General Comments

The article, "Sensitivity Studies of Four-Dimensional Local Ensemble Transform Kalman Filter Coupled With WRF-Chem Version 3.9.1 for Improving Particulate Matter Simulation Accuracy," written by Lin et al., applies the 4D-LETKF data assimilation technique to the WRF-Chem CTM model. Sensitivity experiments regarding ensemble size and assimilation time-window were conducted. The optimal ensemble size and time-window were determined for high-concentration PM cases and further validated with moderate pollution scenarios.

The authors logically designed numerical experiments to draw conclusions that support their research objectives. This study highlights the importance of surface observations in PM modeling and provides optimal ensemble sizes and time-windows through experimental results. Furthermore, it includes an in-depth discussion of ensemble spread, a critical aspect of ensemble-based data assimilation, providing valuable insights for researchers in this field.

However, some parts of the experimental methods lack clarity, and the modeling approach focuses on reproduction rather than prediction, which should be elaborated further. Additionally, specific modifications and improvements needed to enhance reader comprehension are discussed below. With sufficient revisions reflecting these suggestions, this paper is worth publishing in GMD.

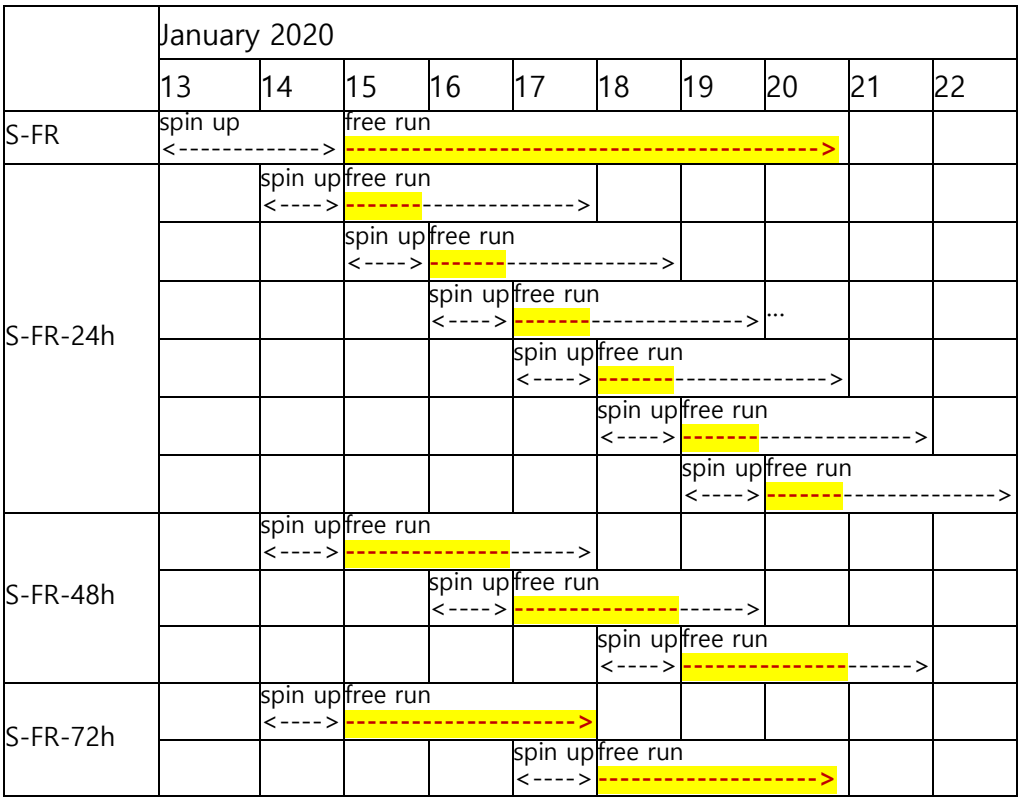
Specific Comments

1. Model Setup and Restart Timings

Has the accuracy difference in meteorological modeling used in WRF-Chem been considered for 24h, 48h, and 72h interval setups in with DA and without DA experiments?

- (Lines 273-278)
  In the free run experiments, restart intervals were set to 24h, 48h, and 72h. These experiments allowed for one day of spin-up followed by three days of modeling. For instance, in the Severe-FR-48h cycles, the first cycle had a spin-up on January 14 and free runs on January 15, 16, and 17. The second cycle spun up on January 16 and performed free runs on January 17, 18, and 19. Here, the restart point is at 00:00 on January 17. Is the spin-up on January 16 meant just for meteorological fields? A diagram illustrating the restart experiments and 4D-LETKF cycles would help clarify time configurations. Based on my understanding, the free run process

can be summarized as follows. Yellow-shaded areas denote durations analyzed in the paper:



2. Analysis vs Forecast Concept

Does the paper focus on reproduction rather than prediction? Clarify this aspect throughout the text.

- Does the 4D-LETKF approach via the equation at (line 152) include all the spatiotemporal information for $\bar{x}^a$ or $\bar{x}^b$ within the specified time-window? For example, if the 48h time-window assimilates 48 hours of 100x100x40 grid data, does it mean the results reflect only the analysis field? If forecasts are presented, the sustained impact of initial conditions should also be discussed.

3. Ensemble Spread

- (Figure 9)
  The ensemble spread represented by the standard deviation in the analysis results (third and fourth row) significantly decreases after assimilation. This is expected in ensemble-based data assimilation. I understood that the $\bar{x}^a + X^a$ is used for the next cycle. Is the inflation technique unnecessary, relying solely on emission perturbations for sufficient spread?

4. Regional Labels

- The paper mentions many Chinese regions. Figures 6, 7, and 9 zoom into BTH, but BTH labels should be added. Additionally, Figure 10 extensively discusses Shijiazhuang, which is not clearly located.

5. Terminology for Ensemble Spread

- Section 3.3 uses various terms to describe the spread, such as divergence, convergence, dispersion, and high standard deviation. Please standardize terminology to avoid confusion and replace "divergence" with "ensemble spread". From the line 560, the paper mentions "filter convergence" instead of "filter divergence", which is well known terminology. As you know, the filter divergence occurs when ensemble spread becomes too small, preventing observations from impacting the model.

Technical Corrections

1. (line 154)
   "X^b is calculated as $x^b(i) - \bar{x}^b$," instead of "X is calculated as $\mathrm{x}(i) - \bar{x}$,"
2. (line 163)
   "$y^b(i) - \bar{y}^b$," instead of "$y^{b(i)} - \bar{y}^b$,"
3. (line 165)
   "$(y^b)^T$" instead of "y^{b^T}"
4. (line 169)
   "$X^a = X^b[(k-1)\widetilde{P^a}]^{1/2} = X^a W^a$" instead of "$X_a = X_b[(k-1)\widetilde{P^a}]^{1/2} = X^a W^a$"
5. (line 170)
   "X^a" instead of "X_a"
6. (line 181)
   "according to corresponding uncertainty in MEIC inventory"
   Please clarify the uncertainty percentages in the MEIC inventory (e.g., $PM_{2.5}$, PM_{10}, BC, and OC at 80%, 50%, 100%, and 250%, respectively, when I roughly calculated based on the standard deviations in the Figure S3).
7. (line 214)
   "$y^b_{PM_{2.5}}$" instead of "$y^f_{PM_{2.5}}$"
8. (line 222)
   "$y^b_{PM_{10}}$" instead of "$y^f_{PM_{10}}$"

9. (line 227)
   "$y^b_{PM_{10-2.5}}$" instead of "$y^f_{PM_{10-2.5}}$"
10. (line 261-263), italic style
   "$\cdots \Delta l \cdots L \cdots L \cdots$" instead of "$\cdots \Delta \mathrm{l} \cdots \mathrm{L} \cdots \mathrm{L} \cdots$"
11. (line 297)
   "where $RMSE^a$ and $RMSE^f$ is" instead of "where $RMSE^f$ and $RMSE^a$ is"
12. (line 307)
   "air quality index"
   Please add references for the air quality index and explain why AQI is used instead of direct concentrations.
13. (line 311 and Figure 3(b))
   Please add labels for (a) and (b) and clarify axes titles for the right panel.
14. (line 339 and Figure 4)
   What the color bar values represent?
15. (line 410-414)
   Since FNL boundary conditions were used, please avoid referencing GFS accuracy for specific regions.
16. (line 448-450)
   Simplify "also avoid the underestimation of model spread and overconfidence in the first-guess" like "avoid underestimation of model spread, which implies overconfidence in the first guess."
17. (line 469)
   "among first guess ($x^b(i)$) and analysis field ($x^a(i)$) in terms of ensemble members" instead of "among first guess and analysis field in terms of ensemble members"
18. (line 506 and Figure 9)
   Does the red dots in the analysis fields (third and fourth row) represent the observation points? I would recommend just circles with a black solid line since the red color overlap with the color scale bar.
19. (line 554)
   "large spread" instead of "high dispersion"
20. (line 555)
   "background error variance" instead of "background covariance"
21. (line 559)
   "spread" instead of "divergence"

22. (line 560)
    “filter divergence” instead of “filter convergence”
23. (line 571)
    “enlarge ensemble spread” instead of "enlarge the deviations between ensemble members”
24. (line 575)
    “ensemble spread” instead of “ensemble dispersion”
25. (line 581)
    “enlarges ensemble spread” instead of "strengthens divergence”