Evaluation of AQ Monitoring Network

Proposed exercise using the interactive webtool - MoNet

Joana Soares, Christoffer Stoll, Islen Vallejo and Leonor Tarrasón

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**: Hierachical 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

NILU

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
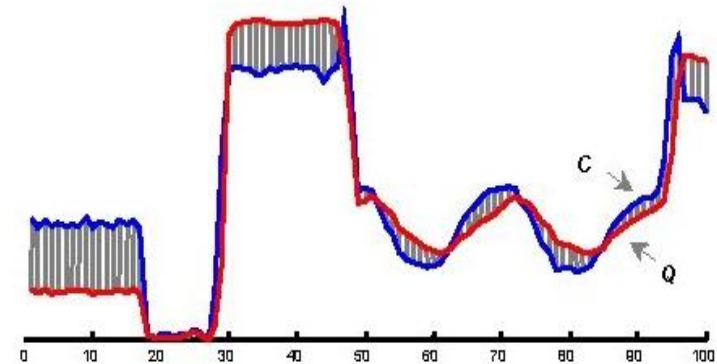
NILU

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**:

  o 1-R, R=Pearson correlation coefficient (Solazzo and Galmarini, 2015): **time variation**

  o Euclidean distance (EuD) (Soares et al, 2018): **magnitude variation**
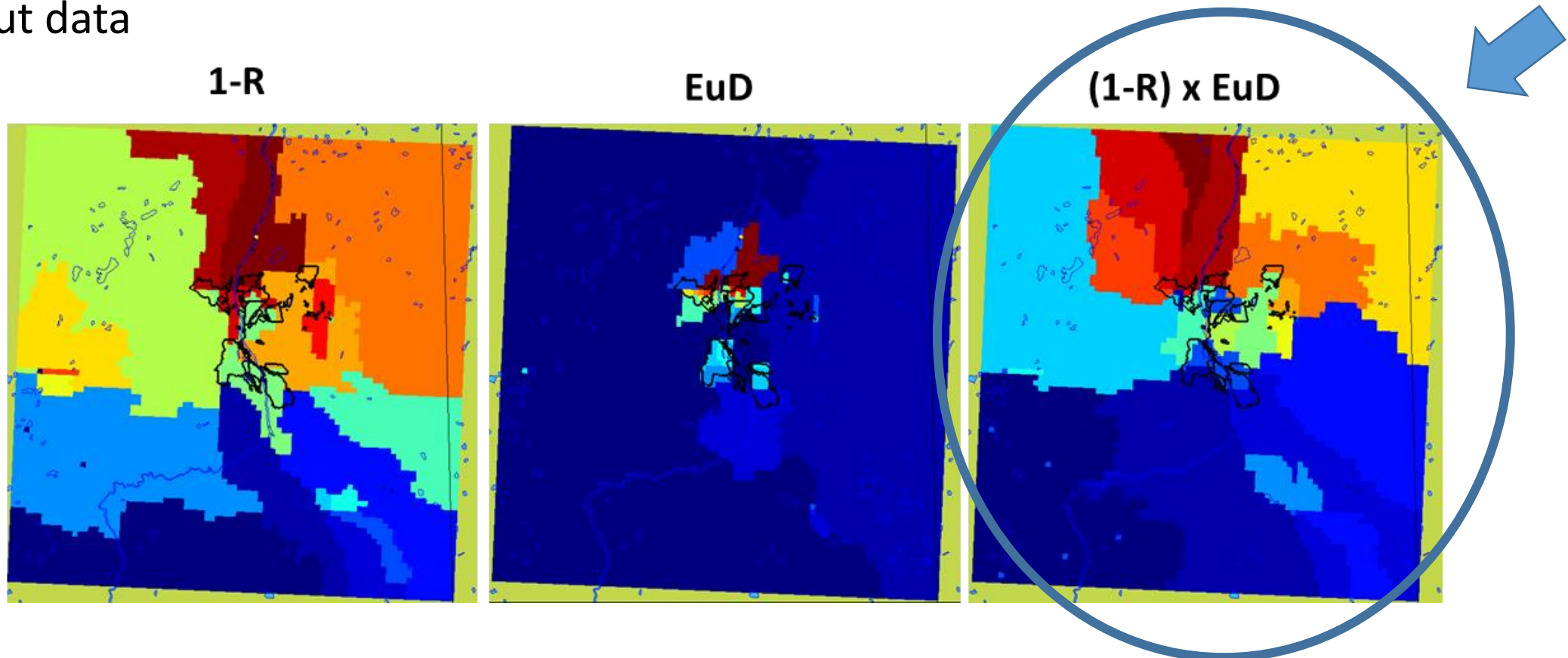
  o **(1-R) x EuD time and magnitude variation**
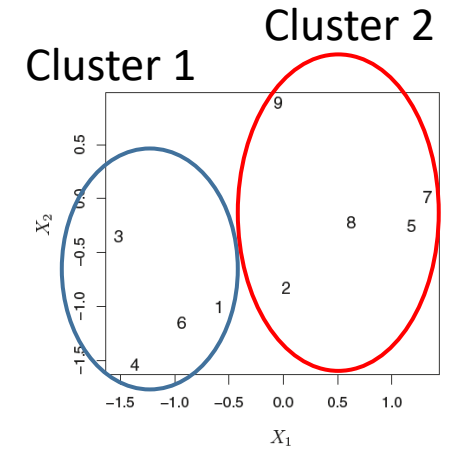
$$r_{cq} = \frac{\sum_j (c_j - m_c) \; (q_j - m_q)}{\sqrt{\sum_j (c_j - m_c)^2 \sum_j (q_j - m_q)^2}}$$

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

Cluster 1

Cluster 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)

NILU

Hierarchical Clustering Representation

- Dendrogram is a 2D representation of the hierarchical clustering process

  o show the pattern of linkages between the data series while clustering occurs

  o provide a quantitative description of the degree of similarity between the data points

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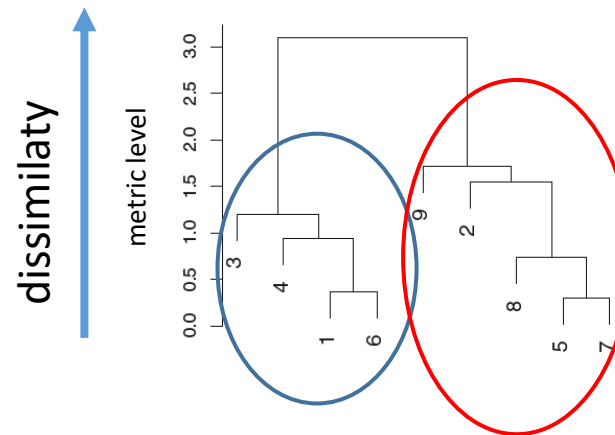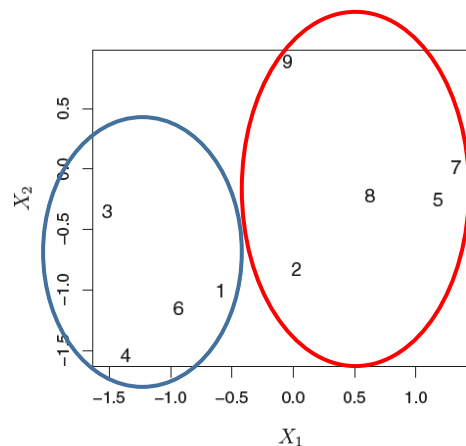
Reading the dendrogram:
- vertical lines: the level of (dis)similarity at every single stage of clustering
- horizontal lines: links between time series
- **less similar observations** are the ones joining up at the top

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.

metric = 11 - 1 cluster

MoNet Webtool(https://dev-monet.nilu.no/)
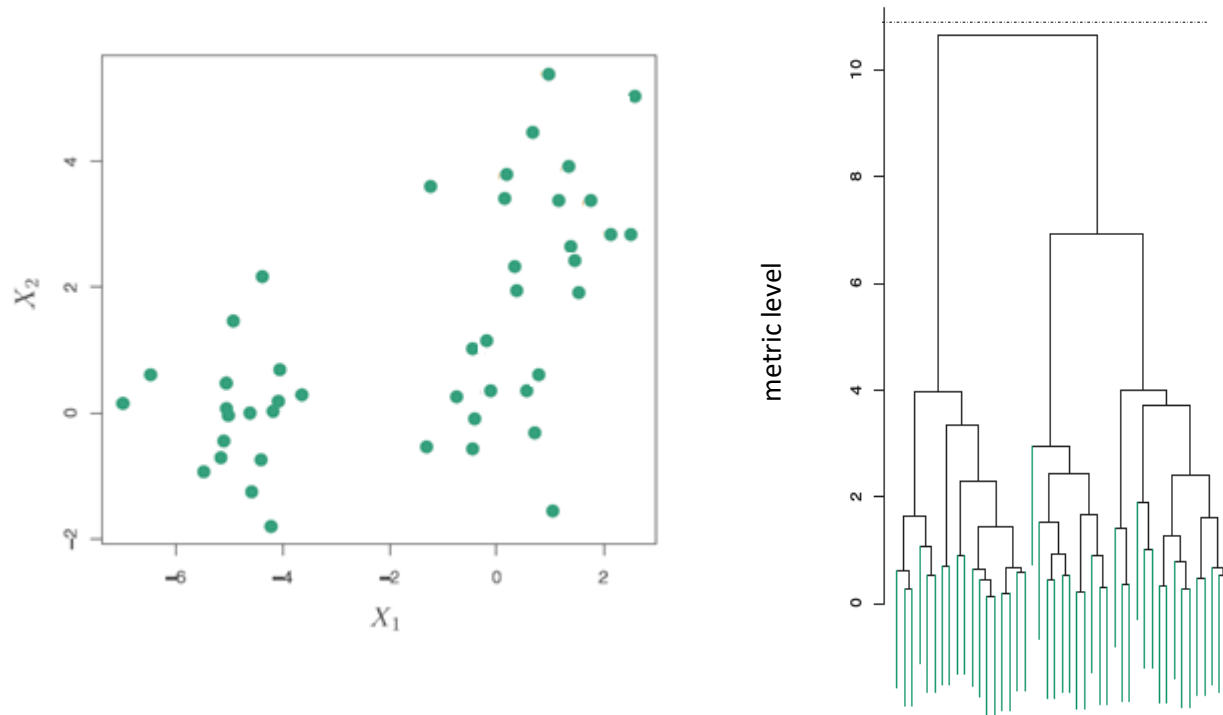
Current Features:
- Input: user- defined observation data or download from EEA
- 3 (di)similarity metrics:
  - 1-R
  - **(1-R)*EuD**
  - Euclidean distance
- Selection of clustering level
- Results on map and dendrogram
- Multiple clustering jobs

Features to come:
- More data source services
- Use gridded model data as input
- Ranking of stations
- Export results as image/ascii
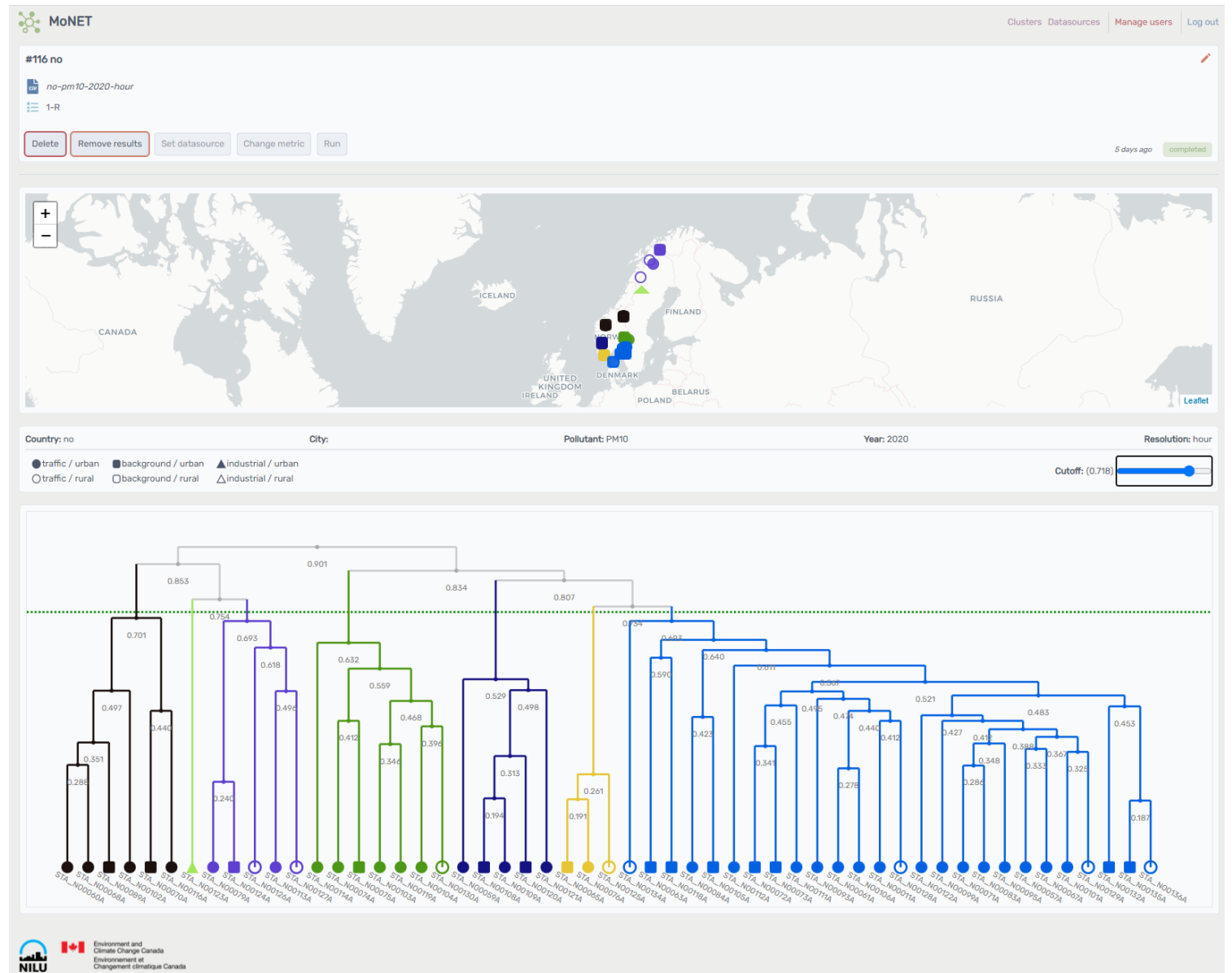
Demonstration of the tool - https://dev-monet.nilu.no/

Visualise and interpret results
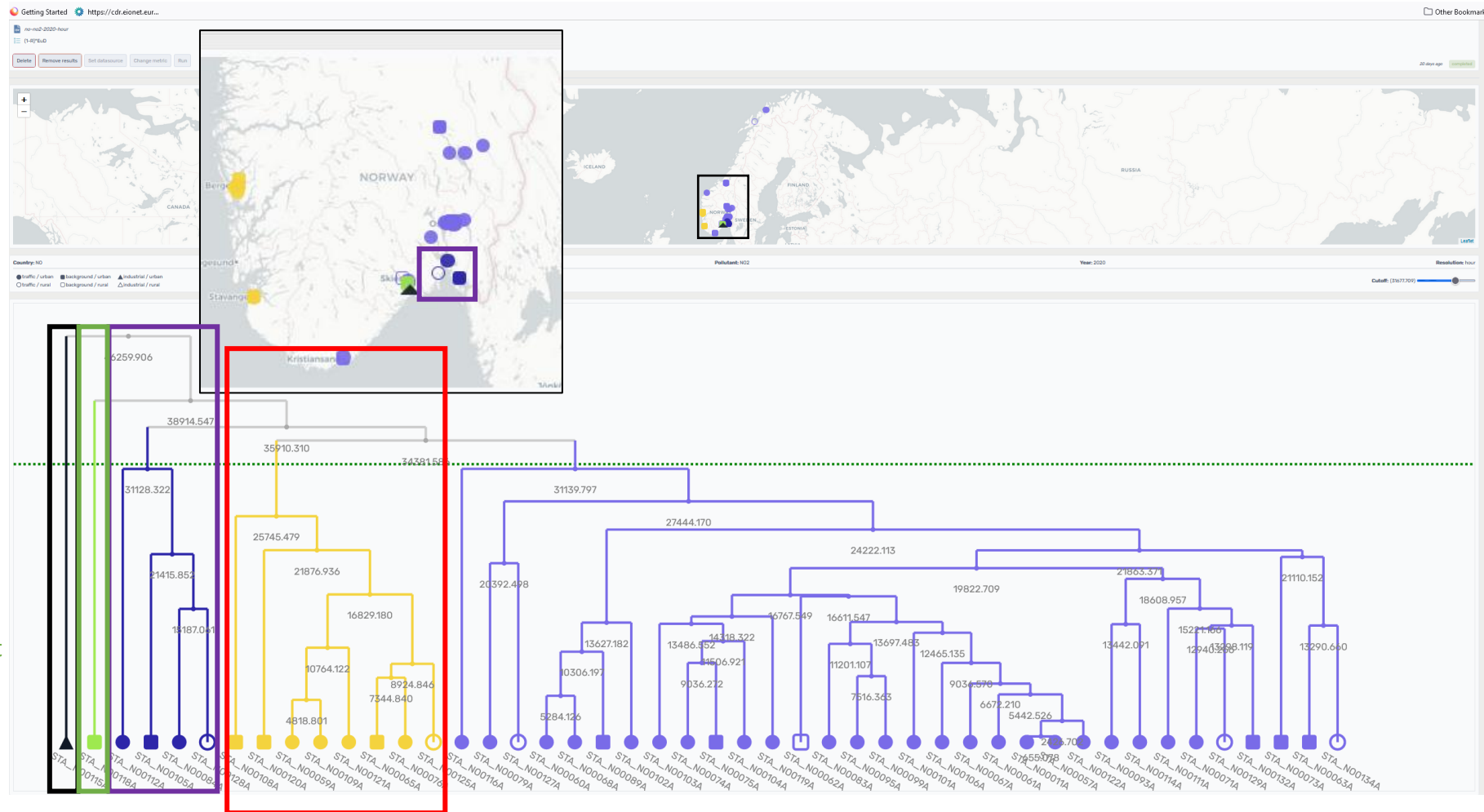
Industrial site with clear signals of NO2 from the industrial source at Norcem Brevik

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 influnced by small shipping; similar orography and meteorology

Meteorology is the key driver

Example for Norway (NO2, 2020)

Visualise and interpret results
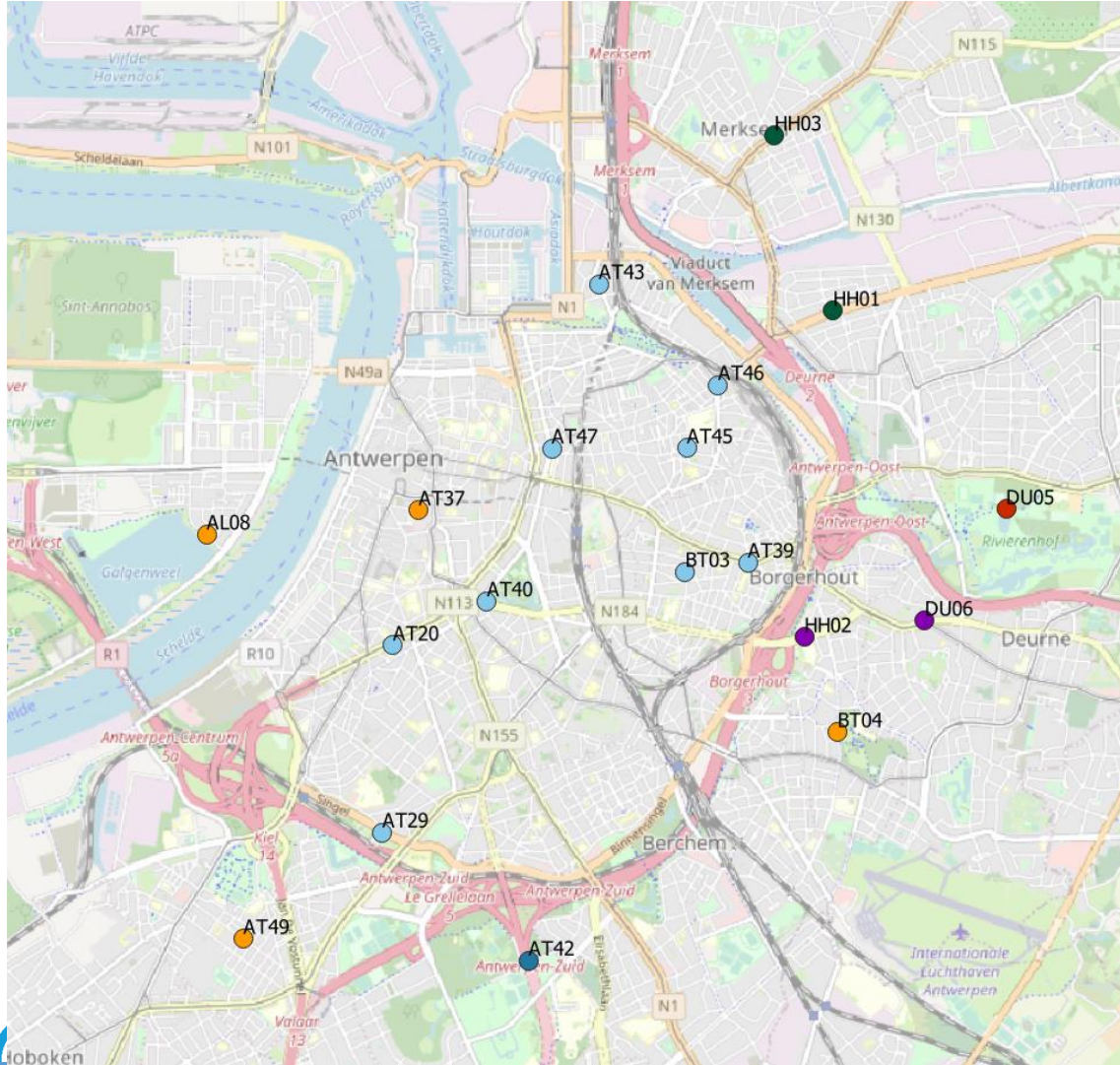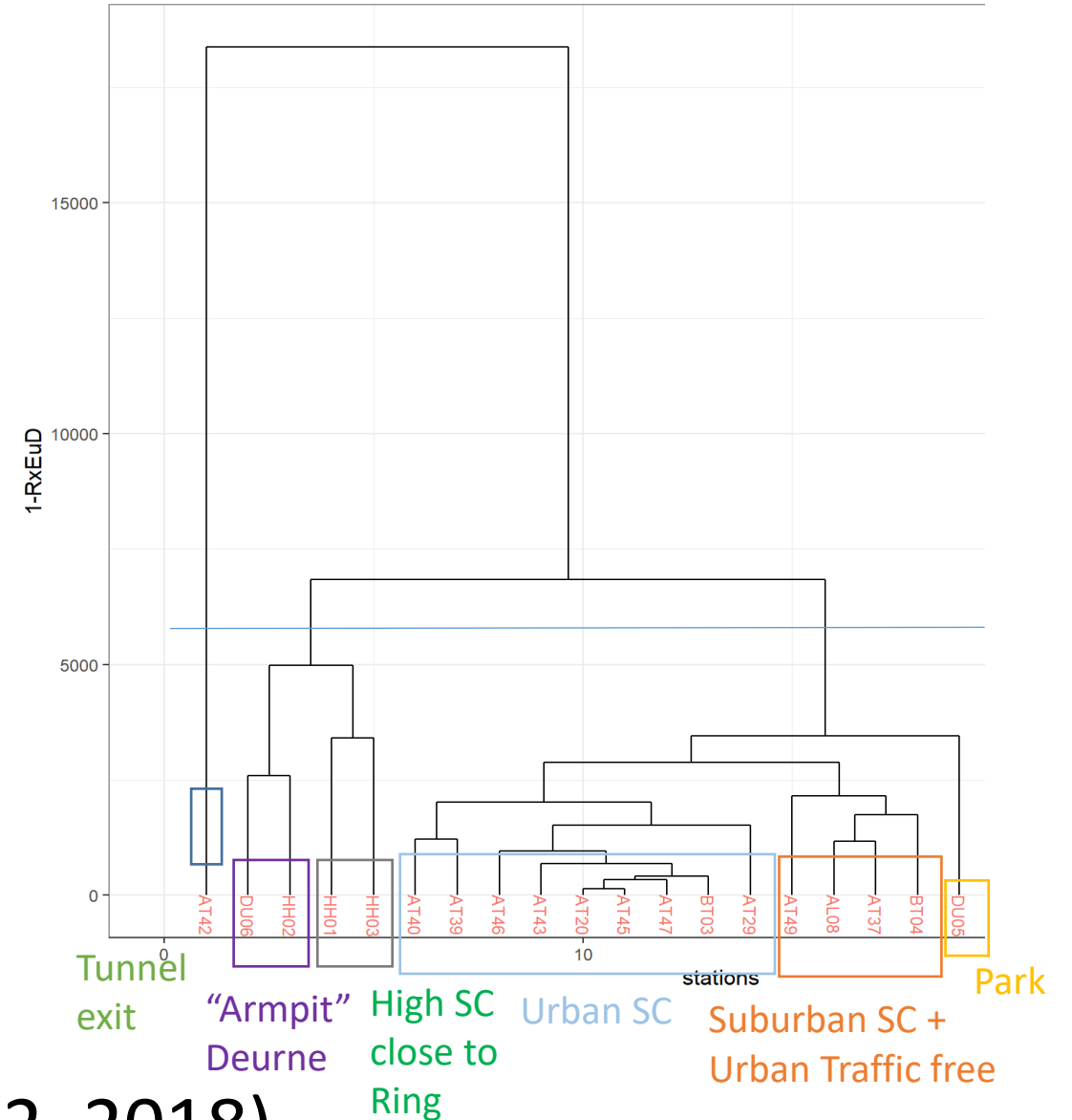
Example for Antwerp (NO2, 2018)

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)

NILU

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)

NILU

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**

NILU

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

NILU

Access to MoNet

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

Send email for username and password

jos@nilu.no

NILU

STEP 1 - Template – input data

queued (red)

completed (green)

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

Exclusion: Some data coverage was below 75%

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

STEP 2– create a new job (2)

STEP 3 - Visualise and interpret results

Information about the run

Delete the entire run or
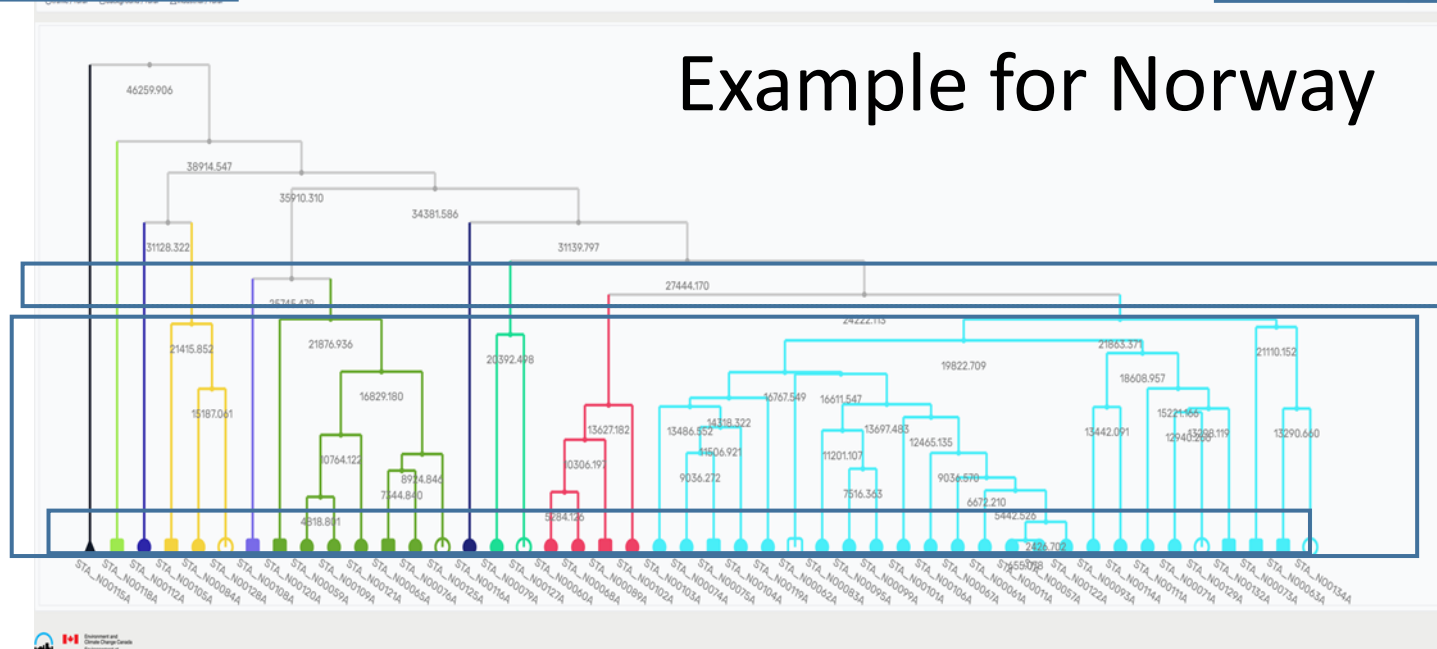delete only the results
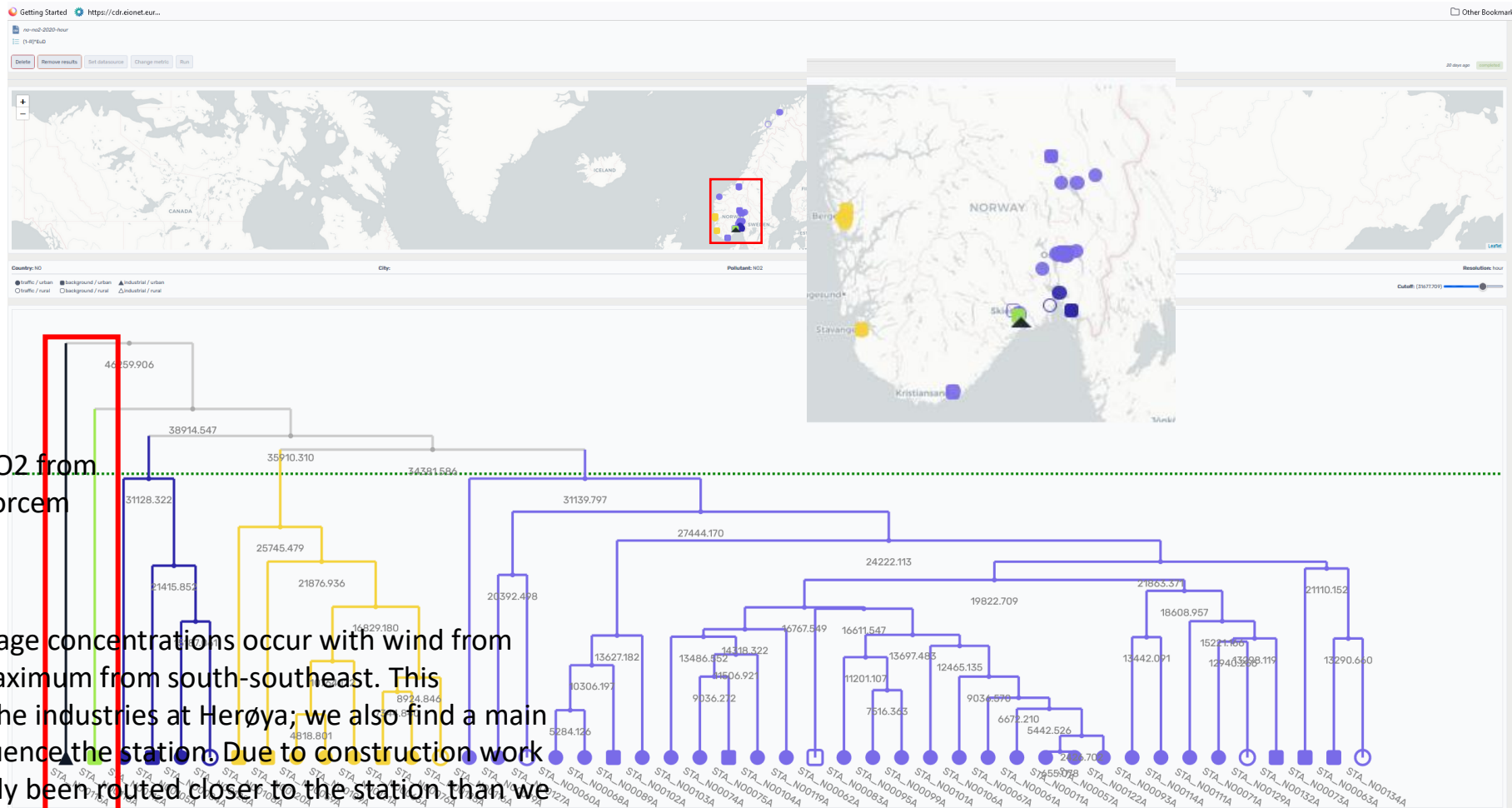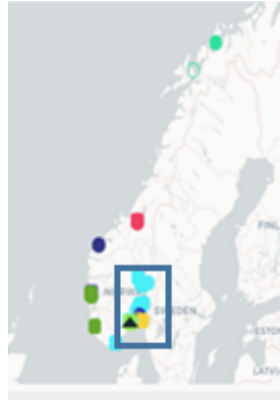
Shape/fill ➜ station type/area

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)