Authors' replies on Referee#1's Comments on ACP 2024-2064

"Downward and upward revisions of Chinese emissions of black carbon and CO in bottom-up inventories are still required: an integrated analysis of WRF/CMAQ model and EMeRGe observations in East Asia in spring 2018"

General Comments

AC: We thank for your time and efforts in reviewing our manuscript. We are grateful for your general comments on recommending it to be published after revisions. We also acknowledge the benefit of writing- and presentation-based revisions for our manuscript. We have thoroughly revised the manuscript, noting the changes and indicating the line numbers in the author comments for each referee comment. Minor textual adjustments and reconstructed text are not highlighted.

High-level comments

RC: 1. Consider reorganizing the manuscript for clearer flow. Current flow does not read well and may be improved by aggregating analyses based on a region, adding an Overview in the Methods, and then a bulleted list of analyses performed. Similar changes in the Results section could be useful. Also considering renaming section titles to be clearer and consistent.

AC: We now have the overview or summary in the Methods **(L100-104)** and Results (**L284-287**) by moving the relevant text to this section, and the Conclusion **(L644-646)**.

We substantially reorganized the manuscript by aggregating analyses based on regions. Section 3 (Results) now includes 2 subsections: 3.1 for three polluted cases (THL, PHL, JPN) including comparisons of concentrations and residual ratios; 3.2 for CHN case including comparisons of concentrations, residual ratios, and emission estimates, with a separate part for uncertainty assessments. **Section titles** were also renamed.

RC: 2. The manuscript needs more focus on the robustness of the results when compared to aircraft observations. Do a few days of aircraft observations represent seasonal or annual magnitudes?

AC **(L461-464, 606-609)**: The analyzed plume originating from a major pollution center in China was one of the most significant air pollution transport events recorded during EMeRGe and by our long-term ground-based observations at Fukue Island in western Japan over the past 10 years. The event was regarded as representative with respect to its transportation route and negligible deposition loss (**L461-464**). The annual and nationwide emission estimates appear to be

extrapolations based on limited aircraft observations. However, this extrapolation is supported by our previous analysis, which used long-term ground-based observations (Kanaya et al., 2020). This point has been clarified in the revised text (**L606-609**). Note that the uncertainty associated with representativeness has been discussed in the original manuscript ("We estimated (3) representation errors…", L582-585).

RC: 2. Or could the uncertainty in the modeled concentrations be more due to the simulated meteorology and aerosol representation alone and not due to the underlying inventory?

AC **(L588-590)**: The difference between the model and observations was investigated using another model (CAMx), which employed different meteorology and aerosol schemes but used emission inventory common to the CMAQ-based simulations. The analysis with the CAMx resulted in a correction factor of 0.54, similar to 0.48 from CMAQ. This suggests that the discrepancy is primarily due to the underlying emission inventory, and not from the meteorology or aerosol scheme. We added some more explanation in the part for model-specific uncertainty assessment (**Sect. 3.2.4**).

RC: 3. While a model-observation study shows potential bias in the inventory, it does not separate those biases by source. Which sources would inventory developers need to tune up or down to match these observations? Can aircraft observations aid in identifying hotspots from a region?

AC (L606, 523-526): It is difficult to point to sectors from our observation only. Nonetheless, the analysis of the BC/CO enhancement ratio implies that the BC and CO emissions in the model need to be reduced and increased, respectively, providing hints that sectors with large contributions or very high BC/CO emission ratio have to be tuned up. They included coal-fuel stoves. A thorough review of individual sources also helps assessing missing updates that should reflect policy changes or new experimental data. We added discussions in **Sect. 3.2.2 and 3.2.4**, supported by our previous analysis based on ground-based observations (Kanaya et al., 2020, 2021).

RC: 4. The model-observation mismatch is attributed to uncertain emissions in the inventory. However, as the HYSPLIT back-trajectories show, the source regions could be a narrow band or can cover multiple regions. How robust are this study's findings in generalising over the entire CHN or other regions?

AC (**L606-610**): We agree that the footprint from two flights covers the key emission regions but

is not very wide, even if it is simulated taking atmospheric diffusion into account across the central transport axis, which was not included in the original HYSPLIT trajectories. As stated earlier, the annual-based and country-wide emission estimation is an apparent extrapolation from the analysis of the limited number of events, but the idea has been supported by our previous analysis based on the long-term ground-based observations (Kanaya et al., 2020). These points have been clarified in the revised text. Note that the associated uncertainty arising from representativeness has been discussed in the original manuscript (L582-585).

RC: 5. Consider consolidating Figures 2-5 into one figure and focus on showing how the model performs in each of the cases. For example, a figure could focus on just the identification of case names based on observed back-trajectories, and another figure could focus on the flight observations and modeled concentrations. Show the relevant flight paths in the figures with back trajectories.

AC: Thank you for the suggestions. We have updated the **figures** and all figure cross-references accordingly.

RC: 6. Aircraft observations have their own merits and demerits. For example, while they aid in isolating non-local sources, comparing them with modeled concentrations is difficult due to higher model uncertainties at those distances from the source region. At aircraft distances, model uncertainties may be higher than emission inventory uncertainties, for example, due to modeled rapid deposition offsetting higher emissions. How robust are aircraft observations at isolating uncertainties in emission inventories?

AC (**L58-59, 206-218, 588-590**): Model uncertainties in the aircraft observation range would be indeed non-negligible. Nonetheless, we demonstrate that the transport/meteorology error is small, by showing that a similar overestimation, occurred with a different model system (CAMx), using the same emission inventory (**L58-59, 588-590**). Note that we have attributed the transport error to be 26% in the original manuscript (L585-588).

The BC/CO emission ratio is free from such transport error and provides an unequivocal indication that the emission used must not be wrong. Consistency with the ratio from ground-based data on Fukue Island ($\approx$ 3.5 ng m^{-3}/ppb) corroborates our conclusion (mentioned in the original manuscript at L517-519).

Deposition during transport is excluded from the captured air mass, given that the mean APT 72 h is only 0.1 $\pm$ 0.3 mm, indicating negligible precipitation influence. Previous studies have shown that a low APT 72 h value (e.g., < 1 mm) for the air mass with typical 40 hours transport time

indicates negligible wet deposition, which does not significantly influence emission-specific characterization (Kanaya et al., 2016, 2020) (**L206-218**). In our study, the transport time of the CHN air masses to the aircraft was ~40 – 120 hours (**L201**).

We have added some text for clarification in **Introduction** part (**L58-59**) together with explanatory text to Section 2.3 and 3.2.4, addressing several related comments.

Specific Comments

RC: Title: It can be made tighter; something along the lines of "Assessing uncertainty in emission estimates from China using EMeRGe aircraft observations and models".

AC: We made change to the **title** as "Assessing BC and CO Emissions from China Using EMeRGe Aircraft Observations and WRF/CMAQ Modeling"

RC: L22: See point 6 in high-level comments. Justify this in the Introduction.

AC (**L58-59, 206-218, 588-590**): We added the discussion to the Introduction and relevant parts (Sect. 2.3 and 3.2.4).

RC: L30: "The results suggested that downward and upward revisions of Chinese emissions of BC (−50%) and CO (+20%), respectively, are required in HTAPv2.2z emission inventory." Also mention the range in other inventories such as CEDS (Hoesly et al., 2018) or the IIASA GAINS.

AC (**L30-34**): We added in the **abstract**.

RC: L34: Mention the lifetimes of SLCFs.

AC (**L36-37**): We added the lifetimes of SLCFs as requested in the **introduction**.

RC: L50: "Our understanding of the responses of SLCF emissions to the establishment of techniques that decrease emissions in the last two decades in fast-growing Asian economies is insufficient (Chen and Chen, 2019; Kanaya et al., 2020; Ikeda et al., 2023; Zhang et al., 2022)." This point was not clear. Emissions decrease as we use more abatement (assuming activity remains the same), and the same is true everywhere.

AC (**L52-53**): It was changed to "The changes in **national** SLCF emissions due **to the potentially uneven adoption of emission reduction techniques** across fast-growing Asian economies over

the past two decades are not well understood (Chen and Chen, 2019; Kanaya et al., 2020; Ikeda et al., 2023; Zhang et al., 2022)."

RC: L52: "Biomass-burning habits in Southeast Asia has become the main contributor of carbon emissions from forest fires in spring (Reid et al., 2013; Heald et al.,2003; Palmer et al., 2006; Johnston et al., 2012)." Consider rewording this to not write it as a 'habit'. People use biomass as a source of residential or heating energy in those regions and while it is a habit, it is more of a need.

AC (**L56):** It was changed to "Moreover, biomass-burning in Southeast Asia has become the primary source of carbon emissions from forest fires in spring (Reid et al., 2013; Heald et al., 2003; Palmer et al., 2006; Johnston et al., 2012), **driven by the regional reliance on biomass for residential and agricultural purposes**."

RC: L80: Good point. Agree!

AC: We appreciate your support on this point.

RC: L83: "Emission inventories therefore need to be tested using independent observational data." I understand this focuses on the EMeRGe dataset, but any mention of ground-based long-term monitoring and its usefulness would also be good to include in the Introduction, and then mentioning how aircraft observations can help provide another perspective will be great to add.

AC (**L58-59, 206-218, 588-590**): We believe the addition in response to the comment on L22 adequately addresses this point.

RC: L132-137: I appreciate the inclusion of underlying inventories in HTAP.

AC (**L155-158**): We appreciate your support on this point.

RC: L164: Why is 1200hrs (or 5 days) a suitable time range for back-trajectory analyses in this work?

AC (**L195-200):** A 120-hour (5-day) period was selected for the back-trajectory analyses in long-range transport studies, to account for the time it takes for air masses to travel from the time of emissions to the time of detection by the aircraft (traveling time for the studied events was

estimated to be 40–120 hours for the CHN case). This approach is consistent with previous studies (Choi et al., 2020; Kanaya et al., 2016; Kanaya et al., 2020; Miyakawa et al., 2017; Zhu et al., 2019).

RC: L189: Describing the common BC/CO, CO/CO2, and BC/CO2 ratios in different combustion sources would aid the reader in understanding that whenever there is a higher BC/CO ratio, it signals a source X, and a lower CO/CO2 ratio signals source Y.

AC (**L234-239)**: We added the description.

RC: L215: Shouldn't this correction factor be applied to only those grids in CHN where the emissions are coming from and were measured in the campaign? Or to the whole of CHN?

AC **(L268-270):** Considering that emission structures are almost uniform over the country, as evidenced from our longer-termed ground-based observations receiving various air masses, we concluded that it is reasonable to apply the correction factor to the entire region of China.

RC: Equations 4, 5, and 6: Should this correction factor based on BC/CO and other ratios be applied to just the emission source with a known BC/CO (and other ratios)? For example, if the aircraft observed a high BC/CO ratio, it signals those emissions from source X dominated the concentrations on that day. Now, the modeled concentrations show a lower BC/CO ratio, indicating either the contribution from source X in the inventory is low, or that the contribution from other sources (with a lower BC/CO ratio at emissions) is higher. How do you differentiate between these two offsetting effects?

AC **(L277-282):** Differentiating between these two offsetting effects is difficult because Chinese emissions are well-mixed before reaching the aircraft, as sector-specific sources exhibit partial spatial and temporal overlap in emission inventories. To estimate total national emissions, we applied averaged ratios and correction factors from captured air masses and the model-prescribed national emission amounts, ensuring a more representative result by mitigating local transport and source biases. To minimize BC/CO ratio discrepancies in aircraft measurements, we used the spring-averaged value from Fukue Island. Further explanation is given in the end of **Sect. 2.5**.

RC: L225: This paragraph should be in Methods.

AC **(L206-218):** We moved a part of it (introduce the precipitation and APT properties of the

selected cases) to the end of Sect. 2.3.

RC: L260: "On the other hand, the IFS-CAMS simulation predicted the maximum CO concentration well in Deroubaix et al. (2024a), possibly because anthropogenic emissions in the IFS-CAMS simulation were taken from the CAMS-GLOB-ANTv4.2 emission inventory (Granier et al., 2019)." This brings up a good point and something that should be discussed in the end and generally --- how does the choice of inventory affect your findings?

AC (**L408-409 and 610-612):** The regional emission distribution pattern is largely consistent across inventories, with a nearly linear concentration-emission response (Ikeda et al., 2022). Therefore, the choice of emission inventory does not impact the findings of this study (**Sect. 3.2.4, L610-612**). However, selecting representative air masses is key to accurately analyzing national emission ratios, while examining co-emitted species concentrations and meteorological parameters can inform whether an air mass is sufficiently indicative of the region (**Sect. 3.1.3, L408-409**).

We added relevant discussions to the final part of the Chinese emission uncertainty (**Sect. 3.2.4**) and the JPN emission ratios (**Sect. 3.1.3**).

RC: L264: A general comment --- what does the observed/modeled ratio refer to? Is it the average observations and average modeled concentrations for the whole flight path? If so, the values closer to the source region better represent emissions uncertainty due to lesser influence by transport processes such as coagulation and deposition. Consider adding any discussion on that as well.

AC (**L151-152, 254-256**): the observed/modeled ratios were calculated using Eqs. 1 and 2. Values closer to the source region represent best emissions uncertainty due to minimal transport influence and deposition. Values far from source are affected by deposition. While coagulation alters particle size distribution, it does not remove BC mass from the air mass (**L151-152**). Wet and dry deposition for this range of transport speeds have been analyzed and found to be negligible in the case where the APT in 72 hours traveled from the source is smaller than 1 mm (Kanaya et al., 2016, 2020). Thus, we assumed linear responses between emissions and concentrations for both BC and ΔCO, which has been verified by Ikeda et al. (2022). This is discussed in **Sect. 2.5**, including an additional explanation in **L254-256** and a small mention about coagulation in **Sect. 2.2 (L151-152)**.

RC: L282: This raises a good point about the temporal resolution in the inventory. Inventories probably do not capture diurnal or hourly emission patterns and thus the bias in aircraft

observations at two time-stamps against modeled concentrations may be due to the lack of such temporal resolution in inventories. Consider mentioning that in perspective of the model-observation differences in this and other sections or in general.

AC (**L340, 366-367**): Emission inventories capture typical hourly and diurnal emissions but not irregular events. Japan's JEI-DB includes diurnal anthropogenic emissions, allowing the 3–6 UTC data to reflect small enhancements from JPN emissions. Thus, observed BC concentrations fluctuated without a clear CO increase, likely due to mixed sources and low enhancement levels rather than the inventory temporal resolution. In contrast, CHN and PHL inventories use monthly emission data, where limited temporal resolution may contribute to discrepancies with observational concentrations (**L340**). However, concentration ratio gaps primarily stem from emission factors or pollutant activity levels, with certain sectoral activity levels requiring more detailed characterisation (**L366-367**). Note that the observed air masses have been influenced from emissions integrated over time along the transport, which must be longer than the observation duration, particularly for the CHN case. Thus the observed concentration differences in the 2–3 hours (from different locations) would be more reasonably attributed to the spatial inhomogeneity (due to temporarily averaged emissions and transport) than to the diurnal variation of emissions itself. We added some short explanation to the PHL case Sect. 3.1.2 (**L340, 366-367**).

RC: Table 2: How is the R calculated here? Why is it so low in some cases? Consider adding Spearman's coefficient since the model and observed trends were similar.

AC: In the THL and JPN cases, the Pearson correlation coefficient (R) is relatively low due to temporal discrepancies between the modeled and observed peak times, as well as the influence of mixed sources. We examined the Spearman coefficients, but they showed only minimal improvement over the Pearson correlation coefficients, so we opted not to include them.

RC: L339: Even if there is no rainfall, higher moisture can lead to rapid aerosol growth from condensation and thus a faster deposition, especially farther from source regions when BC gets activated due to sulfate reactions.

AC (**L207-215, 306-311, 322, 461-464)**: Excluding data based on stricter rainfall/moisture criteria had to be avoided unless necessary, as the available data for each case is already limited. A detailed analysis of Accumulated Precipitation along Trajectories (APT) was conducted, applying APT criterion for the air masses from CHN, THL (APT 72 h < 1 mm), PHL E-AS-06 and E-AS-10 (APT 24 h = 0 mm), PHL E-AS-03 S2 (APT 10 h = 0 mm), JPN (APT 36 h = 0 mm) based on

their traveling time to the aircraft (Table S6). The APT 72 h < 1 mm criterion has been previously analyzed and showed no large influence of wet deposition on the emission-concentration relationship (Kanaya et al., 2016, 2020), whereas the APT = 0 mm criterion aims to completely exclude the effect of rain. Compared to the full dataset, applying APT criterion reduced data across all cases except PHL (E-AS-03 S2) and JPN. In CHN and PHL (E-AS-06), statistical metrics showed only minor changes. In PHL, metrics for E-AS-10 changed due to substantial data reduction and biogenic flux complexity, though the residual ratios remained largely stable (**Sect. 2.3, L211-214**).

In THL, R values increased and observational concentrations decreased (**Sect. 2.3, L214**). The observed BC/CO ratio has decreased, bringing it closer to the modeled value, with regression slopes of 5.21 ± 0.68 ng m^{-3} ppb^{-1} (observation) and 5.37 ± 0.07 ng m^{-3} ppb^{-1} (model). While observed BC, CO, and CO_2 concentrations declined, simulated values remained unchanged, suggesting that the air mass selection, not the APT criterion, is responsible. The exclusion of long-range transport from Myanmar, northern Laos, Vietnam, and southern China over the East Ocean, while retaining local air masses from Cambodia, Thailand, and southern Laos, may enhance the model's representation of regional air masses (Figure S5e, f). CO/CO_2 and BC/CO_2 ratios dropped to 24 ppb ppm^{-1} and 112 ng m^{-3} ppm^{-1}, respectively, though correlation coefficients worsened. However, these ratios remain consistent with reference values for dominant fire types in THL (**Sect. 3.1.1, L.306-311, 322**)

For CHN case, the HYSPLIT back trajectory model indicates that specific humidity and relative humidity remained low throughout the three-day transport period. The mean specific humidity was 3.6±1.1 g/kg, while relative humidity averaged 57±13%. These values suggest that high moisture levels or cloud formation were unlikely to influence the airmass during transport, under the high-pressure system (**Sect. 3.2.1, L461-464**).

We briefly added some relevant discussion in Method Sect. 2.3 (**L207-215)**, Sect. 3.1.1 for THL (**L. 306-311, 322**), Sect. 3.2.1 for CHN (**L461-464**).

RC: L377: Similar to the comment above on Equations 4, 5, and 6 --- should not this correction be based on source-specific BC/CO ratios? This can also explain some of the differences observed in Figures 7a and 7b.

AC **(L277-278)**: We explained above that this study aims to estimate national total emissions to facilitate comparisons with other references. Chinese emissions are well-mixed before reaching the aircraft or Fukue Island, as sector-specific sources (e.g., power plant, transport, domestic, industries) exhibit largely spatial and temporal overlap in emission inventories. So, the explanation in Sect. 2.5 still applies for this issue.

RC: L460-465: While this is a great finding, consider including some process-specific discussion --- which specific source needs to be tuned up or down in inventories? Do the observed BC/CO/CO2 ratios help in differentiating source signals?

AC **(L523-526)**: The ratios provide limited guidance on source sector revisions due to complexities, including errors from smaller footprints and highly variable emission factors. Since this study focuses on national total emissions, it highlights the national emission gaps rather than sector-specific gaps. In CHN, BC emissions, which require reduction, are revised by identifying the dominant sources based on magnitude and activity levels. For CO and CO_2, whose model gaps are harder to identify the source attributions, we consider that an approach involving updating the emission factors in sectors where BC needs revision. Additionally, results for the CHN case point to missing CO_2 sources beyond those shared with BC. This point was added to Sect. 3.2.2.

RC: Figure 6: Any explanation on why the model does not produce as much variability (concentration range) as observations will be good to add.

AC **(L338-341, 425-428)**: The model exhibits much less variability than the observations in PHL, JPN (BC), and CHN (CO) cases, which could be due to several factors. The most evident reason is the coarse resolution as the model operates at ~0.5°, whereas the aircraft observations have a much finer footprint (~0.01° per 15-second time step). Additionally, input emission inventories may lack detailed temporal variations, e.g. HTAPv2.2z in CHN case and REASv2.1 in PHL case, leading to differences in concentration ranges. Furthermore, atmospheric dispersion and mixing processes might be smoothed in the model, reducing sharp concentration contrasts observed in real measurements, especially when measurement locations are near emission hotspots or areas with stronger variability, as in the PHL case. Errors in modeled meteorological factors also impact pollutant accumulation and dispersion, contributing to variability gaps in JPN, where air masses are mixed and not solely from JPN sources.
We included a detailed explanation for the PHL case (**L338-341**), as it shows the most significant difference between the model and observed concentration range, and in the caption of Figure 4 (**L425-428**).

RC: Figure 8: How did you calculate the uncertainty in red and green boxes?

AC **(L632)**: As discussed in Sect. 3.2.4 (Sect. 3.1.5 in the original manuscript), the uncertainties of the emission estimates were propagated from (1) the uncertainties in the data of observation-

to-model ratios and emission ratios for multiple species, (2) the systematic errors of the instrument, (3) representation errors caused by the limited opportunities for aircraft observations to be made in terms of seasonal and spatial variabilities by comparing the spring and annual data, CEC region vs. the large footprint found at Fukue Island, and (4) an error for model-specific transport determined by comparing the BC concentrations simulated by the CMAQ model with the CAMx model system using the same emission data. These points have been included in the original manuscript (L580-588). We added a citation to **Sect. 3.2.4** in the caption of **Figure 8 (L632)**.

RC: Sec. 3.2: This section touches upon some of the points I raise above but a richer discussion in identifying the sources will be good to add. Consider using information from this section to inform inventory updates in the previous section. This does not need to be big changes but a short circling back would be useful.

AC **(L523-526, 613-615)**: As discussed above, while emission ratios are expected to reflect source characteristics and are used to diagnose national scale emissions, sector-specific information is often difficult to obtain. This is due to the complexity of regional anthropogenic activities, leading to greater errors from smaller footprints and highly variable emission factors. Therefore, we basically have to remain conservative. Nonetheless, several implications from this and our previous studies were drawn. Fire emissions, such as in the THL case, tend to exhibit a more uniform regional pattern, making them relatively easier to predict. A practical approach for inventory updates is to focus on dominant sources. In CHN, BC inventory can be improved by targeting major sources, specifically raw coal used in traditional cooking and heating stoves, based on their high emission factors and activity levels in the ECLIPSEv6b inventory (discussed in Sect. 3.2.2). Similarly, CO and CO_2 emission factors could be updated in the same sectors where BC requires revision, as well as other potentially missing sources.
We add a brief summary at the end of **Sect. 3.2.2** and **3.2.4**.

RC: Sec. 4: Consider adding recommendations on **how inventory-developers can directly utilize** such campaigns' information to tune up or down in this section.

AC **(L644-646)**: We added brief summary sentences to the Conclusion.

RC: L653: Add discussion on whether there is any chance or reason why the BC/CO ratios in the model and observations are similar due to completely different reasons, such as rapid BC deposition in the model offset by higher BC emissions?

AC **(L454-456)**: Though considered, it would be difficult to reproduce the observed BC/CO ratios in the model by changing deposition parameters in a reasonable range. The horizontal and vertical distributions of BC and CO at layers below the measurement altitudes (not shown) showed no indication of increased emissions followed by rapid deposition in the model. The similarity in the BC/CO ratio between the model and observations (e.g., the PHL case in E-AS-03), despite concentration gaps, likely results from appropriate BC and CO emission factors but insufficient activity levels in REASv2.1. We added some sentences for clarification in the caption of **Figure 6**.