PRACE Final Report Form – Project Access

1. General information

1.1. Proposal ID

2012060992

1.2. Type of proposal granted: Project Access

1.3. Period of access to the PRACE facilities

Extended access: 1st November 2013 – 31st March 2014

1.4. Name of the PRACE facility assigned

MareNostrum III Barcelona Supercomputing Center, Spain

1.5. Name of the Principal Investigator

Dr. Colin Jones, Swedish Meteorological and Hydrological Institute,
report written by successor Dr. Ralf Döscher, Swedish Meteorological and Hydrological Institute

2. Project information

2.1. Project name

HiResClim: High Resolution Ensemble Climate Modeling

2.2. Main research field (in brackets the corresponding ERC fields)

- Economics, Finance and Management (SH1, SH2)
- Linguistics, Cognition and Culture (SH3, SH4, SH5, SH6)
- Biochemistry, Bioinformatics and Life sciences (LS1, LS2, LS8, LS9)
- Physiology and Medicine (LS3, LS4, LS5 LS6, LS7)
- Mathematics and Computer Sciences (PE1, PE6)
- Fundamental Physics (PE2, PE3)
- Chemical Sciences and Materials (PE4, PE5)
- Engineering (PE7, PE8)
- Universe Sciences (PE9)
- Earth System Sciences (PE10)

2.3. Institutions and research team members

Laurent Terray, CERFACS, France, Sophie Valcke, CERFACS, France, Eric Maisonnave, CERFACS, France, Christophe Cassou, CERFACS, France, Klaus Wyser, Swedish Meterological and Hydrological Institute (SMHI), Sweden, Uwe Fladrich, Swedish Meterological and Hydrological Institute (SMHI), Sweden, Muhammad Asif, Catalan Institute of Climate (IC3) Sciences, Spain, Domingo Manubens, Catalan Institute of Climate Sciences (IC3), Spain, Francisco Doblas-Reyes, Catalan Institute of Climate Sciences (IC3), Spain, Chandan Basu, Linkoping University, Sweden, Torgny Faxen, Linkoping University, Sweden, Wilco Hazeleger, Royal Netherlands Meteorological Institute (KNMI), Netherlands, Richard Bintanja, Royal Netherlands Meteorological Institute (KNMI), Netherlands, Camiel Severijns, Royal Netherlands Meteorological Institute (KNMI), Netherlands

2.4. Summary of the project *(Maximum 500 words)*

Please fill in the field with the same text used in the application form.

HiResClim aims to make major advances in the science of climate change modelling. This will be achieved by addressing the dual requirements of; increased climate model resolution and increased number of ensemble realizations of future climate conditions for a range of plausible socio-economic development pathways. Increased model resolution aims to deliver a significant improvement in our ability to simulate key modes of climate and weather variability and thereby provide reliable estimates of future changes in this variability. A large ensemble approach acknowledges the inherent uncertainty in estimating long-term changes in climate, particularly in phenomena that are highly variable and, of which, changes in the occurrence of the rare but intense events are those impacting society and nature most strongly. To provide credible risk assessment statistics on future change in phenomena such as; extra-tropical and tropical cyclones, heatwaves, droughts and flood events, the combination of high climate model resolution and a large ensemble approach is unavoidable. In HiResClim we attack both of these requirements in a balanced approach, which, as well as being the most efficient way to utilise the most advanced HPC systems of today, is also the only path to providing more robust and actionable estimates of future climate change.

2.5. Description of the results obtained from the scientific point of view, future perspectives, benefits to our society, and the benefits of using computer resources. *(Maximum of 1000 words)*

The consortium carried out simulations for the final purpose to improve our capability to perform initialized climate predictions on time scales from months to a decade. This can be achieved by improving the numerical resolution, increasing the number of relevant ensemble members, identifying necessary processes, exploring the impact of different initialization procedures and by generally applying constantly improved and adjusted climate models.

While HighResClim is largely approaching the problem by increasing resolution, the coupled models and the component models need to be adjusted when applied in new set-ups and specific time scales. As an example, the sensitivity of the orographic gravity wave drag (GWDSE=0.01, 0.001 and 0.0001) in the ARPEGE model was explored. Fig. 1 explains the outcome and the choice of GWDSE=0.001 for decadal runs.

Important questions to high resolution climate prediction include:

- How does small-scale SST in the Gulf Stream region impact the atmosphere locally and remotely?
- Does the coupling affect the small-scale air-sea interaction?

CERFACS carried out coupled and atmospheric forced high-resolution experiments to answer these questions and the most relevant conclusion are:

- small-scale air-sea interaction generates an increase by 30% of turbulent heat fluxes through SST-induced wind-speed anomaly and MABL destabilization.
- Small-scale SST in the Gulf Stream region have a remote impact on winter storm tracks, especially on intense Northern Mediterranean storm tracks.
- Coupled experiments have shown that there are no significant differences between the coupled model and the forced mode with respect to the small-scale interaction between SST and wind-speed at 10 m.

IC3, SMHI and KNMI are using the EC-Earth model in different configurations. HighResClim is most interested in the effects of increased model resolution in the ocean and atmosphere on our predictive capacity with respect to the predicted mean climate and monthly-seasonal variability and extremes. Different model set-ups with standard and high resolution grid (EC-Earth3 T255/ORCA1 and T511/ORCA025) are initialized using ERA-Interim reanalysis and the GLORYS initial conditions for the month of May, and prediction skill for the following three months (June, July, August, JJA) is examined. The test period covers 1993-2009. As a result, the higher resolution case gives distinctly improved mean conditions, in particular over the North Atlantic and North Pacific Ocean. Fig. 2 shows the systematic error of the Sea Surface Temperature (SST) reflecting the improvements.

Oscillating teleconnection patterns and the model description of associated mechanism are key ingredients for a skilful prediction. The ENSO phenomenon can recently be predicted by seasonal prediction systems with non-negligible skill. Fig. 3 documents the RMS error of El-Nino(3.4) SST compared to reanalysis and observational products as well as the ensemble spread. All those indicators taken from the EC-Earth predictions show clear skill improvement with enhanced resolution. The ENSO skill can further be enhanced by a stochastic physics approach in combination with low resolution. In a 10-member ensemble, best prediction skills are achieved by high resolution and no stochastic physics (Fig. 4).

Another teleconnection pattern of key importance for European climate is the North Atlantic Oscillation (NAO). Predictions of the winter (DJF) NAO with EC-Earth 3 in low and high resolution set-ups and with the ECMWF seasonal prediction system 4 started each year in the month of November over the period 1993-2009 with ERA-Interim and GLORYS data as initial conditions and as five-member ensembles. While EC-Earth 3 in the low resolution shows a correlation close to zero, EC-Earth in the high resolution configuration gives the clearly best anomaly correlation (AC=0.36) and beats the current ECMWF seasonal prediction system 4 (AC=0.28) (Fig.5).

Experiments for climate prediction can be carried out following the principle of full-field initialization or the anomaly initialization. Predictions described above are performed by full-field initialization, which comes with the advantage of not prerequiring the simulated mean climate. The method of anomaly initialization implies knowledge of the simulated mean climate conditions, thus depending on spin-up and historical simulations, which are utilized to generate a mean climate. Initialization fields of e.g. ocean potential temperatures and sea ice concentrations are derived from observed anomalies and applied to simulated anomalies. The advantage of the anomaly method is seen in the fact that the initial conditions are consistent with the coupled model's preferred mean state (the model's attractor) hence diminishing the model drift towards its own climatological state. Initialization shocks are thereby largely avoided.

In order to establish mean climate conditions, a spin-up of the EC-Earth model in T255-ORCA1-LIM3 configuration has been carried with a length of about 400 years. Time series of ocean heat content in different depth intervals are shown in Fig. 6. The heat content of the deep ocean does not stabilize in the entire simulation although the rate of change decreases slightly with the simulation length. Ocean states picked from this spin-up/control simulation were later used for the initialization of the ocean in transient climate simulations, which in turn form the base for anomaly-initialized prediction ensembles.

Despite serious and remarkable problems in running high resolution model configurations, analysis of short simulations lead the EC-Earth group to develop strategies to deal with the unavoidable mismatch between the ocean and atmosphere grids in high latitudes, especially at very high resolutions, which required correcting the coupling strategies. This work would not have been possible without the HiResClim resources and also benefits the numerical stability standard resolution experiments.

2.6. Expected future work in the area

A second year of the project is currently ongoing, framing a more ambitious set of simulations in the seamless climate paradigm, i.e. covering a range of time scales from seasons to a decade. HighResClim 2 is focusing on spin-up, historical simulations and seasonal and decadal predictions in very high resolution. Work in that project and further on will focus on the specific effects of improving resolution, initialization methods and process descriptions relevant for lifting prediction skill on a range of prediction time scales.

2.7. Images of the results including description or caption *(Minimum resolution of 300 dpi)*

Seven figures are attached to the report.

2.8. Publications or reports regarding the developed project

(Format: Author(s). "Title". Publication, volume, issue, page, month year)

n.a.

The project has ended just a few weeks ago and results have not yet been published in scientific journals yet.

2.9. Patents registered in relation with the developed project
(Format: patent identifier, title and description) (Maximum of 850 characters)

n.a.

2.10. Name and Surname of the students that have deployed their thesis collaborating in the developed project and title of the thesis.

n.a.

2.11. Talks given in the area of the project

- U. Fladrich: "EC-Earth 3" at the Second Workshop on Coupling Technologies for Earth System Models (CW2013)
- Francisco J. Doblas-Reyes: "The HiResClim and SPECS Projects" at the Exascale Technologies & Innovation in HPC for Climate Models conference (Hamburg, 17-19 March 2014)
- Marie Piazza: "High resolution modeling of small-scale air-sea interaction in the Gulf Stream region and impact on North Atlantic storm-tracks", EGU 2014
- Laurent Terray & Sophie Valcke: CERFACS: "Small-scale ocean-atmosphere interaction: does the coupling matter? DRAKKAR/MYOCEAN 2014 Annual Workshop (http://lgge.osug.fr/meom/Events/Drakkar2014/)
- Sophie Valcke: First results of CERFACS HR coupled model (ARPEGE T359 - NEMO 3.4 ORCA025 – LIM2), CONVERGENCE kick-off (French ANR project)
- M.Caian, K. Wyser, LP Caron, C. Jones (SMHI Rossby Centre): ,2013, April: "Improving anomaly initialisation for decadal prediction" EGU 2013.

2.12. Other information (URLs, logos, photographies, etc.)
Please fill in the box with the regarding information and attach the photographies or logos to this form.

n.a.

2.13. Any further funding obtained as a result of the developed project

No further funding has been obtained, although most of the EU projects linked to this proposal (IS-ENES2, SPECS) have already started. Partners are involved in ongoing proposals to the EU's Horizon 2020 research programme.

3. Feedback and technical deployment

3.1. Feedback on the centers/PRACE mechanism *(Maximum 500 words)*

- A uniform utilisation of the resource over the allocation time, as preferred by PrACE, does not seem a realistic assumption. This demand, though understandable, conflicts with the computational requirements of the project's scientific experiments (spin-up runs followed by historical and ensemble experiments), the need for a technical start-up phase (e.g. analysis and tuning of computational performance with respect to the specific experiments) and the unavoidable appearance of technical problems. In our case problems were intensifying due to the platform availability first about 10 weeks after the start of the project (see 3.2).
- We experience repeated problems with unexplained hanging for specific configurations for very high resolution. The project had to focus on capacity configurations rather than on capability in order to identify efficient set-ups with reasonable speed rather than maximum set-ups.
- The complete work-flow of the use of coupled climate models should be taken into account and be reflected in technical provisions. This includes the interaction of computationally intensive experiments with post-processing phases and data storage.
- A better appreciation of the technical needs for coupled climate models is needed. These models constitute complex software and an involved development process. Hence, it is not desirable to adapt the models to a single purpose or a single computing platform. Scientific flexibility is a key issue for climate models, which implies rather bulky software, complex runtime environment, and diverse work-flow.
- A number of nodes or a dedicated machine with large memory is needed to deal with post-processing of the output, specially for rebuilding the high resolution ocean data.

3.2. Explanation of how the computer time was used compared with the work plan presented in the proposal. Justification of discrepancies, especially if the computer time was not completely used. *(Maximum 500 words)*

The project spent about 36 Mcore hours out of a granted allocation of 38 Mcore hours (~95 %) during this project until March 31st. Under the conditions of substantial technical and infrastructure issues as described below, HighresClim used the given allocation adequately.

The HighResClim project was granted starting 1st November 2012, however, access to the MareNostrum III machine could not be provided by BSC until the beginning of 2013, which initially delayed the project start for about 10 weeks. A number of issues have appeared in the course of the activities. All of them have been worked on with the help of the BSC PrACE support, but some remained and hindered to forge ahead with the actual experiments of the project.

Activities around the ARPEGE-NEMO climate model were focusing on physically adjusting the high-resolution model (ARPEGE Climat V5.3 T359L31, NEMO V3.5 ORCA025+LIM2, coupled with oasis3-mct) and its component models. This effort is a necessary condition to achieve skilful climate prediction capability. Examples are given in section 2.5. It was also necessary to optimize the coupled set-up with respect to load balancing. As a result, CERFACS gets an overall coupled model performance of 2.8 simulated-years/day using 4000 cores-hours/simulated year. Additional scalability studies on MN3 were helpful to understand and improve the performance. Interaction with the PRACE support was needed to solve bugs in the installation (e.g. for the parNetcdf installation). ARPEGE-NEMO simulation were also run with climatological sea-ice for an analysis of the impact of high-resolution on small scale interactions between SST and surface wind with results discussed in section 2.5.

After an extension of the HighResClim project was granted starting 1st November 2013, spin-up and historical experiments in the new EC-Earth standard resolution (T255ORCA025-LIM3, high compared to other climate models) have been carried out, as well as first high resolution T511-ORCA025-LIM2 seasonal hindcast ensembles.

Moving further to multi-annual and longer experiments in high resolution (T255ORCA025) resulted again in issues. Again, all of them have been worked on with the help of the BSC PrACE support, with limited success. Here we list the most severe type of problems:

- **Computational speed and work flow of the coupled model**

The achieved performance of the high resolution set-up of EC-Earth (T511ORCA025) did not surpass about 2.5 SYPD (*simulated years per day).* The actual SYPD metric (ASYPD), which includes queue waiting time, as well as necessary post-processing and data transfer was much lower (Fig. 7). Climate models can be run efficiently only with well dimensioned post-processing and data transfer. Identified bottlenecks in long queue waiting times, limited resources for post-processing, and insufficient data transfer rates do further limit the model efficiency. Predicting the model speed at MN3 has thus shown to be a difficult task.

- **Performance analysis of the coupled model**

In an attempt to improve the model performance on MN3, the project received valuable support from the PrACE team at BSC, nevertheless, this interdisciplinary type of work is time consuming, and the complexity of running a climate model requires a significant initial effort from the PrACE support staff. Furthermore, tools for computational performance analysis are not necessarily suitable for full-scale problems. In practise, performance analysis results were limited to few (64) cores of single model components, which deviates from the multi-executable model on many cores, that showed the actual performance problems.

- **Frequent, temporary service disturbances and disruptions**

The project has experienced, over the complete duration, system malfunctions of many kinds. One problem with these interruptions is that they can not easily be identified as system issues, which often triggers time-consuming investigations about the model configuration. Another problem is that such events disrupt the complex work-flow of the modelling experiments. For example, file system issues have caused the model to crash a number of times. Recovering from such a failure requires a manual procedure, including the inspection and relocation of a large number of restart files, as well as a reconfiguration of the experiment. These problems are exacerbated for multi-member ensemble experiments, which rely on automatic re-submission of jobs.

- **MPI**

It turned out that Intel MPI, in its default configuration, was de-functional when used for the coupled model (multi executable) on a number of cores higher than 500. BSC has proposed a modified configuration, by means of environment variables, which allows to run some of the intended experiments. Nevertheless, the high-resolution configuration of EC-Earth 3 in particular was limited to core numbers smaller than 615 (shared by the atmosphere, ocean, and coupler components), and currently smaller than 2000. OpenMPI proved not to be a suitable alternative due to substantially reduced (by a factor of two) computational performance. This was a main obstacle for running the high-resolution spin-up runs as planned in the proposal.

- **GPFS issues**

A number of runs experience occasional (random) patterns of severe performance degradation, which are, according to BSC information, associated with GPFS problems. This affects the experiment work-flow as jobs are running out of time and fail before completion. In particular, the number of files created by the climate model seems to negatively affect performance and stability of the file system for the planned model runs.

- **Post-processing facilities**

Another bottleneck was experienced due to limited large memory availability to deal with post-processing of the output, specially for the high resolution ocean data.

As a result of the listed issues on MareNostrum III, the spin-up runs have been delayed substantially and consumed more resources in comparison to the HiResClim application. Consequently, less computer time has been used for applications that were originally planned towards the end of the project, such as e.g. high resolution historical simulations and predictions.

Our original 2-year proposal envisaged the first year primarily consisting of pre-industrial spin-up runs of EC-Earth and ARPEGE-NEMO in standard and high resolution. While we have proceeded with the standard resolution spin-ups, the high resolution spin-ups proved to be unfeasible given the technical problems described above and the time required to scientifically tune the coupled systems.

Output from the lower resolution EC-Earth spin-up runs was used to initialize equivalent-resolution 20th century and projection/prediction experiments on partners national HPC systems. Following this strategy we maximized the outcomes of HighResClim year 1.5 and, are able to address a number of important questions with respect to the role of both ocean and atmospheric model resolution. This includes simulating future anthropogenic-induced climate change and seasonal and decadal climate prediction. Furthermore, we will use the lower resolution spin-up runs to investigate methods to accelerate the production of spun-up pre-industrial states for use in the high-resolution coupled model.

As the full-field initialized seasonal and decadal predictions (originally planned for year 2 of HiResClim) could be started before pre-industrial spin-ups completed, those simulations were moved forward to the extension period (3rd half year) of the project, allowing some of the year 2 allocation to be reserved for extending the high-resolution EC-Earth spin up run. Such a rescheduling was only feasible for the full-field initialization technique.

3.3. Please, let us know if you plan to apply for a regular PRACE project? If not, explain us why. *(Maximum 500 words)*

A second-year proposal has been granted to the 7th PrACE tier-0 open call. We anticipate further applications in the field of seamless climate prediction using updated versions of high resolution climate models.