State of Delaware
Final Report: Ozone Observations and Forecasts in 2018

A Report Prepared for the Delaware Department of Natural Resources and Environmental Control

Dr. Amy K. Huff
William F. Ryan
Department of Meteorology and Atmospheric Science
The Pennsylvania State University

February 28, 2019

Table of Contents

List of Tables

List of Figures

Executive Summary

- 2018 was the third season using the latest 8-hour average ozone (O_3) exceedance threshold of 71 ppbv.
- In 2018, there were 8 days with observed 8-hour average $O_3 \geq 71$ ppbv, no days with observed 8-hour average $O_3 \geq 86$ ppbv, and no days with observed 8-hour average $O_3 \geq 106$ ppbv. This translated into 8 Code Orange days (71-85 ppbv) on the air quality index (AQI) scale.
- The downward trend since 2013 in the number of O_3 exceedance days in Delaware continued in 2018. The 8 observed O_3 exceedance days (defined as observed 8-hour average $O_3 \geq 71$ ppbv) in 2018 tied the 8 exceedance days observed in 2017, 2015, and 2014 and was essentially equal to the 2013-2017 average of 8.2 exceedance days per year.
- Summer 2018 was slightly warmer than normal in Delaware but about average in terms of precipitation; much higher than normal precipitation at neighboring upwind locations may have helped to offset the slightly higher than average summer temperatures, resulting in an average O_3 season in Delaware.
- The 2018 seasonal (May 1 to September 30) mean and median of the maximum observed 8-hour average O_3 (49 ppbv) are equivalent to the values from 2017, which were the lowest observed in Delaware going back to 1997.
- The period 2013-2018, termed the New Normal, is characterized by the decrease in observed O_3 exceedance days and average O_3 relative to the Post-NO_x SIP period of 2003-2012. These consistent decreases are attributed to continuing reductions in regional NO_x emissions, particularly from the upwind Ohio River Valley source region. This decrease in O_3 is not unique to Delaware; the same decrease is evident across the entire Mid-Atlantic region.
- Key trends that have emerged during the New Normal period continued in 2018, including the growing impact of smoke transported from wildfires in Mexico, Canada, and the western U.S.; the breakdown in the relationship between hot weather and O_3 exceedance days; regular occurrence of isolated (only one monitor location) and single day O_3 exceedances; and a shift in climatology that clusters most of the O_3 exceedance days early in the summer season with a late season surge in September.
- Expert forecasts of O_3 exceedance days in 2018 were skillful, similar to recent years.
    - Hit rate in 2018 (0.63) improved relative to 2017 (0.50) and 2016 (0.55), which is attributable to the increased frequency of multi-day high O_3 events in 2018. All 5 of the correctly identified O_3 exceedances in 2018 fell during regional multi-day O_3 episodes, which are easier to predict than single day events.
    - False alarm ratio in 2018 (0.58) deteriorated relative to 2017 (0.43) and 2016 (0.33) due to 4 false alarm forecasts issued for days during regional multi-day high O_3 events; although incorrect in Delaware, exceedances were observed on these days in neighboring metropolitan areas.
- Overall expert forecast skill (all days) for 2018 was comparable to recent years, with a median absolute error of 5.0 ppbv and an over-forecasting bias of 1.9 ppbv.
- Similar to recent years, numerical air quality models struggled to predict O_3 exceedance days in 2018; the hit rate of expert forecasts (0.63) was higher than that of the models (0.38-0.50).
- Also similar to recent years, the skill of the individual air quality models was not consistent with the previous year, making the models unreliable forecast predictors.

- Since forecasters can no longer depend on hot weather, O_3 persistence, climatology, or numerical air quality models as reliable forecast predictors, a return to the consistently high hit rates (> 0.60) of the pre-2013 period is unlikely. Nevertheless, the general trend of higher expert hit rates and lower false alarm ratios, which has emerged over the past four seasons, is expected to continue.

1. Ozone Observations in 2018

Summer 2018 was the third season under the 8-hour average ozone (O_3) exceedance threshold of **71 ppbv**, set by the 2015 daily O_3 National Ambient Air Quality Standard (NAAQS). Figure 1 compares the 2018 O_3 season (May 1 to September 30) to previous years in the State of Delaware forecast area, which includes Kent, New Castle, and Sussex Counties. In 2018, there were **8 days** with observed 8-hour average $O_3 \geq 71$ ppbv, no days with observed 8-hour average $O_3 \geq 86$ ppbv, and no days with observed 8-hour average $O_3 \geq 106$ ppbv. This translated into 8 Code Orange days (71-85 ppbv) on the air quality index (AQI) scale.

The 8 observed O_3 exceedance days in Delaware in 2018 tied the 8 exceedance days observed in 2017, 2015, and 2014 and was essentially equal to the 2013-2017 average of 8.2 exceedance days per year. Summer 2018 (June, July, and August) was slightly warmer than normal in Delaware (Figure 2) but about average in terms of precipitation (Figure 3). Precipitation was much higher than normal at neighboring upwind locations in the Philadelphia, PA and Baltimore, MD metropolitan areas, however, which may have helped to offset the slightly higher than average summer temperatures, resulting in an average O_3 season in Delaware.

Figure 1 also illustrates that the recent drop in the number of O_3 exceedance days (defined as maximum observed 8-hour average $O_3 \geq 71$ ppbv) in Delaware, which began in 2013, continued in 2018. This period, termed the "New Normal," is attributable to continued reductions in regional emissions of nitrogen oxides (NO_x) relative to the Post-NO_x SIP period (2003-2012), particularly from the upwind Ohio River Valley source region. These changes were described in detail in the 2016 O_3 season final report for Delaware. The observed O_3 decreases that have occurred during the New Normal period are not unique to Delaware; the same trend is evident across the entire Mid-Atlantic region.

The seasonal (May 1 to September 30) mean and median of maximum observed 8-hour average O_3 (Figure 4) have the same downward trend as the O_3 exceedance days in Delaware. The 2018 seasonal mean and median (49 ppbv) are equivalent to the values from 2017, which were the lowest observed in Delaware going back to 1997. Thus, not only have the highest O_3 values dropped during the New Normal period, but the average O_3 values have as well.

The overall changes in observed O_3 during the New Normal period relative to the Post-NO_x SIP period are encapsulated in Figure 5. The polynomial best fit line of maximum observed 8-hour average O_3 in Delaware during the Post-NO_x SIP period (black line) demarcates the historical seasonal O_3 cycle, which was characterized by an increase in May, a peak in June and July, a downward trend in August, and a drop off in September. Figure 5 shows that the average day-to-day variations of maximum observed 8-hour average O_3 in Delaware during the New Normal period (purple dots/line) fall almost completely below the polynomial best fit line for the Post-NO_x SIP period, reflecting the decreasing trend in O_3. The exceptions are in early May and September, when the peak of observed O_3 in the New Normal period is slightly higher than the values on the best fit line for the Post-NO_x SIP period. These changes are due to a temporal shift in the climatology of the peak of the O_3 season, which is discussed in more detail in Section 2.4. The polynomial best fit trend line for the New Normal period (red dashed line) also demonstrates how the seasonal O_3 climatology has flattened, making high O_3 equally likely from approximately mid-May through late August.

2. Continuation of Key Trends in 2018

Details about the 8 O_3 exceedance days in Delaware during 2018, listed in Table 1, highlight the recent trends in observed O_3 that have emerged during the New Normal period. These trends continued in 2018 and include the growing impact of smoke transported from wildfires in Mexico, Canada, and the western U.S (Section 2.1).; the breakdown in the relationship between hot weather and O_3 exceedance days (Section 2.2); regular occurrence of isolated (only one monitor location) and single day O_3 exceedances (Section 2.3); and a shift in climatology that clusters most of the O_3 exceedance days early in the summer season, with a late season surge in September (Section 2.4).

In Table 1, grey shading indicates the 5 O_3 exceedance days that were accurately forecasted in 2018, with alerts issued to the public. One change in reported data for 2018 is that maximum temperature (T_{max}) is taken from Wilmington-New Castle County Airport (KILG), not from Dover Air Force Base (DOV) as in years past. Daily weather observations from DOV are missing for May 1-31 and September 1-30, 2018, necessitating substitution with data from KILG for the entire O_3 season.

2.1 Wildfire Smoke

Smoke transported from wildland fires continued to be an important factor for local O_3 production in Delaware in 2018. Smoke contains high concentrations of NO_x and reactive volatile organic compounds (VOCs), which are precursors for O_3 production. Most of the Mid-Atlantic region, including Delaware, is in a "NO_x-limited" regime during summertime. Consequently, formation of O_3 is enhanced by increases in ambient NO_x. The presence of additional NO_x from dilute wildfire smoke can amplify local O_3 production substantially. Because regional anthropogenic NO_x emissions have been significantly reduced during the New Normal period, the most widespread and highest magnitude O_3 exceedances now occur more often on days when dilute transported smoke is present.

In 2018, the most striking example of the O_3-generating power of transported smoke occurred on May 1 (Figure 6), which was the first day of the 2018 O_3 season. An O_3-conducive Bermuda High pressure pattern was in place (Figure 7), with surface high pressure centered over North Carolina and the Atlantic coast of the Carolinas, promoting sunny skies, subsidence, and light westerly surface winds across Delaware. T_{max} at KILG only reached 80 °F on May 1, however, which is not typically hot enough to support a widespread O_3 exceedance. Nevertheless, Delaware was the hotspot of O_3 production in the Mid-Atlantic, with exceedances at 5 of the 6 O_3 monitors in the state. The difference was dilute smoke transported from fires in Mexico and from agricultural burning in the Southeast and Mid-Atlantic regions of the U.S. This smoke was caught up in the westerly circulation on the west side of the Bermuda High and concentrated in Delaware on May 1. The presence of this smoke, in conjunction with strong sun and light westerly surface winds, compensated for the relatively low T_{max} on May 1. The smoke lingered on May 2 and promoted a second consecutive day of O_3 exceedances in the northern half of the state; a more southwesterly component to the synoptic winds brought O_3 back into the Moderate range in southern Delaware (not shown). By May 3, breezier southwesterly winds blew the smoke northward, returning O_3 to the Moderate range across the entire state.

2.2 Hot Weather

Observations during 2018 reinforced the recent trend that hot weather, while still necessary for O_3 exceedances, is no longer sufficient during the New Normal period. Figure 8 shows the percent of hot days in Delaware with maximum observed 8-hour average $O_3 \geq 71$ ppbv (orange dots/line) and ≥ 86 ppbv (red dots/line) for 1997-2018. After a temporary surge in the percent of hot days that corresponded with an O_3 exceedance in 2017 (30%), the value in 2018 (17%) dropped and aligned with the rest of the years in the New Normal period (12-22%).

Historically, the only other year with < 30% of hot days corresponding to an O_3 exceedance was the Great Recession year of 2009, which was analogous to the New Normal period in terms of low regional NO_x emissions due to the economic recession and an unusually cool summer. In the current New Normal period, most hot days ($\geq$ 70-80%) are not O_3 exceedance days in Delaware, and a forecast of $T_{max} \geq 90$ °F is no longer a reliable predictor of an O_3 exceedance.

2.3 Persistence O_3

Persistence O_3 – using the current day maximum observed 8-hour average O_3 value as the next day forecast – continued to be a weak O_3 predictor in 2018. Historically, persistence O_3 was a useful forecast tool because the lifetime of O_3 in the atmosphere is long enough to allow for multi-day transport. During regional O_3 events in the Pre- and Post-NO_x SIP periods, synoptic weather conditions promoted the build-up of O_3, making persistence O_3 an informative indicator of next day O_3 levels.

Figure 9 shows the linear relationship between maximum observed 8-hour average O_3 and local 1-day lag (persistence) maximum observed 8-hour average O_3 for Delaware in the New Normal period, including 2018. The coefficient of determination (R^2 value) of the persistence regression equation indicates that persistence O_3 only explains about 29% of the total variation in observed O_3 during the New Normal period, compared to 39% in the Pre-NO_x SIP period (1997-2002) and 33% in the Post-NO_x SIP period (Figure 10). Although this drop reinforces the current limits of persistence as an effective forecast tool in the New Normal period, an increased frequency of multi-day high O_3 events in 2018 relative to the rest of the years in the New Normal period, discussed in Section 3.1, drove the R^2 value of 0.29 for 2013-2018 up from the value of 0.24 for 2013-2017. This increase highlights that persistence can still be a useful forecast tool during multi-day high O_3 events, but its overall role is limited given the trend in the New Normal period toward more isolated single day exceedance "spikes" and exceedances at only one monitor location.

Even though multi-day events were more frequent in 2018, geographically and temporally isolated O_3 exceedance events also occurred. The exceedances on July 11 and September 5 were both single-day events that were observed by only one monitor, and while the June 30 exceedance, though part of a regional multi-day event, was observed by only one monitor in Delaware. Single day and isolated exceedances are difficult to forecast, as they are typically geographically isolated and depend on mesoscale weather features that are less accurately predicted by numerical weather models. The July 11 exceedance at Lewes underscores the difficulty that forecasters face in accurately identifying isolated single day exceedances. 2018 was the second consecutive year that an isolated, sea breeze-driven exceedance occurred in July at the Lewes monitor. This was a "missed" forecast, in that Code Orange O_3 (72 ppbv) was observed, but Code Yellow O_3 (68 ppbv) was forecasted. Figure 11 shows how isolated the Lewes exceedance was; observed 8-hour average O_3 ranged from Code Green in northern Delaware, northeastern Maryland, and southern Pennsylvania to Code Yellow across the rest of central and southern Delaware, southern New Jersey, and southern Maryland.

2.4 Climatology

As explained in Section 1, the historical seasonal O_3 cycle has virtually disappeared in the New Normal period. Figure 12 illustrates the historical climatological O_3 cycle, as shown by the daily time series of observed 8-hour average O_3 in Delaware during the Post-NO_x SIP period (blue dots/line), overlaid with the best polynomial fit (black line). This plot is similar to Figure 5 but focuses on the average day-to-day variations in peak O_3 in the New Normal versus Post-NO_x SIP periods and highlights the 2018 daily observations. The daily time series of maximum observed 8-hour average O_3 in Delaware during 2018 (red dots/line in Figure 12) departs from the historical O_3 seasonal cycle, consistent with other years in the New Normal period. This change makes O_3 climatology a much weaker foundation for the daily O_3 forecast in the New Normal period compared to previous periods, limiting its usefulness as a foundation to anchor the daily O_3 forecast, with the 25th percentile, median, 75th percentile, and 95th percentile values as benchmarks.

The recent changes in the climatology of the peak of O_3 season also continued in 2018 (Figure 13). During the Pre-NO_x SIP (blue bars) and Post-NO_x SIP (red bars) periods, O_3 season peaked during June, July, and August, with the highest percentage of O_3 exceedance days occurring in July. There was a smaller but consistent percentage of exceedance days in May, and a marked decrease in September. In the New Normal period (green bars), however, the historic distribution of O_3 exceedances has shifted at the beginning and end of the season. Figure 13 illustrates that the percent of seasonal exceedances in June and July is approximately the same, but the peak in August has dramatically decreased (down to ~6% in the New Normal from ~ 22% in the Pre- and Post-NO_x SIP), while there has been a slight increase in the percent of exceedances in May (up to ~17% from 12.5-14% in the Pre- and Post-NO_x SIP) and a surge in September (up to ~18.5% in the New Normal from ~7% in the Pre- and Post-NO_x SIP). It is still too early to know the exact cause of this shift, but it is likely related to a combination of meteorology (perhaps a shift towards more humid conditions in August and hotter conditions in September) and NO_x emissions (possibly due to shifts in energy demand or cap-and-trade program credit distributions). The end result is that the majority of O_3 exceedance days are now clustered early in the summer season, during May, June, and the beginning of July, with a late season surge in September.

3. Skill of Ozone Forecasts in 2018

Despite the continuing challenges associated with the deterioration of historical forecast predictors like hot weather, O_3 persistence, and climatology, expert O_3 forecasts in 2018 were skillful, similar to recent years. Figure 14 shows the time series of 8-hour average O_3 forecasts and observations for 2018. Throughout the season, O_3 forecasts closely matched the day-to-day variations in observed O_3 values. Note that the time series of forecasted values (purple dots/dashed line) is discontinuous due to gaps associated with two-day lead forecasts of AQI color codes for most Sundays and Mondays.

3.1 Hit Rate

Expert skill in predicting O_3 exceedance days in 2018 improved compared to 2017. Figure 15 illustrates the false alarm ratio (previously called false alarm rate; red bars) and hit rate (blue bars) for expert forecasts of O_3 exceedance days in 2011-2018. Details on the calculation of threshold skill measures are given in Appendix A; the exceedance threshold was 76 ppbv prior to 2016.

Hit rate in 2018 (0.63) improved compared to 2017 (0.50) and 2016 (0.55) and was equivalent to 2012 (0.63) during the Post-NOx SIP period, when the much higher number of O_3 exceedance days (31 in 2012) drove up hit rates. This increase in skill in 2018 relative to the past two years was due to the increase in the number of hits in 2018 (5) compared to 2017 (4) (Figure 16).

The higher hit rate in 2018 is at least partially attributable to the increased frequency of multi-day high O_3 events, which were much more common in the years prior to 2013, as discussed in Section 2.3. All 5 of the correctly identified O_3 exceedances in 2018 fell during regional multi-day O_3 episodes: May 1-3, June 17-18, June 30 to July 3, and July 9-10. Note that not all days in each of these 2018 episodes were exceedance days in Delaware, but on days that Delaware did not observe exceedances, there were exceedances in the neighboring Baltimore, Maryland and/or Philadelphia, Pennsylvania forecast areas. Thus, the episodes were regional and multi-day in that there were continuous observed exceedances during these periods at locations along the Interstate-95 Corridor from Baltimore to Philadelphia. Exceedances during regional multi-day O_3 episodes, driven by synoptic scale meteorology, are much easier for forecasters to predict than single day events.

3.2 False Alarm Ratio

Although hit rate improved in 2018 relative to 2017, there was a deterioration in false alarm ratio (Figure 15). The false alarm ratio in 2018 was 0.58, which is still a marked improvement from the extremely high values in 2013-2015 (0.71-1.0), but is a decline from the relatively lower values in 2017 (0.43) and 2016 (0.33). Figure 16 shows there were 7 expert false alarm forecasts in 2018, compared to an average of 3.8 during the 2013-2017 period.

Table 2 lists details about the 7 expert false alarm forecasts for Delaware in 2018. More than half of the false alarms (4/7) were issued for days that fell during regional multi-day high O_3 events: May 3, June 17, and July 1-2. As illustrated in Figure 17, the neighboring metropolitan areas of Baltimore and/or Philadelphia observed O_3 exceedances on all of these days, even though O_3 only reached the upper Code Yellow range in Delaware. Thus, Code Orange exceedance day forecasts were not unreasonable for Delaware in this context. Setting aside these 4 false alarms that fell during regional multi-day O_3 episodes, the remaining 3 false alarm forecasts in 2018 are comparable in number and magnitude to recent years.

It is not clear if the lower peak O_3 in the Wilmington metropolitan area in 2018, relative to the Philadelphia and Baltimore metropolitan areas, is the beginning of a trend or is a unique aspect of the 2018 O_3 season. Two of the Delaware false alarm forecast days fell on a Sunday (June 17 and July 1), which is the only day of the week that has statistically lower observed O_3, attributed to lower mobile emissions. Possibly mobile emissions were just low enough in Wilmington on June 17 and July 1 to preclude O_3 exceedances, despite very favorable synoptic meteorological conditions. In addition, Delaware seems to be becoming very sensitive to synoptic surface winds with a southerly component (e.g., false alarm forecast on May 3); even southwesterly winds seem to be sufficient to bring cleaner air into the state and keep O_3 below the Code Orange range.

Despite the increase in false alarm ratio in 2018 relative to 2016-2017, the fact that the hit rate was larger than the false alarm ratio for the fourth year in a row underscores the overall skill of expert forecasts of O_3 exceedance days.

3.3 Overall Forecast Skill

Overall forecast skill (all days) for 2018 was comparable to recent years. Figure 18 shows that forecasts in 2018 were consistent with 2017 in terms of accuracy (median absolute error of 5.0 ppbv). Expert forecasts in 2018 showed a slight tendency towards over-forecasting, with a bias of 1.9 ppbv (unbiased forecasts have a bias of 0.0 ppbv). The 2018 bias value is an increase from the relatively low bias of 0.70 ppbv in 2017, but it is comparable to 2016 (bias of 2.1 ppbv).

4. Performance of Ozone Numerical Forecast Models in 2018

Numerical air quality models struggled to predict O_3 exceedance days in 2018, similar to recent years. Figure 19 shows the false alarm ratios (red bars) and hit rates (blue bars) of the available numerical models in 2018, compared to the expert threshold forecasts for Delaware in 2018. The models used in 2018 were the same as in 2017, and included the NOAA-EPA model (NAQFC CMAQ model), the North Carolina Division of Air Quality model (NC CAMx model), and two versions of the Baron Advanced Meteorological Services (BAMS) model, the CMAQ and MAQSIP models.

4.1 BAMS and North Carolina Models

As has been the case in recent years, the skill of the individual air quality models was not consistent from the previous year. In 2017, the NC model was extremely unskillful, with a very high false alarm ratio of 0.86 and a low hit rate of only 0.13. In contrast, the NC model performed much more skillfully in 2018, with a false alarm ratio and hit rate both equal to 0.50. Despite this improvement, the hit rate of the expert threshold forecasts (0.63) was superior. Similarly, while the BAMS models were surprisingly skillful in 2017, they had relatively high false alarm ratios of 0.76 (MAQSIP) and 0.69 (BAMS) in 2018. The BAMS models' hit rates also fell in 2018 (0.50) relative to 2017 (0.75 for the RT and 0.63 for the CMAQ), and were lower than the expert threshold forecasts. The inconsistent year-to-year performance of the BAMS and NC models makes it difficult to recognize when they may be providing skillful versus unskillful guidance during the O_3 season. For example, the slightly lower false alarm ratio of the NC model in 2018 may have been able to eliminate one of the expert false alarms, if forecasters had known they could rely on the guidance. In the New Normal period, it has become much easier to recognize the skill of a numerical model at the end of the season rather than in an operational setting in the midst of the season, making O_3 numerical models less reliable forecast predictors compared to the Post-NO_x SIP period.

4.2 NAQFC Model

As has been the case throughout the New Normal period, the NAQFC model had relatively poor skill in predicting the O_3 exceedance days in 2018, with the lowest hit rate (0.38) of the available models (Figure 19). This relatively low hit rate has been essentially consistent (0.33-0.38) since 2014 (Figure 20), aside from 2015 when the NAQFC had zero hits, making the hit rate undefined (Figure 21). The NAQFC had the same number of hits in 2018 (3) as 2017, compared to 5 hits for the expert forecasts. Unlike in 2017, when the false alarm ratio of NAQFC predictions temporarily dropped (0.25), the false alarm ratio in 2018 rose to 0.40. As with the BAMS and NC models, the inconsistent year-to-year performance of the NAQFC model in regard to false alarms makes it an unreliable forecast tool.

The performance of the NAQFC model for all days (Figure 22) showed a slight decline in accuracy in 2018, with a median absolute error of 6.0 ppbv compared to 5.0 ppbv in 2017. The bias of the model was similar to 2017, showing a slight tendency to under-forecast, with a bias of -1.7 ppbv compared to -1.3 ppbv in 2017. This was the second consecutive season that the model tended to under-predict; the NAQFC model had a small tendency to over-forecast in 2014-2016 (~ +2.0 ppbv). The overall skill of the NAQFC model in the New Normal period demonstrates that it provides useful forecast guidance in general, but it is not reliable for differentiating between high Code Yellow and Code Orange O_3 days.

5. Outlook for 2019

An extension of the New Normal period is expected in 2019, with a range of 6-12 O_3 exceedance days in Delaware. The same forecast challenges that have emerged since 2013 will continue in 2019, with added uncertainty regarding the possible trend in lower observed peak O_3 in the Wilmington metropolitan area during regional multi-day air quality events compared to the Philadelphia and Baltimore metropolitan areas. If this trend from 2018 extends into 2019, another season with a higher than average number of false alarm forecasts is possible. Every effort will be made to recognize and to compensate for this trend, however, if it does continue. Since forecasters can no longer depend on hot weather, O_3 persistence, climatology, or numerical air quality models as reliable forecast predictors, a return to the consistently high hit rates (> 0.60) of the pre-2013 period is unlikely. Nevertheless, the trend of higher expert hit rates and lower false alarm ratios, which has emerged over the past four seasons, should continue.

An operational version of NOAA's "bias corrected" NAQFC model guidance, generated using an analog Kalman filter method, should be available during the 2019 season. This post-processed product may have greater skill in identifying O_3 exceedance days in Delaware compared to the direct model output.

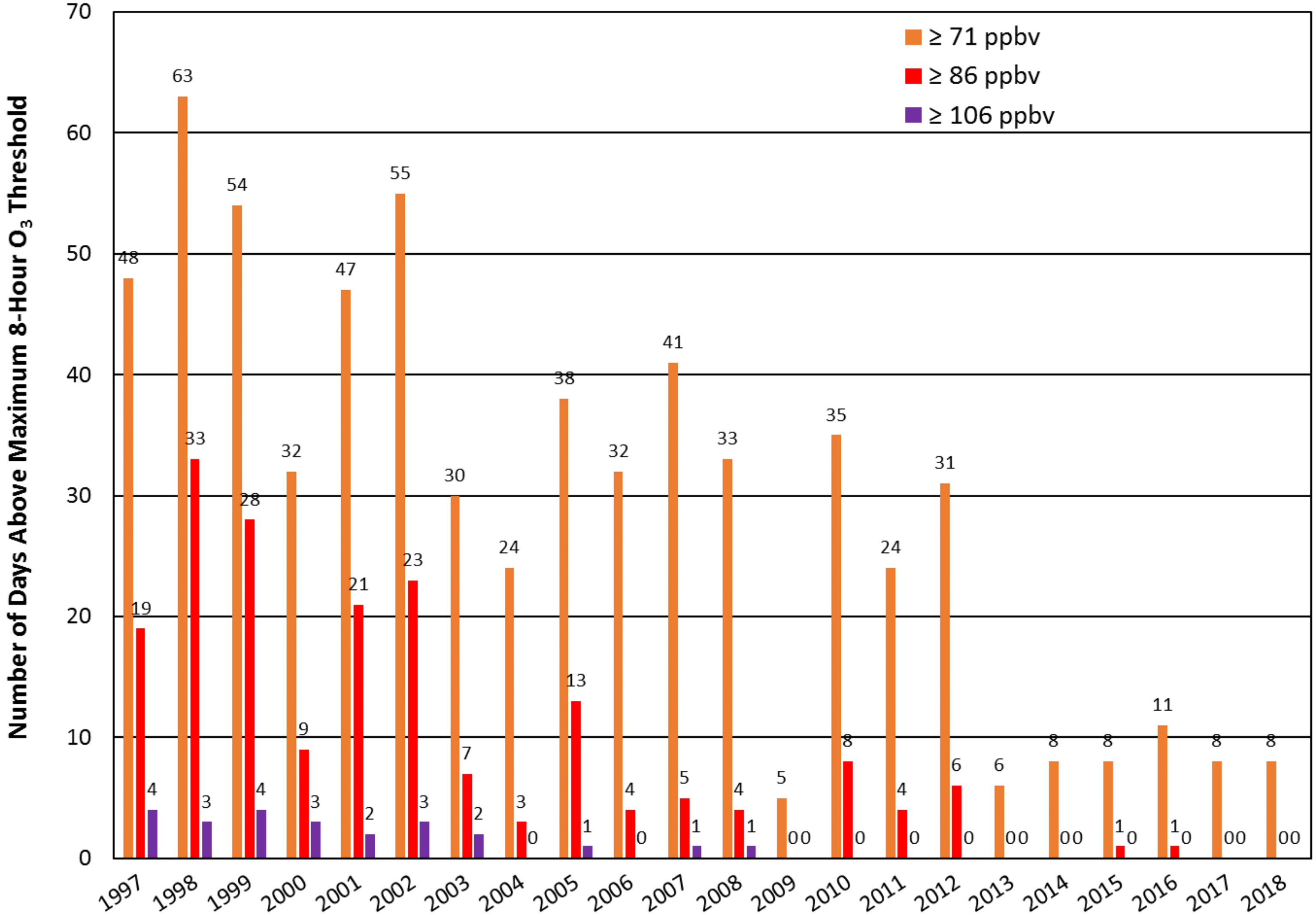

Figure 1. Number of days with maximum observed 8-hour average O_3 at or above thresholds of 71 ppbv (orange bars), 86 ppbv (red bars), and 106 ppbv (purple bars) in Delaware for 1997-2018

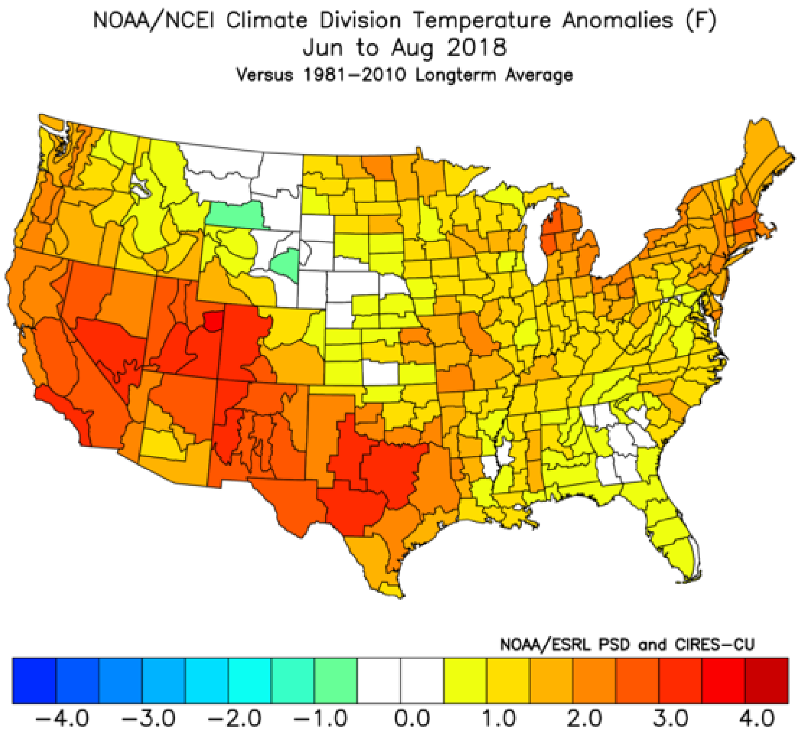

Figure 2. Temperature anomalies (in °F) in the U.S. for June-August 2018 compared to the 1981-2010 long-term average (courtesy of NOAA/ESRL).

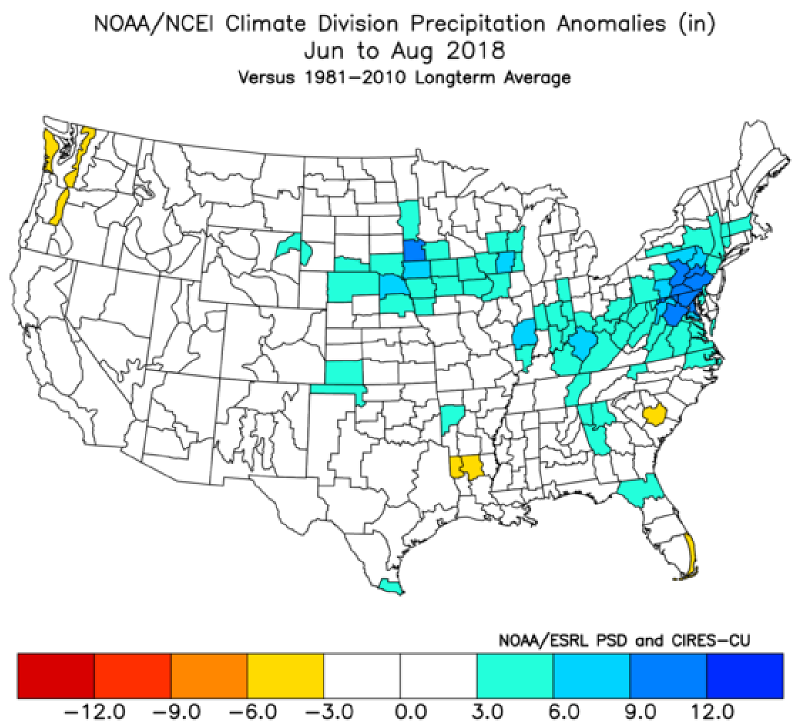

Figure 3. Precipitation anomalies (in inches) in the U.S. for June-August 2018 compared to the 1981-2010 long-term average (courtesy of NOAA/ESRL).

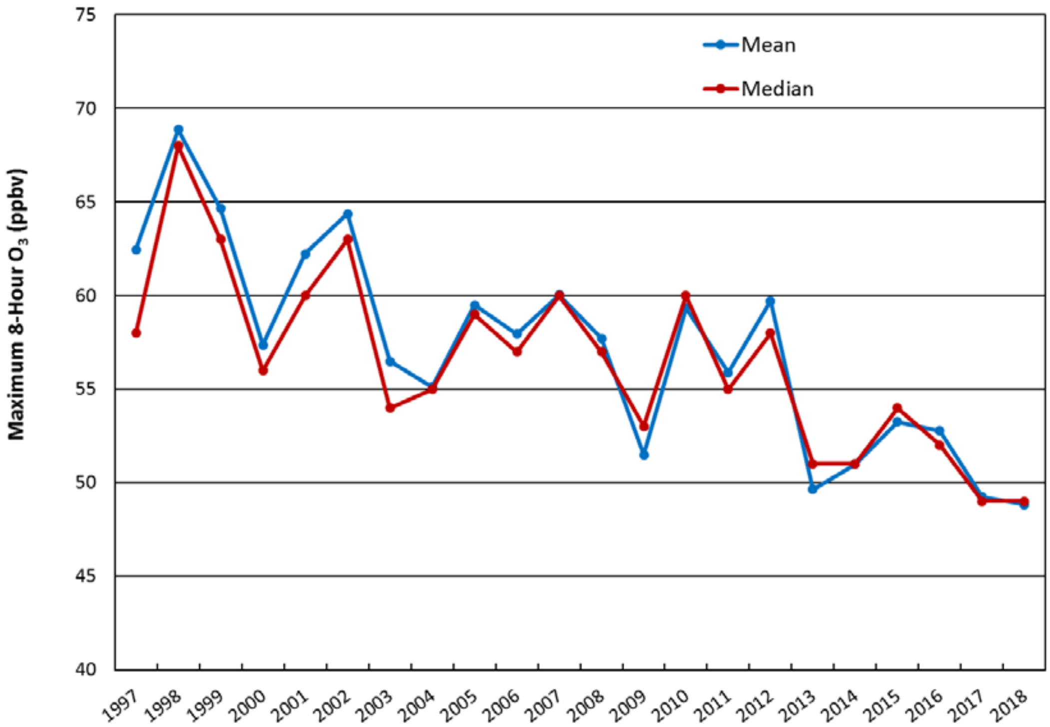

Figure 4. Seasonal (May 1 to September 30) mean and median of maximum observed 8-hour average O_3 in Delaware for 1997-2018.

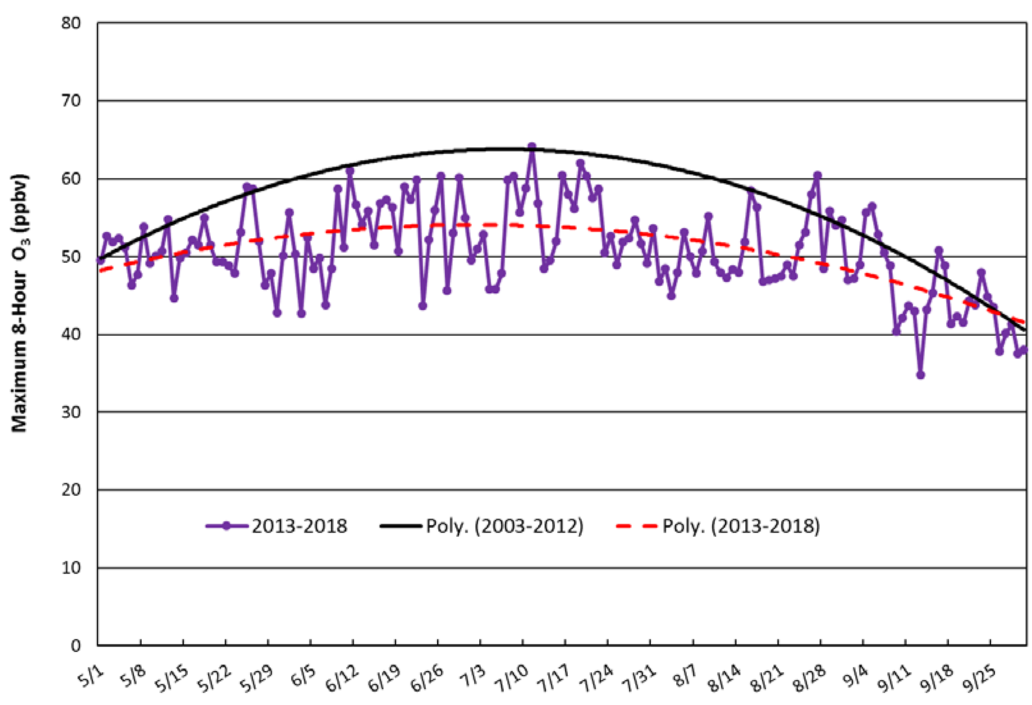

Figure 5. Daily time series of average maximum observed 8-hour average O_3 in Delaware during the "New Normal" period (2013-2018; purple dots/line) and best polynomial fit (red dashed line) compared to the best polynomial fit to the 2003-2012 average (black line).

Table 1. Details regarding observed O_3 exceedance days in Delaware in 2018. Grey shading indicates days that O_3 exceedances were correctly forecasted, with alerts issued to the public. T_{max} indicates maximum temperature at Wilmington-New Castle County Airport (KILG).

Date	Day of Week	T_{max} KILG (°F)	Max 8-Hr O_3 (ppbv)	Number of Monitors	Conditions
5/1	Tue	80	75	5	Bermuda High (full sun, low humidity); smoke likely
5/2	Wed	88	74	3	Bermuda High (full sun, low humidity); smoke likely
6/18	Mon	92	76	3	Bermuda High (hot, sunny, humid)
6/30	Sat	97	71	1	Strong upper level ridge overhead; middle of heat wave
7/9	Mon	87	73	2	Mid-level ridge overhead (stagnant, sunny, low humidity)
7/10	Tue	95	85	6	Mid-level ridge overhead (stagnant, hot, sunny, low humidity, light westerly surface winds)
7/11	Wed	92	72	1	Sea breeze at Lewes (clean everywhere else)
9/5	Wed	93	77	1	Upper level ridge overhead (hot and stagnant)

Table 2. Details regarding expert false alarm forecasts of O_3 exceedances in Delaware in 2018. T_{max} indicates maximum temperature at Wilmington-New Castle County Airport (KILG); "over-forecast" is the difference between forecasted and observed maximum 8-hour average O_3.

Date	Day of Week	T_{max} KILG (°F)	Max 8-Hr O_3 (ppbv)	Over-Forecast (ppbv)	Conditions
5/3	Thu	90	66	6	Too windy?; exceedances in PHL, BAL
5/25	Fri	85	66	5	Regional smoke but too much southerly flow in DE
6/17	Sun	88	67	4	Lower Sunday emissions?; exceedances in PHL
7/1	Sun	94	68	5	Lower Sunday emissions?; exceedances in PHL, BAL
7/2	Mon	95	65	9	Exceedances in PHL, BAL
8/28	Tue	93	57	15	Thick smoke scattered sunlight, reduced O_3 production
9/6	Thu	93	64	7	Previous day missed exceedance; no change in weather

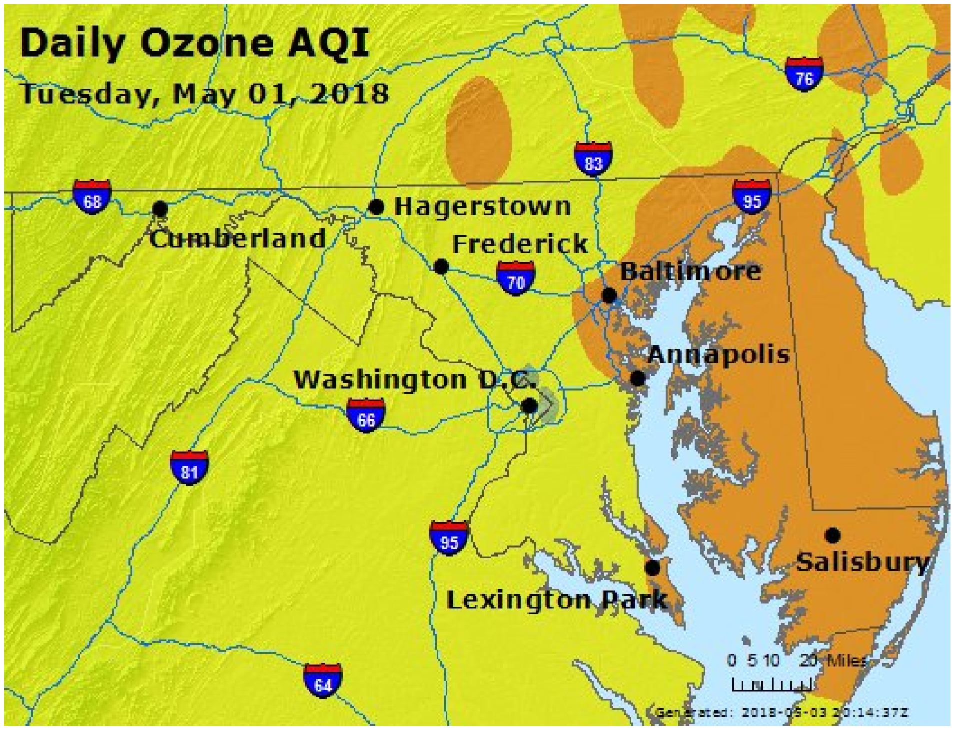

Figure 6. Observed O_3 air quality index (AQI) color codes on May 1, 2018, showing widespread Code Orange O_3 across almost the entire Delmarva Peninsula.

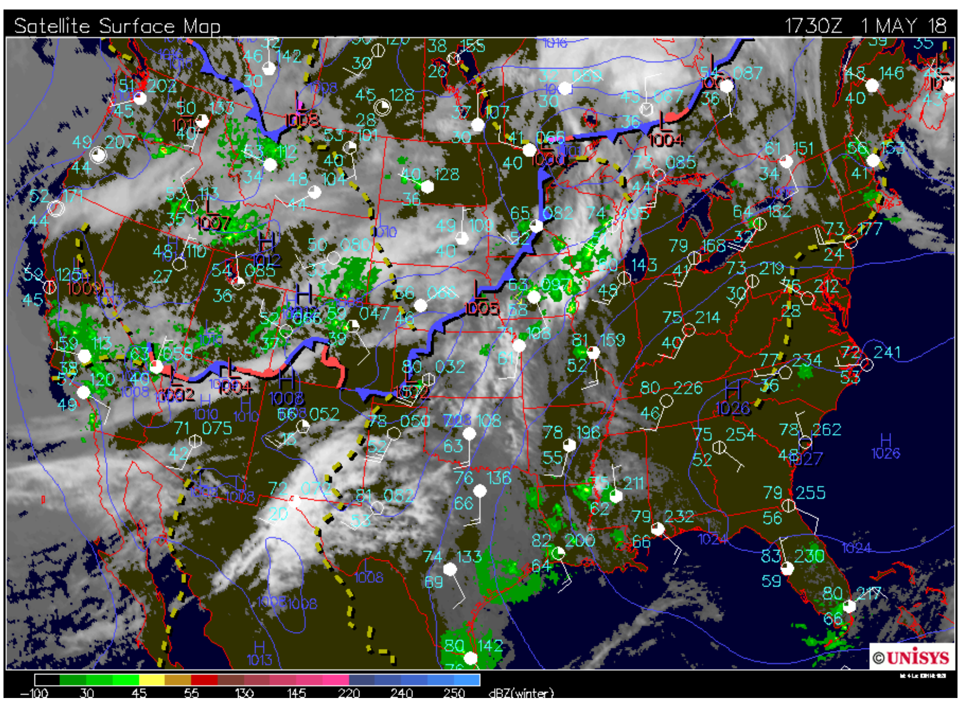

Figure 7. Combined visible satellite, radar, and surface analysis map of the continental U.S. on May 1, 2018 at 17:30 UTC, showing clear skies and light westerly surface winds across the Mid-Atlantic region due to the location of the Bermuda High pressure over NC and the coastal Atlantic Ocean.

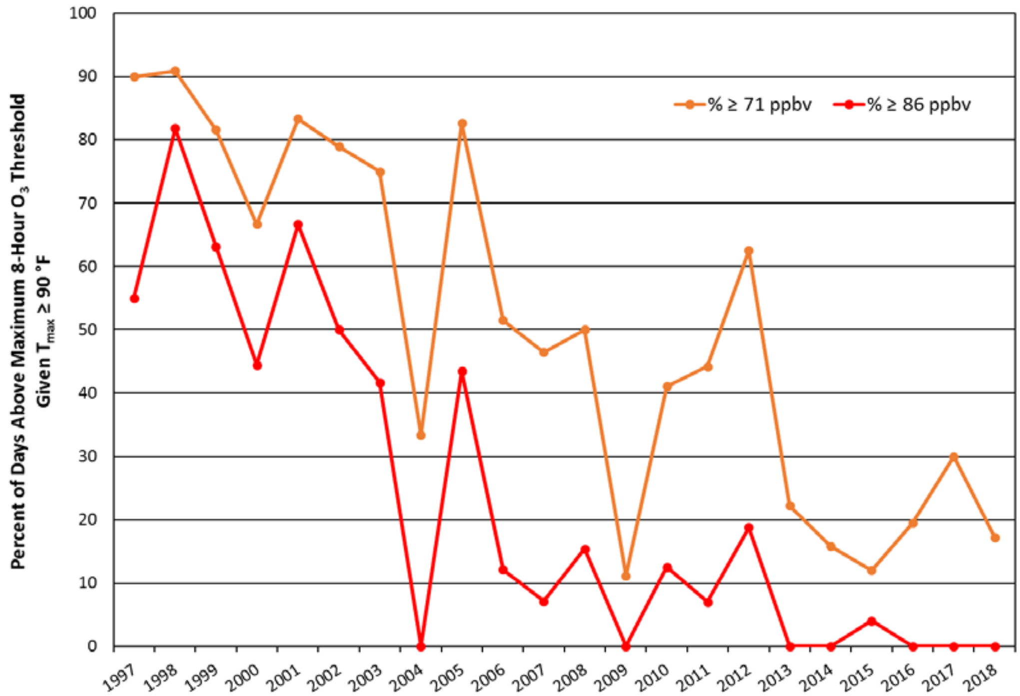

Figure 8. Percent of days with maximum temperature (T_{max}) ≥ 90 °F and maximum observed 8-hour average O_3 ≥ 71 ppbv (orange dots) or ≥ 86 ppbv (red dots) in Delaware for 1997-2018.

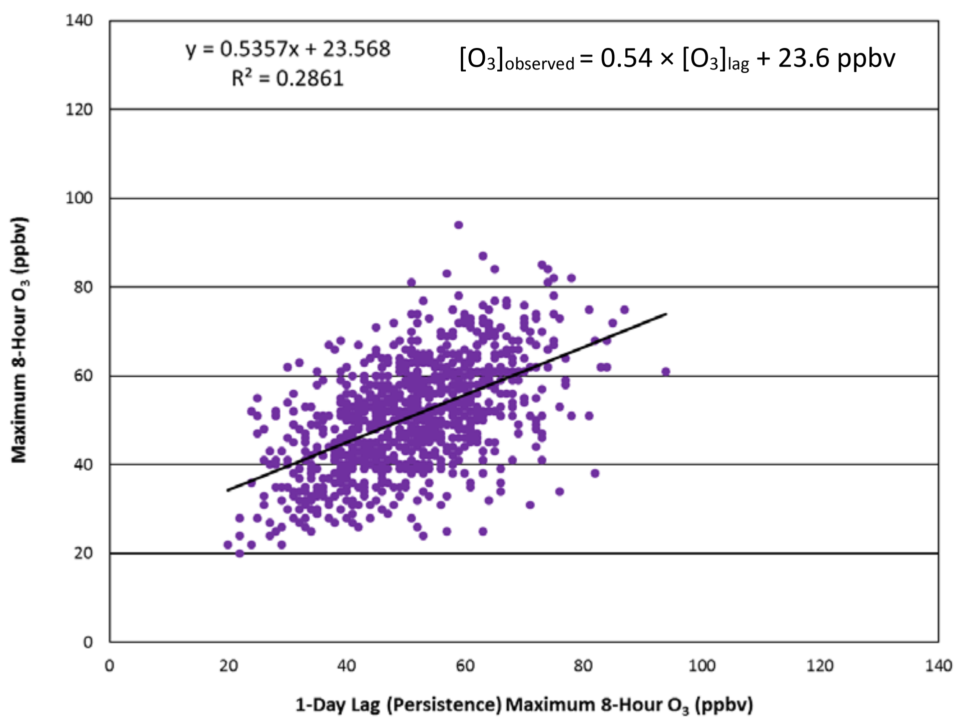

Figure 9. Linear relationship between seasonal (May 1 to September 30) maximum observed 8-hour average O_3 and local 1-day lag (persistence) maximum observed 8-hour average O_3 in Delaware for 2013-2018.

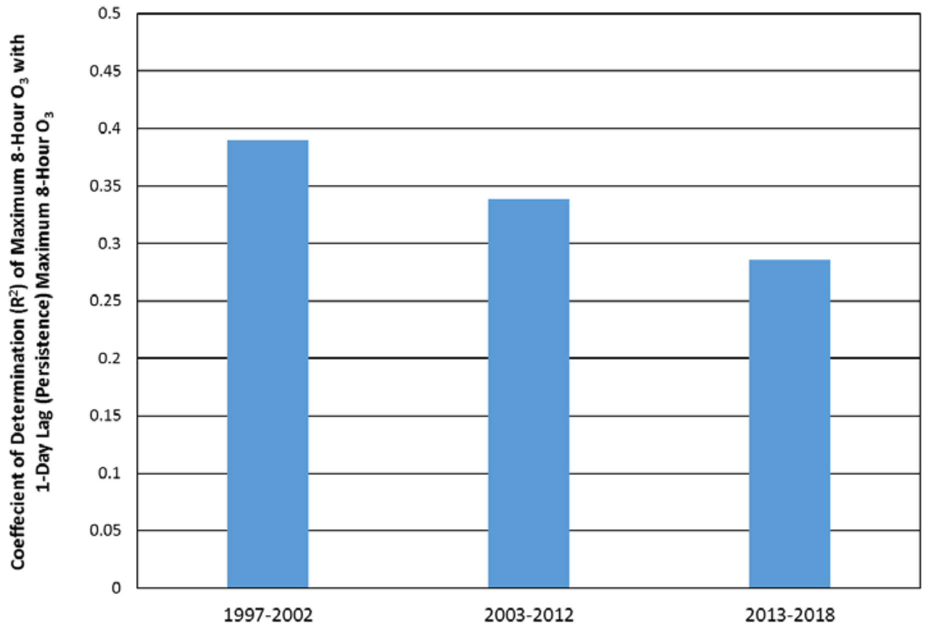

Figure 10. Coefficient of determination (R^2) of linear relationship between seasonal (May 1 to September 30) maximum observed 8-hour average O_3 and local 1-day lag (persistence) maximum observed 8-hour average O_3 in Delaware for given periods.

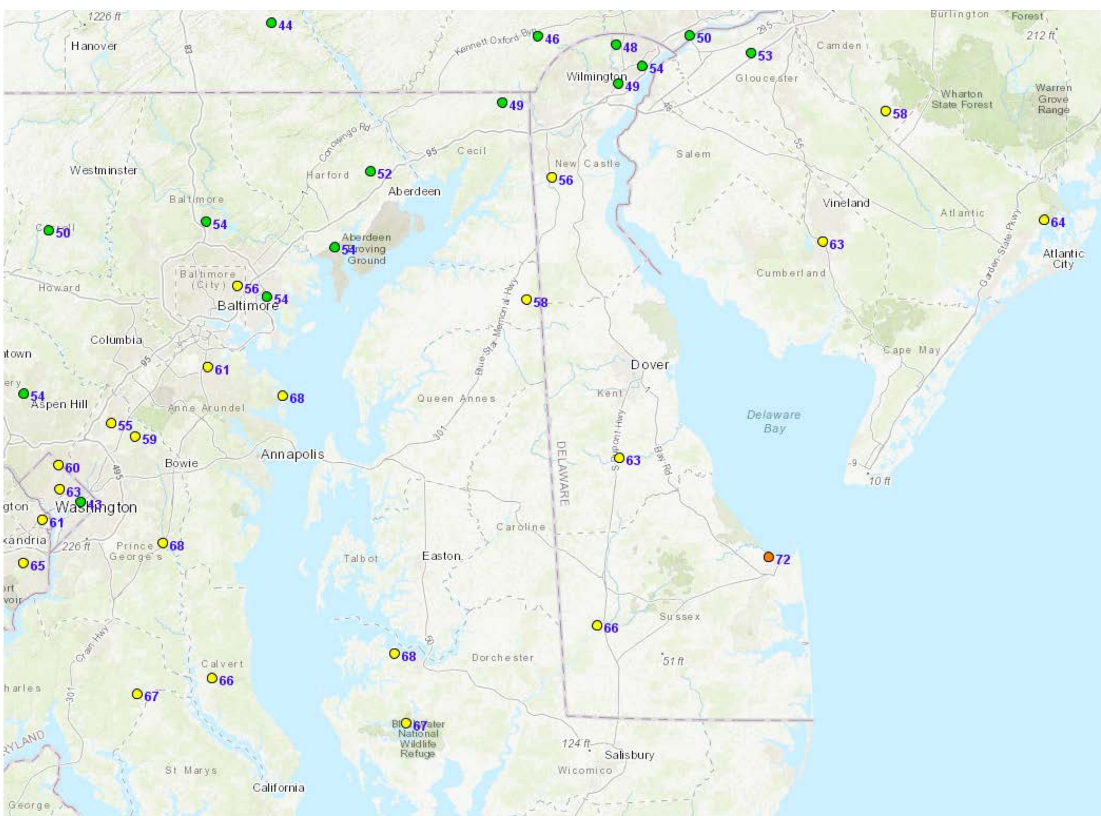

Figure 11. Map of mximum observed 8-hour average O_3 on July 11, 2018, showing the isolated O_3 exceedance at the Lewes, Delaware monitor (courtesy of AirNow-Tech Navigator).

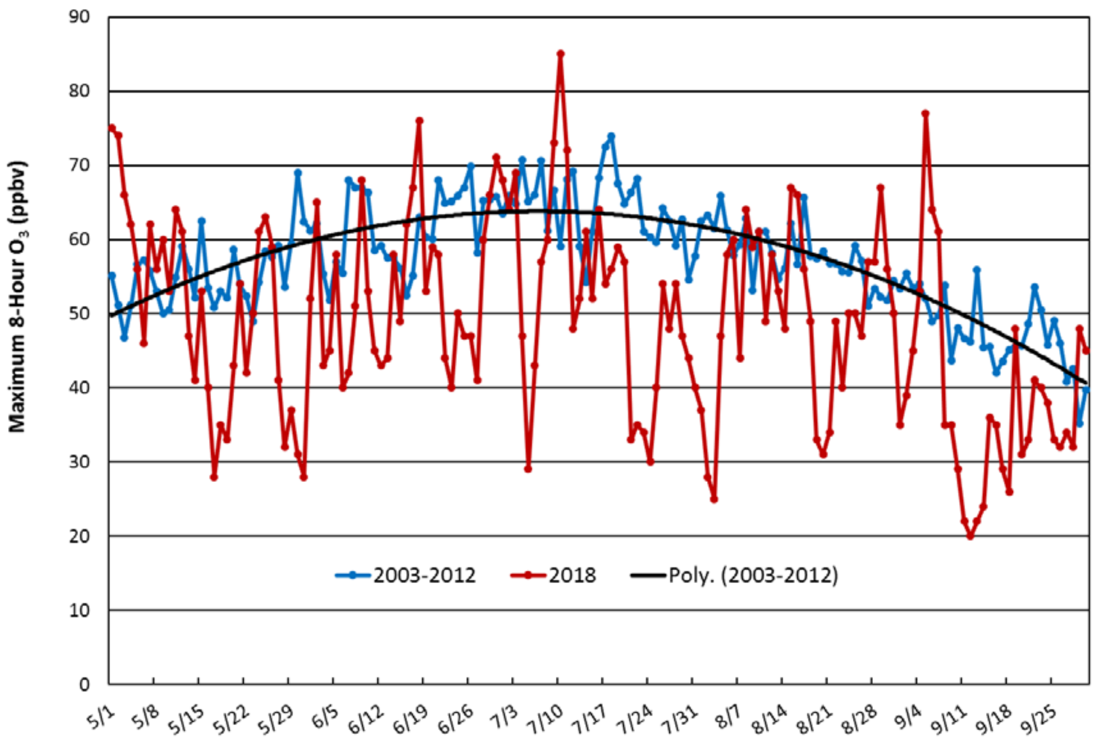

Figure 12. Daily time series of maximum observed 8-hour average O_3 in Delaware for 2018 (red line) compared to the 2003-2012 average (blue line). The black line is the best polynomial fit to the 2003-2012 average.

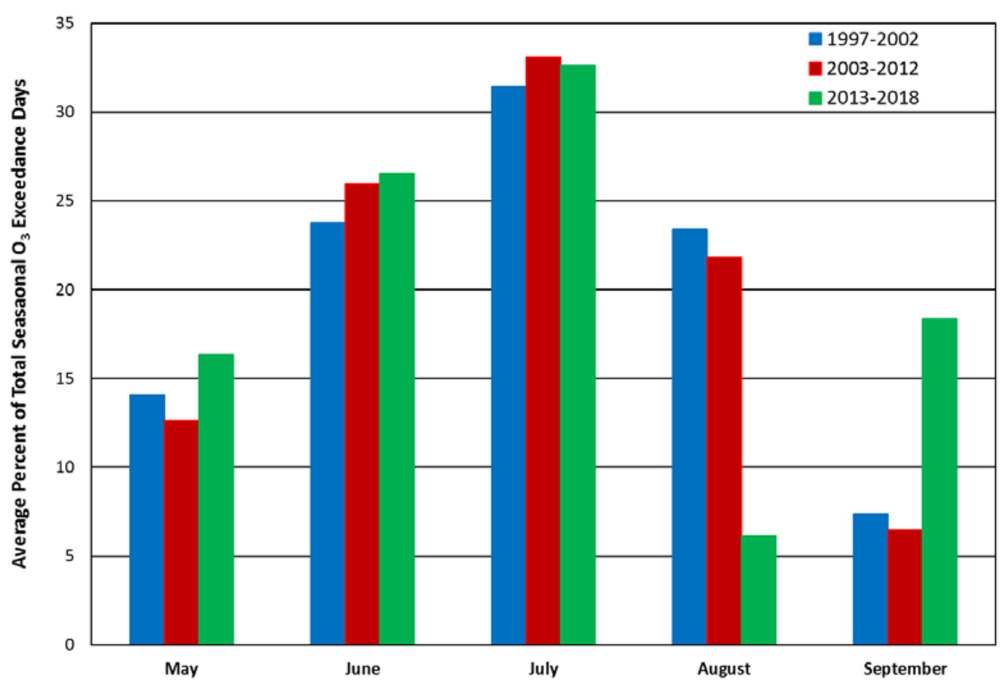

Figure 13. Average percent per month of total seasonal number of days that maximum observed 8-hour average O_3 exceeded the threshold of ≥ 71 ppb in Delaware for the given periods.

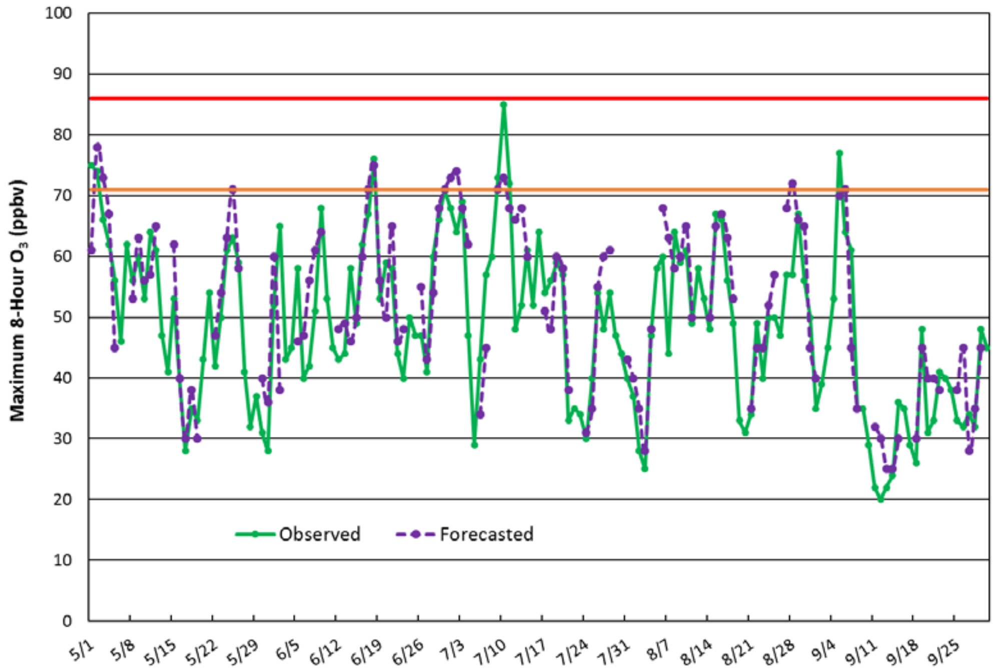

Figure 14. Maximum 8-hour average O_3 forecasts and observations for Delaware during May 1 to September 30, 2018. The orange and red lines indicate the Code Orange and Code Red thresholds, respectively.

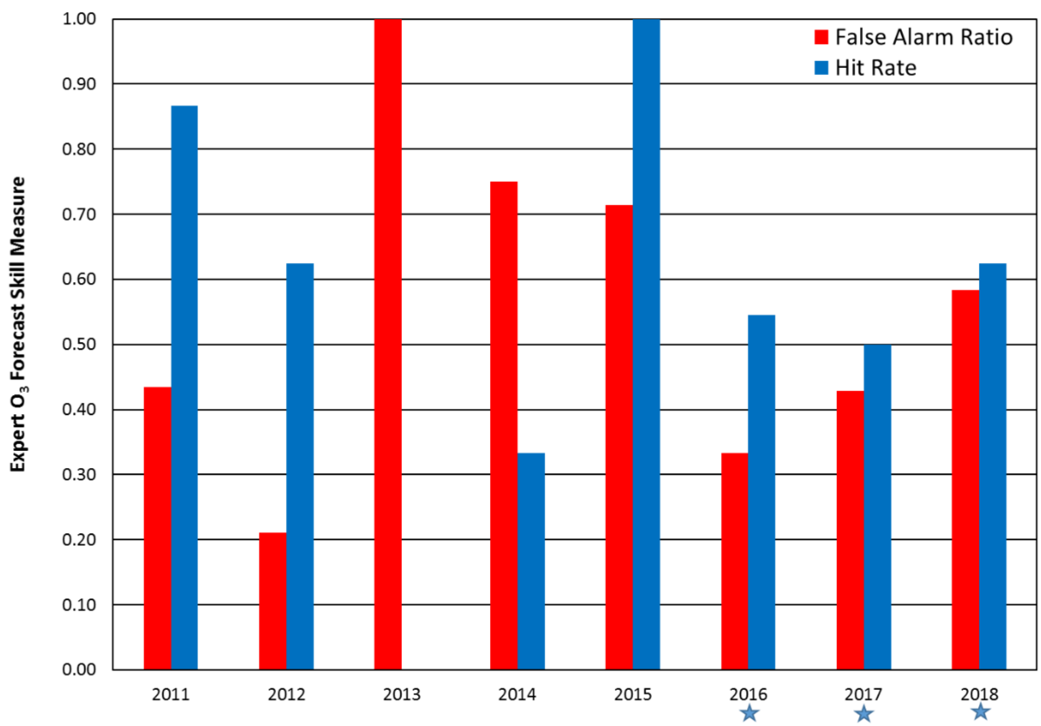

Figure 15. False alarm ratio and hit rate for threshold forecasts of maximum 8-hour average O_3 in Delaware for 2011-2018. Note that for 2011-2015, the threshold was 76 ppbv, but in 2016, it lowered to 71 ppbv (designated by the blue stars).

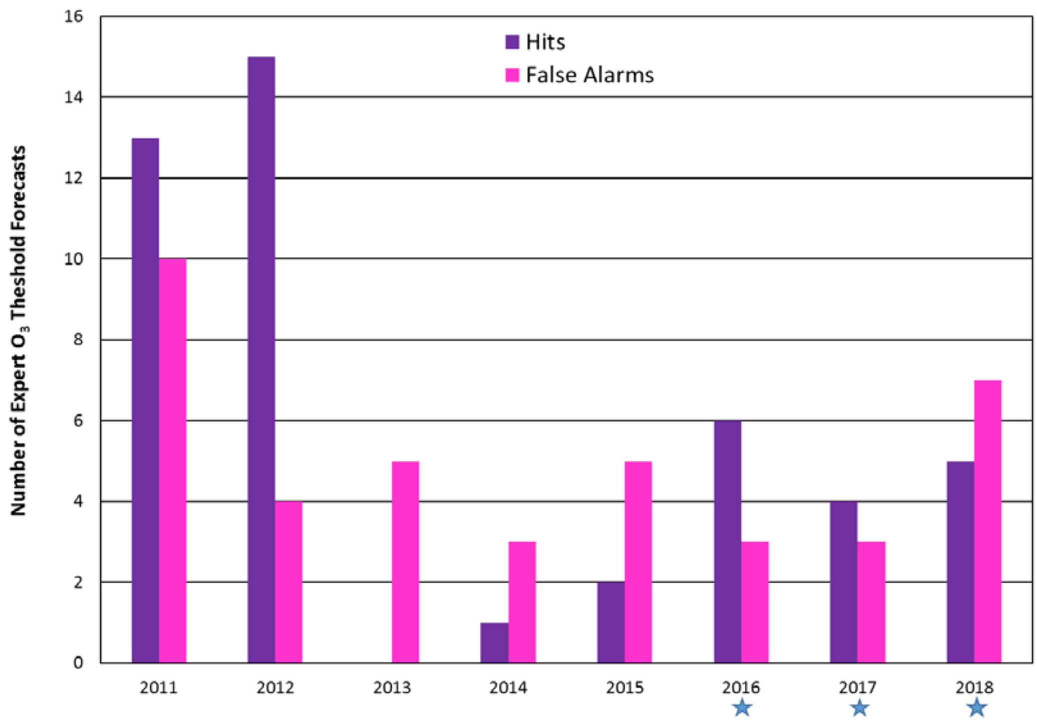

Figure 16. Number of false alarms and hits for threshold forecasts of maximum 8-hour average O_3 in Delaware for 2011-2018. Note that for 2011-2015, the threshold was 76 ppbv, but in 2016, it lowered to 71 ppbv (designated by the blue stars).

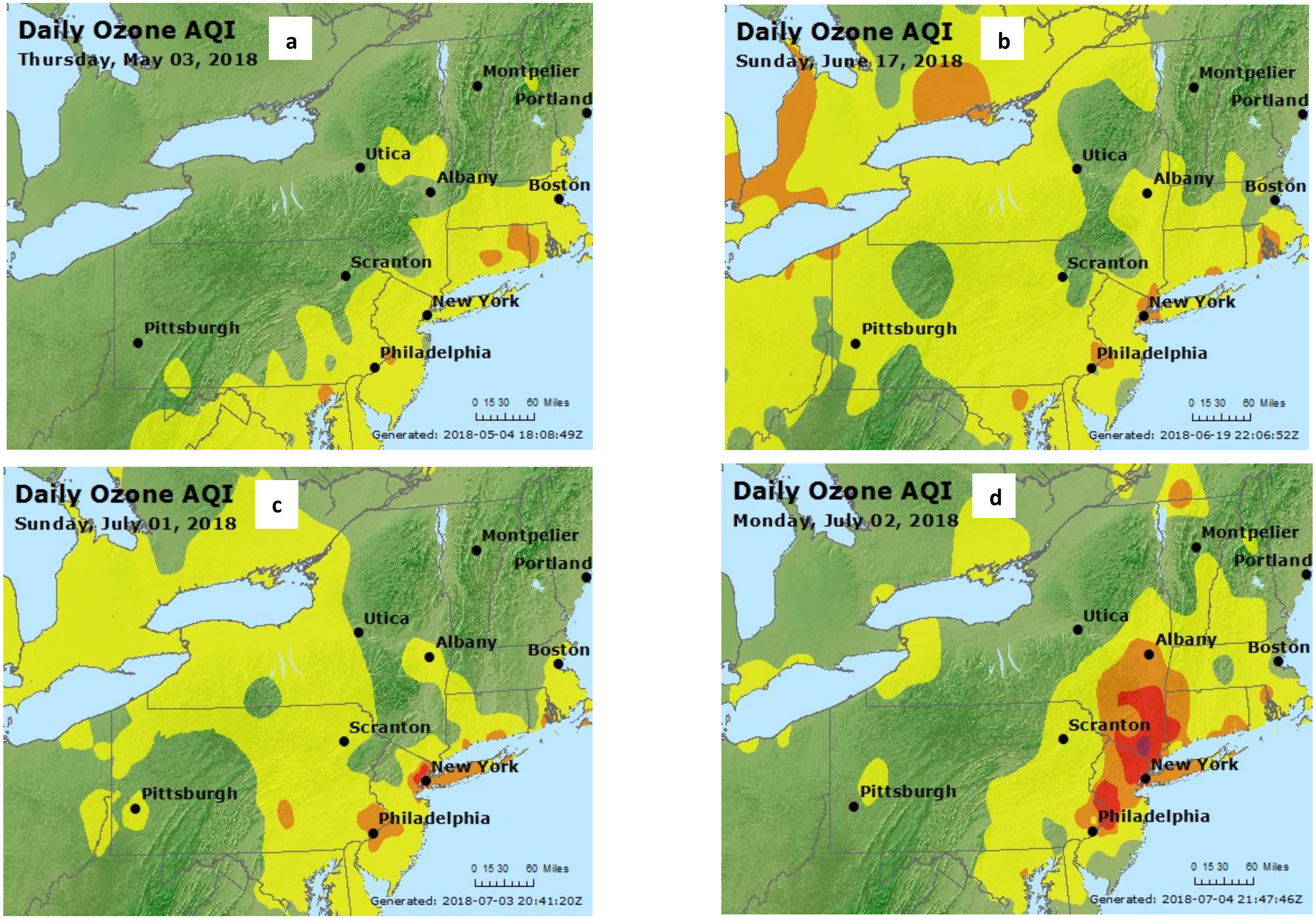

Figure 17. Observed O_3 air quality index (AQI) color codes for the northern Mid-Atlantic and southern New England regions on a) May 3, b) June 17, c) July 1, and d) July 2, 2018. All of these days were expert false alarm forecast days in Delaware, but neighboring metropolitan areas observed O_3 exceedances (Code Orange or Code Red).

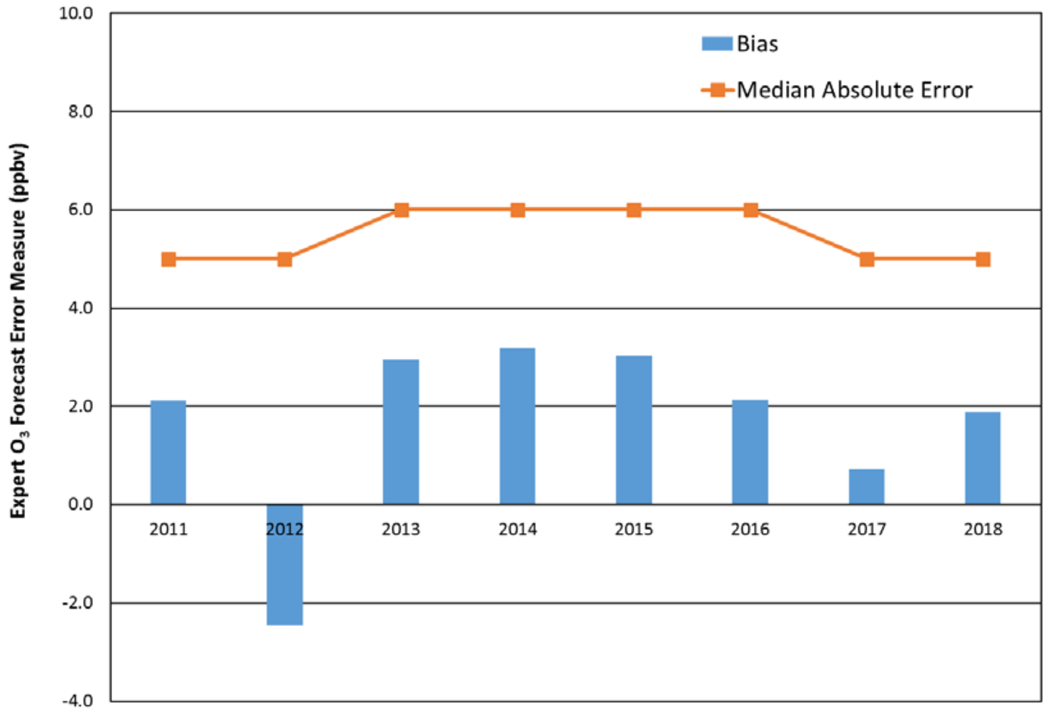

Figure 18. Error statistics for all maximum 8-hour average O_3 forecasts, May 1 to September 30, in Delaware for 2011-2018.

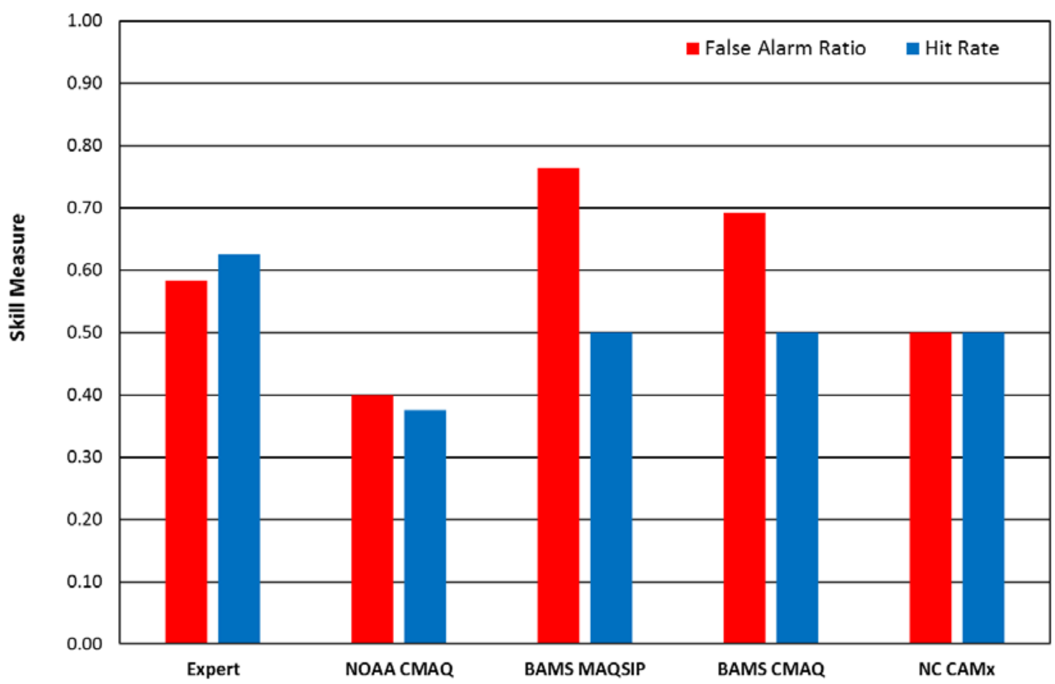

Figure 19. False alarm ratio and hit rate for expert and numerical air quality model threshold predictions of maximum 8-hour average O_3 in Delaware for 2018. Results from the NAQFC model (NOAA CMAQ), two variations of the Baron Meteorological Services (BAMS) models (MAQSIP and CMAQ), and the North Carolina Division of Air Quality (NC CAMx) model are shown.

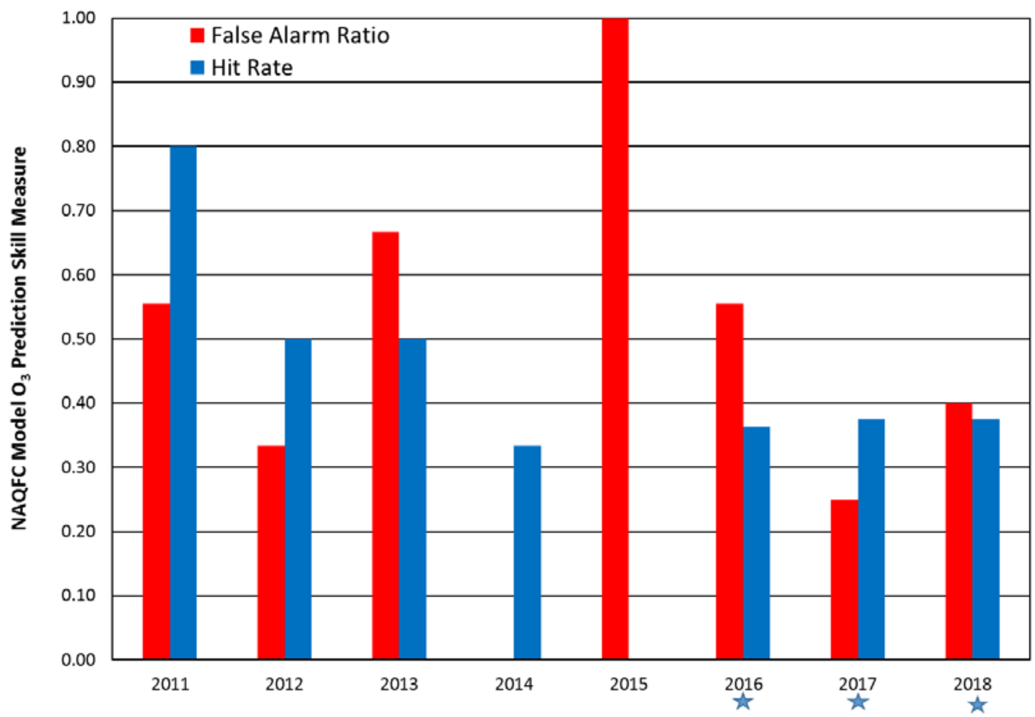

Figure 20. False alarm ratio and hit rate for threshold predictions of maximum 8-hour average O_3 by the NAQFC model in Delaware for 2011-2018. Note that for 2011-2015, the threshold was 76 ppbv, but in 2016, it lowered to 71 ppbv (designated by the blue stars).

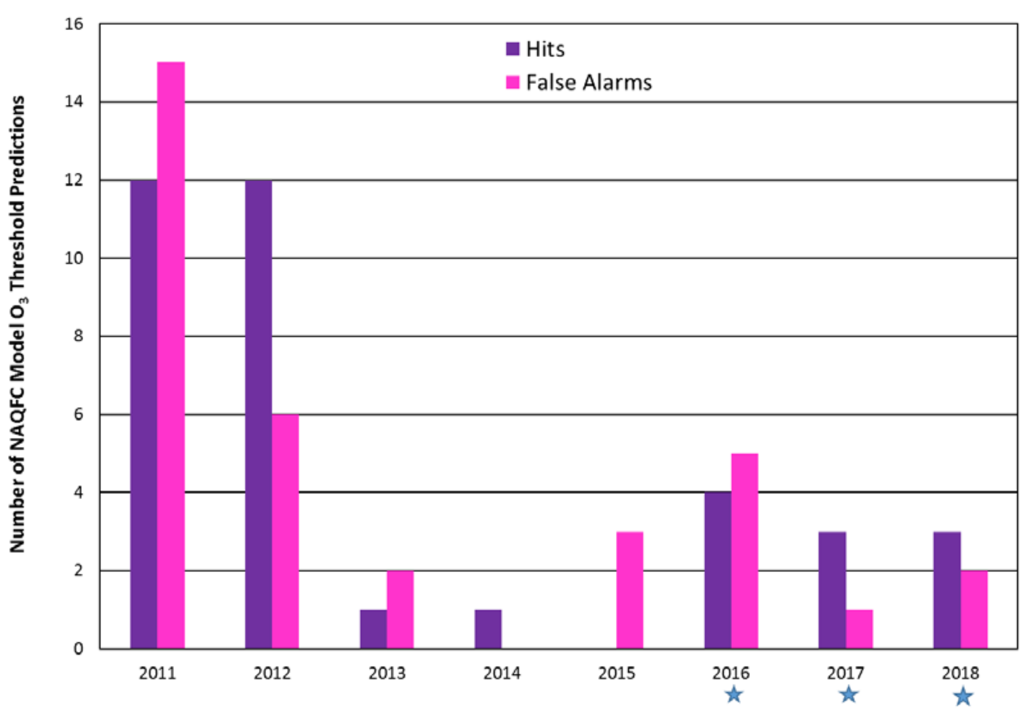

Figure 21. Number of false alarms and hits for threshold predictions of maximum 8-hour average O_3 by the NAQFC model in Delaware for 2011-2018. Note that for 2011-2015, the threshold was 76 ppbv, but in 2016, it lowered to 71 ppbv (designated by the blue stars).

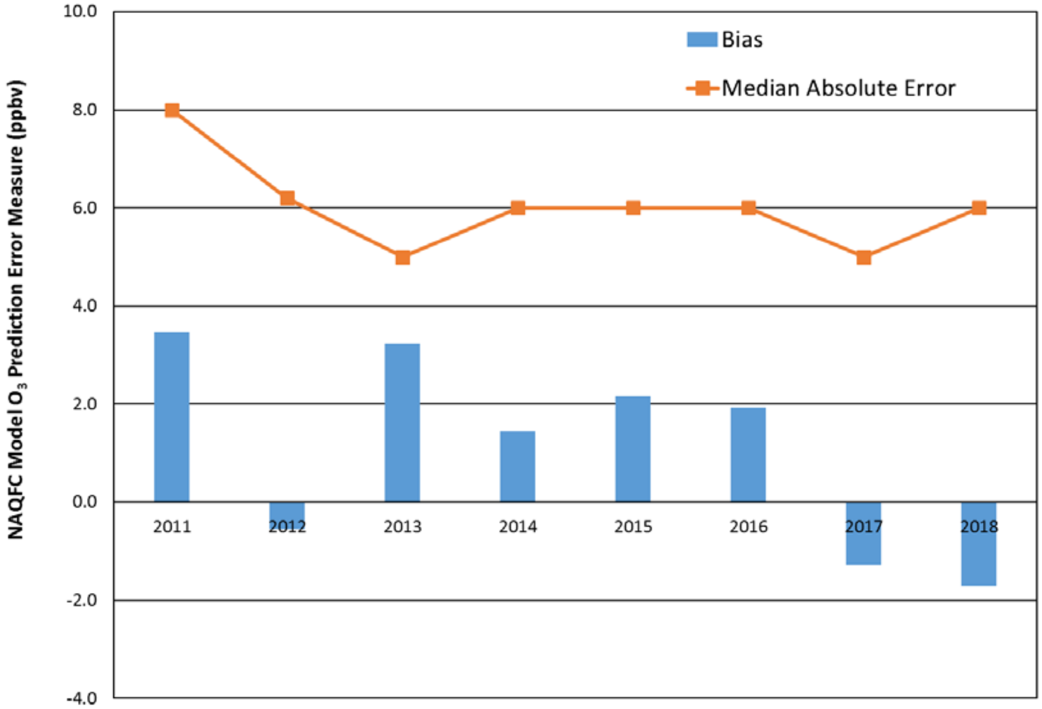

Figure 22. Error statistics for all maximum 8-hour average O_3 predictions by the NAQFC model, May 1 to September 30, in Delaware for 2011-2018.

Appendix A. Skill Measures for Threshold Forecasts

The determination of the skill of threshold forecasts (e.g., O_3 exceedances) begins with the creation of a 2x2 contingency table of the form of Figure 21. For example, if an O_3 exceedance is both observed and forecasted ("hit"), then one unit is added to "a." If an O_3 exceedance is forecasted but not observed ("false alarm"), then one unit is added to "b."

		Observed	
		Yes	No
Forecasted	Yes	a	b
	No	c	d

Figure 23. 2 x 2 contingency table for threshold forecasts.

Basic Skill Measures

Two common scalar skill measures are calculated using the values from the 2x2 contingency table: false alarm ratio, formerly termed false alarm rate, and hit rate.

False alarm ratio (FAR) is the fraction of threshold forecasts which are made but do not verify.

$$FAR = \frac{b}{a+b} \qquad (1)$$

Hit rate (H), also called probability of detection (POD), is the ratio of correct threshold forecasts to the total number of O_3 exceedances that were observed.

$$H = \frac{a}{a+c} \qquad (2)$$

Reference

Wilks, D. S., *Statistical Methods in the Atmospheric Sciences*, Third Edition, Academic Press, 676 pp., 2011.