REQUEST FOR A SPECIAL PROJECT 2024–2026

MEMBER STATE: Spain

Principal Investigator[1]: Etienne Tourigny

Affiliation: Barcelona Supercomputing Center – Earth Sciences Department

Address: Plaça Eusebi Güell, 1-3, 08034 Barcelona (Spain)

Other researchers: Markus Donat, Roberto Bilbao, Rashed Mahmood, Pablo Ortega Montilla, Miguel Castrillo, Vladimir Lapin, Aude Carreric

Project Title: Understanding inter-annual to decadal predictability in the EC-Earth3 model and the potential benefits from perfect initialisation

If this is a continuation of an existing project, please state the computer project account assigned previously.	**SPESICCF**	
Starting year:	2024	
Would you accept support for 1 year only, if necessary?	YES x	NO ☐

Computer resources required for project year:	**2024**	**2025**	**2026**
High Performance Computing Facility [SBU]	120 million		
Accumulated data storage (total archive volume)[2] [GB]	20480		

EWC resources required for project year:	**2024**	**2025**	**2026**
Number of vCPUs [#]	0		
Total memory [GB]	0		
Storage [GB]	0		
Number of vGPUs[3] [#]	0		

Continue overleaf.

[1] The Principal Investigator will act as contact person for this Special Project and, in particular, will be asked to register the project, provide annual progress reports of the project's activities, etc.
[2] These figures refer to data archived in ECFS and MARS. If e.g. you archive x GB in year one and y GB in year two and don't delete anything you need to request x + y GB for the second project year etc.
[3] The number of vGPU is referred to the equivalent number of virtualized vGPUs with 8GB memory.

Principal Investigator: Etienne Tourigny

Project Title: Understanding inter-annual to decadal predictability in the EC-Earth3 model and the potential benefits from perfect initialisation

Extended abstract

I. Scientific plan

A major challenge for the advance of seasonal-to-decadal climate predictions is related to model initialization towards the observed state of the climate. In addition to imperfect estimates of the observed initial state due to limited observations, current initialization methods can cause 'initialization shocks' (1,2). These are rapid adjustments, due to inconsistencies between the observations and the model climatology, or between model components that have been initialized with different observational data products, and can interfere with the models' ability to evolve real climate signals (3). Partly related to the initialization shock, the initialized simulations may also be affected by climate drift (4), which can mask useful information and reduce prediction skill.

In the recently granted Europa Excelencia project funded by the Agencia Estadal de Investigación (PRECEDE, grant number EUR2022-134059) we proposed to explore the effects of these initialisation-related issues, and to what extent they affect the model-specific predictability of climate. To this end, we will perform so-called "perfect model" prediction experiments which are useful to study predictability in the absence of such artifacts (5–7). These experiments aim to predict a climate realization of the same climate model used to make the predictions, rather than attempting to predict the real-world climate, and therefore avoid the inconsistencies affecting real-world predictions. Perfect-model predictability has usually been investigated using different experimental setups and measures (8–10) than those used for the evaluation of real-world predictions (11). Therefore previous investigations of perfect-model predictability are largely unsuitable for comparison to real-world predictions.

While these perfect-model predictions do not provide information for the predictability of the real world, these experiments are useful to better understand the climate predictions and to what extent they may be deteriorated by the initialisation-related shortcomings (12). To generate this novel understanding, the behaviour of the perfect-model predictions will be compared to real-world predictions initialized with observations. By using the same external forcing as in the real-world predictions, and running predictions with annual initialisation and same ensemble size as used for real-world decadal predictions, we will produce and analyze perfect-model prediction experiments that are directly comparable to (already existing) real-world climate predictions performed with the EC-Earth climate model. In comparison to initialized real-world predictions, these simulations can help us answer the question of how the imperfect initialization affects predictability in the climate model. In particular, we will compare the behaviour of the simulations with ideal and non-ideal initialisation (evaluating how well the model predicts other initialised simulations), to understand whether and how the predictability estimates are affected by the initialisation-related shortcomings.

While it is known that models exhibit limitations in predicting the real-world climate (13,14), such idealised prediction experiments can teach us important lessons as to which improvements to our prediction systems (i.e. initialisation, targeted observations in key regions, increased process realism) can be most effective towards developing predictions of highest skill. Such a large set of standardized perfect-model experiments covering predictions on seasonal to decadal time scales requires large computational resources and has not been produced before in this size. These perfect-model experiments can in the future be further exploited to analyse predictability of other climate features and extremes such as heatwaves, which will also contribute to the Horizon Europe ASPECT project. Finally, comparison to (existing but smaller) perfect-model experiments with other models (12) will provide new insights into the model-dependence of predictability.

II. Justification of the computing resources requested

We will perform a set of coupled climate model simulations to examine predictability in the absence of initialization shocks and model drift, and avoiding imperfect assimilation of the observed climate state (due to lack of observations), using the EC-Earth climate model. To this end, simulations will be started from each year of a historical climate simulation covering the 20th and early 21st century (same as the 'historical' forcing simulations in CMIP6). The climate model states at a specific time (e.g. 1st November – same as the initialization date of the real-world decadal predictions with EC-Earth) of the historical simulation will be used to initialize an ensemble of decadal simulations, all using identical ocean states but with small perturbations to the atmosphere temperature field as initial errors that can then grow and define an uncertainty range for each prediction (15). The concept of these idealized prediction experiments is to estimate predictability from almost identical initial conditions between the different ensemble members, with perturbations orders of magnitude smaller than the accuracy of observations used to initialize real-world predictions (12).

These idealised predictions will then be compared to existing prediction experiments initialised with (imperfect) observations. Besides an investigation of general predictability estimates (11,12), our analysis will in particular focus on comparing the simulations with perfect and imperfect initialisation. Previous studies have highlighted the so-called signal-to-noise paradox, i.e. that current climate models predict the real-world climate with higher skill than they can predict the model climate itself (16). These estimates of model-predictability, however, have used simulations initialised with (imperfect) observations, and these simulations have therefore also suffered from the shocks and drift affecting the real-world predictions. Based on the new perfect-model predictions, we will estimate to what extent the model-predictability is affected by those initialisation-related shortcomings used for previous estimates, and whether perfect initialisation can avoid some of the inconsistencies between model and real-world predictability.

In our research group, we use EC-Earth (17) as the model for coupled climate and earth system simulations, and substantial expertise and technical support is available to ensure best possible computing performance. The experiment will be run with the version 3.3 of EC-Earth using its standard configuration: the atmospheric component is the IFS (from the ECMWF) with a T255 horizontal resolution (approximately 80 km) and 91 vertical levels, and the ocean component is NEMO3.6 and the LIM3 sea ice model, both run with an ORCA1 configuration (1° horizontal nominal resolution) and 75 vertical levels. This model version is used in the BSC Decadal Prediction System (18) and has been used for most of the EC-Earth contributions to the sixth phase of the Climate Model Intercomparison Project (CMIP6) (19). While in our group we have extensive experience performing experiments with the EC-Earth model in its standard configuration on the Marenostrum4 HPC, the major challenge of the proposed experiment is its size in terms of simulated years (as detailed below), which requires a dedicated computing project. Additionally, the computationally-expensive simulations must be done concurrently and in a short time, in order to be completed before the planned shutdown of Marenostrum4 in summer 2023.

Using EC-Earth version 3.3, we will perform the largest existing sets of perfect-model prediction experiments to robustly and comprehensively estimate predictability. Starting idealized predictions each year in November from 1960 to 2014 (i.e. 55 different initial states; 1960 is chosen as first year because this is typically the first initialization year also in real-world decadal prediction experiments (20)), we will run ensembles of 10 members (matching the ensemble size of real-world decadal predictions contributing to the CMIP6-DCPP project) for up to 11 years after initialization – amounting to a total of 6,050 simulated years.

We will use the Autosubmit workflow manager (21) developed by the BSC Computational Earth Science (CES) group to submit the simulation, post-processing and cleanup jobs required for such a complex workflow. The Autosubmit workflow includes tasks to send processed model output to the BSC data storage and clean the HPC storage space within reasonable time limits. We will require a minimum of 30 TiB of scratch space, consisting of model code, input files and temporary output files, to guarantee that we can run several experiments in parallel (each experiment consisting of a given start date and member occupying 200GiB of temporary disk space).

Significant effort has been put into finding the optimal computing setup on the Marenostrum4 supercomputer located at the BSC, using a modified oasis coupler optimization tool (lucia-lite) and employing machine files to ensure optimal load-balancing and minimizing the waste of computing resources. As a result of these optimizations, we identified the best configuration which requires 768 cores (16 nodes), and the time taken to simulate a model year is approximately 1.25 hours, consuming on average 950 CHPSY (Computing Hours Per Simulated Year). The EC-Earth3 model has already been ported to the ECMWF HPC2020 system by EC-Earth consortium partners, and we have completed the porting of the required utilities to run EC-Earth3 with the Autosubmit workflow manager. We are in the process of finding the optimal computing setup on the ECMWF HPC2020 system and expect a similar computational cost as that on Marenostrum4.

The computing resources required to run the 6,050 simulated years (at a cost of 950 CHPSY) amount to 5,747,500 CPU hours, or a total of 6,322,250 CPU hours, when accounting for an extra 10% for run failures and serial post-processing. Using the CPUH-to-SBU conversion factor of 18,91, we estimate a consumption of approximately 120,000,000 SBU on the ECMWF HPC2020 system.

III. Technical characteristics of the code to be used

The EC-Earth3 GCM (Global Climate Model) comprises three major components: the atmospheric model IFS (Integrated Forecasting System) Cy36r4, the ocean model NEMO 3.6, which also includes the LIM3 sea-ice model, and OASIS3 that couples the main components. IFS is an operational global meteorological forecasting model developed and maintained by the European Centre of Medium-Range Weather Forecasts (ECMWF). NEMO is a state-of-the-art modelling framework for the ocean used for oceanographic research, operational oceanography, seasonal forecasting and climate research studies. In this activity we will use the standard T255-ORCA1 configuration, which corresponds to a spatial resolution of approximately 80 km in the atmosphere/land and around 100 km in the ocean. This activity will be carried out with EC-Earth version v3.3.3 released on 25 Jun 2020, for consistency with the CMIP6 simulations.

As the experiment will have a complicated workflow in certain phases, the Autosubmit software will be used to manage the workflow and ensure a uniform and optimal use of the resources. The jobs will be managed, and packed in groups in a single big job whenever required, by Autosubmit to better manage the I/O system while maximising the use of the machine.

IV. Relevant publications from the Special Project researchers

Bilbao, R., S. Wild, P. Ortega, J. Acosta-Navarro, T. Arsouze, P.-A. Bretonnière, L.-P. Caron, M. Castrillo, R. Cruz-García, I. Cvijanovic, F. J. Doblas-Reyes, M. Donat, E. Dutra, P. Echevarría, A.-C. Ho, S. Loosveldt-Tomas, E. Moreno-Chamarro, N. Pérez-Zanon, A. Ramos, Y. Ruprich-Robert, V. Sicardi, E. Tourigny, J. Vegas-Regidor (2021), Assessment of a full-field initialized decadal climate prediction system with the CMIP6 version of EC-Earth, Earth Syst. Dynam., 12, 173–196, https://doi.org/10.5194/esd-12-173-2021

Liu, Y., M. G. Donat, H. W. Rust, L. V. Alexander, M. H. England (2019), Decadal predictability of temperature and precipitation means and extremes in a perfect-model experiment, Climate Dynamics, 53, 3711–3729, https://doi.org/10.1007/s00382-019-04734-z

Liu, Y., M. G. Donat, A. S. Taschetto, F. J. Doblas-Reyes, L. V. Alexander, M. H. England (2019). A framework to determine the limits of achievable skill for interannual to decadal climate predictions. Journal of Geophysical Research: Atmospheres, 124, 2882– 2896. https://doi.org/10.1029/2018JD029541

Delgado-Torres, C., M. G. Donat, N. Gonzalez-Reviriego, L.-P. Caron, P. J. Athanasiadis, P.-A. Bretonnière, N. J. Dunstone, A.-C. Ho, K. Pankatz, A. Paxian, N. Pérez-Zanón, M. S. Cabré, B. Solaraju-Murali, A. Soret, F. J. Doblas-Reyes (2022), Multi-model forecast quality assessment of CMIP6 decadal predictions, Journal of Climate, 35(13), 4363-4382, https://doi.org/10.1175/JCLI-D-21-0811.1

V. References

(1) W. Hazeleger, et al., Geophys. Res. Lett. 40, 1794–1798 (2013).
(2) D. M. Smith, R. Eade, H. Pohlmann, Clim. Dyn. 41, 3325–3338 (2013).
(3) M. J. McPhaden, S. E. Zebiak, M. H. Glantz, Science 314, 1740–5 (2006).
(4) V. V. Kharin, G. J. Boer, W. J. Merryfield, J. F. Scinocca, W.-S. Lee, Geophys. Res. Lett. 39, L19705 (2012).
(5) G. J. Boer, Clim. Dyn. 16, 469–477 (2000).
(6) M. Collins, M. R. Allen, J. Clim. 15, 3104–3109 (2002).
(7) G. J. Boer, Clim. Dyn. 36, 1119–1133 (2011).
(8) H. Pohlmann, et al., J. Clim. 17, 4463–4472 (2004).
(9) M. Collins, et al., J. Clim. 19, 1195–1203 (2006).
(10) G. Branstator, H. Teng, J. Clim. 23, 6292–6311 (2010).
(11) L. Goddard, et al., Clim. Dyn. 40, 245–272 (2012).
(12) Y. Liu, M.G. Donat et al., J. Geophys. Res. Atmos. 124, 2882–2896 (2019).
(13) D. M. Smith, et al., Nature 583, 796–800 (2020).
(14) M. A. Walz, M. G. Donat, G. C. Leckebusch, J. Geophys. Res. Atmos. 123, 11,518-11,535 (2018).
(15) W. A. Müller, et al., Geophys. Res. Lett. 39, L22707 (2012).
(16) A.A. Scaife, Smith, D., npj Clim Atmos Sci 1, 28 (2018). https://doi.org/10.1038/s41612-018-0038-4
(17) Döscher, et al.Geosci. Model Dev., gmd-15-2973-2022 (2022).
(18) Bilbao et al., Earth Syst. Dynam., esd-12-173-2021
(19) Eyring, V. et. al,Geosci. Model Dev., gmd-9-1937-2016, (2016) .
(20) J. Marotzke, et al., Bull. Am. Meteorol. Soc., BAMS-D-15-00184.1 (2016).
(21) Manubens-Gil, D., et. al, Proceedings of the 2016 International Conference on High Performance Computing & Simulation (HPCS) (2016).