Statistical Considerations for Comparison to Background

John W. Kern PhD.
Kern Statistical Services, Inc.
Sauk Rapids MN

Factors Controlling Selection of Statistical Methods

- Objectives
  - Baseline Comparisons
  - Injury quantification
  - Adaptive management triggers
  - Demonstration of restoration success
- Sampling design
  - Random, Systematic, Stratified, Biased
- Metric or variable types
  - Continuous Variables (Sediment concentration, Biomass)
  - Discrete Counts (Stem density, Species richness, Abundance)
  - Binary Outcomes (Presence absence, Percent survival)
  - Categorical Scores (Red green blue)
  - Indices (HIS, Habitat value, Resource selection)
  - Qualitative Scoring Systems
- Potential for confounding factors

Primary Comparison of Interest for HEA (Baseline)

1) Document and estimate the duration and extent of injury, from the time of injury until the resource recovers to baseline, or possibly to a maximum level below baseline condition[1].

[1]*Habitat Equivalency Analysis: An Overview. Damage Assessment and Restoration Program National Oceanic and Atmospheric Administration Department of Commerce. Revised May 23., 2006.*

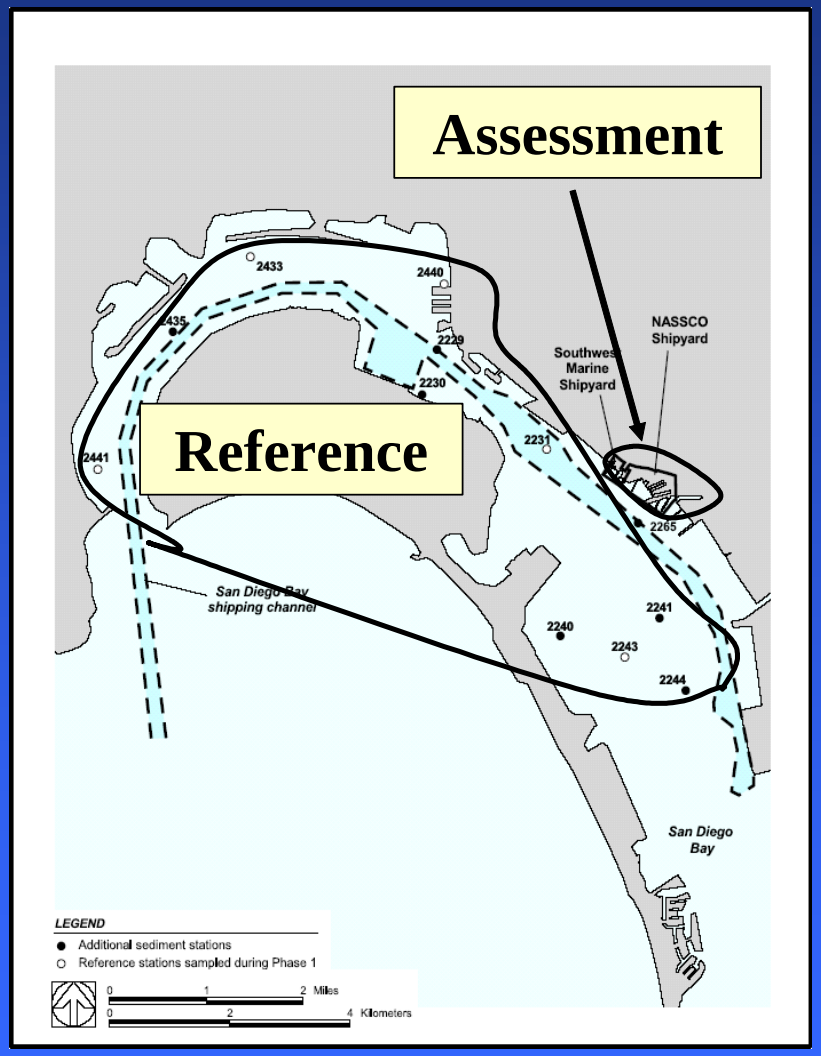
Difficulties With Site vs. Baseline Comparisons

- Baseline data may not be available for the site.
- Ecological conditions vary temporally, so baseline is often not a single condition.
- Reference sites are typically used as a surrogate for baseline, creating the potential for site vs. baseline comparisons to be confounded.

Some Examples:

- Chromium concentration in San Diego Bay sediment.
- Sex ratios in Horseshoe Crabs at Cape Cod Bay and Monomoy National Wildlife Refuge, MA.
- Species richness in salt marsh in Hatches Harbor, MA.
- Water Quality (Surface Coal Mine, UT)
- Fish PCB Concentrations (Kalamazoo River, MI)

Confounding Factors:



San Diego Bay Chromium

Site	Mean Cr (mg / kg)	Mean % Fines
Reference (N=12)	25.2	33.4
Assessment (N=66)	60.3	68.0

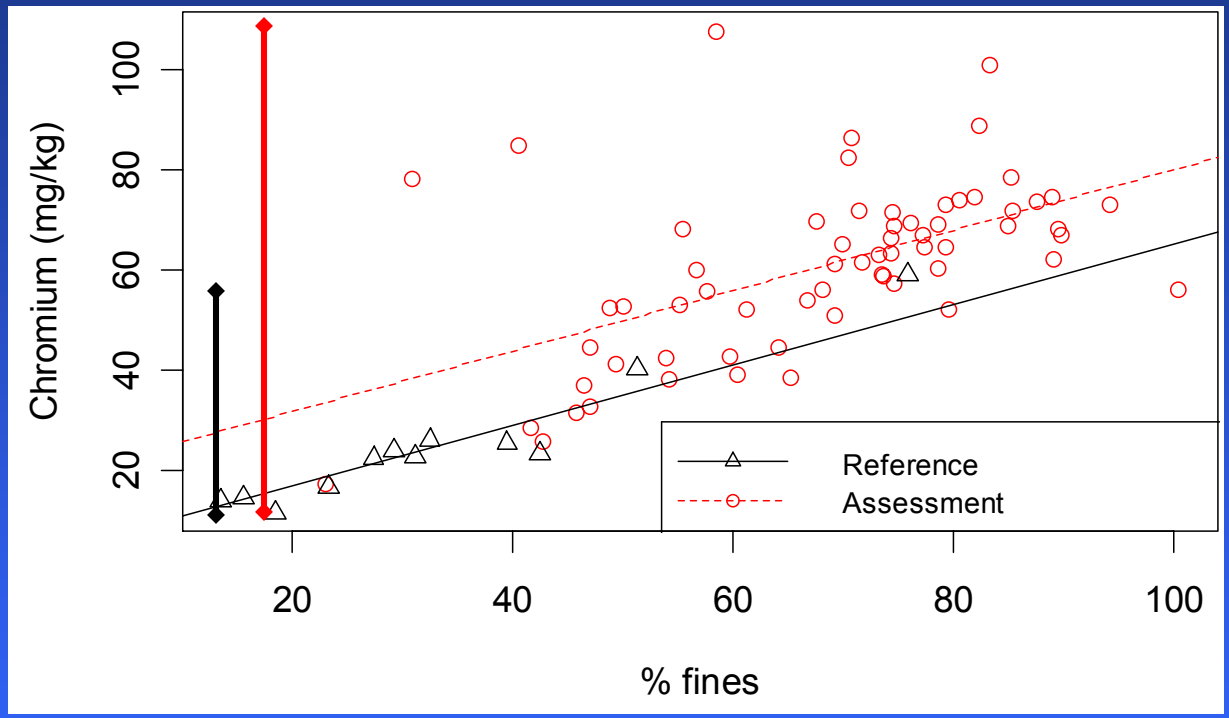
- Percent fines may be confounded with site effects.
- This may limit the strength of the data to support regulatory decisions.
- Interpretation of data is conditioned on the effects of percent fines on Chromium concentration.

Confounding Factors Are Often Handled Qualitatively

- From NASSCO and Southwest Marine Detailed Sediment Investigation: Volume I (Exponent Inc., October 2003):

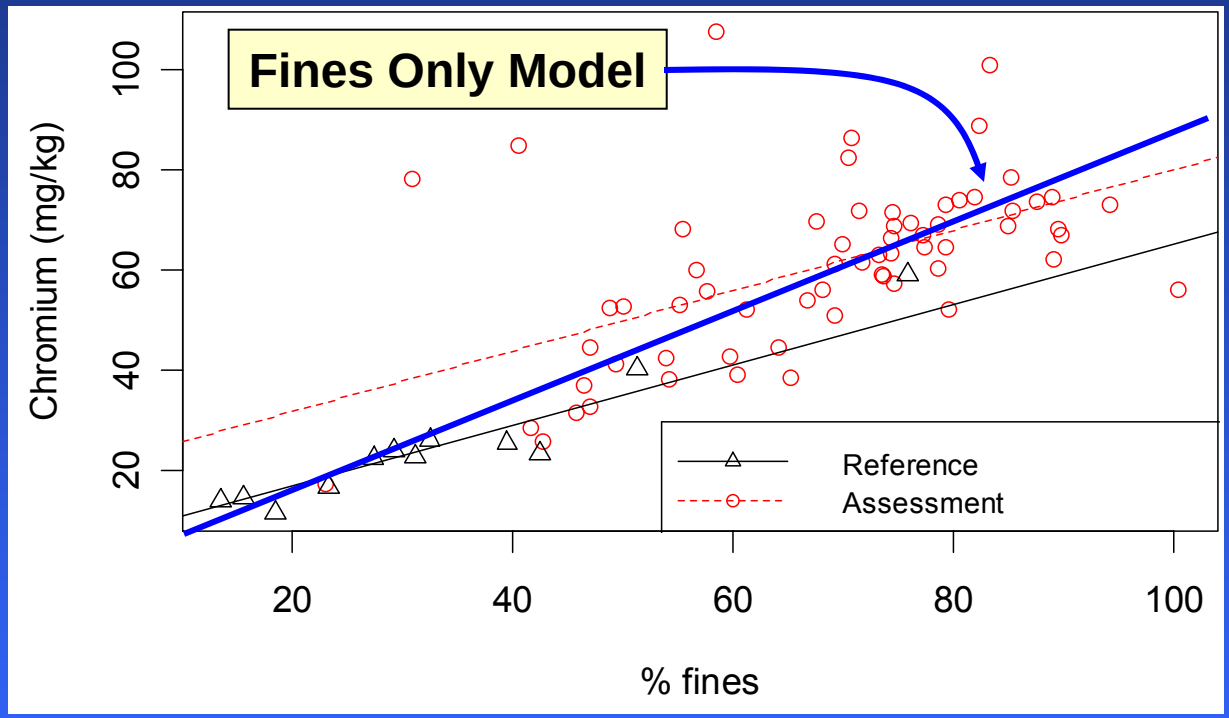
- "Because higher chemical concentrations are ordinarily associated with finer grain sizes and higher TOC content, and because benthic macroinvertebrate community composition also depends on these variables, the physical differences between reference and shipyard stations are expected to result in different chemical and biological conditions regardless of any influences of the shipyards."

Regression Approach To Adjust For Confounding Factors.



$$Cr = \beta_0 + \beta_1 \times Fines + \beta_2 \times Site + \beta_3 \times Fines \times Site$$

Akaike Information Criterion Can Be Used to Select Best Model.



$$Cr = \beta_0 + \beta_1 \times Fines + \beta_2 \times Site + \beta_3 \times Fines \times Site$$

Model Fit and Analysis of Parameters

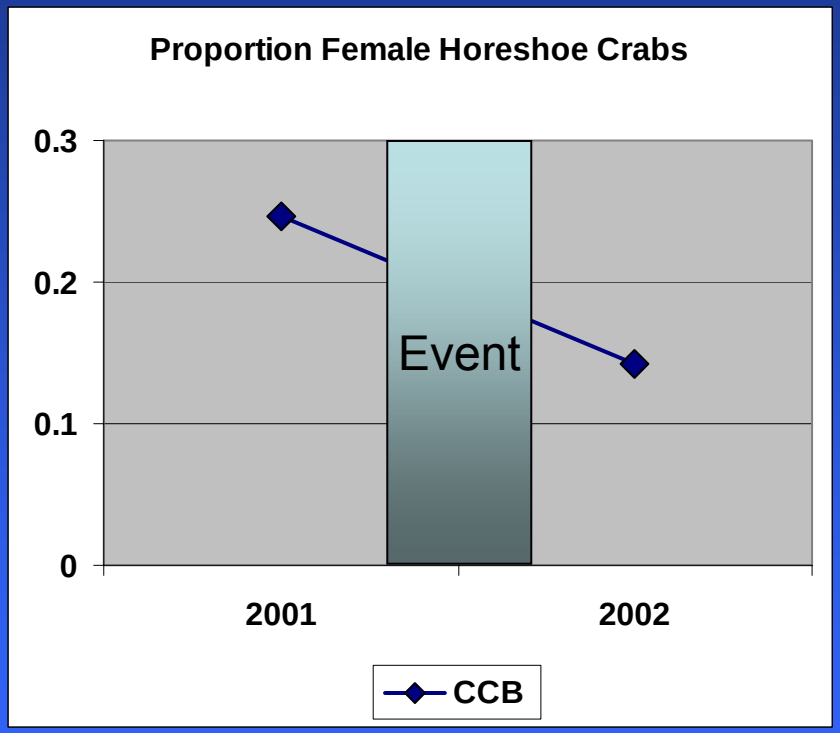
Model	AIC
Chromium ~ % Fines	623.2
Chromium ~ % Fines + Site	**617.9**

	Estimate	Std Error	t value	Pr(>|t|)
Intercept	5.07	5.10	0.99	0.324
% Fines	0.60	0.01	6.21	≤0.001
Site	14.73	5.46	2.70	0.009

Site effect of 14.7 mg/kg. 95% UCL (13.7, 25.7)

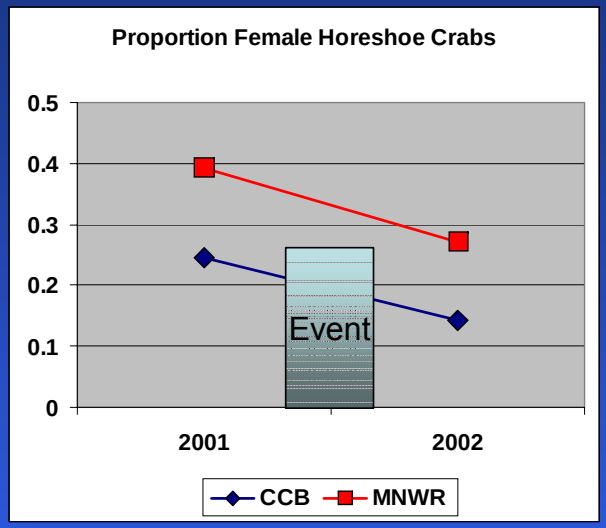
Pre-Post Evaluation of Cape Cod Bay Horseshoe Crabs

- Metric of Interest: Proportion of Females (Binary)
- There is an apparent decline in 2002 after an adverse event.
- Is the observed decline due to an event or due to natural temporal variation.
- It could be argued that the proportion of females in 2001 represents baseline condition.



Reference Data From the Monomoy National Wildlife Refuge Changes the Picture (BACI)

- Temporal trends appear to be similar at Monomoy National Wildlife Refuge and Cape Cod Bay.
- Event effects are tested by looking at interaction between time and location
- Test for difference of differences.



Before After Control Impact (BACI) With Logistic Regression

$$logit\left(p\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2$$

$$x_1 = \begin{cases} 1 \ for \ Cape \ Cod \\ 0 \ for \ Monomoy \ \ NWR \end{cases}; \quad x_2 = \begin{cases} 1 \ for \ 2001 \\ 0 \ for \ 2000 \end{cases}$$

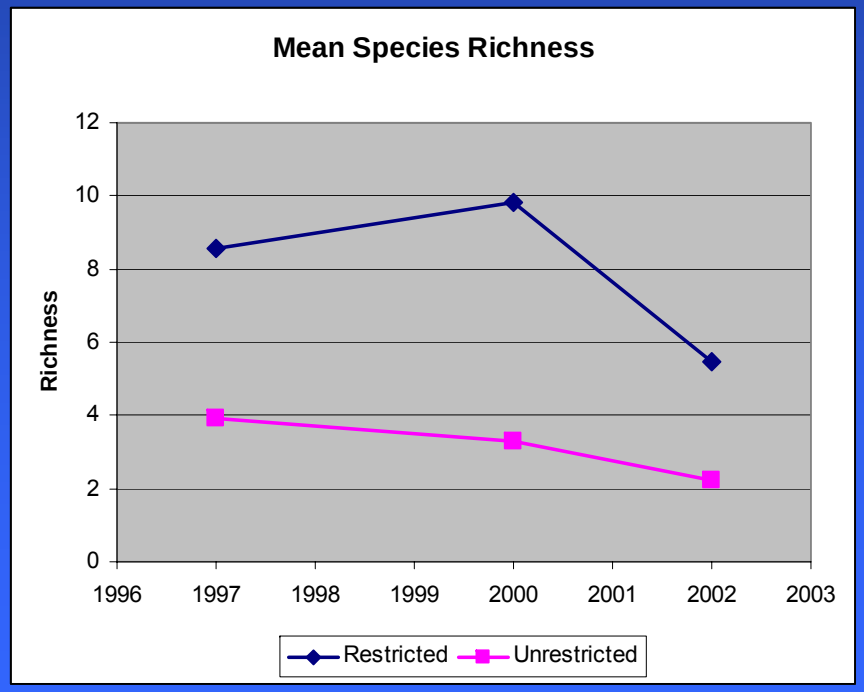
Species Richness of Salt Marsh Vegetation
Another BACI Design

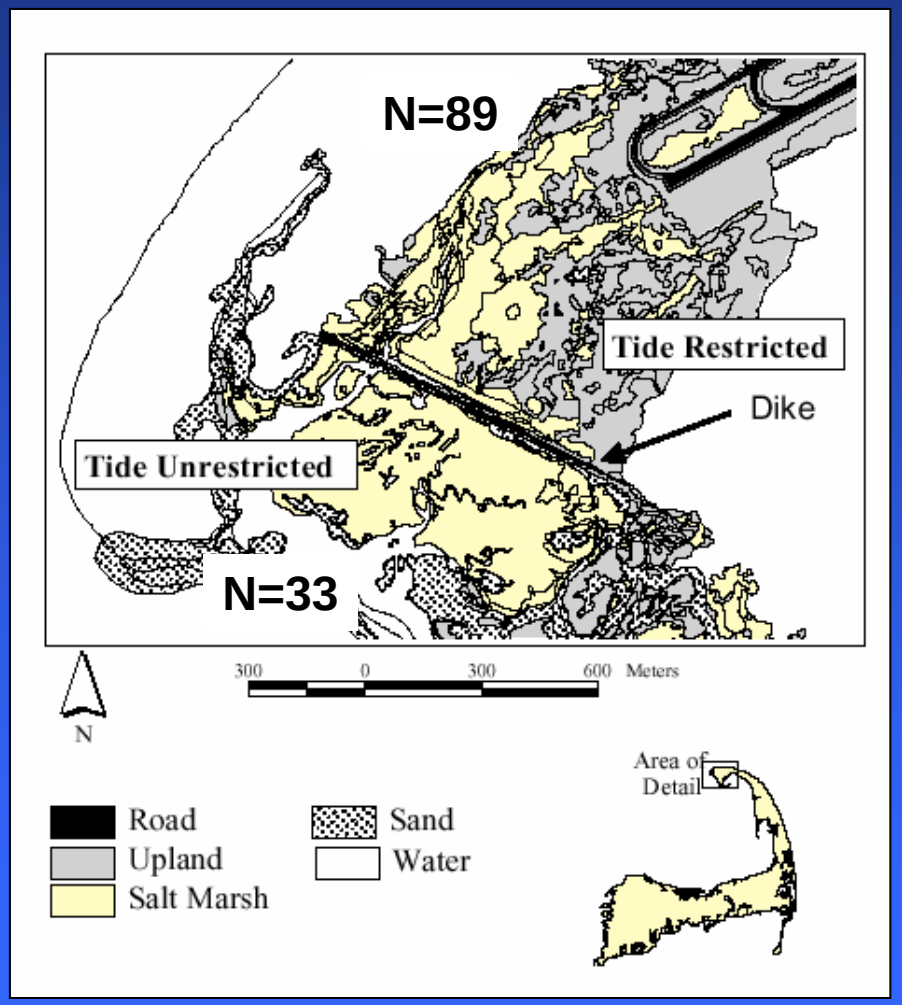
Stewart-Oaten, A., W.W. Murdoch, and K.R. Parker. 1986. Environmental impact assessment: “pseudoreplication” in time? *Ecology* 67: 929-940.

Underwood, A.J. 1992. Beyond BACI: the detection of environmental impacts on populations in the real, but variable, world. *J. of Experimental Marine Biology and Ecology* 161: 145-178.

Species Richness
Hatches Harbor, Provincetown, MA

Mean Species Richness			
Year	Restricted	Unrestricted	Ratios
1997	8.6	3.9	2.2
2000	9.8	3.3	3.0
2002	5.5	2.2	2.4





Poisson Regression Model for Counts
BACI Design

- Count data are typically not normally distributed when counts are "*small*".
- The Poisson distribution is a reasonable alternative for count data.
- Log of the mean response is assumed to be a linear function of the predictors.
- Parameter estimation based on maximum likelihood.
- Statistical inferences are based on the likelihood and deviance as with logistic regression.
- Over dispersion is accommodated

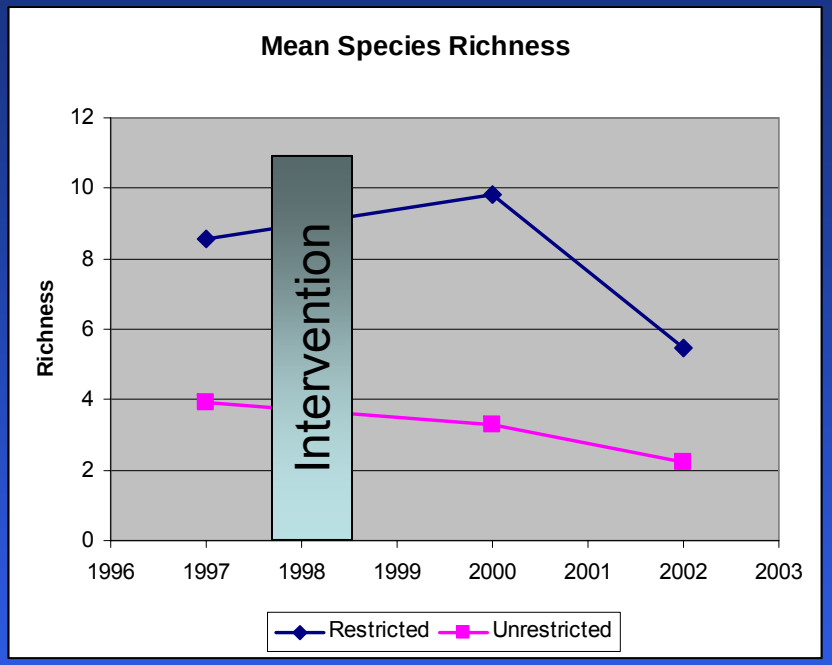
$$E(y) = \mu = e^{\beta_0 + \beta_1 year_1 + \beta_2 site + \beta_3 Year \times Site}$$

$$log(\mu) = \beta_0 + \beta_1 Year_1 + \beta_2 Site + \beta_3 Year \times Site$$

$$prob(Y = y) = \frac{\mu^y e^{-\mu}}{y!}; \quad y = 0, 1, 2, \ldots$$

BACI Analysis of Species Richness in Salt Marsh

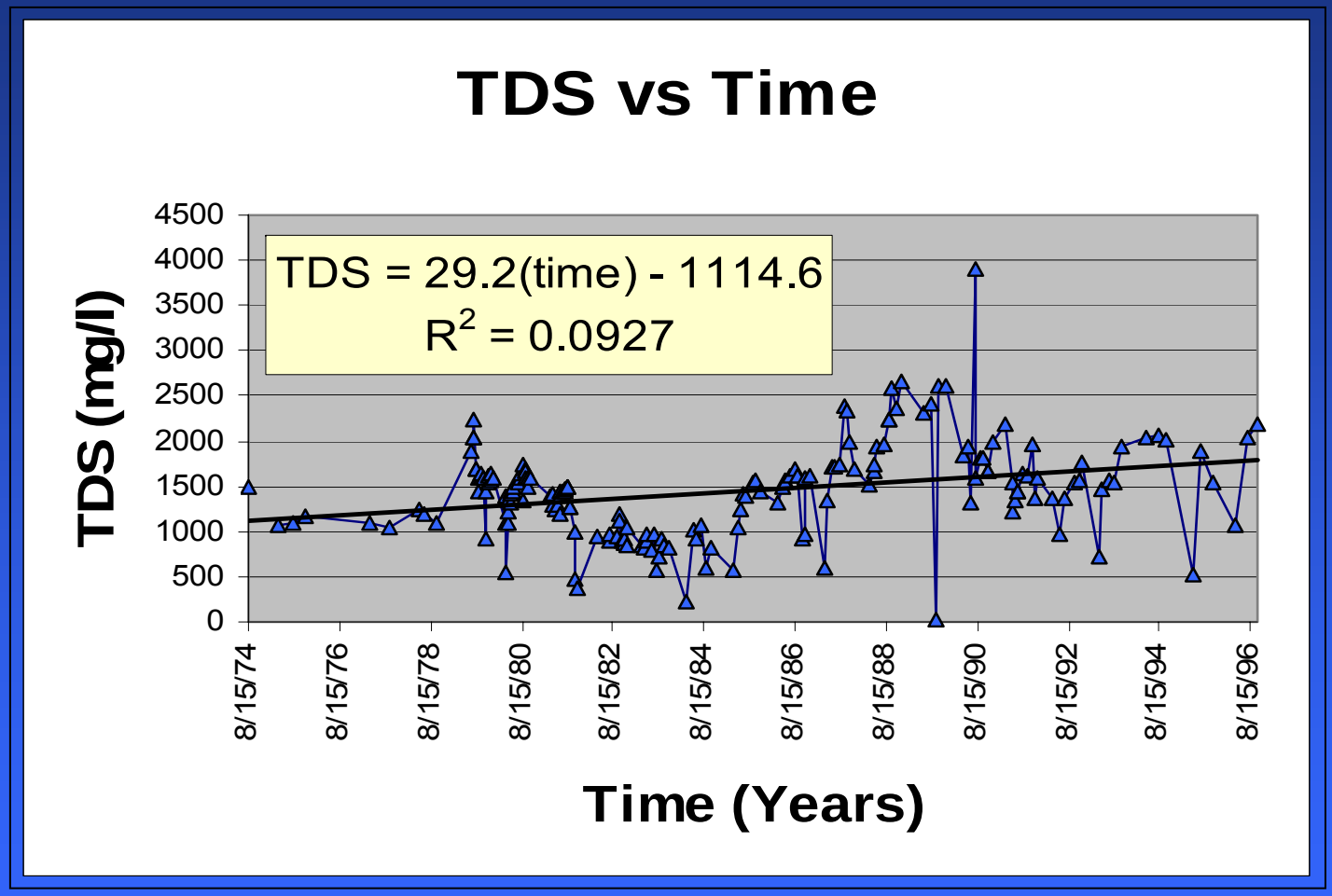
- Richness was higher at the restricted site (p=0.0001)
- Temporal trajectories were similar at restricted and unrestricted sites.
- Verdict still out.



Likelihood Ratio Tests			
Source	**df**	**F**	**Significance**
Year	2	14.4	0.0001
Study Area	1	145	0.0001
Year x Study Area	2	1.4	0.2579

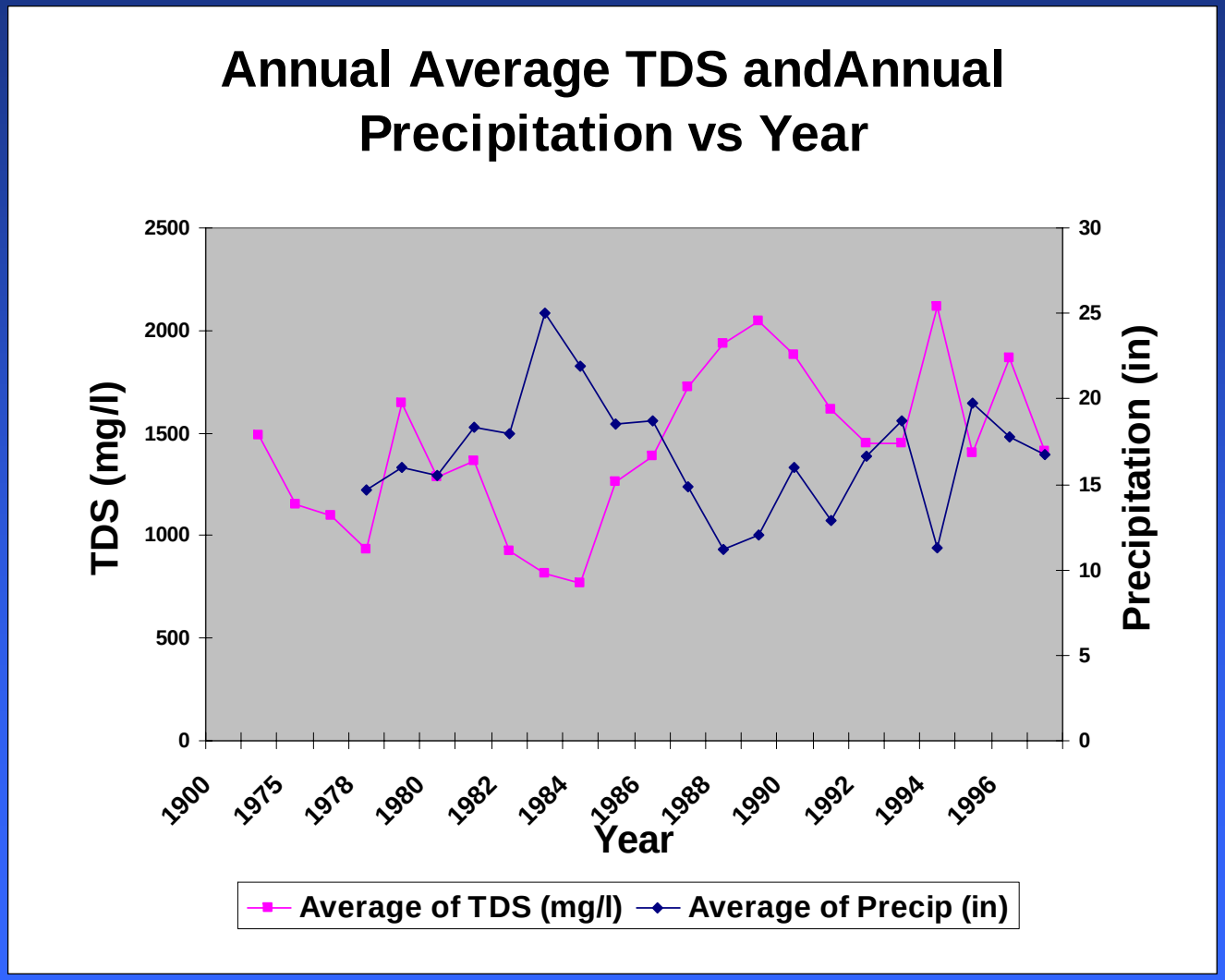
Temporal Variation and Confounding Factors

Water Quality Trends

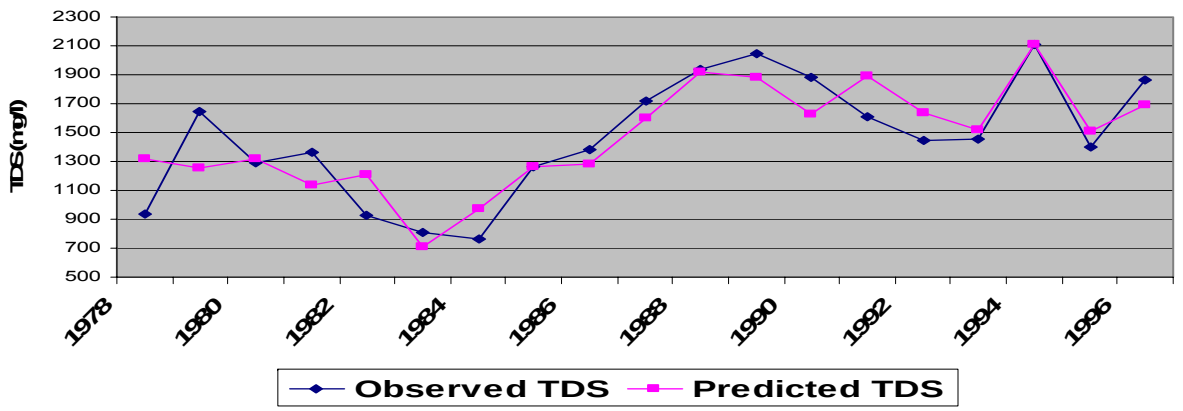


$$TDS = B_0 + B_1 (Time) + Error$$

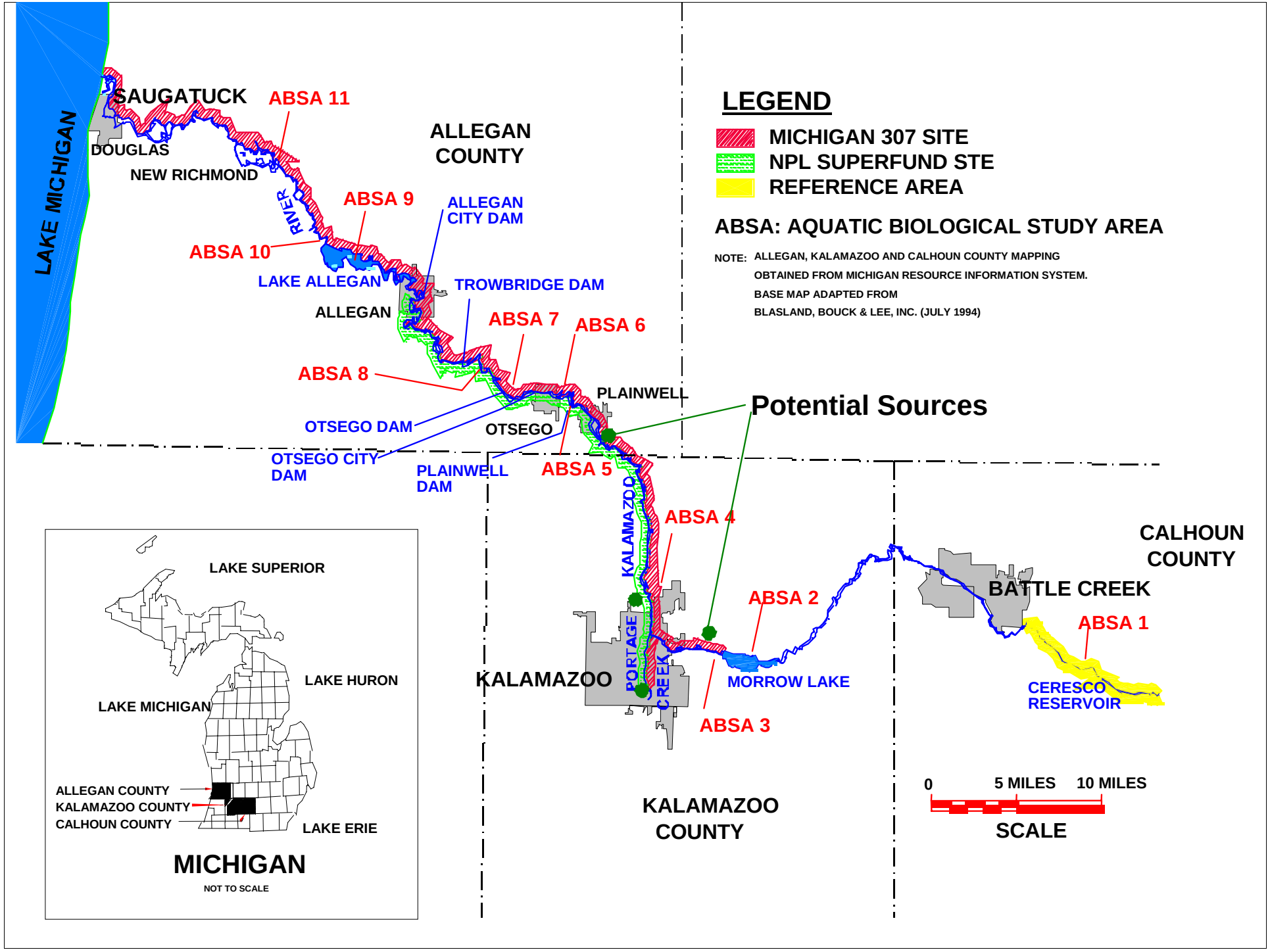
Annual precipitation data are also available.



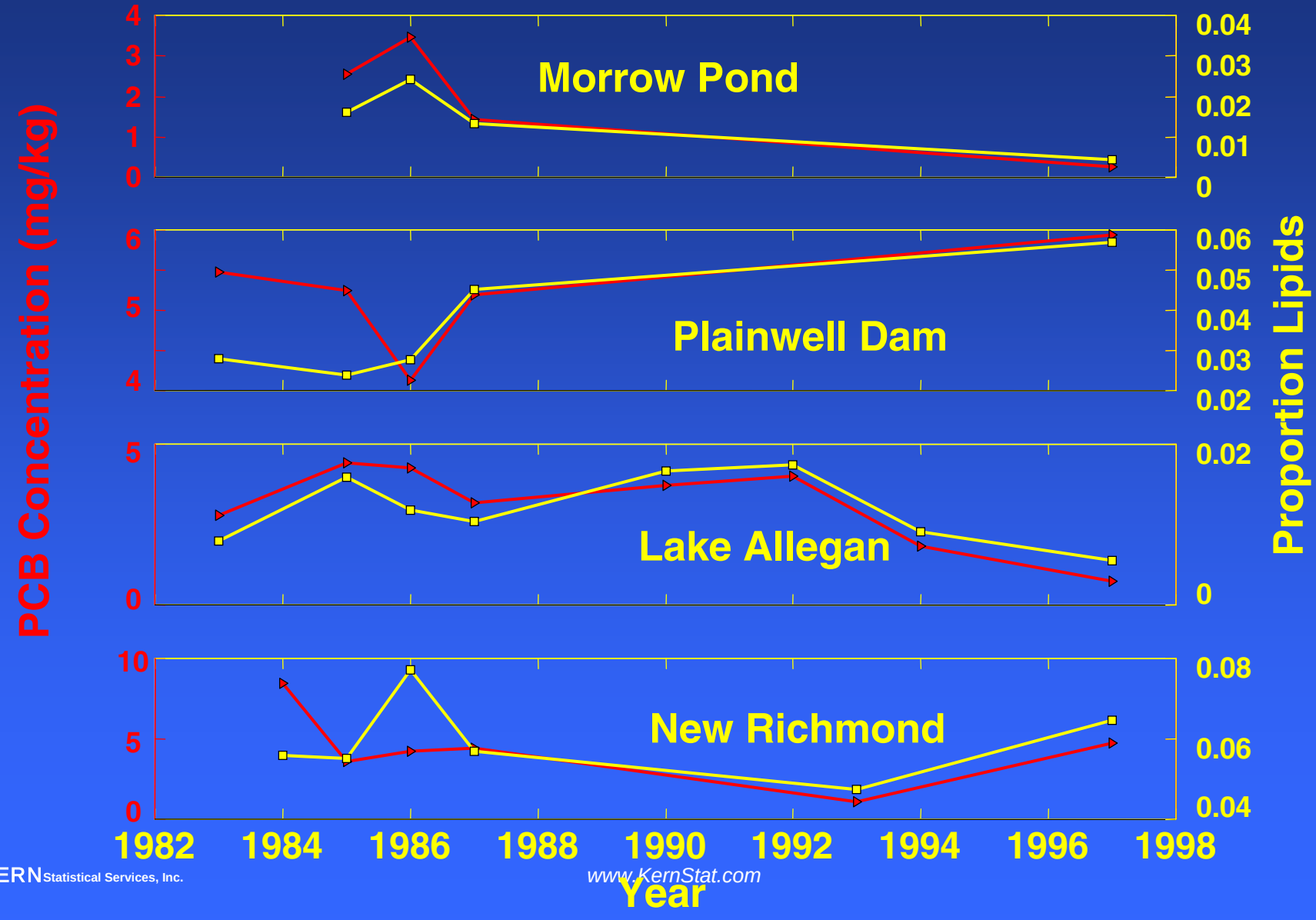
$TDS = B_0 + B_1 (Precip) + B_2 (Time) + Error$



Regression Statistics				
Multiple R	0.86			
R Square	0.75			
Adj. R Square	0.72			
	Coefficients	Standard Error	t Stat	P-value
Intercept	-64533.5	18505.7	-3.5	0.003
Time (Years)	33.8	9.3	3.6	0.002
Precip (in)	-75.1	14.7	-5.1	0.000


LAKE MICHIGAN
SAUGATUCK
DOUGLAS
NEW RICHMOND
ABSA 11
ALLEGAN COUNTY
ABSA 9
ALLEGAN CITY DAM
ABSA 10
KALAMAZOO RIVER
LAKE ALLEGAN
TROWBRIDGE DAM
ALLEGAN
ABSA 7
ABSA 6
ABSA 8
PLAINWELL
OTSEGO DAM
OTSEGO
OTSEGO CITY DAM
PLAINWELL DAM
ABSA 5
ABSA 4
KALAMAZOO RIVER
PORTAGE CREEK
KALAMAZOO
ABSA 2
ABSA 3
MORROW LAKE
CALHOUN COUNTY
BATTLE CREEK
ABSA 1
CERESCO RESERVOIR
KALAMAZOO COUNTY
0
5 MILES
10 MILES
SCALE
LEGEND
MICHIGAN 307 SITE
NPL SUPERFUND STE
REFERENCE AREA
ABSA: AQUATIC BIOLOGICAL STUDY AREA
NOTE: ALLEGAN, KALAMAZOO AND CALHOUN COUNTY MAPPING OBTAINED FROM MICHIGAN RESOURCE INFORMATION SYSTEM.
BASE MAP ADAPTED FROM BLASLAND, BOUCK & LEE, INC. (JULY 1994)
Potential Sources
LAKE SUPERIOR
LAKE HURON
LAKE MICHIGAN
LAKE ERIE
ALLEGAN COUNTY
KALAMAZOO COUNTY
CALHOUN COUNTY
MICHIGAN
NOT TO SCALE

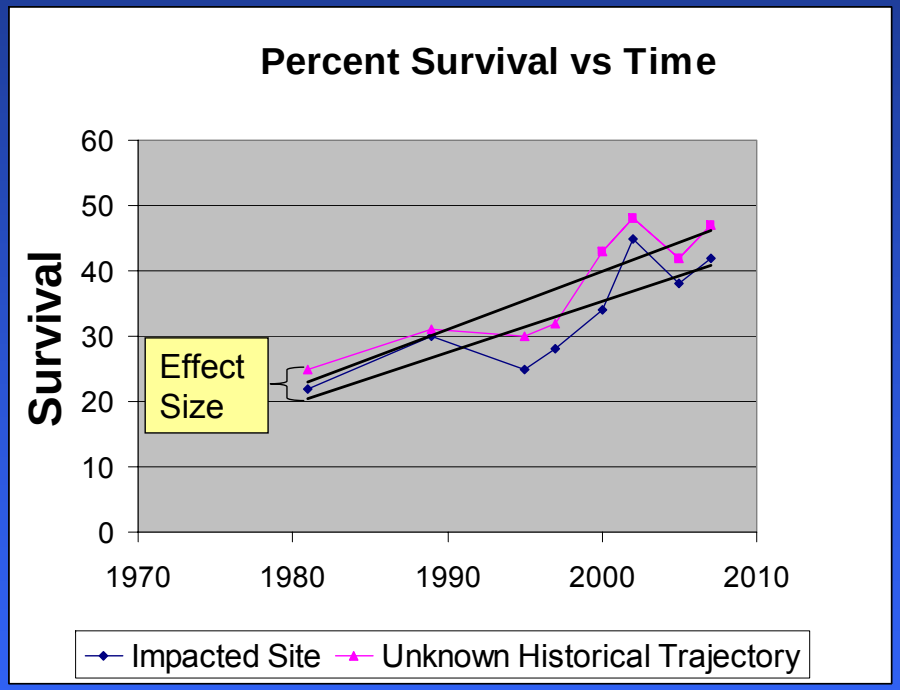
Mean PCB Concentration and Percent Lipids in Carp Fillets (Skin Off)

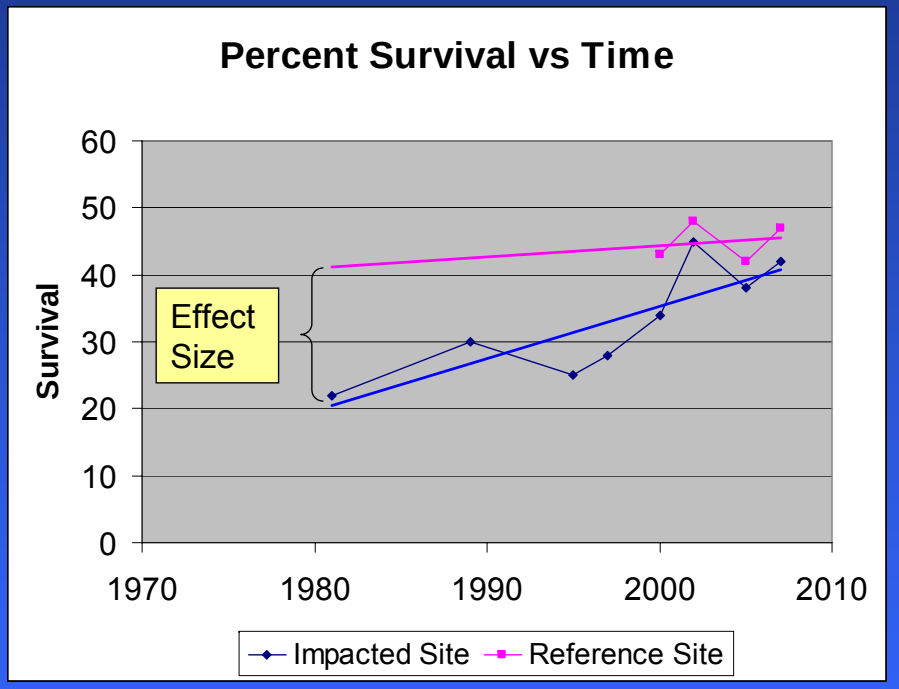


KERN Statistical Services, Inc.

www.KernStat.com

Which is Right?





Data Sources and Acknowledgements

- Don McDonald
  - NOAA
- Charles Roman National Park Service
  - National Park Service North Atlantic Coast Cooperative Ecosystem Studies Unit at the University of Rhode Island
- Mary Jane James-Pirri
  - Marine Research Associate Graduate School of Oceanography University of Rhode Island Narragansett, RI 02882
- Joe Galetovic
  - Office of Surface Mines and Reclamation, Denver CO
- Michigan Department of Environmental Quality