FAIRMODE

National Institute for Public Health and the Environment
Ministry of Health, Welfare and Sport

Benchmark of Data Fusion using sensor data

FAIRMODE WG6, Dublin, October 7, 2024

FAIRMODE

Agenda WG6 Monday

- Quick process recap
- General benchmark setup
- Presentation of results participants including questions
  - INERIS Alicia Gressent, remote
  - MetNo Lewis Blake, remote
  - ISSeP Fabian , recorded/remote
  - CERC Jenny Stocker
  - RIVM Joost Wesseling
- First comparison between preliminary results, Joost Wesseling
- Discussion

FAIRMODE

National Institute for Public Health and the Environment
Ministry of Health, Welfare and Sport

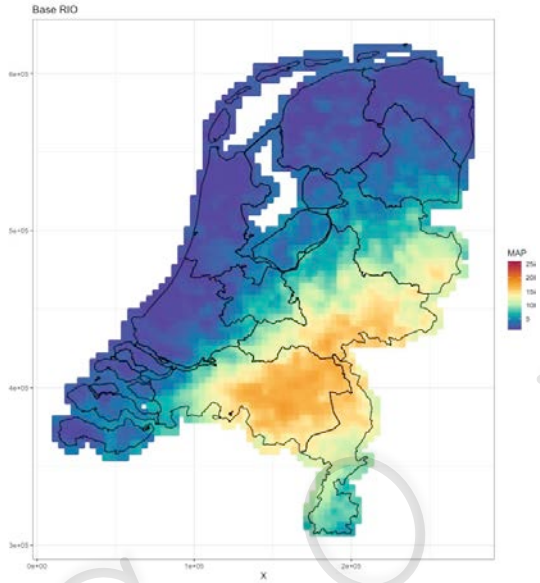
Benchmark of Data Fusion using sensor data

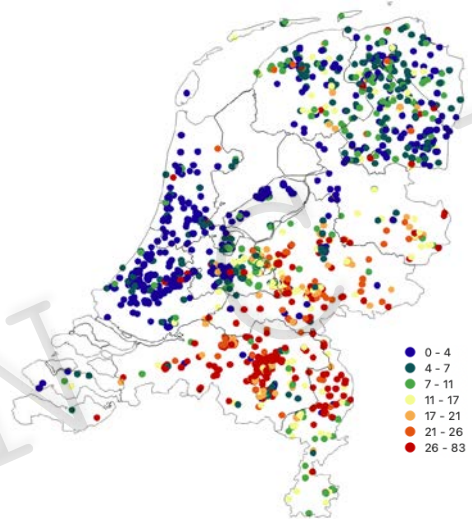
Setup

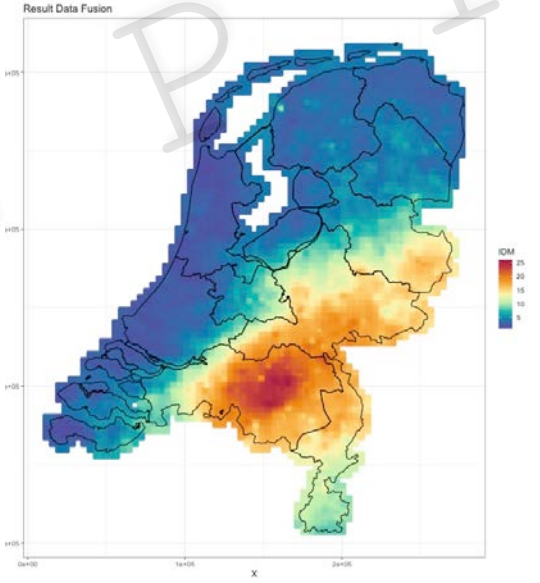
FAIRMODE WG6, Dublin, October 7, 2024
Sjoerd van Ratingen, Joost Wesseling, RIVM

Data fusion with sensors

Starting from a concentration field based on official measurements and many sensors, what is the best combination?







CONCEPT

FAIRMODE

Since Paris...

- Benchmark for data fusion of **~ 1800 sensors** in the **Netherlands** together with **RIO model results**
- Continued search for organizations, interested the sensor data fusion benchmark.
- Four live meetings to converge on benchmark setup
- Outcome is six organizations now actively testing fusion models within the benchmark
  VITO, INERIS, MetNO, ISSeP, CERC, RIVM
  Other interested parties up to date via meetings and correspondence.
- Discussions
  - Input /output formats, Most interesting data sets to use which period
  - Variants / validation, Uncertainties
  - Use of reference stations
- Testing with preliminary data from participants

FAIRMODE

Options DF runs

Run	Official measurements	Sensor measurements	RIO map	Goal
Basic run	**All**	Raw and calibrated using **all** measurements	Based on **all** official measurements	Compare results from different data fusion methods
Option 1a	Leave out 1-2 official measurements in each province	Raw and calibrated using **all** measurements	Leave out the same 1-2 measurements in each province	Can **good** sensors compensate for less official measurements?
Option 1b	Leave out 1-2official measurements in each province	Leave out the same 1-2 measurements in each province in the calibration	Leave out the same 1-2 measurements in each province	Can **not so good** sensors compensate for less official measurements?
Option 2a	Leave out **all** official measurements in **one** province	Raw and calibrated using **all** measurements in **all** provinces	Based on **all** official measurements in the **other** provinces	Can **good** sensors compensate for a gap in the official measurements?
Option 2b	Leave out **all** official measurements in **one** province	Raw and calibrated using **all** measurements in the **other** provinces	Based on **all** official measurements in the **other** provinces	Can **not so good** sensors compensate for a gap in the official measurements?

Analysis of 1st results Jan/Aug 2024

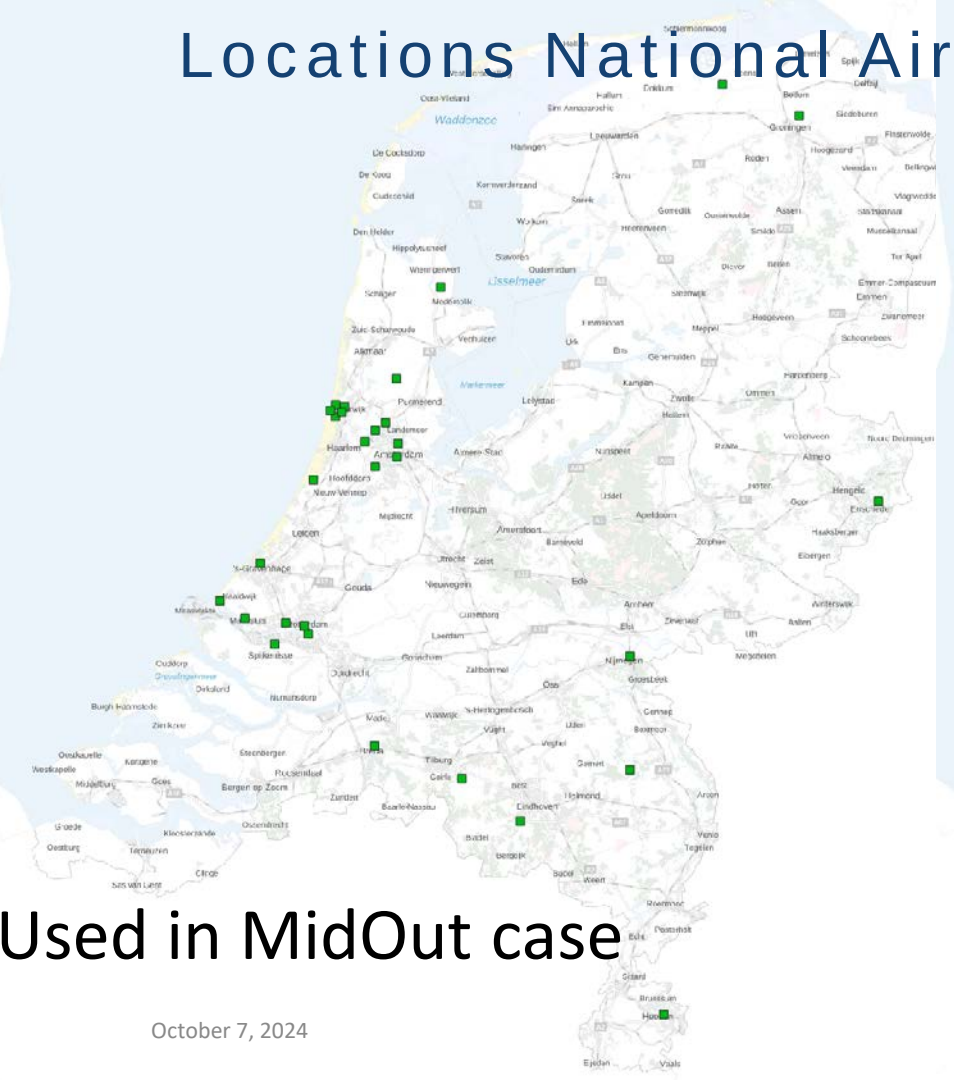
- Available independent official data to test the results of RIO and the data fusion.
- Interesting periods to test the data fusion.
- Analyses …
- Results.
- Conclusions.

Locations National Air Quality network (LML)

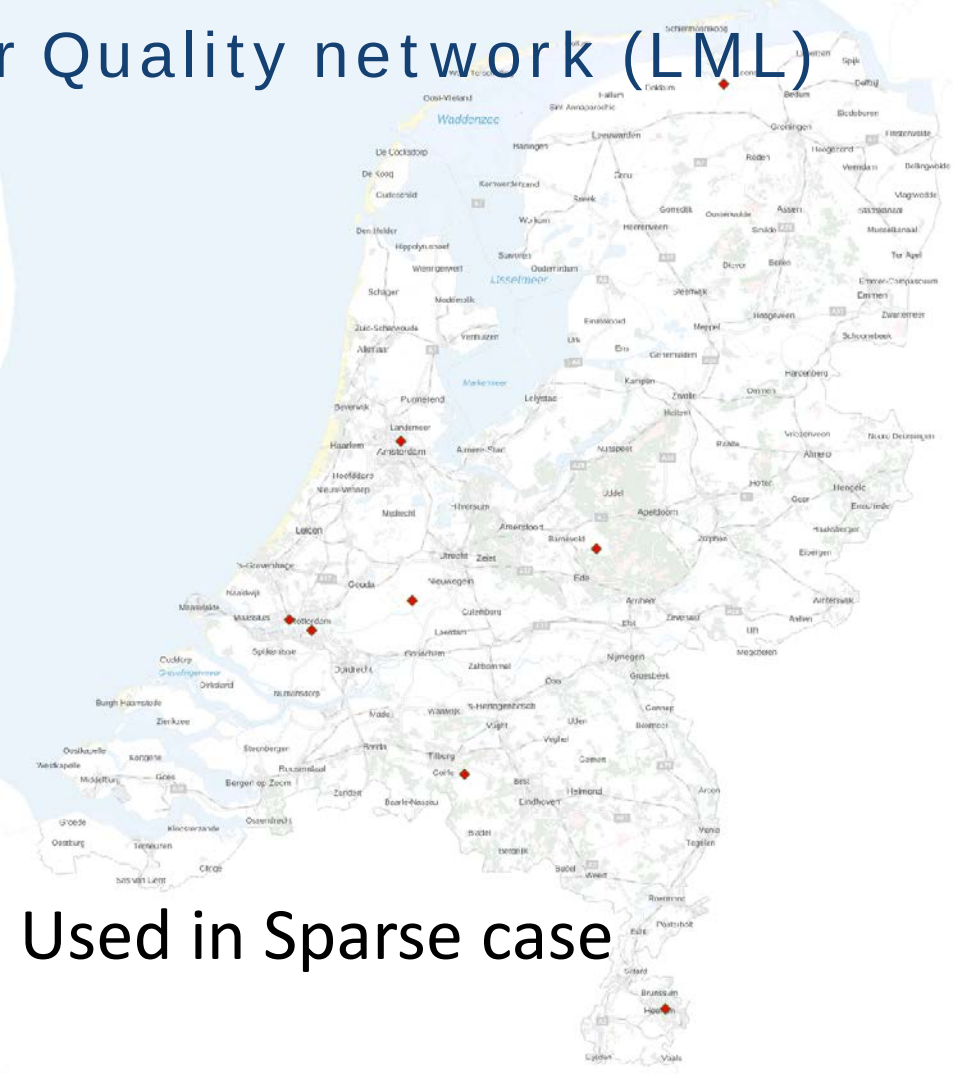
Used in base case

<u>Not</u> used in RIO

Locations National Air Quality network (LML)



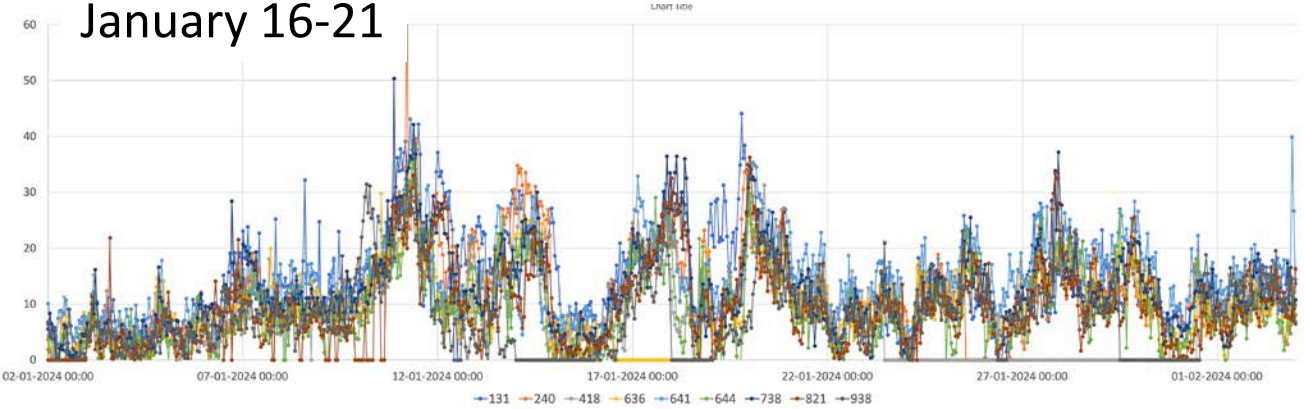
Used in MidOut case



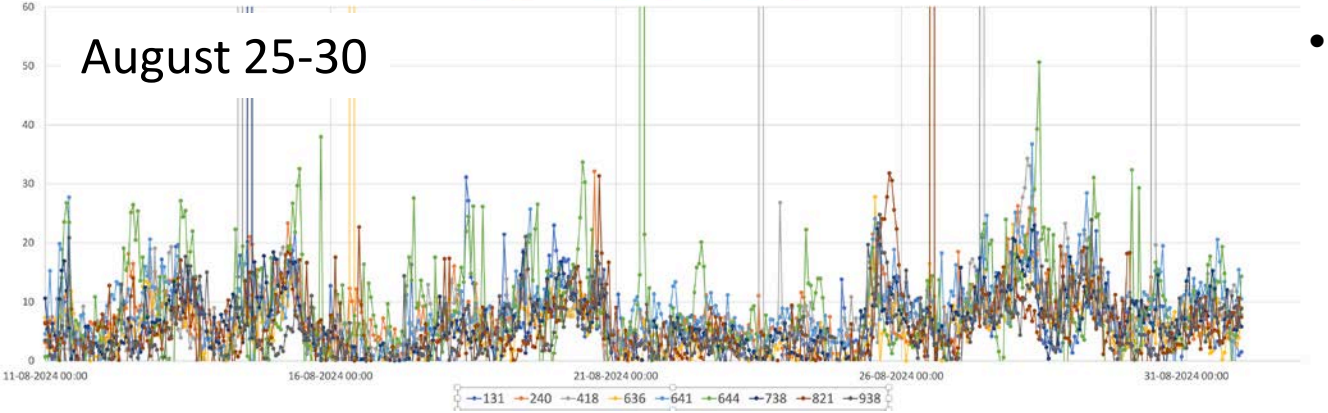
Used in Sparse case

Interesting periods

January 16-21



August 25-30



- PM2.5 concentrations at various locations spread over the country.
- Most other days in the period Jan-Sep, 2024 show low concentrations and little variations.
- Two weeks were chosen to test and compare the data fusion methods.

Input data

For every hour in the test periods, RIVM provided several files:

2024032909_official.txt	Official network	ID X Y PM2.5
2024032909_sensors.txt	Sensor data	ID X Y CALIB Ucal Urnd
2024032909_knmi.txt	Meteo data	(description follows)
2024032909_pm25_uur_11.txt	Output of RIO model	X Y CONC
2024032909_pm25_uur_11_err.txt	Uncertainty of RIO model	X Y CONC

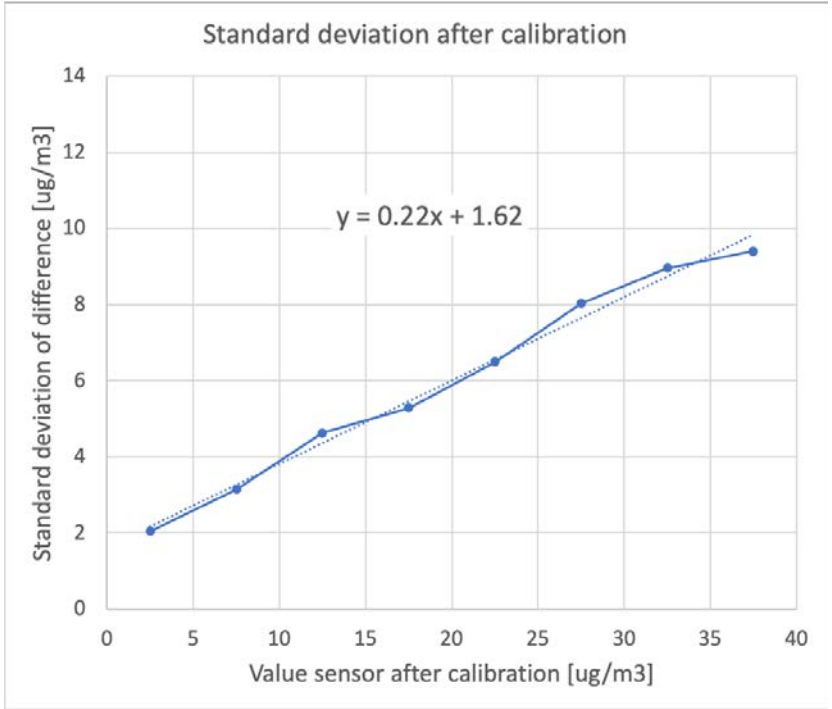
ID : Name of location or sensor
X, Y : Official Dutch coordinates, “Amersfoort /RD New”, EPSG:28992
CALIB : Calibrated Nova/SDS011 sensor data for PM2.5
Ucal : Calibration uncertainty from bootstrap
Urnd : Random uncertainty estimated from comparison to official data
CONC : RIO concentration and uncertainties

Calibration uncertainty sensors

- The calibration of the sensors is determined by comparing the average values of groups of sensors to those of neighbouring official measurements . By performing many such calibrations, each with a random selection of both available sensors and official measurements, a range of calibrated values is obtained for each sensor.

- The span of values between the 0.025 and 0.975 percentiles of the range of calibrated values is interpreted as the 95% uncertainty.

Random uncertainty sensors



- Every individual calibrated sensor has an individual random uncertainty estimated at $\sigma_{rnd,sensor}$ = 1.62 + 0.22 C ug/m3 (standard deviation), with C the calibrated concentration of the sensor.
- The uncertainty is estimated using results from many collocated sensors.

Analyze results ...

First analysis:

1. Compare the **time series** at a set of locations for the values of the RIO map and results from all the models in the benchmark.
2. Correlate model results at a set of locations with official measurements.
3. Compare average RMSE at relevant official locations that were not used in the data fusion.

Results from individual models

FAIRMODE

National Institute for Public Health and the Environment
Ministry of Health, Welfare and Sport

Benchmark of Data Fusion using sensor data

Setup

FAIRMODE WG6, Dublin, October 8, 2024
Sjoerd van Ratingen, Joost Wesseling, RIVM

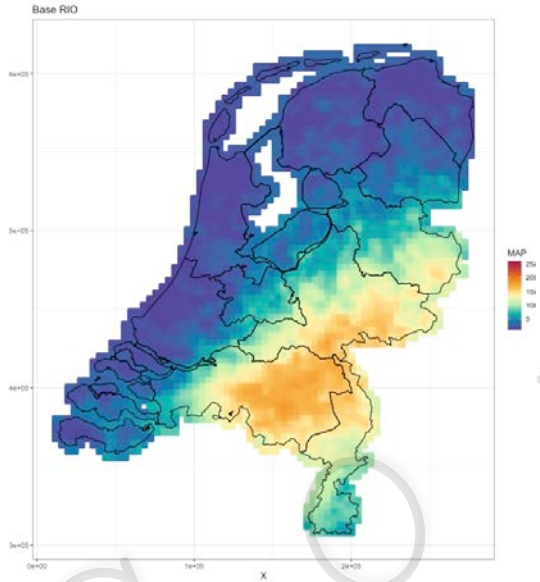
FAIRMODE

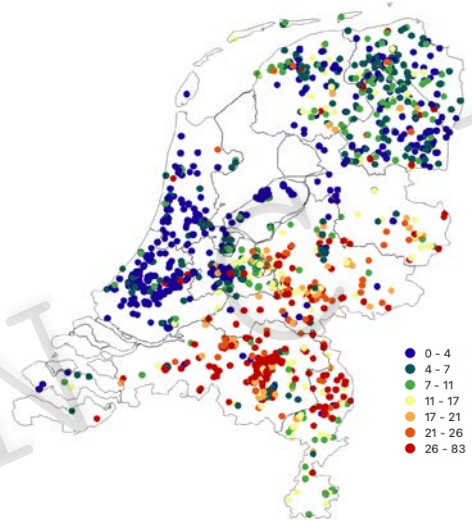
Agenda WG6 Tuesday session

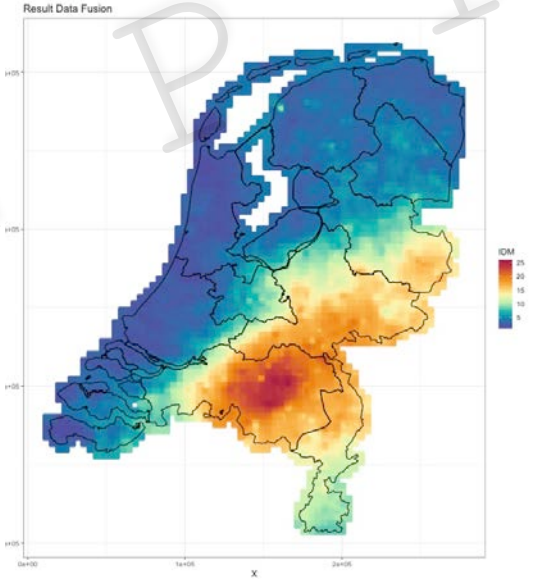
- Introduction benchmark
- Presentation of RIVM –DF results
- Comparison between preliminary results.
- Discussion/Questions

Data fusion with sensors

Starting from a concentration field based on official measurements and many sensors, what is the best combination?







CONCEPT

Analysis of 1st results Jan/Aug 2024

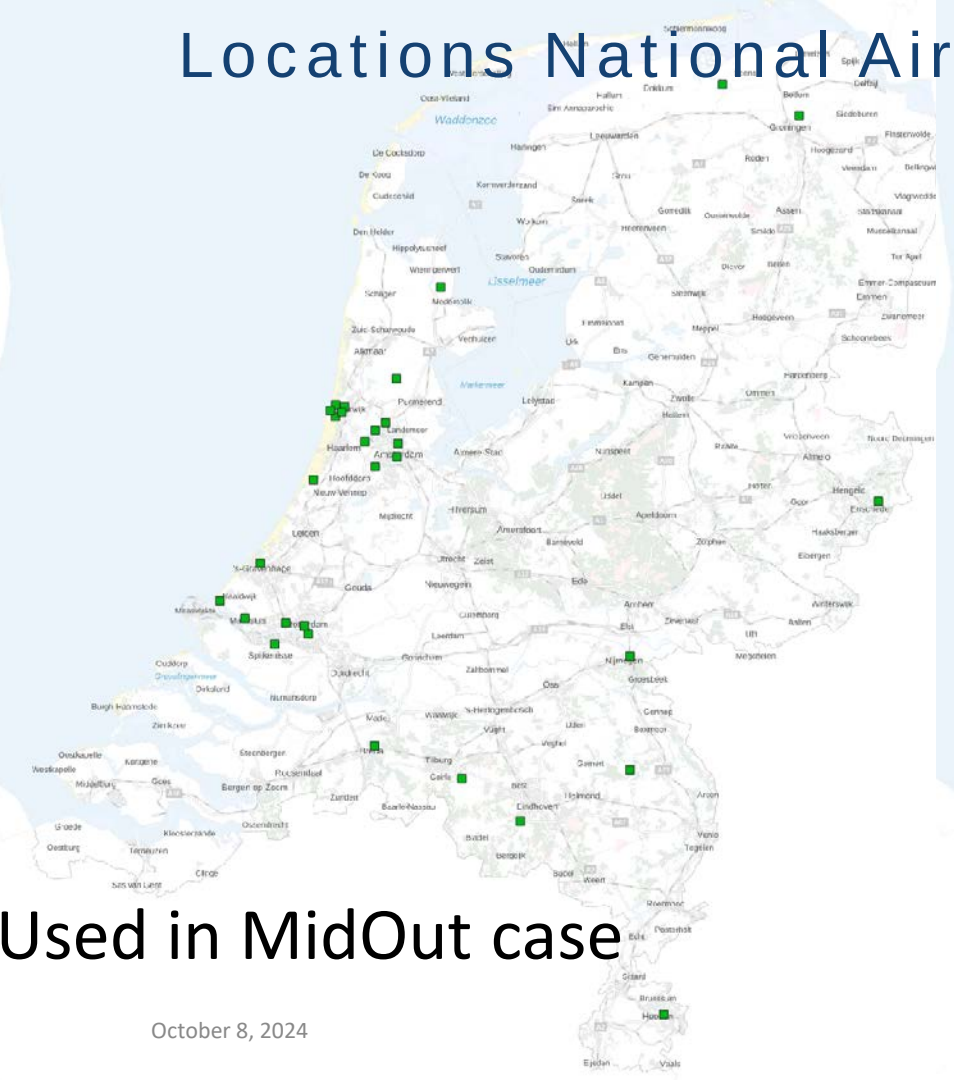
- Available independent official data to test the results of RIO and the data fusion.
- Interesting periods to test the data fusion.
- Six active participants have shared their model results
- Analyses …
- Results.
- Conclusions.

Locations National Air Quality network (LML)

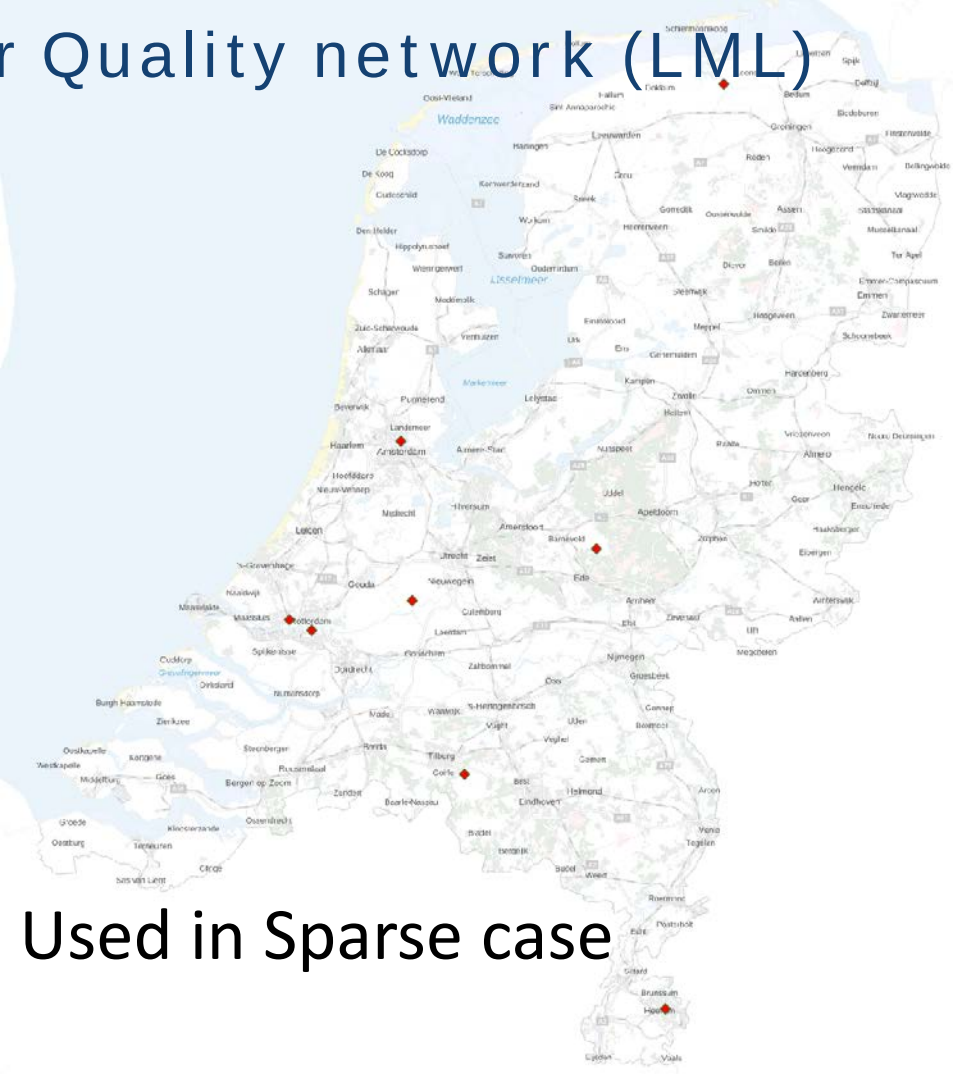
Used in base case

<u>Not</u> used in RIO

Locations National Air Quality network (LML)



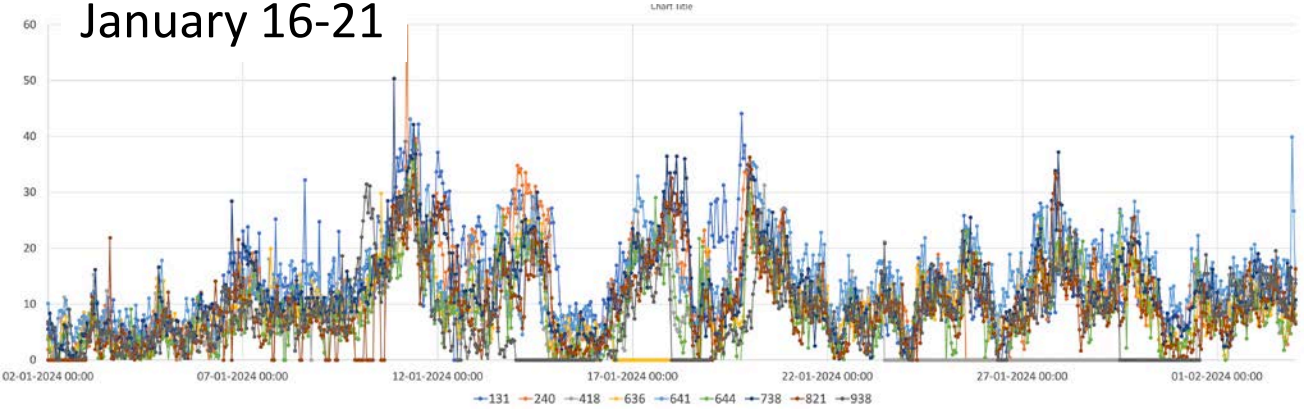
Used in MidOut case



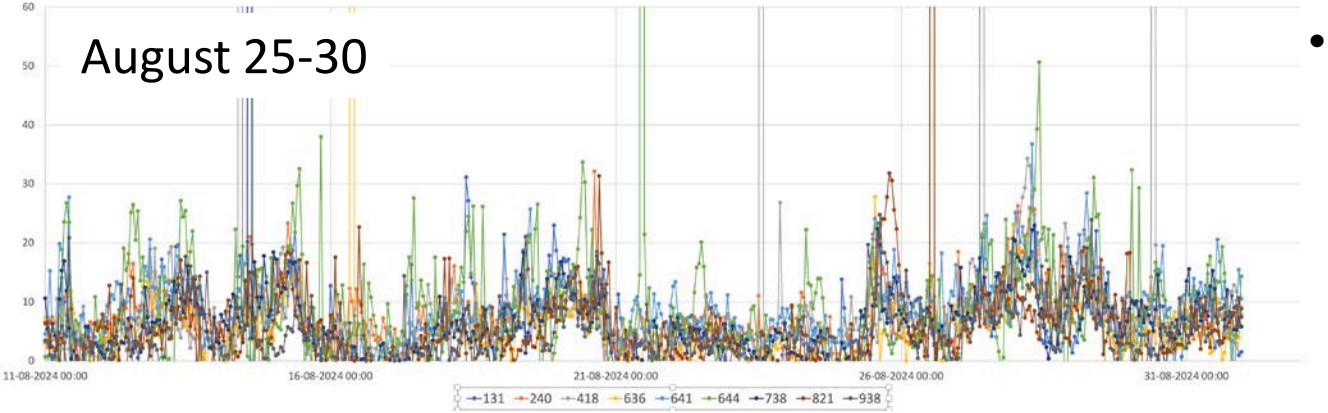
Used in Sparse case

Interesting periods

January 16-21

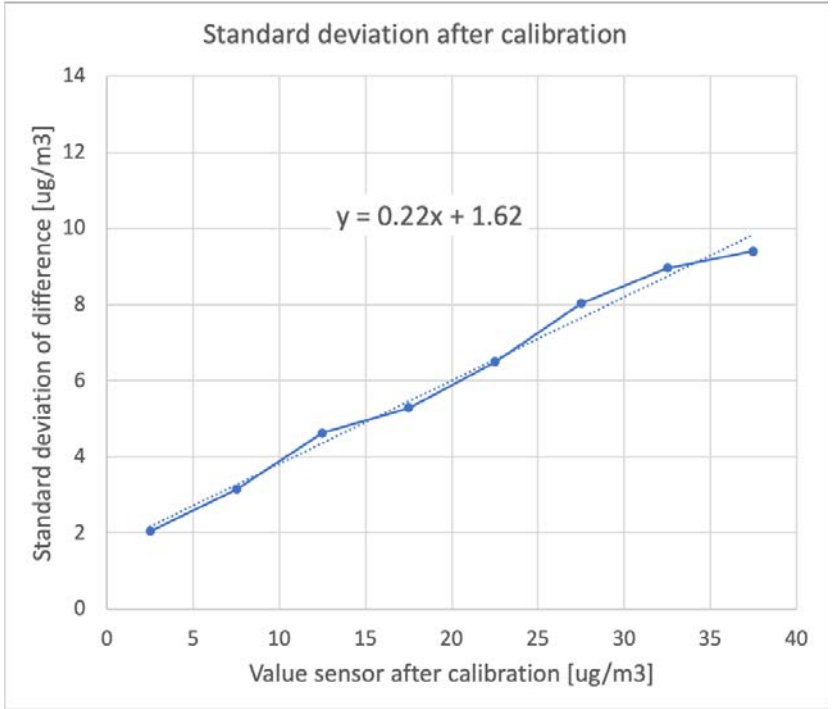


August 25-30



- PM2.5 concentrations at various locations spread over the country.
- Most other days in the period Jan-Sep, 2024 show low concentrations and little variations.
- Two weeks were chosen to test and compare the data fusion methods.

Random uncertainty sensors



- Every individual calibrated sensor has an individual random uncertainty estimated at $\sigma_{rnd,sensor}$ = 1.62 + 0.22 C ug/m3 (standard deviation), with C the calibrated concentration of the sensor.
- The uncertainty is estimated using results from many collocated sensors.

Results from individual models

- Individual results presented Monday
  - INERIS
  - MET Norway
  - ISSep
  - CERC
  - VITO
- Individual results presented Today
  - RIVM

Analyze results ...

First analysis:

1. Compare the **time series** at a set of locations for the values of the RIO map and results from all the models in the benchmark.
2. Correlate model results at a set of locations with official measurements.
3. Compare average RMSE at relevant official locations that were not used in the data fusion.

Calibration uncertainty sensors

- The calibration of the sensors is determined by comparing the average values of groups of sensors to those of neighbouring official measurements . By performing many such calibrations, each with a random selection of both available sensors and official measurements, a range of calibrated values is obtained for each sensor.

- The span of values between the 0.025 and 0.975 percentiles of the range of calibrated values is interpreted as the 95% uncertainty.

Input data

For every hour in the test periods, RIVM provided several files:

2024032909_official.txt	Official network	ID X Y PM2.5
2024032909_sensors.txt	Sensor data	ID X Y CALIB Ucal Urnd
2024032909_knmi.txt	Meteo data	(description follows)
2024032909_pm25_uur_11.txt	Output of RIO model	X Y CONC
2024032909_pm25_uur_11_err.txt	Uncertainty of RIO model	X Y CONC

ID : Name of location or sensor
X, Y : Official Dutch coordinates, “Amersfoort /RD New”, EPSG:28992
CALIB : Calibrated Nova/SDS011 sensor data for PM2.5
Ucal : Calibration uncertainty from bootstrap
Urnd : Random uncertainty estimated from comparison to official data
CONC : RIO concentration and uncertainties

FAIRMODE

Since Paris...

- Benchmark for data fusion of **~ 1800 sensors** in the **Netherlands** together with **RIO model results**
- Continued search for organizations, interested the sensor data fusion benchmark.
- Four live meetings to converge on benchmark setup
- Outcome is six organizations now actively testing fusion models within the benchmark
  VITO, INERIS, MetNO, ISSeP, CERC, RIVM
  Other interested parties up to date via meetings and correspondence.
- Discussions
  - Input /output formats, Most interesting data sets to use which period
  - Variants / validation, Uncertainties
  - Use of reference stations
- Testing with preliminary data from participants

FAIRMODE

Agenda WG6 Tuesday session

- Introduction benchmark
- Presentation of RIVM –DF results
- Comparison between preliminary results.
- Discussion/Questions

Discussion

Are we using the best metrics for evaluating the effects of the methods

- Now: Improvement of RMSE at validation locations
- More detail: Estimate up bias and random part of RMSE?
- ……
- More focus on interesting episodes (peaks) and apply validation metrics to these situations
- Focus on metrics, fit for picking up peaks
- Focus on weak model aspects

Discussion

The RIO model that is used as a start of the DF performs already quite good. It would be interesting to use another model/map as a start, suggestions?

- CAMS(ensemble/models)
- Other variants of RIO interpolation, leaving out larger part of country
- Other countries/regions than the Netherlands?

- CAMS
- Hourly varying constant field (and estimate its uncertainty!)
- Don’t forget higher resolution than 1x1kmL
- Large variance within km cells could be indicative of local phenomena.

Discussion

Use of additional information on top of a map that is already based on these measurements?

- How would this extra information be (significantly more) useful?
- Reference stations
- Meteorological data
- Land use data

- Focus on comparing DF (per station) models that explicitly do and don’t use reference data.
- Check the effects of adding one extra reference station in an area that did not contain in.

Discussion

What practical alternatives are there for DF with sensors?

- Data assimilation ?
- DF using Machine learning ?

- DF Machine Learning , RF would be interesting to add to the benchmark

Discussion

What are the most interesting practical configurations of maps / Low Cost Sensors to test?

- Use case other than the Netherlands?

- Dense vs sparse sensor networks

- Larger vs smaller domains…

- Combining with more detailed models?
- Croatia (sparse sensornetwork), Ireland (more dense, raw measurements present?)

- Romania → Urban areas with approx. 400 sensor and 8.b. Boekarest.
- Maybe use (few) CAMS cells covering urban area as starting point construct a refined grid only using sensors.
- Consider having specific sensors for e.g. wood burning..

Discussion

Can real time sensor data be used to adjust air quality forecasts?

- Together with a meteo-forecast?
- Extrapolate corrections into next 2 -3 days ?
- Yes. Use them to correct current hour → persistence model.

Discussion

Best combinations for reference / maps / Low Cost Sensors / … ?

- Sensors still added value ?
- Which typical locations ?
- Optional ?