Long-Term Watershed Monitoring
Statistical Models & Examples

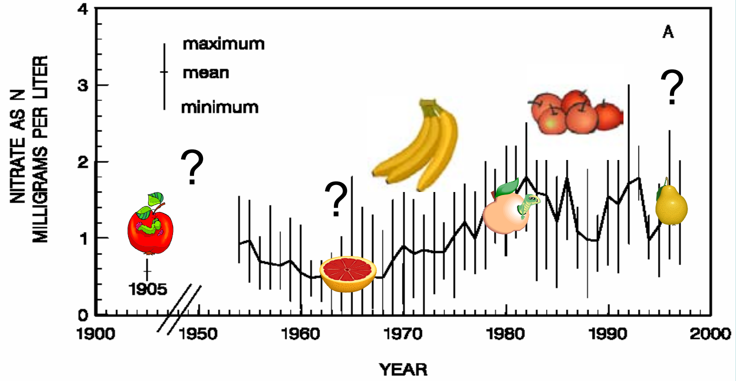
Watershed Monitoring & Reporting Session
Gulf of Mexico Hypoxia Workshop
St Louis
October 16-18, 2002

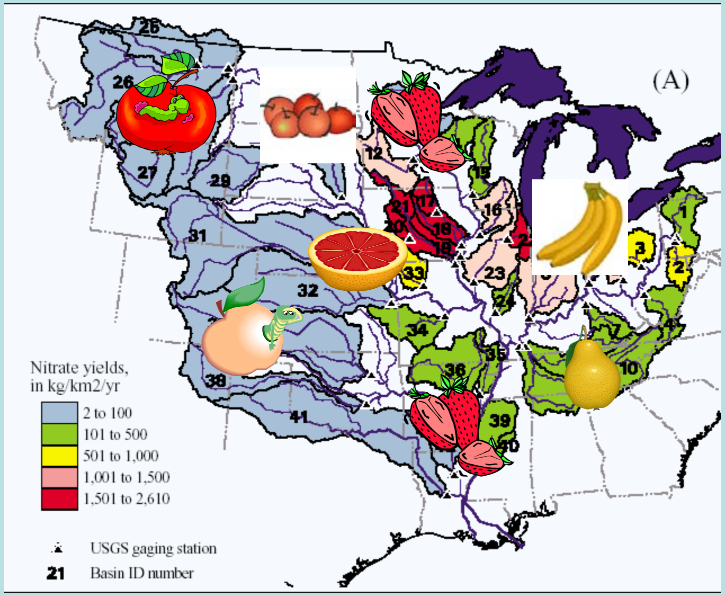
William W. Walker
Environmental Engineer
wwwalker.net

Topics

- Gulf Effort: Challenges & Suggestions
- Statistical Models
  - Allocating Sampling Effort
  - Computing Loads
  - Estimating Power for Detecting Trends
  - Measuring Long-Term Progress
- Examples
  - Onondaga Lake, New York
  - Everglades

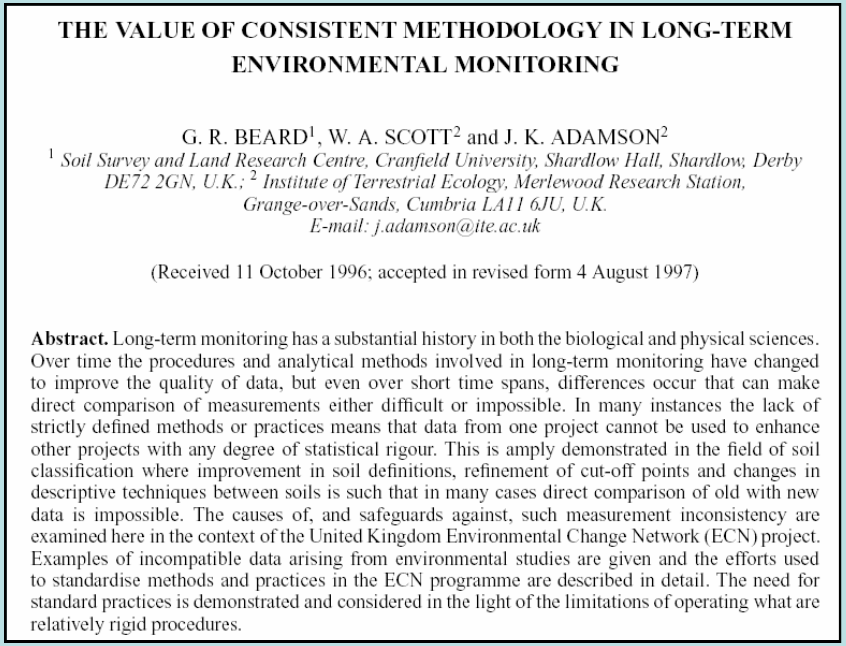
Developing a Consistent Long-Term Dataset





- Challenge to Gulf Effort
- Multiple Agencies
- Multiple Sampling Objectives
- Compatibility
  - Over Time
  - Across Sub-Basins
- Potential Inconsistencies
  - Different Baseline Periods
  - Sampling Sites
  - Sampling Methods
  - Sampling Frequencies
  - Analytical Methods
  - Load Computation Methods

Evolution of Long-Term Monitoring Programs



THE VALUE OF CONSISTENT METHODOLOGY IN LONG-TERM ENVIRONMENTAL MONITORING

G. R. BEARD[1], W. A. SCOTT[2] and J. K. ADAMSON[2]
[1] *Soil Survey and Land Research Centre, Cranfield University, Shardlow Hall, Shardlow, Derby DE72 2GN, U.K.;* [2] *Institute of Terrestrial Ecology, Merlewood Research Station, Grange-over-Sands, Cumbria LA11 6JU, U.K.*
E-mail: j.adamson@ite.ac.uk

(Received 11 October 1996; accepted in revised form 4 August 1997)

Abstract. Long-term monitoring has a substantial history in both the biological and physical sciences. Over time the procedures and analytical methods involved in long-term monitoring have changed to improve the quality of data, but even over short time spans, differences occur that can make direct comparison of measurements either difficult or impossible. In many instances the lack of strictly defined methods or practices means that data from one project cannot be used to enhance other projects with any degree of statistical rigour. This is amply demonstrated in the field of soil classification where improvement in soil definitions, refinement of cut-off points and changes in descriptive techniques between soils is such that in many cases direct comparison of old with new data is impossible. The causes of, and safeguards against, such measurement inconsistency are examined here in the context of the United Kingdom Environmental Change Network (ECN) project. Examples of incompatible data arising from environmental studies are given and the efforts used to standardise methods and practices in the ECN programme are described in detail. The need for standard practices is demonstrated and considered in the light of the limitations of operating what are relatively rigid procedures.

Environmental Monitoring and Assessment **54:** 239–258, 1999.

Competing Objectives

“Improve” Methods
- Increase Precision
- Reduce Cost
- Support Other Programs

Vs.

Maintain Consistency
- vs. Historical Data
- vs. Station Network

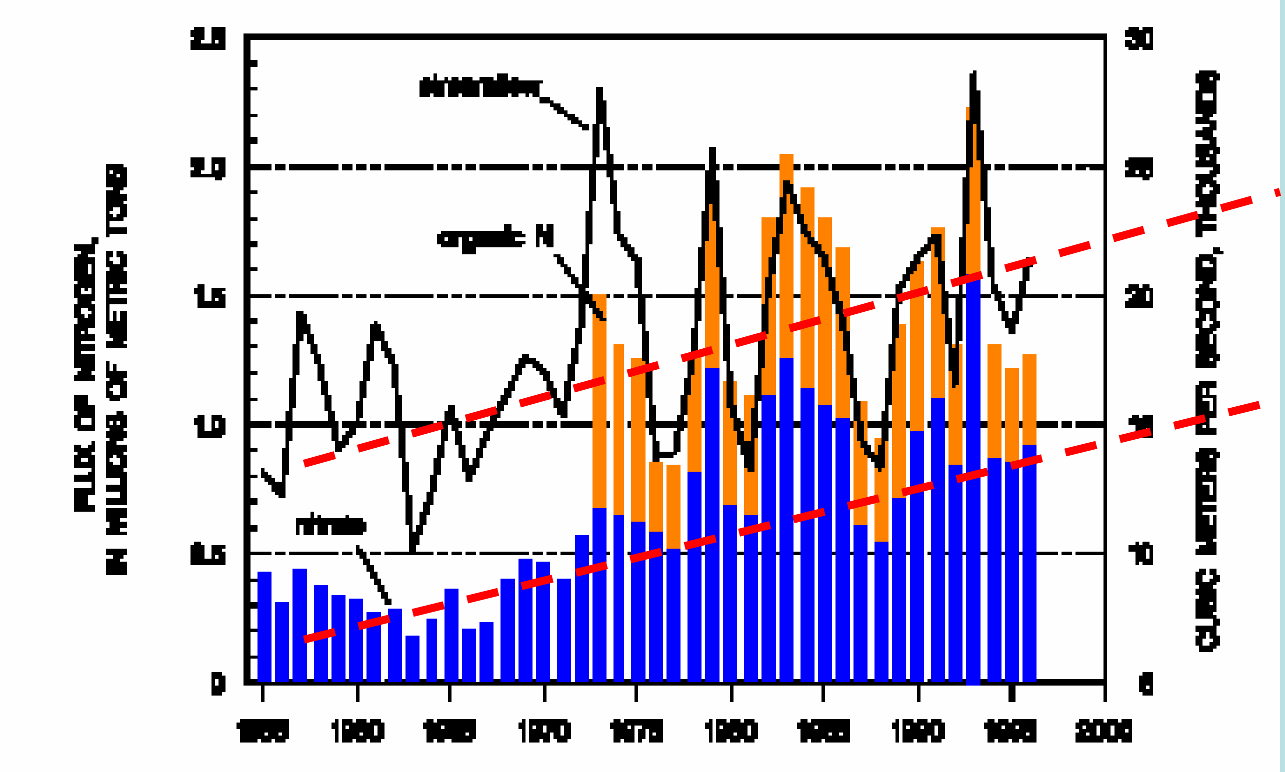
Suggestions

- Define Monitoring Objectives
- Develop & Document Consistent Sampling Protocols
- Establish Precision Targets for Yearly Load Estimates
- Use Historical Data to Design/Refine Programs
  - Identify Stations with Baseline Data
  - Assess Variability
  - Estimate Required Sampling Frequencies
- Central Database, Data Reduction, & Reporting
- Consistency Checks

Potential Consistency Checks

- Water Balances (Basin = Σ Sub Basins)
- Mass Balances (Chloride, Nitrate)
- Laboratory Round Robins
- Paired Sampling
- Sensitivity to Agency / Data Source
- Sensitivity to Load Calculation Method
- Trends at Mouth vs. Sub-Basins
- Watershed Models / Cause-Effect

Interpretation Difficulties – Collinearity

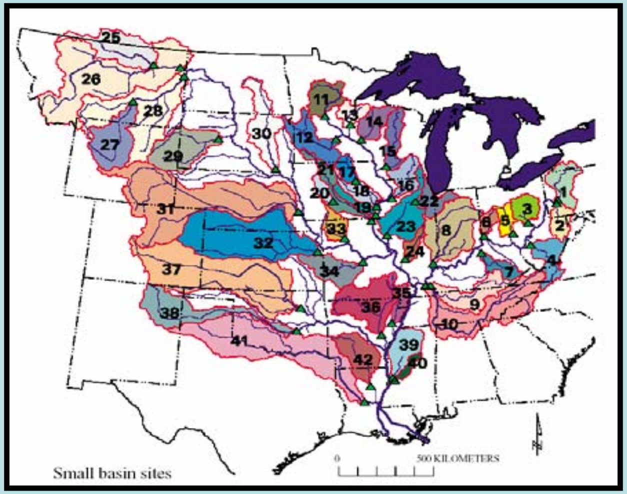


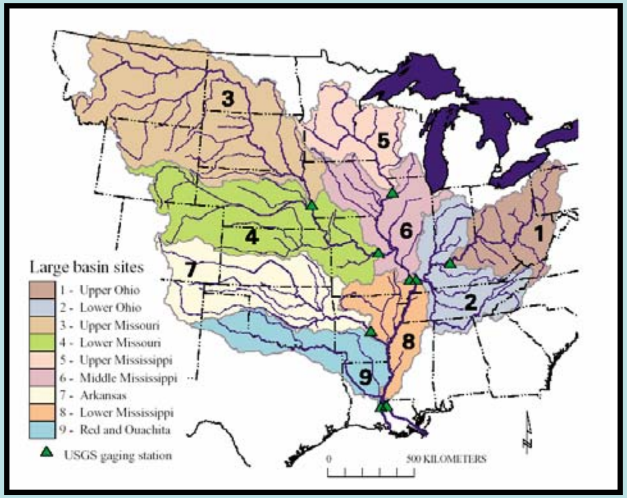
Trend in Flow
Short-Term?
Climate Change?

Trend in Load
Climatologic?
Anthropogenic?
Reversible?

MRB Nitrate Flux, Goolsby et al, 1999

Allocating Sampling Effort Across Basins





Design Objectives

- Basin Mass Balances
- Support Modeling
- Identify Sources
- Measure Progress

Design Variables

- Agency
- Methods
- Location/ Stream Order
- Frequency

Optimal Allocation of Sampling Effort Across Sub-Basins to Estimate Total Load

T = Total Load = $\Sigma \, L_i$

$SE^2(T)$ = Uncertainty (T) = $\Sigma \, S_i^2 / N_i$

N = Fixed Total Samples = $\Sigma \, N_i$

S_i = Standard Deviation of Load for Sub-Basin i

Objective: Find N_i to Minimize $SE(T)$ for Fixed N

Solution: $N_i / N = S_i / T$

Optimal Sampling Freq. $\propto$ Std Deviation of Load

- Load Magnitude
- Variability (Coef. of Variation)

Sampling Program Design for Measuring Loads at a Given Site

Factors

- Load Magnitude
- Concentration & Load Variability
- Concentration/Flow Dependence
- Desired Time Step
- Species (Nutrient, Dissolved vs. Particulate)
- Budget !

Options

- Periodic Grab Sampling
- Additional Grab Sampling at High Flow
    - Seasonal
    - Storm-Event
- Composite Sampling (Spatial, Temporal)

Computing Loads – FLUX Program

Stratified Estimates

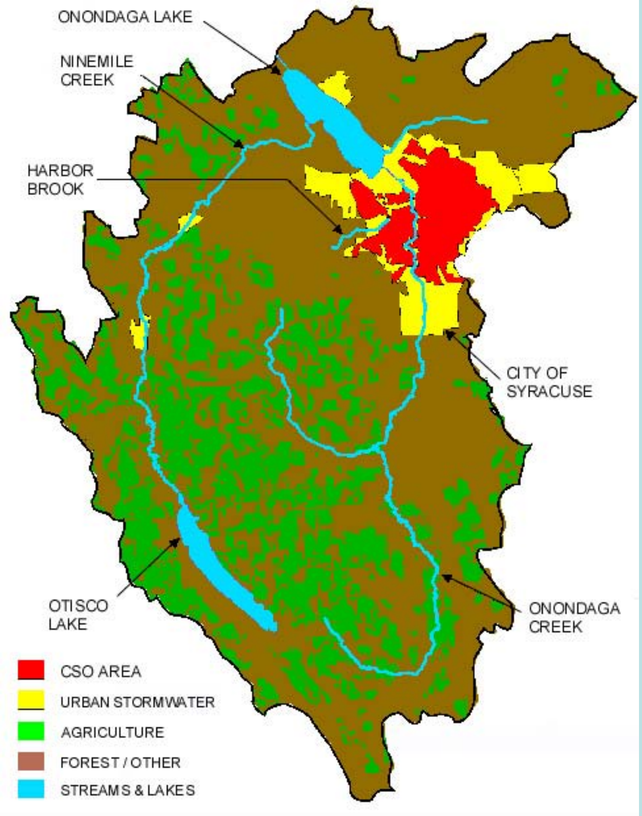
- >= Yearly Time Step
- Stratification Variables (Flow, Season)
- Precision Estimated
- Limitations
  - Sparse Data in Some Years and/or Strata
  - Hysteresis in Flow/Conc Relationship
  - Seasonality in Flow/Conc Relationship

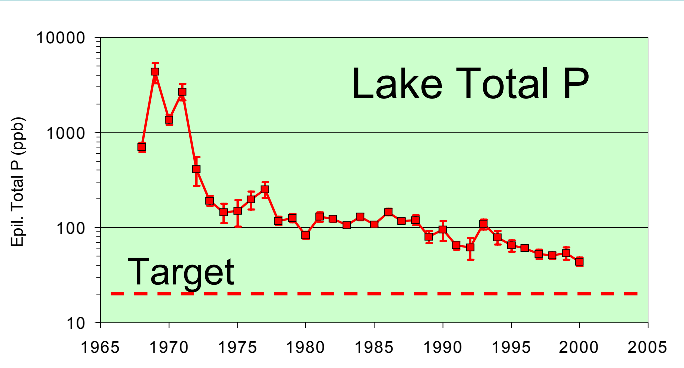
Computing Loads – FLUX Program

Time Series Methods

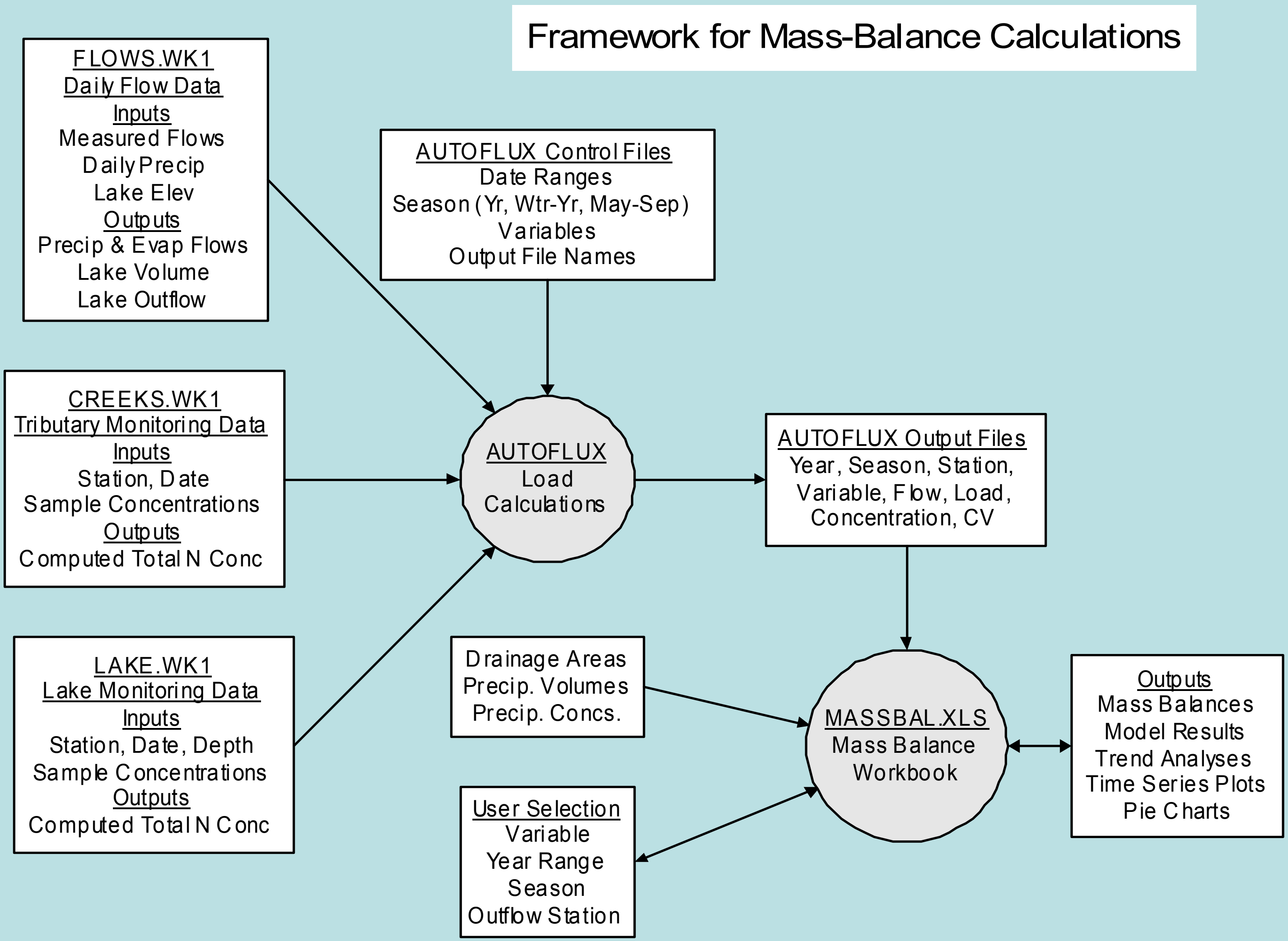
- >= Daily Time Step
- Algorithms
  - Interpolation Between Sampling Dates
  - Regression vs. Flow, Season, Trend (NOAA, 1999)
  - Regression + Interpolation (Walker & Havens, 2002)
- Limitations
  - Precision Difficult to Estimate
  - Data Gaps (Long Interpolation Intervals)
  - Serial Correlation in Residuals (Regression)
  - Change in Regression Slopes

Onondaga Lake, Syracuse, NY
Long-Term Monitoring (1968 – Present)





Mass Balance Framework (1986-2002)
Onondaga Lake, New York



Framework for Mass-Balance Calculations

FLOWS.WK1
Daily Flow Data
Inputs
Measured Flows
Daily Precip
Lake Elev
Outputs
Precip & Evap Flows
Lake Volume
Lake Outflow

AUTOFLUX Control Files
Date Ranges
Season (Yr, Wtr-Yr, May-Sep)
Variables
Output File Names

CREEKS.WK1
Tributary Monitoring Data
Inputs
Station, Date
Sample Concentrations
Outputs
Computed Total N Conc

AUTOFLUX
Load
Calculations

AUTOFLUX Output Files
Year, Season, Station,
Variable, Flow, Load,
Concentration, CV

LAKE.WK1
Lake Monitoring Data
Inputs
Station, Date, Depth
Sample Concentrations
Outputs
Computed Total N Conc

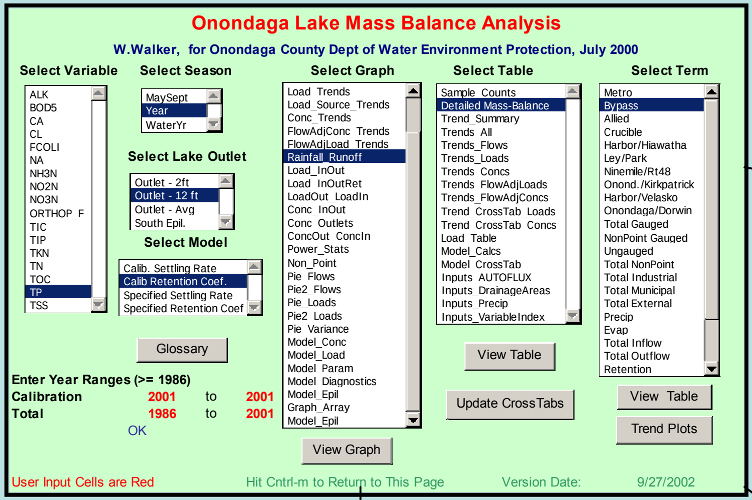
Drainage Areas
Precip. Volumes
Precip. Concs.

MASSBAL.XLS
Mass Balance
Workbook

Outputs
Mass Balances
Model Results
Trend Analyses
Time Series Plots
Pie Charts

User Selection
Variable
Year Range
Season
Outflow Station

Mass Balance Framework
Onondaga Lake, New York



Onondaga Lake Mass Balance Analysis

W.Walker, for Onondaga County Dept of Water Environment Protection, July 2000

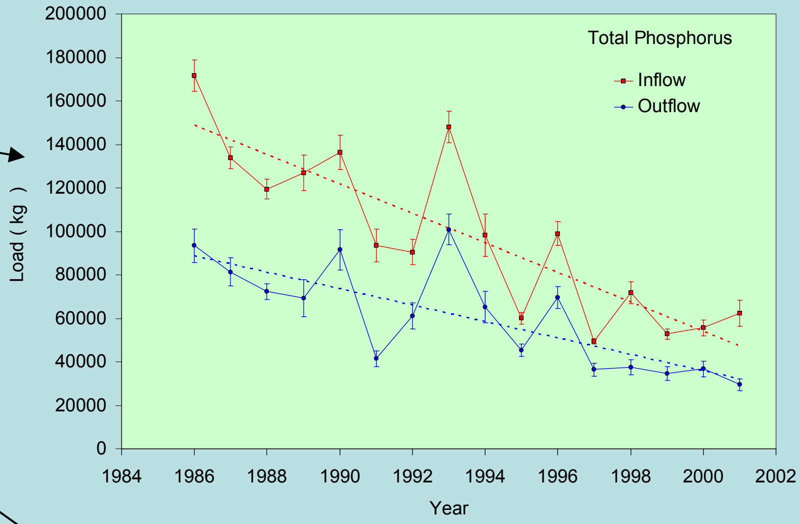
Enter Year Ranges (>= 1986)

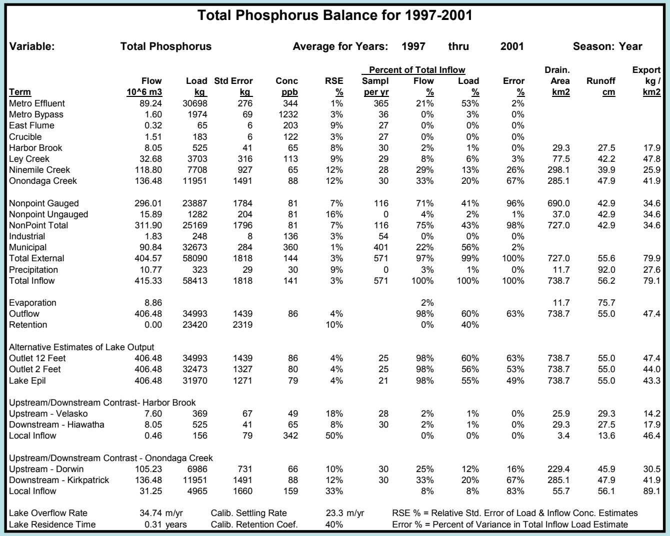
Calibration	2001	to	2001
Total	1986	to	2001

User Input Cells are Red

Hit Cntrl-m to Return to This Page

Version Date: 9/27/2002



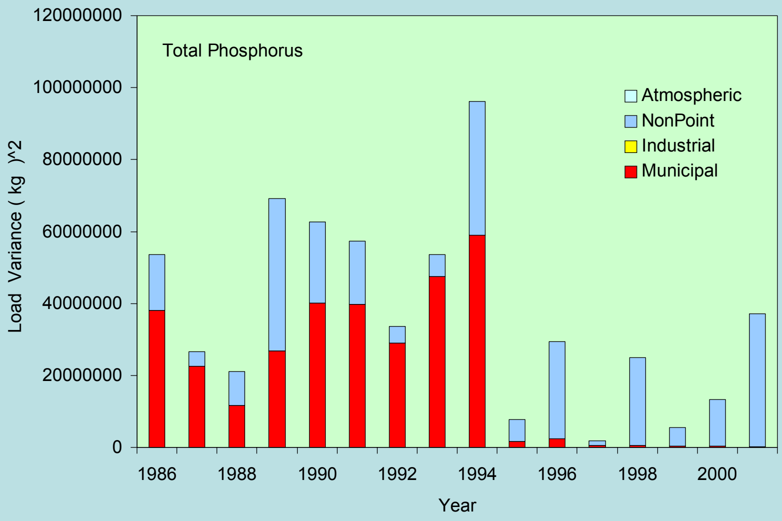


Total Phosphorus Balance for 1997-2001

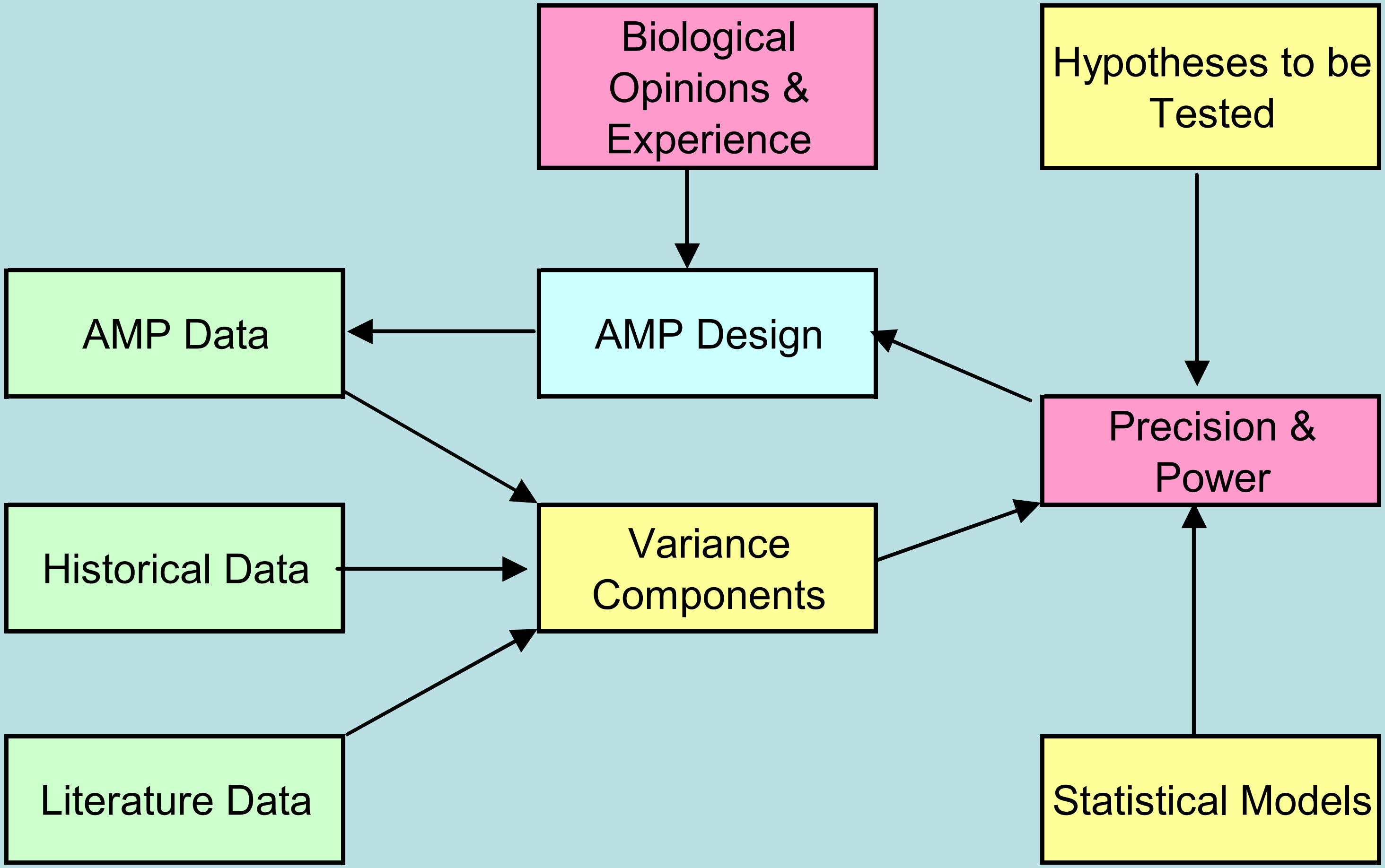
Variable: **Total Phosphorus** **Average for Years:** **1997** **thru** **2001** **Season: Year**

						Percent of Total Inflow				Drain.		Export
Term	Flow 10^6 m3	Load kg	Std Error kg	Conc ppb	RSE %	Sampl per yr	Flow %	Load %	Error %	Area km2	Runoff cm	kg / km2
Metro Effluent	89.24	30698	276	344	1%	365	21%	53%	2%			
Metro Bypass	1.60	1974	69	1232	3%	36	0%	3%	0%			
East Flume	0.32	65	6	203	9%	27	0%	0%	0%			
Crucible	1.51	183	6	122	3%	27	0%	0%	0%			
Harbor Brook	8.05	525	41	65	8%	30	2%	1%	0%	29.3	27.5	17.9
Ley Creek	32.68	3703	316	113	9%	29	8%	6%	3%	77.5	42.2	47.8
Ninemile Creek	118.80	7708	927	65	12%	28	29%	13%	26%	298.1	39.9	25.9
Onondaga Creek	136.48	11951	1491	88	12%	30	33%	20%	67%	285.1	47.9	41.9
Nonpoint Gauged	296.01	23887	1784	81	7%	116	71%	41%	96%	690.0	42.9	34.6
Nonpoint Ungauged	15.89	1282	204	81	16%	0	4%	2%	1%	37.0	42.9	34.6
NonPoint Total	311.90	25169	1796	81	7%	116	75%	43%	98%	727.0	42.9	34.6
Industrial	1.83	248	8	136	3%	54	0%	0%	0%			
Municipal	90.84	32673	284	360	1%	401	22%	56%	2%			
Total External	404.57	58090	1818	144	3%	571	97%	99%	100%	727.0	55.6	79.9
Precipitation	10.77	323	29	30	9%	0	3%	1%	0%	11.7	92.0	27.6
Total Inflow	415.33	58413	1818	141	3%	571	100%	100%	100%	738.7	56.2	79.1
Evaporation	8.86						2%			11.7	75.7	
Outflow	406.48	34993	1439	86	4%		98%	60%	63%	738.7	55.0	47.4
Retention	0.00	23420	2319		10%		0%	40%				
Alternative Estimates of Lake Output												
Outlet 12 Feet	406.48	34993	1439	86	4%	25	98%	60%	63%	738.7	55.0	47.4
Outlet 2 Feet	406.48	32473	1327	80	4%	25	98%	56%	53%	738.7	55.0	44.0
Lake Epil	406.48	31970	1271	79	4%	21	98%	55%	49%	738.7	55.0	43.3
Upstream/Downstream Contrast- Harbor Brook												
Upstream - Velasko	7.60	369	67	49	18%	28	2%	1%	0%	25.9	29.3	14.2
Downstream - Hiawatha	8.05	525	41	65	8%	30	2%	1%	0%	29.3	27.5	17.9
Local Inflow	0.46	156	79	342	50%		0%	0%	0%	3.4	13.6	46.4
Upstream/Downstream Contrast - Onondaga Creek												
Upstream - Dorwin	105.23	6986	731	66	10%	30	25%	12%	16%	229.4	45.9	30.5
Downstream - Kirkpatrick	136.48	11951	1491	88	12%	30	33%	20%	67%	285.1	47.9	41.9
Local Inflow	31.25	4965	1660	159	33%		8%	8%	83%	55.7	56.1	89.1

Lake Overflow Rate 34.74 m/yr Calib. Settling Rate 23.3 m/yr RSE % = Relative Std. Error of Load & Inflow Conc. Estimates
Lake Residence Time 0.31 years Calib. Retention Coef. 40% Error % = Percent of Variance in Total Inflow Load Estimate



Statistical Framework for Onondaga L. Monitoring Program



Biological Opinions & Experience

Hypotheses to be Tested

AMP Data

AMP Design

Precision & Power

Historical Data

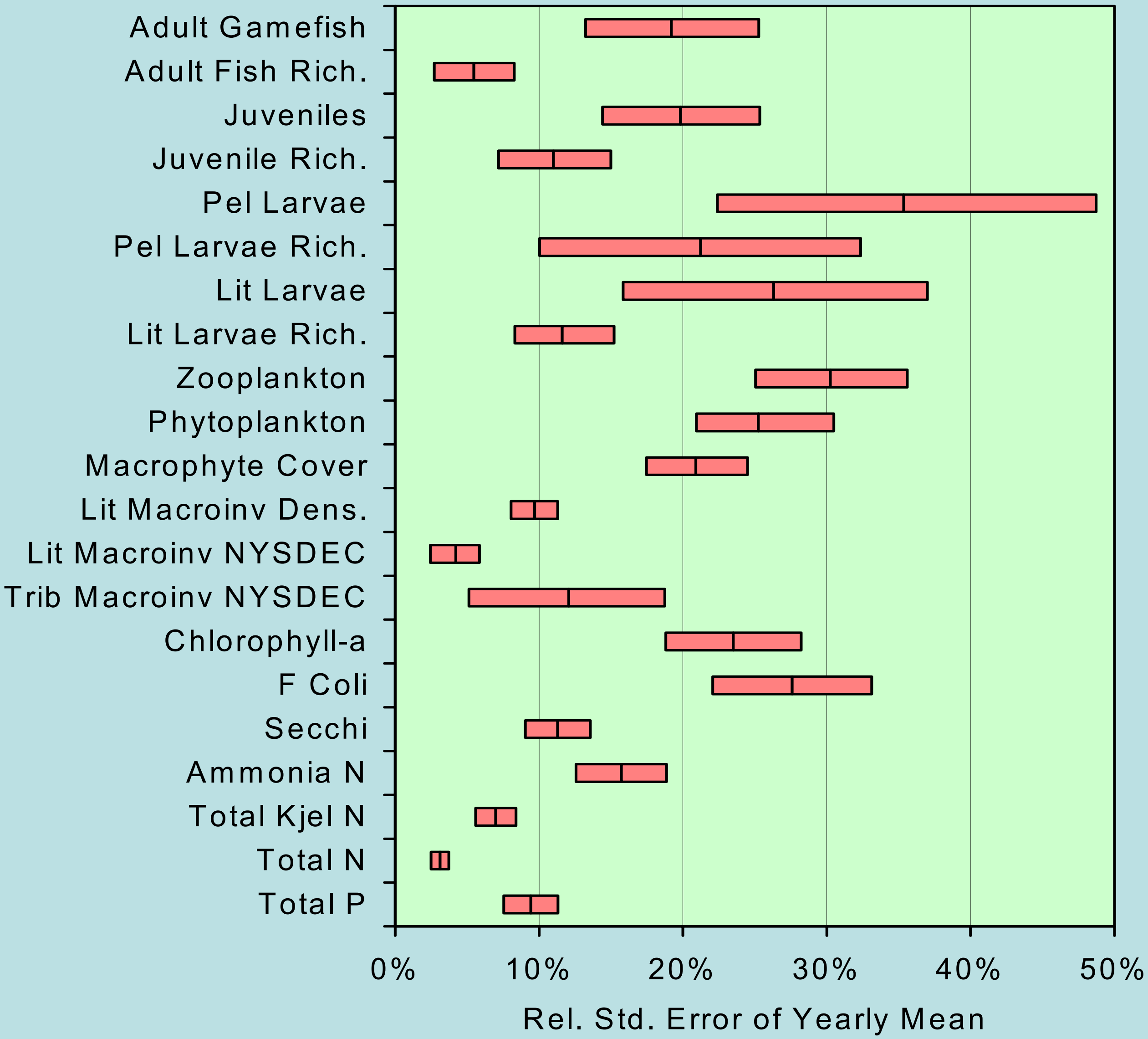
Variance Components

Literature Data

Statistical Models

Onondaga Lake Statistical Framework

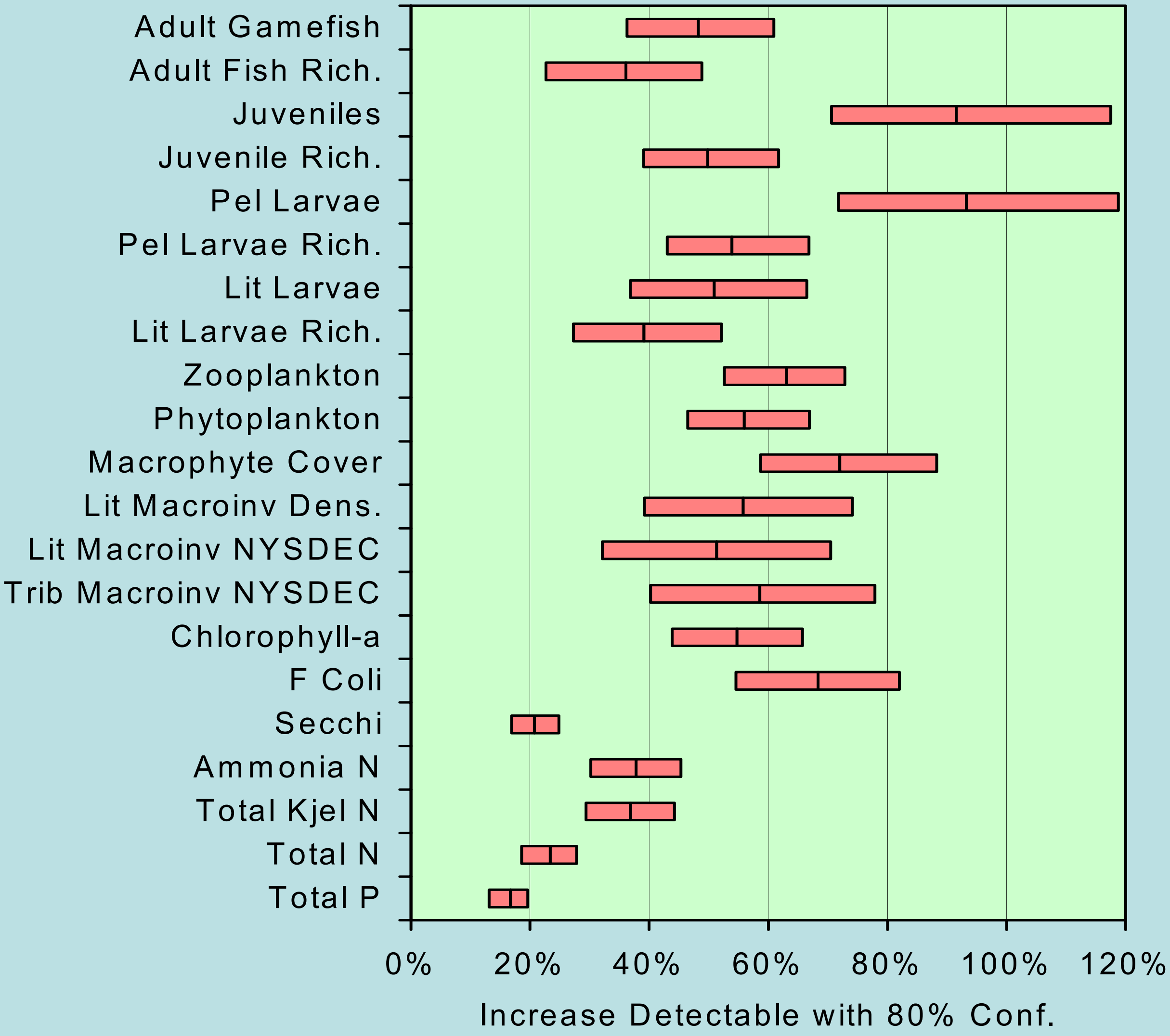
Precision Estimates: Relative Std. Error of Yearly Mean



Onondaga Lake Statistical Framework

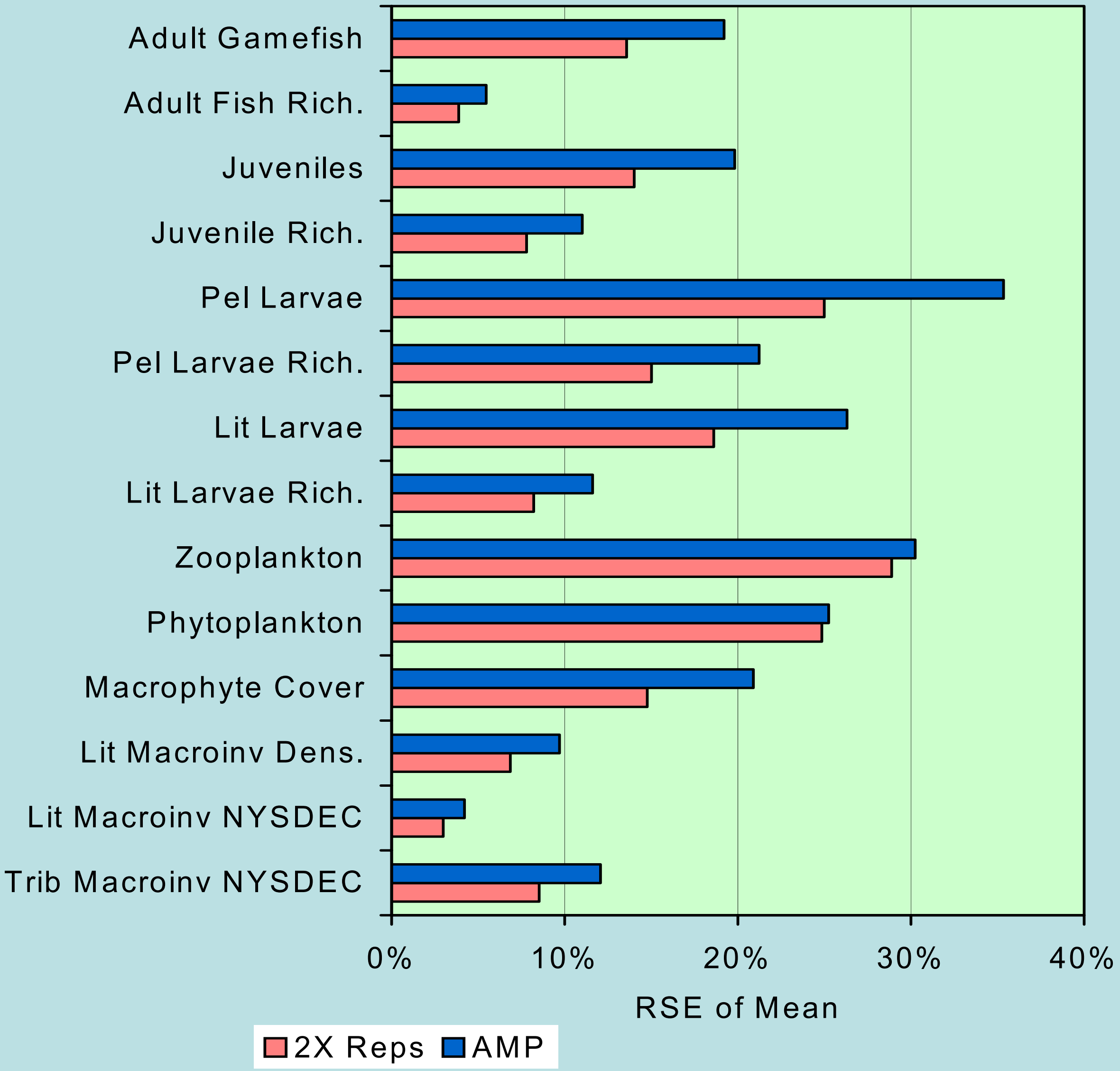
Power Estimates: Increases Detectable with 80% Confidence

5 Years of Baseline & 5 Years of Post-Baseline Data



Onondaga Lake Statistical Framework

Precision Sensitivity to Doubling Number of Sample Replicates



Onondaga Lake Statistical Framework

Precision of Fish Abundance, Richness, & Diversity Indices

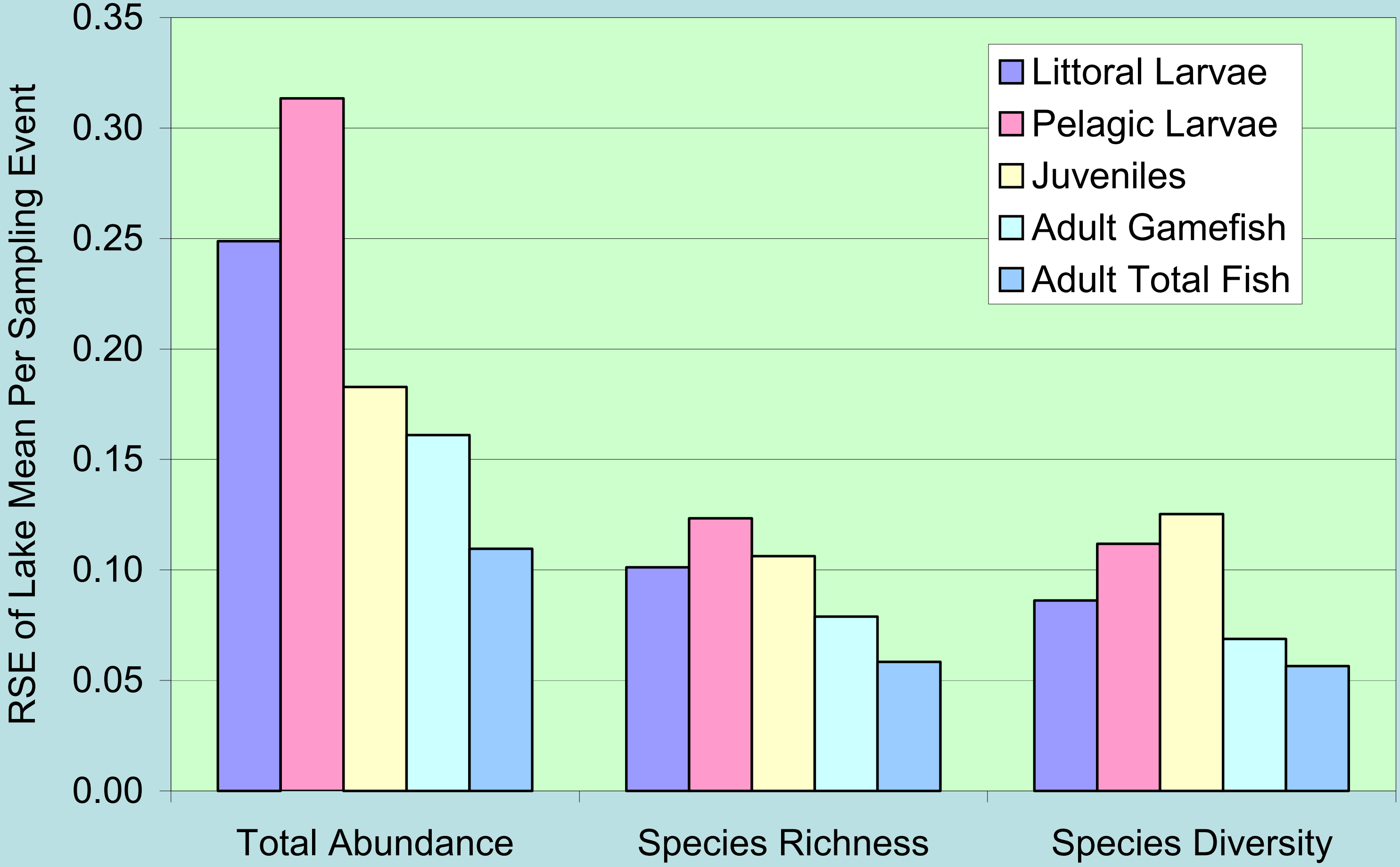
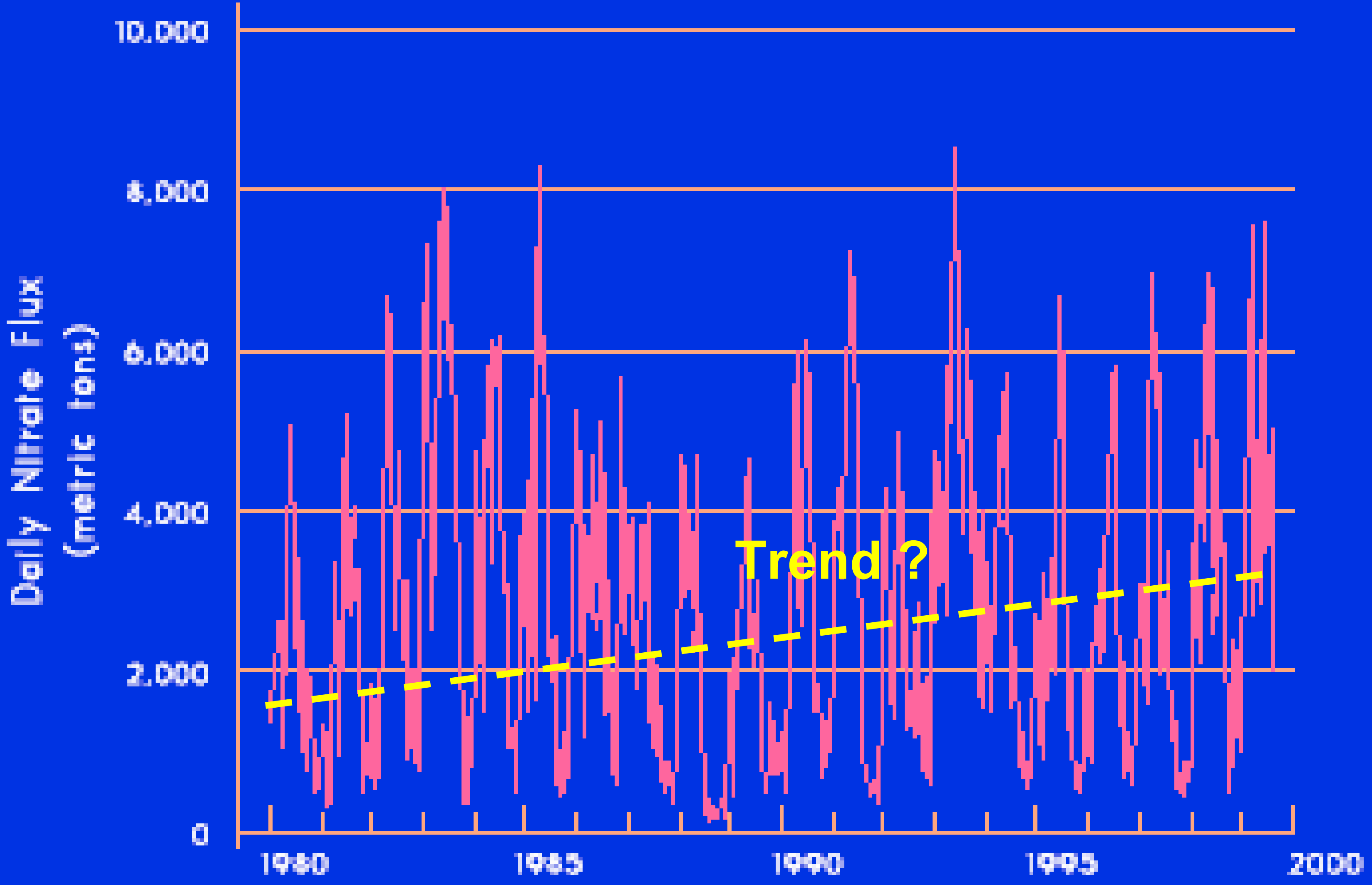


FIG. 2.8 Daily Flux of Nitrate from the Basin to the Gulf: August 1980–1999



Trend ?

Power for Detecting Trends

Given: N Years of Monitoring Data, H_0 = No Trend

Power = Probability of Rejecting H_0 when a Trend Exists

Power = Function (α, T)

α = Significance Level of Hypothesis Test

T = (Trend Magnitude) x $N^{1.5}$ / S

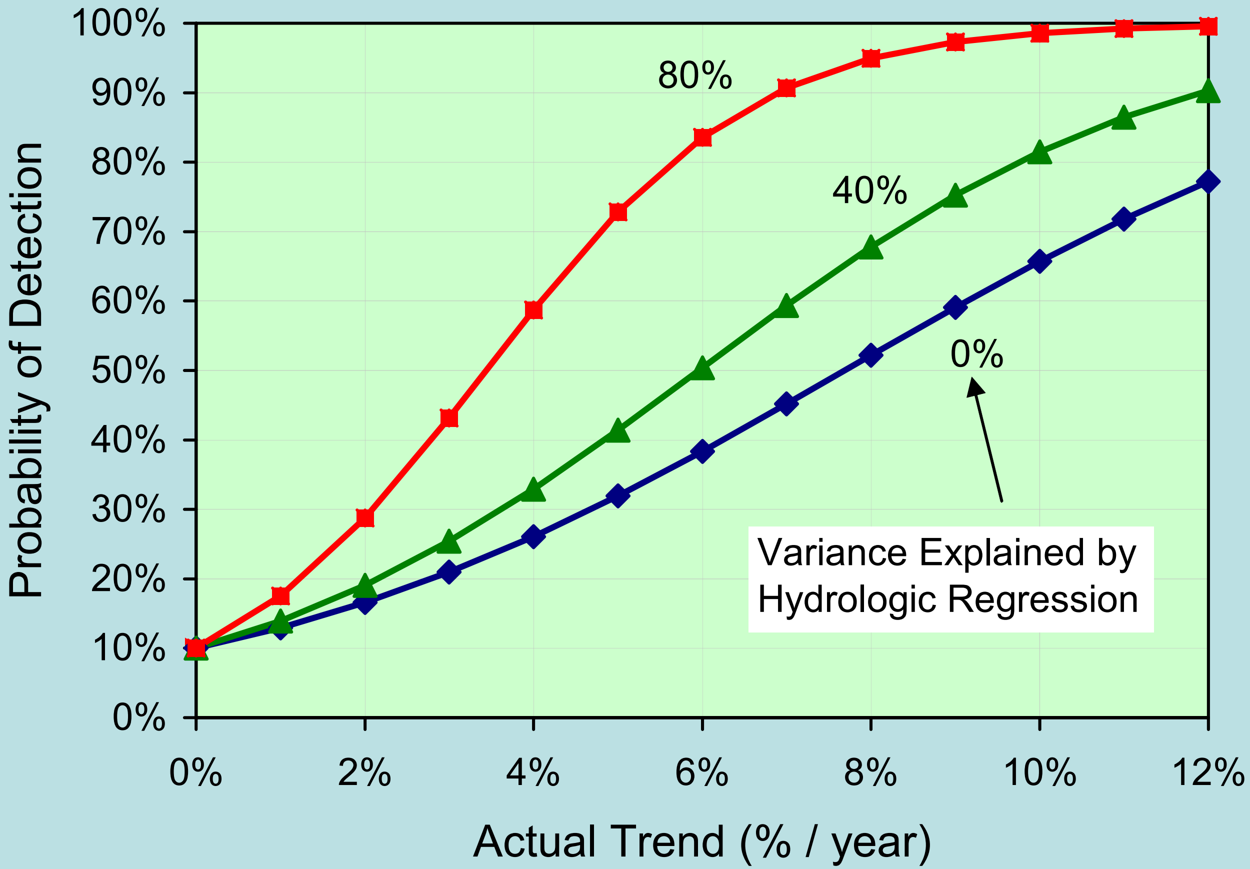
Factors Influencing S = Std. Dev. of Yearly Means

True Variability in System

Measurement Precision = F (Sampling Design)

Variance Explained by Hydrologic Factors

Power for Detecting a Linear Trend



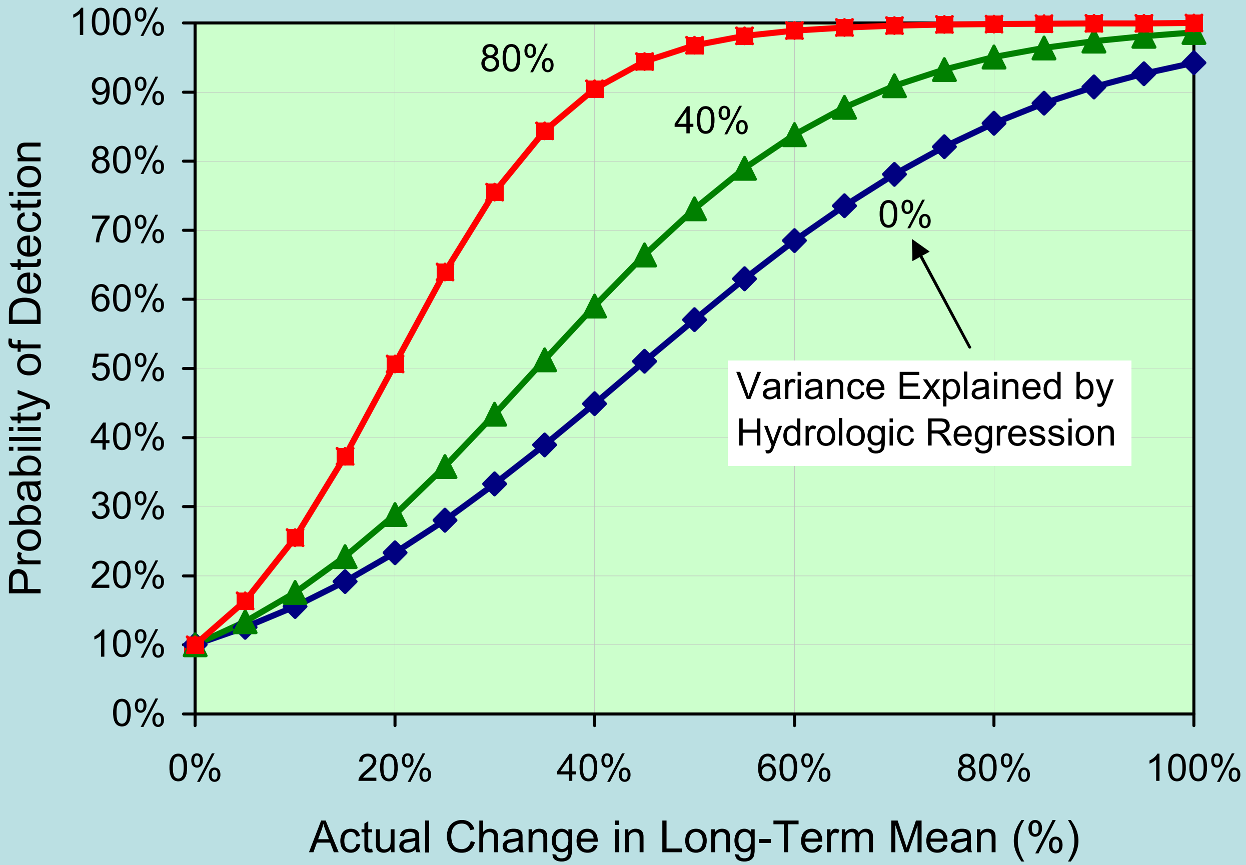
Probability of Rejecting Null Hypothesis of No Trend

Based upon Yearly Data from a 10-Year Period

Using Linear Regression (alpha = 0.1)

Year-to-Year Standard Deviation = 50%

Power for Detecting Step Change in Long-Term Mean



Probability of Rejecting Null Hypothesis of No Change in Mean

Comparing Yearly Means from Two 5-Year Periods

Using Student's t-test (alpha = 0.1)

Year-to-Year Standard Deviation = 50%

Model for Tracking Management Effects in the Presence of Hydrologic & Other Variations

Given: “Consistent” Long-Term Dataset

Regression Model:

……..Sources of Variation………

Meas. Load = Mean + Management + Hydrology + Random

$$L = M + \Delta M? + B\,(Q - Q_M) + E$$

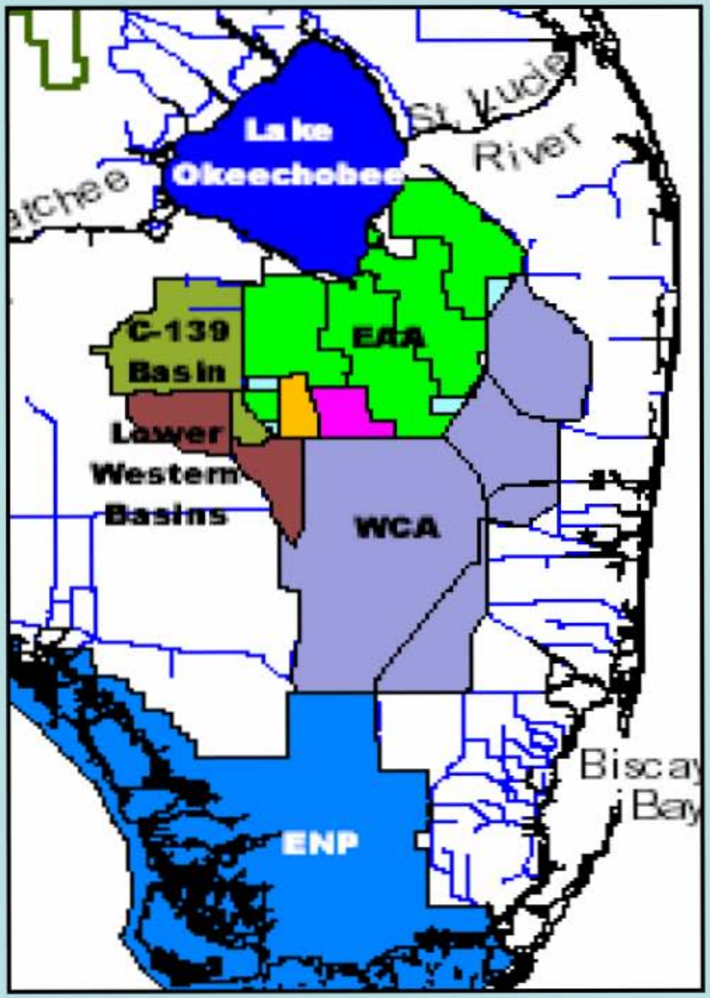
Q = Relevant Hydrologic Variable (Flow, Rainfall, Stage, etc.)

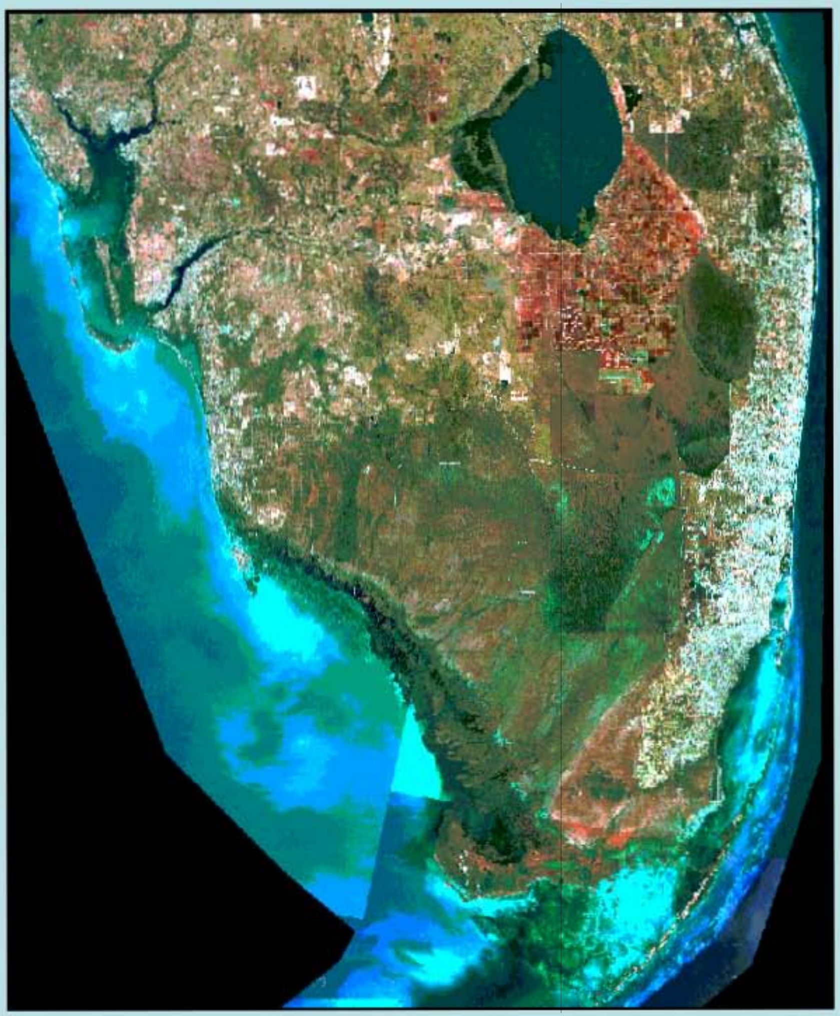
Hydrologically Adjusted Load:

$$L_A = L + B\,(Q_M - Q) = M + \Delta M? + E$$

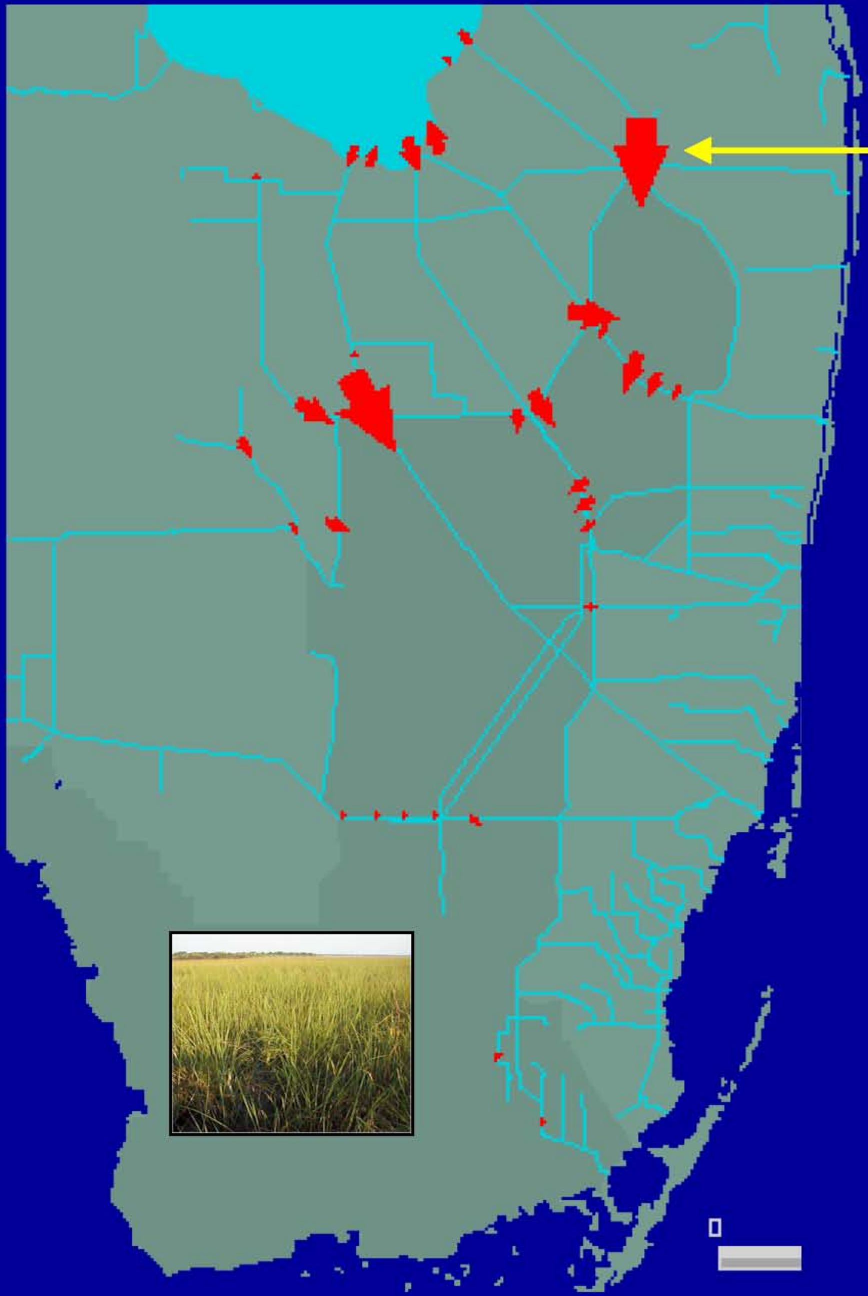
Removing Hydrologic Variation Increases Power for Detecting Change

Tracking Phosphorus Loads in the Everglades



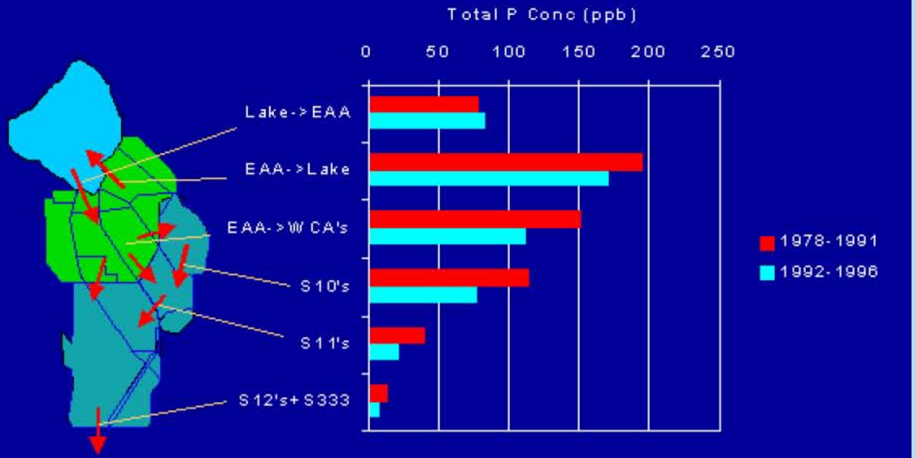


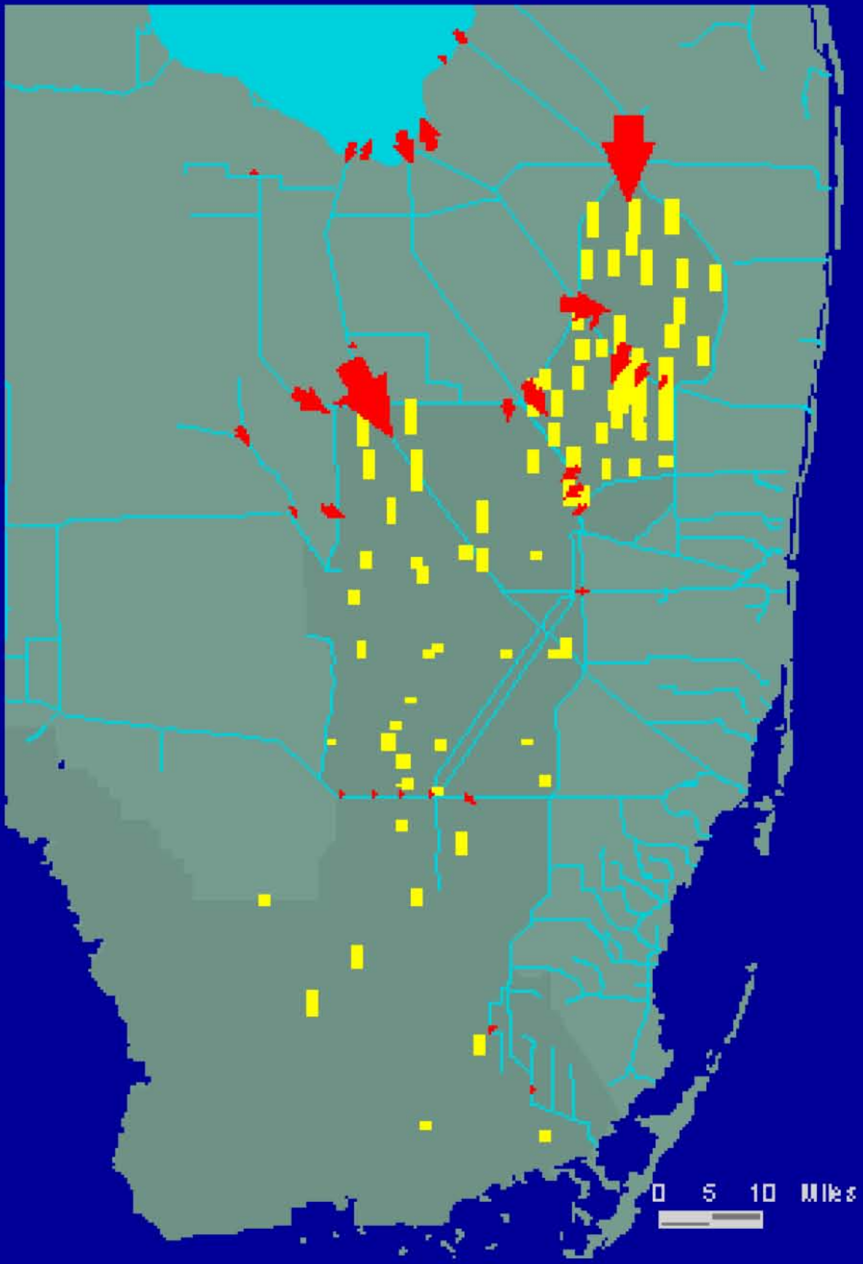
Structure TP Loads
WY 1978 - 1991



Flow-Weighted-Mean P Concentrations

Water Years 1978-1991 vs. 1992-1996





Structure TP Loads

&

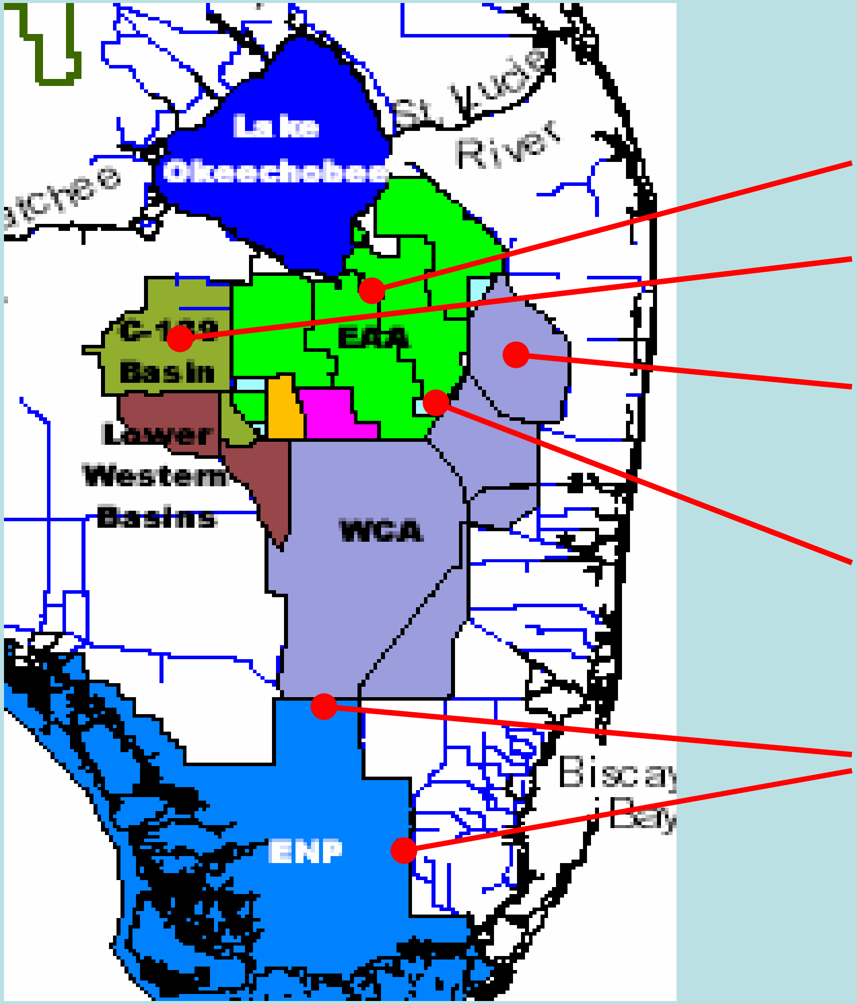
Marsh Frequencies
TP > 10 ppb

WY 1978 - 1991

Time Line for Everglades P Controls

		Inflow P (ppb)	
		Plan	Actual
1991	Law Suit Settlement	170	170
1995	Phase 1 BMP's 25% Reduc	130	100
1999-2003	Phase 1 Stormwater Treatment Areas	50	20-40
1995-2003	Treatment Technology Research		
1995-2002	Marsh P Threshold Research		
2003	Adopt Numeric P Standard (~ 10 ppb ?)		
2006	Enhanced BMP's / STA's	10 ?	?

Tracking Model Applications to Everglades



EAA BMP Rule

C139 BMP Rule

Loxahatchee Refuge
Marsh P Levels

Stormwater Treatment Area
Discharge Permits

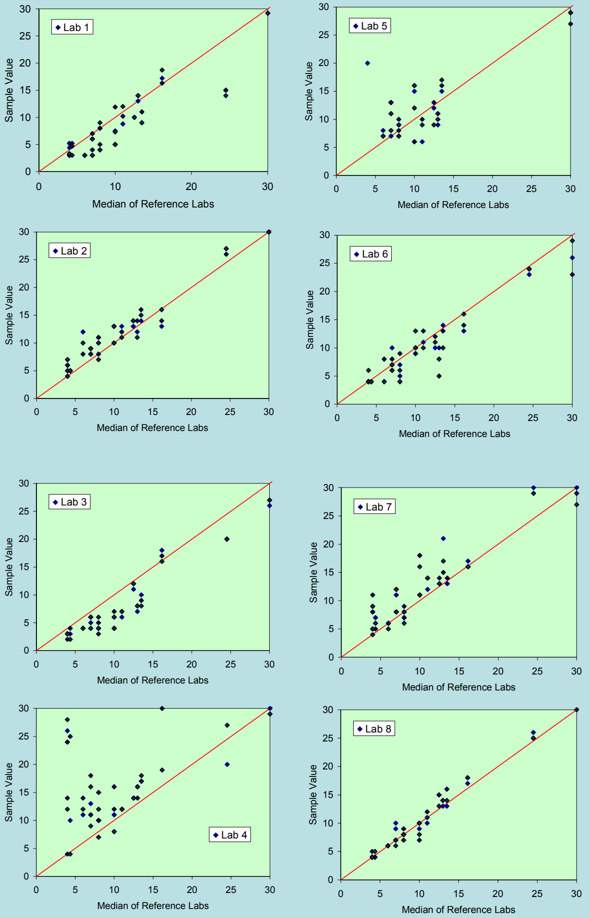
ENP Inflow P Limits

Total P Analysis - Everglades Round Robin

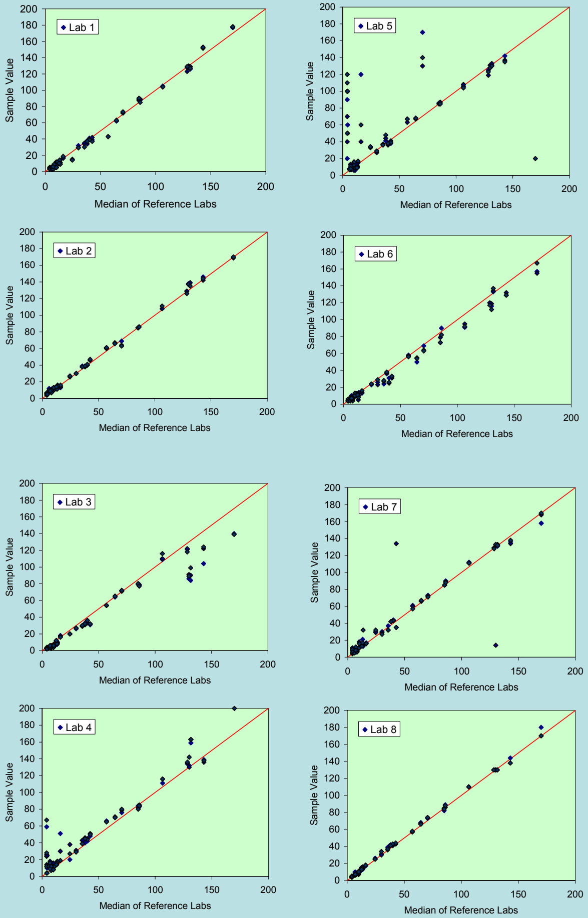
8 Labs, 33 Samples Collected on 7 Dates, Triplicate Analyses

Lab Result Compared with Median Value for Reference Labs

Low Range (0 – 30 ppb)

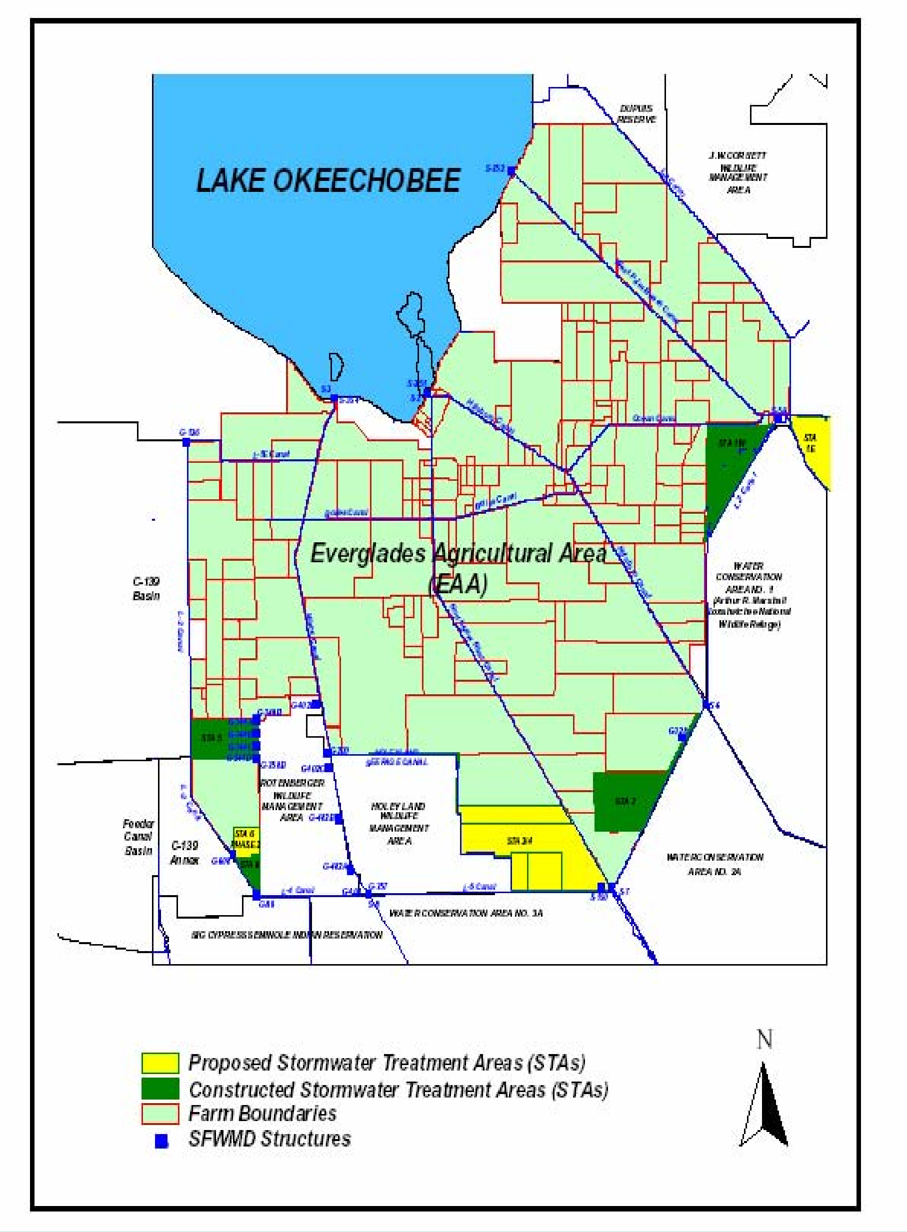


High Range (0 – 200 ppb)



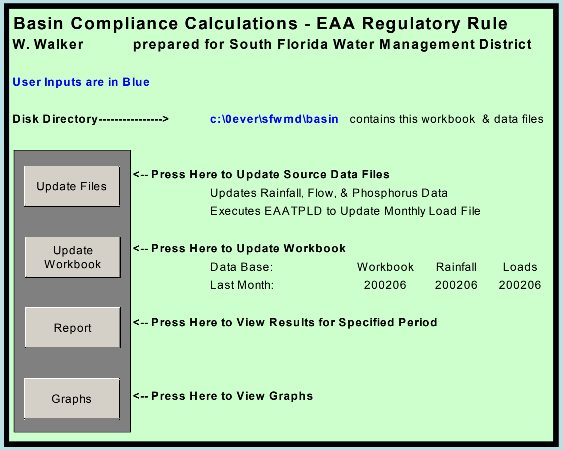
EVERGLADES
BEST MANAGEMENT PRACTICES
PROGRAM
SOUTH FLORIDA WATER MANAGEMENT DISTRICT

- Basin Area ~500,000 Acres
- Objectives
    - Implement BMP’s!
    - 25% Reduction in Basin P Load
    - 1979-1988 Baseline
- Regulatory Rule Effective 1995
- Monitoring Program
    - Weekly Composite Sampling
    - Basin-Scale ~35 Sites
    - Farm-Scale ~200 Sites



Software for Tracking EAA & WCA Phosphorus Loads

EAA Basin Compliance



Basin Compliance Calculations - EAA Regulatory Rule

W. Walker **prepared for South Florida Water Management District**

User Inputs are in Blue

Disk Directory----------------> **c:\0ever\sfwmd\basin** contains this workbook & data files

<-- Press Here to Update Source Data Files
Updates Rainfall, Flow, & Phosphorus Data
Executes EAATPLD to Update Monthly Load File

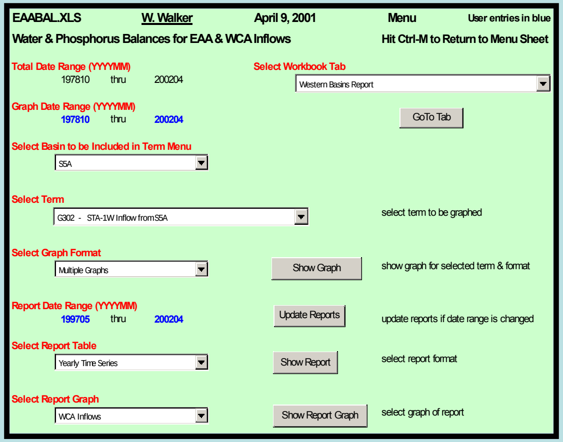
<-- Press Here to Update Workbook

Data Base:	Workbook	Rainfall	Loads
Last Month:	200206	200206	200206

<-- Press Here to View Results for Specified Period

<-- Press Here to View Graphs

System-Wide Trends



EAABAL.XLS **W. Walker** **April 9, 2001** **Menu** **User entries in blue**

Water & Phosphorus Balances for EAA & WCA Inflows **Hit Ctrl-M to Return to Menu Sheet**

Total Date Range (YYYYMM)
197810 thru 200204

Select Workbook Tab
Western Basins Report

Graph Date Range (YYYYMM)
197810 thru 200204

Select Basin to be Included in Term Menu
S5A

Select Term
G302 - STA-1W Inflow from S5A
select term to be graphed

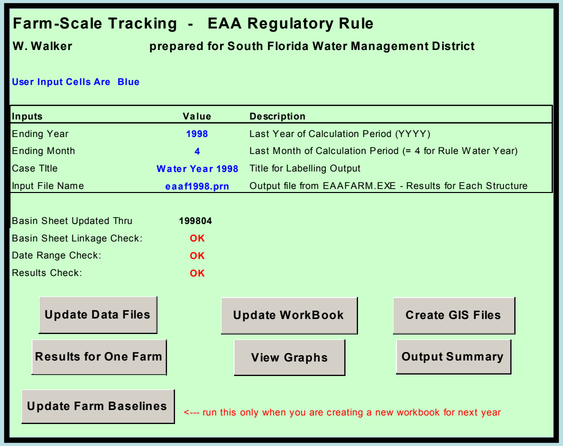
Select Graph Format
Multiple Graphs
show graph for selected term & format

Report Date Range (YYYYMM)
199705 thru 200204
update reports if date range is changed

Select Report Table
Yearly Time Series
select report format

Select Report Graph
WCA Inflows
select graph of report

Farm-Scale Data Analysis



Farm-Scale Tracking - EAA Regulatory Rule

W. Walker **prepared for South Florida Water Management District**

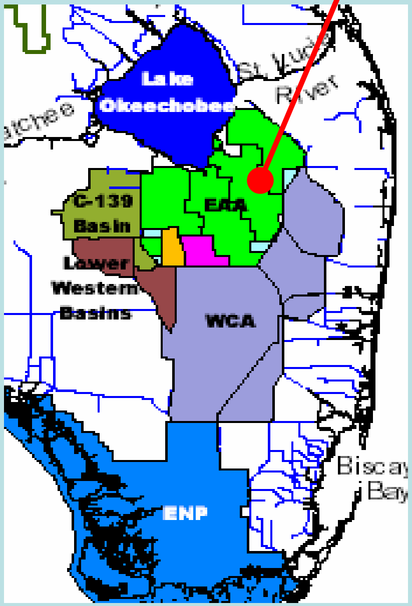
User Input Cells Are Blue

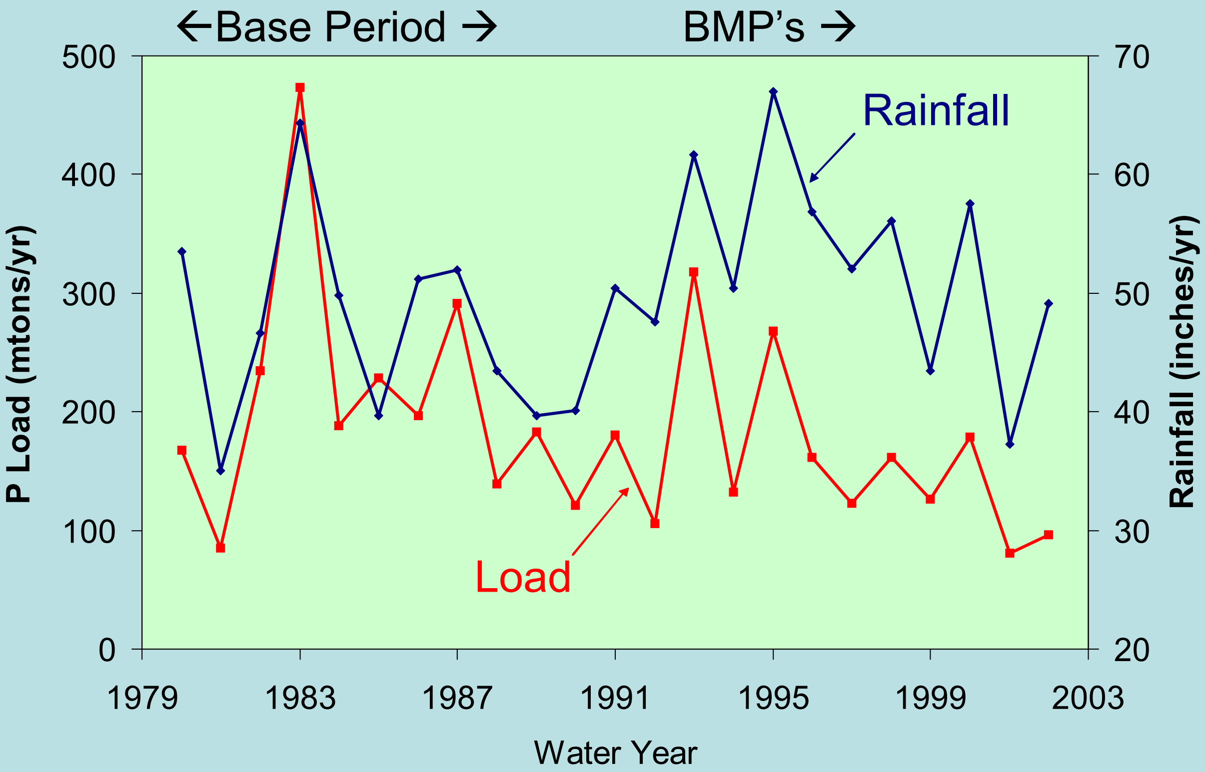
Inputs	Value	Description
Ending Year	1998	Last Year of Calculation Period (YYYY)
Ending Month	4	Last Month of Calculation Period (= 4 for Rule Water Year)
Case TItle	Water Year 1998	Title for Labelling Output
Input File Name	eaaf1998.prn	Output file from EAAFARM.EXE - Results for Each Structure

Basin Sheet Updated Thru	199804
Basin Sheet Linkage Check:	OK
Date Range Check:	OK
Results Check:	OK

<--- run this only when you are creating a new workbook for next year

Tracking EAA Total P Loads



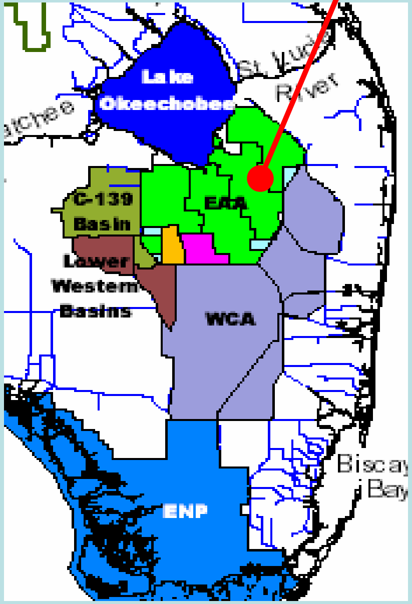


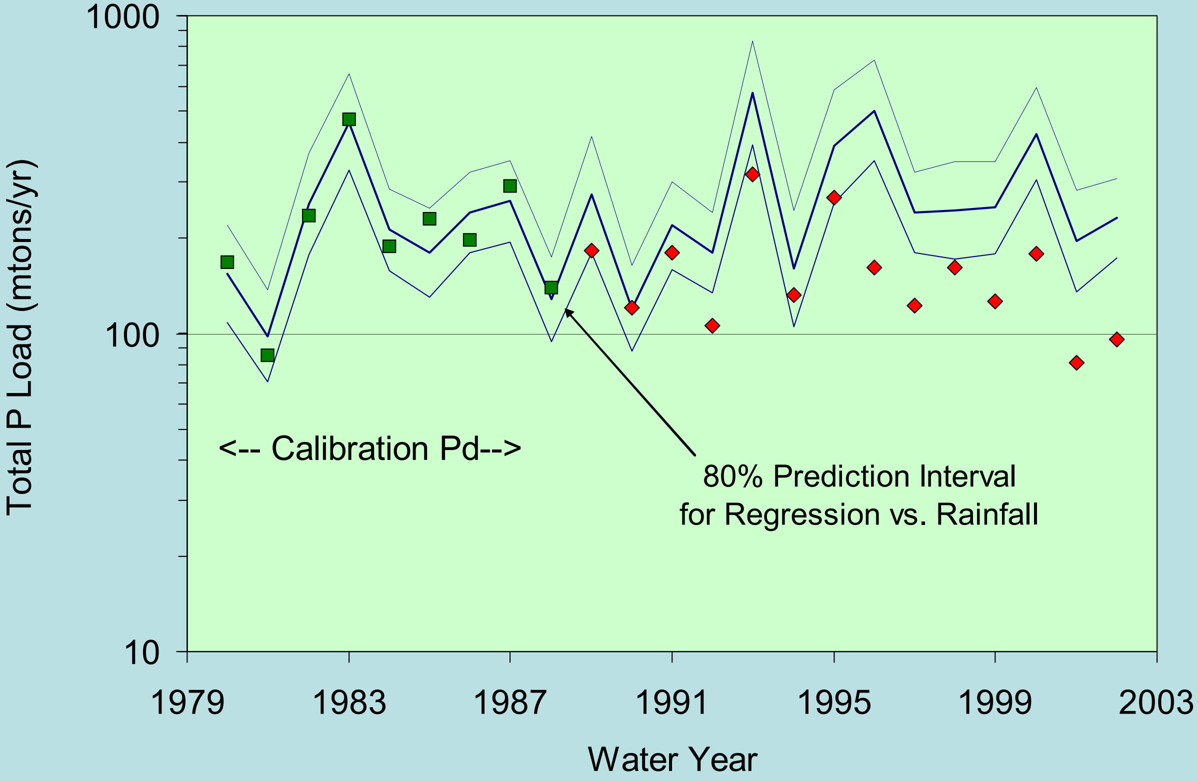
Model:

TP Load = Reduction + Rainfall-Effect + Random

Objective: 25% Load Reduction vs. 1979-88

Tracking EAA Total P Loads





<-- Calibration Pd-->

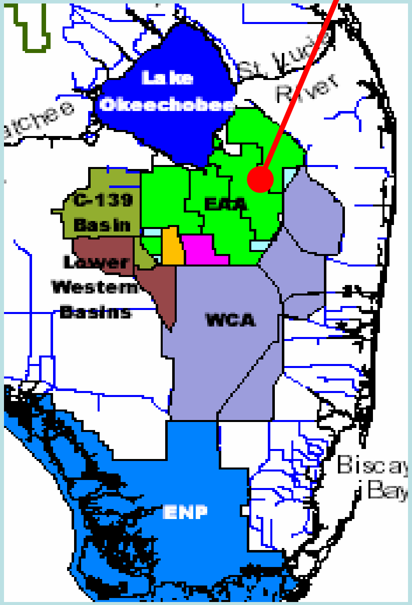
80% Prediction Interval
for Regression vs. Rainfall

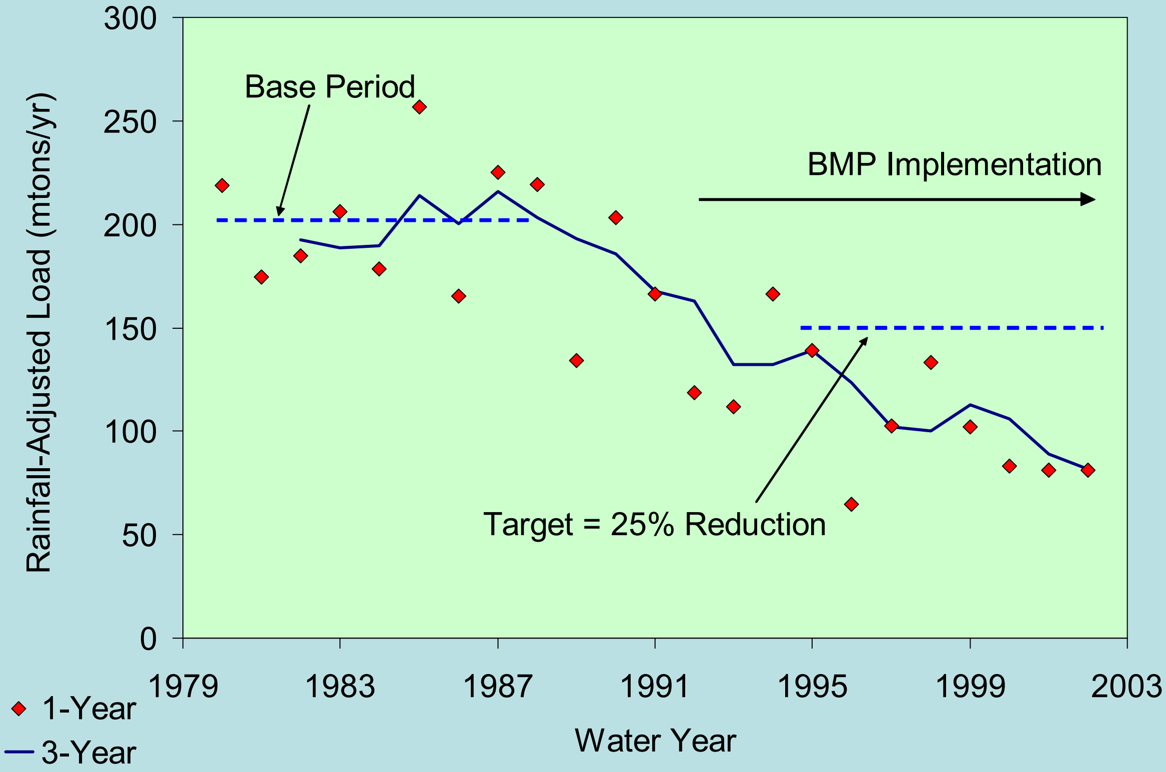
Model Calibration: $R^2 = 0.91$

TP Load = Reduction + Rainfall-Effect + Random

Objective: 25% Load Reduction vs. 1979-88

Tracking EAA Total P Loads

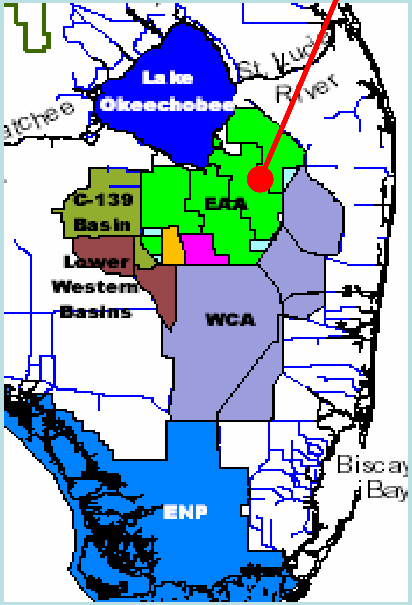


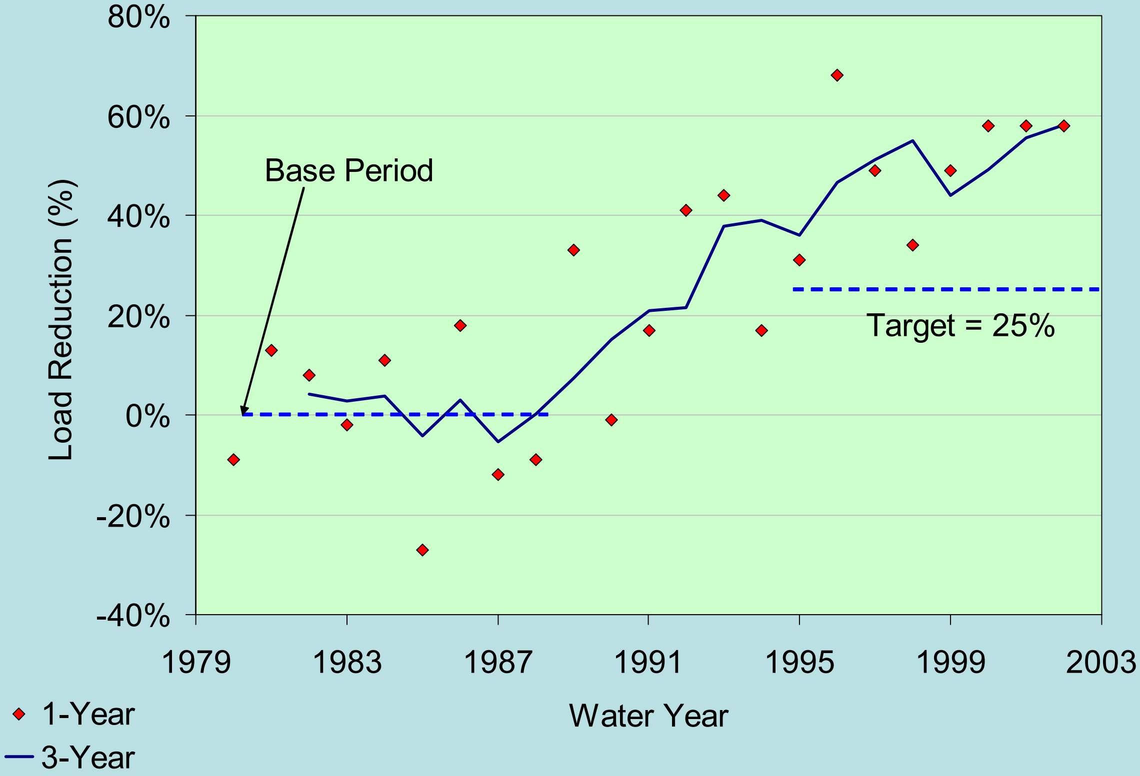


TP Loads Adjusted to Average Rainfall

Objective: 25% Load Reduction vs. 1979-88

Tracking EAA Total P Loads



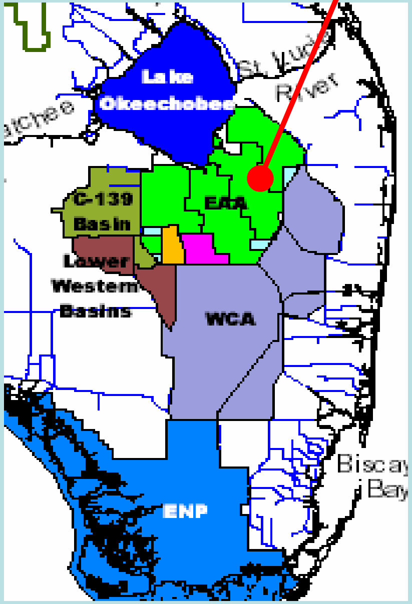


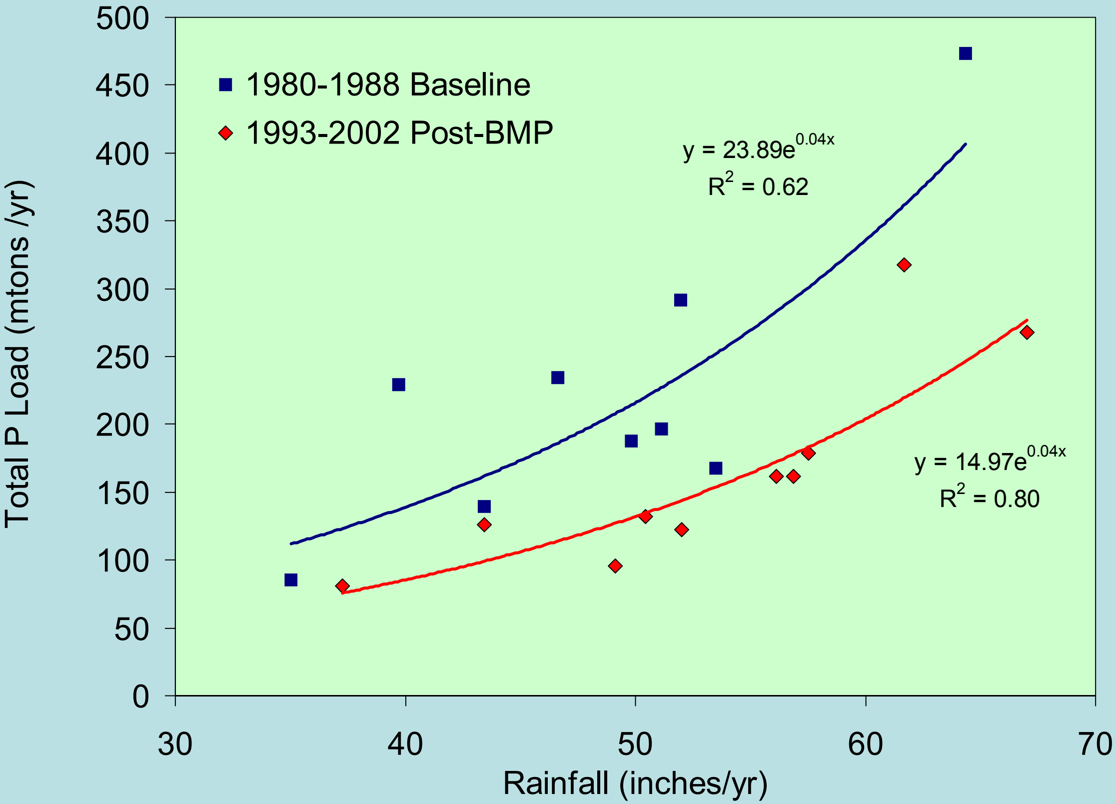
TP Load Reduction = 1 - LO / LP

LO = Observed Load

LE = Predicted from Rainfall, 1979-1988

Tracking EAA Total P Loads



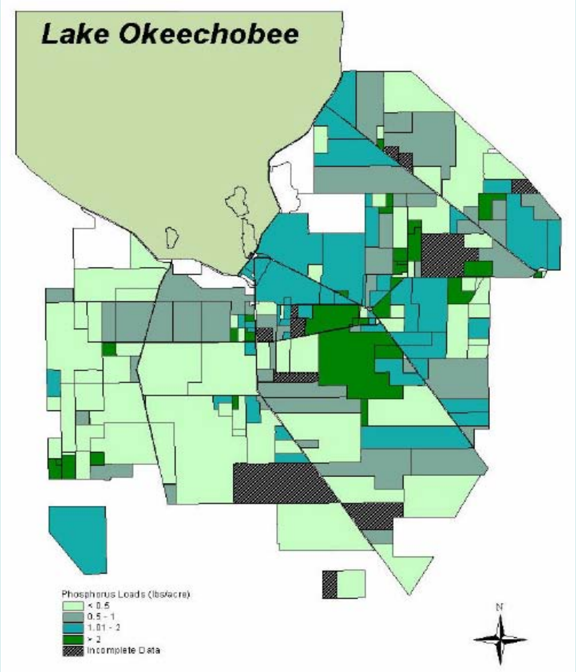


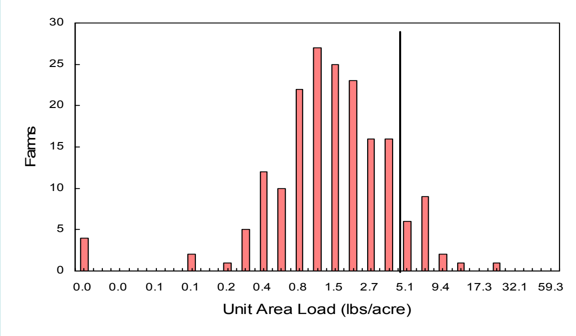
Model:

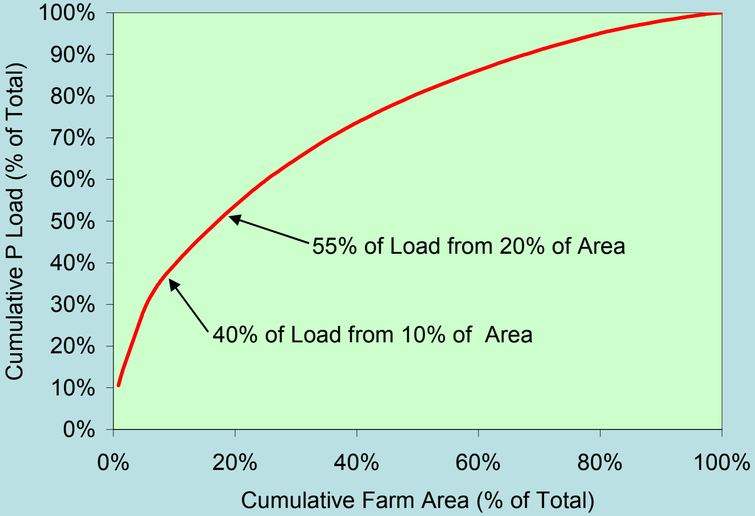
TP = Reduction + Rainfall-Effect + Random

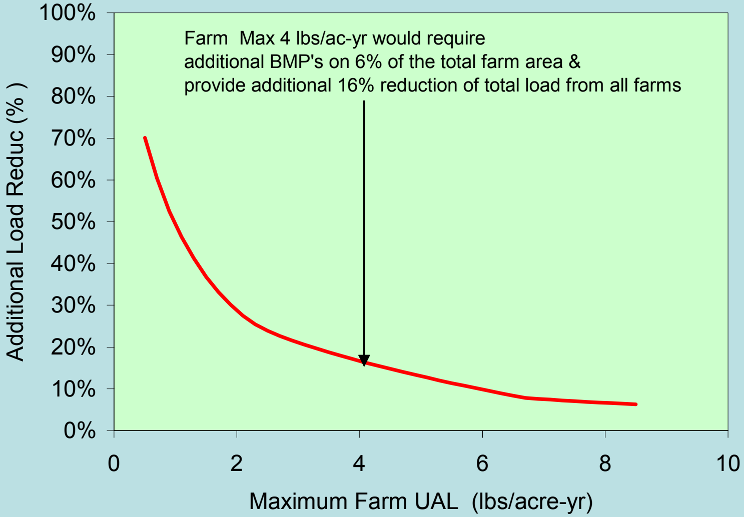
Objective: 25% Load Reduction vs. 1979-1988

Deriving Sub-Basin Targets to Achieve Basin Target
EAA Farm-Scale Unit Area Loading Rates, ~200 Hydrologic Units

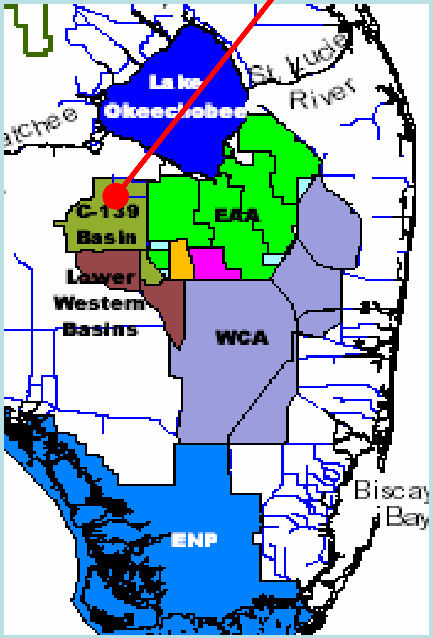


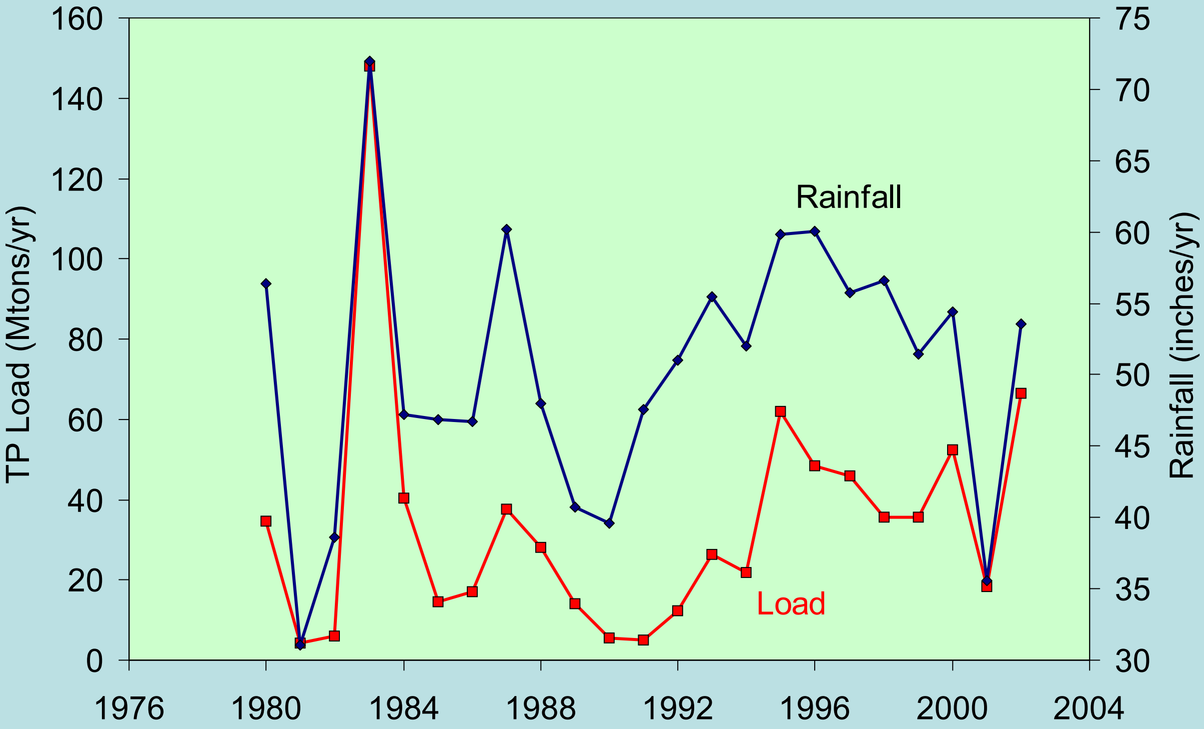






Tracking C139 Basin P Loads



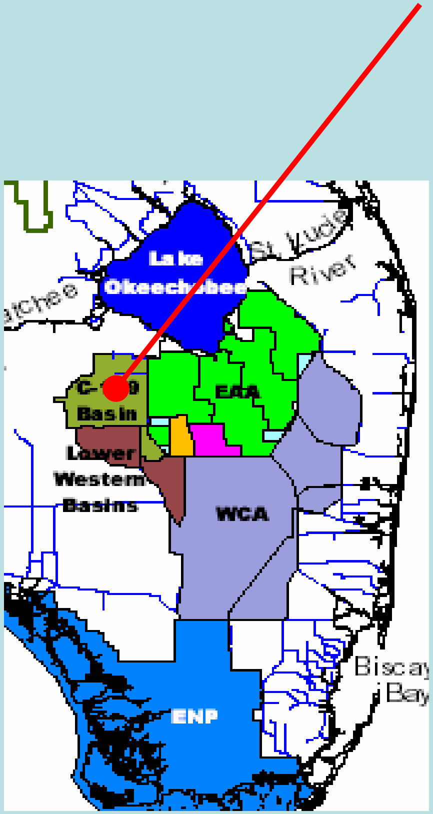


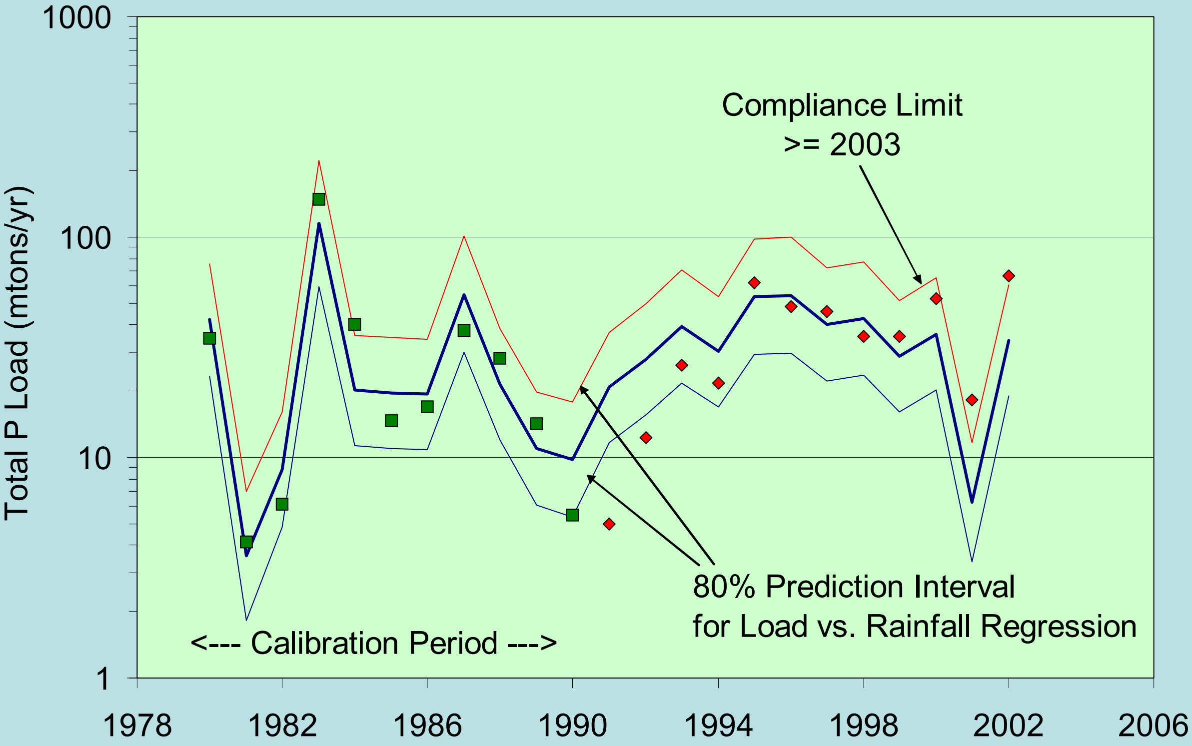
Model: Effective 2003

TP Load = Rainfall-Effect + Random

Objective: No Increase vs. 1979-1988

Tracking C139 Basin P Loads



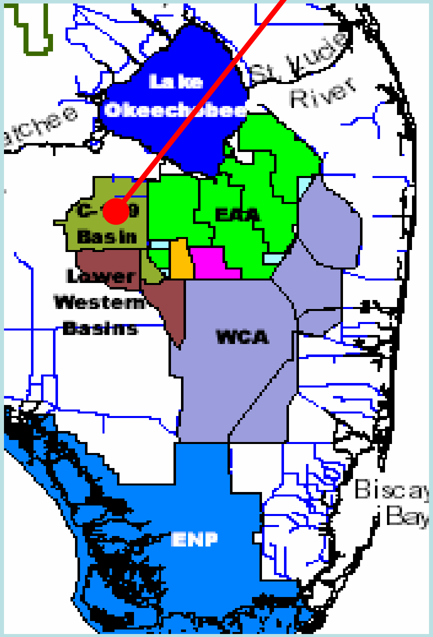


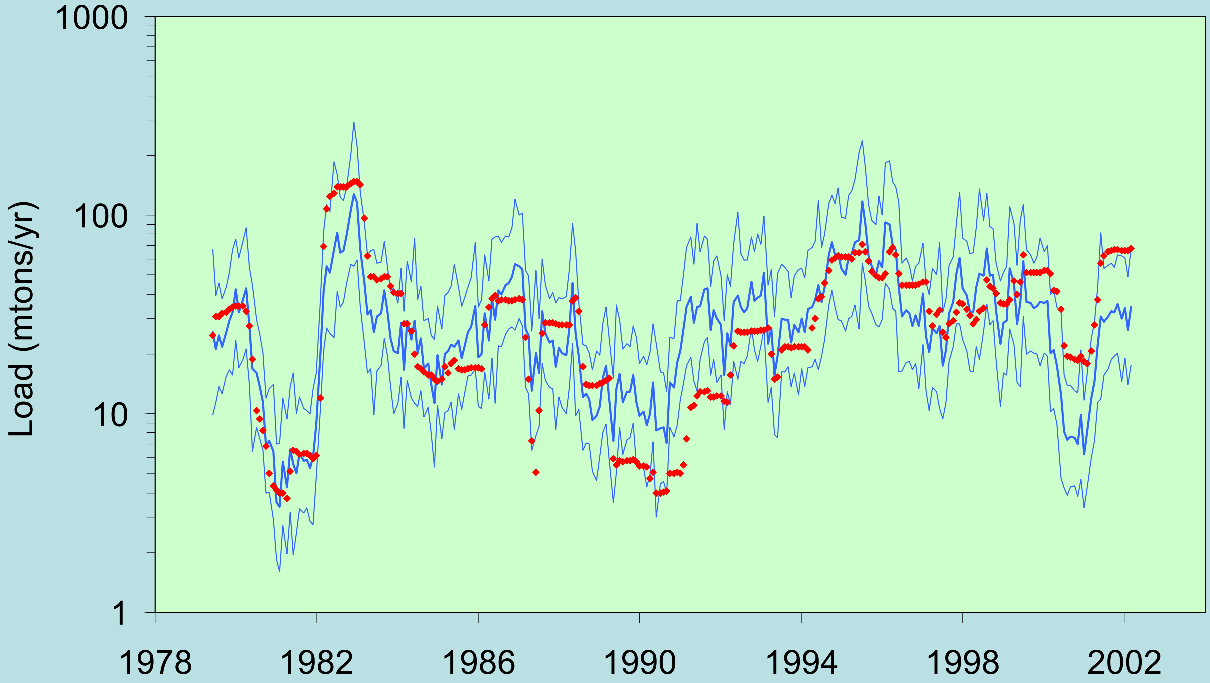
Model: $R^2 = 0.89$

TP Load = Rainfall-Effect + Random

Objective: No Increase vs. 1979-1988

Tracking C139 Basin P Loads

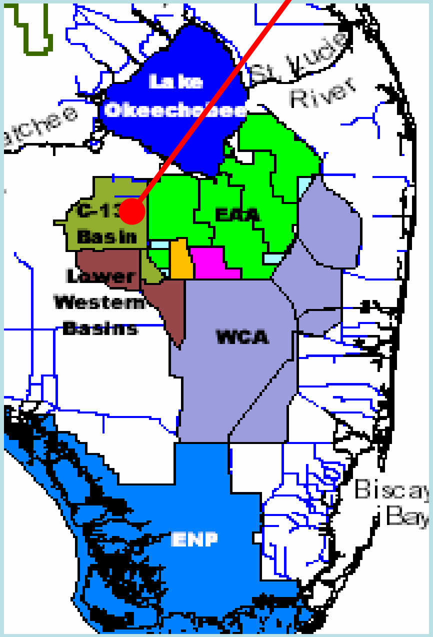


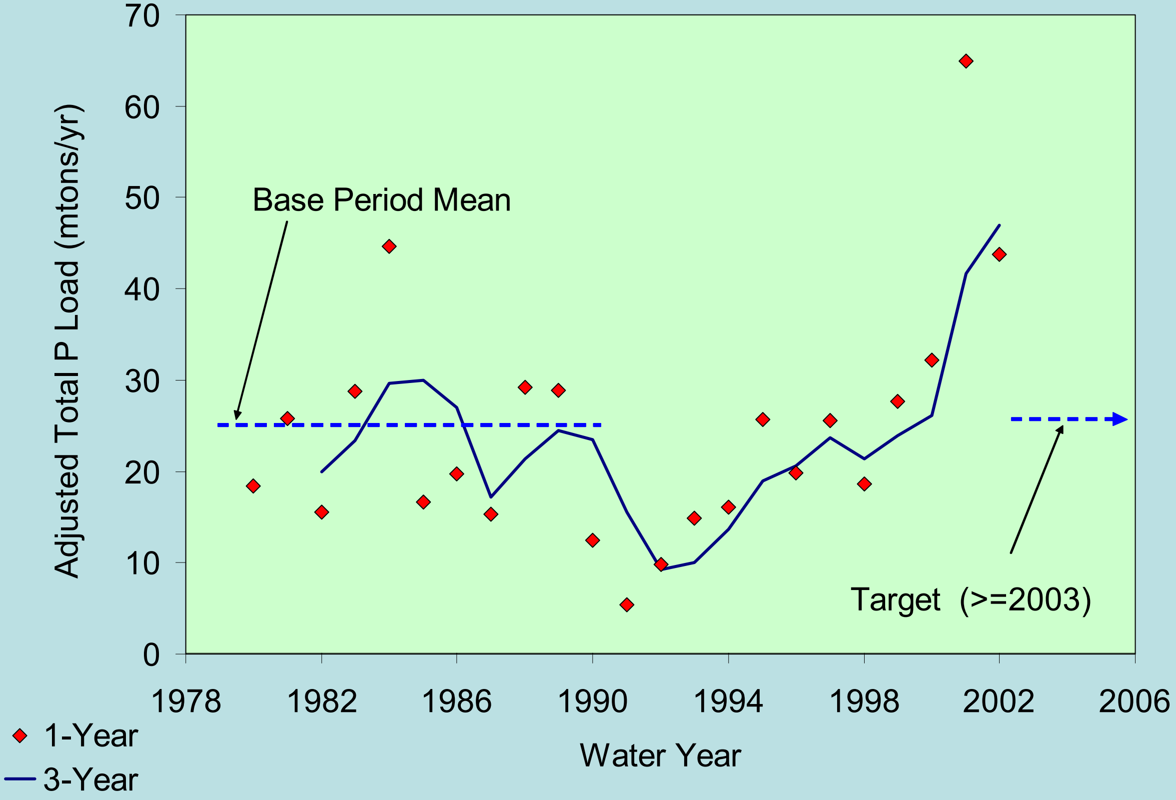


12-Month Rolling Average Loads

Observed & Predicted from Rainfall

Tracking C139 Basin P Loads

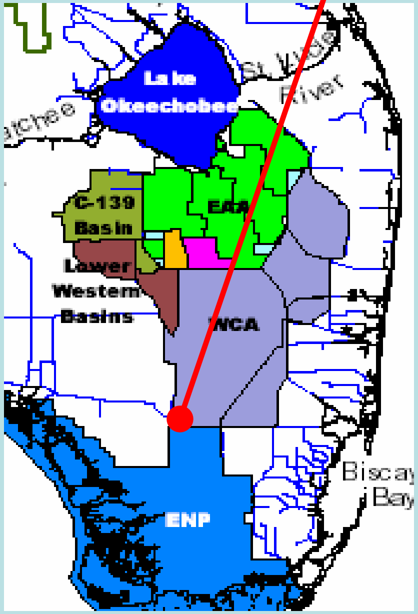


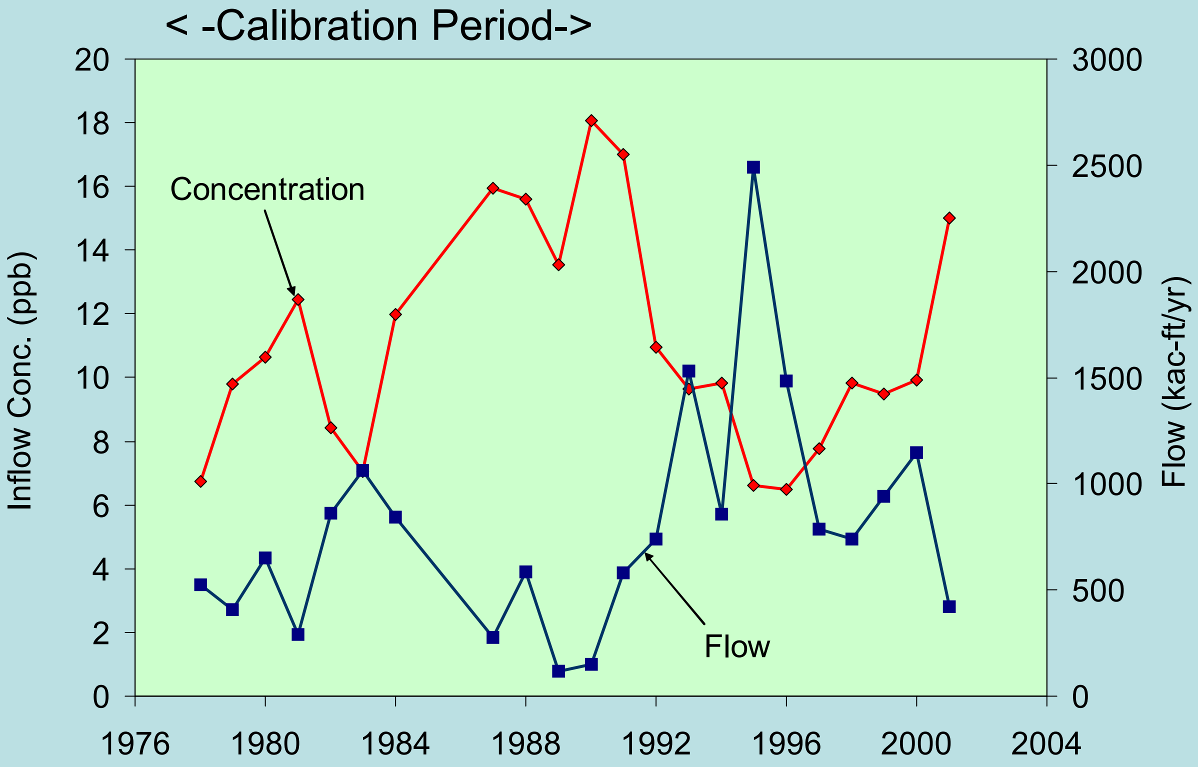


TP Loads Adjusted to Average Rainfall

Objective: No Increase vs. 1979-1988

Tracking ENP Inflow P Concentration



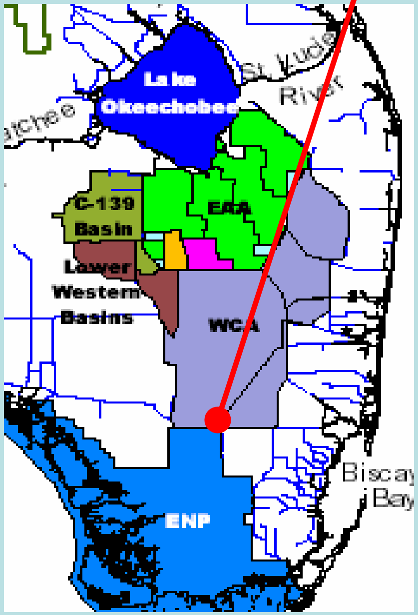


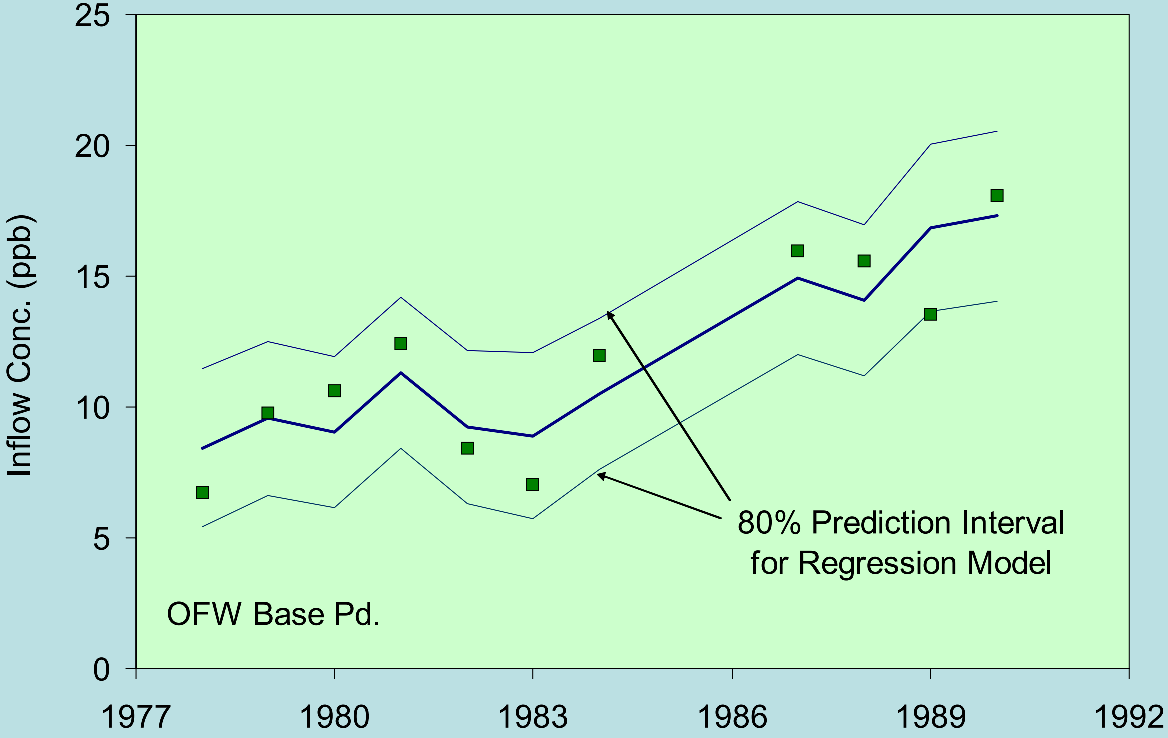
Model: Effective 2003

TP = Trend + Flow-Effect + Random

Objective: Restore 1978-1979 Levels

Tracking ENP Inflow P Concentration





80% Prediction Interval for Regression Model

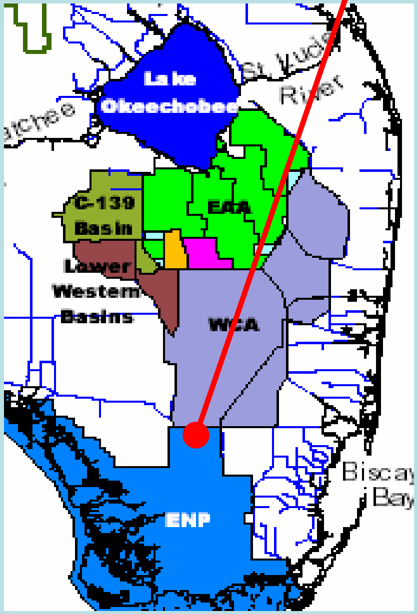
OFW Base Pd.

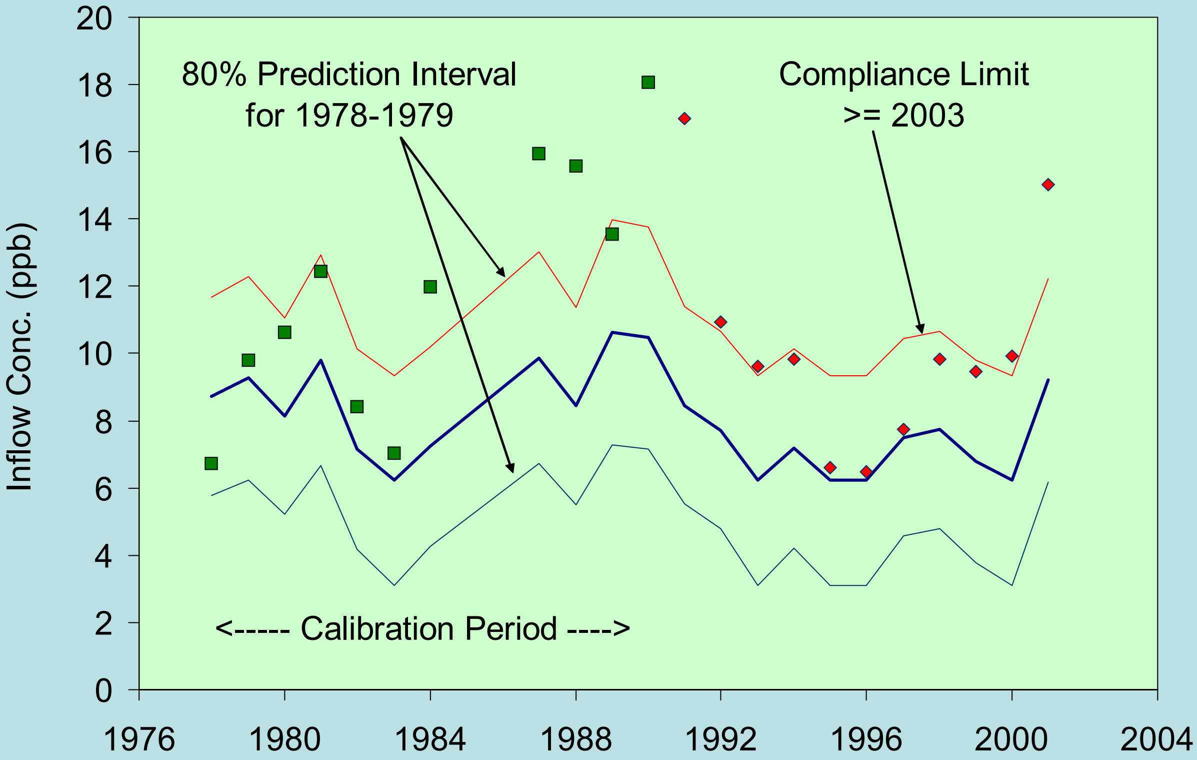
Model Calibration: $R^2 = 0.80$

TP = Trend + Flow-Effect + Random

Objective: Restore 1978-1979 Levels

Tracking ENP Inflow P Concentration



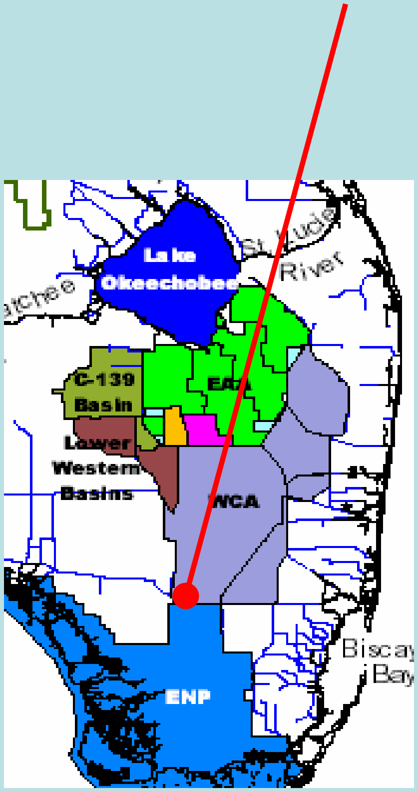


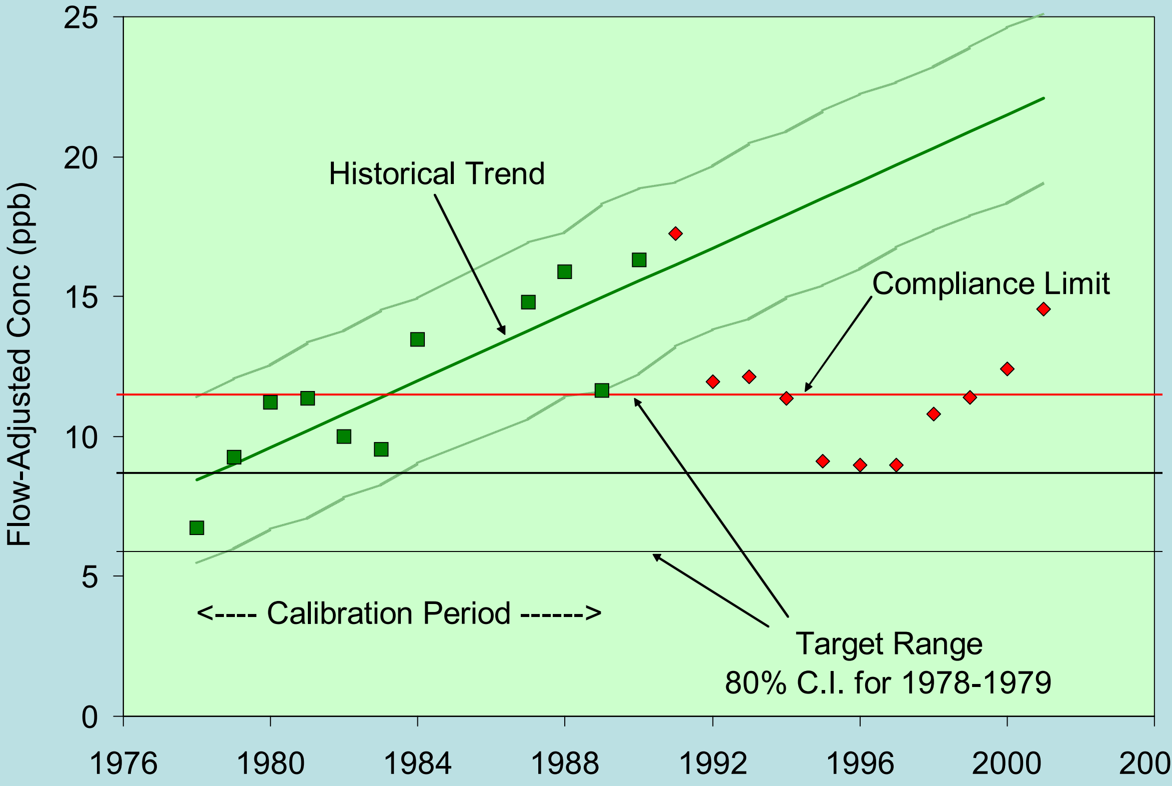
Flow-Dependent Compliance Limits

TP = Trend + Flow-Effect + Random

De-trended to 1978-1979 Conditions

Tracking ENP Inflow P Concentration





Inflow P Adjusted to Average Basin Flow

Objective: Restore 1978-1979 Levels

Related Links

This Presentation: wwwalker.net

FLUX – Load Calculation Software:

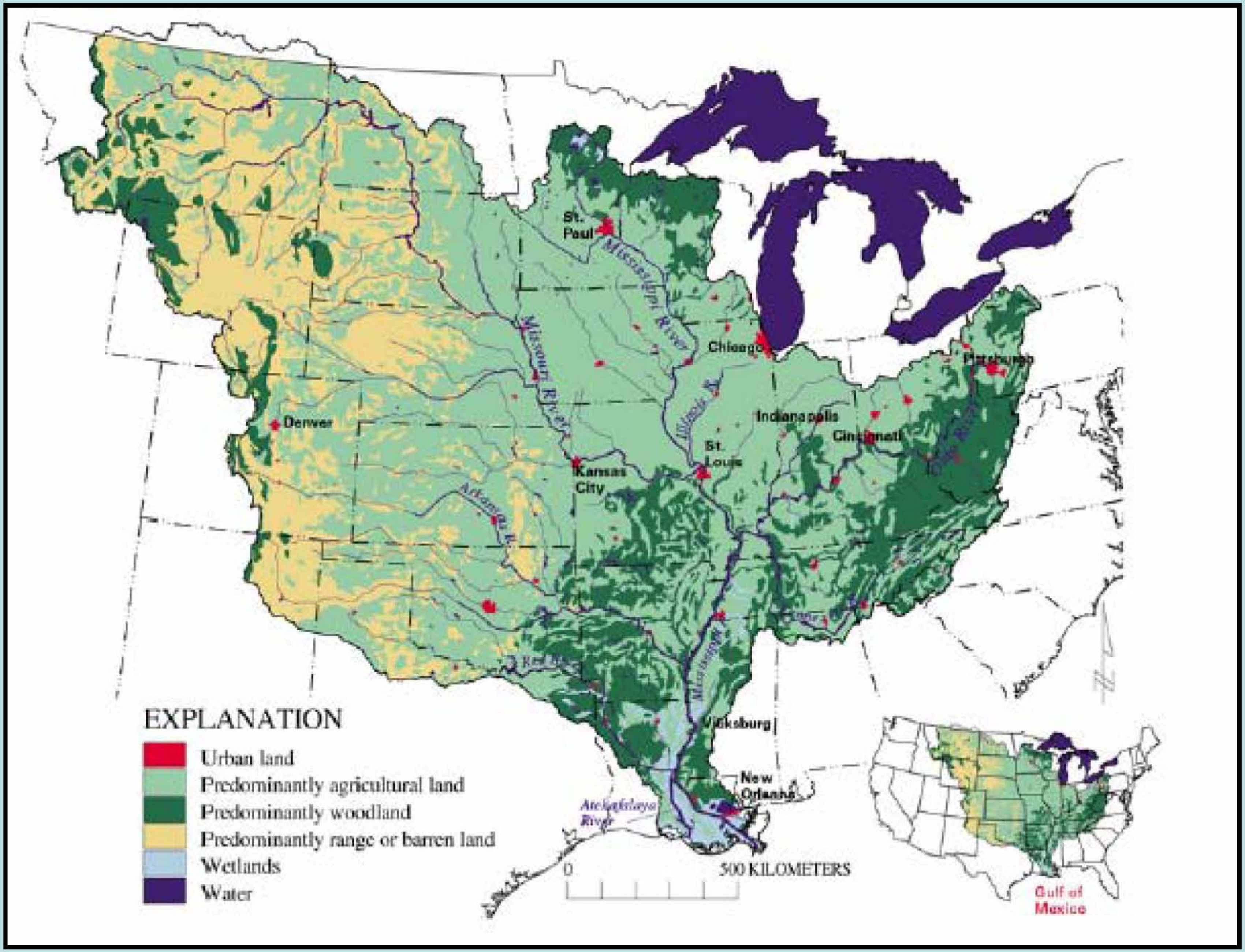
http://www.wes.army.mil/el/elmodels/emiinfo.html
http://www.wes.army.mil/el/elmodels/index.html

Everglades:

http://www.sfwmd.gov/org/wrp/
http://www.sfwmd.gov/org/wrp/wrp_evg/2_wrp_evg_glades/2_wrp_evg_glades.html
http://www.sfwmd.gov/org/ema/everglades/index.html
http://www.sfwmd.gov/org/ema/reports/settlement/sa_current.pdf
http://www.sfwmd.gov/org/reg/evg/eaareports/index.htm
http://www.wwwalker.net/pdf/pcriteria_everglades_epa2000.pdf
http://www.wwwalker.net/clearwtr/index.htm
http://www.wwwalker.net/pdf/wqtrends91.pdf

Onondaga Lake:

http://www.lake.onondaga.ny.us/


St. Paul
Mississippi River
Chicago
Pittsburgh
Missouri River
Illinois R.
Indianapolis
Cincinnati
Ohio River
Denver
St. Louis
Kansas City
Arkansas R.
Red River
Mississippi River
Vicksburg
New Orleans
Atchafalaya River
EXPLANATION
Urban land
Predominantly agricultural land
Predominantly woodland
Predominantly range or barren land
Wetlands
Water
0
500 KILOMETERS
Gulf of Mexico