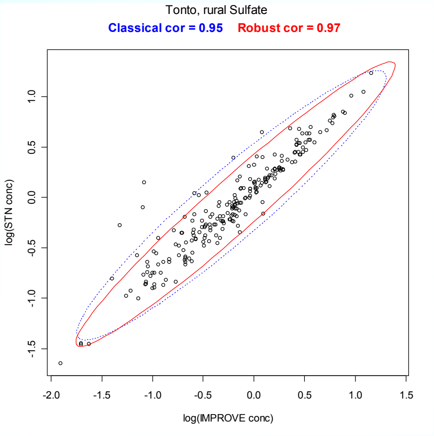
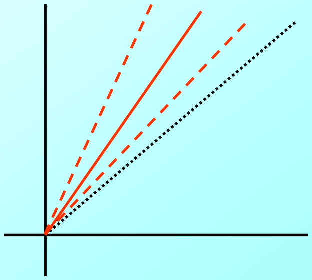
Combining Classical Analyses with Practicality

A Technical Approach for Method Comparison Studies using the STN and IMPROVE Intercomparison Data



Louise Camalier and Joann Rice*
U.S. EPA, OAQPS, AQAD
*project lead



H_0: b_0 =1
$H_{A:}$ b_0 ≠1

EPA Monitoring Conference
Data Validation & Analysis Session
Las Vegas, November 6, 2006

Critical Elements of the Study

1. STN and IMPROVE Intercomparison
2. Description of networks and study design
3. Data treatment and selection
4. The use of a Deming regression technique
5. Accounting for Measurement Error
6. Bootstrap Analysis

STN vs. IMPROVE Intercomparison

- STN is an urban companion network to the rural IMPROVE and is often used for urban versus rural comparison purposes
- Analysis was needed on the intercomparison data to determine whether organic carbon (OC) and elemental carbon (EC) sampling and analytical methods were comparable for trends and assessment of excess urban concentrations
- Analysis was also necessary to evaluate differences in OC and EC methods between networks

Description of networks and study design

STN

Speciation Trends Network:

- 54 sites, established in 2000
- Implemented to characterize the primary chemical components of mass in urban areas and to provide data for trends across the United States
- STN instruments are maintained by state and local monitoring technicians

IMPROVE Network

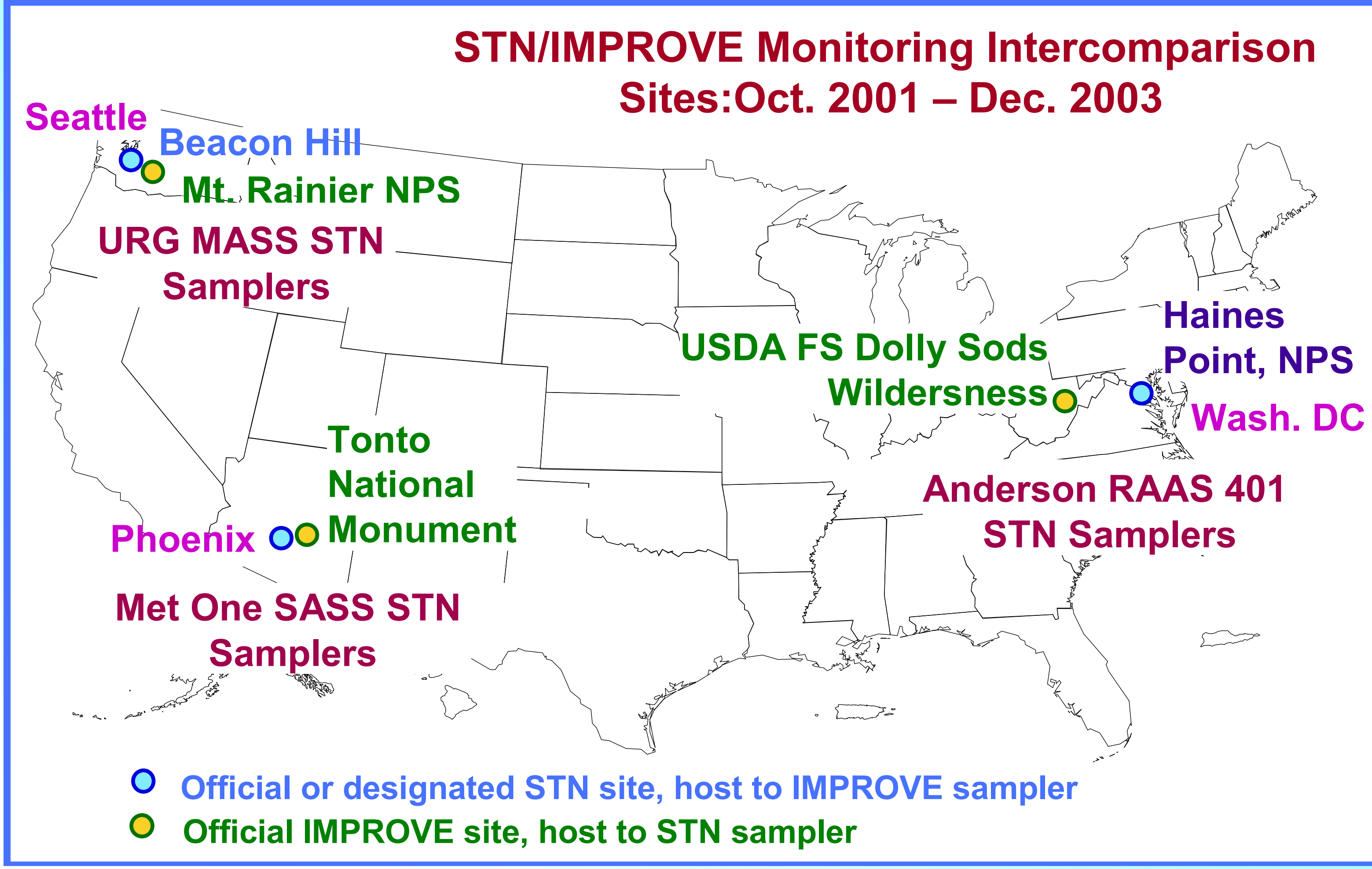
“Interagency Monitoring of Protected Visual Environments” network:

- 110 sites, established in 1985
- Supports the Regional Haze Rule, supports assessment and tracking of visibility impairment in Class I areas
- IMPROVE sampler filters are collected by park rangers, volunteers, etc.

Method Sampling and Analytic Differences for OC, EC

	STN	IMPROVE
Sampler type	MetOne, URG, Andersen	IMPROVE sampler
Filter size	46.2 mm	25 mm
Flow rate	6.7, 16.7, 7.3 L/min	22.8 L/min
Filter type (carbon)	Quartz	Quartz
Shipping temp.	Cooler	Ambient
Blank correction (for sampling artifacts)	Not reported with blank correction	Reported with blank correction
Thermal optical method	TOT (thermal optical transmittance)	TOR (thermal optical reflectance)
Operator	Monitoring technician	volunteer

Operated According to Each Network's Protocols



Technical Approach

- Data selection - remove outliers
- Determination of regression lines (slope and intercept)
- Definition of the measurement error structure
- Test for significant difference in slopes (shown) and intercepts (not shown)

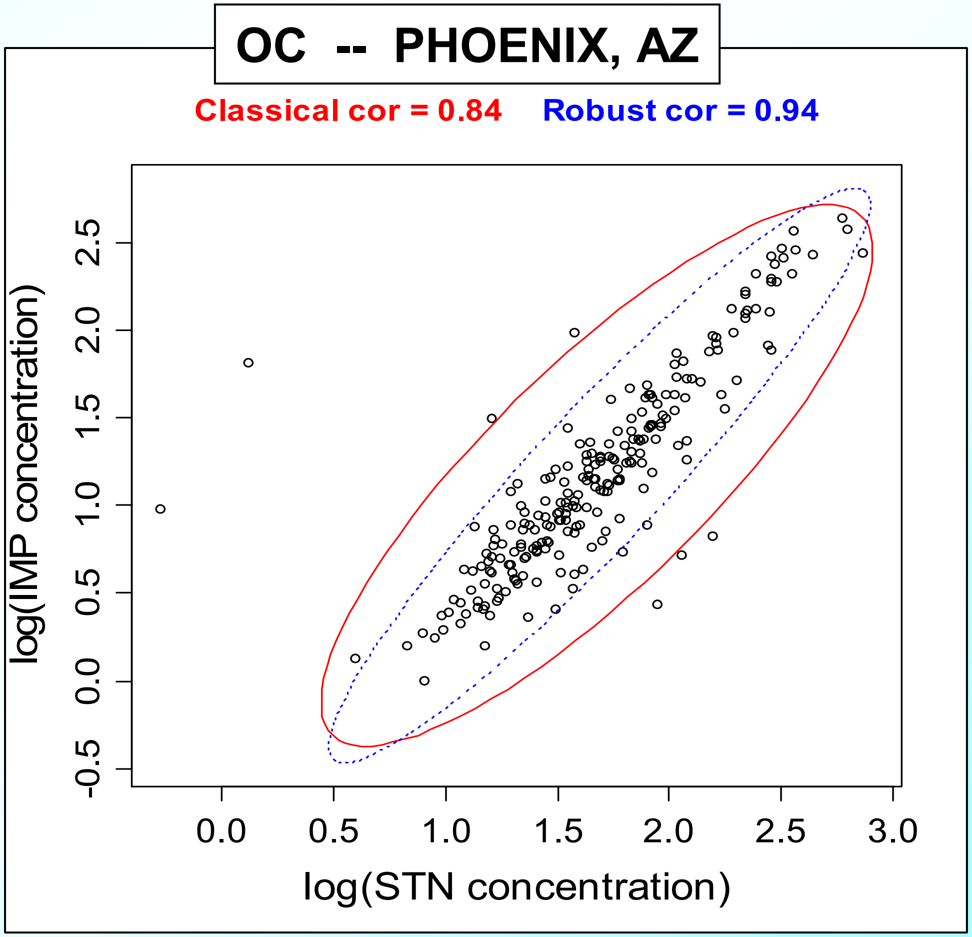
Data Selection

- Add blank to IMPROVE OC concentrations

- Log-transform data to improve linearity

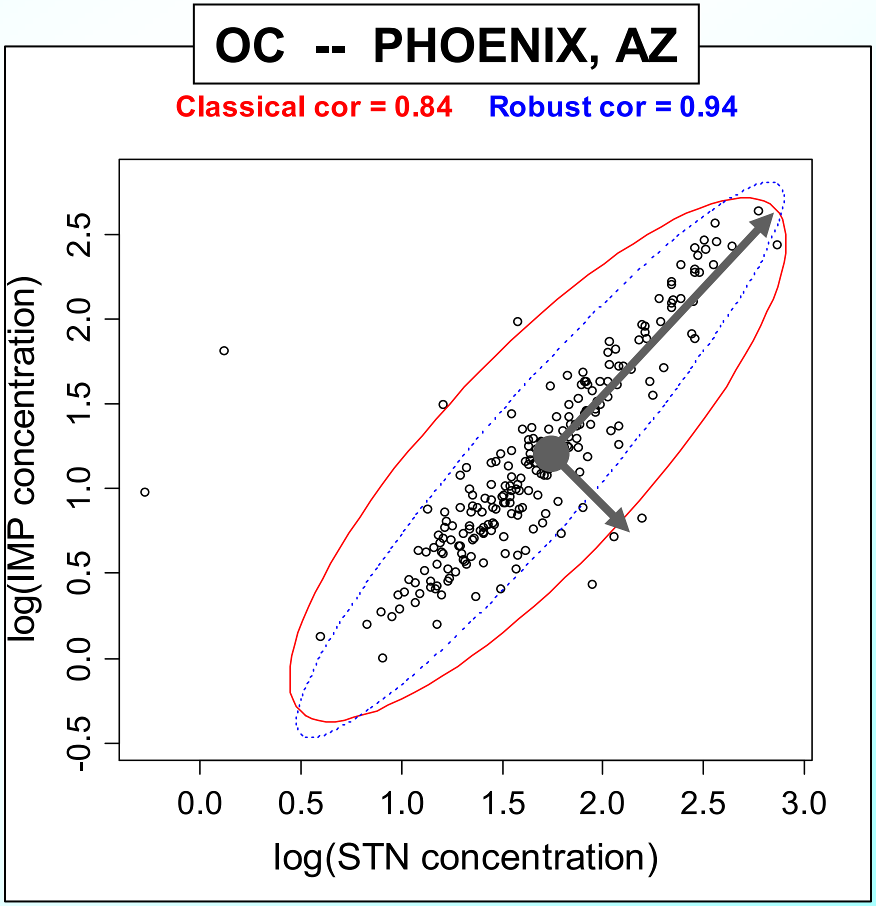
- Calculate *Mahalanobis' distance* for each parameter and site

- Reject data points at the *α=0.01 level*



Mahalanobis' Distance

- Method to detect bivariate outliers
- Uses the variable populations' covariance matrix to calculate a Mahalanobis' distance from the mean of the data
- Region of constant Mahalanobis' distance around the mean forms a two-dimensional ellipse around the most representative data points

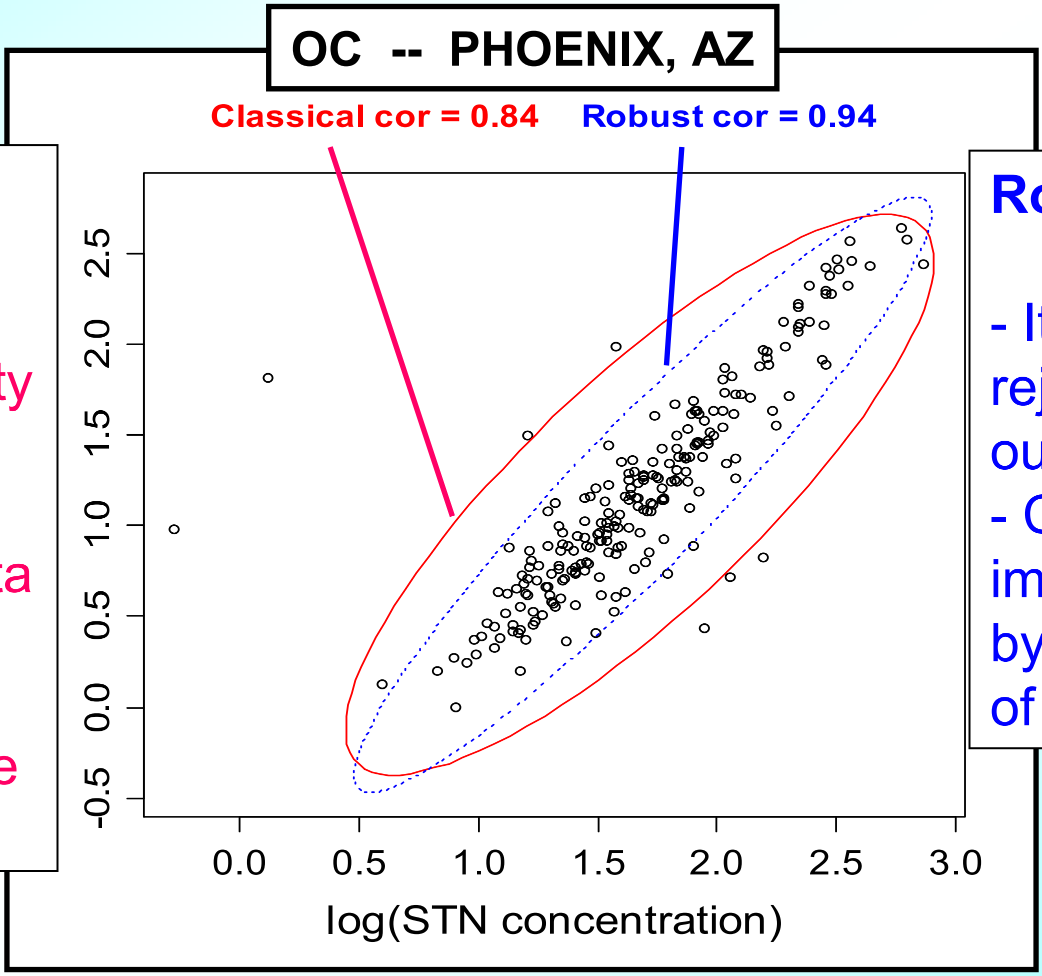


Classical vs. Robust Mahalanobis'

Which is better?

Classical

- Allows for more natural uncertainty
- More conservative (rejects fewer data points)
- Does not artificially improve correlation



Robust

- Iterative fitting rejects more outliers
- Can artificially improve correlation by restricting range of variability

Regression Technique

- Determine a technique that accounts for errors in both x and y
- Assess and assign measurement error in x and y
- Discussion of actual “Deming” regression
- Estimate standard errors for slope and intercept

Determine regression line through “errors-in-variables” regression

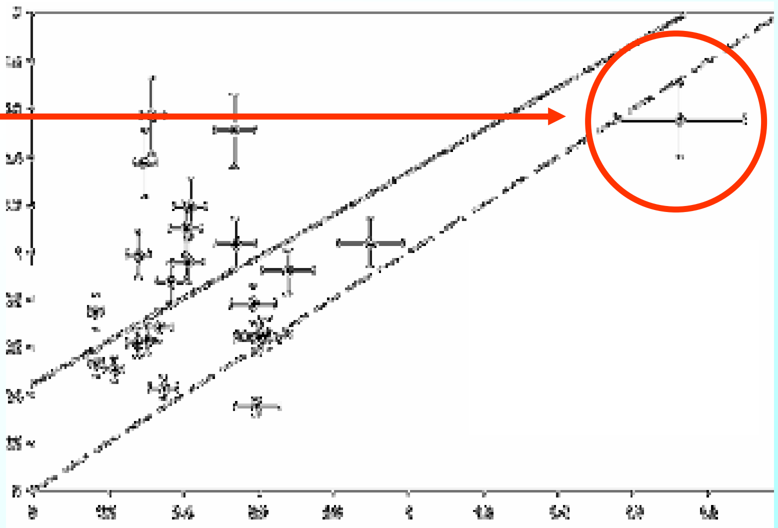
While classical regression assumes error in only one variable…

Deming Regression incorporates error in both x and y

Deming was used because both methods had error (neither method was “true”)

Requires a specified 2-dimensional measurement error structure for both variables

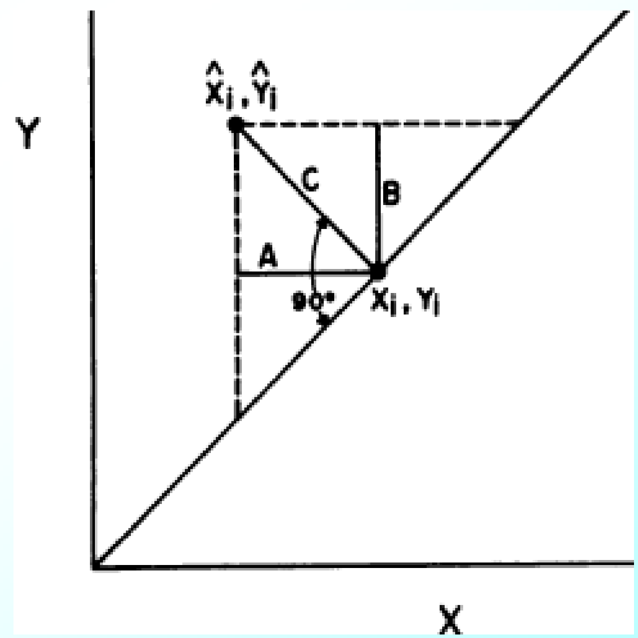
This takes into account proportional error with respect to concentration



… and weights data points in relation to this proportional error

How does errors-in-variables regression work?

- One method of Deming minimizes the perpendicular distances from the line to come up with slopes and intercepts
- This errors-in-variables ("Deming") method comes up with slopes and intercepts via an optimizing routine that minimizes the sums of squares on x and y



Given the specified error structure…

- "Deming" uses measurement error at each concentration to calculate slope and intercept to individually weight data points that feed into regression line

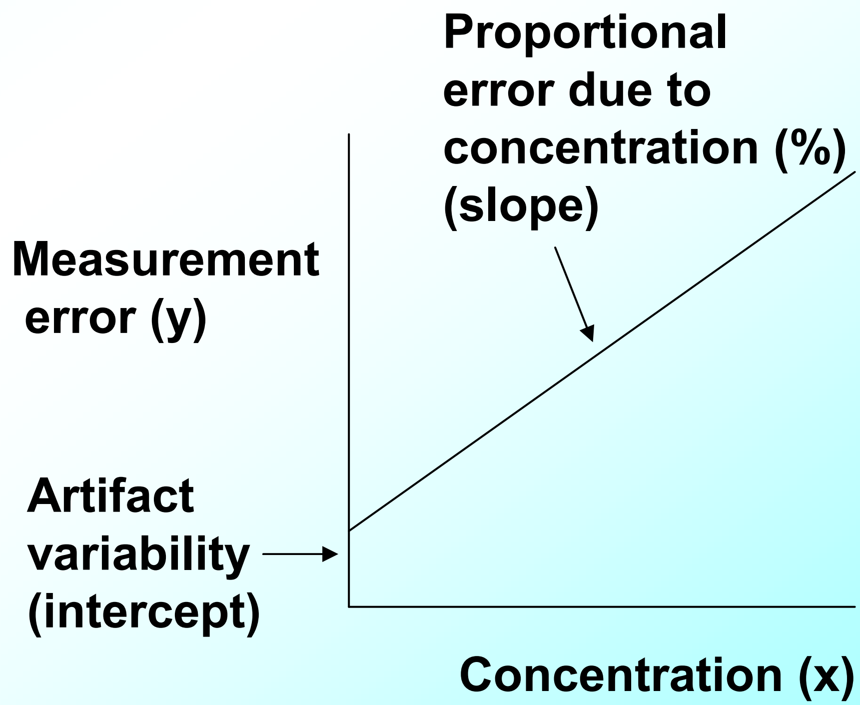
Depresses the dominance of high-leverage data points on slope

Defining the Error Structure

To best approximate the actual measurement error for both IMPROVE and STN…

An error structure was designed to consist of:

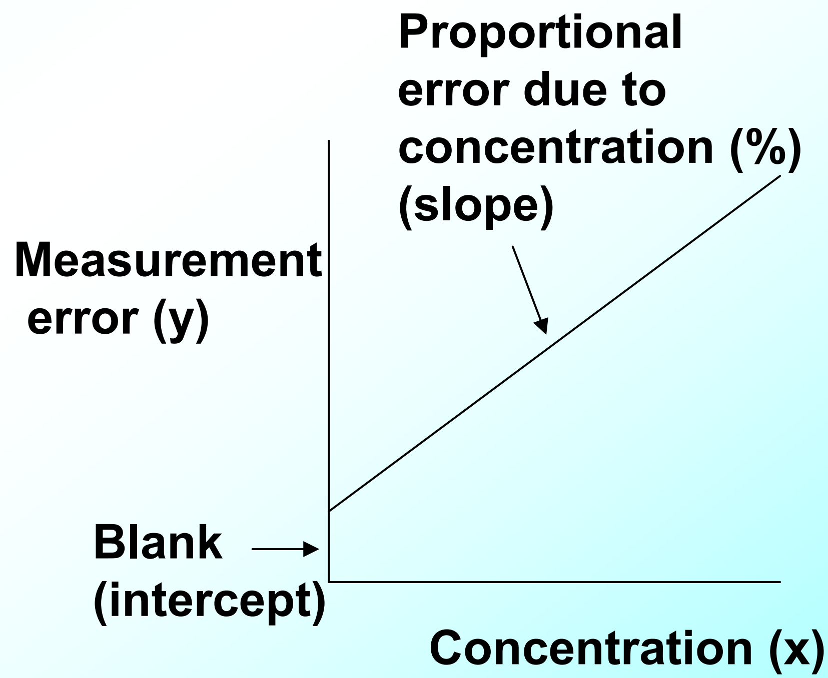
- Artifact (intercept) variability
- Reasonable estimate of proportional error due to concentration (slope variability)



Measurement Error: intercept component

Artifact variability:

- The STN and IMPROVE blank values incorporates error from both the lab and sampling setting (field and trip blanks)
- Variability in the STN blank was used as a surrogate for IMPROVE artifact variability in order to remain consistent with the "within method" variability defined (next slide)



Measurement Error: slope component

Estimate of proportional error due to concentration (slope):

- Collocated measurements were only available for STN samplers
- The "within method" variability procured from the collocated STN data was used to approximate the "within method" variability for IMPROVE
- Relative percent difference (RPD) calculated for this data

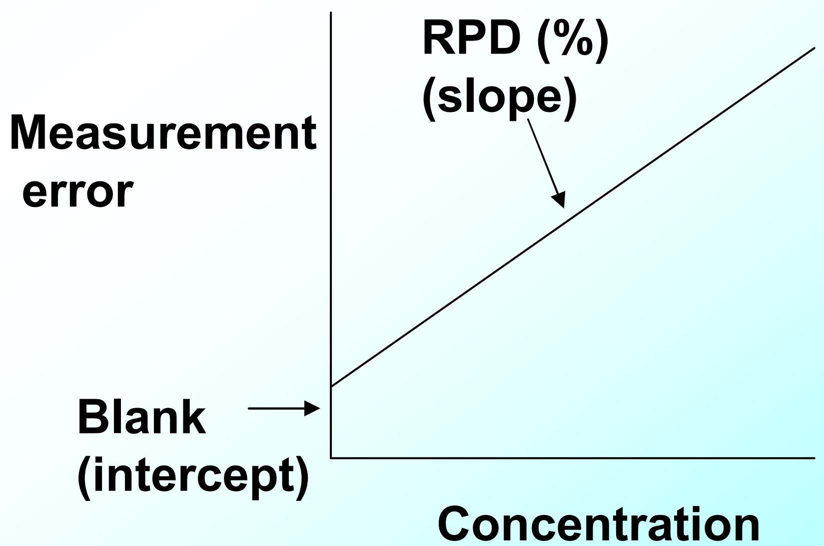
RPD = [(X-Y)/((X+Y)/2)] * 100
(using collocated STN monitors)

Error Structure

error(y) = std(STN blank value by area) + median(RPD of colloc STN instr.)*x

error(x) = std(STN blank value by area) + median(RPD of colloc STN instr.)*y

(where y is the IMPROVE concentration and x is the STN concentration)



Estimating standard error of slope and intercept

No classical methods are available to feasibly calculate standard error, therefore bootstrapping was needed to produce an approximation

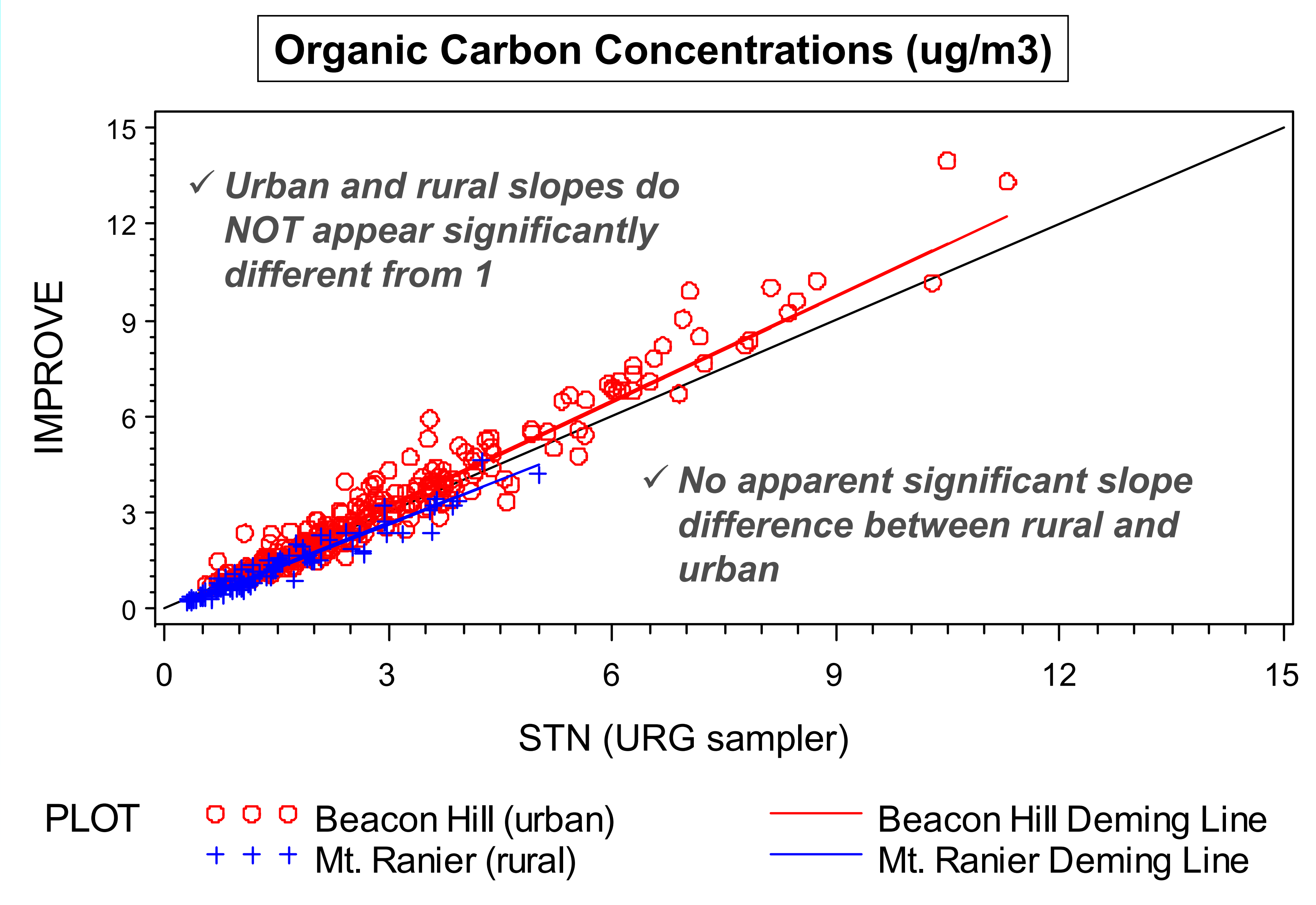
Bootstrapping is a way to estimate the sampling distribution of an estimator by resampling with replacement from the original sample many times to appropriately estimate the standard error of the slopes and intercepts

- ✓ The Bootstrapping was done 500 times to calculate 99% CIs on slopes and intercepts
- ✓ 99% CIs on slopes were used for hypothesis testing to assess differences between methods among urban and rural sites

OC Results

S1

Comparison of Beacon Hill and Mt. Rainier

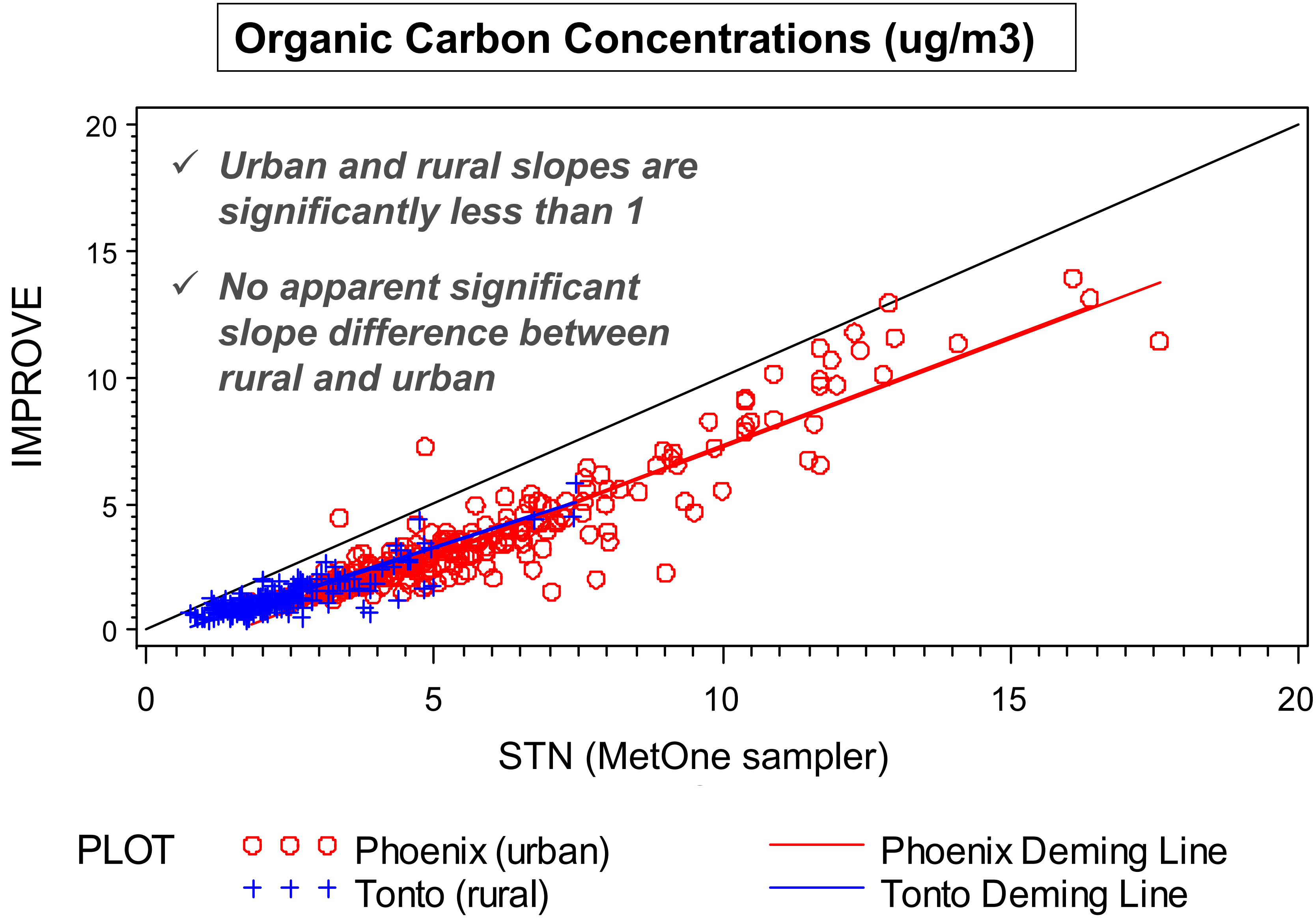


Slide 22

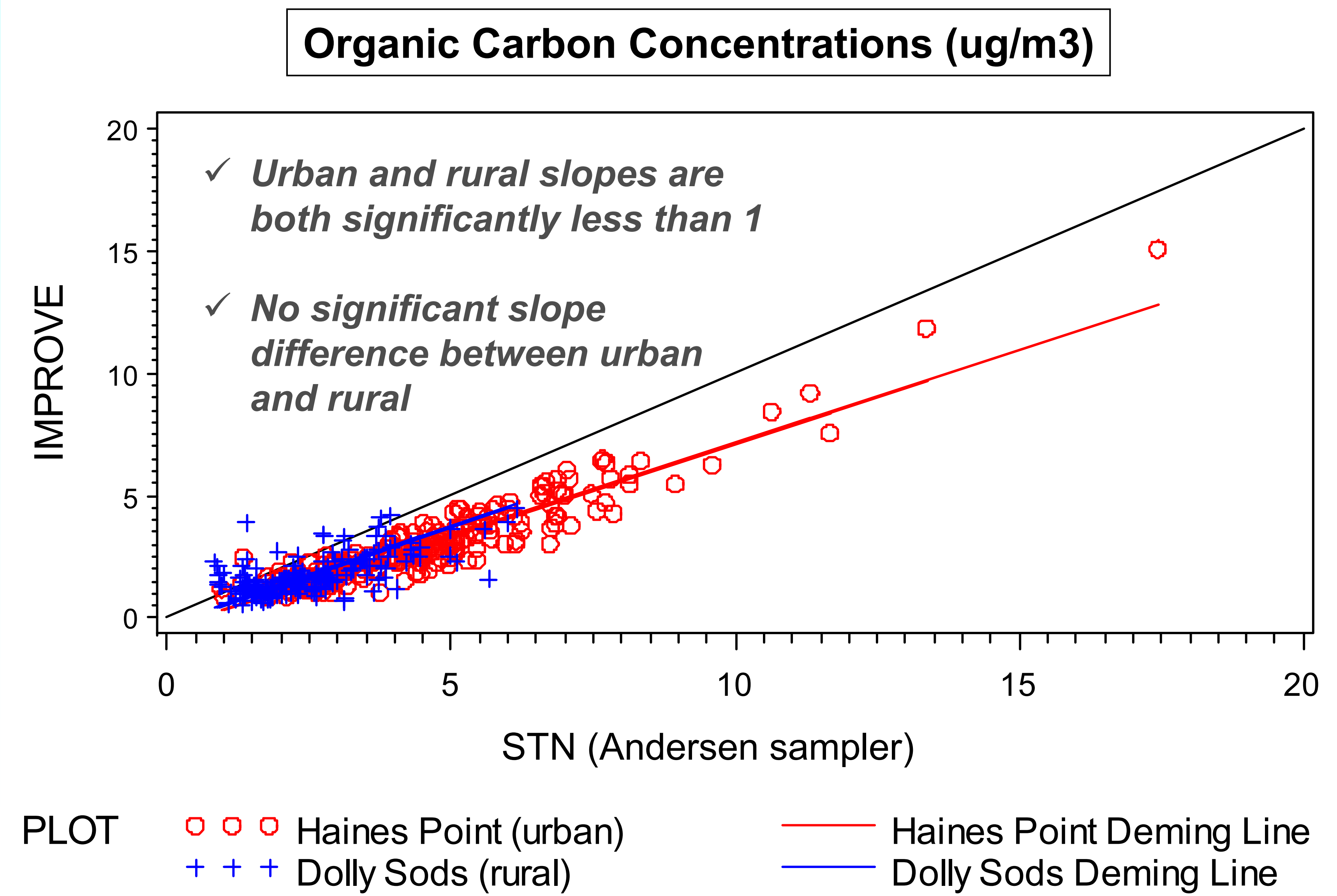
S1

explain plot
imp vs stn and urban vs rural for 3 yrs
explain the dots and lines and the 1to1 line
SMTG#1, 4/12/2006

Comparison of Phoenix and Tonto

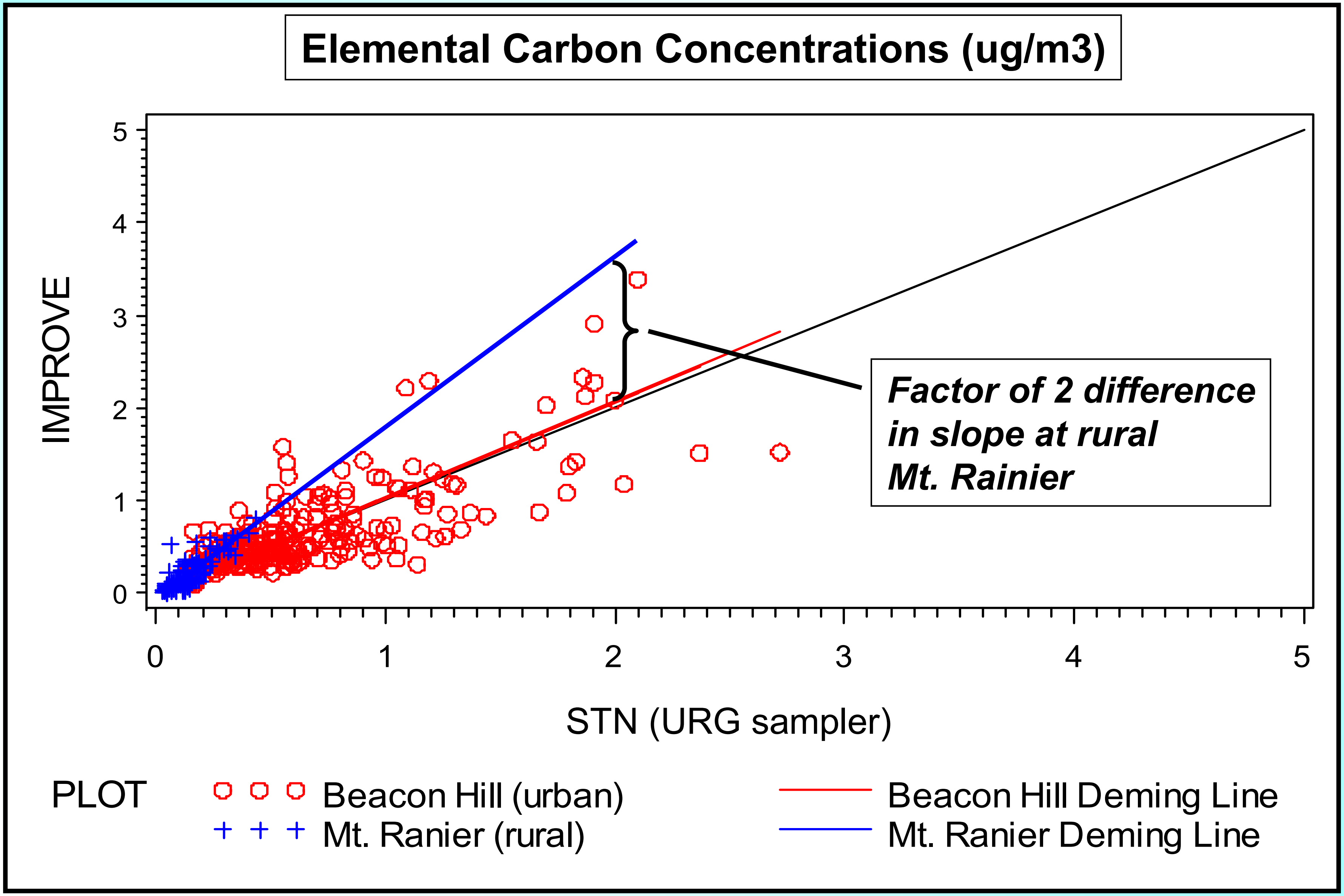


Comparison of Haines Point and Dolly Sods

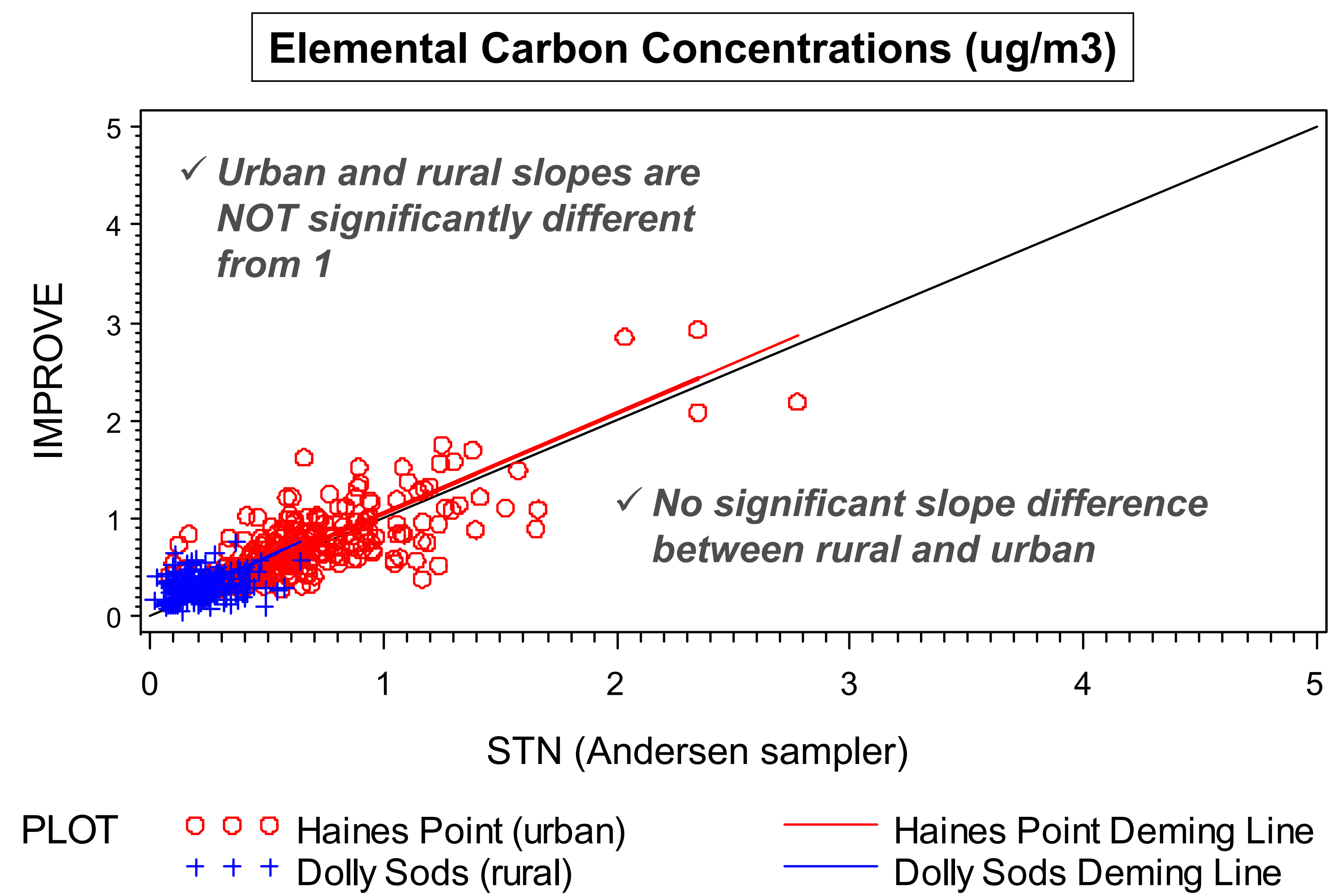


EC Results

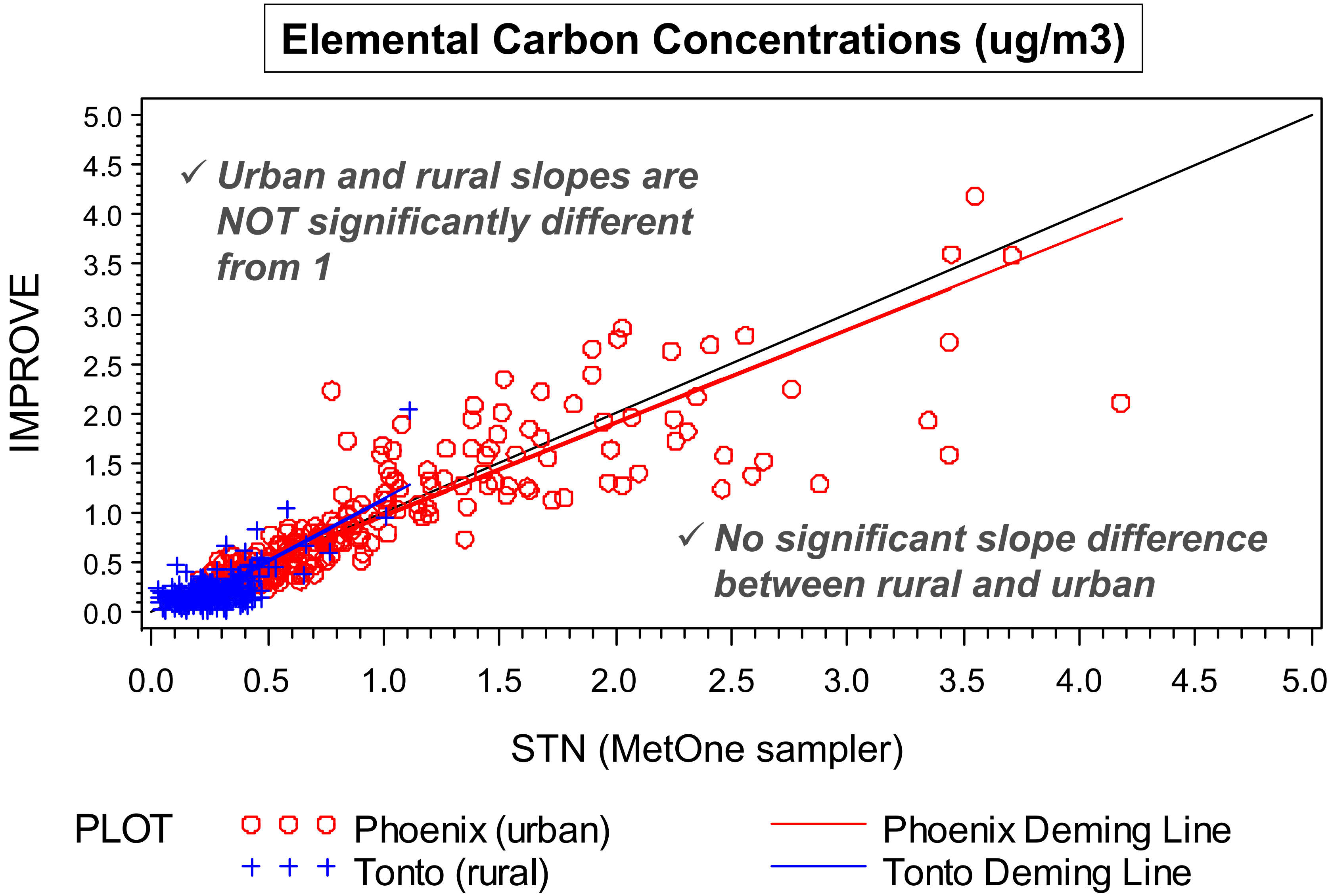
Comparison of Beacon Hill and Mt. Rainier



Comparison of Haines Point and Dolly Sods



Comparison of Phoenix and Tonto



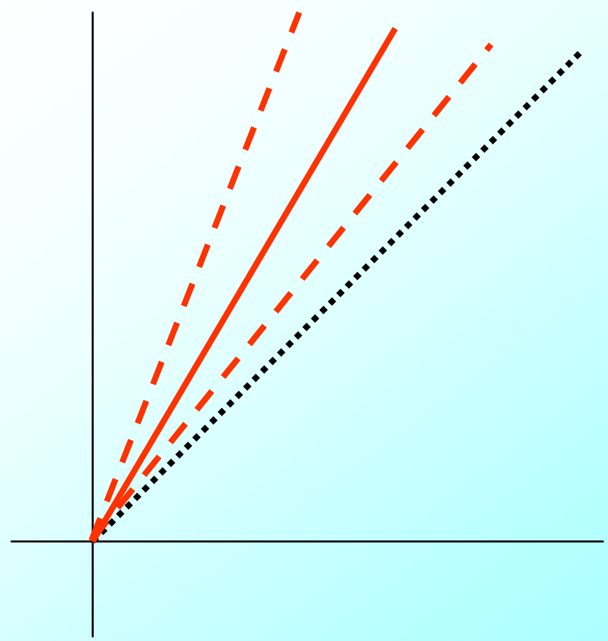
How did we test for significance?

Testing whether slope is significantly different from one

Is one method consistently measuring higher (or lower) than the other?

H_0: $b_0 = 1$
H_A: $b_0 \neq 1$

Use 99% Confidence intervals to determine whether 1 is contained

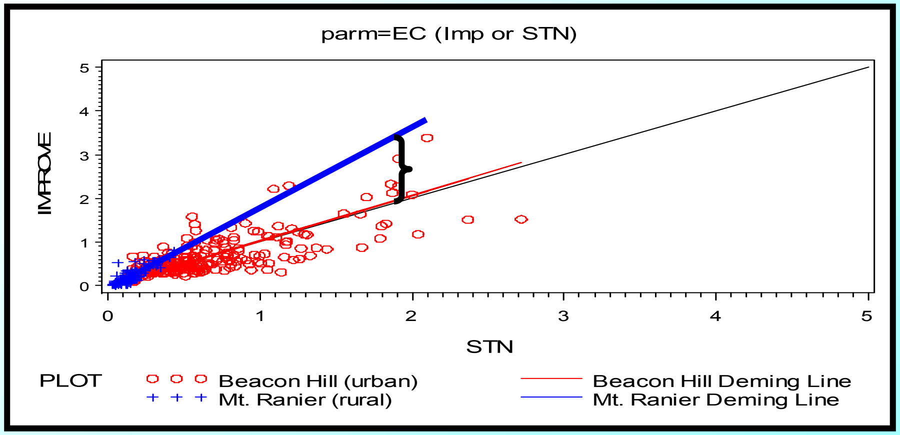


99% Confidence Intervals on slopes

	Arizona – 99% CI on slope		Washington, D.C. – 99% CI on slope		Washington State – 99% CI on slope	
Param	Urban	Rural	Urban	Rural	Urban	Rural
OC	(0.79, 0.86)	(0.66, 0.82)	(0.70, 0.82)	(0.71, 0.90)	(1.03, 1.15)	(0.84, 0.99)
EC	(0.81, 1.07)	(0.91, 1.59)	(0.87, 1.17)	(0.94, 1.19)	(0.85, 1.23)	(1.66, 2.06)

Value shows significant differences from 1

Flow rate is an important factor

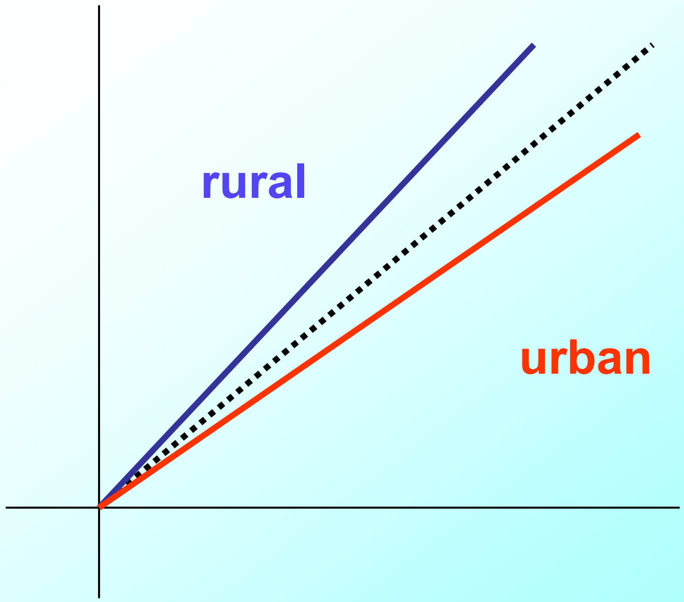


Testing for significant difference between urban and rural Slopes

Is the relationship between STN and IMPROVE consistent in rural vs. urban in an area?

$$H_0: b_{0_{diff}} = 0$$
$$H_A: b_{0_{diff}} \neq 0$$

Use the t-value to test for an urban versus rural slope difference



Variance is obtained by pooling the errors from both urban and rural slopes (from bootstrap analysis)

T-values for Urban vs. Rural Slopes

	Arizona area				Washington, D.C. area			Washington State area		
Param	urban slope	rural slope	se	t-value	urban slope	rural slope	t-value	urban slope	rural slope	t-value
OC	0.86	0.74	0.04	2.83	0.76	0.8	-1.09	1.09	0.92	4.86
EC	0.94	1.25	0.1	-2.18	1.02	1.06	-0.54	1.04	1.86	-7.7

Values shows significant urban versus rural slope differences

Values beyond ~+/-2.33 indicate significant difference at α=0.01

Summary and Conclusion

- Developed a systematic, technical approach for comparing methods and testing for significant differences
- Determine “practical” differences from test results
- Applied the approach to collocated STN and IMPROVE samplers in urban and rural areas
- May be useful in other method comparison studies

References

Tracy Klamser-Williams, US EPA, ORIA, Las Vegas, NV; Paul A. Solomon, Peter Egeghy, US EPA, ORD, Las Vegas, NV; Dennis Crumpler, Joann Rice, James Homolya, Neil Frank, OAQPS, RTP Marc Pitchford, US EPA/NOAA, OAQPS, Las Vegas, NV; Lowell Ashbaugh, Charles McDade, UC Davis, Sacramento, CA; James Orourke, James Flanagan, Edward Rickman, Research Triangle Institute, RTP, NC; *"Multi-Site Comparison Of Mass And Major Chemical Components Obtained By Collocated STN and IMPROVE Chemical Speciation Network Monitors"*, presented at the AAAR Annual Conference in Atlanta, GA; 2004.

Cornbleet, J., Gochman, N. "Incorrect Least-Squares Regression Coefficients in Method-Comparison Analysis," *Clinical Chemistry*, Vol. 25, No. 3, pp. 432-438, 1979.

Chow, J., et. al. "The DRI Thermal/Optical Reflectance Carbon Analysis System: Description, Evaluation and Applications in U.S. Air Quality Studies," *Atmospheric Environment*, Vol. 27a, No. 8, pp. 1185-1201, 1993, Great Britain.

Watson, J.G., Chow, J.C., Antony Chen, L.-W. "Summary of Organic and Elemental Carbon/Black Carbon Analysis Methods and Intercomparisons," *Aerosol and Air Quality*, Vol. 5, No. 1, pp. 65-102, 2005.