**Interactive comment on “Finding candidate locations for aerosol pollution monitoring at street level using a data-driven methodology” by V. Moosavi et al.**

Anonymous Referee #1

Received and published: 22 June 2015

This second version of the paper, "Finding candidate locations for aerosol pollution monitoring at street level using a data-driven methodology," has been edited in a minor way by the authors. However, certain fundamental, important issues have not been addressed. These include: (1) the need to apply the data driven model to a range of conditions beyond 2 hour evening limits of the initial measurements to account for longer term averages that account for a range of emissions contributions and meteorological variability. Longer term averages, of course, are the manner in which air pollution regulations for health risk are written and (2) the impact of conditions outside the local area which will affect air quality inside the area of interest. While the local land use variables for this very complex urban condition are accounted for, the model does not include any meteorological variability other than temperature and humidity. It is well known that conditions outside of a local area in a city are key to air quality conditions within that area. Local street canyon air flow variability as well as larger scale wind conditions averaged over a suitable period of time will be a strong factor in the monitoring data. The model being locally data driven may implicitly account for such influences, but this should be tested, perhaps with application of a grid-based conventional air quality model. This aspect of the modeling should be discussed qualitatively to provide for a broader critical discussion of the strengths and weaknesses of the methodology. The authors have not included a rationale for the pathway chosen for the measurements; the areal coverage of measurements is much smaller than the region modeled for relying on a data driven spatial extrapolation. The authors have not discussed the importance of integrating gas and particle data to represent air pollution conditions in the model. It is likely that measurements of different gases would yield different spatial extrapolations in the community compared with the aerosols. The results shown for the clusters in Figure 10 are puzzling. The clusters appear to cover spatial regimes that are disconnected from one another. One would expect the conditions in such a complex topographical environment to be driven by concentrations gradients that are linked with one another or to similar emission sources. The authors should discuss the cluster results in more detail—trying to interpret their physical meaning from the model results. Since the link with emissions is not really discussed, the authors could add a short qualitative narrative that would link motor vehicle or cooking or other elevated emission densities to the cluster results to make sense of them. The revised paper provides a summary of this approach to a data based model for spatial extrapolation but does not extend insight for the reader beyond the initial draft. The siting options of the hypothetical monitoring sites is interesting, but the results appear unconvincing that the siting of three added stations for aerosols will improve exposure risk estimates beyond the single government station now in operation. I recommend that the authors look more closely at the strengths and weaknesses of the method at
this stage of development and offer more insight about the methodology in the broader context of hypothetical air quality characterization specified by the model vs. current practice, and the needs for exposure assessment.