OpenSense: City-Scale Air Quality Monitoring with Wireless Sensor Nodes

D. Hasenfratz, O. Saukh, C. Walser, C. Hueglin, M. Fierz, Lothar Thiele

Computer Engineering and Networks Laboratory
ETH Zurich, Switzerland

18.12.2013
COST Action TD1105, Cambridge, UK

ETH Zürich EMPA n|w Fachhochschule Nordwestschweiz nano-tera.ch

Urban Air Quality

- Major concern in many cities worldwide
- Responsible for respiratory and cardiovascular illnesses

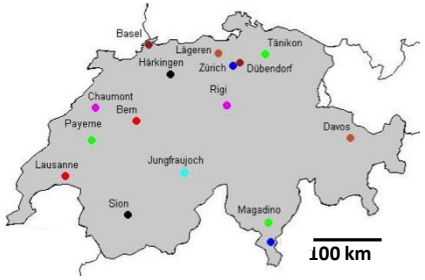
Air Pollution Monitoring

- Static measurement networks operated by national authorities
- Measurements are highly reliable and very accurate
- High acquisition and maintenance costs limit number of installations

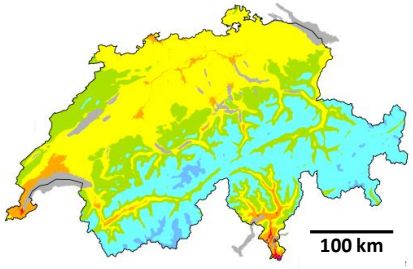
NABEL station in Zurich, Switzerland

Air Pollution Maps

- Today: Limited spatial or temporal resolution

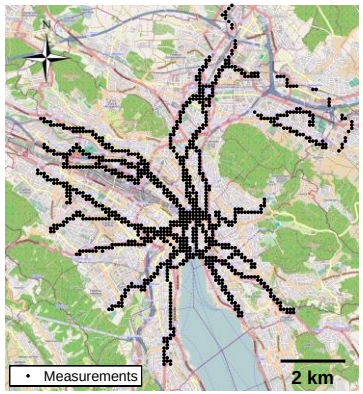


Static measurement network

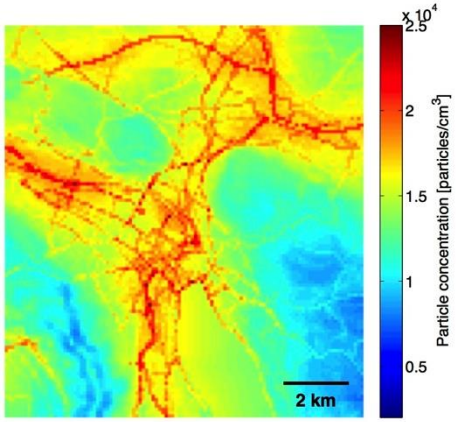


Fine particle pollution map

- Goal of OpenSense: Increase spatio-temporal resolution



City-wide coverage



Fine-grained pollution map

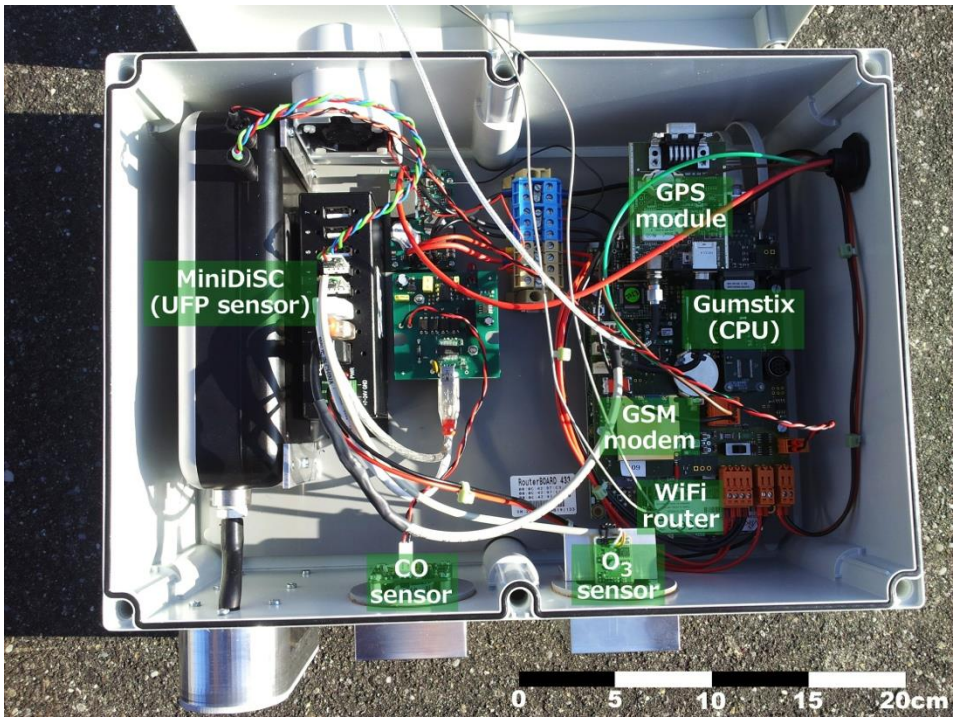
Mobile Sensor Nodes

- Sensors: ozone, carbon monoxide, ultrafine particles, temperature, humidity
- Location: GPS
- Communication:

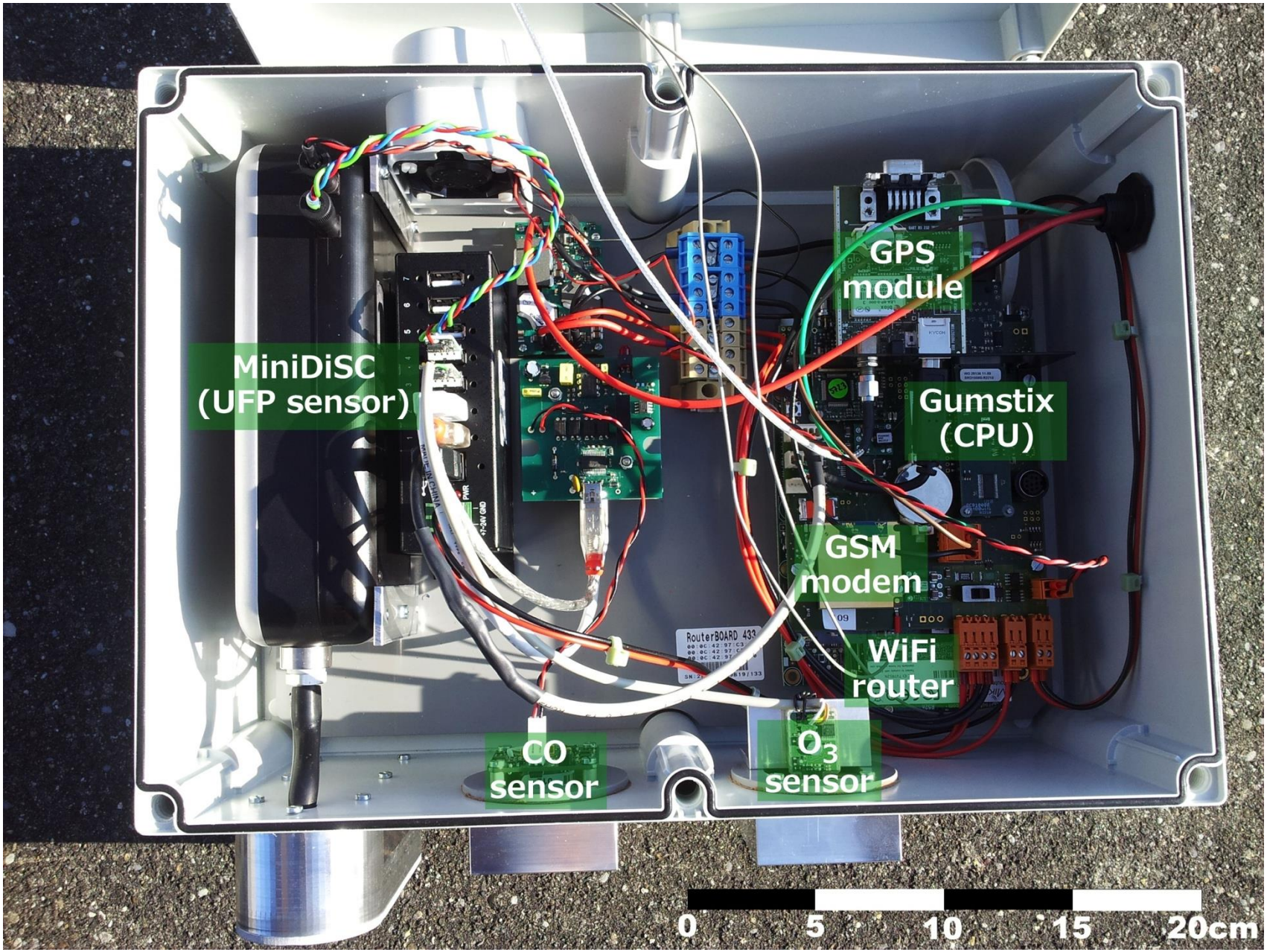
  GSM (cellular network)
- Interfaces:

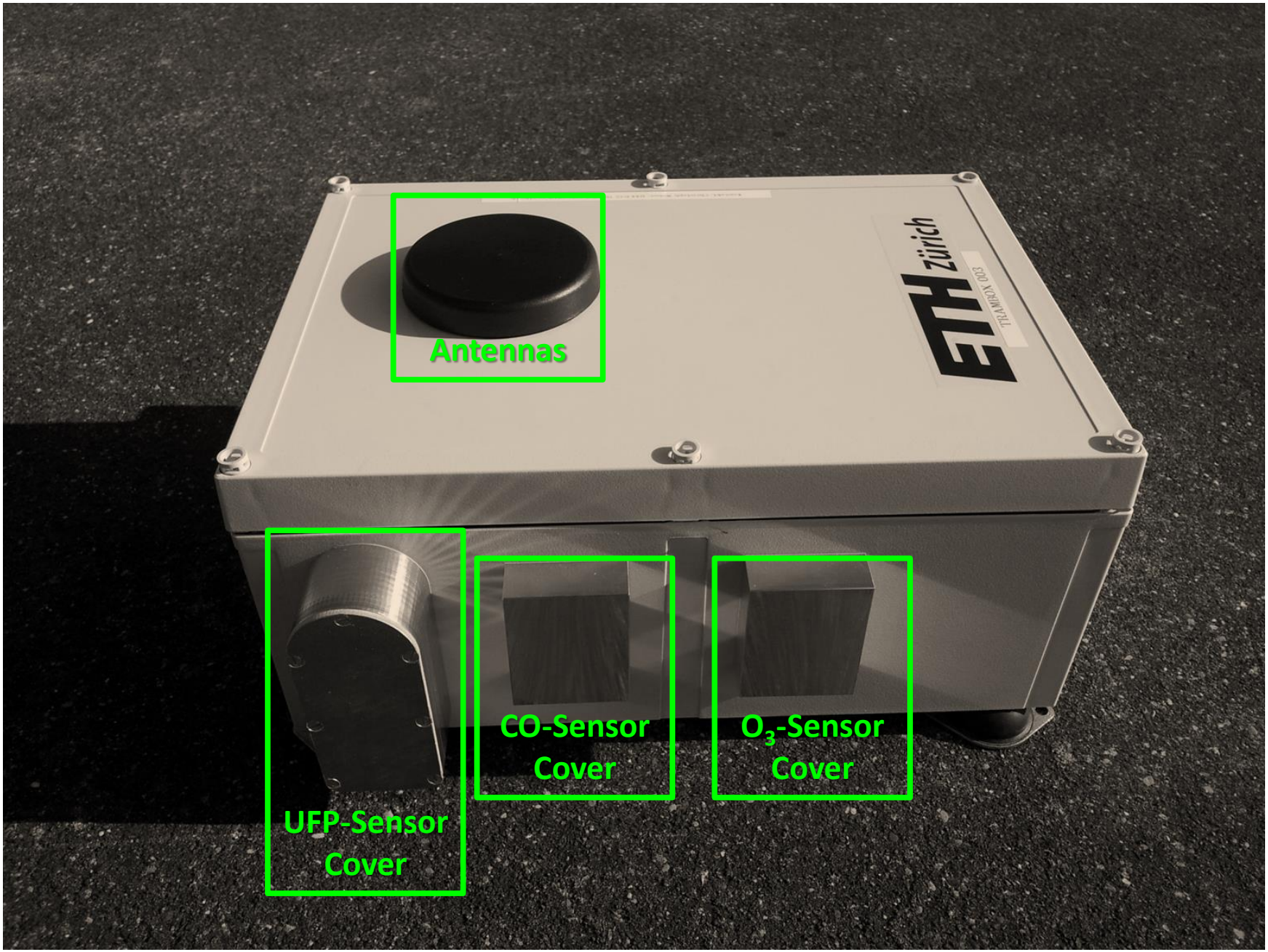
  USB, analog/digital
- Power supply:

  external (streetcar)



Air quality sensor node


GPS module
MiniDiSC (UFP sensor)
Gumstix (CPU)
GSM modem
WiFi router
CO sensor
O_3 sensor
0
5
10
15
20cm


Antennas
UFP-Sensor Cover
CO-Sensor Cover
O_3-Sensor Cover
ETH zürich

Deployments

- Static deployment

On top of a static measurement station

- Testing new sensors: stability, accuracy
- Long-term sensor tests

- Mobile deployment

On top of streetcars in the city of Zurich

- 10 nodes
- 20 hours per day in operation
- 2 years of measurements
- > 50 million data points

Deployments

- Static deployment

On top of a static measurement station

- Testing new sensors: stability, accuracy
- Long-term sensor tests

- Mobile deployment

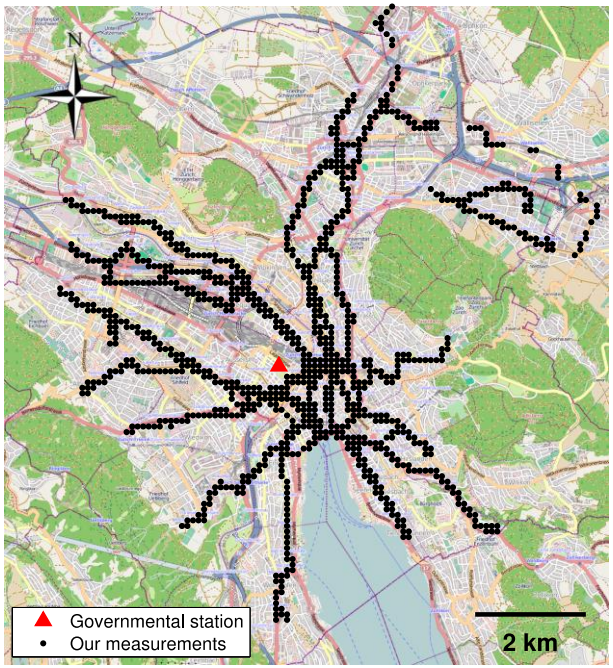
On top of streetcars in the city of Zurich

- 10 nodes
- 20 hours per day in operation
- 2 years of measurements
- > 50 million data points

Mobile Sensor Network

- 10 sensor nodes on top of 10 streetcars
- Streetcars are not bounded to specific tracks

Sensor nodes are installed on top of streetcars



Measurement coverage

Monitoring Ultrafine Particles (UFPs)

- Particles with a diameter < 100nm
- Most countries have mass emission limits for particulate matter PM_{10} and $PM_{2.5}$, but no restrictions for UFP
- Adverse health effects of UFP most probably underestimated

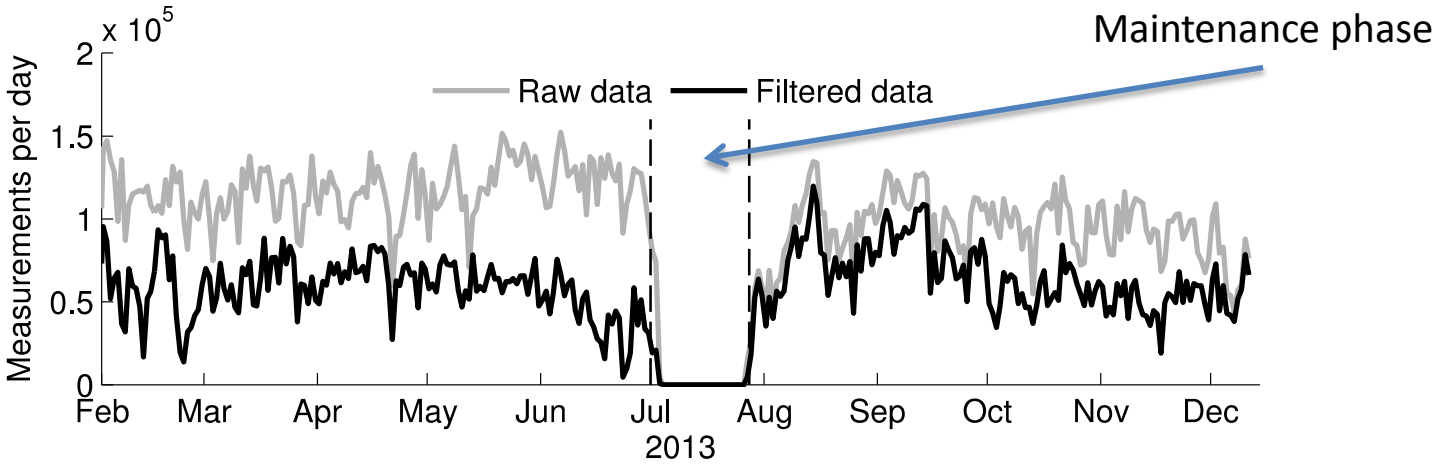
UFP sensor (MiniDisc)

Problem: Lack of spatially resolved exposure data, lack of epidemiological studies

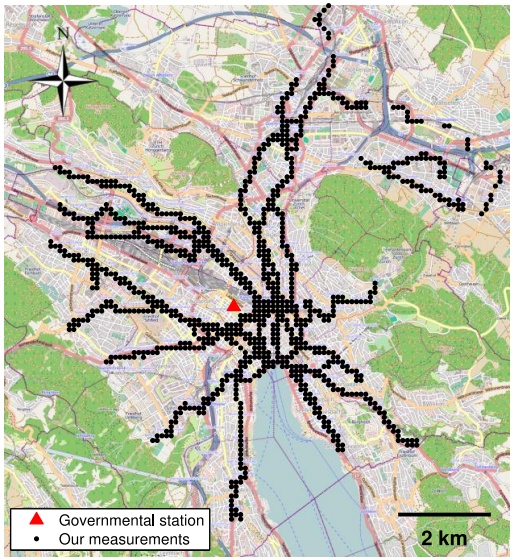
UFP Data Set

- 10 MiniDiscs are installed on top of 10 streetcars
  - Five installed in April 2012, five more in January 2013
  - Sampling rate of 20Hz, aggregated to one packet per 5s
  - Collected more than **40 million measurements**

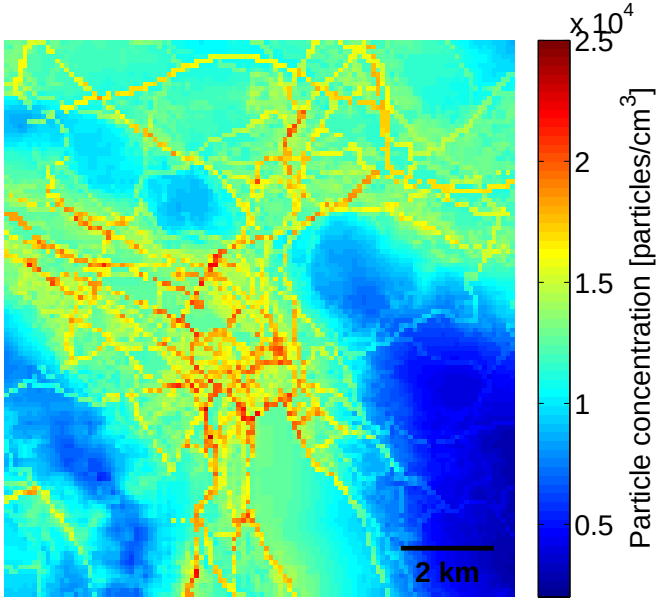
    (after filtering around 25 million remaining)



From Single Measurements to Fine-Grained Pollution Maps



Single measurements



Fine-grained pollution map

Processing steps:

Raw data → Data filtering & calibration → Data validation → **LUR model** → Pollution map → Map validation

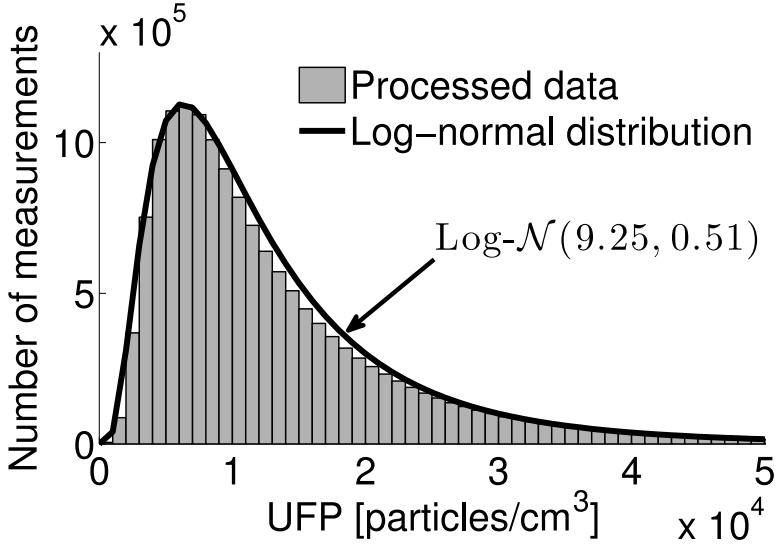
Data Validation

- Has the harsh deployment setting on top of streetcars an impact on data quality?
  - Long-term unattended operating times
  - Mobility and constant vibrations
  - High temperature and humidity variations

- Good data quality is a must for the development of reliable pollution maps
- **Challenge: only very sparse ground truth is available**

Statistical Distribution

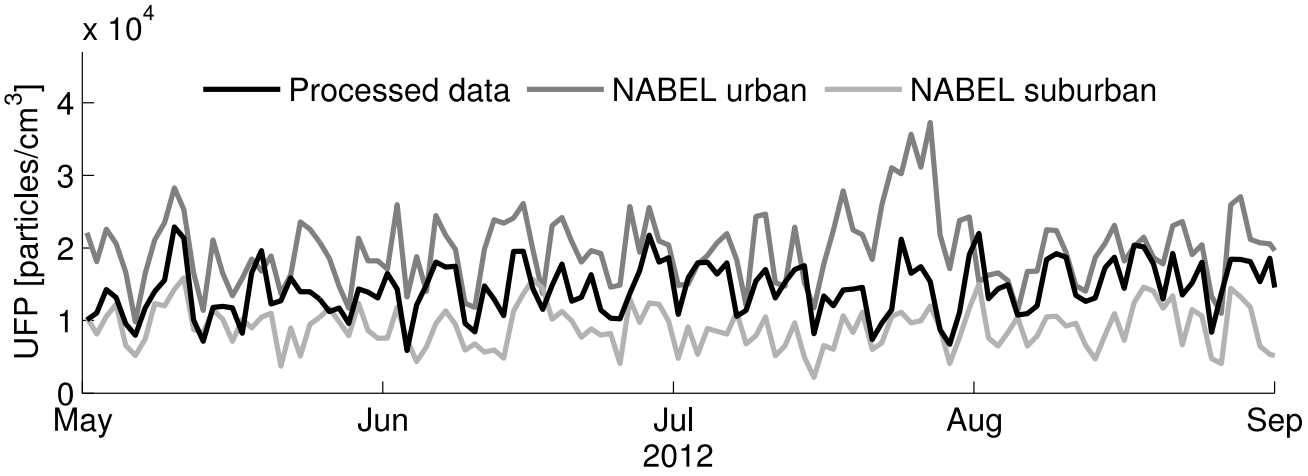
- UFPs are approximately **log-normally distributed**
- Log-normal distribution (black) with mean and standard deviation of the UFP data (gray)



- Distribution of the processed data closely follows a log-normal distribution

Comparison to High-Quality Data

- Comparison to UFPs measured by static stations of the Swiss National Air Pollution Monitoring Network (NABEL)
- Locations: **urban heavy traffic** and **suburban**



- Daily average measured UFP concentration corresponds well to the measurements of the two static stations

Developing Land-Use Regression Models to Create Pollution Maps

- Land-use regression (LUR) models widely used to assess spatial variation of air pollutants
- Use land-use and traffic characteristics (explanatory variables) to model pollution levels:
  1. Evaluate dependency between explanatory variables and monitored pollution levels
  2. Model pollution levels with the found relationships at locations without measurements but land-use data

Explanatory Variables

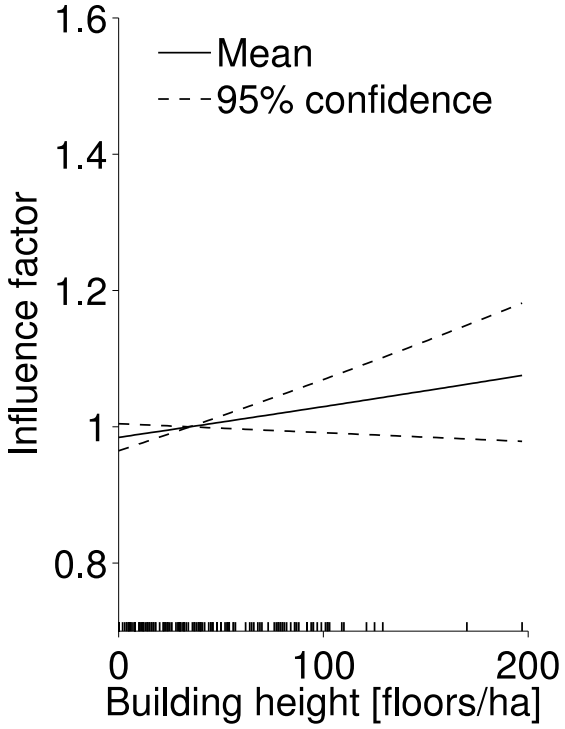
- Resolution: **100m x 100m (1 hectare)**
- **12 variables**:

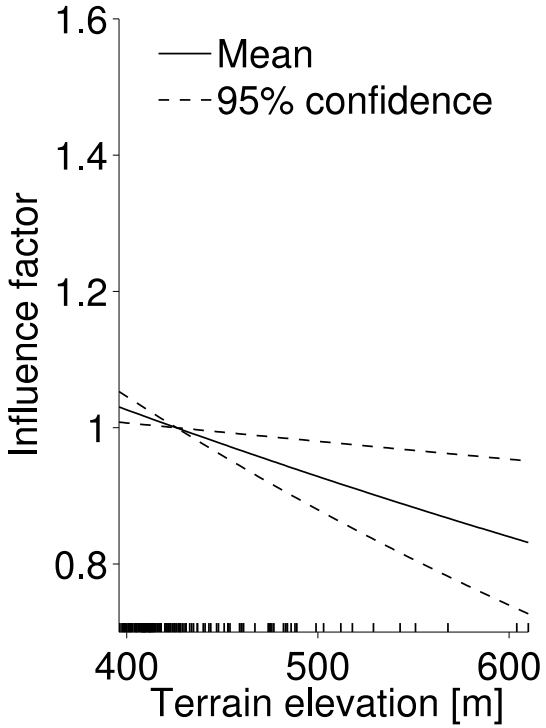
Variable [unit]	Variable [unit]
Population [inhabitants/ha]	Industry [industry buildings/ha]
Building height [floor levels/ha]	Heating [oil and gas heatings/ha]
Terrain elevation [average m (asl)/ha]	Road type [busiest road type/ha]*
Distance to next road [m]	Distance to next large road [m]†
Terrain slope [average degree/ha]	Terrain aspect [average degree/ha]
Traffic volume [vehicles per day/ha]	Distance to next traffic signal [m]

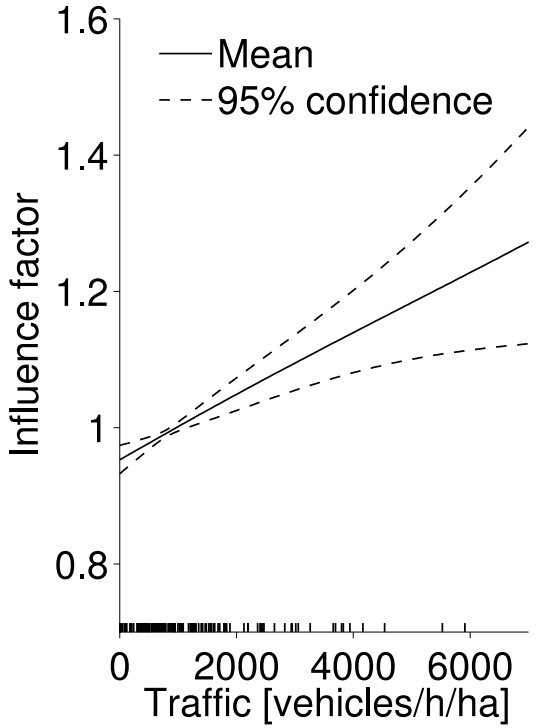
*Five road types: residential, tertiary, secondary, primary, and freeway.
†Road types classified as large: secondary, primary, and freeway.

Influence Factor

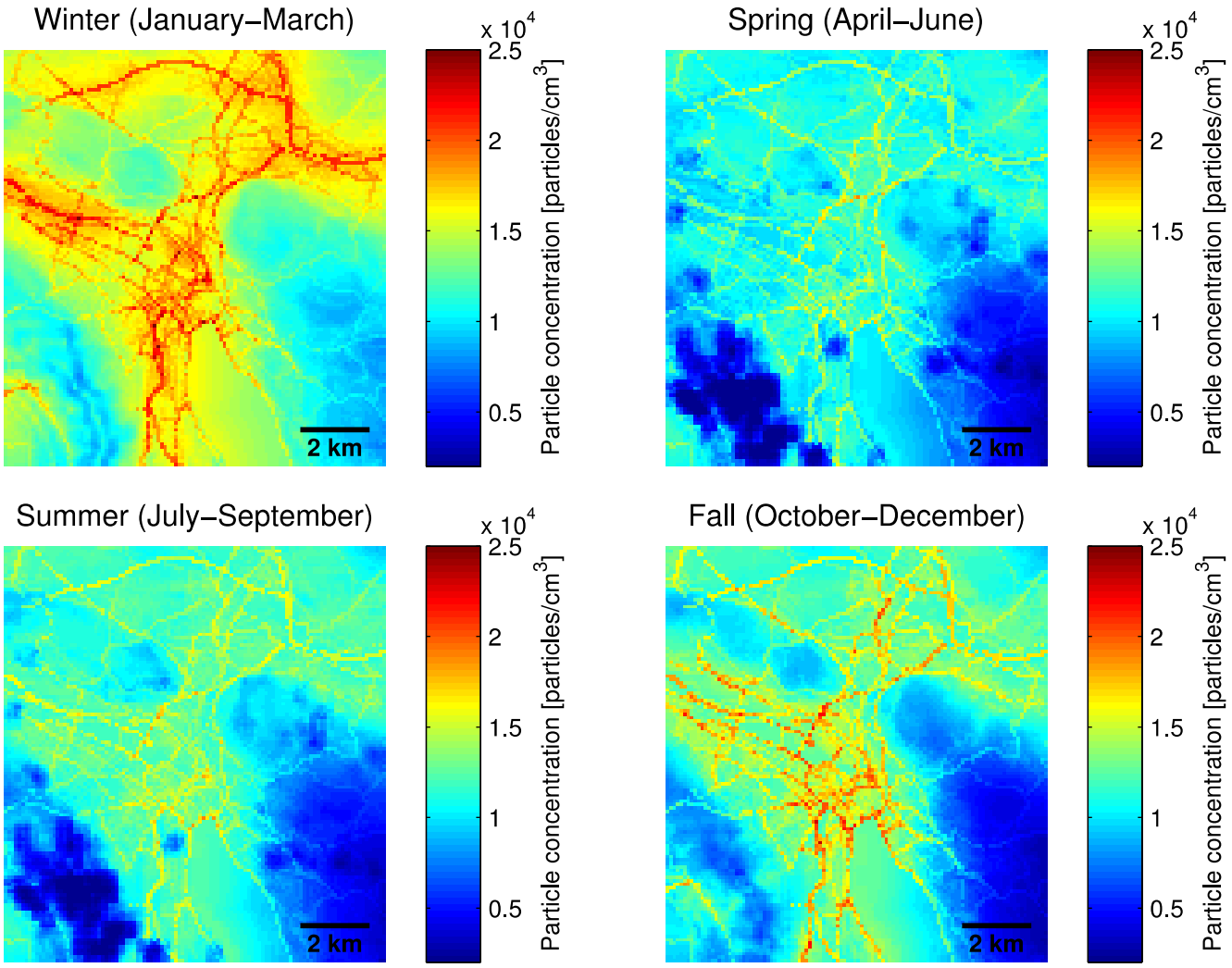
- Every variable has an **influence factor** on the modeled pollution level







Pollution Maps of Ultrafine Particles

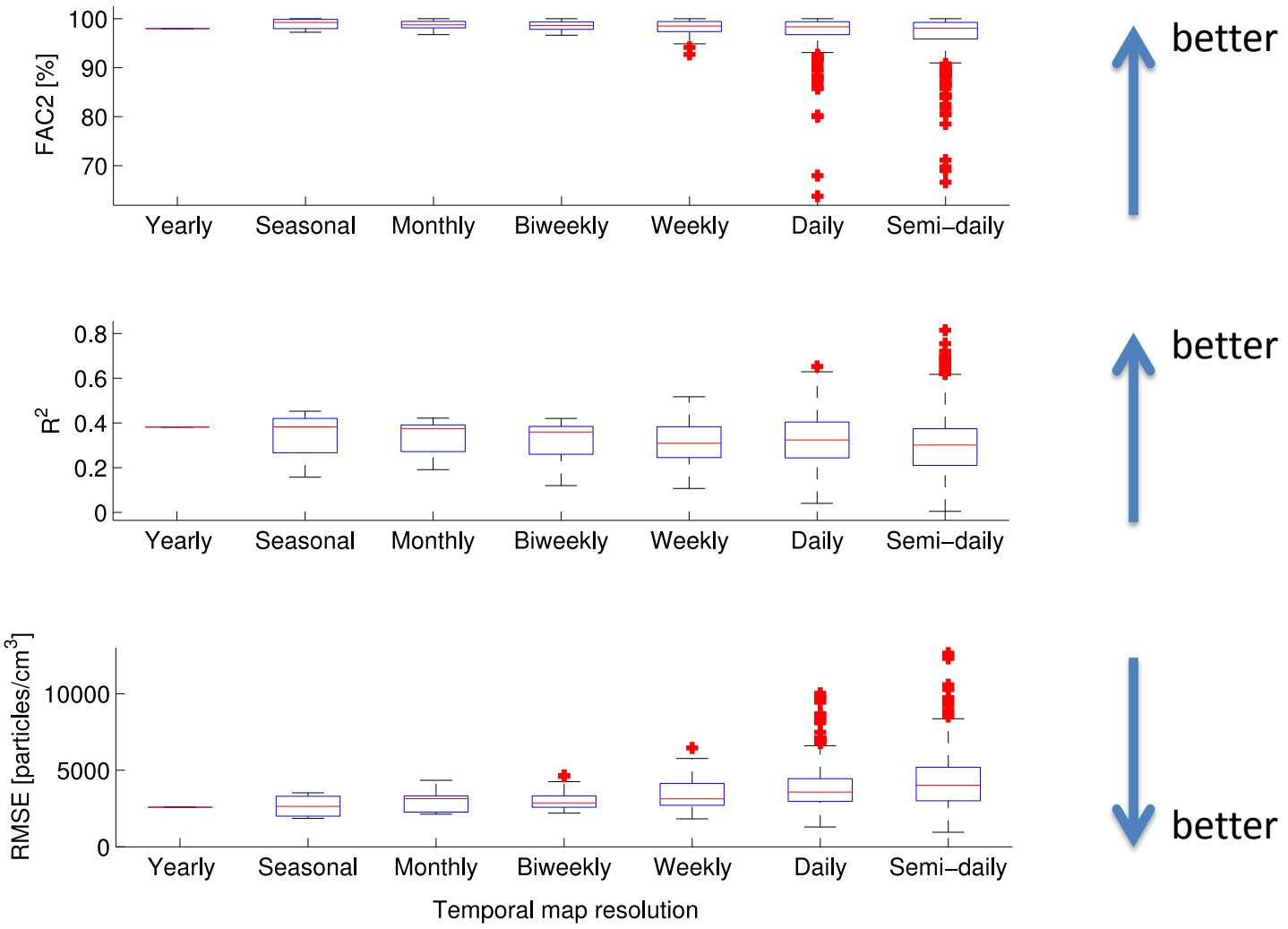


Model Performance (1/2)

- Generated **989 air quality models**:
  - 1 year of measurements: April 2012 to March 2013
  - Spatial resolution: 100m
  - Temporal resolution: **yearly to semi-daily**

- Metrics
  - Factor of 2 measure (**FAC2**): Fraction of predicted concentrations with an error less than a factor of two
  - Coefficient of determination (**R^2**): Indicates how well predicted concentration fits measurements
  - Root-mean-square error (**RMSE**): Quantifies difference between predicted and measured concentrations

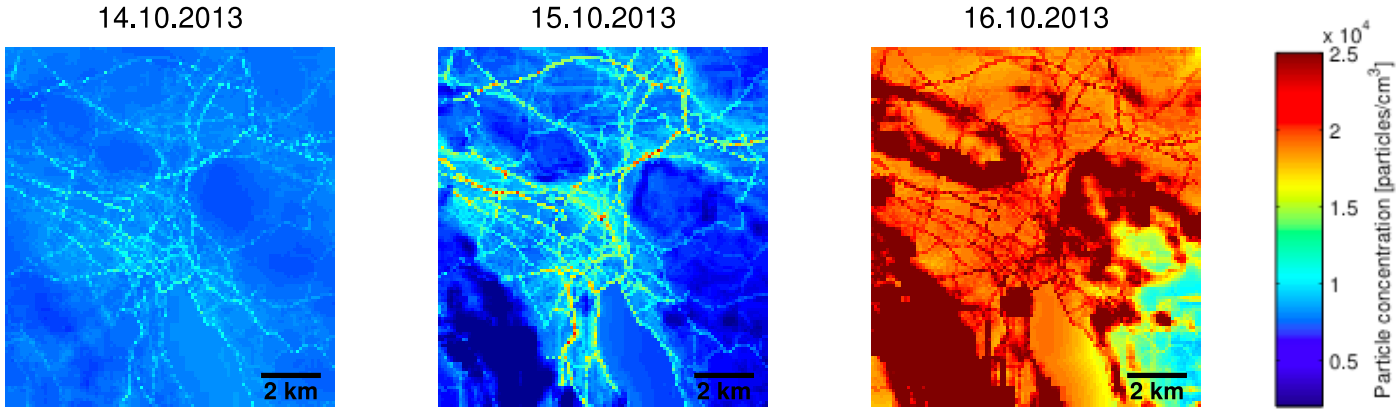
Model Performance (2/2)



Model Performance (2/2)

Challenges

- Lower number of measurements available to calculate the relationships to the explanatory variables
  - Limited temporal and spatial coverage of the measurements
  - Single erroneous and inaccurate measurements (e.g., outliers) have higher impact on the model



Outlook

- Increase spatial resolution (e.g., 10m resolution)
- Increase temporal resolution towards the goal of real-time pollution maps
- Analyze other pollutants, such as ozone (O_3), carbon monoxide (CO), and nitrogen dioxide (NO_2)
- Extend the mobile sensor network with measurements from locations not covered by the streetcar tracks

Website: www.opensense.ethz.ch

Backup Slides

Carbon Monoxdie (CO)

- Electrochemical gas sensor from Alphasense