Mobile vs. static sensor networks for urban air quality assessment

Martin Fierz[1], Michael Müller[2] and David Hasenfratz[3] [1]FHNW [2]EMPA [3]ETHZ contact: martin.fierz@fhnw.ch

Introduction

Mobile monitoring networks offer advantages over traditional static monitoring stations, as they allow much better spatial coverage of an area than static stations. The better spatial coverage comes at the price of a lower temporal resolution, as each "site" is monitored much less often. Obvious questions are: is it worth losing time resolution to increase spatial resolution? Which type of network delivers "better" results? What properties should mobile sensor networks have in order to perform well?

Here, we look at results from ultrafine particulate (UFP) measurements from the Opensense network; a streetcar-based mobile monitoring network operated by ETH Zürich. From early 2013 on, it contained 10 UFP sensor nodes of FHNW which collected data until April 2014.

Sensor nodes

Each sensor node was equipped with GPS for accurate time/position measurement, and with a miniDiSC for UFP detection. The miniDiSC measures particle number, lung-deposited surface area (LDSA) and average particle diameter simultaneously with the high time resolution necessary for mobile sensing.

The sensors were only accessible for service at night and distributed over different streetcar depots. Therefore, sensor reliability was very important from a practical point of view.
In the standard version, the miniDiSC proved too unreliable for this type of deployment and the instruments had to be modified in multiple ways to allow for longer service intervals.

Data & data reduction

The network generated large amounts of data – over 50 million UFP measurements were collected during the deployment. A grid of hectare-sized cells was used to aggregate the data, i.e. data in the same hectare-cell was averaged to a mean value of this cell. Raw data was filtered to remove data points where the instruments reported errors.

Modelling

The filtered and averaged data was used for land-use regression modelling (LUR). In this model, the UFP data (either particle number or LDSA) is explained by predictor variables P_i, such as traffic count and building density per hectare cell. The simplest model is linear:

$$UFP = c_1 \cdot P_1 + c_2 \cdot P_2 + \cdots + c_n \cdot P_n$$

Results

The following figures show a typical path from raw data to a model. Figure 1 shows average LDSA concentrations for a two-month period in Winter 2013.

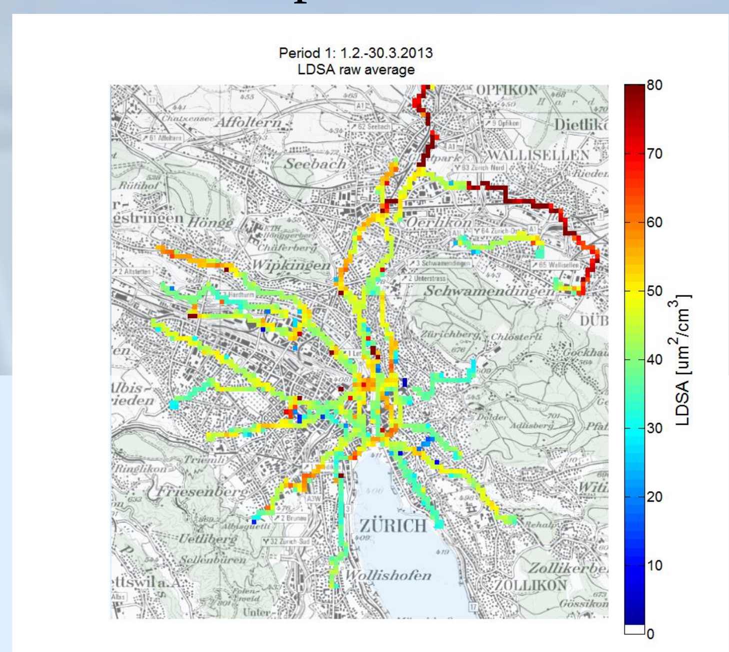


Figure 1: raw LDSA averages

For modelling, we required a minimal data coverage - in figure 2, only cells with "enough" coverage are retained – defined as at least 40 measurements per cell and day, and for at least 1/3 of the days in the measurement period:

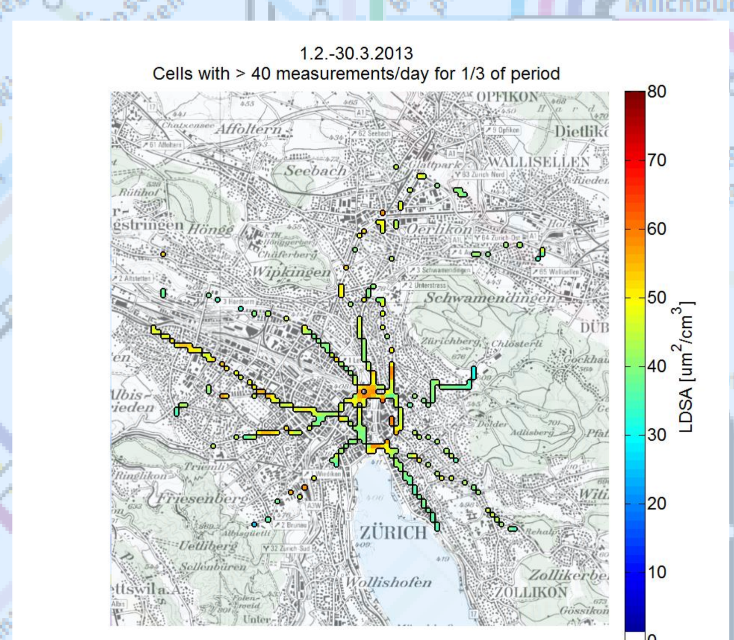


Figure 2: Cells with enough data

The remaining cells are used to create a model with land-use regression:

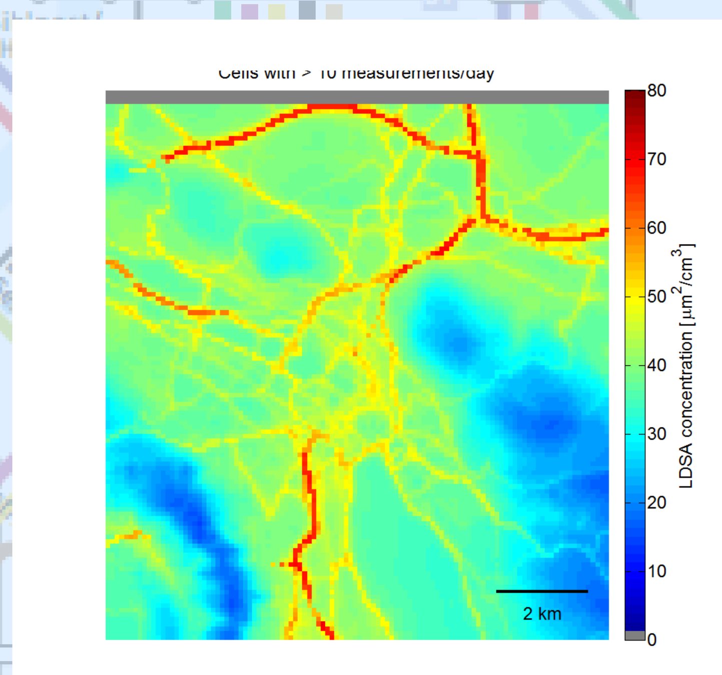


Figure 3: modelled LDSA

A further interesting result is that the interday temporal variation of UFP concentrations is much higher than the spatial variation. We define a "contrast" as the ratio of the 90th to the 10th percentile of UFP measurements. In the above period, the temporal contrast of daily averages is 2.88, whereas the spatial contrast is only 1.42. This has important implications for the analysis.

Discussion

The stochastic nature of the data collection in the mobile network leads to a very unequal distribution of the measurements over cells which makes corrections to the data necessary.
Figure 4 shows a cell where a many more data points per day are available on cleaner days. Naive averaging over all data points would give a too low estimate of UFP concentrations. By first computing daily averages, and then averaging these, one gets a 1.25 times higher UFP concentration.

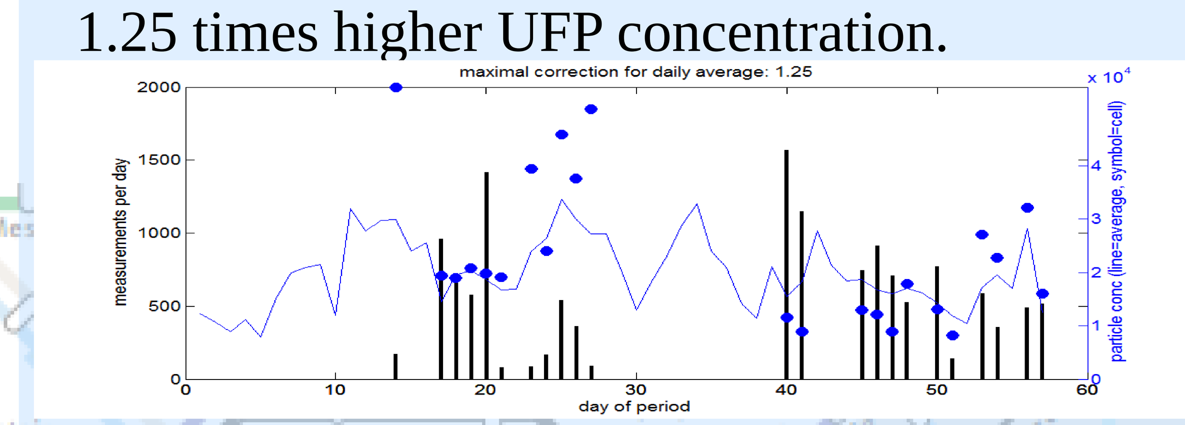
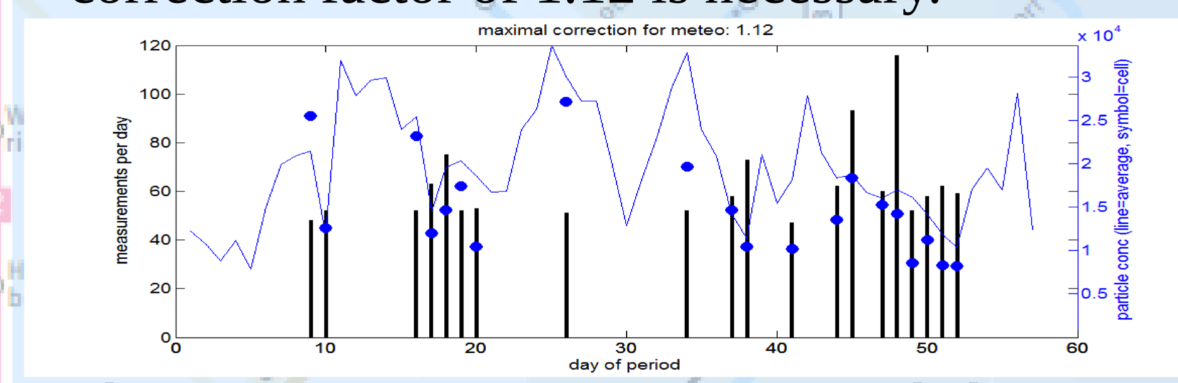


Figure 4: correction for different numbers of data points per day

Figure 5 shows a cell where more generally clean days are covered than polluted days. Again, the UFP concentration of the cell is underestimated, because of this sampling bias. This bias is corrected for by adjusting daily averages of each cell by the global daily average of all cells. In this example, a correction factor of 1.12 is necessary.



These are two extreme cases; nevertheless, the magnitude of the correction factors (1.12 and 1.25) is already very relevant compared to the overall spatial contrast of 1.42; the corrections are thus definitely necessary.

Conclusions

The concept of mobile monitoring is attractive but gives rise to problems not seen in static networks:

- Here, a too low number of sensors was deployed, leaving lots of gaps in the data. This makes complex correction algorithms necessary, which can never be perfect.
- Sensor accessibility is usually low, thus sensors need to be extremely reliable (6-12 months service intervals).
- Another issue with this specific mobile network is its limitation to streetcar tracks, i.e. to locations mostly on streets, and never in quiet neighbourhoods, parks or near highways. This limits the predictive power of the models to areas that are similar to those measured.