Reply to Anonymous Referee #3

We thank the reviewer for the careful reading of the manuscript and helpful comments. We have revised the manuscript following the suggestion, as described below.

The paper by Li et al. quantifies the contributions of residential coal combustion to air quality in BTH in January 2014 to show the importance of controlling residential sources in reducing air pollution in the region. I have following major concerns needed to be addressed before the publication of the paper.

1 Comment: The paper lacks description of inputs to the model. For example, emission information is only mentioned in Table 1 with two papers cited. The two papers actually are two different inventories, which one is used. The emissions are annual/month averages. How did the authors do speciation, temporal allocation for different sources? The resolution of the emission is 0.25deg for MIX, how do you re-allocate them to 6km resolution for the model and what are the uncertainties regarding that?

Response: We have clarified in Section 2.1 "*The WRF-Chem model adopts one grid with a horizontal resolution of 6 km centered at 39°N, 117°E, and 35 sigma vertical levels with a stretched vertical grid with spacing ranging from 30 m near the surface, to 500 m at 2.5 km and 1 km above 14 km, and the grid cells used for the domain are 150 × 150. The physical parameterizations employed in the simulation include the microphysics scheme of Hong and Lim (2006), the unified Noah Land-surface model (Chen and Dudhia, 2001), the Goddard longwave scheme (Chou and Suarez, 2001), and the Goddard shortwave scheme (Chou and Suarez, 1999). The National Centers for Environmental Prediction (NCEP) 1°×1° reanalysis data are used for the meteorological initial and boundary conditions, and the meteorological simulations are not nudged in the study. The chemical initial and boundary conditions are interpolated from the 6 h output of MOZART (Horowitz et al., 2003). The spin-up time of the WRF-Chem model is 28 h. The monthly average anthropogenic emissions with a 6 km horizontal resolution in the North China Plain are developed by Zhang et al. (2009) with the base year of 2013, including contributions from agriculture, industry, power generation, residential, and transportation sources, and the volatile organic compounds (VOCs) speciation based on the SAPRC chemical mechanism. The* *temporal allocation for different sources follows those in Zhang et al. (2009). The biogenic emissions are calculated online using the MEGAN (Model of Emissions of Gases and Aerosol from Nature) model developed by Guenther et al. (2006).*".

2 Comment: The paper lacks validation of the model performance on meteorological conditions and air pollutants. The meteorological performance was not mentioned at all. For air pollutants, we can see from Figure 2, most sites have large differences between observation and predictions, mostly under-prediction. Figure 2 actually cannot show if the model performance is acceptable. Statistics should be based on those suggested by previous studies (see refs) are needed and comparison with other studies over China is important. Refs:

EPA, U.S.: Guidance on the Use of Models and Other Analyses in Attainment Demonstrations for the 8-hour Ozone NAAQS, EPA454/R-05-002, 2005.

J.W. Boylan, A.G. Russell. PM and light extinction model performance metrics, goals, and criteria for three-dimensional air quality models Atmos Environ, 40 (2006), pp. 4946-4959.

Recommendations on statistics and benchmarks to assess photochemical model performance C Emery, Z Liu, AG Russell, MT Odman, G Yarwood, N Kumar Journal of the Air & Waste Management Association 67 (5), 582-598.

Response: We have made further validations of meteorological fields and air pollutants in Figure 2 and Table 2 as suggested and clarified in Section 3.1.1 "*Considering the key role of meteorological fields in determining the formation transformation, diffusion, transport, and removal of the air pollutants, Figure 2 presents the diurnal profiles of the observed and simulated temperature, relative humidity (RH), wind speed and direction at meteorological sites in Beijing, Tianjin, and Shijiazhuang during the simulation period. The WRF-Chem model reasonably well predicts the diurnal variations of the temperature in the three cities against observations, with IOAs of around 0.80. The model also well yields the temporal variation of the RH in Beijing when compared with observations, but it tends to underestimate the RH in Tianjin and Shijiazhuang with IOAs less than 0.70, and generally fails to capture the high RH exceeding 80%. The temporal variations of the wind speed and direction in BTH are also reasonably reproduced, but the model biases are still rather large.*".

"*Table 2 presents the further validation of WRF-Chem model simulations of air* *pollutants based on statistics methods suggested by previous studies (US EPA, 2005; Boylan and Russell, 2006; Emery et al., 2017). Compared to the suggested model performance criteria of air pollutants, the WRF-Chem model performs well in simulating the air pollutants and aerosol species in this study. The FB, FE, NMB, and NME of $PM_{2.5}$ and O_3 are generally within the benchmarks, with the correlation coefficients approaching 0.90, showing good consistency between the simulations and observations. As for the aerosol species, except for sulfate, the differences between the observed and simulated organic aerosol, nitrate, and ammonium are all less than the reference criteria. The FB and FE of sulfate are reasonable, but the NMB of 37.6% and NME of 67.8% are slightly higher than the suggested criteria.*".

3 Comment: The brute force method and the uncertainties are not mentioned. Due to the non-linear processes of atmospheric processes. Sensitivity methods such as brute force used in this study can tell the importance of the sources but have major flaws in quantifying the source contributions.

Response: We have clarified the brute force method in Section 2.1 "*The brute force method is used to quantify the contribution of the RCC emission in BTH and its surrounding areas to the air quality (Dunker et al., 1996). It is worth noting that, although the method can evaluate the importance of the certain emission source, it still has flaws in quantifying the source contribution, considering the complicated non-linear processes in the atmosphere (Zhang and Ying, 2011).*".

4 Comment: Not enough credits are given to previous source apportionment studies in China, especially the source-oriented models. See below examples:
https://doi.org/10.1016/j.envpol.2015.08.037.
Source apportionment of $PM_{2.5}$ for 25 Chinese provincial capitals and municipalities using a source-oriented Community Multiscale Air Quality model. X Qiao, Q Ying, X Li, H Zhang, J Hu, Y Tang, X Chen. Science of the Total Environment, 612, 462-471.

Response: We have clarified in Section 1 "*Using the source-oriented CMAQ model, Qiao et al. (2017) have conducted simulations to evaluate source apportionment of $PM_{2.5}$ in 25 Chinese provincial capitals and municipalities and concluded that industrial and residential sources are predicted to be the largest contributor to $PM_{2.5}$*

for all the city groups, with annual fractional contributions of 25.0%-38.6% and 9.6%-27%, respectively.".

5 Comment: The writing can be improved, please go through more times. Limited examples are given:

a. WRF-Chem is used officially.

Response: We have changed "WRF-CHEM" to "WRF-Chem" as suggested.

b. Line 18, 9 to 25 January should a persistent episode.

Response: We have changed it to "a persistent episode".

c. "the" can be better used. For example, the Beijing-Tianjin-Hebei region, the Yangtze River Delta region.

Response: We have revised the manuscript as suggested.

d. The reference method includes first name initial, not sure if ACP requires this now.

Response: We have revised the manuscript as suggested.