
Anonymous Referee #3

Received and published: 13 January 2020

Low-cost sensors are important components in environmental Internet of Things (IoT) for high resolution monitoring pollutant distribution over an area. This manuscript presents a method for addressing the challenge of low-cost PM sensor (i.e., questionable data quality) by using machine learning (ML) technologies. This study is interdisciplinary and is novel in terms of the application of random search techniques in ML configuration, as well as the application of ML on ambient PM monitoring over multi-seasons. The results show ML has great potential for calibrating low-cost PM sensor for ambient aerosol monitoring. It would merit publication if the authors address/explain the following comments/questions below.

1. Lines 147-150, the authors stated that random search is more efficient, which is also a unique part of this study. Please explain why it is more efficient than manual or grid search in principle and if possible, give some quantitative information.

2. Figure 1, I would appreciate a geographical map showing where is the monitoring location.

3. CRAZ also monitors NOx, NMHC, ozone, and wind data, which may also influence the PM concentration. Why these data were not included in the machine learning?

4. Line 195: there is a typo: “SHAPR” should be “SHARP”.

5. SHARP was used as the reference method for PM monitoring. How often was SHARP calibrated to ensure its data quality.

6. Line 228: how the hyperparameters were determined?

7. Figure 3: will you explain what the shape of erlenmeyer flask means in the plot?

8. One aspect of the uniqueness of this study is that its study covers different seasons. I would like to see a brief discussion how season influence the results of low-cost sensors.