Project name	Glob15km (global 15km coupled climate simulations)
Research field	PE10_3 Climatology and climate change

Project leader

Title	Dr
Last name	Virginie
First name	Guemas
Organisation name*	Barcelona Supercomputing Center (BSC)
Department*	Earth Sciences Department
Group*	Climate Prediction group
Country	Spain

Abstract (max. half a page)
Recent studies have established that the typical atmospheric and oceanic resolutions used for the CMIP5 coordinated exercise (Coupled Model Intercomparison Project, phase 5), i.e around 40km-150km globally, are a limiting factor to correctly reproduce the climate mean state and variability. BSC has developed a coupled version of the EC-Earth 3.2 climate model at a groundbreaking horizontal resolution of about 15km in each climate system component. The HighResMIP coordinated exercise, as part of the Sixth Phase of the Coupled Model Intercomparison Project (CMIP6), offers a framework for building a large multi-model ensemble of high resolution simulations with a low resolution counterpart following a common experimental protocol. This coordinated exercise will allow for identifying the robust benefits of increased model resolution based on multi-model ensemble simulations. The Glob15km project proposes to follow the entire HighResMIP protocol for coupled climate simulations with this ground-breaking resolution configuration of EC-Earth 3.2. Its experimental protocol consists in running a 50-year spinup under perpetual 1950 conditions followed by: 1) a historical simulation covering the 1950-2050 period, 2) a control simulation under perpetual 1950 conditions run for 100 years. These simulations will represent an extensive source of information for the writing of the next Assessment Report of the Intergovernmental Panel on Climate Change. Our main scientific objective will be to pin down physical and dynamical reasons behind differences in model representation induced by resolution change. Process-based assessment will focus on the representation of mean state, variability and teleconnections on a wide range of timescales.

1. **Research project**

Recent studies have established that the typical atmospheric and oceanic resolutions used for the CMIP5 coordinated exercise (Coupled Model Intercomparison Project, phase 5), i.e around 40km-150km globally, 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. The typical resolution that will be used during HighResMIP will reach about 20-50km. This is the first time ever such a largely demanding coordinated 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, including BSC participation with EC-Earth 3.2 model which corresponds to the latest available model version of the community model EC-Earth.

Within the framework of the european PRIMAVERA project funded by the H2020 program, BSC has developed 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. Accounting for the strongly reduced time step imposed by the higher resolution (6 minutes against 45 minutes at 1° resolution), this corresponds to an increase in computing resources by two orders of magnitude. To date, BSC has been able to run two years of simulation on marenostrum 3 and work in progress includes the development of suitable postprocessing set of tools to diagnose the large amount of output produced.

This project proposes to follow the entire HighResMIP protocol for coupled climate simulations (Tier 2; Haarsma et al, 2016) with this ground-breaking resolution. Its experimental protocol consists in running a 50-year spinup under perpetual 1950 conditions followed by: 1) a historical simulation covering the 1950-2050 period, 2) a control simulation under perpetual 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. This makes a total of 250 years of simulation. BSC will therefore contribute to HighResMIP with simulations at 3 different horizontal resolutions : a standard resolution version (ORCA1-T255, approximately 100km), a high resolution version (ORCA025-T511, approximately 25km) resolution and an extremely high resolution version (ORCA12-T1279, approximately 15km). As such, BSC will not only contribute with the highest resolution used for HighResMIP but also with the set of simulations allowing for the most robust assessment of the impact of increasing resolution worldwide. These simulations will represent an extensive source of information for the writing of the next Assessment Report of the Intergovernmental Panel on Climate Change which will take place in 2018. Given the substantial computational cost of this type of simulations (see sections 3 and 5) and the the small time-integration speed,, a 2-year period will be necessary to complete these simulations.

Relying on process-based diagnostics and metrics tools developed within PRIMAVERA, which supports most European participation to HighResMIP, our main scientific objective will be to pin down physical and dynamical reasons behind differences in model representation induced by resolution change. The HighResMIP experimental protocol focuses 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 an extensive range of scientific expertise which will allow focusing a correspondingly large range of processes that can be affected by resolution: sea ice dynamics and thermodynamics (Neven Fuckar and Juan Camilo), El Nino Southern Oscillation and tropical-extratropical teleconnections (Javier Garcia Serrano and Etienne Tourigny), 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).

2. **Methodology**

EC-Earth3 comprises three major components: IFS, NEMO and OASIS3. It is essential to configure and build a separate executable for each one of them. The resolution proposed (T1279: ~2140702 horizontal grid points, ORCA12: 13220998 horizontal grid points) will help efficiently share calculations between 2,521 processes, increasing the range of efficient compute-core usage per model executable. IFS and NEMO fully support a parallel

environment, while OASIS3 supports a pseudo-parallel environment. OASIS3 requires Cray pointers. 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).
The Integrated Forecasting System (IFS) is an operational global meteorological forecasting model developed and maintained by the European Centre of Medium-Range Weather Forecasts (ECMWF). The dynamical core of IFS is hydrostatic, two-time-level, semi-implicit, semi-Lagrangian and applies spectral transformations between grid-point space and spectral space. Vertically, the model is discretized using a finite-element scheme. A reduced Gaussian grid is used in the horizontal. The IFS cycle used in EC-Earth 3.2 if 36r4.
NEMO is an state-of-the-art modelling framework for the ocean, based in the Navier Stoke equations, used for oceanographic research, operational oceanography, seasonal forecast and climate studies. It is used by a large community with more than 1000 registered users (half in Europe, and half elsewhere) and since 2015 it has been used in more than 100 projects. The NEMO version used in EC-Earth 3.2 is v3.6 stable. The core of the NEMO model is OPA, it is a primitive equation model adapted to regional and global ocean circulation problems down to kilometric scale. Prognostic variables are the three-dimensional velocity field, a linear or non-linear sea surface height, the temperature and the salinity. In the horizontal direction, the model uses a curvilinear orthogonal grid and in the vertical direction, a full or partial step z-coordinate, or s-coordinate, or a mixture of the two. The distribution of variables is a three-dimensional Arakawa C-type grid. Various physical choices are available to describe ocean physics, so as various HPC functionalities to improve performances.
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), 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 hardware that best fits the needs of the model is a general purpose core. Due to lack of source code adaptations, any usage with an accelerator or any other computing device, would not take full advantage of these resources.
Related to the performance of the model, all the model components (IFS, NEMO and OASIS), are regularly benchmarked and analyzed using a methodology based on extracting traces using Extrae from real executions. These traces are displayed using Paraver software and processed to discover possible bottlenecks (Acosta et al 2016). As it is detailed in the corresponding chapter, BSC actively contributes to improve the performance of EC-Earth, having presence both in the EC-Earth Technical Working Group and in the NEMO System

Team HPC group. In particular, BSC has recently developed different optimizations for the EC-Earth coupling (Acosta, M.C. et al., 2016) and for the NEMO model (Tintó Prims, O. et al., 2015).
The Autosubmit software (Manubens-Gil et al., 2016), briefly described below, 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 if required, by Autosubmit to optimize the use of the machine and avoid collapsing the I/O system. The data storage and data transfer can be organized with a disk space of 3 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.
The Earth Sciences department has developed a suite of tools to evaluate and validate the model outputs against observational data from different input sources, that runs every time the results are generated. These tools includes R packages like "s2dverification" or Python tools like "Earth Diagnostics". This evaluation is completely integrated in the workflow, allowing early detection of any possible issues and therefore reducing waste of resources in those cases.

3. **Resource management**

As already mentioned above, we will produce a 50-year spinup under perpetual 1950 conditions followed by: 1) a historical simulations covering the 1950-2050 period, 2) a control simulation under 1950 conditions run for 100 years. The spinup will be initialised from the EN4 ocean climatology. Table 1 summarises the resources requested and detailed in section 5.

	Simulations (years)	Total core-hours (millions)	Total archive (TB)
Spinup	50	10.9	75
Control	100	21.9	150
Transient	100	21.9	150
Total	250	54.8	375

Table 1: Resources requested. The estimates have been made on the basis of a cost of 7h15 wallclock hours (by using 2,521 cores) and 1.5 TB of output per year of simulation with the T1279L91-ORCA12L75 configuration. These estimates have been obtained by running EC-Earth3.2 on MareNostrum III.

These experiments will follow the schedule indicated in Table 2. Using 2512 cores which is the basis taken to build Table 1, one month will run in about 7h15. Therefore more than 3 months will be necessary to complete 50 years of simulation.

	M1-M4	M5-M8	M9-M12	M13-M16	M17-M20	M21-M24
Spinup						
Control						
Transient						

Table 2: Approximate schedule of the experiments to be performed. M5-M8 stands for Month 5 to 8.

4. **Justification of Tier-0 needs**

Core number	SYPD	Core number	SYPD
768	0.16	5120	0.61
1024	0.23	5376	0.64
1280	0.28	5632	0.65
1536	0.33	5888	0.65
1792	0.37	6144	0.64
2048	0.40	6400	0.66
2304	0.42	6656	0.65
2560	0.47	7168	0.65
2816	0.52	7424	0.66
3072	0.52	7680	0.65
3328	0.54	7936	0.64
3584	0.56	8192	0.67
3840	0.60	8448	0.63
4096	0.60	8704	0.63
4352	0.59	8960	0.62
4608	0.61	9216	0.62
4864	0.61	9728	0.62
		9984	0.63

Table 3: NEMO 3.6 scalability on MN3 using an ORCA12 grid.

For this project, we will be running the most recent version of EC-Earth using the highest resolution coupled configuration currently available. Running EC-Earth with such configuration requires a system with high-level resources, which requires computational resources available only on a small subset of machines worldwide. Ec-Earth has already been deployed on MareNostrum III (a tier-0 system with 3056 nodes, 2x Intel SandyBridge-EP E5-

2670/1600 8-core processors at 2.6 GHz, 8x4GB DDR3-1600 DIMMS (2GB/core) per node and Infiniband FDR10 network) and within the previous PRACE projects "HiResClim", "HiResClim2" and "LSIHP".

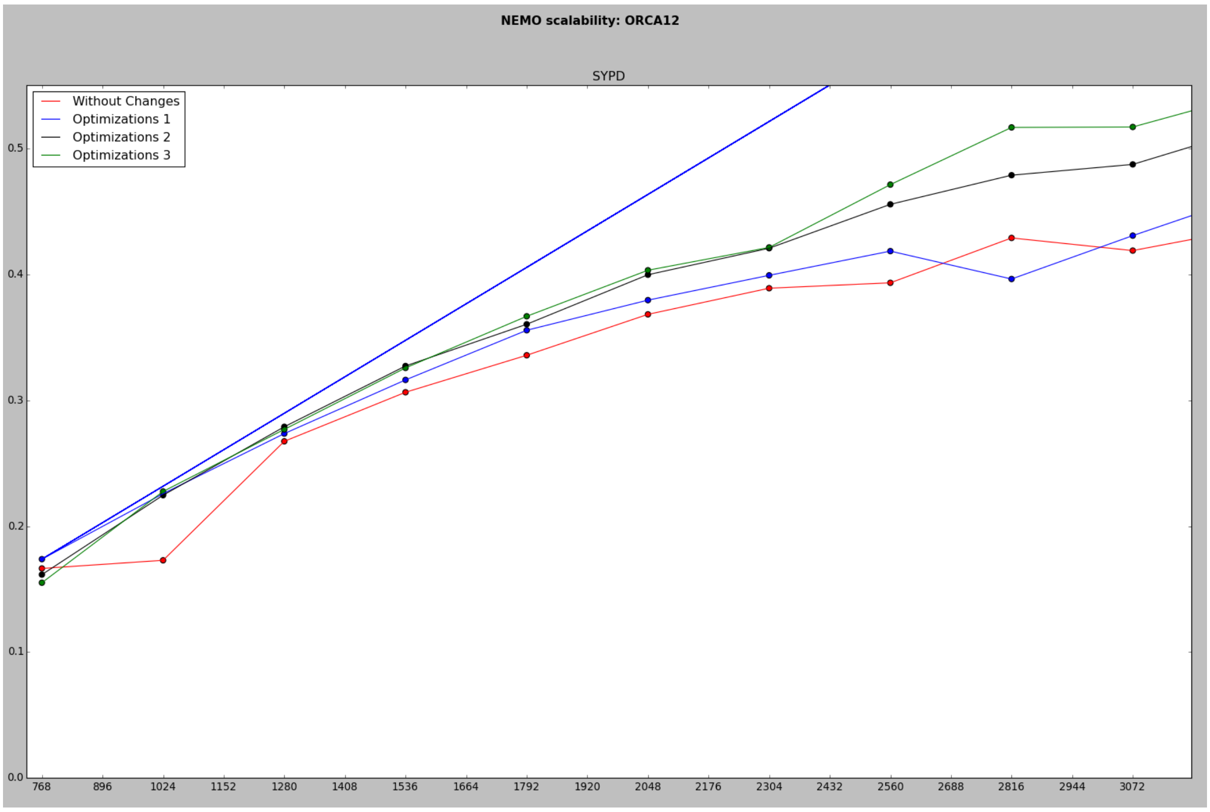

Figure 1. NEMO scalability using the proposed ORCA12 resolution, run on Marenostrum III supercomputer. For this experiment LIM sea-ice model was coupled every two timesteps. The different series correspond to version 3.6 stable, without and with the optimizations developed at the BSC, which are cumulative and have an aggregated impact.

The proposed configuration uses 2,521 cores, devoting 1,040 cores to the atmospheric component (IFS), 1,449 to the oceanic component (NEMO) and 32 cores for the I/O server (XIOS). We still have to add one more core for the Runoff Mapper, which uses a separate binary. From both component models, the slower one is NEMO. Figure 1 illustrates how the proposed assignation is efficient enough and ensures a throughput of more than 0.23 SYPD. In the coupled executions the ELPiN (Exclude Land Processes in NEMO) is used to find an optimal domain decomposition for NEMO with computation of only ocean subdomains, which significantly improves the throughput and efficiency, allowing to achieve the 0.28 SYPD (7h15 per month) specified in Section 5. Although the scalability plot exposed in figure 1 only represents the scalability of the ocean component, in the proposed configuration the ocean model is slower than the atmospheric and shows a worse scaling behavior, so the the whole coupled model can efficiently use more than 4000 cores.

5. **Justification of the amount of resources requested**

Table 4 lists the experiments described in section 3, which comprise 3 simulations of 50 or 100 years organized into 600 or 1200 chunks of 1 month length respectively. The two 100-year long simulations can be run independently and simultaneously depending on the machine load by Autosubmit. These experiments use the extreme high resolution configuration of EC-Earth3. Its cost, for a one-month simulation requires a wall-clock time of 7.15 hours. A benchmarking performed on MareNostrum III suggests that, taking into account the average load of its queues, optimum performance is obtained using 2,521 cores.

Run type	# Runs	# Steps/ Run	Walltime/ Step	#CPUcores	Total core hours (year 1)	Total core hours (year 2)	Total core hours/Type Run
Spinup	1	600	7h15	2521	10,966,350	0	10,966,350
Control	1	1200	7h15	2521	5,483,175	16,449,525	21,932,700
Transient	1	1200	7h15	2521	5,483,175	16,449,525	21,932,700
					21,932,700	32,899,050	54,831,750

Table 4: Cost of the experiments proposed.

The final estimate is for a total request of 58 million core-hours, which includes the numbers described in Table 4 plus a small buffer of 5% to account for failing jobs that will need to be repeated. The experiments will be run using Autosubmit, the launching and monitoring solution developed by the group of the applicant 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 estimates for the archive are an absolute upper value in case the automatic download does not perform as expected. It is very likely that this number will approach a figure ten times smaller.

6. **Data management plan**

The type of simulations conducted during this project requires hosting a set of configuration and initial conditions files, together with the code and namelists. This amounts to a negligible storage space requirements compared to the outputs to be generated. For each year of simulations, we estimated a total output of 1.5 Tb. These outputs will be post-processed on Marenostrum IV to compute climate indices and area-averaged quantities. Thanks to a well-organized workflow structure, the data produced by the integrations will be downloaded to the local BSC-ES servers within 24 hours after completion. Hence, a standard two-week delay between the end of the project and the closing of the accounts is more than enough to let us clean the HPC repositories.
The datasets will be stored in the BSC-ES data archive server, devoted to that purposes. The data will be managed and curated by the Data and Diagnostics Team of BSC-ES, which develops a framework to make available experiments and observational data to the BSC investigators. It will be be accessible through a THREDDS server, installed for this specific

purpose.

7. Past Experience

The BSC Earth Sciences Department has been involved in the IS-ENES2/PRACE-1IP working group focusing on the EC-Earth3 adaptation to Tier-0 machines. It has tested a range of EC-Earth3 configurations, in the atmospheric resolutions T255/511/799/1579 on several HPC systems: SGI Altix 3500, NEC-SX6, Linux cluster with Intel Xeon, Dell PowerEdge 2900, IBM pSeries 575 Power6 and IBM Power PC. Finally, members of the BSC Earth Sciences Department coordinated the recent HiResClim and HiResClim2 projects supported by PRACE. The PI of Glob15km has been PI of the recent PRACE LHSIP project.
The BSC Earth Sciences Department has the privilege to have an expert team on parallel model's performance, which is specialized in analyzing parallel programming model codes using cutting-edge performance tools. This advantage in resources and expertise allows the department to always have the latest performance results and therefore be able to determine the optimum configurations for EC-Earth, IFS and NEMO. Currently the BSC Earth Sciences Department is collaborator of the NEMO development team, and member of the NEMO HPC working group, providing the Consortium with performance reports and code optimizations. As a result of this collaboration two optimization branches were created in the NEMO code repository by members of the BSC Earth Sciences Department; both branches were merged into NEMO 3.6 stable version and trunk, and are being taken as base for further improvements in the communications within the model. Following the success of this cooperation, a similar collaboration has been established with the OpenIFS developers at ECMWF. To set up this partnership, a two month visit from Dr. Mario Acosta to ECMWF has been done from January to March 2017. During this time, BSC performance tools methodology has been applied for the analysis of the IFS model, leading to identify and apply two optimizations in the code. In parallel, an analysis to compare two different coupling approaches in EC-Earth (with and without OASIS coupled) has started recently.
Figures 2 & 3 illustrate an example of the performance tool's output. These are two different views provided by the Paraver tool from one single EC-Earth model execution, focused on its different components. Both images display the communications pattern along time, representing time on the horizontal axis, while the vertical correspond to the different processes executing the model. The different colours correspond to different MPI communication functions, except the light blue which corresponds to no communication. This Paraver view is very useful to determine the communications within the model and explain where the possible bottlenecks come from, especially when dealing with coupled models which have to communicate between components. Figure 2 focuses on the IFS component model, which has reserved most of the processors used for the execution. Simplifying, it can be said that the first half has less MPI communication, with more computation-only regions, while the second half contains a section performing lots of broadcasting message, which can impact the scalability of the code. Figure 3 focuses on the NEMO component model, which in this configuration takes the last 32 processors. The image clearly shows the importance of the communications in NEMO, which for a large percentage of time is not performing effective computations. It also shows that a big portion of this time is devoted to global communications, which are painted in pink colour. We know that those

collective communications belong to the horizontal diffusion routine, inside the ice model (LIM) used in NEMO, and they were one of the main targets of our optimizations, so this problem was solved to a large extent.

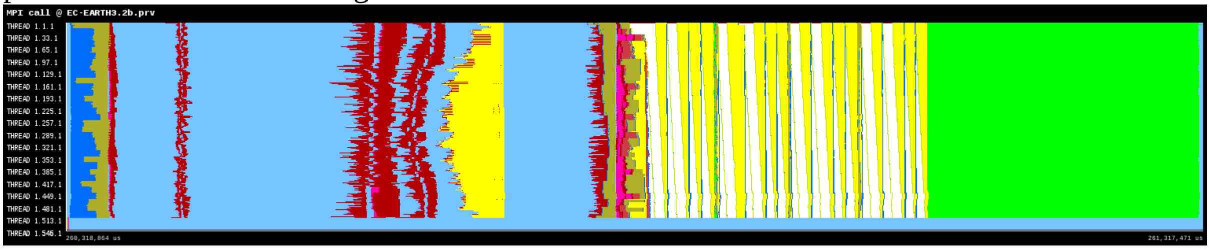

Figure 2: MPI call Paraver view of the IFS component model in an EC-Earth's model execution. The horizontal lines give the behaviour of the different processes (1 to 512) along time. Each colour corresponds to one different MPI communication function.

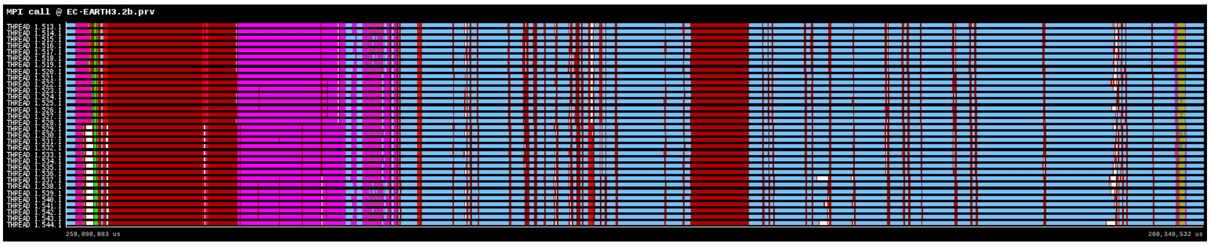

Figure 3: MPI call Paraver view of the NEMO component model in an EC-Earth's model execution. The horizontal lines give the behaviour of the different processes (513 to 544) along time. Each colour corresponds to one different MPI communication function.

The proposal is a new project that has not been submitted to any other HPC platform and that is not planned to be submitted elsewhere, to date.

8. Previous results and communication plan

The applying team has already assessed extensively the added-value of increased resolution using lower resolution simulations than the one planned to be used under the Glob15km project, namely increasing resolution from about 80km to about 40km in the atmosphere and 100km to 25km in the ocean globally. The added-value on climate prediction performance has been assessed within Prodhomme et al (2016). Most members of the applying team were also members of the HighResClim and HighResClim2 teams and therefore heavily involved in the assessment of the added-value of increased resolution on a variety of climate simulations including historical and climate change projections as well as climate predictions with various initialization strategies. Glob15km builds on previous experience from the applying team but aims at a leap forward with a substantial new increase in global resolution.
The outputs of model simulations produced during Glob15km will be made available publicly through an ESGF node, in the framework of CMIP6. These outputs will be exploited in multi-model analyses to be performed during the HighResMIP CMIP6 project as well as the PRIMAVERA project funded by the European commission through the H2020 Framework programme in support of HighResMIP. At least one publication in an international peer-

reviewed journal will be submitted by the BSC members based on the results from Glob15km. This article will provide a thorough assessment of the added-value of increasing the Ec-Earth3 resolution on the representation of mean state, variability and teleconnections. The Glob15km results will also be made publicly available through a wiki website which will summarize the experimental setup and provide the key figures illustrating the main results. These results will be widely presented in scientific conferences including PRACEdays and will feature at the meetings of the Scientific Steering Group (SSG) CLIVAR (Climate and Ocean Variability Predictability and Change) project, of which Virginie Guemas is a member. Finally, the results will be presented to the PRIMAVERA consortium during its General Assembly. These dissemination activities will be supported by the PRIMAVERA funding for human resources and travels.

9. **Additional needs**

The use of NCO, CDO, TOTALVIEW and NCVIEW is required to allow debugging in case a problem arises on the HPC. CDFTOOLS and Python (with a set of identified modules) will be essential for postprocessing and data reduction before transfer to the local storage space (in this sense bandwidth will be required to move data from the HPC facility to our storage facilities).

10. **Reviewers** (optional)

None

11. **Track record**

Dr Virginie Guemas is currently Head of the Climate Prediction Group within the Earth Sciences department at BSC. She is member of the WCRP (World Climate Research Program) CLIVAR (Climate and Ocean Variability, Predictability, and Change) SSG (Scientific Steering Group), principal investigator (PI) of seven European projects funded under the FP7 (PREFACE), H2020 frameworks (IMPREX, APPLICATE, INTAROS), the European Space Agency (CMUG2) or Copernicus (C3S-MAGIC), one MINECO-funded project (HIATUS), one PRACE-funded project (LSHIP) which final report is due by 10th June 2017, one ECMWF-funded project (HighResMIP_BSC) and she is WP leader in the H2020 PRIMAVERA project. She is also the tutor of 2 Marie-Sklodowska-Curie grants (NetNPPAO, SPFireSD). She was a contributing author to the IPCC (Fifth Assessment Report) and has participated in 14 EU FP7 and national projects. She is author of 46 articles published in international peer-reviewed journals, among which 17 as first author, 1 in Sciences, 1 in Nature Climate Change, 1 in Nature Communication and 5 in Bulletin of the American Meteorology Society. She has supervised 1 PhD student in the past and she is currently supervising 2 PhD students and 12 postdoctoral scientists.

Mr Kim Serradell is currently the Head of the Computational Earth Sciences group within the Earth Sciences department. The CES group is a multidisciplinary team of 19 members with different IT profiles that interacts closely with all the other groups of the Earth Sciences Dept. The group has among its tasks providing help and guidance to the scientists with the technical issues related to their work and developing a framework for the most efficient use

of HPC resources. Kim Serradell is the PI of the ESIWACE European project funded by the H2020 program and which focuses on porting and optimizing the code on various platforms.
Mr Miguel Castrillo is an expert on performance profiling and optimization applied to NEMO & EC-Earth with a strong experience in software development. He is coordinator of the Models & Workflows team within the Earth Sciences department, and has experience in deploying and running the NEMO model in several platforms. He is part of the Technical Issue Working Group for EC-Earth and member of the NEMO model HPC Working Group.
Mr Pablo Echevarria has a master in computer science, the final work was a implementation of a Data assimilation cycle using LETKF and NOAA WW3 wave model. He worked since 2012 in modelling and HPC in R&D in the national weather service of Argentina. He is currently member of the Computational Earth Sciences group within the Earth Sciences department. He is the responsible of the auto-ecearth / auto-nemo project.
Dr. Mario C. Acosta is a postDoctoral fellow in the Computational Group of the Earth Sciences Department at BSC. He is the leader of the performance team within the Earth Sciences department and the supervisor of one PhD student and one master student. His research interests and expertise include a wide knowledge in numerical models (governing equation, numerical algorithms and computational implementation), performance analysis to highlight the main bottlenecks of the models and how to adapt and optimize them efficiently to current and new High Performance Computing Platforms. He is author of 4 peer-reviewed articles in international journals and one member of the group of experts of European eXtreme Data and Computing Initiative: Weather, Climatology and Solid Earth Sciences (EXDCI/WCES).
Dr. Etienne Tourigny is a postdoctoral fellow with a strong background in atmospheric and computer sciences. He is currently working for the Climate Prediction Group of the Earth Sciences Department at BSC and has been awarded a Marie Skłodowska-Curie Individual Fellowship (SPFireSD, seasonal prediction of fire risk). He is very active in development of the high resolution version of the coupled EC-Earth model, developed within the PRIMAVERA project and has been working on workflow and performance aspects of the EC-Earth model. He is also involved in the EC-Earth Technical and Climate Prediction Working Groups.
Dr. Louis-Philippe Caron is a Juan de la Cierva postdoctoral fellow and co-leader of the Climate Prediction group of the BSC Earth Sciences Department. He is also a member of the seasonal Tropical Cyclone Prediction Panel of the Working Group on Tropical Meteorology Research (WGTMR). He is an expert on hurricane activity, their representation in climate models as well as their variability and predictability at seasonal to decadal timescales. He is involved in a number of European projects, including PRIMAVERA, and works in collaboration with scientist in the insurance industry to develop applications relevant to this sector of activity.
Dr. Eleftheria Exarchou is currently a postdoctoral scientist in the Climate Prediction group within the Earth Sciences Department. She has a strong background in physical oceanography and in ocean climate modeling. She has experience in seasonal forecasting and climate projections, with a particular focus on ocean dynamics. She has investigated the model biases of the coupled EC-Earth climate model in the Tropical Atlantic within the PREFACE project. She is involved in the EC-Earth Ocean working group, which, among other tasks, coordinates the tuning of the ocean model component of EC-Earth. She will

investigate the simulations that will be performed with the high resolution version of EC-Earth within the PRIMAVERA project in order to address the involvement of the ocean heat uptake on surface climate changes.

Mr Pierre-Antoine Bretonnière holds a Masters Degree in "Mathematical and Mechanical Modelling" from the Matmeca engineer school in Bordeaux (France). Graduated in 2010, he has worked in several climate research institutes (CERFACS - Toulouse - France, Catalan Institute of Climate Sciences - Barcelona - Spain and the Earth Sciences Department of the Barcelona Supercomputing Center). His work focuses on climate models outputs and diagnostics, data management and model coupling, and is in charge of the Data and diagnostics team of the Earth Sciences department. He was the person in charge of the data management plan and data conventions definitions in the SPECS FP7 project and has participated in several other European projects. He is also involved in the Research Data Alliance (RDA) framework as chairman of the "Weather, climate and air quality" interest group.

Dr. Neven S. Fučkar is a Juan de la Cierva Postdoctoral Fellow in the Climate Prediction Group of the Department of Earth Sciences at the Barcelona Suprecomputing Center. He obtained a Ph.D in Atmospheric and Oceanic Sciences at Princeton University, Princeton, NJ, USA. Dr. Fučkar research focus is on sea ice and climate dynamics and predictions from seasonal to decadal time scales. He is Principal Investigator of 3 national (RES) supercomputing projects and the convenor of Polar Climate Predictability and Prediction session at the EGU general assembly in Vienna. Dr. Fučkar is author or co-author of 16 manuscripts in international peer-reviewed journals, among which 1 in Nature Geoscience and 3 in Bulletin of the American Meteorology Society (with total of 527 citations on Google scholar, and h-index and i10-index of 9).

Dr Martin Ménégoz is an expert in aerosol and climate sciences. After a PhD in atmospheric physics and chemistry that he got at Météo-France in 2009, he investigated the aerosols interactions with the climate system during several contracts with French CNRS institutes. He conducted both model developments and field campaigns to understand how the atmospheric pollution affects the cryosphere. Now at BSC, he leads the research line devoted to stratosphere-troposphere coupling processes in the Earth Sciences Department. He is largely involved in the development of the EC-Earth European model that he applies in initialised simulations to make quasi-operational forecasts at seasonal-to-decadal timescales. In addition to his strong involvement in several European projects, he coordinates the VOLCADEC project (MINECO, 2016-2018, 168Keuros) to investigate the climate variability to stratospheric aerosols and to volcanic eruptions. He has an H-index of 9 (Google Scholar), with 17 articles in peer-review journals (6 as first author), 3 articles in revision and 500 citations. He presented his results trough 37 conferences and international project meetings.

Dr Chloé Prodhomme is an expert in climate modelling and ocean-atmosphere, land-atmosphere interactions. Her PhD took place at the LOCEAN-IPSL, in collaboration with the University of Tokyo on the interactions between the Indian Monsoon and tropical Sea Surface Temperatures (SSTs). After her PhD defense, she joined the team of the Pr. Francisco Doblas-Reyes as a post-doctoral fellow, first at the Institut Català de Ciències del Clima (IC3) and then at the Barcelona Supercomputing Center (BSC). In these two positions, she has been working on the sources of skill at seasonal time scale. She has been in several

European Projects: SPECS (FP7), PREFACE (FP7), IMPREX (H2020), APPLICATE (H2020), QA4SEAS (Copernicus) and MAGIC (Copernicus). In this context, she managed to create international collaborations with more than 10 European Institutions. She is the author of 9 peer review papers (4 as first authors, all of them first quartile).

Mr Oriol Tintó Prims is a PhD candidate at the Earth Sciences department from the Barcelona Supercomputing Center in collaboration with the Computer Architecture and Operating Systems department from the Universitat Autònoma de Barcelona with the specialization of High Performance Computing. He received his degree in physics from the Universitat Autònoma de Barcelona and his Masters degree in Modelling for Science and Engineering from the Universitat Autònoma de Barcelona. He has experience in analysis and optimization of Earth System models in HPC infrastructures actively collaborating with the NEMO development team and the NEMO HPC group.

12. **References**

Acosta, M.C., X. Yepes-Arbós, S. Valcke, E. Maisonnave, K. Serradell, O. Mula-Valls and F.J. Doblas-Reyes, 2016: Performance analysis of EC-Earth 3.2: Coupling. BSC-CES Technical Memorandum 2016-006, 38 pp.

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.

Haarsma RJ, Roberts M, Vidale PL, Senior CA, Bellucci A, Corti S, Fučkar NS, Guemas V, von Hardenberg J, Hazeleger W, Kodama C, Koenigk T, Leung LR, Lu J, Luo JJ, Mao J, Mizielinski MS, Mizuta R, Nobre P, Satoh M, Scoccimarro E, Semmler T, Small J, von Storch JS, 2016, High Resolution Model Intercomparison Project (HighResMIP). Geosci. Model Dev. Discuss., 9, 4185-4208, doi:10.5194/gmd-9-4185-2016.

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

Manubens-Gil, D., J. Vegas-Regidor, C. Prodhomme, O. Mula-Valls and F. J. Doblas-Reyes, 2016, Seamless management of ensemble climate prediction experiments on HPC platforms, International Conference on High Performance Computing & Simulation (HPCS), Innsbruck, 2016, pp. 895-900. doi: 10.1109/HPCSim.2016.7568429

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

Prodhomme, C., L. Batté, F. Massonnet, P. Davini, O. Bellprat, V. Guemas, and F.J. Doblas-Reyes (2016). Benefits of increasing the model resolution for the seasonal forecast quality in EC-Earth. Journal of Climate, 29, 9141-9162, doi:10.1175/JCLI-D-16-0117.1.

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.

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

Tintó Prims, O., M. Castrillo, K. Serradell , O. Mula-Valls and F.J. Doblas-Reyes, 2015: Optimization of an ocean model using performance tools. BSC-CES Technical Memorandum 2015-002, 16 pp.

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.

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.

Appendix – CV of the Principal Investigator with list of publications (<= 5 years)