EPA & CDC Symposium on Air Pollution Exposure and Health

Air Quality Estimates for Public Health Surveillance

(aka, PHASE Modeling for EPHT)

Fred Dimmick, David Holland, Alice Gilliland[1] and Tim Watkins

Environmental Protection Agency

Office of Research and Development

National Exposure Research Laboratory

September 19, 2006

Research Triangle Park, NC

[1] National Oceanic and Atmospheric Administration

Focus of Presentation

Modeling ambient PM, O3 and ~~air toxics~~ concentrations … though combining air quality modeling data with air quality monitoring and satellite data.

Questions to Address

- What are the strengths and weaknesses of models used for estimating individual or population level air pollution exposures in health effects research or for public health tracking?
  - Availability, coverage and reliability of information required by these models (e.g., emissions inventories, traffic/roadway information, ambient monitoring measurements, etc.)
  - Cross-sectional vs. longitudinal modeling issues
  - Differences between models or methods used for predicting exposures in air pollution health effects research and public health tracking studies, and their potential implications to interpretation of results
  - Uncertainties in modeled estimates
  - Evaluation of model results by comparison with ambient concentration measurements
- What are the key programmatic and research needs of the existing ambient pollution modeling approaches?
- What new modeling activities are being planned or anticipated that could support future air pollution health effects, accountability or public health tracking programs?

Public Health Air Surveillance Evaluation

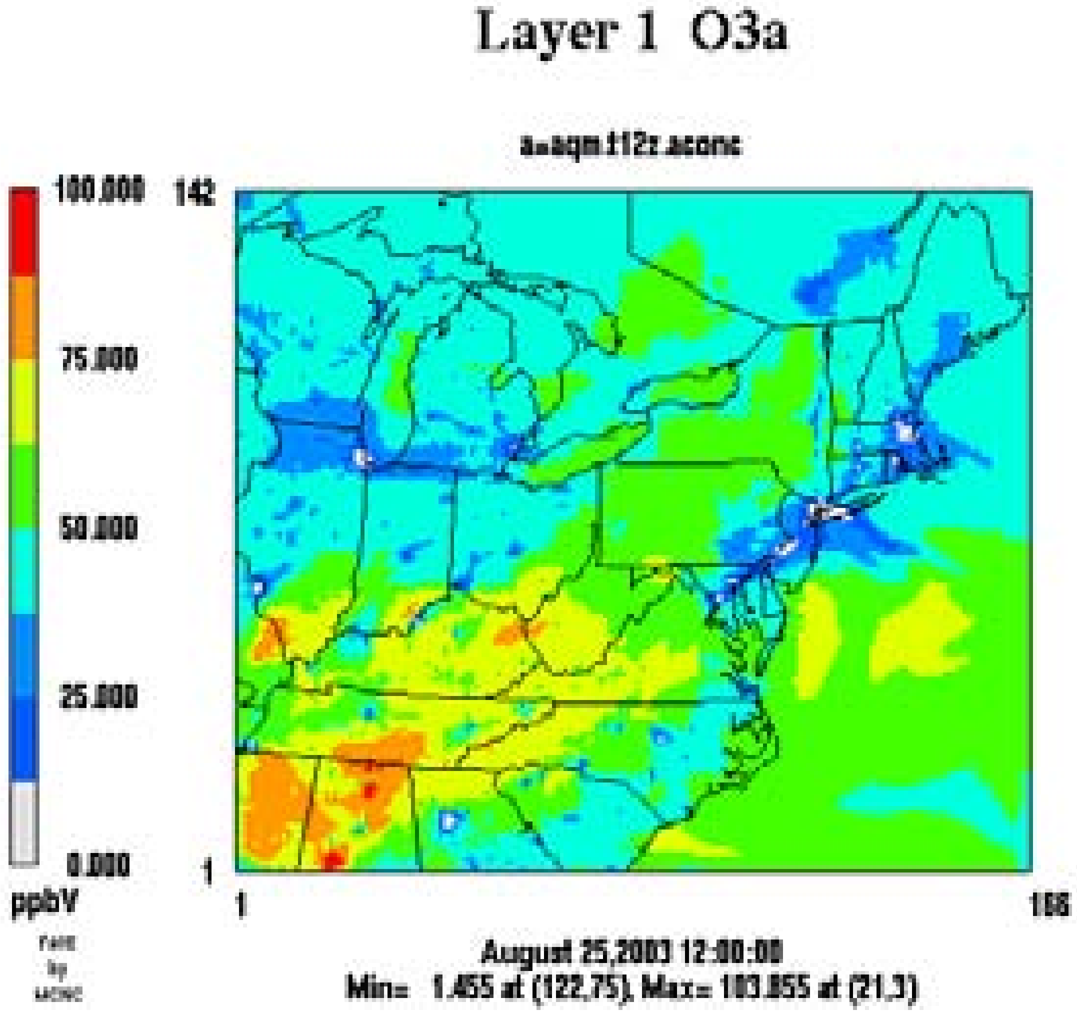
Sources of Air Quality Data

- Ambient Air Monitoring
  - **AQS: PM2.5 and Ozone for 2001**
- Air Quality Modeling
  - **CMAQ: PM2.5, Ozone and Met for 2001**
- Satellite Data
  - **NASA**
- Spatial Predictions: Combined Data
  - **EPA/National Exposure Research Lab**

Air Quality Modeling

- Estimate of air quality
  - High spatial and temporal resolution
  - 36km × 36km or 12km ×12km horizontal grid typically
  - 1-hr time steps
  - *but* higher error relative to monitoring data

- Air Quality Forecasting
  - Emerging source of routine model estimates



Improved Spatial Prediction with Combined Sources of Data

Issue: Cannot monitor at all locations, but want to know pollution everywhere

- Typical Solution: use kriging to interpolate air monitoring data, but
  - Monitoring data is spatially sparse, some areas have no monitors
  - Use of classical kriging techniques may introduce arbitrarily large prediction errors in these areas

- New Solution: Consider Combined Prediction Approaches

Improved Spatial Prediction with Combined Sources of Data

What Does the Combined Approach Provide ?

- Monitoring Data and CMAQ model output can be used simultaneously to predict the pollutant surface (i.e., continuous gridded concentration field)

- Draw on strengths of each data source:
  - Give more weight to accurate monitoring data in monitored areas
  - Rely on model output in non-monitored areas
  - Statistically model underlying spatial and temporal dependence, and measurement errors of each data source

- Leads to more accurate daily predictions and provides prediction errors

Availability, coverage and reliability of information required by these models

- Information for Air Quality Modeling (CMAQ)
  - Emissions inventories: readily available for a given year, requires projections for other years
    - Power plant *(most certain – use stack measured emissions)*
    - Traffic/roadway information *(less certain – based on mobile model)*
    - Area (i.e., more dispersed) sources *(most uncertain – based on static emission factors and estimated activity)*
  - Meteorology – readily available
    - includes assimilation of observed meteorology
  - Chemistry – under continuous improvement
    - Ozone chemistry has longest history of observation and study
    - $PM_{2.5}$: sulfate production best understood, secondary organic aerosols most uncertain

Availability, coverage and reliability of information required by these models

- Statistical Predictions
  - Air Quality – AQS (via State/Fed monitoring)
    - Ozone data robust – reference method & QA'd
    - PM2.5 recently robust -- reference method & QA'd
    - Metropolitan areas ++ // Rural areas –
  - CMAQ – existing & special applications
    - Need years not developed for EPA regulatory purposes
    - CONUS – 36km, Eastern US – 12km

Spatial and Temporal Characteristics

(aka cross-sectional and longitudinal issues)

- <u>CMAQ</u>
  - Grids @ 36, 12 and 4km
  - Hourly output, annual to multi-year simulations
- <u>Statistical</u>
  - Grid - same as CMAQ
  - Time periods limited by AQ data
- <u>Longitudinal studies</u> require long time-series (e.g., ten years) not currently available from AQ models (*but coming along*)
- <u>Cross-section studies</u> require daily/hourly data available from model and observations *(if available in that area)*

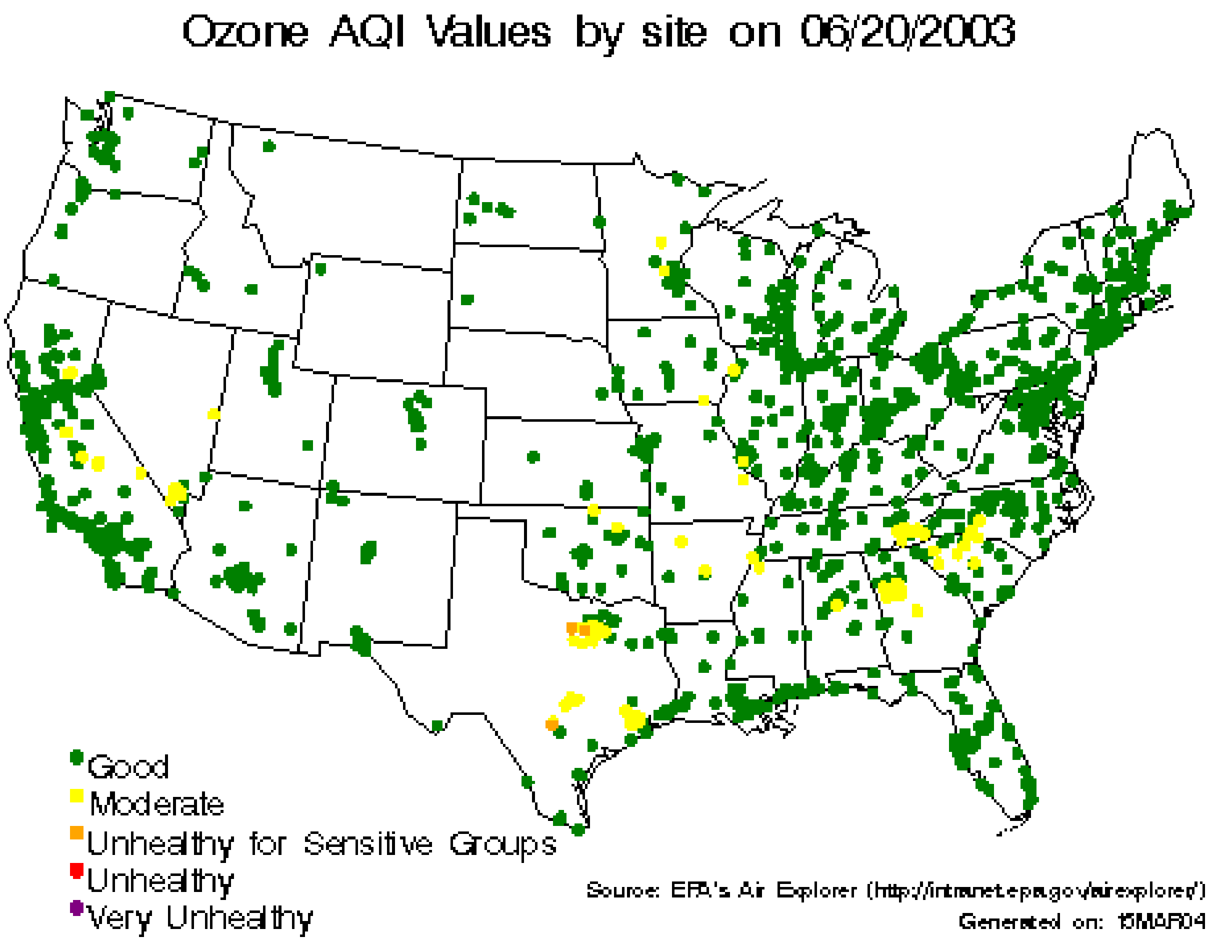
Differences between models or methods used for predicting exposures in air pollution health effects research and public health tracking studies, and their potential implications to interpretation of results

- Spatial resolution
  - Grid sizes (36, 12, 4 km)
- Temporal resolution
  - Hourly, Daily, Every Third-Day,
- Relation to models to link AQ & health outcomes
  - Street addresses, zip codes, county
  - Monday spike in HD

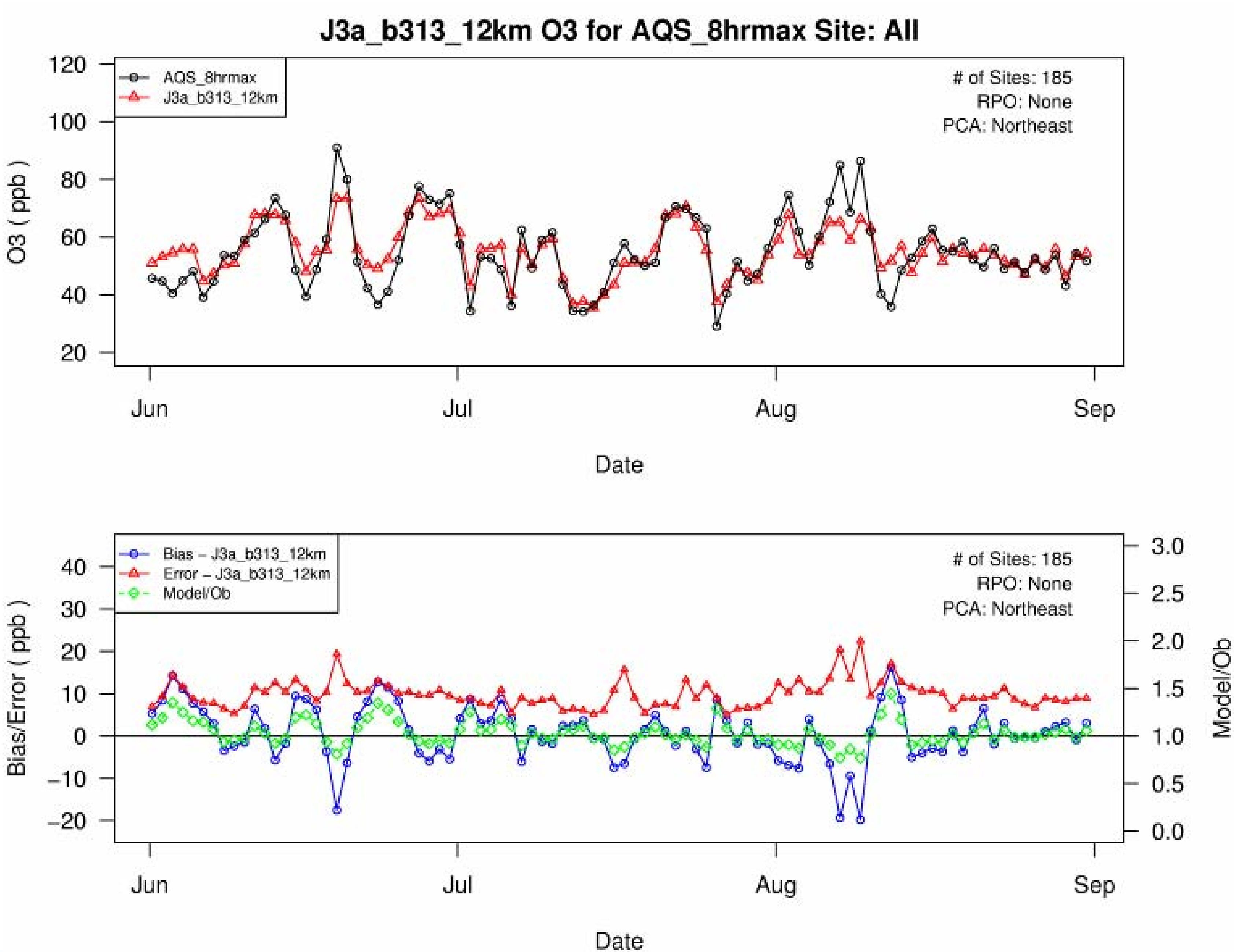
Uncertainties in modeled estimates

Evaluation of model results by comparison with ambient concentration measurements

- "True" measure of air quality
- Spatial gaps - rural areas have few sites
- Temporal quality varies – hourly to weekly
- Emerging with routinely available information (hourly - daily)

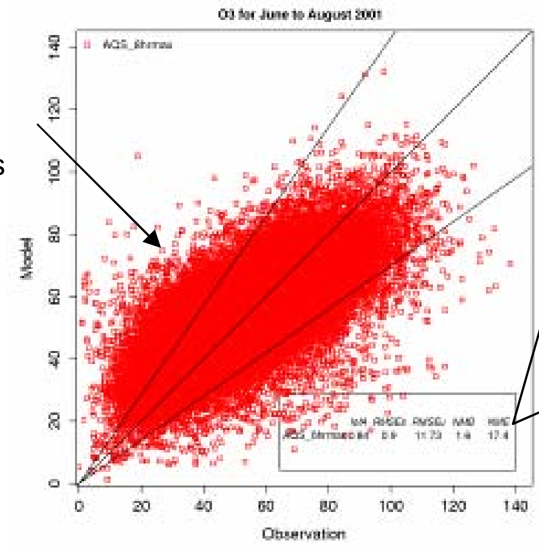


CMAQ and Monitored Ozone Estimates

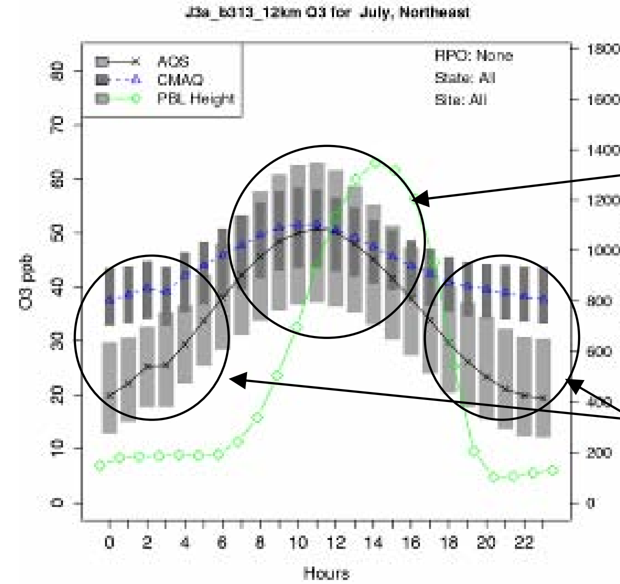


CMAQ and Monitored Ozone Estimates

High bias at low concen- trations 10-50 ppb

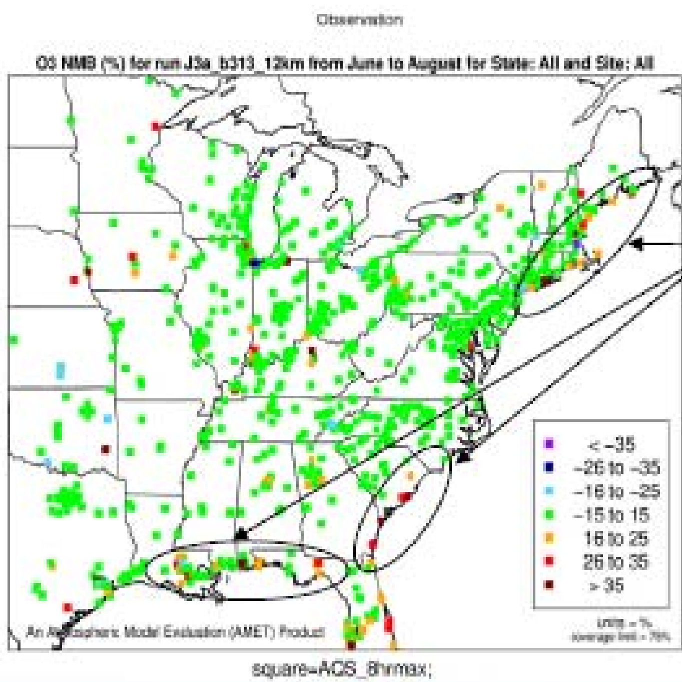


RMSEs = 0.9
RMSEu = 11.73
NMB = 1.6
NME = 17.4

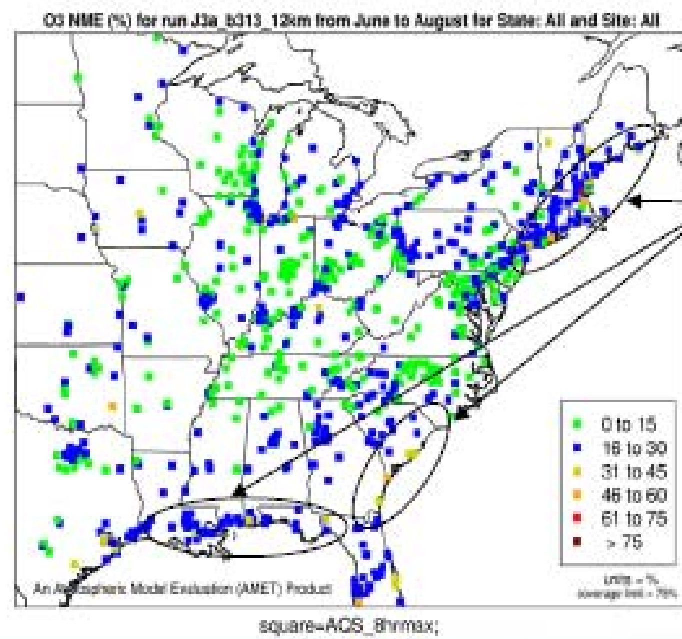


Good Agreement During Mid-day

Model Not Capturing Overnight lows



Higher biases along the coast



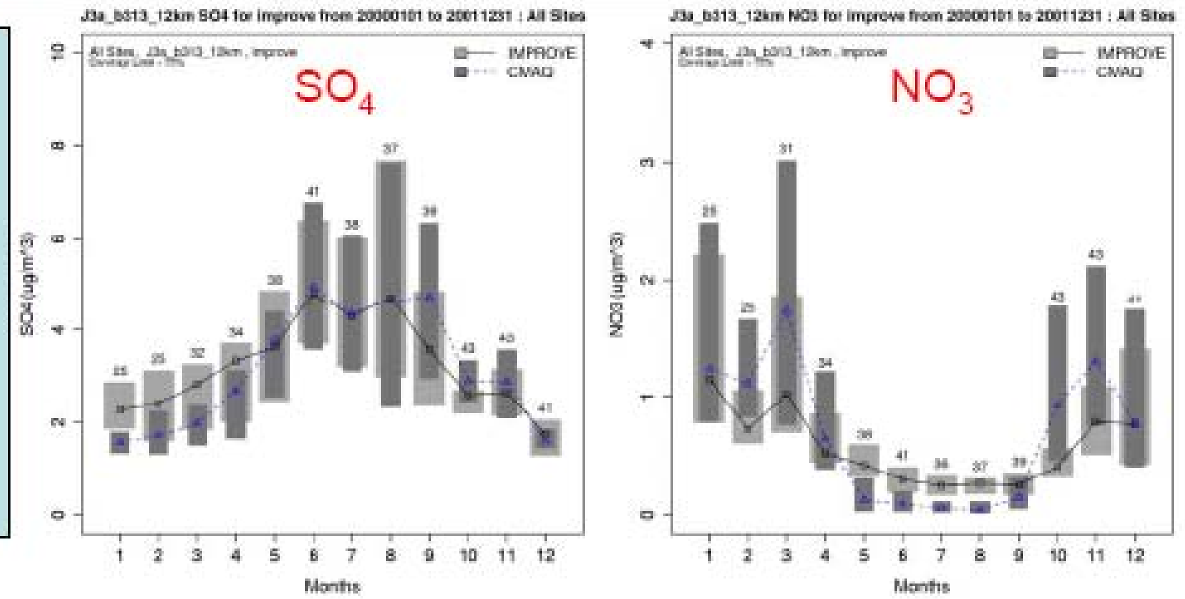
Higher errors along the coast

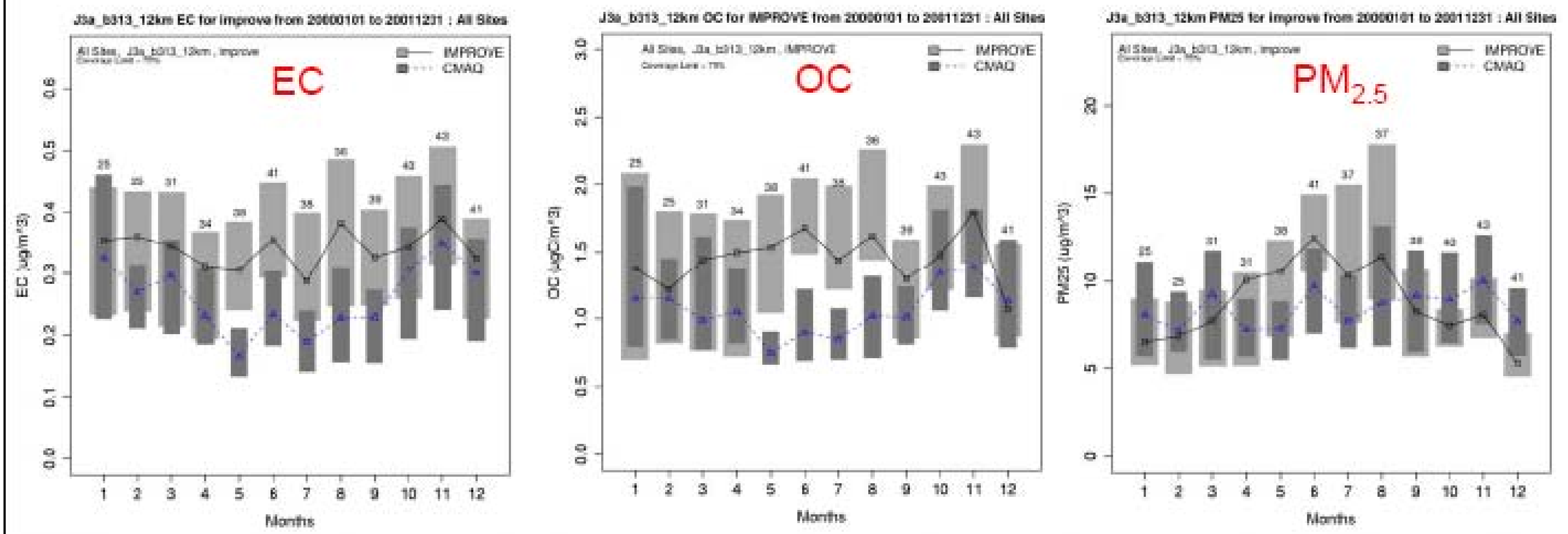
CMAQ and Monitored Fine Particle Estimates

IMPROVE (v4.5, 12km)

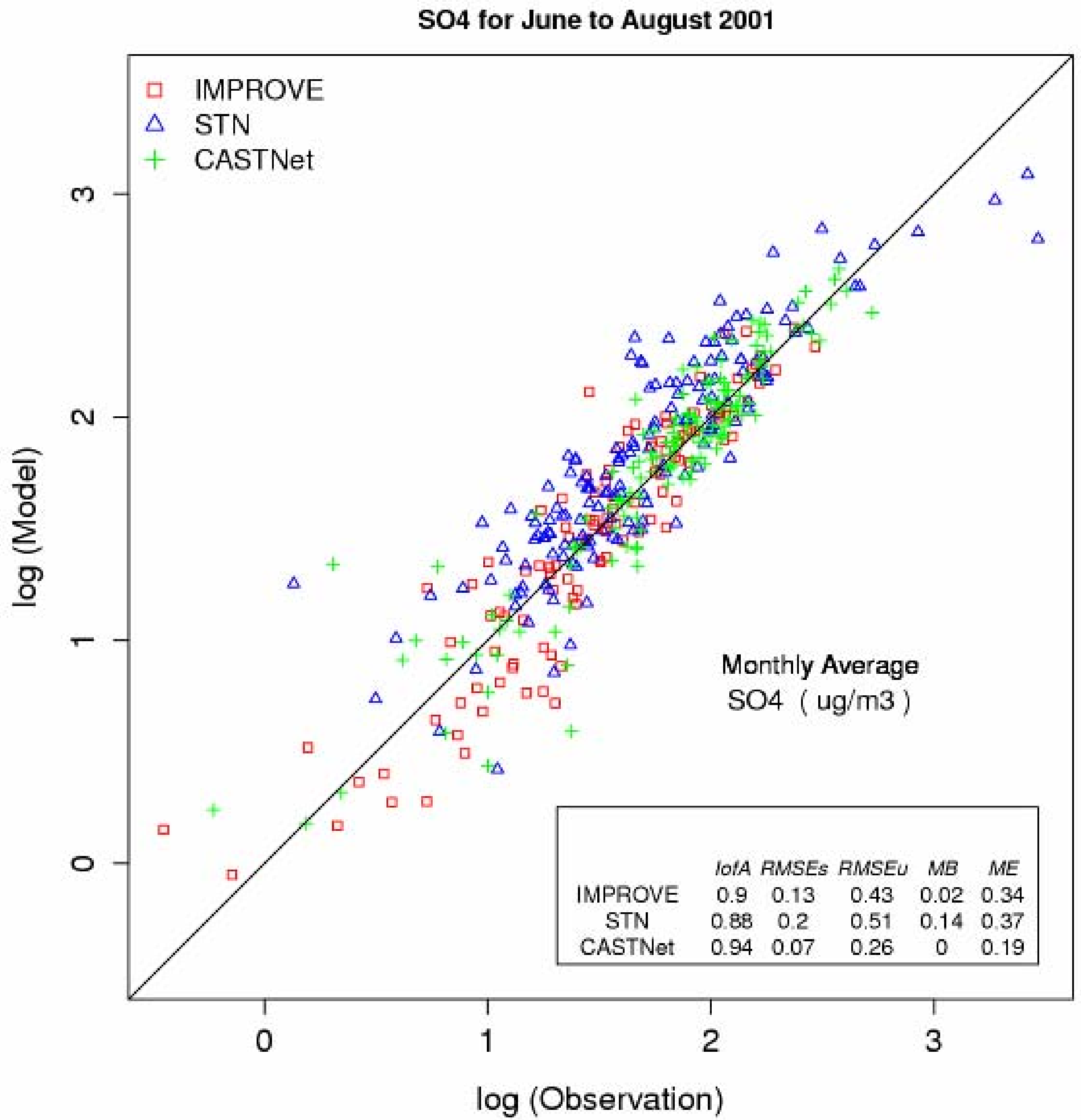
Under-predictions in SO_4, NO_3, EC and OC contribute to under-predictions in $PM_{2.5}$ in the spring and summer.

Over-predictions in SO_4 and NO_3 contribute to over-predictions in $PM_{2.5}$ in the fall.

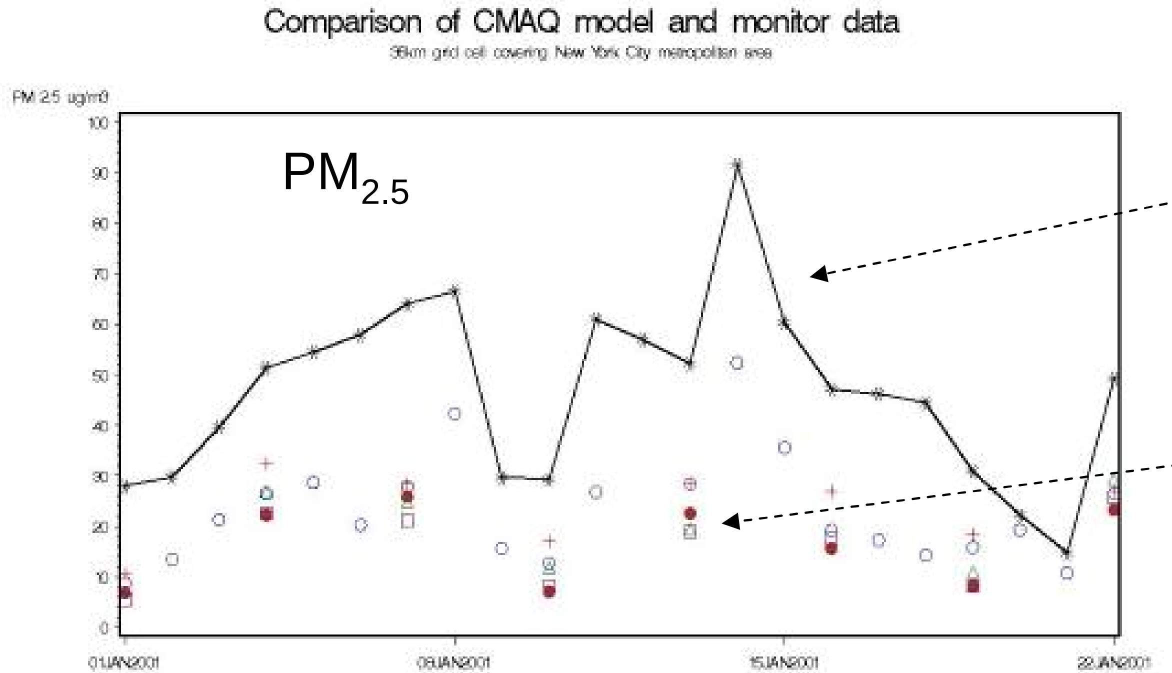


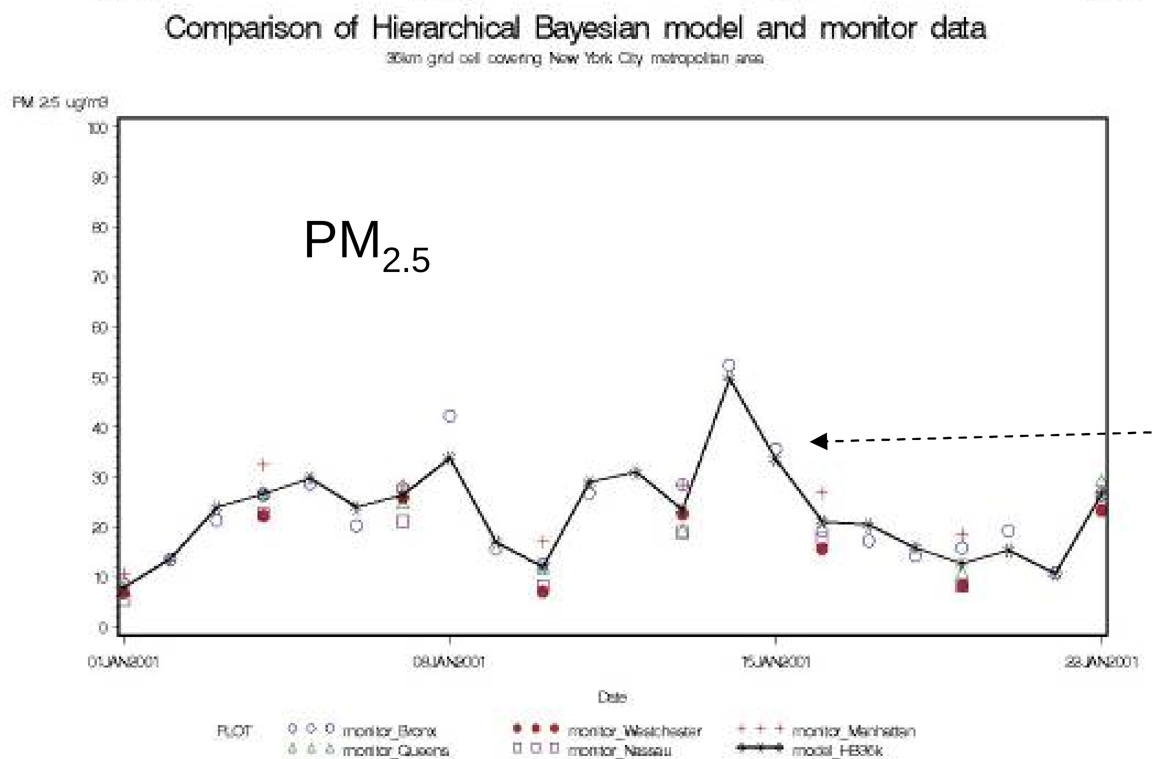


CMAQ and Monitored Fine Particle Sulfates



	IofA	RMSEs	RMSEu	MB	ME
IMPROVE	0.9	0.13	0.43	0.02	0.34
STN	0.88	0.2	0.51	0.14	0.37
CASTNet	0.94	0.07	0.26	0	0.19





CMAQ Prediction

AQ Monitoring Data

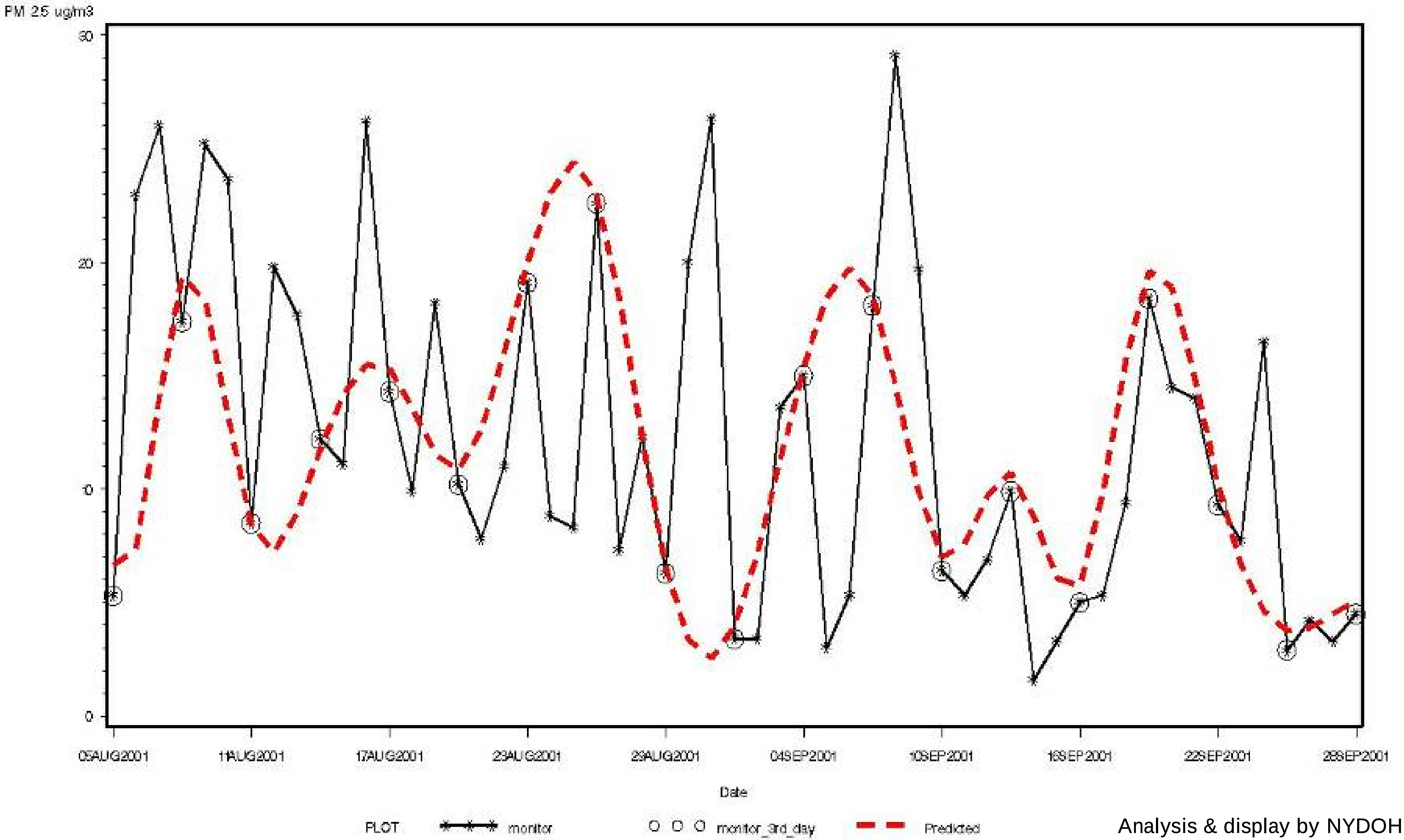
Improving the usefulness of observations or modeled results alone through "fusion" techniques.

After application of fusion technique

Courtesy of NYSDOH

Comparison fine particulate monitor data and interpolated 36km model

Corning, New York



Analysis & display by NYDOH

Root Mean Squared Prediction Errors (RMSPE) Using Three Prediction Surfaces

	O_3			$PM_{2.5}$		
		Combined Model Improvement (%)			Combined Model Improvement (%)	
Prediction Surface	**Overall Estimate**	**Sites 11**	**Days 245**	**Overall Estimate**	**Sites 60**	**Days 78**
CMAQ Output	0.525	91%	64%	0.277	90%	87%
Kriged Monitor Data	0.521	91%	57%	0.165	58%	69%
Bayesian Combined	0.501			0.127		

Remaining uncertainties

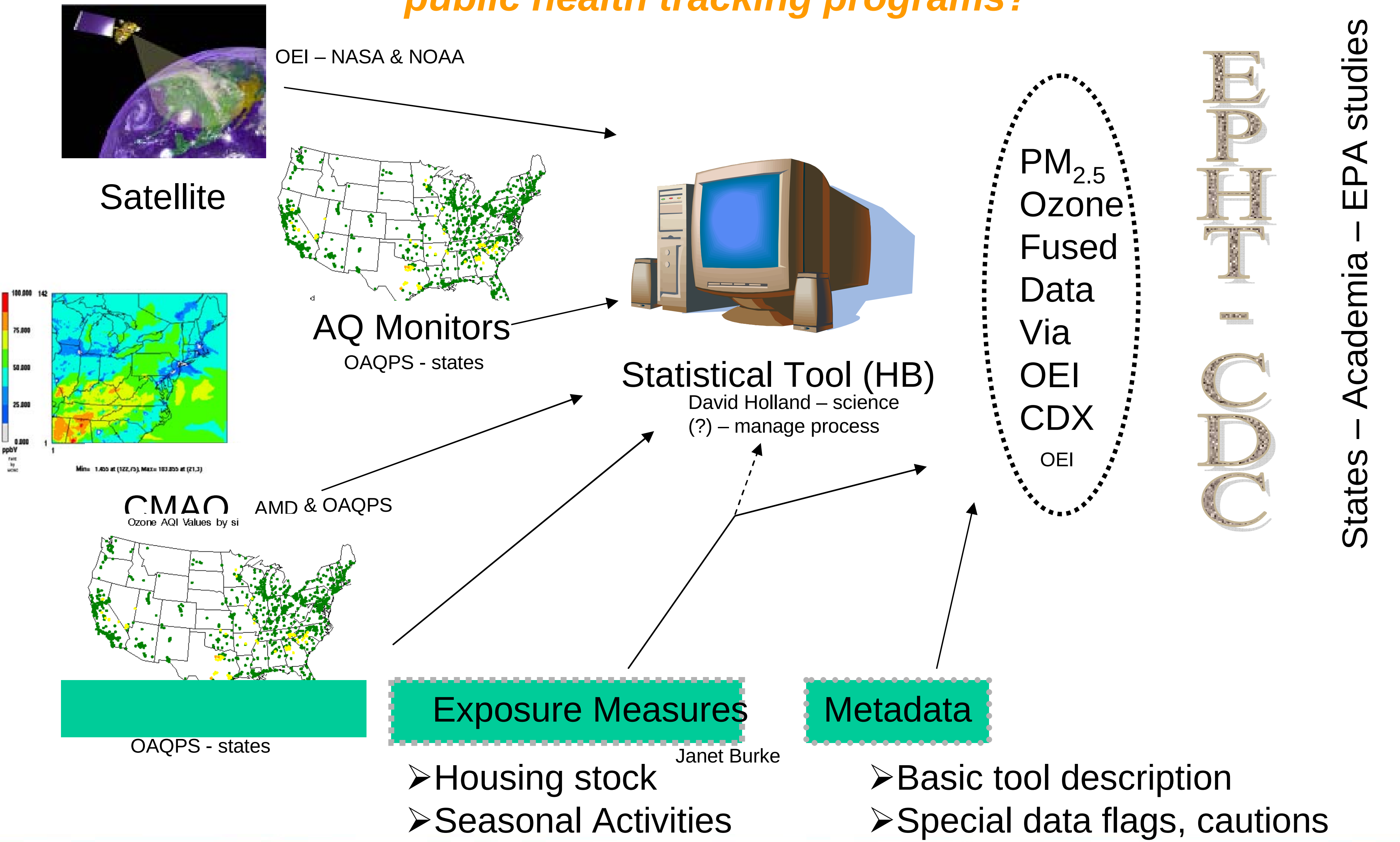
- Ambient vs Personal Exposure
  - Methods estimate ambient concentrations, but . . .
  - People experience health impacts from the air they breathe (i.e., their personal exposure)
  - How do the outputs from the various air quality characterization methods relate to personal exposure?
- Health Outcome
  - Temporal Resolution? => Chronic vs Acute
  - Outcome Measured? => Effect/Endpoint

Personal Monitor

What are the key programmatic and research needs of the existing ambient pollution modeling approaches?

- CMAQ
  - Near real-time forecasting
  - Fire emissions
  - Atmospheric chemistry
  - Other pollutants
- Statistical Models
  - Proper grid size
  - Standardization of best estimate
  - Incorporating other data
  - Shrinking ambient monitoring network
- Availability of Metadata for Interpretations

What new modeling activities are being planned or anticipated that could support future air pollution health effects, accountability or public health tracking programs?



Summary

- PHASE selected …
  - AQ = HB (CMAQ + AQ)
  - Produce metadata for interpretations
- Further research is needed
  - Appropriate grid size
  - Efficiency with AQ-Health Methods

Non reviewed material – do not cite or quote