CRC Report No. A-124

EVALUATION OF OZONE PATTERNS AND TRENDS IN 8 MAJOR METROPOLITAN AREAS IN THE U.S.

Final Report

March 2021

COORDINATING RESEARCH COUNCIL, INC.
5755 NORTH POINT PARKWAY • SUITE 265 • ALPHARETTA, GA 30022

The Coordinating Research Council, Inc. (CRC) is a non-profit corporation supported by the petroleum and automotive equipment industries. CRC operates through the committees made up of technical experts from industry and government who voluntarily participate. The four main areas of research within CRC are: air pollution (atmospheric and engineering studies); aviation fuels, lubricants, and equipment performance; heavy-duty vehicle fuels, lubricants, and equipment performance (e.g., diesel trucks); and light-duty vehicle fuels, lubricants, and equipment performance (e.g., passenger cars). CRC's function is to provide the mechanism for joint research conducted by the two industries that will help in determining the optimum combination of petroleum products and automotive equipment. CRC's work is limited to research that is mutually beneficial to the two industries involved. The final results of the research conducted by, or under the auspices of, CRC are available to the public.

EVALUATION OF OZONE PATTERNS AND TRENDS IN 8 MAJOR METROPOLITAN AREAS IN THE U.S.

Prepared by:

Daniel A. Jaffe
University of Washington
School of STEM and Department of Atmospheric Sciences
Bothell, WA 98011
djaffe@uw.edu

Prepared for:

Coordinating Research Council, Inc.
5755 North Point Parkway, Suite 265
Alpharetta, Georgia 30022

Table of Contents

List of Figures

List of Tables

Note on report organization and statistics

The goals, methods, results and conclusions for the 8 metropolitan regions considered are discussed in the main report. The appendix includes a series of graphs and tables that are presented in a consistent way for each of the regions studied. Figures and tables in the Appendix are numbered as S1, S2, etc. Page numbers in the Appendix are listed as A1, A2, etc.

Statistical parameters were calculated using a combination of Excel (Office 2019 version) and IBM SPSS, version 27. Generalized Additive Modeling was completed using R software, version 4.03. For this report, a statistically significant result implies a p value <0.05. For the correlation coefficient (R) we use the Pearson correlation coefficient. This provides a measure of association between two variables. Positive values of R represent a direct correlation, whereas negative values of R are inversely correlated. Note that the R^2 can be interpreted as to the degree to which the x variable explains the variance in the y variable. While there is no universal definition of “strong” or “weak” correlation, I consider correlations with an absolute value of R <0.5 to be “weak.” Finally, it should be noted that even weak correlations can still be statistically significant, if there is a sufficient number of observations.

Executive Summary

In this project, I examined O_3 and related data in 8 U.S. metropolitan regions for 2006-2018. These 8 regions currently do not meet the U.S. National Ambient Air Quality Standards (NAAQS) for O_3 and progress in the past decade has been slow. To understand the reasons for this, I evaluated the role of NO_x, temperature and smoke influence on these 8 regions. The primary conclusions from this work are:

1. NO_x remains one of the most important controls on the Maximum Daily 8-hour Average (MDA8) O_3 concentration at most locations considered. The long-term trend in annual mean NO_2 is well correlated with surface O_3 for most sites. In addition, NO_2 concentrations and the frequency of O_3 exceedances are lower on weekend days in 5 of the 8 cities considered. In most regions, daily variations in NO_2 are significantly correlated with the MDA8. As NO_2 concentrations have declined, the annual fourth highest MDA8 O_3 values have also declined in most regions. However the relationship with NO_x is not uniform across all metropolitan areas and in some regions, the concentrations of NO_2 appear to be only weakly linked to the highest O_3 days.
2. Smoke days are identified through a combination of the NOAA HMS Fire and Smoke Product and enhanced surface $PM_{2.5}$. For each region, the average number of identified smoke days per year for 2006-2018 ranges from 5-9 per year. For Salt Lake City and San Francisco, the number of smoke days per year in the 2016-2018 was ~17, much higher than the average for the entire period.
3. Using surface and radiosonde meteorological data, and satellite observations, I have used Generalized Additive Modeling (GAM) for sites in each Combined Statistical Area (CSA) to predict the MDA8 for May-Sept. The GAM results are trained using 90% of the non-smoke data and tested against the remaining 10%. These yield R^2 (observed compared with predicted) values between 0.52-0.77.
4. Smoke days have an average MDA8 that is 15 ppb higher than non-smoke days, but some of this effect is due to higher temperature on smoke days. Since the GAM calculations take into account temperature variations and the residuals are unbiased against temperature, the residuals are indicative of the smoke contribution to the MDA8. Based on the GAM results, the presence of smoke increases the <u>mean</u> MDA8 value by between 2-8 ppb, depending on the site.
5. If smoke days are excluded from the calculation, the 2016-2018 O_3 Design Value (ODV) would be reduced by up to 10 ppb for some locations. The largest impacts are seen for sites in San Francisco, Salt Lake City and, to a lesser extent, sites in New York and Detroit. **Air quality managers in these regions may wish to investigate these smoke days more thoroughly, so as to decide if these cases have policy relevance.**
6. In two of the cities considered, Phoenix and Salt Lake City, background O_3 has a much stronger contribution to the MDA8, compared to other locations. This is evident from the high MDA8 values seen in these cities, even on the cleanest days and is consistent with other published work.

1. Introduction and project goals

Surface ozone (O_3) is a criteria air pollutant that is formed from reactions of nitrogen oxides ($NO_x = NO+NO_2$) and volatile organic compounds (VOCs) in the presence of sunlight. O_3 has serious health impacts up to and including premature mortality. In the U.S., reductions in the precursor emissions, NO_x and VOCs, over the past several decades have reduced peak O_3 concentrations considerably (Simon et al., 2015), but at present, there are still more than 40 regions in the U.S. that exceed the current 8-hour O_3 standard, so more than 130 million Americans live in areas that do not meet the U.S. National Ambient Air Quality Standards (NAAQS). The current standard is met when the O_3 design value (ODV), defined as the annual fourth highest maximum daily 8-hour average (MDA8) averaged over 3 years, is 0.070 ppm or less. This standard has become stricter several times over the last few decades.

The chemistry of O_3 production is complex and non-linear. While both NO_x and VOCs are required for O_3 production, VOCs have significant biogenic sources that cannot be controlled. As a result there is strong evidence that most high O_3 regions in the U.S. are now or are close to being “NO_x limited” (see discussion in Nussbaumer and Cohen 2020). This means that we would expect O_3 production to decline in a nearly linear fashion with NO_x reductions. I will explore this concept further in the results.

In addition to urban photochemistry, O_3 can also come from background and uncontrollable sources, such as stratospheric intrusions, wildfire emissions or transported from international sources. In general, elevated regions in the Western U.S. are exposed to higher levels of background O_3 (Jaffe et al 2018). In addition, due to the increase in large wildfires in California and the intermountain west in recent years (Jaffe et al 2020), sites in the western U.S. have likely experienced greater contributions from background O_3 sources. Along with the significant variability due to meteorology, these factors given rise to substantial year-to-year variations in O_3 concentrations that are not directly related to emissions.

Given that NO_x and VOC emissions have declined in the last 3 decades, we want to evaluate if the changes in O_3 are consistent with the observed emission trends. Many areas of the country have shown little or no change in annual fourth highest O_3 in the last 5 years and it is not clear if these are random variations, influence from smoke or evidence of a more systematic effect that we need to understand. In general terms, the goals of this project are to understand the relationship between O_3, NO_x, smoke and temperature in each of the 8 metropolitan regions considered. Specific goals are:

1. Evaluate the relationship between daily MDA8 O_3, temperature, NO_x and the presence/absence of smoke for 8 large metropolitan areas in the U.S. that have a high ODV.
2. Develop Generalized Additive Models to understand the relationship between daily MDA8 values and key meteorological predictors for these 8 large metropolitan areas in the U.S.
3. Examine patterns and trends in the GAM results to evaluate the causes for trends or lack of trends at each of the 8 regions. In particular, I will examine the role that smoke, temperature and NO_x play on long term O_3 changes.

2. Data sources

In this project I examined the trend and controlling factors on O_3 in 8 U.S. metropolitan areas. Our region focus is set by the Combined Statistical Areas (CSAs) in these 8 regions, as defined by the United States Office of Management and Budget (OMB). These regions often include multiple adjoining cities. For example, for the analysis of O_3 in the New York City region, I use observations from the New York-Newark, NY-NJ-CT-PA CSA. Note that some of the highest O_3 concentrations in the region are seen at the Fairfield CT and Leonia NJ monitoring locations, and these are clearly associated with the NY metropolitan area. This is evident in the plot of O_3 concentrations vs transport direction (Figure S6 for the Fairfield site). Throughout this report, I use the city name or a 2 or 3 digit letter abbreviation to be synonymous with the CSA.

I integrated data from a wide variety of data sources. This includes surface pollution data from the Environmental Protection Agency's Air Quality System (AQS), surface and radiosonde meteorological data from NOAA, and several products from NASA, including the Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2) global assimilation model and UV and NO_2 tropospheric column observations from the Ozone Monitoring Instrument (OMI) onboard the Aura satellite. Specific sites used and coordinates for the MERRA-2 and OMI products are given in Table 1. Note that the coordinates for the NASA products (MERRA-2 and OMI) that are used to define each region, are given in the specific format needed for the Giovanni data access platform (see https://giovanni.gsfc.nasa.gov/ for more details). Note that while NO and NO_2 are in rapid photochemical equilibrium, most urban areas only have long-term observations of NO_2, since it is the criteria pollutant. This should not have any bearing on our study, since NO_2 is usually the dominant form.

Table 1. List of sites used in each CSA and data sources.

Region	O_3 sites (name and AQS id)	NO_2 site (AQS id)	Surface met site	Radiosonde site (station id, code)	MERRA-2 coordinates	OMI coordinates (UV and NO_2)
Atlanta CSA (ATL)	United Avenue (UA)-131210055	130890002	Hartsfield Airport	Peachtree City (72215, FFC)	-84.46,33.62, -84.26,33.82	-84.61, 33.47, -84.11, 33.97
Chicago CSA (CHI)	Chiwaukee (CHIW) 550590019	170310063 170310072 170310076 170313103 170314002 170314201 171971011 180890022	Midway Airport	Davenport (74455, DVN)	-87.91, 42.40, -87.71, 42.60,	-88.06, 42.25, -87.56, 42.75,
Dallas CSA (DAL)	Grapevine Fairway (GRP)- 484393009	481130069 481130075 481130087 481210034 481390016 482311006 482570005 484391002 484393009 484393011	Dallas Love Field Airport	Ft. Worth (72249, FWD)	-97.16, 32.88, -96.96, 33.08	-97.31, 32.73, -96.81, 33.23
Detroit CSA (DET)	New Haven (NEW) 260990009 Oak Park (OAK) 261250001 Port Huron (PORT) 261470005 East 7 Mile (EAST 261630019	261630019 261630094	Pontiac Oakland Airport	White Lake (72632, DTX)	-82.89, 42.63, -82.69,42.83	-83.04, 42.48, -82.54, 42.98
NY-NJ-CT-PA CSA (NYC)	Leonia (LEO) 340030006 Fairfield (FAIR) 090019003	360050133 360810124	NY Central Park	Upton (72501, OKX)	-74.09, 40.77, -73.89, 40.97	-74.24, 40.62, -73.74, 41.12
Phoenix CSA (PHX)	North Phoenix-(NP) 040131004	40130019 40133002 40133003 40133010 40139997	Deer Valley Airport	Flagstaff (72376, FGZ)	-112.17, 33.46, -111.97, 33.66	-112.32, 33.31, -111.82,33.81
Salt Lake City CSA (SLC)	Hawthorne-(HAW) 490353006 Herriman (HERR) 490353013 Erda-490450004 Bountiful (Bountiful) 490110004	490353006	SLC Intl Airport	Salt Lake City (72572, SLC)	-111.97, 40.63, -111.77, 40.83	-112.12, 40.48, -111.62, 40.98
San Francisco CSA (SF)	Livermore-(LIV) 060010007	060010007	Livermore Munic. Airport	Oakland (72493, OAK)	-121.88, 37.59, -121.68, 37.79	-122.03, 37.44, -121.53, 37.94

3. Results

3a. Recent changes in O_3 and the role of NO_x

For most regions of the country, NO_x and O_3 concentrations have continued to go down, although the rate of reduction has slowed. Table 2 shows these changes for the 8 cities in this study. Nationally, NO_2 emissions have declined by 53% since 2006, the observed 1-hour daily max NO_2 for May-September has declined by 44% and the fourth highest MDA8 values have declined by 14% in the 40 U.S. metropolitan regions that exceed the standard (D.Jaffe, unpublished analysis of EPA data). Table 2 shows the changes for the 8 cities in this study for both 2006-2019 and 2010-2019. For some regions, like Salt Lake City, the fourth highest MDA8 has barely moved in the past decade. Figures S1 and S2 show these changes for each of the 8 metropolitan regions considered.

Table 2. Change in annual May-September 1-hour daily maximum NO_2 and annual fourth highest O_3 MDA8 for 2006-2019 and 2010-2019.

Metro region	% reduction in annual fourth highest O_3 MDA8 2006-2019.	% reduction in observed NO_2 2006-2019.	% reduction in annual fourth highest O_3 MDA8 2010-2019.	% reduction in observed NO_2 2010-2019.
Atlanta	18.5	35.8	6.3	21.9
Chicago	15.2	42.5	17.3	25.2
Dallas	25.3	49.3	14.5	29.3
Detroit	12.8	37.7	10.0	0.6
New York	12.0	40.1	7.6	20.5
Phoenix	14.1	35.1	7.6	26.9
Salt Lake City (Hawthorne)	11.0	50.3	0.0	30.7
San Francisco	19.1	46.8	2.7	32.0

But despite this slowing trend, Figures S3 and S4 in the appendix show that for most regions, there is an approximately linear relationship between the annual observed NO_2 and the annual fourth highest MDA8 in each city. Table 3 below shows the slope, intercept and R^2 for the linear fits between the observed annual May-September average 1-hour daily maximum NO_2 and the annual fourth highest MDA8. The slopes range from 0.3 to 2.3 ppb O_3 per ppb of NO_2. SLC and San Francisco have weaker correlations, a higher intercept and, for SLC, a much lower slope. These are all connected to the influence of smoke on O_3, which was especially strong in these two cities for 2016-2018, as described in a later section of this report.

Table 3. Parameters for linear fit between annual averaged (May-Sept) daily 1-hour max NO_2 concentration and fourth highest O_3 MDA8.

	Slope (ppb O_3 per ppb NO2)	**Intercept**	**R^2**
Atlanta	2.29	30.1	0.73
Chicago	0.9	49.7	0.47
Dallas	1.38	58.9	0.60
Detroit (East 7th site)	0.84	56.8	0.55
New York (two sites)	1.04	55.5	0.84
Phoenix	0.69	55.7	0.74
Salt Lake City (Hawthorne site)	0.32	67.7	0.28
San Francisco	0.92	62.3	0.36

I can evaluate the hypothesis that these regions are experiencing NO_x-limited chemistry, by examining the day of week pattern for O_3 exceedances. Figure S11 shows the daily pattern of average NO_2 concentrations and probability of an O_3 exceedance by day of week for each city. For all cities, the lowest observed NO_2 concentrations are on Sunday, with Saturday usually the second lowest. At the same time, Sunday has the lowest probability of an O_3 exceedance days (defined as a day with an MDA8>70 ppb) for 5 of the 8 regions considered (Atlanta, Dallas, NYC, Phoenix and SF). This pattern is not seen for Chicago, Detroit or SLC. At least for SLC, this is likely due to frequent smoke influence and enhanced background O_3 (discussed later in this report).

The relationships between annual NO_2 concentrations and the fourth highest MDA8 O_3 values shown in Table 3 suggest that we can use this relationship to estimate at what NO_2 concentration the O_3 standard would be reached. This is done by extrapolation of the linear function to an O_3 concentration of 70 ppb. The results are shown in Table 4. But importantly for some regions the results are not realistic. For example, SLC has a very low slope and the highest intercept. Both of these relate to the fact that SLC has some of the strongest influence from wildfire emissions and significant background O_3 influence, which acts to decouple the relationship between O_3 and local NO_2 concentrations.

I note that the NO_2 concentrations shown in Table 4 cover a wide range from 9.8 ppb (SF) to 25.4 ppb (Phoenix). We should not read too much into to the differences between cities. This is because each city uses a different number of NO_2 monitors and over different time periods. Because my goal was to evaluate the long term relationships, I chose the monitors in each region that gave the most consistent dataset over the timeframe of 2006-2018. In some cases, this includes only a single monitor (like Atlanta), whereas in other cities, such as Chicago, Phoenix and Detroit, multiple NO_2 datasets are available from around the metropolitan region.

But given that I found that the percent reductions across all monitors in each CSA were similar, the exact number of monitors used does not seem to have had an impact on this analysis.

Table 4. Estimated NO_x reductions needed to reach an annual fourth highest O_3 MDA8 of 70 ppb.

	2019 NO_2 (ppb)	Predicted NO_2 to meet O_3 standard (ppb)	% NO_2 reductions needed
Atlanta	21.3	17.4	18.2
Chicago	25.0	22.6	9.8
Dallas	12.0	8.0	33.0
Detroit (East 7th site)	23.0	15.7	31.3
New York (two sites)	18.9	13.9	26.2
Phoenix	25.4	20.7	18.4
Salt Lake City* (Hawthorne site)	17.8	7.2	59.6
San Francisco*	9.8	8.4	14.6

*****The calculated reductions for SF and SLC are influenced by recent wildfire smoke (see text).

3b. Role of temperature on O_3

It is well known that O_3 concentrations are higher on warmer days (Pusede et al 2015), although at extreme temperatures (above approximately 40°C), O_3 concentrations may be suppressed (Steiner et al 2010). Table 5 shows the Reduced Major Axis (RMA) slope for the MDA8 O_3-temperature relationship. In all cases the relationship is positive and statistically significant, although the R^2 values for some locations are rather weak. This is especially true for the warmer cities in our study; Phoenix, Dallas and Atlanta. For these locations, daily variations in daily maximum temperature have a weak relationship to variations in MDA8 values.

Table 5. Relationship between the MDA8 and daily max temperature for May-September data using Reduced Major Axis regression.

	RMA Slope (MDA8 ppb/°C)	R^2
ATL	1.39	0.20
CHI	0.80	0.35
DAL	1.09	0.13
DET-NEW	0.83	0.39
NYC-FAIR	1.01	0.42
PHX	0.74	0.07
SLC-Bountiful	0.49	0.32
SF	0.71	0.43

Figures S12 shows the annual 98th percentile daily max temperature (DMT) for each city along with the fourth highest MDA8 value. For some regions there is a clear enhancement and statistically significant relationship between the 98th percentile of DMT and the fourth highest MDA8 (Atlanta, Chicago, Detroit and SLC). For the other cities, no relationship is apparent at the seasonal level.

Given the role of temperature, it is important to examine whether systematic changes in temperature, due to climate change, may be making it more difficult to achieve the O_3 standard. Using the data shown in Figure S12 (98th percentile of DMT), we see no significant trends in any of the 8 cities considered. This is not surprising given the large year-to-year variability. Keep in mind that global climate change is currently pushing land surface temperatures higher by about 0.04°C per year, so over the 13 years considered in this study (2006-2018) this equates to a change of 0.52° C. (See for example the NASA-GISS temperature records at https://data.giss.nasa.gov/gistemp/.) Given that the year-to-year variability in these cities is much larger, the change due to global climate change is not apparent in these short records. So while temperature is an important control on the daily and inter-annual patterns of O_3, there is no evidence that current temperature trends are increasing the frequency of high O_3 days in these 8 cities at this time.

3c. Identification of smoke days

Since there are no clear markers for smoke in the EPA database, I have developed my own method to identify smoke days. This is based on a combination of the NOAA-HMS Fire and Smoke Product (hereafter simply HMS) and surface $PM_{2.5}$ data. This approach has been successfully used in a number of analyses concerning the smoke influence on O_3 (Gong et al 2017; Jaffe et al 2020; Jaffe 2020a). It is essential to use both the HMS satellite product and surface $PM_{2.5}$ data, as neither provides a clear signature of **surface** smoke by themselves.

A day is considered a smoke day if there was both identified overhead HMS smoke and the observed surface $PM_{2.5}$ was greater than a defined threshold. To identify the $PM_{2.5}$ threshold, first the data were segregated based on the HMS data. Then I calculate the monthly mean $PM_{2.5}$ for non-HMS days for all months in the study and determine if there is a significant trend with year (2006-2018), in the monthly means. If there is no trend, the $PM_{2.5}$ threshold is set to the average monthly mean plus one standard deviation of the daily values. If there is a significant trend, I apply a linear correlation to the monthly means $PM_{2.5}$ concentration vs year, and use the linear fit plus one standard deviation of the daily means as the threshold. Removing the impact of a trend is important since if this is not done and there is a significant trend due to local emission reductions, then we would be incorrectly identifying smoke days. The procedure described above sets a conservative threshold to use for identifying smoke days. It is possible that some smoke days will be missed if they have a $PM_{2.5}$ concentration that is enhanced, but not as high as the threshold. This is discussed further in section 3g. Table S1 shows the monthly means, trends and standard deviation of the daily data for each city.

3d. Generalized Additive Modeling (GAM) for each region

Generalized Additive Modeling (GAM) is powerful tool that can incorporate linear, non-linear and categorical variables (Wood 2017). I have previously used GAM results to predict the MDA8 O_3 concentrations based on daily variations in meteorology (Gong et al 2017; 2018; McClure and Jaffe 2018; Jaffe et al 2018) and this method was also successfully used in our 2019 CRC project A-118 (Jaffe 2020a). Typical GAM results demonstrate an ability to predict MDA8 O_3 values with an R^2 of between 0.5-0.8. This is a type of "machine learning" which uses a training dataset to understand the complex and inter-related patterns in the data. The difference between the GAM statistical prediction and the observed value is called the residual. For more background on GAM, please see our final report for CRC project A-118 (Jaffe 2020a) or our earlier scientific publications. For GAM, I use the "mgcv" package in R software.

GAM results were computed on the MDA8 for the primary O_3 season May-September, 2006-2018. For each site, I examined a large number of predictor variables for inclusion in the GAM (see Table 6 for a list of the variables considered). Each predictor was evaluated based on the degree to which it had explanatory power for the MDA8 and the degree to which it improved the overall model fit, without overfitting. To evaluate this, I used the Akaike information criterion (AIC), as described by Wood (2017). Lower values of the AIC represent a superior model. Individual predictors were also evaluated using the approximate p-value as computed by the mgcv program. Correlation between meteorological variables is acceptable and nearly impossible to avoid. For example, the daily maximum surface temperature and afternoon radiosonde temperature typically show a strong positive correlation. Inclusion of both variables in a GAM is acceptable and may improve the model performance, as each variable could potentially explain different aspects of the O_3 relationship. For every site, at least one variable related to daily max temperature was included in the GAM. On the other hand, inclusion of essentially duplicate variables (e.g. daily max temperature from two nearby stations) is not likely to improve model performance. If two nearly identical variables are included, one variable will typically show a high p-value and the AIC will increase, compared to inclusion of just one of these. Ultimately the goal is to find the optimum combination of the fewest variables that can explain the greatest amount of variance in the predicted variable.

Table 6. Variables considered for the GAM calculations

Variable name	Details	Source
Surf_NO2	See specific sites for each CSA in Table 1	AQS
Merra_tmax	See coordinates used for each CSA in Table 1	NASA Giovanni
Merra_tavg	See coordinates used for each CSA in Table 1	NASA Giovanni
OMI UV (filled)	OMI measured UV flux. Missing data are filled with long term monthly mean. See coordinates used for each CSA in Table 1.	NASA Giovanni

Variable name	Details	Source
OMI NO2 (filled)	OMI measured UV flux. Missing data are filled with monthly mean from same year. See coordinates used for each CSA in Table 1	NASA Giovanni
Surface Tmin	Local met data, see station i.d. in Table 1	NOAA-CDO
Surface Tmax	Local met data, see station i.d. in Table 1	NOAA-CDO
Traj distance	Point to point distance from 12-hour HYSPLIT back-trajectories with GDAS 1^o x 1^o met data.	NOAA Hysplit
Traj direction	Point to point direction from 12-hour HYSPLIT back-trajectories with GDAS 1^o x1^o met data.	NOAA Hysplit
850 Hpa temp-morning	From nearest radiosonde location (see Table 1). Morning sondes are launched at 12 GMT, which is between 4-7 am local standard time, depending on the time zone. Afternoon sondes are launched at 0 GMT, which is 4-7 pm local standard time, depending on the time zone.	University of Wyoming website
700 Hpa temp-morning	From nearest radiosonde location (see Table 1).	University of Wyoming website
1000 Hpa temp-morning	From nearest radiosonde location (see Table 1).	University of Wyoming website
WD-morning	Wind direction in lowest 1000 meters above sonde site. From nearest radiosonde location (see Table 1).	University of Wyoming website
WS-morning	Wind speed in lowest 1000 meters above sonde site. From nearest radiosonde location (see Table 1).	University of Wyoming website
RH-1000 -morning	Relative humidity (RH) in lowest 1000 meters above sonde site. From nearest radiosonde location (see Table 1).	University of Wyoming website
Surface mixing ratio-morning	Water vapor mixing ratio in lowest 1000 meters above sonde site. From nearest radiosonde location (see Table 1).	University of Wyoming website
Cape-morning	Calculated Convective Available Potential Energy (CAPE) from nearest radiosonde location (see Table 1).	University of Wyoming website
1000 Hpa temp-afternoon	Temperature in lowest 1000 meters above sonde site. From nearest radiosonde location (see Table 1).	University of Wyoming website
WD-afternoon	Wind direction in lowest 1000 meters above sonde site. From nearest radiosonde location (see Table 1).	University of Wyoming website
WS-afternoon	Wind speed in lowest 1000 meters above sonde site. From nearest radiosonde location (see Table 1).	University of Wyoming website

Variable name	Details	Source
RH-1000 Hpa-afternoon	RH in lowest 1000 meters above sonde site. From nearest radiosonde location (see Table 1).	University of Wyoming website
Surface mixing ratio-afternoon	Water vapor mixing ratio in lowest 1000 meters above sonde site. From nearest radiosonde location (see Table 1).	University of Wyoming website
Day of week	Calculated variable	

For fitting and testing the GAM, I used the following steps:

1. Using the non-smoke days, I randomly selected 90% as the training data and 10% to be used as the cross validation data.
2. The models were run and optimized using the training dataset.
3. The optimized GAM for each city was then rerun and used to predict both the cross-validation dataset and the smoke day dataset.

The variables selected for the GAM for each site are shown in Table 7 and Table S22 gives the GAM equations that were used in the R program.

Table 7. Variables included in each GAM by site.

	ATL	**CHI**	**DAL**	**DET**				**NYC**		**PHX**	**SLC**				**SF**
Site	UA	CHIW	GRP	EAST	PORT	OAK	NEW	LEO	FAIR	NP	HAW	HERR	ERDA	Bountiful	LIV
DOY	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Surf NO2*	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y			Y		Y
Merra_tmax		Y	Y					Y	Y					Y	Y
Merra_tavg		Y						Y							
OMI UV (filled)		Y		Y	Y	Y	Y			Y					
OMI NO2 (filled)	Y	Y	Y							Y	Y	Y	Y		
Surface Tmin														Y	Y
Surface Tmax	Y	Y		Y		Y	Y	Y	Y	Y		Y	Y		Y
Traj distance	Y		Y	Y		Y	Y	Y	Y	Y	Y	Y			Y
Traj direction		Y		Y	Y		Y	Y	Y		Y	Y	Y	Y	Y
850 Hpa temp-morning									Y						
700 Hpa temp-morning										Y				Y	
1000 Hpa temp-morning			Y							Y	Y			Y	
WD-morning	Y		Y	Y	Y	Y			Y	Y	Y				
WS-morning	Y									Y		Y			
RH-1000 Hpa-morning														Y	
Surface mixing ratio-morning	Y			Y	Y	Y	Y		Y		Y				
Cape-morning		Y													
1000 Hpa temp-afternoon				Y	Y						Y				Y

	ATL	CHI	DAL	DET				NYC		PHX	SLC				SF
WD-afternoon		Y	Y				Y		Y		Y				
WS-afternoon	Y			Y		Y		Y			Y	Y		Y	Y
RH-1000 Hpa-afternoon	Y		Y			Y		Y	Y		Y	Y		Y	Y
Surface mixing ratio-afternoon							Y				Y			Y	
Day of week	Y		Y			Y				Y	Y			Y	Y
# Variables in final model	11	10	10	10	7	10	9	9	11	11	13	7	5	11	11

*For the Detroit sites, I used an interaction term between surface NO_2 and year. This was found to give a slight improvement in the model predictions.

GAM works by fitting spline functions to the observations using the training dataset. Figure 1 shows an example of the spline fit for the Detroit New Haven site for back trajectory direction (0-360°). Figure 2 shows the observations for the MDA8 vs back trajectory direction. Since the New Haven site is northeast of downtown Detroit and north of the major emission sources in the region, it is not surprising that the MDA8 is highest when trajectories are from the southeast direction (210°). Figures S6 and S7 show similar plots of observed MDA8 vs trajectory distance and direction for all sites. This is a good example of one of the key advantages of GAM; incorporation of non-linear and categorical variables.

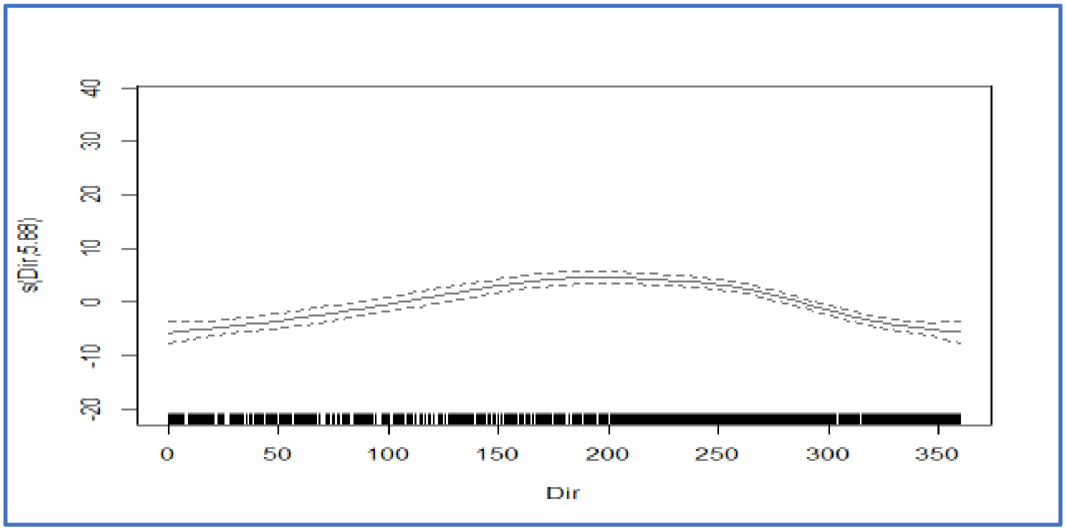

Figure 1. Spline fit from GAM for Trajectory direction vs MDA8 for DET New Haven site.

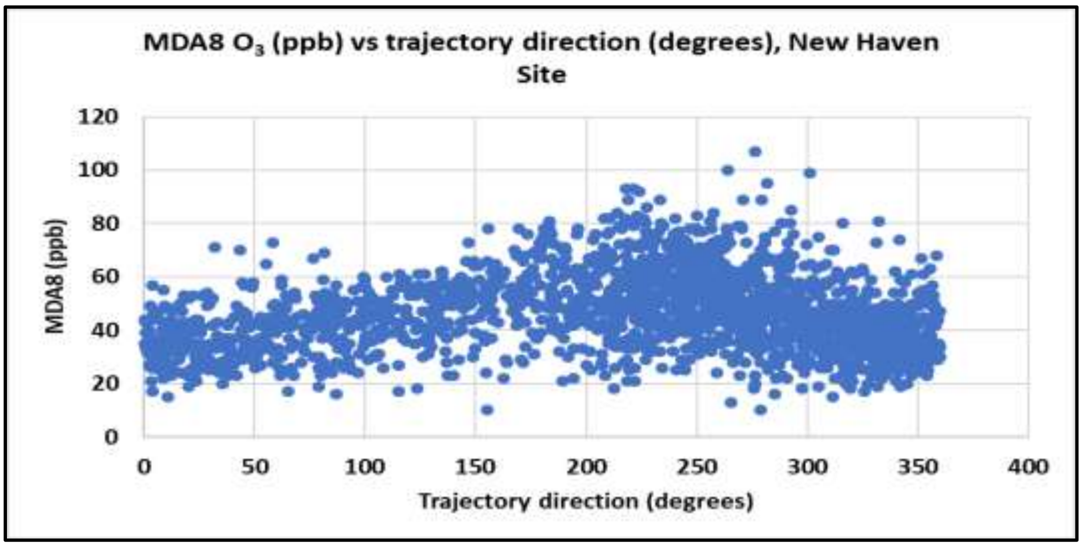

Figure 2. Observed MDA8 vs trajectory direction for DET New Haven site.

As with any model, it is important to evaluate the underlying assumptions and fidelity of the model predictions. To examine the quality of the model, I looked at a number of factors:

1. The residuals (observed value minus fit value) should be normally distributed and not exhibit heteroscedasticity (non-uniform variance relative to the predictor variables).
2. The residuals should not show bias with respect to any of the model predictors.
3. The cross-validation predictions should demonstrate similar skill at predictions as the training dataset and should have no bias with O_3 or any of the predictors.

I used the gam.check code in R software to check the QQ plots (sample quantiles against theoretical quantiles), scatterplots of the observed versus the predicted values and the residual versus the predicted values.

3e. GAM Results

Table 8 show the R^2 values for the correlation of the observed and predicted MDA8 for each site using the training, cross-validation and smoke datasets. The R^2 values range from 0.52-0.77 using the training datasets and are fairly similar using the cross-validation datasets. The R^2 for the smoke datasets tend to be significantly lower. This arises both because there are far fewer smoke cases (see Table S5) and because of the greater variability associated with the influence of smoke (see for example Jaffe and Wigder 2012).

The detailed GAM results for each site are shown in Figures S8-S10 and Tables S2-S6. Table S6 shows the residuals (observed – GAM predicted) by year and separated by training, cross-val and smoke datasets. Figures S8 and S9 show the residuals vs the predicted MDA8 values. For all sites, the training and cross-val datasets have a mean residual that is very close to zero, whereas the smoke datasets have a mean that ranges from 1.7 ppb (Phoenix) to 7.3 ppb (SF-Livermore). This residual is the average contribution to the MDA8 due to smoke. I conducted a t-test to compare the residuals for the cross validation and smoke datasets. In all cases, with the exception of Phoenix, the differences were significant with a P value of <0.01. For Phoenix, the average smoke influence is relatively small and the number of identified smoke days is low, partly due to the fact that $PM_{2.5}$ are not available until 2011.

Table 8. R^2 for observed vs GAM predicted MDA8 for each site.

CSA	Site	R^2-training	R^2-cross val	R^2-smoke days
Atlanta	United Ave	0.73	0.70	0.60
Chicago	Chiwaukee Prairie	0.68	0.68	0.49
Dallas	Grapevine Fairway	0.71	0.72	0.72
Detroit	East 7 Mile	0.74	0.73	0.52
	Port Huron	0.69	0.69	0.50
	Oak Park	0.76	0.75	0.54
	New Haven	0.76	0.73	0.54
NY/NJ/CT	Leonia	0.68	0.69	0.54
	Fairfield	0.72	0.72	0.46
Phoenix	North Phoenix	0.52	0.44	0.19
SLC	Hawthorne	0.65	0.63	0.23
	Herriman	0.59	0.50	0.35
	Erda	0.52	0.49	0.11
	Bountiful	0.62	0.64	0.52
SF	Livermore	0.77	0.75	0.55

3f. Influence of smoke

The daily max temperature is a good predictor for the MDA8 in 5 of the 8 cities (see Table 5). Figure 3 shows the relationship between MDA8 and surface tmax for the SLC Bountiful site. Table S3 shows that smoke days have substantially higher MDA8 values compared to non-smoke days at all sites, in some cases by up to 25 ppb. But in all cases, the average temperatures on smoke days are also higher. It is important to recognize that the GAM for each site incorporate some measure of daily high temperatures. This could be any one of the variables Merra_Tmax, Surface_tmax or the 1000 Hpa temp-afternoon. These three variables are highly correlated at all sites (e.g. R^2 of 0.8-0.9). Thus, the GAM residuals should show no bias with respect to daily maximum temperature. Figure 4 shows the GAM residuals vs temperature for the SLC Bountiful site, where the temperature is segregated into 5^{o}F bins. The mean residuals for each bin are all around 0 and the standard deviations are 6-7 ppb. This confirms that the GAM are adequately accounting for temperature and therefore the residual must be due to another factor.

Figure 5 compares the temperature-O_3 relationship for smoke and non-smoke days for the SLC Bountiful site. There are two points to be made with this figure. First, it is clear that smoke days have a higher daily max temperature, compared to non-smoke days. This pattern holds independently within each month (May-September). Second, it is apparent is that on smoke days, there is a stronger impact of temperature compared to non-smoke days. In other words, photochemical processes that generate O_3 are operating more efficiently on hotter smoke days, likely due to the greater abundance of O_3 precursors. Taken together, this analysis indicates that while smoke days are warmer than non-smoke days, the GAM results correctly account for temperature, so that the GAM residual is indicating the additional contribution to the MDA8 from smoke chemistry alone.

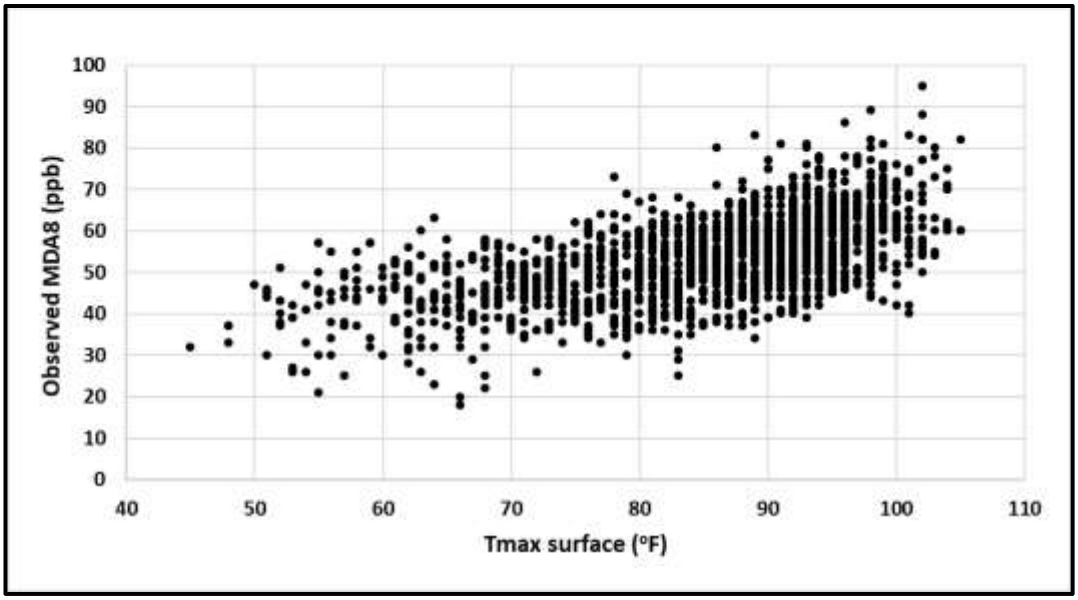

Figure 3. MDA8 O_3 vs daily max temperature for the SLC-Bountiful site for non-smoke cases.

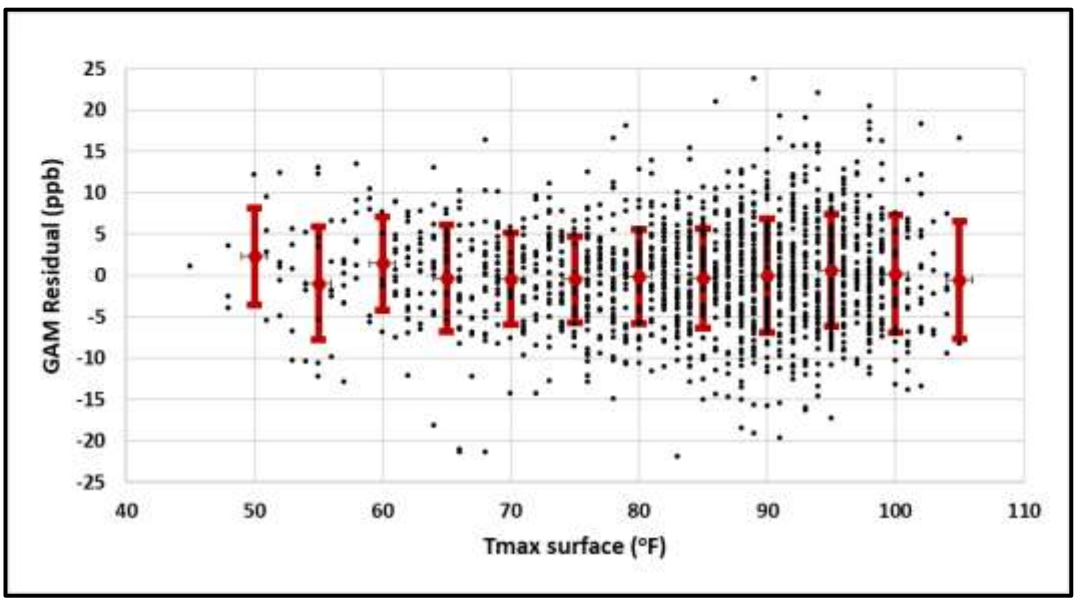

Figure 4. GAM residuals for SLC-Bountiful site for non-smoke cases. The black dots show the individual points and the red circles and bars show the mean and one s.d., respectively, segregated into 5°F bins.

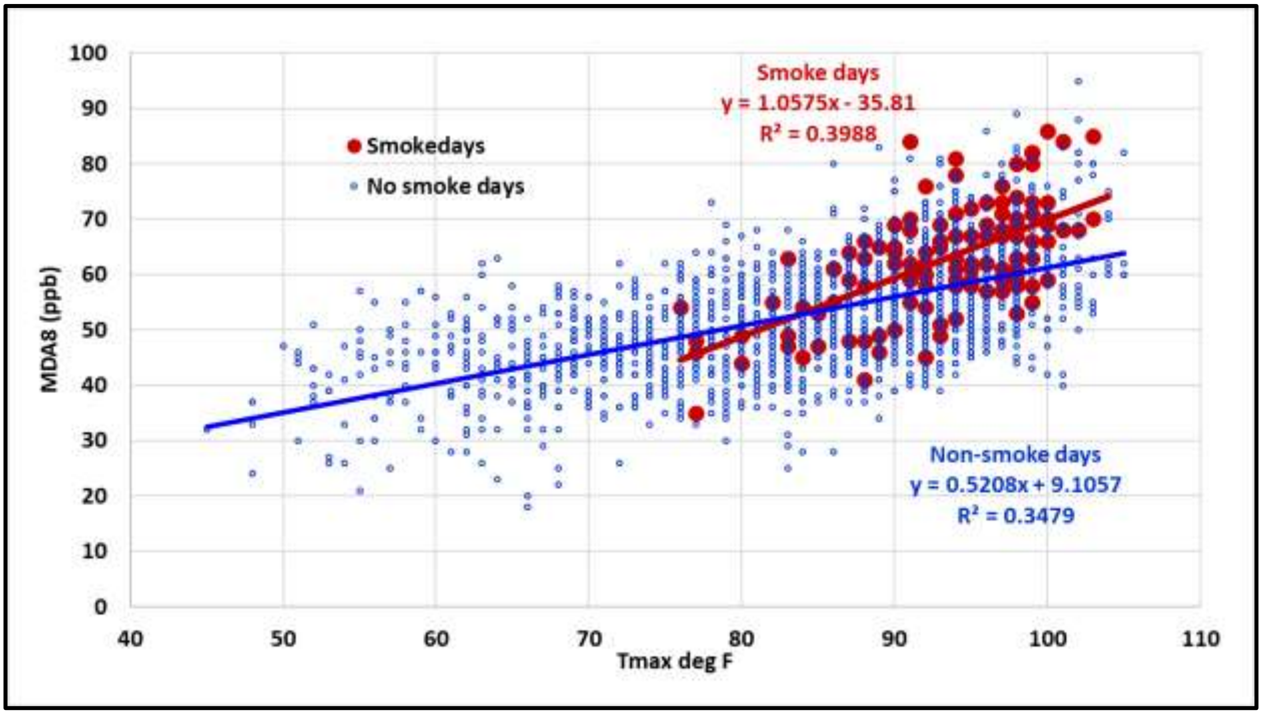

Figure 5. O_3 temperature relationship for smoke and non-smoke days.

3g. Estimated influence of smoke on design values

The ODV is calculated by averaging the fourth highest value from the three previous years in ppm and truncating the result to 3 decimals. Since smoke days have a higher probability of leading to an exceedance day (MDA8>70 ppb), these can significantly influence the ODV. For example, using the SLC Bountiful site data, I find that 16% of smoke days have an MDA8 >70 ppb, compared to 5% for non-smoke days. This is true even if I examine the data within each temperature bin. Smoke influenced days may be considered for exclusion from the design value calculation under the EPA's "Exceptional events" rule. However, the process is complex. To help guide air quality managers in this process, I examine the possible influence of smoke days on the 2016-2018 design values for each site considered in our study. I do this using several different assumptions:

1. I calculate the ODV for the 2016-2018 period using all available observations at each site.
2. I calculate the ODV using only non-smoke days, as defined by **both** the HMS overhead smoke data and the surface $PM_{2.5}$ data (above the threshold, as described in Section 3c).
3. I calculate the ODV using non-smoke days, as defined by **only** the HMS overhead smoke product, without regard to the surface $PM_{2.5}$ concentrations.

While this last definition for smoke is a looser definition for smoke days, it does provide some indication of the importance of smoke days and our ability to accurately identify these with the available data. Because I use a 24-hour average $PM_{2.5}$ concentration, it is possible that we have missed a smoke day with the more stringent requirement. For example, a day might have very low $PM_{2.5}$ during the evening and morning hours, and then have elevated $PM_{2.5}$ during the afternoon, when smoke moved into the region. Thus smoke might have still increased O_3 concentrations in the afternoon, even if the 24-hour daily $PM_{2.5}$ average concentration did not reach the threshold. The use of this "HMS only" smoke day definition allows us to identify days that warrant further investigation, if they turn out to be policy relevant.

Table 9 shows the ODV calculation for 2016-2018 and Figures S13 (and S14-S16 for additional sites) show how the annual fourth highest MDA8 changes under each assumption. Again, it is important to emphasize that this is only an example calculation. Exclusion of monitoring data as an exceptional event requires additional evaluation and EPA concurrence. But nonetheless, these calculations give an idea of the implications of the magnitude due to smoke influence on the ODV for each monitoring site.

Table 9. Summary of 2016-2018 O_3 design values using the existing data and using various hypothetical assumptions to exclude data.

Metro region	**Site**	**Avg # smoke days per year (2016-2018)**	**2016-2018 Design Value**	**2016-2018 Design Value if "smoke days" excluded**	**2016-2018 Design Value if all HMS days excluded**
ATL	UA	4.7	73	73	73
CHI	CHIW	4.7	79	79	77
DAL	GRP	4.0	76	76	73
DET	EAST	5.0	74	73	71
	PORT	5.0	72	68	68
	OAK	5.0	73	72	72
	NEW	5.0	72	71	71
NYC	LEO	5.0	76	74	74
	FAIR	5.0	84	80	80
PHX	NP	4.0	76	75	75
SLC	HAW	16.7	76	75	72
	HERR	16.7	77	74	72
	ERDA	16.7	74	72	70
	Bountiful	16.7	80	76	70
SF	LIV	18.3	73	67	64

Sites that show the strongest impact of smoke days on the ODV include Port Huron (DET), Fairfield (NYC), Herriman and Bountiful Valley (SLC), and Livermore (SF). In some cases, the ODV would drop below 70 ppb if these smoke days were excluded. SLC and SF were particularly hard hit by smoke days in 2016-2018 with an average number of smoke days per year of ~17-18. As such, air quality managers in these regions especially, should consider additional work to identify smoke days and possible exceptional event designations.

4. Summary for each city

4a. Atlanta

Atlanta data show a good relationship between MDA8 and the observed NO_2 concentration. The R^2 value for the correlation between the annual fourth highest MDA8 and annual averaged May-September daily maximum NO_2 is 0.73 and the correlation for the daily data is 0.43. The probability of an exceedance day and NO_2 concentrations are both lowest on weekend days (Figure S11). All of these are consistent with a NOx-limited regime, where further reductions in NOx will lead to significant lowering of the highest MDA8 values. Based on the 2019 May-September mean daily maximum NO_2 value of 21.3 ppb, I estimate that an 18% reduction in NO_2 concentrations will bring Atlanta into compliance with the O_3 NAAQS (see Table 4 and Figure S4).

The GAM results for Atlanta are good, with an R^2 value of 0.73 for the correlation of the observed and GAM predicted MDA8. The average GAM residual on smoke days is 5.0 ppb, suggesting a contribution of this amount to the MDA8 on smoke days (on average). However, the contribution to the 2016-2018 ODV due to smoke in Atlanta appears to be rather minimal (Table 9 and Figure S13).

4b. Chicago

Chicago data show a moderate relationship between MDA8 and the observed NO_2 concentration. The R^2 value for the correlation between the annual fourth highest MDA8 and annual averaged May-September daily maximum NO_2 is 0.47 and the correlation for the daily data is 0.19. While the NO_2 concentrations are lowest on the weekends, the probability of an exceedance day does not follow this pattern (Figure S11). Thus, at least as of the 2013-2018 time period, Chicago does not appear to be in a clear NOx-limited regime. However, they are probably very close, so it is likely that as NO_x is further reduced, this will occur. Based on the 2019 May-September mean daily maximum NO_2 value of 25.0 ppb, I estimate that a 10% reduction in NO_2 concentrations will bring Chicago into compliance with the O_3 NAAQS (see Table 4 and Figure S4).

The GAM results for Chicago are good, with an R^2 value of 0.68 for the correlation of the observed and GAM predicted MDA8. The average GAM residual on smoke days is 4.8 ppb, suggesting a contribution of this amount to the MDA8 on smoke days (on average). I estimate the contribution to the 2016-2018 ODV due to smoke in Chicago is 0-2 ppb. (Table 9 and Figure S13).

4c. Dallas

Dallas data show a moderate relationship between MDA8 and the observed NO_2 concentration. The R^2 value for the correlation between the annual fourth highest MDA8 and annual averaged May-September daily maximum NO_2 is 0.60 and the correlation for the daily data is 0.35. The probability of an exceedance day and NO_2 concentrations are both lowest on weekend days (Figure S11). All of these are consistent with a NOx-limited regime, where further reductions in NOx will lead to significant lowering of the highest MDA8 values. Based on the 2019 May-September mean daily maximum NO_2 value of 12.0 ppb, I estimate that a 33% reduction in NO_2 concentrations will bring Dallas into compliance with the O_3 NAAQS (see Table 4 and Figure S4).

The GAM results for Dallas are good, with an R^2 value of 0.71 for the correlation of the observed and GAM predicted MDA8. The average GAM residual on smoke days is 5.6 ppb, suggesting a contribution of this amount to the MDA8 on smoke days (on average). I estimate the contribution to the 2016-2018 ODV due to smoke in Dallas is 0-3 ppb. (Table 9 and Figure S13).

4d. Detroit

Detroit data show a moderate relationship between MDA8 and the observed NO_2 concentration. The R^2 value for the correlation between the annual fourth highest MDA8 and annual averaged May-September daily maximum NO_2 is 0.59 and the correlation for the daily data is 0.21 (both for the East 7 Mile site). While the NO_2 concentrations are lowest on the weekends, the probability of an exceedance day does not follow this pattern (Figure S11). Thus, at least as of the 2013-2018 time period, Detroit does not appear to be in a clear NOx-limited regime. However, they are probably very close, so it is likely that as NO_x is further reduced, this will occur. Based on the 2019 May-September mean daily maximum NO_2 value of 23.0 ppb, I estimate that a 31% reduction in NO_2 concentrations will bring Detroit into compliance with the O_3 NAAQS (see Table 4 and Figure S4).

The GAM results for Detroit are good, with an R^2 value of 0.76 for the correlation of the observed and GAM predicted MDA8 (New Haven site). The average GAM residual on smoke days is very similar at all sites, ranging from 5.1-5.7 ppb, suggesting a contribution of this amount to the MDA8 on smoke days (on average). I estimate the contribution to the 2016-2018 ODV due to smoke in Detroit is 1-4 ppb, depending on the site and assumptions made (Table 9 and Figure S13).

4e. New York City

NYC data show a good relationship between MDA8 and the observed NO_2 concentration. The R^2 value for the correlation between the annual fourth highest MDA8 and annual averaged May-September daily maximum NO_2 is 0.84 (Fairfield site) and the correlation for the daily data is 0.17. The probability of an exceedance day and NO_2 concentrations are both lowest on Sunday (Figure S11). All of these are consistent with a NOx-limited regime, where further reductions in NOx will lead to significant lowering of the highest MDA8 values. Based on the 2019 May-September mean daily maximum NO_2 value of 18.9 ppb, I estimate that an 26% reduction in NO_2 concentrations will bring NYC into compliance with the O_3 NAAQS (see Table 4 and Figure S4).

The GAM results for NYC are good, with an R^2 value of 0.72 (Fairfield site) for the correlation of the observed and GAM predicted MDA8. The average GAM residual on smoke days is 5.6 ppb, suggesting a contribution of this amount to the MDA8 on smoke days (on average). The contribution to the 2016-2018 ODV due to smoke in NYC are modest at around 2-4 ppb (Table 9 and Figure S13).

4f. Phoenix

Phoenix data show a good relationship between MDA8 and the observed NO_2 concentration. The R^2 value for the correlation between the annual fourth highest MDA8 and annual averaged May-September daily maximum NO_2 is 0.74 and the correlation for the daily data is 0.15. The probability of an exceedance day and NO_2 concentrations are both lowest on weekends (Figure S11). All of these are consistent with a NOx-limited regime, where further reductions in NOx will lead to significant lowering of the highest MDA8 values. Based on the 2019 May-September mean daily maximum NO_2 value of 25.4 ppb, I estimate that an 18.4% reduction in NO_2 concentrations will bring PHOENIX into compliance with the O_3 NAAQS (see Table 4 and Figure S4).

The GAM results for PHOENIX are good, with an R^2 value of 0.52 for the correlation of the observed and GAM predicted MDA8. The average GAM residual on smoke days is 1.7 ppb, suggesting a contribution of this amount to the MDA8 on smoke days (on average). The contribution to the 2016-2018 ODV due to smoke in PHOENIX is very modest at around 1 ppb (Table 9 and Figure S13).

Phoenix and Salt Lake City also appears to be strongly influenced by background O_3, where background O_3 is defined as any source that is not from local or regional photochemical production. While background O_3 is present everywhere, certain regions are exposed to higher concentrations of background due to elevation, wildfire smoke, long-range transport of pollutants from domestic or international sources, or due to influence from upper tropospheric/lower stratospheric (UTLS) O_3. Langford et al (2017) describe several events of strong background O_3 influence in the southwestern U.S. in May 2013. These events led to elevated O_3 (MDA8 >70 ppb) at several monitoring sites in Clark County, NV. Using the NOAA Geophysical Fluid Dynamics Laboratory AM3 model, they further show that these events typically incorporate both stratospheric air and O_3 from anthropogenic sources in Asia. In particular, the southwestern U.S. is most strongly influenced by these events due to the persistence of deep convective boundary layers. The AM3 model, suggests that the combined background influence from the UTLS and Asian sources can add up to 25 ppb to the monthly mean O_3 concentrations in the southwestern U.S. This results in what is, probably, the highest background O_3 (non-local) concentration seen anywhere in the U.S. Based on the AM3 model results (shown in Langford et al 2017), both Phoenix and SLC are strongly influenced by these sources. Phoenix, and to a lesser extent SLC, may also be influenced by O_3 from non-local domestic sources, particularly emissions from California.

Figure 6 shows the MDA8 distribution by decile cut points (10-evenly spaced groups) for May-September, 2015-2018. Both Phoenix and SLC have very high MDA8 values for the lowest decile. In other words, in the cleanest air, these regions have MDA8 values that are 10-15 higher than any other location in our study. This is consistent with the strong influence due to background O_3 for these locations as described in Langford et al (2017). Jaffe et al (2020b) show two examples of how background O_3 impacts high O_3 days for sites in EPA regions 2 (NY and NJ) and Region 8 (CO, MT, ND, SD, UT, WY). What is clear from Figure 6 and these prior published analyses is that sites in the southwestern U.S. and the intermountain west, experience a much higher level of background O_3, compared to most other parts of the U.S. Local photochemical production of O_3 is then added to this higher background level.

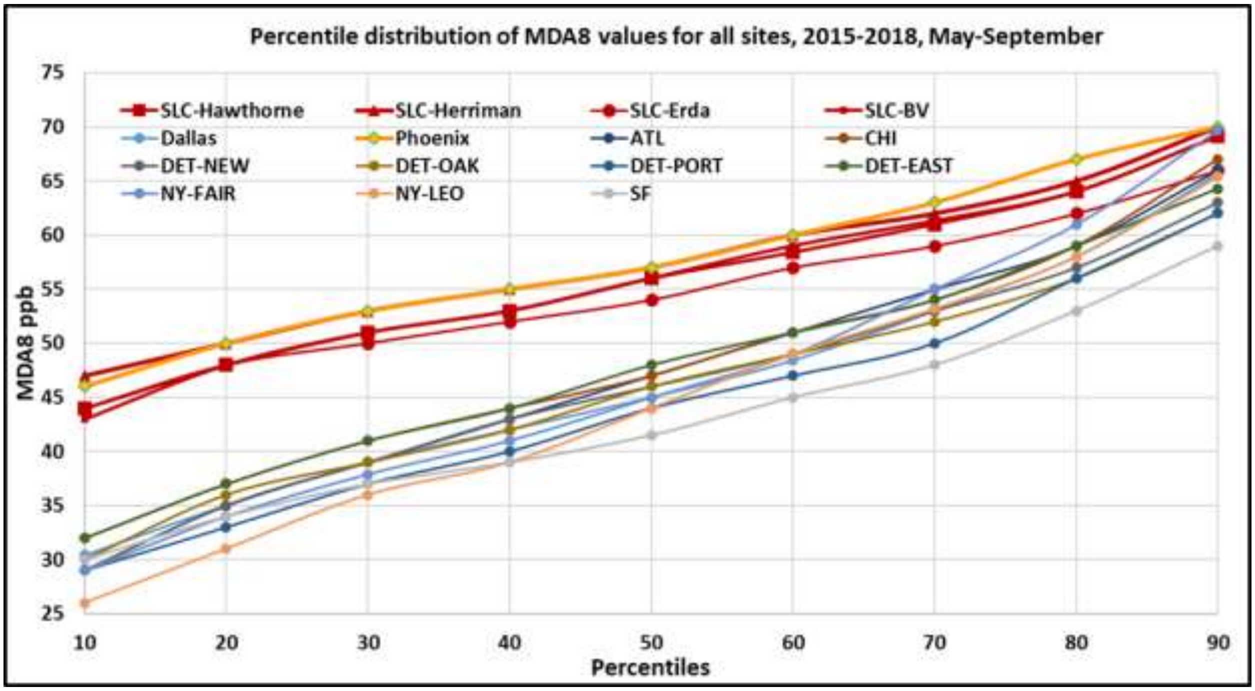

Figure 6. Distribution of MDA8 values (ppb) by deciles for all sites in this study. The y-axis shows the MDA8 for the decile indicated on the x-axis. This plot uses MDA8 data for May-September, 2015-2018.

4g. Salt Lake City

SLC data show a poor relationship between MDA8 and the observed NO_2 concentration. The R^2 value for the correlation between the annual fourth highest MDA8 and annual averaged May-September daily maximum NO_2 is 0.28 and the correlation for the daily data is 0.04. The NO_2 concentrations are lowest on weekends, but the probability of an exceedance day is not clearly lower on weekends (Figure S11). This raises questions about whether SLC is NO_x limited or not. Nonetheless, I did estimate the NO_2 concentration that would bring SLC into compliance with the O_3 standard using the same approach as the other cities. Based on the 2019 May-September mean daily maximum NO_2 value of 17.8 ppb, I estimate that a 59.6% reduction in NO_2 concentrations will bring SLC into compliance with the O_3 NAAQS (see Table 4 and Figure S4). However this value should be viewed with some skepticism due to a number of uncertainties (described below).

The GAM results for SLC are good, with an R^2 value of 0.65 (Hawthorne site) for the correlation of the observed and GAM predicted MDA8. The average GAM residual on smoke days are 3.2-5.6 ppb for the four sites, suggesting contributions of this amount to the MDA8 on smoke days (on average). The contribution to the 2016-2018 ODV due to smoke in SLC is significant and ranges from 2-10 ppb, depending on the site and assumptions (Table 9 and Figure S13).

As described above, Phoenix and Salt Lake City also appear to be strongly influenced by background O_3, where background O_3 is defined as any source that is not from local or regional photochemical production. See the discussion in the previous section (Phoenix) for more information and evidence concerning the strong background O_3 influence in SLC.

Data from SLC are unique among the 8 cities considered due to:

1. The weak observed relationship between O_3 and NO_2.

2. The strong influence from smoke on the ODV in SLC.
3. The strong influence from background O_3 in SLC.

While it is expected that further reductions in NO_2 would decrease O_3 in SLC, the combined contributions from background O_3 and smoke mean that this region must overcome significant challenges to meet the standard. Given the ambiguity on the role of NO_x in SLC, further research on photochemistry in the region, both smoke and non-smoke conditions, would provide a better understanding of the conditions that will be required to achieve the O_3 standard.

4h. San Francisco

San Francisco data show a modest relationship between MDA8 and the observed NO_2 concentration. The R^2 value for the correlation between the annual fourth highest MDA8 and annual averaged May-September daily maximum NO_2 is 0.36 and the correlation for the daily data is 0.35. The probability of an exceedance day and NO_2 concentrations are both lowest on Sunday (Figure S11). All of these are consistent with a NOx-limited regime, where further reductions in NOx will lead to significant lowering of the highest MDA8 values. Based on the 2019 May-September mean daily maximum NO_2 value of 9.8 ppb, I estimate that an 14.6% reduction in NO_2 concentrations will bring SF into compliance with the O_3 NAAQS (see Table 4 and Figure S4).

The GAM results for SF are good, with an R^2 value of 0.77 for the correlation of the observed and GAM predicted MDA8. The average GAM residual on smoke days is 7.3 ppb, suggesting a contribution of this amount to the MDA8 on smoke days (on average). The contribution to the 2016-2018 ODV due to smoke in SF is significant at around 6-9 ppb (Table 9 and Figure S13).

5. References

Gong X., Kaulfus A., Nair U. and Jaffe D.A. Quantifying O_3 impacts in urban areas due to wildfires using a Generalized Additive Model. Envir. Sci. Tech. DOI: 10.1021/acs.est.7b03130, 2017.

Gong X., Hong S. Jaffe D.A. Ozone in China: Spatial distribution and leading meteorological factors controlling O_3 in 16 Chinese cities. Aer. Air Qual. Res, DOI: 10.4209/aaqr.2017.10.0368, 2018.

Jaffe D. Role of Meteorology, Emissions and Smoke on Ozone in the South Coast Air Basin. Final project report for CRC Project A-118, Coordinating Research Council, Alpharetta, GA, January 2020. Available at: http://crcao.org/wp-content/uploads/2020/01/CRC-Project-A-118-Final-Report_Jan2020.pdf, 2020a.

Jaffe D.A., Fiore A.M. and Keating, T.J., Importance of Background O_3 for Air Quality Management. EM November 2020, 2020b.

Jaffe D.A., O'Neill S.M., Larkin N.K., Holder A.L, Peterson D.L., Halofsky J.E. and Rappold A.G. Wildfire and prescribed burning impacts on air quality in the United States, J. Air and Waste Mgt. Assn., DOI: 10.1080/10962247.2020.1749731, 2020.

Jaffe DA, Cooper OR, Fiore AM, Henderson BH, Tonneson GS, Russell AG, Henze DK, Langford AO, Lin M and Moore T. Scientific assessment of background ozone over the U.S.: Implications for air quality management. Elem Sci Anth. 2018;6(1):56. doi:http://doi.org/10.1525/elementa.309, 2018.

Jaffe, D.A. and Wigder, N.L., Ozone production from wildfires: A critical review. Atmos, Envir., doi:10.1016/j.atmosenv.2011.11.063, 2012.

Langford, A. O., et al. (2016), Entrainment of stratospheric air and Asian pollution by the convective boundary layer in the southwestern U.S., J. Geophys. Res. Atmos., 122, doi:10.1002/ 2016JD025987.

McClure C.D. and Jaffe D.A. Investigation of High Ozone Events due to Wildfire Smoke in an Urban Area. Atmos. Envir. https://doi.org/10.1016/j.atmosenv.2018.09.021, 2018.

Nussbaumer C.M. and Cohen R.C. The Role of Temperature and NOx in Ozone Trends in the Los Angeles Basin. Envir. Sci.Tech. 54 (24), 15652-15659 DOI: 10.1021/acs.est.0c04910, 2020.

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. *Environ Sci Technol* 49(1): 186–195. DOI: https://doi. org/10.1021/es504514z, 2015.

Wood, S.N. Generalized Additive Models, An Introduction with R, Second Edition. Chapman and Hall/CRC, Boca Raton, FL 2017.

CRC-A-124: Appendix

In this appendix, I show a number of common graphs and tables for each region and each site considered in the project. For ease of viewing, the tables and graphs all use a consistent formatting and are organized by geographic region. Note that most plots include data for 2006-2018, except for Figures S3 and S4, which use data for 1995-2019.

Figures

S1. Trend in 4th highest O_3 for each site.

This plot only shows the data from O_3 monitors considered in our analysis.

S2. Trend in NO_2 surface daily max and OMI column, 2006-2018.

This plot compares the May-Sept.NO_2 trends as observed by the surface monitors and the OMI satellite. One plot per CSA.

S3. Trend in 4th highest O_3 and daily max NO_2, 1995-2019.

This plot compares the May-Sept surface NO_2 trends with the annual 4th highest MDA8 O_3

S4. O_3 vs NO_2, 1995-2019.

This plot shows an X-Y scatterplot of the data in the previous figure.

S5. MDA8 vs daily max NO_2 using May-Sept data for one site in the CSA.

This plot shows how the O_3 MDA8 responds to the daily max surface NO_2 value for one site in the CSA.

S6. MDA8 vs trajectory direction for one site in the CSA.

This plot shows the relationship of the MDA8 to daily trajectory direction.

S7. MDA8 vs trajectory distance for one site in the CSA.

This plot shows the relationship of the MDA8 to daily trajectory distance.

S8. GAM predicted vs Observed MDA8 using the training dataset only.

S9. GAM residuals vs the predicted MDA8 using the training dataset (blue points) and the cross-validation dataset (red points). *Statistics on the residuals are shown in Table S6.*

S10. GAM residuals vs predicted, training dataset (blue points) and smoke dataset (red points). *Statistics on the residuals are shown in Table S6.*

S11. Probability (P) of an O_3 exceedance day (MDA8>70 ppb) across all sites that we considered and averaged surface NO_2 concentrations both as a function of day of week (Sunday=1, Monday =2, etc). Both are calculated using data from May-September 2013-2018 data. The probabilities are computed by summing the number of exceedance days for all sites and dividing by the number of observations.

S12. Annual 98th percentile of daily maximum temperature (DMT, °F) and fourth highest MDA8.

S13. Annual fourth highest for all data (blue), smoke days excluded (red) and HMS days excluded (black).

This figure shows how the annual fourth highest MDA8 would change with different assumptions. For this analysis a "smoke day" requires that there be both overhead HMS smoke <u>and</u> elevated $PM_{2.5}$. Thus the smoke day criteria is a more stringent requirement than excluding HMS smoke days alone. As such, the "HMS days excluded fourth highest" will always be similar or lower than the other fourth highest values. It's important to keep in mind that EPA would require a rigorous evaluation to designate a day as an exceptional event. But nonetheless, this analysis gives us some idea as to how important that might be for each CSA.

Figures S14-S16: Same as Figure S12, but for other sites that were considered in each CSA.

Tables:

S1. Average (µg m^{-3}), trend (slope, µg m^{-3}-yr^{-1}) and standard deviation (µg m^{-3}) of daily $PM_{2.5}$ by month for days without overhead HMS smoke for 2006-2018. If the trend is not statistically significant, the slope is set to zero. The 2016-2018 monthly mean $PM_{2.5}$ concentration plus one standard deviation of the daily means or the linear fit to the trend in $PM_{2.5}$ (if the trend is significant) plus one standard deviation of the daily values, is used as a criterion to determine if a day qualifies as a smoke day.

S2. Number of days in each category (training, cross-validation, smoke).

S3. Average MDA8 (ppb) in each category(training, cross-validation, smoke).

S4. Averages surface daily maximum temperature (^{o}F) in each category (training, cross-validation, smoke).

S5. Number of days in each category for the GAM analysis (training, cross-validation, smoke).

S6. Average residual from the GAM analysis (training, cross-validation, smoke). Note that some days can not be included due to some missing data.

S7-S21: *Same as Tables S2-6, but for each additional site in the CSA.*

S22. R code GAM equations for each site.

Table of contents for appendix

Atlanta CSA

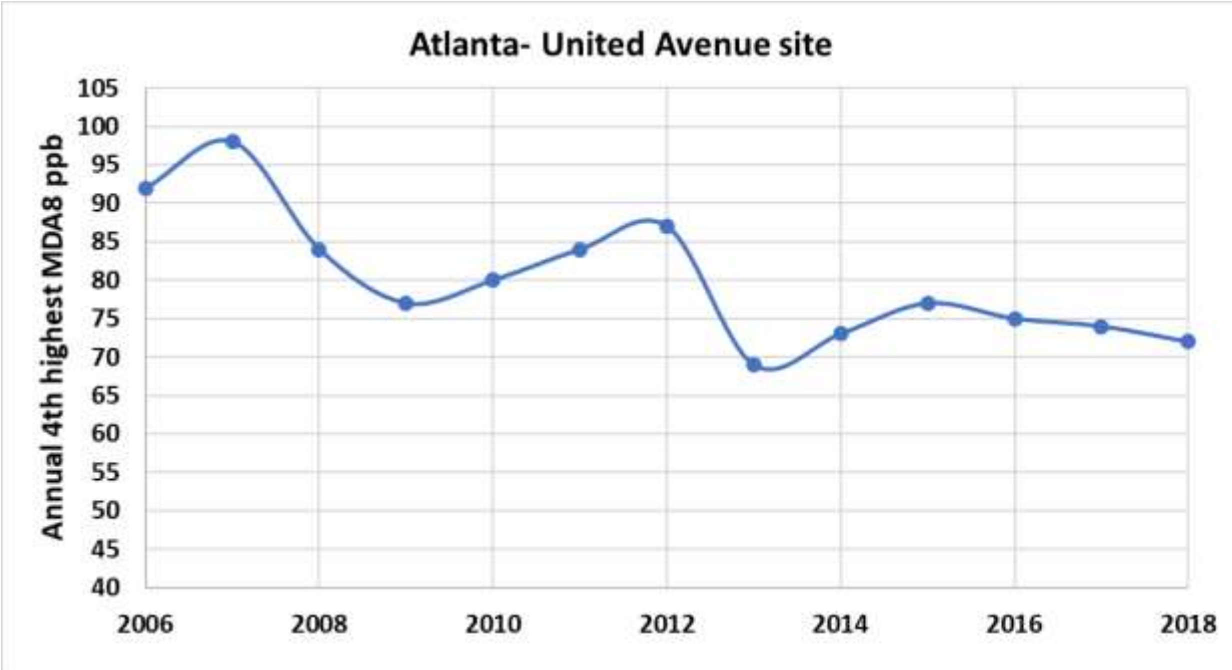

Figure S1. Trend in 4th highest MDA8 O_3 for each site.

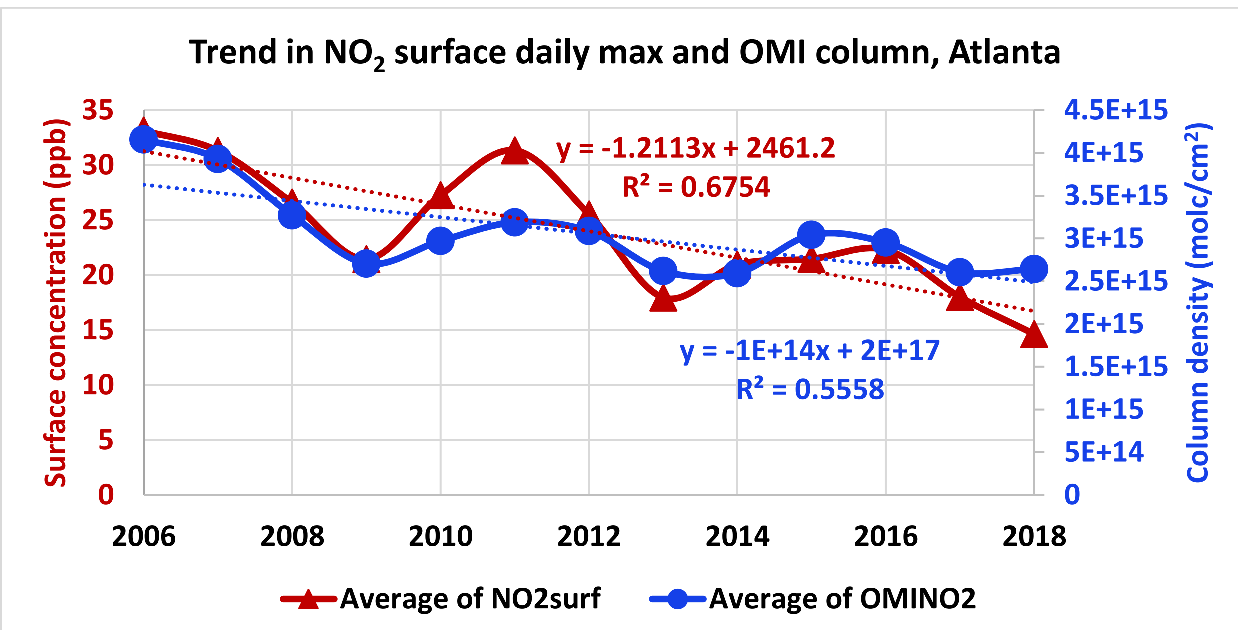

Figure S2. Trend in surface NO_2 daily max and OMI column, 2006-2018.

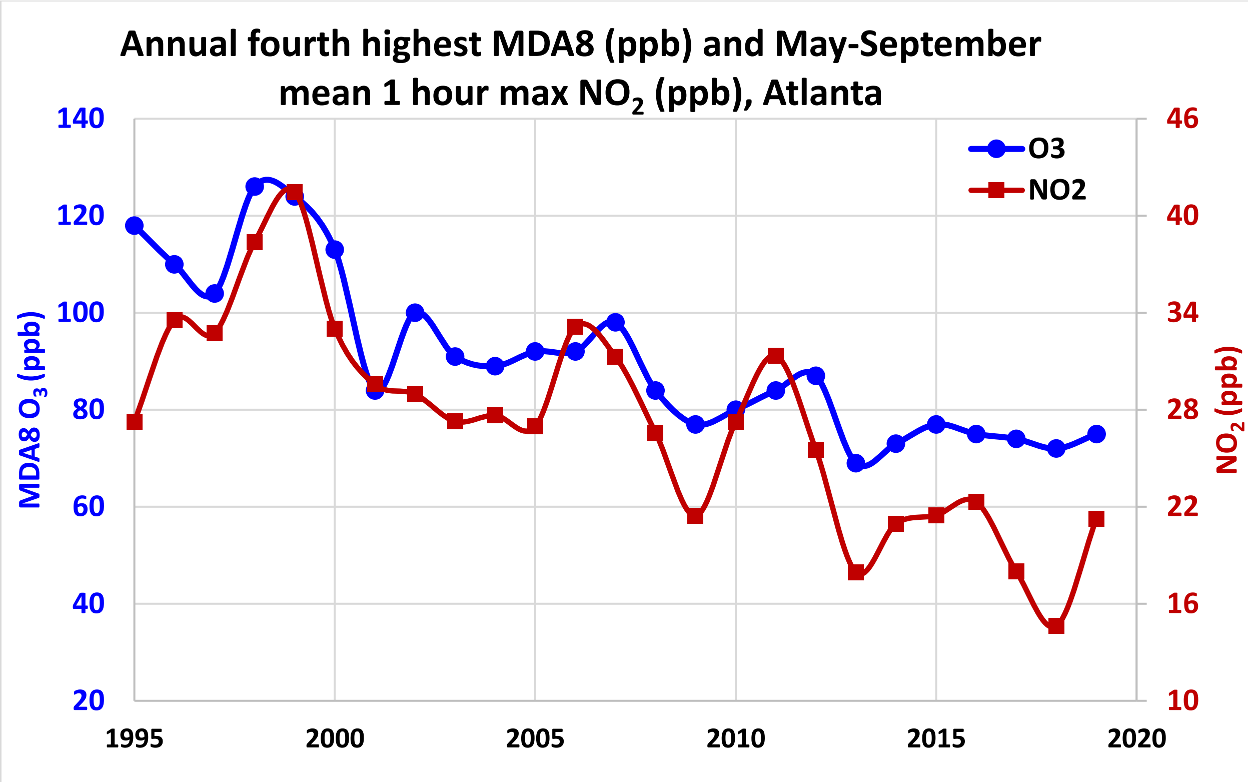

Figure S3. Trend in 4th highest O_3 and daily max NO_2, 1995-2019. For this plot and the next, we use NO_2 data from site 130890002.

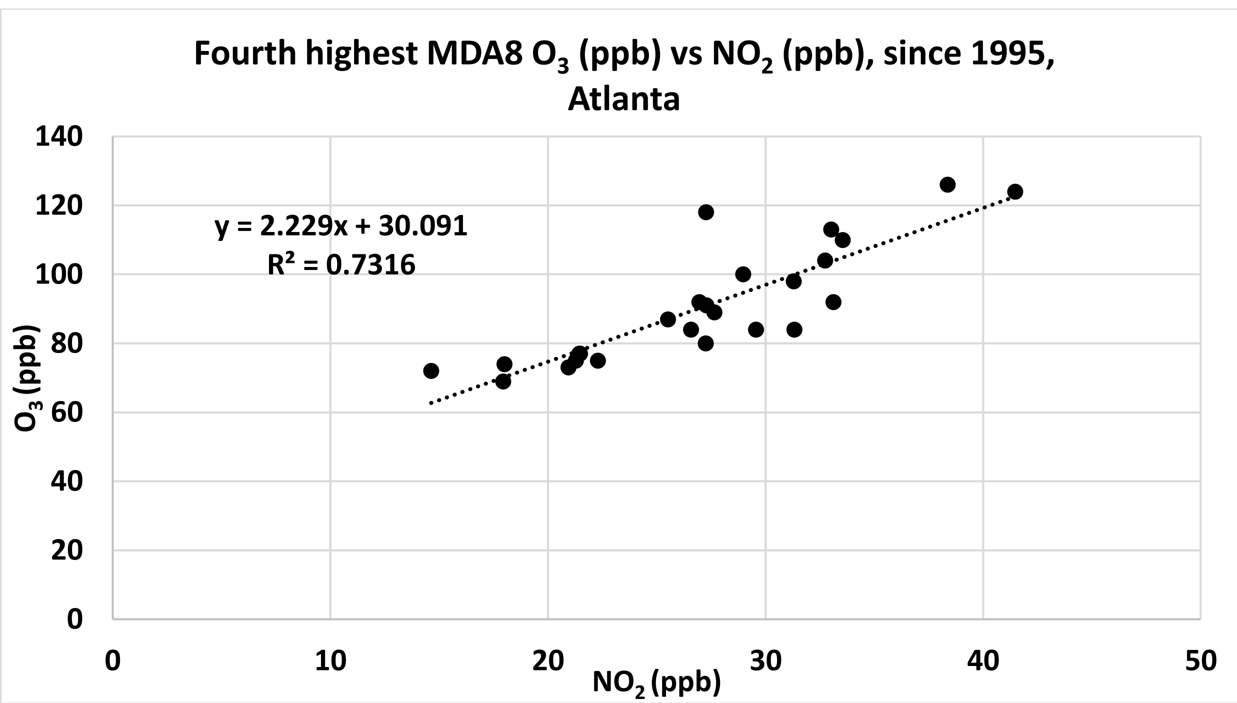

Figure S4. O_3 vs NO_2, 1995-2019.

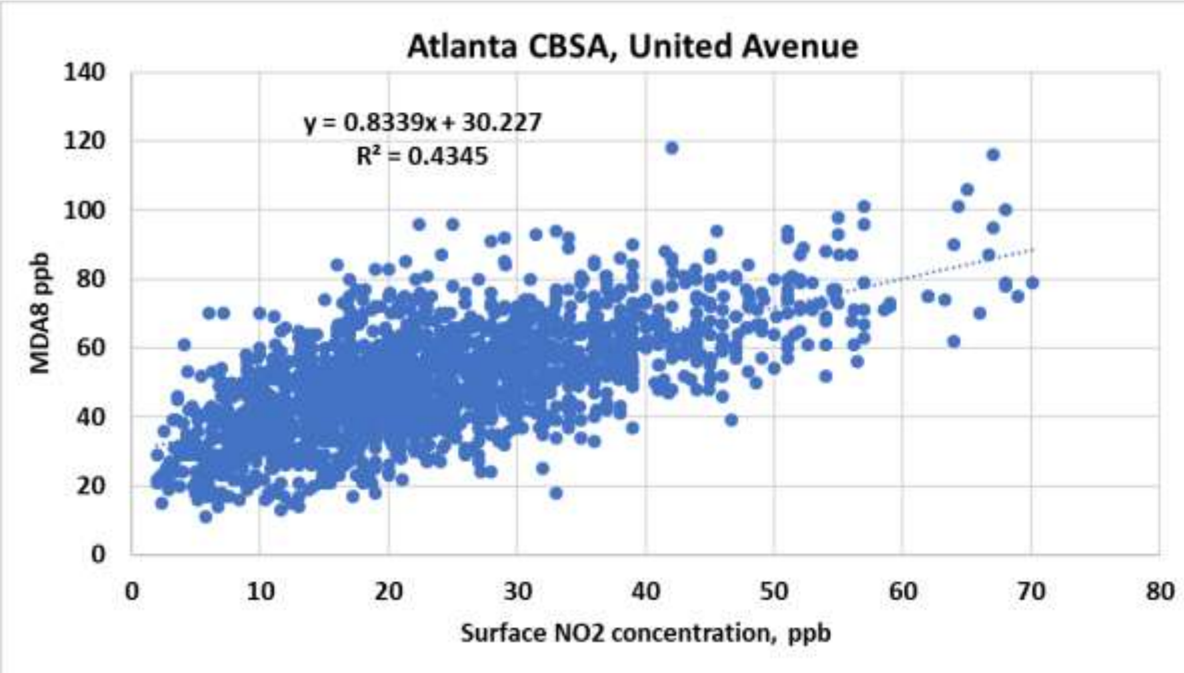

Figure S5. MDA8 vs daily max NO_2 for United Avenue site.

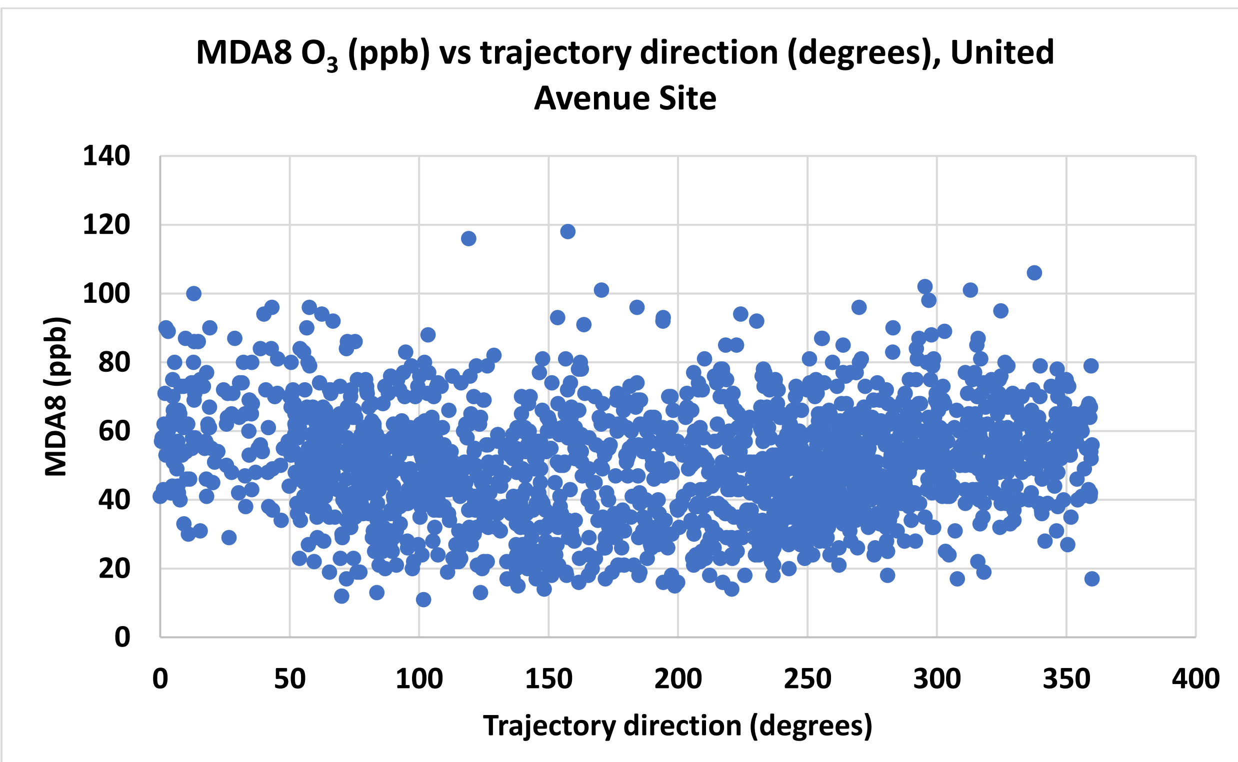

Figure S6. MDA8 vs trajectory direction for United Avenue site.

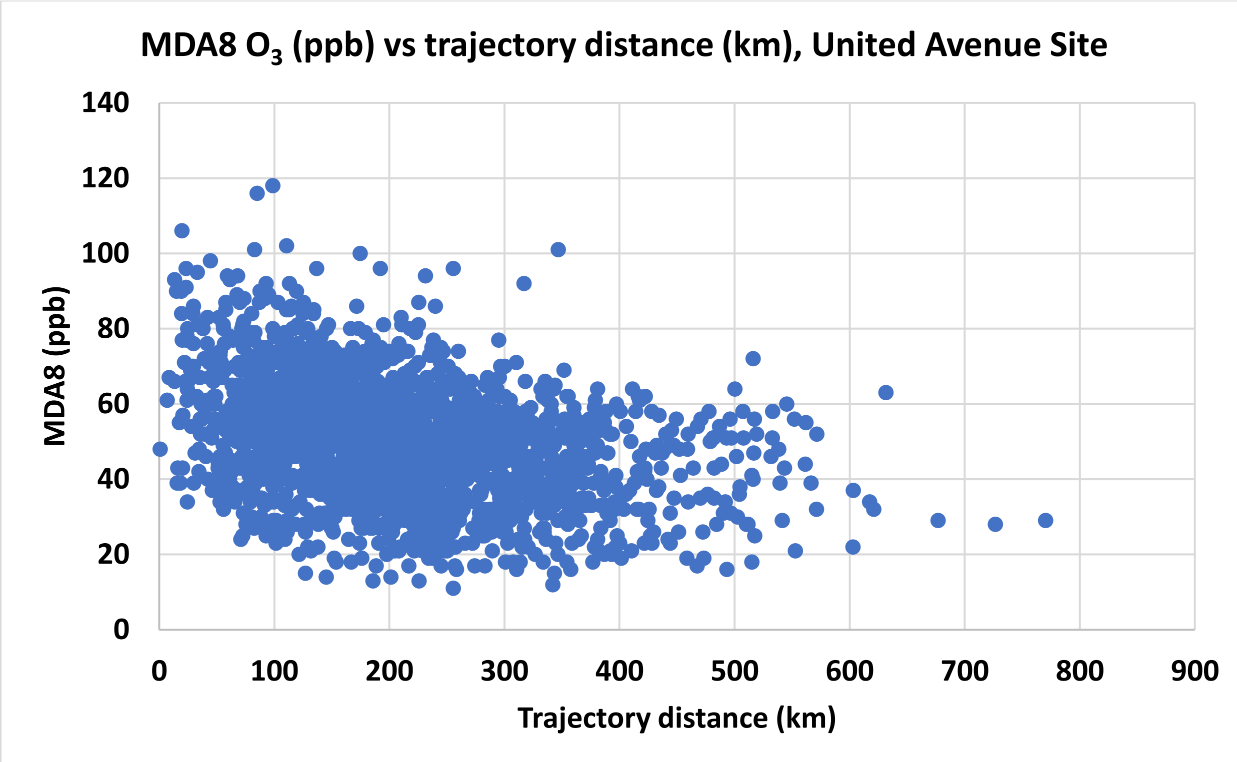

Figure S7. MDA8 vs trajectory distance for United Avenue site.

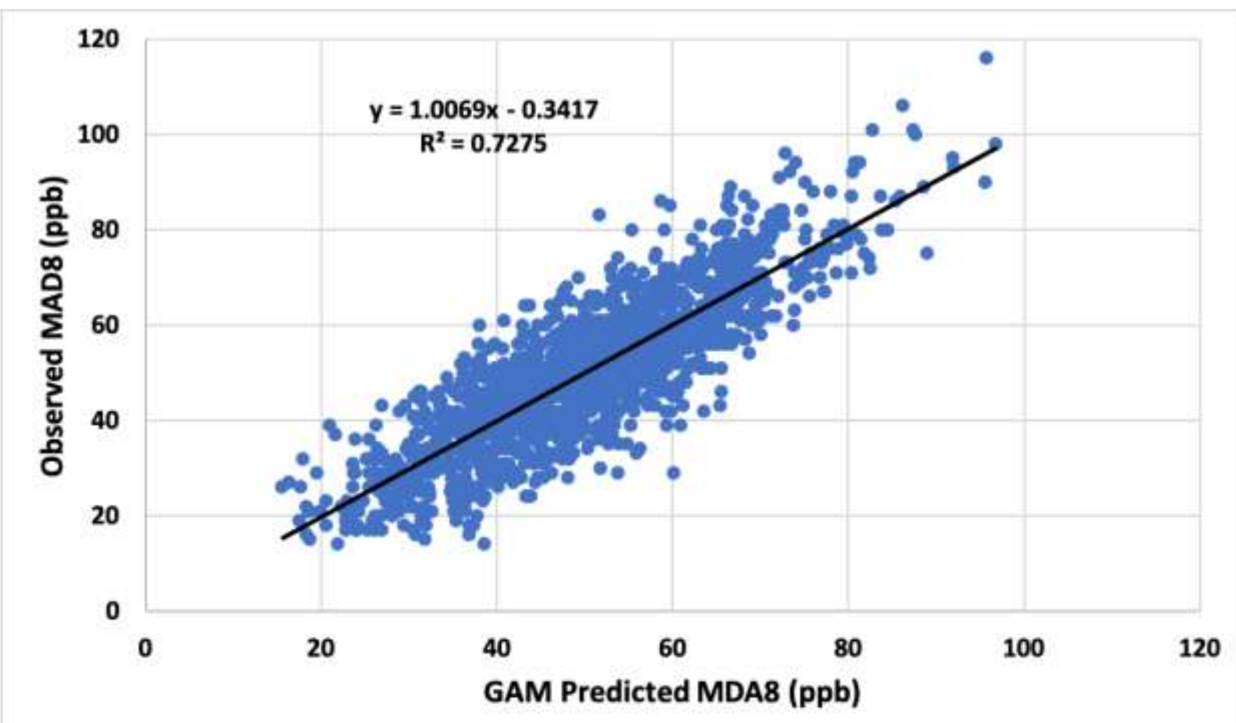

Figure S8. GAM predicted vs Observed, training data only for United Avenue site.

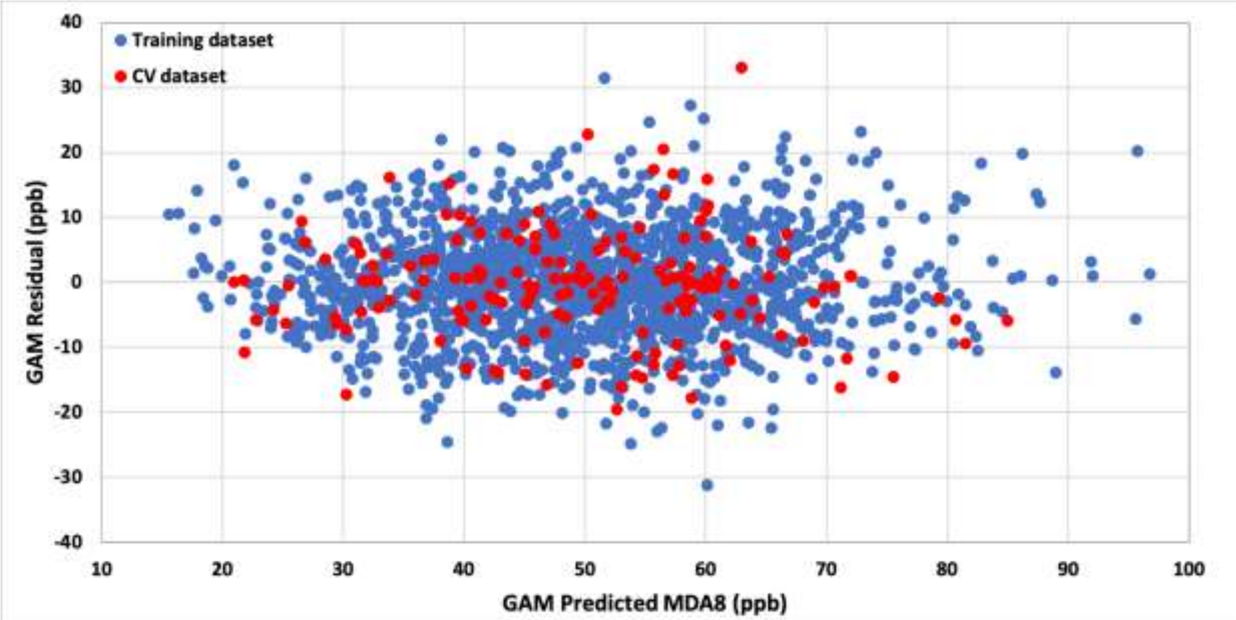

Figure S9. GAM residuals vs predicted, training and cross-validation data for United Avenue site.

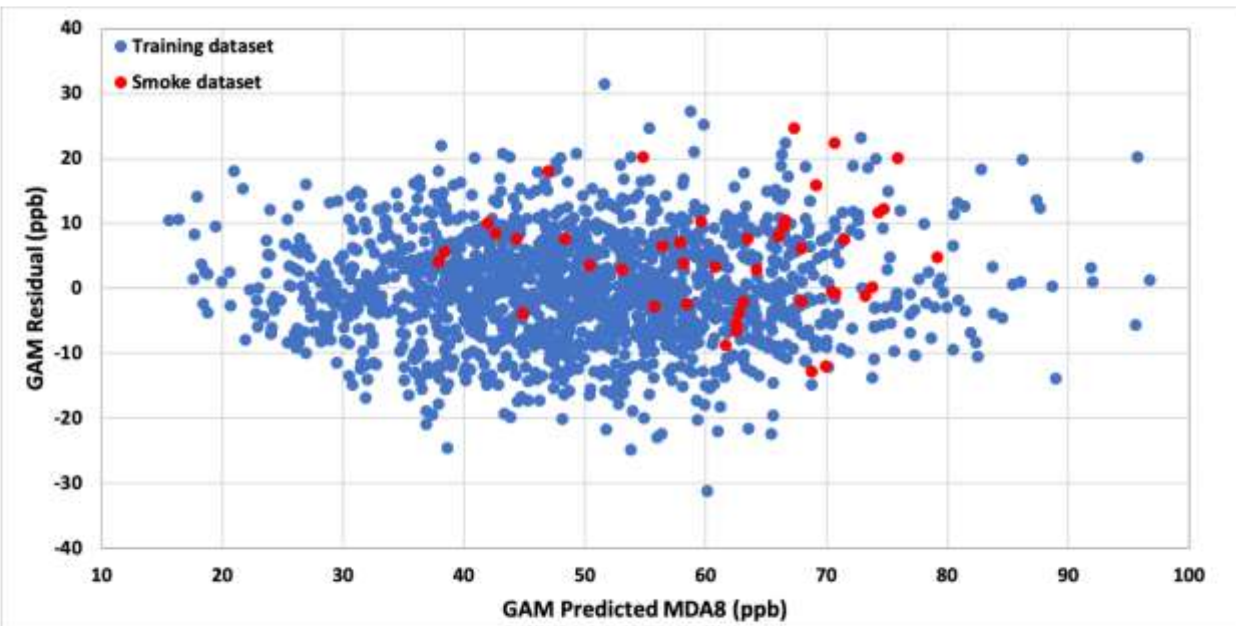

Figure S10. GAM residuals vs predicted, training and smoke data for United Avenue site.

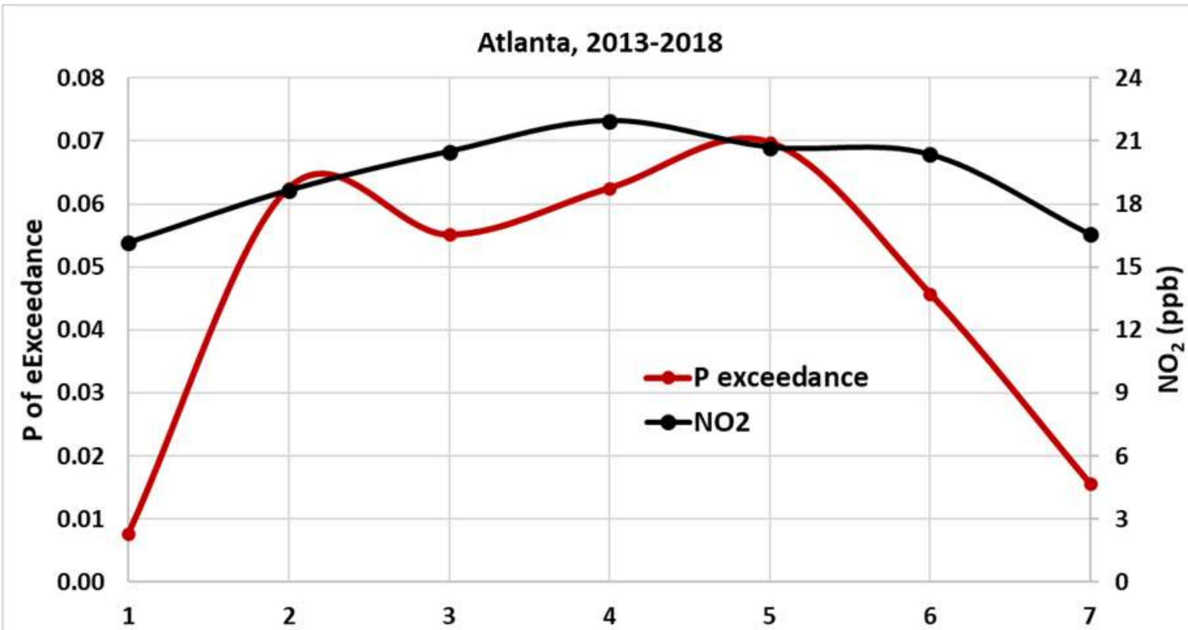

Figure S11. Probability (P) of an O_3 exceedance day (MDA8>70 ppb) across all sites that we considered and averaged surface NO_2 concentrations as a function of day of week (Sunday=1, Monday =2, etc). Both are calculated using May-September 2013-2018 data.

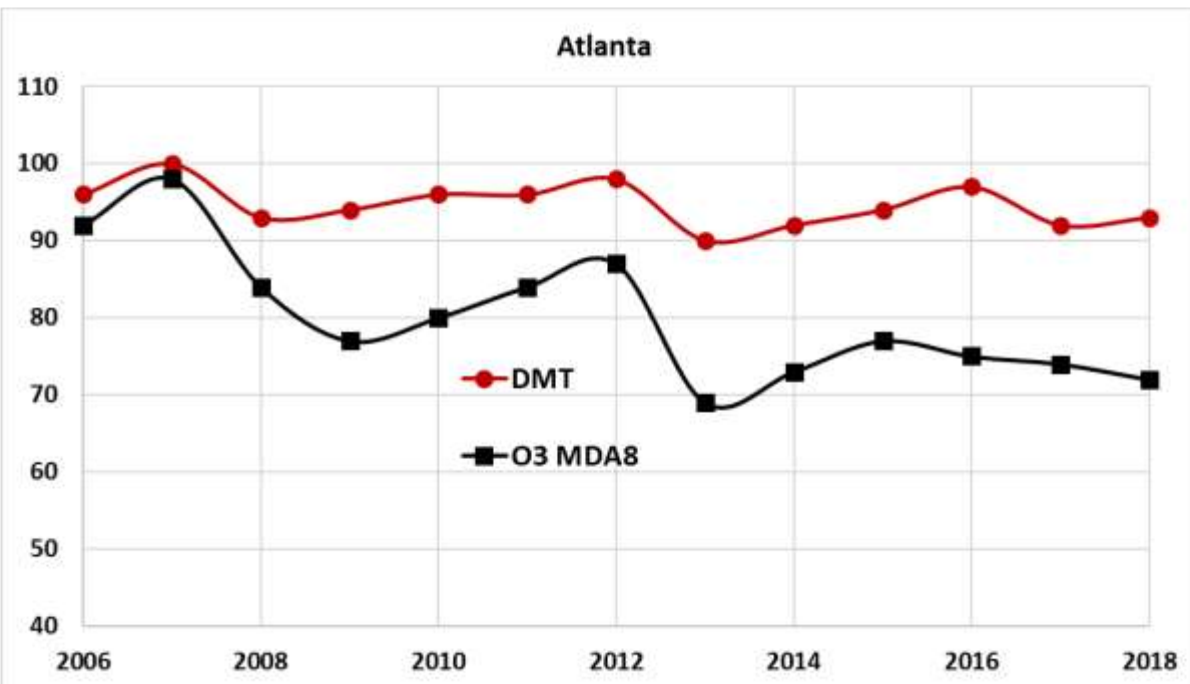

Figure S12. Annual 98th percentile of daily maximum temperature (^oF) and fourth highest MDA8 (ppb).

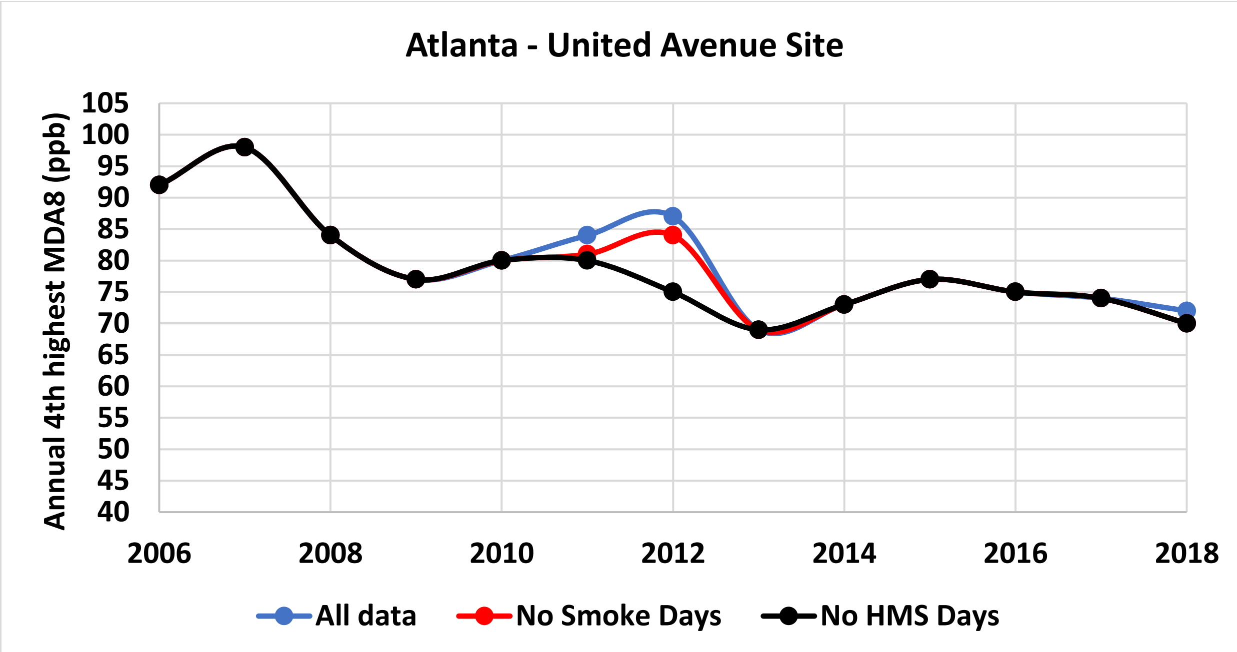

Figure S13. Annual fourth highest for all data, smoke days excluded, and HMS days excluded for United Avenue site.

Table S1. Average, standard deviation and trend in $PM_{2.5}$ for ATL by month for 2006-2018 ($\mu g\ m^{-3}$). Only statistically significant trends are shown. If the trend is not statistically significant, the slope is set to 0. The standard deviation is based on the variation in daily $PM_{2.5}$ observations.

Month	Avg	Slope	SD
1	8.6	-0.1	3.7
2	9.3	-0.3	4.3
3	10.6	-0.5	5.2
4	10.3	-0.4	4.4
5	11.3	-0.6	5.4
6	13.2	-0.9	6.0
7	13.5	-1.0	6.1
8	13.0	-1.0	6.0
9	11.5	-0.7	5.3
10	9.9	-0.3	4.5
11	10.5	0.0	4.4
12	9.4	0.0	4.5

Table S2. Number of days in each category by year, May-September only.

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	136	116	132	135	135	121	131	137	134	136	133	130	130	1706
Cross-val	15	13	14	16	13	13	15	15	14	16	15	14	15	188
Smoke	2	11	3		2	18	3		1	1	2	2	5	50

Table S3. Average MDA8 in each category (ppb)

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	57.6	58.6	54.0	46.6	52.1	54.3	51.4	42.1	45.1	48.2	48.8	46.3	45.0	49.9
Cross-val	57.3	57.4	47.1	45.0	53.3	52.8	50.5	42.7	49.0	53.0	44.7	49.1	44.1	49.5
Smoke	74.5	69.9	61.7		62.5	70.2	73.0		61.0	53.0	61.5	43.0	60.8	66.8

Table S4. Average surface Tmax (ºF) in each category

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	85.5	87.2	85.5	84.3	88.8	87.8	87.1	83.3	85.1	86.2	89.0	85.1	88.0	86.3
Cross-val	85.8	88.1	85.1	84.1	90.1	87.5	86.5	81.9	86.9	86.6	89.7	85.3	87.0	86.5
Smoke	93.0	92.8	92.0		86.0	92.7	96.0		90.0	94.0	96.0	82.5	90.4	92.1

Table S5. Number of days included in the GAM

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	125	97	106	132	127	120	131	133	125	120	132	117	119	1584
Cross-val	15	10	12	16	13	12	14	15	14	14	15	10	14	174
Smoke	2	8	3		2	17	3		1		2	2	4	44

Table S6. Average of residuals from the GAM

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	0.11	-0.87	0.77	0.42	-0.41	0.03	-0.37	0.95	-0.61	1.07	-2.63	0.86	0.81	0.00
Cross-val	0.01	-2.82	-0.12	-0.74	-2.47	-2.60	-2.46	3.30	-0.43	3.41	-4.11	3.96	-0.61	-0.45
Smoke	9.57	3.91	2.57		5.60	8.42	0.18		-2.11		-3.72	4.87	1.43	4.98

Chicago CSA

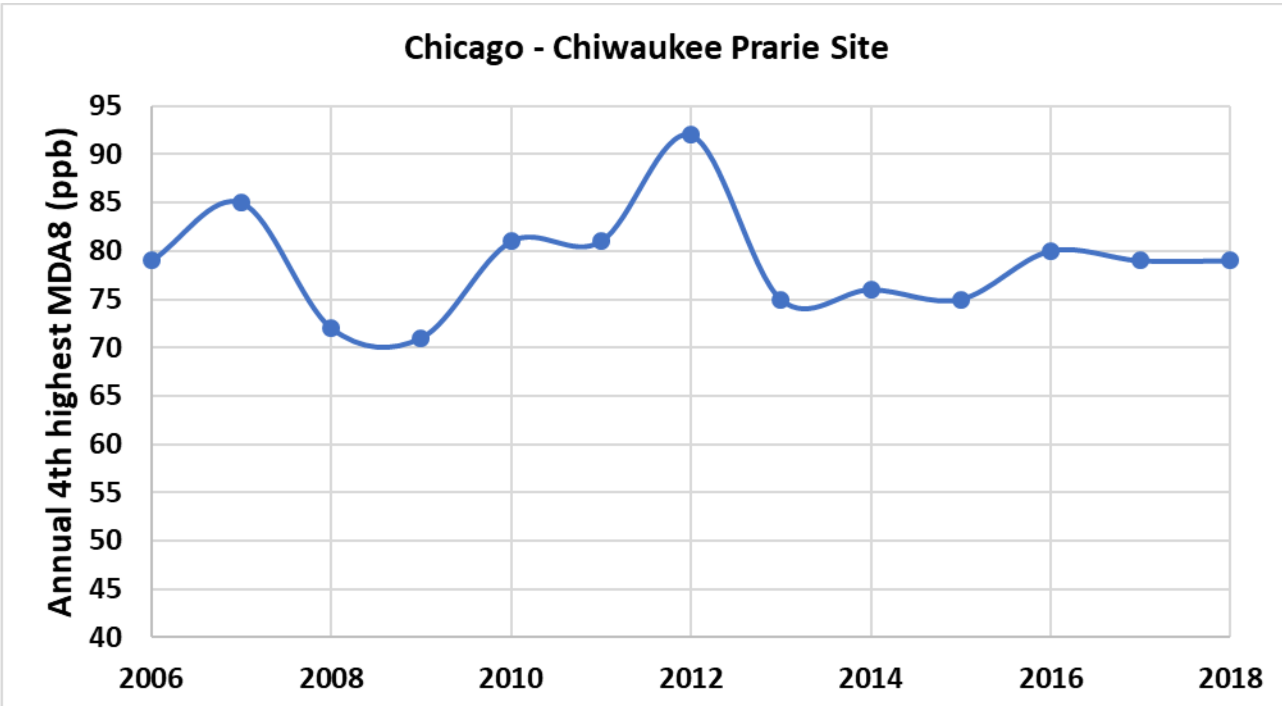

Figure S1. Trend in 4th highest MDA8 O_3 for each site.

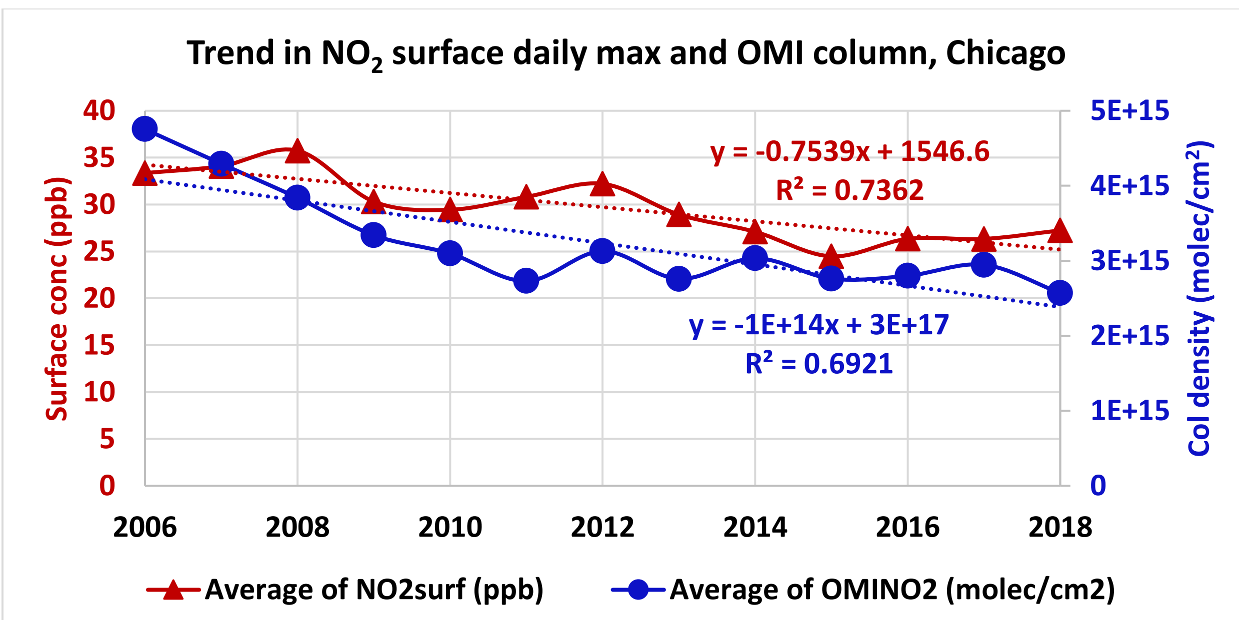

Figure S2. Trend in surface NO_2 daily max and OMI column, 2006-2018.

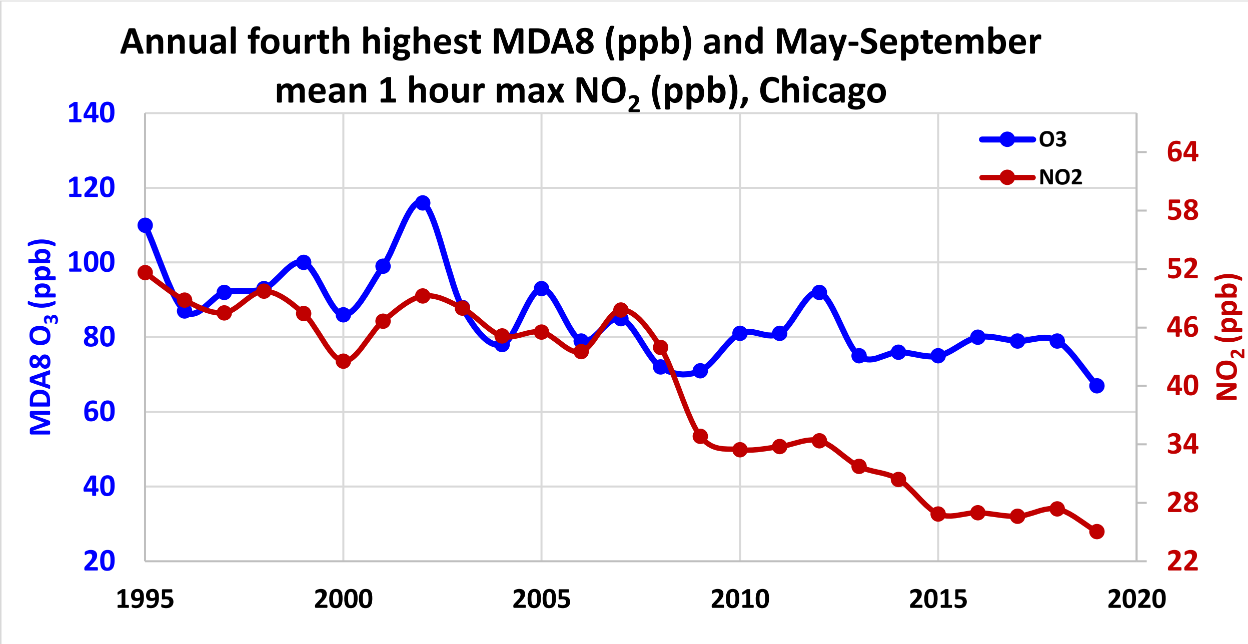

Figure S3. Trend in 4th highest O_3 and daily max NO_2, 1995-2019. For this plot and the next one, I use NO_2 data from AQS sites 170310063, 170313103, 170314002 and 180890022.

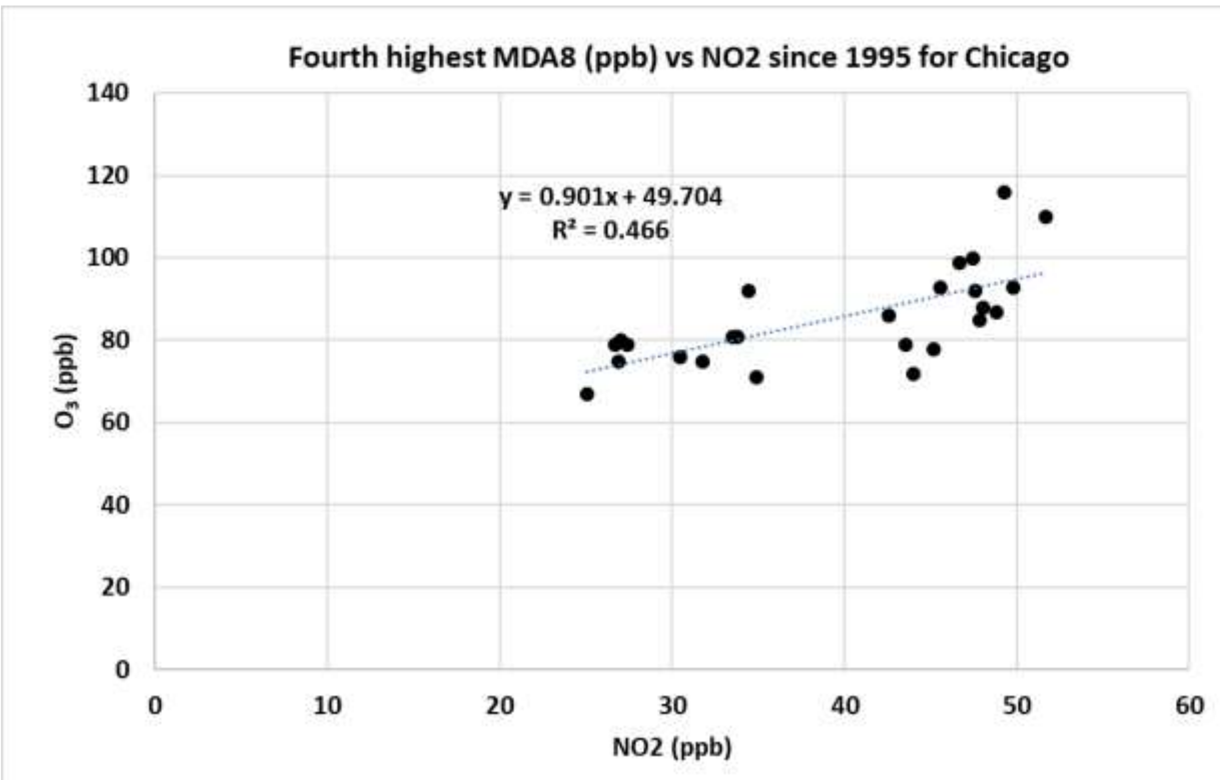

Figure S4. O_3 vs NO_2, 1995-2019.

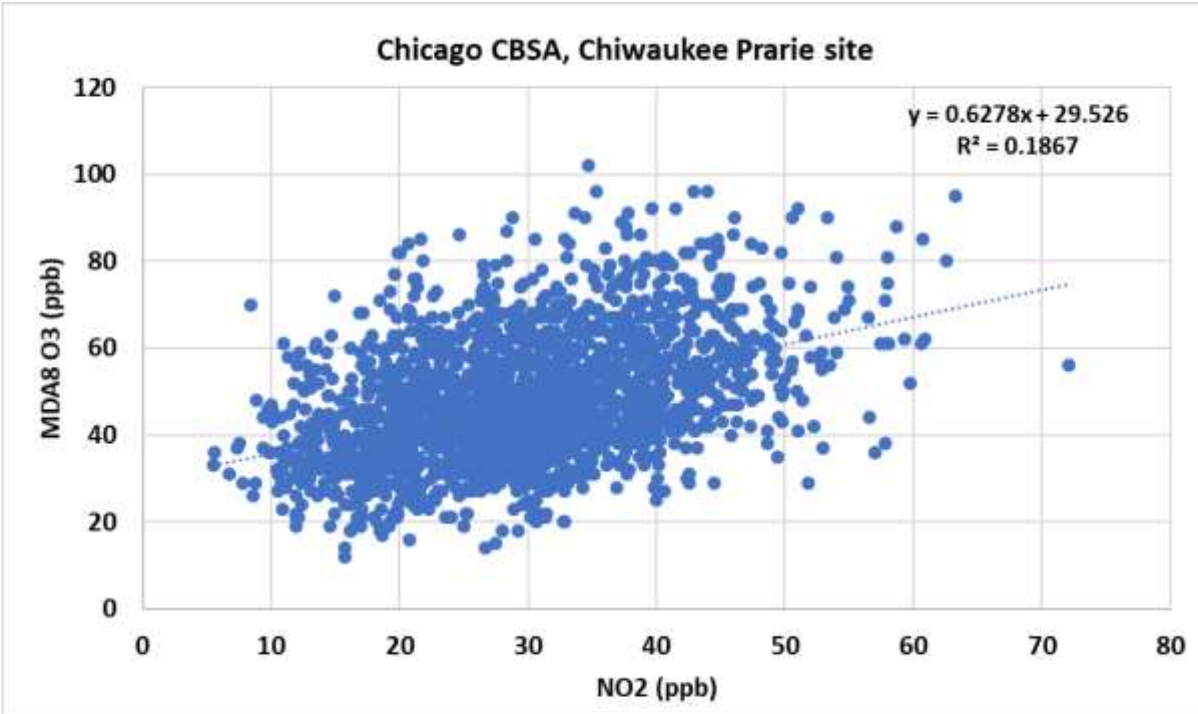

Figure S5. MDA8 vs daily max NO_2 for Chiwaukee Prairie Stateline site.

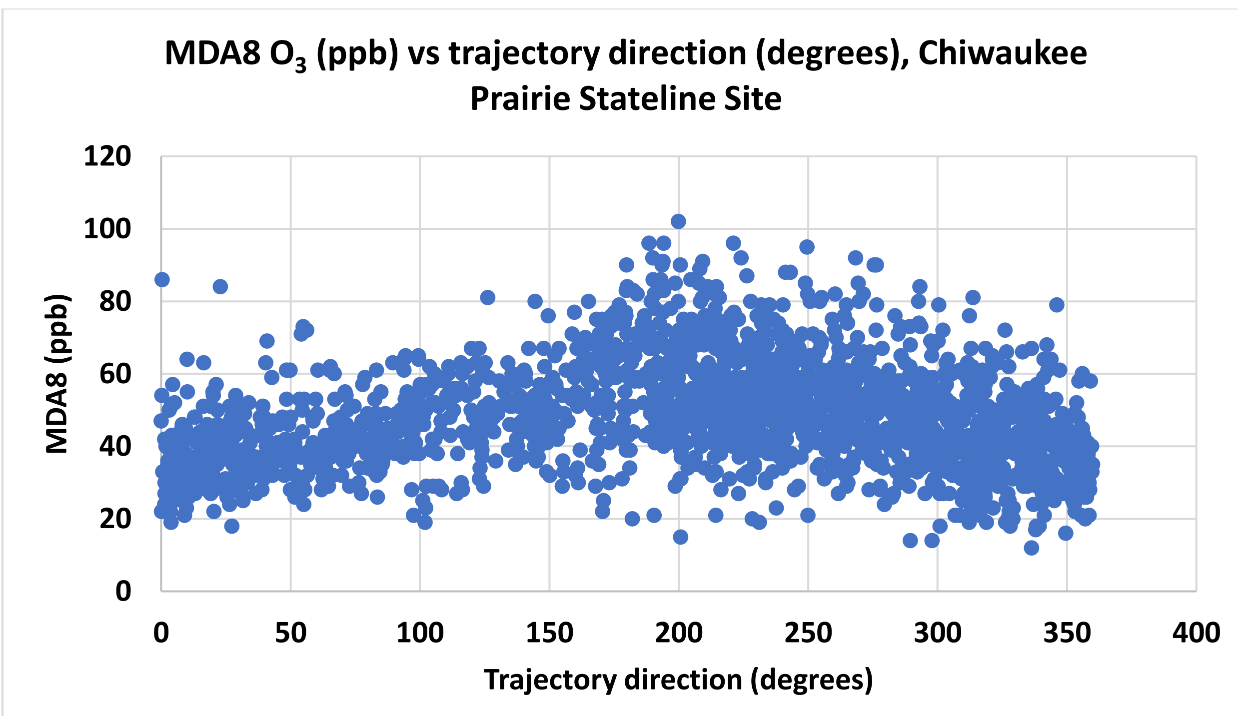

Figure S6. MDA8 vs trajectory direction for Chiwaukee Prairie Stateline site.

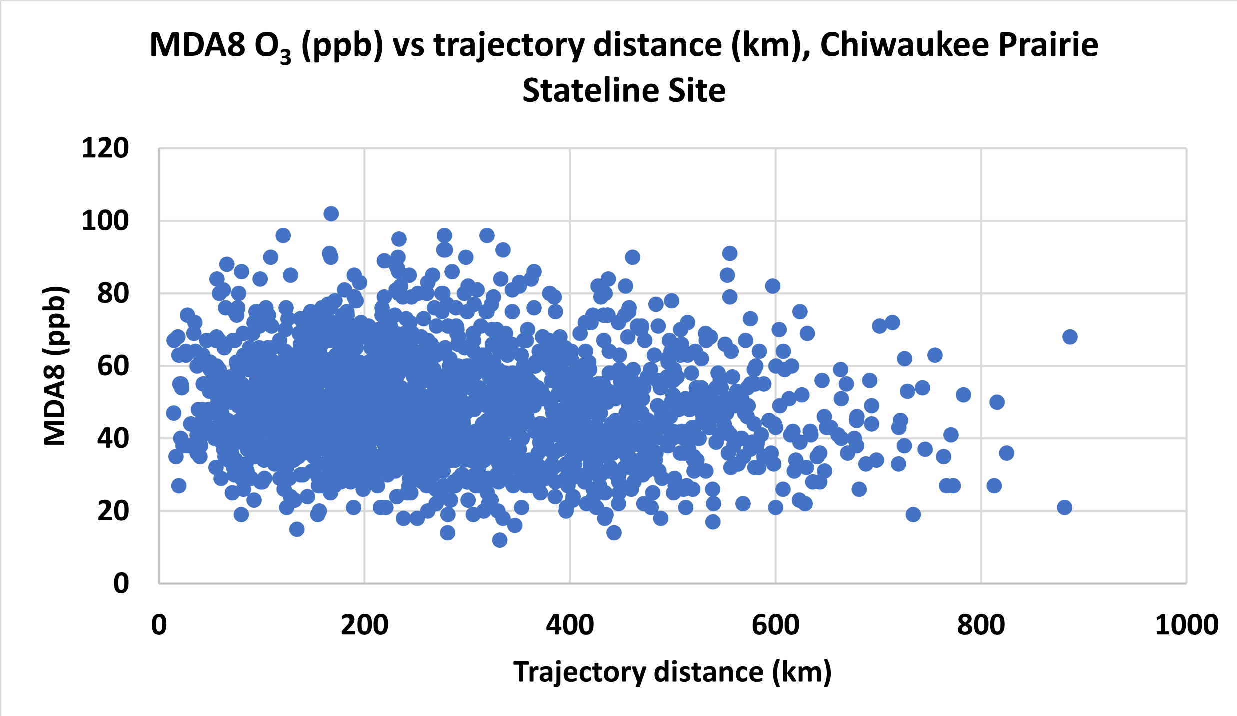

Figure S7. MDA8 vs trajectory distance for Chiwaukee Prairie Stateline site.

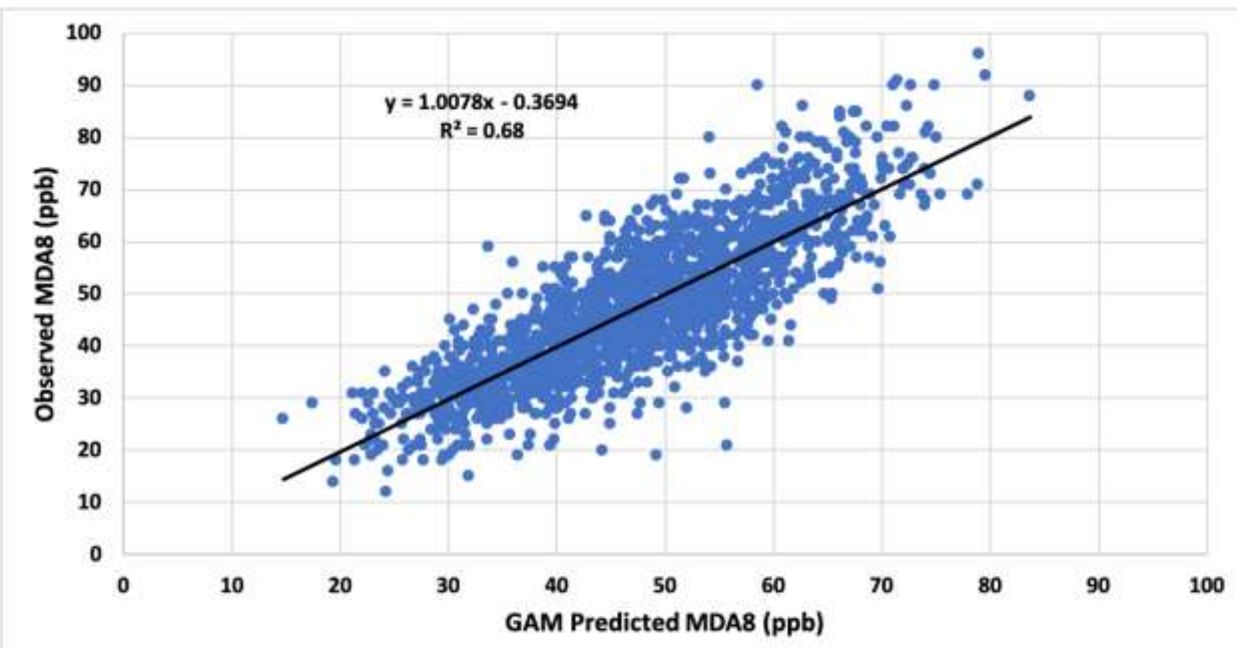

Figure S8. GAM predicted vs Observed, training data only for Chiwaukee Prairie Stateline site.

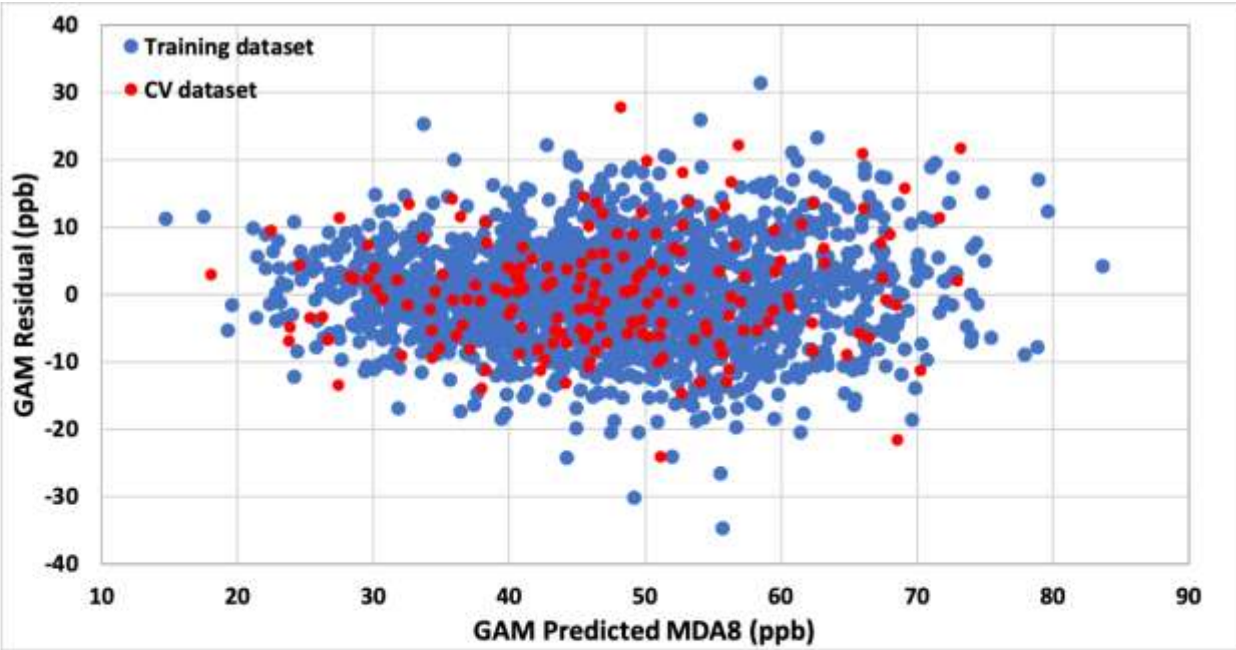

Figure S9. GAM residuals vs predicted, training and cross-validation data for Chiwaukee Prairie Stateline site.

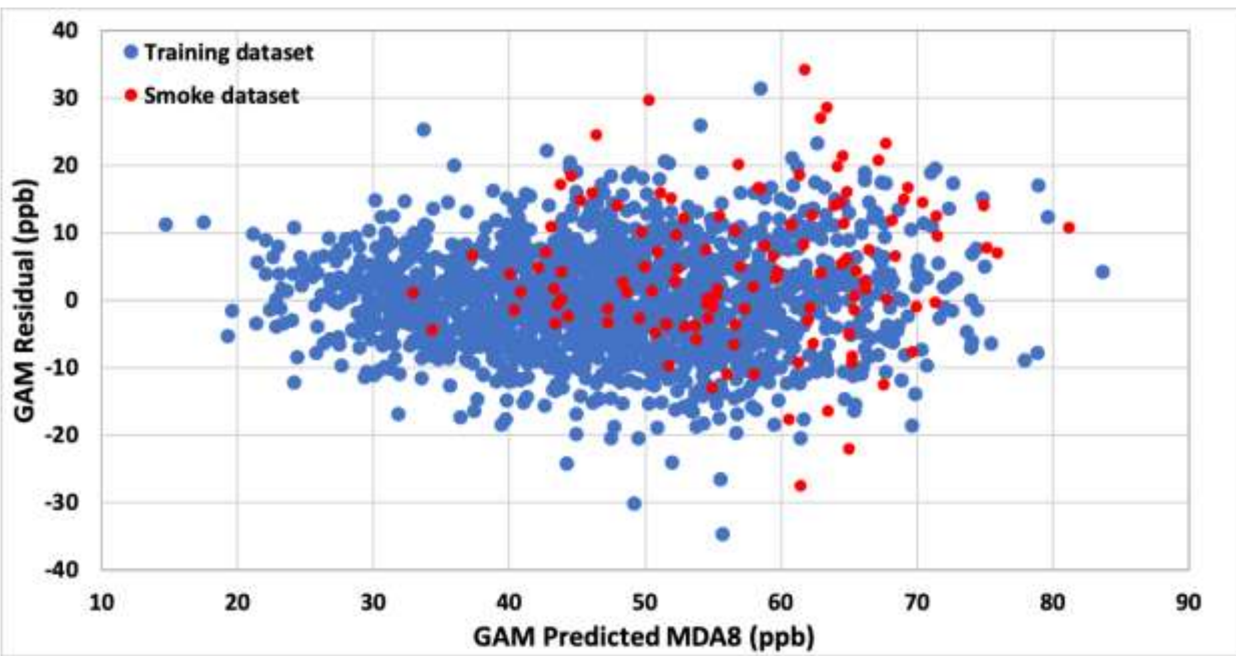

Figure S10. GAM residuals vs predicted, training and smoke data for Chiwaukee Prairie Stateline site.

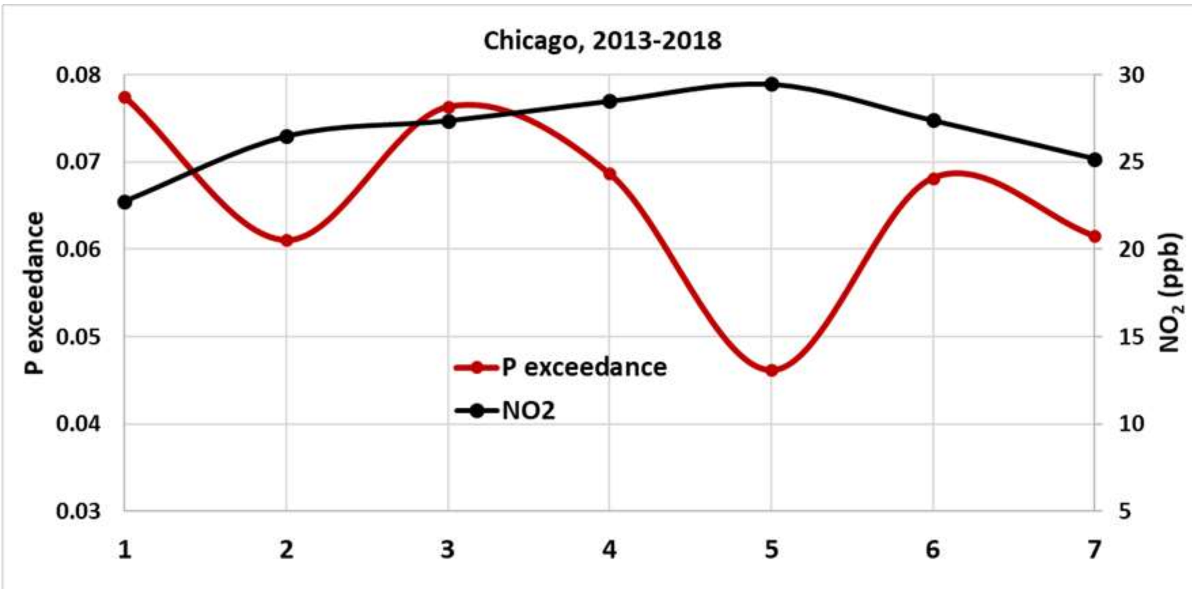

Figure S11. Probability (P) of an O_3 exceedance day (MDA8>70 ppb) across all sites that we considered and averaged surface NO_2 concentrations as a function of day of week (Sunday=1, Monday =2, etc). Both are calculated using May-September 2013-2018 data.

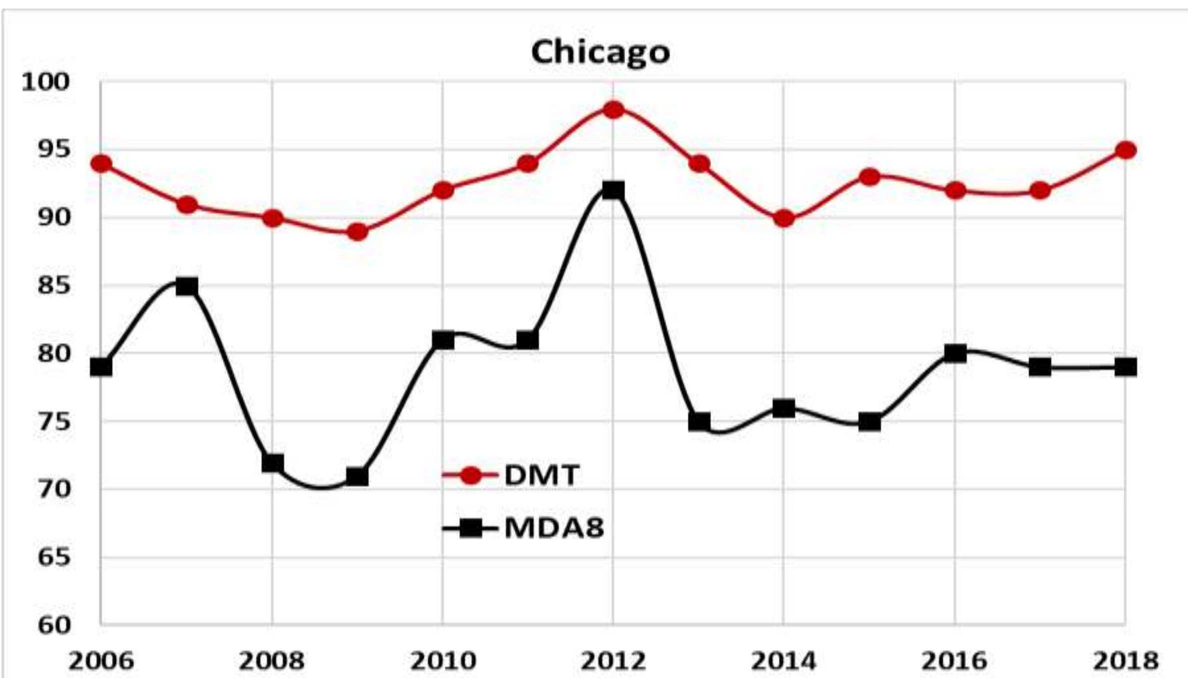

Figure S12. Annual 98th percentile of daily maximum temperature (ºF) and fourth highest MDA8 (ppb).

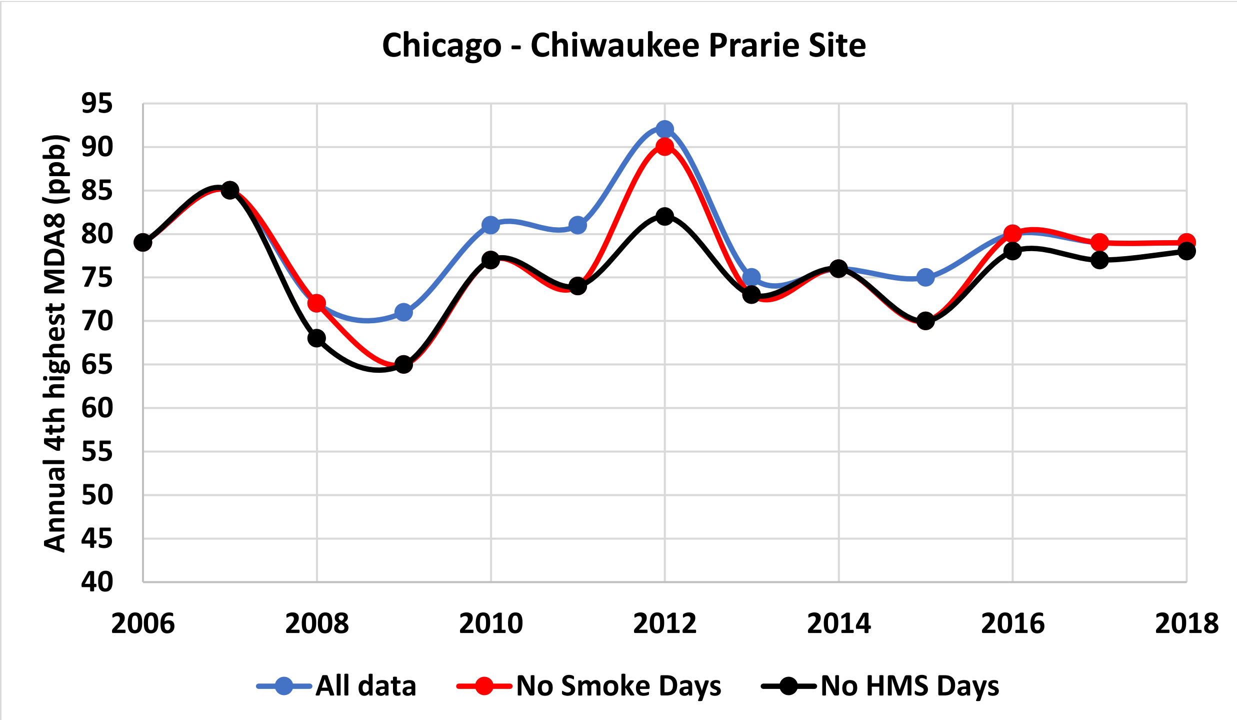

Figure S13. Annual fourth highest for all data, smoke days excluded, and HMS days excluded for Chiwaukee Prairie Stateline site.

Table S1. Average, standard deviation and trend in $PM_{2.5}$ for CHI by month for 2006-2018 ($\mu g\ m^{-3}$). Only statistically significant trends are shown. If the trend is not statistically significant, the slope is set to 0. The standard deviation is based on the variation in daily $PM_{2.5}$ observations.

Month	Avg	Slope	SD
1	12.1	-0.4	5.4
2	12.6	-0.4	6.5
3	10.8	-0.4	5.3
4	9.7	-0.4	4.7
5	10.1	-0.4	4.9
6	10.7	-0.4	4.8
7	11.9	-0.5	5.1
8	10.8	-0.4	5.6
9	9.7	-0.4	4.6
10	8.7	-0.4	3.8
11	11.4	-0.5	5.8
12	11.8	-0.5	5.9

Table S2. Number of days in each category by year, May-September only.

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	133	129	132	120	133	119	121	132	123	123	137	136	129	1667
Cross-val	13	15	14	15	14	13	14	15	14	15	15	14	13	184
Smoke	5	9	7	13	6	18	17	6	16	12		3	11	123

Table S3. Average MDA8 in each category

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	45.8	50.7	47.2	40.3	49.4	43.4	52.6	46.7	45.4	45.5	49.1	49.1	47.9	47.2
Cross-val	49.7	56.1	43.4	37.3	46.9	43.2	50.4	52.1	43.6	44.3	48.7	52.4	51.5	47.6
Smoke	59.6	70.3	58.1	52.3	55.7	67.8	76.5	63.8	57.9	59.8		58.3	52.5	62.2

Table S4. Average surface Tmax (ºF) in each category

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	77.0	79.8	77.7	75.7	79.7	76.7	80.9	78.4	78.8	79.4	80.8	78.8	80.7	78.8
Cross-val	79.5	81.3	76.2	75.1	76.6	75.5	78.6	81.3	74.9	76.9	81.3	80.0	84.2	78.6
Smoke	84.5	88.0	82.4	80.4	87.7	90.1	94.2	86.3	83.5	85.0		89.0	89.5	87.1

Table S5. Number of days included in the GAM

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	127	124	131	120	131	117	117	128	119	121	135	133	128	1631
Cross-val	12	13	14	15	14	13	14	15	14	15	15	14	13	181
Smoke	5	9	7	13	6	18	16	6	16	12	0	3	10	121

Table S6. Average of residuals from the GAM

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	-1.04	0.43	-0.44	-3.27	1.46	-1.68	0.76	0.18	0.20	0.42	-0.03	3.24	-0.62	0.00
Cross-val	2.44	3.77	-2.15	-4.91	0.19	-1.54	2.61	2.59	0.33	-0.16	-1.12	4.83	-0.42	0.44
Smoke	3.20	7.85	4.49	1.70	-4.15	5.60	10.4	6.76	6.56	7.34		2.69	-3.74	4.84

Dallas CSA

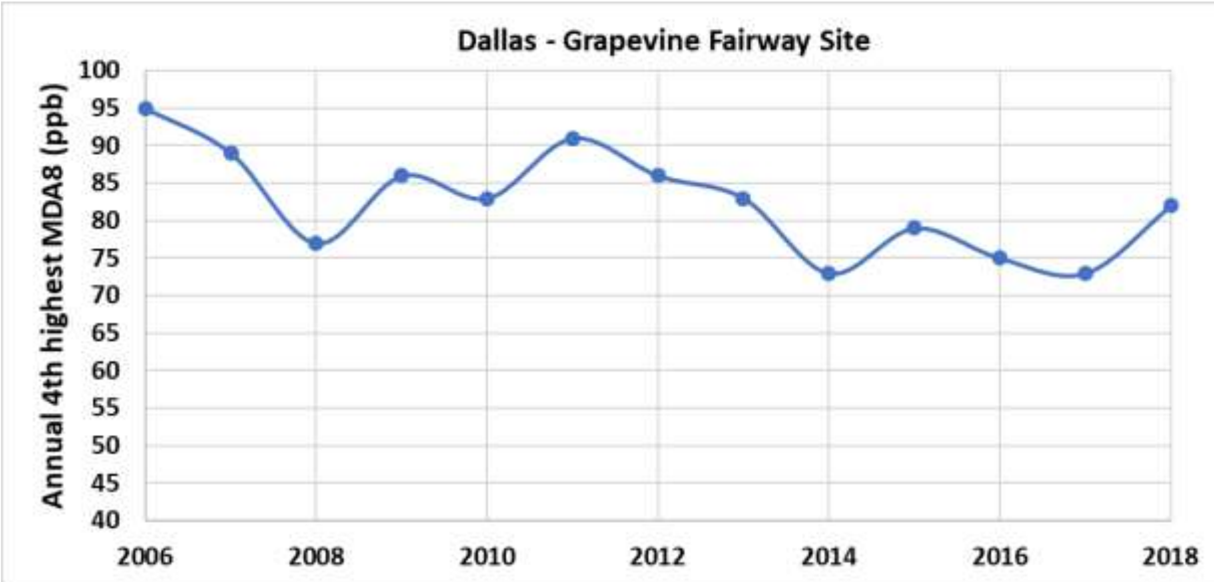

Figure S1. Trend in 4th highest MDA8 O_3 for each site.

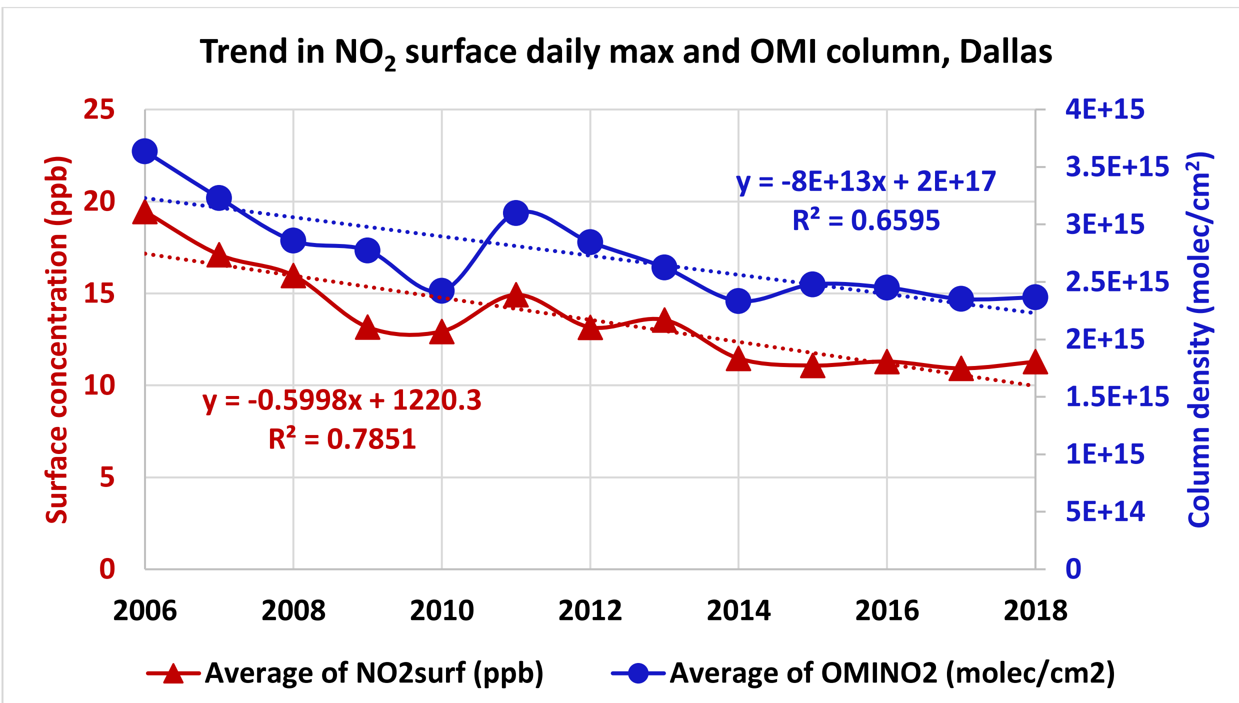

Figure S2. Trend in surface NO_2 daily max and OMI column, 2006-2018.

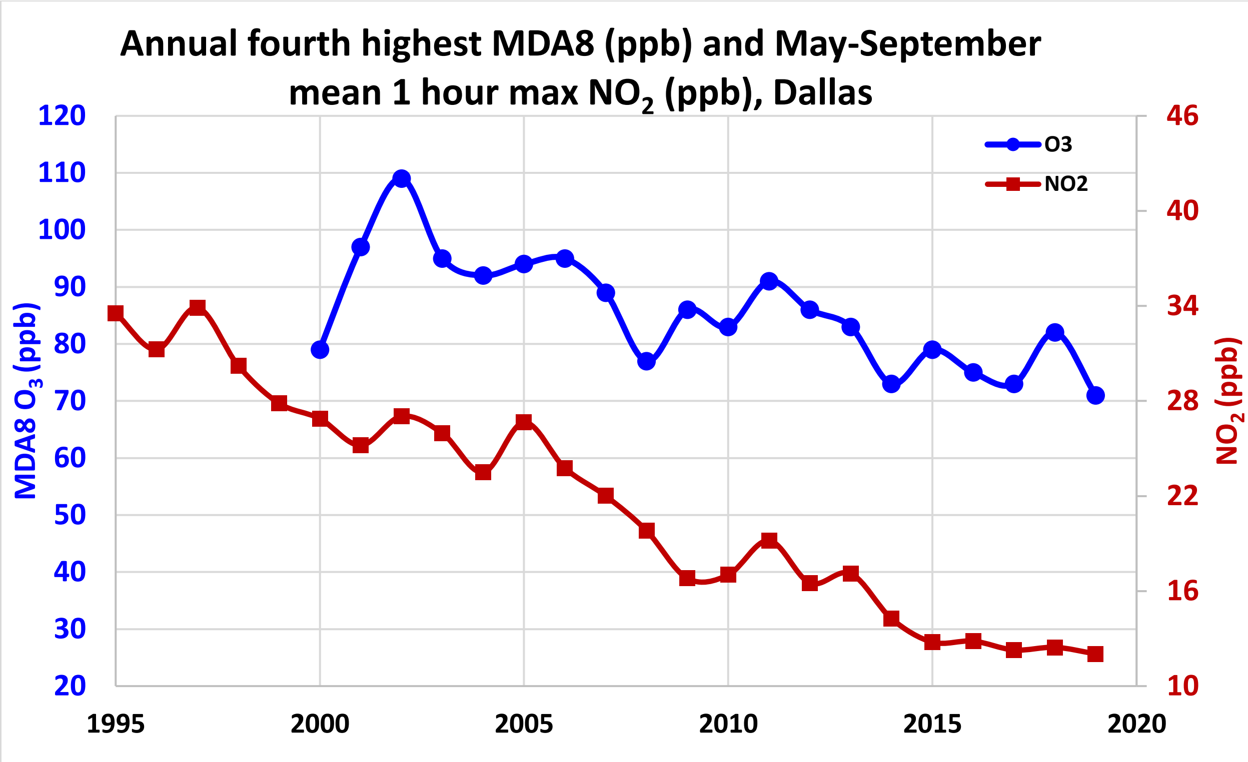

Figure S3. Trend in 4th highest O_3 and daily max NO_2, 1995-2019. For this plot and the next one, I use NO_2 data from AQS sites 481130069, 481130075, 481130087, 481210034 and 484391002.

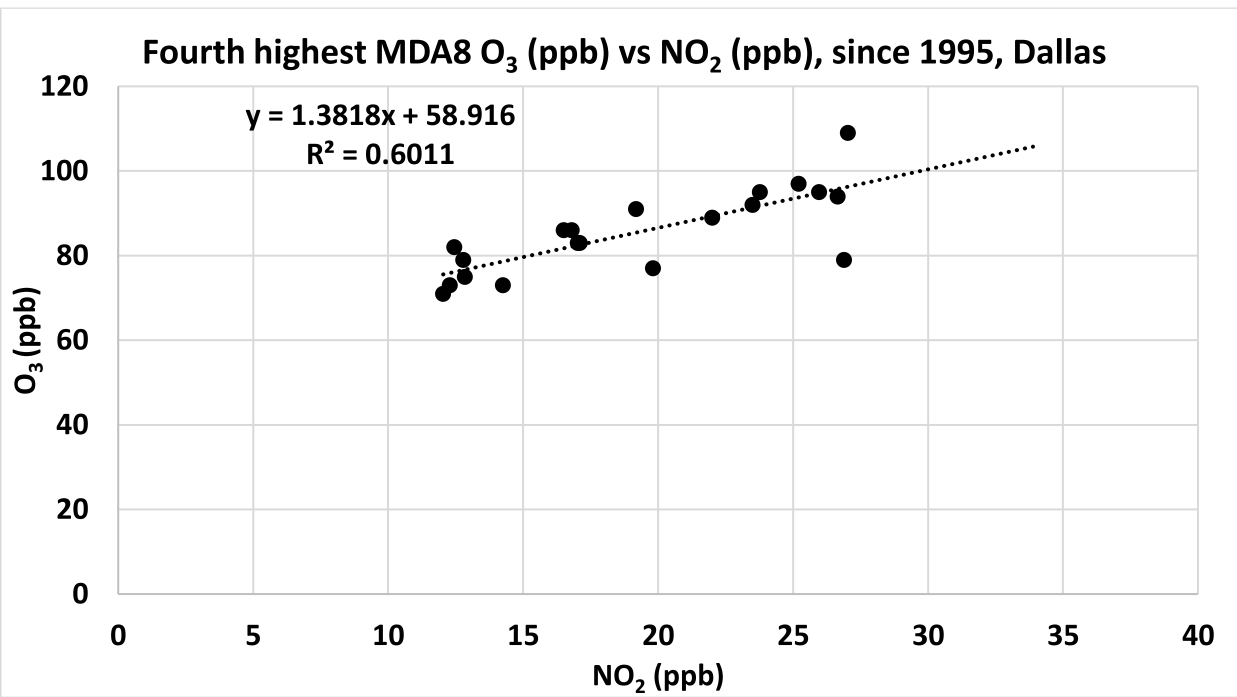

Figure S4. O_3 vs NO_2, 1995-2019.

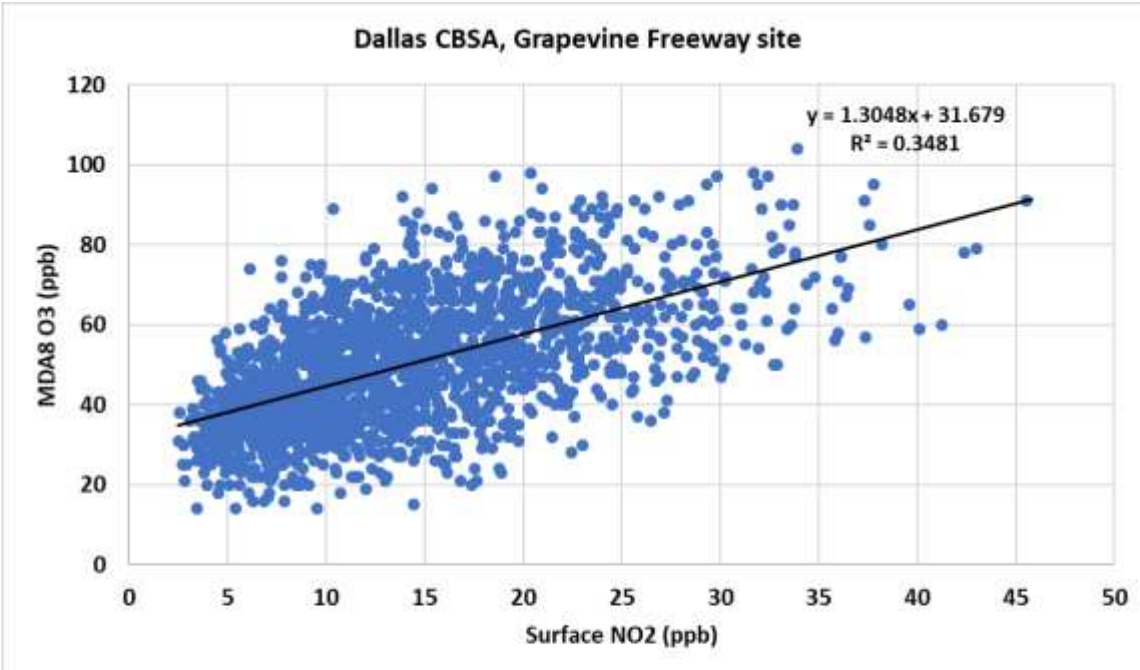

Figure S5. MDA8 vs daily max NO_2 for Grapevine Fairway site.

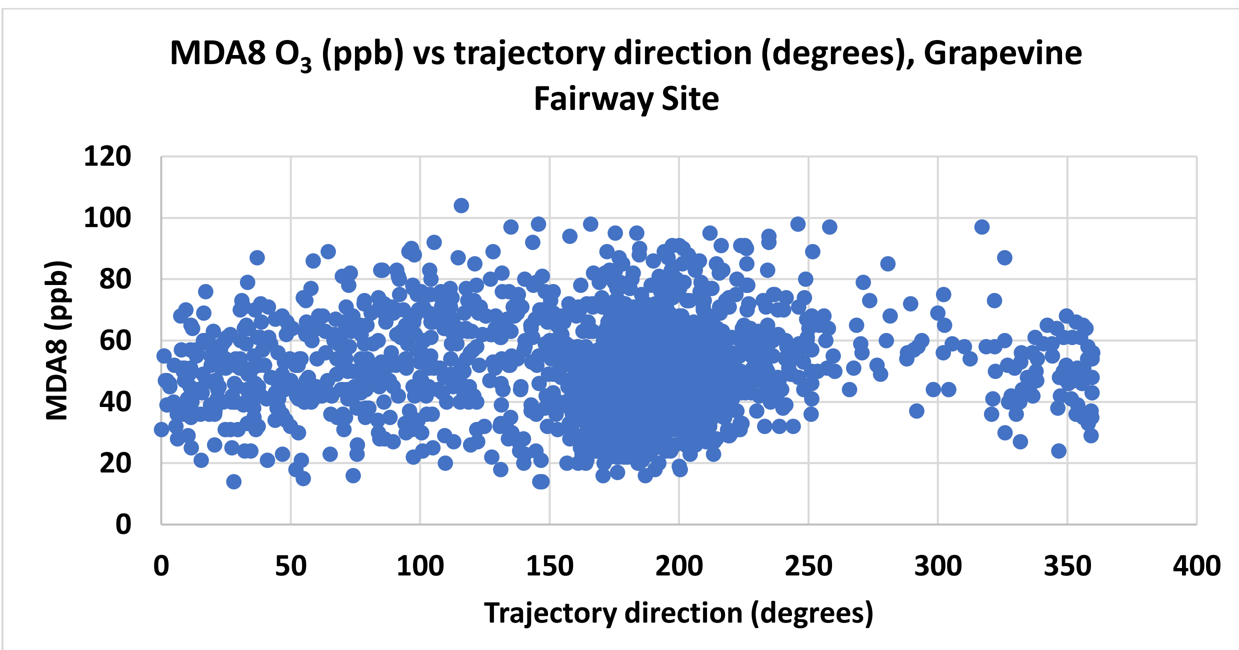

Figure S6. MDA8 vs trajectory direction for Grapevine Fairway site.

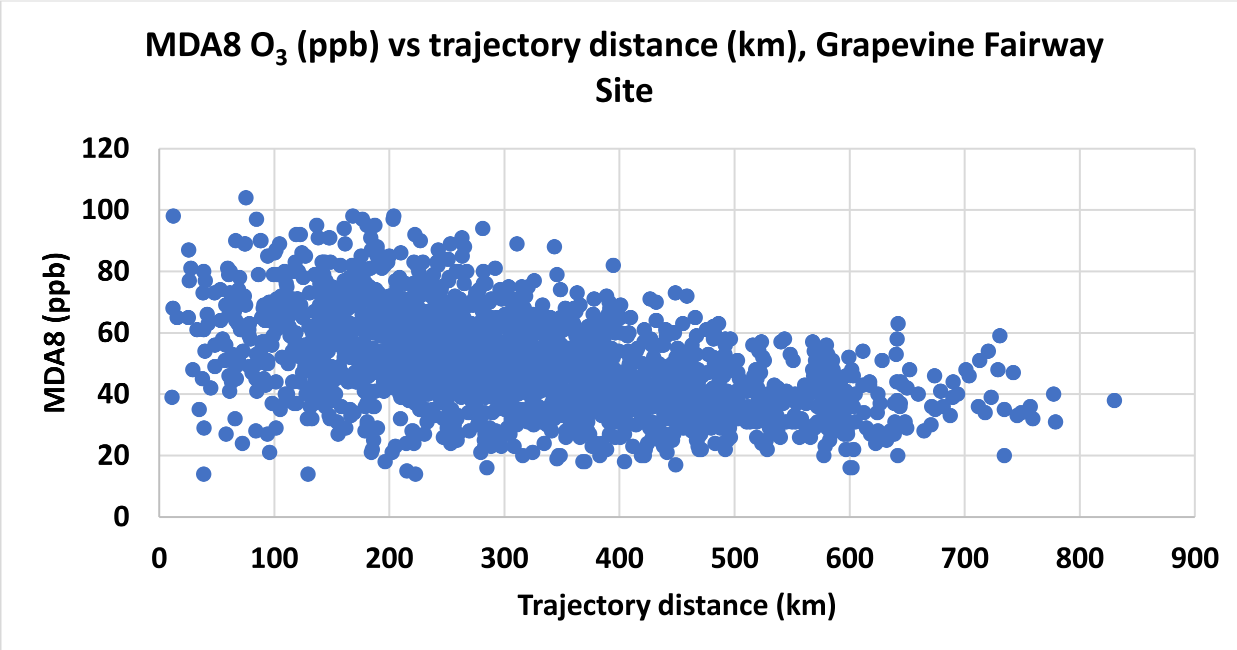

Figure S7. MDA8 vs trajectory distance for Grapevine Fairway site.

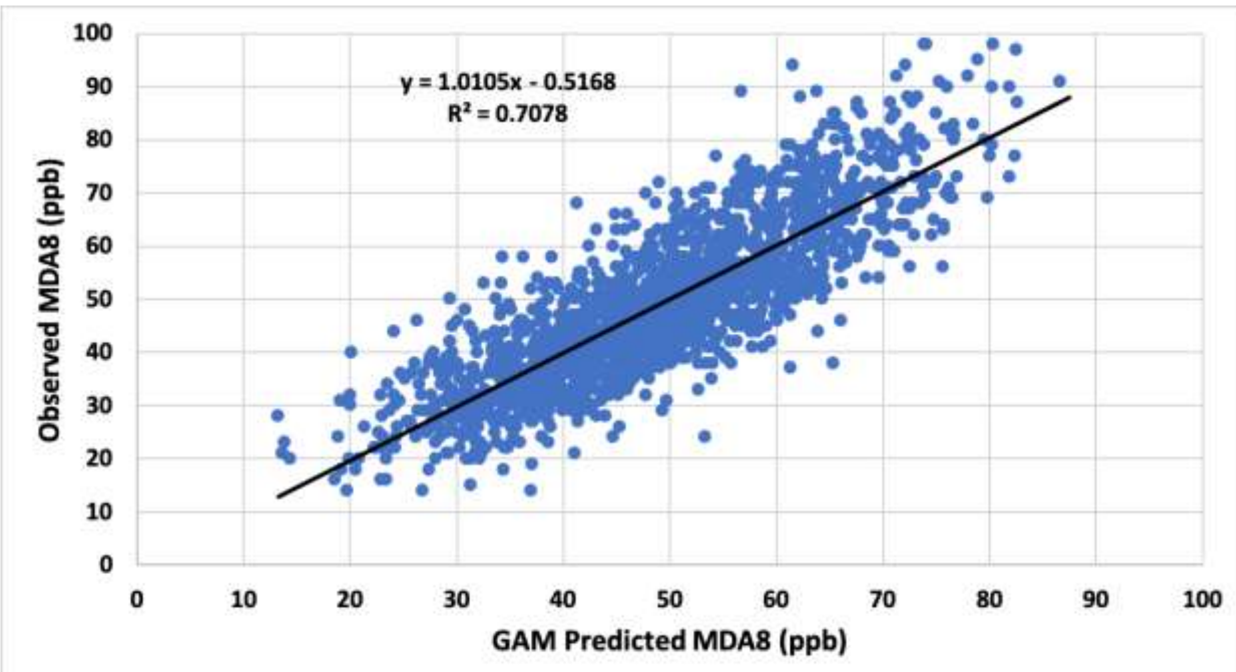

Figure S8. GAM predicted vs Observed, training data only for Grapevine Fairway site.

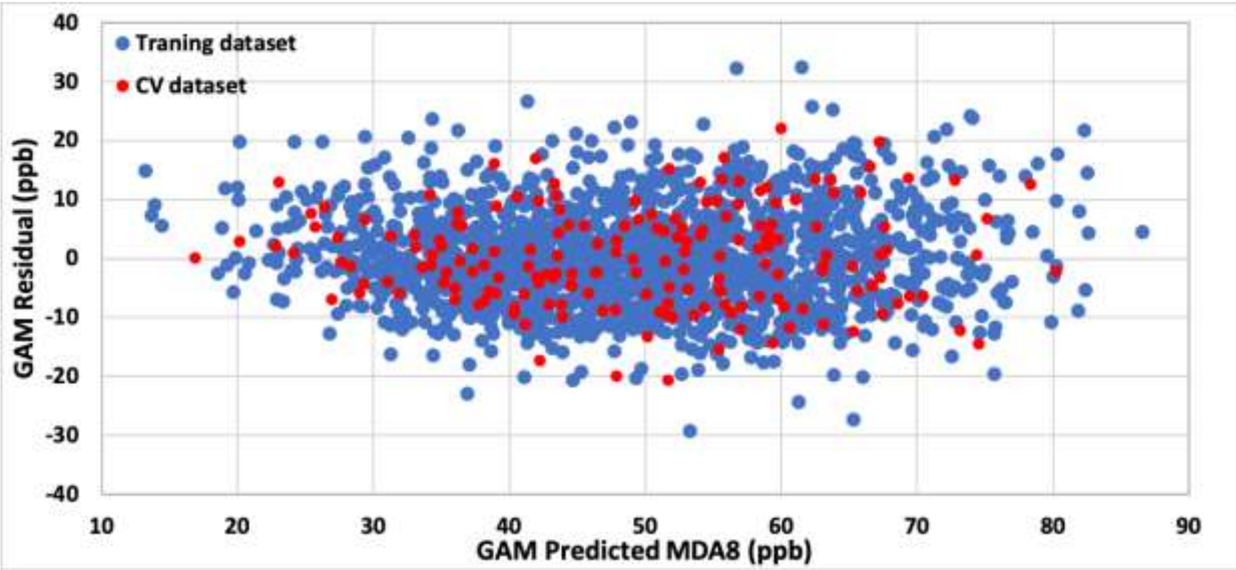

Figure S9. GAM residuals vs predicted, training and cross-validation data for Grapevine Fairway site.

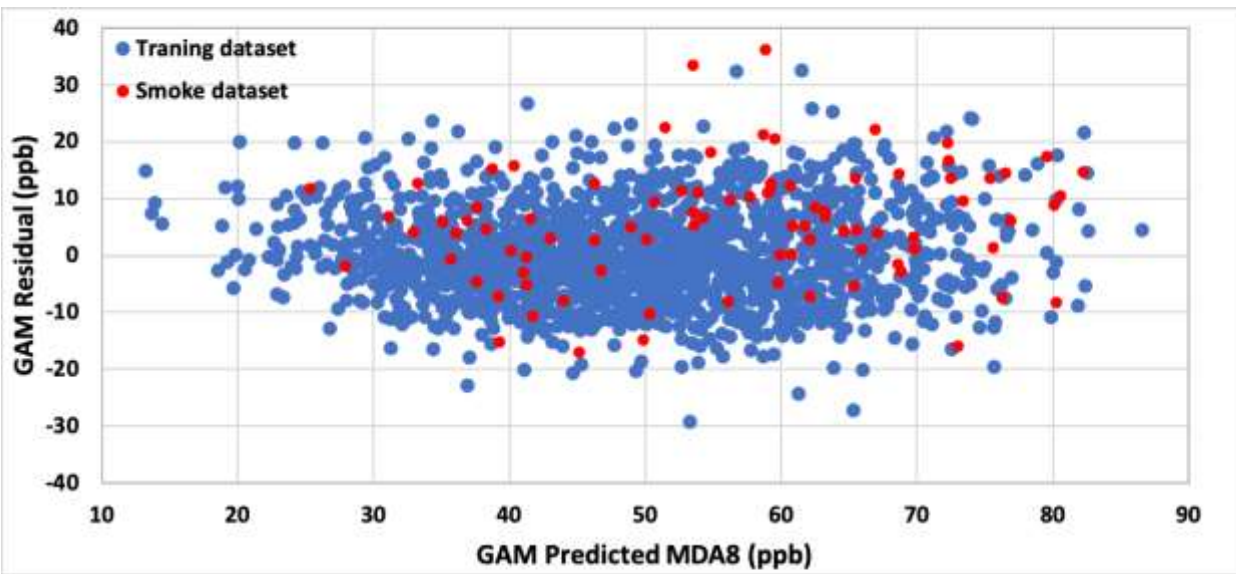

Figure S10. GAM residuals vs predicted, training and smoke data for Grapevine Fairway site.

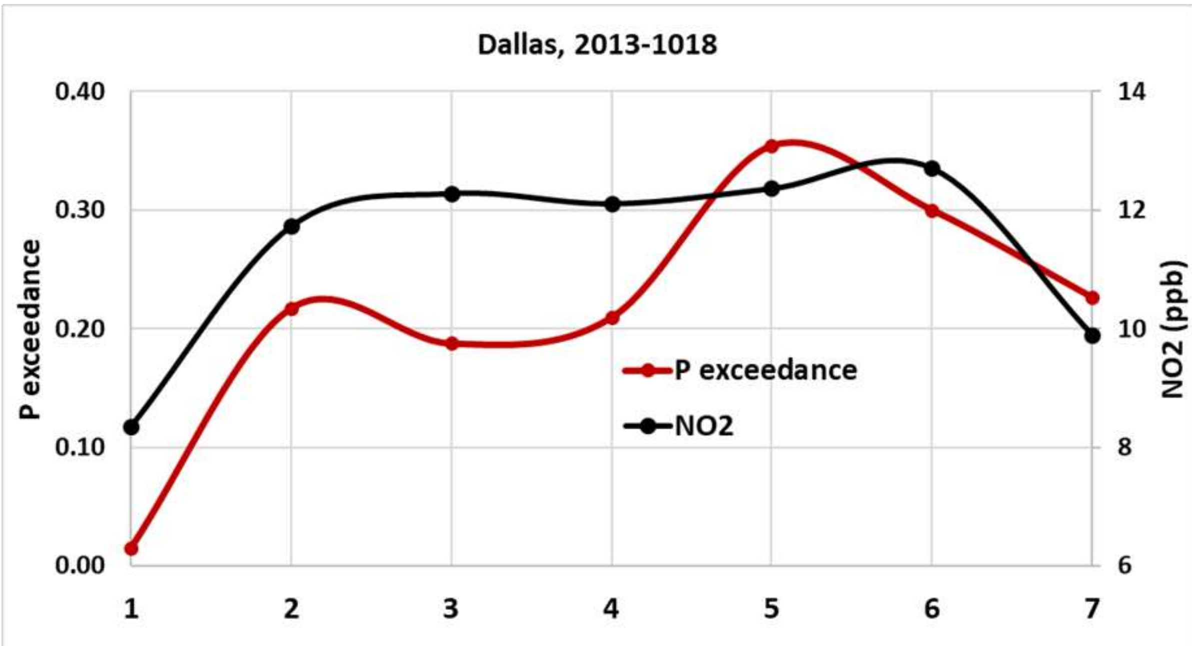

Figure S11. Probability (P) of an O_3 exceedance day (MDA8>70 ppb) across all sites that we considered and averaged surface NO_2 concentrations as a function of day of week (Sunday=1, Monday =2, etc). Both are calculated using May-September 2013-2018 data.

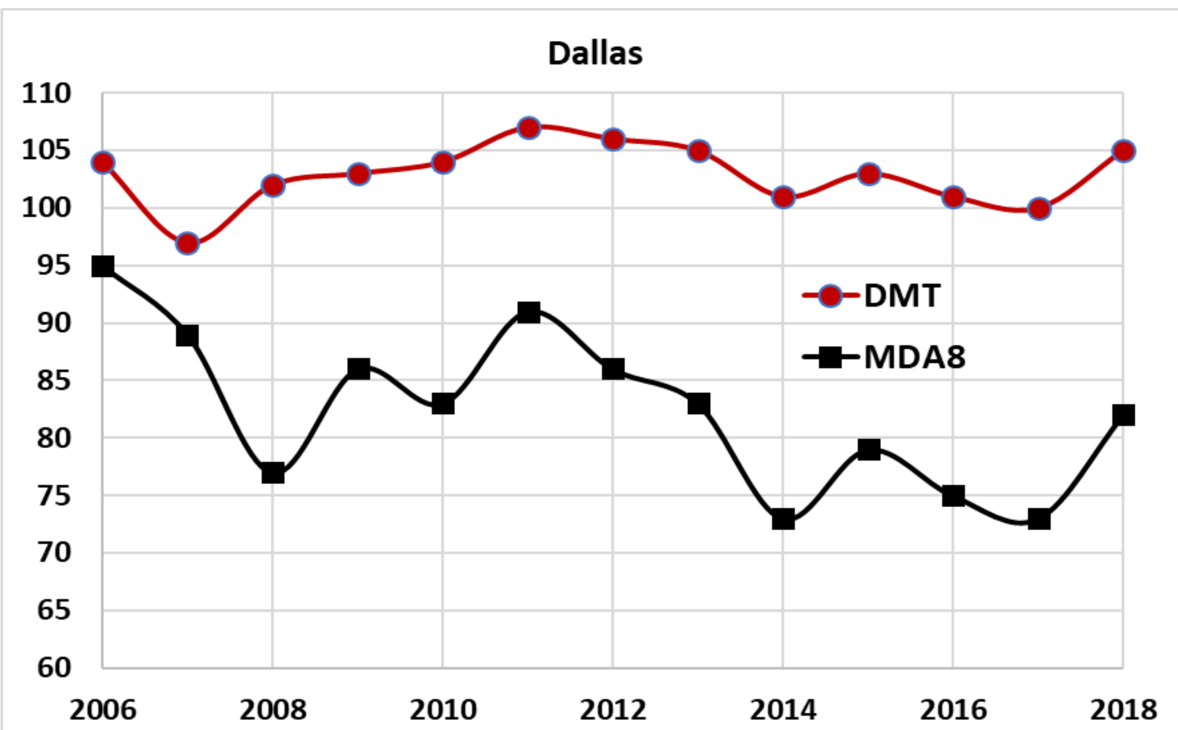

Figure S12. Annual 98th percentile of daily maximum temperature (°F) and fourth highest MDA8 (ppb).

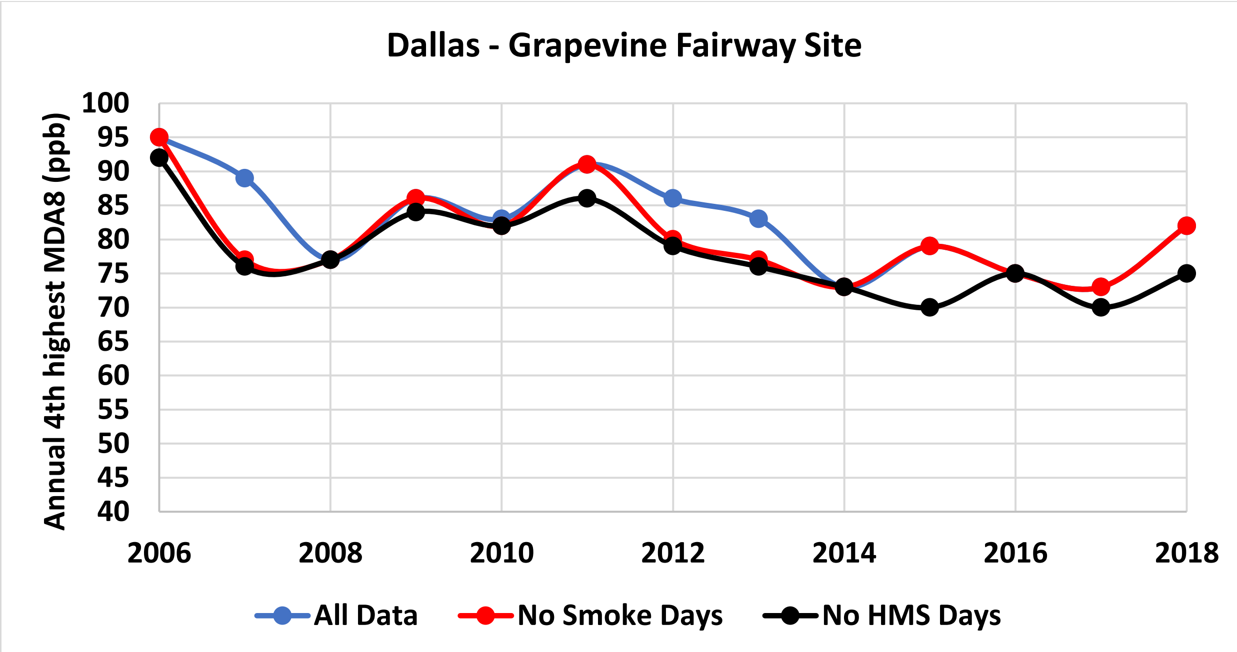

Figure S13. Annual fourth highest for all data, smoke days excluded, and HMS days excluded for Grapevine Fairway site.

Table S1. Average, standard deviation and trend in $PM_{2.5}$ for DAL by month for 2006-2018 (μg m^{-3}). Only statistically significant trends are shown. If the trend is not statistically significant, the slope is set to 0. The standard deviation is based on the variation in daily $PM_{2.5}$ observations.

Month	Avg	Slope	SD
1	7.3	0.0	2.7
2	8.2	0.0	3.3
3	8.5	-0.2	3.3
4	9.1	-0.3	3.9
5	9.2	-0.3	4.0
6	10.1	0.0	4.4
7	12.3	0.0	5.0
8	11.3	0.0	4.3
9	10.0	-0.3	4.5
10	8.2	0.0	3.2
11	8.1	0.0	3.3
12	8.1	0.0	3.2

Table S2. Number of days in each category by year, May-September only.

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	133	131	136	130	130	108	123	116	133	130	134	135	132	1671
Cross-val	14	15	14	15	15	14	16	13	14	16	13	12	14	185
Smoke	3	6	3	5	4	18	10	20	4	5	2	4	6	90

Table S3. Average MDA8 in each category

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	57.2	44.2	46.2	51.8	47.2	56.5	53.9	51.1	46.5	46.7	44.9	47.6	45.8	49.1
Cross-val	57.6	47.9	45.1	45.9	50.9	57.1	59.1	51.3	50.5	45.0	48.6	49.1	45.7	50.3
Smoke	70.0	75.3	39.3	66.4	60.3	54.9	73.8	60.9	63.8	68.6	41.0	50.5	58.2	61.4

Table S4. Average surface Tmax (ºF) in each category

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	94.0	89.0	91.4	90.8	93.7	96.7	94.0	93.0	91.0	92.7	91.6	90.9	94.2	92.5
Cross-val	92.9	88.5	92.9	89.9	93.8	95.9	92.4	93.5	90.1	91.9	92.1	93.5	93.0	92.3
Smoke	98.0	90.5	93.3	97.8	96.0	94.7	97.9	98.4	94.3	90.6	90.5	94.3	101.8	96.0

Table S5. Number of days included in the GAM

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	125	128	130	128	128	108	121	116	133	127	133	134	131	1642
Cross-val	12	13	14	15	15	14	16	13	14	16	12	12	14	180
Smoke	2	6	2	5	4	18	10	20	4	5	2	4	6	88

Table S6. Average of residuals from the GAM

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	0.44	0.25	-2.67	3.15	-1.03	-1.03	1.49	-0.75	-0.08	0.25	-0.49	0.74	-0.36	0.00
Cross-val	-0.64	0.86	-5.72	0.21	-0.96	-1.73	4.98	-1.44	3.38	0.66	0.11	3.34	0.59	0.31
Smoke	3.87	13.22	-2.95	6.31	2.06	1.22	9.35	4.16	7.39	16.33	6.90	7.66	2.23	5.55

Detroit CSA

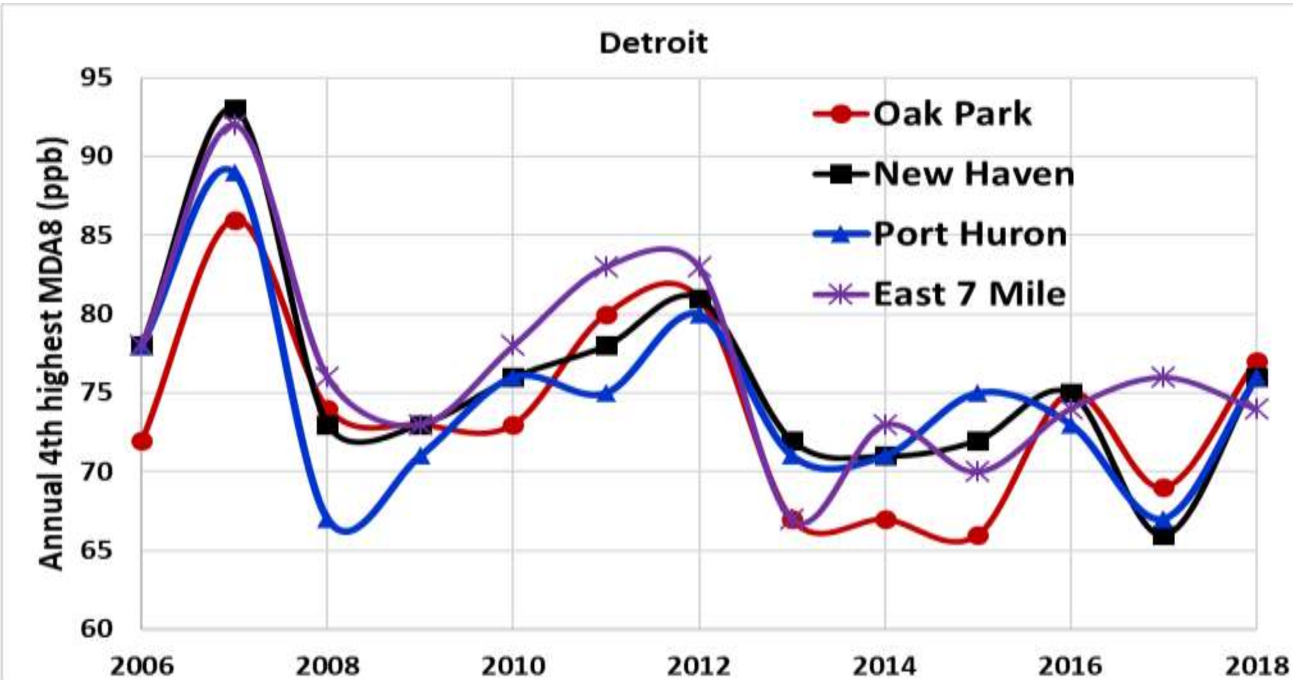

Figure S1. Trend in 4th highest MDA8 O_3 for each site.

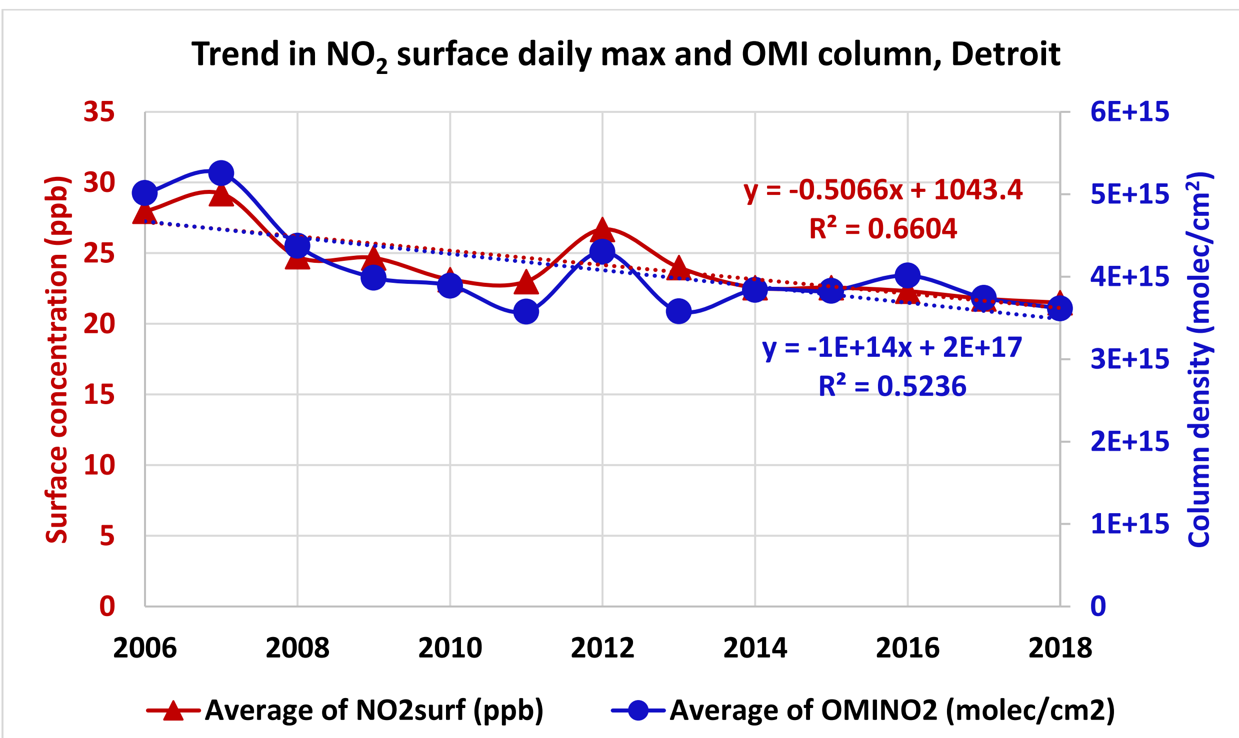

Figure S2. Trend in surface NO_2 daily max and OMI column, 2006-2018.

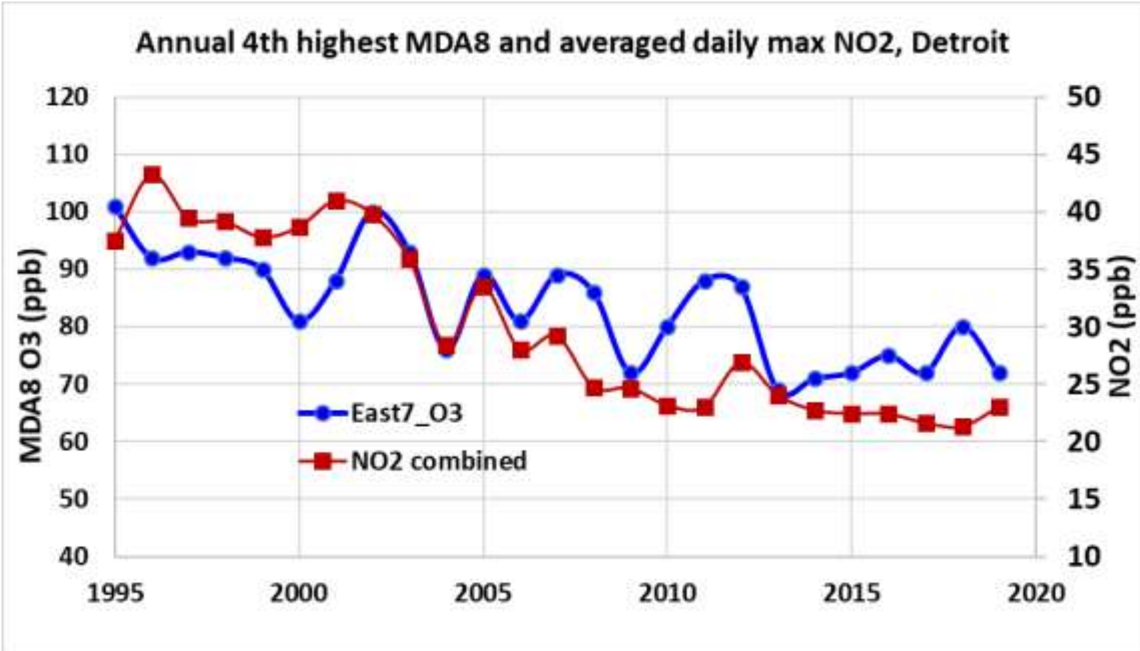

Figure S3. Trend in 4[th] highest O_3 (East 7 mile site) and daily max NO_2, 1995-2019. For NO_2, we use a combination of sites as East 7 mile was not available for 2019 at this time. The sites used were: AQS ids: 261630015, 261630016, 261630019, 261630093, 261630094, 261630095, 261630098, 261630099, 261630100, 261631010, 261631011.

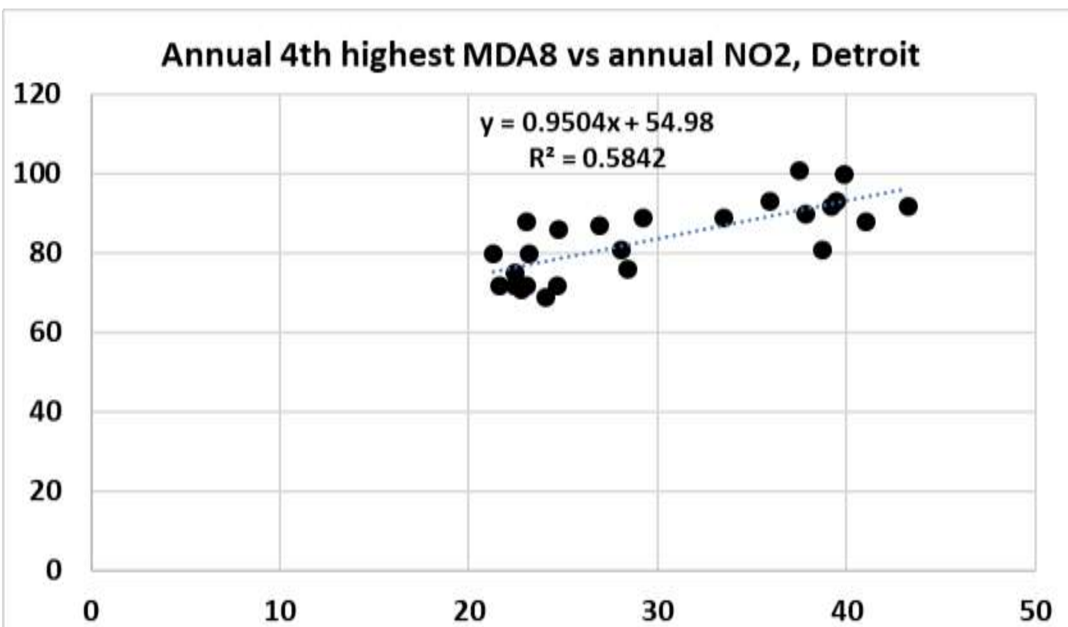

Figure S4. O_3 vs NO_2, 1995-2019 using same data as in previous figure.

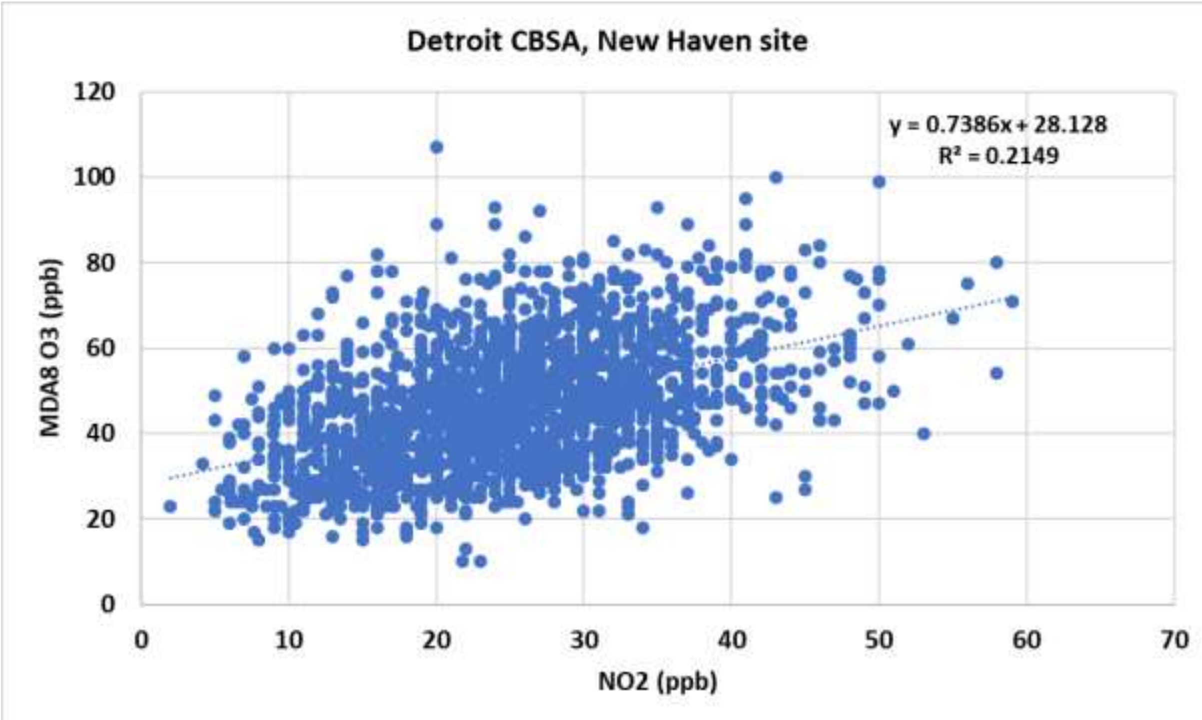

Figure S5. MDA8 vs daily max NO_2 for New Haven site.

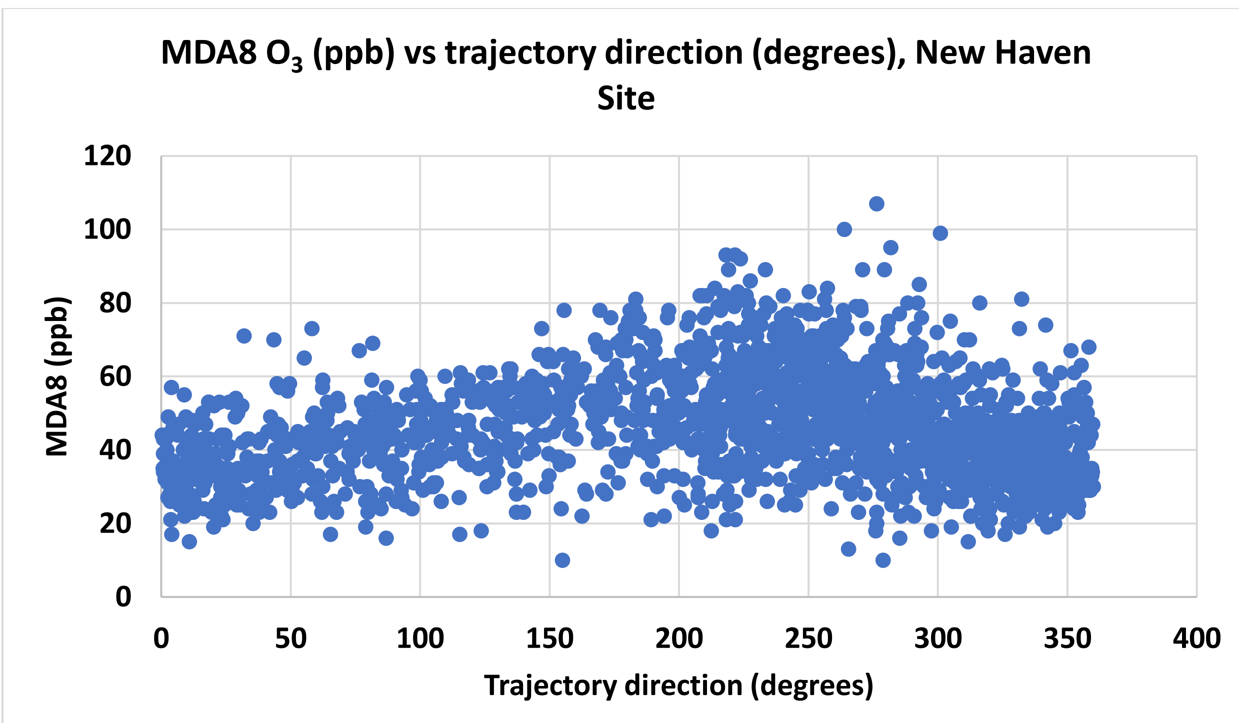

Figure S6. MDA8 vs trajectory direction for New Haven site.

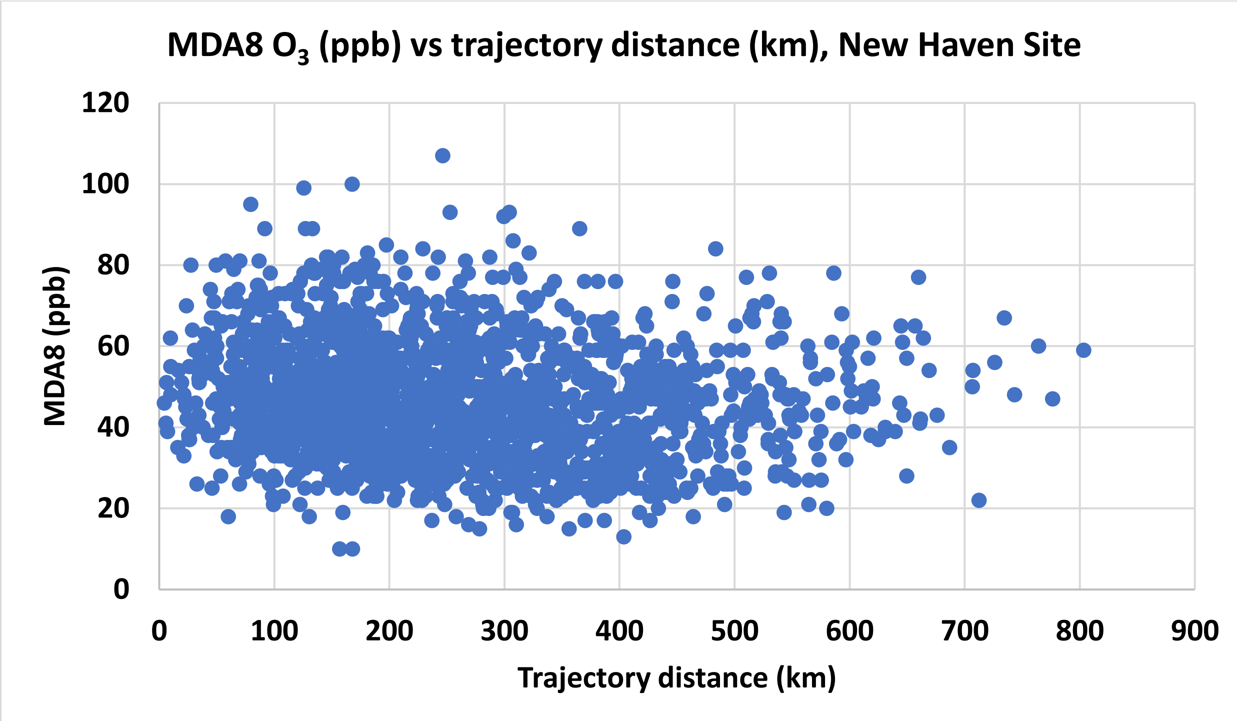

Figure S7. MDA8 vs trajectory distance for New Haven site.

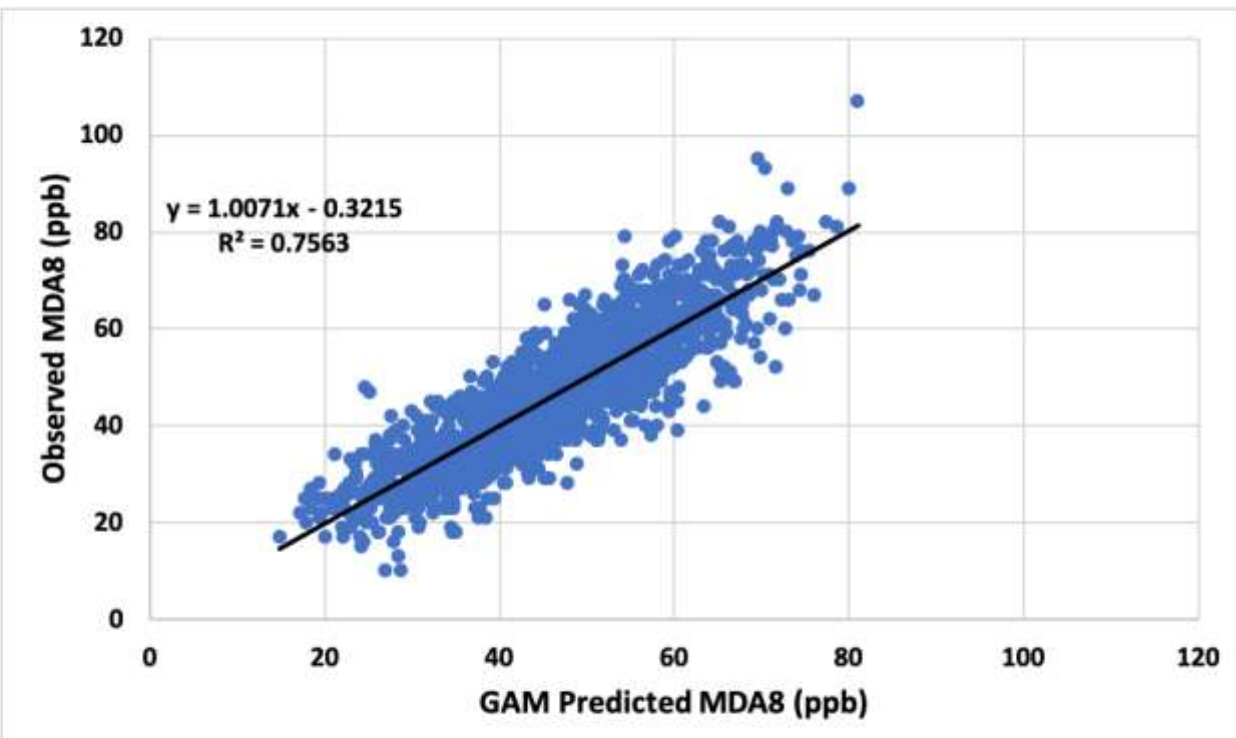

Figure S8. GAM predicted vs Observed, training data only for New Haven site.

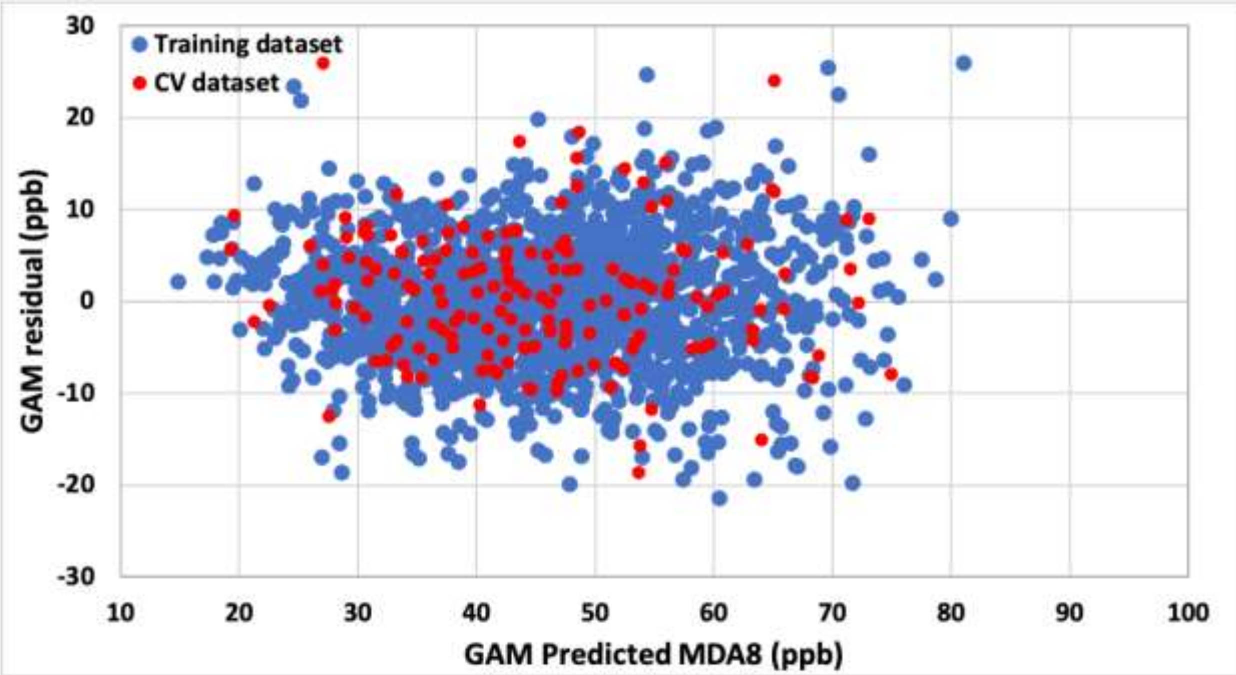

Figure S9. GAM residuals vs predicted, training and cross-validation data for New Haven site.

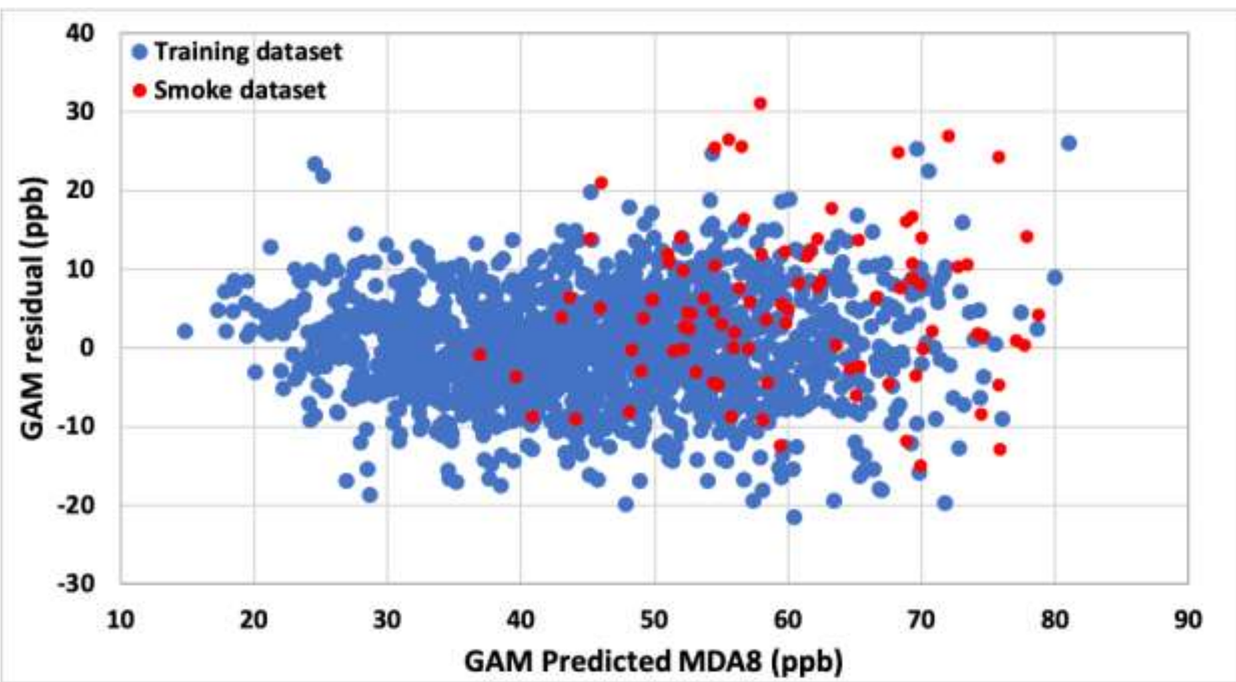

Figure S10. GAM residuals vs predicted, training and smoke data for New Haven site.

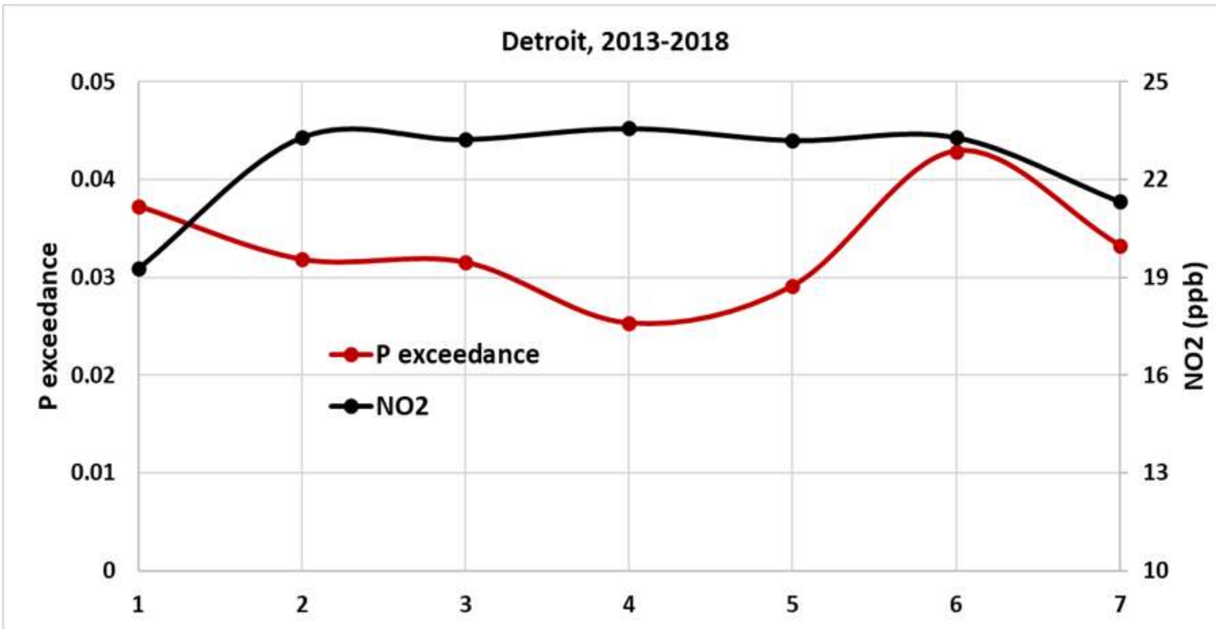

Figure S11. Probability (P) of an O_3 exceedance day (MDA8>70 ppb) across all sites that we considered and averaged surface NO_2 concentrations as a function of day of week (Sunday=1, Monday =2, etc). Both are calculated using May-September 2013-2018 data.

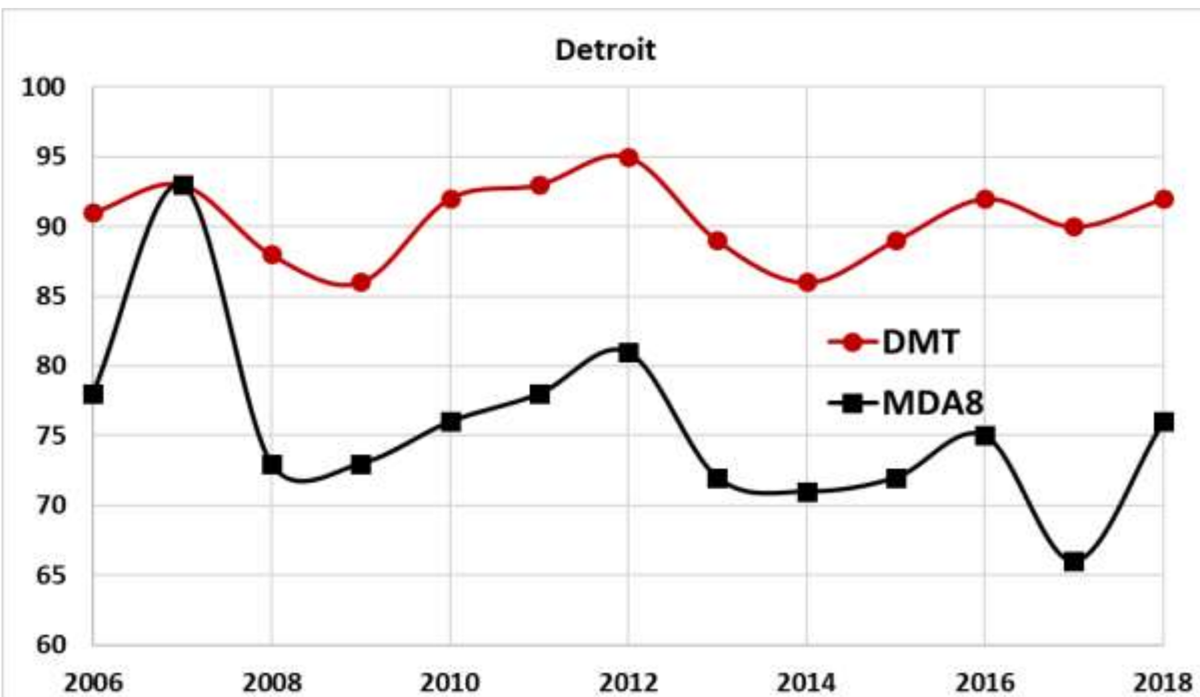

Figure S12. Annual 98th percentile of daily maximum temperature (°F) and fourth highest MDA8 (ppb) for New Haven site.

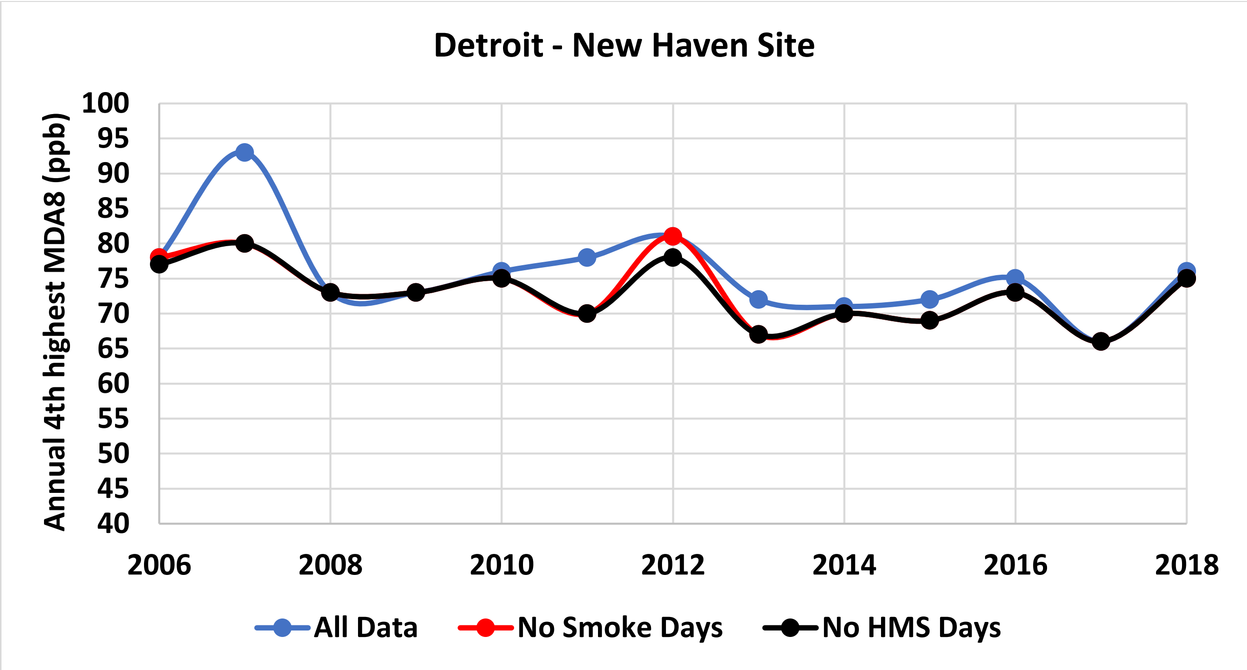

Figure S13. Annual fourth highest for all data, smoke days excluded, and HMS days excluded for New Haven site.

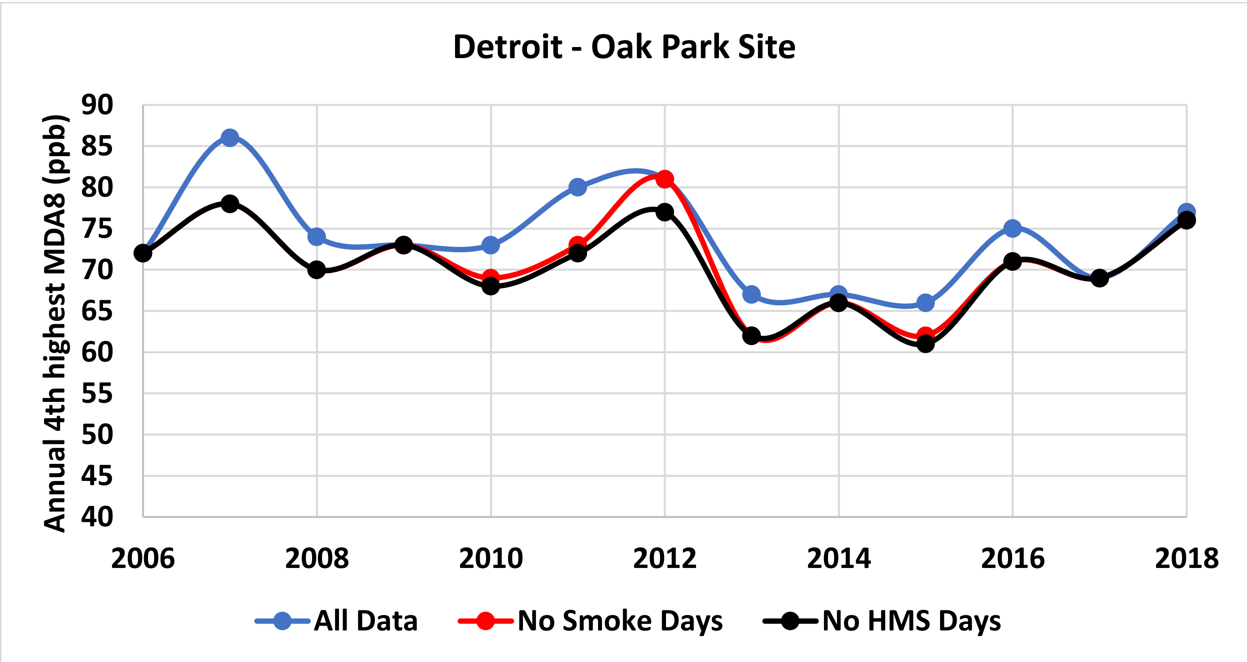

Figure S14. Annual fourth highest for all data, smoke days excluded, and HMS days excluded for Oak Park site.

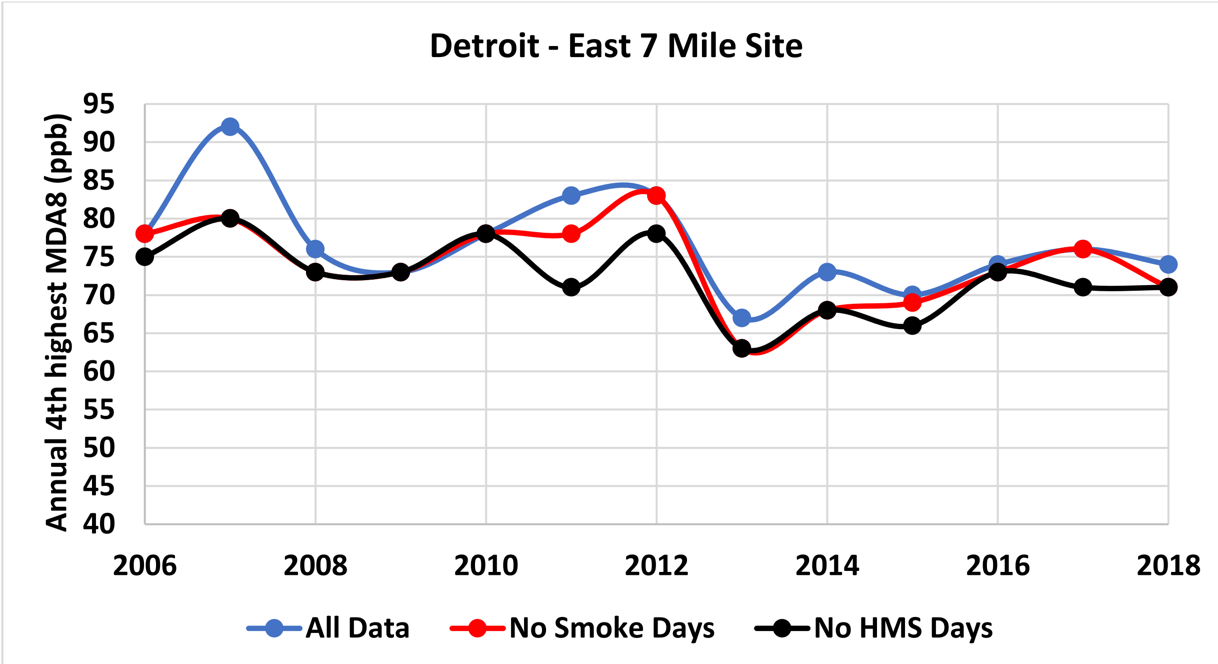

Figure S15. Annual fourth highest for all data, smoke days excluded, and HMS days excluded for East 7 Mile site.

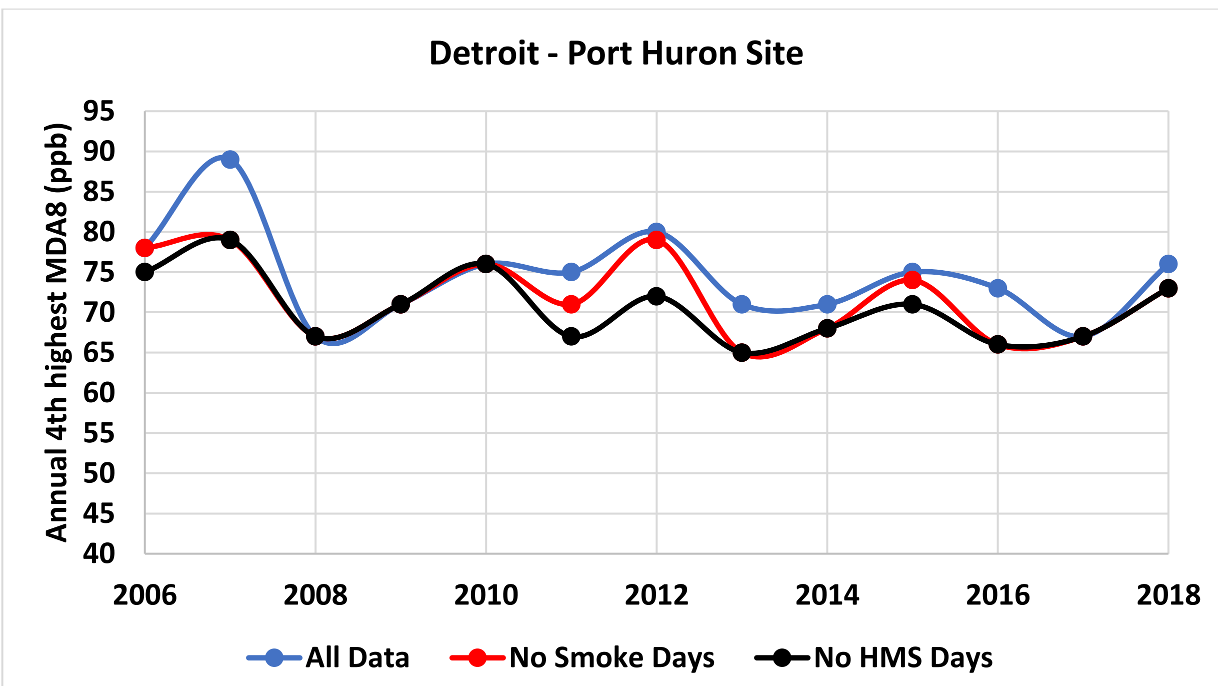

Figure S16. Annual fourth highest for all data, smoke days excluded, and HMS days excluded for Port Huron site.

Table S1. Average, standard deviation and trend in $PM_{2.5}$ for DET by month for 2006-2018 (μg m^{-3}). Only statistically significant trends are shown. If the trend is not statistically significant, the slope is set to 0. The standard deviation is based on the variation in daily $PM_{2.5}$ observations.

Month	Avg	Slope	SD
1	11.8	-0.3	6.0
2	11.7	-0.5	6.5
3	10.2	-0.5	6.4
4	8.3	-0.2	4.1
5	9.2	0.0	4.8
6	9.7	-0.3	4.4
7	11.9	-0.4	5.5
8	10.8	-0.3	5.7
9	9.5	0.0	5.2
10	8.4	0.0	4.4
11	11.2	-0.5	7.0
12	11.4	-0.5	6.4

Detroit CSA -New Haven

Table S2. Number of days in each category by year, May-September only.

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	135	124	131	131	130	128	124	127	128	130	135	135	125	1683
Cross-val	15	13	15	16	14	13	16	13	14	13	14	15	13	184
Smoke	2	15	6	5	5	10	9	11	7	9	3	1	11	94

Table S3. Average MDA8 in each category (ppb)

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	45.2	47.1	45.7	42.2	47.6	44.7	50.7	43.3	44.7	46.5	46.8	44.5	44.8	45.7
Cross-val	48.3	49.9	40.4	44.2	47.7	43.5	48.6	46.1	43.6	45.1	42.3	46.9	48.8	45.8
Smoke	62.5	74.9	57.2	54.6	69.0	73.0	68.1	64.0	58.9	62.3	73.0	56.0	56.3	65.1

Table S4. Average surface Tmax (ºF) in each category

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	75.6	78.1	75.7	74.2	77.7	75.9	78.9	75.1	74.9	77.0	78.7	76.1	78.3	76.6
Cross-val	78.6	76.3	75.1	73.7	75.4	74.5	76.8	76.5	71.9	73.7	77.1	77.5	80.6	76.0
Smoke	85.5	87.5	82.2	81.2	87.6	90.4	92.1	85.0	83.3	83.2	86.7	91.0	88.5	86.7

Table S5. Number of days included in the GAM

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	125	111	124	120	114	125	124	127	124	129	135	133	125	1616
Cross-val	13	10	15	15	13	13	16	12	14	13	14	15	13	176
Smoke	2	15	6	5	4	10	9	11	7	9	3	1	11	93

Table S6. Average of residuals

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	0.22	-1.15	0.81	-0.89	0.99	-0.40	0.01	-0.09	0.96	0.53	-0.63	0.19	-0.59	0.00
Cross-val	0.97	1.62	-2.46	1.13	4.50	0.16	1.58	0.60	1.90	0.92	-0.88	1.09	-1.08	0.74
Smoke	-0.76	10.54	6.09	6.91	9.34	9.09	-1.11	3.26	5.97	6.30	3.41	0.06	-0.46	5.22

Detroit CSA -Oak Park

Table S7. Number of days in each category by year, May-September only.

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	131	124	131	131	131	125	122	123	128	123	135	136	121	1661
Cross-val	15	13	15	16	15	13	15	14	14	11	14	14	13	182
Smoke	2	15	6	5	5	10	9	11	6	8	3	1	8	89

Table S8. Average MDA8 in each category (ppb)

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	43.5	46.5	45.3	43.3	46.3	45.5	51.8	42.6	43.3	44.7	46.3	44.4	45.7	45.3
Cross-val	47.2	45.5	40.1	44.1	46.2	45.7	48.3	42.6	43.9	46.1	42.3	48.8	49.2	45.3
Smoke	59.0	72.1	57.0	56.0	64.4	70.4	65.9	62.4	57.5	60.9	72.3	59.0	56.8	63.9

Table S9. Average surface Tmax (ºF) in each category

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	75.6	78.1	75.7	74.2	77.7	75.9	78.9	75.1	74.9	77.0	78.7	76.1	78.3	76.6
Cross-val	78.6	76.3	75.1	73.7	75.4	74.5	76.8	76.5	71.9	73.7	77.1	77.5	80.6	76.0
Smoke	85.5	87.5	82.2	81.2	87.6	90.4	92.1	85.0	83.3	83.2	86.7	91.0	88.5	86.7

Table S10. Number of days included in the GAM

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	122	111	124	120	115	122	122	123	124	122	135	134	121	1595
Cross-val	13	10	15	15	14	13	15	13	14	11	14	14	13	174
Smoke	2	15	6	5	4	10	9	11	6	8	3	1	8	88

Table S11. Average of residuals

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	0.44	-1.53	0.86	-0.14	-0.16	0.68	0.13	-0.41	0.47	0.00	-0.82	0.16	0.22	0.00
Cross-val	1.74	-0.01	-2.44	1.51	0.84	1.50	0.10	-1.64	2.80	0.22	-1.48	1.83	0.19	0.40
Smoke	-3.51	10.15	5.81	7.70	3.54	8.18	-1.72	4.35	5.10	7.09	5.17	2.70	0.28	5.17

Detroit CSA -East 7 Mile

Table S12. Number of days in each category by year, May-September only.

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	132	124	124	130	122	126	124	122	126	127	128	131	125	1641
Cross-val	14	13	14	16	14	13	16	14	13	12	14	14	13	180
Smoke	2	15	6	5	3	10	8	11	7	9	2		11	89

Table S13. Average MDA8 in each category (ppb)

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	45.5	48.3	46.3	42.5	48.4	45.9	52.0	43.9	46.0	47.5	48.4	49.3	44.2	46.8
Cross-val	48.6	48.3	41.3	43.1	47.7	45.5	48.4	44.2	45.8	47.8	45.9	51.9	48.0	46.6
Smoke	63.5	77.1	60.5	53.4	73.7	73.7	67.9	64.7	65.1	64.0	77.0		55.3	66.5

Table S14. Average surface Tmax (ºF) in each category

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	75.6	78.1	75.7	74.2	77.7	75.9	78.9	75.1	74.9	77.0	78.7	76.1	78.3	76.6
Cross-val	78.6	76.3	75.1	73.7	75.4	74.5	76.8	76.5	71.9	73.7	77.1	77.5	80.6	76.0
Smoke	85.5	87.5	82.2	81.2	87.6	90.4	92.1	85.0	83.3	83.2	86.7	91.0	88.5	86.7

Table S15. Number of days included in the GAM

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	123	111	118	119	113	124	124	122	122	126	128	130	125	1585
Cross-val	12	10	14	15	13	13	16	13	13	12	14	14	13	172
Smoke	2	15	6	5	3	10	8	11	7	9	2		11	89

Table S16. Average of residuals

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	0.35	-0.93	0.50	-1.53	1.21	0.31	0.20	-0.94	0.56	0.41	-0.63	3.51	-3.18	0.00
Cross-val	2.62	-0.29	-2.79	-1.22	2.15	1.24	0.06	-2.45	2.13	1.26	0.09	4.39	-3.30	0.27
Smoke	-1.95	11.45	5.92	4.34	9.81	9.61	-2.90	3.87	8.97	6.54	4.17		-4.38	5.08

Detroit CSA-Port Huron

Table S17. Number of days in each category by year, May-September only.

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	135	124	130	131	132	127	127	126	131	127	134	137	128	1689
Cross-val	15	13	15	16	15	13	16	14	14	13	14	15	14	187
Smoke	2	15	6	5	5	10	9	11	7	9	3	1	11	94

Table S18. Average MDA8 in each category (ppb)

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	41.4	43.2	42.5	39.0	44.5	40.9	46.9	39.9	41.8	46.4	43.1	43.0	43.0	42.7
Cross-val	45.7	47.1	36.6	40.8	41.9	40.2	44.2	44.3	42.4	42.8	36.8	46.2	48.3	42.8
Smoke	67.0	70.5	48.7	52.0	64.6	70.4	62.6	62.4	58.7	64.2	70.7	46.0	56.2	62.6

Table S19. Average surface Tmax (ºF) in each category

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	75.6	78.1	75.7	74.2	77.7	75.9	78.9	75.1	74.9	77.0	78.7	76.1	78.3	76.6
Cross-val	78.6	76.3	75.1	73.7	75.4	74.5	76.8	76.5	71.9	73.7	77.1	77.5	80.6	76.0
Smoke	85.5	87.5	82.2	81.2	87.6	90.4	92.1	85.0	83.3	83.2	86.7	91.0	88.5	86.7

Table S20. Number of days included in the GAM

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	125	111	123	120	116	125	127	126	127	126	134	136	128	1624
Cross-val	13	11	15	15	14	13	16	14	14	13	14	15	14	181
Smoke	2	15	6	5	4	10	9	11	7	9	3	1	11	93

Table S21. Average of residuals

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	0.15	-0.42	0.64	-1.14	1.07	-0.67	0.45	-0.94	0.04	2.32	-1.26	0.16	-0.32	0.00
Cross-val	1.32	3.54	-2.64	-0.90	-0.09	-0.11	0.93	1.26	1.45	0.07	-3.93	1.30	1.74	0.24
Smoke	3.01	10.36	-1.19	7.94	7.34	10.13	-2.20	4.11	9.15	12.69	2.72	-2.12	-0.65	5.67

NY/NJ CSA

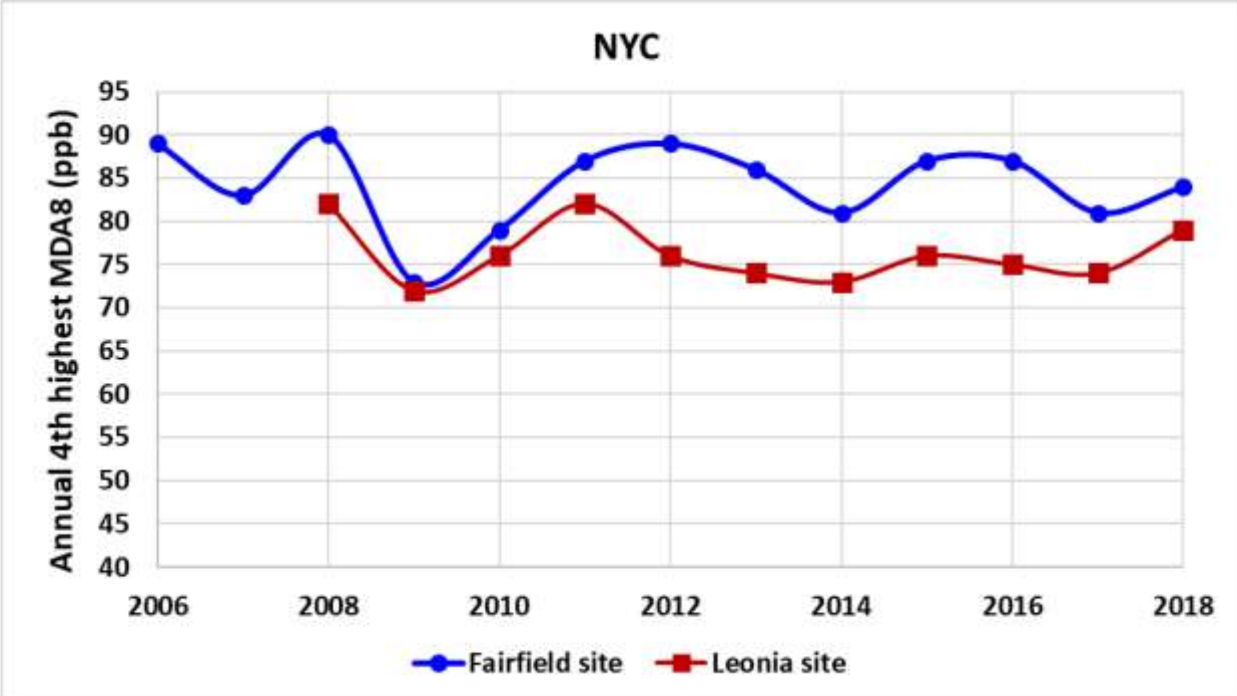

Figure S1. Trend in 4th highest MDA8 O_3 for each site.

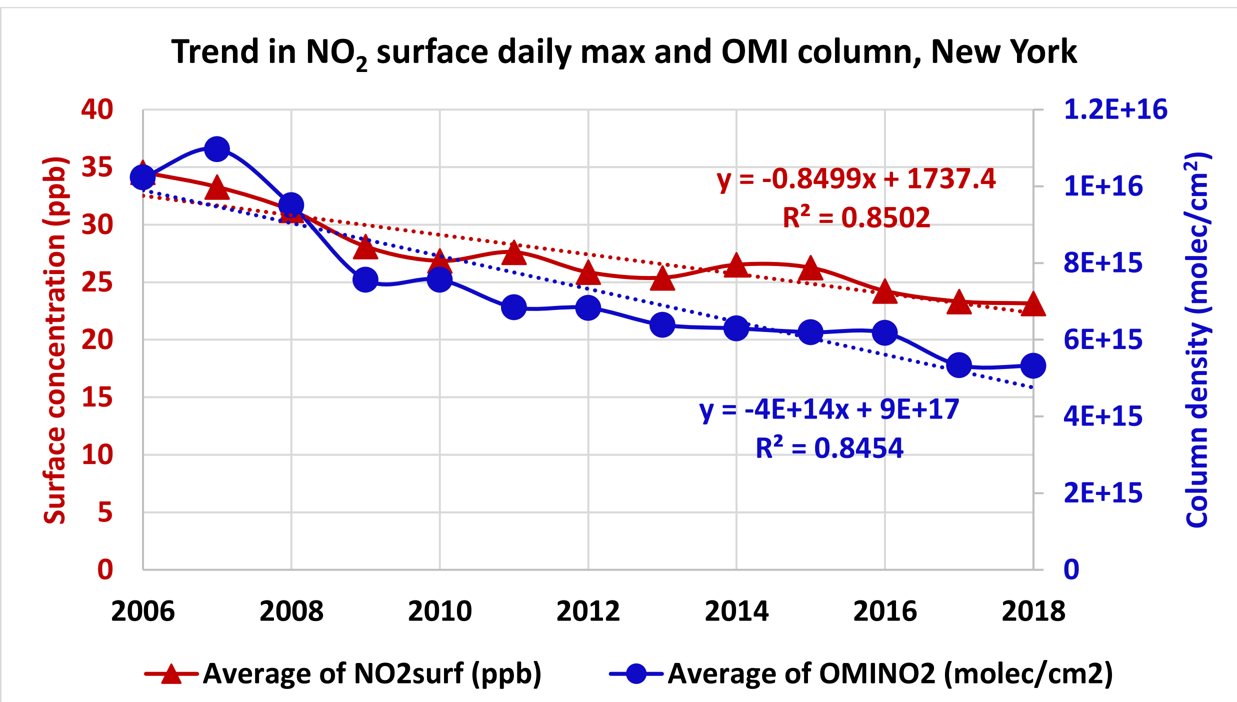

Figure S2. Trend in surface NO_2 daily max and OMI column, 2006-2018.

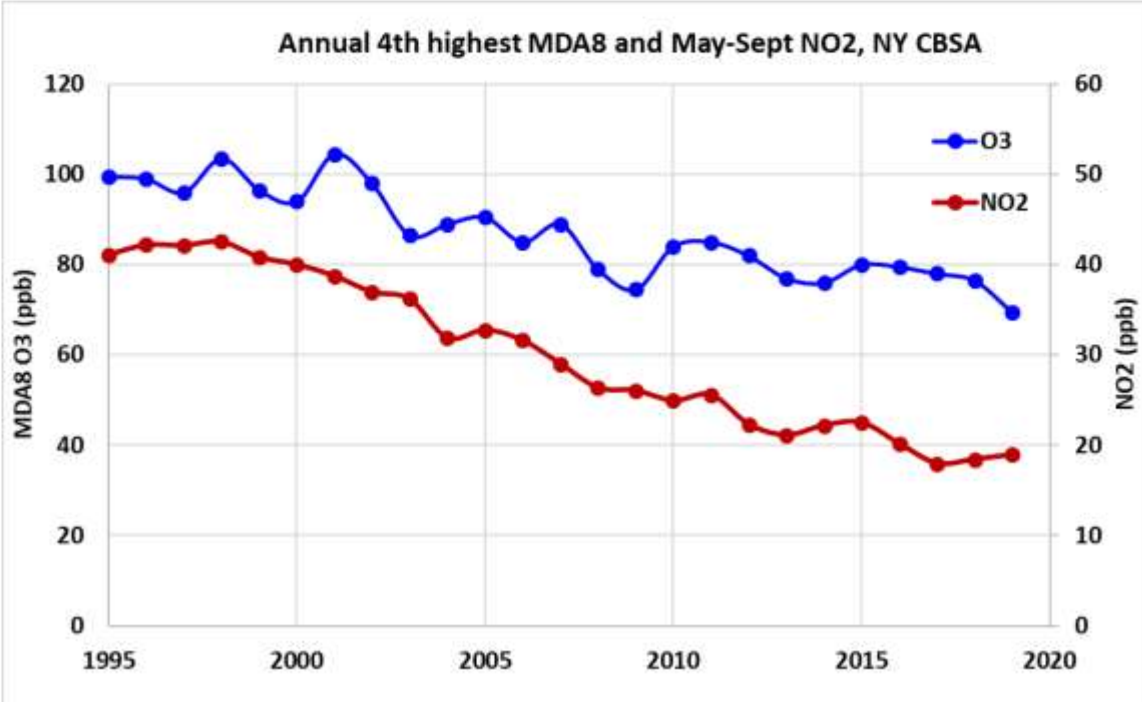

Figure S3. Trend in 4th highest O_3 (two sites) and daily max NO_2, 1995-2019. For this plot and the next I use fourth highest MDA8 averaged from two sites; Fairfield CT and Flemington NJ (AQS i.d. 340190001). For NO_2, I use data from AQS sites: 340131003, 340170006, 340230011, 340273001 and 340390004.

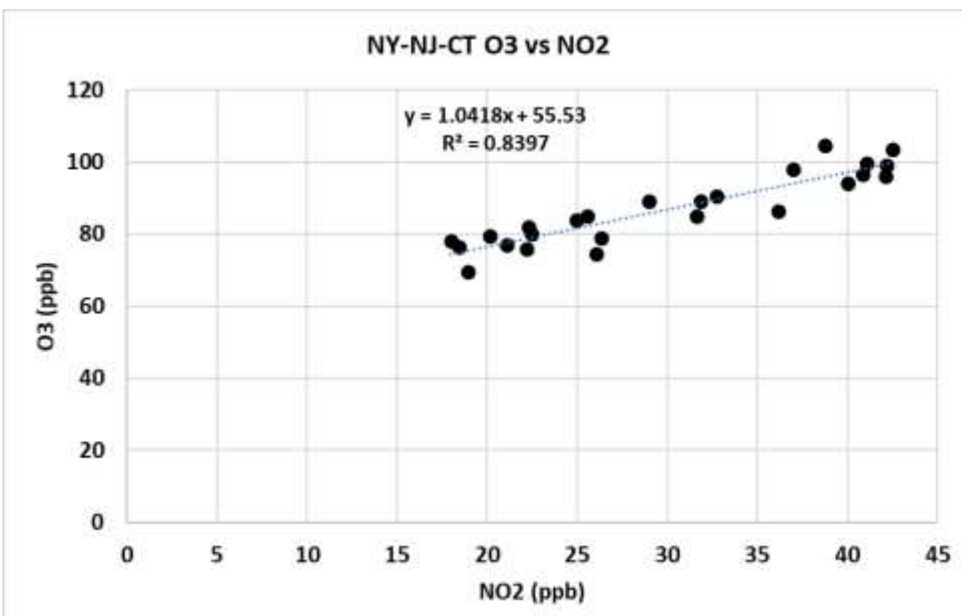

Figure S4. O_3 vs NO_2, 1995-2019.

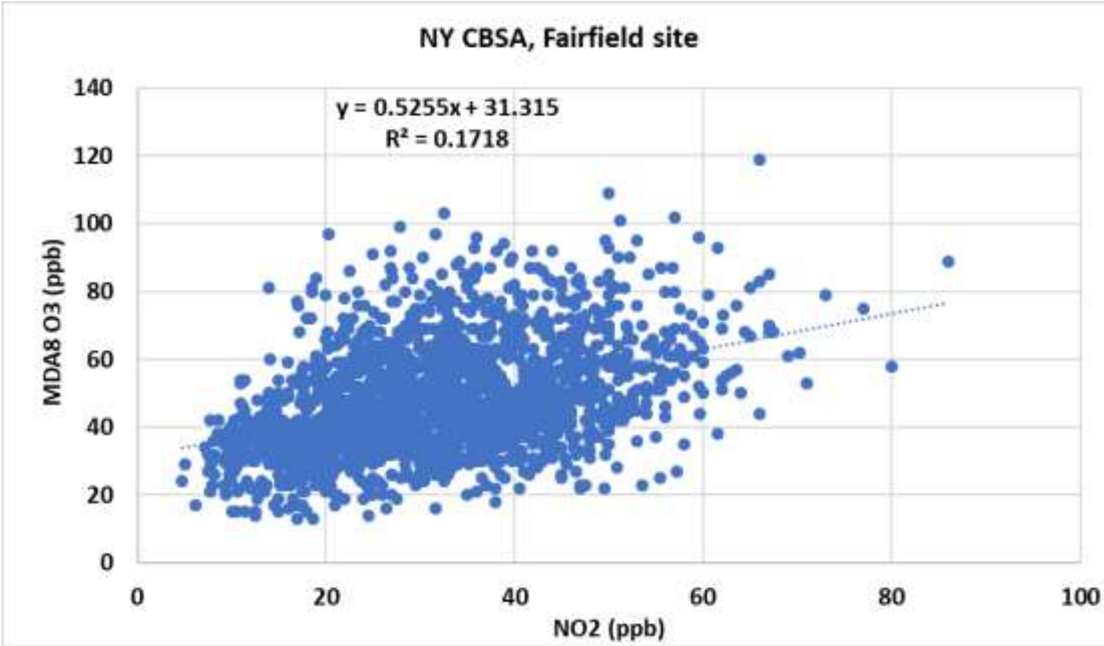

Figure S5. MDA8 vs daily max NO_2 for Fairfield site.

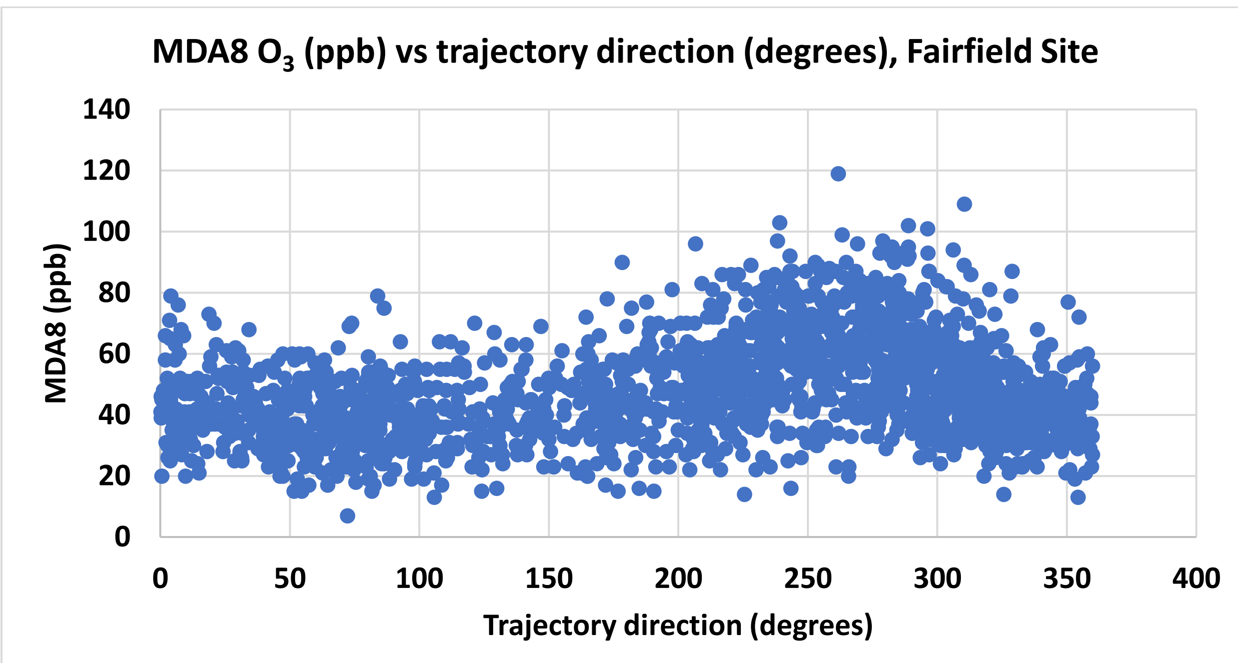

Figure S6. MDA8 vs trajectory direction for Fairfield site.

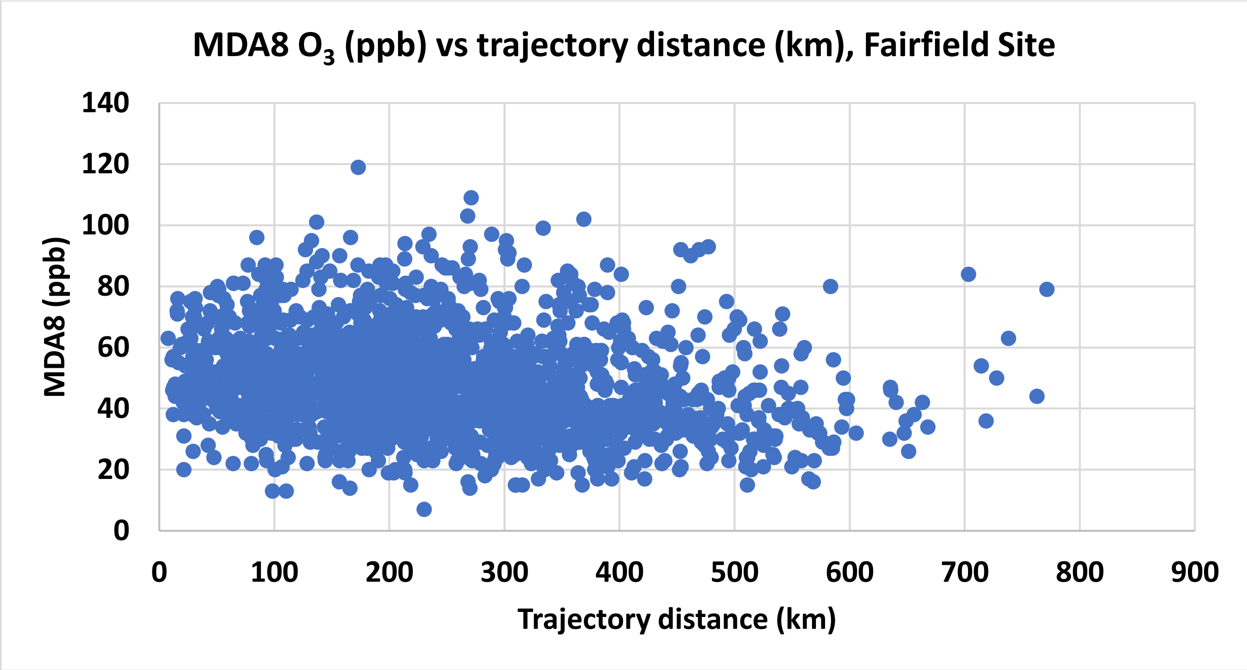

Figure S7. MDA8 vs trajectory distance for Fairfield site.

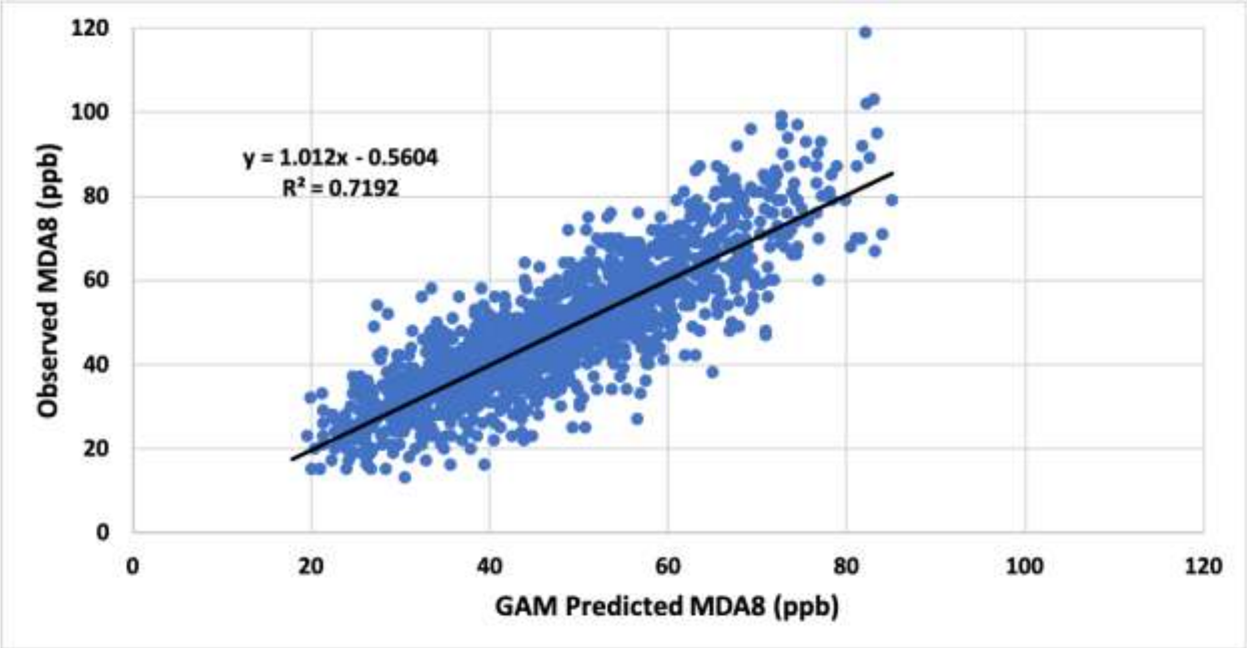

Figure S8. GAM predicted vs Observed, training data only for Fairfield site.

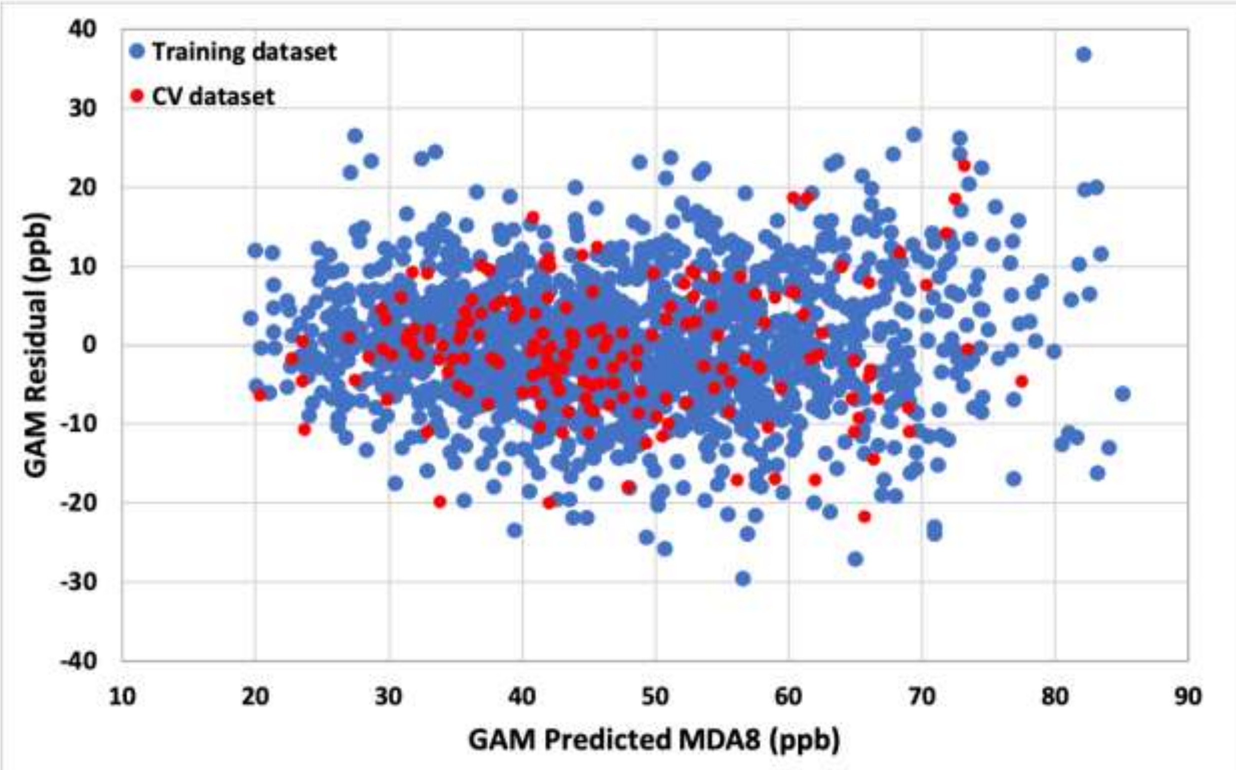

Figure S9. GAM residuals vs predicted, training and cross-validation data for Fairfield site.

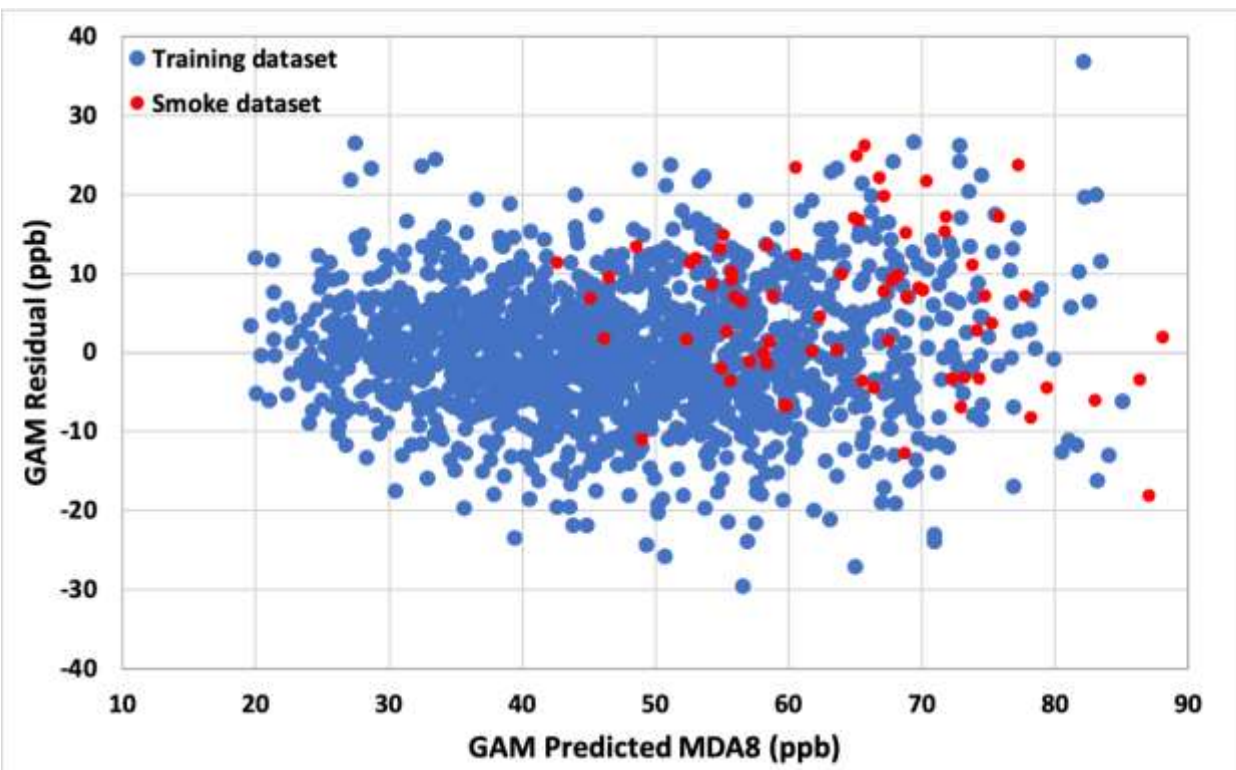

Figure S10. GAM residuals vs predicted, training and smoke data for Fairfield site.

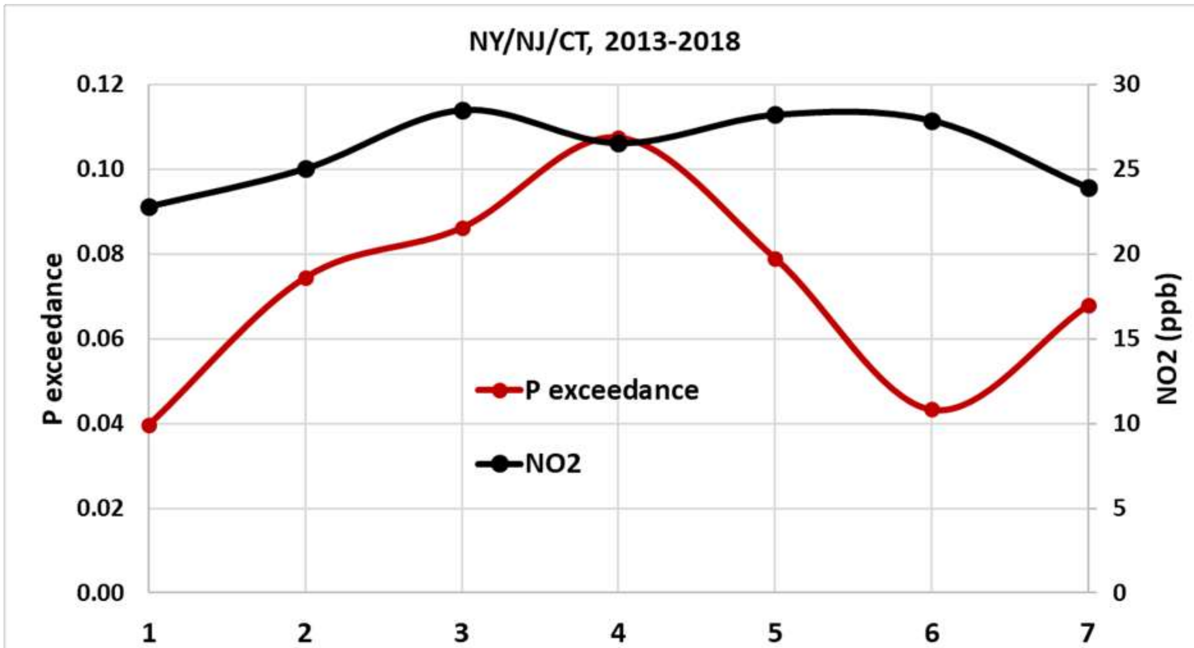

Figure S11. Probability (P) of an O_3 exceedance day (MDA8>70 ppb) across all sites that we considered and averaged surface NO_2 concentrations as a function of day of week (Sunday=1, Monday =2, etc). Both are calculated using May-September 2013-2018 data.

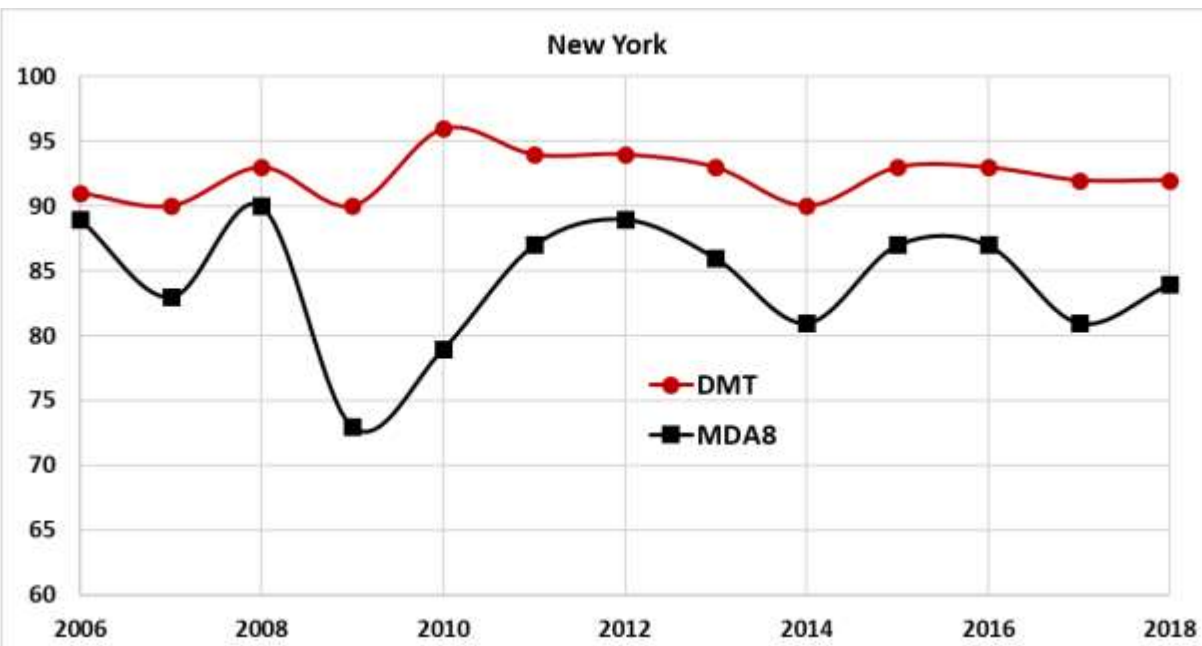

Figure S12. Annual 98th percentile of daily maximum temperature (ºF) and fourth highest MDA8 (ppb) for Fairfield site.

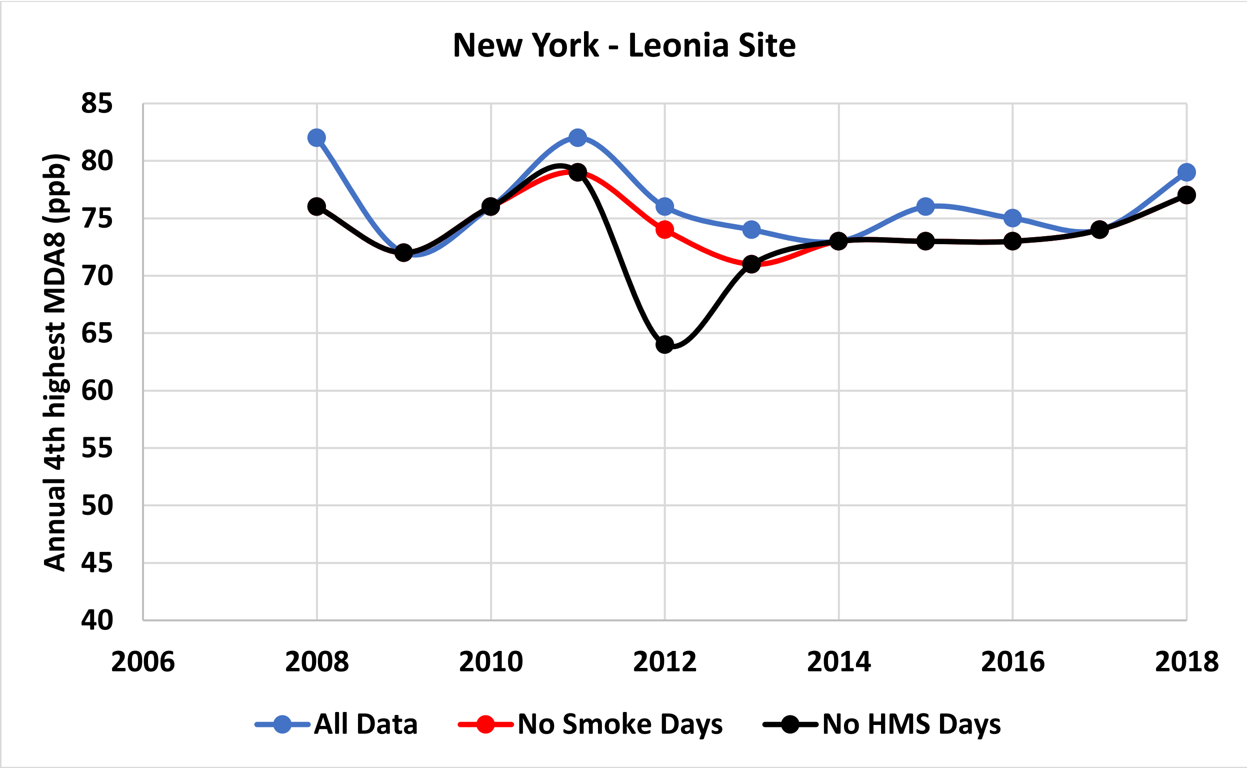

Figure S13. Annual fourth highest for all data, smoke days excluded, and HMS days excluded for Leonia site.

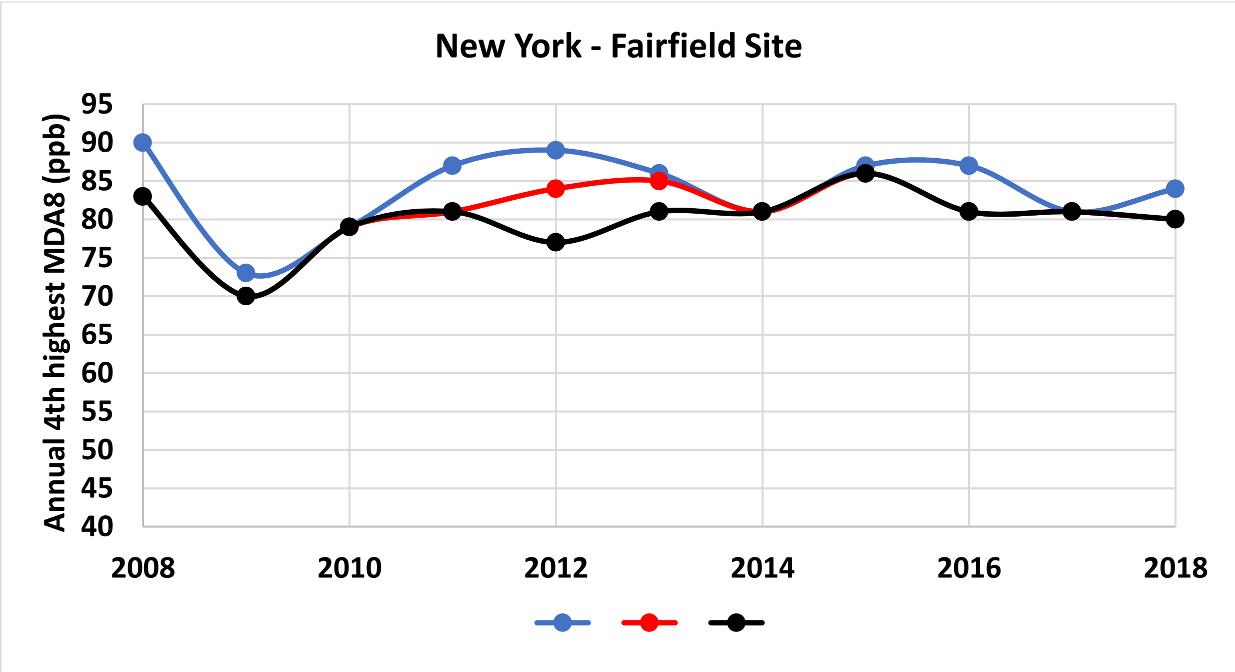

Figure S14. Annual fourth highest for all data, smoke days excluded, and HMS days excluded for Fairfield site.

Table S1. Average, standard deviation and trend in $PM_{2.5}$ for NY by month for 2006-2018 ($\mu g\ m^{-3}$). Only statistically significant trends are shown. If the trend is not statistically significant, the slope is set to 0. The standard deviation is based on the variation in daily $PM_{2.5}$ observations.

Month	Avg	Slope	SD
1	11.8	-0.3	5.1
2	10.8	-0.4	5.2
3	8.5	-0.4	4.8
4	7.5	-0.3	4.0
5	8.3	-0.4	4.9
6	9.2	-0.7	6.3
7	11.5	-1.0	7.2
8	9.1	-0.5	5.7
9	7.3	-0.4	4.7
10	7.6	-0.4	4.4
11	9.0	-0.5	5.2
12	10.3	-0.3	5.0

NY/NJ/CT CSA -Leonia*

Table S2. Number of days in each category by year, May-September only.

	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	128	134	136	127	57	132	133	124	136	134	127	1368
Cross-val	14	16	14	15	7	15	14	14	15	15	15	154
Smoke	3	2	3	11	4	4		10	2	3	10	52

Table S3. Average MDA8 in each category (ppb)

	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	45.2	38.1	45.8	44.5	44.2	44.8	40.5	46.4	45.6	43.0	43.1	43.7
Cross-val	46.8	40.4	42.6	41.9	41.4	45.0	35.6	39.9	44.5	39.3	40.7	41.7
Smoke	78.0	70.0	66.0	68.4	74.0	72.3		62.6	85.5	56.3	66.9	68.2

Table S4. Average surface Tmax (ºF) in each category

	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	78.9	76.0	82.3	79.6	79.1	78.8	79.3	82.2	80.7	78.0	79.5	79.3
Cross-val	79.9	76.9	81.4	79.1	76.1	79.1	78.8	79.3	79.9	76.3	78.7	78.9
Smoke	91.3	87.0	91.0	87.6	90.6	91.8		90.5	89.0	84.7	90.4	89.0

Table S5. Number of days included in the GAM

	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	123	118	134	126	56	131	132	123	134	134	125	1336
Cross-val	13	15	13	15	7	15	14	14	15	15	15	151
Smoke	3	1	3	10	4	4		10	2	3	10	50

Table S6. Average of residuals

	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	-0.60	-1.15	-1.88	1.47	1.58	2.10	-2.46	-1.03	-0.19	1.09	1.95	0.00
Cross-val	2.65	-0.63	-2.92	-1.10	1.65	1.22	-4.49	0.83	1.98	-0.35	2.42	0.07
Smoke	8.94	18.76	1.54	10.44	9.74	2.92		3.59	21.84	4.92	7.85	7.56

*Data for Leonia site start in 2008. There is also a significant amount of missing data in 2012.

NY/NJ/CT CSA-Fairfield

Table S7. Number of days in each category by year, May-September only.

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	132	125	126	132	130	123	125	127	135	122	113	132	126	1648
Cross-val	14	14	14	16	13	15	15	15	14	13	13	15	15	186
Smoke	3	13	3	2	3	11	12	4		10	2	3	8	74

Table S8. Average MDA8 in each category (ppb)

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	47.9	45.5	49.3	43.2	49.1	46.2	48.3	49.2	47.9	49.6	49.8	45.2	43.8	47.3
Cross-val	51.3	49.5	46.0	42.6	47.8	44.3	44.8	49.6	42.1	41.5	46.9	42.5	40.5	45.3
Smoke	72.3	67.0	78.3	72.5	68.7	65.5	75.7	83.3		67.6	88.5	56.7	67.3	70.2

Table S9. Average surface Tmax (ºF) in each category

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	77.6	78.5	78.9	76.0	82.3	79.6	79.1	78.8	79.3	82.2	80.7	78.0	79.5	79.3
Cross-val	80.5	80.2	79.9	76.9	81.4	79.1	76.1	79.1	78.8	79.3	79.9	76.3	78.7	78.9
Smoke	87.3	86.5	91.3	87.0	91.0	87.6	90.6	91.8		90.5	89.0	84.7	90.4	89.0

Table S10. Number of days included in the GAM

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	105	112	118	113	127	122	123	126	134	120	109	131	123	1563
Cross-val	12	13	13	15	12	15	15	15	14	13	13	15	15	180
Smoke	3	11	3	1	3	10	12	4		10	2	3	8	70

Table S11. Average of residuals

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	1.54	-2.74	-1.01	0.22	-2.15	-0.53	1.02	2.35	1.04	-0.90	0.94	0.51	-0.37	0.00
Cross-val	3.92	-0.72	-0.91	-3.21	-2.69	-2.17	3.66	2.25	-1.12	-2.30	-1.33	-1.58	-2.54	-0.69
Smoke	6.34	5.03	2.15	13.61	-4.50	5.50	10.56	9.38		3.76	22.38	0.46	5.58	6.12

Phoenix CSA

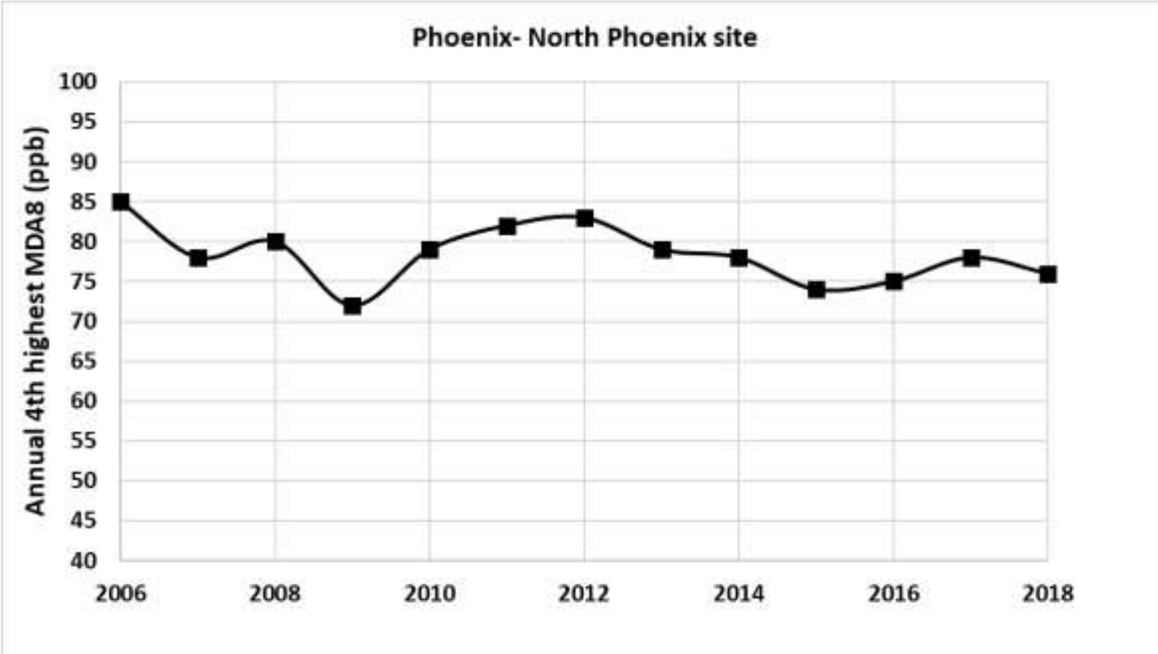

Figure S1. Trend in 4th highest MDA8 O_3 for each site.

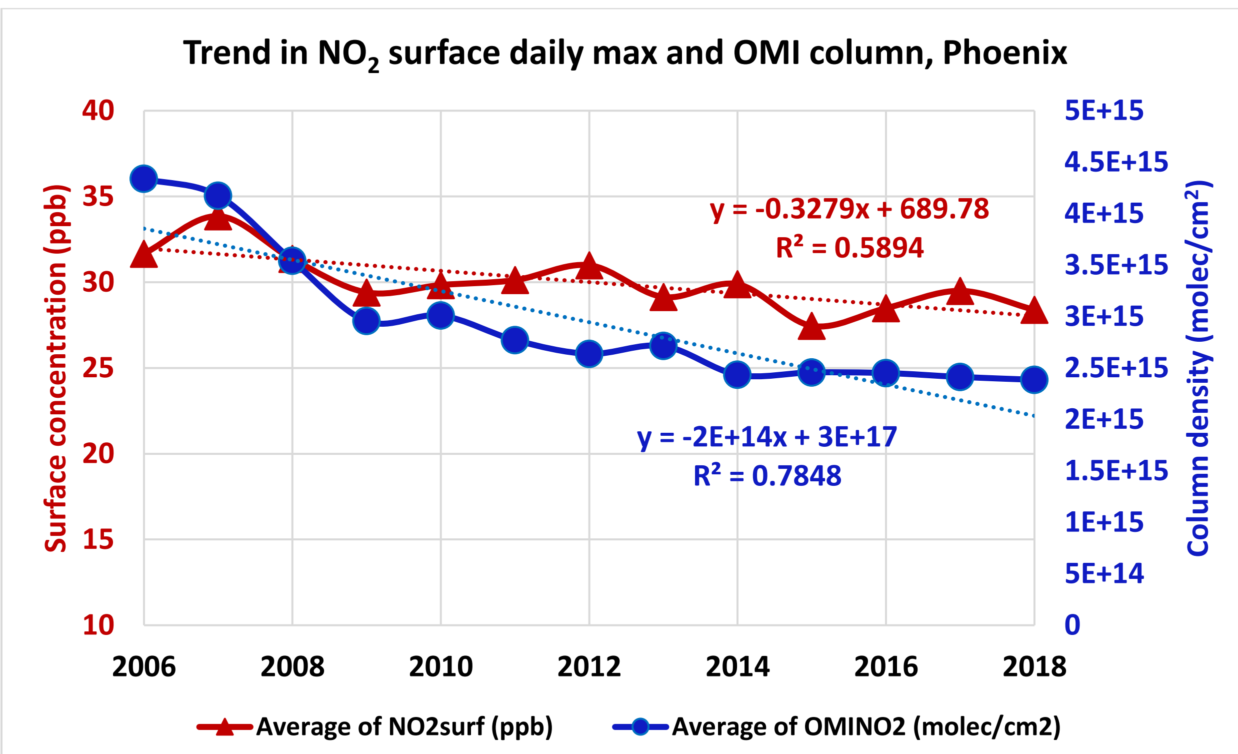

Figure S2. Trend in surface NO_2 daily max and OMI column, 2006-2018.

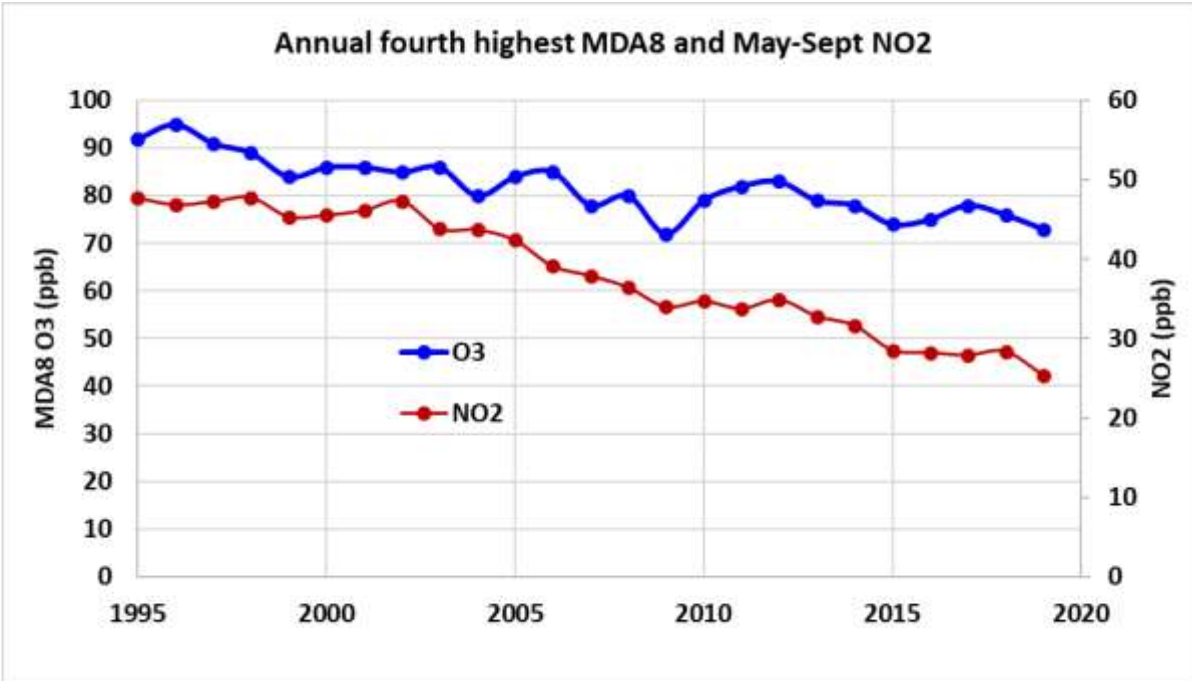

Figure S3. Trend in 4th highest O_3 and daily max NO_2, 1995-2019.

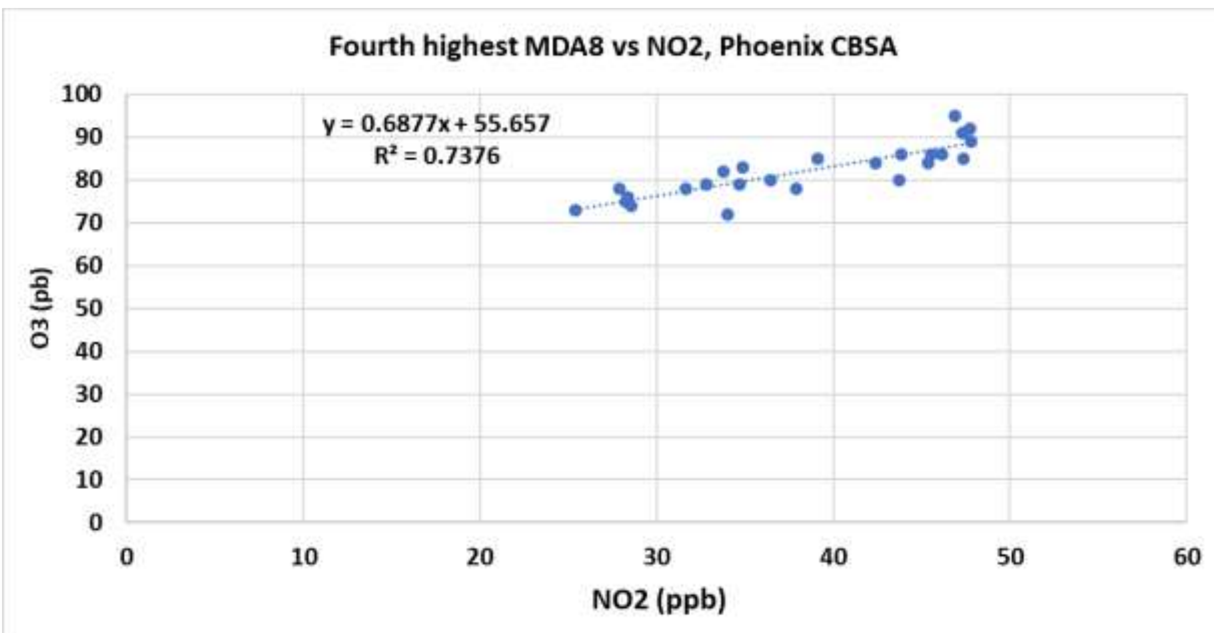

Figure S4. O_3 vs NO_2, 1995-2019.

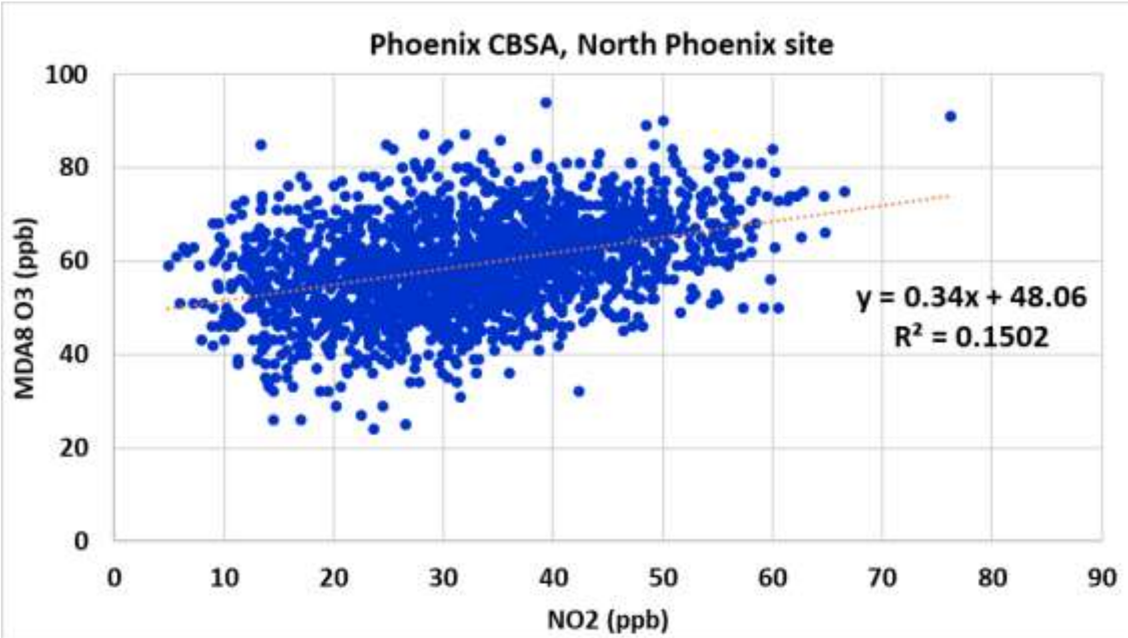

Figure S5. MDA8 vs daily max NO_2 for North Phoenix site.

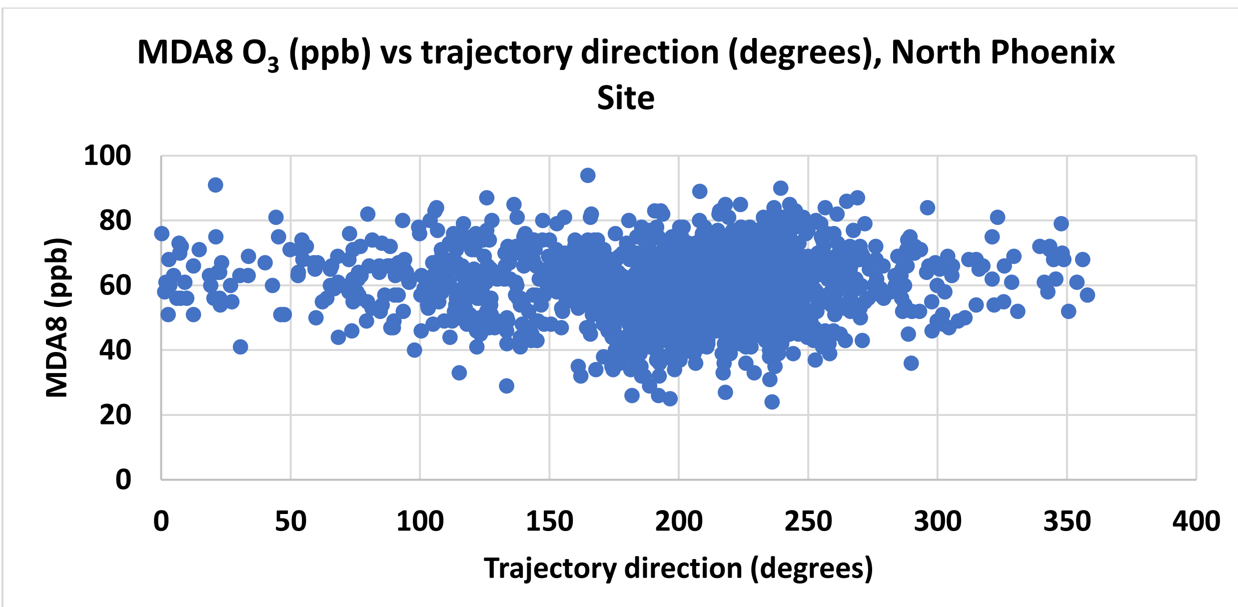

Figure S6. MDA8 vs trajectory direction for North Phoenix site.

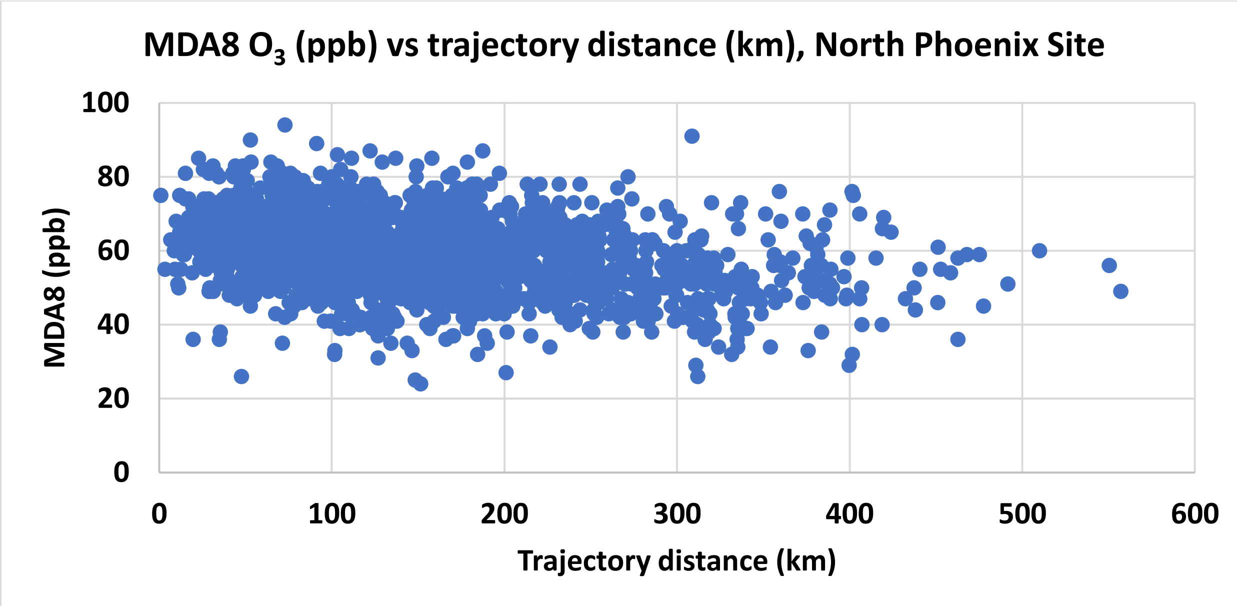

Figure S7. MDA8 vs trajectory distance for North Phoenix site.

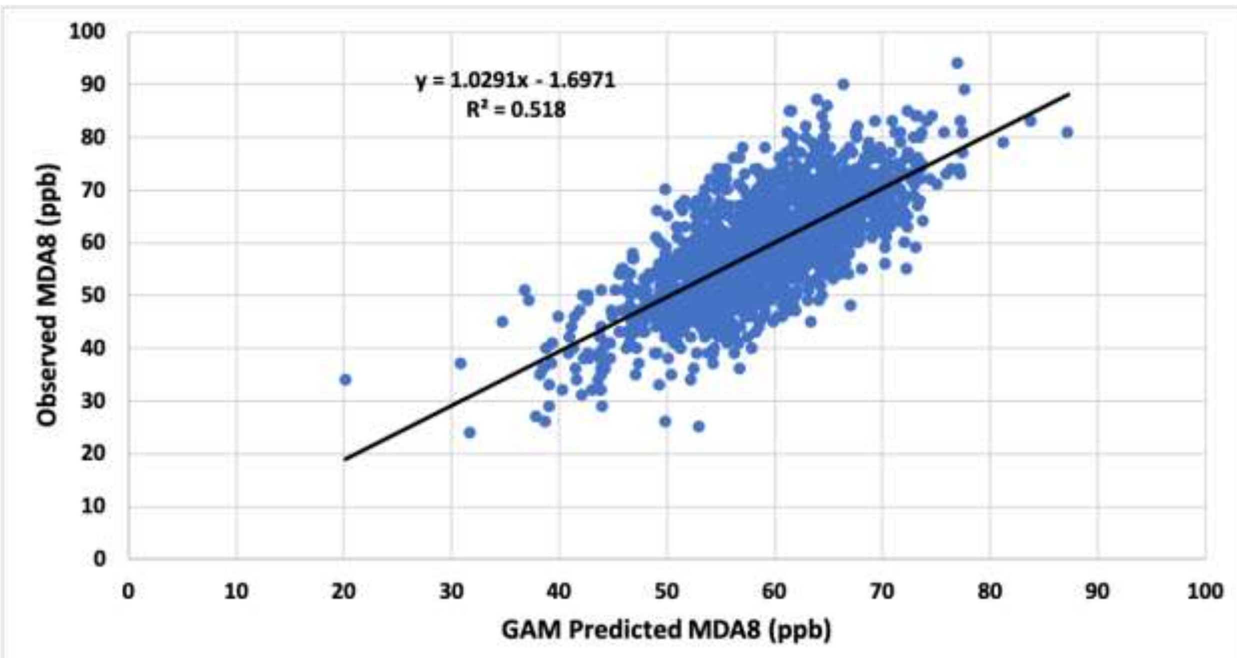

Figure S8. GAM predicted vs Observed, training data only for North Phoenix site.

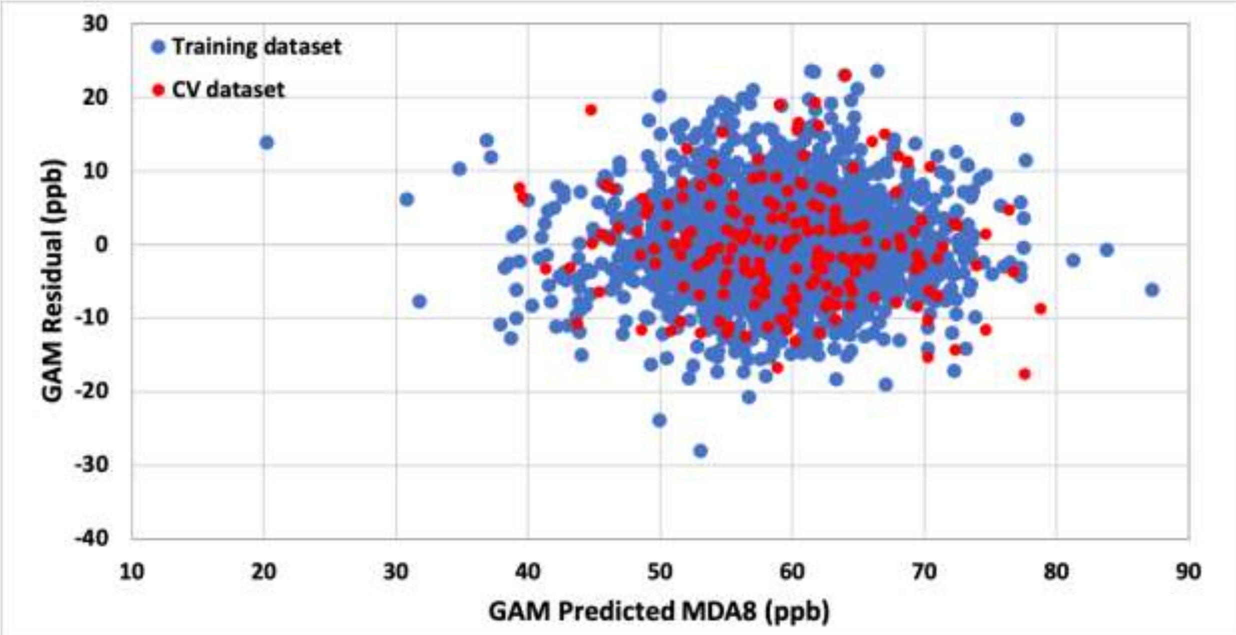

Figure S9. GAM residuals vs predicted, training and cross-validation data for North Phoenix site.

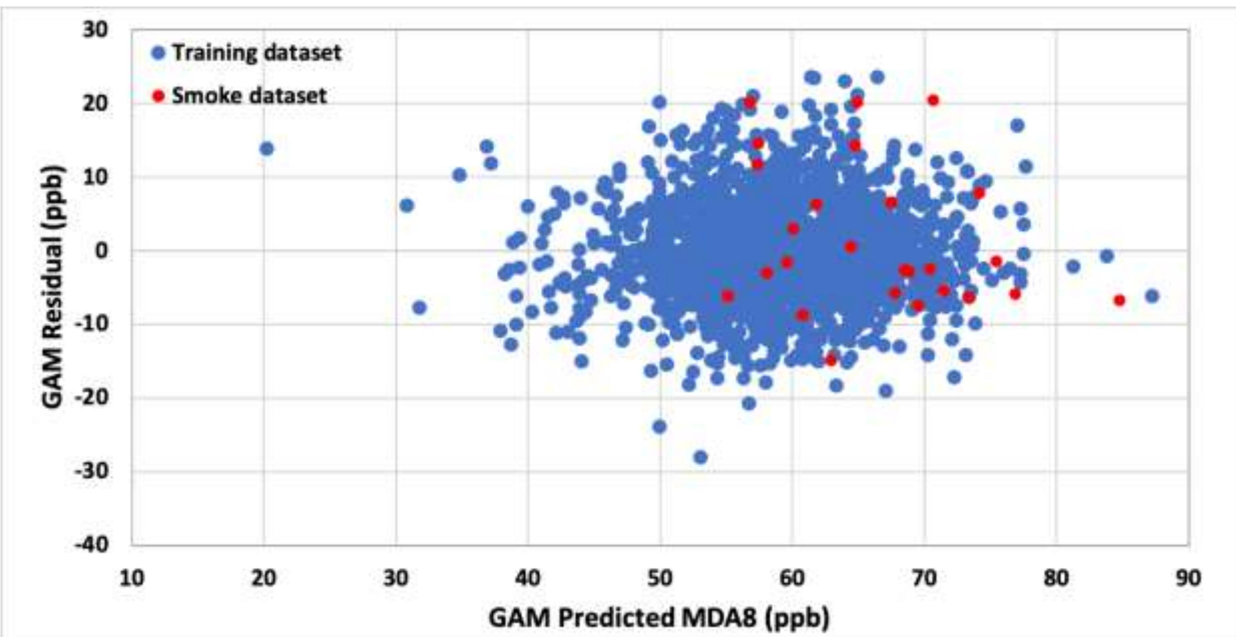

Figure S10. GAM residuals vs predicted, training and smoke data for North Phoenix site.

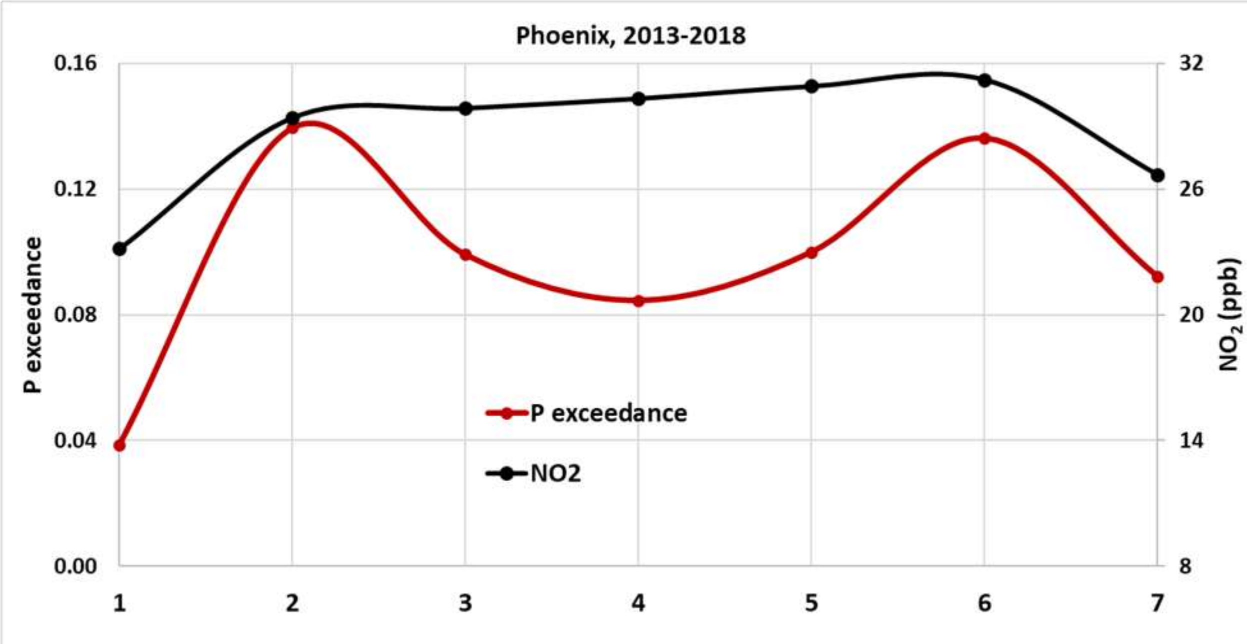

Figure S11. Probability (P) of an O_3 exceedance day (MDA8>70 ppb) across all sites that we considered and averaged surface NO_2 concentrations as a function of day of week (Sunday=1, Monday =2, etc). Both are calculated using May-September 2013-2018 data.

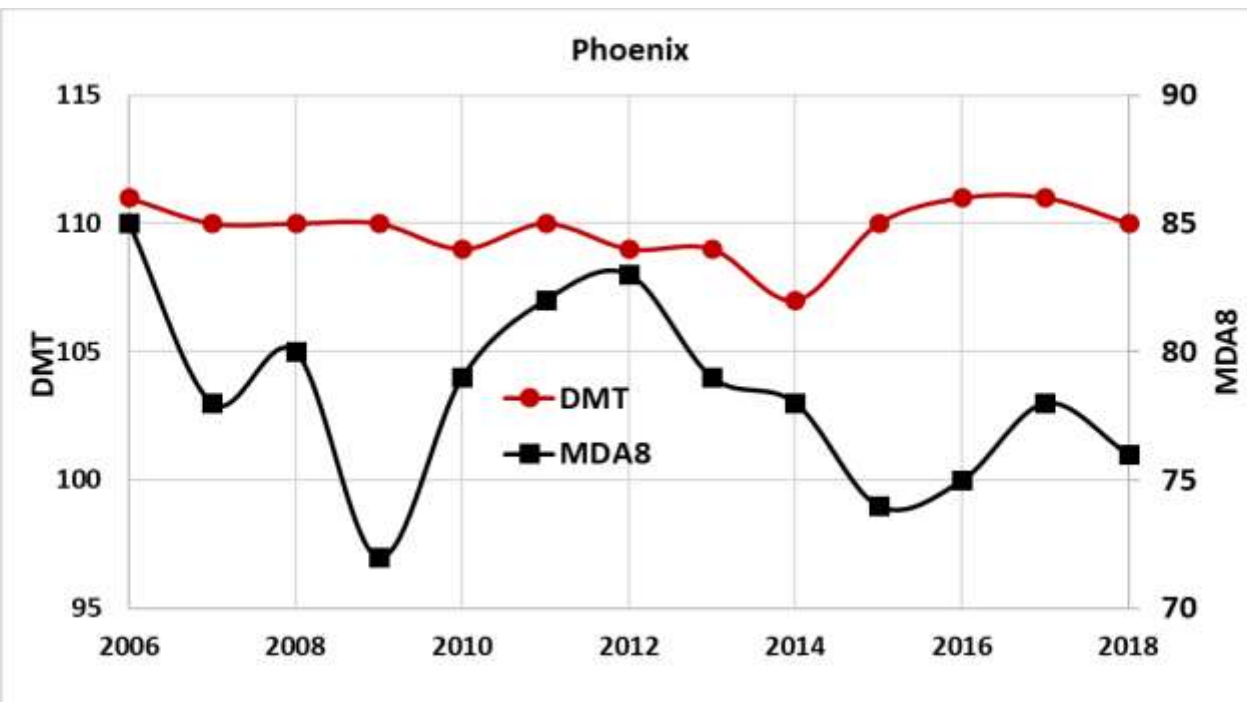

Figure S12. Annual 98th percentile of daily maximum temperature (^{o}F) and fourth highest MDA8 (ppb).

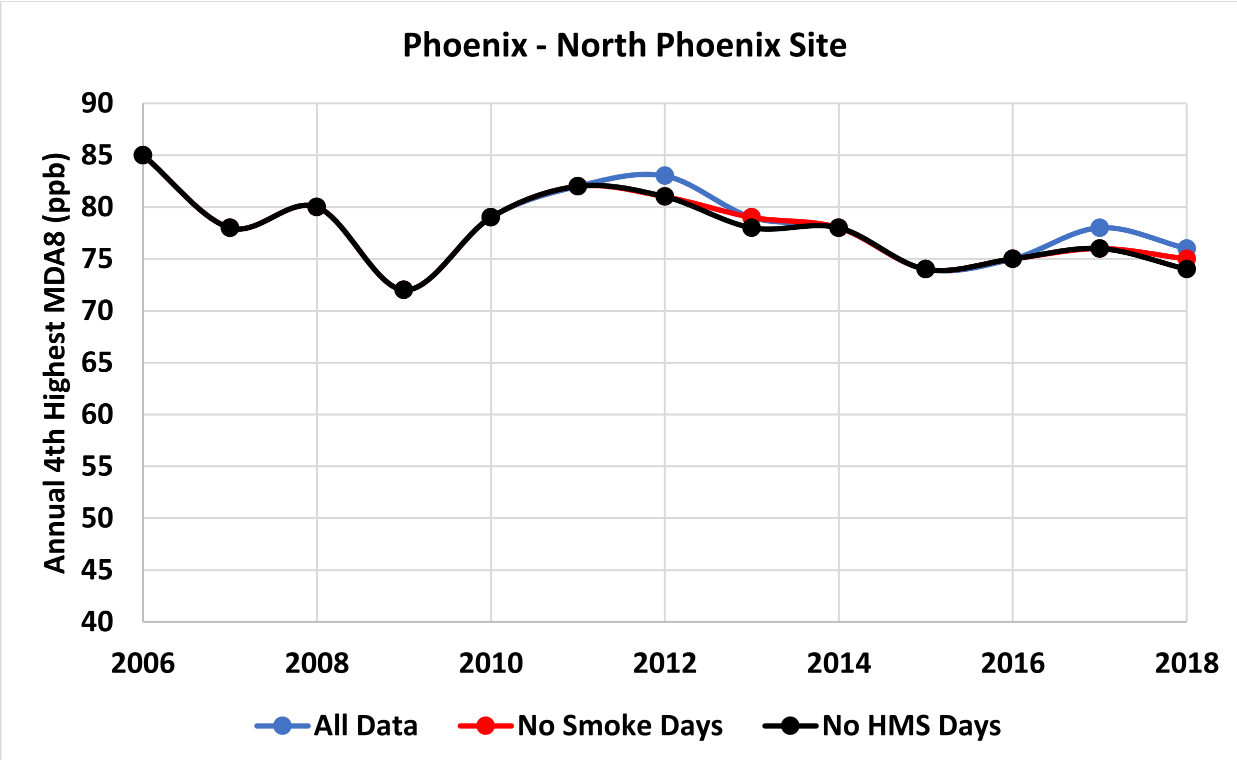

Figure S13. Annual fourth highest for all data, smoke days excluded, and HMS days excluded for North Phoenix site.

Table S1. Average, standard deviation and trend in $PM_{2.5}$ for PHX by month for 2006-2018 ($\mu g\ m^{-3}$). Only statistically significant trends are shown. If the trend is not statistically significant, the slope is set to 0. The standard deviation is based on the variation in daily $PM_{2.5}$ observations.

Month	Avg	Slope	SD
1	11.5	0.0	8.9
2	8.6	0.0	3.4
3	7.5	0.0	2.4
4	7.9	0.0	2.7
5	8.0	-0.3	2.6
6	8.4	0.0	2.7
7	7.9	0.0	4.2
8	7.1	0.0	2.9
9	7.2	0.0	2.5
10	8.4	0.0	2.4
11	10.9	-0.3	3.7
12	13.5	0.0	6.4

Table S2. Number of days in each category by year, May-September only*.

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	136	137	136	137	138	133	132	134	138	137	136	130	133	1757
Cross-val	15	15	15	16	15	15	15	15	15	16	14	15	14	195
Smoke						5	6	4				7	4	26

For Phoenix, $PM_{2.5}$ data is not available before 2011, thus smoke days can not be identified before then.

Table S3. Average MDA8 in each category (ppb)

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	62.7	59.2	59.9	55.8	60.2	60.4	60.0	59.2	57.7	55.6	56.5	59.5	58.8	58.9
Cross-val	64.1	63.7	59.1	54.1	57.7	59.4	64.1	61.9	59.9	56.9	57.3	55.2	57.1	59.3
Smoke						55.8	76.3	61.3				72.4	69.5	68.0

Table S4. Average surface Tmax (ºF) in each category

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	100.2	100.7	99.1	100.3	98.8	99.3	99.2	98.5	98.6	98.7	99.0	99.9	100.6	99.4
Cross-val	101.7	99.1	97.6	99.5	100.5	100.3	100.6	101.8	99.3	98.4	100.1	98.5	101.7	99.9
Smoke						100.4	103.0	110.0				110.9	101.5	105.5

Table S5. Number of days included in the GAM

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	132	131	132	137	137	132	129	133	136	136	136	128	132	1731
Cross-val	13	13	15	16	15	15	15	14	15	16	14	15	14	190
Smoke						5	6	4				7	4	26

Table S6. Average of residuals

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	0.12	-0.47	1.11	-1.33	0.43	0.11	0.32	-0.39	1.10	-1.42	-0.04	0.81	-0.27	0.00
Cross-val	0.69	4.47	1.03	-3.98	-1.73	-2.08	2.34	0.42	0.81	0.64	0.94	-1.57	-3.16	-0.15
Smoke						-3.78	3.67	-1.24				2.78	6.59	1.69

Salt Lake City CSA

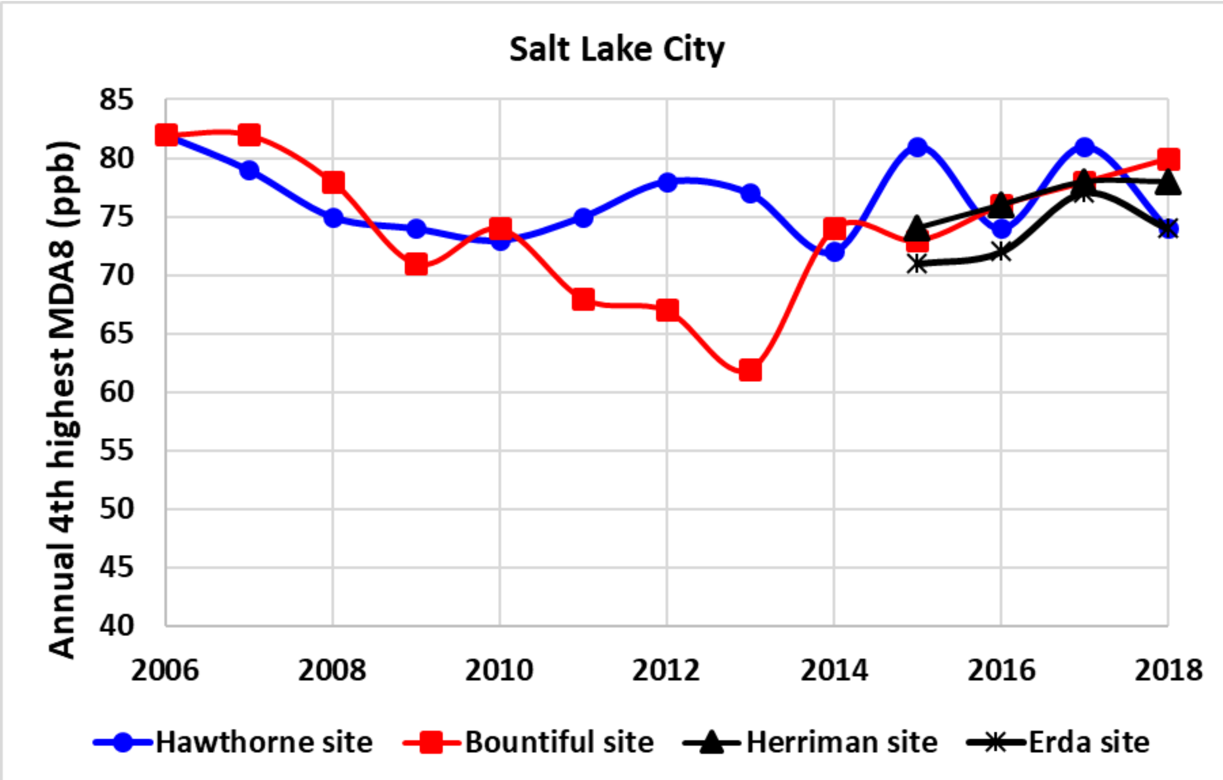

Figure S1. Trend in 4th highest MDA8 O_3 for each site.

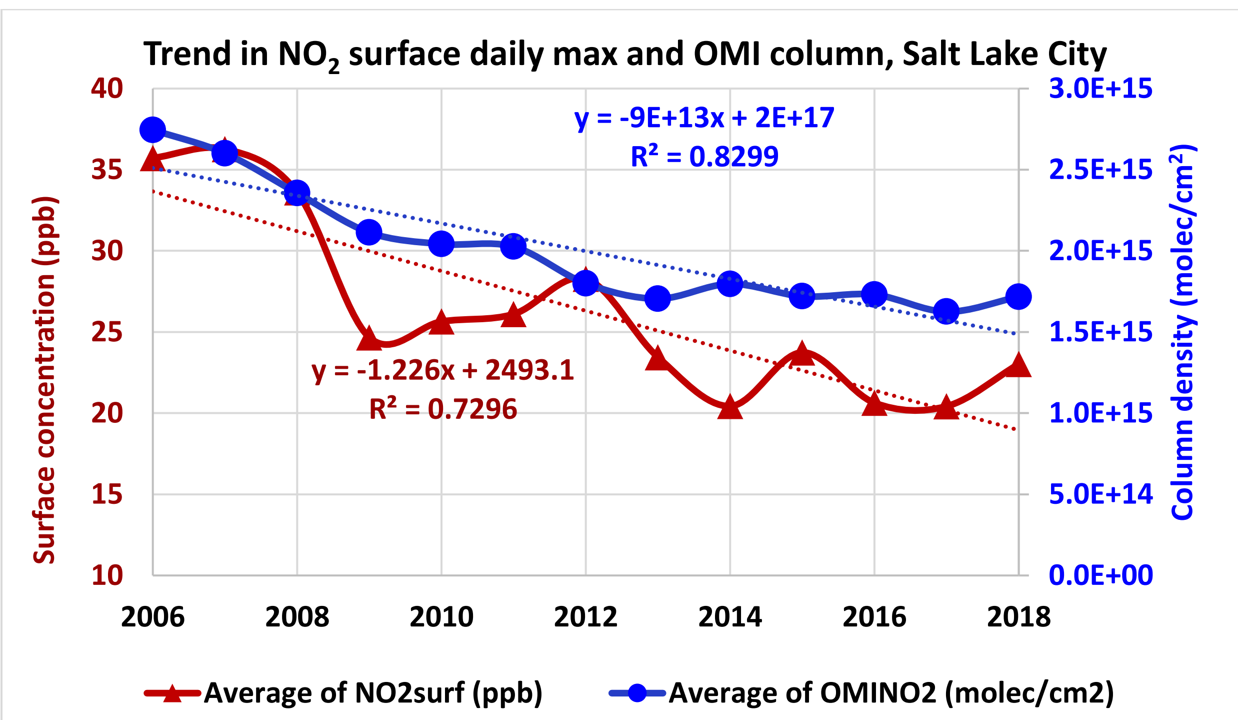

Figure S2. Trend in surface NO_2 daily max and OMI column, 2006-2018.

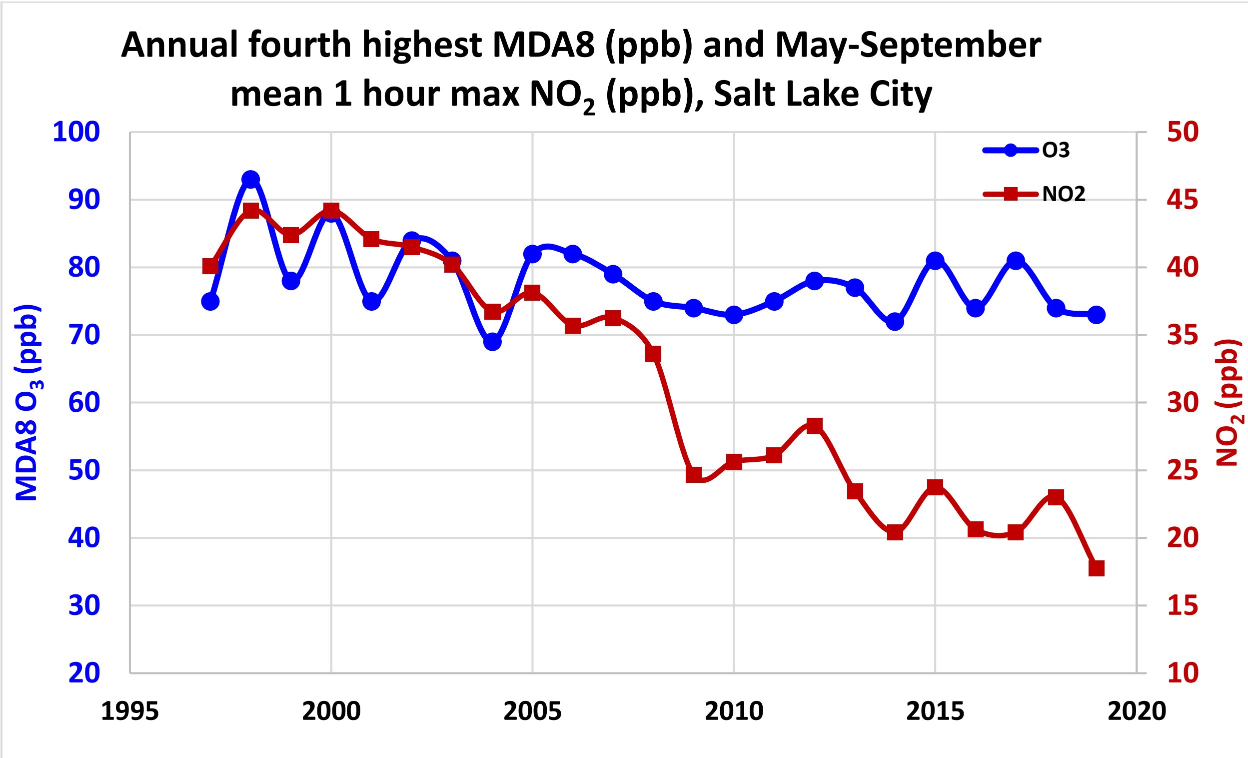

Figure S3. Trend in 4th highest O_3 and daily max NO_2, 1995-2019 for Hawthorne site.

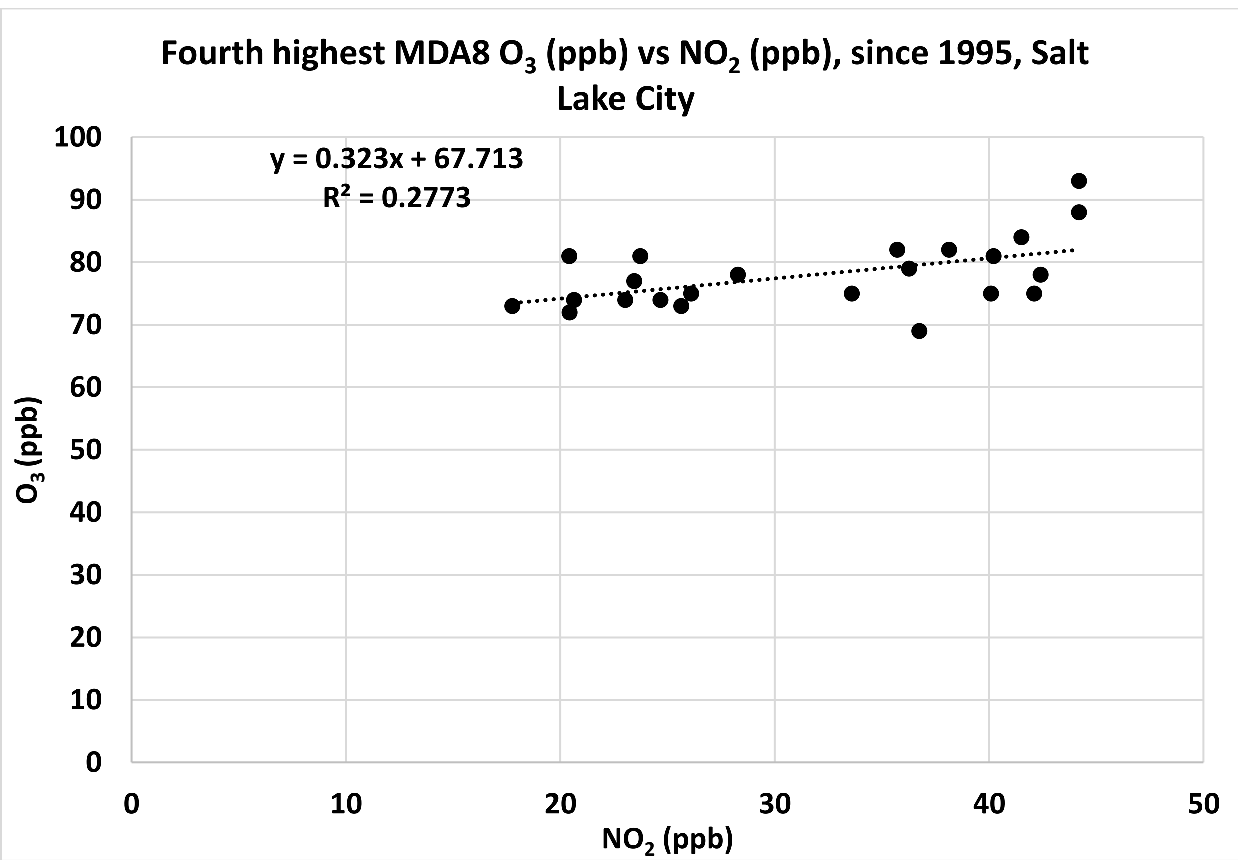

Figure S4. O_3 vs NO_2, 1995-2019.

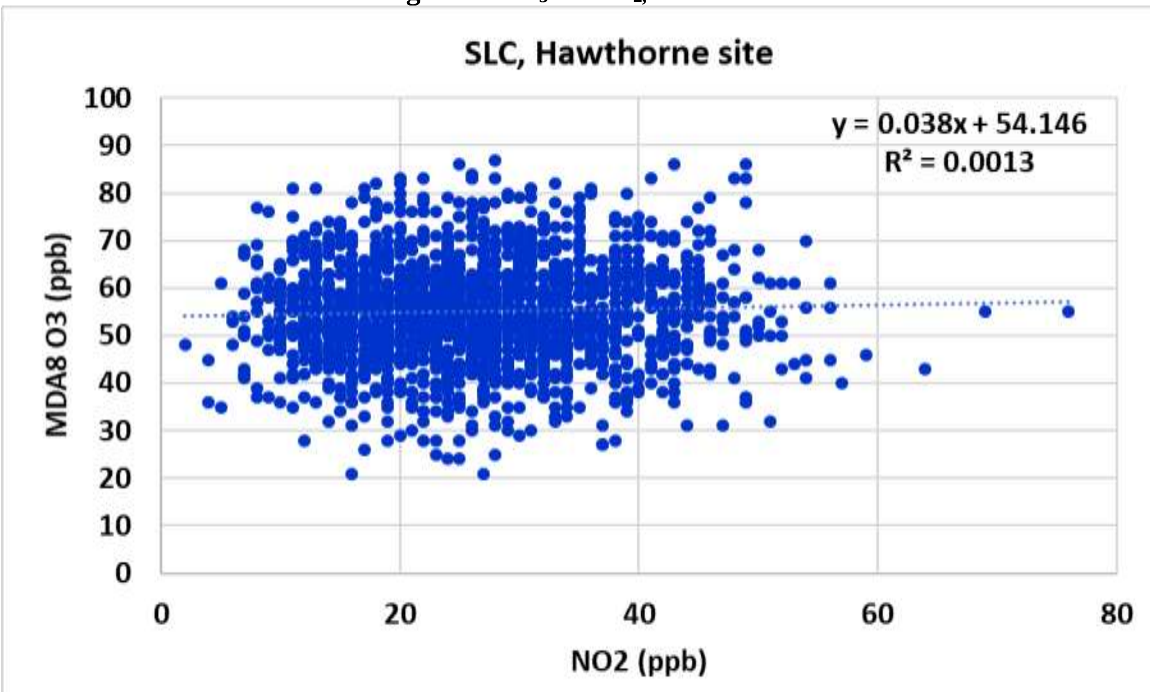

Figure S5. MDA8 vs daily max NO_2 for Hawthorne site.

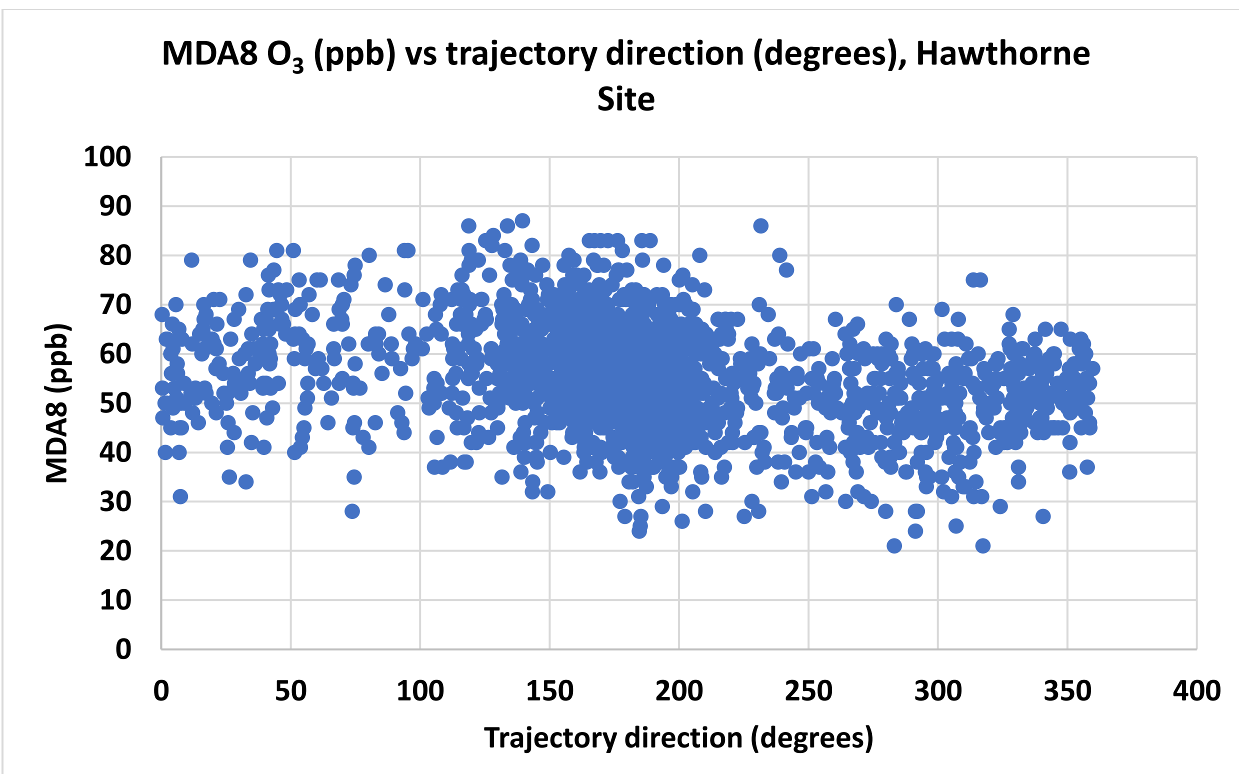

Figure S6. MDA8 vs trajectory direction for Hawthorne site.

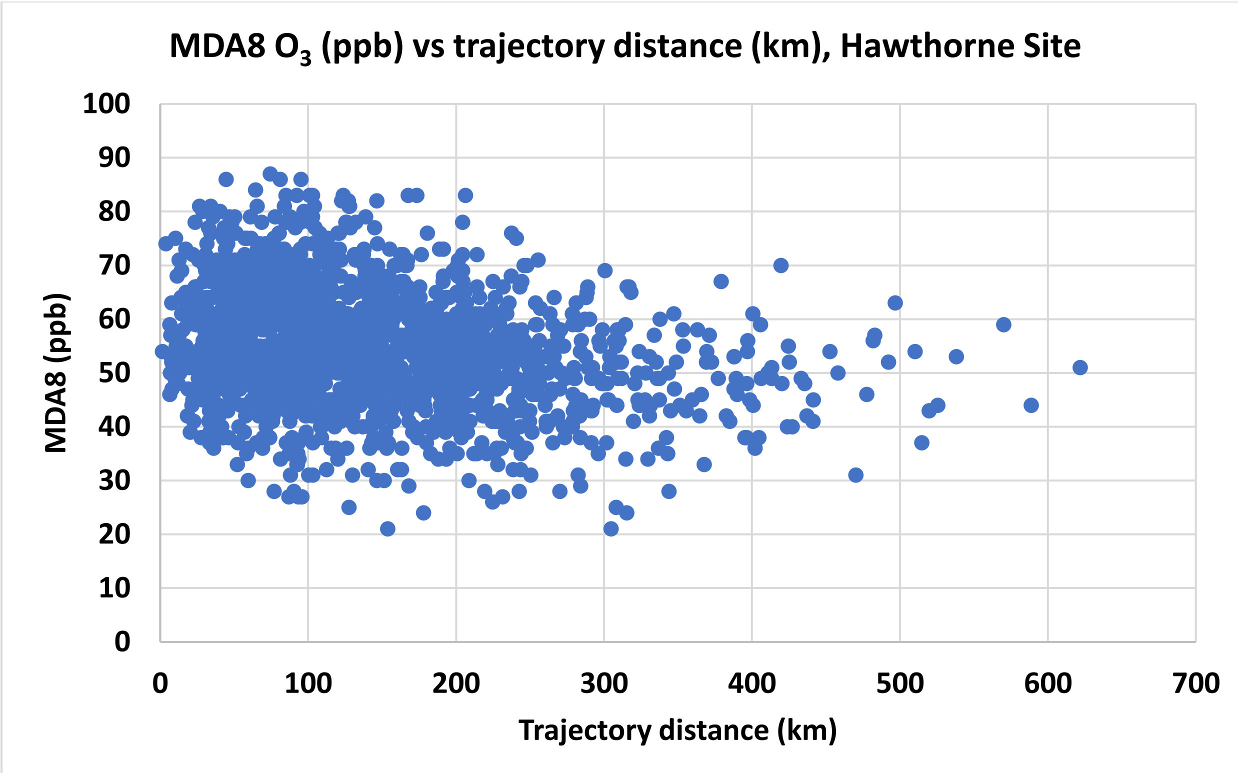

Figure S7. MDA8 vs trajectory distance for Hawthorne site.

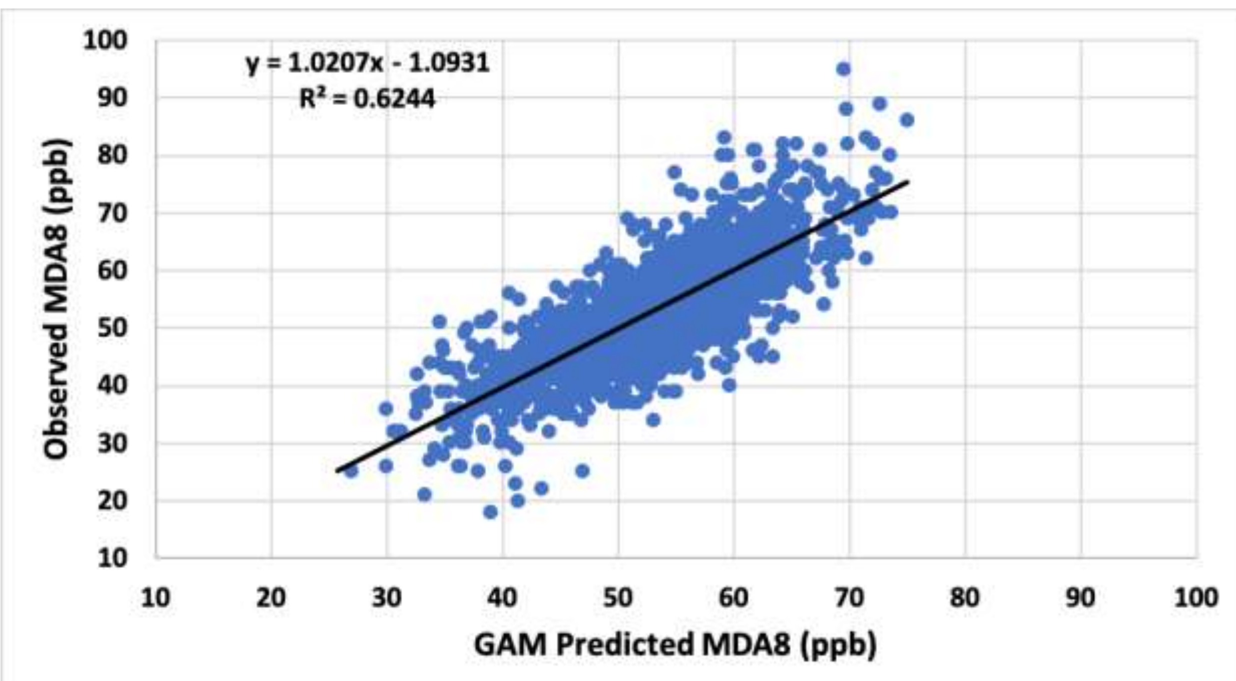

Figure S8. GAM predicted vs Observed, training data only for Bountiful site.

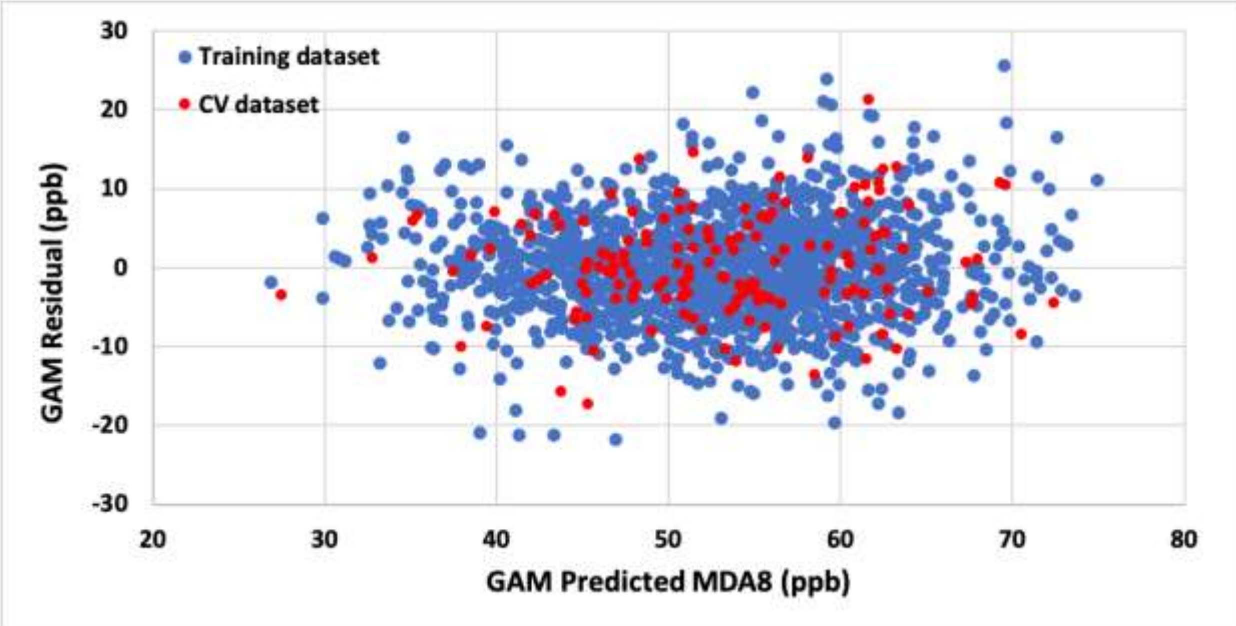

Figure S9. GAM residuals vs predicted, training and cross-validation data for Bountiful site.

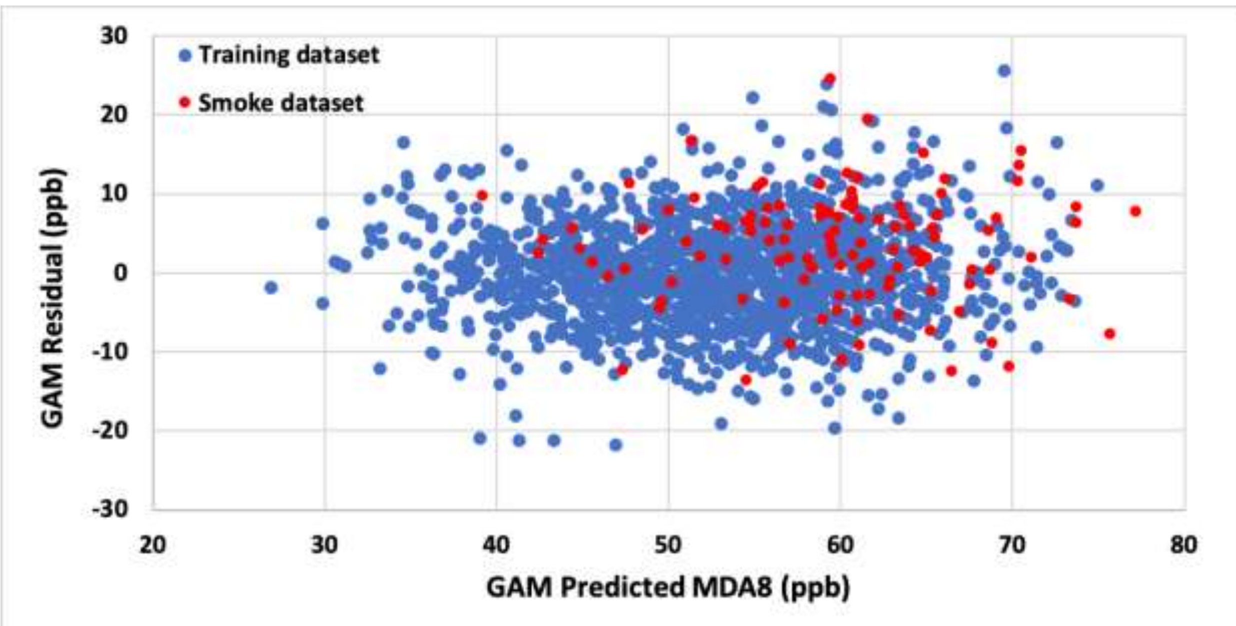

Figure S10. GAM residuals vs predicted, training and smoke data for Bountiful site.

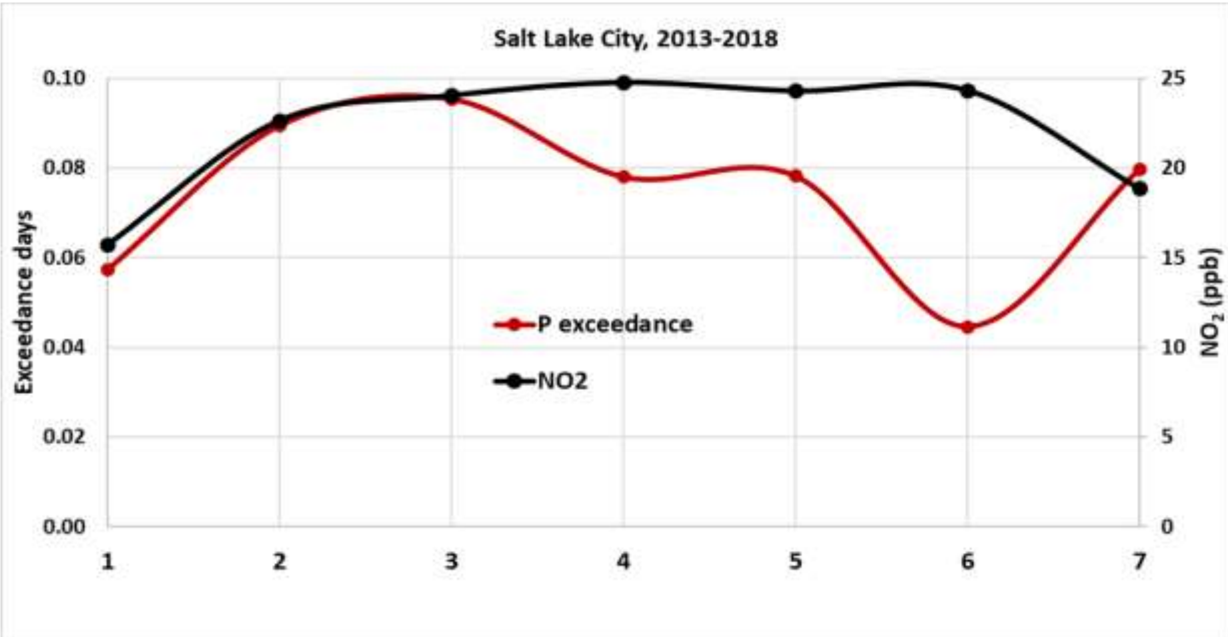

Figure S11. Probability (P) of an O_3 exceedance day (MDA8>70 ppb) across all sites that we considered and averaged surface NO_2 concentrations as a function of day of week (Sunday=1, Monday =2, etc). Both are calculated using May-September 2013-2018 data.

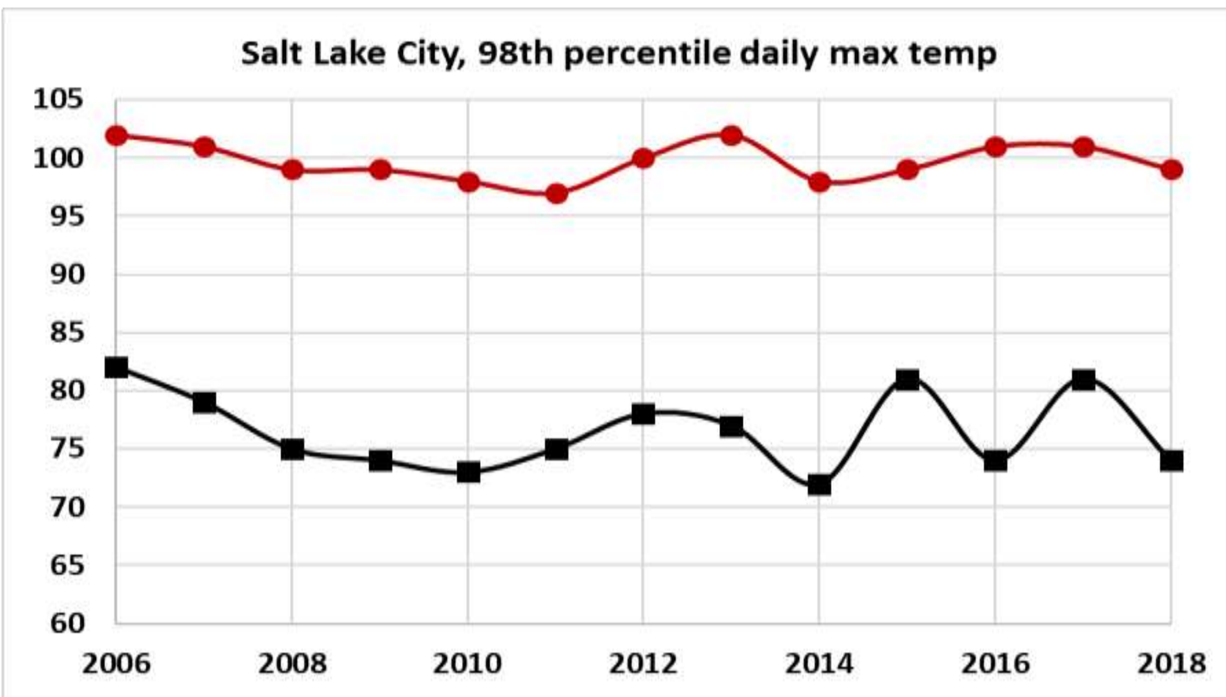

Figure S12. Annual 98th percentile of daily maximum temperature (ºF) and fourth highest MDA8 (ppb) for Hawthorne site.

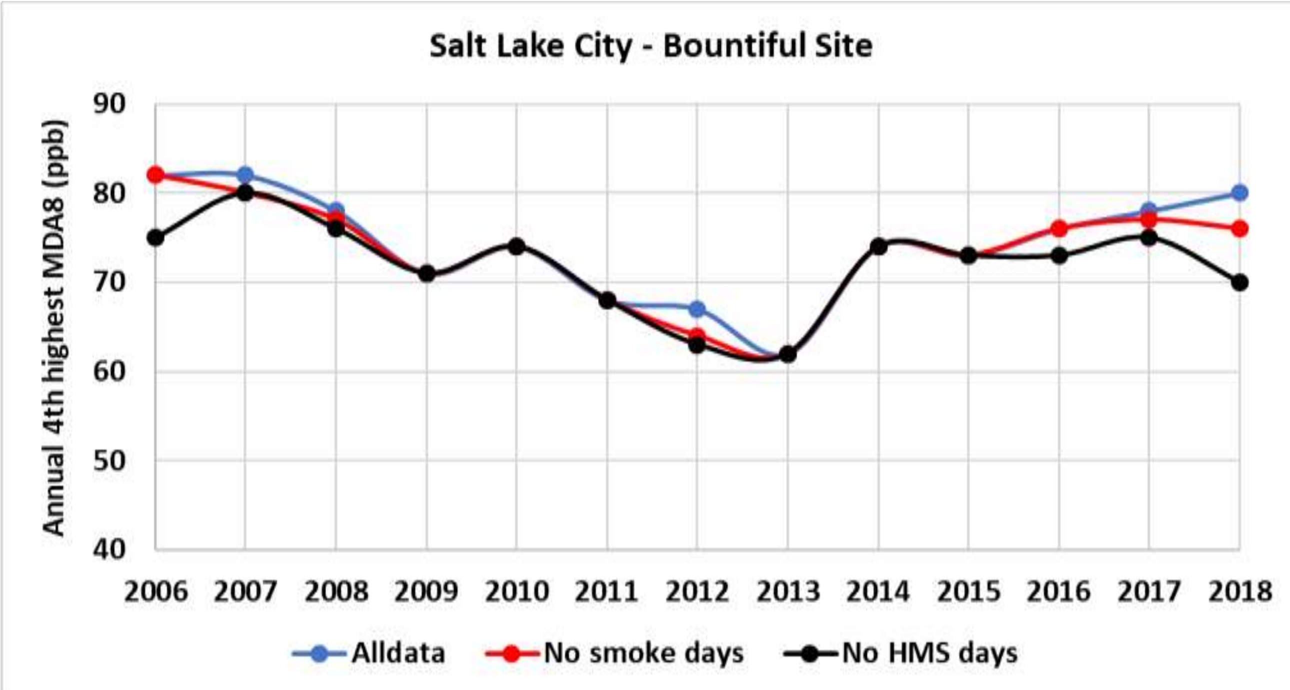

Figure S13. Annual fourth highest for all data, smoke days excluded, and HMS days excluded for Bountiful site.

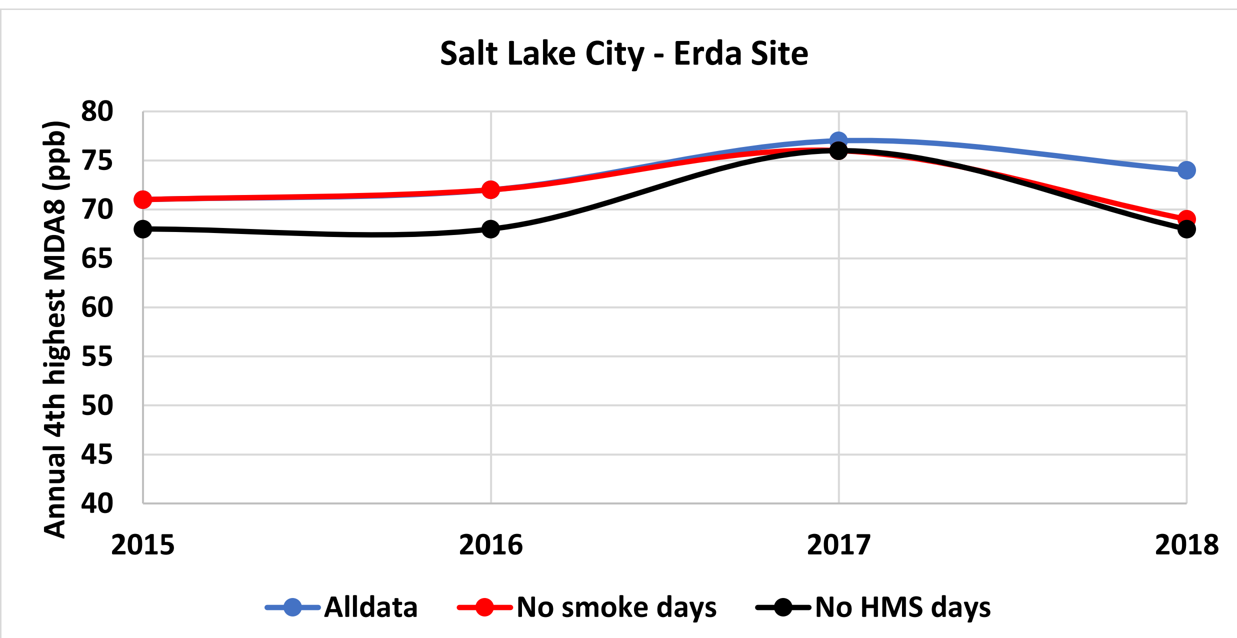

Figure S14. Annual fourth highest for all data, smoke days excluded, and HMS days excluded for Erda site.

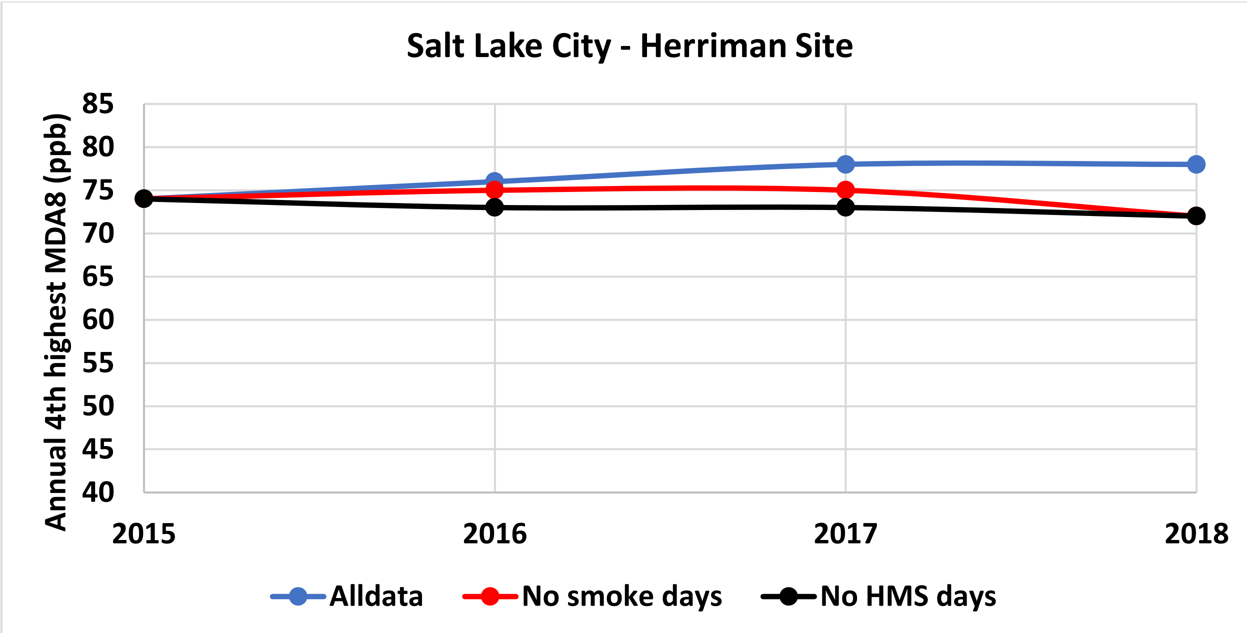

Figure S15. Annual fourth highest for all data, smoke days excluded, and HMS days excluded for Herriman site.

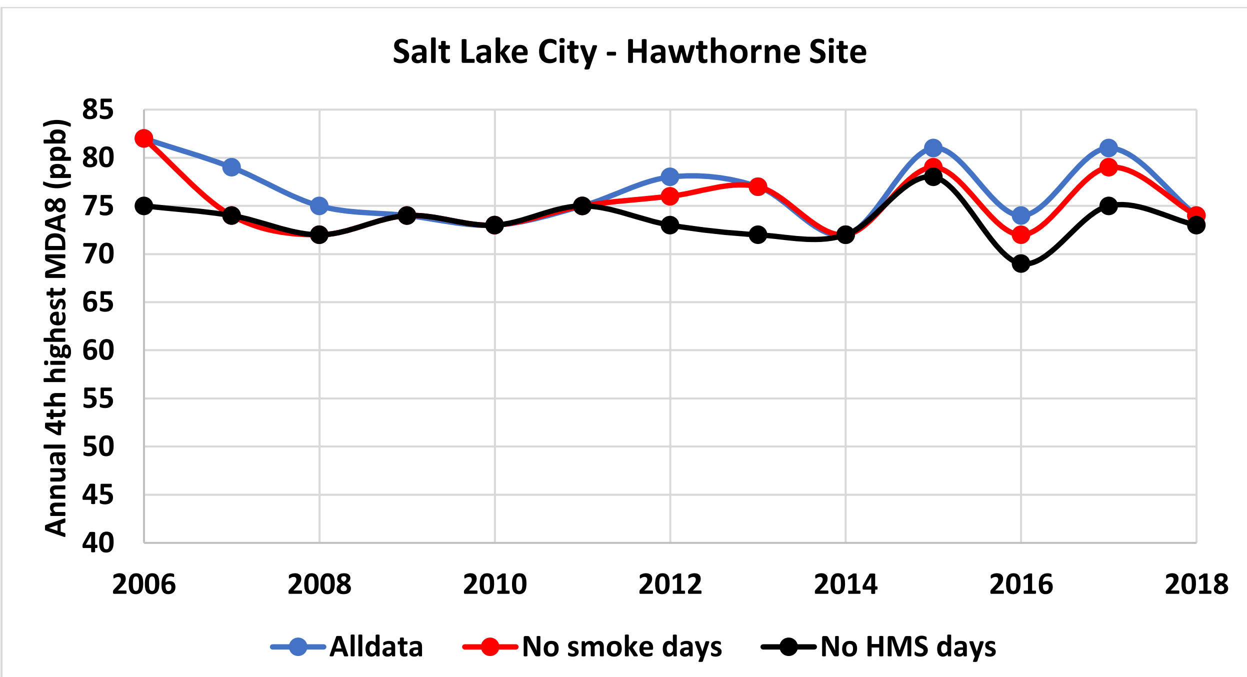

Figure S16. Annual fourth highest for all data, smoke days excluded, and HMS days excluded for Hawthorne site.

Table S1. Average, standard deviation and trend in $PM_{2.5}$ for SLC by month for 2006-2018 ($\mu g\ m^{-3}$). Only statistically significant trends are shown. If the trend is not statistically significant, the slope is set to 0. The standard deviation is based on the variation in daily $PM_{2.5}$ observations.

Month	Avg	Slope	SD
1	22.0	0.0	17.7
2	10.7	0.0	10.6
3	5.8	-0.3	4.5
4	5.3	-0.3	3.8
5	5.3	-0.3	2.3
6	6.0	0.0	2.2
7	7.8	-0.3	3.4
8	7.2	0.0	3.0
9	7.3	0.0	3.9
10	6.5	-0.4	3.8
11	9.2	-0.5	6.3
12	15.9	0.0	12.6

Salt Lake City CSA-Bountiful

Table S2. Number of days in each category by year, May-September only.

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	132	122	119	128	136	133	119	87	134	73	132	127	108	1550
Cross-val	12	14	13	16	15	14	13	9	14	8	15	13	13	169
Smoke	8	16	16	7		1	17	1	3	1	5	12	31	118

Table S3. Average MDA8 in each category (ppb)

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	57.0	55.1	52.1	51.1	50.4	49.2	49.9	48.8	53.3	53.0	52.8	58.2	56.1	52.9
Cross-val	58.2	56.9	52.2	51.9	49.3	49.9	51.9	48.0	55.1	52.3	53.2	59.0	59.1	53.7
Smoke	64.4	63.6	63.9	56.0		61.0	57.6	55.0	57.0	59.0	65.0	66.5	64.4	62.6

Table S4. Average surface Tmax (ºF) in each category

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	84.9	86.8	83.3	83.5	82.1	81.2	85.8	87.3	83.3	83.8	86.0	85.7	85.0	84.5
Cross-val	81.7	86.4	79.3	82.8	83.9	83.9	87.6	87.1	85.1	85.3	88.2	84.4	86.2	84.8
Smoke	94.6	91.1	94.3	87.9		86.0	93.5	101.5	93.7	91.0	97.6	95.6	93.1	93.2

Table S5. Number of days included in the GAM

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	125	119	117	128	136	131	117	87	134	72	130	127	108	1531
Cross-val	11	12	13	16	15	13	13	9	14	8	14	13	13	164
Smoke	7	15	16	7		1	17	1	3	1	5	12	31	116

Table S6. Average of residuals

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	0.70	-0.82	-0.11	0.12	0.10	0.55	-0.66	-2.22	1.40	1.68	-2.25	2.25	-1.03	0.00
Cross-val	2.99	0.53	0.90	1.11	-0.98	-0.93	-0.48	-1.71	1.91	0.22	-2.50	2.92	0.03	0.32
Smoke	1.74	4.77	3.68	5.43		6.32	4.43	-5.97	-4.48	5.67	0.68	4.64	2.07	3.21

Salt Lake City CSA-Erda*

Table S7. Number of days in each category by year, May-September only.

	2015	2016	2017	2018	Grand Total
Training	62	133	125	107	427
Cross-val	9	15	14	12	50
Smoke	3	5	12	32	52

Table S8. Average MDA8 in each category (ppb)

	2015	2016	2017	2018	Grand Total
Training	54.7	51.9	56.4	54.8	54.4
Cross-val	50.7	50.9	54.3	57.8	53.5
Smoke	61.0	60.8	61.6	62.0	61.7

Table S9. Average surface Tmax (ºF) in each category

	2015	2016	2017	2018	Grand Total
Training	83.8	86.0	85.7	85.0	84.5
Cross-val	85.3	88.2	84.4	86.2	84.8
Smoke	91.0	97.6	95.6	93.1	93.2

Table S10. Number of days included in the GAM

	2015	2016	2017	2018	Grand Total
Training	62	73	119	101	355
Cross-val	9	7	14	12	42
Smoke	3	2	11	32	48

Table S11. Average of residuals

	2015	2016	2017	2018	Grand Total
Training	-1.18	-1.90	1.13	0.77	0.00
Cross-val	-5.34	-3.26	-0.53	3.10	-0.98
Smoke	0.55	-3.08	0.72	4.99	3.40

*Data for Erda site start in 2015.

Salt Lake City CSA-Herriman*

Table S12. Number of days in each category by year, May-September only.

	2015	2016	2017	2018	Grand Total
Training	128	130	125	101	484
Cross-val	15	15	14	12	56
Smoke	10	5	12	32	59

Table S13. Average MDA8 in each category (ppb)

	2015	2016	2017	2018	Grand Total
Training	55.1	55.7	59.4	58.2	57.0
Cross-val	55.3	56.5	56.8	58.7	56.7
Smoke	64.2	67.0	66.8	66.4	66.1

Table S14. Average surface Tmax (ºF) in each category

	2015	2016	2017	2018	Grand Total
Training	83.8	86.0	85.7	85.0	84.5
Cross-val	85.3	88.2	84.4	86.2	84.8
Smoke	91.0	97.6	95.6	93.1	93.2

Table S15. Number of days included in the GAM

	2015	2016	2017	2018	Grand Total
Training	127	128	125	101	481
Cross-val	15	14	14	12	55
Smoke	10	5	12	32	59

Table S16. Average of residuals

	2015	2016	2017	2018	Grand Total
Training	-1.00	-0.81	1.48	0.45	0.00
Cross-val	-2.85	-1.17	2.13	0.68	-0.64
Smoke	4.83	5.25	5.65	6.47	5.65

*Data for Herriman site start in 2015.

Salt Lake City CSA-Hawthorne

Table S17. Number of days in each category by year, May-September only.

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	133	119	123	127	87	126	120	136	136	127	112	123	102	1571
Cross-val	12	13	13	15	10	15	14	15	13	14	12	14	12	172
Smoke	8	16	16	7		1	18	2	3	10	4	12	26	123

Table S18. Average MDA8 in each category (ppb)

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	56.5	55.0	50.2	54.1	51.2	54.8	57.5	53.7	51.7	55.2	52.0	57.5	57.2	54.4
Cross-val	57.0	58.1	51.4	53.7	51.3	55.5	57.9	54.1	53.3	54.7	51.2	56.1	60.3	55.0
Smoke	63.4	63.9	59.2	53.6		73.0	65.4	67.0	56.7	68.1	63.5	68.3	57.7	62.3

Table S19. Average surface Tmax (ºF) in each category

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	84.9	86.8	83.3	83.5	82.1	81.2	85.8	87.3	83.3	83.8	86.0	85.7	85.0	84.5
Cross-val	81.7	86.4	79.3	82.8	83.9	83.9	87.6	87.1	85.1	85.3	88.2	84.4	86.2	84.8
Smoke	94.6	91.1	94.3	87.9		86.0	93.5	101.5	93.7	91.0	97.6	95.6	93.1	93.2

Table S20. Number of days included in the GAM

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	126	115	121	127	87	125	118	136	136	126	110	123	102	1552
Cross-val	11	11	13	15	10	14	14	15	13	14	11	14	12	167
Smoke	7	15	16	7		1	18	2	3	10	4	12	26	121

Table S21. Average of residuals

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	0.40	0.17	-1.53	1.39	-1.95	0.58	1.44	-0.64	-1.32	1.89	-1.58	0.57	-0.01	0.00
Cross-val	0.12	3.45	0.35	1.84	-1.16	1.14	0.23	3.09	-2.98	0.78	-4.77	1.78	1.89	0.55
Smoke	1.97	6.93	-1.73	4.56		13.95	7.62	7.50	-4.62	11.18	1.95	7.32	-3.92	3.14

San Francisco CSA

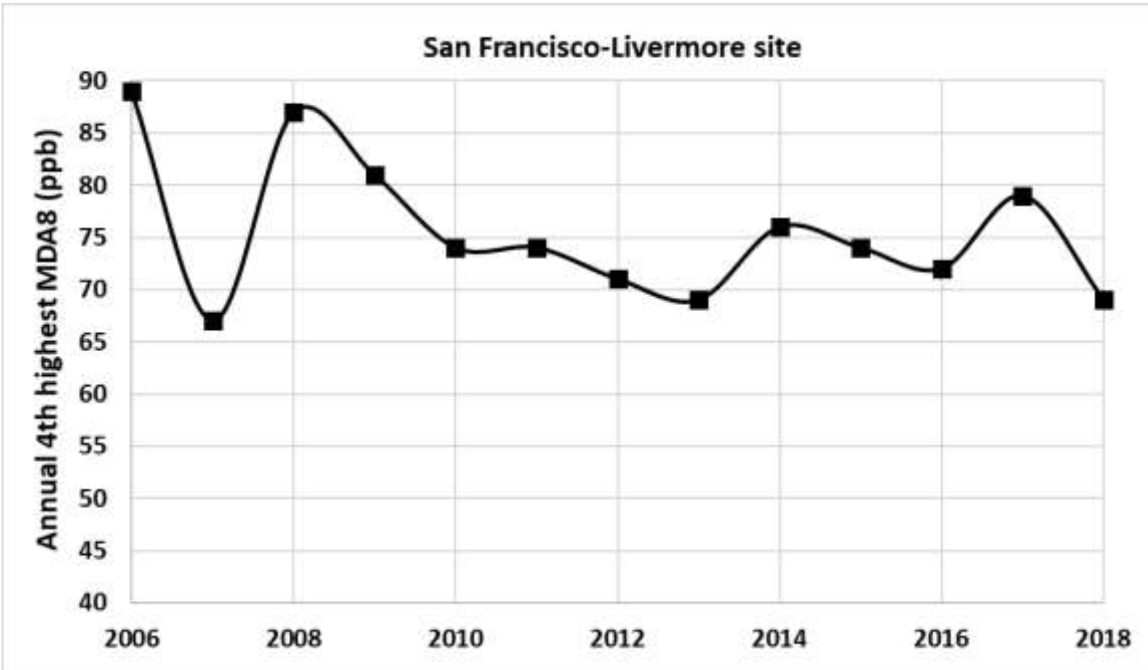

Figure S1. Trend in 4th highest MDA8 O_3 for each site.

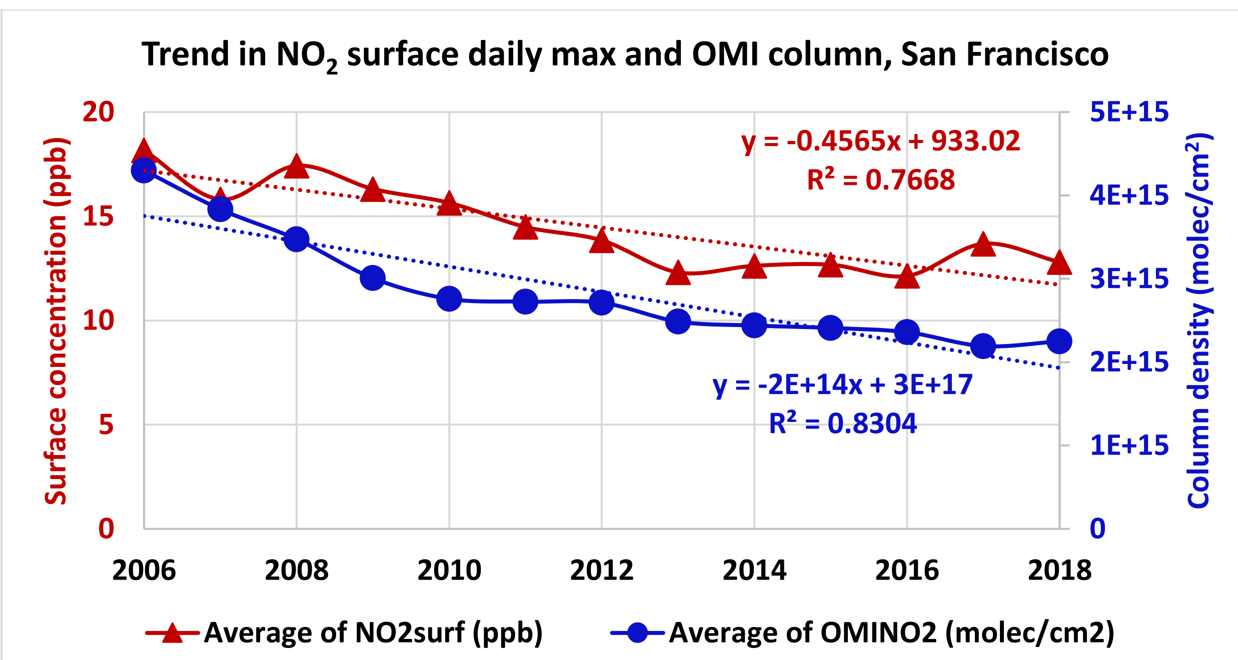

Figure S2. Trend in surface NO_2 daily max and OMI column, 2006-2018.

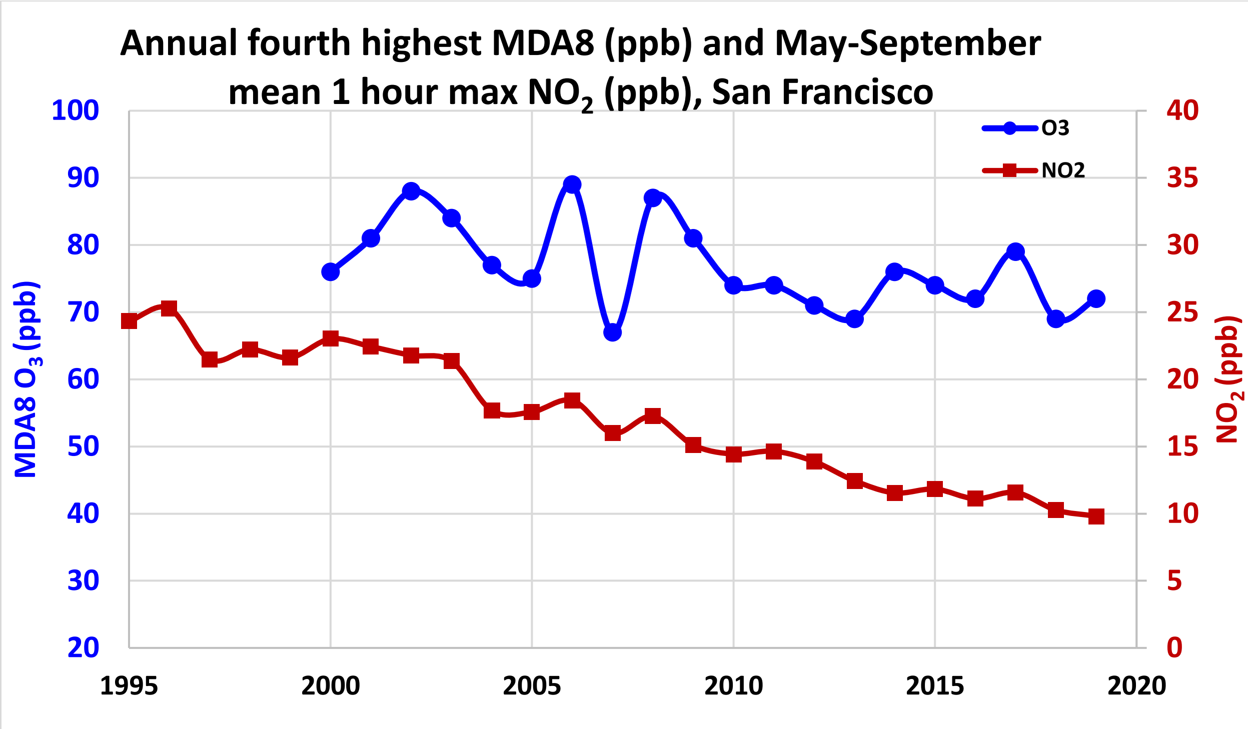

Figure S3. Trend in 4th highest O_3 and daily max NO_2, 1995-2019. For NO_2 I use data from these sites: 060010007, 060130002, 060131002, 060811001, 060750005 and 060410001.

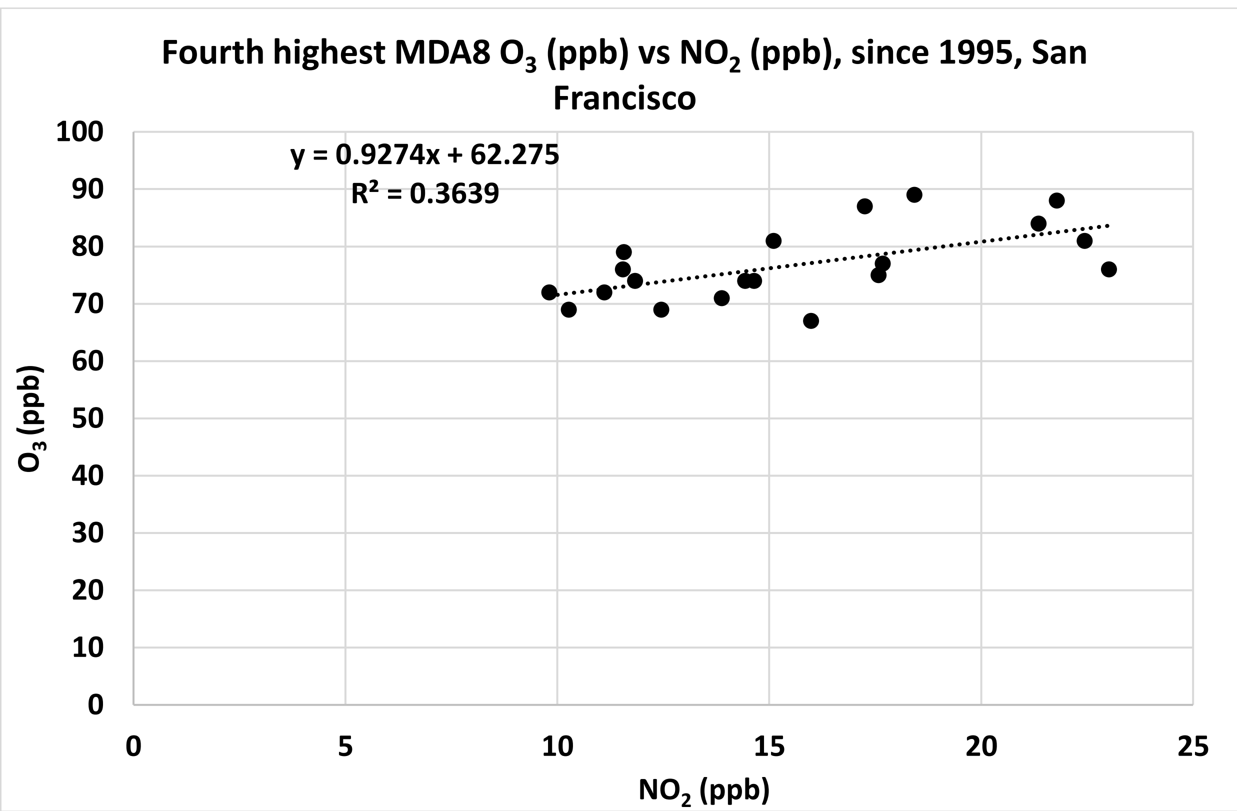

Figure S4. O_3 vs NO_2, 1995-2019.

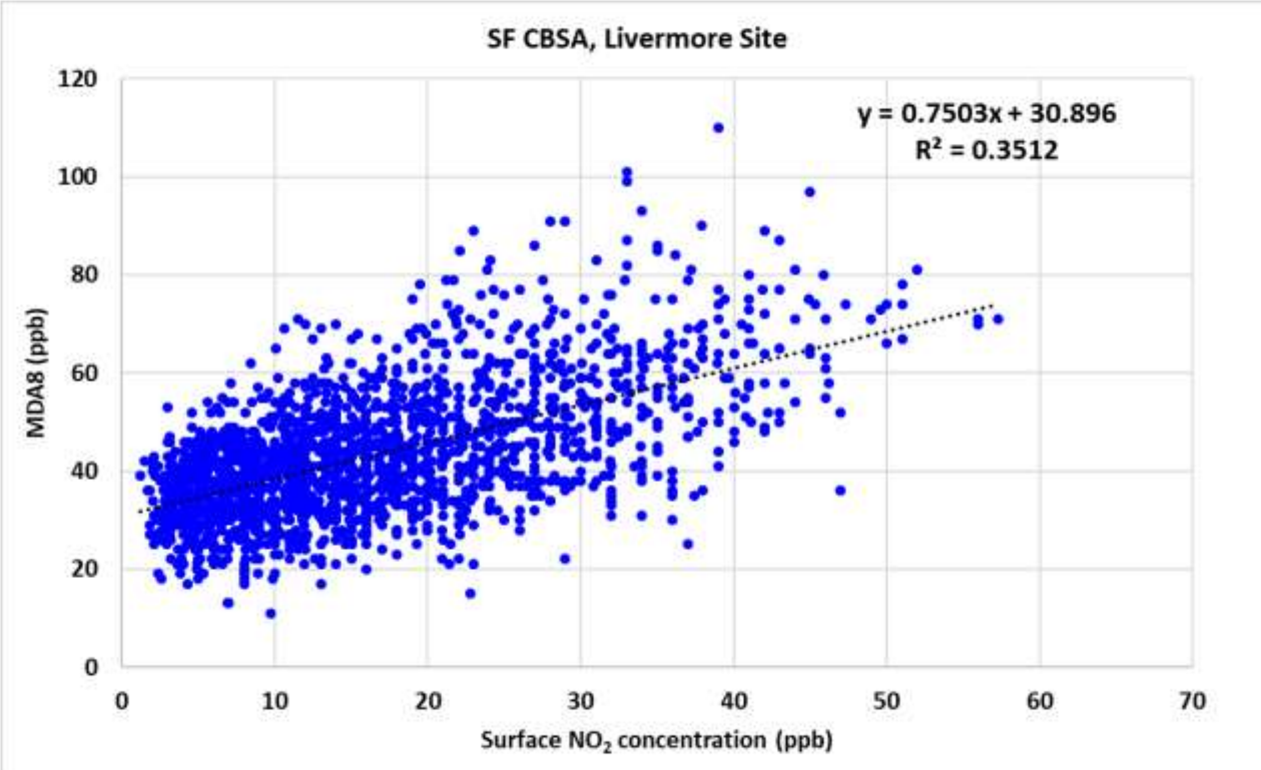

Figure S5. MDA8 vs daily max NO_2 for Livermore site.

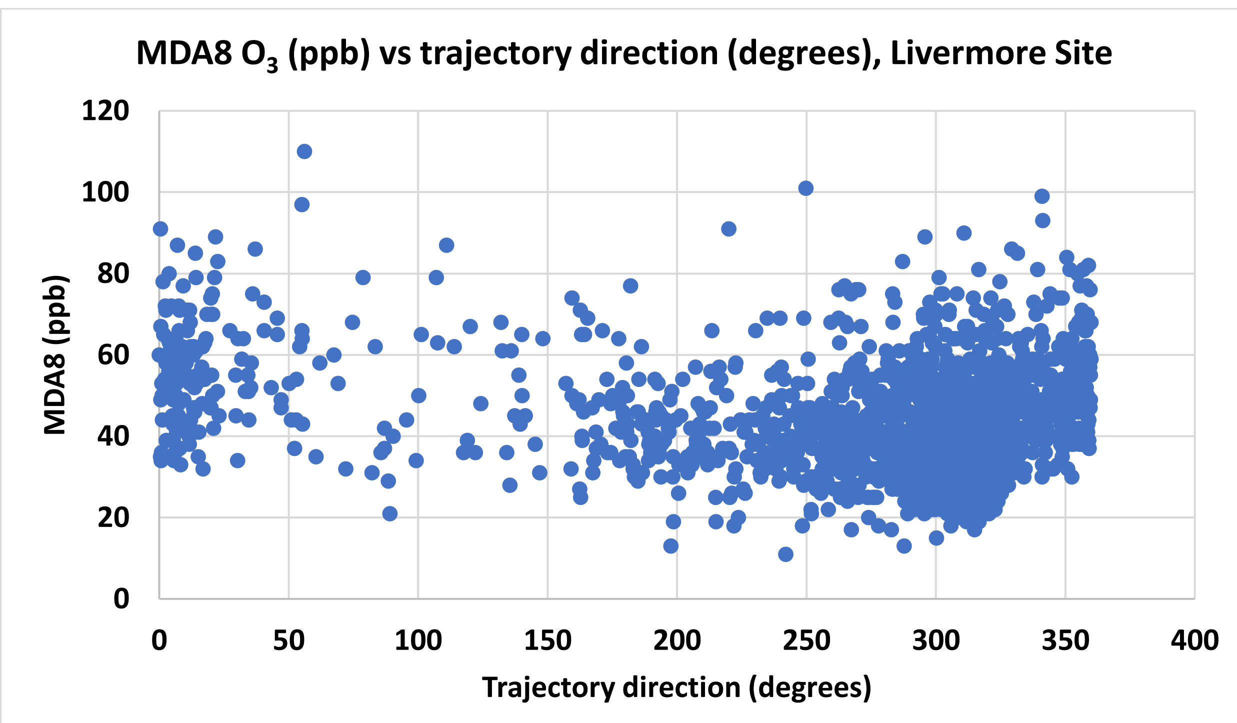

Figure S6. MDA8 vs trajectory direction for Livermore site.

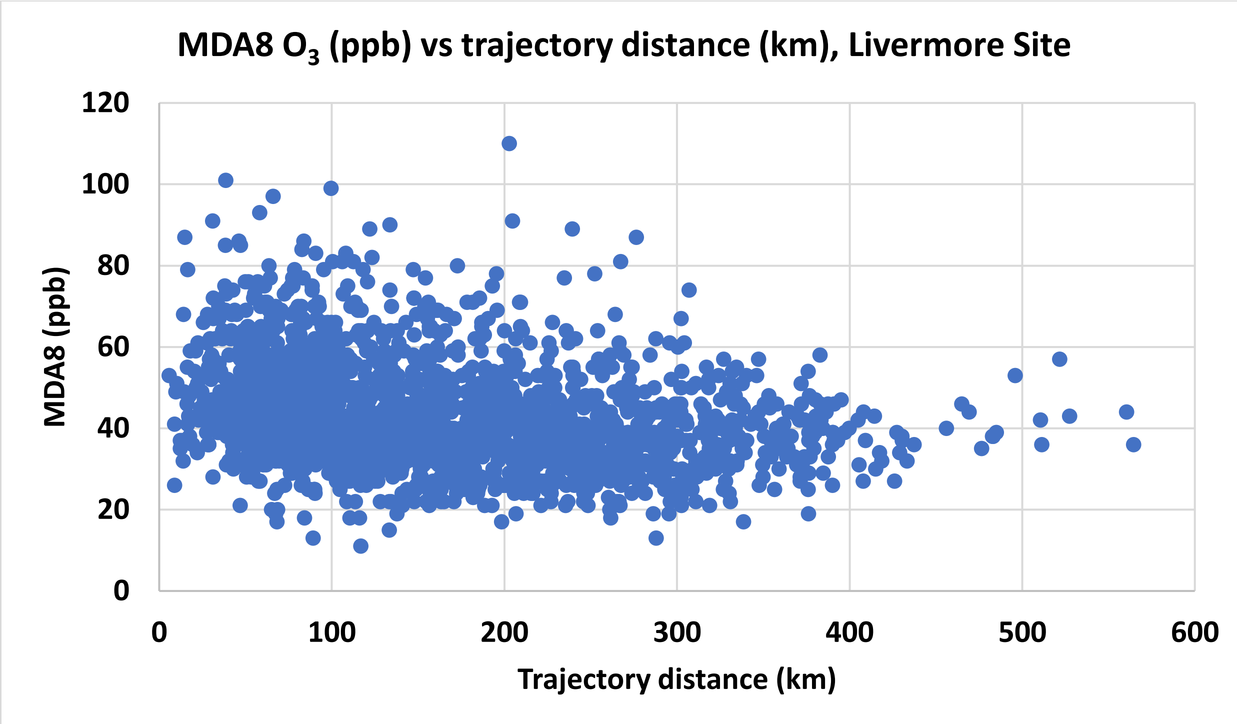

Figure S7. MDA8 vs trajectory distance for Livermore site.

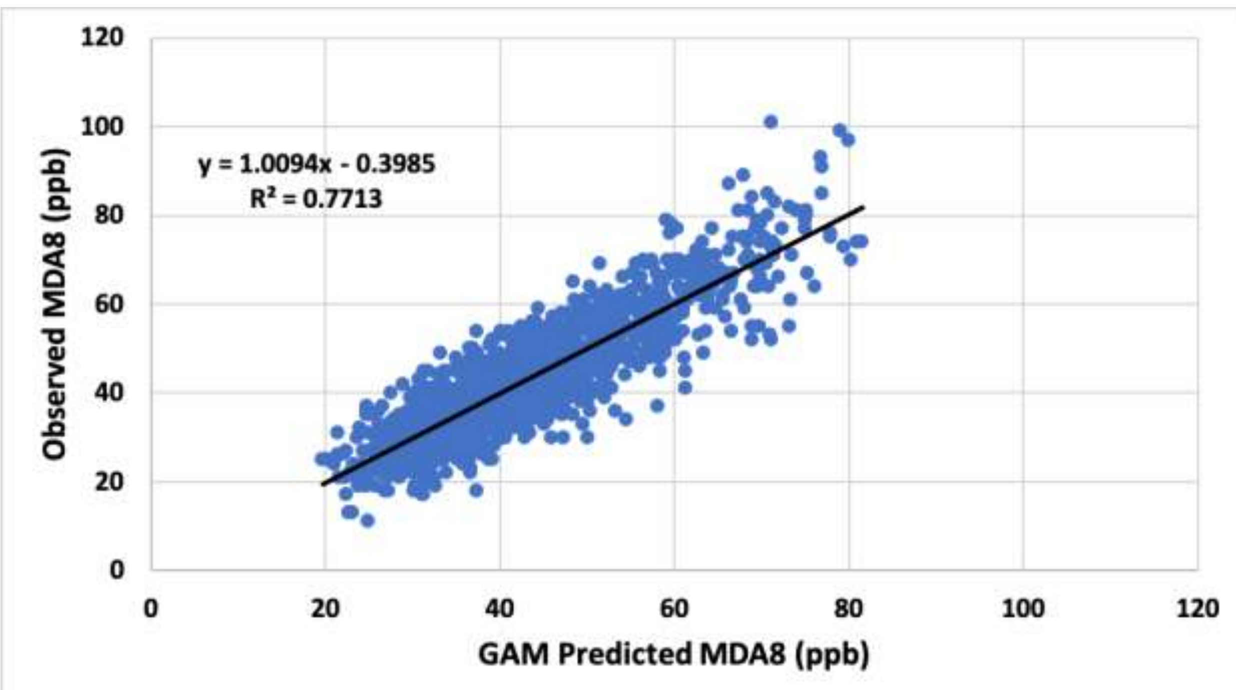

Figure S8. GAM predicted vs Observed, training data only for Livermore site.

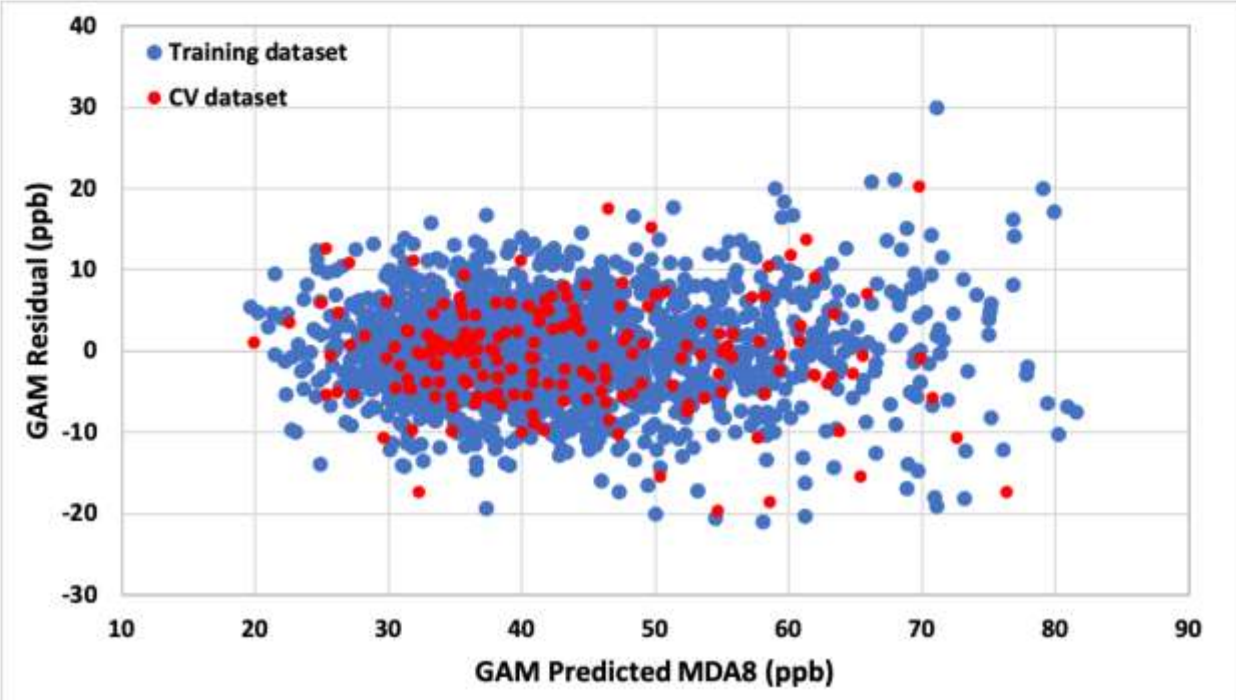

Figure S9. GAM residuals vs predicted, training and cross-validation data for Livermore site.

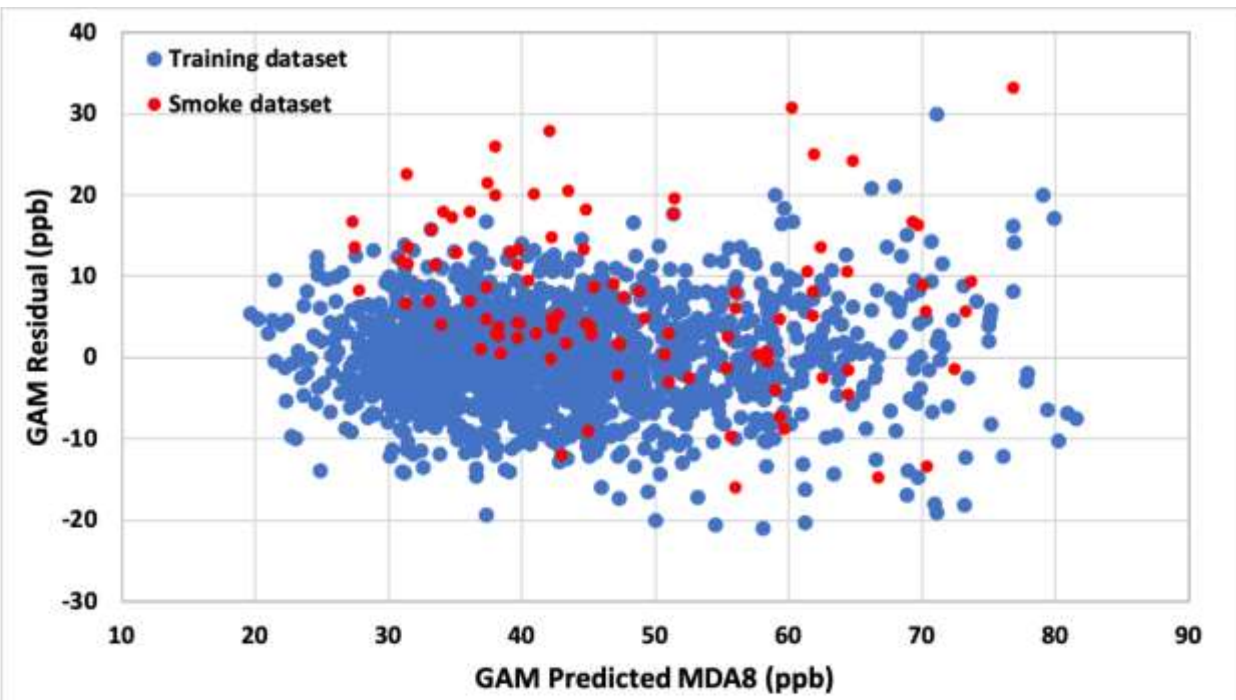

Figure S10. GAM residuals vs predicted, training and smoke data for Livermore site.

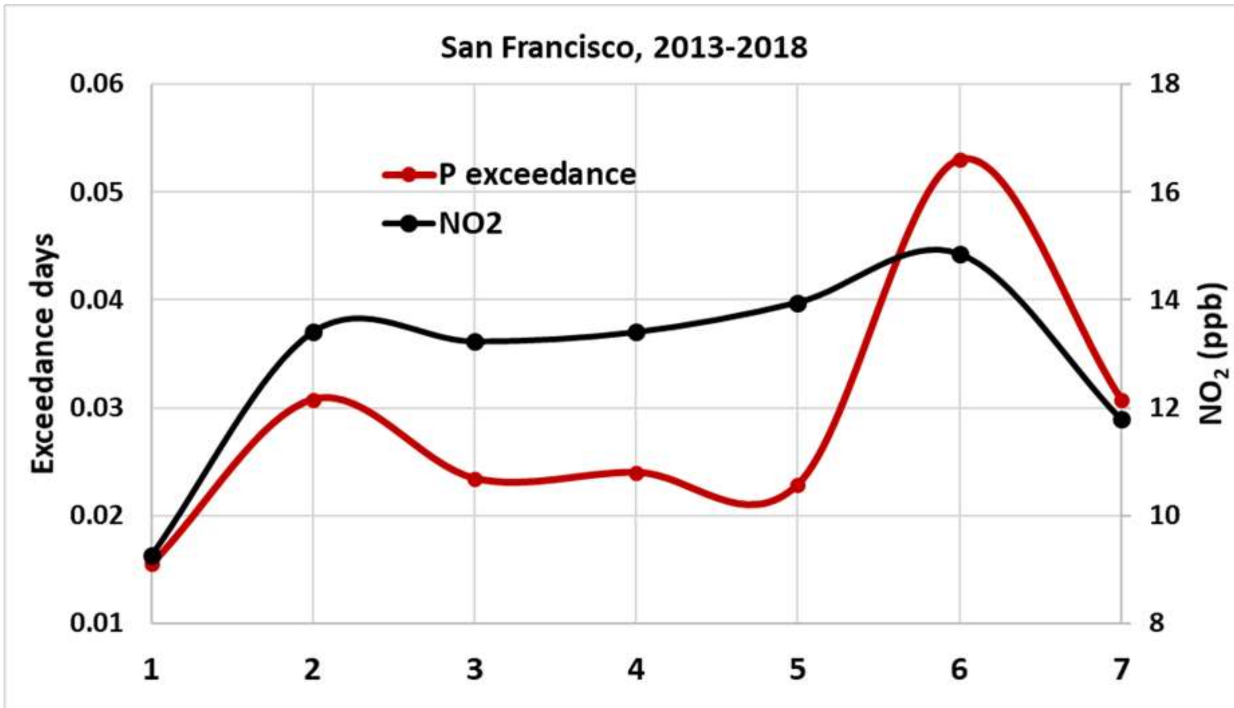

Figure S11. Probability (P) of an O_3 exceedance day (MDA8>70 ppb) across all sites that we considered and averaged surface NO_2 concentrations as a function of day of week (Sunday=1, Monday =2, etc). Both are calculated using May-September 2013-2018 data.

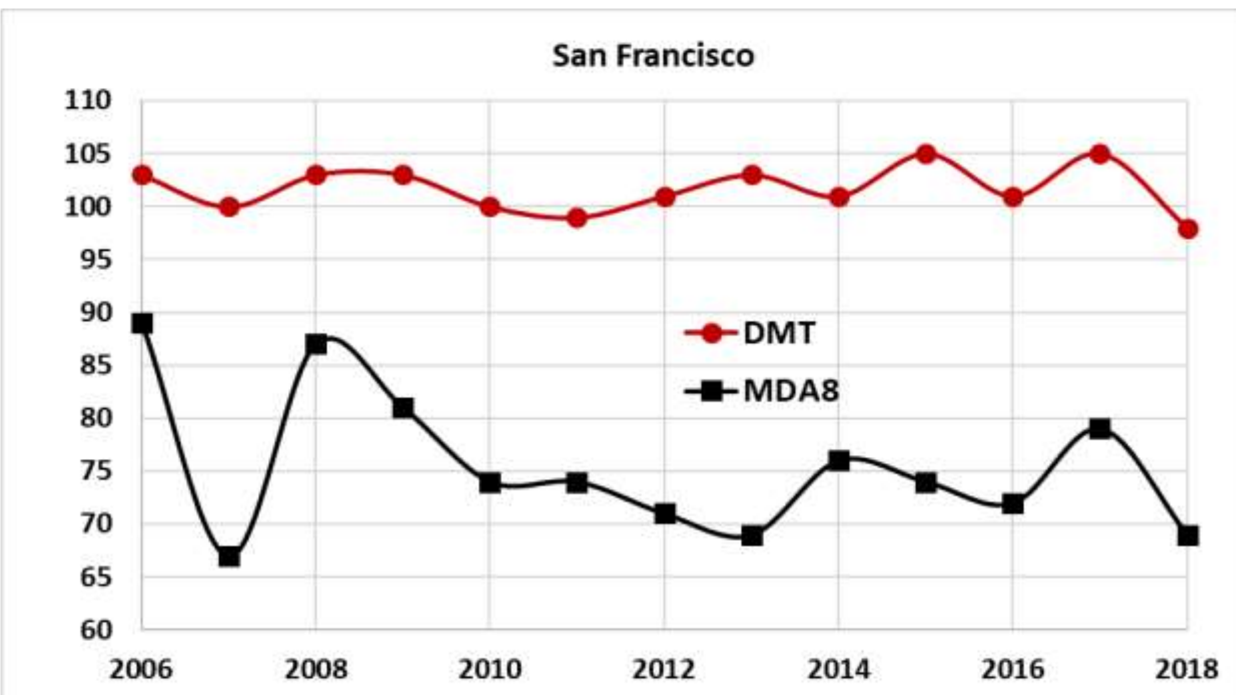

Figure S12. Annual 98th percentile of daily maximum temperature (ºF) and fourth highest MDA8 (ppb).

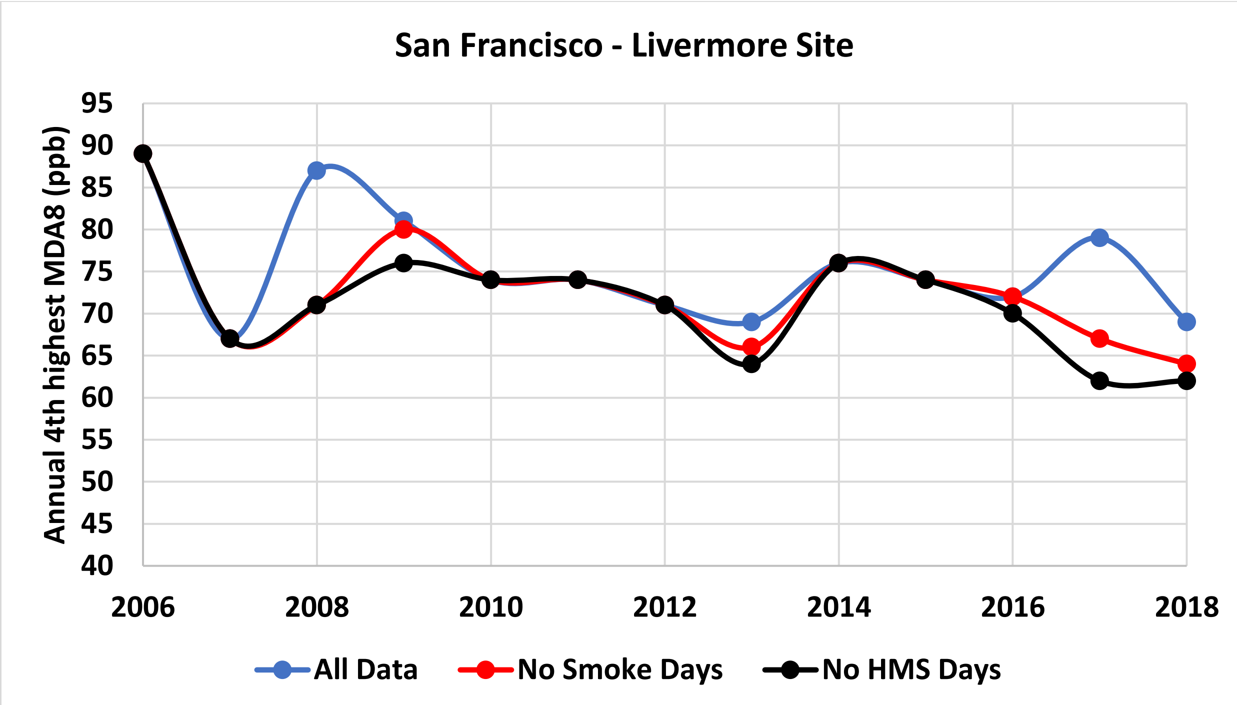

Figure S13. Annual fourth highest for all data, smoke days excluded, and HMS days excluded for Livermore site.

Table S1. Average, standard deviation and trend in $PM_{2.5}$ for SF by month for 2006-2018 (µg m^{-3}). Only statistically significant trends are shown. If the trend is not statistically significant, the slope is set to 0. The standard deviation is based on the variation in daily $PM_{2.5}$ observations.

Month	Avg	Slope	SD
1	13.2	0.0	7.7
2	8.5	0.0	5.3
3	6.4	0.0	2.5
4	7.3	0.0	2.9
5	8.2	0.0	3.6
6	8.4	0.0	3.8
7	7.5	0.0	3.6
8	8.0	0.0	3.8
9	8.5	0.0	3.0
10	8.5	0.0	3.6
11	10.1	0.0	5.4
12	12.7	0.0	7.9

Table S2. Number of days in each category by year, May-September only.

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	131	133	113	131	137	136	135	129	133	135	132	116	118	1679
Cross-val	14	15	11	15	15	15	16	15	15	16	13	14	12	186
Smoke	7	4	29	7				6	3	1	5	21	21	104

Table S3. Average MDA8 in each category (ppb)

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	47.6	39.9	43.5	42.0	40.9	42.5	44.0	38.5	43.0	43.3	44.2	39.6	43.4	42.5
Cross-val	46.9	36.5	40.6	39.5	41.7	48.7	48.1	39.5	43.4	41.8	42.9	38.7	47.2	42.7
Smoke	51.0	49.5	57.4	63.3				55.3	55.7	63.0	51.4	58.7	49.9	55.4

Table S4. Average surface Tmax (ºF) in each category

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	84.7	83.1	84.6	84.6	81.5	81.9	84.8	84.5	86.1	86.1	85.1	84.4	83.5	84.2
Cross-val	86.8	79.9	82.5	82.7	83.3	83.9	87.2	86.0	86.8	84.8	86.5	84.0	87.1	84.7
Smoke	92.3	87.5	89.6	96.3				90.2	95.0	101.0	93.6	94.8	85.9	90.9

Table S5. Number of days included in the GAM

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	128	130	110	130	134	132	135	128	133	134	131	115	115	1655
Cross-val	14	15	11	15	15	15	16	14	15	16	13	14	12	185
Smoke	7	4	28	7				6	3	1	5	20	20	101

Table S6. Average of residuals

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Grand Total
Training	0.51	-0.96	0.31	0.66	-0.81	0.26	0.65	-0.99	0.45	-0.26	0.96	-1.45	0.60	0.00
Cross-val	-1.19	-0.16	0.17	-0.05	-1.94	1.68	1.58	-1.81	0.60	-1.01	-1.47	-2.85	-0.15	-0.48
Smoke	-6.96	5.68	13.09	6.95				6.68	-0.77	-1.48	-2.18	8.48	7.45	7.26

Table S22: R code GAM equations for each site.

Site	GAM equation (R code)
Atlanta -UA	Oz = gam(O3~s(DOY,bs="cr",k=10)+s(NO2surf, bs = "cr", k = 10)+s(OMINO2,bs="cr",k=10)+s(Tmaxsurf,bs="cr",k=10)+s(Dist,bs="cr",k=10)+ +s(WD1000M,bs="cr",k=10)+s(WS1000M,bs="cr",k=10)+s(MIX1000M,bs="cr",k=10)+ +s(WS1000A,bs="cr",k=10)+s(RH1000A,bs="cr",k=10)+ s(DOW,y),data=dat1,action=na.exclude)
Chicago-CHIW	Oz = gam(O3~s(DOY,bs="cr",k=10)+s(NO2surf, bs = "cr", k=10)+s(Merratmax,bs="cr",k=10) +s(Merratavg,bs="cr", k=3)+s(OmiUV,bs="cr",k=10)+s(OMINO2,bs="cr",k=10)+ s(Tmaxsurf,bs="cr",k=10) +s(Dir,bs="cr",k=10)+ s(CAPEM,bs="cr",k=10)+s(WD1000A,bs="cr",k=10),data=dat1,action=na.exclude)
Dallas-GRP	Oz = gam(O3~s(DOY,bs="cr",k=10)+s(NO2surf, bs = "cr", k = 10)+s(Merratmax,bs="cr",k=10) +s(OMINO2,bs="cr",k=10)+s(Dist,bs="cr",k=10)+s(T1000M,bs="cr",k=10)+s(WD1000M,bs="cr",k=10)+ s(WD1000A,bs="cr",k=10)+s(RH1000A,bs="cr",k=10)+s(DOW,y),data=dat1,action=na.exclude)
Detroit-EAST	Oz = gam(O3~s(DOY,bs="cr",k=10)+s(NO2surf, y)+s(OmiUV,bs="cr",k=10)+s(Tmaxsurf,bs="cr",k=10) +s(Dist,bs="cr",k=10)+s(Dir,bs="cr",k=10)+ s(WD1000M,bs="cr",k=10)+ s(MIX1000M,bs="cr",k=10)+ s(T1000A,bs="cr",k=10)+s(WS1000A,bs="cr",k=10),data=dat1,action=na.exclude)
Detroit-PORT	Oz = gam(O3~s(DOY,bs="cr",k=10)+s(NO2surf, y)+s(OmiUV,bs="cr",k=10)+s(Dir,bs="cr",k=10)+ s(WD1000M,bs="cr",k=10)+s(MIX1000M,bs="cr",k=10)+ s(T1000A,bs="cr",k=10),data= dat1,action=na.exclude)
Detroit-OAK	Oz = gam(O3~s(DOY,bs="cr",k=10)+s(NO2surf,y)+s(OmiUV,bs="cr",k=10)+s(Tmaxsurf,bs="cr", k=10)+s(Dist,bs="cr",k=10)+ s(WD1000M,bs="cr",k=10)+s(MIX1000M,bs="cr",k=10)+ s(WS1000A, bs="cr",k=10)+ s(RH1000A,bs="cr",k=10)+ s(DOW,y),data=dat1,action=na.exclude)
Detroit-NEW	Oz = gam(O3~s(DOY,bs="cr",k=10)+s(NO2surf, y)+s(OmiUV,bs="cr",k=10)+s(Tmaxsurf,bs="cr",k=10) +s(Dist,bs="cr",k=10)+s(Dir,bs="cr",k=10)+ s(MIX1000M,bs="cr",k=10)+ s(WD1000A,bs="cr",k=10)+ s(MIX1000A,bs="cr",k=10) ,data=dat1,action=na.exclude)
NYC-LEO	Oz = gam(O3~s(DOY,bs="cr",k=10)+s(NO2surf, bs = "cr", k = 10)+s(Merratmax,bs="cr",k=10)+ s(Merratavg,bs="cr", k=3)+s(Tmaxsurf,bs="cr",k=10)+s(Dist,bs="cr",k=10)+s(Dir,bs="cr",k=10)+ s(WS1000A,bs="cr",k=10)+s(RH1000A,bs="cr",k=10),data=dat1,action=na.exclude)
NYC-FAIR	Oz = gam(O3~s(DOY,bs="cr",k=10)+s(NO2surf, bs = "cr", k = 10)+s(Merratmax,bs="cr",k=10) +s(Tmaxsurf,bs="cr",k=10)+s(Dist,bs="cr",k=10)+s(Dir,bs="cr",k=10)+ s(T850M,bs="cr",k=10)+ s(WD1000M, bs="cr",k=10)+s(MIX1000M,bs="cr",k=10)+ s(WD1000A,bs="cr",k=10)+ s(RH1000A,bs="cr",k=10),data=dat1,action=na.exclude)
PHX-NP	Oz = gam(O3~s(DOY,bs="cr",k=10)+s(NO2surf, bs = "cr", k = 10)+s(OmiUV,bs= "cr",k=10) +s(OMINO2, bs="cr",k=10)+s(Tmaxsurf,bs="cr",k=10)+s(Dist,bs="cr",k=10)+ s(T700M,bs= "cr",k=10)+s(T1000M, bs="cr",k=10)+s(WD1000M,bs="cr",k=10)+s(WS1000M,bs="cr",k=10)+s(DOW,y),data=dat1,action=na.exclude)
SLC-HAW	Oz = gam(O3~s(DOY,bs="cr",k=10)+s(OMINO2,bs="cr",k=10)+s(Dist,bs="cr",k=10)+s(Dir,bs="cr",k=10)+ s(T1000M,bs="cr",k=10)+s(WD1000M,bs="cr",k=10)+s(MIX1000M,bs="cr",k=10)+s(T1000A,bs="cr",k=10) +s(WD1000A,bs="cr",k=10)+s(WS1000A,bs="cr",k=10)+s(RH1000A,bs="cr",k=10)+s(MIX1000A,bs="cr",k=10)+s(DOW,y),data=dat1,action=na.exclude)
SLC-HERR	Oz =gam(O3~s(DOY,bs="cr",k=10)+s(OMINO2,bs="cr",k=10)+s(Dist,bs="cr",k=10)+s(Dir,bs= "cr",k=10)+ s(WS1000M,bs="cr",k=10)+s(WS1000A,bs="cr",k=10)+s(RH1000A,bs="cr",k=10) ,data=dat1,action=na.exclude)
SLC-ERDA	Oz = gam(O3~s(DOY,bs="cr",k=10)+s(NO2surf, bs = "cr", k = 10)+s(OMINO2,bs="cr",k=10)+ s(Tmaxsurf,bs="cr",k=10)+s(Dir,bs="cr",k=10) ,data=dat1,action=na.exclude)
SLC-Bountiful	Oz = gam(O3~s(DOY,bs="cr",k=10)+s(Merratmax,bs="cr",k=10)+s(Tminsurf,bs="cr",k=10)+ s(Dir, bs="cr",k=10)+ s(T700M,bs="cr",k=10)+s(T1000M,bs="cr",k=10)+s(RH1000M,bs="cr",k=10)+s(WS1000A,bs= "cr", k=10)+s(RH1000A,bs="cr",k=10)+s(MIX1000A,bs="cr",k=10)+ s(DOW,y),data=dat1,action=na.exclude)
SF-LIV	Oz = gam(O3~s(DOY,bs="cr",k=10)+s(NO2surf, bs = "cr", k = 10)+s(Merratmax,bs="cr",k=10)+ s(Tminsurf,bs="cr",k=10)+s(Tmaxsurf,bs="cr",k=10)+s(Dist,bs="cr",k=10)+s(Dir,bs="cr",k=10)+s(T1000A,bs= "cr", k=10)+s(WS1000A,bs="cr",k=10)+s(RH1000A,bs="cr",k=10)+ s(DOW,y),data=dat1,action=na.exclude)