Forecasting Pollution: Observations, Models and Value of Information

Anne M Thompson, PSU Meteo

Students J E Yorks (MS, 2007), G G Garner (PhD candidate)

Co-Advisors: W F Ryan, K Keller

Talk Outline

- Mid-Atlantic Pollution Problem
  - Observations
  - Models & Tang et al (2008) study with IONS-06
  - Over-prediction & residual-layer – Beltsville data
- Air Quality (AQ) Forecasts
  - Monitoring data used in model evaluation (Eder et al, 2009, approach)
  - Types & measures of quality
  - Value of Information (VOI) as metric for evaluating quality

IONS-04 for INTEX-NA, July-Aug 2004
IONS-06 for TEXAQS (Spring & Summer)

IONS-06 Sites

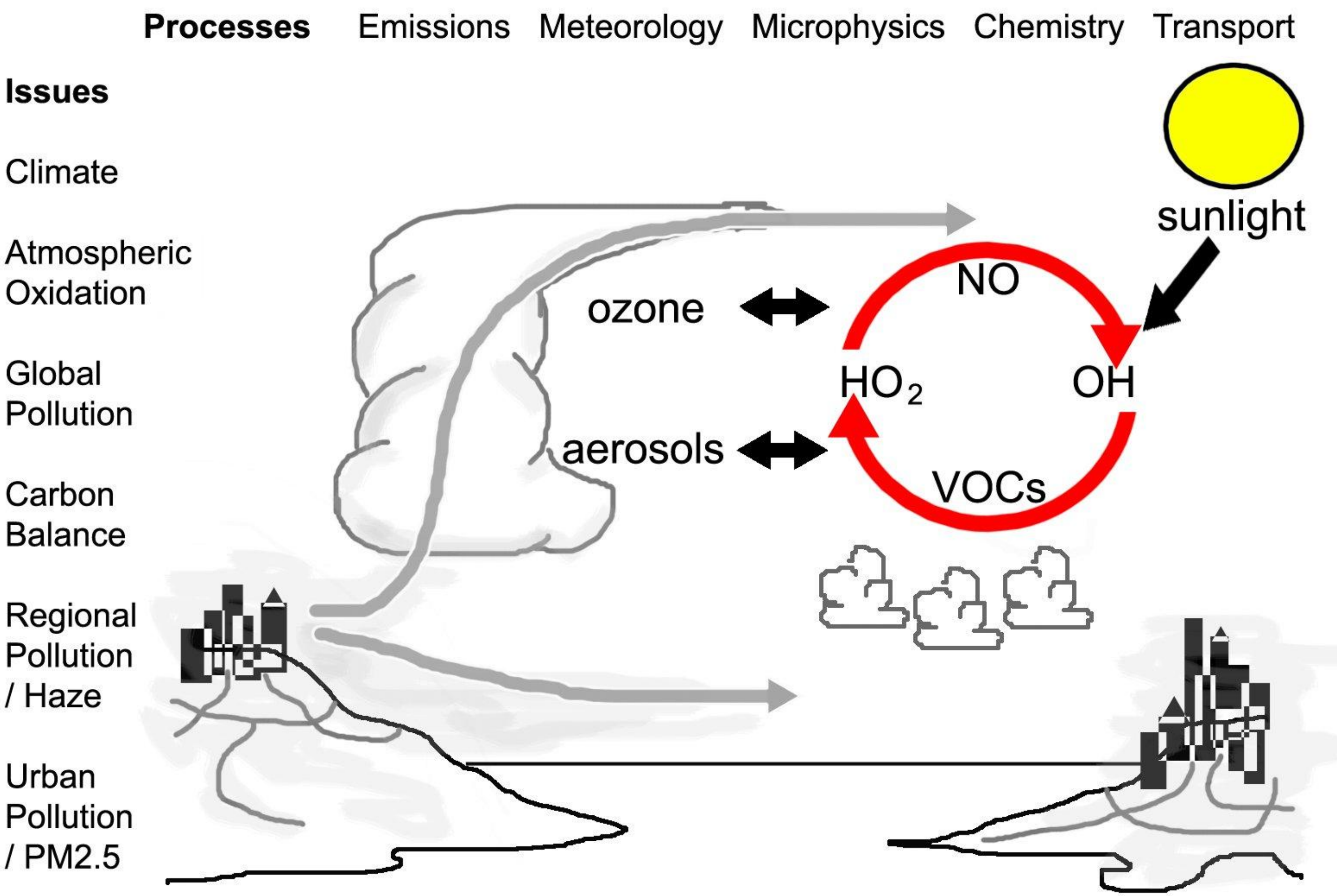
1 - Kelowna, BC
2 - Stonyplain, AB
3 - Bratt's Lake, Sask.
4 - Egbert, ON
5 - Walsingham, ON
6 - Paradox, NY
7 - Narragansett, RI
8 - Yarmouth, NS
9 - Sable Is. NS
10 - Valparaiso, IN
11 - Beltsville, MD
12 - Wallops Is., VA
13 - Trinidad Head, CA
14 - Table Mountain, CA
15 - Imperial Valley, CA
16 - Boulder, CO
17 - Los Alamos, NM
18 - Huntsville, AL
19 - Houston, TX
20 - R/V Ron Brown
21 - Barbados
22 - Tecamec, Mexico

IONS-04 daily sondes [Thompson et al, JGR, 2007a,b].

Right: IONS-06 design

http://croc.gsfc.nasa.gov/intexb/ions06

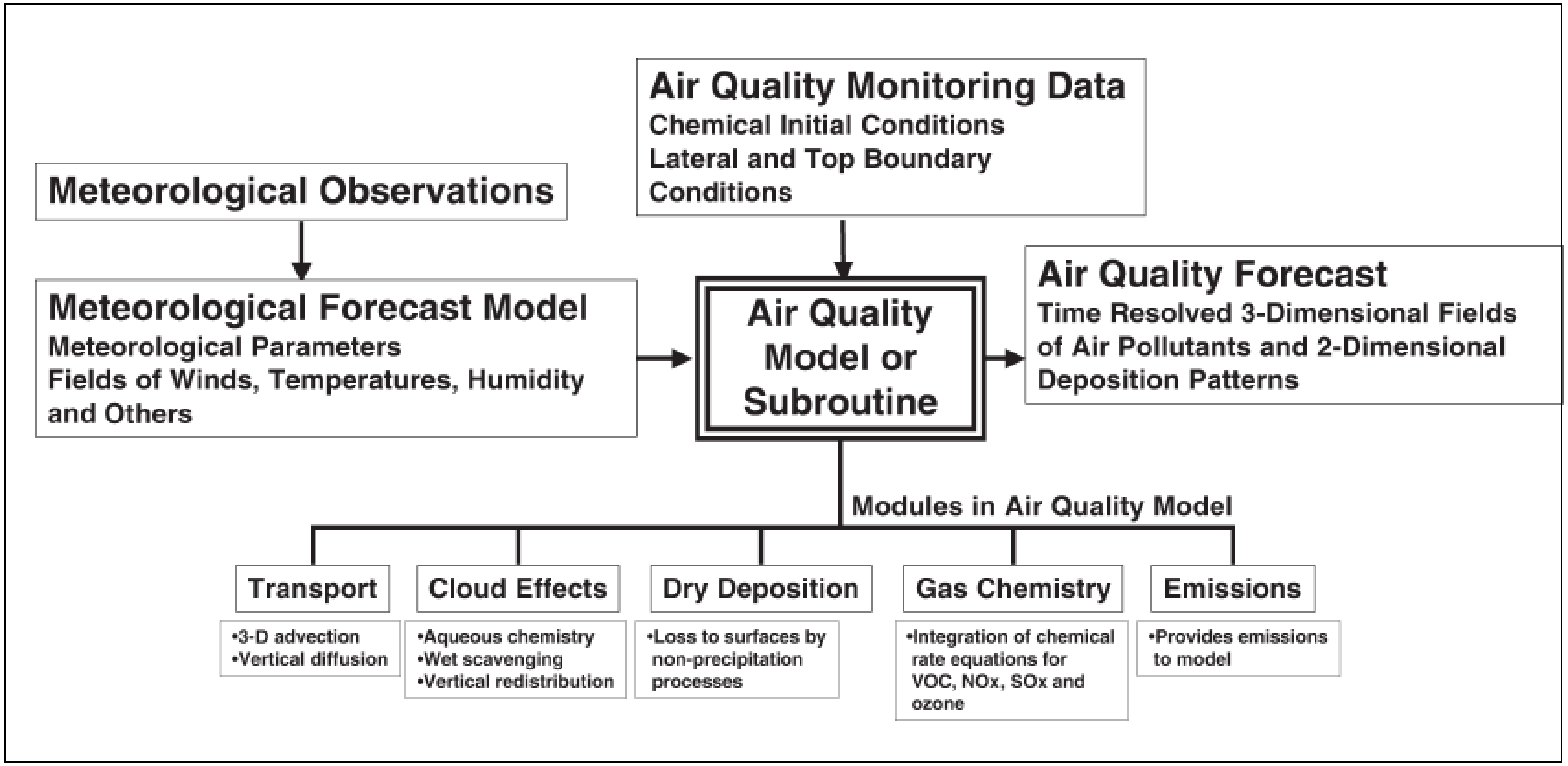
Reference: A. M. Thompson et al., *Atmos. Chem. Phys*. 8, 5113-5125, 2008.



Issues: - AQ Models require accurate simulation of emissions, reactions, dynamic processes

Elements of NAQFC: Natl AQ Forecast Capability

- The system consists of 2 coupled models:
  - NAM-12 (2005) & WRF (updated) provide meteorology data
  - EPA CMAQ model predicts tropospheric ozone production and transport (chemistry)

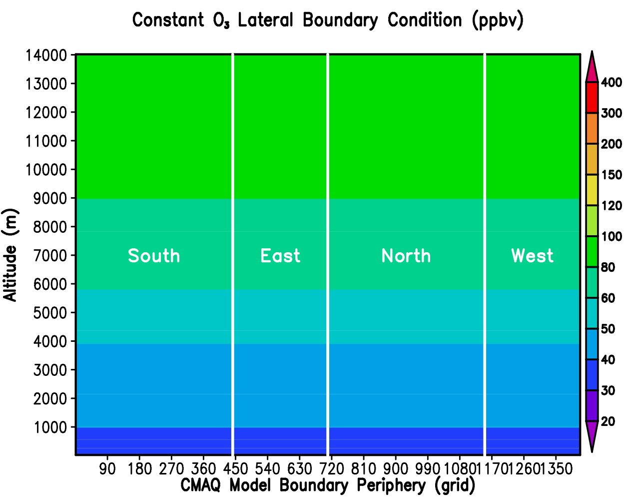


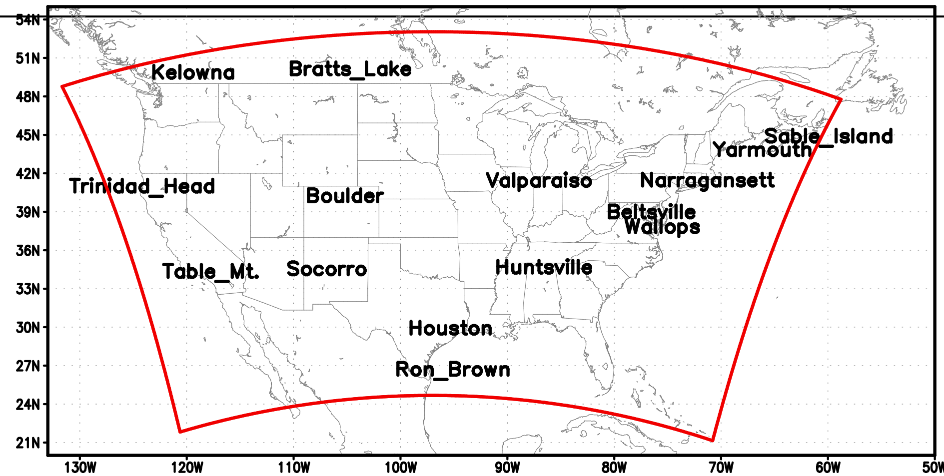
Ozonesondes with CMAS (Community Models for AQ/Systems) – Y. Tang et al., *Environ. Fluid Mech.,* 2008

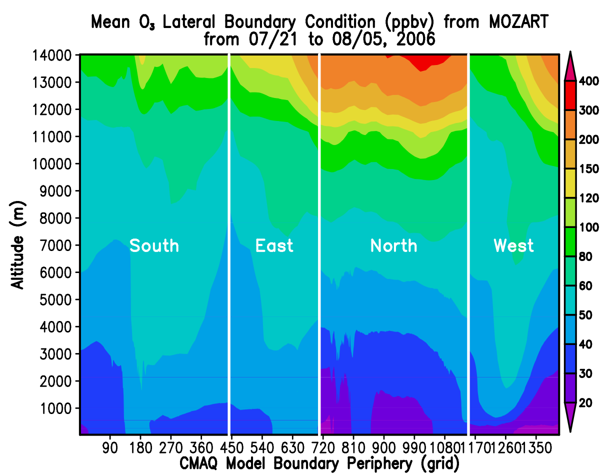
- Basic premise – most sensitive components of AQF – lateral boundary conditions (LBC), chemical reaction model 'package'
- Similar result, IONS-04: Tarasick et al., *JGR*, 2007
- Tang et al (2008) study with IONS-06 data (Thompson et al., 2008)
  - Run (1) standard "fixed," static "LBC" (2) different model-based LBC varying in region and time; (3) LBCs based on IONS-06 sonde data
  - *Evaluate* model ozone profiles with IONS-06 data

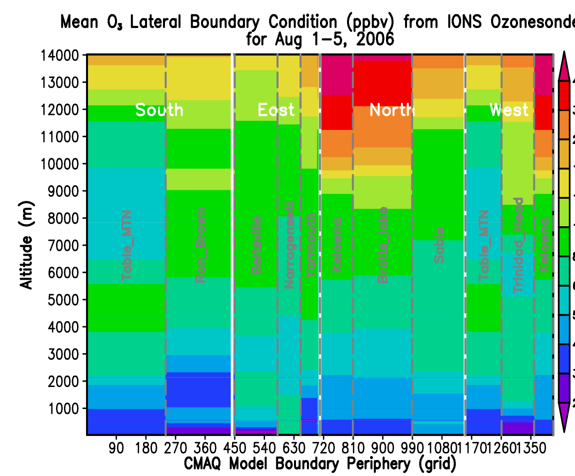
CMAQ & IONS-06: IONS-06 Network (upper). Three of six LBC tested (lower). From Tang et al., *EFM,* 2008

- IONS-06 Network,
- Thompson et al., *ACP,* 2008



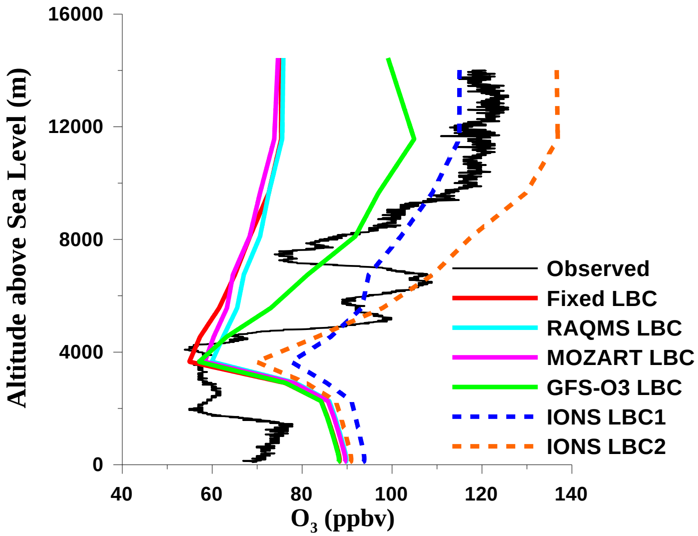




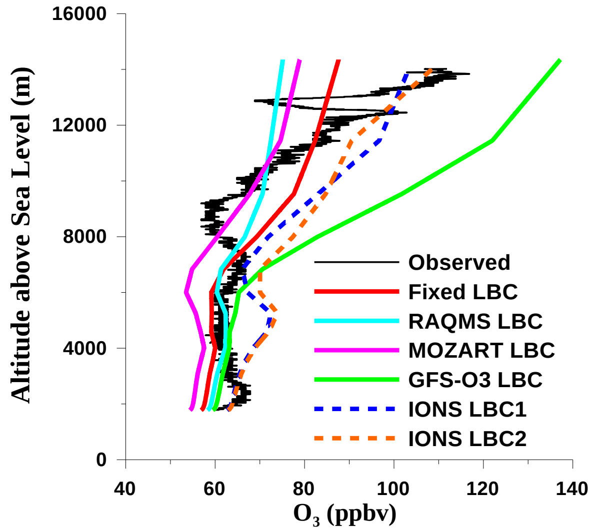


CMAQ Result: Individual IONS-06 Comparisons with Six LBCs 3 Aug 06. Pollution aloft, BV, HSV; surface not very polluted From Tang et al., *EFM,* 2008

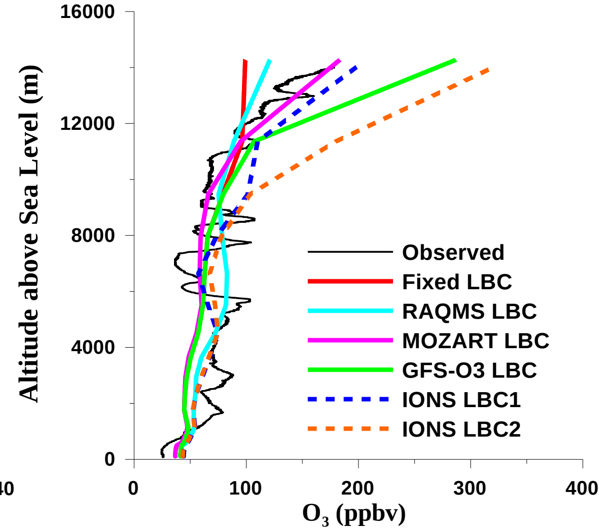
Beltsville 20060803 17.92 UTC



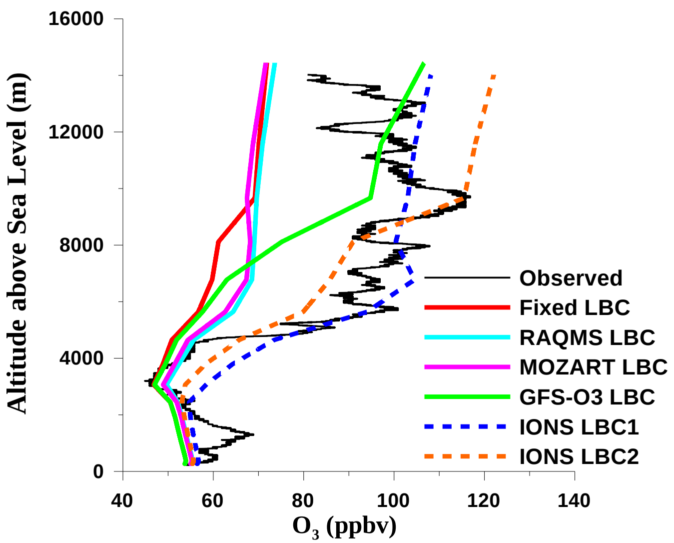
Boulder 20060803 19.33 UTC



Trinidad Head 20060803 21.02 UTC



Huntsville 20060803 17.53 UTC

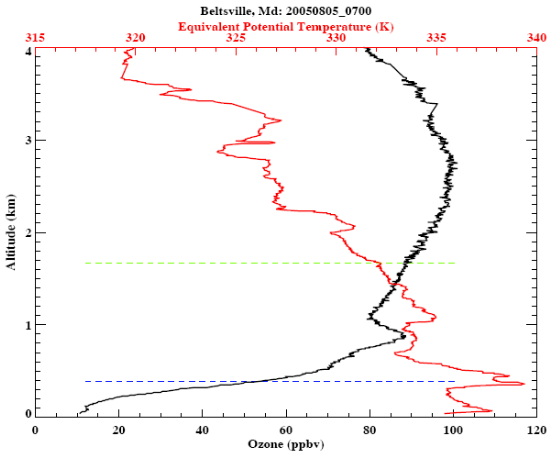


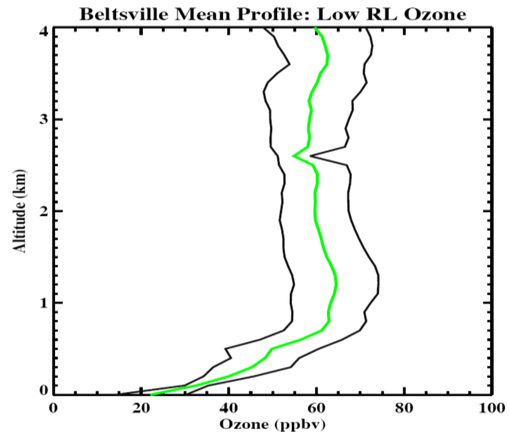
> Free tropospheric ozone improved with IONS LBC.

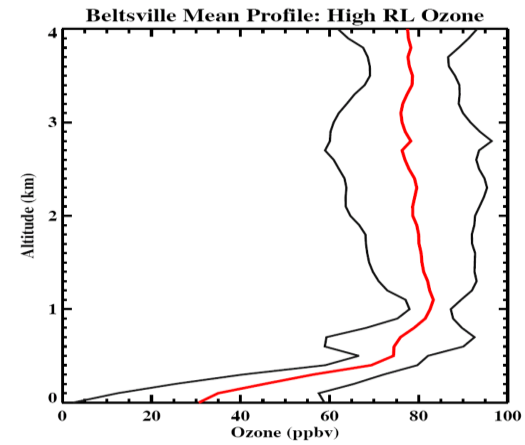
> Surface AQF not accurate, esp at Beltsville.

- Surface AQF not very LBC-sensitive
- Mixing issue with model?
- Chemical reaction set limitations?

Factor in Surface Pollution: Residual Layer. BeltsvilleSonde Statistics, from JJA 2005-2007 Data





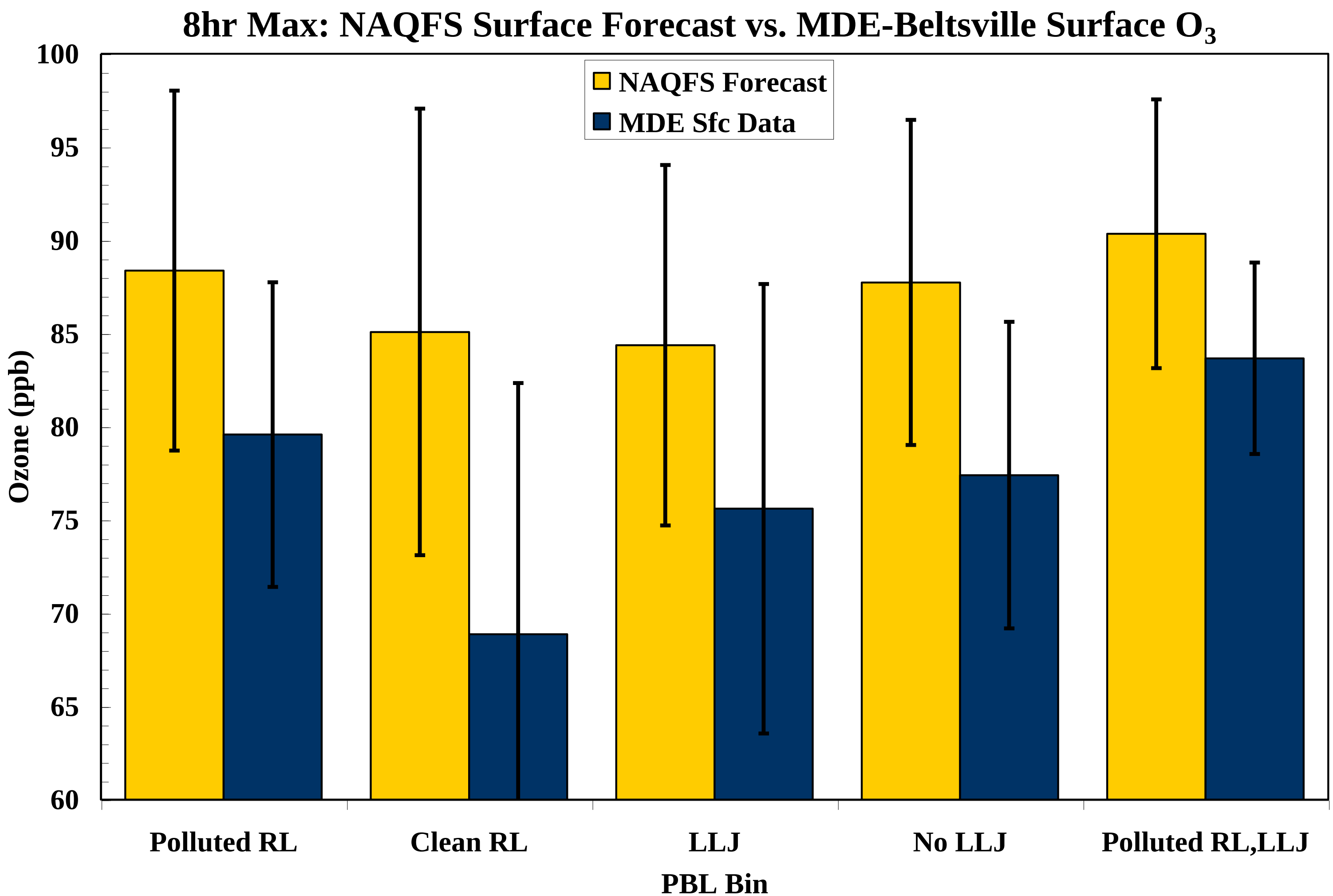


- **Define a "residual layer" with equiv PT inflections. Distinct from Surface Layer – RL can be "local" or transported ozone**
- **Two distinct sets of RL profiles emerge (center, right)**
- **Consider an Alternative Empirical Air Quality Forecast**
  - **Compare nighttime sonde data with surface ozone data**
  - **Assumption: mixing from residual layer above (RL – yesterday's ozone, maybe from upwind!) critical to surface ozone evolution**
  - **Does high ozone RL predict high surface ozone better than a model?**

J E Yorks, MS Thesis, 12/07.

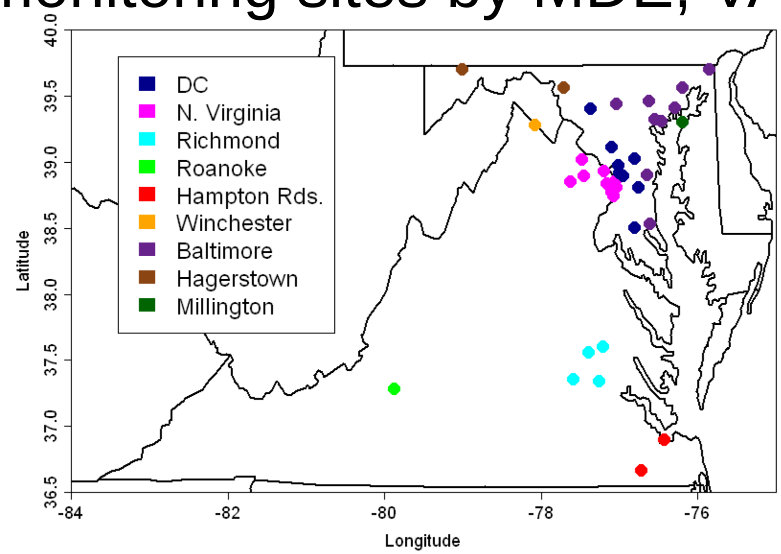
- Sonde Evaluation of NAQFS Forecast – For days with nighttime sondes, classified by RL and whether or not a low-level jet, mixing RL ozone to SL
- **Persistent over-prediction in all cases**
- **O_3 avg. observed below NAAQS threshold, avg. forecast above threshold**

Only in one case did surface ozone actually exceed threshold

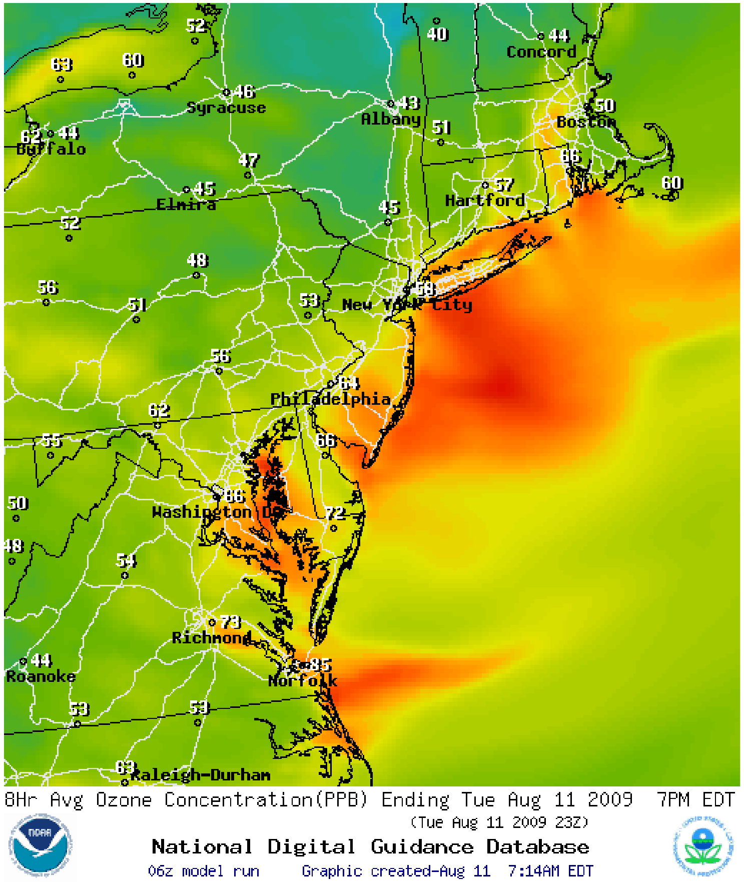


- National Air Quality Forecast Capability (NAQFC). AQI Color Code – Ozone-based

- Focus on Maryland, DC, Virginia forecast regions. Note 41 monitoring sites by MDE, VADEQ



Air Quality Index (AQI) Values	Levels of Health Concern
When the AQI is in this range:	...air quality conditions are:
0 to 50	Good
51 to 100	Moderate
101 to 150	Unhealthy for Sensitive Groups
151 to 200	Unhealthy
201 to 300	Very Unhealthy
301 to 500	Dangerous



AQ Forecasts

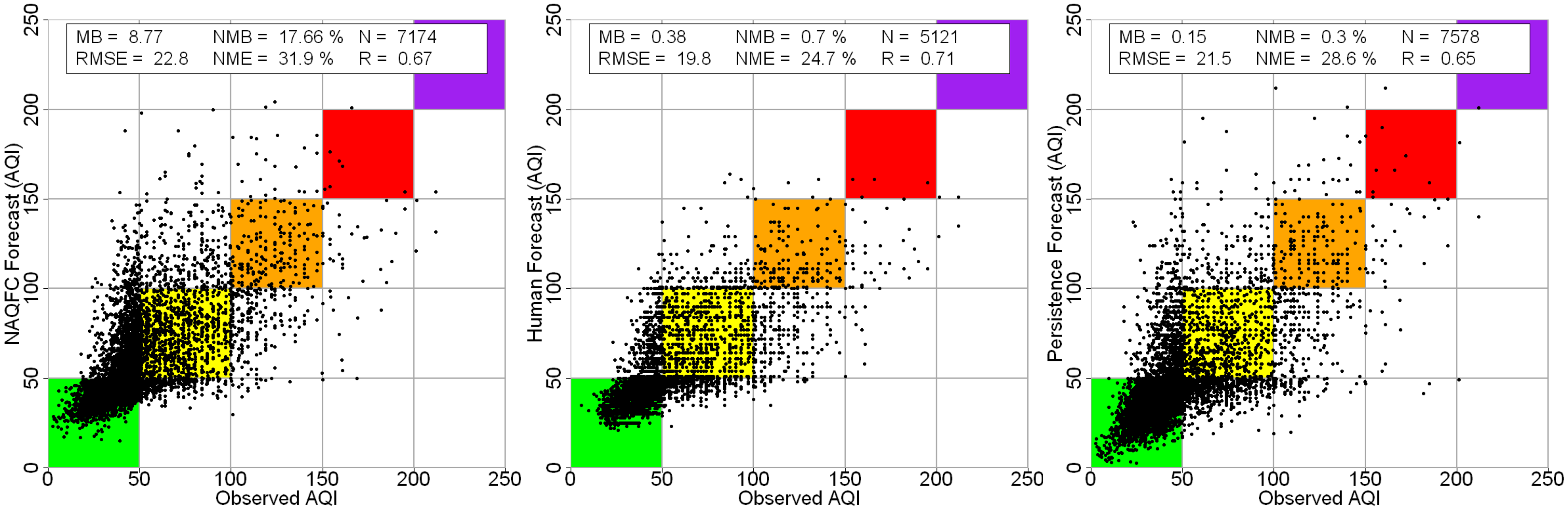
- Three types evaluated
  - Persistence – today → tomorrow
  - Human: assembles various forecast info
  - Model – NAQFC evolving in US (ozone now, aerosols coming)

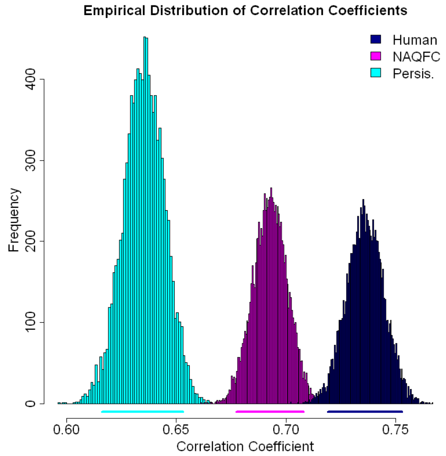
- Measures of quality
  - Mean bias (absolute, normalized), RMSE
  - Also Critical Hit rate, False alarm rate, etc
  - Value of Information is mathematical relationship based on differencing climatology and forecast with a measure of costs / losses of prevention or not

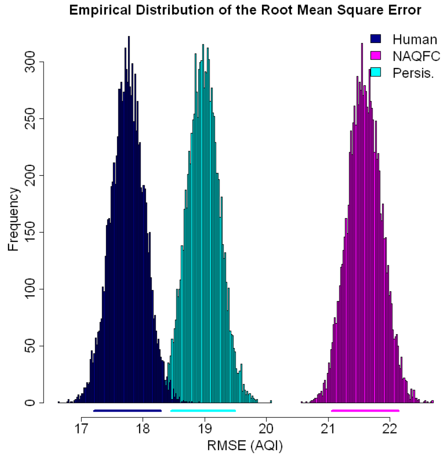
Value Analysis*

- Data for evaluating models
  - 2005-2009 hourly, 8-hr mean surface ozone from MD-VA-DC monitoring stations, June-July-August
  - "Value" – use cost-loss ratios – financial data rarely available
- Value based on simple 'umbrella' concept
  - There is 'loss' if preventive measures are not taken on bad AQ days (tourism, health, etc)
  - Costs are incurred when measures taken (driving alternatives, free buses, curtail other activities)
  - * **Garner & Thompson, submitted, 2011**

Forecast Comparisons (y- forecasts; x- observed AQI) . Goal – Get Orange 'Correct'







‘Umbrella Problem’ – Rain Forecast Accuracy vs Decision to Protect.

	Protect	Do Not Protect
Good	Cost	null
Bad	Cost	Loss

- Goal: Maximize loss avoided relative to cost of measure

Cost

Loss

How Much is Lost on Given Day

- If one Protects:

  (Prob {Good} x Cost) + (Prob {Bad} x Cost)

  = (Prob {Good} + Prob {Bad}) x Cost

  = Cost

- If no protections:

  (Prob {Good} x 0) + (Prob {Bad} x Loss)

  = Prob {Bad} x Loss

- GOAL:

  Minimum (Cost, Prob {Bad} x Loss)

Probabilities with/without Forecast

- With no forecast – use climatology

  Prob {Bad} → Climatological probability that a Bad event occurs

$$\pi(\theta \in \Theta_i)$$

  - With Forecast – define probability with "y"

  Prob {Bad} → Probability that a Bad event occurs with the given forecast

$$\pi(\theta \in \Theta_i \mid y_t)$$

Goal is to Minimize Loss – Maximize Value

- You expect to lose with climatology…

$$l_o = min\{C, \pi(\theta \in \Theta_i) \cdot L\}$$

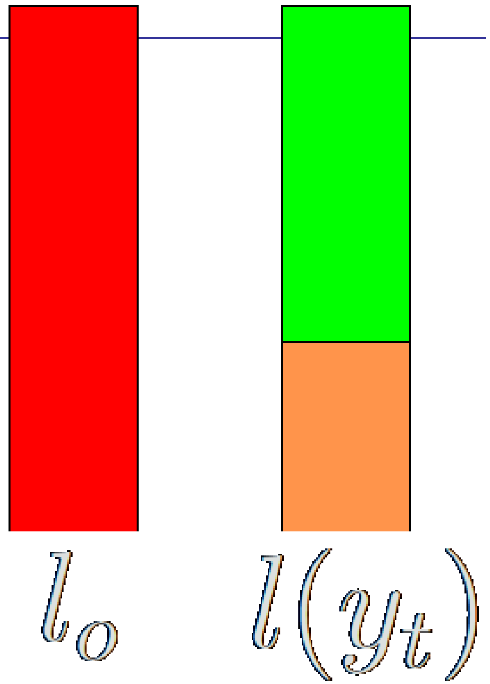
- You expect to lose with the forecast…

$$l(y_t) = min\{C, \pi(\theta \in \Theta_i \mid y_t) \cdot L\}$$

- Difference between the two is "value"

$$l_o - l(y_t)$$

Defining Value

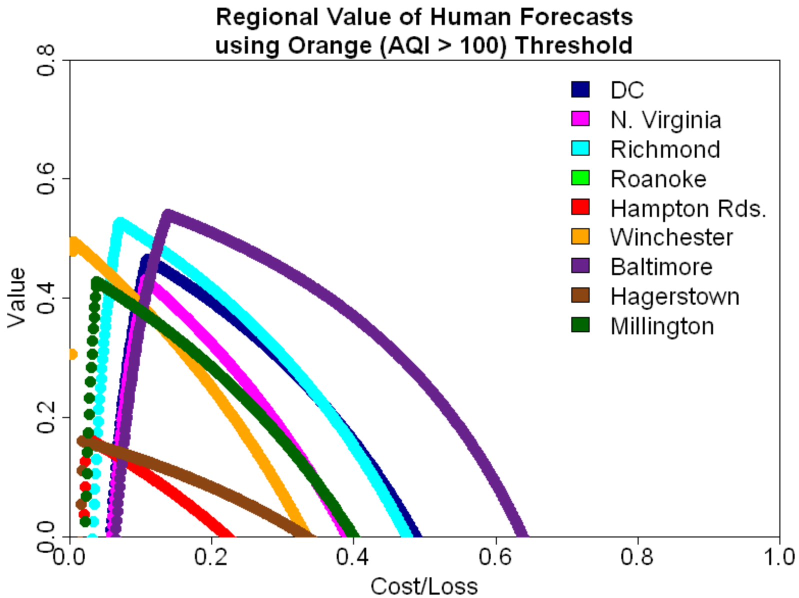


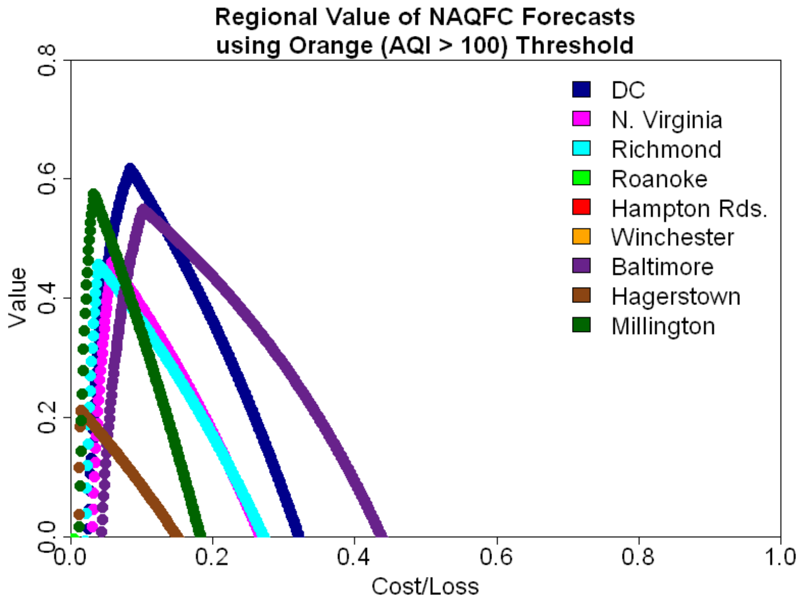
- Expected “Value” of forecast system depends on probability of forecast “Bad” event
- Sum up over all levels above threshold (AQI=100):

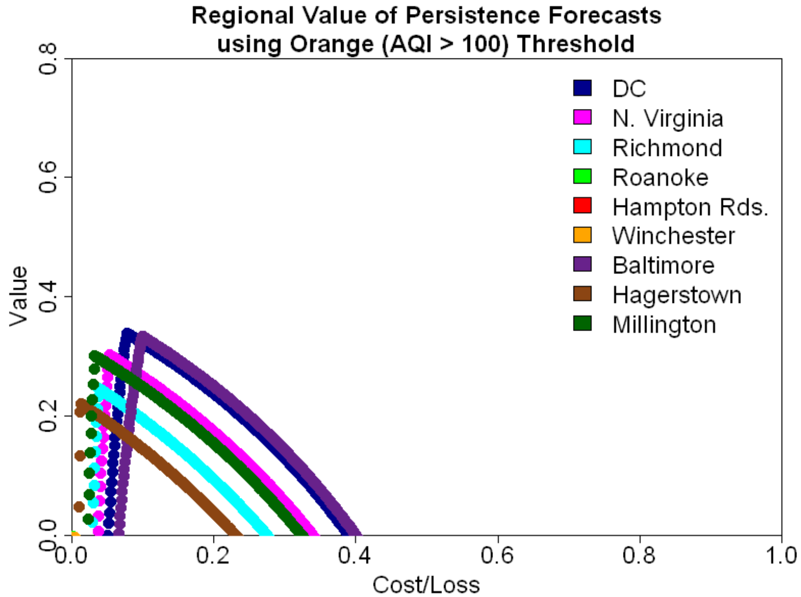
$$V = \sum_i [Prob\{y_t \in \Theta_i\} \cdot (l_o - l(y_t))]$$

C/L – protective measures. Goal is wide range

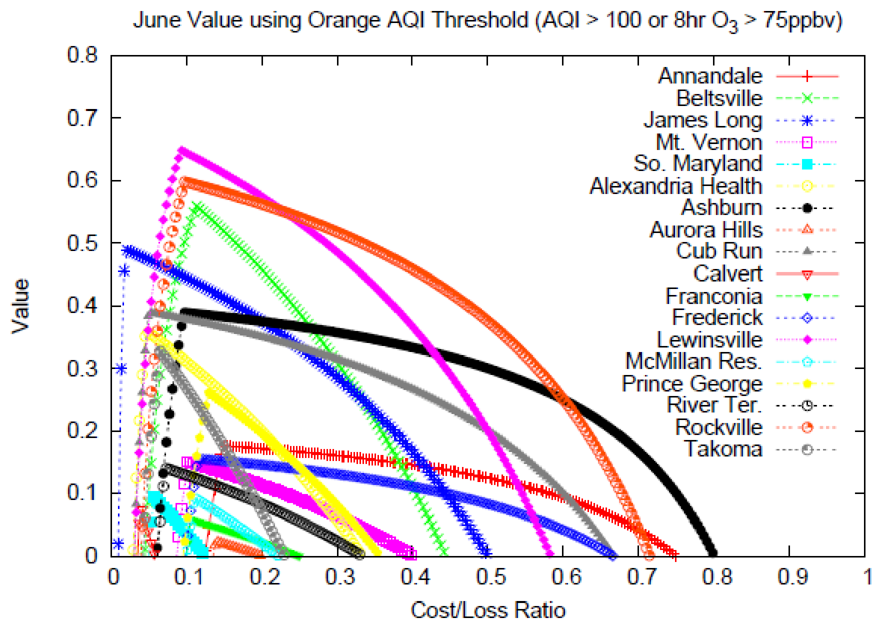
Forecast Value

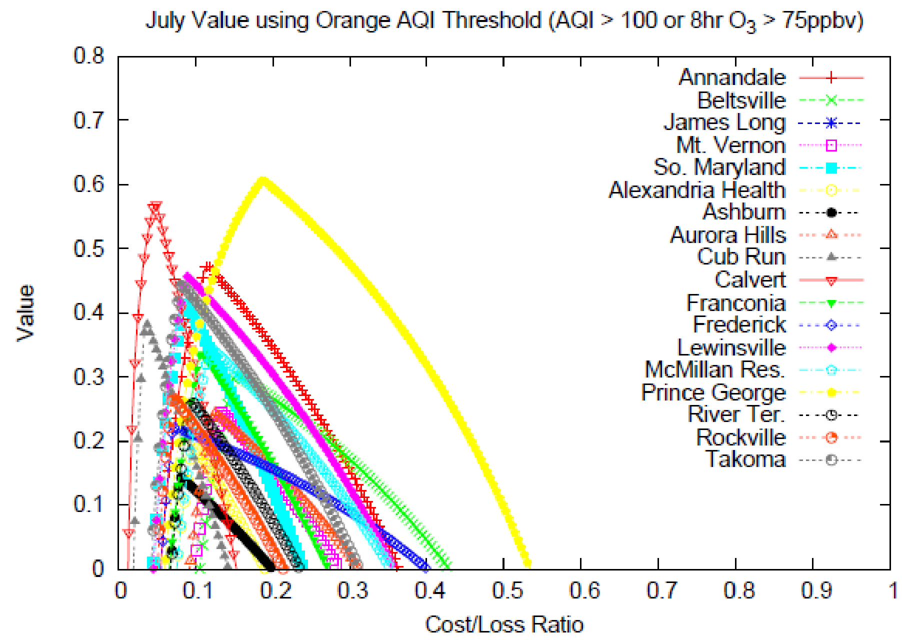




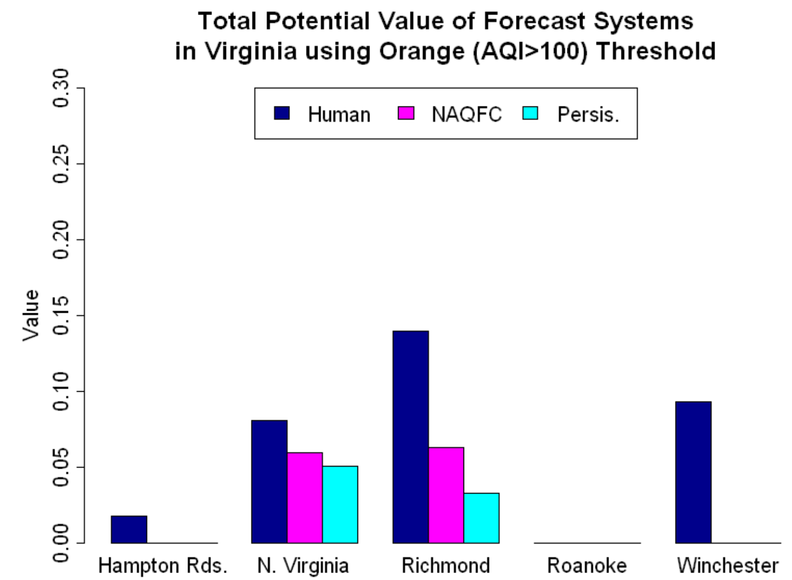


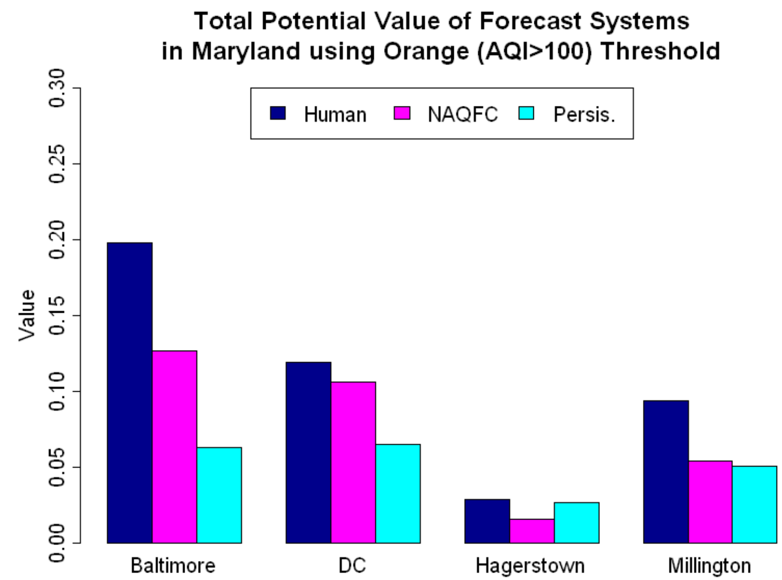
Spatial & Temporal Variation in Value

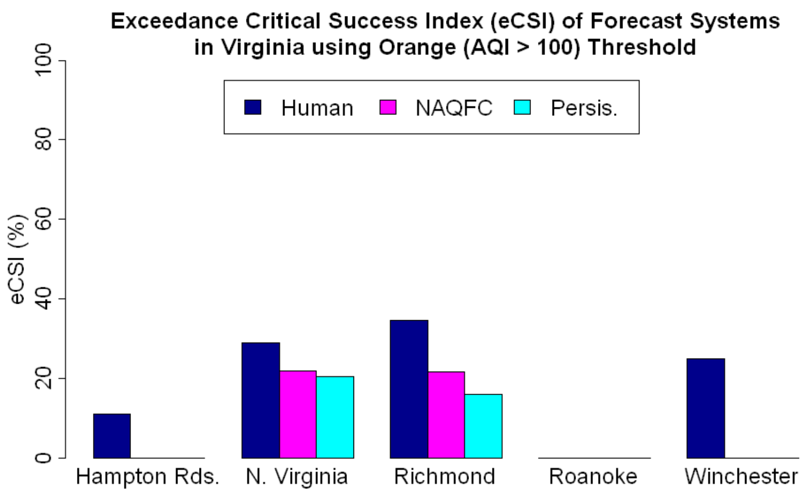


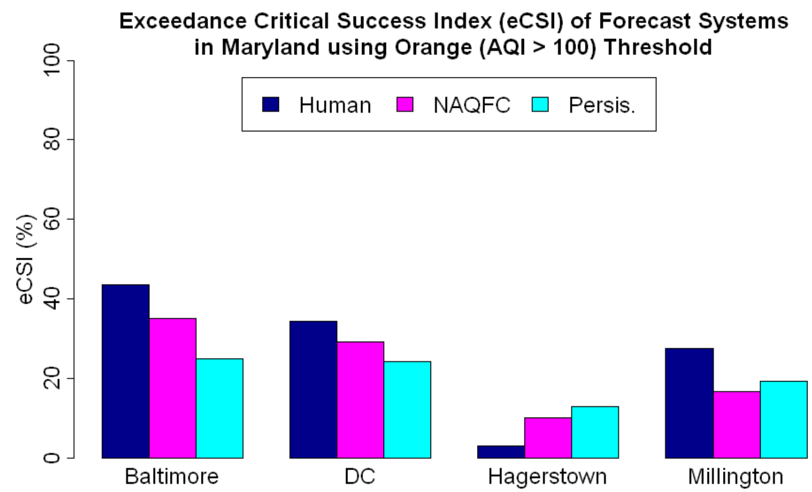


Value as a metric…









Summary

- Soundings valuable for AQ model evaluation. Nighttime profiles good AQ predictor (Morris et al., 2009, Houston)
- Evaluation of 3 forecasts shows human better than model (NAQFC) at this point
- Value of Information approach
  - Forecasts evaluated with economic value in mind
  - Probabilities used to assign value
  - Practical applications – DISCOVER Expt in Maryland-VA, July 2011. Siting tests for Monitoring stations

Acknowledgements

- Data provided by
  - Maryland Department of the Environment
  - District Department of the Environment
  - Metropolitan Washington Council of Governments
  - Virginia Department of Environmental Quality
  - NOAA – National Operational Model Archive and Distribution System

- Duc Nguyen (MDE), and Laura Landry (MDE)
  - Valuable insight into AQ forecasting operations

- Funded by NSF DRU Program Award I.D. 0729413 with added support from NASA Tropospheric Chemistry Program and Aura Validation.