Interactive comment on "Development and application of the WRFDA-Chem 3DVAR system: aiming to improve air quality forecast and diagnose model deficiencies" *by* Wei Sun et al.

Anonymous Referee #2

This manuscript presented a data assimilation system in the WRF-Chem air quality model by assimilating the surface observations of aerosol and gas pollutants to improve the model initial conditions and subsequent 24-h forecast simulations. This study tested and evaluated the data assimilation system (WRFDA-Chem) using surface observations of six pollutants over China in January 2017. The results showed that assimilation of surface observations could largely reduce the bias and errors of the model initial conditions and improve the 24-h forecast performances. The impacts of assimilating frequency and other model uncertainties (e.g., heterogeneous chemical mechanism) on the assimilation results were also discussed in the manuscript.
This study presents a very nice attempt demonstrating how assimilating the regular

surface observations in China would improve the air quality forecast. The methodology and results sound solid, and the study is well conducted and organized. I recommend publish on ACP but after the following comments been addressed.

Specific comments:

1. Page 14, Line 267- Can you provide possible reasons why the model significantly underestimated PMcoarse? Missing emissions?

2. Page 16, Line 305:
Does Figure 4 show the vertical profiles of pollutant concentrations in the model simulations? It is not clear why the ozone levels are so low in the upper troposphere (9 km or above 15 km). Does the model account for stratospheric ozone boundary conditions? Please clarify.

3. Page 18, Line 356-360:
I suggest move the definition of the threat score (TS) from the Supplement to the main text here. Also here in the text I suggest explain what the values of TS represent.

4. Page 19, Line 379:
Need to explain here "ALL_6h" is the "ALL" simulation in Table 2, right?

5. Page 20, Line 397-406:
I do not think the discussion on ozone performance here and other places (e.g., abstract, conclusions) is convincing. The study also simulated a winter month (January 2017) when ozone photochemistry is very weak. Therefore, I do not think that the ozone photochemistry or NOxVOC ratios would explain the decreased forecast skill for ozone when increasing assimilation frequency. Since January is not an ozone pollution season, the conclusion that "assimilate O3 and NO2 every 6 h" would not be

robust only based on results of this month. Please clarify and discuss the limits.

6. Page 23, Line 460-462:
As a future development, is it possible to directly constrain the coefficients of heterogeneous reactions using the data assimilation system?

7. Page 50, Figure 12:
Need to add the unit in the figure or in the caption.

8. Page 51, Figure 13:
The titles say "Used 79" and "Used 80". What do they mean?