Diagnostic and In-Depth Model Evaluation

Robin Dennis*
Atmospheric Sciences Modeling Division
NOAA - Air Resources Laboratory
Research Triangle Park, NC

* On assignment to the National Exposure Research Laboratory, U.S. EPA.

CMAQ Model Peer Review Meeting
R.T.P., NC
December 17, 2003

Objectives of Diagnostic/In-Depth Model Evaluation

- Test the model to check
  - Reliability of the Predictions (Right Reason)
    - Right answer for the right reason
    - Wrong answer for the right reason or understandable reason
  - Right Response
    - Reasonably accurate response (a major focus of the work)
- Separate sources of error
  - Discern among:
    - Emissions input error
    - Meteorological error
    - Chemistry/aerosol physics and chemistry error
- Aid model developers in identifying and treating problem areas

Approach: Thinking and What is Behind It

- Advance the science underpinning for diagnostic or investigative evaluations
- Assess the model through evaluations

Advance the Science

- Indicator Measures to probe processes
  - Development
  - Application and testing
- “Instrumented” Models
  - Process Analysis
  - Tracking model versions
    - Sulfate tracking
    - OC source apportionment
- Statistical quantification of the testing
  - Tests associated with indicator metrics
  - Tests associated with space-time comparisons (new)

Assess the Model

- CMAQ (3-D) Model Comparisons
  - Comparisons against special (supersite) data sets
    - Complete species set (simultaneous multiple species)
    - High temporal resolution (hourly)
  - Sensitivity analyses
    - Model against data
    - Model against model
      - Process testing focus
      - Control strategy response focus
- Box-model Comparisons
  - Photochemical box models
  - PM box models
- Apply CMAQ to gain experience with what it does

Approach (cont.)

- Limitations: Only so much we can cover at one time. Try to pick important areas for testing that will provide insight into the model processing

- Focus on predictive use of CMAQ for control strategy assessment
  - What are the implications for control strategy predictions?

- There has been a progression over time from Ozone to PM and then towards in-depth PM + Ozone accountability.

- Effort has transitioned from evaluators working alone to working in concert with development efforts

Lessons Garnered / Guidance Developed

What are some of the lessons we have garnered or guidance we have developed out of the diagnostic or in-depth evaluation efforts?

Photochemistry

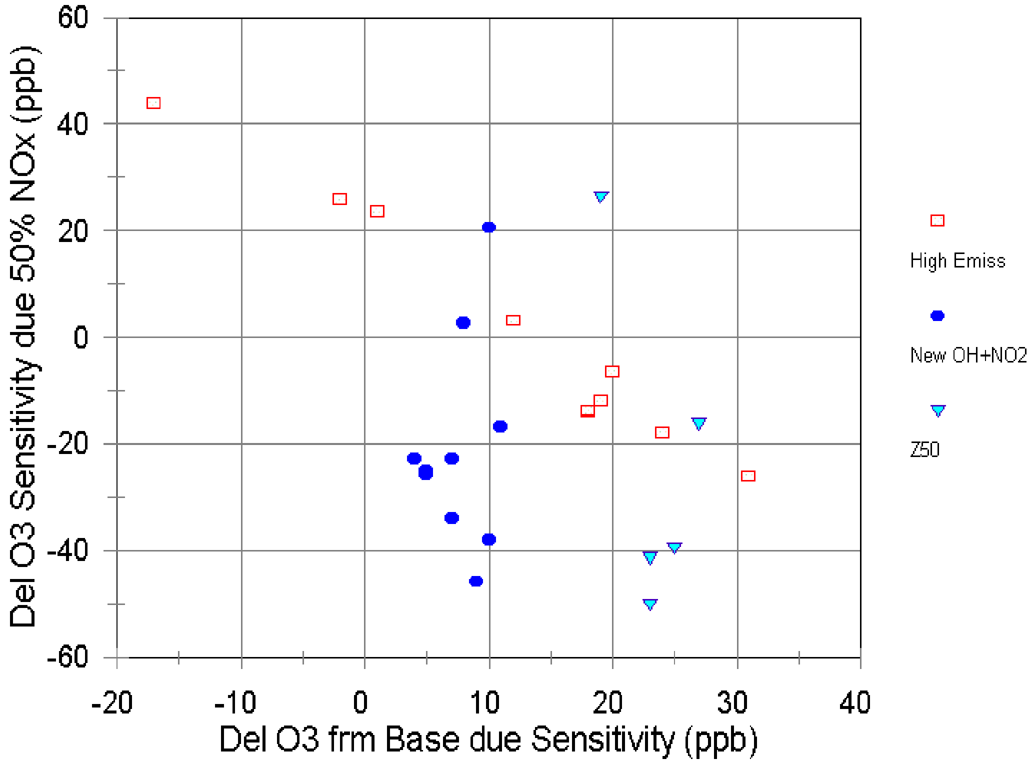
Fine Particle Predictions

Photochemistry: sensitivity vs. control

Clear sense now that $\Delta O_3)_{sens} \neq \delta O3/\delta E)_{sens}$

Therefore, we must test the sensitivity to the control response directly through sensitivity studies

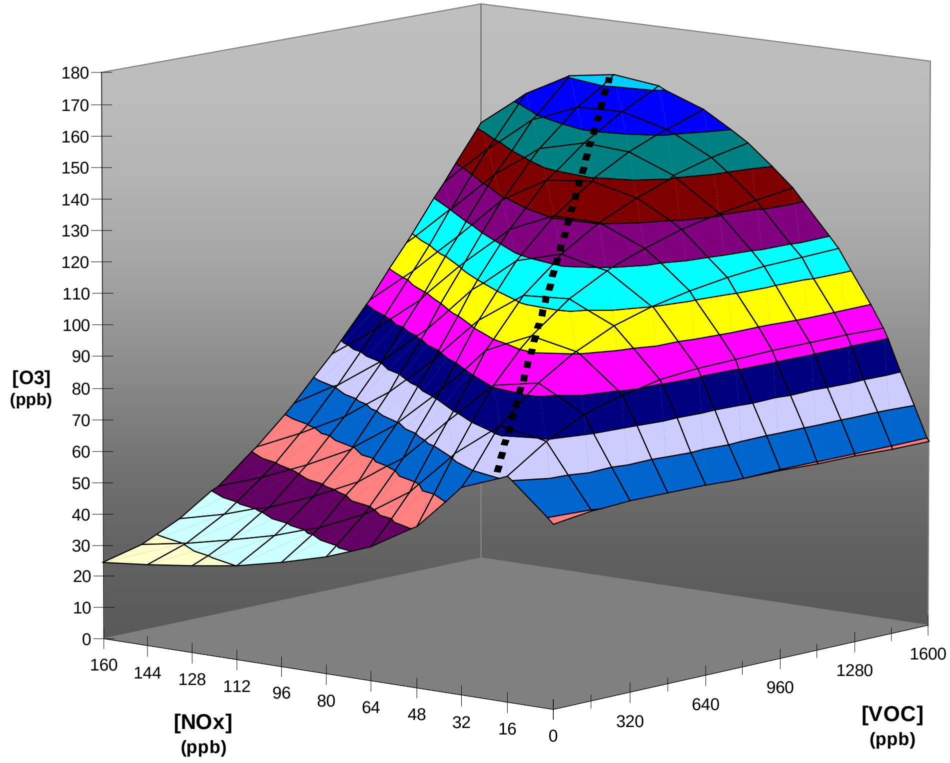
We cannot assume that better performance on O3 base cases translates to better performance relative to control strategy predictions



Photochemistry: Indicators of process functioning

Establish use of indicators for photochemical functioning through diagnostic studies; to see how well chemistry in model is replicating chemical processing we observe in the atmosphere

- O_3/NO_X to indicate system location relative to the ridgeline



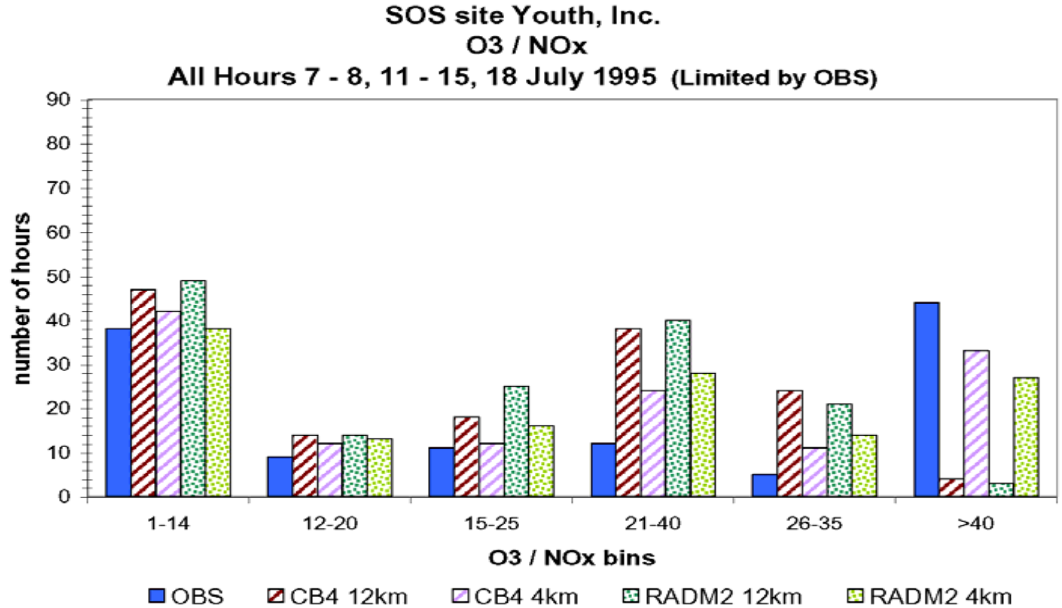
- O_3 vs. NO_Z to indicate degree of photochemical processing (NO_X cycle)

Photochemistry: O_3/NO_X bins

Exploring system state:

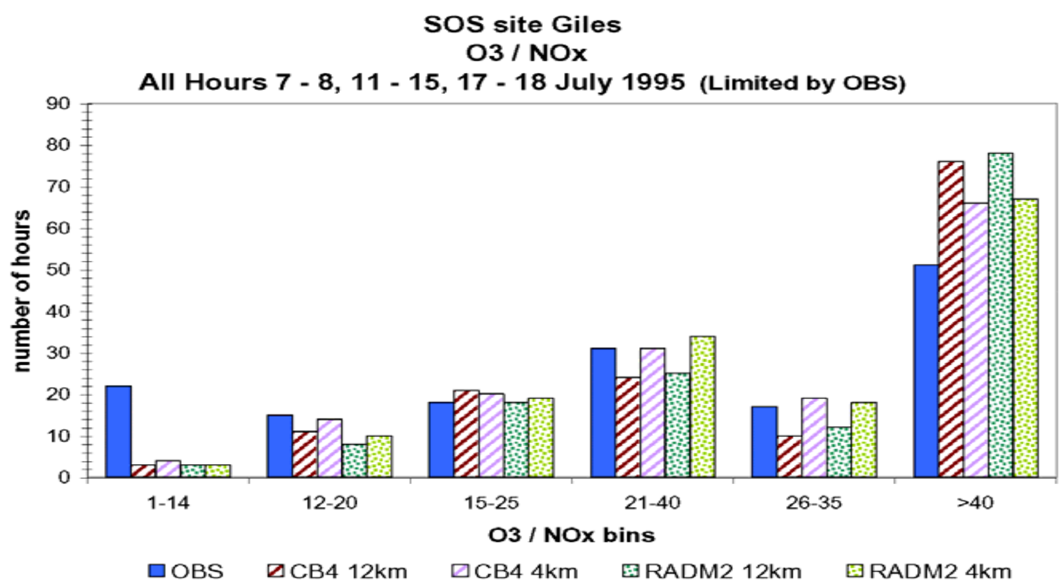
Not much difference between RADM2 & CB4

4-km model is better



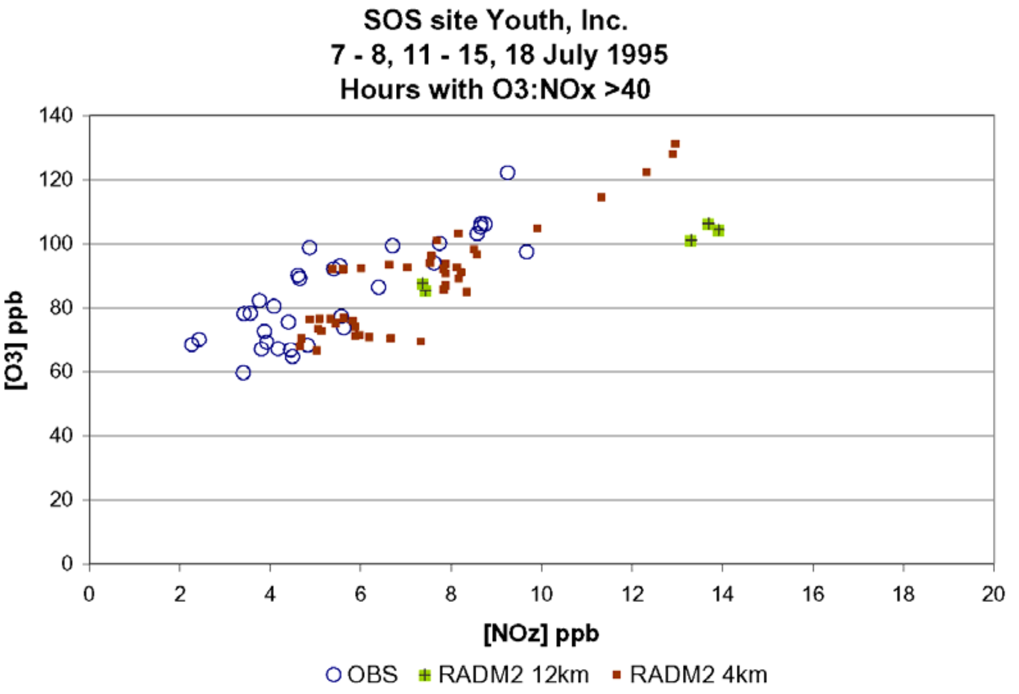
Local NO_X emissions are missing in the model

May throw model off for control predictions

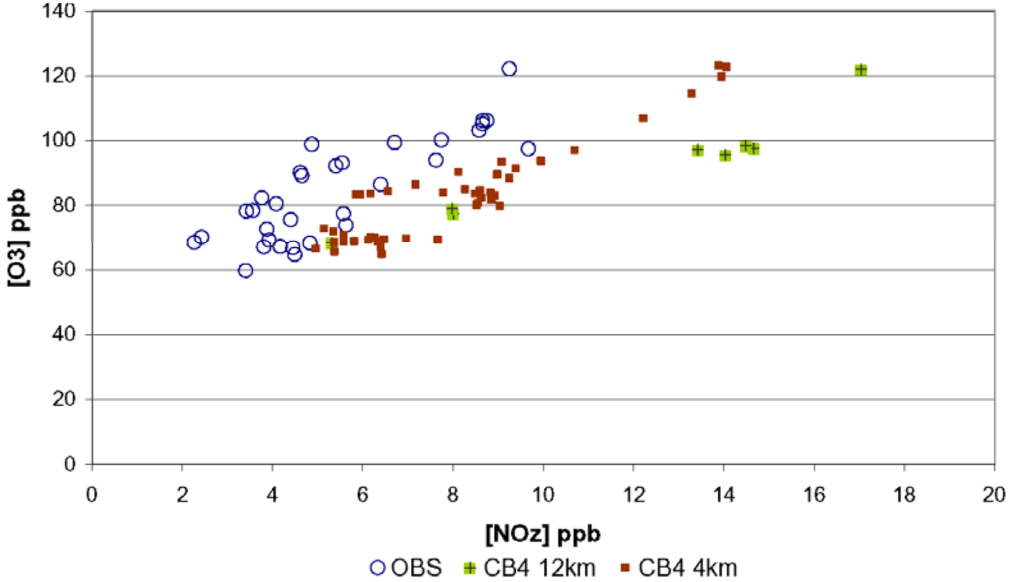


Photochemistry: O3 vs. NOZ   (NOX cycling)

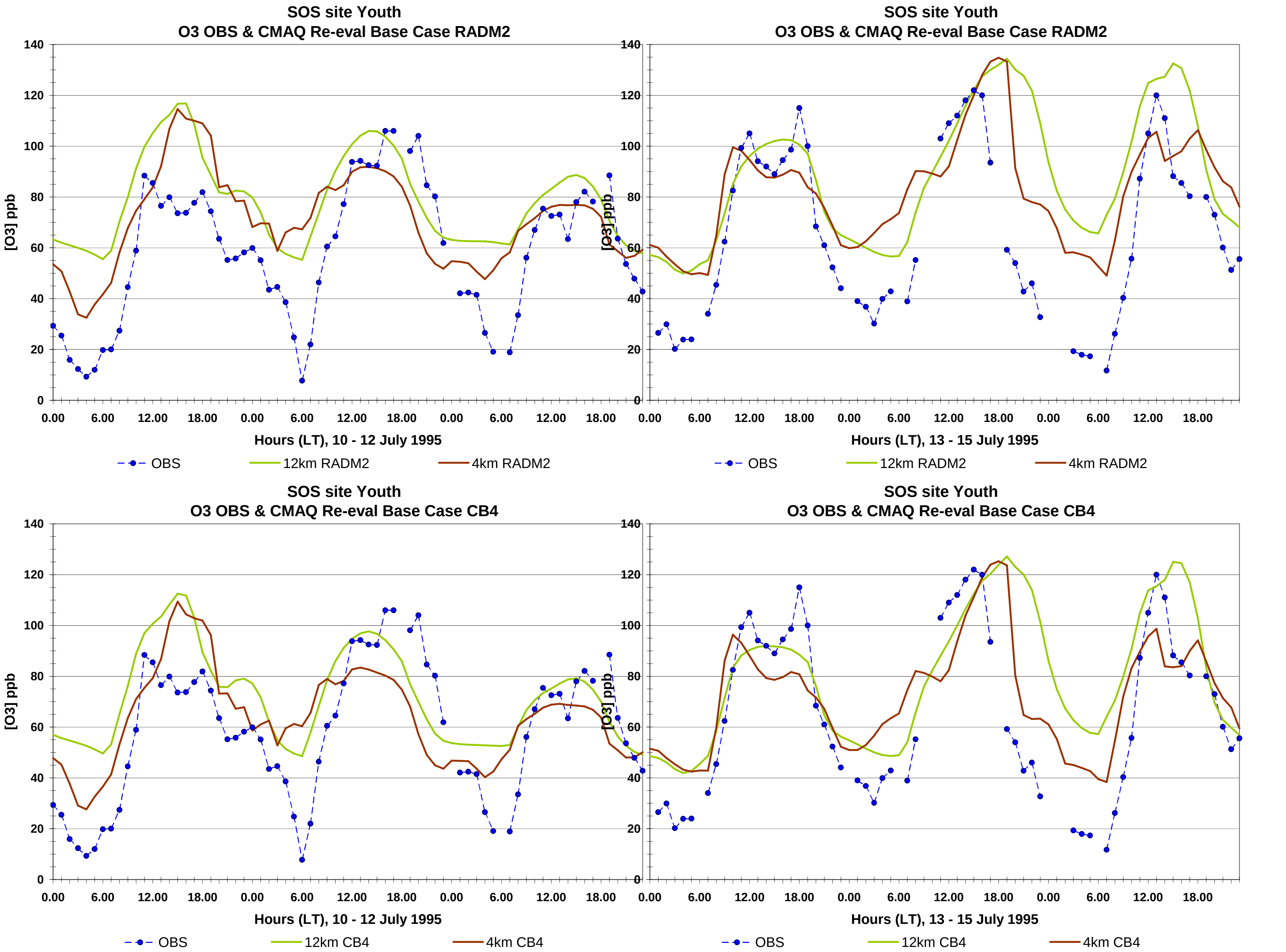
Exploring differences between RADM2 and CB4



4-km is better



RADM2 looks better

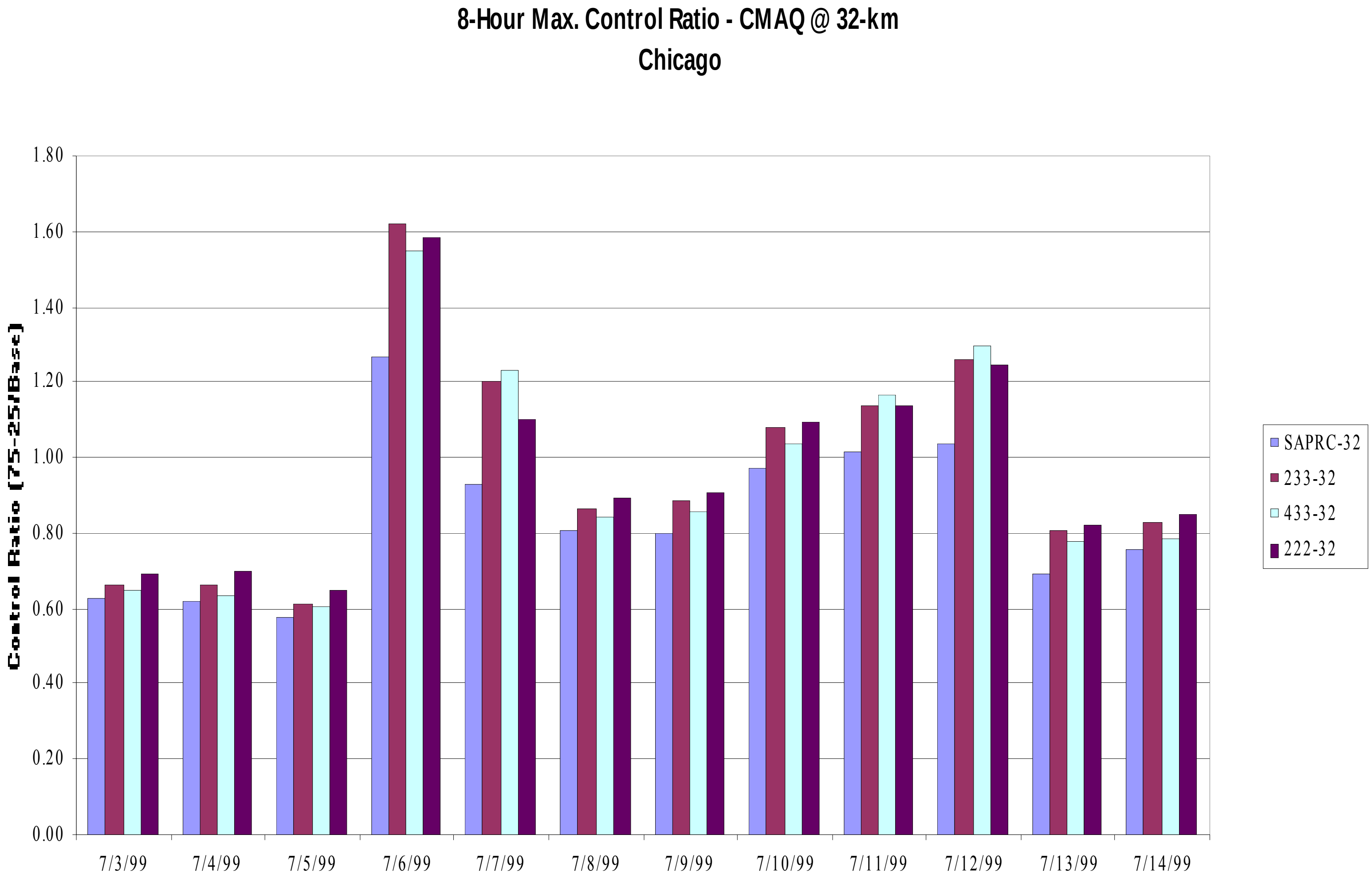

SOS site Youth
O3 OBS & CMAQ Re-eval Base Case RADM2
[O3] ppb
Hours (LT), 10 - 12 July 1995
OBS
12km RADM2
4km RADM2
SOS site Youth
O3 OBS & CMAQ Re-eval Base Case RADM2
[O3] ppb
Hours (LT), 13 - 15 July 1995
OBS
12km RADM2
4km RADM2
SOS site Youth
O3 OBS & CMAQ Re-eval Base Case CB4
[O3] ppb
Hours (LT), 10 - 12 July 1995
OBS
12km CB4
4km CB4
SOS site Youth
O3 OBS & CMAQ Re-eval Base Case CB4
[O3] ppb
Hours (LT), 13 - 15 July 1995
OBS
12km CB4
4km CB4

Photochemistry: Indicators of process functioning

We have been providing feedback to the measurement community about the need for NO_Y and true-NO_2 measurements (for NO_X) to support these diagnostic metrics to support better examination of the photochemical processing in air quality models

Now pushing on more reliable HNO_3

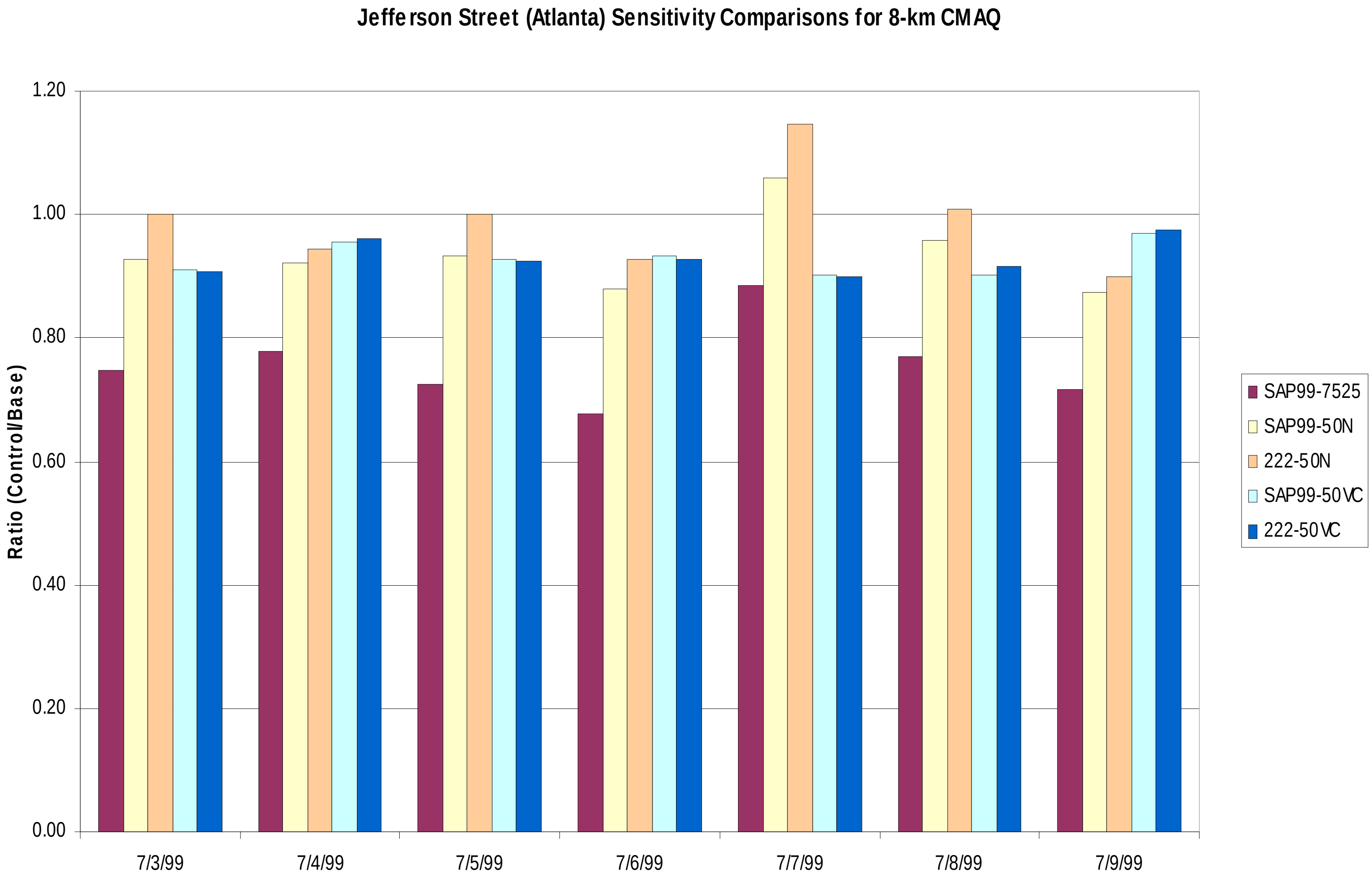
Photochemistry: NO_X Control strategy response



Differences among SAPRC99, RADM2 and CB4

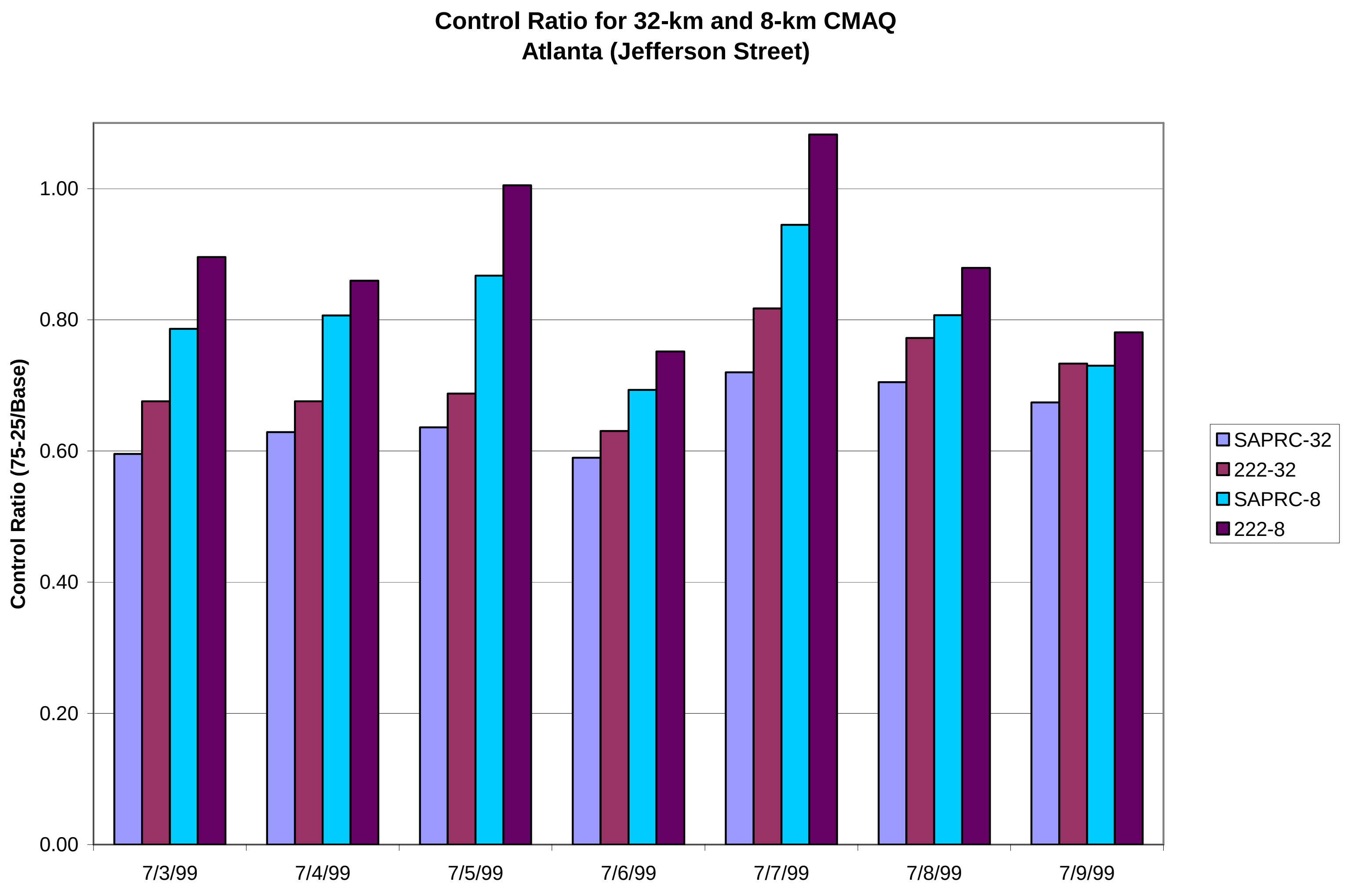
Differences more a function of “age”, not related to $OH+NO_2$ rate

Photochemistry: Control strategy response



Differences are in the NO_X control response, not the VOC response

Photochemistry: NO_X Control strategy response



NO_X Control effectiveness goes down with smaller grid size
Difference between old and new mechanism increases

Photochemistry: Indicators of process functioning

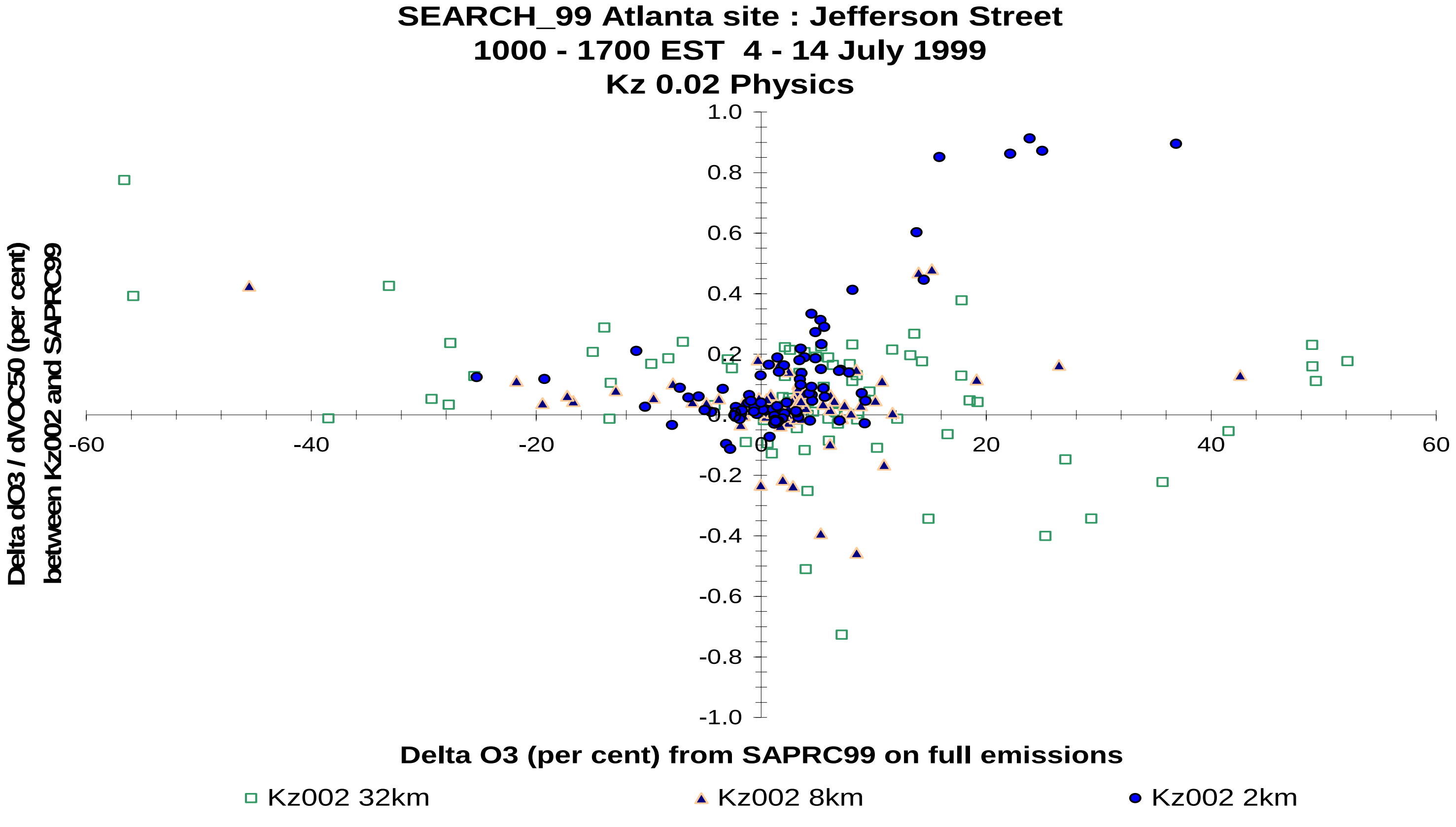
We want to close the loop between these differences in control strategy response and the O_3 vs. NO_Z test, using our matrix of '99 sensitivity runs

Grid	SAPRC99 Chemistry								CB4	RADM 2	RADM 2a
	Base Kz Mixing				ACM		Kz_{lim}=0.002		Mix-ing	Mix-ing	Chem
	Base	50% NOx	50% VOC	75%N 25%V	Base	Control Sensitiv	Base	Control Sensitiv	Base +Sens	Base+ Sens	Base Kz +Sens
32km	15d	15d	15d	15d	15d	45d	15d	45d	180d	180d	60d
8km	15d	15d	15d	15d	15d	45d	15d	45d	180d	180d	60d
2km Nash	15d	15d	15d	15d	15d	45d	15d	45d	180d	180d	60d
2km Atl	15d	15d	15d	15d	15d	45d	15d	45d	180d	180d	60d

Total = 2,400 simulations days (= 6-3/4 years of days) to analyze

Photochemistry: Mixing & Control strategy response

Differences in the nighttime vertical mixing rate (K_Z) can make a large difference in the ozone performance statistics at night, especially in the a.m.. The day-time O3 levels can be changed by +/– 20% or more (horizontal axis). Sensitivity results show that the relative change in daytime ozone due to a control strategy (in %) is generally modified by less than half of a percentage point (vertical axis)..

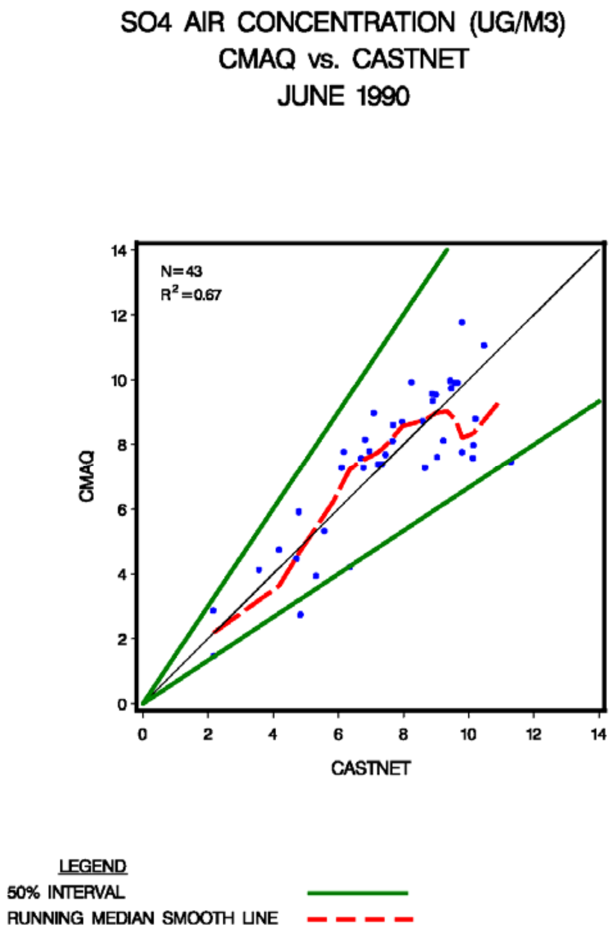


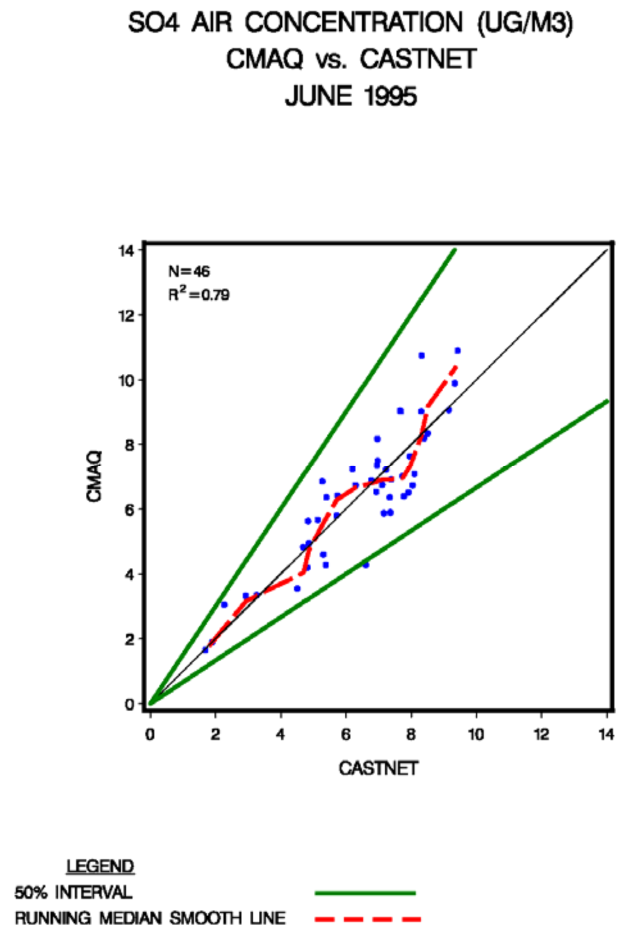
Fine Particle Predictions

Stepping Through Issues

- Sulfate response to SO_2 reductions
- Ammonia inverse to establish appropriate NH_3 emissions
- Separate chemical and physical issues using Atlanta99 supersite data
- Probe nitrate aerosol overprediction problem
- Box-model examination of nitrate formation
- Contemporary vs. old carbon and carbon emissions inventory

Sulfate: Nonlinear Response

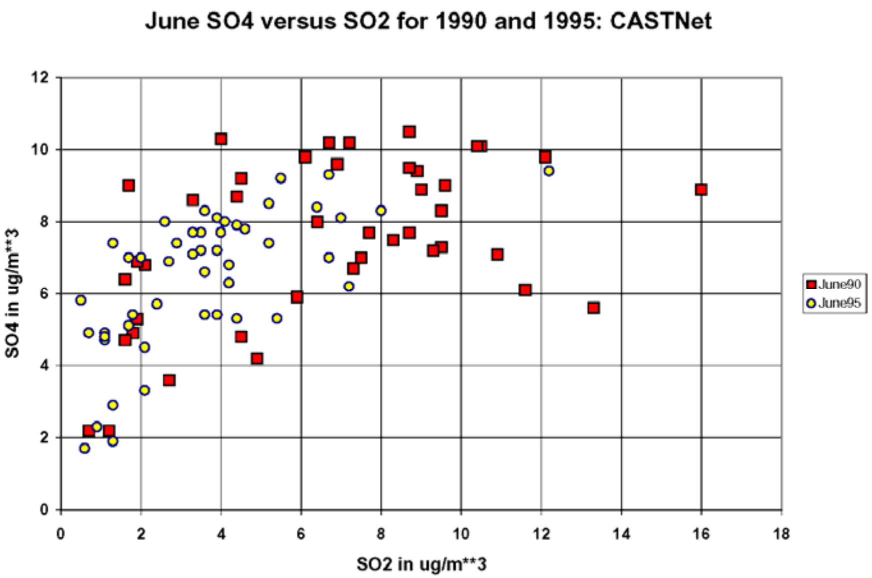


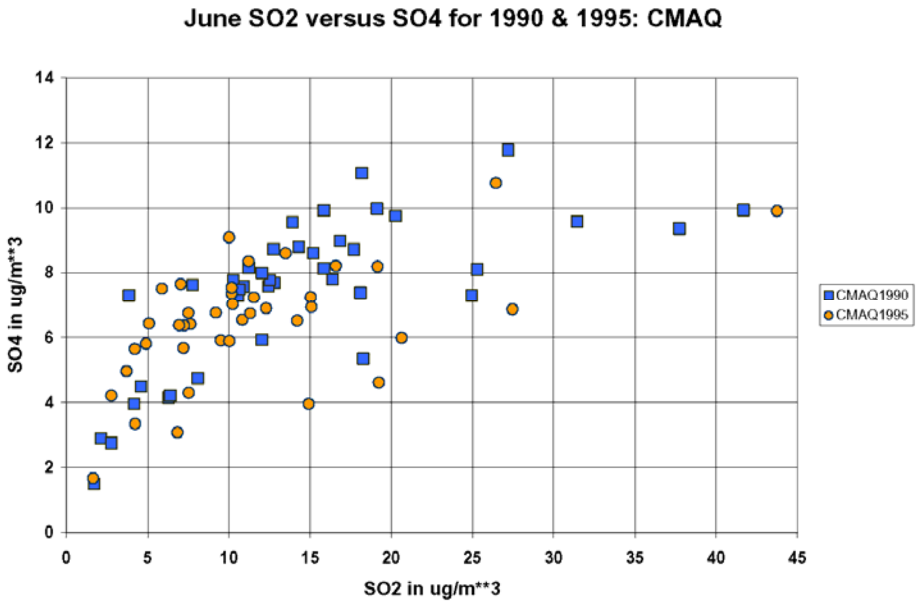


Model correctly predicted $SO_4^{=}$ response = $0.7*SO_2$ reduction

Two sources of nonlinearity: chemistry and meteorology

Sulfate: Nonlinear Response - Chemistry

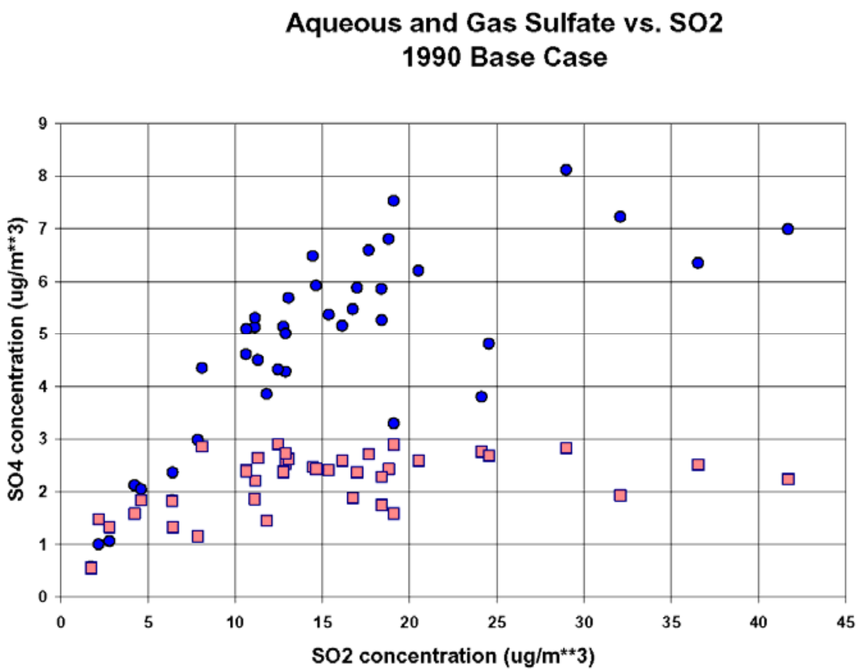


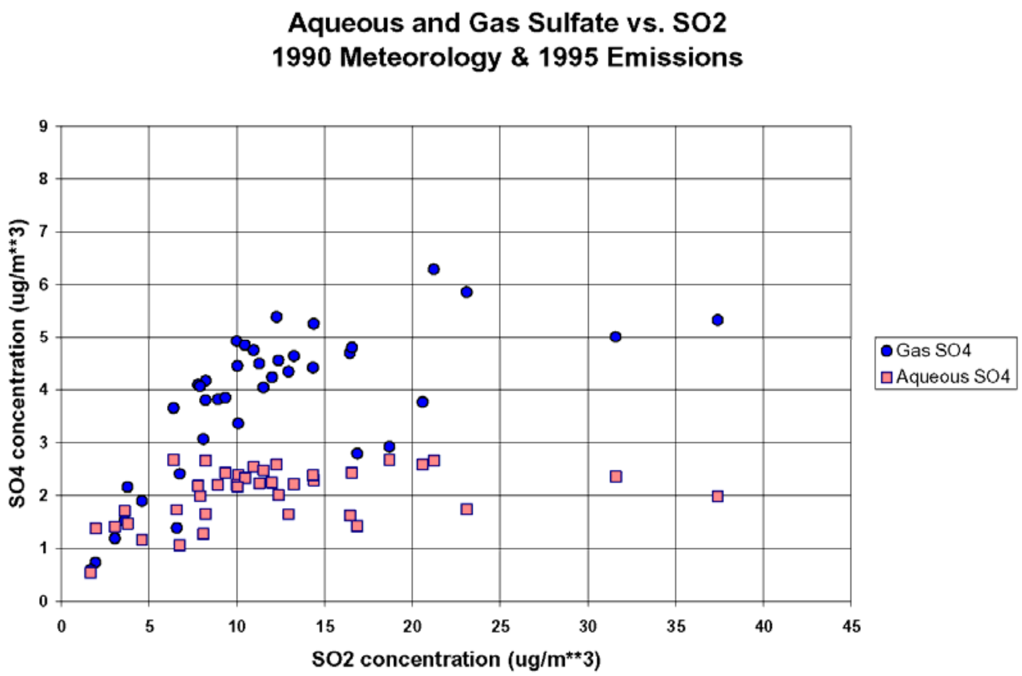


Production of $SO_4^{=}$ rolls off as get to higher concentrations of SO_2

Observed in both model and measurements

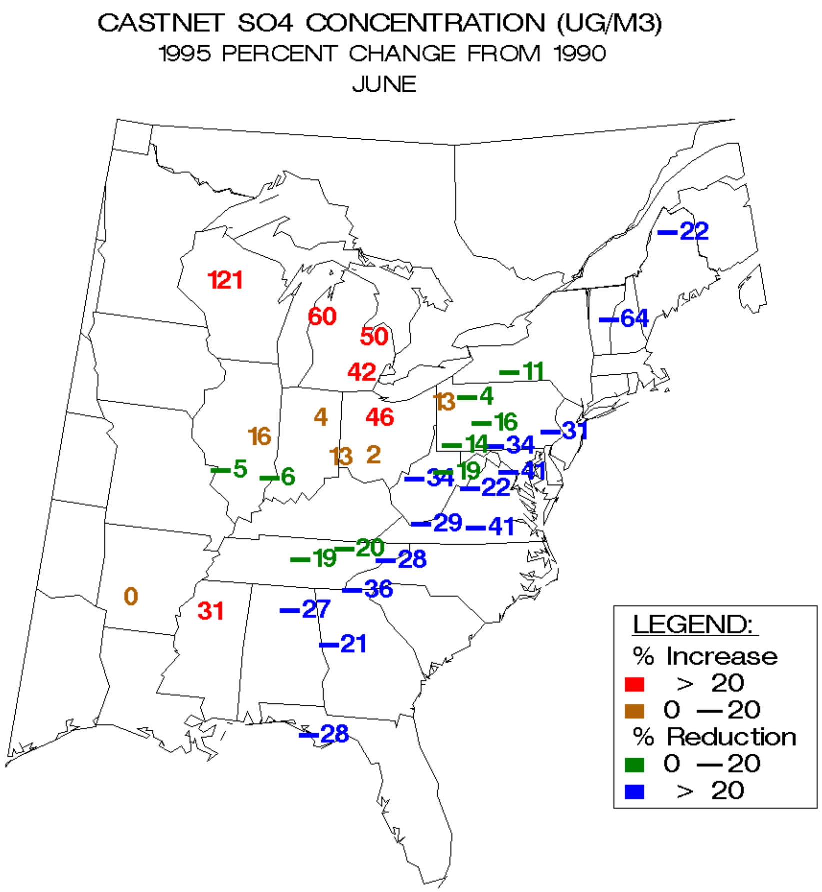
Sulfate: Nonlinear Response - Chemistry

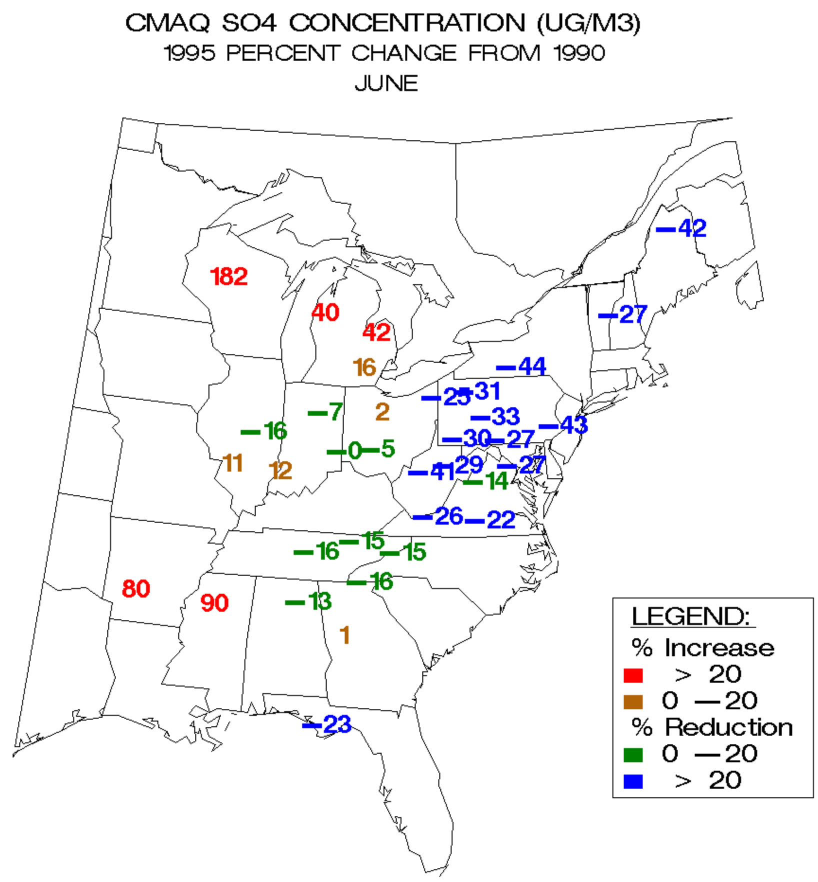




The sulfate tracking model shows that, indeed, the $SO_4^{=}$ roll-off is due to the oxidant limitation in the aqueous phase chemistry

Sulfate: Nonlinear Response - Meteorology

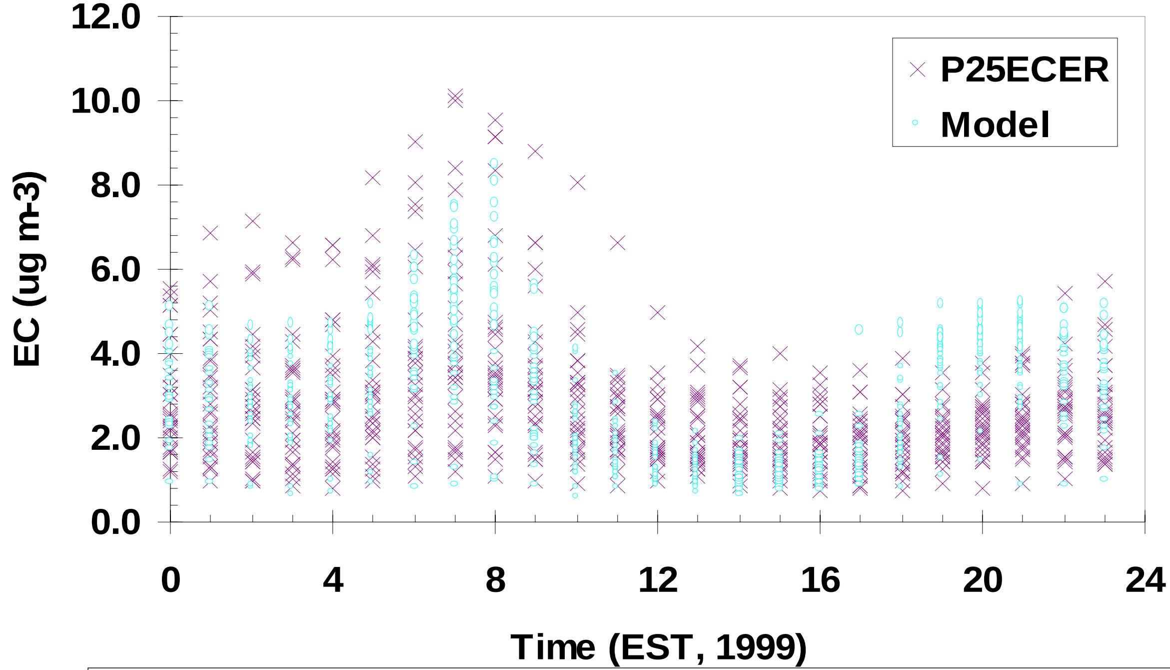




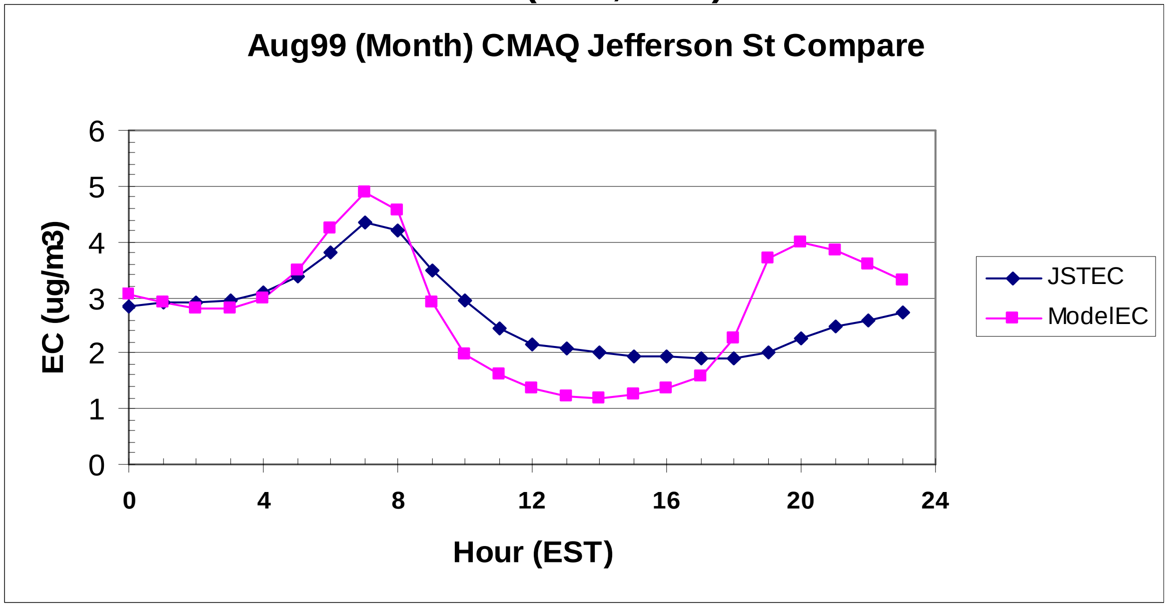
Differences in monthly-average precipitation (cleansing) can result in $SO_4^{=}$ increases even though the SO_2 emissions decreased and CMAQ + MM5 is able to capture this phenomenon

Separating Chemical and Physical Issues

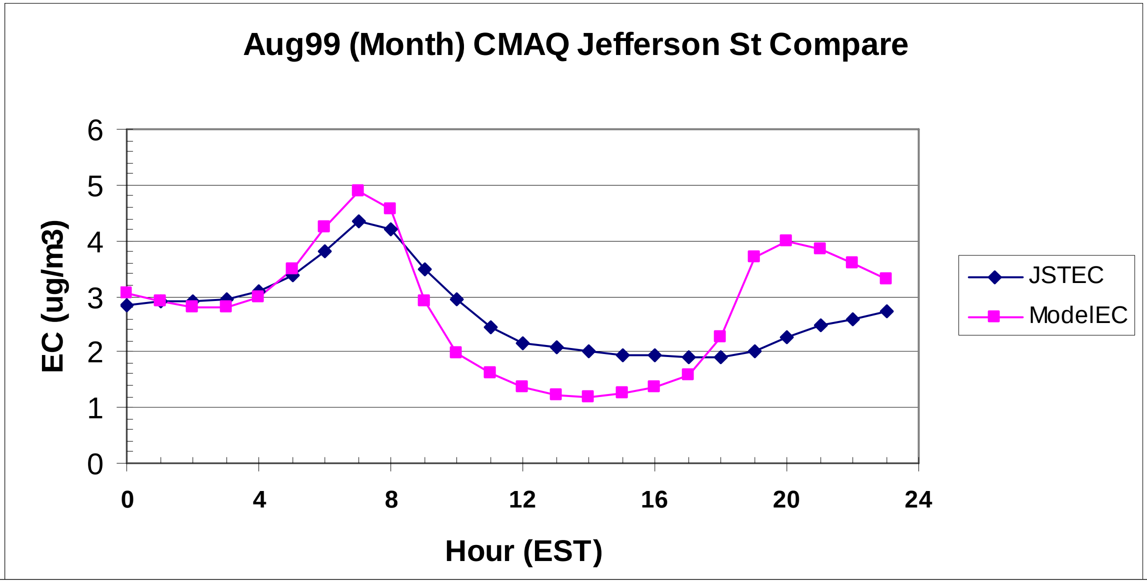
Found again our persistent issue with the nighttime pbl

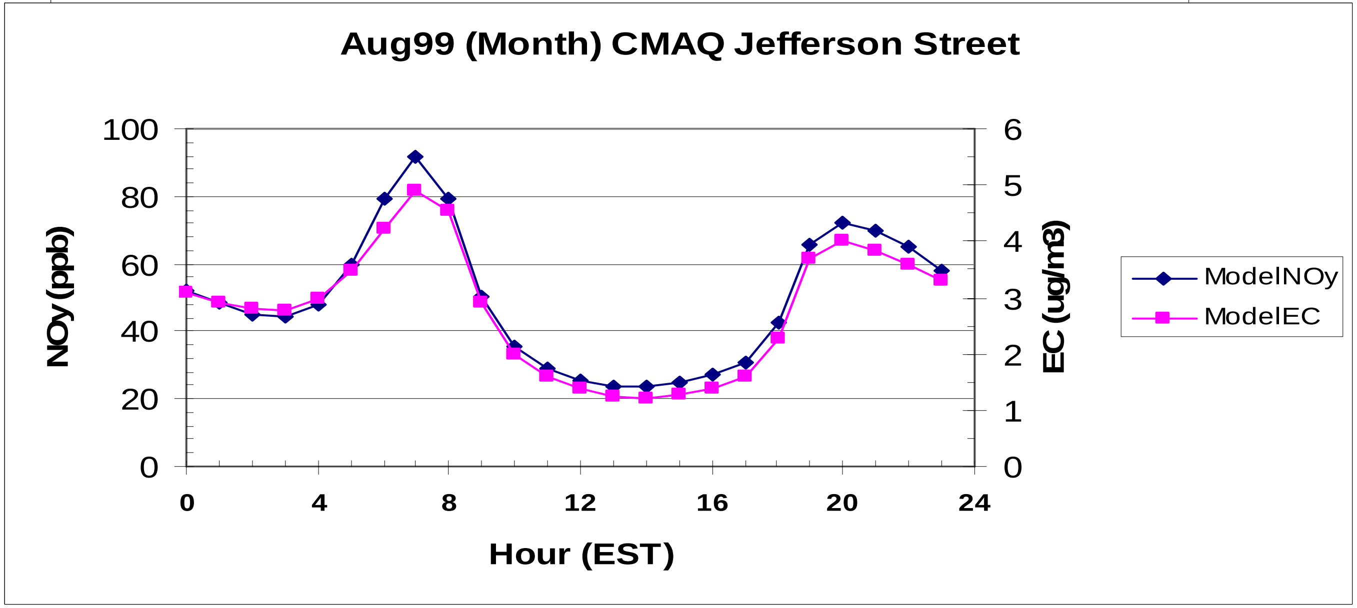


Supports earlier recommendation regarding Kz_{lim}=1 as current default



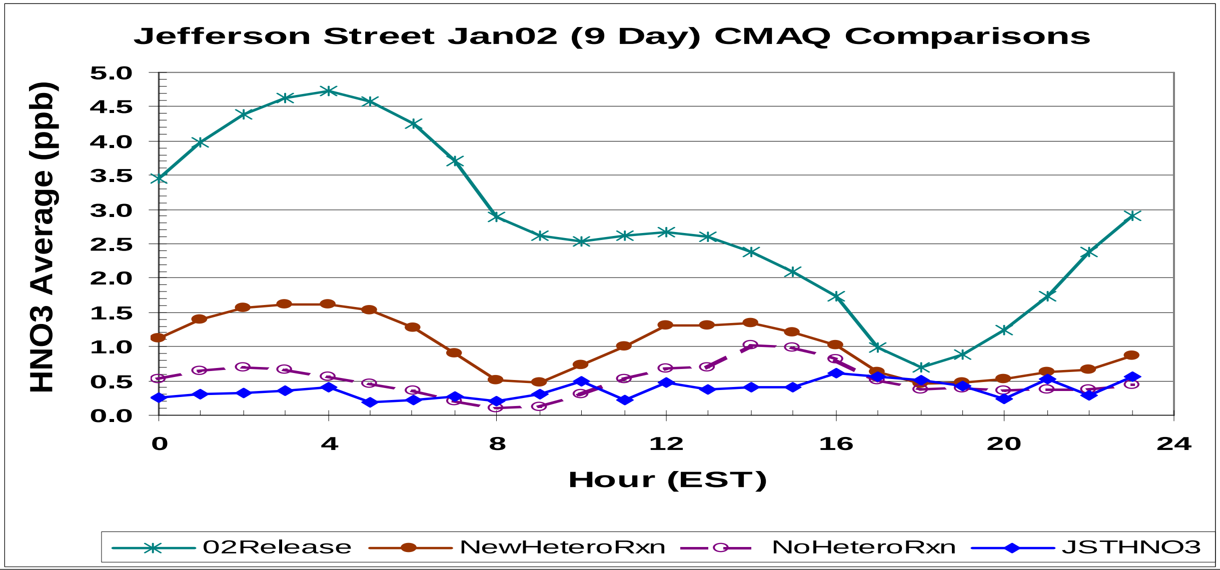
Separating Chemical and Physical Issues

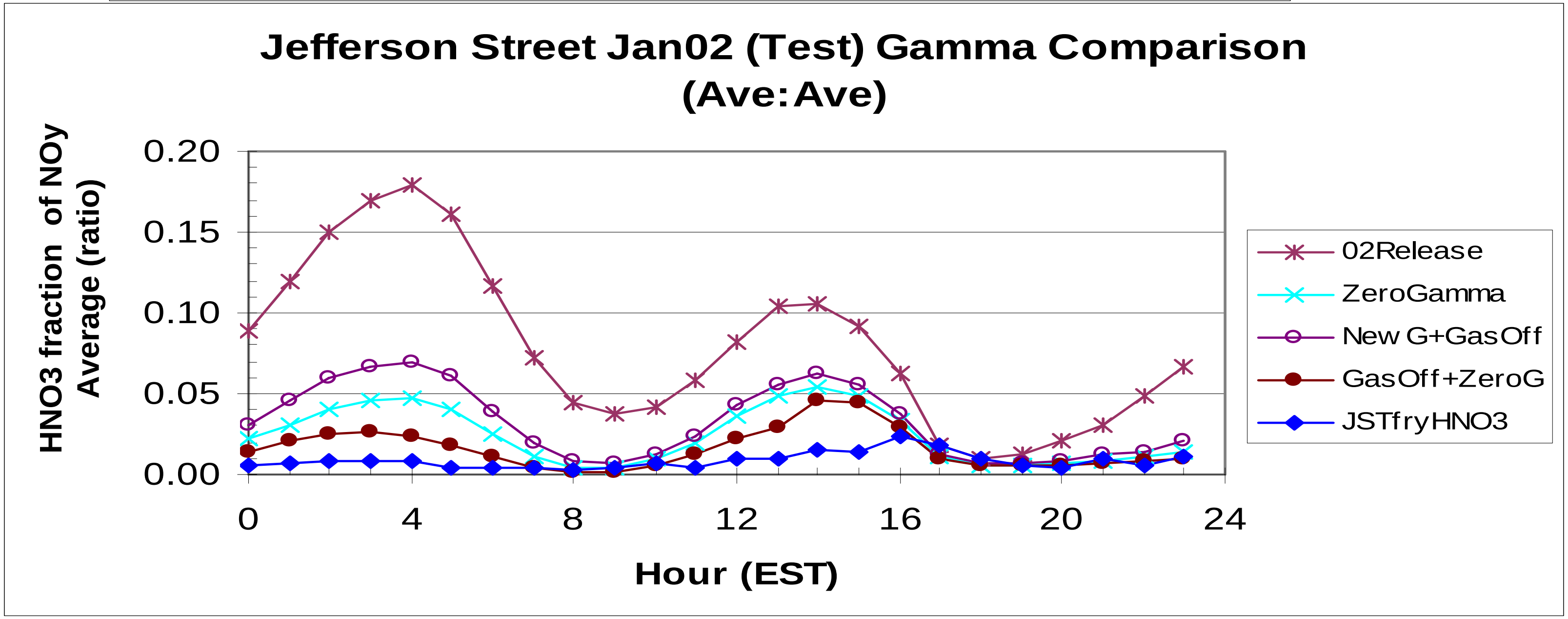




Behavior is consistent for conservative surface emission sources

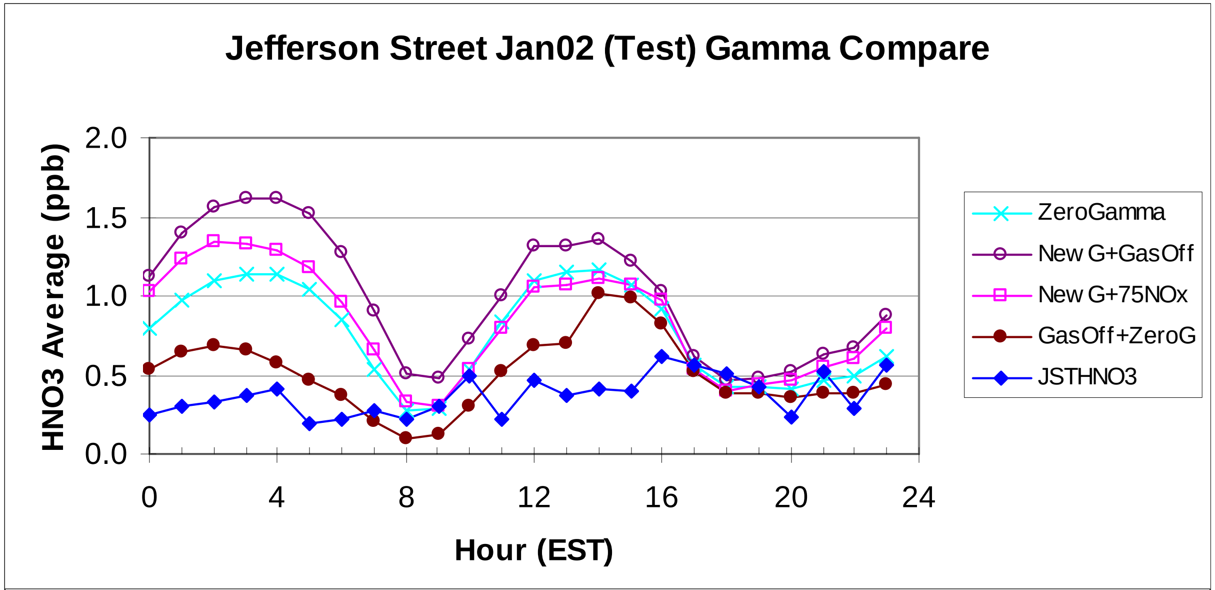
Nitrate Aerosol Issue

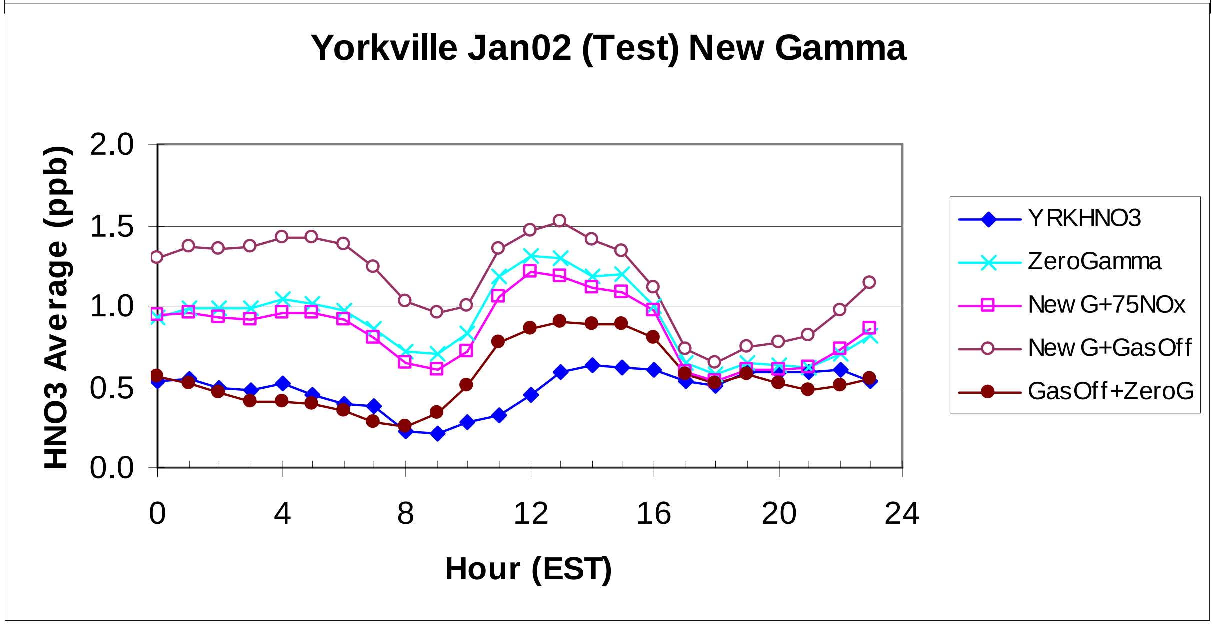




Determined serious over prediction of HNO3 stemmed from nighttime heterogeneous N_2O_5 chemistry

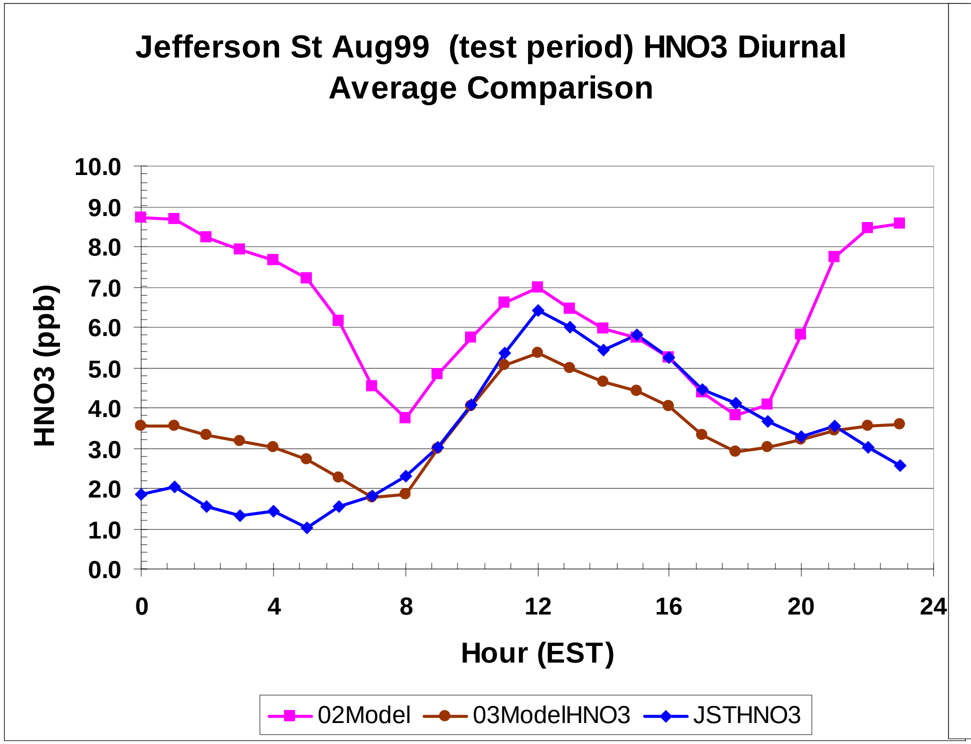
Nitrate Aerosol Issue

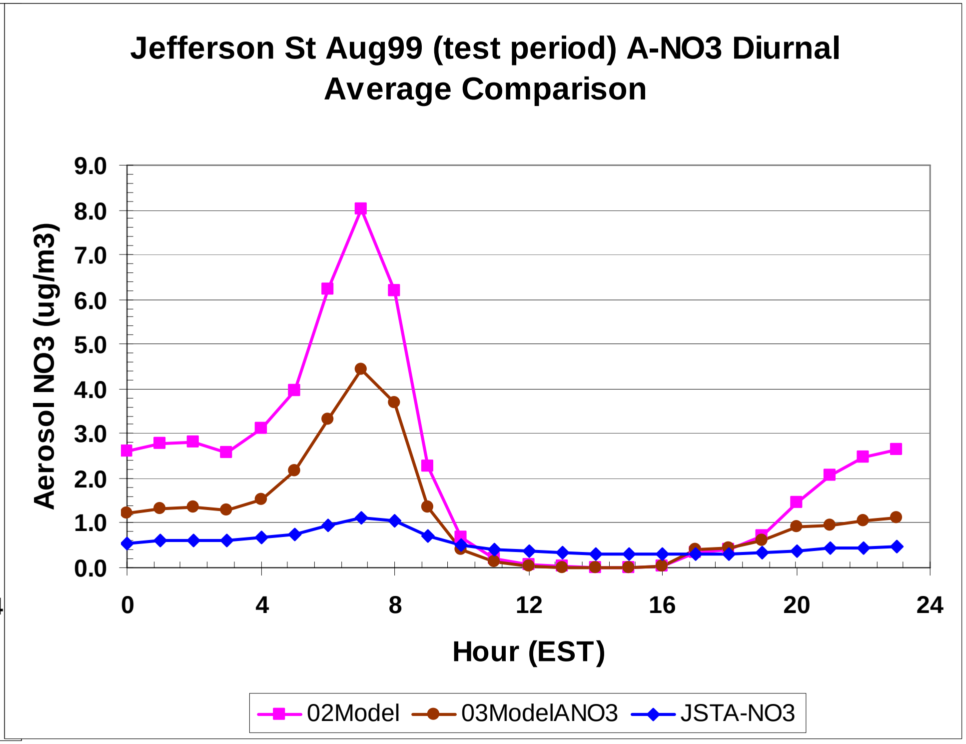




Even with no N_2O_5 conversion there would still be a daytime HNO_3 production issue

Nitrate Aerosol Issue





Re-check found that a larger portion of HNO_3 bias than expected was due to nighttime heterogeneous N_2O_5 chemistry rather than to pbl errors. This implies we need a more robust diagnosis.

Current Plans for the Future

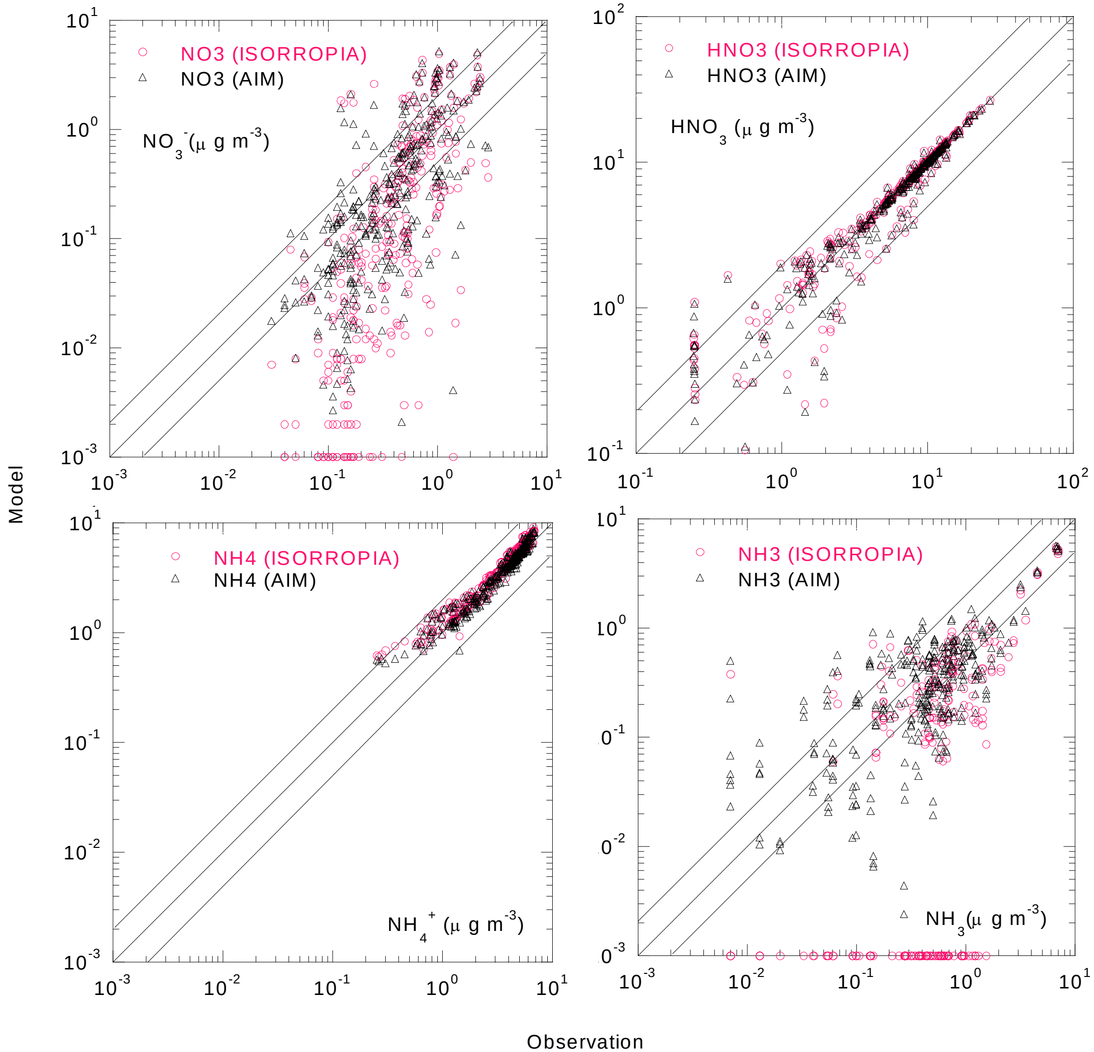
- Photochemistry - current
  - Physics/Mixing sensitivity: write up
  - Chemical Mechanism sensitivity: close the loop and write up
- Photochemistry – near term
  - NO_X accountability initiative (next slide)
  - Work with Gail Tonnesen and UCR chamber testing of indicators
- Particulate Matter – current
  - Difficulty in prediction of aerosol nitrate
  - OC source apportionment work with special CMAQ version
- Particulate Matter – near term
  - Nitrate replacement for sulfate control strategy sensitivity
  - Comparisons against new July 2001 Supersite data
  - Test new section model, CMAQ-AIM, especially for sea salt effects
  - Test sulfate production as a function of grid resolution influence on meteorological simulation and other factors

Photochemistry – NOX Accountability

Looking for truth in the real world. We expect a signal in next years stemming from the NO_X SIP Call, Tier II and HDDV regulations

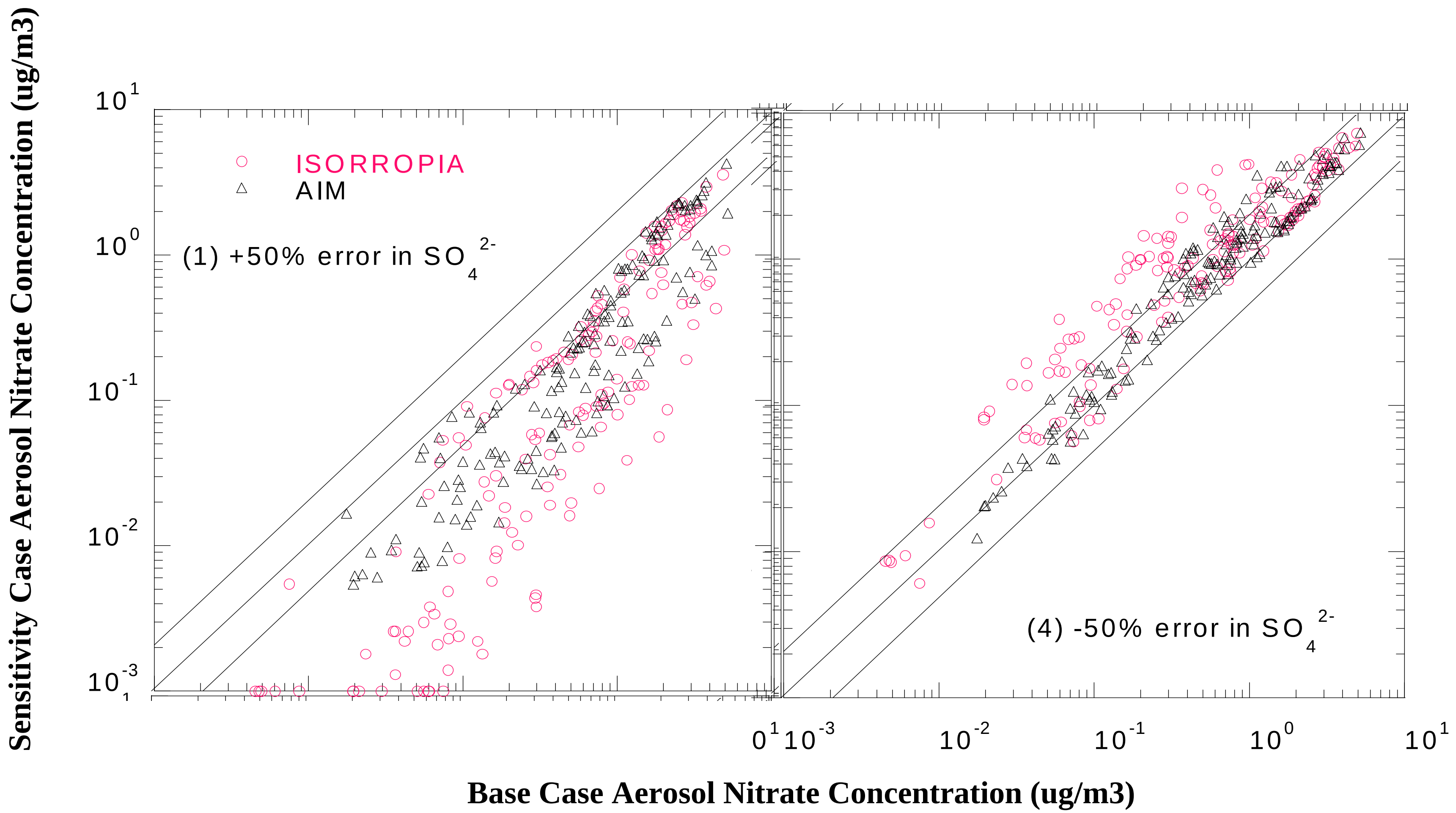
- What is the expected signal and where is it? How does this relate to available monitoring sites? Are there critical spatial gaps?
- Which sector reductions produce the signal? Is the signal spatially different for the different sector contributions? How does this provide guidance to statistical analyses?
- What additional indicator species would be most valuable?
- How much uncertainty can chemical mechanism variation introduce into the expected signal (SAPRC99 vs. CB4 vs. CB403?
- How much uncertainty can differences in base meteorological periods introduce into the expected signal?

Box-Model Examination of Nitrate Formation



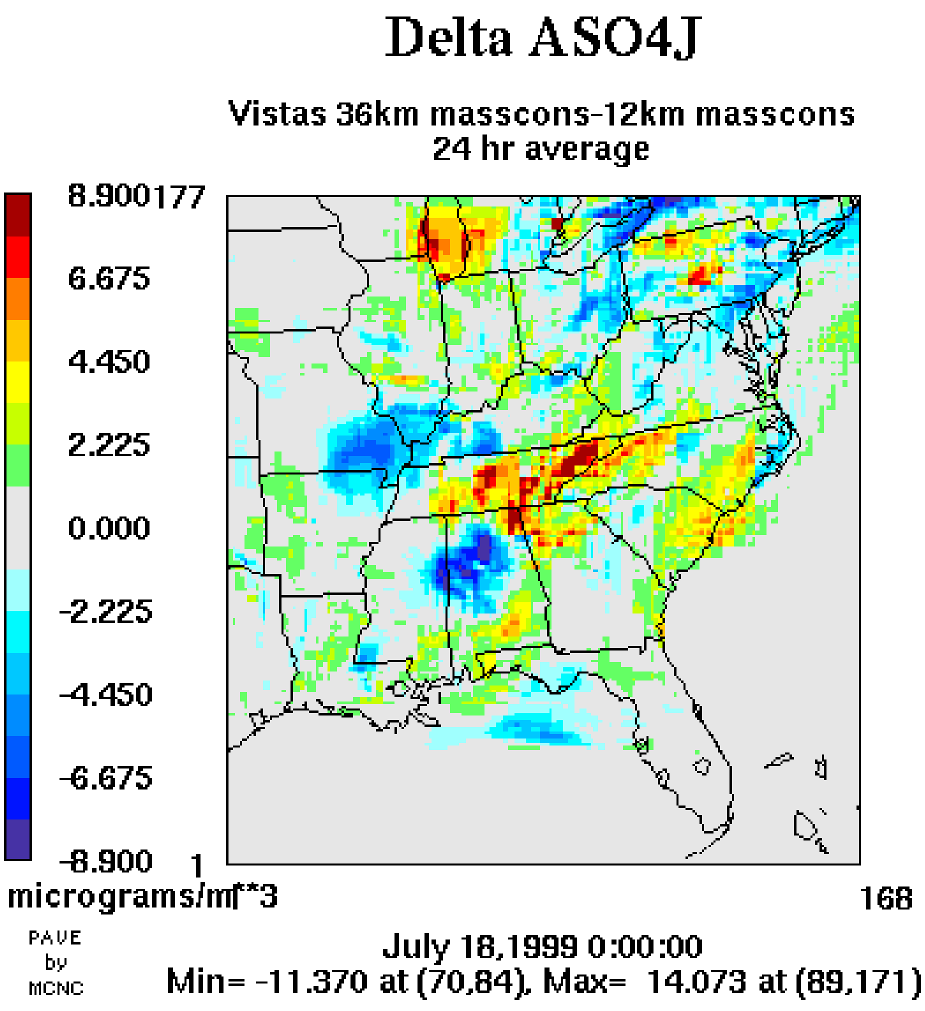
Using actual measurement data (i.e., no input error) to test only the equilibrium partitioning indicates nitrate prediction will have large uncertainty. AIM=gold std. We need to school our expectations.

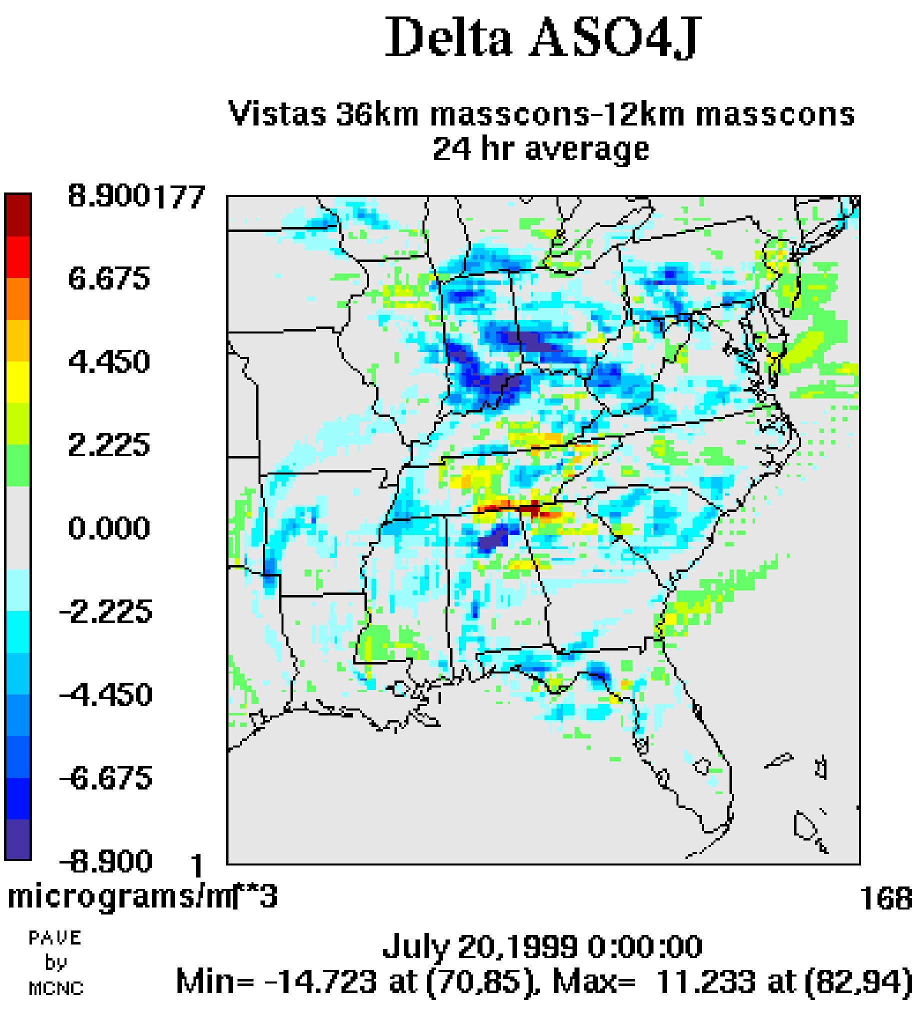
Examination of Nitrate Replacement of Sulfate



There is a major difference in the NO_3^- response to an $SO_4^=$ reduction between the gold standard AIM and ISORROPIA that is used in CMAQ. This is a concern and needs testing.

Examination of Sulfate Production as Function of Grid Size





Observations

- Only so much we can cover at one time. Lots to do. There is a need to publish. Not all work is publishable.
- Diagnostic evaluations depend on close coordination with model developers. We are coming into conflict with demands on them to improve the model as rapidly as possible.
- Post docs are an invaluable help. We have suffered a large discretionary budget reduction and have lost basic support for model evaluation post docs.
- We get “captured” by special data sets. We hope that interpretations generalize, but place or season does not always do so.
- Inverse modeling raises a ticklish issue. Do we run the model with “blessed” inputs or do we run the model with inputs that give the best control strategy response? Some charge inverse modeling is “tuning”.