Authors' Response

Dear Editor,

We thank the referee for the thorough reevaluation of our revised manuscript and the many constructive comments. We have revised the manuscript accordingly. Below we respond point-by-point; all line numbers mentioned refer to the revised manuscript (tracked-changes version).

Anonymous Referee #1

Review of Li et al. (2025), "Drivers and implications of declining fossil fuel CO_2 in Chinese cities revealed by radiocarbon measurements"

I reviewed a previous version of this article for publication in ACP. My major concern was that the estimated trend in C_{ff} concentration was unreliable because the $\Delta^{14}C$ observations had been taken at different locations/times of the day, and were compared to different backgrounds. Furthermore, given the high variability in atmospheric mixing, it had not been demonstrated how representative the flask observations with an integration time of 15-20 min are.

In this revised version of the manuscript, the authors have addressed the comments I made. They have provided additional information, such as meteorological data and back-trajectory simulations, to help assess the representativeness of their observations. However, what I'm still missing is to discuss the direct impact of potential differences in atmospheric mixing on the observed concentrations. For instance, the authors present standardized anomalies (z) of the planetary boundary layer height (PBLH) during the flask sampling times in Dec 2022. They obtained a z-value of about 0.2, indicating that the PBLH during the flask sampling period was higher than the average PBLH in winter 2021/22 and the climatological winter mean (see Fig. G1). What implications does this have for the observed concentrations? A higher PBLH indicates a larger mixing volume, resulting in more diluted emissions. Therefore, how much would a z-value of 0.2 alter the concentrations if emissions remained constant over time? When using the observed trend in C_{ff} concentration to draw conclusions about changes in emissions, it is essential to carefully discuss which part of the observed C_{ff} trend can be explained by changes in atmospheric mixing.

Response: We thank the reviewer for raising this important point. We agree that variability in atmospheric mixing, particularly changes in the planetary boundary-layer height (PBLH), can directly modulate near-surface enhancements and therefore needs to be considered when interpreting inter-period differences in observed C_{ff} and when discussing implications for emission changes.

To provide a quantitative, first-order constraint on this effect, we added a simple mixed-layer ("box") scaling argument. For predominantly surface-derived tracers under well-mixed boundary-layer conditions, the near-surface enhancement scales approximately as

$$C_{\mathrm{ff}} \propto \frac{1}{\mathrm{PBLH}},$$

which implies (to first order)

$$\frac{\Delta C_{\mathrm{ff}}}{C_{\mathrm{ff}}} \approx -\frac{\Delta \mathrm{PBLH}}{\overline{\mathrm{PBLH}}}.$$

We relate the standardized PBLH anomaly z to the relative PBLH deviation via

$$z = \frac{\mathrm{PBLH} - \overline{\mathrm{PBLH}}}{\sigma_{\mathrm{PBLH}}} \Rightarrow \frac{\Delta \mathrm{PBLH}}{\overline{\mathrm{PBLH}}} = z \cdot \frac{\sigma_{\mathrm{PBLH}}}{\overline{\mathrm{PBLH}}}.$$

This scaling provides an order-of-magnitude estimate of the mixing-volume effect and should be interpreted as a first-order sensitivity rather than a full attribution, as horizontal transport and mesoscale circulations can also influence the observed C_{ff}.

Using ERA5 for the Guangzhou sites and the same local-time sampling window as the flasks, we obtain $z \approx 0.17$. For transparency, we report the corresponding winter reference mean and standard deviation used to define z (Fig. G1), which yields a relative PBLH increase of ~11%. Under the inverse scaling above (and assuming emissions and other transport factors remain unchanged), this would imply an expected dilution of C_{ff} of ~11% during those sampling hours. For typical wintertime C_{ff} levels in Guangzhou, this would translate into a concentration difference of only a few μmol mol^{-1}, i.e. substantially smaller than the observed inter-period difference.

We then explicitly compare this mixing-driven effect to the observed inter-period difference. The harmonized wintertime C_{ff} in Guangzhou decreases from 28.7–34.9 μmol mol^{-1} in 2010 (afternoon-equivalent range derived from the observed nocturnal enhancement) to 12.5 ± 3.4 μmol mol^{-1} in 2022, i.e. an absolute difference of 16.2–22.4 μmol mol^{-1}. Therefore, the PBLH anomaly implied by $z \approx 0.17$ would be expected to affect C_{ff} at the ~10% level (i.e.,1-3 μmol mol^{-1}), and is thus unlikely to account for the full observed inter-period difference (16.2–22.4 μmol mol^{-1}).

Finally, we clarified the manuscript language to distinguish observed concentration changes from emission changes: our analysis quantifies inter-period changes in observed C_{ff}, whereas any inference about emission trends must be interpreted with caution in the context of transport/mixing variability (including PBLH) and the limited sampling in the target period.

Corresponding manuscript changes

Section 3.4.1: Meteorological typicality of the sampling months

"These results demonstrate that both sampling periods were representative of their respective seasonal transport conditions. Consequently, atmospheric transport variability is unlikely to bias the reported C_{ff} trend or the inferred changes in $R_{CO/CO2ff}$ ratios and fuel-type contributions." was changed to

"Overall, these diagnostics suggest that the sampling windows in both 2010 and 2022 were not associated with anomalous large-scale transport. Nevertheless, variability in mixing and transport at sub-monthly scales may still contribute to uncertainty, especially given the limited number of winter flasks in 2022. Accordingly, we treat transport/mixing variability as an uncertainty in the inter-period comparison rather than assuming it to be negligible."

Section 3.4.2: Representativeness of weekly flask samples

"This demonstrates that the weekly flask observations are meteorologically and dynamically representative of their respective seasonal backgrounds, minimizing the potential bias from short-term synoptic variability. Consequently, the

derived C_{ff} concentrations from these samples can be regarded as seasonally robust, and the subsequent interannual comparisons mainly reflect emission-driven rather than sampling-driven differences." was changed to

"We therefore consider the weekly flask observations to be broadly representative of their seasonal backgrounds in terms of large-scale transport, while noting that the discrete nature of flask sampling (and the small winter 2022 sample size) limits the ability to fully average out synoptic-scale variability."

Section 3.4.3: Historical variation of C_{ff} concentrations

"Because both August and December 2022 were meteorologically typical, the observed inter-annual differences in C_{ff} are attributed mainly to emission rather than transport variability. To ensure comparability, all available historical datasets (Table H1) were harmonized to identical sites, seasons, and local-time windows, and recalculated using unified background references (Table H2, Fig. 4a). These harmonized datasets minimize transport and spatial biases, allowing the remaining differences in C_{ff} mole fractions to be interpreted as primarily emission driven." was changed to

"To ensure comparability, all available historical datasets (Table H1) were harmonized to identical sites, seasons, and local-time windows, and recalculated using unified background references (Table H2, Fig. 4a). This harmonization reduces methodological differences (e.g., background choice and sampling-window differences) and facilitates an inter-period comparison of C_{ff} mole fractions, while transport and mixing variability remains a source of uncertainty."

"This adjustment yields 27.3 ± 16.9 μmol mol^{-1} for 2010 and 11.6 ± 3.4 μmol mol^{-1} for 2022, indicating a pronounced reduction in fossil-fuel-derived CO_2 over the decade." was changed to

"This adjustment changes C_{ff} from 45.6 ± 5.3 μmol mol^{-1} to 44.2 ± 5.3 μmol mol^{-1} for 2010, and from 16.8 ± 3.4 μmol mol^{-1} to 12.5 ± 3.4 μmol mol^{-1} for 2022."

"Even after harmonizing the background and correcting for sampling-time bias, C_{ff} in Guangzhou declined by ≈ 34–46 % between 2010 and 2022, confirming a genuine and statistically significant ($p < 0.01$) decrease in fossil-fuel CO_2. FLEXPART footprint analyses for 2010 and 2022 show similar source-sensitivity patterns centered on the Guangzhou urban core, confirming that GZ7 remains spatially representative of Guangzhou's urban domain." was changed to

"In addition to harmonizing background $\Delta(^{14}C)$ and sampling times, we explicitly evaluated the impact of changes in boundary-layer mixing between 2010 and 2022 (Sect. 3.4.1, Table G1). To assess how much of the inter-period difference could plausibly be explained by changes in boundary-layer mixing, we provide a first-order estimate of the sensitivity of near-surface C_{ff} to PBLH variations. Under a well-mixed boundary-layer "box" approximation, the surface enhancement of predominantly surface-emitted tracers scales approximately as $C_{ff} \propto 1/\mathrm{PBLH}$, implying $\Delta C_{ff}/C_{ff} \approx -\Delta\mathrm{PBLH}/\mathrm{PBLH}$. Using ERA5 PBLH at the actual flask sampling hours (Fig. G1), the standardized anomaly of PBLH in Dec 2022 at Guangzhou sites is $z \approx 0.17$, corresponding to a relative PBLH increase of 11% (based on the local winter mean and standard deviation used to define z). If emissions and other factors were unchanged, this would translate into an expected dilution of C_{ff} by 11% (i.e., 1–3 μmol mol^{-1} for typical wintertime C_{ff} levels). This indicates that the modestly higher PBLH in Dec 2022 would tend to reduce the observed C_{ff}, but its magnitude is smaller than the observed inter-period difference (16.2–22.4 μmol mol^{-1}).

Taken together, after harmonizing the $\Delta(^{14}C)$ background and accounting for sampling-time differences, the observations indicate an indicative inter-period decrease in wintertime C_{ff} in Guangzhou between 2010 and 2022. Using the CO-based diurnal scaling (21–35% nighttime enhancement), the 2010 value corresponds to an afternoon-equivalent

C_{ff} of 28.7–34.9 μmol mol^{-1}, compared to 12.5 ± 3.4 μmol mol^{-1} in 2022 (i.e., 56–64% lower; the range reflects uncertainty in the diurnal scaling). This percentage refers to the observed concentration change and may include a modest contribution from differences in boundary-layer mixing; our first-order PBLH-based scaling suggests that the Dec 2022 mixing anomaly would affect C_{ff} at the ~10% level (1–3 μmol mol^{-1}). Given the limited number of winter flasks in 2022, we performed a leave-one-out sensitivity test (Appendix H), which shows that the inferred 2010–2022 decrease remains negative for all subsets, although the magnitude varies. Accordingly, we interpret the Guangzhou 2010–2022 difference as an indicative inter-period change rather than a robustly quantified long-term trend. FLEXPART footprint analyses for 2010 and 2022 show similar source-sensitivity patterns centered on the Guangzhou urban core, supporting that GZ7 remains representative of Guangzhou's urban influence domain in both periods."

What has been improved is the harmonization of the dataset. The authors revised their analysis to make the C_{ff} estimates more comparable, e.g. by using a consistent $\Delta^{14}C$ background and observations from the same sites when deducing the trend in C_{ff} concentration. However, this data harmonization means that there are far fewer observations available for the C_{ff} trend analysis. If I understand correctly, one of their key statements in the abstract ("We found distinct regional trends: megacities like Guangzhou show significant C_{ff} declines (34–46 % decrease from 2010 to 2022) along with their source regions") is now based solely on four winter flask observations in 2022. Given the poor statistics and the high variability in atmospheric mixing, I wonder how reliable and robust this statement is.

Response: We thank the reviewer for highlighting the trade-off between improved harmonization and reduced sample size. We agree that, after applying consistent site/season/time-window selection and a unified $\Delta(^{14}C)$ background, the number of winter observations available for Guangzhou in 2022 is small (n = 4). Consequently, the 2010–2022 difference is more sensitive to residual variability in transport and boundary-layer mixing, and it should not be presented as a fully robust, statistically well-constrained "trend".

To address this concern, we have taken the following actions in the revised manuscript:

1. **Increased transparency on sample size and statistics.**
   We now explicitly report the number of flasks used for each period (n) together with the harmonized C_{ff} values and uncertainties in the main text, and provide full details in Table H2.

2. **Reframed the result as an inter-period difference (indicative) rather than a robust trend.**
   We revised the Abstract and Conclusions to avoid over-strong wording such as "distinct regional trends" and "significant declines". The Guangzhou result is now described as an *indicative inter-period decrease* in observed wintertime C_{ff}, explicitly noting the limited winter 2022 flask sampling.

3. **Explicitly discussed transport/mixing uncertainty and provided a first-order quantitative estimate.**
   Using ERA5 diagnostics and HYSPLIT back trajectories, we show that the sampling windows are not associated with anomalous large-scale transport. In addition, we include a first-order PBLH-based scaling to estimate the magnitude of mixing-driven modulation (order of ~10% for the Dec 2022 PBLH anomaly, corresponding to 1–3 μmol mol^{-1} for the wintertime C_{ff} levels observed here). This mixing-driven effect may contribute to the observed inter-period difference but is smaller than the observed 2010–2022 afternoon-equivalent C_{ff} decrease (16.2–22.4 μmol mol^{-1}, corresponding to 56–64%; the range reflects uncertainty in the diurnal scaling).

4. **Added a simple sensitivity test for the small 2022 sample size.**
   To further assess robustness given limited samples in winter 2022, we added a leave-one-out sensitivity analysis (Appendix H). The inferred 2010–2022 decrease remains negative for all subsets of the 2022 winter flasks, while the magnitude varies, supporting that the result is qualitatively robust but quantitatively uncertain. We therefore interpret the Guangzhou 2010–2022 difference as indicative rather than a precisely quantified long-term trend.

Overall, the harmonized datasets provide improved comparability and suggest an indicative decrease in observed wintertime C_{ff} in Guangzhou between 2010 and 2022 rather than a statistically well-constrained long-term trend. However, we acknowledge that additional $\Delta(^{14}CO_2)$ observations and/or transport modelling would be required to robustly quantify a long-term trend and to attribute it quantitatively to emission changes.

Corresponding manuscript changes

Abstract

"We found distinct regional trends: megacities like Guangzhou show significant C_{ff} declines (34–46 % decrease from 2010 to 2022) along with their source regions, while smaller cities have yet to demonstrate similar reductions. These improvements can be attributed to a 23 % coal consumption reduction, 17 % increased natural gas use (evidenced by stable isotope analysis), and improved combustion efficiency (indicated by 63 % falling $R_{CO/CO2ff}$ ratios)."

was changed to

"We found regional differences in C_{ff} and co-emission characteristics: megacities like Guangzhou show an indicative inter-period decrease in wintertime C_{ff} concentrations, of roughly 56–64 % lower in 2022 than in 2010 in afternoon-equivalent terms, while smaller cities have yet to demonstrate comparable decreases. These changes are consistent with a 23 % reduction in coal use, a 17 % increase in the natural-gas contribution (evidenced by stable isotope analysis), and improved combustion efficiency (indicated by a 63 % decline in $R_{CO/CO2ff}$ ratios)."

"These findings indicate nationwide progress toward C_{ff} emission peaks, with megacities leading the transition."

was changed to

"These findings are consistent with progress toward mitigating C_{ff} and co-emitted CO in major Chinese cities."

Besides that, I'm concerned about the strong impact of the $\Delta^{14}C$ background. When the "NL air" background is replaced by the "NL tree-ring" background, the C_{ff} estimate at the GZ7 site changes from 16.8 to 11.6 ppm (see Tab. H2). Why are the "NL air" and the "NL tree-ring" $\Delta^{14}C$ values so different? This needs to be resolved and understood in order to draw conclusions about the C_{ff} concentration and its trend.

Response: We thank the reviewer for raising this key issue. We agree that inferred C_{ff} is sensitive to the choice of the $\Delta^{14}C$ background and that the difference between the "NL air" and "NL tree-ring" background products must be explained. We also agree that the implications of this background sensitivity should be made explicit when interpreting the absolute C_{ff} level and any apparent interannual change at GZ7.

Why can "NL air" and "NL tree-ring" $\Delta^{14}C$ differ?

In our case, the December 2022 "NL air" background (-10.6 ± 0.8 ‰; Table A1) is higher (less negative) than the 2022 "NL tree-ring" value (-20.8 ‰). Importantly, these two values do not represent the same atmospheric signal in time: the former reflects discrete wintertime atmospheric observations, whereas the latter is a growing-season-integrated biospheric proxy and (for 2022) an extrapolated estimate. Given these differences in integration window and representativeness, we do not expect the two products to match quantitatively; we therefore address the issue by clarifying their intended use and explicitly propagating background-choice sensitivity into the Guangzhou interpretation. Differences can arise for the following reasons:

1. **Discrete wintertime air samples versus growing-season-integrated proxy.**
   "NL air" refers to atmospheric $\Delta^{14}C$ observations at Nanling during **December 2022**, which can reflect synoptic transport and boundary-layer variability on event time scales. In contrast, the NL tree-ring record integrates carbon assimilated during the **local growing season (March–October)** and therefore reflects a growing-season-weighted $\Delta^{14}C$ proxy with strongly reduced high-frequency variability. It is thus not expected to represent wintertime background conditions or day-to-day variability.

2. **Tree-ring $\Delta^{14}C$ is not a direct atmospheric measurement and can deviate from the regional atmospheric mean.**
   Even within the growing season, tree-ring $\Delta^{14}C$ does not necessarily equal the regionally mixed atmospheric background. Potential deviations can arise from seasonal weighting of assimilation, near-canopy/local influences (including local fossil-fuel CO_2 or biospheric respiration), and post-photosynthetic storage/mixing processes that blur the exact time window represented. Therefore, the tree-ring product is best interpreted as an internally consistent, growing-season-integrated reference rather than a point-in-time atmospheric background.

3. **The 2022 tree-ring value is an extrapolated estimate.**
   The 2022 "NL tree-ring" $\Delta^{14}C$ value (-20.8 ‰) is not a direct 2022 observation; it is linearly extrapolated from the 2011–2020 NL tree-ring record (Li et al., 2025b). This introduces additional uncertainty and may contribute to the difference from the December 2022 atmospheric samples.

4. **Physical rather than analytical origin.**
   Both products were produced within the same laboratory and reporting framework, so the discrepancy is unlikely to reflect an inter-laboratory calibration offset; it more plausibly reflects differences in temporal representativeness and, for the tree-ring product, proxy/extrapolation uncertainty.

How do we address this in the revised manuscript?

(i) **Clarified the intended use of each background product.**
We retain "NL air" (contemporaneous atmospheric $\Delta^{14}C$ observations in December 2022) as the primary background for deriving C_{ff} for the 2022 Guangzhou flasks. The NL tree-ring product is used only as a harmonized reference baseline for the inter-study inter-period comparison (2010 vs. 2022), to avoid background-definition inconsistencies across studies. Specifically, Ding et al. (2013) used a corn-leaf background of 37.5 ± 3.0 ‰ for 2010, whereas the NL tree-ring $\Delta^{14}C$ reference for 2010 is 33.9 ‰ (Li et al., 2025b). Using the NL tree-ring record as a common reference baseline therefore reduces methodological differences in background definition. We explicitly note that the tree-ring record is growing-season-integrated and extrapolated to 2022, and therefore it is not intended to represent wintertime background variability.

(ii) **Propagated background choice into the interpretation.**
We now state explicitly that absolute C_{ff} levels and the inferred magnitude of the 2010–2022 difference are background-sensitive. Accordingly, we interpret the 2010–2022 comparison as indicative rather than a quantitatively robust long-term trend. We also report the sensitivity of C_{ff} to background choice in Appendix H (Table H2): while the qualitative direction of change (lower wintertime C_{ff} in 2022 than in 2010) is preserved, the magnitude depends on the background assumption.

Corresponding manuscript changes

(A) Replace the "directly comparable" sentence (Section 3.4.3)

Original: "To make the datasets directly comparable, the winter 2010 C_{ff} values from Ding et al. (2013) and the winter 2022 values from this work were recalculated using the NL tree-ring $\Delta^{14}C$ record (Li et al., 2025b) as a common background reference."

Revised (replace with): "To *harmonize the background reference used in the C_{ff} calculation between studies, the winter 2010 C_{ff} values from Ding et al. (2013) and t*he winter 2022 values from this work were recalculated using the NL tree-ring $\Delta^{14}C$ record (Li et al., 2025b) as a common reference baseline. *The NL tree-ring $\Delta^{14}C$ represents a growing-season (March–October) integrated proxy and the 2022 value is linearly extrapolated from the 2011–2020 record; it is therefore not intended to represent wintertime background variability and is used here only to provide an internally consistent baseline for inter-study comparison*."

(B) Add a brief clarification in Figure 4 (caption)

"*For Guangzhou, the inter-study harmonization in Table H2 uses a common NL tree-ring $\Delta(^{14}C)$ reference baseline (growing-season integrated; extrapolated to 2022 from the 2011–2020 record; Li et al., 2025b) to harmonize background definitions across studies (used for harmonization only, not as a winter background)*."

(C) Footnotes (d, e)

"[d] corrected to 14:00 sampling and *recalculated using a common NL tree-ring $\Delta(^{14}C)$ background as a harmonized reference baseline for inter-period comparison. The NL tree-ring $\Delta(^{14}C)$ represents a growing-season (March–October) integrated proxy at Nanling, and the 2022 value is linearly extrapolated from the 2011–2020 tree-ring record (Li et al., 2025b); it is therefore not intended to represent wintertime background variability*. [e] *recalculated using the same NL tree-ring $\Delta(^{14}C)$ harmonized reference baseline as in footnote d*."

Overall, one must be very careful when attributing the trends in C_{ff} concentration to changes in emissions. To reliably estimate trends in emissions, an atmospheric transport model is required, e.g. in combination with inverse modelling. This should really be emphasized throughout the manuscript, including in the title (add "concentration" ager "declining fossil fuel CO_2"). Therefore, conclusions regarding trends in emissions should be tentative rather than quantitative in this manuscript. As also suggested by the other reviewer, the present study could focus more on trends in the $R_{CO/CO2ff}$ ratios, which are less influenced by atmospheric mixing (see my comment on this in the previous report), and could lead to more robust conclusions.

Response: We thank the reviewer for this overarching and important recommendation. We fully agree that trends in

observed C_{ff} **concentrations** cannot be quantitatively attributed to **emission** trends without an atmospheric transport framework (e.g. a chemistry-transport model combined with inverse modelling). In the absence of such modelling, emission-related statements must remain tentative and qualitative, and we do not derive quantitative emission trends in this manuscript.

In response, we revised the manuscript to (i) consistently frame our results as changes in observed concentrations rather than as emission trends, (ii) explicitly state that robust emission-trend estimation requires transport/inverse modelling, and (iii) place greater emphasis on diagnostics that are less sensitive to mixing, such as $R_{CO/CO2ff}$.

Specifically, we:

1. Revised the title to explicitly refer to concentrations, as suggested by the reviewer:
   New title: "Drivers and implications of declining fossil fuel CO_2 concentrations in Chinese cities revealed by radiocarbon measurements".

2. Softened the wording throughout the Abstract and Conclusions by:

   - Replacing emission-trend language (e.g. "emission trends", "indicate nationwide progress toward C_{ff} emission peaks") with formulations such as "observed decreases in C_{ff} concentrations", "consistent with" or "indicative of emission reductions"; and

   - Explicitly mentioning that transport and mixing variability (including boundary-layer changes) limit our ability to infer emission trends from concentration changes alone.

3. Added a dedicated limitation statement in the Results and Discussion clarifying that we do not perform an inversion and therefore cannot quantitatively separate emission changes from transport/mixing effects. At the end of the first paragraph in Section 3.4.3, we now write: "We emphasize that the following comparison addresses observed near-surface C_{ff} concentrations. Without an atmospheric transport model and inverse modelling, we cannot quantitatively attribute the observed inter-period concentration differences to emission changes."

4. Following the reviewer's suggestion, we place greater emphasis on trends in the $R_{CO/CO2ff}$ ratios, which are less influenced by atmospheric mixing. In the revised text (Sections 3.3 and 3.4), we explicitly highlight $R_{CO/CO2ff}$ as a more robust indicator of changes in combustion efficiency and co-emission characteristics, and we frame the corresponding conclusions more strongly than those based solely on C_{ff} concentration trends.

Together, these changes ensure that the manuscript focuses on observed concentration changes, treats any emission-related interpretation as tentative, and clearly acknowledges that robust emission-trend estimation would require an atmospheric transport model and inverse modelling.

Specific comments (mainly focusing on the revised parts of the manuscript):

l. 185-187: This sentence is slightly misleading because, ultimately, you use the correction estimate from Turnbull et al. (2009), which does not take into account biomass burning. Similar is true for the second sentence in the conclusions (l. 630-632).

Response: We thank the reviewer for this helpful comment and agree that the current wording may be misleading, as it could be interpreted to mean that the *applied* correction itself explicitly includes biomass burning (BB). In our study, BB emissions (and their uncertainties) are explicitly included in the simulation framework and full error analysis for C_{ff}, but for the *final* C_{ff} values we adopt the correction estimate from Turnbull et al. (2009), which does not explicitly account for BB, in order to maintain consistency and comparability with previous studies.

To clarify this, we revised the text as follows.

At lines 183–191, we replaced the current text with:

"This study is the first to explicitly account for BB emissions within a C_{ff} estimation framework, allowing us to quantify their contribution and associated uncertainty relative to Rh under our assumptions. To maintain methodological consistency and comparability with previous work, the final C_{ff} values reported here adopt the correction estimate from Turnbull et al. (2009), which does not explicitly include BB. Nevertheless, our simulations, which incorporate BB emissions and their uncertainties, indicate that the magnitude of the required correction (≤ -0.5 µmol mol^{-1}) is broadly consistent with Turnbull et al. (2009), and that our main conclusions are robust across this range of potential corrections."

At lines 683–685 in the conclusions, we replaced the first point with a shorter and clearer statement:

"First, we provide a comprehensive error analysis framework for C_{ff} estimation, including contributions from air–sea exchange, nuclear facilities, and particularly biomass burning."

We believe these changes make it clear that (i) BB is fully considered in our error analysis and simulations, and (ii) the applied correction for the reported C_{ff} values follows Turnbull et al. (2009) for comparability, while our conclusions are robust to the range of corrections suggested by our BB-inclusive analysis.

l. 194-201: Why don't you use these FLEXPART simulations to assess the impact of transport variability on the trends in the C_{ff} concentrations?

Response: We thank the reviewer for this suggestion. FLEXPART was used in this study to characterize the footprint and source–receptor sensitivity of the observations (i.e., to support the interpretation of the sampled upwind regions and potential source influences). A quantitative assessment of transport-driven variability in long-term trends in C_{ff} would, however, require a dedicated meteorological normalization framework (e.g., multi-year simulations with consistent meteorological input, careful treatment of changes in model configuration and resolution through time, and an explicit separation of emission changes from meteorology). Implementing and validating such a normalization is beyond the scope of this minor revision and would constitute a substantial additional analysis rather than a minor change.

To address the reviewer's concern without adding new analyses, we have clarified in the manuscript (Sect. 2.5) that FLEXPART is used here for qualitative footprint characterization and not for meteorological normalization of C_{ff} trends. We also note that the C_{ff} trends reported are derived from $\Delta(^{14}CO_2)$-based source separation and are therefore primarily constrained by the measurements; a full quantification of the transport contribution to the trend will be addressed in future work.

Corresponding manuscript changes

Section 2.5: C_{ff} footprint by FLEXPART model

"Surface flux sensitivity of C_{ff} were conducted using the FLEXible PARTicle (FLEXPART) dispersion model (version 10.4) (Pisso et al., 2019)." was changed to

"Surface flux sensitivity simulations for C_{ff} were performed using the FLEXible PARTicle (FLEXPART) dispersion model (version 10.4) (Pisso et al., 2019). In this study, FLEXPART is used to characterize source–receptor sensitivities ("footprints") to support qualitative interpretation of the sampled upwind regions and potential source influences; it is not used to meteorologically normalize the long-term trends in C_{ff}."

l. 226-230: $R_{CO/CO2ff}$ is not an "emission ratio" because it is based on the ΔCO and C_{ff} concentrations.

Response: We thank the reviewer for this helpful clarification. We agree that $R_{CO/CO2ff}$, as used in our study, is not an emission ratio in the strict sense, since it is derived from observed ΔCO and C_{ff} concentrations rather than directly from emissions. In response, we have revised the wording in Sections 2.7 and 3.5 to avoid calling $R_{CO/CO2ff}$ an "emission ratio" and now explicitly define it as an "observational concentration ratio". For example, the phrase 'emission ratio $R_{CO/CO2ff}$' in the original text has been revised to 'observed concentration ratio of ΔCO to C_{ff} ($R_{CO/CO2ff}$)'. We have also checked the manuscript to ensure that this terminology is used consistently wherever $R_{CO/CO2ff}$ is discussed.

l. 229-230: To avoid any confusion, please state here that this approach is applied for sites/times without $\Delta^{14}C$ observations.

Response: We thank the reviewer for this helpful comment. To avoid confusion, we have clarified in the text that the 20 % correction for non-fossil CO_2 is applied only at sites and during periods without concurrent $\Delta^{14}CO_2$ observations. The relevant passage now reads:

"To correct for the contribution of non-fossil CO_2 in the observed enhancement, the concentration ratio $R_{CO/CO2ff}$ was estimated by dividing observed $R_{CO/CO2}$ by 0.8 *for sites and times without $\Delta^{14}CO_2$ observations*. Previous $\Delta^{14}CO_2$ and CO–CO_2 studies (Turnbull et al., 2011; Lopez et al., 2013; Newman et al., 2016; Miller et al., 2020) have shown that ~10–30 % of the total CO_2 enhancement above background during daytime is typically of non-fossil origin, while CO is emitted almost exclusively from fossil-fuel combustion. Thus, the 20 % correction represents a reasonable first-order approximation for well-mixed afternoon conditions."

l. 316-318: What do you mean by "the atmospheric trapping of emissions is higher than local emissions"? It seems that the analysis in this section relies heavily on the assumption that the emission inventory is accurate. What if some sources are missing or at the wrong place in the inventory? I suggest removing this analysis or providing clearer, more quantitative validation for this statement.

Response: We thank the reviewer for raising this important point. We agree that the phrase "the atmospheric trapping of emissions is higher than local emissions" was unclear and could be interpreted as an inappropriate comparison between a transport/dispersion process and an emission magnitude. Our intention was to convey that reduced atmospheric ventilation (e.g., a shallow boundary layer and stagnant conditions) and transport from upwind source

regions can amplify observed C_{ff} enhancements and shape the spatial gradients, rather than to quantify "trapping" versus "local emissions".

We also agree that any footprint-weighted inventory attribution of local versus non-local contributions is conditional on the completeness and spatial allocation of the emission inventory, and may be biased if sources are missing or mislocated. Because a full quantitative validation of the inventory and a meteorological-normalization analysis are beyond the scope of this minor revision, we have revised Sect. 3.3 to (i) remove the ambiguous statement, (ii) rephrase the discussion in physically consistent terms (ventilation/transport affecting observed enhancements), and (iii) explicitly acknowledge the inventory-related uncertainties and treat the ODIAC/MEIC comparisons as qualitative support rather than a definitive attribution.

Corresponding manuscript changes

Section 3.3: Spatial distribution and seasonal variations

"This was further supported by significant positive correlations between the C_{ff} measurements and the corresponding 1x1 km gridded ODIAC (Oda and Maksyutov, 2011; Oda and Maksyutov, 2024) inventory emissions in GZs ($r = 0.53$, $p = 0.1$) and SGw ($r = 0.91$, $p = 0.03$)." was changed to

"This was further supported by a positive correlation between the C_{ff} measurements and the corresponding 1 × 1 km gridded ODIAC (Oda and Maksyutov, 2011; Oda and Maksyutov, 2024) inventory emissions in GZs ($r = 0.53$, $p = 0.1$), and a significant positive correlation in SGw ($r = 0.91$, $p = 0.03$). These correlations are used here as qualitative support and should be interpreted cautiously given uncertainties in the emission inventory (e.g., missing or spatially misallocated sources)."

"This was likely attributable to atmospheric trapping of emissions from the Pearl River Delta (PRD) urban agglomeration, as illustrated by air mass back trajectories (HYSPLIT, Fig. F1a) and emission footprints (FLEXPART, Fig. 3g), rather than high local emissions due to lower summer emissions from inventories." was changed to

"Trajectory and footprint analyses suggest that summer observations at Shaoguan were frequently influenced by air masses arriving from the Pearl River Delta (PRD) urban agglomeration (HYSPLIT, Fig. F1a; FLEXPART, Fig. 3g), consistent with a larger upwind contribution under prevailing transport conditions."

"The higher winter concentrations found in Guangzhou, Shenzhen, and Zhanjiang in this study likely resulted from atmospheric trapping of emissions in the shallow planetary boundary layer, and high local emissions, because ODIAC (MEIC) indicates that winter emissions were 8 %, 10 %, and 11 % (17 %, 22 %, and 14 %) higher, respectively, than those in summer (Oda and Maksyutov, 2024; Meic, 2023). The atmospheric trapping of emissions is higher than local emissions during winter in Guangzhou (GZw) and Shenzhen (SZw), which is supported by higher C_{ff} concentrations occurring in downwind areas (GZ2, GZ6, and GZ10; SZ3 and SZ4) compared with upwind areas (GZ1 and GZ3; SZ8 and SZ9)." was changed to

"The higher winter C_{ff} levels likely reflect a combination of (i) reduced ventilation (e.g., a shallower planetary boundary layer) and (ii) higher wintertime emissions suggested by ODIAC/MEIC (winter emissions 8%, 10%, and 11% [ODIAC] and 17%, 22%, and 14% [MEIC] higher than summer for Guangzhou, Shenzhen, and Zhanjiang, respectively; Oda and Maksyutov, 2024; MEIC, 2023), noting the associated inventory uncertainties. Within Guangzhou and Shenzhen, wintertime spatial gradients show higher C_{ff} at downwind sites (GZ2, GZ6, and GZ10; SZ3 and SZ4) than at upwind

sites (GZ1 and GZ3; SZ8 and SZ9), suggesting an important role of transport/accumulation in shaping the observed enhancements.”.

l. 361-363: How do the meteorological parameters for the flask samples from 2022 and 2010 compare? In order to make the statement that the 'interannual comparisons mainly reflect emission-driven rather than sampling-driven differences', the meteorological parameters for the 2010 flask samples must also be investigated.

Response: We thank the reviewer for highlighting that a direct comparison of meteorological conditions between the 2010 and 2022 flask sampling periods was missing in the original manuscript. We have now explicitly analysed the meteorological parameters for the December 2010 flask samples and compared them with those for December 2022 at GZ7.

First, as shown in the revised Fig. G1e–f, all five standardized anomalies (U10, V10, T2M, SP, PBLH) for December 2010 at GZ7 lie within $|z| \leq 1$ relative to both the DJF 2010 mean and the 2010–2022 DJF climatology, indicating that the 2010 winter sampling month was meteorologically typical. ERA5 wind roses for December 2010 and DJF 2010 (Fig. G2g–h) further show dominant northerly to north-easterly flow, consistent with the climatological East Asian winter monsoon.

Second, we have added a quantitative comparison of ERA5 diagnostics on flask sampling days between December 2010 and December 2022 at GZ7 (new Tables G1–G2). Both periods are dominated by northerly to north-easterly winds, with winds in the 0–45° sector accounting for 61.7 % and 82.7 % of occurrences in December 2010 and December 2022, respectively (Table G2). However, December 2022 exhibits significantly stronger winds and deeper boundary layers than December 2010 (mean wind speed: 3.6 vs. 2.6 m s^{-1}; mean PBLH: 476.4 vs. 377.0 m; mean ventilation: 2024.5 vs. 1258.0 m^2 s^{-1}), with differences that are statistically significant ($p < 0.01$ for both Student’s t-test and Mann–Whitney U test; Table G1). These changes in wind speed and PBLH would tend to dilute near-surface enhancements in 2022 relative to 2010.

We have revised Sect. 3.4.1 to summarize this comparison (new paragraph in lines 373–380):

“To directly compare the meteorological environments of the two sampling years, we further analyzed ERA5 diagnostics on flask sampling days at GZ7 (Tables G1–G2). Both December 2010 and December 2022 were dominated by northerly to north-easterly flow, with winds in the 0–45° sector accounting for 61.7 % and 82.7 % of occurrences, respectively (Table G2). However, December 2022 exhibited stronger winds and deeper boundary layers than December 2010 (mean wind speed: 3.6 vs. 2.6 m s^{-1}; mean PBLH: 476 vs. 377 m; mean ventilation: 2024.5 vs. 1258.0 m^2 s^{-1}), and these differences are statistically significant ($p < 0.01$; Table G1). These conditions would tend to dilute near-surface enhancements in 2022 relative to 2010, implying that the observed decreases in C_{ff} and $R_{CO/CO2ff}$ between the two periods are, if anything, conservative with respect to emission changes.”

In Sect. 3.4.3 we further provide a first-order box-model estimate, showing that the December 2022 PBLH anomaly would affect C_{ff} at the ~10 % level (1–3 µmol mol^{-1}), much smaller than the observed 56–64 % decrease in afternoon-equivalent C_{ff}. Together, these additions demonstrate that we have explicitly investigated and quantified the meteorological differences between 2010 and 2022, and that these differences cannot explain the magnitude of the observed inter-period decrease, which we therefore interpret as being consistent with a primarily emission-driven

change.

l. 365-366: Again, this statement is not valid. The representativeness of the flask samples from 2010 has not been demonstrated.

Response: We agree that the representativeness of the 2010 flask samples needed to be demonstrated more explicitly in the original manuscript. We have now added a dedicated analysis of the meteorological and transport conditions during the December 2010 sampling period.

As noted above, Fig. G1e–f shows that standardized anomalies of U10, V10, T2M, SP and PBLH for December 2010 at GZ7 are within ±1σ of both the DJF 2010 seasonal mean and the 2010–2022 DJF climatology, indicating that the month as a whole was meteorologically typical. ERA5 wind roses for December 2010 and DJF 2010 (Fig. G2g–h) show dominant northerly to north-easterly flow, closely matching the DJF climatological regime and confirming that the 2010 winter sampling period was embedded in the canonical East Asian winter monsoon pattern. New Tables G1–G2 further demonstrate that the ERA5 diagnostics on flask sampling days in December 2010 are consistent with the corresponding monthly statistics and with the prevailing monsoon sectors.

We have revised Sect. 3.4.2 to explicitly state the representativeness of the 2010 samples (lines 392–395):

"For the December 2010 flask sampling at GZ7, ERA5 diagnostics and wind roses (Fig. G1e–f and G2g–h) likewise show that sampling-day conditions were consistent with the DJF 2010 seasonal mean and the 2010–2022 DJF climatology, indicating that these earlier samples were also collected under typical winter transport regimes."

We also removed overly strong wording that could be interpreted as claiming that transport variability is negligible, and instead emphasize that transport and mixing differences are treated as a source of uncertainty in the inter-period comparison (Sects. 3.4.1–3.4.3). With these changes, we believe that the representativeness of both the 2010 and 2022 flask samples is now more clearly demonstrated and appropriately caveated.

l. 382-383: A part of the C_{ff} concentration difference between afternoon and evening could maybe also be explained by a diurnal cycle in the emissions?

Response: We thank the reviewer for this insightful comment. We agree that, in addition to weaker nocturnal boundary-layer mixing, a diurnal cycle in emissions could also contribute to the difference in C_{ff} concentrations between afternoon and evening. In the revised manuscript, we now explicitly acknowledge this point and clarify that our interpretation likely primarily emphasizes the role of boundary-layer dynamics, while recognizing the potential influence of emission variability. The sentence has been modified to:

"These findings suggest that the evening C_{ff} level is typically 21–35 % higher than the well-mixed afternoon value, likely primarily due to weaker nocturnal boundary-layer mixing*, although a diurnal cycle in emissions may also contribute to this difference*."

Fig. 4: Are the scattered measurements from Guangzhou missing in panel (a)? I cannot see them.

Response: We thank the reviewer for pointing this out. In the original version, the individual Guangzhou measurements were indeed difficult to distinguish in Fig. 4a due to the plotting style. In the revised manuscript, we have ensured that the Guangzhou individual data points are clearly visible in panel (a) by adjusting the symbol size, contrast, and plotting order. We also revised the caption to more explicitly describe that large symbols indicate annual/multi-year/winter means and smaller symbols show the corresponding individual measurements, including those from Guangzhou.

Here is the revised caption of Fig. 4a:

Figure 4: (a) Harmonized comparison of C_{ff} mole fractions at the same sites and seasons, after applying consistent sampling time and background assumptions. C_{ff} concentrations are compiled from atmospheric measurements (Wang et al., 2022b; Zhou et al., 2022; Ding et al., 2013; Zhou et al., 2020; Wang et al., 2018) in Beijing, Xi'an, and Guangzhou. Large symbols indicate annual means, multiyear averages, or winter means (w) of the harmonized C_{ff} values listed in Table H2; small symbols represent the corresponding individual measurements. C_{ff} is calculated as enhancements over the regional background (Nanling for Guangzhou; Waliguan for Beijing and Xi'an). For Guangzhou, the inter-study harmonization in Table H2 uses a common NL tree-ring $\Delta(^{14}C)$ reference baseline (growing-season integrated; extrapolated to 2022 from the 2011–2020 record; Li et al. (2025b)) to harmonize background definitions across studies (used for harmonization only, not as a winter background). The y-axis error bars indicate uncertainty, and the x-axis error bars represent the observed period.

Fig. 5: Does the top figure in panel (b) show the C_{ff} estimates for the harmonized dataset?

Response: Thank you for raising this point. In the original version, the top panel in Fig. 5b was based on the full (non-harmonized) set of available C_{ff} estimates. In the revised manuscript, we have updated Fig. 5b so that the top panel now shows the C_{ff} estimates from the harmonized dataset used for the inter-study comparison (i.e., applying consistent sampling-time and background assumptions). To avoid ambiguity, we also revised the caption to explicitly state this and to clarify what is shown in each panel:

"**Figure 5:** (a) Coal, oil, and natural gas fractions of C_{ff} in Guangdong Province from the MEIC inventory from 1990 to 2021 (points), and in the cities of Guangzhou, Shenzhen, Zhanjiang, and Shaoguan from measurements in this study in 2022 (triangles). (b) (Top) Comparison of reductions in C_{ff} inventory emissions (blue) and **harmonized** measured C_{ff} concentrations in Guangzhou (red**; harmonized by applying consistent sampling-time and background assumptions**) **associated with** (Bottom) reduced coal usage and increased natural gas usage in Guangzhou. The vertical dashed line indicates the year 2013 when China's Clean Air Action Plan was implemented."

l. 546-552: What impact does the background have on the $R_{CO/CO2ff}$ ratios? When a regional background site is used, the $R_{CO/CO2ff}$ ratio may also be affected by emissions from outside the target city. Could this explain some of the discrepancies compared to the $I_{CO/CO2ff}$ ratios from the emission inventories?

Response: We thank the reviewer for this thoughtful and important comment. We agree that the choice of background can influence the inferred $R_{CO/CO2ff}$ ratios, particularly when a regional background site is used. In this case, the calculated ΔCO and C_{ff} enhancements are defined relative to a regional baseline rather than a pristine background, and thus may include contributions from emissions outside the target city. As a result, the observational $R_{CO/CO2ff}$ ratios can

represent a mixture of urban and regional emission signatures, rather than a purely city-scale signal, and this may contribute to some of the discrepancies when comparing $R_{CO/CO2ff}$ with $I_{CO/CO2ff}$ ratios derived from city-level emission inventories.

In the revised manuscript, we have clarified this point in both the "Data and methods" and the "Results and discussions" sections:

1. **Data and methods (The first paragraph in Sect. 2.7):**
   At the end of the first paragraph in Sect. 2.7, We have added the following sentences to explicitly describe the nature of the background and its implications for $R_{CO/CO2ff}$:

"The CO and CO_{2ff} enhancements were defined relative to a regional background site, which is intended to represent upwind regional conditions rather than a completely remote, pristine background. Consequently, the inferred ΔCO and C_{ff}, and thus the derived $R_{CO/CO2ff}$ ratios, may include contributions from emissions outside the target city. We do not explicitly correct for this potential bias because the relative contributions of urban versus regional sources cannot be robustly constrained with the available data, but we consider it as an additional source of uncertainty when comparing observational $R_{CO/CO2ff}$ with city-level $I_{CO/CO2ff}$ ratios from emission inventories."

2. **Results and discussions (The fourth paragraph in Sect. 3.5.2):**
   At the end of the paragraph discussing the comparison between observational and inventory-based ratios, we have added a sentence to explicitly acknowledge the impact of the background choice. The paragraph now reads:

"When a regional background site is used, however, the inferred $R_{CO/CO2ff}$ ratios may be influenced by emissions from outside the target city, so that the observed ratios represent a mixture of urban and regional emission signatures rather than a purely city-scale signal. This background effect may therefore contribute to some of the discrepancies between $R_{CO/CO2ff}$ and $I_{CO/CO2ff}$."

l. 727-728: When a regional wind climatology is used the impact of nuclear contamination might be smoothed out. The nuclear contamination may be higher for individual flask events when the power plant plume hits the observation site directly. It is important to note this, particularly as only a small number of flask samples are used for the C_{ff} trend analysis in Guangzhou.

Response: We thank the reviewer for this insightful comment. We fully agree that a purely climatological treatment of winds can smooth out episodic plume events and may, in principle, underestimate the impact of individual nuclear power plant (NPP) plumes on $\Delta^{14}CO_2$ and inferred C_{ff}. This concern is especially relevant when only a small number of flask samples are available, as is the case for the C_{ff} trend analysis in Guangzhou.

However, in our specific setting the potential impact of NPP emissions on the Guangzhou site is expected to be very small for two reasons:

1. **Large distances to NPPs for Guangzhou:**
   As summarized in Table C1, all major NPPs in Guangdong that could affect Guangzhou are located more than 100 km away from the Guangzhou observation sites (e.g. nearest distances > 100 km), whereas the closest

distances of 6–7 km apply only to the Shenzhen site (SZ10), not to Guangzhou.

2. **Very low $^{14}CO_2$ emission rates and dispersion effects:**
   In Guangdong Province, Zazzeri et al. (2018) estimated $^{14}CO_2$ emissions from the Daya Bay, Ling'ao and Yangjiang NPPs to be 0.111, 0.233 and 0.166 TBq yr^{-1}, respectively, values comparable to or smaller than those reported for Diablo Canyon. For the Shenzhen sites, which are only 6–7 km from Daya Bay and Ling'ao, we combined these emission rates with Gaussian plume scaling and regional wind climatology and found that the contribution of reactor $^{14}CO_2$ to measured $\Delta^{14}CO_2$ would be < 0.1 ‰, corresponding to an effect on inferred C_{ff} below 0.05 ppm. Under prevailing southeasterly winds in summer and northeasterly winds in winter, dispersion within the coastal boundary layer further dilutes any potential $^{14}CO_2$ plumes before they reach the sampling locations. For Guangzhou, where all NPPs are located at distances > 100 km, the nuclear influence on $\Delta^{14}CO_2$ and C_{ff} is therefore expected to be even smaller than for these closest stations.

To address the reviewer's concern more explicitly, we have revised the discussion in the manuscript to (i) stress that our quantitative estimate for the closest Shenzhen sites provides an upper bound on potential nuclear contamination, and (ii) state that, given the much larger distances for Guangzhou, any nuclear influence on individual flask samples and on the derived C_{ff} trend is expected to be negligible compared to other sources of uncertainty.

Specifically, we have revised the following text to the Appendix C2:

"In Guangdong Province, Zazzeri et al. (2018) estimated $^{14}CO_2$ emissions from the Daya Bay, Ling'ao, and Yangjiang NPPs to be 0.111, 0.233, and 0.166 TBq yr^{-1}, respectively, values comparable to or smaller than those reported for Diablo Canyon. Although Daya Bay and Ling'ao are located only 6–7 km from the nearest observation site (SZ10) (Table C1), their emission rates remain extremely low. Under prevailing southeasterly winds in summer and northeasterly winds in winter, dispersion within the coastal boundary layer further dilutes any potential $^{14}CO_2$ plumes before they reach the sampling locations. Based on Gaussian plume scaling and regional wind climatology, we estimate that *even under typical plume conditions,* the contribution of local reactor $^{14}CO_2$ to measured $\Delta^{14}CO_2$ at these Shenzhen urban sites would be <0.1 ‰, corresponding to an effect on inferred C_{ff} below 0.05 ppm. *This estimate can be regarded as an upper bound for potential nuclear contamination at our sites, because all NPPs that could influence Guangzhou are located at distances >100 km from the Guangzhou observation site (Table C1)*.

Therefore, even for the closest stations, the impact of nearby nuclear facilities on $\Delta^{14}CO_2$ measurements is considered negligible and does not affect our radiocarbon-based source partitioning. *For Guangzhou in particular, any nuclear influence on individual flask samples, and thus on the derived C_{ff} trend, is expected to be even smaller than this upper bound and negligible compared to other sources of uncertainty*."

Anonymous Referee #2

The authors have significantly improved the manuscript by conducting additional analysis. They have also addressed the reviewers' comments comprehensively.
The paper does rely on quite a few assumptions and extrapolations, but most are well-reasoned and explained in enough detail so the reader can follow the methodology.
Overall, the study is presented in a clear way with an easy to follow structure. Key findings are highlighted and figures

are mostly straightforward.
The data itself is of great interest to the scientific community as urban fossil fuel emissions and mitigation policies have regional and global impact.
There are a few remaining minor issues that should be resolved before publication.

Minor comments:

Throughout the manuscript the authors refer to 3 decades of observational data. However, observational data is from1998 to 2022 listed in tables, i.e. 24 years. Furthermore, error bars/standard deviations are reported in the manuscript in figures, but it is unclear if they have been fully considered when estimating trends and their uncertainties.

Response: Thank you for pointing this out. We agree that our wording was imprecise. The observational record compiled in this study spans 1998–2022 (i.e., 24 years) and should not be described as "three decades". We have corrected the text throughout the manuscript to refer to a "24-year record (1998–2022)", and we now reserve "three decades" only for contexts that explicitly refer to the early-1990s **inventory** time series.

We also agree that the uncertainty treatment needed to be stated more explicitly, because the error bars shown in different figures represent different quantities. We have revised the manuscript (text and figure captions) to clarify the following:

1. **Fig. J1:** error bars denote ±1 SD of the seasonal $R_{CO/CO2ff}$ ratios at each site.
2. **Fig. J2:** vertical and horizontal error bars denote the propagated uncertainties in C_{ff} and CO, respectively, obtained by error propagation combining measurement and calculation uncertainties.
3. **Fig. 6 (including the four-city winter 2022 points):** for observation-based $R_{CO/CO2ff}$, vertical error bars denote the uncertainty of the fitted ΔCO-C_{ff} regression slope. Horizontal error bars indicate the time span of each observation period, and the symbol is plotted at the median time.

Where we quantify changes between two observation periods, we assess robustness by comparing the change to the combined 1σ uncertainty (added in quadrature, using the reported uncertainties for the two periods). We further clarify that, due to the sparse and episodic nature of the compiled observational record (discrete campaigns rather than a continuous annual time series), we do not fit a single formal linear trend with one slope uncertainty across 1998–2022; instead, we focus on period-to-period differences and interpret them as indicative long-term changes.

Corresponding manuscript changes

(1) Text change (Sect. 3.5.2):

We inserted the following sentences immediately after: "We retrieved historical $R_{CO/CO2ff}$ data from observations in China by estimation from $\Delta(^{14}C)$ measurements and correction from $R_{CO/CO2}$ (increased by 20 %) (Table J1), and $I_{CO/CO2ff}$ data from the MEIC, MIXv2, and EDGAR inventories (Fig. 6)."

"Because the observational record consists of discrete campaigns, for the observations we assess changes using inter-period differences (rather than fitting a single 1998–2022 linear trend), and we test robustness by comparing the inferred change with the combined 1σ uncertainties (added in quadrature, using the reported vertical 1σ uncertainties for each period)."

We revised "Observational and inventory data show a sustained decline in China's $R_{CO/CO2ff}$ and $I_{CO/CO2ff}$ ratios over the past 30 years (Fig. 6a), demonstrating that efforts to improve fossil fuel combustion efficiency are effective (Wang et al., 2010; Lee et al., 2020), which is another factor contributing to the reduction in C_{ff} concentrations." to be "The compiled observations (1998-2022) and inventories (1990-2022) both indicate $R_{CO/CO2ff}$ and $I_{CO/CO2ff}$ ratios tend to be lower in recent years than in earlier periods (Fig. 6a), consistent with improved combustion efficiency (Wang et al., 2010; Lee et al., 2020), which is another factor contributing to the reduction in C_{ff} concentrations.".

(2) Caption changes:

We revised the captions of Figs. J1, J2 and 6 to explicitly define the error bars as follows:

- **Fig. J1:** "Error bars denote ±1 SD of the seasonal $R_{CO/CO2ff}$ ratios at each site."
- **Fig. J2:** "Vertical and horizontal error bars denote the propagated uncertainties in C_{ff} and CO, respectively, obtained by error propagation combining measurement and calculation uncertainties."
- **Fig. 6:** "For observation-based $R_{CO/CO2ff}$, vertical error bars denote the uncertainty of the fitted ΔCO-C_{ff} regression slope. Horizontal error bars indicate the time span of each observation period, and the symbol is plotted at the median time."

(In addition, we replaced "three decades" with "24-year record (1998–2022)" throughout the manuscript, and we use "three-decade" only when explicitly referring to inventory time series starting in the early 1990s.)

Specific comments:

Line 79: Please add reference here, e.g. https://doi.org/10.1016/j.rse.2019.111353

Response: We thank the reviewer for this helpful suggestion. We have now added the suggested reference "(Taubenböck et al., 2019)" at line 79 of the revised manuscript and included the full citation in the reference list:

Taubenböck, H., Weigand, M., Esch, T., Staab, J., Wurm, M., Mast, J., Breunig, M., and Siedentop, S.: A new ranking of the world's largest cities – Do administrative units obscure morphological realities?, Remote Sens. Environ., 232, 111353, https://doi.org/10.1016/j.rse.2019.111353, 2019.

Line 84: Please provide a reference for the information here.

Response: We thank the reviewer for pointing this out. In response to the comment, we have revised the sentence at line 84 to include explicit data sources and updated the percentages accordingly. The sentence now reads:
"Zhanjiang features extensive cultivated land (31.7 %) and coastal ports (Zmbs, 2025), while Shaoguan is distinguished by 74.5 % forest coverage (Smbs, 2024)."

Here, Zmbs and Smbs refer to the Zhanjiang Statistical Yearbook and Shaoguan Statistical Yearbook, respectively. We have also added the corresponding references to the reference list as:

Zhanjiang Municipal Bureau of Statistics: Zhanjiang Statistical Yearbook 2024, China Statistics Press, Beijing, 2025.
Shaoguan Municipal Bureau of Statistics: Shaoguan Statistical Yearbook 2023, China Statistics Press, Beijing, 2024.

Line 222 and elsewhere: Why was a factor of 0.8 (or 20%) chosen for this correction based on other studies? You actually calculate and report C_{ff}/C_{xs} from this study in section 3.2.

Response: We thank the reviewer for this constructive suggestion and agree that the rationale for the "20% correction" should be stated more clearly and consistently, and that our own C_{ff}/C_{xs} constraints (Sect. 3.2) should be explicitly linked to this step.

Purpose of the correction. Because CO is emitted almost exclusively from fossil-fuel combustion in urban environments, while C_{xs} (the CO_2 enhancement above background) can include both fossil-fuel and non-fossil contributions, we apply a correction to approximate $R_{CO/CO2ff}$ from ratios based on C_{xs} for those site–time subsets where paired $\Delta(^{14}CO_2)$-based C_{ff} is not available.

Why 0.8 (20%) was chosen. Our intention is not to suggest that a universal non-fossil fraction applies across all cities/seasons. Rather, the factor 0.8 represents a pragmatic first-order normalization for long-term R compilation, motivated by previous urban $\Delta(^{14}CO_2)$ studies showing that, under well-mixed daytime/afternoon conditions, the non-fossil contribution to C_{xs} is commonly in the range of ~10–30% (e.g., Turnbull et al., 2011; Lopez et al., 2013; Newman et al., 2016; Miller et al., 2020; see also Table E1). We therefore adopt 20% as a representative mid-range value and treat 10–30% as a plausible uncertainty range.

Connection to our Sect. 3.2 results. Importantly, our $\Delta(^{14}CO_2)$-based source separation (Sect. 3.2) yields C_{ff}/C_{xs} values that are broadly consistent with the above literature range but exhibit clear city/season variability. We therefore revised the manuscript to explicitly state that 0.8 is an approximation applied only when C_{ff} is not directly available from $\Delta(^{14}CO_2)$.

Revision implemented. In the revised manuscript we (i) explained the literature basis and the rationale for selecting 20% as a mid-range value, and (ii) explicitly cross-referenced Sect. 3.2 to show consistency with our observation-based constraints.

Corresponding manuscript change (Methods, Sect. 2.7; revised text with marked changes)

Old text:

"To correct for the contribution of non-fossil CO_2 in the observed enhancement, the concentration ratio $R_{CO/CO2ff}$ was estimated by dividing observed $R_{CO/CO2}$ by 0.8 for sites and times without $\Delta(^{14}CO_2)$ observations. Previous $\Delta(^{14}CO_2)$ and CO–CO_2 studies (Turnbull et al., 2011; Lopez et al., 2013; Newman et al., 2016; Miller et al., 2020) have shown that ~10–30 % (Table E1) of the total CO_2 enhancement above background during daytime is typically of non-fossil origin, while CO is emitted almost exclusively from fossil-fuel combustion. Thus, the 20 % correction represents a reasonable first-order approximation for well-mixed afternoon conditions."

Revised text (replace the old paragraph with the following):

To correct for the contribution of non-fossil CO_2 in the observed enhancement (C_{xs}), the concentration ratio $R_{CO/CO2ff}$ was estimated by dividing observed $R_{CO/CO2xs}$ by 0.8 for sites and times without $\Delta(^{14}CO_2)$ observations. *Equivalently, we assume C_{ff}/C_{xs} = 0.8 (i.e., 20% of C_{xs} is non-fossil), so that $R_{CO/CO2ff} = R_{CO/CO2xs} / (C_{ff}/C_{xs}) = R_{CO/CO2xs} / 0.8$ for those subsets.* Previous urban $\Delta(^{14}CO_2)$ studies (Turnbull et al., 2011; Lopez et al., 2013; Newman et al., 2016; Miller et al., 2020) have shown that ~10–30% (Table E1) of the total CO_2 enhancement above background during daytime/ afternoon is typically of non-fossil origin, while CO is emitted almost exclusively from fossil-fuel combustion. Thus, the 20% correction represents a reasonable first-order approximation for well-mixed afternoon conditions. *Our $\Delta(^{14}CO_2)$-based*

source separation (Sect. 3.2) provides city/season-dependent C_{ff}/C_{xs} constraints that are broadly consistent with this range.

Figure 4 (line 415): If we take the error bars into account, is this trend still robust and significant?

Response: Thank you for this important point. The apparent decrease from the winter 2010 value (34.9 ± 4.2 µmol mol^{-1}) to the winter 2022 value (12.5 ± 3.4 µmol mol^{-1}) is indeed substantial. When taking the error bars into account, the difference between the two winter means remains much larger than the combined 1σ uncertainty (added in quadrature), so the direction of change (lower wintertime C_{ff} in 2022 than in 2010 at GZ7) is robust.

At the same time, we note that the *magnitude* of this decrease is sensitive to the assumed $\Delta^{14}C$ background (NL air vs. NL tree-ring), as summarized in Table H2. Given the limited number of winter observations and this background sensitivity, we interpret the decrease as an indicative reduction rather than a statistically rigorous long-term trend. We have added explicit wording to this effect in Sect. 3.4.3 and in the caption of Fig. 4.

Line 430: The information in the paragraph is likely better presented in a table

Response: We appreciate the reviewer's constructive suggestions. The detailed information discussed in the paragraph at line 430 is already summarized in Table I1. To make this clearer and improve readability, we now explicitly refer to this table at the end of the paragraph ('… (Table I1)').".

Sincerely,
Gan Zhang
On behalf of all authors