



---

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

This pilot aims to “better simulate” climate change, at seasonal to decadal time scale and forecast air quality using both existing and locally developed models EC-Earth (global circulation model, GCM) and NMMB-BSC (air quality model). By “better simulate”, we mean making a better use of the huge amount of raw data generated by these models. That includes data transfer between the different research institutes using the data that are disseminated all over the world, but also curation, and data discovery on portals where different projects store their data.

### The Scientific Challenge

In the latest version of the climate models, that will be used in the next Coupled Model Inter-comparison Project (CMIP6) between other projects, the resolution used has increased up to 25km in the ocean with 75 vertical levels and 40km in the atmosphere (T511/ ORCA025) and the trend is to go to T1279/ORCA012, doubling the resolution.

The time frequency at which the outputs are saved is also increasing and the size of the outputs consequently explodes: for example, one year of a typical experiment can occupy 1TB, knowing that in a climate experiment, hundreds of years are simulated for each experiment. Once this raw data is produced by a local institute, it needs to be shared among the whole community. We can estimate that a community of several hundreds of scientists disseminated in more than 30 research institutes around the world will use this raw data. The sharing and (multiple) transfers of such an amount of data is one of the first technical obstacles we have to cope with.

The other technical challenge is how to get meaningful information from this huge amount of data for climate scientists but also downstream communities such as health (impact of climate and aerosols on health) and climate services (renewable energy industry, policy makers). Simple diagnostics such as time means or calculations of indices along the time series can become almost impossible (or at least extremely time consuming) if one needs to explore the whole dataset, retrieve the data and compute the output needed.

A more technical challenge that climate scientists are facing with the increase in volume but also in data sources (satellites, observations, many instances of models) is the data discovery and indexing part. The Earth System Grid Federation (ESGF) is an example of web portal that serves this kind of data.

### Who benefits and how?

The first and direct research communities targeted are the climate sciences and atmospheric modelling ones. Developing and using innovative solutions provided by the EUDAT community will allow them to store more efficiently their data as well as transferring/sharing data among users widely disseminated around Europe who will benefit an easier access to the results of the whole community, making the scientific work more collaborative.

Downstream communities such as health (impact of climate and aerosols on health) and climate services (renewable energy industry, policy makers, etc.) will also be directly impacted by the expected outcome of this pilot as users of the data and products generated.

### Technical Implementation

In the framework of our data pilot, we planned to install and work on two EUDAT services: B2SAFE and B2STAGE. For these two services, during a timeframe of 6 months, we performed the following technical tasks:

---

EUDAT receives funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 654065.



- B2SAFE included the local installation of the IRODS server for the pilot and federation of the server with the service provider (BSC). Additionally, data has been copied from dust models outputs (initially stored in a local storage) to the B2SAFE storage (around 1TB). Several options to connect the IRODS server to an external web server where the data are presented “to the outside world” (work in progress) has begun.
- In terms of B2STAGE to date the x509 certificates have been purchased and tests of GridFTP have been performed.

## Preliminary Results

The pilot final objective is to create a backend repository for the project “Sand and Dust Storm Warning Advisory and Assessment System”. This research project is a collaboration between the BSC and the Spanish meteorological agency and aims to improve the numerical modelling of sand and dust storms through the continuous evaluation of a set of numerical models outputs provided by several institutions around the world (BSC, NASA, ECMWF, etc.), collecting observational data proceeding from various sources (satellites, in situ observations, etc.).

It also provides services to the community such as daily forecast plots, time series of model comparisons, statistics of models performances and datasets download. Currently 12 models output of 11 institutions of 9 different countries are participating to the project. The total estimated storage occupied is about 1T Band growing.

At the end of the pilot the result expected is that users (scientific community) will be able to download forecast datasets stored in B2SAFE storage directly through the web portal interface.

## Contacts

- Pierre-Antoine Bretonnière, BSC, pierre-antoine.bretonniere(a)bsc.es
- Massimo Fares, INGV, massimo.fares(a)ingv.it

## Further Information

[Learn more about the "air quality modelling pilot"/EUDAT collaboration here](#)  
[Scientific Challenges behind the Pilot](#)

[Read more](#)