SPECIAL PROJECT PROGRESS REPORT

Progress Reports should be 2 to 10 pages in length, depending on importance of the project. All the following mandatory information needs to be provided.

Reporting year	2011
Project Title:	Assessment of the limit of initial-condition useful skill in interannual climate prediction
Computer Project Account:	**SPESICCF**
Principal Investigator(s):	Dr. Francisco J. Doblas-Reyes
Affiliation:	IC3 - Institut Català de Ciències del Clima
Name of ECMWF scientist(s) collaborating to the project (if applicable)	……………………………………………………………… ………………………………………………………………
Start date of the project:	January 2010
Expected end date:	December 2011

Computer resources allocated/used for the current year and the previous one

(if applicable)

Please answer for all project resources

		Previous year		**Current year**	
		Allocated	Used	Allocated	Used
High Performance Computing Facility	(units)	3,135,000	3,135,000	1,461,000	105
Data storage capacity	(Gbytes)	5,225	4,500	2,435	

Summary of project objectives
(10 lines max)
The goal of the project consists in exploring the usefulness of the EC-Earth model as a seamless system. While for EC-Earth both climate predictions and climate-change projections are or will be made available soon by the consortium, this project will perform shorter-term (from seasonal to interannual) climate predictions that will allow to cover the full spectrum of simulations that can be performed with a seamless system. Besides, EC-Earth will be used as a test-bed to understand the possible improvements to the ECMWF operational seasonal forecast system coming from a climate model. Among the processes that will benefit from this collaboration are the initialization of the sea-ice component, the treatment of the boundary forcings and the assessment of the relevance of new physical modules such as interactive vegetation. The project considers a comprehensive set of ensemble interannual hindcasts with different initialization strategies.

Summary of problems encountered (if any)
(20 lines max)
The ECMWF systems worked as planned and no major technical problems were encountered. Many experiments, some of whose results are described below, have been carried out. This work allowed learning many things about the EC-Earth performance as a seasonal-to-interannual forecast system, as well as the relevance of using realistic sea-ice initial conditions. However, it was found a few months ago that there were major errors in the way anthropogenic and natural aerosols were implemented in the version of IFS used in EC-Earth v2.2. This had a strong impact on some of the results, in particular those beyond the seasonal time scale, which is one of the main objectives of this special project.
There is a corrected EC-Earth version available since a few weeks ago and is being extensively tested. However, many of the results presented here can not be published in peer-review journals due to the effect of the regional distribution of aerosols on forecast quality. Hence, some of the experiments will be repeated with the remaining SBUs. The selection of the experiments is currently being decided, although it is likely that they will include the sensitivity to the sea-ice initial conditions and the two-year experiments started in May and November of each year.

Summary of results of the current year (from July of previous year to June of current year)
This section should comprise 1 to 8 pages and can be replaced by a short summary plus an existing scientific report on the project

A set of seven-month, five-member hindcasts have been performed with EC-Earth v2.2 over the period 1981-2005. Four start dates per year, i.e. the first of February, May, August and November, were used. This experiment has been compared in terms of forecast quality with the eleven-member ECMWF's operational seasonal forecast System 3 (Stockdale et al., 2011). It is important to remind that, even if EC-Earth is originally based on the same cycle as System 3, 31R1, EC-Earth has implemented the soil scheme, convection and some boundary layer features from cycle 33R1. Weisheimer et al. (2011) have shown the relevance of some of these changes in predicting extreme summer temperature over Europe.
The atmosphere and land surface initialization was taken from the ERA-40 reanalysis (Uppala et al., 2005) for all start dates before 1989 and ERA-Interim afterwards. The atmospheric perturbations are applied to all members except for the first one and are based on the operational singular vectors (Magnusson et al., 2008). The perturbations are added at the initial time to all the prognostic variables except humidity. The ocean initial conditions have been taken from the 3D-Var five-member ensemble ocean re-analysis known as NEMOVAR-COMBINE (Balmaseda et al., 2010). The re-analysis has been performed with a variational data assimilation system based on the ocean

model NEMO v3.0 where profiles of temperature and salinity from a quality controlled EN3_v2a data set (Ingleby and Huddleston, 2007) were assimilated. The assimilation cycle is 10 days and a bias correction method has been included. The NEMO model is forced by ERA40 fluxes from 1957 to 1988 and by ERA-Interim thereafter (Balmaseda et al. 2010) and includes a strong relaxation to observed SSTs. The reanalysis covers the period of 1958-2009. The five members are generated by perturbing the wind stress, the ocean initial conditions and the amount of observations assimilated. The comparison with independent observations shows that the assimilation of ocean data helps constraining the uncertainty introduced by ocean models and forcing fluxes, mostly in the upper ocean. The sea-ice initial conditions are produced from a NEMO v2.0 coupled to LIM2 driven by the DFS4.3 ocean forcing data (Brodeau et al. 2009). This forcing dataset is derived from ERA40/ERAInterim with corrections on tropical surface air humidity, Arctic sea surface temperature and global wind field based on high quality observations.
Figures 1 and 2 show the mean bias of the two-metre air temperature and the precipitation for the boreal summer and winter of the one-month lead seasonal forecasts. Biases are often comparable in magnitude to the anomalies that we seek to predict. There are larger temperature errors over the tropics and the Antarctic circumpolar region in EC-Earth than in System 3, even from forecast month one (not shown).EC-Earth is generally cooler over the ocean, in particular the equatorial Pacific, than System 3, although EC-Earth has the benefit of a substantial reduction of the warm bias in the Northern Hemisphere high-latitude land areas in winter (the relevant areas are marked with red ellipses). The land air temperature bias is likely reduced by the introduction of a more sophisticated snow scheme in EC-Earth with respect to System 3 (Dutra et al., 2010). Precipitation also shows differences between the two forecast systems, in particular over the equatorial Atlantic and Pacific in boreal summer. This leads to a better West African monsoon precipitation representation in EC-Earth.

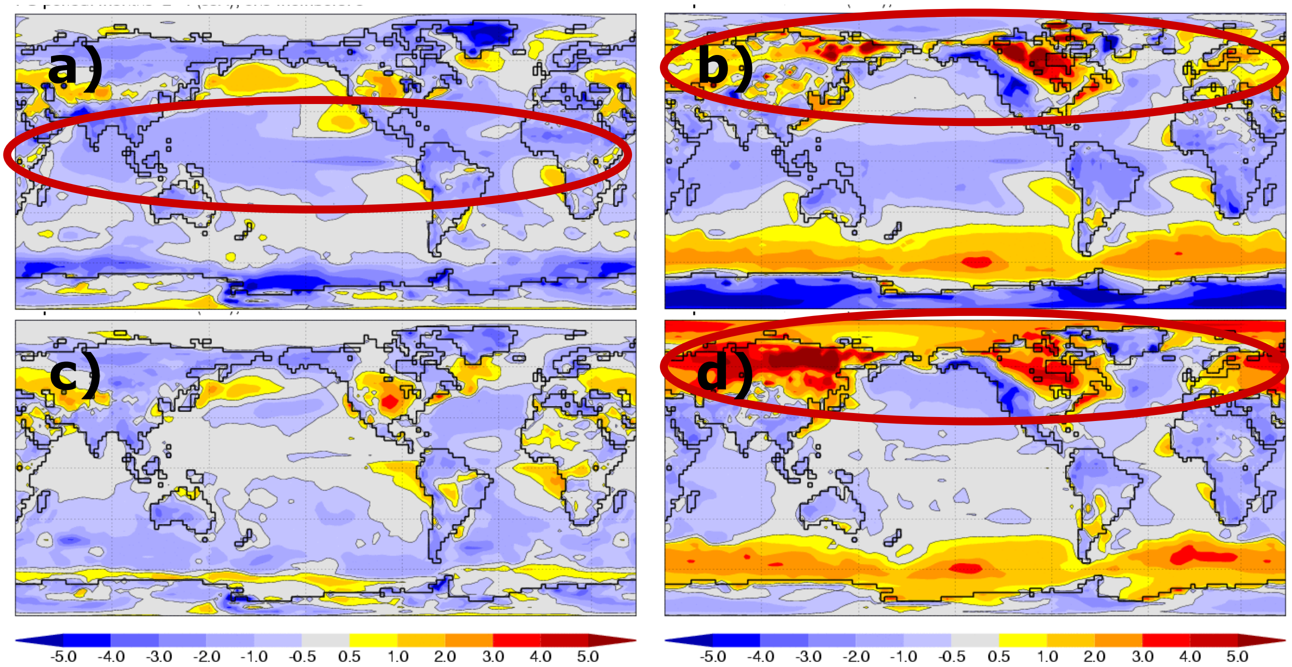

Figure 1: Mean bias of the one-month lead, 2-4 month near-surface air temperature re-forecasts for EC-Earth (a, b) and System 3 (c, d) with the May (a, c) and November (b, d) start dates. Re-forecasts for the period 1981-2005 have been used. The reference is taken from ERA40/Int.

ENSO being the most relevant phenomenon at seasonal-to-interannual time scales, a basic validation of a forecast system should start by estimating the ability to predict its variability. Figure 3 displays time series of Niño 3.4 sea surface temperature time series for the one-month lead boreal summer seasonal predictions. Both forecast systems have a similar forecast quality in terms of ensemble-mean correlation and RPSS. An important difference between both systems is the larger amplitude of the simulated anomalies in EC-Earth, which leads to an overestimation of the signal of the large event produced in 1997. This is typical of systems with a strong cold bias and is a feature that disappears with longer lead times (not shown).

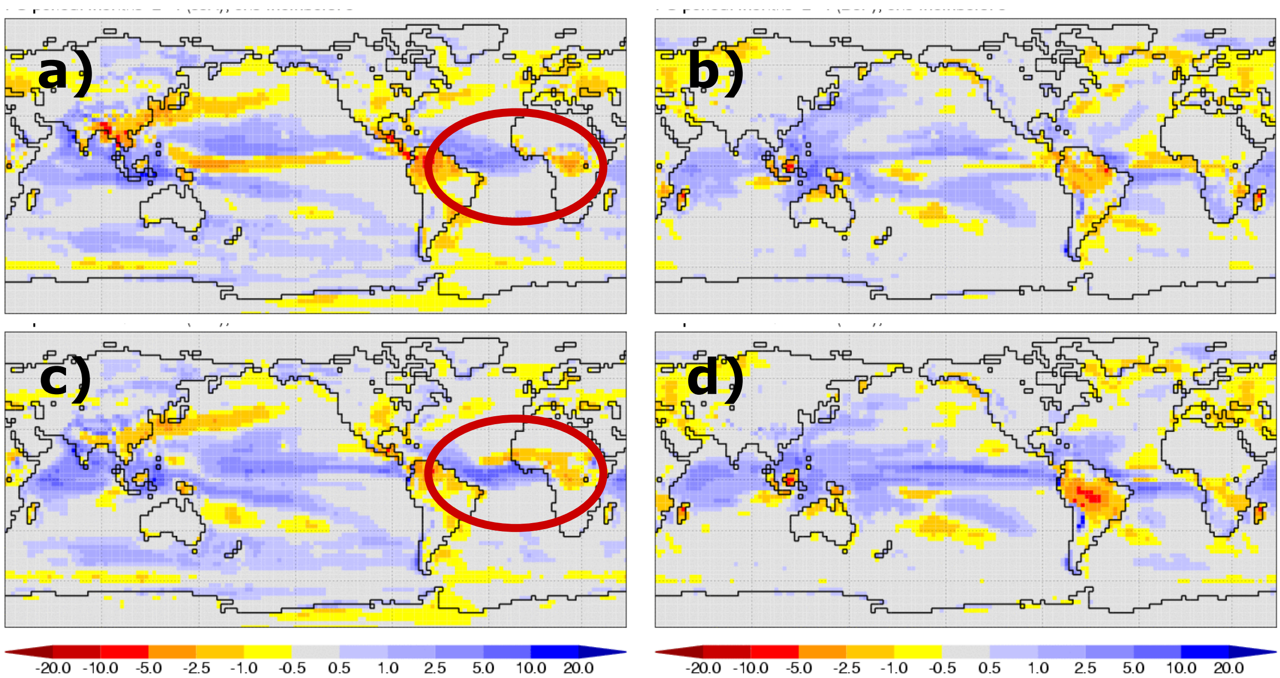

Figure 2: Mean bias of the one-month lead, 2-4 month precipitation re-forecasts for EC-Earth (a, b) and System 3 (c, d) with the May (a, c) and November (b, d) start dates. Re-forecasts for the period 1981-2005 have been used. The reference is taken from GPCP.

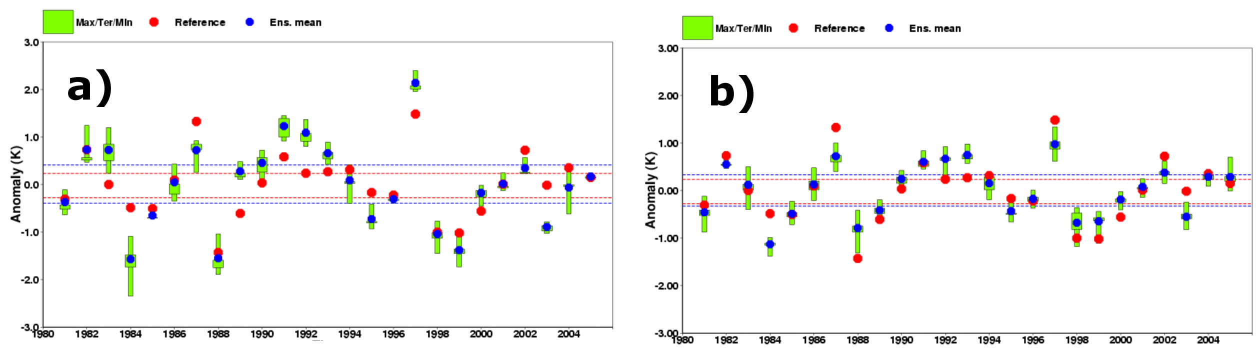

Figure 3: Niño3.4 sea surface temperature time series for ERA40/Int (red dots), the ensemble range (green box-and-whisker plots) and the ensemble mean (blue dots) for the one-month lead, 2-4 month re-forecasts for EC-Earth (a) and System 3 (b) with the May start dates. Re-forecasts for the period 1981-2005 have been used. The ratio between the model and reference interannual standard deviation for EC-Earth (System 3) is 1.34 (0.84), while the ensemble mean correlation is 0.82 (0.86).

The similar performance of both systems at the seasonal time scale can also be seen in Figure 4. The ensemble-mean anomaly correlation of precipitation computed over large areas shows very similar scores and uncertainty ranges between the two systems. This conclusion is even more obvious in the summary shown in Figure 5. The proportion of scores above and below the diagonal is very close for both systems, although this depends of the region and the score considered. The scores are close enough for the uncertainty intervals to largely overlap in most cases. This produces a very low number of cases where one system is significantly better than the other. Hence, it can be concluded that both systems have an equivalent forecast quality at the seasonal time scale.

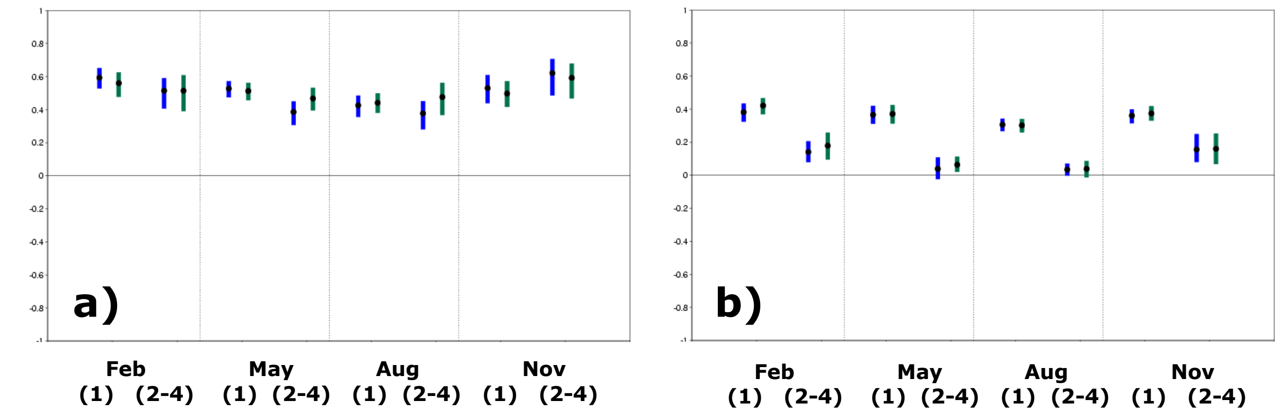

Figure 4: Anomaly correlation coefficient of the 2-4 month ensemble-mean re-forecasts of precipitation for EC-Earth (blue bars) and System 3 (green bars) over a) the tropical band and b) the northern extratropics as a function of the start date and forecast time. The first set of bars in each panel corresponds to the results for the first month (1) and the average 2-4 months (2-4) predictions. The black dot shows the sample value while the bars around it depict the bootstrapped 95% confidence interval. Re-forecasts for the period 1981-2005 have been used. The reference is taken from GPCP.

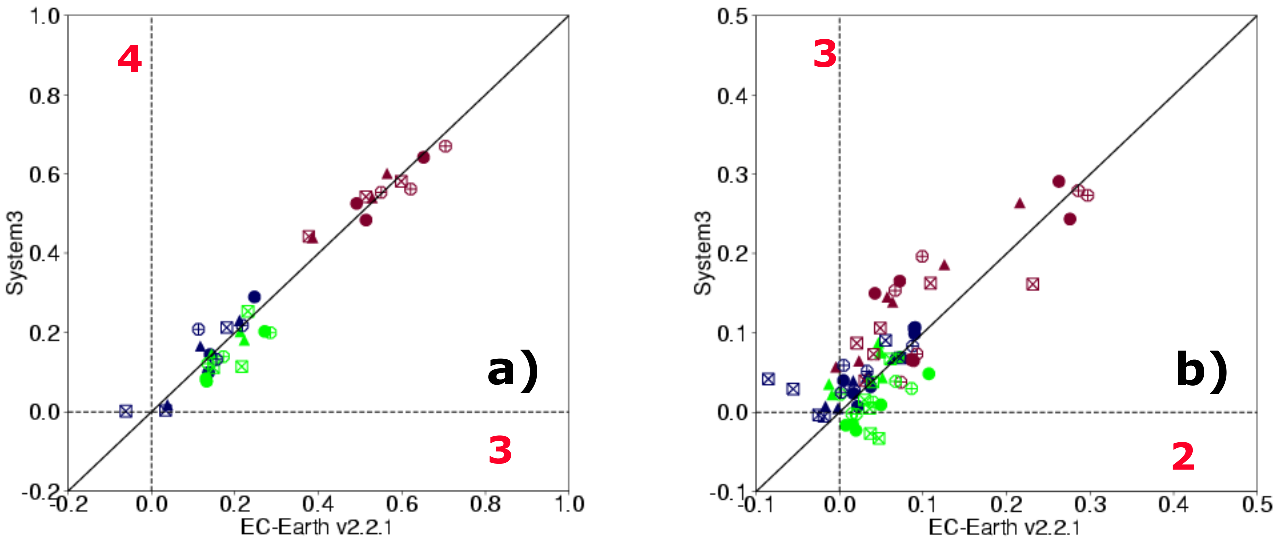

Figure 5: a) Anomaly correlation coefficient and b) Brier skill score for infinite-size ensembles for 2-4 month ensemble-mean re-forecasts of precipitation, two-metre temperature and mean sea level pressure for three regions (Northern Hemisphere in blue, tropical band in brown and Southern Hemisphere in green) and four start dates per year (each one represented with a different symbol) for EC-Earth in abscissa and System 3 in ordinates. The Brier skill score is computed for the events anomalies above (below) the upper (lower) tercile. The inset red numbers correspond to the number of cases where one of the systems is significantly better than the other with 95% confidence (lower right for EC-Earth being better than System 3 and vice verse for upper left). Re-forecasts for the period 1981-2005 have been used. The reference is taken from GPCP and ERA40/Int.

As EC-Earth is a competitive seasonal forecast system, a sensitivity experiment to the sea-ice initialization has been carried out. A set of re-forecasts initialized with climatological sea-ice initial conditions has been performed over the period 1991-2005 using the May and August start dates. The May start date re-forecasts have been run for up to seven months, while the August ones have been run only four months into the future. Figure 6 summarizes some of the results, where the ensemble-mean correlation for the 5-7 month seasonal (boreal autumn) predictions obtained from the May start date are shown. While substantial differences can be found over the Arctic, no major changes can be observed elsewhere. Actually, some climate features show a degraded forecast quality when using realistic sea-ice initial conditions. Figure 6c shows the ensemble-mean correlation for the North Atlantic Oscillation index of the 5-7 month seasonal (September-to-October) predictions from the start date re-forecasts, where the EC-Earth predictions initialized with time-varying sea-ice initial conditions has lower correlation than both System 3 and the EC-Earth experiment initialized with climatological sea-ice. This suggests that a more in-depth analysis of the regional impact of the sea-ice initial conditions is necessary.

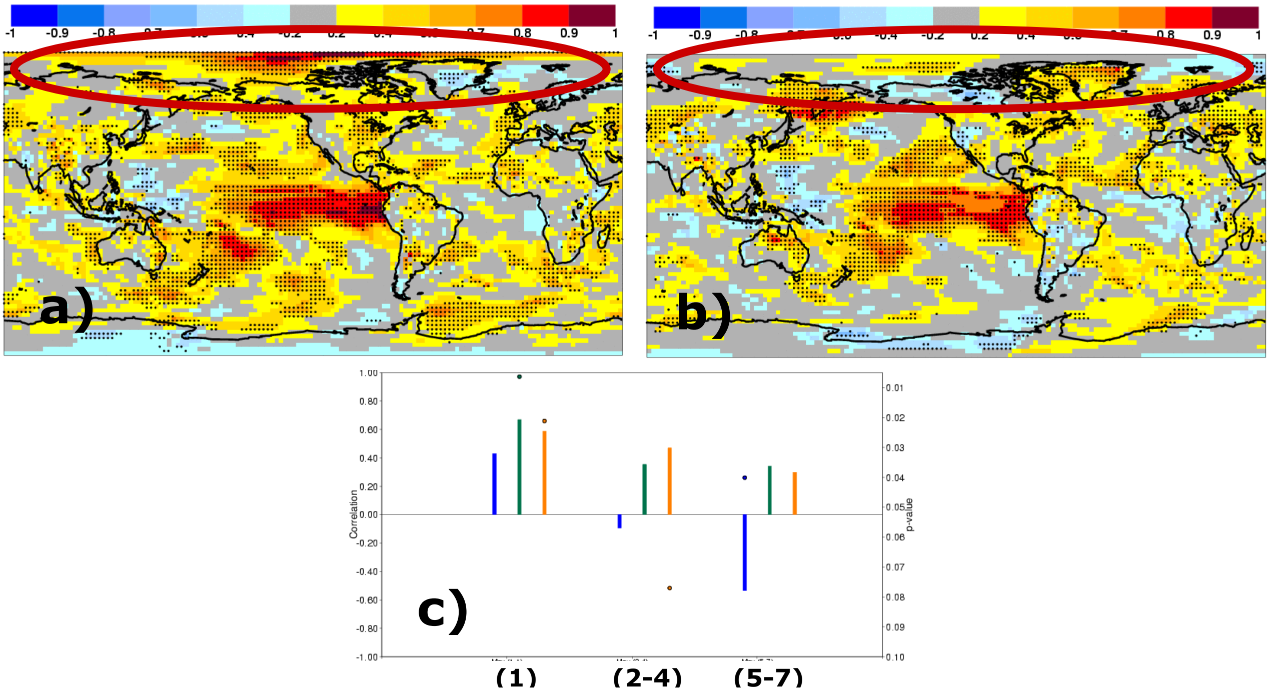

Figure 6: Ensemble-mean correlation for 5-7 month seasonal re-forecasts of two-metre temperature from the May start date of the EC-Earth experiments initialized with 1) realistic and b) climatological sea-ice initial conditions. Dotted areas have correlations statistically significant with 95% confidence. c) Ensemble-mean correlation for the mean boreal autumn (September-to-November) NAO index for EC-Earth (blue), System 3 (green) and EC-Earth with climatological sea-ice initial conditions (orange) as a function of the forecast time in months. The right axis marks the p-value of the correlation, which is shown by the dots on the vertical of the coloured bars. When no dot is shown, the p-value is larger than 0.10. Re-forecasts for the period 1991-2005 have been used. The reference is taken from ERA40/Int.

As the objective of this special project is to explore the evolution of the impact of the initialization along the forecast time, the seasonal forecast experiments mentioned above were extended. The five-member ensemble re-forecasts for the May and November start dates were expanded up to 13 forecast months, over the period 1976-2005. The results are compared to those obtained from the five-member System 3 annual outlook re-forecasts. The forecasts drifted from the initial conditions towards a cooler climate, along the lines of what has been shown in Figure 1. The warm bias in boreal winter over the Northern Hemisphere becomes worse though. The drift compares well with what is found in System 3, although the System 3 re-forecasts show an overall warmer tropical band that enhances the strong warm bias in the eastern part of the tropical oceans, typical of boreal summer. Over the southern circumpolar region the biases are very similar in both systems, slightly cooler in austral winter than in summer, a period when both systems have a consistent warm bias south of 40ºS. However, a large difference between both systems is the overestimation of the temperature variability in that area in EC-Earth, while System 3 strongly underestimates it. This is a consequence of the lack of dynamical sea-ice model in System 3, while the overestimation of the EC-Earth variability has also been found in historical long-term integrations.
ENSO is known to have skill beyond the first few forecast months (Doblas-Reyes et al., 2009). This is also applicable to the EC-Earth forecasts, as it is illustrated in Figure 7. The Niño3.4 SST probability forecasts are skilful in terms of both accuracy (measured by the ROC skill score) and reliability (measured by the reliability skill score). This is more obvious for the event “anomalies below the lower tercile” than for the event “anomalies above the upper tercile”, which suggests a certain asymmetry in the forecast quality of the probability forecasts, probably related to the well known ENSO skewness. System 3 is slightly more skilful, especially for the boreal summer and autumn predictions obtained from the May start date. The systems are much more similar in terms of reliability. Furthermore, the confidence intervals for the reliability are much larger than for the ROC, so that it is likely that any substantial difference will not be statistically significant.

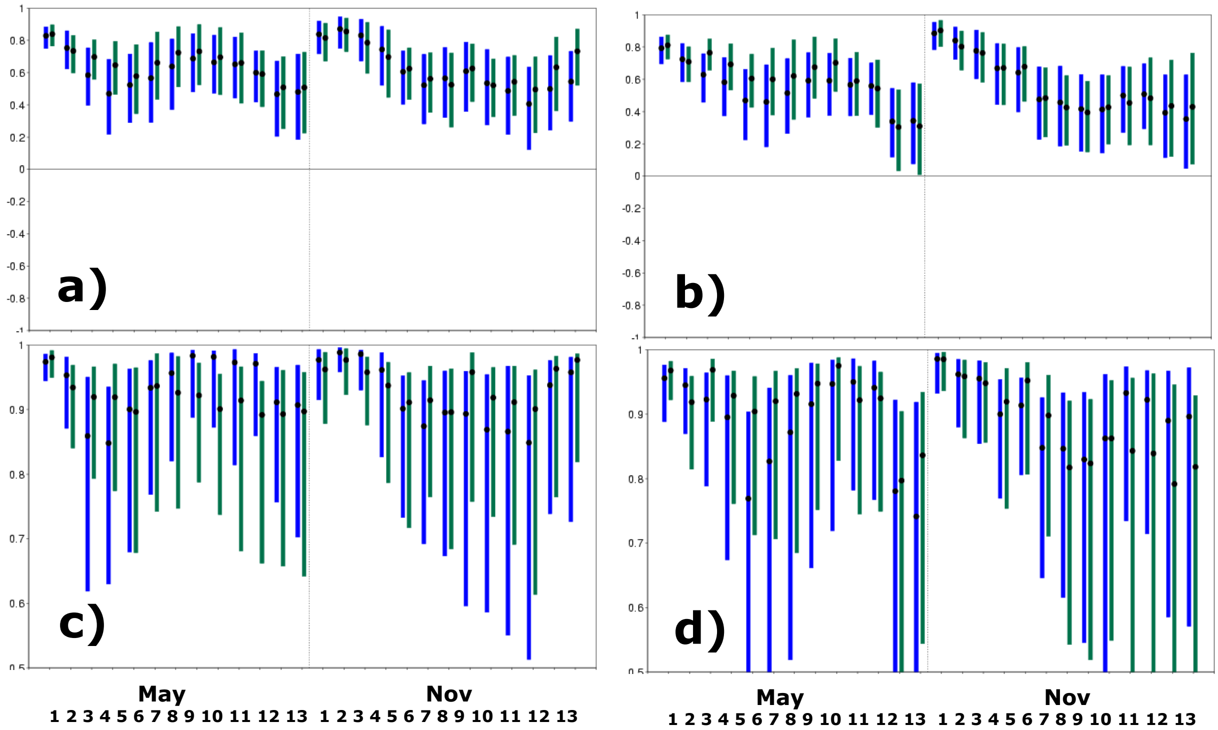

Figure 7: Niño3.4 SST index ROC (panels a and b) and reliability (panels b and d) skill scores of the EC-Earth (blue) and System 3 (green) re-forecasts for the event "anomalies below the lower tercile" (panels a and c) and "anomalies above the upper tercile" (panels b and d) as a function of the start date and forecast time. The two sets of 13 pairs of bars in each panel correspond to the results for the 13 forecast months of the May and November start dates. The black dot shows the sample value while the bars around it depict the bootstrapped 95% confidence interval. Re-forecasts for the period 1976-2005 have been used. The reference is taken from ERA-40/Int.

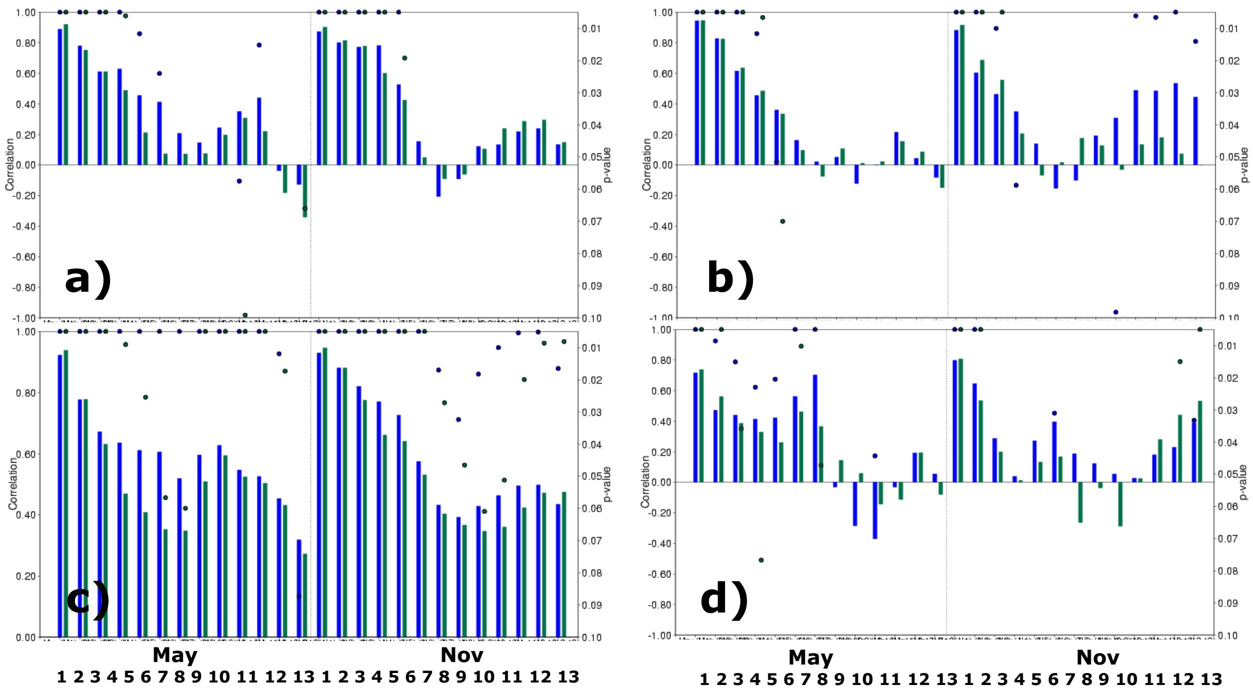

Figure 8: Ensemble-mean correlation for the a) Atlantic3, b) tropical Atlantic meridional mode, c) average tropical Atlantic and d) Indian ocean dipole SST indices of the EC-Earth (blue) and System 3 (green) re-forecasts as a function of the start date and forecast time. The two sets of 13 pairs of bars in each panel correspond to the results for the 13 forecast months of the May and November start dates. The right axis marks the p-value of the correlation, which is shown by the dots on the vertical of the coloured bars. When no dot is shown, the p-value is larger than 0.10. Re-forecasts for the period 1976-2005 have been used. The reference is taken from ERA-40/Int.

As further examples of the skill of the EC-Earth forecast system to predict tropical SSTs, Figure 8 shows the ensemble-mean correlation for Atlantic and Indian Ocean indices. Both System 3 and EC-Earth display an interesting ability to predict the tropical Atlantic average SSTs. EC-Earth shows a slightly higher correlation for the boreal summer and autumn in the Atlantic3 and meridional mode indices. The Indian Ocean dipole has its maximum skill typically in boreal autumn, an aspect that can be seen in Figure 8, with a similar performance in both forecast systems. The forecast quality of the two forecast systems is compared with the skill of a five-member ensemble EC-Earth experiment where the atmospheric component employed prescribed observed SSTs, which to a certain measure gives an upper estimate of the skill, although this interpretation is limited by the lack of ocean-atmosphere coupling in such experiment. Except in a few cases, both System 3 and EC-Earth have a similar forecast quality. The prescribed-SST experiment shows higher correlation for long forecast times and the annual mean over the tropics, a consequence of the reduced skill in predicting SSTs by the coupled forecast systems. By contrast, the prescribed-SST experiment displays a similar correlation over the extratropics, which suggests that improving the forecast systems does not necessarily rely on better predicting the local SSTs but rather on increasing the realism of the coupled processes and of the atmospheric component itself.

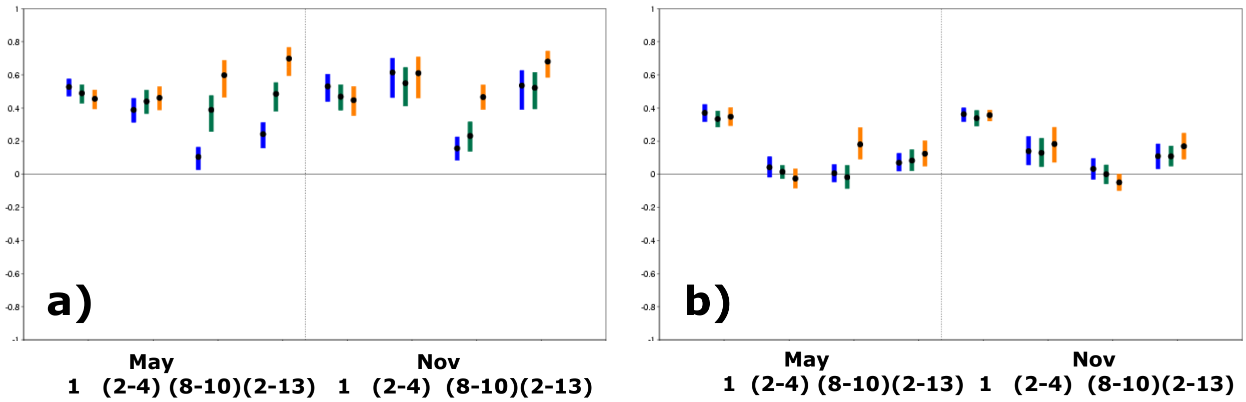

Figure 9: Anomaly correlation coefficient of the ensemble-mean re-forecasts of precipitation for EC-Earth (blue bars), System 3 (green bars) and EC-Earth forced by observed SSTs (orange bars) over a) the tropical band and b) the northern extratropics as a function of the start date and forecast time. The first set of bars in each panel corresponds to the results for the first month (1), the average of the 2-4 , 8-10 and 2-13 months predictions. The black dot shows the sample value while the bars around it depict the bootstrapped 95% confidence interval. Re-forecasts for the period 1976-2005 have been used. The reference is taken from GPCP.

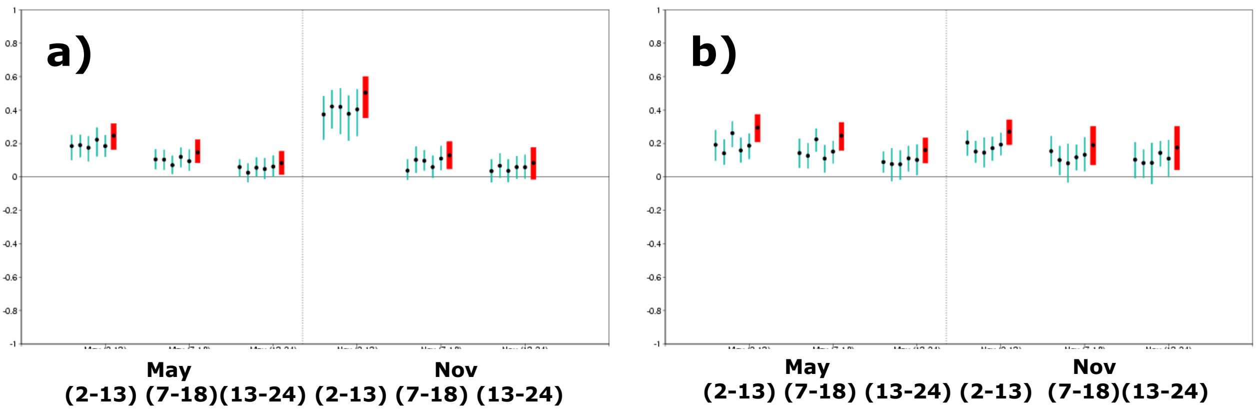

Figure 10: Anomaly correlation coefficient of the ensemble-mean re-forecasts of a) tropical precipitation and b) Northern Hemisphere two-metre temperature for the individual members (light blue bars) and ensemble-mean (red bars) EC-Earth predictions as a function of the start date and forecast time. The first set of bars in each panel corresponds to the results for the first annual mean prediction discounting the initial month (2-13), and the annual means for the 7-18 and 13-24 month predictions. The black dot shows the sample value while the bars around it depict the bootstrapped 95% confidence interval. Re-forecasts for the period 1976-2005 have been used. The reference is taken from GPCP and ERA-40/Int.

The impact of the initialization on forecast quality has been explored further by extending the five-member EC-Earth re-forecasts to two years. In this case, no comparison with System 3 is possible because the longest forecast time available in that system is 13 months. Figure 10 shows the anomaly correlation for each ensemble member and for the ensemble mean for different annual averages of tropical precipitation and Northern Hemisphere two-metre temperature. The tropical precipitation skill is always positive and, as expected, higher for the ensemble mean than for the individual members. It is interesting that the first-year skill is higher for the November than for the May start date. This is a consequence of a better prediction, due mostly to higher SST persistence, of the boreal winter precipitation when the predictions are close to the start date. The Northern Hemisphere temperature shows a more stable level of skill, less dependent of the lead time used to formulate the predictions. As for the tropical precipitation, the ensemble-mean skill is in most cases statistically significantly different from zero and higher than for the individual members. It should be mentioned that the skill for the individual members is very similar, which is a pre-requisite for a well-generated ensemble.

The performance of EC-Earth for such long forecast times is illustrated in Figure 11 in comparison with the single-model systems contributing to the ENSEMBLES multi-model and the perturbed parameter DePreSys system (Doblas-Reyes et al., 2011). While for tropical temperatures successive averaging forecast times show most models close to the same correlation, the temperature predictions over the Northern Hemisphere depict DePreSys and the ENSEMBLES multi-model above the single-model systems and EC-Earth close to the worse one. It has been recently discovered that the version 2.2 of EC-Earth used in these experiments treated the natural and anthropogenic aerosols incorrectly. This makes difficult to interpret some of the results presented here as representative of the actual potential of EC-Earth as a climate forecast system, although they already give a broad idea of how promising this system is a research tool to explore the limits and help in the development of a new generation of forecast systems.

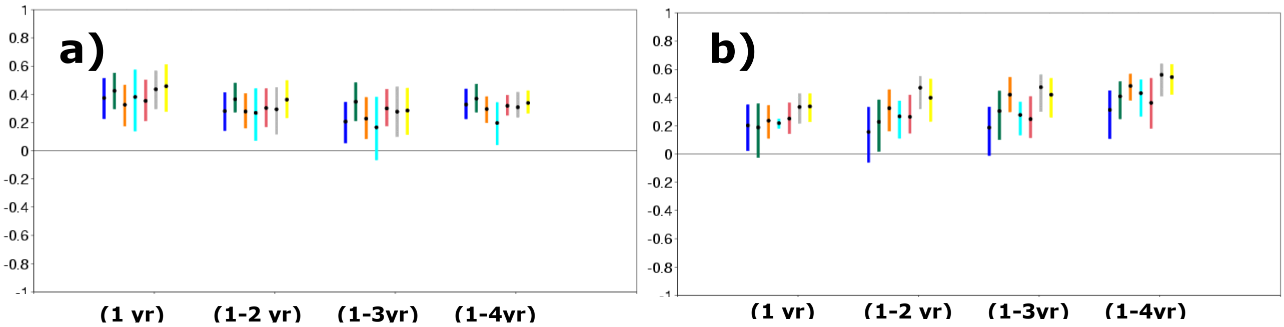

Figure 11: Ensemble-mean anomaly correlation coefficient of the ensemble-mean re-forecasts of a) tropical and b) Northern Hemisphere near-surface temperature for the predictions of EC-Earth (blue bars), the contributors to the ENSEMBLES multi-model (green, orange, cyan and pink bars), DePreSys_PP (grey bars) and the multi-model (yellow bars) as a function of the forecast time. The first set of bars in each panel corresponds to the results for the first annual mean prediction discounting the initial month (1), and the average predictions of the 1 and 2, 1-to-3 and 1-to-4 forecast years. The black dot shows the sample value while the bars around it depict the bootstrapped 95% confidence interval. Re-forecasts for the period 1960-2005 with five-year intervals between start dates have been used. The reference is taken from ERA-40/Int.

References

Balmaseda, M.A., K. Mogensen, F. Moteni and A.T Weaver, 2010. The NEMOVAR-COMBINE ocean re-analysis. NEMOVAR Technical reports No. 1, available from http://www.combine-project.eu/Technical-Reports.1668.0.html.

Brodeau L. B. Barnier, A.-M. Treguier, T. Penduff and S. Gulev, 2009. An ERA40-based atmospheric forcing for global ocean circulation models. Ocean Modelling, doi:10.1016/j.ocemod.2009.10.005 Key: citeulike:6067348.

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. Quart. J. Roy. Meteorol. Soc., 135, 1538-1559, doi:10.1002/qj.464.

Doblas-Reyes, F.J., G.J. van Oldenborgh, J. García-Serrano, H. Pohlmann, A.A. Scaife and D. Smith, 2011. CMIP5 near-term climate prediction. CLIVAR Exchanges, in press.

Dutra, E., G. Balsamo, P. Viterbo, P.M.A. Miranda, A. Beljaars, C. Schär and K. Elder, 2010. An improved snow scheme for the ECMWF land surface model: description and offline validation. J. Hydrometeorol, 11, 899-916. doi: 10.1175/2010JHM1249.1.

Ingleby, B. and M. Huddleston, 2007. Quality control of ocean temperature and salinity profiles - historical and real-time data. Journal of Marine Systems, 65, 158-175, doi:10.1016/j.jmarsys.2005.11.019.

Magnusson, L., M. Leutbecher and E. Kallen, 2008. Comparison between singular vectors and breeding vectors as initial perturbations for the ECMWF Ensemble Prediction System. Mon. Weather Rev., 134, 4092-4104.

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., doi: 10.1007/s00382-010-0947-3.
Uppala, S. M., and 45 others, 2005. The ERA-40 reanalysis. Quart. J. Roy. Meteorol. Soc., 131, 2961-3012.
Weisheimer, A., F.J. Doblas-Reyes, T. Jung and T.N. Palmer, 2011. On the predictability of the extreme summer 2003 over Europe. Geophys. Res. Letters, 38, L05704, doi:10.1029/2010GL046455.

List of publications/reports from the project with complete references

Doblas-Reyes, F.J., 2011. Seasonal and interannual forecasting with the EC-Earth model. Climate Dyn., submitted.
Du, H., F.J. Doblas-Reyes, J. García-Serrano, V. Guémas, Y. Soufflet and B. Wouters, 2011. Impact of initial perturbations on decadal predictions. Climate Dyn., submitted.
Doblas-Reyes, F.J., G.J. van Oldenborgh, J. García-Serrano, H. Pohlmann, A.A. Scaife and D. Smith, 2011. CMIP5 near-term climate prediction. CLIVAR Exchanges, in press.

Summary of plans for the continuation of the project

(10 lines max)

An updated EC-Earth version with corrections in the natural and anthropogenic aerosol treatment, and some other minor bug corrections, has been recently released to the consortium. Some of the experiments described in this report will be repeated to assess the sensitivity to this correction. Although it is likely that some results will change, it is expected that many of the conclusions will still hold. As a consequence of having to repeat a substantial amount of the experiment, it is likely that we will have to decline performing the five-member ensemble continuous simulations using prescribed SSTs over the period 1975 to 2009 initialized from the "historical" EC-Earth simulation carried out for CMIP5, along with the re-forecasts initialized from the restarts of these simulations.
New sets of sea-ice restart files will be available in the coming weeks and some sensitivity tests to the sea-ice initialization will be carried out.