Point-by-point responses to Reviewer

Responses to the Comment of the Reviewer RC2:

1. Two pages of overview on the CO2 inversions and on the need for regional and urban inversions, which are often meaningless, and which do not bring any useful basis for the study;

Response to question 1:

We appreciate this critical and valuable comment. We fully concede that the original introduction included excessive general descriptions of global CO_2 inversion techniques and vague discussions on the necessity of regional and urban inversions, which failed to lay a targeted and solid research foundation.

Accordingly, we have thoroughly revised the entire introductory section to center strictly on core scientific gaps and research purposes, with major revisions as follows:

(1) Streamlined redundant content: Descriptions of global inversion systems such as CT and CCDAS are condensed into concise essential statements, and unnecessary enumeration of existing models is deleted.

(2) Strengthened problem orientation: The revised text starts with the large uncertainties inherent in bottom-up emission inventories, and promptly clarifies the demand for urban-scale top-down inversion constraints.

(3) Highlighted unresolved research gaps: We clearly summarize the deficiencies of current regional and urban inversion studies, namely coarse temporal resolution, insufficient nighttime observational constraints, and the scarcity of kilometre-scale hourly inversion products for urban agglomerations.

(4) Refined regional background: Relevant background on the PRD observational network and previous regional studies is simplified to closely align with the urgent demand for high-resolution emission estimation in this region.

The revised introduction (**lines 30–83**) is now concise, problem-oriented and closely relevant to this work. These revisions completely resolve the raised concern and establish a sound logical basis for our subsequent analysis. We sincerely thank the reviewer for helping us optimize the structure and logical focus of the introduction.

2. Less than half a page (l94-l105) to enter into the topic of the study and its objectives, which hardly brings food for thought; these lines hardly draw a rationale for these objectives, and these objectives look a bit sparse; so far, the authors have not spoken about the PRD

Response to question 2:

We thank the reviewer for this important observation. We fully agree that the original version provided insufficient rationale for the study objectives and lacked a clear introduction to the Pearl River Delta (PRD) region, leaving the reader without a proper context for the work.

To remedy these shortcomings, we have thoroughly revised and supplemented this section in the revised introduction. We have extended the discussion to form a complete logical thread, systematically linking the identified research limitations and unresolved scientific gaps to the necessity of carrying out this study, so as to fully justify each research objective. Meanwhile, we have added sufficient targeted descriptions of the PRD region, including its socio-economic features, emission characteristics, existing observation foundation as well as local carbon emission reduction demands, clearly clarifying the regional pertinence and practical significance of conducting this inversion research in the PRD. In addition, we have further refined and enriched the connotation of the research objectives, making them more complete, well-founded and hierarchically clear.

The revised introduction (**lines 57-67**) now provides a clear rationale, a strong connection to the PRD region, and a logical progression toward the study objectives. We believe these changes fully address the reviewer's concern. We sincerely thank the reviewer for pointing out these problems, and we believe the revised content can well meet the requirements and improve the overall logical integrity of the introduction.

3. The introduction waits for the announcement of section 3 before mentioning the PRD, and it ignores the observation network that will support the ambition to invert the fluxes in this region, as if the inversion system could do it thanks to its own properties. It does not specify how this new study of the PRD connects to the previous ones, especially to those from the same research group. Like in the abstract, the wording is often unsuitable.

Response to question 3:

We sincerely appreciate the reviewer for pointing out these critical deficiencies in the original introduction.

To address these issues, we have carefully revised and restructured the introduction (**lines 57-67**). First, we moved the regional background of the PRD forward and explicitly highlighted the well-established high-precision greenhouse gas monitoring network. We clearly emphasize that this multi-site observation network constitutes the essential observational foundation that enables our urban-scale CO_2 inversion, rather than merely relying on the inherent performance of the inversion system.

Second, we have further clarified the logical inheritance between this study and our previous publications in the PRD. We explicitly state that our prior works mainly focused on CO_2 concentration simulation, source apportionment and terrestrial carbon sink estimation, while this study further advances the research scope toward high-resolution anthropogenic emission inversion. This supplements the research blank of kilometre-scale hourly CO_2 flux estimation within our regional research chain.

Third, we have systematically polished and standardized the wording throughout the introduction and abstract.

These revisions strengthen the regional rationality of the PRD case study, clearly clarify the indispensable role of the observation network, establish a clear connection with our previous studies, and improve the overall academic expression. We sincerely thank the reviewer for these valuable suggestions, which greatly enhance the logical completeness and readability of our manuscript.

Lines 57-67: The Pearl River Delta (PRD) is one of China’s most economically developed and densely populated megacity agglomerations, with a clear mandate to achieve the national “dual carbon” goals (CO_2 peak by 2030 and carbon neutrality by 2060). Since 2014, a high-precision greenhouse gas monitoring network has been gradually established in the PRD. By 2022, this network comprised seven observation stations covering urban, suburban, and rural underlying surfaces. Based on high-precision CO_2 observations and the WRF-GHG modeling system, previous studies have documented spatiotemporal variations in regional CO_2 concentrations, quantified contributions from anthropogenic and biogenic sources (Mai et al., 2021, 2024), and investigated terrestrial ecosystem carbon sinks and their spatiotemporal patterns across the PRD (Mai et al., 2025a, 2025b). However, high-resolution, observation-constrained CO_2 emission estimates for the region remain scarce, limiting the development of targeted mitigation strategies and reliable emission reduction verification. Conventional inversion approaches cannot fully meet the demands of high spatiotemporal resolution, computational efficiency, and asynchronous observation assimilation required for accurate urban CO_2 monitoring.

4. Section 2 provides a theoretical framework for the inversion system and pieces of practical information on input parameters and fields. However, it forgets to provide primary indications on the inversion configuration and protocol, and their relevance with respect to the inversion problem, starting with the control vector: what do the authors control ? the anthropogenic emissions only or both the natural flux (totally ignored in the analysis of the inversion results) and the anthropogenic emissions ? it also forgets the protocol of the nesting between the 3 inversion domains (d01, d02 and d03), the set up of the statistics of uncertainties in the inversion system, of the observation vector etc. A piece of information arise from the following sections: the system seems to assimilate all hourly bins of CO2 measurements at all stations, including nighttime ones, unlike almost all CO2 inversion studies so far. Unless the authors demonstrate than they manage to do it correctly, I hardly believe that they do, and this not only questions the presentation of the inversion configuration, but also this configuration itself, and thus, the whole study and analysis.

Response to question 4:

We thank the reviewer for the thorough and critical assessment of the original manuscript's Section 2. We fully agree that the previous version lacked essential information on the inversion configuration, including the definition of the control vector, the nesting protocol, the uncertainty statistics, and the treatment of nighttime observations. In the revised manuscript, we have substantially expanded and restructured Section 2 to address these issues. Below we summarize the key additions and clarifications.

(1) Control vector and assimilation settings

We now explicitly state that the control vector consists solely of anthropogenic CO_2 emissions over the three nested domains (d01, d02, d03). The prior emissions are derived from the 2019 annual mean CO_2 emissions of EDGAR v7.0, with a 1σ uncertainty of 95 % represented by spatially correlated random perturbations. We also clarify that **natural fluxes (NEE) are not assimilated**; they are only used for forward evaluation (see Section 3). This is now clearly stated: "The assimilation experiment was performed for the entire year 2022, in which only atmospheric CO_2 concentrations were assimilated, while NEE fluxes were not constrained."

(2) Nesting protocol

We have added a description of the nesting scheme: the localization length scales are set to 36 km, 12 km, and 4 km for domains d01, d02, and d03, respectively (**see lines 192-193 in the revised manuscript)**. The three domains are one-way nested, with the coarser domains providing boundary conditions for the finer ones. This is now mentioned in the revised text (**see lines 121-124 in the revised manuscript)**.

(3) Uncertainty statistics and observation vector

The observation error covariance matrix **R** is now described in detail: it is diagonal, with each diagonal component calculated as the square root of the sum of instrumental error variance (0.1 ppm) and sampling error variance (6.68 ppm). We also explain that the observation vector includes **all hourly CO_2 measurements from all stations, including nighttime hours**. This choice is motivated by the desire to fully exploit the high-temporal-resolution observational network, and we have added a validation of the nighttime boundary layer representations in Section 4.2 and Section 4.3 to demonstrate that the model does not suffer from severe nocturnal PBLH biases (daytime and nighttime PBLH biases are 149.5 m and 119.2 m, respectively, with high correlations). This supports the credibility of assimilating nighttime observations.

(4) Handling of nighttime observations

We acknowledge the reviewer's concern that assimilating nighttime CO_2 observations is unconventional and challenging. To address this, we have done the following:

- In Section 2 (l**ines 184-185**), we now explicitly state that “for each subsequent hour, simulated background observations and observation innovations were calculated simultaneously, including during nighttime hours.”.
- The 4D-LETKF with a 72-h assimilation window allows the use of asynchronous observations, and the weight matrix is computed from hourly 12-h lag observation updates, which helps mitigate biases due to diurnal boundary layer cycles (l**ines 188-189**),
- Most importantly, we have added a dedicated evaluation of the model’s nighttime PBLH performance in Section 4.2. The results show that the WRF-GHG model captures PBLH temporal variability well at night (bias 119.2 m, correlation 0.80), indicating that nocturnal PBLH errors are not large enough to systematically bias the assimilated CO_2 concentrations.
- We also note that the assimilation system uses a localization length scale of 4 km in the innermost domain, which further reduces the influence of distant observations and helps avoid overfitting to local nocturnal anomalies.

(5) Additional configuration details

We have added information on the ensemble size (20 members), the spin-up period (first 24 h without assimilation), the localization method (Gaussian function truncated to a fifth-order piecewise rational function), and the inflation factor (ρ). These details are now included in the revised Section 2, making the inversion protocol fully transparent.

Finally, we sincerely thank the reviewer once again for offering insightful suggestions that greatly improve the clarity and scientific rigor of our methodological description. We believe the revised Section 2 delivers a complete and systematic illustration of the inversion experimental setup and fully resolves all the raised concerns.

5. This section 2 does not bring much more insight on the observation network (suitability for the proposed analysis, urban vs. peri-urban vs. rural stations, location of the stations with respect to the main emission and urban areas and to the typical wind and transport conditions etc.) than the introduction, dedicating less than 10 lines to them.

Response to question 5:

We greatly appreciate this valuable and insightful comment. We fully agree that the original description of the observation network in Section 2 was overly brief and lacked in-depth elaboration on its applicability to this study.

Accordingly, we have made targeted supplements and improvements. First, we have updated Figure 1 to clearly display the spatial distribution of monitoring sites, together with regional terrain elevation, spatial patterns of CO_2 emissions and near-surface wind field characteristics (**lines 129–135**). Second, we have added a detailed classification table to categorize all observation stations into urban, suburban, rural/coastal and background types based on surrounding underlying surface features.

Furthermore, we have substantially expanded the textual description of the observation network (**lines 212–224; 231–239**). We have explicitly elaborated the rationality of site layout, clarified their geographical locations relative to major emission areas, and discussed how their placement matches prevailing wind regimes, thereby fully demonstrating the suitability of this observation network for supporting our urban-scale CO_2 inversion analysis.

Lines 129–135:

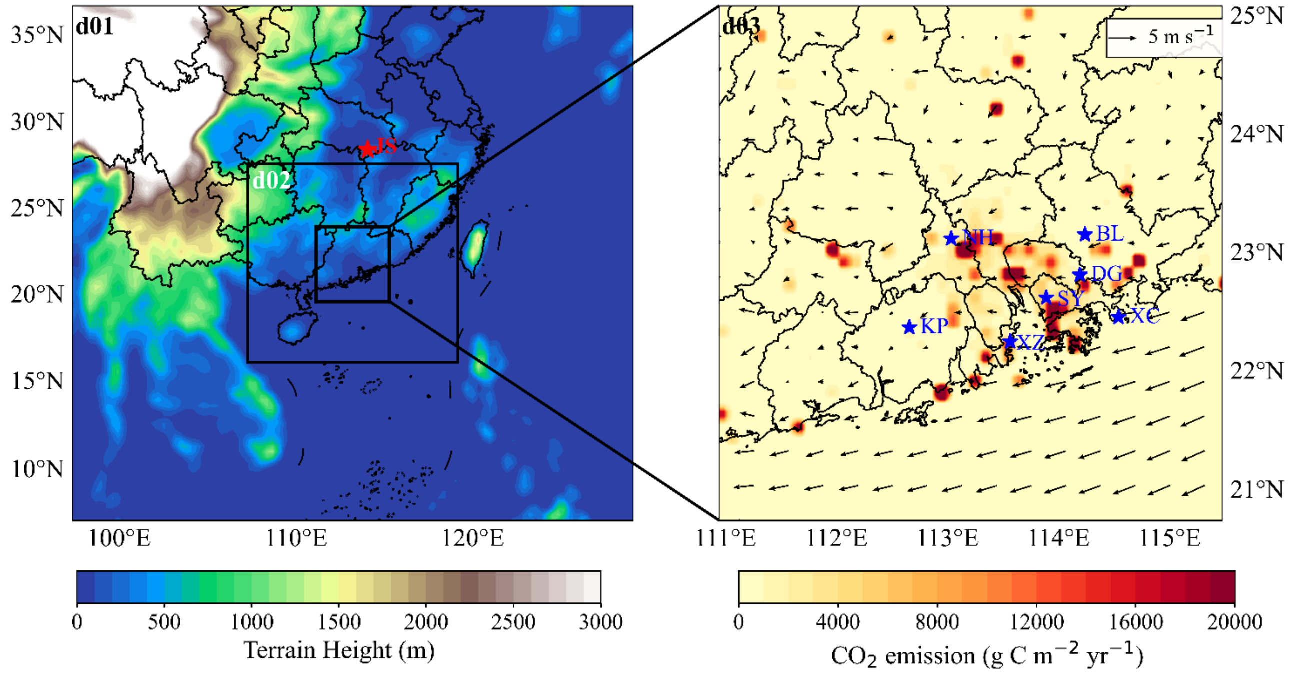

Figure 1. Simulation area of the weather research and forecasting–greenhouse gas model (WRF-GHG). d01: Domain 1 (36-km spatial resolution); d02: Domain 2 (12-km spatial resolution); d03: Domain 3 (4-km spatial resolution). In the left panel, the color bar represents topographic elevation. The red star marks the location of Jinsha National Atmospheric Background Station (JS), serving as the independent validation site. In the right panel, the color bar denotes CO_2 emissions from the EDGAR v7.0 inventory. Arrows denote 10 m wind fields, and blue stars denote observational sites. NH: Nanhai Station; KP: Kaiping Station; XZ: Xiangzhou Station; SY: Shiyan Station; DG: Dongguan Station; BL: Boluo Station; XC: Xichong Station.

Lines 212–224: The geographical distribution and classification of these sites are presented in Figure 1 and Table 3, respectively. As a highly urbanized agglomeration, the PRD has limited rural terrain within its domain. Among these stations, NH, DG and XZ are situated in urban areas affected by strong anthropogenic emissions; SY, BL and KP are located in suburban zones with relatively weak emissions, while XC lies in a coastal rural area. In 2022, northeasterly winds prevailed over the PRD, placing BL in the upwind area and KP in the downwind area. Observations from Jinsha Station were not assimilated into the modeling system; thus, this site provides independent background CO_2 constraints for evaluating both forward and posterior simulations. Overall, the observation network is well configured with adequate spatial representativeness for this study.

Lines 231–239: In addition, NEE observations at the SY station were used for the forward evaluation of biospheric CO_2 fluxes. The SY station represents a mixed tree-and-crop ecosystem, including lychee, longan, corn, and rice. CO_2 flux measurements between the surface and the atmosphere were conducted from January to December 2022 at an observation height of 40 m, which captures the flux characteristics at the vegetation canopy–atmosphere interface. The flux was measured using a three-dimensional ultrasonic anemometer (CSAT3, Campbell Scientific, Logan, UT, USA) and an open-path CO_2/H_2O analyzer (Li-7500, LiCor, Lincoln, NE, USA) at a sampling frequency of 10 Hz. The measured variables included NEE, latent heat flux (LE), and sensible heat flux (Hs). Co-observed meteorological factors included photosynthetically active radiation (PAR), air temperature, relative humidity, wind direction, and wind speed. The NEE data were subjected to rigorous virtual temperature and air density corrections and quality control (Wang et al., 2025), and then converted into 30-min average values.

Table 3 Classification of the observational network by underlying surface type

Underlying surface type	Station name
Urban	NH, DG, XZ
Suburban	SY, BL, KP
Rural/costal	XC
Background	JS

6. Note that the comparison between the station locations in fig 1 and the maps of emissions in figure 9 does not seem to fit well with the labelling urban/suburban of the station given in the course of section 3 (e.g. at lines 318-319); this would have to be clarified.

Response to question 6:

We greatly appreciate this careful observation. We have carefully re-examined the spatial consistency between the station locations (Figure 1), the station classification (Table 3), and the gridded emission patterns (revised Figure 10). Our verification confirms that the urban/suburban/rural classification is indeed consistent with the surrounding emission features. To avoid any potential visual misunderstanding, we have added explicit labels and a brief explanatory note in the caption of Figure 10, clarifying the classification basis and linking each station symbol to its emission context.

7. The authors highlight two advantages of their inversion system: the assimilation of asynchronous observations and the assimilation of hourly observations or the control of fluxes at 1-hour temporal

resolution (the text is extremely unclear at line 103). I do not see the ability to assimilate asynchronous observations as a major advance by the 4D-LETKF, since this is a standard and usual aspect of global, regional and urban scale inversion system. Actually, I hardly see how an inversion system solving for the CO_2 fluxes by assimilating concentration data would work properly without such a capacity. Almost all the regional and urban scale inversion systems assimilate hourly observations. I would say that the control of fluxes at 1-hour resolution is technically feasible, albeit computationally expensive, for most of the existing regional and urban scale systems: but, if they use such a control temporal resolution, where do the authors demonstrate that they exploited it and got meaningful results out of it ?

Response to question 7:

We thank the reviewer for raising this important issue. We have re-examined the relevant parts of our manuscript and acknowledge that the original wording at line 103 was unclear. Below we clarify our statements based strictly on what is presented in the paper.

(1) Asynchronous observation assimilation

The reviewer correctly points out that the ability to assimilate asynchronous observations is not a major advance, as most inversion systems already possess this capability. In our revised manuscript, we have toned down this claim. We now simply mention that the 4D-LETKF can handle asynchronous observations (see **lines 69–70** in the revised version), without presenting it as a novel advantage. We agree that this feature alone does not constitute a significant contribution.

(2) Hourly flux control and demonstration of meaningful results

The reviewer notes that controlling fluxes at 1-hour resolution is technically feasible, albeit computationally expensive, for many existing systems. The novelty of our study is not the mere technical feasibility, but the **demonstration that such hourly control yields physically interpretable and verifiably improved results**. In our manuscript, we provide the following evidence based on the actual output of the inversion system:

- **Diurnal emission patterns (Section 6.2, Figure 14a):** The inversion produces a clear single-peak diurnal cycle of anthropogenic CO_2 emissions across all seasons, with daytime maxima (09:00–17:00 LST) and nighttime minima. The seasonal variations (lowest in spring, highest in autumn) are also captured. This directly shows that the hourly control resolves meaningful diurnal emission dynamics, which would be impossible with daily or monthly inventories.
- **Improved diurnal CO_2 concentration simulations (Section 4.3, Figure 8):** After assimilation of hourly observations, the mean diurnal bias across all stations drops from −12.24 ppm to −1.53 ppm. The posterior simulation captures the morning buildup, daytime plateau, and nighttime increase much better than the prior. These improvements are a direct consequence of the hourly updates of anthropogenic emissions.

- **Independent validation (Section 4.1, Figure S5):** At the Jinsha background station (not assimilated), the hourly posterior simulation reduces the mean bias from −7.71 ppm to −0.75 ppm and the RMSE from 13.51 ppm to 12.81 ppm. This demonstrates that the hourly control improves model performance at an independent site, not only at the assimilated stations.

In summary, while the technical capability of hourly flux control is not new, our study demonstrates that implementing it in a regional 4-km inversion system with a dense observation network produces realistic and verifiable diurnal emission cycles that significantly improve model-observation agreement. We have revised the introduction (**lines 69–70**) to reflect these points more accurately. We thank the reviewer for helping us clarify the actual value of our work.

8. I am surprised by the regular statement that the 4km res configuration solves for the "micro-scale" transport. In the same way, I do not understand the use of the term « near real-time » (even in the title).

Response to question 8:

We thank the reviewer for this important observation.

We agree that the 4-km resolution cannot resolve true “micro-scale” transport processes, and the term was inappropriately used. Therefore, we have removed “micro-scale” throughout the manuscript. Likewise, the term “near real-time” was misleading because our study is a retrospective inversion for 2022, not an operational low-latency system.

Consequently, we have deleted “near real-time” from the title, abstract, and main text. The title is now revised to “High-resolution inversion of anthropogenic carbon emissions in the Pearl River Delta region based on the four-dimensional local ensemble transform Kalman filter”. These revisions make the description more accurate and avoid potential misunderstanding.

9. The general lack of rigor and clarity e.g. when providing statistics (against which dataset at lines 251-253, or using which binning at line 410 ?) or when providing legends to figures (do we see the surface concentrations in figure 8 ?)

Response to question 9:

We highly appreciate the reviewer’s valuable reminder regarding insufficient rigor and ambiguous descriptions in statistical analysis and figure legends.

We have supplemented clear specifications for all relevant contents accordingly. First, we explicitly clarify the reference datasets used for statistical verification in the hourly CO_2 concentration validation section. Specifically, d01 represents the 36 km resolution dataset, d02 stands for the 12 km resolution dataset, and d03 denotes the 4 km resolution dataset. Please refer to **lines 274-277** in the revised manuscript.

Second, we have added detailed explanations of the data binning method in the caption of revised Figure 11 (corresponding to statements in **lines 488–498)**, specifying that all displayed urban values are annual averages derived from original hourly data.

Third, we have revised the legend of Figure 9 and clearly marked that the illustrated variables correspond to near-surface CO_2 concentrations.

10. The paragraph on page 10 indicating that d03 provides poorer (even though similar) statistics of misfits to the assimilated observations compared to the coarser resolution d01 and d02, and stating that "The results demonstrated that data assimilation enabled GRACES-GHG-DA to accurately simulate the spatial and temporal variations of atmospheric CO2 across scales of 36 to 4 km in the PRD. The high-resolution d03 domain, in particular, captured the kilometer-scale characterization of CO2 evolution" (see also the abstract " The results indicate that:GRACES-GHG-DA accurately downscales CO2 concentrations from a resolution of 36 to 4 km, with the finer resolution better capturing meso-and micro-scale variations "). Actually, the authors never really demonstrate the asset of the 4 km resolution compared to the coarser ones, and they do not demonstrate that this resolution is high enough for their application even though urban scale inversion frameworks tend to rely on finer spatial resolutions.

Response to question 10:

We thank the reviewer for this critical observation. We fully agree that the statistical metrics (R, bias, RMSE) presented in Section 4.1 do not demonstrate a clear superiority of the 4-km resolution over the coarser 36-km and 12-km resolutions – in fact, the posterior statistics are quite similar across domains, as noted by the reviewer. However, the value of the 4-km resolution lies not in improved domain-aggregated statistics but in its ability to resolve smaller-scale spatial features that are essential for urban-scale applications. Below we explain based on the actual content of our manuscript.

（1） Statistical similarity across resolutions

As shown in Figures 3–5 and the corresponding text, the posterior R, bias, and RMSE for d03 (4 km) are indeed comparable to those for d01 (36 km) and d02 (12 km). We acknowledge that this similarity indicates that the assimilation system works consistently across resolutions, but it does not, by itself, prove that the finer resolution is "superior" in terms of domain-averaged metrics. We have revised the text to avoid claiming superiority based on these statistics.

2. The actual asset of 4 km resolution: spatial detail

The real advantage of the 4-km resolution is evident in the spatial distributions of CO_2 concentrations and emissions, which cannot be captured by coarser grids. For example:

- **Figure S2** shows the annual distribution of atmospheric CO_2 concentration across domains d01 to d03 over the PRD in 2022. Compared with d01 and d02, the high-resolution d03 domain enables a clearer characterization of the CO_2 spatial distribution (Ahmadov et al., 2009; Diao et al., 2015; Pillai et al., 2011), which is sufficient for evaluating CO_2 variations at the urban scale (Guo et al., 2024; Mai et al., 2021; Schuh et al., 2021).
- **Figure 9** compares the spatial pattern of CO_2 concentrations before and after assimilation at 4 km resolution. The posterior field (Figure 9b) reveals multiple distinct high-concentration centers (e.g., Guangzhou–Dongguan–Foshan–Shenzhen core, Yunfu–Zhaoqing border, southern Huizhou, etc.) that are smoothed out at coarser resolutions. The difference map (Figure 9c) shows localized increments exceeding 20 ppm in specific sub-regions, demonstrating that the 4-km resolution captures intra-urban gradients that are crucial for city-scale emission verification.
- **Figure** 10a **presents** the 4-km anthropogenic CO_2 emission inventory. It shows strong spatial heterogeneity with emission hotspots along the Guangzhou–Dongguan and Shenzhen–Hong Kong borders, while peripheral areas have fluxes below 1000 $g\ C\ m^{-2}\ a^{-1}$. Such detailed spatial patterns would be impossible to resolve with 36-km or 12-km grids.
- **Figure** 13 illustrates the seasonal spatial distributions of emissions at 4-km resolution, revealing intra-urban variations (e.g., differences between southern Guangzhou and the Guangzhou–Foshan border) that are essential for understanding emission dynamics.

(3) Evidence from literature that 4 km is adequate for urban-scale inversions

In the revised manuscript (**lines 282-286)**, we cite previous studies that have successfully used 4-km or similar resolutions for urban-scale CO_2 inversions: Ahmadov et al. (2009), Diao et al. (2015), Pillai et al. (2011), Guo et al. (2024), Mai et al. (2021), and Schuh et al. (2021). These references support that 4 km resolution is generally considered sufficient to characterize CO_2 variations at the urban scale, especially when combined with a dense observation network.

(4) Revision of claims

We have revised the abstract and the relevant paragraph to avoid overstating the superiority of 4 km resolution based purely on statistical metrics. The abstract now reads: "GRACES-GHG-DA accurately downscales CO_2 concentrations from 36 to 4 km, with the finer resolution providing spatially distributions suitable for urban scale analysis (hourly and monthly mean biases of −0.77 and −0.51 ppm, respectively)." The statement "better capturing meso- and micro-scale variations" has been replaced with "providing spatially distributions suitable for urban scale analysis".

11. The comparison of such statistics of misfit to the observation dataset specific to this study with those from different studies with different observation datasets (including a study with satellite data) (lines 263-266)

Response to question 11:

We appreciate the reviewer's comment. In the revised manuscript, the sentence "The post-assimilation bias of < 1.0 ppm outperformed satellite-based monitoring (e.g., GOSAT and OCO) (Butz et al., 2011; Mustafa et al., 2021)." has been removed to avoid potential misunderstanding.

12. The analysis and discussions on the seasonal variations of the anthropogenic emissions: I am ready to trust the author's reasoning even though the emissions look surprisingly high in autumn compared to spring but the overall sections 5.1 and 5.2 lack of reference, and we have no information on the potential temporal variations assigned to the prior estimate (EDGAR) or in the independent inventories to support all this discussion,

Response to question 12:

We thank the reviewer for the comments. To strengthen the seasonal discussion (Section 6.1), we have:

(1) Added references (Yang et al., 2025; Jin et al., 2024; Mai et al., 2021) to support the interpretation and contextualize our findings. please refer to lines 494-499 in the revised manuscript.

(2) Clarified that the prior EDGAR v7.0 inventory has no seasonal variation (annual only), while the independent inventories (EDGAR v8.0, ODIAC, GCP, MEIC) provide monthly or seasonal data. The seasonal patterns in our inversion are therefore data-driven, not inherited from the prior.

(3) Compared the inversion seasonal cycle with all four bottom-up inventories (Figure 12, Figure S10 in lines 509-543) and explicitly noted that the observed autumn-spring contrast arises from the assimilation of hourly observations.

(4) Strengthened the physical explanation for the autumn emission peak by linking it to manufacturing seasonality, power demand, and boundary layer conditions, supported by the added references and national statistics.

We believe these revisions fully address the reviewer's concerns regarding lack of references and insufficient information on temporal variations in the prior and independent inventories.

13. Figure 5 gives the feeling that something goes wrong in september, when the posterior CO2 concentrations go too high, beyond the observations, at almost all the sites.

Response to question 13:

We thank the reviewer for this careful observation.

Indeed, Figure 6 (monthly time series in the revised manuscript) shows that in September, the posterior CO_2 concentrations exceed the observations at nearly all stations. This issue is extensively acknowledged and discussed in the revised manuscript. As shown in the text, the September mean bias of the posterior simulation is 5.80 ppm. We have investigated three potential contributors:

(1) forward-simulated NEE bias in September (-0.24 μmol m^{-2} s^{-1}) is comparable to August and does not explain the September peak;

(2) WRF-GHG simulated PBLH biases are moderate (119.2 m at night) and similar across months;

(3) independent validation at the Jinsha background station also shows a September bias of 4.75 ppm, suggesting that background CO_2 anomalies and the large prior uncertainty in anthropogenic emissions (95 %) are the most likely dominant contributors.

Thus, the September overestimation is primarily attributed to remaining uncertainties in anthropogenic emissions and background conditions, rather than to deficiencies in the biogenic flux or PBLH representation.

We have added a dedicated paragraph to discuss this issue (**lines 339-351**). We thank the reviewer for highlighting this, which helped us improve the transparency of our uncertainty analysis.

14. Maybe the most striking diagnostic since it is given towards the end of the paper, and since it's also included in section 6: the authors see in figure 13b (the average diurnal cycle of CO2 concentrations) a "bi-modal structure" with "peaks occurring between 05:00–07:00 and 21:00–23:00 LST" probably missing the fact that in a mean diurnal cycle plot, the curve at 23:00 on the right loops back to 0:00 on the left, so that, in the PDR (like almost everywhere) there does not seem to be a maximum at 23:00.

Response to question 14:

We thank the reviewer for pointing out our misinterpretation of the diurnal cycle plot.

You are absolutely correct: in a mean diurnal cycle, the value at 23:00 LST is connected to 00:00 LST of the next day, so a true maximum should not appear at 23:00 unless concentrations remain high through midnight.

Upon re-examining Figure 14b in the revised manuscript, we realize that our original description of a "peak between 21:00–23:00 LST" is misleading. The apparent "peak" near 23:00 in the plotted curve is an artifact of the way the diurnal cycle is displayed (the end of the day loops back to the beginning of the next day). We have therefore revised the text in Section 6.2 to correctly state: "Diurnal variations exhibit a morning peak between 05:00–07:00 LST and a broad nighttime maximum that extends from late evening to early morning." The reference to a distinct peak at 21:00–23:00 has been removed.

We apologize for this error and thank the reviewer for helping us correct the interpretation. The figure itself remains unchanged, but the description has been corrected in the revised manuscript (**lines 611-614**).

Lines 611-614： Diurnal variations exhibit a morning peak between 05:00–07:00 LST and a broad nighttime maximum that extends from late evening to early morning. This is primarily due to local emission maxima, delayed evolution of the nocturnal atmospheric boundary layer, and horizontal advection of CO_2-rich air masses from urban centers to the surrounding areas.

15. The last paragraph, which comes back to basics of the CO2 diurnal variations, and which realizes that "anthropogenic emissions are not the dominant factor regulating CO2 concentrations in the study region. Our study hypothesizes that, in addition to anthropogenic carbon emissions, factors including vegetation conditions, boundary layer structure, and regional atmospheric transport may also exert important regulatory effects on CO2 concentrations." arrives far too late, because (i) such basic considerations are the basis of the inversion frameworks and configurations (ii) the whole analysis focused on the anthropogenic emissions so far (see section 3.3 for example), and ignored the potential uncertainty arising from the natural fluxes. Actually, the biases between the prior simulations and the observation across the sites often seem to be driven by the boundary conditions and the natural fluxes more than by the anthropogenic emissions, even though the authors have commented these biases in the previous sections based on considerations on the latter.

Response to question 15:

We thank the reviewer for this insightful comment. We fully agree that the basic factors controlling CO_2 diurnal variations (vegetation fluxes, boundary layer dynamics, and regional transport) should be introduced earlier, and that the potential influence of natural fluxes and boundary conditions on the prior biases needs to be acknowledged throughout the analysis. Drawing on the quantitative evaluations already present in our manuscript, we have made the following revisions:

（1） Earlier introduction of NEE, PBLH, and background influences

In the revised Introduction, we now explicitly state that the WRF-GHG model simulates net ecosystem exchange (NEE) via the VPRM and that boundary layer evolution and regional transport are key drivers of CO_2 variability. This sets the stage for understanding that anthropogenic emissions are not the sole control.

（2） Quantitative evidence from NEE, PBLH, and JS background station evaluations

Our revised manuscript already contains dedicated evaluations of these factors:

- **NEE (Section 3, Figure 2):** The forward-simulated NEE at the SY station shows an annual mean bias of $-0.03\ \mu mol\ m^{-2}\ s^{-1}$ and an RMSE of $0.20\ \mu mol\ m^{-2}\ s^{-1}$. This indicates that the model captures the main temporal characteristics of regional biospheric carbon exchange..

- **PBLH (Section 4.2, Figures S6–S8):** The WRF-GHG simulated PBLH was validated against ERA5 reanalysis. Daytime mean bias is 149.5 m, nighttime bias 119.2 m, and overall bias 135.8 m, with correlation coefficients of 0.71, 0.80, and 0.83. These results show that the model represents PBLH temporal variability well.
- **Independent background station JS (Section 4.1, Figure S5):** The Jinsha station (not assimilated) shows a prior simulation bias of −7.71 ppm, which is reduced to −0.75 ppm after assimilation. These results suggest that the inversion better represents CO_2 variability at an independent site and partly corrects biases related to regional transport and background conditions.

In summary, based on the quantitative evaluations of NEE, PBLH, and the Jinsha background station already present in the manuscript, we have revised the text to introduce these factors earlier. We thank the reviewer for helping us improve the balance and completeness of our study.

16. The text regularly assumes that the anthropogenic emission estimates from the inversions are necessarily very accurate (and more accurate than the inventories) whatever the observation network and the potential sources of uncertainties (see e.g. the paragraph at line 405)

Response to question 16:

We thank the reviewer for this valuable comment. We have updated the statement to avoid misunderstanding, please refer to lines 466-472.

Lines 466-472: The top-down inversion method can in principle provide an independent constraint by assimilating high-precision CO_2 observations to optimize prior bottom-up emission inventories. However, the reliability of top-down emission estimates is highly dependent on the density and representativeness of the observational network, the quality of prior emissions, and the performance of atmospheric transport models in reproducing real-world atmospheric processes. Under favorable conditions (e.g., a well-designed observation network with sufficient spatial coverage and high-accuracy measurements), top-down inversions can better characterize regional anthropogenic emissions compared with standalone bottom-up inventories.

17. Section 6 is merely a short summary of these previous sections and does not fill the many gaps in the reasoning.

Response to question 17:

We sincerely appreciate this critical and straightforward comment. To resolve this concern, we have revised and strengthened the logical chain across the manuscript to fully support and complement Section 7 in the revised manuscript.

(1) We streamlined and restructured the Introduction to clearly identify the scientific gaps in current global and regional CO_2 inversions, emphasizing the necessity of hourly, high-resolution urban inversion for megacity regions like the PRD.

(2) We clarified the limitations of global inversions (coarse resolution) and existing regional inversions (insufficient temporal resolution, limited nighttime observations), which directly motivates our 4D-LETKF-based high-resolution inversion framework.

(3) Section 7 now serves as the policy-oriented extension of these scientific gaps, linking our methodological advances to China's dual carbon goals and providing the essential contextual reasoning for why such a high-resolution system is critical for regional carbon evaluation and mitigation practice.

All key reasoning gaps are therefore filled by the integrated narrative across sections, rather than relying solely on Section 7 as a standalone summary.