Urban Ozone Trends in Europe and the USA (2000-2021)

Beth S. Nelson[1,2,*] and Will S. Drysdale[1,2,*]

[1]Wolfson Atmospheric Chemistry Laboratories, Department of Chemistry, University of York, Heslington, York, YO10 5DD, UK
[2]National Centre for Atmospheric Science, University of York, York, UK
[*]These authors contributed equally to this work.

Correspondence: Beth S. Nelson (beth.nelson@york.ac.uk) and Will S. Drysdale (will.drysdale@york.ac.uk)

Abstract.

Trends in urban $MDA8O_3$, alongside annual O_3 episodic and exposure metrics, across Europe and the United States of America were explored between 2000-2021. Using surface monitoring site data from the TOAR-II and European Environment Agency databases, piecewise quantile regression (PQR) analysis was performed on 353 $MDA8O_3$ time series (205 European, 149 USA). The PQR analysis permitted 2 break points over period to balance the intent to describe changes over a large time period, while still capturing the abrupt changes that can occur in urban atmospheres. We found that there were many sites across Europe with high certainty increasing trends in the 5^{th} and 50^{th} quantiles, whereas the majority of high certainty trends in the 95^{th} quantile were found to be decreasing. A similar pattern was observed across the USA, with 5^{th} quantile trends increasing and 95^{th} quantile trends decreasing, though a small but increasing number of sites showed a return to increasing trend at τ = 95. To group trends, hierarchical clustering with dynamic time warping was employed and these groups used to guide analysis. Clustering was typically regional across Europe and the USA, and increasing trends were identified across Southern and the Central alpine regions of Europe, and in California and the Intermountain West of the USA. Recent high certainty increasing trends in $MDA8O_3$ in the Intermountain West may be related to warmer summers and increased wildfire events in the region, highlighting the need to monitor changing trends in $MDA8O_3$ with climate change, to assess human exposure risk to elevated levels of O_3.

1 Introduction

Tropospheric ozone (O_3) is a greenhouse gas and air pollutant harmful to human health, and plant growth (Fleming et al., 2018; Mills et al., 2018; Szopa et al., 2021). It is a secondary air pollutant, formed from the photochemical reactions of primary pollutants NO_x (NO + NO_2) and volatile organic compounds (VOCs). The chemistry of O_3 formation is non-linear and the effect of changing precursor concentrations can cause O_3 concentrations to increase or decrease depending on the photochemical regime found at a given location (Sillman et al., 1990; Sillman, 2002). Despite global successes in reducing primary pollutant emissions over the past few decades, global exposure to O_3 has been increasing throughout the 21st century. This is particularly observed in urban areas, where the vast majority of the global population live, projected to increase to 68 % in 2050 from 55 % in 2018 (UN, 2019). In a study of 12946 cities located worldwide, the average mean weighted O_3

concentration increased by 11 % between 2000 and 2019, and the number of cities exceeding the 2021 WHO peak season O_3 standard (60 μg m^{-3}) increased from 89 % to 96 % (Malashock et al., 2022).

Due to the complexity of O_3 production, its trend direction, magnitude, and significance varies by location. The Tropospheric Ozone Assessment Report Phase I (TOAR I) was the first global review of trends in surface ozone, covering 1970 - 2014 (Fleming et al., 2018; Gaudel et al., 2018). As part of the TOAR-I review, trends in two O_3 metrics were calculated globally over the period of 2000 - 2014: 4th highest daily maximum 8-hour O_3 (4MDA8), and the number of days with $MDA8O_3 > 70$ ppb O_3 (NDGT70). The study used data from 4801 global monitoring sites over this time period. For both of these metrics, downward trends were observed for most of the USA, and some sites in Europe. However, whilst trends were decreasing, over the period of 2010-2014 (2,600 sites utilised), some of these sites were located in regions that had the highest 4MDA8 and NDGT70 values, particularly California and Southern Europe. These corresponded with the regions with the highest O_3 precursor emissions illustrating that downward trends had not yet reduced these peak metrics.

Since the 1990s, a general downward trend in urban O_3 pollution has been observed in the United States (He et al., 2020). This reducing trend has been linked to stricter limiting regulations on the emissions of primary pollutants such as NO_x and VOCs. Although NO_x and VOC emissions in Europe have also been declining since the late 1980s, the trend in O_3 is less clear due to large inner-annual variation, driven by climate variability and the dispersion and transport of pollutants from other regions (Jonson et al., 2006; Yan et al., 2018). Between 1995 and 2014, negative trends in the highest O_3 levels across urban sites in Europe were identified due to pollutant emission restrictions across Europe. However, increasing background levels, particularly in northern and eastern Europe, make it difficult to identify strong trends in urban O_3 when transboundary effects are considered (Yan et al., 2018). Despite this, a study of 93 suburban and urban sites across Europe identified notable enhancements in O_3 seasonal and annual means between 1995 - 2012, even with the continuous downward trend in anthropogenic emissions across the continent (Yan et al., 2018).

In this study, we aim to investigate a longer time-series of 21st century trends (2000 - 2021). Here, we focus on the trends in urban O_3, using urban monitoring site data from Europe and the United States of America across the 22 year period. We employ quantile regression and change point detection to construct trends that capture the broad structure of a complex time series, while remaining explainable with concise statistics. To gain a clearer understanding of the group behaviour of sites, we use hierarchical clustering (HC) with dynamic time warping (DTW), to group together 22 year time series with comparable time-series structures within Europe and the United States of America independently. This is then used to investigate whether trends in MDA8, and both episodic and exposure-relevant O_3 metrics, are varying regionally. Our analysis focuses comparisons between warm-season and cold-season $MDA8O_3$ values, as well as probing different changes across quantiles, to allow for an assessment of whether low or high O_3 distributions have different behaviours, across Europe and the USA, and between hierarchical clusters.

2 Methodology

2.1 Data Preparation

Hourly O_3 data were obtained for urban sites in the USA and Europe using the TOAR-II (Schröder et al., 2021) and European Environment Agency (EEA) (European Environment Agency, a, b) databases. Those obtained from the EEA are categorised as urban traffic, urban background and urban industrial, while those from the TOAR database are categorised as urban and suburban. A full list of sites can be found in the supplementary information (table S1). TOAR-II data was retrieved between 2000-01-01 and 2021-12-31 (the latest available) and EEA data between 2000-01-01 and 2023-12-31, the latter being extended allowing break points to be fit around the changes due to the COVID-19 pandemic, though we remain focused on 2001-2021 in this study. Time series were quality controlled via threshold, persistence and variance checks similar to those by Wang et al. (2023): O_3 concentrations were required be between 0 and 500 ppbv, periods where an identical value was repeated 8 or more times in a row were removed as were days where the difference between the minimum and maximum concentration was ≤ 2 ppbv.

After quality control, time series were required to have > 80 % data coverage between 2000-01-01 and 2021-12-31 and for the calculation of annual metrics each individual year was retained if it had more than 60 % coverage. Additionally, a small number of time series were removed following inspection due to large changes in the mean of data before and after a missing period, indicating possible sensor issues - these are listed in table S2. This resulted in 353 O_3 time series (204 Europe, 149 USA).

2.1.1 Metrics

In addition to the series described above, we calculated six common health related O_3 metrics which are explored in section 3.4 and used in conjunction with calculated trends in 3.5. More information on these metrics can be found in Lefohn et al. (2018) and Fleming et al. (2018).

- 4MDA8 - The 4th highest MDA8 value in the warm season (ppbv)
- NDGT70 - Number of days in a year where MDA8 > 70 ppbv (days)
- SOMO35 - Subtract 35 ppb from the MDA8 value and sum all postive values (ppbv days)
- 3MMDA1 - Annual maximum value of 3 month rolling mean of MDA1 (maximum daily 1 hour O_3) (ppbv)
- 6MMDA1 - As 3MMDA1 with a rolling 6 month mean (ppbv)
- AVGMDA8 - Mean of the warm season MDA8 (ppbv)

2.2 Trend Analysis

Trends were calculated on the maximum daily 8-hour average O_3 concentrations ($MDA8O_3$), which is calculated by taking a rolling 8-hour mean, then selecting the highest value per day between 0700 - 2300 L. The 8 hour rolling average is required

to have 6 of the 8 hours present to be valid. Additionally, subgroups of warm (April - September) and cold (October - March) season were created. It is common to remove the seasonal component of an ozone time series to improve the accuracy of trends (Cooper et al., 2020) and here we subtract monthly mean climatologies for each time series resulting ozone anomalies.

These were calculated following the methodology in, and using code provided by, Chang et al. (2023). QR calculates a linear model that seeks to minimise the residuals with a defined proportion (τ) of the points above and below the fit line. For example, the scenario τ = 0.5 splits the data 50:50 above and below the line and can be considered analogous to a "median" trend line. QR has the advantage of being insensitive to outliers and the which is desirable for the longer-term trends being investigated here. The 1-sigma uncertainly for the QRs were calculated via a moving block bootstrapping method, where the data are subdivided into overlapping blocks that are $n^{1/4}$ points wide (or ~20 points for 20 years of hourly data). These blocks are shuffled to generate a new time series, and replicated 1000 times, and the standard deviation of these replicates calculated. This uncertainty was subsequently used in the determination of the p-value, providing a metric for the significance of the trends.

Quantile regressions were calculated piecewise with 0 - 2 breakpoints to allow the trends to capture large non-linearities. As urban concentration trends can change sharply (e.g. with policy intervention), this gives the model the freedom to represent these. Limiting the model to a maximum of two breakpoints strikes a balance between capturing sufficiently large-scale changes while still being able to describe a ~20 year time series with a small number of coefficients.

To determine what break points to use on each time series, a range of candidate models were constructed, each with zero, one or two break points. These break points were restricted to be unable to occur within the first or last two years of the time series, nor could they occur within 5 years of each other. They were arbitrarily set to occur on the 1st January in a given year, which was deemed an appropriate degree of freedom given the 20 year span of the time series and locating the break points sub annually does not have a large effect on the resulting trends. For a time series that fully spanned the range of 2000-01-01 - 2021-12-31, 110 models would be created - one with zero break points, 18 with one break point, and 91 with two. All regressions were calculated at τ = 0.05, 0.10, 0.25, 0.50, 0.75, 0.95 and 0.95.

To select 'change points' from the array of break points available, the model performance was evaluated via Akaike information criterion (AIC). AIC was chosen as the evaluation criteria oweing to its penalty term penalising models with more break points that do not appreciably improve the model fit over the simpler case, The model with the minimum AIC was selected from the candidate pool as the best fit for a time series at a given τ.

Trends have been collated into significance categories: $p \leq 0.05$ (high certainty), $0.05 < p \leq 0.10$ (medium certainty), $0.10 < p \leq 0.33$ (low certainty) and $p > 0.33$ (very low certainty or no evidence) based off of the guidance of Chang et al. (2023). For the most part slopes where the p-value is > 0.33 are treated as 'no trend' regardless of their magnitude (generally we observe that as the magnitude of the trend decreases so does its significance), though sometimes are given with a direction when required by a visualisation.

2.3 Clustering

To aid analysis it was desirable to group similar time series, clusters were calculated using hierarchical clustering (HC) with dynamic time warping (DTW) used as a distance measure (Aghabozorgi et al., 2015). DTW provides a distance measure between two time series that allows for some deviation in the relation of features in time and one-to-many mapping of features, as opposed to use of e.g. the euclidean distance where mapping between series is one-to-one and cannot undergo deformation (Berndt and Clifford, 1994; Mueen and Keogh, 2016). We adopt a similar method to Reed et al. (2025) using the R package *dtwclust* (Sarda-Espinosa, 2024) for both DTW and HC steps, and used to calculate cluster validity indices (CVIs). Clustering was performed by region and time series type (e.g sites in Europe for the MDA8-O3 warm season). Time series were normalised and then, to provide clusters related to the range of τ values used in the QRs, aggregated monthly by calculating the τ^{th} quantile per month. Metrics were not aggregated as they are all annual values. Clustering was calculated with 5 - 75 clusters, and then evaluated with 6 internal CVIs (Silhouette, Dunn, Calinski-Harabasz, COP, Davies-Bouldin and Modified Davies-Bouldin) where the latter 3 were inverted such that all can be maximised. The final selection for the number of clusters was determined as the mean of the optimal number of clusters suggested by each of the CVIs.

3 Results and Discussion

As an initial overview, figure 1 shows the annual mean concentrations across the range of all the urban sites in this study. Additional daily mean aggregations are shown to produce day (0800 - 1900 L) and night (2000 - 0700 L) subgroups. European O_3 concentrations show a steady increase over the last 20 years, with the average across all data increasing from 21.2 ppbv in 2000 to 25.9 ppbv in 2019, the $MDA8O_3$ increasing from 30.8 to 35.3 ppbv and warm season $MDA8O_3$. In the USA average O_3 has increased from 27.7 to 29.7 ppbv, but $MDA8O_3$ and warm season $MDA8O_3$ both decreased, from 41.5 to 40.9 and 49.9 to 45.1 ppbv respectively. Annual average urban NO_2 is also shown to provide some context to precursor concentrations - these sites were selected under the same critera for site type and coverage as the O_3 data. Both Europe and the USA have seen reductions in NO_2, 19.0 in 2000 to 13.0 ppbv in 2019 in Europe and 19.7 to 9.8 ppbv in the USA. The USA did see a NO_2 minima in 2020 of 9.3 ppbv, but increased to 11.8 ppbv in 2022, whereas Europe continued its decrease to 10.6 ppbv in 2022.

The other features that do stand out appear in the sub groups containing the warm season, but are most clear in the warm $MDA8O_3$ series correspond to years with significant heatwaves - 2003, 2006, 2015 and 2018 in Europe and 2012 in the USA. Heat waves are historically linked to enhanced O_3 concentrations, and are expected to increase in frequency (Schär et al., 2004; Russo et al., 2015; Shen et al., 2016; Otero et al., 2016; Gouldsbrough et al., 2022). The continued impact of heat waves on high O_3 is however, dependant on the future levels of O_3 precursors and reductions and indeed this can be observed at some sites (Meehl et al., 2018; Otero et al., 2021; Li et al., 2025; Chang et al., 2025), but urban sites may take longer to reach levels where O_3 - temperature sensitivity is reduced, owing to a higher initial concentrations (Vazquez Santiago et al., 2024).

While QRs were calculated across a wider range of τ values, this discussion mostly focuses on τ = 0.05, 0.5 and 0.95 as the omitted values did not sufficiently expand the results presented.

3.1 Change Point Assignment

Regional differences were observed in the types (QR, PQR_1, PQR_2) of change point assigned. Figure 2 shows that $MDA8O_3$ series in Europe were rarely described by QR (1 - 3 % of series), whereas the United States of America was described by it 30 - 62 % of the time. As the European time series extend to 2023, some of these additional change points are likely due to the COVID pandemic, and indeed examining the location of change points by year (figures S1 and S2) shows an increased number of change points in 2019 - 2021 in Europe - this is also reflected in the (fewer) change points seen in the USA. This also highlights that there are more change points focused at the beginning and end of the series generally (for both O_3 and NO_2) which may indicate some edge effects from the selection process. However, we don't expect this to have impacted our results as the method demonstrates the ability to choose regressions with zero change points, we do not observe large swings in trends overall at these times and that there is a good amount of other change points assigned within the time period. We use this information to inform the selection of "snapshot" years presented in this manuscript, avoiding years close the end beginning or end of the series

3.2 Trends in $MDA8O_3$ Across Quantiles

The number of sites with increasing and decreasing trends for a selection of quantiles (5^{th}, 50^{th} and 95^{th}), annually, as well as the degree of certainty of the trends, in the $MDA8O_3$, $MDA8O_3$ warm season and $MDA8O_3$ cold season data is shown in Figure 3. We use p-values to describe the certainty of each trend, as defined by Chang et al. (2023). Generally, there are a large proportion of high certainty positive trends in the 5^{th} and 50^{th} quantile in the Europe data set (range of 34 - 72 % of trends across the 20-year period). In contrast there are fewer trends of high certainty in the 95^{th}, the majority of which are decreasing (between 5 - 25 % across the 20-year period). The warm-season data shows a similar pattern, but with fewer high certainty increasing trends in the 5^{th} and 95^{th} quantile. Many more high certainty increasing trends are observed in the cold season only data, particularly in the 5^{th} and 50^{th} quantile data (range of 29 - 73 % across the 20-year period). The proportion of high certainty increasing trends in the cold season appears to be increasing with time.

In the US $MDA8O_3$ trend data, the majority of 5^{th} quantile trends are high certainty increasing trends (35 - 53 %) and the 95^{th} quantile is dominated by high certainty negative trends (46 - 76 %), with a mixture found in the 50^{th} quantile. This is consistent with the findings of Simon et al. (2015) who observed increasing O_3 trends across urban sites in the US (1998 - 2013) at the lower end of the O_3 distribution, and decreasing trends at the upper end, leading to an observed compression of the O_3 range. In the warm-season only data, there is a larger contribution of decreasing trends across all quantiles. This is particularly true in the 95^{th} quantile data, where the vast majority of trends are high certainty decreasing trends (55 - 84 %). However, there is a notable change year-on-year in the US 95^{th} warm-season dataset that is worth some discussion. Although decreasing trends dominate, and the number of high certainty increasing trends are low (1-11 %), we observed a step-change. Between 2000 - 2008, between 1-3 % of trends are high certainty increasing trends. However, after 2008, the number of sites showing a high certainty increasing trend increases rapidly from 2 to 16 by 2016, with the majority of these sites located in Western and Central US, and the top three largest magnitude slopes (medium-large), located in the South. In the cold season data, there is very little year-on-year change in the distribution of increasing and decreasing trends in the 5^{th}, 50^{th} and 95^{th} quantile.

While certainty and magnitude tend to follow one another, to probe the trends more specifically we define a threshold of "medium magnitude" of 0.5 ppbv yr^{-1}, and explore the moderate or higher certainty trends with and without this threshold. Figure S5 shows full distribution of trends in the groups discussed. A summary of the proportion increasing and decreasing trends, highlighting the proportion of moderate or higher certainty and medium or higher magnitude trends in 2004 and 2018 is presented in Figure 4.

Across the Europe data set at the start of the century (2004), 85 % and 78 % of slopes showed an increasing trend in the 5^{th} and 50^{th} quantiles respectively (73 % and 57 % of total at moderate certainty or higher respectively), compared to 74 % showing a decreasing trend in the 95 % quantile (31 % of total at moderate certainty or higher). By 2018, the number of increasing slopes increased across the three quantiles, to 90 % and 86 % for the 5^{th} and 50^{th} quantile (74 % and 69 % of total at moderate certainty or higher), and from 26 % to 39 % in the 95^{th} quantile (from 7 % to 15 % of total at moderate certainty or higher). This supports an increase in the number of Europe sites with increasing O_3 trends at the end of the two-decadal

period compared to the beginning, for which the majority of sites showed increasing trends in the low-mid quantiles during both time periods. In the 95th quantile, the majority of sites were are showing decreasing trends in both time periods, but a larger proportion of increasing trends is apparent in 2018 compared to 2004.

In comparison, 82 % and 49 % of slopes showed an increasing trend in the 5th and 50th quantiles of the US data set at the start of the century (56 % and 36 % of total at moderate certainty or higher). Similarly to the Europe, 95th quantile trends were decreasing the in US (87 % - 68 % of total at moderate certainty or higher). However, in contrast to Europe, by 2018 fewer trends were increasing, down to 68 % in the 5th quantile (40 % of total at moderate certainty or higher), from 82 % in 2004 (56 % of total at moderate certainty or higher). In the 50th quantile, more trends are decreasing (51 % - 30 % of total at moderate certainty or higher) than increasing (49 % - 36 % of total at moderate certainty or higher), although more trends are increasing than decreasing when the data is filtered for moderate certainty or higher. In the 95th quantile, there was an increase in the number of sites with increasing $MDA8O_3$ in 2018 (26 %, 12.1 % of total at moderate or higher certainty) compared to 2004 (14 %, 3.9 % of total at moderate or higher certainty), and, whilst the majority of sites were still decreasing trends, a decrease in the number of these from 87 % (68 % of total at moderate or higher certainty) to 74 % (60 % of total at moderate or higher certainty) between 2004 and 2018 was observed. To summarise, between 2004 and 2018 a reduction in the number of increasing trends was observed at the 5th and 50th quantiles, but the high extremes (95th quantile) saw an increase in the number of increasing trends, coupled with a decrease in the number of decreasing trends.

Separating the beginning (2004) and end (2018) of time series $MDA8O_3$ trends into warm (April - September) and cold (October - March) seasons reveals much clearer trends, and reveals a regionality in the direction of $MDA8O_3$ trends. Generally across Europe early century (2004) data, there more increasing trends in the cold season (72 %; 37 % of total at of moderate or higher certainty) compared to the warm season (25 %; 7 % of total at moderate or higher certainty) in the 95th quantile. Although there are a high number of increasing trends in the warm and cold season for the 5th quantile (88 % and 78 % respectively), there are many more increasing trends with a moderate or higher certainty and medium or greater magnitude in the warm season (15 %) compared to the cold season (4 %). This pattern is retained in the 2018 slope data, with an increase to 19 % of 5th quantile trends increasing at moderate or higher certainty and medium or higher magnitude in the warm season. A smaller increase in the cold-season 5th quantile was observed, and a larger increase in the 50th quantile (from 7 % to 17 % of total at moderate or higher certainty and medium or higher magnitude). In summary, although we are generally seeing increasing trends across the quantiles in both 2004 and 2018, this is particularly prevalent for the lower extremes in the warmer months, and the 50th quantile cold-season data, where the trends are not only increasing, but are of higher certainty and magnitude.

In the US cold season at the start of the century, 92 %, 83 % and 45 % of trends are increasing in the 5th, 50th, and 95th quantiles respectively, the majority of which are of moderate or high certainty in the 5th and 50th quantile. In comparison, there are fewer increasing trends in the warm season for the 5th (52 %), 50th (17 %) and 95th (9 %) quantiles. In the warm season, the majority of trends are decreasing in the 95th quantile (90 %), with 63 % of all trends in this quantile having a moderate or higher degree of certainty, at a medium or larger magnitude. The majority of 50th quantile trends are also decreasing (82 %; 70 % at moderate or higher certainty), but with a smaller proportion with both moderate or higher certainty and medium or higher magnitude (14 %). Comparable trends are observed in the 2018 slopes, but with a small increase in the number of increasing

trends of moderate certainty or higher in the warm season 50th (from 12 % to 16 %) and 95th (from 3 % to 10 %) quantile, and in the 95th quantile cold season (from 23 % to 32 %). In the cold season, the number of increasing trends with a moderate or high certainty and also medium or higher magnitudes in the 95th quantile has increased from 5 % to 14 % of the trends.

3.3 Hierarchical Clustering of MDA8O_3 time series

So far, trends have only be separated by region, but it is clear when displaying the trends on a map there is more structure. Figures 5 and 6 show this for the warm season MDA8O_3 trends (those for MDA8O_3 and cold season MDA8O_3 are shown in figures S6 - S9). From these there appear difference in the north and south of Europe, and the coasts versus central United States. However, it is not trivial to chose sub-regions through inspection alone, so dynamic time warping coupled with hierarchical clustering (DTW, HC, described in section 2.3) was implemented to identify time series with similar features. Here, we discuss the clustering observed in the MDA8O_3 data at the 5th, 50th and 95th quantile. HC was performed for each of the three types of MDA8O_3 data set, separately for the United States of America and Europe.

The HC of the European data set indicates that most sites are well described by one cluster only, in the 5th and 50th quantiles across all three types of MDA8O_3 data aggregation. However, a few additional clusters are introduced in the 5th and 50th quantile cold-season data, as well as the 50th quantile full data set and warm-season only data set. In the cold-season, these clusters don't appear to have any particular regionality in the 5th quantile. However, in the 50th quantile two new clusters are introduced, localised around South Central and East Europe (clusters 3 and 4). In the 95th quantile (high extreme) data, more regional clustering is observed in the full MDA8O_3 data, and the warm-season only data. This regional clustering is more clearly observed in the MDA8O_3 data, with clusters location in North Eastern Europe (cluster 1); across the continent from South and Central Western France to East Europe (cluster 2); some selected sites from across this same area, but also including sites in Spain and the UK (cluster 3); Northern France (cluster 4), the Far East (cluster 5), and North East (cluster 6). In the warm season, the majority of clustered sites are located across Europe, with some exceptions in the North East (cluster 2), North West (cluster 3), and Spain (cluster 4).

For all clusters in the 95th quantile MDA8O_3 European data, an enhancement in the mean AVGMDA8, 4MDA8, NDGT70, SOMO35, 3MMDA1 and 6MMDA1 metrics is observed in 2003, when an extensive European heatwave occurred. Enhancement in other significant heatwave years are also visible in some data clusters in 2006 (clusters 1,2 and 5); 2015 (clusters 1 and 2); and 2018 (clusters 1,2, and 6). This relates to the features seen in the (warm season) MDA8O3$_3$ averages in figure 1, the 2003 heat wave is notably larger in the average presented there. Later enhancements may be less clear due to those heat waves being less wide spread, or reductions in O_3 - temperature sensitivities in those locations. This is indicated by the heat waves not being present in all clusters as the decades progress, but from this study it is not possible to say definitively whether one or both effects are causing this.

The HC of the United States of America data set has only one large cluster for the full MDA8O_3 data set, cold-season, and warm-season only data in the 5th quantile. However, some clear regional clusters appear in the 50th and 95th quantile data sets across all three data types. The main secondary cluster sites are located in Florida, also including some sites along the Gulf of Mexico and the South in the cold-season (cluster 2, or cluster 3 in the MDA8O_3 full 95th quantile data). Some clustering

around the Intermountain West and the South West is observed in the full $MDA8O_3$ data set in the high extremes (cluster 2, 95^{th} quantile). The East Coast generally falls into the largest cluster which doesn't appear to have a strong regonality, except for the in full $MDA8O_3$ data 95^{th} quantile, where there is an independent North East coast cluster (cluster 5). Also in this data type, and in the 95^{th} quantile data, there is a cluster located in South Central USA (cluster 1).

For the US 95^{th} $MDA8O_3$ data, clusters behave much less uniformly than they do in Europe. Whilst all $MDA8O_3$ metrics are generally decreasing across all clusters (clusters 1,3,4,5 and 6). However, many metrics are either clearly elevated, increasing across the 20-year period, or both, for cluster 7, located in and around Los Angles, in Southern California. This is also observed to a lesser extent in cluster 2, located more generally in the West and in the Intermountain West region of the United States. This affect can be seen by looking at the trend data. For cluster 7, increasing trends are observed in the full $MDA8O_3$ data set. However, the $MDA8O_3$ trends in cluster 7 are generally of low certainty. In cluster 2, mean $MDA8O_3$ is elevated compared to other clusters, excluding cluster 7, and the slope in $MDA8O_3$ trends is generally quite small. However, these trends typically have a high degree of certainty ($p < 0.05$), which are typically negative in the first half of the vicennial, and positive in the second.

3.4 Spatio-temporal distribution of O_3 metrics, relevant to health and exposure

As the health/exposure O_3 metrics are annually aggregated, the DTW/HC method was not directly applied to these data sets, but observations from inspection of figures S10 - S21 is presented here.

The NDGT70 and 4MDA8 metrics are generally used to investigate variations in the highest values of O_3. High values of these metrics are typically an indicator of episodes of high photochemical O_3 production. Here, we explore the frequency of high values of these metrics over the 20-year time period, separately in Europe and the United States of America.

Across Europe, patterns in high 4MDA8 (>= 85 ppb) and NDGT70 (>= 25 days) frequencies and site locations are similar. In general, the only sites that consistently exceed these metrics are located in Southern Europe. Exceptions to this can be seen in 2003, 2006 and 2018 for both metrics, known heatwave years, when a large number of high 4MDA8 and NDGT70 values occurs across the continent. However, the number of sites exceeding the prescribed limits during these heatwave events generally reduces with time, with 89, 45, 27, and 15 exceedances in 4MDA8 and 68, 19, 6, and 8 exceedances in NDGT70 for 2003, 2006, 2015, and 2018 respectively.

For the US metrics, a high number of sites across the country are found to have high 4MDA8 and NDGT70 values in the early years of the 20-year decade, from 2000 up until 2007. From 2008, there are few sites with exceedances, generally located in the West, and occasionally around the Intermountain West region of the US. Due to a lack of data availability in 2012, a reported heatwave year, it is not possible for us to comment here on whether this event led to elevated 4MDA8 and NDGT70 frequencies.

The SOMO35, AVGMDA8, 3MMDA1, and 6MMDA1 O_3 metrics are used to describe the general O_3 exposure, not just high O_3 events, important for assessing O_3 health burdens, rather than policy compliance. Across Europe, and across the 20-year period, values of SOMO35 are generally fairly low across the board (typically < 4000 ppb day). SOMO35 values of > 4000 ppb day are more widespread across Europe, excluding Northern Europe, in 2003; and at a few sites across Southern

Europe over the years, typically in Spain, Greece, Switzerland and Italy. This is reflected in the AVGMDA8 metric, where the lowest values are typically found in the UK, Scandinavia, and North of mainland Europe (typically < 40 ppb). Slightly higher values of AVGMDA8 are observed in Central and Eastern Europe (41 - 45 ppb) in the early years (2000 - 2014). However, an uptick in AVGMDA8 is observed in Eastern Europe in multiple years from 2015 onward, with many sites observing levels of 46-50 ppb, or even higher in 2018 (51 - 55 ppb). Few sites exceed 56 ppb, and those that do are located in Southern Europe. The exception to this is once more in 2003, when AVGMDA8 were typically > 56 ppb across Europe, reaching 72 ppb in the alpine region of Central Europe. Both the 3MMDA1 and 6MMDA1 metrics are low across Europe for the 20-year time period, with maximum values of 52 ppb and 44 ppb respectively observed in Northern Italy in 2003. Despite lower values across the continent, the highest of these (31 - 40 ppb) is consistent observed in Southern Europe and the central alpine region, consistent with previous studies (Wang et al., 2024).

For the US data, these general O_3 exposure metrics are generally higher than the observed values for Europe. At the start of the 20-year period, the lowest values of SOMO35 were observed in the Central and Eastern US (typically 1000 - 4000 ppb day), with higher values observed in the South (typically 4000 - 6000 ppb day), and the highest values found on the West coast (up to 11981 ppb day). This is consistent with previous studies, which have observed that Southern California is a hotspot for O_3 pollution (Fleming et al., 2018; Wang et al., 2024). Over the 20-year period, SOMO35 levels in Central, Eastern, and Southern USA generally reduce. However, values in the West and Intermountain West are found to increase or stay high, with SOMO35 values of 10973 ppb day observed in one site in 2020. Across the vicennial, all sites exceeding 7000 ppb day are located in the West, with a general decrease in the number of sites observing this limit over the 20 years. Similarly to the European data, the AVGMDA8 metric shows a similar pattern to the SOMO35 data, and by 2014 there is a clear divide between Central, Eastern and Southern US, where the lowest values are observed (up to 50 ppb), and the West and Southwest, including the Intermountain West (51 - 80 ppb). This narrative is also reflected in the 3MMDA1 and 6MMDA1 metrics.

3.5 Coupling MDA8O_3 trends with 6MMDA1 O_3 exposure metrics

To investigate how trends are developing in different clustered regions across Europe and the US variations in the annual exposure metric (6MMDA1) are compared to the slopes and significance of trends in MDA8O_3 in 2004 and 2018. Here, we only look at sites that have been assigned a cluster (where the number of sites > 3), using the HC clusters analysis described in section 2 to explore whether regional clusters were showing similar trends. Figures 9 - 11 show the MDA8O_3 trend against 6MMDA1, coloured by the clusters shown in section 3.3. Figures S22 - S24 show these same scatters coloured by significance.

In the European MDA8O_3 data and in the 5th quantile where sites are behaving more uniformly (one bulk cluster), we do not see a clear reduction in high certainty increasing slopes between 2004 and 2018, but we do see fewer high certainty decreasing slopes in 2018. Across the cluster, which includes a broad range of sites across Europe, we observe similar maximum values of annual 6MMDA1 in the two years (c.a. 68 ppbv), but fewer 6MMDA1 values in the lower range in 2018 (20 - 40 ppbv). For the 50th quantile data, where there are more clusters, observe that a cluster located across Southern mainland Europe (cluster 2), includes several sites with the higher 6MMDA1 values (> 60 ppbv), in both 2004 and 2018. However, the certainty and direction of the trends in MDA8O_3 is mixed. A smaller cluster of sites located in Spain observed lower values of 6MMDA1

in 2004 (20 - 40 ppbv), which increases to 40 - 50 ppbv by 2018. These sites also have high certainty increasing $MDA8O_3$ trends in both 2004 and 2018. More broadly across all clusters, we again observe fewer high certainty decreasing slopes in 2018 than 2004. In the 95 th quantile, where the clearest regional clustering can be observed, generally higher 6MMDA1 values are seen in 2018 than 2004. There is no notable grouping of the clusters their 6MMDA1 values in either 2004 and 2018, but there is generally a broader spread of values in 2004 compared to 2018. The location of sites with high certainty increasing or decreasing values is also mixed, showing no clear regionality. When we compare 6MMDA1 values to the 95 th quantile warm-season only $MDA8O_3$ trends, we again observed a compression of the range of 6MMDA1 values, at the higher mixing ratio end (c.a. 25 - 70 ppbv in 2004, vs 40 - 70 ppbv in 2018). We also observe that clusters located in Northern Europe have the smallest 6MMDA1 values in both years, and trends are generally increasing but with low certainty.

Average 6MMDA1 values across the USA are generally higher than those reported across Europe. Some clearer regional clustering can be observed in the 50 th quantile $MDA8O_3$ United States data, with a cluster located in Florida (cluster 2) generally showing lower 6MMDA1 values (40 - 60 ppbv) in both 2004 and 2018. Additionally, these sites are generally high certainty decreasing trends, and this is reflected in both the cold-season and warm-season only $MDA8O_3$ trends. In contrast, sites located in the West show the highest 6MMDA1 mixing ratios in both 2004 and 2018 (c.a. 70 - 90 ppbv) (cluster 7), and these are typically high certainty increasing trends in both years. These sites also have the highest 6MMDA1 values in the 95 th quantile (cluster 7), but increasing trends in this quantile are generally of low certainty. More interestingly, the 95 th quantile cluster located in the West including the Intermountain West region of the USA (cluster 2), typically observes 6MMDA1 values of 55 - 70 ppbv in both 2004 and 2018, but there is a upward shift in the slope magnitude of this cluster from approximately -1 - 0.5 ppb per year in 2004, to -0.5 - 1.25 ppbv per year. In addition, most of these trends shift from being of low certainty in 2004, to high certainty increasing trends in 2018. These trends are also reflected in the warm-season $MDA8O_3$ trends. Over the 21st century, the Intermountain West has experienced both increased wildfires, and hotter springs and summers, which can enhance photochemical activity forming O_3 from regional anthropogenic emissions (Lin et al., 2017; Li et al., 2021; Peterson et al., 2021; Iglesias et al., 2022). The higher magnitude and high certainty increasing slopes in $MDA8O_3$ may be linked to increases in one or both of these factors.

4 Conclusions

Trends in urban $MDA8O_3$ were calculated from de-seasoned monitoring site data across both Europe and the USA. Piecewise quantile regression analysis was utilised to assess long-term trends. Piecewise trends were limited to two change points per time series to keep the focus on long-term trends and the largest changes. Generally more sites in the United States of America were described by a trend line with no change points. In comparison, sites in Europe almost excursively described by trends with 1 or 2.

By varying the selected quantile, quantile regression allows trends in different areas of the $MDA8O_3$ distribution to be quantified. The analysis presented here focuses on the 0.05, 0.50 and 95th quantiles, relating to low, medium and high values of $MDA8O_3$. Additionally these trends were calculated for warm season and cold season subgroups. Across the 22 year period,

there are many sites in Europe with high certainty increasing trends in the 5th and 50 th quantiles, but the majority of medium-high certainty 95th quantile trends in Europe were found to be decreasing. Separating the trends by season shows that the majority of these decreasing trends are found in the warm season in all quantiles and very few trends were increasing at τ = 0.95 with p < 0.33. In the USA the main time where increasing trends are observed is in the cold season, in all quantiles but to a lesser extent at τ = 0.95. Overall τ = 0.95 trends are dominated by sites decreasing, driving by the warm season trends, though a small number of these trends at higher quantiles do appear to be increasing in the Intermountain region.

A primary challenge of interpreting monitoring station data is how to identity sites showing similar behaviour, without grouping based on pre-existing geographic regions. Hierarchical clustering with dynamic time warping was shown to be effective in grouping similar time series and returned differing clusters when time series were aggregated at different quantiles which agreed with those observed qualitatively when mapping calculated trends. Clusters tended to be more homogeneous at lower values of τ, with differing subregional behaviour becoming more apparent at the τ = 0.95. In the high extreme trends, a cluster located in California was showing increasing trends across the 22 year period, but these were typically of lower certainty. Higher certainty increasing trends were observed in a cluster located in the Intermountain West in the latter half of the century, which were of lower magnitude. O_3 exposure metrics, SOMO35, AVGMDA8, 3MMDA1, and 6MMDA1 were typically higher in the United States of America than Europe, and the European sites that where AVGMDA8 > 56 ppbv were located in Southern and the central alpine region of the continent. In the USA, sites located in the West observed SOMO35 > 7000 ppbv day more consistently over the 22 year period, though the number of sites exceeding this limit decreased in frequency with time. More interestingly, although 6MMDA1 values in the Intermountain West were similar in 2004 and 2018, the MDA8O_3 trends shifted from low certainty in 2004 to high certainty increasing trends in 2018. This was also observed in the warm-season trends, and may be linked to an observed increase in wildfire events in the area over the 22 year period, or warmer summers and springs as reported in previous studies (Lin et al., 2017; Li et al., 2021; Peterson et al., 2021; Iglesias et al., 2022). With the risk of increasing occurrences of warm summers, wildfires, and heatwaves, it is crucial that we continue to understand trends in MDA8O_3, as this study has identified increasing trends in locations in the USA (the West and Intermountain West), and Europe (Southern and Central alpine region), that already experience hot and dry weather conditions, that may exacerbate human risk to O_3 exposure.

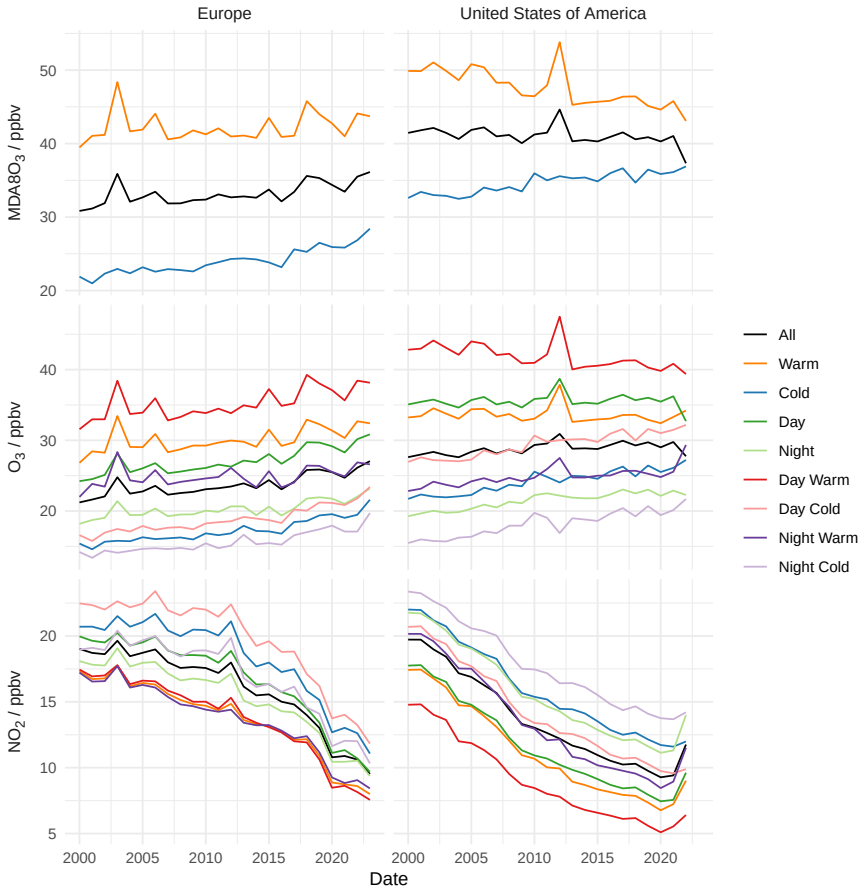

Figure 1. Time series of annual average $MDA8O_3$, O_3 and NO_2 for urban sites in Europe and the USA. Day and night subgroups correspond to the hours of 0800 L - 1900 L and 2000 - 0700 L respectively, warm and cold refer to the months April - September and October - March respectively.

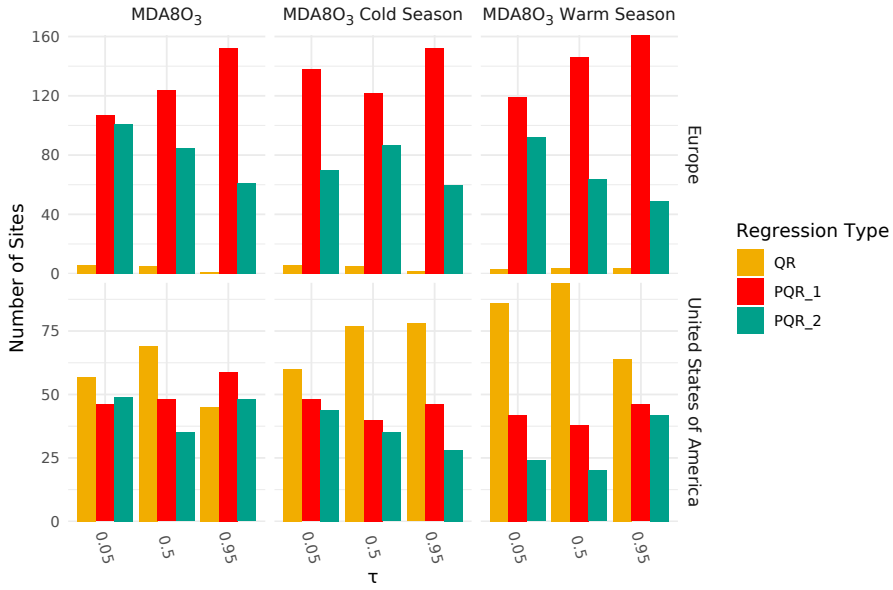

Figure 2. Number of sites with no change point (QR, gold), one change point (PQR_1, red) and two change points (PQR_2, green) in the Daily Mean, $MDA8O_3$ and $MDA8O_3$ warm season series, separated into Europe and United States of America regions.

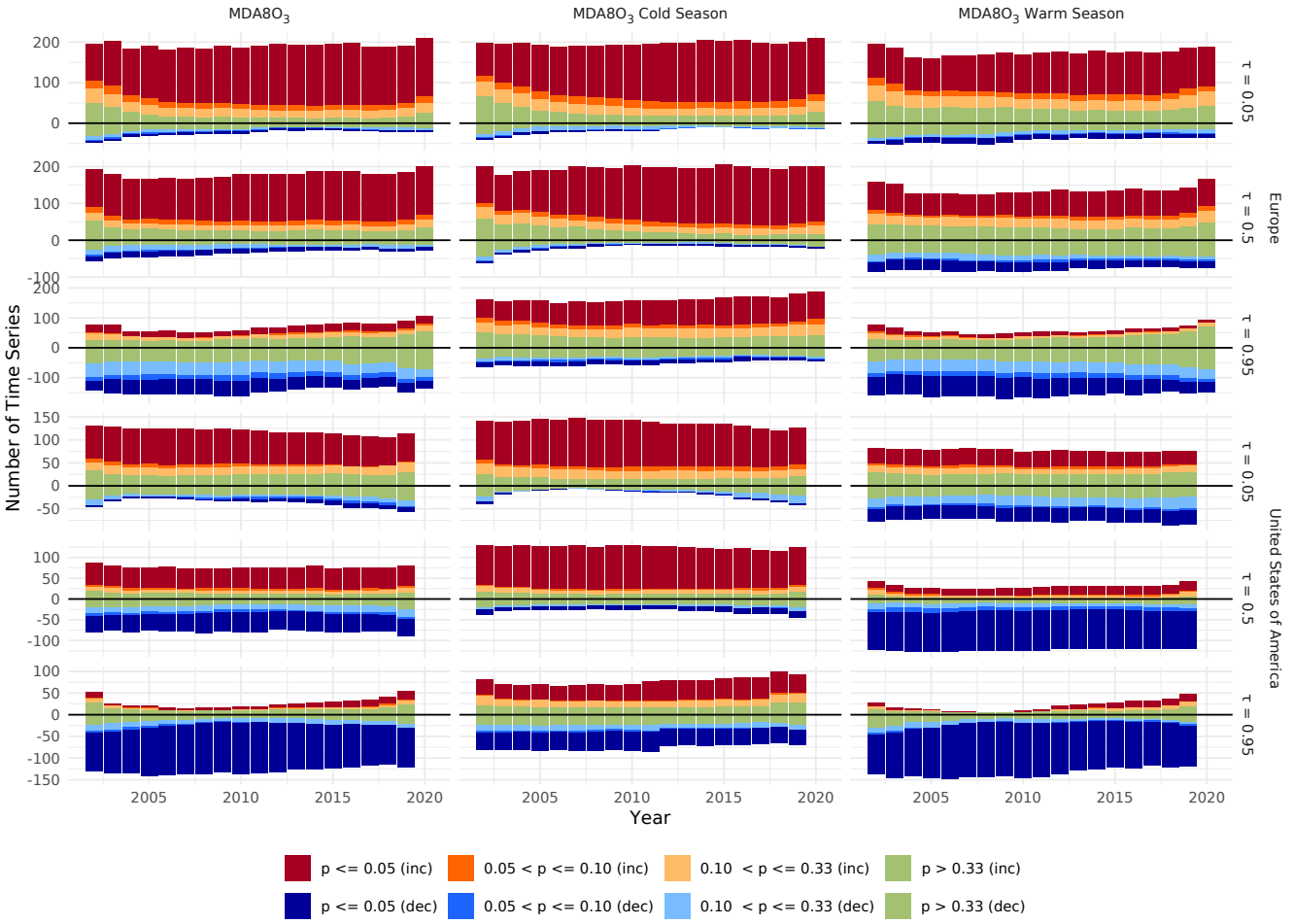

Figure 3. Time series of trends in $MDA8O_3$ for (left) annual, (middle) warm season, and (right) cold season, for Europe (top) and the USA (bottom). Within each region the panels are grouped by $\tau = 0.05$ (top), 0.5 (middle), 0.95 (bottom)

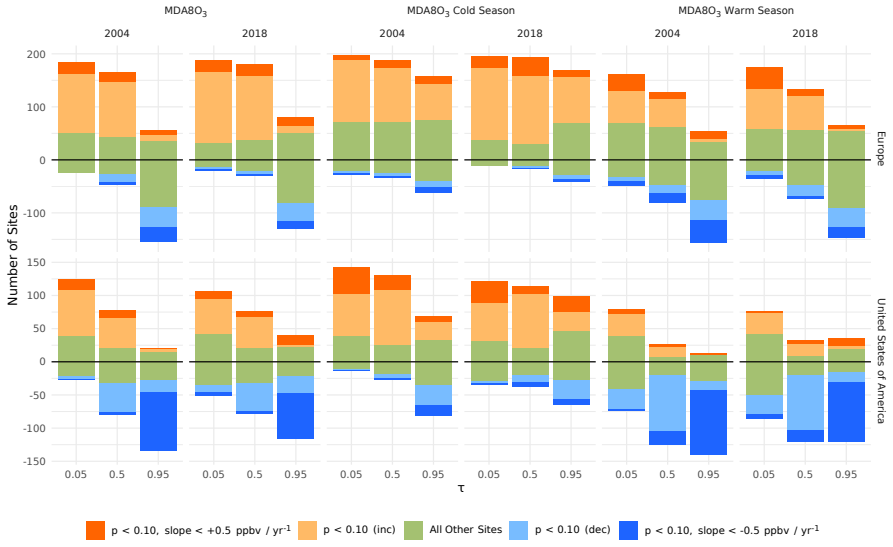

Figure 4. Number of O_3 series across slope significance and slope magnitude categories. Those of medium significance ($p > 0.10$) and medium significance and medium magnitude slope are highlighted.

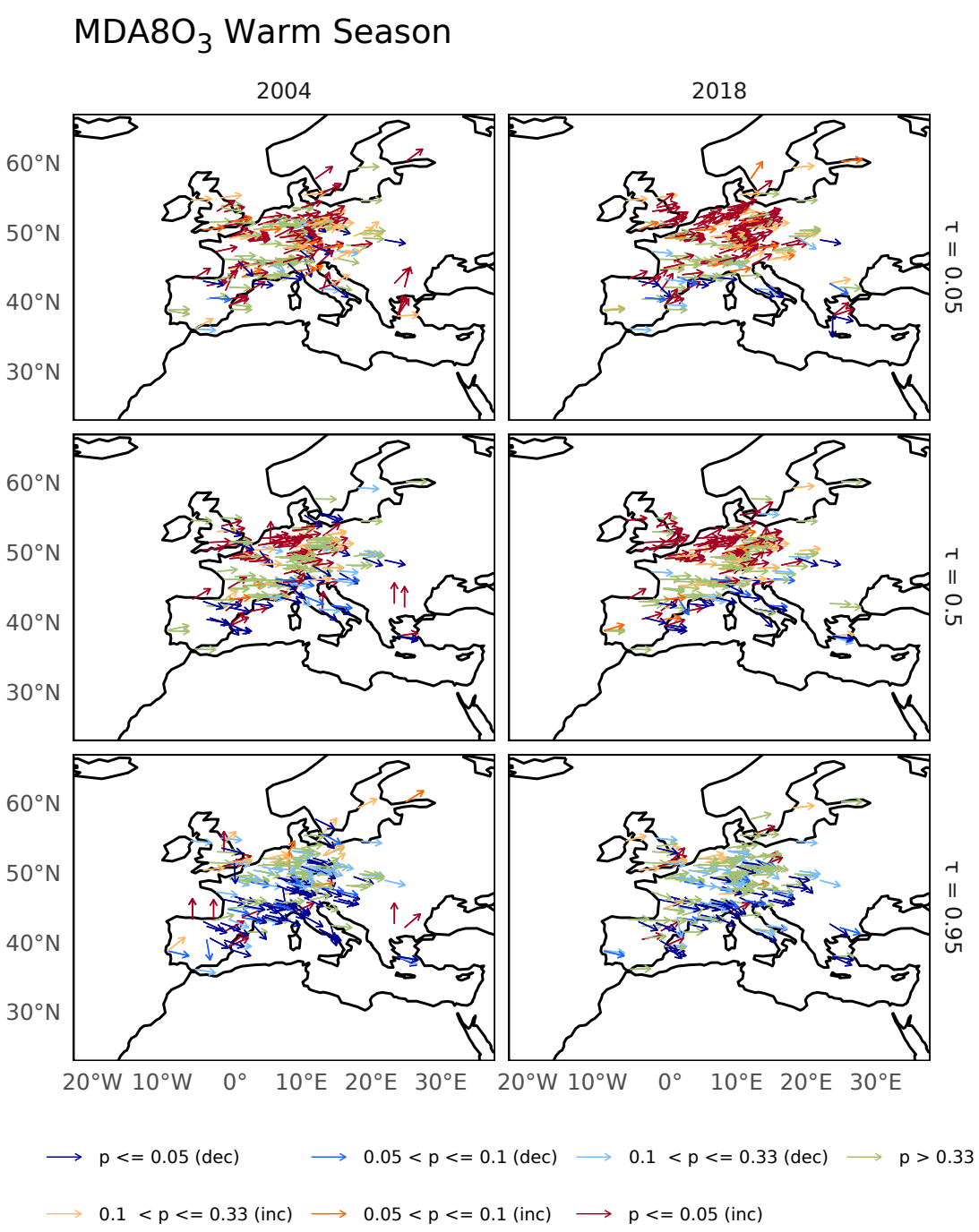

Figure 5. Trends in warm season $MDA8O_3$ in Europe in 2004 and 2018. Trends are shown by an arrow where the angle between vertically up and horizontal corresponds to 2.5 - 0 ppbv / yr^{-1}, and between horizontal and vertically down corresponds to 0 - -2.5 ppbv / yr^{-1}. The few trends that have a magnitude greater than 2.5 ppbv have been clamped. Colour corresponds to the direction and significant of the trend.

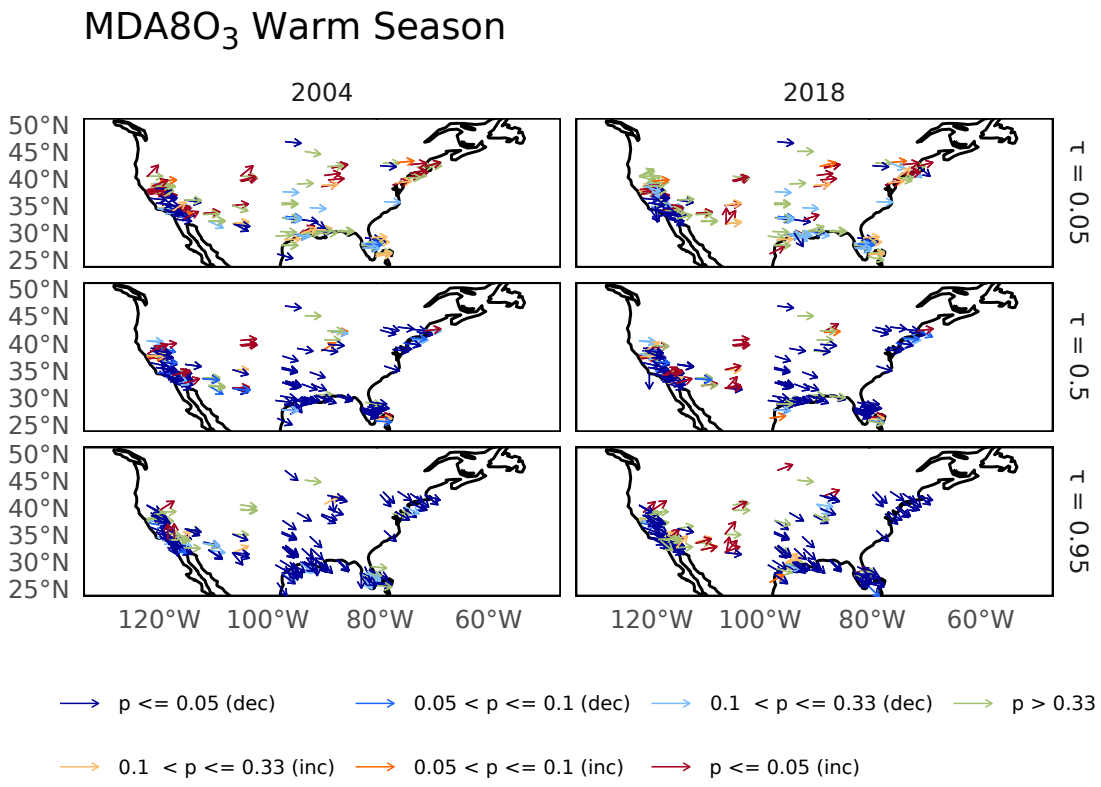

Figure 6. Trends in warm season $MDA8O_3$ in the United States of America in 2004 and 2018. Trends are shown by an arrow where the angle between vertically up and horizontal corresponds to 2.5 - 0 ppbv / yr^{-1}, and between horizontal and vertically down corresponds to 0 - -2.5 ppbv / yr^{-1}. The few trends that have a magnitude greater than 2.5 ppbv have been clamped. Colour corresponds to the direction and significant of the trend.

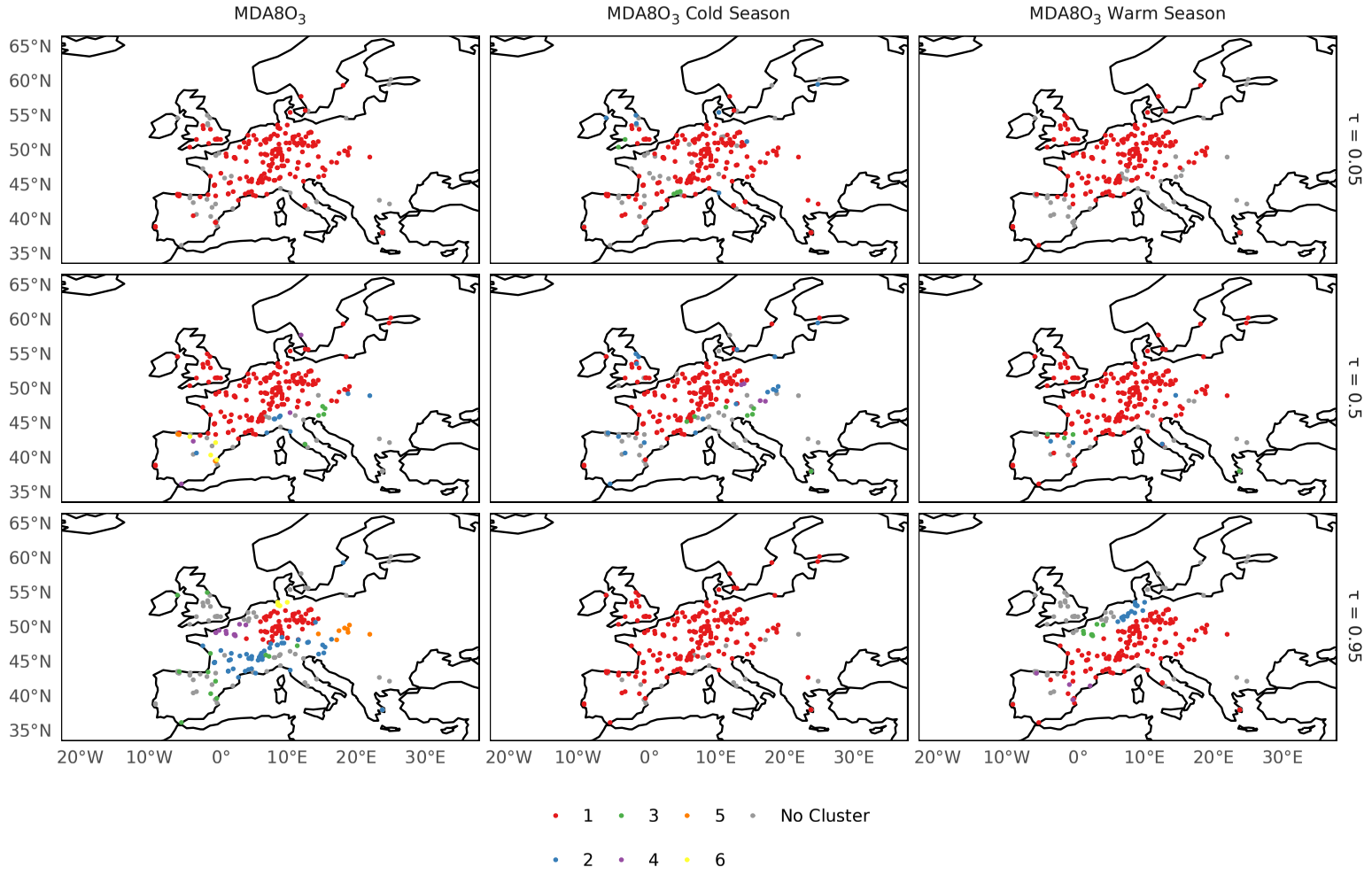

Figure 7. Clusters determined for $MDA8O_3$, $MDA8O_3$ warm season and $MDA8O_3$ cold season at τ 0.05, 0.50 and 0.95 in Europe. Clusters are coloured when there are > 3 sites within a cluster, otherwise these are assigned 'No Cluster'. Clusters are numbered starting with the cluster containing the most sites as 1 within each pane.

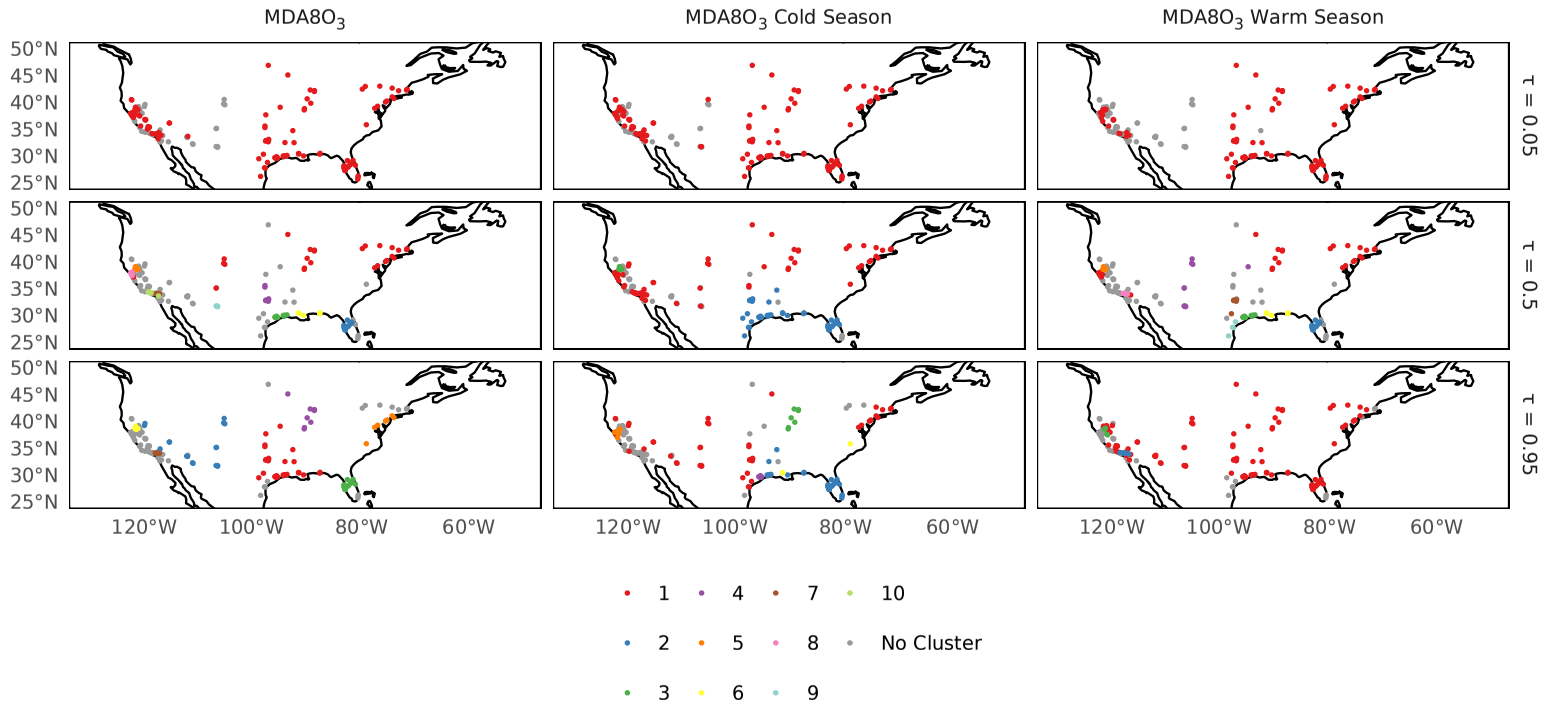

Figure 8. As figure 7 for the United States of America

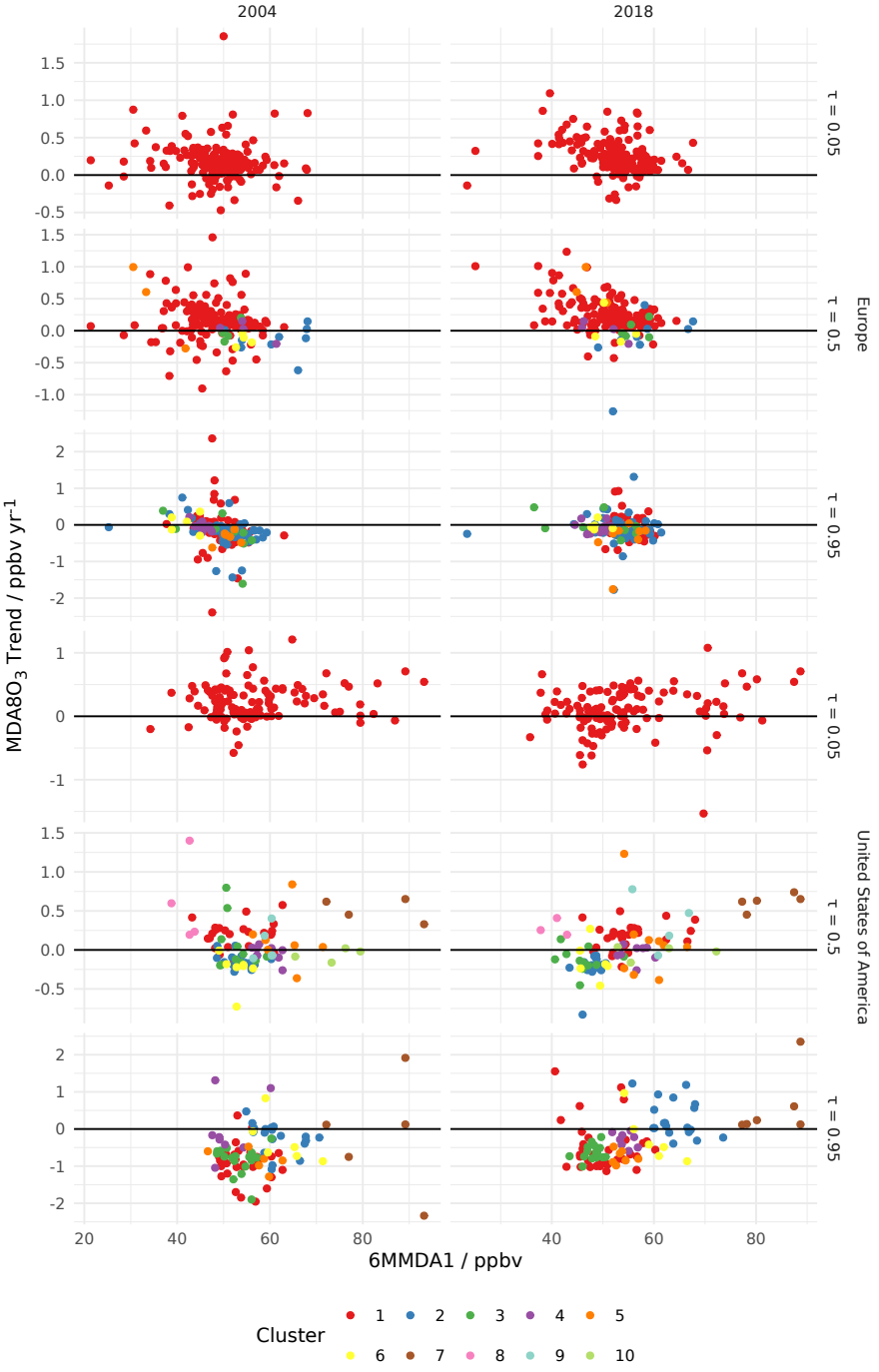

Figure 9. For the years 2004 (left) and 2018 (right) scatter plots show that years' $MDA8O_3$ trend in (top to bottom) Europe τ = 0.05, 0.5, 0.95, United States of America τ = 0.05, 0.5, 0.95 vs the 6MMDA1 for the same year. Colours correspond to clusters presented in figures 7 and 8.

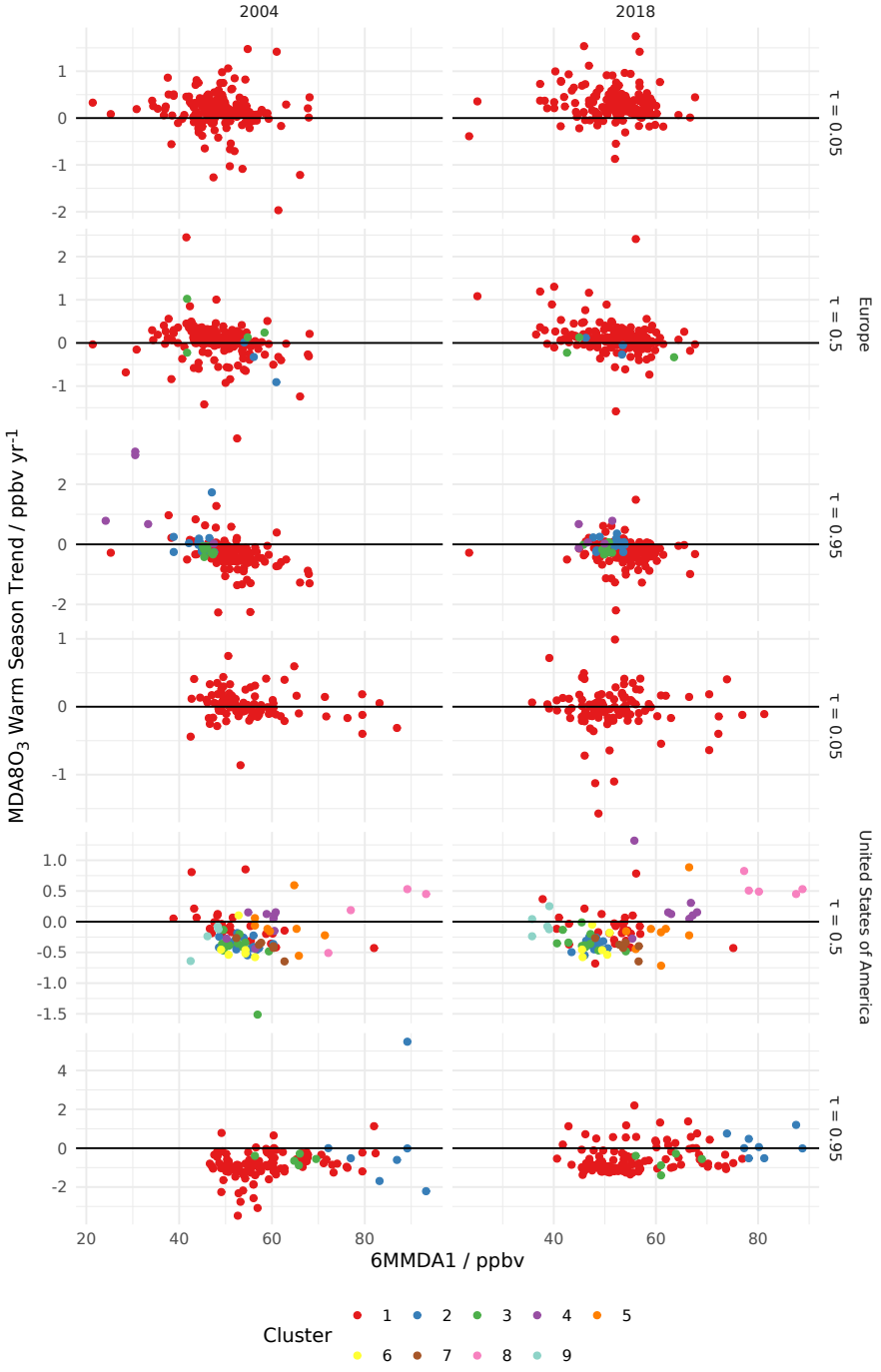

Figure 10. For the years 2004 (left) and 2018 (right) scatter plots show that years' warm season $MDA8O_3$ trend in (top to bottom) Europe τ = 0.05, 0.5, 0.95, United States of America τ = 0.05, 0.5, 0.95 vs the 6MMDA1 for the same year. Colours correspond to clusters presented in figures 7 and 8.

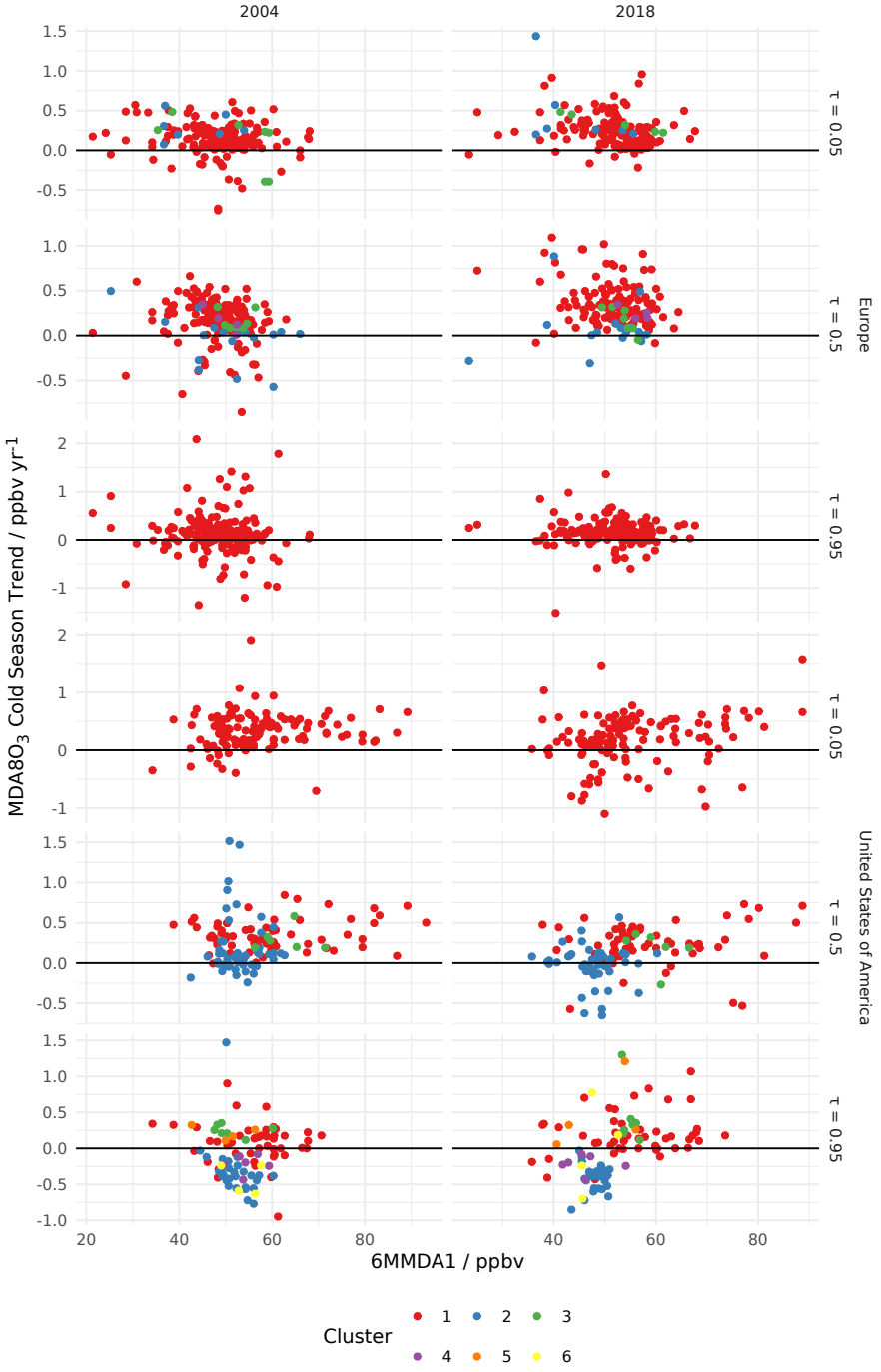

Figure 11. For the years 2004 (left) and 2018 (right) scatter plots show that years' cold season $MDA8O_3$ trend in (top to bottom) Europe τ = 0.05, 0.5, 0.95, United States of America τ = 0.05, 0.5, 0.95 vs the 6MMDA1 for the same year. Colours correspond to clusters presented in figures 7 and 8.

Code and data availability. All data for this study can be downloaded from the relevant databases , this study used the r-packages *toarR* (Drysdale, 2024) and *saqgetr* (Grange, 2019). Code for performing the download as well as the analysis can be found at 10.5281/zenodo.14538198 (Drysdale and Nelson, 2024)

Author contributions. BSN and WSD equally contributed to all aspects of this manuscript's production

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

Acknowledgements. The Viking cluster was used during this project, which is a high performance compute facility provided by the University of York. We are grateful for computational support from the University of York, IT Services and the Research IT team.

The Authors acknowledge Prof. James Lee for their scientific advice and Prof. David Carslaw and Dr Stuart Lacy for their advice on the statistical analysis. We also thank Dr Stuart Lacy for their help with SQL and very useful suggestions on managing large datasets.

References

Aghabozorgi, S., Seyed Shirkhorshidi, A., and Ying Wah, T.: Time-series clustering – A decade review, Information Systems, 53, 16–38, https://doi.org/https://doi.org/10.1016/j.is.2015.04.007, 2015.

Berndt, D. J. and Clifford, J.: Using dynamic time warping to find patterns in time series, AAAIWS'94, p. 359–370, AAAI Press, 1994.

Chang, K.-L., Schultz, M. G., Koren, G., and Selke, N.: Guidance note on best statistical practices for TOAR analyses, https://arxiv.org/abs/2304.14236, 2023.

Chang, K.-L., McDonald, B. C., Harkins, C., and Cooper, O. R.: Surface ozone trend variability across the United States and the impact of heat waves (1990–2023), Atmospheric Chemistry and Physics, 25, 5101–5132, https://doi.org/10.5194/acp-25-5101-2025, 2025.

Cooper, O. R., Schultz, M. G., Schröder, S., Chang, K.-L., Gaudel, A., Benítez, G. C., Cuevas, E., Fröhlich, M., Galbally, I. E., Molloy, S., Kubistin, D., Lu, X., McClure-Begley, A., Nédélec, P., O'Brien, J., Oltmans, S. J., Petropavlovskikh, I., Ries, L., Senik, I., Sjöberg, K., Solberg, S., Spain, G. T., Spangl, W., Steinbacher, M., Tarasick, D., Thouret, V., and Xu, X.: Multi-decadal surface ozone trends at globally distributed remote locations, Elementa: Science of the Anthropocene, 8, 23, https://doi.org/10.1525/elementa.420, 2020.

Drysdale, W. S.: wacl-york/toarR: v0.1.0, https://doi.org/10.5281/zenodo.14537446, 2024.

Drysdale, W. S. and Nelson, B. S.: wacl-york/Urban-Ozone-Trends-in-Europe-and-the-USA-2000-2021: v0.1.0 preprint, https://doi.org/10.5281/zenodo.14538198, 2024.

European Environment Agency: Air Quality e-Reporting, https://www.eea.europa.eu/en/datahub/datahubitem-view/, a.

European Environment Agency: EIONET, https://cdr.eionet.europa.eu/, b.

Fleming, Z. L., Doherty, R. M., von Schneidemesser, E., Malley, C. S., Cooper, O. R., Pinto, J. P., Colette, A., Xu, X., Simpson, D., Schultz, M. G., Lefohn, A. S., Hamad, S., Moolla, R., Solberg, S., and Feng, Z.: Tropospheric Ozone Assessment Report: Present-day ozone distribution and trends relevant to human health, Elementa: Science of the Anthropocene, 6, 12, https://doi.org/10.1525/elementa.273, 2018.

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., Eboijie, F., Foret, G., Garcia, O., Granados-Muñoz, M. J., Hanningan, J. W., Hase, F., Hassler, B., Huang, G., Hurtmans, D., Jaffe, D., Jones, N., Kalabokas, P., Kerridge, B., Kulawik, S., Latter, B., Leblanc, T., Le Flochmoën, 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., Shultz, M. G., Schwab, J., Siddans, R., Smale, D., Steinbacher, M., Tanimoto, H., W., T. D., 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, Elem. Sci. Anth., 6, 1–58, https://doi.org/10.1525/elementa.291, 2018.

Gouldsbrough, L., Hossaini, R., Eastoe, E., and Young, P. J.: A temperature dependent extreme value analysis of UK surface ozone, 1980–2019, Atmospheric Environment, 273, 118 975, https://doi.org/https://doi.org/10.1016/j.atmosenv.2022.118975, 2022.

Grange, S.: Technical note: saqgetr R package, https://doi.org/10.13140/RG.2.2.15533.44001, 2019.

He, H., Liang, X.-Z., Sun, C., Tao, Z., and Tong, D. Q.: The long-term trend and production sensitivity change in the US ozone pollution from observations and model simulations, Atmospheric Chemistry and Physics, 20, 3191–3208, https://doi.org/10.5194/acp-20-3191-2020, 2020.

Iglesias, V., Blach, J. K., and Travis, W. R.: U.S. fires became larger, more frequent, and more widespread in the 2000s, Sci. Adv., 8, 1–10, https://doi.org/10.1126/sciadv.abc0020, 2022.

Jonson, J. E., Simpson, D., Fagerli, H., and Solberg, S.: Can we explain the trends in European ozone levels?, Atmospheric Chemistry and Physics, 6, 51–66, https://doi.org/10.5194/acp-6-51-2006, 2006.

Lefohn, A. S., Malley, C. S., Smith, L., Wells, B., Hazucha, M., Simon, H., Naik, V., Mills, G., Schultz, M. G., Paoletti, E., De Marco, A., Xu, X., Zhang, L., Wang, T., Neufeld, H. S., Musselman, R. C., Tarasick, D., Brauer, M., Feng, Z., Tang, H., Kobayashi, K., Sicard, P., Solberg, S., and Gerosa, G.: Tropospheric ozone assessment report: Global ozone metrics for climate change, human health, and crop/ecosystem research, Elementa: Science of the Anthropocene, 6, 27, https://doi.org/10.1525/elementa.279, 2018.

Li, S., Wang, H., and Lu, X.: Anthropogenic emission controls reduce summertime ozone–temperature sensitivity in the United States, Atmospheric Chemistry and Physics, 25, 2725–2743, https://doi.org/10.5194/acp-25-2725-2025, 2025.

Li, Z., Angerer, J. P., and Wu, X. B.: Temporal Patterns of Large Wildfires and Their Burn Severity in Rangelands of Western United States, Geophysical Research Letters, 48, 1–10, https://doi.org/10.1029/2020GL091636, 2021.

Lin, M., Horowitz, L. W., Payton, R., Fiore, A. M., and Tonnesen, G.: US surface ozone trends and extremes from 1980 to 2014: quantifying the roles of rising Asian emissions, domestic controls, wildfires, and climate, Atmos. Chem. Phys., 17, 2943–2970, https://doi.org/10.5194/acp-17-2943-2017, 2017.

Malashock, D. A., Delang, M. N., Becker, J. S., Serre, M. L., West, J. J., Chang, K.-L., Cooper, O. R., and Anenberg, S. C.: Global trends in ozone concentration and attributable mortality for urban, peri-urban, and rural areas between 2000 and 2019: a modelling study, The Lancet Planetary Health, 6, e958–e967, https://doi.org/https://doi.org/10.1016/S2542-5196(22)00260-1, 2022.

Meehl, G. A., Tebaldi, C., Tilmes, S., Lamarque, J.-F., Bates, S., Pendergrass, A., and Lombardozzi, D.: Future heat waves and surface ozone, Environmental Research Letters, 13, 064 004, https://doi.org/10.1088/1748-9326/aabcdc, 2018.

Mills, G., Pleijel, H., Malley, C. S., Sinha, B., Cooper, O. R., Schultz, M. G., Neufeld, H. S., Simpson, D., Sharps, K., Feng, Z., Gerosa, G., Harmens, H., Kobayashi, K., Saxena, P., Paoletti, E., Sinha, V., and Xu, X.: Tropospheric Ozone Assessment Report: Present-day tropospheric ozone distribution and trends relevant to vegetation, Elementa: Science of the Anthropocene, 6, 47, https://doi.org/10.1525/elementa.302, 2018.

Mueen, A. and Keogh, E.: Extracting Optimal Performance from Dynamic Time Warping, KDD '16, p. 2129–2130, Association for Computing Machinery, New York, NY, USA, ISBN 9781450342322, https://doi.org/10.1145/2939672.2945383, 2016.

Otero, N., Sillmann, J., Schnell, J. L., Rust, H. W., and Butler, T.: Synoptic and meteorological drivers of extreme ozone concentrations over Europe, Environmental Research Letters, 11, 024 005, https://doi.org/10.1088/1748-9326/11/2/024005, 2016.

Otero, N., Rust, H. W., and Butler, T.: Temperature dependence of tropospheric ozone under NOx reductions over Germany, Atmospheric Environment, 253, 118 334, https://doi.org/https://doi.org/10.1016/j.atmosenv.2021.118334, 2021.

Peterson, G. C. L., Prince, S. E., and Rappold, A. G.: Trends in Fire Danger and Population Exposure along the Wildland-Urban Interface, Environ. Sci. Technol., 55, 16 257–16 265, https://doi.org/10.1021/acs.est.1c03835, 2021.

Reed, D. E., Chu, H., Peter, B. G., Chen, J., Abraha, M., Amiro, B., Anderson, R. G., Arain, M. A., Arruda, P., Barron-Gafford, G. A., Bernacchi, C., Beverly, D. P., Biraud, S. C., Black, T. A., Blanken, P. D., Bohrer, G., Bowler, R., Bowling, D. R., Bret-Harte, M. S., Bretfeld, M., Brunsell, N. A., Bullock, S. H., Celis, G., Chen, X., Classen, A. T., Cook, D. R., Cueva, A., Dalmagro, H. J., Davis, K., Desai, A., Duff, A. J., Dunn, A. L., Durden, D., Edgar, C. W., Euskirchen, E., Bracho, R., Ewers, B., Flanagan, L. B., Florian, C., Foord, V., Forbrich, I., Forsythe, B. R., Frank, J., Garatuza-Payan, J., Goslee, S., Gough, C., Green, M., Griffis, T., Helbig, M., Hill, A. C., Hinkle, R., Horne, J., Humphreys, E., Ikawa, H., Iwahana, G., Jassal, R., Johnson, B., Johnson, M., Kannenberg, S. A., Kelsey, E., King, J., Knowles, J. F., Knox, S., Kobayashi, H., Kolb, T., Kolka, R., Krauss, K. W., Kutzbach, L., Lamb, B., Law, B., Lee, S.-C., Lee, X., Liu, H., Loescher, H. W., Malone, S. L., Matamala, R., Mauritz, M., Metzger, S., Meyer, G., Mitra, B., Munger, J. W., Nesic,

Z., Noormets, A., O'Halloran, T. L., O'Keeffe, P. T., Oberbauer, S. F., Oechel, W., Oikawa, P., Olivas, P. C., Ouimette, A., Pastorello, G., Perez-Quezada, J. F., Phillips, C., Posse, G., Qu, B., Quinton, W. L., Reba, M. L., Richardson, A. D., Picasso, V., Rocha, A. V., Rodriguez, J. C., Ruzol, R., Saleska, S., Scott, R. L., Schreiner-McGraw, A. P., Schuur, E. A., Silveira, M., Sonnentag, O., Spittlehouse, D. L., Staebler, R., Starr, G., Staudhammer, C., Still, C., Sturtevant, C., Sullivan, R. C., Suyker, A., Trejo, D., Ueyama, M., Vargas, R., Viner, B., Vivoni, E. R., Wang, D., Ward, E. J., Wiesner, S., Windham-Myers, L., Yannick, D., Yepez, E. A., Zenone, T., Zhao, J., and Zona, D.: Network of networks: Time series clustering of AmeriFlux sites, Agricultural and Forest Meteorology, 372, 110 686, https://doi.org/https://doi.org/10.1016/j.agrformet.2025.110686, 2025.

Russo, S., Sillmann, J., and Fischer, E. M.: Top ten European heatwaves since 1950 and their occurrence in the coming decades, Environmental Research Letters, 10, 124 003, https://doi.org/10.1088/1748-9326/10/12/124003, 2015.

Sarda-Espinosa, A.: dtwclust: Time Series Clustering Along with Optimizations for the Dynamic Time Warping Distance, https://doi.org/10.32614/CRAN.package.dtwclust, r package version 6.0.0, 2024.

Schär, C., Vidale, P. L., Lüthi, D., Frei, C., Häberli, C., Liniger, M. A., and Appenzeller, C.: The role of increasing temperature variability in European summer heatwaves, Nature, 427, 332–336, https://doi.org/10.1038/nature02300, 2004.

Schröder, S., Schultz, M. G., Selke, N., Sun, J., Ahring, J., Mozaffari, A., Romberg, M., Epp, E., Lensing, M., Apweiler, S., Leufen, L. H., Betancourt, C., Hagemeier, B., and Rajveer, S.: TOAR Data Infrastructure, https://doi.org/10.34730/4D9A287DEC0B42F1AA6D244DE8F19EB3, 2021.

Shen, L., Mickley, L. J., and Gilleland, E.: Impact of increasing heat waves on U.S. ozone episodes in the 2050s: Results from a multimodel analysis using extreme value theory, Geophysical Research Letters, 43, 4017–4025, https://doi.org/https://doi.org/10.1002/2016GL068432, 2016.

Sillman, S.: Chapter 12 The relation between ozone, NOx and hydrocarbons in urban and polluted rural environments, in: Air Pollution Science for the 21st Century, edited by Austin, J., Brimblecombe, P., and Sturges, W., vol. 1 of *Developments in Environmental Science*, pp. 339–385, Elsevier, https://doi.org/https://doi.org/10.1016/S1474-8177(02)80015-8, 2002.

Sillman, S., Logan, J. A., and Wofsy, S. C.: The sensitivity of ozone to nitrogen oxides and hydrocarbons in regional ozone episodes, Journal of Geophysical Research: Atmospheres, 95, 1837–1851, https://doi.org/https://doi.org/10.1029/JD095iD02p01837, 1990.

Simon, H., Reff, A., Wells, B., Xing, J., and Frank, N.: Ozone Trends Across the United States over a Period of Decreasing NOx and VOC Emissions, Environmental Science & Technology, 49, 186–195, https://doi.org/10.1021/es504514z, pMID: 25517137, 2015.

Szopa, S., Naik, V., Bhupesh, A., Artaxo, P., Berntsen, T., Collins, W. D., Fuzzi, S., Gallardo, L., Kiendler-Scharr, A., Klimont, Z., Liao, H., Unher, N., and Zanis, P.: Short-lived Climate Forcers, p. 817–922, Cambridge University Press, 2021.

UN: World urbanization prospects 2018 highlights, UN, http://digitallibrary.un.org/record/3828520, 2019.

Vazquez Santiago, J., Hata, H., Martinez-Noriega, E. J., and Inoue, K.: Ozone trends and their sensitivity in global megacities under the warming climate, Nature Communications, 15, 10 236, https://doi.org/10.1038/s41467-024-54490-w, 2024.

Wang, H., Wang, H., Lu, X., Lu, K., Zhang, L., Tham, Y. J., Shi, Z., Aikin, K., Fan, S., Brown, S. S., and Zhang, Y.: Increased night-time oxidation over China despite widespread decrease across the globe, Nature Geoscience, 16, 217–223, https://doi.org/10.1038/s41561-022-01122-x, 2023.

Wang, H., Lu, X., Palmer, P. I., Zhang, L., Lu, K., Li, K., Nagashima, T., Koo, J.-H., Tanimoto, H., Wang, H., Gao, M., He, C., Wu, K., Fan, S., and Zhang, Y.: Deciphering decadal urban ozone trends from historical records since 1980, Nat. Sci. Rev., 11, 1–13, https://doi.org/10.1093/nsr/nwae369, 2024.

Yan, Y., Pozzer, A., Ojha, N., Lin, J., and Lelieveld, J.: Analysis of European ozone trends in the period 1995–2014, Atmospheric Chemistry and Physics, 18, 5589–5605, https://doi.org/10.5194/acp-18-5589-2018, 2018.