

DATA ANALYSES

Experimental Design

In this study, annually for 3 yr we will use a stratified random sampling design to sample benthic fish populations in the Missouri River. As defined by Schaeffer et al. (1990),

“A stratified random sample is one obtained by separating the population elements into nonoverlapping groups, called strata, and then selecting a simple random sample of each stratum.”

For this study, our strata are the six macrohabitats we chose to sample with segments and years being the experimental blocking in which we are randomly sampling those strata.

Stratified random sampling is a commonly used and effective method for evaluating measurement parameters when easily identifiable strata are present (Bhattacharyya and Johnson 1977, Schaeffer et al. 1990, Guetreuter 1993, Brown and Austen 1996). The effectiveness in using the stratified random design we chose comes from: 1. the ability to assure that what we have defined as the most critical, common macrohabitats found in segments are adequately sampled; 2. the enhancement of our ability to assess differences in segments based on these common, critical macrohabitats by accounting for differences in macrohabitat types and thus decreasing the error term used to evaluate segment differences.

Fish Attributes

Community Structure

Definition and Importance

An effective discussion of fish community structure evaluation is given by Gutreuter et al. (1995):

“Fish community structure refers to the relative abundance of fishes of each species within a multispecies assemblage of fishes (Carline 1986). Relative abundance is traditionally measured by catch in numbers per unit of sampling effort, but measures based on weight are also commonly used. A “fish community” theoretically includes all of the fish that use a defined area over a given period of time. The best overall method for measuring fish community structure is the one that is most effective (samples the largest number of specimens) and least selective (captures species in proportion to their occurrence in the sampled area). Given common

available levels of time and personnel, no single method routinely satisfies both criteria (Starrett and Barnikol 1955; Funk 1957; Hayes 1983; Hubert 1983).”

Given such limitations, for this study the analysis procedures for fish community structure among segments include the use of two or three sampling gears for each macrohabitat.

For those macrohabitat types for which the same sampling gear is used comparisons among segments are straight forward; the same is not true for comparisons within macrohabitat types with different mesohabitat types, among macrohabitat types, and among segments over all macrohabitat types due to differences in gear used. There are two ways we will address this issue of different sampling gears to build a strong basis for evaluating differences among segments: (1) weight of evidence based on catch-per-unit-effort (CPUE); and (2) conversion to catch-per-unit-area (CPUA). The weight of evidence approach is basically consistency of differences in CPUE among segments from one gear-macrohabitat analysis to the next. The conversion to CPUA will allow for the combination of data from various sampling gears in a form that should adequately represent each macrohabitat sampled. The CPUA data can then be used to make comparisons among segments using all fish collections. The usage of combined gear data to evaluate differences in fish communities among sites has been demonstrated by Weaver et al. (1993) to be more powerful than individual gear comparisons. Since this study uses a stratified random design (see previous section) to sample the six primary macrohabitats found in the Missouri River, the combination of macrohabitat and CPUA comparisons among segments should provide an effective evaluation of the Missouri River benthic fish community.

We will calculate CPUA as follows:

A. Beach Seine ($CPUA_{BS}$)

$$CPUA_{BS} = \frac{N}{\left(\frac{\pi \times r^2}{2}\right)}$$

where:

N = number of fish collected

r = radius of the seine hall

/ B. Benthic Trawl ($CPUA_{BT}$)

$$CPUA_{BT} = \frac{N}{W_{BT} \times D_{BT}}$$

where:

N = number of fish collected

W_{BT} = width of the benthic trawl (i.e., 2 m)

D_{BT} = distance the trawl was pulled

C. Drift Trammel Net ($CPUA_{DNT}$)

$$CPUA_{DNT} = \frac{N}{W_{DNT} \times D_{DNT}}$$

where:

N = number of fish collected

W_{DNT} = width of the trammel net (i.e., 22.9 m)

D_{BT} = distance the trammel net was drifted

D. Electrofishing ($CPUA_{EF}$)

$$CPUA_{EF} = \frac{N}{W_{EF} \times D_{EF}}$$

where:

N = number of fish collected

W_{EF} = width of the effective electrical field of the shocking sphere (i.e., 3.5 m)

D_{EF} = distance covered by the electroshocking boat

E. Stationary Gill Net ($CPUA_{SNG}$)

$$CPUA_{SNG} = \sum_{i=1}^{N_s} \frac{n_i}{W_{SNG} \times T_{SNG} \times s_i}$$

where:

N = number of species collected

n_i = number fish collected from species i

W_{SNG} = width of the gill net (i.e., 30.5 m)

T_{SNG} = time (sec) for gill net set

s_i = average swimming speed (m/sec) for species i (e.g. 0.03 m/sec)

To try and validate the use of CPUA we will use a form of sensitivity analysis to assure that the CPUA estimate is not unevenly affected by inclusion of particular gear types; parallel analyses will be done for CPUE for comparison to CPUA and to results from other researchers. As long as the CPUA is not swamped by inclusion of particular gear types, differences in CPUA among segments should account for actual differences. Any differences noted in CPUA will be a result of a combination of overall population differences and/or dominance of macrohabitat types within a segment.

Relative Abundance

Community structure is often evaluated using relative abundance. “It is called relative abundance to stress the fact that virtually every sampling method is somewhat selective and therefore produces a biased view of true abundance,” (Gutreuter et al. 1995). In this study, effective comparisons of segments based on relative abundance will be accomplished by the use of multiple sampling gears in each macrohabitat type to strengthen the weight of evidence approach and by the use of an area standardized measure (i.e., CPUA) that combines multiple sampling gear data for macrohabitat characterization.

In this study relative abundance will be assessed for all fish combined, the benthic fish species as a group, and individual species. Relative abundance will be measured in units appropriate to the method (i.e., per m^3 for benthic trawl, drifting trammel net, and beach seine or per min for electroshocking and stationary gillnet) used for each macrohabitat type analysis. Conversion of electroshocking and stationary gillnet to some form of area sampled measure will be used to try and standardize data for input into the CPUA calculations.

Species Richness

For this study, species richness refers to the total number of fish species and the total number of target benthic fish taken in a segment of the river. As was the case for the LTRMP, species richness will not include hybrids (e.g., goldfish x common carp). Taxonomic categories

higher than species will not be included in species richness calculations if a lower level taxonomic category within the higher category is identified in that segment (e.g., *Ictiobus sp.* if *Ictiobus bubalus* has already been found to be present).

In this study, as in the study by Gutreuter et al. (1995), “species richness is a component of the overall diversity of the fish community. Because the sample species richness increases with increasing sampling effort, comparison of species richness estimates requires either constant sampling effort or formal estimation methods. Estimation of species richness is an important but difficult (Bunge and Fitzpatrick 1993) task.” For this study, the standardization of macrohabitats, gears, and sampling effort for all segments of the Missouri River and the number of macrohabitat replicates sampled should provide for a fairly accurate measure of species richness that can be used to make valid segment comparisons. In order to validate this assumption we will use the method of rarefaction discussed by Ludwig and Reynolds (1988) which they present as the most effective way to assess species richness when sample sizes differ. Despite the balanced design of this study, it is expected that we will not collect the same number of fish in all segments and all habitats within and among segments in this study which necessitates the validation of any segment comparisons of species richness using rarefaction analysis.

The method of rarefaction, as describe by Hurlbert (1971) and James and Rathburn (1981), is the plotting of the expected number of species in a sample of n individuals drawn from a population of N total individuals in order to evaluate species richness at a constant sample size among factors being compared (i.e., segments, macrohabitat). The rarefaction equation is:

$$E(S_n) = \sum_{i=1}^s \left[1 - \frac{\binom{N-N_i}{n}}{\binom{N}{n}} \right]$$

where:

$E(S_n)$ is expected number of species

n is sample size

N is total number of individuals collected

N_i is total number of individuals in species i

S is total number of species collected

In this study, the smallest sample size collected for the level (i.e., segment or macrohabitat) of the comparison being made will be used for the calculation of $E(S_n)$ for all equivalent levels among which comparisons are to be made.

Population Structure

Definition and Importance

In this study, as was the case for Gutreuter et al. (1995), “Population structure refers to the distribution of individuals of a single species among size or age groups.” For this study, population structure evaluations will be done on each species from the list of target benthic fish.

Size Distribution

For this study, as it was for Gutreuter et al. (1995), “the size distribution for a given species is the vector (list) of the numbers of specimens taken in a collection or a unit of effort that fall into selected size categories.” In this study, the unit of effort is the macrohabitat since the combination of the multiple gears used in each macrohabitat is designed to, as much as possible, sample the entire fish population within a macrohabitat. As was the case for species richness, the standardization of macrohabitats, gears, and sampling effort for all segments of the Missouri River and the number of macrohabitat replicates sampled should provide for a fairly accurate estimate of the population size structure (i.e., length and weight) that can be used to make valid segment comparisons using the Kolmogorov-Smirnov goodness-of-fit test (Daniel 1978, Steele and Torrie 1980, SAS 1990).

Recruitment

From Everhart and Youngs (1981), “recruitment can be defined as the addition of new members to the aggregate under consideration. In a fishery recruitment is the supply of fish that becomes available at some particular stage in their life history...” In this study, we are mainly interested in YOY and Age-1 recruitment to gear and potentially the ratio of YOY to Age-1; the definition of age classes based on lengths will be determined from the results of the age and growth work described in the next section. As was the case for relative abundance, effective comparisons of segments based on recruitment will be accomplished by the use of multiple sampling gears in each habitat type to strengthen the weight of evidence approach and by the use of a area standardized measure (i.e., C_{PUA}) that combines multiple sampling gear data for macrohabitat characterization.

Individual Characteristics

Growth

In this study, we are concerned with the growth of each species of benthic fish for which we will collect aging structures. Our analyses will include average growth per year both in length and weight as back calculated from aging structures (Everhart and Youngs 1981, Jearld, Jr. 1985).

Body Condition

In this study, for those species for which the equations are available or for which we develop equations, we will calculate relative weight (Wr) (Anderson and Gutreuter 1985):

$$Wr = \frac{W}{Ws} \times 100$$

where:

W = weight (g)

$Ws = a'L^b$

L = length (mm)

a' , b' = constants from the weight length relationship for a species that ideally account for genetically determined shape differences.

For those species for which no relative weight equation is available, we will calculate relative condition factor (Kn) (Anderson and Gutreuter 1985):

$$Kn = \frac{W}{aL^b}$$

where:

W = weight (g)

L = length (mm)

a , b = constants from the weight length relationship for a species

Physical Factors

Definition and Importance

In this study numerous physical factors, along with the already determined habitat categories, will be measured that describe the physical environment within which the fish were collected. These other physical factors will be measured in an attempt to understand the relationship between fish attributes and the physical environment.

Water

The following water parameters and the associated units will be measured in this study:

Depth in meters (m)

Velocity in meters per second (m/sec)

Water temperature in Celsius (C)

Conductivity in microohms per centimeter (μ ohms/cm)

Turbidity in Nephelometric Turbidity Units (NTU)

Bottom Features

Since this study is focused on benthic fish it seems obvious that what the bottom shape is (bedform) and what it is made of (substrate composition) are important physical factors to consider. In this study, we will examine bedform using the slope per distance, mean, frequency per distance, and variance of trough depth with a fish subsample to give us effective measures of bottom relief and thus, potential current refugia for the fish. The measurements will be made by randomly spacing a predetermined number of points per meter of bottom covered along the bedform chart and measuring the water depth at each point. The number of points used will be calculated by dividing the half the average adult length of the species of concern into a meter and multiplying by the number of meters in a bedform trace. A trough will be defined at any point where the points on both sides of that point are higher. The slope of the regression for troughs should give us an index of how directly the subsample area is affected by flow. Mean trough depth will be calculated by taking the mean of the difference between the average of the two points on either side of a trough point and the trough point. Trough frequency per distance will be calculated by calculating the number of trough points per meter and should give us an estimate of the number of troughs in the subsample area. The variance of trough depth should produce a measure of the evenness of trough depths. Also we will examine substrate in separate size class percentages, the geometric mean of particle size (Orth 1985) and the fredle index (Chapman 1988); the last two indices are the recommended measures for evaluating substrate composition (McMahon et al. 1996):

A. Geometric Mean of Substrate Size (G_s)

$$G_s = D_1^{w_1} \times D_2^{w_2} \times \dots \times D_i^{w_i}$$

where:

D_i is the mean diameter of size range i .

w_i is the proportion by weight of the sample in size range i .

B. Fredle Index (F)

$$F = \frac{G_s}{S} \text{ where } S = \sqrt{\frac{d_{75}}{d_{25}}}$$

and

d_{75} substrate diameter below which 75% of the sample lie.

d_{25} substrate diameter below which 25% of the sample lie.

Hypotheses

Definition and Importance

In formulating hypotheses that can be tested using the appropriate statistics, the standard procedure is to state a null hypothesis that can be rejected at a given level of confidence (i.e., α); for this study we will use the standard practice of rejecting a null hypothesis using $\alpha < 0.05$. By using $\alpha < 0.05$ we are accepting the fact that there would be a 1 in 20 chance that any rejection of a null hypothesis we observe could have randomly occurred. We will also use power analysis (i.e., β) in support of any null hypothesis we can not reject using $\alpha < 0.05$; again, we will use the standard practice which in this case is to use $\beta = 0.8$ (Steel and Torrie 1980). A $\beta = 0.8$ translates into a 1 in 5 (i.e. $1 - \beta = 0.2$ Type II error) chance that we accepted a null hypothesis when in actuality it should have been rejected. For accepted null hypotheses, power analysis will allow us to estimate how large observed differences would have had to be in order for us to reject any null hypothesis; concurrently, we will be able to estimate how many more samples would have been necessary in order for us to have rejected any such null hypothesis. By using the combination of α and β we will be able to discuss any rejected or accepted null hypothesis from as defensible a position as possible. The calculation we will use for power analysis comes from Snedecor and Cochran (1982):

$$n = (Z_{\alpha} + Z_{\beta})^2 \times \frac{2 \times S^2}{D^2}$$

$$\text{thus, } |D| = \sqrt{2} \times (Z_{\alpha} + Z_{\beta}) \times SE$$

$$\text{with } SE = \sqrt{\frac{S^2}{n}}$$

where:

n = sample size

$Z_{\alpha} = Z_{0.05} = 1.96$

$Z_{\beta} = Z_{0.2} = 0.84$

S^2 = sample variance

|D| = difference between sample means

Along with any null hypothesis there is always alternate hypothesis(es) that may represent possible explanation(s) when the null hypothesis is rejected with our accepted level of confidence. For example:

Null hypothesis: There is no difference in segments in the CPUA of the identified benthic fish group caught in main channel habitat using the benthic trawl.

Alternate hypotheses: The CPUA of the identified benthic fish group caught in main channel habitat using the benthic trawl increases as we move upstream.

We will use any appropriate statistical methods necessary to try to determine any alternate hypothesis the data may support.

In the end, following the procedures just outlined will allow us to present as strong a case for the final set of conclusions we make from the results of this study.

At this point, we will try to specify the hypotheses that we may want to address from this study; these hypotheses apply both within and among years. It must be kept in mind that the amount and quality of data that is actually collected, despite the use of an effective sampling design, will determine if we have the ability to accept or reject these hypotheses in the end.

Among Segment Type Hypotheses

(Example: Unchannelized versus channelized segments)

I. For each species of fish, the benthic fish group, and all fish combined:

A. Community structure

1. Relative abundance

a. CPOA

i. no difference using all subsamples

ii. no difference using subsamples for each macrohabitat

b. CPUE

i. no difference using all subsamples

ii. no difference using subsamples for each macrohabitat

iii. no difference using subsamples for each gear-macrohabitat combination

2. Species richness

a. no difference using all subsamples

b. no difference using rarefaction values

B. Population structure

1. Size distribution

a. no difference using all subsamples

b. no difference using subsamples for each macrohabitat

c. no difference using subsamples for each gear-macrohabitat combination

2. Recruitment

a. CPOA YOY and Age-1

i. no difference using all subsamples

ii. no difference using subsamples for each macrohabitat

b. CPUE YOY and Age-1

i. no difference using all subsamples

ii. no difference using subsamples for each macrohabitat

iii. no difference using subsamples for each gear-macrohabitat combination

c. YOY/Age-1

i. no difference using all subsamples

ii. no difference using subsamples for each macrohabitat

iii. no difference using subsamples for each gear-macrohabitat combination

- C. Individual characteristics
 - 1. Growth rate per year in length and weight
 - a. no difference using all subsamples
 - 2. Body condition
 - a. no difference using all subsamples
 - b. no difference using subsamples for each macrohabitat
 - c. no difference using subsamples for each gear-macrohabitat combination
- II. Physical factors
 - A. Depth, velocity, water temperature, conductivity, turbidity, and bedform
 - 1. no difference using all subsamples
 - 2. no difference using subsamples for each macrohabitat
 - 3. no difference using subsamples for each gear-macrohabitat combination
- III. Correlation of fish measures to physical factors
 - A. Community structure
 - 1. Relative abundance
 - a. CPUA
 - i. no relationship using all subsamples
 - ii. no relationship using subsamples for each macrohabitat
 - b. CPUE
 - i. no relationship using all subsamples
 - ii. no relationship using subsamples for each macrohabitat
 - iii. no relationship using subsamples for each gear-macrohabitat combination
 - 2. Species richness
 - a. no relationship using all subsamples and means of segment physical factors
 - b. no relationship using rarefraction values and means of segment physical factors
 - B. Population structure
 - 1. Recruitment
 - a. CPUA YOY and Age-1
 - i. no relationship using all subsamples
 - ii. no relationship using subsamples for each macrohabitat
 - b. CPUE YOY and Age-1
 - i. no relationship using all subsamples
 - ii. no relationship using subsamples for each macrohabitat
 - iii. no relationship using subsamples for each gear-macrohabitat combination

- c. YOY/Age-1
 - i. no relationship using all subsamples
 - ii. no relationship using subsamples for each macrohabitat
 - iii. no relationship using subsamples for each gear-macrohabitat combination
- C. Individual characteristics
 - 1. Growth rate per year in length and weight
 - a. no relationship using all subsamples and means of segment physical factors
 - 2. Body condition
 - a. no relationship using all subsamples
 - b. no relationship using subsamples for each macrohabitat
 - c. no relationship using subsamples for each gear-macrohabitat combination
- IV. Correlations among physical factors
 - A. Depth, velocity, water temperature, conductivity, turbidity, and bedform
 - 1. no relationship using all subsamples
 - 2. no relationship using subsamples for each macrohabitat
 - 3. no relationship using subsamples for each gear-macrohabitat combination
- V. Correlations among fish species
 - A. Community structure
 - 1. Relative abundance
 - a. CPUA
 - i. no relationship using all subsamples
 - ii. no relationship using subsamples for each macrohabitat
 - b. CPUE
 - i. no relationship using all subsamples
 - ii. no relationship using subsamples for each macrohabitat
 - iii. no relationship using subsamples for each gear-macrohabitat combination
 - B. Population structure
 - 1. Recruitment
 - a. CPUA YOY and Age-1
 - i. no relationship using all subsamples
 - ii. no relationship using subsamples for each macrohabitat
 - b. CPUE YOY and Age-1
 - i. no relationship using all subsamples
 - ii. no relationship using subsamples for each macrohabitat
 - iii. no relationship using subsamples for each gear-macrohabitat combination

- c. YOY/Age-1
 - i. no relationship using all subsamples
 - ii. no relationship using subsamples for each macrohabitat
 - iii. no relationship using subsamples for each gear-macrohabitat combination

C. Individual characteristics

- 1. Growth rate per year in length and weight
 - a. no relationship using all subsamples
- 2. Body condition
 - a. no relationship using all subsamples
 - b. no relationship using subsamples for each macrohabitat
 - c. no relationship using subsamples for each gear-macrohabitat combination

Among Segments Hypotheses

I. For each species of fish, the benthic fish group, and all fish combined:

A. Community structure

- 1. Relative abundance
 - a. CPUA
 - i. no difference using all subsamples
 - ii. no difference using subsamples for each macrohabitat
 - b. CPUE
 - i. no difference using all subsamples
 - ii. no difference using subsamples for each macrohabitat
 - iii. no difference using subsamples for each gear-macrohabitat combination

B. Population structure

- 1. Size distribution
 - a. no difference using all subsamples
 - b. no difference using subsamples for each macrohabitat
 - c. no difference using subsamples for each gear-macrohabitat combination
- 2. Recruitment
 - a. CPUA YOY and Age-1
 - i. no difference using all subsamples
 - ii. no difference using subsamples for each macrohabitat
 - b. CPUE YOY and Age-1
 - i. no difference using all subsamples
 - ii. no difference using subsamples for each macrohabitat
 - iii. no difference using subsamples for each gear-macrohabitat combination

- c. YOY/Age-1
 - i. no difference using all subsamples
 - ii. no difference using subsamples for each macrohabitat
 - iii. no difference using subsamples for each gear-macrohabitat combination
 - C. Individual characteristics
 - 1. Growth rate per year in length and weight
 - a. no difference using all subsamples
 - 2. Body condition
 - a. no difference using all subsamples
 - b. no difference using subsamples for each macrohabitat
 - c. no difference using subsamples for each gear-macrohabitat combination
- II. Physical factors
- A. Depth, velocity, water temperature, conductivity, turbidity, and bedform
 - 1. no difference using all subsamples
 - 2. no difference using subsamples for each macrohabitat
 - 3. no difference using subsamples for each gear-macrohabitat combination
- III. Correlation of fish measures to physical factors
- A. Community structure
 - 1. Relative abundance
 - a. CPUA
 - i. no relationship using all subsamples
 - ii. no relationship using subsamples for each macrohabitat
 - b. CPUE
 - i. no relationship using all subsamples
 - ii. no relationship using subsamples for each macrohabitat
 - iii. no relationship using subsamples for each gear-macrohabitat combination
 - B. Population structure
 - 1. Recruitment
 - a. CPUA YOY and Age-1
 - i. no relationship using all subsamples
 - ii. no relationship using subsamples for each macrohabitat
 - b. CPUE YOY and Age-1
 - i. no relationship using all subsamples
 - ii. no relationship using subsamples for each macrohabitat
 - iii. no relationship using subsamples for each gear-macrohabitat combination

- c. YOY/Age-1
 - i. no relationship using all subsamples
 - ii. no relationship using subsamples for each macrohabitat
 - iii. no relationship using subsamples for each gear-macrohabitat combination
- C. Individual characteristics
 - 1. Body condition
 - a. no relationship using all subsamples
 - b. no relationship using subsamples for each macrohabitat
 - c. no relationship using subsamples for each gear-macrohabitat combination
- IV. Correlations among physical factors
 - A. Depth, velocity, water temperature, conductivity, turbidity, and bedform
 - 1. no relationship using all subsamples
 - 2. no relationship using subsamples for each macrohabitat
 - 3. no relationship using subsamples for each gear-macrohabitat combination
- V. Correlations among fish species
 - A. Community structure
 - 1. Relative abundance
 - a. CPUA
 - i. no relationship using all subsamples
 - ii. no relationship using subsamples for each macrohabitat
 - b. CPUE
 - i. no relationship using all subsamples
 - ii. no relationship using subsamples for each macrohabitat
 - iii. no relationship using subsamples for each gear-macrohabitat combination
 - B. Population structure
 - 1. Recruitment
 - a. CPUA YOY and Age-1
 - i. no relationship using all subsamples
 - ii. no relationship using subsamples for each macrohabitat
 - b. CPUE YOY and Age-1
 - i. no relationship using all subsamples
 - ii. no relationship using subsamples for each macrohabitat
 - iii. no relationship using subsamples for each gear-macrohabitat combination
 - c. YOY/Age-1
 - i. no relationship using all subsamples
 - ii. no relationship using subsamples for each macrohabitat
 - iii. no relationship using subsamples for each gear-macrohabitat combination

- C. Individual characteristics
 - 1. Growth rate per year in length and weight
 - a. no relationship using all subsamples
 - 2. Body condition
 - a. no relationship using all subsamples
 - b. no relationship using subsamples for each macrohabitat
 - c. no relationship using subsamples for each gear-macrohabitat combination

Among Macrohabitats Hypotheses

- I. For each species of fish, the benthic fish group, and all fish combined:
 - A. Community structure
 - 1. Relative abundance
 - a. CPOA
 - i. no difference using subsamples for each macrohabitat
 - b. CPUE
 - i. no difference using subsamples for each macrohabitat
 - ii. no difference using subsamples for each gear-macrohabitat combination
 - B. Population structure
 - 1. Size distribution
 - a. no difference using subsamples for each macrohabitat
 - b. no difference using subsamples for each gear-macrohabitat combination
 - 2. Recruitment
 - a. CPOA YOY and Age-1
 - i. no difference using subsamples for each macrohabitat
 - b. CPUE YOY and Age-1
 - i. no difference using subsamples for each macrohabitat
 - ii. no difference using subsamples for each gear-macrohabitat combination
 - c. YOY/Age-1
 - i. no difference using subsamples for each macrohabitat
 - ii. no difference using subsamples for each gear-macrohabitat combination
 - C. Individual characteristics
 - 1. Body condition
 - a. no difference using subsamples for each macrohabitat
 - b. no difference using subsamples for each gear-macrohabitat combination

II. Physical factors

A. Depth, velocity, water temperature, conductivity, turbidity, and bedform

1. no difference using subsamples for each macrohabitat
2. no difference using subsamples for each gear-macrohabitat combination

III. Correlation of fish measures to physical factors

A. Community structure

1. Relative abundance

a. C_{PUA}

- i. no relationship using subsamples for each macrohabitat

b. C_{PUE}

- i. no relationship using subsamples for each macrohabitat
- ii. no relationship using subsamples for each gear-macrohabitat combination

B. Population structure

1. Recruitment

a. C_{PUA} YOY and Age-1

- i. no relationship using subsamples for each macrohabitat

b. C_{PUE} YOY and Age-1

- i. no relationship using subsamples for each macrohabitat
- ii. no relationship using subsamples for each gear-macrohabitat combination

c. YOY/Age-1

- i. no relationship using subsamples for each macrohabitat
- ii. no relationship using subsamples for each gear-macrohabitat combination

C. Individual characteristics

1. Body condition

- a. no relationship using subsamples for each macrohabitat

- b. no relationship using subsamples for each gear-macrohabitat combination

IV. Correlations among physical factors

A. Depth, velocity, water temperature, conductivity, turbidity, and bedform

1. no relationship using subsamples for each macrohabitat
2. no relationship using subsamples for each gear-macrohabitat combination

V. Correlations among fish species

A. Community structure

1. Relative abundance

a. C_{PUA}

- i. no relationship using subsamples for each macrohabitat

- b. CPUE
 - i. no relationship using subsamples for each macrohabitat
 - ii. no relationship using subsamples for each gear-macrohabitat combination
- B. Population structure
 - 1. Recruitment
 - a. CPUA YOY and Age-1
 - i. no relationship using subsamples for each macrohabitat
 - b. CPUE YOY and Age-1
 - i. no relationship using subsamples for each macrohabitat
 - ii. no relationship using subsamples for each gear-macrohabitat combination
 - c. YOY/Age-1
 - I. no relationship using subsamples for each macrohabitat
 - ii. no relationship using subsamples for each gear-macrohabitat combination
- C. Individual characteristics
 - 1. Body condition
 - a. no relationship using subsamples for each macrohabitat
 - b. no relationship using subsamples for each gear-macrohabitat combination

Statistical Analyses

Testing Assumptions

Since our first attempt at analyzing our data will focus on parametric techniques, we must make sure we meet the two primary assumptions of parametric techniques (i.e., normality of the data and homogeneity of variance).

Test for Normality

Using a standard univariate technique, we will plot the frequency distribution of the data and calculate a test statistic for normality (i.e, Shapiro-Wilk W, Shapiro Wilk (1965)). Our decision for whether or not a variable is normally distributed will be based mainly on examination of the frequency distribution; the test statistic is only a guide since it is highly sensitive to any deviations from normal (i.e., one point may cause the test to indicate non-normal data). When the frequency distribution appears close to normal and the value of the test statistic is close to what would be defined as normal for the test statistic (i.e., ignore the probability associated with the test statistic unless it indicates normality) we will assume that the data is normally distributed.

Note: This will have to be done for every variable at the level at which it is being analyzed (i.e., test for normality of combined subsample values for a gear-habitat collection)

Homogeneity of Variance

We will test for homogeneity of variance based on a gear-habitat level by plotting the residuals based on habitat for a gear and testing for a significant correlation between means and variances for the gear-habitat level.

Transformations

If the normality and/or the homogeneity of variance assumptions are violated then the data will have to be transformed to try and meet these assumptions. For our data, the transformation used in most cases will likely be either $\log(x)$ or $\log(x + 1)$, the latter allowing zeros (Green 1979, Steel and Torrie 1980). Then, the transformed data must again be tested to evaluate if the assumptions of normality and homogeneity of variance are being met. If after transformation the assumptions of normality and homogeneity of variance are still being violated we must shift to either rank ANOVA or an appropriate nonparametric test.

Analyses of Variance

For the purpose of understanding within and among years the differences among segments, habitats, and habitats within segments we will analyze each species of fish, the benthic fish group, all fish combined, and physical factor measurements separately (i.e., two-way ANOVA) and in combinations of fish species, fish measurements, and physical factors (i.e., two-way multivariate analysis of variance--MANOVA) (Johnson and Wichern 1992). The MANOVA is a multivariate technique that will allow us to state whether or not there are overall significant differences within and among years among segments, among habitats, and among habitats within segments. The ANOVA will allow us to test which of the separate abiotic and biotic factors are different within and among years among segments, among habitats, and among habitats within sections.

Theoretical Main-Effects Model for the Overall Study

I. Model

Year (3 total or 2 degrees of freedom--df)

Segment (17 total or 16 df)

Year x Segment (2 x 16 or 32 df)

Habitat (6 total or 5 df)

Year x Habitat (2 x 5 or 10)

Segments x Habitat (16 x 5 or 80 df)

Year x Segment x Habitat (2 x 16 x 5 or 160 df)

Error (corrected for the mean with $1530 - 1 - 2 - 16 - 32 - 5 - 10 - 80 - 160 = 1224$ df)

Note: total observations (i.e., n) is:

$$3 \text{ years} \times 17 \text{ segments} \times 6 \text{ habitats} \times 5 \text{ habitat replicates} = 1530$$

As the heading states, the actual model used to test each of the hypotheses listed in the previous section is only theoretical in nature. As we all already know, the actual number of observations collected and the number of replicates of each habitat sampled varies from segment to segment. The actual number of total observations will be something less than 1530. Also, we already know that many of the tests of habitat differences will be limited to comparisons among samples collected by the same gear or gears in the habitats being compared.

II. Comparisons Between Sections

Comparison of specific sections will be accomplished using contrast statements along with the ANOVA being applied at the time.

Actual Main-Effects Models to be Used in this Study

I. At the smallest scale the tests will be done by gear type:

A. Boat Electroshocking

1. Habitats compared will be:

Outside Bend

Inside Bend (deep bank and rocks)

Tributary Mouth (small)

2. Main-Effects Model:

Year (df equal to number of years with following described segments - 1)
Segment (df equal to number of segments with all three habitats - 1)
Year x Segment (df of year x df segment)
Habitat (df of 2)
Year x Habitat (df of year x 2)
Segment x Habitat (df of segment x 2)
Year x Segment x Habitat (df of year x df segment x 2)
Error (df equal total number observations from years with all segments
with all three habitats - df of year - df of segment - df of year x
segment - 2 - df year x 2 - df of segment x 2 - df of year x segment
x 2 - 1)

B. Drift Trammel Net and Benthic Trawl

1. Habitats compared will be:

Main Channel Crossover
Outside Bend
Inside Bend (channel border)
Tributary Mouths (large)
Secondary Channel Connected (deep)

2. Main-Effects with Split Plots Model:

Year (df equal to number of years with following described segments - 1)
Segment (df equal to number of segments with all five habitats - 1)
Year x Segment (df of year x df segment)
Habitat (df of 4)
Year x Habitat (df of year x 4)
Segment x Habitat (df of segment x 4)
Year x Segment x Habitat (df of year x df segment x 4)
Error A (df equal total number observations from years with all segments
with all five habitats - df of year - df of segment - df of year x
segment - 4 - df year x 4 - df of segment x 4 - df of year x segment
x 4 - 1)

Gear (df 1 since two gears were used)
 Gear x Year (df year)
 Gear x Segment (df of segment)
 Gear x Year x Segment (df year x df of segment x 4)
 Gear x Habitat (df of 4)
 Gear x Year x Habitat (df years x 4)
 Gear x Segment x Habitat (df of segment x 4)
 Gear x Years x Segment x Habitat (df of year x df segment x 4)
 Error B (df total number of subsamples collected for the two gears
 combined - df of year - df of segment - df of year x segment - 4 -
 df year x 4 - df of segment x 4 - df of year x segment x 4 - df Error
 A - df of gear - df gear x year - df of gear x segment - df gear x
 year x segment - df of gear x 4 - df gear x year x 4 - df gear x
 segment x 4 - df gear x year x segment x 4 - 1)

C. Stationary Gill Net

1. Habitats compared will be:

Inside Bend (pool)
 Tributary Mouth (small)
 Secondary Channel Non-connected

2. Main-Effects Model:

Year (df equal to number of years with following described segments - 1)
 Segment (df equal to number of segments with all three habitats - 1)
 Year x Segment (df of year x df segment)
 Habitat (df of 2)
 Year x Habitat (df of year x 2)
 Segment x Habitat (df of segment x 2)
 Year x Segment x Habitat (df of year x df segment x 2)
 Error (df equal total number observations from years with all segments
 with all three habitats - df of year - df of segment - df of year x
 segment - 2 - df year x 2 - df of segment x 2 - df of year x segment
 x 2 - 1)

D. Beach Seine

1. Habitats compared will be:

Inside Bend (sandbar)
 Secondary Channel Connected (shallow)
 Secondary Channel Connected (deep)
 Secondary Channel Non-connected

2. Main-Effects Model:

Year (df equal to number of years with following described segments - 1)

Segment (df equal to number of segments with all four habitats - 1)

Year x Segment (df of year x df segment)

Habitat (df of 3)

Year x Habitat (df of year x 3)

Segment x Habitat (df of segment x 3)

Year x Segment x Habitat (df of year x df segment x 3)

Error (df equal total number observations from years with all segments with all four habitats - df of year - df of segment - df of year x segment - 3 - df year x 3 - df of segment x 3 - df of year x segment x 3 - 1)

Principal Components Analysis

Principal Components Analysis (PCA), like MANOVA, is a multivariate technique and it will be used to help us better understand the relationships among physical factors and fish measurements (Johnson and Wichern 1992). The object of PCA is to produce a shorter list of variables than you started with that are each some linear combination of the original variables. We will attempt to use PCA to generate a shorter list of variables (i.e, principle components) from the full list of variables that we will then try to correlate with fish measurements.

Regression and Correlation Analyses

Again to help us better understand the relationships among physical factors and fish measurements, we will use a stepwise regression approach. We will regress physical factors against fish measurements using a stepwise approach where the most significant physical factors are added to the regression model until no more significant factors are found. We will use correlation estimates among physical factors to help us discern when a factor was not added due to its close correlations with a factor already part of the regression model.

Analyses When Normality and/or Homogeneity of Variance Lacking

Rank ANOVA

The Rank ANOVA (RANOVA) is a method by which non-normal data can be analyzed using parametric statistics (Conover and Iman 1981). For this process the data are ranked within the lowest level of organization which in this study is the gear-habitat level. The ranks are then analyzed using the same models list above for ANOVA, MANOVA, PCA, regression, and correlation.

Permutation Techniques

Another alternative to the RANOVA is to use permutation techniques (Slausen et al. 1994; www.nbs.gov/mesc/swprods.html). Permutation techniques are distribution free (i.e., no assumption of normality). Permutation techniques are based on randomization theory. Permutation techniques test main-effects differences by calculating the probability of obtaining a Euclidean distance less than or equal to the distance represented by the data from the set of all possible combinations of the data.

Analyses When Many Zeros are Present

It should be apparent that the presence of many zeros in the data set will automatically produce non-normal data, but even RANOVA or other non-parametric techniques are not appropriate or necessary. If we assume that our sampling techniques and effort are adequate for collection of the targeted fish present in the system then the zeros produced by our efforts can be considered useful data for comparisons. There are several options for making the comparisons we are after:

- I. Exclusion of gear-habitat and/or segments
 - A. Exclude from the data when the great majority of the data are zeros.
 - B. Analyze the remaining data in the same manner as previously described
 - C. Compare remaining data to zero (i.e., the assumed value of the excluded data)
- II. Presence/Absence as a main effect
 - A. create new variable where gear-habitat level data is classified as present (i.e., fish species being analyzed for was collected) or absent (i.e., fish species being collected for was not collected)
 - B. include presence/absence as a main effect in the models listed above
- III. Summing to a higher level
 - A. sum data for a replicate of the segment (i.e., sum data for one replicate of each gear-habitat collection within a segment) producing five replicate sums for each segment
 - B. analyze these replicate sums among segments

Literature Cited

- Anderson, R.O. and S.J. Gutreuter. 1985. Length, weight, and associated structural indices, p. 283-300. *In* L.A. Nielsen and D.L. Johnson (eds.). *Fisheries Techniques*. American Fisheries Society, Bethesda, Maryland.
- Bhattacharyya, G. K. and R.A. Johnson. 1977. *Statistical Concepts and Methods*. John Wiley and Sons, Inc., New York, and New York.
- Brown, M.L. and D. J. Austen. 1996. Data management and statistical techniques. p. 17-62. *In* B.R. Murphy and D.W. Willis (eds.). *Fisheries Techniques: Second Addition*. American Fisheries Society, Bethesda, Maryland.
- Bunge, J. and M. Fitzpatrick. 1993. Estimating the number of species: a review. *Journal of the American Statistical Association* 88:364-373.
- Carline, R.F. 1986. Assessment of fish populations and measurement of angler harvest: indices as predictors of fish community traits, p. 46-56. *In* G.E. Hall and M.J. Van Den Avyle (eds.) *Reservoir Fisheries Management: strategies for the 80's*. Southern Division, American Fisheries Society, Bethesda, Maryland.
- Chapman, D.W. 1988. Critical review of variables used to define effects of fines in redds of large salmonids. *Transaction of the American Fisheries Society* 117(1):1-21.
- Conover, W.J. and R.L. Iman. 1981. Rank transformation as a bridge between parametric and nonparametric statistics. *The American Statistician* 35(3):124-133.
- Daniel, W.W. 1978. *Applied Nonparametric Statistics*. Houghton Mifflin Company, Boston, Massachusetts.
- Everhart, W.H. and W.D. Youngs. 1981. *Principles of Fishery Science*. Cornell University Press Ltd., London, United Kingdom.
- Funk, J.L. 1957. Relative efficiency and selectivity of gear used in the study of fish populations in Missouri streams, p. 31-43. *In* *Symposium on Evaluation of Fish Populations in warm water streams*. Iowa State College, Ames, Iowa.
- Green, R. 1979. *Sampling Design and Statistical Methods for Environmental Biologist*. John Wiley and Sons, Inc., New York, New York.
- Gutreuter, S. 1993. *A Statistical Review of Sampling of Fishes in the Long Term Resource Monitoring Program*. National Biological Survey Environmental Technical Center, Onalaska, Wisconsin.
- Gutreuter, S., R. Burkhardt, and K. Lubinski. 1995. *Long Term Resource Monitoring Program Procedures: Fish Monitoring*. National Biological Service, Environmental Management Technical Center, Onalaska, Wisconsin. LTRMP 95-P002-1. 42 pp. + Appendixes A-J.
- Hayes, M.L. 1983. Active fish capture methods, p. 123-145. *In* L.A. Nielsen and D.L. Johnson (eds.). *Fisheries Techniques*. American Fisheries Society, Bethesda, Maryland.
- Hubert, W.A. 1983. Passive capture techniques, p. 95-111. *In* L.A. Nielsen and D.L. Johnson (eds.). *Fisheries Techniques*. American Fisheries Society, Bethesda, Maryland.
- Hurlbert, S.H. 1971. The non-concept of species diversity: a critique and alternative parameters. *Ecology* 52:577-586.
- James, F.C. and S. Rathbun. 1981. Rarefaction, relative abundance, and diversity of avian communities. *Auk* 98:785-800.

- Jearld, Jr., A. 1985. Age determination. *In* L.A. Nielsen and D.L. Johnson (eds.). Fisheries Techniques. American Fisheries Society, Bethesda, Maryland.
- Johnson, R.S. and Wichern, D. W. 1992 Applied Multivariate Statistical Analysis. Prentice Hall, Englewood Cliffs, New Jersey.
- Ludwig, J.A. and J.F. Reynolds. 1988. Statistical Ecology: a Primer on Methods and Computing. John Wiley and Sons, Inc., New York, New York.
- McMahon, T.E., A.V. Zale, and D.J. Orth. 1996. Aquatic habitat measurements. *In* B.R. Murphy and D.W. Willis (eds.). Fisheries Techniques: Second Edition. American Fisheries Society, Bethesda, Maryland.
- Orth, D.J. 1985. Aquatic habitat measurements. *In* L.A. Nielsen and D.L. Johnson (eds.). Fisheries Techniques. American Fisheries Society, Bethesda, Maryland.
- SAS. 1990. SAS System. SAS Institute Incorporated, Cary, North Carolina.
- Schaeffer R.L., W. Mendenhall, and L. Ott. 1990. Elementary Survey Sampling. PWS-KENT Publishing Company, Boston, Massachusetts.
- Slausen, W.L., B.S. Cade, and J.D. Richards. 1994. User Manual for the Blossom Statistical Software. Midcontinent Biological Science Center, U.S. Geological Survey, Biological Resources Division, Fort Collins, Colorado.
- Shapiro, S.S. and M.B. Wilk. 1965. An analysis of variance test for normality (complete samples). *Biometrika* 52:591-611.
- Snedecor, G.W. and W.G. Cochran. 1982. Statistical Methods: Seventh Edition. The Iowa State University Press, Ames, Iowa.
- Steel, R.G.D. and J.H. Torrie. 1980. Principles and procedures of statistics: a biometrical approach. McGraw-Hill Book Company, New York, New York.
- Starrett, W.C. and P.G. Barnikol. 1955. Efficiency and selectivity of commercial fishing devices used on the Mississippi River. *Illinois Natural History Survey Bulletin* 26:325-366.
- Weaver, M.J., J.J. Magnuson, and M.K. Clayton. 1993. Analyses for differentiating littoral fish assemblages with catch data from multiple sampling gears. *Transactions of the American Fisheries Society* 122:1111-1119.