USGS
science for a changing world

Assessment of Regression-Model for Estimating Solute Loads

Nitrogen Processes in Large Rivers Workshop –New Orleans, La

August 24 and 24, 2005

This presentation was prepared by Brent Aulenbach, USGS, Atlanta, Ga.

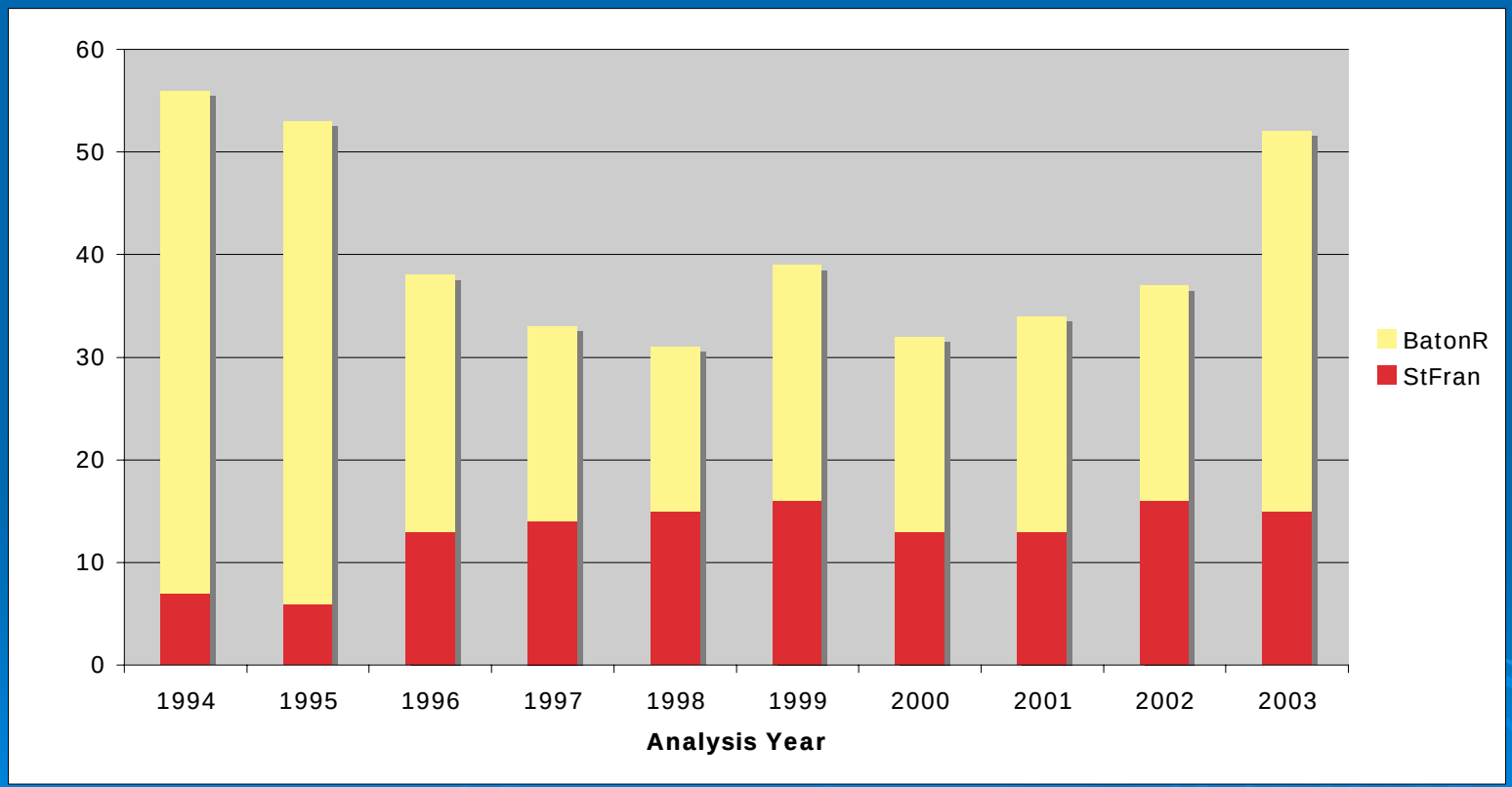
770-903-9148

Data Set

- St. Francisville and Baton Rouge data combined
- $NO_3 + NO_2$
- 10-years: 7/93 - 6/03
- 405 samples
- 68% samples from Baton Rouge

USGS
science for a changing world

Sampling Coverage



QW Compatibility

- St. Francisville: EWI
- Baton Rouge: Composite upper 20’ of water column taken at end of pier
- 16 samples taken same day at both sites
  - 5/91 - 9/03
  - Baton Rouge samples 0.03 ppm higher
  - Variability: Standard deviation 0.09 ppm
  - Average concentration 1.4 ppm
  - Significant figures often 0.1 ppm

USGS
science for a changing world

Regression Model Methodology

- Sampling
- Model Structure
- Calibration Period
- Ends vs. Middle of Calibration Period

Subsampling

- Quarterly (100% coverage)
- Bi-monthly (100%)
- Monthly (99%)
- Bi-weekly (90%)
- All (405 samples, 1 sample / 9 days)

Regression Model Approach to Estimate Loads

- ESTIMATOR code
- Flow terms lnQ and lnQ^2
- Time trend terms t and t^2
- Flow and time terms centered
- Seasonal sine and cosine terms

Regression Model Structures

- Flow terms (lnQ and lnQ^2)
  - At site: Tarbets Landing + Old River Outflow
  - At site + upstream: Lagged upstream flows Ohio River at Metropolis and Mississippi River at Thebes (10-day lag)
- Time terms
  - No time terms
  - Linear time term t
  - t and t^2 terms

Calibration Period

- Vary between 3 and 10-years
- Increment through 10-year period

Accuracy

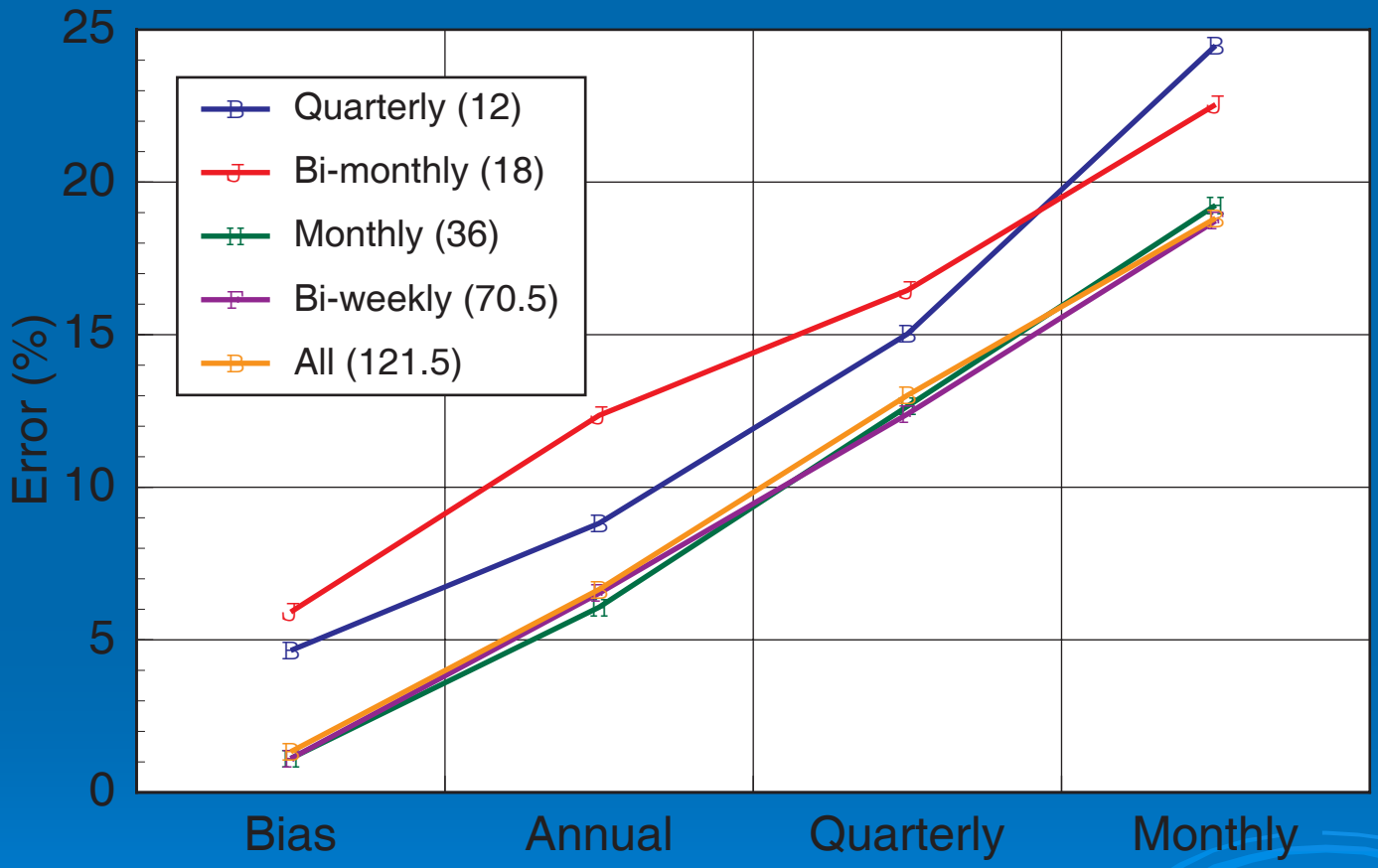
- Bias determined by comparing with composite method, all samples
- Precision determined for different time intervals
  - Annually
  - Quarterly
  - Monthly

Regression Model Performance

- Concentration model R^2 0.23 - 0.90 (most between 0.4 and 0.65)
- Load model R^2 0.77 - 0.98

Sampling Frequency

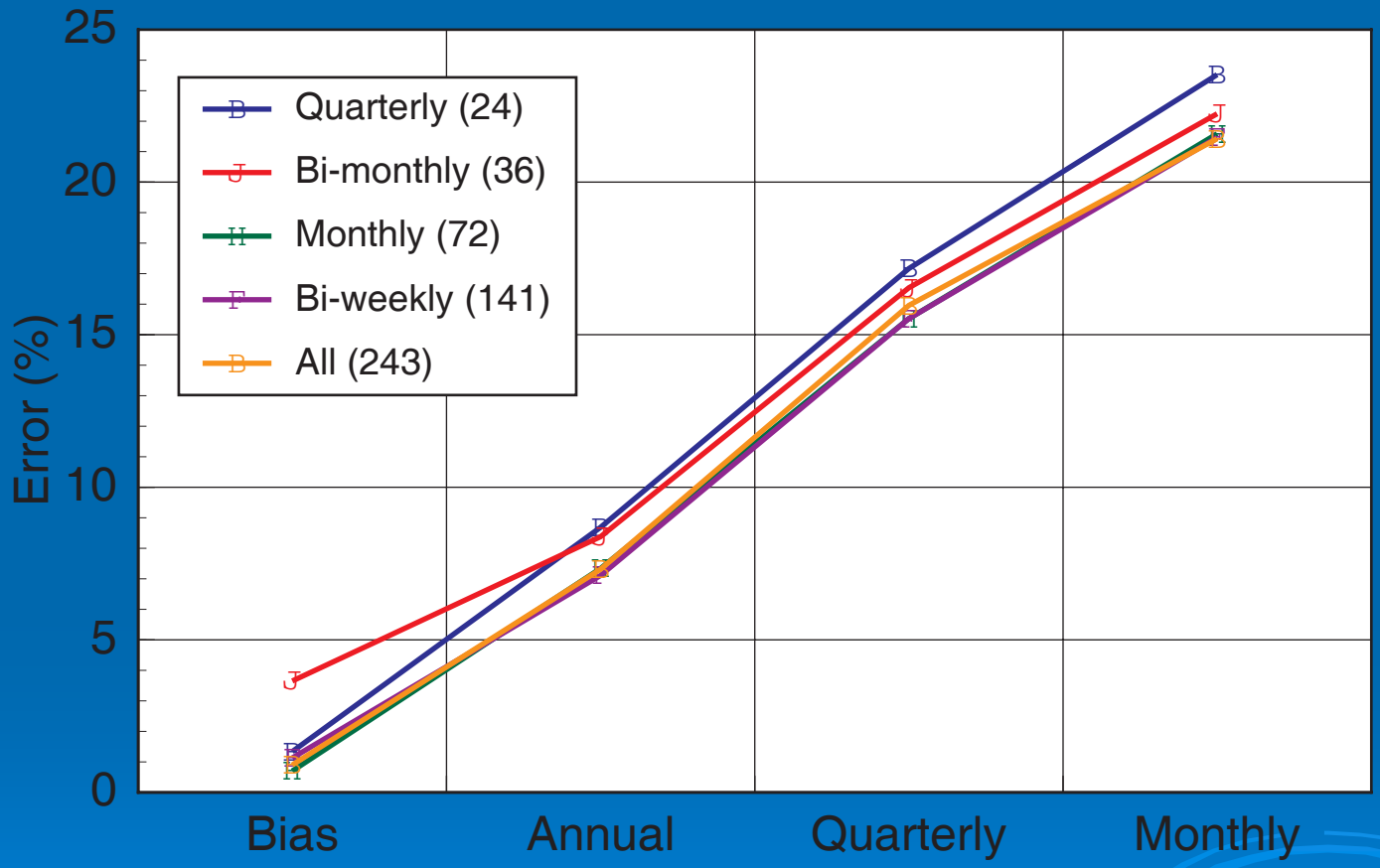
(Regression-model method; at site, t^2; 3-year cal)



- Bias for quarterly, bi-monthly sampling
- No improvement > monthly sampling

Sampling Frequency

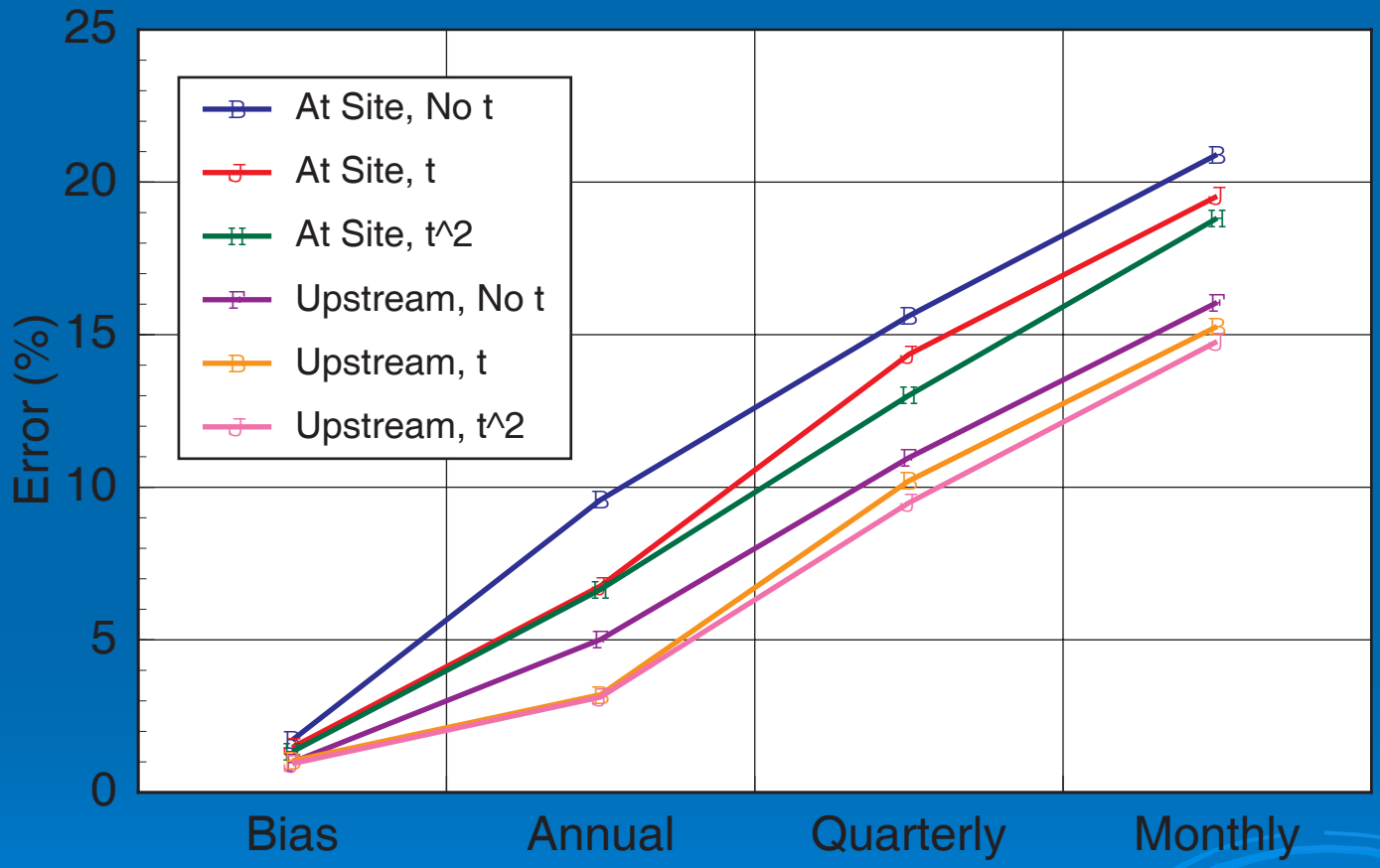
(Regression-model method; at site, t^2; 6-year cal)



- Sampling not important once have enough samples
- 6-year calibration period less precise

Model Structure

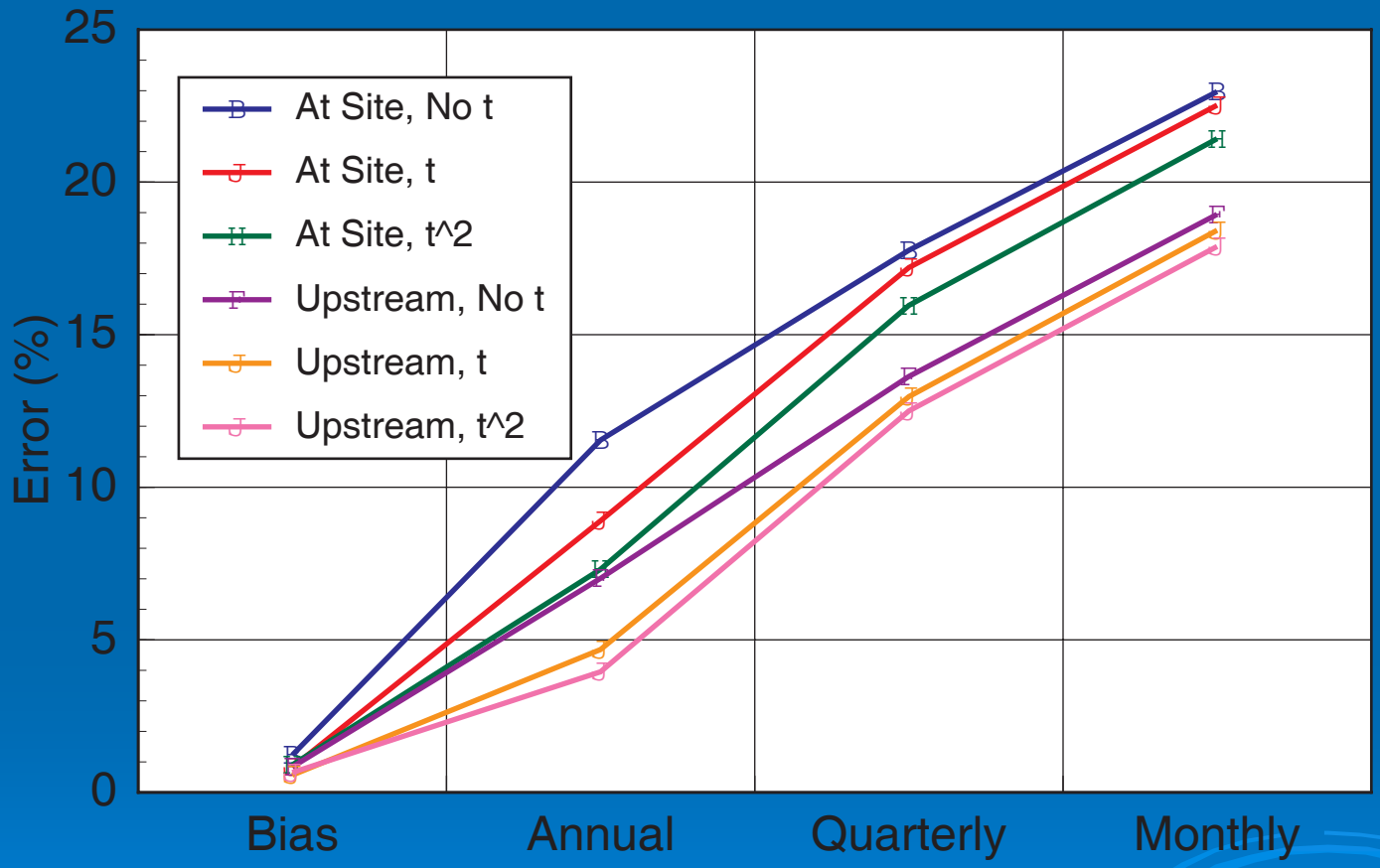
(Regression-model method; all samples; 3-year cal)



- Upstream improves more than time terms
- t^2 term not much improvement over linear t term

Model Structure

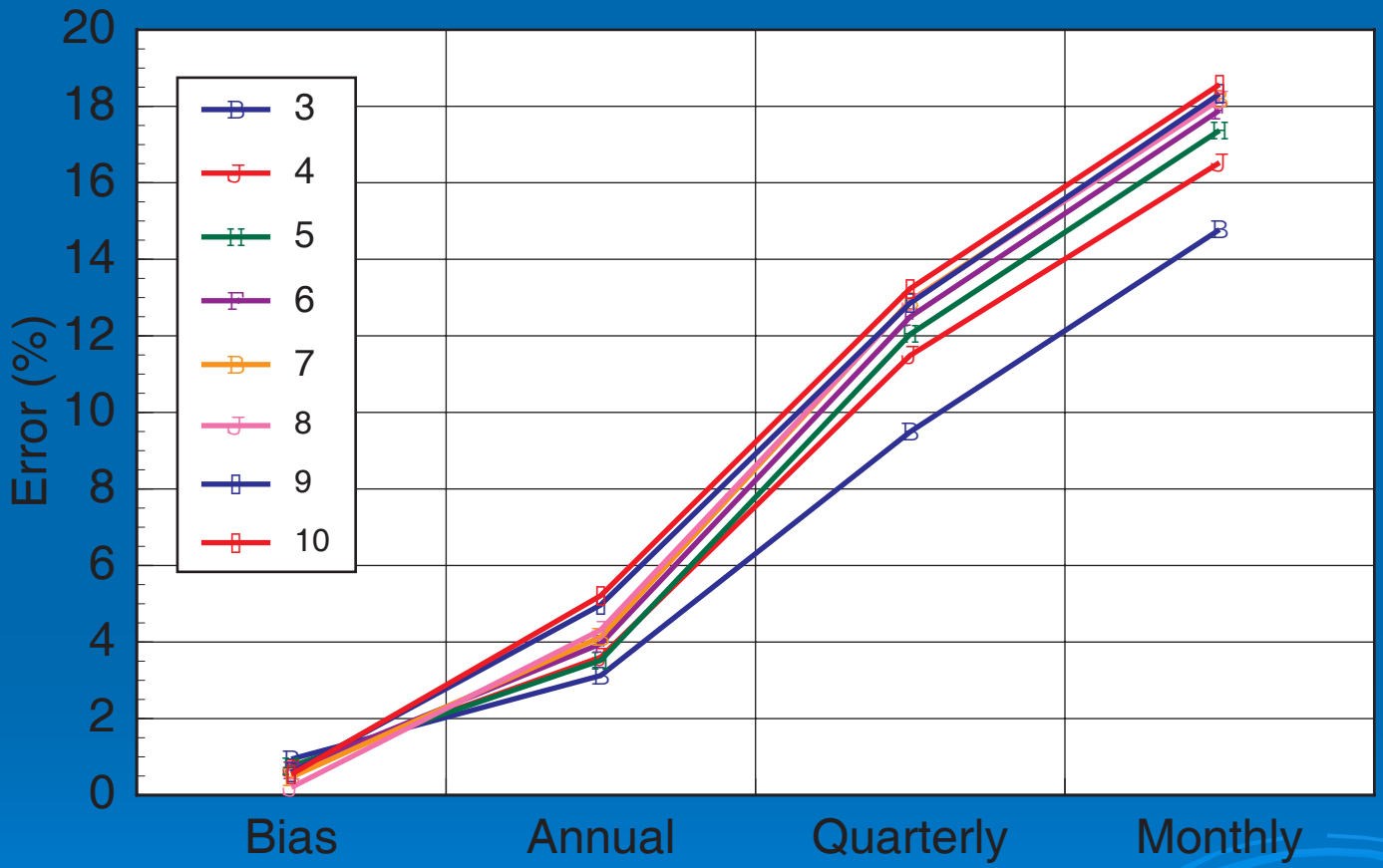
(Regression-model method; all samples; 6-year cal)



- Upstream improves more than time terms
- t^2 term helps more, especially for at site model

Calibration Period

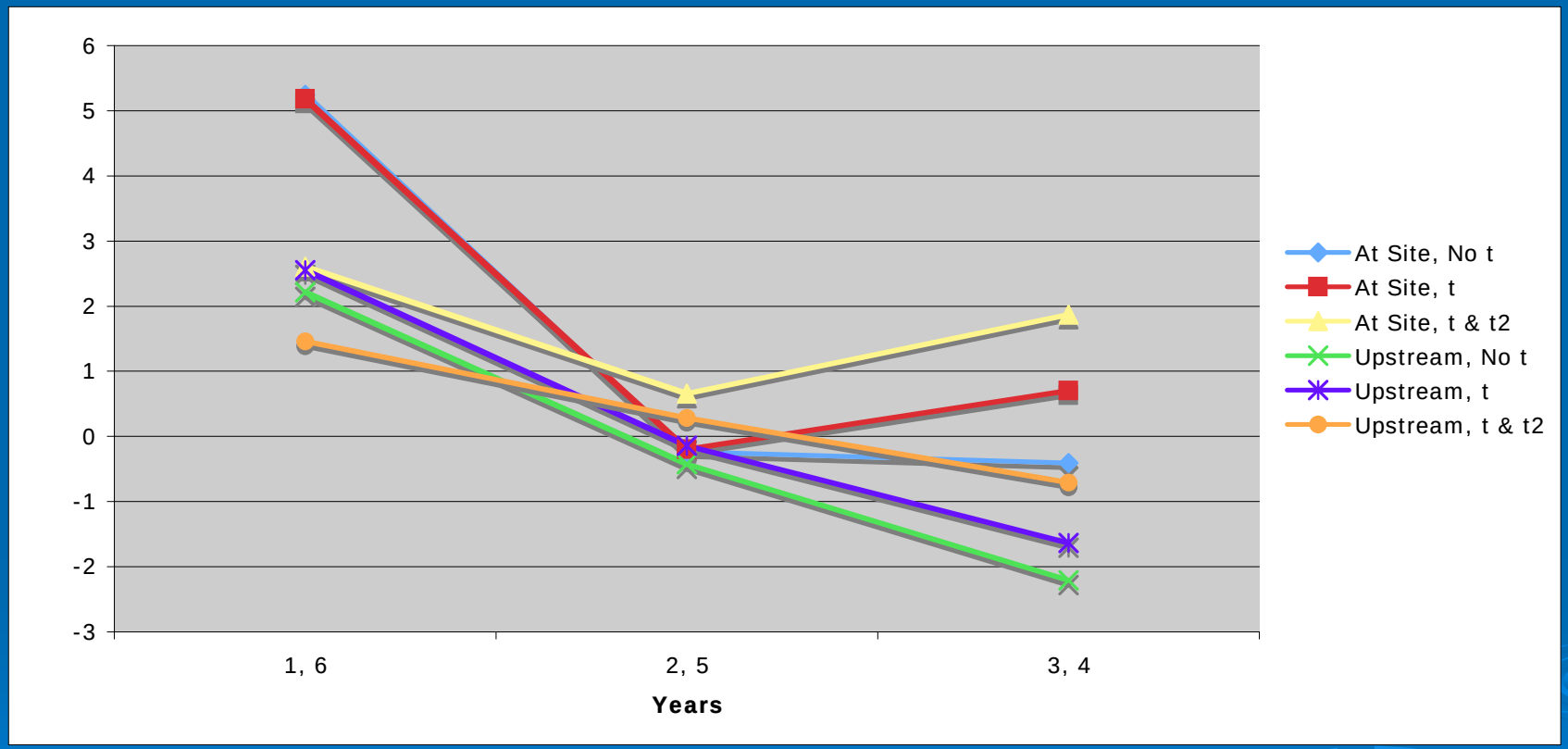
(Regression-model method; Upstream, t^2; all samples)



- 3-year calib. period more bias, better precision
- Longer calib. Period, less precision

End vs. Middle of Cal. Period

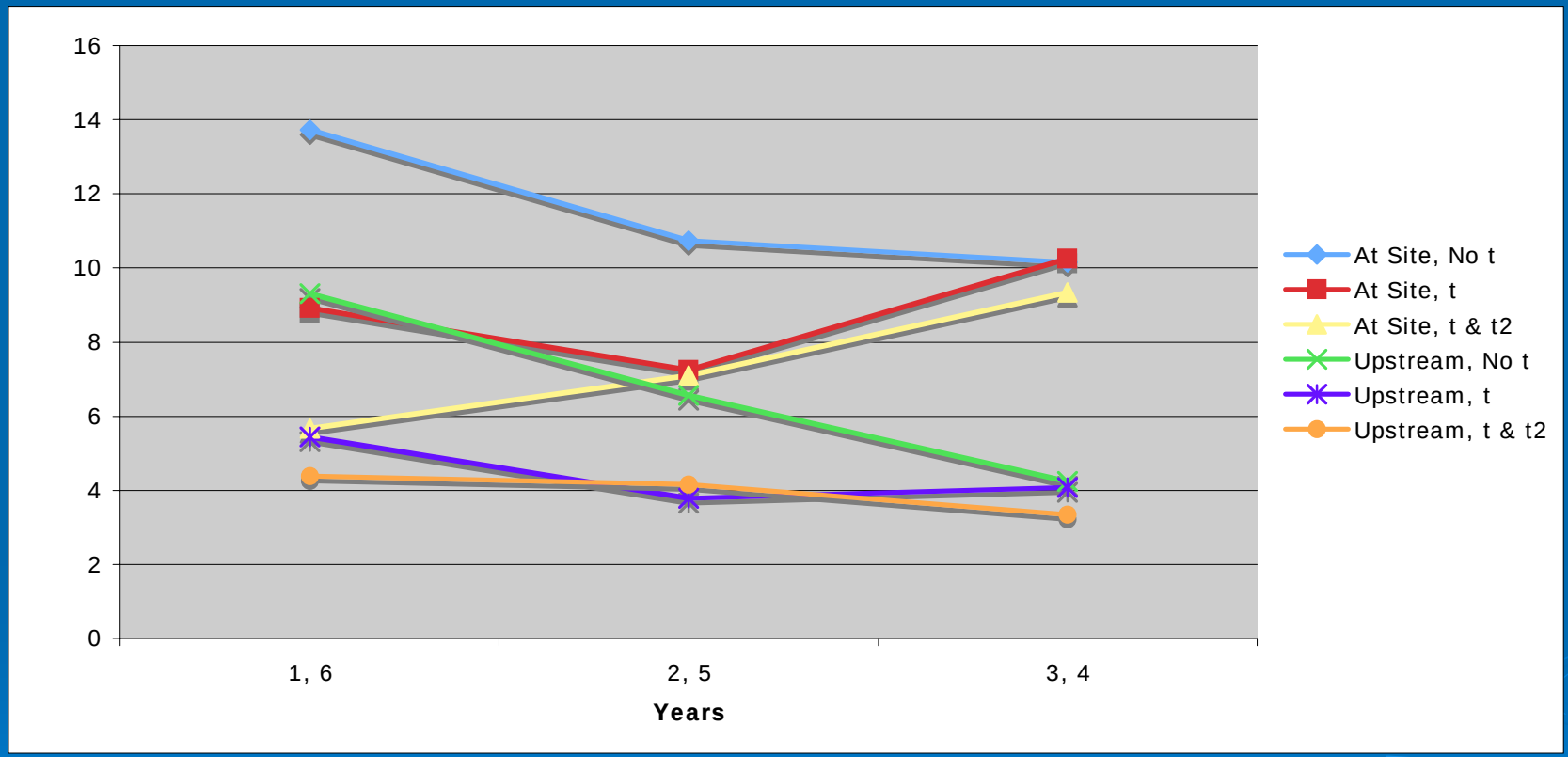
(Regression-model method; 6-year cal; all samples)



- Positive bias for end years
- Not as pronounced for more complex models

End vs. Middle of Cal. Period

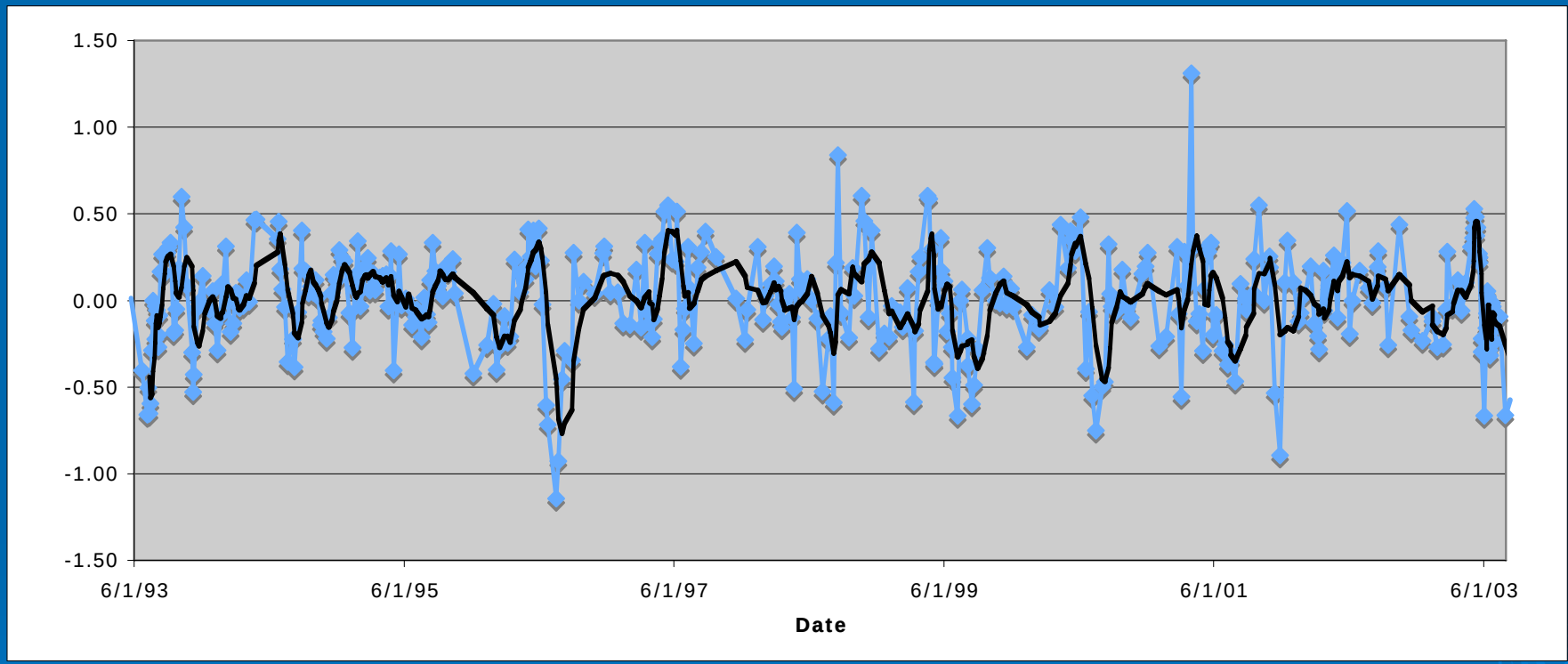
(Regression-model method; 6-year cal; all samples)



- No t models, less precise end years?
- Results from only 5 model runs

Regression Residuals

(Upstream, t^2; 3-year cal; centered loads)



- Short-term patterns in residuals (not random)
- Can't be modeled with long-term trend terms

Summary - Regression Model Method

- Sampling does not improve loads after about 24 - 36 samples in calibration set
- Upstream flows help more than time terms
- Longer calibration periods result in:
  - less overall bias
  - less precision at shorter time intervals >3 years

USGS
science for a changing world

Summary - Regression Model Method (cont.)

- Have not fully assessed accuracy at beginning and end of calibration period versus middle - some possible patterns
- Precision not good at short time intervals - Annually to Monthly
- Regression model does not model short term deviations well

USGS
science for a changing world