Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2019-410-RC1, 2020

Atmospheric Measurement Techniques
Discussions

Interactive comment on "High-resolution mapping of urban air quality with heterogeneous observations: a new methodology and its application to Amsterdam" *by* Bas Mijling

Anonymous Referee #1

Received and published: 13 January 2020

The manuscript describes an air pollution modelling and mapping framework for cities that combines two elements, a dispersion model and a measurement network. The dispersion model utilizes, among others, meteorological measurements, background pollutant concentrations from external modelling and three types of spatially distributed emission sources (highways, urban roads, residential). In principle, the framework is based on three steps that eventually lead to hourly concentration maps: First, the dispersion model is run. Second, the measurement network is used to calibrate the model, i.e. to estimate the emission factors of the three source types. Third, the measurements are assimilated to the model in order to locally improve the pollutant concentration field.

The manuscript is well written and clear. The underlying concept of the described framework convinces, especially its modularity. Many activities are going on that look for meaningful applications for low-cost sensors. Using such data for the assimilation in a dispersion model is an attractive option. However, some computation steps outlined in the manuscript are based on approximations and assumptions that are computationally fast but also simple. Also the input data used for the dispersion model are partly only proxies. The manuscript does not present in enough detail the consequences of these simplifications. I favor the publication of the manuscript in AMT after the following two main issues and the specific comments have been carefully addressed and the manuscript has been substantially revised.

1. Due to the modularity of the modelling framework it is necessary that the accuracy and limitations of the individual components are determined and described in more detail and a little bit more critical. For example, the used dispersion model is computationally fast but it seems that it has clear limitations when applied in an urban environment as e.g. building geometries are not an input. Similarly, the emissions used as input for the dispersion model are derived in a simple way that similar data is probably available also for many other cities but the accuracy of the derived emission inventory seems by far not to be optimal. I think that the gain in accuracy of a spatially variable pollutant field by assimilating measurement data strongly depends on the model's capability to resolve small scale structures. It should be made more clear in the manuscript if measurements adjust local deviations in emission source activity or only general model deficiencies. In the latter case the assimilation of measurements does not necessarily lead to throughout improved results in a local environment around the sensor.

2. The validation part should be extended. The model uses a proxy for residential emissions as input data. Residential emissions probably have a distinct seasonality. Therefore, I strongly recommend to use an additional winter period to validate the modelling framework and to analyse the resulting alpha_pop. Actually, the simulation

of a whole year would be best. No low-cost sensor data are required for this analysis.

Specific comments

- Introduction (section 1) or Setting up an urban air quality model (section 2): For completeness, a short discussion/list of other dispersion models that could be used as an alternative to AERMOD should be included.

- Traffic emissions (section 2.2.1): The author writes that traffic emissions are the dominant source of NO2 in the Amsterdam domain. Hence, traffic emissions are an important input factor for the simulation of the NO2 concentration. The interpolation of the vehicle counts for arbitrary locations based on the counting sites using IDW is practical but, at least for urban roads, possibly limitedly accurate as network characteristics are neglected. When I think of parallel roads in close distance, IDW would assign them similar vehicle counts, but in reality the true counts can be very different. An analysis/description/discussion of the accuracy of the resulting traffic input data should be added.

- Population data (section 2.2.2): The indication of the magnitude of different contributions (heating, cooking, others) to the total residential emissions in Amsterdam would be helpful. For the all-season applicability of the model: does the population database also include the spatial distribution of employees to account for heating emissions of office buildings? And for the selected period: Are heating emissions substantial in this summer period?

- Calibrating the model (section 3): First, lines 217 to 219 are unclear for me. After reading these sentences I was confused if cj(t) in Eq. 6 is the measured NOX concentration. But it is not, correct? Second, can you comment on how worse the model is performing when residential emissions are omitted? I guess that in the selected summer period heating emissions are nearly zero. Are the estimated two-hourly alpha_pop plausible and can you show that the temporal pattern of the values are related e.g. to cooking emissions? In Figure 3b, the contribution of residential emissions to the NO2

is surprisingly large given the fact that residential emissions are only 1/3 of the traffic emissions (stated in section 2).

- Assimilation of observations (section 4): The described algorithm is applied by using the pollutant concentrations transformed into the log-space. Here, one has to be aware that the distribution of the pollutant concentration at a particular location is not equal the measurement error that is required for the algorithm in this section. The measurement error is described in the manuscript as being dependent on the concentration (section 2.3). So, the reasoning of this transformation is not correct. However, I suppose that the transformation of the measurements into the log-space has a positive effect on the stability of the results as the modeling framework becomes less sensitive to (less frequent) measurements of high concentrations by reducing their impact. The transformation into the log-space is fine, but the respective paragraph should be reformulated.

- Modelling the model error covariance matrix (section 4.1): The interpolation of the model error by IDW might result in an error field that is too smooth in the urban environment given that the model is limitedly capable to represent small-scale structures (e.g. buildings). At least a comment should be included in the manuscript that points out this issue.

- Validation (section 5): The first paragraph is, as I understand, only an example where the modelling framework works well. It can be removed. Start with the second paragraph ("overall assessment"). Figure 6 can be presented directly after the overall assessment by discussing sites where the model performs well (NL49019) and where it performs less optimal (e.g. NL49002, NL49014). Figure 6 should include examples for both types. The time period of the used data in this Figure should correspond to table 3. Omit in the Figure the performance analysis of low-cost sensors but extend chapter 6. Moreover, add a file to the manuscript with supplementary materials where the scatter plots of the remaining, in the manuscript not presented air quality monitoring sites are shown analogue to Fig. 6 in order to provide the reader a clear picture of the model performance.

- Added value of low-cost sensors (section 6): The material presented in this chapter is only qualitative. The two average concentration maps presented are not validated and so the accuracy of their differences is unclear. The single example of the "Oude Schans" site is not sufficient to show that low-cost sensors add value.

I have some questions here: What is the reasoning of largely adjusting the results of the dispersion model by low-cost sensors when there is also the option to improve the input data for the dispersion model? Is it possible to generate traffic input data for the dispersion model for each month based on the traffic counts you have access to? Moreover, NO2 concentrations also depend on meteorology. What fraction of the differences in the monthly aggregated concentration fields is related to different weather conditions?

As I understand the main benefit of the low-cost sensors for the modelling framework is the increased spatial resolution of the measurements. Here, I miss some sensitivity analyses or similar material regarding measurement network design. What options exist when using this modelling framework in reducing traditional air quality monitoring sites and adding low-cost sensors? The accuracy of low-cost sensors is reported to be about 30%. Is this enough for adding substantial information?

Technical comments

- page 4, lines 102-103. How are the parallel distances of 75 and 125 m related to the grid? Maybe reformulate this sentence to make it clearer.

- page 5, lines 143-144. Refers traffic "climatology" to counting sites?

- page 7, line 197. Mijling (2018) instead of Mijling (2017)?

- page 7, lines 204-205. Pik, Pki: keep consistent.

- page 7, lines 213-214. I would not say that b(t) is observed. It is rather the output of

another modelling system.

- page 9, line 266. Section 3 instead of Section 2.3?

- page 10, line 285-286. "Isotropic" is the wrong word here as it is not isotropic.

- page 11, line 336. Change to "lower accuracy".

- Figure 1. Add units to the x and y axes.

- Figure 2. What do the depicted lines show? Sample week, yearly aggregation? Add more information.

- Figure 3a. Add north arrow and scale in one of the four Figures. The meaning of the three dots should be explained already in the caption of Figure 3a.

- Figure 3b. Indicate in the Figure in an appropriate manner the weekday the dates refer to. 2016-07-07 –> Thu, 2016-07-07. In Figure 2 the distinct traffic pattern is shown. It is interesting if there is a clear relation between traffic and NO2 in the modelling results.

- Figure 5. Add units for x and y axes and scale in all Figures. Remove the first "and" in the second sentence of the caption.

- Figure 7. Add scale in both Figures. Moreover, the visibility of the points could be better in all the presented maps.

- Figure 8. Add location of IJ-tunnel and of the historic center.

- Table 3. Indicate more precisely the date period the analysed measurements refer to.

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2019-410, 2019.