Response to Reviewer 3

General Reply
We thank Reviewer 3 for the careful reading of our manuscript, *"Air Pollution in the Upper Troposphere: Insights from In-Situ Airplane Measurements (1991–2018)"*, and for providing constructive and detailed feedback. Below, we offer a point-by-point response clarifying the changes we have made (or will make) in the manuscript. The reviewer's suggestions have significantly helped us improve the clarity, consistency, and overall scientific rigor of our work.

A complete replies to the reviewer's comments point by point will be uploaded in the next two weeks.

---

1. Overall Structure, Clarity, and Key Results

Reviewer Comment:
The manuscript is difficult to follow, the results are not always clear, and the main findings should be more explicitly stated in the abstract and conclusions. The paper would benefit from more concise organization, improved figures, and a deeper discussion/interpretation of the data.

Response:

1. **Abstract & Conclusions**: We have **rewritten both the abstract and the conclusions** to present the main findings concisely:
   - We have revised the conclusion.
   - We have completely revised the abstract.
2. **Reorganized Paper Structure**:
   - We have strengthen the paper with detailed data suggested by the reviewers.
3. **Improving Figures**: We have reduced the number of main-text figures by combining or moving some to the **Supplement**. The remaining figures have **larger labels**, clearer legends, and consistent sub-figure labeling (e.g., (a), (b), etc.). We also ensure each figure caption explicitly states:
   - The **time periods** included,

- The **symbols/lines** (e.g., what each color or marker indicates),
- The **statistical** or sampling details, if applicable (e.g., monthly medians vs. campaign averages). We have included 95% confidence intervals in the figure.

We believe these changes significantly enhance the manuscript' s readability and coherence.

---

2. Title Specificity

Reviewer Comment:
The title implies a general view on upper tropospheric air pollution, yet the focus is really on CO.

Response:
We have revised the title.

"Long-Term Trends of Carbon Monoxide Over the North Pacific Upper Troposphere From In-Situ Airplane Measurements 1991-2018"

This revision clarifies that we focus primarily on CO, while we do occasionally refer to O3_33 and H2_22O for supporting context.

---

3. Relationship to Wang et al. (2024)

Reviewer Comment:
Is this manuscript an extension/follow-up of Wang et al. (2024)? Clarify how the two papers differ.

Response:

- In the **Introduction**, we now explicitly note that Wangetal.,2024Wang et al., 2024Wangetal.,2024 (under review) focuses on **short-term back-trajectory analyses** using HYSPLIT to identify pollution origins during a subset of flights in 2012–2013. That study is distinct from **our current manuscript**, which examines **multi-**

decadal changes (1991 – 2018) using multiple datasets (e.g., GTE campaigns, IAGOS, MOPITT, IMS). We include a parenthetical "(under review)" in references to make clear the status of [Wang et al., 2024].

---

4. Consistency of Time Frames

Reviewer Comment:
Various periods are introduced (e.g., 2012 – 2023, 1991 – 2019, 2001 – 2018, 2004 – 2022…), which is confusing.

Response:

- We have included an infographics to show the time-series of the dataset used in this work.

---

5. Clarifying the Impact of Chemistry on CO

Reviewer Comment:
Please summarize the main chemical processes influencing CO profiles and trends and specify what "significant impact of chemistry" means.

Response:

- In the **Introduction**, we added a short paragraph describing CO's atmospheric lifetime (weeks to months), its oxidation by the OH radical, and how photochemistry can shape upper tropospheric CO.
- We also highlight that **vertical transport** and **chemical destruction/production** (via oxidation of hydrocarbons, etc.) can significantly alter CO concentrations aloft, particularly when outflow from strong convection or intrusions of stratospheric air occur.
- When we say "significant impact of chemistry on CO profiles," we now explain that we refer to the roles of **OH** availability and **photochemical production** from VOC precursors, both of which can drive region-specific trends.

---

6. Introduction Organization

Reviewer Comment:
The introduction contains preliminary results from the current manuscript as if they were from another source. Clarify references and expand the background on CO’ s role, sources, and known trends.

Response:

- We **reorganized** the Introduction:
    1. A new paragraph on **CO’ s significance**: sources, lifetime, and chemical role in tropospheric O3_33 formation (lines 23 – 3323 – 3323 – 33).
    2. A subsequent paragraph **reviews prior studies**, focusing on major CO trend results from satellite and in situ measurements. We added references to, e.g., **Cohen et al. (2018)**, **Gaudel et al. (2020)**, **Worden et al. (2013)**, and **Smoydzin and Hoor (2022)**.
    3. Then we **state the scope** of our study (lines 45 – 5545 – 5545 – 55) without mixing current data results into the introduction.

We have included a table to summarize previous studies on the studies of CO trends in the troposphere.

---

7. Data and Methods: Details, Precision, and Processing

Reviewer Comment:
The data descriptions are insufficient (instruments, sampling frequency, quality checks) and there is no explicit “Methodology” section. More thorough explanations are needed about how data were averaged, how UT is defined, how satellite data are compared to in situ, etc.

Response:

1. **Data Description (Section 2)**:
    - We expanded each subsection (2.1 for aircraft, 2.2 for MOPITT, 2.3 for Mauna Loa, etc.) to include:
        - **Instrument type** and **measurement principle** (e.g., IAGOS uses a COmeasuringdeviceCO measuring deviceCOmeasuringdevice with ±2 – 5 ppbv accuracy).
        - **Sampling frequency** (IAGOS typically logs measurements every few seconds/minutes during cruise).

        - **Quality control** references (e.g., Petzold et al., 2015, Nédélec et al., 2015).
2. **New Methodology (Section 3)**:
    - **Study Region**: We define latitudinal/longitudinal bands (with a new Table listing them).
    - **Definition of the Upper Troposphere (UT)**: We use a pressure/altitude threshold (e.g., altitude >8 – 9 km or pressure <300 hPa) to label data as UT. We specify the typical cruise altitudes for IAGOS flights (~10 – 12\sim10 – 12~10 – 12 km).
    - **Data Processing**: We describe how we **bin** data monthly or seasonally, compute **percentiles** (25%, 50%, 75%), and note the removal of outliers or flagged values.
    - **Trend Analysis**: We have included 95% confidence intervals in the trend analysis.
    - **Comparison with MOPITT**: We now clarify that MOPITT partial-column retrievals are interpolated to a coarse vertical grid, and we focus on the "upper-tropospheric layer" retrieval product. We discuss averaging kernels and acknowledge that mismatch may arise from representativeness differences (footprint vs. point in situ).

These additions ensure readers can understand precisely how each dataset is used and how we handle uncertainties. We have also included a table showing the uncertainties, methods, measurement periods associated with each dataset used in this work.

---

8. Figures (2, 3, 4, 5, etc.)

Reviewer Comment:
Many figures are crowded; legends and sub-figure labels are missing or unclear. Trend lines are sometimes in mismatched colors, and there is no clear explanation of symbols or time periods.

Response:

- We have **redesigned** Figures 2 – 5 with larger panels, sub-figure labels (a), (b), (c), etc. in the **top-left corners**, and more concise figure captions.
- **Legends** now explicitly state:
    - What each symbol/line color represents (e.g., red = lower troposphere, blue = upper troposphere).

    - The regression equations (with confidence intervals) when plotted.
- In **Figure 2**, we simplified the approach by showing monthly/seasonal medians with **shaded interquartile ranges** rather than multiple percentile lines (25%, 50%, 75%, etc.) as separate sub-panels.
- Where multiple datasets overlap, we use a single multi-panel figure with consistent coloring for each dataset (e.g., IAGOS in black, older NASA GTE in gray, MOPITT in green, IMS in dashed lines, etc.).

Below are some specific improvements:

1. **Figure 2**: We have revised this figure to new Figure 3..
2. **Figure 3**: We have revised this figure to new Figure 6.
3. **Figure 4**: We have revised this figure to new Figure 7 and Figure 8.
4. **Figure 5**: We have revised this figure to new Figure 9.

---

9. Statistical Significance and Error Bars

Reviewer Comment:
Trends are shown without error bars, and p-values often indicate no significant correlation. The text sometimes draws conclusions about positive/negative trends despite R or p suggesting otherwise.

Response:

- **Error Bars**: We have included 95% confidence intervals in the figures and in the discussions.

---

10. Vertical Pumping and Negative Correlations (CO vs. O3, etc.)

Reviewer Comment:
More explanation is needed for why negative correlation among CO, O3_33, and H2_22O implies vertical pumping or other processes. Expand the discussion, especially for regional differences.

Response:

- We added a short paragraph in Section 3.23.23.2 explaining that:

- **Negative CO – O3_33 correlation** in the upper troposphere can arise when stratosphere-influenced air (high O3_33, low CO) and tropospheric outflow (high CO, lower O3_33) mix.
- The term "vertical pumping" refers to deep convection or monsoon circulation carrying boundary-layer CO to higher altitudes. Concomitantly, O3_33 can be titrated or replaced by fresh emissions.
- **Regional Contrasts** (e.g., Western Europe vs. North Pacific) may stem from different meteorological regimes, anthropogenic sources, or chemical environments.

---

11. Comparisons with MOPITT and IMS Model

Reviewer Comment:
More details on how MOPITT data are prepared, what altitude level is used, and how biases might arise. The IMS model altitude range must be stated explicitly. The model's representativeness vs. in situ data and other available models (GEOS-Chem, etc.) should be discussed.

Response:

- **MOPITT:** In Section 3.1–3.23.1–3.23.1–3.2, we now describe that we primarily use MOPITT's TIR retrievals for the **upper troposphere** layer, acknowledging the partial-column averaging kernel and potential smoothing. We do not expect an exact 1:1 match with IAGOS, so we only compare large-scale seasonal/annual patterns and trend directions.
- **IMS Model:**
  - We specify the vertical levels relevant for UT (pressure <300 hPa).
  - We emphasize that IMS runs were limited to 2003 with fixed anthropogenic emissions, so direct comparisons after 2003 are more of a conceptual check on model skill.
  - We added references to other models (e.g., GEOS-Chem) that also handle global CO, clarifying that IMS is one modeling system among several, and that our choice was motivated by pastreferencespast referencespastreferences as well as existing expertise with IMS.

---

12. Further Discussion and Flow

Reviewer Comment:
Section 3.4 lumps together comparisons of MOPITT, in situ, and the IMS model, then abruptly moves to CO budget from IMS. A clearer linkage is needed.

Response:

- We restructured **Section 4** (formerly 3.4) into clearer sub-sections:
  - **4.1 Model – Satellite Overlap (pre-2003)**: acknowledging limited overlap but showing broad patterns.
  - **4.2 MOPITT vs. IAGOS**: discussing partial column vs. in situ.
  - **4.3 IMS-Based CO Budget**: explaining the historical perspective from 1948 – 2003, including how convective transport or emissions changes might shape global patterns.
- Transitional sentences note that **the IMS budget** helps interpret the broader trends, but we stress it is not valid for analyzing post-2003 emission changes.

---

13. Conclusions and Limitations

Reviewer Comment:
The conclusions are too general. Summarize specific findings, mention data/analysis limitations, and suggest future work.

Response:

- **Conclusions** (Section 555) now succinctly list:
  1. **We have included 95% confidence intervals in the results.**
  2. **Regional differences**: possible reasons (emission changes vs. chemistry).
  3. **Acknowledged uncertainties**: data sparsity in the early 1990s, short time windows (2012 – 2018), MOPITT sensitivity, IMS emission assumptions.
  4. **Future studies**: better modeling with updated emission inventories, more synergy with additional satellite data (IASI, TROPOMI), or Lagrangian trajectory analysis for source attribution.

We explicitly state that **some trends are not statistically significant** and that more in-depth modeling or longer records could clarify the influence of emissions vs. chemistry.

---

14. Technical Corrections and Typos

Reviewer Comment:
Multiple spelling and labeling issues, such as "pptv" vs. "ppbv," references to "UAS" vs. "USA," missing references, figure caption typos, etc. Also, HYSPLIT is mentioned in the acknowledgments even though it is not used here.

Response:

- We have **proofread** the entire manuscript carefully to eliminate typographical errors (e.g., "dwonstream" → "downstream," "routues" → "routes," "pptv" → "ppbv," etc.).
- We corrected the **longitude references** (e.g., "123°E – 142°W" instead of "E – E").
- We removed **unnecessary** references to HYSPLIT in the acknowledgments, clarifying that HYSPLIT was utilized in the separate [Wang et al., 2024] paper, not in this one.
- We added a **Data Availability** section to comply with journal requirements and clarify where readers can obtain the datasets.

Conclusion

We appreciate Reviewer 3's thorough suggestions, which prompted us to improve our manuscript's structure, clarity, and scientific detail. The revised paper now:

1. Features a **more consistent** discussion of timescales and datasets,
2. Provides **detailed methodological** explanations (instruments, data processing, uncertainties, trend calculations),
3. Includes a more **rigorous approach** to significance testing and interpretation of trends,
4. Presents **refined figures** and **concise conclusions** aligned with standard practices for multi-dataset CO trend studies.

We hope these revisions address all concerns raised and make our manuscript suitable for publication in *Atmospheric Chemistry and Physics*.