Session 2 Part 1: Using Objective Tools in Daily Forecasts

How to Use Objective Forecast Tools in Conjunction with Subjective Tools to Produce an Air Quality Forecast

Bill Ryan

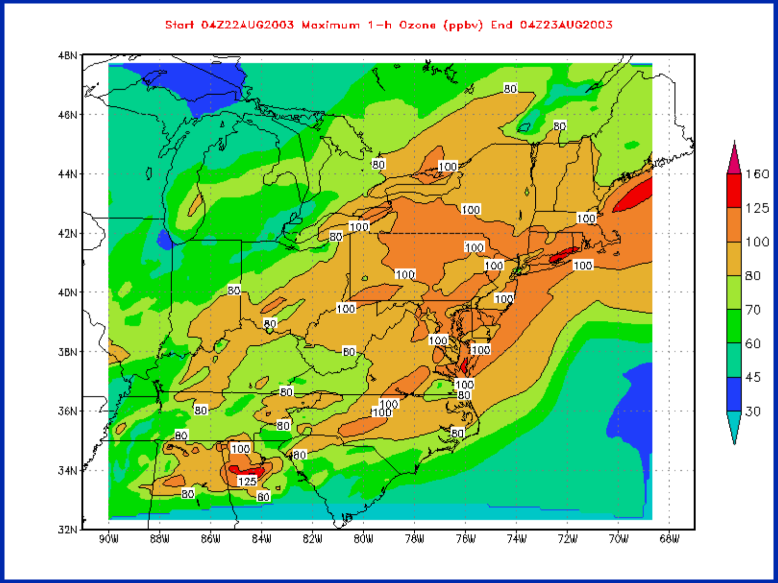
Pennsylvania State University

Outline

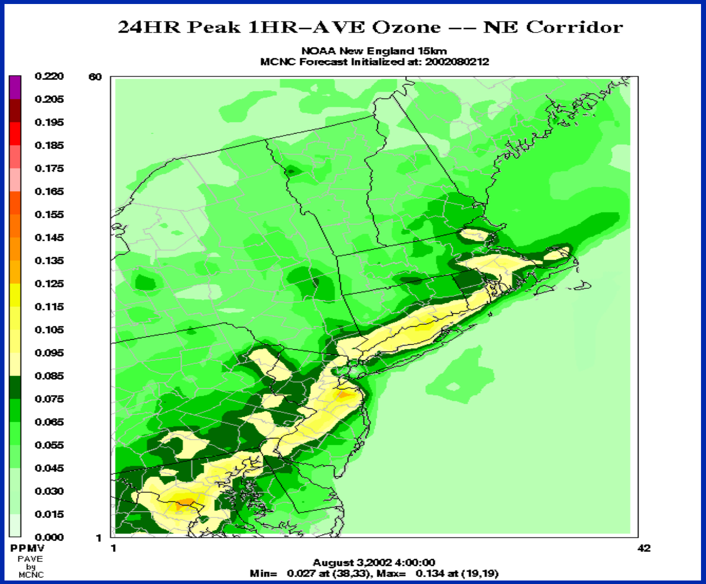
- Forecast Techniques
- Regression-based Statistical Models
- Archive of Forecast Data
- Selection of Input Predictors
- Use of Numerical Weather Models
- Analysis of Statistical Model Guidance
- Modification of Guidance for Final Forecast
- Example Case – Mid-Atlantic Ozone Episode

Forecast Techniques – Numerical Guidance

Numerical air quality forecast models are available and improving. In the short term, however, statistical models will remain most common.



NOAA Air Quality Forecast Model:
http://wwwt.emc.ncep.noaa.gov/mmb/aq/AQ.html



MAQSIP:
http://www.baronams.com/products/maqsip

Forecast Techniques – Statistical Forecast Methods

- Statistical methods currently in use include

  Regression

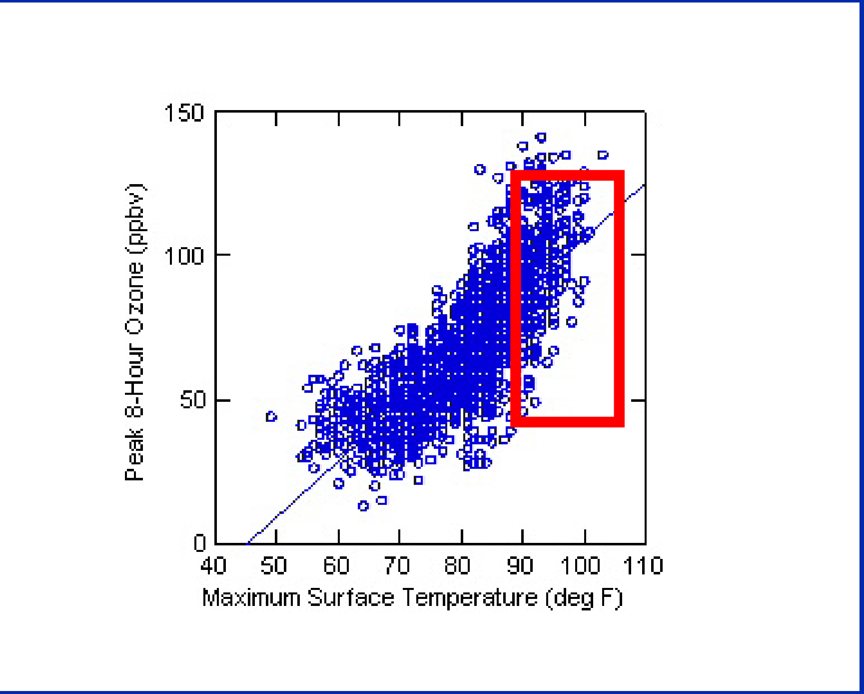
  CART

  Neural Networks

- These common methods have similar accuracy.
- Regression methods are most transparent to the user and will be used for the examples discussed here.

Regression-based Statistical Models (1 of 4)

- Observed air quality data (e.g. 8-hour O_3) is related to a forecast predictor.
- For linear regression, the relation is of the common form: y = mx + b.
- At right, maximum temperature (T_{max}) is a good predictor for peak O_3.
- Other predictors can be added to improve accuracy in the higher O_3 cases.



$$[O_3] = 1.92*T_{max} - 86.8$$

$$r = 0.77 \qquad r^2 = 0.59$$

Regression-based Statistical Models (2 of 4)

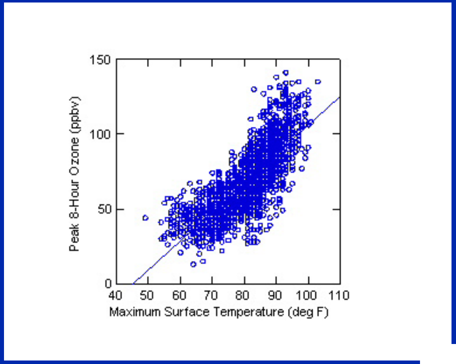
- More predictors can be added (“stepwise regression”) so that the equation looks like this:

  $$y = m_1x_1 + m_2x_2 + m_3x_3 + \ldots\ldots m_nx_n + b$$

- Each predictor (x_n) has its own “weight” (m_n) and the combination leads to better forecast accuracy.
- The mix of predictors varies from place to place. For mid-Atlantic O_3, the typical predictors include T_{max}, wind speed (surface and aloft), relative humidity (RH), day length (or solar zenith angle), previous day O_3, and upper-air T.

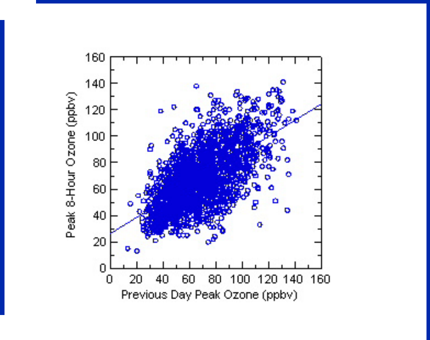
Regression-based Statistical Models (3 of 4)

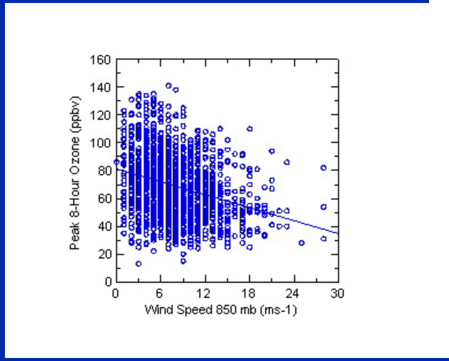
- The various predictors are not equally weighted, some are more important than others.
- It is essential to identify the strongest predictors and work hardest on getting those predictions right.
- Examples from Philadelphia are shown at right.



T_{max} vs. O_3

Previous day O_3 vs. O_3



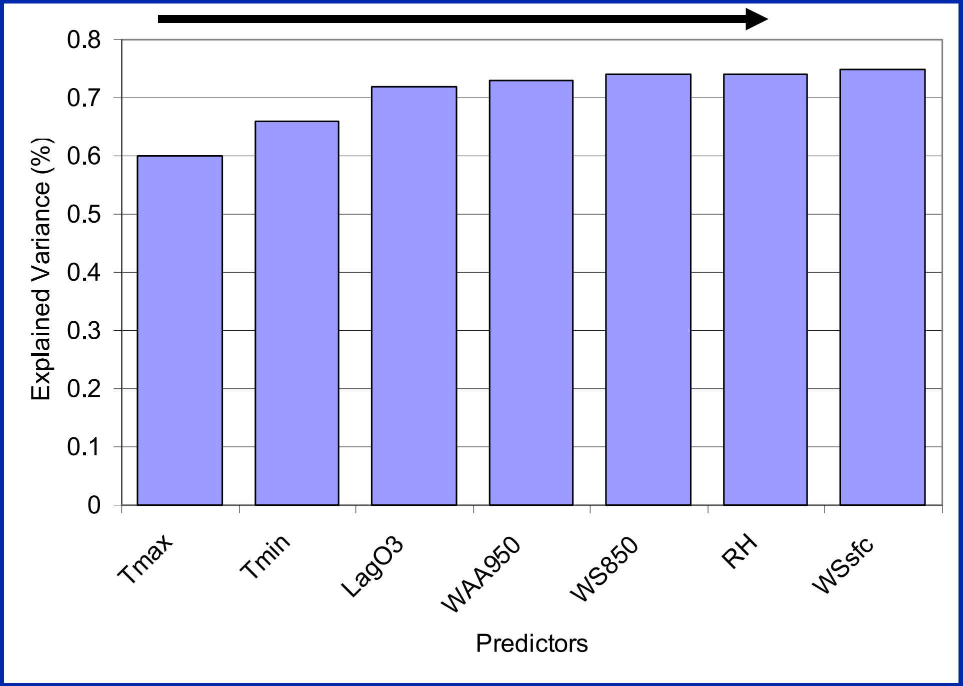


Wind Speed vs. O_3

Regression-based Statistical Models (4 of 4)

- In our example case, most of the variance in O_3 is explained by T_{max} (60%), with the additional predictors adding ~ 15%.
- Overall, 75% of the variance in observed O_3 is explained by the forecast model.
- Our job as forecasters is to fill in the additional 25% using the statistical model output as guidance.

Accumulated explained variance



Archive of Forecast Data (1 of 2)

- The MOST IMPORTANT job to ensure improvement in forecast accuracy is the maintenance of a good archive of forecast performance.
- The archive should include observed and forecast O_3 (or $PM_{2.5}$), critical meteorological data, and room for a quick comment. (AIRNowforecaster.org can help store information.)
- ALWAYS make a ppbv (or µg/m^{-3}) forecast even if only color codes are released to the public. Performance cannot be measured accurately otherwise.

Archive of Forecast Data (2 of 2)

- Air Quality Data: http://www.epa.gov/airnow/2004/
- Surface Hourly Observations:
  http://www.uswx.com/us/stn/Kxxx/

  (where xxx = station ID)
- Upper-Air Data: http://raob.fsl.noaa.gov/
- All kinds of useful data are archived at

  http://www.arl.noaa.gov/ready/ametus.html

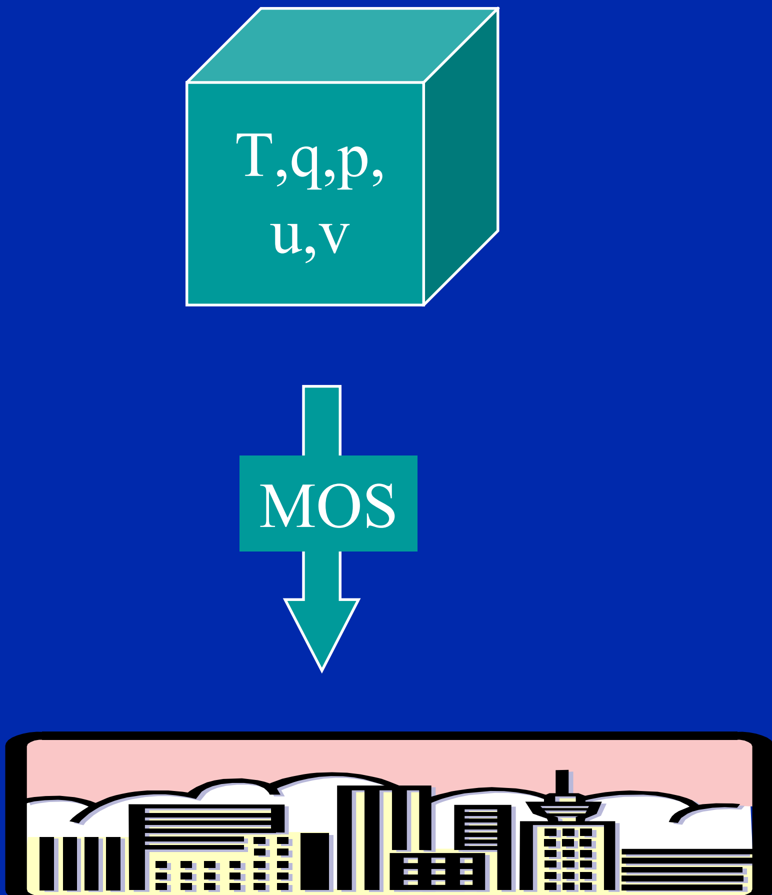
Selection of Input Predictors (1 of 5)

- The first step for good statistical forecast guidance is to select accurate inputs to the model.
- Some predictors use observed data (e.g., solar zenith angle) but most require forecast data.
- Where do you obtain forecast data? Discussed in next session.

Selection of Input Predictors (2 of 5)

- Model output statistics (MOS) relates output from numerical weather models (T, p, moisture, wind) to sensible weather observed at the surface.
- Each forecast model (Eta, GFS, NGM) has its own MOS or set of equations that relate model output to surface conditions.
- MOS can be used as input to statistical models for surface meteorological data.

Selection of Input Predictors (3 of 5)



Weather forecast models provide a small set of weather variables for each grid cell in the model.

MOS takes the model data and relates it to sensible weather at the surface – e.g., surface winds, cloud cover, chance of rain, using regression techniques.

Selection of Input Predictors (4 of 5)

- We can use MOS forecasts to provide input to statistical air quality forecast models (e.g., T_{max}) :

  http://www.nws.noaa.gov/mdl/synop/products.shtml

  http://www.nws.noaa.gov/om/tpb/486body.html
- MOS forecasts are only as good as the model predictions that underlie them.
- The new NWS digital forecasts mimic MOS formats so that you can compare local forecasters to the MOS:

  e.g., http://www.srh.noaa.gov/data/PHI/AFMPHI

Selection of Input Predictors (5 of 5)

- For upper-air data to input to the air quality statistical model, use output directly from the forecast models.
- Forecast model soundings are the easiest way to access the data:
  http://www.arl.noaa.gov/ready.html
- The BUFKIT software is also extremely useful for visualization:
  http://www.wdtb.noaa.gov/resources/projects/BUFKIT

Use of Numerical Weather Models (1 of 6)

- There are a number of forecast models available each day – how many should you look at?
- In operational time limits, you are usually limited to two models – which models are worth your time?
- The choice of models depends on a variety of factors – how does the model handle key issues for air quality forecasts?
- This is a complex subject – perhaps another Short Course?

Use of Numerical Weather Models (2 of 6)
Description of Models

- Description of the main numerical weather models:

  http://meted.ucar.edu/nwp/pcu2/

- The Eta is reliable and has good resolution. However, it is still evolving. This evolution can affect critical air quality parameters and MOS output. Model change bulletin discussing the July 2003 changes can be found at:

  http://meted.ucar.edu/nwp/pcu2/EtaMay2003update.htm

Use of Numerical Weather Models (3 of 6)
Interpreting the Model

- Consensus is often the best forecast, but what if there is no consensus among models (e.g., timing of frontal passage)?
- Five things to consider

    NCEP diagnostic discussion

    Local NWS forecast office discussion

    Run-to-Run consistency

    Trend in forecast (d{prog}/dt)

    Short range (6-hr) forecast accuracy

Use of Numerical Weather Models (4 of 6)
Picking the Right Model

- NCEP Diagnostic Discussion:

  http://www.hpc.ncep.noaa.gov/html/model2.shtml

- The diagnostic discussion tells you which model is expected to perform best in certain regions.
- Local NWS Forecast Office Discussion: May not always comment on model performance. The best discussion is found in the 0700 UTC discussion:

  http://weather.gov/
  (click on location, find discussion)

Use of Numerical Weather Models (5 of 6)
Picking the Right Model

- Run-to-Run consistency – Has the model solution changed much over the past 12-24 hours? A consistent forecast is often a good forecast.
- d{prog}/dt – How are the models trending? E.g., have they slowed the front? The expectation is that the trend will continue.
- Short range forecast – Our forecasts are often made ~ 1800 UTC, so the model has run 6 hours. Is it accurate so far? E.g., is rain well-forecast?

Use of Numerical Weather Models (6 of 6)

- For many air quality forecasters, time is always in short supply. Forecast “triage” directs attention to those inputs of most weight in the statistical model (e.g., T_{max}) and of most concern to the key air quality issue of the day (more details later).
- A quick morning look at the 0700 UTC NWS forecast discussions and the 0000 UTC run is suggested.

Analysis of Statistical Guidance (1 of 8)

- Use a worksheet to collate forecast data from meteorological models.
- Select statistical model inputs from the worksheet.
- Run model and produce guidance forecast.
- Now what?

2003 Worksheet: Philadelphia Ozone Forecast Regression Model

Valid Date: _____________ , 2003 Forecaster:

Surface Variables

	Fcst	NGM	CCF	Eta	GFS NWS
T_{max} (°F)					
T_{min} (°F)					
RH_{19-21} (%)					
WS_{07-09} (kts)					
WS_{12-17} (kts)					

Upper Air Variables Height (mb) 12 UTC

	Fcst	Eta	GFS	MM5	
T_{950} (Current, °C)					
T_{950} (+24 h, °C)					
WS_{850} (ms^{-1})					

Other Variables

		Persistence	1300 EDT	
SZA (deg)		$1h\text{-}O_3$		
Julian Date		$8h\text{-}O_3$		

Regression Results

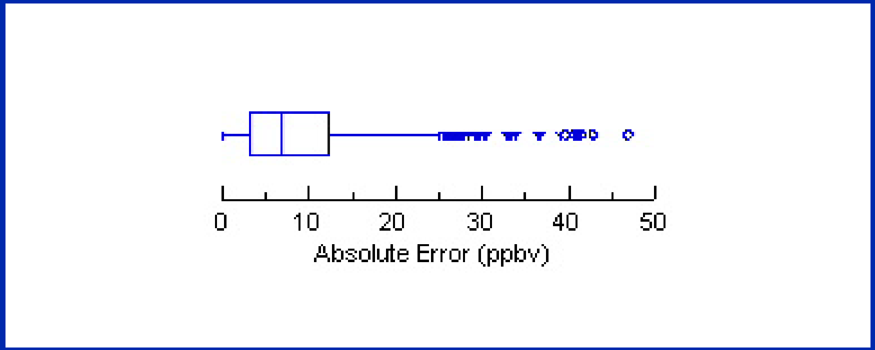
Run 01	
Run 02	
Run 03-01	
Run 03-02	
Ln 2001	

Forecast

ppbv	
Code	

Analysis of Statistical Guidance (2 of 8)
Envelope of Uncertainty

- What is the "envelope of uncertainty" for statistical model output? If you choose to modify the forecast, it should usually be within the envelope.



<u>PHL 8-hr O_3 example</u>
median error: 6.9 ppbv
mean error: 8.6 ppbv
bias: -0.5 ppbv

- This is an 8-hr O_3 forecast algorithm example. It shows the range of error from the training data set.

Analysis of Statistical Guidance (3 of 8)
Envelope of Uncertainty

- How do you determine the range of uncertainty of the statistical guidance?
  - Use the database that created the model.
  - Use a subset of data held out when the model was created (“cross validation”).
  - Use the previous year’s forecasts (archives come in handy here).
- Once you have this information – park it on your bulletin board, you’ll need it every day.

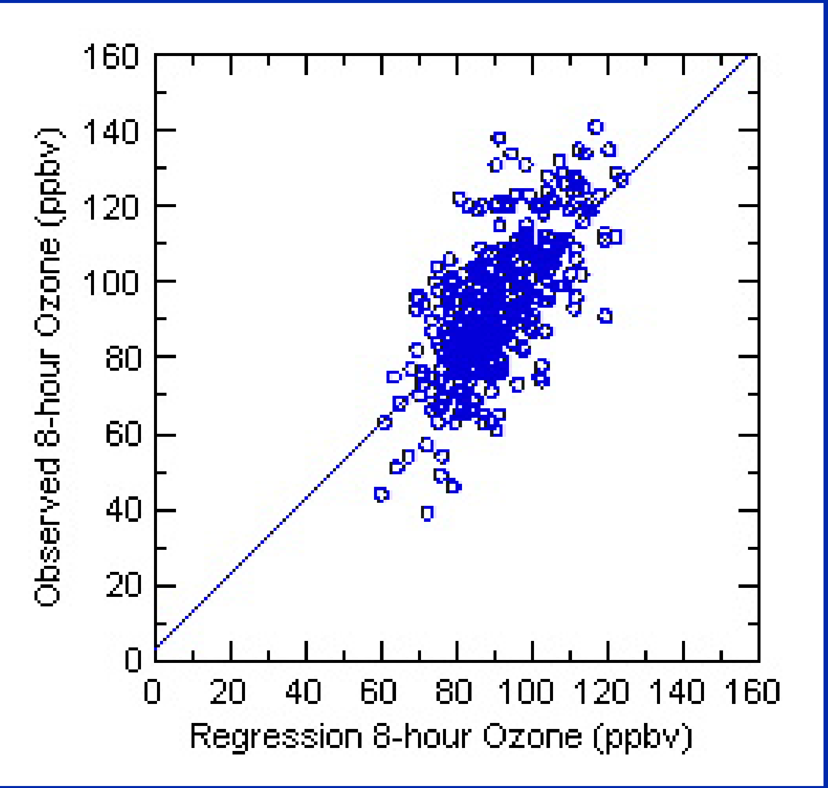
Analysis of Statistical Guidance (4 of 8)
Performance in Key Cases

- In addition to overall skill, you also need to know the range of skill in critical cases.
- For O_3, this often includes high O_3 and high temperature cases. For $PM_{2.5}$, this can include stagnation or strong inversion cases.
- You also need to know bias in categorical forecasts, e.g., does the model tend to issue a large number of false alarms?

Analysis of Statistical Guidance (5 of 8)
Example From Philadelphia

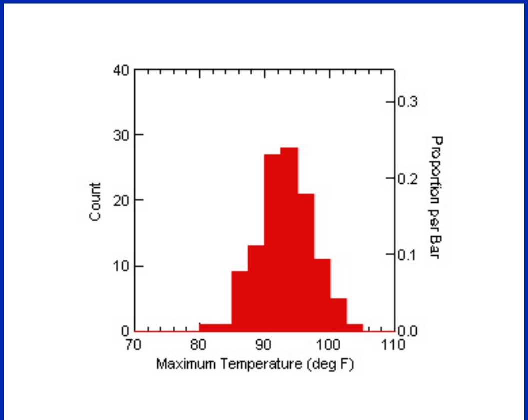
High temperature cases:

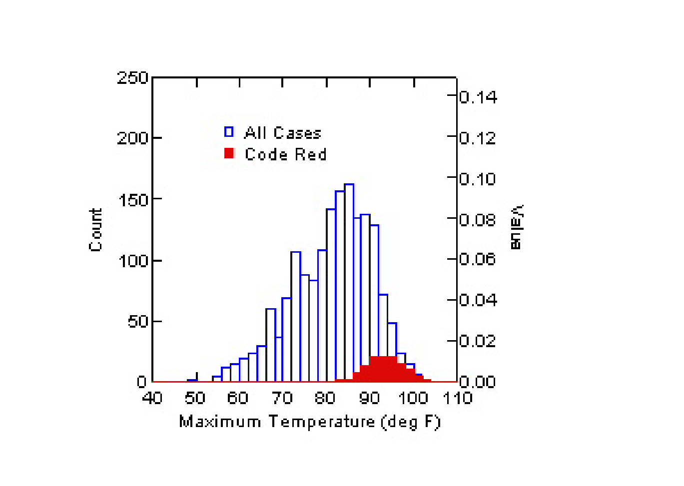
- Error is slightly higher than average (by ~2 ppbv)
- Bias is to underpredict by 3 ppbv, which is not much different from overall performance.



Analysis of Statistical Guidance (6 of 8)
Understanding the Strongest Predictor

- For O_3, temperature is a strong predictor.
- Hot weather is necessary (top panel, T_{max} for all ozone action days [OAD]) but not sufficient for Code Red (bottom panel, T_{max} for all cases).
- Know the likely frequency of codes when temperature is hot.





Analysis of Statistical Guidance (7 of 8) Color Code Performance

- Does statistical guidance tend to issue more false alarms or misses?
- For the PHL 8-hr forecast example:
  - Forecast Red and observe Red: 83%
  - Forecast > 85 ppbv and observe > 85 ppbv: 81%
  - Observed Red and forecast Red only: 30%
  - Observe > 85 ppbv and forecast > 85 ppbv: 58%

Analysis of Statistical Guidance (8 of 8)
Color Code Performance

- Results just shown – few false alarms of high O_3 but too many misses – are common with regression methods.
- This tells us the critical forecast question will be to resolve cases that are forecast Orange but observed Red.
- What additional information can help to distinguish these cases?

Modification of Statistical Guidance (1 of 7)
Forecast Critical Factors

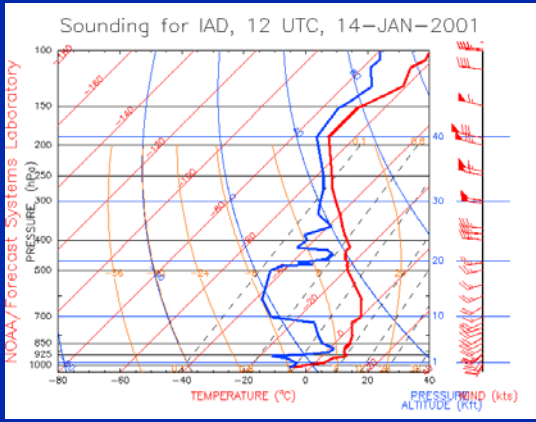
- Goal – to determine when to vary official forecast from statistical guidance.
  - What weather/other factors are critical to observed conditions?
  - How well does the statistical model handle them?
- These factors vary by location and by forecast parameter (e.g., O_3, $PM_{2.5}$). Some are only known through experience; others can be understood from basic principles or historical data.

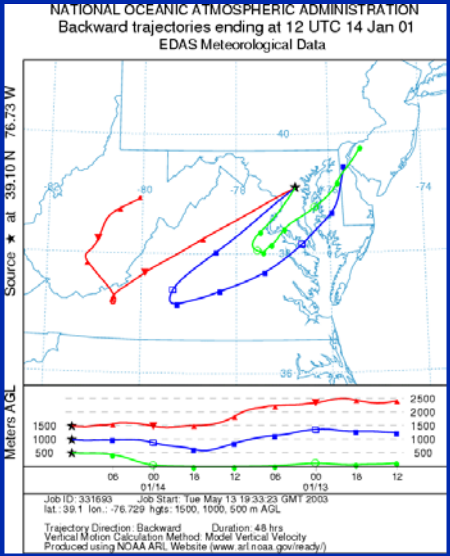
Modification of Statistical Guidance (2 of 7)
Critical Factors for Ozone

- How much sun: UV = photochemistry
- Volume of air mass: Inversion strength
- Boundary conditions: Transported O_3
- Emissions: Vary by day?
- Thunderstorms: Timing?
- Mesoscale effects: Bay/land breeze

Modification of Statistical Guidance (3 of 7)
Critical Factors for $PM_{2.5}$

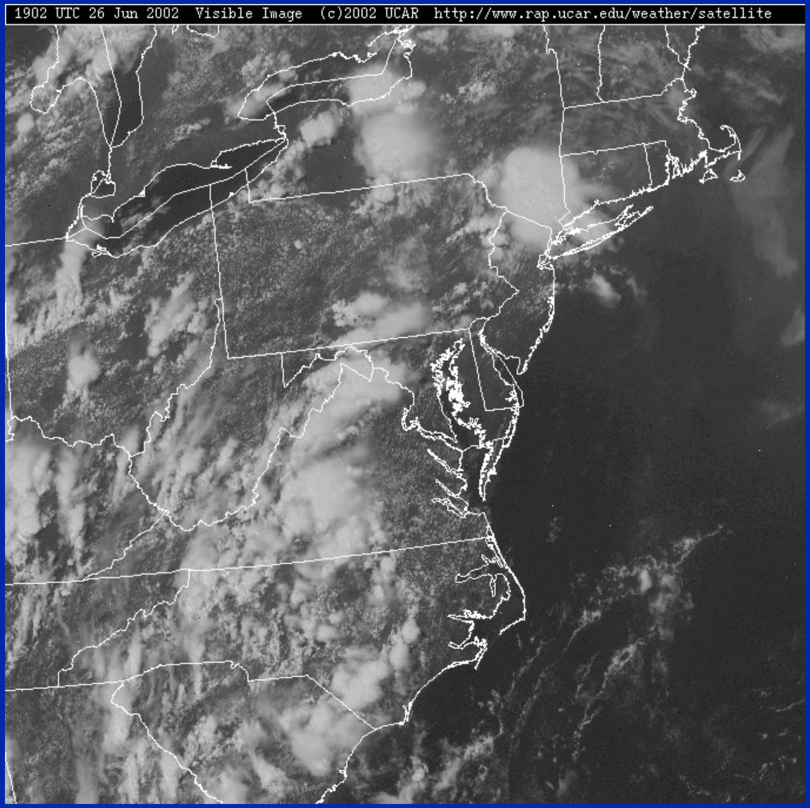
- Although with many of the same parameters as O_3, $PM_{2.5}$ is often more sensitive to stagnation and the strength of the low level inversion.
- Right: Examples from high $PM_{2.5}$ event on January 14, 2001.





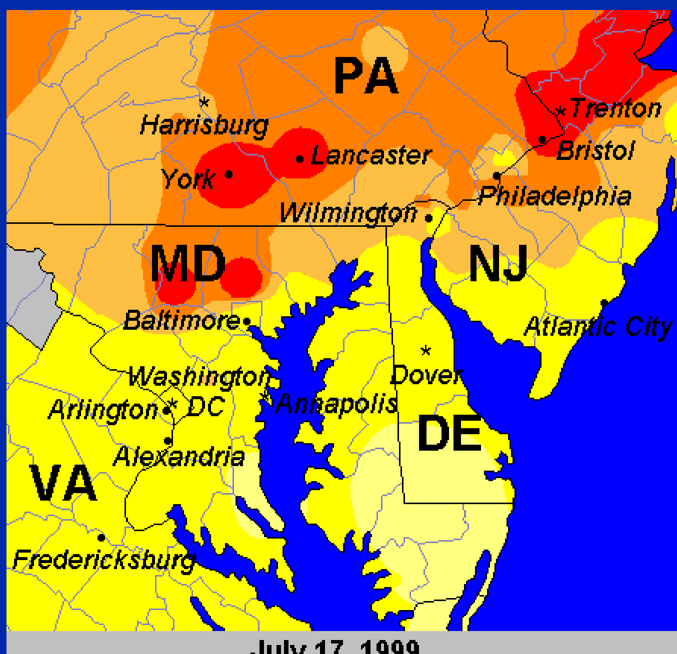
Modification of Statistical Guidance (4 of 7)
How Are Critical Factors Handled?

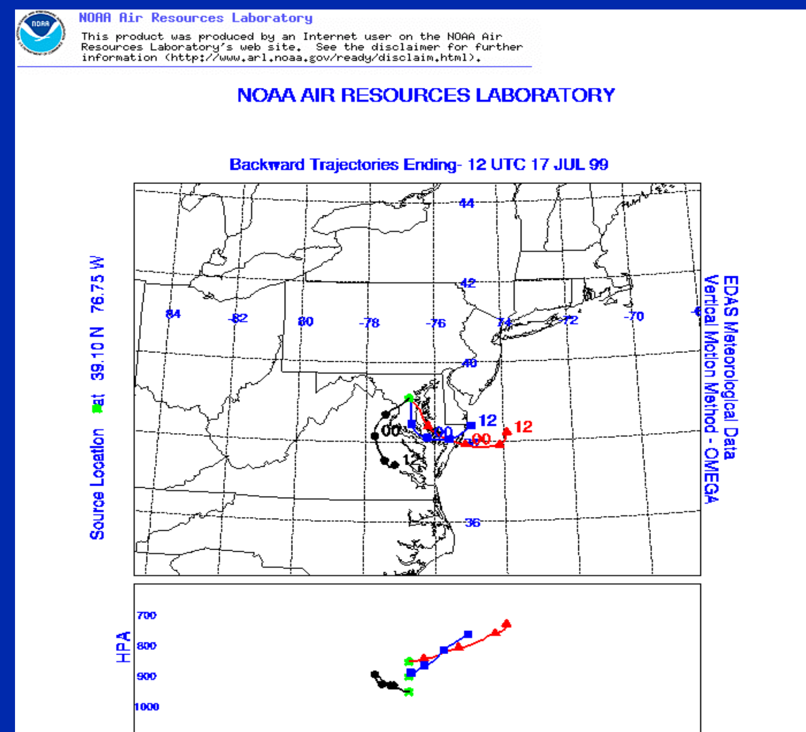
- How are critical physical effects handled by the statistical model – directly or through proxies? For example,
  - UV: Clouds (ASOS?), RH
  - Convection
  - Boundary layer depth



Modification of Statistical Guidance (5 of 7)
How Are Critical Factors Handled?

- Is transported O_3 a factor? Is it captured in the model by current- or previous-day O_3?
- How strongly weighted is regional-scale O_3 in the statistical model?



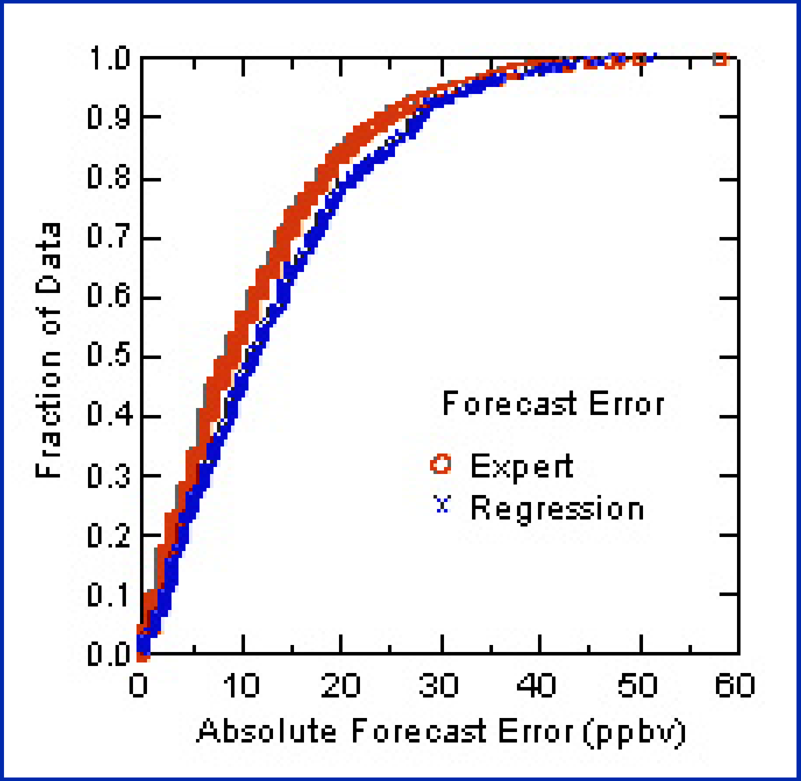


Modification of Statistical Guidance (6 of 7)
The Strategy

- Run statistical guidance using best available model inputs
- Adjust regression output with reference to key factors operating that day that may not be fully resolved by the statistical model
- Example follows

Modification of Statistical Guidance (7 of 7)
What Improvement Can be Expected?

- Experience in the mid-Atlantic, Philadelphia (98-02) – mean absolute error improved 13% or 2 ppbv overall.
- Expert forecasts are much better at resolving OADs.
- Twice as likely to predict an observed OAD.
- False alarm rate for expert forecast is slightly better and much less likely to predict an OAD when only moderate concentrations are observed (6% for expert forecast compared to 17% for regression).

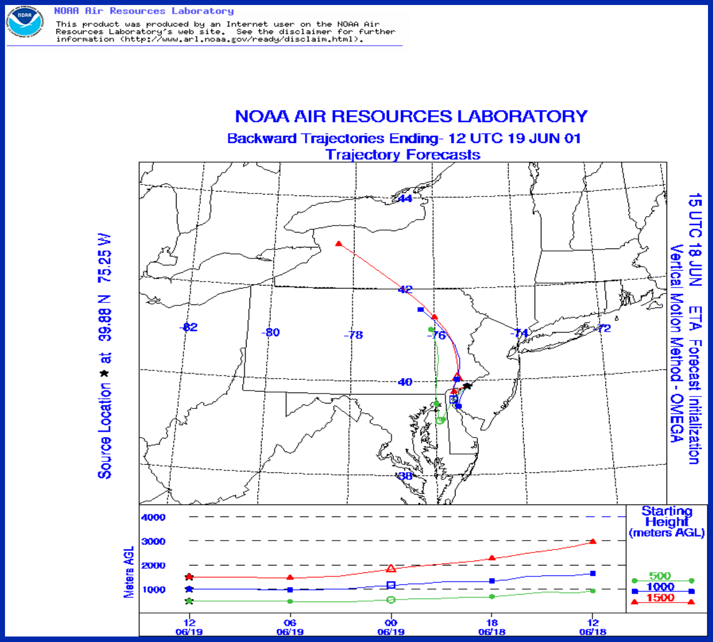


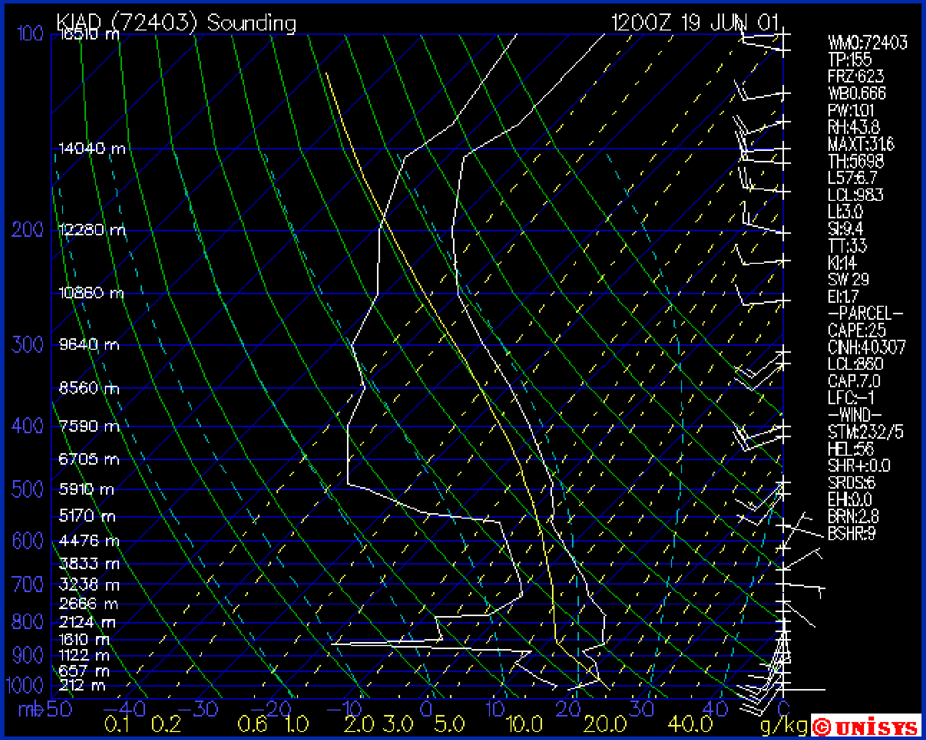
Example Case: Mid-Atlantic Ozone Episode (1 of 7)
June 18-22, 2001

- An O_3 event that is "not exceedingly warm, humid, or uncomfortable".
- Key forecast issues are stagnation and re-circulation, depth of the mixed layer, frontal position, convection, and cloud cover.
- Concentrations rise to the upper Moderate AQI on June 18 as high pressure settles overhead.

Example Case: Mid-Atlantic Ozone Episode (2 of 7)
June 19 – Guidance Adjusted Upward

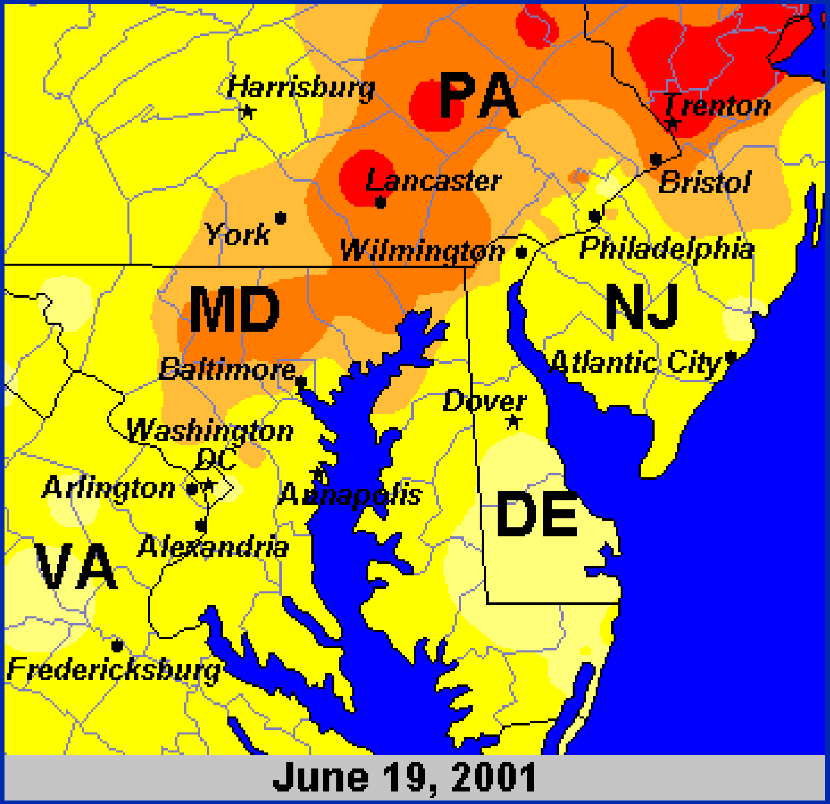
- Regression forecast 119 ppbv (T_{max} only 88°F) but forecast adjusted upward to the Code Red range (125 ppbv) due to re-circulation, strong inversion, and day length

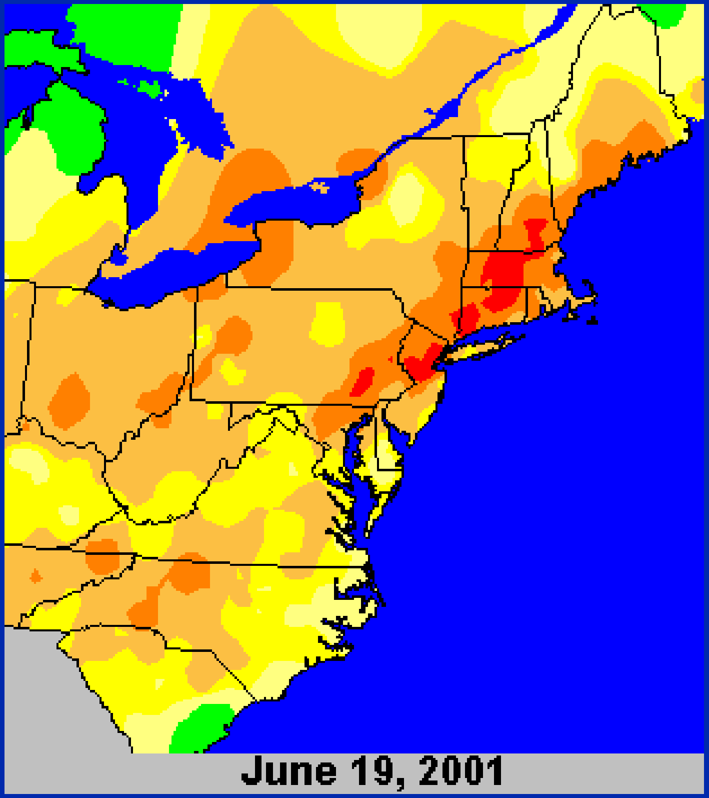




Example Case: Mid-Atlantic Ozone Episode (3 of 7)
June 19 – Code Red Observed

- Code Red observed in PHL with high O_3 region-wide.



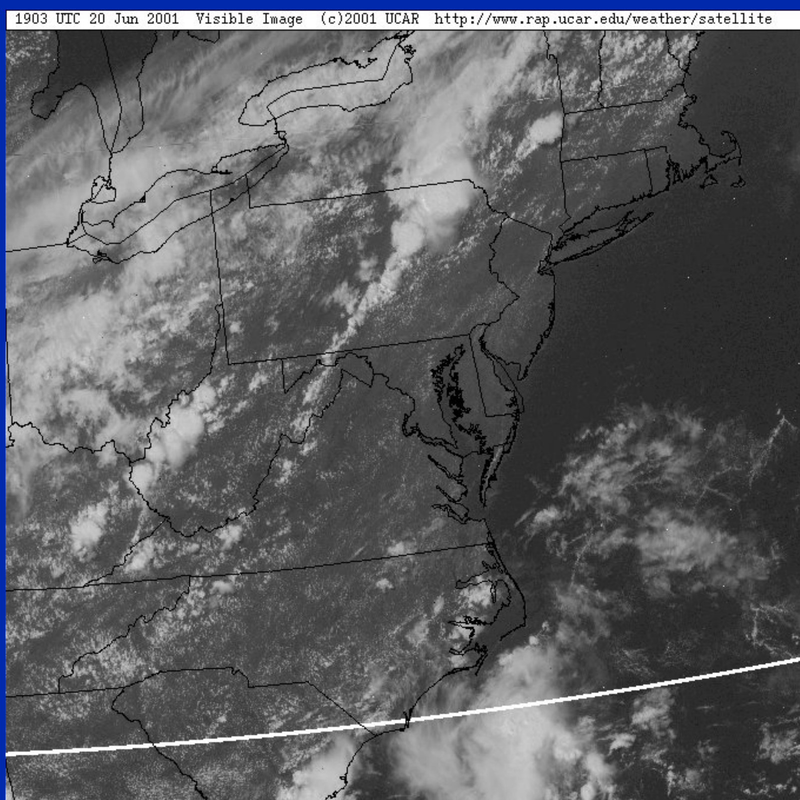


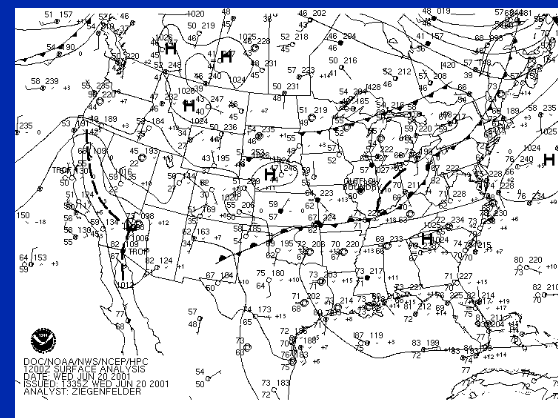
Example Case: Mid-Atlantic Ozone Episode (4 of 7)
June 20 – Second Ozone Action Day

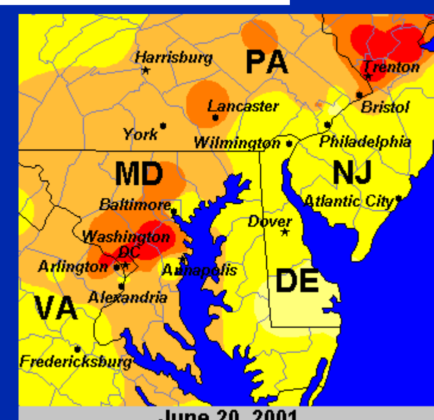
- Regression forecast remains 119 ppbv (Code Orange); T_{max} increases to 90°F but wind speeds increase as well.
- Forecast is again adjusted upward due to high O_3 upstream (WSW wind) and continued shallow mixed layer.
- But, what about convection?

Example Case: Mid-Atlantic Ozone Episode (5 of 7)
June 20

- Convection remains west of the forecast area. Forecast of 125 ppbv verified by observed maximum of 128 ppbv.





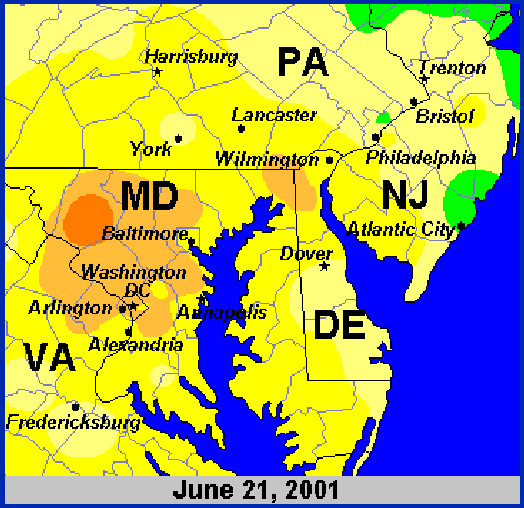


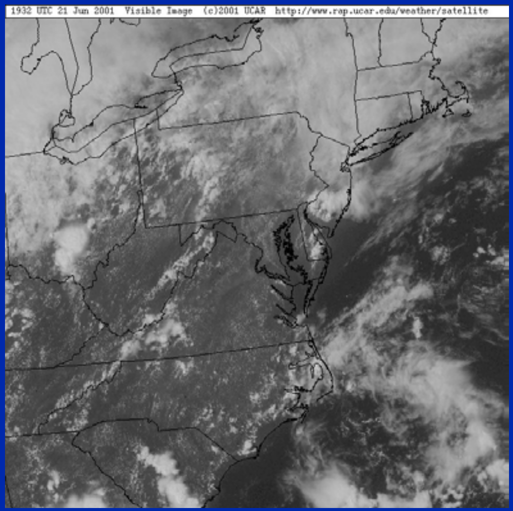
Example Case: Mid-Atlantic Ozone Episode (6 of 7) June 21 – Convection is Key

- Front stalls, a tricky situation, but inversion is weaker so deeper mixing expected.
- Winds still weak but widespread convection expected.
- Statistical guidance trends lower as T_{max} falls in anticipation of thunderstorms. No modification of statistical forecast guidance is made.

Example Case: Mid-Atlantic Ozone Episode (7 of 7) June 21

- Thunderstorms are observed, particularly near PHL, but lingering regional O_3 load keeps concentrations in the Code Orange range in areas not hit by thunderstorms, as forecast.
- Widespread convection and rain ends episode on June 22.





Conclusions (1 of 2)

- Keep a good forecast archive.
- Know the strongest statistical predictors and techniques to forecast them.
- Know the “envelope of uncertainty” for the statistical model, overall and for key subsets.
- Know the key factors producing the pollutant.

Conclusions (2 of 2)

- Understand how key factors are handled, or not handled, by the statistical model.
- Adjust the forecast to reflect conditions affecting pollutant that are not fully resolved by the statistical forecast.
- Take the time to do post-mortems on pollution events.