ECLISEA

European advances on **CLI**mate services for coasts and **SEA**s

WP 3 - Predictability and uncertainty of SSDs on seasonal, decadal and long-term future climate projections.

D3.B - Synthesis report on the research about seasonal prediction of waves and surge.

Start date of project: 1th November 2017 Duration: 36 months

Due Date of Deliverable: May 2019 (M20) Actual Submission Date: April 2019

Authors: S. Herrera, P. Camus, I. Losada

Table of Contents

1. Introduction

ECLISEA aims to advance coastal and marine climate science and associated services through developing innovative research of sea surface dynamics (SSD). To this aim, the project's consortium (five multidisciplinary European leader RPOs on coastal, marine and climate services) has identified three main aspects in which the project should be focused on (i) collecting and developing new data and climate information by supporting harmonized analysis at European scale, (ii) investigation on climate variability, predictability and long-term projections, and (iii) of the complexity and uncertainties of the modeling frameworks.

ECLISEA proposes an integral, and applicable throughout Europe, research plan that starts with assessing the needs of the stakeholders at different time-scales (seasonal, decadal and climate change), and ends up with the development of a prototype of a coastal climate service in Europe.

Working Package 3 (WP3) aims to obtain predictions/projections of SSD and to assess the corresponding uncertainties in specific topics, such as regional mean sea level rise, seasonal-to-decadal prediction and performance of climate model outputs required to simulate SSDs. This report presents an overall understanding of the current state of the seasonal climate forecasting: description, data available, forecast validation and uncertainty communication.

2. Seasonal climate forecast

Many sectors such as agriculture, health or energy, among others, rely on accurate forecasts to make management decisions. Forecasts provide a qualitative tool for the assessment of weather and climate risk on a range of time scales from days to decades. Seasonal climate forecast provides information on the average seasonal weather conditions which can be expected from a few months up to one year in advance (Doblas-Reyes et al., 2013). Due to the chaotic nature of the climatic system, accurate daily predictions, as those given by the weather forecast system (e.g. which significant wave height is going to be in Santander the next May 15th), are impossible beyond 15 days, so the seasonal forecast aims to predict deviations from the mean seasonal climatology some months in advance (e.g., next season will be above-, near- or below- normal if tercile-based categories are used).

The feasibility of seasonal climate prediction largely rests on the existence of slow, and predictable, variations in the storage of heat and moisture by the ocean and the land, the presence or absence of snow and sea ice, and how the atmosphere interacts and is affected by these boundary conditions (Doblas-Reyes et al., 2013). Predictability at this longer time-scale is influenced by components of the global climate system that change at slower rate than weather events, especially sea surface temperature (SST), which can influence the weather at some regions. The Niño-Southern Oscillation (ENSO) is a good example of a climatic phenomenon that contributes to the forecast quality on seasonal time scale (Palmer et al., 2005, Harrison et al., 2008a). A warm SST anomaly in the tropical Pacific Ocean leads to increased heat flux from the ocean to the atmosphere (coupled system). The extra latent heat release will impact the atmospheric circulation leading to climatic anomalies in remote regions of the globe (atmospheric teleconnections). Regarding the North Atlantic sector, it is

known that wintertime mean wind and wave conditions are statistically associated with large scale atmospheric circulation patterns such as the North Atlantic Oscillation (NAO), as well as the East Atlantic and Scandinavian patterns (EA, SCAND), so it would be expectable that some skill in predicting these circulation patterns should consequently translate into skill in predicting wind speed and wave heights in this region. However, the seasonal predictability of the winter climate in the Northern Hemisphere extratropics is generally assumed to be low.

Both statistical/empirical and dynamical methods are used to generate seasonal climate predictions. The first approach is commonly based on statistical methods which relate the observed seasonal anomalies of the variable of interest (e.g., precipitation) with observed lagged SST anomalies (based on atmospheric teleconnections), usually by means of multivariate linear regression methods. The dynamical method is based on complex dynamical numerical models of the different components of the earth system. Mixed methodologies are also employed because statistical post-processing of the dynamical predictions is required by the users. Both methods are complementary because advances in statistical prediction are often associated with enhanced understanding, which leads to improve dynamical prediction, and vice versa.

In the last years, the advances in seasonal forecasting have related to the use of the dynamical methods based on fully coupled general circulation models (GCMs) (Troccoli, 2010), numerically solving the system of partial differential equations (PDEs) describing the processes that occur in the different components of the global climate system (e.g. the atmosphere and the oceans) and their interactions. To this aim, on the one hand the three-dimensional space should be discretized in grid boxes, with a typical horizontal resolution of about hundreds of kilometers, and on the other hand the equations are also simplified by making assumptions (e.g. hydrostatic approximation) and/or using parameterizations for modelling some processes at spatial scales smaller than the grid box size, which are commonly different from model to model. GCMs are in continually development in complexity due to the increase of the scientific understanding and computational resources.

Due to the non-linear character of the PDEs and the model-dependence of the parameterizations and the numerical schemes (Palmer and Anderson, 1994) used by the GCM, at these time-scales (weather and seasonal forecast) the prediction is very sensitive to both the initial conditions (Lorenz, 1963) and the model considered. The first is usually overcome through the use of an ensemble of predictions, consisting of running the model several times from slightly different initial conditions (members). The second is taken into account by considering a multimodel approach where different GCMs are combined. Multimodels generally produce more skillful forecasts due to a reduction of the ensemble error mean and a reduction of overconfidence because the ensemble spread is widened (Palmer et al., 2004, Wang et al., 2009).

A particularly valuable aspect of ensemble forecasting is its capacity to yield information about the magnitude and nature of the uncertainty in a forecast. In this case, the full set of deterministic forecasts is used for obtaining a more reliable and skilful estimate of the forecast probability.

The capability for seasonal climate predictions has been evolving in the last decades through different multimodel ensemble systems in order to obtain results not dependent on

the model formulation and initial conditions. The improved collaboration between different centers and the increasing computing resources have promoted several initiatives for developing multi-model seasonal forecast systems. In Europe, the European Center for Medium–range Weather Forecasts (ECMWF) has pioneered the multimodel ensemble approach in the development of consecutive European projects (PROVOST, DEMETER, ENSEMBLES). The potential for the use of the multimodel ensemble in seasonal climate prediction was first addressed in PROVOST. These initiatives motivated the collaboration at international level in the last years in the European Operational Seasonal to Interannual Prediction (EUROSIP). The most recent initiative is the Climate Change Service (C3S) as part of the Copernicus Programme that plans to provide data from several state-of-the-art seasonal prediction systems.

Despite the recent developments on the seasonal forecast, Only thirteen centres around the world produce seasonal forecast on a global scale, all of them only focused in providing atmospheric (wind, sea level pressure, precipitation) and ocean (sea surface temperature) dynamics, so any operational seasonal forecast system generates marine predictions (waves, storm surge or currents).

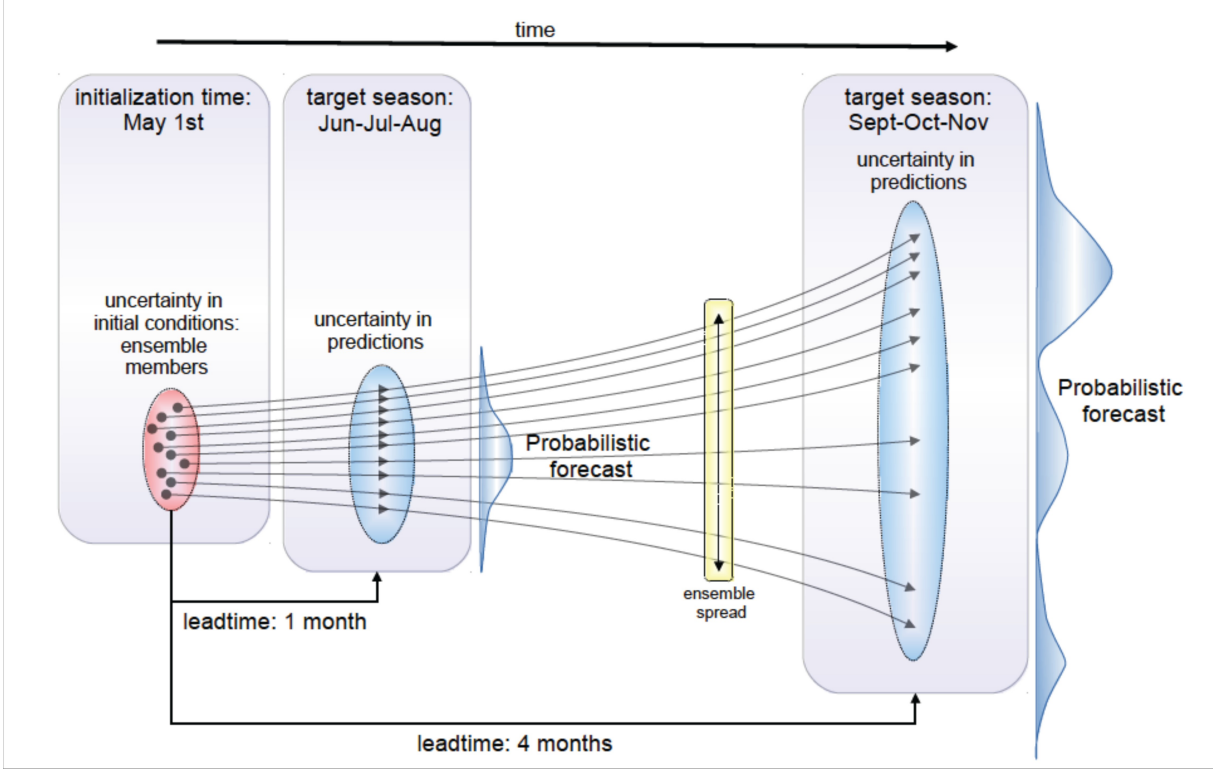

Figure 1. Probabilistic Seasonal Forecast obtained from different initial conditions (Manzanas 2016).

Particularities of the seasonal forecasting

As is shown in Figure 1, seasonal forecast presents some particularities:

1) GCMs are initialized several times (members) per month and run forward in time from six months to one year. Two terms are defined in this context: initialization time (moment in which the model is initialized) and lead-time (the time passed from the initialization moment to the beginning of the target season to be predicted).

2) Seasonal climate forecasts are intrinsically probabilistic. An ensemble of simulations is generated by using a set of slightly different initial conditions which provides forecasts with different atmospheric trajectories (members) compatible with the underlying slow variables.

Seasonal forecast verification

Despite the potential value of seasonal forecasts, as has been pointed out previously, there are still a number of problems which limit the practical application of this type of predictions. Note that the current models for seasonal forecasting are not globally reliable and, then, for each variable and season, their usefulness is limited to certain regions of the world, mainly the tropics. As a result, seasonal forecast verification is mandatory to ensure when, where and how much the predictions can be trusted based on past performance (Doblas-Reyes et al., 2013). The validation of the seasonal forecast (forecast verification) is a multi-faceted task in which several performance measures providing complementary information on different aspects of forecast quality (association, accuracy, reliability, etc.) should be considered to provide users with information about possible skill (quality) in future forecasts (Murphy, 1993). In this sense, the probabilistic communication of uncertainty in seasonal forecast products has been highlighted as a major challenge in the practical application of seasonal predictions in different economic sectors (see e.g. Mason, 2008; Lemos et al., 2012; Raftery, 2016).

In order to properly evaluate these probabilistic forecasts, for each seasonal forecast system a retrospective seasonal forecast, known as hindcast, covering a long period (ideally, over 30 years) is needed. The hindcast is used to build trust (or not) in the probabilistic forecast system issuing an actionable future forecast. A strong signal in a future forecast (i.e. very high or low probability of occurrence for an event) is useless or even dangerously misleading, if the forecast system has no skill in predicting such events, something revealed by analysing the historical hindcast performance. For this reason it is recommended to blend forecast information with the verification results of the forecast system, as an indication of its past performance.

The main properties considered for the forecast verification in this deliverable are briefly described:

- **Association** reflects the strength of the relationship between the forecasts and the corresponding observations. In this case the Pearson correlation between the ensemble mean and the observed interannual time-series.

- **Accuracy** measures the average distance between forecasts and observations. We consider here the Ranked Probability Score (RPS, Epstein 1969) that measures the sum of squared differences in cumulative probability space for a multi-category probabilistic forecast, in our case terciles, being thus its perfect value 0.

- **Discrimination** measures the ability of the forecasts to distinguish between an event and the corresponding non-event, which is assessed here by means of the area under the ROC curve Kharin and Zwiers (2003) (simply referred to as ROC hereafter).

3. Climate data

In the framework of this project, the current seasonal climate forecasts given by the Climate Data Store (CDS) from the Copernicus Climate Change Service (C3S) have been

considered. C3S is designed to provide the data from the hindcast and forecast of several state-of-the-art seasonal prediction systems, among other datasets. It is defined as one of the six thematic streams of the Copernicus Programme coordinated and managed by the European Commision (http://www.copernicus.eu/). Table 1 shows the seasonal forecast systems considered in this deliverable and their main properties. Note that all the models are distributed in a common resolution and spatial grid of 1º.

For each seasonal forecast system, the variables identified as relevant for the SSD and used as predictors for the statistical downscaling methods have been downloaded and linked through the Santander User Data Gateway (http://meteo.unican.es/udg-tap/home) to be accessible through the Climate4R tools (Iturbide et al. 2019; http://meteo.unican.es/en/climate4R). In particular, the 6-hourly Sea Level Pressure (SLP) and Sea Surface Temperature (SST) were downloaded for the initializations corresponding to the standard seasons: February → Spring; May → Summer; August → Autumn; November → Winter. In addition, also the output for the forecast of the Spring of 2019 was considered when available in order to illustrate the proper visualization of the seasonal forecast.

Table 1. List of Seasonal Forecast Systems considered in this deliverable.

Institution	Model	Code	Members	Period
ECMWF	SEAS5	ECMWF - System 5	25 members	1993-2016
UK Met Office	GloSea5	UKMO – System 12	12 members	1993-2015
Météo France	System5	Meteo France - System5	15 members	1993-2014
DWD	GCFS2.0	DWD – System 2	30 members	1993-2016

Note that previous studies (Camus et al. 2019) have been done using the Climate Forecast System version 2 (CFSv2) from the National Centers for Environmental Prediction (NCEP) due to his availability and that this seasonal forecast is freely distributed by the NCEP, including both a long hindcast, needed for forecast verification, and the operational forecast. The CFSv2 (Saha et al., 2011) is an ensemble retrospective seasonal climate forecast dataset with 24 members that extends along 28 years from 1982 to 2009. Beginning at January 1st, 9-month hindcasts were initiated every 5 days with 4 cycles on those days. For each calendar month, the hindcasts with initial dates after the 7th of that month were used as the ensemble members of the next month. For instance, the starting dates for the February ensemble members are the January 11th, 16th, 21st, 26th, 31st, and the February 5th. In general, there are at least 16 members having initial dates before the 1st day of the target month. The CFSv2 used in the reforecast consists of the NCEP Global Forecast System at T126 (~0.937°) resolution, the Geophysical Fluid Dynamics Laboratory Modular Ocean Model version 4.0 at 0.25–0.5° grid spacing coupled with a two-layer sea ice model, and the four-layer NOAH land surface model.

In order to evaluate the skill of the predictors considered by the statistical downscaling methods (see Section 4) also the reanalysis ERA-Interim (Dee et al. 2014 ***) was considered. Both the SLP and SST were loaded for the hindcast period (1993-2016) and regridded to the resolution of the seasonal forecast (1º).

Data access

All the data processing has been done using the recently developed climate4R tool (Iturbide et al., 2018), an R-based climate services oriented framework that integrates a number of R packages implementing data access (loadeR, Cofiño et al., 2018), data collocation (transformeR package, Bedia and Iturbide, 2016), bias adjustment and downscaling (downscaleR package, Bedia et al., 2017b), as well as blended visualization of probabilistic forecasts and quality measures, (visualizeR, Frías et al., 2018). All the packages are available through the SantanderMetGroup public GitHub Repository (https://github.com/SantanderMetGroup/). In the Appendix A of this deliverable we have included the main code used to obtain the results included in this report in order to ensure the reproducibility of the results shown.

4. Statistical Downscaling

The low spatial resolution of these models is commonly insufficient for most practical applications, so their raw outputs cannot be directly used as prediction tool for resource management and planning decisions. A postprocess is needed to translate the coarse global seasonal forecast to the useful local-scale (downscaling). In the case of marine conditions, the downscaling must be performed to infer local wave conditions from seasonal atmospheric predictions because these physical process and variables are not simulated/considered by the seasonal forecast systems.

The problem of downscaling is addressed by applying two complementary methodologies. On one hand, the dynamical downscaling increases the spatial resolution of global models by means of nesting regional model in the area of interest (e.g. Europe). In the case of wave climate, the atmospheric fields at global scale define the boundary conditions of the global wave generation model. Regarding the dynamical process to transfer wave climate from deep water to shallow water, a wave propagation model at local scale is nested to the wave boundary conditions from the wave generation model, simulating the wave transformation processes due to its interaction with the bathymetry. The dynamical downscaling is the best solution in reproducing waves in shallow water but requires enormous computational effort and is affected by systematic biases, being necessary to apply some post-process (bias-correction methods) to the model output.

On the other hand, statistical downscaling is based on empirical models that link meteorological variables from the global models to local variables related to the atmospheric phenomenon (see Gutiérrez et al. 2018). Some assumptions are inherent in the statistical downscaling approach: (i) variability of the local variable should be explained by the statistical connection, (ii) changes in the mean climate should lie within the range of its natural variability, and (iii) the relationships should be stationary. Long observation time series, physical explanation of the relation between the large-scale predictor and the local predictand, and reliable predictor simulations by Global Circulation Models (GCMs) fulfill these conditions (Wilby et al., 2004). Note that the main shortcoming of statistical

downscaling in climate change applications is the assumption of the stationarity between global and local variables (Gutiérrez et al., 2013) but in the case of seasonal forecast predictions not significant changes from historical data are expected.

Other relevant aspects which determine the skill of the SD method are the predictor choice, regarding variables and the spatial domain (Fowler et al., 2007). In the case of sea surface waves, sea level pressure (SLP) fields and the squared SLP gradient fields have demonstrated to be good predictors (Wang et al., 2012; Casas-Prat et al., 2014).

In the case of climate change projections, the statistical downscaling methods are usually based on multivariate regression models. Significant wave height has been predicted at 6-hourly resolution from sea level pressure fields (SLP) at a global [Wang et al., 2014a, 2014b] or regional (Casas-Prat et al., 2014) scale, or at seasonal to interannual time scale (Martínez-Asensio et al., 2016). Mori et al. (2013) projected wave height using an empirical formula as a function of sea surface winds. Wave climate changes from multimodel ensemble over Europe are obtained using a statistical downscaling approach based on SLP weather types (WTs) and for an optimal ensemble of models selected according to a skill criterion (Perez et al., 2015). This statistical downscaling method was extended to obtain global wave climate projections by Camus et al. 2017 and, more recently, adapted to be applied for seasonal forecast.

First, the wave predictors, SLP and SST, are obtained from the retrospective seasonal forecast and their skill at this time-scale is evaluated in order to establish the benchmark given by the input of the statistical downscaling method.

The predictor defined by daily sea level pressure (SLP) fields from the reanalysis CFSR atmospheric database over the wave generation area of the local waves (predictand) is classified in a reduced number (100 in this work) of WTs. A regression guided classification is applied to a combination of the weighted predictor and predictand estimations from a regression model, which links the SLP fields with local marine climate. First, the statistical relationship is established identifying hourly sea states parameters at each location of interest during each daily predictor field within the corresponding cluster. Then, the empirical probability distribution of each sea state parameter (e.g., significant wave height) associated with each WT is calculated. Finally, the complete distribution of this variable for a time period can be estimated as the sum of the probability of each WT during that period multiplied by the corresponding empirical distribution. As a result, different statistics (e.g., mean, 95th percentile) can be derived from the estimated distribution.

The spatial domain is based on the wave generation patterns obtained in Camus et al. (2017). In particular, the predictor domain considered extends from 64°W to 16°E and from 0°N to 76°N. The predictor is defined as the 3-daily mean SLP, i.e., the predictor at each specific day is calculated as the average of the SLP at the same day and the previous 2 days through the historical time period. The predictor is defined by the leading principal components (PCs) explaining 95% of the entire predictor variance. PCs are calculated for the seasonal forecasts by projecting the corresponding standardized fields onto the empirical orthogonal functions obtained from the reanalysis, used for the calibration of the method.

Following Manzanas (2016), who obtained a more skillful statistical downscaling model for seasonal precipitation forecast using season-specific data in the model calibration, a particular regression-guided classification is performed at every wave GOW2 grid node at 1.0^{o} resolution taking into account multivariate wave conditions (Hs, Tp, θ) in each season

independently. 100 WTs of the SLP fields are obtained for every GOW2 grid node. The seasonal empirical distribution of hourly significant wave height associated with each WT at every grid node of the GOW2 wave database is calculated..

The most similar semi-guided WT is identified for each 3-daily mean SLP field from the hindcast database to calculate the probability of WTs and infer seasonal empirical distribution of the significant wave height at each grid node at the target season. The seasonal predictions of the mean and the 95th and 99th percentiles of the significant wave height are obtained to assess the seasonal forecast quality.

5. Results

As has been reflected in the previous sections the proper evaluation of a seasonal forecast system of waves and surge should include:

- A verification of the predictors, in this case SLP and SST, given by the different hindcasts.

- A verification of the predictions obtained by applying the statistical downscaling method to the hindcast.

According to this, we have divided this section in different subsections associated with the previously defined points. In addition, we include as an example the forecast of the Spring corresponding to this year 2019 in order to reflect the proper way to introduce the uncertainty of the forecast in the predictions.

For the sake of the brevity, a more detailed description of the results has been included in the Appendix A of this deliverable, giving here a summary with the main results and conclusions.

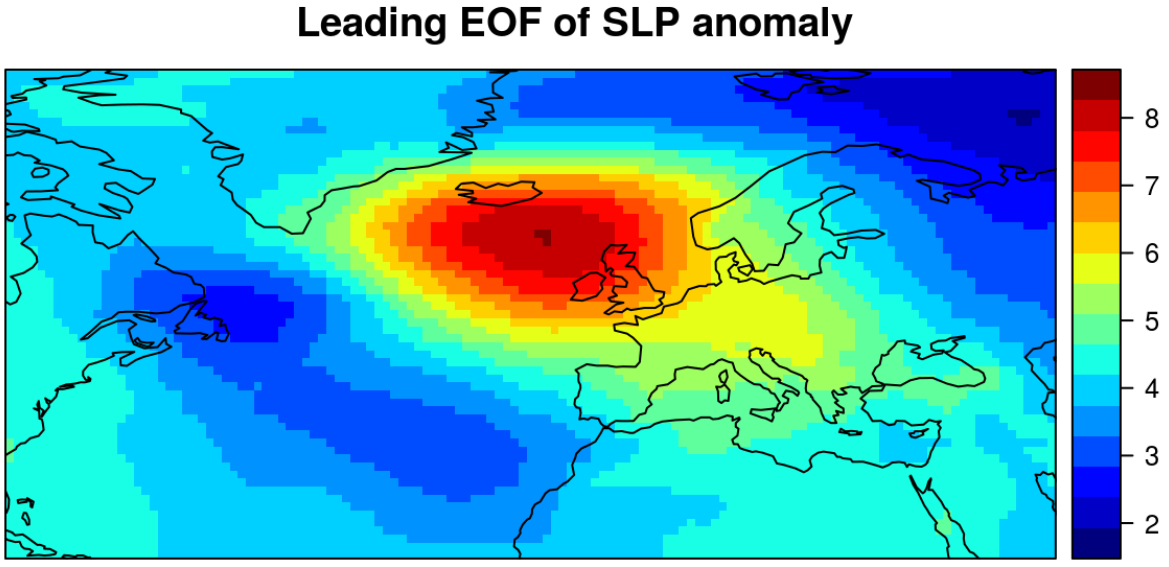

Figure 2. Winter North Atlantic Oscillation pattern.

Verification of the Climatic Predictors:

Taking into account the limitations reflected in the literature about the predictability of

the seasonal forecast in Europe and the association of the mean wind and wave conditions in Winter with the North Atlantic Oscillation (NAO) pattern (see Figure 2), we first evaluate the capability of the hindcast to properly reproduce the interannual variability of this pattern.

Figure 3 shows the limitations of the ECMWF – System 5 seasonal forecast system to reproduce the interannual variability of the NAO-Index and, as a consequence, the positive and negative phases of this index and their effect on the local climate in Europe. The same results have been obtained for the others hindcast (see Appendix A) reflecting the difficulties of the GCMs to reproduce the North Atlantic Oscillation Pattern.

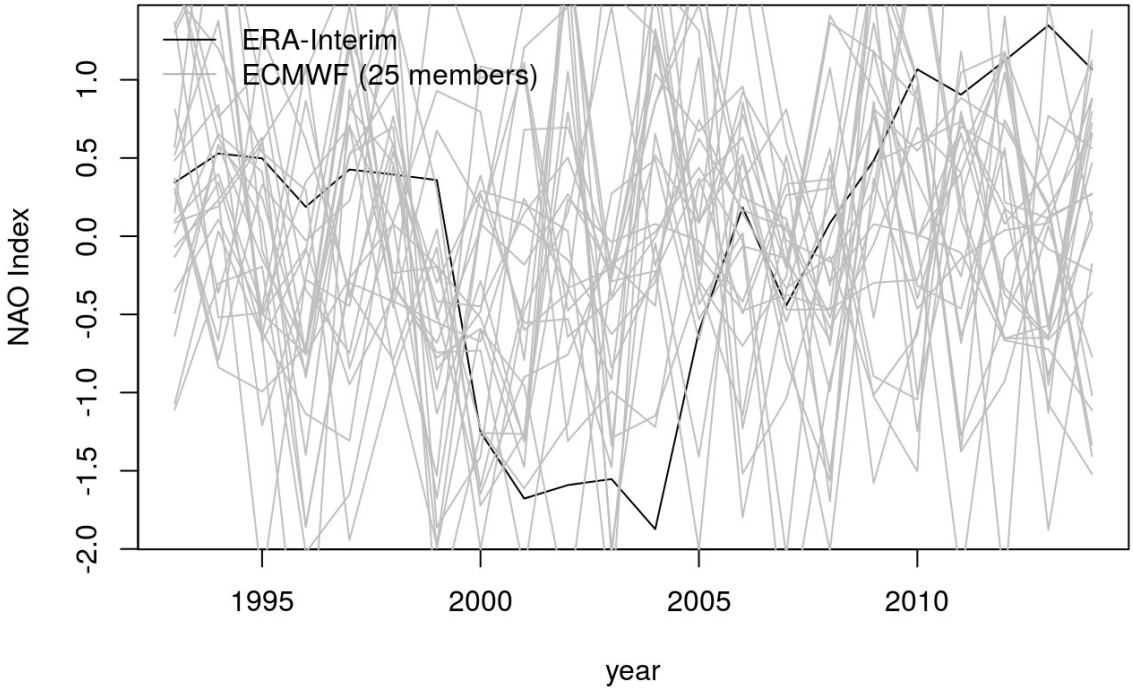

Figure 3. Winter NAO-Index for ERA-Interim (black line) and the 25 members (gray lines) of the ECMWF – System5 seasonal forecast system for the hindcast period (1993-2016).

As has been reflected in Section 2 – Seasonal Forecast Verification, several validation measures should be considered to evaluate the different properties of the seasonal forecast system. In particular, we consider the Pearson correlation coefficient, the Ranked Probability Score and the Area under the ROC curve to evaluate the association, accuracy and discrimination capabilities of the seasonal forecast system.

Figures 4 y 5 show the Pearson correlation of the seasonal Sea Level Pressure and the Sea Surface Temperature between the ensemble mean of the seasonal forecast and the observed interannual time-series. First, strongest correlations have been found for all the hindcast and season for the SST than the SLP. Regarding the spatial pattern of the Pearson correlation for the SLP, relevant differences exist between hindcasts with the largest values located in the north of the domain, although the sign depends on the season and hindcast. For SST the strongest correlation is found over the upper half of the spatial domain, being it more stable between hindcast and seasons.

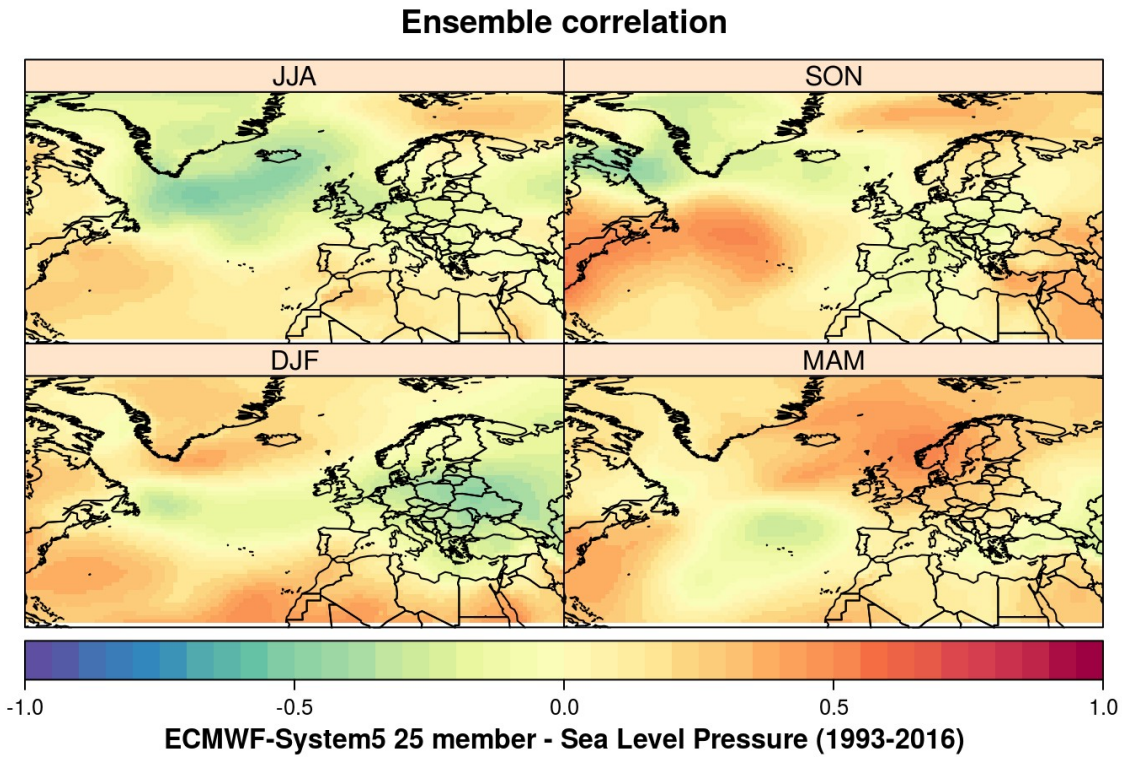

Figure 4. Pearson correlation between the Sea Level Pressure of the ensemble mean of the ECMWF – System5 hindcast and ERA-Interim for the period 1993-2016.

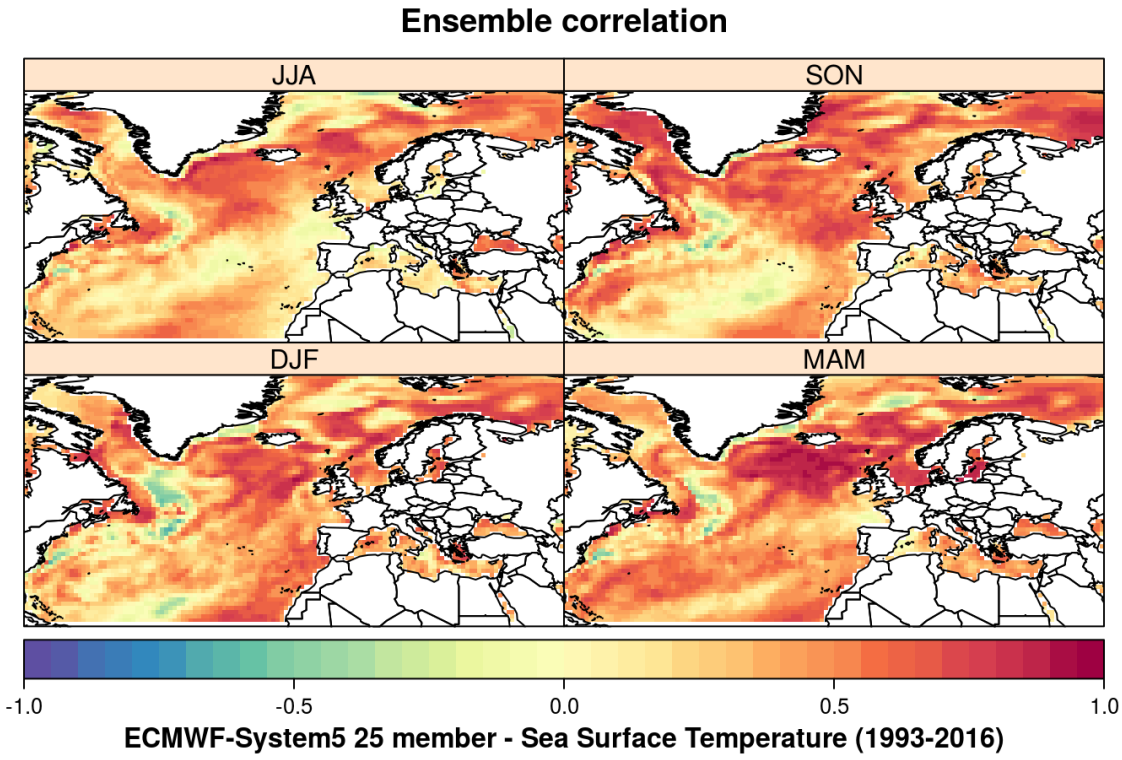

Figure 5. Pearson correlation between the Sea Surface Temperature of the ensemble mean of the ECMWF – System5 hindcast and ERA-Interim for the period 1993-2016.

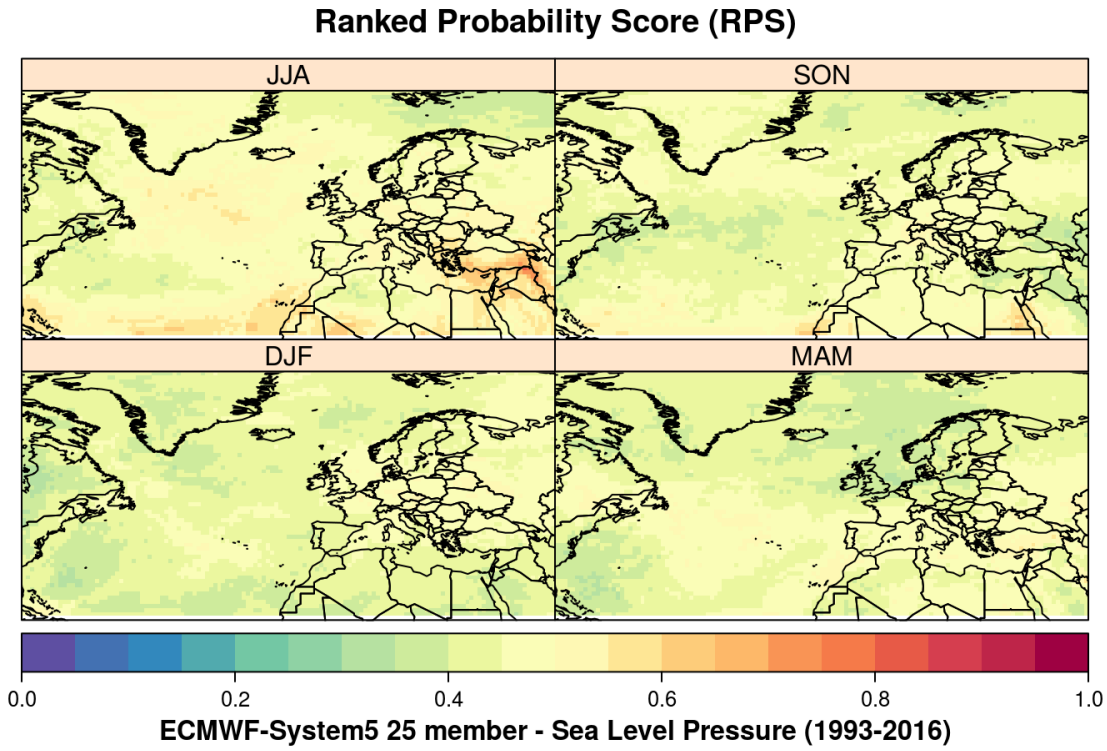

Figure 6. Ranked Probability Score of the seasonal Sea Level Pressure for the ECMWF – System5 hindcast for the period 1993-2016.

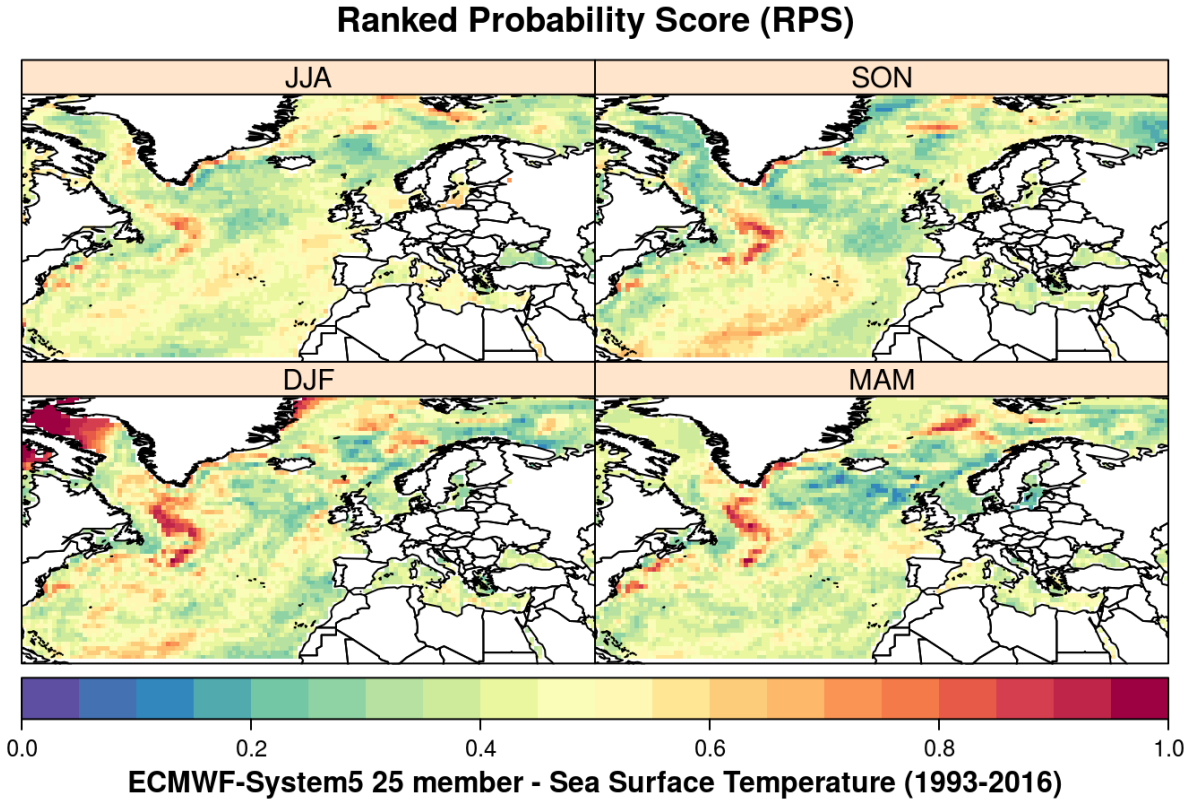

Figure 7. Ranked Probability Score of the seasonal Sea Surface Temperature for the ECMWF – System5 hindcast for the period 1993-2016.

Regarding the accuracy of the model Figures 6 y 7 show the Ranked Probability Score of the the ECMWF – System 5 for seasonal Sea Level Pressure and the Sea Surface Temperature. First, note that almost all the pattern ranges between 0.3 and 0.7 reflecting a very low accuracy of this model for both variables. Only some isolated areas present values close to 0 or 1 for the SST, reflecting some sort of predictability although, in this case, the model tends to predict the opposite tercile, in agreement with the results shown in Figure 5. Same results and conclusions are obtained for the rest of the models (see Appendix A) in agreement with previous results.

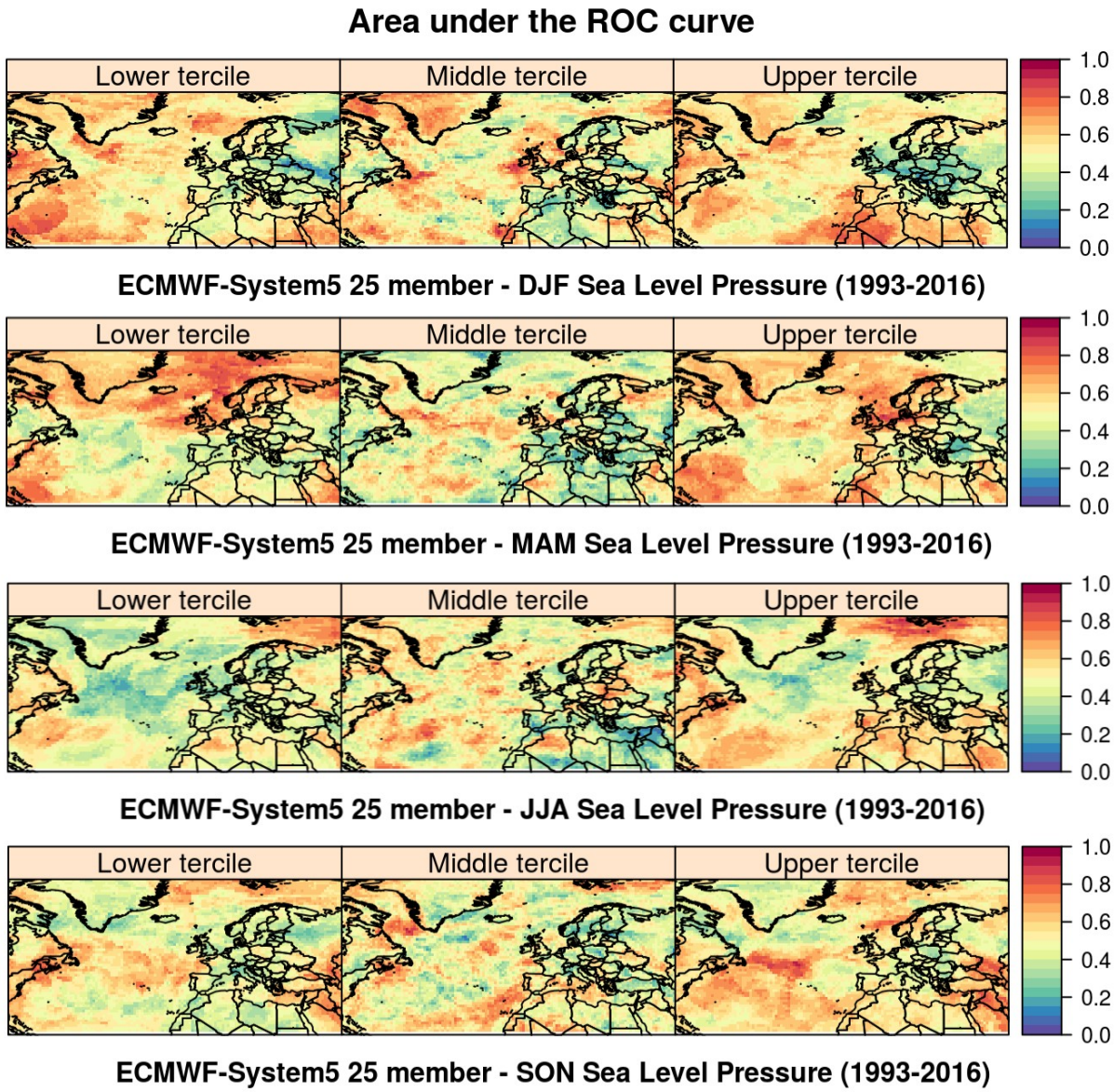

Figure 8. Area under the ROC curve of the seasonal Sea Level Pressure for the ECMWF – System5 hindcast for the period 1993-2016.

Finally, in order to evaluate the discrimination capability of the hindcast, the Area under the ROC curve is considered for each tercile, season and model. Figure 8 shows the results obtained for the Sea Level Pressure and the ECMWF – Sytem 5 hindcast. First, note that there is a great seasonal variability in the pattern. Second, there is not a clear signal for any region or season, being the non occurrence of a particular tercile and season the most significant. In this case, although the spatial pattern is very similar, the rest of models are

able to re-inforce the signal shown in Figure 8 in most of the domain. Some models present regions with high ROC values (e.g. the upper tercile in Africa for Météo-France System 5) but without a clear agreement between the different GCMs, pointing out to a great model uncertainty in these regions.

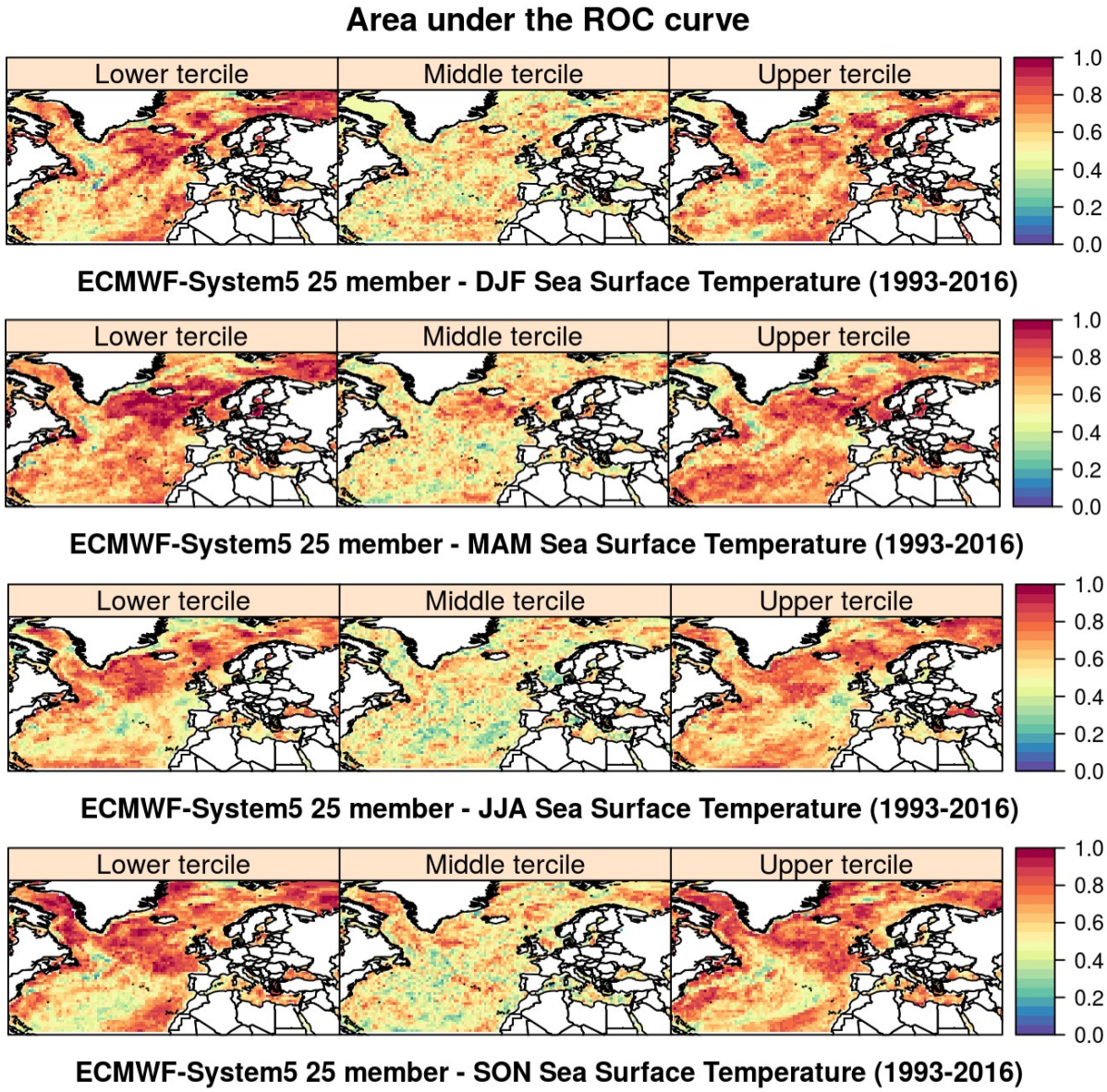

Figure 9. Area under the ROC curve of the seasonal Sea Surface Temperature for the ECMWF – System5 hindcast for the period 1993-2016.

As for Pearson correlation, a more coherence between the different hindcast has been found for the Sea Surface Temperature. Moreover, the areas with high correlation are more or less the same reflecting high ROC values. Note that for the middle tercile the obtained ROC values correspond to a random forecast in most of the domain, affecting the rest of validation measures, in particular the RPS. However, the high ROC values obtained for upper/lower terciles for all the seasons and most of the upper half of the domain reflect a great discrimination capability for these cases.

Visualization of the Seasonal Forecast: Spring 2019

From the previous Sections we have obtained some conclusions about the skill of our seasonal forecast that should be reflected in the operational seasonal forecast in order to reflect the uncertainties associated to the predicted values. In this subsection, we introduce some graphs to visualize the prediction, the uncertainty and both, in order to illustrate the need of the inclusion of the uncertainty to the final products developed by a seasonal forecast service. To this aim we have considered the forecast given by the ECMWF – System 5 for the Spring 2019 and the initialization of February, corresponding to a lead month 1.

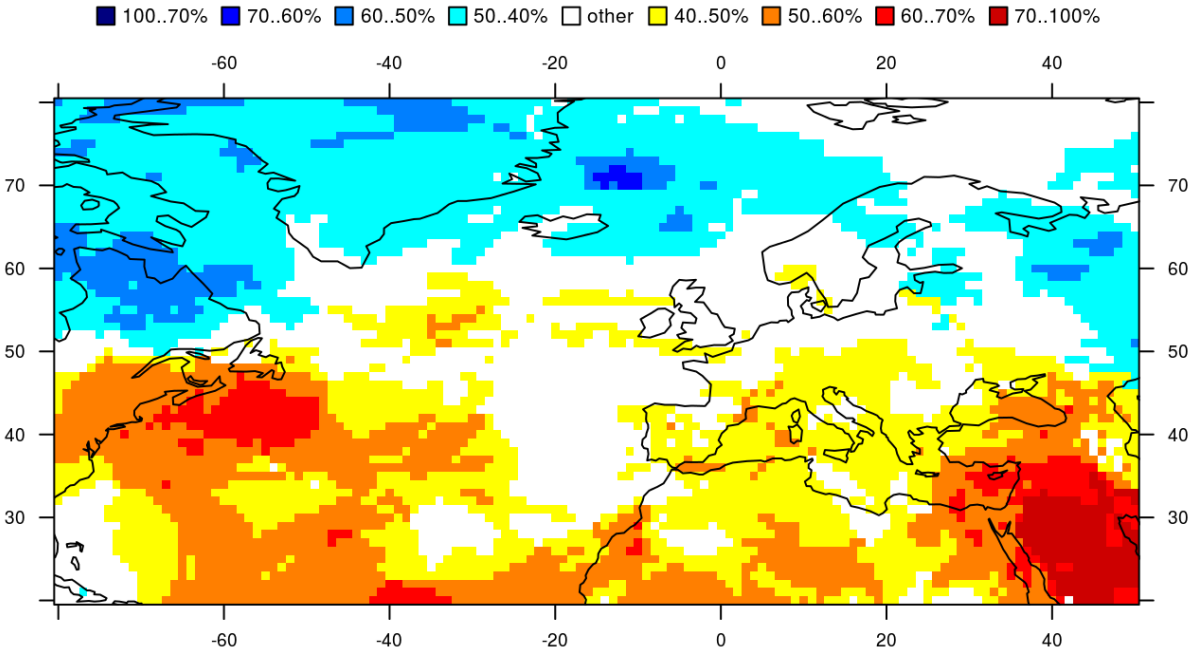

Figure 10. Tercile map of the forecast of SLP for Spring 2019 given by the ECMWF – System5.

Figure 10 shows the distribution of the members between the different terciles, reflecting a dipole between the upper and lower halves, with below- and upper- than normal Sea Level Pressure, respectively. Blank areas correspond to the normal tercile, in this case associated to regions of transition. However, the only information the user has is the color scale reflecting the percentage of members predicting each tercile. Another way to represent the same information is shown in Figure 11 in which the dot's size represents the percentage and the colour the tercile predicted. Again the information given to the user is partial and incomplete.

A simple modification of the latest plot let us to introduce the skill of our prediction. Figure 12 shows the same bubble plot but introducing a degree of transparency based on the ROC skill score (ROCSS) and, as a result, reflecting the confidence on the forecast based on the verification done using the corresponding hindcast. Note that, in this case, in almost all the domain the confidence, as given by the ROCSS, is very low. A more clear example is given in Figure 13 in which the same plot but for the Sea Surface Temperature is shown. In this case, the first two plot would give you a completely red image but, once we include the ROCSS, some regions disappear due to the lack of skill on them.

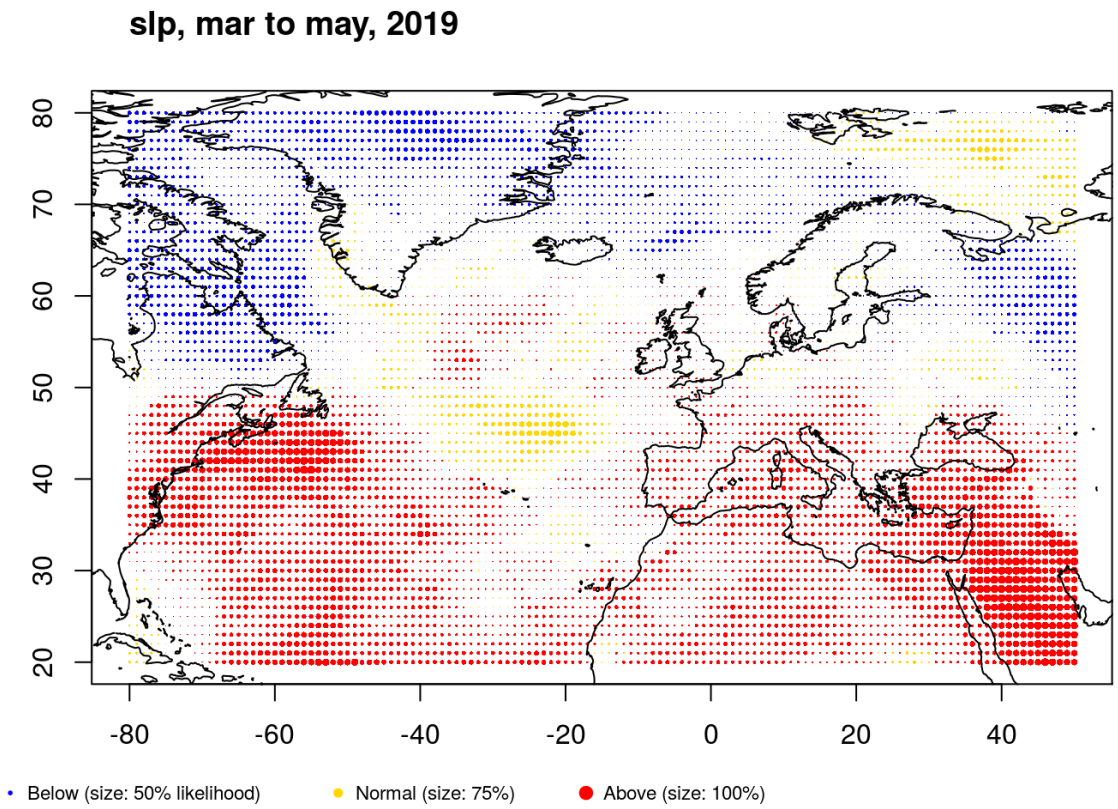

Figure 11. Bubble plot of the forecast of SLP for Spring 2019 given by the ECMWF – System5.

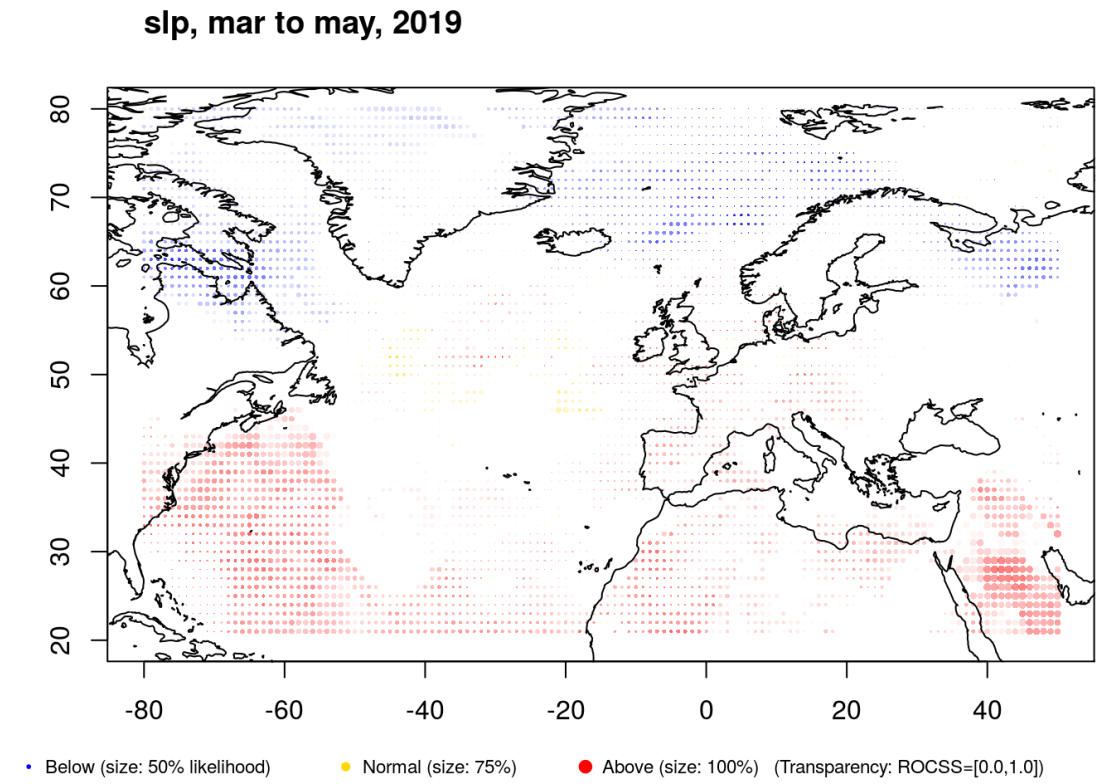

Figure 12. Bubble plot of the forecast of SLP for Spring 2019 given by the ECMWF – System5. The transparency of the dots reflects the ROCSS.

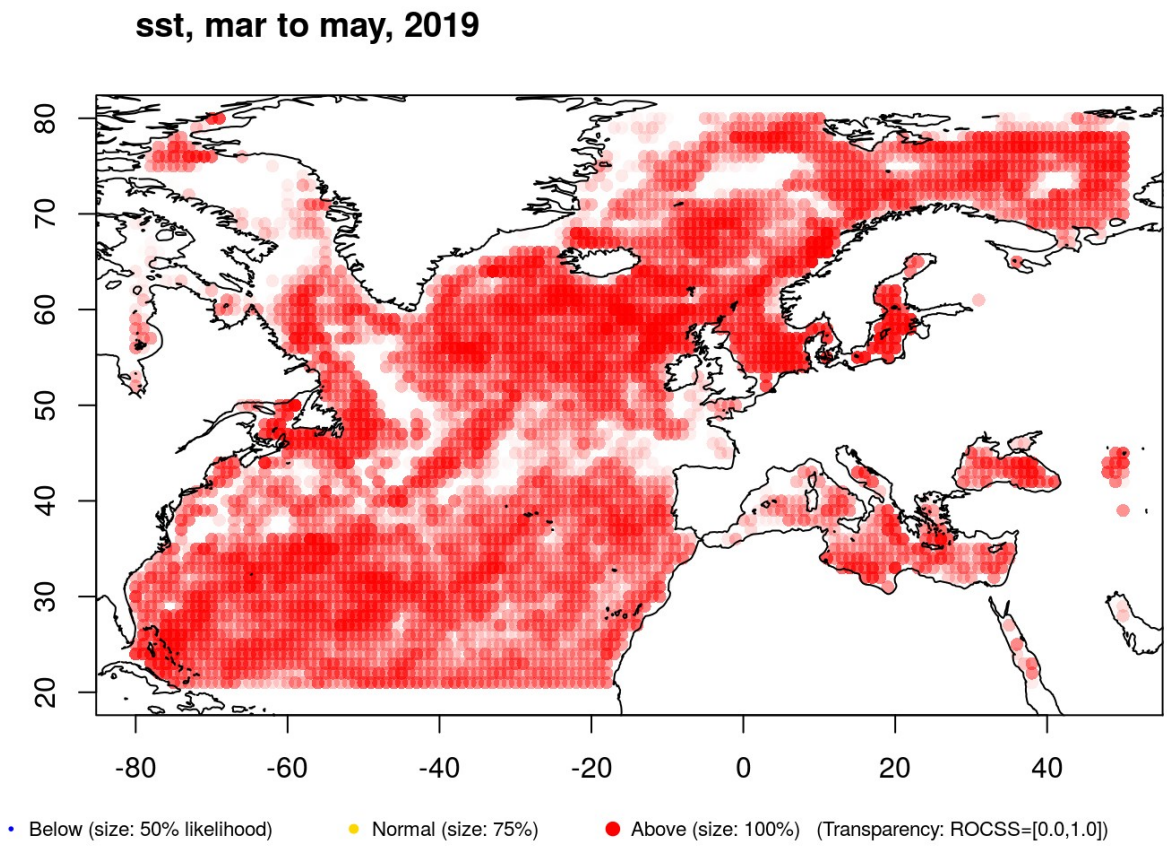

Figure 13. Bubble plot of the forecast of SST for Spring 2019 given by the ECMWF – System5. The transparency of the dots reflects the ROCSS.

Finally, in many applications the impact user only need the prediction for a particular region instead for all the country or continent. In this case, the previous maps don't make sense and other type of diagram reflecting the same information should be considered. To this aim, the tercile diagram is a simple but powerful tool to visualize both the forecast and the uncertainty associated to it. Figure 13 shows the tercile diagram obtained for the North Sea for both the SLP (upper diagram) and the SST (lower diagram). Note that in rows is shown each tercile, with the colour reflecting the probability of the tercile, and in columns each year, with the first block corresponding to the hindcast period (1993-2016) and the second to the forecast (2019). The dots represent for each year which was the tercile observed. Finally, the numbers of the right correspond to the ROCSS of each tercile reflecting the uncertainty associated to the forecast. In this case, as could be expected taking into account the previous maps, the SLP shows a very low skill, with the probability distributed between all the tercile, whilst for the SST all the members predict the upper tercile.

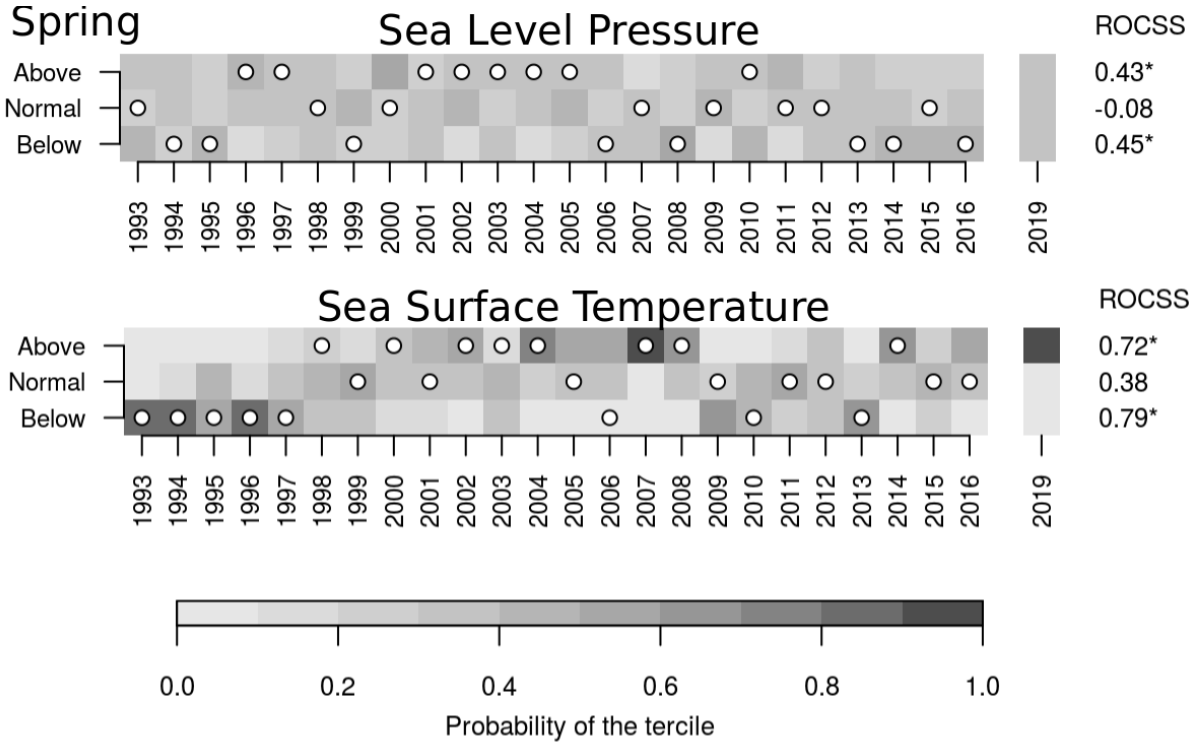

Figure 14. Tercile plot of the North Sea for each season (rows). The grayscale shows the probability of each tercile given by the ratio between the number of members falling in each tercile and the total number of members. White dots show, for each year, the observed tercile. The asterisk indicates statistically significant ROCSS values.

6. Conclusions and discussion.

This document presents a complete analysis of the different elements needed to develop a climate service of seasonal forecast of waves and surge based on the state-of-the-art of both the climate models and visualization tools including the uncertainties related with the forecast. The particularities of this kind of forecasts and the relevant role of the verification and the calibration of the data have been highlighted along the manuscript as an initial step to integrate these long term predictions as input in the different lake models.

First, the climate predictors identified in the literature as related with sea surface dynamics were evaluated in order to establish the raw signal given by the predictors. Moreover, the main atmospheric pattern driven the wind and wave conditions in Europe at this time-scale, the North Atlantic Oscillation, was considered to evaluate if the current seasonal forecast models were able to reproduce its interannual variability and its negative/positive phases.

Once the skill of the predictors has been obtained, an statistical downscaling method was applied to the hindcast to obtain a retrospective seasonal forecast of waves and apply the same verification done to the climatic variables. In this way, the added value, if exists, of the statistical downscaling method, in terms of skill, can be evaluated and the final skill can be used to properly interpret new predictions.

Finally, for the sake of the reproducibility of the results obtained all the code has been included in the Appendix A of this document.

7. References

1. Bedia, J., Iturbide, M., 2016. *transformeR: an R Package for Climate Data Manipulation and Transformation*. R Package Version 1.0.1. http://github.com/SantanderMetGroup/transformeR/wiki. (Accessed 24 August 2017).
2. Bedia, J., Iturbide, M., Baño-Medina, J., Herrera, S., Manzanas, R., Gutiérrez, J.M., 2017. *downscaleR: An R Package for Statistical Downscaling.* http://github.com/SantanderMetGroup/downscaleR/wiki. R package version 3.0.3.
3. Buontempo, C., Hewitt, C., 2018. *EUPORIAS and the development of climate services*. Clim. Serv. 9: 1-4. https://doi.org/10.1016/j.cliser.2017.06.011.
4. Christensen, J. H., Boberg, F., Christensen, O. B., Lucas-Picher, P., 2008: *On the need for bias correction of regional climate change projections of temperature and precipitation*. *Geophys. Res. Lett.*, 35, L20709, doi:10.1029/2008GL035694.
5. Cofiño, A.S., Bedia, J., Iturbide, M., Vega, M., Herrera, S., Fernández, J., Frías, M.D., Manzanas, R., Gutierrez, J.M., 2018. *The ECOMS User Data Gateway: towards seasonal forecast data provision and research reproducibility in the era of Climate Services*. Clim. Serv. 9: 33-43https://doi.org/10.1016/j.cliser.2017.07.001
6. Déqué, M., 2007. *Frequency of precipitation and temperature extremes over France in an anthropogenic scenario: Model results and statistical correction according to observed values*. Global and Planetary Change 57, 16–26. doi:10.1016/j.gloplacha.2006.11.030.
7. Doblas-Reyes, F.J., García-Serrano, J., Lienert, F., Biescas, A.P., Rodrigues, L.R.L., 2013. *Seasonal climate predictability and forecasting: status and prospects*. Wiley Interdiscip. Rev. Clim. Change 4, 245e268.
8. Frías, M.D., Herrera, S., Cofiño, A.S., Gutiérrez, J.M., 2010. *Assessing the skill of precipitation and temperature seasonal forecasts in spain: Windows of opportunity related to ENSO events*. J Clim 23(2): 209–220. doi: 10.1175/2009JCLI2824.1
9. Frías, M.D., Iturbide, M., Manzanas, R., Bedia, J., Fernández, J., Herrera, S., Cofiño, A., Gutiérrez, J.M., 2017. *VisualizeR: Visualizing and communicating uncertainty in seasonal climate prediction*. Environmental Modelling and Software 99, 101–110.
10. Frieler, K., et al., 2017. *Assessing the impacts of 1.5 °C global warming – simulation protocol of the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP2b)*. Geoscientific Model Development, 10(12), 4321–4345. doi:10.5194/gmd-10-4321-2017
11. Gutiérrez, J.M., Maraun, D., Widmann, M., Huth, R., Hertig, E., Benestad, R., Roessler, R., Wibig, T., Wilcke, R., Kotlarski, S., San-Martín, D., Herrera, S., Bedia, J., Casanueva, A., Manzanas, R., Iturbide, M., Vrac, M., 2018. *An intercomparison of a large ensemble of statistical downscaling methods over Europe: Results from the VALUE perfect predictor cross-validation experiment*. International Journal of Climatology. doi:10.1002/joc.5462.
12. Harrison, M., Troccoli, A., Anderson, D., Mason, S., Coughlan, M., Williams, J.B., 2008a. *A Way Forward for Seasonal Climate Services*. In: Troccoli, A., Harrison, M., Anderson, D., Mason, J. (eds.) Seasonal Climate: Forecasting and Managing Risk. NATO Science Series: Springer.
13. Harrison, M., Troccoli, A., Coughlan, M., Williams, J.B., 2008b. *Seasonal forecasts in decision making*. In: Troccoli, A., Harrison, M., Anderson, D., Mason, J. (eds.) Seasonal Climate: Forecasting and Managing Risk. NATO Science Series: Springer.
14. Iturbide, M., Bedia, J., Herrera, S., Baño-Medina, J., Fernández, J., Frías, M.D., Manzanas, R., San-Martín, D., Cimadevilla, E., Cofiño, A.S., Gutiérrez, J.M., 2019. *climate4R: An R-based Open Framework for Reproducible Climate Data Access and Post-processing.* Environmental Modelling and Software 111: 42-54. https://doi.org/10.1016/j.envsoft.2018.09.009
15. Jolliffe, I.T., Stephenson, D.B., 2003. *Forecast Verification: a Practitioner's Guide in Atmospheric Science*. John Wiley & Sons

16. Lange, S., 2016. *EartH2Observe, WFDEI and ERA-Interim data Merged and Bias-corrected for ISIMIP (EWEMBI)*, GFZ Data Services, doi:10.5880/pik.2016.004

17. Lemos, M.C., Kirchhoff, C.J., Ramprasad, V., 2012. *Narrowing the climate information usability gap*. Nature Climate Change 2 (11), 789–794.

18. Lorenz, E.N., 1963. *Deterministic nonperiodic flow*. J Atmos Sci, 20(2):130–141

19. Manzanas, R., Frías, M.D., Cofiño, A.S., Gutiérrez, J.M., 2014. *Validation of 40 year multimodel seasonal precipitation forecasts: the role of ENSO on the global skill.* J Geophys Res Atmos 119(4):1708–1719. doi: 10.1002/2013JD020680

20. Manzanas, R., Gutiérrez, J.M., Fernández, J., van Meijgaard, E., Calmanti, S., Magariño M., Cofiño, A., Herrera, S., 2017. *Dynamical and statistical downscaling of seasonal temperature forecasts in Europe: added value for user applications*. Clim. Serv. https://doi.org/10.1016/j.cliser.2017.06.004.

21. Manzanas, R., Gutiérrez, J.M., Ghend, J., Hemri, S., Doblas-Reyes, F.J., Torralba, V., Penabad, E., Brookshaw, A., 2018. *Bias adjustment and ensemble recalibration methods for seasonal forecasting: A comprehensive intercomparison using the C3S dataset*. Submitted to Climate Dynamics.

22. Mason, S.J., 2008. *Understanding forecast verification statistics*. Meteorol. Appl. 15 (1), 31e40.

23. Murphy, A.H., 1993. *What is a good forecast? an essay on the nature of goodness in weather forecasting*. Weather Forecast. 8 (2), 281-293

24. Ogutu, G.E.O., Franssen, W.H.P., Supit, I., Omondi, P., Hutjes, R.W., 2016. *Skill of ECMWF system-4 ensemble seasonal climate forecasts for East Africa*. Int. J. Climatol. 37, 2734e2756.

25. Palmer, T.N., Anderson, D.L.T., 1994. *The prospects for seasonal forecasting-A review paper*. Quart J Roy Meteor Soc, 120:755–793.

26. Palmer, T.N., Alessandri, A., Andersen, U., Cantelaube,P., Davey, M., Délécluse, P., Déqué, M., Díez, E., Doblas-Reyes, F.J., Feddersen, H., Graham, R., Gualdi, S., Guérémy, J., Hagedorn, R., Hoshen, M., Keenlyside, N., Latif, M., Lazar, A., Maisonnave, E., Marletto, V., Morse, A.P., Orfila, B., Rogel, P., Terres, J., Thomson, M.C., 2004: *Development of a European multimodel ensemble system for seasonal-to-interannual prediction (DEMETER). Bull. Amer. Meteor. Soc.,* 85,853–872, https://doi.org/10.1175/BAMS-85-6-853

27. Palmer, T., Doblas-Reyes, F.J., Hagerdorn, R., Weisheimer, A., 2005. *Probabilistic prediction of climate using multi-model ensembles: from basics to applications*. Philosophical Transactions of the Royal Society B: Biological Sciences, 360, 1991.

28. Primo, C., Ferro, C.A.T., Jolliffe, I.T., Stephenson, D.B., 2009. *Calibration of probabilistic forecasts of binary events*. Mon. Weather Rev. 137 (3), 1142e1149.

29. Raftery, A.E., 2016. *Use and communication of probabilistic forecasts*. Stat. Anal. Data Min. ASA Data Sci. J. 9 (6), 397e410.

30. Shi, W., Schaller, N., MacLeod, D., Palmer, T. N. and Weisheimer, A., 2015. *Impact of hindcast length on estimates of seasonal climate predictability*. *Geophys. Res. Lett.*, 42: 1554–1559. doi:10.1002/2014GL062829.

31. Siegert, S., Stephenson, D.B., Sansom, P.G., Scaife, A.A., Eade, R., Arribas, A., 2016. *A bayesian framework for verification and recalibration of ensemble forecasts: how uncertain is NAO predictability?* J. Clim. 29 (3), 995-1012.

32. Troccoli, A. 2010. *Seasonal climate forecasting*. Meteorological Applications, 17, 251-268.

33. Wang, B., et al., 2009. *Advance and prospectus of seasonal prediction: Assessment of the APCC/CliPAS 14-model ensemble retrospective seasonal prediction (1980–2004)*. Climate Dyn., 33,93–117.

34. Weedon, G. P., G. Balsamo, N. Bellouin, S. Gomes, M. J. Best, and P. Viterbo (2014), The

WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data, *Water Resour. Res.*, 50, 7505–7514, doi: 10.1002/2014WR015638.

35. Wilcke, R.A.I., Mendlik, T. & Gobiet, A. Climatic Change (2013) 120: 871. https://doi.org/10.1007/s10584-013-0845-x