Project Name	Monitoring the Arctic Climate (MAC)	
Project Representative	Affiliation : Barcelona Supercomputing Center (BSC)	Name: Virginie Guemas

1 Significance of the Research

MAC is a proposal for the "General Projects of the K computer use" call that aims at refining estimates of the sea ice state over the last three decades and its impact on the atmosphere as a general purpose. Its specific purpose is to harvest the benefits of the latest high-resolution and high-quality satellite products available to generate a sea ice reanalysis at a resolution never explored up-to-date and examine the impacts of this sea ice reanalysis on numerical weather prediction (NWP) using a global atmospheric data assimilation system in RIKEN AICS.

1.1 Purpose of the Research

1.1.1 Scientific background for the research

The dramatic decline in Arctic sea ice has been an emblematic sign of ongoing global climate change. Profound reductions in sea ice areal coverage (Cavalieri and Parkinson, 2012) and thickness (Rothrock et al., 2008), among others, have already had devastating impacts on local ecosystems (Tynan, 2015; Kovacs et al., 2011), indigenous populations (Meier et al., 2014) and possibly lower-latitude climate (Cohen et al., 2015). These rapid changes are also unlocking economic and industrial opportunities. Thinner, younger ice facilitates operations of icebreakers in the High North. Increased marine accessibility promotes polar shipping (Stephenson et al., 2013; Smith and Stephenson, 2013) as an economically viable alternative to existing commercial routes (Liu and Kronbak, 2010). Ecotourism, resources extraction and industrial fishing are other examples of activities that can take place in an open Arctic Ocean. Further reductions in the sea ice cover (Massonnet et al., 2012; Stroeve et al., 2012) are expected in the near future and predicting the first Arctic-free summer is one of the current challenges of the scientific community. The rapid sea ice decline contributes to polar temperature amplification (Manabe and Stouffer 1980; Serreze et al, 2009; Screen and Simmonds 2010), which affects the atmosphere in the Arctic region and beyond. The suggested impacts of sea ice decline include, among others, a polar stratospheric cooling (Screen et al., 2013), a weakening of the midlatitude jet (Francis et al., 2009; Francis and Vavrus, 2012), and an increase in the frequency of cold Northern Hemisphere midlatitude winter events (Cohen et al., 2012; Liu et al., 2012; Yang and Christensen, 2012).

A better understanding of the interactions between the long-term externally forced climate trend and the internal variability, which is essential to accurately estimate the amplitude of upcoming sea ice losses, requires long and continuous monitoring of polar climate variables. Unfortunately, before 1973, Arctic sea-ice data are limited to monthly estimates of sea ice extent (SIE), with complete cover assumed within the ice pack and a necessary treatment of missing data in the marginal seas (Walsh and Johnson, 1979). The situation is worse in the Antarctic where sea ice data is limited to estimates of extent climatologies over two distinct periods: 1929-1939 (Deutsches Hydrographisches Institute, 1950) and 1947-1962 (Tolstikov, 1966). From 1973, the US Navy, Canadian and Danish aerial reconnaissance provided quasi-weekly estimates of sea ice concentration (SIC; Knight, 1984). The advent of satellite microwave imagery in 1978 allowed for the retrieval of SIC at roughly 25 km resolution and a 2-day frequency, which was increased to a daily frequency in 1987 (Cavalieri et al., 1996).

The publicly distributed datasets which are currently available, i.e. National Snow and Ice Data Center (NSIDC; Cavalieri et al., 1996), OSI-SAF (1978-present; Eastwood et al., 2015) and HadISST (1870-present; Rayner et al. 2003) sea ice concentration fields stand as the best estimates of sea ice concentration obtained from combination, homogenization and extrapolation of these sparse observational data. High-quality and high-resolution satellite sea ice concentration products have also been distributed recently by the European Space Agency (ESA) and cover the 1993-2008 period (Ivanova et al., 2015). The sea ice thickness (SIT) data are much scarcer (Kwok and Rothrock, 2009): the first unified dataset (Lindsay, 2010) was released in 2010. It combines Arctic observations by submarines from 1975, moored upward-looking sonar from 1990 and airborne or satellite electromagnetic measurements from the past decade.

Filling these critical gaps in sea ice observations is essential for monitoring the long-term evolution of the sea ice system but also to provide a complete description of the sea ice state to initialize seasonal-to-decadal climate predictions. Indeed, the knowledge of the initial climate system state has been shown to be a major source of information in seasonal forecasts (Balmaseda and Anderson 2009) as well as in decadal climate predictions (Smith et al. 2007; Keenlyside et al. 2008; Pohlmann et al. 2009). In particular, the spring Arctic sea ice thickness distribution has been shown to be a precursor of the September sea ice cover (Chevallier and Salas-Melia, 2011) in a model study. A summer-to-summer re-emergence mechanism has been suggested whose memory lies in the sea ice thickness (Blanchard-Wrigglesworth et al. 2011). Wang et al. (2013) also found a dependence of the sea ice prediction skill on the initial sea ice thickness. Refined estimate of the sea ice cover are also essential to provide trustworthy lower boundary conditions for Numerical Weather Prediction (NWP). Polar storms, for example, can interact on hourly timescales with the sea ice cover (Simmonds and Rudeva, 2012). It is also known that the Arctic observations affect mid-latitude NWP (Inoue et al. 2009 and follow-on studies), so that the Arctic sea-ice state may have impact on mid-latitude atmospheric analyses and forecasts.

A complete and coherent description of the sea-ice state can only be obtained through a physical extrapolation of the sparse observations, relying on the equations that describe the sea-ice dynamics and thermodynamics. This is, in short, the goal of data assimilation. The first attempt at providing such a complete description of the sea-ice state was performed by Lisaeter et al. (2003) by assimilating SIC data in a coupled ice–ocean model every week over more than 1 year using an ensemble Kalman filter (Evensen, 1994). Sea-ice velocity is well observed in the recent decades, and its assimilation can be used to improve the simulated SIT, as shown by Zhang et al. (2003) by applying an optimal interpolation procedure to assimilate buoy motion and satellite ice-motion data in an ice–ocean model. Lindsay and Zhang (2006) combined the velocity assimilation scheme by Zhang et al. (2003) with a nonlinear nudging scheme for SIC that substantially corrects the SIC when large differences between the model and the observational data are seen, i.e., along the sea-ice edges. The three-dimensional variational assimilation technique was tested by Caya et al. (2010) on a regional ice–ocean model. They showed significant improvements over a simple nudging technique when combining multiple sources of observational data. The propagation of the analysis update from the observed fields to the unobserved fields, such as the SIT, remains an issue. Using the extended Ensemble Kalman Filter, the propagation of the SIC information is performed through a model error covariance matrix which is estimated thanks to an ensemble of sea ice reanalyses (Massonnet et al 2013). This approach has two advantages: 1) the technique to propagate the information is fairly simple and allows for a propagation to a large number of sea ice variables, 2) an ensemble of reanalyses is generated which allows to sample the uncertainty on the sea ice state.

Up to now, the use of the Ensemble Kalman filter to assimilate sea ice observational data has only been applied at a typical resolution of about 1° (Massonnet et al, 2014). However, resolving mesoscale ocean eddies would allow for a more realistic representation of the ice drift and deformation and, consequently, of the Arctic open water percentage (Zhang et al. 1999; Gent et al, 2010). A better representation of the ocean circulation and associated heat transport toward the Arctic thanks to high resolution as well as a better representation of the western boundary currents/frontal areas and their impact on the generation of storms which break the Arctic sea ice and lead to its melting, can produce a more realistic representation of Arctic sea ice processes (Delworth et al. 2012).

<u>1.1.2 What will be elucidated and to what extent will it be pursued during the research period ?</u>

Here, we propose to generate a 25-member ensemble of sea ice reanalyses at a resolution never achieved up to now in a data assimilation context. Such a large ensemble size will provide a robust evaluation of the related uncertainty. This reanalysis will use the deterministic Ensemble Kalman Filter (Sakov and Oke, 2008) to assimilate sea ice data from the OSISAF and ESA products and will cover the 1978-present period, thus becoming the longest sea ice reanalysis available. ESA products stand as the highest quality and highest resolution product available nowadays, thus we'll favour this product over its availability period (1993-2008) and use OSISAF, the second best product for the rest of the reanalysis (Massonnet et al 2016). The version 3 of the Louvain-La-Neuve sea Ice Model (LIM3, Vancopenolle et al. 2009) is embedded in the version 3.6 of the Nucleus for European Modelling of the Ocean (NEMO), which is the latest version incorporated in several coupled climate models participating to the next Coupled Model Intercomparison Project Phase 6 (CMIP6). The grid resolution will be of 0.25° globally (approximately 25 km) with 75 vertical levels, with the thickness increasing from 1m below surface up to 500m in the deep ocean. The Drakkar Forcing Set Version 5.2 will be used as atmospheric surface forcing with perturbations to generate the different members of the sea ice reanalysis (Guemas et al. 2014). This sea ice reanalysis will then be exploited as lower boundary conditions within the global atmospheric data assimilation system developed in RIKEN AICS, the system known as the NICAM-LETKF consisting of the Nonhydrostatic Icosahedral Atmospheric Model (NICAM) and the Local Ensemble Transform Kalman Filter (LETKF) (Terasaki et al. 2015). Here we propose to assess to what extent using the high-resolution sea ice data in the NICAM-LETKF affects the quality of atmospheric analyses and forecasts for high-impact events such as polar storms and potentially in lower latitudes.

MAC will generate the longest sea ice reanalysis and counting the largest number of ensemble members available up to date, at a ground-breaking resolution and exploiting the latest high quality and high resolution satellite products available. **MAC** will allow for the most robust estimate of the Arctic sea ice changes over the last decades and its uncertainty performed up to date and therefore deliver crucial information about Arctic changes monitoring, being hence a project of scientific excellence. **MAC** will also provide the most trustworthy sea ice initial conditions to train dynamical climate prediction systems and estimate future Arctic changes. The final results will be of utmost interest for local ecosystems and populations as well as for industries willing to invest in this region. Finally, **MAC** will assess to what extent NWP can benefit from using the high quality sea ice data for atmospheric analyses and forecasts, particularly for high-impact weather events with drastic human and economic consequences such as polar storms and even in mid latitudes. Therefore these results will be of key *social significance*.

1.1.3 Scientific characteristics and originality of the research in the area

MAC proposes to harvest the benefits of the highest quality highest resolution sea ice satellite products, through the use of the highest resolution explored up to date in sea ice reanalysis. This project could lead to a breakthrough in the monitoring and prediction of changes in the Arctic region, the most vulnerable to climate change. Its execution requires the use of the K computer because of the exceptionally large amount of computing resources required to carry out successfully such project.

References:

Balmaseda MA, Anderson D (2009) *Geophys Res Lett,* 36:L0170. doi:10.1029/2008GL035561.

Blanchard-Wrigglesworth E, Armour KC, Bitz CM, DeWeaver E (2011) *J Clim,* 24:231–250.

Cavalieri DJ, Parkinson CL, Gloersen P, Zwally HJ (1996) National Snow and Ice Data Center. Digital media, Boulder.

Cavalieri DJ, Parkinson CL (2012) *The Cryosphere*, 6:881-889, doi:10.5194/tc-6-881-2012.

Chevallier M, Salas-Melia D (2011) *J Clim,* 25:3025–3038, doi:10.1175/JCLI-D-11-00209.1.

Caya A, Buehner M, Carrieres T (2010) *J Atmos Oceanic Technol,* 27: 353–369.

Cohen J, Furtado JC, Barlow MA, Alexeev VA, Cherry JE (2012) *Environ. Res. Lett,* 7: 014007.

Cohen J, Screen J, Furtado JC, Barlow M, Whittleston D, Coumou D, Francis J, Dethloff K, Entekhabi D, Overland J, Jones J (2014) *Nature Geoscience*, 7:627-637, doi:10.1038/ngeo2234.

Delworth TL, and Coauthors (2012) *J. Clim*, 25:2755–2781.

Deutsches Hydrographisches Institute (1950) Atlas of ice conditions in the north atlantic ocean and general charts of the ice conditions of the north and south polar regions. Publ 2335, Hamburg.

Eastwood S, Jenssen M, Lavergne T, Sorensen A, Tonboe R (2015) *Technical Report, Norwegian and Danish Meteorological Institutes, Oslo, Norway and Copenhagen, Denmark*, 5526, 552.

Evensen G (1994) *J Geophys Res*, 99: 10143–10162, doi: 10.1029/94JC00572.

Francis JA, Vavrus SJ (2012) *Geophys Res Lett*, 39: L06801, doi:10.1029/2012GL051000.
Francis JA, Chan W, Leathers DJ, Miller JR, Veron DE (2009) *Geophys Res Lett,* 36: L07503.
Gent PR, Yeager SG, Neale RB, Levis S, Bailey DA (2010) *Clim Dyn*, 34:819–833.
Guemas V, Doblas-Reyes FJ, Mogensen K, Keeley S, Tang Y (2014) *Clim Dyn*, 43: 9-10, 2813-2829.
Ivanova N, and Coauthors (2015) *The Cryosphere*, 9:1797-1817. doi: 10.5194/tc-9-1797-2015.
Inoue J, Enomoto T, Miyoshi T, Yamane S (2009) *Geophys. Res. Lett.*, 36, L08501, doi:10.1029/2009GL037380.
Keenlyside NS, Latif M, Jungclaus J, Kornblueth L, Roeckner E (2008) *Nature* 453:84–88.
Knight RW (1984) *Ann Glaciol* 5:81–84.
Kovacs KM, Lydersen C, Overland JE, Moore SE (2011) *Marine Biodiversity*, 41, 181-194.
Kwok R, Rothrock DA (2009) *Geophys Res Lett,* 36: L15501, doi: 10.1029/2009gl039035.
Lindsay RW (2010) *EOS Trans AGU,* 91(44):405–416.
Lindsay RW and Zhang J (2006) *J Atmos Oceanic Technol,* 23: 742-749.
Lisaeter KA, Rosanova J, Evensen G (2003) *Ocean Dyn,* 53: 368–388, doi: 10.1007/s10236-003-0049-4.
Liu M, Kronbak J (2010) *Journal of Transport Geography*,18, p. 434-444.
Liu J, Curry JA, Wang H, Song M, Horton RM (2012) *Proc. Natl. Acad. Sci. U.S.A.* 109: 4074–4079.
Manabe S, Stouffer RJ (1980) *J Geophys Res,* 85: 5529–5554, doi: 10.1029/JC085iC10p05529.
Massonnet F, Fichefet T, Goosse H, Bitz CM, Philippon-Berthier G, Holland MM, Barriat PY (2012) *The Cryosphere*, 6:1383-1394.
Massonnet F, Mathiot P, Fichefet T, Goosse H, König T, Beatty C, Vancoppenolle M, Lavergne T (2013) *Ocean Modelling*, 64:67-75, doi:10.1016/j.ocemod.2013.01.003.
Massonnet F, Goosse H, Fichefet T (2014) *Ocean Modelling,* 88:16-25, doi:10.1016/j.ocemod.2014.12.013.
Massonnet F, Bellprat O, Guemas V, 2016, *Sciences*, Oct 28;354(6311):452-455.
Meier WN, Hoverlsrud GK, van Oort BEH, Key JR, Kovacs KM, Michel C, Haas C, Granskog MA, Gerland S, Perovich D, Akshtas A, Reist JD (2014) *Review of Geophysics*, 52(3):185-217.
Pohlmann H, Jungclaus JH, Kohl A, Stammer D, Marotzke J (2009) *J Clim,* 22:3926–3938.
Rayner NA, Parker DE, Horton EB, Folland CK, Alexander LV, Rowell CP, Kent EC, Kaplan A (2003) *J Geophys Res* 108:D14, doi:10.1029/2002JD002670.
Rothrock DA, Percival CB, Wensnahan M (2008) *J Geophys Res*, C05003, doi:10.1029/2007JC004252.
Sakov P, Oke P (2008) *Tellus A*, doi:10.1111/j.1600-0870.2007.00299.x.
Screen JA, Simmonds I (2010) *Nature,* 464: 1334–1337, doi: 10.1038/nature09051.
Screen JA, Simmonds I, Deser C, Tomas R (2013) *J Clim,* 26: 1230–1248.
Serreze MC, Barnett AP, Stroeve JC, Kindig N, Holland MM (2009) *Cryosphere,* 3: 11–19.
Simmonds I, Rudeva I (2012) *Geophys Res Lett,* 39, L23709, doi:10.1029/2012GL054259.
Smith D, Cusack A, Colman A, Folland C, Harris G, Murphy JM (2007) *Science* 317:796–799.
Smith LC and SR Stephenson (2013) *Proc. Natl. Acad. Sci. U.S.A.*, 110: E1191-E1195.
Stephenson SR, Brigham LW, Smith LC (2013) *Polar Geography*, doi: 10.1080/1088937X.2013.845859.
Stroeve J, Kattsov V, Barret A, Serreze M, Pavlova T, Holland M, Meier W (2012) *Geophys Res Lett*, 39: L16502, doi:10.1029/2012GL052676.
Terasaki K, Sawada M, Miyoshi T (2015) *SOLA,* 11, 23-26, doi:10.2151/sola.2015-006.
Tolstikov EI (1966) Atlas of the Antarctic. Glav Uprav Geod Kart MGSSSR, Moscow, vol 1, 126 pp.
Tynan E (2015) *Nature Climate Change*, 5, 621, doi:10.1038/nclimate2708.
Vancopenolle, M, Fichefet T, Goosse H, Bouillon G, Madec G, Maqueda MAM (2009) *Ocean Modelling*, 27(1):33–53.
Walsh JE, Johnson CM (1978) *J Phys Oceanogr,* 9(3):580–591.
Wang W, Chen M, Kumar A (2013) *Mon Weather Rev,* doi:10.1175/MWR-D-12-00057.1.
Yang S, Christensen JH (2012) Geophys Res Lett, 39: L20707, doi: 10.1029/2012GL053338.
Zhang Y, Maslowski W, Semtner AJ (1999) *J Geophys Res*, 104, 18409-18429, doi:10.1029/1999JC900158.
Zhang J, Thomas D, Rothrock DA, Lindsay RW, Yu Y, Kwok R (2003) *J Geophys Res*, 108: 3170, doi: 10.1029/2001JC001041.

1.2 Expected achievements and Ripple effects

This project will generate an ensemble of sea ice reanalyses at a resolution never achieved up to now, counting as many as 25 members which will provide a robust evaluation of the related uncertainty. This ground-breaking sea ice reanalysis will be publicly distributed and extensively exploited for the analysis and understanding of internal climate variability, for the estimation of the rate of externally forced sea ice loss, and for the source of low boundary conditions for NWP. Through close collaboration with the Earth Services Group from the Earth Department of the BSC, we will ensure that the resulting progress feeds the climate services team and is exploited to generate user-relevant climate information, as a *contribution to society*.

The results of the project will be of the research type and, hence, will not be the subject of commercial use. Instead, the results of the project will be widely presented in scientific conferences and will feature at the meetings of the Scientific Steering Group (SSG) CLIVAR (Climate and Ocean Variability Predictability and Change) project from which the project representative (Virginie Guemas) is a member. These results will contribute to the INTAROS (understanding of polar processes), APPLICATE (linkages between the polar and mid-latitude climate), PRIMAVERA (added-value from increasing climate model resolution) and ESIWACE (porting and optimization of model codes on different platforms) projects funded by the European Commission in which the applying team (BSC members in particular) is participating, as well as to the FLAGSHIP2020 priority issue #4 (improving NWP through big data assimilation) and JAXA Precipitation Measuring Mission (use of satellite data with NICAM-LETKF)

MAC will also be the opportunity to trigger a collaboration on data assimilation between the team of Takemasa Miyoshi on data assimilation at RIKEN, and the group of climate prediction of Virginie Guemas and the group of Computational Earth Sciences of Kim Serradell at the Barcelona Supercomputing Center (BSC, Barcelona, Spain).

The results of this project will benefit the climate community and therefore contribute to the state-of-the-art knowledge on Arctic climate change. At least one paper will be published in a peer-reviewed journal to summarize and disseminate the main results of this project. A press release will be prepared at the time of publishing this article to ensure an optimal media outreach. Being of *scientific excellence and social significance*, this project could lead to a *breakthrough*, i.e. a refined and robust estimate of the rate of Arctic sea ice loss that can be expected in the coming years and its impact on the atmosphere. Its execution requires the use of the K computer because of the exceptionally large amount of computing resources required to carry out successfully such project.

2 Research Plan and Method

2.1 Research Plan and Method

Our research plan will include : 1. the porting and speed up of the model code to be used to generate our sea ice reanalysis on the K computer (section 2.1.1), 2. the generation of a model spinup simulation to equilibrate the model climate (section 2.1.2), 3. the generation of our high resolution ensemble sea ice reanalysis assimilating different sea ice observational products (section 2.1.2), 4. the production of atmospheric data assimilation and forecast experiments exploiting the high resolution high quality sea ice reanalysis from MAC as lower boundary conditions (section 2.1.3), 5. a refined estimation of the rate of Arctic sea ice loss and its characterization over the last decades and an identification of Arctic sea ice variability modes and their oscillations, as well as an assessment of the impact of the most realistic Arctic sea ice prescription achievable up-to-date on the quality of NWPs (section 2.1.4).

2.1.1 Porting, performance analysis and speed up of the NEMO3.6-LIM3 model code

The very first step will be to port the NEMO3.6 (http://www.nemo-ocean.eu/) ocean model in which the LIM3 sea ice model is embedded, running at the same resolution. The configuration that we plan to use has a grid resolution of about 0.25degree globally (about 25km) with 75 vertical levels which thickness increases from 1m below the surface to 500m in the deep ocean. The code is written in Fortran and the parallelization uses MPI. Several optimizations (Tintó et al., 2016) have been applied to the code by the performance team of the department in order to ensure a larger parallel efficiency than 50%. Although K compilers have not been tested to run NEMO model, the building process NEMO executable is fully configurable and allows to use a wide range of compilers and libraries. Having deployed the model in many different HPC facilities, we can demonstrate a large experience in using different environments. Some developments could be necessary to adapt Autosubmit workflow manager (briefly described later) to use the K computer scheduler. Finally, optimizations based on compilation flags or K system administrator's indication will be applied. Additional tests will be done in order to optimize the efficiency of the model, applying a methodology to do a scalability test of the parallel execution of NEMO. These tests will ensure the optimal use of the machine and the adaption to the specific hardware of the platform (access memory or communication among nodes could be optimized for example), in the same way that was applied to other platforms (Acosta et al., 2016).

2.1.2 Sea ice reanalysis and control experiment

A single-member spin-up experiment will be first started from the EN4 climatology (Good et al. 2013) and an ocean at rest with a sea ice thickness of 1m in the Antarctic and 3m in the Arctic. This spin-up will be carried out with NEMO3.6-LIM3 forced by DFS5.3 and will cover the 1958-1977 period. Two experiments will then be initialized from this spin-up, both also run with NEMO3.6-LIM3 forced by DFS5.2 and covering 1978-present: a 5-member control experiment without any sea ice data assimilation (but perturbation in the atmospheric forcing) and a 25-member sea ice reanalysis assimilating sea ice concentration from OSISAF over the 1978-1992 and 2009-present periods and ESA over the 1993-2008 period. The ESA product will be chosen over OSISAF whenever available due to its high resolution, which will benefit our high-resolution reanalysis. To assess whether the heterogeneity of the observational datasets introduces undesired effects in the reanalysis, the OSISAF data will be assimilated in an additional 25-member simulation over the period 1993-2008. The reanalyses assimilating ESA and OSISAF will be compared over that period, offering an additional estimate of the impact of the usually underestimated observational uncertainty. While the minimum number of members to use in ensemble data assimilation is a topic of scientific debate, running 25 should be enough to obtain a good sampling and estimation of the model error covariance matrix. In order to reduce the necessary amount of computational resources, only 5 members will be run for the control experiment without data assimilation. Table 1 summarises the resources requested and detailed in section 2.3. These experiments will follow the schedule indicated in Table 2.

2.1.3 Atmospheric experiments

Using the high resolution sea ice reanalysis products generated within MAC, we will carry out global atmospheric data assimilation experiments with NICAM-LETKF (Terasaki et al. 2015) at 112-km horizontal resolution with 100 ensemble members for a three-month period. The NICAM-LETKF has been developed by the Data Assimilation Research Team in RIKEN AICS and used for cutting edge data assimilation studies (e.g. Miyoshi et al. 2015, Kotsuki et al. 2017a, b). The NICAM-LETKF uses the sea ice input from the National Centers for Environmental Prediction (NCEP)'s final analysis (FNL) data. As a first step, we will examine the impact of using the MAC sea ice reanalysis products instead of the NCEP FNL with the 112-km resolution NICAM-LETKF. The present NICAM-LETKF experiments show warm temperature biases in the Arctic region. We investigate if using more accurate sea ice products mitigates the warm biases and improves the atmospheric analyses in the Arctic region. Two experiments using the MAC reanalyses and NCEP FNL will be performed from August to October 2007, including the period when the lowest Arctic sea ice was recorded in September 2007 over the 1993-2008 period that the ESA product covers. Indeed, the ESA product is assimilated into the MAC sea ice reanalysis when available and it stands as the highest quality highest resolution product therefore ensuring the highest possible quality of sea ice reanalysis over the 1993-2008 period. As a second step, we will conduct high-resolution NICAM-LETKF experiments at 28-km horizontal resolution 40 ensemble members to examine the benefits of high-resolution sea-ice reanalysis data. We will conduct the high-resolution experiments for 18 days in September 2007.

2.1.4 Analyses

Disentangling externally-forced Arctic sea ice loss and internally-generated climate variability is a highly challenging task due to the short observational records and the limitations of statistical tools at hand. To make the most of our exceptionally long sea ice reanalysis, we will use the most advanced machine learning techniques for data mining. In particular, we will use the k-means clustering technique to identify Arctic sea ice modes of variability and assess their recurrence and persistence over the last decades (Fuckar et al, 2015). Emergent constraints on the future rate of Arctic sea-ice loss can be detected using techniques such as causal neural network (Kretschmer at al 2016). Typical forecast skill assessment scores will be applied to assess the added-value of our high-resolution sea ice reanalysis on NWP. The atmospheric forecast fields such as temperature and winds will be validated relative to the ECMWF (European Center for Medium range Weather Forecasting) reanalysis data.

2.1.5 Risk assessment and mitigation actions

The applying team has a strong experience in performing sea ice reanalyses (Massonnet et al, 2014; Guemas et al, 2014) at standard resolution and running high resolution experiments (Prodhomme et al, 2017) with the model we plan to use for the sea ice reanalysis. Therefore, the main risk for this project is a delay in porting the model code on the K computer. If this issue was to be hit, the number of members generated for our sea ice reanalysis would have to be reduced to allow the project to be achieved on time. This sea ice reanalysis would still be at the highest resolution used up to date for reanalysis as well as the longest existing sea ice reanalysis and using the highest quality highest resolution satellite products (ESA). The atmospheric assimilation system routinely runs on the K computer so that there is no risk related to this part of the project.

References:

Acosta MC, Yepes X, Valcke S, Maisonnave EM, Serradell K, Mula O, Doblas-Reyes FJ (2016). Performance Analysis of EC-Earth 3.2. *Tech. Report.*
Fuckar N S, Guemas V, Johnson N C, Massonnet F, Doblas-Reyes F J (2015) *Clim Dyn,* 45:1-17.
Good SA, Martin MJ, Rayner NA (2013) *J Geophys Res*, 118: 6704-6716, doi:10.1002/2013JC009067.
Guemas V, Doblas-Reyes FJ, Mogensen K, Keeley S , Tang Y (2014) *Clim Dyn*, 43(9-10):2813-2829.
Kotsuki S, T Miyoshi T, Terasaki K, Lien GY, Kalnay E (2017a) *J Geophys Res*, 122:631-650.
Kotsuki S, Ota Y, Miyoshi T, (2017b) *Quart J Roy Meteorol Soc*, in press. doi:10.1002/qj.3060
Kretschmer M, Coumou D, Donges JF, Runges J (2016) *J Climate*, 29:4069-4081.
Manubens-Gil D, Vegas-Regidor J, Prodhomme C, Mula-Valls O, Doblas-Reyes,FJ (2016) *2016 International Conference on High Performance Computing & Simulation (HPCS)*, doi:10.1109/hpcsim.2016.7568429.
Massonnet F, Goosse H, Fichefet T (2014) *Ocean Modelling,* 88:16-25, doi:10.1016/j.ocemod.2014.12.013.
Prodhomme C, Batte L, Massonnet F, Guemas V, Davini P, Doblas-Reyes F (2017) *J Climate*, doi/10.1175/JCLI-D-16-0117.1.
Miyoshi T, Kondo K, Terasaki K (2015) *Computer,* 48, 15-21, doi:10.1109/MC.2015.332.
Terasaki K, Sawada M, Miyoshi T (2015) *SOLA*, 11, 23-26. doi:10.2151/sola.2015-006.

2.2 Utilization plan of computational resources
2.2.1 The requested computational resources (for each computational resource)
Provider name/Resource name: RIKEN/K computer

Program name	Input data set name	Data size	#time steps	Normal execution confirmed?
NEMO3.6-LIM3	25km-grid resolution with 25 ensemble members + ESA sea ice concentration	1,475,000 grid points x 25 members	55.7 million	Yes
NICAM, LETKF	110-km-NICAM-LETKF with 100 ensemble members	409,600 grids x 100 members	3600 cycle x 2 times	Yes
NICAM, LETKF	28-km-NICAM-LETKF with 40 ensemble members	6,553,600 grids x 40 members	72 cycle x 2 times	Yes

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. Whereas the 25 members of the reanalysis need to run simultaneously since the ensemble Kalman Filter is applied to all members at the end of each month of simulation, the control simulation offer flexibility as it comprises 5 independent 39-year long simulations (description in section 2.1.2) which can run in parallel for an optimal use of the computing 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 10 TB in the "scratch" file system. This required scratch space is motivated by the large amount of output to be generated. These output data will be transferred immediately locally. Around 500 GB of "home" space will be required to host the code and its modified versions.

	Simulation years	Jobs / year	Cores / job Nodes / job	Walltime / job	Buffer to cope with potential errors	Total node-hours (millions)	Total archive (TB)
Porting	1					~0	0
Spin up	20	1	2048 cores = 256 nodes	3.61 hours	10%	0.0203	0.72
Control	195	1	1024 cores = 128 nodes	4.92 hours	20%	0.1461	7.02
Reanalysis	1375	1	1024 cores = 128 nodes	4.92 hours	30%	1.116	49.5
Atmosphere (112-km)	0.5	12	13440 cores = 1680 nodes	20 hours	0%	0.4032	20.
Atmosphere (28-km)	36 days	6 Jobs / 36 days	22400 cores = 2800 nodes	24 hours	0%	0.4032	30.
Total						**2.0889**	**107.24**

<u>Table 1</u>: Resources requested. These estimates have been obtained by running NEMO3.6-LIM3 on MareNostrum III and the resources have been multiplied by 2 to obtain an estimate on the K computer because the K computer nodes are half bigger than Marenostrum ones. The estimates for the atmospheric experiments are obtained from tests on the K computer.

(Utilization plan for the first half-year)

According to Table 2 planning the usage of resources, the first half will be dedicated to porting the code (1), running the spinup (2), the control experiment (3) and starting the sea ice reanalysis (4):

(1) Porting the code

To port and test the model version chosen for MAC, very few computing resources will be required. The plan is: 1) To install or get installed essential tools like SZIP, HDF5, NETCDF4, with respect to the available compiler and MPI environment, 2) To compile the model, 3) Perform a few trivial modifications of the available testing scripts that have been used on other HPC environments to adapt them to the K Computer environment, 4) To make and optimize a few of short runs of the ORCA025L75 configuration required for this project, in order to optimize some execution parametrizations. Porting and testing the model on the KComputer will require three to fours weeks in total. For safety, we make the planning considering this work will take 2 months instead of 1. The computing hours needed for this step are negligible with respect to what has been estimated for the experiments described below so that we do not account for these in the total requested computing time.

(2) Spinup

As described in section 2.1.2, the spinup experiment will comprise 20 years of NEMO3.6-LIM3 simulation. Based on the model performance at the Marenostrum 3 platform on which the model has already been tested, we estimate a need of ***0.0185*** million node hours to complete these 20 years of simulation using 2048 cores, that we inflate by 10% to account for possible errors, leading to ***0.0203*** million node hours. We choose 2048 cores for this simulation since all of the other simulations are dependent on this one so we want to make it as fast as possible. This configuration is less efficient but 50% faster than with 1024 cores that we use for the other experiments.

<u>Total :</u> 256 nodes * 20 years* 3.61 hours/year = 18,506 node-hours + (18,506 node-hours * 0,1) = 20,357 node-hours

(3) Control

As described in section 2.1.2, the control experiment will comprise 195 years of NEMO3.6-LIM3 simulation. Based on the model performance at the Marenostrum 3 platform, we estimate a need of ***0.1218*** million node hours to complete these using 1024 cores, that we inflate by 20% to account for possible errors, leading to ***0.1461*** million node hours. We choose here 1024 cores as well as in the next simulations because it is a good trade-off between efficiency (about 70%) and speed.

Total : 128 nodes * 195 years * 4.92 hours/year = 121,756 node-hours + (121,756 node-hours * 0,2) = 146,107 node-hours

(4) Sea ice reanalysis

As described in section 2.1.2, the sea ice reanalysis will comprise 1375 years of NEMO3.6-LIM3 simulation with the Ensemble Kalman Filter activated. Based on the model performance at the Marenostrum 3 platform, we estimate a need of ***0.8585*** million node hours to complete these. We inflate this amount by as much as 30% because the dependence between the 25 members might force us to rerun several members in case of errors, which leads to ***1.116*** million node hours using 1024 cores. About 30% of these resources will be used in the first half year, i.e. ***0.3348*** million node hours.

Total: 128 nodes * 1375 years * 4.92 hour/year = 858,537 node-hours + (858,537 node-hours * 0,3) = 1,116,008 node-hours

(Utilization plan for the second half-year)

According to Table 2 planning the usage of resources, the second half will be dedicated to completing the sea ice reanalysis (5) and performing atmospheric sensitivity experiments:

(5) Sea ice reanalysis

The remaining 70% of the resources required to complete the sea ice reanalysis will be used in the second half year, i.e. ***0.7812*** millions node hours.

(6) Atmosphere experiments

The atmospheric experiments need 403,200 node hours for each of the 112-km and 28-km experiments. We plan to conduct two (control and test) experiments to assess the added-value of MAC sea-ice reanalysis product.

NICAM (112-km): 280 node hours per cycle ; LETKF (112-km) : 280 node hours per cycle

NICAM-LETKF (112-km): {(280+280) node hours / cycle} x {360 cycles / 3 months } = 201,600 node-hours / half year

Total = 403,200 node hours

NICAM (28-km): 1400 node hours per cycle ; LETKF (28-km) : 1400 node hours per cycle

NICAM-LETKF (28-km): {(1400+1400) node hours / cycle} x {72 cycles / 18 days} = 201,600 node-hours / 15 days

Total = 403,200 node hours

	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12
Porting	■	■										
Spin up			■									
Control				■	■	■						
Reanalysis				■	■	■	■	■	■	■	■	
Atmosphere											■	■

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

==

Requested resource amount
Provider name / Resource name:
First half-year (Oct. 2017-Mar. 2018): ***501 266*** node-hours
Second half-year (Apr. 2018-Sep. 2018): ***1 587 606*** node-hours
Total (Oct. 2017-Sep. 2018): ***2 088 872*** node-hours

==

References:

Asif M,Cencerrado A, Mula-Valls O, Manubens D, Doblas-Reyes FJ, Cortés A (2014) *Procedia Computer Science*, 29;2370-2379, doi:10.1016/j.procs.2014.05.221.

Balaji V, Maisonnave E, Zadeh N, Lawrence BN, Biercamp J, Fladrich U, Wright G (2016) *Geoscientific Model Development Discussions*, 1-25. doi:10.5194/gmd-2016-197.

Tintó-Prims O, Castrillo M, Serradell K, Mula-Valls O, Doblas-Reyes FJ (2015) *BSC-CES Technical Memorandum,* 2015-002.

2.3 Preparation status of the application program
Application program name: a) NEMO3.6-LIM3 ; b) NICAM, LETKF

1) Has the development of the program been finished?
Have you verified that the calculation results of the program are correct, irrespective of the computer and the kind of data?

a) The development of the NEMO3.6-LIM3 model code has finished and this model is routinely used by the applying institute, BSC. It has already been well deployed on MareNostrum III (a tier-0 system; having 2x Intel SandyBridge-EP E5-2670/1600 8-core processor at 2.6 GHz, 8x4GB DDR3-1600 DIMMS (2GB/core) and Infiniband FDR10 network). The BSC has been involved in the IS-ENES/PRACE-1IP working group focusing on the NEMO adaptation to Tier-0 machines. The BSC is also involved in the NEMO System Team, and is active member of the NEMO HPC Group, having developed and implemented several optimizations to improve the parallel efficiency of the model (Tintó et al, 2016). These optimizations are now part of the stable and development versions, and are core for ongoing and future developments. BSC has also tested a range of NEMO configurations on several HPC systems: SGI Altix 3500, NEC-SX6, Linux cluster with Intel Xeon, Dell PowerEdge 2900, IBM pSeries 575 Power6, IBM Power PC and IBM Blue Gene/Q. The BSC leads the development of Autosubmit, a python-based wrapper software that can manage any type of climate simulation workflow. It can also bundle numerous jobs into a single big job for performing job-control on parallel sets of simulations throughout the execution. The BSC has also expertise in analyzing parallel programming model codes using profiling tools. This allows continuously carrying out performance analyses to reach the optimum configurations for EC-Earth and NEMO.

b) Yes. The development of the NICAM-LETKF code has finished. The NICAM-LETKF system routinely runs on the K computer under ongoing project IDs hp160229 and hp170178.

2) Have you verified that the calculation results of the program are correct on the requested computers?

a) NEMO3.6-LIM3 has never run on the K-computer. Hence, two months of this one-year project will be dedicated to porting the model code and scaling it up to optimize the run time as described in section 2.2.1. Afterwards, a short test will run at the K computer will be compared to an equivalent run at Marenostrum 3, briefly described under section 2.4, to verify that the differences between the two tests arise only from the machine precision. If not, a thorough analyses of the model code and the compilation option that could lead to such differences will be carried out during the one-month buffer we have kept in the first half-year period of the project to sort this out.

b) Yes. The NICAM-LETKF system routinely runs on the K computer. The system has been verified to run correctly under ongoing project IDs hp160229 and hp170178

3) Have you verified that the calculation results of the program are correct on the requested computers using the input data set of the similar kind and scale to those for the target?

a) We have already verified that the calculation results of NEMO3.6-LIM3 are correct on

Marenostrum 3 machine, which is briefly described under section 2.4.

b) Yes. The NICAM-LETKF system has been tested on K-computer under ongoing project IDs hp160229 and hp170178

4) Is the estimated parallel efficiency for the target value of the number of nodes sufficient in carrying out the project? (50% or more expected.)

a) The target value of the number of nodes: 128 nodes
Estimated parallel efficiency: 69,84 %

b.1) LETKF) The target value of the number of nodes: 1,800 node
Estimated parallel efficiency: 56.689906 %

b.2) NICAM) The target value of the number of nodes: 60,000 node
Estimated parallel efficiency: 93.387423 %

The estimated parallel efficiency for the target value of the number of nodes is sufficient to carry out this project.

2.4 The necessity of the requested computational resources

Given the large amount of simulation years and the groundbreaking high resolution of these simulations, such a project would only be feasible with a HPC machine such as the K computer.

	ECMWF-CCA	MareNostrum3	KComputer
Motherboard	Cray XC40 system	IBM dx360 M4	Fujitsu
Processor	Dual 18-core Intel Xeon EP E5-2695 V4 (Broadwell) series processors (2.10 GHz), 36 cores per node	Intel SandyBridge-EP E5-2670 (2.6 GHz), 16 cores per node	SPARC64 VIIIfx 8C (2GHz), 8 cores per node
Main memory	128 GB per node	32 GB per node	16 GB per node
Interconnect	Cray Aries interconnect links all compute nodes in a Dragonfly topology	Infiniband (IB)	Tofu Interconnect (6D Mesh/Torus)
Operating system	Cray Linux Environment (CLE)	Linux - SuSe Distribution 11 SP2	Linux
NEMO3.6-LIM3 (one month test run with ORCA025 configuration)	0.57 wall-clock hours using 1024 procs	0.40 wall-clock hours using 1024 procs	?

2.5 Organizational plan for the project execution, and staffing plan

(Organization Chart)

Organization	Roles
BSC	Sea ice data assimilation and generation of an ensemble sea ice reanalysis
RIKEN	Impact of sea ice reanalysis on weather forecast quality

(The above table may replace to a diagram.)

(Project Members)

Project member	Affiliation	Specialization	Whether to use the resources	Roles
Virginie Guemas	BSC	Sea ice	no	Principal Investigator (PI)
Takemasa Miyoshi	RIKEN	Data assimilation	no	co-PI
Kim Serradell	BSC	High Performance Computing	no	co-PI
Miguel Castrillo	BSC	High Performance Computing	yes	Porting NEMO3.6-LIM3 code
Pablo Echevarria	BSC	High Performance Computing	yes	Porting NEMO3.6-LIM3 code
Mario Acosta	BSC	High Performance Computing	yes	Porting NEMO3.6-LIM3 code
Juan Acosta	BSC	Sea ice	yes	Running sea ice reanalysis
Neven Fuckar	BSC	Sea ice	yes	Running sea ice reanalysis
Shunji Kotsuki	RIKEN	Data assimilation	yes	Atmospheric simulations
Koji Terasaki	RIKEN	Data assimilation	yes	Atmospheric simulations
Keiichi Kondo	RIKEN	Data assimilation	yes	Atmospheric simulations

(Experience and Achievement)

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), one ECMWF-funded project (HighResMIP_BSC) and she is WP leader in the H2020 PRIMAVERA project. 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.

Dr Takemasa Miyoshi is the Team Leader of the Data Assimilation Research Team in RIKEN Advanced Institute for Computational Science (AICS). Dr. Miyoshi's scientific achievements include more than 80 peer-reviewed publications and more than 50 invited conference presentations including the Core Science Keynote at the American Meteorological Society Annual Meeting (2015). Dr. Miyoshi has been recognized by several prestigious awards such as the Yamamoto-Syono Award by the Meteorological Society of Japan (2008) and the Young Scientists' Prize by the Minister of Education, Culture, Sports, Science and Technology (2014), the Japan Geosciences Union Nishida Prize (2015), and the Meteorological Society of Japan Award (2016).

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 Juan Camilo Acosta Navarro is currently a scientific researcher in the Climate Prediction group within the Earth Sciences department at BSC. His work is focused on sea ice process and how sea ice may affect climate remotely. He has published nine scientific peer reviewed articles, being first author in three of them. He has been adviser to two MSc students.
Dr Neven Fuckar is a Juan de la Cierva Postdoctoral Fellow in the Climate Prediction Group of the Department of Earth Sciences at the Barcelona Supercomputing Center. He obtained a Ph.D in Atmospheric and Oceanic Sciences at Princeton University, Princeton, NJ, USA. Dr. Fuckar research focus is on sea ice and climate dynamics and predictions from seasonal to decadal time scales. He has been the Principal Investigator of 3 national (RES) supercomputing projects and the convener of Polar Climate Predictability and Prediction session at the EGU general assembly in Vienna. Dr. Fuckar 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 523 citations on Google scholar, and h-index and i10-index of 9).
Drs Koji Terasaki, Shunji Kotsuki and Keiichi Kondo are Research Scientist and Postdoctoral Researchers, respectively, in the Data Assimilation Research Team, RIKEN AICS, and conducted many atmospheric data assimilation experiments using the K computer.

(Staffing plan & training)
October : Miguel Castrillo, Pablo Echevarria and Mario Acosta will be given the “K-computer users' Lecture”
December : Neven Fuckar and Juan Acosta will be given the “K-computer users' Lecture”

3 Research information relevant to this proposal
3.1 Ongoing Projects/Continuing past project

3.2 Related projects
Research project using the HPCI Systems.
Ongoing project ID: hp170246
Ongoing project name: Advancement of meteorological and global environmental predictions utilizing observational “Big Data”
Status of achievement toward the goal of the project: This project aims to develop new-generation data assimilation systems using observational Big Data such as satellite observations. The NICAM-LETKF system has been developed and advanced under this project.

Ongoing project ID: hp170178

Ongoing project name: "Big Data Assimilation" studies for predicting sudden severe rainstorms
Status of achievement toward the goal of the project: This project aims to predict sudden severe rainstorms by assimilating the dense and frequent phased array radar observations into the regional NWP systems. The project is going to carry out quasi-real-time NICAM-LETKF experiments to provide boundary conditions to the regional NWP systems.

Research plan for collaborating with large experimental facility projects.
FLAGSHIP2020 Priority Issue #4 : The NICAM-LETKF was selected to be the target software of the co-design to improve the computer hardware, the system software and applications to resolve their bottlenecks and achieve best performance for post-K supercomputer.

4 Research achievements references

Selected publications:

1. **Kotsuki S,** Ota Y, **Miyoshi T,** (2017) Adaptive covariance relaxation methods for ensemble data assimilation: Experiments in the real atmosphere. *Quart J Roy Meteorol Soc,* in press. doi:10.1002/qj.3060.
2. **Kotsuki S, Miyoshi T, Terasaki K,** Lien GY, Kalnay E (2017) Assimilating the Global Satellite Mapping of Precipitation Data with the Nonhydrostatic Icosahedral Atmospheric Model NICAM. *J Geophys Res,* 122, 631-650. doi:10.1002/2016JD025355.
3. Massonnet F, Bellprat O, **Guemas V,** Doblas-Reyes F (2017) Using climate models to estimate the quality of global observational data sets. ***Sciences***, doi : 10.1126/science.aaf6369.
4. Bellprat O, Massonnet F, García-Serrano J, **Fuckar N, Guemas V,** Doblas-Reyes F (2016) The role of Arctic sea ice and sea surface temperatures on the cold 2015 February over North America. ***Bull. Amer. Meteor. Soc***, 97, S36-S41, doi:10.1175/BAMS-D-16-0159.1.
5. **Fuckar N**, Massonnet C, **Guemas V**, Garcia-Serrano J, Bellprat O, Doblas-Reyes F, Acosta M (2016) Record low northern hemisphere sea ice extent in March 2015. ***Bull. Amer. Meteor. Soc,*** 97, S136-S140, doi:10.1175/BAMS-D-16-0153.1.
6. **Guemas V,** Chevallier M, Deque M, Bellprat O, Doblas-Reyes F J (2016) Impact of sea ice initialisation on sea ice and atmosphere prediction skill on seasonal timescales. *Geophys Res Lett,* 43 (8), 3889-3896, doi:10.1002/2015GL066626.
7. **Miyoshi T, Kondo K, Terasaki K** (2015) Big Ensemble Data Assimilation in Numerical Weather Prediction. *Computer,* 48, 15-21. doi:10.1109/MC.2015.332.
8. **Terasaki K,** Sawada M, **Miyoshi T** (2015) Local Ensemble Transform Kalman Filter Experiments with the Nonhydrostatic Icosahedral Atmospheric Model NICAM. *SOLA*, 11, 23-26. doi:10.2151/sola.2015-006.
9. **Fuckar N S, Guemas V,** Johnson N C, Massonnet F, Doblas-Reyes F J (2015) Clusters of interannual sea ice variability in the Northern Hemisphere. *Clim Dyn,* 45, 1-17, doi:10.1007/s00382-015-2917-2.
10. Massonnet F, **Guemas V, Fuckar N S,** Doblas-Reyes F J (2015) The 2015 high record of Antarctic sea ice extent [in "Explaining Extreme Events of 2014 from a Climate Perspective"]. ***Bull. Amer. Meteor. Soc.***, 96 (9), S163-S167, doi:10.1175/BAMS-D-15-00093.1.
11. Stroeve J, Blanchard-Wrigglesworth E, **Guemas V,** Howell S, Massonnet F, Tietsche S (2015) Improving Predictions of Arctic Sea Ice Extent. ***EOS,*** 96, doi:10.1029/2015EO031431.
12. Jung T, Doblas-Reyes FJ, Goessling H, **Guemas V,** Bitz C, Buontempo C, Caballero R, Jokobsen E, Karcher M, Koenigk T, Matei D, Overland J, Spengler T, Yang S (2015) Polar-lower latitude linkages and their role in weather and climate prediction. ***Bull. Amer. Meteor. Soc.***, 96, ES197-ES200, doi:10.1175/BAMS-D-15-00121.1.
13. **Guemas V,** Blanchard-Wrigglesworth E, Chevallier M, Day J J, Déqué M, Doblas-Reyes F J, **Fučkar N**, Germe A, Hawkins E, Keeley S, Koenigk T, Salas y Mélia D, Tietsche S (2015) A review on Arctic sea ice predictability and prediction on seasonal-to-decadal timescales, *Quart. J. Roy. Meteorol. Soc.*, doi:10.1002/qj.2401.
14. **Guemas V,** Auger L, Doblas-Reyes FJ, Rust H, Ribes A (2014) Dependencies in Statistical Hypothesis Tests for Climate Time Series. ***Bull. Amer. Meteor. Soc,*** 95 (11), 1666-1667.
15. **Guemas V,** Doblas-Reyes F J, Mogensen K, Keeley S, Tang Y (2014) Ensemble of sea ice initial

conditions for interannual climate predictions. *Clim. Dyn.,* 43(9-10), 2813-2829, doi:10.1007/s00382-014-2095-7.

16. **Guemas V,** Doblas-Reyes F, Germe A, Chevallier M, Salas y Mélia D (2013) September 2012 Arctic sea ice minimum : Discriminating between sea ice memory, the August 2012 extreme storm and prevailing warm conditions [in "Explaining Extreme Events of 2012 from a Climate Perspective"], ***Bull. Amer. Meteor. Soc.,*** 94 (9), S20-S22.

17. **Guemas V,** Doblas-Reyes F J, Andreu-Burillo I, Asif M (2013) Retrospective prediction of the global warming slowdown in the past decade. ***Nature Climate Change,*** 3, 649-653, doi : 10.1038/nclimate1863.

18. Doblas-Reyes F J, Andreu-Burillo I, Chikamoto Y, García-Serrano J, **Guemas V,** Kimoto M, Mochizuki T, Rodrigues L R L, van Oldenborgh G J (2013), Initialized near-term regional climate change prediction. ***Nature Communications,*** 4, 1715, doi:10.1038/ncomms2704.

5 Validity to perform the project

We confirm that the execution of the project and utilization of the results will be limited to peaceful purposes and proper in terms of Basic Act on Science and Technology and social standards. MAC is a research project which results are not intended to commercial use but rather intended to contributing to society through the provision of user-relevant climate information.

Eligibility confirmation of Junior Researcher Promotion Project Applicant
* To apply for the junior researcher promotion project, it must be less than or equal to 39 years of age at the time Oct 1, 2017 (born on and after Oct 2 1977). Only if you want to apply to it, please fill in the age of the project representative.

Age	(as of October 1, 2017)