

Review Comments for "Development of a Real-time Hazardous Air Quality Ensemble System for Improved Wildfire Air Quality Forecasting" submitted by Li et al.

General Comments:

Wildfire events present a significant challenge to air quality (AQ) predictions due to model system incompleteness and uncertainties in fire emission treatment. The manuscript presents the development of the Hazardous Air Quality Ensemble System (HAQES) and its success in improving PM_{2.5} predictions during wildfire events over the CONUS. The approach of using a weighted ensemble forecast contributes to improved forecasting accuracy, along with the application of quantile regression and weighted regression methods. The manuscript effectively compares the HAQES ensemble forecast with individual models, showcasing a substantial enhancement in forecast accuracy. The inclusion of quantitative measures, such as fractional bias, false alarm rate, and hit rate, strengthens the study's credibility. Overall, the manuscript is well-structured, providing a clear overview of the research, including the problem statement, methodology, and key findings. However, several concerns need further clarification, and the manuscript is recommended for publication in BAMS upon addressing these comments.

Specific Comments:

1. While the manuscript successfully addresses the impact of heavy wildfire events on air quality predictions, the use of 2022 as an example, being a fire-inactive year, raises questions. It might be beneficial to consider other fire-active years to better demonstrate the HAQES's advantages over individual models during intense wildfire events.
2. In the abstract, the authors assert that **weighted ensemble** reduced fractional bias by 34%, false alarm rate by 72%, and increased hit rate by 17%. However, in Section 3.2, it is noted that MLR reduces the fractional bias by 34%, increases the hit rate by 17%, and reduces the false alarm rate by 72%. It is recommended to explicitly specify the Multilinear Regression (MLR) in the abstract to avoid confusion since MLR is one of the weighted ensembles.
3. In the text, it is stated that the first 9-month simulations serve as the training data, with the subsequent three months designated for testing (see Lines 187-188). However, there appears to be a discrepancy, as Table 2 displays results for October-December, while Figures 1-3 showcase outcomes for the entire year of 2022. To avoid potential confusion for readers, please provide clarification on this inconsistency.
4. Should the results for quantile regression (QR) and weighted regression (WR) be incorporated into Figure 4 for a comprehensive visual representation, or would it be more suitable to present them in a separate figure in the appendix to ensure clarity and enhance the visualization of these specific QR and WR findings?
5. In the abstract (Line 23), the authors assert the introduction of a new real-time Hazardous Air Quality Ensemble System (HAQES) in the manuscript. However, it appears that the presented results pertain to the year 2022 rather than real-time outcomes. Please clarify how the five

operational forecasts are integrated into the HAQES system to generate ensemble forecast products in real-time.

6. Are you sure that the WRF-Chem's atmospheric aerosol chemistry is used in the GEFS-Aerosol model (Line 133)?
7. Consider using the number of Airnow observation sites rather than the number of cities for more meaningful representation (Line 176).
8. Please make sure abbreviations are consistently and appropriately defined. For example, GOCART and MMA are both defined twice, while NPP and MODIS are never defined.
9. Provide clarification on why β_o is included in Eq. (1) and what value of β_o is used in this study (Line 183).
- 10.
11. Explain the choice of $20 \mu\text{g}/\text{m}^3$ in Eq. (7) instead of the EPA standards for $\text{PM}_{2.5}$ ($35 \mu\text{g}/\text{m}^3$ for the 24-hour average and $12 \mu\text{g}/\text{m}^3$ for the annual average).
12. Please correct the inconsistencies in units in Lines 233 and 236 to be $\mu\text{g}/\text{m}^3$.
13. Please keep consistency in notation, using K_α in Lines 246-247 to match Eq. (A2).
14. Please specify the parameter presented on the Y-axis in Figure 3's caption and clarify why negative values indicate that the ensemble forecast is more skillful (Lines 279-280).
15. Line 335: $\text{PM}_{2.5}$ should be $\text{PM}_{2.5}$