Main Comments:

The authors develop a powerful parameterization of ozone production rate using satellite-derived columns of NO2 and HCHO along with modeled photolysis rates. Overall, an excellent, important, manuscript, although it can be difficult to follow at times.

We thank this reviewer for their constructive comments. Our response is as follows:

L206: At the beginning of the methods section, add a few sentences explaining how sections 3.1-3.4 fit together.

Response
We added a few sentences to describe the sections.
Modifications
"In this section, we begin by discussing a robust regression model specifically developed for feature selection in the parameterization of PO_3. We then describe the training dataset created for this purpose. Following that, we introduce a clustering technique utilized to organize the training data, which enables us to identify the key drivers of PO_3 variability. Finally, we provide a comprehensive overview of the PO_3 estimates algorithm by integrating data from the TROPOMI retrievals, ground-based remote sensing, and various models."

L274-L296: How does the clustering described in section 3.3 fit into the rest of the paper? Is it part of the coefficient determination or just an analysis tool. Please explain more clearly.

Response
The clustering algorithm was an auxiliary tool to pinpoint the major drivers of PO_3 variability as well as to show that a wide range of atmospheric conditions has been covered in our study.
Modifications
We modified the section by starting: "The aim of using a classifier to group the large quantity and types of aircraft data into similar features is to allow us to study the primary contributors to PO_3 under different chemical, solar, and meteorological conditions. Additionally, this approach will help us understand the range of atmospheric conditions included in the training dataset."

L400: How did you end up with 7 distinct classes after your clustering analysis? Was it trial and error based on how the deviations of observations from the centroids of the 11 features looked?

Response
While some statistical tools (such as the silhouette metric) can help find the optimum number of classes, we found the number of classes sufficient to explain their distinctive characteristics with respect to solar radiation, FNR, FNP, and altitude. Almost every class has a unique feature, allowing us to explain their differences quickly.

L524-543. It appears that you adjust the TROPOMI NO2 and HCHO to remove biases with respect to MAX-DOAS and FTIR observations. How important is this result to your bottom line coefficients for PO3 and wouldn't these biases be regional and subject to change with new versions of TROPOMI data?

Response
Thanks for the comment; the coefficients determined for the PO_3 parametrization rely only on the training dataset obtained from the observationally-constrained F0AM model. They do not depend on the satellite dataset. What can change that coefficient is the inclusion of a new air

quality campaign or different configurations in F0AM. The statistics we gain from TROPOMI errors can change with new updates to the retrieval algorithm or new benchmarks; however, as far as we know, the present work is the most comprehensive validation of the most recent TROPOMI dataset. It is also important to recognize that some biases associated with these data are systematic and do not change drastically from a specific version of the product to another. For instance, the large underestimation of both HCHO and NO2 columns has been widely recognized in literature based on various subsets of ground-based remote sensing measurements (please see Table 1 in https://acp.copernicus.org/articles/21/18227/2021/).

Minor Comments

L93 Break this paragraph into two with the last paragraph previewing what you are doing in this manuscript.

Response

Thanks, we divided them into two pieces.

L127: Is there a version number for these recently reprocessed fields?

Response

Yes, we added it.

Modifications

We use the recently reprocessed daily level-2 (L2) TROPOMI tropospheric NO_2 and total HCHO columns (v2.4) derived from UV-visible radiances onboard the European Space Agency's (ESA's) Sentinel-5 Precursor (S5P) spacecraft (~328-496 nm).

L199-200: I don't understand the meanings of the colons within the parentheses. Are these threshold values for the look up tables? If yes, why do the values jump around so much such as 100:50:600?

Response

Sorry for the confusion! The first and the last numbers are the boundaries and the middle number is the interval.

Modifications

We modified the sentence to: "This look-up table is based on the calculation of more than 20,064 solar spectra over a wide range of solar zenith angle (SZA) (the range [0, 90] in steps of 5°), altitude (the range [0, 15] in steps of 1 km), overhead total ozone column (the range [100, 600] in steps of 50 DU), and surface UV albedo (the range [0, 1] in steps of 0.2) using NCAR's Tropospheric Ultraviolent and Visible radiation model (TUV v5.2) and cross sections and quantum yields from IUPAC and JPL (Wolfe et al., 2016)."

L268: Do you really mean equation 3 here?

Response

Corrected.

Figure 2. Perhaps italicize SZA, ambient temperature, and Pressure as they are dropped?

Response

Done.

Figure 2: Why is the left column labeled “Input Candidates from Aircraft” when it uses model and satellite data?

Response
The inputs for the parametrization (i.e., training data set) come from the F0AM model constrained by aircraft dataset. For the prediction (the right part of the flowchart) we use satellites and models.

L325: M2GMI (be sure to define somewhere)

Response
Defined.

L359-L364: Here or perhaps in section 3.2, Add some background on the role dilution factors play in box model calculations.

Response
Thanks, we had already dedicated few sentences to talk about the role of dilution factor: *…As a result, we have simplified the physical loss by employing a first-order dilution rate set to 1/86400 s-1, equivalent to a lifetime of 24 hours. This approach ensures that unconstrained trace gases that take longer to break down do not accumulate over time. Exact knowledge of dilution factors requires knowing molecular and turbulent diffusion, entrainment and detrainment, and deposition rates, all of which are unknown at the micro-scale level of aircraft observations. Nonetheless, studies of Brune et al. (2022) and Souri et al. (2023) showed that HO2, OH, NOx, and HCHO are relatively immune to the choice of the dilution factor, whereas RO2 mixing ratios can depart introducing some biases in PO3 estimates….*

L391-397: The last 3 sentences of the Figure 3 caption contain information that is also in the main body of the article. Perhaps delete. You probably should mention which field campaigns had the most observations and therefore played the largest role in determining the statistics.

Response
Removed.

Figure 4: I notice you use log(FNR) and log(FNP) here. Could you explain the benefits of this transformation.

Response
Both FNRs and FNP can have extremely large values making it more difficult to put both low and high values in the same plot. Therefore, the use of log() was helpful to have all of them on the same plot with minimal spaces.
Modifications
We added the following sentence to Figure 4: “Both FNR and FNP are scaled using the logarithmic function to enable the simultaneous visualization of low and high values within a single plot.”

L449-451. Did ambient T, H2O vapor, pressure and/or SZA add any additional insights? Preview the results here.

Response

H2O is known to influence ozone through O1D+H2O->2OH, and many chemical reactions rely on temperature and pressure. As mentioned in the paper, we can't say if they add new information based on the clustering algorithm. Still, we decided to include them in the LASSO estimate so the L1-regularization could determine if they are helpful at better predicting PO3. The LASSO algorithm didn't consider them for several reasons: i) we think there is a strong correlation between HCHO and temperature, so HCHO data already have temperature information included; ii) SZA and photolysis rates are highly correlated; and iii) H2O has non-linear effect on PO3 due to generation of 2OH molecules (please see https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2019GL084486) ; since we did not separate the regression into different humid regions, the LASSO algorithm was unable to use H2O in a linear form. This is an inherent limitation of the LASSO algorithm, and for this reason, we had mentioned in the conclusion that we may need to explore the capabilities of a more sophisticated algorithm (i.e., deep neural network) to consider all non-linearities without having to linearize the problem using various ozone indicators.

L467-469: Earlier you mention that SZA, pressure, and temperature were dropped. Here you also include H2O vapor.

Response

Thanks for noticing this! We added water vapor to the text.

L621: Be sure to expand the Benelux acronym the first time it is introduced.

Response

Added.

Figure 17. You may want to change the order of the contributions so that the third listed contribution in the legend (jNo2) is also the third in the Figure (it is currently the second from the top).

Response

The legend follows the order of the "area" function in MATLAB starting from the bottom to the top part of the charts. So we decided to leave it as is.

L756-773: The financial support section lists numerous measurements some of which seem to have little relation to this project. Would it be possible to tighten this section up by eliminating data sets that are only peripherally related to this study while adding more information on how particular measurements were important for this study.

Response

While we fully understand the reviewer's concern, it is mandatory for us to include all FTIR and MAX-DOAS contributions in the acknowledgment. It is part of their terms of use. The correction factors derived from these datasets had a significant effect on our results because the slopes were far from one.

Grammatical Comments:

L96: use degrees symbol.

Response
Corrected.

Section 2.4. …. But how do you convert the VCDS?

Response
Thanks, we added the equation.
Modifications
To carry out the conversion, we apply the following conversion factor (γ) to the TROPOMI VCDs:
$$\gamma = \frac{\bar{q}_{PBLH}}{\frac{NA}{g \times Mair}\sum qdp} \quad (4)$$
where $\bar{q}_{PBLH}$ is the average of the target trace gas mixing ratios in the PBLH, *g* is the acceleration of the gravity (assumed 9.81 m/s^2), *NA* is the Avogadro constant, *Mair* is the air molecular weight (assumed 28.96 g/mol), *q* is the target trace gas mixing ratio at a given altitude, and *dp* is the thickness of each model vertical grid box in hPa. The denominator in Eq. 4 represents the modeled VCD. We integrate modeled partial VCDs up to top of the atmosphere for HCHO, and up to the tropopause pressure layer for NO_2.

L197: To estimate photolysis rates of JNO2 and JO1d-◊ To estimate the photolysis rates, JNO2 and JO1d), we

Response
Corrected.

L216: --> (Tibshirani, 1996). They consider a regression,

Response
Corrected.

L292: These features include --> These features are

Response
Corrected.

L317: are based on converted the bias-corrected --> are derived by converting the bias-corrected

Response
Corrected.

Figure 2: Typo. Should be M2GMI Conversion Factor within the diamond.

Response
Thanks for noticing this! Fixed.

L480: more photolysis rates --> higher photolysis rates

Response
Corrected.

L487: by random dropping --> by randomly dropping

Response
Corrected.

L513: predictor power --> predictive power

Response
Corrected.

L596: making NO2 levels --> meaning NO2 levels

Response
Corrected.

L686: maps of within the PBL --> PBL maps

Response
We fixed this sentence.
Modifications
"In this study, we generated PO_3 maps within the planetary boundary layer (PBL), constrained by bias-corrected TROPOspheric Monitoring Instrument (TROPOMI) observations, using a piecewise regularized regression model."