ERA4CS Joint Call on Researching and Advancing Climate Services Development – Topic B
(GRANT AGREEMENT 689029)

European Climate Observations, Monitoring and Services initiative (2)

Deliverable D2.2

Assessment of the ability of state-of-the-art climate prediction systems to represent mechanisms of predictability in the Mediterranean and role of systematic errors

Version	Date	Modified by	Comments
1	30/06/2019	Chloé Prodhomme	Submitted version with each partners' contribution

Deliverable Title	Assessment of the ability of state of the art climate prediction systems to represent mechanisms of predictability in the Mediterranean and role of systematic errors	
WP number	2	
Lead Beneficiary	*Chloé Prodhomme, BSC/UB*	
Contributors	*Lauriane Batté, Constantin Ardilouze (CNRM)* *Stefano Materia, Marianna Benassi, Paolo Ruggieri (CMCC)* *Chloé Prodhomme, Javier García Serrano (BSC/UB)* *Esteban Rodríguez Guisado (AEMET)*	
Deliverable Due Date	*30/06/2019*	
Nature of the Deliverable	R	*R – Report*
		SP – Software Package
		D – Dataset
		O - Other

Table of contents

EXECUTIVE SUMMARY

The present deliverable summarizes the ability of the state-of-the-art forecast systems to represent mechanisms of predictability in Europe and more specifically in the Mediterranean region. We have analyzed the C3S seasonal forecasts systems (CMCC, DWD, ECMWF and MeteoFrance). These systems show very limited skill for precipitation, however for temperature and more specifically heatwave indices the systems reach significant and potentially useful skill. El Niño-Southern Oscillation (ENSO) is an important source of skill for atmospheric circulation in the Northern Hemisphere, the relationship between ENSO and Sea Level Pressure (SLP) is relatively well reproduced by the models. However, in the case of the Siberian Snow Cover (SCC), while a relationship between SLP and SCC is clearly seen in the observation, this relationship is not represented by the models. Soil moisture is also an important driver of temperature over Europe. C3S models represent relatively well the weather regimes over the Mediterranean region and the relationship between North Atlantic Oscillation and temperature is fairly realistic. Finally, models show a good ability in representing the sudden stratospheric warming frequencies and this seems to be independent of the zonal wind biases in the stratosphere.

1. Skill analysis over the Mediterranean region for empirical and dynamical prediction system

1.1 Temperature and precipitation (AEMET)

One of the main issues seasonal forecast faces over the Mediterranean is the low skill over the region. However, levels of skill change among seasons, models and areas. In this section, the distribution of skill is assessed, to provide insight on the matter. As most seasonal forecasting systems publish their forecasts as tercile probabilities, and the most common variables are surface temperature and precipitation, several scores have been calculated over the Mediterranean domain for the different models. Here, for sake of briefness, only Ranked Probability Skill Score (RPSS) data is plotted, but all data will be uploaded to MEDSCOPE's website, to provide information on skill when producing forecast for a given season and area.
Verification scores are computed and visualized following Sánchez et al. (2018). For every predictand (temperature and precipitation), season, score and verifying sub region, results from a selection of seasonal forecasting systems based on dynamical models and the empirical system are put together in a table for easy comparison. Figure 1 shows results for RPSS over a selection of sub regions.
Red colors indicate that the forecast has lower skill than climatology, while orange indicates similar or slightly higher skill. Generally, skill seems to be higher for the eastern part of the domain, and higher for temperature than for precipitation. Multi-model shows more regular results, reaching better skill values for a higher number of seasons than any particular model on its own, although for one particular season and area, scores can be lower than those of the best model. Generally, there is large variability among seasons and areas.

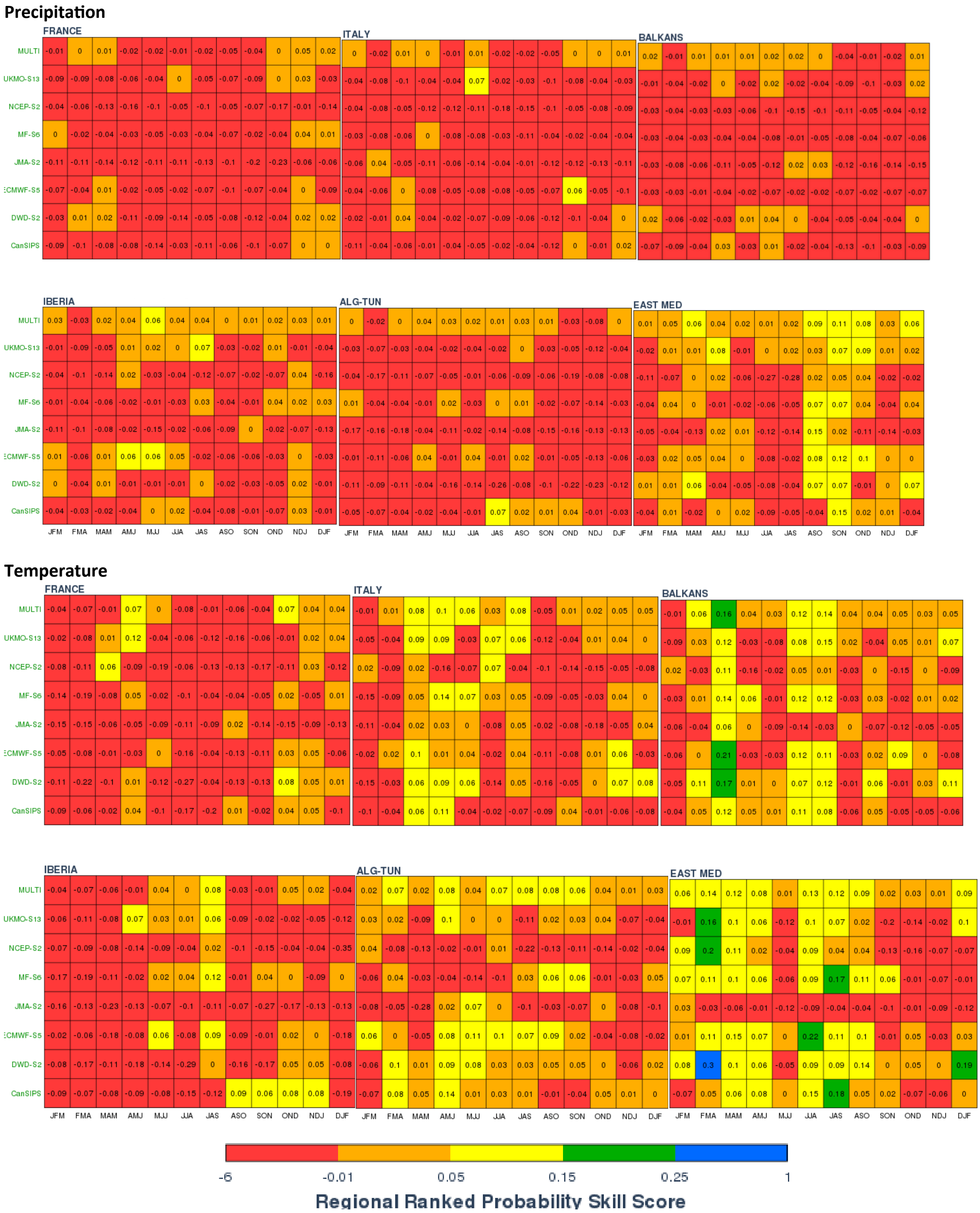
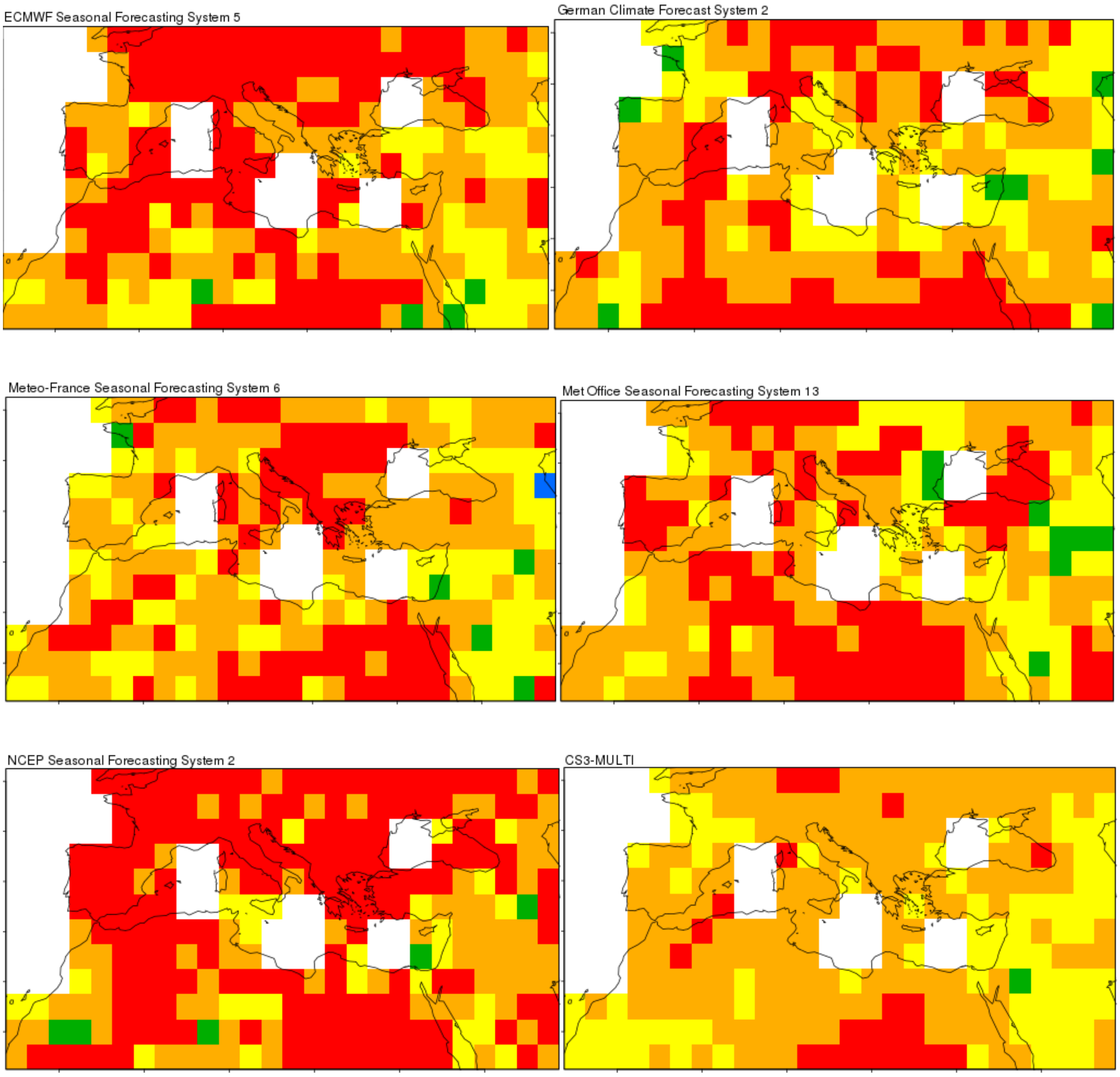


Figure 1. *Ranked Probability Skill Score (RPSS) (Terciles) for a) Precipitation forecasts and b) Temperature forecasts. Skill is evaluated over a selection of sub regions on the Mediterranean domain (France (42.5-45N, 10W-12.5E), Italy (32.5-47.5N, 2.5-20E), Balkans (32.5-47.5N, 12.5-30E), Iberia (32.5-47.5N,10W-7.5E), Algeria-Tunisia (17.5-37.5N,5W-12.5E) and Eastern Mediterranean (27.5-37.5N, 20-40E)). Every table contains information of an individual sub region, representing each cell the average RPSS value over the selected area for a particular model (y- axis) and 1-month lead time forecasted season (x-axis). The uppermost row in all tables corresponds to a multimodel composed from CS3 system models with complete hindcast available (ECMWF-S5, MF-S6, DWD-S2, UKMO-S13). GPCC precipitation data and ERA-Interim T2m are used as verifying observations. Verification period is 1994-2015.*

a)Precipitation



b)Temperature

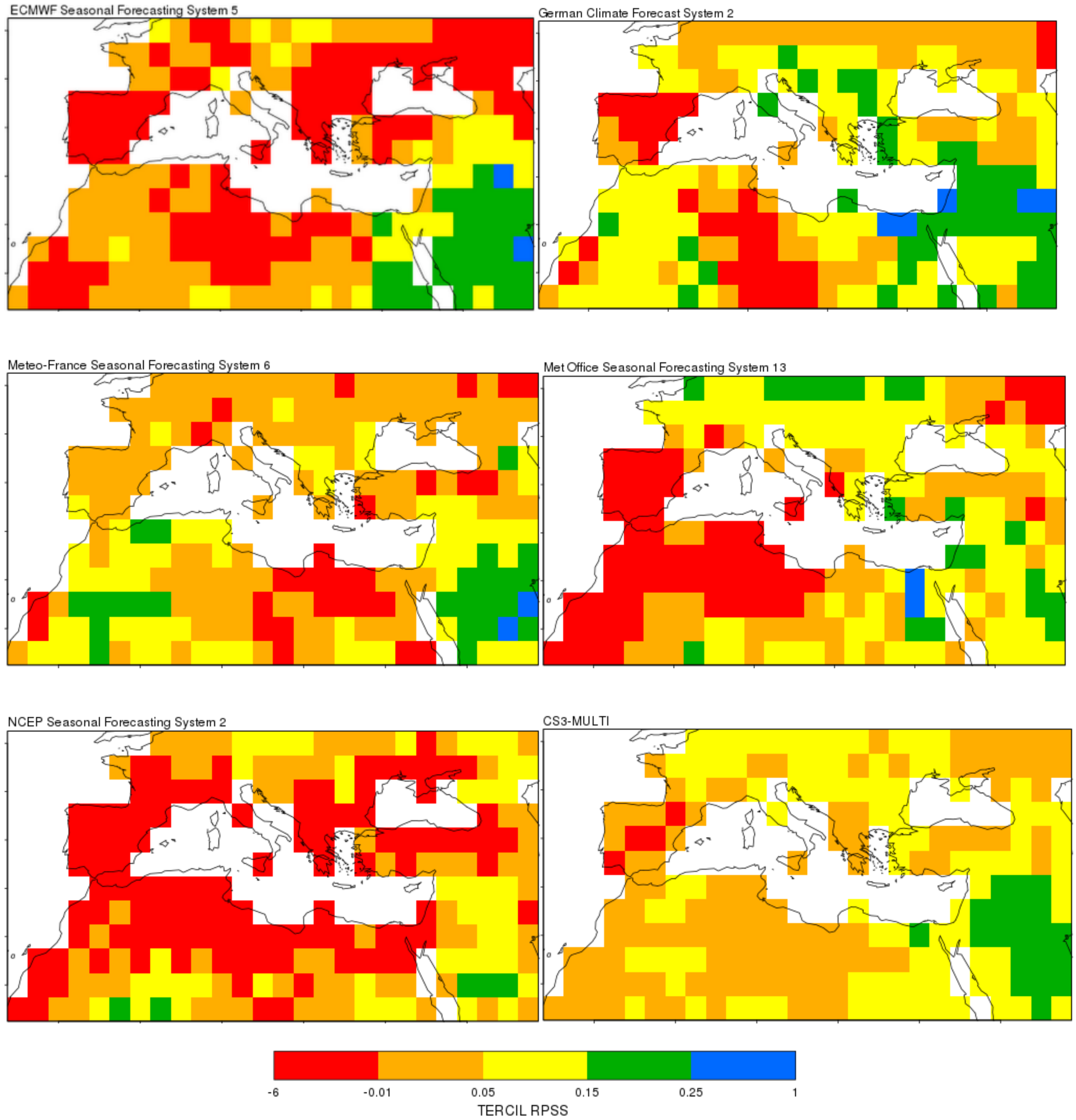


Figure 2. *Ranked Probability Skill Score (RPSS) (Terciles) for a) Precipitation forecasts and b) Temperature forecasts. Every map is calculated using hindcast data from one Seasonal Forecast Model, for 1-month lead time DJF forecast. The last map (bottom right) corresponds to a multi-model composed from CS3 system models with complete hindcast available (ECMWF-S5, MF-S6, DWD-S2, UKMO-S13). GPCC precipitation data and ERA-Interim T2m are used as verifying observations. Verification period is 1994-2015.*

To further explore the spatial variability of skill, Figure 2 shows spatial maps of RPSS are plotted for DJF (1-month lead forecasts). Spatial differences are even more striking than Figure 1 when considering grid-point skill compared to regional average. Huge differences can be seen among models: for example, for temperature in the Balkans, DWD seems to perform better, when, over northern part of the domain, MetOffice model reaches higher values. MF-S6 performs better over western Africa and Iberia. Again, the multi-model seems more consistent, but doesn't reach the highest value over most regions. The eastern part of the domain is the region where the skill is generally the highest.

Although some general remarks can be made looking at maps, the overall conclusion is the high variability of score values among regions, models and seasons. Although skill is low, for a particular season and area, there can be some models that reach relatively high and potentially useful skill. This suggests the importance of consulting skill maps when producing forecasts or deciding the source of the forecast for testing a prototype product, and to choose those which display higher skill for the region of interest.

1.2 Heatwave predictability (BSC, CMCC)

Due to global warming, heatwave frequency and magnitude is expected to increase severely in the coming decades (Russo et al. 2015). The predictability at the seasonal time scale of extreme temperature events appears to be crucial for climate services, adaptation and risk management (Thomson et al. 2006, Ceglar et al. 2017). We measure heatwave using the Heat Wave Magnitude Index (HWMI) defined by Russo et al. (2015), the used index presents several advantages compared to its counterparts, first it is a robust estimator that allows to measure only “real heatwaves” avoiding considering extreme temperatures of very short duration (which cannot be considered as heatwaves). In addition, heatwaves are normalized for each grid point, allowing to consider heatwaves relative to local climate. Another advantage of this method is that it gives only one number for a summer season, including both length and magnitude of the strongest heatwave independently of the moment it happens, this makes it easily comparable with seasonal forecasts which could hardly predict the exact duration and timing of the heatwave. We analyse the HWMI computed over June, July and August (JJA) for three C3S systems: EMWF S5, MeteoFrance S6 and DWD S2 over the period 1993-2009 for the May start dates. The verification is done considering the ERA-Interim reanalysis as the reference dataset, interpolation is done toward the model grid, using the Inverse Weighted Distance method, which is a method appropriate to interpolate data presenting strong spatial gradients as it is the case for the HWMI.

Figure 3 shows the anomaly correlation computed for two quantities: the ensemble mean (Figure 3a-c) and the ensemble maximum (Fig 3d-f). Generally, this figure shows that the models have some skills in predicting heatwaves, especially over Eastern Europe. As for temperature (see section 1.1), the areas of skill are different in every system, it is interesting to note that considering ensemble maximum instead of ensemble mean, seems to give a better prediction of the HWMI, in terms of correlation.

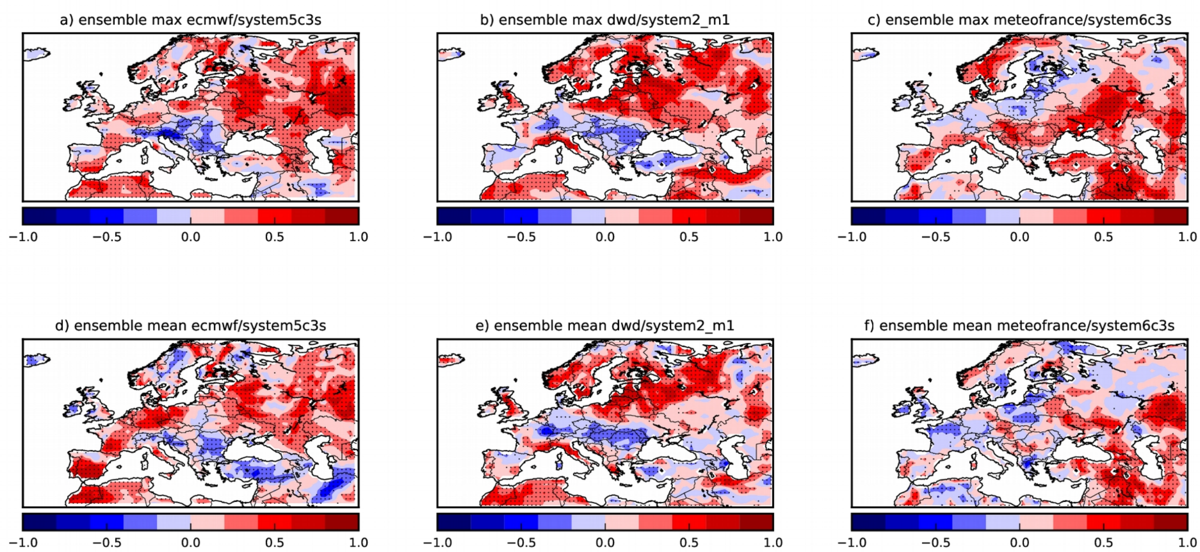


Figure 3: *Correlation of JJA HWMI computed over the period 1993-2009 with ERA-Interim in the C3S systems, dots shows correlation significant at the 95% confidence level.*

In order to evaluate the ability of the model to forecast large heatwave, we defined a large heatwave for a given grid point where the HWMI exceeds the percentile 80. We compute the contingency table (false alarm, missed event, hit event and correct rejection), and then compute the two following quantities:

- Hit rate = hit event/(hit event + missed event)
- False alarm rate = False alarm/(False Alarm + correct rejection)

Figure 4 shows for each grid point the difference between hit rate and false alarm rate, positive values indicate that the forecast systems might give a useful prediction of large heatwaves. As for figure 3, we compare the ensemble mean and maximum. We see that C3S systems generally display positive values, showing that they are relatively skilful in predicting large heatwaves with this metric. Ensemble mean seems to perform slightly better than ensemble maximum.

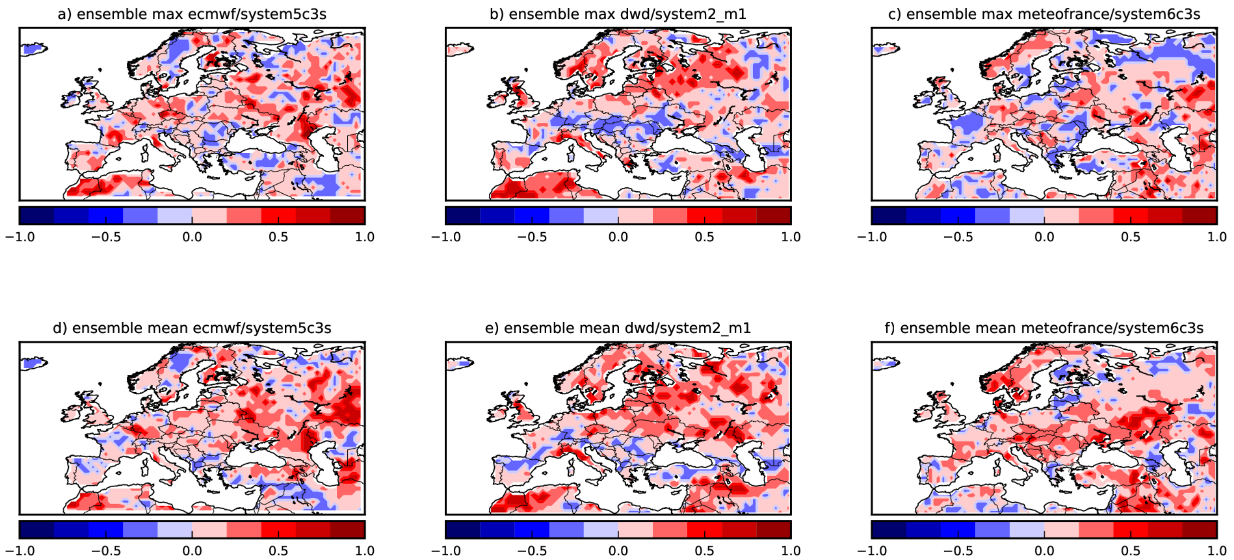


Figure 4: *Difference between Hit rate and False alarm rate for JJA HWMI computed over the period 1993-2009 with ERA-Interim in the C3S systems.*

	ecmwf/system5c3s		dwd/system2_m1		meteofrance/system6c3s	
Iberian peninsula	0.45	0.74*	0.47	0.47	0.12	0.35
France	0.32	0.46	0.64*	0.50*	0.42	0.47
Alps	0.62*	0.38	0.38	0.43	-0.25	0.26
Central Europe	0.80*	0.58*	0.43	0.64*	-0.06	0.35
Western Europe	0.54*	0.54*	0.72*	0.59*	0.38*	0.58*
Mediterranean	0.68*	0.61*	0.42	0.61*	-0.01	0.33
Northern Europe	0.63*	0.64*	0.63*	0.64*	0.57*	0.63*
Eastern Europe	0.57*	0.58*	0.33	0.53*	0.01	0.35
Europe	0.64*	0.72*	0.77*	0.68*	0.51*	0.68*

Table 1: *Correlation for the HWMI averaged in different European Region. Value in red show the correlation computed over ensemble maximum and values in grey over ensemble mean. * marks correlation significant at the 95% confidence level.*

In table 1, we show correlation for HWMI averaged in several European regions. Results confirm again that models present significant skill for the HWMI, especially when we consider large scale regions.

To conclude, C3S systems might be skilful in predicting heatwaves, however results should be considered carefully since the analysed period is relatively short.

2. Potential sources of predictability in C3S seasonal forecast systems. (CMCC, MF)

We investigate potential sources of predictability for the Euro-Mediterranean winter season (DJF) in the current generation of C3S seasonal forecast systems. We focus our attention on two remote processes: El Nino Southern Oscillation (ENSO) and on the Siberian Snow Cover (SSC) and one local process: the soil moisture.

2.1 El Niño-Southern Oscillation

El Niño-Southern Oscillation (ENSO) is known to be one of the major drivers of global climate variability at the seasonal to interannual time scale, and it is considered as the strongest predictor for the Euro-Atlantic atmospheric circulation (*e.g.* Scaife et al., 2014, Toniazzo and Scaife, 2008). Here we focus on the hindcast period (1993/2016), looking at the reforecasts issued in November and considering the average winter circulation (Mean Sea Level Pressure averaged over DJF). Each forecast system is represented by its ensemble mean. As observational benchmark, we include ERA-Interim reanalysis data (Dee et al., 2011). We characterize ENSO using the canonical NINO3.4 index. The NINO3.4 has been defined as the normalized area average SST anomaly compared to the reference climatology 1981/2010. The November values have been considered, under the hypothesis that the initialization of the forecast is including this observed signal into the forecast system.
An assessment of the skill of the C3S systems may be done by looking at the pattern correlation coefficient in the hindcast period. In figure 5a, the pattern correlation coefficient for the whole Northern Hemisphere is reported, while in figure 5b the focus is on the Euro-Mediterranean region. At the hemispheric level, the pattern correlation coefficient is generally higher than the regional one, and the agreement across the different systems is good. The maxima are found when a strong El Nino event takes place (i.e. in 1997 and in 2015). In general, the skill increases with the intensity of the ENSO events. On the other hand, at the regional level, the agreement across systems is much less pronounced, and even in correspondence of strong ENSO event the system skill is generally lower than the hemispheric case. The hemispheric skill in representing ENSO teleconnection is dominated by the PNA pattern, with a weaker signal over the Euro-Mediterranean sector.

Figure 6 shows the regression maps and the correlation patterns of the reanalysis and of the C3S system ensemble averages with the NINO3.4. The canonical PNA pattern is indeed well captured by all the C3S systems. The circulation response to ENSO is active also in the North Atlantic, with the canonical signal characterized by a dipole with increased sea level pressure over the Arctic, and a decrease of sea level pressure over mid latitude. Figure 6 shows less agreement across different systems on the location and amplitude of these anomalies. If we focus on the Mediterranean area, the correlation values suggest the ENSO teleconnection to be stronger in the models than in the reanalysis. This is consistent with the fact that the reanalysis pattern comes from a single realization of the system, and hence is much more affected by internal variability than the patterns from the model ensemble averages.
From this analysis, some skill of ENSO as a potential predictor for winter conditions over the Euro-Atlantic sector has been found. The time period considered here does not allow to sample a significant number of events with high/low snow cover anomalies, and hence to make a robust assessment.

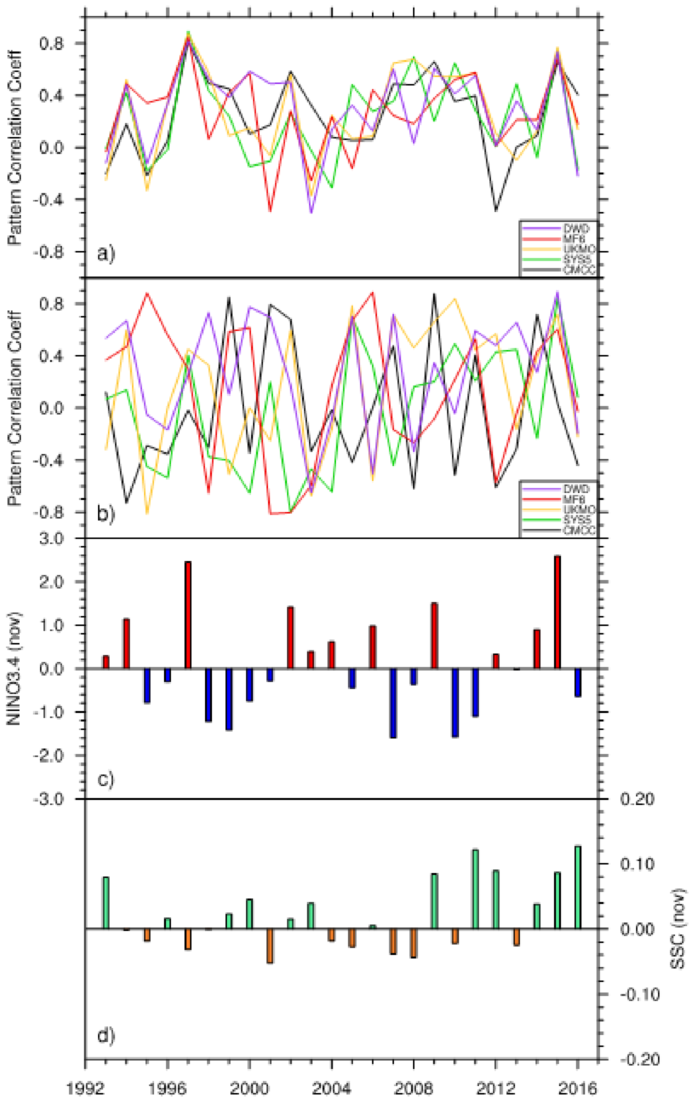


Figure 5: *Panel (a): Northern Hemisphere DJF mean sea level pressure pattern correlation coefficient between ERA-Interim and C3S system ensemble mean. Panel (b) as panel (a) for the Euro-Mediterranean region (defined as a box from 20 to 75 degrees N in latitude, and from -17 to 55 degrees E in longitude). Panel (c): NINO3.4 index computed with HadISST reanalysis (Rainer et al., 2003). Panel (d): SSC index computed with Rutgers GSL snow cover extent data (Robinson et al., 2012)*

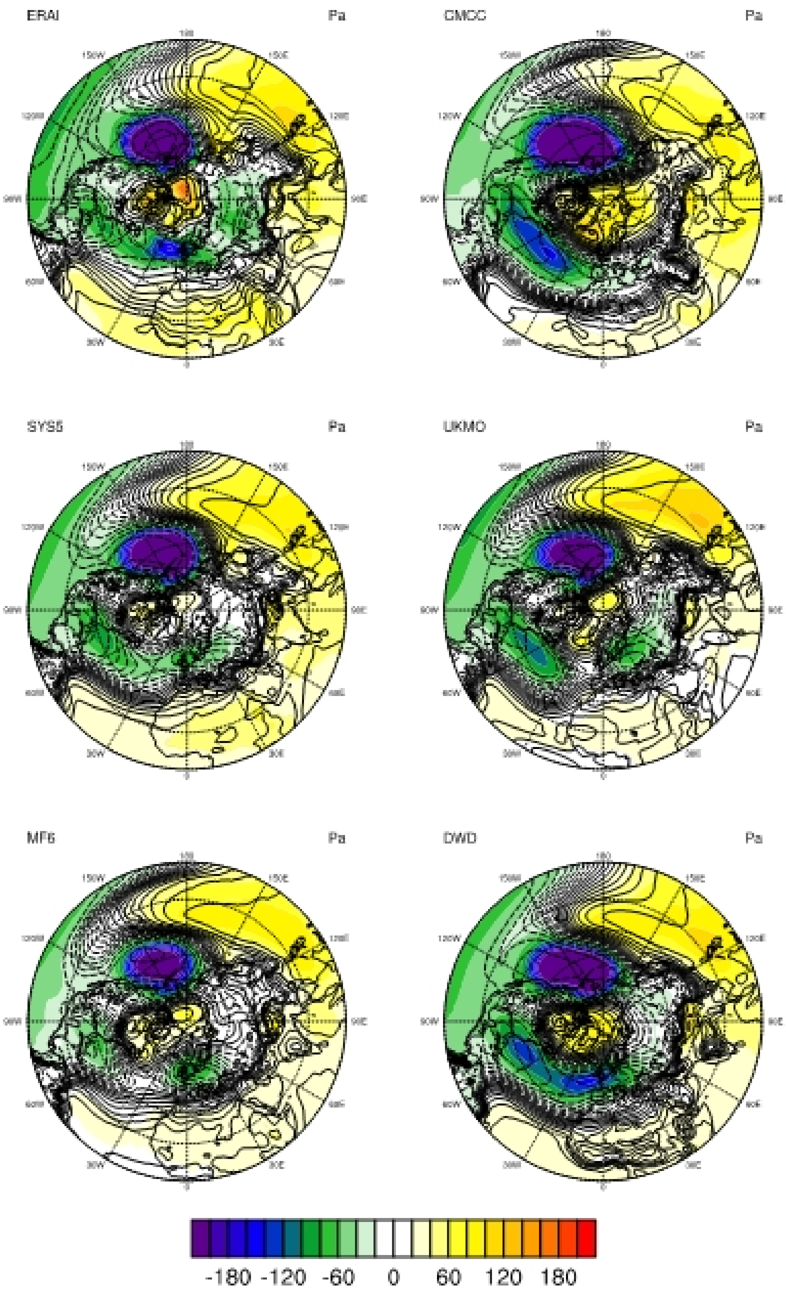


Figure 6: *Regression pattern (in shadings) and correlation map (in contour) between mean sea level pressure and NINO3.4 index. The correlation contours range from 1.0 to -1.0.*

2.2 Siberian Snow Cover

The Siberian Snow Cover (SSC) during fall has been suggested as a potential source of predictability for the Arctic Oscillation in the following winter (*e.g.* Cohen et al. 2007). The SSC index has been defined as the fraction of anomalous snow cover, compared to the 1981/2010 climatology, averaged over the Siberian region (as defined in Douville et al., 2017, Figure 5c). The November values have been considered, under the hypothesis that the initialization of the forecast is including this observed signal into the forecast system.
In Figure 7, the regression maps and the correlation patterns of the reanalysis and of the C3S system ensemble averages with the SCC are reported. In reanalysis data, results show some evidence of the paradigm linking snow cover during fall to the Arctic Oscillation, while in the models there is no consensus on the response, which in general appears much weaker. While the Aleutian anomaly is found in most of the model regression patterns, only UK MetOffice and DWD systems seem to partly capture the positive signal over the Arctic.
From this analysis, the impact of SSC is less clear than the one of ENSO. Even if an increasing SCC trend can be noticed in the last years of the time series, such behaviour is not reflected in the forecast skill, either at the hemispheric or at the regional level.
From this analysis, the actual role of Siberian Snow Cover remains questionable. The time period considered here does not allow to sample a significant number of events with high/low snow cover anomalies, and hence to make a robust assessment.
Moreover, it should be noticed that this discussion is based on deterministic metrics, comparing reanalysis and ensemble averages. A possible future development is to fully exploit the ensemble, performing a probabilistic skill analysis, and evaluating the sensitivity of our results to the ensemble size.

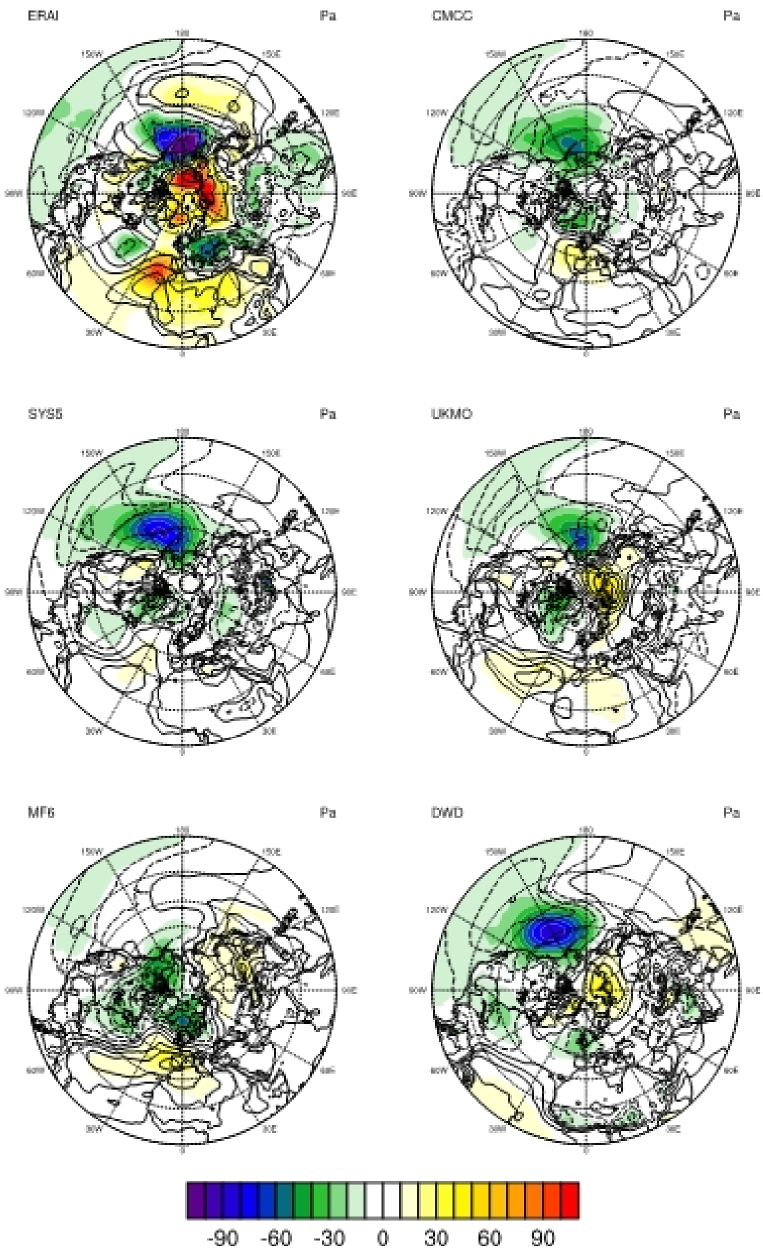


Figure 7: *Regression pattern (in shadings) and correlation map (in contour) between mean sea level pressure and SCC index. The correlation contours range from 1.0 to -1.0.*

2.3 Soil moisture

We have carried out an idealized numerical study to assess the potential impact of soil moisture on summer temperature predictability. The reference experiment (G-REF) consists in an ensemble re-forecast of the JJA season over the 1993-2012 period, based on the CNRM-CM climate model. All the components of the model (atmosphere, land, ocean, sea-ice) are initialized on May 1st with conditions derived from reanalyses. In the perturbed experiment (G-SOIL), the climate components are initialized likewise, but here, soil moisture is prescribed daily towards pseudo-observed values, derived from the ERA-Interim/Land surface reanalysis. The figure 8 shows the mean JJA 2-meter maximum temperature correlation between simulated and observed (CRUTS4) data, as well as the correlation difference between the two simulations. We find a considerable increase of correlations in G-SOIL with respect to G-REF over most of North hemisphere mid and high latitudes. Comparable results have been found for precipitation (not shown). This set-up is idealized and does not take into account the predictability of soil moisture in the first place. However, it shows that soil moisture is potentially a major source of climate predictability at seasonal scale, at least in those regions exhibiting enough soil moisture persistence. More details on the setup and results at regional scale over Europe can be found in Ardilouze et al. (2019a) and in the PhD manuscrit of Ardilouze (2019b)

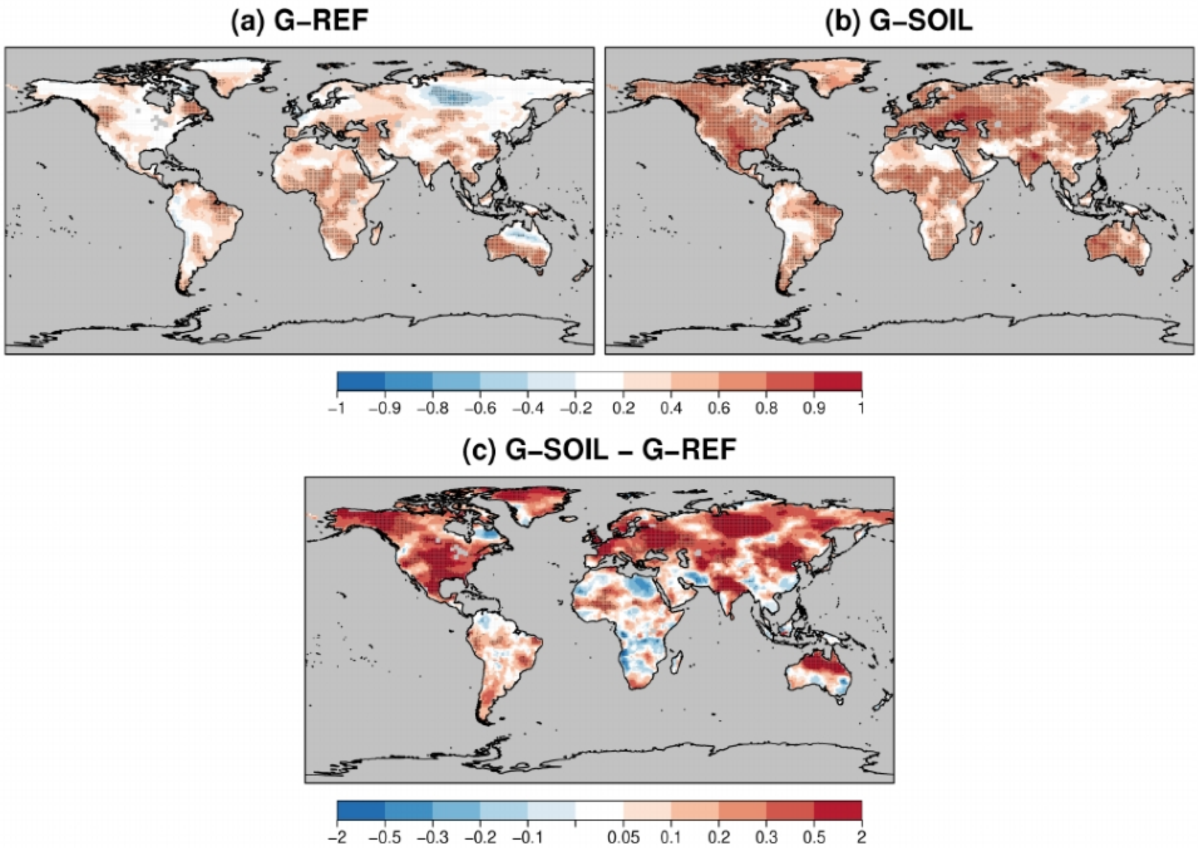


Figure 8: *Mean JJA Maximum temperature correlation over the 1993-2012 period against CRU TS4.01 for (a) G-REF and (b) G-SOIL, and (c) correlation difference G-SOIL minus G-REF. Stippling highlights grid points with significant values at a 95% confidence level.*

3. Seasonal prediction of atmospheric circulation in the C3S

3.1 Weather regimes

We explore potential sources of predictability in a framework of weather regimes (WRs) analysis. The concept of WR, which has been extensively used in the Euro-Atlantic region (e.g. Dawson et al. 2012, Ferranti et al. 2015), is applied here to the Mediterranean. Mediterranean Wrs (MWRs) are defined as large-scale quasi-stationary atmospheric patterns that can last from a few days to two or three weeks. They are computed by decomposing daily 500 hPa geopotential height field anomalies with the Empirical Orthogonal Functions (EOFs), restricting the analysis to four principal components (explaining 80% of the variance) to which a k-means clustering algorithm is applied. In agreement with Rohas et al. (2013), four MWRs were identified for the winter (DJF) in the reanalysis (ERA-Interim) dataset in the period 1993-2016. The same methodology has been used to identify Mediterranean Weather Regimes (MWRs) in the ensemble hindcasts of four operational seasonal predictions systems. A metric based on anomaly pattern correlation and the ratio between model and reference patterns' standard deviation has been applied to evaluate the quality of Mediterranean weather regime simulation in such systems.

Here below, a summary of the main feature of the four Mediterranean Winter Regimes shown in figure 9 (left panels) is provided.

- MWR1, with a frequency of 30.1%, shows a dipole NW-positive / SE-negative geopotential height anomalies that corresponds to general cold conditions over most of the Mediterranean region, while for precipitation it corresponds to a dipole NW dry-SE wet. This regime seems to be associated with the Atlantic Ridge.

- MWR2, with a frequency of 26.6%, shows a MWR1 inverted dipole (NW-negative / SE-positive), whose corresponding seasonal temperature pattern displays widespread dry conditions over the Mediterranean, with a western half in wet conditions and the southeast of the domain in dry conditions.

- MWR3, with a frequency of 22% shows a prominent positive geopotential height anomaly over the whole domain and corresponds to warmer than usual conditions over the NW and, in terms of precipitation, to a generally dry Mediterranean region. This regime is associated to the positive phase of the NAO.

- MWR4, with a frequency of 21.4%, refers to a cyclonic pattern. In opposition to MWR3, this regime depicts a generally negative anomaly over most of the domain, with a positive anomaly limited to the southeast. In terms of effects of this regime on temperature and precipitation we find the opposite of MWR3, with a cold NW and hot SE, and a generalized wet Mediterranean. This regime is associated to the negative phase of the NAO.

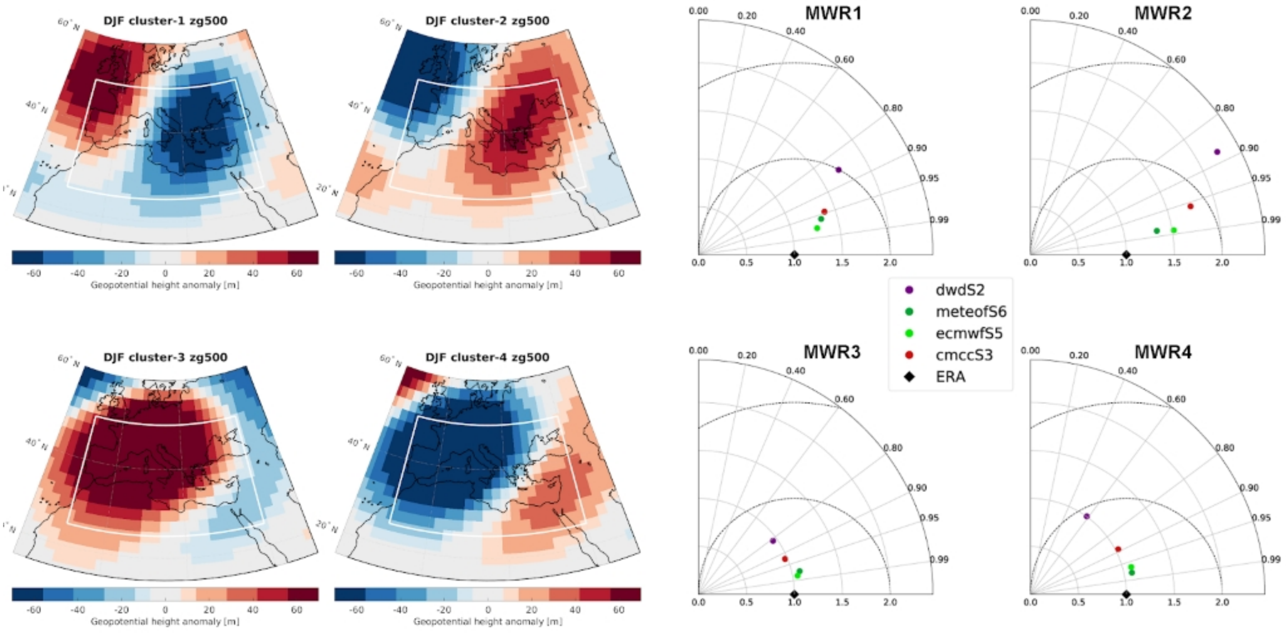


Figure 9: *Mediterranean DJF Weather Regimes obtained using ERA-Interim (left panels); Taylor diagrams of the regimes comparing ERA-Interim (black diamond) to the prediction systems (colored dots).*

As mentioned above the same regimes have been computed in the 1993-2016 November starting dates of hindcast ensembles for the following operational Seasonal Prediction Systems: ECMWF S5, CMCC S3, CNRM S6, DWD S2. These systems have a varying number of ensemble members (25 to 40), however for the sake of consistency, 25 ensemble members have been considered for each of them. Overall, the Prediction Systems considered exhibit a good consistency in reproducing MWR patterns. The Taylor diagrams shown in figure 9 above provide information on how well the prediction systems capture the spatial pattern of ERA-Interim (black dot) for the different regimes: in terms of standard deviation MWR3 and MWR4 show good alignment with ERA-Interim and a good correlation too. MWR1 and MWR2, although with generally good correlation values, show higher values of standard deviation compared to ERA-Interim. Regime frequencies were generally underestimated for the MWR1 and overestimated for MWR3 compared to ERA-Interim, but the overall ranking of the occurrence of the regimes was fairly well captured by the prediction systems (ranking unanimously MWR4 as least frequent and the majority ranking MWR1 as most frequent).

A similar analysis was carried out on the JJA summer season using 5 clusters (with 74% of variance explained by the EOFs) revealing, as seen for DJF, good agreement in the spatial patterns of the regimes reproduced by the prediction systems. The Five Mediterranean Summer Regimes (MSRs) (not shown) are overall consistent with those found by Zampieri et al. (2017).

Overall we found that the Operational Seasonal Forecast Systems analyzed are able to simulate the observed Mediterranean Weather Regimes with good accuracy, especially in the winter season.

3.2 North Atlantic Oscillation

The North Atlantic Oscillation (NAO, Hurrell, 2003) is the main mode of variability over the North Atlantic Ocean and neighbouring regions at monthly to decadal time scales. Previous studies have suggested that state-of-the-art seasonal forecasting systems can now reasonably predict the NAO index (e.g. Athanasiadis et al. 2017), although with some levels of uncertainty. However, associated patterns in terms of temperature and precipitation over Europe are not always well represented.

We evaluate here how seasonal forecast systems from the C3S database represent the impacts of a low NAO (lower tercile) or a high NAO (upper tercile) on temperature and precipitation anomalies, notwithstanding skill levels for the NAO index. The NAO index is computed as the leading EOF in 500 hPa geopotential height over a North Atlantic - Europe region. The composites of temperature and precipitation anomalies (using the 1993-2016 reference period of the C3S seasonal re-forecasts) for ERA-Interim (GPCP for precipitation) are shown in the left-hand side of the figures 10 and 11. The corresponding composites for ECMWF SEAS5 (center) and Météo-France System 6 (right), computed member by member, irrespective of the ensemble mean NAO index, are plotted for direct comparison.

Figure 10 shows the composites for near-surface temperature. Both seasonal forecast systems show strikingly similar patterns for lower (b,c) and upper (e,f) NAO index terciles. The pattern amplitudes are weaker than in ERA-Interim, and models exhibit a higher linearity in the NAO response than observed. Note however that in the case of seasonal re-forecasts, since each ensemble member is used to construct these composites, the sample size is 25 times higher than in reanalysis data, which may contribute to smoothing the signal. The robustness of results in ERA-Interim for Northern Canada is questionable, since the same sign of positive anomalies in both upper and lower terciles of the NAO index is found.

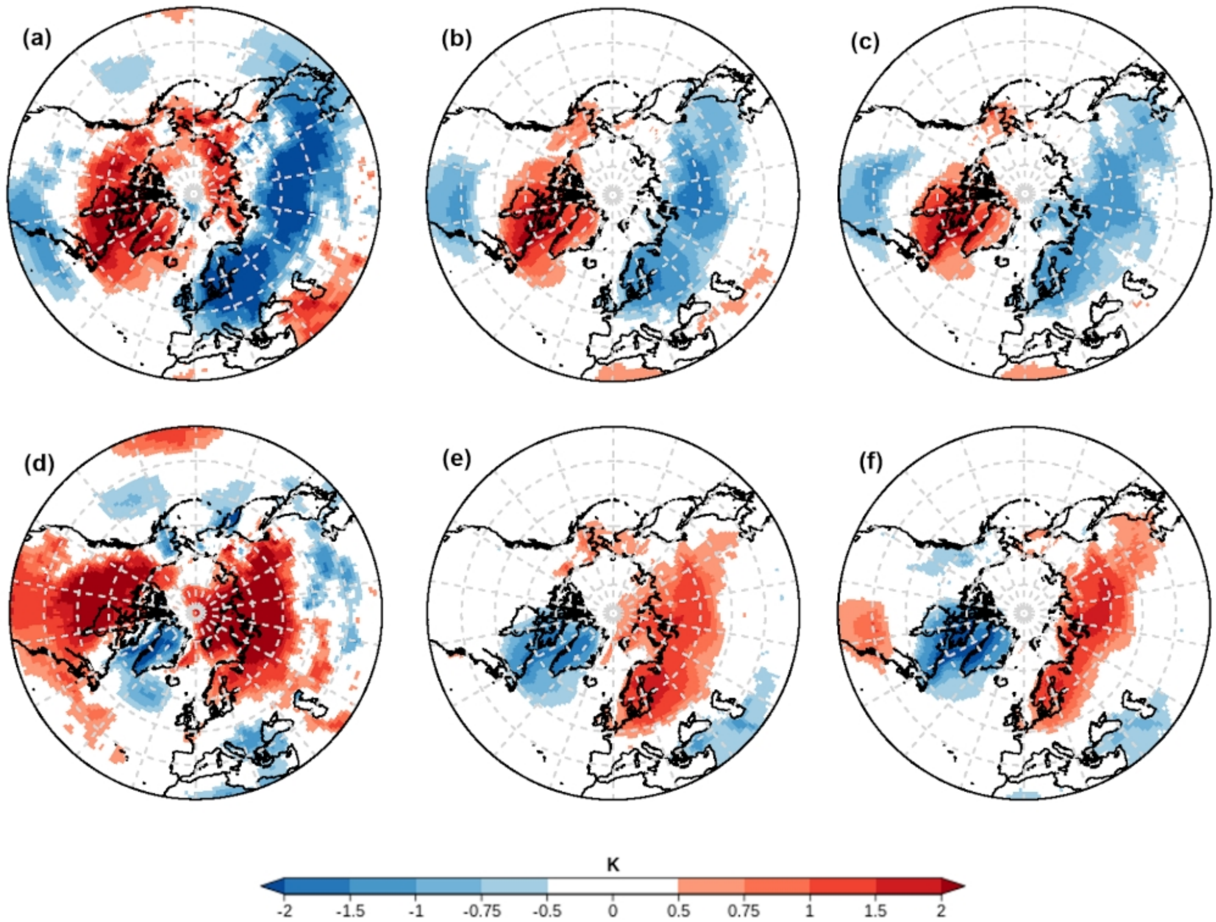


Figure 10: *Composites of DJF near-surface air temperature (in K) for the lower tercile (a-c) and upper tercile (d-f) of the NAO index over 1993-2016 in ERA-Interim reanalysis (a,d), ECMWF SEAS5 re-forecasts (b,e) and Météo-France System 6 (c,f) initialized in November.*

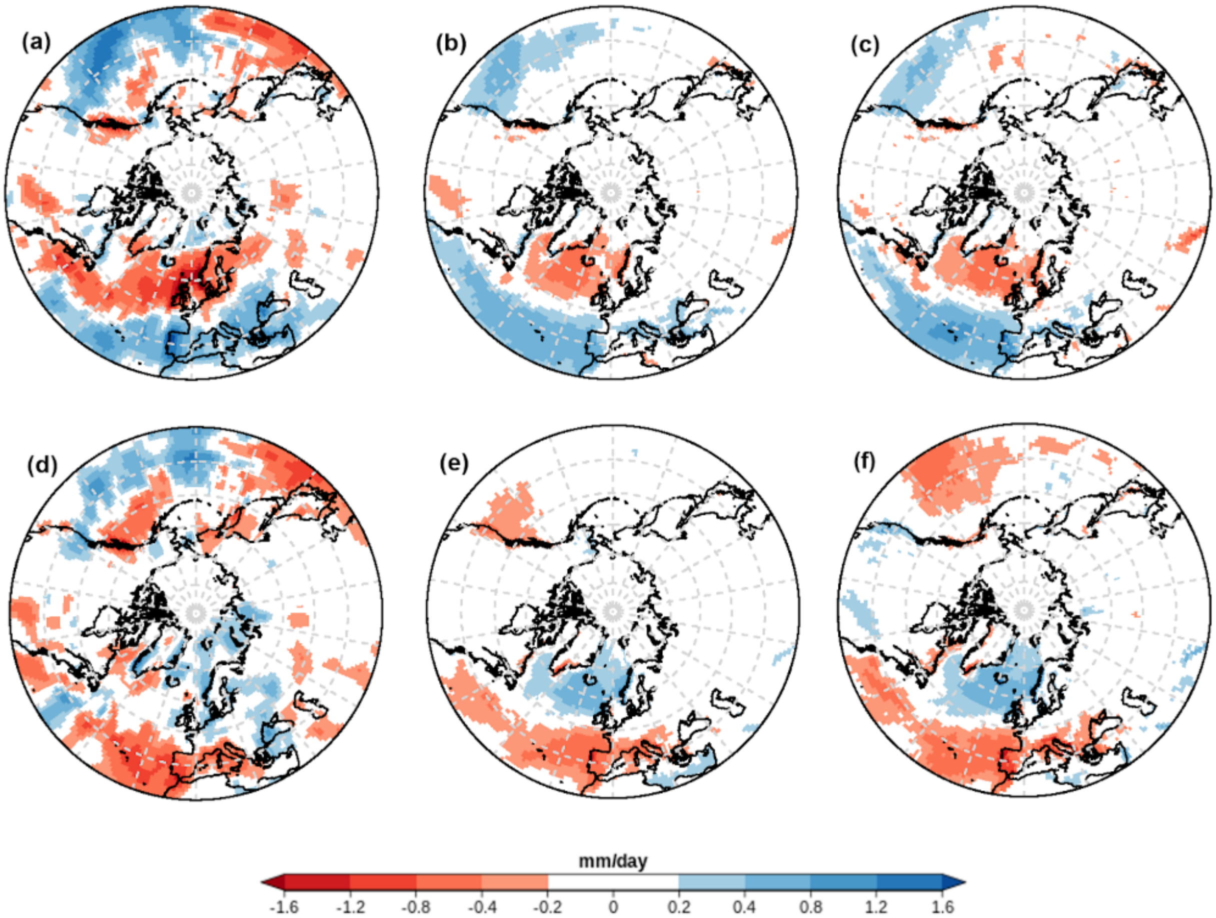


Figure 11: *Same as figure 10 but for total precipitation (mm/day) using GPCP.*

4. Seasonal forecast skill in the winter stratosphere

The variability of the winter stratosphere is dominated by the Quasi-Biennial Oscillation (QBO) in the tropics and by the modulation of the intensity of the polar vortex (PV) in the extra-tropics. Changes in the state of the PV are associated with a signal of the North Atlantic Oscillation (NAO, Scaife et al. 2016) and sudden stratospheric warmings (SSWs) are harbingers of extreme surface events and the negative NAO (Palmeiro et al. 2015). We present an analysis of the seasonal predictability of the stratospheric circulation in winter (DJF). Data used are obtained from the Copernicus Climate Change service multi-model seasonal forecast. We use forecasts initialised on November 1st for 4 modelling centres (CMCC, DWD, ECMWF, Meteo-France), and forecasts initialised on October 25th, November 1st, November 9th for the UKMO model. The analysis is based on the zonal mean zonal wind at 60 °N. The anomaly correlation coefficient is computed for the zonally averaged zonal wind at 10 hPa as function of latitude. A persistence forecast is computed assigning to the DJF forecast anomaly the observed anomaly of November. The spread is computed by the standard deviation of all members. The mean spread is the square root of the squared spread averaged across all start dates. The interannual variability is the standard deviation of the ensemble means across all start dates.

From November to March, SSW are detected as zonal-mean zonal wind reversals at 10 hPa and in a range of latitudes from 55 to 70 o N so each latitude is evaluated individually (Palmeiro et al. 2015). Final warmings are discarded as in Charlton and Polvani (2007).

In figure 12a, we can see the model climatology compared to ERA-Interim. Figure 12b shows that the variability of the tropical winds is captured by the interannual variability of the models. Inside the stratospheric polar vortex, interannual variability is low compared to the spread. This has implications for potential predictability if the mean spread is used as a proxy of the total variability. Figure 12c shows the skill for the zonal mean zonal wind compared with a persistence forecast. In the tropics the skill is very high, but the forecast is essentially initialised and persisting and this feature is arguably explained by the flipping direction of the wind due to the QBO. In the extratropics the drop of skill in the midlatitudes is noticeable and inside the polar vortex the skill always beat the (poorly performing) persistence forecast, but ranges from near-zero values to 0.6. Finally, figure 12d shows how the skill at 60 N scales with the size of the ensemble. It is unlikely that the wide range of skill in polar latitudes is explained by the size of the ensemble only.

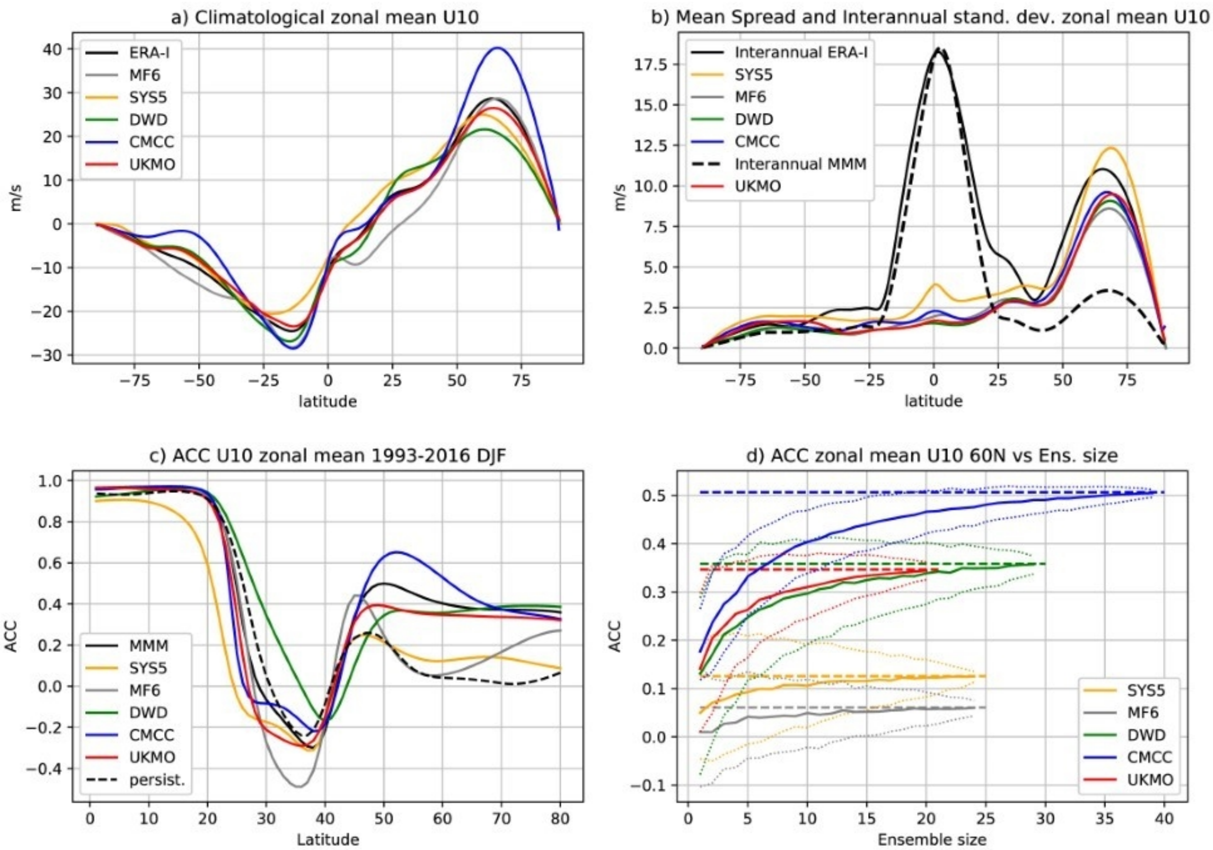


Figure 12: *a) DJF climatology of zonal mean U10 for ERA-Interim and C3S models. b) Mean spread and interannual standard deviation of zonal mean U10 c) Anomaly correlation coefficient for zonal wind at 10 hPa as a function of latitude and d) at 69 N as a function of the ensemble size. Dashed lines in panel d) indicate the interquartile range of distributions obtained subsampling without repetition.*

Since SSW occurrence is a potential tool for seasonal forecasting (Sigmond et al. 2013) it is desirable that seasonal forecast models represent them realistically. Most of the models considered herein show SSW frequencies similar to Era-Interim, and the seasonal distribution of the events during winter is not far from the observed (Figure 13). Interestingly, CMCC, which shows the best skill on the stratosphere (Figure 12d), has the lowest SSW frequency, and a clear shift of the SSW distribution to late winter. This is consistent with the strongest wind speeds shown for CMCC (Figure 12a).

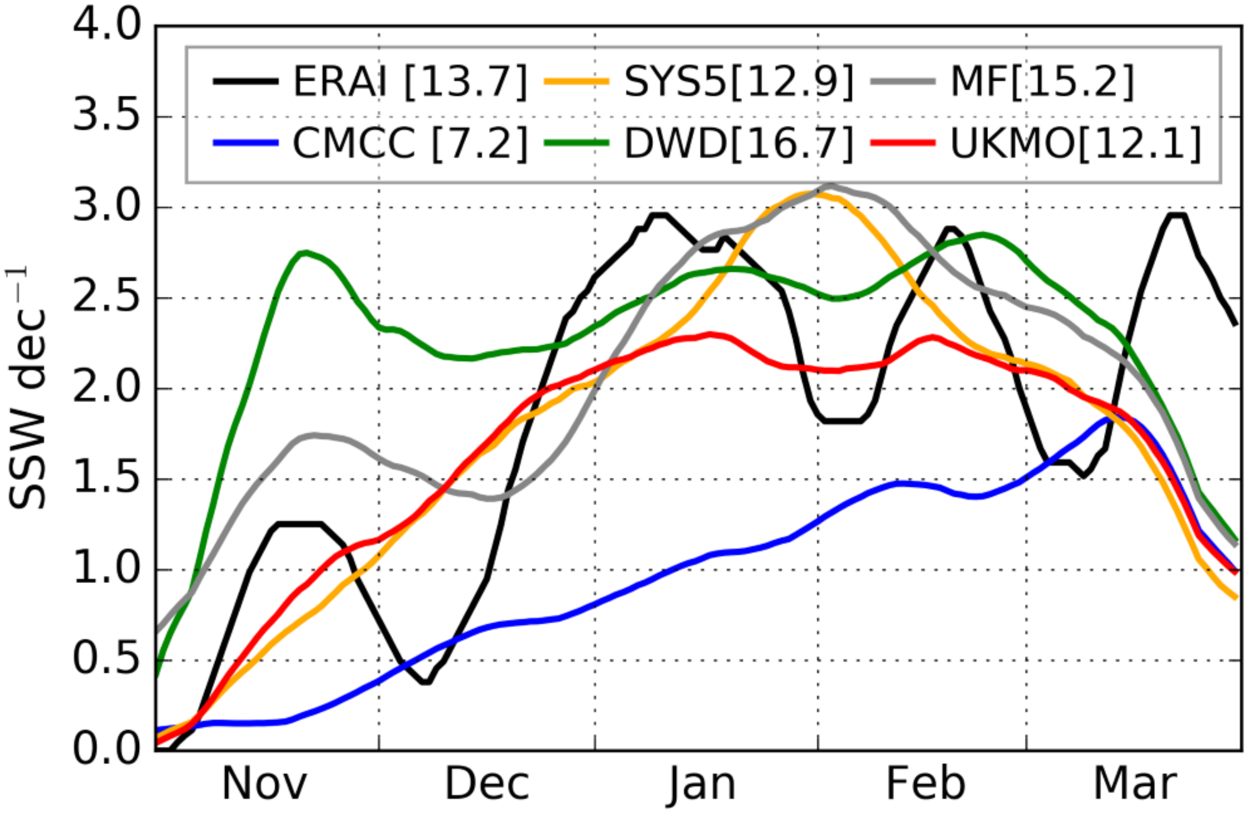


Figure 13: *November to March intra-seasonal distribution of SSWs per decade in a [-10, 10]-day window for Copernicus Seasonal Forecasts (colors) and ERA-Interim (black) in the 1993-2016 time period. Time series are smoothed with a 10-day running mean. Total decadal frequencies are shown in brackets.*

References:

Ardilouze, C., Batté, L., Déqué, M., van Meijgaard, E., & van den Hurk, B. (2019a). Investigating the impact of soil moisture on European summer climate in ensemble numerical experiments. *Climate Dynamics*, *52*(7-8), 4011-4026.

Ardilouze, C. (2019b) PhD Manuscript (in French) Impact de l'humidité du sol sur la prévisibilité du climat estival aux moyennes latitudes

Ceglar, A., A. Toreti, C. Prodhomme, M. Zampieri, M. Turco, and F. J. Doblas-Reyes, 2018: Land-surface initialisation improves seasonal climate prediction skill for maize yield forecast. Sci. Rep., 8, https://doi.org/10.1038/s41598-018-19586-6.

Charlton, A. J. and Polvani, L.M. (2007) A new look at stratospheric sudden warmings. part I: Climatology and modeling benchmarks. Journal of Climate, 20(3):449–469, 2007.

Cohen, J. and C. Fletcher, 2007: Improved Skill of Northern Hemisphere Winter Surface Temperature Predictions Based on Land–Atmosphere Fall Anomalies. *J. Climate,* **20**, 4118–4132, https://doi.org/10.1175/JCLI4241.1

Dawson, A., T. N. Palmer, and S. Corti 2012, *"Simulating regime structures in weather and climate prediction models*", Geophys. Res. Lett., 39, L21805, doi:10.1029/2012GL053284

Dee, D. P. et al. (2011), The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q.J.R. Meteorol. Soc.*, 137: 553–597. doi: 10.1002/qj.828

Douville, H. et al. (2017): Snow-(N)AO relationship revisited over the whole twentieth century. *Geophys. Res. Lett.*, 44, 569–577, doi:10.1002/2016GL071584.

Ferranti, L., Corti, S. and Janousek, M. 2015, "*Flow-dependent verification of the ECMWF ensemble over the Euro-Atlantic sector*." Q.J.R. Meteorol. Soc., 141: 916–924. doi: 10.1002/qj.2411

Hurrell, J. W., Y. Kushnir, G. Ottersen, and M. Visbeck, 2003: An overview of the North Atlantic Oscillation. American Geophysical Union (AGU), 1–35 http://www.agu.org/books/gm/v134/134GM01/134GM01.shtml

Sánchez García, E., Voces Aboy, J., Rodríguez Camino, E. Verification of six operational seasonal forecast systems over Europe and Northern Africa. Technical note, AEMET, 2018. Available at http://medcof.aemet.es/index.php/models-skill-over-mediterranean

Palmeiro F. M., Barriopedro, D., García-Herrera, R., and Calvo, N (2015) Comparing sudden stratospheric warming definitions in reanalysis data. Journal of Climate, 28(17):6823–6840, 2015.

Rojas, M., L. Z. Li, M. Kanakidou, N. Hatzianastassiou, G. Seze, and H. Le Treut (2013), Winter weather regimes over the Mediterranean region: their role for the regional climate and projected changes in the twenty-first century, *Clim. Dyn.*, *41*(3–4), 551–571, doi:10.1007/s00382-013-1823-8.

A. A. Scaife, A. Yu. Karpechko, M. P. Baldwin, A. Brookshaw, A. H. Butler, R. Eade, M. Gordon, C. MacLachlan, N. Martin, N. Dunstone, and D. Smith. Seasonal winter forecasts and the stratosphere. Atmospheric Science Letters, 17(1):51–56, 2016.

M Sigmond, JF Scinocca, VV Kharin, and TG Shepherd. Enhanced seasonal following stratospheric sudden warmings. Nature Geoscience, 6(2):98, 2013.

Rayner, N.A. et al. (2003): Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res.: Atmospheres*, *108*(D14), https://doi.org/10.1029/2002JD002670

Robinson, D. A. et al. (2012): NOAA Climate Data Record (CDR) of Northern Hemisphere (NH) Snow Cover Extent (SCE), Version 1.NOAA National Centers for Environmental Information. doi: 10.7289/V5N014G9.

Russo, S., and Coauthors, 2014: Magnitude of extreme heat waves in present climate and their projection in a warming world. J. Geophys. Res. Atmos., 119, 12,500-12,512, https://doi.org/10.1002/2014JD022098.
Scaife, A. A., et al. (2014), Skillful long-range prediction of European and North American winters, *Geophys. Res. Lett.*, 41, 2514– 2519, doi:10.1002/2014GL059637.
.
Toniazzo, T., and Scaife, A. A. (2006), The influence of ENSO on winter North Atlantic climate, *Geophys. Res. Lett.*, 33, L24704, doi:10.1029/2006GL027881.

Zampieri, M., A. Toreti, A. Schindler, E. Scoccimarro, and S. Gualdi (2017), Atlantic multi-decadal oscillation influence on weather regimes over Europe and the Mediterranean in spring and summer, *Glob. Planet. Change*, *151*, 92–100, doi:10.1016/j.gloplacha.2016.08.014.