Imputing Air Quality Estimates for Alternative Meteorological Years Using Relative Response Values from a Single Meteorological Year

George Bowker and Timothy Sharac

US Environmental Protection Agency, Office of Air & Radiation, Clean Air Markets Division

Air quality modeling projections often rely on determining the relative change in ozone concentrations resulting from differences in emissions between a base year and a future year while holding meteorology constant. Here, we describe an interpolation method to impute a modeled response for alternative years with different meteorology derived from the modeled estimates (base year and future year) based on meteorology from a single year. For each day in the alternative year, we match a day from the base year data set based on similarity of ozone concentrations. This matching occurs after the base year data set is directly scaled to the alternative year using linear interpolation between the modeled base and future years. Then, we assume that the matched day in the alternative year will respond in a similar way to its comparable base year day. We can bootstrap estimates from the base year modeling data for each day across many different years, including policy-relevant days such as the 4th highest annual value. We demonstrate the uncertainty of this method using estimates derived from 2011 and 2012 base year modeling with a 2023 future year (with 2011 and 2012 meteorology).

Introduction

Projecting air quality to the future is critically important for assessment and design of environmental regulation. It is complicated by day to day and year to year changes in emissions and meteorology (e.g., surface temperature, wind speed, relative humidity, and precipitation) that can affect ozone formation (Camalier *et al.*, 2007). Over time, ozone precursor emissions have decreased within the contiguous United States, as emission controls and updated technology have reduced emissions across point, area, and mobile sources. The response in ozone concentrations to these emission reductions is not linear, with the result dependent on the starting ozone concentrations. Projecting air quality concentrations to the future relies on regional chemical transport air quality models. Air quality modeling is performed for the base year and for a future year. The future year modeling relies on adjustments of base year emission inventories while holding meteorology constant. The change estimated by the air quality modeling is applied to air quality monitoring data centered around the base year (Foley *et al.*, 2015).

The monitoring data in and around the base year are affected by the particular meteorological patterns experienced in those years. Consequently, atypical concentrations resulting from unusual patterns in meteorology that would persist after application of the modeled change.

The approach here utilizes the modeled response from meteorologically different days within a single base year simulation to impute responses for similar meteorological days in other years, to estimate how those other years would respond. Consequently, in calculating an air quality estimate for the future, this approach allows many more years to be averaged together, which provides a more-stable estimate of the future air quality and provides estimates of the uncertainty.

The goal of this poster is twofold: first, to utilize modeled changes in daily concentrations from two separate meteorological base years to create two estimates of future air quality using monitoring data from a wide range of years; and second, to compare these estimates to a more-typical approach for projecting design values based on a set of 5-years of monitoring data centered around a single year.

Methods

The methodology depends on two data sets; the first required data set are modeled base year and future year monitor-level ozone concentrations using different emissions where the meteorology is held constant and the second data set needed uses multiple years of air quality monitoring data. We have the modeled response in concentration for each day as it is projected to the future as a function of reduced ozone precursor emissions. The daily concentration in the future divided by the ozone concentration in the base year is that day's daily relative response factor (dRRF) (Figure 1). Once an array of different days has been modeled, given a particular concentration in the base year, its dRRF is relatively predictable.

Next, to transform the daily base data to the future, the dRRFs are applied to the monitoring data. While there are several alternatives to identify an appropriate dRRF for each day, we elected to match the base monitored data with the "closest" modeled day in the base using concentration. The base monitoring data are multiplied by the matched dRRF to complete the projection to the future.

To extend the modeling data to other years that were not explicitly modeled, we linearly interpolate each day's dRRF value between the base and future years to estimate how that day could have responded in other years assuming that emissions change linearly from one year to the next (Figure 2). We also assume that the ozone response is linear over the range of emission changes between the base and future years. These assumptions could be refined by using additional emission inventories and model simulations.

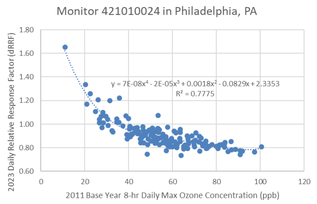


Figure 1. Relationships between the 2011 base year modeled daily ozone concentrations and the daily Relative Response Factor for the 2023 future year.

Camalier, L., W. Cox, and P. Dolwick (2007), The effects of meteorology on ozone in urban areas and their use in assessing ozone trends, Atmos. Environ., 41, 7127–7137, doi:10.1016/j.atmosenv.2007.04.061.

Foley, K.M., P. Dolwick, C. Hogrefe, H. Simon, B. Timin, N. Possiel (2015), Dynamic evaluation of CMAQ part II: Evaluation of relative response factor metrics for ozone attainment demonstrations, Atmos. Environ., 103, 199-195, doi.org/10.1016/j.atmosenv.2014.12.039.

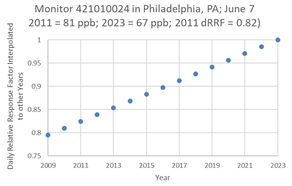
Wayland, R.A., (2018) Modeling Guidance for Demonstrating Air Quality Goals for Ozone, PM2.s and Regional Haze. Nov. 29, 2018. U.S. Environmental Protection Agency, Memorandum to Regional Air Division Directors, Regions 1-10. https://www.epa.gov/scram/state-implementation-plan-sip-attainment-demonstration-guidance.

Disclaimer: The views expressed in this poster are those of the authors and do not necessarily represent the views or policies of the U.S. Environmental Protection Agency.

The result is a series of estimated dRRFs and ozone concentrations for each site and modeled day that are an interpolation between the modeled base and future year (Figure 3). Alternate year model projections between the base and future year can likely be adequately represented provided that the two following assumptions are met: one, base year has a sufficiently wide range of ozone concentrations representing various types of meteorology and emissions patterns, and two, there is a reasonable amount of emission change between the base and future year.

Finally, using the technique we developed to match each of the daily monitored concentration values in the base year to the closest daily modeled concentration in the base year, we match each daily monitored concentration to its closest modeled day from the interpolated year. In this way, we identify and apply a year-specific dRRF value for each monitored concentration, thereby projecting that monitored day to the future.

This process was repeated for each monitored daily value from 2009 through 2018, using the 2011 base year (and extrapolations of other years from 2009-2018) and 2023 future year air quality modeling. The result was a series of estimated daily concentrations that have all been projected to a common future year of 2023. We repeated the technique using a 2012 base year for emissions and meteorology that was also projected to 2023. We then compared these two independent sets of estimates, and then compared the estimates with estimates based on the current design value projection technique.



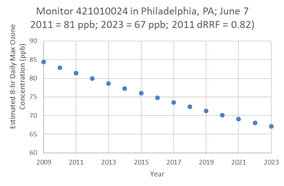


Figure 2. Relationship between the 2011 base year dRRF and scaled dRRF values for other future years.

Figure 3. Relationship between the 2011 base year daily ozone concentration and interpolated values for other year.

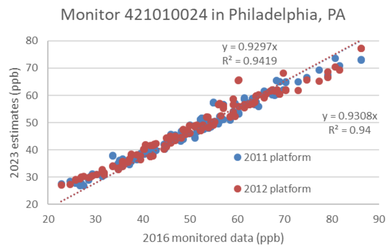


Figure 4. Comparison between 2023 projections of 2016 8-hr daily maximum ozone monitoring data using the 2011 and 2012 base year modeling platforms.

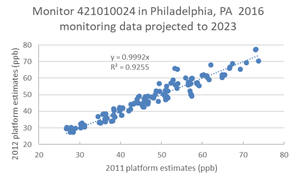


Figure 5. Direct comparison between 2023 projections of 2016 monitoring data using the 2011 and 2012 base year modeling platforms.

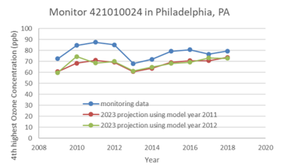


Figure 6. 4th highest monitoring concentrations vary considerably from year to year in the monitoring and projected data.

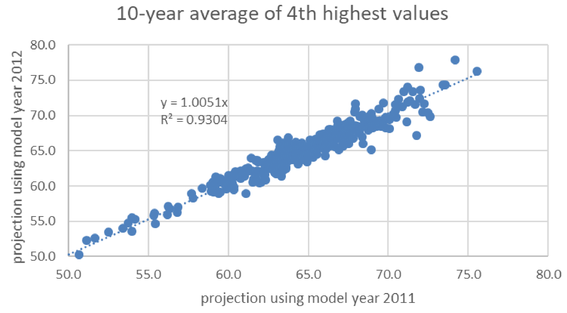


Figure 7. Examined across 314 monitoring locations, the average 4th highest monitoring concentrations from 2009-2018 using the 2011 and 2012 modeling platforms are highly correlated.

Results and Discussion

For each location and day within a particular year, when we use the modeling data in a relative way and apply the change to the monitoring data, the results are consistent. Figure 4 shows the 2023 estimates relative to the 2016 based monitoring data for a site in Philadelphia, PA. Directly comparing the 2011 with 2012 based modeling projections for this site the estimates are highly correlated, with differences between these two projections averaging 0.1 ppb, a standard deviation of 3.1 ppb, and an average absolute difference of 2.4 ppb (Figure 5).

Focusing on the policy-relevant 4th highest values over time at any individual site, the actual monitoring data shows substantial variation from one year to the next and a general downward trend over time as emissions have decreased. When the monitoring data are projected to the future, the estimated 2023-based projections from each year are lower and are detrended, though the year-to-year patterns of variation persist. Since there are now ten separate estimates of the concentration in 2023 derived from 2009-2018 monitoring data, a "typical" value could be identified from these data, such as the average or the median of the ten 4th highest values. Figure 7 shows the average 4th highest value in 2023 extrapolated from the 2011 and 2012 platforms using monitoring data from 2009-2018 for 314 monitoring sites. Despite being derived from completely different modeled meteorology, the values are relatively consistent, with a slope around 1 and a coefficient of regression value of 0.93. The average difference was 0.34 ppb, with a standard deviation of 1.2 ppb. This stability of "typical" concentration values contrasts with the high year-to-year variation in the values. The standard deviation of the 4th highest projected values from one year to the next was 4.4 ppb, on average, for 314 monitoring sites.

In recent air quality modeling efforts, 3-year and 5-year weighted design values are projected based on the 5 years of monitoring data centered around the base year. Each 3-year design value is projected to the future using the average of the dRRFs from the top 10 modeled days in the base year modeling (Wayland, 2018). The three design values are then averaged to give a 5-year weighted "average" design value. We approximated this methodology, projecting the 5-year weighted average values using the 5 years of monitoring data centered around each year and the 10 dRRF values from the highest days in that same year using the scaled data from the 2011 and 2012 modeling platforms. The results from this method can be compared to the 10-year average values calculated using the method developed here and to a 5-year weighted average of the 4th highest values calculated using the method developed here. Comparing the sets of results, we observe that the results are similar between the 10-day average RRF method and the 5-year weighted average value, and that both vary around the long-term average value (Figure 8).

For both the 5-year weighted methods, the values are quite dependent on the particular patterns of monitoring data. When the base year monitoring values are higher than "typical", the resulting projected values may also be elevated. Consequently, when attempting to identify expected typical concentrations in the future, care must be taken to select representative monitoring data in the base time period. The method developed here offers the opportunity to incorporate substantially more monitoring data in estimating the expected "typical" value than other methods. This, in turn, results in more-stable air quality estimates that are more-independent of the base year monitoring data and meteorology used.

This technique could be improved by applying such factors as variation in the emission inventories between 2009 and 2018, such as by accounting for differences in emissions between types of days (weekend vs. weekday), between sectors, and between states. These additional factors could lead to improvements to the method developed here could lead to better estimates of the extrapolated modeled daily values, probably resulting in improved estimates of the future year. This, in turn, could lead to more-robust estimates of the effects of potential rules, by accommodating the natural variation in air quality resulting from variation in meteorology.

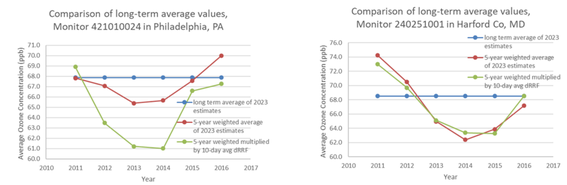


Figure 8. 4th highest monitoring concentrations vary considerably from year to year (blue line). The patterns of variability in the monitoring data persist in the projected data, even though the concentrations are lower (red and green lines).

Take-home Messages

- We were able to successfully create 10-year average projections of the 4th highest 8-hour daily maximum ozone concentrations based on monitoring data from 2009-2018.
- While the estimates use the model-predicted change in concentration as a result of emissions changes from a single year, the 10-year average projected concentrations are not dependent on the base year meteorology and pattern in the monitoring data used in the air quality modeling.