

Response to RC1:

Zhang et al. present a novel framework for partitioning urban eddy covariance CO₂ fluxes into anthropogenic and biogenic components in mixed-source environments where conventional isotopic or multi-species approaches are infeasible. The method is applied to Guangzhou and demonstrates good consistency with independent tracers (NO_x, CO), wind-sector source area analyses, and traffic congestion patterns. Overall, the manuscript is well structured and clearly written, and it benefits from extensive sensitivity analyses. I only have a few minor comments that I hope will strengthen the paper.

Response:

We sincerely thank the reviewer for the positive assessment of our manuscript and for recognizing the novelty and overall quality of the study. We also appreciate the constructive and insightful suggestions, which have been very helpful in further improving the manuscript. Guided by these comments, we have revised the manuscript to clarify the methodological assumptions and applicability of the framework, strengthen the interpretation of the results, and more explicitly communicate relevant uncertainties and limitations. Below, we provide point-by-point responses to each comment and indicate the corresponding locations of the revisions in the revised manuscript. The revised text has been highlighted in the manuscript for ease of reference.

The selection of the CH–PY site pair is reasonably justified on the grounds of shared regional climate and similar evergreen broadleaf forest composition. However, the manuscript would benefit from a more explicit discussion of the ecological comparability between the two sites. This would help readers assess whether the methodology is transferable to other cities. In particular, urban environments often host vegetation types that differ markedly from those in surrounding natural areas, and urban soils may exhibit distinct respiration characteristics due to irrigation, management

practices, and other disturbances. Could these factors affect the general applicability of the approach to other urban settings?

Response:

We thank the reviewer for this insightful comment. We agree that ecological comparability between donor and target sites is a key prerequisite for transferring ecological response parameters and that the original manuscript did not sufficiently discuss the conditions and limitations for applying the framework to other urban environments.

In the revised manuscript, we have added a new Section 3.6 “Transferability and implications for urban carbon monitoring” to explicitly discuss this issue. We clarify that the transferability of the donor–target framework depends on ecological comparability rather than ecological identity between sites. In the present study, CH and PY share a humid subtropical climate and broadly similar evergreen vegetation phenology, while CH experiences substantially weaker anthropogenic disturbance and therefore provides a practical reference for identifying ecological response structures. Importantly, the framework does not directly transfer absolute GPP or ecosystem respiration (R_{eco}) fluxes from CH to PY. Instead, CH-derived ecological response parameters are normalized using footprint-weighted vegetation information (EVI_{fp}) and reconstructed at PY using PY-specific meteorological forcing and dynamic vegetation conditions.

We further added discussion on potential limitations for application to other urban areas. Specifically, differences in vegetation functional types, phenology, irrigation and management practices, and soil temperature, moisture, and substrate conditions may introduce additional uncertainties, particularly for R_{eco} , because satellite vegetation indices constrain photosynthetic activity more directly than below-canopy and soil respiration processes. We also clarify that applying the framework to other cities requires explicit donor-site selection and target-site constraints rather than direct reuse of the CH–PY parameter relationships. Where a single donor site cannot represent the ecological heterogeneity of an urban footprint, multiple donor sites, vegetation-specific parameterizations, or additional local ecological observations may be needed.

We also clarify that the framework has so far been evaluated only for the CH–PY site pair, and that its performance in other climatic and ecological settings requires local evaluation. These discussions have been added in Sect. 3.6 (**Pages 46–47, Lines 1241–1268**), and the corresponding statement in the Conclusions has been revised to emphasize that application to other urban sites depends on sufficient ecological comparability between donor and target sites (**Page 50, Lines 1334–1335**).

More generally, I would encourage the authors to offer broader guidance on how eddy covariance systems, when equipped with this fossil-fuel separation technique, could be deployed to advance our understanding of urban carbon dynamics. At present, while all the results appear reasonable, the approach provides only limited insight into what is not known or what is uncertain about the city-wide fossil-fuel CO₂ emissions, except for the inventory-based proxies used here are not applicable at fine spatial scales. A more explicit discussion of the potential and limitations of this technique for informing urban emission monitoring and policy would greatly enhance the paper's impact.

Response:

We appreciate the reviewer's suggestion to further clarify the broader implications and limitations of this framework for urban carbon monitoring. We agree that the original manuscript focused primarily on demonstrating the feasibility of EC-based fossil-fuel CO₂ separation, while the potential application of this approach within broader urban carbon monitoring systems required further discussion.

In the revised manuscript, we have expanded Section 3.6 “Transferability and implications for urban carbon monitoring” to discuss how EC observations combined with fossil-fuel separation can contribute to urban carbon assessments. We now emphasize that the main value of this framework is not to replace bottom-up inventories or city-scale atmospheric inversions, but to provide a footprint-scale observational constraint that separates anthropogenic release from biogenic exchange within dynamically varying urban source areas.

The revised discussion highlights several potential applications. First, the framework can help identify the timing and source-area dependence of anthropogenic

CO₂ signals and quantify the extent to which biological uptake masks urban fossil-fuel emissions in EC observations. Second, it provides an observational perspective for evaluating whether spatial proxies used in gridded emission inventories are consistent with local source-area conditions. Third, an EC network covering contrasting urban functional zones could help determine whether the temporal emission patterns observed at one tower also occur elsewhere in the city. Fourth, sustained observations across such a network could help track multi-year trends in anthropogenic emissions and net biogenic uptake, supporting assessment of emission-reduction and urban-greening measures. Such assessments would need to account for meteorological variability, changing footprints, and vegetation dynamics.

We have also clarified the limitations of the approach. A single EC tower samples only a subset of the urban landscape, and the derived F_{ff} represents a footprint-scale estimate rather than a direct estimate of city-wide fossil-fuel emissions. In addition, the current framework does not independently separate individual anthropogenic sectors such as traffic, buildings, and industry. Extending the framework to assess city-wide emission trends or mitigation outcomes would require integration with multiple EC sites, concentration networks, and atmospheric transport or inversion modelling.

These additions have been included in Section 3.6 (**Pages 47–48, Lines 1269–1290**). The revised discussion clarifies both the potential of EC-based fossil-fuel separation to support urban emission monitoring and mitigation assessment and the limitations of extending footprint-scale estimates to city-wide emission trends. The final sentence of the Conclusions has also been revised to highlight the relevance of multi-year trend monitoring to assessing local emission-reduction and urban-greening measures (**Page 50, Lines 1350–1352**).

In Section 3.4.3, the authors note that the PY footprint includes a substantial proportion of water bodies (~10 %), which are commonly treated as zero-emission areas in emission inventories but may contribute to the total flux. How different assumptions regarding the treatment of water fractions might influence the inventory comparison?

Response:

Thank you for this important comment. We agree that the treatment of water/wetland areas can affect the footprint-weighted inventory comparison and should be evaluated explicitly. We therefore added a sensitivity analysis that distinguishes two assumptions for aquatic pixels identified from the 10 m GLC_FCS10 land-cover product. The baseline and both sensitivity scenarios were recalculated using the same pixels with valid footprint, inventory, and land-cover data.

First, aquatic pixels were retained within the EC footprint but assigned zero direct anthropogenic CO₂ emissions, representing the common zero-emission treatment in anthropogenic inventories. This treatment preserves the contribution of these pixels to the averaging denominator. Under this assumption, the monthly footprint-weighted EDGAR, MEIC and ODIAC estimates decreased by mean values of 22.21 %, 22.09 % and 21.20 %, respectively. Nevertheless, the resulting monthly inventory ranges remained 5.30–9.70, 7.87–13.08 and 12.11–22.40 $\mu\text{mol m}^{-2} \text{s}^{-1}$, respectively, compared with a monthly mean PY F_{ff} range of 2.49–4.23 $\mu\text{mol m}^{-2} \text{s}^{-1}$. Thus, treating aquatic areas as zero anthropogenic-emission surfaces narrows, but does not remove, the inventory– F_{ff} difference.

Second, we excluded aquatic pixels and renormalized the remaining terrestrial footprint weights. This land-only treatment changed the mean EDGAR, MEIC and ODIAC estimates by only +0.10 %, +0.29 %, and +1.44 %, respectively. This treatment yields a terrestrial-only mean and is therefore a spatial-support sensitivity test rather than an equivalent estimate for the complete EC source area. The different responses arise because the zero-emission treatment retains aquatic pixels in the averaging denominator, whereas the land-only treatment excludes them and renormalizes the terrestrial weights.

We also clarify that the approximately 10 % water/wetland fraction reported in Sect. 2.1 is an area fraction within the fixed PY computational domain, rather than a footprint-weighted fraction. The inventory comparison uses month-specific footprint climatologies, so the contribution of water/wetland pixels varies among months and need not equal this domain-wide area fraction.

We note that actual air–water CO₂ exchange is physically distinct from the treatment of aquatic pixels in anthropogenic inventories. Any air–water exchange not represented by the ecological correction may remain in the inferred F_{ff} . Because this component was not independently measured or separated at PY, we acknowledge its unquantified potential contribution as a limitation of the present flux partitioning.

We have revised Sect. 3.4.3 of the main text (**Pages 38–39, Lines 1027–1040**) and Sect. S5 and Table S2 of the Supplement accordingly (**Supplement: Sect. S5, Pages 9–10, Lines 230–261**).

In Section 3.2.1, the temperature-stratified GPP scheme is found to outperform both the pure saturating and VPD-inhibition structures, and is consequently adopted as the baseline. This is an interesting finding. However, the discussion could usefully acknowledge that temperature and VPD are often correlated, making it difficult to isolate their individual effects. It is plausible that the temperature-stratified scheme may indirectly capture VPD influences; this point merits brief consideration.

Response:

Thank you for this helpful comment. We agree that the better performance of the temperature-stratified scheme should not be interpreted as evidence that temperature exerts an independent or stronger physiological control than VPD. Temperature and VPD commonly covary under daytime conditions, while VPD itself can regulate stomatal conductance and photosynthetic uptake (Leuning, 1995; Oren et al., 1999; Lasslop et al., 2010).

We have therefore revised Sect. 3.2.1 (**Pages 24–25, Lines 672–682**) to clarify that the temperature-stratified scheme may partly capture variation associated with VPD and atmospheric dryness. In this context, the near-zero fitted k_{VPD} in the explicit VPD-inhibition structure suggests that this particular term contributed little to the fit under the present dataset and model formulation; it does not imply that VPD effects were absent. Accordingly, the temperature-stratified scheme is retained as the baseline because of its better empirical fitting performance, whereas the pure saturating and VPD-inhibition structures remain as structural sensitivity scenarios.

Leuning, R.: A critical appraisal of a combined stomatal-photosynthesis model for C3 plants, *Plant Cell Environ.*, 18, 339–355, <https://doi.org/10.1111/j.1365-3040.1995.tb00370.x>, 1995.

Oren, R., Sperry, J. S., Katul, G. G., Pataki, D. E., Ewers, B. E., Phillips, N., and Schäfer, K. V. R.: Survey and synthesis of intra- and interspecific variation in stomatal sensitivity to vapour pressure deficit, *Plant Cell Environ.*, 22, 1515–1526, <https://doi.org/10.1046/j.1365-3040.1999.00513.x>, 1999.

Lasslop, G., Reichstein, M., Papale, D., Richardson, A. D., Arneth, A., Barr, A., Stoy, P., and Wohlfahrt, G.: Separation of net ecosystem exchange into assimilation and respiration using a light response curve approach: critical issues and global evaluation, *Global Change Biol.*, 16, 187–208, <https://doi.org/10.1111/j.1365-2486.2009.02041.x>, 2010.

Figure 6a: it would be helpful to include monthly variations in F_{cr} as well, as what is done in Fig. 5.

Response:

Thank you for this helpful suggestion. We have revised Fig. 6a to include the monthly variation in the corrected net CO₂ flux (F_{cr}), together with F_{ff} , GPP, and R_{eco} , thereby providing a more direct comparison between the observed net exchange and the partitioned flux components at the monthly scale, consistent with Fig. 5. The panel presents monthly medians of daily aggregated values, with error bars showing the interquartile range. The figure caption has been updated accordingly (**Page 30, Lines 814–817**).

We also added a brief statement in Sect. 3.3.2 (**Pages 29–30, Lines 806–808**) to clarify the monthly relationship between F_{cr} and F_{ff} : “Monthly F_{cr} was also positive but remained consistently lower than F_{ff} , reflecting the offset of anthropogenic and respiratory sources by photosynthetic uptake at the net-flux level.”

Figure 7: given that both CO and CO₂ exhibit high background concentrations, the comparison based on percentage changes may be misleading.

Response:

Thank you for this important comment. We agree that percentage changes in raw concentrations can be affected by the magnitude of the background concentration, particularly for CO and CO₂, and therefore may not provide a fair basis for cross-variable comparison.

We have replaced the percentage-change metric in Fig. 7c–d with standardized peak enhancement. Each variable was standardized using the mean and standard deviation of its own 24-point mean diurnal profile. Peak enhancement was then calculated as the difference between the mean z-score in the peak window and that in the corresponding pre-peak window. This is equivalent to dividing the peak–pre-peak difference in the original profile by that profile’s standard deviation, rather than by an absolute concentration level. A constant additive background changes neither the peak–pre-peak difference nor the standard deviation. The revised metric is therefore invariant to such a background offset and avoids its influence on percentage-based comparisons.

This treatment does not, however, constitute a physical subtraction of atmospheric background. Background variations with local hour, boundary-layer evolution, and atmospheric transport can still affect the mean diurnal profile and its standard deviation. The metric thus expresses peak enhancement relative to each variable’s own diurnal variability; it does not quantify background-corrected local emissions or provide a direct comparison of source strengths.

With the revised metric, NO_x and CO remain positively enhanced alongside F_{ff} during the morning and evening peak windows. We therefore interpret these results as supporting evidence of diurnal consistency. The metric is defined in the revised Fig. 7 caption, and its rationale and interpretive limits are clarified in Sect. 3.4.1 (**Pages 34–35, Lines 914–943**).

In addition, we have revised the wording in the first paragraph of Sect. 3.4.1 to distinguish the factors influencing CO₂ concentration from the additional processes affecting PM_{2.5}, avoiding a combined description that could incorrectly associate secondary aerosol formation with CO₂ (**Page 33, Lines 908–910**).

I am unclear on the purpose of showing the illustrative 100 % uncertainty range in Figure 10. Unless this serves a specific analytical or pedagogical role, I would suggest removing it or providing a clearer justification for its inclusion.

Response:

Thank you for this comment. We agree that the illustrative ± 100 % envelope did not represent an inventory-specific uncertainty estimate and could be misinterpreted as a formal confidence range. We have therefore removed this envelope from Fig. 10 and now show only the footprint-weighted EDGAR, MEIC, and ODIAC estimates together with the monthly means and interquartile ranges of the observation-inferred PY F_{ff} estimates (**Page 39, Line 1041**). The figure caption has been revised accordingly.

We also removed the corresponding discussion that relied on the illustrative ± 100 % range, including the statement that PY F_{ff} remained within these envelopes. The revised discussion in Sect. 3.4.3 retains relevant literature on scale-dependent inventory differences and possible spatial allocation biases, without assigning a uniform uncertainty bound to the three inventories. The comparison provides context for flux magnitudes and possible spatial allocation differences, rather than a formal validation of either the inventories or F_{ff} (**Pages 40–41, Lines 1071–1080**).

We have also corrected the PY F_{ff} range reported in Sect. 3.4.3 from 2.59–4.29 to 2.49–4.23 $\mu\text{mol m}^{-2} \text{s}^{-1}$ to match the monthly means shown in Fig. 10 (**Page 38, Line 1026**). This correction concerns only the range reported in the text; the PY F_{ff} data and their presentation in Fig. 10 remain unchanged. The previously reported range corresponds to the monthly medians of daily aggregated fluxes presented in Sect. 3.3.2 and Fig. 6. Sect. 3.4.3 now explicitly identifies the corrected range as monthly means, whereas Sect. 3.3.2 and Fig. 6 retain the monthly medians of daily aggregated fluxes. This numerical correction does not alter the conclusion that the footprint-weighted inventory estimates are substantially higher than PY F_{ff} .