Authors' Response to interactive comments by 2 Anonymous Referees on "Estimating background contributions and U.S. anthropogenic enhancements to maximum ozone concentrations in the northern U.S." by David D. Parrish and Christine A. Ennis

The authors greatly appreciate the additional comments regarding our paper and the Editor's recommendations. The reviews of this paper have required much more time and effort than is usually the case, and the authors appreciate everyone's willingness to make these efforts. Responses to all new comments follow, and where appropriate the manuscript has been revised as described herein, and indicated in the "tracked changes" manuscript copy at the end of this response.

Point-by-Point Responses

Notes:

- The comments from the Referees are reproduced below in black regular font with our responses in *blue italic font.*
- The revised draft that was the subject of the present review included some changes that were not incorporated into Figure 14. That figure has now been updated in this new revision, and the discussion revised accordingly; this change did not lead to any changes in interpretation.

Report #3, by Anonymous Referee #4

August 2019 Review of acp-2018-1174

"Estimating background contributions and U.S. anthropogenic enhancements to maximum ozone concentrations in the northern U.S." by D. D. Parrish and C. A. Ennis

This manuscript is much improved! My concerns have been adequately addressed, and the readability is an order of magnitude better than the last version. This work is now presented in a way that others can understand and evaluate it. I look forward to the substantive and reasoned discussion which may follow in ACP and in the community at large.
Thank you for these positive comments, and the very great effort that you devoted to your two reviews.

Given the substantial amount of new text that the authors have generated, I have just a few comments that may improve the presentation.

A) The new section 2.3 describing the additional analyses was confusing to me.

- Should line 11 read "we also derive y0 through THREE somewhat different approaches that… " ?
- I recommend you remove the last sentence of the first paragraph, as it is repeated in the third paragraph.

These changes have been made. The first paragraph now reads:

"Acknowledging the uncertainty introduced by the assumptions required to implement the exponential analysis described in Section 2.2, we derive y_0 through three additional, somewhat different approaches that also provide two estimates of τ appropriate for the northeastern states."

• The third and fourth paragraphs seem to jump around a bit and are not presented in the same order as they are used later in the manuscript. Please see if this suggested re-organization describes what you intended:

Two additional approaches can approximately quantify the value of τ in the northeastern states; both of these approaches assume that constant values of y0 and τ are appropriate for all ODV time series included in each analysis. First, a linear fit to the initial period of decreasing ODVs provides direct information regarding the magnitude of τ and y0. The absolute value and the time derivative of Equation 1 when evaluated at year 2000 are y0 + A and –A/τ, respectively. Fits to two ODV time series provide four parameters (τ, y0, A1 and A2) if the τ and y0 values are the same for the two time series. Algebraic manipulation gives τ = - Δintercept / Δslope, where Δ indicates the difference in the subscripted parameter between the two linear fits, and y0 = (Σintercept + τ * Σslope)/2, where Σ indicates the sum of the subscripted parameter from the two fits. A complication with this approach is that the linear fits to time periods of significant length give biased measures of the derivative and year 2000 value of Equation 1; however, this bias can be corrected to first order through numerical comparison of a linear fit to the selected period of the exponential fit. The second approach is described in Section 2.4 of Parrish et al. (2017a) and is adapted here to the northeastern U.S. ODV time series. It uses an iterative, non-linear regression analysis that simultaneously derives values for τ and y0, plus the A parameter for each ODV time series included in the analysis. These two additional approaches help to constrain the uncertainty of the assumed value of τ (21.9 years).
Thank you very much for this suggested reorganizing; it has been incorporated exactly as suggested.

B) When the results of the method shown in Fig 13 are described in section 3.3.2, there is insufficient information to allow the reader to evaluate your "bias correction". The description of that "first order" correction "through numerical comparison of a linear fit" in section 2.3 does not shed much additional light. If it is possible to give a bit more information (or a reference?), the reader can have more confidence that your "corrections" reduce uncertainty, rather than increase it.
A 3-sentence explanation of the bias correction has been added to Section 3.3..2; see tracked changes manuscript for those changes.

And for clarity of communication, I would suggest that "intercept" is not strictly the correct noun here, as the plots you show in Figure 13 would have fitting parameters with an intercept at Year 0. I recommend you either replot with the horizontal axis as Year Minus 2000 instead of Year, or stick to using the descriptor "value" or "absolute value at Year 2000" (e.g., revised manuscript, page 8, line 4). This will also remove any confusion about if the "intercept" for the Maine data set is to be chosen at 2000 or at 1991.
Corrected by consistently using the descriptive phrase "absolute value at Year 2000" in place of "intercept" throughout the discussion of the method and its results.

C) "Variance" is discussed often in this manuscript, and it seems that the meaning is slightly different in different contexts. On page 6, variance (in units of ppb) is

described in terms of RMSD between a dataset and its fit. But then on page 9 and in Table 1, variance has units of ppb^2 and appears to be (but is never actually*) defined as the square of the standard deviation of a dataset. Then both types are used back and forth through the remainder of the manuscript. It's especially tricky to parse at the top of page 19, where you are using both the ppb^2 values (like 251 and 13.4), but are also referring to Figure 9, which gives RMSDs of 3.5 and 5.6 ppb. I don't have a nice, tidy suggestion, but I would ask that the authors take a few minutes to think about how they might make their dual-use of this word a little bit easier for the reader.
Thank you. On page 6 we have added the definition of variance as the square of the standard deviation of a dataset and we discuss that the square of the RMSD is an estimate of the variance not captured by a fit to Equation 1. We have gone through the paper to ensure that the use of "variance" is now consistent throughout.

*It is given quite succinctly on page 6 of the Response document, but I couldn't find it in the manuscript itself. I recommend including it in the manuscript, especially since the rounding effects are just enough to make it questionable (e.g., pg 18: 3.7 * 3.7 = 13.69 = 13.7, not 13.4).
And by the way, the Figure 3 legend says 252 ppb^2, not 251 ppb^2.
Thank you for the very careful reading. The definition of variance is now included (see previous response). However, we cannot repair the roundoff errors without giving an undue number of significant figures in the numbers. However, we now have 252 ppb^2 in the text as well as the figure legend.

D) And a few small suggestions:
Thank you for the very detailed reading evidenced by your suggestions below.
• Pg 9, line 14, add a comma: "Figures 3 and 4, and averages with standard deviations…"

Comma added.

• Pg 13, line 15 change to: Figure 10 plots the time series of these state maximum ODVs recorded in each year with respective fits over the 2010-2017 period.

Suggested change made.

• Is that really supposed to be 2010? In Fig 10, the solid fit lines start at 2000.

Thank you. It has now been changed to 2000-2017.

• Pg 13, line 19 change to: only the largest of the state's ODVs in a given year… {to match singular "across the state" later in the sentence"}

Suggested change made.

• Pg 13, line 27: is this "NYC urban maximum" a new subset? Or is it the same as the data which generated the red dashed line in Fig 7? This is not a big deal, but I got distracted for a while trying to figure it out.
This is a new data subset. I apologize for the confusion, but it seems necessary to include both the subset that generated the curve in Figure 7 (since this was the choice made when the subsets of data were initially chosen) and this "NYC urban maximum", which is designed to have as little

interannual variability about the exponential fit to optimize the analysis described in Section 3.3.1.

• And a related question: In Fig S13, Does the fact that the 1:1 line goes through NJ data points at all values of "NYC urban max" indicate that the "NYC urban max" data is actually entirely from NJ? (I understand the geography; that's not my point. If all the "reference" data is contained in the data plotted against it in Fig S13, that seems a bit circular.)

Actually many but not all "NYC urban max" are from the NJ data points. I do not think that this is circular. Rather it makes the fit in this figure more uncertain, because the method relies on the intercept of the fitted slope with the 1:1 line, which requires a slope differing from unity. The close correspondence between the "NYC urban max" and the NJ data points causes the fitted slope to approximate unity, giving an uncertain intercept. This comparison is nevertheless included for completeness.

• Pg 13, line 28, remove comma after "is selected".

Comma removed.

• Suggested clarification and nuance regarding spatial variability of ODVs on page 16: the derived US background ODV has significant variability on a continental scale. Within….. significantly smaller than in any of the western US regions, but shows no discernable spatial variability within this region. For context… NAAQS of 70 ppb. In contrast, in the northeastern US the A…

Clarifications added.

• Pg 18, top line has an extra space

Corrected.

• Pg 18, line18, strike "by" before (Fiore et al., 2015)

Corrected.

• Pg 18, line 21: change "limit" to "limited"

Corrected.

• Competing Interests section needs to be updated, as does the last line of the Acknowledgements.

Corrected.

• Fig 9 caption: The explanation of dashed, dotted, and y0 is awkward. Maybe try rewriting with dashed, then dotted, then y0 last.

Revised.

• Fig 14, vertical axis label: Is this really US background, not NAB?

Good point; axis relabeled.

• Supplemental, pg 2, line 11: One…area

Corrected.

• Supplemental, line 21: some sites in the NY…

Corrected.

• Supplemental, line 32: neighboring states until 2013.

Corrected.

• Figs S3-S10: the gray circles are a bit too faint to really see. Could they be darker?

The lightest symbols have been darkened.

• Fig S10 legend, lower panel: green sites are called "interior" or "rural inland" elsewhere. I recommend choosing one of those in the lower legend for consistency.

Corrected.

• Fig S16: the blue dashed line is different from all the other blue dotted lines.

Corrected.

Report #1, by Anonymous Referee #3

First, I would like to thank the author for taking the time to consider many different pieces of feedback.

Thank you for this comment. Indeed the feedback has been very valuable to us in improving the manuscript.

This reviewer cannot overlook the assumptions necessary to support the implications and conclusions. In many ways, the additional mathematical and uncertainty estimates obfuscates the real problem. In my estimation, the value of the author's work is the predictive power and not the physical interpretation. Yet, the paper's implications and conclusions focus on the interpretation. The authors application and interpretation is not unlike Hubbert's peak oil curve, which likewise sparked much controversy around the pop cultural interpretation.

The similarity of your fit to Hubbert's peak oil curve has both strengths and weaknesses. Even when fitting observations, both curves equally support two competing interpretations. The first interpretation is, as the author focuses, y0 occurs when US contribution is zero. By analogy, Hubbert's peak oil curve can be interpreted as approaching zero when there is no oil is left in the US (only other countries). The second interpretation is that y0 is when US regulations have been applied to the sectors where and the extent to which it will be applied -- leaving some residual (USRES). By analogy, Hubbert's peak oil curve can be interpreted as no *practically recoverable* oil is left in the US. In the second interpretation, y0 is really the sum of y_{USB} + y_{USRES}. The strength of such the simple mathematical fits is their predictive power, and not the conjecture about the composition of the intercept.

Thank you for this interesting analogy, which we appreciate because we believe that it supports, rather than calls into question, our analysis and its interpretation. The prediction of the temporal evolution toward ultimate total possible petroleum production is analogous to the prediction of the temporal evolution toward complete elimination of the anthropogenic contributions to ozone in an urban area. However, Hubbert's peak oil hypothesis suffers from a major difficulty - the total recoverable petroleum reserves is quite uncertain, and is subject to technological or exploration driven "surprises". This quantity has increased by a significant factor over past decades; Wikipedia reports that "The ratio between proven oil reserves and current production has constantly improved, passing from 20 years in 1948 to 35 years in 1972 and reaching about 40 years in 2003." This is primarily due to improved petroleum recovery technology, and discovery of some new oil fields. When the ultimate total possible petroleum production (i.e., the proven reserves) is not accurately known, prediction of the temporal evolution toward that ultimate limit is very likely to be inaccurate.

In contrast, elimination of the anthropogenic ozone contribution requires the reduction of anthropogenic emissions of NOx emissions to zero. (Reduction of VOC emissions may also be important, but the presence of important quantities of biogenic VOCs implies that ultimately the

NOx emissions must be reduced to zero to completely eliminate the anthropogenic ozone contribution.) We do know very accurately the "proven NOx emission reserves" simply from measurements of ambient NOx or NOy measurements. For example, Pollack et al. (2013) show that NOx concentrations in California's South Coast Air Basin have decreased by a factor of ~4 between 1960 and 2010. Thus, we are at least about ¾ of the way toward complete elimination of all anthropogenic NOx emissions in at least that area (and similar relative decreases are likely in all U.S. regions, although we are not aware of similar analyses for other regions). Since ambient NOx concentrations in rural areas of the U.S. are small with respect to urban areas, there is no possibility of discovering unknown reservoirs of NOx emissions (and certainly no likelihood of improving technology increasing NOx emissions.) In summary, our predictive task is much less vulnerable to surprises than was Hubbert's.

There is value in the predictive power of the exponential model proposed. The exponential model does fit recent data and one may conjecture that it will fit near-term future reductions. As the author points out, the curve may represent diminishing reductions of US contribution. The predictive power is independent of the physical interpretation of the composition of y0. Parrish and Ennis project an attainment date of 2021 and that prediction is not predicated upon the physical interpretation y0, which does seem valuable -- as was their prediction for SoCal.
Thank you for these supportive thoughts.

Regarding your response to my critique of fitting to data for Massachusetts (old Fig 7, new Fig 6). If in 2003 you been applying your exponential model, you would have only had the historical record to that date. My visual interpretation of the historical record between 1984-2003 is that your curve would have had some limited explanatory power. If I am not much mistaken, your y0 would have been substantially higher (80 ppb?) because the record as a local plateau between 1995 and 2003. This illustrates clearly that the physical interpretation of y0 predicted from that segment (1984-2003) would have been better interpreted as y_{USB} + y_{USRES}. For the 2003 inflection point, it is easy to guess that is caused by the NOx SIP call. How can we justify assuming there are not other controllable sectors that would produce a similar inflection in the future?

Thank you for this clarification. Your visual interpretation is at least qualitatively correct. Figure 1 below reproduces Figure 3 from the previous response with the fits that you specify above replacing the earlier fits. The y_0 derived from the 1984-2003 record is substantially higher (68 ± 9 ppb), at least partly due to the reason you suggest. The y_0 derived from the total 1984-2017 record is also is substantially higher (60 ± 4 ppb).

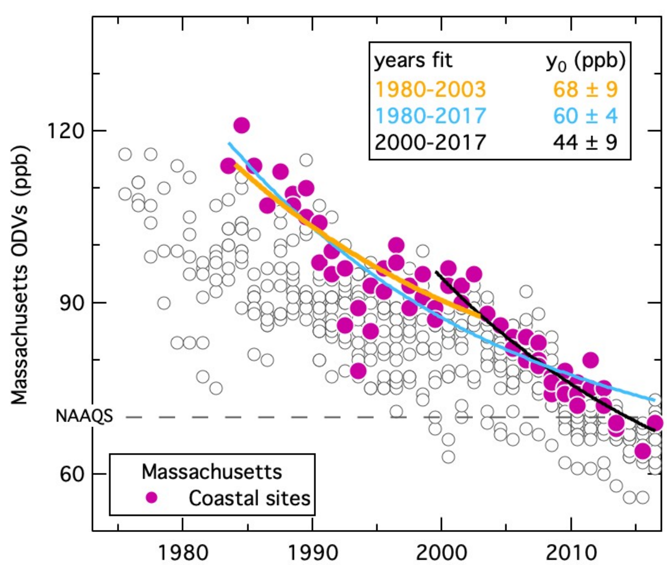


Figure 1. Analysis of the Massachusetts coastal data set (violet points from Figure 6 of the revised manuscript) over the time periods indicated in the figure annotation. The fits to Equation 1 for the different time periods are indicated by the superimposed, color-coded curves

We struggled with this issue throughout the analysis. Comparisons of all of the data sets, including the urban and rural areas of the northern U.S. as well as other regions of the country, indicated that the 2000-2017 period does give consistent results over nearly two decades of measurements for all data sets considered. This choice is thoroughly discussed in the manuscript.

The additional sections with all the analysis suggest that this fundamental problem of interpretation can be solved by more math. The largest uncertainty in y_{USB} is how well it is approximated by y_0. If y_{USRES} is large, then the uncertainty is swamped by the interpretation. If y_{USRES} is small, then the interpretation is reasonable. y_{USRES} is some combination of the technologically challenging reductions and, as illustrated by the previous paragraph, the sectors that may not yet have been the focus of controls. No effort is made to consider or estimate the magnitude of US emission sectors that are not decreasing. In response to a previous revision, this point was made by highlighting specific sectors that are not decreasing. No amount of regression uncertainty analysis will address the uncertainty associated with a physical interpretation.
As discussed above, we do have ambient NOx and VOC measurements that provide guidance regarding the magnitude of y_{USRES}, and all indications are that it is relatively small with the possible exceptions of intense, fertilized agricultural regions (e.g., California's Imperial Valley) and very densely populated urban centers, which may lead to the Connecticut and NYC ODV deviations from exponential decreases, as we discuss in the paper.

In my view, this manuscript has valuable analysis mixed with what seems like an arbitrary interpretation. The author has made token statements that allow for an alternative interpretation, but the conclusions and implications are so intertwined with the interpretation and those have not changed. This manuscript would be far easy to accept if the conclusions and implications it promotes/asserts were not dependent on an arbitrary physical interpretation of a parameter with an unknowable composition. Given that there are two equally likely physical interpretations ($y0 = y_{USB}$ or $y0 = y_{USB} + y_{USRES}$), this reviewer cannot support a manuscript whose conclusions and implications are based on a seemingly arbitrary choice between the two.
We believe that the two alternative physical interpretations are not equally likely. As discussed above and in much of the previous responses to reviews of this paper, all indications are that y_{USRES} is generally quite small, with the possible exceptions discussed in the response to the previous comment. Of course, only future monitoring results can definitively resolve this issue.

References:

Pollack, I. B., T. B. Ryerson, M. Trainer, J. A. Neuman, J. M. Roberts, and D. D. Parrish (2013), Trends in ozone, its precursors, and related secondary oxidation products in Los Angeles, California: A synthesis of measurements from 1960 to 2010, J. Geophys. Res. Atmos., 118, 5893–5911, doi:10.1002/jgrd.50472.

Estimating background contributions and U.S. anthropogenic enhancements to maximum ozone concentrations in the northern U.S.

David D. Parrish[1,2,3], Christine A. Ennis[4]

[1] Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, USA

[2] NOAA/ESRL Chemical Sciences Division, Boulder, Colorado, USA

[3] David.D.Parrish, LLC, Boulder, Colorado, USA

[4] 2B Technologies, Boulder, Colorado, USA

Correspondence to: David D. Parrish (David.D.Parrish@noaa.gov)

Abstract. U.S. ambient ozone concentrations have two components: U.S. background ozone and enhancements produced from the country's anthropogenic precursor emissions; only the enhancements effectively respond to national emission controls. We investigate the temporal evolution and spatial variability of the largest ozone concentrations, i.e., those that define the ozone design value (ODV) upon which the National Ambient Air Quality Standard (NAAQS) is based, within the northern tier of U.S. states. We focus on two regions: rural western states, with only small anthropogenic precursor emissions, and the urbanized northeastern states, which include the New York City urban area, the nation's most populated. The U.S. background ODV (i.e., the ODV remaining if U.S. anthropogenic precursor emissions were reduced to zero) is estimated to vary from 54 to 63 ppb in the rural western states, and to be smaller and nearly constant (45.8 ± 3.0 ppb) throughout the northeastern states. These U.S. background ODVs correspond to 65 to 90% of the 2015 NAAQS of 70 ppb. Over the past two to three decades U.S. emission control efforts have decreased the U.S. anthropogenic ODV enhancements at an approximately exponential rate with an e-folding time constant of ~22 years. These ODV enhancements are relatively large in the northeastern U.S. with state maximum ODV enhancements of ~35-64 ppb in 2000, but are not discernible in the rural western states. The U.S. background ODV contribution is significantly larger than the present-day ODV enhancements due to photochemical production from U.S. anthropogenic precursor emissions in the urban as well as the rural regions investigated. Forward projections of past trends suggest that average maximum ODVs in northeastern U.S. will drop below the NAAQS of 70 ppb by about 2021, assuming that the exponential decrease of the ODV enhancements can be maintained and the U.S. background ODV remains constant. This estimate is much more optimistic than in the Los Angeles urban area, where a similar approach estimates ~2050 for the maximum ODV to reach 70 ppb (Parrish et al., 2017a). The primary reason for this large difference is the significantly higher U.S. ODV background (62.0 ± 2.0 ppb) estimated for the Los Angeles urban area. The approach used in this work has some unquantified uncertainties that are discussed. Models can also estimate U.S. background ODVs; some of those results are

shown to correlate with the observational-based estimates derived here (r^2 values for different models are ~0.31 to 0.90), but they are on average systematically lower by 4 to 13 ppb. Further model improvement is required until their output can accurately reproduce the time series and spatial variability of observed ODVs. Ideally, the uncertainties in the model and observational based approaches can then be reduced through additional comparisons.

Deleted: 85

Deleted: 12

1 Introduction

The U.S. has a long-standing air quality problem associated with elevated ozone concentrations (e.g., NRC, 1991). Fortunately, this problem has been greatly improved over the past 3 to 5 decades, particularly in urban areas. For example, through the 1960s and 1970s the Los Angeles urban area (i.e., California's South Coast Air Basin – SoCAB) endured maximum 1-hr average and maximum daily 8-hr average (MDA8) ozone mixing ratios that exceeded 500 and 300 ppb, respectively (ppb = nmole ozone per mole air) (Parrish and Stockwell, 2015). The National Ambient Air Quality Standard (NAAQS) is based on the ozone design value (ODV), which is defined as the 3-year average of the annual fourth-highest daily maximum 8-hour average (MDA8) ozone concentration; in 2015 the NAAQS was lowered, now requiring that ODVs not exceed 70 ppb. A fit to the long-term trend of the maximum ODVs recorded in the SoCAB indicates that these highest ozone concentrations decreased from 289 to 102 ppb over the 36-year, 1980 to 2015 period (Parrish et al., 2017a). This decrease demonstrates that controls on U.S. ozone precursor emissions have been remarkably effective in reducing maximum ambient ozone concentrations. However, much additional emission reduction effort is required to reach the NAAQS of 70 ppb. A critical question has relevance to policy development for managing U.S. ozone concentrations: What is the limit to which ODVs can be reduced by controlling U.S. anthropogenic emissions? One goal of this work is to provide an observation-based estimate of this limit.

Both natural and anthropogenic processes interact to determine the temporal and spatial distribution of surface ozone concentrations in both urban and rural areas. Thus, even if U.S. anthropogenic emissions of ozone precursors were completely eliminated, ambient ozone concentrations throughout the U.S. would still be well above zero due to contributions from natural sources of ozone, enhanced by anthropogenic contributions from other countries. Parrish et al. (2017a) estimate that this remaining ODV (denoted as U.S. background ODV) would be 62.0 ± 1.9 ppb in the Los Angeles urban area. This contribution is the limit to which the ODVs can be reduced by U.S. emission controls alone; it is so large that there is little margin for enhancement of ambient ozone concentrations by photochemical production from U.S. anthropogenic precursor emissions before the NAAQS of 70 ppb is exceeded.

Two northern U.S. regions (maps in Figures 1, 2 and S1) are the focus of this work: eight northeastern states, which include the most populated U.S. urban area (New York City metropolitan area), and three sparsely populated, rural western states (Montana, North Dakota and South Dakota), containing no cities with >260,000 population. The temporal histories of ODVs measured in these two regions (Figure 3) correlate with the degree of urbanization - in the rural western states they remained approximately constant at relatively small values over the 39 years of measurements, while the largest ODVs with temporally decreasing values have been in the northeastern states. The northern tier of U.S. states also includes three Pacific Northwest states and three midwestern states (map in Figure S1) with intermediate ODV behavior (Figure S2); these regions are not examined in detail but are included here for comparison. Notably, none of the ODVs in these regions have approached the maximum ODVs recorded in the SoCAB (indicated by blue lines in Figures 3 and S2). There are three designated ozone nonattainment areas in the northern U.S. states (based on the 2015 ozone NAAQS - U.S. EPA's "Green Book" https://www.epa.gov/green-book, last accessed 8 July 2019), which include 38 counties in three of northeastern states - Connecticut, New Jersey and New York.

In this paper we apply the approach of Parrish et al. (2017a) to examine the temporal and spatial variability of the highest ozone concentrations (i.e., the ODVs) observed over the past three to four decades in the two contrasting regions of the northern U.S. representing extremes in anthropogenic influence. We separately estimate the U.S. background ODVs and the enhancements of the ODVs above that background contribution due to photochemical production from U.S. anthropogenic precursor emissions. The U.S. background ODV estimates quantify the maximum ozone concentrations that would exist in these regions in the absence of U.S anthropogenic precursor emissions. We also aim to quantify the temporal evolution and spatial variability of the US anthropogenic ODV enhancements, and based on past trends, project the expected time required for the maximum ozone concentrations to decrease to the 70 ppb NAAQS in the northeastern U.S.

Photochemical modeling systems are generally utilized for quantifications and projections of ODVs (e.g., Dolwick et al., 2015; Emery et al., 2012; Fiore et al., 2014). However, present model quantifications of U.S. ozone concentrations have large uncertainties (Jaffe et al., 2018; Guo et al., 2018). An observational-based approach such as presented here provides useful comparisons for the results of modeling efforts, and differences between the two approaches identify needs for further research.

The analysis approach in this paper relies on differences in the temporal behavior of the U.S. background ODV (demonstrated in this work to be approximately constant) and ODV enhancements resulting from US anthropogenic precursor emissions; these enhancements have greatly decreased over recent decades in response to U.S. emission controls. Previously published studies have identified a multitude of additional processes that potentially can make systematic contributions on a variety of time scales to the variability of ozone concentrations at U.S. surface sites; however there has been little in the way of systematic, quantitative analysis of their effects on ozone concentrations across the U.S. In this work, we first quantify the U.S. background ODVs and the temporal decrease of U.S. anthropogenic ODV enhancements, and then discuss the influence of other processes through examination of the fraction of the ODV variance not accounted for by decreasing U.S. anthropogenic ODV enhancements.

Papers investigating U.S. surface ozone trends (see Lin et al., 2017 and references therein) have treated a variety of statistics (medians, means, and various percentiles) to characterize ozone concentrations. In this work all trends are based on ODVs. The reason for this choice is that the NAAQS is based on this statistic, and thus it is most relevant for policy considerations. The ODV corresponds to ~98th percentile of the MDA8 concentrations during the ozone season. As a consequence, the U.S. background ODVs that we discuss are significantly larger than average or median background ozone concentrations examined in other studies. Given these different choices, care must be taken in comparing trends derived in this work with those from other analyses.

The sources of data and the analysis methods are discussed in the next section, followed by the applications of those methods to quantify the U.S. background ODVs and the U.S. anthropogenic enhancements in the rural western region (Section 3.1) and the northeastern U.S. (Section 3.2). The larger temporal ODV trends and the greater spatial variation of those trends in the northeastern U.S. provide the basis for the elucidation of several features of regional ozone concentrations. Section 3.3 examines the uncertainty of the analysis approach used in Section 3.2. Section 4 gives a summary of the approach and the results, discusses implications of those results, and identifies needs for further research.

2 Data and Methodology

2.1 Ozone Design Values analyzed

This work considers Ozone Design Values (ODVs) reported from the beginning of U.S. ozone monitoring in the mid-1970s through 2017 in seventeen northern U.S. states. An ODV, the statistic upon which the U.S. NAAQS is based, is calculated every year for each ozone monitoring station in the U.S. if the measurements achieve the specified completeness criteria. Each year all recorded ODVs are added to EPA's AQS data archive (https://www.epa.gov/aqs last accessed 23 June 2019). All ODVs reported for the northern states were downloaded from this archive; only the ODVs marked as valid were retained for analysis. Exceptional events that have concurrence from the U.S. EPA were excluded. Table S1 summarizes these archived ODVs for each state, including the number of monitoring sites, the years spanned by the reported ODVs, and their maximum and minimum values. The reported ODVs span the range from 169 ppb to 41 ppb. Yellowstone National Park (NP) in another state (Wyoming) is also included because its measurement record has been examined in previous analyses of long-term trends of U.S. background ozone concentrations (e.g., Lin et al., 2017). It should be noted that very few sites have continuous measurements over the indicated time spans, and that many sites operated for only short periods. All reported ODVs are included in this analysis, even if only a single ODV was reported for a particular site. It is implicitly assumed that the temporal discontinuities associated with initiation or termination of individual sites does not prevent an accurate quantification of temporal trends of ODVs within the regions selected for analysis.

2.2 Exponential ODV trend analysis

A well-established conceptual model (e.g., Parrish et al., 1986) guides our analysis. Ambient ozone concentrations at U.S. surface sites are composed of two contributions: 1) background ozone and 2) enhancements resulting from ozone produced from photochemical processing of U.S. anthropogenic emissions of ozone precursors. The first contribution is the ozone that would be present in the absence of U.S. emissions of ozone precursors from anthropogenic sources; this ozone is transported into the U.S. or produced over the U.S. from naturally emitted precursors. The U.S. Environmental Protection Agency (EPA) has defined this contribution as U.S. background ozone (e.g., Dolwick et al., 2015). The first contribution has remained relatively constant, while the second contribution has greatly decreased over the past 2 to 4 decades in response to reductions in anthropogenic emissions of ozone precursors.

In this work we focus on the time period of decreasing ODVs. Fitting observational data to a simple functional form is a common tool utilized for quantitative observational analysis; linear trend analysis (i.e., fitting observational data to a linear function) is one example. Here we choose to fit observed ODVs to Equation 1,

$$\text{ODV} = y_0 + A\ \exp\{-(\text{year-2000})/\tau\}, \tag{1}$$

with three undetermined parameters. This equation is the simplest possible functional form consistent with the guiding conceptual model of a background contribution and a consistently decreasing anthropogenic contribution. (A linear fit with only two undetermined parameters – slope and intercept – is simpler, but cannot fit a positive background contribution, as a decreasing linear fit will eventually go negative.) We identify the first term of Equation 1, y_0, as an estimate of the ODV that would result from U.S. background ozone alone (i.e., consistently called U.S. background ODV), and the second term as an estimate of the enhancement of observed ODVs above y_0 (i.e., consistently called U.S. anthropogenic ODV enhancement) due to contributions from photochemical processing of U.S. anthropogenic precursor emissions. This second term decreases exponentially with a time constant of τ, and equals A in the reference year, which we choose as 2000.

A simple intuitive argument suggests that an exponential decrease in the anthropogenic ozone contribution is expected to be a reasonable approximation for the response of maximum ozone concentrations to implementation of emission controls. When controls are initiated, early progress can be rapid since large existing emission sources evolved without planning for their control. With time, reducing emissions will become progressively more difficult, since the most easily controlled emissions will likely be addressed first, and the smaller, remaining emissions will be more difficult and/or expensive to control. This expected increasing difficulty in reducing emissions may well lead to an approximately constant fractional decrease in anthropogenic ozone enhancements, which corresponds to an approximately exponential decrease in these enhancements.

A previous analysis (Parrish et al., 2017a) quantified the temporal evolution of the maximum ODVs in seven southern California air basins over the 1980-2015 period (shorter periods beginning later and ending in 2015 in two basins). That work utilized fits to Equation 1 (with the reference year 1980 instead of 2000), and showed that a single value of $\tau = 21.9 \pm 1.2$ years, a single value of $y_0 = 62.0 \pm 1.9$ ppb, and a different value of A in each air basin provided an excellent fit ($r^2 = 0.984$) to the ODVs in all of those air basins.

As we will see in the following analysis, in the northeastern states the period of consistently decreasing ODVs (generally 2000 and later, hence the choice of 2000 as the reference year in Equation 1) is too short to allow precise determinations of all three parameters of Equation 1 from fits to individual ODV time series. In face of this difficulty, our primary analysis approach is to assume that the τ value (21.9 years) derived for southern California is also appropriate for the northeastern states. Uncertainty in the value of τ is then the greatest source of uncertainty in the analysis results; this impact of this uncertainty will be addressed in Section 3.3.2.

Equation 1 assumes that decreasing U.S. anthropogenic ODV enhancements is the only cause of ODV variability at a particular location. Other factors (e.g., rising anthropogenic emissions in Asia, variable occurrences of wild fires, interannual meteorological and climate variability, etc.) can also potentially affect observed ODVs. The approach taken here is to interpret the observed ODVs initially on the basis of Equation 1, and to examine the fraction of the ODV variance captured by that interpretation. The remaining fraction of the variance is then attributed to other factors, including those listed above. We use three statistics to quantify the variance in the total data set and the fraction not captured by Equation 1. The total variance in a data set is the square of the standard deviation of those data (in units of ppb^2). The root-mean-square deviation (RMSD) between the derived fit and the observed ODVs gives an absolute measure (in ppb) of the ODV variability about the fit; the square of the RMSD (in units of ppb^2) gives an estimate of the variance not captured by the fit. The square of the correlation coefficient (r^2) between the observed ODVs and the values derived from the fit to Equation 1 gives a measure of the fraction of the total variance that is captured by that fit; the difference between unity and the r^2 value is then a relative measure (as a fraction) of the ODV variance not captured by Equation 1. In the southern California air basins (Parrish et al., 2017a), the derived $r^2 = 0.984$ and RMSD ≈ 4 ppb indicate that all factors not included in Equation 1 account for no more than 1.6% of the total variance in the basin maximum ODVs analyzed in that work, and contribute a RMSD to those ODVs of no more than ~ 4 ppb.

Deleted: two

Formatted: Superscript

Deleted: not captured by Equation 1

Deleted: ODV

A potential complication in the interpretation of the two terms of Equation 1 arises if there is a significant fraction of U.S. anthropogenic ozone precursor emissions that has not been reduced by emission controls. Ozone produced from such emissions will not have decreased in the same manner as that produced from most U.S. anthropogenic emissions, which could raise the derived value of y_0 above the actual U.S. background ODV. Parrish et al. (2017a) have discussed this issue with regard to the emissions associated with the intense agricultural activity in the Imperial Valley of the Salton Sea air basin, where the derived

y_0 is higher than in other southern California air basins. The final section of this paper briefly considers the possible impact of this complication in the northeastern U.S. states.One difference between the application here and that of Parrish et al. (2017a) should be noted. The former work chose 1980 as the reference year, while here we choose the year 2000. The curves derived from the fits to Equation 1 and the values derived for the y_0 parameter do not depend on the choice of reference year, while the values derived for the A parameter do. Consequently, comparing the A parameters derived here with those given for California by Parrish et al. (2017a) requires adjustments for this difference, which can be provided through the second term of Equation 1.

2.3 Additional observation-based analyses of ODV time series

Acknowledging the uncertainty introduced by the assumptions required to implement the exponential analysis described in Section 2.2, we derive y_0 through three additional, somewhat different approaches that also provide two estimates of τ appropriate for the northeastern states.

An independent analysis approach discussed in Section 2.3 of Parrish et al. (2017a) can estimate U.S. background ODVs without assuming any specific functional form for the time dependence of the ODV enhancements. Different assumptions underlie this analysis - namely that all of the ODV time series under consideration follow the same functional form, but not necessarily an exponential decrease, and that all time series are approaching a common U.S. background ODV (i.e., y_0 value). These assumptions imply that all of the time series will converge to a common ODV as anthropogenic precursor emissions are reduced to zero; this common ODV is necessarily the regional U.S. background ODV. In practice this analysis uses correlations between time series of ODVs with U.S. anthropogenic ODV enhancements that differ as much as possible. One time series is selected as a reference; in the examples discussed here the time series with the largest U.S. anthropogenic ODV enhancements is selected. Other time series are then linearly correlated with this reference. The intercept of each linear correlation with the 1:1 line then provides an estimate of the U.S. background ODV; at that point the ODVs from the two time series are equal. Parrish et al. (2017a) show that the results of this approach for seven southern California air basins are nearly identical to the results from fits to Equation 1. We apply this approach to estimate U.S. background ODV in the northeastern U.S. and compare the results to those from the exponential analysis.

Two additional approaches can approximately quantify the value of τ in the northeastern states; both of these approaches assume that constant values of y_0 and τ are appropriate for all ODV time series included in each analysis. First, a linear fit to the initial period of decreasing ODVs provides direct information regarding the magnitude of τ and y_0. The absolute value and the time derivative of Equation 1 when evaluated at year 2000 are $y_0 + A$ and $-A/\tau$, respectively. Fits to two ODV time series provide four parameters (τ, y_0, A_1 and A_2) if the τ and y_0 values are the same for the two time series. Algebraic manipulation gives $\tau = -\Delta_{\text{year 2000 value}} / \Delta_{\text{slope}}$, where Δ indicates the difference in the subscripted parameter between the two linear fits, and

Deleted: also

Deleted: two

Deleted: These τ values offer insight into the uncertainty of the value of 21.9 years assumed for τ in the exponential analysis.

Formatted: Space Before: 0 pt, After: 0 pt

Moved (insertion) [1]

Deleted: B

Moved up [1]: Both of these approaches assume that constant values of y_0 and τ are appropriate for all ODV time series included in each analysis.

Deleted: – analysis of linear fits to segments of ODV time series, and simultaneous least square regression fits to several ODV time series; these approaches help to constrain the uncertainty of the assumed value of τ (21.9 years). Both of these approaches assume that constant values of y_0 and τ are appropriate for all ODV time series included in each analysis. Section 2.4 of Parrish et al. (2017a) describes an iterative, non-linear regression analysis that simultaneously derives values for τ and y_0, plus the A parameter for each ODV time series included in the analysis. This analysis will be adapted to the northeastern U.S. ODV time series. ¶
Finally,

Deleted: intercept

$y_0 = (\Sigma_{\text{year 2000 value}} + \tau * \Sigma_{\text{slope}})/2$, where Σ indicates the sum of the subscripted parameter from the two fits. A complication with this approach is that the linear fits to time periods of significant length give biased measures of the derivative and year 2000 value of Equation 1; however, this bias can be corrected to first order through numerical comparison of a linear fit to the selected period of the exponential fit. The second approach is described in Section 2.4 of Parrish et al. (2017a) and is adapted here to the northeastern U.S. ODV time series. It uses an iterative, non-linear regression analysis that simultaneously derives values for τ and y_0, plus the A parameter for each ODV time series included in the analysis. This analysis will be adapted to the northeastern U.S. ODV time series. These two additional approaches help to constrain the uncertainty of the assumed value of τ (21.9 years).

Deleted: intercept

Deleted:

Formatted: Font color: Text 1

2.4 Confidence limits and uncertainties

Formatted: Line spacing: 1.5 lines

In this work we consistently give 95% confidence limits for derived parameters, unless indicated otherwise. Most of the analysis in this work is based on non-linear, least-squares regression fits of the archived ODVs to Equation 1, and interpretation of the derived values for y_0 and A. In this interpretation it is important to properly consider the uncertainty of these values. We begin with the 95% confidence limits given by the least-squares fitting routines, which are then adjusted to account for the known covariance between the recorded ODVs. Each ODV is a three-year running mean; therefore only every third ODV is independent from the others determined at a given site. Consequently, the number of independent ODVs in each fit is approximately a factor of three smaller than the number of reported ODVs. Thus, all confidence limits derived from the fitting routines have been increased by a factor of $3^{1/2}$ to account for this covariance. Note that the confidence limits are typically 1 to a few ppb; thus results and their confidence limits are often given to 0.1 ppb precision, even though the last significant figure is likely not justified.

There are additional sources of covariance between the ODVs included in any particular fit. The ODVs from different sites within a region can co-vary due to regionally coherent interannual variability, and interannual variability may lead to covariance between ozone concentrations measured in successive years. We are not able to account for the effect of this additional covariance; the derived confidence limits are thus lower limits for the true confidence limits of the derived parameters. However, as discussed in the next section, we can find no indication that additional regional or temporal covariance of the ODVs makes significant contributions to the uncertainties of the results. The influence of often cited major drivers of temporal variation of ozone, which could possibly cause such covariance, is discussed in Section 4, and found to be small.

3 Results

Here we examine the time series of ODVs from the western rural states (Section 3.1), fit the time series of ODVs from the northeastern states to Equation 1 (Section 3.2) and discuss the results in the context of the conceptual model introduced above. This model considers the recorded ODVs to comprise two contributions: 1) an approximately constant U.S. background ODV identified with y_0 in Equation 1, and 2) U.S. anthropogenic ODV enhancements, which are approximated by the second term in Equation 1. Section 3.2 and Section S1 of the Supplement discuss further details of the spatial and temporal variability of ODVs in the northeastern states. Section 3.3 uses the alternative approaches described in Section 2.3 to examine the uncertainty inherent in the parameter determinations from the exponential analysis using Equation 1.

3.1 ODVs in rural western states

The sparsely populated, three-state, rural western region generally lies on the Northern U.S. Great Plains downwind of more mountainous terrain to the west. Figure 1 shows a topographical map of the region, with the locations of the ozone monitoring sites indicated. This area gradually slopes to the east and north. All of the monitoring sites lie below 1.55 km elevation, with the exception of Yellowstone NP at 2.43 km.

The histories of the ODVs recorded in the region are illustrated in Figures 3 and 4, and averages with standard deviations and variances are given in Table 1. The gaps in the Montana and South Dakota records were caused by extended periods when no valid ODVs were recorded at any site within the respective states. Throughout the ODV record there is little variability due to any cause. The 283 tabulated ODVs recorded over 39 years at 35 sites in the three states average 59.3 ppb with a standard deviation of 3.7 ppb (corresponding to a variance of 13.4 ppb^2) – strong evidence that the ODVs correspond to an approximately constant U.S. background ODV within this region with no evidence for significant U.S. anthropogenic ODV enhancements. At the individual sites and within each state the entire measurement records are all well described by averages with small standard deviations (Table 1): < 3 ppb in Montana and North Dakota, and < 4 ppb in South Dakota, the state whose sampling sites span the largest elevation range (0.34 to 1.55 km). U.S. background ODVs generally increase with the elevation of the sampling site (e.g., see discussion in Jaffe et al., 2018), so larger variability is expected when the monitoring sites within a state span a larger range of elevations. The state averages in Table 1 lie within a range of ~6 ppb, but there are some significant differences: a maximum in South Dakota (61.5 ppb) and a minimum in Montana (55.4 ppb), with North Dakota intermediate (59.3 ppb). Consistent with the site elevation differences, the average ODV at Yellowstone NP is significantly larger than that at Glacier NP: 64.0 ppb at 2.43 km and 54.5 ppb at 0.96 km, respectively. The variances of the data sets vary from 2 to 15 ppb^2; these values indicate that only small variance in long-term ODV records can arise from variation of U.S. background ozone alone, at least in this particular region of the country.

Deleted: very

3.2 Exponential fits to ODVs in northeastern states

A topographical map showing the networks of ozone monitoring sites in the eight northeastern U.S. states is given in Figure 2. All of the ODVs recorded in four of the eight states are plotted in Figures 5 and 6, along with curves showing fits of Equation 1 to the ODVs from selected groups of sites over selected time periods. These ODV time series are in striking contrast to those in the rural western states (compare Figures 5 and 6 with Figure 4) with much larger concentrations showing strong decreases over the past two to three decades, and much greater variability of ODV values. We attribute this contrast to the much greater influence of U.S. anthropogenic ODV enhancements in the northeastern states. The greater variability is quantitatively reflected in the ODV variance in this region (252 ppb^2), which is nearly a factor of 20 larger than that seen in the rural western states; this comparison shows the dominant influence of the U.S. anthropogenic ODV enhancements in the northeastern states.

Deleted: 251

The four states included in Figures 5 and 6 are shown for illustrative purposes, with Figures S3-S10 of the Supplement showing detailed ODV temporal plots and fitted curves to the selected groups of sites in all eight states. These groups of sites were selected to represent different environments within each state, with the expectation that similar temporal ODV trends will be found at all sites within each group. The strategy adopted is to fit the ODVs recorded at all sites within each group over the time period beginning when a clear, consistent decrease in ODVs is first established, and continuing through 2017, the most recent ODVs available. This strategy is required since Equation 1 is designed to provide fits to ODVs only during such periods of consistent decreases. In all cases these fits begin by 2000, with some beginning earlier - either at the start of measurement record, 1990 or 1995 determined by the best, consistent fit to the functional form of Equation 1. Figures S3-S10 include maps indicating the locations of all selected groups of sites. In all, seventeen groups within the eight states were selected; they are listed in Table 2 along with the parameters derived from the fits of Equation 1.

There are some consistent general features of the ODV time series and the corresponding fits that inform the following analysis:

- Throughout the measurement record, the largest ODVs are found in the states that contain the New York City metropolitan area (New York, New Jersey and southwestern Connecticut), or that lie directly downwind (coastal Connecticut and Long Island, New York). Such sites compose two of the selected groups of sites in New York and Connecticut (see highlighted points in that area in the map of Figure 2), whose ODVs and fits of Equation 1 are highlighted in Figure 5.
- In several states, the largest ODVs are recorded at coastal sites (i.e., Connecticut, Massachusetts, New Hampshire and Maine in Figures 5, 6, S5, S7, S8 and S10). The large ODVs at coastal sites emphasize the important, widely-discussed (e.g., Wolff and Lioy, 1980; Wilcox, 1996) role of transport in bringing high ozone concentrations from the major East Coast urban areas far downwind, particularly when that transport occurs over the waters of Long Island Sound and the Coastal Atlantic Ocean. Two relatively isolated Massachusetts coastal sites on the offshore island of Martha's Vineyard and near the tip of Cape Cod record some of the highest ODVs within that state (see Figure S7). Dukes County, which includes only Martha's Vineyard, with a total population of ~17,000 was once designated as a marginal non-attainment area for ozone.
- In the past, ODVs at rural, generally upwind sites on the western border of New York (green symbols in on the left in Figure 2) were significantly smaller than in the northeastern U.S. urban areas, although in recent years that difference has diminished (Figure 5). These upwind rural areas in New York, and similar sites in Vermont (Figure S9), experienced ozone concentrations exceeding 80 ppb throughout the measurement record until about 2005. These high concentrations caused Chautauqua County, N.Y., with a population of ~95,000, to also once be designated as a marginal non-attainment

area, again emphasizing the importance of ozone transport in the northeastern U.S., although in this case the source of the transported ozone is not as clearly established.

Additional systematic features of the ODV time series in the northeastern U.S. are discussed in Section S1 of the Supplement.

All of the curves derived from the fits of Equation 1 to the long-term trends of the ODVs shown in Figures 5, 6 and S3-S10 are compared in Figure 7, with the corresponding parameters included in Table 2. Except for the four fits denoted by the colored dotted and dashed curves, all fits are similar in the sense that they exhibit the same relative long-term decrease and are asymptotically approaching approximately the same value of y_0. The same relative long-term decrease is necessarily forced by the use of the same value of $\tau = 21.9$ years in all fits. However, the derived A and y_0 values do provide information regarding the spatial and temporal variation of ODVs over the past two to three decades. Three of the four curves with noticeably different behavior are from fits to the groups of sites with the highest recently reported ODVs (Connecticut, especially the coastal sites, and the New York sites highlighted in Figures 2 and 5); these are discussed further in Section S1 of the Supplement. The fourth exception is the one high elevation site (Mt. Washington in New Hampshire at an elevation of 1.9 km), which is also discussed separately in in Section S1. The parameters in Table 2 provide the basis for quantitatively comparing the fits throughout the northeastern U.S. in the next two sections.

3.2.1 Estimation of U.S. background ODV in northeastern states from exponential fits

All y_0 values in Table 2 (excluding the four exceptions indicated in Figure 7) agree with each other within their indicated confidence limits. The arithmetic mean of these y_0 values is 45.9 ppb with a standard deviation of 3.2 ppb. The average of these y_0 values weighted with the inverse square of the respective confidence limits is 45.8 ± 1.7 ppb, where the 95% confidence limit of this average is indicated. All of the y_0 values in Table 2 agree (again excluding the four exceptions noted above) with these average values within their indicated confidence limits. Figure S11 of the Supplement shows the distribution of the y_0 determinations; 13 of the 17 derived y_0 values approximately define a normal distribution with a median of 47.7 ppb and a standard deviation of 4.5 ppb. The median is interpreted as representing a common regional y_0 value, and the standard deviation as reflecting the uncertainty in determining each y_0 value. This median is consistent with the above averages. The highest 4 of 17 derived y_0 values define a high value tail; these are the 4 four exceptions indicated in Figure 7.

Recalling earlier discussion, we identify the average $y_0 = 45.8$ as the best estimate of the U.S. background ODV throughout the northeastern U.S.; there is no discernable spatial variability within this region. This value is significantly smaller than the value of 62.0 ± 1.9 ppb derived for southern California (Parrish et al., 2017a); however even at this smaller value, the U.S. background ODV in the northeastern U.S. amounts to 65% of the 70 ppb NAAQS.

3.2.2 Estimation of U.S. anthropogenic ODV enhancements in northeastern states from exponential fits

The fits to Equation 1 with $\tau = 21.9$ years provide estimates of A, the U.S. ODV enhancement in the reference year 2000; Table 2 lists these values for the 17 selected groups of sites from two-parameter fits, i.e., fits with y_0 and A as independent parameters determined from the least-squares fits themselves. However, the results above show that a constant value of $y_0 = 45.8 \pm 1.7$ ppb is characteristic of the entire northeastern U.S. region. Using this result allows fits of Equation 1 to all groups of sites without the larger uncertainty in the y_0 derived from the individual fits. Consequently, results of one parameter fits of Equation 1 (i.e., with y_0 held constant at the value of 45.8 ppb) are included in Table 1 as the A^* values. (Such a fit is not included for the Mt. Washington results, since U.S. background ODV is evidently greater than 45.8 ppb as discussed in Section S1 of the Supplement). The A^* values generally agree with the A values from the two-parameter fits within their confidence limits, which are smaller since only one parameter need be derived. The exceptions to the agreement between A and A^* are the fits to the exceptions discussed earlier - the two groups of Connecticut sites and the New York maximum ozone sites, which are the upper three colored curves in Figure 7. In Table 2 the A values for these three groups of sites are anomalously small compared to the results from neighboring groups of sites (i.e., New Jersey, Rhode Island, Massachusetts/coastal); the A^* values for all of these neighboring groups of sites agree more closely. In the following discussion we take these A^* values as the best estimate for the U.S. anthropogenic ODV enhancements in the northeastern states.

A contour plot (Figure 8) derived from the A^* values in Table 2 provides an overview of the spatial variation of the U.S. anthropogenic ODV enhancements across the northeastern U.S. The groups of selected sites fit to Equation 1 give only coarse spatial resolution across the region, so the contour plot has uncertainties not apparent from the smooth spatial variability of this figure. This uncertainty has been mitigated in deriving the contour plot by including duplicate A^* values at the site locations in each selected group of sites; these additions ensure that the contouring program reproduces a more nearly constant value over the sometimes large regions covered by the selected groups of sites. Despite the uncertainties, the contour plot does give a useful, semi-quantitative representation of the magnitude and regional variation of the U.S. anthropogenic ODV enhancements in the region. Note that the contour plot and the A and A^* values of Table 2 describe the ODV enhancements in the year 2000. As is apparent from Equation 1 and the illustrated temporal trends in the figures, the ODVs have decreased throughout the last two to three decades. The e-folding time of $\tau = 21.9$ years implies that between the reference year of 2000 and 2017, the ODV enhancements decreased by a factor of 2.2. Hence, dividing the year 2000 ODVs in the contour plot by that factor gives an approximation of the 2017 U.S. anthropogenic ODV enhancements.

The ability of Equation 1 to accurately reproduce observed ODVs can be judged by comparing the observed ODVs with the values predicted from the fits derived with $y_0 = 45.8$ ppb and $\tau = 21.9$ years. Figure 9a shows this comparison as a correlation plot. The fits for ODVs recorded at all sites in the eight northeastern states over the entire measurement period are calculated from the A^* values at each site interpolated from the contour plot of Figure 8. The correlation is high ($r^2 = 0.71$) for the 1719 separate ODV values recorded at the 148 sites over the 2000-2017 period, but significantly lower for earlier years as expected from the figures illustrating the derived fits. A general decrease in ODVs throughout the region did not begin until 2000, which is about the time that the U.S. EPA "NOx SIP Call" began reducing power plant NOx emissions across much of the eastern U.S. (Aleksic et al., 2013). There is significant scatter about the 1:1 line in the comparison in Figure 9a; the RMSD between observed and calculated ODVs is 5.6 ppb for the 2000-2017 period. Much of this scatter is due to variability in ODVs recorded at different sites within a given region, which arises from differences in local photochemical ozone production and transport patterns. This variability can be reduced by comparing state maximum ODVs (Figure 9b), rather than individual site ODVs. Figure 10 plots the time series of these state maximum ODVs recorded in each year with respective fits over the 2000-2017 period. The derived A^* values (given in Table 3) are somewhat larger than would be expected from the contour plot in Figure 8, consistent with consideration of only the maximum ODVs recorded in each state. Stronger correlation ($r^2 = 0.89$) is found for the fits to the state maximum ODVs as expected, since considering only the largest of the state's ODVs in a given year removes much of the regional variability across the state.

Deleted: for the eight northeastern states

Deleted: 2010

Deleted: '

3.3 Evaluation of uncertainty of the exponential fits to ODVs in northeastern states

Here the methods described in Section 2.3 are applied to investigate the uncertainty of the results from the exponential fits presented in Section 3.2. Section 3.3.4 provides an overall assessment of this uncertainty.

3.3.1 Alternative approach for estimating U.S. background ODVs in the northeastern U.S. states

The independent analysis approach introduced in Section 2.3 can estimate U.S. background ODVs through correlations between separate ODV time series. The ODVs from each of the 13 groups of sites that give the black lines in Figure 7 are included in this analysis. The reference ODV time series chosen is the maximum observed ODVs in the New York City urban area (NYC urban maximum), which is equated to the maximum ODV observed each year in either New York or New Jersey. These maxima (plotted in Figure 11a) are all recorded near the New York urban area. This reference is selected because these are among the largest ODVs recorded in the northeastern U.S., and after 2000 this time series closely follows an exponential decrease with little interannual variability. Figure 12 shows three example linear correlations (the ODVs recorded at the three sets of Massachusetts sites) with that reference. Figures S12-S18 of the Supplement show all of the linear regressions for the 13 regional data, Figure S19 compares all of the fits, and Table S2 collects the results. These results are quite variable (25 to 62 ppb) due to the relatively short 2000-2017 data records and because the slopes are not widely different from unity, preventing a precise determination of the intercepts of the correlations with the 1:1 line. However, the average of the derived background ODVs (49.2 ± 3.9 ppb for ordinary linear regressions and 42.5 ± 5.7 ppb for reduced major axis regressions, where 95% confidence limits of the averages are indicated) bracket the result derived from the exponential fits, and neither average is statistically significantly different from that earlier result. The agreement between these two approaches for estimating U.S. background ODVs shows that the assumption of an exponential decrease in the ODV enhancements is not essential for estimating the background ODV (although that approach does give more precise results), and increases our confidence in the

Deleted: ,

results of each approach.

3.3.2 Estimate of τ and y_0 from linear fits to ODV trends in the northeastern U.S. states

Linear fits to the period of decreasing ODVs for three ODV time series are shown in Figure 13. These three series were chosen so that one (NYC urban maximum introduced in the previous section) includes the largest ODVs, and two have some of the smaller ODVs in the northeastern U.S.; this choice gives the largest contrast in the absolute year 2000 values and fitted slopes in order to provide the most precise τ and y_0 determinations. Table S3 gives the year 2000 values and slopes of those fits, which give zero-order estimates of $\tau = -\Delta_{\text{year 2000 value}} / \Delta_{\text{slope}}$ and $y_0 = (\Sigma_{\text{year 2000 value}} + \tau * \Sigma_{\text{slope}})/2$. However, as is apparent in Figure 13, the year 2000 value and slope derived from each linear fit over the 18- or 26-year period are biased with respect to the instantaneous value and slope of Equation 1 in year 2000. This bias can be estimated from linear fits over those same time periods to the exponential curves defined by the zero-order estimates of τ and y_0. Table S3 gives year 2000 values and slopes corrected to first order for this bias. These corrected values give τ = 21.1 ± 5.9 and 21.7 ± 5.0 years and y_0 = 48.7 and 47.0 ppb for the fit parameters from the upper and lower Figure 13 panels, respectively. These τ values compare favorably with the assumed California value (21.9 years), while the y_0 values are larger than derived in the analysis using exponential fits to Equation 1 (45.8 ± 1.7 ppb).

Deleted: and year 2000 intercepts

Deleted: and intercepts

Deleted: We calculate τ and y_0 after correcting for the bias between the instantaneous value and slope of Equation 1 in year 2000, and the intercept and slope derived from the linear fits over the 18- and 26-year periods. Table S3 gives t

Deleted: , which in turn

Deleted: quite

Deleted: to

3.3.3 Simultaneous least-squares regression to fits northeastern U.S. state ODV maxima

An iterative, non-linear regression analysis similar to that described in Section 2.4 of Parrish et al. (2017a) and introduced in Section 2.3 is applied here to simultaneously fit seven ODV time series to Equation 1 to determine nine parameter values. The data sets are the 2000-2017 maximum ODVs recorded in seven states plotted in Figure 10. A simultaneous fit to multiple ODV time series improves the precision of the parameter determinations. Values of τ and y_0 (assumed the same for all seven states) and values of A for each of seven states are optimized in an iterative process that minimizes the sum of the squares of the deviations between the fit and the original time series. The resulting parameter values are given in Table S4, and Figure S20 compares the fit results with the original ODVs. The derived τ value (26.0 ± 6.0 years) □□ larger than the southern California value of (21.9 ± 1.2 years), although it agrees within the derived 95% confidence limit. Correspondingly, the derived y_0 value (41.8 ± 3.0 ppb) is smaller than derived earlier (45.8 ± 1.7 ppb), and the A values are larger (compare to Table 3), but again all agree within the derived confidence limits. This fit captures 89.6 % of the variance in the seven ODV time series, comparable to the result shown in Figure 9b. (Note that since its recent ODV behavior is different from the other states, as discussed in Section 3.2, Connecticut is not included in this analysis.)

3.3.4 Assessment of uncertainty of the results

Section 3.2 presents fits of Equation 1 to ODV time series in the northeastern U.S. derived with an assumed value for τ; all confidence limits given for the derived parameters are lower limits due to this assumption. The above analyses in this Section

3.3 investigate alternative approaches to better constrain the overall uncertainty of the results. With regard to the value of τ, the analysis of Section 3.3.2 gives two values (21.1 ± 5.9 and 21.7 ± 5.0 years) that agree well with the assumed value (21.9 ± 1.2 years) derived by Parrish et al. (2017a) from analysis of ODVs in southern California, while the analysis of Section 3.3.3 gives a larger value (26.0 ± 6.0). Importantly, all of these derived τ estimates agree within their indicated confidence limits, indicating that there is no evidence for a different exponential rate of decrease of U.S. anthropogenic ODV enhancements between southern California and the northeastern states.

With regard to the value of the U.S. background ODV (y_0), Section 3.2 gave 45.8 ± 1.7 ppb using the assumed fixed τ value. The alternative approach of Section 3.3.1 gives two results, 49.2 ± 3.9 ppb and 42.5 ± 5.7 ppb, depending upon the linear fitting approach used, Section 3.3.2 gives two estimates of 48.7 and 47.0 ppb (without easily defined confidence limits), and Section 3.3.3 gives 41.8 ± 3.0 ppb. The average of these five results is 45.8 ± 3.0 ppb, which agrees well with the Section 3.2 result. This average value with the wider confidence limit is taken as the best estimate of y_0.

4 Discussion and Conclusions

The analysis presented in this paper is applied to ODVs from eight northeastern U.S. states, and contrasted with ODVs from three sparsely populated rural western states in the northern U.S. (maps in Figures 1 and 2); it has two complementary parts. First, time series of the highest ozone concentrations (i.e., the ODVs, the statistic upon which the NAAQS is based) in the northeastern states are fit to Equation 1. This equation has two terms - one constant and one exponentially decreasing – with two variable parameters: y_0, the magnitude of the constant term; and A, the year 2000 magnitude of the decreasing term. The fits are limited to the most recent two to four decades, when the ODVs are consistently decreasing, and we assume an e-folding time of τ = 21.9 years in Equation 1 in these fits. The success of the fitting process is judged through standard statistical tests that quantify how well the fits capture the variability of the ODV time series, and quantify the uncertainty of the derived parameter values. The second part of the analysis is the physical interpretation of the parameters derived from the fits to Equation 1; y_0 is taken as an estimate of the U.S. background ODV (i.e., the ODV that would exist in the absence of U.S. anthropogenic emissions of ozone precursors), and the second term is interpreted as an estimate of the regional U.S. anthropogenic ODV enhancement (i.e., the amount that ODVs are enhanced above the U.S. background ODV by photochemical production of ozone from existing U.S. anthropogenic precursor emissions). Several alternative analyses are presented to compare with the primary analysis of the exponential fits.

The northeastern states contain major urban centers while the western rural states contain no large cities, leading to marked differences in the ODV time series. In the rural western states the ODVs recorded at 35 sites over a 39-year period show remarkably little variability (Figures 3 and 4), with an overall standard deviation of 3.7 ppb (variance of 13.4 ppb^2). In contrast, the ODVs recorded in the northeastern states vary from >160 to <50 ppb (Figures 3, 5, 6 and S3-S10) with an overall standard deviation of 16 ppb (variance of 252 ppb^2). The derived U.S. background ODV has significant spatial variability on a continental scale. Within the rural western states ODV averages (Table 1) quantify the U.S. background ODV; the values for the three states (55 to 62 ppb) are similar to the value of 62.0 ± 1.9 ppb derived by Parrish et al. (2017a) for large areas of southern California, including the Los Angeles urban area. The U.S. background ODV in the northeastern U.S. states (45.8 ± 3.0 ppb) is significantly smaller than in any of the western U.S. regions, but shows no discernible spatial variability within this region. For context, these U.S. background ODVs account for 65 to 90% of the 2015 NAAQS of 70 ppb. In contrast, in the northeastern U.S. the *A* parameter (representing the U.S. anthropogenic ODV enhancement) varies spatially as shown by the contour plot in Figure 8, with the largest values (>54 ppb) immediately downwind of New York City decreasing to <22 ppb over northeastern Maine. Importantly, these derived *A* parameters quantify the U.S. anthropogenic ODV enhancements in the year 2000. By 2017 these enhancements had decreased by a factor of 2.2 according to our analysis; thus the largest ODV enhancements immediately downwind of New York City have decreased to ~25 ppb. No signficant anthropogenic ODV enhancements are present in the rural western states.

Deleted: 251

4.1 Implications of the results for air quality

The analysis presented here and the results of Parrish et al. (2017a) demonstrate that throughout diverse regions of the country (i.e., rural western states, northeastern U.S., and southern California) the U.S. background ODV contribution is significantly larger than the present-day ODV enhancements due to photochemical production from U.S. anthropogenic precursor emissions. This comparison is true not only in rural areas, but also in the two most populous U.S. urban areas, New York City and Los Angeles. Since these ODVs, upon which the NAAQS is based, represent the largest observed ozone concentrations, degraded air quality due to elevated ozone concentrations is attributed primarily to the U.S. background ODV, with local and regional photochemical production from U.S. anthropogenic precursor emissions enhancing that background by a significant, but smaller amount.

Forward projections of the fits to the maximum ODVs (shown in Figures 10 and 11a) allow an estimate of future trends of ODVs in the northeastern U.S., assuming that the U.S. background ODV (i.e., y_0) remains constant at 45.8 ppb throughout the region, and that the exponential decrease of the U.S. anthropogenic ODV enhancements can be maintained with an e-folding time, τ, of 21.9 years, by means of continued emission reduction efforts. These projections suggest that the maximum ODVs throughout the northeastern U.S. will drop below the 2015 NAAQS of 70 ppb by about 2021. However, these projections do not account for the variability of observed maximum ODVs (i.e., RMSD of 3.9 ppb in the northeastern U.S.) about the fitted curves, so that even after 2021 this variability will likely result in the occasional recording of ODVs above 70 ppb.

These forward projections cannot account for any systematic deviations of the ODVs from the behavior given by Equation 1. The recent temporal evolution of ODVs in Connecticut appears to differ significantly from the general regional behavior (see

Figures 5-7 and 10). In the discussion of the fit to Equation 1 of the Connecticut ODVs this difference was noted (see dashed colored curves in Figure 7), but nevertheless the temporal evolution was forced with y_0 = 45.8 ppb in deriving the A^* values given in Table 2 and in deriving the contour plot of Figure 8. The different behavior and fits for Connecticut are due to the most recent five years of ODV values lying above the expected trend, as most clearly shown in Figure 10. The cause of this difference is not understood. Whether this difference is simply a statistical fluctuation cannot be determined at this time; however, random fluctuations of similar magnitude are only rarely apparent in the temporal records of ODVs in the states discussed. McDonald et al. (2018) have recently discussed a class of ozone precursor emissions, i.e., volatile chemical products - including pesticides, coatings, printing inks, adhesives, cleaning agents, and personal care products - that have not been addressed by emission controls to the same extent as other emission sectors. The impact of this emission sector on ODVs has not been quantified, but is expected to be most significant in areas of largest population density, exactly the regions where the significant differences in temporal evolution of ODVs are noted.

The higher U.S. background ODV (y_0) in southern California of 62.0 ± 1.9 ppb (Parrish et al., 2017a) compared to the value of 45.8 ± 3.0 ppb derived here for the northeastern U.S. implies much less difficulty in achieving the 2015 ozone NAAQS of 70 ppb in the New York City (NYC) urban area compared to Los Angeles (LA), because the northeastern U.S. has a much larger margin for U.S. anthropogenic enhancement of ODVs while still attaining the NAAQS. Figure 11 compares the U.S. background ODVs and the maximum ODVs in these two urban areas. In 2015 these curves indicated maximum ODVs of 78 and 102 ppb in NYC and LA, respectively. To lower the maximum ODVs to 70 ppb would require respective decreases in total ODVs of 10% in NYC and 31% in LA. However, only the U.S. anthropogenic ODV enhancements can be addressed by local and regional controls of ozone precursor emissions. In 2015 these enhancements were about 25% larger in LA than in NYC (40 and 32 ppb, respectively). To reach a maximum ODV of 70 ppb requires ODV enhancement reductions of 25% in NYC and 80% (i.e. a factor of 5 reduction) in LA. The exponential term of Equation 1 projects that such reductions of the 2015 ODV enhancements will require 5 years in NYC and 35 years in LA; hence the projected years of 2021 and 2050 in NYC and LA, respectively. From the perspective of lowering maximum ODVs to the ozone NAAQS, the most important difference between NYC and LA urban areas is the higher U.S. background ODV in LA, although the 25% larger anthropogenic ODV enhancements in LA play a secondary role. This comparison provides an insightful context for the consideration of relative anthropogenic enhancements of ozone concentrations across the country.

Finally, it is important to note that from a human health perspective , continuing efforts to reduce ambient ozone concentrations are beneficial, despite the difficulty of achieving the NAAQS. Recent studies establish human health impacts from long-term ozone exposure over several years (Turner et al., 2016; Di et al., 2017; Berger et al., 2017). Therefore, any reduction in ozone concentrations below present levels will benefit U.S. human health, regardless of whether or not ODVs remain above 70 ppb.

Deleted:

4.2 Implications for our understanding of surface ozone concentrations

In this work we have used Equation 1 to quantify the temporal evolution of ODVs in the northeastern U.S.; this equation incorporates a constant U.S. background ODV and decreasing U.S. anthropogenic ODV enhancements, but makes no attempt to account for any other process that affects observed ODVs. Previously published studies have identified a multitude of additional processes that potentially can make time-varying contributions to ozone concentrations at U.S. surface sites, including: stratospheric intrusions, which can bring particularly high ozone concentrations to the surface (Langford et al., 2009, 2014; Lin et al., 2012a, 2015); increasing Asian anthropogenic emissions, which are believed to raise ozone concentrations over the U.S. (Jacob et al., 1999; Lin et al., 2012b); increasing frequency of wildfires, which can produce episodic ozone enhancements (McKeen et al., 2002; Jaffe, 2008, 2013; Pfister et al., 2016); variable meteorological conditions, which can lead to changes in transport patterns (Wang et al., 2016) or changes in the conditions conducive to photochemical ozone production (Shen and Mickley, 2017; Shen et al., 2017); increasing methane, which is argued to increase global ozone concentrations (Fiore et al., 2008, and references therein); and a warming climate, which has been argued may partially offset air quality improvement from regional emission controls (Fiore et al., 2015). However, there has been little in the way of systematic, quantitative analysis of the effects of these additional processes on ODVs across the U.S. Parrish et al. (2017b) show that baseline ozone concentrations transported ashore at the U.S. west coast have systematically varied over a limited range, presumably due to some of the above mentioned processes. Also, any systematic departure of average ODV trends from the purely exponential decrease incorporated in Equation 1 could contribute ODV variability not captured by our analysis. Here we approximately quantify the total influence of all these additional processes and effects by equating that influence to the ODV variance in the rural western states and the ODV variance in the northeastern states not captured by fits of Equation 1 to the ODVs.

Deleted: by

In the rural western states all ODVs reported from 35 sites over 39 years of measurements have a standard deviation of 3.7 ppb, corresponding to a variance of 13.4 ppb^2. At the individual sites and within each state the ODV records are all well described by averages with generally smaller standard deviations (Table 1). For example, Glacier NP is a single site with a 27-year measurement record that is often utilized for characterizing background ozone concentrations (see Lin et al., 2017 and references therein); the ODVs at this site have a standard deviation of only 1.4 ppb. The northeastern U.S. states contrast sharply with the rural western states, because here variation in the anthropogenic ODV enhancements dominates the much larger variance (252 ppb^2 for the entire 1975-2015 period). Fits of Equation 1 capture the large majority of this variance in this region; in Figure 9 the r^2 values for 18 years (2000-2017) indicate that Equation 1 captures more than two-thirds of the variance of the individual site ODVs, and 89% of the variance of the maximum ODVs in the eight states. The difference between these percentages is attributed to interannual variability in the spatial distribution of ODVs within the states, plus spatial variability in the ODV enhancements not accurately represented by the contour plot of Figure 8. The RMSD between observed and calculated state maxima ODVs is 3.5 ppb (variance = 12 ppb^2), which is similar to the standard deviation of 3.7 ppb (13.4 ppb^2 variance) of the average ODVs in the rural western states. The analyses in the two regions agree that the total influence of all factors affecting ODVs over the regions accounts for RMSD ≤ 3.7 ppb, or no more than ~11% of the total ODV variance over

Deleted: 251

Deleted: very

the 2000-2017 period in the northeastern states. In summary, Equation 1 is remarkably successful at capturing a large fraction of the ODV variability in the northeastern U.S. states. Guo et al. (2018) discuss a contrasting result; they suggest that monthly regional mean U.S. background MDA8 ozone concentrations vary by up to 15 ppb from year to year, and that a 3-year averaging period (as is used to define the ODV) is not long enough to eliminate interannual variability in background ozone on the days of highest observed ozone. This is not a direct comparison, but it suggests that Guo et al. (2018) overestimate the actual variability of the observed ODVs in the two northern U.S. regions examined in this work and in southern California examined by Parrish et al. (2017a).

The estimates derived in this work for the U.S. background ODV can be compared with model results. Fiore et al. (2014) compare calculations of the fourth highest MDA8 North American background (NAB) ozone (also called policy-relevant background (PRB) ozone) from two global models. The NAB concentration is that which would be present if anthropogenic emissions were reduced to zero throughout North America, not just in the U.S. NAB ozone concentrations are therefore somewhat smaller than U.S. background ozone concentrations, but for the purposes of this comparison, we can ignore this difference. The color scales in their Figures 2 and 10 allow estimates of the U.S. background ODV from the GEOS-Chem and AM3 models, respectively. Similarly, the color scale in Figure 6 of Emery et al. (2012) allows estimates of results from a different version of the GEOS-Chem model for the fourth highest MDA8 PRB. Figure 14 and Table S5 compare the model results with the observationally based estimates of U.S. background ODV derived in this work. These model results do have some skill in calculating the U.S. background ODVs. For five regions (three western rural states, the northeastern U.S. region, and the South Coast Air Basin) the model-observation correlations give r^2 values varying from 0.31 to 0.90, but the model results are on average systematically lower by 4.4 to 13 ppb. Importantly, the model results disagree with each other, as well as with the observationally based results.

Deleted: 85

Deleted: 12

4.3 Possible shortcomings of the analysis

An uncertainty in the fits of the ODV time series to the exponential decay of the ODV enhancement term in Equation 1 is the determination of the time constant, τ. The clear decrease in ODVs across the entire northeastern U.S. did not begin until about 2000; the 18-year period of consistent decreases is not long enough for fits of Equation 1 to accurately derive all three parameters. The primary approach we have taken is to use τ = 21.9 years, the value determined for southern California (Parrish et al., 2017a) in the northeastern U.S. as well. It is not clear how the time scales of reductions in U.S. anthropogenic ODV enhancements compare between California and the northeastern U.S. In California, precursor emission reductions may have been faster, because that state may have had more aggressive emission control measures, but they may also have been slower because controls on eastern coal-fired power plants dramatically reduced NOx emissions. This latter reduction would not have occurred in California where such power plants are located downwind, out-of-state. On the other hand, emission reduction rates could be roughly the same, as most northeastern U.S. states have adopted the California on-road light-duty motor vehicle emission control program, and this is a large source sector both in California and the Northeast. The alternative analysis approaches described in Section 2.3 with results discussed in Section 3.3 do not show evidence for a different exponential rate of decrease of U.S. anthropogenic ODV enhancements between southern California and the northeastern states, but uncertainty in the value of τ remains a source of uncertainty in all of the results. The y_0 and A values derived from the fits are sensitive to the selected τ value, with a larger value of τ attributing a smaller fraction of the ODV time series to y_0 and yielding a larger A value.

Finally, Equation 1 implicitly assumes that all sectors of anthropogenic U.S. ozone precursor emissions have been reduced by emission controls at approximately the same rate. However, in some respects this is a poor approximation in that some emission sectors have received lesser efforts than others. Any emissions that have not been reduced would tend to lead to an overestimate in the U.S. background ODV, since ozone produced from those emissions would not have decreased. For example, Parrish et al. (2017a) note that continuing agricultural emissions in the Salton Sea Air Basin may account for the anomalously high y_0 value derived for that region. Here, the possible influence of volatile chemical products (McDonald et al., 2018) in the northeastern U.S. is mentioned above. It is not possible to account for uncertainties in the results that may arise from this issue.

4.4 Needs for further research efforts

Accurately quantifying the U.S. background contribution to ODVs (i.e., the limit to which ODVs can be reduced through U.S. anthropogenic emission reductions alone) is important from the perspective of determining the extent of emission reductions required to attain the ozone NAAQS. In this work we have determined the value of the parameter y_0 of Equation 1 within relatively small uncertainties (estimated 95% confidence limits of ~3 ppb). These uncertainties are derived from the scatter in the observed ODVs about the fits to Equation 1. However, identifying the value of y_0 as the U.S. background ODVs brings in additional possible uncertainties (see discussion in the preceding section) that have not been quantified. Traditionally, models have been used to estimate U.S. background ozone (see Jaffe et al., 2018 and references therein), but the models utilized in these efforts have significant shortcomings (e.g., see discussion in Parrish et al., 2017a), that lead to large uncertainties in the results. Jaffe et al. (2018) estimate an uncertainty in modeled seasonal mean U.S. background ozone of about ±10 ppb, with greater uncertainty for individual days (such as those that define the ODV), and Guo et al. (2018) find biases as high as 19 ppb in modeled seasonal mean MDA8 ozone. Thus, modeling and the observational based approach discussed in this paper are both available for estimating U.S. background ODVs, but each has significant, poorly quantified uncertainties.

In summary, effective air quality management can be usefully informed by quantification of U.S. background ODVs. However, given the relatively small differences between estimated U.S. background ODVs and the 2015 ozone NAAQS of 70 ppb, these quantifications will be of more utility if they are accurate to within a couple of ppb (see Figure 11 and associated discussion). Currently, two general approaches are available for estimating U.S. background ODVs (the observational based method discussed here and in Parrish et al. (2017a), and a variety of modelling approaches), but the limited comparisons of results from these two approaches and between the different model results indicate differences much larger than ideal. However, the magnitudes of these disagreements are within the uncertainty of the model estimates as discussed by Jaffe et al. (2018) and Guo et al. (2018). Further improvement is required in modeling systems until their output can accurately reproduce the magnitude and variability of the time series of observed ODVs discussed here; these model calculations could then provide accurate determination of the U.S. background ODVs, the ODV enhancements from U.S. anthropogenic emissions, and robust interpretations of the parameters y_0 and A derived in this work. Until that model improvement is accomplished, the observationally based approach utilized in this work can provide useful estimates for air quality management guidance, as well as for comparison with evolving model calculations.

Competing interests. The authors declare that they have no conflict of interest. David Parrish is the sole proprietor of David.D.Parrish, LLC, which has had contracts funded by several state and federal agencies and private companies. One of those contracts funded some of this work.

Deleted: s
Deleted: has
Deleted: He

Acknowledgements. Some of the content of this paper was originally developed as a report submitted in fulfillment of the Technical Services Agreement between the Northeast States for Coordinated Air Use Management (NESCAUM) and David.D.Parrish, LLC funded under Agreement No. 101132 from the New York State Energy Research and Development Authority (NYSERDA). NYSERDA has not reviewed the information contained herein, and the opinions expressed in this report do not necessarily reflect those of NYSERDA or the State of New York. The authors appreciate the comments and discussion provided by Paul Miller of NESCAUM and Tom Ryerson, Fred Fehsenfeld, Owen Cooper and Andrew Langford of NOAA ESRL CSD. This paper has greatly benefited from extensive efforts of four anonymous referees and the editor. The scientific results and conclusions, as well as any views or opinions expressed herein, are those of the authors, and do not necessarily reflect the views of NESCAUM, NOAA, or the Department of Commerce.

Deleted: s

References

Aleksic, N., Ku, J.-Y., Sedefian, L.: Effects of the NOx SIP Call program on ozone levels in New York. J. Air Waste Manage. Assoc. 63, 1335–1342, 2013.

Berger, R.E., Ramaswami, R., Solomon, C.G., Drazen, J.M.: Air pollution still kills, N. Engl. J. Med., 376, 2591-2592, 2017.

Di, Q., Wang, Y., Zanobetti, A., Wang Y., Koutrakis P., Choirat C., Dominici F., Schwartz, J.D., Air Pollution and Mortality in the Medicare Population, N. Engl. J. Med., 386, 2513-2522. 2017.

Dolwick, P., Akhtar, F., Baker, K.R., Possiel, N., Simon, H., Tonnesen, G.: Comparison of background ozone estimates over the western United States based on two separate model methodologies, Atmos. Environ., 109, 282-296, doi.org/10.1016/j.atmosenv.2015.01.005, 2015.

Emery, C., Jung, J., Downey, N., Johnson, J., Jimenez, M., Yarwood, G., Morris, R.: Regional and global modeling estimates of policy relevant background ozone over the United States, Atmos. Environ., 47, 206-217, doi:10.1016/j.atmosenv.2011.11.012, 2012.

Fehsenfeld, F. C., et al.: International Consortium for Atmospheric Research on Transport and Transformation (ICARTT): North America to Europe—Overview of the 2004 summer field study, J. Geophys. Res. 2006, 111, D23S01, doi:10.1029/2006JD007829.

Fiore, A.M., West, J.J., Horowitz, L.W., Naik, V., and Schwarzkopf, M.D.: Characterizing the tropospheric ozone response to methane emission controls and the benefits to climate and air quality, J. Geophys. Res., 113, D08307, doi:10.1029/2007JD009162, 2008.

Fiore, A.M., et al.: Estimating North American background ozone in U.S. surface air with two independent global models: Variability, uncertainties, and recommendations, Atmos. Environ., 96, 284-300, doi.org/10.1016/j.atmosenv.2014.07.045, 2014.

Fiore, A.M., Naik, V., and Leibensperger, E.M.: Air Quality and Climate Connections, J. Air Waste Manage., 65, 645–685, doi:10.1080/10962247.2015.1040526, 2015.

Guo, J.J., Fiore, A.M., Murray, L.T., Jaffe, D.A., Schnell, J.L., Moore, C.T., and Milly, G.P.: Average versus high surface ozone levels over the continental USA: model bias, background influences, and interannual variability. Atmos. Chem. Phys., 18, 12123–12140, doi.org/10.5194/ac acp-18-12123-2018, 2018.

Jacob, D.J., Logan, J.A., Murti, P.P.: Effect of rising Asian emissions on surface ozone in the United States, Geophys. Res. Lett., 26, 2175–2178, doi:10.1029/1999gl900450, 1999.

Jaffe, D., Chand, D., Hafner, W., Westerling, A., and Spracklen, D.: Influence of fires on ozone concentrations in the western US, Environ. Sci. Technol., 42, 5885–5891, doi:10.1021/es800084k, 2008.

Jaffe, D., Wigder, N., Downey, N., Pfister, G., Boynard, A., and Reid, S. B.: Impact of Wildfires on Ozone Exceptional Events in the Western US, Environ. Sci. Technol., 47, 11065–11072, doi:10.1021/es402164f, 2013.

Jaffe, D.A., Cooper, O.R., Fiore, A.M., Henderson, B.H., Tonneson, G.S., Russell, A.G., Henze, D.K., Langford, A.O., Lin, M., Moore, T.: Scientific assessment of background ozone over the U.S.: Implications for air quality management. Elem. Sci. Anth. 6 56 doi.org/10.1525/elementa.309, 2018.

Langford, A.O., Aikin, K.C., Eubank, C.S., Williams, E.J.: Stratospheric contribution to high surface ozone in Colorado during springtime, Geophys. Res. Lett., 36, L12801, doi:10.1029/2009gl038367, 2009.

Langford, A.O., Senff, C., Alvarez II, R., Brioude, J., Cooper, O., Holloway, J., Lin, M., Marchbanks, R., Pierce, R., Sandberg, S., Weickmann, A., Williams, E.: An overview of the 2013 Las Vegas Ozone Study (LVOS): Impact of stratospheric intrusions and long-range transport on surface air quality, Atmos. Environ., 109, 305–322, doi:10.1016/j.atmosenv.2014.08.040, 2014.

Lin, C.-Y. C., Jacob, D. J., Fiore, A. M.: Trends in violations of the ozone air quality standard in the continental United States, 1980-1998, Atmos. Environ., 35, 3217–3228, 2000.

Lin, M., Fiore, A.M., Cooper, O.R., Horowitz, L.W., Langford, A.O., Levy, H., Johnson, B.J., Naik, V., Oltmans, S.J., Senff, C.J.: Springtime high surface ozone events over the western United States: Quantifying the role of stratospheric intrusions, J. Geophys. Res., 117, D00V22, doi:10.1029/2012jd018151, 2012a.

Lin, M., Fiore, A.M., Horowitz, L.W., Cooper, O.R., Naik, V., Holloway, J., Johnson, B.J., Middlebrook, A.M., Oltmans, S.J., Pollack, I.B., Ryerson, T.B., Warner, J.X., Wiedinmyer, C., Wilson, J., Wyman, B.: Transport of Asian ozone pollution into surface air over the western United States in spring, J. Geophys. Res., 117, D00V07, doi:10.1029/2011jd016961, 2012b.

Lin, M.; Fiore, A.M., Horowitz, L.W., Langford, A.O., Oltmans, S.J., Tarasick, D., Rieder, H.E.: Climate variability modulates western US ozone air quality in spring via deep stratospheric intrusions, Nat. Commun., 6, 7105, doi:10.1038/ncomms8105, 2015.

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

McDonald, B.C., et al.: Volatile chemical products emerging as largest petrochemical source of urban organic emissions, Science, 359, 760-764, 2018.

McKeen, S.A., Wotawa, G., Parrish, D.D., Holloway, J.S., Buhr, M.P., Hübler, G., Fehsenfeld, F.C. Meagher, J.F.: Ozone production from Canadian wildfires during June and July of 1995, J. Geophys. Res., 107 (D14), doi:10.1029/2001JD000697, 2002.

National Research Council (NRC), 1991. Rethinking the Ozone Problem in Urban and Regional Air Pollution. National Academy Press, Washington, DC.

Oltmans, S.; Schnell, R.; Johnson, B.; Pétron, G.; Mefford, T.; Neely III, R.: Anatomy of wintertime ozone associated with oil and natural gas extraction activity in Wyoming and Utah, Elem. Sci. Anth., 2, 24, doi.org/10.12952/journal.elementa.000024, 2014.

Parrish, D.D., Fahey, D.W., Williams, E.J., Liu, S.C., Trainer, M., Murphy, P.C., Albritton, D.L., Fehsenfeld, F.C.: Background Ozone and Anthropogenic Ozone Enhancement at Niwot Ridge, Colorado, J. Atmos. Chem., 4, 63-80, 1986.

Parrish, D.D., and Stockwell, W.R.: Urbanization and air pollution: Then and now, Eos: Earth & Space Science News, 96, 10-15, 2015.

Parrish, D.D., Young, L.M., Newman, M.H., Aikin, K.C., and Ryerson; T.B.: Ozone Design Values in Southern California's Air Basins: Temporal Evolution and U.S. Background Contribution, J. Geophys. Res. Atmos., 122, 11,166–11,182, doi.org/10.1002/2016JD026329, 2017a.

Parrish, D.D., Petropavlovskikh, I., and Oltmans, S.J.: Reversal of long-term trend in baseline ozone concentrations at the North American West Coast, Geophys. Res. Lttrs., 44, 10,675–10,681, doi.org/10.1002/2017GL074960, 2017b.

Pfister, G.G., Wiedinmyer, C., and Emmons, L.K.: Impacts of the fall 2007 California wildfires on surface ozone: Integrating local observations with global model simulations, Geophys. Res. Lttrs., 35, L19814, doi:10.1029/2008GL034747, 2008.

Shen, L., and Mickley, L.J.: Effects of El Niño on summertime ozone air quality in the eastern United States, Geophys. Res. Lttrs., 44, 12,543-12,550, doi.org/10.1002/2017GL076150, 2017.

Shen, L., Mickley, L.J., Leibensperger, E.M., and Li, M.: Strong dependence of U.S. summertime air quality on the decadal variability of Atlantic sea surface temperatures, Geophys. Res. Lttrs., 44, 12,527–12,535, doi.org/10.1002/2017GL075905, 2017.

Turner, M.C., Jerrett, M., Pope III, C.A., Krewski, D., Gapstur, S.M., Diver, W.R., Beckerman, B.S., Marshall, J.D., Su, J., Crouse, D.L., Burnett, R.T.: Long-Term Ozone Exposure and Mortality in a Large Prospective Study, Am. J. Respir. Crit. Care Med., 193(10), 1134-1142, 2016.

Wang, Y., Jia, B., Wang, S.-C., Estes, M., Shen, L., and Xie, Y.: Influence of the Bermuda High on interannual variability of summertime ozone in the Houston–Galveston–Brazoria region, Atmos. Chem. Phys., 16, 15265-15276, https://doi.org/10.5194/acp-16-15265-2016, (2016).

Wilcox, G.L.: New England and the Challenge of Interstate Ozone Pollution Under the Clean Air Act of 1990. B.C. Envtl. Aff. L. Rev., 24, 1-101, http://lawdigitalcommons.bc.edu/ealr/vol24/iss1/2, 1996.

Wolff, G. T., Lioy, P. J.: Development of an ozone river associated with synoptic scale episodes in the eastern United States. Environ. Sci. Technol., 14, 1257–1260, 1980.

Table 1. ODV statistics from the rural western states.

State/Site	Avg. ± Std. Dev. (ppb)	Variance (ppb^2)	years of ODV record
Montana	55.4 ± 2.2	4.8	1979-2017
Glacier NP	54.5 ± 1.3	2.0	1991-2017
Yellowstone NP	64.0 ± 2.1	4.4	1999-2017
North Dakota	59.3 ± 2.7	7.3	1982-2017
South Dakota	61.5 ± 3.8	14.6	1990-2017

Table 2. Results of least-squares fits to Equation 1 illustrated in Figures 5-7 and S3-S10; RMSD indicates the root-mean-square deviation between the observed ODVs and the derived fit.

State/sites	y_0 (ppb)	A (ppb)	RMSD (ppb)	A^* (ppb)	years fit
New York/maximum O_3	53 ± 6	43 ± 9	3.9	53 ± 2	2000-2017
New York/rural upwind	42 ± 7	50 ± 10	5.1	44 ± 2	2000-2017
New Jersey/all sites	43 ± 4	57 ± 6	4.6	54 ± 2	2000-2017
Connecticut/all sites	56 ± 5	40 ± 7	5.0	55 ± 2	2000-2017
Connecticut/coastal	61 ± 6	36 ± 8	4.1	57 ± 3	2000-2017
Rhode Island/all sites	49 ± 8	44 ± 12	4.0	49 ± 3	2000-2017
Massachusetts/Boston	46 ± 6	27 ± 6	3.1	27 ± 2	1990-2017
Massachusetts/suburban	41 ± 10	52 ± 14	3.3	45 ± 3	2000-2017
Massachusetts/coastal	44 ± 9	52 ± 13	3.2	49 ± 3	2000-2017
New Hampshire/coastal	49 ± 6	35 ± 8	3.7	38 ± 2	1995-2017
New Hampshire/northwest	45 ± 6	29 ± 9	3.7	28 ± 2	2000-2017
New Hampshire/Mt. Washington	66 ± 7	8 ± 8	2.9	---	1993-2017
Vermont /all sites	46 ± 7	34 ± 10	2.7	33 ± 2	2000-2017
Maine/interior	44 ± 8	23 ± 10	5.8	21 ± 3	1990-2017
Maine/NE coast	47 ± 5	22 ± 5	2.0	23 ± 2	1991-2017
Maine/SW coast	49 ± 5	36 ± 5	4.1	39 ± 2	1990-2017
Maine/Cadillac Mtn.	52 ± 16	36 ± 20	5.2	44 ± 5	1997-2017

Table 3. Results of least-squares fits of Equation 1 to the state maximum ODVs illustrated Figure 10; y_0 and τ were held constant at 45.8 ppb and 21.9 years, respectively. The absolute root-mean-square deviations between the observed ODVs and the derived fits are indicated. $Year_{NAAQS}$ indicates the projected year that the fit to the state maximum ODV drops to the NAAQS of 70 ppb.

State	A^* (ppb)	RMSD (ppb)	$Year_{NAAQS}$
Connecticut	61 ± 7	5.8	2021
Maine	48 ± 4	3.2	2015
Massachusetts	53 ± 5	3.9	2017
New Hampshire	43 ± 4	3.0	2013
New Jersey	64 ± 5	3.7	2021
New York	58 ± 4	3.0	2019
Rhode Island	52 ± 4	3.4	2017
Vermont	35 ± 3	2.1	2008

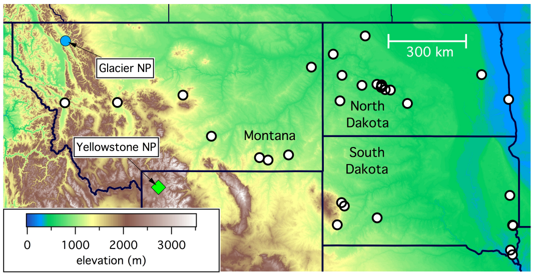


Figure 1: Topographical map of the three rural western states with symbols indicating the locations of the monitoring sites. The two colored symbols indicate two long-term sites in national parks that are discussed in detail. Note that Yellowstone National Park is located in Wyoming, but is nevertheless considered here.

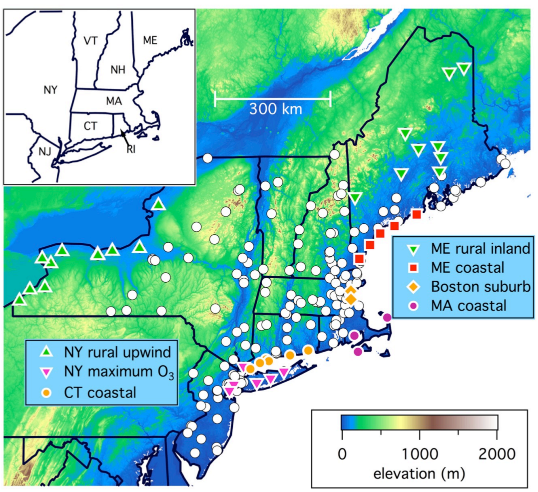


Figure 2: Topographical map of the eight northeastern states with symbols indicating the locations of the ozone monitoring sites. Seven groups of colored symbols indicate groups of sites that are discussed in detail. The inset gives the abbreviations for each of the eight states.

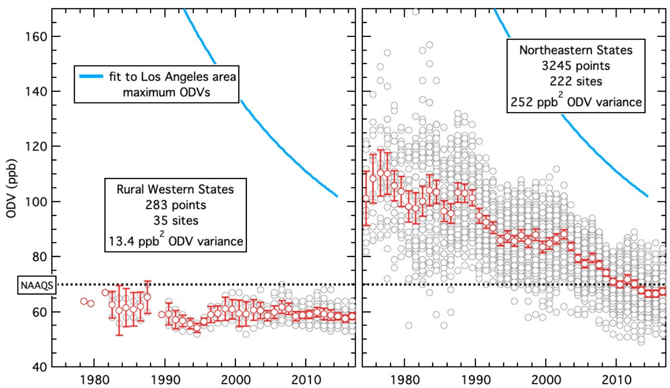


Figure 3: Time series of all ODVs (grey symbols) reported from all monitoring sites in the two northern U.S. regions shown in Figures 1 and 2. The numbers of monitoring sites and reported ODVs and the ODV variance are annotated for each region. The red symbols give the averages and 2-σ confidence limits for all ODVs reported in each year. For comparison, the blue curve in each panel indicates a fit to the maximum ODVs recorded in the Los Angeles urban area (Parrish et al., 2017a). The dotted line indicates the 2015 NAAQS of 70 ppb.

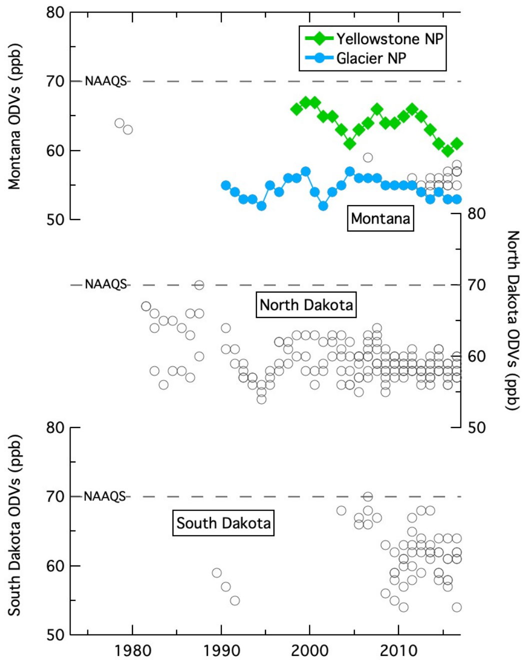


Figure 4: Time series of all ODVs (grey symbols) reported from all monitoring sites in three rural western states, plus Yellowstone NP located in Wyoming. The two sets of colored symbols are results from two long-term sites in national parks.

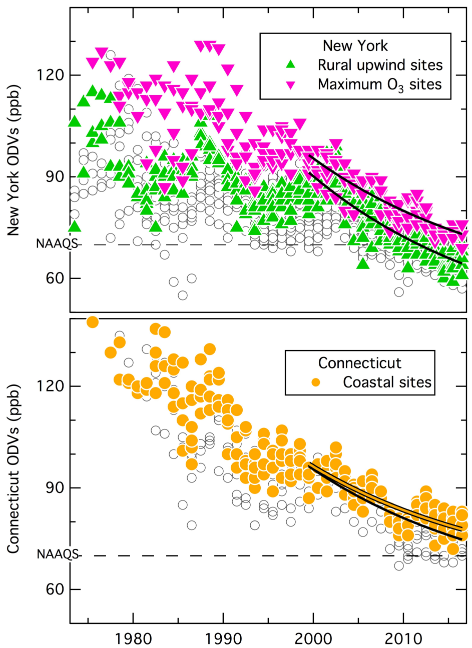


Figure 5: Time series of all ODVs (grey symbols) reported from all monitoring sites in New York and Connecticut. The three sets of colored symbols indicate the results from groups of sites that are discussed in detail. The curves are fits of Equation 1 to respective colored symbols, and to all data points for Connecticut.

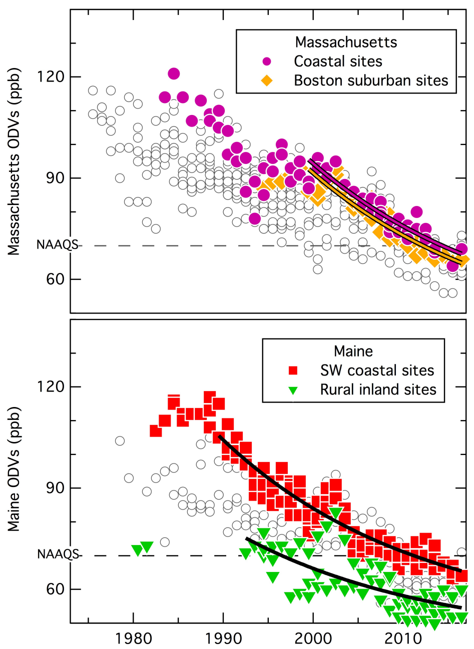


Figure 6: Time series of all ODVs (grey symbols) reported from all monitoring sites in Massachusetts and Maine. The four sets of colored symbols indicate the results from groups of sites that are discussed in detail. The curves are fits of Equation 1 to respective colored symbols.

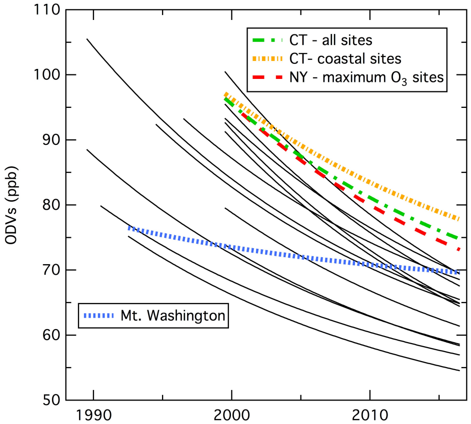


Figure 7: Comparison of fits of the ODVs from 17 groups of sites in 8 northeastern U.S. states shown in Figures 5, 6 and S3-S10 to Equation 1. The parameters of these fits are included in Table 2.

5

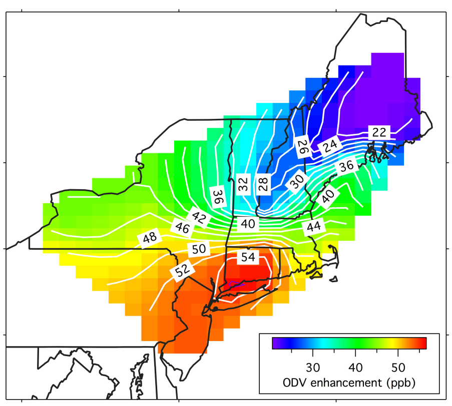


Figure 8: Approximate contour plot of the U.S. anthropogenic ODV enhancement due to photochemical production from precursor emissions in the year 2000, estimated from the *A values given in Table 2.**

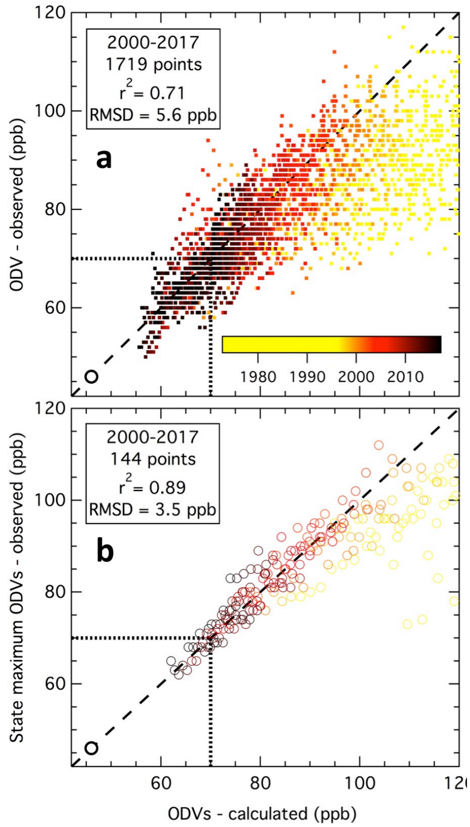


Figure 9. Comparison of observed ODVs color-coded by year with those calculated from Equation 1 for a) all monitoring sites and b) for the maximum observed in each state. The dashed lines indicate the 1:1 relationships with y_0 near the origin indicated by the larger circle. The dotted lines indicate the NAAQS. The number of data, square of the correlation coefficient, and the root-mean-square difference between the observed and calculated ODVs for 2000-2017 are annotated.

Deleted: ,

Deleted: and t

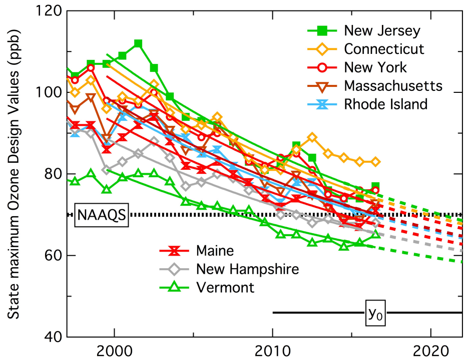


Figure 10: Time series of maximum ODVs reported from any site within each of the eight northeastern states. The solid curves are fits of Equation 1 to the respective colored symbols for the 2000-2017 period. The derived *A values from these latter fits are given in Table 3. The dashed lines are projections of the solid curves.**

Deleted: 2010

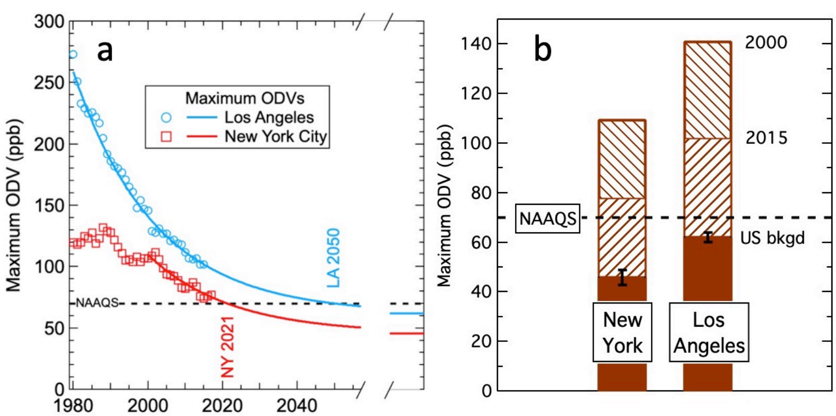


Figure 11. Comparison of maximum observed ODVs in the New York City and Los Angeles urban areas. a) Temporal trend of observations (symbols) and fit to Equation 1, including extrapolation to infinite time; annotations indicate year that extrapolations decrease to 70 ppb. The New York City results are the maxima from either the states of New York or New Jersey, and the Los Angeles results are those for the South Coast Air Basin (Figure 8 of Parrish et al., 2017a). b) Bar graph indicating maximum ODVs in 2000 and 2015 (hatched bars) and the estimated U.S. background ODV (solid bars); the maximum ODVs are derived from the fits to Equation 1 included in a).

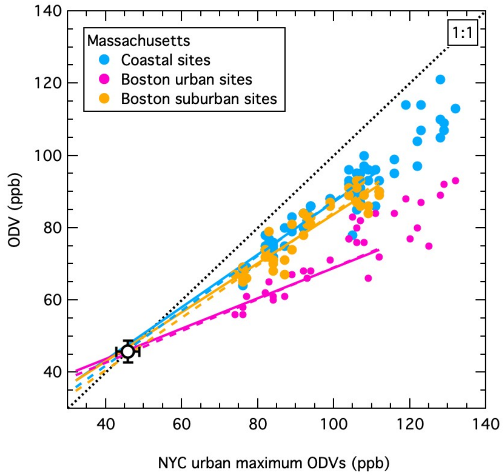


Figure 12. Correlation between the ODVs from three sets of Massachusetts sites and the maximum ODVs recorded in the New York city urban area. Lines of corresponding color show ordinary linear regression (solid) and reduced major axis regression with equal weighting (dashed) fits of the correlated data sets for the ODVs recorded in 2000-2017. The black symbol shows the mean U.S. background ODV derived from the exponential fits to the ODV time series and the dotted line indicates the 1:1 relationship.

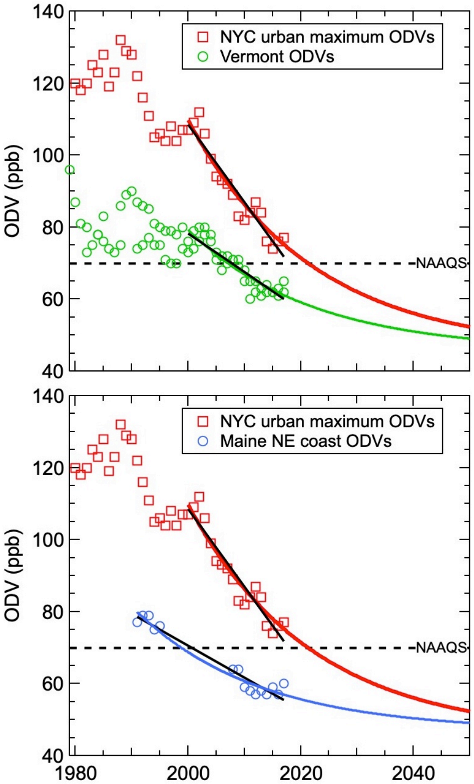
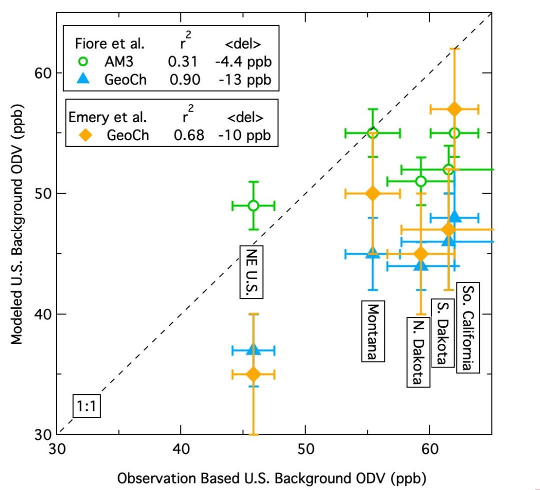


Figure 13: Comparison of linear and exponential fits to three ODV data sets. The black line segments are linear fits to 2000-2017 for two, and 1991-2017 for one. The colored curves are the exponential fits to these time series, also shown in Figures 11, S9 and S10.



Deleted:

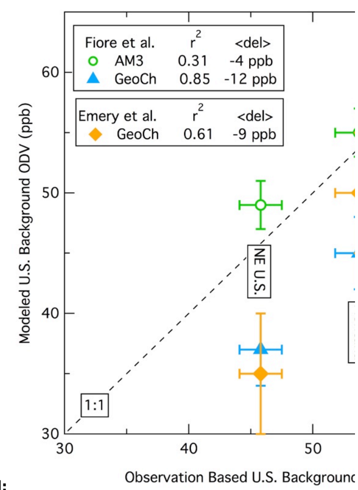


Figure 14: Comparison of U.S. Background ODV estimates from model calculations with those derived in this work from observations. The r^2 values of the correlations and the average differences (<del>) are annotated.