Response to Reviewer 1
Retrieval of Bulk Hygroscopic from PurpleAir Sensors, Psotka et al., (egusphere–2024–3618)

Major Comments

Reviewer:
This paper presents an interesting approach for extracting aerosol properties from low-cost air quality sensor data. While the method is novel and the results are promising, I would suggest a few additional analyses and clarifications in the work.

First, and most significantly, it seems that the same data are used for both calibrating the OEM method (i.e., extracting the aerosol properties) and for evaluating the performance of this method. The same holds true for the MLR method against which the OEM is being compared. I don't think that this is a fair way to do these comparisons. If possible, I would suggest separating the data from each season into distinct sets for calibration and evaluation. This separation could be done randomly or (my preference) using e.g. one month of data to run the OEM (and to calibrate the MLR) and the other months to evaluate the calibration performance. I prefer this latter approach since it is more realistic to how these sensors might be used in practice, i.e., calibrated near a reference for some period of time and then deployed to another site. Overall, I think this will be a more realistic evaluation of the method and its strengths/limitations compared to the "traditional" MLR.

Response:
We apologize for the confusion we caused with both reviewers. We should not have said we were deriving a calibration for the measurements. We did not do that. Rather, we are correcting local measurements from a single site to demonstrate the potential for our method to be applied to correct single-day measurements (e.g. on the University of Northern British Columbia Air Quality map (https://cyclone.unbc.ca/aqmap/v2/)), or in the future used to calibrate networks of PurpleAir sensors as have been done in works like Mailings et al. 2019 and Barkjohn et al. 2021. Then, what we propose above (e.g. a training set) would be the appropriate way to develop a calibration.

The procedure of using a training set for a retrieval is not typically done when solving inverse problems, as the forward model is a complete (as possible) description of the geophysical situation and the instrument function. What is problematic in OEM is how good is the model fit, that is, the value of the cost function. If the cost function is less than 1, then the model is overfitting the measurements (bad!). If the cost function is too high, say (arbitrarily) 10, the model is missing something. A perfect fit of the model to the data would be a value of 1 for the cost function. In our case, the cost values are around 5, meaning we are doing a good job of describing the measurements without overfitting them.

Reviewer:
Second, there are some comments that the method is quite sensitive to the relative humidity, and that the PurpleAir relative humidity data are of insufficient quality. Can more be said about this? I would suggest, for example, testing the method with the relative humidity as measured by PurpleAir, but using a higher relative uncertainty for these measurements in the covariance matrices within OEM. Is the OEM method still able to resolve the aerosol properties in such cases? While accurate humidity data from a nearby whether station may be available, it would be useful to also see how well the calibration can perform using only the information from the PurpleAir itself (while understanding the relatively lower quality of these data).

Response: The Jacobian of the forward model with respect to relative humidity is now presented in Figure 1 to quantitatively show that the $PM_{2.5}$ correction given by the forward model is sensitive to relative humidity. We also included more discussion in section 3.5.1 on the retrieved bulk hygroscopicity error, which we found to be less sensitive to relative humidity. A benefit of using an optimal estimation method is that the sensitivity of the retrieved parameters (here hygroscopicity) to a model parameter (here temperature, particle diameter, relative humidity) can be explicitly calculated.

Reviewer:
Third, the OEM method requires appropriate prior terms, and especially appropriate uncertainties (parameterized in covariance matrices) for these terms. While there is some discussion of the sensitivity of the method to various parameters, there is little quantitative information here which would help other researchers understand the applicability of these findings to their work. I would suggest describing how the measurement error covariance matrix and the state vector covariance matrix were defined in more detail. Furthermore, I would suggest presenting the results of any sensitivity analyses conducted for these terms, possibly in a supplement or appendix.

Response: The a priori state vector and its error are now given in Table 2. The Jacobian for each of the parameters (temperature, particle diameter, and relative humidity) are now shown in Figures 1-3. Section 3.1 describes the sensitivity analysis that we performed.

Other Comments:
I also have several specific comments and suggested corrections, listed here:

Line 20: Note that these data are still publicly available, but not freely available.

Response: The data is freely available on the PurpleAir website as shown here: https://api.purpleair.com/#api-sensors-get-sensors-data.

Line 25: Please provide a source or reference for the description of the operating principles.

Response: Ardon-Dryer et al. (2020) and Ouimette et al. (2022) were added as references describing the Plantower sensors used by the PurpleAir device.

Line 30: “was” should be “were”.

Response: Done.

Line 43: The calibration is not only applicable to “their” sensor, but also any other sensor, within the range of conditions under which the calibration was created and validated.

Response: Done.

Lines 71-72: The meaning of this sentence is unclear. I think you might intend to say that the physics-based correction model which was just described was also compared with a purely statistical correction model incorporating multiple linear regression terms. Please consider rephrasing if this is indeed what you meant.

Response: Done.

Line 87: What is meant by “clearly erroneous readings”? Is this the result of the same quality control procedure just described, or are these data erroneous for different reasons?

Response: This is separate from the quality control procedure described. This was clarified in the text in the second paragraph of Section 2.2.1. This was a one-time event likely due to the sensor being jammed with insects.

Section 2.2.3: Not a necessary change, but I would suggest moving most of this information to the introduction, as it is background detail and motivation for the study. The specific model of reference instrument used can be mentioned in Section 2.2.2 instead.

Response: Done.

Section 2.3: Values for the a-priori state vector and the error covariance matrices should be provided; this can be done in a supplement or appendix. Some argument should be provided about how these values were selected as well, so that others can make appropriate decisions when replicating this.

Response: The a priori are now listed in Table two and the sources for these values are referenced in Section 2.3.1.

Section 3.1: Are there any quantitative results which can be presented based on this analysis to support the conclusions that particle diameter and temperature are of lesser importance? For example, relative magnitudes of terms in the error covariance matrix?

Response: As described previously, the Jacobians have been added to section 3.1.

Line 176: remove “from”.

Response: Done.

Section 3.2.1: Are the same results observed for the empirical calibration, or is this over-correction unique to the physical calibration? There is a comment addressing this in lines 229-230; maybe this could be moved up and expanded on. The comment seems to suggest that, based on the data, the calibration dependence on humidity should in fact be linear, as opposed to nonlinear as in the physical calibration. This could potentially mean that humidity is impacting the sensor performance in more complex ways than just the hygroscopicity of the particles.

Response: This over-correction is more prominent in our physical calibration, but it has been observed in other studies with statistical corrections as well. Figure 5 was added to make this observation more clear. We also added a reference to a recently published paper, Mathieu-Campbell et al. (2024), describing a new way to correct high humidity measurements in Section 3.3.

Line 196: I would suggest using the measured humidity from the PurpleAir as your input and calibrating the uncertainty for this term using your comparison to a nearby weather station. This could give a sense of how robust the measurements are to the data quality of the internal humidity sensor, which you have noted is not high.

Response: We now present the Jacobian for relative humidity and expand upon how bulk hygroscopicity is affected by the error in relative humidity as previously described.

Lines 218-219: This argument seems to contradict the previous statements that the method is not sensitive to particle diameter.

Response: This argument is saying that if particles are less than 300 nm in diameter, they are not detected. This means that if there are many small particles, then we are undercounting. Once particles are above the detection limit, then we are not sensitive to particle size.

Line 221: “2.5” should be subscripted.

Response: Done.

Table 2: Why is the R-squared of the raw data not reported? Reporting the biases across the different methods could also be useful.

Response: R^2 is now reported in Table 3.

Line 254: The bias of the linear correction being zero indicates that it is being assessed on the same data on which it is calibrated. I'd suggest defining separate calibration and evaluation datasets for each month; see my general comment on this from above.

Response: See response to first major comment.