MODIFIED REQUEST FOR A SPECIAL PROJECT 2017–2019

MEMBER STATE: SPAIN ……….………………………………………...…

Principal Investigator[1]**:** Louis-Philippe Caron
………………………………………….……

Affiliation: Barcelona Supercomputing Center – Earth Sciences Department

Address: Carrer Jordi Girona, 31, 08034 Barcelona

……………………………………………………………….……

E-mail: louis-philippe.caron@bsc.es………………………………….……

Other researchers: Eleftheria Exarchou, Roberto Bilbao, Javier Garcia-Serrano, ~~Neven Fuckar~~, Francois Massonnet, Louis-Philippe Caron, Mario Acosta, Oriol Tinto, Miguel Castrillo, Chloe Prodhomme, Martin Menegoz, Etienne Tourigny, Domingo Manubens, Pierre-Antoine Bretonniere, Thomas Arsouze

……………………………………………………………….……

Project Title: CMIP6 BSC contribution to HighResMIP (HighResMIP_BSC)

……………………………………………………………….……

If this is a continuation of an existing project, please state the computer project account assigned previously.	**SP** ________________	
Starting year: (Each project will have a well-defined duration, up to a maximum of 3 years, agreed at the beginning of the project.)	2017	
Would you accept support for 1 year only, if necessary?	YES X	NO ☐

Computer resources required for 2017-2019:

		2017	2018	2019
High Performance Computing Facility	(millions SBU)	41.5	41.5	41.5
Data storage capacity (total archive volume)	(Terabytes)	30	60	60

An electronic copy of this form ***must be sent*** *via e-mail to:* *special_projects@ecmwf.int*

Electronic copy of the form sent on (please specify date):

…………29 June 2018 ……………

Additions to the original proposal are in blue

Principal Investigator: Louis-Philippe Caron

---

[1] The Principal Investigator will act as contact person for this Special Project and, in particular, will be asked to register the project, provide an annual progress report of the project's activities, etc.

Extended abstract

It is expected that Special Projects requesting large amounts of computing resources (500,000 SBU or more) should provide a more detailed abstract/project description (3-5 pages) including a scientific plan, a justification of the computer resources requested and the technical characteristics of the code to be used. The Scientific Advisory Committee and the Technical Advisory Committee review the scientific and technical aspects of each Special Project application. The review process takes into account the resources available, the quality of the scientific and technical proposals, the use of ECMWF software and data infrastructure, and their relevance to ECMWF's objectives. - Descriptions of all accepted projects will be published on the ECMWF website.

I – Scientific plan

Recent studies have established that the typical atmospheric and oceanic resolutions used for the CMIP5 coordinated exercise (Coupled Model Intercomparison Project, phase 5) are a limiting factor to correctly reproduce climate mean state and variability. Noteworthy improvements when increasing resolutions have been obtained in the simulation of El Niño Southern Oscillation (ENSO) (Shaffrey et al. 2009, Masson et al. 2012), Tropical Instability Waves (Roberts et al. 2009), the Gulf Stream and its influence on the atmosphere (Chassignet and Marshall 2008; Kuwano-Yoshida et al. 2010), the global water cycle (Demory et al. 2014), jet stream (Lu et al., 2015), storm tracks (Hodges et al. 2011), Euro-Atlantic blockings (Jung et al 2012), tropical cyclones (Zhao et al. 2009, Bengtsson et al. 2007), tropical-extratropical interactions (Baatsen et al. 2014, Haarsma et al. 2013), monsoons (Sperber et al. 1994; Lal et al. 1997; Martin 1999), heat waves and droughts (Van Haren et al. 2015) and sea ice drift and deformation (Zhang et al 1999; Gent et al, 2010).

The main obstacle to running climate models at higher resolution is computational. Up to now, relatively few research centres have carried out such highly-demanding simulations. These were relatively short experiments, with very few or only one member and not performed in a coordinated way. The HighResMIP coordinated exercise, as part the Sixth Phase of the Coupled Model Intercomparison Project (CMIP6), offers a framework for increasing synergies and building a large multi-model ensemble of high resolution simulations with a low resolution counterpart following a common experimental protocol, i.e. a common integration period, forcing and boundary conditions. This coordinated exercise will allow for an optimization of the computing resources usage toward a common goal of identifying the robust benefits of increased model resolution based on multi-model ensemble simulations. This is the first time ever such a largely demanding computing exercise will be implemented and the largest ensemble ever built at such high resolution. The ensemble spread will allow diagnosing uncertainties and attributing them to model formulation. The CMIP6 framework ensures a high level of participation to this project. Conclusions will be highly relevant to the seasonal prediction science in general and to ECMWF activities in particular, in which the question of the added-value of increased resolution remains open.

Process-based assessment will be central to HighResMIP. The main goal is to pin down physical and dynamical reasons behind differences in model representation induced by resolution change. Such a process-based assessment is supported by the PRIMAVERA project, funded by the European commission under the H2020 framework programme in which BSC is highly involved. New diagnostics and metrics will be developed which fulfil essential criteria such as: 1) focusing on processes or feedbacks rather than large-scale performance metrics that can hide compensation of errors; 2) relying on observable variables; 3) being stable with respect to internal variability on different timescales, i.e. insensitive to the time period selected; 4) being interpretable physically; 5) being distributed in publicly released tools to ensure reproducibility. The HighResMIP simulations will focus on present-day and near-term future climate. Hence, process-based assessment will not focus on climate sensitivity but on the representation of mean state, variability and teleconnections on a wide range of timescales. The primary goal will be to determine which processes can be represented reliably at typical CMIP5 resolutions and what is the minimum resolution required for an adequate representation of other processes as well as what are the limitations of representing such processes in lower resolution models. A wide variety of processes will be considered from global to regional scales and from the upper atmosphere down to mesoscale eddies. The applying team covers a wide range of scientific expertise which will allow focusing a wide range of process that can be affected by resolution: sea ice dynamics and thermodynamics (Neven Fuckar and Francois Massonnet), El Nino Southern Oscillation and tropical-extratropical teleconnections (Javier Garcia Serrano), monsoons and heat waves and droughts (Chloe Prodhomme), Tropical Instability Waves (Eleftheria Exarchou), the Gulf Stream and its influence on the atmosphere (Roberto Bilbao), the jet streams and Euro-Atlantic blockings (Etienne Tourigny), tropical cyclones (Louis-Philippe Caron). Within the framework of the H2020 PRIMAVERA project, strong collaboration with the MetOffice (UK), the University of Reading (UK), KNMI (Holland), SMHI (Sweden), CERFACS (France), Max Planck Institute for Meteorology (Germany), the Universite Catholique de Louvain La Neuve (Belgium), CMCC (Italy), AWI (Germany), CNR (Italy), ECMWF (UK) and the NERC (UK) will be highly beneficial to our project.

Through the HighResMIP_BSC project, BSC plans to take part in HighResMIP using EC-Earth3.2, which corresponds to the latest available model version of the community model EC-Earth. A spectral truncation of the atmospheric model (IFS) at T511 (approx. 40 km globally) and 91 vertical levels will be used as well as grid resolution of the ocean model (NEMO3.6) of 0.25° globally (approximately 25 km) with 75 vertical levels which thickness increases from 1m below surface up to 500m in the deep ocean. Compared to the resolution of the model used in CMIP5 simulations, the horizontal resolution is increased by a factor of up to 4, with an increased number of vertical levels in both ocean and atmosphere by factors of 1.5 to 2. Accounting for the reduced time step necessary for stability, this corresponds to an increase in computing resources by a factor of about 30. BSC will be involved in the Tier 2 of the HighResMIP project. Its experimental protocol consists in running a 50-year spinup under 1950 conditions followed by: 1) a historical simulation covering the 1950-2050 period, 2) a control simulation under 1950 conditions run for 100 years. This experimental protocol has been chosen as a compromise between limiting the computational cost to ensure that a maximum number of participating institutes can afford it and allowing for a minimization (thanks to the spinup) and subtraction (thanks to the control) of the model drift. The spinup will be initialised from the EN4 ocean climatology.

Additional Material

Within the framework of the European project PRIMAVERA, funded within the H2020 program, BSC is developing a coupled version of EC-Earth 3.2 at a groundbreaking resolution. In the atmosphere the horizontal domain is based on a spectral truncation of the atmospheric model (IFS) at T1279 (approx. 15 km globally, i.e. the highest resolution we can use with the standard IFS - higher resolutions would require e.g. non-hydrostatic parameterizations) together with 91 vertical levels. The ocean component (NEMO), is run on the so-called ORCA12 tripolar (cartesian) grid, at a horizontal resolution of about 1/12° (approximately 16 km); with 75 vertical levels which thickness increases from 1m below surface up to 500m in the deep ocean. Compared to the resolution of the model used in CMIP5 simulations, the horizontal resolution is increased by one order of magnitude, with an increased number of vertical levels in both ocean and atmosphere by factors of 1.5 to 2. At this resolution, we expect many smaller scale phenomena such as tropical cyclones and deepwater formation to be significantly improved.

For the extension of this project, we would like to produce a continuous simulation with this new configuration in order to further study the benefits of increasing climate model resolution. It would be unrealistic to follow the HighResMIP protocol at this stage (i.e. follow the protocol described above). Instead, we plan to run some scalability test (similar to what has been done previously with the T511/ORCA025 configuration) to determine the optimum performance of this new configuration. Once this test has been completed, we plan to run a spin up simulation. Ideally, we would like to perform a spinup similar to the one performed with the T511-ORCA025 configuration as it would allow a direct comparison of processes at 3 different resolutions, but based on our estimate, 41.5 SBU (which we are requesting) will not be sufficient to do so. However, KNMI, who is part of the EC-Earth consortium and also has a special project on a topic very close to this one (EC-Earth high resolution simulations) has offered their collaboration to run this very high resolution configuration. Depending on the number of hours available, we might be able to complete the 50-year spin-up. We will also consider applying for additional resources to obtain additional computing hours in 2019, as we have done in year 1 of this project, if necessary.

	M1-2	M3-4	M5-6	M7-8	M9-10	M11-12	M13-14	M15-16	M17-18	M19-20	M21-22	M23-24
Spinup	■	■										
Control			■	■	■	■	■					
Historical								■	■	■	■	■

Table 1: Approximate schedule of the experiments to be performed during year 1-2. M5 stands for Month 5.

	M25-26	M27-28	M29-30	M31-32	M33-34	M35-36
Historical	■					
Scalability test of the VHR configuration		■				
Spinup - VHR		■	■	■	■	■

Table 1a: Approximate schedule of the experiments to be performed during year 3 (2019).

II – Justification of the computing resources requested

The HighResMIP Tier 2 experimental protocol makes 50 years of spinup, 100 years of control simulation and 100 years of historical simulation, i.e. a total of 250 years of simulation. These experiments will follow the schedule indicated in Table 1. They will be run with EC-Earth 3.2 in configuration ORCA025L75 – T511L91 (described in section I). The low resolution counterpart will be performed using national resources.

Table 2 summarized the computing resources needed to run these experiments, which comprise between 150 and 300 chunks of 4 month length, to reach 50 or 100 years of simulations respectively. The two 100-year long simulations can be run independently and simultaneously depending on the machine load. These experiments use the HR EC-Earth3.2 configuration. Its cost, for a four-month simulation requires a wall-clock time of 4 hours (14400 seconds).

Several tests have been done within the framework of the SPESICCF project in order to evaluate the performance of EC-Earth, evaluating different metrics such as SpeedUp, Efficiency and Energy consumed. The results of these tests show that optimum performance is obtained when using 1730 for the high resolution coupled experiments. These values provide a compromise between the required amount of resources and the computational cost, taking into account the scalability of the parallel application and the average load of the platform.

Run type	# Runs	# Steps/Run	Walltime/Step	# SBU / Step	Total SBU
Spinup	1	150	14400 (sec)	112000	16800000
Control	1	300	14400 (sec)	112000	33600000
Historical	1	300	14400 (sec)	112000	33600000
Total					79000000

Table 2: Cost of the experiments proposed.

The final estimate is for a total request of 83 million SBU, which includes the numbers described in Table 2 plus a small buffer of 5% to account for failing jobs that will need to be repeated. Regarding storage space, we estimate about 1Tb of output per year of simulation, given the substantial requirements of HighResMIP in terms of output variables and their frequency. This would make a total of 250Tb generated by the project. The experiments will be run using Autosubmit, the launching and monitoring solution developed by the applying team that allows the remote submission of EC-Earth and NEMO experiments. Autosubmit includes in the workflow of the experiments a job that retrieves the data back to the Department data storage as soon as each chunk of simulation has completed. This means that the actual storage space that will be required will approach a figure ten times smaller than the total outputs generated, i.e. about 30Tb.

We are requesting an additional 41.5M SBU, which is in line with the amount of resources of the first two years of the project. Table 2a provides some information, based on local estimate, on the experiment, which comprises 1 simulation of 12.5 years organized into 150 chunks of 1 month length. A one-month simulation requires a wall-clock time of ~7 hours. With the current estimate, we should be able to complete around 12.5 years, but we might be able to push this number further based on the results of the scalability test. And as mentioned previously, we could use the resources from a different special project to extend this spinup further.

Run type	# Runs	# Steps/Run	Walltime/Step	# SBU / Step	Total SBU
Spinup	1	150	25,000 (sec)	280,000	42,000,000

Table 2a: Cost of the new experiment.

If possible, we would like to have an additional 30Tb of storage, as we expect the amount of data to be produced to increase significantly.

III – Technical characteristics of the code to be used

The technical characteristics of the code remain the same, except for the atmosphere and ocean resolutions. The atmosphere will be run at T1279 (~2140702 horizontal grid points) while the ocean will use the ORCA12 configuration (13220998 horizontal grid points).

EC-Earth3 is a global coupled climate model, which integrates a number of components in order to simulate Earth systems. EC-Earth3 consists of two main components which are coupled using a library known as OASIS3-CMT. These two main components are the IFS model for the atmosphere, developed by the ECMWF, and NEMO for the ocean, developed by the European consortium. EC-Earth uses the MPI paradigm in order to distribute the workload on a supercomputer, using a specific number of task or processes for NEMO and IFS, the OASIS3-CMT coupling is used to communicate directly IFS and NEMO. It is essential to configure and build a separate executable for each one of them. The resolution proposed (T511: ~425,000 grid points, ORCA025: ~1,475,000 grid points) will help efficiently share calculations between 1000-1500 sub-domains, increasing the range of efficient compute-core usage per model executable. For IFS there is a possibility to activate an OpenMP switch but, in this case, the implemented MPI should be thread-safe. IFS generates the output in GRIB format and NEMO in NetCDF, while OASIS3 does not generate any output. At the end of a simulation the three components always generate restarts separately (IFS in binary, and NEMO and OASIS3 in NetCDF format).

For configuring and building the model executables, GNU make 3.81 or 3.81+, FORTRAN 77/90/95 compliant compiler with pre-processing capabilities and NetCDF4 deployed with HDF5 and SZIP are needed. A newly designed tool for automatic build configuration called "ec-conf" can be used. This useful tool requires Python 2.4.3 or 2.4.3+ (although it does not work yet with Python 3.0+). For NEMO, the FCM bash and perl mechanism is essential, as it is the I/O GRIB_API 1.9.9 or 1.9.9+ and GRIBEX 370 mechanism that are needed for IFS. To test the model with the run scripts, GNU date (64-bit) is also required.

The simulations will require MPI libraries and runtime facilities (MPICH2, MPICH-MX, HP-MPI, OpenMPI, INTEL-MPI), optimization and data handling tools, such as BLAS, LAPACK, HDF4, HDF5, NETCDF, PARMETIS, SCALAPACK, P-NETCDF, UDUNITS, GRIB_API, CDFTOOLS v2, CDO, NCO and general configurations tools, such as PERL, PYTHON, AUTOCONF and AUTOMAKE.

The Autosubmit software (Asif et al., 2014) will be used to manage the workflow and ensure a uniform and optimal use of the resources. Our simulations offer flexibility as they comprise 2 independent 100-year long simulations which can run in parallel for an optimal use of the computing resources. The data storage and data transfer can be organized with a disk space of 10 TB in the "scratch" file system. This required scratch space is motivated by the large amount of output to be generated. These output data will be transferred immediately locally. Around 500 GB of "home" space will be required to host the code and its modified versions.

IV – References

Asif, M., A. Cencerrado, O. Mula-Valls, D. Manubens, F.J. Doblas-Reyes and A. Cortés (2014). Impact of I/O and data management in ensemble large scale climate forecasting using EC-Earth3. Procedia Computer Science, 29, 2370-2379, doi:10.1016/j.procs.2014.05.221.

Baatsen, M., R. J. Haarsma, A. J. Van Delden, H. De Vries, 2014: Severe autumn storms in future western Europe with a warmer Atlantic ocean. Clim. Dyn., 45, 949-964. DOI:10.1007/s00382-014-2329-8

Bengtsson, L., K. I. Hodges, M. Esch, N. Keenlyside, L. Kornblueh, J.-J. Luo, and T. Yamagata, 2007: How may tropical cyclones change in a warmer climate? Tellus, 59A, 539-561.

Chassignet, E.P., and D. P. Marshall, 2008: Gulf stream separation in numerical ocean models. In: Hecht, M., Hasumi, H. (eds.), Eddy-Resolving Ocean Modeling, AGU Monog. Ser., 39-62.

Demory, M.-E., et al. (2014). The role of horizontal resolution in simulating drivers of the global hydrological cycle. Clim. Dyn., 42, 2201-2225.

Gent, P. R., S. G. Yeager, R. B. Neale, S. Levis, and D. A. Bailey, 2010: Improvements in a half degree atmosphere/land version of the CCSM. Clim. Dyn., 34, 819–833, doi:10.1007/s00382-009-0614-8.

Haarsma, R.J., W. Hazeleger, C. Severijns, H. de Vries, A. Sterl, R. Bintanja, G.J. van Oldenborgh and H.W. van den Brink, 2013: More hurricanes to hit Western Europe due to global warming. Geophys. Res. Lett. doi:10.1002/grl.50360.

Hodges, K. I., et al., 2011: A comparison of extratropical cyclones in recent re-analyses ERA-Interim, NASA MERRA, NCEP CFSR, JRA-25. J. Clim., 24, 4888-4906.

Jung, T., et al., 2012: High-resolution global climate simulation with the ECMWF model in project Athena: Experimental design, model climate, and seasonal forecast skill. J. Clim., 25, 3155-3172.

Kuwano-Yoshida A, Minobe S, Xie S-P, 2010: Precipitation response to the Gulf Stream in an Atmospheric GCM. J. Clim 23, 3676–3698.

Lal, M., U. Cubasch, J.P. Perlwitz, and J. Waszkewitz, 1997: Simulation of the Indian monsoon climatology in ECHAM3 climate model: Sensitivity to horizontal resolution. Int. J. Climatol., 17, 847-858

Martin, G. M., 1999: The simulation of the Asian summer monsoon, and its sensitivity to horizontal resolution, in the UK meteorological office unified model. Q.J.R. Meteorol. Soc., 125: 1499–1525. doi:10.1002/qj.49712555703

Masson, S., P. Terray, G. Madec, J.-J. Luo, T. Yamagata, and K. Takahashi, 2012: Impact of intra-daily SST variability on ENSO characteristics in a coupled model. Climate Dynamics, 39, 681-707

Roberts, M. J., et al., 2009: Impact of Resolution on the Tropical Pacific Circulation in a Matrix of Coupled Models. J. Clim., 22, 2541-2556.

Zhao, M., et al., 2009: Simulations of Global Hurricane Climatology, Interannual Variability, and Response to Global Warming Using a 50km Resolution GCM. J. Climate, 33, 6653-6678.

Shaffrey, L. C., et al., 2009: U.K. HiGEM: the new U.K. High-Resolution Global Environment Model-model description and basic evaluation. J. Clim., 22,1861-1896.

Sperber, K. R., H. Sultan, G. L. Potter, and J. S. Boyle, 1994: Simulation of the Northern Summer Monsoon in the ECMWF Model: Sensitivity to Horizontal Resolution. Mon. Wea. Rev., 122, 2461–2481. doi:10.1175/1520-0493(1994)122<2461:SOTNSM>2.0.CO;2

Van Haren, R., R. J. Haarsma, H. de Vries, G. J. van Oldenborgh and W. Hazeleger, 2015: Resolution dependence of circulation forced future central European summer drying. Env. Res. Lett., 10, 055002. DOI: 10.1088/1748-9326/10/5/055002.

Zhang, Y., Maslowski W., Semtner, A.J, 1999. Impact of mesoscale ocean currents on sea ice in high-resolution Arctic ice and ocean simulations. J Geophys Res, 104, 18409-18429, doi:10.1029/1999JC900158.