Response to Reviewer 2

Thank you very much for your comments on our paper. We appreciate the time that you have taken to read our manuscript and give us your comments and suggestions. Our replies to your comments are given below in blue.

Anonymous Referee #2
General comments:
This research presented the WRF-Chem model evaluation over East and South Asia. For model evaluations, the authors uses EANET observation network, Kathmandu valley site, Mt Lulin site in Taiwan. Moreover, API was used as a proxy of PM10 concentration in China. Such synergetic analysis over Asia, where the air quality is concerned issue, are needed to further promote our understanding. Before the model evaluation, the authors provided the sensitivity simulations on the basis of two different emission inventories of REAS and EDGAR. This is also the attractive point of this study. However, some discussion points would not be robust. I would like to recommend the authors to revise the manuscript.

Response: We thank reviewer 2 for the positive comments. We have revised the manuscript based on your suggestions.

1) 'South Asia' is included in the title of this manuscript, and discussed (P9386. L17-25). However, the model evaluation was done at only one site of Kathmandu valley and for only PM10. It would not be appropriate to include 'South' in title of this manuscript. Are there no observation dataset as network or literature over India or any other countries in South or Southeast Asia?

Response: In this manuscript, we are focusing on the observation dataset in China and thus we have taken out "South Asia" from the title as suggested.

2) The concluding remarks of 'The findings suggest that future model development and evaluation of emission inventories and models are needed for particulate matter and gaseous pollutants in East and South Asia' on Abstract and Conclusions section is general and ambiguous. This sentence will not provide the authors findings through this study.

Response: We have expanded on this sentence and changed the abstract and conclusion as follows:

Abstract:
Our study highlights the importance of constraining emissions at the provincial level for regional air quality modeling over East Asia. Our results suggest that future work should focus on the improvement of provincial-level emissions especially estimating primary PM, SO_2 and NO_x.

Conclusion section, paragraph 3:
Quantifying uncertainties of simulated air quality at the provincial level due to emission inputs reveals that the uncertainty in emissions inventories leads to significant differences in simulated levels of air pollutants, especially PM_{10}, SO_2 and NO_2. For O_3, on the contrary, different emissions inventories lead to only a moderate variability, showing agreement with the findings of previous studies (Ma and van Aardenne, 2004; Amnuaylojaroen et al. 2014). Our study highlights the importance of better constraining emissions at the provincial level for regional air quality modeling over East Asia, where anthropogenic emissions are high and air pollution is a major environmental and public health challenge. Model evaluation results also indicate that emissions inventories that do not consider local emissions control

policies could result in large model discrepancies. Our results suggest that future work should focus on the improvement of provincial-level emissions especially estimating primary PM, SO_2 and NO_x.

3) Discussion on the sensitivity simulation with different two emissions of REAS and EDGAR will not be robust. The authors concluded that the REAS inventory showed better performance, and the further study was conducted with REAS. However, this conclusion was simply from the sensitivity simulation on the 14-day simulation in July 2007. What about other period? At least, I recommend the authors to provide the specific reason to select this 14-day simulation to evaluate two different inventories.

Response: The purpose of the 14-day simulation in July 2007 was to show that the REAS simulation exhibited a smaller discrepancy between observed and simulated data than the EDGAR simulation did. For a different study, we conducted simulations in 2008, January and July, using REAS and EDGAR inventories. Therefore, we have compared these model simulations with 2007 observations since we do not have 2008 observational data. The comparison is shown in the following table. We find that the REAS simulation has less bias and less error than the EDGAR simulation in January and July for PM_{10} and NO_2, while EDGAR has slightly less bias and error for SO_2 only in January in terms of NMB and MFB. EDGAR and REAS simulations are similar for reproducing O_3. These two-month comparisons agree with what we find using 14-day simulation in July. This is why we argue that a 14-day simulation is able to represent other periods.

We add the following sentence in section 3.2 paragraph 2 in the revised manuscript to justify the use of 14-day simulation.

We have conducted additional sensitivity simulations using REAS and EDGAR in January and July and compared the simulated air pollutants and observation. The results of these two-month simulations (not shown here) agree with what we find here.

Table 1: Statistical measures for REAS and EDGAR simulations in January and July.

Pollutants	month	REAS					EDGAR				
		r	NMB	MFB	MFE	NMSE	*r*	NMB	MFB	MFE	NMSE
PM_{10}	Jan	0.07	-17.92	-24.12	62.61	0.81	0.02	-47.15	-59.86	77.89	1.34
	Jul	0.29	-11.64	-23.58	57.97	0.47	0.18	-47.56	-64.25	75.47	0.91
O_3	Jan	0.24	19.75	20.58	25.71	0.09	0.34	22.03	22.51	26.49	0.09
	Jul	0.43	19.21	18.20	36.50	0.20	0.41	21.85	21.53	36.91	0.21
SO_2	Jan	0.52	-18.78	-20.19	77.79	1.84	0.55	-0.35	6.34	71.52	1.51
	Jul	0.55	64.43	42.18	95.89	1.97	0.43	93.60	69.15	103.28	2.63
NO_2	Jan	0.40	-31.33	-43.21	59.63	0.56	0.31	-45.89	-58.88	69.14	0.87
	Jul	0.59	-26.18	-23.25	60.81	0.47	0.52	-52.46	-59.20	82.80	1.20

Specific comments:
4) P9377, L9-11: Does this sentence refer the finding from this study or other study? If the former, it should not be mention here. If the latter, please insert the appropriate references.

Response: This is the finding from this study. We have revised this sentence and added the following text in section 1, paragraph 3. It reads as follows:

Several emissions inventories for Asia have been developed by different groups, each with different purposes and characteristics (JRC and PBL, 2010; Kurokawa et al., 2013; Streets et al., 2003; Klimont et al., 2009). Comparison of the emissions inventories has revealed large differences in these emission

estimates. Kurokawa et al. (2013) compared different emissions inventories for several provinces in China and found that the difference in primary PM_{10} emissions can be as high as 140%.

5) P9378, L9-12: What is the necessity to use WRF-Chem model? What are the advantages?

Response: We have added the following sentences to explain the advantages of using WRF-Chem in Section 1, paragraph 5.

WRF-Chem is an online-coupled meteorology and chemistry model, simulating meteorological quantities and air pollution concentrations simultaneously and allowing two-way interactions between meteorological and chemical constituents. In regions with high PM loading, meteorology-chemistry interaction significantly improves model performance in simulating air pollutant concentrations. (Kong et al. 2015)

6) P9378, L21: Was the aerosol-radiation feedback tuned on in the WRF-Chem simulation?

Response: Yes, we include aerosol-radiative feedback in our model setup. We have included the following descriptions in Section 2.1, paragraph 1.

We include the aerosol-radiative feedback in our simulation. The rapid radiative transfer model (RRTM) scheme (Mlawer et al., 1997) is used to represent both shortwave and longwave radiation.

7) P9379, L13-21: Why the lateral boundary condition was not corresponded to analyzed period?

Response: We performed an additional simulation in July 2007 with boundary conditions obtained from MOZART-4 for year 2007. The comparisons of simulated PM_{10}, O_3, NO_2 and SO_2 using two different boundary conditions are shown in Table 2 below. We find that the boundary conditions have a small influence on simulated monthly mean O_3 and NO_2 mixing ratios. For these two pollutants, 2007 MOZART-4 boundary leads to lower MFB and NMB, while 2010 AM3 shows a stronger correlation. Based on this analysis we conclude that our model simulated pollutant concentrations are not significantly sensitive to boundary conditions.

Table 2. Statistical measures calculated from two simulations using 2010 AM3 and 2007 MOZART-4 as boundary inputs.

Pollutants	Boundary	Obs	Model	*r*	NMB	MFB	MFE	NMSE
PM_{10}	2010 AM3	82.45	80.40	0.37	-2.49	-9.86	48.13	0.36
	2007 MOZART	82.45	79.02	0.39	-4.16	-13.70	50.72	0.38
O_3	2010 AM3	35.18	43.58	0.78	23.87	25.98	30.16	0.10
	2007 MOZART	35.18	39.29	0.64	11.68	15.20	29.28	0.11
SO_2	2010 AM3	4.82	8.25	0.74	71.14	30.77	74.02	1.90
	2007 MOZART	4.82	8.74	0.73	81.35	42.71	77.11	1.99
NO_2	2010 AM3	17.54	12.17	0.73	-30.60	-54.99	62.01	0.47
	2007 MOZART	17.54	13.24	0.63	-24.50	-27.61	49.97	0.51

8) P9380, L3: Does 'PM_{10}' in emissions (shown in Table 1 and Fig. 2) include PM2.5? Also, '$PM_{2.5}$' in emissions include BC, OC, and other primary particulate matter which aerodynamic diameter less than 2.5 micrometer?

Response: Yes, PM_{10} emissions in Table 1 and Fig. 2 include all primary particulate matter with an aerodynamic diameter of less than 10μm. $PM_{2.5}$ emissions from REAS include BC, OC, and other species.

9) P9380, L17-19: So, what is the treatment of BC, OC, and PM2.5 on the simulation using EDGAR? Taken from REAS?

Response: For the simulation using EDGAR, we only read PM_{10} emissions as provided by EDGAR, which include BC, OC, and other unspeciated primary $PM_{2.5}$ and coarse PM_{10}. We did not add additional BC, OC, and $PM_{2.5}$ from other emissions inventories. We have conducted a sensitivity simulation where PM_{10} of EDGAR was split to each species as mentioned above, using weight factors calculated from REAS. The results are provided in Table 5 below. As you can see, they do not affect our simulation results and conclusions of this paper.

Table 5. Comparisons of statistic measures for simulations using EDGAR inputs with and without decoupled PM_{10}

Treatment of PM_{10}	r	NMB	MFB	MFE	NMSE
BC, OC, $PM_{2.5}$, coarse PM_{10} (sensitivity simulation)	0.15	-32.82	-45.93	64.20	0.72
Total PM_{10} (original)	0.20	-27.28	-37.34	56.70	0.58

10) P9380, L22: Please specify the temporal resolution (e.g., hourly, daily, monthly, annual) on each emission inventory used in this study.

Response: REAS emissions are monthly, while EDGAR emissions are yearly. We considered diurnal variation when we processed the emissions. We have included the description of seasonal and diurnal variation in section 2.2 paragraph 2. It reads as:

We apply the same diurnal variation to both REAS and EDGAR. REAS emissions inventory provides monthly emissions for each pollutant, while the EDGAR emissions inventory provides only yearly emissions estimates.

11) P9386, L14-17: However, better performance in weather will be the advantage of the 'online' coupled WRF-Chem model compared to other model (e.g., CMAQ).

Response: We agree with the reviewer that online coupled model is better than 'offline' models in simulating meteorology. However, as reported by many other studies, WRF-Chem still has difficulties in simulating stagnant weather conditions. We further compared the simulated and observed wind speed in four months of 2007 and found that wind speed in the winter tends to be overestimated, as shown in Table 6. This result illustrates that WRF-Chem also has difficulties simulating stagnant conditions in winter.

Table 6. Statistic measures for wind speed in different months.

Month	Obs.	Model	r	MB	ME	NMB	NMSE
Jan	1.97	3.38	0.49	1.41	1.59	71.40	0.62
Apr	2.55	3.88	0.56	1.33	1.53	52.13	0.37
Jul	2.13	3.24	0.39	1.11	1.34	51.97	0.41
Oct	2.08	3.42	0.57	1.34	1.51	64.04	0.51

12) P9387, L13-15: But Happo in Japan is also mountainous site (Table S2b). Any other reasons?

Response: The reviewer is correct that both Happo (2860m) and Lulin (1850m) are mountainous sites. However, ozone levels at these two sites are very different. The 4-month average O_3 is 60ppbv at Happo and 32 ppbv at Lulin. Our model largely overestimates O_3 at Lulin in October as shown in the table below. As suggested by the study of Ou Yang et al (2012), Lulin has more pronounced mountain valley circulation in the fall, which leads to low O_3 mixing ratios in October. Due to our model resolution, WRF-Chem is not able to capture such local meteorological characteristics. We have included Table S7 in supporting material and added the following text in section 4.2, paragraph 2. It reads as:

Statistical analysis of O_3 in different seasons at the Lulin site (Table S7) reveals that such high bias is mainly caused by overestimation in October (MFB = 63 %). A previous study by Ou Yang et al (2012) suggested that Lulin has more pronounced mountain valley circulation in fall, which leads to low observed O_3 mixing ratios in October. Our model with a horizontal resolution of 20 km x 20 km is not able to capture such local meteorology.

Table S7. Statistical measures for O_3 at the Lulin site in different months.

month	Obs.	Model	r	NMB	MFB	MFE	NMSE
Jan	32.97	48.39	0.63	46.80	37.99	37.99	0.17
Apr	50.80	53.40	0.38	5.11	4.95	17.63	0.05
Jul	21.14	35.16	0.48	66.35	49.60	49.60	0.33
Oct	23.40	44.42	0.35	89.84	62.80	62.92	0.51

Additionally, while analyzing O_3 in Lulin, we realized an error in reading the data in January at this site. We have corrected Tables 4 and S6 and Figures 7 and 8, which used O_3 data in January in Lulin. This correction does not affect our previous findings and discussion.

13) P9388, L2-4: What is the source of coal combustion?

Response: We checked the monthly emissions of SO_2 in China in January, and realized that this is not the highest emission month among the four. We have deleted the sentence about coal combustion.

14) P9390, L7: Correct the typo ‘mode’.

Response: Corrected.

15) Figures 8 and 11: X-axis have different months, so these figures have some difficulty to understand.

Response: Based on the suggestion, we have added grey lines between months in the figures. Figures 8 and 11 have been updated as shown below.

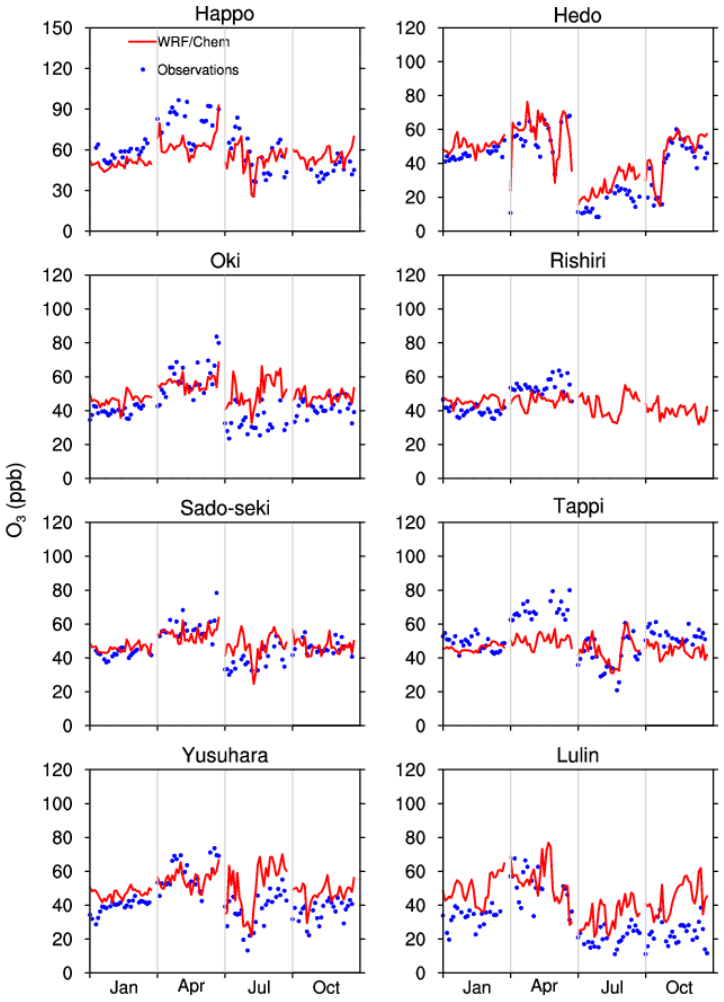

Figure 8. Comparisons of observed (blue dots) and modeled (red lines) daily mean O_3 (ppbv) at seven sites in Japan and one site in Taiwan.

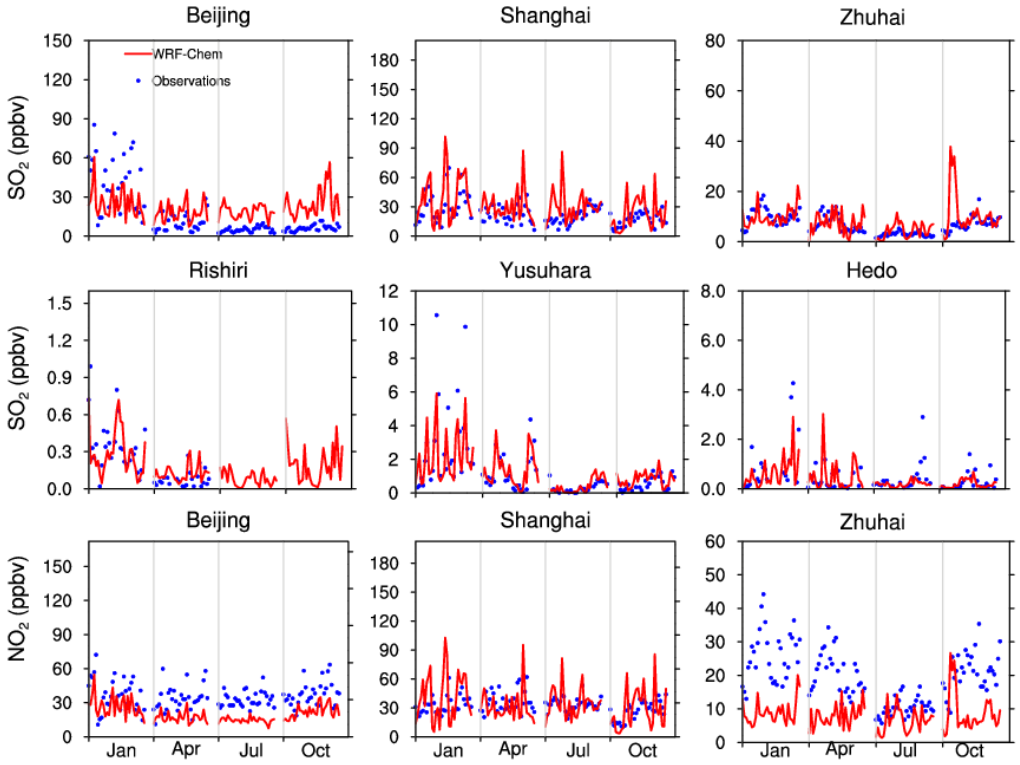

Figure 11. Comparisons of observed (blue dots) and modeled (red lines) daily mean SO_2 (ppbv) at six sites in China and Japan and NO_2 (ppbv) at three sites in China.