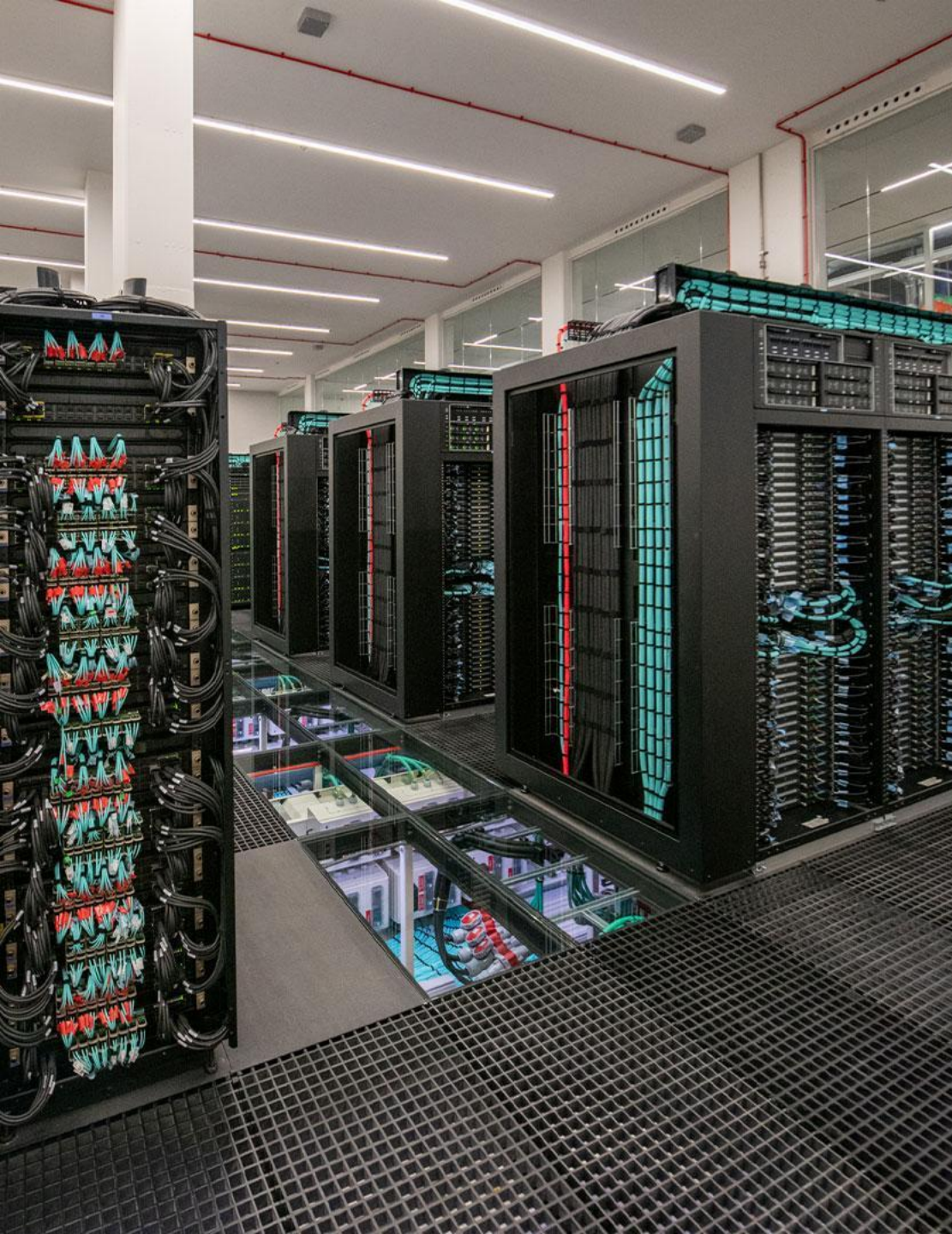




**Barcelona  
Supercomputing  
Center**  
Centro Nacional de Supercomputación



EXCELENCIA  
SEVERO  
OCHOA

## **Providentia**

An evaluation software package for the  
atmospheric modelling community

**Alba Vilanova Cortezón**

Research Engineer at Earth Sciences Department of  
Barcelona Supercomputing Center

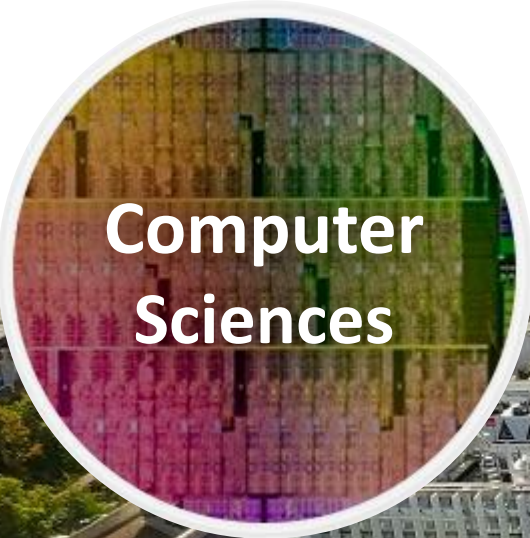
June 30, 2026 · RIKEN AIP Open Seminar

# Earth Sciences Department

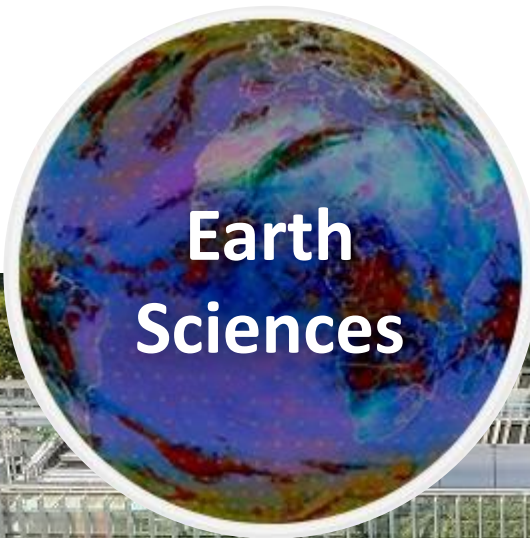


**Barcelona  
Supercomputing  
Center**  
*Centro Nacional de Supercomputación*

# Scientific Departments at BSC



**Computer  
Sciences**



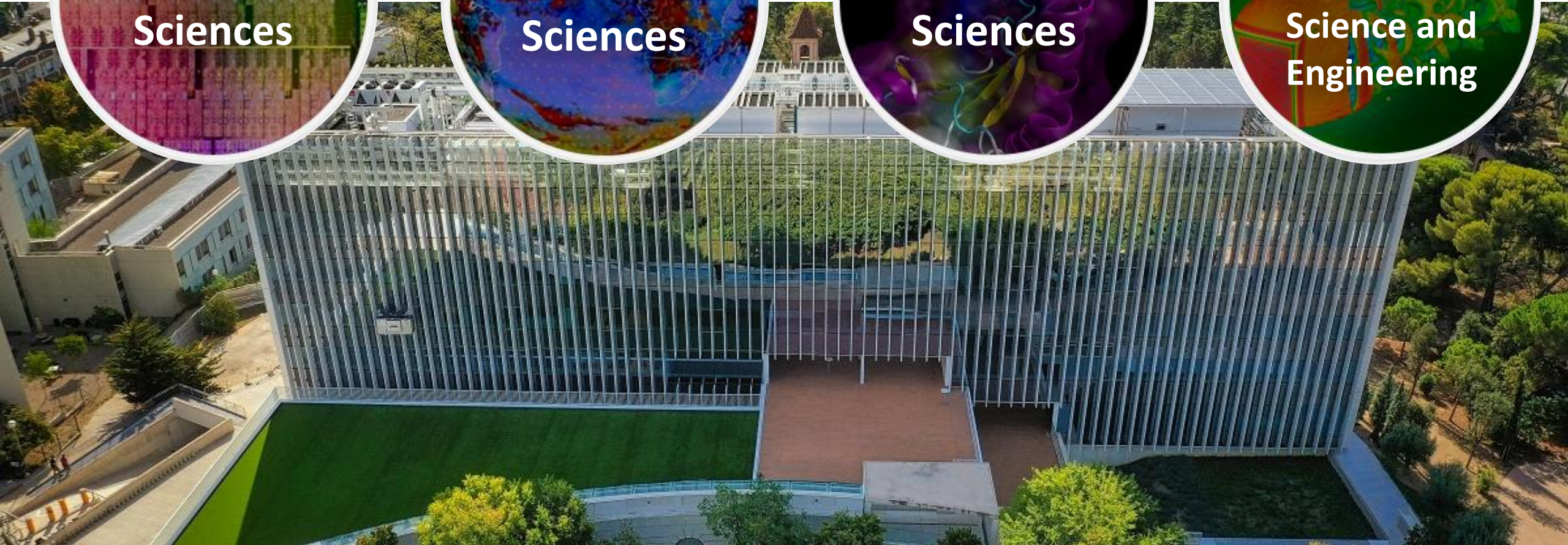
**Earth  
Sciences**



**Life  
Sciences**



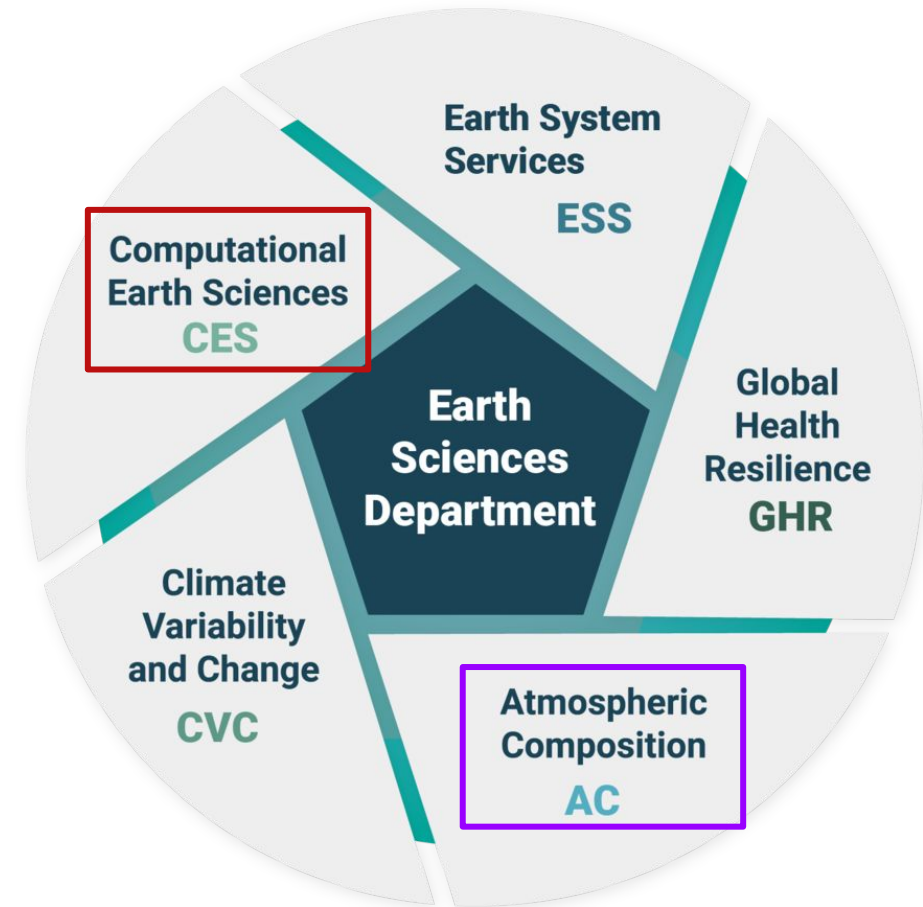
**Computer  
Applications in  
Science and  
Engineering**



# Earth Sciences Department

**Environmental modelling and forecasting** using models, with a particular focus on **climate, air quality, atmospheric composition and climate-related impacts**.

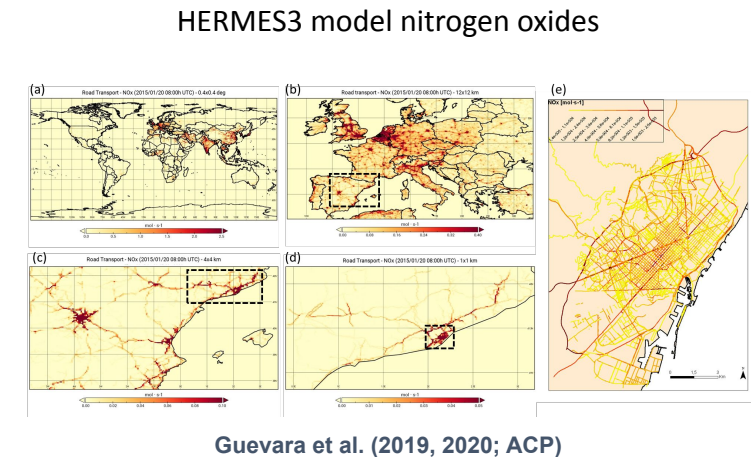
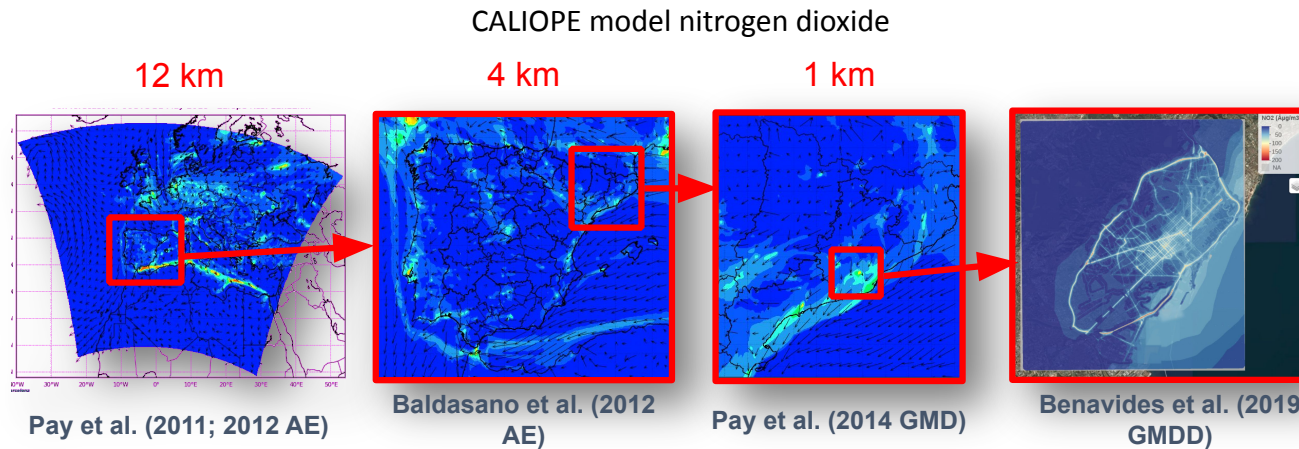
This includes **transferring solutions** to support the main societal environmental challenges through data applications.



# Atmospheric Composition Group

It produces **daily air quality forecasts and emission inventories** at several scales:

- **MONARCH** model: Advanced chemistry and aerosol modelling system representing the atmospheric chemistry cycle. This data is provided since 2018 to the Copernicus Atmosphere Monitoring Service.
- **HERMES3** emission model: Modelling framework for gas and aerosols.

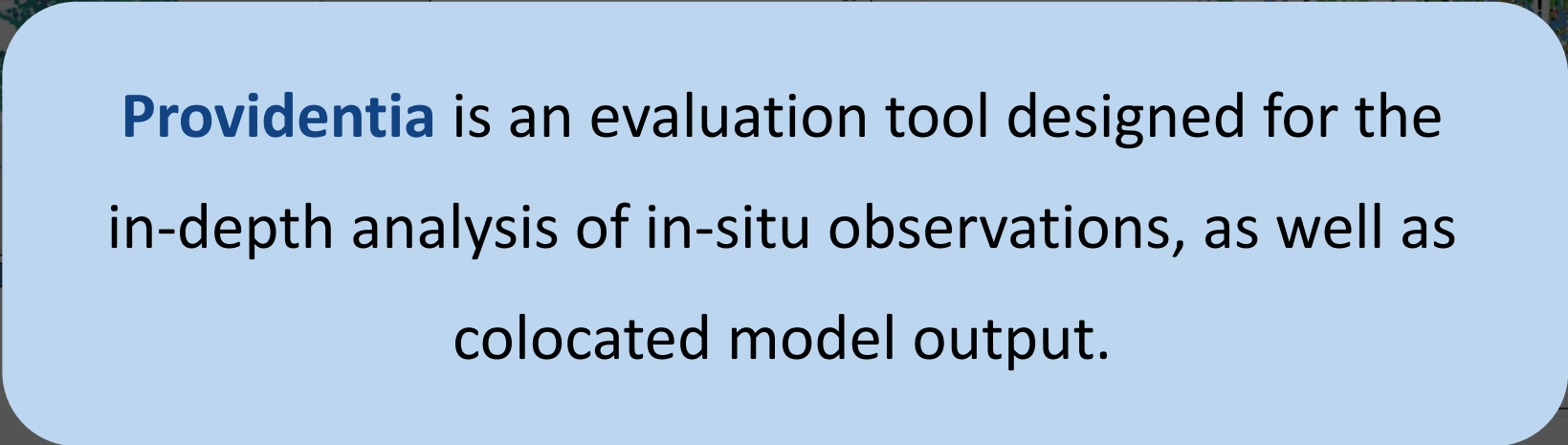


It also contributes to improving the global and regional systems and to the **evaluation of both reanalyses and forecasts**, as well as to global and regional emission inventories.

# Providentia

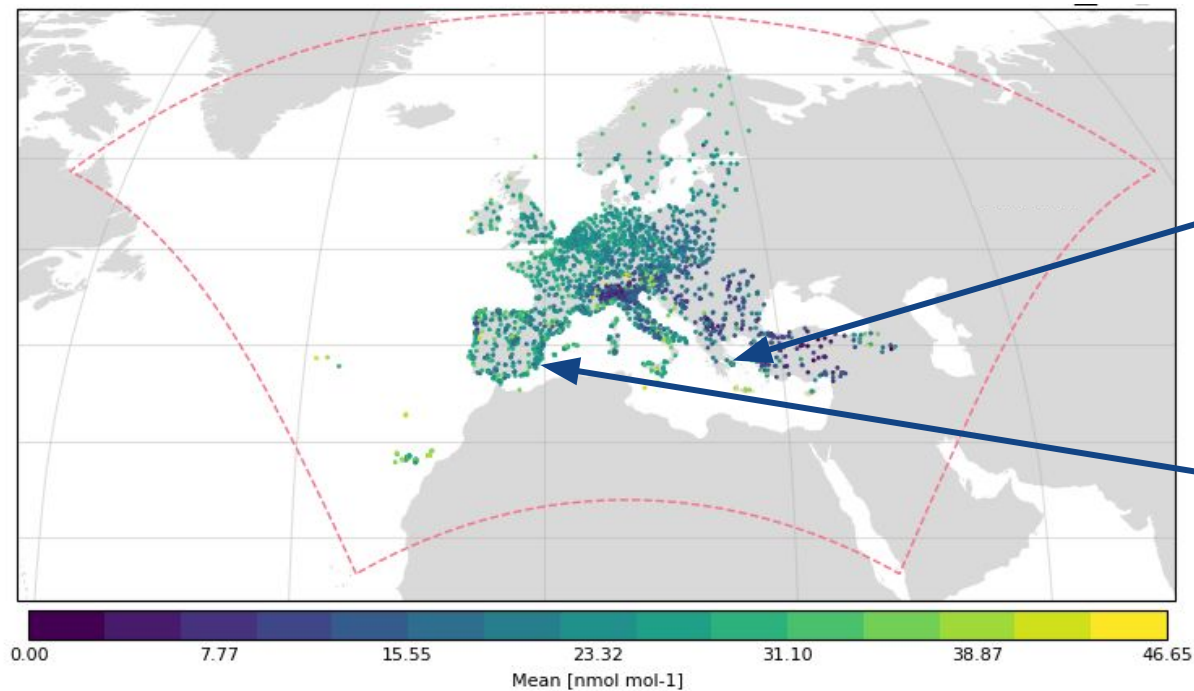


**Barcelona  
Supercomputing  
Center**  
*Centro Nacional de Supercomputación*



# Why was Providentia created?

**Reason 1:** Evaluate ground-based station data



Mean of the observed surface concentration of ozone (January 2023)



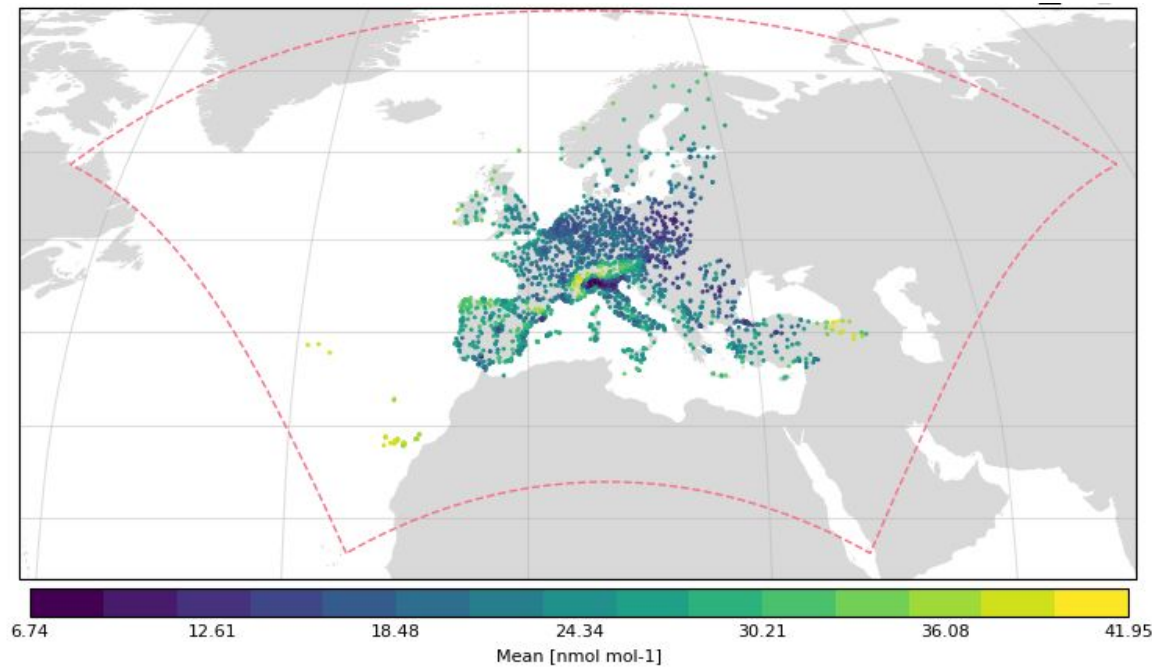
Teledyne 400 ozone  
measuring instrument



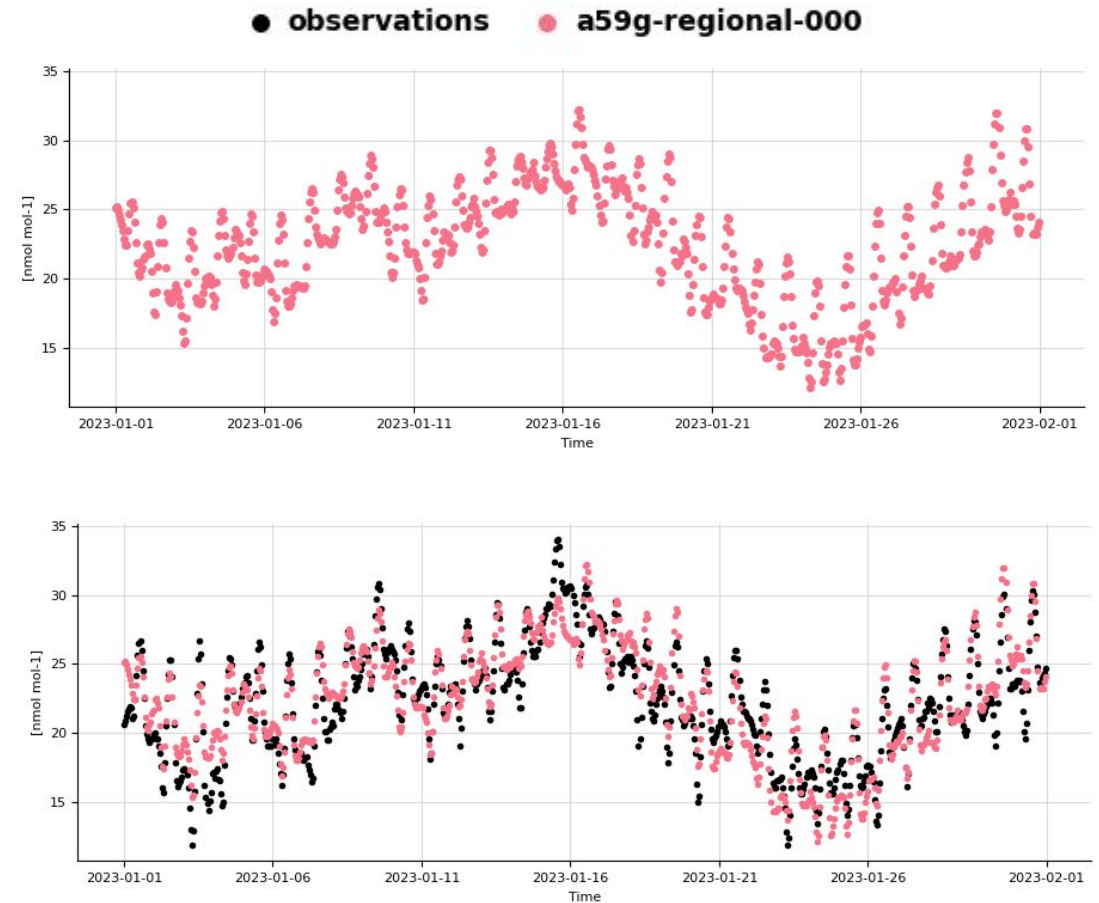
Ecotech serinus 10  
ozone analyzer

# Why was Providentia created?

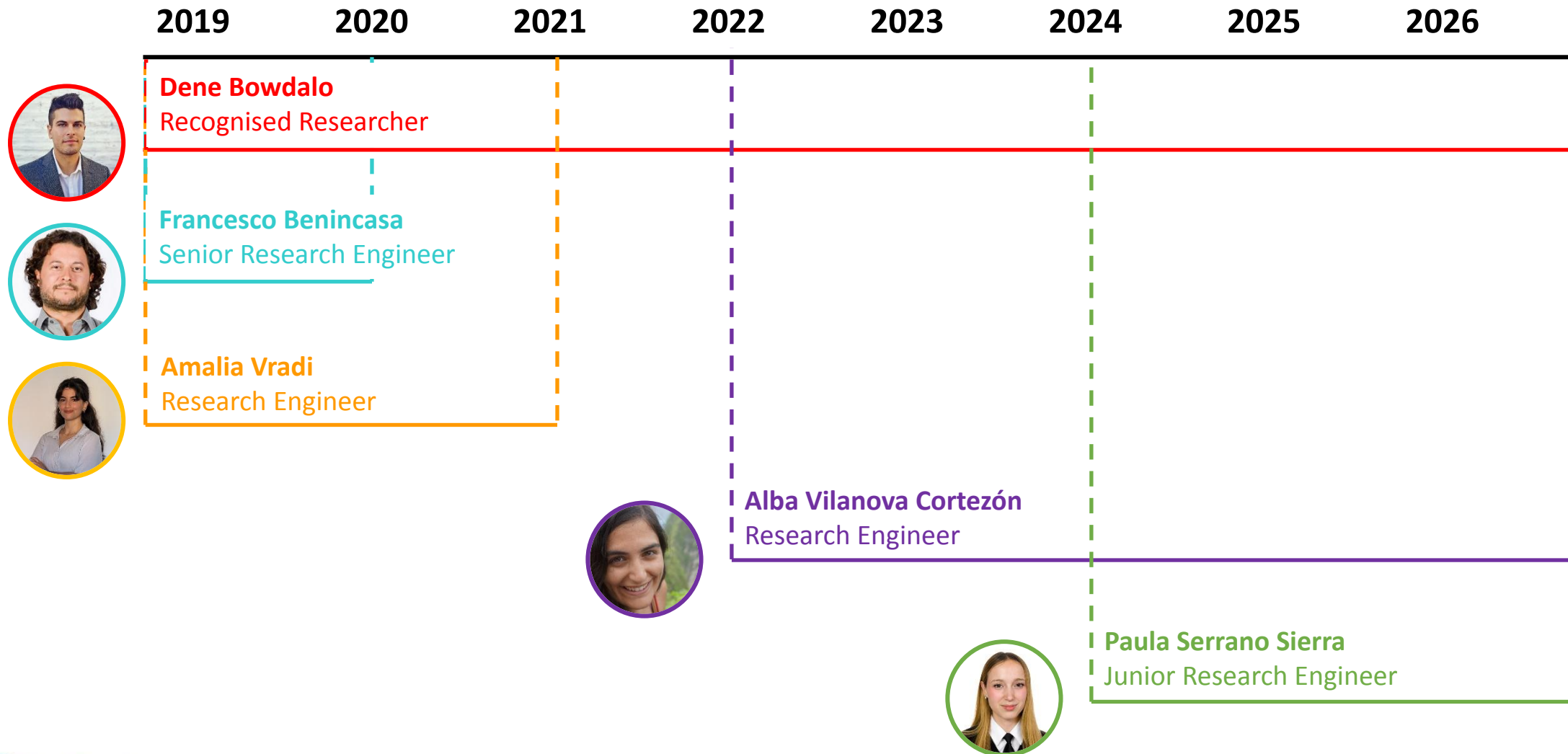
**Reason 2:** Evaluate model data using ground-based station data to **improve forecasts**



Mean of the modelled surface concentration of ozone (January 2023)



# Development timeline



# Use of Providentia



## Internal users at BSC:

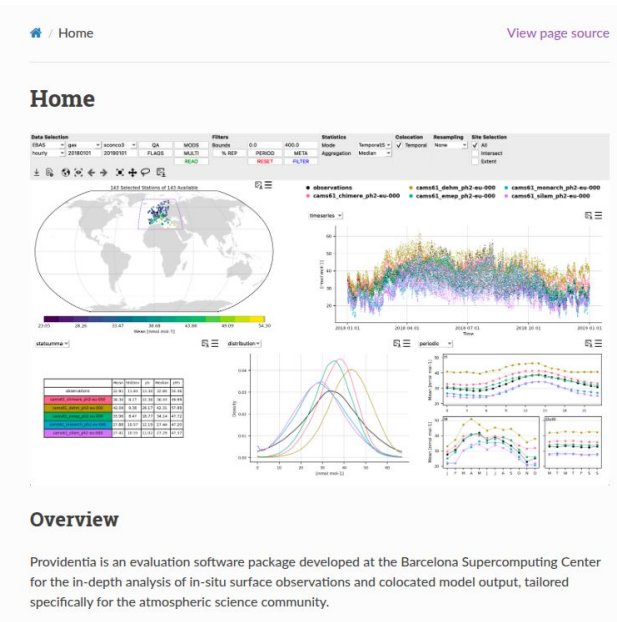
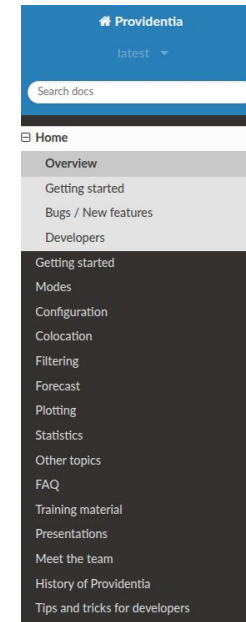
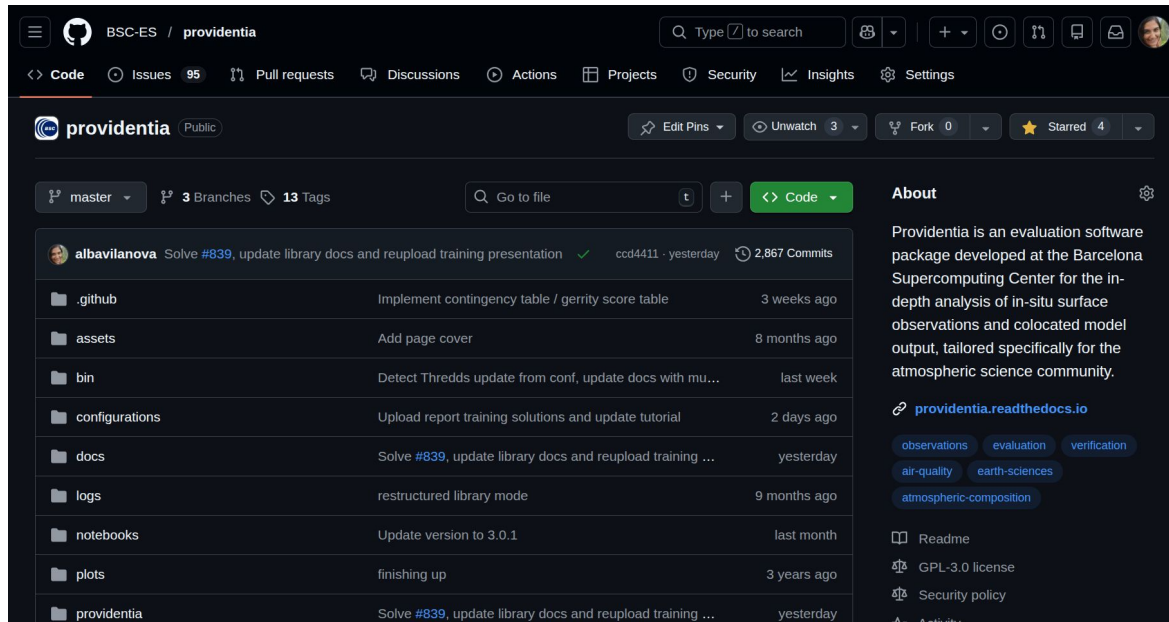
- Used by 25 atmospheric modellers

## External users:

- Spanish National Research Council (Spain)
- University of Granada (Spain)
- University of Zaragoza (Spain)
- Norwegian Institute for Air Research (Norway)
- Charles University (Czech Republic)
- ERATOSTHENES Centre of Excellence (Cyprus)
- Directorate General of Civil Aviation (Kuwait)



# Access



**Repository:** <https://github.com/bsc-es/providentia>

**Documentation:** <https://providentia.readthedocs.io/en/latest/>

**Stack:** Python and Bash

**Machines:** HPC (MN5 and Nord4) and local

**OS:** Linux and Mac. On Windows through WSL and virtual machines.

# Structure of Providentia

# Structure

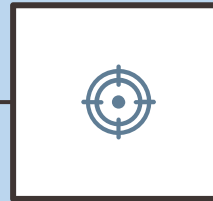
## Analysis

### Pre-processing



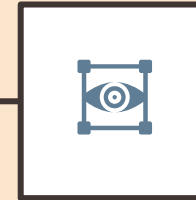
#### Download

Downloads and formats the models and observations.



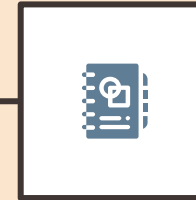
#### Interpolation

Matches spatially models to observation networks.



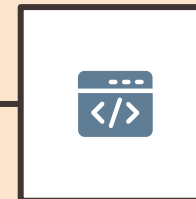
#### Dashboard

Performs interactive analyses of model results, using observational data.



#### Report

Creates detailed pdf reports for in-depth analysis.



#### Library

Uses backend functions within scripts.

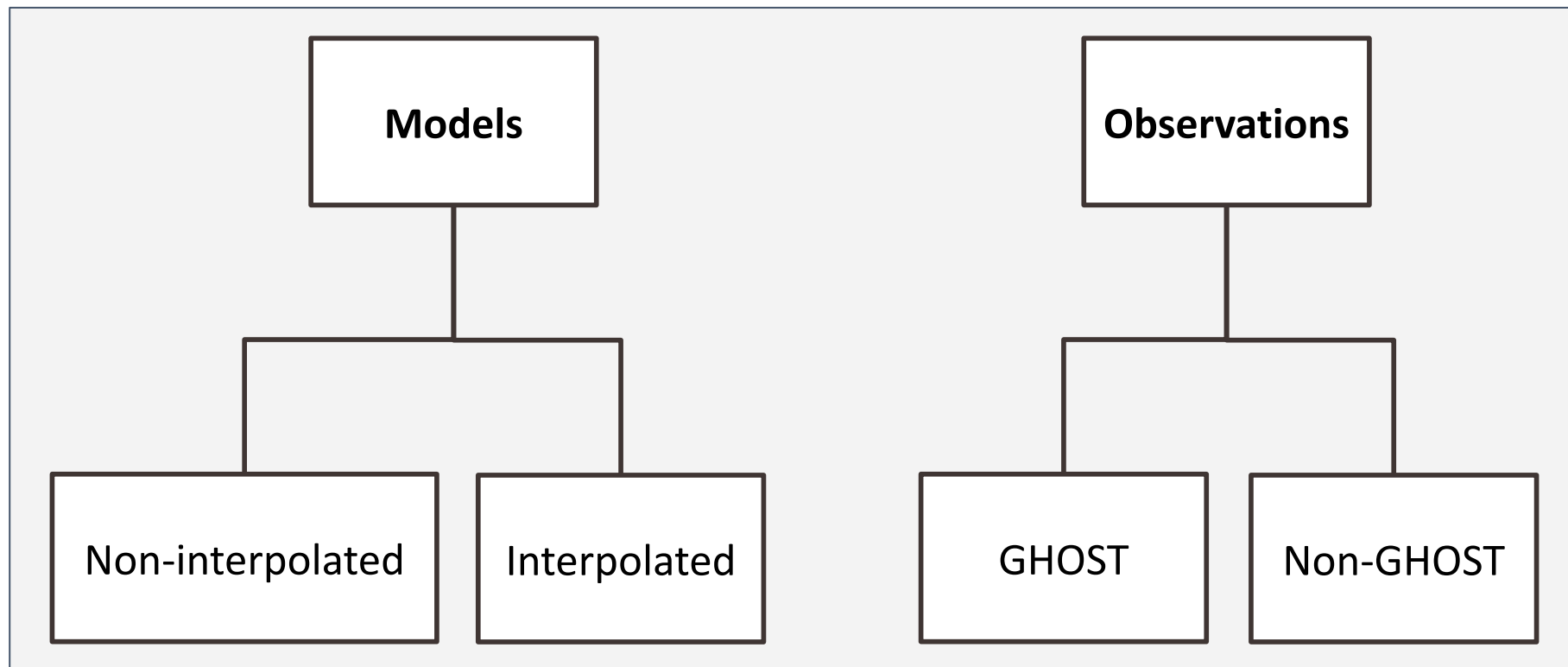
# Data



**Barcelona  
Supercomputing  
Center**  
*Centro Nacional de Supercomputación*

# Data classification

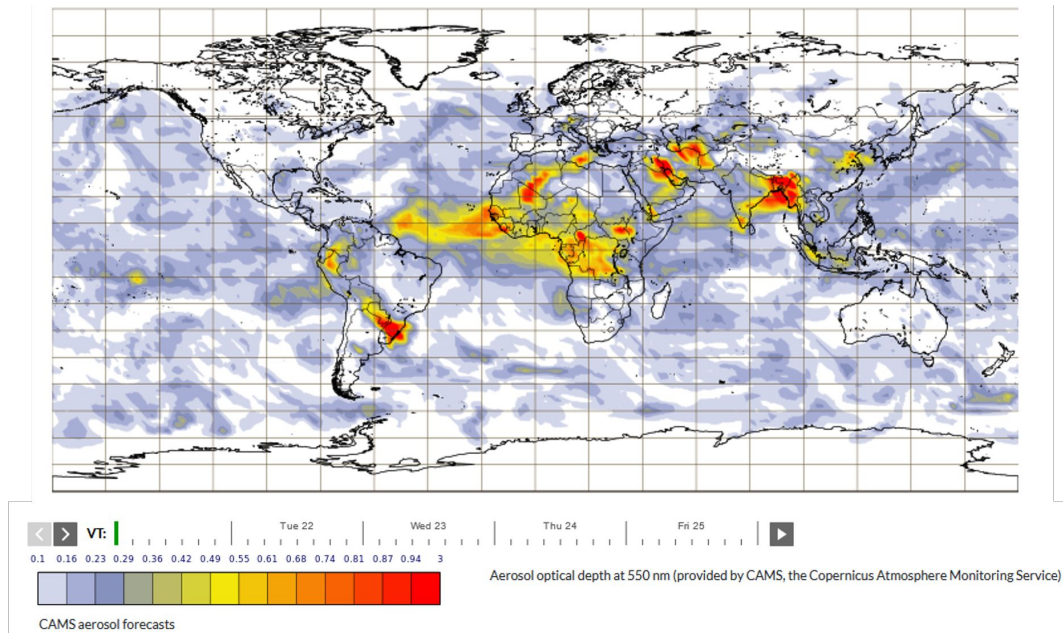
To help you understand the data used in Providentia, we classify it into the following groups and subgroups:



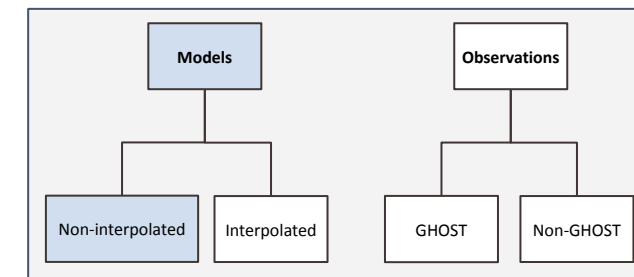
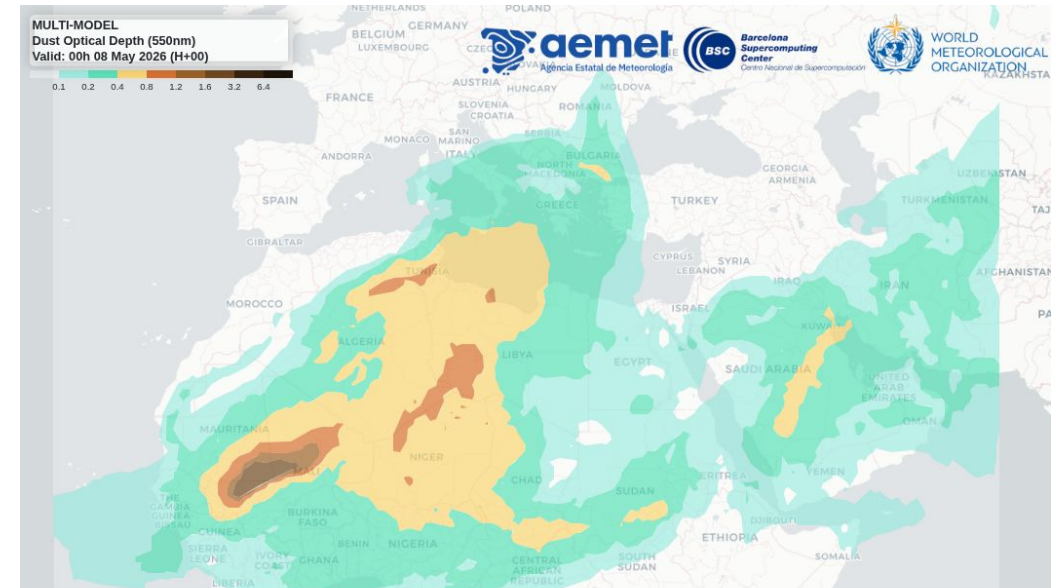
# Non-interpolated model

These are gridded datasets simulating the atmosphere at different geographical domains

## CAMS Aerosol Optical Depth

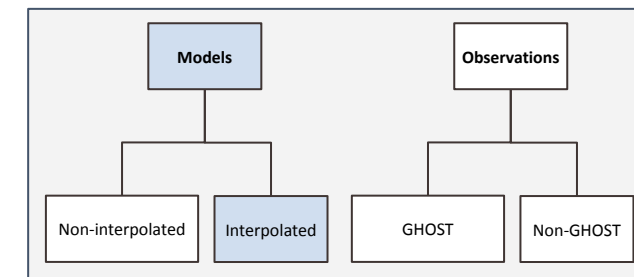
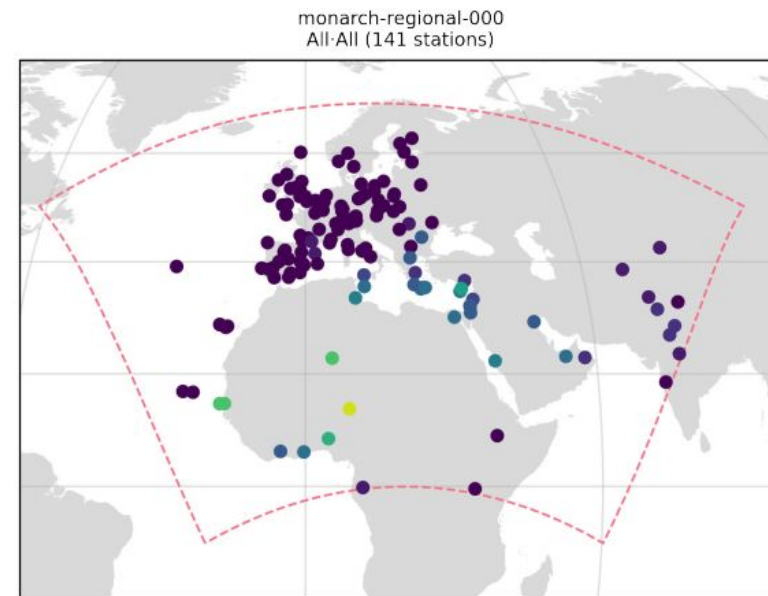
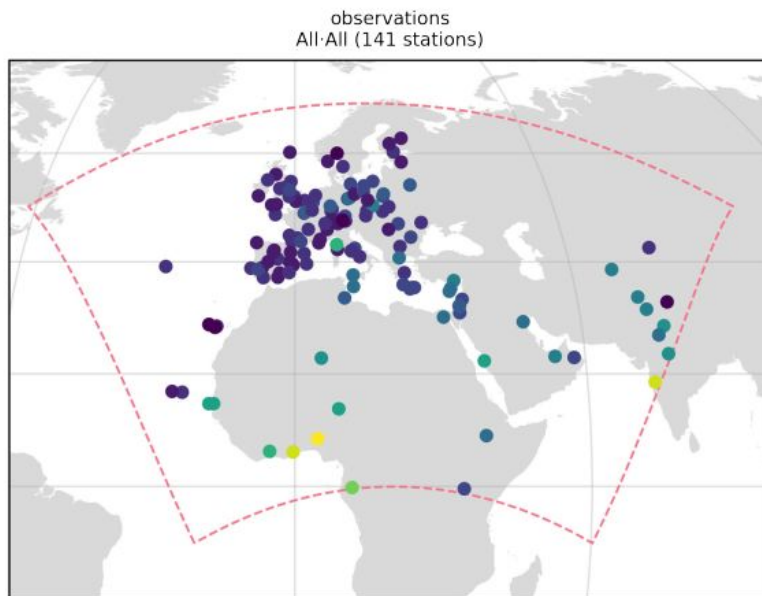


## BDRC Dust Optical Depth



# Interpolated model

Interpolation in Providentia refers to the mathematical process of estimating the model variable at the exact location of an observational station. *More to come in the following slides.*



# GHOST:

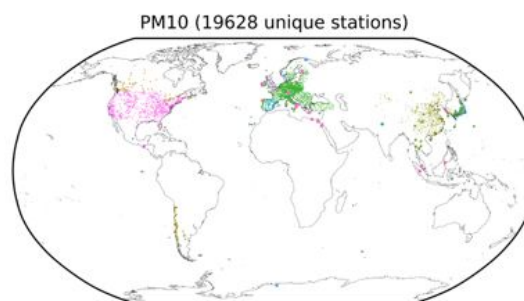
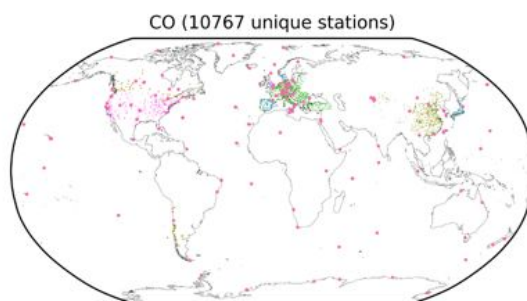
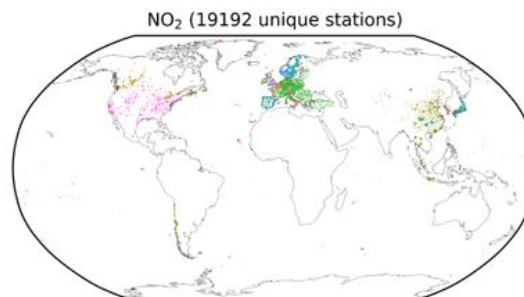
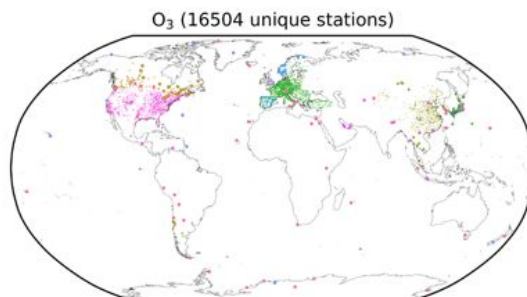
## Globally harmonised observations in space and time

N° Processed  
Measurements:  
**~7.3 billion**

N° Components:  
**227**

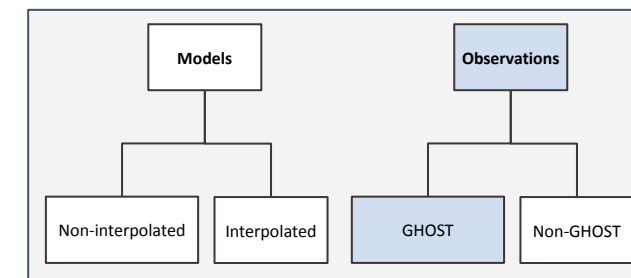
N° Networks:  
**38**

Temporal range:  
**1970-2023**



- |                        |                        |
|------------------------|------------------------|
| ● ACTRIS               | ● AERONET v3 Level 1.5 |
| ● BJMEMC               | ● CAMP                 |
| ● Chile SINCA          | ● CNEMC                |
| ● EEA AirBase          | ● EEA AQ e-Reporting   |
| ● EUSAAR               | ● HELCOM               |
| ● Independent (EBAS)   | ● Japan NIES           |
| ● NADP AMNet           | ● NADP AMoN            |
| ● NOAA-GGGRN           | ● OECD                 |
| ● US EPA AirNow DOS    | ● US EPA AQS           |
| ● WMO GAW WDCGG        | ● WMO GAW WDCRG        |
| ● AERONET v3 Level 2.0 | ● AMAP                 |
| ● Canada NAPS          | ● CAPMoN               |
| ● COLOSSAL             | ● EANET                |
| ● EMEP                 | ● EUCAARI              |
| ● HTAP                 | ● IMPACTS              |
| ● Mexico CDMX          | ● MITECO               |
| ● NILU                 | ● NOAA-ESRL            |
| ● UK AIR               | ● UK DECC              |
| ● US EPA CASTNET       | ● WMO GAW WDCA         |

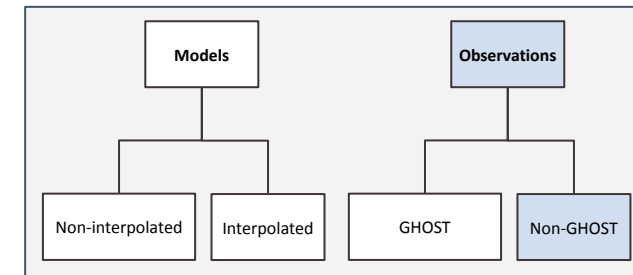
Bowdalo, D., Basart, S., Guevara, M., Jorba, O., Pérez García-Pando, C., Jaimes Palomera, M., Rivera Hernandez, O., Puchalski, M., Gay, D., Klausen, J., Moreno, S., Netcheva, S., and Tarasova, O.: GHOST: a globally harmonised dataset of surface atmospheric composition measurements, *Earth Syst. Sci. Data*, 16, 4417–4495, <https://doi.org/10.5194/essd-16-4417-2024>, 2024.



# non-GHOST

Providentia can read data non-harmonised observational data, named as non-GHOST, the limitation is the data will be have less filtering features, specifically:

- No GHOST QA flags
- No network flags
- No native representativity filters
- No period filters
- Limited metadata

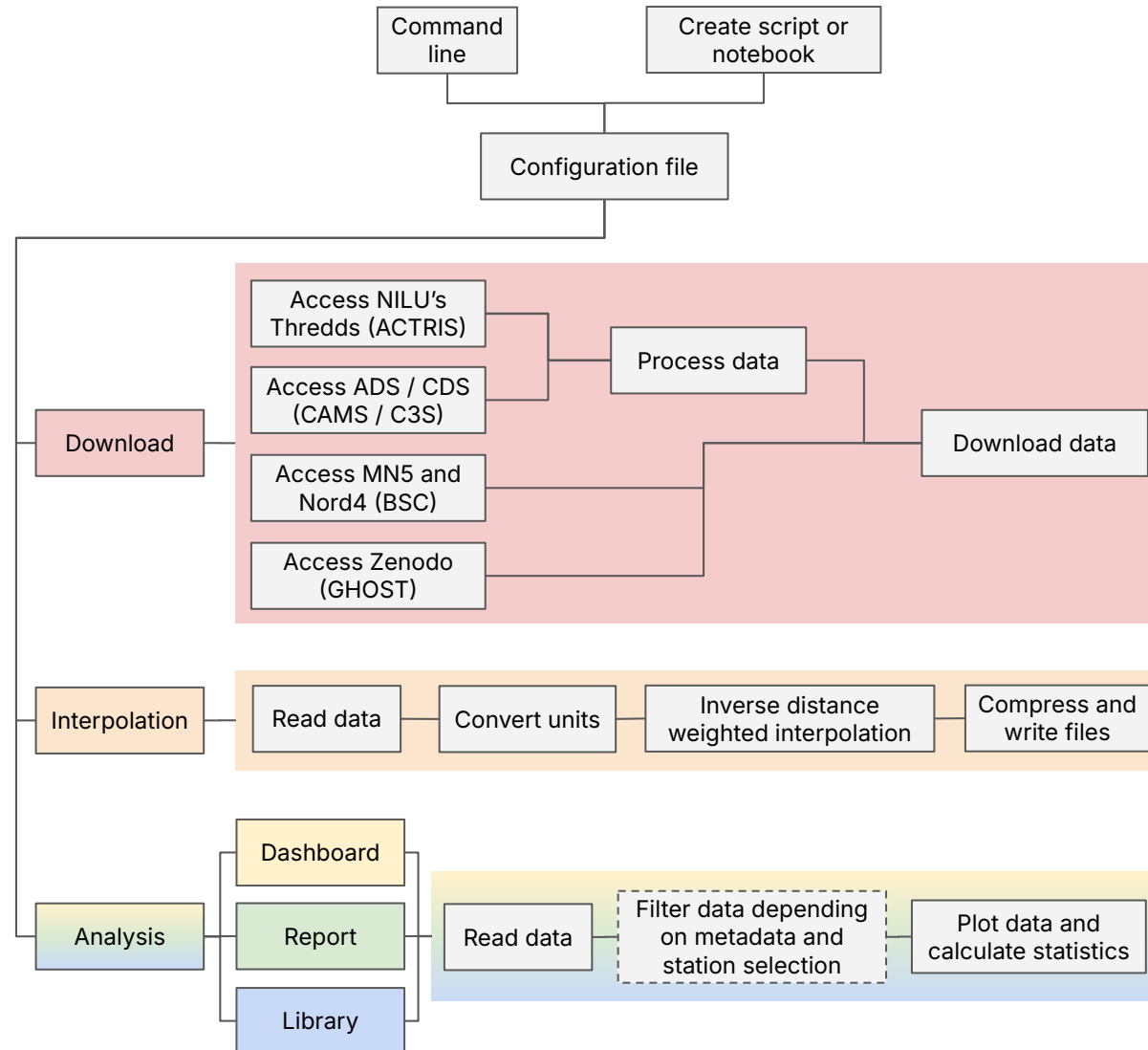


# Workflow



**Barcelona  
Supercomputing  
Center**  
*Centro Nacional de Supercomputación*

# Workflow



# Configuration files

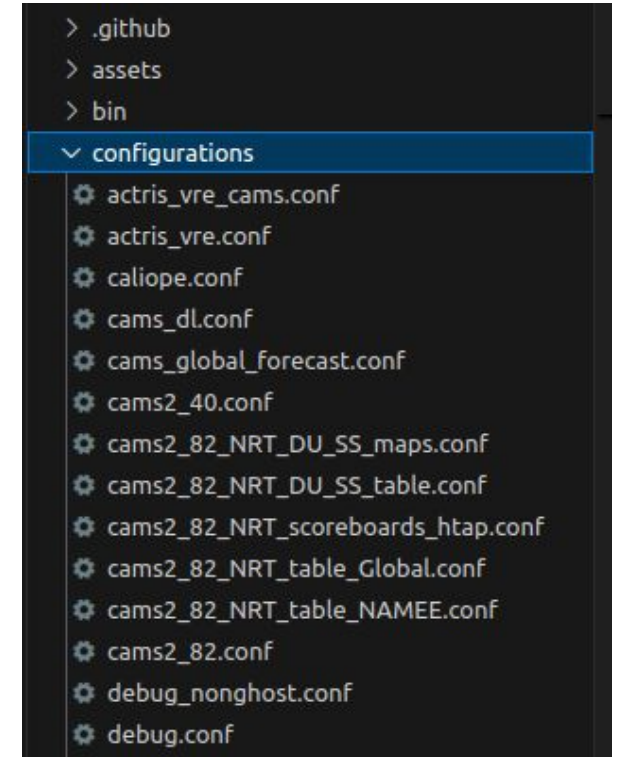
Configuration files are files where the user **defines the parameters** used to specify which data to:

- Download (download mode)
- Interpolate (interpolation mode)
- Read and plot (dashboard, report and library modes)

By default, Providentia reads these files from the **configurations/** directory.

For more details, see the official documentation:

<https://providentia.readthedocs.io/en/latest/Configuration-files.html>



# Main fields in a configuration file

The following fields are commonly defined in configuration files across all the different modes:

|                      | Description                                                                                       |
|----------------------|---------------------------------------------------------------------------------------------------|
| <b>network</b>       | Observational data. (e.g. <b>EBAS</b> )                                                           |
| <b>models, model</b> | Model IDs. Domain and ensemble information can also be included. (e.g. <b>a4j1-regional-000</b> ) |
| <b>species</b>       | Atmospheric components to analyse in BSC nomenclature. (e.g. <b>sconco3</b> )                     |
| <b>resolution</b>    | Temporal resolution of the observations (e.g. <b>hourly</b> )                                     |
| <b>start_date</b>    | Start date. In YYYYMMDD format or YYYYMM for interpolation. (e.g. <b>20210101</b> )               |
| <b>end_date</b>      | End date. In YYYYMMDD format or YYYYMM for interpolation. (e.g. <b>20220101</b> )                 |

# Structure of a configuration file

**Section1 ([name])**

**Section1·Subsection1 ([[name]])**

**Section1·Subsection2 ([[name]])**

**Section2 ([name])**

**Section2·Subsection1 ([[name]])**

configurations/multiple\_configuration\_name.conf

```
[PRV_sconco3_CHIMERE]
network = EBAS
species = sconco3
resolution = hourly
start_date = 20180101
end_date = 20180601
models = cams61_chimere_ph2-eu-000 (CHIMERE)
```

```
[[Spain]]
latitude = 35, 44
longitude = -9, 2
area_classification = rural
```

```
[[British_Isles]]
latitude = 50, 58
longitude = -12, 2
```

```
[PRV_sconco3_EMEP]
network = EBAS
species = sconco2
resolution = hourly
start_date = 20180101
end_date = 20190101
model = cams61_emep_ph2-eu-000 (EMEP)
```

```
[[France]]
country = keep: France | |
```

# Download



**Barcelona  
Supercomputing  
Center**  
*Centro Nacional de Supercomputación*

# Download data to your local machine

Providentia provides direct access to observational and model data that is publicly available via a dedicated download mode.

## Observational data

- GHOST (Zenodo)
- ACTRIS (via NILU THREDDS)

## Model data

- BSC models (HPC users)
- CAMS regional forecast, analysis, and reanalysis
- CAMS global forecast, and reanalysis
- ERA5 reanalysis

For more details, see the official documentation:

<https://providentia.readthedocs.io/en/latest/Download-mode.html>

# Interpolation

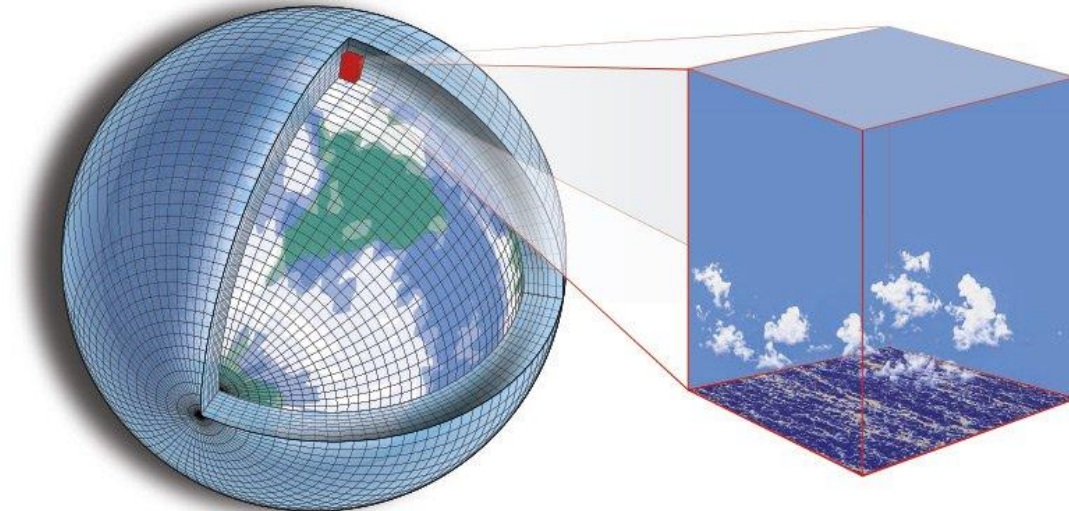


**Barcelona  
Supercomputing  
Center**  
*Centro Nacional de Supercomputación*

# What does interpolation do?

Providentia makes use of in-situ observations from specific measurement stations, at fixed locations. When you have model data, its output is typically provided on a regular model grid at a fixed resolution.

Interpolation in Providentia refers to the mathematical process of estimating the model variable at the exact location of an observational station.

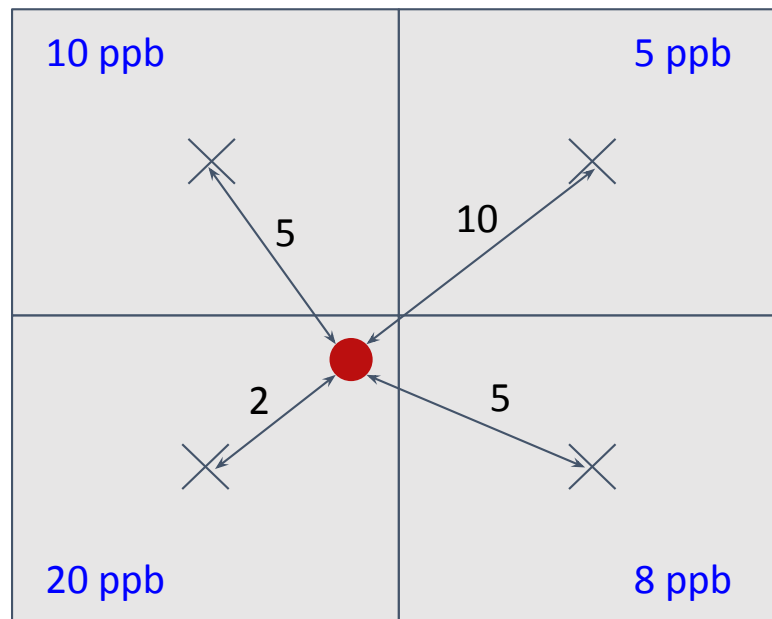


# Inverse distance weighted interpolation

**Inverse Distance Weighted** interpolation estimates the value at an observational station by averaging model data in **N closest neighbour** grid cells, giving higher weight to grid cells that are closer.

N neighbours can be set via **interp\_n\_neighbours = N** in the configuration file. The default is **4**.

For 4 closest neighbours (default):



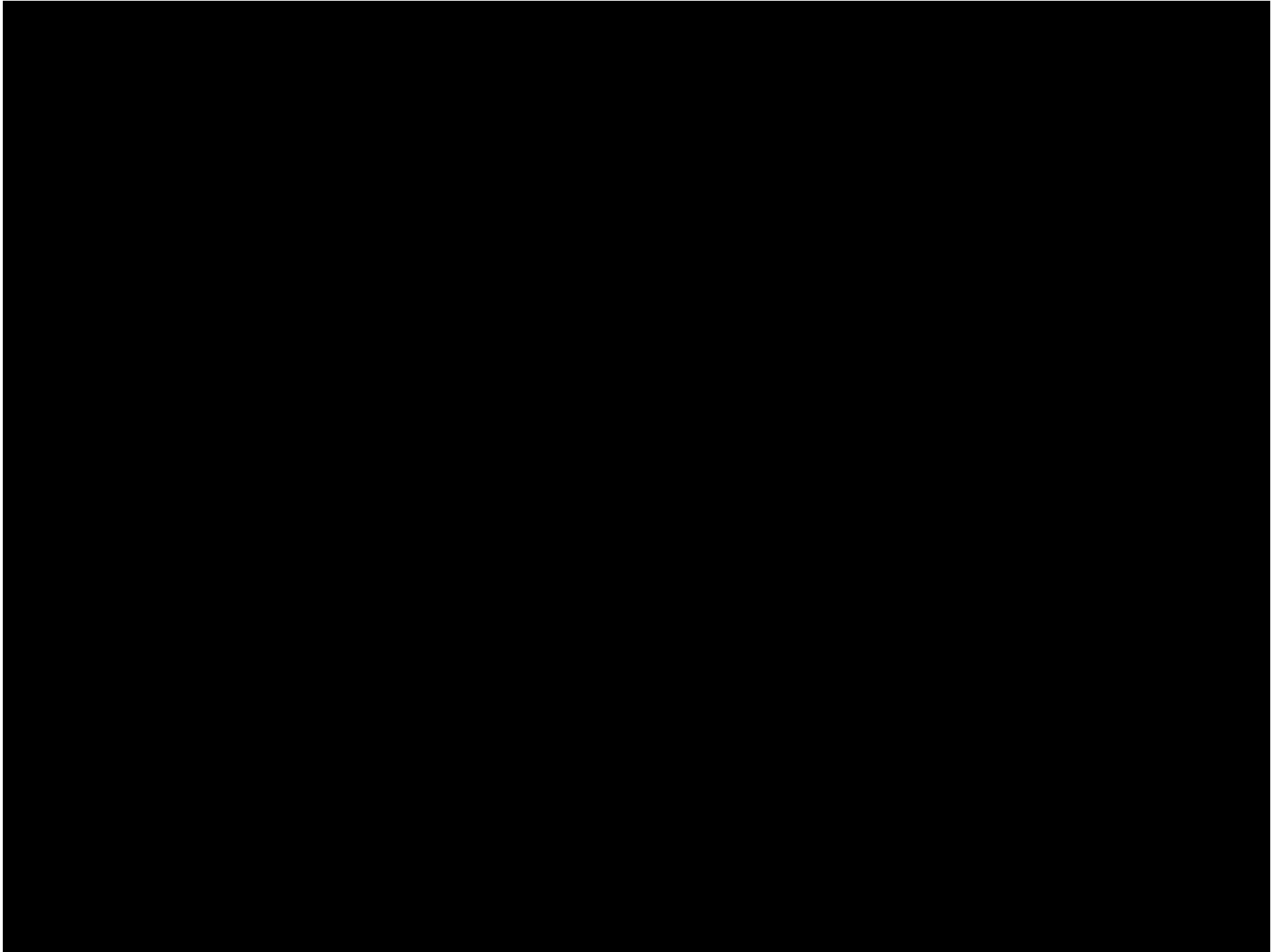
## Weights

$$\frac{1}{2}, \frac{1}{5}, \frac{1}{5}, \frac{1}{10} \\ = 0.5, 0.2, 0.2, 0.1$$

Weighted average = **14.1 ppb**

Nearest neighbour = **20 ppb**

# Dashboard



# Report



**Barcelona  
Supercomputing  
Center**  
*Centro Nacional de Supercomputación*

# Report

## EVALUATION REPORT

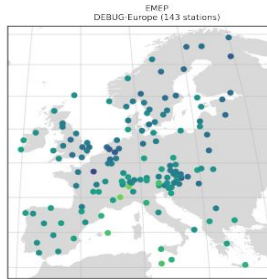
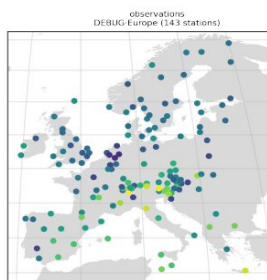
Network: EBAS  
Species: O<sub>3</sub>  
Temporal resolution: Hourly  
Date range: 01/01/2018 - 01/01/2019  
Models: EMEP, MONARCH  
Subsections: DEBUG-Europe,

Earth Sciences  
Department

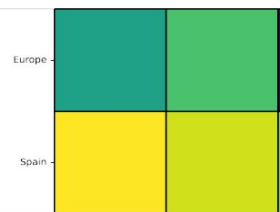
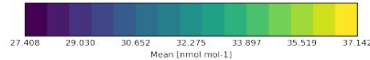


**Barcelona  
Supercomputing  
Center**  
Centro Nacional de Supercomputación

Map Median (Summary)  
EBAS|sconco3

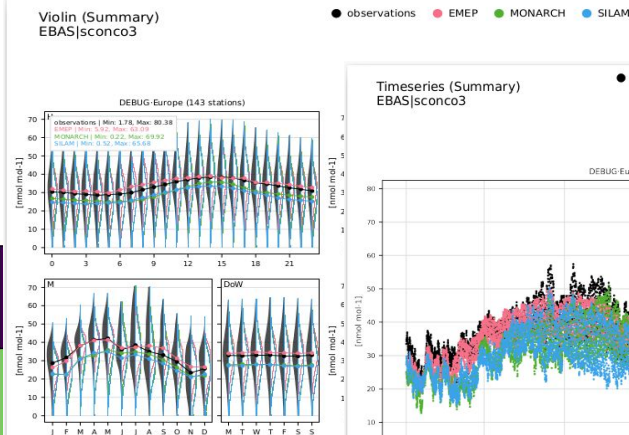


Heatmap Mean (Summary)  
multispecies

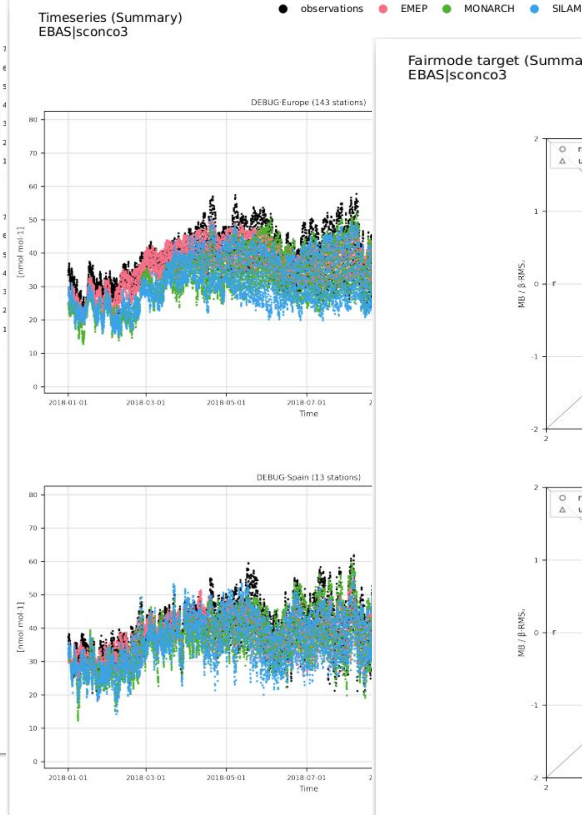


observations EMEP

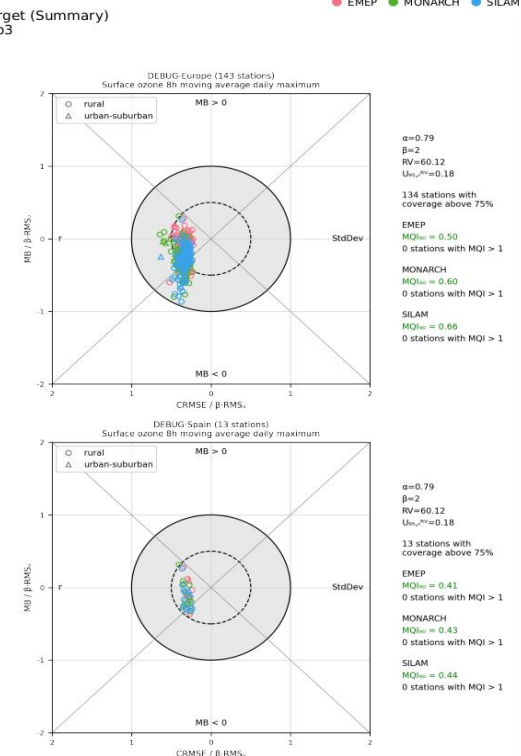
Violin (Summary)  
EBAS|sconco3



Timeseries (Summary)  
EBAS|sconco3



Fairmode target (Summary)  
EBAS|sconco3



# Library



**Barcelona  
Supercomputing  
Center**  
Centro Nacional de Supercomputación

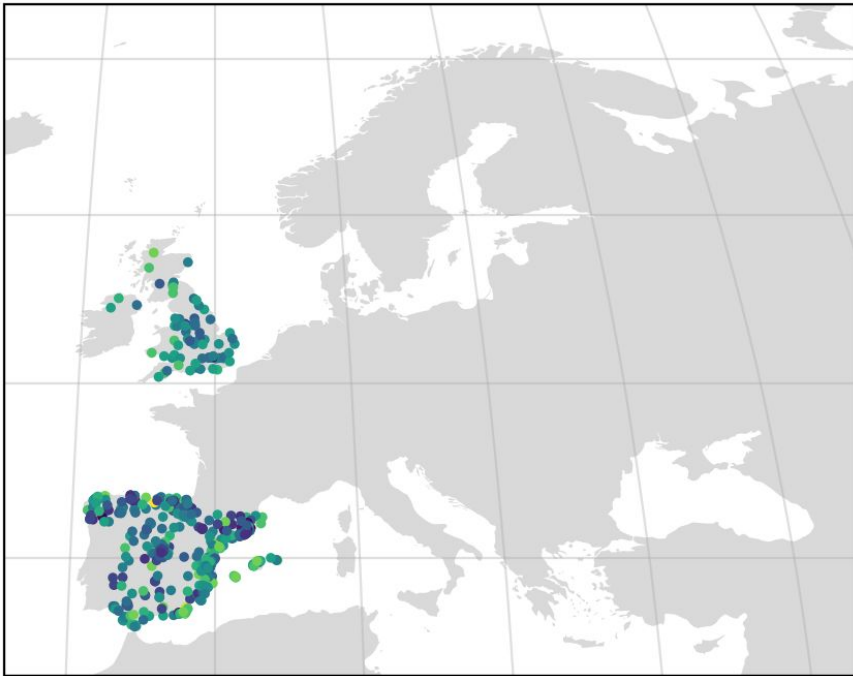


# Other features

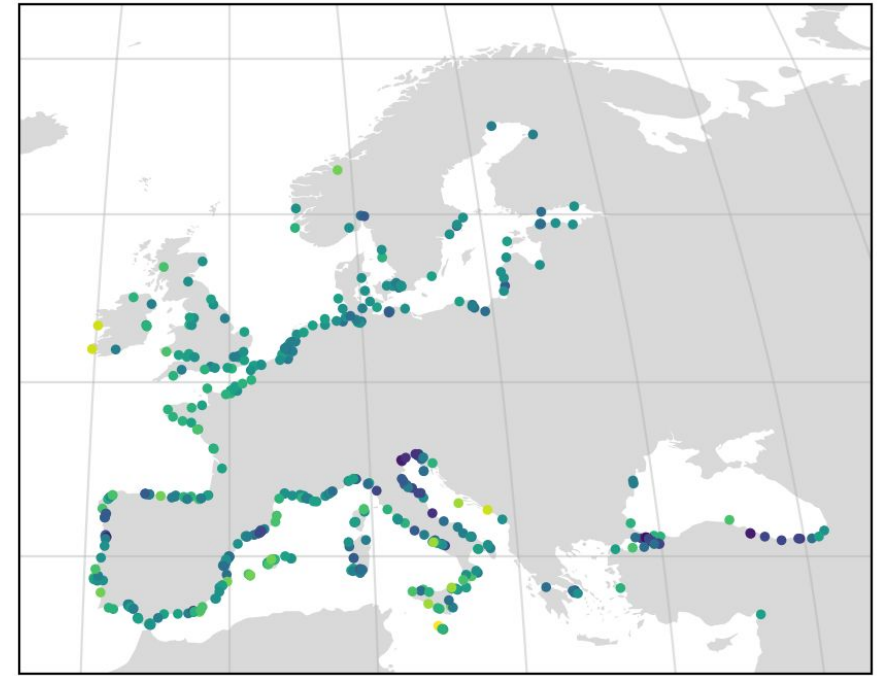
# Filtering by metadata

Providentia provides multiple way of filtering data, each of which can be applied on-the-fly. One of the most typical is by metadata, either text or numerical.

**Keep stations in Spain & UK**

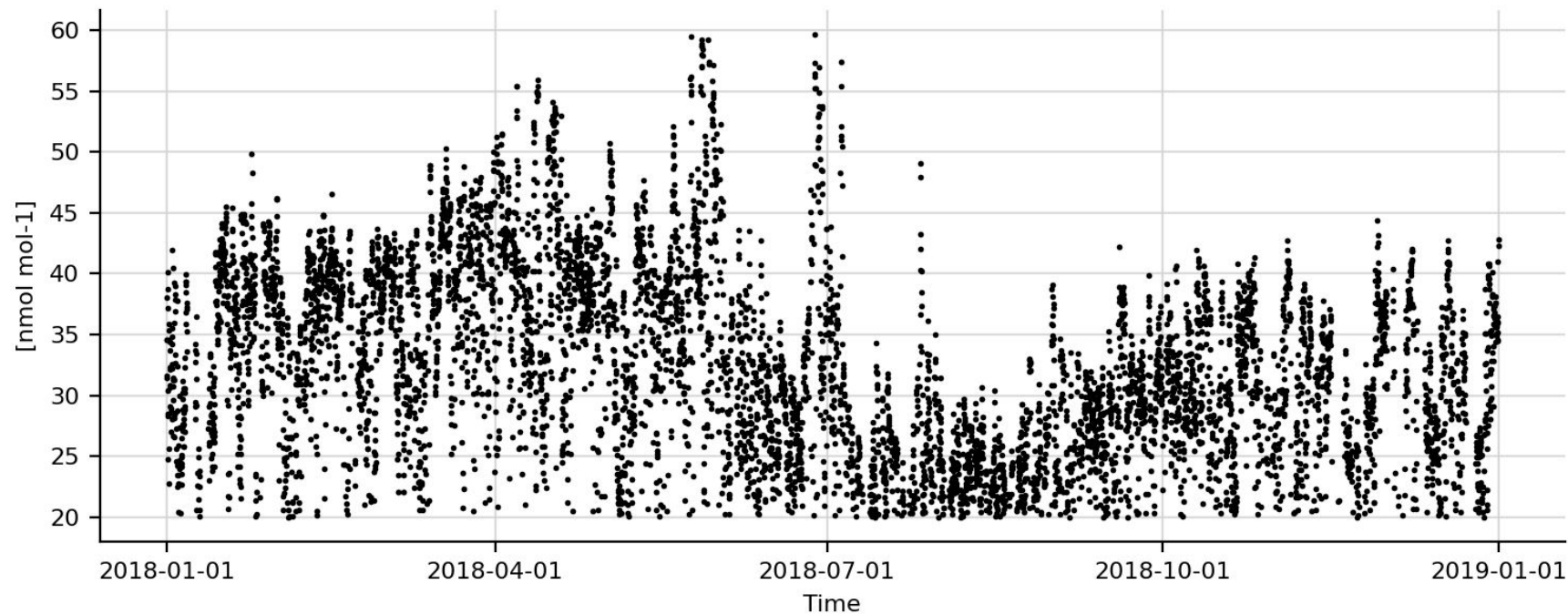


**Keep stations <10km from coast**



# Filtering by data bounds

|        |                                   |                                   |
|--------|-----------------------------------|-----------------------------------|
| Bounds | <input type="text" value="20.0"/> | <input type="text" value="60.0"/> |
|--------|-----------------------------------|-----------------------------------|

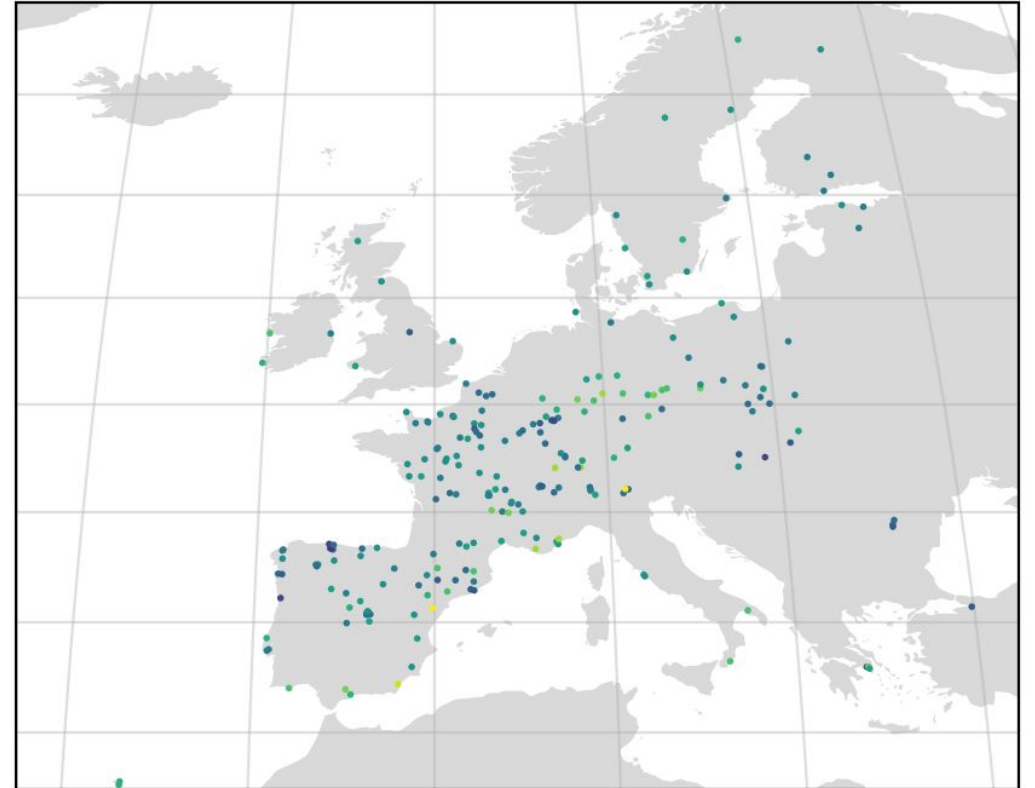


Example of a station time series with data bounds imposed

# Filtering by data coverage



**Minimum 0% Data Coverage**  
**2307 Stations**



**Minimum 99% Data Coverage**  
**253 Stations**

# Filtering by QA flags

Select standardised data reporter provided flags to filter by

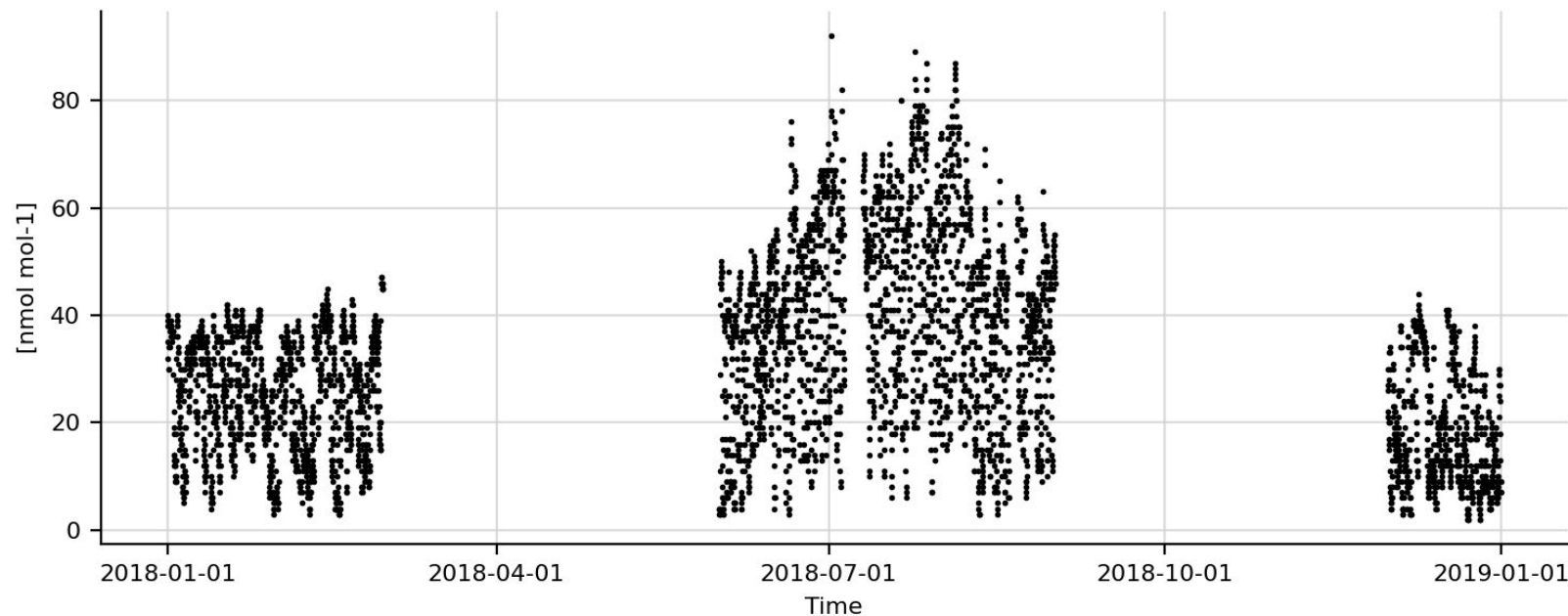
Select All Clear All Select Default

|                                                                                       |                                                                                 |                                                                         |                                                                        |
|---------------------------------------------------------------------------------------|---------------------------------------------------------------------------------|-------------------------------------------------------------------------|------------------------------------------------------------------------|
| <input type="checkbox"/> Valid Data                                                   | <input type="checkbox"/> Physical Issue With Filter                             | <input type="checkbox"/> Freezing Drizzle - Moderate                    | <input type="checkbox"/> Ice Pellets Shower/s - Unspecified Intensity  |
| <input type="checkbox"/> Preliminary Data                                             | <input type="checkbox"/> Power Failure                                          | <input type="checkbox"/> Freezing Drizzle - Heavy                       | <input type="checkbox"/> Ice Pellets Shower/s - Light                  |
| <input type="checkbox"/> Missing Data                                                 | <input type="checkbox"/> Sample Diluted for Analysis                            | <input type="checkbox"/> Rain - Unspecified Intensity                   | <input type="checkbox"/> Ice Pellets Shower/s - Moderate               |
| <input type="checkbox"/> Invalid Data - Unspecified                                   | <input type="checkbox"/> Unmeasured Key Meteorological Parameter                | <input type="checkbox"/> Rain - Light                                   | <input type="checkbox"/> Ice Pellets Shower/s - Heavy                  |
| <input type="checkbox"/> Un-Flagged Data                                              | <input type="checkbox"/> Sample Not Analysed                                    | <input type="checkbox"/> Rain - Moderate                                | <input type="checkbox"/> Snow Pellets - Unspecified Intensity          |
| <input type="checkbox"/> Estimated Data - Unspecified                                 | <input type="checkbox"/> Operational Maintenance - Unspecified                  | <input type="checkbox"/> Rain - Heavy                                   | <input type="checkbox"/> Snow Pellets - Light                          |
| <input type="checkbox"/> Estimated Data - Measured Negative Value                     | <input type="checkbox"/> Calibration                                            | <input type="checkbox"/> Rain Shower/s - Unspecified Intensity          | <input type="checkbox"/> Snow Pellets - Moderate                       |
| <input type="checkbox"/> Estimated Data - No Value Detected                           | <input type="checkbox"/> Accuracy Check                                         | <input type="checkbox"/> Rain Shower/s - Light                          | <input type="checkbox"/> Snow Pellets - Heavy                          |
| <input type="checkbox"/> Estimated Data - Value Below Detection Limit                 | <input type="checkbox"/> Blank Check                                            | <input type="checkbox"/> Rain Shower/s - Moderate                       | <input type="checkbox"/> Snow Pellets Shower/s - Unspecified Intensity |
| <input type="checkbox"/> Estimated Data - Value Above Detection Limit                 | <input type="checkbox"/> Detection Limits Check                                 | <input type="checkbox"/> Rain Shower/s - Heavy                          | <input type="checkbox"/> Snow Pellets Shower/s - Light                 |
| <input type="checkbox"/> Estimated Data - Value Substituted from Secondary Monitor    | <input type="checkbox"/> Precision Check                                        | <input type="checkbox"/> Freezing Rain - Unspecified Intensity          | <input type="checkbox"/> Snow Pellets Shower/s - Moderate              |
| <input type="checkbox"/> Estimated Data - Multiple Parameters Aggregated              | <input type="checkbox"/> Retention Time Check                                   | <input type="checkbox"/> Freezing Rain - Light                          | <input type="checkbox"/> Snow Pellets Shower/s - Heavy                 |
| <input type="checkbox"/> Extreme/Irregular Data - Unspecified                         | <input type="checkbox"/> Span Check                                             | <input type="checkbox"/> Freezing Rain - Moderate                       | <input type="checkbox"/> Snow Grains - Unspecified Intensity           |
| <input type="checkbox"/> Data Does Not Meet Internal Network Quality Control Criteria | <input type="checkbox"/> Zero Check                                             | <input type="checkbox"/> Freezing Rain - Heavy                          | <input type="checkbox"/> Snow Grains - Light                           |
| <input type="checkbox"/> High Variability of Data                                     | <input type="checkbox"/> Instrumental Inspection                                | <input type="checkbox"/> Freezing Rain Shower/s - Unspecified Intensity | <input type="checkbox"/> Snow Grains - Moderate                        |
| <input type="checkbox"/> Irregular Data Manually Screened and Accepted                | <input type="checkbox"/> Instrumental Repair                                    | <input type="checkbox"/> Freezing Rain Shower/s - Light                 | <input type="checkbox"/> Snow Grains - Heavy                           |
| <input type="checkbox"/> Irregular Data Manually Screened and Rejected                | <input type="checkbox"/> Quality Control Audit                                  | <input type="checkbox"/> Freezing Rain Shower/s - Moderate              | <input type="checkbox"/> Diamond Dust - Unspecified Intensity          |
| <input type="checkbox"/> Negative Value                                               | <input type="checkbox"/> Data Formatting/Processing Issue                       | <input type="checkbox"/> Freezing Rain Shower/s - Heavy                 | <input type="checkbox"/> Diamond Dust - Light                          |
| <input type="checkbox"/> No Value Detected                                            | <input type="checkbox"/> Corrected Data Formatting/Processing Issue             | <input type="checkbox"/> Snow - Unspecified Intensity                   | <input type="checkbox"/> Diamond Dust - Moderate                       |
| <input type="checkbox"/> Reconstructed/Recalculated Data                              | <input type="checkbox"/> Aggregation/Representation Issue - Unspecified         | <input type="checkbox"/> Snow - Light                                   | <input type="checkbox"/> Diamond Dust - Heavy                          |
| <input type="checkbox"/> Value Close to Detection Limit                               | <input type="checkbox"/> Data Window Completeness < 90%                         | <input type="checkbox"/> Snow - Moderate                                | <input type="checkbox"/> Glaze                                         |
| <input type="checkbox"/> Value Below Acceptable Range                                 | <input type="checkbox"/> Data Window Completeness < 75%                         | <input type="checkbox"/> Snow - Heavy                                   | <input type="checkbox"/> Rime                                          |
| <input type="checkbox"/> Value Above Acceptable Range                                 | <input type="checkbox"/> Data Window Completeness < 66%                         | <input type="checkbox"/> Snow Shower/s - Unspecified Intensity          | <input type="checkbox"/> Thunderstorm                                  |
| <input type="checkbox"/> Value Below Detection Limit                                  | <input type="checkbox"/> Data Window Completeness < 50%                         | <input type="checkbox"/> Snow Shower/s - Light                          | <input type="checkbox"/> Funnel Cloud/s                                |
| <input type="checkbox"/> Value Above Detection Limit                                  | <input type="checkbox"/> Data Window Completeness < 25%                         | <input type="checkbox"/> Snow Shower/s - Moderate                       | <input type="checkbox"/> Squalls                                       |
| <input type="checkbox"/> Measurement Issue - Unspecified                              | <input type="checkbox"/> >= 75% of Measurements in Window Below Detection Limit | <input type="checkbox"/> Snow Shower/s - Heavy                          | <input type="checkbox"/> Tropical Cyclone (Cyclone/Hurricane/Typhoon)  |
| <input type="checkbox"/> Chemical Issue                                               | <input type="checkbox"/> >= 50% of Measurements in Window Below Detection Limit | <input type="checkbox"/> Hail - Unspecified Intensity                   | <input type="checkbox"/> Duststorm                                     |
| <input type="checkbox"/> Erroneous Sampling Operation                                 | <input type="checkbox"/> No Significant Weather                                 | <input type="checkbox"/> Hail - Light                                   | <input type="checkbox"/> Sandstorm                                     |
| <input type="checkbox"/> Extreme Internal Instrument Meteorological Conditions        | <input type="checkbox"/> Precipitation - Unspecified Intensity                  | <input type="checkbox"/> Hail - Moderate                                | <input type="checkbox"/> Dust/Sand Whirls                              |
| <input type="checkbox"/> Extreme Ambient Laboratory Meteorological Conditions         | <input type="checkbox"/> Precipitation - Light                                  | <input type="checkbox"/> Hail - Heavy                                   | <input type="checkbox"/> High Winds                                    |
| <input type="checkbox"/> Extreme External Meteorological Conditions                   | <input type="checkbox"/> Precipitation - Moderate                               | <input type="checkbox"/> Hail Shower/s - Unspecified Intensity          | <input type="checkbox"/> No Atmospheric Obscuration                    |
| <input type="checkbox"/> Extreme Sample Transport Conditions                          | <input type="checkbox"/> Precipitation - Heavy                                  | <input type="checkbox"/> Hail Shower/s - Light                          | <input type="checkbox"/> Atmospheric Obscuration - Unknown             |
| <input type="checkbox"/> Invalid Flow Rate                                            | <input type="checkbox"/> Drizzle - Unspecified Intensity                        | <input type="checkbox"/> Hail Shower/s - Moderate                       | <input type="checkbox"/> Dust                                          |
| <input type="checkbox"/> Human Error                                                  | <input type="checkbox"/> Drizzle - Light                                        | <input type="checkbox"/> Hail Shower/s - Heavy                          | <input type="checkbox"/> Blowing Dust                                  |
| <input type="checkbox"/> Matrix Effect                                                | <input type="checkbox"/> Drizzle - Moderate                                     | <input type="checkbox"/> Ice Pellets - Unspecified Intensity            | <input type="checkbox"/> Drifting Dust                                 |
| <input type="checkbox"/> Mechanical Issue/Non-Operational Equipment                   | <input type="checkbox"/> Drizzle - Heavy                                        | <input type="checkbox"/> Ice Pellets - Light                            | <input type="checkbox"/> Sand                                          |
| <input type="checkbox"/> No Technician                                                | <input type="checkbox"/> Freezing Drizzle - Unspecified Intensity               | <input type="checkbox"/> Ice Pellets - Moderate                         | <input type="checkbox"/> Blowing Sand                                  |
| <input type="checkbox"/> Operational Maintenance Check Issue                          | <input type="checkbox"/> Freezing Drizzle - Light                               | <input type="checkbox"/> Ice Pellets - Heavy                            | <input type="checkbox"/> Drifting Sand                                 |

# Filtering by period

## Select Data Periods

|           | <i>K</i>                            | <i>R</i>                 |
|-----------|-------------------------------------|--------------------------|
| Daytime   | <input type="checkbox"/>            | <input type="checkbox"/> |
| Nighttime | <input type="checkbox"/>            | <input type="checkbox"/> |
| Weekday   | <input type="checkbox"/>            | <input type="checkbox"/> |
| Weekend   | <input type="checkbox"/>            | <input type="checkbox"/> |
| Spring    | <input type="checkbox"/>            | <input type="checkbox"/> |
| Summer    | <input checked="" type="checkbox"/> | <input type="checkbox"/> |
| Autumn    | <input type="checkbox"/>            | <input type="checkbox"/> |
| Winter    | <input checked="" type="checkbox"/> | <input type="checkbox"/> |

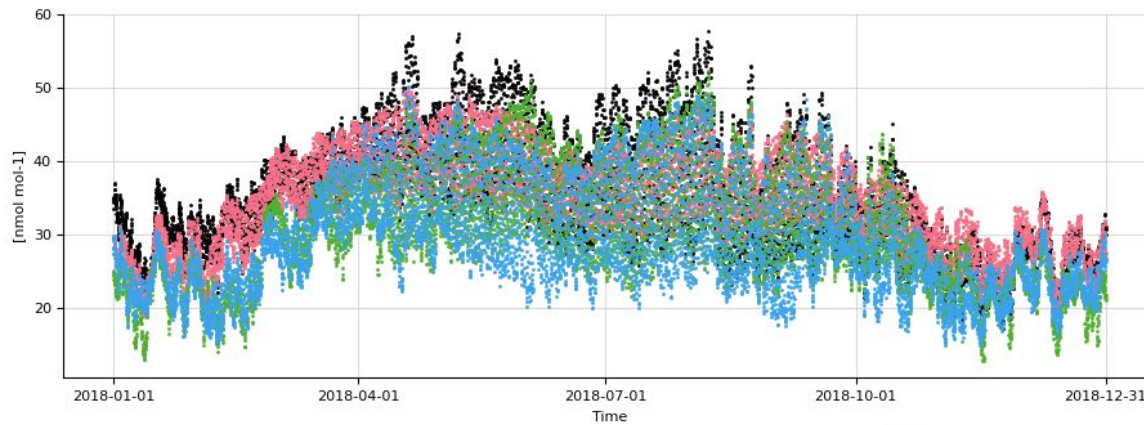


Filtering to keep Winter and Summer seasons

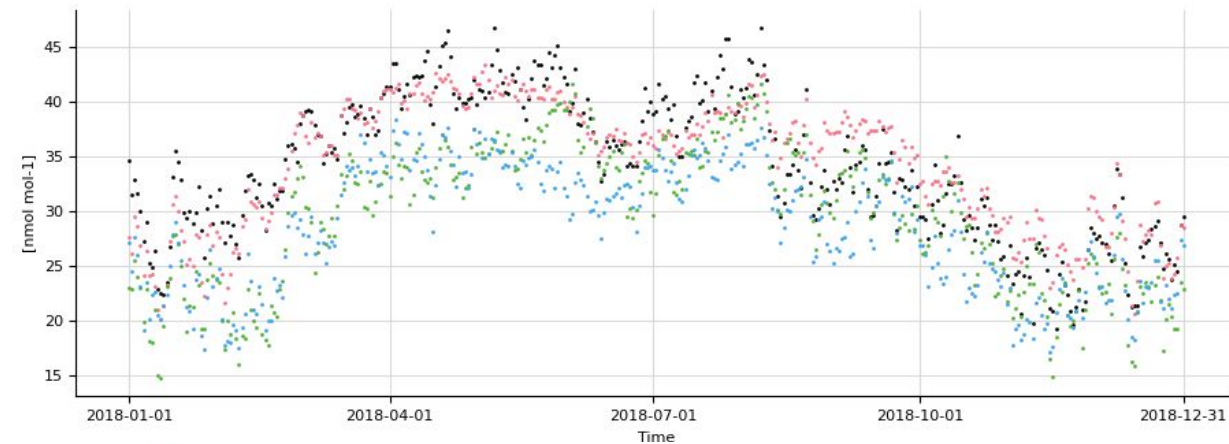
# Temporal resampling

The temporal resolution of data can be resampled on the fly, going from finer to coarser resolutions, e.g. hourly to daily.

## Hourly



## Daily



● observations    ● cams61\_match\_ph2-eu-000    ● cams61\_monarch\_ph3-eu-000

# Spatial collocation

Ozone: 1731 stations

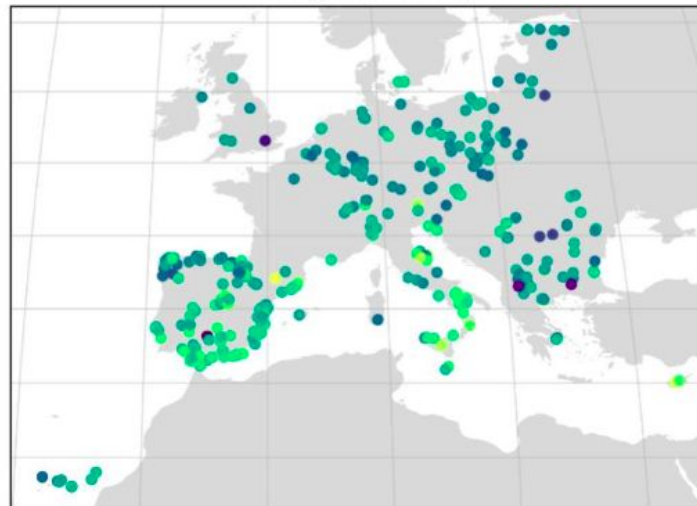


CO: 635 stations

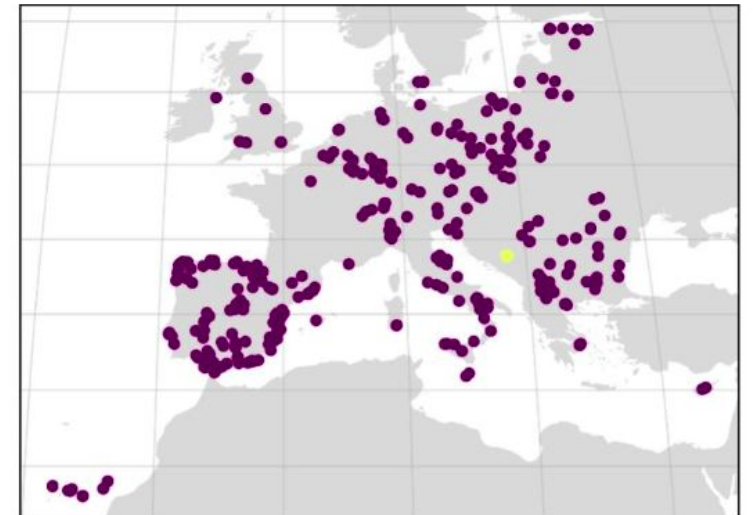


Without spatial  
collocation

Ozone: 366 stations



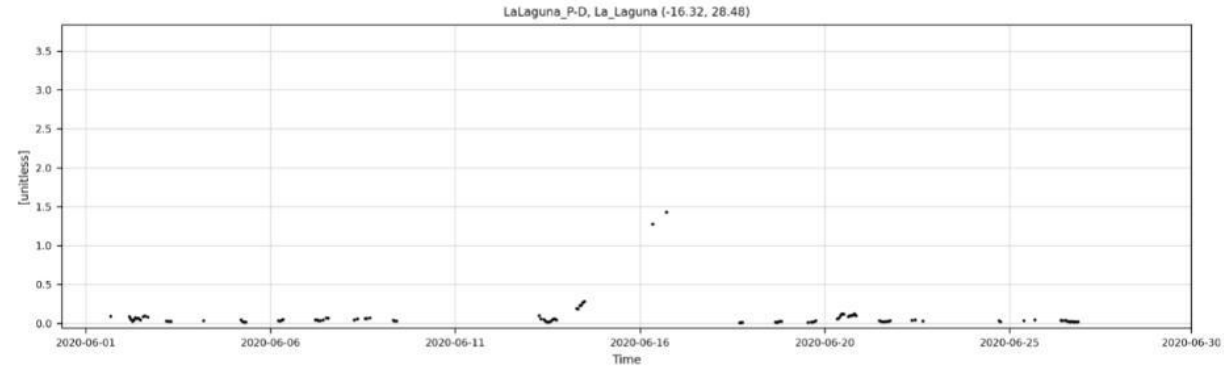
CO: 366 stations



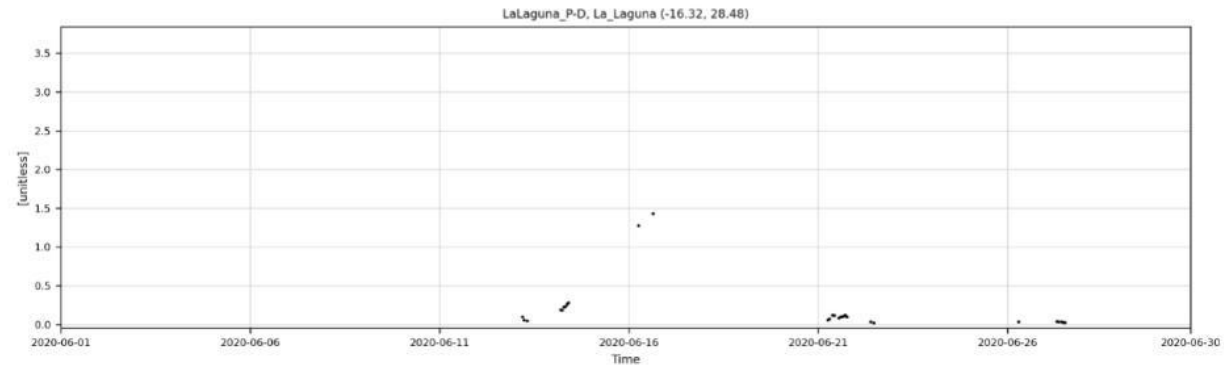
With spatial  
collocation

# Multispecies filtering

**AOD without  
multispecies  
filtering**

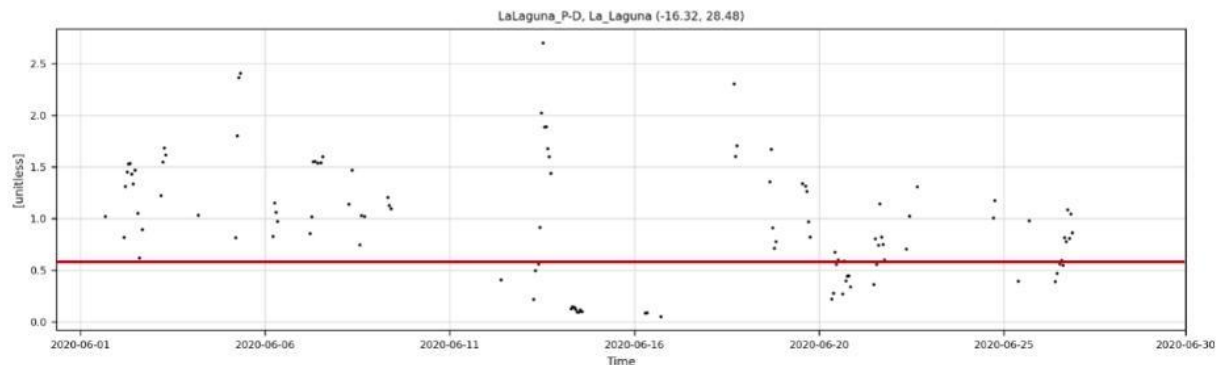


**AOD with  
multispecies  
filtering**

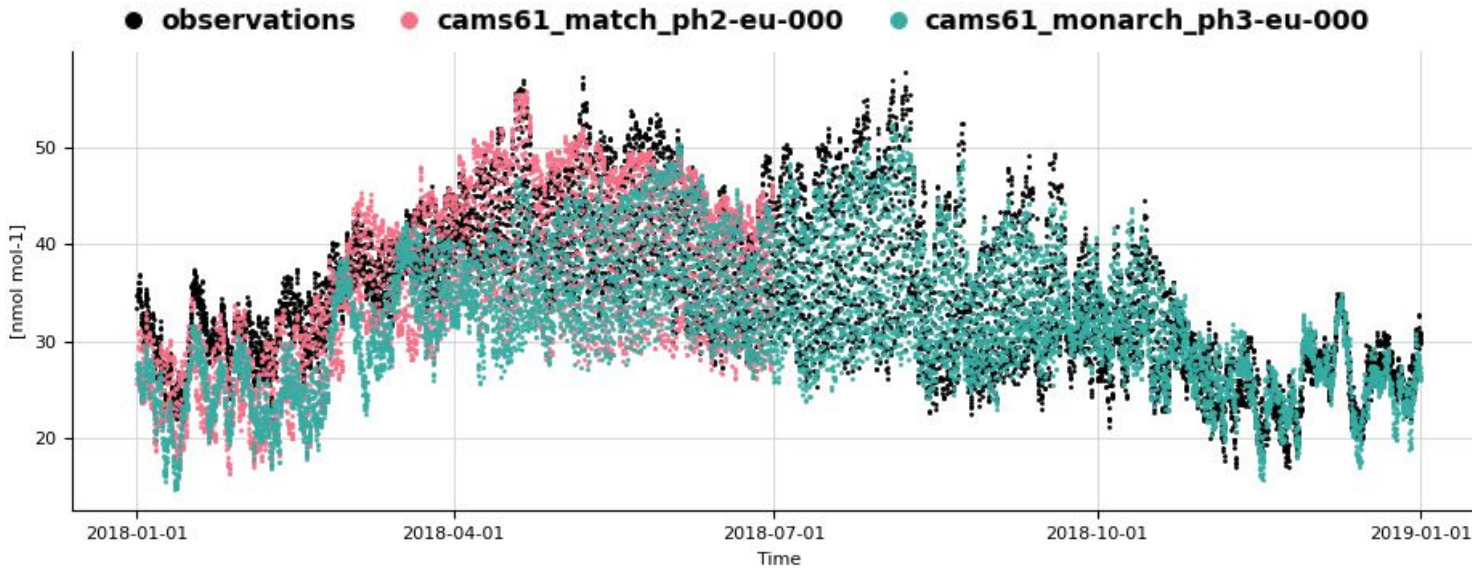


If Angstrom exponent is  $> 0.6$ ,  
AOD is filtered out

**Angstrom  
exponent**

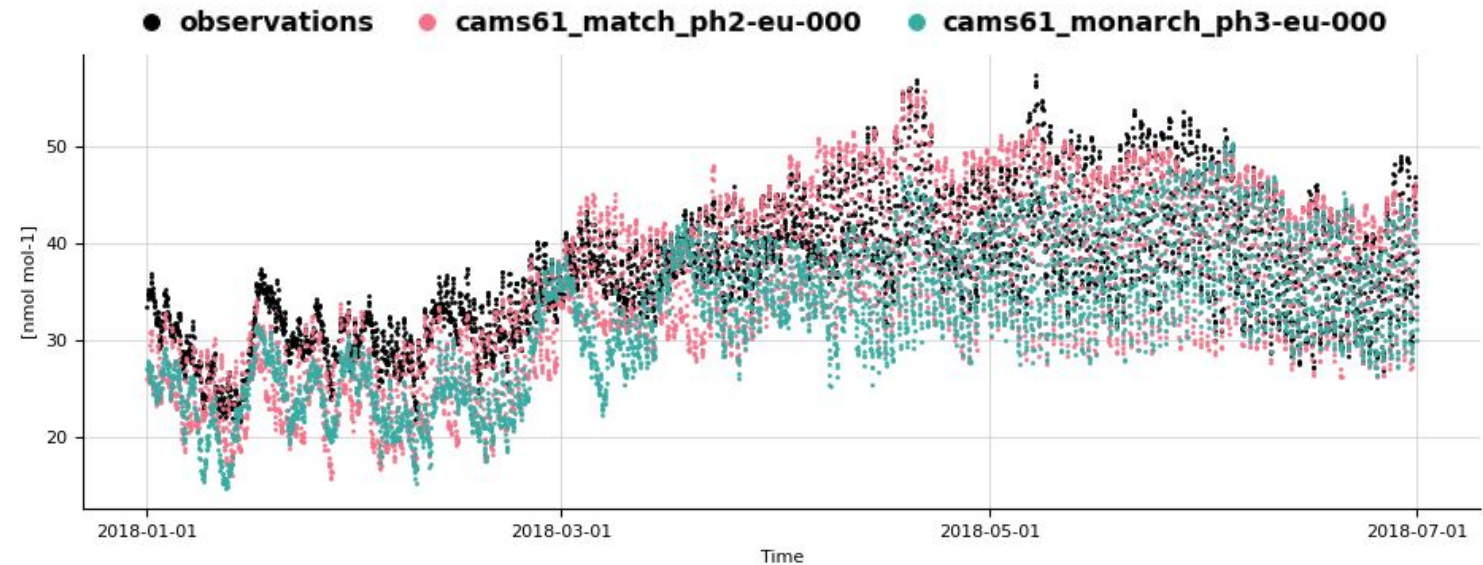


# Temporal colocation



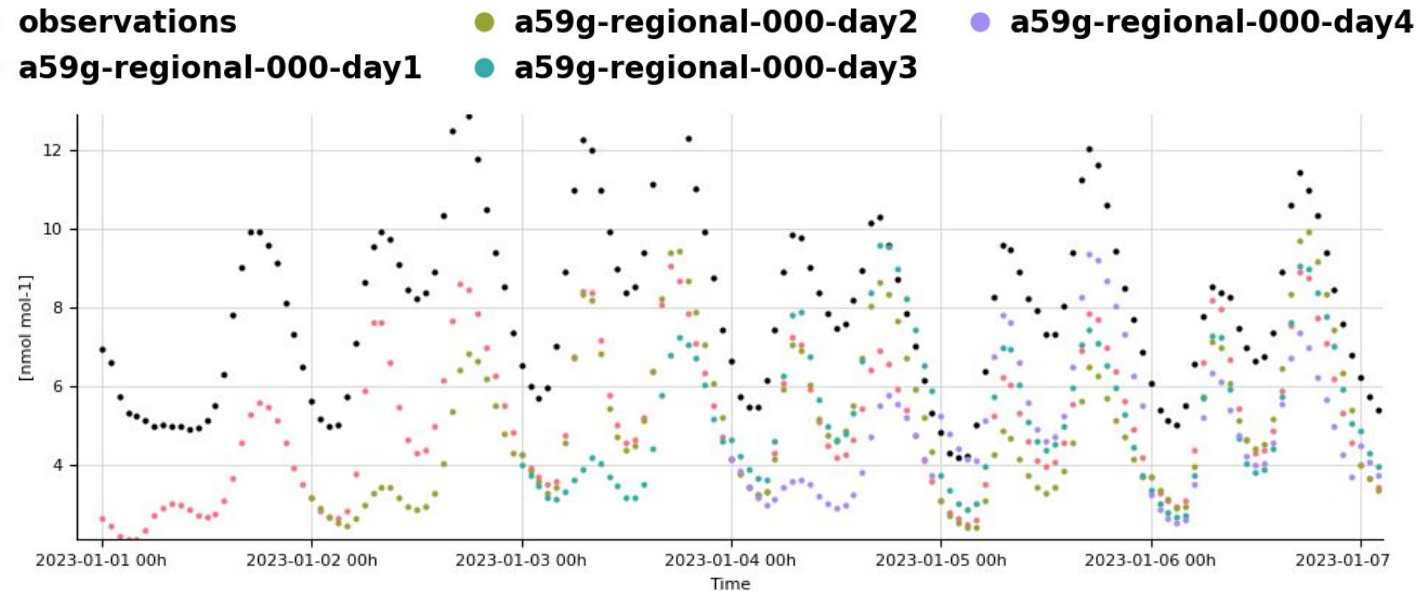
Without temporal  
colocation

With temporal  
colocation



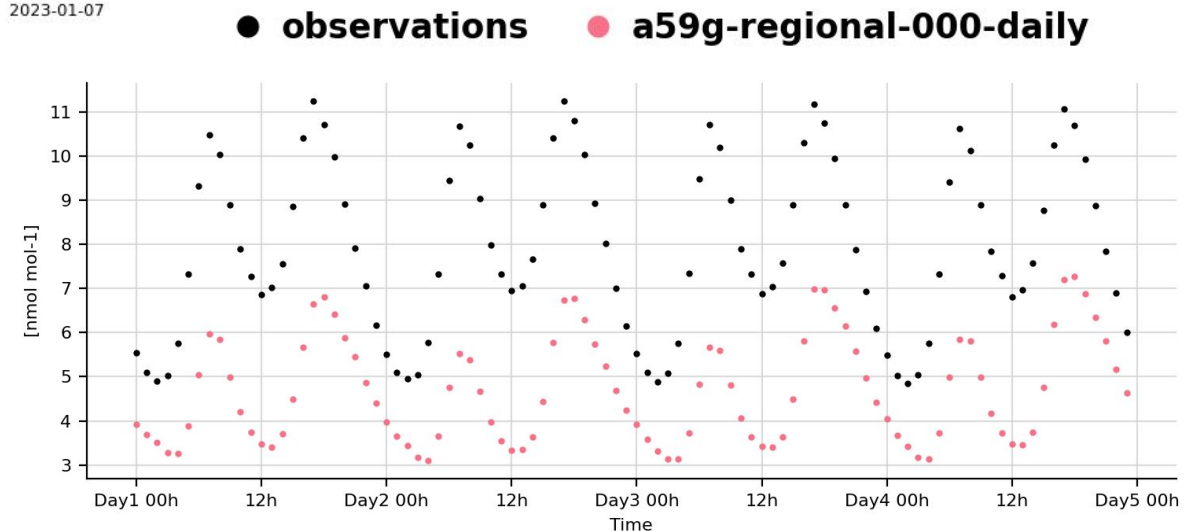
# Forecast

Providentia can be used to evaluate forecast model data, either by forecast day, or mean across forecast days.



By forecast day

Mean across forecast  
days



# Available statistics

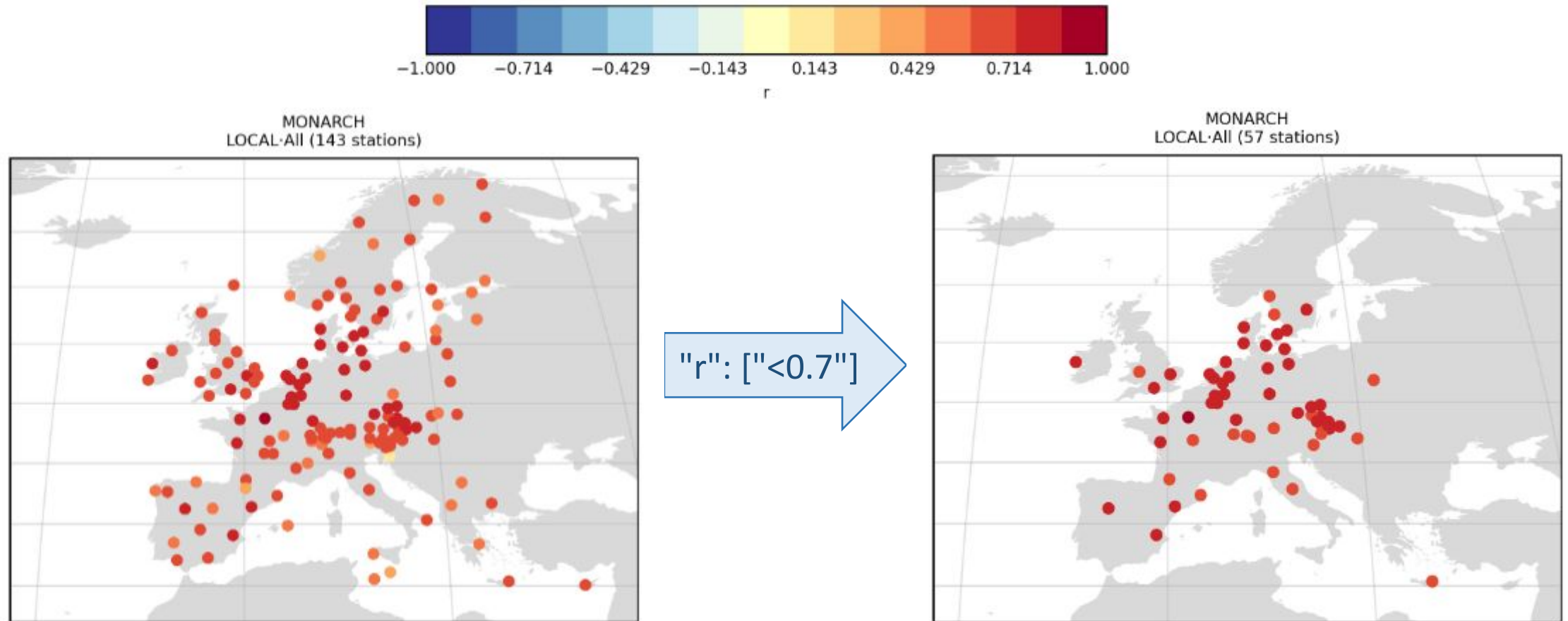
## Basic statistics

| Statistic                                 | Meaning                      |
|-------------------------------------------|------------------------------|
| Mean                                      | Mean                         |
| StdDev                                    | Standard deviation           |
| Var                                       | Variance                     |
| Min                                       | Minimum                      |
| Max                                       | Maximum                      |
| Data%                                     | Data availability            |
| Exceedances                               | Number of exceedances        |
| p1, p5, p10, p25, p50, p75, p90, p95, p99 | Percentiles                  |
| NStations                                 | Number of stations           |
| MDA8                                      | Daily maximum 8 hour average |

## Model bias statistics

| Statistic      | Meaning                                                            |
|----------------|--------------------------------------------------------------------|
| MB             | Mean bias                                                          |
| NMB            | Normalised mean bias                                               |
| ME             | Mean error                                                         |
| NME            | Normalised mean error                                              |
| MNB            | Mean normalised bias                                               |
| MNE            | Mean normalised error                                              |
| MFB            | Mean fractional bias                                               |
| MFE            | Mean fractional error                                              |
| RMSE           | Root mean square error                                             |
| NRMSE          | Normalised root mean square error                                  |
| COE            | Coefficient of efficiency                                          |
| FAC2           | Fraction of model values within a factor of two of observed values |
| IOA            | Index of agreement                                                 |
| R              | Pearson correlation coefficient                                    |
| R <sup>2</sup> | Coefficient of determination                                       |
| UPA            | Unpaired peak accuracy                                             |

# Remove stations by statistics



# Plot customisation

To update the style of plots in the dashboard, reports and library, edit the plot characteristics in **settings/plot\_characteristics.yaml**.

settings/plot\_characteristics.yaml

All

Dashboard

Report

Library

```
"timeseries":
{
  "grid": { "axis": "both", "color": "lightgrey", "alpha": 0.8 },
  "threshold_line":
    { "linestyle": "--", "linewidth": 1.0, "color": "red", "zorder": 0 },
  ...
  "dashboard":
    {"remove_spines": ["top", "right"], ...},
  "report":
    {"orientation": "landscape", ...},
  "library":
    {"figsize": [17, 6], ...},}
```

Reference: <https://providentia.readthedocs.io/en/latest/Plot-customisation.html>

# Conclusions

# Conclusions

- Providentia is an open-source Python framework for evaluating and analyzing in-situ observations and atmospheric model outputs.
- It is widely used by the atmospheric composition group at Barcelona Supercomputing Center.
- Its main goal is to streamline the processing, visualization, and evaluation of large observational and modelling datasets within a unified platform.
- It supports multiple usage modes, including an interactive dashboard, automated PDF report generation, and integration into custom Python workflows.
- It is highly customizable and extensible, with support for multiple data sources, enabling its use across diverse atmospheric and environmental applications.

# Thank you for your attention!

More information at:

<https://github.com/bsc-es/providentia>

**Contact:** [alba.vilanova@bsc.es](mailto:alba.vilanova@bsc.es)