# *Ranking Stressors: Prevalence & Relative Risk*

**Based on work by**

**John Van Sickle**

**US EPA NHEERL**

**Western Ecology Division**

***Problem:***  *Assessing the relative importance of multiple stressors.*

*Previous Approach:*

-- Compare regional prevalence of each stressor.

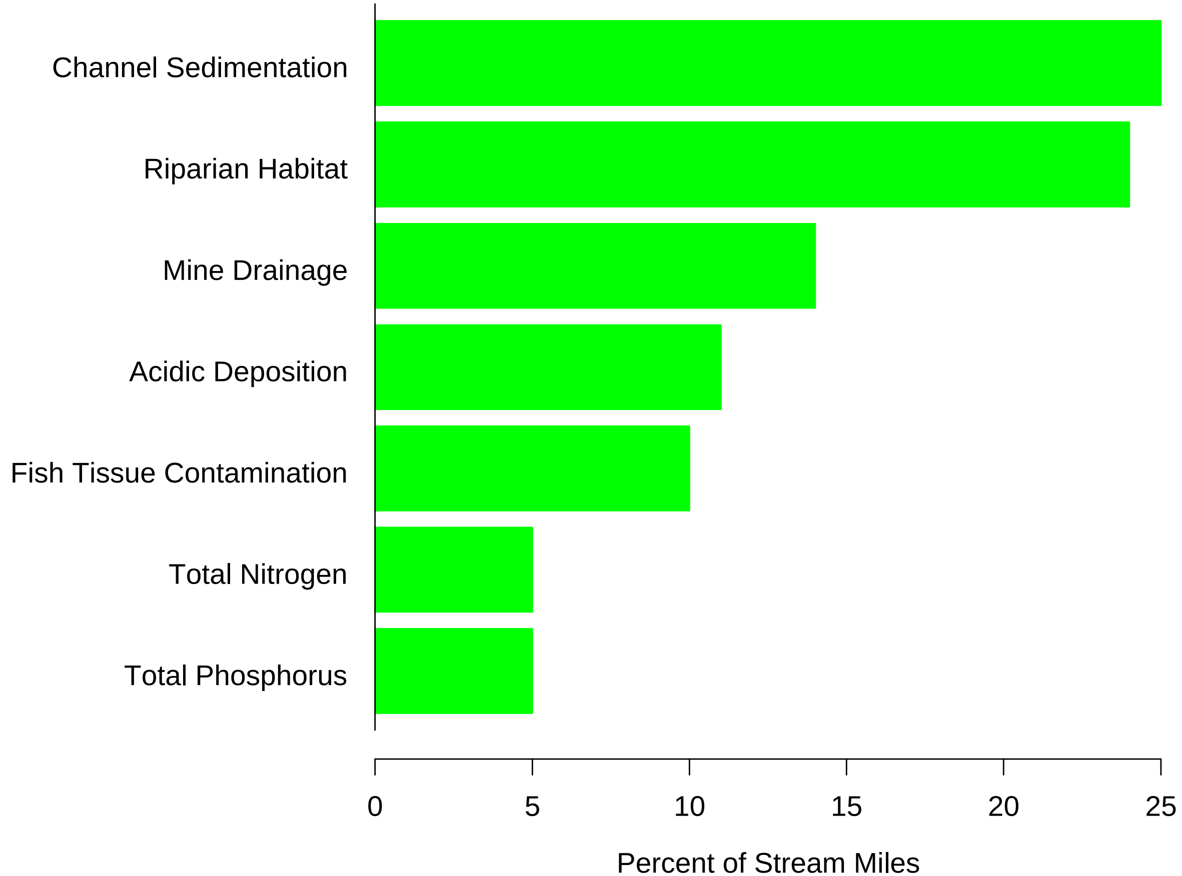
-- ***Define "Poor" condition for each stressor.***

-- ***Estimate percent of stream miles in Poor condition.***

-- Example:

Mid-Atlantic Highlands Assessment (MAHA) streams.
(EPA/903/R-00/015)

# Percent of stream miles in Poor condition



*Limitations of previous approach:*

***1) Stressor "importance" should <u>also</u> be based on the <u>severity</u> <u>of</u> <u>its</u> <u>effects</u> on biological endpoints.***

***2) Definitions of "Poor" and "Good" condition may be arbitrary, either for stressors or endpoints.***

**To move forward:**

1) Assess the <u>strength</u> <u>of</u> <u>association</u> between stressors and endpoints, as a surrogate for "effect severity".

2) Explore association methods for <u>continuous</u>, as well as <u>class-based</u>, stressors and endpoints.

## *Explorations – Modeling associations between stressors and the EPT richness of macroinvertebrates.*

**Continuous methods: stres.ranking.cont.ppt**

-- More complex: Multiple stressors.
-- Language/methods of *correlation* and *multiple regression*.
-- Confronts stressor collinearity.
-- Does not use EMAP sample weights (Yet!)

**Class-based methods: This presentation.**

**--** Simpler: One stressor at a time.
-- Language/methods of *probability*, *risks*, and *odds.*
-- Uses EMAP sample weights.

# Stressor Ranking: Classes, Odds and Risk

**Goal:** To rank stressors, based on their strength of association with biological response indicators.

**Approach:** Use stressor and response classes (MAHA report).

**Responses:** EPT Richness and Fish IBI

**Stressors:**

- Excess sediment
- Riparian condition
- Acid mine drainage
- Acid deposition
- Total P
- Total N

## *Basic tool -- 2-way table*
## *Example: EPT Richness vs. Excess Sediment,*
## *("Base grid" sites, n=80)*

**Site counts**

| | SED GOOD | SED MARG | SED POOR | Total |
|---|---|---|---|---|
| EPT GOOD | 14 | 8 | 0 | 22 |
| EPT MARG | 13 | 18 | 5 | 36 |
| EPT POOR | 2 | 8 | 12 | 22 |
| Total | 29 | 34 | 17 | 80 |

**Percent of Stream Length**

| | SED GOOD | SED MARG | SED POOR | Total |
|---|---|---|---|---|
| EPT GOOD | 17 | 12 | 0 | 29 |
| EPT MARG | 15 | 21 | 7 | 43 |
| EPT POOR | 3 | 7 | 18 | 28 |
| Total | 35 | 40 | 25 | 100 |

# *Combining classes, to simplify.*

| Old | New |
|---|---|
| "Good" → | "OK" |
| "Marginal" → | "OK" |
| "Poor" → | "Poor" |

***Also:***

| | |
|---|---|
| "Not Impaired" → | "OK" |
| "Impaired" → | "Poor" |

# *Association strength:* Calculate the *Relative Risk* of "Poor" EPT richness, in streams having "Poor" sediment, versus streams having "OK" sediment.

$$RR = \frac{\Pr(\text{Poor EPT, given Poor SED})}{\Pr(\text{Poor EPT, given OK SED})}$$

$$RR = \frac{.18/.25}{.10/.75} = 5.4$$

Proportion of stream length
(Pearson $X^2$ = 24.7)

| | SED OK | SED POOR | Total |
|---|---|---|---|
| EPT OK | .65 | .07 | .72 |
| EPT POOR | .10 | .18 | .28 |
| Total | .75 | .25 | 1.00 |

So: **"The risk of Poor EPT is 5.4 times greater in streams with Poor SED than in streams with OK SED."**

# *Odds, and Odds Ratios*

**Odds** of finding Poor EPT, in streams having Poor sediment

$$= \frac{\text{Pr(Poor EPT, given Poor SED)}}{\text{Pr(OK EPT, given Poor SED)}} = \frac{.18/.25}{.07/.25} = 2.60$$

Odds of finding Poor EPT, in streams having OK sediment = 0.15.

**Odds Ratio**, of Poor-SED odds to OK-SED odds $= \frac{2.60}{0.15} = 16.7$

So: **"The odds of finding Poor EPT are 16.7 times greater in Poor-sediment streams than they are in OK-sediment streams."**

| | SED OK | SED POOR | Total |
|---|---|---|---|
| EPT OK | .65 | .07 | .72 |
| EPT POOR | .10 | .18 | .28 |
| Total | .75 | .25 | 1.00 |

## *Stressor ranks, based on odds ratio of "Poor" EPT, for "Poor" vs. "OK" conditions of each stressor.*

| Stressor | Odds Ratio | Stressor Rank | Relative Risk | Pearson $X^2$ | Data set |
|---|---|---|---|---|---|
| Excess Sediment | 14.3 | 1 | 5.1 | 24.7 | Base (n=80) |
| Acid Mine Drainage | 14.1 | 2 | 4.0 | 71.0 | Enhanced (n=355) |
| Total P | 4.8 | 3 | 2.4 | 12.0 | Enhanced |
| Riparian Condition | 4.6 | 4 | 2.7 | 8.2 | Base |
| Total N | 2.0 | 5 | 1.6 | 1.7 | Enhanced |
| Acid Deposition | 0.7 | 6 | 0.7 | .84 | Enhanced |

# *“Effects” <u>and</u> “Prevalence” of Stressors*

| Stressor | “Effect” of Poor Stressor (Odds Ratio of Poor EPT) | “Prevalence” of Poor stressor (% stream length) |
|---|---|---|
| Excess Sediment | 14.3 | 24 |
| Acid Mine Drainage | 14.1 | 14 |
| Total P | 4.8 | 5 |
| Riparian Condition | 4.6 | 25 |
| Total N | 2.0 | 4 |
| Acid Deposition | 0.7 | 9 |

# *“Correlation” between stressor classes?*

YES –

*Consider the top 2 stressors:*

-- Odds of “Poor” Acid mine drainage are only 11%, over all sites.

-- But the odds are 8.5 higher in Poor-sediment streams
than in OK-sediment streams.

◆ “Poor” sediment and “Poor” Acid mine drainage co-occur.

◆ Cannot clearly separate their effects on EPT.

## *Modeling the effects of multiple stressors and natural gradients:*

-- Use multiple logistic regression.

-- Example: A “best model” is:

**log(Odds of Poor EPT) =**
**-8.11 + 3.14(SED) + 2.52(RIP) + 6.36(logSIO2)**

where SED = 1, if “Poor” sediment
= 0, if “OK” sediment

RIP = 1, if “Poor” Riparian condition
= 0, if “OK” Riparian condition

-- Model explains 42% of “deviance” (like $R^2$).

## *Interpreting the logistic regression model*

**log(Odds of Poor EPT) =**
**-8.11 + 3.14(SED) + 2.52(RIP) + 6.36(logSIO2)**

**- Effect of SED**: If RIP and SIO2 are held fixed, then:

$$\frac{\text{Odds of Poor EPT, for Poor Sediment}}{\text{Odds of Poor EPT, for OK Sediment}} = e^{3.14} = 23.1$$

**- Effect of SIO2 :**

"If RIP and SED are held fixed, then a doubling of SIO2 yields a predicted 7-fold increase in the odds of Poor EPT."

# *Issues for class-based associations*

**Sample sizes**

-- Strong constraint on estimates and their uncertainty.

**Defining classes.**

-- Strive for only 2 classes per variable.

-- Avoid rare classes.

**Need statistical methods that use sample weights.**

-- For testing $H_0$: "No association"

-- For CI's on RR and odds.

-- For logistic regression modeling.