

Comments to the Authors

General comments:

The study by Wang et al. provides a PM_{2.5} data assimilation technique in a regional air quality model (AQMv7) and apply the modelling system to September 2020 western US wildfires. The authors conducted several experiments including a control run without any assimilation and runs with AirNow and PurpleAir data assimilated to demonstrate the improved forecast performance with data assimilation. The manuscript fits in the scope of the journal. However, there are several unaddressed issues that limits the strength of the manuscript for the reader in its current form. Hence I strongly suggest that the following comments are addressed before consideration of the manuscript for publication.

Major comments:

- (1) How is this study different than previous similar studies conducted over US (for example, compared to US based studies mentioned in lines 79-80). Novelty part should be articulated clearly.
- (2) Usually for studies utilizing the air quality models, physics and chemistry options selected are usually reported. Does this model use a fixed suite of options or there are several options available for physical and chemical processes. In any case, the options used should be reported somewhere. Also, how are meteorological initial and boundary conditions prepared?
- (3) There is no clear information on emission inputs? Which anthropogenic and natural emissions are included in the model.
- (4) It would help the reader if there is a flow chart to explain the complete system used in this study.
- (5) Lines 245-246 - Why are the error standard deviations different for AirNow and PurpleAir and how the results would change if error standard deviations are kept same for both AirNow and PurpleAir (to say 10%). In other words, would it make more sense to compare DA_PA and DA_AN if error standard deviations are kept same?
- (6) Line 255-256 - Explain more clearly about these different thresholds for large PM_{2.5} observations mentioned for AirNow and PurpleAir?
- (7) Equation 9 is difficult to understand. Also, what is T in superscript?
- (8) Is there any possible explanation for the hourly changes (for example - first increase and then decrease in bias) in figure 3 and other similar figures like figure 10. Also, put axis title on the x-axis in figure 3?
- (9) Line 418-421 - Can you briefly explain about performance diagram with some references? How are these thresholds chosen?
- (10) Line 455-459 and figure 7 – I think it will be better if bias is shown for DA_AN, DA_PA and DA_ANPA with respect to AirNow (as a reference set of observations) for ease of comparison between figures 7 (b,d,f,h) for the forecast performance.

(11) Line 482-486 – Can the authors also discuss the performance EPA region-wise (as in lines 517-522) for 1h and 24h forecast. This may help to understand the performance in wildfire impacted regions more clearly.

(12) Lines 531-532 – Mention some statistics to demonstrate improvement.

Minor comments:

(1) Line 22 – particles with **aerodynamic** diameters

(2) Line 40-41 – “PM_{2.5} is a major contributor to poor air quality in US” - Provide some references to this line.

(3) Line 44 – Kindly avoid the usage of term ‘primary pollutants’ in this context.

(4) Line 76 - replacing the previous offline-coupled ~~the~~ Global Forecast System

(5) Line 77-78 – “modeling system” repeated twice.

(6) Line 88 – Put 2.5 here (and in other places in manuscript) as subscript

(7) Line 170 – Check the sentence. Should it be: *It is noted that a recipe ~~that~~ uses temperature...*

(8) Line 203 - ~~M~~mountain ~~R~~region?

(9) Line 213 – Add the last access date.

(10) Line 242 – **and** Texas.

(11) Lines 309-312 – Reduce the sentence length. Break into two sentences for ease of read.

(12) Lines 345-346 – Is it “the ratio of the particle number ~~concentration~~ to total particulate volume”? Particle number concentration is already number per unit volume.

(13) Line 422 - Figure 6 presents ~~shows~~ time series of

(14) Line 580 – September 2025 or 2020?

(15) Line 593 – There are two full stops before and after the reference.