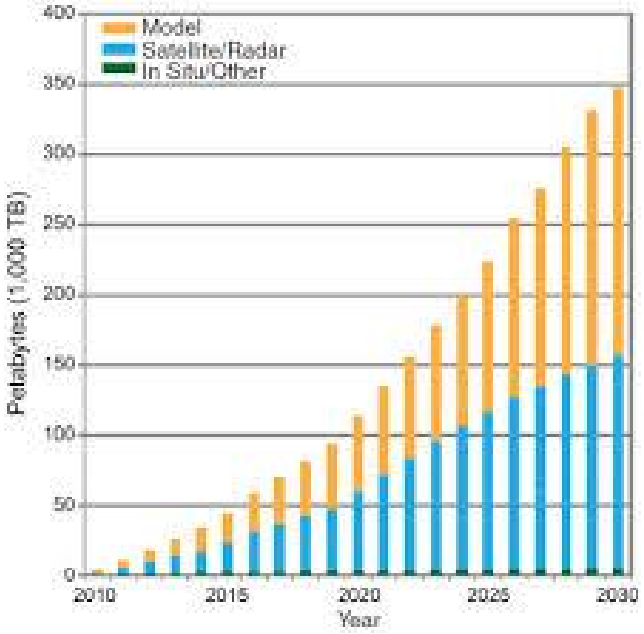
# Support to scientific research on seasonal-to-decadal climate and air quality modelling

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

# Scientific challenges

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



# Scientific challenges

## 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<br>60km/100km | 26 GB |
| **High Resolution** | T511/ORCA025<br>40km/25km | 120 GB |
| **Ultra High Resolution** | T1279/ORCA012<br>25km/12km | 1TB |

# Scientific challenges

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)

# Expected future impact

- 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

EUDAT

# 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