Evaluation of PM Model Performance

James W. Boylan

Georgia Department of Natural Resources

3rd Particulate Matter/Regional Haze/Ozone Modeling Workshop

New Orleans, LA

May 19, 2005

Outline

- Performance Statistics
- Graphical Displays
- Performance Goals and Criteria
- Diagnostic Tests
- Model Performance and Model Response

Species to Evaluate

- PM Species
  - Sulfate (SO4), Nitrate (NO3), Ammonium (NH4), Organic Carbon (OC), Elemental Carbon (EC), Soils, Coarse Mass (CM), PM2.5, and PM10
- Gaseous and Wet Deposition Species
  - Ozone, HNO3, NO2, NOy, PAN, NH3, SO2, CO, H2O2
- Wet Deposition Species
  - SO4, NO3, NH4
- Indicator Ratios
  - (NH4+NH3)/(TNO3+SO4)
  - NO3/TNO3

Performance Statistics

Statistical Evaluations

- Monitoring Networks
  - IMPROVE: SO4, NO3, NH4, OC, EC, Soils, CM, PM2.5, PM10
  - STN: SO4, NO3, NH4, OC, EC, TCM, Soils, PM2.5, PM10
  - CASTNet: SO4, NO3, NH4, SO2, HNO3, TNO3
  - AQS: Ozone, CO, SO2, NO2, PM2.5, PM10
  - NADP: SO4, NO3, NH4
  - Regional and Local Monitors (e.g., SEARCH)
- Evaluation Metrics
  - Mean and Normalized Bias and Error
  - r^2
- Spatial and Temporal Evaluations
  - Domain-Wide and by Sub-Regions (e.g., RPO, State, NAA, etc.)
  - Episodic or Month-by-Month for Annual Simulations

Performance Metrics	Equation
Mean Bias ($\mu g/m^3$)	$MB = \frac{1}{N}\sum_{i=1}^{N}(C_m - C_o)$
Mean Error ($\mu g/m^3$)	$ME = \frac{1}{N}\sum_{i=1}^{N}\lvert C_m - C_o\rvert$
Mean Normalized Bias (%) (-100% to +∞) Mean Normalized Error (%) (0% to +∞)	$MNB = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{C_m - C_o}{C_o}\right)$ $MNE = \frac{1}{N}\sum_{i=1}^{N}\left\lvert\frac{C_m - C_o}{C_o}\right\rvert$
Normalized Mean Bias (%) (-100% to +∞) Normalized Mean Error (%) (0% to +∞)	$NMB = \frac{\sum_{i=1}^{N}(C_m - C_o)}{\sum_{i=1}^{N}C_o}$ $NME = \frac{\sum_{i=1}^{N}\lvert C_m - C_o\rvert}{\sum_{i=1}^{N}C_o}$
Mean Fractional Bias (%) (-200% to +200%) Mean Fractional Error (%) (0% to +200%)	$MFB = \frac{1}{N}\sum_{i=1}^{N}\frac{(C_m - C_o)}{\left(\frac{C_o + C_m}{2}\right)}$ $MFE = \frac{1}{N}\sum_{i=1}^{N}\frac{\lvert C_m - C_o\rvert}{\left(\frac{C_o + C_m}{2}\right)}$

Standard Performance Metrics

Mean Normalized Bias and Error

- Usually associated with observation-based minimum threshold
  - Some components of PM can be very small making it difficult to set a reasonable minimum threshold value without excluding a majority of the data points
- Without a minimum threshold, very large normalized biases and errors can result when observations are close to zero even though the absolute biases and errors are very small
  - A few data points can dominate the metric
- Overestimations are weighted more than equivalent underestimations
- Assumes observations are absolute truth
- Most biased and least useful of the three metrics

Standard Performance Metrics

Normalized Mean Bias and Error

- Biased towards overestimations
- Assumes observations are absolute truth

Mean Fractional Bias and Error

- Bounds maximum bias and error
- Symmetric: gives equal weight to underestimations and overestimations
- Normalized by average of observation and model
- **Least biased and most useful of the three metrics**

Model ($\mu g/m^3$)	Obs. ($\mu g/m^3$)	MB ($\mu g/m^3$)	NMB (%)	MNB (%)	MFB (%)	ME ($\mu g/m^3$)	NME (%)	MNE (%)	MFE (%)
0.05	1.0	-0.95		-95	-180.95	+0.95		+95	+180.95
1.0	0.05	+0.95		+1900	+180.95	+0.95		+1900	+180.95
1.0	0.01	+0.99		+9900	+196.04	+0.99		+9900	+196.04
0.683	0.353	+0.33	+93.4	+3901.7	+65.3	0.96	272.9	3965.0	186.0

VISTAS Model Performance

Species	Network	US (36km)				Vistas States (36km)				Vistas States (12km)			
		CB4		SAPRC99		CB4		SAPRC99		CB4		SAPRC99	
		FE(%)	FB (%)	FE(%)	FB (%)	FE(%)	FB (%)	FE(%)	FB (%)	FE(%)	FB (%)	FE(%)	FB (%)
OC	IMPROVE	70.4	30.8	61.7	4.2	63.0	-39.5	70.7	-56.3	63.5	-46.9	71.7	-63.9
	SEARCH	43.4	43.4	30.4	30.4	--				--			
	STN	52.5	-18.3	55.5	-24.6	50.4	-42.4	55.3	-50.3	43.3	-24.4	46.9	-31.3
EC	IMPROVE	56.5	4.3	56.7	-4.8	49.6	-12.7	50.4	-15.4	49.4	-26.5	50.9	-29.3
	SEARCH	58.6	-38.7	61.1	-41.7	--				--			
	STN	71.8	59.1	71.9	58.6	55.3	30.7	55.3	30.3	70.4	64.5	70.5	64.5
TCM	SEARCH_H	86.8	-26.5	87.8	-37.3	--				--			
	STN	48.6	-4.6	51.0	-9.9	42.7	-32.4	44.7	-38.9	34.1	-8.2	35.7	-13.9
SOIL	IMPROVE	81.6	34.5	85.4	44.9	104.1	41.5	106.4	46.4	99.7	35.0	101.3	40.0
	SEARCH	148.3	135.9	150.4	138.4	--				--			
	STN	129.6	121.1	131.2	122.9	128.4	107.5	129.8	109.0	136.5	118.3	137.3	119.2
CM	IMPROVE	86.0	-52.3	137.1	-131.5	78.9	-37.8	125.1	-121.6	82.1	-43.3	127.4	-124.6
PM25	IMPROVE	46.0	6.5	46.2	-16.0	46.8	-26.8	52.5	-39.8	44.4	-30.1	55.2	-45.2
	SEARCH	29.2	-2.0	30.7	-11.7	--				--			
	SEARCH_H	56.0	-10.2	57.4	-21.1								
	STN	42.3	19.7	41.5	12.0	31.5	-12.2	33.8	-22.3	34.3	0.0	37.9	-9.9
PM10	IMPROVE	52.2	-19.0	67.5	-54.6	51.4	-33.7	63.8	-56.6	52.0	-38.3	68.6	-62.0

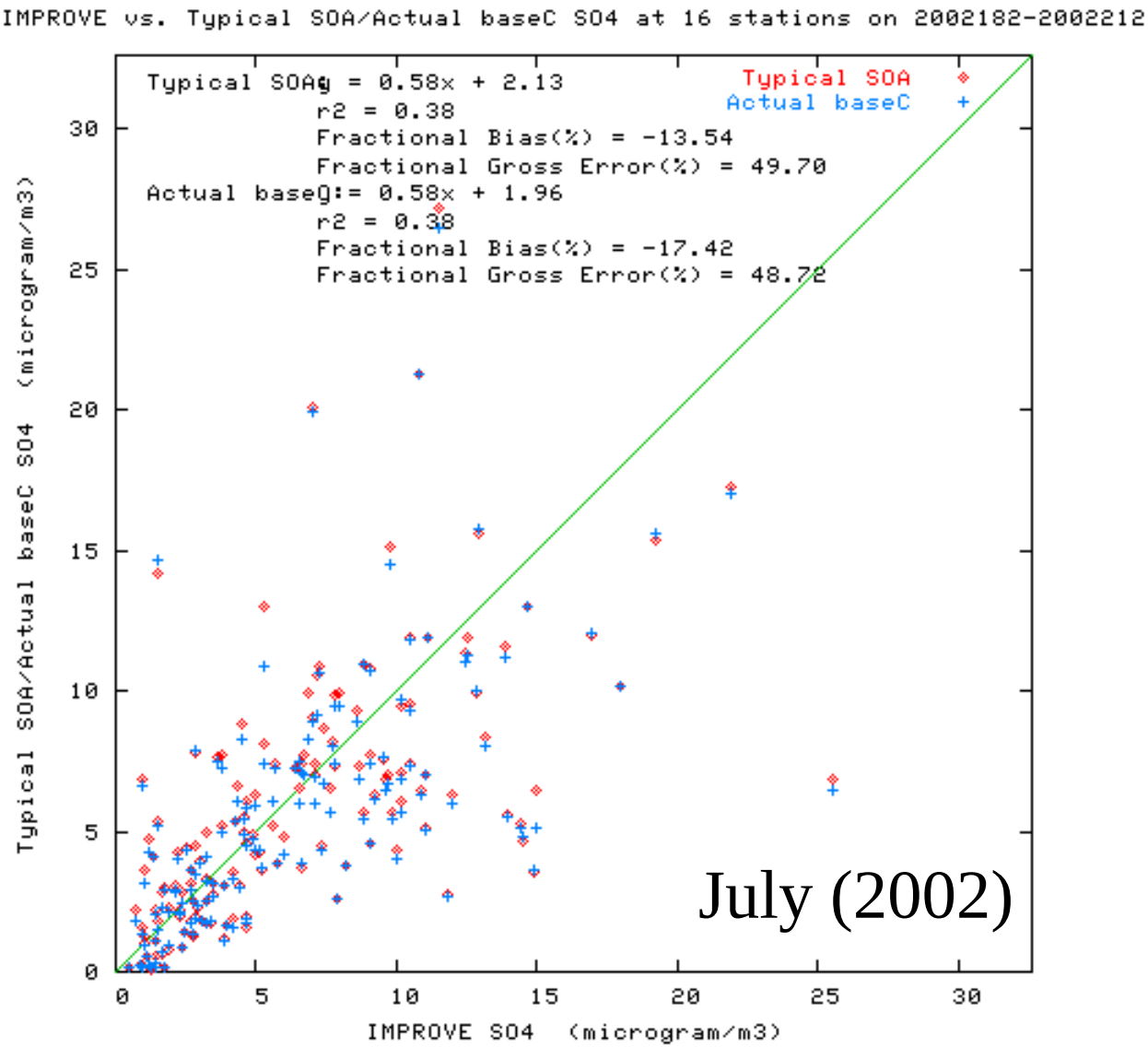
Species	Network	US (36km)				Vistas States (36km)				Vistas States (12km)			
		CB4		SAPRC99		CB4		SAPRC99		CB4		SAPRC99	
		FE(%)	FB (%)	FE(%)	FB (%)	FE(%)	FB (%)	FE(%)	FB (%)	FE(%)	FB (%)	FE(%)	FB (%)
O3	AQS	61.6	-14.7	76.8	-35.2	58.7	7.9	67.8	-7.3	61.2	-4.8	73.4	-22.9
NO2	AQS	75.8	-11.7	73.8	-17.1	74.8	-13.0	73.7	-15.2	76.0	1.3	75.2	-5.3
SO2	AQS	88.6	8.9	88.1	12.6	80.5	-0.3	79.5	5.4	89.1	12.7	88.0	18.4
	CASTNET	54.9	6.2	57.5	15.1	49.1	33.4	59.5	48.6	52.7	27.0	63.7	44.0
CO	AQS	82.7	-55.5	81.7	-52.8	85.8	-72.7	84.0	-70.0	81.0	-64.4	79.4	-62.0
SO4	IMPROVE	42.8	11.2	51.8	-30.1	35.9	2.4	44.7	-21.5	38.5	2.7	50.4	-25.5
	CASTNET	31.4	-12.1	47.1	-34.6	24.6	-23.6	35.8	-35.7	21.6	-12.3	31.2	-27.0
	SEARCH	36.9	21.7	34.8	5.1	--				--			
	SEARCH_H	45.4	7.2	44.2	-9.7								
	STN	35.0	15.0	37.0	-6.1	30.5	14.1	30.4	-3.9	33.8	17.0	36.9	-3.3
	NADP	81.5	23.9	79.8	8.8	75.1	35.7	75.0	27.0	85.3	21.5	85.9	13.4
NO3	IMPROVE	142.7	-125.7	145.9	-132.4	127.8	-100.0	127.5	-102.6	143.0	-123.6	144.2	-125.8
	CASTNET	119.7	-38.5	115.8	-38.3	82.2	11.5	83.1	4.6	77.9	-27.0	78.7	-33.7
	SEARCH	126.6	-97.5	129.8	-105.7	--				--			
	SEARCH_H	144.6	-79.6	142.6	-84.2								
	STN	97.9	-56.0	96.6	-59.1	106.6	-77.5	103.6	-86.2	100.3	-79.5	101.5	-82.0
	NADP	96.5	-61.7	97.0	-63.3	77.8	-41.2	80.6	-48.2	94.5	-57.2	95.7	-62.9
G_HNO3	CASTNET	45.4	-25.1	47.4	-32.2	35.5	-19.1	43.2	-29.7	34.7	-19.8	42.7	-29.9
Total_NO3	CASTNET	52.8	-31.7	54.8	-37.6	37.0	-20.9	42.3	-30.2	38.3	-22.5	44.2	-31.0
NH4	IMPROVE	46.6	-19.1	59.8	-49.8	60.4	-51.4	65.4	-58.9	70.1	-64.1	76.7	-72.1
	CASTNET	37.8	-20.7	48.6	-35.4	39.5	-37.6	42.2	-41.5	36.6	-36.2	40.4	-40.4
	SEARCH	36.4	-17.8	40.8	-25.0	--				--			
	SEARCH_H	68.8	25.6	66.7	15.5								
	STN	51.0	33.8	44.2	20.2	49.4	43.7	42.7	34.3	58.3	52.0	47.7	39.6
	NADP	79.3	-18.3	80.1	-22.9	70.4	17.2	70.6	14.5	82.5	4.7	82.4	2.2

Graphical Displays

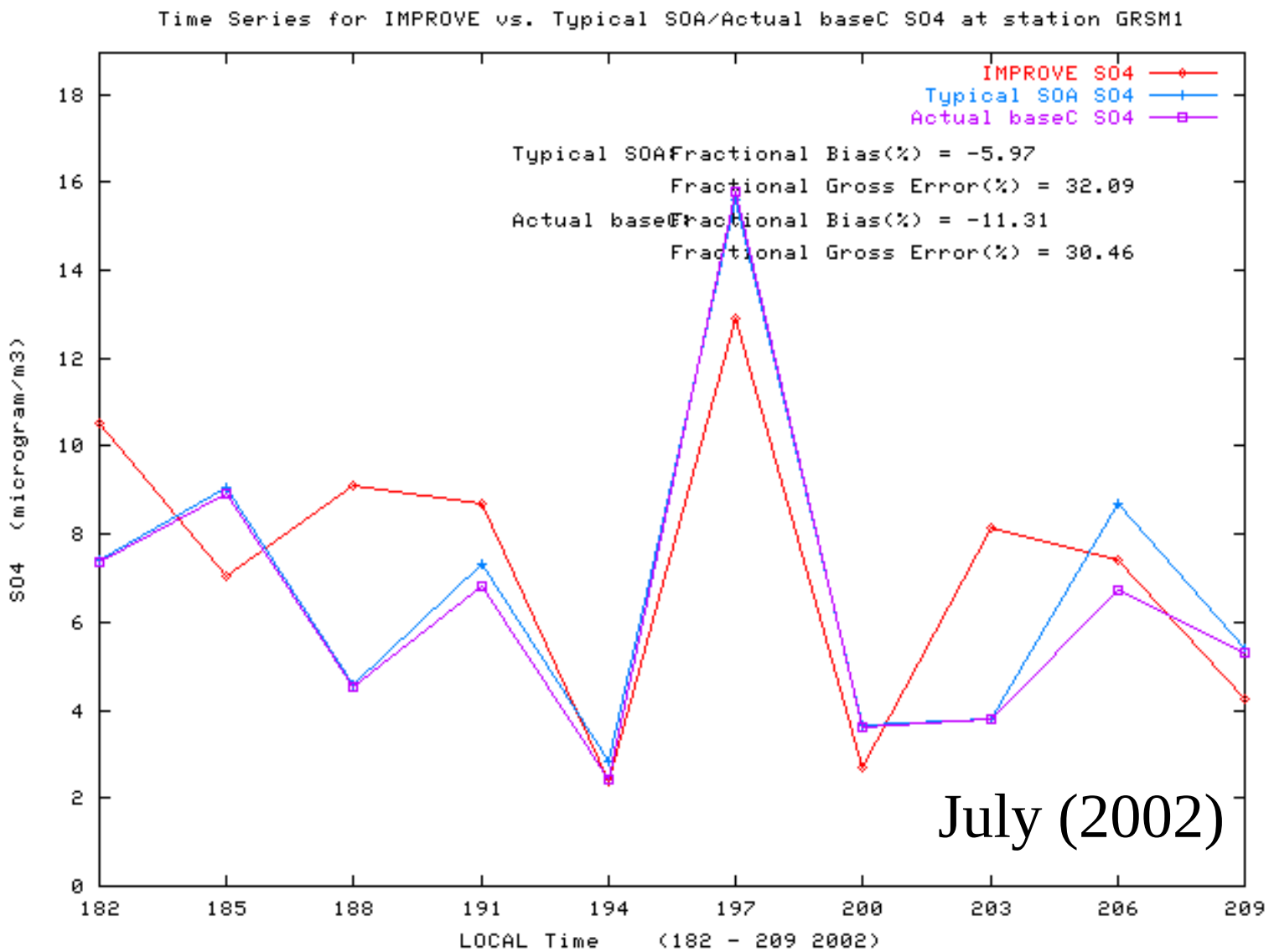
Standard Graphical Displays

- Scatter Plots
  - All Sites_All Days
  - All Sites_One Day
  - One Site_All Days
- Time Series Plots
  - Individual Sites and All Sites
  - Daily and Hourly (if available)
- Stacked Bar Charts and Stacked Time Series
- Q-Q Plots
- Daily Tile Plots
- Animations

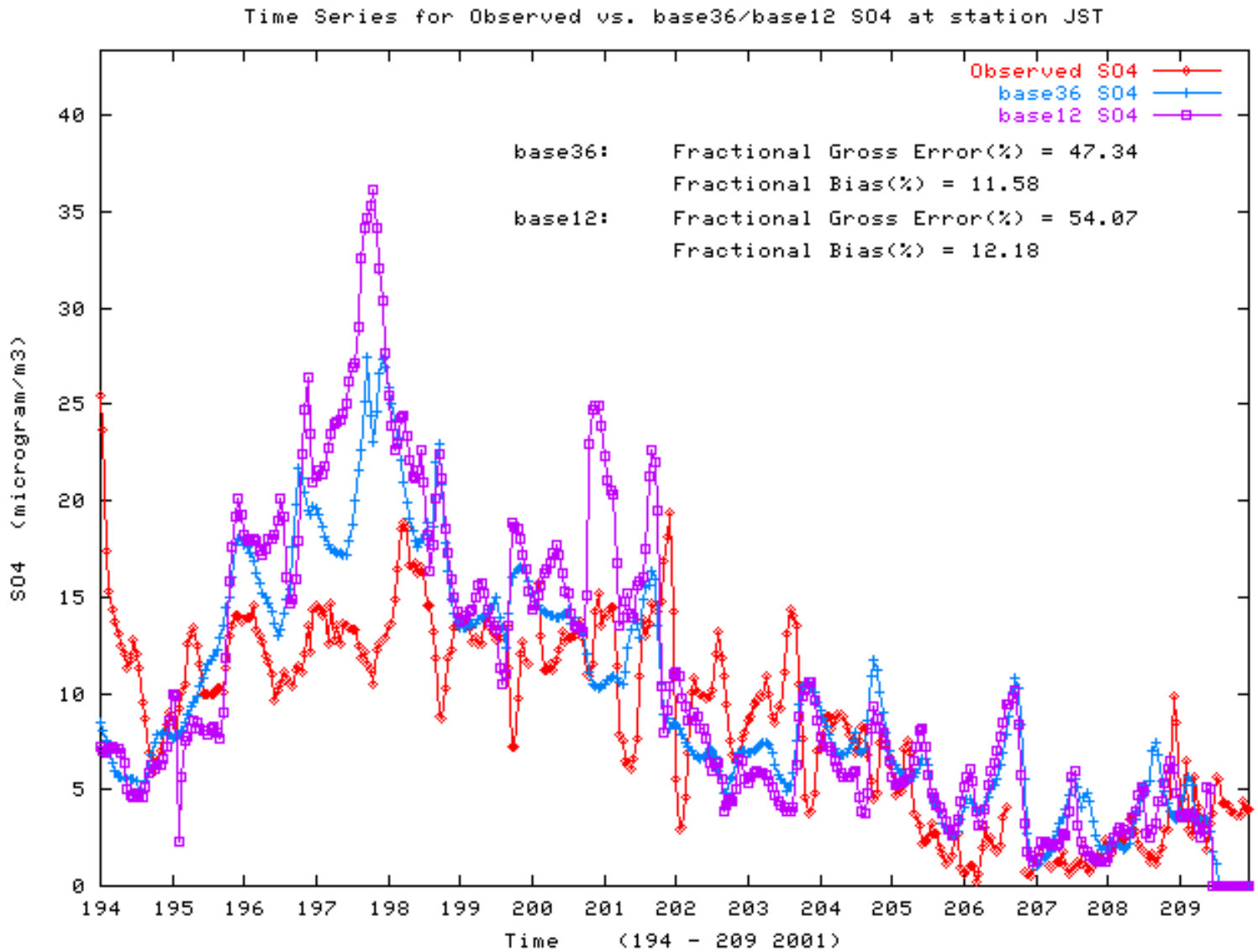
IMPROVE SO4 (All Sites_All Days)



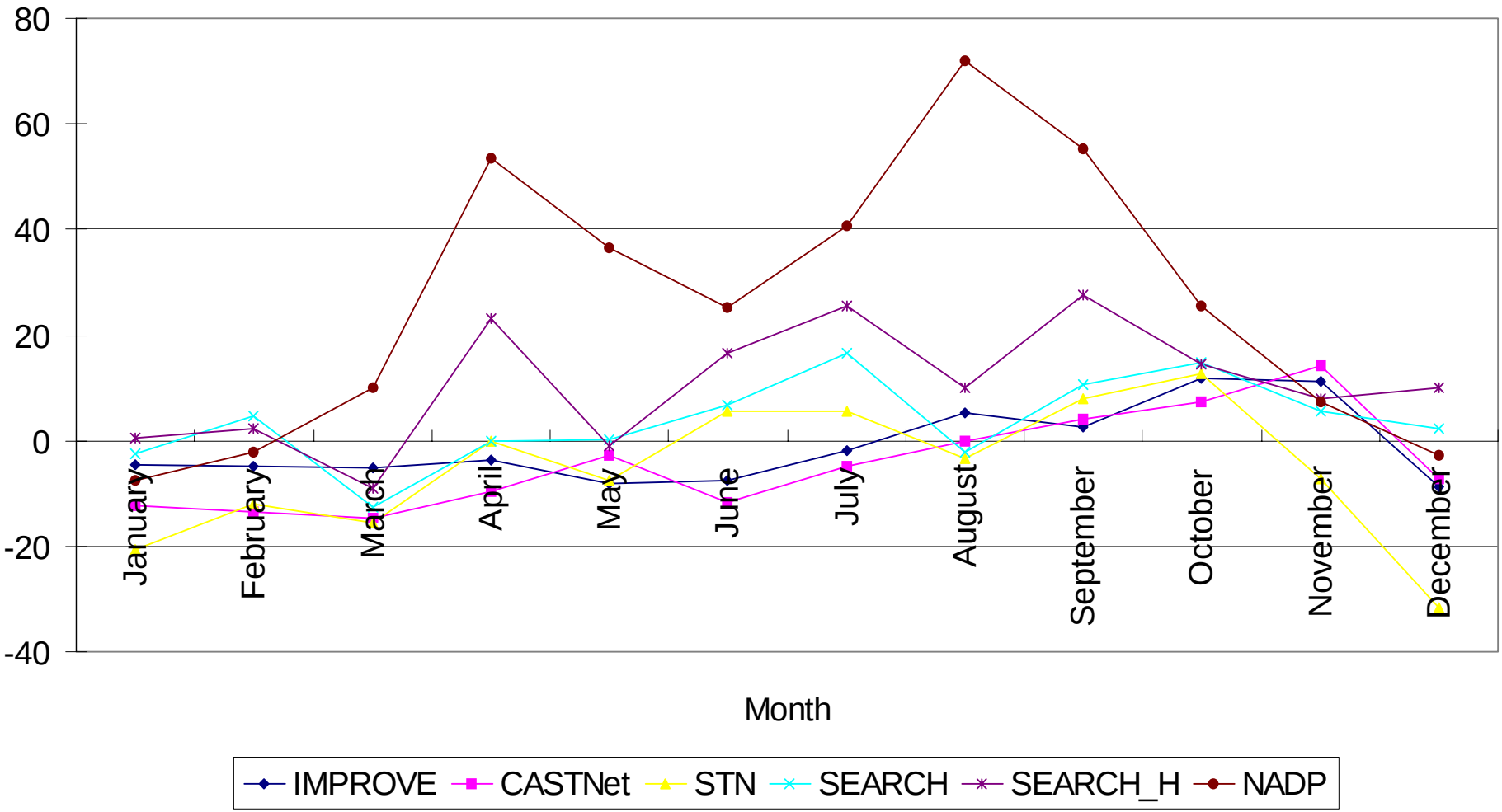
Daily IMPROVE Sulfate (GRSM)



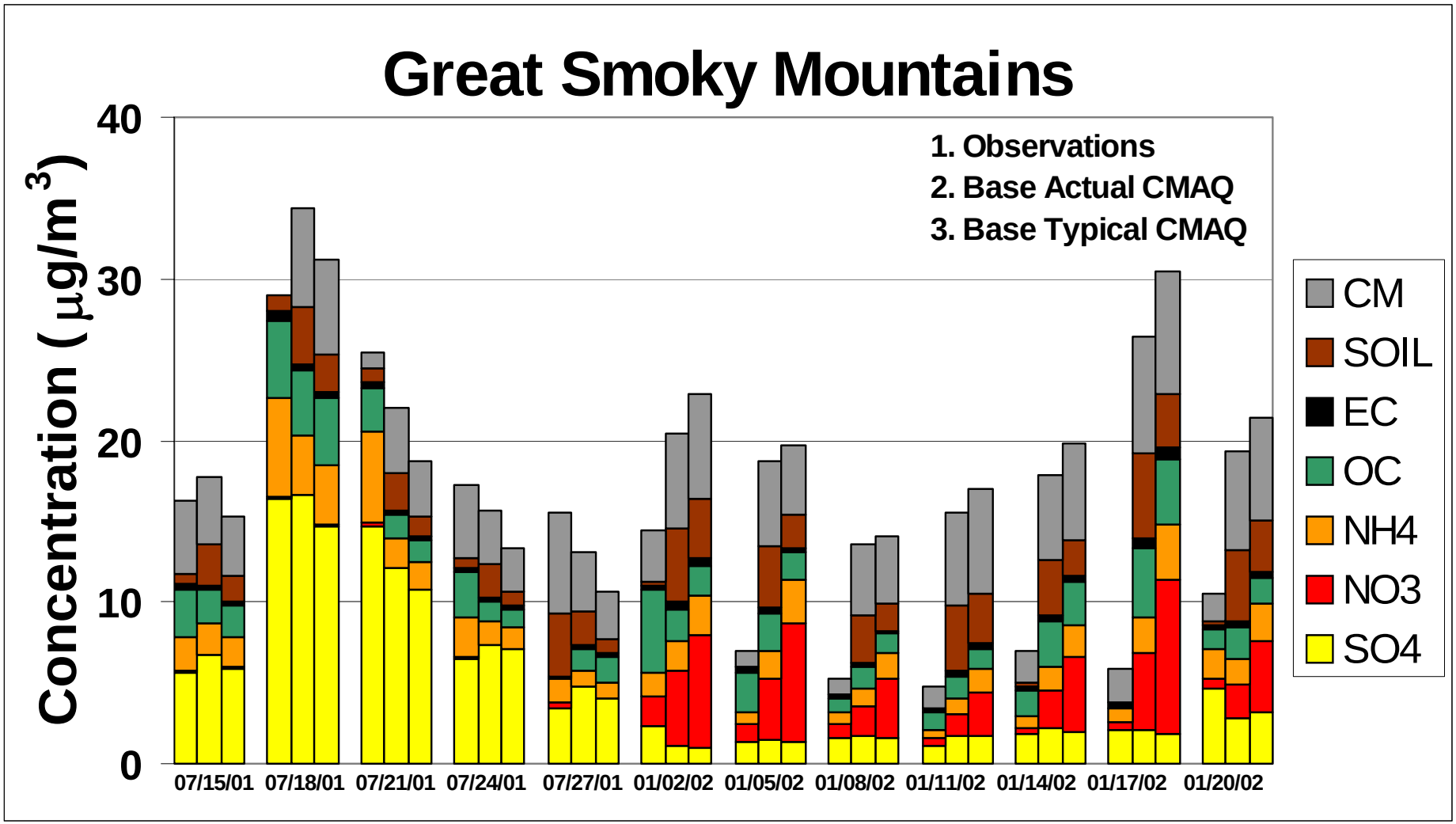
Hourly SEARCH Sulfate (JST)



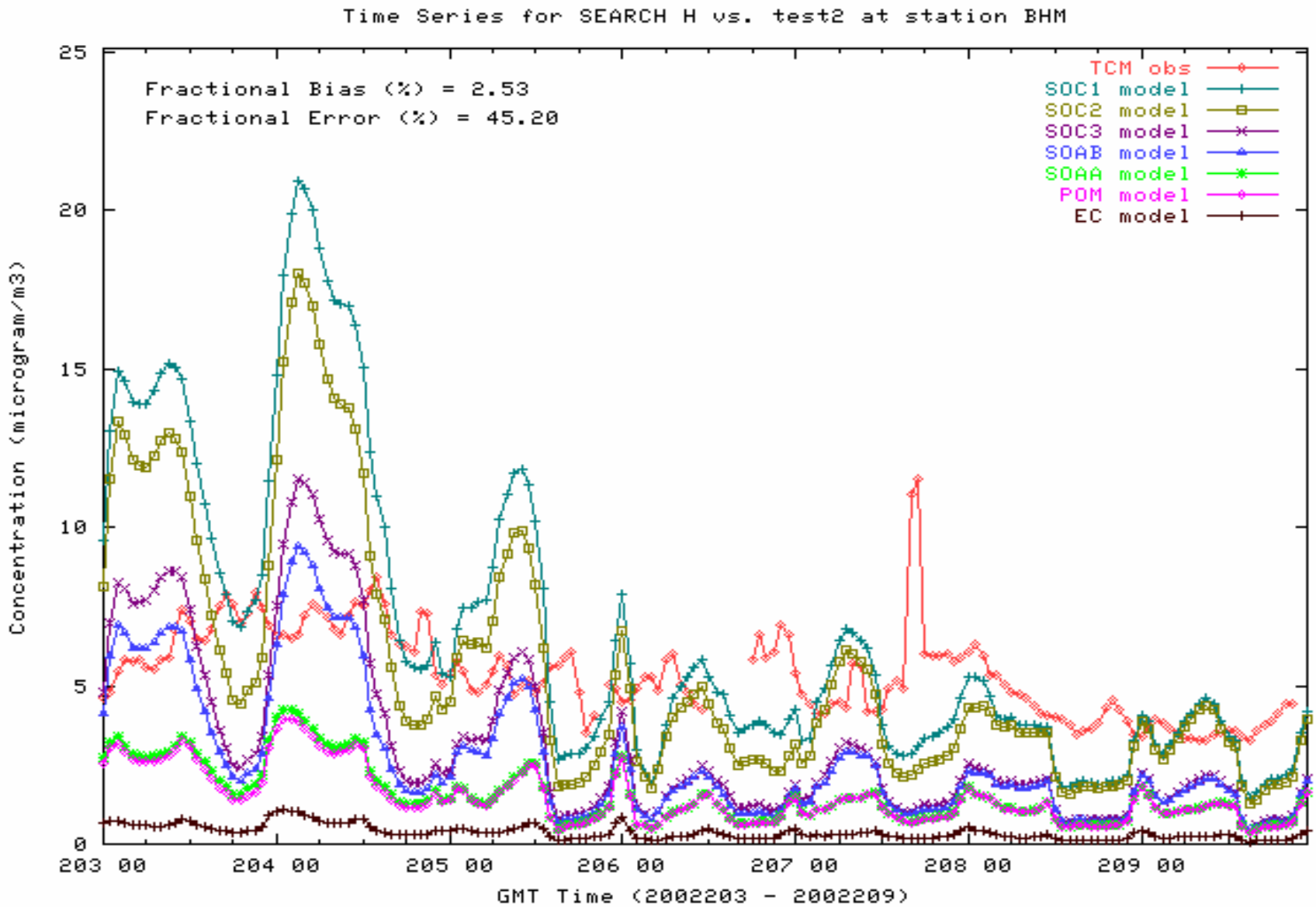
Sulfate (All Sites_All Networks)



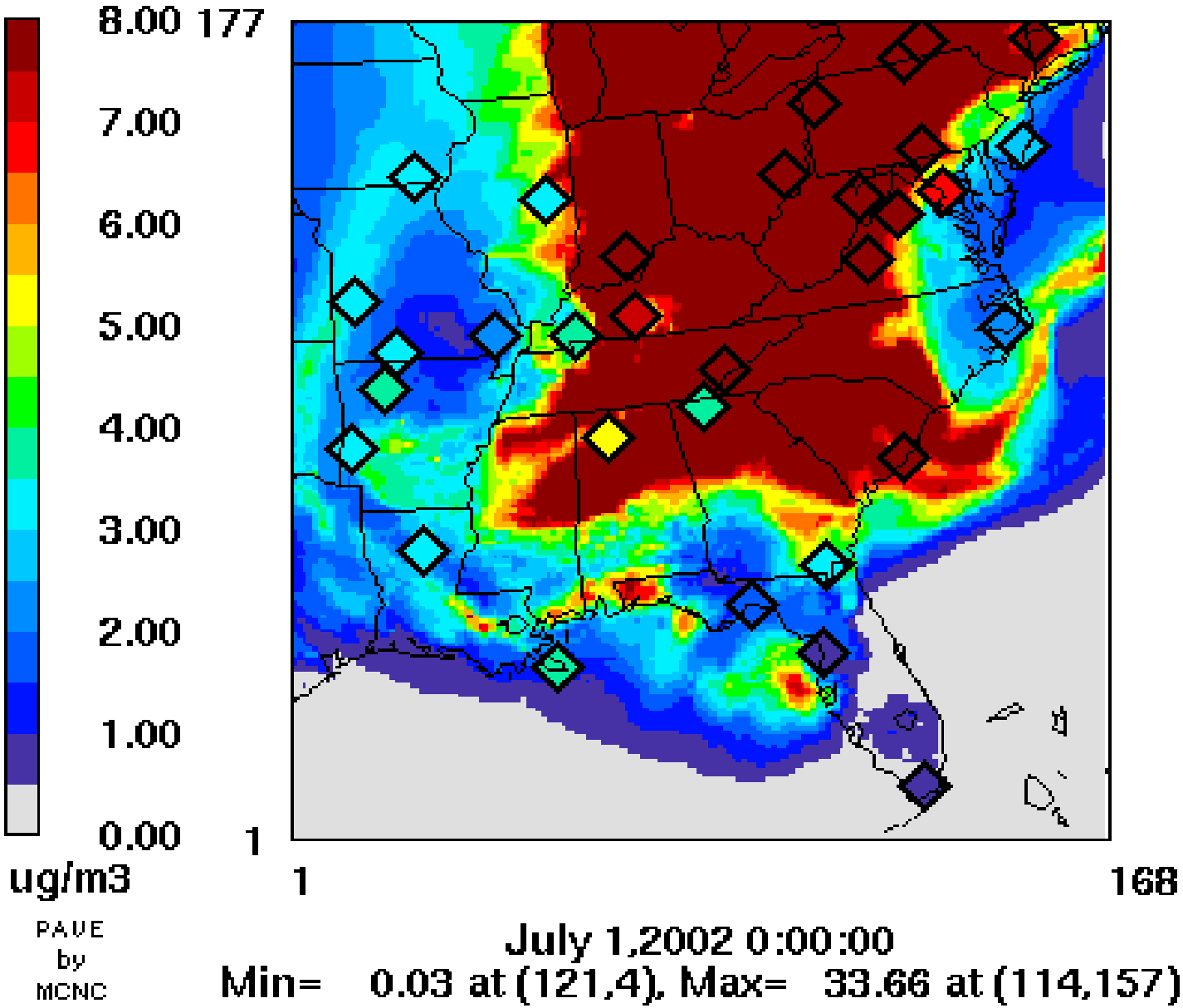
PM Stacked Bar Chart



TCM Stacked Time Series Plot



Daily Sulfate Tile Plot (Animation)

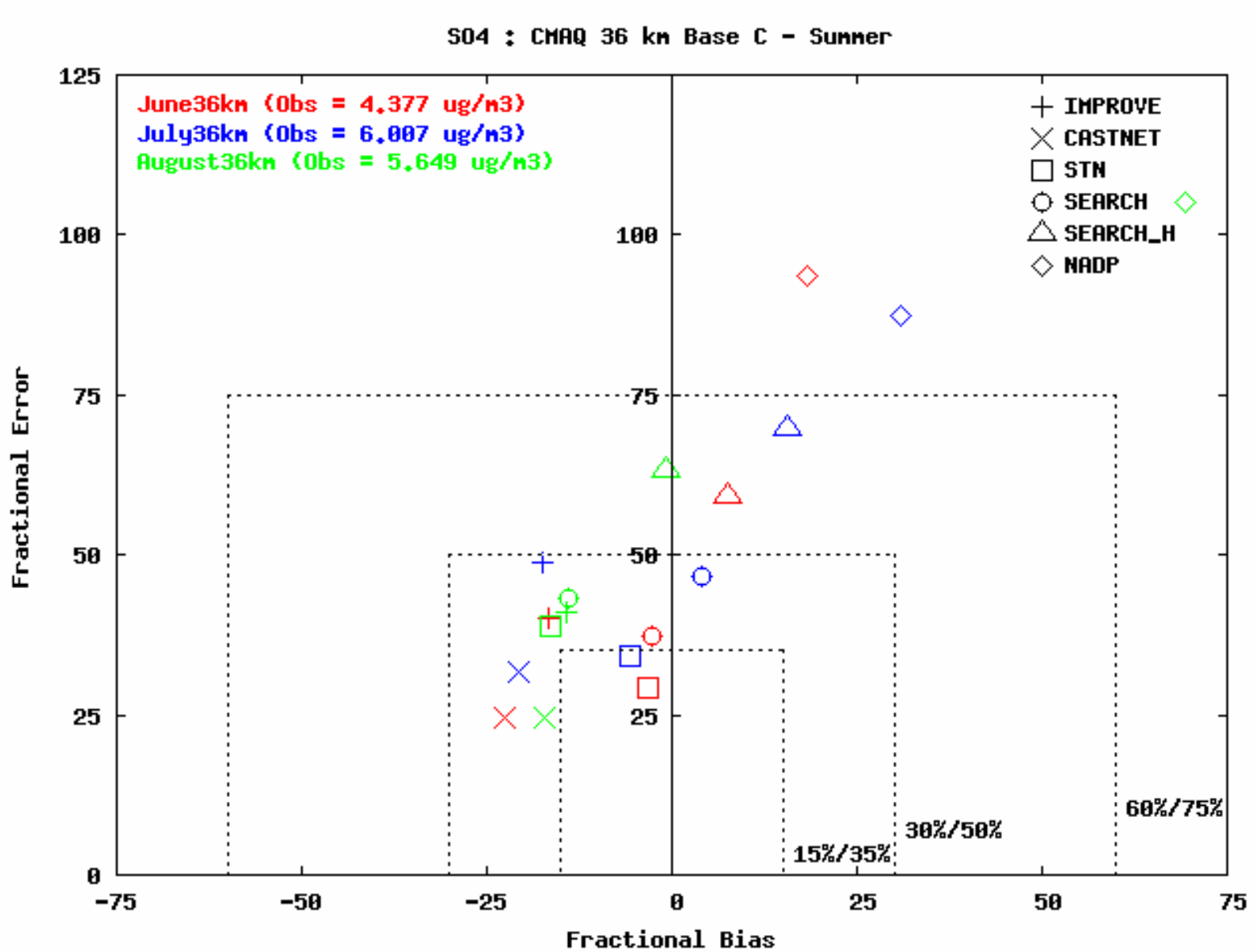


Performance Goals and Criteria

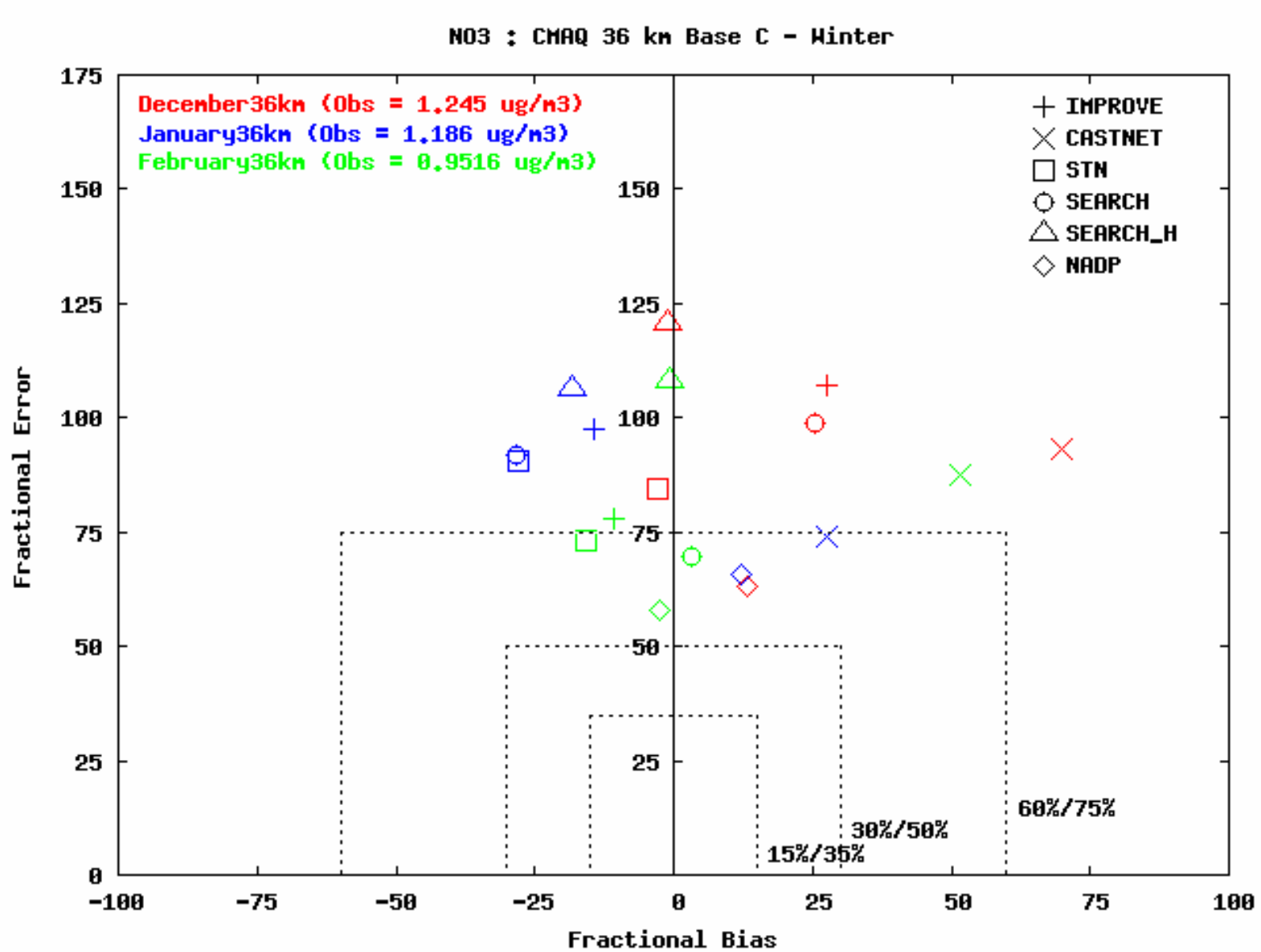
Performance Goals and Criteria

- EPA has issued minimal guidance on PM and visibility model performance evaluation metrics, goals, and criteria.
- **Performance Goals:** Level of accuracy that is considered to be close to the best a model can be expected to achieve
- **Performance Criteria:** Level of accuracy that is considered to be acceptable for regulatory applications
- Soccer Goal Plots
  - Developed by ENVIRON
- Bugle Plots
  - Developed by GA DNR/Georgia Tech

Sulfate Bias and Error



Nitrate Bias and Error



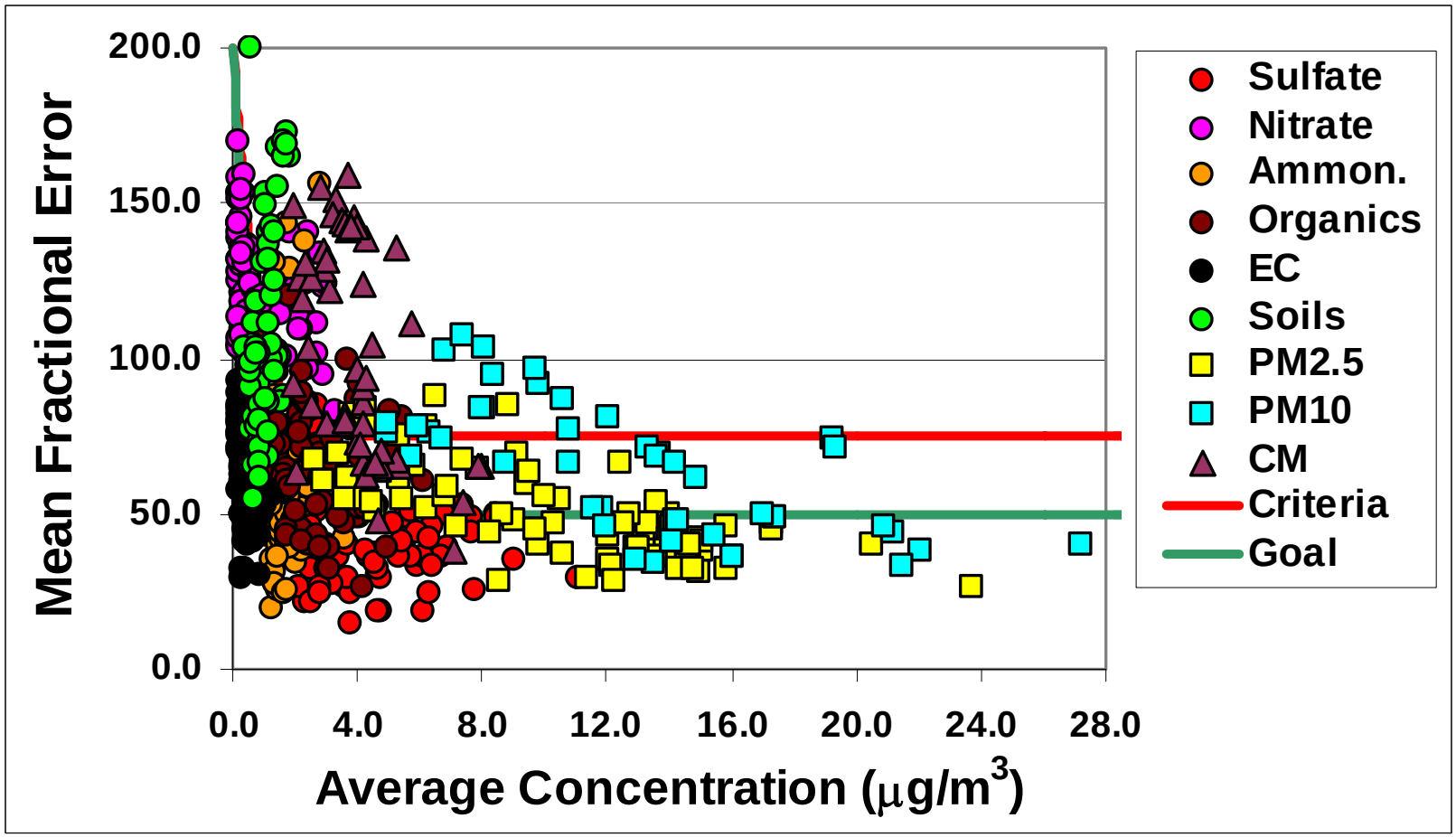
Alternative Approach….

- It has been suggested that different performance goals and criteria are needed for:
  - Different Components of PM2.5 and light extinction
  - Different Seasons of the Year
  - Different Parts of the Country
  - Urban vs. Rural Sites
  - 20% Haziest and 20% Cleanest Days
- **Answer: performance goals and criteria that vary as a function of concentration and light extinction.**

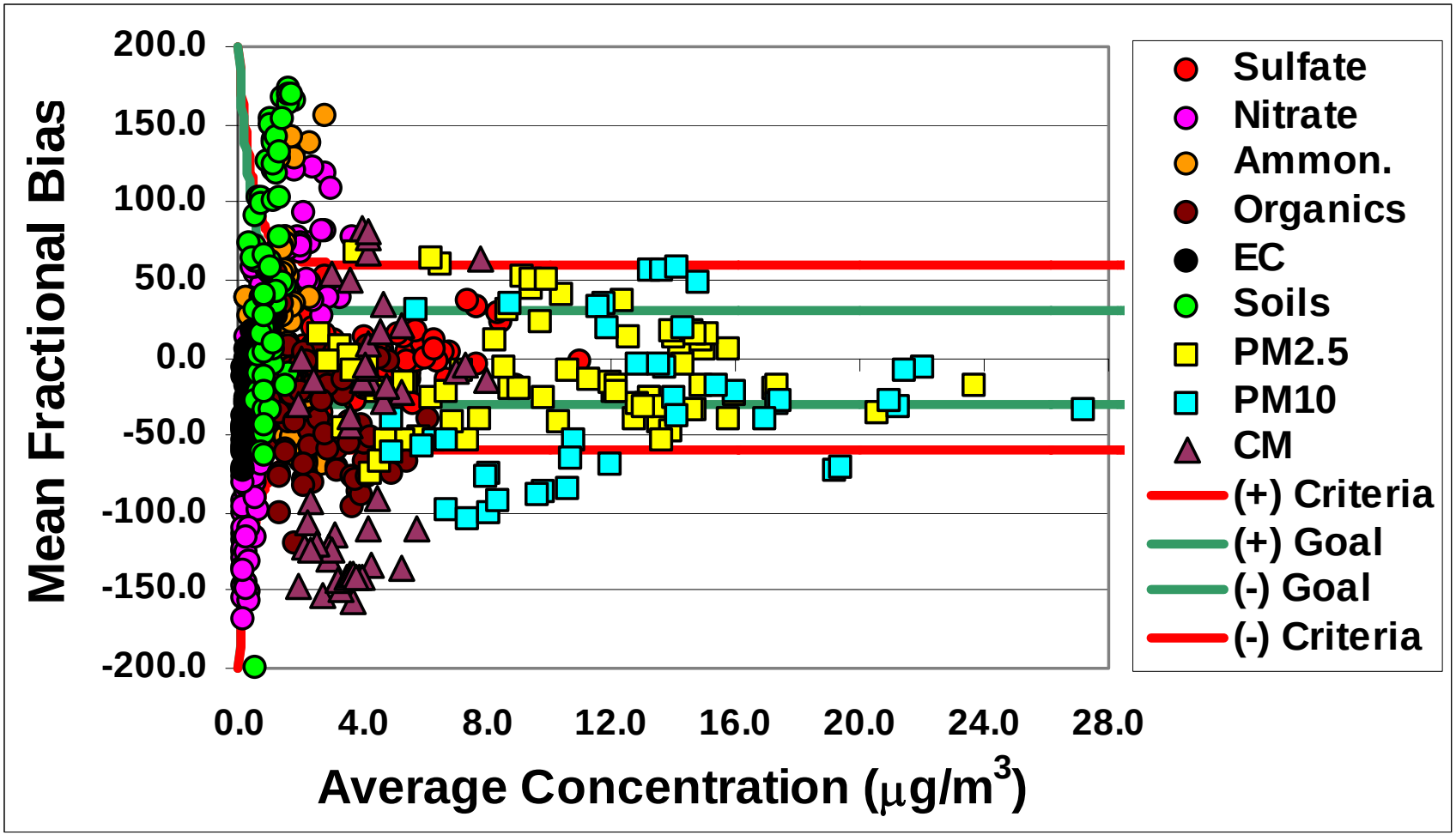
Proposed Goals and Criteria

- Based on Mean Fractional Error and Mean Fractional Bias calculations
- Based on PM modeling benchmarks
  - **SAMI (GT) – 6 Episodes – URM – IMPROVE**
  - **WRAP 309 (UCR/CEP/Environ) – 12 months – CMAQ – IMPROVE**
  - **WRAP 308 (UCR/CEP/Environ) – 2 Seasons – CMAQ – IMPROVE**
  - **Clear Skies (EPA) – 4 Seasons - REMSAD – IMPROVE**
  - **MANE-VU (GT) – 2 Episodes – CMAQ - IMPROVE/SEARCH**
  - **MRPO (LADCO) – 2 Episodes - CMAQ/CAMx/REMSAD – IMPROVE**
  - **EPRI (AER/TVA/Environ) – 1 Episode – CMAQ/MADRID/CAMx**
  - **VISTAS (UCR/AG/Environ) – 3 Episodes - CMAQ – IMPROVE**
  - **VISTAS (UCR/AG/Environ) – 12 months - CMAQ – 4 Networks**

Mean Fractional Error



Mean Fractional Bias



Proposed Goals and Criteria

- Continuous functions with the features of:
  - Asymptotically approaching proposed goals and criteria when the mean of the observed and modeled concentrations are greater than 2.25 $\mu g/m^3$ for PM2.5 evaluations and greater than 10.0 Mm^{-1} for light extinction evaluations
  - Approaching +200% MFE and ±200% MFB when the mean of the observed and modeled concentrations are extremely small
- Vary as a function of species concentrations
  - Less abundant components that contribute less to the total PM2.5 and total light extinction should have less stringent performance goals and criteria
  - "Bugle" Plots

Proposed PM Goals and Criteria

- **Goal: MFE ≤ +50% and MFB ≤ ±30%**

$$MFE \le 150e^{\frac{-0.5(\overline{C_o}+\overline{C_m})}{0.75\,\mu g/m^3}} + 50 \qquad MFB \le \pm 170e^{\frac{-0.5(\overline{C_o}+\overline{C_m})}{0.5\,\mu g/m^3}} + 30$$

- **Criteria: MFE ≤ +75% and MFB ≤ ±60%**

$$MFE \le 125e^{\frac{-0.5(\overline{C_o}+\overline{C_m})}{0.75\,\mu g/m^3}} + 75 \qquad MFB \le \pm 140e^{\frac{-0.5(\overline{C_o}+\overline{C_m})}{0.5\,\mu g/m^3}} + 60$$

Proposed B_{ext} Goals and Criteria

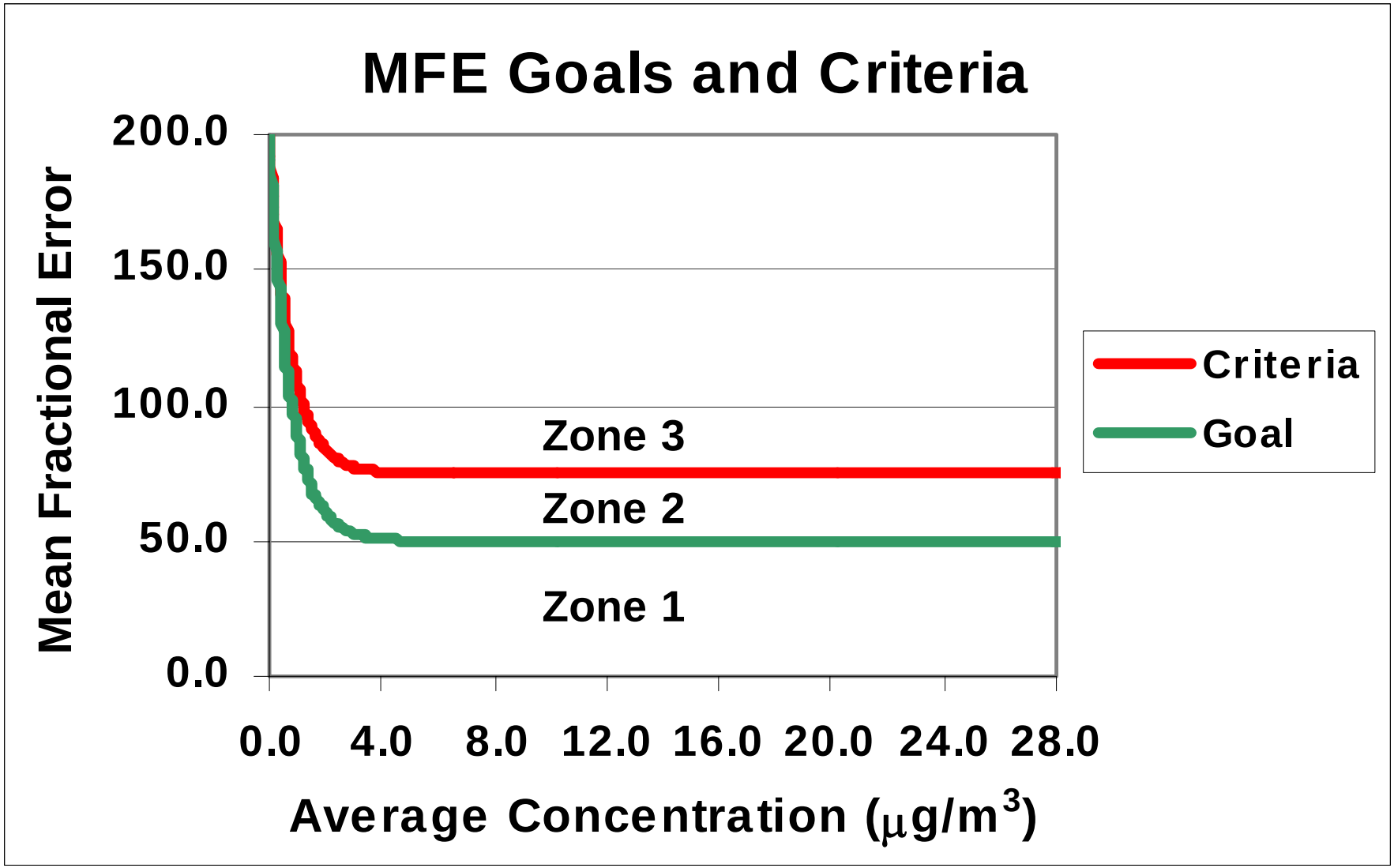
- **Goal: MFE ≤ +50% and MFB ≤ ±30%**

$$MFE \le 150e^{\frac{-0.5(\overline{C_o}+\overline{C_m})}{3.75Mm^{-1}}} + 50 \qquad MFB \le \pm 170e^{\frac{-0.5(\overline{C_o}+\overline{C_m})}{2.5Mm^{-1}}} + 30$$

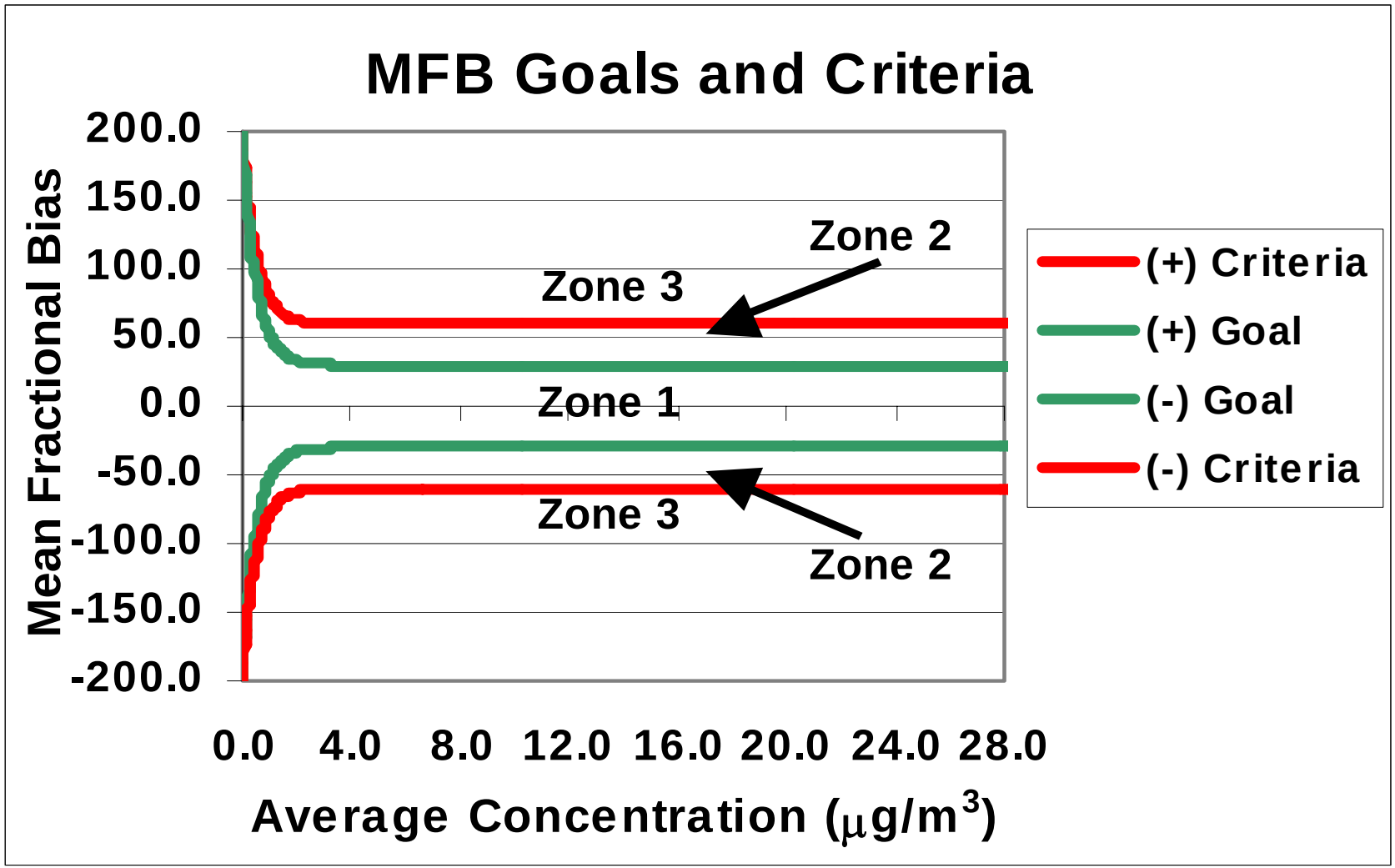
- **Criteria: MFE ≤ +75% and MFB ≤ ±60%**

$$MFE \le 125e^{\frac{-0.5(\overline{C_o}+\overline{C_m})}{3.75Mm^{-1}}} + 75 \qquad MFB \le \pm 140e^{\frac{-0.5(\overline{C_o}+\overline{C_m})}{2.5Mm^{-1}}} + 60$$

Performance Zones (MFE)



Performance Zones (MFB)



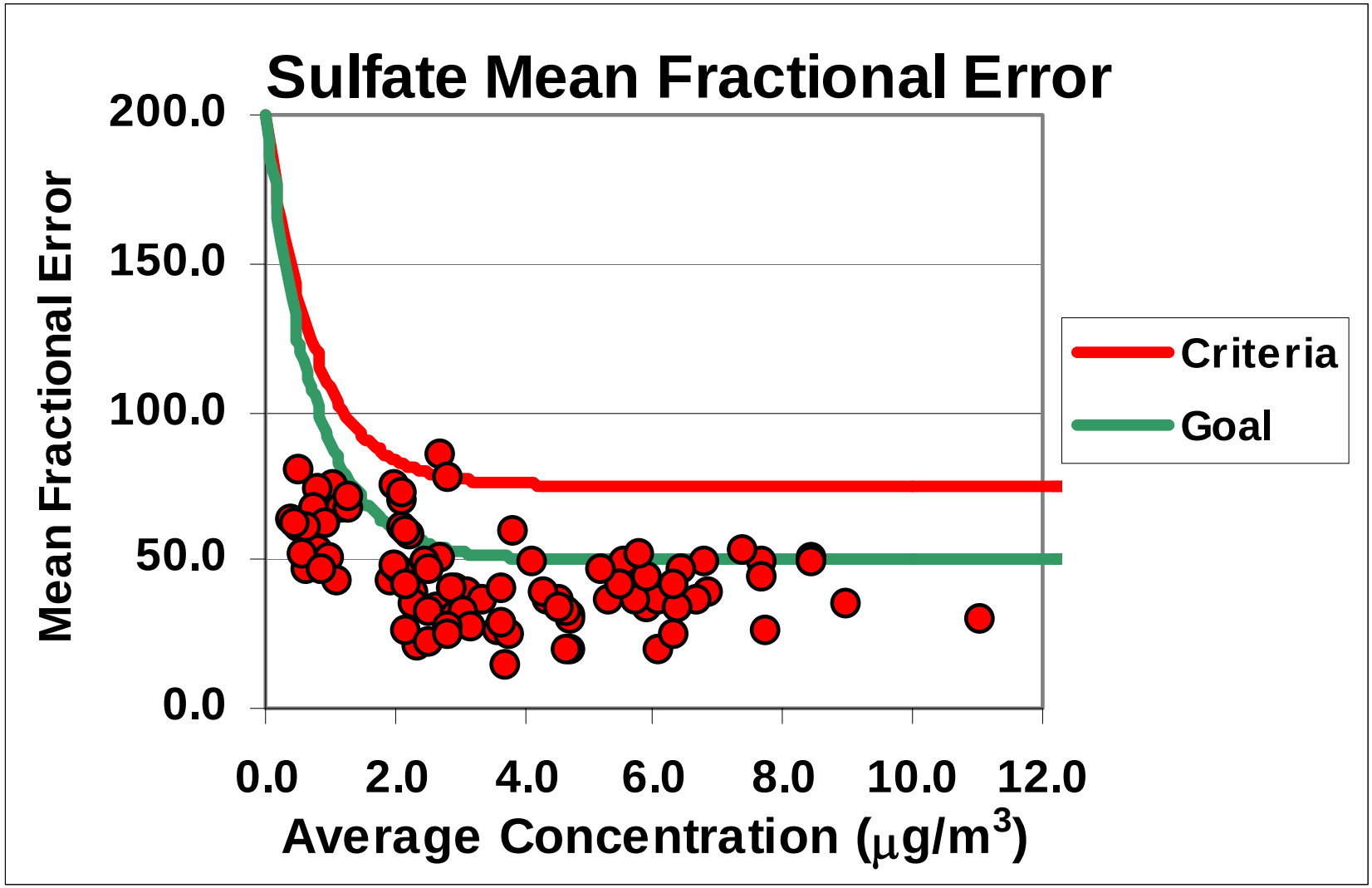
Model Performance Zones

- Zone I
  - Good Model Performance
  - Level I Diagnostic Evaluation (Minimal)
- Zone II
  - Average Model Performance
  - Level II Diagnostic Evaluation (Standard)
- Zone III
  - Poor Model Performance
  - Level III Diagnostic Evaluation (Extended) and Sensitivity Testing

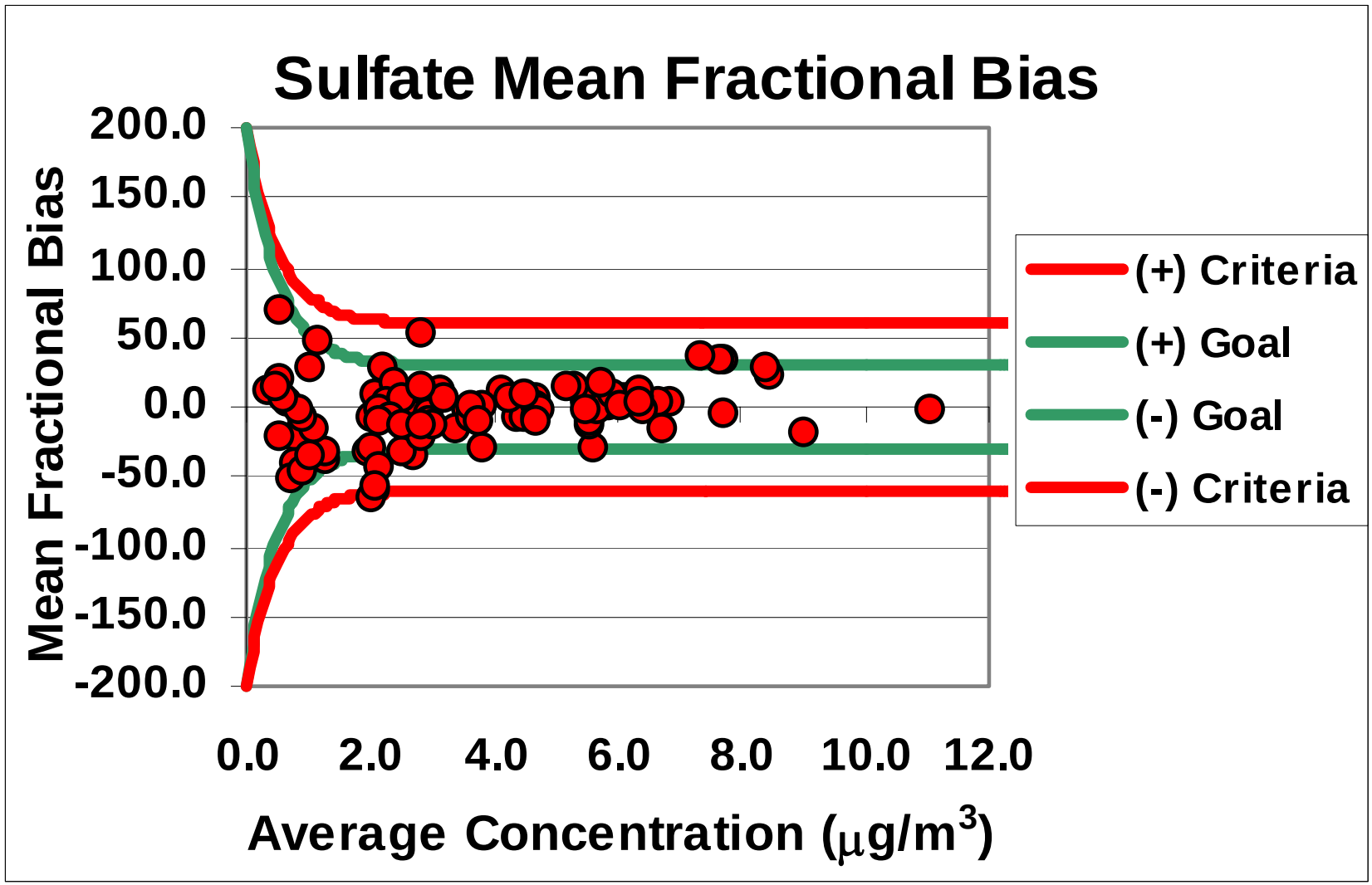
Flexibility in Goals and Criteria

- Failure to meet the proposed criteria should not necessarily mean that the modeling can not be used for regulatory applications.
  - Consideration should be given to the extent that modeling results fall outside the criteria.
  - Also, if it can be demonstrated that the modeling results can still be used with confidence for certain components of PM even though other components of PM are modeled poorly, there is no reason to deem the modeling as a whole as unacceptable.
- As PM models mature and performance improves, performance zones can be made more restrictive by simply adjusting the coefficients in the performance goals and criteria equations.

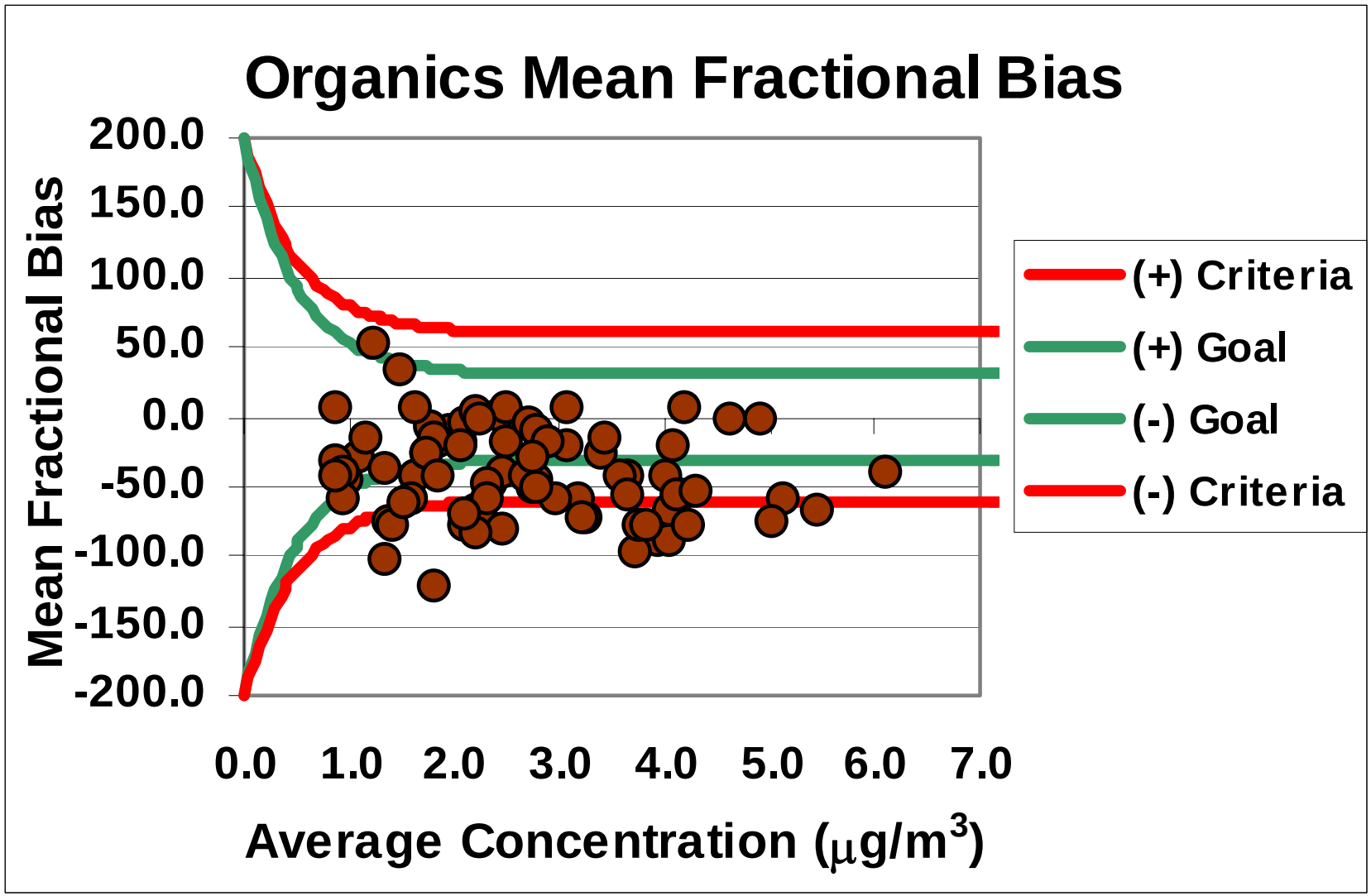
Sulfate (PM) Error



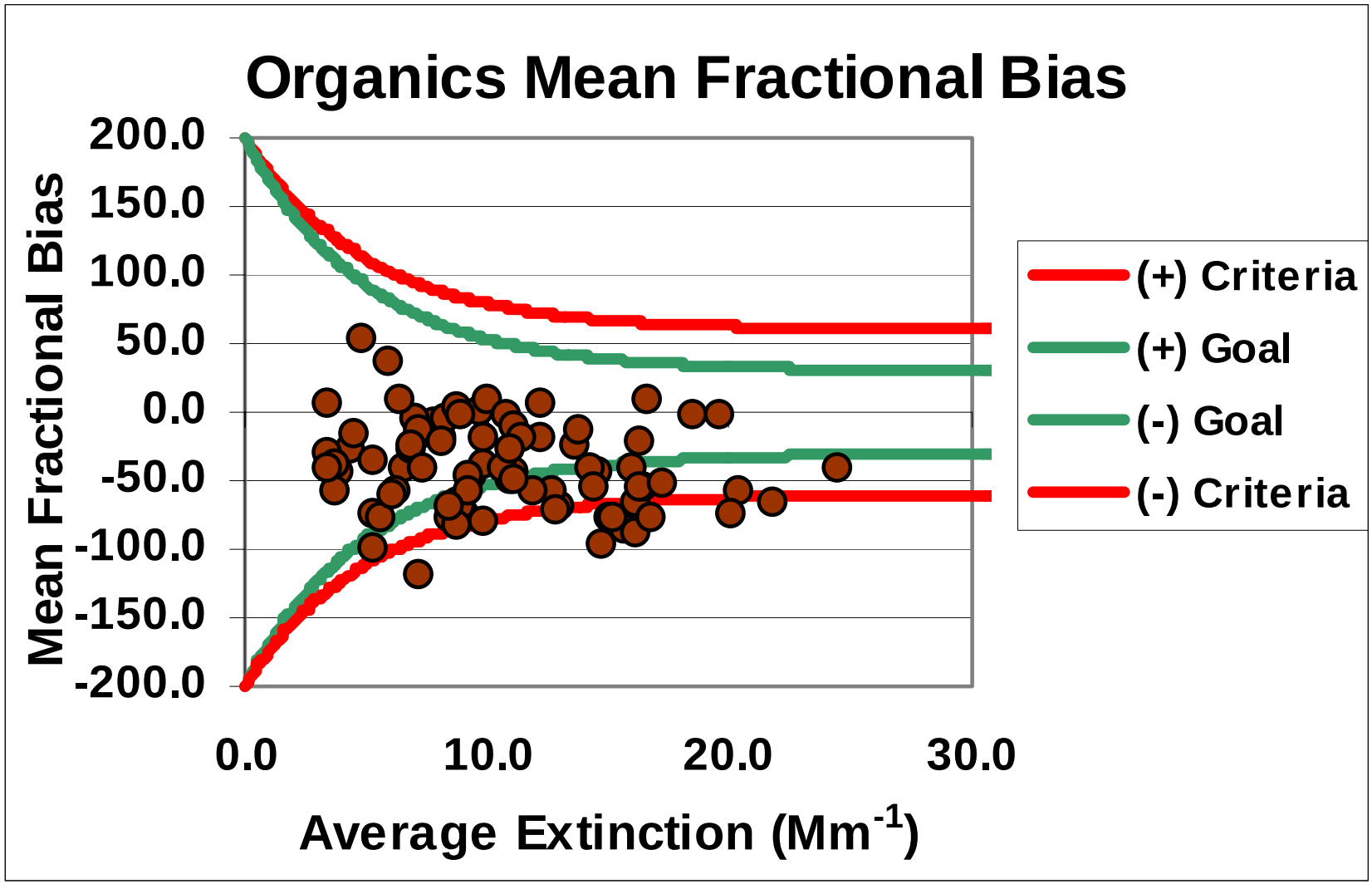
Sulfate (PM) Bias



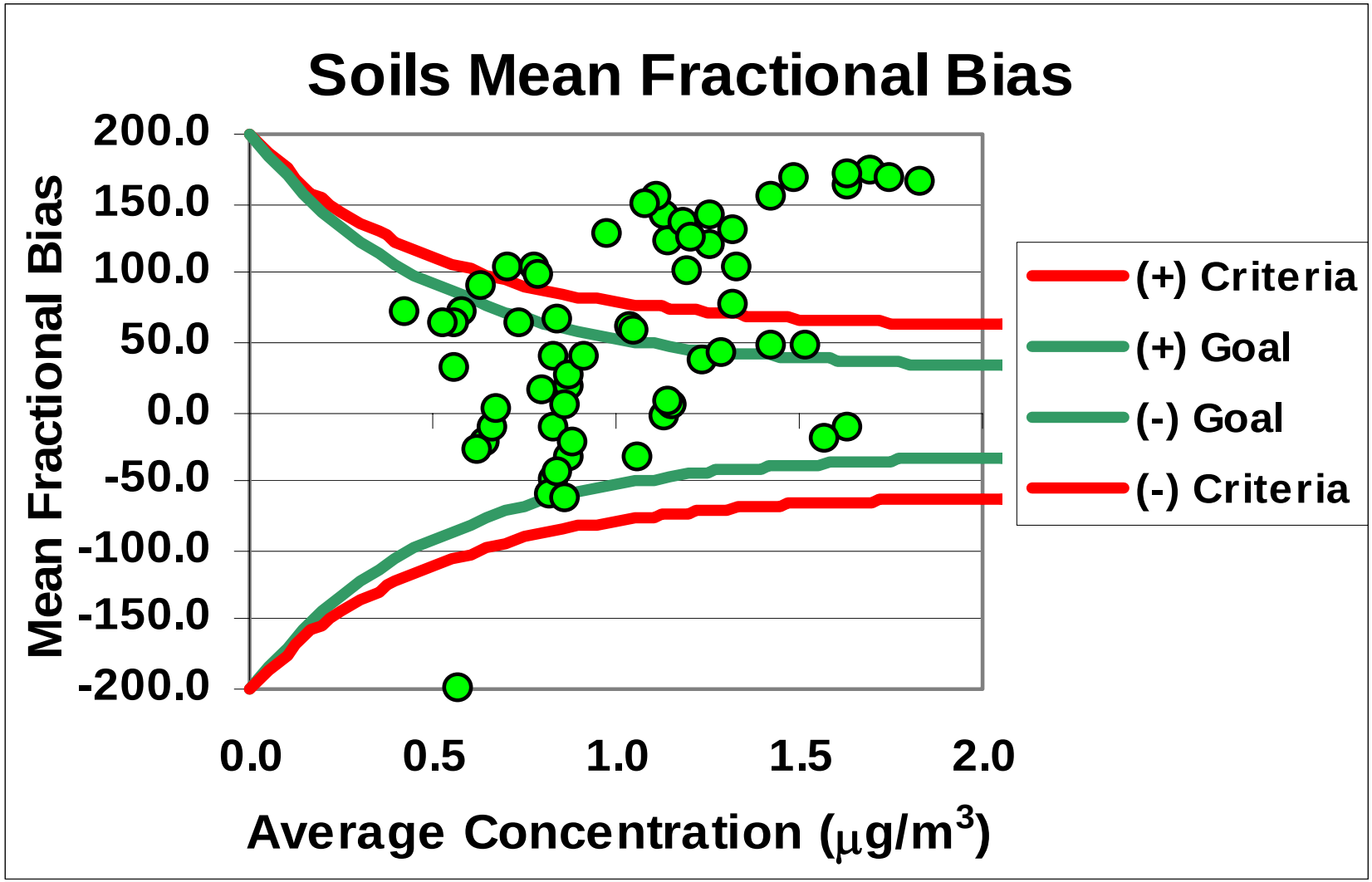
Organic Carbon (PM) Bias



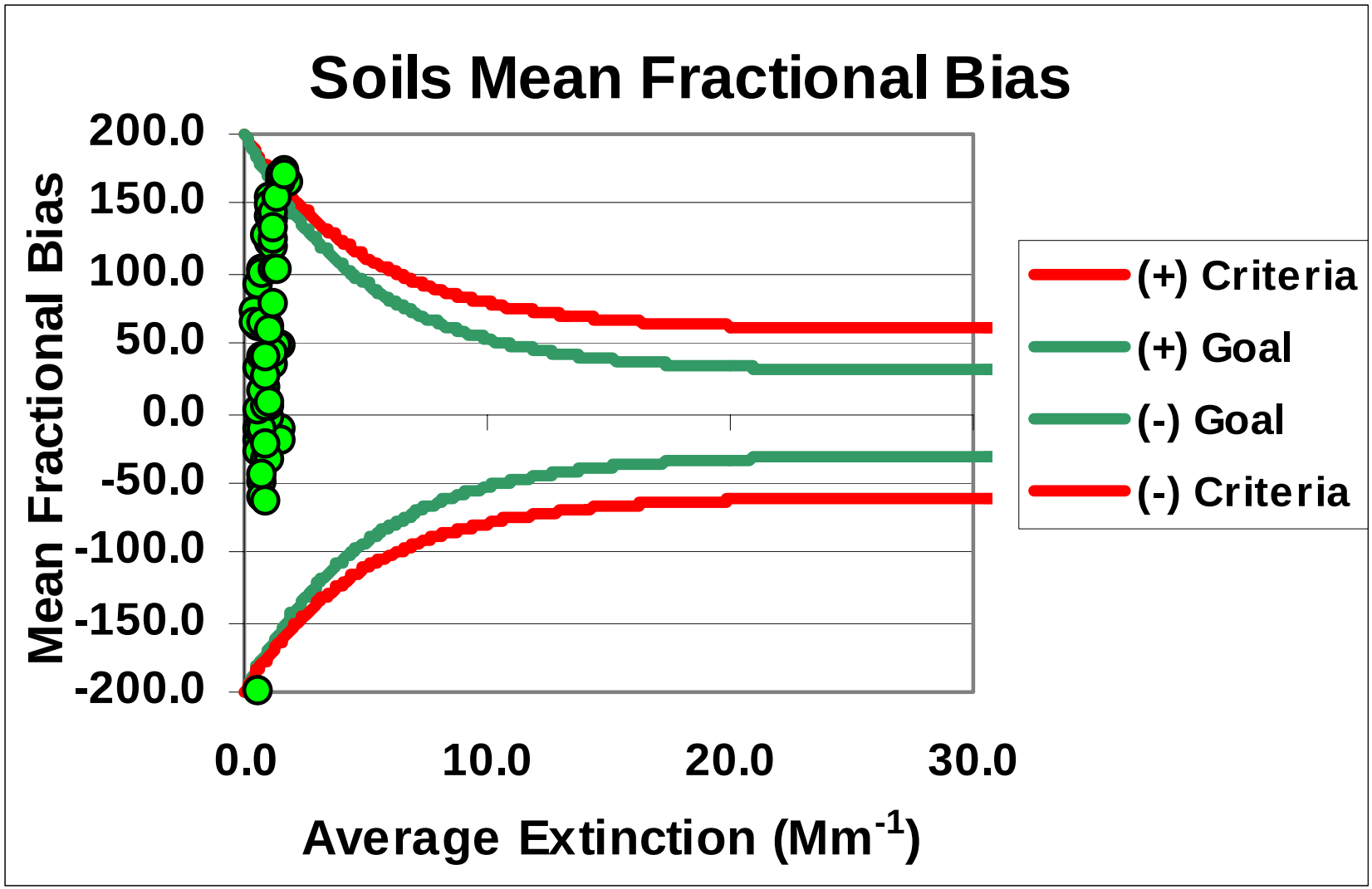
Organic Carbon (B_{ext}) Bias



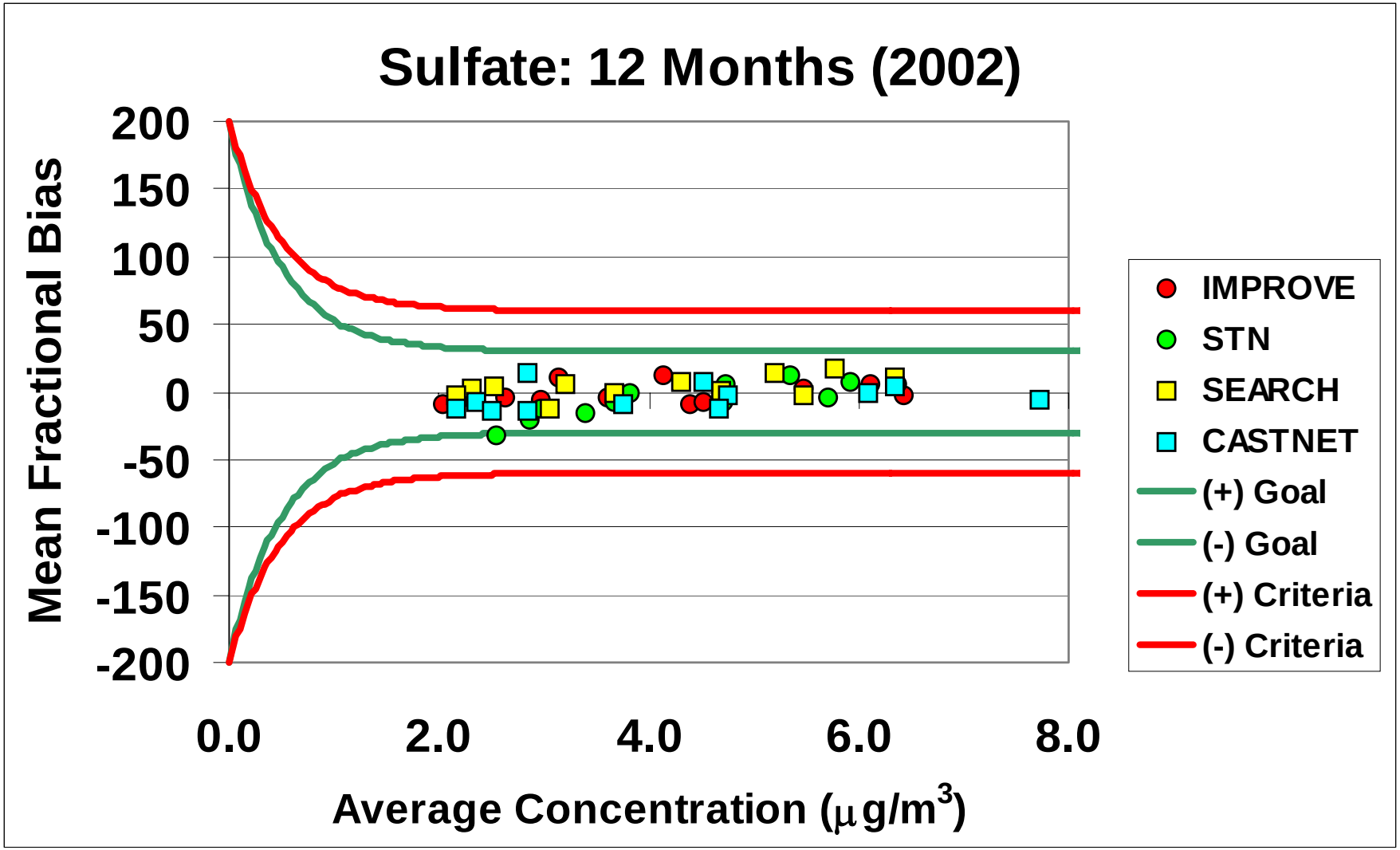
Soils (PM) Bias



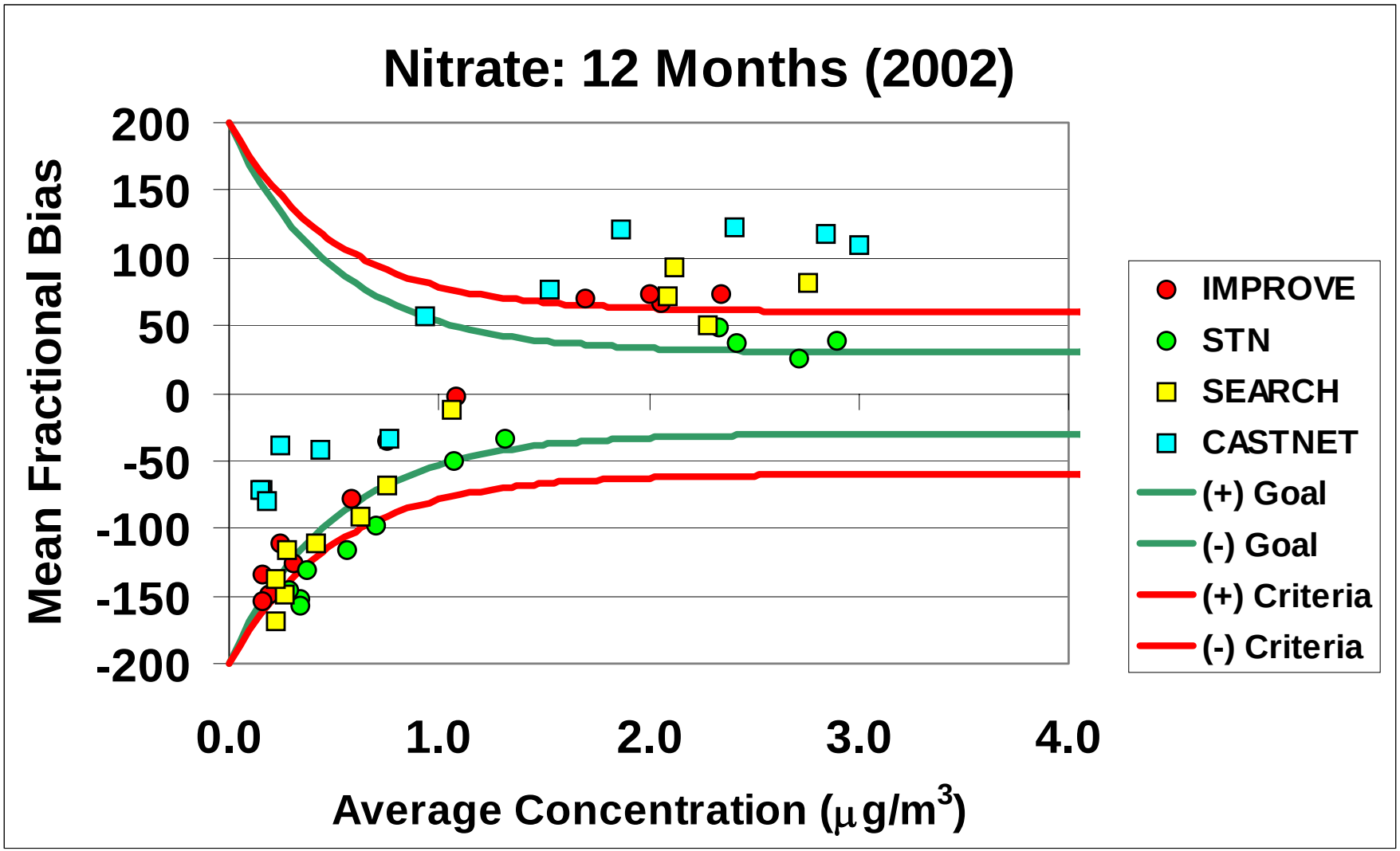
Soils (B_{ext}) Bias



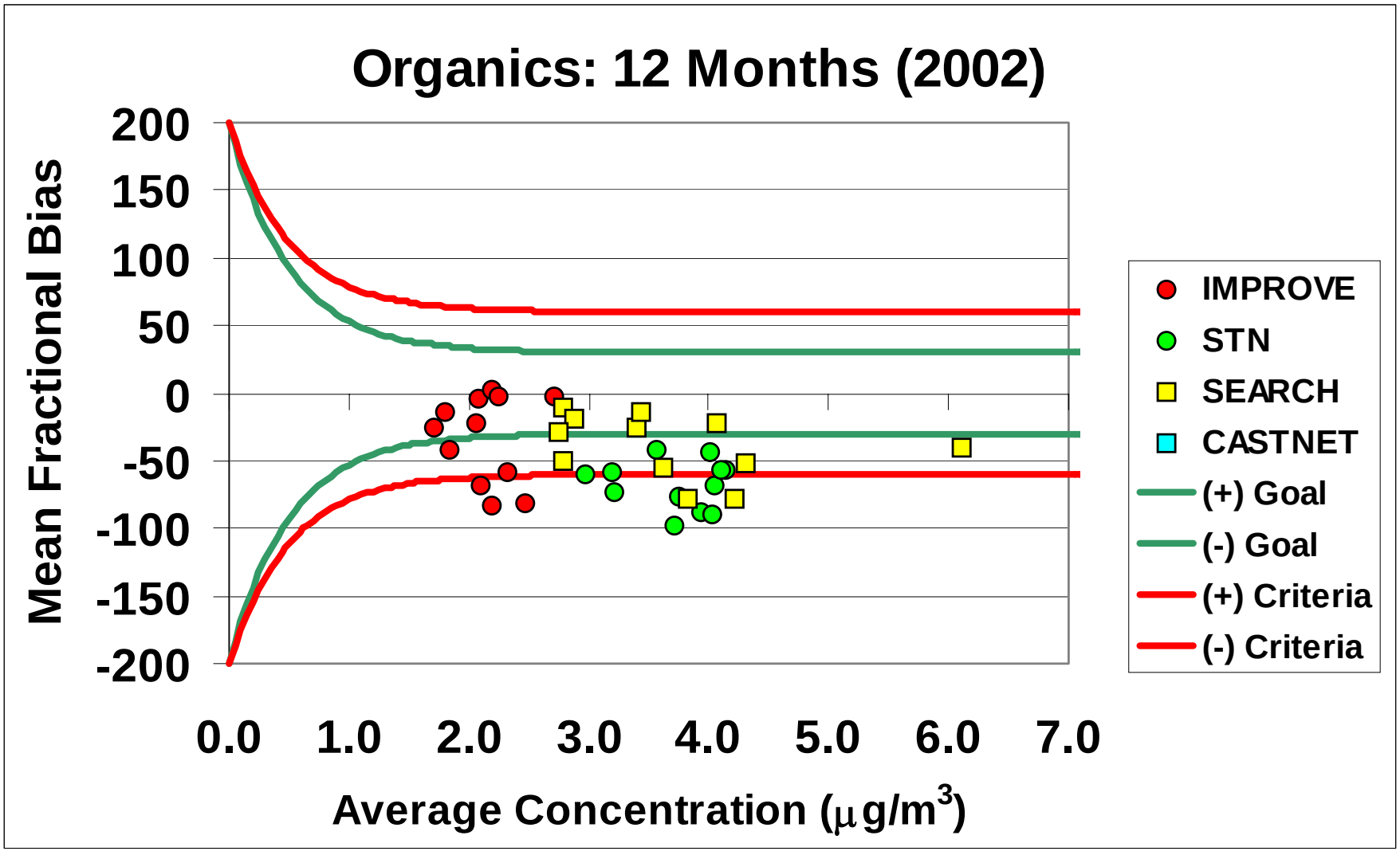
Sulfate Performance



Nitrate Performance



Organic Carbon Performance



Diagnostic Tests

Diagnostic Tests

- Sensitivity Analysis
  - Test if model performance is sensitive to a particular input or combination of inputs
  - Prioritize additional data gathering efforts
  - Assess the robustness of a control strategy
- Process Analysis
  - Assess relative importance of various model assumptions as well as physical and chemical phenomena contributing to a predicted concentration at a particular time and location

VISTAS Sensitivity Tests

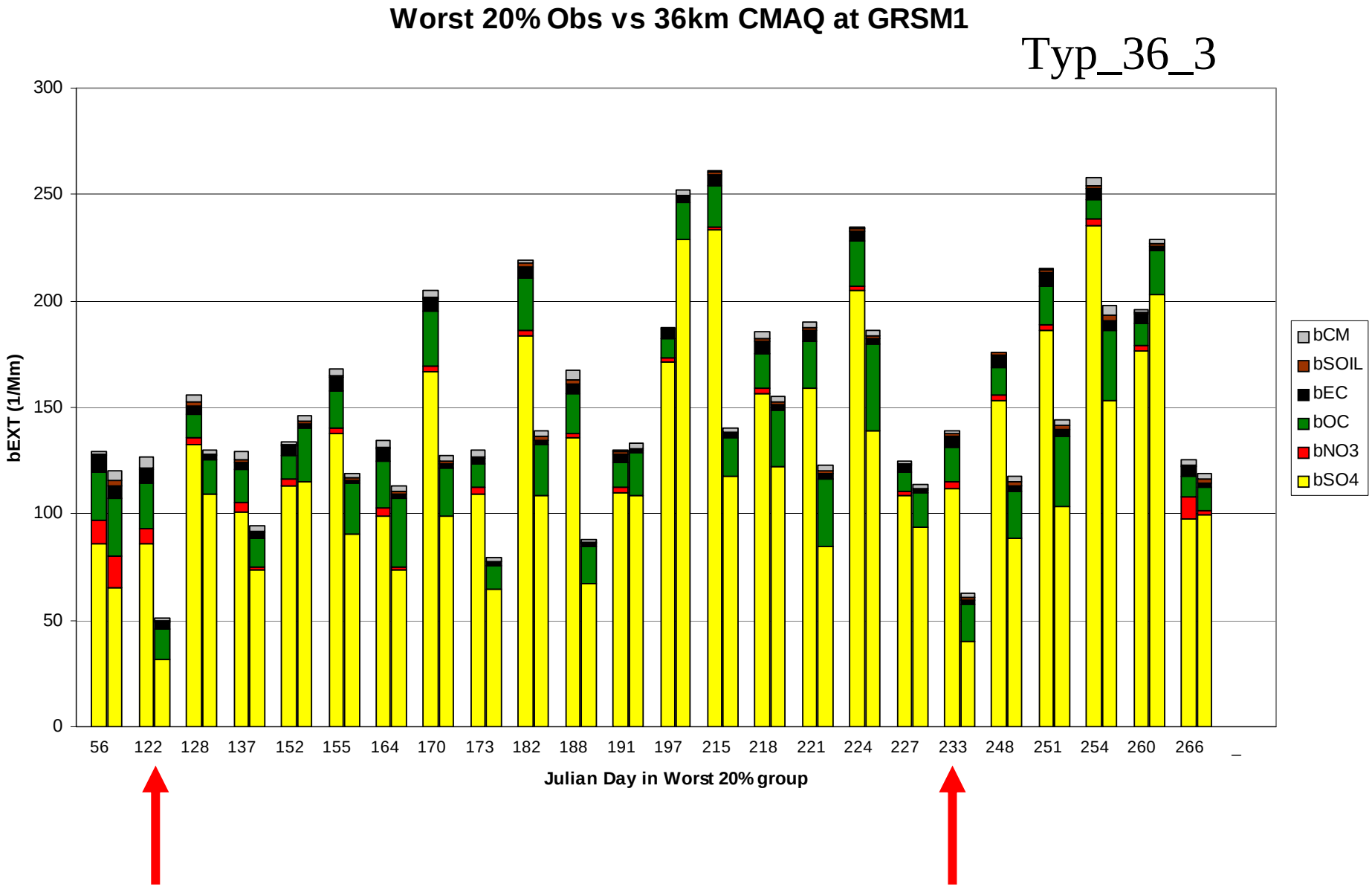
1) Fugitive Dust Transport Factor
   - FDTF=1.0 vs. FDTF=0.25 vs. FDTF=0.05
2) Number of Vertical Layers
   - NLAYS=34 vs. NLAYS=19
3) Vertical Diffusivity - Minimum Kz
   - Kz_min=1.0 vs. Kz_min=0.1
4) Ammonia Emissions (Winter Episode)
   - 0% Reduction vs. 50% Reduction
   - Standard Diurnal Pattern vs. Revised Diurnal Pattern
5) Mexican/Canadian Emissions
   - MX/CAN Emissions vs. No MX/CAN Emissions
6) Boundary Conditions
   - EPA Default vs. GEOS-CHEM
7) Boundary Layer Heights – Minimum PBLs
   - Standard PBL Code vs. Revised PBL Code

VISTAS Sensitivity Tests (cont.)

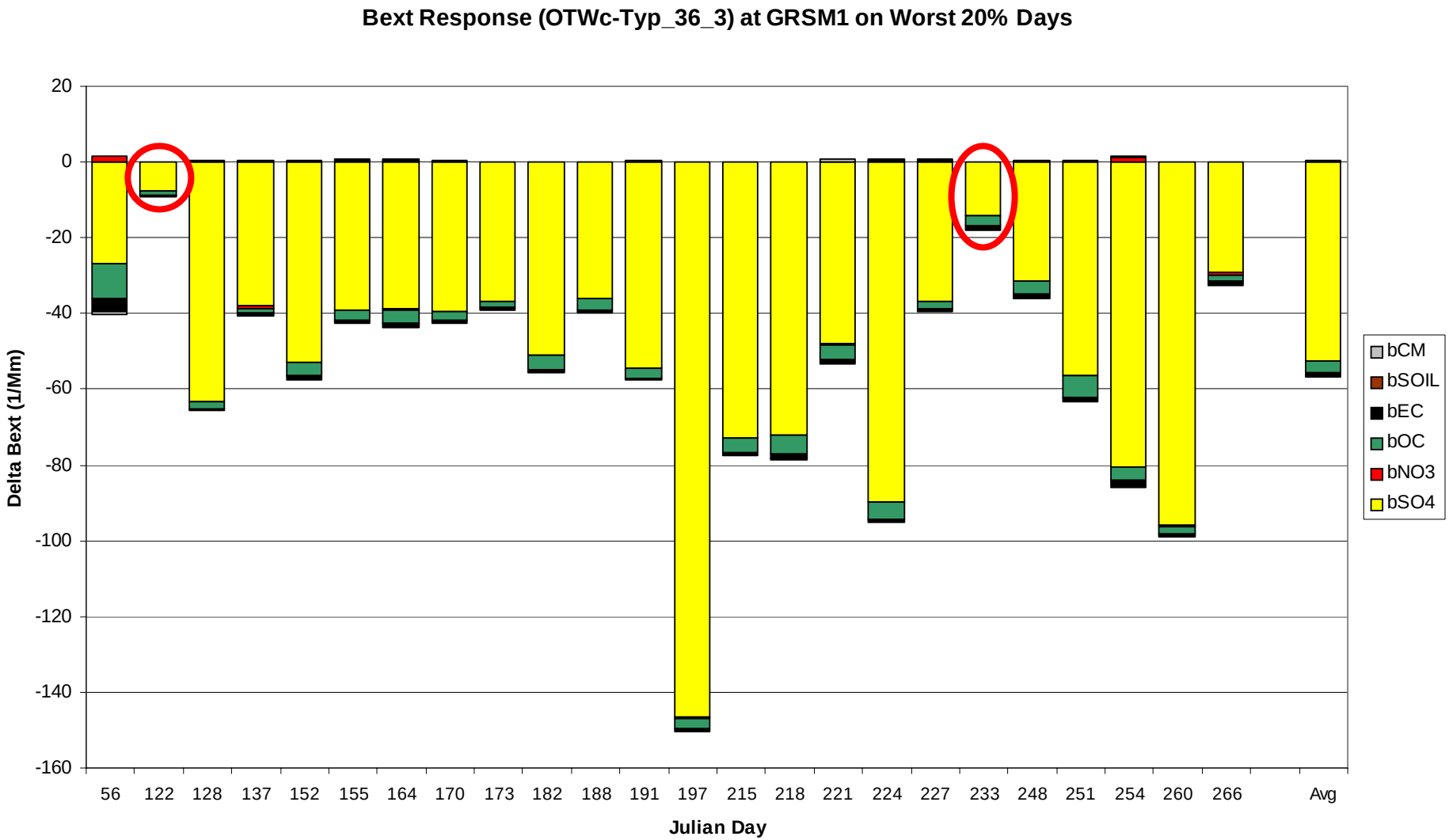
8. Alternative MM5 Configuration
   - Pleim-Xiu vs. NOAH-ETA-MY
9. Aerosol Mass Conservation
   - No Patch vs. GT Patch
10. Alternative Chemical Mechanisms
    - CB-IV vs CB4-2002 vs. SAPRC-99
11. Alternative Aerosol Module
    - AE3/ISORROPIA vs. CMAQ–AIM
12. Grid Resolution
    - 36 km vs. 12 km
13. Alternative Air Quality Model
    - CMAQ vs. CAMx

Model Performance and Model Response

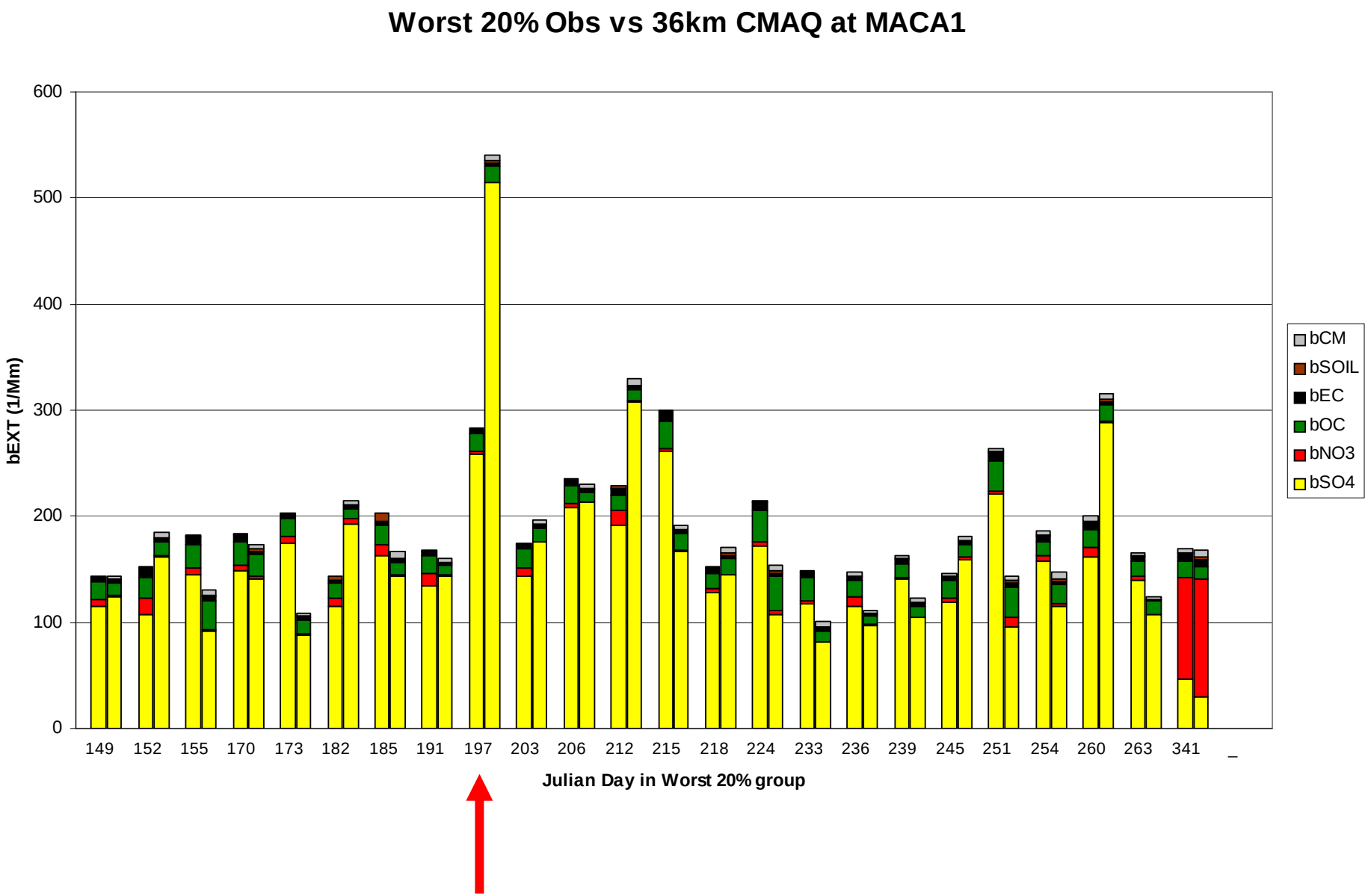
Performance at GRSM



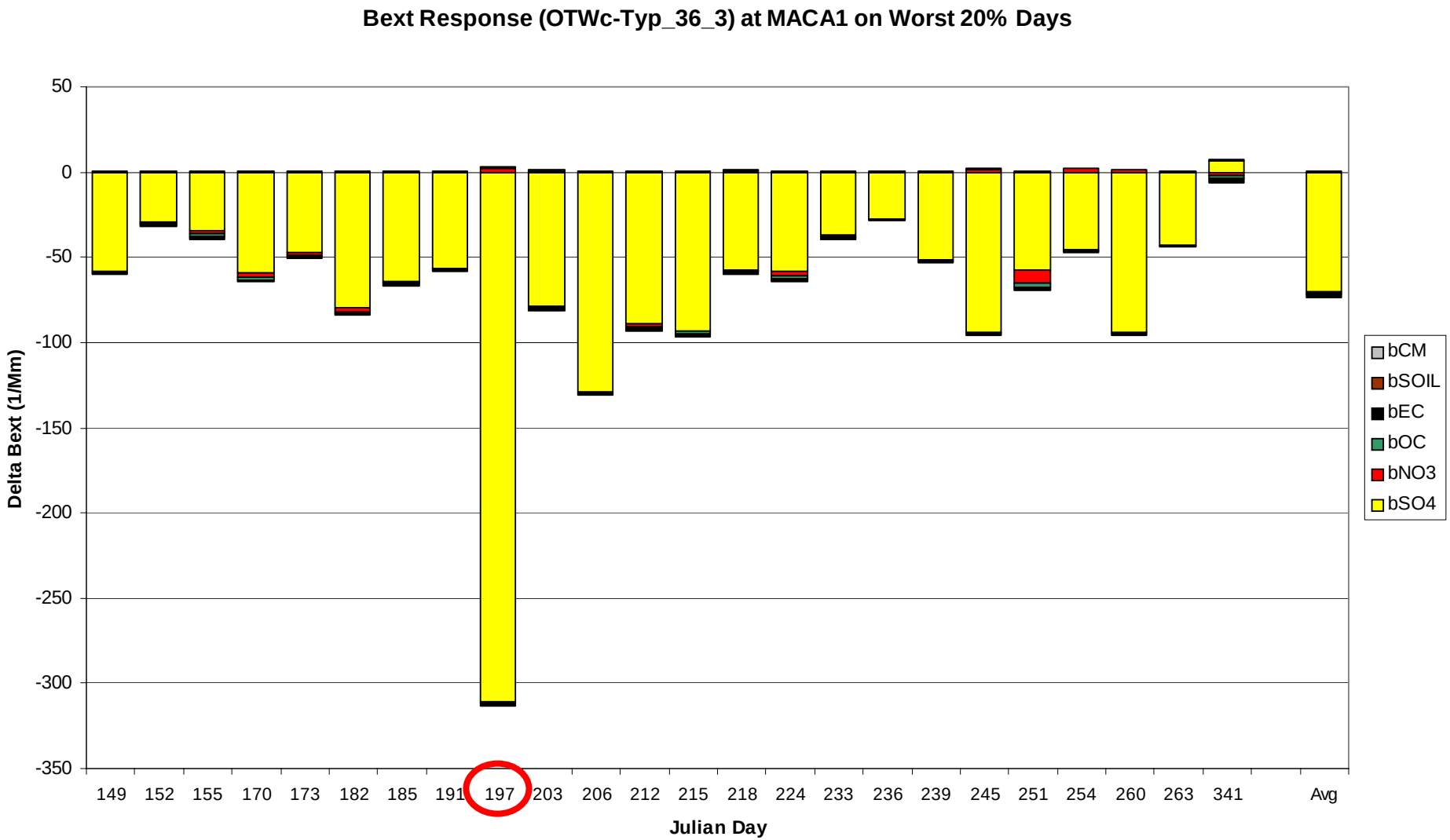
ΔB_{ext} at GRSM



Performance at MACA



ΔB_{ext} at MACA



Model Performance Software

- UC Riverside Package
- Alpine Geophysic's MAPS Package
- CENRAP's Package developed by AER
- Others….

Questions?