Seasonal prediction over Europe

F. J. Doblas-Reyes [(1)(2)]

(1) Institut Català de Recerca i Estudis Avançats, Barcelona, Spain
(2) Institut Català de Ciències del Clima (IC3), Doctor Trueta 203, 08005 Barcelona, Spain
f.doblas-reyes@ic3.cat

1. Introduction

Seasonal forecasting is a field with enormous potential influence in different socioeconomic sectors, such as agriculture (Challinor et al., 2005), health (Thompson et al., 2006) and energy (García-Morales and Dubus, 2007).

The feasibility of seasonal prediction largely rests on the existence of slow, and predictable, variations in the Earth's boundary conditions of soil moisture, snow cover, sea-ice and ocean surface temperature (Shukla and Kinter, 2006), and somehow also the stratosphere (Marshall and Scaife, 2009; Marshall and Scaife, 2010), and how the atmosphere interacts and is affected by these boundary conditions. For example, a warm sea surface temperature (SST) anomaly, in say the tropical Pacific Ocean will lead to increased heat flux from the ocean to the atmosphere. This increased flux, if sufficiently large in magnitude and spatial scale will alter the atmospheric boundary layer and ultimately change the structure of the rainfall and the release of latent heat in the free troposphere. The change in tropospheric latent heat release, in turn, will perturb the circulation leading to climatic anomalies in remote regions of the globe. Of course the atmospheric response itself interacts and affects the SST. Hence, the seasonal predictability or the quality of the seasonal prediction is a coupled problem.

A major portion of predictability at monthly to seasonal time scales is attributed to anomalies in tropical sea surface temperatures (SST), in particular those related to El Niño-Southern Oscillation (ENSO, e.g., Chang et al. 2006) events (Kirtman and Pirani, 2009). This makes seasonal forecast quality reach a maximum over most tropical regions. Although the predictability is lower at extratropical latitudes, some positive skill has been found in those regions (e.g., in North America) associated with ENSO teleconnections (e.g., Quan et al., 2006), and also with other sources of seasonal predictability, such as the persistence of the North Pacific decadal oscillation (Gershunov and Cayan, 2003; Lienert et al., 2011) or the moisture content (Douville, 2004). The predictability of temperature and precipitation in Europe differs significantly from that over North America, due to different characteristics of the variability and remote influences of the local climate (Rodwell and Doblas-Reyes, 2006).

In most of the extratropics, the signals predicted by general circulation models are weak and do not add valuable information over a climatological forecast. However, some preliminary studies have revealed some signals of skill for particular European regions, periods, and variables (Doblas-Reyes et

al., 2000; Feddersen and Andersen, 2005; Frías et al., 2005; Shongwe et al., 2007). Moreover, some traces of seasonal skill induced by ENSO events have been reported in Spain (Sordo et al., 2008).

In previous decades, a considerable effort has been made to improve the understanding of the physical phenomena responsible for the observed seasonal variability and to transfer the advances to the operational numerical forecasting systems, e.g. Saha et al. (2006). This transfer requires an appropriate assessment of the forecast skill achieved in different regions for different variables to evaluate future model improvements in terms of forecast quality. The aim of the present study is to illustrate the forecast quality over Europe of a state-of-the-art seasonal forecasting system for both temperature and precipitation. A brief summary of the experiment follows in Section 2. Results from a simple persistence model that will be used as a benchmark are described in Section 3. The most relevant characteristics in terms of model drift and forecast quality results are given in Sections 4 and 5. A summary and a brief discussion is offered in Section 5.

2. Data

2.1. Experimental set-up

The ECMWF seasonal forecast System 3 (S3 henceforth, Stockdale et al., 2011) is the system used in this study. It has been operational since 2007 until recently and consists of the ECMWF atmospheric model IFS coupled to the HOPE ocean model. The cycle 31R1 of the atmospheric IFS model is used in S3. This cycle was operational in medium-range weather forecasting in 2006. The only change to the IFS model in seasonal forecasting compared to the medium-range forecasts concerns its reduced resolution, where S3 runs the IFS with 62 vertical levels, extending to ~5 hPa, and a T_L159 spectral horizontal resolution, with a corresponding grid mesh resolution of 1.125° or about 125km. Important features of the atmosphere model include two-time level semi-lagrangian numerics with a finite element discretization in the vertical, the RRTM (Rapid Radiation Transfer Model) scheme for longwave and a six spectral interval scheme for shortwave radiation, mass-flux convection, prognostic clouds, a boundary layer scheme with an eddy-diffusivity mass-flux framework, the TESSEL tiled surface scheme with six land tiles and a four-level representation of soil, turbulent orographic form drag, and sub-grid scale orographic drag. A comprehensive model of the ocean surface waves and their interaction with the atmosphere is also included. The model numerical code allows for a one-hour time-step. The ocean-model resolution remains effectively 1^o x 1^o in mid-latitudes, with a 0.3^o meridional resolution at the equator. There are 29 levels in the vertical with the highest resolution (10 m) near the surface.

S3 does not have a physically-based model of sea ice. Instead, for both forecasts and hindcasts, the ocean model specifies persistence for ten days of the initially specified fractional ice cover, taken from the ECMWF operational NWP analyses or ERA-40 reanalyses, as appropriate. After a forecast time of ten days, the specified fractional ice cover is a linear combination of the initial sea ice cover and the climatological ice cover valid at the specification date. Beyond a forecast time of 30 days, the specified sea ice cover is simply defined by linear interpolation of climatological monthly mean values.

The forecasts are initialised in the atmosphere and soil with the ERA-40 re-analysis before 2002 and with the ECMWF operational analysis afterwards. Every simulation has a start date of the first of the month and is seven months long. The ocean is initialized with the ORA-S3 reanalysis (Balmaseda et al., 2008). They consist of an ensemble of eleven members generated by sampling uncertainties in the initial conditions of both the atmosphere and ocean. The initial atmospheric conditions are perturbed with atmospheric singular vectors, calculated as per the ECMWF medium-range ensemble forecast system for this atmospheric model version (a combination of initial singular vectors, evolved singular vectors, and targeted singular vectors in the tropics). Stochastic physics (stochastic perturbation of physical tendencies with a six-hour decorrelation time scale) is active throughout the forecast period, again as in the medium-range ensemble forecast system. The ocean initial conditions in S3 are provided not from a single ocean analysis but from a 5-member ensemble of ocean analyses, created by adding perturbations to the wind forcing used in the analysis. The ocean initial conditions are further perturbed by adding sea surface temperature perturbations to the five member ensemble of ocean analyses.

All the simulations analysed in this study are retrospective forecasts (also known as re-forecasts or hindcasts). The focus has been on the one-month lead seasonal averages for March-to-May (MAM), June-to-August (JJA), September-to-November (SON) and December-to-January (DJF). These re-forecasts are initialized on the 1st of February, May, August and November of each year, respectively.

There are several sources of variation of climate which are not modelled interactively in S3, but which can be specified. One of them is the time-variation of greenhouse gases in the atmosphere model. Changes in greenhouse gas concentration have a substantial impact on seasonal forecasts, even at a time range of only a few months, when comparing forecasts and hindcasts made decades apart (Doblas-Reyes et al., 2006). An approximate time-history of CO2, methane and CFCs is specified, based on observed values up to 2000 and values derived from the IPCC A1B scenario beyond this. In the re-forecasts the year-to year variability in the solar constant is specified, although a fixed value is used after 2000. It was decided to switch off the option for time variation of volcanic aerosol in the hindcasts, since real-time volcanic aerosol analyses are not available for the operational forecasts.

2.2. Reference data

Various measures of forecast quality have been used to assess the differences between the experiments. All forecast quality measures have used ERA-Interim as the atmospheric reference dataset, except for near-surface air temperature and precipitation for which the Global Historical Climatology Network (GHCN; Peterson and Vose, 1997) and the Global Precipitation Climatology Centre (GPCC; Schneider et al., 2008) have been used, respectively.

Every forecast quality measure has been computed taking into account the systematic error of the forecast systems. The reader should be aware that this linear method assumes that there is no relationship between the model drift and the anomalies.

3. Benchmark predictions: persistence

Persistence of the observed anomalies is perhaps the simplest form of prediction. Persistence-based predictions will be used as benchmarks to assess the forecast quality of the dynamical predictions are given. The predictor is the monthly mean observed anomaly of the variable ahead of the start of the prediction.

The spatial distribution of the persistence-based skill is shown in Figure 1 for the four standard climatological seasons. Red circles appear over the areas where there is skill. Skill appears in summer over Southern Europe and the Middle East. This skill could arise from long-term trends and changes in the Bowen ratio associated with soil moisture anomalies in the more arid Mediterranean area. There is substantial skill over central and Northern Europe in spring along the snow line into central Europe, in agreement with the results described in Shongwe et al. (2007). Other candidates for the skill may include the influence of coastal SSTs, the snow albedo and the latent heat effects of snow melting. Positive skill appears over much of western Europe in autumn. It is less clear what processes could lead to this skill although both long-term trends and the role of Atlantic SSTs may be important. In winter the atmospheric internal variability is strong, which could explain the reduced skill of the persistence forecasts. These results agree well with those of Rodwell and Doblas-Reyes (2006). The skill shown decreases fast with lead time.

The persistence skill for precipitation is much lower than for temperature, except in Southern Europe in late winter, and it is hardly statistically significant in any area. One of the reasons for precipitation showing a lower skill due to persistence might be due to the lower precipitation long-term trends observed in the area (not shown).

More sophisticated methods to empirically predict seasonal climate variability have been developed, but they hardly give higher skill than the one shown here (Folland et al., 2012; Smith et al., 2012).

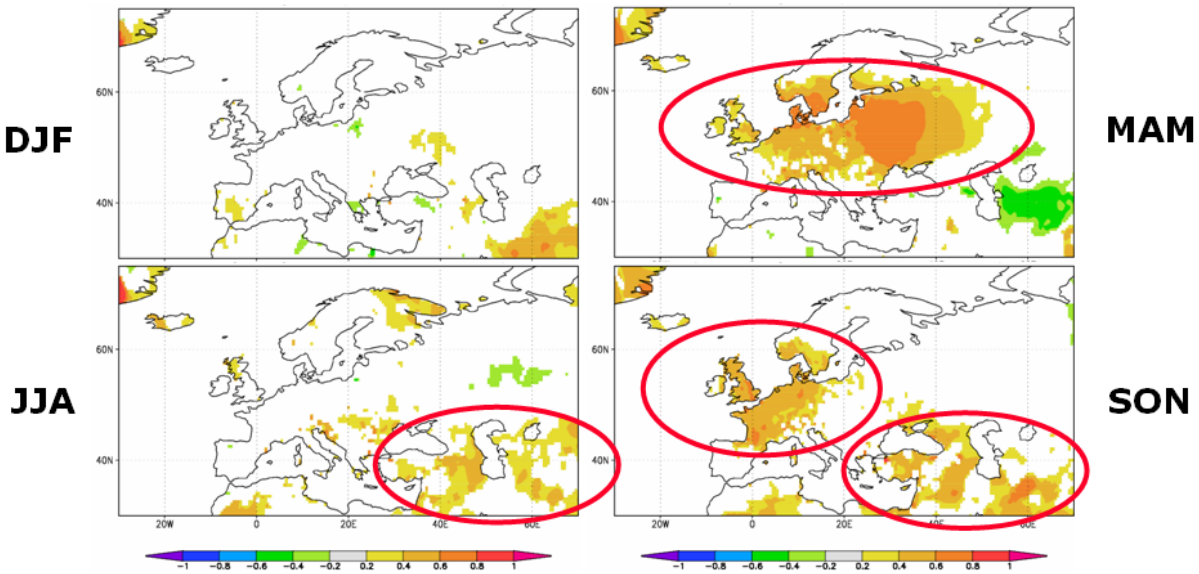

Figure 1: Correlation of one-month lead seasonal predictions from a persistence model based on the persistence of the anomalies of GHCN near-surface air temperature. Predictions have been computed over 1981-2008. Results are plotted only where the correlations are significant with 80% confidence. Red circles correspond to the areas where consistently positive skill is found.

4. Model drift and systematic error

Model inadequacy causes forecasts to drift away from the observed climate towards an imperfect model climate. The drift is the evolution of the systematic error with forecast time and is the cause of the model systematic error to depend on the start date. The systematic error is assessed as the difference in the features of the climatological distribution estimated from the model climate, which are the result of averaging predictions from all the years and ensemble members made from a specific start date, and of the observed or reference climate for the same calendar period.

Figure 2 shows that the mean sea level pressure systematic error has a strong seasonality. While in winter the model enhances the north-south gradient by deepening the Icelandic low and strengthening the Azores high leading to a stronger zonal flow over the Atlantic and Europe, the summer mean systematic error has a stronger and northward shifted Azores high. The equinoctial seasons are closer to the situation in winter. The mean systematic errors for temperature (Fig. 3) and precipitation agree well with the mean errors in circulation. This is because a variable like precipitation is dominated in most parts of Europe by atmospheric advection of moisture from the Atlantic (e.g. Ent et al., 2010). There is a warm and wet bias over northern Europe and dry and cold bias over the southern Mediterranean region in winter in agreement with the zonal flow excess. The warm and dry bias in summer over southeast Europe fits well with the collocated reduction in mean sea level pressure. An important wet bias is found in spring across northern Europe in agreement with the negative bias in mean sea level pressure.

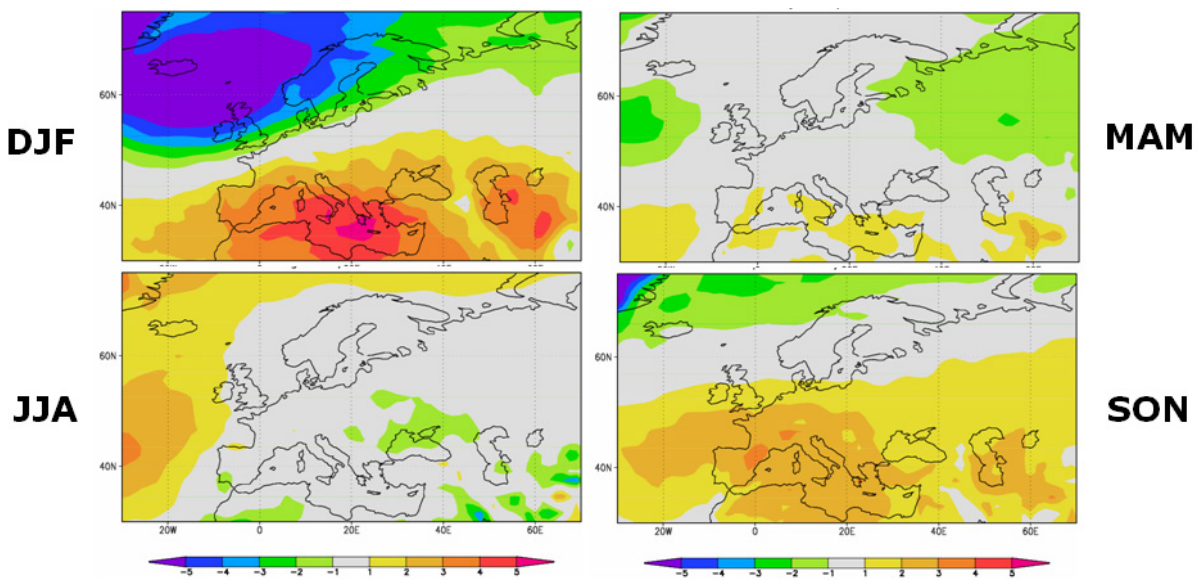

Figure 2: Mean sea level pressure systematic error in the mean for the System 3 one-month lead re-forecasts over the period 1981-2005. ERA Interim data have been used as reference.

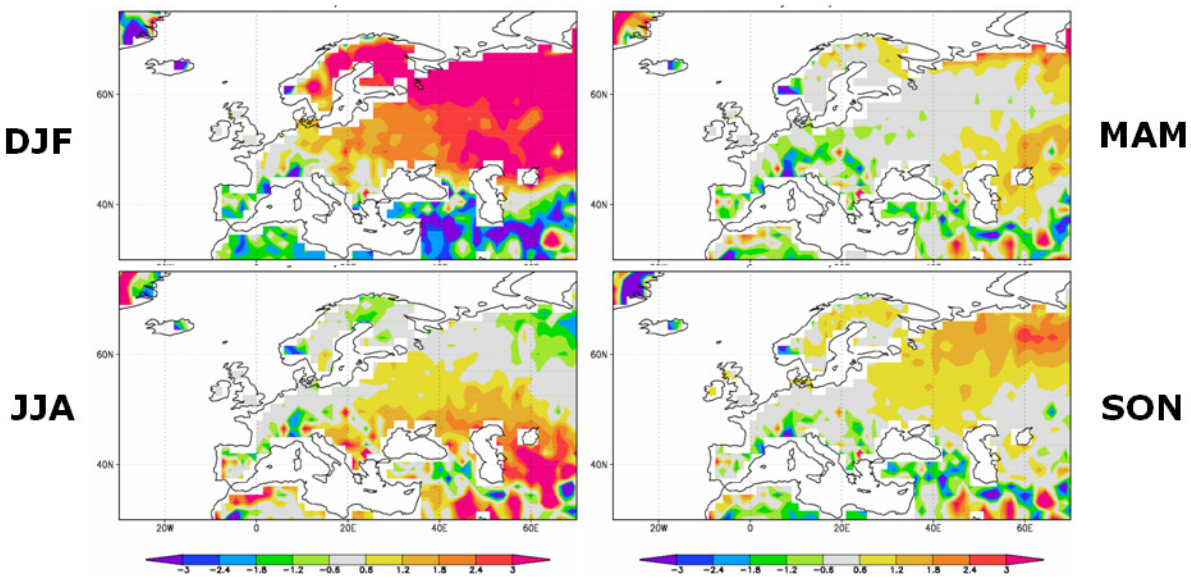

Figure 3: Near-surface air temperature systematic error in the mean for the System 3 one-month lead re-forecasts over the period 1981-2005. GHCN data have been used as reference.

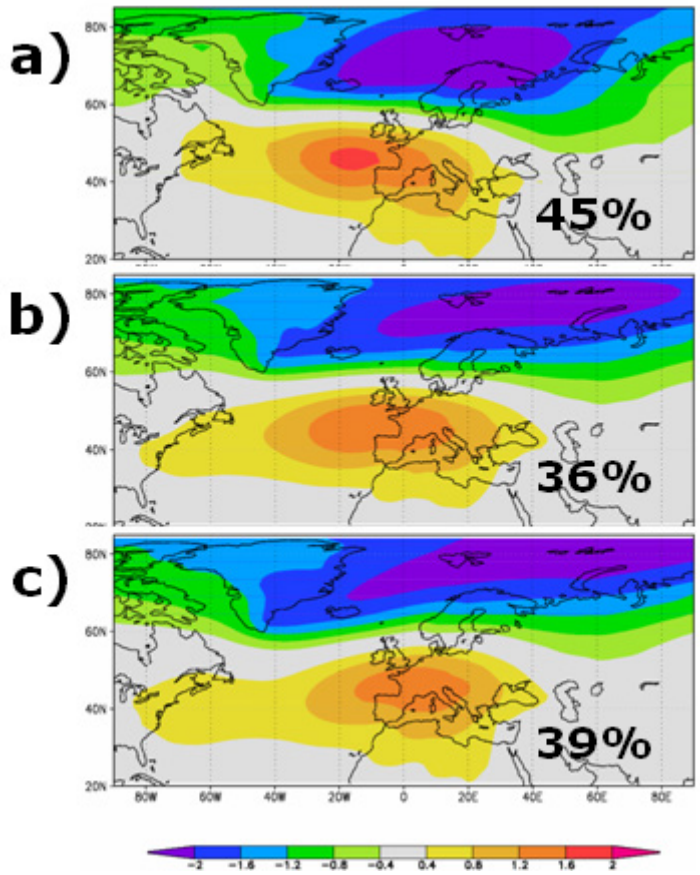

Figure 4: Winter leading empirical orthogonal function of mean sea level pressure for a) ERA Interim, b) one-month lead and c) three-month lead System 3 re-forecasts. Computations have been performed over the period 1981-2005. The fraction of explained variance appears at the bottom right corner of each panel. Units are arbitrary.

The drift is established in the first few months of the simulation. An analysis of the drift shows that the one-month and four-month lead time seasonal mean systematic error is very similar.

There are also important systematic errors in the variability. These errors are typically taken into account when formulating probabilistic predictions. Some systematic errors also appear in the spatial distribution of the variability, as shown in Figure 4. The leading empirical orthogonal function of the mean sea level pressure is typically used as a proxy of the North Atlantic Oscillation (NAO; Portis et al., 2001). While S3 reproduces all the NAO spatial features, it underestimates the variance associated, even with short lead times.

5. Forecast quality

Figure 5 illustrates the S3 near-surface air temperature ensemble-mean skill. The skill is significantly different from zero in the same regions where the persistence-based predictions described in Section 3 have skill, plus some additional regions such as central Europe in winter, the western Mediterranean region in summer and western Russia in autumn. In spite of the substantial systematic errors, S3 is able not only of persisting the slow variability associated with persistence but to propagate some of the initial-condition signal. Instead, the ensemble-mean skill of precipitation is as low as for the persistence-based predictions.

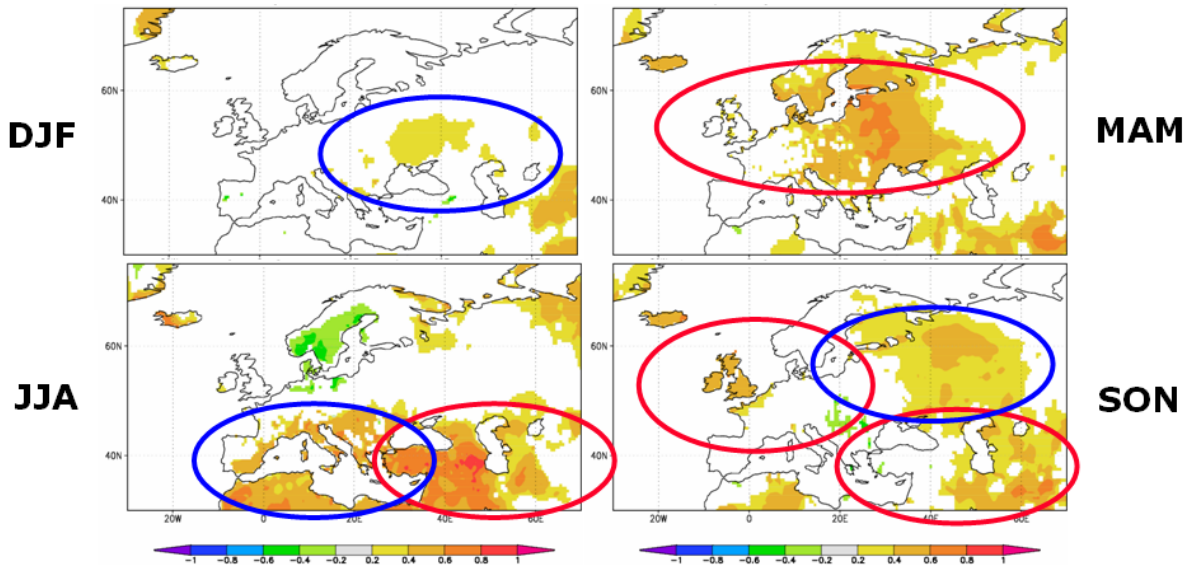

Figure 5: Correlation of one-month lead System 3 seasonal predictions of near-surface temperature computed over the 1981-2005 period. The GHCN dataset has been used as reference. Results are plotted only where the correlations are significant with 80% confidence. Blue (red) circles correspond to the areas where consistently positive skill is found for the dynamical (persistence-based and dynamical) predictions.

The skilful summer temperature predictions for southern Europe are illustrated in Figure 6. Part of the skill can be associated with the slow warming trend in the region that the forecast system reproduces correctly (Weisheimer et al., 2011a). The impact on the skill of the global warming trend is shown in Figure 7. The ensemble-mean correlation of temperature is reduced almost everywhere when the predictions are linearly detrended, except in the areas where the skill can be linked to variations in the snow cover. However, large skill improvements could still be achieved by correctly reproducing the global-warming effect on the predictions. The forecast system is far from correctly predicting the variations in the global-mean temperature, for which the one-month lead ensemble-mean skill is 0.69 in summer and 0.44 in winter. Besides, the local impact of global warming is not correctly represented in the forecast system as found with the regression of the near-surface air temperature on the global-mean temperature (not shown). The problem is particularly obvious over the eastern North Atlantic, as it has already been found in historical uninitialized experiments (Oldenborgh et al., 2009), and is another reason to keep improving the dynamical models. A similar analysis performed for the precipitation re-forecasts suggests that the local variability linked to the global-warming effect is much smaller than for temperature.

Beyond the impact of global warming on European skill, the NAO could be considered as another source of seasonal predictability. The S3 skill for the NAO is low compared to what is obtained for other main modes of variability like ENSO. In spite of the satisfactory representation of the NAO pattern in the forecast system, the highest correlation is not higher than 0.35 and appears in winter.

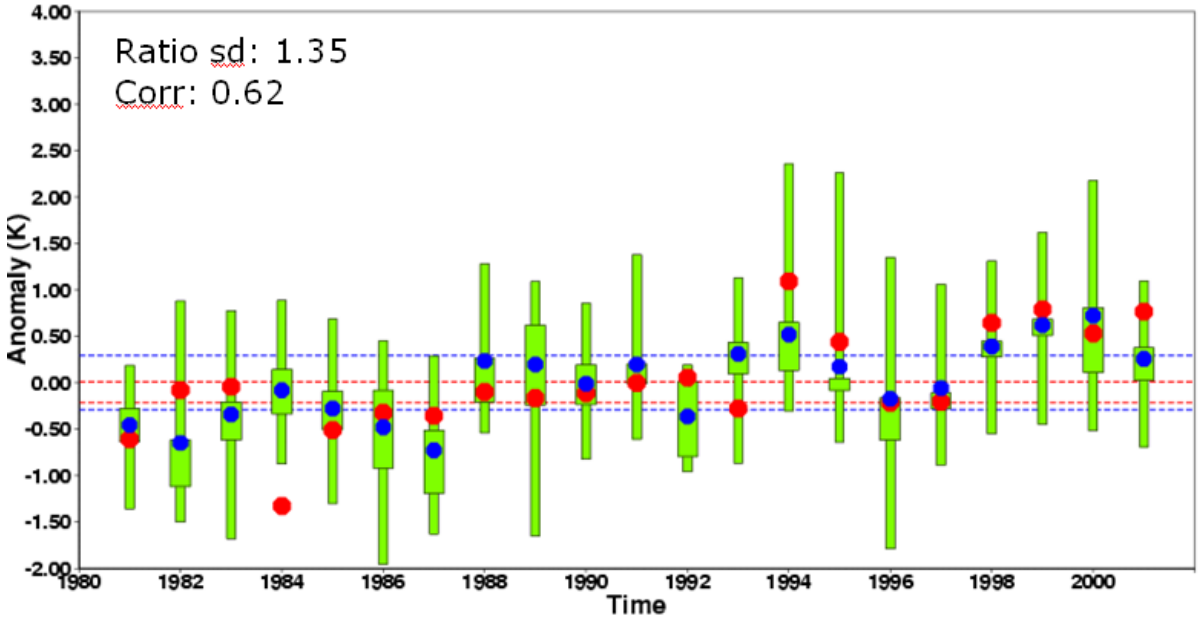

Figure 6: One-month lead System 3 summer seasonal predictions of temperature for southern Europe (30°-45°N, 10°W-40°E). The green box-and-whisker plots represent the range of the 11-member ensemble, with the central box corresponding to the interquartile range, the blue dot the ensemble mean and the red dot the observational reference (ERA Interim). The ensemble-mean correlation with the reference is included in the box. The blue (red) horizontal lines are for the upper and lower climatological terciles for the re-forecasts (observational reference). The variance overestimation indicated in the box agrees well with the larger range between the terciles of the predictions than between the reference terciles.

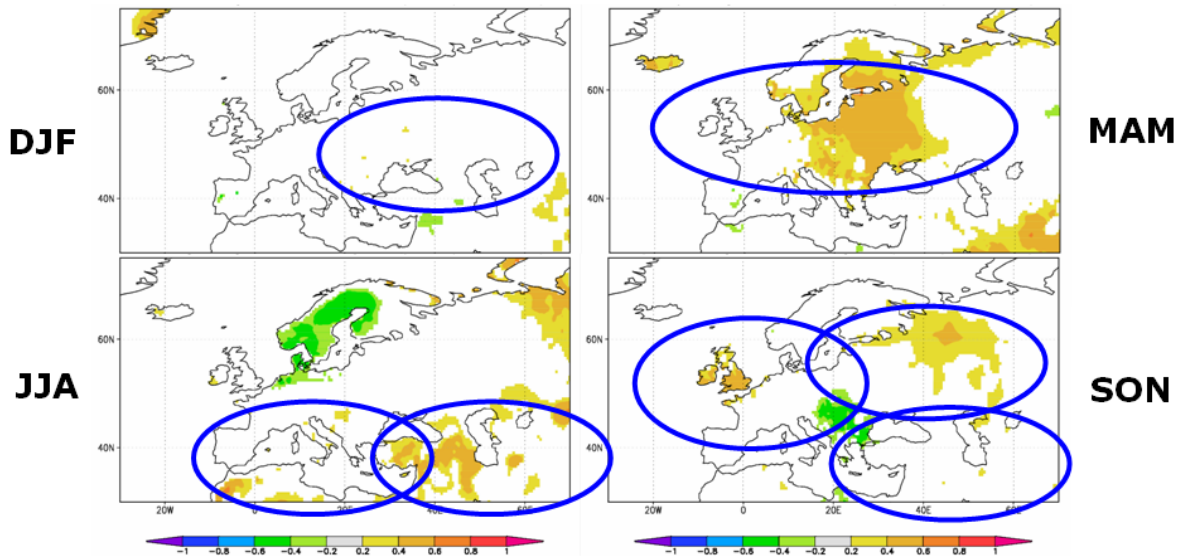

Figure 7: Correlation of linearly detrended one-month lead System 3 seasonal predictions of near-surface temperature computed over the 1981-2005 period. To detrend the forecasts, the linear regression against the predicted and observed global-mean temperature has been removed from the anomalies. The GHCN dataset has been used as reference. Results are plotted only where the correlations are significant with 80% confidence. Blue circles correspond to the areas where consistently positive skill is found for the dynamical predictions in Figure 4.

The impact of ENSO on the European region as a skill source deserves a bit more understanding. S3 is one of the best seasonal forecast systems predicting ENSO (Stockdale et al., 2011). However, the teleconnections between ENSO and the European climate are most often than not, wrongly reproduced in the model. Figure 8 shows an example for one of the best known teleconnections between the European climate and ENSO: the link between the western Mediterranean spring precipitation and La Niña events (Oldenborgh et al., 2000). The teleconnections have the opposite sign to that found with observational data, especially in western Europe. This has been found for several variables, seasons and lead times and is one of the main limitations that hampers the production of more skilful European forecasts. It is possible that ENSO, as a major source of predictability at seasonal time scales, is only active in the European region under certain conditions, offering windows of opportunity for European seasonal prediction (Frías et al., 2010).

Climate forecasts are essentially probabilistic. This means that a statistical model is required to transform the ensemble of predictions into a probability density function. Figure 9 shows a cartoon of some of the approaches typically used. While the simplest option consists in using a histogram based on the individual members, probability density functions can be estimated in either a parametric (as with fitting a Gaussian distribution to the ensemble) or non-parametric (e.g. with a kernel dressing) way. The quality of these predictions, when considered for dichotomous events, can be assessed with the reliability diagram (Fig. 10). The probabilistic scores for the European region, as happens with the ensemble-mean skill measures, are lower than for most of other regions. The reliability diagrams in Figure 10 show that most of the predictions are clustered around the climatological frequency, with a flat curve for the most populated bins. The flatness of the curve suggests that the system is unable to reliably distinguish between events and non-events when the probability forecast is close to the climatological frequency. For more extreme probabilities (i.e. close to zero or one), the curve depicts a slope, suggesting that the system is more skilful when issuing probabilities within that range.

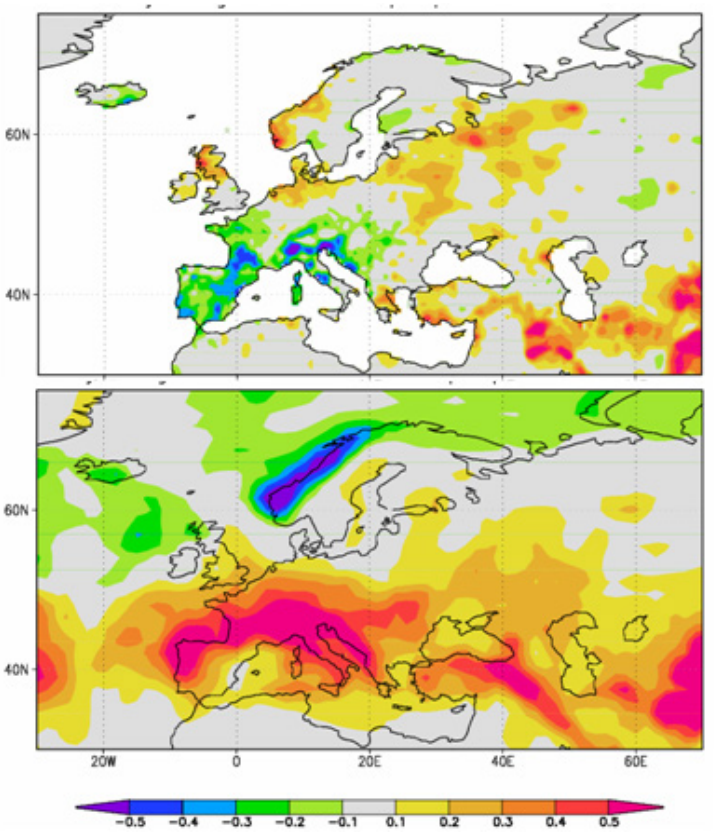

Figure 8: Regression (mm/day/K) of (top) GPCC spring precipitation on HadISST1 Niño3.4 time series and (bottom) S3 one-month lead spring precipitation on the predicted Niño3.4 time series. The computations have been performed over 1981-2005.

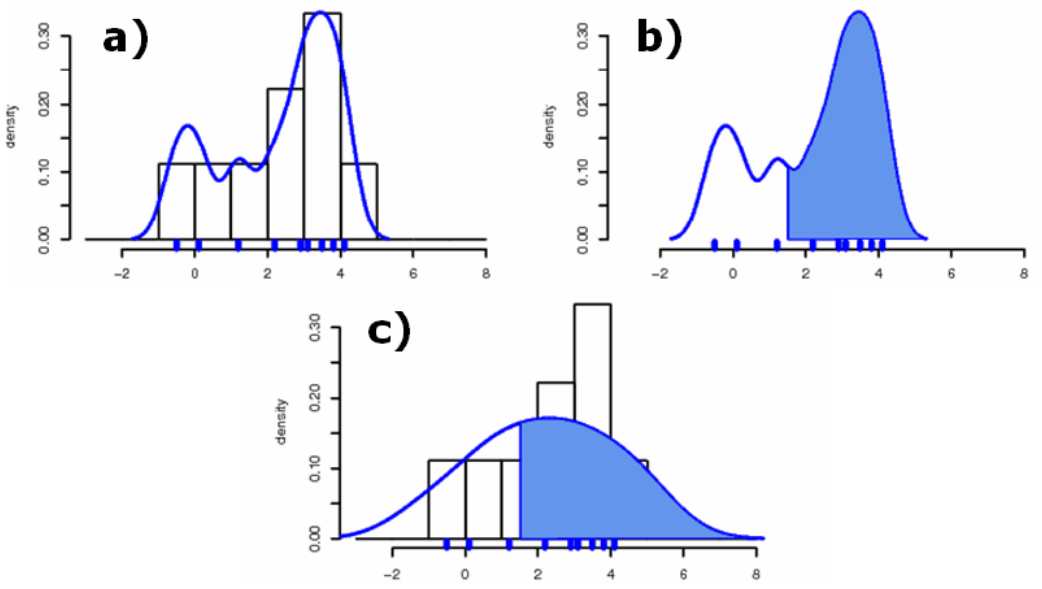

Figure 9: Illustration of different approaches to obtain probability forecasts from an arbitrary ensemble forecast. The ensemble-member values are shown with blue short lines. (a) A histogram can be constructed from the ensemble values and relative frequencies estimated from it. The blue line corresponds to a kernel estimate of the probability density function, and matches quite closely the histogram. (b) For dichotomous events, the probability forecast is computed with respect to a threshold, which in the figure has been chosen as 1.75. The blue area corresponds to the forecast probability of the variable being above the chosen threshold. (c) An alternative method to estimate the probability density function consists in fitting a parametric function, in this case a Gaussian, to the ensemble values. The blue area is the forecast probability of the variable being above the chosen threshold, and is different from that shown in panel b.

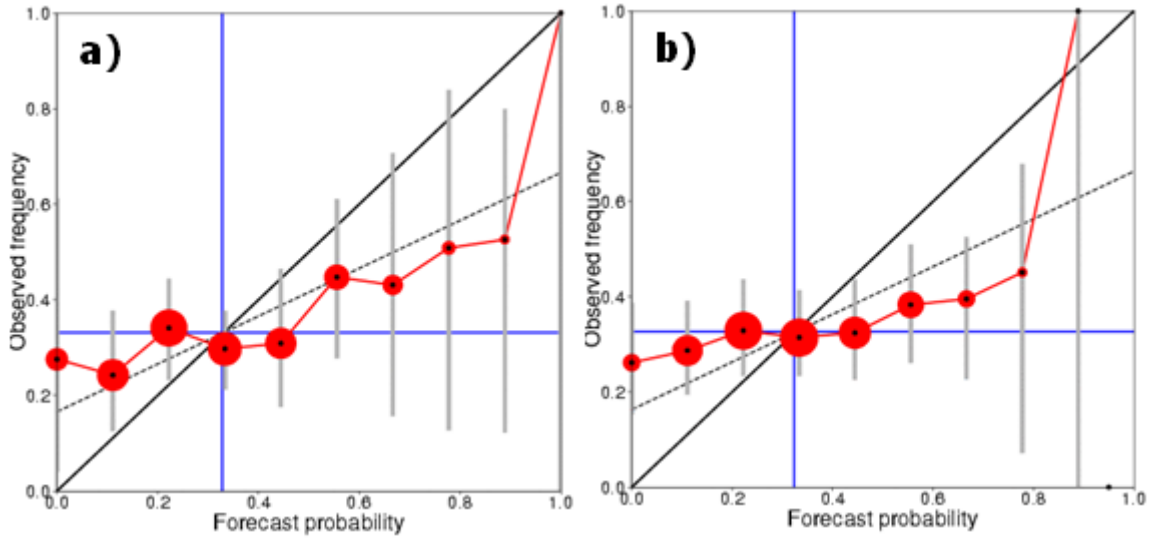

Figure 10: Reliability diagrams for one-month lead System 3 summer seasonal predictions of a) near-surface temperature and b) precipitation above the upper tercile for southern Europe (30°-45°N, 10°W-40°E) computed over the 1981-2005 period. The size of the circles corresponds to the relative frequency each forecast probability bin is populated. The grey bars are for the 95% confidence intervals of the observed relative frequency at each bin obtained with a 1,000 sample bootstrap method. The vertical blue line indicates the average probability while the horizontal one is for the climatological frequency of the event (theoretically one third for this type of events based on climatological terciles). The dashed grey line determines the areas of the diagram with no skill.

Unfortunately, those probabilities are issued less frequently. As a result, the Brier and ROC skill scores for the event "anomalies above the upper tercile" are -0.078, with a 95% confidence interval of (-0.287,0.093) and 0.148 (-0.073,0.379) for temperature and -0.088 (-0.167,-0.038) and 0.075 (-0.027,0.175) for precipitation, respectively. Similar results have been found for northern Europe.

Multi-model, perturbed parameters and stochastic physics approaches to deal with model uncertainty have been found to improve both reliability and skill of probability predictions (Doblas-Reyes et al., 2009). This is also the case for Europe, as described in Weisheimer et al. (2011b).

6. Summary and discussion

This study offers an illustration of the current level of skill over Europe of a state-of-the-art global seasonal forecast system. While the forecast quality of European seasonal predictions is limited, there are regions with significant skill linked to global warming and soil processes such as snow cover. The skill goes beyond what can be achieved by using a simple persistence model and is found in spite of the important systematic errors of the system. ENSO being the main source of global seasonal skill, it has been found that this forecast system has important drawbacks as far as the ENSO teleconnections is concerned. The NAO, another main player in determining European climate variability, has limited skill, mainly in winter.

Apart from ENSO, other sources of skill linked to SST variability are the tropical and extratropical Atlantic. However, the processes linked to such predictability seem to need an adequate representation in the forecast system. In a maximum covariance analysis between the North Atlantic winter mean sea level pressure and the preceding northern tropical Atlantic SSTs (not shown), the leading pair of patterns links an NAO-like winter circulation to a warming of the tropical Atlantic in autumn, while the second pair links a wave train across the North Atlantic to a dipole in the tropical Atlantic. When the maximum covariance analysis is applied between the observed October SSTs and the one-month lead time mean sea level pressure forecasts, the patterns over the tropical Atlantic are very similar to those obtained with the observational data while the winter pressure patterns barely resemble those found in the observations. This suggests that the forecast system might not be taking advantage of the predictability from the tropical Atlantic SSTs. As for the extratropical SSTs, Rodwell and Folland (2002) and Iwi et al. (2006) found a link between a tripolar North Atlantic SST pattern in May and an NAO-like signal the following winter. As in the previous case, the forecast system has problems to reproduce this link in a maximum covariance analysis between the observed May SSTs and the one-month lead winter mean sea level pressure, where the simulated pattern is degenerated beyond the leading pair and shows large differences with the one obtained from the observations. The difficulty for S3 to take advantage of these sources of skill, a problem that can easily be extrapolated to many other global seasonal forecast systems, could be linked to the many systematic errors in the SSTs (for instance, in the tropical Atlantic) and the tropical-extratropical circulations. Improvements in the models that go beyond resolution increases will be fundamental to go beyond the current levels of European seasonal skill.

Many different initiatives are ongoing to improve seasonal prediction over Europe. Important modifications of and sensitivity experiments to the initialization of soil moisture (Koster et al., 2011), snow cover (Orsolini et al., 2012) and sea ice (Chevallier and Salas-Melià, 2012) are being explored. Fundamental topics like the commonality of the errors of the different dynamical forecast systems or the formulation of more process-based empirical models (Smith et al., 2011) will be addressed more widely in the near future. Special mention deserves the efforts that will be necessary to reduce the different aspects of the systematic error and, among them, the contentious role of the stratosphere to modulate the action of the different signals playing a role in determining the European climate (Ineson and Scaife, 2009) and the correct representation of the impact of the natural and anthropogenic changes in atmospheric composition. There seems to be a long way to go yet before producing seasonal predictions useful for a wide range of users, but the present predictions can already be of use for some well-trained ones.

Acknowledgements

This work was supported by the QWeCI project (GOCE-CT-2003-505539) and the MICINN-funded project RUCSS. The author acknowledges the significant contributions by L. Rodrigues, B. van den Hurk, G.J. van Oldenborgh, A. Weisheimer and the seasonal forecast section at ECMWF.

References

Balmaseda, M.A., A. Vidard and D. Anderson, 2008. The ECMWF ORA-S3 ocean analysis system. *Mon. Weather Rev*, **136**, 3018-3034.

Challinor, A.J., J.M. Slingo, T.R. Wheeler and F.J. Doblas-Reyes, 2005. Probabilistic simulations of crop yield over western India using the DEMETER seasonal hindcast ensembles. *Tellus*, **57A**, 498–512.

Chang, P., T. Yamagata, P. Schopf, S. K. Behera, J. Carton, W. S. Kessler, G. Meyers, T. Qu, F. Schott, S. Shetye and S.-P. Xie, 2006. Climate fluctuations of tropical coupled systems - The role of ocean dynamics. *J. Climate*, **19**, 5122-5174.

Chevallier, M. and D. Salas-Melià, 2012. The role of sea ice thickness distribution in the Arctic sea ice potential predictability: A diagnostic approach with a coupled GCM. *J. Climate*, **25**, 3025-3038.

Doblas-Reyes, F.J., M. Déqué and J.P. Piedelievre, 2000. Multimodel spread and probabilistic seasonal forecast in PROVOST. *Quart. J. Roy. Meteorol. Soc*., **126**, 2069-2087.

Doblas-Reyes, F.J., R. Hagedorn, T. N. Palmer, and J.-J. Morcrette, 2006. Impact of increasing greenhouse gas concentrations in seasonal ensemble forecasts. *Geophys. Res. Lett*., **33**, L07708, doi:10.1029/2005GL025061.

Doblas-Reyes, F.J., A. Weisheimer, M. Déqué, N. Keenlyside, M. McVean, J.M. Murphy, P. Rogel, D. Smith and T.N. Palmer, 2009. Addressing model uncertainty in seasonal and annual dynamical seasonal forecasts. Q*uart. J. Roy. Meteorol. Soc.*, **135**, 1538-1559, doi:10.1002/qj.464.

Douville, H., 2004. Relevance of soil moisture for seasonal atmospheric predictions: Is it an initial value problem? *Climate Dyn.*, **22**, 429-446.

Ent, R.J. van der, H.H.G. Savenije, B. Schaefli and S.C. Steele-Dunne, 2010. The origin and fate of atmospheric moisture over 2 continents. *Water Resour. Res*., **46**, W09525, doi:10.1029/2010WR009127.

Feddersen, H. and U. Andersen, 2005. A method for statistical downscaling of seasonal ensemble predictions. *Tellus*, **57A**, 398-408.

Folland, C.K., A.A. Scaife, J. Lindesay and D.B. Stephenson, 2012. How potentially predictable is northern European winter climate a season ahead? *Int. J. Climatol*., **32**, 801-818, doi:10.1002/joc.2314.

Frías, M.D., S. Herrera, A.S. Cofiño and J.M. Gutiérrez, 2010. Assessing the skill of precipitation and temperature seasonal forecasts in Spain: Windows of opportunity related to ENSO events. *J. Climate,* **23**, 209-220, doi: 10.1175/2009JCLI2824.1.

García-Morales, M. and L. Dubus, 2007. Forecasting precipitation for hydroelectric power management: How to exploit GCM's seasonal ensemble forecasts. *Int. J. Climatol*., **27**, 1691-1705.

Gershunov, A. and D.R. Cayan, 2003. Heavy daily precipitation frequency over the contiguous United States: Sources of climatic variability and seasonal predictability. *J. Climate*, **16**, 2752–2765.

Ineson, S. and A.A. Scaife, 2009. The role of the stratosphere in the European climate response to El Niño. *Nature Geoscience*, **2**, 32-36, doi:10.1038/ngeo381.

Iwi, A.M., R.T. Sutton and W.A. Norton, 2006. Influence of May Atlantic Ocean initial conditions on the subsequent North Atlantic winter climate. *Quart. J. Roy. Meteorol. Soc*., **132**, 2977–2999, doi: 10.1256/qj.05.62.

Kirtman, B.P. and A. Pirani, 2009. The state of the art of seasonal prediction: Outcomes and recommendations from the First World Climate Research Program Workshop on Seasonal Prediction. *Bull. Amer. Meteorol. Soc*., **90**, 455-458.

Koster, R.D., S.P.P. Mahanama, T.J. Yamada, G. Balsamo, A.A. Berg, M. Boisserie, P.A. Dirmeyer, F.J. Doblas-Reyes, G. Drewitt, C.T. Gordon, Z. Guo, J-H. Jeong, W-S. Lee, Z. Li, L. Luo, S. Malyshev, W. J. Merryfield, S. I. Seneviratne, T. Stanelle, B.J.J.M. van den Hurk, F. Vitart and E.F. Wood, 2011. The second phase of the global land-atmosphere coupling experiment: Soil moisture contributions to subseasonal forecast skill. *J. Hydrometeorol*., **12**, 805-822, doi:10.1175/2011JHM1365.1.

Lienert, F., J.C. Fyfe and W.J. Merryfield, 2011. Do climate models capture the tropical influences on North Pacific sea surface temperature variability? *J. Climate*, **24**, 6203-6209.

Marshall, A.G. and A.A. Scaife, 2009. Impact of the QBO on surface winter climate. *J. Geophys. Res.*, **114**, D18110, doi:10.1029/2009JD011737.

Marshall, A.G. and A.A. Scaife, 2010. Improved predictability of stratospheric sudden warming events in an atmospheric general circulation model with enhanced stratospheric resolution. *J. Geophys. Res*., **115**, D16114, doi:10.1029/2009JD012643.

Oldenborgh, G.J. van, G. Burgers and A. Klein Tank, 2000. On the El-Niño teleconnection to spring precipitation in Europe. *Int. J. Climatology*, **20**, 565-574.

Oldenborgh, G.J. van, S.S. Drijfhout, A. van Ulden, R. Haarsma, A. Sterl, C. Severijns, W. Hazeleger and H. Dijkstra, 2009. Western Europe is warming much faster than expected. *Climate Past*, **5**, 1-12, doi:10.5194/cp-5-1-2009.

Orsolini, Y.J., R. Senan, G. Balsamo, F.J. Doblas-Reyes, F. Vitart, A. Weisheimer, A. Carrasco and R.E. Benestad, 2012. Impact of snow initialization on sub-seasonal forecasts. *Climate Dyn*., submitted.

Peterson, T.C. and R.S. Vose, 1997. An overview of the Global Historical Climatology Network temperature database. *Bull. Amer. Meteorol. Soc*., **78**, 2837-2849.

Portis, D.H., J.E. Walsh, M. El Hamly and P.J. Lamb, 2001. Seasonality of the North Atlantic Oscillation. *J. Climate*, **14**, 2069-2078.

Quan, X.W., M.P. Hoerling, J.S. Whitaker, G.T. Bates and T.Y. Xu, 2006. Diagnosing sources of U.S. seasonal forecast skill. *J. Climate*, **19**, 3279–3293.

Rodwell, M.J. and C.K. Folland, 2002. Atlantic air–sea interaction and seasonal predictability. *Quart. J. Roy. Meteorol. Soc*., **128**, 1413–1443.

Rodwell, M. and F.J. Doblas-Reyes, 2006. Predictability and prediction of European monthly to seasonal climate anomalies. *J. Climate*, **19**, 6025-6046.

Saha, S., S. Nadiga, C. Thiaw, J. Wang, W. Wang, Q. Zhang, H. M. van den Dool, H.-L. Pan, S. Moorthi, D. Behringer, D. Stokes, M. Pena, S. Lord, G. White, W. Ebisuzaki, P. Peng and P. Xie, 2006. The NCEP climate forecast system. *J. Climate*, **19**, 3483–3517.

Schneider U., T. Fuchs, A. Meyer-Christoffer and B. Rudolf, 2008. Global precipitation analysis products of the Global Precipitation Climatology Centre (GPCC). Deutscher Wetterdients (DWD), 12 pp. Available from www.ksb.dwd.de.

Shongwe, M., C. Ferro, C. Coelho and G.J. van Oldenborgh, 2007. Predictability of cold spring seasons in Europe. *Mon. Weather Rev*., **135**, 4185-4201.

Shukla, J. and J.L.I. Kinter, 2006. Predictability of seasonal climate variations a pedagogical review. P*redictability of Weather and Climate*, T. Palmer, and R. Hagedorn Eds., Cambridge University Press.

Smith, K.L., P.J. Kushner and J. Cohen, 2011. The role of linear interference in Northern Annular Mode variability associated with Eurasian snow cover extent. *J. Climate*, **24**, 6185-6202.

Smith, D.M., A.A. Scaife and B.P. Kirtman, 2012. What is the current state of scientific knowledge with regard to seasonal and decadal forecasting? *Environ. Res. Lett*., **7**, doi:10.1088/1748-9326/7/1/015602.

Sordo, C.M., M.D. Frías, S. Herrera, A.S. Cofiño and J.M. Gutiérrez, 2008. Interval-based statistical validation of operational seasonal forecasts in Spain conditioned to El Niño-Southern Oscillation events. *J. Geophys. Res*., **113**, D17121, doi:10.1029/2007JD009536.

Stockdale, T.N., D.L.T. Anderson, M.A. Balmaseda, F.J. Doblas-Reyes, L. Ferranti, K. Mogensen, T.N. Palmer, F. Molteni and F. Vitart, 2011. ECMWF seasonal forecast system 3 and its prediction of sea surface temperature. *Climate Dyn*., **37**, 455-471, doi: 10.1007/s00382-010-0947-3.

Thompson, M.C., F.J. Doblas-Reyes, S.J. Mason, R. Hagedorn, S.J. Connor, T. Phindela, A.P. Morse and T.N. Palmer, 2006. Malaria early warnings based on seasonal climate forecasts from multi-model ensembles. *Nature*, **439**, 576–579.

Weisheimer, A., F.J. Doblas-Reyes, T. Jung and T.N. Palmer, 2011a. On the predictability of the extreme summer 2003 over Europe. *Geophys. Res. Letters*, **38**, L05704, doi:10.1029/2010GL046455.

Weisheimer, A., T.N. Palmer and F.J. Doblas-Reyes, 2011b. Assessment of representations of model uncertainty in monthly and seasonal forecast ensembles. *Geophys. Res. Letters*, **38**, L16703, doi:10.1029/2011GL048123.