Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2020-413-RC1, 2020

Atmospheric Measurement Techniques Discussions

Interactive comment on "Development and Application of a United States wide correction for $PM_{2.5}$ data collected with the PurpleAir sensor" *by* Karoline K. Barkjohn et al.

Anonymous Referee #3

Received and published: 9 December 2020

The manuscript presents a development of a generalized correction equation for the PurpleAir PM2.5 sensors in the USA. The used data set includes 39 different sites in 16 states and 50 sensor units altogether. The number of recorded datapoints was 12,635 (24h-average, from 8 to 3762 per state). Only sites which had reference level data available no further than 50m away were included in the analysis.

Major comment:

The chosen analysis approach does not utilize the data as well as it could, and there is a strong case to be made that the generalized correction equation is not representative: in fact, a single generalized and true correction equation is most probably

Printer-friendly version

Discussion paper

impossible to form due to the site-specific differences in aerosol composition. Is it not obvious that two different sites entailing two completely different environments (e.g. rural background vs. urban city center) are not representative of each other? To add, the analysis does not consider possible seasonality, which may have a significant impact on aerosol composition: the shortest sensor data is composed of only 8 data points (MT, Table 1). When these issues are coupled with the notion that not all the factors affecting the sensor behavior are fully understood (line 296), generalizations should be made with extreme caution. In my opinion, the only way to address this issue is to generate correction equations for a few, most generalizable environments (e.g. urban, urban background, detached housing area, regional background) with known aerosol sources. This would also lend for a further investigation regarding the underlying reasons affecting sensor responses.

Technical comment:

In multiple linear regression independent variables should be independent. In my opinion, the fact that previous studies have misused linear regression does not warrant for a new research to continue misusing it. It reinforces the bad habit and undermines the quality and significance of the whole research line of low-cost aerosol sensors. Besides the relative humidity and temperature analysis, this applies also for the analysis regarding the binned particle counts (line 270).

Another misuse particularly characteristic for low-cost sensor studies is the use of R2 as goodness-of-fit indicator in nonlinear regression. R2 is not valid for nonlinear regression (line 295).

Recommendation: reject

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2020-413, 2020.