Review of “Impact of aerosol-radiation interaction on meteorological forecast over northern China by offline coupling the WRF-Chem simulated AOD into WRF: a case study during a heavy pollution event” by Yang et al.

General summary
This paper assesses the impact of incorporating aerosol-radiation interactions in the NWP models on surface radiation and weather forecasts during a heavy pollution episode in North China Plain. Hourly AOD fields simulated using WRF-Chem model are fed offline into the radiation schemes of a WRF based NWP system called RMAPS-RT. The inclusion of aerosols in the NWP system reduced overestimation of daytime surface radiation magnitude and budget, and improved forecasts of temperature and wind speed. The results highlight the importance of including aerosols in the NWP system and are interesting. However, the paper lacks detailed evaluation of AOD and PM2.5 (see my specific comments on improving the evaluation part). Additionally, the paper does not discuss whether or not aerosol induced changes in the weather forecast are statistically significant or not. If changes are not statically significant, it may not be worthwhile to incorporate more realistic aerosol information in the NWP models and just a climatological aerosol representation in the radiation routines may be sufficient. Thus, I recommend major revisions of the paper before publication in ACP.

Specific comments

Line 123: change “accessed” to “assessed”.

Line 195-196: why RRTMG was not used for WRF-Chem simulations. Are aerosol-radiation interactions turned off purposely in the WRF-Chem simulations?

Line 203: Why FNL data were used in WRF-Chem experiments and ECMWF data used as met IC/BC in WRF forecast? What is the sensitivity of meteorological parameters to different driving datasets?

Lines 205-206: Did you run WRF-Chem continuously for 10 days? If yes, did you use any kind of nudging to limit the drift of meteorological fields from the large-scale reanalysis fields?

Lines 213-214: I do not agree that MODIS AOD retrievals are not available during this episode. I did a quick average AOD plot in Giovanni and the resulting images are shown below in Figures R1 and R2 for both MODIS Terra and Aqua satellites. While AOD is not available everywhere in the domain but I think the datasets is still useful for validation of the model simulated spatial distribution of AOD. I encourage the authors to use Level 2 MODIS AOD retrievals for comparison with WRF-Chem.

Lines 249-250: In addition to my above comment, the authors should consider using other satellite-based products such as MISR and MAIAC AOD, and aerosol extinction coefficient retrievals from CALIPSO.
Figure R1: Time averaged MODIS Terra AOD map for 6-9 December 2015.

Figure R2: Time averaged MODIS Aqua AOD map for 6-9 December 2015.

Figure 3 and Lines 253-259: This discussion is very qualitative and I recommend the authors to include some quantitative information about the evaluation. I suggest plotting time series of
hourly averaged observed and modeled PM2.5 mass concentrations over the Henen and Hebei provinces (similar to Fig. 4 for the three cities). Maps of bias, root mean square error, and correlation coefficient for each site for the heavy pollution and cleaner periods will also be useful to understand model skill in reproducing the heavy pollution event.

Line 279: Change “were overlay” to “were overlaid”.

Figure 5: Why does the AOD peak before the reduction in SW especially on 6th June? At Taiyuan, there is not much difference between Aero and NoAero simulations which may be because AOD at this site is likely not captured well by the model.

Line 351: change “biases” to “biased”.

Line 355: change “leaded” to “led”

Line 391: change “shown” to “showed”.

Section 3.2.2 and related figures: Are the changes in different meteorological parameters statistically significant?