

CT6: Data fusion

Future steps on data fusion analysis

Jorge Sousa and Stijn Janssen

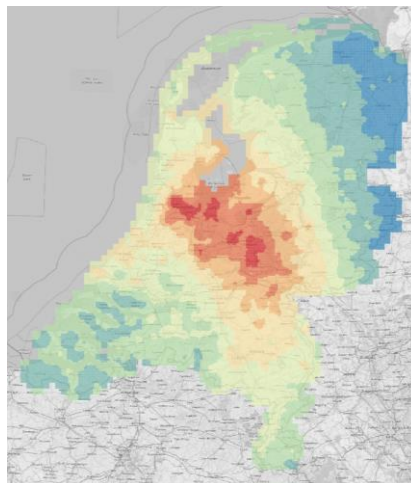
**Short overview of what has been done in CT6.
Open questions about the data fusion step.**

Observations



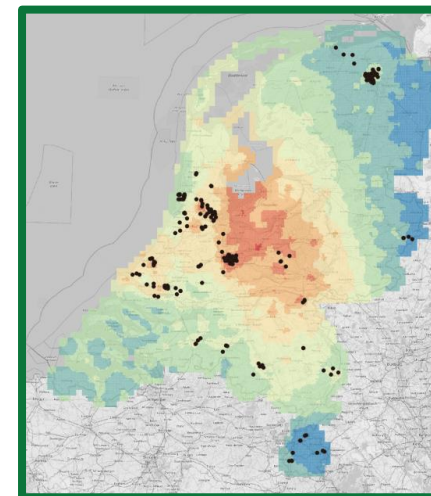
+

Map



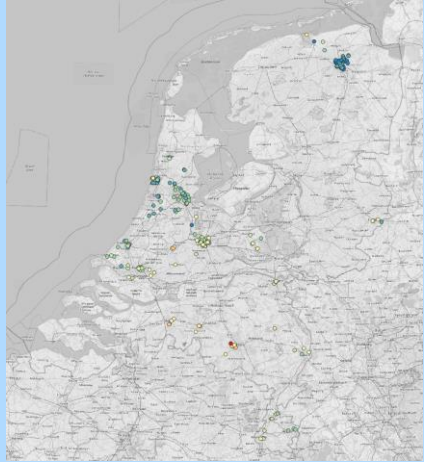
=

Fused map



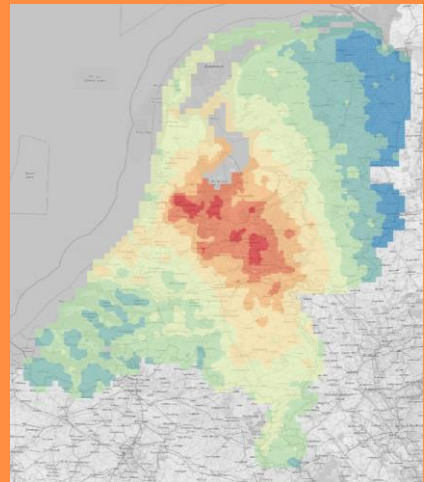
CT6: DATA FUSION

Measurements



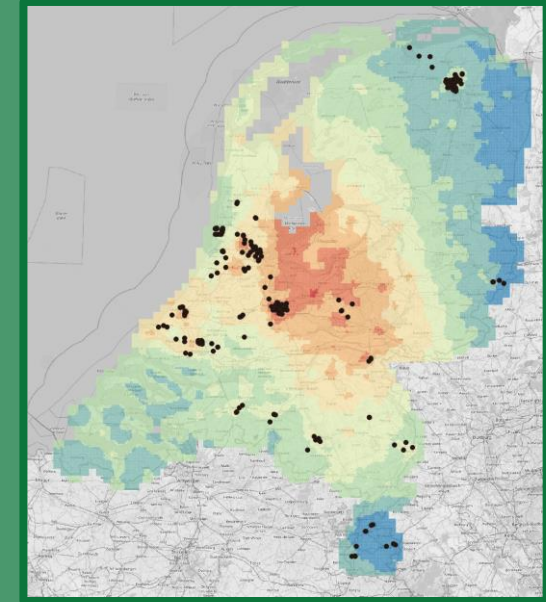
- Network sensor calibration
- Outlier detection
- Procedure to create synthetic data
- Sensitivity to sensor uncertainty
- Sensitivity to calibration procedure

Model



- Model uncertainty characterization (RIO)
- Procedure to create synthetic model data
- Sensitivity to different types of models
- Sensitivity to location/ scale of interest

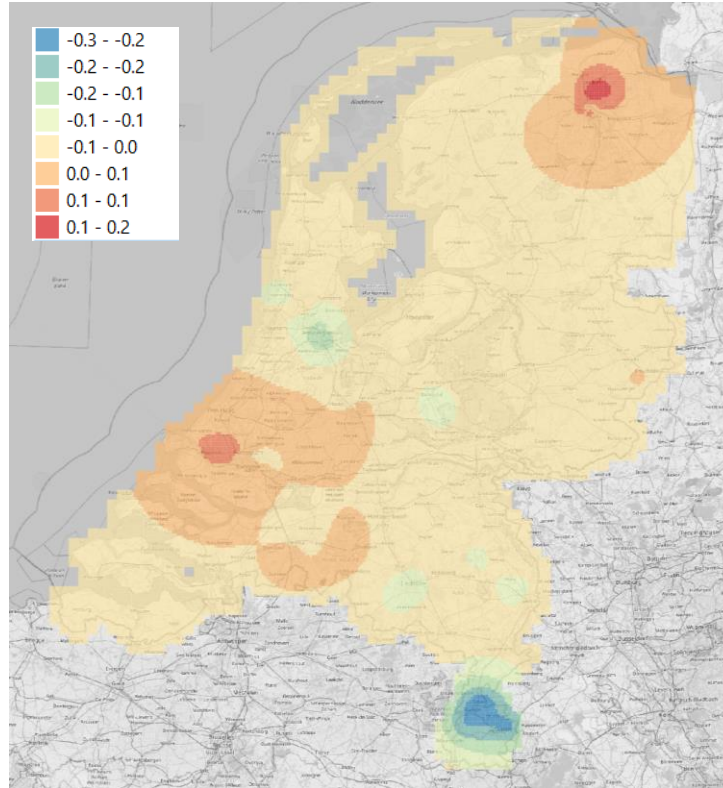
Fusion



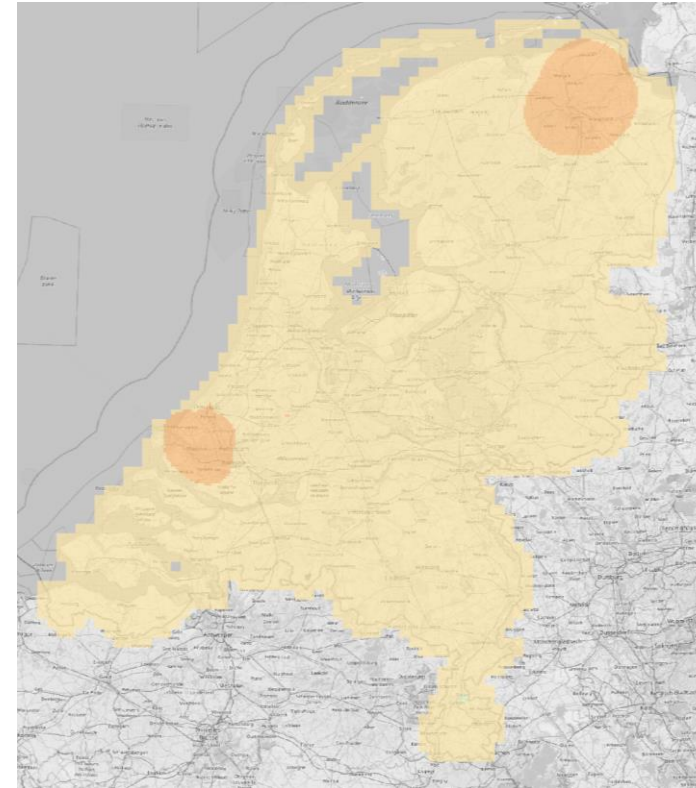
- Sensitivity to different fusion methods
- Sensitivity to sensor density
- Sensitivity to uncertainty knowledge

Impact of different levels of uncertainty in the sensors:

10% Random uncertainty in sensors



50% random uncertainty in sensors



Open Question: What is the required sensor network accuracy to have a relevant impact on the air quality map?

Erroneous characterization of the sensor/ model uncertainty:

- If we assume a too high obs. uncertainty:  *We ignore completely the sensor data
(We might lose some relevant information)*

- If we assume a too low obs. uncertainty:  *We nudge too much to the sensor data
(we might add noise to our map)*

- If we do it right:
(test with synthetic data)



Still extract some useful information from the network of sensors

Open Question: To which extent do we need a good characterization of sensor performance?

Assimilation/ Fusion further steps:

Address some of the open questions:

- Sensitivity to sensor uncertainty. **What accuracy is sufficient to have relevant impact on the map?**
- Sensitivity to sensor calibration strategies. **Does the calibration algorithm have a large impact?**
- Sensitivity to different fusion methodologies. **Is there a recommended methodology?**
- Uncertainty characterization. **How well do we need to know the uncertainties?**
- Evaluation procedure: **Synthetic data? Or real data by splitting the network?**

