

Reply to Reviewer 2

General comment:

I agree with all the comments provided by Referee #1 and would like to add the following comments:

I agree that methodology presented in the manuscript has a certain degree of novelty. However, I have several concerns, ranging from minor to major regarding publication of the manuscript in its current form.

Overall response:

We sincerely thank the reviewer for the constructive assessment of our manuscript, particularly recognizing the methodological novelty of this study. In revising the manuscript, we have carefully considered all the comments and have addressed the methodological and interpretative issues raised here. Our point-by-point responses below explain how each concern has been addressed and identify the corresponding changes made to the manuscript.

Comment 1:

1. Line 38-39: I am not fully convinced by the statement that a "high-resolution aviation emission inventory serves as a critical link connecting aviation activities with atmospheric environmental responses." The manuscript should provide a clearer justification for why such high spatial and temporal resolution is necessary and what specific scientific or practical advantages it offers. This point is particularly important because the conclusion states that "the inventory provides second-level temporal resolution and nationwide coverage, offering a basis for near-airport air quality assessment and for evaluating the assumptions underlying existing lower-resolution aviation emission inventories." I suggest that the Introduction be revised to better align with this more specific motivation. For example, Yim et al. (2015), which is cited in the manuscript, employed an atmospheric model with a horizontal resolution of approximately $4^{\circ} \times 5^{\circ}$. When an emission inventory with a resolution of $0.1^{\circ} \times 0.1^{\circ}$ is used in such a coarse-resolution model, much of the benefit of the fine spatial resolution may be lost through model-grid aggregation. Therefore, the manuscript should more clearly explain under what applications the proposed high-resolution

inventory provides meaningful added value. If the primary advantages of the inventory are improved estimation of total emissions and/or its applicability to regional- or local-scale air quality assessments, particularly near airports, I find these motivations reasonable. However, the current Introduction does not articulate these points sufficiently clearly and should be revised accordingly.

Response 1:

We fully agree that the original statement was overly broad, and we appreciate the critical point regarding model-grid aggregation, a $0.1^\circ \times 0.1^\circ$ high-resolution inventory indeed loses its spatial advantages when applied to coarse-grid global models like the $4^\circ \times 5^\circ$ grid used in Yim et al. (2015). Following the exact recommendations, we have thoroughly revised the Introduction to explicitly articulate the specific applications where our high-resolution inventory provides meaningful added value: (1) reducing aggregation biases in total emission estimates, and (2) supporting regional and near-airport local air quality assessments. We have also aligned the Conclusion to ensure perfect consistency with these specific motivations.

In Section 1, Paragraph 1 was revised to explicitly differentiate the spatial scales of emission impacts, establishing the foundation for high-resolution requirements, reading as follows:

"...Furthermore, they pose serious threats to public health by deteriorating air quality across multiple spatial scales (Yim et al., 2013). While emissions at cruise altitudes contribute to large-scale background air pollution and premature mortality globally (Barrett et al., 2010), low-altitude emissions during landing and take-off (LTO) operations directly degrade near-surface air quality in densely populated airport vicinity areas (Yim et al., 2015). Given the pronounced spatiotemporal heterogeneity of aviation activities across flight phases and altitudes, accurately quantifying aircraft emissions and clarifying their complex distributions serve as the scientific foundation for targeted emission control and environmental impact assessment (Lu et al., 2024)."

In Section 1, Paragraph 2 was revised to transition smoothly from these impact scales to the specific application domains where bottom-up, high-resolution modeling is strictly necessary, reading as follows:

"Aviation emission inventories provide the fundamental basis for quantifying emission loads,

evaluating mitigation strategies, and supporting atmospheric modeling (Simone et al., 2013; Stettler et al., 2013). Depending on data availability and target application scales, inventory compilation methods are generally categorized into top-down and bottom-up approaches. The top-down approach estimates aggregate emissions primarily from global or national jet fuel sales statistics (e.g., from the IEA). While computationally straightforward and sufficient for macro-level carbon accounting or coarse-grid global modeling, top-down inventories cannot resolve sub-grid spatial gradients or operational dynamics. Consequently, higher spatiotemporal resolution becomes essential when the objective shifts toward reducing aggregation biases in total emission estimates, characterizing emission hotspots along busy flight corridors, or supporting regional and local air quality assessments around airport terminal areas. To meet these demands, compilation methods have increasingly shifted toward the bottom-up approach, which estimates fuel consumption and emissions by aggregating individual flight activities across space and time, thereby providing both improved estimation fidelity and fine-grained spatiotemporal resolution (Cui et al., 2022)."

Furthermore, we refined the final paragraph of the Introduction to explicitly summarize how our dataset fulfills these precise demands:

"To address these limitations, this study constructs a national-scale aviation emission inventory for China's domestic civil aviation sector using 1-Hz QAR records from approximately 4.52 million flights operated in 2023. Synthesizing measured single-engine fuel flow with airborne thermodynamic parameters and reanalysis humidity profiles, we derive a four-dimensional inventory of CO₂, NO_x, SO₂, PM, CO, and HC emissions, and examine how input uncertainties propagate across QAR, ADS-B, and schedule-based frameworks. This observation-based dataset constrains activity and fuel-flow uncertainties in national emission accounting, resolves phase-level and terminal-area emission patterns relevant to regional and near-airport air-quality assessments, thereby establishing an empirical constraint for regional exposure assessments and evaluating legacy assumptions in lower-resolution models."

Finally, in Section 5 (Conclusion), we explicitly solidified these two core practical advantages to ensure full consistency with the Introduction, reading as follows:

"Ultimately, this dataset provides a robust empirical baseline for improving national total emission

estimates and offers a high-resolution physical foundation for regional and near-airport air-quality assessments."

Comment 2:

2. Lines 100 and 143-144: The manuscript states that the dataset covers approximately 4.52 million flights operating to and from civil airports across China in 2023 and later refers to a comprehensive dataset of civil aviation operations in mainland China. However, the exact spatial scope of the dataset is not sufficiently clear. From a data-coverage perspective, it appears that the inventory may be intended to represent aviation emissions over mainland China, but this should be stated explicitly. In particular, the authors should clarify how flight trajectories and emissions were treated at the geographical boundaries of the study domain. For example, for international flights departing from or arriving at airports in mainland China, were emissions included only while the aircraft was located within the defined mainland China domain, or were the full flight trajectories included? Similarly, how were flights crossing the domain boundary handled?

Response 2:

We fully agree that the exact geographical domain and flight inclusion criteria must be explicitly defined to avoid ambiguity. In the revised manuscript, we have clearly specified that the study domain is strictly bounded to purely domestic flights operating exclusively within mainland China. Flights with departure or arrival airports located outside mainland China, including all international flights as well as domestic flights involving Hong Kong, Macao, and Taiwan, were entirely excluded prior to the emission calculation. Consequently, because the retained dataset consists exclusively of domestic flights connecting two airports within mainland China, their full gate-to-gate trajectories were retained and computed in their entirety, meaning no artificial boundary clipping or partial-flight truncations were required during the emission quantification.

In Section 2.2.1, the text was revised to explicitly define the geographical boundary and the strict exclusion criteria for international and regional flights, reading as follows:

"The initial archive contained approximately 4.85 million flight records for the entire year of 2023. As this study aims to construct a civil aviation emission inventory for strictly domestic flights within mainland China, flights with departure or arrival airports outside mainland China, as well as flights involving Hong Kong, Macao, and Taiwan, were excluded. After flight-scope screening and the removal of invalid records, approximately 4.52 million valid domestic flight records were retained, accounting for 93.2% of the original archive (details of the initial data screening are provided in Table S1)."

Table S1. Summary of QAR Flight Screening and Data Retention

Data Processing Stage	Proportion of Raw Records (%)
Raw QAR records	100.0%
Flights beyond the study scope	6.2%
Domestic flight candidate records	93.8%
Invalid records	0.6%
Final valid domestic flights	93.2%

Comment 3:

3. Line 168: Does the term "dynamic environmental correction model" refer to the methodology described in Section 2.3.2, "Atmospheric Correction of Emission Indices"? If so, I suggest using consistent terminology throughout the manuscript to avoid potential confusion.

Response to Comment 3:

We confirm that the term originally referred to as the "dynamic environmental correction model" describes the exact same methodology detailed in the section for atmospheric corrections. To avoid any potential confusion and ensure clarity, we have standardized the terminology throughout the manuscript, ensuring that all references to this procedure consistently align with the specific title and framework of "Atmospheric Correction of Emission Indices".

In Section 2.3.3, the section heading was explicitly standardized to reflect the exact terminology, reading as follows:

"2.3.3 Atmospheric Correction of Emission Indices"

In Section 2.3.4, the subsequent text was aligned to consistently use this standardized terminology when referring back to the preceding step, reading as follows:

"Following the atmospheric correction of emission indices, pollutant emissions were quantified at the individual-engine level using the 1-Hz QAR observations."

Comment 4:

4. Line 214, Equation (5): Please check Equation (5). It appears that the exponent should be 0.51 and 1.65. It is currently unclear whether this is only a typo in the manuscript or whether the equation was implemented incorrectly in the actual calculations. This should be carefully verified, as it may directly affect the resulting emissions estimates.

Response 4:

We have verified our calculation code and confirm that this discrepancy was purely a typesetting error in the original manuscript, where the outer exponent of 0.5 in the BFFM2 pressure and temperature correction term was inadvertently formatted as a multiplication factor. In the actual emission calculations, the standard BFFM2 formulation with the correct mathematical scaling was properly implemented. In the revised manuscript, the NO_x atmospheric correction equation is now explicitly written in its standard form, retaining the outer exponent of 0.5 applied to the baseline parameters (which is mathematically equivalent to the expanded exponents of 0.51 and 1.65 noted by the reviewer). We have also eliminated the duplicate typographical error to ensure consistency. In Section 2.3.3, the NO_x atmospheric correction equation was explicitly corrected to accurately display the standard BFFM2 mathematical formulation, reading as follows:

"The NO_x emission index was corrected for both temperature and pressure effects and ambient humidity:

$$H = -19(\omega - 0.00634) \quad (5)$$

$$EI_{NO_x, flight} = EI_{NO_x, ref} \left(\frac{\delta_{amb}^{1.02}}{\theta_{amb}^{3.3}} \right)^{0.5} e^H \quad (6)$$

where ω is the ambient specific humidity."

Comment 5:

5. As also noted by RC #1, I believe that several parts of the manuscript go beyond what can be directly supported by the presented analysis. For example, the discussion beginning around Line 298 appears to attribute the differences between the emission inventories to specific causes that have not been explained in the study. While these explanations look plausible, they remain hypotheses unless authors apply the stated assumptions to the QAR data. Subsequently, the authors should reconstruct the emission inventory accordingly and then compare the resulting estimates with the CAAC inventory. Unless this statement may be removed from the document.

Response 5:

We fully agree that attributing discrepancies to specific operational factors remains speculative without dedicated sensitivity reconstructions. Following your suggestion, we removed all unsupported causal attributions from the manuscript. As detailed in our response to Reviewer 1, after strictly harmonizing the system boundary to purely domestic flights within mainland China, our estimated fuel consumption of 23.7 Mt yields a 5.3% difference relative to the CAAC statistical estimate of 25.03 Mt. In the revised manuscript, we now present the CAAC figure strictly as an independent macroscopic benchmark of overall magnitude. We limit our discussion to the factual observation that our 96.8% flight-count coverage does not directly translate to an identical fuel-burn coverage due to variations in aircraft type, stage length, and payload, deliberately avoiding any further speculative hypotheses.

In Section 3.1, the discussion comparing the QAR-derived estimates with the CAAC inventory was comprehensively revised to remove speculative causal attributions, reading as follows:

"To evaluate the macroscopic consistency of the inventory, the QAR-based fuel consumption was compared against official industry statistics released by the Civil Aviation Administration of China

(CAAC, 2024). Based on the annual transport turnover and average fuel efficiency reported by the CAAC, national domestic fuel consumption in 2023 is estimated at approximately 25.03 Mt, yielding a relative difference of 5.3% compared with our bottom-up summation. Here, the CAAC statistical value serves as an independent national-scale benchmark of overall magnitude. While the retained QAR archive captures approximately 96.8% of the commercial flight movements reported by the CAAC, flight-count coverage does not correspond directly to fuel-burn coverage because individual flights vary in aircraft type, stage length, and payload. Overall, the 5.3% difference indicates good agreement in magnitude between the integration of 1-Hz onboard fuel-flow measurements and top-down statistical accounting. The statistical distribution of major engine models and departure airports comprising this domestic fleet is detailed in Figure S2. Furthermore, monthly variations in flight activity and emission totals across the six species are presented in Figure S3."

Comment 6:

6. A similar issue appears in Lines 329-330. In this case, appropriate citations may help support "general" validity of the proposed explanation. However, there should still be a clear distinction between "general" statement and what has been presented by the results of the current manuscript.

Response 6:

We fully agree that cross-study comparisons must be supported by appropriate citations and that general domain knowledge should be clearly distinguished from the specific empirical findings of our study. Accordingly, we have comprehensively revised and condensed the comparative discussion in Section 3.1. We now explicitly discuss the observed magnitude differences across the three pollutant groups by referencing established operational and combustion mechanisms grounded in published literature, strictly avoiding unsupported speculation across external inventories. Meanwhile, quantitative claims regarding error propagation in this model are reserved strictly for the empirical Monte Carlo and Shapley analyses in Section 3.2.

In Section 3.1, the comparative discussion was revised to explicitly support all general domain

knowledge and cross-study explanations with robust literature citations, reading as follows:

"Building upon this validated baseline, meaningful divergences emerge in thermodynamically sensitive pollutants and incomplete combustion products (Table 4). Specifically, the estimated CO emissions (73.5 kt) are higher than previous full-flight values (39.7–59.6 kt), whereas NO_x emissions (311.84 kt) are substantially lower than trajectory-modeled upper estimates (up to 477.2 kt). Rather than mere methodological noise, these discrepancies highlight the inherent limitations of relying on idealized operating profiles. Schedule- and trajectory-based models often underrepresent prolonged low-power ground operations, such as extended taxiing and holding which disproportionately elevate CO generation due to incomplete combustion. Conversely, generic performance models frequently assume maximum rated thrust during takeoff and climb, leading to a systematic overestimation of peak combustor temperatures and thermal NO_x formation (Koudis et al., 2017). By coupling 1-Hz measured fuel flow with actual in-flight ambient temperature, pressure, and specific humidity, our QAR-based framework directly captures real-world transient operations (e.g., reduced-thrust takeoffs) and dynamic environmental effects. This approach transitions the emission estimates from model-parameterized approximations to a physically constrained operational baseline, with the corresponding uncertainty propagation formally evaluated in Section 3.2."

Comment 7

7. Is NO_x mass NO₂ equivalent or nitrogen mass?

Response 7:

Throughout this study, all reported NO_x emission indices and total NO_x emission masses are explicitly expressed in terms of NO₂ mass equivalent, strictly adhering to the standard conventions of ICAO Annex 16 Volume II and the ICAO Engine Emissions Databank (EEDB). To ensure there is no ambiguity for readers, we have added explicit clarifications to the manuscript text to formally define this metric.

In Section 3.1, the text was revised to explicitly state the NO₂ equivalent metric alongside the

reported totals, reading as follows:

"Under this system boundary, total annual fuel consumption is estimated at 23.7 Mt, corresponding to total emissions of 74.6 Mt for CO₂, 311.84 kt for NO_x (reported as NO₂ mass equivalent), 73.5 kt for CO, 6.02 kt for HC, 23.69 kt for SO₂, and 2.00 kt for PM."

Comment 8:

8. The dataset currently provided is incomplete. As also raised by the editor, the authors should either revise the manuscript to clearly reflect the actual scope of the available dataset or provide the complete dataset described in the manuscript.

Response 8:

We fully agree that the originally provided two-dimensional GeoTIFF files did not adequately represent the full four-dimensional structure of the inventory described in the manuscript. To rectify this, we have comprehensively reorganized the data products and released the complete 4-D gridded emission dataset, incorporating longitude, latitude, altitude, and temporal dimensions, via a finalized Zenodo record. Furthermore, we expanded the methodological documentation to explicitly detail how the high-frequency QAR observations are mathematically aggregated into this four-dimensional spatiotemporal grid.

In Section 2.3.4, the aggregation methodology was expanded to transparently define the four-dimensional grid structure, reading as follows:

"The second-by-second emission records were subsequently aggregated into a four-dimensional spatiotemporal emission grid. The horizontal domain was discretized at a resolution of 0.1°×0.1°, while the vertical dimension was divided into 1000-m altitude layers. The temporal dimension was aggregated by calendar month. For each 1-Hz QAR observation, the calculated emission mass was assigned to the corresponding horizontal grid cell, altitude layer, and calendar month based on the actual aircraft longitude, latitude, and pressure altitude. Emissions from all trajectory observations falling within the same spatiotemporal grid cell were then summed.

The resulting dataset contains monthly four-dimensional emission fields for CO₂, SO₂, NO_x, PM,

CO, and HC. This bottom-up aggregation retains the spatial and temporal distribution of individual flight trajectories while converting the high-frequency emission estimates into a standardized 4D emission inventory."

In Section 4, the Data Availability statement was updated to reflect the finalized 4-D dataset and the permanent repository link, reading as follows:

"The high-resolution 4D gridded aviation emission inventory dataset generated in this study is freely available at Zenodo via <https://doi.org/10.5281/zenodo.23334964> (Lu et al., 2026). The dataset is provided in NetCDF format with a horizontal spatial resolution of $0.1^{\circ} \times 0.1^{\circ}$ and a vertical resolution of 1000 m, including emission fields for CO₂, NO_x, SO₂, PM, CO, and HC."

Comment 9:

1. There are few instances throughout the manuscript where spaces are either missing or duplicated (maybe more than indicated below, please check!)

a. Space

Line 90, 100, 172, 295, 440, 467

b. Double space

Line 355, 442, 553

Response 9:

We have carefully proofread the entire revised manuscript and eliminated all instances of missing or duplicated spaces, including but not limited to the specific locations pointed out by the reviewer.

Comment 10:

2. Table 3 (also there is no space here)

Font size are different within the same table.

Tab 4 -> no space here again in Emission Index(NO_x/others)

Response 10:

We have unified the font sizes and corrected the spacing issues in Table 3 to ensure a standardized presentation. Regarding the original Table 4, please note that as part of the fundamental restructuring of our emission calculation methodology (specifically the revised emission-index parameterization described in Section 2.3.1), this table has been removed from the revised manuscript.

Furthermore, we have conducted a comprehensive proofreading of all remaining tables, figures, and captions across the entire manuscript to eliminate any instances of missing or duplicated spaces and to ensure rigorous academic consistency.

Comment 11:

3. As a reader, I found that there are several points where the terminology or intended reference has to be inferred from the surrounding context. I encourage the authors to review the manuscript with this in mind and define terms and references more explicitly where needed. For example, the "typical flight" denoted in the caption of Figure 5 presumably refers to the B737-800 flight described in Line 422.

Response 11:

We fully agree that readers should not have to infer intended references from the surrounding context. We have thoroughly reviewed the manuscript and made all ambiguous terms explicitly clear. Specifically, the vague reference to "a typical flight" has been updated to explicitly identify the aircraft model, ensuring rigorous and unambiguous presentation.

In Section 3.3, the text and corresponding figure caption were updated to explicitly identify the aircraft type, reading as follows:

"To examine the temporal variations in aircraft emissions under changing flight conditions, a representative B737-800 flight was selected to construct a second-by-second emission profile (Figure 5)."

"Figure 5: High-frequency emission profiles of a representative Boeing 737-800 flight."