
Response to Reviewer 2

General Reply

We thank Reviewer 2 for the thorough and insightful review of our manuscript, *"Air Pollution in The Upper Troposphere: Insights from In-Situ Airplane Measurements (1991–2018)"*. The review raises important questions about (1) dataset selection and comparability, (2) methodological rigor in analyzing trends, (3) clarity of model vs. observations comparisons, and (4) the broader context of existing literature. We have undertaken substantial revisions to address these comments. Below, we reply to each point in detail.

---

1. General Comments

1.1 "This paper tries to do too much with disparate data sets."

Comment: The manuscript includes datasets from multiple sources (older campaigns, more recent IAGOS flights, MOPITT satellite retrievals, and a model) without adequately explaining how they can legitimately be combined or compared.

Response:

Given the intensive industrial processes over the Eastern Asia for the past 30 years, what have happened in the upper troposphere over the North Pacific for the short-lived air pollutants CO? In order to solve this question, we have started with very good in-situ IAGOS data. During the processes of research, we found that more data are needed, from the 1990s, other in-situ measurements, satellite measurements, and modelling. So, the effort has grown in size. But the key question is the same: CO in the upper troposphere over the North Pacific.

- We recognize that combining multiple datasets with different spatial/temporal coverage requires careful justification. We have **restructured Section 2 ("Data and Methods")** to clearly explain the rationale and constraints for each dataset (e.g., NASA GTE campaigns, IAGOS, Mauna Loa, MOPITT, and IMS model outputs).

- In the **revised Introduction**, we articulate why each data source is included: older campaigns provide historical snapshots of CO in the upper troposphere, while IAGOS flights offer more recent, regular sampling. MOPITT is used to gain a broader spatial perspective, and the IMS model provides context for long-term chemical/dynamical processes.
- We have added two new **tables** in the revised manuscript to summarizing each dataset's temporal coverage, altitude range, and measurement uncertainties. We also discuss how data sparsity in the early years affects the significance of any "trend" conclusions.

1.2 "Some of the datasets are not referenced adequately."

Comment: Datasets need proper citation and methodological details, including referencing NASA GTE missions, IAGOS, MOPITT, etc.

Response:

- We have revised **Section 2** to include explicit references for GTE campaigns (PEM-West A/B, TRACE-P), IAGOS (e.g., Nédélec et al., 2015; Cohen et al., 2018), MOPITT retrieval algorithms (e.g., Worden et al., 2013), and NOAA/ESRL (now Global Monitoring Laboratory) for Mauna Loa.
- We added specific references where the GTE mission data and associated publications are described in detail (e.g., Bey et al., 2001; and NASA GTE mission websites).
- The data sources are also listed in the revised **References** section to ensure transparency.

1.3 "Authors do not explain why it is legitimate to compare older campaigns with more recent/denser datasets."

Comment: Are the older campaigns truly comparable to IAGOS flights, given their short duration and sparser coverage?

Response:

- We **clarified** in Section 2.1.2 that the NASA GTE campaigns (PEM-West, TRACE-P) are treated as **snapshots** representing specific months/years. We **do not** claim these data provide full annual means. Instead, we discuss them as indicative baselines for early 1990s/2000s conditions in the Pacific upper troposphere.

1.4 "It is not clear that some analysis/results, especially for trends, are robust."

Comment: Trend analysis is questionable due to limited data overlap, seasonality, and non-uniform temporal sampling.

Response:

- In the **revised Sections 3.1 – 3.2**, we now carefully describe our **trend-fitting approach**, including:
  1. Computing **standard errors** as 95% confidence intervals on the slope and intercept.
  2. Reporting **p-values** to indicate whether a trend is statistically significant.

1.5 "The model simulations do not go beyond 2003 ... limited value in the comparisons."

Comment: The IMS simulations end in 2003, do not incorporate changing emissions, and have minimal temporal overlap with the latest IAGOS and MOPITT data.

Response:

- We acknowledge in Section 2.4 that the **IMS model** used here is primarily a historical simulation with **constant anthropogenic emissions** (EDGAR-based) to 2003. Our original motivation was to assess long-term chemical/dynamical controls from the 1950s to early 2000s.
- In the **revised manuscript**, we now clarify that we use the IMS results **only** to illustrate (i) possible baseline patterns and (ii) sensitivity to climate/meteorology, not to derive 2010+ trends. We no longer attempt to linearly extrapolate model results beyond 2003.
- We have substantially **toned down** claims about direct model – MOPITT comparisons and discuss the model's overlap primarily with earlier aircraft datasets (Section 3.4).

1.6 "How can you compare MOPITT partial column retrievals with point measurements?"

Comment: MOPITT retrievals represent partial columns, while IAGOS and Mauna Loa are point or in situ measurements. Previous studies have addressed such comparisons; authors should follow that approach or reference it.

Response:

- In **Section 2.2**, we now describe the MOPITT retrieval characteristics (e.g., weighting functions, typical vertical resolution). We cite **Worden et al. (2013)** and **Deeter et al. (2019)** (or relevant references) for standard practices in validating partial column CO with in situ data.
- We limit comparisons of MOPITT vs. in situ to **broad monthly/seasonal means** and discuss their representativeness differences (Section 3.3). We also note that, for the upper troposphere, the MOPITT retrievals have reduced sensitivity and thus we present them **qualitatively**, rather than claiming a precise 1:1 match.
- We have include in a table showing the uncertainties associated with the MOPITT and the IAGOS data. We believe that MOPITT data is very good in obtaining a global estimates of the CO trends over the upper troposphere as most the upper troposphere are covered by in-site CO measurements. By comparing MOPITT with IAGOS we can calibrated the data measured from the MOPITT.

1.7 "The authors do not discuss their findings in the context of previous publications."

Comment: There is insufficient reference to the large body of work on CO trends in Asia, the Northern Hemisphere, and specifically in the upper troposphere over the Pacific.

Response:

- We revised the **Introduction** (Section 1) and the **Discussion** (Sections 3.3 – 3.4) to reference more publications, such as:
  - **Smoydzin and Hoor (2022)** on Asian contributions to Pacific UT CO.
  - **Wang et al. (2022)** for global tropospheric ozone trends (which also analyze IAGOS CO).

    - **Cohen et al. (2018)** for long-term IAGOS CO over northern mid-latitudes.
    - **Worden et al. (2013)** for MOPITT-based CO trend analyses.
- We have added a table to summarize previous studies. Our work aim to continue the great works done from previous studies.

1.8 "The paper is very long, has 14 figures, and there are typos."

Comment: The manuscript is lengthy and some figures are cluttered. Reviewer suggests a careful re-reading and possibly streamlining.

Response:

- We regret that the revised manuscript is now even longer, with inclusion of more tables and figures to validated the uncertainties of datasets used in this work as requested by the reviewers. However, we aim to make this work a good study to show the current status of the CO trends over the North Pacific upper troposphere from the best in-situ measurements done by the IAGOS and previous NASA GTE experiments.

1.9 "Overlap with Wang et al. (2024) submitted to another journal. How are these two papers complementary?"

Comment: Reviewer notes that [Wang et al., 2024] uses HYSPLIT for source attribution, yet HYSPLIT is only acknowledged (not used) in the present manuscript. There appears to be overlap in authorship and content.

Response:

- We clarified that [Wang et al., 2024] focuses specifically on **source attribution** for short episodes in 2012 – 2013 using HYSPLIT, while **this** manuscript focuses on **long-term** CO data (1991 – 2018). We found the pumping theory for the measurements in 2012-2013. With extensive measurements to 2018, the pumping processes are validated.
- We cite [Wang et al., 2024] only where relevant (e.g., to mention that HYSPLIT is a useful tool for back-trajectory analysis of specific events), but we do **not** use HYSPLIT in the present study. We emphasize the differences in **scope** and **temporal coverage**: the other paper addresses a narrow time window with detailed

transport modeling, whereas this paper examines multi-decadal CO changes and model comparisons up to 2003 (IMS) or 2018 (IAGOS, MOPITT).

---

2. Detailed Comments

2.1 Title and Affiliations

Comment:

1. The title should emphasize that the main focus is CO.
2. Affiliations are repeated inconsistently in the author list.

Response:

- We have revised the **title** and the abstract.

2.2 Abstract

Comment:

1. Remove overly general statements; be specific about the datasets, their coverage, and the meaning of "short-term fluctuations."
2. Clarify which MOPITT product and time period.
3. Explain why the IMS model ends at 2003, given that other measurements go beyond 2012–2023.

Response:

- We have completely revised the abstract.

2.3 Introduction

Comment:

1. Lines 18–20: "Complex interactions" are too broad; reduce or remove.

2. Provide references for UV-B, temperature dependence of CO+OH.
3. Satellite "emissions" vs. "retrievals."
4. Expand acronyms PGGM, IAGOS.
5. More references needed for older campaigns and CO emission inventories.
6. Clarify how the GTE campaigns are used, given their short duration.

Response:

- We **streamlined** the opening paragraphs, focusing on CO's role in atmospheric chemistry and referencing known CO–OH reaction temperature dependences (e.g., Sander et al., 2011) and UV-B influences on radical production.
- We replaced "satellite emissions" with "satellite retrievals" to be precise.
- Acronyms PGGM (Pacific Greenhouse Gases Measurement) and IAGOS (In-service Aircraft for a Global Observing System) are now spelled out at first use.
- We cite additional references for GTE campaigns (e.g., Bey et al., 2001), clarify they are short-lived measurement periods, and now label them as "episodic snapshots."

2.4 Data and Methods

Section 2.1.2 and 2.1.3 (Aircraft, Mauna Loa)

Comment:

1. Why mention CO2 at Mauna Loa?
2. Provide altitude and references for Mauna Loa CO measurements.
3. NOAA ESRL now GML.
4. Don't call it "verifying satellite measurements."

Response:

- Mauna Loa data has good data quality for validating the MOPITT data and the IMS model results over the North Pacific lower troposphere.
- Added **Mauna Loa** coordinates (19.54° N, 155.58° W) and altitude (~3397 m), plus the calibration method references (e.g., Novelli et al., 1998; or updated NOAA website references).
- Replaced "verifying" with "evaluating" regarding satellite retrievals.
- Used "NOAA Global Monitoring Laboratory (GML)" consistently.

Section 2.2 (MOPITT)

Comment:
Specify which MOPITT version, sensitivity, and partial column retrieval altitudes.

Response:

- We now state we use, for example, **MOPITT Version 8** (TIR/NIR retrievals), valid from 2000 onward. We describe weighting functions and typical degrees of freedom for signal in the lower vs. upper troposphere.
- We have also added references: Worden et al. (2013), Deeter et al. (2019).

Section 2.3 (Anthropogenic Emissions)

Comment:
Which EDGAR version? Provide references and justify how accurate it is for 1970 – 2020.

Response:

We specify **EDGARv6 used in this work:** Emissions Database for Global Atmospheric Research (EDGAR), release EDGAR v6.1 (1970 - 2018) of May 2022,
European Commission, Joint Research Centre (EC-JRC)/Netherlands Environmental Assessment Agency (PBL), \url{https://edgar.jrc.ec.europa.eu/index.php/dataset_ap61}, 2022.

Section 2.4 (IMS Model)

Comment:

1. Provide the CTM resolution, meteorological fields, chemistry scheme, etc.
2. Why end in 2003?
3. Why not account for changing emissions?

Response:

- We updated the text to say: IMS uses **48x40 grid resolution**, 19 vertical layers (surface to ~10 hPa), driven by **NCEP reanalysis** (Kalnay et al., 1996), with a complete tropospheric chemistry.
- The model was configured for **historical runs** (1950s – 2003) with largely **constant** anthropogenic emissions after ~1990. We clarify in the revised text (Section 2.4) that this approach cannot capture post-2003 emission changes and is thus used to

explore pre-2003 variability and general dynamical processes, and to compare with Mauna Loa measurements.

2.5 Results (Sections 3.1 – 3.4)

3.1 "Long-term Time Series"

Comment:

1. The term "long-term" is questionable for these composite campaign data.
2. Clarify how min, max, and quartiles are derived.
3. Explain how trends are computed with such irregular coverage.

Response:

- Long-term means 27 years (1991-2018) of CO trends over the North Pacific upper troposphere.
- We specify that the 25th, 50th, and 75th percentiles are calculated from each month's dataset.
- Our "trend" fits are strictly limited to intervals with sufficient data density (e.g., post-1994 IAGOS). Where data are too sparse, we only mention "possible tendencies" rather than formal trends. We have included 95% confidence intervals of calculated trends in the revised manuscript.

3.2 Trend Computations

Comment:

1. Clarify altitude ranges for UT vs. LT.
2. Figure 4 is cluttered.
3. Slope of 0.02 ppbv/yr is basically no trend.

Response:

- In **Section 2.1**, we define UT as ~8 – 12 km (depending on tropopause height) and LT as ~0 – 3 km or 0 – 2 km, depending on data coverage.
- We re-labeled subplots (a), (b), (c), etc. in Figure 4 and increased font sizes. We also corrected missing "+" signs in the regression equations.

- We have included 95% confidence intervals for the CO trends in the figures.

3.3 Annual Profiles and Short Periods (2012 – 2018)

Comment:

1. Are 7-year trends (2012 – 2018) reliable?
2. Could other factors (e.g., ENSO, fires) cause interannual variability?

Response:

- We have included 95% confidence of the CO trends in the figures.
- We added references to known interannual drivers (e.g., wildfires, meteorological variability) in Section 3.3. The Mauna Loa data does indeed show the ENSO/fires over the tropical southeast Asia has impact on the CO measurements over the Mauna Loa. We have compared trends from Mauna Loa measurements and the IMS model. The 95% confidence intervals in the CO trends from the model are within the 95% confidence intervals of the CO trends from Mauna Loa measurements.

3.4 Model Comparisons

Comment:

1. Why end model in 2003 and then compare with data up to 2018?
2. Emissions are constant, so how can the model inform "long-term trends?"
3. Provide details on initial conditions, meteorological fields.

Response:

- As noted, we now **limit** the comparison primarily to the **overlapping period** (pre-2003) for NASA GTE campaigns, or to highlight broad climatological differences (Section 3.4).
- The chemical lifetime for CO is short, about weeks. Hence, integration of a full three-dimensional chemistry model will exhibit the effect of photochemistry controlling the variations of short-lived chemical species such as CO over 30 years (1948-1978) and 20 years (1984-2003) of continuous integration. The results help us understand the long-term trends of short-lived species CO from observations. The model integration like a control run with constant emissions. These controlled

runs are then compared with real-word observations which contains varied CO emissions. The results are discussed in Figure 12 (now new Figure 16).

- IMS initialization from early 1950s uses NCEP reanalysis fields. We have included this description in the section of the IMS model.

2.6 Figures 7, 10, and Others

Comment:

1. Plot definitions not always clear.
2. MOPITT partial columns vs. in situ point measurements.
3. Hard to see data points and lines in Figures 10 – 13.

Response:

- We have included 95% confidence intervals of CO trends in the revised figures.
- We have revised Figures 7/10.
- As shown in Figure 4 (now new Figure 7), the upper troposphere are calculated at altitudes between 8.25 and 14.25 km, while the lower troposphere are calculated at altitudes between 0.75 and 2.25 km.

Conclusion

We sincerely thank Reviewer 2 for bringing these critical points to our attention. In summary, we have:

1. **Clarified the scope**: specifying which data are used for what purpose and the limitations of combining older and newer measurements.
2. **Expanded methodological detail**: describing our approach to trend fitting, deseasonalization, significance testing, and acknowledging uncertainties.
3. **Improved references**: citing relevant studies (Smoydzin and Hoor, 2022; Wang et al., 2022; etc.) for context.
4. **Restructured figures and text** to reduce clutter, clarify legends, and properly present comparisons between MOPITT, IAGOS, Mauna Loa, and IMS.
5. **Reduced overstatements** about model-based conclusions post-2003, acknowledging limitations of constant emissions in IMS.

These revisions should make the manuscript more transparent, rigorous, and aligned with prior literature. We trust that the updated version will address Reviewer 2’ s concerns and improve the scientific clarity and overall quality of our paper.

Thank you again for your detailed review and constructive suggestions.