Supplementary Information

This documents presents several plots to support the material displayed in the manuscript.

Figure S1 displays the annual mean climatology for the selected TOR data sets. In the title of each subplot, we report the mean global TrOC values and their standard deviations.
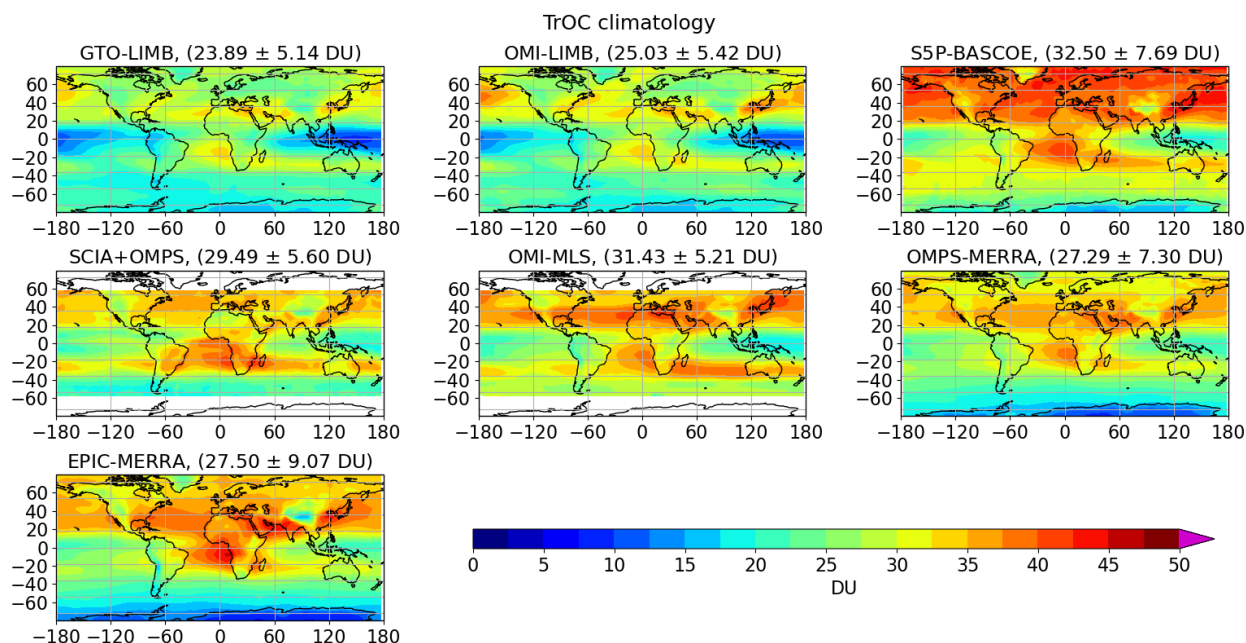

Figure S1: Tropospheric ozone climatology of the seven satellite data sets used for this study. In parenthesis the respective global averages and standard deviations are indicated.

We notice an overall bias present between the maps: global average ozone values are reported in the titles of the subplots and they differ by up to 8-10 DU (from 23.9 DU for GTO-LIMB to 32.5 for S5P-BASCOE, with high TrOC values especially in the northern hemisphere). These biases have several possible reasons: e.g., the TOC and SOC biases between different satellites, the tropopause definition adopted to construct the TOR product, the climatology used to fill stratospheric profile gaps, the criteria used for the subtraction between TOC and SOC. Fig. 1 in the manuscript showcases the biases in terms of zonal means.

Regarding the drop in TrOC occurred after 2020, especially in OMPS-MERRA and EPIC-MERRA, Fig. S2 shows differences of MERRA-2 temperature (panels a and b) and ozone (panels c and d) w.r.t. ERA5 and MLS at 20.5 km. We chose this altitude to check discontinuities in the lower stratosphere (large SOC contribution). No jump in 2020 is visible in these plots, indicating that the jump is not directly related to the reanalysis used to construct these data sets.
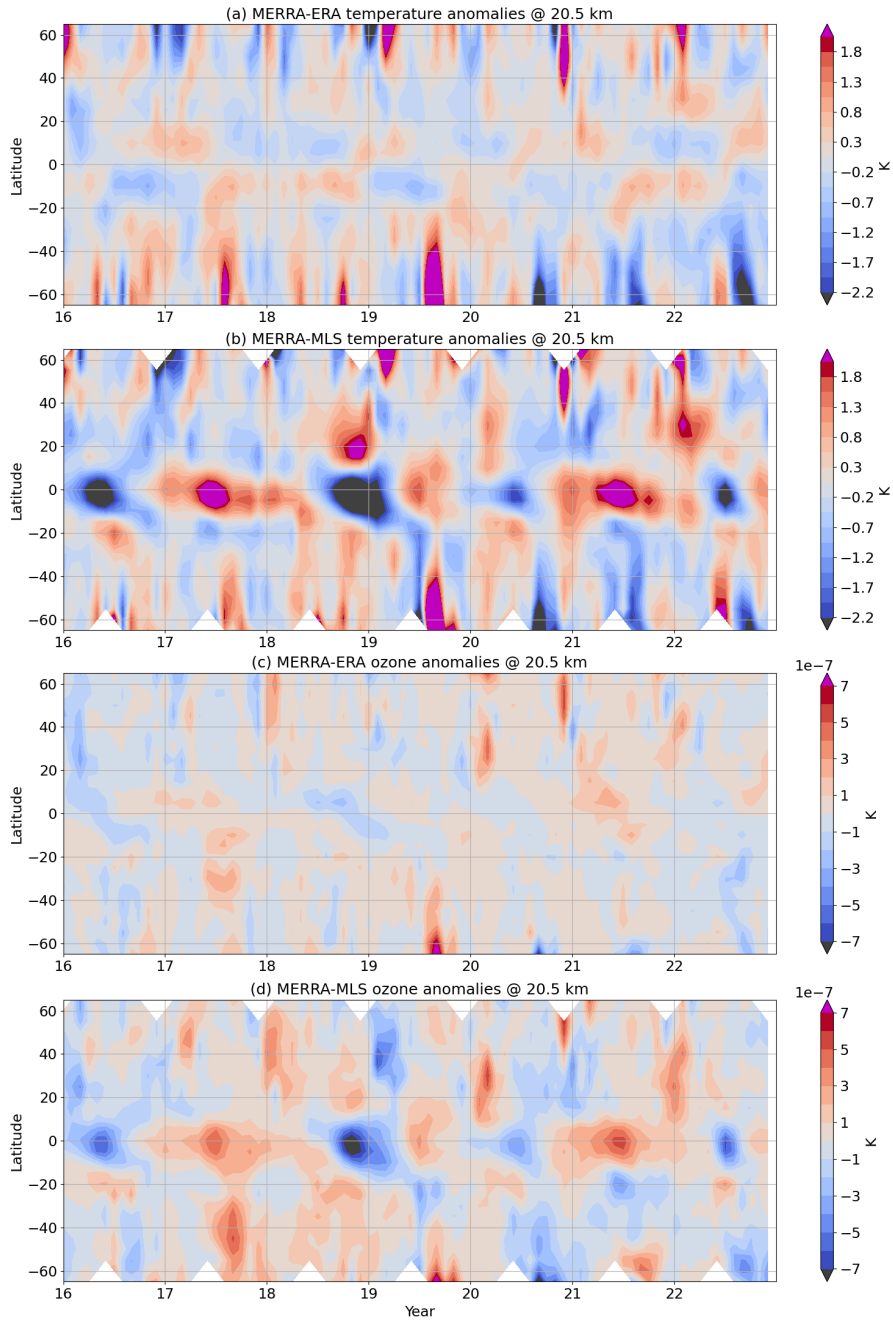
Figure S2: Time series of differences between MERRA-2 monthly temperature and ozone w.r.t. ERA5 (panels a and c) and MLS (panels b and d).

Figure S3 shows TrOC anomalies as in Fig.4 of the manuscript, when the seasonal cycle used to obtain anomalies is computed over the 2016-2019 period only, instead that over each respective whole time frame. Figure S4 displays the time series of the data sets at 42.5°N in terms of anomalies, when subtracting the seasonal cycle computed over the respective time period (top panel) or only over 2016-2019 (bottom panel). We can better identify the drop in 2020 and the lower values of the anoamlies over the 2020-2022 period, especially if the 2016-2019 seasonal cycle is subtracted. The only data set not showing any drop is S5P-BASCOE.
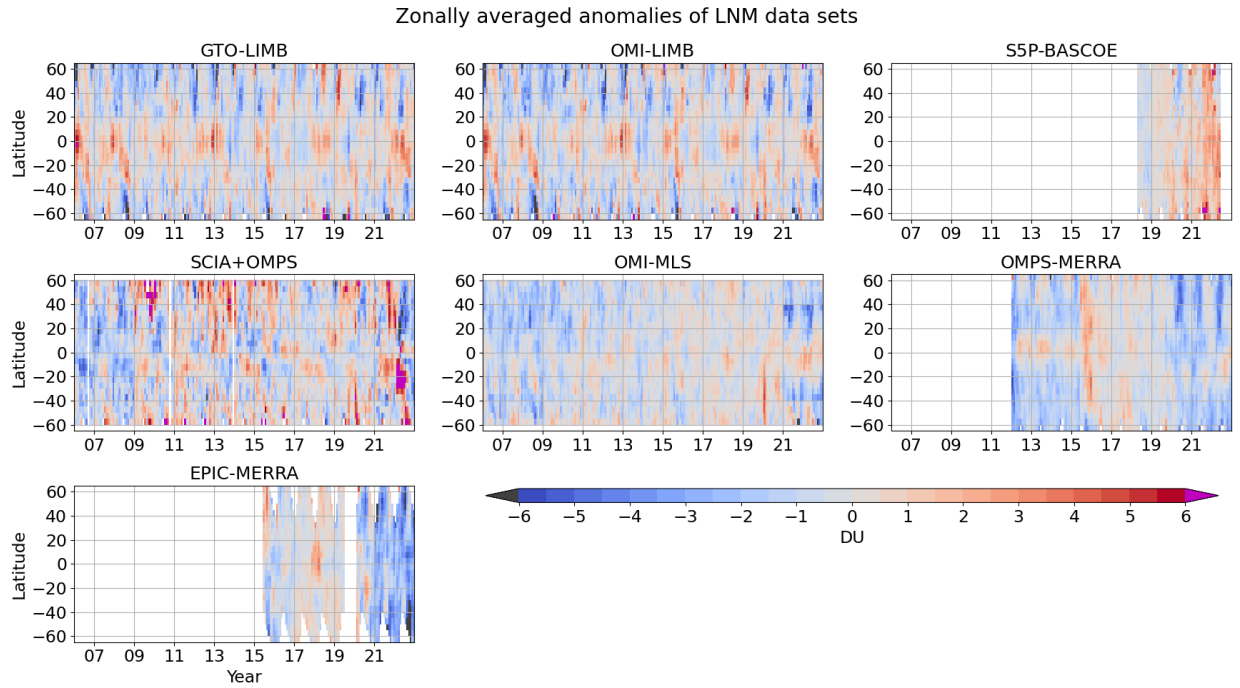
Zonally averaged anomalies of LNM data sets

Figure S3: De-seasonalized time series (using the 2016-2019 seasonal cycles) as a function of latitude for the seven data sets.
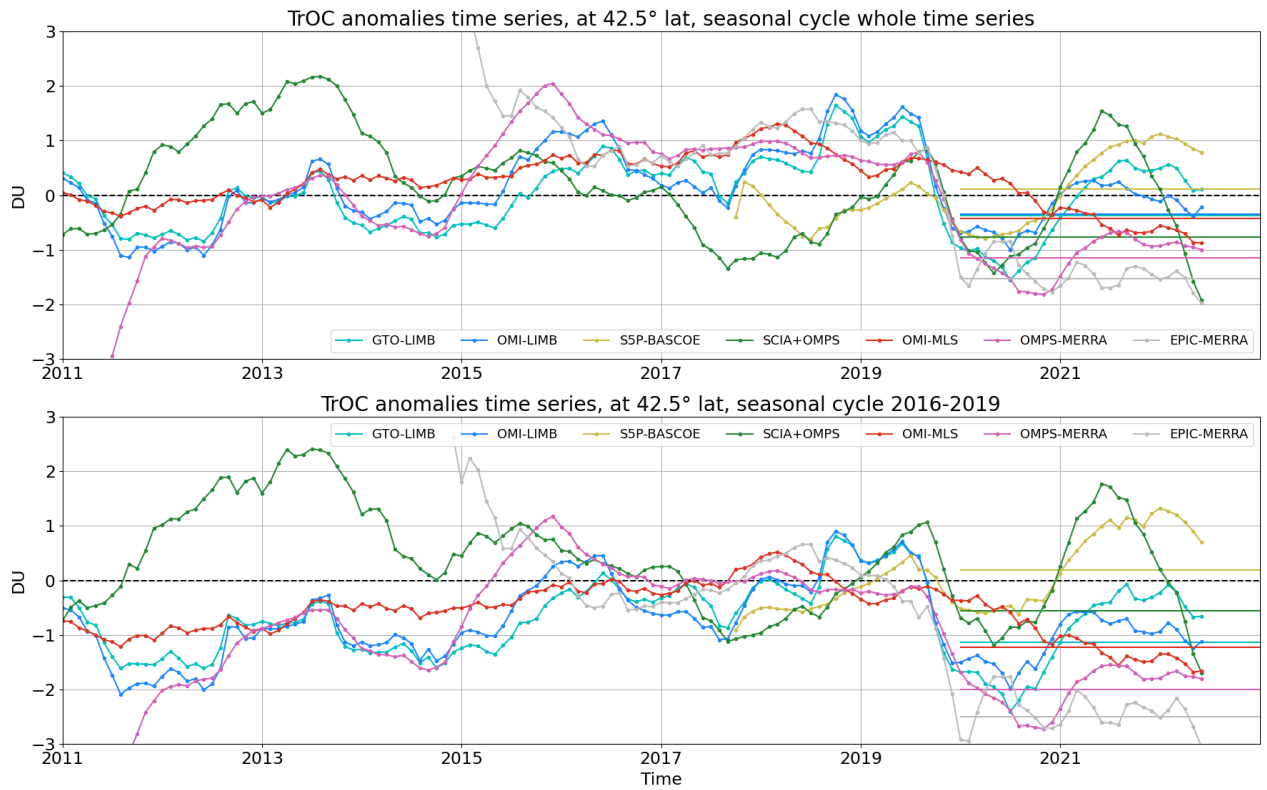

Figure S4: Deseasonalized time series at 42.5°N, using the seasonal cycles computed over the respective time series (top panel) or over 2016-2019 (bottom panel).

Figure S5 shows the time series of the TPH absolute anomalies for the selected data sets as a function of latitude, yearly averaged to have a better visualization of the main patterns. Very similar features are found in the data sets, with the largest discrepancy found for OMPS-MERRA (MERRA-2 meteorology), which shows pattern with a smaller amplitude w.r.t. the others. SCIA+OMPS displays large anomalies at mid-latitudes due to a discrepancy between the TPH seasonal cycle extracted from ERA5 for SCIAMACHY and OMPS observations, related to different sampling patterns between SCIAMACHY and OMPS LNM datasets.
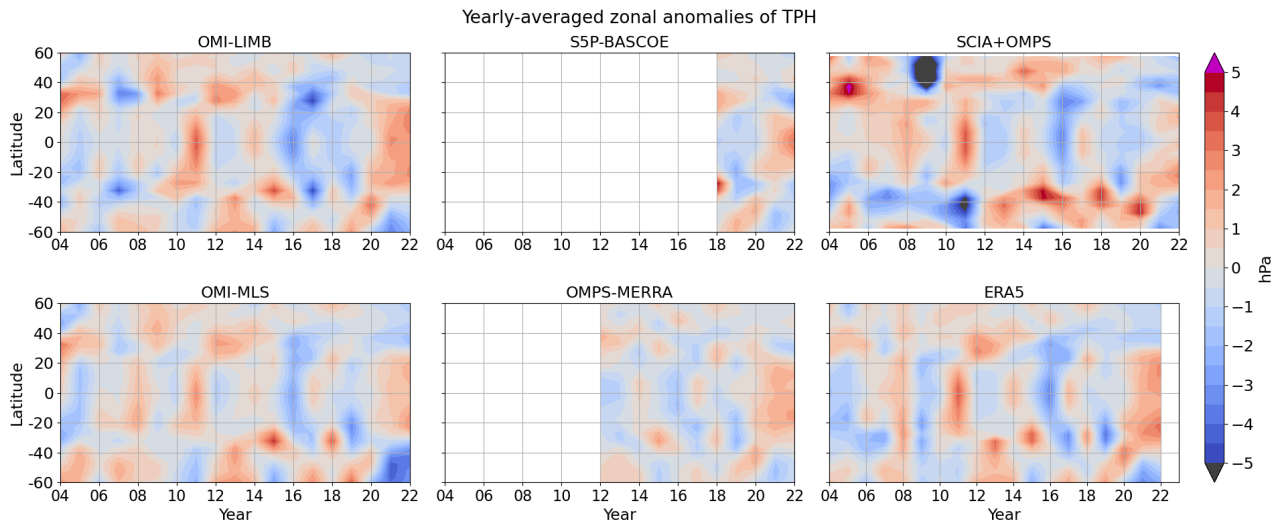

Figure S5: Time series of yearly-averaged zonal absolute anomalies of TPH.

To add some more details about the procedure to correct the datasets and account for their TPH differences, Fig. S6 reports the average TPH bias on the left side, for each data set with respect to ERA5 data. The corresponding column ozone biases are shown on the right panels.
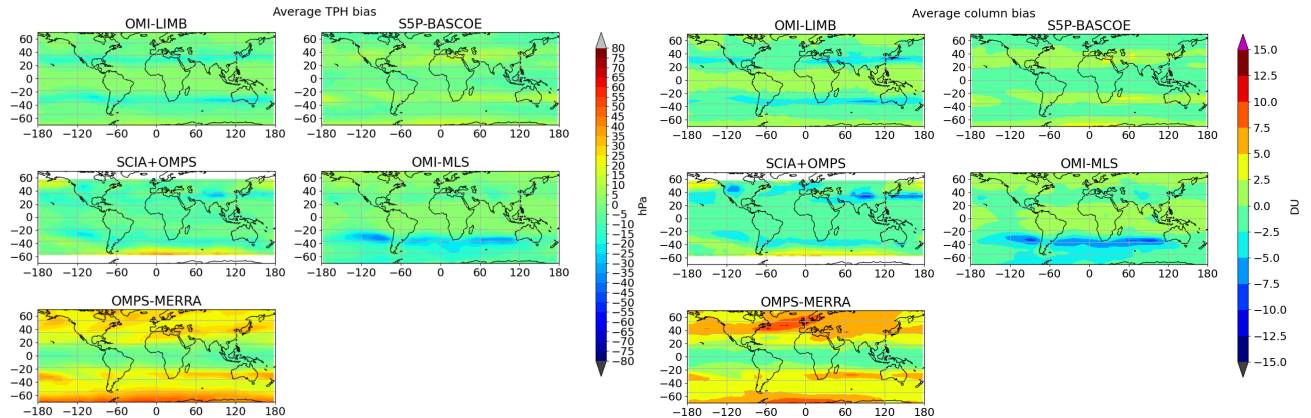

Figure S6: Left panel: TPH bias with respect to the ERA5 thermal TPH; on the right panels the corresponding ozone column between the TPHs of the datasets and ERA5.

In addition, Fig. S7 shows a comparison of the results of the correction of TPH-related differences. In particular, we compare the original time series, and the effect of removing a mean TPH-related bias or a monthly averaged bias (seasonal cycle bias), at 45°N latitude, 11°E longitude. In black the time series from HEGIFTOM data at Hohenpeissenberg is shown. Neither of the two debiasing procedures bring a substantial improvement in the agreement between the data sets. However, OMPS-MERRA and EPIC-MERRA display a better agreement with the HEGIFTOM time series after correcting for the TPH bias, due to their substantially different TPH definition adopted.

Figure S8 showcases the difference between a TPH correction performed at L2 or at L3. The left panel shows a map of the differences between two SCIAMACHY TrOC data sets obtained using different TPH definitions during the LNM process (the one used for the SCIA+OMPS product and the TOAR one, i.e. fixed pressure level depending on latitude). The SCIAMACHY TrOC differences are multi-year averages and are compared to ozonesonde sub-columns (in colored dots), which correspond to the ozone content between the two TPH definitions. The right panel shows the results obtained by adjusting the L3 data to take into account this TPH difference, by using a monthly climatology ("IUP 2018"). In this case the correction in TrOC are significantly different to the ozonesonde sub-columns (same as left panel).
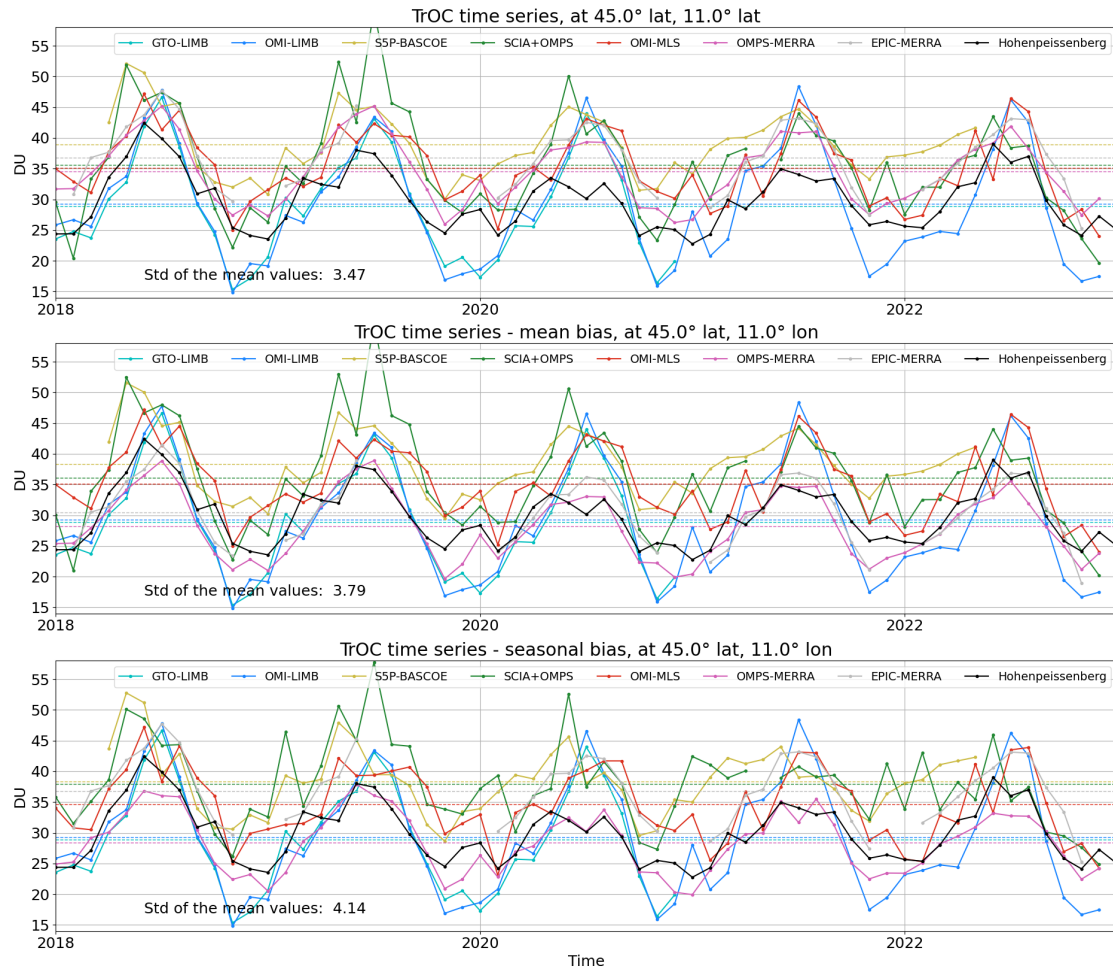
Figure S7: TrOC time series at Hohenpeissenberg (TrOC from HEGIFTOM sondes in black); top panel before correcting for TPH difference, middle panel after removing the mean bias, lower panel after removing monthly-resolved TPH-related bias.
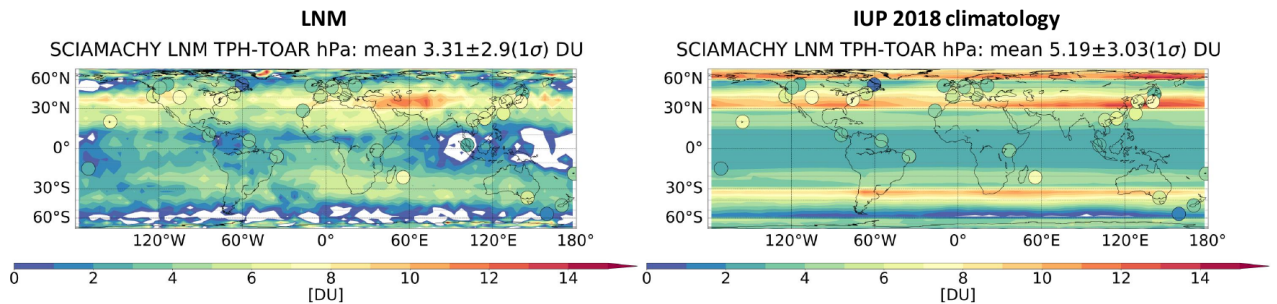

Figure S8: Differences between two SCIAMACHY TrOC data sets obtained using different TPH definitions at the LNM process (the one used for SCIA+OMPS product and the TOAR one). In the left panel the correction is done based on the L2 data, on the right L3 monthly data are corrected for the different TPH using a climatology. The circle show the ozone sub-columns from ozonesonde data between the two TPH definitions.

Table S1 reports the list of ozonosonde stations used for this work, including their locations and data providers. All ozonesonde data used for the study are HEGIFTOM-homogenised.

Table S1: List of the used sonde stations used in this study.

Station	Latitude	Longitude	Data provider
Churchill	58.74	-93.82	WOUDC
Edmonton	53.55	-114.10	WOUDC
Goose Bay	53.29	-60.39	WOUDC
Legionowo	52.40	20.97	NASA-AMES
De Bilt	52.10	5.18	WOUDC
Uccle	50.80	4.35	WOUDC
Kelowna	49.47	-119.38	WOUDC
Hohenpeissenberg	47.80	11.01	WOUDC
Payerne	46.49	6.57	WOUDC
Haute-Provence	43.94	5.71	NASA-AMES
Yarmouth	43.87	-66.11	WOUDC
Trinidad Head	40.80	-124.16	NOAA
Madrid	40.47	-3.58	WOUDC
Boulder	40.13	-105.24	NASA-AMES
Izania	28.31	-16.50	WOUDC
Hanoi	21.20	105.80	SHADOZ
Hilo	19.58	-155.07	SHADOZ
Costarica	9.98	-84.21	SHADOZ
Paramaribo	5.80	-55.21	SHADOZ
Kuala Lumpur	2.73	101.27	SHADOZ
Nairobi	-1.27	36.80	SHADOZ
Natal	-5.42	-35.38	SHADOZ
Ascension Island	-7.97	-14.40	SHADOZ
Samoa	-14.23	-170.56	SHADOZ
La Reunion	-21.06	55.48	SHADOZ
Lauder	-45.00	169.68	WOUDC

Fig. S9 displays the time series of the satellite data sets collocated at four ozonesonde stations with corresponding HEGIFTOM data, with and without applying "AK-like" correction, i.e. the weighting in Eq. 3 of the manuscript.
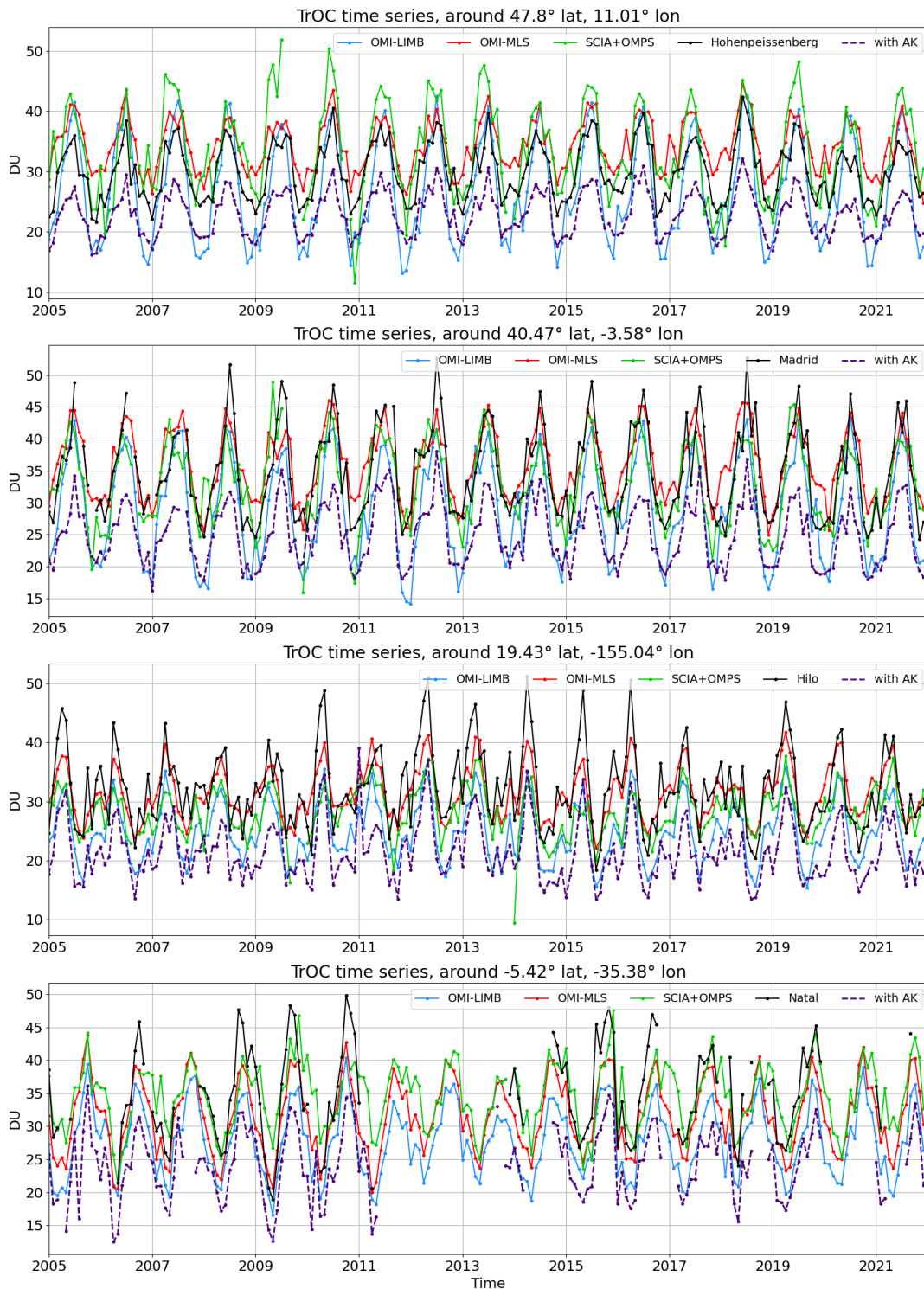

Figure S9: Collocated TrOC at Hohenpeissenberg, Madrid, Hilo and Natal with HEGIFTOM data. In black original HEGIFTOM TrOC, in purple re-calculated applying nadir-like AK.

As we can see, the application of TOC-like AK to HEGIFTOM sonde profiles, leads to the expected overall reduction of the TrOC content, which in some case better agrees only with the OMI-LIMB time series. This however doesn't have a large impact on the seasonality, although it reduces summer-time TrOC peaks at some stations, such as Madrid.

Regarding the drift assessment, Fig. S10 shows the time series of the differences between TOR products and HEGIFTOM sondes, with their respective linear trends, averaged over the two selected latitude bands. In contrast to the figure in the manuscript, the AK-like correction was not applied in this case. Both panels display the monthly time series and their 13-months running averages for a less noisy visualization of long-term tendency. We notice, the more pronounced artefact in the tropics, with negative difference w.r.t. sonde data in the 2012-2013 period, and the HEGIFTOM mean anomalies (in gray, reversed in sign) being strongly correlated with the difference TOR-HEGIFTOM.
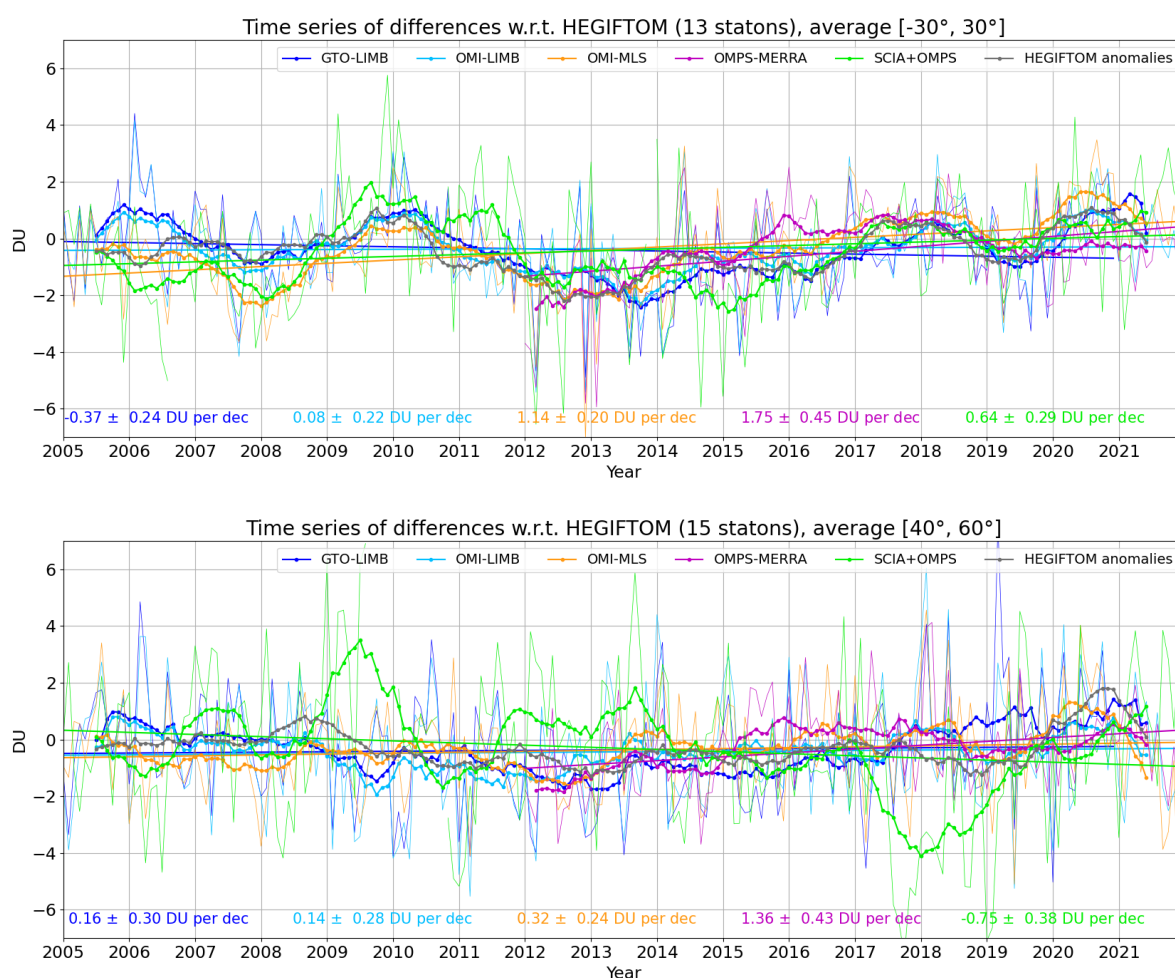

Figure S10: Linear trend in DU per decade of the difference between satellite and HEGIFTOM sonde anomalies (without applying the AK-like weighting). The titles include also the number of sonde stations used for the analysis.

9

Figure S11 shows a point-to-point correlation map of TrOC values. This was used to define geographical regions where the SC of the several data sets is coherent. In this case the reference location is off-coast of west Africa, identified with a white cross.
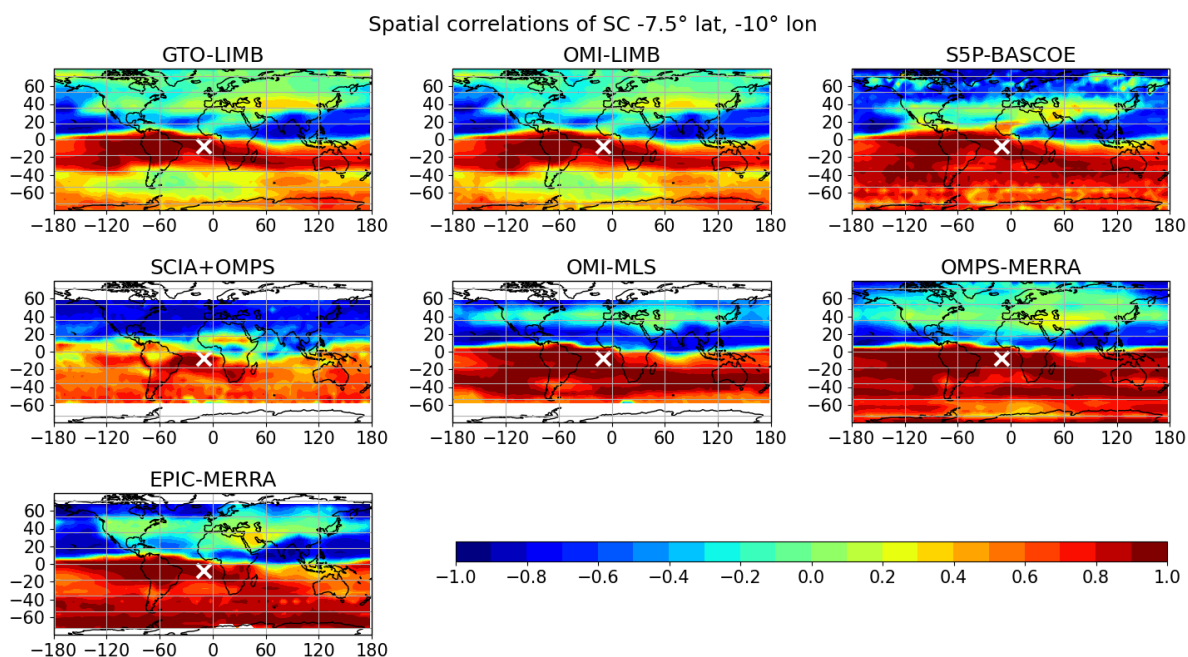

Figure S11: Correlation of the seasonal cycle (SC) of all data sets w.r.t. the location at 7.5°S, 10°W identified with a white cross on the maps.

To better understand the behaviour of the three regression models in different situations, Fig. S12 shows OMI-LIMB time series at two locations, one in the tropics and one at southern mid/high-latitudes. The regression is performed on the absolute anomalies monthly time series (grey line); the 13-month average is over-plotted to visualize a smoothed time series. In the top panel, the contribution of ENSO is relevant, with the MLR fit (dashed green line) reproducing the 13-month average well. This results in the quantile regression having the largest negative trend value, whereas MLR and LIN indicate a closer to zero value. In the second panel, the presence of some outliers towards the end of the time series is properly taken into consideration only by the quantile regression approach. The chosen proxies for MLR do not play a relevant role here.
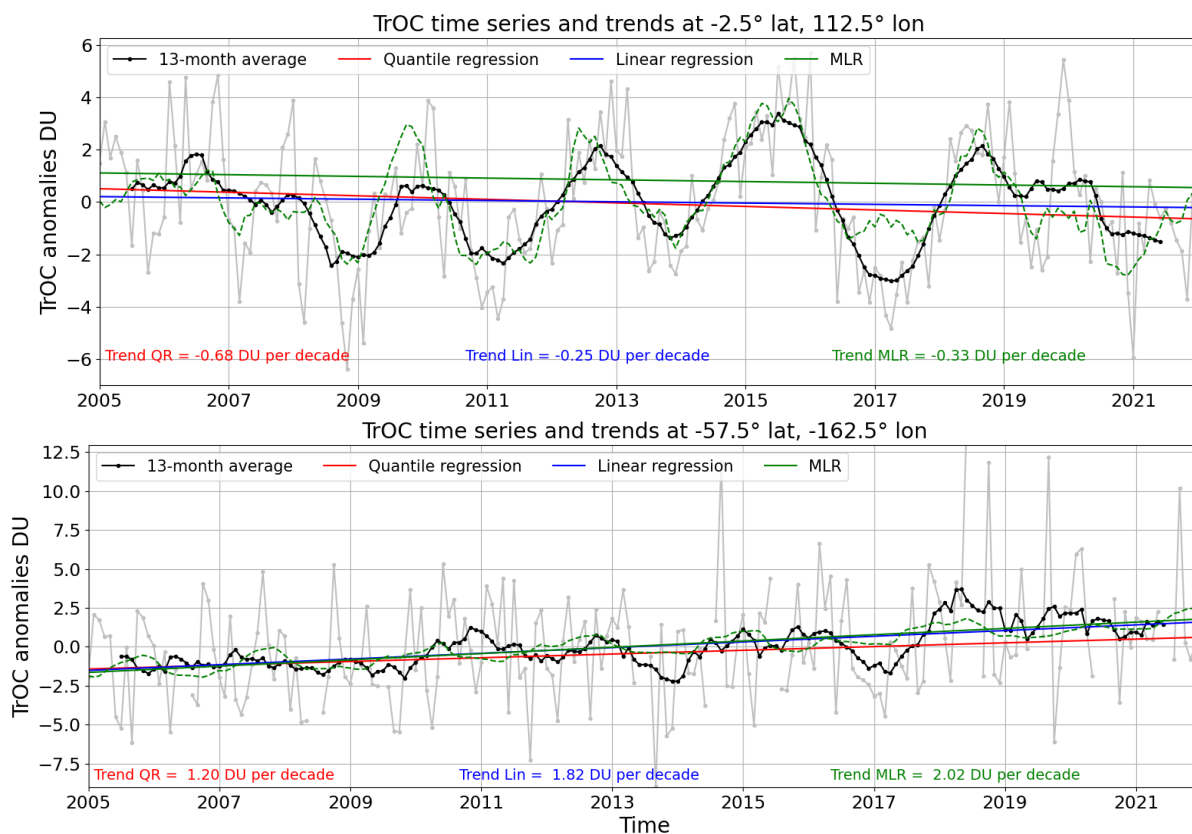

Figure S12: Comparison of the trend results from three regression models applied to OMI-LIMB data at two locations. The MLR fit is also included as dotted green line.

Figure S13 shows the signature of QBO and ENSO proxies in OMI-MLS time series, as an example, at two locations where these proxies show high correlations to ozone. We think that the QBO signature is mostly related to the contamination of the stratospheric signal when computing limb-nadir residuals, although, as shown in the bottom panel of this figure, the HEGIFTOM time series for Nairobi also shows a noticeable QBO signature.
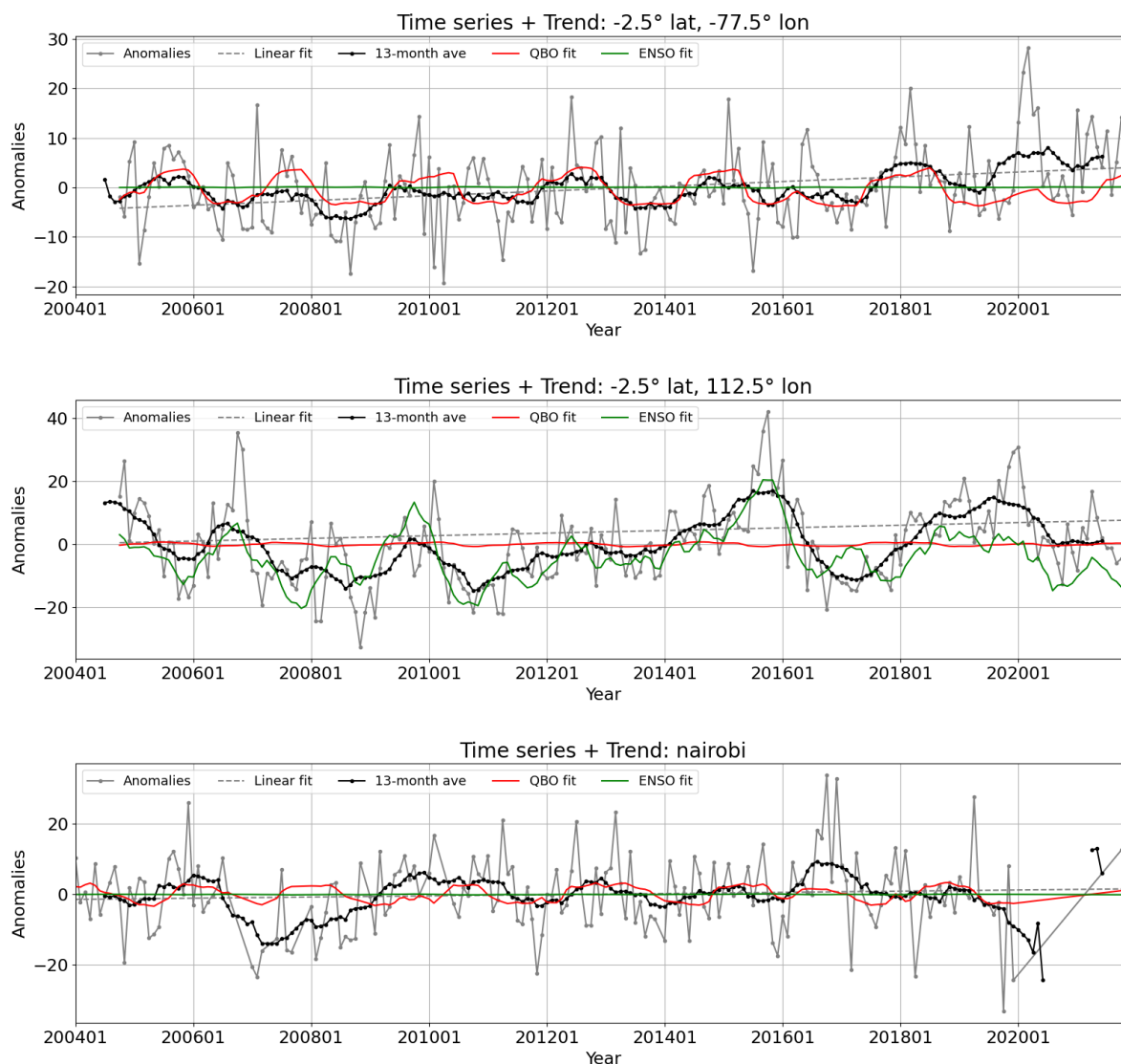

Figure S13: In the first two panels, OMI-MLS time series at the reported latitude-longitude bin is regressed using QBO and ENSO proxies: the respective fit contributions are over-plotted. The bottom panel shows the HEGIFTOM TrOC time series from Nairobi (1.27° S, 36.8° E) with QBO and ENSO regression contributions.

Figure S14 shows the trend results from MLR regression applied to the OMI-MLS data set, using standard proxies (top panel), and including in addition AOD and TPH (bottom panels). Differences are rather minor.
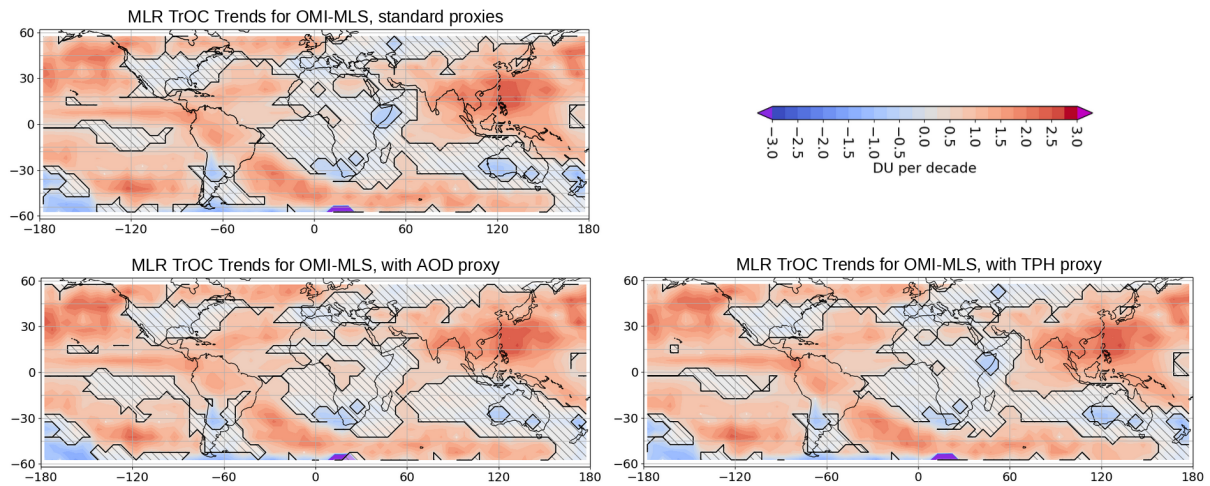
Figure S14: Comparison of the trend results from MLR regression using standard proxies (top panel), and including in addition AOD and TPH (bottom panels).

Figure S15 shows the time series and respective trends for three data sets over Europe and the Mediterranean. In the top panel trends from the monthly time series are computed, showing slightly negative mostly non-significant trends. In the second panel summer-time trends are computed, with negative significant trends evident in by the OMI-LIMB data set. Over the Mediterranean, negative and significant summer-time trends are shown by all data sets.
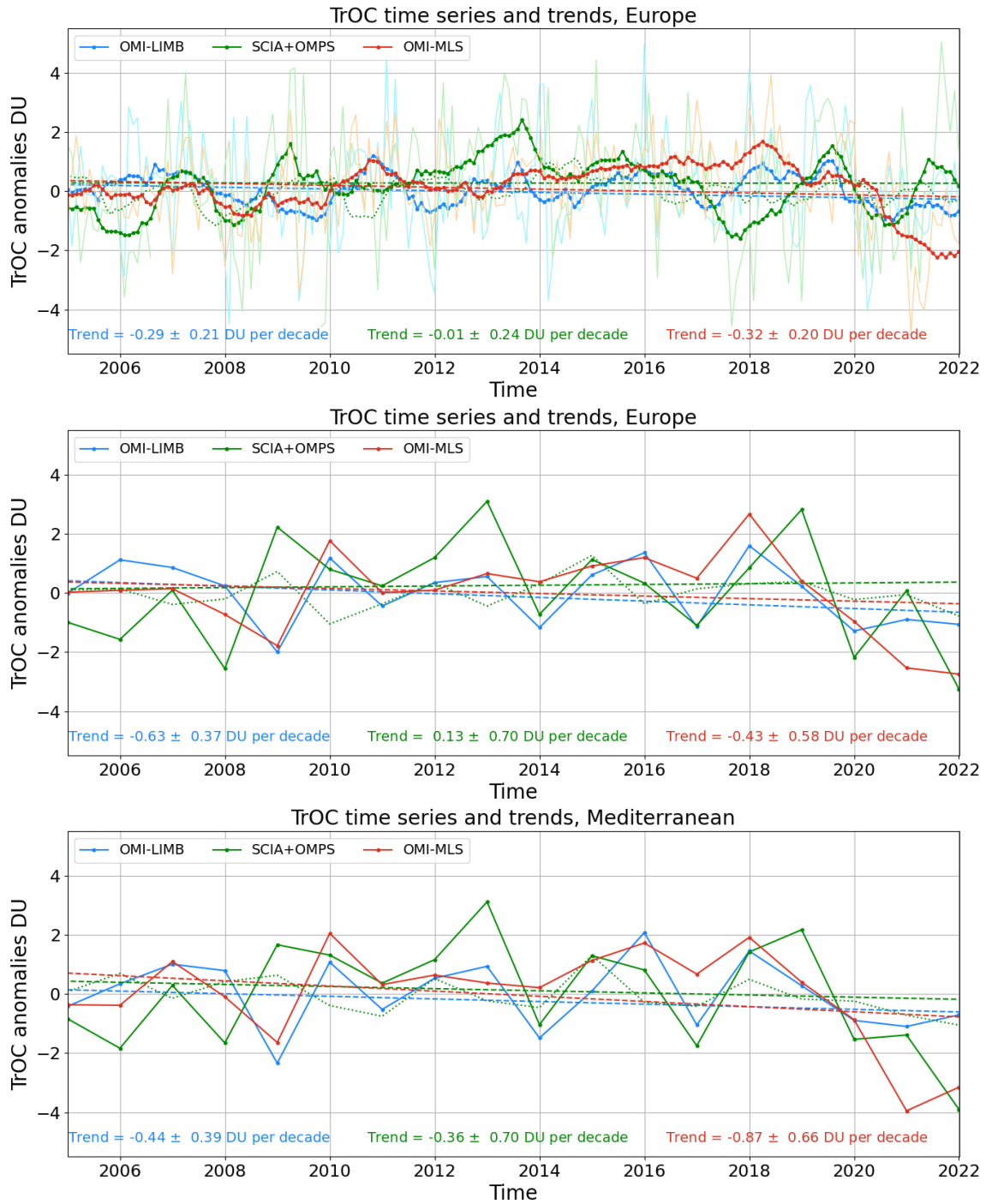
Figure S15: Top panel: time series and trends of the monthly time series averaged over Europe. Middle and bottom panels: summer-time time series and trends over Europe and the Mediterranean, respectively.