Coeur d'Alene, Idaho
31 March – 4 April, 2003

Index 101

Use of Linear Discriminant Models to Determine Life Use Attainment

Tom Danielson, Susan Davies, Leon Tsomides, and Dave Courtemanch; Maine DEP

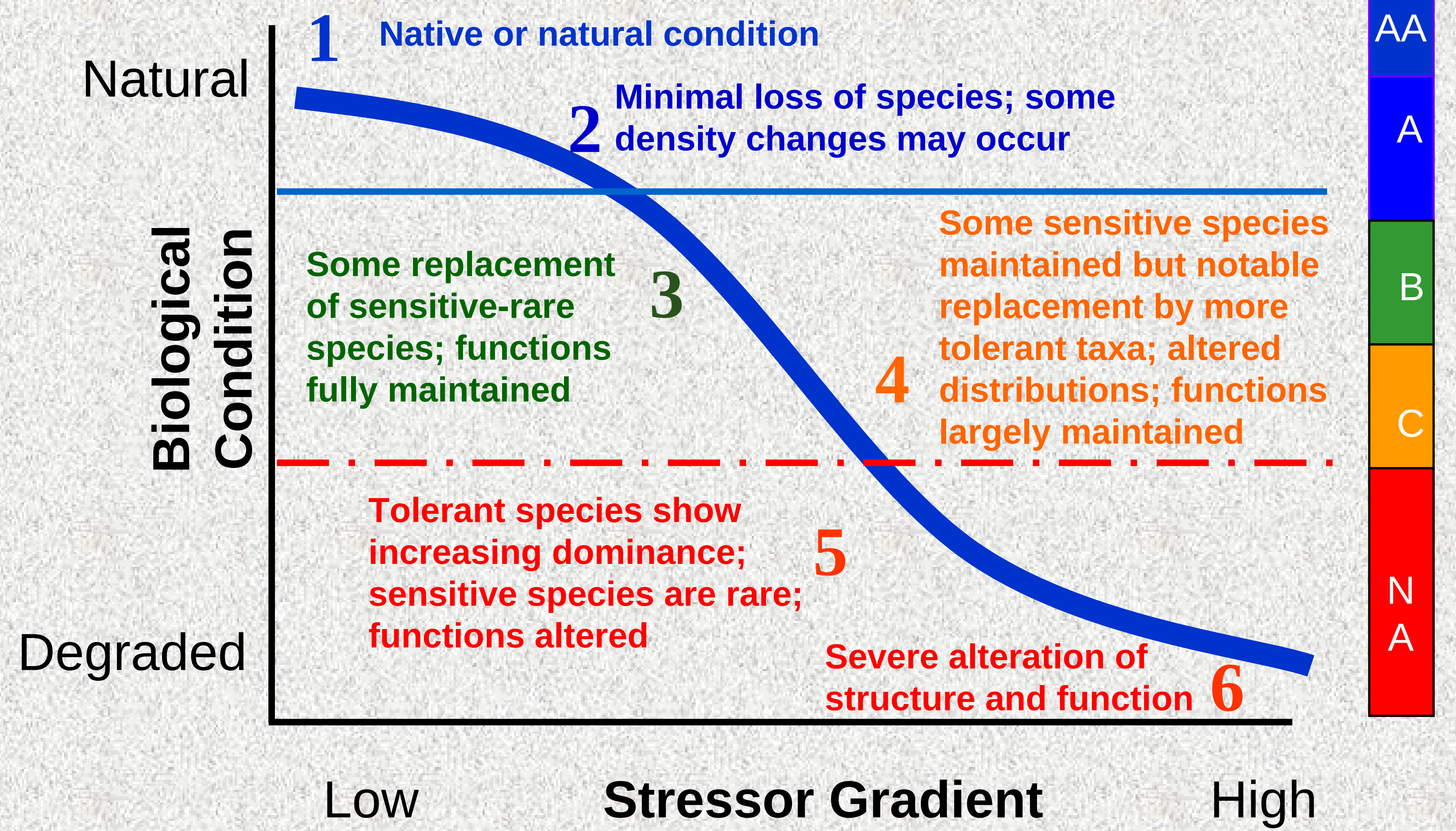
Outline

- Maine's Water Classification System
- Macroinvertebrate Sampling Methods
- Linear Discriminant Models
- Advantages and Considerations

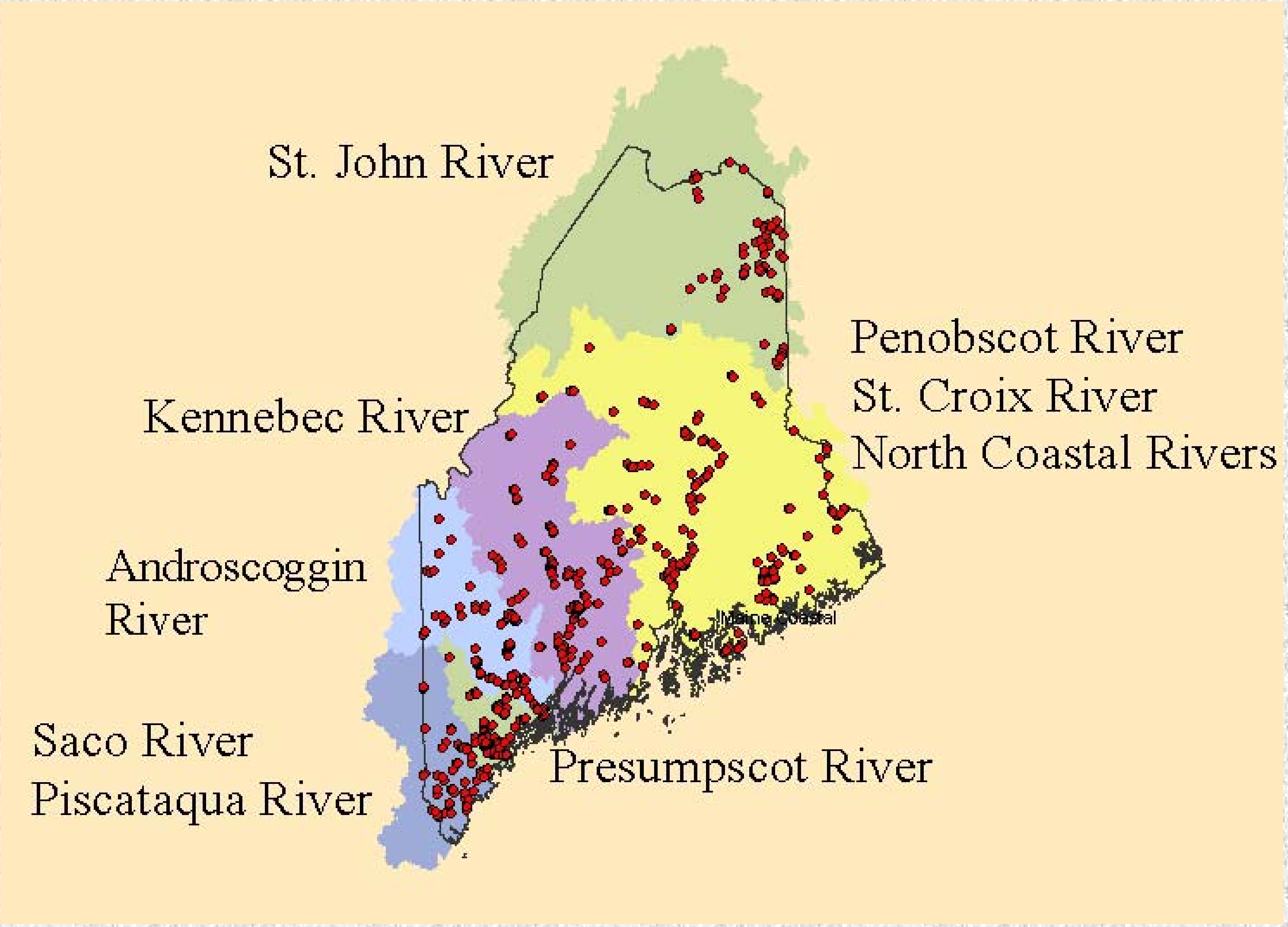
Maine's Water Classification System for Rivers and Streams

- Classes A and AA (treated same for aquatic life use)
  - Aquatic life shall be as naturally occurs.
- Class B
  - no detrimental changes in the resident biological community
  - maintain all indigenous species
- Class C
  - maintain structure and function of resident biological community
- Non-attainment (**NA**)
  - does not meet minimum criteria

Tiered Aquatic Life Use Support (TALUS)



Sampling Stations



Sampling Methods

- Rock bags or baskets
  - Standard volume of cobble
- Usually 3 replicates
- Placed in riffle or run of wadable stream or river
- Left in stream for 4 weeks to allow macroinvertebrates to colonize rocks
- Standard sampling window between July and September

Sampling Methods for Deep Rivers

- 3 or 4 cones filled with standard amount of rocks.
- Cones have attached rope and buoy to facilitate retrieval.
- During retrieval, staff slide a "hat" down the rope to cover cone during retrieval and minimize loss or organisms.
- Divers help retrieve cones if problems arise.

Sampler Retrieval

- Sampler collected with D-frame dipnet to avoid losing critters
- Sampler emptied into sieve bucket
- Sampler and rocks are cleaned inside bucket to remove macroinvertebrates and detritus
- Macroinvertebrates are picked from detritus in the lab

Data Manipulation

- Subsampling and identification
  - <500 individuals - all individuals identified
  - >500 individuals - subsampling is allowed (e.g., 1/2, 1/4)
- Level of taxonomic identification
  - 88% of taxa identifications have been to genus or species
  - 12% of taxa identifications have been to a higher taxonomic level because of early instar or damaged specimens.
  - Taxa counts from replicates are averaged
- Taxa counts are standardized to genus level before model variables are calculated

Development of Linear Discriminant Models

- In 1999, DEP biologists assigned 376 blind samples to one of four *a priori* groups -
  - Class A (n = 120)
  - Class B (n = 117)
  - Class C (n = 72)
  - Non-attainment (NA) of minimum criteria (n = 67)
- DEP biologists included Dave Courtemanch, Susan Davies, and Leon Tsomides
- Assignment of samples was based on abundance, richness, community structure, and ecological theory.

Consistency of *a priori* Assignments

- Consistency of MDEP biologists
  - 96% of independent assignments were unanimous OR majority agreement (2 out of 3)
- Three non-MDEP biologists independently assigned *a priori* classes to samples
  - 80% of independent assignments concurred with MDEP biologists' consensus assignments
- Interpretations did not differ by more than one class in either direction

Development of Linear Discriminant Models

- LDMs are multivariate predictive models that use biological variables to predict a new sample's probability of membership in the four *a priori* groups (A, B, C, & NA).

- For example,
  - Given a set of biological variable values, what is the probability that a sample belongs to the Class A group?

Series of Four
Linear Discriminant Models

First Stage Model
(4-way test)
A vs. B vs. C vs. NA

"C or Better" Model
(2-way test)
A/B/C vs. NA

"B or Better" Model
(2-way test)
A/B vs. C/NA

"A" Model
(2-way test)
A vs. B/C/NA

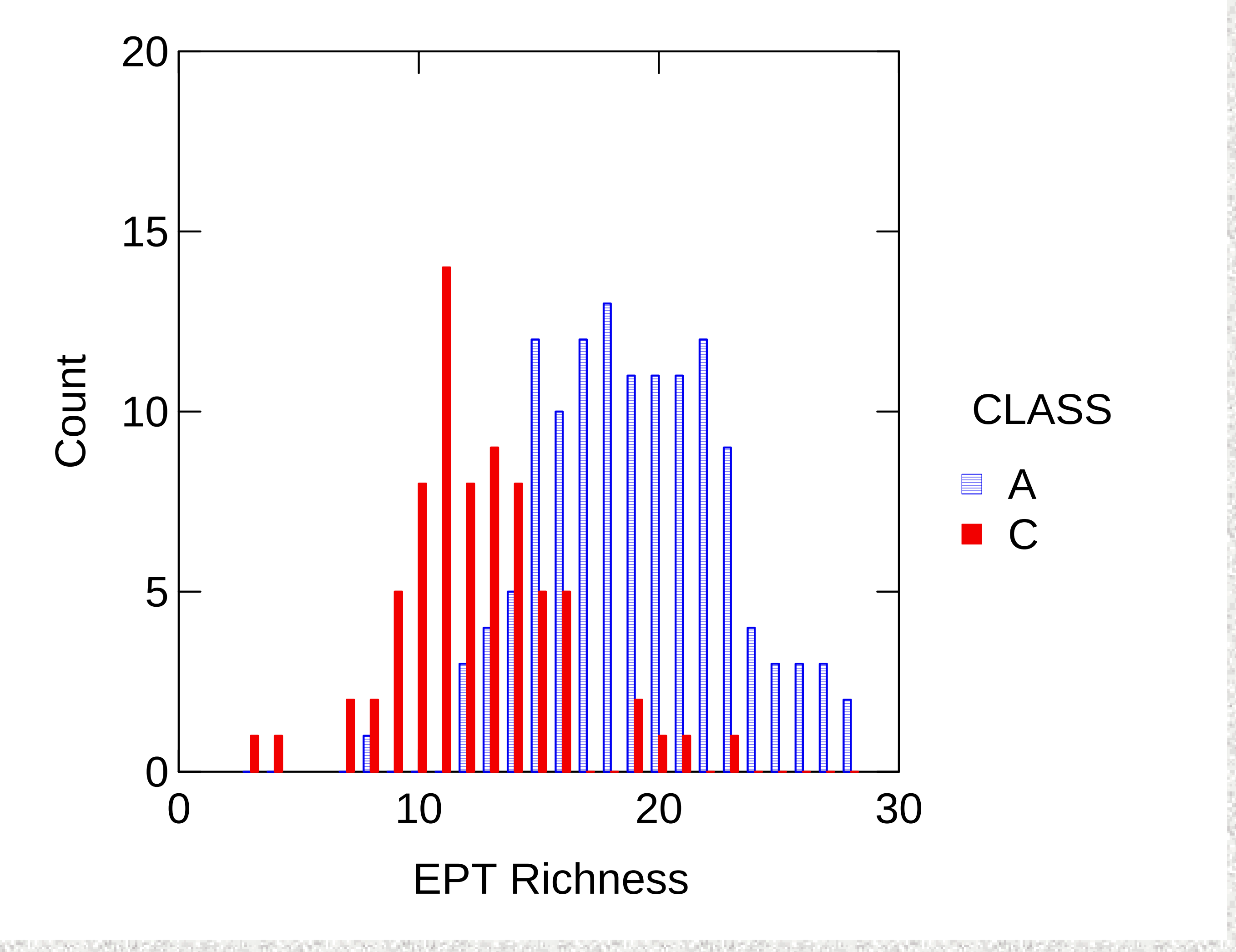
* Aquatic life use attainment decisions are based on the three 2-way tests.

First Stage Model (4-way test)

- Example: 0.30 A, 0.54 B, 0.16 C, 0.00 NA
  - Based on 9 variables
    - Total Abundance of Individuals
    - Generic Richness
    - Plecoptera Abundance
    - Ephemeroptera Abundance
    - Shannon-Weiner Diversity
    - Hilsenhoff Biotic Index
    - Relative Abundance of Chironomidae
    - Relative Generic Richness of Diptera
    - *Hydropsyche* Abundance

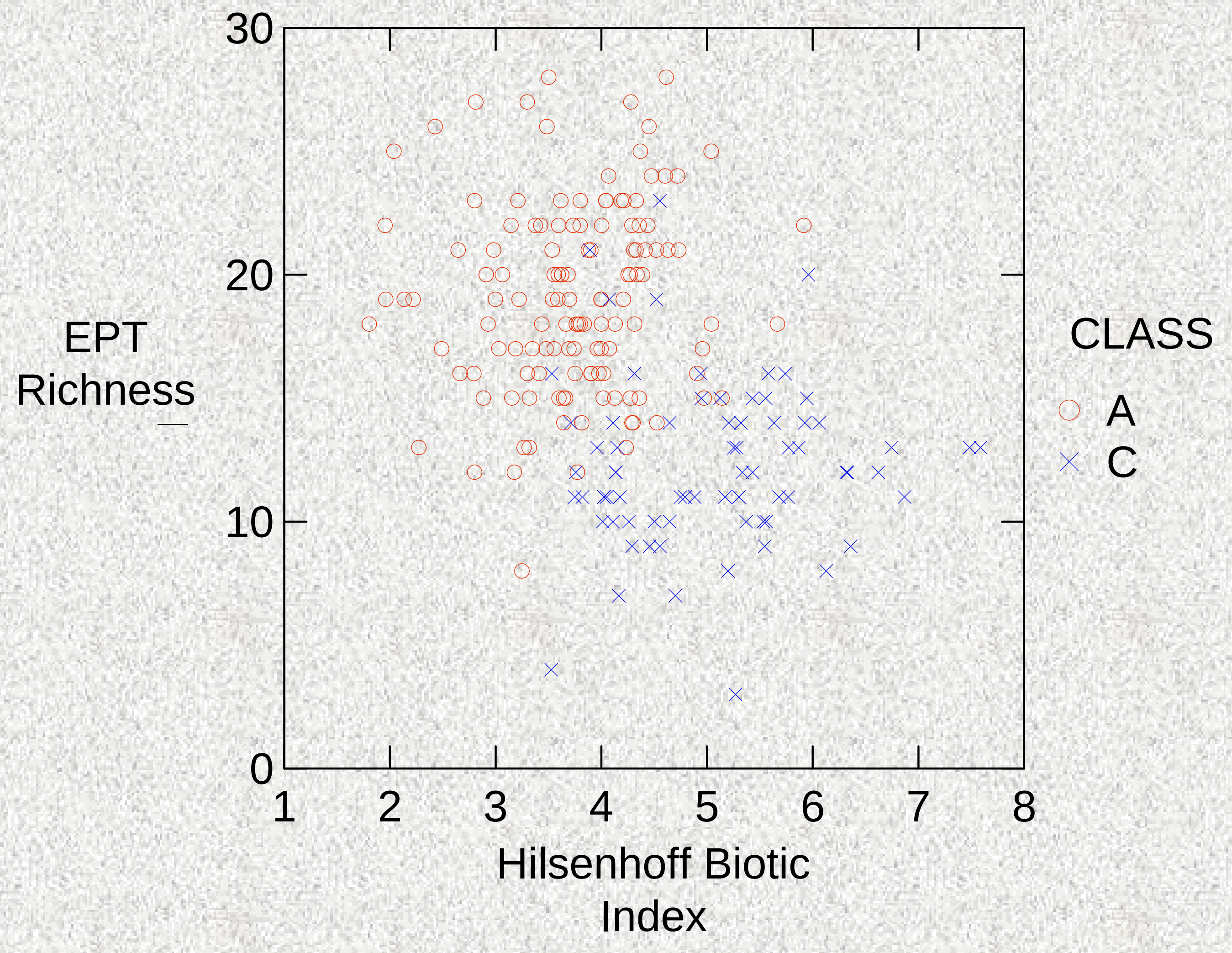
Advantages of Multivariate Analysis

Separation of Class A and Class C samples using 1 variable.



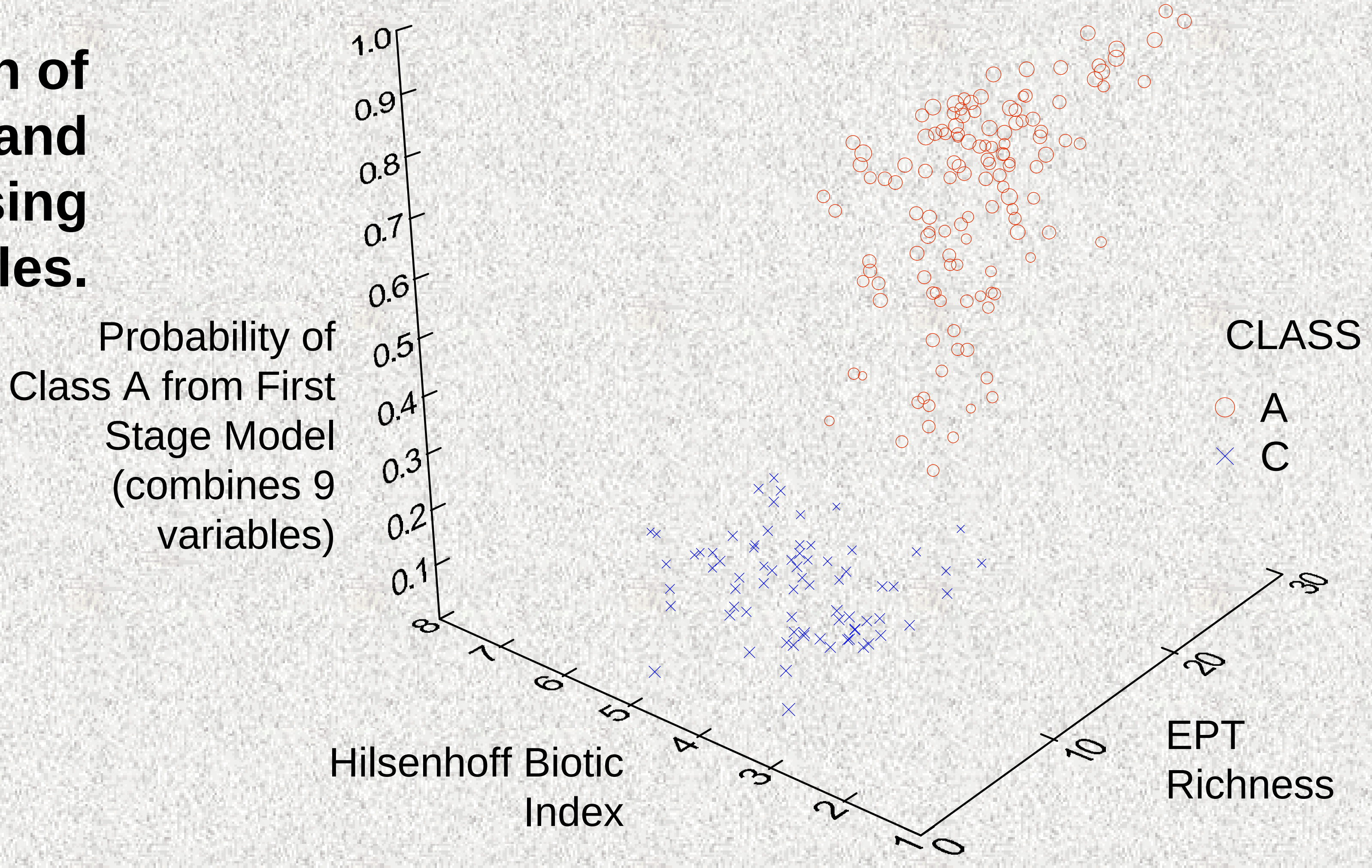
Advantages of Multivariate Analysis

Separation of Class A and Class C samples using 2 variables.



Advantages of Multivariate Analysis

Separation of Class A and Class C using 11 variables.



“C or Better” Model (2-way test)

- Example: 1.00 A/B/C 0.00 NA
  - Based on 4 variables
    - Probability A+B+C from First Stage Model
    - *Cheumatopsyche* Mean Abundance
    - EPT Richness / Diptera Richness
    - Relative Oligochaeta Abundance

"B or Better" Model (2-way test)

- Example: 0.99 A/B 0.01 C/NA
  - Based on 7 variables
    - Probability A+B from First Stage Model
    - Perlidae Mean Abundance
    - Tanypodinae Mean Abundance
    - Chironomini Mean Abundance
    - Relative Ephemeroptera Abundance
    - EPT Generic Richness
    - Sum of Mean Abundances of *Dicrotendipes*, *Micropsectra*, *Parachironomus*, and *Helobdella*

“A” Model (2-way test)

- Example: 0.05 A 0.95 B/C/NA
  - Based on 6 variables
    - Probability A from First Stage Model
    - Relative Plecoptera Richness
    - Sum of Mean Abundances of *Cheumatopsyche*, *Cricotopus*, *Tanytarsus*, and *Ablabesmyia*
    - Sum of Mean Abundances of *Acroneuria* and *Stenonema*
    - Ratio EP Generic Richness
    - Ratio of Class A Indicator Taxa (*Brachycentrus*, *Serratella*, *Leucrocuta*, *Glossosoma*, *Paragnetina*, *Eurylophella*, and *Psilotreta*)

Results of Linear Discriminant Models

First Stage Model
0.30 A   0.16 C
0.54 B   0.00 NA

"C or Better"
1.00 A/B/C
0.00 NA

"B or Better"
0.99 A/B
0.01 C/NA

"A"
0.05 A
0.95 B/C/NA

* Based on p=0.60 threshold, result is Class B.

Model Performance

Class A Model			
		Model Prediction	
		A	B,C,NA
A Priori	A	**87%**	13%
	B,C,NA	9%	91%

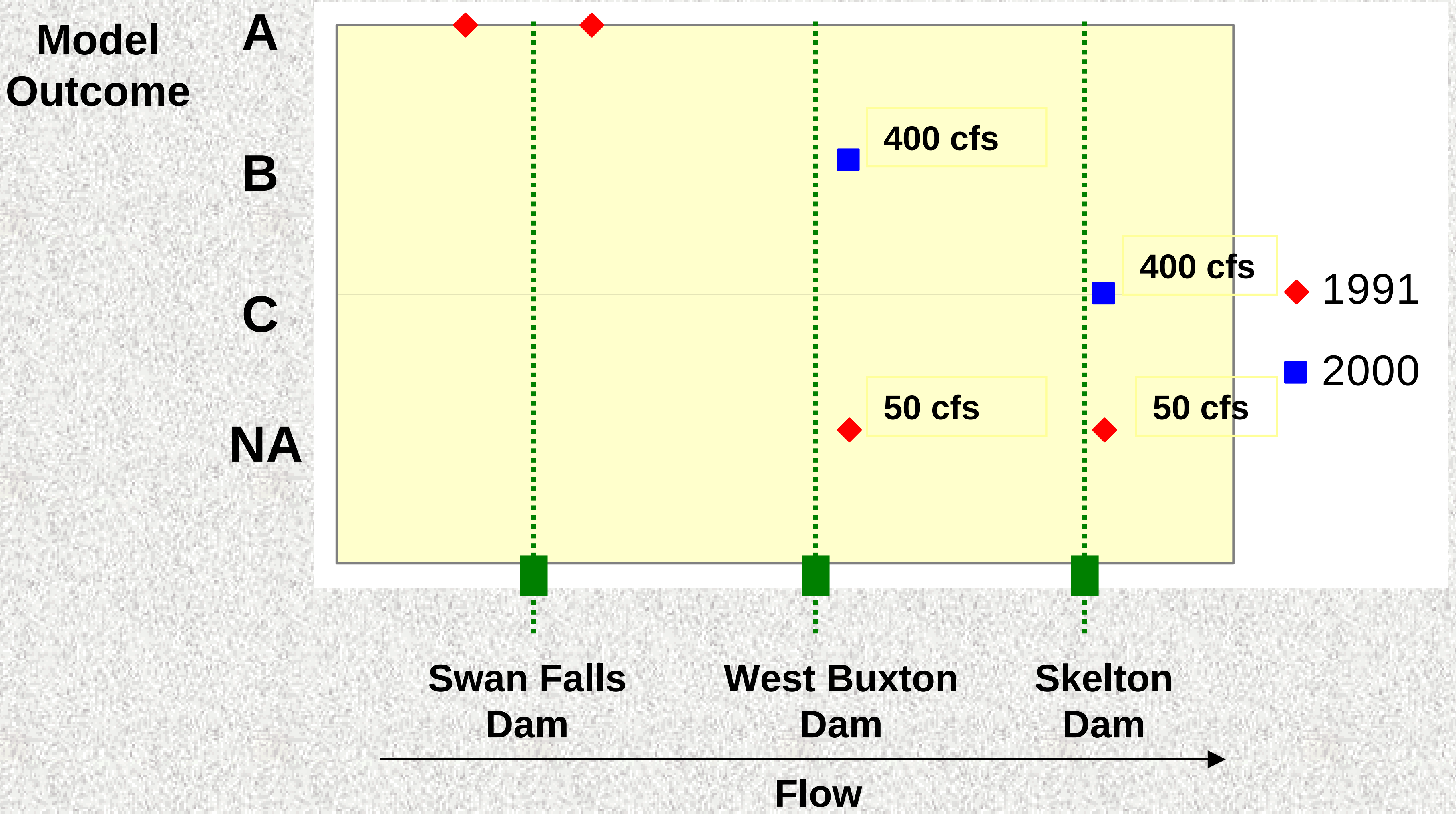
B or Better Model			
		Model Prediction	
		A,B	C,NA
A Priori	A,B	**94%**	6%
	C,NA	6%	94%

C or Better Model			
		Model Prediction	
		A,B,C	NA
A Priori	A,B,C	**96%**	4%
	NA	12%	88%

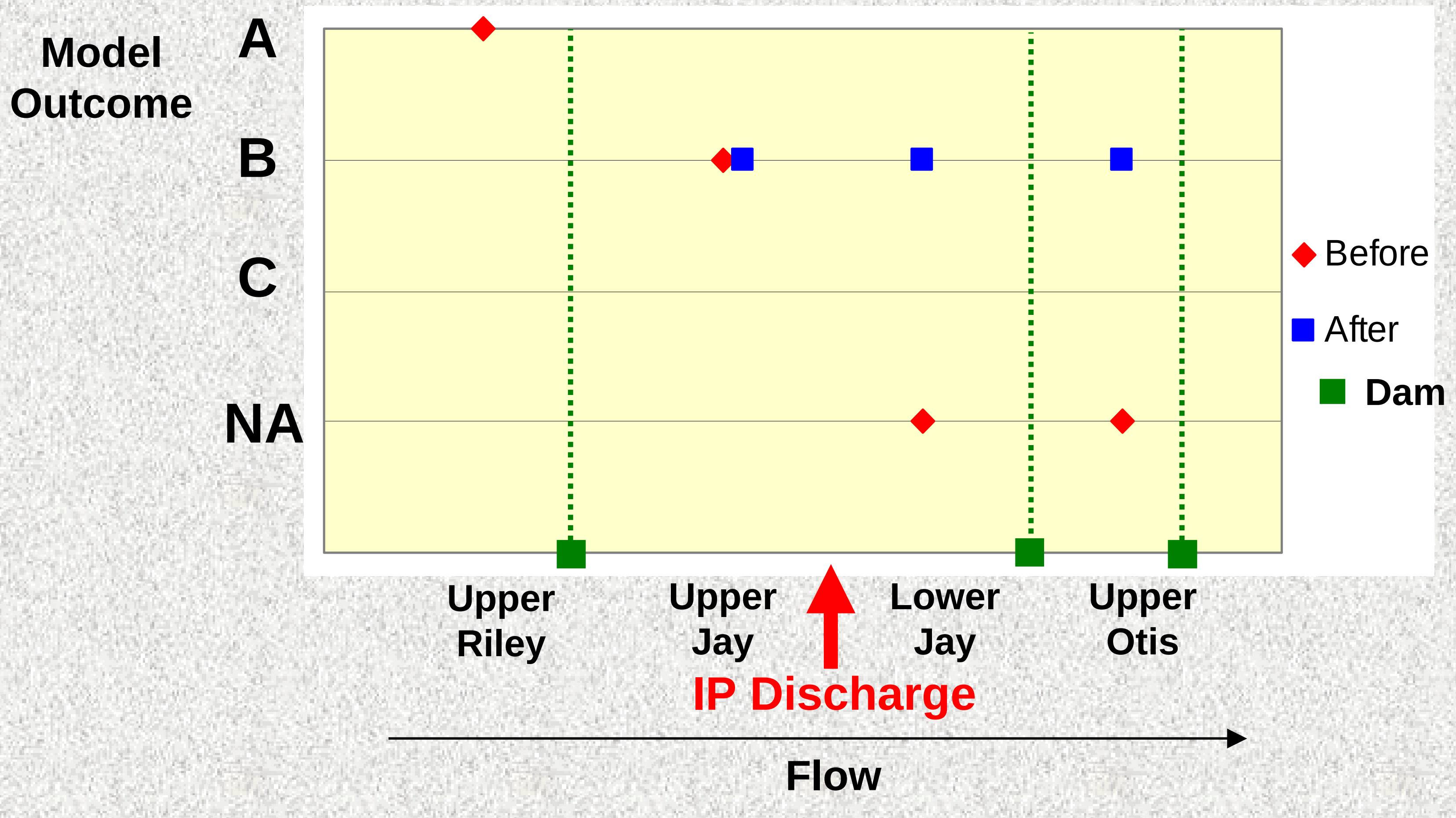
Advantages of Approach

- Direct relationship between model outcomes and aquatic life uses.
  - Translates broad resource goals and objectives to scientifically defensible, quantitative thresholds
- Based on ecological theory and demonstrated to reflect changes in resource condition.
- Statistically based with known probability of error.

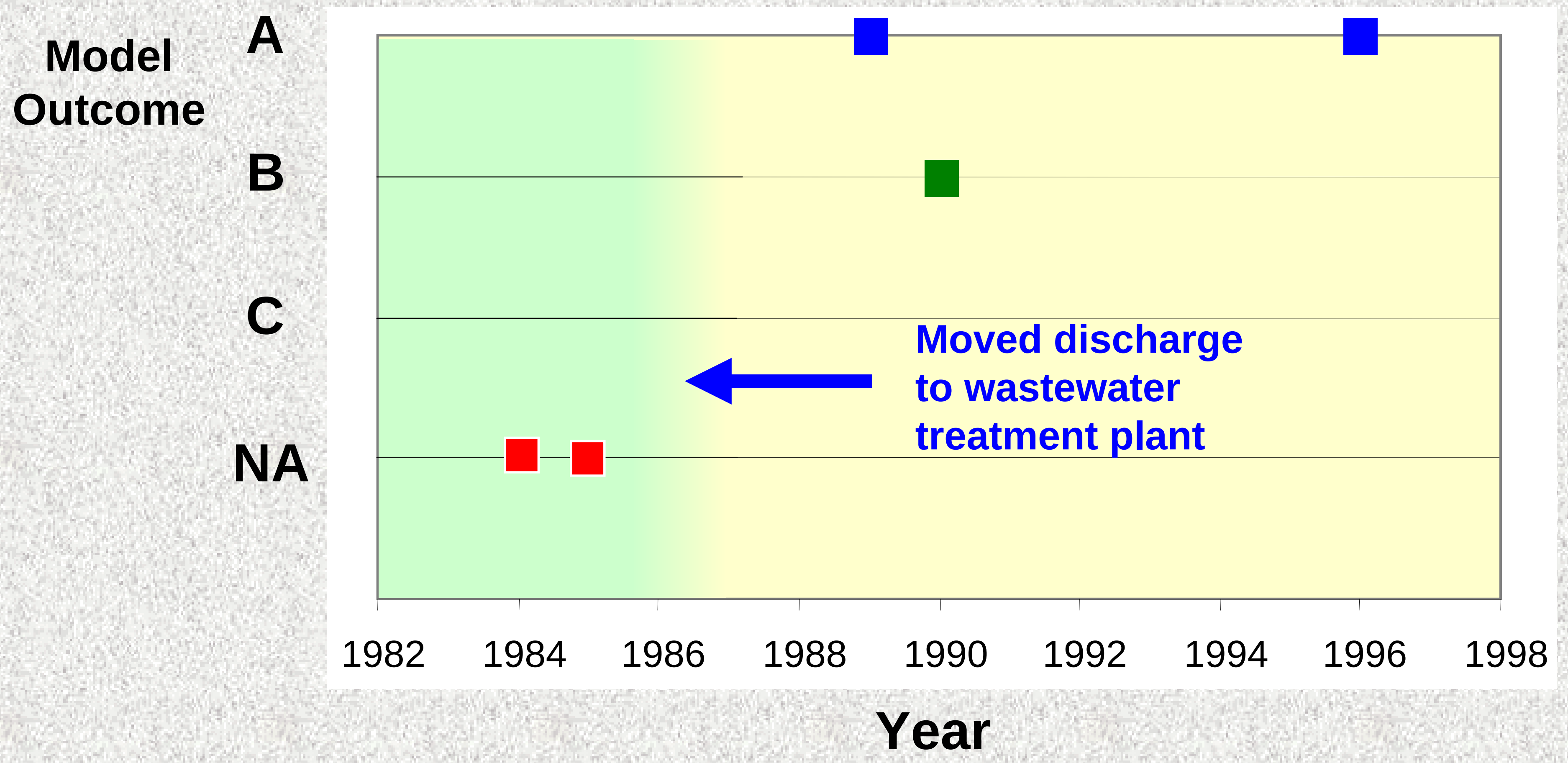
Effects of Increasing Flow below Dams on the Saco River



Effects of Removing TSS Discharge on Androscoggin River Impoundments



Reducing Discharges from Guilford Industries into Piscataquis River



Considerations of Approach

- Process of assigning *a priori* classes requires experienced biologists
  - but classification steps in developing multimetric indexes and predictive models also greatly benefit from having experienced biologists
- Requires periodic recalibration as number of samples in database increases.
- Possible circularity based on *a priori* classification
  - Do Class A model outcomes represent minimally-disturbed reference conditions?

Does the model accurately classify minimally disturbed streams?

- **27** samples were selected with following criteria:
  - not used to build the model
  - no known point sources
  - average % of upstream watershed
    - 94% forested
    - 3% logged
    - 2% crop
    - 1% residential
    - <1% urban/industrial/commercial
- **24 (89%)** of samples had model outcomes of class A

For More Information

- Biomonitoring Web Site
  - http://www.state.me.us/dep/blwq/docmonitoring/biomonitoring/index.htm
- Methods Manual
  - http://www.state.me.us/dep/blwq/docmonitoring/finlmeth1.pdf
- Fifteen Year Retrospective
  - http://www.state.me.us/dep/blwq/docmonitoring/biomonitoring/biorep2000.htm
- E-mail
  - biome@maine.gov