8.1

THE EFFECT OF RECENT REGIONAL EMISSIONS REDUCTIONS ON AIR QUALITY FORECASTING IN THE MID-ATLANTIC

William F. Ryan
Jeannine E. Bradbury
The Pennsylvania State University, University Park, Pennsylvania

1. INTRODUCTION

Significant reductions in emissions of oxides of nitrogen, a precursor of both ozone (O_3) and fine particles ($PM_{2.5}$), from large power generation facilities occurred in the eastern United States beginning in the early 2000's as a result of the so-called "NO_x SIP Call" rule. This reduction has lowered regional O_3 concentrations and may also have lowered $PM_{2.5}$ concentrations although the limited observation network for the latter prior to 2003 limits any conclusions in this regard. These secular changes in emissions can have a significant impact on air quality forecast skill. While forecast guidance from operational coupled weather and chemistry models have been available since the mid-2000's, the skill of these models vary by location. The complexity of the $PM_{2.5}$ formation process further limits the skill of numerical model guidance. As a result, many forecasters continue to rely on empirical and statistical methods that require a "training" database of years of historical data. These models are expected to lose skill as the emission base changes over time. In this paper, we investigate changes in O_3 concentrations in the mid-Atlantic and the effect of these changes on the skill of statistical and numerical forecast models in the Philadelphia area. Because numerical models are able to adjust, year to year, to changes in power plant NO_x emissions, their forecast skill is found to be less affected by secular changes in O_3 concentrations. However, for forecasts in the high end of the O_3 distribution, numerical model output requires additional post-processing to adjust to land-water boundary effects, day of week effects and seasonal drift in bias.

2. DATA AND METHODS

This paper analyzes forecast guidance from the National Air Quality Forecast Capability (NAQC) O_3 operational forecast model (Otte et al., 2005; and, see, http://weather.gov/aq). The forecast metric of interest is domain wide peak 8-hour average O_3. The forecast guidance data was extracted in real time from model grid cells closest to the ~ 15-20 O_3 monitors in the Philadelphia (PHL) forecast area and the domain maximum determined from this set. Because of the strong forecast O_3 gradient near land-sea boundaries, using a "true" domain maximum, extracted from all locations within the domain, resulted in an unacceptable number of false alarms of high O_3 (Ryan, 2010).

The period May-September, 2007-2010 is used for the bulk of the analysis. The NAQC model has undergone a number of changes during that period. The primary changes are yearly updates to model emissions with significant changes in electrical generating unit (EGU) NO_x emissions. A full description of the yearly changes to the NAQC can be found at the EMC Change Log (NCEP, 2010).

Observed O_3 data for the PHL forecast area was extracted from the EPA AQS archive for the period 1990-2009 with 2010 data taken from the AirNowTech archive operated in conjunction with EPA AirNow (http://www.airnow.gov). As with model data, O_3 concentrations are expressed as 8-hour averages. This is in concordance with the National Ambient Air Quality Standard (NAAQS) for O_3 which is used as the threshold for forecast air quality warnings.

3. CHANGES IN EMISSIONS AND REGIONAL OZONE CONCENTRATIONS

This study assesses the impact of recent NO_x emissions reductions on O_3 forecasting. These reductions followed EPA regulations usually referred to as the "NO_x SIP Call" promulgated under Section 110 of the Clean Air Act. The SIP Call regulations required 22 states in the eastern US, along with the District of Columbia, to submit State Implementation Plans (SIPs) that included NO_x emissions reductions in order to mitigate ozone transport in the eastern US. In order to achieve the reduction targets, most states accepted the EPA voluntary Model Rule under which EGUs, and large industrial boilers, were controlled under an EPA-administered emissions trading program. The Model Rule

* *Corresponding author address*: William F. Ryan, Pennsylvania State University, Department of Meteorology, 401 Walker Building, University Park, PA, 16802, email: wfr1@psu.edu.

established state wide caps on total NO_x emissions during the O_3 season (May-September) as well as an allowance trading system. Sources of NO_x could reduce emissions directly, typically through fuel switching or application of new controls, or acquire allowances from other states under the NO_x budget trading program. Beginning in 2003, and continuing until the current date, significant NO_x emissions reductions have occurred with the greatest reductions west of the mid-Atlantic region (EPA, 2010; Bloomer, et al., 2009). While overall motor vehicle and industrial NO_x emissions have decreased only slightly over the past decade, power plant NO_x has decreased approximately 43% as of 2006 with continuing decreases occurring since that date. Through 2007, hourly O_3 concentrations have decreased in response to the NO_x reductions by approximately 10% in the eastern U.S. with larger decreases occurring in the highest concentrations ranges with decreases in 95th percentile O_3 of 16 ppbv (Bloomer et al., 2010). For air quality forecasting, changes in upper end of the O_3 distribution is of most interest due to its affect on the frequency of public health warnings.

4. CHANGES IN OZONE CONCENTRATIONS IN THE PHILADELPHIA METROPOLITAN AREA

The Philadelphia metropolitan area, as defined for this study, consists of a network of 16 monitors in southeastern PA, southern NJ and extreme northern DE (Figure 1). While there have been NO_x emissions reductions in the PHL area (EPA, 2010), the most significant reductions have occurred further west – the predominant upwind direction in the O_3 season. The impact on regional O_3 in the mid-Atlantic from these reductions can be seen at the regional scale O_3 monitor at Big Meadows in Shenandoah National Park, Virginia (38.55ºN, 78.45ºW). The monitor at Big Meadows is located near a ridge top at approximately 1000 m above sea level and has been in place since the late 1980's (Poulia, O., et al, 1991). Its elevation and remote location make it well situated to assess regional scale O_3 changes. Comparing the period 1990-2002 with 2003-2010, the frequency of days with 8-hour peak O_3 in excess of 70 ppbv has been reduced by 67% (from 38 days per season to 12 days). It is also notable that the extremely warm summer of 2010 resulted in only 13 days above 70 ppbv.

The reduction in regional O_3 concentrations carries over to the PHL metropolitan area – particularly when the higher end of the O_3 distribution is analyzed. For the period 1993-2002, compared to 2003-2010, the frequency of Code Orange (≥ 76 ppbv 8-hour average) cases is reduced by 43% and the frequency of Code Red (≥ 96 ppbv) cases is reduced by 76%. At the 75th percentile level, the reduction in peak O_3 is 14%. This reduction is steady across the temperature range. For cases with maximum temperature (T_{max}) in excess of 80ºF, peak O_3 concentrations in PHL, binned by ranges of T_{max}, show a decrease on the order of 10 ppbv. In addition, hot days in PHL are now much less likely to also be high O_3 days. For 1993-2002, 50% of "hot" days (≥ 90ºF) reached the Code Red threshold. For 2003-2010, this percentage is reduced to 16%.

5. IMPACT ON FORECASTS

Traditional forecast methods in the mid-Atlantic region have relied on the strong association of T_{max} with O_3 as a basis for stand-alone statistical forecast models using a mix of meteorological, climatic and chemistry predictors (see, e.g., Ryan, 2000). For the 1993-2002 period, for example, the explained variance of peak 8-hour O_3 using T_{max} in a single predictor regression model was 0.58 (r^2 value). Statistical methods, utilizing multiple linear regression (MLR), neural network and Classification and Regression Tree (CART) methods, were reasonably successful, particularly for forecasts of the occurrence of high O_3. The T_{max}-O_3 relationship appears to be fraying in the post-NO_x SIP Call environment. For the period 2003-2010, r^2 for the single predictor (T_{max}) model was reduced from 0.58 to 0.44, a 24% reduction. Of more concern to forecasters, the bias of the statistical model in use in PHL increased steadily from +1 ppbv in 2003 to +9.5 ppbv in 2009 (Figure 2). A portion of this increase is related to the use of a MLR model trained on pre-NO_x SIP Call data (1993-2001) but even when the MLR model was updated to use a training set from 2003-2007, the over prediction bias in 2009-2010 was still +7 ppbv. The impact of the steady decline in skill of the statistical guidance model can be seen in the official forecasts, which is informed by statistical model guidance, where bias also increased, though at a slower rate, through 2006.

As seen in Figure 2, official forecast bias error improved beginning in 2007. This improvement was driven primarily by the increasing use of numerical O_3 forecast model guidance as part of the standard forecast routine. The NAQC model was available beginning in 2004 and became operational until 2007. Due to a variety of issues related to emissions and other sub modules, the results from the NAQC did not become "steady" enough for routine forecast guidance use in PHL until the 2007 O_3 season. The advantage the NAQC model has over traditional statistical models in the changing chemical environment since 2002 is that it is updated yearly with respect to NO_x emissions from EGUs. The NAQC guidance, in terms of overall median absolute error, improved steadily compared to the statistical guidance from a 9% improvement in 2007 to 39% by 2010. In fact, for the 2010 forecast season, skill scores for the NAQC exceeded even the official (expert) forecast skill scores for the first time ever – though the NAQC does continue to have a slightly higher false alarm rate.

At this stage, it is clear that the continued usefulness of stand-alone statistical models is highly doubtful, and the use of numerical O_3 forecast model guidance is preferred for operational forecasting. However, there remain a number of known difficulties with the NAQC model, some of which are shared with other numerical models, which can limit its usefulness. It is to be emphasized that numerical model performance varies across the eastern US. Locations in the SE US, for example, experience different results than those in the mid-Atlantic. For the mid-Atlantic region, using PHL as an example, there are three major challenges using the NAQC forecast model guidance to inform operational forecasts. First, the presence of strong forecast O_3 gradients near land-sea boundaries – particularly when high O_3 is forecast; (2) the manner with which the model handles the recently emerging weekday and weekend differences in peak O_3 concentrations; and, (3) seasonal drift in forecast bias.

For a variety of reasons, including rapid changes in the vertical structure of the modeled planetary boundary layer, as well as differences in surface deposition and emissions rates, numerical O_3 models tend to develop strong gradients in O_3 concentrations near land-sea boundaries. This is particularly evident near embayments such as the Chesapeake and Delaware Bays as well as in Long Island Sound (Figure 3). These effects can be significant at the Code Orange forecast threshold (Ryan, 2010). The effect of these gradients varies with the manner in which forecasters extract peak O_3 concentrations from the model. For the NAQC, automated techniques, based on zip codes, are available to extract peak O_3 from all land areas within a metropolitan forecast area via the AirNow data clearinghouse service, AirNowTech. For the PHL forecast area, using this technique results in a large number of forecast false alarms of Code Orange O_3, primarily along the southern NJ coast. An alternative method to extract peak domain O_3 information from the NAQC is to extract forecasts only at the grid cells closest to the location of existing regulatory O_3 monitors. This retrieval can be quite easily automated using standard web browsers. Monitors in PHL are situated well away from the immediate coast and forecasts for these locations are less influenced by the land-sea effect. The result of this simple post-processing technique is quite striking. In 2010, for example, model forecasted false alarms of Code Orange O_3 in PHL were reduced by greater than 50% using only the monitor specific forecasts. This resulted in an overall model false alarm rate of 23%, coupled with an 86% hit rate (Figure 4). Application of this technique was not as successful in forecast areas more strongly impacted by bay-land boundaries, such as Baltimore and Connecticut, but further steps based on this approach did show promise (Ryan, 2010).

One of the more interesting changes brought about by recent regional emissions reductions is the emergence of day of week differences in peak O_3 in the PHL area. Prior to the NO_x SIP Call reductions, there was little day of week variation in peak O_3. At the 75^{th} percentile, for example, the difference in day of week peak O_3 was only 4 ppbv, with Saturday the highest O_3 day on average. From 2003-2010, however, the difference from the highest day (Wednesday), to the lowest (Sunday) was 10 ppbv with weekend concentrations lower than weekday concentrations (Figure 5). This emerging day of week difference has posed a problem for forecast skill in the air quality warning range (≥ 76 ppbv). For the period 2003-2010, official forecasts had a weekday false alarm rate of 21% compared to a weekend false alarm rate of 42%. This problem is also seen in the numerical model forecast. For the period 2007-2010, the weekend forecast bias is more than double the weekday bias (5.9 ppbv compared to 2.6 ppbv). Interestingly, the Code Orange false alarm rate for the NAQC is only significantly higher on Sunday (0.70 compared to 0.30 for the remainder of the week). At this point, due to yearly changes in the NAQC and its sub models, there is no specific correction that can be applied. In practice, however, forecasters have become very skeptical of high O_3 model forecasts for Sunday in the PHL area.

The final challenge for air quality forecasters is the existence of a defined seasonal drift in NAQC guidance bias. This has been noted in regions other than the mid-Atlantic (see, e.g., Byun, et al., 2009). For the PHL area, for the period 2007-2010, NAQC peak O_3 forecasts are essentially unbiased through early July and then develop an over-prediction bias for the remainder of the forecast season (Figure 6). The causes of this bias have not been determined.

6. DISCUSSION AND CONCLUSIONS

Significant changes in O_3 concentrations have occurred in the eastern US following NO_x emissions reductions at EGU's in the early 2000's. The impact on air quality forecasting has also been significant. In the PHL forecast area, the utility of forecast guidance from stand-alone statistical forecast models for peak O_3 has been substantially reduced. An over-prediction bias, beginning in 2003, has steadily increased through 2010. In addition, continuing changes in emissions, due to new regulations as well as fuel switching and equipment upgrades, make it unlikely that a stable historical O_3

data base sufficient to train statistical models will be available in the near future. Forecasters must now rely on numerical forecast model guidance. For the PHL area, using the NAQC forecast model, updated yearly for emissions changes from EGUs, has resulted in good skill scores. There are, however, several ongoing challenges using the NAQC forecast guidance. First, some degree of post-processing is needed to reduce the effect of strong land-sea boundary O_3 gradients. Second, it appears that there are increasing differences in day of week O_3 concentrations in the PHL forecast area that can lead to excessive false alarms of high ozone on Sundays. Finally, the NAQC continues to show a seasonal drift in forecast guidance making forecasts less skillful later in the season. The land-sea problem is easily solved for the PHL area while the remaining issues require additional research before quantitative methods can be utilized to reduce their effect.

ACKNOWLEDGEMENTS

The authors acknowledge the support of the Delaware Valley Regional Planning Commission and the States of Delaware and New Jersey, as well as the Commonwealth of Pennsylvania, for this research as well as operational air quality forecasting in Philadelphia and the State of Delaware. The support of the Air and Radiation Management Administration of the Department of the Environment of the State of Maryland (David Krask, Michael Woodman and Duc Nguyen) is also gratefully acknowledged.

REFERENCES

Bloomer, B. J., et al., 2009, Observed relationships of ozone air pollution with temperature and emissions, *Geophys. Res. Letters*, **36**, L09893, doi: 10.1029/2009GL037308.

Bloomer, B.J., K. Vinnikov, and R.R. Dickerson, 2010, Changes in seasonal and diurnal cycles of ozone and temperature in the eastern U.S., Atmos. Environ., 2543-2551.

Byun, D., et al., 2009: NOAA National Air Quality Forecast Guidance for Ozone and Particulate Matter, http://airquality.ucdavis.edu/pages/events /2009/iama/Byun.pdf

EPA, 2010, Clean Air Markets, http://www.epa.gov/airmarkets/progress/interactive mapping.html.

NCEP, 2010: Operational Air Quality Forecast Change Log, http://www.emc.ncep.noaa.gov/mmb /aq/AQChangelog.html.

Otte, T. L., et al., 2005, Linking the Eta model with the Community Multiscale Air Quality (CMAQ) modeling system to build a national air quality forecasting system, *Wea. Forecasting*, **20**, 367-384.

Poulida, O., et al, 1991, Trace gas concentrations and meteorology in rural Virginia: 1. Ozone and carbon monoxide, *J. Geophys. Res*., **96,** 22,461-22,475. Doi: 1029/91JD02353.

Ryan, W. F., C. A. Piety and E. D. Luebehusen, 2000, Air quality forecasting in the mid-Atlantic region: Current practice and benchmark skill, *Wea. Forecasting*, **15**, 46-60.

Ryan, W. F., 2010, Post-processing of numerical ozone model forecasts: The land-sea problem, Second International Workshop on Air Quality Forecasting, Quebec, http://www.meteo.psu.edu /~wfryan /quebec/iwaqf-2010.pptx

Stephenson, D. B., 2000: Use of the “odds ratio” for diagnosing forecast skill, *Wea. Forecasting*, **15**, 221-232.

Wilks, D. S., *Statistical Methods in the Atmospheric Sciences*, Academic Press, 467pp., 1995.

Figure 1. The Philadelphia metropolitan air quality forecast area (blue box) with the location of surface O_3 measurement monitors given by black triangles. Figure courtesy of AirNowTech (http://www.airnowtech.org/).

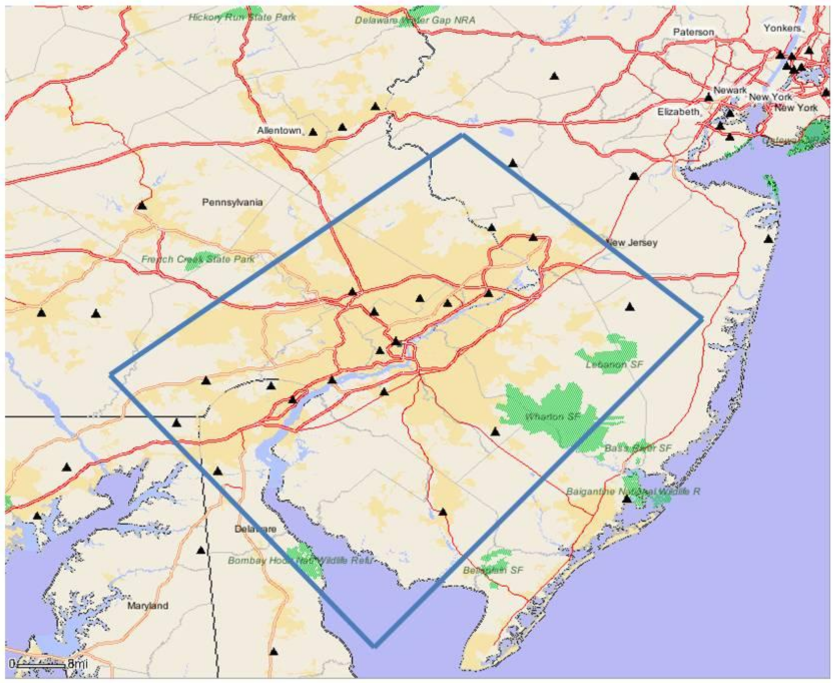

Figure 2. Yearly bias of the forecast guidance models for peak 8-hour O_3 in the PHL area. "Forecast" is the public, expert forecast, "NAQFS" is the post-processed guidance from the NOAA-EPA numerical model, as described in the text, "R0302" is a statistical model using training data from the pre-NO_x SIP Call era (1993-2002) and "R2009" is a similar statistical model using training data from the post-NO_x SIP Call era (2003-2010). All values given in ppbv.

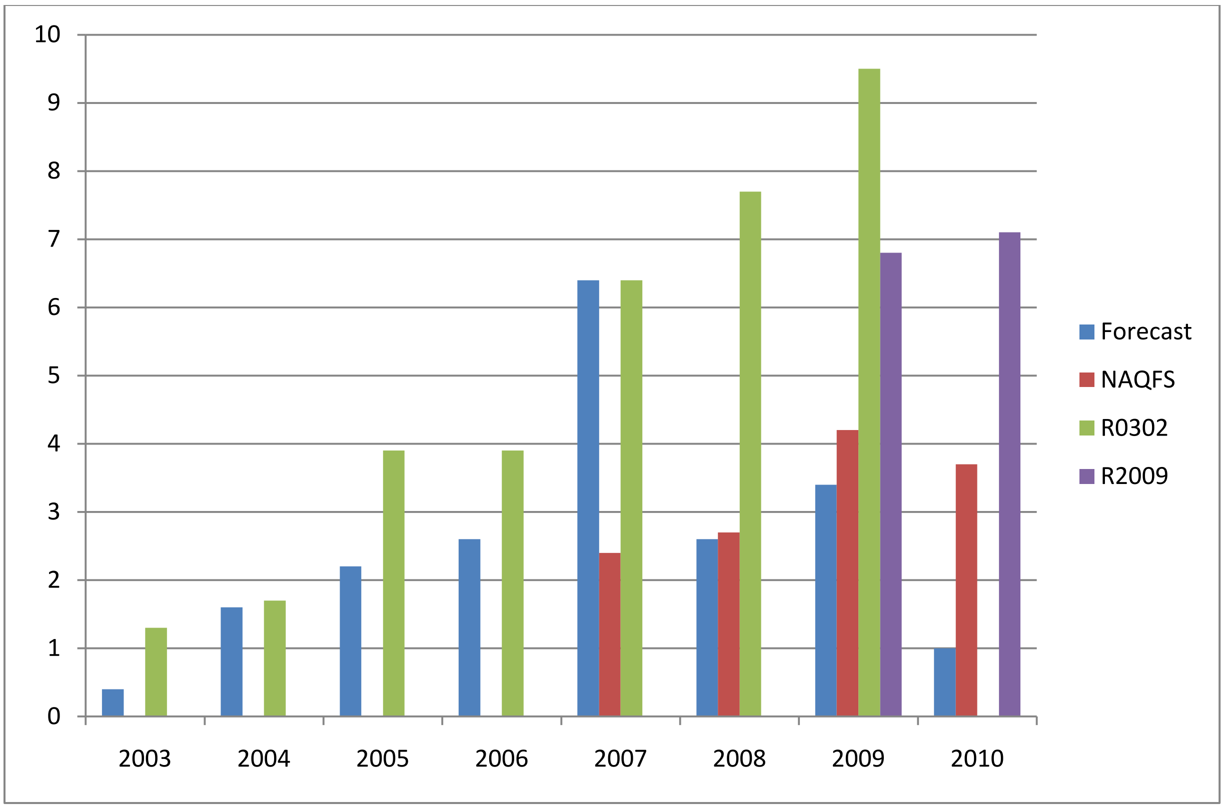

Figure 3. An example of land-sea forecast O_3 gradients in the mid-Atlantic. Peak 8-hour average O_3 forecast (in ppbv) for August 5, 2009 from the NAQC model initialized at 1200 UTC on August 4. Figure courtesy of NOAA EMC Mesoscale Modeling Branch. (http://www.emc.ncep.noaa.gov/mmb/aq/).

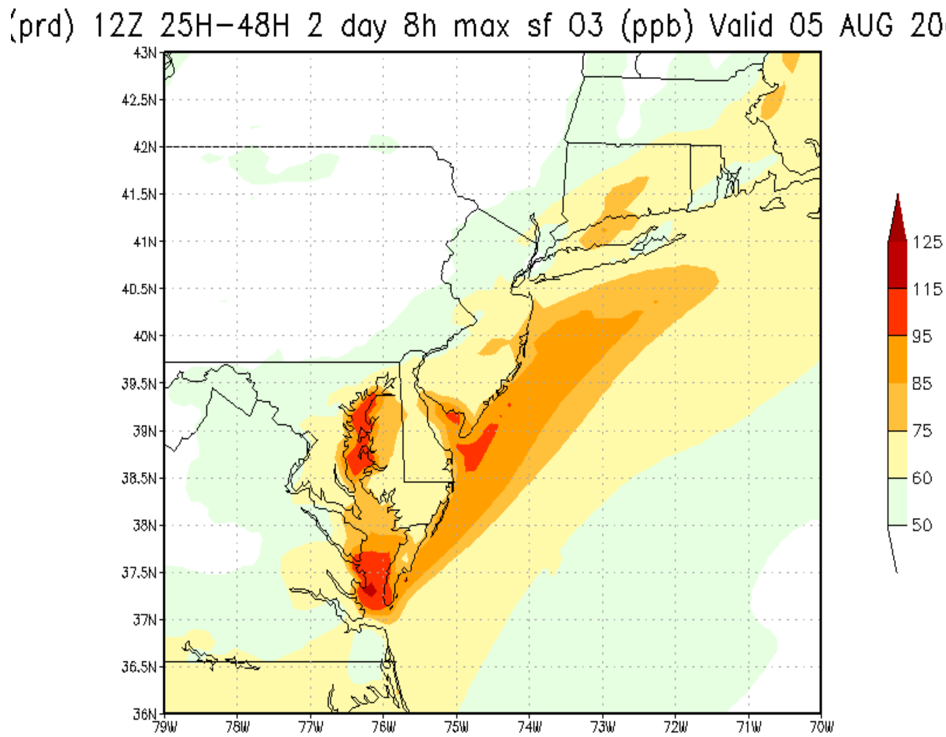

Figure 4. Skill measures for PHL O_3 forecast methods in 2010. “FC” refers to the public, expert forecast, “NAQC” the post-processed NOAA-EPA numerical model, as described in the text, and “R2009” refers to the statistical model using training data from the post-NO_x SIP Call era (2003-2010). All measures are given for forecasts of peak domain-wide 8-hour average O_3 concentrations. The skill score measures are based on a 2 x 2 contingency table using 76 ppbv (Code Orange) as the threshold (Stephenson, 2000; Wilks, 1995). “AE” refers to absolute, or unsigned, error. The “Threat” score is also known as the Critical Success Index (CSI). A summary of skill score measures can be found at: http://www.meteo.psu.edu/~wfryan/skill-score-appendix-2010.docx

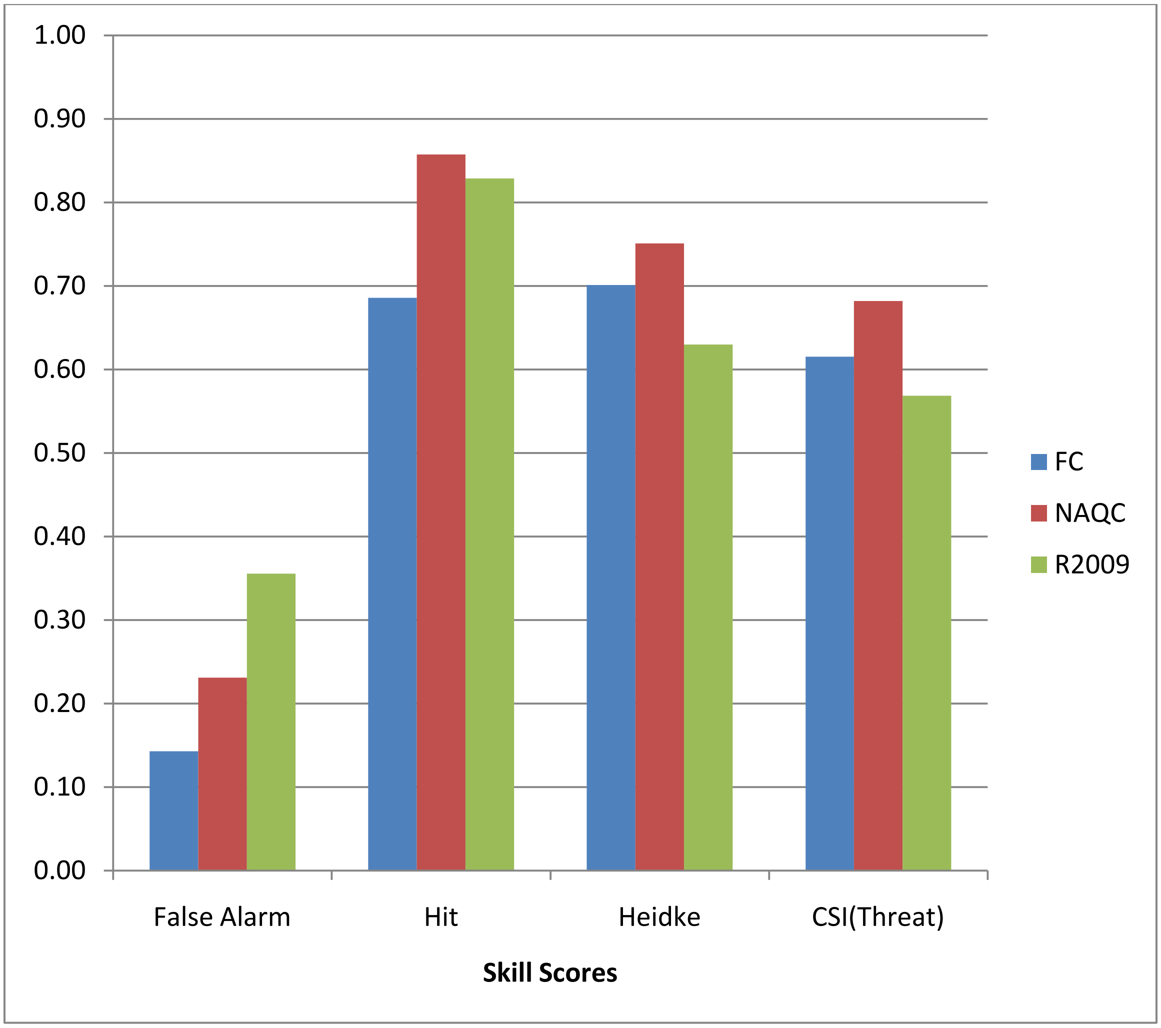

Figure 5. The 75th percentile of average daily maximum 8-hour O_3 for the PHL metropolitan area by day of week.

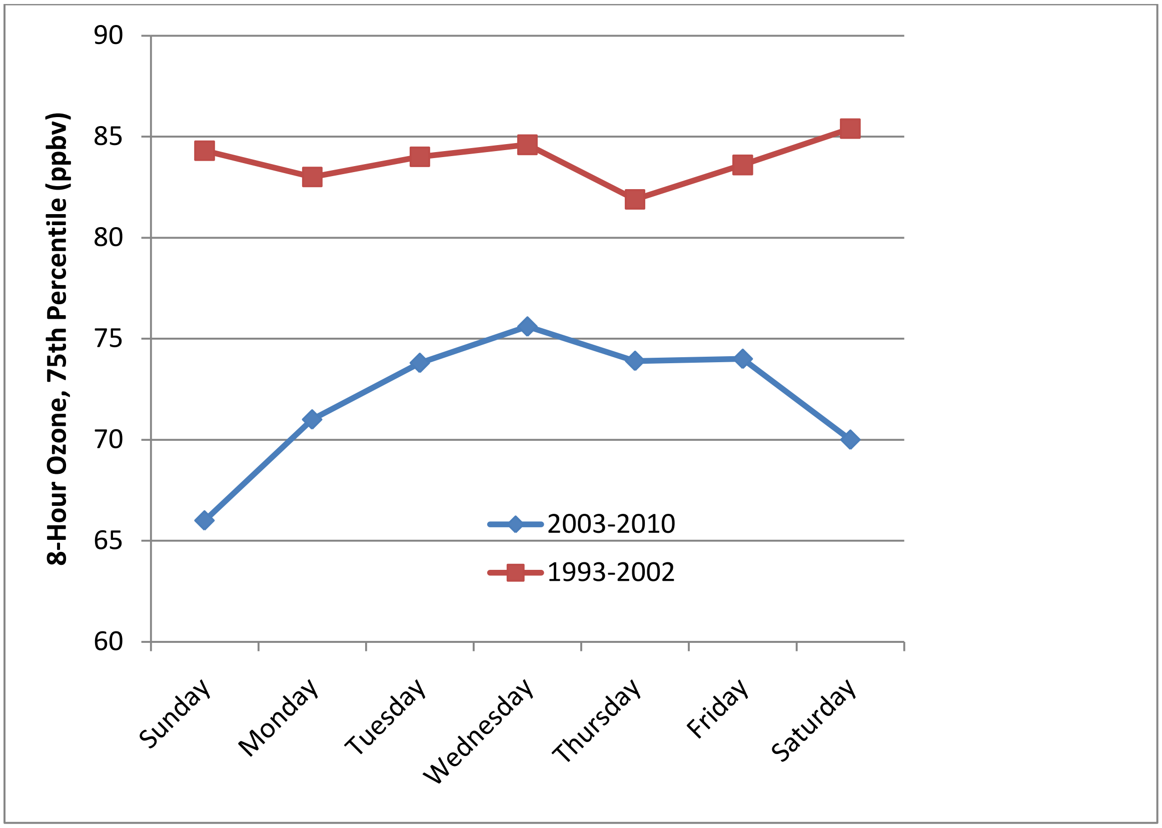

Figure 6. Time series of mean and median forecast bias for the NOAA-EPA model for the period May-September, 2007-2010.