The authors thank the editor for the comments. We have modified the manuscript based on these comments to improve and clarify the text. Please find below our detailed responses in bold blue text (with direct quotes from the revised manuscript shown in "bold, quoted and italic" text). The editor's comments are shown in black unformatted text. All line numbers in our responses correspond to the "clean" version of the revised manuscript.

RESPONSE TO THE COMMENTS FROM THE REVIEWER

There are a few minor issues that need to be addressed prior to final publication.

1) One of the reviewers made a last recommendation: The MLR model parameters have a significantly high precision. It will be helpful to understand the rounding errors this model causes if the parameter precision is reduced. Will the R-square decrease with parameters that have fewer precisions?

Response: Thank you for the question. We tested the model for different levels of precision for the MLR coefficients. R^2 was robust to these changes. However, RMSE increased slightly as coefficient precision was decreased from 8 to 4 to 3 significant figures, while MAE increased with decreased precision for Model 3. Changes in these parameters were at the hundredth place. Please see Table S14 in the Supplemental Material.

"Table S14: Tests of model coefficient precision

Number of significant figures	RMSE	MAE	R^2
Model 3			
8	2.318026	1.674111	0.7717575
4	2.318178	1.674819	0.7717097
3	2.320251	1.664604	0.7724541
Model 4			
8	2.236673	1.595438	0.7871297
4	2.237311	1.597079	0.7870842
3	2.244286	1.610342	0.7865547
"

2) Please label the sites from Fig 2 in Fig 1. Or at least provide a figure/table in the SI that defines the sites and their metadata (lat, lon, elevation, etc.). Also, define PA in the caption of Fig 1, although most can figure out this is PurpleAir.

Response: Thank you for the comment. We added the corresponding longitude and latitude to each PurpleAir to Table S13 in the Supplemental Material.

3) Lines 155-157 - Which NOAA database? Add a better descriptor and perhaps a URL.

Response: Thank you for the comment. We edited the text to add the specific NOAA database. The following sentence provides the reference of the database (lines 156-158).

"We compared hourly RH from the PurpleAir with the corresponding hourly RH from the National Oceanic and Atmospheric Administration (NOAA) Integrated Surface Database (ISD). The NOAA data were downloaded using the R package worldmet (worldmet: Import Surface Meteorological Data from NOAA ISD)."

4) Section 2.4.2 - It is not clear where the other models came from. Based on a statement farther down, it appears they were developed/proposed by the authors. Please provide some more details as to how these were formulated. Also, please define all parameters that were not defined in the section above (i.e., the betas and epsilon).

Response: Thank you for the comment. We edited the text to provide more details about the models (lines 180-181) and to define the parameters (lines 187-188).

"Based on the evaluated predictors, we developed Model 1-4. The four proposed models and the Barkjohn model were structured as follows:"

"For each model, β_0 represents the intercept, β_1 - β_3 are the coefficient of the predictors C_{PA}, RH and T respectively, and ε is the error term."

5) Fig 2 - I assume the axes are 0-1 but that needs to be defined in the caption for clarity. Also, SD should be shown in panel b, perhaps as a shaded area, so the variability in the measurements is clear.

Response: Thank you for the comment. We edited the text to clarify the y-axis, and we added the SD for RH in the caption (lines 257-259).

"Figure 2: (a) Summary statistics and time series (yellow lines) of daily average RH for each PurpleAir site showing the presence of data (green) and missing data (red). The y axis represents RH scaled from zero to the maximum daily value. The percentage of data captured per year is also provided. (b) Time series of daily average RH for the entire dataset with a SD of 10.56 %."

6) Lines 285-286 - Be clear that the improved performance is due to slightly reduced errors and not necessarily better correlation coefficients.

Response: Thank you for the comment. We want to draw your attention to sentences that follow that statement. They provide detailed information about the comparison between the models in regard to their R^2 and error metrics. (lines 288-293).

"Our dataset illustrates improved predictive performance for our four MLR-fitted models compared with the Barkjohn model (Table 1). The Barkjohn model presented a higher R^2, as a measure of the goodness of fit, than Model 1, however Model 1 is improved with respect to all error metrics. The Barkjohn model resulted in a higher MAE than the four models developed

for this study. The best model fit was observed for Model 4, incorporating C$_{PA}$, T, and RH, with substantially better prediction performance metrics compared with the other models (Table 1)."

7) Lines 360-361 - What is the odd text? "Click here or tap to enter text." I assume some sort of error…?

Response: Thank you for the comment. This erroneous text has been removed.

8) Table 4. Please provide the R^2 values since they are provided elsewhere.

Response: Thank you for the comment. We edited the table to provide the R^2 values.

"

Table 4: Application of MLR- Model 4 and SSC model to individual state. The SSC combined clusters result is the result obtained after applying each cluster to the hourly data, then added together.

States	MLR				SSC combined Clusters			
	R^2 (%)	RMSE ($\mu g\ m^{-3}$)	MAE ($\mu g\ m^{-3}$)	R (%)	R^2 (%)	RMSE ($\mu g\ m^{-3}$)	MAE ($\mu g\ m^{-3}$)	R (%)
SC	56	3.41	1.92	75	57	3.40	1.87	75
NC	80	2.81	1.82	89	80	2.76	1.76	90
VA	88	2.70	2.36	94	87	2.77	2.42	93
FL	65	2.63	1.64	81	65	2.58	1.62	81
TN	75	3.11	2.21	87	75	3.10	2.19	87

"