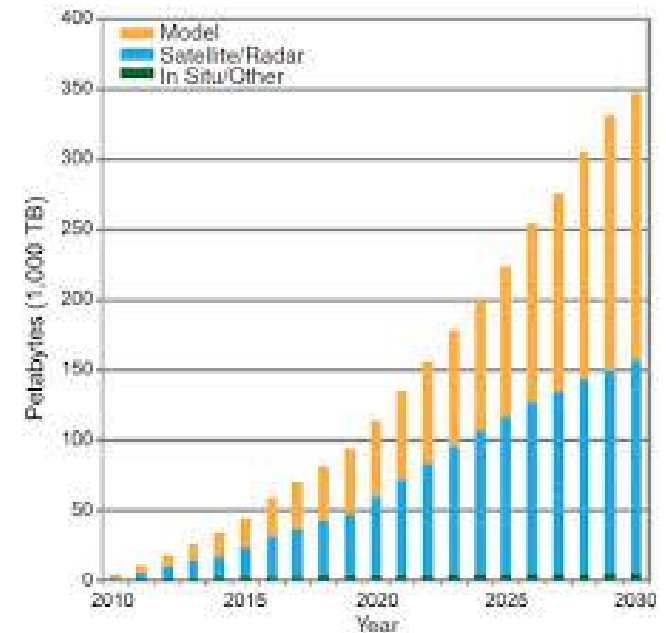


Support to scientific research on seasonal-to-decadal climate and air quality modelling

User Forum

3-4 February 2016, Rome, Italy

- Cope with exponentially growth of volume because of:
 - increase of spatial and temporal resolutions
 - sources of data: observations, model outputs, reanalysis, sensors,...



Air Quality Model (NMMB/BSC-CTM)

	Horizontal resolution (grid cell size)	Output size of a one year simulation, global fields (including meteorology, aerosols and gas phase chemistry)
Standard Resolution	10 km	2.3 Pb
High Resolution	4 km	9.1 Pb
Ultra High Resolution	1 km	36.5 Pb

Climate Model (EC-EARTH)

	Horizontal resolution (atmosphere/ocean)	Output sizes of one year monthly simulation
Standard Resolution	T255/ORCA1 60km/100km	26 GB
High Resolution	T511/ORCA025 40km/25km	120 GB
Ultra High Resolution	T1279/ORCA012 25km/12km	1TB

Increase of number of centers involved in Model Intercomparison Projects (MIP)

	CMIP	CMIP2	CMIP3	CMIP5
Number of experiments	1	2	12	110
Centers participating	16	18	15	24
Number of models	19	24	21	45
Total dataset size	1GB	540GB	36TB	3.3PB

Potential consequences without a solution



- Having data “stuck” locally and impossible to share among institutions
- Data repositories too big to be indexed/explored
- Too much memory demanding softwares to calculate physical diagnostics

Why EUDAT services



- Need for new tools for data management to tackle weather, climate and air quality issues
- Sharing data knowledge with other communities



Specific benefits expected from EUDAT services



- Increase velocity and efficiency (B2SHARE) in data transfers
- Develop “ESGF type” innovative solutions for data indexing and discovery (B2FIND)
- “Bring the compute to the data”: improve the global data workflow (pre-processing -> HPC simulation -> transfer -> post-processing), using the computation power where it is (B2STAGE)

A screenshot of the Earth System Grid Federation (ESGF) search interface. The page features the ESGF logo at the top left, the PCMDI logo at the top right, and a navigation bar with links for Home, Search, Tools, Login, and Help. Below the navigation bar, there is a search input field with a magnifying glass icon and a "Search" button. To the right of the search field, there are links for "Temporal Search", "Clear search constraints and datacart", "Search Help", "Search Controlled", and "Vocabulary". Below the search field, there are checkboxes for "Search All Sites", "Show All Replicas", and "Show All Versions". A pagination bar shows "< 1 2 3 ... 816 817 >" and "displaying 1 to 10 of 8162 search results". Below the pagination bar, there is a "Display" dropdown menu set to "10" and "datasets per page", with links for "Add All Displayed to Datacart" and "Remove All Displayed from Datacart". On the left side, there is a "Current Selections" section with links for "remove all", "(x) project:specs", and "(x) experiment family:Decadal". Below that is a "Search Categories" section with a table listing various categories and their counts. On the right side, there is a "Results" section with a "Data Cart" tab. The results list includes several entries for "specs_output.IPSL_IPSL-CM5A-LR.decadal.S19610101.day.atmos.day.clt.r1i1p1" with "Data Node: esgf-data1.ceda.ac.uk" and "Version: 20141013". Each entry also includes "No description available." and "Further options: Add To Cart".

ESGF
Earth System Grid Federation

Home Search Tools Login Help

Current Selections
[remove all](#)
[\(x\) project:specs](#)
[\(x\) experiment family:Decadal](#)

Search Categories

Category	Count
Project	
Institute	
IPSL (8162)	
Model	
IPSL-CM5A-LR (8162)	
SubModel	
Instrument	
Experiment Family	
Experiment	
decadal (8162)	
SubExperiment	
Time Frequency	
Product	
Realm	
Variable	
al (159)	
clt (318)	
goro (53)	
hfls (159)	

Search

Examples: temperature, "surface temperature", climate AND project:CMIP5 AND variable:hus.
To download data: add datasets to your Data Cart, then click on *Expand* or *wget*.

Search All Sites Show All Replicas Show All Versions

< 1 2 3 ... 816 817 > displaying 1 to 10 of 8162 search results

Display 10 datasets per page

[Add All Displayed to Datacart](#) [Remove All Displayed from Datacart](#)

Results Data Cart

[specs_output.IPSL_IPSL-CM5A-LR.decadal.S19610101.day.atmos.day.clt.r1i1p1](#)
Data Node: esgf-data1.ceda.ac.uk
Version: 20141013
No description available.
Further options: [Add To Cart](#)

[specs_output.IPSL_IPSL-CM5A-LR.decadal.S19610101.day.atmos.day.clt.r2i1p1](#)
Data Node: esgf-data1.ceda.ac.uk
Version: 20141013
No description available.
Further options: [Add To Cart](#)

[specs_output.IPSL_IPSL-CM5A-LR.decadal.S19610101.day.atmos.day.clt.r3i1p1](#)
Data Node: esgf-data1.ceda.ac.uk
Version: 20141013
No description available.
Further options: [Add To Cart](#)

[specs_output.IPSL_IPSL-CM5A-LR.decadal.S19610101.day.atmos.day.pr.r1i1p1](#)
Data Node: esgf-data1.ceda.ac.uk
Version: 20141013
No description available.
Further options: [Add To Cart](#)

[specs_output.IPSL_IPSL-CM5A-LR.decadal.S19610101.day.atmos.day.pr.r2i1p1](#)

- Interest of our pilot to other communities in our scientific domain:
 - The data transfer and replica issues found in the Earth Sciences community are very common to many communities sharing data. Even if the indexing and file organization are very specific to the community (variables, models, start dates,...), solutions could be easily extrapolated to other types of files.
- Further expected support from EUDAT solving the problem/s we may envisage in the future:
 - The development of generic tools for data transfer and staging and research for cross-communities tools by EUDAT can allow to think “out of the box” and import in our specific domain innovative solutions



INSTITUT CATALÀ
DE CIÈNCIES DEL
CLIMA



**Barcelona
Supercomputing
Center**

*Centro Nacional
de Supercomputación*



Thank you!

User Forum

3-4 February 2016, Rome, Italy

Pierre-Antoine Bretonnière - Francesco Benincasa
IC3-BSC - Spain

pierre-antoine.bretonniere@ic3.cat francesco.benincasa@bsc.es