Investigation of satellite vertical sensitivity on long-term retrieved lower tropospheric ozone trends

Richard J. Pope[1,2], Fiona M. O'Connor[3,4], Mohit Dalvi[3], Brian J. Kerridge[5,6], Richard Siddans[5,6], Barry G. Latter[5,6], Brice Barret[7], Eric Le Flochmoen[7], Anne Boynard[8,9] , Martyn P. Chipperfield[1,2], Wuhu Feng[1,10], Matilda A. Pimlott[1], Sandip S. Dhomse[1,2], Christian Retscher[11], Catherine Wespes[12] and Richard Rigby[1,13]

1: School of Earth and Environment, University of Leeds, Leeds, United Kingdom

2: National Centre for Earth Observation, University of Leeds, Leeds, United Kingdom

3: Met Office Hadley Centre, Exeter, United Kingdom

4: Department of Mathematics & Statistics, Global Systems Institute, University of Exeter, United Kingdom

5: Remote Sensing Group, STFC Rutherford Appleton Laboratory, Chilton, United Kingdom

6: National Centre for Earth Observation, STFC Rutherford Appleton Laboratory, Chilton, United Kingdom

7: Laboratoire d'Aérologie/OMP, Université de Toulouse, Toulouse, France

8: LATMOS/IPSL, Sorbonne Université, UVSQ, CNRS, Paris, 75005, France

9: SPASCIA, Ramonville-Saint-Agne, 31520, France

10: National Centre for Atmospheric Science, University of Leeds, Leeds, United Kingdom

11: European Space Agency, ESRIN, Frascati, Italy

12: Université libre de Bruxelles (ULB), Spectroscopy, Quantum Chemistry and Atmospheric Remote Sensing, Brussels, Belgium

13: Centre for Environmental Modelling and Computation, University of Leeds, Leeds, United Kingdom

In preparation for *Atmospheric Chemistry and Physics*

Correspondence to: Richard J. Pope (r.j.pope@leeds.ac.uk)

Key Points

- Satellite lower tropospheric column ozone ($LTCO_3$) records in the northern hemisphere show small trends with large uncertainty ranges between 2008 and 2017.~~Satellite lower tropospheric column ozone ($LTCO_3$) trends in the northern hemisphere show small scale trends with large uncertainty ranges between 2008 and 2017~~.
- Modelled $LTCO_3$ over that period is temporally stable and application of the satellite averaging kernels (AKs), accounting for the vertical sensitivity, to the model yields little impact on the simulated trends.

Formatted: Font colour: Auto

Abstract:

Ozone is a potent air pollutant in the lower troposphere and an important short-lived climate forcer (SLCF) in the upper troposphere. Studies investigating long-term trends in tropospheric column ozone (TCO_3) have shown large-scale spatiotemporal inconsistencies. Here, we investigate the long-term trends in lower tropospheric column ozone ($LTCO_3$, surface-450 hPa sub-column) by exploiting a synergy of satellite and ozonesonde datasets and an Earth System Model (UKESM) over North America, Europe and East Asia for the decade 2008-2017. Overall, we typically find small $LTCO_3$ linear trends with large uncertainty ranges from the Ozone Monitoring Instrument (OMI) and the Infrared Atmospheric Sounding Interferometer (IASI), while

model simulations indicate a stable $LTCO_3$ tendency. Trends in the satellite apriori datasets show negligible trends indicating that any year-to-year changes in spatiotemporal sampling over of these satellite data sets over the period concerned has not influenced derived trends inTrends in the satellite a priori datasets show negligible trends indicating year-to-year sampling is not an issue. The application of the satellite averaging kernels (AKs) to the UKESM ozone profiles, accounting for the satellite vertical sensitivity and allowing for like-for-like comparisons, has a limited impact on the modelled $LTCO_3$ tendency in most cases. While, in relative terms, this is more substantial (e.g. in the order of 100%), the absolute magnitudes of the model trends show negligible change. However, as the model has a near-zero tendency, artificial trends were imposed on the model time-series (i.e. $LTCO_3$ values rearranged from smallest to largest) to test the influence of the AKs but simulated $LTCO_3$ trends remained small. Therefore, the $LTCO_3$ tendency between 2008 and 2017 in northern hemispheric regions are likely small, with large uncertainties, and it is difficult to detect any small underlying linear trends due to inter-annual variability or other factors which require further investigation. (e.g. the radiative transfer scheme (RTS) used and/or the inputs (e.g. meteorological fields) used in the RTS).

Formatted: Font colour: Auto

Formatted: Font colour: Auto

1. Introduction

Tropospheric ozone (TO_3) is a short-lived climate forcer (SLCF) and an important greenhouse gas (GHG; Myhre et al., 2013; Forster et al., 2021). TO_3 is also a hazardous air pollutant with adverse impacts on human health (Doherty et al., 2017; WHO, 2022) and agricultural/natural vegetation (Sitch et al., 2007; Hollaway et al., 2012). Since the pre-industrial (PI) period, anthropogenic activities have increased the atmospheric loading of ozone (O_3) precursor gases, most notably methane (CH_4) and nitrogen oxides (NO_x) resulting in an increase in TO_3 of 25-50% since 1900 (Gauss et al., 2006; Lamarque et al., 2010; Young et al., 2013). The PI to present day (PD) radiative forcing (RF) from TO_3 is estimated by the Intergovernmental Panel on Climate Change (IPCC) to be 0.47 Wm^{-2} (Forster et al., 2021) with an uncertainty range of 0.24-0.70 Wm^{-2}.

During the satellite-era (i.e. since the mid-1990s), extensive records of TO_3 have been produced, e.g. by the European Space Agency Climate Change Initiative (ESA-CCI; ESA, 2019). However, the large presents presence of stratospheric O_3, coupled with the different vertical sensitivities and sources of error associated with observations in different wavelength regions (e.g. Eskes and Boersma 2003; Ziemke et al., 2011; Miles et al., 2015) means large-scale inconsistencies in time and space exist between the records of satellite tropospheric column ozone (TCO_3) (as shown by Gaudel et al., 2018).

The work by Gaudel et al. (2018) was part of the Tropospheric Ozone Assessment Report (TOAR), which represented a large global effort to understand spatio-temporal patterns and variability in TO_3. However, tTheir investigation of ozonesondes (2003-2012) and products from nadir viewing satellites in polar orbits (three from the Ozone Monitoring Instrument (OMI) (2005-2015/6) and two from the Infrared Atmospheric Sounding Interferometer (IASI) (2008-2016)) displayed discrepancies in the spatial distribution, magnitude, direction and significance of the TCO_3 trends. They noted that the records cover slightly different time periods but were unable to provide any definitive reasons for these discrepancies beyond briefly suggesting that differences in measurement techniques and retrieval methods were likely to be causing the observed spatial inconsistencies. The range of potential definitions of the tropopause height used to derive TCO_3 from these nadir-viewing profile products could also lead to differences between the satellite product absolute values and their temporal evolution. While the 5 products discussed above use the same definition (i.e. World Meteorological Organisation (WMO) 2 K/km lapse rate; WMO, 1957), several of the other products analysed by Gaudel et al. (2018) did use other definitions.

Formatted: Font colour: Auto

The vertical sensitivity of each retrieved product (function of measurement technique and retrieval methodology) used by Gaudel et al. (2018) will have had an impact on which part of the troposphere the O_3 signal is weighted towards. This is potentially one of the drivers behind the different OMI and IASI TCO_3 trends, where OMI showed predominantly positive trends between 60°S and 60°N while the opposite was the case for IASI.This was evident in the OMI and IASI TCO_3 trends, where OMI showed predominantly positive trends between 60°S and 60°N while the opposite was the case for IASI. The vertical sensitivity is represented by the “averaging kernel” (AK), which provides the relationship between perturbations at different levels in the retrieved and true profiles (Eskes and Boersma, 2003). Typically, for the products used by Gaudel et al., (2018), the peak AK sensitivities for TO_3 are in the 0-6 km range for OMI (Miles et al., 2015) and around 11-12 km for IASI (Keim et al., 2009), while there is a secondary peak at approximately 5 km (Boynard et al., (2009). In the case of the Rutherford Appleton Laboratory (RAL) Space OMI data, used in Gaudel et al., (2018),TCO_3 values were derived from retrieved surface – 450hPa layer average mixing ratios applied also to the overlying 450hpa – tropopause layer using ERA-Interim profiles. As the TO_3 values were derived from different (UV and IR) sensors and methodologies whose vertical sensitivities differ, they were likely representing O_3 controlled by different contributions of atmospheric processes (e.g. precursor emissions from the surface and stratosphere-troposphere exchanges). Therefore, TCO_3 trends from the different satellite products are not necessarily expected to be similar. The determination of the linear trend in a satellite TCO_3 record(s) can also be difficult as many factors (e.g. chemistry, emissions, deposition and transport) control ozone interannual variability, especially on time-periods of a decade or less (Barnes et al., 2016; Change et al., 2020; Fiore et al., 2022).

In this study, we undertake the first assessment of spatio-temporal variability in satellite-derived lower tropospheric column ozone ($LTCO_3$, surface-450 hPa) from three instruments over a consistent decade (2008-2017). In combination with an Earth System Model (ESM), we aim to quantify the impact of year-to-year spatiotemporal samplingyear-to-year sampling, the satellite instrument uncertainties and the instrument vertical sensitivity on long-term $LTCO_3$ trends. We focus our analysis on North America, Europe and East Asia given their large emissions of ozone precursor gases and temporal variability. In our manuscript, **Section 2** discusses the satellite/ozonesonde datasets and model used, **Section 3** presents our results, and our discussion/ conclusions are summarised in **Sections 4 and 5**.

2. Methodology and Datasets

2.1. Satellite Datasets

The satellite products (see **Table 1**) used here are from nadir-viewing polar-orbiting platforms providing ozone sub-column profiles. This includes ozone profile data from the OMI product developed by the RAL Space and the IASI products from the Laboratoire d'aérologie (IASI-SOFRID) and the Université Libre de Bruxelles, in collaboration with the Laboratoire Atmosphères, Observations Spatiales (ULB-LATMOS) (IASI-FORLI). OMI and IASI are on NASA’s Aura and Eumetsat's MetOp-A satellites in sun-synchronous low Earth orbits with local overpass times of 13.30 and 9.30, respectively. OMI and IASI are ultraviolet-visible (UV-Vis) and infrared (IR) sounders with spectral ranges of 270-500 nm (Boersma et al., 2008, Boersma et al., 2011) and 645-2760 cm^{-1} (Illingworth et al., 2011), respectively. OMI has a spatial footprint at nadir of 24 km × 13 km, while IASI measures simultaneously in four fields of view (FOV, each circular at nadir with a diameter of 12 km) in a 50 km x 50 km square which are scanned across track to sample a 2200 km-wide swath (Clerbaux et al., 2009).

The OMI retrieval scheme is based on an optimal estimation (OE) approach, produced by RAL Space, which is described in detail by Miles et al., (2015). The retrieval schemes for IASI-FORLI and IASI-SOFRID O_3 are discussed in detail by Boynard et al., (2018) and Barret et al., (2020). The lowest sub-column in the OMI sub-column profile represents the surface-450 hPa layer (i.e. $LTCO_3$). For the IASI products, there were several

Formatted: Font colour: Auto

Formatted: Font colour: Auto

Formatted: Font colour: Auto

Formatted: Font colour: Auto

sub-columns spanning the surface to 450 hPa range. Therefore, the IASI sub-columns were totalled up between the surface and the layer beneath or equal to the 450 hPa level. Where the 450 hPa level was located within a sub-column (i.e. was located between its bounding upper and lower pressure levels), the sub-column proportion between the lower pressure barrier and the 450 hPa level was determined and added to the sub-columns below (i.e. towards the surface). For the ozone a priori profile, the RAL Space and FORLI schemes use the ozone latitude vs month of year climatology of McPeters et al. (2007), while IASI-SOFRID uses the dynamical ozone climatology described in Sofieva et al. (2014). However, the FORLI scheme uses a single ozone profile (Boynard et al., 2018) derived from the McPeters et al. (2007) dataset, so has no seasonality nor latitude dependence unlike the other retrieval schemes.

In this work, the OMI data were filtered for good quality retrievals where the geometric cloud fraction was <0.2, the sub-column O_3 values were > 0.0, the solar zenith angle < 80.0°, the retrieval convergence flag = 1.0 and the normalised cost function was < 2.0. The IASI-FORLI data were filtered for a geometric cloud fraction <0.13 (pre-filtered), degrees of freedom > 2.0, O_3 values > 0.0, solar zenith angle < 80.0° and the surface to 450 hPa sub-column O_3 / total column O_3 < 0.085. The IASI-SOFRID data were provided on a 1.0°×1.0° horizontal grid (i.e. level 3 product, but at a daily temporal resolution – we use the daytime data in this study) with filtering already applied in Barret et al., (2020). Here, only O_3 values > 0.0 were used. To remove systematic biases between the satellite records, while maintaining the long-term inter-annual variability of each record, ozonesondes were used to generate bias correction offsets (BCOs) (2008-2017) to help harmonise the data sets (i.e. subtraction term in units of Dobson units, DU - as done in Russo et al. (2023) and Pope et al. (2024)) and is discussed in the Supplementary Material (SM) (i.e. S1)To remove systematic biases between the satellite records, ozonesondes were used to generate bias correction factors (2008-2017) to help harmonise the data sets.

Here, each ozonesonde profile was co-located with the nearest satellite retrieval within 500 km and 6 hours to reduce spatiotemporal sampling biases (e.g. Keppens et al., 2019). The ozonesonde sonde profile was then interpolated in the vertical onto the satellite pressure grid where the sub-columns between pressure levels were determined. The ozonesonde sub-column profiles were then convolved by the satellite averaging kernels (AKs), which represent the satellite's sensitivity to retrieval ozone as a function of altitude. Thus, allowing for a robust like-for-like comparison between the ozonesondes and the retrieved $LTCO_3$. The application of AKs to ozonesonde profiles to evaluate satellite ozone products is discussed in detail by Pope et al. (2023). The application of the AKs to the ozonesondes (and the model) is outlined in **Equation 1**:

$$sonde_{AK} = AK(sonde_{int} - apr) + apr \qquad (1)$$

where $\boldsymbol{sonde_{AK}}$ is the modified ozonesonde sub-column profile (Dobson units, DU), $\boldsymbol{AK}$ is the averaging kernel matrix, $\boldsymbol{sonde_{int}}$ is the sonde sub-column profile (DU) on the satellite pressure grid and $\boldsymbol{apr}$ is the apriori (DU). The application of the AKs to the ozonesondes is discussed in more detail in the SM **S1**.This is discussed in the Supplementary Material (i.e. **S1**). The application of the satellite AKs to the ozonesondes and the model is also discussed in **S1**.

To investigate long-term trends over North America, Europe and East Asia, the Hemispheric Transport of Air Pollution (HTAP) regional sea-land mask (European Commission (2016); see **S2**, **Figure S2**), is used to sub-sample the gridded satellite data for the respective regions and then generate average monthly time-series for each product over each region of interest. To investigate long-term trends over North America, Europe and East Asia, the Hemispheric Transport of Air Pollution (HTAP) regional sea-land mask (European Commission (2016); see **S2**, **Figure S2**), is used to generate average monthly time-series for each product over each region of interest. For the ozonesonde time-series for each HTAP region investigated, only ozonesonde sites which are located within each HTAP region are selected. This results in 15, 13 and 6 ozonesonde sites for North America, Europe and East Asia, respectively. As ozonesonde data for East Asia are all from Japan, Taiwan and Hong Kong, trends in ozone $LTCO_3$ will likely be different to satellite/model trends over all East Asia.

Formatted: Font colour: Auto
Formatted: Font colour: Auto
Formatted: Font colour: Auto
Formatted: Font colour: Auto
Formatted: Font colour: Auto
Formatted: Font colour: Auto
Formatted: Font colour: Auto
Formatted: Font colour: Auto
Formatted: Font colour: Auto
Formatted: Font: Bold
Formatted: Font: Bold
Formatted: Font colour: Auto

In Section 3.2, where we discuss the impact of satellite retrieval errors on derived $LTCO_3$ linear trends, the OMI and IASI-FORLI retrieval errors are provided in their product files, but are not available for IASI-SOFRID. Therefore, while not a perfect metric to represent the error in the IASI-SOFRID data, we use the standard deviation in the monthly-spatial average of the regional time-series.

2.2. United Kingdom Earth System Model (UKESM)

The UK's Earth System Model, UKESM1.0, is a state-of-the-art ESM with fully interactive coupled component models (e.g. atmosphere, ocean, land surface, atmospheric chemistry), which has been developed by the UK Met Office and the Natural Environment Research Council (NERC). The detailed coupling of all the Earth System components is described by Sellar et al. (2019). However, in this study, we run UKESM1.0 in an atmosphere only configuration (e.g. similar to Archibald et al., (2020)). The aim is to use UKESM1.0 to investigate long-term trends in TO_3 and help explore inconsistencies between satellite records, so it is computationally more time efficient as only the atmospheric dynamics and chemistry components are simulated. Over the 2008-2017 time period (with a 1-year spin up), the UKESM1.0 model tracers and diagnostics (e.g. ozone, pressure) are output as 3D fields at sub-daily (6-hourly) time steps to allow robust comparisons between the model and satellite data sets (i.e. model-satellite spatio-temporal co-location to reduce representation biases and application of the satellite AKs to map the instrument vertical sensitivity onto the model yielding like-for-like comparisons). The satellite AKs from OMI and IASI-FORLI are provided in the level-2 files (i.e. an AK matrix per retrieval). However, the IASI-SOFRID AKs are provided from the gridded level-3 data product (i.e. an AK matrix for each 1°×1° grid box).

Formatted: Font colour: Auto

Here, the UKESM1.0 land and atmosphere share a regular latitude–longitude grid with a resolution of 1.25° ×1.875° with 85 vertical levels on a terrain-following hybrid height coordinate with a model lid at 85 km above sea level (50 levels are below 18 km). All the key inputs to the model from other Earth system components (e.g. sea surface temperature (SST) and land surface vegetation) were prescribed from ancillary files. The ocean and ice forcing are represented by the monthly Reynolds sea ice and SSTs data from the National Oceanic and Atmospheric Administration (NOAA, https://climatedataguide.ucar.edu/climate-data/). Solar forcings are provided by Phase 6 of the Coupled Model Intercomparison Project (CMIP6; Matthes et al., 2017; Eyring et al., 2016), as is the stratospheric aerosol climatology to represent contributions from volcanic eruptions (Sellar et al., 2019). The land cover is provided from output from the land surface component of the ESM (JULES; Wiltshire et al., 2021) from a fully coupled historical simulation. Anthropogenic and biomass burning emissions from Hoesly et al. (2018) and van Marle et al. (2017) are prescribed for the period 2008 to 2014. After 2014, anthropogenic and biomass burning emissions are from the Shared Socioeconomic Pathway (SSP, Rao et al., 2017) 2-4.5 (i.e. a middle-of-the-road climate and emissions scenario).

Biological emissions are a climatology between 2001 and 2010 from the MEGAN-MACC data base (Sindelarova et al., 2014), while natural emissions are from the Precursors of Ozone and their Effects in the Troposphere (POET, http://accent.aero.jussieu.fr/database_table_inventories.php) based on 1990. Dry deposition of O_3 to the land surface is represented by the Wesley scheme, which is applied as in O'Connor et al., (2014). The model is also in a nudged or "specified dynamics" configuration (i.e. meteorological analyses are used to "nudge" the model's meteorological variables, i.e. u- and v-wind components, and potential temperature, towards reality; Telford et al., 2008) using 6-hourly reanalysis data from the European Centre for Medium-Range Weather Forecasts (ECMWF) ERA-Interim product. A similar configuration of UKESM1.0 was used by Archibald et al., (2020), in which a thorough evaluation against multiple observations (e.g. surface, aircraft and satellite) was carried out.

2.3. Trend Approach

$LTCO_3$ trends are calculated using the linear least squares fit approach of van der A et al., (2006; 2008), and utilised by Pope et al., (2018) who investigated $LTCO_3$ trends. Here, the monthly $LTCO_3$ time-series are represented by the function:

$$Y_t = C + BX_t + A\sin(\omega X_t + \phi) + N_t \qquad (~~1~~2)$$

where Y_t is the observed monthly $LTCO_3$ for month t, X_t is the number of months since the start of the record, *C* is the first monthly mean $LTCO_3$ value of the ~~*C* is the first month of the record~~record, *B* is the monthly linear trend and $A\sin(\omega X_t + \phi)$ is the seasonal model component (Weatherhead et al., 1998). *A* is the amplitude, ω is the frequency (set to 1 year; ω=π/6) and ϕ is the phase shift. *C*, *B*, *A* and ϕ are the fit parameters from the linear least squares fit. N_t represents the model errors/residuals. The linear trend uncertainty, σ_B, represents the trend precision and is calculated as:

$$\sigma_B = \left[\frac{\sigma_N}{n^{\frac{3}{2}}}\sqrt{\frac{(1+\alpha)}{(1-\alpha)}}\right] \qquad (~~2~~3)$$

where n is the number of years, α is the autocorrelation in the residuals (N_t) and σ_N is the standard deviation in the residuals. As in van der A et al., (2006) and Pope et al., (2018), we calculate the autocorrelation for each time-series using a lag of one-time step (i.e. one month). The autocorrelation in **Equation ~~1~~ 2** is not accounted for directly, so is factored into the trend uncertainty (**Equation ~~2~~3**), as used and discussed by van der A et al., (2006) and Weatherhead et al., (1998), respectively.

3. Results

A detailed evaluation of UKESM1.0 $LTCO_3$ through comparisons with the three satellite products and ozonesondes is presented in **S3**. Overall, UKESM1.0 robustly simulates $LTCO_3$ spatially and seasonally in comparison to the ozonesondes and satellite instruments (i.e. typically within the ozonesonde variability and satellite uncertainty range).

3.1. UKESM1.0 and Satellite $LTCO_3$ Trends

3.1.1. North America

$LTCO_3$ trends from OMI, IASI-FORLI, IASI-SOFRID and ozonesondes are derived between 2008 and 2017 (i.e. consistent time record for all instruments) using the linear-seasonal trend model (**Equation ~~1~~2**). For each satellite product, the corresponding UKESM1.0 time-series (with and without AKs) are analysed as well as the satellite apriori. For the North America OMI metrics (**Figure 1 – top left, Table 2**), there is clear seasonality in the apriori ranging between approximately 17.0 and 22.0 Dobson Units (DU). As this is based on the climatology of McPeters et al., (2007), there is no trend and there is a very good model fit (i.e. R^2=1.0). The key point is that, as a climatology, the apriori will have no trend but if there are substantial temporal sampling differences between years, then an artificial trend could be introduced. OMI $LTCO_3$ ranges between 20.0 and 27.0 DU with substantial variability. There is a drop in $LTCO_3$ to 19.0 DU in 2009 before peaking at 25.0-27.0 DU between 2010 and 2015. Peak $LTCO_3$ then drops to 22.0-24.0 DU in 2016 and 2017. As a result, the linear-seasonal trend model, which does not account for interannual variations such as this, only has a fit skill of R^2=0.59. The corresponding OMI $LTCO_3$ trend is -0.79 (-7.07, 5.48; 95% confidence interval~~, p-value = 0.80~~) DU/decade showing a negligible trend with a large uncertainty range. Here, -0.79 DU/decade is the trend while the -7.07 and 5.48 DU/decade values are the 95% confidence interval. The UKESM1.0 $LTCO_3$ time-series ranges between 17.0 and 22.0 DU with clear seasonality, though somewhat less inter-annual variation than OMI, and the linear-seasonal trend model therefore has a considerably better fit with R^2=0.95. The model trend has the opposite sign at 0.21 (-0.37, 0.78~~; p = 0.59~~) DU/decade. Here, the

Formatted: Font: Italic

Formatted: Font: Bold

Formatted: List Paragraph, Outline numbered + Level: 3 + Numbering Style: 1, 2, 3, … + Start at: 1 + Alignment: Left + Aligned at: 0.63 cm + Indent at: 1.9 cm model trend is near-zero with a relatively large uncertainty range (though not as sizable as OMI). When the AKs are applied to the model, the trend switches sign to -0.57 (-1.58, 0.45~~, p = 0.98~~) DU/decade and the linear-seasonal trend model fit decreases in skill to R^2=0.90. The trend switch of sign, though small, is potentially linked to the application of the AKs, which also increases $LTCO_3$ by 2.0-3.0 DU in general.

We also investigated the satellite degrees of freedom of signal (DOFS) over the lower troposphere (i.e. surface to 450 hPa), which provides an estimate of the number of independent pieces of information in the $LTCO_3$. The DOFS are calculated by taking the trace of the AK matrix over the lower tropospheric levels in the satellite vertical grid. Overall, we found that the products for the three regions had negligible trends in their time-series (i.e. within ±1.0 %/year) meaning that the information content of satellite $LTCO_3$ had remained stable with time (see **S3**).

The IASI-FORLI $LTCO_3$ time-series (**Figure 1 – top right**) tends to be lower than OMI and range between 17.0 and 22.0 DU. There is a substantial negative IASI-FORLI trend (-1.42 (-2.35, -0.50~~; p = 0.00~~) DU/decade); **Table 2**) though as stated by Boynard et al., (2018) and Wespes et al., (2018)~~though as suggested by Boynard et al., (2018) and Wespes et al., (2018)~~, the input IASI Level-1 data sets into the FORLI retrieval are not consistent with time; they suffer from a specific discontinuity in September 2010 which degrades the robustness of this trend. While we are aware of the artificial trend in the IASI-FORLI dataset, it is still a valuable long-term product allowing us to quantify multiple factors (e.g. impact of AKs on model tendencies/absolute values and year-to-year spatiotemporal sampling~~year-to-year sampling~~ stability – i.e. near-zero trend in the apriori). The apriori has a negligible trend but there is no clear seasonality in the apriori time-series. As a result, the linear-seasonal trend model has a more limited fit skill (i.e. R^2=0.67). The impact of the satellite AKs appears to have less impact for IASI-FORLI as both UKESM1.0 and UKESM1.0+AKs have time-series ranging between approximately 17.0 and 21.0 (though slightly smaller UKESM1.0+AKs range) and linear-seasonal trend model fits of R^2=0.93 and R^2=0.92, respectively. The corresponding trends are small at -0.13 (-0.75, 0.49~~; p = 0.67~~) and -0.32 (-0.82, 0.20~~; p = 0.22~~) DU/decade, but the introduction of the AKs does move the UKESM1.0 trend slightly towards that of the satellite. Interestingly, while the application of the IASI-FORLI AKs to UKESM marginally pushes the convolved model trend in $LTCO_3$ towards that of the satellite (which has a substantial negative trend), the IASI-FORLI DOFS actually have small positive trends (0.37-0.57 %/year – see **S3**). Therefore, there is minor scale, yet contrasting, discrepancy in how the vertical sensitivity is influencing the long-term $LTCO_3$ trends.

For IASI-SOFRID (**Figure 1 – bottom left**), there is little difference between any of the time-series as they all range between 16.0 and 21.0 DU with corresponding linear-seasonal trend model fits of R^2=0.94 to 0.98 and negligible trends. The IASI-SOFRID and apriori trends are 0.12 (-0.59, 0.82; p = 0.74) and 0.11 (-0.17, 0.39~~; p = 0.43~~) DU/ decade; **Table 2)**, respectively, with the model showing near-zero trends in both cases. Given the close agreement between the satellite and apriori time series and fit metrics, it is suggestive that IASI-SOFRID TO_3 is more closely confined to the apriori profile than are the other products. The ozonesondes show a substantial trend of -1.15 (-2.0, -0.10~~; p = 0.03~~) DU/decade, while the model trend sampled as the sondes is -0.16 (-1.67, 1.35; p =0.63) DU/decade. The co-located model and ozonesonde linear-seasonal trend model fits are R^2=0.62 and 0.64, respectively. The noise and lack of seasonality in the ozonesonde time-series is slightly unexpected given the reasonable density of stations over North America, though the spatial coverage and temporal sampling is much less than the satellite products.

3.1.2. Europe

Formatted: Font colour: Auto

Formatted: Font: Not Bold

Formatted: Font colour: Auto

Formatted: Font colour: Auto

Formatted: Font colour: Auto

Formatted: Outline numbered + Level: 3 + Numbering Style: 1, 2, 3, … + Start at: 1 + Alignment: Left + Aligned at: 0.63 cm + Indent at: 1.9 cm

In Europe, the OMI $LTCO_3$ values are larger than in North America, ranging between 19.0 and 30.0 DU (**Figure 2 – top left**). The same inter-annual variability exists, peaking between 2010 and 2015 with the minimum in 2009. Hence, the linear-seasonal trend model, which does not represent interannual variation, does not have high skill and R^2=0.72. The corresponding trend is -0.80 (-7.29, 5.69~~; p = 0.80~~) DU/decade, so has a similar direction and magnitude to that for North America, though is not substantial. The apriori ranges between 17.0 and 22.5 DU with a trend of -0.12 (-0.26, 0.03~~; p = 0.10~~; **Table 2**) DU/decade. Given the relatively small trend and uncertainty range, unlike the OMI equivalent, it suggests there is unlikely to be an artificial trend arising through year-to-year spatiotemporal sampling ~~year-to-year~~ changes in geographical sampling across the European region. UKESM1.0 $LTCO_3$ ranges between approximately 19.0 and 22.0 DU with a good linear-seasonal trend model fit of R^2=0.99 and a trend of -0.11 (-0.50, 0.29~~; p = 0.59~~) DU/decade. As for North America, when the OMI AKs are applied, the UKESM $LTCO_3$ values systematically increase by 2.0-3.0 DU, move further away from the satellite apriori and more closely follow the variability of OMI (although R^2 decreases slightly to 0.95). The trend tends towards that of OMI at -0.72 (-1.77, 0.32~~; p = 0.16~~) DU/decade. As in the case of North America, the European IASI-FORLI apriori has no seasonal cycle (and moderate R^2 of 0.48 in the linear-seasonal trend model fit) with a near-zero trend (0.09 (-0.09, 0.27~~; p = 0.32~~) DU/decade) (**Figure 2 – top right**, **Table 2**).

The IASI-FORLI data exhibit a substantial negative trend of -1.83 (-2.78, 0.89~~; p = 0.00~~) DU/decade, again due to ~~again potentially attributable to~~ step changes in the IASI Level-1 processor, with a good linear-seasonal trend model fit of R^2=0.92. UKESM1.0 $LTCO_3$ trends, without and with AKs applied, are -0.28 (-0.77, 0.20~~; p = 0.25~~) and -0.43 (-1.21, 0.35~~; p = 0.27~~) DU/decade. Again, though a small change, the application of the AKs introduces a slight perturbation of the model trend compared to IASI-FORLI. The IASI-SOFRID apriori, ranging between 17.0 and 21.0 DU, has a trend of 0.17 (-0.12, 0.45~~; p =0.24~~) DU/decade with good fit skill of R^2=0.98 (**Figure 2 – bottom left**).

The IASI-SOFRID and UKESM1.0 metrics, with and without averaging kernels applied, are similar, with $LTCO_3$ trends of 0.05 (-0.91, 1.01;~~ p = 0.92~~), -0.27 (-0.72, 0.19~~; p = 0.24~~) and 0.08 (-0.33, 0.49~~; p = 0.69~~) DU/decade, respectively, and with R^2 values between 0.93 and 0.98. The ozonesonde monthly regional means (**Figure 2 – bottom right**) has a more pronounced time-series than North America, yielding a less noisy time-series of $LTCO_3$. Here, there is clear seasonality ranging between 17.0 and 24.0 DU with a large R^2 value of 0.95. The ozonesonde trend is relatively small at -0.61 (-1.39, 0.17~~; p = 0.12~~) DU/decade while the UKESM1.0 equivalent is more substantial at -0.96 (-1.56, 0.35~~; p = 0.00~~) DU/decade.

3.1.3. East Asia

For East Asia, OMI $LTCO_3$ again has both a pronounced seasonal cycle and inter-annual variability (19.0-27.0 DU), consistent with the other two regions discussed above (**Figure 3 – top left**). This yields a moderate skill fit to the linear-seasonal trend model of R^2=0.52 and near-zero trend (-0.09 (-7.88, 7.70~~; p = 0.98~~) DU/decade). The apriori has a trend of -0.25 (-0.71, 0.22~~; p = 0.29~~) DU/decade, so year-to-year spatiotemporal sampling ~~year-to-year sampling~~ changes could be influencing the robustness of OMI retrieved time-series in this region. However, both the instrument and apriori trend uncertainties intersect with 0.0. UKESM1.0 $LTCO_3$ ranges between approximately 16.0 and 22.0 DU with a good fit R^2 of 0.98. Like the other regions, the application of the OMI AKs increases the model values systematically by several DUs. The UKESM1.0 $LTCO_3$ trend is -0.16 (-0.94, 0.62~~; p = 0.67~~) DU/decade, which is small, but the AKs increase the trend magnitude to -0.62 (-2.24, 1.00~~; p = 0.44~~) DU/decade, which moves it away from the OMI trend.

IASI-FORLI (**Figure 3 – top right**, **Table 2**), like the other two regions, has a substantial negative trend of -1.52 (-2.16, 0.88~~; p = 0.00~~) DU/decade. The apriori again exhibits virtually no seasonal cycle (low fit skill of

Formatted: Font colour: Auto

Formatted: Font colour: Auto

Formatted: Outline numbered + Level: 3 + Numbering Style: 1, 2, 3, … + Start at: 1 + Alignment: Left + Aligned at: 0.63 cm + Indent at: 1.9 cm

Formatted: Font colour: Auto

R^2=0.21) ~~but non-zero~~ and negligible year-to-year spatiotemporal sampling differences yielding a ~~year-to-year variation so a~~ ~~low fit skill of R^2=0.21 and shows a~~ near-zero trend of -0.03 (-0.22, 0.16; p = 0.76) DU/decade. For UKESM1.0, the East Asian seasonal range is much larger than other regions, ranging between 17.0 and 27.0 DU (i.e. seasonal amplitude of approximately ±5.0 DU). When the AKs are applied, this range shrinks to approximately 19.0 to 23.0 DU, more in-line with the IASI-FORLI $LTCO_3$ values. The corresponding model trends are -0.03 (-0.62, 0.56~~; p = 0.93~~) DU/decade and -0.29 (-0.80, 0.22~~; p = 0.25~~) DU/decade, so the AKs are pushing the model tendency towards that of the instrument, though the impact is small in absolute terms (large in relative terms).

IASI-SOFRID and its apriori $LTCO_3$ seasonality are again very similar, ranging between 16.0 and 21.0 DU with very little interannual variability and with linear seasonal trend model fit skills of R^2=0.96 and 0.98 (**Figure 3 – bottom left**, **Table 2**). The IASI-SOFRID and apriori linear trends are therefore also consistent at -0.19 (-1.01, 0.63~~; p = 0.65~~) and -0.15 (-0.73, 0.58~~; p = 0.82~~) DU/decade. The UKESM1.0 seasonal variability is again large, between 17.0 and 26.0 DU, and, as in the case of IASI-FORLI, when the instrument AKs are applied to the model, the seasonal range shrinks (i.e. 16.0-22.0 DU) to be much closer to those of the retrieval and its prior. The model trends are -0.42 (-0.97, 0.13~~; p =0.12~~) and -0.24 (-0.67, 0.20~~; p = 0.28~~) (with AKs) DU/decade, where there is a minor shift in the model tendency towards that of IASI-SOFRID and its prior. For the ozonesondes (**Figure 3 – bottom right**), there is a substantial $LTCO_3$ trend of 3.17 (0.16, 6.17; p = 0.04) DU/decade with a fit skill of R^2=0.79, which is larger than those for North America and Europe. $LTCO_3$ increases from 18.0-25.0 in 2008 to 21.0-28.0 in 2011. This remains similar in 2012 and 2013 before dropping by several DUs between 2014 and 2017. The UKESM1.0 sampled as the ozonesondes has considerably less inter-annual variability with a smaller trend of 0.37 (-0.90, 1.64~~; p = 0.56~~) DU/decade. Therefore, UKESM1.0 and the satellite product trends are generally smaller (in magnitude) than the ozonesonde tendencies. However, it is worth considering that there are only a few sites (e.g. Hong Kong and Taiwan) where ozonesonde data is available in East Asia.

~~1.1.~~3.2. Influence of Satellite Averaging Kernels on UKESM1.0 $LTCO_3$

Formatted: Indent: Left: 0.63 cm, Hanging: 0.63 cm, Outline numbered + Level: 2 + Numbering Style: 1, 2, 3, … + Start at: 1 + Alignment: Left + Aligned at: 0 cm + Indent at: 0.63 cm

To investigate the impact of applying the satellite averaging kernels to UKESM1.0, and thus learn something about vertical sensitivity influence on retrieved $LTCO_3$, three different metrics are considered for the 2008 to 2017 time-period. These are the absolute $LTCO_3$ value, amplitude of the $LTCO_3$ seasonal cycle and the linear trend. These metrics are compared for the satellite, the satellite ± error term, the apriori, UKESM1.0 and UKESM1.0+AKs for the three regions discussed above.

From **Figure 4**, average OMI $LTCO_3$ is approximately 22.0, 24.0 and 23.0 DU for North America, Europe and East Asia, respectively. This represents a substantial deviation away from the apriori values of 17.5, 20.0 and 16.0 DU, respectively. However, the average error term for OMI $LTCO_3$ is sizeable at approximately ±8.0 to ±9.0 DU for all regions. The average UKESM1.0 value for each region is approximately 19.5, 21.5 and 19.0 DU but the application of the AKs increases this by several DU to 22.0, 24.0 and 21.0 DU. In comparison, mean values for both IASI products vary less between the three geographical areas: IASI-FORI (IASI-SOFRID) $LTCO_3$ values are 20.0 (18.5), 19.0 (18.5) and 22.0 (18.0) DU, respectively. The corresponding error ranges, in comparison with OMI, are smaller between 17.0 and 23.0 (16.0 and 21.5), 16.0 and 21.5 (16.0 and 21.0) and 18.0 and 23.5 (14.5 and 21.5) DU for North America, Europe and East Asia, respectively. With the IASI-FORLI AKs applied to UKESM1.0, $LTCO_3$ decreases from 19.5 to 19.25 DU, 21.25 to 19.5 DU and 22.75 to 21.25 DU for the three regions. For IASI-SOFRID, there is a decrease from 21.0 to 19.5 DU in Europe and a decrease from 22.0 to 19.5 DU in East Asia, while no change occurs in North America. Overall, OMI has the largest error range and the application of the AKs to UKESM1.0 systematically increases the model $LTCO_3$ time-series by several DU. The opposite occurs for the IASI products where there is a smaller decrease to UKESM1.0 $LTCO_3$ of 1.0-2.0 DU. The error ranges are also smaller than that of OMI.

In terms of the $LTCO_3$ seasonal amplitude (**Figure 5**), OMI (including the error terms) is approximately 2.6 (for all) DU, 3.3-3.8 DU and 2.3-2.6 DU for North America, Europe and East Asia. The apriori seasonal amplitude ranges from 2.7 to 2.9 DU across the regions. The IASI-FORLI averages (including the error terms) tend to be lower than OMI but have similar seasonal ranges. North America, Europe and East Asia have amplitudes of 2.3-2.5 DU, 2.3-2.5 DU and 1.6-1.8 DU, respectively. It is noteworthy that this seasonal cycle is despite the IASI-FORLI prior exhibiting virtually no seasonal cycle at all. IASI-SOFRID has a European range of 2.4-2.6 DU, and comparable ranges for North America and East Asia at 1.8-2.5 DU and 2.3-3.0 DU. Therefore, seasonal amplitude in IASI-SOFRID is more sensitive to the error metric but as the “error” term is based on the $LTCO_3$ standard deviation, given the lack of an error term in the product, it is unsurprising that there is more variability in the seasonal amplitude. For the OMI comparisons, the application of the AKs to UKESM1.0 shifts the simulated amplitude slightly upwards from 2.0 to 2.1 DU, 3.1 to 3.3 DU and 4.0 to 4.4 DU for the respective regions. The IASI-FORLI AK impacts are a decrease from 1.9 to 1.4 DU, 3.0 to 2.1 DU and 4.2 to 1.9. For IASI-SOFRID, the corresponding impact on UKESM1.0 is 2.2 to 2.4 DU, 3.3 to 2.9 and 4.5 to 3.2 DU. Therefore, the OMI AKs have a minimal impact, increasing the model seasonal amplitude by 0.1-0.3 DU, but the IASI products suppress the simulated amplitude by 1.0-2.0 DU at the most extreme.

The impact of the satellite $LCTO_3$ error terms on the derived linear trends are shown in **Figure 6**. For OMI, the range in trends calculated (i.e. satellite ± error term) is approximately -1.50 (-7.04, 4.04~~; p = 0.59~~) to -0.09 (-6.98, 6.81~~; p = 0.98~~) DU/decade, -1.65 (-6.92, 3.62~~; p = 0.53~~) to 0.05 (-7.44, 7.53~~; p = 0.99~~) DU/decade and -1.05 (-6.61, 4.52~~; p = 0.70~~) to 0.87 (-8.24, 9.98~~; p = 0.85~~) DU/decade for North America, Europe and East Asian, respectively. The IASI-FORLI trends (i.e. satellite ± error term) are substantial , ranging from -1.50 (-2.51, -0.50~~; p =0.00~~) to -1.34 (-2.21, -0.47~~; p = 0.00~~) DU/ decade, -1.87 (-2.87, -0.87~~; p =0.00~~) to -1.80 (-2.72, -0.88~~; p = 0.00~~) DU/decade and -1.62 (-2.27, -0.98~~; p =0.00~~) to -1.42 (-2.06, -0.78~~; p = 0.00~~) for the three regions, respectively. The corresponding IASI-SOFRID trends were 0.09 (-0.48, 0.66~~; p =0.75~~) to 0.14 (-0.59, 0.88~~; p = 0.70~~) DU/decade, -0.07 (-0.91, 0.78~~; p = 0.87~~) to 0.16 (-0.74, 1.07~~; p = 0.72~~) DU/decade and -0.30 (-1.02, 0.42~~; p = 0.41~~) to -0.08 (-0.73, 0.58~~; p =0.82~~), respectively. Therefore, only the IASI-FORLI trends (i.e. satellite ± error term) are substantially different from zero (i.e. $p < 0.05$). However, that is likely due in part to discontinuities in the Level-2 input meteorological data used to generate this version of the product (Boynard et al., 2018).

The application of the OMI AKs to UKESM1.0 had the largest impacts on the simulated trends with changes in a negative direction from of 0.21 (-0.37, 0.78~~; p = 0.59~~) to -0.57 (-1.58, 0.45~~, p = 0.98~~) DU/decade, -0.11 (-0.50, 0.29~~; p = 0.59~~) to -0.72 (-1.77, 0.32~~; p = 0.16~~) DU/decade and -0.16 (-0.94, 0.62~~; p = 0.67~~) to -0.62 (-2.24, 1.00~~; p = 0.44~~) DU/decade for the respective regions. IASI-FORLI AKs introduced small decreases from -0.13 (-0.75, 0.49~~; p = 0.67~~) to -0.32 (-0.82, 0.20~~; p = 0.22~~) DU/decade, -0.28 (-0.77, 0.20~~; p = 0.25~~) to -0.43 (-1.21, 0.35~~; p = 0.27~~) DU/decade and -0.03 (-0.62, 0.56~~; p = 0.93~~) to -0.29 (-0.80, 0.22~~; p = 0.25~~) DU/decade. IASI-SOFRID AKs introduced small increases in the $LTCO_3$ trend from -0.24 (-0.85, 0.37~~; p = 0.44~~) to -0.04 (-0.53, 0.45~~; p = 0.87~~) DU/decade, -0.27 (-0.72, 0.19~~; p = 0.24~~) to 0.08 (-0.33, 0.49~~; p = 0.69~~) DU/decade and -0.42 (-0.97, 0.13~~; p = 0.12~~) to -0.24 (-0.67, 0.20~~; p = 0.28~~) DU/decade.

As the absolute model trends are small, it is difficult to determine the impact of the AKs on the simulated trends. In relative terms, it can have impacts of several 100% but the model and model+AK trend ranges (95% confidence interval) always intersect. Therefore, in an attempt to derive more substantial UKESM1.0 $LTCO_3$ trends (without and with AKs applied), to assess the maximum impact the AKs can have on UKESM

$LTCO_3$ trends, the modelled data were sorted from lowest to highest and the trend re-calculated. In North America, this approach forced positive model trends, sub-sampled to OMI, IASI-FORLI and IASI-SOFRID, of 0.73 (0.22, 1.25~~; p = 0.00~~), 0.64 (-3.50, 4.77~~; p = 0.76~~) and 0.80 (0.41, 1.19~~; p = 0.00~~) DU/decade. When the AKs were applied, it yielded trends of -0.74 (-1.89, 0.40~~; p = 0.20~~), 0.55 (0.08, 1.03~~; p =0.02~~) and 0.58 (0.24, 0.92~~; p =0.00~~) DU/decade. In Europe, this forced positive trends model trends, of 0.62 (0.14, 1.10~~; p = 0.01~~), 0.37 (-0.05, 0.79~~; p = 0.08~~) and 0.46 (0.09, 0.84~~; p = 0.01~~) DU/decade, respectively. With the AKs applied, the trends become 0.47 (-0.51, 1.44~~; p = 0.34~~), 0.28 (-0.38, 0.94~~; p = 0.40~~) and 0.10 (-0.32, 0.51~~; p = 0.64~~) DU/decade. Finally, in East Asia, the forced model trends are 0.90 (0.34, 1.47~~; p = 0.00~~), 0.66 (0.15, 1.17~~; p = 0.01~~) and 0.63 (0.26, 1.00~~; p = 0.00~~) DU/decade. The application of the AKs introduced model trends of 1.02 (-0.04, 2.09~~; p = 0.05~~), 0.08 (-0.44, 0.61~~; p = 0.75~~) and 0.20 (-0.20, 0.61~~; p = 0.31~~) DU/decade.

Even with forced trends in the UKESM1.0 regional time-series, the trends are relatively small (i.e. typically less than 1.0 DU/decade in magnitude). Therefore, the application of the AKs to the forced UKESM $LTCO_3$ time-series still yields small scale change tendencies and there is overlap in the two model trend uncertainty ranges (i.e. 95% confidence level). However, in relative terms, the trend changes are larger (e.g. >100% in multiple cases) and there is often a shift of the modelled $LTCO_3$ trend uncertainty range either intersecting or no longer intersecting with zero (i.e. a shift in p-value regime from <0.05 to >0.05). Therefore, in modelled and satellite datasets with more substantial trends, the impacts of the AKs, and thus the satellite vertical sensitivity, on $LTCO_3$ trends would be much greater and potentially help pinpoint sources of differences between satellite products in their TO_3 temporal evolution.

Section 3.3. Diurnal Variability on Regional $LTCO_3$ and Temporal Evolution

As TO_3 varies diurnally due to meteorological and photochemical processes (e.g. Gaudel et., 2018), the different satellite overpass times (i.e. Aura and MetOp-A daytime overpasses are around 13:30 and 09:30 local time, respectively) will likely influence the spatial distributions of TO_3 which OMI and IASI will retrieve. In principle, this could therefore explain some differences between the two sensors and their long-term $LTCO_3$ trends. The model is a useful tool to help investigate this and we used the 6-hourly output to derived the UKESM simulated $LTCO_3$ spatial distributions at the Aura (13.30 LT) and MetOp-A (09.30 LT) day-time overpasses. These model fields were then used to calculate regional time-series for North America, Europe and East Asia. For each region and month, between 2008 and 2017, we calculated the regional average absolute difference (i.e. from the selection of model grid cells which fell within the HTAP-2 mask for a specific month) and the standard deviation of the absolute differences between the overpass times. Here, across all months and regions, we found the peak average absolute difference (13:30 LT – 09:30 LT) and standard deviation to be small at 2.03 and 2.56%, respectively. For the long-term trends, across all regions and overpass times, all of the UKESM trends were smaller than ±0.5 DU/decade. Therefore, the model $LTCO_3$ regional trends are negligibly different between overpass times. This might not be surprising given the negligible model trends in the satellite spatio-temporal trend comparisons (see **Section 3.1**), but the actual absolute differences (average and range) in simulated $LTCO_3$ are also small supporting the argument that on the regional scale, the day-time diurnal cycle differences between satellite overpass times has limited influence on the reported satellite trend discrepancies (e.g. in Gaudel et al., 2018).

~~2.~~4. Discussion

Investigation of satellite $LTCO_3$ focussed on 2008 to 2017, representing a decade of overlap of the OMI and IASI records. The analysis focussed on North America, Europe and East Asia as these regions are subject to large emissions of and temporal changes in O_3 precursor gases. $LTCO_3$ is typically spatially homogeneous with shallow gradients between background and source-induced O_3 concentrations. Secondly, individual retrievals of $LTCO_3$ are often associated with large uncertainties (e.g. random and systematic uncertainties).

Formatted: Font colour: Auto

Formatted: Indent: Left: 0.63 cm, Hanging: 0.63 cm, Outline numbered + Level: 1 + Numbering Style: 1, 2, 3, … + Start at: 1 + Alignment: Left + Aligned at: 0.63 cm + Indent at: 1.27 cm

Formatted: Font colour: Auto

There are multiple contributory factors concerning both instrumental attributes (notably spectroradiometric noise and calibration accuracy) and variability in geophysical variables which influence radiative transfer and vertical sensitivity (e.g. stratospheric ozone, cloud and aerosol, water vapour, surface spectral reflectivity/emissivity and pressure and temperature profile) which can result in $LTCO_3$ time-series with substantial variability/noise when derived at high spatial resolution (e.g. when deriving time-series from data gridded at 0.5° or 1.0°). Therefore, we undertake our analysis at the regional (e.g. continental) scale where more satellite retrievals are included in time-series monthly means yielding a reduction in the random error component of the sample.Secondly, individual retrievals of $LTCO_3$ are subject to multiple issues (e.g. influences on radiative transfer and vertical sensitivity of stratospheric ozone, cloud and other particulates, surface spectral reflectivity/emissivity and temperature profile) which can result in noisy $LTCO_3$ time-series at high resolution (e.g. when gridded on a scale of 0.5° or 1.0°). Both of these factors supported analysis at a regional scale (e.g. continental scale).

Ideally, this analysis would have utilised several more records (e.g. several UV-Vis and IR products) to quantify long-term trends in $LTCO_3$ and investigate the potential reasons for any discrepancies, as shown by Gaudel et al., (2018) for TCO_3. While RAL Space, and other providers, have generated UV-Vis profile O_3 products for more instruments, e.g. from the Global Ozone Monitoring Experiment 1 & 2 (GOME-1 & GOME-2) and the SCanning Imaging Absorption spectroMeter for Atmospheric CartograpHY (SCIAMACHY), the GOME-1 and SCIAMACHY records do not overlap for as long with IASI and step changes in the GOME-2A Level-1 processing scheme used to produce the available $LTCO_3$ Level-2 version mean it is not sufficiently homogeneous (see Pope et al., (2023)). For the IR instruments, other potential sensors include the Tropospheric Emissions Spectrometer (TES; Richards et al., 2008) and the RAL Space IASI Extended Infrared Microwave Sounding (IMS; Pimlott et al., 2022) scheme applied to IASI. Unfortunately, the TES record only covers 2005 to 2013, with decreasing spatial coverage with time, and at the time of this work the IASI-IMS product had only been processed on a sub-sampled basis of 1 in 10 days.

In this work, we some find discrepancies in the observed long-term tendencies from the utilised $LTCO_3$ products in these northern hemispheric regions. The OMI product is subject to large-scale interannual variability over the 2008-17 decade, in comparison with which the underlying linear trends are small in absolute terms with large confidence ranges (i.e. 95% confidence intervals) intersecting with zero. . However, the OMI $LTCO_3$ product has been shown to be stable over this period relative to ozonesondes by Pope at el., (2023). IASI-FORLI has substantial negative $LTCO_3$ tendencies, but this is driven by a specific discontinuity in 2010 due to inhomogeneity in Eumetsat (water vapour, temperature) data used in IASI-FORLI Level-2 processing (Boynard et al., 2018; Wespes et al., 2018). It induces an artificial drift that explains the substantial negative $LTCO_3$ trends reported here and in Gaudel et al., (2018). The IASI-SOFRID $LTCO_3$ and apriori are very similar, with little inter-annual variability, which suggests that the IASI-SOFRID O_3 retrieval in this height-range is more constrained by the apriori (i.e. i.e. less TO_3 sensitivity than the other products – see **S3**less TO_3 sensitivity than the other products). Importantly, analysis of the three products' apriori $LTCO_3$ records show negligible trends meaning that year-to-year spatiotemporal sampling differences (i.e. the number of retrievals used in the spatial-monthly regional averages) are not skewing long-term satellite trends. In summary: any underlying linear trend in $LTCO_3$ occurring during the decade 2008-17 was masked by interannual variability in the OMI retrieval and by constraint to the apriori in the IASI-SOFRID retrieval and, although substantial for IASI-FORLI retrieval, that is that is due to changing meteorological inputs to the data processing (Boynard et al., 2018; Wespes et al., 2018).believed to be due to changing meteorological input to the data processing.

Formatted: Font colour: Auto

Formatted: Font colour: Auto

Formatted: Font colour: Auto

For UKESM1.0, the model exhibits negligible temporal variability in $LTCO_3$ for all regions and instruments' samplings. Modelled $LTCO_3$ trends never exceeded 1.0 DU/decade in magnitude, all of which were deemed to be insignificant due to large associated p-values by the linear-seasonal trend model detailed in **Section 2.~~4~~** 3 and **Equations** ~~4~~ 2 & ~~5~~3. The ozonesondes for each region were included to ground truth the model and satellite trends. The North American sites' $LTCO_3$ time-series was relatively noisy and exhibited considerable inter-annual variability in its seasonal cycle. The comparatively low level of inter-annual variability in the European UKESM1.0 record of $LTCO_3$ was in good agreement with the ozonesondes, however, and so was its low trend, providing confidence in the model over that region. For East Asia, the interannual variability differed substantially between UKESM1.0 and ozonesondes and the reported ozonesonde trend was significantly much larger than for UKESM1.0. Therefore, when considering UKESM1.0 and the ozonesondes, no consistent $LTCO_3$ trends can be determined for any of the regions. Overall, taking all data sets into account, LTO_3 appears to have neither increased nor decreased markedly over these three regions between the beginning and end of the study decade (i.e. 2008 to 2017).

One key aspect of this work was to exploit UKESM1.0 to determine the importance of vertical sensitivity on retrieved LTO_3 and how this influences the reported long-term tendency. In terms of the absolute model trends (with and without the satellite AKs), the impact on $LTCO_3$ was small with typically near-zero tendencies and large uncertainty ranges (i.e. the 95% confidence interval). In relative terms, the changes in model trend values were more substantial in the order of 100%. To explore this further, the UKESM1.0 $LTCO_3$ time-series (with and without the satellite AKs) were sorted from lowest to highest (based on annual averages) to impose the most substantial trend in the model data. When the trends were re-calculated, the largest model $LTCO_3$ trends ranged between 0.37 and 0.90 DU/decade. When the AKs were applied, the $LTCO_3$ trends ranged from -0.74 to 1.02 DU/decade. Again, in relative terms, this represents a relatively large impact of the AKs on simulated LTO_3 tendencies but in absolute terms, these are small changes. Though, it should be noted that many of the 95% confidence intervals for these trends either shifted to intersect with zero or vice versa once the AKs were applied to the model. Gaudel et al., (2018) suggested two potential reasons for the TCO_3 trend discrepancies in their study:

- Time varying instrument biases/drift.
- The impact of satellite vertical sensitivity.

A further two important reasons are:

- Changes over time in latitude/longitude domains sampled by satellite sensors (e.g. GOME-1 has substantial issues after 2003).
- The time-period used for the trend analysis.

As stated by Boynard et al., (2018) and Wespes et al., (2018), the IASI-FORLI-v20151001 product ~~According to Boynard et al., (2018) and Wespes et al., (2018), the IASI-FORLI-v20151001 products~~ has an artificial negative drift with time explained by a discontinuity found in the Level-2 meteorological inputs taken from Eumetsat. However, in the near future, a new consistent IASI-FORLI ozone climate data record will be available using homogeneous Level-1 and Level-2 Eumetsat meteorological data. Analysis of OMI $LTCO_3$ by Pope et al., 2023 showed OMI $LTCO_3$ to be temporally stable against ozonesondes. A similar analysis (not shown here) indicates IASI-SOFRID $LTCO_3$ to also be temporally stable with near-zero drift in bias. For the satellite vertical sensitivity, some of our results were unexpected. While the application of the AKs to UKESM1.0 can substantially shift the simulated absolute LTO_3 values and squash/stretch the seasonal amplitude, the impact on the simulation $LTCO_3$ tendencies are small in absolute terms. In relative terms, the impacts can be large (e.g. 100% change in trend rate). However, as the UKESM1.0 simulated $LTCO_3$ trends are generally near-zero, it is difficult to confidently say either way if the vertical sensitivity, when retrieving $LTCO_3$, is important for influencing long-term tendencies, even when a more substantial trend is forced upon UKESM1.0. Future work on this would probably need to look at artificial model data which already has substantial TO_3 trends in it (e.g. 5.0 or 10.0 DU/decade). This will obviously not match reality but would provide some further quantification on how important vertical sensitivity is from different instruments/sounders in LTO_3 trend determination.

As for year-to-year spatiotemporal sampling, our results suggest negligible trends for the product $LTCO_3$ apriori time-series and thus monthly sampling biases are unlikely to be introducing artificial trends as the apriori datasets are trendless. Finally, the time-period over which the trend analysis is undertaken is critically important. Gaudel et al., (2018), using the available data at the time, focussed on 2005-2015/6 and 2008-2015/6 for the OMI and IASI products they used. For the IASI products, using a slightly extended time-period, the trends show similar tendencies. However, for OMI, 2016 and 2017 represent lower years of TO_3. As a result, this dampens the strong significant positive trends reported by Gaudel et al., (2018) in TCO_3. It is notable that the substantial positive increase in tropical LTO_3 between 1995 and 2017 reported by Pope et al., (2023) from a series of UV-Vis sounders, included the same OMI global dataset as that is used here, further suggests the selection of time period and geographical region to be crucial in regard to the role of interannual variability on linear trend detection.

~~3.~~5. Conclusions

Gaudel et al., (2018) undertook a multi-satellite analysis of long-term trends in tropospheric column ozone (TCO_3). They found large scale differences between these products with no clear consensus on the signs or drivers of these TCO_3 trends. To avoid complications with tropopause definition and reduce influence of stratospheric ozone on retrieved values, this study has undertaken a detailed follow-up assessment of decadal trends in $LTCO_3$ (surface – 450 hPa layer) rather than TCO_3 exploiting ozonesonde records, model simulations and accounting carefully for satellite O_3 metrics (e.g. averaging kernels, AKs, apriori information and satellite uncertainties). We have focussed on $LTCO_3$ data sets from Ozone Monitoring Instrument (OMI) produced by the RAL Space scheme and from Infrared Atmospheric Sounding Interferometer produced by the IASI-FORLI and IASI-SOFRID schemes, for which there were consistent records from 2008-2017.

Evaluation of satellite LTO_3 from these three products over the North American, European and East Asian regions resulted in linear trends which varied over a small range close to zero and with confidence intervals intersecting with zero. This was consistent with simulations from the UK Earth System Model (UKESM1.0). There were no large-scale trends in the apriori information, so changes in satellite year-to-year spatiotemporal sampling has not been driving inconsistencies between products. When convolving UKESM1.0 with the satellite AKs (i.e. to assess the impact of the satellite vertical sensitivity) it did change the size of the model trend, and in some instances, the direction of the trend, but as the simulated LTO_3 trends were small and insignificant, they had limited influence. Overall, our results show that changes in LTO_3 during the decade 2008-2017 in North America, Europe and East Asia were dominated by variability in processes which control LTO_3 on shorter timescales.

In the near future, the new European polar orbiting mission MetOp Second Generation will include IASI Next Generation and Sentinel-5 UV/VIS sounders to provide height-resolved ozone products to extend current missions through to the mid-2040s. This will be supplemented by the new USA Near Earth Orbit Network (NEON) series as a replacement for the Joint Polar Satellite System (JPSS). The Geostationary Environment Monitoring Spectrometer (GEMS) and Tropospheric Emissions: Monitoring of Pollution (TEMPO) have also recently been launched and there will be new geostationary platforms: the Infrared Sounder (IRS) and

Formatted: Indent: Left: 0.63 cm, Hanging: 0.63 cm, Outline numbered + Level: 1 + Numbering Style: 1, 2, 3, … + Start at: 1 + Alignment: Left + Aligned at: 0.63 cm + Indent at: 1.27 cm

Formatted: Font colour: Auto

Sentinel-4 UV/VIS sounder on Europe's Meteosat-Third Generation (MTG-S), again through to the mid-2040s, and the USA Geostationary Extended Observations (GeoXO) series. Overall, these platforms will provide large volumes of data (e.g. diurnal observations) and over a long-time scale on tropospheric ozone for future regional trend analyses.~~In the future, new polar orbiting missions including the IASI Next Generation and Sentinel-5 UV/VIS sounders on the MetOp Second Generation will provide tropospheric ozone products to extend current missions through to the mid-2040s. There will also be the new geostationary platforms like the Infrared Sounder (IRS) and Sentinel-4 UV/VIS sounder on Meteosat-Third Generation (MTG-S) and the already in orbit Geostationary Environment Monitoring Spectrometer (GEMS) and Tropospheric Emissions: Monitoring of Pollution (TEMPO), which will provide large volumes of data (e.g. diurnal observations) and over a long-time scale on tropospheric ozone for future regional trend analyses.~~

Acknowledgements

This work was funded by the UK Natural Environment Research Council (NERC) by providing funding for the National Centre for Earth Observation (NCEO, award reference NE/R016518/1) and funding from the European Space Agency (ESA) Climate Change Initiative (CCI) post-doctoral fellowship scheme (award reference 4000137140). For UKESM1.0 model runs, we acknowledge use of the Monsoon2 system, a collaborative facility supplied under the Joint Weather and Climate Research Programme, a strategic partnership between the Met Office and NERC. IASI is a joint mission of EUMETSAT and the Centre National d'Etudes Spatiales (CNES, France). The IASI-SOFRID research was conducted at LAERO with some financial support from the CNES French spatial agency (TOSCA–IASI project). The authors acknowledge the AERIS data infrastructure for providing access to the IASI-FORLI data, ULB-LATMOS for the development of the FORLI retrieval algorithm, and the AC SAF project of the EUMETSAT for providing IASI-FORLI data used in this paper. Anna Maria Trofaier (ESA Climate Office) provided support and advice throughout the fellowship.

Data Availability

The IASI-FORLI and IASI-SOFRID data can be obtained from https://iasi.aeris-data.fr/O3 and https://iasi-sofrid.sedoo.fr/. The RAL OMI data is available via the NERC Centre for Environmental Data Analysis (CEDA) Jasmin platform subject to data requests. However, the RAL Space satellite data, as well as the UKESM1.0 simulations, will be uploaded to the Zenodo open access portal (https://zenodo.org/) if this manuscript is accepted for publication in ACP after the peer-review process. The ozonesonde data for WOUDC, SHADOZ and NOAA is available from https://woudc.org/, https://tropo.gsfc.nasa.gov/shadoz/ and https://gml.noaa.gov/ozwv/ozsondes/.

Author Contributions

RJP conceptualised, planned and undertook the research study. BB, ELF, BJK, RS, BGL, AB and CW provided the OMI and IASI ozone data and advice on using the products and their analysis. FO and MD provided advice and expertise on using and running UKESM. CR provided advice and help during RP's ESA CCI fellowship. Scientific and technical contributions came from MPC, WF, MAP, SSD and RR. RJP prepared the manuscript with input from all co-authors. ~~RJP conceptualised, planned and undertook the research study. BB, ELF, BJK, RS, BGL, LJV, AB and CW provided the OMI and IASI ozone data and advice on using the products. FO and MD provided advice and expertise on using and running UKESM. CR provided advice and help during RP's ESA CCI fellowship. RJP prepared the manuscript with scientific and technical contributions from all co-authors.~~

Formatted: Font colour: Auto

Conflicts of Interest

The authors declare no conflicts of interest.

References:

Archibald, A.T., O'Connor, F.M., Abraham, N.L., Archer-Nicholls, S., Chipperfield, M.P., Dalvi, M., Folberth, G.A., Dennison, F., Dhomse, S.S., Griffiths, P.T., Hardacre, C., Hewitt, A.J., Hill, R.S., Johnson, C.E., Keeble, J., Kohler, M.O., Morgenstern, O., Mulcahy, J.P., Ordonez, C., Pope, R.J., Rumbold, S.T., Russo, M.R., Savage, N.H., Sellar, A., Stringer, M., Turnock., S.T., Wild, O. and Zeng, G.: Description and evaluation of the UKCA stratosphere–troposphere chemistry scheme (StratTrop vn 1.0) implemented in UKESM1. *Geoscientific Model Development*, **13**, 1223–1266, doi: 10.5194/gmd-13-1223-2020, 2020.

Barret, B., Emili, E., Le Flochmoen, E. 2020. A tropopause-related climatological a priori profile for IASI-SOFRID ozone retrievals: improvements and validation. *Atmospheric Measurement Techniques*, **13**, 5237–5257, doi: 10.5194/amt-13-5237-2020.

Boersma, K.F., Jacob, D.J., Eskes, H.J., Pinder, R.W., Wang, J. and van der A, R.J.: Intercomparison of SCIAMACHY and OMI tropospheric NO_2 columns: Observing the diurnal evolution of chemistry and emissions from space. *Journal of Geophysical Research: Atmospheres*, **113 (D16S26)**, doi: 10.1029/2007JD008816, 2008.

Boersma, K. F., Eskes, H. J., Dirksen, R. J., van der A, R. J., Veefkind, J. P., Stammes, P., Huijnen, V., Kleipool, Q. L., Sneep, M., Claas, J., Leitão, J., Richter, A., Zhou, Y., and Brunner, D.: An improved tropospheric NO_2 column retrieval algorithm for the Ozone Monitoring Instrument, *Atmospheric Measurement Techniques*, 4, 1905–1928, doi: 10.5194/amt-4-1905-2011, 2011.

Boynard, A., Clerbaux, C., Coheur, P.-F., Hurtmans, D., Turquety, S., George, M., Hadji-Lazaro, J., Keim, C., and Meyer-Arnek, J.: Measurements of total and tropospheric ozone from IASI: comparison with correlative satellite, ground-based and ozonesonde observations, Atmos. Chem. Phys., 9, 6255–6271, https://doi.org/10.5194/acp-9-6255-2009, 2009.

Boynard, A., Hurtmans, D., Garane, K., Goutail, F., Hadji-Lazaro, J., Koukouli, M. E., Wespes, C., Vigouroux, C., Keppens, A., Pommereau, J.-P., Pazmino, A., Balis, D., Loyola, D., Valks, P., Sussmann, R., Smale, D., Coheur, P.-F., and Clerbaux, C.: Validation of the IASI FORLI/EUMETSAT ozone products using satellite (GOME-2), ground-based (Brewer–Dobson, SAOZ, FTIR) and ozonesonde measurements, *Atmospheric Measurement Techniques*, **11**, 5125–5152, https://doi.org/10.5194/amt-11-5125-2018, 2018.

Chang, K. L., Cooper, O. R., Gaudel, A., Petropavlovskikh, I., & Thouret, V.: Statistical regularization for trend detection: an integrated approach for detecting long-term trends from sparse tropospheric ozone profiles. Atmospheric Chemistry and Physics, 20(16), 9915-9938, 2020.

Clerbaux, C., Boynard, A., Clarisse, L., George, M., Hadji-Lazaro, J., Herbin, H., Hurtmans, D., Pommier, M., Razavi, A., Turquety, S., Wespes, C. and Coheur, P.F.: Monitoring of atmospheric composition using the thermal infrared IASI/MetOp sounder, *Atmospheric Chemistry and Physics*, 9 (16), 6041–6054, doi:10.5194/acp-9-6041-2009, 2009.

Doherty, R. M., Heal, M. R., and O'Connor, F. M.: Climate change impacts on human health over Europe through its effect on air quality, *Environmental Health*, 16(1), 33–44, doi:10.1186/s12940-017-0325-2, 2017.

ESA. 2019. Climate Change Initiative. http://cci.esa.int/ozone (last accessed 01/09/2022).

Eskes HJ and Boersma KF. 2003. Averaging kernels for DOAS total column satellite retrievals. *Atmospheric Chemistry and Physics*, **3**, 1285–1291, doi: 10.5194/acp-3-1285-2003.

European Commission, Joint Research Centre, Dentener F, et al. 2016. Hemispheric Transport of Air Pollution (HTAP): specification of the HTAP2 experiments: ensuring harmonized modelling, Publications Office, https://data.europa.eu/doi/10.2788/725244.

Formatted: Line spacing: single

Formatted: Font:

Formatted: Font colour: Auto

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geosci. Model Dev., 9, 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016, 2016.Gaudel A, et al. 2018. Tropospheric Ozone Assessment Report: Present day distribution and trends of tropospheric ozone relevant to climate and global atmospheric chemistry model evaluation. *Elementa*, **6 (39)**, 1-58, doi: 10.1525/elementa.291.

Fiore, Arlene M., et al. (2022), Understanding recent tropospheric ozone trends in the context of large internal variability: A new perspective from chemistry-climate model ensembles, Environmental Research: Climate, https://doi.org/10.1088/2752-5295/ac9cc2.

Forster, P., Storelvmo, T., Armour, K., Collins, W., Dufresne, J.- L., Frame, D., Lunt, D. J., Mauritsen, T., Palmer, M. D., Watanabe, M., Wild, M., and Zhang, H.: The Earth's Energy Budget, Climate Feedbacks, and Climate Sensitivity, in: Climate Change 2021: The Physical Science Basis, Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M. I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J. B. R.,Maycock, T. K., Waterfield, T., Yelekçi, O., Yu, R., and Zhou, B., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 923– 1054, doi:10.1017/9781009157896.009, 2021.

Gaudel, A., Cooper, O.R., Ancellet, G., Barret, B., Boynard, A., Burrows, J.P., Clerbaux, C., Coheur, P.F., Cuesta, J., Cuevas, E., Doniki, S., Dufour, G., Ebojie, F., Foret, G., Garia, O., Granados-Munoz, M.J., Hannigan, J.W., Hase, F., Hassler, B., Huang, G., Hurtmans, D., Jaffe, D., Jones, N., Kalabokas, P., Kerridge, B., Kulwaik, S., Latter, B., Leblanc, T., Le Flochmoen, E., Lin, W., Liu, J., Liu, X., Mahieu, E., McClure-Begley, A., Neu, J.L., Osman, M., Palm, M., Petetin, H., Petropavlovskikh, I., Querel, R., Rahpoe, N., Rozanov, A., Schultz, M.G., Schwab, J., Siddans, R., Smale, D., Steinbacher, M., Tanimoto, H., Tarasick, D.W., Thouret, V., Thompson, A.M., Trickl, T., Weatherhead, E., Wespes, C., Worden, H.M., Vigouroux, C., Xu, X., Zeng, G. and Ziemke, J..: Tropospheric Ozone Assessment Report: Present day distribution and trends of tropospheric ozone relevant to climate and global atmospheric chemistry model evaluation. *Elementa*, 6(39), 1-58, doi:10.1525/elementa.291, 2018.

Gauss, M., Myhre, G., Isaksen, I. S. A., Grewe, V., Pitari, G., Wild, O., Collins, W. J., Dentener, F. J., Ellingsen, K., Gohar, L. K., Hauglustaine, D. A., Iachetti, D., Lamarque, F., Mancini, E., Mickley, L. J., Prather, M. J., Pyle, J. A., Sanderson, M. G., Shine, K. P., Stevenson, D. S., Sudo, K., Szopa, S., and Zeng, G.: Radiative forcing since preindustrial times due to ozone change in the troposphere and the lower stratosphere, *Atmospheric Chemistry and Physics*, **6**, 575–599, https://doi.org/10.5194/acp-6-575-2006, 2006.

Hoesly, R. M., Smith, S. J., Feng, L., Klimont, Z., Janssens-Maenhout, G., Pitkanen, T., Seibert, J. J., Vu, L., Andres, R. J., Bolt, R. M., Bond, T. C., Dawidowski, L., Kholod, N., Kurokawa, J.-I., Li, M., Liu, L., Lu, Z., Moura, M. C. P., O'Rourke, P. R., and Zhang, Q.: Historical (1750–2014) anthropogenic emissions of reactive gases and aerosols from the Community Emissions Data System (CEDS), *Geoscientific Model Development*, 11, 369–408, https://doi.org/10.5194/gmd-11-369-2018, 2018.

Hollaway, M.J., Arnold, S.R., Challinor, A. J. and Emberson, L.D: Intercontinental trans-boundary contributions to ozone-induced crop yield losses in the North Hemisphere, *Biogeosciences*, *9*, 271–2929, doi: 10.5194/bg-9-271-2012, 2012.

Hubert, D., Lambert, J-C., Verhoelst, T., Granville, J., Keppens, A., Baray, J-L., Bourassa, A.E., Cortesi, U., Degenstein, D.A., Froidevaux, L., Godin-Beekmann, S., Hoppel, K.W., Johnson, B.L., Kyrola, E., Leblanc, T., Lichtenberg, G., Marchand, M., McElroy, C.T., Murtagh, D., Nakane, H., Portafaix, T., Querel, R., Russell, J.M.,

Formatted: Font colour: Auto

Salvador, J., Smit, H.G.J., Stebel, K., Steinbrecht, W., Strawbridge, K.B., Stubi, R., Swart, D.P.J., Taha, G., Tarasick, D.W., Thompson, A.M., Urban, J., van Gijsel, J.A.E., Van Malderen, R., von der Gathen P., Walker, K.A., Wolfram, E. and Zawodny, J.M.: Ground-based assessment of the bias and long-term stability of 14 limb and occultation ozone profile data records. *Atmospheric Measurement Techniques*, 9, 2497-2534, doi: 10.5194/amt-9-2497-2016, 2016.

Illingworth, S. M., Remedios, J. J., Boesch, H., Moore, D. P., Sembhi, H., Dudhia, A., and Walker, J. C.: ULIRS, an optimal estimation retrieval scheme for carbon monoxide using IASI spectral radiances: sensitivity analysis, error budget and simulations, *Atmospheric Measurements Techniques*, 4, 269–288, doi: 10.5194/amt-4-269-2011, 2011.

Keim, C., Eremenko, M., Orphal, J., Dufour, G., Flaud, J.-M., Höpfner, M., Boynard, A., Clerbaux, C., Payan, S., Coheur, P.-F., Hurtmans, D., Claude, H., Dier, H., Johnson, B., Kelder, H., Kivi, R., Koide, T., López Bartolomé, M., Lambkin, K., Moore, D., Schmidlin, F. J., and Stübi, R.: Tropospheric ozone from IASI: comparison of different inversion algorithms and validation with ozone sondes in the northern middle latitudes, *Atmospheric Chemistry and Physics*, **9**, 9329–9347, doi: 10.5194/acp-9-9329-2009, 2009.

Keppens, A., Lambert, J-C., Graville, J., Hubert, D., Verhoelst, T., Compernolle, S., Latter, B., Kerridge, B., Siddans, R., Boynard, A., Hadji-Lazaro, J., Clerbaux, C., Wespes, C., Hurtmans, D.R., Coheur, P-F., van Peet, J.C.A., van der A, R.J., Garane, K., Koukouli, M.E., Balis, D.S., Delcloo, A., Kivi, R., Stubi, R., Godin-Beekmann, S., Van Roozendael, M. and Zehner, C.: Quality assessment of the Ozone_cci Climate Research Data Package (release 2017) – Part 2: Ground-based validation of nadir ozone profile data products. *Atmospheric Measurement Techniques*, 11, 3769-3800, doi: 10.5194/amt-11-3769-2018, 2018.

Keppens, A., Compernolle, S., Verhoelst, T., Hubert, D., and Lambert, J.-C.: Harmonization and comparison of vertically resolved atmospheric state observations: methods, effects, and uncertainty budget, Atmos. Meas. Tech., 12, 4379–4391, https://doi.org/10.5194/amt-12-4379-2019, 2019.

Lamarque, J.-F., Bond, T. C., Eyring, V., Granier, C., Heil, A., Klimont, Z., Lee, D., Liousse, C., Mieville, A., Owen, B., Schultz, M. G., Shindell, D., Smith, S. J., Stehfest, E., Van Aardenne, J., Cooper, O. R., Kainuma, M., Mahowald, N., McConnell, J. R., Naik, V., Riahi, K., and van Vuuren, D. P.: Historical (1850–2000) gridded anthropogenic and biomass burning emissions of reactive gases and aerosols: methodology and application, *Atmospheric Chemistry and Physics*, **10**, 7017–7039, doi: 10.5194/acp-10-7017-2010, 2010.

Matthes, K., Funke, B., Andersson, M. E., Barnard, L., Beer, J., Charbonneau, P., Clilverd, M. A., Dudok de Wit, T., Haberreiter, M., Hendry, A., Jackman, C. H., Kretzschmar, M., Kruschke, T., Kunze, M., Langematz, U., Marsh, D. R., Maycock, A. C., Misios, S., Rodger, C. J., Scaife, A. A., Seppälä, A., Shangguan, M., Sinnhuber, M., Tourpali, K., Usoskin, I., van de Kamp, M., Verronen, P. T., and Versick, S.: Solar forcing for CMIP6 (v3.2), *Geoscientific Model Development*, 10, 2247–2302, https://doi.org/10.5194/gmd-10-2247-2017, 2017.

McPeters, R.D., Labow, G.J., and Logan, J.A. 2007. Ozone climatological profiles for satellite retrieval algorithms. *Journal of Geophysical Research*, **112** (D05308), https://doi.org/10.1029/2005JD006823.

Miles, G.M., Siddans, R., Kerridge, B. J., Latter, B. G., and Richards, N. A. D.: Tropospheric ozone and ozone profiles retrieved from GOME-2 and their validation, *Atmospheric Measurement Techniques*, 8, 385–398, doi:10.5194/amt-8-385-2015, 2015.

Monks, S.A., Arnold, S.R., Hollaway, M. J., Pope, R.J., Wilson, C., Feng, W., Emmerson, K.E., Kerridge, B.J., Latter, B.L., Miles, G.M., Siddans, R. and Chipperfield, M.P.: The TOMCAT global chemistry transport model v1.6: Description of chemical mechanism and model evaluation, *Geoscientific Model Development*, 10 (8), 3025–3057, doi:10.5194/gmd-10-3025-2017, 2017.

Formatted: Font colour: Auto

Formatted: Font: +Body (Calibri), 11 pt, Font colour: Auto

Formatted: Default, Line spacing: single, Adjust space between Latin and Asian text, Adjust space between Asian text and numbers

Myhre, G., Shindell, D., Bréon, F.-M, Collins, W., Fuglestvedt, J., Huang, J., Koch, D., Lamarque, J.-F., Lee, D., Mendoza, B., Nakajima, T., Robock, A., Stephens, G., Takemura, T. and Zhang, H.: Anthropogenic and Natural Radiative Forcing, in: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 659–740, 2013.

O'Connor, F. M., Johnson, C. E., Morgenstern, O., Abraham, N. L., Braesicke, P., Dalvi, M., Folberth, G. A., Sanderson, M. G., Telford, P. J., Voulgarakis, A., Young, P. J., Zeng, G., Collins, W. J., and Pyle, J. A.: Evaluation of the new UKCA climate-composition model – Part 2: The Troposphere, *Geoscientific Model Development*, 7, 41–91, doi: 10.5194/gmd-7-41-2014, 2014.

Pimlott, M.A., Pope, R.P., Kerridge, B.J., Latter, B.G., Knappett, D.S., Heard, D.E., Ventress, L.J., Siddans, R., Feng, W. and Chipperfield, M.P.: Investigating the global OH radical distribution using steady-state approximations and satellite data. *Atmospheric Chemistry and Physics,* **22**, 10467-10488, doi: 10.5194/acp-22-10467-2022, 2022.

Pope, R.J., Arnold, S.R., Chipperfield, M.P., Latter, B.G., Siddans, R. and Kerridge, B.J.: Widespread changes in UK air quality observed from space. *Atmospheric Science Letters*, **19:e817**, doi: 10.1002/asl.817.

Pope, R. J., Kerridge, B. J., Siddans, R., Latter, B. G., Chipperfield, M. P., Feng, W., Pimlott, M. A., Dhomse, S. S., Retscher, C., and Rigby, R.: Investigation of spatial and temporal variability in lower tropospheric ozone from RAL Space UV–Vis satellite products, *Atmospheric Chemistry and Physics*, 23, 14933–14947, https://doi.org/10.5194/acp-23-14933-2023, 2023.

Pope, R. J., Rap, A., Pimlott, M. A., Barret, B., Le Flochmoen, E., Kerridge, B. J., Siddans, R., Latter, B. G., Ventress, L. J., Boynard, A., Retscher, C., Feng, W., Rigby, R., Dhomse, S. S, Wespes, C. and Chipperfield, M. P.: Quantifying the tropospheric ozone radiative effect and its temporal evolution in the satellite era, Atmospheric Chemistry and Physics, 24, 3613-3626, doi: 10.5194/acp-24-3613-2024, 2024.

Rao, S., Klimont, Z., Smith, S.J., Van Dingenen, R., Dentener, F., Bouwman, L., Riahi, K., Amann, M., Bodirsky, B.L., van Vuuren, D.P., Reus, L.R., Calvin, K., Drouet, L., Fricko, O., Fujimori, S., Gernaat, D., Havlik, P., Harmsen, M., Hasegawa, T., Heyes, C. and Tavoni, M.,: Future air pollution in the shared socio-economic pathways. *Global Environmental Change*, **42**, 346-358, doi: 10.1016/j.gloenvcha.2016.05.012.

Richards, N.A.D, Osterman, G.B., Browell, E.V., Hair, J.W., Avery, M. and Li, Q.: Validation of tropospheric emission spectrometer ozone profiles with aircraft observations during the intercontinental chemical transport experiment–B. *Journal Geophysical Research*, **113**(D16S29), doi: 10.1029/2007JD008815, 2008.

Russo, M. R., Kerridge, B. J., Abraham, N. L., Keeble, J., Latter, B. G., Siddans, R., Weber, J., Griffiths, P. T., Pyle, J. A., and Archibald, A. T.: Seasonal, interannual and decadal variability of tropospheric ozone in the North Atlantic: comparison of UM-UKCA and remote sensing observations for 2005–2018, Atmos. Chem. Phys., 23, 6169–6196, https://doi.org/10.5194/acp-23-6169-2023, 2023.

Sellar, A.A., Jones, C.G., Mulcahy, J.P., Tang, Y., Yool, A., Wiltshire, A., O'Connor, F.M., Stringer, M., Hill, R., Palmieri, J., Woodward, S., de Mora, L., Kuhlbrodt, T., Rumbold, S.T., Kelley, D.I., Ellis, R., John, C.E., Walton, J., Abraham, N.L., Andrews, M.B., Andrews, T., Archibald, A.T., Berthou, S., Burke, E., Blockley, E., Carslaw, K., Dalvi, M., Edwards, J., Folbert, G.A., Gedney, N., Griffiths, P.T., Harper, A.B., Hendry, M.A., Hewitt, A.J., Johnson, B., Jones, A., Jones, C.D., Keebie, J., Liddicoat, S., Morgenstern, O., Parker, R.J., Predoi, V., Robertson, E., Siahaan, A., Smith, R.S., Swaminathan, R., Woodhouse, M.T., Zeng, G. and Zerroukat, M.: Description and Evaluation of the UK Earth System Model. *Journal of Advances in Modeling Earth Systems*, **11**, 4513-4558, doi: 10.1029/2019MS001739, 2019.

Formatted: Font colour: Auto

Formatted: Line spacing: single

Sindelarova, K., Granier, C., Bouarar, I., Guenther, A., Tilmes, S., Stavrakou, T., Müller, J.-F., Kuhn, U., Stefani, P., and Knorr, W.: Global data set of biogenic VOC emissions calculated by the MEGAN model over the last 30 years, Atmos. Chem. Phys., 14, 9317–9341, doi: 10.5194/acp-14-9317-2014, 2014.

Sitch, S., Cox, P.M., Collins, W.J., & Huntingford, C.: Indirect radiative forcing of climate change through ozone effects on the land carbon sink, *Nature*, 448, 791–795, doi:10.1038/nature06059, 2007.

Sofieva, V. F., Tamminen, J., Kyrölä, E., Mielonen, T., Veefkind, P., Hassler, B., and Bodeker, G. E.: A novel tropopause-related climatology of ozone profiles, Atmos. Chem. Phys., 14, 283–299, https://doi.org/10.5194/acp-14-283-2014, 2014.

Telford, P. J., Braesicke, P., Morgenstern, O., and Pyle, J. A.: Technical Note: Description and assessment of a nudged version of the new dynamics Unified Model, *Atmospheric Chemistry and Physics*, 8, 1701–1712, https://doi.org/10.5194/acp-8-1701-2008, 2008.

van der A, R.J., Peters, D.H.M.U., Eskes, H., Boersma, K.F., Van Roozendael, M., De Smedt, I. and Kelder, H.M.: Detection of the trend and seasonal variation in tropospheric NO_2 over China. *Journal of Geophysical Research*, **11**, D12317, doi: 10.1029/2005JD006594, 2006.

van der A, R.J., Eskes, H.J., Boersma, K.F., van Noije, T.P.C., Van Roozendael, M., De Smedt, I., Peters, D.H.M.U. and Meijer, E.W.: Trends, seasonal variability and dominant NO_x sources derived from a ten year record of NO_2 measured from space. *Journal of Geophysical Research*, **113**, 1–12, doi: 10.1029/2007JD009021, 2008.

van der Werf, G. R., Randerson, J. T., Giglio, L., van Leeuwen, T. T., Chen, Y., Rogers, B. M., Mu, M., van Marle, M. J. E., Morton, D. C., Collatz, G. J., Yokelson, R. J., and Kasibhatla, P. S.: Global fire emissions estimates during 1997–2016, Earth System Science Data, 9, 697–720, https://doi.org/10.5194/essd-9-697-2017, 2017.

van Marle, M. J. E., Kloster, S., Magi, B. I., Marlon, J. R., Daniau, A.-L., Field, R. D., Arneth, A., Forrest, M., Hantson, S., Kehrwald, N. M., Knorr, W., Lasslop, G., Li, F., Mangeon, S., Yue, C., Kaiser, J. W., and van der Werf, G. R.: Historic global biomass burning emissions for CMIP6 (BB4CMIP) based on merging satellite observations with proxies and fire models (1750–2015), Geoscientific Model Development, 10, 3329–3357, https://doi.org/10.5194/gmd-10-3329-2017, 2017.

Weatherhead, E.C., Reinsel, G.C., Tiao, G.C., Meng, X., Choi, D., Cheang, W., Keller, T., De Luisi, J., Wuebbles, D.J., Kerr, J.B., Miller, A.J., Oltmans, S.J. and Frederick, J.E.: Factors affecting the detection of trends: statistical considerations and applications to environmental data. *Journal of Geophysical Research*, **103(D14)**, 17149–17161, doi: 10.1029/98JD00995, 1998.

Wespes, C., Hurtmans, D., Clerbaux, C., Boynard, A., and Coheur, P.-F. 2018. Decrease in tropospheric O_3 levels in the Northern Hemisphere observed by IASI. *Atmospheric Chemistry and Physics*, **18**, 6867–6885, doi:10.5194/acp-18-6867-2018.

WHO (World Health Organisation), Ambient (outdoor) air pollution, available at: https://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health (last accessed 2nd January 2024), 2022.

Wiltshire, A.J., Burke, E.J., Chadburn, S.E., Jones, C.D., Cox, P.M., Davies-Barnard, T., Friedlingstein, P., Harper, A.B., Liddicoat, S., Sitch, S. and Zaehle, S.: JULES-CN: a coupled terrestrial carbon--nitrogen scheme (JULES vn5.1). *Geophysical Model Development*, 14 (4), 2161-2186, doi: 10.5194/gmd-14-2161-2021, 2021.Yool A, et al. 2013. MEDUSA-2.0: an intermediate complexity biogeochemical model of the marine

Formatted: Font colour: Auto

Formatted: Font colour: Auto

Formatted: Font colour: Auto

Formatted: Font colour: Auto

Formatted: Font colour: Auto

Formatted: Font colour: Auto

carbon cycle for climate change and ocean acidification studies. *Geoscientific Model Development*, **6**, 1767–1811, doi: 10.5194/gmd-6-1767-2013.

WMO, Meteorology – A three-dimensional science, World Meteorological Organisation, Bulletin 6, (Oct), 134–138, 1957.

Young, P.J., Archibald, A.T., Bowman, K.W., Lamarque, J-F., Naik, V., Stevenson, D.S., Tilmes, S., Voulgarakis, A., Wild, O., Bergmann, D., Cameron-Smith, P., Cionni, I., Collins, W.J., Dalsoren, S.B., Doherty, R.M., Eyring, V., Faluvegi, G., Horowitz, L.W., Josse, B., Lee, Y.H., MacKenzie, I.A., Nagashima, T., Plummer, D.A., Righi, M., Rumbold, S.T., Skeie, R.B., Shindell, D.T., Strode, S.A., Sudo, K., Szopa, S. and Zeng. G.: Pre-industrial to end 21st century projections of tropospheric ozone from the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP). *Atmospheric Chemistry and Physics*, 13, 2063-2090, doi: 10.5194/acp-13-2063-2013, 2013.

Ziemke, J.R., Chandra, S., Labow, G.J., Bhartia, P.K., Froidevaux, L. and Witte, J.C.: A global climatology of tropospheric and stratospheric ozone derived from Aura OMI/MLS measurements, *Atmospheric Chemistry and Physics*, 11, 9237-9251, doi: /10.5194/acp-11-9237-2011, 2011.

Formatted: Font colour: Auto

Formatted: Font colour: Auto

Formatted: Font colour: Auto

Figures & Tables:

Data Provider	Satellite Profile Products & Version	Product Link	Data Range	Data Size
RAL Space	OMI–fv214	http://www.ceda.ac.uk/	2004-2018	1442 GB
ATMOS-ULB	IAS-FORLI-v20151001	https://iasi.aeris-data.fr/catalog/	2008-2019	9.1 TB
Université de Toulouse	IASI-SOFRID vn3.5	https://iasi-sofrid.sedoo.fr/	2008-2017	3.0 TB

Table 1: *List of the satellite ozone profile data sets.*

Satellite	Quantity	Trend	Trend Lower	Trend Upper	p-value	Fit (R^2)
OMI	Trend	-0.79	-7.07	5.48	0.80	0.58
	Trend Error 1	-1.50	-7.04	4.04	0.59	0.68
	Trend Error 2	-0.09	-6.98	6.81	0.98	0.50
	Apriori Trend	-0.05	-0.21	0.11	0.56	1.00
	UKESM Trend	0.21	-0.37	0.78	0.47	0.95
	UKESM+AKs Trend	-0.57	-1.58	0.45	0.26	0.90
	UKESM Trend Forced	0.73	0.22	1.25	0.00	0.95
	UKESM+AKs Trend Forced	-0.74	-1.89	0.40	0.20	0.89
FORLI	Trend	-1.42	-2.35	-0.50	0.00	0.93
	Trend Error 1	-1.34	-2.21	-0.47	0.00	0.93
	Trend Error 2	-1.50	-2.51	-0.50	0.00	0.93

Formatted Table

	Apriori Trend	0.00	-0.11	0.12	0.94	0.67
	UKESM Trend	-0.13	-0.75	0.49	0.67	0.93
	UKESM+AKs Trend	-0.32	-0.83	0.20	0.22	0.92
	UKESM Trend Forced	0.64	-3.50	4.77	0.76	0.46
	UKESM+AKs Trend Forced	0.55	0.08	1.03	0.02	0.93
SOFRID	Trend	0.12	-0.59	0.82	0.74	0.94
	Trend Error 1	0.14	-0.59	0.88	0.70	0.90
	Trend Error 2	0.09	-0.48	0.66	0.75	0.94
	Apriori Trend	0.11	-0.17	0.39	0.43	0.98
	UKESM Trend	-0.24	-0.85	0.37	0.44	0.95
	UKESM+AKs Trend	-0.04	-0.53	0.45	0.87	0.97
	UKESM Trend Forced	0.80	0.41	1.19	0.00	0.97
	UKESM+AKs Trend Forced	0.58	0.24	0.92	0.00	0.98
OMI	Trend	-0.80	-7.29	5.69	0.80	0.71
	Trend Error 1	-1.65	-6.92	3.62	0.53	0.76
	Trend Error 2	0.05	-7.44	7.53	0.99	0.67
	Apriori Trend	-0.12	-0.26	0.03	0.10	1.00
	UKESM Trend	-0.11	-0.50	0.29	0.59	0.99
	UKESM+AKs Trend	-0.72	-1.77	0.32	0.16	0.95
	UKESM Trend Forced	0.62	0.14	1.10	0.01	0.98
	UKESM+AKs Trend Forced	0.47	-0.51	1.44	0.34	0.94
FORLI	Trend	-1.83	-2.78	-0.89	0.00	0.92
	Trend Error 1	-1.80	-2.72	-0.88	0.00	0.93
	Trend Error 2	-1.87	-2.87	-0.87	0.00	0.92
	Apriori Trend	0.09	-0.09	0.27	0.32	0.48
	UKESM Trend	-0.28	-0.77	0.20	0.25	0.98
	UKESM+AKs Trend	-0.43	-1.21	0.35	0.27	0.94
	UKESM Trend Forced	0.37	-0.05	0.79	0.08	0.98
	UKESM+AKs Trend Forced	0.28	-0.38	0.94	0.40	0.93
SOFRID	Trend	0.05	-0.91	1.01	0.92	0.93
	Trend Error 1	0.16	-0.74	1.07	0.72	0.91
	Trend Error 2	-0.07	-0.91	0.78	0.87	0.93
	Apriori Trend	0.17	-0.12	0.45	0.24	0.98
	UKESM Trend	-0.27	-0.72	0.19	0.24	0.98
	UKESM+AKs Trend	0.08	-0.33	0.49	0.69	0.98
	UKESM Trend Forced	0.46	0.09	0.84	0.01	0.99
	UKESM+AKs Trend Forced	0.10	-0.32	0.51	0.64	0.98
OMI	Trend	-0.09	-7.88	7.70	0.98	0.51
	Trend Error 1	-1.05	-6.61	4.52	0.70	0.66
	Trend Error 2	0.87	-8.24	9.98	0.85	0.38
	Apriori Trend	-0.25	-0.71	0.22	0.29	0.98
	UKESM Trend	-0.16	-0.94	0.62	0.67	0.98
	UKESM+AKs Trend	-0.62	-2.24	1.00	0.44	0.95
	UKESM Trend Forced	0.90	0.34	1.47	0.00	0.99

	UKESM+AKs Trend Forced	1.02	-0.04	2.09	0.05	0.97
FORLI	Trend	-1.52	-2.16	-0.88	0.00	0.93
	Trend Error 1	-1.42	-2.06	-0.78	0.00	0.93
	Trend Error 2	-1.62	-2.27	-0.98	0.00	0.92
	Apriori Trend	-0.03	-0.22	0.16	0.76	0.21
	UKESM Trend	-0.03	-0.62	0.56	0.93	0.98
	UKESM+AKs Trend	-0.29	-0.80	0.22	0.25	0.95
	UKESM Trend Forced	0.66	0.15	1.17	0.01	0.98
	UKESM+AKs Trend Forced	0.08	-0.44	0.61	0.75	0.93
SOFRID	Trend	-0.19	-1.01	0.63	0.65	0.96
	Trend Error 1	-0.08	-0.73	0.58	0.82	0.90
	Trend Error 2	-0.30	-1.02	0.42	0.41	0.93
	Apriori Trend	-0.15	-0.39	0.09	0.21	0.98
	UKESM Trend	-0.42	-0.97	0.13	0.12	0.99
	UKESM+AKs Trend	-0.24	-0.67	0.20	0.28	0.98
	UKESM Trend Forced	0.63	0.26	1.00	0.00	0.99
	UKESM+AKs Trend Forced	0.20	-0.20	0.61	0.31	0.98

Table 2: *$LTCO_3$ trends (DU/decade) for the satellite trend (Trend), the satellite-uncertainty trend (Trend Error 1), the satellite+uncertainty trend (Trend Error 2), the satellite apriori trend (Apriori Trend), UKESM trend (UKESM Trend), UKESM with AKs applied trend (UKESM+AKs Trend), UKESM forced trend (UKESM Trend Forced) and UKESM with AKs applied forced trend (UKESM+AKs Trend Forced). The trends from OMI, IASI-FORLI and IASI-SOFRID are for North America (red), Europe (blue) and East Asia (green). The trend lower and trend upper represent the trend 95% confidence interval. R^2 is the trend fit skill (i.e. correlation squared) and the p-value is also shown.*

Satellite	Quantity	Trend	Trend Lower	Trend Upper	p-value	Fit (R^2)
OMI – North America	Trend	-0.79	-7.07	5.48	0.80	0.58
	Trend Error 1	-1.50	-7.04	4.04	0.59	0.68
	Trend Error 2	-0.09	-6.98	6.81	0.98	0.50
	Apriori Trend	-0.05	-0.21	0.11	0.56	1.00
	UKESM Trend	0.21	-0.37	0.78	0.47	0.95
	UKESM+AKs Trend	-0.57	-1.58	0.45	0.26	0.90
	UKESM Trend Forced	0.73	0.22	1.25	0.00	0.95
	UKESM+AKs Trend Forced	-0.74	-1.89	0.40	0.20	0.89
FORLI – North America	Trend	-1.42	-2.35	-0.50	0.00	0.93
	Trend Error 1	-1.34	-2.21	-0.47	0.00	0.93
	Trend Error 2	-1.50	-2.51	-0.50	0.00	0.93
	Apriori Trend	0.00	-0.11	0.12	0.94	0.67
	UKESM Trend	-0.13	-0.75	0.49	0.67	0.93
	UKESM+AKs Trend	-0.32	-0.83	0.20	0.22	0.92
	UKESM Trend Forced	0.64	-3.50	4.77	0.76	0.46

	UKESM+AKs Trend Forced	0.55	0.08	1.03	0.02	0.93
SOFRID – North America	Trend	0.12	-0.59	0.82	0.74	0.94
	Trend Error 1	0.14	-0.59	0.88	0.70	0.90
	Trend Error 2	0.09	-0.48	0.66	0.75	0.94
	Apriori Trend	0.11	-0.17	0.39	0.43	0.98
	UKESM Trend	-0.24	-0.85	0.37	0.44	0.95
	UKESM+AKs Trend	-0.04	-0.53	0.45	0.87	0.97
	UKESM Trend Forced	0.80	0.41	1.19	0.00	0.97
	UKESM+AKs Trend Forced	0.58	0.24	0.92	0.00	0.98
OMI -Europe	Trend	-0.80	-7.29	5.69	0.80	0.71
	Trend Error 1	-1.65	-6.92	3.62	0.53	0.76
	Trend Error 2	0.05	-7.44	7.53	0.99	0.67
	Apriori Trend	-0.12	-0.26	0.03	0.10	1.00
	UKESM Trend	-0.11	-0.50	0.29	0.59	0.99
	UKESM+AKs Trend	-0.72	-1.77	0.32	0.16	0.95
	UKESM Trend Forced	0.62	0.14	1.10	0.01	0.98
	UKESM+AKs Trend Forced	0.47	-0.51	1.44	0.34	0.94
FORLI - Europe	Trend	-1.83	-2.78	-0.89	0.00	0.92
	Trend Error 1	-1.80	-2.72	-0.88	0.00	0.93
	Trend Error 2	-1.87	-2.87	-0.87	0.00	0.92
	Apriori Trend	0.09	-0.09	0.27	0.32	0.48
	UKESM Trend	-0.28	-0.77	0.20	0.25	0.98
	UKESM+AKs Trend	-0.43	-1.21	0.35	0.27	0.94
	UKESM Trend Forced	0.37	-0.05	0.79	0.08	0.98
	UKESM+AKs Trend Forced	0.28	-0.38	0.94	0.40	0.93
SOFRID - Europe	Trend	0.05	-0.91	1.01	0.92	0.93
	Trend Error 1	0.16	-0.74	1.07	0.72	0.91
	Trend Error 2	-0.07	-0.91	0.78	0.87	0.93
	Apriori Trend	0.17	-0.12	0.45	0.24	0.98
	UKESM Trend	-0.27	-0.72	0.19	0.24	0.98
	UKESM+AKs Trend	0.08	-0.33	0.49	0.69	0.98
	UKESM Trend Forced	0.46	0.09	0.84	0.01	0.99
	UKESM+AKs Trend Forced	0.10	-0.32	0.51	0.64	0.98
OMI – East Asia	Trend	-0.09	-7.88	7.70	0.98	0.51
	Trend Error 1	-1.05	-6.61	4.52	0.70	0.66
	Trend Error 2	0.87	-8.24	9.98	0.85	0.38
	Apriori Trend	-0.25	-0.71	0.22	0.29	0.98
	UKESM Trend	-0.16	-0.94	0.62	0.67	0.98

	UKESM+AKs Trend	-0.62	-2.24	1.00	0.44	0.95
	UKESM Trend Forced	0.90	0.34	1.47	0.00	0.99
	UKESM+AKs Trend Forced	1.02	-0.04	2.09	0.05	0.97
FORLI – East Asia	**Trend**	-1.52	-2.16	-0.88	0.00	0.93
	Trend Error 1	-1.42	-2.06	-0.78	0.00	0.93
	Trend Error 2	-1.62	-2.27	-0.98	0.00	0.92
	Apriori Trend	-0.03	-0.22	0.16	0.76	0.21
	UKESM Trend	-0.03	-0.62	0.56	0.93	0.98
	UKESM+AKs Trend	-0.29	-0.80	0.22	0.25	0.95
	UKESM Trend Forced	0.66	0.15	1.17	0.01	0.98
	UKESM+AKs Trend Forced	0.08	-0.44	0.61	0.75	0.93
SOFRID - East Asia	**Trend**	-0.19	-1.01	0.63	0.65	0.96
	Trend Error 1	-0.08	-0.73	0.58	0.82	0.90
	Trend Error 2	-0.30	-1.02	0.42	0.41	0.93
	Apriori Trend	-0.15	-0.39	0.09	0.21	0.98
	UKESM Trend	-0.42	-0.97	0.13	0.12	0.99
	UKESM+AKs Trend	-0.24	-0.67	0.20	0.28	0.98
	UKESM Trend Forced	0.63	0.26	1.00	0.00	0.99
	UKESM+AKs Trend Forced	0.20	-0.20	0.61	0.31	0.98

Table 2: *LTCO$_3$ trends (DU/decade) for the satellite trend (Trend), the satellite-uncertainty trend (Trend Error 1), the satellite+uncertainty trend (Trend Error 2), the satellite apriori trend (Apriori Trend), UKESM trend (UKESM Trend), UKESM with AKs applied trend (UKESM+AKs Trend), UKESM forced trend (UKESM Trend Forced) and UKESM with AKs applied forced trend (UKESM+AKs Trend Forced). The "trend lower" and "trend upper" represent the trend 95% confidence interval based on the trend precision calculated from **Equation 3**. R^2 is the trend fit skill (i.e. correlation squared) and the p-value is also shown.*

Formatted: Font colour: Auto

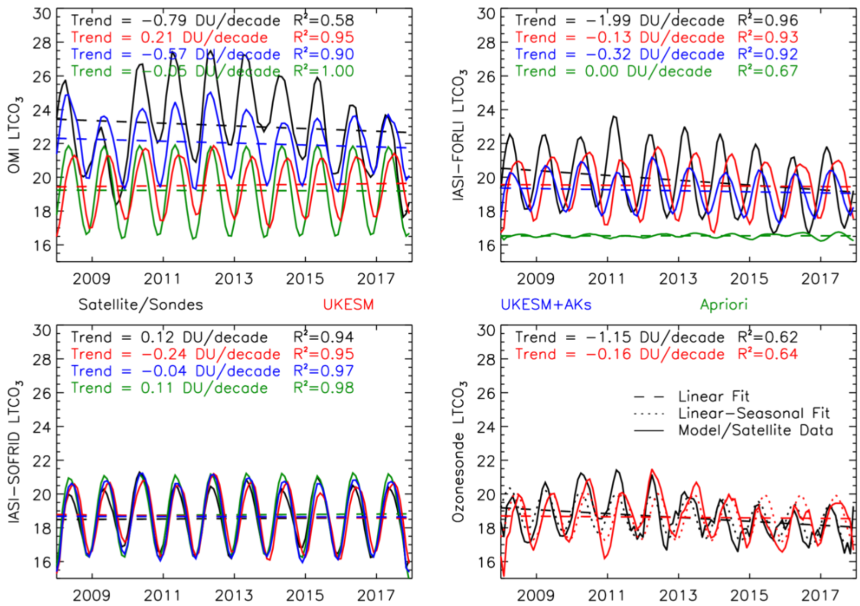

Figure 1: *Lower tropospheric column ozone (LTCO$_3$, surface to 450 hPa, DU) regional time-series for North America, based on the HTAP land mask, from OMI (top-left), IASI-FORLI (top-right), IASI-SOFRID (bottom-left) and ozonesondes (bottom-right) are shown by the black lines in the respective panels. UKESM simulations without and with satellite averaging kernels (AKs) applied are shown in red and blue lines. Green lines show the satellite apriori. Dashed lines show the LTCO$_3$ linear trend which are labelled in the top of each panel. The R^2 squared values show the linear-seasonal trend model fit to the corresponding LTCO$_3$ time-series (i.e. correlation squared).*

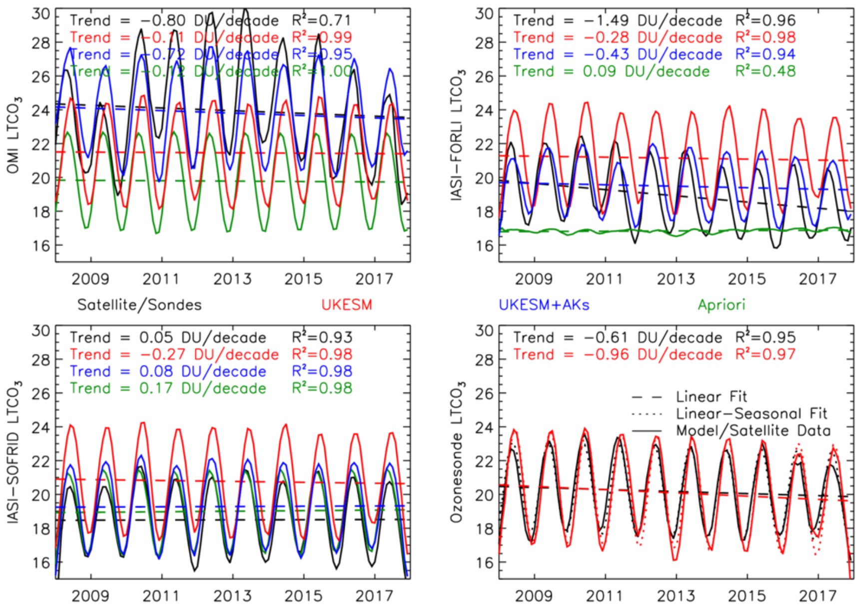

Figure 2: *$LTCO_3$ (DU) regional time-series for Europe, based on the HTAP land mask, from OMI (top-left), IASI-FORLI (top-right), IASI-SOFRID (bottom-left) and ozonesondes (bottom-right) are shown by the black lines in the respective panels.. UKESM simulations without and with satellite AKs applied are shown in red and blue lines. Green lines show the satellite apriori. Dashed lines show the $LTCO_3$ linear trend which are labelled in the top of each. The R^2 squared values show the linear-seasonal trend model fit to the corresponding $LTCO_3$ time-series (i.e. correlation squared).*

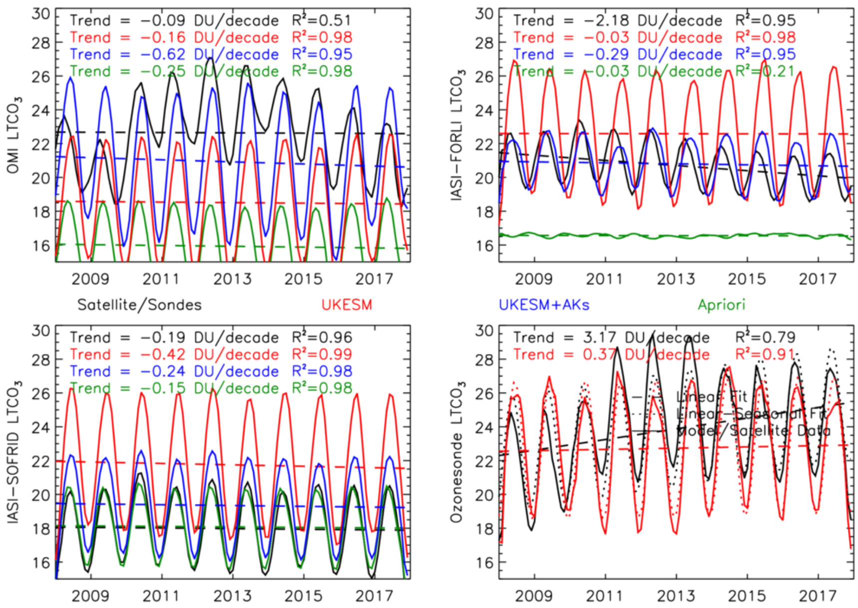

Figure 3: *$LTCO_3$ (DU) regional time-series for East Asia, based on the HTAP land mask, from OMI (top-left), IASI-FORLI (top-right), IASI-SOFRID (bottom-left) and ozonesondes (bottom-right) are shown by the black lines in the respective panels.. UKESM simulations without and with satellite AKs applied are shown in red and blue lines. Green lines show the satellite apriori. Dashed lines show the $LTCO_3$ linear trend which are labelled in the top of each panel. The R^2 squared values show the linear-seasonal trend model fit to the corresponding $LTCO_3$ time-series (i.e. correlation squared).*

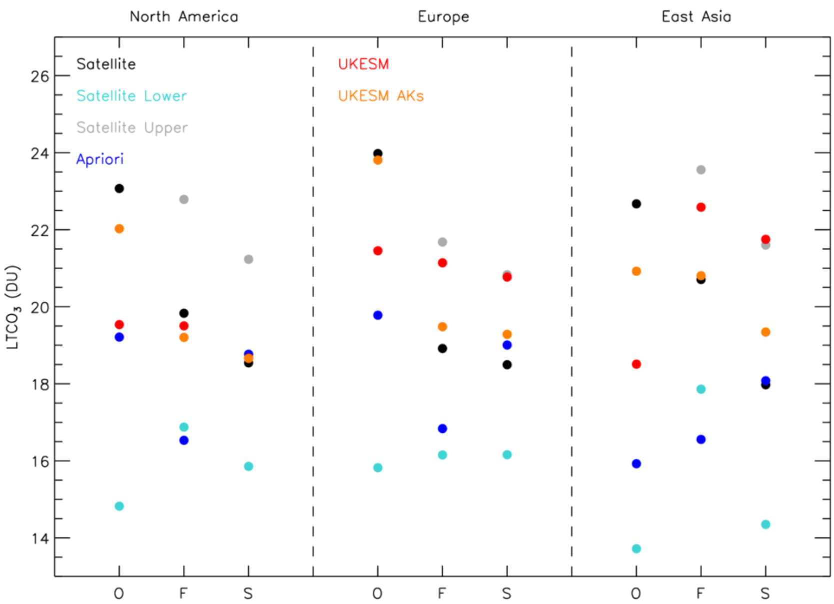

Figure 4: *Average $LTCO_3$ (DU) values across the 2008-2017 time-period for the satellite (black), satellite-lower (cyan), satellite-upper (grey), apriori (blue), UKESM (red) and UKESM+AKs (orange). The satellite-lower and satellite-upper values are the average of the satellite ± its error term time-series (note: these values do not always fit in the y-axis range). O, F and S represent OMI, IASI-FORLI and IASI-SOFRID for North America (left), Europe (centre) and East Asia (right).*

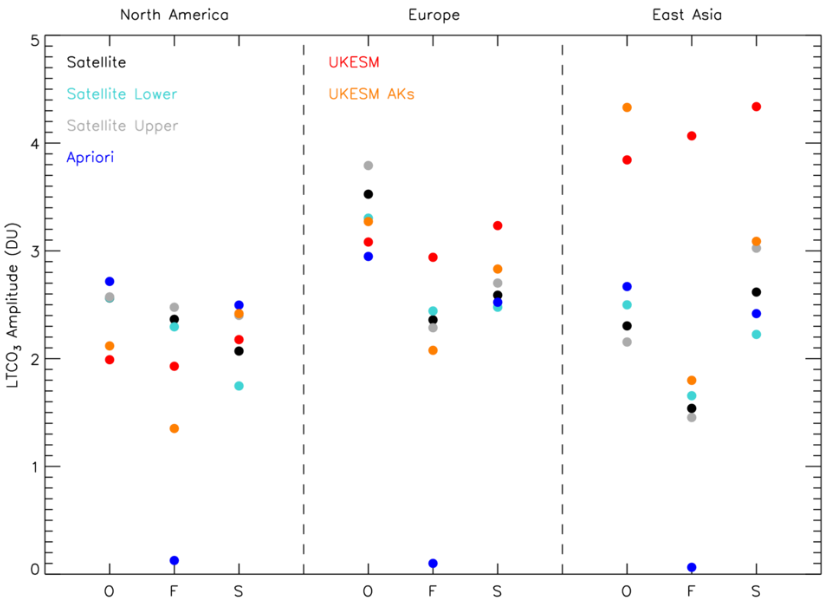

Figure 5: *Average $LTCO_3$ seasonal cycle amplitude (DU) values across the 2008-2017 time-period for the satellite (black), satellite-lower (cyan), satellite-upper (grey), apriori (blue), UKESM (red) and UKESM+AKs (orange). The satellite-lower and satellite-upper values are the average of the satellite ± its error term time-series (note: these values do not always fit in the y-axis range). O, F and S represent OMI, IASI-FORLI and IASI-SOFRID for North America (left), Europe (centre) and East Asia (right).*

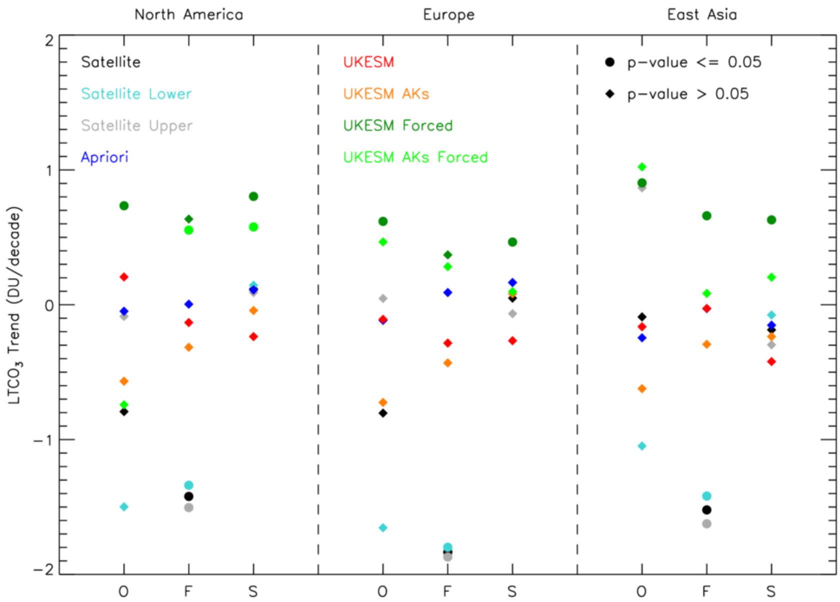

Figure 6: *Average LTCO$_3$ linear trends (DU/decade) values across the 2008-2017 time-period for the satellite (black), satellite-lower (cyan), satellite-upper (grey), apriori (blue), UKESM (red), UKESM+AKs (orange), UKESM forced (dark green) and UKESM+AKs forced (light green). The satellite-lower and satellite-upper values are the average of the satellite ± its error term time-series (note: these values do not always fit in the y-axis range). O, F and S represent OMI, IASI-FORLI and IASI-SOFRID for North America (left), Europe (centre) and East Asia (right). Triangle and circular symbols represent linear trends with p-values > 0.05 or p <= 0.05, respectively.*