**Response to the Reviewer #2:**

This paper analyzed the declining trend in black carbon (BC) emissions from China, based on the long-term measurement data at a remote observation site in Japan. Combining air mass transport and air quality models, the authors made reasonable data filtering and simulation experiments. They drew a conclusion that China's BC emissions were clearly reduced in recent years, consistent with the big and continuous efforts of air pollution control by the government. In general the paper is well organized and written. Before it can be accepted for publication in Atmos Chem Phys, however, I have some concerns that should be more stressed or discussed. Some more detailed information should be provided as well, mainly in the measurement data and comparison between observation and modeling. The details follow.

We thank the reviewer very much for reading our paper carefully and giving us valuable comments. Detailed responses to the comments are given below.

1. The paragraph in Pages 3-4. The authors described the method of unifying the observations from COSMOS and MAAP, and stressed that the datasets correlate each other well. I suggest they provide the detailed correlation analysis between the two datasets with a figure, and indicated quantitatively the gaps between the two.

In the revised supplementary material, time series plots during the whole period (2009–2019) and a correlation plot between the two data sets (i.e., COSMOS and MAAP) will be provided as Figures S1 and S2. In relation, the statements on the uncertainty estimation for the BC observations will be revised:
" From the average and a standard deviation of the monthly MAAP/COSMOS ratios, systematic and random uncertainties were estimated as ± 14% and ± 17%, respectively (± 22% in total)." (Page 4, Lines 8-9 in the track change document)

2. The third paragraph in Page 4. Does the author mean that the determination of influencing regions depend on the air mass transport modeling (HYSPLITT)? If so, I would suggest the authors provide the time-series or temporal variation of influencing regions within the research period, at least in the supplement. Some more discussions should also be given on the information.

Yes, we meant that HYSPLIT-based backward trajectories were used for the judgement of the influencing regions. In Fig. S1 of the revised manuscript, the assignment of the influencing regions during the whole study period will be shown. Figure S5a (formerly Figure S2a) showed statistics and trends of number of hourly cases of observed air masses from different origin areas. In text, we mentioned as follows (Page 7, Lines 17-18):

Changes in large-scale flow patterns could also be a potential contributor to this trend in BC observations. However, this is unlikely, as the frequencies for various air-mass origin areas were almost unchanged during the study period (Fig. S5).

3. Lines 11-16, Page 5. I cannot quite agree with the authors that the gap between observation and modeling indicates only the emission change, without the full evaluation of the model performance on 2008 (for which the emission data were applied). The determination of E(y)/REAS2.1(2008) thus seems problematic. How did the authors evaluate the model and recognize the modeling uncertainty for BC besides the wet deposition?
5. Relevant with Question 3, I feel the authors need first to evaluate the model performance based on the observation, emission data and meteorology for the same year. The deviation between simulation and observation should be carefully studied to understand the uncertainty of modeling. Such bias should be excluded in the following step of determination of E(y)/REAS2.1(2008).

First, it should be noted that REAS2.1(2008) emission flux values were used as working references, and they were not necessarily true values for 2008 (Page 6, Lines 5-6 in the revised manuscript, track-change version); the true emission flux for 2008 is very likely different from REAS2.1(2008) and therefore difficult to perform the "evaluation" that the reviewer suggested. On the other hand, the superior performance of the WRF/CMAQ model incorporating REAS2.1(2008) emission in simulating peaks and relative variations is demonstrated in Fig. 4 (for spring 2018). Similarly good performance will be shown for the other periods (Fig. S1 of the revised supplementary material). Therefore, except for the systematic and random uncertainties that we consider for the observations and model simulations, there will be no other missing factors which could rival to the emission rates.

We agree that the uncertainties of observations and models are important and will clarify this point in the revised manuscript:

"Therefore the only factor that the model failed to replicate the observation, <u>except for their uncertainties</u>, was the emission trend." (Page 5, Lines 33-34)

And the methodologies and results of the estimation of systematic and random uncertainties in the observations and models will be clarified:
" From the average and a standard deviation of the monthly MAAP/COSMOS ratios, systematic and random uncertainties were estimated as ± 14% and ± 17%, respectively (± 22% in total)." (Page 4, Lines 8-9)
" We estimated the model uncertainties from meteorology to be ~± 16%, in simulating surface BC concentrations under conditions with negligible wet deposition, considering both horizontal and vertical inhomogeneities in the model and the spread of multi model simulations of CO over the East China Sea (Kong et al., 2019). The uncertainty was assumed to be contributed equally from systematic and random terms (± 12 %)." (Page 5, Lines 25-28)

Then, the uncertainties in the absolute emission correction factors (E(y)/REAS2.1(2008)) and their trends will be clearly mentioned as follows:

"Overall uncertainty in the estimated E(y)/REAS2.1(2008) values was estimated to be ± 27%, including those random and systematic from both model and observation (see Sect. 2). On the other hand, the uncertainty in its trend was estimated to be ± 21%, as influenced only by random uncertainties." (Page 9, Lines 9-11)

In Fig. 7a, the uncertainty range (a band with pale red color) will be expanded accordingly, to cover the overall uncertainty.

4. Figures 3 and 4. Why stress spring and select spring 2018 for comparing the modeling and observation results? Any special reasons?

We just selected this period as an example and did not have any particular intention with the selection. We will provide time series plots during the whole period (2009–2019) as Fig. S1 in the revised supplementary material.

6. Small issue. What are the meanings of the dots with two colors in Figure 10b?

In legend of the figure we will show that they are yearly and 3-y running mean values in the revised manuscript.

We again thank the reviewer for the important suggestions.