Response to Referee #2 (Anonymous)

We thank the reviewer for the thorough review of the paper and for recommending publication subject to minor comments. We have addressed these specific points below. Our replies are in *##blue italics.*

General Comments:

The manuscript provides an overview of two newly developed CFC-11 and 12 products, which combine the TOMCAT chemistry transport model capabilities with solar occultation observations to obtain a gap-free data set over the last two decades.

The manuscript is well-written, the methodology is clearly exposed, and the results are overall logically presented. A couple of sections need some reworking to improve the description of the results, and some points need to be clarified, particularly Section 5.4.

I recommend the publication with minor comments, which are discussed below.

##Thank you

Minor comments

1. I found the mention of Version 1 of the presented algorithm in the ACE-FTS section out of place. I would move this to the end of Sect. 3, where "Version 1" is already mentioned. In other parts of the paper version 1 is also discussed: was this version developed in Dhomse and Chipperfield (2023), or was it an internal processing? Can you give some more references/introduction to this versioning in Sect. 3?

##Clarification is added in Abstract, Introduction and Section 3. In the revised version, we clarify that methodology described in Dhomse and Chipperfield (2021) was for version 1.0 and primarily focussed on N2O and CH4 using both HALOE and ACE-FTS data. Using V1.0 methodology, profile data sets for other species were constructed (e.g. O3, H2O, HCl, HF) and were released via Zenodo. We note that a revised methodology and constructed data sets are labelled as V2.0. Here, we use ACE-FTS data only and the manuscript provides updates to the V1.0 methodology (such as using more features, improved preprocessing and use of absolute values instead of TOMCAT-observation differences as target variable).

2. One point in the explanation of the XGBoost training is not clear to me: the ML is said to be trained at each latitude band and altitude independently. However, from line 210, it seems that the CFC concentration time series is derived for all TOMCAT latitude grid points at the same time. Can you clarify this point?

##We have updated Section 4.3 of the revised manuscript to clarify the distinction between the independent training process and the subsequent global dataset construction.

The XGBoost model is indeed trained independently for each 1 km altitude level and within each of the five specified latitude bins to ensure the machine learning captures local chemical and dynamical characteristics. The confusion regarding global fields is addressed as follows:

1) *Training Phase: Models are developed for specific "regimes" (e.g., 14 km altitude in the SHpol bin) using co-located data points.*
2) *Prediction Phase: Once a specific model is trained, it is applied to the continuous TOMCAT feature matrix at that altitude across all 128 longitudes and all latitudes belonging to that specific bin.*
3) *Merging Phase: These independent regional and vertical predictions are then merged onto the global TOMCAT grid (64 latitudes × 128 longitudes) to create the final 3D gap-free fields. For overlapping latitude regions, a simple averaging is used to ensure spatial continuity.*

These details have been explicitly added to the text in Section 4.3 to ensure the methodology is transparent and reproducible.

3. Sect. 5.1: it seems to me that the three bullet points are explaining the reasons for the performance differences between tropics and mid latitudes, and not the contradictory results (small RMSE and small R2 values at the same time) found in the tropics, as stated at line 226. Can this point be better discussed, maybe in 5.1.1?

We have added a dedicated summary paragraph at the end of Section 5.1 to specifically address the coexistence of low RMSE and low R2 values in the tropical latitude band. We modified the text as

"Overall, the small RMSE values demonstrate that the absolute prediction errors remain low because the fundamental dynamical processes driving CFC concentrations are reasonably well-represented in TOMCAT. Conversely, the lower R2 values are a result of the high dynamical variability and limited data availability in the tropics. In particular, at lower altitudes (below 17 km) ACE-FTS measures upper tropospheric VMRs, which make it statistically more difficult for the model to explain a high proportion of the observed variance even when the absolute predictions are accurate."

4. Sect. 5.2: I found confusing the description of SHAP values for Figs. 5 and 6. For example, at line ~316 the authors say that high AoA and low CFC-11 concentrations are

associated with positive SHAP values, but it seems to be the opposite case from Fig. 5: red AoA and blue CFC-11 dots correspond to negative SHAP values. Can you please clarify this point that seems contradictory?

We apologise for the confusion in the SHAP value description. The reviewer is correct: in Figure 5, high AoA (red) and low CFC-11 (blue) dots correspond to the influence direction shown in the plot. We have corrected these descriptions to ensure they are consistent with Figure 5 and Figure 6, where high Age-of-Air (red dots) and low TOMCAT concentrations (blue dots) are clearly located on the negative side of the SHAP axis, indicating a downward adjustment to the model output.

The exact changes in the revised manuscript are as follows:

For CFC-11 (Section 5.2), we modified the text as :

" The analysis of the SHAP values in Figure 5b for CFC-11 at 14 km in the SHpol region provides a detailed look at how XGBoost functions as a transport corrector by applying specific adjustments based on the chemical and dynamical characteristics of different air masses. The results indicate that XGBoost is able to identify distinct regimes to rectify biases in the underlying TOMCAT fields. Age-of-Air is identified as the most influential dynamical feature, where high values representing photochemically aged air that has descended from the mid-to-upper stratosphere are associated with negative SHAP values reaching −50 pptv. Conversely, younger air masses receive positive adjustments of up to +40 pptv. The simulated TOMCAT CFC-11 tracer concentration itself also plays a major role; air masses where TOMCAT-simulated VMRs are relatively high, typically representing air transported from lower latitudes, receive positive corrections of approximately +40 pptv, while very little adjustment occurs at lower concentration levels. Other features also help to characterise these regimes. For example, airmasses with high water vapour and warmer temperatures receive negative adjustments, whereas corresponding air with low water vapour (a sign of dehydrated vortex air), and colder conditions receive positive adjustments."

Similarly analysis for CFC-12 is added in the revised manuscript (Section 5.2) detailing influence of most prominent features.

5. Sect 5.4: the description of the results in this section is not very clear. I see the difficulty in condensing several time series at different latitudes in the same plot, but it is hard to identify in Figs. 9 and 10 features that are described in the text. For example, I can clearly spot a negative TOMCAT bias only in the tropics (0° lat) at both 50 hPa and 100 hPa. At mid-latitudes, TOMCAT seems to be in good agreement with observations at 50 hPa and with TCOM, whereas TCOM introduces positive/negative corrections at 100 hPa in the N/S hemisphere, respectively. This is quite different from what is written in

the text at lines 374-378. Another example at line ~391: I can hardly see any bias removed in the 32.5° N/S panels at 50 hPa, for either CFC-11 or CFC-12. Again, at line ~398, I don't see how TCOM is successfully tracking the MIPAS data in the polar regions in Fig. 9, as the orange line is very different from the green dots. It is not always clear if the sentences are referring to CFC-11 or CFC-12. I suggest revising the whole description of these figures.

We agree with the reviewer that condensing multiple time series into a single figure made it difficult to distinguish specific features and identify the biases discussed in the text. To resolve this, we have completely revised Figures 9 and 10. The new figures now focus on three specific latitude bins (60°S, 30°S, and 0°) across three pressure levels (100, 50, and 10 hPa), resulting in a clearer 9-panel layout. We have also rewritten Section 5.4 to provide a systematic, panel-by-panel discussion.
Specifically:

- *Tropics (0°): As noted by the reviewer, a negative TOMCAT bias is most prominent here at 50 hPa and 100 hPa, which TCOM successfully corrects to match ACE-FTS observations.*
- *Subtropics (30°S): The differences between TOMCAT and TCOM are much smaller in this region. TCOM values are slightly adjusted to align with SPARC ACE-FTS data within uncertainty bounds.*
- *High latitudes (60°S): We observe high biases (overestimations) in TOMCAT at the 100 hPa and 50 hPa levels, which TCOM accurately rectifies with downward adjustments as described in the revised Section 5.4.1 and 5.4.2.*
- *Inter-instrument comparisons: We now explicitly discuss the high bias of SPARC MIPAS values relative to both ACE-FTS and TCOM.*

The revised text now explicitly refers to CFC-11 or CFC-12 in each paragraph to avoid confusion.

6. Can you mention in the conclusion the future of this data set without ACE-FTS observations? And possibly in the introduction, if any upcoming limb mission is going to have the capabilities of measuring these molecules?

We have added discussion about NASA's STRIVE mission, as

"Finally, while the current version of the TCOM dataset relies on the high-quality constraints provided by ACE-FTS, the methodology is designed to be adaptable to future observational streams. The upcoming ALICE instrument on NASA STRIVE mission (to be launched in early 2030s), with its ability to measure vertical profiles for a wide range of trace gases, including both CFCs and other species currently included in the TCOM datasets (e.g. O3, CH4, N2O, HN O3, H2O), will offer a critical opportunity to maintain and extend these gap-free data records well into the next decade. Within current framework, we have shown that even with a data gap (e.g. 2000 to 2003 in this study) profile data sets can be constructed for most of the long-lived species based on the chemical-dynamical information derived during the observed period. Extrapolations based on gradually varying source gas loadings, e.g. obtainable from ground-based observations, can be accommodated. The most problematic issue would be sudden atmospheric perturbations such as the changes in the stratospheric aerosol properties due to volcanic eruptions or wildfires for which direct aerosol observations are needed. Nonetheless, future trace observations from missions such as STRIVE are essential to provide reference anchor points for the end of a gap-filled period."

Typos and technical corrections

7. L56: for February →from February;

Done

8. L85: add space between "the" and "SCISAT";

Done

9. ~L135: please add that you show some results in Figs. S1 and S2.

Done

10. L153: "...that are small"→ "...are small";

Done

11. L190: The sentence "These updates have improved..." would fit better at the end of 4.2 section.

Done

12. L235: stratospheres → stratosphere;

Done

13. L257: "...in XGBoost indicating that the discrepancies..." → "...in XGBoost, the discrepancies...";

Modified as "in XGBoost, so the discrepancies may be"

14. Figs. 7 and 8: why are the color bars moving upwards with the subplots?

Modified so that 0 is close to middle of each subplot.

15. Caption of Fig. 9 : "e.g." → "from";

Done

16. L376: I would delete "observed";

Done

17. L421: I would remove that the datasets "will be publicly released in Dec 2025".

Updated and expanded as "The TCOM version 2.0 datasets are publicly available on Zenodo. While the initial methodology for version 1.0 was detailed in Dhomse and Chipperfield (2023) focusing on CH4 and N2O, the current v2.0 products described in this study have been expanded to include a wide range of stratospheric species. Specifically, the datasets for CFC-11 and CFC-12 can be accessed at https://zenodo.org/records/18145730 (Dhomse, 2026a) and https://zenodo.org/records/18147392 (Dhomse, 2026b), respectively. Using the same methodology, profile data set (v2.0) for the following species are also publicly available:

– O3: https://zenodo.org/records/18199586

– H2O: https://zenodo.org/records/18199962

– CH4: https://zenodo.org/records/18197333

– N2O: https://zenodo.org/records/18197444

– HF: https://zenodo.org/records/18184779455

– HCl: https://zenodo.org/records/18184430

– HNO3: https://zenodo.org/records/18199002

– COF2: *https://zenodo.org/records/18201786*