Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2019-55-RC3, 2019

Atmospheric Measurement Techniques Discussions

Interactive comment on "Gaussian Process regression model for dynamically calibrating a wireless low-cost particulate matter sensor network in Delhi" *by* Tongshu Zheng et al.

Anonymous Referee #3

Received and published: 25 June 2019

General Comments

This manuscript presents findings from a deployment of 10 low-cost PM monitors (based on the Plantower light-scattering sensor) in Delhi, a dense, polluted, urban environment – and methods used to calibrate the network leveraging 22 reference sites. The calibration method described relies on a blended approach using both kriging (Gaussian process regression) and linear regression. These types of novel approaches are of increasing importance given the emphasis on lower-cost sensing systems globally. These systems are often calibrated by collocation with reference monitors, which can be time-consuming and expensive. Findings from the study indi-

cate some success with this new calibration method, though dealing with the perceived (and likely real) heterogeneity in emissions and sources across the vast, varied Delhi landscape proved challenging.

Major Comments 1/ Data are interpolated for both monitor types (LCS and reference). Why not perform analyses to validate your interpolation? For instance, by removing data of similar size to what is missing from non-missing periods and applying the same interpolation? How much consecutive data is interpolated? An hour here or there, or larger chunks of time?

2/ Speaking of interpolation and missingness: Are data missing in any specific pattern? That is, are areas that are typically reading higher levels of pollution more likely to have missing data? Do missing data occur most often on certain days (weekdays vs weekends)? Is missingness associated with ambient temperature, time of day, etc? Are certain monitors more prone to missingness?

3/ Relatedly, QAQC procedures for reference monitors are not described. While this data can be hard to obtain from the relevant Indian agencies, it is important to more strongly highlight this as a potential shortcoming or to find out more data on how and how often reference monitors are maintained and calibrated.

4/ Is any correction – of raw signal or for temperature and/or humidity – performed by the LCS platform? Are any filters applied at the LCS station or in the cloud? Describe more fully.

5/ Can you provide and compare data from the India Meteorology Department for average temp and RH across the period you performed measurements and for the 59 days of data you used? Are they statistically distinct?

Minor Comments P1, L15 – insert comma after "sites is questionable" P1, L19 – insert comma after Delhi P1, L20 – rephrase – perhaps "available for 59 days. . ." If you elect to keep the word "valid", describe what makes the data valid P2, L15 – add "with" between

"follow-up" and "routine" P3, L18 – whole sentence is very long, but specifically, for item 3), rephrase "auto-detect the faulty and auto-correct drift nodes" to (perhaps) "auto-detect faulty and auto-correct nodes with drift" P4, L15 – replace "our" with "the" P5, L12-13 – describe the API, or remove mention of it P5, L16 – add " 's " to location P6 – describe more fully the "standardization" that occurs P8, L14 – rephrase "Spatially, the global average. . ." – is this the average across all LCS and reference monitors? Or? And if it is the average across all, then does "spatially" apply? P9, L3 – insert comma after "decent"; consider rephrasing (what does decent mean in this context?) P9, L8-13 – While I understand that GPR would have done better absent local sources, is that realistic for these types of urban environments in places like India or China? Or even in the US, in places like Queens, Oakland, or Atlanta? Isn't the spatial heterogeneity exactly why many are considering more spatially and temporally resolved monitoring networks? P10, L14 – Neither of these sites are really background sites in the way they are traditionally thought of. P12, L10 – Perhaps rephrase to "The following questions remain:" or somesuch

Figure 2 – consider different shapes and colors (in B&W, the colors are not distinguishable) Figure 3 – a more elaborate caption may help better explain the flow (for instance, a sentence for each step)

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2019-55, 2019.