Referee comment on "Implementation of aerosol data assimilation in WRFDA (V4.0.3) for WRF-Chem (V3.9.1) using the RACM/MADE-VBS scheme" by Soyoung Ha, Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2021-74-RC1, 2021

The present work of Ha demonstrates an interface between the WRF-Chem model and the WRFDA system. Said interface is used for surface assimilation of PM2.5, PM10 and four gas species (SO2, NO2, O3 and CO). A comparison against independent surface observations recorded over the Korean Peninsula indicates that 3DVAR PM2.5 forecasts produced by the WRFDA system have a lower RMSE than that of a non-DA baseline, although this does not appear to be true for all chemical species under consideration. 3DVAR increases the overall accuracy of PM2.5 forecasts (for both time series and categorized events, see Figures 12 and 13). However, the experiments do not appear to show a statistical significant improvement in the "false alarm" rates over those of the non-DA baseline in either of the two model domains (cf. right panels in Figure 13). Nevertheless, the manuscript is well written and the work of Ha should facilitate further developments on top of the existing WRFDA implementation (as described in the "Conclusions" section). I believe it warrants publication in GMD after minor revisions, as follows:

1/ Abstract: "co" -> "CO". "And the effects" -> "The effects". Also I feel like the abstract is too "optimistic" re. the improvements in forecast skill over the non-DA baseline (particularly the final sentence). This should be qualified to be consistent with the actual results presented in sections 3 - 4.

2/ l. 60 (p. 3): "readers refer to" -> "readers are referred to"

3/ l. 139 (p.5): "so2, no3, o3 and co" - > should be all capitalized (there are other instances of inconsistent capitalization throughout the manuscript)

4/ l. 231 (p. 8): define "(o-f)'s" as "observations-minus-forecasts". This should be consistent with the labels in e.g. Figure 6 ("omb", "oma", ...)

5/ Figure 8: The differences between the averaged analysis ("Mean") and the May 26 analysis is difficult to interpret, I suggest using a different layout to display the information (not a pie chart).

6/ l. 379 - 380: The "significant reduction" in false alarm rates is not evident to this
reviewer from the results presented, particularly for domain D1 (Fig. 13).