

Evaluation of AQ Monitoring Network

Proposed exercise using the interactive webtool - MoNet

Joana Soares, Christoffer Stoll, Islen Vallejo and Leonor Tarrasón

FAIRMODE Technical Meeting 18th – 20th October 2022



Norsk institutt for luftforskning
Norwegian Institute for Air Research

CT8 - Proposal for a monitoring network evaluation exercise

Evaluation of the air quality (AQ) monitoring network in a given area

- 1) **Goal:** identify how well the current monitoring network represents the different pollution regimes in a given area

- 2) **Method:** Hierarchical Clustering method through NILUs Webtool MoNet
 - Identification of outliers and possible reasons for the unique performance of specific stations
 - Evaluation of station type in terms of pollution regime

What is MoNet?

A web-based **screening tool** developed to assist on the AQ monitoring network optimization and decision-making process:

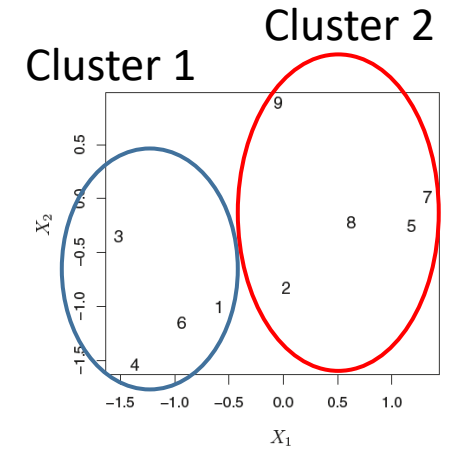
- **assessing (dis)similarity** between time series
- **Identifying outliers** (unique signal or, e.g., errors in Measurements)
- **ranking stations** according to their similarity and indicate potential redundancies (TIER 2)
- **inferring the area of representativeness** for a single station in a large area (TIER 3)

The **tool can also be used for other purposes:**

- **model validation purposes** by comparing the (dis)similarity analysis based on modelling data and observations (TIER 3) - This can be tested in future FAIRMODE exercises, not now
- **(Re)designing** monitoring networks based on modelling data

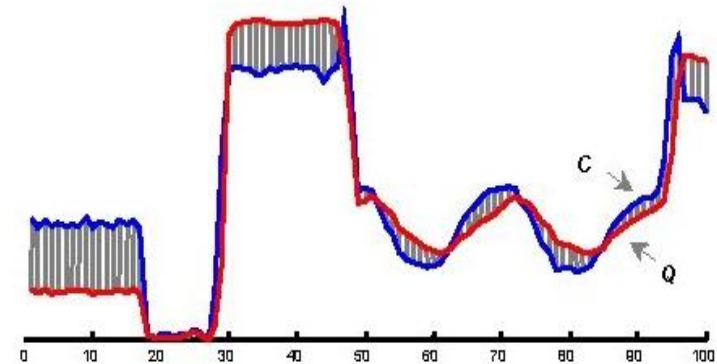
Monitoring network evaluation tool

- **Tool built-on a statistical method (hierarchical clustering)** to determine the natural groupings of data (clusters) done on basis of *(dis)similarities*.
- The assessment compares time series based on a **metric**:
 - 1-R, R=Pearson correlation coefficient (Solazzo and Galmarini, 2015): **time variation**
 - Euclidean distance (EuD) (Soares et al, 2018): **magnitude variation**
 - **(1-R) x EuD** **time and magnitude variation**



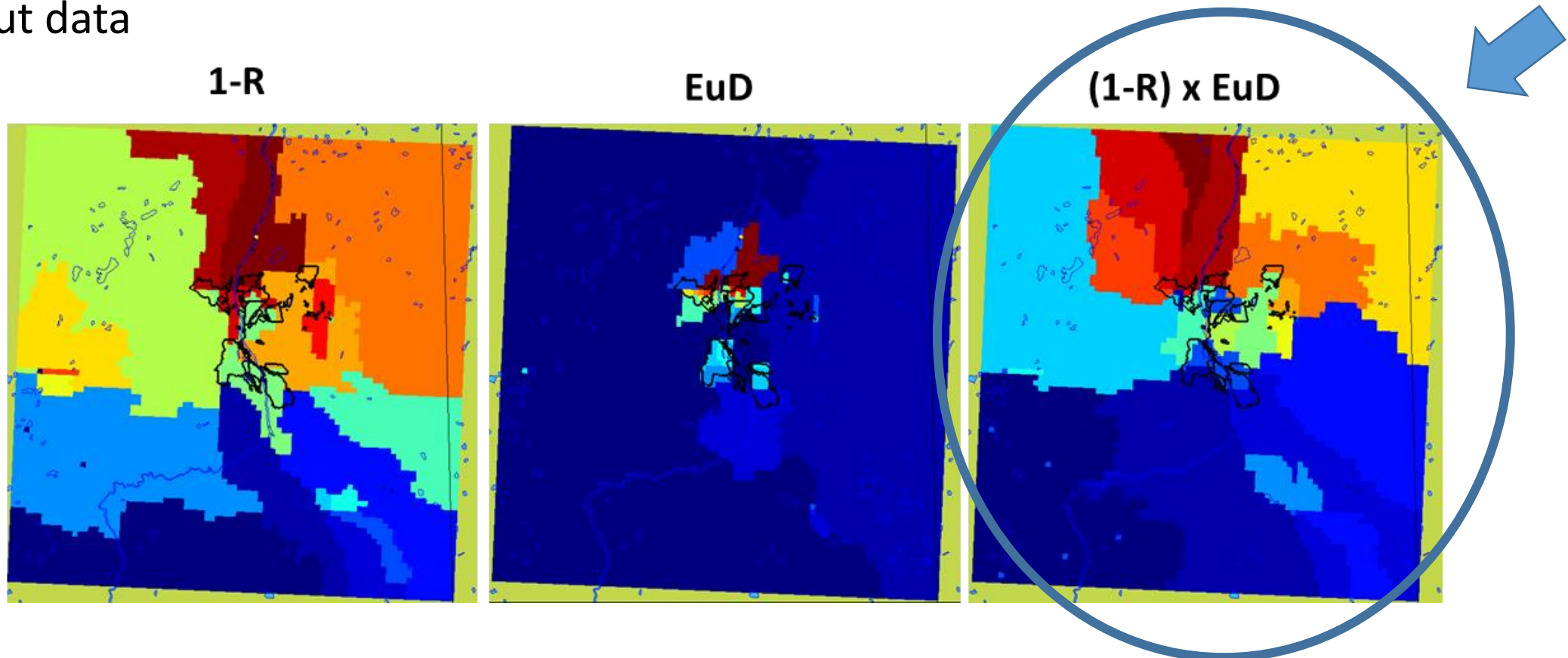
$$r_{cq} = \frac{\sum_j (c_j - m_c) (q_j - m_q)}{\sqrt{\sum_j (c_j - m_c)^2 \sum_j (q_j - m_q)^2}}$$

$$EuD_{cq} = \sqrt{\sum_j (c_j - q_j)^2}$$



Choice of Metric

- Depends on the purpose of the monitoring network.
- Using different metrics allows the assessment of different properties of the input data



Monitoring network evaluation tool

- **Tool built-on a statistical method (hierarchical clustering)** to determine the natural groupings of data (clusters) done on basis of (dis)similarities.

Advantage:

- does not depend on the no. of clusters to be searched or a starting configuration assignment.

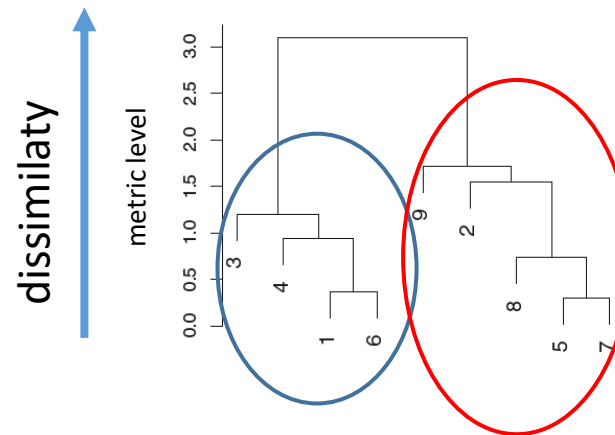
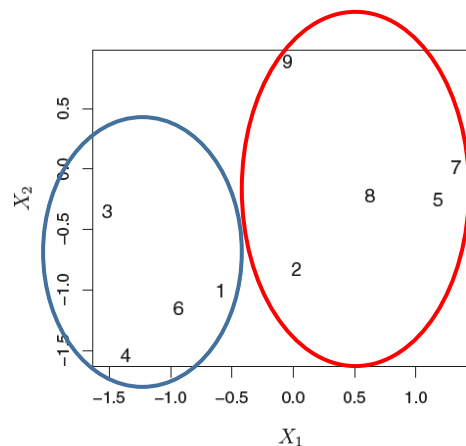
Caveats:

- Outcome depends on the quality of the data: error in measurements, data accuracy different sampling technologies, outliers, temporal coverage
- Computationally demanding for large number of observations (> 1e4 hourly time series)
- Requires temporal and spatial continuity (when used in combination with modelling data)

Hierarchical Clustering Representation

- Dendrogram is a 2D representation of the hierarchical clustering process
 - show the pattern of linkages between the data series while clustering occurs
 - provide a quantitative description of the degree of similarity between the data points
 - At the lowest level, each cluster contains a single observation. At higher levels of hierarchy, clusters are created by merging clusters at the lower level.

helps on the interpretation of the results

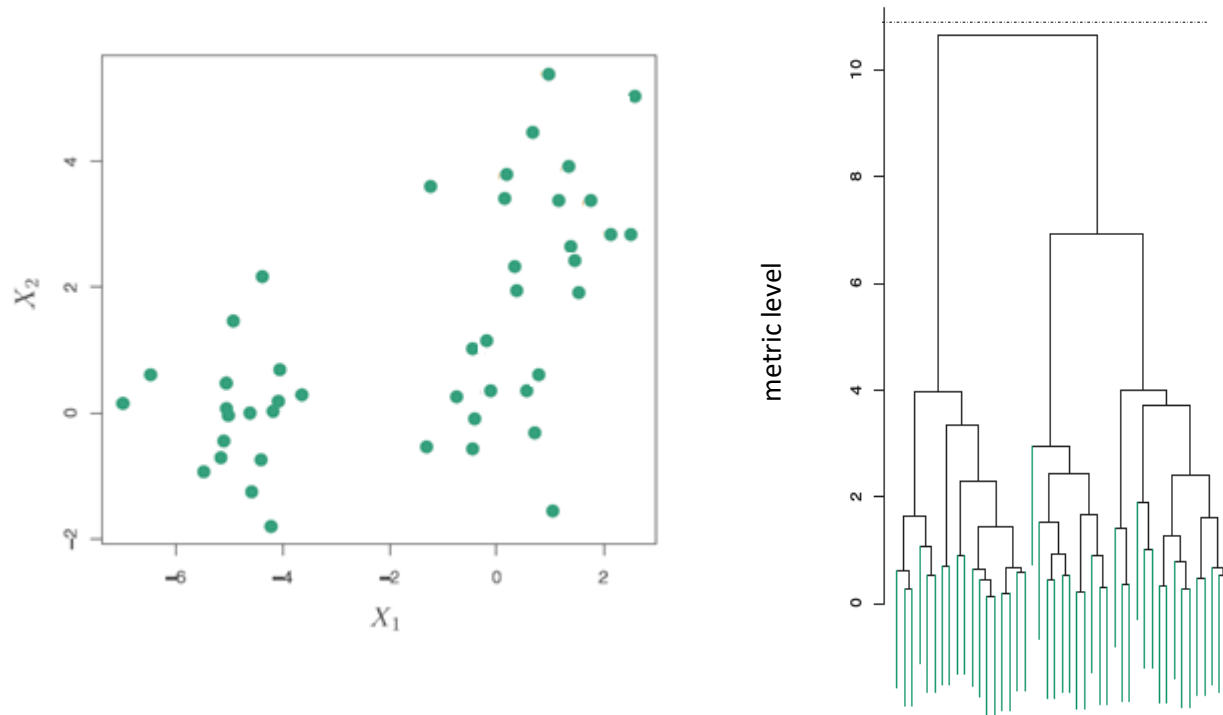


Reading the dendrogram:

- vertical lines: the level of (dis)similarity at every single stage of clustering
- horizontal lines: links between time series
- **less similar observations** are the ones joining up at the top

Interpreting a dendrogram

Qualitative analysis: the metric level is dependent of the type of metric, number of observation points available, number of records and period (time series), etc.



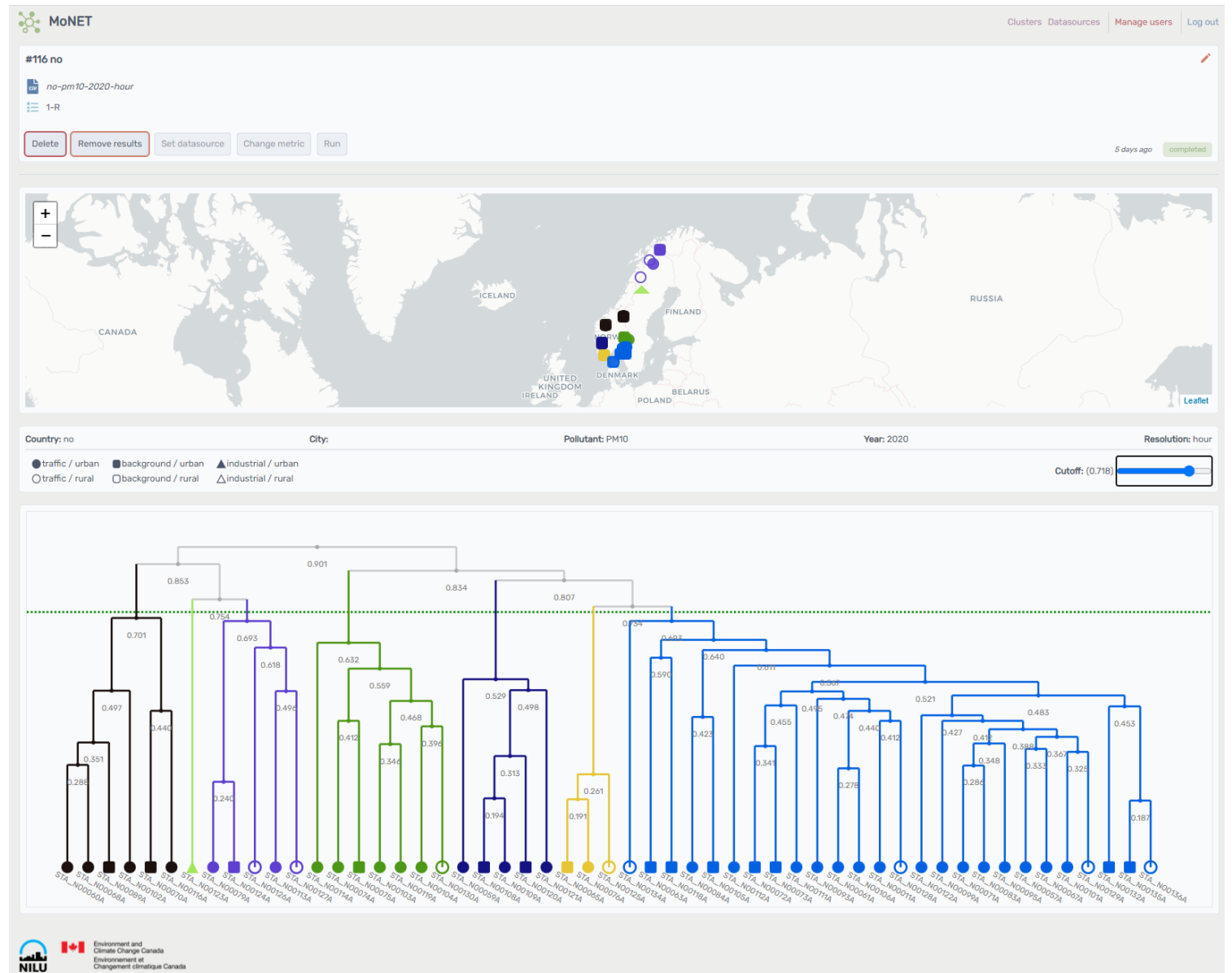
MoNet Webtool(<https://dev-monet.nilu.no/>)

Current Features:

- Input: user- defined observation data or download from EEA
- 3 (di)similarity metrics:
 - 1-R
 - $(1-R)*EuD$
 - Euclidean distance
- Selection of clustering level
- Results on map and dendrogram
- Multiple clustering jobs

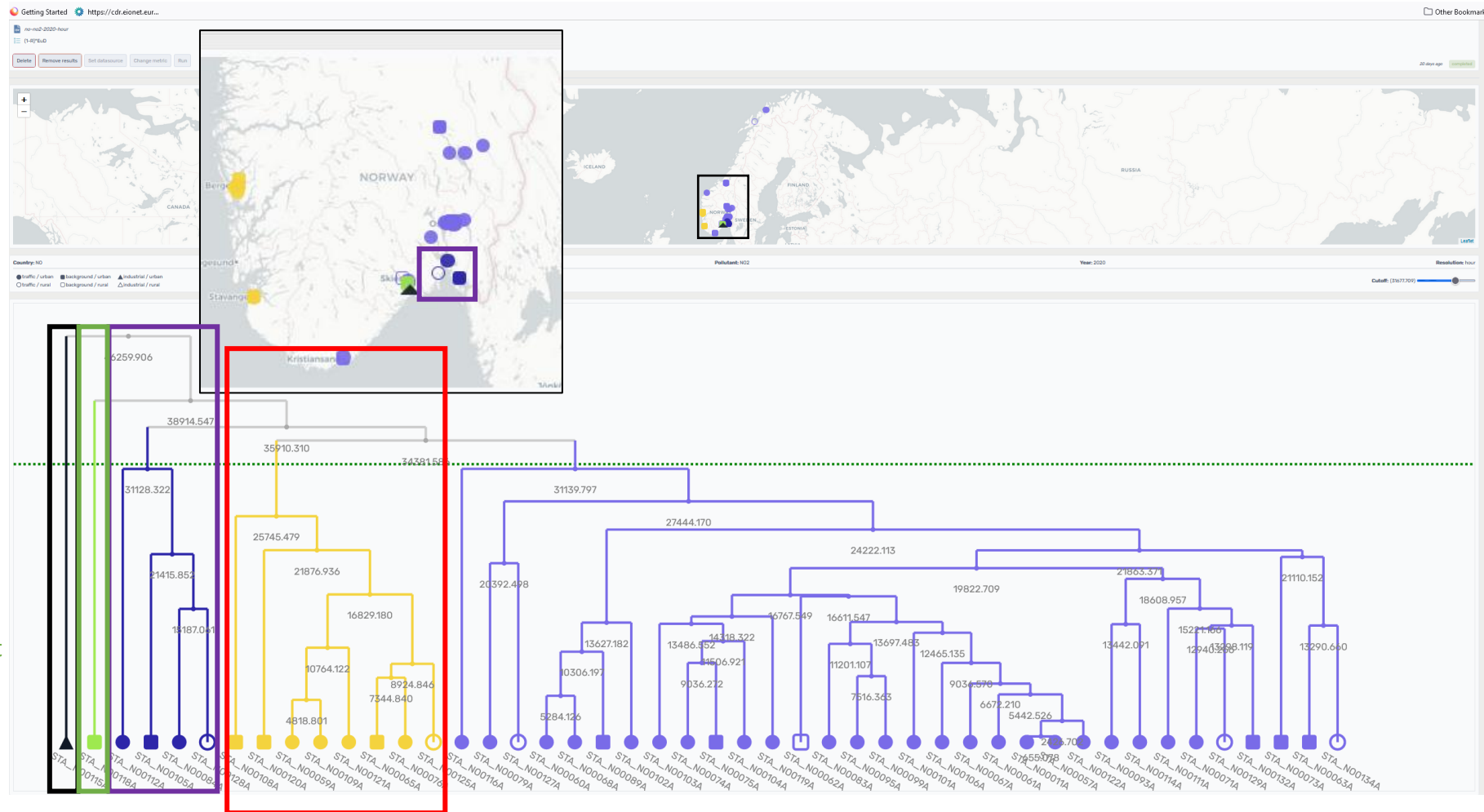
Features to come:

- More data source services
- Use gridded model data as input
- Ranking of stations
- Export results as image/ascii



Demonstration of the tool - <https://dev-monet.nilu.no/>

Visualise and interpret results



Industrial site with clear signals of NO₂ from the industrial source at Norcem Brevik

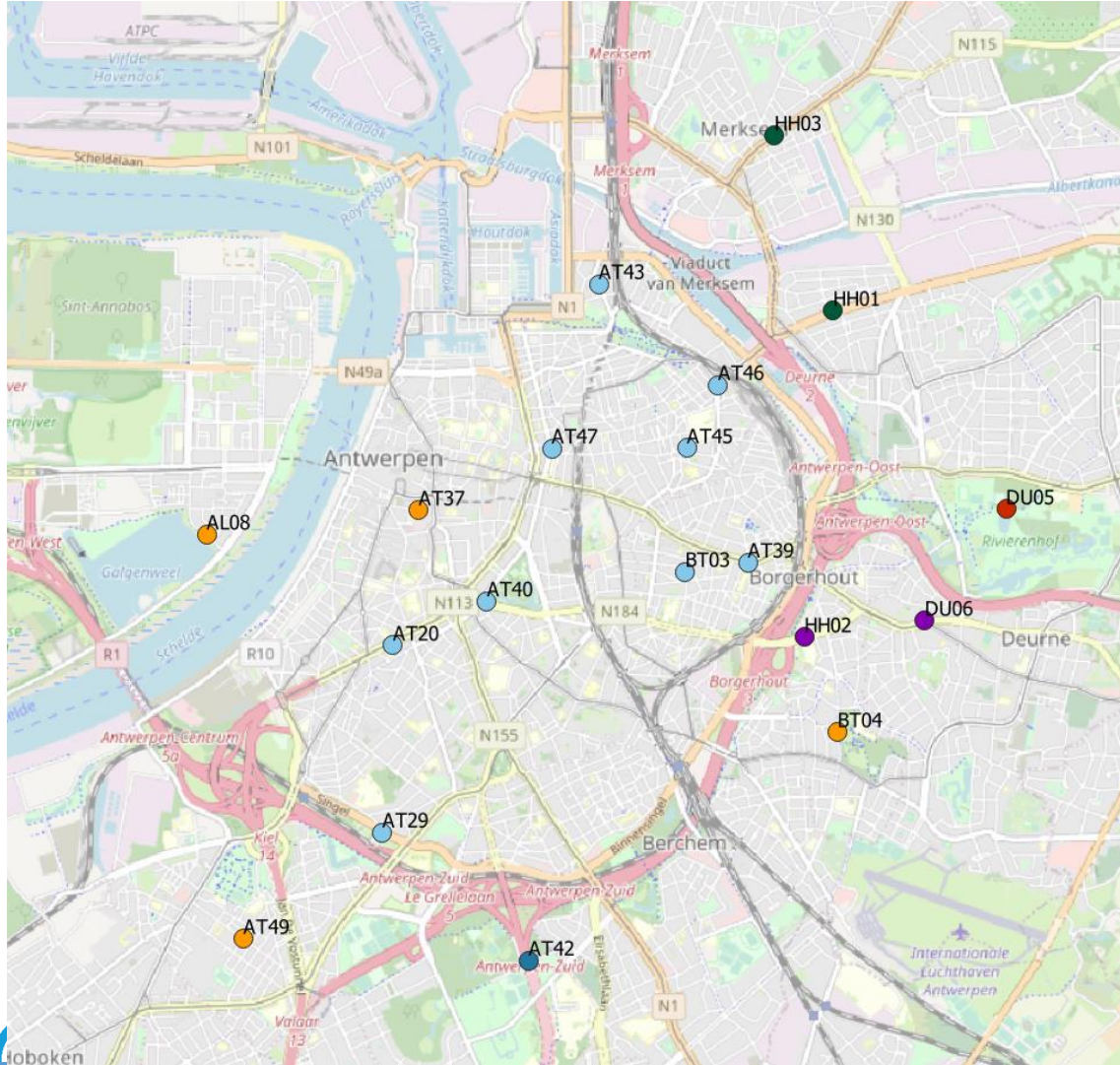
The highest average concentrations occur with wind from east to south, with the maximum from south-southeast (industries at Herøya); construction work the traffic

small cities mostly influenced by small shipping; similar orography and meteorology

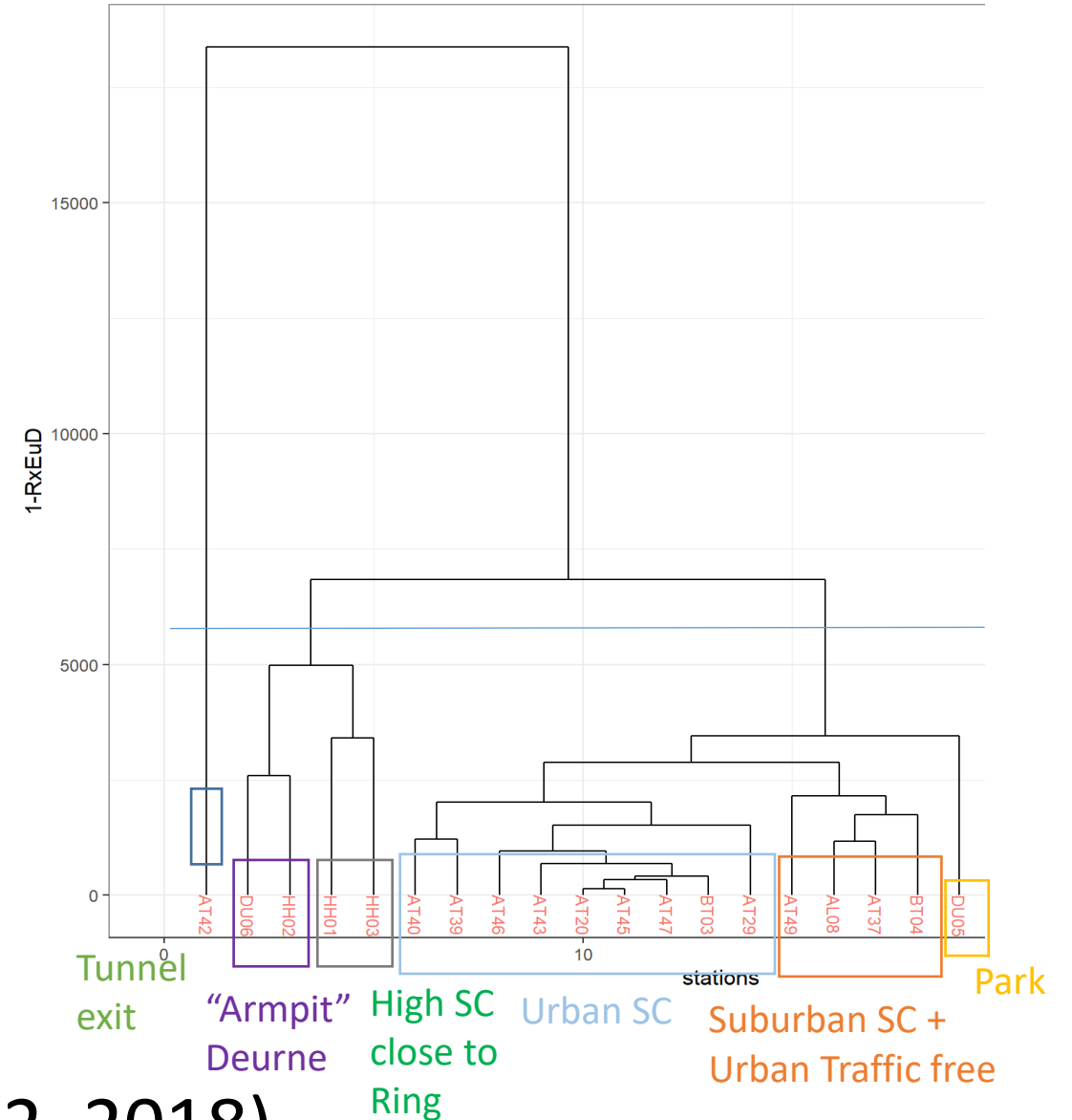
Meteorology is the key driver

Example for Norway (NO₂, 2020)





NO2 model at sampler locations (1year), metric: 1-RxEuD



Example for Antwerp (NO2, 2018)

Proposal for monitoring network evaluation exercise

Evaluation of monitoring network – proposed choices

- Components: PM2.5 , PM10, NO2 (free choice)
- Regions : city, air quality zone or country (choose two areas)
- Input data : Monitoring station hourly timeseries data for the whole year (75% temporal coverage)

Proposal for monitoring network evaluation exercise

Evaluation of monitoring network – proposed choices

- Components: PM2.5 , PM10, NO2 (free choice)
- Regions : city, air quality zone or country (**choose two areas**)
- Input data : Monitoring station hourly timeseries data for the whole year (75% temporal coverage, < 168 hours gaps)

CT8 - Proposal for monitoring network evaluation exercise

Evaluation of monitoring network - how to proceed

- **Indicate your interest** to participate via email to Joana (jos@nilu.no)
- You will then receive a **user name and password**
- Identify your input data and upload it in the Webtool
 - Use the Template for input data
 - Use the download capability at the NILUs webtool (retrieves from EEAs portal but only for a given city or country)
- Run the tool by **mid-November 2022**
- **Meet us at the result evaluation on-line workshop end of November/beggining of December 2022**

Curious?

Soares, J., Makar, P. A., Aklilu, Y., and Akingunola, A.: The use of hierarchical clustering for the design of optimized monitoring networks, *Atmos. Chem. Phys.*, 18, 65043-6566, <https://doi.org/10.5194/acp-18-6543-2018>.

Soares, J., Makar, P. A., Aklilu, Y., and Akingunola, A.: Hierarchical Clustering Network Analysis of Ambient Air Monitoring in Alberta: Phases 1 and 2, OSM technical report series, <http://environmentalmonitoring.alberta.ca/wp-content/uploads/2018/10/3.0-Summary-Evaluation-and-Integration-of-Atmospheric-Deposition-Monitoring-in-the-Athabasca-Oil-Sands-Region.pdf>

Leonor Tarrasón, Claudia Hak, Joana Soares, Håvard Vika Røen, Rune Ødegård and Leif Marsteen 2020, Assessing the spatial representativeness of air quality sampling point. Application of siting criteria and sampling point classification – Task 3 interim report Service Request 5 under Framework Contract ENV.C.3/FRA/2017/0012 Specific Contract: 07.0203/2018/793545/SFRA/ENV.C.3

Janssen, S., Tarrasón, L., Viaene, P., Hooyberghs, H., and Brookes, D., J. Soares 2020. Assessing the spatial representativeness of air quality sampling point – Sensitivity and feasibility tests for a tiered approach – Interim Report Task 1. Service Request 5 under Framework Contract ENV.C.3/FRA/2017/0012 Specific Contract: 07.0203/2018/793545/SFRA/ENV.C.3

Access to MoNet

<https://dev-monet.nilu.no/>

Send email for username and password

jos@nilu.no

STEP 1 - Template – input data

The screenshot displays the MoNET web interface across three browser windows. The top window shows the main dashboard with a list of clusters (#122, #120, #119, #103) and a search bar. The middle window shows the 'Add datasource' page with a list of data sources (#66, #60, #58, #52, #48) and their completion status. The bottom window shows a detailed view of the data sources, with a table listing the cluster name, data source name, and completion status. A red box highlights the 'Exclusion: Some data coverage was below 75%' message, and a green box highlights the 'completed' status. Text annotations explain that clustering requires 75% data coverage (gap-filling applied) and that 'queued (red)' indicates data below 75% coverage, while 'completed (green)' indicates data above 75% coverage.

Cluster ID	Data Source Name	Age	Completion Status
#66	es-no2-2020-hour	8 months ago	Some data below 75% coverage was excluded
#60	fi-pm10-2020-hour	10 months ago	Some data below 75% coverage was excluded
#58	no-oslo-pm10-2020-hour	10 months ago	completed
#52	se-pm10-2018-hour	10 months ago	Some data below 75% coverage was excluded
#48	europa-pm10-2020-hour	10 months ago	completed
#71	europa-pm10-2020-hour	about 1 hour ago	Exclusion: Some data coverage was below 75%
#70	no-no2-2020-hour	about 1 hour ago	Exclusion: Some data coverage was below 75%
#66	es-no2-2020-hour	8 months ago	Some data below 75% coverage was excluded
#60	fi-pm10-2020-hour	10 months ago	Some data below 75% coverage was excluded
#58	no-oslo-pm10-2020-hour	10 months ago	completed
#52	se-pm10-2018-hour	10 months ago	Some data below 75% coverage was excluded
#48	europa-pm10-2020-hour	10 months ago	completed

Clustering requires 75% data coverage (gap-filling applied)

queued (red) completed (green)



STEP 2 – create a new job /run the hierarchical cluster analysis

The image is a collage of five overlapping screenshots from the MoNET web application, illustrating the steps to create a new job for hierarchical cluster analysis. The screenshots are arranged in a cascading fashion from top-left to bottom-right.

- Top-left screenshot:** Shows the MoNET dashboard with a sidebar on the left containing a list of jobs (#71, #70, #66, #60, #58, #52, #48) and a search bar. A mouse cursor is pointing at the search bar.
- Second screenshot:** Shows the 'Add datasource' form. The 'Cluster name' field is highlighted, and an 'Add' button is visible next to it. A mouse cursor is pointing at the 'Add' button.
- Third screenshot:** Shows a table of existing jobs. The table has columns for job ID, cluster name, location, and status. The job #132 with cluster name 'NO-FAIRMODE' is highlighted. A mouse cursor is pointing at the job entry.
- Fourth screenshot:** Shows the configuration page for job #132. It displays the job name '#132 NO-FAIRMODE', a message 'No datasource found', and a '1-R' configuration. At the bottom, there are buttons for 'Delete', 'Remove results', 'Set datasource', 'Change metric', and 'Run'. A mouse cursor is pointing at the 'Run' button.
- Bottom-most screenshot:** Shows the 'Task 132' page. It displays the job name '#132 NO-FAIRMODE', a message 'No datasource found', and a '1-R' configuration. At the bottom, there are buttons for 'Delete', 'Remove results', 'Set datasource', 'Change metric', and 'Run'. A mouse cursor is pointing at the 'Run' button.

The MoNET logo and the Environment and Climate Change Canada logo are visible in the bottom-left corner of each screenshot.

STEP 2— create a new job (2)

The image displays three overlapping screenshots of the MoNET web application interface, illustrating the process of creating a new job configuration. Each screenshot shows a browser window with the URL `https://dev-monet.nilu.no/task/132`, `https://dev-monet.nilu.no/task/135`, and `https://dev-monet.nilu.no/task/135` respectively.

Top Screenshot (Task 132): Shows a job configuration for "#132 NO-FAIRMODE". The "No datasource found" message is present. The "Set datasource" button is highlighted with a mouse cursor.

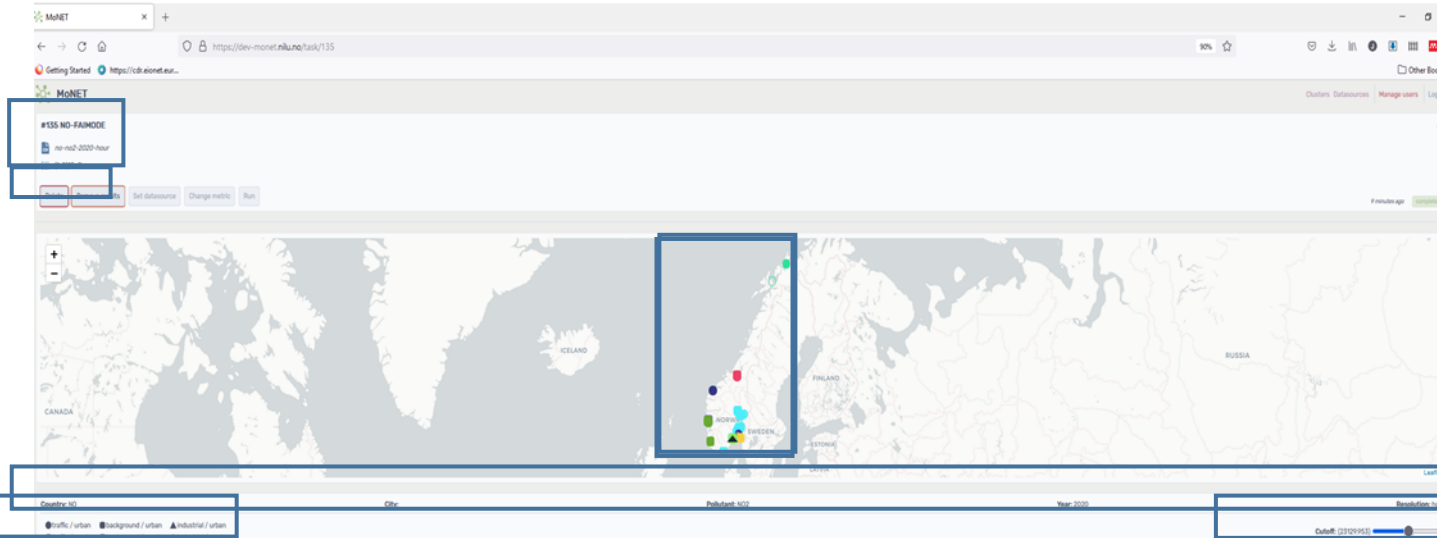
Middle Screenshot (Task 135): Shows a job configuration for "#135 NO-FAIRMODE". The "no-no2-2020-hour" metric is selected. The "Change metric" button is highlighted with a mouse cursor.

Bottom Screenshot (Task 135): Shows the final job configuration for "#135 NO-FAIRMODE". The "no-no2-2020-hour" metric is selected, and the "Run" button is highlighted with a mouse cursor.

The MoNET logo and the text "Environment and Climate Change Canada / Environnement et Changement climatique Canada" are visible in the bottom left corner of each screenshot.

STEP 3 - Visualise and interpret results

Delete the entire run or delete only the results



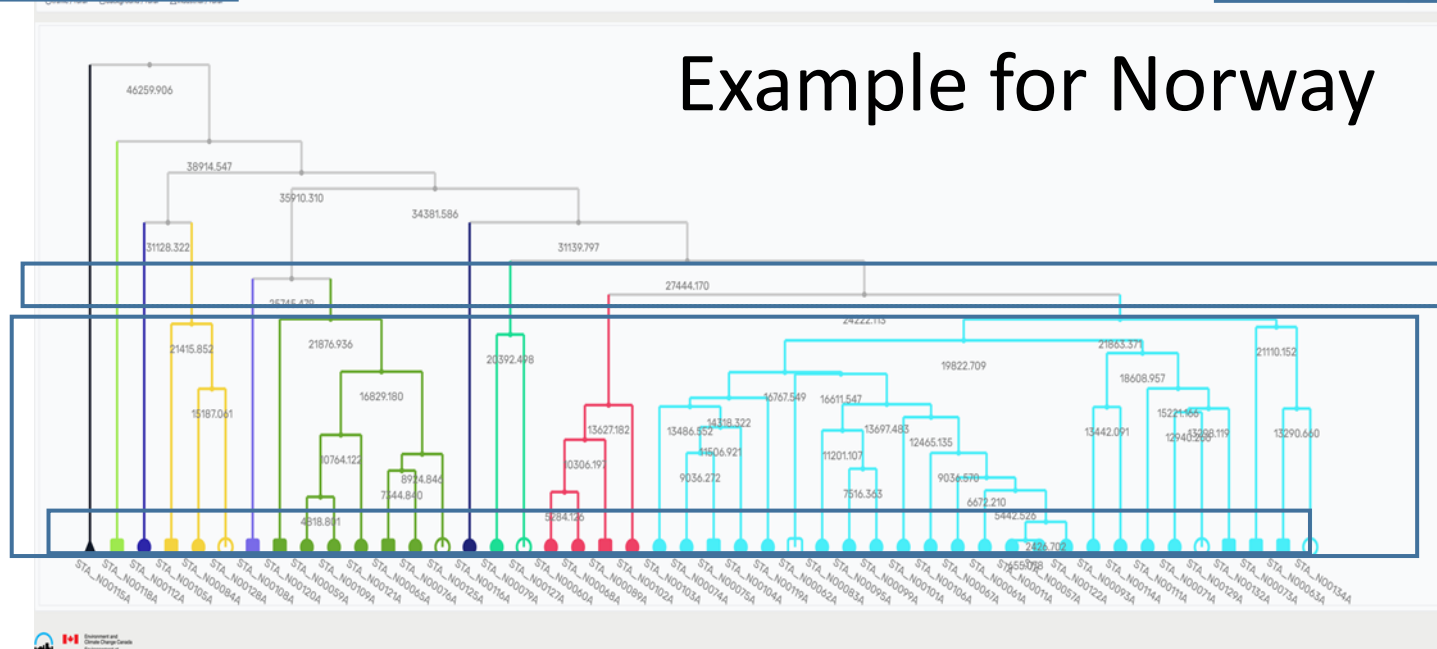
Information about the run

Shape/fill → station type/area

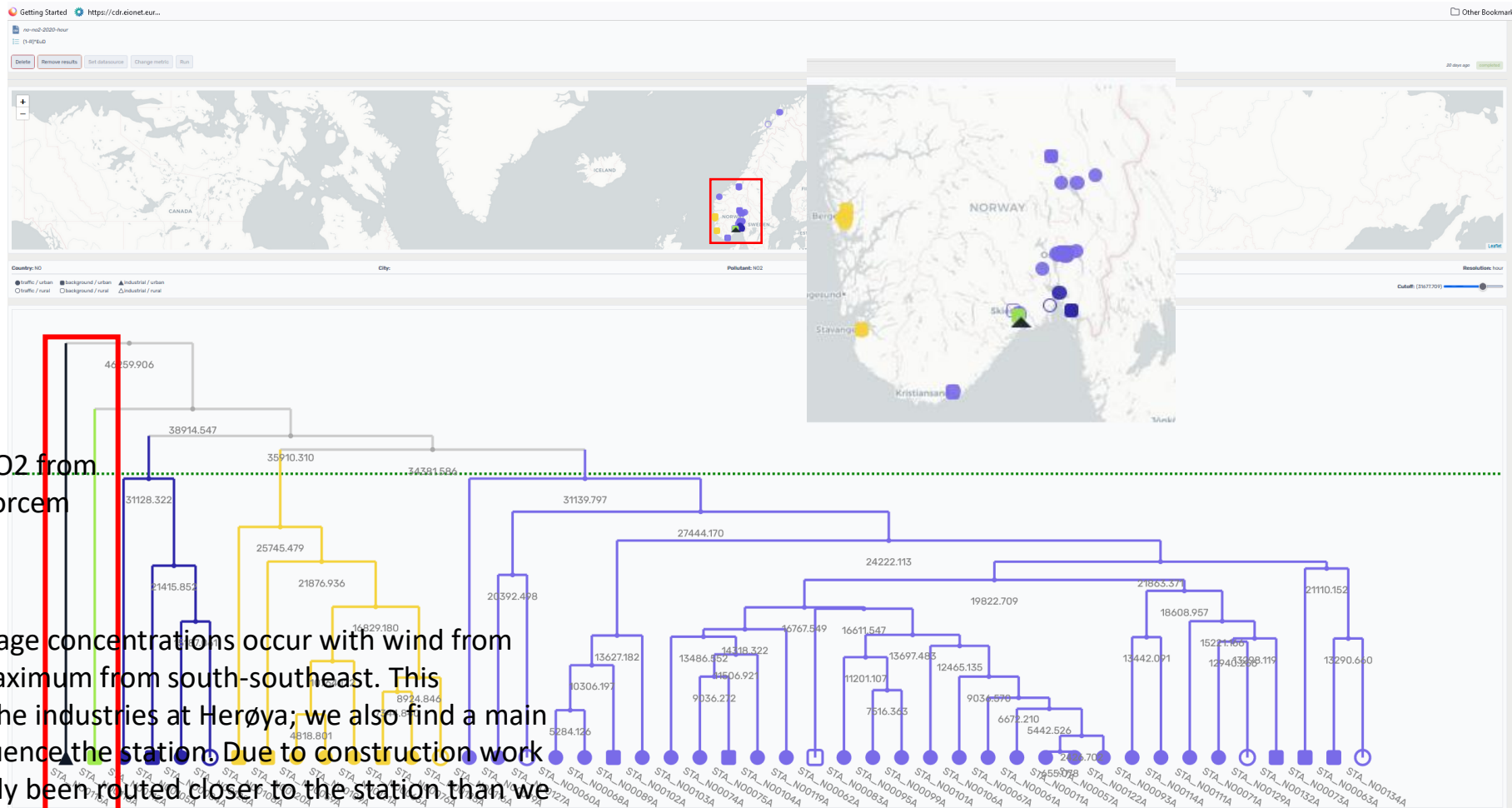
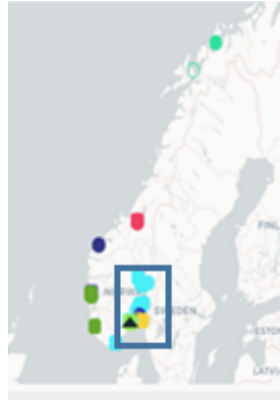
Slider: level chosen for analysing the results (stations clustered under the line)

Example for Norway

colour → clusters



STEP 3 - Visualise and interpret results



Industrial site (black)

Expert: "clear signals of NO2 from the industrial source at Norcem Brevik"

Urban background (green)

Expert: «The highest average concentrations occur with wind from east to south, with the maximum from south-southeast. This indicates influence from the industries at Herøya; we also find a main road, which also may influence the station. Due to construction work the traffic have periodically been routed closer to the station than we expected when the location was decided



Example for Norway (NO2, 2020)