APPLICATION FORM
DECI 10 call

General Information

Type of proposal	DECI-10 PROJECT
Start date	1st May 2013

Project Title:	**Seasonal Prediction of the Arctic Ice and Tropical Atlantic Cyclones**
Project acronym:	**SPAITAC**
Research field	Earth Sciences and Environment

Project leader (personal data and contact)

Gender	Male
Title	ICREA Research Professor
First name	Francisco
Last name	Doblas-Reyes
Initials	F. J. D.-R.
Date of birth	June 7th, 1968
Nationality	Spanish
e-mail	francisco.doblas-reyes@ic3.cat
Phone number	+34 93 567 99 76
Fax number	+34 93 309 76 00

Project leader (organisation and job title)

Job title	Climate Forecasting Unit Team Leader
Organisation name	Institut Català de Cienciès del Clima (IC3)
Group	Climate Forecasting Unit (CFU)
Department	
Address	Carrer del Doctor Trueta, 203
City	Barcelona
Postal code	08005
Country	Spain

Contact person for all correspondence (if different from above)

Name	Louis-Philippe Caron

E-mail	caron@misu.su.se

Organisation name	Stockholm University
Department	Meteorology
Address	
City	Stockholm
Postal code	106 91
Country	Sweden
Job title	Post-doctoral fellow
Gender	Male
Title	Dr.
First name	Louis-Philippe
Last name	Caron
Initials	
Date of birth	July 31st, 1978
Nationality	Canadian
e-mail	caron@misu.su.se
Phone number	+46 727 733 549
Fax number	+34 93 309 76 00

Collaborators (add more as required)

Organisation name	Swedish Meteorological and Hydrological Institute (SMHI)
Department	Rossby Centre
Address	Folkborgsvagen 1
City	60176
Postal code	Norrköping
Country	Sweden
Job title	Head of Research Group
Gender	Male
Title	Dr.
First name	Colin
Last name	Jones
Initials	G.
Date of birth	13th October 1967
Nationality	British
e-mail	colin.jones@smhi.se
Phone number	+46 11 495 8032
Fax number	+46 11 495 8001

Organisation name	Institut Català de Cienciès del Clima
Department	Climate Forecasting Unit
Address	Carrer del Doctor Trueta, 203
City	08005
Postal code	Barcelona
Country	Spain
Job title	Researcher
Gender	Female
Title	Dr.
First name	Virginie
Last name	Guemas
Initials	
Date of birth	November 21st, 1983
Nationality	French
e-mail	vguemas@ic3.cat
Phone number	+34 93 567 99 76
Fax number	+34 93 309 76 00

Organisation name	Swedish Meteorological and Hydrological Institute (SMHI)
Department	Rossby Centre
Address	Folkborgsvagen 1
City	60176
Postal code	Norrköping
Country	Sweden
Job title	Research Scientist
Gender	Male
Title	Dr.
First name	Klaus
Last name	Wyser
Initials	
Date of birth	November 15th, 1965
Nationality	Swiss
e-mail	klaus.wyser@smhi.se
Phone number	+46 11 495 8464
Fax number	+46 11 495 8001

Abstract of the project: (Maximum 500 words):

SPAITAC: Seasonal Prediction of the Arctic Ice and Tropical Atlantic Cyclones

The Arctic Ocean has experienced a sharp decline in sea ice extent and thickness in the recent decades, which threatens to leave the Arctic ice-free during the summer months in a foreseeable future. Insights of pan-Arctic sea-ice cover a few months in advance would be of key interest for the marine accessibility of the Arctic Seas and for the Arctic population whose livehoods depend on fishing and hunting. In parallel, the tropical Atlantic is currently experiencing a remarkably high level of hurricane activity: the last three hurricane seasons have produced at least 17 tropical cyclones, making it the most active 3-year period since the beginning of the hurricane record in 1851. The recent passage of Sandy over New York serves as a powerful reminder to the incredible destructive power these storms can wreck on exposed populations.

If both Arctic sea ice and Atlantic hurricanes show a clear trend over the recent past, they both also display relatively large internal natural variability which can induce changes from one year to another of the same order of magnitude as the changes associated with the long-term trend. Previous studies suggest that such observed changes are predictable a few weeks to a few months ahead. Accordingly, we plan to evaluate the ability of the most recent version of a state-of-the-art Earth system model, EC-Earth, which participated in the CMIP5 (Fifth Climate Model Intercomparison Project) project at the basis of the IPCC (Intergovernmental Panel on Climate change) AR5 (Fifth Assessment Report), to predict these two important features of the climate system at the seasonal scale.

In CMIP5, EC-Earth was run using a resolution of 1.125° in the atmosphere and about 1° in the ocean with enhanced resolution in the tropics (ORCA1 configuration). A new version, with significantly higher resolution, is now available and will become the standard version for the upcoming years. This new version can be run with either 1) a resolution of ~0.7° in the atmosphere and 1° in the ocean, or 2) a resolution of ~0.35° in the atmosphere and 0.25° in the ocean. In the second configuration, the enhanced atmospheric resolution significantly improves the realism of climate simulations, in particular, of the hurricane activity. Similarly, the high ocean resolution improves the representation of oceanic fronts, eddies and convection among other oceanic features, and potentially the representation of the Artic sea ice shrinking rate. In two series of retrospective seasonal forecasts over the last two decades using these configurations, we will assess the skill of EC-Earth in predicting tropical Atlantic cyclones and Arctic sea ice, and the benefits derived from increased resolution.

This project represents a key step in the development of an operational high-resolution seasonal forecast system. It will contribute to the European-funded SPECS (Seasonal-to-decadal climate Prediction for the improvement of European Climate Services) project and to the PICA-ICE (Previsión Interanual de la Cubierta de hielo marino del Ártico y su Impacto en el Clima de Europa) project funded by the Spanish Ministry of Economy and Competitiveness.

Computer resources requested
Total number of core hours requested: 5,000,000

Further details
(NOTE: if multiple codes to be used, give the specifications for each of them)

We are interested in each of the codes you intend to run. Please provide the following details for ***each*** of your codes, making a copy of this page to describe each code in turn.

Name of code #1: EC-Earth3

Please rank your preferred target PRACE architecture
(1=first choice, 2=second choice, xxx, x=not suitable)

Class of machine	Ranking
IBM BlueGene/Q	
IBM Power 7	2
Cray XE6/XC30	1
Intel cluster without GPU	
Intel cluster **with** GPU	x
Don't know	

Please give the motivation for your preferences above:

Our preference is to run on the Lindgren machine (PDC, Sweden) since both the code and the running environment, including the monitoring and post-processing systems, have already been installed and thoroughly tested on that platform. Being able to run on that particular machine would significantly cut down the amount of enabling work required to get started.

Total core-hours requested for this code:	5,000,000			
In case of only a partial grant, what will be the minimum requested core-hours	4,500,000			
Select machine on which this estimate is based				
IBM BlueGene/P ☐	IBM Power 6 ☐	Cray XT4/5/6, XE6 ☒	Cluster with GPUs? ☐	Other architecture ☐ CPU type and clock rate:
Select all the machines on which the code has been ported				
IBM BlueGene/P ☐	IBM Power 6 ☒	Cray XT4/5/6, XE6 ☒	Cluster with GPUs? ☐	Other architecture ☐ Details:
Number of planned runs during the whole project	High Resolution (H): 680 Low Resolution (L): 680			
Maximum number of cores per run	H: 1543 and L: 343			
Minimum number of cores per run	H: 775 and L: 151			

Total RAM for the run with maximum number of cores (Gbyte)	H: ~2500 and L: ~500
Total RAM for the run with minimum number of cores (Gbyte)	H: ~1200 and L: ~200
Maximum job wall clock time	H: 48 hours and L: 5 hours
Minimum job wall clock time	H: 14 hours and L: 2 hour
Please check the box if max wall clock time>12hours and your runs have a checkpoint/restart feature	☒
Total storage/hard disk for duration of the project (Gbyte)	For model outputs: H:19924 and L: 4284 For model restarts: H: 12920 and L: 1564
Storage/hard disk needed for one production run (Gbyte)	For model outputs: H: 29.3 and L: 6.3 For model restarts: H: 19 and L: 2.3
What applications and/or libraries does this code require	**GNU make 3.81 or above, Fortran 77/90/95, C, C pre-processor, Python 2.4.3 or above, Bash, Perl, BLAS/LAPACK, NetCDF, Grib_API, GRIBEX, GNU date (64-bit version), CDO, NCO, CDFTOOLS**
How is your application parallelized (MPI, OpenMP, Hybrid etc.). If GPU, state whether Cuda/OpenCL.	**Application could work with MPI or OpenMP**

Please check the box if this code I/O intensive? ☒
If checked, please describe your strategy concerning I/O (for example usage of I/O libraries, MPI I/O, netcdf, HDF5 or other approaches):

Each MPI rank writes its own restart file. Restart files for the atmosphere are raw binary files while those of the ocean model and the coupler are saved as NetCDF files. One complete set of restart files in an experiment using the T255L91/ORCA1L46 configuration results in about 2.3 GB of data, which is a small amount when compared to the typical output for a simulation. For a T511L91/ORCA025L46 experiment, the restarts are estimated to be approximately 19 GB. Typically restart files are saved at the end of a simulation chunk, which ranges in length between one and six months depending on the use of the machine recommended. The numbers provided in the table above correspond to chunks of 3-month length.

The number of output climate variables stored in GRIB for the atmospheric and land surface models and NetCDF files for the ocean models will be kept to a strict minimum to reduce the I/O volume and the size of the files to post-process. The outputs, for a chunk of 3-month length, in an experiment using the T255L91/ORCA1L46 configuration, results in about 6.3 GB of data, while for a T511L91/ORCA025L46 experiment, they are estimated to be approximately 29.3 GB.

Our 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.

Project Data

In this section, we are interested in the overall data requirements for your project as a whole. Please provide the following information.

What is the amount of data to be transferred to the target platforms before production runs can start:

The transfer of initial-condition files from the applicants' local machines to the production system is necessary prior to the start of the 680 (340 with High-resolution and 340 with Low-resolution configurations) seasonal forecasts from 680 different initial climate states. These files represent about ~3 Tb for the atmosphere, land, ocean and sea-ice initial conditions. Instead of transferring all the initial conditions at a time, the transfer could be organized in successive sessions depending on the production rate as the initial conditions for a forecast already produced are not necessary any longer for performing other forecasts. The model code and associated configuration files represent an additional 100 Gb approximately to be transferred.

What is the amount of data to be transferred from the target platform after all production runs are finished:

The transfer of the output files from the production system to the applicant institutions' local machines will be performed after horizontal interpolation and/or time averaging, which implies a substantial reduction in size of the output files in two orders of magnitude with respect to the original files produced during the simulation. The output files will be first locally stored and then transferred back to the applicant institutions for further analysis. The transfer of each of the 680 seasonal forecasts can be started as soon as the forecast ends. Summing the values provided in the table in the previous section, we reach a total amount of produced data to be transferred of 39 Tb, though most of those data would have been transferred before all the production runs are finished. The most adequate options for transfer (for instance, gridftp) will be discussed with the local technical support. At the end of the project, all data will have been transferred towards the applicant institutions and erased from the offline storage of the production platforms.

How much disk space do we need to reserve for your project (Tbyte)? As an indication please consider 0,5 Tbyte as small allocation and over 10Tbyte is a large allocation.

Since the transfer of the data will be done as the forecasts are performed, we can organize our data storage data transfer with a disk space of 3 Tb for data storage, when the total amount of data to be produced amounts to 39 Tb.

However, a large amount of work space will be required to perform several simulations at the same time (this is possible because there is no dependency between the members of an ensemble or the predictions for different start dates, i.e. each simulation can be treated as discrete). As a high-resolution six-month long forecast with the high-resolution configuration produces around 200 Gb of raw data, five ensemble members for five independent starting dates, which is a realistic number of the maximum number of simulations that can be performed simultaneously, will produce ~5 Tb. These data have to be held in the work space while it is post-processed to reduce the data volume to be stored offline and transferred to the applying institutions down to ~2.5Tb. On top of the outputs and restarts produced during the forecasts, the work space has to host the initial conditions, which represent an additional 500 Gb. Hence, we request 5 Tb of work space.

Around 20 Gb of home space will be required to host the code.

Please check the box if it is possible to start transferring data before all production runs are finished? ☒

Where will the data eventually be stored? (HPC facility, local storage at university, institute):

Institut Català de Ciències del Clima (IC3) in Barcelona

If any planned runs require extensive amounts of memory, hard disk space, etc., please provide details on how the project (in terms of CPU hours) may be split across multiple platforms. ____________

Since splitting the project between multiple platforms will increase significantly the amount of enabling work to start the project (as the amount of work required to port the code and to set up the monitoring and post-processing systems will almost double) and therefore might compromise its achievement on time, we prefer request such a splitting. However, it might be preferable to use either a smaller machine attached to the HPC or dedicated nodes for the post-processing of the model output before it is sent back to the applicant's institution.

Describe what work has already been done to develop the codes

This should include the following: describing the main algorithms, how they have been implemented and parallelized, and their main performance bottlenecks and the solutions to the performance issues you have considered. Please provide the name and version of all codes to be used in the project. For each code that needs to be optimized, please provide the details listed below.

1. Name and version.
2. Webpage and other references.
3. Licensing model.
4. Contact information of the code developers.
5. Your relationship to the code (developer, collaborator to main developers, end user, etc.).

The EC-EARTH earth-system model (http://ecearth.knmi.nl, https://dev.ec-earth.org) is the tool selected to perform the investigation, of which we are end users, as part of the EC-EARTH consortium. The development of EC-Earth started with ECMWF's Integrated Forecast System (IFS) as a well-tested atmospheric module, with different components being added over time. The oceanic (NEMO), sea ice (LIM) and land surface (HTESSEL) components have been coupled to the IFS through the OASIS3 coupler. A special issue on EC-Earth has recently been published in Climate Dynamics. Additionally, EC-Earth is an important contributor to the next Coupled Model Intercomparison Project (CMIP5). Since EC-Earth is a collaborative project between a range of institutes, it is run on a series of different platforms:

- SGI Altix 3500, Intel Itanium 2 processors by the KNMI
- NEC-SX6 by the DMI
- Linux cluster, 129 CPU nodes, each with 64-bit dual Intel Xeon EM64T CPU @ 3.40GHz by the SMHI
- Dell PowerEdge 2900 (18 nodes) by the CGUL
- IBM pSeries 575, Power6 processors by the KNMI
- 64 IBM p575+ 16-way nodes, 1.9GHz dual-core cpus, 2080 GB memory by the KNMI
- IBM Power PC 970MP processors at 2.3 GHz by the IC3

For this project, we plan to run the latest version of EC-Earth (EC-Earth3). A description of the EC-Earth model can be found in the following papers:

Hazeleger W. et al., EC-Earth: A Seamless Earth-System Prediction Approach in Action. Bull. Amer. Meteor. Soc. 91:10, 1357-1363. doi: 10.1175/2010BAMS2877.1 (2010)
Hazeleger, W. et al., EC-Earth V2.2: description and validation of a new seamless Earth system prediction model. Clim. Dyn., doi:10.1007/s00382-011-1228-5 (2012).

Klaus Wyser (Klaus.Wyser@smhi.se), Wilco Hazeleger (hazelege@knmi.nl) and Richard Bintaja (Richard.Bintanja@knmi.nl) are the main coordinators of the code development.

a) Describe what enabling work will need to be completed before production runs can begin.
b) Please make it clear what work will need to be done by your own group and what you are requesting to be done by PRACE staff.
c) Please make it clear if any enabling work can be done in parallel with production runs
d) Please indicate how many months of FTE will be required

a) The enabling work required to run EC-Earth3 will strongly depend on the platform on which we will be allowed to run. On our preferred platform, Lindgren, the enabling work will be relatively small as it is one of the platforms used for the development of the code by other EC-Earth partners. However, having the code installed on the platform is only covering the ability to perform a standard simulation. The simulations proposed here require the preparation of a complex environment:
- for the model to read the initial conditions,
- the interaction between the launching and monitoring system used to perform the climate predictions, Autosubmit,
- and the time-consuming task of post-processing the model output for its transfer to the applicant's institution.
Autosubmit is a tool, developed by IC3, to manage and monitor remotely climate forecasting experiments. It is designed with the following goals:

- Supercomputer independent framework to perform experiments
- Efficient utilization of available computing resources at supercomputers
- User-friendly interface to start, stop and monitor experiments
- Auto restarting the experiment or some part of experiment
- Ability to reproduce the completed experiments

The post-processing, storing and transfer scripts need to be carefully tested to efficiently handle the large volume of files and data that will be generated.

b) All the work described above will be done by the applicants. We may ask for assistance by the HPC support team if the installation of additional software facilitates our work. We will appreciate also an expert opinion to use the HPC resources as efficiently as possible.

c) The production runs can not be started before the enabling work is finished.

d) All this work will require 2 months FTE in the preferred platform and at least three times as much in any other platform.

Discuss the routes that you will use for dissemination of the project and for any appropriate knowledge transfer. This should include any resources that you will be using to support this. (Maximum 500 words):

The dissemination of the results will be done through the standard communication channels, i.e.:

1) through participation in international conferences such as yearly the European Geoscience Union (EGU) meeting, and
2) through publication of manuscripts in specialized peer-review journals.

The topics of which will include, but will not be limited to,

1) the general performance of EC-Earth high-resolution seasonal forecasts in comparison to other systems (preferably operational) and lower resolution EC-Earth forecasts
2) the ability of the system to predict highly important climate features, such as hurricane activity and Arctic sea ice losses

The knowledge and experience gained by running the system in this particular configuration will also benefit the entire EC-Earth community, as well as other European ESM groups using components of the EC-Earth system, such as IFS (atmospheric model), OASIS (atmosphere-ocean coupler) and NEMO (ocean model).

Finally, the results of the project will be featured at the meetings of WCRP's Working Group on Seasonal- to-Interannual Prediction, of which the team leader is a panel member, and the SPECS FP7 European project of which the team leader is the coordinator.

Confidentiality

Is any part of the project covered by confidentiality? No

If YES, please give the reasons for confidentiality:

Recent bibliographic references that are relevant to the project:

Blanchard-Wrigglesworth E, K C Armour, C M Bitz and E DeWeaver (2011a) Persistence and Inherent Predictability of Arctic Sea Ice in a GCM Ensemble and Observations. *J. Clim.*, **24,** 231–250.

Blanchard-Wrigglesworth E, C M Bitz and M M Holland (2011b) Influence of initial conditions and climate forcing on predicting Arctic sea ice. *Geophys. Res. Lett*., **38** (18), L18503. doi:10.1029/2011GL048807.

Caron L-P, C.G.Jones and K.Winger (2011) Impact of resolution and downscaling technique in simulating recent Atlantic tropical cyclone activity. *Clim Dyn, **37**, 869-892*

Caron, L.-P., C.G. Jones, P.A. Vaillancourt, and K. Winger (2012) On the relationship between cloud-radiation, atmospheric stability and Atlantic tropical cyclone activity in a variable-resolution climate model. *Clim Dyn*. DOI: 10.1007/s00382-012-1311-6

Delworth, Thomas L., and Coauthors (2012) Simulated Climate and Climate Change in the GFDL CM2.5 High-Resolution Coupled Climate Model. *J. Clim*, **25**, 2755–2781.

Emanuel, K. (2012) Potential Economic Value of Seasonal Hurricane Forecasts. *Wea Clim Soc*, **4**, 110-117.

Jung, T., and Coauthors (2012) High-Resolution Global Climate Simulations with the ECMWF Model in Project Athena: Experimental Design, Model Climate & Seasonal Forecast Skill. *J Clim*, **25**, 3155–3172.

Sakamoto T.T. and co-authors (2012) A new High-Resolution Atmosphere-Ocean Coupled General Circulation Model. *J Japan Met Soc,* **90***, 3, 325-359*

Serreze M C, M M Holland and J Stroeve (2007) Perspectives on the Arctic's Shrinking Sea-Ice Cover *Science,* **315** (5818), 1533–1536.

Stroeve J C, M C Serreze, M M Holland, J E Kay, J Malanik and A P Barrett (2012) Arctic sea ice decline: Faster than forecast. *Climatic Change,* **110** (3-4), 1005-1027.

Zhang, Y., W. Maslowski, and A. J. Semtner (1999), Impact of mesoscale ocean currents on sea ice in high-resolution Arctic ice and ocean simulations. *J. Geophys. Res*., **104**(C8), 18,409–18,429, doi:10.1029/1999JC900158.

Do you have any other support for this application e.g. from your national funding council, the EC or international collaborations? Please give details of this below:

The computing time requested will go towards supporting the IC3 and SMHI commitments in the context of the SPECS FP7 European project. Led by IC3, this 4-year project, which started in November 2012, aims at building a new generation of climate forecast systems based on the latest scientific progress in climate modelling and in operational forecasting.

The assessment of the forecast skill in the Artic region will additionally contribute the PICA-ICE project, which will start in January 2013 and last three years. This project, funded by the Spanish Ministry of Economy and Competitiveness, aims at investigating the mechanisms driving the interannual variability of the Artic sea ice cover and the mechanisms by which the Artic sea ice cover influence the Northern hemisphere climate, as well as to assess the extent of their predictability. It also aims at refining the predictions of the externally forced Artic sea ice loss and internally-generated variability of the sea ice cover for the next decade.

This work will substantially contribute to raise the profile of the EC-Earth climate community, in particular in the field of climate prediction, both at the European and international level. Given its interest, it has received the explicit support of a number of international partners such as NOAA and ECMWF.

Please use the remainder of this form to give a detailed description of the project and its relevance for PRACE.

Please provide 3-4 pages, including:
1. scientific objectives
2. scientific and technical innovation potential
3. current profile and performance of code(s), including scalability, requirements on interconnect, I/O, architecture, clarification of how requested core-hours was calculated, are jobs independent, chained and/or workflows, etc.
4. computational objectives
5. specific benefits expected from PRACE
6. summary

1. Scientific objectives

The main scientific objective is to explore the potential benefits of increased resolution of dynamical climate forecast systems on seasonal forecast quality. This objective works towards a longer-term aim of the climate prediction community at pursuing the creation of a new generation of climate forecast systems to produce actionable local climate information over land at seasonal time scales.

The standard horizontal resolution of typical climate forecast systems is ~1-2° for both atmosphere and ocean components. This is the resolution currently used operationally and also for the assessment of the impact of long-term anthropogenic climate change. This is also the resolution at which forecast systems have pervasive systematic errors. Increasing the model resolution has been suggested to improve many aspects of climate simulations. Jung et al (2012) found significant improvements in the simulation of tropical precipitation and the frequency/intensity of both tropical and mid-latitude cyclones in IFS-only simulations (the atmospheric component of EC-Earth). Two recent studies report the first use of coupled climate models in the resolution range ~50km (atmosphere) and ~10-30km (ocean), identifying large improvements in a number of systematic biases seen in lower resolution versions of the same models. These include: a reduction in the double ITCZ problem (GFDL CGCM; Delworth et al. 2012), improved intensity distribution of tropical rainfall (MIROC4h CGCM, Sakamoto et al. 2012), improved coastal upwelling and ENSO variability (Sakamoto et al. 2012) and improved structure of the North Atlantic Ocean circulation (Delworth et al 2012). Caron et al (2011) using a variable resolution AGCM found significant improvements in the representation of Atlantic tropical cyclones from 150km to 30km resolution. Resolving the mesoscale ocean eddies allow for a more realistic representation of the ice drift and deformation and subsequently of the Arctic open water percentage (Zhang et al 1999).

In this proposal, we intend to investigate to which extent the improvements seen in historical climate simulations with increased resolution translate into a better seasonal forecast quality, with a special focus on the Tropical Atlantic cyclones on the one hand and on the Arctic sea ice cover on the other hand. The Atlantic hurricane activity is expected to be better predicted for two main reasons: 1) ~30km (T511) resolution simulations have already been shown to produce realistic geographic distribution and numbers of Atlantic tropical cyclones, and 2) the Atlantic cyclones are strongly modulated by ENSO, the prediction of which lies at the heart of seasonal predictions. Furthermore, the Arctic sea ice is expected to be better predicted thanks to: 1) a better representation of the ocean circulation and associated heat transport toward the Arctic (Delworth et al 2012), 2) a better representation of the western boundary currents/frontal areas and their impact on the generation of storms which can break the Artic sea ice and favour its melting, 3) a better representation of the ocean eddies which affect the sea ice drift, deformation and the formation of open water areas (Zhang et al, 1999).

This is a fundamental question still not comprehensively addressed by any research or operational community. To successfully achieve such progress is a major challenge both scientifically and technically, while also being computationally demanding.

2. Scientific and technical innovation potential

This proposal aims to make major advances in the science of climate prediction by addressing the dual requirements of increased climate-model resolution while performing enough predictions for the past that allow a robust estimation of forecast quality. Increased model resolution aims to deliver a significant improvement in our ability to simulate the key modes of climate variability at the origin of climate predictability, while the sample of past predictions is needed to estimate forecast quality. The ensemble method used acknowledges the inherent uncertainty in estimating changes in climate.

The computer cost necessary to perform this type of analysis is very high and exceed the resources locally available to most research groups, and in particular the two applicants. As such, the computer time requested here would allow the development of the EC-Earth state-of-the-art seasonal forecast system, allow the EC-Earth consortium to be at the forefront of seasonal forecasting research and contribute to improving the current European operational climate prediction capabilities.

Since the results of the project will be of the research type and not the subject of commercial use, the results will be widely presented in scientific conferences, courses and peer-reviewed literature. Results derived from this experiment will be made available to other research groups interested in the development of a high resolution forecast system.

Moreover, EC-Earth being run at different timescales and resolutions by additional European research groups beyond the ones involved in this particular project (DMI, KNMI, …), the experience gained here will benefit the EC-Earth community at large. Results will be disseminated through the EC-Earth community through bi-annual meetings where the latest developments relating to EC-Earth are shared and through relative frequent direct communications between the different members.

Finally, the data will be made available for download to all SPECS member on a project specific platform, allowing further analysis to be done according to the expertise of individual research groups.

3. Current profile and performance of code(s), including scalability, requirements on interconnect, I/O, architecture, clarification of how requested core-hours was calculated, are jobs independent, chained and/or workflows, etc.

The EC-Earth3 coupled climate model comprises three main computational components: the atmospheric model IFS, the ocean model NEMO, and the OASIS3 coupler. All components are mainly written in Fortran (with some units implemented in C) and parallelized for distributed memory architectures using MPI. Model I/O is being processed in GRIB and NetCDF formats using the appropriate external libraries. Output post-processing is primarily done with the cdo tools (code.zmaw.de/projects/cdo/wiki).

EC-Earth, as many other climate models, faces a number of challenges regarding its scalability on massively parallel architectures. Among its siblings, however, EC-Earth3's components show distinguished scalability. Moreover, the development has, over the last years, had a focus on computational efficiency and scalability both from a per-component and overall coupled model perspective. Bottlenecks, such as parallel I/O and load balancing have being identified in the course of preceding and ongoing projects and performance tuning has become an integral part of the software development cycle. Although the scalability of EC-Earth is suboptimal beyond a few thousand processors, the ensemble strategy considered in this proposal and that can be easily handled by Autosubmit offers the possibility of obtaining a manageable time-to-solution.

A typical climate prediction experiment consists in running a climate model with a forecast length ranging between a few months to a few years. An experiment has several independent start dates and several independent members. All the start dates and all the members being independent, the jobs can be run simultaneously if necessary. The forecast period is typically divided into a number of sections, known as chunks, of fixed length that exploits the option of model restart. Each chunk has two big sections; a parallel one where the simulation is performed using a number of processors, and a serial one to perform all the necessary operations to post-process the model output, select and archive the output to be retained and cleaning the disk space for the smooth proceeding of the experiment.

These tasks are controlled by Autosubmit (http://ic3cfu.wikispot.org/Autosubmit), a tool running remotely that manages and monitors the experiment, achieving the following goals: efficient handling of highly dependant jobs, optimum utilization of available computing resources, ease of starting, stopping and live monitoring of experiments, auto restarting the experiment or some part of experiment in case of failure, use of database for experiment creation and assigning automatic experiment identity and the ability to reproduce the completed experiments fully or partially. The workflow of the post-processing tasks will also be handled accordingly by Autosubmit.

The experiments planned do not need any pre-processing of the data, but the post-processing of the output will include an interpolation toward a coarser grid and/or time averaging to reduce the cost of both storage and transfer. The output will be in NetCDF for the ocean and sea-ice components and in GRIB for the atmosphere and land-surface components, although the GRIB files could be converted to NetCDF in situ if necessary for exchange with other groups. This, and other actions, will require the availability of the cdo tools. Autosubmit will handle the submission of the post-processing jobs taking into account the scheduling of the simulation jobs.

The required computer resources for this project are described in the next section in detail together with how they were estimated.

4. Computational objectives

Two experiments will be carried out with the two different configuration labelled EC-Earth3 and EC-Earth3hr in the table below. We will conduct ensemble experiments with four start dates a year (1st of February, May, August and November) over the 1993-2009 period. For each experiment, this makes a total of 68 independent start dates (4 start dates per year over an 17 year period) and a total of 170 years of simulation when using five-member ensembles and producing six-month long forecasts. Sets of initial conditions for the two resolutions considered will be required, and these initial conditions will be prepared by the leading institution. Coordination with other members of the EC-Earth community will be undertaken to ensure that the most suitable techniques for ensemble generation and the latest high-resolution atmosphere, sea-ice and ocean reanalyses (that act as initial conditions for the predictions) are used. A successful assessment should make sure that the experiments have sufficient forecast spread, which up a certain point depends on the way the ensemble is generated. The predictions will be initialized with the European Centre for Medium-Range Weather Forecasts (ECMWF) most recent atmosphere and land surface reanalysis (ERAInterim, Dee et al, 2011) interpolated to the two atmospheric resolutions. The ocean and sea-ice initial conditions will be produced with a set of five simulations of the ocean and sea-ice model component included in EC-Earth forced with atmospheric fluxes from the ECMWF reanalysis and nudged toward the GLORYS2v1 ocean reanalysis (Ferry et al, 2010) which covers the 1993-2009 period. The ensemble of initial conditions will be produced by introducing wind-stress perturbations during the offline ocean simulations, which will be performed on the applying institutions own HPC resources

Model Version	Atmosphere and Soil	Ocean and Sea-ice
EC-Earth2	T159 (1.125°) and 62 levels	1° and 46 levels
EC-Earth3	T255 (0.7°) and 91 levels	1° and 46 levels
EC-Earth3hr	T511 (0.35°) and 91 levels	0.25° and 46 levels

The required computer resources for both experiments are described below. Those estimates have been obtained from one-month long simulations with the corresponding configuration on Lindgren.

Low-resolution (EC-Earth3) experiment:

CPU per simulated year: 1,098 CPU-hours
Offline storage per year: 34.4 Gb
Processors (per member): 343
Number of simulated years: 170
Total CPU: 186,660 CPU hours
Total offline storage: 5.9 Tb

High-resolution (EC-Earth3hr) experiment:

CPU per simulated year: 26,231 CPU-hours
Offline storage per year: 193.2 Gb
Processors (per member): 1543
Number of simulated years: 170
Total CPU: 4,459,270 CPU hours
Total offline storage: 32.9 Tb

This leads to a total of 4,645,930 CPU hours. Adding a small allowance for unavoidable mistakes in the setup of the experiments, a total of 5,000,000 CPU-hours is requested.

5. Specific benefits expected from PRACE

This project is particularly innovative and demanding in computational resources. Recognizing that forecast skill is likely to be dependent on specific atmosphere-ocean-ice initial conditions and the improvement in model fidelity associated with the increase in resolution, and that a robust forecast quality assessment requires large enough samples, we consider the construction of a hindcast dataset as long as possible rather than the analysis of a few case studies. PRACE would offer the opportunity to run at cutting-edge high horizontal resolution for climate prediction and assess the potential benefits of such an increase in resolution.

6. Summary

The goal of this "**Seasonal Prediction of the Arctic Ice and Tropical Atlantic Cyclones"** proposal is to assess to which extent an increased model resolution improves the forecast quality for the Arctic sea ice cover and the Tropical Atlantic hurricane activity, among other climate features. The EC-Earth Earth System Model will be used in a set of ensemble seasonal forecast experiments initialized with the best analyses of the climate system available up to date. This project will also contribute to a better understanding of how the initial-condition information propagates in the climate system, a better understanding of the mechanism underlying the predictability of the Arctic sea ice cover and the tropical Atlantic cyclones. Predicting such high impact climate features stands as a key scientific challenge of high public and socio-economic interest.