

# 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** 2014

**Project Title:** Seasonal climate forecast quality with EC-Earth: role of initialization and stochastic physics

**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 2012

**Expected end date:** December 2014

**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
<b>High Performance Computing Facility</b>	(units)	3,960,000 + 1,500,000 extension	approx. 5,000,000	5,880,000	5,600,000
<b>Data storage capacity</b>	(Gbytes)	5,060 + 2,000	4,510	5,060	3,748

A system to automatically download data to IC3 was implemented in 2013 to limit the use of data storage capacity at ECMWF. The 4.5 Tb of storage capacity used in 2013 contains mainly the reference files to perform atmospheric nudging with EC-Earth and are made available to the whole EC-Earth consortium. Part of the data storage capacity currently in use for this year will be transferred back to IC3.

## Summary of project objectives

(10 lines max)

The goal of the project consists in analyzing the impact of improved sea-ice and soil-moisture initial conditions, as well as of the implementation of a simplified stochastic physics scheme, on the forecast quality of seasonal predictions performed with the EC-Earth forecast system. EC-Earth is used as a test-bed to transfer improvements to the ECMWF operational seasonal forecast system coming from a full climate model. The project considers a comprehensive set of ensemble interannual hindcasts with different initialization strategies. By developing the sub-seasonal and seasonal forecast capability of EC-Earth, this project expects to continue implementing the seamless approach, whose basic premise is that there are fundamental physical processes in common across the climate forecast time scales and with the climate-change projections.

## Summary of problems encountered (if any)

(20 lines max)

None.

## 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

Seasonal-to-interannual (s2i) predictions deal with a time horizon ranging from a season to several years. On these time scales, the storage of heat by the ocean and moisture by the land, together with the presence or absence of snow and sea ice become important factors in determining the atmospheric variability. Based on knowledge of the initial conditions, important aspects of climate are predictable up to a year ahead (Kirtman and Pirani, 2009). This predictability is primarily, though not solely, associated with the El Niño Southern Oscillation (ENSO). Besides, the natural variability of temperature and other climate variables at the s2i time scale should be considered as superimposed on externally forced low-frequency variability due to external forcings: human-induced changes in greenhouse gas and aerosol (GHGA) concentrations, land-use changes as well as natural variations in solar activity and volcanic eruptions.

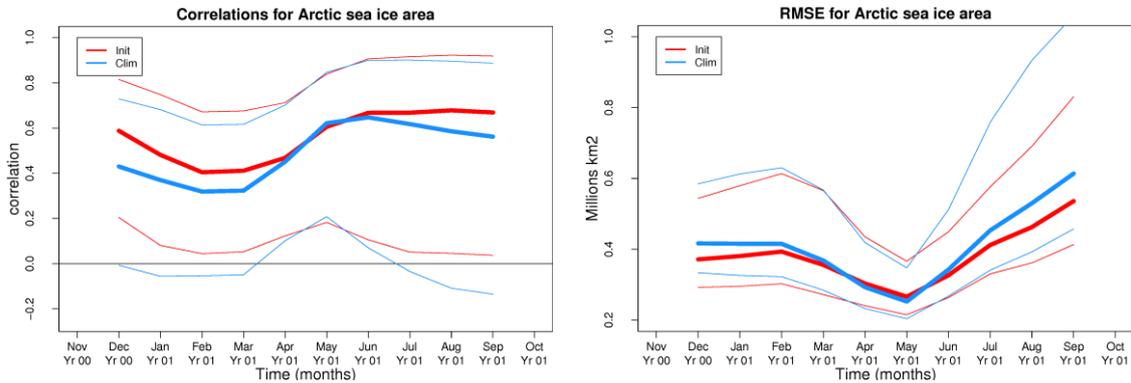
The ocean anomalies associated with ENSO events and other ocean phenomena, soil moisture and snow and sea-ice cover anomalies are taken into account when initializing dynamical s2i predictions. Unfortunately, less information is available about the state of the ocean, sea-ice, snow and land than about the atmosphere (Saha et al., 2010), and often the forecasts are penalized by a lack of understanding of the interactions among the subsystems (Pegion and Kirtman, 2008). Initial-condition and GHGA concentration uncertainties are not the only sources of forecast error. Model inadequacy, a result of the lack of knowledge about relevant processes and the limitations in available computational resources (Palmer et al., 2005), also limits the ability to make predictions on time scales longer than two weeks.

The goal of this project is threefold: to assess the impact of sea-ice and soil initialisation on s2i predictions, as well as to test stochastic physics techniques to account for model inadequacies at the seasonal time scale, using the EC-Earth Earth system model (ESM; Hazeleger et al., 2012). The following paragraphs present results about the sensitivity of s2i forecast quality to the sea-ice and soil initialisation datasets and methods and to the stochastic physics schemes tested in the last 12 months.

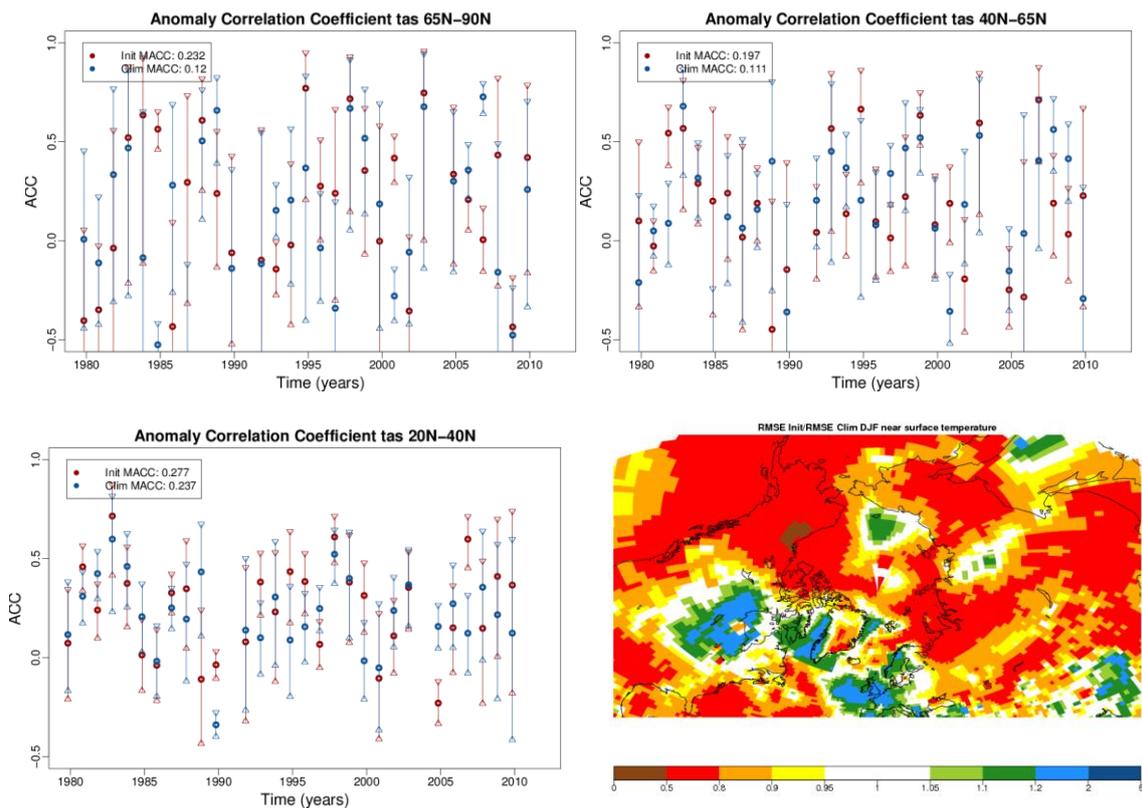
## Impact of sea-ice initialisation

The tuning of EC-Earth3 by the EC-Earth consortium lasted longer than initially planned. To be able to satisfy our commitment in this proposal we chose to perform our sea-ice and land-surface initialisation experiments with EC-Earth2.3 instead. Ensembles of 5-member, 1-year long climate predictions have been initialized on 1<sup>st</sup> November, 1<sup>st</sup> May and 1<sup>st</sup> August every year from 1979 to 2011 with EC-Earth 2.3. The atmospheric initial conditions were taken from the ERA-interim reanalysis (Dee et al., 2011), the ocean

initial conditions from the ORAS4 reanalysis (Mogensen et al., 2011; Balmaseda et al., 2012) and the sea ice initial conditions from a sea ice reconstruction (Guemas et al., 2014) produced by forcing the ocean and sea ice components of EC-Earth 2.3 with ERA-interim and nudging the ocean thermodynamic state towards ORAS4. This set of climate predictions will be referred to as Init in the following.



**Figure 1:** Correlation (left) and root mean square error (RMSE, right) of the ensemble mean predictions initialised in November for the total Arctic sea-ice area as a function of the forecast time computed after applying a bias correction via the per-pair method (Garcia-Serrano and Doblas-Reyes, 2012) and after a smoothing with a 3-month running mean. The reference dataset is the National Snow and Ice Data Center (Cavalieri et al., 1996). The confidence intervals shown by thin lines are computed using a Student  $t$  distribution for the correlation and a  $\chi^2$  distribution for the RMSE. The autocorrelation of the data is accounted for following Von Storch and Zwiers (2001).

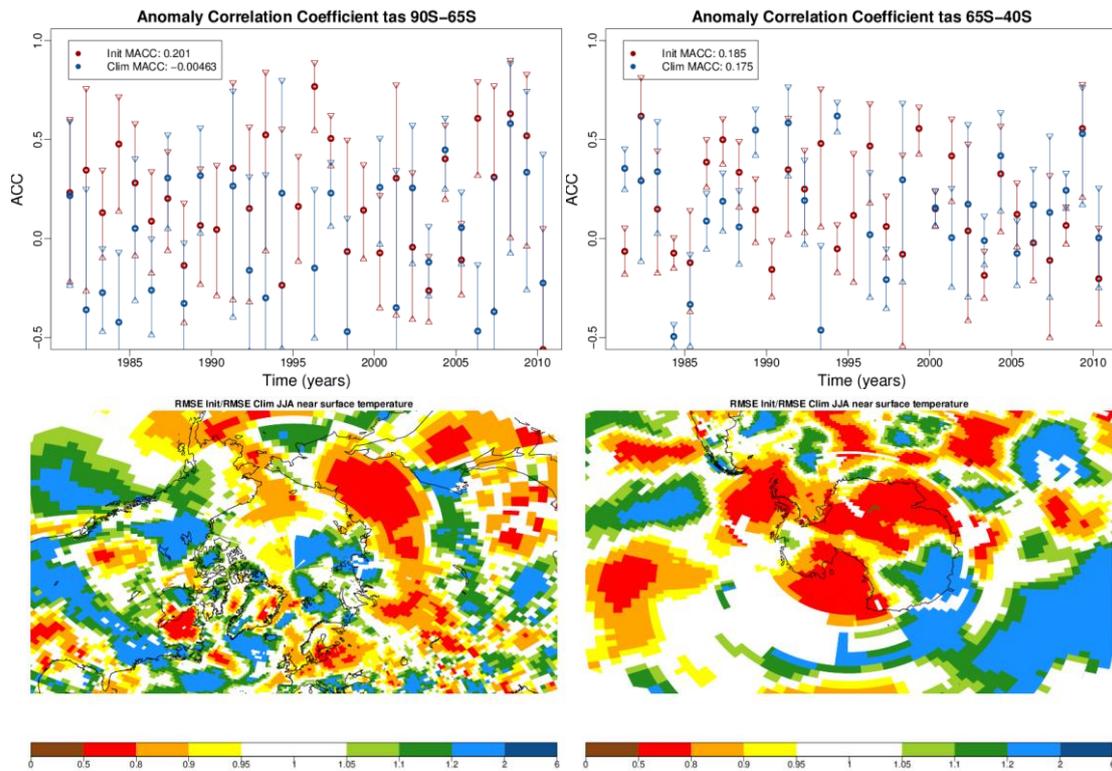


**Figure 2:** Anomaly correlation coefficient (ACC) of the ensemble mean predictions initialised in November for winter (December-January) bias-corrected near-surface air temperature north of  $65^{\circ}\text{N}$  (top left), in the  $40^{\circ}\text{N}$ - $65^{\circ}\text{N}$  band (top right) and in the  $20^{\circ}\text{N}$ - $40^{\circ}\text{N}$  band (bottom left). The ratio between the RMSE of the Init and Clim experiments is shown in the bottom right panel. The reference data comes from the merged dataset containing the GHCN/CAMS land surface temperature (Fan and van den Dool, 2008) and the ERSST v3b sea surface temperature (Smith et al., 2008) inside the  $60^{\circ}\text{S}$ - $60^{\circ}\text{N}$  latitude band and the GISSTEMP with 1200 km decorrelation scale (Hansen et al., 2010) in the polar regions. The confidence intervals for the ACC are computed using a Student  $t$  distribution.

Three sensitivity experiments, referred to as Clim in the following, have been performed by initialising the sea-ice component from climatologies of the 1st November, 1st May and 1st August respectively and computed over the 1981-2010 period from the same sea-ice reconstruction mentioned above. Given the strong seasonality of Arctic sea ice, the results are described separately for each start date but for reasons of space only for the hindcasts initialised in May and November.

### Forecasts initialised in November

The Init experiment shows a higher skill in total Arctic sea ice area than Clim in winter and summer but not in spring (Figure 1). Initializing from a sea ice reconstruction allows for an increase in mean anomaly correlation coefficient (MACC) of Arctic winter (December-January) near surface air temperature from 0.12 to 0.23 North of 65°N, from 0.11 to 0.20 in the 40°N-65°N latitude band and from 0.24 to 0.28 in the 20°N to 40°N latitude band (Figure 2) compared to initializing from a sea-ice climatology. The root mean square error (RMSE) in winter near-surface temperature is significantly reduced in the Central Arctic, over the Barents, Kara, Laptev, Beaufort, Bering Seas and the Gulf of Alaska but also over continental regions such as Alaska and Northern Asia (Figure 2). No improvement is found in the Tropics nor in the Southern Hemisphere.



**Figure 3:** As in panel bottom right of Figure 2 but for seasonal forecasts initialised in 1<sup>st</sup> May and for summer (June-August) near-surface temperature. Latitude bands are 90°S-65°S (top left) and 65°S-40°S (top right). Areas are the North Pole (bottom left) and South Pole (bottom right).

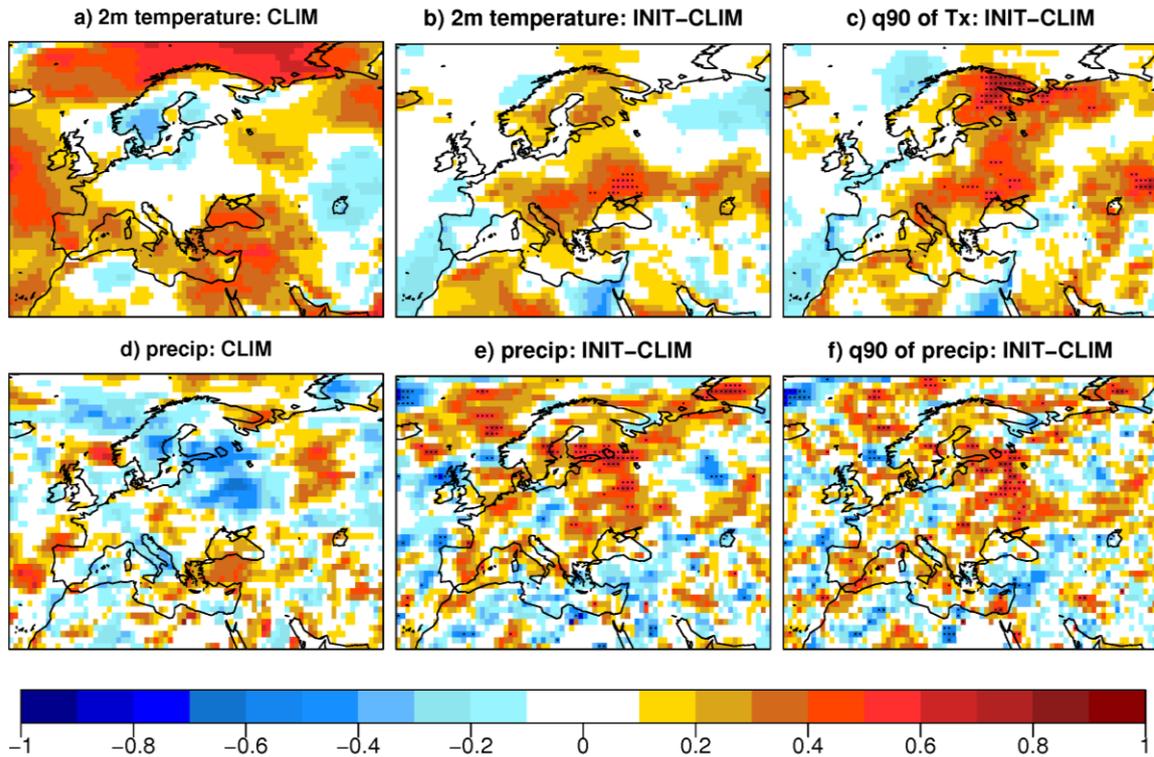
### Forecasts initialised in May

The RMSE in summer (June-to-August) for near-surface air temperature is significantly reduced only over north-eastern Asia (Figure 3, bottom left) in Init relative to Clim. However, whereas no significant impact of sea-ice initialization was obtained in the Southern Hemisphere for forecasts from 1<sup>st</sup> November, a significant reduction in the RMSE in summer near surface temperature over Antarctica is obtained in the forecasts from 1<sup>st</sup> May (Figure 3, bottom right). The MACC increases from 0 to 0.20 South of 65°S and from 0.17 to 0.18 in the 65°S-40°S latitude band (Figure 3, top row) in Init compared to Clim.

These experiments tend to illustrate the added-value for near-surface temperature skill of initializing the sea-ice component of seasonal forecast systems even from a sea-ice reconstruction instead of using sea-ice climatologies at least for the first few forecast months. The data of the sea-ice reconstruction has been shared with the seasonal forecast group at ECMWF.

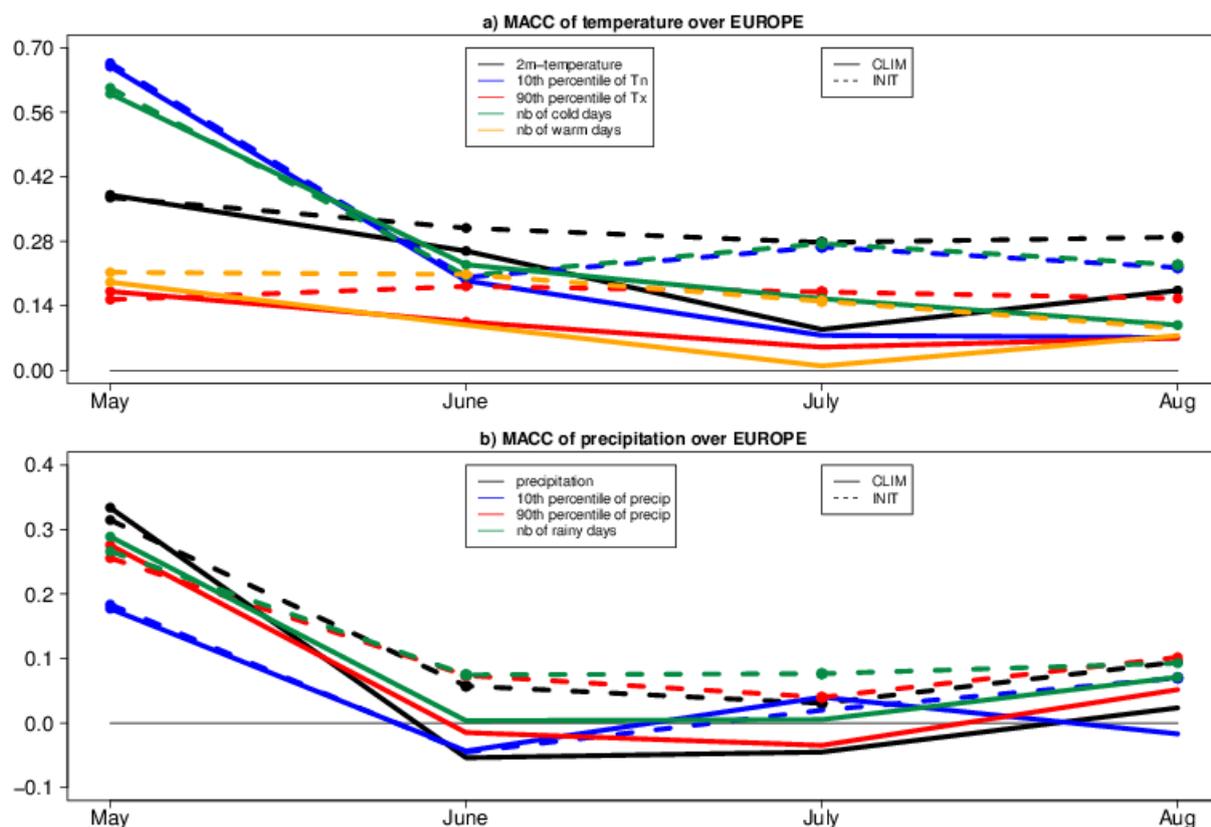
## Impact of land-surface initialisation

The impact of a realistic land-surface initialization on sub-seasonal and seasonal forecasts was also assessed with EC-Earth2.3, by comparing two sets of 10-member four-month long hindcasts over the 1981-2010 period starting the predictions on the 1<sup>st</sup> of May, June, July and August. For simplicity, only results for the May start dates are discussed here. In the first set, the Clim experiment, the land surface is initialised with climatological conditions taken from ERAInt-Land (Balsamo et al., 2013), while in the second set, the Init experiment, the soil moisture is initialized with the simultaneous ERAInt-Land values. The ocean, sea-ice and atmospheric components are initialized with ORAS4, a sea-ice reconstruction and ERA-Interim (Dee et al., 2011), respectively. However, the sea-ice initial conditions of these experiments are different from the ones used in the previous section as the IC3 sea-ice reconstruction was not available until the land-surface experiments were already started. Seasonal forecast quality is assessed using ERA-Interim as a reference for temperature and GPCP for precipitation (Huffman et al., 2009).



**Figure 4:** a) Correlation of the mean near-surface air temperature in boreal summer (JJA) in the Clim experiment. The correlation is calculated after linearly detrending the ensemble-mean averaged JJA values of the Clim experiment and the observations over the 1981-2010 period. The dots indicate the areas where the correlation is significant at the 95% level based on a Student *t*-test. b) Difference in correlation of the JJA near-surface air temperature in JJA between Init and Clim. The dots indicate the areas where the correlation difference is significant at the 95% level, according to the Olkin and Finn (1995) methodology. c) Same as b) but for the 90<sup>th</sup> percentile of maximum daily temperature of each single season. d) Same as a) for precipitation. e) Same as b) for precipitation. f) Same as b) but for the 90<sup>th</sup> percentile of the daily precipitation.

Figure 4 shows the correlation of the boreal summer near-surface air temperature and precipitation in the Clim experiment initialised in May, as well as differences in correlation found with a realistic soil moisture initialisation in Init. Substantial improvements are found for both variables as well as for the 90<sup>th</sup> percentile of their daily values, especially over Eastern Europe.



**Figure 5:** a) Mean anomaly correlation coefficient calculated between the ensemble-mean prediction over land in Europe ( $10^{\circ}\text{W}$ - $40^{\circ}\text{E}$ ;  $35^{\circ}\text{N}$ - $75^{\circ}\text{N}$ ) and observations for the Clim (solid lines) and Init (dotted lines) experiments for the monthly mean near-surface air temperature (black), the monthly 90<sup>th</sup> percentile of maximum daily temperature ( $T_x$ , red), the monthly 10<sup>th</sup> percentile of the minimum daily temperature ( $T_n$ , blue), the number of days when the maximum daily temperature is above the climatological 90<sup>th</sup> percentile (orange) and the number of days when the minimum daily temperature is under the climatological 10<sup>th</sup> percentile (green). The dots mark the values significant at a 95% level, estimated with a spatial block bootstrap of 1000 samples. b) Same as a) but for monthly mean precipitation (black), the monthly 90<sup>th</sup> percentile of daily precipitation (red), the monthly 10<sup>th</sup> percentile of daily precipitation (blue), and the number of days when the daily precipitation is above the climatological 90<sup>th</sup> percentile (green).

Figure 5 shows the mean anomaly correlation coefficient over European land grid points for surface temperature and precipitation as a function of the forecast time. Skill improvements with a realistic initialization of soil moisture are found throughout the season for both variables as well as the different extreme indices considered, even at forecast months 3 and 4. To our knowledge, it is the first time this kind of consistent improvements has been found over Europe.

### Impact of a stochastic parameterisation scheme

Two types of stochastic parameterisations have been considered in this project: stochastic physics and stochastic dynamics. The stochastically perturbed parameterisation tendency scheme (SPPT; Palmer et al., 2009) was implemented in EC-Earth3 because it was part of IFS Cy36r4 that was used for the most recent version of EC-Earth. This also meant that this part of the project had to be carried out last. This multiplicative stochastic parameterisation scheme applies in-run univariate Gaussian perturbations to the wind, humidity and temperature tendencies. The perturbation pattern varies smoothly in time and space, and the same pattern is applied to the tendency of each variable. Several patterns defined with different space and time decorrelation scales can be combined linearly. The importance of each scale in the total perturbation is defined by choosing the standard deviation  $\sigma$ . The stochastic dynamics (SDYN) method (Batté and Déqué, 2012) is an additive stochastic perturbation technique, first implemented in the ARPEGE-Climat atmospheric component of the CNRM-CM climate model. We implemented a technique similar to Batté and Déqué (2012) in IFS Cy36r4. Prognostic variables  $T$ ,  $u$ ,  $v$ , and  $q$  are perturbed by adding random draws of initial tendency error corrections of IFS, estimated using atmospheric relaxation. To our knowledge, this is the first time that two completely independent stochastic approaches are tested simultaneously in IFS and we expect these results to be relevant to ECMWF's activities for the introduction of stochastic parameterisations.

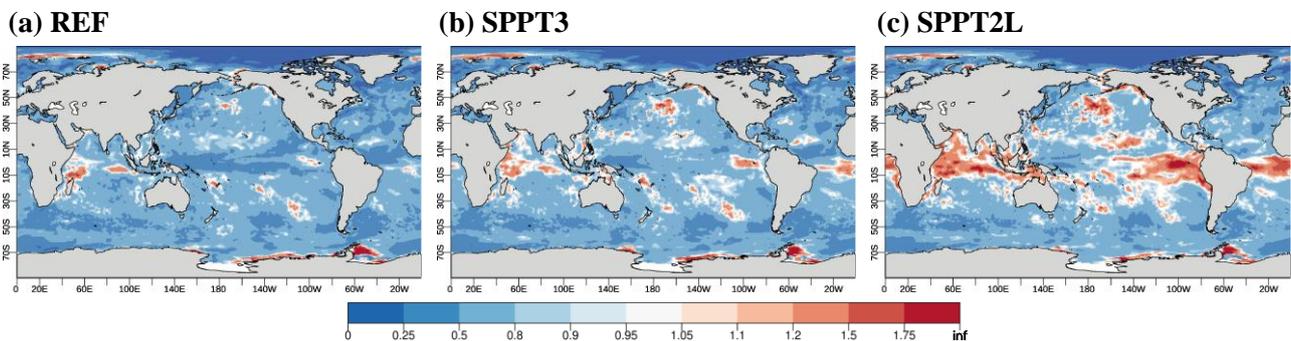
## Impact of SPPT perturbations

Different sets of space and time scales and weights have been tested with EC-Earth3 using the T255L91-NEMO 3.3 ORCA1L46-LIM2 configuration. The experiments are described in Table 1. For each experiment, ten ensemble members were run for four months starting from the first of May and November over the 1993-2009 hindcast period, using the ERA-Interim and GLORYS v2.1 reanalyses to initialise the atmospheric and ocean/sea-ice components.

Exp	Scales	Time scale	Spatial scale	$\sigma$	Time scale	Spatial scale	$\sigma$	Time scale	Spatial scale	$\sigma$
REF	0	-	-	-	-	-	-	-	-	-
SPPT3	3	6 hours	500 km	0.5	3 days	1000 km	0.25	30 days	2000 km	0.125
SPPT2L	2	10 days	1000 km	0.173	30 days	2000 km	0.288	-	-	-

**Table 1:** Space and time scales tested for SPPT in EC-Earth3.

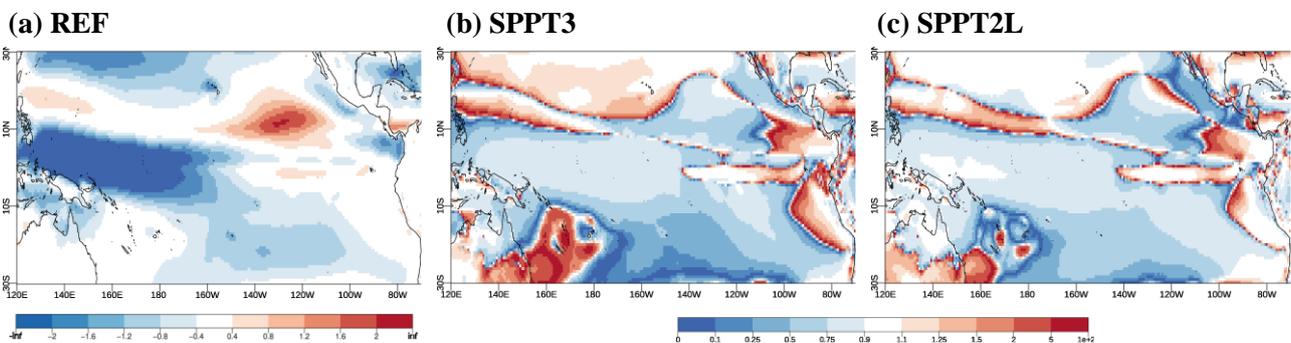
Forecast quality is evaluated in terms of systematic error, spread-skill ratio, correlation of the ensemble mean and probabilistic skill using the Brier Score, focusing on seasonal averages for the forecast months 2 to 4. We use ERAInt as reference data for SST and sea-level pressure, and GPCP v2.2 for precipitation.



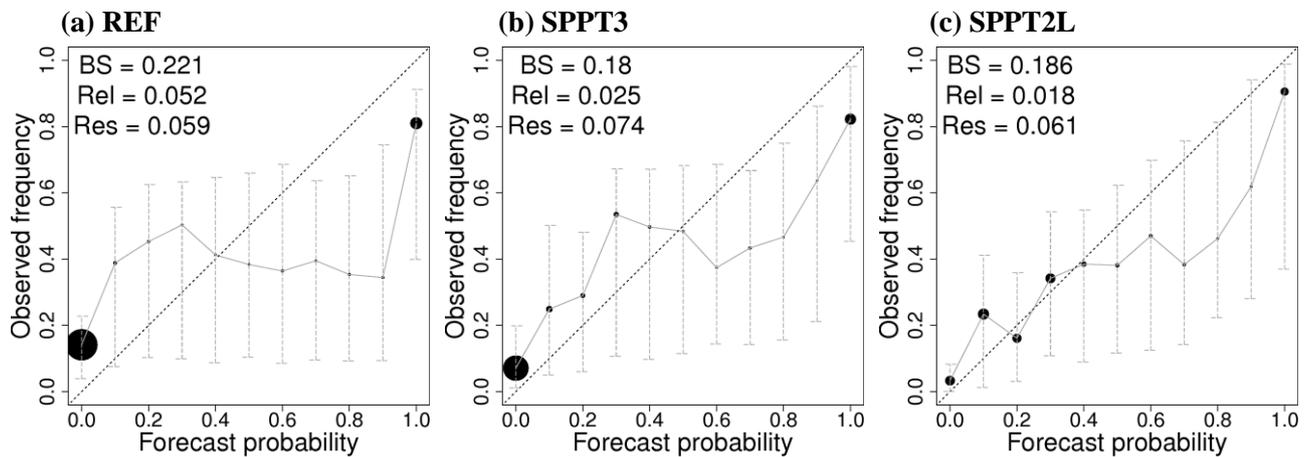
**Figure 6:** Spread-skill ratio for DJF SST over the 1993-2010 period for REF and both SPPT experiments (from left to right). Spread is computed as the standard deviation around the ensemble mean, whereas skill is the RMSE of the hindcast with respect to ERAInt data.

Figure 6 shows the impact of SPPT on the spread-skill ratio for SST hindcasts in DJF using the hindcast with November start dates. The REF ensemble is highly under-dispersive over most of the regions. The introduction of SPPT perturbations increases the spread-skill ratio over most of the tropical oceans, due to a significant increase in ensemble spread without affecting the model RMSE. With the larger scale perturbations (SPPT2L), the ensemble is over-dispersive over the tropical Atlantic and Indian oceans.

In most of our analyses we find that introducing SPPT in EC-Earth3 helps reducing some of the systematic errors over the tropical Pacific, and lead to an improvement of the model mean state in terms of precipitation over the Maritime Continent. These improvements are consistent with a reduction of the near-surface wind biases in DJF over the Western Tropical Pacific, corresponding to a correction of the excessively easterly trade winds over the region (Figure 7).



**Figure 7:** (a) Systematic error for the near-surface zonal-wind component in the DJF REF hindcasts started in November. (b-c) Reduction in absolute systematic error with respect to the REF experiment in the SPPT3 and SPPT2L DJF hindcasts, respectively. Blue regions correspond to the regions where the systematic error is reduced, regardless of its sign.



**Figure 8:** Reliability diagrams for the event JJA SST exceeding the second climatological tercile over the Niño3.4 region for the REF, SPPT3 and SPPT2L experiments started in May. The size of the dots shows the population of each forecast probability bin over the hindcast period, and vertical bars show the range of uncertainty in the reliability diagram based on 1000 bootstrap samples over the hindcast period. The insets provide the Brier score and its decomposition in reliability and resolution.

Figure 8 shows an evaluation of the probabilistic skill of the JJA SST forecasts over the Niño 3.4 region for the hindcasts started in May. The reliability diagrams display the observed relative frequency of SST exceeding the second tercile of the climatology against the forecast probabilities calculated by counting the number of ensemble members forecasting the occurrence of the event. In a well-calibrated ensemble, when the forecast gives a 40% chance of the event occurring, it should be observed about 40% of the time. The points should therefore be close to the diagonal line. The weighted probabilistic bias is measured by the “Rel” value shown in the top left corner of the figures, and should be as small as possible. The diagrams in Figure 8 show that the introduction of SPPT perturbations improves the reliability, especially in the case of SPPT2L. Another aspect of probabilistic forecast skill is the forecast resolution. Climatological forecasts are perfectly reliable, but give no indication as to the inter-annual variability of the probabilistic event. Resolution evaluates the ability of the system to separate different predictions, and is measured by the “Res” value which should be as large as possible. The Brier Score combines both aspects of probabilistic skill and measures a mean square skill score in probabilistic space with respect to using climatology. It should be as small as possible. SPPT perturbations tend to slightly improve the Brier score, mainly due to significant improvements in the Rel value. In contrast, over higher latitudes SPPT tends to increase model systematic error without significantly changing seasonal forecast skill.

#### Impact of a stochastic dynamics technique

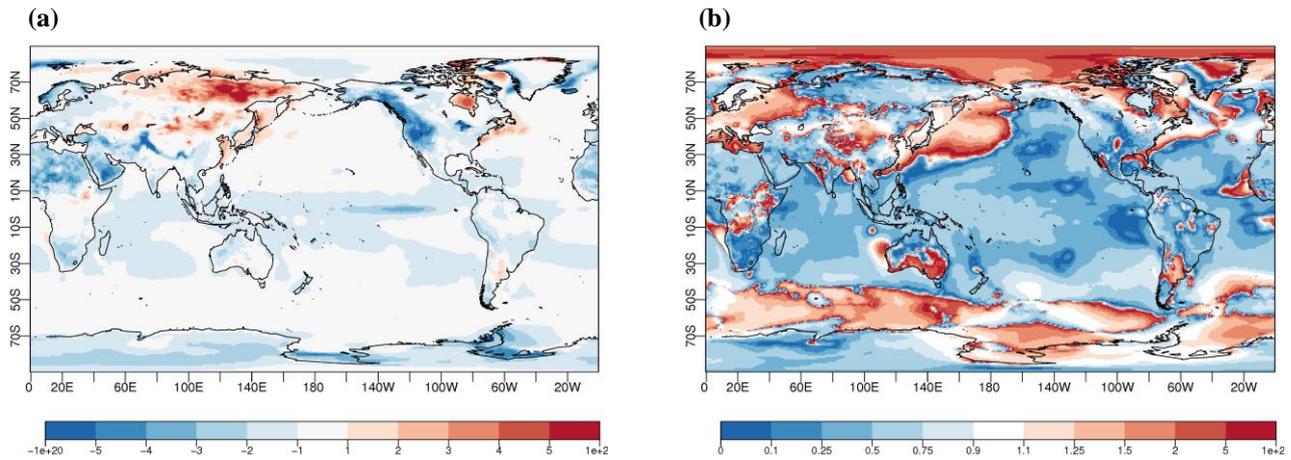
The SDYN implementation, which has also been carried out on EC-Earth3 to allow an appropriate comparison with the SPPT results, follows two main steps:

- Atmospheric relaxation run: the atmospheric component is relaxed towards ERAInt reanalysis data. Corrections towards ERAInt are saved daily and are the basis for the perturbation population  $\{\delta X\}$ . The implementation of the atmospheric relaxation in EC-Earth3 is a new tool available to the EC-Earth community, for whom the reference files have been made available in ECFS.
- Hindcasts: a seasonal re-forecast is run perturbing each ensemble every 6 hours with random draws of the  $\delta X$  set for the corresponding calendar month (in cross-validation mode).

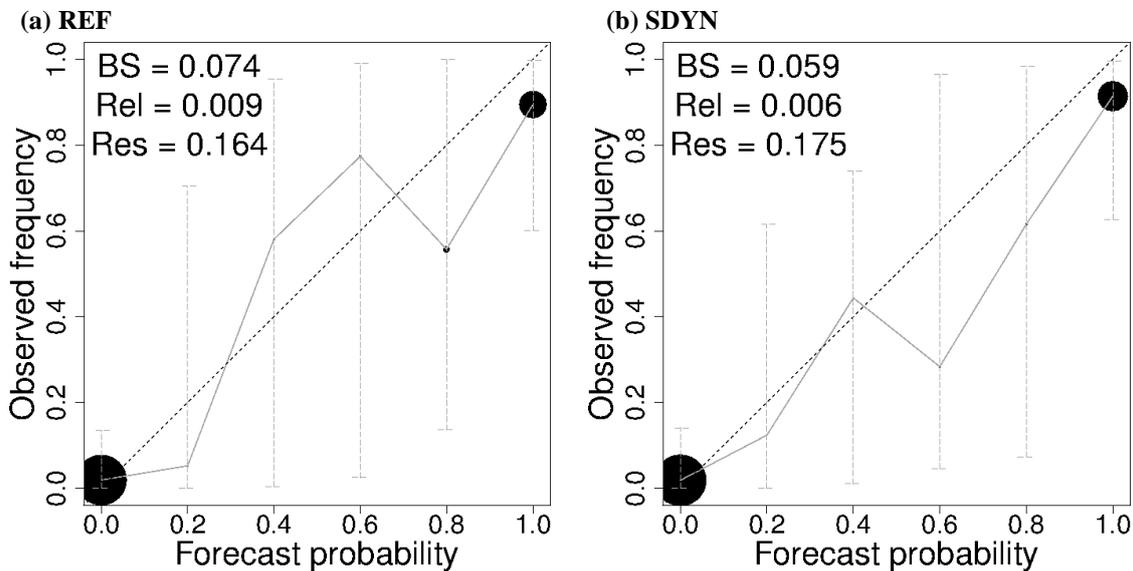
The first step was implemented in the IFS Cy36r4 available in EC-Earth3, where the atmospheric relaxation routines were adapted from the atmospheric relaxation branch available in perforce for a more recent cycle, Cy38r1 as recommended by Linus Magnusson (ECMWF). The modifications were tested with several atmospheric relaxation coefficients for temperature, u and v winds, humidity, surface pressure, cloud cover and cloud water fractions. Two nudged IFS runs over the 1993-2009 time period were completed, relaxing atmospheric fields to ERAInt data with relaxation coefficients of 0.1 (strong nudging) and 0.01 (weak nudging, corresponding to a nudging time scale of approximately 4 days). An analysis of the perturbation populations estimated with these nudged runs showed that the perturbation terms for the strong nudging were mainly made of intra-month variability, whereas the weak nudging terms contained a larger proportion of systematic corrections.

Five-member, four-month long ensemble seasonal hindcasts were performed with SDYN (second step of the method) started on the first of November over 1993-2009, using the perturbation population provided by the

weak nudging run and random draws of  $\delta X$  for each member every six hours. The hindcasts are compared to a reference five-member ensemble REF mentioned above.



**Figure 9:** (a) Systematic error for DJF 1993-2009 near-surface air temperature in REF with respect to ERA-Interim data and (b) relative systematic error for the same period for SDYN with respect to REF.



**Figure 10:** Reliability diagrams for the event near-surface air temperature exceeding the second climatological tercile over the Niño3.4 region in DJF for (a) REF and (b) SDYN hindcasts started in November over 1993-2009. The size of the dots shows the population of each forecast probability bin over the hindcast period, and vertical bars show the range of uncertainty in the reliability diagram based on 1,000 bootstrap samples over the hindcast period.

Figure 9 shows the impact of the stochastic dynamics perturbations on the systematic error of near-surface air temperature for winter hindcasts. The relative systematic error in the SDYN ensemble (figure (b)) highlights that temperature bias is noticeably reduced over most of the tropics, as well as over Northern Eurasia where a strong warm bias was found in the REF experiment. Some areas do exhibit an increase in bias, such as most of Europe and the Mediterranean, the Arctic and the northeast Pacific Ocean.

Unlike SPPT, the stochastic dynamics method has no impact on ensemble spread (it even appears to reduce spread for some variables, although with five ensemble members the confidence intervals are very large). Figure 10 shows, as in Figure 8, the reliability diagrams but for near-surface air temperature over the Niño3.4 region in winter for the November start dates. As for SPPT, the stochastic dynamics technique improves both reliability and resolution in the forecasts and leads to a better Brier Score. However, the confidence intervals are very large and improvements cannot be considered significant.

Future work on the subject includes running larger ensembles and longer hindcast periods to possibly draw more robust conclusions on the impact of stochastic perturbations in IFS on the EC-Earth3 seasonal forecast quality. For this purpose, we plan to use the NEMOVAR ocean initial conditions in future tests.

## Some conclusions

This special project has demonstrated the relevance for the forecast quality of seasonal-to-interannual predictions of producing accurate initial conditions for components of the climate system that traditionally do not receive as much attention as the atmosphere and the ocean. While the impact of the sea ice and land surface on the forecast quality is much lower than the impact of the other components, their relevance at local and regional scales as well as for predicting extreme, high-impact events like the summer 2012 sea-ice minimum or the summer 2010 eastern European heat wave is now much clearer thanks to these experiments. Special mention should be made of the experiments performed to assess the impact of different approaches to introduce stochastic perturbations during the forecasts. EC-Earth is the only system currently available where both SPPT and the Météo-France stochastic dynamics methods can be compared on an even foot. The results of these experiments, as well as of those that are being performed on different HPC platforms with computing time granted by other programmes, are directly relevant to similar efforts undertaken by ECMWF. The closeness between EC-Earth and the ECMWF seasonal forecast system and the continuous collaboration between the IC3 scientists involved and some ECMWF staff will benefit in the medium term the development of their operational sub-seasonal and seasonal systems. Further analysis of these experiments and the dissemination of the results in peer-reviewed journals already demonstrate and will clearly illustrate in the future the impact, not just for ECMWF's objectives but also for the climate-prediction community at large, of the resources granted by ECMWF through its special project programme.

## References

- Balmaseda MA, Mogensén KS, Weaver AT (2012) Evaluation of the ECMWF Ocean Reanalysis ORAS4. *Q J R Meteorol Soc.* doi:10.1002/qj.2063.
- Balsamo G et al. (2013) ERA-Interim/Land: a global land water resources dataset. *Hydrol. Earth Syst. Sci. Discuss.* 10:14705-14745. doi:10.5194/hessd-10-14705-2013.
- Batté L, Déqué M (2012) A stochastic method for improving seasonal predictions. *Geophys. Res. Lett.* 39. doi:10.1029/2012GL051406.
- Cavalieri DC, Parkinson C, Gloersen P, Zwally HJ (1996) Sea ice concentrations from Nimbus-7 SMMR and DMSP SSM/I passive microwave data. *National Snow and Ice Data Center. Digital media, Boulder.*
- Dee DP et al (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q J R Meteorol Soc* 137:553-597.
- Fan Y, van den Dool H (2008) A global monthly land surface air temperature analysis for 1948-present. *J Geophys Res* 113:D0110. doi:10.1029/2007JD008470.
- García-Serrano J, Doblas-Reyes F J (2012) On the assessment of near-surface global temperature and North Atlantic multi-decadal variability in the ENSEMBLES decadal hindcast. *Climate Dyn.* 39:2025-2040. doi:10.1007/s00382-012-1413-1.
- Guemas V, Doblas-Reyes F J, Mogensén K, Keeley S, Tang Y (2014) Ensemble of sea ice initial conditions for interannual climate predictions. *Climate Dyn.* doi:10.1007/s00382-014-2095-7.
- Hansen J, Ruedy R, Sato M, Lo K (2010) Global surface temperature change. *Rev Geophys* 48:RG4004. doi:10.1029/2010RG000345.
- Hazeleger W et al (2012) EC-Earth V2.2: description and validation of a new seamless Earth system prediction model. *Climate Dyn.* 39:2611-2629. doi: 10.1007/s00382-011-1228-5.
- Huffman GJ, Adler RF, Bolvin DT, Gu G (2009) Improving the global precipitation record: GPCP version 2.1. *Geophys. Res. Lett.* 36. doi:10.1029/2009GL040000.
- Kirtman B, Pirani A (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.
- Mogensén KS, Balmaseda MA, Weaver A (2011) The NEMOVAR ocean data assimilation as implemented in the ECMWF ocean analysis for System 4. *ECMWF Tech. Memo* 668.
- Olkin I and Finn J D (1995) Correlations redux. *Psychol. Bull.* 118:155-164. doi:10.1037/0033-2909.118.1.155.
- Palmer TN, Doblas-Reyes F J, Hagedorn R, Weisheimer A (2005) Probabilistic prediction of climate using multi-model ensembles: from basics to applications. *Phil. Trans. Roy. Soc. B* 360:1991-1998.
- Palmer TN et al (2009) Stochastic parametrization and model uncertainty. *ECMWF Tech. Memo* 598.

- Pegion K, Kirtman BP (2008) The impact of air-sea interactions on the simulation of tropical intraseasonal variability. *J. Climate*, 21:6616-6635.
- Saha S et al (2010) The NCEP Climate Forecast System Reanalysis. *Bull. Amer. Meteorol. Soc.*, 91:1015-1057.
- Smith TM, Reynolds RW, Lawrimore TCPH (2008) Improvements to NOAA's historical merged land-ocean surface temperature analysis (1880–2006). *J. Climate* 21:2283-2296.
- Von Storch H, Zwiers FW (2001) Statistical analysis in climate research. Cambridge University Press, Cambridge.

### **List of publications/reports from the project with complete references**

Guemas V., Doblas-Reyes F.J., Mogensen K., Tang Y., Keeley S., 2014. Ensemble of sea ice initial conditions for interannual climate predictions. *Climate Dyn*, doi:10.1007/s00382-014-2095-7.

Two more publications, one on the first SPPT results and another one with the preliminary description of the impact of the land-surface initialisation are about to be submitted.

### **Summary of plans for the continuation of the project**

(10 lines max)

At this stage more than 95% of the allocated resources have been consumed. We plan to apply for additional resources in this final year, which would be extremely useful to continue exploring some of the vast range of options of the stochastic parameterisations tested in EC-Earth3.

A new project has been submitted for the period 2015-2016.