May 26, 2022
Ref.: acp-2022-202

Author Response (AR): We thank all three reviewers for providing thoughtful guidance and insight, improving the current manuscript. We are excited to share these results with the greater community.

AR: All of the below comments have been addressed in the final manuscript, unless otherwise noted.

Interactive comment on "Measurement Report: Tropospheric and Stratospheric Ozone Profiles during the 2019 TROpomi vaLIdation eXperiment (TROLIX-19)" by John Sullivan et al.,

Anonymous Referee #1

The paper submitted to ACP by J.F. Sullivan and co-workers is a measurement report of data collected during a joint NASA-KNMI field experiment designed to validate ozone retrieval by the TROPOMI spaceborne mission. On 11 days, lidar O3 observations in the troposphere and the stratosphere are carried out for several hours at the Cabauw site. The two NASA lidars are indeed very valuable tools for the assessment of TROPOMI. Observations of total ozone column (TOC), stratospheric limb sounders (MLS), and nearby ECC launched at De Bilt are also analyzed in this work, in addition to the daily ozone simulations of the GEOS-CF model. The time evolution of TOC and tropospheric O3 partial columns (0-2km, 0km-Tropopause level) is used to assess TROPOMI observations and GEOS-CF model ozone mapping. Considering the value of satellite validation exercises and comparisons between model and observations, the paper is appropriate for publication in ACP (or AMT considering the focus on experimental data analysis). It is well written with some remaining errors in figure legends. I fully agree that there is a general good agreement between the different observations and that the GEOS-CF model performs quite well.

AR: Thank you for understanding the value of utilizing many different platforms to better evaluate the satellite products and their physical interpretation. We hope to continue using the lidars, models, and other ground-based observations to evaluate future satellites as well. We also agree this is well within the scope of ACP (especially the special issue) and think has broad enough scope to be quite impactful.

My main criticism would be that detailed discussions of the differences observed in the troposphere are sometimes missing and that these differences need to be acknowledged in the conclusions: TROPOMI overestimate in the lowermost stratosphere (10-15 km), GEOS-CF low performance during the 9/20-21 episode (not only at 4 km but also below the tropopause leading to small differences of tropospheric columns without a good agreement in the vertical profile), ECC/lidar positive differences in the free troposphere on 9/12,17,21 (De Bilt/Cabauw spatial differences ?).

AR: We appreciate the thoughtful review and have answered the below comments that should be more than adequate to alleviate any concerns.

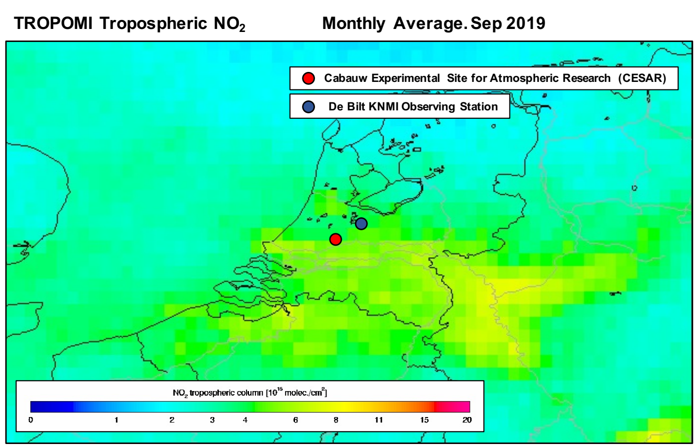
Specific comments

Line 45 : Lidar is also sensitive to cloud cover. Are missing lidar data during the campaign related to cloud cover ?

AR: Yes and this has now been mentioned in the paragraph beginning on Line 116.

Fig. 1 : Fig. 1 is not referenced in the text. What is the purpose of this figure ? Please also show De Bilt and Cabauw on the NO2 map.

AR: Figure 1 is now reference in line 50. The purpose of this figure is to provide the read with geographic context for the site. Rather than focus on a specific day in the original submission (Sep 14, 2019), the image now serves as a monthly mean for September 2019 (the TROLIX-19 period).



Line 108 : Addition of lidar O3 measurement accuracy and vertical resolution would be useful in Table 1. Since partial tropospheric columns and TOCs are discussed in section 3.2 and 5, the expected accuracy of the lidar retrieval on these columns would help.

AR: The below text has been added to the paper near line 125 to better characterize the uncertainty estimates.

A brief description and community standardized definitions of the uncertainty budget of the lidar measurements presented in this paper can be found in Sullivan *et al.,* 2014, Leblanc *et al.,* 2016 and Leblanc *et al.,* 2018. The maximum statistical uncertainties for the two GSFC lidars vary from night to night depending on atmospheric conditions and laser power fluctuations. They are mostly within 10-20% for 5 min and 5-8% for 30 min integrations throughout the atmosphere within their measurable ranges although they are different at the same altitude due to laser performance and telescope/detector efficiency differences.

Line 151 : Give a more recent reference i.e. Smit and Thompson 2021 GAW report 268 describing ASOPOS 2.0 error calculations.

AR: This reference has been added and we appreciate the attention to detail relating to contemporary references on this subject matter.

Line 171 : What is the expected difference when using the 2.5 PVu instead the 3 PVu tropopause definition taken for the lidar tropospheric ozone column calculation ?

AR: The tropopause height was calculated from the GEOS-CF and was then used consistently for all measurements in the trop. ozone column calculations. We have confidence that this is appropriate by visually comparing the tropopause heights compared to the model and lidar in Figure 2 and later on with ozonesonde values. Within the community, there are several definitions of the tropopause height, and we have done our best effort to stay consistent with the 2.5 PVU value and have described the method used to calculate this in the literature.

Line 206 : What is the typical GEOS-CF model vertical resolution in the UTLS ?

AR: There are 72 total vertical layers in the GEOS-CF model, scaled by atmospheric pressure From 10 mb to 375mb there are 20 vertical layers. In general this is roughly ~1.1-1.2 km and a visual reference is shown in Fig. 3 here :https://acp.copernicus.org/articles/17/1417/2017/acp-17-1417-2017.pdf

You can find much of the details (also cited in paper as Keller et al., 2021) pertaining to the GEOS-CF system here:
https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2020MS002413

Line 225 : Add 2f in the list of panels with large lidar/model differences.

AR: Done.

Line 229 : It is also true on 9/13-14.

AR: Added.

Line 233 : What is the thickness of the 4-km layer used for plotting ozone in Fig. 3 ?

AR: The layer was calculated to match the closest vertical layer of the GEOS-CF for consistent intercomparison. The top of model layer 53 is at 3.94 km (~605 mb, for the specific files used in this work) and the top of model layer 54 is 4.44 km (~570 mb, for the specific files used in this work). This is roughly a 500m or 40mb layer thickness.

Figure 2 : Ozone unit is not specified in the caption or color scale. Please add the pink tropopause altitude on the ozone plot and the vertical limits of the 4-km layer used in Fig. 3.

What is the time end of panel e ? This plot is very nice. It's a shame not including the two days with ECC at De Bilt on 9/12 and 9/17.

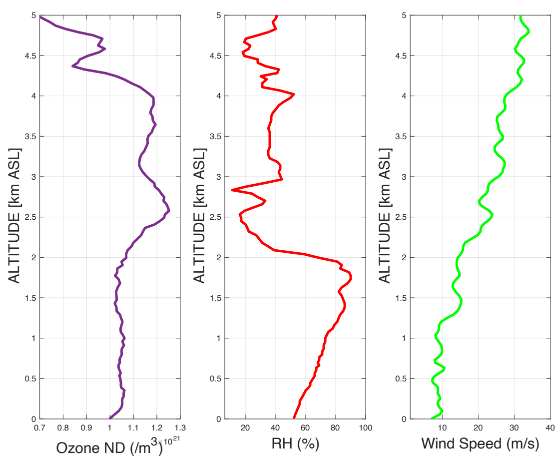
AR: The ozone unit in Figure 2 has been added – Ozone mixing ratio [ppbv]. Also the tropopause heights have been added to the ozone lidar measurements. The details of the layer thickness have also been added to the caption.

Line 243 : The O3 vertical structure in the UTLS is also missed by the model (no low O3 at 11 km). It is then likely that the mesoscale ozone 3D transport in the frontal system is not very well resolved by the model for this specific event.

AR: A comment reflecting this has been added to the revised manuscript starting near Line 286.

Line 246 : Why are 9/17 ECC ozone concentrations significantly larger than model, TROPOMI and lidar O3 values ? Did you see any horizontal O3 gradients near Cabauw in the model or TROPOMI mapping ?

AR: Below is a quick plot of the 9/17 ECC (however, not added to current manuscript). You can see the sampled ozone, RH, and wind speed in their respective units. There is no sign of any pump current overheating or other QC/QA issues that may cause invalid data within the sonde chamber. In particular, this feature at ~4km near 1.2e21 mol/m3 is associated with a much drier air mass and is notable traveling upwards of 30 m/s during the sampling period. The wind direction was consistently coming from the WNW during this period.
At this speed and direction, the sonde (already launched at De Bilt) may have just been sampling a similar but not identical air mass. The differences begin to decrease in this profile compared to the lidar near about 11km (as shown in the current manuscript Figures).



Line 271-274 : This discussion is unclear. The difference between the tropospheric columns are in fact not too bad but are fortuitous because the effect of the missing layer at 4 km is cancelled by the high model ozone at 11 km. Please reformulate this part of the discussion.

AR: A clarification regarding this has been added to the text near Line 408.

Line 279 Figure 4 low panel (not top panel)

AR: Done.

Line 282 I do not understand this sentence.

AR: A clarification regarding this has been added to the text and discussion in the same paragraph has been added from the previous comment (starting near line 408).

Line 285 What are the reasons for low 0-2 km column in some lidar data on 9/12 and 14 and in TROPOMI data on 9/17 ?

AR: These values were driven by cloud contamination in the lower levels of the troposphere for all observations.

For the 09/12 and 09/14 lidar data, this is due to some data points being cloud contaminated. A stricter (and more manual) cloud screen using lidar backscatter SNR has been implemented now.

For the low TROPOMI data on 09/17 (and by extension a few other days in the second half of the domain) are also likely due to cloud interference. There were several TROPOMI pixels that were no retrieved because they did not match the quality requirements (noise too large). By using the OMPS reflectivity data, I have screened additionally for the tropospheric and 0-2km column data in Figure 4. This was changed from a 0.8 reflectivity to 0.6 reflectivity and a note has been made in the caption.

Line 288 A discussion on the TROPOMI retrieval of the partial column is needed in this section. While the 0-2 km columns remain within the range of the diurnal variability, several full tropospheric columns (9 /12,13, 18,19,20) are well above this diurnal variability. It is likely related to the limited vertical resolution in the UTLS. The TROPOMI overestimates are also quite clear in the ozone profiles shown in Fig. 5.

AR: This is an important point, thank you for this. A short discussion has been added, and the reader is directed to the Mettig et al., work that systematically goes through the retrieval from first principles within the spectral range of the ozone profile retrieval (https://amt.copernicus.org/articles/14/6057/2021/amt-14-6057-2021.pdf).

Line 299 Reading this sentence, I am not sure which altitude range is critically needed for monitoring of the column

AR: This has been specified in the context of the previous comment.

Line 308 What is OMPS-LP ? The stratospheric MERRA-2 profile used for the OMPS tropospheric ozone column calculation ? Do hybrid lidar profiles combine daytime TROPOZ data with nighttime STROZ-LITE data or are nighttime profiles only considered in the hybrid version ? Please clarify this point.

AR: From Section 2.3.1, this is the OMPS Limb Profiler ozone profile product and a clarifying statement has been added to start Sec 2.3.1.

The vertical distribution of ozone in the stratosphere and lower mesosphere is obtained from the OMPS Limb-Profiler (LP) sensor on the Suomi-NPP satellite merging the UV (29.5-52.5 km) and VIS (12.5-35.5 km) bands to provide a full profile from 12.5km to 52.5km (Kramarova *et al.*, 2018).

AR: I have specified that this is daytime or nighttime TROPOZ data, and only nighttime measurements for the STROZ. This was ultimately manually selected each day since this was such a focused intensive campaign and we were able to do so. Laser performance and cloud considerations also drove the decisions.

Fig. 5 I believe that the titles of the bottom panels are wrong otherwise Fig.3 and Fig.5 are not consistent.

AR: Great catch – the ozonesonde read file was pointing to opposite directories for those two days (so the titles were correct, but the ECC profile is swapped). This has been corrected and actually shows a much better comparison to the lidar and model, while simultaneously highlighting the feature described in Fig 3.

Line 320 Comparisons of the stratospheric profiles are extensively discussed while the general agreement is quite good in the stratosphere. The discussion of the tropospheric differences is however limited while there are some interesting differences (TROPOMI in the UTLS, lidar ECC differences).

AR: A clarification regarding this has been added to the text throughout Lines 352-373 and is now reiterated in the concluding remarks.

Fig. 7 Color scale is missing

AR: This has been added and this was a major oversight to miss on initial submission.

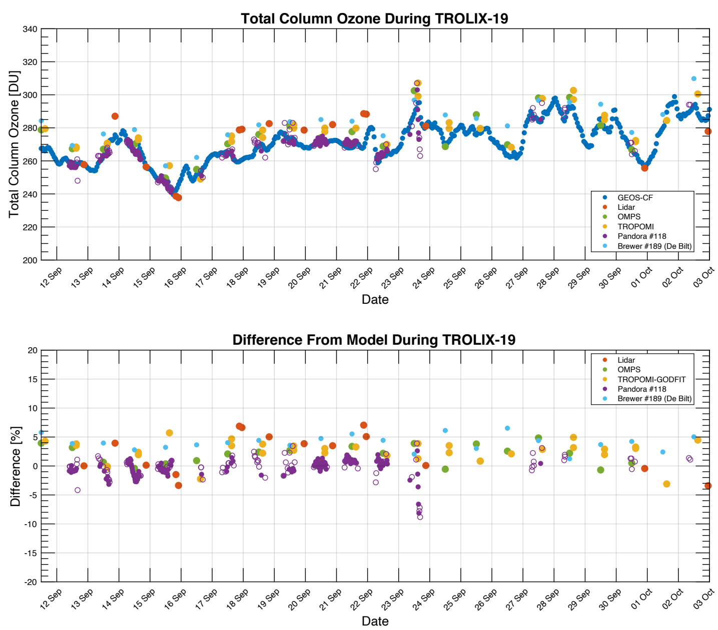
Line 339: Why do you say that OMPS-LP underestimates O3 concentrations at altitudes below 20 km ? I see both positive and negative differences. Regarding TROPOMI, it looks like TROPOMI versus MLS is better between 10 and 20 km than the large lidar/TROPOMI differences in the UTLS. Is it related to similar UTLS vertical resolutions for MLS and TROPOMI ?

AR: A clarification regarding this has been added to the text throughout Lines 352-373 and is now reiterated in the concluding remarks.

Line 358: A 7% difference on TOC is already quite significant. What are the reasons for the low TOCs given by the hybrid lidar retrieval on 9/14-15 or by Pandora on 9/23 ?

In looking further at the mismatch in lidar total column ozone values, we discovered the calculation was performed and plotted with the corresponding TROPOZ integration time of day which was predominantly during the daytime observations. For this specific section describing total column ozone, the majority of that information is coming from the STROZ LITE nighttime data. Therefore, more optimal merges (particularly in the UTLS) were performed to include as much STROZ data as possible and shift the red data points to a more representative time of the lidar observations. This results in a much closer comparison to the observations as compared to the previous figure for the time in question.

Pandora flags equal to 10 (purple filled dots) and 11 (purple open dots) are shown. On 09/23 there were many observations mixed with both flags throughout the day, likely due to cloud contamination. Only data with the highest QC/QA level (i.e. 10) is plotted in the updated final figure. This results in less data overall.



Line 371: GEOS-CF indeed performs well reproducing the ozone downward transport in the UTLS, but a sentence could be added about GEOS-CF failure to resolve some high resolution laminae related to specific mesoscale ozone 3D transport from the UTLS, e.g. 09/19-20.

AR: A clarification regarding this has been added to the text starting near Line 288.

Line 374: The overestimate of the TROPOMI retrieval between 10 and 15 km needs to be mentioned in the conclusions.

AR: A clarification regarding this has been added throughout the previous reccomendations and line numbers above as well as additional context in the conclusions.

We thank you for such a thoughtful review (including a few typos/errors) and know the manuscript has improved greatly from your suggestions.

Anonymous Referee #2

Review of "Measurement Report: 1 Tropospheric and Stratospheric Ozone Profiles during the 2019 TROpomi vaLIdation eXperiment (TROLIX-19)" by J.T. Sullivan et al.

In this manuscript the authors describe the many ozone measurements made from multiple platforms during the TROpomi vaLIdation eXperiment (TROLIX-19) campaign in fall 2019. The campaign was designed to support satellite validation efforts with the emphasis on understanding the vertical profile retrievals of ozone. Instrumentation included ozone lidars, pandora spectrometer, Brewer Spectrophotometer and ozonesondes. Satellite data used include OMPS, MERA-2, MLS and TROPOMI. In addition, measurement data is compared with GEOS-CF model output.

The analysis focuses on two main goals; one to evaluate ozone retrievals in relation to current and future (TEMPO) satellites. And the combination of the tropospheric ozone lidar and the stratospheric lidar providing hybrid ozone profiles from ~0.2 km to ~50 km.

The article is clearly written and provides a comprehensive presentation of data from a number of measurement platforms. With figures and discussion of the comparisons across measurement platforms and model within the full column ozone and 0-2 km tropospheric column ozone. The study makes several important observations as to the structure of the ozone within the atmospheric column and the differences in instrument/model performances.

AR: Thank you for understanding the value of utilizing many different platforms to better evaluate the satellite products and their physical interpretation. We hope to continue using the lidars, models, and other ground-based observations to evaluate future satellites as well. We also agree this is well within the scope of ACP (especially the special issue) and think has broad enough scope to be quite impactful.

Specific comments:

I would suggest that the conclusions section needs to be expanded upon. This section emphasizes the importance of observations and for the site itself (which I agree on) but I would like to see more quantitative evaluation of the data here to back up the statements, adding some percentage differences etc would make this an easy reference source for future readers. For example, there is a statement (line 74-75): " TROPOMI ozone profile products are able to accurately reproduce ozone quantities in the lower troposphere…" with reference to Figure 3 but this only shows TROPOMI compared with observations vertical layer at 4km. In addition, Figure 7 indicates that TROPOMI generally overestimates, especially within the troposphere. Expanding the conclusions to include some of the quantitative results would help to firm up the concluding statements.

AR: The conclusion section has been broadened beyond the importance of observational platforms. We have added several discussion clarification on TROPOMI overestimations per the previous reviewer that are intended to improve the communication surrounding this issue.

Figure 7 was intended to be the quantitative resultant figure that could be referenced or looked up easily and this has been elaborated on in the conclusions. We feel this figure better illustrates both the temporal (over the course of the campaign) and vertical differences observed in the retrievals and reducing that down to a simple number or XX % may underserve the push for the vertical profiling needed for these types of efforts. (And hopefully the colorbar now added helps the reader too!)

Figure 1. Add the CESAR site also be added to the image on the right or add lat/lon information to maps.

AR: CESAR has been added.

Figure 7. Needs a color bar with values for the differences in ozone number densities

AR: Done, thanks!

Reviewer 3:

Review of "Measurement Report: Tropospheric and Stratospheric Ozone Profiles during the 2019 TROpomi vaLIdation eXperiment (TROLIX-19)" by J.T. Sullivan et al, submitted to ACP

In this manuscript, a measurement report is given on the various ground-based measurements deployed during the TROLIX-19 field campaign in Cabauw, Netherlands during September 2019. The focus of the campaign and this paper was to provide comprehensive validation of the TROPOMI ozone profile retrievals using ozone lidar measurements, the measurement focused on in this paper, as well as ozonesondes, Pandora, Brewer, etc. Ozone lidar measurements were also compared against other satellite and model datasets, such as OMPS (satellite), MLS (satellite), and GEOS-CF (model). The authors also analyzed the temporal variability of full tropospheric and 0-2 km ozone columns using GEOS-CF, TROPOMI, ozonesondes, and ozone lidar.

The article provides a comprehensive analysis of the measurements and ancillary datasets used to demonstrate the variability of tropospheric and stratospheric ozone profiles during TROLIX-19. The authors clearly and concisely state the impact of the ground-based measurements have on understanding the capabilities and limitations of current polar orbiting and future geostationary ozone retrievals.

Specific comments:

Line 95: It would be worth mentioning how many stations are in the network

AR: Done.

Line 145: Correct spelling of campaign

AR: Done.

Line 147: Specify which radiosonde manufacturer was used and the specific model

AR: Done.

Line 159: Put space between "are" and "used"

AR: Done.

Figure 2: Add label to the color bar, include units

AR: Done.

Line 233: It can be assumed that 4 km is chosen to investigate the model and observation difference observed between 3-5 km on 20-21 September. However, it should be specifically mentioned at the beginning of this paragraph why 4 km is chosen.

AR: A clarification has been added.

Line 279: Correct to say Figure 4, bottom panel

AR: Done.

~~Measurement Report:~~ Tropospheric and Stratospheric Ozone Profiles during the 2019 TROpomi vaLIdation eXperiment (TROLIX-19)

John T. Sullivan[1], Arnoud Apituley[2], Nora Mettig[3], Karin Kreher[4], K. Emma Knowland[1,5], Marc Allaart[2], Ankie Piters[2], Michel Van Roozendael[6], Pepijn Veefkind[2], Jerry R. Ziemke[1,5], Natalya Kramarova[1], Mark Weber[3] Alexei Rozanov[3], Laurence Twigg[1,7], Grant Sumnicht[1,7], and Thomas J. McGee[1*]

[1] NASA Goddard Space Flight Center, Greenbelt, MD 20771
[2] Royal Netherlands Meteorological Institute (KNMI), De Bilt, Netherlands
[3] Institute of Environmental Physics, University of Bremen, Bremen, Germany
[4] BK Scientific GmbH, Mainz, Germany
[5] Morgan State University/GESTAR-II, Baltimore, MD 21251
[6] Belgian Institute for Space Aeronomie (BIRA), Ukkel, Belgium
[7] Science Systems and Applications Inc., Lanham, MD, 20706
*Now Emeritus

Correspondence to: John Sullivan (john.t.sullivan@nasa.gov)

Abstract A TROPOspheric Monitoring Instrument (TROPOMI) validation campaign was held in the Netherlands based at the CESAR (Cabauw Experimental Site for Atmospheric Research) Observatory during September 2019. The TROpomi vaLIdation eXperiment (TROLIX-19) consisted of active and passive remote sensing platforms in conjunction with several balloon-borne and surface chemical (e.g. ozone and nitrogen dioxide) measurements. The goal of this joint NASA-KNMI geophysical validation campaign was to make intensive observations in the TROPOMI domain in order to be able to establish the quality of the L2 satellite data products under realistic conditions, such as non-idealized conditions with varying cloud cover and a range of atmospheric conditions at a rural site. The research presented here focuses on using ozone lidars from NASA's Goddard Space Flight Center to better evaluate the characterization of ozone throughout TROLIX-19. Results of comparisons to the lidar systems with balloon, space-borne, and ground-based passive measurements are shown. In addition, results are compared to a global coupled chemistry meteorology model to illustrate the vertical variability and columnar amounts of both tropospheric and stratospheric ozone during the campaign period.

1 Introduction

In September 2019, a joint Royal Netherlands Meteorological Institute (KNMI) and the U.S. National Aeronautics and Space Administration (NASA) field campaign was performed in the Netherlands, based at the Cabauw Experimental Site for Atmospheric Research (CESAR, 51.97° N, 4.93° E), to provide the scientific community with additional information to further understand and evaluate the Copernicus Sentinel-5 Precursor mission (S-5P) TROPOspheric Monitoring Instrument (TROPOMI) instrument (https://sentinels.copernicus.eu/web/sentinel/missions/sentinel-5p). The main objective of the Copernicus Sentinel-5P mission is to perform atmospheric measurements with high spatio-temporal resolution, to be used for scientific studies and monitoring of air quality and chemical transport (https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-5P).

To properly support satellite evaluation, the 2019 TROpomi vaLIdation eXperiment (TROLIX-19) campaign was designed to bring together many active and passive remote sensing platforms in conjunction with several balloon-borne, airborne and surface measurements. Specifically, the observations were established to provide geophysical verification in order to establish the quality of TROPOMI Level 2 (L2) main data products under realistic non-idealized conditions with varying cloud cover and a wide range of atmospheric conditions. Cabauw, using its comprehensive in-situ and remote sensing observation program in and around the 213 m meteorological tower (https://ruisdael-observatory.nl/trolix19-tropomi-validation-experiment-2019/) was the main site of the campaign with focus on vertical profiling using lidar instruments for aerosols, clouds, water vapor, tropospheric and stratospheric ozone, as well as balloon-borne sensors for nitrogen dioxide (NO_2) and ozone (Figure 1). Although this work focuses primarily on the ozone lidar profiling during the study, the larger campaign overview, background, and motivation can be found in Apituley *et al.* (2019; 2020) or Kreher *et al.* (2020a).

One main goal of this work is also to understand ozone profile retrievals as they relate to upcoming satellite endeavors. As NASA prepares to launch its first geostationary air quality satellite "Tropospheric

Emissions: Monitoring of POllution" (TEMPO~~),~~) this work also specifically establishes a paradigm of evaluation for TEMPO-derived products such as tropospheric ozone columns and a 0-2km tropospheric ozone product. An analogous geo-stationary air quality satellite, the Copernicus Sentinel-4 mission (S-4), https://sentinels.copernicus.eu/web/sentinel/missions/sentinel-4), will provide hourly data on tropospheric constituents over Europe and the CESAR site is directly within the satellite's field of regard. Due to the finer spatial footprint, increased temporal frequency and vertical extent of ~~TEMPO's~~the TEMPO tropospheric ozone retrievals, ozone lidars are an ideal platform to perform future evaluations of the products, which builds from recent work done in Johnson *et al.,* 2018. Specifically, this work will investigate the results from the combination of having a co-located NASA tropospheric (Sullivan et al., 2014) and stratospheric ozone lidars (McGee *et al.,* 1991) in order to obtain an entire vertical profile of ozone from ~0.2km to 50km.

For the first time, this transportable combination of lidars is able to explicitly derive diurnally varying tropospheric and total ozone columns, which are compared directly to measurements obtained by ground-based passive sensors, current satellite instrumentation and chemical transport models. In Section 2 we present all available data and methods used in this work, across the various platforms during the TROLIX-19 study. Section 3 focuses on comparisons of the tropospheric ozone retrievals of the vertical profiles of ozone within the troposphere and columnar reductions of 0-10 km and 0-2 km. Comparisons of lidar data with available complete ozone profiles (Sec 4) and columnar amounts (Sec 5) from several platforms and chemical transport models are also presented to further understand the quality of satellite derived ozone profiles during the TROLIX-19 period.

2 Data and Methods

Descriptions of the various observational and model data sets and used in this study are below, including a summary table (Table 1).

2.1 NASA Ozone Differential Absorption Lidars (DIAL)

NASA deployed and operated two ozone lidars during TROLIX-19 at the Cabauw site near the CESAR tower to observe temporal and vertical gradients in tropospheric and stratospheric ozone. This was the first dual-deployment of these lidars, in which the tropospheric ozone lidar measured between the near surface (about 0.2 km) to a height of about 18 km and the stratospheric lidar during night-time from 15 km upwards to nearly 50 km, providing complete hybrid ozone profiles for the campaign period. Measurements were made during periods of mostly clear skies, although occasional cloud cover did enter the measurement period.

The NASA GSFC Mobile Stratospheric Ozone Lidar Trailer Experiment (STROZ-LITE) has been a participant in the Network for the Detection of Atmospheric Composition Change (NDACC) since its inception and is housed in a 12.5m container allowing for transport around the world (McGee et al., 1991). The lidar instrument transmits two wavelengths, 308 nm from a XeCl excimer laser, and 355 nm from a ND:YAG laser to derive ozone number density profiles, which have historically served as an intercomparison data set for other NDACC ozone lidars (recent intercomparison can be found at Wing *et al.,* 2020; 2021).

The NASA GSFC TROPOZ has been developed in a transportable 13.5m trailer to take routine measurements of tropospheric ozone near the Baltimore–Washington, D.C. area as well as various campaign locations (Sullivan *et al.,* 2014; 2015,2019, Leblanc *et al.*, 2018). This instrument, which utilizes a ND:YAG laser and Raman cell, has been developed as part of the ground-based Tropospheric Ozone Lidar NETwork (TOLNet, https://www-air.larc.nasa.gov/missions/TOLNet/), which currently consists stations across the North America (http://www-air.larc.nasa.gov/missions/TOLNet/). The primary purposes of the instruments within TOLNet are to provide regular, high-fidelity profile measurements of ozone within the troposphere for satellite and model evaluation. This lidar also operates routinely for the Network for the Detection of Atmospheric Composition Change (NDACC).

More than thirty NDACC ground-based Lidar instruments (https://lidar.jpl.nasa.gov/ndacc/index_ndacc.php) deployed worldwide from Pole to Pole are monitoring atmospheric ozone, temperature, aerosols, water vapor, and polar stratospheric clouds. Both lidars collect backscattered radiation with a large primary telescope and a 10cm telescope for near field channels. Spectral separation is accomplished using dichroic beam-splitters and interference filters. For the stratospheric system, five return wavelengths are recorded: the two transmitted wavelengths, and the nitrogen Raman scattered radiation from each of the transmitted beams 332 nm and 382 nm, and the 408 nm water vapor channel. In this arrangement for TROLIX-19, the tropospheric system pumped the Raman cell with the fourth harmonic (266 nm), which resulted in conversion to 289 nm and 299 nm using a single hydrogen/deuterium Raman cell. All of the signals are further split to improve the dynamic range of the respective lidar optical detection chains and are then amplified, discriminated and recorded using photon counting techniques.

During TROLIX-19, the STROZ-LITE was operated on cloud free nights, with measurements lasting between 2-4 hours to obtain enough signal to properly retrieve the entire stratospheric ozone profile. The TROPOZ was operated during daytime and night time to provide tropospheric ozone profiles. For instances of TROPOMI overpasses, campaign ozonesondes, or coincident stratospheric ozone lidar measurements, the TROPOZ reported data is averaged for 30 minutes, centered around the satellite overpass or launch time. This temporal period of averaging has been optimized in several cases to avoid cloud contamination. For all other times during the TROPOZ operation, the data has been averaged to 10 minutes, which is suitable under most clear sky conditions to retrieve ozone information within the entire troposphere. A brief description and community standardized definitions of the uncertainty budget of the lidar measurements presented in this paper can be found in Sullivan *et al.*, 2014, Leblanc *et al.*, 2016 and Leblanc *et al.*, 2018. The maximum statistical uncertainties for the two GSFC lidars vary from night to night depending on atmospheric conditions and laser power fluctuations. They are mostly within 10-20% for 5 min and 5-8% for 30 min integrations throughout the atmosphere. Within overlapping measurement regions in the upper troposphere/lower stratosphere, they are different at the same altitude due to laser performance and telescope/detector efficiency differences, and are therefore joined manually for this work based on appropriate signal to noise and uncertainty estimates.

2.2 Ground Based Passive Sensors and Ozonesondes

2.2.1 Pandora Spectrometer Instrument

A Pandora spectrometer instrument (#118) has been used to measure columnar amounts of trace gases in the atmosphere at 3–5-minute resolution at the Cabauw site since 2016 and previously used for the second Cabauw Intercomparison of Nitrogen Dioxide (CINDI-2) campaign (Kreher et al., ~~2020b~~2020). Using the theoretical solar spectrum as a reference, Pandora determines trace gas amounts using differential optical absorption spectroscopy (DOAS). This attributes in principal these differences in spectra measured by Pandora to the presence of trace gases within the atmosphere (*i.e.* the difference between the theoretical solar spectrum and measured spectrum is caused by absorption of trace gas species). For this study, L2 direct sun columnar values of ozone are used, although retrievals of nitrogen dioxide are also operationally acquired. Data used passed the strictest QC/QA estimate (Flags = 10) and was obtained from the Pandonia Global Network (http://data.pandonia-global-network.org/).

2.2.2 Brewer MKIII Spectrophotometer

A Brewer MKIII spectrometer instrument (#189) has been used to measure daily columnar amounts of ozone in the atmosphere at the KNMI/De Bilt (30km NE of Cabauw~~) site since 2007. ~~, 52.10° N, 5.18° E). Brewer #189 has been operated continuously since 1 October 2006. It replaced Brewer #100 which provided observations since 1 January 1994. De Bilt has the longest continuous record of ozone measured with an MKIII instrument in the World Ozone and Ultraviolet Radiation Data Centre (WOUDC) database.

The Brewer is specifically designed to provide high accuracy measurement of spectrally resolved UV for satellite evaluation, climatology monitoring and public health to international standards. Similar to Pandora spectrometers, these measurements of total column of trace gases are compared to the measured UV spectrum with the known solar output, and modeling the scattering properties of the atmosphere and have been historically used to evaluate columnar satellite products (McPeters *et al.,* 2007; Wenig *et al.,* 2008; Garane et al., 2019). The Brewer is the standard instrument used in the World Meteorological Organization ozone monitoring network and for NDACC. This data was obtained at the NDACC website (https://www-air.larc.nasa.gov/missions/ndacc/data.html).

2.2.3 Ozonesondes

Ozonesondes have been used to measure vertical profiles of ozone in the atmosphere at the KNMI/De Bilt (30km NE of Cabauw) site since November 1992, and measurements are made weekly, historically at 12 UTC on Thursdays. Description of the Electro Chemical Cell (ECC) details and metadata are summarized in Malderen et al., 2016, which also describes the importance of understanding and reporting changes in ozonesonde operation procedures. During the ~~campagin~~campaign, in situ measurements of ozone were made using a balloon-borne payload consisting of an ECC ozonesonde (Science Pump Corporation, Serial Numbers: 6A35438, 6A35447, 6A35448, 6A35441) coupled with a radiosonde (Vaisala RS41) and have been used to evaluate TROPOMI tropospheric ozone products in the tropics (Hubert et al., 2021). The ECC technique is widely used for the high vertical resolution measurements of O_3. The ECC consists of two chambers with platinum electrodes immersed in potassium iodide (KI) solutions at different concentrations. The accuracy in the O_3 concentration measured by an ECC ozonesonde is ± 5%–10% up to an altitude of 30 km (Smit et al., 2007~~).~~; Smit and Thompson et al., 2021). This data was obtained at the NDACC website (https://www-air.larc.nasa.gov/missions/ndacc/data.html).

2.3 Satellite Observations and Products

Satellite data used in this work was selected based on the closest retrieval (*i.e.* column, profile) to the CESAR station within +/-2.5 degrees latitude and +/-10 degrees in longitude.

2.3.1 Ozone Mapping and Profiling Suite (OMPS) and MERRA-2 products

~~Daily total column ozone overpasses over Cabauw station from the~~ The Ozone Mapping and ~~Profiling~~Profiler Suite (OMPS) ~~Nadir-Mapper (NM) instrument~~ on the Suomi National Polar-orbiting Partnership (S-NPP) platform ~~areused~~consists of three sensors to measure the total column and the vertical distribution of ozone with high spatial and vertical resolutions (Flynn et al., 2006). Daily total column ozone overpasses over Cabauw station from the OMPS Nadir-Mapper (NM) instrument are used in this study. The vertical distribution of ozone in the stratosphere and lower mesosphere is obtained from the OMPS Limb-Profiler (LP) sensor on the Suomi-NPP satellite merging the UV (29.5-52.5 km) and VIS (12.5-35.5 km) bands to provide a full profile from 12.5km to 52.5km (Kramarova *et al*., 2018). Variations of this merged OMPS-LP retrieval were considered, however the work shown in Arosio *et al.,* 2018, indicates the same overall conclusions would be reached. Further work beyond this manuscript may involve comparing this TROLIX-19 measurement data set to specific experimentally performed satellite retrievals.

The Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2) provides data beginning in 1980 and since August 2004 assimilates NASA's satellite ozone profile observations from Aura Microwave Limb Sounder (MLS) (Livesey et al, 2008) to more comprehensively characterize stratospheric ozone abundance. A residual tropospheric ozone product (Ziemke *et al.,* 2019) is derived using the OMPS NM total column ozone minus the co-located MERRA-2 stratospheric column ozone. Tropopause pressure is derived from MERRA-2 potential vorticity (2.5 PVU) and potential temperature (380 K).

2.3.3 MLS

NASA's Aura Microwave Limb Sounder (MLS) uses microwave emission to measure stratospheric and upper tropospheric constituents, such as ozone. Ozone data (v5) used in this study is binned on various vertical grids and are converted from volume mixing ratio to number density using the coincident MERRA-2 atmosphere state parameters. Both daytime and nighttime data are used in this study and the corresponding closest profile is utilized for comparison.

2.3.4 TROPOMI

In October 2017, the Sentinel-5 Precursor (S5P) mission was launched, carrying the TROPOspheric Monitoring Instrument (TROPOMI), which is a nadir-viewing 108° Field-of-View push-broom grating hyperspectral spectrometer. Starting in August 2019, Sentinel-5P TROPOMI along-track high spatial resolution (approximately 5.5 km at nadir) has been implemented and total ozone columns values used in this work are subsetted from the NASA GES DISC (https://tropomi.gesdisc.eosdis.nasa.gov/data/S5P_TROPOMI_Level2/S5P_L2__O3_TOT_HiR.1/) to provide the Offline 1-Orbit L2 (S5P_L2__O3_TOT_HiR), which is based on the Direct-fitting algorithm (S5P_TO3_GODFIT), comprising a non-linear least squares inversion by comparing the simulated and measured backscattered radiances.

Tropospheric Ozone vertical profiles were retrieved using the TOPAS (Tikhonov regularized Ozone Profile retrievAl with SCIATRAN) algorithm and were applied to the TROPOMI L1B spectral data version 2, using spectral data between 270 and 329 nm for the retrieval (Mettig *et al.,* 2021). This data set will cover the TROLIX-19 period from 09 September until 28 September; however, it is available outside of this work for specific weeks between June 2018 and October 2019. Since the ozone profiles are very sensitive to absolute calibration at short wavelengths, a re-calibration of the measured radiances is required using comparisons with simulated radiances with ozone limb profiles from collocated satellites used as input. The a priori profiles for ozone are taken from the ozone climatology of Lamsal et al. (2004) and the calibration correction spectrum is determined using the radiances modelled with ozone information from collocated MLS/Aura measurements as described in depth throughout Mettig *et al.,* 2021.

2.4 Coupled Chemistry and Meteorology Model

The GEOS Composition Forecasting (GEOS-CF, https://gmao.gsfc.nasa.gov/weather_prediction/GEOS-CF/, Keller *et al.,* 2021, Knowland et al., 2021) system was chosen to serve as a comparison simulation for this effort, based on its altitude coverage (up to 80 km) and implications for future geostationary satellite use. The system produces global, three-dimensional distributions of atmospheric composition with a spatial resolution of 25km. Using meteorological analyses from other GEOS systems, the GEOS-CF products include a running atmospheric replay to provide near-time estimates of surface pollutant distributions and the composition of the troposphere and stratosphere. Individual case study evaluations using ozone lidar of the GEOS-CF meteorological replay have recently been performed in Dacic et al. (2020), Gronoff et al. (2021) and Johnson et al. (2021). These results will also be used to better evaluate the GEOS-CF as the source of a priori ozone profiles for use in the TEMPO tropospheric ozone retrievals. Model output for this work is used from the closest GEOS-CF model grid cell to the CESAR observatory.

3 Tropospheric Ozone Comparisons

3.1 Vertical Profiles

Example tropospheric ozone profile observations are presented in **Figure 2** for 7 individual observation periods during the TROLIX-19 campaign. Each of the panels show the cloud screened TROPOZ lidar retrievals (top panels) and the corresponding GEOS-CF model output (bottom panels). Pink dots are overlaid to indicate the simulated tropopause altitude based on a blended estimate (TROPPB) which meets criteria of the lowest altitude bin corresponding with either a pressure level above the thermal tropopause (380K) or dynamical (3 PVU) tropopause.

In general, the observations and simulations agree quite well in characterizing the broad features that impacted the CESAR site during the TROLIX-19 campaign. However, in each panel there are ozone laminae within the lower troposphere that are not replicated in the model simulation, most notably the underestimation of ozone during the September 20-21 period from 3-5km (Figure 2d-~~e). The observations indicate increased ozone levels as compared to the model during this period, centered around the 3-5km and 8-10km region of the atmosphere (this is explored in more detail below~~f, black dashed box). However, the model does simulate well the lowered tropopause height and abundance of lower stratospheric ozone observed in the 2 October observations (see Fig. 2g), which is an indication of the model ~~well representing the dynamical variability that affects the lowering of the tropopause height~~representing the dynamical variability that affects the lowering of the tropopause height. This suggests the model is appropriately capturing the complex dynamics during this period near the upper troposphere, but may not have been initialized with the correct boundary conditions or is too spatially coarse to allow for simulation of the layer emphasized with the black dashed box. However, this is an important altitude region for identifying long range transport of aged stratospheric air and inter-continental transport that may be downward mixing towards the surface layer and will be explored in more detail below.

To bring in additional platforms and to better understand these differences throughout the campaign at discrete altitudes, **Figure 3** shows the ozone number density values for the TROPOZ lidar, GEOS-CF model, TROPOMI and ECC ozonesondes at the average 4 km vertical level for the entire TROLIX-19 campaign period. Within ~~the 4km~~this layer, the platforms are all characterizing the general ozone features throughout the campaign at an altitude that frequently is associated with aged transported ~~layering~~laminae. There is a noticeable difference between the observations and model during the previously described 20-21 September period. On 21 September at 12 UT, the lidar and ECC sonde quantify an elevated layer (1.2-1.3 x 10^{21} molecules m^{-3}) ~~-~~into the region that is not simulated by model (0.75-0.9 x10^{21} molecules m^{-3}), resulting in an approximately 30% ~~difference in ozone abundance within layer. Since the model correctly simulated many other ozone features during this time period within the upper tropospheric region, this is likely aged transport into the domain that was not available during model initialization. Back-trajectories were performed to better identify the source of this air mass, however nothing conclusive can be reported. The layer is not associated with any increase in lidar attenuated backscatter within the associated altitude, suggesting it was not urban in origin and therefore more likely aged stratospheric air mixing down to the lower free troposphere. Outside of this Sep 21 period, there is generally good agreement between the observations and model, indicating the combination of observations and modeling are able to represent the rural conditions and ozone perturbations at the CESAR site.~~underestimation in ozone abundance within the layer.

3.2 Columnar Data Reduction

There continues to be a need within the atmospheric and satellite community to understand the variability of ozone as it pertains to both the tropospheric column (*i.e.* the Earth's surface to the tropopause height) and the 0-2km tropospheric column (*i.e.* the Earth's surface to the 2 km height). The 0-2 km region is of particular interest as it is projected to be delivered hourly from the North American geo-stationary satellite: Tropospheric Emissions: Monitoring of Pollution (TEMPO). Due to the increased temporal frequency and vertical extent of TEMPO's tropospheric ozone retrievals, ozone lidars, such as those from TOLNet (https://www-air.larc.nasa.gov/missions/TOLNet/) used in this work, are an ideal platform to perform future evaluations of the products~~, which build from Johnson *et al.*, 2018~~.

Full tropospheric columns (**Figure 4, top panel**) are consistently calculated from each platform using the blended tropopause height (TROPPB) produced by the GEOS-CF and described above (*c.f.* pink dots in **Figure 2**) and are then converted to Dobson Units (DU). The tropospheric and 0-2km columns are calculated explicitly by integrating the ozone number density from the lowest data bin of usable data to the TROPPB or 2km layer height produced in the nearest model temporal output. The exception to this is the OMPS/MERRA-2 tropospheric column using the residual method described above (subtracting the MERRA-2 stratospheric column from the OMPS-NM total ozone column). ~~For the 0-2km tropospheric column (**Figure 4, bottom panel**), there were no major surface layer pollution events at the CESAR observatory during TROLIX-19.~~

For the full tropospheric column (Figure 4, top panel), the campaign variability ranges from approximatively 20-55 DU~~.~~ based on the lidar observations. The model, lidar, and ECC sonde observations agree quite well throughout the 12 Sep to 23 Sep time frame when looking at day-to-day variability. However, when assessing the variability on a single day for 21 Sep, full tropospheric columns reduced from the lidar observations are some of the largest observed during this TROLIX-19 period (reaching ~~nearly 50~~46 DU), ~~while the~~ but mostly staying between 34-40 DU. During this time, the model mainly ranges between 35-37 DU~~ or a difference~~, resulting in differences within 10% for most of ~~upwards~~the observations (albeit closer to 30% for the peak on this day).

When looking at Figure 2d-f, the lower ozone values just below the tropopause during this period are not simulated in the model, which may be an indication that the mesoscale ozone transport in the frontal system is not very well resolved by the model for this specific event. Since the model correctly simulated many other ozone features during this time period within the upper tropospheric region, this may also be attributed to aged transport into the domain that was not available during model initialization. Back-trajectories were performed to better identify the sources of ~~40%. Therefore, this~~ these air masses, however nothing conclusive can be reported. The layer is not associated with any increase ~~identified and discussed in Fig 3, not only results an~~ in lidar attenuated backscatter within the associated altitude~~ specific difference, but ultimately results in a large overall impact to the full~~, suggesting it was not urban in origin and therefore more likely aged stratospheric air mixing down to the lower free troposphere. Outside of this Sep 21 period, there is generally good agreement between the observations (including the OMPS-MERRA2 product) and model, indicating the combination of observations and modeling is able to represent the rural conditions and ozone perturbations at the CESAR site.

When assessing these tropospheric column values from the TROPOMI ozone profile observations, it is important to mention the vastly different vertical resolution or averaging kernel schemes as compared to the independent observations near the tropopause. The ECC samples an instantaneous observation with a vertical resolution generally less than 100m, while the lidar is averaging over 500-750m of atmosphere for each data point near the tropopause. However, the vertical resolution near the tropopause for TROPOMI using the TOPAS algorithm (Mettig *et al.*, 2021) is nearly 6 km, indicating it is not able to completely represent sharp gradients that may occur near the tropopause layer and the lower stratosphere (where ozone content sharply increases). This lack of degrees of independent information is evident in the relatively higher TROPOMI tropospheric column ozone values as compared to the other independent measurements presented in this work. This suggests ground-based profiling observations are still critically needed to confirm large deviations from a priori and climatology in order to evaluate the atmospheric chemistry models, especially in the upper tropospheric region and within the boundary layer.

There exists both diurnal and day-to-day variability of the 0-2 km ozone, ranging from 4-10 DU (Figure 4, ~~top~~bottom panel). In the 0-2 km ozone reduction, the lidar and model are critically needed to understand ozone variability on a continuous scale. For instance, on 15 Sep the 0-2 km ozone column was near 9 DU at 03 UT and finished near 5.5 DU at 16 UT, resulting in a -60% change in DU within ~~a single day. Furthermore, the gradient of the 21 Sep ozone column change was similar in scale to the entire campaign variability, indicating that there is a significant amount of information gained in the understanding of the variability in ozone from continuous measurements. Although a daily snapshot of OMPS-MERRA-2 residuals and TROPOMI ozone profile observations are critical for their vast spatial coverage, ground-based observations such as ozone lidar and ECC sondes are critically needed to~~

~~quantify measurement gaps.~~13 hours. The need for continuous measurements during highly variable days are further emphasized by the fact that this gradient in 0-2km ozone for this single day (15 Sep, 5.5 DU – 9 DU) was comparable to the variance of 0-2 km ozone values throughout the entire campaign.

In summary, we find that the ozone columns evaluated in this study generally reproduced the structure of the TROLIX-19 ozone lidar observations for N=835 coincidences. For the full tropospheric column, the lidar calculated median was 30.9 ± 4.7 DU, compared to 33.4 ± 3.9 DU for the GEOS-CF. This indicates a difference of 2.5 DU or 7.9 %, which is well within the lidar uncertainty of around 10 % throughout the tropospheric column, and as we described above is likely driven by select days rather than an overall bias between the measurements~~.~~. For the 0-2 km tropospheric column, the lidar calculated median was 5.8 DU ± 0.9 DU, compared to ~~6.9~~7.8 DU ± 0.7 DU for the corresponding GEOS-CF measurements. ~~This indicates a difference between observations and model of 1.1 DU or 18.9 %, which is higher than the lidar uncertainty of around 5-10 % throughout the column.~~ For the TROLIX-19 campaign, a 0-2 km tropospheric column accounts for approximately 20% of the tropospheric column as detailed in Figure 4 (top panel), indicating measurements above the surface are critically needed at understanding ozone variability at rural sites such as Cabauw, NL, where free tropospheric ozone features dominate the column.

4 Full Profile Ozone Comparisons

4.1 Hybrid Tropospheric/Stratospheric Ozone Comparisons

To better understand differences in ozone retrievals from multiple platforms, it is important to assess the entire vertical distribution of ozone. To characterize the vertical distribution throughout the entire troposphere and stratosphere, hybrid ozone profiles were created from longer (integrations of 60-120 ~~minute~~minutes vs 10 minutes in Sec 3) temporal retrievals from the closed co-located daytime/nighttime TROPOZ and nighttime STROZ ~~lidars~~lidar data, which were then interpolated to the GEOS-CF model vertical grid levels. **Figure 5** compares these results to the GEOS-CF, OMPS-LP, TROPOMI, MLS and the ECC ozonesonde profiles for 12 Sep, 17 Sep, 19 Sep, and 21 Sep 2019. These days were selected as days within the campaign that had an ECC launch from De Bilt, NL (30 km from Cabauw).

For each observation period in **Figure 5**, all platforms manage to characterize a similar shape and extent of the ozone maxima between 2.5-4.5 molecules m^{-3} throughout the vertical layer between 20-25 km. In each case, there are differences between the platforms in characterizing the vertical variability and extent of the ozone maxima, which will be quantified in the following section. One notable feature that emphasizes the cross-platform ability to illustrate ozone variability in the stratosphere is from the 19 and 21 Sep profiles. A dual ozone maximum is observed quite remarkably by the merged lidar, ECC, MLS, OMPS-LP and simulated by the GEOS-CF centered around 20 km and then again at 25km. The wind observations from the ozonesonde payload (not shown) indicate a wind shear within the two ozone layers, suggesting this feature was dynamically driven. The TROPOMI retrieval is not able to retrieve this vertical features due to its coarser vertical resolution and appears to average through the layers.

4.2 Difference Profiles

To quantitatively compare the ozone retrievals and simulations, **Figure 6** displays the ozone values for the TROLIX-19 time period from the hybrid lidar dataset (**Figure 6a**), GEOS-CF (**Figure 6b**), OMPS-LP (**Figure 6c**), MLS (**Figure 6d**) and TROPOMI (**Figure 6e**). This double ozone maxima, starting after 20 September serves as a geophysical marker to visually compare the ozone products. The lidar, model, and OMPS-LP all capture this feature, but with varying ozone abundances and altitudes. From **Figure 6d**, it appears as if TROPOMI retrievals are not able to resolve this feature. The percent differences, as compared to the lidar observations, are displayed in **Figure 7a-d**. These percent differences are calculated using (1)

$$\text{(1) Percent Difference} = \frac{(E_1 - E_2)}{\frac{1}{2}(E_1 + E_2)} \times 100$$

where E_2 are the lidar observations and E_1 are the respective ozone values from the various platforms in **Figure 6**.

The percent differences in **Figure 7a** indicate the GEOS-CF model from 20-45 km generally represents the lidar observations, but are generally 0-10 % lower in abundance. The percent differences in **Figure 7b** indicate OMPS-LP is also representing the ozone maxima and altitude above 25 km. There are larger differences below 20 km, which indicates the OMPS-LP retrieval ~~is generally underestimating the~~worsens (in both directions) as compared to the ozone abundance below 20 km as shown in the profiles in **Figure 5**. The percent differences in **Figure 7c** indicate the MLS data, especially that within the 20-40km region, perform quite well as compared to the lidar observations. The percent differences in **Figure 7d** indicate the TROPOMI retrieval is generally over representing the ozone concentrations throughout the atmosphere, which worsens within the troposphere~~.~~ and has been discussed earlier for the tropospheric ozone column as a result of a much larger vertical resolution in this region. In all cases, the most variability in the differences occur within the active region from 10-20 km that is driven by the dynamical tropopause height and lower stratospheric ozone abundance. Within each satellite dataset, we find larger biases in the lower stratosphere and upper troposphere below 18km, which has been previously described in the literature for the OMPS-LP dataset in Kramarova *et al.*, 2018 and were improved in the updated version 2.5 algorithm used in this work.

5 Total Column Ozone

Similar to the troposphere, to better understand to what extent the vertical distribution of ozone impacts the atmospheric column, **Figure 8 (top panel)** shows the various platforms and their retrieved total column ozone. For this analysis, the GEOS-CF, lidar, OMPS-NM, TROPOMI (GODFIT) are shown, in addition to local ground-based measurements from a Pandora instrument and Brewer. The total column values range from 230-300 DU throughout the campaign period, with the median total column ozone of 271 DU. With the previous analyses from Sec 3.2, this indicates the median total tropospheric column of

33 DU and 0-2km boundary layer column of 6 DU result in percentages of the entire ozone column of 12% and 2.3%, respectively. Similar to the full tropospheric ozone columns, larger total ozone columns were observed towards the end of the TROLIX-19 period, suggesting this variability was partly due to a larger abundance of ozone in the lower stratosphere.

Figure 8 (bottom panel) shows the various platforms as a percent differences from the model. In general, the various platforms are all mostly within ~~10~~5 % of each other, with most differences being within ±~~5~~3%. This analysis emphasizes the stability and maturity of the Pandora and Brewer systems for monitoring the total column ozone amounts. Interestingly, the double maxima feature in vertical ozone distribution in the stratosphere (with local minima between) described in Sec 4.1 on 21 Sep does not severely impact the total column ozone.

6 Conclusions

This work has highlighted the various differences in retrieved ozone quantities during the TROLIX-19 campaign. This has emphasized the importance of ground-based ozone lidars and other measurements in understanding the vertical variability of ozone and how it relates to the column reduction. This work also shows the first effort to directly resolve both tropospheric columns and 0-2km ozone columns from the NASA TROPOZ lidar. Other TOLNet lidars are able to perform this data reduction and future work will be to expand this effort to the other TOLNet locations. This work indicates the level of performance of the GEOS-CF modeling system as compared to the other platforms, which ultimately performs extremely well both in the stratosphere (Figure 6 and Figure 7) and within the troposphere~~, as emphasized~~ (Figure 2 and Figure 4).

One takeaway message or point of caution for future efforts is that although there are situations identified where the vertical profile and the model disagree in ~~**Figure 6** and **Figure 7**.~~ a certain altitude range (Figure 3), when the data is reduced to a columnar product, compensating over/under-estimations may cancel out and produce a more accurate value when only looking at the resultant as compared to observations. For this reason, it is essential when doing data columnar reduction for the troposphere, and even more so in the 0-2km column or planetary boundary layer, that observations of the vertical profile be used to evaluate the representativeness of the model and auxillary data sets.

In looking towards the NASA TEMPO ~~mission~~, this work indicates that the GEOS-CF, with its global coverage, hourly resolution, and adequate vertical information to resolve most atmospheric features, is an appropriate choice for the a priori profiles for the TEMPO ozone retrievals. Continued investigations are needed with high resolution observations, as presented in this work, to better evaluate the GEOS-CF, especially in these common transport regions of the atmosphere. Although the GEOS-CF performed well in reproducing the ozone downward transport throughout the upper troposphere and lower stratosphere, the model did fail to resolve some high-resolution laminae deeper into the lower troposphere related to specific mesoscale ozone transport in this region as evidenced in Figure 2 and Figure 3.

This work shows the TROPOMI TOPAS ozone profile algorithm products are able to accurately reproduce ozone quantities in the lower troposphere at various atmospheric levels. In particular, **Figure 3** ~~shows~~and Figure 4 show promising results that indicate the TROPOMI satellite observations compare well with the observations from ground-based measurements (lidar, sonde) of specific elevated ozone features. However, there is an observed overestimate of the TROPOMI retrieval in the upper troposphere and lower stratosphere (between 10 and 15 km) associated with a larger vertical resolution that needs to also be further evaluated to better understand the representativeness of the retrieval in this region.

Figure 7 was presented as a quantitative resultant figure to illustrate both the temporal (*i.e.,* throughout the course of the TROLIX-19 campaign) and vertical differences observed in the retrievals from each observational platform. This serves as a rare opportunity to cross-evaluate multiple satellited based observations, a global chemical transport model, ozonesondes and a high-resolution ozone lidar suite.

The authors feel that this figure has served to point out the strengths of each platform and present careful considerations for areas of under/over estimation, Furthermore, we feel reducing these comparisons down to a specific percentage may underserve the community push for supporting the vertical profiling needed for these types of efforts.

The CESAR Observatory continues to be a critical landmark for campaigns that revolve around atmospheric composition measurements for satellite validation and evaluation beyond this effort, such as CINDI and CINDI-2 (Kreher *et al.*, 2020; Wang et al., 2020; Tirpitz *et al.,* 2021). As the European Commission (EC) in partnership with the European Space Agency (ESA) continues to launch tropospheric composition satellites, including the upcoming geo-stationary Sentinel-4 satellite, we expect this observatory will continue to host and maintain critical atmospheric sampling for future validation efforts.

Data Availability.

1. MLS ozone profiles can be downloaded from the NASA Goddard Space Flight Center Earth Sciences Data and Information Services Center (GES DISC; Schwartz et al., 2020, https://doi.org/10.5067/Aura/MLS/DATA2516, last access: 29 March 2022).
2. The Pandora data is available at the Pandonia Global Network Archive (http://data.pandonia-global-network.org/Cabauw/, last access 29 March 29, 2022).
3. The OMPS LP version 2.5 ozone profiles can be downloaded from the NASA Goddard Space Flight Center Earth Sciences Data and Information Services Center (GES DISC; at https://doi.org/10.5067/X1Q9VA07QDS7 (Deland, 2017, last access: 29 March 2022).
4. The tropospheric ozone lidar data used in this publication were obtained from the Cabauw Experimental Site for Atmospheric Research (CESAR) as part of a campaign involving the

Network for the Detection of Atmospheric Composition Change (NDACC) and NASA's Tropospheric Ozone Lidar Network (TOLNet) and are publicly available (https://www-air.larc.nasa.gov/cgi-bin/ArcView.1/TOLNet?NASA-GSFC=1, last access: 29 March 2022).

5. The ozonesonde and Brewer data used in this publication were obtained from the De Bilt, NL site as part of a campaign involving the Network for the Detection of Atmospheric Composition Change (NDACC) and are publicly available (ftp://ftp.cpc.ncep.noaa.gov/ndacc/station/debilt/, last access: 29 March 2022).
6. The stratospheric ozone lidar data used in this publication were obtained from the Cabauw Experimental Site for Atmospheric Research (CESAR) as part of a campaign involving the Network for the Detection of Atmospheric Composition Change (NDACC) and are publicly available (ftp://ftp.cpc.ncep.noaa.gov/ndacc/station/cabauw/, last access: 29 March 2022).
7. The TROPOMI TOPAS Ozone Profile data and source codes are available upon request from Nora Mettig (mettig@iup.physik.uni-bremen.de) or Mark Weber (weber@uni-bremen.de). The L1B version of the S5P data is available upon request to the S5P Validation Team.
8. The Tropospheric Ozone Column from OMPS-NM/MERRA-2 Daily measurements data are available upon request from Jerry Ziemke (Jerald.r.ziemke@nasa.gov).
9. The NASA GEOS-CF simulations are available at the data sharing portal (https://portal.nccs.nasa.gov/datashare/gmao/geos-cf/v1/forecast/, last access 29 March 2022).

Author contributions. JS drafted the original manuscript. JS, LT, GS, and TM deployed and operated the NASA ozone lidars and provided expertise on use of measurements. NM, AR, and MW provided TOPAS ozone profile data and guidance on how best to use the measurements. AA and KK provided overall context as principal investigators of the TROLIX-19 campaign and coordinated science team meetings to foster this collaboration. KEK provided GEOS-CF data and insight on its use in this work. MA, AP, MvR, and PV provided expertise and data for the ground-observations for ozonesondes, Brewer, and historical data for the Cabauw site. JZ provided data for the OMPS-MERRA-2 tropospheric column data. NK provided Aura MLS, OMPS-LP merged data and further insight into the use of the data.

Competing interests. The authors declare that they have no conflict of interest.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Acknowledgements. NASA data has been provided through the Tropospheric Composition and Upper Atmosphere Research Programs. We acknowledge all additional data providers and their funding agencies for performing regular measurements and retrievals.

References

Apituley, Arnoud, Karin Kreher, Michael Van Roozendael, John Sullivan, Thomas J. McGee, Marc Allaart, Ankie Piters et al. "Overview of activities during the 2019 TROpomi vaLIdation eXperiment (TROLIX'19)." In *AGU Fall Meeting Abstracts*, vol. 2019, pp. A43J-2958. 2019

Apituley, Arnoud, Karin Kreher, Ankie Piters, John Sullivan, Michel vanRoozendael, Tim Vlemmix, Mirjam den Hoed et al. "Overview of the 2019 Sentinel-5p TROpomi vaLIdation eXperiment (TROLIX)." In *EGU General Assembly Conference Abstracts*, p. 10539. 2020.

Arosio, Carlo, Alexei Rozanov, Elizaveta Malinina, Kai-Uwe Eichmann, Thomas von Clarmann, and John P. Burrows. "Retrieval of ozone profiles from OMPS limb scattering observations." Atmospheric Measurement Techniques 11, no. 4 (2018): 2135-2149.

Copernicus Sentinel data processed by ESA, German Aerospace Center (DLR) (2019), Sentinel-5P TROPOMI Total Ozone Column 1-Orbit L2 5.5km x 3.5km, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed: [10 December 2021], 10.5270/S5P-fqouvyz

Dacic, Natasha, John T. Sullivan, K. Emma Knowland, Glenn M. Wolfe, Luke D. Oman, Timothy A. Berkoff, and Guillaume P. Gronoff. "Evaluation of NASA's high-resolution global composition simulations: Understanding a pollution event in the Chesapeake Bay during the summer 2017 OWLETS campaign." *Atmospheric Environment* 222 (2020): 117133.

Flynn, L. E., Seftor, C. J., Larsen, J. C., and Xu, P.: The Ozone Mapping and Profiler Suite, in: Earth Science Satellite Remote Sensing, edited by: Qu, J. J., Gao,W., Kafatos, M., Murphy, R. E., and Salomonson, V. V., Springer, Berlin, 279–296, doi:10.1007/978-3-540-37293-6, 2006.

Garane, Katerina, Maria-Elissavet Koukouli, Tijl Verhoelst, Christophe Lerot, Klaus-Peter Heue, Vitali Fioletov, Dimitrios Balis et al. "TROPOMI/S5P total ozone column data: global ground-based validation and consistency with other satellite missions." *Atmospheric Measurement Techniques* 12, no. 10 (2019): 5263-5287.

Gronoff, G., T. Berkoff, K. E. Knowland, L. Lei, M. Shook, B. Fabbri, W. Carrion, and A. O. Langford. "Case study of stratospheric Intrusion above Hampton, Virginia: lidar-observation and modeling analysis." *Atmospheric Environment* (2021): 118498.

Hubert, Daan, Klaus-Peter Heue, Jean-Christopher Lambert, Tijl Verhoelst, Marc Allaart, Steven Compernolle, Patrick D. Cullis et al. "TROPOMI tropospheric ozone column data: geophysical assessment and comparison to ozonesondes, GOME-2B and OMI." *Atmospheric Measurement Techniques* 14, no. 12 (2021): 7405-7433.

Johnson, M. S., Liu, X., Zoogman, P., Sullivan, J., Newchurch, M. J., Kuang, S., Leblanc, T., and McGee, T.: Evaluation of potential sources of a priori ozone profiles for TEMPO tropospheric ozone retrievals, Atmos. Meas. Tech., 11, 3457–3477, https://doi.org/10.5194/amt-11-3457-2018, 2018.

Keller, Christoph A., K. Emma Knowland, Bryan N. Duncan, Junhua Liu, Daniel C. Anderson, Sampa Das, Robert A. Lucchesi et al. "Description of the NASA GEOS Composition Forecast Modeling System GEOS-CF v1. 0." *Journal of Advances in Modeling Earth Systems* 13, no. 4 (2021): e2020MS002413.

Kramarova, Natalya A., Pawan K. Bhartia, Glen Jaross, Leslie Moy, Philippe Xu, Zhong Chen, Matthew DeLand et al. "Validation of ozone profile retrievals derived from the OMPS LP version 2.5 algorithm against correlative satellite measurements." *Atmospheric Measurement Techniques* 11, no. 5 (2018): 2837-2861.

Kreher, Karin, Michel Van Roozendael, Francois Hendrick, Arnoud Apituley, Ermioni Dimitropoulou, Udo Frieß, Andreas Richter et al. "Intercomparison of NO_2, O_4, O_3 and HCHO slant column measurements by MAX-DOAS and zenith-sky UV–visible spectrometers during CINDI-2." *Atmospheric Measurement Techniques* 13, no. 5 (2020): 2169-2208.

Lamsal, L. N., M. Weber, S. Tellmann, and J. P. Burrows. "Ozone column classified climatology of ozone and temperature profiles based on ozonesonde and satellite data." Journal of Geophysical Research: Atmospheres 109, no. D20 (2004).

Leblanc, Thierry, Mark A. Brewer, Patrick S. Wang, Maria Jose Granados-Muñoz, Kevin B. Strawbridge, Michael Travis, Bernard Firanski et al. "Validation of the TOLNet lidars: the Southern California Ozone Observation Project (SCOOP)." *Atmospheric measurement techniques* 11, no. 11 (2018): 6137-6162.

Leblanc, Thierry, Robert J. Sica, Joanna AE Van Gijsel, Sophie Godin-Beekmann, Alexander Haefele, Thomas Trickl, Guillaume Payen, and Gianluigi Liberti. "Proposed standardized definitions for vertical resolution and uncertainty in the NDACC lidar ozone and temperature algorithms–Part 2: Ozone DIAL uncertainty budget." Atmospheric Measurement Techniques 9, no. 8 (2016): 4051-4078.

Livesey, N. J., M. J. Filipiak, L. Froidevaux, W. G. Read, A. Lambert, M. L. Santee, J. H. Jiang et al. "Validation of Aura Microwave Limb Sounder O3 and CO observations in the upper troposphere and lower stratosphere." Journal of Geophysical Research: Atmospheres 113, no. D15 (2008).

Malderen, Roeland Van, Marc AF Allaart, Hugo De Backer, Herman GJ Smit, and Dirk De Muer. "On instrumental errors and related correction strategies of ozonesondes: possible effect on calculated ozone trends for the nearby sites Uccle and De Bilt." *Atmospheric Measurement Techniques* 9, no. 8 (2016): 3793-3816.

McGee, Thomas J., David N. Whiteman, Richard A. Ferrare, James J. Butler, and John F. Burris. "STROZ LITE: stratospheric ozone lidar trailer experiment." *Optical Engineering* 30, no. 1 (1991): 31-39.

McPeters, Richard D., Gordon J. Labow, and Jennifer A. Logan. "Ozone climatological profiles for satellite retrieval algorithms." *Journal of Geophysical Research: Atmospheres* 112, no. D5 (2007).

Mettig, N., Weber, M., Rozanov, A., Arosio, C., Burrows, J. P., Veefkind, P., Thompson, A. M., Querel, R., Leblanc, T., Godin-Beekmann, S., Kivi, R., and Tully, M. B.: Ozone profile retrieval from nadir TROPOMI measurements in the UV range, Atmos. Meas. Tech., 14, 6057–6082, https://doi.org/10.5194/amt-14-6057-2021, 2021.

Mettig, Nora, Mark Weber, Alexei Rozanov, John P. Burrows, Pepijn Veefkind, Nadia Smith, Anne M. Thompson et al. "Combined UV and IR ozone profile retrieval from TROPOMI and CrIS measurements." *Atmospheric Measurement Techniques Discussions* (2021): 1-33.

Piters, A. J. M., Boersma, K. F., Kroon, M., Hains, J. C., Van Roozendael, M., Wittrock, F., Abuhassan, N., Adams, C., Akrami, M., Allaart, M. A. F., Apituley, A., Beirle, S., Bergwerff, J. B., Berkhout, A. J. C., Brunner, D., Cede, A., Chong, J., Clémer, K., Fayt, C., Frieß, U., Gast, L. F. L., Gil-Ojeda, M., Goutail, F., Graves, R., Griesfeller, A., Großmann, K., Hemerijckx, G., Hendrick, F., Henzing, B., Herman, J., Hermans, C., Hoexum, M., van der Hoff, G. R., Irie, H., Johnston, P. V., Kanaya, Y., Kim, Y. J., Klein Baltink, H., Kreher, K., de Leeuw, G., Leigh, R., Merlaud, A., Moerman, M. M., Monks, P. S., Mount, G. H., Navarro-Comas, M., Oetjen, H., Pazmino, A., Perez-Camacho, M., Peters, E., du Piesanie, A., Pinardi, G., Puentedura, O., Richter, A., Roscoe, H. K., Schönhardt, A., Schwarzenbach, B., Shaiganfar, R., Sluis, W., Spinei, E., Stolk, A. P., Strong, K., Swart, D. P. J., Takashima, H., Vlemmix, T., Vrekoussis, M., Wagner, T., Whyte, C., Wilson, K. M., Yela, M., Yilmaz, S., Zieger, P., and Zhou, Y.: The Cabauw Intercomparison campaign for Nitrogen Dioxide measuring Instruments (CINDI): design, execution, and early results, Atmos. Meas. Tech., 5, 457–485, doi:10.5194/amt-5-457-2012, 2012.

Smit, H. G. J., Thompson, A. M., & the Panel for the Assessment of Standard Operating Procedures for Ozonesondes, v2.0 (ASOPOS 2.0) (2021). Ozonesonde Measurement Principles and Best Operational Practices. World Meteorological Organization, GAW Report 268. [Available at https://library.wmo.int/doc_num.php?explnum_id=10884].

Smit, Herman GJ, Wolfgang Straeter, Bryan J. Johnson, Samuel J. Oltmans, Jonathan Davies, David W.Tarasick, Bruno Hoegger et al. "Assessment of the performance of ECC-ozonesondes under quasi-flight conditions in the environmental simulation chamber: Insights from the Juelich Ozone Sonde Intercomparison Experiment (JOSIE)." Journal of Geophysical Research: Atmospheres 112, no. D19 (2007).

Sullivan, J. T., T. J. McGee, G. K. Sumnicht, L. W. Twigg, and R. M. Hoff. "A mobile differential absorption lidar to measure sub-hourly fluctuation of tropospheric ozone profiles in the Baltimore–Washington, DC region." *Atmospheric Measurement Techniques* 7, no. 10 (2014): 3529-3548.

Sullivan, John T., Thomas J. McGee, Anne M. Thompson, R. Bradley Pierce, Grant K. Sumnicht, Laurence W. Twigg, Edwin Eloranta, and Raymond M. Hoff. "Characterizing the lifetime and occurrence of stratospheric-tropospheric exchange events in the rocky mountain region using high-resolution ozone measurements." *Journal of Geophysical Research: Atmospheres* 120, no. 24 (2015): 12410-12424.

Sullivan, John T., Timothy Berkoff, Guillaume Gronoff, Travis Knepp, Margaret Pippin, Danette Allen, Laurence Twigg et al. "The ozone water–land environmental transition study: An innovative strategy for understanding Chesapeake Bay pollution events." Bulletin of the American Meteorological Society 100, no. 2 (2019): 291-306.

Tirpitz, Jan-Lukas, Udo Frieß, François Hendrick, Carlos Alberti, Marc Allaart, Arnoud Apituley, Alkis Bais et al. "Intercomparison of MAX-DOAS vertical profile retrieval algorithms: studies on field data from the CINDI-2 campaign." *Atmospheric Measurement Techniques* 14, no. 1 (2021): 1-35.

van Geffen, J.H.G.M., Eskes, H.J., Boersma, K.F., Maasakkers, J.D. and Veefkind, J.P., TROPOMI ATBD of the total and tropospheric NO2 data products, Report S5P-KNMI-L2-0005-RP, KNMI, De Bilt, The Netherlands; see TROPOMI ATBD list for the latest available version.

Wang, Yang, Arnoud Apituley, Alkiviadis Bais, Steffen Beirle, Nuria Benavent, Alexander Borovski, Ilya Bruchkouski et al. "Inter-comparison of MAX-DOAS measurements of tropospheric HONO slant column densities and vertical profiles during the CINDI-2 campaign." Atmospheric Measurement Techniques 13, no. 9 (2020): 5087-5116.

Wenig, Mark O., A. M. Cede, E. J. Bucsela, E. A. Celarier, K. F. Boersma, J. P. Veefkind, E. J. Brinksma, J. F. Gleason, and J. R. Herman. "Validation of OMI tropospheric NO2 column densities using direct-Sun mode Brewer measurements at NASA Goddard Space Flight Center." *Journal of Geophysical Research: Atmospheres* 113, no. D16 (2008).

Wing, Robin, Wolfgang Steinbrecht, Sophie Godin-Beekmann, Thomas J. McGee, John T. Sullivan, Grant Sumnicht, Gérard Ancellet, Alain Hauchecorne, Sergey Khaykin, and Philippe Keckhut. "Intercomparison and evaluation of ground-and satellite-based stratospheric ozone and temperature profiles above Observatoire de Haute-Provence during the Lidar Validation NDACC Experiment (LAVANDE)." *Atmospheric Measurement Techniques* 13, no. 10 (2020): 5621-5642.

Wing, Robin, Sophie Godin-Beekmann, Wolfgang Steinbrecht, Thomas J. Mcgee, John T. Sullivan, Sergey Khaykin, Grant Sumnicht, and Laurence Twigg. "Evaluation of the new DWD ozone and temperature lidar during the Hohenpeißenberg Ozone Profiling Study (HOPS) and comparison of results with previous NDACC campaigns." *Atmospheric Measurement Techniques* 14, no. 5 (2021): 3773-3794.

Ziemke, Jerry R., Luke D. Oman, Sarah A. Strode, Anne R. Douglass, Mark A. Olsen, Richard D. McPeters, Pawan K. Bhartia et al. "Trends in global tropospheric ozone inferred from a composite record of TOMS/OMI/MLS/OMPS satellite measurements and the MERRA-2 GMI simulation." *Atmospheric Chemistry and Physics* 19, no. 5 (2019): 3257-3269.

Table 1: Instrument platforms, associated products, and short description used in this work during the TROLIX-19 campaign.

Instrument	Products	Platform	Description
GSFC TROPOZ [NASA]	Profiles [0.2 – 18 km]	Ground-based Lidar	10 min integration; 30-90-min avg around ECC or Satellite Overpass
GSFC STROZ [NASA]	Profiles [15 - 48 km]	Ground-based Lidar	~2-4-hr avg between (20-23 UT)
ECC Ozonesondes [KNMI]	Profiles [0 – 33 km]	Balloonborne	Balloonborne, Launched at 12 UT from De Bilt (~30 km from Cabauw) on 4 days
Pandora [NASA/KNMI]	Column [TCO]	Spectrometer	L2 Pandora 118s, Data Used has QC/QA Flags = 10
Brewer [KNM]	Column [TCO]	Spectrophotometer	L2 Brewer #189m, MKIII, Located in De Bilt
S5P/TropOMI [ESA]	Column [TCL]	Satellite	L2 TOPAS Product, Overpass between 12-14 UT (5.5x3.5 km, nadir)
S5P/TropOMI [KNMI]	Column [TCO]	Satellite	L2 GODFIT v4 TO3 Product, Overpass between 12-14 UT (5.5x3.5 km, nadir)
OMPS [NASA]	Column [TCO]	Satellite	L3 NM Product, Version 2, Daily Overpass between 12-14 UT (50x50 km, nadir)
OMPS-LP [NASA]	Profiles [12-60km]	Satellite	Merged L2 v2.5 Daily Merged Product, Overpass between 12-14 UT (1km vertical bins)
OMPS/MERRA-2 [NASA]	Trop. Columns	Satellite/Assimilation	L4 Derived Product, OMPS-NM daily Overpass, MERRA-2
AURA MLS [NASA]	Profiles [12-60km]	Satellite	Merged L2 v5 Daily Daytime/Nighttime Products, Overpass between 12-14 UT (1km vertical bins) and 01-03 UT.
GEOS-CF [NASA]	Profiles [0-80km]	Global 3-D CCMM	1-Hr, 72 lev, Met. Replay, (25x25km) gmao.gsfc.nasa.gov/weather_prediction/GEOS-CF/

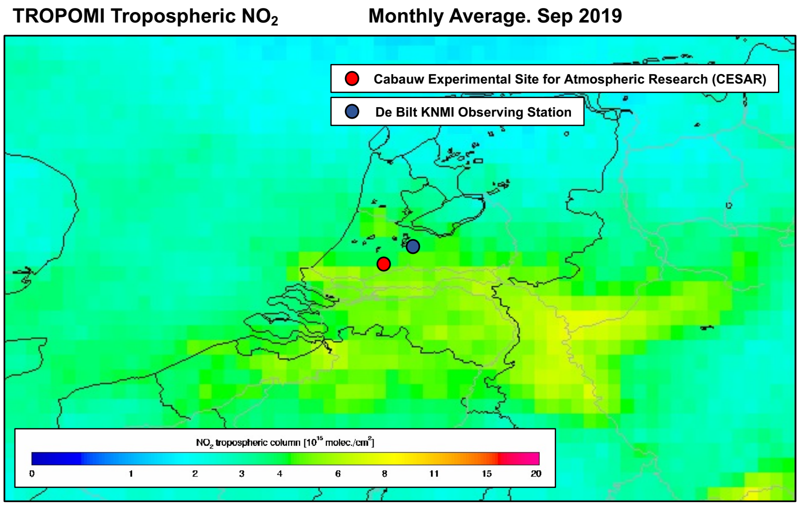


Figure 1: ~~Aqua/MODIS True Color Corrected Reflectance (left) and~~ TROPOMI ~~Tropospheric~~monthly-mean tropospheric NO_2 ~~(right~~column (version 1.0) for ~~14~~ September 2019. The CESAR ~~site is~~and De Bilt, NL sites are indicated in the image ~~on the left.~~ .

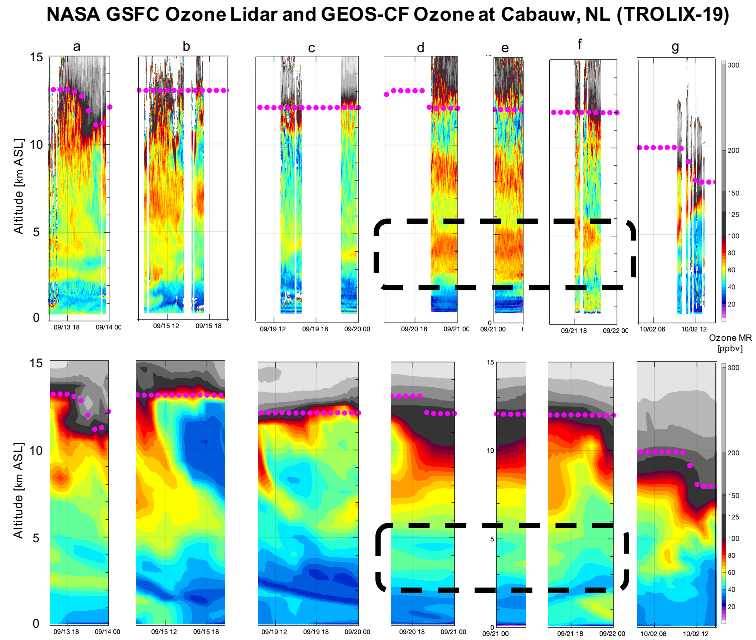


Figure 2: Cloud screened TROPOZ lidar retrievals (top panel) and the corresponding GEOS-CF model output (bottom panel) from the closest model grid cell to the CESAR observatory during TROLIX-19 for a) 13 Sep 14-00 UTC, b) 15 Sep 09-21 UTC, c) 19 Sep 10-00 UT, d) 20 Sep 16-00 UT, e) 21 Sep 0-3 UT, f) 21 Sep 16-00UT, and g) 02 Oct 04-14 UT. Pink dots are overlaid to indicate the simulated tropopause altitude based on a blended estimate (TROPPB).

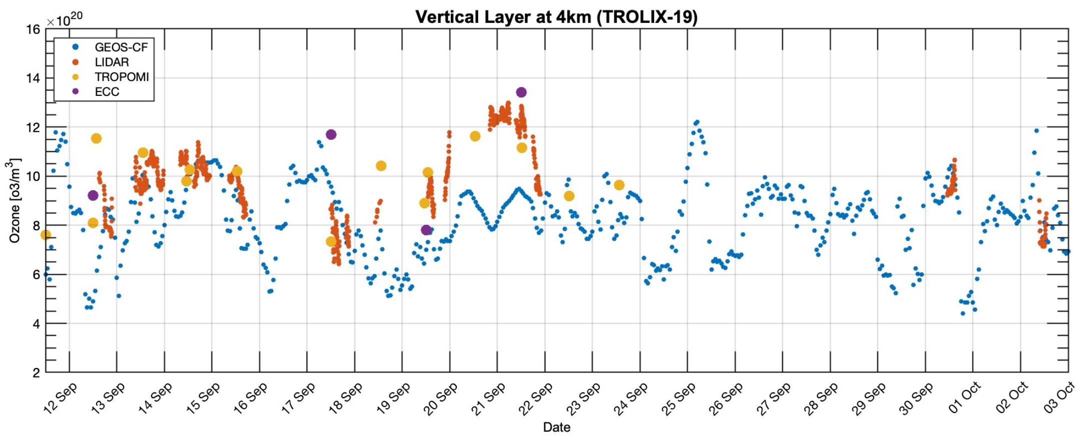


Figure 3: Ozone number density values for the TROPOZ lidar, GEOS-CF mode, TROPOMI and electro-chemical cell (ECC) ozonesondes at the 4km layers/levels. The layer was calculated to match the closest representative vertical layer of the GEOS-CF for consistent intercomparison. Data is averaged in a 500m layer from 3.94 km to 4.44 km AGL.

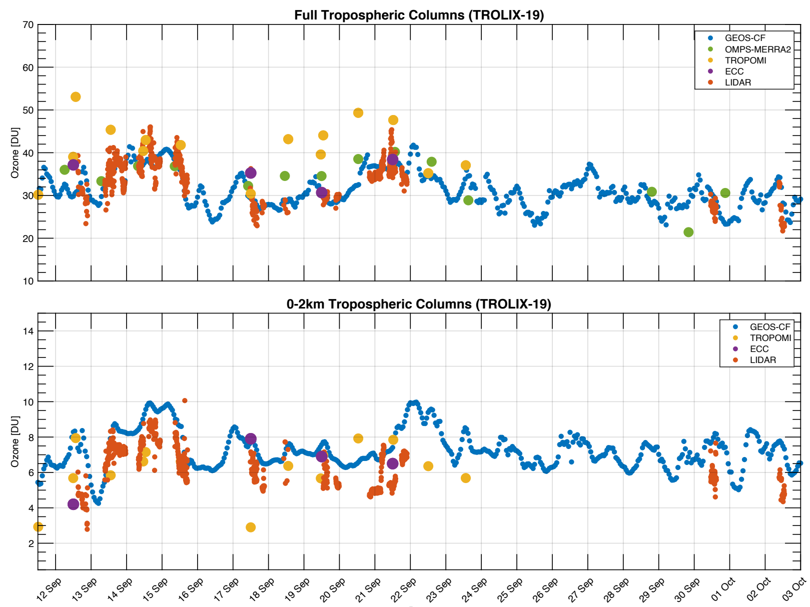


Figure 4: Full tropospheric columns (top panel) and 0-2km tropospheric columns (bottom panel) calculated from GEOS-CF, OMPS-MERRA2 (full column only), TROPOMI, Lidar and ECC. Data where reflectivity was greater than 0.6 was excluded to remove cloud interference.

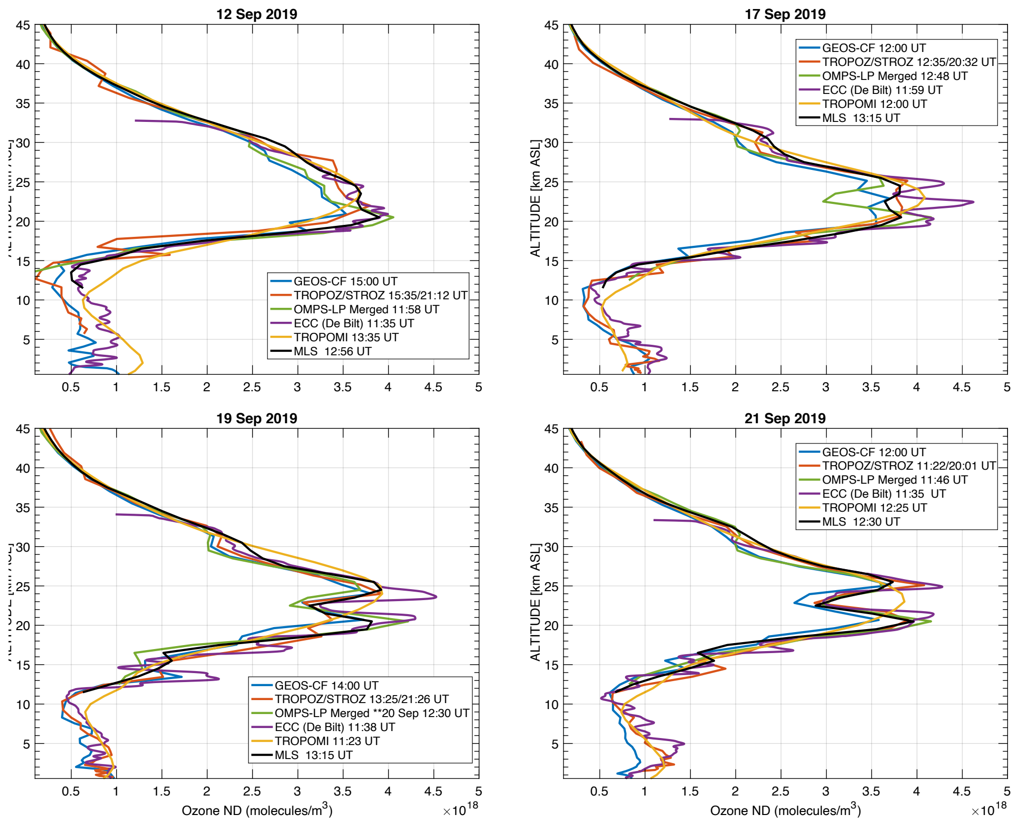


Figure 5: GEOS-CF, Lidar, OMPS-LP, ECC, TROPOMI, and MLS ozone profile comparisons for 12 Sep, 17 Sep, 19 Sep, and 21 Sep 2019. These days were selected as days within the campaign that had an ECC launch from De Bilt.

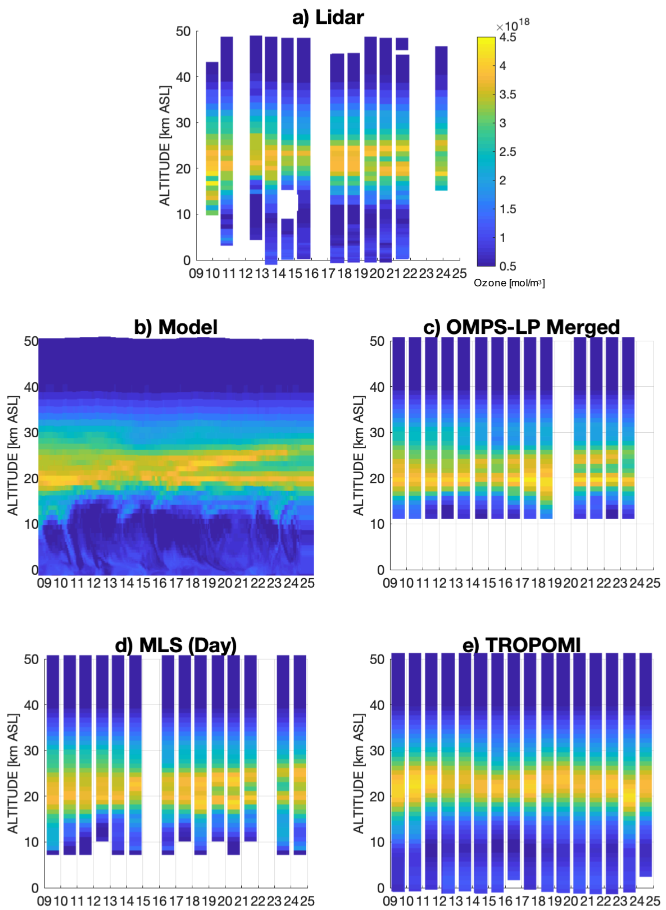


Figure 6: Ozone number densities across all platforms for the TROLIX-19 time period from the hybrid lidar dataset (Figure 6a), GEOS-CF (Figure 6b), OMPS-LP (Figure 6c), MLS (Figure 6d), TROPOMI (Figure 6e). The x-axis as day of September 2019.

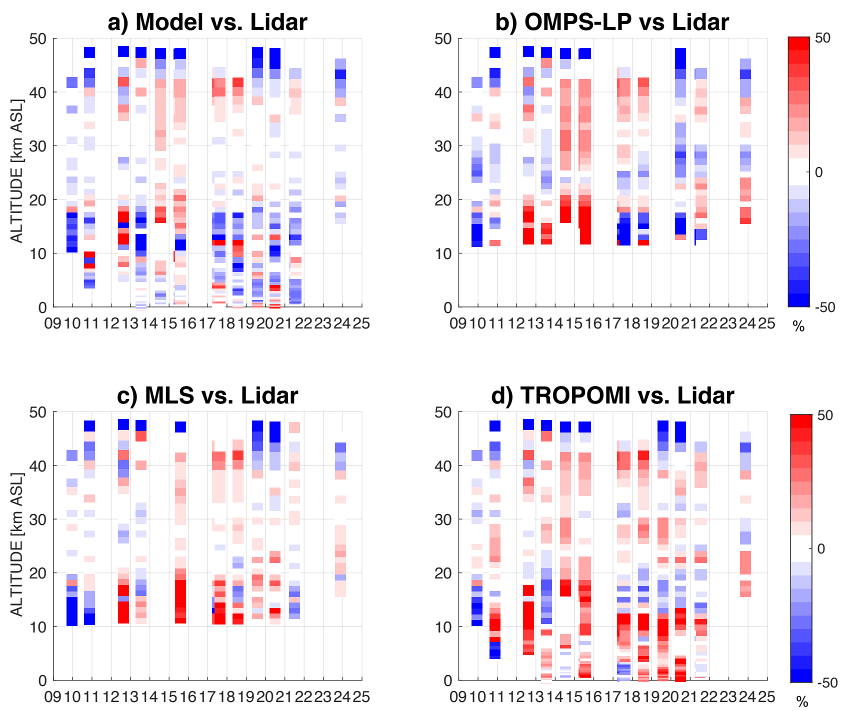


Figure 7: Differences in ozone number densities across all platforms for the TROLIX-19 time period for Model (Figure 7a), OMPS-LP (Figure 7b), MLS (Figure 7c), and TROPOMI (Figure 7d). The x-axis as day of September 2019.

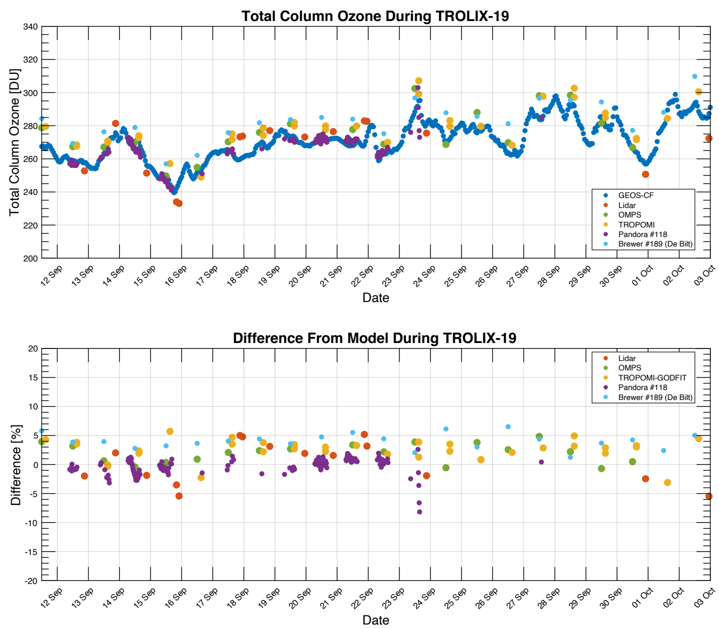


Figure 8: Total Ozone columns (top panel) and percent differences (bottom panel) as compared to the model observations for GEOS-CF, lidar, OMPS, TROPOMI, Pandora, and Brewer.