Microsensor meeting at RIVM,
13 February 2017, Utrecht, NL

Mapping urban air quality in near real-time

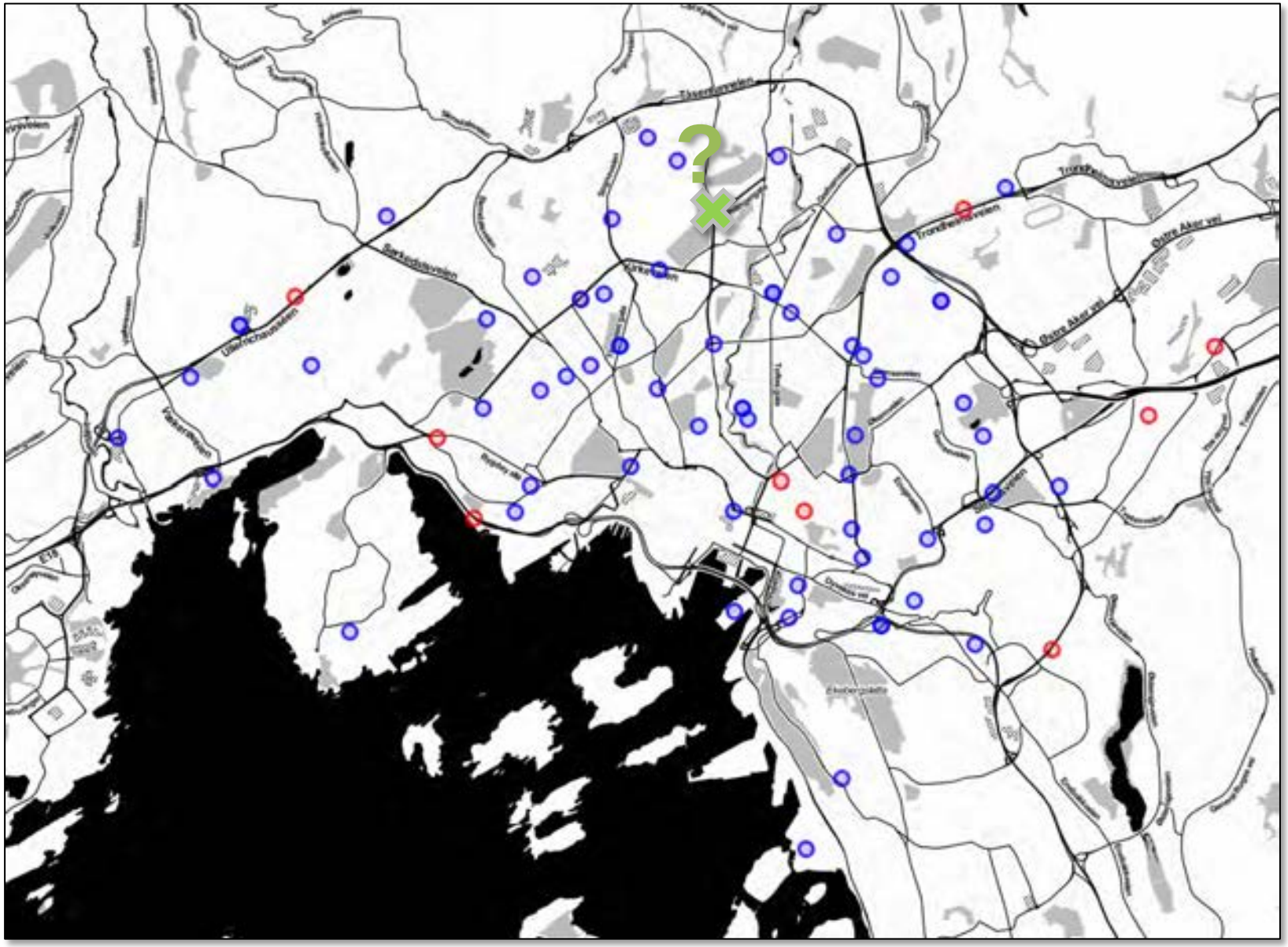
Data fusion of observations from low-cost sensors and model information

Philipp Schneider
Nuria Castell
Matthias Vogt
William Lahoz
Alena Bartonova
and the entire CITI-SENSE team

NILU – Norwegian Institute for Air Research

Introduction

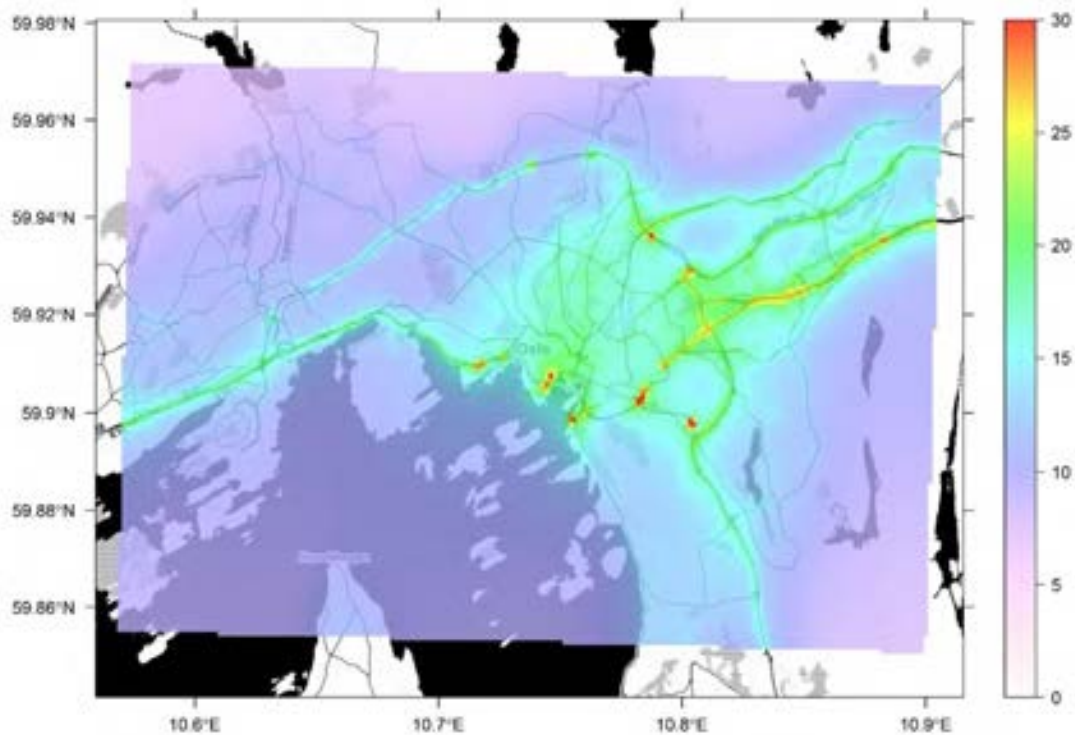
- Low-cost microsensors can provide air quality measurements throughout the city at much higher density than is possible with traditional reference equipment
- This opens the opportunity for creating unprecedented high-resolution urban-scale maps of air quality based on observations
- Such maps can then be used to provide citizens with a wide variety of services, e.g. health-aware routing, personal exposure etc.
- To achieve this we need to combine the sensor observations with model information (either dispersion or land-use regression) to map concentrations onto a high-resolution grid



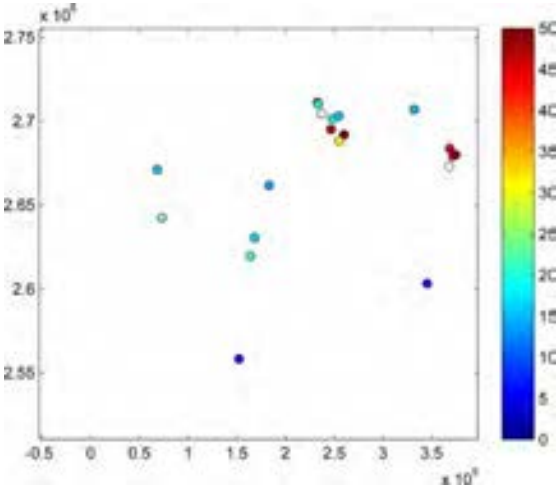
Red markers: Locations of Air Quality Monitoring stations for NO_2
Blue markers: Deployment sites of low-cost microsensors

Combination with model output

- To map the observations from the low-cost sensors onto a high-resolution grid in a scientifically meaningful way we need to use a spatial auxiliary dataset that guides the interpolation
- Two primary ways of combining observations with model data (with much overlap)
  - (Geo-)Statistical techniques: Data fusion
  - Traditional data assimilation: E.g. EnKF, 3D-VAR, 4D-VAR
- We use here the output from the EPISODE air quality model (high-resolution long-term average concentration maps)

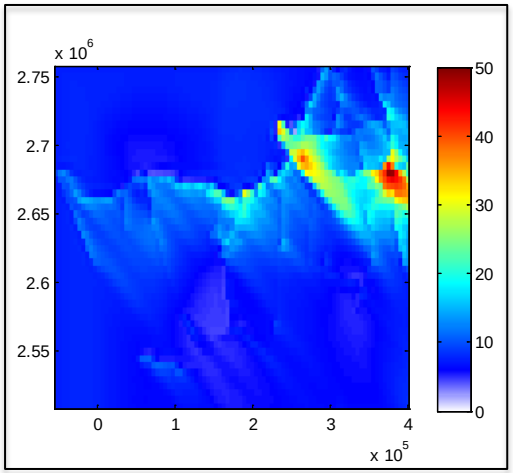


Annual average concentration of NO_2 over Oslo for 2014 as computed by the EPISODE air quality model.

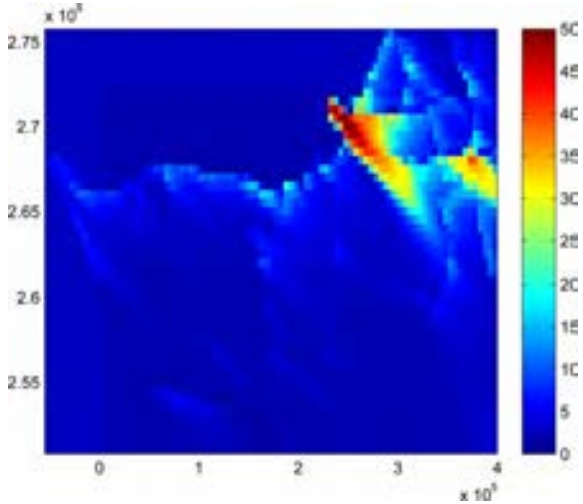


Observations

DATA FUSION



Combined map



Modelling results or other auxiliary data

Data fusion (as a subset of data assimilation) creates a value-added product by

a) Interpolating the observations in an objective way
b) “correcting” the model estimates with true observations

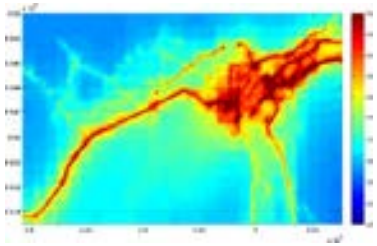
Data fusion method used here provides a combined concentration field by separately interpolating the observational residuals from a regression model and then combining both.

Mapping Methodology

- Theoretical basis
  - Data fusion is a subset of data assimilation techniques (Lahoz and Schneider, 2014)
  - We use geostatistical framework: Universal kriging approach
  - Analysis performed entirely in log-space
  - Explicit automated modelling of spatial autocorrelation
- In practice
  - Create static basemap for each mapping location
  - Retrieve crowdsourced sensor observations at each hour
  - Modify basemap based on latest observations using geostatistical data fusion
  - Final result are hourly maps with the current best guess for the NO_2/PM_{10}/$PM_{2.5}$ concentration field at all locations

Lahoz, W. A., and P. Schneider (2014), Data assimilation: making sense of Earth Observation, *Front. Environ. Sci.*, *2*(16), 1–28, doi:10.3389/fenvs.2014.00016.

Static basemap
(for each species and location)



Basemap: Provides information about general spatial patterns

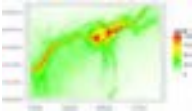
\+ Obs 15:00 CET

\+ Obs 16:00 CET

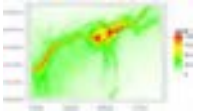
\+ Obs 17:00 CET

AQMesh observations: Provide information about current state of atmosphere at a few sampling locations

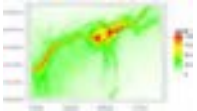
Fused 15:00 CET



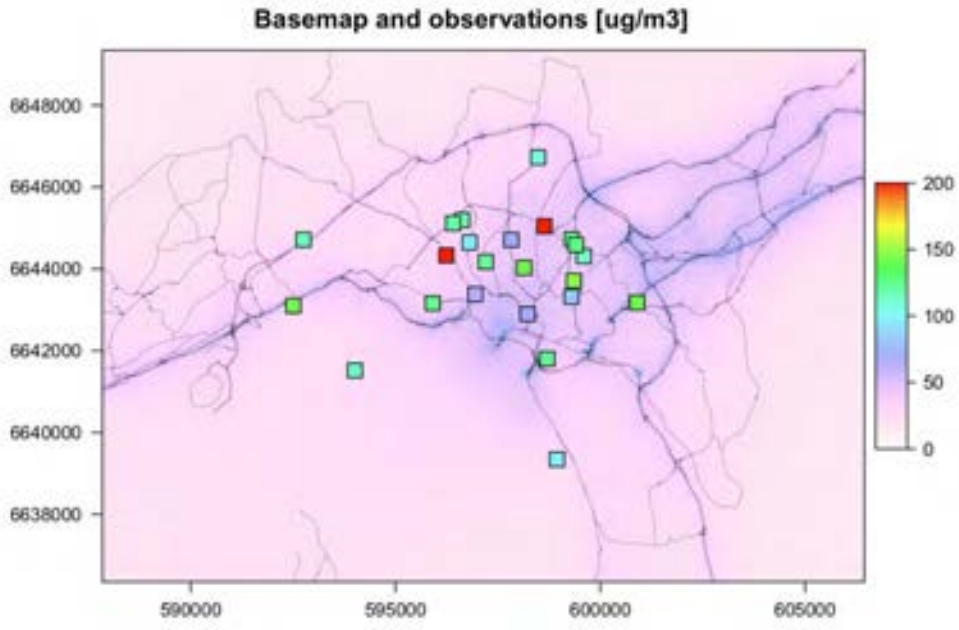
Fused 16:00 CET



Fused 17:00 CET

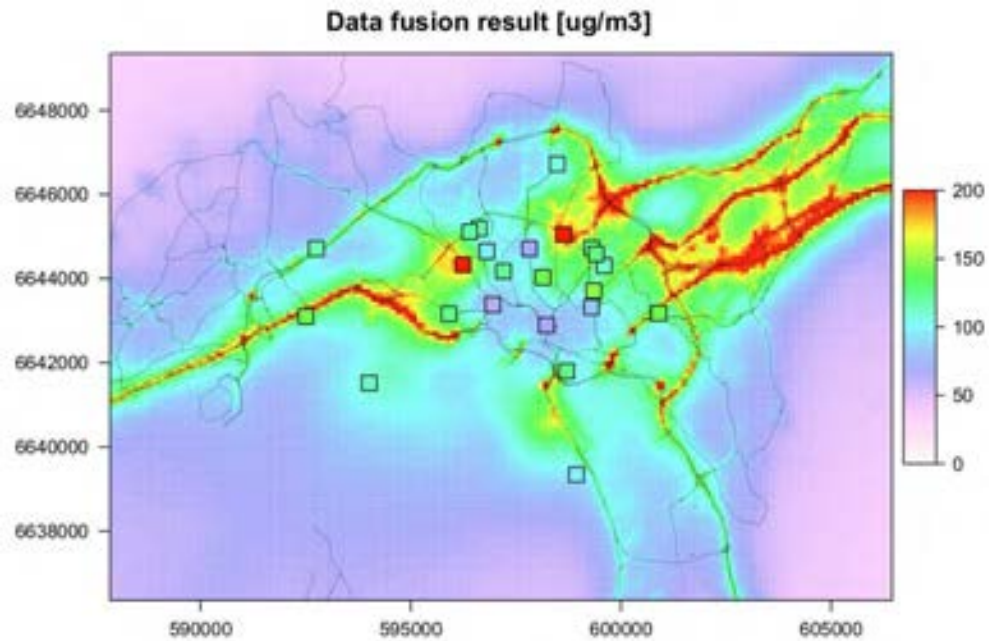


Fused map: Value-added product providing a best guess of current state of atmosphere for the entire domain

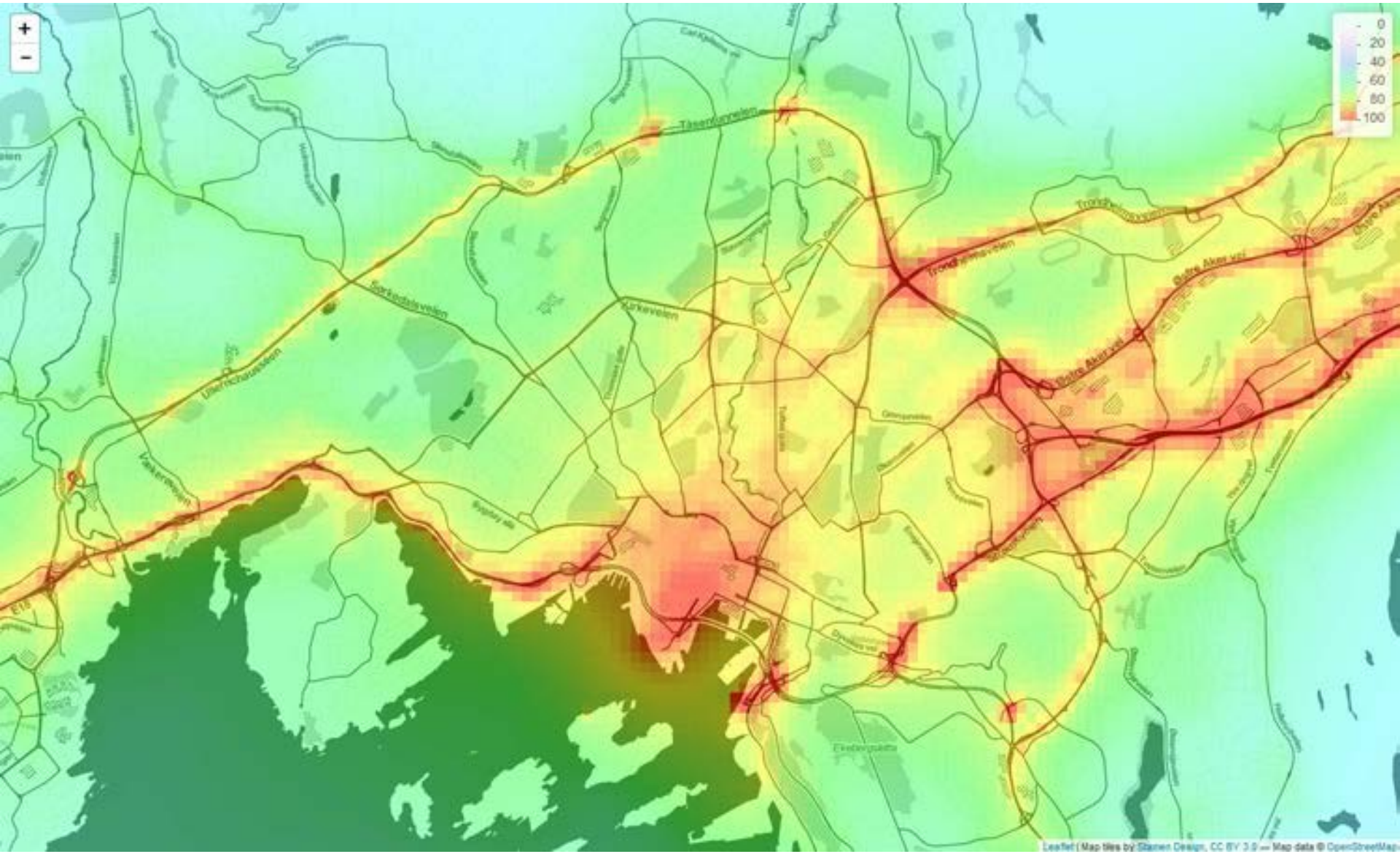


Oslo NO_2 model-derived annual average basemap (background) and observations from AQMesh nodes (markers) on 6 January 2016 at 9:00 UTC. Units in µg/m³.

Note: The sensors used here were all **co-located for several weeks** at the Kirkeveien AQ monitoring station **before deployment** and are thus **field-calibrated**!



Result of data fusion process (background) and observations from AQMesh nodes (markers). Units in µg/m³.

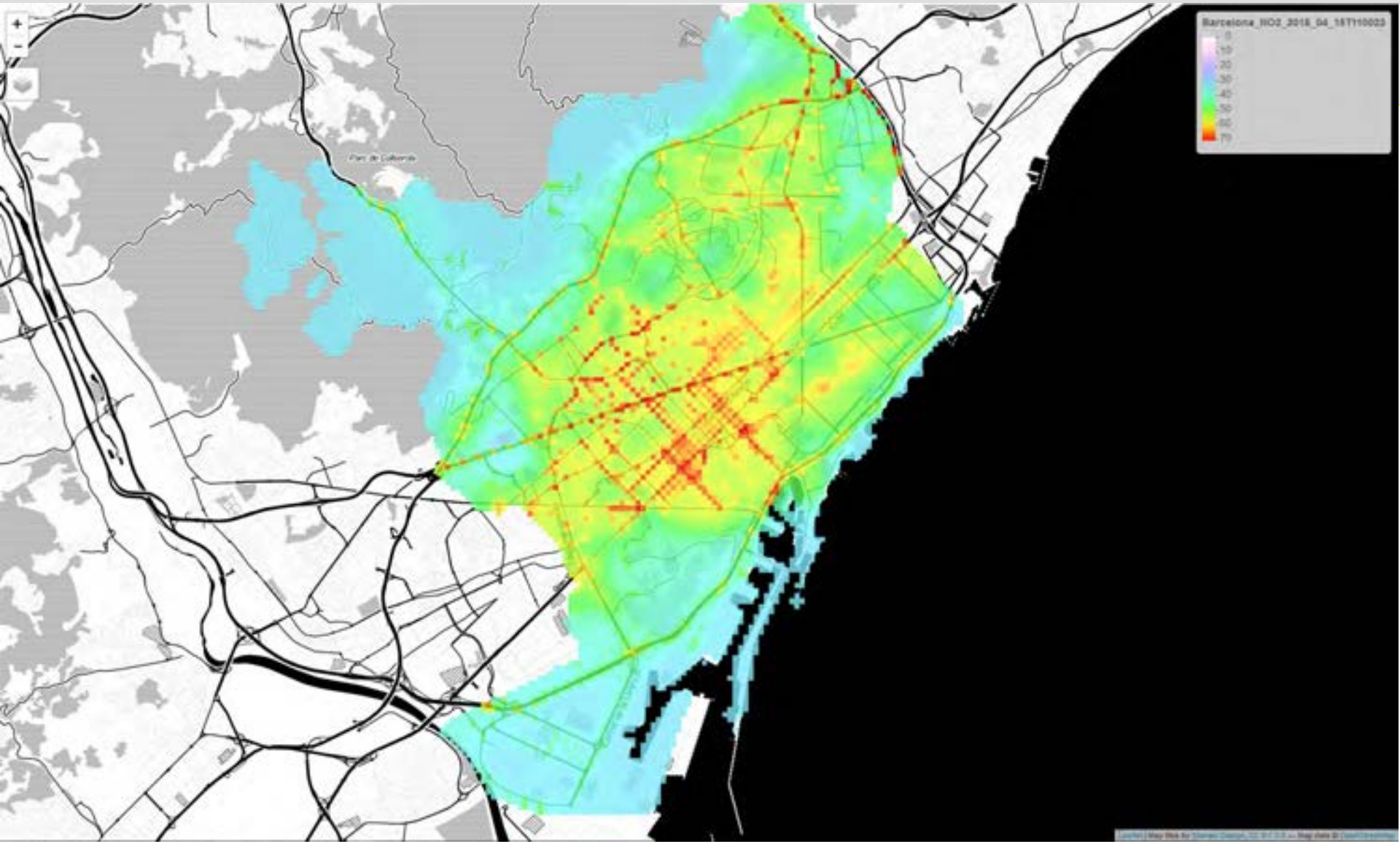


Example of a data fusion-based surface concentration field of NO_2 for Oslo, Norway, at 100 m spatial resolution (link).

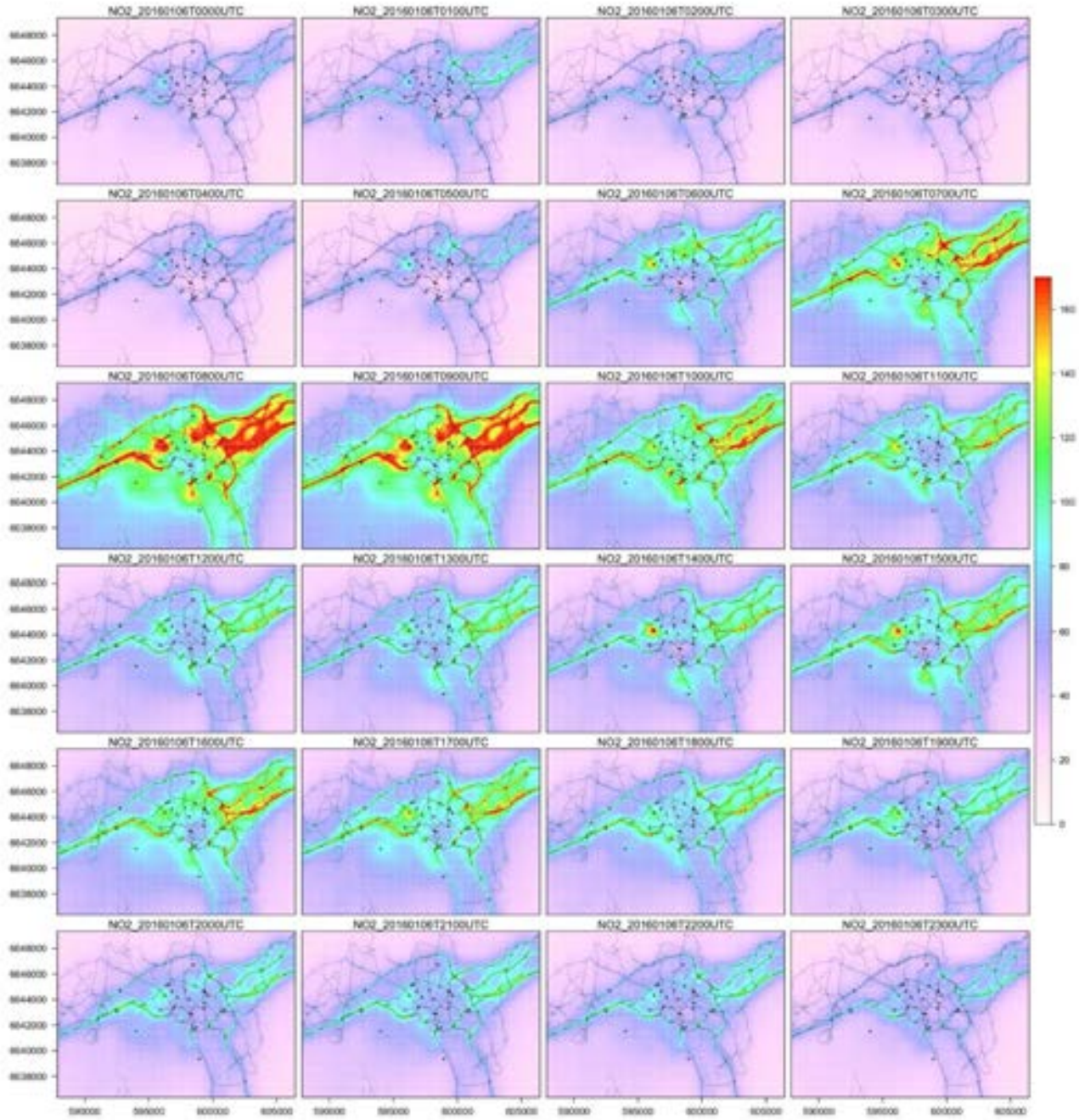
Applications of data fusion maps

Estimated real-time NO_2 concentrations along major Oslo bike paths, extracted from a data fused map.

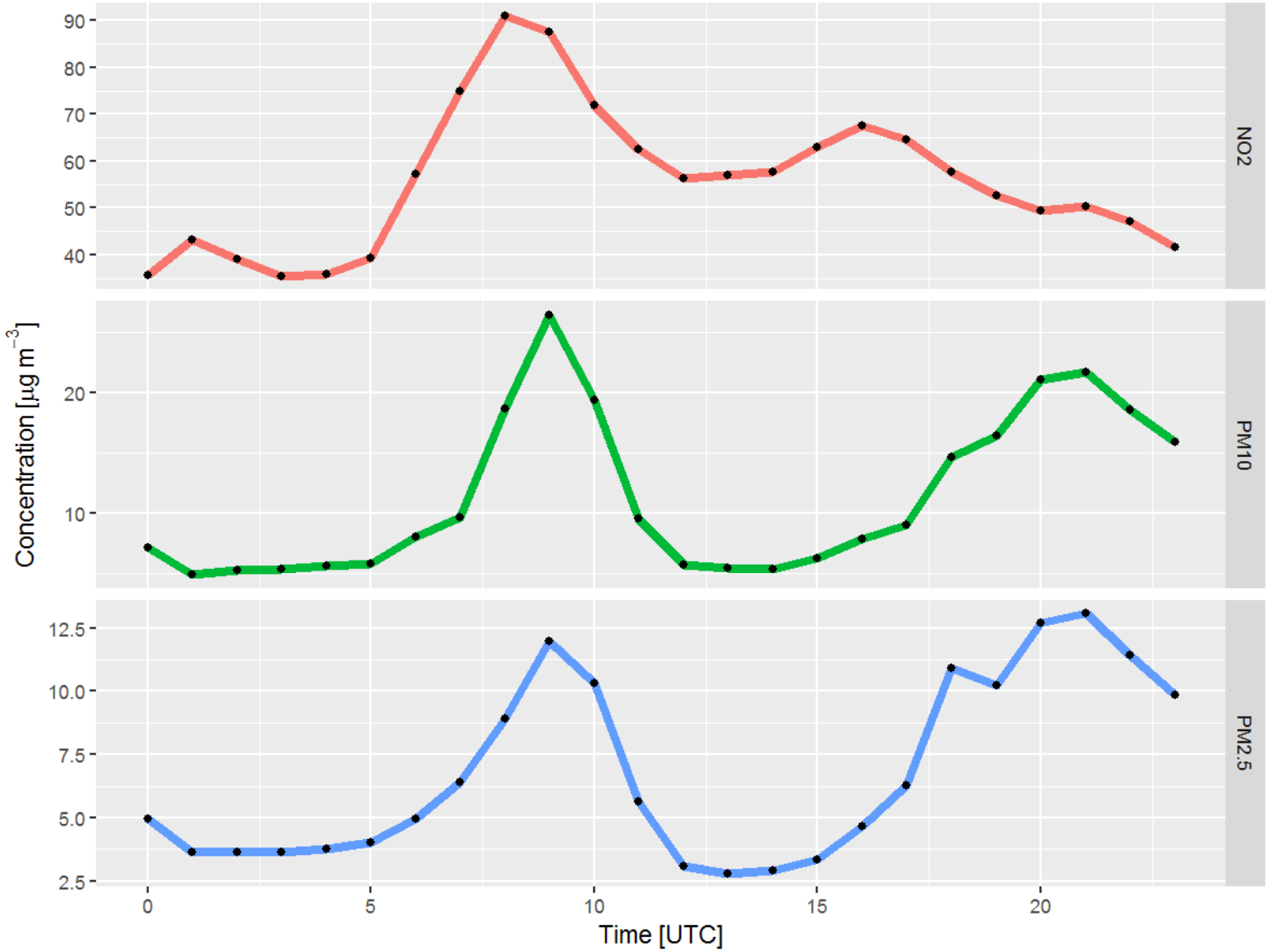
Schneider et al. (2016). Mapping urban air quality in near real-time. Microsensor meeting at RIVM, 13 February 2017, Utrecht, NL



Example of a data fusion-based surface concentration field of NO_2 for Barcelona, Spain, at 100 m spatial resolution (link).

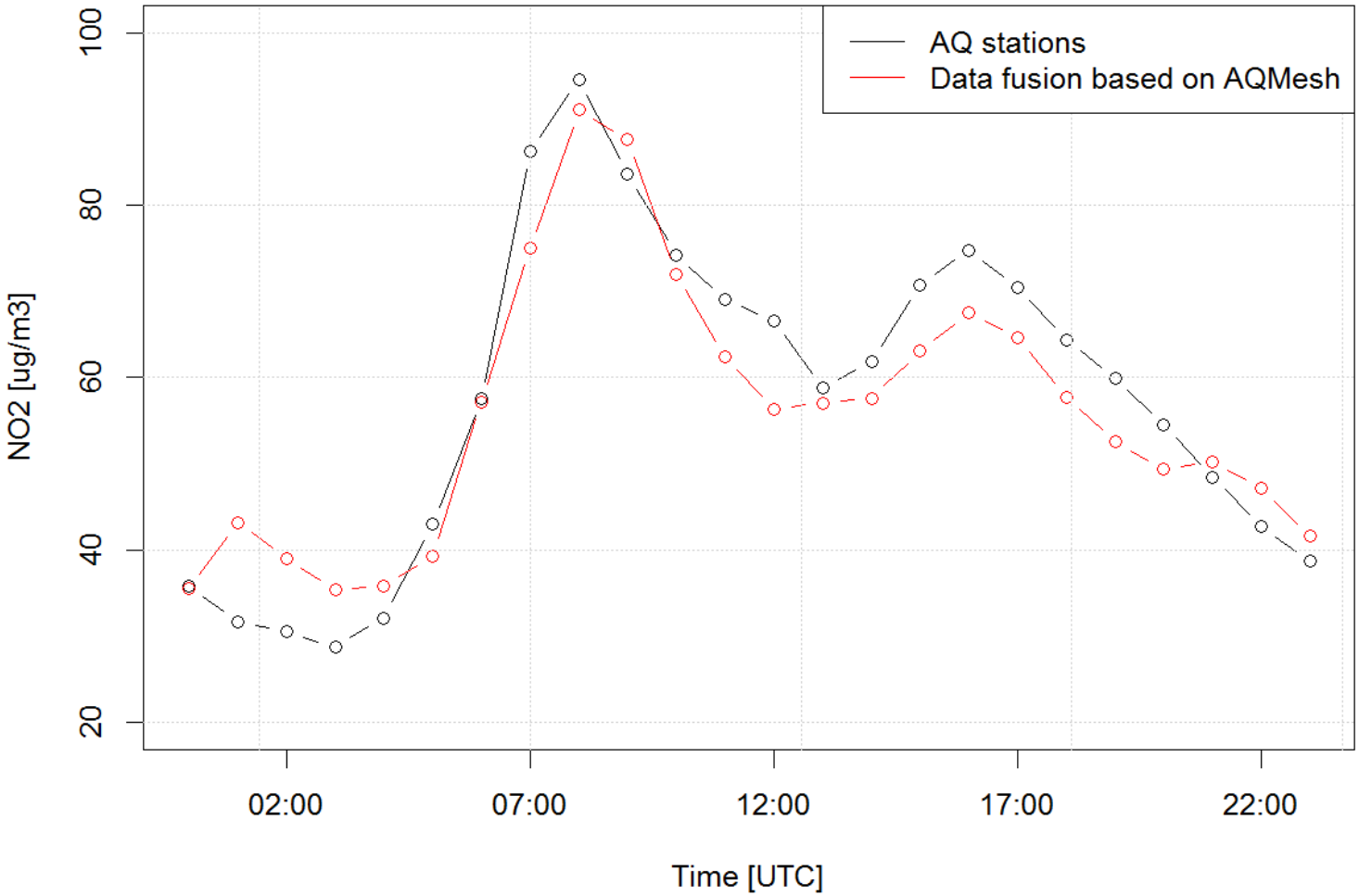


Example of 24 hours of data fusion results in Oslo, combining NO_2 measurements from the AQMesh units with a long-term average basemap derived from the EPISODE model, here shown for 6 January 2016



Data fusion maps: Daily cycle of NO_2, PM_{10}, and $PM_{2.5}$ for Oslo on January 6 2016 (NO2) and 22 March 2016 (PM).

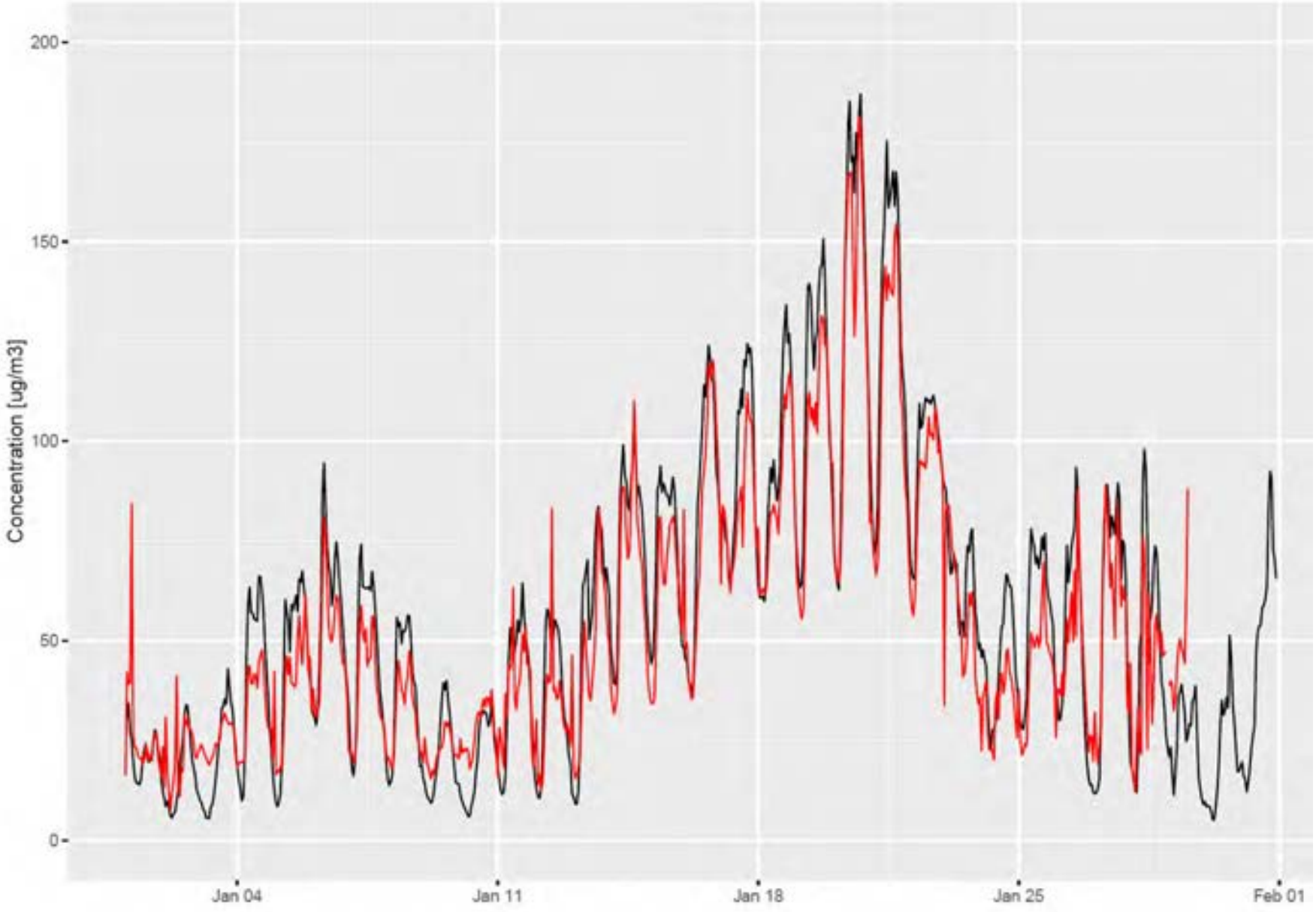
Comparison to AQ monitoring stations



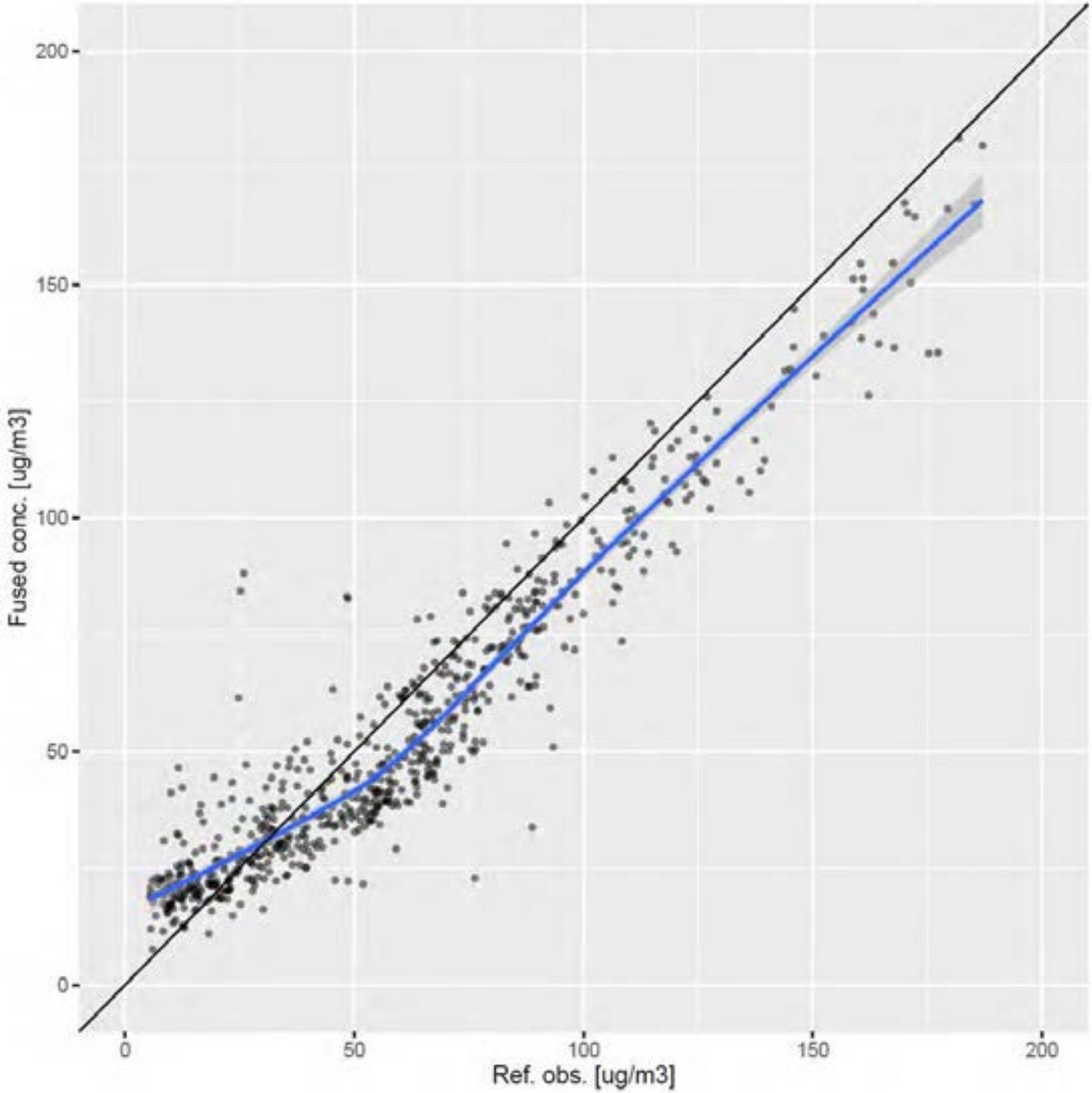
The fused maps not only replicate the patterns of the typical daily cycle, but are able to reproduce the overall magnitude in terms of actual concentrations.

This shows that despite high uncertainty at the individual sensor level, we can tease out a **useful and realistic signal** from an **entire network** sensor nodes.

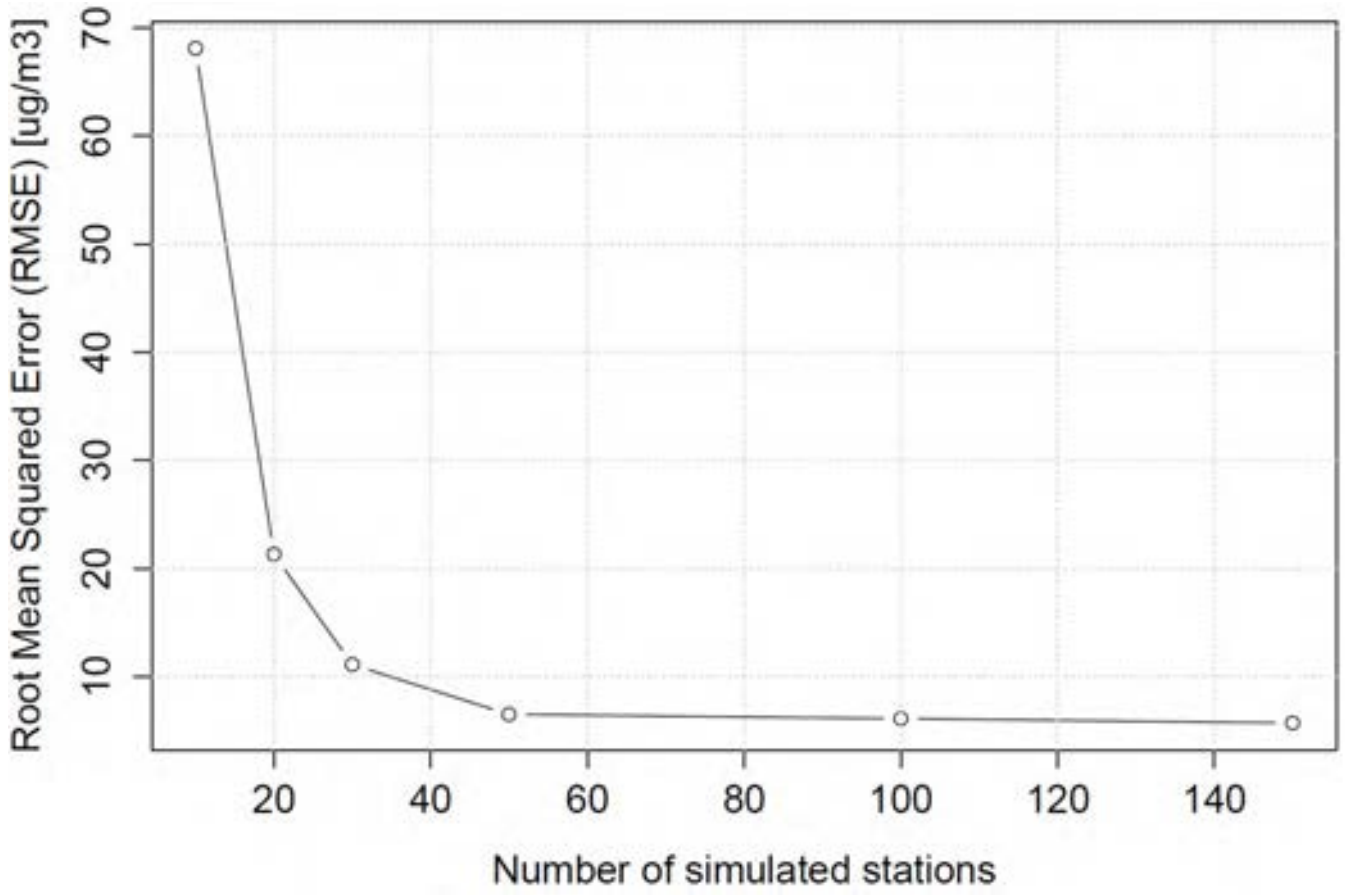
Entire daily cycle of NO_2 as measured by the reference air quality monitoring stations versus the NO_2 concentrations provided by the data fusion map.



Average NO2 concentration in Oslo in January 2016
Black line: Reference AQM stations
Red line: Data fusion of AQMesh low-cost sensor network and EPISODE


Fused conc. [ug/m3]
Ref. obs. [ug/m3]
0
50
100
150
200

Dependency of map quality on network size



Relationship between accuracy of AQ mapping based on data fusion and the network density/size using simulated observations.

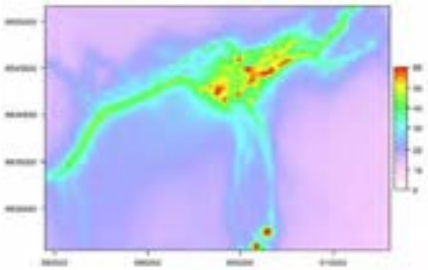
Note that this is plot only indicates the relationship for a very specific simulated example in Oslo, and is not representative for networks in other locations.

Take-home messages

- A method was developed for creating up-to-date urban-scale air quality maps from a network of static low-cost AQ sensors
- Resulting maps reproduce the overall spatial patterns of AQ in the city and at the same time quantitatively reproduce the observations
- Despite many challenges at the individual sensor level, low-cost microsensors allow for detailed high-resolution urban-scale mapping of air quality if several conditions are met:
  - the sensors are calibrated in the field and ideally calibration drift is avoided
  - A sufficient number of nodes is deployed
  - the sensor observations are combined with output from an air quality model in a suitable fashion
  - the “swarm knowledge” of the entire network is used
- Future advances in sensor technology and deployment density will significantly increase the usefulness of the sensors for mapping purposes

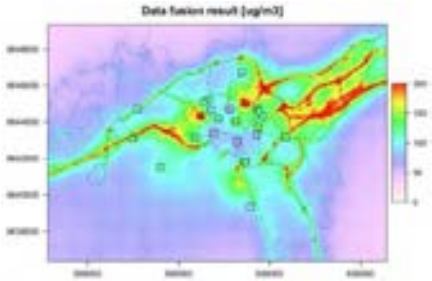
Observations

+



Model

=



Urban AQ Map

Thank you for your attention!

Contact: Philipp Schneider

Email: ps@nilu.no

www.nilu.no

www.citi-sense.eu