sea surface temperature and wind anomalies, 1946–1993. Reviews of Geophysics, 36(3), 353-399.

Hasselmann, K. (1976). Stochastic climate models part I. Theory. Tellus, 28(6), 473-485.

Haylock, M. R., Peterson, T. C., Alves, L. M., Ambrizzi, T., Anunciação, Y. M. T., Baez, J., Vincent, L. A. (2006). Trends in total and extreme South American rainfall in 1960-2000 and links with sea surface temperature. Journal of climate, 19(8), 1490-1512.

Hsieh, W. W., Tang, B. (1998). Applying neural network models to prediction and data analysis in meteorology and oceanography.

Hsieh, W. W. (2001). Nonlinear canonical correlation analysis of the tropical Pacific climate variability using a neural network approach. Journal of Climate, 14(12), 2528-2539.

Janicot, S. (1992). Spatiotemporal variability of West African rainfall. Part I: Regionalizations and typings. Journal of Climate, 5(5), 489-497.

Janicot, S., Moron, V., Fontaine, B. (1996). Sahel droughts and ENSO dynamics. Geophysical Research Letters, 23(5), 515-518.

Janicot, S., Harzallah, A., Fontaine, B., Moron, V. (1998). West African monsoon dynamics and eastern equatorial Atlantic and Pacific SST anomalies (1970-88). Journal of Climate, 11(8), 1874-1882.

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

Janowiak, J. E. (1988). An investigation of interannual rainfall variability in Africa. Journal of

Climate, 1(3), 240-255.

Ji, M., Kumar, A., Leetmaa, A. (1994a). A multiseason climate forecast system at the National Meteorological Center. Bulletin of the American Meteorological Society, 75(4), 569-577.

Ji, M., Kumar, A., Leetmaa, A. (1994b). An experimental coupled forecast system at the National Meteorological Center. Tellus A, 46(4), 398-418.

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

Keenlyside, N. S., Ding, H., & Latif, M. (2013). Potential of equatorial Atlantic variability to enhance El Niño prediction. Geophysical Research Letters, 40(10), 2278-2283.

Klein, W. H., Glahn, H. R. (1974). Forecasting local weather by means of model output statistics. Bulletin of the American Meteorological Society, 55(10), 1217-1227.

Klein, S. A., Soden, B. J., Lau, N. C. (1999). Remote sea surface temperature variations during ENSO: Evidence for a tropical atmospheric bridge. Journal of Climate, 12(4), 917-932.

Knutti, R., Stocker, T. F., Joos, F., Plattner, G. K. (2003). Probabilistic climate change projections using neural networks. Climate Dynamics, 21(3-4), 257-272.

Korecha, D., Barnston, A. G. (2007). predictability of June-September rainfall in Ethiopia. Monthly weather review, 135(2), 628-650.

Kovats, R. S. (2000). El Niño and human health. Bulletin of the World Health Organization, 78(9), 1127-1135.

Kovats, R. S., Bouma, M. J., Hajat, S., Worrall, E., Haines, A. (2003). El Niño and health. The Lancet, 362(9394), 1481-1489.

Latif, M., Barnett, T. P. (1995). Interactions of the tropical oceans. Journal of Climate, 8(4), 952-964.

Legates, D. R., & Willmott, C. J. (1990). Mean seasonal and spatial variability in gauge‐corrected, global precipitation. International Journal of Climatology, 10(2), 111-127.

Legler, D. M., Bryant, K. J., O'Brien, J. J. (1999). Impact of ENSO-related climate anomalies on crop yields in the US. Climatic Change, 42(2), 351-375.

Li, Z., Kafatos, M. (2000). Interannual variability of vegetation in the United States and its relation to El Nino/Southern Oscillation. Remote Sensing of Environment, 71(3), 239-247.

Lin, J. L. (2007). The double-ITCZ problem in IPCC AR4 coupled GCMs: Ocean-atmosphere feedback analysis. Journal of Climate, 20(18), 4497-4525.

Li, G., & Xie, S. P. (2012). Origins of tropical‐wide SST biases in CMIP multi‐model ensembles. Geophysical Research Letters, 39(22).

Li, G., & Xie, S. P. (2014). Tropical Biases in CMIP5 Multimodel Ensemble: The Excessive Equatorial Pacific Cold Tongue and Double ITCZ Problems*. Journal of Climate, 27(4), 1765-1780.

Linthicum, K. J., Anyamba, A., Chretien, J. P., Small, J., Tucker, C. J., Britch, S. C. (2010). The role of global climate patterns in the spatial and temporal distribution of vector-borne disease. In Vector Biology, Ecology and Control (pp. 3-13). Springer Netherlands.

Livezey, R. E., Chen, W. Y. (1983). Statistical field significance and its determination by Monte Carlo techniques. Monthly Weather Review, 111(1), 46-59.

López‐Parages, J., Rodríguez‐Fonseca, B. (2012). Multidecadal modulation of El Niño influence on the Euro‐Mediterranean rainfall. Geophysical Research Letters, 39(2).

López-Parages, J., Rodrígez-Fonseca, B., Terray, L. (2014). A mechanism for the multidecadal modulation of ENSO teleconnections with Europe. Climate Dynamics, 1-14.

Losada, T., Rodríguez-Fonseca, B., Polo, I., Janicot, S., Gervois, S., Chauvin, F., Ruti, P. (2010 a). Tropical response to the Atlantic Equatorial mode: AGCM multimodel approach. Climate dynamics, 35(1), 45-52.

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

Losada, T., Rodríguez‐Fonseca, B., Mohino, E., Bader, J., Janicot, S., Mechoso, C. R. (2012). Tropical SST and Sahel rainfall: A non‐stationary relationship. Geophysical Research Letters, 39(12).

Lu, J. (2009). The dynamics of the Indian Ocean sea surface temperature forcing of Sahel drought. Climate dynamics, 33(4), 445-460.

Maia, A. H., Meinke, H., Lennox, S., & Stone, R. (2007). Inferential, nonparametric statistics to assess the quality of probabilistic forecast systems. Monthly Weather Review, 135(2), 351-362.

Majda, A. J., Timofeyev, I., Eijnden, E. V. (1999). Models for stochastic climate prediction. Proceedings of the National Academy of Sciences, 96(26), 14687-14691.

Martín-Rey, M., Polo, I., Rodríguez-Fonseca, B., Kucharski, F. (2012). Changes in the interannual variability of the tropical Pacific as a response to an equatorial Atlantic forcing. Scientia Marina, 76(S1), 105-116.

Martín-Rey, M., Rodríguez-Fonseca, B., Polo, I., Kucharski, F. (2014). On the Atlantic–Pacific Niños connection: a multidecadal modulated mode. Climate Dynamics, 1-16.

Martín-Rey, M., B. Rodríguez-Fonseca, B., Polo, I. (2015), Atlantic opportunities for ENSO prediction, Geophys. Res. Lett., 42, 6802–6810, doi:10.1002/2015GL065062.

Mason, S. J., Goddard, L., Graham, N. E., Yulaeva, E., Sun, L., Arkin, P. A. (1999). The IRI seasonal climate prediction system and the 1997/98 El Niño event. Bulletin of the American Meteorological Society, 80(9), 1853-1873.

Michaelsen, J. (1987). Cross-validation in statistical climate forecast models. Journal of Climate and Applied Meteorology, 26(11), 1589-1600.

McMichael, A. J., Woodruff, R. E., Hales, S. (2006). Climate change and human health: present and future risks. The Lancet, 367(9513), 859-869.

Mokhov, I. I., & Smirnov, D. A. (2006). El Niño–Southern Oscillation drives North Atlantic Oscillation as revealed with nonlinear techniques from climatic indices. Geophysical research letters, 33(3).

Mohino, E., Janicot, S., Bader, J. (2011). Sahel rainfall and decadal to multi-decadal sea surface temperature variability. Climate Dynamics, 37(3-4), 419-440.

Naylor, R. L., Falcon, W. P., Rochberg, D., Wada, N. (2001). Using El Nino/Southern Oscillation climate data to predict rice production in Indonesia. Climatic Change, 50(3), 255-265.

Newman, M., Sardeshmukh, P. D. (1995). A caveat concerning singular value decomposition. Journal of Climate, 8(2), 352-360.

Nnamchi, H. C., & Li, J. (2011). Influence of the South Atlantic Ocean dipole on West

African summer precipitation. Journal of Climate, 24(4), 1184-1197.

Nnamchi, H. C., Li, J., & Anyadike, R. N. (2011). Does a dipole mode really exist in the South Atlantic Ocean?. Journal of Geophysical Research: Atmospheres (1984–2012), 116(D15).

Palmer, T. N. (1986). Influence of the Atlantic, Pacific and Indian oceans on Sahel rainfall.

Patz, J. A. (2002). A human disease indicator for the effects of recent global climate change. Proceedings of the National Academy of Sciences, 99(20), 12506-12508.

Patz, J. A., Campbell-Lendrum, D., Holloway, T., Foley, J. A. (2005). Impact of regional climate change on human health. Nature, 438(7066), 310-317.

Penland, C., Sardeshmukh, P. D. (1995). The optimal growth of tropical sea surface temperature anomalies. Journal of climate, 8(8), 1999-2024.

Penland, C., Matrosova, L. (1998). Prediction of tropical Atlantic sea surface temperatures using linear inverse modeling. Journal of Climate, 11(3), 483-496.

Phillips, J. G., Cane, M. A., Rosenzweig, C. (1998). ENSO, seasonal rainfall patterns and simulated maize yield variability in Zimbabwe. Agricultural and Forest Meteorology, 90(1), 39-50.

Podestá, G. P., Messina, C. D., Grondona, M. O., Magrin, G. O. (1999). Associations between grain crop yields in central-eastern Argentina and El Niño-Southern Oscillation. Journal of applied meteorology, 38(10), 1488-1498.

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

Polo , I., Martin-Rey, M., Rodriguez-Fonseca, B., Kucharski, F., & Mechoso, C.R. (2015)- Processes in the Pacific La Niña onset triggered by the Atlantic Niño. Climate Dynamics, 44(1-2), 115-131.

Rasmusson, E. M., Carpenter, T. H. (1982). Variations in tropical sea surface temperature and surface wind fields associated with the Southern Oscillation/El Niño. Monthly Weather Review, 110(5), 354-384.

Rasmusson, E. M., Carpenter, T. H. (1983). The relationship between eastern equatorial

Pacific sea surface temperatures and rainfall over India and Sri Lanka. Monthly Weather Review, 111(3), 517-528.

Rayner, N. A., Parker, D. E., Horton, E. B., Folland, C. K., Alexander, L. V., Rowell, D. P., Kaplan, A. (2003). Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. Journal of Geophysical Research: Atmospheres (1984–2012), 108(D14).

Recalde-Coronel, G. C., Barnston, A. G., Muñoz, Á. G. (2014). Predictability of December-April Rainfall in Coastal and Andean Ecuador. Journal of Applied Meteorology and Climatology, (2014).

Richter, I., Xie, S. P. (2008). On the origin of equatorial Atlantic biases in coupled general circulation models. Climate Dynamics, 31(5), 587-598.

Richter, I., Xie, S. P., Wittenberg, A. T., Masumoto, Y. (2012). Tropical Atlantic biases and their relation to surface wind stress and terrestrial precipitation. Climate dynamics, 38(5-6), 985-1001.

Rimbu, N., Lohmann, G., Felis, T., & Pätzold, J. (2003). Shift in ENSO teleconnections recorded by a northern Red Sea coral. Journal of Climate,16(9), 1414-1422.

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

Rodríguez‐Fonseca, B., Janicot, S., Mohino, E., Losada, T., Bader, J., Caminade, C., Voldoire, A. (2010). Interannual and decadal SST‐forced responses of the West African monsoon. Atmospheric Science Letters, 12(1), 67-74.

Rodríguez-Fonseca B., Mohino E., Mechoso CR., Caminade C., Biasutti M., Gaetani M., García-Serrano J., Vizy EK., Cook K., Xue Y., Polo I., Losada L., Druyan L,. Fontaine B., Bader J., Doblas-Reyes FJ., Goddard L., Janicot S., Arribas A., Lau W., Colman A., Vellinga M., Rowell DP., Kucharski F., Voldoire A. (2015), Variability and Predictability of West African Droughts. A review on the role of Sea Surface Temperature Anomalies. Journal of Climate.doi:jcliD1400130

Roe, G. H., & Steig, E. J. (2004). Characterization of millennial-scale climate variability. Journal of climate, 17(10), 1929-1944.

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

Rowell, D. P. (2003). The impact of Mediterranean SSTs on the Sahelian rainfall season. Journal of Climate, 16(5), 849-862.

Rudolf, B., Becker, A., Schneider, U., Meyer-Christoffer, A., Ziese, M. (2010). The new "GPCC Full Data Reanalysis Version 5" providing high-quality gridded monthly precipitation data for the global land-surface is public available since December 2010. GPCC status report December.

Saravanan, R., & Chang, P. (2000). Interaction between tropical Atlantic variability and El Nino-southern oscillation. Journal of Climate, 13(13), 2177-2194.

Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Ziese, M., Rudolf, B. (2014). GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle. Theoretical and Applied Climatology, 115(1-2), 15-40.

Schurer, A. P., Hegerl, G. C., Mann, M. E., Tett, S. F., & Phipps, S. J. (2013). Separating forced from chaotic climate variability over the past millennium. Journal of Climate, 26(18), 6954-6973.

Shin, S. I., Sardeshmukh, P. D., Webb, R. S. (2010). Optimal tropical sea surface temperature forcing of North American drought. Journal of Climate, 23(14), 3907-3917.

Smith, T. M., Reynolds, R. W. (2003). Extended reconstruction of global sea surface temperatures based on COADS data (1854-1997). Journal of Climate, 16(10), 1495-1510.

Smith, T. M., Reynolds, R. W. (2004). Improved extended reconstruction of SST (1854-1997). Journal of Climate, 17(12), 2466-2477.

Smith, T. M., Reynolds, R. W., Peterson, T. C., Lawrimore, J. (2008). Improvements to NOAA's historical merged land-ocean surface temperature analysis (1880-2006). Journal of Climate, 21(10), 2283-2296.

Shukla, R. P., Tripathi, K. C., Pandey, A. C., & Das, I. M. L. (2011). Prediction of Indian summer monsoon rainfall using Niño indices: a neural network approach. Atmospheric Research, 102(1), 99-109.

Tang, B., Hsieh, W. W., Monahan, A. H., Tangang, F. T. (2000). Skill comparisons between neural networks and canonical correlation analysis in predicting the equatorial Pacific sea surface temperatures. Journal of Climate, 13(1), 287-293.

Tao, F., Yokozawa, M., Zhang, Z., Hayashi, Y., Grassl, H., Fu, C. (2004). Variability in climatology and agricultural production in China in association with the East Asian summer monsoon and El Niño Southern Oscillation. Climate Research, 28(1), 23-30.

Toniazzo, T., Woolnough, S. (2013). Development of warm SST errors in the southern tropical Atlantic in CMIP5 decadal hindcasts. Climate Dynamics, 1-25.

Travasso, M. I., Magrin, G. O., Grondona, M. O., Rodríguez, G. R. (2009). The use of SST and SOI anomalies as indicators of crop yield variability. International journal of climatology, 29(1), 23-29.

Trenberth, K. E., Caron, J. M., Stepaniak, D. P., Worley, S. (2002). Evolution of El Niño–Southern Oscillation and global atmospheric surface temperatures. Journal of Geophysical Research: Atmospheres (1984–2012), 107(D8), AAC-5.

Van den Dool, H. M. (1994). Searching for analogues, how long must we wait?. Tellus A, 46(3), 314-324.

Van Oldenborgh, G. J., & Burgers, G. (2005). Searching for decadal variations in ENSO precipitation teleconnections. Geophysical Research Letters, 32(15).

Vannière, B., Guilyardi, E., Madec, G., Doblas-Reyes, F. J., Woolnough, S. (2013). Using seasonal hindcasts to understand the origin of the equatorial cold tongue bias in CGCMs and its impact on ENSO. Climate dynamics, 40(3-4), 963-981.

Verdin, J., Funk, C., Klaver, R., Roberts, D. (1999). Exploring the correlation between Southern Africa NDVI and Pacific sea surface temperatures: results for the 1998 maize growing season. International Journal of Remote Sensing, 20(10), 2117-2124.

Vimont, D. J. (2012). Analysis of the Atlantic meridional mode using linear inverse modeling: Seasonality and regional influences. Journal of Climate, 25(4), 1194-1212.

Vislocky, R. L., Fritsch, J. M. (1995). Improved model output statistics forecasts through model consensus. Bulletin of the American Meteorological Society, 76(7), 1157-1164.

Wahl, S., Latif, M., Park, W., Keenlyside, N. (2011). On the tropical Atlantic SST warm bias in the Kiel Climate Model. Climate Dynamics, 36(5-6), 891-906.

Wallace, J. M., Smith, C., Bretherton, C. S. (1992). Singular value decomposition of wintertime sea surface temperature and 500-mb height anomalies. Journal of climate, 5(6), 561-576.

Wang, S. Y., L'Heureux, M., & Chia, H. H. (2012). ENSO prediction one year in advance using western North Pacific sea surface temperatures. Geophysical Research Letters, 39(5).

Ward, M. N. (1998). Diagnosis and short-lead time prediction of summer rainfall in tropical North Africa at interannual and multidecadal timescales. Journal of Climate, 11(12), 3167-3191.

Widmann, M. (2005). One-dimensional CCA and SVD, and their relationship to regression maps. Journal of climate, 18(14), 2785-2792.

Xue, Y., Chen, M., Kumar, A., Hu, Z. Z., Wang, W. (2013). Prediction skill and bias of tropical Pacific sea surface temperatures in the NCEP Climate Forecast System version 2. Journal of Climate, 26(15), 5358-5378.

Zebiak, S. E., Cane, M. A. (1987). A Model El Niño-Southern Oscillation. Monthly Weather Review, 115(10), 2262-2278.

Table 1. Input parameters used to reproduce the first case study. Left column represents the statements reproduced by the model with the same format as in the simulation. Right column represents the input parameters entered by the user.

Statements reproduced by the model	Input parameters entered by the user
Enter the NetCDF file containing the predictand data in the path /S4CAST_v2.0/data_files/predictand/ Press enter to continue PREDICTAND data available from Jan-1901 to Mar-2015	
Enter the NetCDF file containing the predictor data in the path /S4CAST_v2.0/data_files/predictor/ Press enter to continue PREDICTOR data available from Jan-1854 to May-2015	
Select a common analysis period The common longest analysis period extends from Jan-1902 to Mar-2015 Do you want to select this period? y/n	'y'
The selected analysis period extends from Jan-1902 to Mar-2015 Select the forecast period Type 1 to select a set of months Type 2 to select one month	1
Enter the forecast period using the initials of the months	'JAS'
2015 forecast available from lead time 1 (monthly lag 4) to lead time 6 (monthly lag 9)	
Enter PREDICTAND spatial domain West longitude from -179.5 to 179.5	-18
East longitude from -179.5 to 179.5	10
South latitude from -89.5 to 89.5	12
North latitude from -89.5 to 89.5	18
Do you want to standardize the predictand? y/n	'y'
Dou you want to apply a Butterworth filter to the predictand? y/n	'n'
Enter PREDICTOR spatial domain West longitude from -180 to 178	-60
East longitude from -180 to 178	20
South latitude from -88 to 88	-20
North latitude from -88 to 88	5
Do you want to standardize the predictor? y/n	'n'
Dou you want to apply a Butterworth filter to the predictor? y/n	'y'
Type 1 to apply a high pass filter Type 2 to apply a low pass filter	1
Introduce the cutoff frequency	7
Select the predictor monthly periods Type 1 to select a set of chronological monthly periods Type 2 to select one monthly period	2
Enter the monthly lag regarding the predictand	3
Select the number of modes for MCA analysis Do you want to select a set of modes? y/n	'n'
Enter the mode number	1
To assess the stationarity the model will analyze 21 years moving correlation windows between the expansion coefficients of the PREDICTOR and PREDICTAND fields obtained from MCA method Indicate delayed, centered or advanced moving correlation windows	'delayed'
To assess the significant stationary periods, indicate the degree of statistical significance from 0 to 100	90
To validate the model skill, indicate the degree of statistical significance from 0 to 100	90

Table 2. Input parameters used to reproduce the second case study. Left column represents the statements reproduced by the model. Right column represents the input parameters.

Statements reproduced by the model	Input parameters entered by the user
Enter the NetCDF file containing the predictand data in the path /S4CAST_v2.0/data_files/predictand/ Press enter to continue PREDICTAND data available from Jan-1854 to May-2015	
Enter the NetCDF file containing the predictor data in the path /S4CAST_v2.0/data_files/predictor/ Press enter to continue PREDICTOR data available from Jan-1854 to May-2015	
Select a common analysis period The common longest analysis period extends from Jan-1855 to May-2015 Do you want to select this period? y/n	'y'
The selected analysis period extends from Jan-1855 to May-2015 Select the forecast period Type 1 to select a set of months Type 2 to select one month	1
Enter the forecast period using the initials of the months	'JAS'
2016 forecast not available	
Enter PREDICTAND spatial domain West longitude from -179.5 to 179.5	120
East longitude from -179.5 to 179.5	-60
South latitude from -89.5 to 89.5	-30
North latitude from -89.5 to 89.5	20
Do you want to standardize the predictand? y/n	'n'
Dou you want to apply a Butterworth filter to the predictand? y/n	'y'
Type 1 to apply a high pass filter Type 2 to apply a low pass filter	1
Enter the cutoff frequency	7
Enter PREDICTOR spatial domain West longitude from -180 to 178	-60
East longitude from -180 to 178	20
South latitude from -88 to 88	-20
North latitude from -88 to 88	5
Do you want to standardize the predictor? y/n	'n'
Dou you want to apply a Butterworth filter to the predictor? y/n	'y'
Type 1 to apply a high pass filter Type 2 to apply a low pass filter	1
Introduce the cutoff frequency	7
Select the predictor monthly periods Type 1 to select a set of chronological monthly periods Type 2 to select one monthly period	2
Enter the monthly lag regarding the predictand	5
Select the number of modes for MCA analysis Do you want to select a set of modes? y/n	'n'
Enter the mode number	1
To assess the stationarity the model will analyze 21 years moving correlation windows between the expansion coefficients of the PREDICTOR and PREDICTAND fields obtained from MCA method Indicate delayed, centered or advanced moving correlation windows	'delayed'
To assess the significant stationary periods, indicate the degree of statistical significance from 0 to 100	90
To validate the model skill, indicate the degree of statistical significance from 0 to 100	90

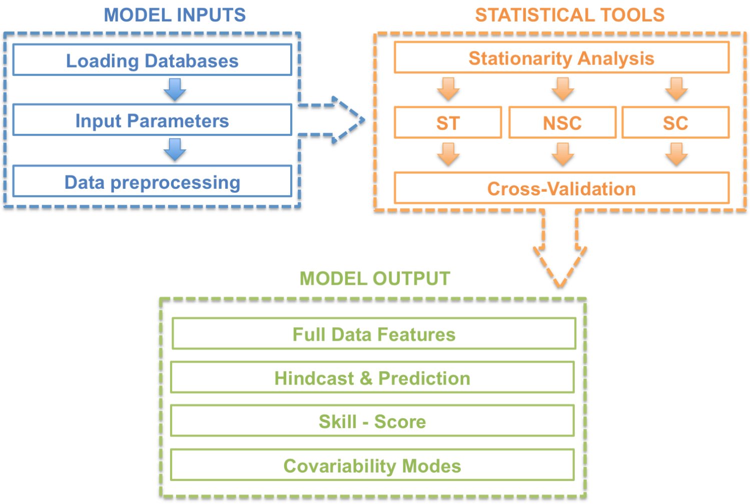

Figure 1. Schematic diagram illustrating the structure of the model.

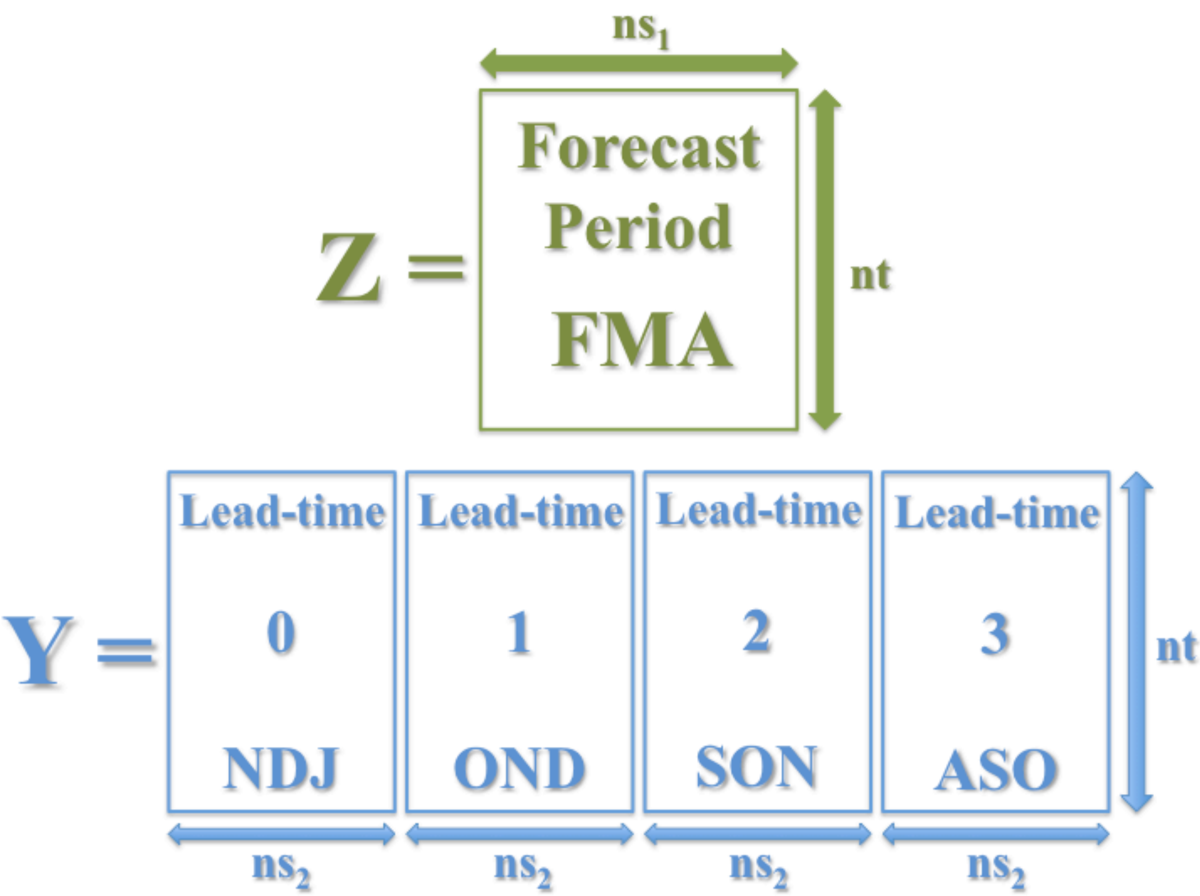

Figure 2. Predictand (Z) and predictor (Y) fields represented by their corresponding data matrices. The illustration relates to an example in which the forecast period covers the months February-March-April (FMA) and the predictor is selected for four distinct seasons: August-September-October (ASO, lead-time=3); September-October-November (SON, lead-time=2); October-November-December (OND, lead-time=1); November-December-January (NDJ, lead-time=0). Each of these sub-matrices for the predictor has the same temporal dimension (nt) and spatial dimension (ns_2). The predictand may have a different spatial dimension (ns_1) but the same temporal dimension (nt) to enable matrix calculations required by MCA methodology.

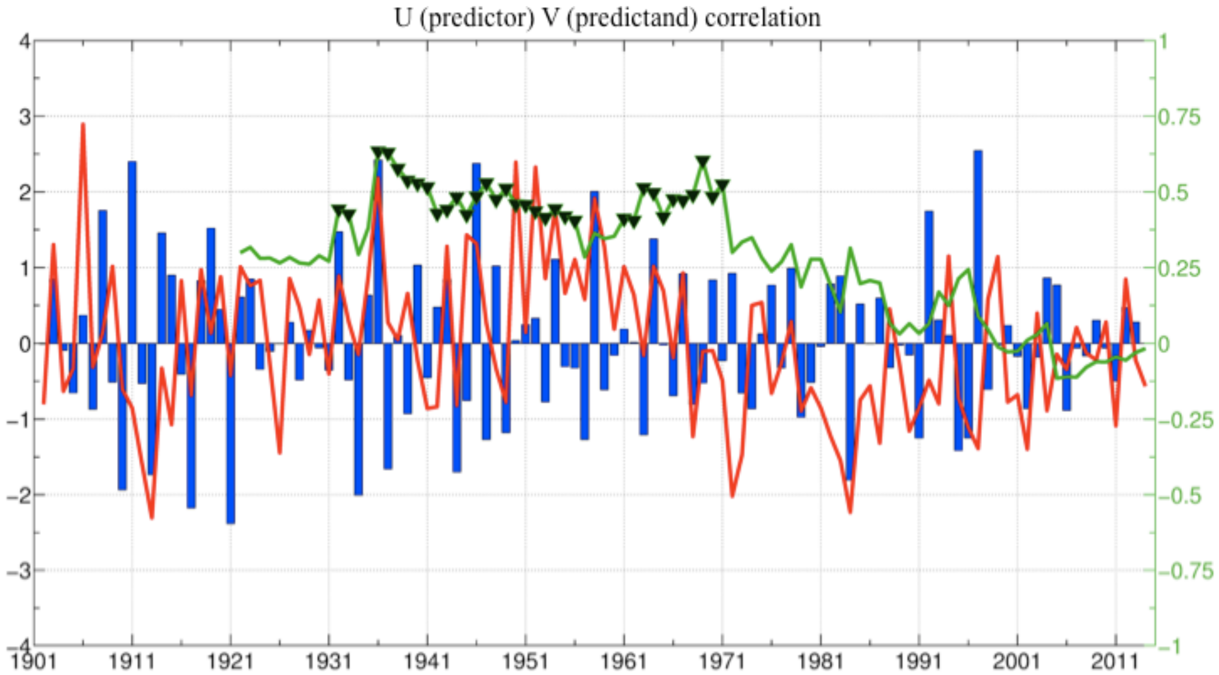

Figure 3. 21 years moving correlation windows (green line) between the expansion coefficients U corresponding to tropical Atlantic SSTA (predictor, blue bars) and V corresponding to Sahelian anomalous rainfall (predictand, red line) obtained for the leading mode of co-variability from MCA analysis. Shaded triangles indicate significant correlation under a Montecarlo Test at 90%.

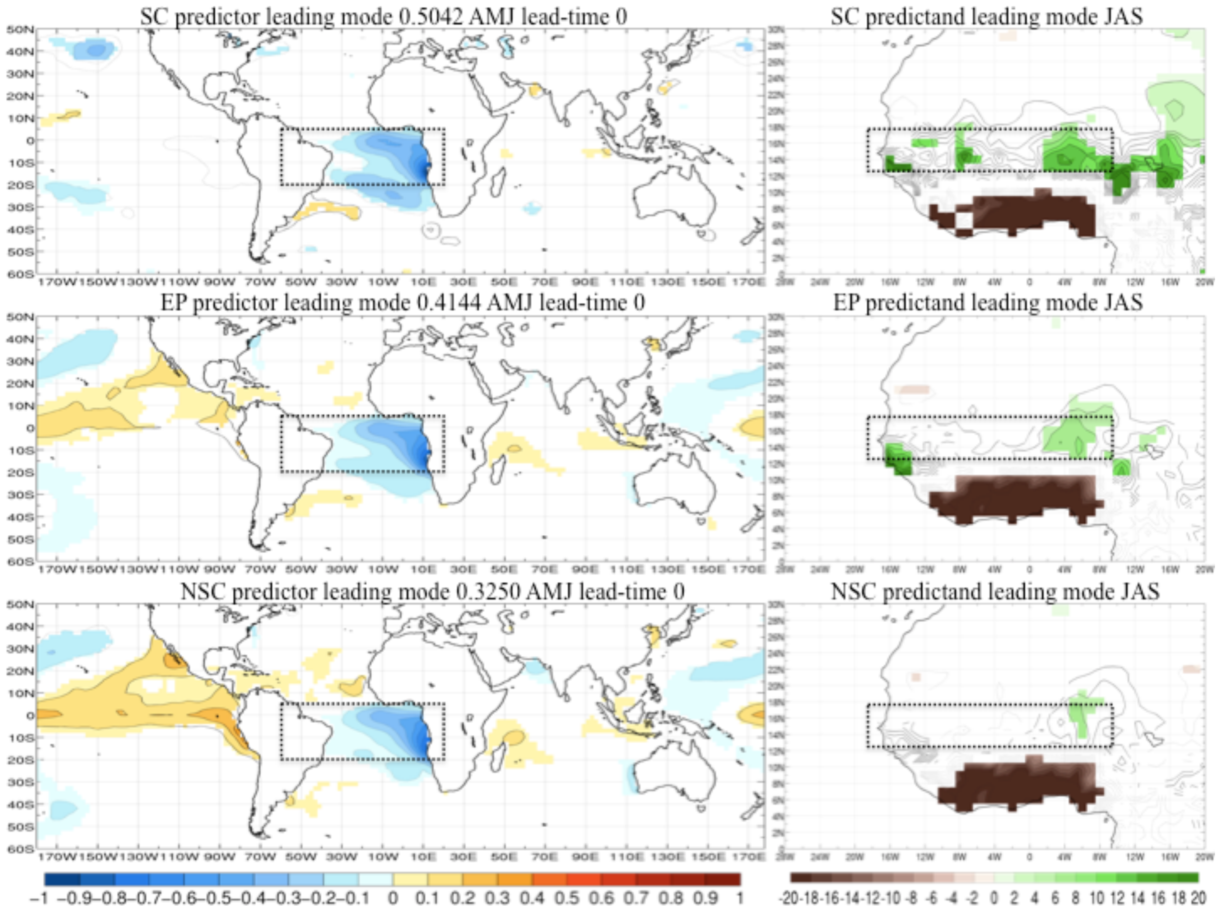

Figure 4. Regression maps obtained for the leading mode by applying MCA between SSTA in the tropical Atlantic (predictor) and western Sahel rainfall (predictand). Left column represents the homogeneous regression map done by projecting the expansion coefficient U onto global SSTA (ºC). Right column represents the heterogeneous regression map done by projecting expansion coefficient U onto the anomalous Sahelian rainfall (mm/day). Period SC (top panels); EP (middle panels) and NSC (bottom panels). Rectangles show the selected regions for predictor and predictand fields considered in the MCA analysis. Values are plotted in regions where statistical significance under a Montecarlo test is higher than 90%.

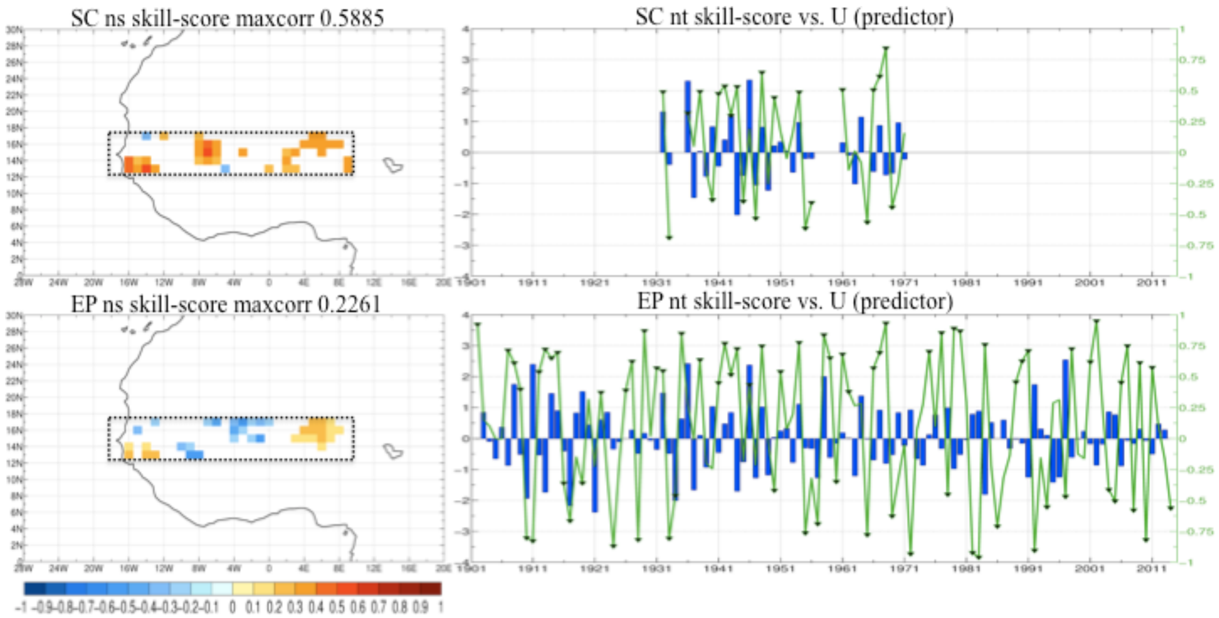

Figure 5. Skill-score validation using Pearson correlation coefficients between observations and hindcasts. Left column corresponds to the spatial validation for each point in space. Right column corresponds to validation time series (green line) between hindcasts and observations considering only the regions indicated by positive significant spatial correlation. Period SC (top panels); EP (bottom panels). Significant correlation values for time series are indicated by shaded triangles. Blue bars correspond to the expansion coefficient (U) of the SSTA (predictor). Significant values are plotted from a 90% statistical significance under a Montecarlo test.

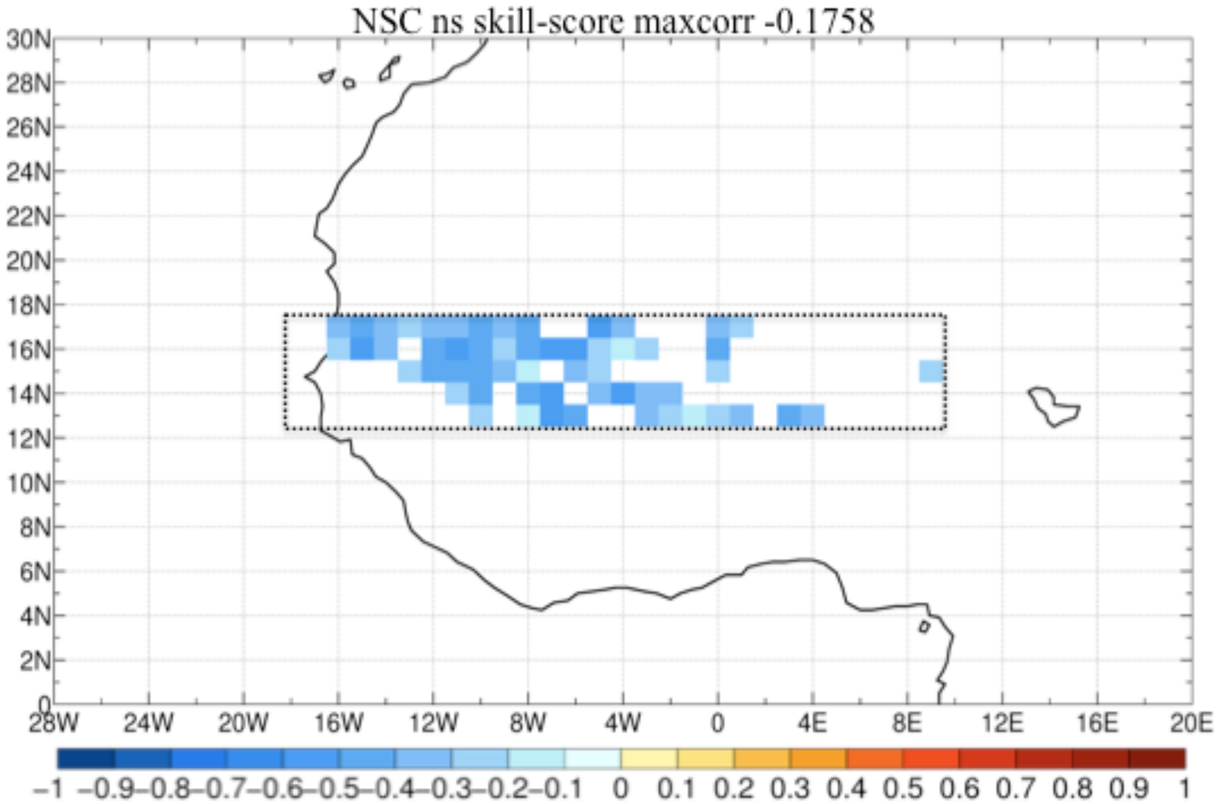

Figure 6. Skill-score validation using Pearson correlation coefficients between observations and hindcasts for each point in space corresponding to NSC period. Significant values are plotted from a 90% statistical significance under a Montecarlo test.

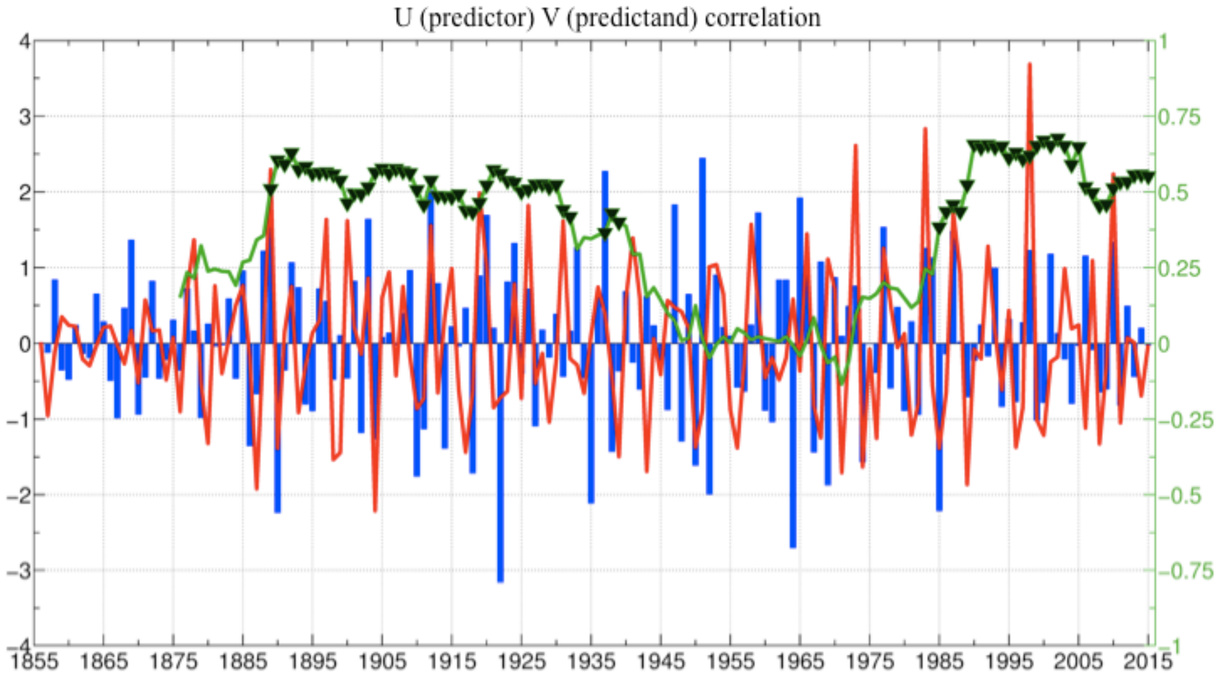

Figure 7. 21 years moving correlation windows (green line) between the expansion coefficients U corresponding to tropical Atlantic SSTA (predictor, blue bars) and V corresponding to tropical Pacific SSTA (predictand, red line) obtained for the leading mode of co-variability from MCA analysis between predictor and predictand fields. Shaded triangles indicate significant correlation under a Montecarlo Test at 90%.

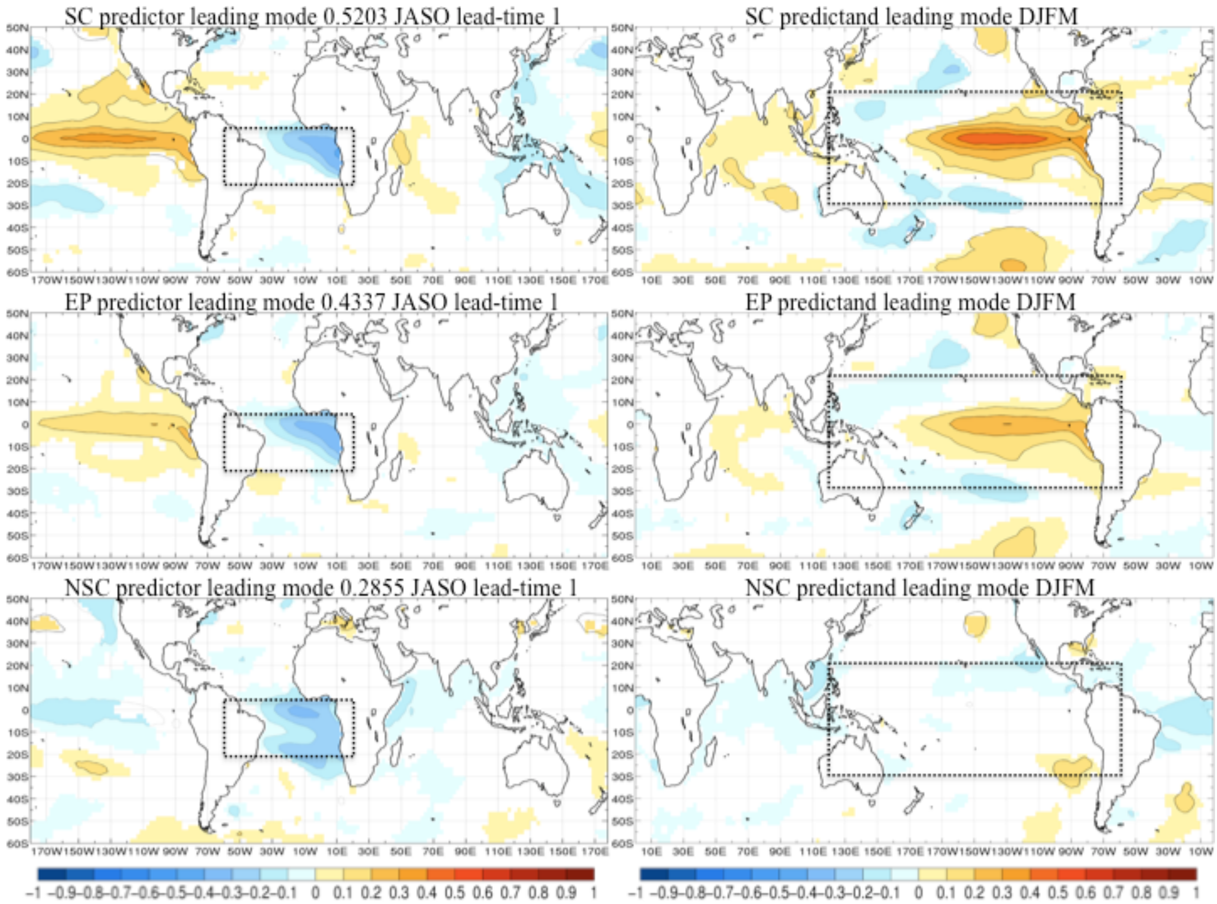

Figure 8. Regression maps obtained for the leading mode by applying MCA between SSTA in the tropical Atlantic (predictor) and SSTA in the tropical Pacific (predictand). Left column represents the homogeneous regression map done by projecting the expansion coefficient U onto global SSTA (ºC) for predictor seasonal period. Right column represents the heterogeneous regression map done by projecting expansion coefficient U onto global SSTA (ºC) for predictand seasonal period. Period SC (top panels); EP (middle panels) and NSC (bottom panels). Rectangles show the selected regions for predictor and predictand fields considered in the MCA analysis. Values are plotted in regions where statistical significance under a Montecarlo test is higher than 90%.

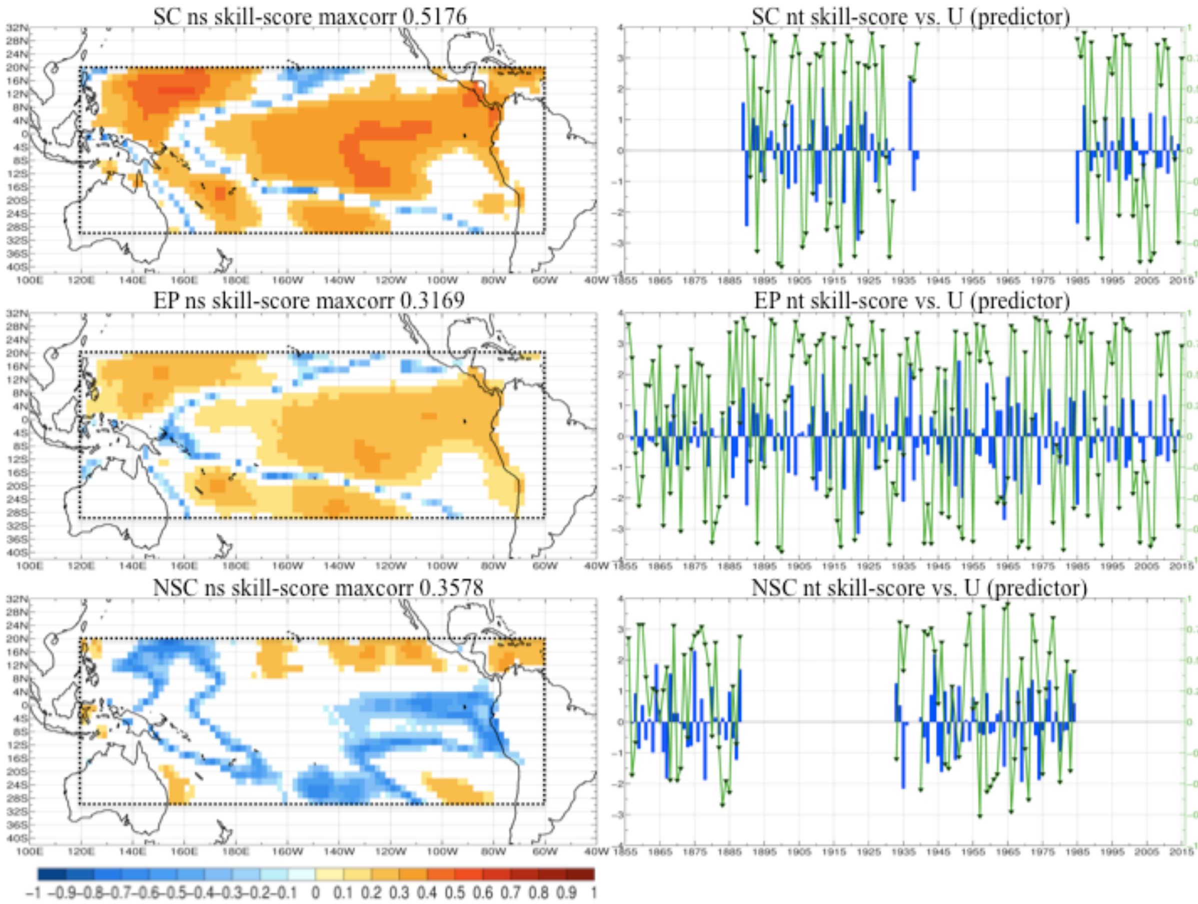

Figure 9. Skill-score validation using Pearson correlation coefficients between observations and hindcasts. Left column corresponds to the spatial validation for each point in space. Right column corresponds to validation time series (green line) between hindcasts and observations considering only the regions indicated by positive significant spatial correlation. Period SC (top panels); EP (middle panels); NSC (bottom panels). Significant correlation values for time series are indicated by shaded triangles. Blue bars correspond to the expansion coefficient (U) of the SSTA (predictor). Significant values are plotted from a 90% statistical significance under a Montecarlo test.