

Response to Editor: We would like to thank the editor for the careful review throughout the paper that help to improve our paper.

Our Reply follows (*the editor's comments are in italics and blue*)

General Comments

Thank you for revising the manuscript. I would like to ask you to address the remaining comments of referee 2 (quoted below) and consider the following in a minor revision:

When discussing different forecast ranges, please make it clearer that the CTM results used are always the same and are based on reanalysis rather than forecasts (even if true CTM forecasts would be used in future applications). Accordingly, please avoid the term "CTM forecast" where it is not appropriate, e.g., in Figs. 4, 6, 7. For example, you could use "CTM ensemble", "CTM prediction", or "CTM result".

Also in the title I suggest replacing "chemical transport model forecasts" by "chemical transport model results", and moreover placing "air quality" more prominently:

"A gridded air quality forecast through fusing site-available machine learning predictions from RFSML v1.0 and chemical transport model results from GEOS-Chem v13.1.0 using EnKF"

Reply: Thanks for the suggestion. The term "*CTM forecast*" is changed into either "*CTM ensemble*" or "*CTM prediction*" depending on the context throughout the paper. More descriptions are now added in page 9 line 1 by saying "*Note that the CTM results utilized in this work remain consistent regardless of changes in the forecast horizon.*" All figures that originally displayed "*CTM forecast*", including Figs. 4, 6, and 7, have been updated to show either "*CTM ensemble*" or "*CTM prediction*". The title is now changed to "*A gridded air quality forecast through fusing site-available machine learning predictions from RFSML v1.0 and chemical transport model results from GEOS-Chem v13.1.0 using EnKF*".

Minor comments

Page 7, line 19: "188 stations that failed" should probably be rephrased to "the 188 stations that were skipped".

Reply: Accepted. The sentence in page 8, line 2-3 is changed to "*For this study, the 188 stations that were skipped in the RFSML model training were used for validating our fused prediction, they are referred to as validation sites B and marked as black rectangles in Fig. 1.*"

Fig. 3: Please mention in the caption that the averages in panel b correspond to the intervals shown in panel a.

Reply: Thanks for the suggestion. The remark “*The averages in panel b correspond to the intervals depicted in panel a.*” is now added to the end of the caption for *Figure 3*.

Fig. 5: The colour scale label is missing.

Reply: The missing colour scale label is added now by “*PM_{2.5} (µg/m³)*”.

Fig. 9: In the panel titles, the units ("h") of the forecast ranges are missing.

Reply: The units "h" is now added in *Fig. 9* and *Figure. S6*.

Response to Referee: We would like to thank the referee for the careful review throughout the paper that helps to improve our paper.

Our Reply follows (*the referee's comments are in italics and blue*)

General Comments

The manuscript has improved after the authors have addressed the previous comments. I still have a few specific comments and suggestions as follows:

Specific comments

1. *Page 4: line 15: top three significant features, what are they?*

Reply: The context concerning the top three significant features for the RFSML prediction was not placed in the right place. They are now added to **Section 2.3 RFSML prediction & uncertainty** in page 7 line 18-20 for better organization and coherence. **“Using an ensemble SAGE selection, we identified the top three significant features for each region, as outlined in Supplement Table S1 (Table 6; Fang et al. (2022)).”**

Table S1. Summary of selected features.

Region	NCP	PRD	SCB	YRD	FWP	REST
	PM _{2.5}	PM _{2.5}	PM _{2.5}	PM _{2.5}	PM _{2.5}	PM _{2.5}
Feature	v10	v10	d2m	v10	d2m	co
	co	pm2p5	tp	pm2p5	co	pm2p5

2. *Page 4: line 16: preceding nine hours. It is unclear what the 9 hours refer to.*

Reply: Sorry for the confusion. The choice concerning the $t_p = 9h$ in **Eq. 9** in **Section 2.3 RFSML prediction & uncertainty** was determined based on the results of auto-correlation and partial auto-correlation. Remarks are now added in page 7, line 23-25, by saying: **“where at any instant t , the input vector storing $s = 3$ individual selected features over the previous $t_p = 9h$ is utilized to forecast the target PM_{2.5} concentrations \hat{y} with a prediction horizon of h h. The choice of $t_p = 9h$ is obtained on the basis of the auto-correlation and partial auto-correlation analysis.”**

3. *It is still not clear how the RFSML prediction works. Yes, it is published in previous paper, but Section 2 should give a brief introduction how the prediction is made. I would suggest updating Section 2.3 to introduce how RFSML works.*

Reply: Thanks for the comment. We provided a more specific introduction to RFSML, with relevant comments on the page 7, line 13-19 by saying: **“In addition to the following steps in machine learning**

prediction, which include data collection of PM_{2.5} observations and datasets, interpolation of missing values in the original dataset, selection of an appropriate machine learning model, reformation of the continuous data time series into the required input structure, repeated training of the model to determine optimal hyperparameters, and making predictions using the trained model. The RFSML utilized ensemble SAGE to obtain the optimal input feature subsets. The total national air quality monitoring stations were divided into six regions. Using a computational efficient ensemble SAGE selection, we identified the top three significant features for each region, as outlined in Supplement Table S1 (Table 6; (Fang et al., 2022)).”

4. Figure 2 is still not clear to explain the EnKF framework. It is hard to follow the flow and the caption does not explain well either. It should be clearly showing what the inputs and outputs are for the system.

Reply: In Figure 2, we have incorporated the inputs and outputs to improve the clarity and understandability of the framework. Additionally, we have included a more specific description in the caption to provide additional context.

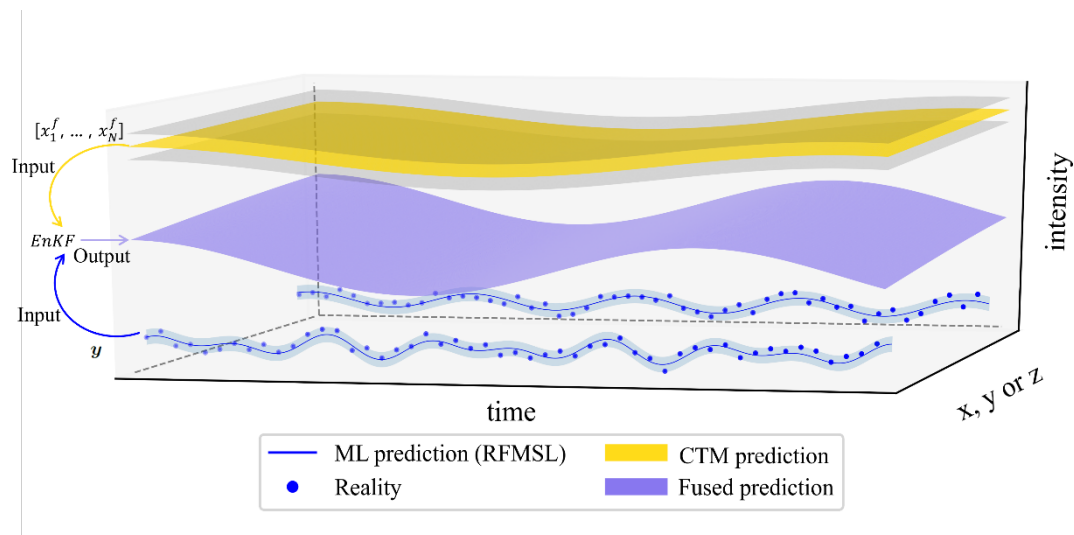


Figure 2. Framework of EnKF-based prediction fusion. The blue lines and their corresponding shaded regions represent the RFSML predictions and their uncertainty at the air quality monitoring stations, which are assumed to be very close to the actual PM_{2.5} concentration values. The golden surface and its surrounding grey surfaces represent the CTM prediction and its uncertainty. The medium slate blue surface represents the fused prediction of the RFSML and CTM prediction. y and $[x_1^f, \dots, x_N^f]$ are the inputs of EnKF, which represent RFSML prediction and ensemble CTM respectively.

5. Figure S3: what does the color bar represent?

Reply: The missing label " PM_{2.5} (µg/m³)" has now been added to the Figure S3.