

Recent Approaches to Freshwater Biomonitoring: A Comparison of Methods Used

Vincent H. Resh

Department of Environmental Science,
Policy & Management

University of California, Berkeley CA 94720-
3112 USA

E-mail: vresh@nature.berkeley.edu



Types of Monitoring Involving Benthic Macroinvertebrates

- Biodiversity or Ecosystem Health Monitoring
- Impact Monitoring
- Target-Taxa Monitoring
- Compliance Monitoring
- Baseline Monitoring
- Effectiveness Monitoring
- Monitoring to provide Feedback for Adaptive Management

Diversity of Approaches in Use for Benthic Macroinvertebrate Biomonitoring

*“There are as many sampling devices as
there are aquatic ecologists!”—1960s*

- Different sampling methods provide different information
- Different measurements provide different information
- Spatial temporal complexity requires adaptation
- Regulatory laws and mandates vary



Potential Criteria to Evaluate the IDEAL Biomonitoring Tool

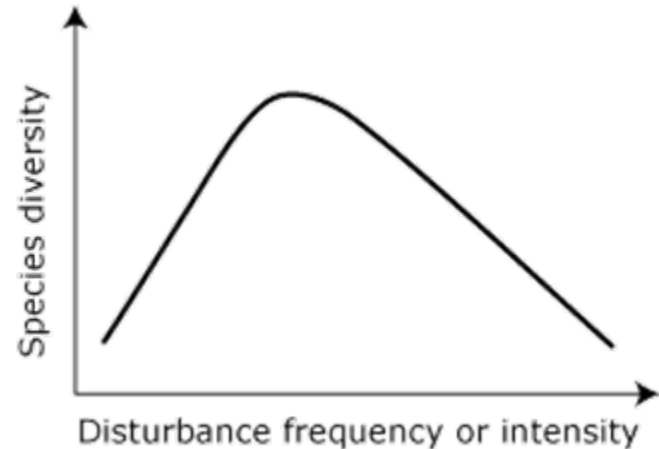
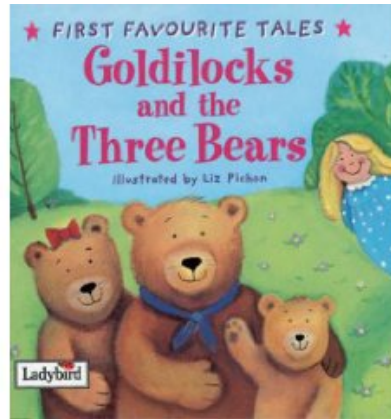
- *Conception*
How did it come about?
- *Realization*
How effectively can it be applied?
- *Performance*
How good does it actually work?



Conception

1. Is it derived from sound theoretical concept?

Intermediate Disturbance Hypothesis



2. Is it *a priori* predictive?

3. Can it be used to assess ecological function?

4. Can it indicate overall human impact?

5. Can it discriminate among different types of human impact?



Realization

How practical is it to actually use this particular biomonitoring tool?

6. Are there low costs for sampling and sorting specimens? Is standardized sampling required, and is it expensive?
7. Is sampling easily done?
8. Are there low costs for taxonomic identifications needed?



Performance

9. Is it applicable across spatial scales
(e.g. all of North America, Europe) with the least
possible regional adaptation?
 10. Does it *reliably* indicate overall impact?
 11. Does it *reliably* indicate specific impacts?
 12. Does it indicate human impact
on a linear scale?
- *Therefore, 12 criteria in three categories will be
used in evaluations...*



What BIOMONITORING TOOLS are available for benthic macroinvertebrates?

- Bioassays
- Biomarkers
- Biotic indices
- Fluctuating Asymmetry
- Multimetric approaches
- Multivariate approaches
- Functional feeding groups
- Multiple biological traits
- Benthic secondary production
- Leaf-litter decay rates

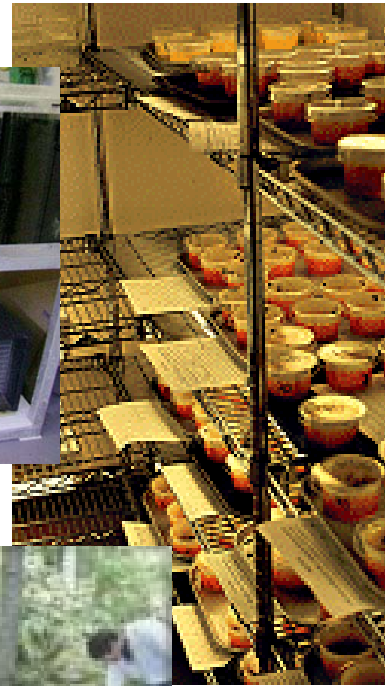
HIERARCHY OF BIOMONITORING APPROACHES

- *SUBORGANISM*: Biomarkers
- *ORGANISM*: Bioassays
- *POPULATION*: Fluctuating asymmetry
- *COMMUNITY*: Biotic indices, multimetric and multivariate approaches, functional feeding groups, multiple biological traits
- *ECOSYSTEM*: Benthic secondary production and Leaf-litter decay

BIOASSAYS--toxicity tests of survival, growth, feeding, behavior or reproduction

- They provide a dose-dependent response to acute or chronic exposures
- Vary depending on reason for testing (toxicity of water or sediment, bioaccumulation), indigenous or standard species, single or multiple species, field or laboratory

Bioassays Come in a Variety of Sizes and Shapes



Conception of Bioassays

- Based on niche theory (suboptimal or outside of niche)
- Predictions are possible ($>$ toxicant, $<$ growth)
- Extrapolation to ecosystem function rare because of monospecific tests; however, decreased feeding rate could reflect decreased processing of detritus in ecosystem
- Appropriate indicators of overall impairment
- Discrimination of different types of pollution not done (but possible if a battery of different species and pollution types were integrated)

Realization of Bioassays

- Low costs for sampling and no standardized sampling needed; may have variable costs (depending on lab or field; standard culture or field collected test organisms)
- Prior knowledge of sensitivity and response measured is essential
- Identification costs low



Performance of Bioassays

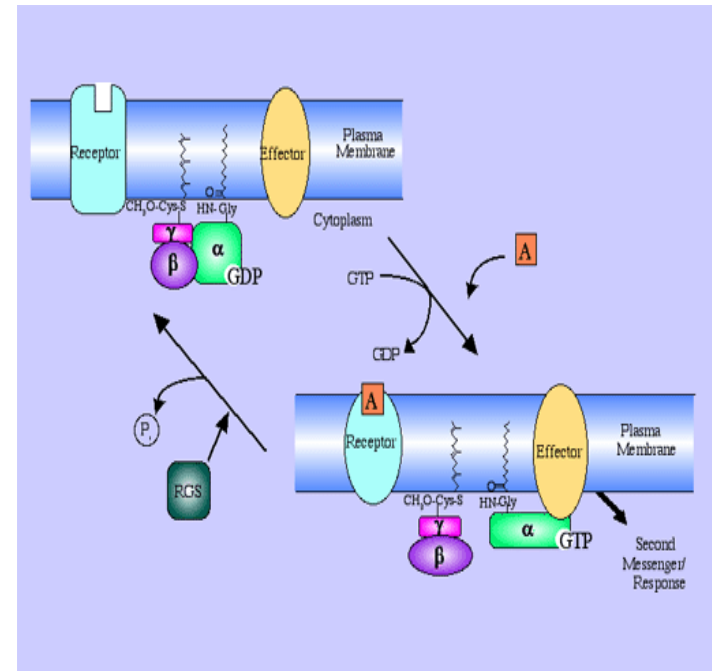
- Large-scale applicability depends on how bioassay is conducted
- Reliability of overall impact unknown—non-toxic impairment not tested
- Reliability of different types of impact is good (but may reflect intrinsic or extrinsic factors)
- Relationships are often linear
- *At best, bioassays match 10 of 12 criteria!*



Biomarkers

- Biochemical measures that indicate sub-lethal responses to toxicants at the molecular, cellular or tissue level

- Mixed-function oxidases, stress proteins, acetylcholinesterase, ion regulation callulase/carbohydrase
- Heavy metals, >HSP, a stress protein that alters lysosomal membrane stability
- Problem: responses aren't necessarily the same at higher organization levels



Conception of Biomarkers

- Biomarkers are based on correlations and experimental tests; not predictive or based on theoretical concepts
- Cannot currently indicate ecological functions
- Low ability to detect overall impact
- High ability to discriminate among different types of impact

Realization of Biomarkers

- Sampling costs are low; no standardized sampling required; lab costs high
- Number of biomarkers selected depends on ecological information available for a site
- Identification costs generally low



Performance of Biomarkers

- Large-scale application unknown (likely to occur but genetic differences and environmental factors like temperature may affect results)
- Reliability to detect overall impact or specific impacts unknown
- Both linear and non-linear patterns observed
- *At best, biomarkers match 6 of 12 criteria!!!*

Biotic Indices

- Oldest is the Saprobian System used in Germany to indicate oxygen deficits >100yrs ago
- Essentially all Biotic Indices are calculated as the sum of Abundance x Tolerance for each taxa, divided by abundance of the total fauna
- Tolerance values are derived for species (in Europe), genera (in North America), families elsewhere
- Extrapolation of values to other areas is questionable

Conception of Biotic Indices— Saprobien System of Germany

- Derive from the Niche Concept (optimum and breadth)
- Derived from observations; not predictable (only one dimension--oxygen, acidity, etc.)
- Cannot assess overall ecological function
- Cannot assess overall impact
- Cannot assess different types of human impact

Realization of Biotic Indices

- Sampling protocols can be complicated or simple
- Standardized sampling is required and costs are high
- Species level identifications are most useful but make identification costs high
- Family level identifications decrease power of index



Performance of Biotic Indices

- Not applicable across broad geographic areas (species occurrence differs)
- Not applicable to different impact types
- Not always applicable to organic pollution; index relies on one niche dimension (oxygen), which is influenced by temperature and flow
- Indices are biased toward high mountain stream
- No linear scale (1000x more oxygen used by cleanest water organisms)
- *At best, biotic indices match 1 of 12 criteria!!!*

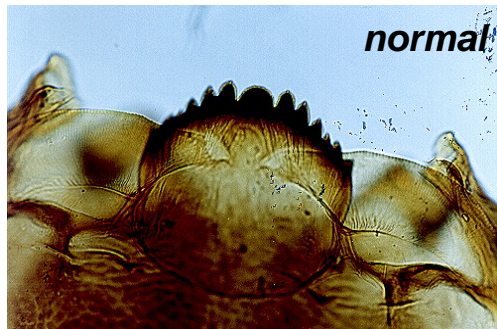


FLUCTUATING ASYMMETRY—Small random alterations in bilateral, morphological symmetry

- Assumed to be the result of human-induced disturbances
- Assessed by measurements of
- Useful morphological features include
- Used with mayflies, damselflies, water striders, midge larvae, and caddisflies

Examples of deformities

Macroinvertebrates



Chironomidae (DIPTERA)

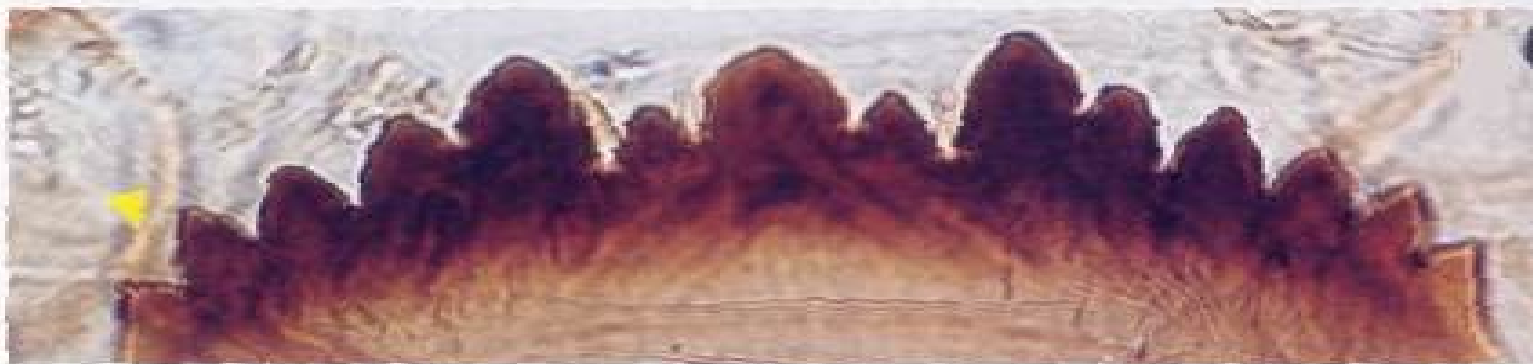
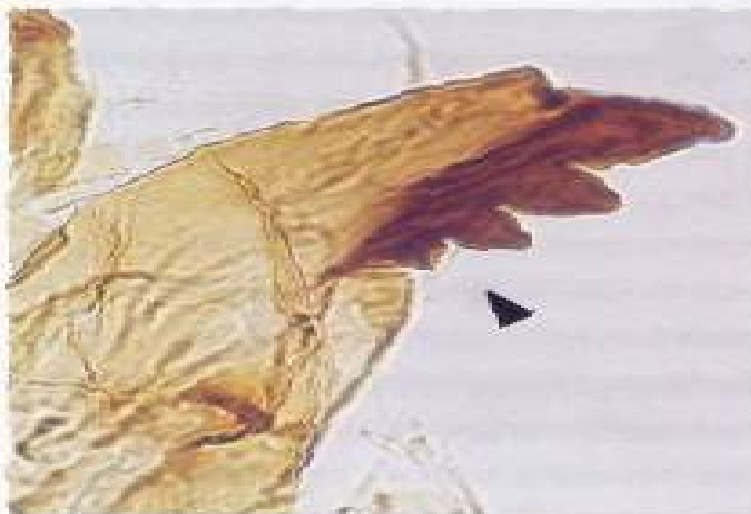
Amphibiians



Fish



Serviá, 2001



Conception of Fluctuating Asymmetry

- Based on the Developmental Instability Concept
- Enables *a priori* predictions
- Cannot assess ecological function
- Viewed as related to overall human impact
- Does not discriminate types of impact

Realization of Fluctuating Asymmetry

- Low costs for sampling, no standardization required
- Uses abundant taxa so identification costs tend to be low
- Higher costs associated with search for fluctuating asymmetry characters and need for relatively large samples to be examined

Performance of Fluctuating Asymmetry

- Unknown as to whether it can be applied across ecoregions
- Unreliable indicator of overall or specific impact (temperature, biological interactions, genetic heterozygosity)
- Non-linear scale
- *At best, Fluctuating Asymmetry matches 6 of 12 criteria!!!*

Multimetric Approaches

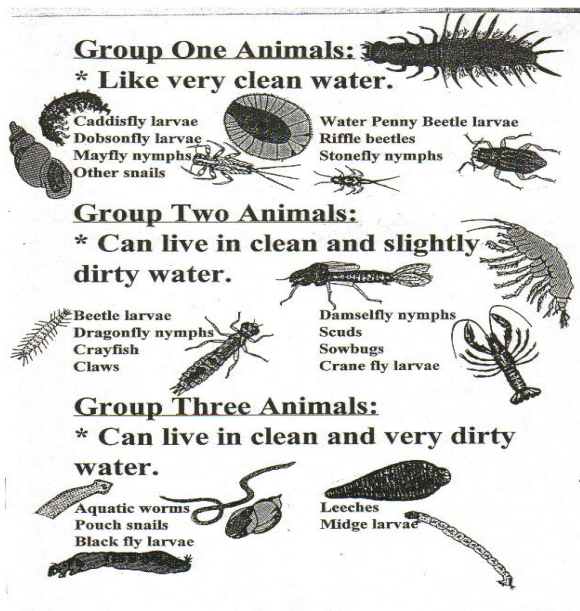
- Metrics are calculated measures that represent some aspect of the structure, function, or other characteristic of the community that change with human impact
- Multimetric indices combine metrics that, presumably, represent a range of macroinvertebrate assemblage responses
- Like economic indices; most widely used by US state agencies
- Approach involves: (1) selection and calibration of metrics applicable to homogeneous sites: (2) determination of threshold levels

Conception of Multimetric Approaches

- Some metrics are based on ecological theory (e.g. Functional Feeding Groups and RCC, but not taxa richness and IDH) but most are based on observations (e.g. EPT)
- Some are *a priori* predictive
- Only FFGs assess ecological function
- They can discern overall impact and some metrics can determine specific impacts

Realization of Multimetric Approaches

- Some protocols have low cost, others not
- Standardized sampling is required
- Taxonomic costs vary depending on levels used (family vs. generic)



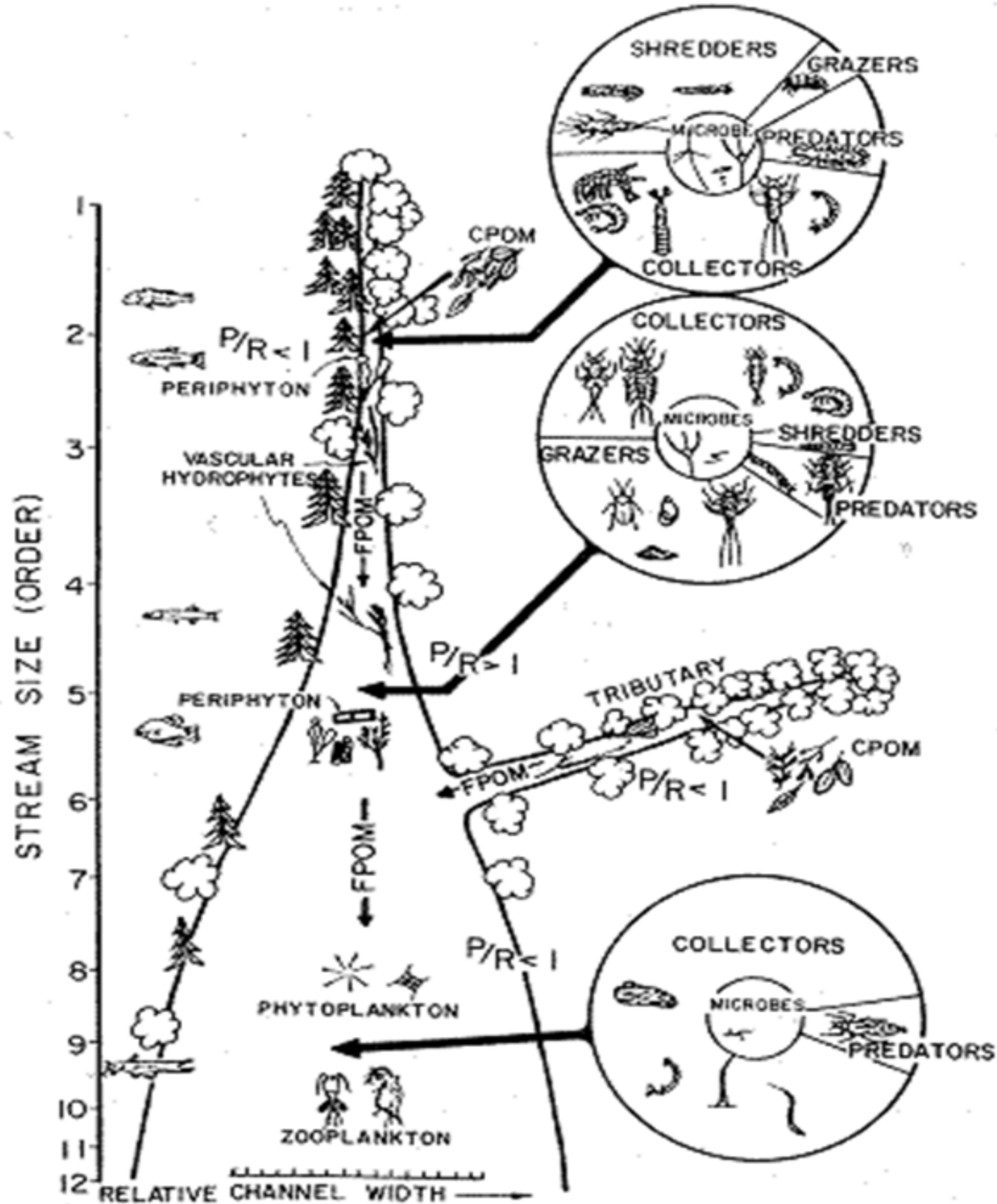
Performance of Multimetric Approaches

- Use of homogenous site groupings does not have large-scale applicability
- Proponents indicate reliable indications of overall human impact; natural variability and regional variation in metrics limits spatial and temporal applicability
- Individual metrics reliably respond to specific impacts
- Not on linear scale because of thresholds
- *At best, multimetric approaches match ~ 10 of 12 criteria!!!*



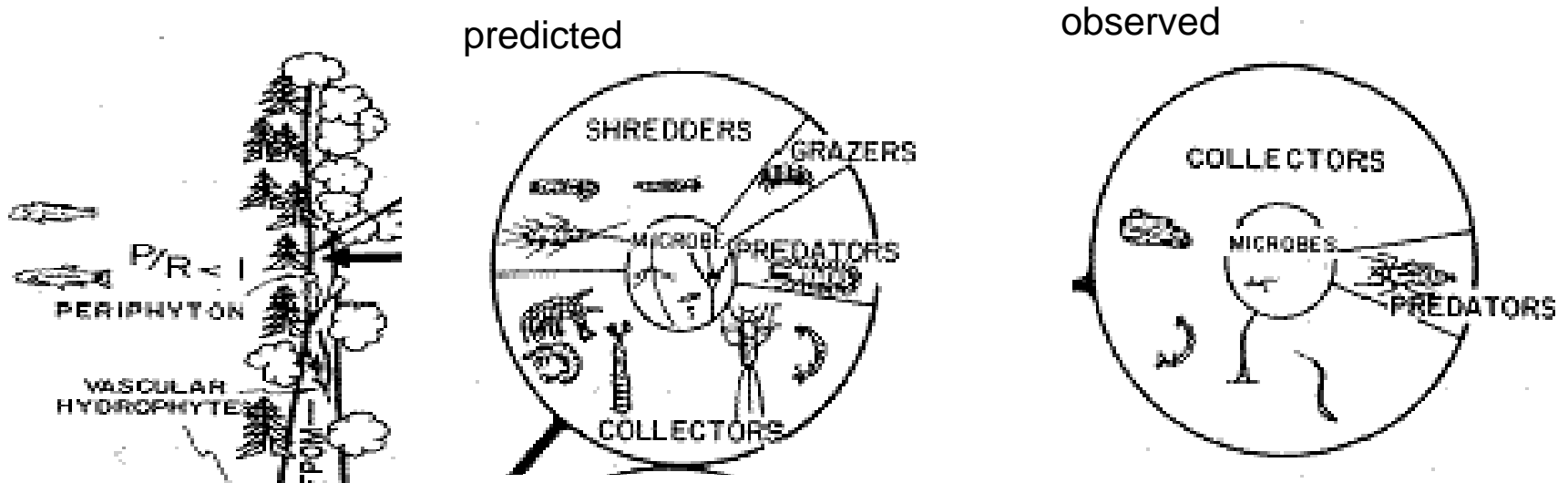
Functional Feeding Groups

- Assignment is based on the type of food source used and the feeding mechanism involved
- Originally intended to be based on mouthpart morphology (reflecting preferred food) not gut contents
- Abundance (% of total) of single groups, ratios between groups, or Index of Trophic Completeness used in assessments
- Based on the River Continuum Concept



Conception of Functional Feeding Groups

- Based on River Continuum Concep
- A priori predictable, can assess ecological function, not discriminate overall impact but rather types of impact



Realization of Functional Feeding Groups

- Low costs for sampling and sorting
- Sampling does not have to be standardized but variation in methods may effect groups collected
- Lower taxonomic levels are preferable but species to family levels are used
- Difficulty in feeding group assignments, diets may shift over life cycles, and omnivory is common among benthic macroinvertebrates



Performance of Functional Feeding Groups

- Applicability across ecosystems is difficult; ffg composition may vary naturally
- Limited reliability to distinguish overall impact, weak discrimination of different types of impact
- Non-linear response along pollution gradient
- *At best, functional Feeding Groups match 8 of 12 criteria!!!*

Multiple Biological Traits (Species traits)

- Size, body form, life cycle, food and feeding habits, reproductive traits
- Most research done on European species
- Aquatic macroinvertebrates can be compared worldwide because of same scale used for traits
- However, consistent description of traits on same scale is difficult

Examples of Traits

- functional feeding-group measures
- % organisms with more than one generation per year
- % long-lived taxa
- % long-lived predators

Conception of Multiple Biological Traits

- Based on Habitat Templet Concept
- Traits confer resistance or resilience to disturbance
- *A priori* predictable, can discriminate overall and specific types of human impact
- Some traits relate to ecological function

Realization of Multiple Biological Traits

- Presence-absence data, and generic level identifications reliably describe trait composition
- Low costs for sampling; no standardization
- Identification costs low because generic keys are available
- Trait Profile available for most European and North American genera of aquatic invertebrates

Performance of Multiple Biological Traits

- Traits stable across France and regional stream types of Europe
- Traits indicate overall human impact and some specific impacts
- Scale of pollution indication is unknown
- *At best, multiple Biological Traits match 10 of 12 criteria!!!*

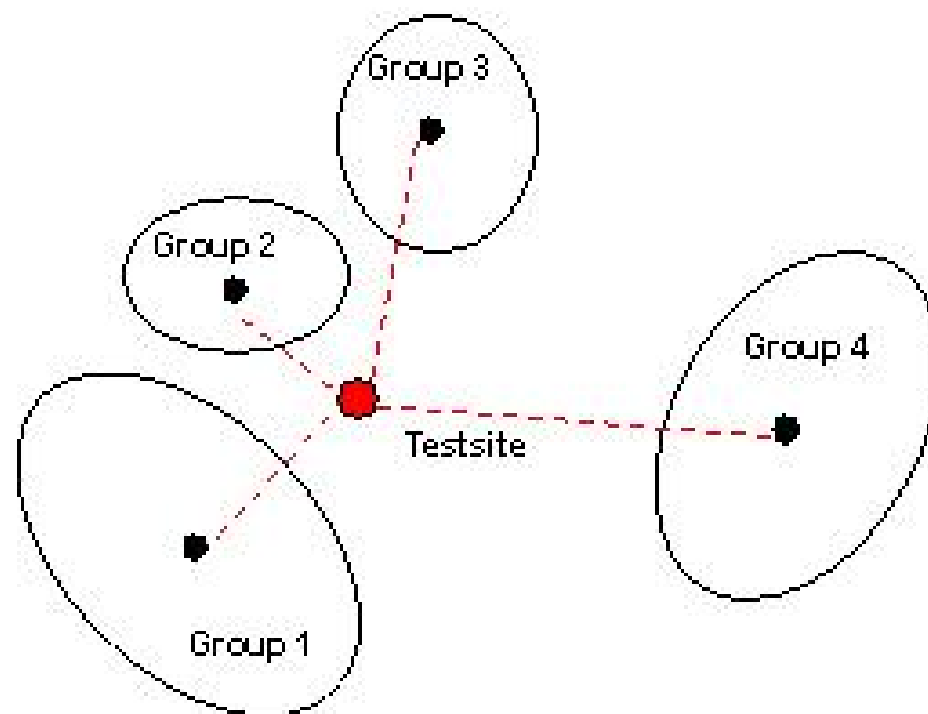
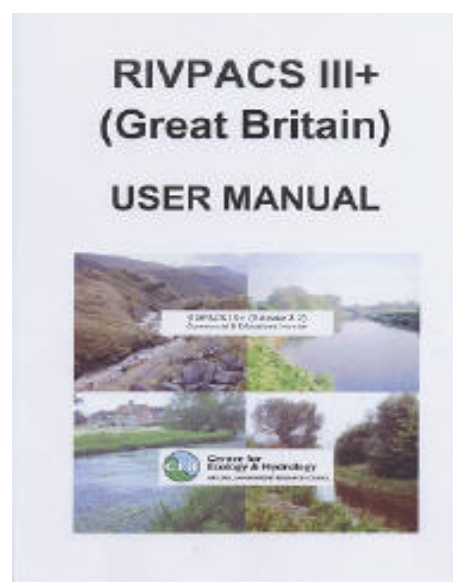


Multivariate Approaches

- Use statistical analyses to compare patterns observed at test and reference sites
- Reference Condition Approach compares test sites with condition representative of a GROUP of minimally disturbed sites
- Selected (usually) by expert opinion; not always available (e.g. industrial areas)
- RIVPACS, AUSRIVAS, BEAST, ANNA are currently used models (the data are really the model!)

How Multivariate Approaches Are Used in Assessments

- Based on predictive, correlative models that compare communities in test sites with reference types
- Reference sites are classified into groups based on faunal composition; correlations with natural environmental variables are made; test sites are assigned to groups based on their environmental characteristics and the faunas compared (e.g. as Observed/Expected species composition)





Conception of Multivariate Approaches

- Approach ~ based on Niche Concept
- *A posteriori* predictive but richness may be *a priori* predictive
- Can evaluate overall human impact but (at least now) do not assess ecological function or specific types of impact



Realization of Multivariate Approaches

- Low costs for sampling and sorting; only sub-samples are usually sorted from timed or set-area sample
- Sampling is standardized
- Identifications range from species to families (depending on richness of families and exclusion of taxa with high costs for identification)



Performance of Multivariate Approaches

- Rely on homogenous reference sites with similar biota; large-scale applicability questionable; adapted to different regions and stream types of UK
- Misclassifications range from 0-66%
- Linear relationship for O/E but non-linear responses in pollution tolerance as well.
- At best, multivariate Approaches match 8 of 12 criteria!!!

Ecosystem-level Processes

- Benthic Secondary Production--The accumulation of biomass produced over time; it links populations and communities to ecosystem-level processes
- Few studies thusfar, but at best *matches 6 of 12 criteria!!!*
- Leaf-litter decay—Measurements of rates that examine functional integrity of streams
- *At best, matches 10 of 12 criteria!!!*

Number of 12 Criteria Met

Lowest-Highest number met

- Biotic Indices **1**
- Biomarkers **3-6**, 3?
- Benthic Secondary Production **3-6**, 5?
- Fluctuating Asymmetry **5-6**, 2?
- Multivariate Approaches **4-8**, 3?
- Functional feeding groups **4-8**, 2?
- Bioassays **4-10**, 2?
- Multiple Biological Traits **10**, 2?
- Multimetric Approaches **3-10**
- Leaf-litter decay **7-10**

? applicability unknown

What is the Ideal Biomonitoring Tool for Benthic Macroinvertebrates?

- No approach matches all 12 criteria; are these criteria for the ideal biomonitoring tool too stringent?

- Which criteria are easiest to meet?

Tools derived from sound theoretical concept in ecology

Tools with low costs for sampling and sorting

- Which criteria are hardest to meet?

Tools with large scale applicability

Tools that are reliable indications of overall and specific impact

Tools indicating impact on a linear scale

Research Needs

- Some need more research before their potential can be realized (Biomarkers)
- ~25% of criteria were scored equivocally—i.e. in some cases they met, but in others they didn't meet criteria
- Managers and society want simple scores or determinations of water quality
- Competition for adoption (and funding) has resulted in formation of “camps” and rigid stances

What Should Be Done by Resource Managers?

- Fiscal consequences of biomonitoring in setting restoration goals are enormous
- Perhaps there is a mismatch between those who create biomonitoring tools and those who must use them for planning restoration
- Societies and government must prioritize which criteria are important
- Clearly, long-term use is not a sufficient reason for continuity of that tool

THANK YOU!!!



Biological Assessment is the process of evaluating the condition of a water body, using biological surveys and other direct measurements of the resident biota (fish, benthic macroinvertebrates, algae, etc.) in lakes and streams

Rapid Bioassessment Protocols (**RBPs**) are an:

- Integrated assessment of habitat (physical structure), water quality, and biological condition
- $\text{Habitat Quality} + \text{Water Quality} = \text{Biological Condition of the Community}$

RBPs provide:

- Cost effective, yet scientifically valid, procedures for surveys
- Opportunities for multiple site investigations
- Quick information for management decisions
- Information readily understood by the public

RBPs can be used to:

- Determine status of water resources
- Evaluate causes of degradation
- Determine effectiveness of control and mitigation plans
- Measure success of management programs

Reference Condition Concept

- Reference conditions can be site specific (e.g. upstream vs. downstream), or
- Regional conditions that are based on a series of unimpaired sites within a uniform region and habitat type
- Multivariate approaches can be used to establishing reference conditions

Advantages of using **benthic macroinvertebrates**

- Good indicators of local conditions
- Integrate effects of short term variations (<1 yr)
- Easy to identify
- Broad range of pollution tolerances
- Broad range of feeding levels
- Sampling is relatively easy
- Food source for fish
- Abundant and widespread occurrence
- Extensive literature

Technical Decisions in Using Benthic Macroinvertebrates

- Choice of sampling universe
- Seasonality of sampling
- Choice of sampling device
- Separating organisms from substrate
- Field or laboratory sorting
- Proportion or number to be examined
- Taxonomic levels for identification
- Percent similarity or statistical analysis
- Other?

Multivariate Approaches

1. Measure physical conditions at unimpaired sites
2. Measure biological conditions at the same sites
3. Create groupings of unimpaired sites based on similar physical characteristics; what organisms commonly occur in these groupings?
4. At test sites, measure physical characteristics to determine appropriate reference-site grouping; do organisms of test site agree with reference group? If not, impairment may have occurred

Other Measures

- ? % parasitized organisms
- ? Morphological deformities
- ? Sub-organismal biochemical changes
- ? Enzyme levels
- ? Paleolimnology
- ? Sub-lethal behavioral changes
- ? Other

RBPs and **Fish**

- Total number of fish species
- Number of pollution-intolerant fish species
- % introduced species
- % hybrids between different species
- % diseased individuals

RBPs and Algae

- Number of taxa
- Number of diatom taxa
- % live diatoms
- % pollution-sensitive diatom taxa
- % motile taxa
- % eutrophic taxa

The Mekong River
Commission can choose the
best of the biomonitoring
approaches that have been
developed for, and tested in,
other large-scale programs

Program Models that can be followed by MRC

- Multimetric approach (United States)
- Multivariate models (UK, Canada, Australia)
- Biotic Indices (Continental Europe, India)
- Pollution-sensitive taxa response (West Africa)
- Species traits (France, Germany, and Eastern Europe)

Organizing an RBP for the Mekong River Commission

1. Develop *classification system* to group reference sites into homogeneous classes
2. Choose ecologically relevant *biological measures* to assess impact
3. Design *Calibration studies* to determine which measures indicate impact when it occurs, and which ones do not indicate impact when it does not occur
4. Establish *thresholds* to differentiate impaired from non-impaired conditions (biological criteria)

How Can We Design a “Defensible” Monitoring Program?

- Have an explicit purpose
- Have a null hypothesis or an expected response
- Know what type of information is needed
- Know how data collected can improve management
- Justify selection of indicators used
- Choose a design that is powerful enough to detect a likely impact?
- Determine that the cost and time budgeted are feasible to meet monitoring goals

Ecological Monitoring: How Does It Differ from Inventories or Surveys?

- It is used to detect the degree of deviation from the expected norm or from the control
- It involves hypothesis testing (e.g. there is no difference between..... at site A and site B)
- It has a purpose—what are the primary consequences of a certain human action
- It uses indicators to detect differences or changes over time

What Monitoring CANNOT Usually Do

- Determine the cause of change
- Decide how much change is acceptable
- Determine threshold values that cause a change in management strategy
- Avoid “missing an impact” or “sounding false alarms”—indicate that impact has not occurred when it actually has occurred, or indicating that impact has occurred when it actually has not occurred (Type I and II errors)

What Must the Experimental Design of Your Monitoring Program Be Able to Do?

- Separate natural, spatial variability from changes caused by impact (a problem in simple impact-control comparisons)
- Separate natural, temporal variability from changes over time caused by impact (a problem in before-after studies)
- Assume that control and impact sites “track” one another perfectly over time (a problem of Before-After-Control-Impact designs)

Deficiencies of Many Monitoring Programs

- Minimum foundation in ecological theory or knowledge
- Monitoring program from one region is incorrectly applied to another region
- No explicit hypotheses
- Little logic in terms of selection of indicators
- No connection to decision making needs
- No analysis of the power of the monitoring design

What is ecological integrity?

- The maintenance of all internal and external processes and attributes that interact with the environment so that
- the biotic community corresponds to the natural state of the type specific aquatic habitat
- and that principles of self-regulation, resilience, and resistance are used in the management of these systems

Costs of Human Health compared to Ecosystem Health

- 1910: control of water borne diseases in Emscher Valley, Germany ~34,000 annual salaries
- 1990: ecological restoration of Emscher Valley ~170,000 salaries



Five-fold increase in cost of ecological compared to human health indicates:

- The importance placed on achieving ecosystem health today
- Our responsibility to develop effective and efficient biomonitoring tools
- Solution to water-related human health problems may be less costly than restoration of human health



BIOMONITORING— the use of biological variables to survey the environment

- The search for an indicator whose presence, abundance, and/or behavior reflect the effect of a stressor on the freshwater biota
- Indicators may range from the suborganismal (molecular and cellular levels) to the ecosystem
- Ideal biomonitoring tool—“a freeze-dried talking fish on a stick”!

Why use **biological communities** for pollution assessments?

- Reflect overall ecological integrity
- Integrate effects of different stresses
- Provide a measure of fluctuations of environmental conditions over time.
- Relatively inexpensive
- Only practical means when non-point sources of pollution occur

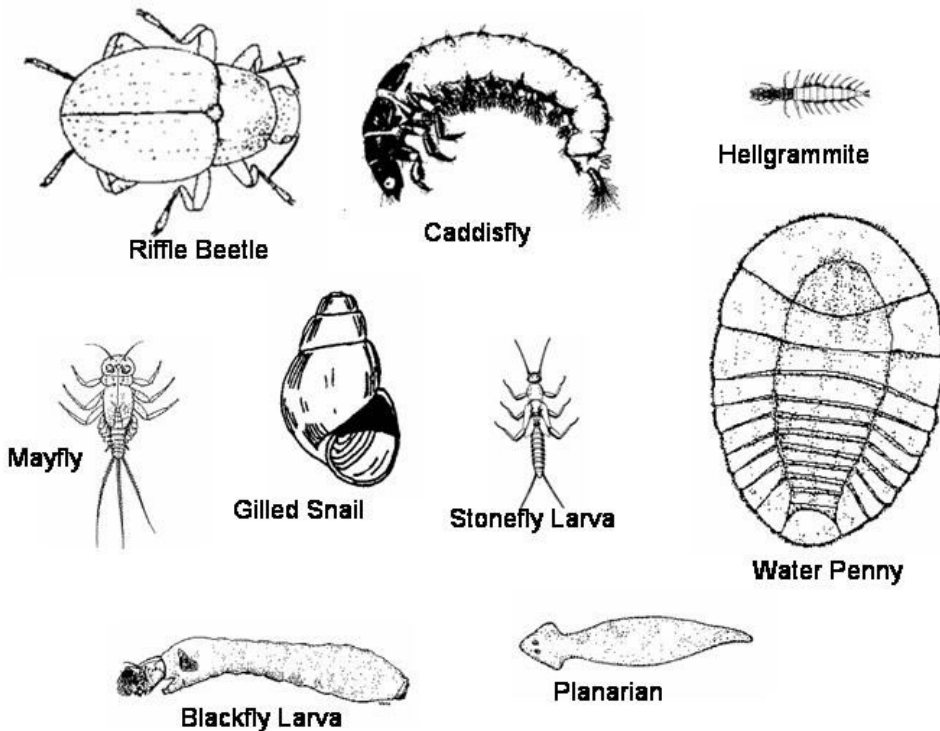
Water chemistry and physical measurements are instantaneous; they reflect conditions at the time of sampling.

Biological measurements reflect past conditions as well.

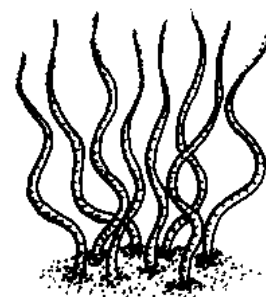
What GROUPS OF ORGANISMS are being used for biomonitoring in USA?

- BENTHIC MACROINVERTEBRATES!

56 State's, territories', and tribe's programs

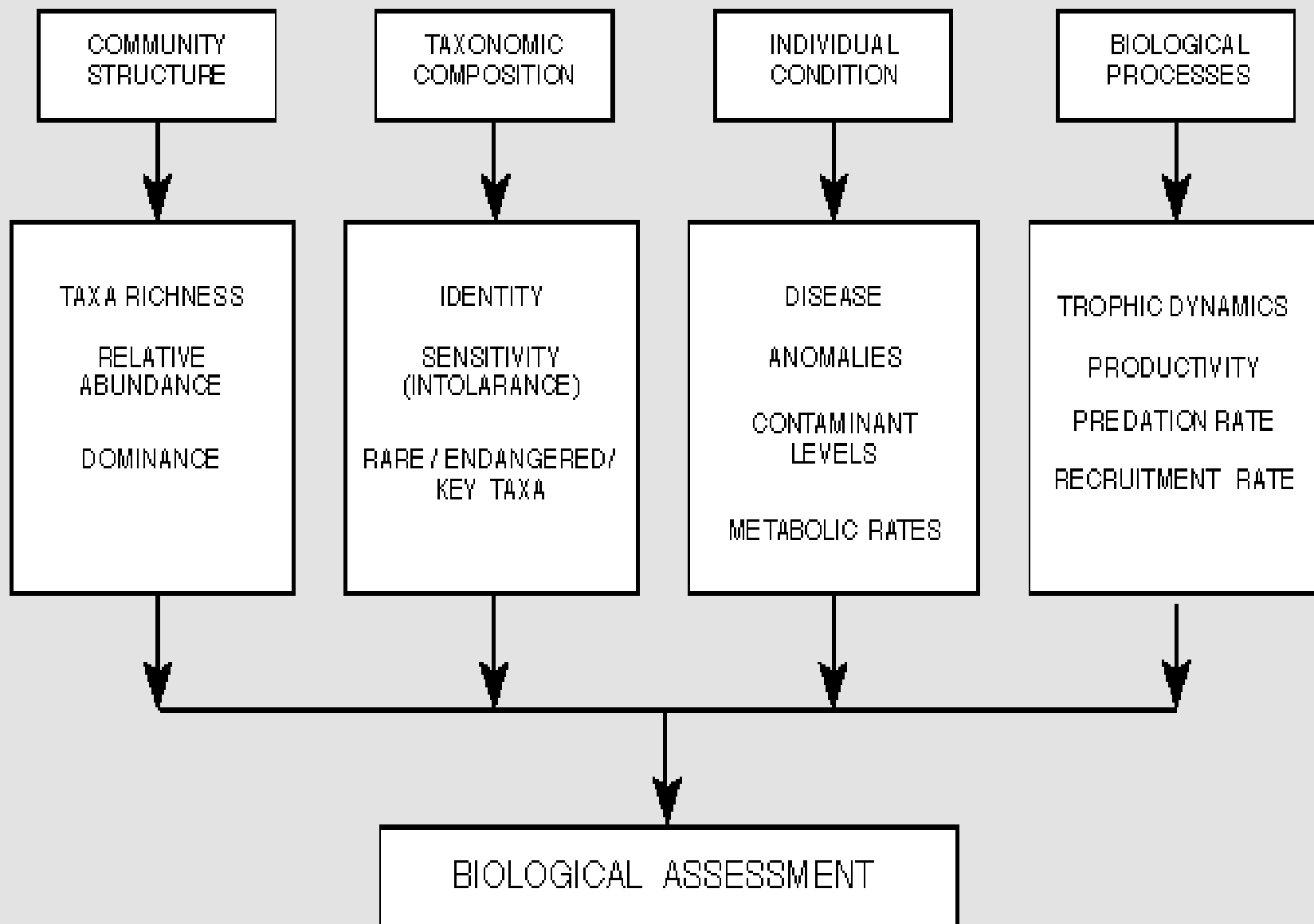


- Benthic Macroinvertebrates 56 State's, territories', and tribe's programs
 - Fish—41 programs
 - Algae—20 programs
- Combinations of groups of organisms—45 programs
- Worldwide— it's clearly BENTHIC MACROINVERTEBRATES!



Why?

- Ubiquitous
- Large number of species and spectrum of responses
- Sedentary nature
- Drift in response to a disturbance
- Long life cycles
- Sample involves simple equipment
- Taxonomy of some groups (EPT) is well known
- Many data analysis tools available
- Responses to specific pollutants known
- Well-suited for experimental approaches
- Extensive literature available
- Biomonitoring adds a temporal component to Physicochemical Monitoring



What types of **measures** or **metrics** can we use in assessing a site?

- Richness
- Proportion of pollution-sensitive organisms
- Pollution tolerant/intolerant organisms
- Functional feeding-groups
- Community diversity
- Biotic indices
- Species traits
- Other?

Richness Measures

- Total number of benthic macroinvertebrate taxa
- Number of Ephemeroptera, Plecoptera, Trichoptera (EPT) taxa, together or as individual groups

Proportion of **Pollution-Sensitive Organisms**

- % Ephemeroptera, Plecoptera, Trichoptera individuals at a site
- % Ephemeroptera
- % Plecoptera
- % Trichoptera

Pollution Tolerant/Intolerant Organisms

- % individuals in the numerically dominant taxa (evenness)
- % intolerant organisms
- Number of intolerant taxa

Functional Feeding-Group Measures

- % filterers (e.g. black flies, net-spinning caddisflies)
- % grazers and scrapers (e.g caddisflies feeding on algae on rock surfaces)
- % shredders (leaf-eating stonefly nymphs)

Community Diversity Measures

- Diversity indices (Shannon, Margalef)
- Similarity Indices (Pinkham-Pearson)
- Evenness or dominance measures

$$\frac{NullDev - ResDev}{NullDev} \times 100\%$$

where $NullDev$ = null deviance of the evaluation data

$ResDev$ = residual deviance of the evaluation data in relation to probabilities predicted by the model

Deviance is calculated as (Hosmer and Lemeshow 1989):

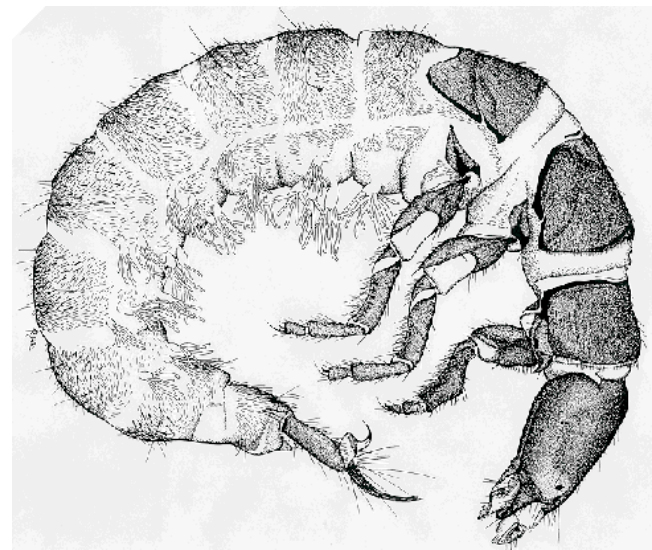
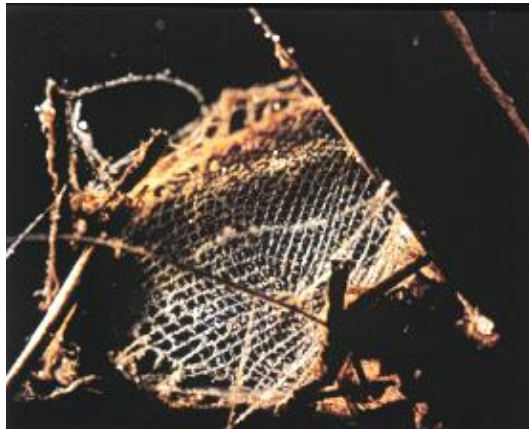
$$-2 \sum_i \{y_i \log(\mu_i) + (1 - y_i) \log(1 - \mu_i)\}$$

where y_i = observed presence (1) or absence(0) at evaluation site i

μ_i = predicted probability of presence at evaluation site i

Biotic Indices

- Pollution tolerance values of species present in the community are weighted by their relative abundances to calculate a score of “average tolerance” at a site
- % Trichoptera that are in the family Hydropsychidae



Habitat Assessment is essential to include because:

- Habitat Quality + Water Quality = Biological Condition
- When *Physical Habitat Quality* at two sites is equal, differences in biological condition are the result of water quality stresses