

IS-ENES2
Project final report



Grant Agreement number: **312979**

Project acronym: **IS-ENES2**

Project title: **Infrastructure for the European Network for Earth System Modelling, Phase 2**

Funding Scheme: **Combination of CP & CSA**

Period covered: **from 01/04/2013 to 31/03/2017**

Name of the scientific representative of the project's coordinator: **Sylvie Joussaume**

Title and Organisation: **Dr, CNRS-IPSL**

Tel: **+33 1 69 08 56 74**

Fax: **+33 1 69 08 30 73**

E-mail: sylvie.joussaume@lsce.ipsl.fr

Project website address: <https://is.enes.org>

Date: 31/05/2017



Partner n°	Partner organisation name	Short name	Country
1	Centre National de la Recherche Scientifique Institut Pierre Simon Laplace	CNRS - IPSL	France
2	Deutsches Klimarechenzentrum GmbH	DKRZ	Germany
3	Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique	CERFACS	France
4	Centro Euro-Mediterraneo per i Cambiamenti Climatici	CMCC	Italy
5	University of Reading National Centre for Atmospheric Science	UREAD - NCAS	United Kingdom
6	Met Office	MetO	United Kingdom
7	Science and Technology Facilities Council	STFC	United Kingdom
8	Sveriges Meteorologiska och Hydrologiska Institut	SMHI	Sweden
9	Koninklijk Nederlands Meteorologisch Instituut	KNMI	Netherlands
10	Max Planck Institute for Meteorology	MPG	Germany
11	Climate System Analysis Group, University of Cape Town	CSAG	South Africa
12	University of Manchester	UNIMAN	United Kingdom
13	Institutul National de Hidrologie si Gospodarie a Apelor	INHGA	Romania
14	Wageningen Universiteit	WU	Netherlands
15	Linköpings Universitet	LiU	Sweden
16	Barcelona Supercomputing Centre	BSC	Spain
17	Universidad de Cantabria	UC	Spain
18	Deutsches Zentrum Für Luft- und Raumfahrt in der Helmholtz Gemeinschaft	DLR	Germany
19	Danish Meteorological Institute	DMI	Denmark
20	Institut Català de Ciències del Clima	IC3	Spain
21	Météo France Centre National de Recherches Météorologiques	MF - CNRM	France
22	Universitetet i Bergen	UiB	Norway
23	Norwegian Meteorological Institute	met.no	Norway

Content

Executive Summary.....	5
Summary description of project context and objectives	6
Description of the main S&T results.....	9
1. Foster the integration of the European climate and Earth system modelling community.....	9
1.1. Strengthening governance	9
1.1.1. ENES governance.....	9
1.1.2. International governance	10
1.2. Community building	10
1.2.1. European training school on climate modelling	10
1.2.2. The ENES Portal	11
1.3. ENES infrastructure strategy.....	11
1.3.1. Evaluation infrastructure strategy.....	11
1.3.2. Mid-term update	12
1.3.3. Planning for a long-term sustained European research infrastructure	13
2. Enhance the development of Earth System Models.....	13
2.1. Services around model environment tools and components	13
2.1.1. Service on OASIS coupler.....	14
2.1.2. Service on CDO Climate Data Operators.....	14
2.1.3. Service on NEMO ocean platform	15
2.1.4. Service on climate models.....	16
2.2. Towards next generation climate models	16
2.2.1. Towards common radiation tools	17
2.2.2. Benchmarking coupling technologies	17
2.2.3. Code/software convergence Chasm workshop.....	18
2.2.4. Computational cores and new parallel approaches for NEMO and ICON	19
2.2.5. Testing revolutionary and evolutionary approaches for NEMO.....	20
2.3. Sharing best practices for model environments.....	21
2.3.1. Workflow tools.....	21
2.3.2. Configuration management	22
2.3.3. Meta-data generation during experiments	23
2.3.4. Governance on common software	23
3. Foster high-end simulations enabling to better understand and predict future climate change	24
3.1. Tracking future computer architecture technologies	24
3.2. Improve model performance for HPC.....	24
3.2.1. Computational performance standards	24
3.2.2. Improving parallel I/O solutions and spreading its usage.....	26
3.2.3. Post-processing.....	27
3.2.4. Impact of OASIS update	27

3.3.	Prepare future high-end experiments	27
3.3.1.	Workflow management system for running ensembles of simulations	28
3.3.2.	M4HR demonstrator.....	28
3.4.	Develop coupled benchmarks	29
3.5.	Support innovation with HPC vendors.....	29
4.	Support the dissemination of Earth system model data to the climate and impact research communities.....	30
4.1.	Service of the ENES climate data infrastructure	31
4.1.1.	The ENES climate data infrastructure	31
4.1.2.	Earth System Grid Federation overview	31
4.1.3.	IS-ENES2 contribution to ESGF/CMIP	32
4.1.4.	IS-ENES2 contribution to ESGF/CORDEX	33
4.1.5.	Key performance Indicators for ESGF European datanodes	34
4.1.6.	Helpdesk service	35
4.1.7.	Service on metadata.....	35
4.1.8.	Service for climate impact communities: the climate4impact portal.....	36
4.2.	Development of core services on ESM data (ESGF)	36
4.2.1.	Contribution to ESGF development	36
4.2.2.	Software developments: Quality control	37
4.2.3.	Service monitoring and dashboard	38
4.3.	Development of Metadata standards and services.....	38
4.4.	Development of data access services for climate impacts.....	39
4.5.	Innovation on data: Linking with climate services and corporates	40
5.	Conclusions and perspectives	41
	Impact of results	42
1.	Impacts on European community on Earth System Modelling.....	42
2.	Impacts on other research communities	43
3.	Impacts at the international level	44
4.	Impacts on innovation.....	44
	Dissemination and exploitation of results.....	46
	Exploitable foreground list	47
	Annex I: Full List of publications and communications	51
	Acronyms.....	57

Executive Summary

IS-ENES2 is the second phase of the distributed infrastructure of the European Network for Earth System Modelling (ENES). This network links and represents the European modelling community working on understanding and predicting climate variability and change. The main objectives of IS-ENES2 were to foster the integration of the global and regional climate modelling community, stimulate common development and sharing of software for models and their environments, foster the execution and exploitation of high-end simulations and support the dissemination of model results to the climate research, climate impact and climate services communities. In doing so, the IS-ENES2 infrastructure supports the European contribution to the international experiments of the World Climate Research Program (WCRP). The resulting data is used extensively in assessments of the Intergovernmental Panel on Climate Change, and provides modelling results on which EU mitigation and adaptation policies are built.

Foster the integration of the European climate and Earth system modelling community

IS-ENES2 has achieved a key step in the integration of the global and regional climate modelling community through support to the WCRP international CMIP and CORDEX programmes (the Coupled Model Intercomparison Projects and Coordinated Regional Downscaling Experiments). IS-ENES2 played a crucial role in the international governance and development of the associated databases and metadata systems. A mid-term update of its 2012-2022 infrastructure strategy confirmed recommendations on models, data, computing, physical network, and people. It added the needs to develop both the infrastructure dimension of model evaluation as well as more sustainable approaches to the infrastructure.

Enhance the development of Earth System Models for the understanding of climate variability and change

Support services for modelling and related environmental software started in IS-ENES1 have been extended to include access to climate models. IS-ENES2 has also fostered common development and sharing of expertise to help improve models and the associated workflow. A common approach to radiative transfer has been agreed upon and a prototype library developed. Dedicated workshops on environmental software tools, such as those needed for configuration management, metadata creation and workflow, have been an opportunity to share experience and best practices between software engineers from different modelling groups. For the first time, different coupling technologies from Europe and USA have been benchmarked using a common approach.

Support high-end simulation to better understand and predict climate variations and change

The climate modelling community has always been facing the challenge of ensuring that models will be able to follow the evolution of computer architectures. In the age of “Exascale”, this will be even more demanding than in the past. IS-ENES2 has helped engender technology tracking mostly through workshops involving computer vendors. A suite of coupled benchmarks has been developed that will ease interaction with vendors and testing of new architectures. A first international metric to measure model performance has been elaborated and tested in partnership with US colleagues. Model performance on parallel systems has been improved through coupler and IO servers, and an ensemble of multi-model multi-member high-resolution models tested successfully. Collaboration within IS-ENES2 has paved the way to the establishment of a Center of Excellence in Simulation for Weather and Climate, ESiWACE, that will specifically address the performance of models for exascale computers.

Support the dissemination of Earth system model simulations

IS-ENES2 teams lead several of the international developments of the CMIP and CORDEX common data infrastructure. IS-ENES2 has led the development of standards for CORDEX data and hosts most of the available data. IS-ENES2 has also had a major contribution in the international upgrade of the metadata services in preparation of the next CMIP Phase (CMIP6). IS-ENES2 expertise on ESGF is now well recognised within Copernicus Climate Change Service and several partners are engaged in preparing ESGF datanodes to access climate projections. Via this service, IS-ENES2 has further developed a dedicated access to climate model data for the community working on climate change impacts. The climate4impact portal has been successfully implemented and new functionalities added, such as the on-line computation of climate indices and downscaled data, data discovery, visualisation and guidance.

Summary description of project context and objectives

IS-ENES2 is the second phase project (04/2013-03/2017) of the infrastructure of the European Network for Earth System modelling (ENES). ENES gathers the European community developing and exploiting models of the Earth's climate system. This community aims to better understand climate variability and change under past, present and future conditions. It is strongly involved in the assessments of the Intergovernmental Panel on Climate Change (IPCC) and provides the predictions on which EU mitigation and adaptation policies are built. The infrastructure of ENES supports the international experiments organised under the auspices of World Climate Research Program (WCRP). These experiments constitute a reference for model evaluation, for better understanding mechanisms and processes and for climate projections of possible future climate change.

During its first phase, the IS-ENES project (FP7, 03/2009-02/2013), referred in the following as IS-ENES1, concentrated on global climate models and on the European contribution to the international WCRP experiments of the fifth phase of the Coupled Modelling Intercomparison Project (CMIP5). The ENES Portal was developed, along with a prototype of a dedicated portal for the impact communities. Service activities on model components and tools as well as on model data archives were launched. A strategy for the ENES infrastructure for 2012-2022, onto which IS-ENES2 is based, was developed. A close collaboration with the Partnership for Advanced Computing in Europe (PRACE) was established.

Project context

ENES faces the challenges to:

- Perform the most up-to-date and accurate climate simulations. This requires sophisticated models, world-class high-performance computers (HPC) and archiving systems, and state-of-the-art software infrastructure to make efficient use of the models, the data and the hardware.
- Better integrate the European climate modelling community in order to speed-up the development of models and the use of high-performance computers, improve the efficiency of the modelling community and improve the dissemination of model results to a large user base, including climate services.

These challenges have increased over the last years with the increasing need to prepare for adaptation, the need to develop reliable regional decadal prediction, the emergence of climate services providing tailored climate information to society, and the technical challenges even of today's computer architectures. IS-ENES1 had already taken key steps towards meeting these challenges but further achievements were still needed.

IS-ENES2 objectives

Following the ENES infrastructure strategy recommendations, IS-ENES2 aimed to strengthen the integration between global and regional modeling by supporting not only CMIP experiments, but also the WCRP CORDEX designed to provide climate information at regional scale. IS-ENES2 also aimed to support the development of international standards for data and metadata within WCRP, following on the FP7 METAFOR (2008-2011) e-infrastructure project for metadata. IS-ENES2 has continued to aim at easing the use of climate model data by the climate impact research community. It has also aimed at strengthening the ENES community capacity to provide more reliable decadal predictions at regional scale for society.

IS-ENES2 has been organised around four main objectives:

- **To foster the integration of the European climate and Earth system modelling community** by strengthening the ENES governance, further developing its strategy, especially with regards to model evaluation and model developments, stimulating interactions between global and regional climate modelling communities, and developing training (WP2/NA1)

- **To enhance the development of Earth System Models for the understanding of climate variability and change** by networking on future model developments required to improve model quality and use of future computing architectures, by stimulating common software developments and by providing a service on models and tools (WP3/NA2, WP4/NA3 and WP7/SA1)
- **To support high-end simulations enabling us to better understand and predict climate variations and change** by preparing for future exascale computing architectures (WP3/NA2), by preparing multi-model high resolution common experiments on the European PRACE high-performance computer facilities (WP9/JRA1), and by stimulating collaboration with ICT companies (WP6/NA5 & WP10/JRA2)

To facilitate the application of Earth system model simulations to better predict and understand the climate system and climate change impacts on society by enhancing the dissemination of model results from both global and regional model experiments (WP8/SA2), by developing an interface dedicated to the climate impact community and improving the quality of information on simulations through metadata developments and guidance to users (WP5/NA4 & WP11/JRA3) and by enhancing interaction between the climate modelling activity and users from companies and the emerging climate services (WP6/NA5).

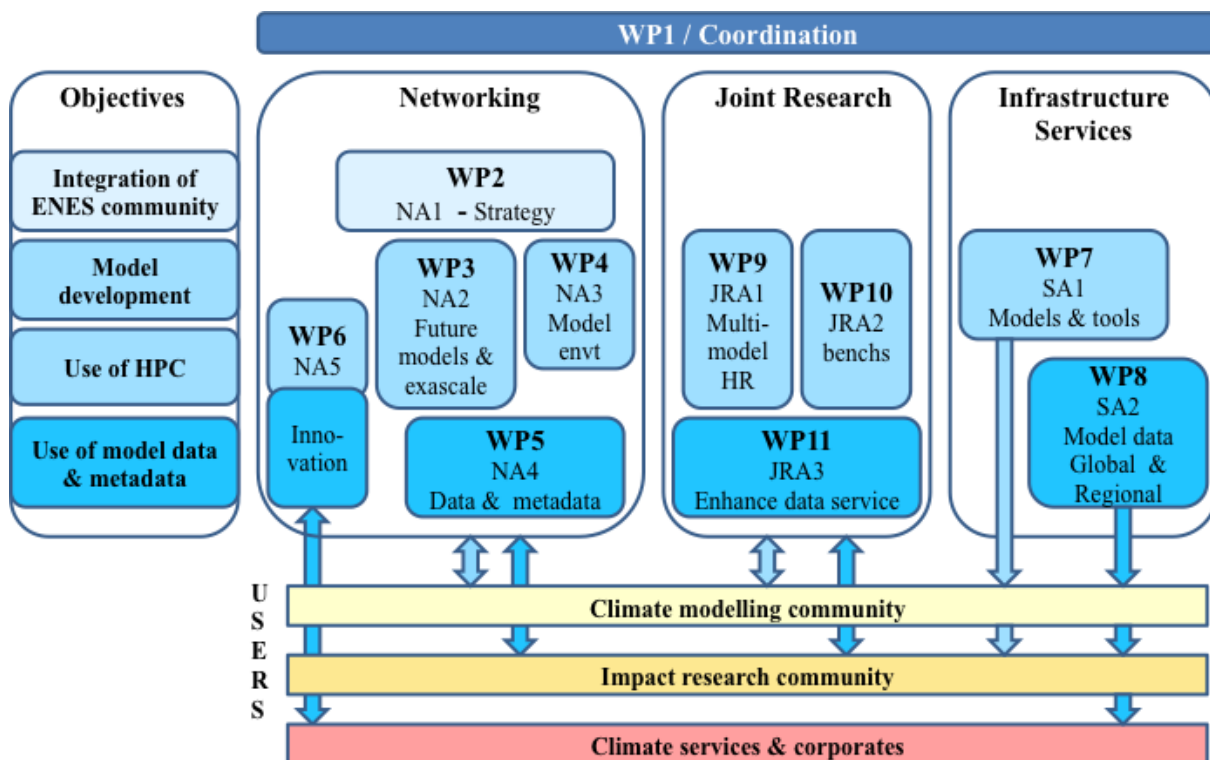


Figure 1: Work package activities with respect to the project's objectives and relations with user communities.

These four objectives have been attended in:

- **Networking activities**, aiming at increasing the cohesion of the European ESM community and at advancing a coordinated European Network for Earth System modelling.
- **Joint research activities** aiming at improving the efficient use of ESMs on high-performance computers, the access to model results in terms of data and metadata and the development of international databases and standards.
- **Services activities** provided on models and model data and metadata to both the climate modelling community and to the users of model results, including the impacts community.

IS-ENES2 also aimed at supporting a specific activity to enhance innovation through ICT technologies and through use of model results for emerging European Climate Services and corporations.

To achieve these objectives, IS-ENES has combined expertise in climate and Earth system modelling, in computational science, and in studies of climate change impacts, gathering about 230 experts from 23 institutions. The ENES Portal (<https://portal.enes.org>) has continued to be the key component to communicate and the IS-ENES2 project web site (<http://is.enes.org>) has delivered the results of the different work packages to the climate community and to users of climate model results.

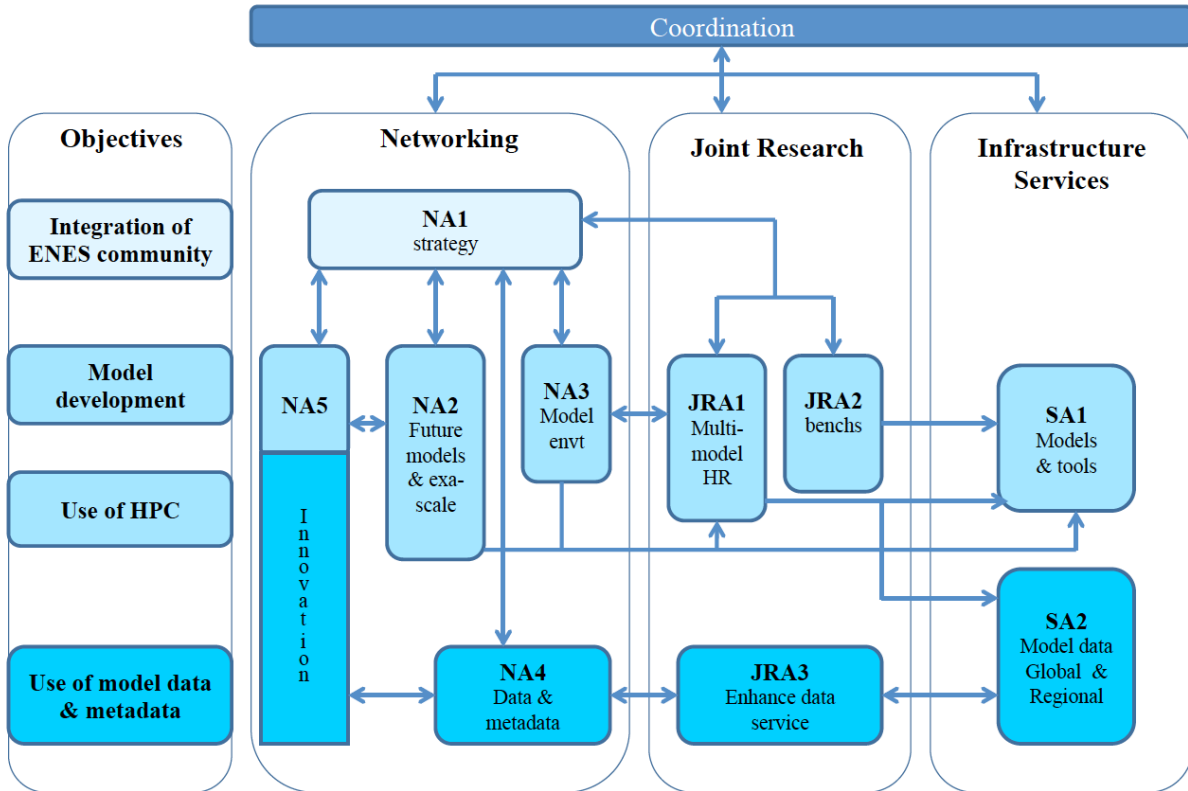


Figure 2: Relationship of work package activities to each other and the major networking, joint research, and infrastructure activities.

Description of the main S&T results

1. Foster the integration of the European climate and Earth system modelling community

The ENES community has a common goal to work towards minimizing duplication in development, while providing up-to-date infrastructure and advancing science in climate modelling. To achieve this goal, ENES aims at strengthening its governance at different levels to foster synergies and agree strategies, especially with respect to model development and evaluation, by stimulating interactions between global and regional climate modelling communities, and developing opportunities for training and exchange.

1.1. Strengthening governance

Emerging challenges for the climate modelling community determine for ENES an increased need for an increased coordination of: model development, access to European high-performance computers (HPC), and climate model data provision, while jointly developing strategies for the future of the infrastructure.

1.1.1. ENES governance

ENES governance activities involve, primarily, the ENES HPC and Data Task Forces (**D2.1**). These task forces advise ENES on all issues relevant to High Performance Computing and to data infrastructure, to support and exploit simulations of the Earth's climate in Europe. For the first time, an ENES Scientific Officer has been recruited to drive the overall ENES governance.

The ENES HPC Task Force (TF) is currently composed of 17 members, primarily representative of the European climate modelling groups although it was recently extended to include representation of the Numerical Weather Prediction community. The HPC TF tackles community issues related to the use of the European HPC ecosystem, in particular with regards to PRACE European facilities, as well as issues related to software development, cooperation with vendors, and future initiatives¹. The HPC TF is not limited to IS-ENES2, it provides an integrating focus for all ENES HPC related activities.

Within IS-ENES2, the HPC-TF used its expertise on the community and technology tracking in support of the organization of the 3rd and 4th (jointly with ESIWACE) ENES HPC workshops. It was also instrumental in supporting the discussion and preparation of possible CMIP6 high-end experiments on PRACE. Learning from CMIP5, the TF has emphasized the needs for multi-year access, stable resources during experiments and community specific storage services. Dedicated access for high-end experiments was discussed with the PRACE Council and led to the preparation of a proposal for CMIP6 high-resolution simulations, which unfortunately could not fit within available resources. The HPC-TF also supported a community statement on the importance of Fortran programming language in Earth system modelling as a reaction to the tendency in the HPC community to propose the use of "new" programming languages needing to re-engineer applications for the ever-changing modern architectures. The identification of current and anticipated future HPC needs was integral to the preparation of the strategy for the whole infrastructure for the next 10 years (see 1.3.2).

The ENES Data Task Force has been created within IS-ENES2 and is currently composed of 10 members. The Data TF deals with data and metadata standards, and the computing and services necessary to exploit the data and information from simulations of the Earth's climate system. Within that it is responsible for a range of activities, including establishing rules, procedures, needed activities, and priorities, as well as identifying necessary tools and services. An important aspect includes coordinating European scale activities (such as IS-ENES2) with national activities, especially in the context of wider international activities such as ESGF. Ongoing initiatives, such as Copernicus Climate Change Services and the European Open Science Cloud, also demand coordination of efforts and contribution. The Task Force also works putting in place a sustainable structure to support the data infrastructure when IS-ENES2 support will cease. The Data TF also

¹ DG Connect initiatives on HPC: ETP4HPC, EXDCI, Extreme Scale Demonstrators.

contributed to the updated ENES strategy for the infrastructure. Several of its members are involved in the international governance (see 1.1.2).

The HPC and Data TFs together collected information on HPC and storage requirements for CMIP6, and in doing so have demonstrated the scale of increases in resources needed over that needed for CMIP5.

The ENES Scientific Officer (ENES-SO), introduced in IS-ENES2, assists the ENES Chair and Board in coordinating ENES. The ENES-SO contributes to the governance tasks as well as to the community building and actively participates in ENES activities such as the development of the infrastructure strategy, the interactions with the scientific community, the dissemination. The ENES-SO plays also an important role in following, coordinating and reporting on activity of IS-ENES2 with other European projects and initiatives, such as the ENVRIPlus cluster of European environmental Infrastructures, the ESiWACE Centre of Excellence on HPC applications in weather and climate, or the ClimatEurope liaison activities. The ENES-SO also supports international governance initiatives (see section 1.1.2.).

D2.1 - Update of ENES governance rules

1.1.2. International governance

The IS-ENES projects have played a crucial role in the development of the international distributed database for international coordinated model results, the Earth System Grid Federation (ESGF). ESGF was initially completely led by US institutions but an international governance has been put in place in April 2015. IS-ENES2 members are fully acknowledged as a major contributor and considered a key influential player along with US partners (DOE, NASA, NOAA), at the Steering (one member) and Executive levels (three members) and lead now half of the sixteen working groups developing, implementing, and sustaining the infrastructure (see section 4). IS-ENES2 has interacted with the CORDEX Scientific Advisory team to define rules for CORDEX implementation within ESGF.

IS-ENES2 has also provided intellectual leadership and played a key role in the evolution of the metadata standards and in the development of the metadata services in collaboration with US institutions in the frame of the Earth System Documentation (ES-DOC), an international collaboration documenting climate models, simulation and conformance to MIP protocols.

CMIP is led by the WCRP Working Group on Coupled Models (WGCM). Following the increasing infrastructure dimension of CMIP activities, WGCM created an infrastructure panel in 2014, the WGCM Infrastructure Panel (WIP), to define the technical requirements defined by CMIP6. IS-ENES2 is engaged through participation of several of its partners and support by the ENES-SO. The WIP ensures the connection with modelling groups and defines recommendations then implemented in ESGF and ES-DOC. The different levels of the international governance for ESGF and ES-DOC are important and IS-ENES2 has been essential to ensure the European contribution to them.

1.2. Community building

To strengthen community building, IS-ENES1 and 2 have supported training for young generation of climate modelling researchers as well as a central information platform, the ENES portal.

1.2.1. European training school on climate modelling

Training schools for doctoral students and/or early career post-doctoral researchers are an important way to both educate young researchers in the complexity of Earth System Model software and to enhance collaboration within the next generation of climate scientists. Following the first IS-ENES1 school, IS-ENES2 has supported the second and third “European Earth System and Climate Modelling School” (E2SCMS). These schools took place in Barcelona in June 2014 and Helsinki in June 2016, respectively. The schools provided 58 young researchers with a series of lectures on the Earth system models, introductions to and hands-on experiences with running up-to-date ESMs, and experience analyzing results and performing intercomparisons. Students could get familiar with three different ESMs from UK, Germany and the EC-Earth Consortium, and provided very positive feedback.

1.2.2. The ENES Portal

The ENES portal (<https://portal.enes.org>) was one of the central outcomes of IS-ENES1; it acted as a communication platform and a central entry point to information for the European Earth System Modelling (ESM) community and provided access to the results achieved in the IS-ENES1 project.

During IS-ENES2 the portal has been further established as a central information gateway for the European ESM community and as unique access point to IS-ENES2 services. The portal's main structure is organized along the main aspects of Earth System Modelling: models & tools, data, and computing, and hosts an additional section on community activities and initiatives. A service section gathers pages in the model & tools and data sections and provides descriptions and direct access to the IS-ENES2 support services.

Following the advices of the internal service reviews and of the Mid-Term Review, improvements were applied to make the access to information and IS-ENES2 support services more effective and to enhance overall consistency. The portal will be continued after IS-ENES2, carrying its legacy in terms of provision of information and access to services for the European modelling community.

D2.7 – Final version of ENES portal

1.3. ENES infrastructure strategy

The ENES infrastructure strategy for the decade 2012-2022, conceived in the timeframe of IS-ENES1, addresses the main challenges for the community in terms of model development, needs for high-performance computers, data infrastructure, and human resources. Many recommendations of this foresight are implemented in IS-ENES2. At mid-term of IS-ENES2, it was crucial to revisit this document and to adapt it to the evolving needs of climate research, software and hardware in the international context. To complement the strategy document, a special focus on model evaluation infrastructure has been prepared.

1.3.1. Evaluation infrastructure strategy

Evaluation of multi-model ensemble results requires a dedicated infrastructure for access, processing, and combining climate model data with observations. Following a workshop organised in Utrecht in May 2014, document **D2.3** presents a strategy for a model evaluation infrastructure. This strategy, developed with international colleagues in the context of the preparation of CMIP6, outlines the infrastructure requirements to ensure a smooth evaluation process through improved and more routine Earth system model evaluation in CMIP (Eyring et al, 2016)².

The ESM community needs to perform much more efficiently many baseline aspects of model evaluation, to enable a systematic, open, and rapid performance assessment of CMIP models and avoid re-writing analysis routines for well-established analysis methods. A wide range of diagnostics and performance metrics should be processed in a systematic approach, as soon as the model output is published in ESGF. Observations (obs4MIPs) and re-analyses (ana4MIPs) for Model Intercomparison Projects, using the same data format and structure as CMIP, largely facilitate routine models evaluation. However, to be able to process the data automatically, the infrastructure needs to be extended with processing capabilities at the ESGF data nodes, where the evaluation tools can be run on a routine basis. The intention is to produce through ESGF a widely accepted quasi-operational evaluation framework for climate models that would routinely execute a series of standardized evaluation tasks and provide a quick and open identification of the strengths and weaknesses of the simulations. This will assist modelling groups in improving their models.

² Available in open access at Earth System Dynamics, <http://www.earth-syst-dynam.net/7/813/2016/>

*D2.3 – White paper on ESM evaluation needs and infrastructure***1.3.2. Mid-term update**

IS-ENES2 started implementing the recommendations of the ENES infrastructure strategy for 2012-2022. At mid-term (2016/17) it was recognized that the strategy might need updating in the light of progress. With this purpose, foresight discussions were held in a dedicated workshop (October 2016, MS2.4) and at the IS-ENES2 Final General Assembly (January 2017).

The HPC-TF also addressed the challenges faced by the climate modelling community in better exploiting current and future HPC technologies (**MS2.5**). It also described how new avenues needed to be considered regarding people, training and relation with vendors.

The mid-term update of the strategy (**D2.6**), as circulated in April 2017, analyses drivers and infrastructure components given the current challenges. It is consistent with previous recommendations on models, HPC, data, networks and people, but brings in new thinking based on the latest scientific and technological challenges. It complements them with guidance on scientific evaluation of models and on how to organizationally tackle the sustainability challenge, reflecting the approach of a more mature community (Table 1). This update will be important to guide further developments of the ENES infrastructure and advice the scientific community as well as national and European stakeholders.

Recommendations	
Models	Accelerate the preparation for exascale computing by exploiting next generation hardware <i>at scale</i> as early as possible, recognising that new algorithms, software infrastructures, and workflows will be necessary and will take substantial time and effort to develop.
High-performance computing	Work through national and European facilities to exploit a blend of high-performance computing facilities, recognizing the need to support both current and next generation science. Sustained access to world-class machines and next generation architectures will be needed to make a step-change in climate science.
Data	The community should invest more in research into data standards, workflow, data handling, high performance data management, and data analytics to meet the challenges of increasing data volumes and complexity. It should ensure that the systems, standards and workflows are in place, with sustainable funding and suitable mechanisms for establishing requirements, so to that data from climate simulations are easily available and well documented and quality assured, especially for downstream users.
Physical network	Work with national and international network providers to maximize the bandwidth between the major European climate data and compute facilities and ensure that documentation and guidance on tools and local network setup are provided for end-users and their local network administrators.
People	The community should be proactive about advertising the intellectual and technical challenges in climate science, both to individuals and to colleagues in other disciplines. Institutions should increase opportunities for training in climate science modelling and underlying technologies, at all levels from undergraduate to doctoral training courses and summer schools, as well as strengthen networking of software engineers.
Model evaluation	NEW: The community needs to put in place shared governance procedures for climate model evaluation software, which covers both the evaluation aims and the software structures. Data infrastructure should be expanded to include the computational resources needed for routine evaluation of new data products as they join the community ensembles.
Organisation	NEW: The European climate science community needs to utilise both national and European funding to develop sustainable community based cooperation that would put both production science and long-term future development activities on a firmer footing. In doing so, it needs to develop clear interfaces with other service providers (e.g. GEANT) and service consumers (e.g. the Copernicus climate services).

Table 1: Recommendations of the mid-term update of the ENES infrastructure strategy

This updated strategy will be of use at three levels: (i) to indicate European funding decisions how European investment in infrastructure to support Earth system simulation could deliver rewards to European society that build on, but do not replace, national investments; (ii) to help colleagues making national funding decisions understand how their decisions impact not only on their national priorities, but also on the

synergies possible at the European level; (iii) to help colleagues across our discipline (including those in relevant national and European institutions) understand: the relationship between scientific goals and the necessary infrastructure, our collective inter-dependencies on infrastructure both now and in the near (up to five year) future, and the relationship between the costs and risks of joint approaches and the potential added value.

MS2.5 – Update of ENES strategy for HPC

D2.6 - Update of the ENES infrastructure strategy

1.3.3. Planning for a long-term sustained European research infrastructure

The European research infrastructure has grown significantly in the last few years in size, services, user community, and international standing, thanks to the IS-ENES projects. As expressed in the new recommendation on “organization”, the ENES community considers it important to sustain its infrastructure. Indeed, sustainability is important to support climate research – support to WCRP experiments, sharing of software and expertise, further integration of the modelling community...- but also to provide the research infrastructure and expertise on which the Copernicus Climate Change Service can broker services on climate projections.

Within IS-ENES2, the first instrument toward sustainability considered was an ESFRI concept (CLiM-ERI, Earth’s CLiMate system Modelling: European Research Infrastructure), which would have encompassed execution and costs of the internationally-coordinated CMIP experiments. However, this attempt raised concerns around the appropriate balance between national and international investments and agreements. The scope of this initiative has thus to be redefined. The mean to pursue sustainability for the infrastructure is now more identifiable in a multi-lateral agreement focusing on the data infrastructure, still to be initiated. Although the ESFRI option was finally not pursued, the discussions led to the creation of CLiMERI-France.

2. Enhance the development of Earth System Models

Supported by IS-ENES2, the climate modelling community aims at accelerating development of Earth system models, fostering common solutions and enhancing the sharing of expertise. Networking and joint research activities on future model developments are necessary to improve the quality of models and tools, ensure efficient use of future computing architectures and increase the cohesion of the European ESM community. Besides stimulating common software developments, IS-ENES2 investigated shared governance models (as for code coupler OASIS), application methodologies (coupling technology benchmarking suites) and provided services on models (those involved in CMIP5, in particular the ocean modelling platform NEMO) and tools (e.g. post-processing software CDO, and OASIS).

2.1. Services around model environment tools and components

IS-ENES2 has sustained the services on common software tools that were started in IS-ENES1: the OASIS coupler, the post-processing tool CDO, and the European ocean modelling platform NEMO. Within IS-ENES2, new services on European models used in CMIP5 have been launched including both level 1 services (provision of contact points for information on each code) and level 2 services (provision on support for access and usage of HADGEM and EC-Earth).

Two reviews of IS-ENES services (**MS7.4**, **MS7.6**) were conducted with external reviewers (**MS7.3**) at mid-term and during the last year of the project. They assessed both service levels, checking reliability of the information, accessibility, efficiency, and technical consistency. They also issued recommendations on these aspects (and validated them). An important improvement is the overall visibility of IS-ENES2 among potential European users.

MS7.3 – Set-up of the review committee

MS7.4 – First review report on the ENES Earth System Model Resources

MS7.6 – Second review report on the ENES Earth System Model resources

2.1.1. Service on OASIS coupler

The OASIS coupler is a software library, developed at CERFACS, which allows synchronized exchanges of coupling information between numerical codes representing different components of the Earth system. It is widely used to couple atmosphere, ocean and land models but also sometimes other components such as atmospheric chemistry or ocean wave components. OASIS is used by about 70 climate modelling groups around the world (40 in Europe).

IS-ENES2 has continued support activities initiated under IS-ENES1, which enabled the users to freely download state-of-the-art versions of the software as well as retrieve relevant information about installation and use. Via the ENES portal, users can access a comprehensive web site (<https://portal.enes.org/oasis>) including up-to-date source download and documentation, a tutorial, technical information with hints for best practices, FAQ and forum, news and events announcements, and dissemination information. Direct user support is also offered: an expert, part of the development team, is available to answer specific questions and provide guidance on how to use the software on specific platforms and in particular configurations. The IS-ENES2 service activity for OASIS is summarized in Table 2.

	1st Period (18 months)	2 nd Period (18 months) ³	3 rd Period (12 months)
Version	OASIS3-MCT	OASIS3-MCT3.0	OASIS3-MCT3.0
Change sets	558	549	231
Downloads	230	>174	78
User support	416	>231	371
Tickets updates (Redmine)⁴	219	475	41

Table 2: Service activity on OASIS

OASIS governance

IS-ENES2 investigated the interest in and possible models of community governance for OASIS. Indeed, the coupling tool OASIS, developed at CERFACS with support from different EU projects, has reached a level of maturity where a 3-tier governance structure is seen to be beneficial (**D4.3**). The agreed structure includes a Developer Team for the day-to-day development and quality control, a User Group to feed back on user requirements and experiences with the software, and an Advisory Board to consider more strategic issues such as long terms requirements and external drivers.

The OASIS development team considers such external governance beneficial for software quality, leading to a stronger offer. This belief is common in a number of successful open source projects. It is noteworthy that OASIS became a popular tool well before any community governance structure was put in place, although community engagement through various EU funded projects has indeed proven useful.

D4.3 - Coupler governance model document

2.1.2. Service on CDO Climate Data Operators

CDO is a collection of command-line operators used to manipulate and analyze climate and numerical weather prediction data as well as other gridded data. The operators (over 600) include simple statistical and arithmetic functions, data selection and subsampling tools, and spatial interpolation and support different standard data formats. CDO is developed at MPG. Thanks to user feedback, the CDO developer is able to successfully identify the (evolving) user requirements and deliver solutions. There is however no formal user engagement nor visible community governance.

³ Due to the computation of new, more meaningful KPI with semi-yearly frequency started during the second period (from month 9 onwards), some values relative to this period are given as a subset of the total as they only cover the last year of the 18 month period.

⁴ <https://inle.cerfacs.fr/projects/oasis3-mct>

CDO has been included in IS-ENES services since the first phase of the project and can be accessed via the ENES Portal (linking to the CDO home page <https://code.zmaw.de/projects/cdo/wiki>). Users can download CDO, up-to-date documentation, and access the FAQ. Services include CDO help desk and webserver set up during IS-ENES1. CDO developers at MPG answer user queries, collect errors, and maintain discussion forums where users have the opportunity to interact on a single platform. All this leads to a fast resolution of problems and efficient exchanges between users and developers on new features. The community of users (~about 1000 users) encompasses members of academia and industry. Source code and binary packages of CDO have been downloaded several thousand times from the download area. Contributions to the different forums on the website amount so far to over 2700. The three reporting periods exhibit a large and growing user community of CDO as can be seen by the growing relative number of downloads (Table 3).

	1st Period (18 months)	2 nd Period (18 months) ⁵	3 rd Period (12 months)
Version	CDO 1.5.9 to 1.6.4	CDO 1.6.5 to 1.7.1	CDO 1.7.1 to 1.8.0
Downloads	> 2000	> 12434	9400
Support requests	> 2700	> 874	947
Tickets opened/closed⁶	-	72/64	76/84

Table 3: Service activity for CDO

2.1.3. Service on NEMO ocean platform

The NEMO (Nucleus for European Modelling of the Ocean) modelling platform is used in five of the seven European ESMs (CNRM-CERFACS, CMCC, EC-EARTH, IPSL, Met-Office Hadley Centre) participating to CMIP6. All the sustainable development of the NEMO platform is organised and shared within the NEMO System Team, i.e. within the experts coming from all of the institutions of the NEMO Consortium: CMCC (Italy), CNRS (France), INGV (Italy), MetOffice (UK), Mercator-Ocean (France), NOC (UK), with contribution from BSC (Spain). NEMO support services aim at facilitating community sharing of experience and expertise. These services are accessible through the ENES portal and organized around NEMO web site <http://www.nemo-ocean.eu/> and the associated forge <https://forge.ipsl.jussieu.fr/nemo/wiki/Users>. The organization of NEMO Collaborative Development Environment was completely reorganised by the NEMO System Team during IS-ENES2 to ensure clearer access levels (first comers, users, developers), more stability and security in a user-friendly environment. Through these improved tools and services, complete and regularly updated information is available so that users can access: the NEMO code reference distributed under free license and its history (using Subversion server), reference manuals, user guides, publications, forums, meetings announcements and news and a ticketing system for developments.

In the framework of IS-ENES2, the NEMO community has developed configurations of NEMO at 1° and ¼° spatial resolutions, incorporating the new sea-ice component LIM3, both to be used for CMIP6 by several climate models (**MS7.7**). First sensitivity experiments were produced and presented to the community (Figure 3). For this, the NEMO System Team has: (i) set up (and administered) the forge project (shaconemo, SHared COnfigurations for NEMO⁷), a community-run service with 80 registered users to share information, expertise and progress on the NEMO configurations for CMIP6; (ii) organized 3 workshops (2016-2017) allowing different groups to confront their experiences and search for optimal solutions; (iii) provided guidance on choices of inputs and tuning of the ESMs coupled system. With these tools, a project can download the whole NEMO platform from the NEMO web site and access configuration-specific items from shaconemo. The NEMO 3_6_STABLE release has been finalized for CMIP6 by the NEMO System team (beyond IS-ENES2 support). Overall, NEMO is currently used in 240 projects in 27 countries (14 in Europe, 13 elsewhere) with more than 1400 registered users.

⁵ Due to the computation of new, more meaningful KPI with semi-yearly frequency started during the second period (from month 9 onwards), all values relative to this period are given as a subset of the total as they only cover the last year of the 18 month period.

⁶ This information is only available for part of the 2nd period and for the whole 3rd period.

⁷ <https://forge.ipsl.jussieu.fr/shaconemo/wiki>

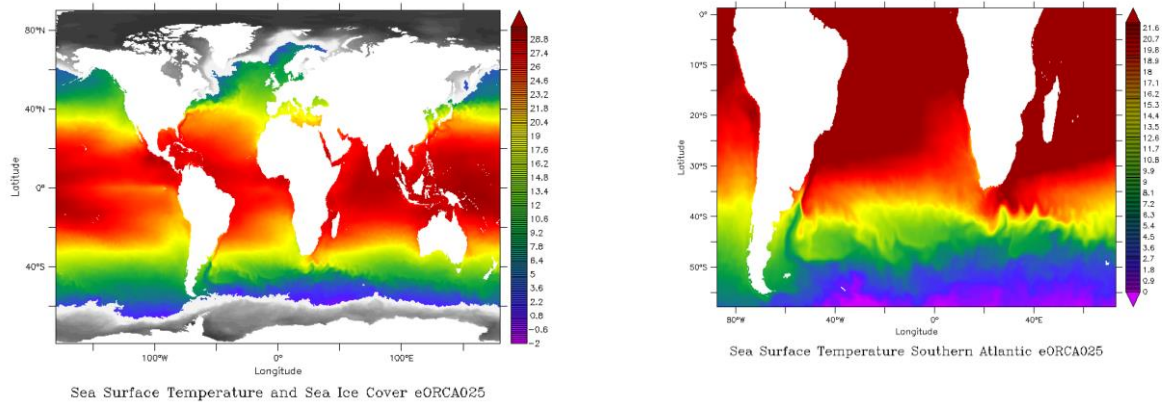


Figure 3: Sea surface temperature and sea ice cover from eORCA025 ($1/4^\circ$) with a zoom on the Southern ocean (courtesy of CNRS-IPSL)

	1st Period (18 months) ⁸	2 nd Period (18 months) ⁹	3 rd Period (12 months)
Changes of source code	> 744	> 1500	2000
Opened tickets/bugs fixed	233/140	322/326	65/45
Edits of the wiki pages	> 482	> 300	150
User support (mail)	> 300	> 140	350

Table 4: Service activity on NEMO

MS7.5 - Documentation of the reference and sensitivity experiments based on ORCA025 configuration of NEMO

MS7.7 – Documentation of the reference and sensitivity experiments based on the activation of LIM3 in ORCA025

2.1.4. Service on climate models

The 7 European ESMs groups have reached different levels regarding the provision of services around their ESM. During IS-ENES1, main focus has been put on model documentation of the European ESMs participating to CMIP5. Metadata using the Common Information Model (CIM) was gathered. Within IS-ENES2, this documentation was maintained and extended. Furthermore, contact persons for each ESM have been identified to access local expertise beyond the available documentation. All this led to the formulation of two levels of service. All European global Earth system modelling groups offer level 1, i.e. a contact address and maintain the model description using the CIM metadata format on the ENES portal (**MS7.1**). The models concerned are: (i) IPSL-ESM (CNRS-IPSL); (ii) NorESM (UiB, met.no), (iii) C-ESM (CMCC); (iv) MPI-ESM1 (MPG); (v) CNRM-CM5 (Météo France-CNRM); (vi) Unified Model ESM (Met Office Hadley Centre); (vii) EC-Earth (EC-Earth consortium, KNMI, SMHI). Met Office Hadley Centre and the EC-Earth Consortium also offer level 2 services (**MS7.2**) for Unified Model and EC-Earth, respectively, i.e. in-person expertise to help in setting up and running the models. As for all IS-ENES2 support services, a service catalogue is available in the ENES portal.

MS7.1 - Contact person identified and contact details available for each European ESM

MS7.2 - Level 2 service on ESM available

2.2. Towards next generation climate models

The climate modelling community needs to prepare for the challenges related to model performance on the upcoming exascale architectures and to better connect those working on climate models by establishing some code convergence and shared understanding of divergent codes. In order to pursue these objectives, best practices in model development need to be defined, coordinated and issued to support the execution of these codes on the massively parallel heterogeneous machines expected in the future. IS-ENES2 tackled different

⁸ Due to the computation of new, more meaningful KPI with semi-yearly frequency started during the second period (from month 9 onwards), all values relative to this period are given as a subset of the total as they only cover the last year of the 18 month period.

areas along the path toward the next generation of climate models: the work toward common radiation tools, the extension of benchmarking activities to coupling technologies, the convergence of hardware-aware libraries and model scientific libraries, and the exploration of computational cores and new parallel approaches for NEMO and ICON.

2.2.1. Towards common radiation tools

The representation of radiation in atmospheric models is handled by parameterisations that use the spectroscopic properties of gases to carry out clear-sky radiative transfer alongside with scattering processes (resulting from the presence of particles and/or clouds). In general the clear-sky radiative transfer is a well-understood problem – whereas spectroscopic behavior of the gases and scattering processes are both under active research.

With the aim of building a community around a common European approach to this problem, IS-ENES2 partners worked together to develop a common radiative transfer library to be used in ESMs, and on observation-system simulators. A prototype has been developed by refactoring the existing and widely adopted Rapid Radiative Transfer Model (RRTM), developed and maintained by the Atmospheric and Environmental Research⁹ (US) since the 1980's, allowing for relatively smooth migration through its backward compatibility. The new code compiles independently from any particular atmospheric model and can therefore be built as a separate, lightweight radiation-only tool. This radiative transfer library can be used directly through C/C++ bindings and a Python wrapper; the latter allows an easy, interactive interface. Additionally, a new ocean surface albedo code has been developed to be more consistent with new observation data and with new ESMs computing marine biota and detrital organic materials.

As compared to other existing libraries, this set of radiation tools developed within IS-ENES2 exhibits increased portability, modularity, and maintainability. The prototype version has successfully been adopted by different research groups, each with different technical constraints and scientific requirements, for which the previously available tools were inadequate.

The comparison of models with models is important to evaluate the success of this approach. The Cloud Feedback Model Intercomparison Project (CFMIP) Observation Simulator Package (COSP)¹⁰ code has therefore also been improved and optimized. A workshop allowed furthermore the emergence of a developers and users community.

D3.3 Report on common radiation tools

2.2.2. Benchmarking coupling technologies

Toward the establishment of a full standard benchmarking suite for coupling technologies, IS-ENES2 partners propose a first benchmark version (available on the ENES portal) implementing one simple test case with 5 different coupling technologies.

The work started in 2013 at the 2nd Workshop on Coupling Technologies (Boulder, US) (Dunlap et al., 2014) with the description of the possible characteristics of coupled Earth System Models. These were then classified (by IS-ENES2 and the Earth System Bridge US project members) in a series of mind-maps¹¹. Priorities in coupling characteristics to benchmark were defined and it was established that the benchmark suite would include a number of pre-coded, stand-alone components, running on different grids, to be assembled via different coupling technologies (**MS10.1**, **MS10.4**) (Valcke et al., 2016). The components would contain neither physics nor dynamics, but their coupling would be representative of real coupled models in terms of coupling characteristics and load. The four stand-alone components currently available run on: 1) a self-generated, regular, latitude-longitude grid; 2) an irregular, stretched, and rotated latitude-longitude mesh (as NEMO ORCA); 3) a quasi-uniform icosahedral mesh (as DYNAMICO), and 4) a quasi-uniform cubed sphere. The component using grid #1 was used to assemble coupled test cases using five different coupling technologies from Europe (OASIS3-MCT, OpenPALM, YAC) and from US (ESMF,

⁹ <http://rtweb.aer.com/>

¹⁰ CFMIP-OBS web site: <http://climserv.ipsl.polytechnique.fr/cfmip-obs/>

¹¹ <https://earthsystemcog.org/projects/es-fdl/mindmaps>

MCT). Test cases were run with different numbers of cores per component and with different grid sizes and the impact on the performance of the coupling technologies was evaluated on different platforms: Bullx (CINES, France), Cray XC40 (Met Office, UK), and the Broadwell partition of Marconi (CINECA, Italy) (**D10.3**)

The performance results, representative only of each test case implemented (e.g. Figure 4), demonstrate more generally that it is indeed possible to build a standard coding environment to objectively compare different coupling technologies. For the very specific cases run, further work is needed to investigate their significance and robustness and to understand why, in few cases, some technologies show a better behavior than others.

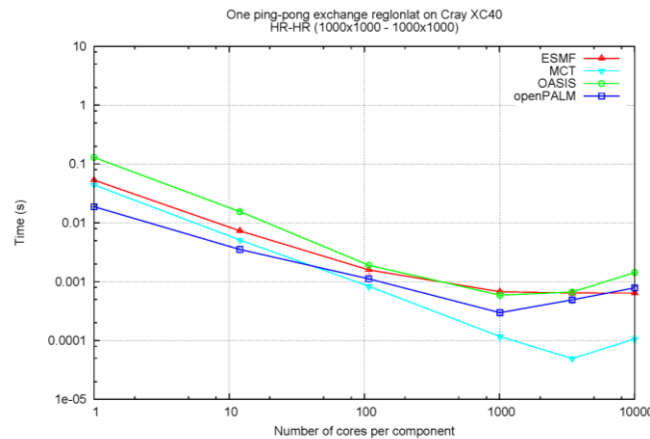


Figure 4: Benchmarking of four coupling technologies on the Cray XC40: time for one back-and-forth exchange between two components running on a regular latitude-longitude grid, representative of today high-resolution atmosphere and ocean models (from D10.3)

This work forms a cornerstone of a wider, international community effort to standardize coupling technology description and evaluation. It shows the maturity the climate modelling community has reached interacting and working together. Possible extensions of this suite are almost infinite. At this point, anyone is welcome to use or extend the existing benchmark suite and to report back on progress and results.

D10.3 – Report on benchmark suite for evaluation of coupling strategies

2.2.3. Code/software convergence Chasm workshop

Weather and climate models are complex pieces of software which include many individual components, each of which is evolving under the pressure to exploit advances in computing to enhance some combination of a range of possible improvements (higher spatio-temporal resolution, increased fidelity in terms of resolved processes, more quantification of uncertainty etc). However, after many years of a relatively stable computing environment with little choice in processing architecture or programming paradigm (basically X86 chips using MPI for parallelism), the existing menu of processor choices includes significant diversity, and more is on the horizon. This computational diversity, coupled with ever increasing software complexity, leads to the very real possibility that weather and climate modelling will arrive at a chasm which will separate scientific aspiration from our ability to develop and/or rapidly adapt codes to the available hardware.

The Reading workshop in October 2016 “Crossing the Chasm: towards common infrastructure software for Earth System Model development” has been an opportunity to review the hardware and software trends which are leading us towards this chasm, and to investigate current progress in addressing some of the tools which we may be able to use to bridge the chasm. An important requirement is to have quality model codes with satisfactory performance and portability, while simultaneously supporting productive scientific evolution. It is likely that the existing method of incremental model improvements employing small steps to adjust to the changing hardware environment will be inadequate for crossing the chasm between aspiration and hardware at a satisfactory pace, in part because institutions cannot have all the relevant expertise in house. Instead, the workshop report (D3.2), which will be submitted as a paper, outlines a methodology based on large community efforts in engineering and standardisation, one which will depend on identifying a

taxonomy of key activities — perhaps based on existing efforts to develop domain specific languages, identify weather and climate key components limiting performance (“dwarfs” or “mini-apps”) and develop community libraries — and then collaboratively building up those key components. Such a collaborative approach will depend on institutions, projects and individuals adopting new interdependencies and ways of working.

D3.2 Report on strategies for developing convergent model codes

2.2.4. Computational cores and new parallel approaches for NEMO and ICON

The next generation computer architectures will enable the resolution and the complexity of climate models to be increased. However, for this to be fully realized, the scalability of current models has to be improved, e.g. through models more able to exploit parallelisms. Indeed, when the number of parallel processes increases up to hundreds of thousands of cores, the main bottlenecks to scalability are the communications overhead and the memory access. More generally, data movement at all levels will increasingly become the main limiting factor.

During IS-ENES2, NEMO and ICON models have been analyzed in order to identify the main bottlenecks to their scalability. The current version of NEMO has not been designed for the high levels of parallelism even on today’s systems. ICON is a next generation Earth system model designed to simulate multiple scales of the atmosphere processes, enabling both climate simulations and numerical weather predictions. It is a joint development of MPI-M and the German Weather Service DWD. The code can be run at very high resolution and is highly scalable. The main computational intensive parts of ICON and NEMO have been identified. A methodology and several metrics have been used to characterize them. Both codes are parallelized using message passing interface (MPI) using horizontal domain decomposition to divide the computational domain across MPI-processes. Moreover, for ICON, thread level parallelism is introduced by employing OpenMP directives to parallelize loops.

Experiments with ICON at very high resolution (**D3.4**) show that the major limiting factor to good scaling is due to the memory issues. This means that with increasing number of MPI-processes the memory consumption per process is increasing as well, due to the replication of some global data structure on all the parallel processes, whereas we need a decrease in memory usage per process with growing number of processes. This is especially important on future architectures with less memory per core as well as less memory bandwidth, typically. New parallel approaches are needed to overcome this problem (Figure 5).

NEMO scalability is highly affected by the communication overhead and some techniques to reduce it are suggested in D3.4. Figure 5 shows how the scalability of routines which perform MPI communications leads to decompose the overall grid with subdomain blocks of 40x40 grid points at minimum, limiting the number of cores used. Routines without MPI exchanges scale up to subdomain blocks of 10x10. A target of 10x10 would enhance the number of cores by 16, improving scalability.

From this and other available research, a roadmap for the NEMO and ICON models re-design has been developed, aiming at increasing the model performance by exploiting the characteristics of the new generation architectures (**D3.4**). The roadmap defines a standard methodology to analyze model performance and both evolutionary and revolutionary approaches to improve model performance. Evolutionary approaches are based on current versions of codes and improve communications for NEMO and memory scaling for ICON. The revolutionary approaches would require a complete rewriting of codes. For NEMO, it is likely to involve a new dynamical core and new solvers that parallelize in time, or new approaches applying a separation of concern (see 2.2.5). ICON is already based on a new dynamical core but decoupling radiation and the dynamics would boost performance. This requires a re-design of the current parallelization concept of ICON, from being purely domain decomposition to including task parallelism.

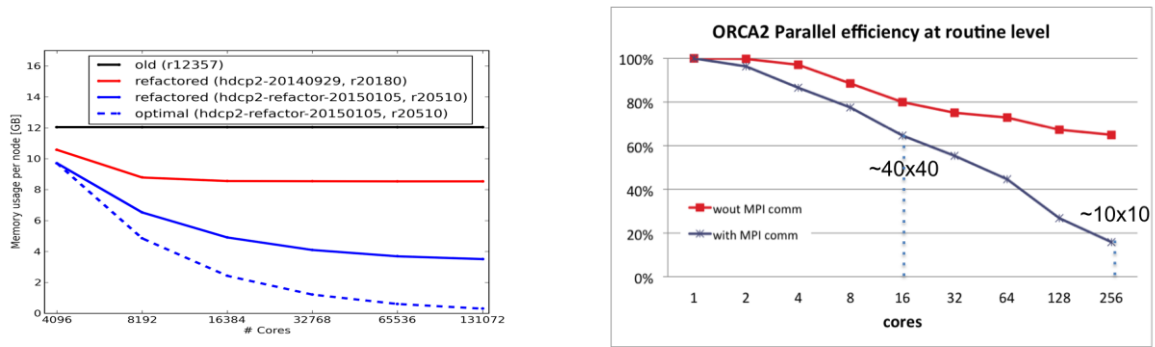


Figure 5: (Left): Optimization of the memory usage per compute node for a local area model with a resolution of 240m on the BlueGene/Q system JUQUEEN (courtesy of DKRZ); (Right): ORCA2 (200 km) parallel efficiency of routines performing (blue) or not (red) MPI communications¹².

D3.4 Report on NEMO and ICON models redesign (added following mid-term review)

2.2.5. Testing revolutionary and evolutionary approaches for NEMO

As mentioned in section 2.2.3, it is likely that existing evolutionary approaches may not be sufficient to adapt to future computer technologies over a longer period. Benefitting from a reallocation of budget within the project, it was proposed by STFC and CMCC to test a new approach promoted in the UK MetOffice's LFRic¹³ project to NEMO. This approach is based on the PSyKAI (Parallel System, Kernel, Algorithm) separation of concerns approach. In this approach, a program is re-structured into three layers: the Algorithm, Parallel System (PSy) and Kernel layers (Figure 6). The Algorithm and Kernel layers are the responsibility of the natural scientist while all code related to parallelism is contained within the PSy layer. Thus the scientists do not need to concern themselves with optimisation issues, such as loop fusion, redundant computation in halos, etc. and parallelism issues, such as placement of directives and halo swap calls. This separation also improves the readability and modularity of code and thus greatly aids with maintenance. This may be one of the approaches that could be used to cross the chasm (2.2.3).

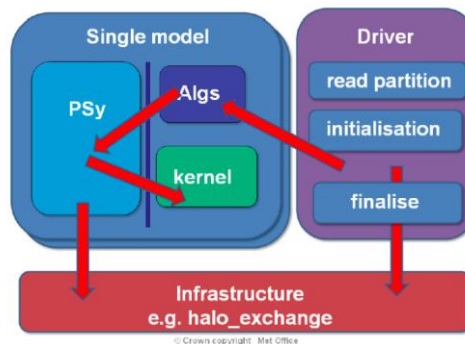


Figure 6: Separation of concerns between the Parallel Systems (the PSy layer) and Algorithms and Kernels¹⁴

A previous UK NERC funded project called GOcean (<https://puma.nerc.ac.uk/trac/GOcean>) demonstrated that the PSyKAI approach could work for a 2D NEMO kernel. IS-ENES2 has extended this work to demonstrate the feasibility of this approach for a 3D NEMO (advection) kernel. A PSyKAI version of the advection kernel from NEMO has been developed which has been shown to be able to run on multi- or many

¹² In the working document available on the project portal <https://portal.enes.org/ISENES2/documents/na2-working-documents/the-nemo-oceanic-model-performance-analysis/view>

¹³ See <https://puma.nerc.ac.uk/trac/GungHo/wiki/LFRic>

¹⁴ See Special report on IS-ENES2 web site “Application of the PsyKAI approach to the NEMO ocean model” Fort R. et al., 2017

core architectures (using OpenMP or GPU-based architectures with OpenACC), with no change to the scientific source code.

However, this initial implementation of the PSyKAI approach means that the code is unrecognisable to developers. Unlike LFRic, NEMO is an existing code and a team with heterogeneous expertise (not only computational scientists) is in charge of its development. Such radical changes to the code base could have a significant impact on both the developers and users. For this reason the NEMO Systems Team are reluctant to take such a major step.

In recognition of this issue, two, less invasive, approaches have been developed and their feasibility examined within the ISENES2 project. They require very little modification of the existing code structure and can be considered to be evolutionary, rather than revolutionary, approaches. After discussion with the NEMO System team it is clear that the natural scientists have a different view of code readability from computational scientists: they consider the introduction of the algorithm layer (in PSyKAI) an overhead for code development, rather than a benefit. Fortunately, NEMO has quite strong code conventions, which can be utilized to extract the necessary information to generate new, transformed source code e.g. using OpenACC directives from unchanged NEMO code. This solution is an attractive proposition since previous work by NVIDIA has shown a two-times speed-up for NEMO 3.4 when running on GPUs. These approaches are scheduled to be presented to the NEMO developers committee to discuss their possible adoption.

2.3. Sharing best practices for model environments

Scientists using Earth System Models do not only need access to models, but also to a rich environment of supporting tools. This software environment includes, among others, configuration management, workflow tools and meta-data tools. Networking occasions are essential to share experiences and facilitate the understanding and exploitation of such tools.

2.3.1. Workflow tools

In the context of climate modelling, a typical workflow is a suite of tasks run in a specific order to complete a climate simulation to an experiment design (Figure 7). Starting with the preparation of model input, the workflow has tasks including completing the integration and finishing with the production of final results. Workflow tools provide a mechanism to define, control and manage these tasks. Like any configurable thing, the definitions are subject to configuration management – see the next section. Typically, the tasks in a suite are submitted to a batch scheduler in a HPC environment and the workflow tool will manage the order in which the tasks are run based on dependencies and it will deal with exceptions such as task failure all based on rules provided by the suite owner. Task parallelism can be exploited within the constraints of tasks dependencies. Workflow, as a discipline, is a rapidly developing area within the climate modelling community as a result of (i) the increased complexity in experimental design, (ii) the use of climate models in an operational context such as seasonal prediction system and (iii) the desire to automate a larger part of the workflow which had previously been done by ad-hoc and manually controlled sequence of scripts. A number of more flexible solutions have and will continue to be evaluated within the climate modelling community. They include generalized workflow tools established within the NWP community, for example ecFlow from ECMWF or Cylc initially developed by NIWA (New Zealand) and now in collaboration with the Met Office. Autosubmit, from BSC, is a tool that was established directly within the climate community.

A first IS-ENES workshop (June 2014, **D4.2**) served to share experiences and best practices regarding existing and new workflow methods and tools. Discussions involved the varied requirements including the role that workflow tools can play in the capture of meta-data to describe the data and to describe the provenance of experiment producing the data (see also 2.3.3). The opportunities and challenges around adoption of tools was also a topic of discussion. Many groups expressed their interest in Cylc with some institutions using it, evaluating it, or planning to evaluate it, and the workshop concluded that it would be worth seeking opportunities for coordinated support and maintenance. Further, this workshop helped identify development priorities for Cylc within WP9/JRA1.

At the second workshop (**MS4.7**), also supported by the Centre of Excellence in Simulation of Weather and Climate in Europe (ESiWACE¹⁵), the community discussed the available post processing solutions in use and their integration into workflows, with special focus on requirements of CMIP6. The workshop also offered Tutorials on Autosubmit (see section 3.3.1) and Cylc. Both the training exercises and the presentation of real-life successes with Cylc and Autosubmit will encourage interest in shared software solutions in the wider community.

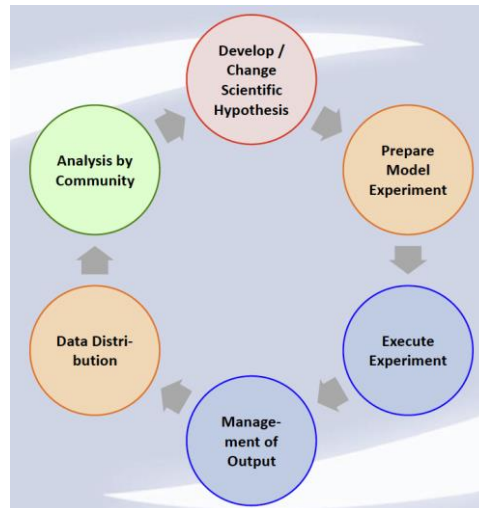


Figure 7: typical workflow for model execution, post-processing and analyses. From MS4.7

D4.2 – Workflow solutions initial workshop report

MS4.7 – Final workshop on workflow solutions

2.3.2. Configuration management

Where workflow tools support the running of experiments through a suite of tasks, the configuration management tools ensure that both the tasks and the suites have proper change control and that we have provenance of both suites and its constituent tasks. A key benefit of configuration management systems in climate science is to ensure scientific repeatability. The use of formalized workflow, rather than manual or ad-hoc tasks is a necessary but not a sufficient condition for good change control. Different configuration management systems are in use at various ESM sites. Two workshops, one physically in UK in 2013 and one virtually in 2016, shared experiences, in particular with one of the most prominent code, FCM (Flexible Configuration Management) developed at the Met Office.

Based on this first workshop, **D4.1** recommends that the scope of existing configuration management tools, which have historically been focused on the model codes, should be extended to ensure that the complete workflow is under configuration management and hence repeatable and traceable through change. Configuration management is about process and governance as well as tools and the processes should also be designed to improve the quality of the code base through change by tracking, testing and code review. From the experience of MetOffice and CNRS-IPSL, institutions that do not have such tools are encouraged to evaluate FCM, whereas for partners having their own effective toolset, there was no compelling reason to change.

Based on advice from the mid-term review, the second virtual workshop resulted in the publication of a “*Configuration management best practice guide*” for climate modelling (**MS4.5**), with contributions from more than 40 individuals from more than 30 institutions, with particular emphasis on technical experts working close to climate scientists. The guide defines configuration management and explains the need to control not only model codes but also the full workflow. Then, it analyses configuration management from the point of view of model developers, system owners, experiment designers, and data consumers and issues

¹⁵ <https://www.esiwace.eu/>

tailored recommendations for each of these roles. The guide will help groups, especially those with less experience, to implement robust configuration management and will inform people about the range of tools and methodologies used in the community. The resulting improvements in configuration management will lead to even more robust science and will better support scientists working on increasingly complex problems.

D4.1 – Configuration management initial workshop

MS4.5 - Configuration Management Best Practice Guide for Climate Science

2.3.3. Meta-data generation during experiments

Metadata capture tools are another example of environment software tool required to ensure that experiments, simulations, and data are properly described and governed. The CMIP5 experiences showed that the overhead of capturing complex metadata can be high, acknowledging the need to build metadata capture into the heart of the experiment process and to drive data provision exercises. This needs to be supported by both software tools and processes (see also workflow and configuration management above).

Current European and global activities address developments in metadata capture of both catalogue metadata and experiment metadata such as model descriptions and user annotations. Metadata capture was discussed at a first workshop (Hamburg, February 2014) and at the “*Joint IS-ENES2 Workshop on Workflows and Metadata Generation*” (Lisbon, September 2016, **D4.4**). The main outcome was that: (i) despite the improvements made between CMIP5 and CMIP6, the community would benefit from further cooperation and investment in more robust standardization of metadata content, structure, formats, and interfaces; (ii) for work with metadata and homogenization of metadata, such standardization efforts should be included with more care into projects at the proposal stage; (iii) for metadata references to external documents like definitions or international standards, the use of persistent identifies (PIDs) should be preferred; (iv) partners should agree on open legal standards at the beginning of a project.

D4.4 – Meta-data capture final workshop

MS4.6 - Final workshop on meta-data generation during experiments

2.3.4. Governance on common software

Software governance is the process of decision-making and quality control for software development. More specifically, community governance is the process by which a wider community, i.e. users of the software outside the developing institutions, can influence decision-making and quality control. Within the context of IS-ENES2, the aims of community governance are to increase the likelihood, sustainability, and efficiency of software sharing. An analysis of examples of community governance concludes that projects benefit from this outside influence. However, the existence of community governance is not always a major factor in the choice of software developed at another site.

The governance structure proposed for OASIS (see 2.2.1) is considered a useful template for mature community software. It encompasses: a Developer Team covering decision making for day-to-day development, a User Group covering user requirements, priorities and experiences, and an Advisory Board covering strategic direction. Note that it is not always possible for the community to mandate a governance structure, especially when there is no sustained funding. **D2.5** also gives advice on both aspects to consider and priorities when setting up community governance structures. More specifically, the issues to be considered for governing bodies include (i) chairmanship, (ii) type of meetings to be preferred, (iii) community consultation and feedback, and (iv) license conditions.

Discussions started at the 4th ENES HPC workshop¹⁶ and continued at the “Crossing the Chasm”¹⁷ workshop, also identified the definition of standards as a new approach that could more effectively lead to sharing codes under community governance. Such standards would be shared across codes out of which

¹⁶ <https://is.enes.org/events/4th-enes-hpc-workshop>

¹⁷ <https://is.enes.org/events/save-the-date-crossing-the-chasm-towards-common-software-infrastructure-for-earth-system-model-development>

dominant codes could emerge supported by the ease of movement between solutions working to a common standard (e.g. for both the interface and the functionality of a calendar library for climate models).

D2.5 - ENES governance report on software sharing

3. Foster high-end simulations enabling to better understand and predict future climate change

High-performance computing is a very critical component of the climate modelling infrastructure. If the provision of computing access is beyond the scope of IS-ENES2, the project has worked in several directions to improve use of HPC in Europe: by tracking technology of future exascale computing architectures (WP3/NA2), by preparing multi-model high resolution common experiments (WP9/JRA1), by stimulating collaboration with ICT companies (WP6/NA5) and by developing coupled benchmarks more adapted to evaluate machine performance for climate models (WP10/JRA2).

Activities within IS-ENES2 identified needs and gaps in the preparation for future exascale computing that led to establishment of ESIWACE in support to Weather and Climate HPC applications.

3.1. Tracking future computer architecture technologies

The ENES infrastructure strategy 2012-2022 emphasized global 1 km climate models as a major challenge for climate modelling. This will require not only to prepare the physics of models but also to be able to use future exascale computers (10^{18} Flops). However, a re-write of climate models needs several years of development. This, in turn, requires us to track technology of future computer hardware to make sure to be able to exploit them.

During IS-ENES2, the third ENES HPC workshop has been organized in Hamburg in March 2014. It was an opportunity to have an overview of technology plans as well as most up-to-date developments of models worldwide. **D3.1** reports on main outcomes from the workshop.

Exascale systems will require codes to scale to millions of cores. When the number of parallel processes increases beyond several thousands of cores, the main bottlenecks to scalability are the communication and memory access overhead and the workload imbalance among the parallel processes. Moreover, new generation HPC systems integrate different types of processors, and are often equipped with co-processors and / or accelerators (e.g. GPGPUs), requiring new programming approaches.

In order to significantly increase their scalability, models will need to (i) integrate new parallel sub-models and algorithms into legacy codes, (ii) tackle the I/O bottleneck for scalability, (iii) take into consideration a co-design approach for the hardware and software development and (iv) adopt new coupling strategies. Other disruptive approaches might be taken. More fine-grained task parallelism, and parallelisation models in time, if proved to be feasible, may increase model scalability. Co-design approaches, joining hardware and software development altogether, may be unique opportunities to optimise use of future hardware. In any case, computer scientists, applied mathematicians, and application scientists will need to work closely together to produce a computational science discovery environment able to exploit the new mathematical and algorithmic approaches.

D3.1 Report on the technology tracking

3.2. Improve model performance for HPC

3.2.1. Computational performance standards

The computational performance of climate models on high performance computers is a crucial parameter since it determines how fast the model can be run as well as how many simulations can be performed for a given amount of resources. Discussions on how to best measure the performance of climate models were initiated at the first ENES HPC workshop supported by IS-ENES1 in 2011, further elaborated at next workshops. Our US colleague, Dr Balaji presented a proposal for a standard set of metrics at the third ENES HPC Workshop in Hamburg in 2014.

The set of metrics measures real computational performance of Earth System Models. The main basic metric, which is becoming a standard, is Simulated Years Per Day (SYPD), which informs on how many years can be simulated in 24-hr of computing. However, more information is required to really understand and compare model performance. Indeed, comparing performance can inform on strengths, weaknesses and opportunities in models and platforms. The metrics include information of the models' resolution and complexity, on the platform used, on the computational cost, on coupling, memory and I/O costs. It is summarized in Table 5 and described in detail in Balaji et al. (2017). It is proposed to use this new set of metrics to launch a Computational Performance Model Intercomparison Project (CPMIP). Data will be collected within CMIP6 model and experiment documentation and will compare, for the first time, the performance of CMIP6 models.

Model and Platform	
Resolution	Number of grid points (latitude x longitude x vertical)
Complexity	Number of prognostic variables per component
Platform	Description of the computational hardware
Computational cost	
SYPD	Simulated years per day
ASYPD	Actual SYPD obtained from a long running simulation
CHSY	Core hours per simulated day
Parallelization	Total number of cores allocated for the run
JPSY	Energy cost of a simulation (Joules per simulated year)
Coupling/memory/I/O	
Coupling cost	Overhead caused by coupling
Memory bloat	Ratio of actual memory size to the ideal memory size (size of complete model)
Data output cost	Cost of performing I/O
Data intensity	Measure of data produced per compute hour

Table 5: Computational Performance Model Intercomparison Project (CPMIP) metrics (Balaji et al. (2017))

As reported in **D9.1**, these metrics have been tested within the IS-ENES2 WP9/JRA1 activities. They have then been used to evaluate model performance throughout the project. A representation of European ESMs is given in Figure 8, displaying performance for the high-resolution versions (HR) used in WP9/JRA1 but also some other configurations. It emphasizes the high impact of resolution but also how complexity of ESMs impacts the performance.

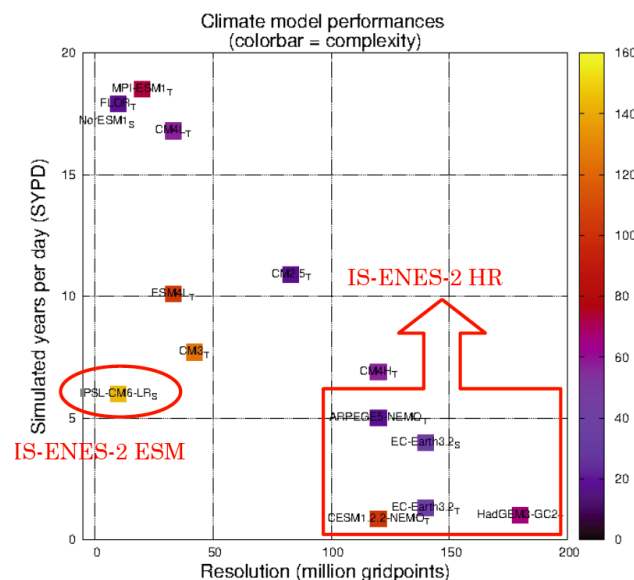


Figure 8: Computational performance in SYPD for European ESMs as a function of resolution (total number of grid points) and model complexity (color coded) (from D9.1)

3.2.2. Improving parallel I/O solutions and spreading its usage

Climate model simulations have to write out fields data throughout the simulation for diagnostic and statistical analysis. The amount of data written can be significant, depending on output frequency and resolution. The common method for writing this data was (and partially still is) to collect all data on a single process and write it to hard disk from there in serial. For today's climate models this is a significant bottleneck. Software have been developed which, instead of collecting the data on a single process, allows processing nodes to offload the IO to dedicated IO nodes that can parallelize the IO without blocking the computation in the model.

During IS-ENES1, two main approaches were considered both including parallel I/O libraries and I/O server but responding to different needs. Two approaches have been followed in the community: (i) the XIOS IO server developed by CNRS-IPSL and now used in the NEMO ocean model and (ii) the extension of the Climate Data Interface (CDI) library by MPG and DKRZ to give CDI-pio. Both use parallel IO and IO servers but in different ways. XIOS writes NetCDF which is more common in the climate community and CDI also writes in the GRIB format which is more common in the weather community.

Within IS-ENES2, XIOS has been applied to Earth system components, not only of the CNRS-IPSL coupled model, but also to the coupled atmosphere-ocean models of CNRM-CERFACS (initially was only planned for sea ice), and EC-Earth, both using NEMO. The three models will use XIOS for the CMIP6 experiments. The Met Office's Hadley Centre model has proposed its use in their next generation climate mode. A new version has been released XIOS2 (**MS9.8**) with a new internal design, enabling more easily the implementation of new functionalities: flexibility on field dimensions, use for input files, grid transformation, improved workflow to allow "in situ" pre and post-treatment of data.

CDI-pio release 1.6.7 (**MS9.5**) has included performance developments from IS-ENES1 and the G8 project ICOMEX. It supports the GRIB and NetCDF data formats and different types of grids. It significantly reduces the I/O overhead of the global coupled model MPI-ESM during the model integration.

In order to prepare for multi-model multi-member high-resolution (M4HR) ensemble experiments, the impact of I/O servers on high-resolution simulations has been evaluated (**D9.4**). Three models, which are planned to be used at very high resolution within the CMIP6 HiResMIP project, have been tested. ECHAM6 at 50 km resolution using CDI-pio shows a data output cost reduction from 43% to 6% when switching from serial to parallel output (Figure 9). On a different machine, with lower data intensities, ARPEGE5-NEMO and EC-EARTH3 at about 25 km (ARPEGE at 50 km), using XIOS2, show a limited data output cost of 1.5% and 1.1% respectively.

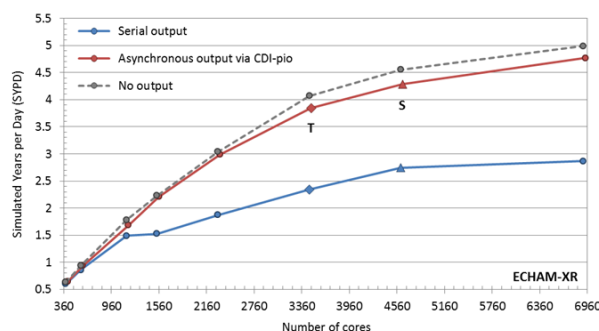


Figure 9: Simulated number of years per day (SYPD) for ECHAM-XR (T255L95) configuration with serial output (blue), asynchronous parallel output via CDI-pio (red) and without output (dashed grey) (From D9.4).

These tests do confirm that parallel I/O and the concept of I/O servers can significantly reduce the costs of data intensive applications such as climate models. Work done in IS-ENES2 also demonstrates that they can be integrated easily in a range of models. Moreover, these software have a strong potential for run time data manipulation for input and output, thus simplifying the workflow.

MS9.3 – Evaluate other ESMs compatibility with IPSL I/O server

MS9.5 - New CDI release with I/O developments from IS-ENES1 and ICOMEX

MS9.8 - New I/O server release

D9.4 – CDI-pio and XIOS I/O servers compatibility with HR climate models

3.2.3. Post-processing

The Climate Data Operators CDO are one of the dominant pre- and post-processing tool used in Earth System Modelling (see section 2.1.2). During IS-ENES2, the CDO have been further developed, with an emphasis on performance, since post-processing becomes more and more a bottleneck for large modelling projects. For example, searches on regular latitude/longitude grids have been improved considerably, leading to a gain by a factor of up to 100 for large grids. Further, support for unstructured grids of the type used in the new dynamical cores has been added (**MS9.6**).

The M4HR demonstrator in WP9/JRA1 (see section 3.3) requires the availability of efficient post-processing tools to allow for the analysis of large volumes of data. The ESMs used in the demonstrator utilise the CDO for their post-processing workflow. The M4HR components produce tens of gigabyte per simulated year, which have to be read from and written to disk by the CDO during the post-processing step. A thorough analysis of the disk transfer rates of the CDO (**D9.5**) demonstrates that the CDO post-processing does not constitute a performance bottleneck, as far as this factor is concerned. The analysis concludes that reading and writing the data for post-processing is orders of magnitude faster than data production of the model itself. Moreover, it shows that sequential processing of output of M4HR experiments does not slow down the overall workflow.

MS9.6 - New releases of CDO with improved parallel performance

D9.5 - CDO performance analysis

3.2.4. Impact of OASIS update

The increasing parallelism of climate models, particularly for high-resolution configurations, necessarily requires each of their components to scale. The OASIS coupling library is one of them and is used by five of the seven European ESMs used in CMIP6. During the first phase of IS-ENES, a new version of OASIS, OASIS3-MCT, was developed to enable the direct communication of coupling fields and the parallelization of interpolations between components of the models. The objective within IS-ENES2 was to further improve OASIS3-MCT and test its implementation in coupled ESMs to be used in CMIP6.

During IS-ENES2, new versions of OASIS3-MCT have been released and tested. WP9/JRA1 results with coupled ESMs emphasized the dominating contribution to coupling cost of the load imbalance between components. Thanks to better set up techniques, the WP9 partners significantly reduced this load imbalance and the associated coupling cost (Table 6).

Models	First measurements (2014)	Final measurement (2016)
ARPEGE5-NEMO	13%	1%
EC-Earth3	24.7%	4%
HADGEM3-GC2	n/a	15%

Table 6: Coupling cost (% of total CPU cost) of HR models

MS9.7 – Model interface updates for compliance with new OASIS coupler

D9.2 - HR ESM performance resulting from OASIS updates

3.3. Prepare future high-end experiments

WP9/JRA1 aimed at supporting the development of multi-model multi-member high-resolution (M4HR) ensembles of models. This was both to support the preparation of high-resolution runs for CMIP6 but also to test the feasibility to run a large ensemble of simulations efficiently, in order to prepare for such applications to be used on PRACE resources.

Initially five models were candidate for this experiment. However, due to lack of computing resources at the end of the project, it was rather decided to implement a simpler M4HR demonstrator on a common computing platform (MareNostrum3 at BSC), using the latest available versions of two ESMs (ARPEGE5-NEMO and EC-Earth3), and controlled by the common workflow management system Autosubmit developed in IS-ENES2.

3.3.1. Workflow management system for running ensembles of simulations

In order to run a M4HR demonstrator efficiently, a dedicated workflow has been prepared. It has been based on the Autosubmit software developed initially at IC3 to manage and run the research group's experiments. Lack of in-house HPC facilities at IC3 led to a software design with very minimal requirements on the HPC that will run the jobs. Autosubmit provides a simple workflow definition capacity that allows running weather, air quality and climate multi-member experiments on more than one supercomputing platform. Autosubmit includes a general-purpose wrapper that packs multi-member simulations into a single executable with suitable job control.

Within IS-ENES2, the Autosubmit software has been tested for multi member HR experiments with one model (**MS9.1**) and then further developed into a general-purpose submission and monitoring tool optimized for M4HR ensembles (**MS9.2**). The final version of Autosubmit (3.0), has then been released to the community through an open source licence (GNU) (**MS9.9**). It is now distributed and maintained at BSC¹⁸

Before applying Autosubmit for M4HR ensembles, a comparison of three different workflow management systems has been conducted. It is based on a comparison between Autosubmit, Cylc and ec-flow. The Cylc suite engine is a workflow engine and meta-scheduler for weather forecasting and climate modelling. It is designed to run operational suites with complex date-time cycling requirement. Cylc was initially created at NIWA (NZ) and is now co-developed with the Met Office and used for a wide range of applications across many platforms worldwide. The workflow package ecFlow enables users to run a large number of programs in a controlled environment. It is used at ECMWF to manage around half their operational suites across a range of platforms. The assessment demonstrated that Autosubmit, Cylc and/or ecFlow are all three suitable options to define, set-up and run such a M4HR experiment (**D9.3**). Based on feedback from the first workflow workshop, Cylc was further developed to meet the communities needs (see foreground for details).

MS9.1 – Multi-member HR prediction experiment using Autosubmit on Tier 1

MS9.2 – Further developments of Autosubmit

MS9.9 – Release of Autosubmit to the ENES community through GNU licence

D9.3 - Assessment report on Autosubmit and Cylc

3.3.2. M4HR demonstrator

Two high-resolution models (ARPEGE-NEMO and EC-Earth) were configured to perform an ensemble of high-resolution coupled simulations (*M4HR demonstrator*). Up-to-date versions of libraries and tools (OASIS coupler, XIOS I/O subsystem, CDO post processing tool and Autosubmit job controller) were included in this demonstrator. The purpose of this exercise was technical but it reproduces the behavior of a realistic workflow of two high-resolution models producing data. Computational performance of the two M4HR demonstrator models, as well as other WP9/JRA1 ESMs outside of the demonstrator, has been measured and documented (**D9.6**), making use of the performance metrics earlier defined in the WP9/JRA1 (D9.1, see section 3.2.1).

The M4HR demonstrator proves the technical feasibility of running an ensemble of multi-model climate simulations. This contributes an important know-how for future model comparison projects. Moreover, it has provided the WP9/JRA1 modelling groups with a valuable experience in the coordinated development and implementation of a shared workflow for climate simulations. Finally, the development of the M4HR demonstrator has clearly shown the requirements, as well as current shortcomings, related to the shared usage of European computational resources at large scale. The M4HR demonstrator was too complex for a

¹⁸ Autosubmit is available at <http://www.bsc.es/projects/earthscience/autosubmit/>

Preparatory Access (as defined by PRACE), but its purpose (technical set up) required too few resources for a comprehensive Project allocation. Hopefully, the involvement of a computing center in the project partnership helped to address the mismatch between our needs and the PRACE offer.

D9.6 - M4HR ESM ensemble performance analysis

3.4. Develop coupled benchmarks

A coupled benchmark suite has been developed within IS-ENES2. This suite includes four coupled ESMs, one component model and two intensive computational kernels. The coupler benchmark suite (see 2.2.2) has been added to the list. All benchmarks are available on the ENES portal (<https://portal.enes.org/computing/benchmarks>) and are listed in Table 7.

Description	Type
MPI-ESM1	Coupled ESM used in CMIP5
IPSL-CM	Coupled ESM used in CMIP5
CMCC-CESM-NEMO	Coupled ESM in preparation to CMIP6
EC-EARTH	Coupled ESM used in CMIP5
ICON	Dynamical core
NEMO tracer advection	Kernel
ICON communication	Kernel
Coupling benchmark suite	Version 1.0.0

Table 7: List of benchmarks available on ENES portal

The benchmarks provide insights into computational characteristics of ESMs. These are important to benchmark HPC systems for procurement, to provide vendors a better way to assess performance characteristics of typical climate applications, to reduce the time needed to port applications to new systems, to monitor and compare performance, provide testbeds for computer scientists.

A framework has been defined (**D10.1**) composed of a service platform on the ENES portal, the benchmark suite, a set of common performance metrics based on the CPMIP metric (see 3.2.1), lightweight performance measurement and validation tools, and a central benchmark repository. The benchmarks have been described in **D10.2** with a more complete report on performance in **D10.4**.

The assembled benchmark suite has been presented to vendors at the IS-ENES2 workshop series of workshops on Innovation in HPC for climate models (see 3.5). Vendors have appreciated the availability and access to benchmarks as they help them get familiarized with the codes outside formal benchmark exercises. However, it was emphasized that the value of benchmarks could be lowered if too complex, hard to execute, or imposed high requirements in human resources. Nevertheless, IS-ENES2 strongly recommends using coupled benchmarks in selecting new HPC systems. However, if not possible, IS-ENES2 recommends to use at least a coupled toy model and to test the real coupled system during the installation phase of a system.

To stay relevant, such coupled benchmark suite would need to keep pace with model upgrades. This is a difficult task. For example, the suite assembled within IS-ENES2 is based on CMIP5 coupled models and not on CMIP6 versions, which are just getting finalized end of 2016. However, several partners have made commitments to continue the work on benchmark development beyond IS-ENES2 project and make up-to-date benchmarks available.

D10.1 - Documentation and user guide for benchmarking framework

D10.2 - Interim report on status of benchmark suite

D10.4 - Report on the suite of base benchmarks and analysis of performance on available platforms

3.5. Support innovation with HPC vendors

IS-ENES2 fostered interactions with HPC vendors and ICT companies at a dedicated session at the 3rd ENES HPC workshop and at a series of talks with vendors organized by IS-ENES2.

At the 3rd HPC Workshop, high-level technical staff from Intel, DDN, BULL/Atos, IBM, Cray, NEC and NVIDIA gave talks to inform the climate modelling community on advances on technology. The topics focused on hardware developments planned for the following few years, on co-design and cooperation with climate scientists or on companies' general perspectives on HPC and extreme scale computing. They are reported in D3.1 (see 3.1).

From November 2016 and January 2017, vendors from NEC, Cray Ltd., Bull / Atos, IBM, Intel and NVIDIA were invited to present an update on their roadmaps. The companies were explicitly asked to talk about planned developments for a longer timeframe - until 2020 – 2025. A core element of the workshop was the intensive interaction of scientists and vendors to get a bidirectional understanding of: 1) the vendors hardware and system strategies, 2) the consequences this may have for software development and the design of next generation climate models, 3) to make the vendors aware of the special requirements that the weather and climate modelling community has, and 4) to encourage the discussion about co-design projects. The community needs to coordinate its approach with vendors to have a significant impact given the external drivers away from our needs, as summarized in Table 8. The major conclusion is: Software development will need to gain even more focussed attention in the community, and according investments into scientists with background knowledge in the field, but also advanced knowledge of Hard- and Software architectures, software engineering and performance programming. It is people that count, and ENES needs to invest here!

Vendors' perspective	Consequences for strategy in ESM
Heterogeneity of systems (mixture of CPU / GPU / accelerators) will increase	Increasing number of abstraction in memory levels have to be expected and will require serious code adaption to use the systems efficiently. This could include delegating system adaption to tools like OpenMP, DSL and/or subdivide models into software layers (abstraction) on top of optimized kernels (tuning level).
Increasing number of cores per socket with decreasing amount of local memory per core. Decreasing Memory Bandwidth and network bandwidth per core. Clock rate of cores cannot be expected to increase in the same rate as in the past. On the contrary, it will probably decrease.	Performance improvements can only be achieved by a higher level of parallelism, which then will increase code complexity, will substantially impede code development and efficient use of the new systems. Higher parallelism with at the same time minimal communication will be required and might narrow the choice of algorithms available depending on respective communication patterns.
Hardware development will increasingly focus the mass market (e.g. mobile phones) and emerging new market segments (e.g. deep learning) and decreasingly serve classical application areas with special requirements (e.g. climate science)	This may lead to reduced engagement in supporting special toolkits, compiler, libraries, higher support costs or to slower response times when it comes to bug fixing etc.

Table 8 summarizes the vendors' perspective on future architectures and their consequences for the strategy to be followed by ESMs.

D6.4 - Report on collaboration between ENES and HPC companies

4. Support the dissemination of Earth system model data to the climate and impact research communities

IS-ENES2 supports the European contribution to the WCRP internationally coordinated climate modelling experiments, CMIP and CORDEX. CMIP5 service access has been essentially developed during IS-ENES1 and continued in IS-ENES2. IS-ENES2 aimed at adding services on metadata, on CORDEX, as well as at supporting the preparation of CMIP6. It also aimed at supporting the dissemination of model data results to a wider community investigating impacts of climate change and developing climate services. IS-ENES2 data activities are highly integrated within the international climate community. IS-ENES represents Europe in the Earth System Grid Federation (ESGF) data distribution, in the Earth System Documentation of metadata

(ES-DOC) and contributes to the international Working Group on Coupled Models Infrastructure Panel (WIP).

Data and metadata activities in IS-ENES2 are delivered through the interaction of three closely interlinked work packages on networking (WP5/NA4), services (WP8/SA2) and development activities (WP11/JRA3). Networking activities gathered requirements and set standards, which fed into the service and product development, and generated the guidance documents that became part of the services delivered. Development was intended to both develop and maintain software and information for deployment in the services.

D11.1 – Delivery plan coordinated with NA4 and SA2

4.1. Service of the ENES climate data infrastructure

4.1.1. The ENES climate data infrastructure

The IS-ENES2 data services are based on the ENES Climate Data Infrastructure. It is a distributed infrastructure that delivers access to climate model and metadata services as well as tools to ease access and use of these results. These services are in open access from the ENES portal (<https://portal.enes.org>), which provides information on how to search and access data using ESGF, how to access metadata using ES-DOC, provides support to users and providers of data and provides processing tools. These data can also be accessed directly from any of the ESGF portals or from ES-DOC.

At the start of IS-ENES2, the service included CMIP5 data from the seven European global climate models. Each model data is stored on a specific data node (seven in Europe for CMIP5) and information can be searched and accessed by users from any of the index nodes (four in Europe). The service also included a more dedicated access for non-climate modelers, such as the community working on climate impacts, through the IPCC Distributed Data Centre (DDC) located at DKRZ as well as a prototype of the climate4impact portal (see 4.1.8).

During IS-ENES2, the service has been increased with the provision of data for CORDEX (see 4.1.4), the metadata services (see 4.1.7) and the climate4impact portal (see 4.1.8). IS-ENES2 also contributed to further developments of ESGF and ESDOC. **D5.1** served as a guide to define the main data archive governance and requirements. Two internal reviews have been conducted that helped improve the description of the service on the ENES portal, especially to simplify the paths to the relevant information.

D5.1 – Data archive governance and requirements document

MS8.2 – Set-up of the review committee

MS8.4 – First review report on the ENES climate data infrastructure

MS8.7 – Second review report on the ENES climate data infrastructure

MS8.9 – Improved data access available from ENES climate data infrastructure

4.1.2. Earth System Grid Federation overview

ESGF is a peer-to-peer (P2P) federated data archive, which provides access for users to distributed data nodes (where model data are stored) through index nodes and identity providers (Figure 10) (Williams et al., 2016). Current data holding, through ESGF data nodes, are also displayed on Figure 10, showing a strong involvement of USA and Europe.



Figure 10: (left) description of the peer-to-peer service of ESGF and (right) location of ESGF datanodes (6th ESGF F2F annual meeting report, 2017 and D11.5).

In December 2016¹⁹, ESGF supports more than 700K datasets and manages over 4.6 petabytes (PB) of data including replicated datasets (about 2PB without replica). IS-ENES2 supports the European contribution to CMIP5 and CORDEX. CMIP5 represents the main part of the ESGF federated data archive in terms of data volume with about 4.3 PB (corresponding to more than 150K datasets).

The number of registered users, which at the time this report is being written, is about 14K. The registered users distribution by continent is reported in Figure 11. The registered European user distribution by country is reported shows a very high degree of geographical coverage. A detailed view about the data downloads, volume of data downloaded, number of active users, and number of files is reported in Section 4.1.5.

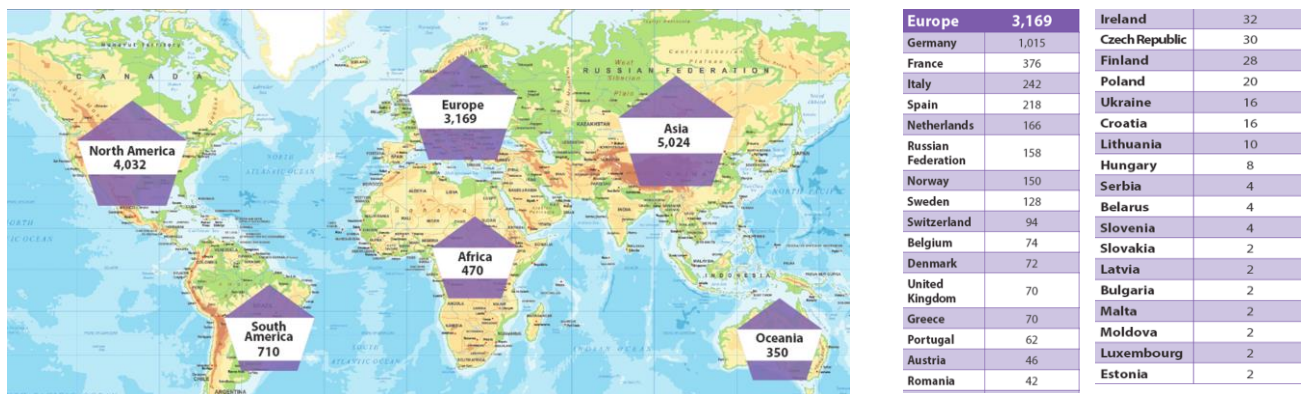


Figure 11: Distribution of ESGF registered users by continent and details for Europe (right) (6th ESGF F2F annual meeting report, 2017 and D11.5).

4.1.3. IS-ENES2 contribution to ESGF/CMIP

The publication of CMIP5 data from European climate models has been supported within IS-ENES1. IS-ENES2 has continued to serve data from CMIP5 and still contribute an important part of the download shown in section 4.1.5.

From July 2015 to December 2015, ESGF has been interrupted following a security breach. All components of the ESGF software have been completely reconditioned, partly rebuilt and extensively tested, with a strong involvement of IS-ENES2 partners (see 4.2.1). The ESGF software release processes have been redefined after the security breach. This led to reorganization of some of the European ESGF datanodes. The

¹⁹ Report of the 6th ESGF F2F annual meeting, December 2016 available at https://esgf.llnl.gov/media/pdf/2017-ESGF_F2F_Conference_Report.pdf

current status of CMIP5 datanodes for Europe is displayed Table 9. Three of the nodes, DKRZ, CEDA and IPSL, provide identity to users. They are ranked ranked 2nd (DKRZ), 4th (IPSL) and 5th (CEDA) among the seven identity providers world-wide (4 in USA, 3 in Europe).

Host institution	ESGF datanode	Number of datasets	Institutes providers of data published
CEDA (UK)	esgf-data1.ceda.ac.uk	10108	MOHC (UK)
DKRZ (DE)	esgf1.dkrz.de	4717	MPI-M (DE)
IPSL (FR)	vesg.ipsl.upmc.fr	3933	IPSL (FR)
IDRIS (FR)	prodn.idris.fr	120	IPSL (FR)
LIU (SE)	esg-dn1.nsc.liu.se	1840	EC-EARTH group
ICHEC (IE)	esgf.ichec.ie	374	EC-EARTH group
Meteo France (FR)	esg.cnrm-game-meteo.fr	1075	CNRM-CERFACS (FR)
NorStore (NO)	noresg.norstore.no	591	NOR-ESM group (NO)
PCMDI (USA)	aims3.llnl.gov	280	CMCC (IT)

Table 9: List of CMIP5 datanodes hosting data from the seven European global climate models (from ESGF on 21/05/2017)

4.1.4. IS-ENES2 contribution to ESGF/CORDEX

CORDEX has for the first time coordinated regional experiments to create a multi-model ensemble. They are downscaled from CMIP5 with higher resolution but limited on regional subdomains, either dynamically using regional climate models forced by CMIP5 global models or using statistical methods applied to CMIP5 results. Europe is one of the 13 domains with two nominal resolutions, at 44 km (EUR-44) and 11 km (EUR-11) coordinated within Euro-CORDEX (Figure 12). European modelling groups have also run simulations on different other domains, with a particular focus on Africa (at 44 km). Europe is also concerned with two other domains, Arctic-CORDEX and Med-CORDEX centered respectively on the Arctic and the Mediterranean areas.

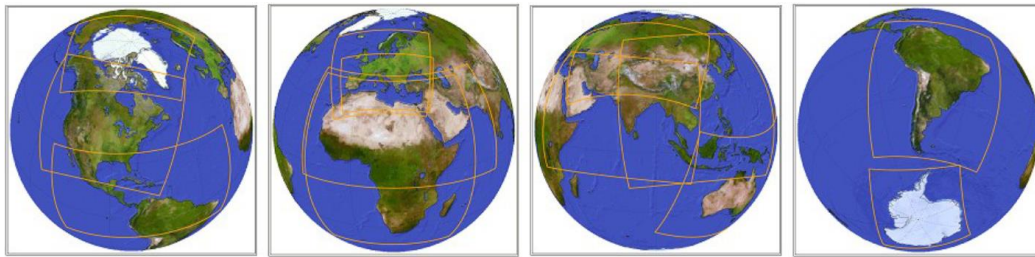


Figure 12: CORDEX domains

IS-ENES2 has promoted the use of ESGF for CORDEX, which has then been adopted by the CORDEX Scientific Advisory Team. Not all domains have been ported onto ESGF yet, but many of the datasets from European groups are there thanks to IS-ENES2 support. IS-ENES2 teams have also been asked to give training in Asia (2014 Nanjing University, 2016 Korea). IS-ENES2 has integrated data from Euro-CORDEX as well as the European contribution to Africa-CORDEX and more recently some Arctic-CORDEX. Med-CORDEX, however, has kept its own database and, although several times incited by IS-ENES2, has been reluctant to port its data on ESGF due to the amount of effort and lack of resources. IS-ENES2 has also supported the development of datanodes for Empirical Statistical Downscaling (CORDEX-ESD) with a node in Spain and the first ESGF node in Africa. They will host ESD data when they will be ready.

The status of CORDEX data on ESGF at the end of IS-ENES2 is largely limited to data published from European modelling groups (Table 10), with a volume of about 60 TB of data and 69 K datasets. Note that

IS-ENES2 datanodes are hosting data from other institutes than IS-ENES2 partners, in Belgium, Croatia, Germany, Hungary, Italy, Switzerland, as well as Russia and Canada.

Indeed, porting CORDEX data on ESGF requires adherence to standards for file and data formats, as well as for archive content. It includes a common naming system, the Data Reference Syntax (DRS), which allows the identification of data sets. The file format (NetCDF-CF), the naming system, as well as the metadata build on controlled vocabularies (CV). A quality control procedure ensures that 100% of the data are compliant with the DRS and CVs. Detailed information is available on the ENES portal at <https://portal.enes.org/data/enes-model-data/cordex> as well as on <http://is-enes-data.github.io/> with access to software at <https://github.com/IS-ENES-Data/IS-ENES-Data.github.io>. IS-ENES2 has supported the development of the CORDEX Archive specifications and CORDEX Variable Requirement table available on these sites.

Institution hosting the datanode	ESGF node	Number of datasets	Institutes providers of data published
NSC-LIU (SE)	<i>esg-dn1.nsc.liu.se</i>	48543	SMHI (SE), ICTP (IT), MGO (RU) & CORDEX-Adjust
DKRZ (DE)	<i>esgf1.dkrz.de</i>	13441	CLM community (DE, CH), AWI (DE) , DHMZ (HR), MPI-GERICS (DE), RMIB-UGent (BE)
DMI (DK)	<i>cordexesg.dmi.dk</i>	12928	DMI (DK), HMS (HU), KNMI (NL), ULg (BE), UQAM (CA)
CEDA (UK)	<i>esgf-data1.ceda.ac.uk</i>	2611	MOHC (UK)
NCI (AU)	<i>esgf.nci.org.au</i>	2550	UNSW (AU)
CCCR (IN)	<i>esg-cccr.tropmet.res.in</i>	2526	IITM (IN)
IPSL (FR)	<i>vesg.ipsl.upmc.fr</i>	929	IPSL-INERIS (FR) & CORDEX-Adjust
Meteo France (FR)	<i>esg.cnrm-game-meteo.fr</i>	422	CNRM (FR)
<i>U. Cantabria (ES)</i>	<i>data.meteo.unican.es</i>	<i>2500 (tbc)</i>	<i>UCAN (ES), CORDEX-Adjust,</i>
<i>CSAG (S. Africa)</i>			<i>CORDEX-ESD</i>

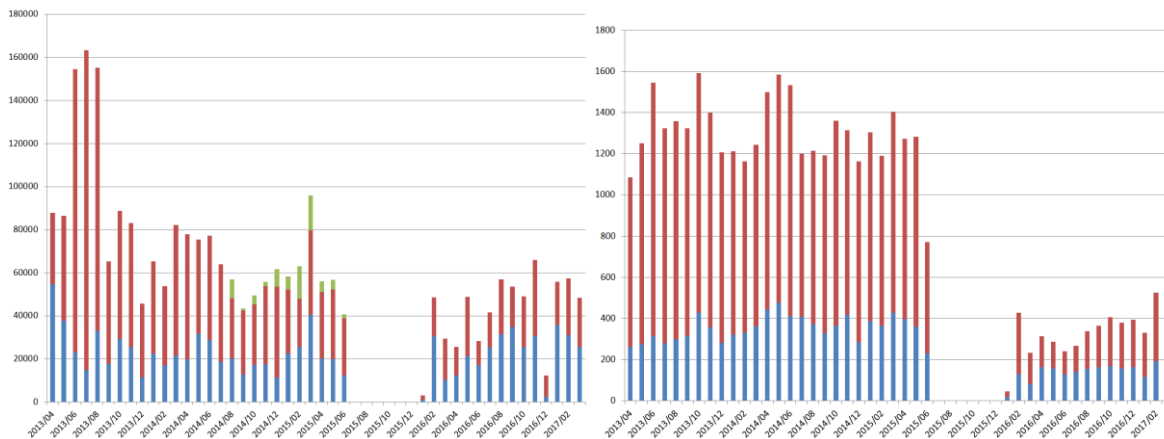
Table 10: Status of ESGF CORDEX datanodes (21/05/2017). In black, groups having run EuroCORDEX and some other domains, in blue institutes running other domains. In bold, IS-ENES data nodes (from ESGF CORDEX, 21/05/2017). Italic, datanodes for CORDEX-ESD ready but still to be opened when data are available.

D5.4 - ENES Data Service Infrastructure requirements

4.1.5. Key performance Indicators for ESGF European datanodes

After the mid-term review, quantitative Key Performance Indicators (KPIs) for IS-ENES2 have been developed to describe the activity of the data infrastructure. They quantify the volume of data downloaded from IS-ENES2 ESGF datanodes, the associated number of files as well as the number of active users on those datanodes (Figures 13).

The statistics show the break during the ESGF overhaul after the security incident. Nevertheless, IS-ENES2 helped users access CMIP5 data by: 1) providing access to the IPCC Data Distribution Centre and to the CERA Long Term Archive hosted at DKRZ and 2) providing a long list of alternative access points (mostly data nodes of the data providers) on the ENES data service website. The latter has been referenced by various ESGF related other websites.



Figures 13: Key performance indicators for ENES data infrastructure from 04/2013 to 03/2017: (Left) Volume of data downloaded (in GB) from ESGF European datanodes, (Right) number of active users on the nodes. Blue color for European users and Red non-European users, Green for use of WDC archive.

4.1.6. Helpdesk service

A helpdesk support service has been put in place to help users of ESGF. Support to users is mainly achieved through a user mailing list. Members of IS-ENES2 work within the ESGF Support Working Team, providing expertise.. Statistics of the ESGF help support shows that answers to requests are given very rapidly, less than a day, and that IS-ENES2 contributes most of them (order of 80%) (Figure 14). Moreover, the DATA tab of the ENES portal provides general information on ESGF which is the only place where such information is available.

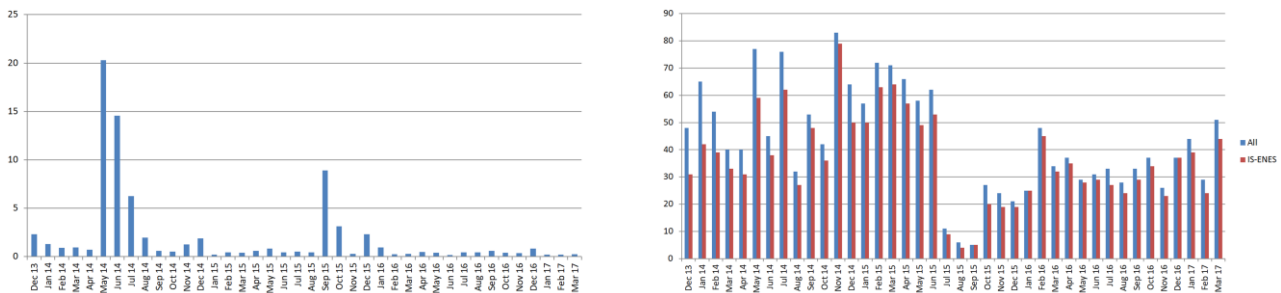


Figure 14: ESGF helpdesk from 12/2013 to 03/2017. (Left) Mean response time of ESGF user support (in days); (Right) Total number of email threads (blue) and IS-ENES2 contribution (red) (KPI).

MS8.1 – ENES climate data infrastructure helpdesk

4.1.7. Service on metadata

An important aspect of CMIP5 is the addition of metadata to document models and simulations. A Common Information Model (CIM) has been developed by the ENES FP7 METAFOR project in collaboration with the US project Curator. During IS-ENES2, various services on CIM metadata have been implemented and further improved. These are CIM Creation Services (ES-DOC python client, CMIP5 questionnaire, ES-DOC questionnaire) and CIM Viewing Services with the CIM viewer plugin. In addition, a CIM Document Comparator (front-end and API) and search capabilities (front-end and API) were established. All those services have been integrated (MS8.3) and can be accessed from <https://search.es-doc.org/>. For each of them, statistics of usage are available on the web since February 2016 (<http://stats.api.es-doc.org/cgi-bin/awstats.pl>) via a monitoring tool and presented at Table 11. In addition, there is a user-oriented description of the services on the ENES portal. The viewer interface provides an easy access to European model documentation.

Period	Unique visitors	Number of page hits
February 2016-December 2017	2290	8000
Jan 2017-March 2017	404	1595

Table 11: Statistics on the usage of ES-DOC model documentation

MS8.3 – Integration of CIM model and experiment repository into ENES CDI

4.1.8. Service for climate impact communities: the climate4impact portal

The climate4impact portal (<http://climate4impact.eu>) is a platform that aims to ease access to ESGF CMIP5 and CORDEX data specifically for the community working on climate impacts. This community, not working on climate models, not only needs access to model data to investigate the impact of climate change on different sectors (e.g. water, health, energy, insurance, agriculture...) but also access to guidance and analysis facilities. During IS-ENES1, a prototype of the portal was developed including a documentation of use cases (Deandreis et al., 2014). The climate4impact portal has been further developed within IS-ENES2 with addition of functionality (see section 4.4).

The climate4impact portal (C4I) is built on the ESGF infrastructure. It uses the data services provided by ESGF and makes them user friendly accessible to this community. The deployed functionality is developed based on user requirements, as stated in requirements surveys of IS-ENES1 and feedback provided by impact users gathered at several conferences and workshops during IS-ENES2.

At the end of IS-ENES2, the climate4impact portal has become operational with 750 registered users. It has evolved from a portal to a platform offering an interface and reusable services to explore data and perform analyses, it includes: sign-in (using ESGF or a Google OpenID); search and access to CMIP5, CORDEX and other project data stored on ESGF with faceted filters; web processing services to perform on-demand computations; visualisation, documentation and guidance; possibility to perform statistical downscaling of data and computation of climate indices and indicators. These are described in section 4.4. Statistics on the usage of the C4I, shows in average for 2015-2016 about 1600 unique visitors per months and 23 GB of data downloaded (Figure 15).

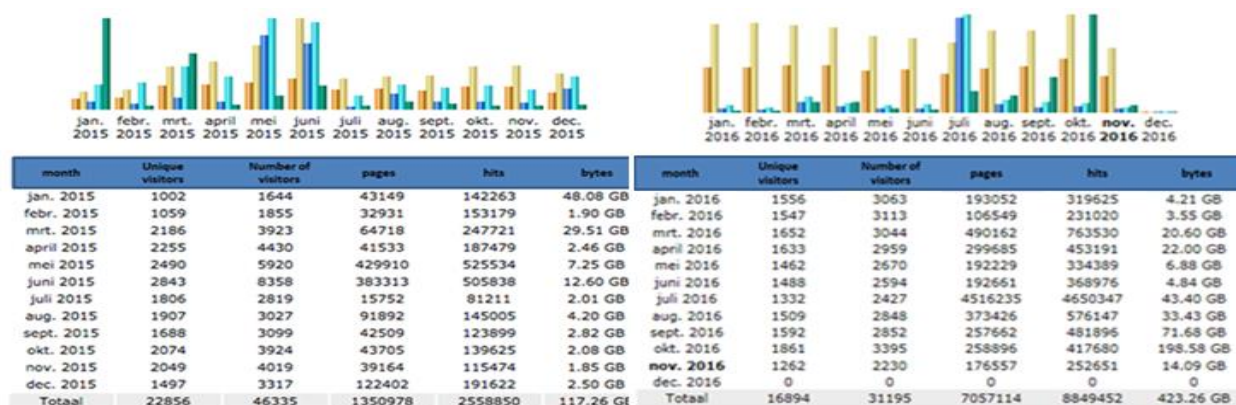


Figure 15: Statistics of usage of the climate4impact portal for 2015 and 2016 (KPI).

MS8.5 – JRA3 developed service deployed on climate4impact

MS8.8 – climate4impact components available for use in tailored impact portal

4.2. Development of core services on ESM data (ESGF)

4.2.1. Contribution to ESGF development

ESGF system requires support and development services. ENES teams, supported by IS-ENES2, are involved in leading or co-leading half of the ESGF sixteen working teams (Table 12), as described in the ESGF First Implementation Plan²⁰. IS-ENES2 has supported developments on many aspects of ESGF, in

²⁰ ESGF First Implementation Plan, Williams et al., May 2016, available at <https://esgf.llnl.gov/>

particular for the overhaul after the ESGF security breach and preparation for CMIP6. As detailed in D11.5, IS-ENES2 is involved in: installation; publishing services; search services; user interface; security; data, transfer, network and replication; computing services; metadata services; provenance, capture, integration and usability; and dashboard monitoring.

Working Team	Working Team Leads	Team Goals
1. CoG User Interface Working Team	Cecelia DeLuca (NOAA) and Luca Cinquini (NOAA)	Improve ESGF search and data cart management and interface
2. Metadata and Search Working Team	Luca Cinquini (NASA)	Implement ESGF search engine based on Solr5 and discoverable search metadata
3. Publication Working Team	Sasha Ames (DOE) and Rachana Ananthakrishnan	Enable capability for publishing CMIP and other project data sets to ESGF
4. Node Manager Working Team	Sasha Ames (DOE) and Prashanth Dwarakanath (IS-ENES)	Manage ESGF nodes and node communications
4a. Tracking/Feedback Notification Working Team	Sasha Ames (DOE) and Prashanth Dwarakanath (IS-ENES)	Implement user and node notification of changed data in the ESGF ecosystem
5. Identity Entitlement Access Management Working Team	Philip Kershaw (IS-ENES) and Rachana Ananthakrishnan (DOE)	Implement ESGF X.509 certificate-based authentication and improved interface
6. Compute Working Team	Charles Douthiaux (DOE) and Daniel Duffy (NASA)	Develop data analytics capability within ESGF
7. Quality Control Working Team	Martina Stockhause (IS-ENES) and Katharina Berger (IS-ENES)	Integrate external information into the ESGF portal
8. Installation Working Team	Nicolas Carenton and Prashanth Dwarakanath (IS-ENES)	Install the components of the ESGF software stack
9. Dashboard Working Team	Paola Nassisi (CMCC) and Sandro Fiore (IS-ENES)	Monitor the Earth System Grid Federation in terms of system metrics and data usage statistics
10. International Climate Network Working Group	Eli Dart (DOE/ESnet) and Mary Hester (DOE/ESnet)	Increase data transfer rates between the ESGF climate data centers
11. Data Transfer Working Team	Lukasz Lacinski (DOE) and Rachana Ananthakrishnan (DOE)	Enhance ESGF data transfer and web-based download
12. Software Security Working Team	George Rumney (NASA) and Dan Duffy (NASA)	Implement security measures to identify vulnerabilities in the ESGF software and provide continuous improvement to the ESGF software development life cycle.
13. Support Working Team	Torsten Rathmann (IS-ENES) and Matthew Harris (DOE)	Develop frequently asked questions regarding ESGF and housed data
14. Documentation Working Team	Matthew Harris (DOE) and Sam Fries (DOE)	Document the use of the ESGF software stack
15. Replication and Versioning Working Team	Stephan Kindermann (IS-ENES) and Tobias Weigel (IS-ENES)	Create replication tool for moving data from one ESGF center to another; in addition, preserve versioning history of the ESGF published data sets
16. Provenance Capture Working	Bibi Raju (DOE)	Enable ESGF provenance capture for

Table 12: ESGF Working teams leadership, from ESGF 1st Implementation Plan (2016)

The scale of European involvement allows project partners to maintain active engagement in all areas of the global collaboration, and hence ensure that potential problems in the timelines of development can be monitored and mitigated. It is highly likely that this collaboration will continue to be productive, held together by a shared commitment to the CMIP process.

D11.5 – Report on Core services package, version 2

4.2.2. Software developments: Quality control

Data quality is very important for ESGF, more especially as the volume of data is large and widely used (**D5.3**). Quality control includes checking that data are compliant with standards of format and content. In CMIP5, the quality of data was mainly to the responsibility of the data producers with some checks done by the data centres and communicated to the producers and consumers. In CMIP6 the role of data nodes will be strengthened with the possibility to accept or refuse the data from any producer not following the agreed standards. The data producers, on the other hand, have access to QC software to ensure the compliance of the data to the most important ESGF standards before delivery.

The core of the Quality Assurance (QA-DKRZ) package (Figure XX and **MS8.6**) has been modularised and additionally expanded to the CORDEX project. NetCDF Climate and Forecast (CF) Metadata Conventions, Directory Reference System (DRS), Controlled Vocabulary, variable requirements, and text based project rules are checked. The software package is commonly accessible from the IS-ENES-Data section on GitHub and a QA-DKRZ user-guide is available on readthedocs.org. At present, QA-DKRZ is used in Europe, Canada, and Asia.

4.2.3. Service monitoring and dashboard

A monitoring system has been developed within IS-ENES2, based on the Federated Archive System Monitoring (FASM). The main goal of the FASM system is to provide a distributed and scalable monitoring framework responsible for capturing usage metrics at the single site level, both at the ENES climate data infrastructure level and at the global ESGF level. The FASM faces these goals through two main modules: the FASM-N managing statistics at each site and the FASM-D, a dashboard module providing a flexible user interface. FASM has been developed in collaboration with PCMDI.

The FASM monitoring system has provided both the overall ESGF usage statistics for CMIP5 and CORDEX emphasized in above section 4.1.2 (Figure 16), as well as the statistics for usage of ENES data infrastructure displayed in IS-ENES2 KPIs (Figures 13 and 14).

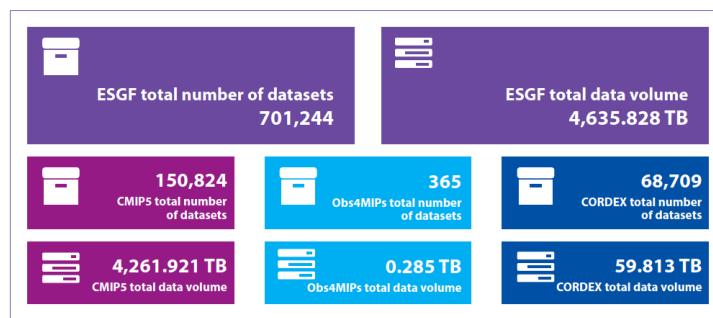


Figure 16: ESGF data archive statistics obtained from the ESGF Console (from the 6th ESGF F2F annual meeting report, 2017 and D11.5). Contributions from CMIP5, Obs4MIP and CORDEX are detailed. IS-ENES2 supports the European contribution to CMIP5 and CORDEX

4.3. Development of Metadata standards and services

During IS-ENES2, the metadata standards and services have been further developed in preparation of CMIP6. **D11.4**, developed internationally through the WIP outlines the items to be documented such as experiments, models and simulations used to generate the CMIP6 datasets. It lists the key properties and features of these documents based on CMIP5 taking into account lessons learned, the underlying tools and workflows as well as what modelling groups should expect and how they should engage with the documentation of their contribution to CMIP6.

For CMIP6, the general approach has been to simplify and streamline the process as much as possible so as to ease the work of the modelling groups and adapt to their internal timelines and activities. Most documents will be created either automatically or by internal ES-DOC effort, with minimal input from the modelling groups (Figure 17). The main effort will remain to document the model but unlike for CMIP5, groups will have several creation tools available and will be able to start from their existing CMIP5 model description (<https://search.es-doc.org/>). These new CMIP6 documents can be created independently and in the order the groups wish. Their connection and access via the further-info-URL global attribute of the netCDF files will be dealt with directly by ES-DOC.

IS-ENES2 has been the major contributor to the international ES-DOC project. It is estimated that about 80% of the development and coordination efforts was directly funded by IS-ENES2, mostly via CNRS-IPSL, University of Reading (UREAD) and MPI-M. Some key software and the overall project management were led by UREAD. CNRS-IPSL was the lead developer on most of the software while MPI-M contributed expertise and links with other aspects of CMIP6. All groups participated in the governance and linked with

the rest of the CMIP6 ecosystem and, most prominently, with standards and ESGF. The strong European membership of the international WGCM/WIP, which oversees the CMIP6 process, is a testimony of the expertise gathered by IS-ENES2.

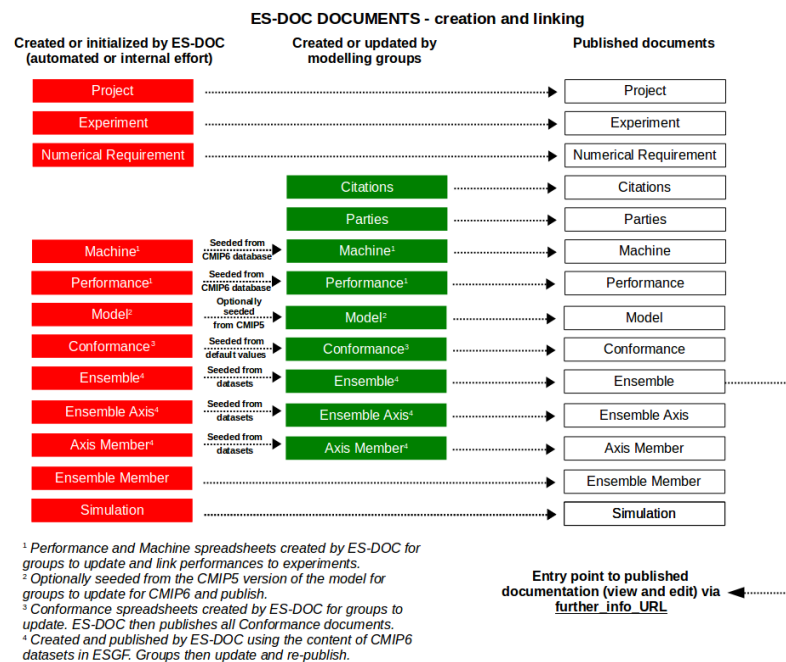


Figure 17: Overview of ES-DOC documents and creation workflow for CMIP6. The right hand side lists all documents and their relationships. The documents listed on the left are generated by ES-DOC (either automatically or via internal effort). The central column lists the documents either created by modelling groups (Citations, Parties, Machine and Model) or updated toward their final version. Latest version and complete CMIP6 documentation ecosystem can be found on <http://es-doc.org/cmip6>.

D11.4 – Report on meta-data services

4.4. Development of data access services for climate impacts

During IS-ENES2, tailored products have been added to the climate4impact portal (**D11.6**) following feedback from users (e.g. **D5.2**). The tailored derived products range from bias-corrected data, to climate indices and indicators, statistical downscaling, specific Use Cases, and also the integration with other portals (Figure 18).

IS-ENES2 provided a major input and support to the Bias Correction Inter-comparison Project (BCIP) initiative that was distributed across several European projects. Four IS-ENES2 partners (CNRS-IPSL, SMHI, UC, MetNo) were involved in the initiative, which focussed on bias adjustment of key variables produced by the Euro-CORDEX collaboration. IS-ENES2 has, together with the CLIPC project, provided essential support in developing the meta-data standards and tools necessary for publishing the data onto ESGF (**D5.4**). These standards have been formally endorsed by WCRP under the ESGF Project heading “CORDEX-Adjust”. In addition, publication of the bias adjusted data benefitted substantially from technical collaboration within IS-ENES2 and the resulting development of software tools and quality control procedures (Table 10).

The users of climate scenario data are very heterogeneous. Having access to pre-computed climate indices and indicators can support many of the needs, but users very often require more flexibility for tailored analyses given their specific needs. The climate4impact (C4I) platform now offers the possibility for users to calculate climate indices and indicators with their own parameters using any input data file accessible in either their C4I Basket, or on any ESGF datanode or OPeNDAP server. The iclim software (see foreground) gives access to the computation of all climate change indices defined by the Expert Team on Climate Change Detection and Indices (iclim also currently includes 9 supplemental climate indices).

The C4I has added a provision of statistical downscaling services, with a tailored and easier interface for the impact communities. It is using the U. Cantabria statistical downscaling portal developed within the previous FP6 ENSEMBLES project, by integrating and accessing downscaling services through a private point of access. The integration of the C4I with the CSAG Climate Information Platform (CIP) has also been investigated, and C4I has benefited from a comprehensive study on climate data portals which helped to prioritize developments. This platform provides observed and statistically downscaled climate data across Africa based largely on point (station) observations. The integration of ESGF/C4I data services into the CIP platform would provide point time series analysis of ESGF datasets through the CIP platform. However, this raises authentication issues that prevent the integration and need further analyses.

These functionalities have been tested in a number of 2-day courses that focussed on the use of climate4impact (D5.2). Feedback from these courses helped to improve the portal, while its successful usage during these courses proved its effectiveness.

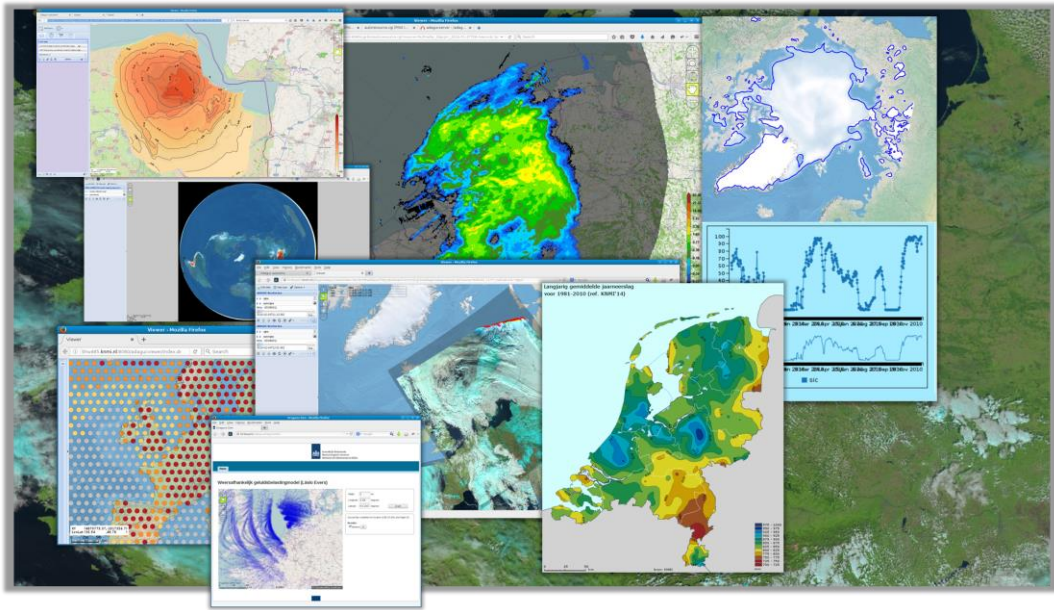


Figure 18: Different functionalities offered by the climate4impact portal (from Page et al., talk at the 6th ESGF F2F meeting, December 2016)

D5.2 – Assessment of impact communities requirements

D5.4 - ENES Data Service Infrastructure requirements

D11.2 – Report on climate4impact, version 1

D11.6 - Report on derived products on climate4impact

4.5. Innovation on data: Linking with climate services and corporates

There is an increasing demand for translating the existing wealth of climate data and information into customised tools, products and information, also known as “**climate services**”. Climate services have the potential to build the bridge between Earth’s climate system model output data, as supported in Europe by the IS-ENES infrastructure, and decision makers, by helping the latter to take informed decisions in order to boost the transition to a climate-resilient and low-carbon society.

Feedback from different workshops and projects (MS6.1 and D6.2) confirms the importance for users to have access to model results concerning climate change projections under different scenarios. It also clearly shows that users are diverse and require a wide range of data, information and support in various forms. The IS-ENES2 expertise on disseminating climate model results through the international ESGF database as well as its expertise in supporting the development of data and metadata standards and the development of tools and guidance for the climate impact communities is a strong asset for IS-ENES2 in the context of the development of the Copernicus Climate Change Service. The IS-ENES2 involvement in the CLIPC Copernicus precursor project has confirmed this.

IS-ENES2 also had the objective to facilitate innovations through the transfer of climate knowledge to SMEs and larger companies providing such climate services. Two entrepreneurship sessions were organized at the 2nd and 3rd ENES summer schools (**MS6.3 and MS6.5**) and a master class developed.

Practically all data for CMIP5, especially all from Europe, are in open access with no restriction for commercial use (**D6.1**). When access for commercial use is requested, this is transferred to CORDEX forced by this climate model. However, this is quite limited since only Japan models are now restricted for non-commercial use. This is a strong advance from CMIP3 where usage was more restricted.

Within IS-ENES2, a master class prototype has been developed aiming at training corporate bodies using CMIP5 and CORDEX data on the application in different decision maker's contexts (**D6.3**). Company representatives, following the master class, enhance their capacity and skills to apply climate model data in a correct way and to assess the possibilities, limitations and uncertainties of data. Climate data providers following the master class benefit through interactions with information users. The master class has an interactive character with dialogues between providers and users on real life decision-making contexts. The exchanges were based on case studies provided by participating companies. Although the first attempt had a limited success, further experiences with international master students and more recently within other projects (such as the Copernicus SWICCA project on water) have proved the efficiency of the concept.

MS6.1 – Report on needs for the EEA Clearinghouse

MS6.3 – Entrepreneurship session at 2nd European school

MS6.5 – Entrepreneurship session at 3rd European school

D6.1 – Report on access rights for CMIP5 and CORDEX for commercial use

D6.2 – Report on needs for climate services

D6.3 – Prototype master classes for SMEs and Corporates

5. Conclusions and perspectives

All the main objectives of IS-ENES2 have been achieved. During IS-ENES2, the European role in international coordinated experiments has been strengthened, and the project has exceeded objectives in providing support for the preparation of CMIP6 and in support of statistical downscaling. The expertise on ESGF is well recognised and IS-ENES2 partners have been delegated the development of ESGF datanodes to access global and regional climate projections within the Copernicus Climate Change Service – this despite having to revise the time-schedule for some activities, particularly those in relation to the ESGF, where a major security breach required significant revision of software and software methodologies (but even then, Europe took a lead in promoting alternative access to the data while the ESGF was unavailable).

The collaboration around high performance computing and related software has paved the way to ESIWACE which will accelerate the preparation of climate models for future exascale architectures (in collaboration with numerical weather prediction community). IS-ENES and ESIWACE now constitute two complementary streams for ENES strategy: IS-ENES focuses on current climate models and their use in WCRP coordinated programmes and to support IPCC Assessment Reports whereas, ESIWACE focuses on future generation climate models.

IS-ENES2 has also strengthened the integration of the climate modelling community, not only within the global and regional modelling communities, but also between researchers and engineers and, via summer schools, within the next generation of researchers. After the two phases of IS-ENES, the need to sustain the research infrastructure cooperation is well expressed by the community in the mid-term update of the strategy. IS-ENES2 ends when CMIP6 is entering its production period. The community has the willingness to continue working together for CMIP6, and is strongly expected to do so by the international community. The HPC and Data Task Forces will help continue the collaboration, however, the lack of dedicated common resources will limit activities, in particular in terms of development and user support.

IS-ENES2 is well integrated in the European landscape of environment research infrastructures, as seen within the ENVRIplus project. It is the only research infrastructure dealing with climate modelling. It concerns the different domains - atmosphere, land and oceans, and needs a good interoperability with all the observation based research infrastructure for model evaluation. Moreover, ENES infrastructure is important for the development of climate services and contributes to the objectives of JPI Climate. The sustainability of IS-ENES is a challenge that will need to be addressed in the coming years.

Impact of results

IS-ENES2 has had an important impact on both global and regional climate modelling in Europe, including, by way of support, on the internationally coordinated CMIP and CORDEX experiments. So, this impact is not limited to Europe. Indeed, European model data are distributed worldwide and IS-ENES2 has played an essential role in supporting the development of the international ESGF database and the ES-DOC metadata standards. IS-ENES2 has also strengthened internal collaboration within the climate modelling community on the different components of its infrastructure: data, models and high-performance computing.

The impact of IS-ENES2 goes beyond the climate modelling community. It also impacts the community working on climate change impacts as well as future computing. IS-ENES2 has been proactive in preparing the 6th Phase of CMIP that will contribute to IPCC 6th Assessment report and on developing an interface with climate services. IS-ENES2 has an important role in innovation, through collaboration with HPC industry but also through societal innovation.

1. Impacts on European community on Earth System Modelling

IS-ENES2 is the infrastructure of the climate modelling community in Europe and therefore its first main impact is clearly on this community.

Serving climate modelling science through international coordinated experiments

The climate research community relies heavily on internationally coordinated experiments organised by the Working Group on Coupled Models (WGCM) of the World Climate Research Program (WCRP). These experiments provide a common reference base for model evaluation, understanding of key processes and projections of possible future climate under different scenarios. IS-ENES in its first and second phases has supported the infrastructure underlying the distribution of model data from these experiments. IS-ENES Phase 1 has supported the global climate modelling experiments of the 5th Phase of the Coupled Model Intercomparison project (CMIP) and IS-ENES2 has supported the additional dissemination of regional climate modelling experiments under the Coordinated Regional Downscaling Experiments (CORDEX) and the preparation of the 6th Phase of CMIP6.

IS-ENES2 has played a strong role in supporting the international database, the Earth System Grid Federation (ESGF), that distributes those data with common standards. IS-ENES2 supports the European datanodes distributing simulations performed in Europe as well as supports the development and maintenance of the ESGF distribution system in collaboration with US colleagues. IS-ENES2 partners are leading or co-leading half of the sixteen working teams of ESGF.

IS-ENES2 has also played an essential role in supporting the further development of metadata standards for CMIP6 under the international Earth System Documentation (ES-DOC) project. It is estimated that about 80% of the support has been provided through work of the IS-ENES2 partners. IS-ENES2 has benefitted from the experience of the previous FP7 METAFOR project.

IS-ENES2 has played a major role in CORDEX. Partners have launched the integration of CORDEX data from regional climate models on ESGF, leading the definition of standards for CORDEX. Most of data available on ESGF are issued from European modelling groups. Having both global and regional climate model data on ESGF strongly eases their usage. Indeed, using the same user interface and standards allows all the tools developed for global models to also be used for regional models. IS-ENES2 has also started to promote ESGF for the second part of CORDEX data, those based on empirical statistical downscaling (ESD).

IS-ENES2 service on model data from CMIP and CORDEX is not limited to Europe. Users of the distributed ESGF database are worldwide, from all the continents. ESGF has over 14 000 users worldwide (37% Asia, 29% North America, 23% Europe, 5% South America, 3% Africa, 3% Oceania). In Europe, users are from 30 countries (including Russia).

Strengthening the integration of the European climate modelling community

IS-ENES1 has started the integration of community working on global models. IS-ENES2 has enlarged this base by integrating also groups working on regional climate modelling, at least main groups involved in

CORDEX. This integration has mainly been focused on the data aspects, although services on models are also of interest to regional climate models.

Following the first ENES infrastructure strategy developed in IS-ENES1 for 2012-2022, IS-ENES2 has organised a mid-term update of the strategy. The researchers and engineers preparing this update clearly expressed their support for a long-term sustained research infrastructure for ENES, although its structure still needs elaboration. This update is an important guide for further collaboration and integration of the community.

IS-ENES2 has strengthened the networking activities among the software engineers working on models and environment tools. Sharing of expertise around tools has been strongly appreciated and a unique opportunity for engineers to meet and exchange. This is hoped to pave the ground for further common developments or sharing of software in the future. For data, coding sprint meetings, where software engineers gather and code in parallel, have been very efficient to perform collaborative work.

IS-ENES2 has continued the work started in IS-ENES1 on capacity building. The second and third schools have confirmed the success of the first summer school launched in IS-ENES1. Training of young researchers is important not only for the gain of expertise on Earth System Models but also to build a collaborative spirit and trust among the future generation of climate modellers. The ENES Portal has been further elaborated as a community information platform.

Fostering common model and tools developments

The service on models has been continued to provide open access and support to key tools (OASIS, CDO) and to the European ocean modelling platform (NEMO). The service has been enhanced to support model access and usage for two European models, showing a need for such a service for countries not developing their own model.

Within CMIP, seven European climate models have provided results for CMIP5. The need for model diversity arises from non-unique representations of the main physical and bio-geophysical processes. However, ENES supports whenever possible shared software. IS-ENES2 partners worked together to develop a common radiative transfer library to be used in Earth System Models, and on common observation-system simulators. Networking activity has also fostered sharing of expertise around environment tool software.

Facilitating access to and optimised use of HPC resources

Access to HPC resources is beyond the scope of IS-ENES2. However, several activities of IS-ENES2 have promoted the optimization of the use of HPC resources by focusing on model performance. In collaboration with US colleagues, a standard set of metrics has been developed to measure model performance; it will be used in CMIP6. The performance of the OASIS coupler, used in six of the seven European coupled climate models, has been further improved in climate models. Parallel IO servers have been implemented in several models, improving the speed of data flow. The possibility to run an ensemble of multi-model multi-member high-resolution ensemble of climate models has been demonstrated, paving the way towards efficient large ensemble experiments on the European HPC facility, PRACE.

IS-ENES2 has also promoted the interaction with PRACE by clarifying specific requirements (storage, stable architectures, need for multi-year projects). However, a dedicated allocation for high-end CMIP6 simulations could not be achieved. Although of interest to Council, large community access did not fit within main objectives and rules of PRACE.

Benchmarks of coupled models have been prepared. They will help interact with vendors and test new systems. They have been presented to vendors.

IS-ENES2 has also promoted and facilitated the development of a Centre of Excellence in Simulation with Weather and Climate ESIWACE, which will have the possibility to go beyond IS-ENES2, aiming at preparing climate models for future exascale computers.

2. Impacts on other research communities

Easing access to model data for the climate impact communities

Users of climate model data are not only from climate research. They also include a wider community working on impacts of climate change for different sectors, e.g. water, agriculture, health, insurance, energy. They are also used by the emerging climate services (see below). These communities are particularly interested by model data downscaled at the regional scale at which impacts are mostly relevant.

IS-ENES1 has developed an interface to ease access to climate model data through the climate4impact portal. The portal has been implemented in IS-ENES2 and functionalities added. The portal has become a platform with possibilities to perform analyses of data. It allows access to data, computations on data such as the computation of widely used climate indices, possibility for the user to downscale data, visualization and guidance. Having both global and regional downscaled results available with the common standard on ESGF, allows using the common climate4impact to access both.

Fostering the interaction between ESM and e-technologies communities

IS-ENES2 has promoted technology tracking and exchange with e-technology experts to better understand how to anticipate future computer architectures. Workshops have been essential to keep track on evolution of technology but also on software developments worldwide. Maintaining climate models ability to run efficiently on rapidly evolving technology is indeed an increasing challenge that will be further addressed in ESIWACE. The HPC Task Force plays a key role of liaison with initiatives on exascale such as the European Exascale Software Initiative and the international Big Data and Extreme-scale Computing initiative.

ENES is also one of the six core communities recognized in the European Data Infrastructure, EUDAT (<http://www.eudat.eu/>), where our community shares its expertise on issues such as data replication, data management and metadata, in strong interaction with activities within IS-ENES2. IS-ENES2 partners are also proactive at participating to the development of the European Open Science Cloud initiative and interact with data cloud initiatives such as INDIGO Datacloud to test technology.

3. Impacts at the international level

IS-ENES2 has further increased the visibility and efficiency of the European contribution to the Earth System Grid Federation. Thanks to the work of IS-ENES2, Europe has become an important contributor to ESGF and the major contributor to the ES-DOC initiative on metadata. ESGF was initially entirely driven by USA but has now become an international collaboration. IS-ENES2 has been engaged in preparing the international governance of ESGF and is involved at both its Steering and Executive levels. IS-ENES2 partners are also strongly involved in the WGCM Infrastructure Panel (WIP) that defines the procedures for CMIP data publication and distribution.

IS-ENES2 has also supported common developments with PCMDI (USA) who is leading the ESGF initiative. The monitoring dashboard developed within IS-ENES2 has been developed in collaboration with PCMDI through visits to PCMDI. Collaboration with other groups such as NCAR and GFDL has also been supported aiming at sharing experience and expertise on issues such as coupler development and model performance metrics.

Collaboration with South Africa has been supported within IS-ENES2, allowing developing the first CORDEX datanode in Africa. This datanode will ease access to ESGF data and contribute to disseminate downscaled data for Africa. IS-ENES2 has also delivered training in Asia to help establish expertise on ESGF and develop their datanodes.

4. Impacts on innovation

IS-ENES2 has focused on two different aspects of innovation: innovation with corporates and societal innovation through the interface with climate services.

Innovation with corporates

IS-ENES2 has promoted interactions with HPC vendors and software companies. This has been done through sessions at HPC workshops and through a series of meetings under Non Disclosure Agreement with individual companies. This was the first time that such meetings have been organised at the European scale,

focusing on vendor strategy and their implications for climate models. This should pave the way for further interactions and possible co-design for the future, in particular within ESIWACE.

Promoting societal innovation through better interfacing with climate services

The need to adapt to climate change raises the challenge of providing society with access to reliable information on climate change. The “Climate Services” community provides such access by transforming climate model data into information tailored for end-users. IS-ENES2 does not provide climate services but has a role to play in the overall chain that go from models to delivery of climate services, as described in the European roadmap for climate services (2015)²¹. The development of the Copernicus Climate Change Service (C3S) during the course of IS-ENES2 has been an important element. IS-ENES2 has contributed to the shaping of the C3S through participation at scoping workshops and through the precursor project CLIPC based on ESGF data access and developments in both IS-ENES1 and IS-ENES2. The expertise of IS-ENES has been recognized and C3S will broker to ESGF to provide access to projections from both global and regional climate models. IS-ENES2 partners are responsible to develop and implement the C3S ESGF datanodes.

IS-ENES2 has also contributed to a better understanding of user needs. This has fed into the climate4impact portal but also into master classes aiming at better informing on how to access climate information. Initially targeting consultancies, the master class has been experimented with success on a range of users of climate model data.

²¹ European roadmap for climate services (2015) :
<http://ec.europa.eu/research/index.cfm?pg=events&eventcode=552E851C-E1C6-AFE7-C9A99A92D4104F7E>

Dissemination and exploitation of results

Dissemination of IS-ENES2 results has first targeted a scientific audience towards an European and International audiences. Some meetings related to HPC and Climate services including corporates. This dissemination has mainly been done through:

- **13 peer-reviewed scientific publications**

Dunlap R. et al., BAMS, 2014, doi:10.1175/BAMS-D-13-00122.1
André J.C. et al., BAMS, 2014, doi: 10.1175/BAMS-D-13-00098.1
Deandreis et al., Climatic Change, 2014, doi: 10.1007/s10584-014-1139-7
Rousset C. et al., Geosci. Model Dev., 2015, doi: 10.5194/gmd-8-2991-2015
Valcke S. et al., Bull. Amer. Meteor. Soc., 2016, doi:10.1175/BAMS-D-15-00239.1
Williams D. et al., BAMS, 2016, doi:10.1175/BAMS-D-15-00132.1
Hewitt, H. et al., Geosci. Model Dev., 2016, doi:10.5194/gmd-9-3655-2016
Roberts, M. J. et al., Geophys. Res. Lett., 2016, doi:10.1002/2016GL070559
Kern, B. and Jöckel, P., Geosci. Model Dev., 2016, doi:10.5194/gmd-9-3639-2016
Eyring, V. et al., Earth System Dynamics, 2016, doi:10.5194/esd-7-813-2016
Balaji, V. et al., Geosci. Model Dev., 2017, doi: 10.5194/gmd-10-19-2017
Webb et al., Geosci. Model Dev., 10, 2017, doi:10.5194/gmd-2016-70
Epico I. et al., International Journal of HPC Applications, 2017, doi: 10.1177/1094342016684930

- **1 paper presented at a conference:**

Manubens-Gil D. et al., HPCS 2016 Conference, Innsbruck, Australia, 18-22 July 2016, doi: 10.1109 / HPCSim.2016.7568429

- **2 Technical reports:**

Valcke S. et al., [OASIS3-MCT User Guide](#), Technical Report URA CERFACS/CNRS No1875, may 2015.
Williams D. et al., [ESGF Implementation Plan v1.0](#), PCMDI, LLNL, USA, May 4, 2016

- **72 Oral presentations at scientific conferences at various international conferences:**

ESGF international face to face annual meetings (30), European Geosciences Union annual meetings (5), other European projects meetings (12), other European and international conferences, workshops, conferences or seminars

- **19 Poster presentations at meetings:**

ESGF annual meetings (7), EGU annual meetings (9), European projects (1), other meetings

- **IS-ENES2 has also organised 32 open conferences and meetings**

Dissemination to larger audiences have also been conducted through:

- **1 large audience Article on IS-ENES2**

S. Joussaume, Climate modelling in support to climate change understanding, Science and Technology 10, March 2014.

- **2 participation to EC policy meetings:**

Joussaume S., HPC for climate, *Strategy Meeting on High Performance Computing (HPC)*, European Commission, April 30th, 2013
Co-organisation of a 1-day meeting on HPC for climate modeling and simulation with DG Connect, 27th February 2014, Brussels, 2014

- **4 Project flyers**

[IS-ENES2 factsheet for the FP7 Climate Change Catalogue](#)

[IS-ENES2 flyer presenting the project](#)

[IS-ENES2, Introducing the European environmental research infrastructures collaborating in ENVRIPUS, booklet](#)

[Autosubmit: a versatile tool to manage weather and climate experiments in diverse supercomputing environments](#)

Exploitable foreground list

IS-ENES2 has developed or contributed to the development of open software available from the web.

Environment tools software

Software library: OASIS3-MCT_3.0 code coupler

OASIS is coupling software developed primarily for the climate community. Implementation and maintenance are managed by Cerfacs and CNRS. It is a portable set of Fortran 77, Fortran 90 and C routines. Low-intrusiveness, portability and flexibility are OASIS key design concepts. The current version of the software, OASIS3-MCT is a coupling library that is compiled and linked to the component models, available at: <ftp.cerfacs.fr/pub/globc/exchanges/distrib-oasis/oasis3-mct.tar.gz>

Cylc workflow-engine (or meta-scheduler)

Cylc (“silk”) is a workflow engine for cycling systems - it orchestrates complex distributed suites of interdependent cycling tasks. IS-ENES2 has improved the pre-existing Cylc software: portability to a wider range of platforms / schedulers; support for climate models (flexible date/time cycling, 360 day calendars; moving functionality from Rose into Cylc; improving efficiency, investigating new approaches to the communication layer in Cylc. Cylc is available at: <https://cylc.github.io/cylc/>

Autosubmit workflow tool for ensemble simulations

Autosubmit is a Python tool further developed into a general-purpose submission and monitoring tool optimized for M4HR ensembles within IS-ENES2. It provides a simple workflow definition capacity that allows running weather, air quality and climate multi-model multi-member experiments in more than one supercomputing platform. Autosubmit is currently being developed at BSC Computational Earth Sciences group. Help about how to install and use Autosubmit and a list of available commands in the online documentation: <http://www.bsc.es/projects/earthscience/autosubmit/>

Configuration Management Best Practice Guide for Climate Science

A configuration management best practice guide for Climate Science has been produced taking contributions from more than 40 individuals from more than 30 institutions. This paper analyses configuration management from the point of view of a number of key roles: model developers, system owners, experiment designers and data consumers. The conclusion is that a wider exploitation of tools to capture the full workflow and to formalize testing and code reviewing would add to the good practices commonly used in the community. Available at: https://portal.enes.org/ISENES2/documents/milestones/is-enes2_ms4-5_configuration-management-best-practice-guide-for-climate-science/view

COSP software to simulate satellite products from climate model variables.

IS-ENES2 has contributed to the optimisation of the current stable version and to the testing of the next major release. The purpose of COSP is to facilitate a consistent comparison of cloud variables from numerical models with observations. COSP variables will be produced in many simulations of CMIP6. It is already routinely used by all the major modelling centres during their model development activities. There is an active community developing new capabilities for COSP. Current research activities include: aerosol signal in the lidar simulator; simulation of ground-based radar; development of radio occultation module. CFMIP webpage: <http://cfmip.metoffice.com/COSP.html>. Code repositories at <https://github.com/CFMIP>

Psrads: software library

The software library provides fast computation of radiative transfer in planetary atmosphere models, as required for weather and climate research in a performance portable manner and with a concise interface.

There are two main audiences for the library. The first are those developing global circulation and weather models, in which most aspects of radiative transfer can be dealt with by the library. The second are scientists interested in investigating radiative transfer, who can use the library as a standalone tool either through its Fortran, C or Python interfaces. The psrad software is open source/access library, not available for commercial purposes. The library will also be part of a climate modelling framework used for research purposes. As such, it will be used in many different applications and be in itself subject to research. The software is currently available as part of the ICON model only, under an open-source-like license for non-commercial purposes from the following link: <https://www.mpimet.mpg.de/en/science/models/icon/>

Computational performance tools

DKRZ-SCT: Performance measurement tool

The SCT (Simple Context Timer) library provides functionality for time and hardware performance counter measurements. The program-parts of interest have to be instrumented manually and recompiled. At runtime the tool can generate performance reports which can include a statistical evaluation over MPI tasks. Measurements are aggregated within a switchable context. SCT allows to evaluate the parallel performance of C or Fortran programs and to identify the parts that become dominant with increasing MPI task count. Available at <https://doc.redmine.dkrz.de/sct/html/>

DKRZ_ICON_COMM_KERNEL: Performance measurement tool for the ICON communication

The purpose of the ICON communication kernel is to simplify further development and analysis of ICON communication. The increased efficiency of working with this kernel results directly from reduced and less complex source code, build process, execution effort and software environment requirements. Within certain limits (e.g., the load imbalance aspect) the kernel should give a good performance estimate of the full model. The communication kernel has been implemented in a way that allows other ICON kernels to use it as software infrastructure. This way the communication functionality is available but the complexity is encapsulated. Available at <https://redmine.dkrz.de/projects/icon-communication-kernel>

CouplingBenchmarks

A standard benchmarking environment for coupling technologies has been set up. It contains four standalone components running on four different grid types and well-defined coupled configurations, or test cases, assembled from the standalone components. Version 1.0.0 of the ENES coupling technology benchmarks contains five coupled configurations running on the regular latitude-longitude grid with different resolutions and using the OASIS3-MCT, OpenPALM, ESMF, MCT-only or YAC coupling technologies. Its purpose is to implement a suite of coupled benchmarks based on simplified model components that capture the essence of the coupling challenges in climate models without the complexities of the science, so to evaluate the performance of different coupling technologies in specific standard configurations. Available from <https://portal.enes.org/computing/benchmarks/coupler-benchmarks>

Data software

FASM

FASM is a set of software packages providing usage and demographics statistics on IS-ENES2/ESGF data nodes. It includes: esgf-dashboard, esgf-dashboard-ui and esgf-stats-api. It was developed by CMCC during the IS-ENES2 project and its current stable implementation has been included into a new release of ESGF software stack to provide federated statistics at ENES and global level. FASM will be installed on the ESGF production data nodes and will support the collection of usage and demographics information at a global scale by providing the users with a simple and usable graphical interface to access the statistics information. New requirements include: addition of sections for the collection of statistics of new project (CMIP6, for example), differentiate statistics on the basis of the type of download the users perform, PerfSONAR statistics, and other minor requirements. The FASM system will support users and node administrators by providing the community with a useful tool for reporting purposes at different levels (i.e. global, by node, by project). Source codes are available at: esgf-dashboard: <https://github.com/ESGF/esgf-dashboard>; esgf-

dashboard-ui: <https://github.com/ESGF/esgf-dashboard-ui>; esgf-stats-api: <https://github.com/ESGF/esgf-stats-api>

QA-DKRZ Quality assurance tool for ESGF model data

A data quality assurance tool has been developed supporting the checking of data with respect to various project specific properties, supporting especially CMIP5, CORDEX and CMIP6 data. The tool has been successfully applied in the CORDEX community to support the data quality assurance before ESGF data publication at different sites in Europe as well as beyond (e.g. Canada). The tool ensures that data from models and data centres conform to CMIP6 regulations and is consistent with e.g. the CF standard (beyond the PrePARE CMIP6 checking module). Additionally a web accessible “spot checker” has been deployed at DKRZ to support end users in testing individual files without the need to install and run the QA-DKRZ (and PrePARE) software packages. Source code: <https://github.com/IS-ENES-Data/QA-DKRZ>; Documentation: <http://qa-dkrz.readthedocs.io/en/latest/>

CF-python

CF-python corresponds to the development of a data model to enable the creation of fully CF-compliant data analysis software, and creation of a reference implementation in the python CF package. The cf-python software library is available at <https://cfpython.bitbucket.io>

Data Request Python API

The Data Request is presented as two XML files whose schema is described in a separate document. A python module is provided to facilitate use of the Data Request. Some users may prefer to work directly with the XML file or with spreadsheets and web page views, but this software provides some support for those who want to use a programming approach. The package provides a library of functions and a command-line interface. It can provide lists of variables filtered according to a range of specifications, and also provide data volume estimates. It is designed to be used by modelling centres contributing to CMIP6.

Available at: <https://earthsystemcog.org/projects/wip/CMIP6DataRequest>

Icclim: Climate indices calculation software

Open-source software to calculate climate indices and indicators has been developed in python. It is called icclim. It is used as a backend to the climate4impact platform and services to enable user to perform on-demand calculations. It has also been used and developed within the FP7-CLIPC project, and by a few external users. Source code at <https://github.com/cerfacs-globc/icclim/>; Documentation at <http://icclim.readthedocs.org>

Synda

Developed since 2011 (IS-ENES1) Synda is a command line tool to search, download and post-process files from the Earth System Grid Federation (ESGF) archive in a highly effective and streamlined way. It is composed of two modules: The synda sdt (synda data transfer) and the synda sdp (synda data processing). Synda will be used for the CMIP6 replication process across the entire ESGF federation (EU/US/AU/ASIA). To date, synda can be used for all projects hosted by ESGF. Source code : <https://github.com/Prodiguer/synda>; Documentation: <http://prodiguer.github.io/synda/>

The esgf-prepare package

The Earth System Grid Federation (ESGF) publication process requires a strong and effective data management, which could also be a burden. The ESGF esgprep toolbox is a piece of software that enables data preparation according to ESGF best practices. esgprep allows the ESGF data providers to easily prepare their data for publishing to an ESGF node. It can be used to fetch required configuration files, apply the Data Reference Syntax on local filesystems or generate mapfiles for ESGF publication. esgprep is designed to follow all requirements from the ESGF Best Practices document. esgprep is built as a full standalone toolbox allowing you to prepare your data outside of an ESGF node.

Source code: <https://github.com/IS-ENES-Data/esgf-prepare>

Documentation: <http://is-enes-data.github.io/esgf-prepare/>

The ES-DOC ecosystem

The Earth System Documentation (ES-DOC) aims to nurture an eco-system of tools & services in support of Earth System documentation creation, analysis and dissemination. Such an eco-system enables the scientific community to better understand & utilize Earth system model data. ES-DOC is coordinated with other community efforts such as CMIP and ESGF via the World Climate Research Programme working group on Climate Modelling (WGCM) and its Infrastructure Panel (WIP). Tools ES-DOC online tools help you to create, search, view & synthesise documents:

Search & View <https://search.es-doc.org>; The ES-DOC search web-application allows a user to quickly search for documentation on a project by project basis. Search results are listed in a tabular format, each row corresponding to a document within the archive. Simply click upon a row to view the document within the ES-DOC documentation viewer.

Online Questionnaire <https://questionnaire.es-doc.org>; The ES-DOC Questionnaire provides a web interface for document creation. Once the Questionnaire has been configured for a particular project, registered users can create, edit, and publish documentation on behalf of that project.

IPython Notebooks: IPython is a highly versatile programming environment that has been adapted by ES-DOC to simplify the process of documenting models. For projects that mandate a set of model description specialisations, ES-DOC can generate & host IPython notebooks with which users can document their models.

Comparator: <https://compare.es-doc.org>; Published model documentation can be compared on a project by project basis. This permits users to perform simple model inter-comparison.

Portals and training

ENES portal

During the IS-ENES2 project, the ENES portal, implemented in IS-ENES1, has been further established as the central information gateway for the European ESM community and as unique access point to the IS-ENES services. Major restructuring during the course of the project has optimized the structure as to make the access of information and services most effective. To ensure the persistence of the ENES portal beyond the IS-ENES2 project, future updates of the contents are distributed, where possible, among community members, whereas large parts of the technical infrastructure are integrated into the standard system and maintenance procedures of DKRZ. Available at: <https://portal.enes.org/>

climate4impact platform

The climate4impact platform is a collection of open services and a tailored intuitive web interface aimed at the climate change impact community end-users and impact modellers. It is also used by graduate and Ph.D. students as well as climate researchers. Features include an intuitive search, fast visualization and download services (as defined by the EU INSPIRE directives). It also includes guidance, use cases, scientific and technical documentation. It is available at: <https://climate4impact.eu/>

climate4impact Services

The climate4impact Services are a collection of operational services accessible through OGC standards and specific APIs. Services include search, access token, combination of datasets, basket requests, WPS requests.

Web site at <https://climate4impact.eu>

Documentation at <https://dev.knmi.nl/projects/impactportal/wiki/API>

Prototype master classes for SMEs and corporates

The master class aims at training SMEs and larger companies using CMIP5 and CORDEX data on the application in different decision maker's contexts, assuming the need for climate data use in such contexts is growing. Company representatives, following the master class, enhance their capacity and skills to apply climate model data in a correct way and to assess the possibilities, limitations and uncertainties of data. Climate data providers following the master class benefit through interactions with information users.

<https://is.enes.org/documents/deliverables/prototype-master-classes-for-smes-and-corporates/view>

Publications on work done within IS-ENES2 by the partners**Reviewed publications on IS-ENES**

- R. Dunlap, M. Vertenstein, S. Valcke, and T. Craig, Second Workshop on Coupling Technologies for Earth System Models, *Bull. Amer. Meteor. Soc.*, 95, ES34–ES38, February 2014, **doi:10.1175/BAMS-D-13-00122.1**
- Jean-Claude André, Giovanni Aloisio, Joachim Biercamp, Reinhard Budich, Sylvie Joussaume, Bryan Lawrence, and Sophie Valcke, High-Performance Computing for Climate Modeling, *Bull. Amer. Meteor. Soc.*, May 2014, **doi: 10.1175/BAMS-D-13-00098.1**
- C. Deandreis, Braconnot P., Pagé C., Bärring L., Bucchignani E., Som de Cerff W., Hutjes R., Joussaume S., Mares C., Planton S., Plieger M., Towards a dedicated impact portal to bridge the gap between the impact and climate communities: lessons from use cases, *Climatic Change*, June 2014, **doi: 10.1007/s10584-014-1139-7**
- Rousset C., M. Vancoppenolle, G. Madec, T. Fichefet, S. Flavoni, A. Barthélemy, R. Benshila, J. Chanut, C. Levy, S. Masson, and F. Vivier: The Louvain-La-Neuve sea ice model LIM3.6: global and regional capabilities. *Geosci. Model Dev.*, 8, 2991–3005, October 2015, **doi: 10.5194/gmd-8-2991-2015**
- Valcke S., A. Craig, R. Dunlap, and G. D. Riley, 2016. Sharing experiences and outlook on Coupling Technologies for Earth System Models, *Bull. Amer. Meteor. Soc.*, March 2016, **doi:10.1175/BAMS-D-15-00239.1**
- Williams D., Balaji V., Cinquini L., Denvil S., Duffy D., Evans B., Ferraro R., Hansen R., Lautenschlager M., Trenham C., A global repository for planet-sized experiments and observations, *Bull. Amer. Meteor. Soc.*, 803–816, May 016, **doi:10.1175/BAMS-D-15-00132.1**
- Hewitt, H. T., M. J. Roberts, P. Hyder, S. E. Belcher, R. Bourdalie-Badie, D. Copsey, A. Coward, C. Guiavarch, C. Harris, R. Hill, J. Hirschi, G. Madec, M. Mizielinski, E. Neininger, A. New, J. Rae, J.-C. Rioual, B. Sinha, D. Storkey, A. Shelly, L. Thorpe, and R. A. Wood, The impact of resolving the Rossby radius at mid-latitudes in the ocean: results from a high-resolution version of the Met Office GC2 coupled model, *Geosci. Model Dev.*, 9, 3655–3670. October 2016, **doi:10.5194/gmd-9-3655-2016**
- Roberts, M. J., H. T. Hewitt, P. Hyder, D. Ferreira, S. A. Josey, M. Mizielinski, and A. Shelly, Impact of ocean resolution on coupled air-sea fluxes and large-scale climate, *Geophys. Res. Lett.*, 43, October 2016, **doi:10.1002/2016GL070559**
- Kern, B. and Jöckel, P.: A diagnostic interface for the ICOSahedral Non-hydrostatic (ICON) modelling framework based on the Modular Earth Submodel System (MESSy v2.50), *Geosci. Model Dev.*, 9, 3639–3654, October 2016, **doi:10.5194/gmd-9-3639-2016**
- Eyring, V., Gleckler P.J., Heinze C., Stouffer R. J., Taylor K. E., Balaji V., Guilyardi E., Joussaume S., Kindermann S., Lawrence B. N., Meehl G. A., Righi M., and Williams D. N, Towards improved and more routine Earth system model evaluation in CMIP, *Earth System Dynamics*, 7, 813–830, November 2016, **doi:10.5194/esd-7-813-2016**
- Balaji, V., Maisonnave, E., Zadeh, N., Lawrence, B. N., Biercamp, J., Fladrich, U., Aloisio, G., Benson, R., Caubel, A., Durachta, J., Foujols, M.-A., Lister, G., Mocavero, S., Underwood, S., and Wright, G., CPMIP: measurements of real computational performance of Earth system models in CMIP6, *Geosci. Model Dev.*, 10, 19–34, January 2017, **doi: 10.5194/gmd-10-19-2017**
- Webb et al., The Cloud Feedback Model Intercomparison Project (CFMIP) contribution to CMIP6, *Geosci. Model Dev.*, 10, 359–384, January 2017, **doi:10.5194/gmd-2016-70**

Epicoco, I., Mocavero, S., Porter, A. R., Pickles, S. M., Ashworth, M., Aloisio, G., Hybridisation strategies and data structures for the NEMO ocean model, *International Journal of High Performance Computing Applications*, January 2017, **doi: 10.1177/1094342016684930**

Technical report

Valcke S., T. Craig and L. Coquart, [OASIS3-MCT User Guide](#), OASIS3-MCT_3.0, TR/CMGC/15/38, *Technical Report* URA CERFACS/CNRS No1875, may 2015.

Paper presented at conference

Manubens-Gil D., J. Vegas-Regidor, C. Prodhomme, O. Mula-Valls and F. J. Doblas-Reyes, Seamless management of ensemble climate prediction experiments on HPC platforms, *HPCS 2016 Conference*, pp. 895-900, Innsbruck, Australia, 18-22 July 2016, **doi: 10.1109 / HPCSim.2016.7568429**

Large audience Article on IS-ENES2

S. Joussaume, Climate modelling in support to climate change understanding, *Science and Technology* 10, March 2014.

Project flyers

[IS-ENES2 factsheet for the FP7 Climate Change Catalogue](#)

[IS-ENES2 flyer presenting the project](#)

[IS-ENES2. Introducing the European environmental research infrastructures collaborating in ENVRIPUS. booklet](#)

[Autosubmit: a versatile tool to manage weather and climate experiments in diverse supercomputing environments](#)

Talks about IS-ENES

Joussaume S., [InfraStructure for the European Network for Earth System modelling « IS-ENES »](#), *European Geosciences Union*, invited talk session NH1.8, April 8th, 2013

Joussaume S., [HPC for climate](#), *Strategy Meeting on High Performance Computing (HPC)*, European Commission, April 30th, 2013

Joussaume S., [InfraStructure for the European Network for Earth System modelling](#), *International Computing in Atmospheric Sciences 2013 (CAS2K13)*, Annecy, France, 9-12 September 2013

Kornblueh L., [Real Applications on Parallel Systems](#), *International Computing in Atmospheric Sciences 2013 (CAS2K13)*, Annecy, France, 9-12 September 2013

Valcke S., [OASIS3-MCT, a coupler for climate modelling](#), *International Computing in Atmospheric Sciences 2013 (CAS2K13)*, Annecy, France, 9-12 September 2013

Lautenschlager M., [CMIP5 Data Management](#), *International Computing in Atmospheric Sciences 2013 (CAS2K13)*, Annecy, France, 9-12 September 2013

Joussaume S., [Climate models: data and computing challenges, invited talk](#), *Joint International Conference on Supercomputing in Nuclear Applications and Monte Carlo*, Paris, France, 31st October 2013

Juckes M., [CORDEX, the Earth System Grid Federation and the European Network for Earth System Modelling](#), WCRP Meeting, Brussels, November 2013

Joussaume S., [InfraStructure for the European Network for Earth System modelling: From « IS-ENES » to IS-ENES2](#), *CLIPC Kick-off meeting*, London, UK, 14-15 January 2014

Joussaume, S., [The European InfraStructure for Earth System modelling](#), *invited seminar at Geophysical Fluid Dynamic Laboratory*, Princeton, USA, 27th January 2014

Joussaume S., [HPC for climate models: Lessons from IS-ENES projects](#), *ECMWF Workshop on Scalability*, Reading UK, April 14-15th 2014

Chris Jack, Update on IS-ENES2, *2nd CORDEX-ESDM Workshop*, Buenos Aires, Brazil, July 30-August 1, 2014

- Page C. et al., [Providing and facilitating climate model data access in Europe: IS-ENES and CLIPC initiatives](#), *EMS/ECAC 2014*, Prague, Czech Republic, 6-10 October 2014
- Foujols M.A., S. Joussaume, Y. Meurdesoif, [Climate modeling : IPSL climate model, HPC challenges, European cooperation and opportunities for collaboration](#), *2nd Joint Laboratory for extreme scale computing meeting*, Chicago, USA, 24-26 November 2014
- Denvil S., [CORDEX and IS-ENES](#), *2014 ESGF F2F*, San Francisco, USA, December 9-11, 2014
- Denvil S., [Modeling groups and Data Center Requirements](#), Sessions Keynote, *2014 ESGF F2F*, San Francisco, USA, December 9-11, 2014
- Dwarakanath P., P. Lundquist, [System for off-line data access](#), *2014 ESGF F2F*, San Francisco, USA, December 9-11, 2014
- Greenslade M., [ES-DOC](#), *2014 ESGF F2F*, San Francisco, USA, December 9-11, 2014
- Kindermann S., S. Wiegel, [Versioning and replication working group](#), *2014 ESGF F2F*, San Francisco, USA, December 9-11, 2014
- Plieger M. et al., [Bridging CMIP5 and CORDEX data infrastructure to impact users](#), *2014 ESGF F2F*, San Francisco, USA, December **9-11**, 2014
- Page C., N. Tatarinova, M. Plieger, W.S. de Cerff, A. Cofiño, M. Vega, [climate4impact: a tailored IS-ENES web portal providing processing and downscaling services through WPS and interfaced with icclim/ocgis](#), *GO-ESSP 2015 Workshop*, Abington UK, 24-25 February 2015
- Joussaume S., [European climate modelling infrastructure: Accessing climate projections](#), *Copernicus C3S Climate Data Store Workshop*, Reading, UK, 3-6 March 2015
- Plieger M. et al., [The climate4impact portal: bridging the CMIP5 and CORDEX data infrastructure to impact users](#), *ESSI2.1, EGU 2015*, Vienna, Austria, April 17th 2015
- Joussaume S., [Access to CMIP climate projections: strengths, weaknesses & perspectives](#), *Copernicus C3S Projection Workshop*, Reading UK, 20-21 April 2015
- Page C., Climate4impact in IS-ENES2, CLIPC 2nd General Assembly, Dortmund, Germany, 2-3 June 2015
- Joussaume S., [Infrastructure for the European network for Earth system modelling](#), *SeRC 6th Annual Meeting*, Stockholm, June 15th 2015
- Page C. et al., [A user friendly portal to give access to climate scenarios and prediction data](#), *SPECS Fourth General Assembly*, Norrköping, Sweden, 15-17 September 2015
- Joussaume S., [Data and computing infrastructures for climate modelling](#), *Research Data Alliance 6th Plenary meeting*, Paris, September 22, 2015
- Joussaume S., [HPC for CMIP6 and beyond](#), *PRIMAVERA-CRESCENDO Workshop*, Exeter, UK, 24-25 November 2015
- Dwarakanath P., N. Carenton, [ESGF Installation working team](#), *2015 ESGF F2F*, Monterey, USA, December 8-11, 2015
- Dwarakanath P., [Software security working team](#), *2015 ESGF F2F*, Monterey, USA, December 8-11, 2015
- Greenslade M. et al., [Earth system documentation \(ES-DOC\) preparations for CMIP6](#), *2015 ESGF F2F*, Monterey, USA, December 8-11, 2015
- Kershaw P. et al., [Identity entitlement access team](#), *2015 ESGF F2F*, Monterey, USA, December **8-11**, 2015
- Kindermann S. et al., [replication and versioning working team](#), *2015 ESGF F2F*, Monterey, USA, December **8-11**, 2015
- Levasseur G. et al., [CMIP6 errata as a new ESGF service](#), *2015 ESGF F2F*, Monterey, USA, December **8-11**, 2015
- Nassissi P., S. Fiore, G. Aloisio, [ESGF dashboard working team](#), *2015 ESGF F2F*, Monterey, USA, December **8-11**, 2015
- Rathmann T., [User support working team](#), *2015 ESGF F2F*, Monterey, USA, December 8-11, 2015,
- Joussaume S., [Earth system complexity](#), *Advancing the climate modelling enterprise in Europe*, Chicheley Hall, Newport Pagnell, UK, January 11-12 2016
- Page C., S. Joussaume, M. N. Juckes, W. S. de Cerff, M. Plieger, L. Barring, M. Kolax, R. Hutjes, F. de Jong, E. Bucchignani, N. Tatarinova, A. Cofiño, and E. de Vreede, [climate4impact: Providing, tailoring and facilitating climate model data access](#), *AMS 2016*, New Orleans, LA, USA, 11-14 January 2016
- Guglielmo F., et al., [ENES the European Network for Earth System modeling and its infrastructure project IS-ENES](#), *EGU 2016*, Vienna, Austria, 17-22 April 2016
- Valcke S., [Code coupling for climate modeling](#), *CEMRACS 2016 Summer School*, Marseille, France, 18th July – 26th August 2016

- Joussaume S., [IS-ENES](#), *Seminar at PCMDI*, Livermore, USA, 30 August 2016
- Valcke S., [OASIS3-MCT, a coupling software for climate modeling](#), *International Fall School on Terrestrial Modeling and High-Performance Scientific Computing*, Bonn, Germany, 10-14 October 2016
- Joussaume S., [Linking climate modeling research infrastructure and climate services](#), *ClimatEurope Webinar*, 16 November 2016
- Joussaume S., Guglielmo F., [IS-ENES](#), *3rd ENVRI week session on linking research infrastructures with Copernicus*, Prague, Czech Republic, 16 November 2016
- Guglielmo F., Joussaume S., [IS-ENES](#), *3rd ENVRI week PICO session on research infrastructures*, Prague, Czech Republic, 14-18 November 2016
- Joussaume S., [Data model intercomparison projects: the experience from climate experiments CMIP](#), *Belmont Forum workshop on e-infrastructure exemplars*, Paris, France, 28-29 November 2016
- Denvil S., [Earth System Grid Federation](#), *Belmont Forum workshop on e-infrastructure exemplars*, Paris, France, 28-29 November 2016
- Joussaume S., [European Network for Earth System modeling Data Infrastructure](#), *ESGF F2F meeting*, Washington DC, USA, 5-9 December 2016
- Rathmann T., [ESGF support working team – Progress update & future roadmap](#), *ESGF F2F meeting*, Washington DC, USA, 5-9 December 2016
- Hollweg H-D., [QA-DKRZ: the Annotation Model](#), *ESGF F2F meeting*, Washington DC, USA, 5-9 December 2016
- Ben Nasser A., Greenslade M., [ES-DOC and ES-DOC Services](#), *ESGF F2F meeting*, Washington DC, USA, 5-9 December 2016
- Plieger M. et al., [Recent Climate4impact Developments: Provenance in Processing and Connection to the CLIPC Portal](#), *ESGF F2F meeting*, Washington DC, USA, 5-9 December 2016
- Pagé, C., Som de Cerff W., Plieger M., Vega M., Cofiño A., Barring L., de Jong F., Hutjes R., Fiore S., [Server-side computing services provided by IS-ENES through the climate4impact platform](#), *ESGF F2F meeting*, Washington DC, USA, 5-9 December 2016
- Nuzzo A., [Stats and Dashboard Working Team](#), *ESGF F2F meeting*, Washington DC, USA, 5-9 December 2016
- Ananthakrishnan R., Kershaw P. et al., [Identity Entitlement Access Management Working Team](#), *ESGF F2F meeting*, Washington DC, USA, 5-9 December 2016
- Levavasseur G., Bennasser A., Greenslade M., Denvil S., [ESGF Errata Service for CMIP6](#), *ESGF F2F meeting*, Washington DC, USA, 5-9 December 2016
- Stockhause M., Berger K., Levavasseur G., Lautenschlager M., [Data Citation Service for CMIP6](#), *ESGF F2F meeting*, Washington DC, USA, 5-9 December 2016
- Dwarakanath P., Carenton N., [ESGF Installation Working Team](#), *ESGF F2F meeting*, Washington DC, USA, 5-9 December 2016
- Kindermann S. et al., [Replication and Versioning Working Team](#), *ESGF F2F meeting*, Washington DC, USA, 5-9 December 2016
- Weigel T. et al., [Persistent Identifier Services](#), *ESGF F2F meeting*, Washington DC, USA, 5-9 December 2016
- Page C., Som De Cerff W., Plieger M., Vega M., Cofino A. S., Barring L., De Jong F., Hutjes R., Fiore S., [Server-side Computing Services provided by IS-ENES through the climate4impact Platform](#), *ESGF F2F meeting*, Washington DC, USA, 5-9 December 2016
- Denvil S., Lautenschlager M., [CMIP6 ESGF Tier 1 and Tier 2 Nodes](#), *ESGF F2F meeting*, Washington DC, USA, 5-9 December 2016
- Denvil S., [Synda \(synchro-data\)](#), *ESGF F2F meeting*, Washington DC, USA, 5-9 December 2016
- Valcke, S., [First results from the IS-ENES2 coupling technology benchmark](#), *Fourth Workshop on Coupling Technologies for Earth System Models*, Princeton, USA, 20-22 March 2017
- Mocavero S., [Towards KNL the NEMO ocean model case study](#), *KNL Workshop*, CINECA, Italy, 22-23 March 2017
- Joussaume S., [European Infrastructure for climate modeling](#), *First ClimatEurope Festival*, Valencia, Spain, 4-7 April 2017
- Parinet M., Guglielmo F., Joussaume S., [IS-ENES2 Project Management, Lessons Learned](#), *EGU 2017*, Vienna, Austria
- Joussaume S., Guglielmo F., [IS-ENES2](#), *ENVRIplus Booth*, *EGU 2017*, Vienna, Austria, 23–28 April 2017

- Joussaume S., [Climate modelling and HPC](#), *International workshop on HPC for Computational Fluid Dynamics*, Toulouse, France, 12 May 2017
- Guilyardi E., [Which community governance of climate model evaluation?](#), *H2020 and FP7 workshop on climate modelling*, Brussels, Belgium, 23-24 May 2017
- Joussaume S., Guglielmo F., [IS-ENES: Overview of the main aims, results, findings](#), *H2020 and FP7 workshop on climate modelling*, Brussels, Belgium, 23-24 May 2017
- Joussaume S., Guglielmo F., [IS-ENES: Lessons learned/recommendations/plans in the context of model evaluation](#), *H2020 and FP7 workshop on climate modelling*, Brussels, Belgium, 23-24 May 2017

Posters

- Berger K., G. Levvasseur, M. Stockhause, M. Lautenschlager, [Integration of external metadata into the Earth System Grid Federation \(ESGF\)](#), *EGU 2015*, Vienna, Austria, April 2015
- Plieger M. et al., [Application of open source standards and technologies in the climate4impact portal](#), *PICO Spot 1, EGU 2015*, Vienna, Austria, April 2015
- Som de Cerff W. et al. [The IS-ENES climate4impact portal: bridging the CMIP5 and CORDEX data to impact users](#), *EGU 2015*, Vienna, Austria, April 2015
- Torsten R., F. Toussaint, S. Kindermann, M. Lautenschlager, [Support to Climate Data Dissemination in a Worldwide Data Federation by IS-ENES2](#), *EGU 2015*, Vienna, Austria, April 2015
- Manubens D., J. Vegas, F. Doblas-Reyes, [Autosubmit: a versatile tool to manage weather and climate experiments in diverse supercomputing environments](#), *EC-Earth meeting*, Reading, UK, May 5th 2015
- Wilson S., [An Investigation of the Timestep of three Atmospheric Models](#), *NCAS Science Conference*, York, UK, 16-17 July 2015
- Guglielmo F., Joussaume S., [IS-ENES](#), *1st ENVRI week*, Prague, Czech Republic, November 16-20, 2015
- Plieger M. et al., [Climate4impact portal](#), *2015 ESGF F2F*, Monterey, USA, December 8-11, 2015
- Som de Cerff W. et al, *Toward a high-performance data analysis platform for impact analysis*, *ESGF F2F meeting*, Washington DC, Washington DC, 5-9 December 2016
- Plieger M. et al, *The Climate4Impact Portal*, *ESGF F2F meeting*, Washington DC, USA, 5-9 December 2016
- Greenslade M. et al., *ES-DOC and ES-DOC Services*, *ESGF F2F meeting*, Washington DC, Washington DC, 5-9 December 2016
- Plieger M., de Vreede E., *ADAGUC open source visualization in climate4impact using OGC standards*, *ESGF F2F meeting*, Washington DC, 5-9 December 2016
- Stockhause M., *Data Citation Service*, *ESGF F2F meeting*, Washington DC, 5-9 December 2016,
- Levvasseur G., *CMIP6 Errata as a New ESGF Service*, *ESGF F2F meeting*, Washington DC, 5-9 December 2016
- Xavier Yepes et al., [Simulation-based performance analysis of EC-Earth 3.2.0 using Dimemas](#), *EGU 2017*, Vienna, Austria, 23–28 April 2017
- Greenslade M., *Documenting CMIP6 with ES-DOC*, *EGU 2017*, Vienna, Austria, 23–28 April 2017, <http://meetingorganizer.copernicus.org/EGU2017/EGU2017-14769.pdf>
- Page et al., *The climate4impact platform: Providing, tailoring and facilitating climate model data access*, *EGU 2017*, Vienna, Austria, 23–28 April 2017, <http://meetingorganizer.copernicus.org/EGU2017/EGU2017-18727.pdf>
- Stockhause M. and Lautenschlager M., *CMIP6 Citation Services and the Data Services of the IPCC Data Distribution Centre for AR6*, *EGU 2017*, Vienna, Austria, 23–28 April 2017, <http://meetingorganizer.copernicus.org/EGU2017/EGU2017-4325.pdf>
- Parinet M., Guglielmo F., Joussaume S., [IS-ENES project management, lessons learned](#), *EGU 2017*, Vienna, Austria, 23–28 April 2017

Opened workshops

- May 28-30th 2013, IS-ENES2 Kick-off meeting, Meudon, France
- September 23-24th 2013, Configuration Management Tools Workshop, Exeter, UK
- January 21-22nd 2014, Metadata Generation Workshop, Hamburg, Germany
- February 12-13th 2014, Workshop on coupling technologies characterization, Exeter, UK
- March 17-19th 2014, Exascale Technologies & Innovation in HPC for Climate Models - 3rd HPC Workshop, Hamburg, Germany
- May 12-13th 2014, Future perspectives for scientific and infrastructural needs in Earth system model evaluation, De Bilt, Netherlands
- May 21st-23rd 2014, Training on OASIS3-MCT, Toulouse, France
- June 3rd-5th 2014, Workflow Solutions workshop, Hamburg, Germany
- June 9-20th 2014, 2nd European Earth System and Climate Modelling school, Barcelona, Spain
- June 11-13th 2014, *IS-ENES2 First General Assembly*, Barcelona, Spain
- November 17-19th 2014, *IS-ENES2-CLIPC joint workshop on Design of scientific portals*, De Bilt, Netherlands
- March 18-20th 2015, *Master class on the use of climate services*, following the prototype design piloted at WU with 48 Academic Consultancy Training students, Wageningen, Netherlands
- April 20-22nd 2015, *3rd Workshop on Coupling Technologies for Earth System Models*, Manchester, UK
- February 2016, *Workshop on Configuration Management in Climate Science* - various virtual meetings through February, Virtual
- February 10-12th 2016, *CLIPC/IS-ENES2 joint workshop on Common CVs for climate indices*, Toulouse, France
- February 16-18th 2016, *IS-ENES2 Second General Assembly*, Hamburg, Germany
- March 08-11th 2016, *IS-ENES2 ES-DOC 2016 Workshop*, Abingdon, UK
- March 18-24th 2016, *Master class on the use of climate services* - Academic Consultancy Training, Wageningen, Netherlands
- April 6-7th 2016, *4th ENES HPC Workshop*, Toulouse, France
- June 9-21st 2016, *Third European Earth System and Climate Modeling School*, Helsinki, Finland
- September 27-29th 2016, *Final workshop on meta-data generation during experiments*, Lisbon, Portugal
- September 27-29th 2016, *Final workshop on workflow solutions*, Lisbon, Portugal
- October 24-25th 2016, *Crossing the Chasm: Towards common Software Infrastructure for Earth System Model development*, Reading, UK
- October 25-27th 2016, *ENES community meeting on Earth System Modelling Infrastructure Strategy*, Reading, UK
- November 15-16th 2016, *2nd Technical ESMValTool Workshop*, Munich, Germany
- November 30th 2016 to January 12th 2017, *Workshop on innovation in HPC for climate models* (several sessions), Hamburg, Germany
- January 16-18th 2017, *IS-ENES2 Final General Assembly*, Paris, France
- January 19th 2017, *Workshop on NEMO in CMIP6*, Grenoble, France
- January 31st 2017 to February 2nd 2017, *IS-ENES2 NEMO Sea Ice working group: designing a strategy for sea ice development in NEMO*, Paris, France
- February 20-22nd 2017, *Workshop on COSP Network Support and Optimisation*, Paris, France
- March 14-15th 2017, *3rd IS-ENES2 Workshop on Metadata for Climate Indices*, De Bilt, Netherlands
- March 20-22nd 2017, *Fourth Workshop on Coupling Technologies for Earth System Models*, Princeton, USA

Acronyms

AWI: Alfred Wegener Institute for Polar and Marine Research, (www.awi.de)

CDO: Climate Data Operators (<http://www.mpimet.mpg.de/cdo>) - collection of about 100 functions developed by the MPG for handling and analyzing data produced by a variety of climate and NWP models - e.g. for file operations, simple statistics, or the calculation of climate indices. The code is used by around 150 groups (220 users) world-wide, including some of the project partners, calling the CDO around 200000 times per day.

CF: Climate and Forecast Metadata Convention (<http://cf-pcmdi.llnl.gov/>) - International standard for model data files format.

CIM: Common Information Model - The FP7 METAFOR project has developed this standard.

CLM: Climate Limited-area Modelling-Community, (www.clm-community.eu)

CLIMATE4IMPACT: ENES Portal for Climate Impact Communities (<http://climate4impact.eu>) developed within IS-ENES and IS-ENES2 to ease access to model data for the climate impact research communities

CMIP5/6: Coupled Model Intercomparison Project Phase 5/6, under the auspices of WCRP to prepare IPCC AR5/6 (<http://cmip-pcmdi.llnl.gov/cmip5/>)

COPERNICUS C3S: European initiative for the implementation of information services dealing with environment and security (previous Global Monitoring for Environment and Security) (<http://www.copernicus.eu>). C3S is the Copernicus Climate Change Service.

CORDEX: “Coordinated Regional downscaling Experiments” under WCRP auspices (http://wcrp.ipsl.jussieu.fr/SF_RCD_CORDEX.html).

CORDEX-ESD: stands for the regional data obtained from Experimental Statistical Downscaling http://www.cordex.org/index.php?option=com_content&view=article&id=222&Itemid=714

COSP: Cloud Observing System Package (<http://cfmip.metoffice.com/COSP.html>) has been developed to simulate cloud satellite observed parameters from climate model data in order to allow a more direct comparison with satellite data. It is available from MetO.

DYNAMICO: (<http://www.lmd.polytechnique.fr/~dubos/DYNAMICO/goals.html>) is a new dynamical core developed at IPSL

EEA: European Environment Agency (<http://www.eea.europa.eu>)

ENES: European Network for Earth System Modelling (<http://www.enes.org>) - A consortium of European institutions aiming at helping the development of use of ESMs for climate and Earth System studies.

EPM: IS-ENES European Project Manager

ES-DOC: Earth System Documentation (<http://es-doc.org>) is an international collaboration led by US on metadata standard for climate models.

ESGF: Earth System Grid Federation (<http://www.earthsystemgrid.org/>) is an international collaboration with a current focus on serving the WCRP CMIP project and supporting climate and environmental science in general. The ESGF grew out of the larger GO-ESSP community.

ESM(s): Earth System Model(s). These models are developed to simulate the climate system in its full complexity, i.e. atmosphere, ocean and land which are the basic components included in climate models together with biogeochemical cycles, i.e., carbon cycle, vegetation, aerosol and chemistry processes.

ESMF: Earth System Modelling Framework (<http://www.esmf.ucar.edu/>)- Devoted to define standards in the designing of climate model components for easier exchange and coupling, US-led

EUDAT: FP7 project “European Data Infrastructure” (<http://www.eudat.eu/>) launched on Oct 1st 2011, aims at providing a pan-European solution to the challenge of data proliferation in Europe's scientific and research communities. MPG and Cerfacs represent ENES in this consortium

Exascale: for exascale computing, refers to computing power corresponding to 10^{18} operations per second. It is thousand times more powerful than present top computing facilities. It is expected to be available around 2018.

GERICS: Climate Service Center Germany, (www.climate-service-center.de)

HPC: High Performance Computing

I/O: Input/Output is the generic process of exchanging data during a simulation, either as input to the model or as output of model simulations

ICON: is a new dynamical core project run by MPG and Deutscher Wetter Dienst based on icosahedral grids for atmosphere and ocean global models (<http://icon.enes.org/>)

ICT: Information & Communication Technology

ICTP: Abdus Salam International Centre for Theoretical Physics, (<https://www.ictp.it>)

IPCC: Intergovernmental Panel on Climate Change (<http://www.ipcc.ch>) - Provides regular scientific assessments reports (AR) on climate change issue under the auspices of UNEP and ICSU. The last one is the AR4 produced in 2007; the next one is AR5 to be issued in 2013

IS-ENES (or IS-ENES1): InfraStructure for the European Network for Earth System Modelling, first phase; **FP7 project** (<http://is.enes.org>)

LIM: Louvain-la-Neuve sea ice model (http://www.astr.ucl.ac.be/index.php?page=LIM_Description) – A Sea Ice Model

MCT: The Model Coupling Toolkit is a US set of open source software tools to create coupled models (<http://www.mcs.anl.gov/research/projects/mct/>)

METAFOR: Common Metadata for Climate Modelling Digital repositories (<http://ncas-cms.nerc.ac.uk/METAFOR/>) - FP7 infrastructure project under ENES, which focuses on developing common standards for data and model information exchange that will be implemented in IS-ENES.

MGO: Voeikov Main Geophysical Observatory, St. Petersburg, Russia

NCAR: National Center for Atmospheric Research in Boulder, USA (<http://www.ncar.ucar.edu/>)

NEMO: Nucleus for European Modelling of the Ocean (<http://www.locean-ipsl.upmc.fr/NEMO/>) - State-of-the-art modelling framework including 3 components: an ocean general circulation model (OPA), a sea-ice model (LIM) and a biogeochemistry model (TOP); NEMO is interfaced with all European atmospheric models via the OASIS coupler.

netCDF: network Common Data Form (<http://www.unidata.ucar.edu/software/netcdf/>) - A set of software libraries and machine-independent data formats that support the creation, access, and sharing of array-oriented scientific data

NOAA: National Oceanographic and Atmospheric Organisation, USA (<http://www.noaa.gov/>)

LFRic: name given to the UK programme of work intended to deliver a replacement for the Unified Model (<https://puma.nerc.ac.uk/trac/GungHo/wiki/LFRic>)

OASIS: Ocean Atmosphere Sea Ice and Soil coupler (<http://www.cerfacs.fr/globc/software/oasis/>) – A software component allowing synchronized exchanges of coupling information between numerical codes representing different components of the climate system.

Obs4MIP: (<https://www.earthsystemcog.org/projects/obs4mips/>) Observations for Model Intercomparisons Project is an activity to make observational products more accessible on ESGF for climate model intercomparisons.

OpenPALM: is a dynamic parallel code coupler developed at CERFACS (http://www.cerfacs.fr/globc/PALM_WEB/)

PCMDI: Program for Climate Model diagnosis and Intercomparison, (<http://www.pcmdi.llnl.gov/>) at Lawrence Livermore National Laboratory (USA) has the responsibility for supporting modelling studies CMIP5.

PRACE: Partnership for Advanced Computing in Europe (<http://www.prace-project.eu/>) - An FP7 infrastructure project devoted to prepare the implementation of world-class high-performance computers in Europe.

RMIB, Royal Meteorological Institute, Brussels, Belgium

SPECS: new Environment FP7 project “Seasonal-to-decadal climate Prediction for the improvement of European Climate Services » (2012-2016)

ULg : University of Liège; Liège, Belgium (MAR model)

UNSW : University of New South Wales, Sidney, AU

WCRP: World Climate Research Programme (<http://www.wmo.ch/pages/prog/wcrp>)

WIP: WGCM Infrastructure Panel (<https://www.earthsystemcog.org/projects/wip/>)

WGCM: Working Group on Coupled Models (<http://www.clivar.org/organization/wgcm/wgcm.php>) - Under WCRP defines the international strategy for climate model evaluation and simulations for IPCC reports

XML: Extensible Markup Language

YAC: YetAnotherCoupler is a coupler software developed at MPG (<https://doc.redmine.dkrz.de/YAC/html/>)