Guidance for Site-Specifically Assessing the Health of Fish Populations with Emphasis on Canada’s Environmental Effects Monitoring Program

Kelly R. Munkittrick,1* Timothy J. Barrett,1 Mark E. McMaster2

1 Canadian Rivers Institute, Department of Biology, University of New Brunswick, P.O. Box 5050, Saint John, New Brunswick E2L 4L5, Canada
2 Environment Canada, National Water Research Institute, Burlington, Ontario, Canada

Techniques have been developed over the past two decades to site-specifically assess effects of contaminants on the health of fish populations using a sentinel species approach. National environment effects monitoring (EEM) programs have been implemented in Canada for pulp and paper effluents since 1992 and liquid metal mining effluents since 2002 to monitor effects of these discharges on the health of fish populations. The major criticisms of past EEM fish population surveys can be separated into concerns about the adequacy of the reference sites, the potential impacts of confounding factors, the ecological relevance of endpoints used, the influences of natural variability, concerns over statistical design issues, and potential genetic influences on species characteristics. This paper provides input to deal with these issues and guidance on the selection of sentinel species, timing of sampling, and nonlethal sampling methods to evaluate the health of fish populations. Sample size requirements, effect sizes, and power analysis are also discussed as well as data analysis guidance needed to obtain reliable results.

Key words: impact assessment, industrial effluents, sentinel fish species, fish populations, environmental effects monitoring

Introduction

Environmental monitoring frameworks to assess the impacts of contaminants on fish populations have been developed and improved over several years (Colby 1984; Ryder and Edwards 1985; Munkittrick and Dixon 1989a, 1989b; Munkittrick 1992; Gibbons and Munkittrick 1994; Power 1997; Munkittrick et al. 2000; Sandström et al. 2005). These frameworks evaluate environmental impacts on fish populations by assessing changes in population characteristics such as age structure, energy expenditure, and energy storage relative to reference site(s). The goal of the analysis is not to provide a definitive assessment of impacts or causes of impacts, but rather to document changes from reference conditions over time and determine the focus of follow-up studies. This iterative monitoring framework was used in the development of the adult fish population survey component of Canada’s national environmental effects monitoring (EEM) program for industrial wastewaters (Ribey et al. 2002).

The EEM program is a mandatory, regulated, cyclical monitoring program for the pulp and paper and metal mining industries to assess whether mills and mines in compliance with their effluent discharge regulations are associated with environmental impacts on fish or benthic invertebrates (Walker et al. 2002). The pulp and paper and the liquid metal mining effluent EEM programs are currently in their fifth and second cycles of monitoring, respectively, and more than 300 fish population surveys have been conducted as part of the programs. The fish population survey provides an assessment of whether there are differences in the growth, reproduction, condition, and survival of the fish populations between exposed and reference areas or within an exposure area where there are gradually decreasing effluent concentrations.

The EEM program was designed in a cyclical nature to determine whether changes are present at a sufficient size such that additional, more detailed studies are needed. The design of EEM is such that this determination requires confirmation in a subsequent cycle of monitoring, and moving to more detailed monitoring or examination requires that effects exceed a critical effect size (CES). To evaluate the effects of stressors on fish populations, the following questions need to be addressed: Is there an effect? Is the effect stressor related? Is the magnitude and extent of the effect known? Is the stressor-related cause of the effect known?

The major criticisms of past EEM studies (and field studies in general), can be separated into concerns about the adequacy of the reference sites, the potential impacts of confounding factors, the ecological relevance of endpoints used, the influences of natural variability, concerns over statistical design issues, and potential genetic influences on species characteristics. This paper will provide input to deal with these issues and discuss study design considerations, including guidance on the selection of sentinel species, timing of sampling, sample size requirements, nonlethal sampling methods, and data analysis.

*Corresponding author; krm@unbsj.ca
Monitoring Level and Endpoints Used

Any “impact” at a biochemical, individual, population, or community level has to be evaluated in terms of the consequences of the change as it affects the sustainability of the ecosystem. Individual organisms may survive numerous biochemical impacts and populations may survive numerous individual impacts. Designing a monitoring program at the fish community level puts constraints on adaptive management (an iterative process of decision making that uses the results of previous studies to provide feedback on whether ecosystem objectives are achieved). Fish community changes are ecologically relevant (but have a long time lag before they are detectable), may be difficult to reverse, and may be the result of changes that are difficult to specifically define. At the other extreme, biochemical responses happen quickly and can more often be related to specific causes, but may have little ecological relevance and be easily reversed. It is important that whatever level is focused on, that the assessment approach balances protection and detection, with reversibility and relevance. The system needs to give enough warning that there is time to respond, but still relates to relevant effects.

Monitoring fish health at the population level has several advantages. The population level approach offers a compromise between the sensitivity and reversibility of biochemical approaches and the relevance of community level endpoints. Monitoring at the community level can potentially miss irreversible, important effects at the population level. Delays in sexual maturity, altered growth, changes in fecundity, and depressions in storage of energy reserves put fish at risk, and knowing this level of risk is important to the management of ecosystems.

In the EEM program, five indicators of fish health or effect endpoints are used to assess fish population health. These are age (indicator of population survival), weight-at-age and relative gonad weight (indicators of energy use), and relative liver weight and condition (indicators of energy storage). Other data such as fecundity and egg weight are also collected and are used as supporting data. It is important to emphasize that the program is iterative and that the outcome of comparisons is to emphasize areas of focus for subsequent monitoring cycles. For example, impacts showing significant decreases in growth, gonad size, liver size, and condition would emphasize challenges with food limitation and suggest increased emphasis on evaluating food and habitat availability (Table 1) (the EEM program also has components requiring assessment of fish habitat and fish use).

Mean age is meant to give an assessment of the relative ages of the reference and exposed fish populations. If size-selective gear such as gillnets are used, and there is a significant difference in mean ages of fish sampled at both sites with identical gear, the difference points to a need in the subsequent cycles to further investigate the population and the reason for the difference. Methods of aging should be consistent at each sampling area and among cycles, and appropriate quality assurance/quality control procedures followed (e.g., independent confirmation). If fish cannot be aged reliably or if it is not cost/time efficient, the age can be determined by using size-frequency distributions. This may be especially useful when sampling small-bodied fish species or when conducting nonlethal sampling techniques. It may also be possible to confirm the size-frequency distributions by aging representative subsamples from each size class. See Nielsen and Johnson (1983) for more information on size-frequency distributions.

Both size-at-age (growth) and the reproductive measures are meant to give an assessment of the ability of the fish to utilize the food available to them. Growth

<table>
<thead>
<tr>
<th>Age distribution</th>
<th>Energy utilization</th>
<th>Energy storage</th>
<th>Generalized pattern</th>
<th>Cause of changes</th>
<th>Follow-up study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shift to younger</td>
<td>Increased</td>
<td>Increased</td>
<td>Exploitation</td>
<td>Decreased competition between adults associated with mortality or eutrophication</td>
<td>Examine food resource availability and population density</td>
</tr>
<tr>
<td>Shift to older</td>
<td>No change</td>
<td>No change</td>
<td>Recruitment failure</td>
<td>Shift to older age classes associated with decreased reproductive success</td>
<td>Detailed examination of spawning habitat and its utilization, and reproductive development</td>
</tr>
<tr>
<td>Shift to older</td>
<td>Decreased</td>
<td>Decreased</td>
<td>Multiple stressors</td>
<td>Simultaneous impacts on food availability and reproductive success</td>
<td>Detailed studies of reproductive development and food resources</td>
</tr>
<tr>
<td>No change</td>
<td>Decreased</td>
<td>Decreased</td>
<td>Food limitation</td>
<td>Increased competition associated with increased reproductive success or decreased food availability</td>
<td>Examine food resource availability and population density</td>
</tr>
<tr>
<td>No change</td>
<td>Decreased</td>
<td>No change</td>
<td>Niche shift</td>
<td>Modest increase in competition for forage base</td>
<td>Examine food base and competition aspects</td>
</tr>
<tr>
<td>Shift to younger</td>
<td>Mixed</td>
<td>Mixed</td>
<td>Metabolic redistribution</td>
<td>Inability to maximally utilize available food resources</td>
<td>Detailed physiological studies of energetics</td>
</tr>
<tr>
<td>Shift to older</td>
<td>Increased or decreased</td>
<td>Increased</td>
<td>Chronic recruitment failure</td>
<td>Shift to small population of older individuals</td>
<td>Detailed study of reproductive performance</td>
</tr>
<tr>
<td>No change</td>
<td>No change</td>
<td>No change</td>
<td>Null response</td>
<td>No obvious changes</td>
<td>Check population size data to see if carrying capacity of the system has changed</td>
</tr>
</tbody>
</table>
is the change in size (weight or length) with time or age. In the case of growth, it may be helpful to collect information on other age classes. For example, it may be important to determine whether there are changes in growth of early life stages. This will assist in determining the magnitude of the effect. Subsequent monitoring should focus on confirming responses detected, and examining the relevance of the changes to other size classes and species.

Reproduction can be expressed as reproductive effort, fecundity, egg weight, or gonad weight relative to body size. Reproduction may be the most sensitive measurement in resident fish, and is of high ecological relevance. Changes in reproductive investment can be evident within a year (Bowron et al. 2009) since the reproductive tissue is generally turned over annually. Fecundity and gonad weight are easy to measure if an appropriate sampling time is chosen (Barrett and Munkittrick 2010). In terms of a change in gonad size, additional work related to magnitude could be focused on determining whether the change occurs at other times of the year (for multiple spawners). Studies related to magnitude could also focus on asking whether the changes are present in other species in the same exposure area.

Measures of stored energy reserves provide valuable information on the efficiency, availability, and quality of food available to the fish. The EEM program uses condition (length to weight relationships) and liver size as indicators for this information. As with other indicators, the consistency in response between indicators is important. Liver size can increase for several reasons, including storage of lipids and glycogen, and enhanced detoxification activity.

Analysis of data shows that there is a high amount of overlap in the information provided by condition and liver size (Fig. 1). Multidimensional scaling ordination of the benthic and fish responses from Cycle 2 of the pulp and paper EEM clearly show that fish population endpoints and benthic endpoints in the program give different information, as evidenced by the differences in the orientation of the principle component correlation vectors. It also shows that gonad size and age pull in opposite directions (i.e., smaller gonads pull in the same direction as younger-aged fish, and vice versa). Differences in the relative responses can be used to highlight response patterns (Hewitt et al. 2005).

**Basic Study Design Concerns**

**Reference Sites**

The most common study designs for fish population health assessments are control/impact designs and gradient designs. It is now common in these assessments to use multiple reference sites. Including additional reference sites increases the ability to evaluate issues related to natural variability, ecological relevance, and confounding factors, and improves the ability to evaluate the adequacy of the chosen reference site(s). However, the increased cost of using multiple reference sites has limited the implementation of better study designs. Over the first three cycles of monitoring in the pulp and paper EEM program, there has been a trend towards using more reference sites. In Cycle 1, 3% of studies used multiple reference sites, in Cycle 2, 9%, and in Cycle 3, 25% used multiple reference sites.

The choice of reference area is a typical criticism of studies. Ideally, the reference site in riverine assessments would be located upstream of the stressor, in similar habitat, free of confounding influences, with a natural barrier that limits movement between sites, but this situation is seldom available. There are several main issues involved in the selection of a reference site, including whether the site is comparable in terms of habitat, is free from the stressor of concern (exposure) and from confounding influences, or is open to movement of fish from the exposure site (fish in an upstream reference area could have been exposed previously or fish in the exposure area could be transient, reducing exposure to potential effects).

Studies that use a gradient approach and multiple reference sites are stronger than studies that depend on a single reference site. There are several new approaches to trying to deal with problems with reference sites, including reference condition approaches (Bailey et al. 1998) and using negative reference sites (using the exposed site as your reference; Vallières et al. 2007). Regardless, the existence of consistent changes over time increases the level of confidence that the changes are real. Follow-up studies must evaluate the adequacy of the reference site(s) especially if consistent results are not found.

**Fig. 1.** Multidimensional scaling ordination of Cycle 2 EEM pulp and paper fish and benthic invertebrate effect endpoints with principle component correlation vectors added. Fish endpoints: Age; G = gonad weight; K = condition; L = liver weight; WA = weight-at-age. Benthic endpoints: Ab = abundance; BC = Bray-Curtis Index; E = evenness; TR = taxon richness. Figure adapted from Lowell et al. (2003).
Confounding Factors

In the second cycle of Canada’s pulp and paper EEM program, almost 80% of the studies that concluded there were effects also recorded the presence of confounding factors in their interpretation of effects (Lowell et al. 2003). A long list of potential confounding factors exist at most sites, including other outfalls, habitat changes, historical uses and contamination, tributaries and nonpoint source inputs, among many other factors. In highly confounded situations, alternative methods should be considered, but it should be emphasized that it is possible to get interpretable field results at most sites with adjustments to the study design. There is no consensus as to what constitutes sufficient data to demonstrate that other discharges or contaminant sources are primarily responsible for observed changes or absence of observed changes. If changes are seen and thought to be potentially contributed to by confounding influences, the objective of subsequent study designs should be to eliminate the confounding influences as being a significant contributor to the issue.

Ecological Relevance of Endpoints Used

There are ongoing concerns expressed about the ecological relevance of the endpoints used in fish population health assessments. Changes in indicators of growth, reproductive performance, age distributions, and body condition are of ecological relevance, but the concern expressed by some is that the relevance of changes in these indicators to the population levels is unknown. While it is possible to mathematically model the population consequences (e.g., Miller et al. 2007; Watanabe et al. 2007), such modelling efforts ignore site-specific factors that determine the population-level consequences of change, including mobility, refugia, species interactions, and seasonal changes in environmental conditions.

Decisions about measurement endpoints require consideration of factors related to the time scale of responses, their reversibility, sensitivity, and their ability to be linked to causative agents. There is general agreement that the loss of an important fish species or the presence of contamination that affects the consumption potential of fisheries resources are changes that are important to people. The retrospective assessment process described here is meant to provide some information on how close we might be to thresholds where we need to become concerned before an important species is lost.

Decisions regarding monitoring level (community, population, individual, biochemical) are a compromise between conflicting concerns over ecological relevance and ease of detection with those of time scale of response, reversibility, and ability to indentify cause. While many jurisdictions and studies focus at the community level (e.g., Hall et al. 2009; Karels and Niemi 2009), the Canadian EEM program did not focus on this level because of a variety of issues, including the difficulty in nationally standardizing an approach, the iterative nature of the program, effort and cost of good community programs, and analytical concerns over the seasonality and ability to isolate the causes of any changes detected.

In Canada, the Fisheries Act requires that we protect fish, fish habitat, and human use of the fisheries resource. The EEM program addresses those three aspects using the adult fish survey (discussed here), an invertebrate community survey (to assess fish habitat), and fish tissue studies (human use of fisheries). Within the program, additional methods are incorporated following the confirmation of effects of the discharge in the investigation of cause and solutions portion of the program or for more research-related studies.

Other jurisdictions (such as Sweden) incorporate impacts at biochemical, individual, population, and community levels (Swedish EPA 1997). This was not done in Canada because of the difficulty in incorporating effect sizes and thresholds at the biochemical level for all of the more than 60 species (Barrett and Munkittrick 2010) that have been used across the country, and at the community level because of the reasons outlined above.

Natural Variability

Natural variability is the tendency for endpoint values to change spatially and/or temporarily from nonanthropogenic causes. Natural variability is contributed to by a number of components, including annual variability in food availability, habitat quality (flow, temperature, etc.), as well as annual variability attributable to sampling design (sampling gear or bias, changes in personnel and training, measurement error, or equipment performance).

There is some overlap and confusion regarding the concerns about adequacy of the reference site, ecological relevance, and natural variability. Natural variability refers to the chance that a specific site difference relative to a specific reference site would reverse in a different year. Concerns about the adequacy of the reference site are better described by concerns about whether an effect reflects a true difference associated with exposure to the stressor(s) of interest, and the chance that a different interpretation would occur with the use of a different reference site. The issue of ecological relevance (within this context of variability) refers to the size of the effect relative to the values that could be seen between a variety of reference sites. A difference smaller than what is seen between numerous reference sites is not interpreted to mean that the site-specific difference is not real relative to a local comparable reference site, but it does provide input into the ecological importance and the relevance of the difference.

Statistical Design Issues

Statistical design issues can be divided into three main criticisms reflecting concerns over the number of comparisons made, power analysis, and pseudoreplication.
Site-Specific Assessing Fish Population Health

There have been many expressed concerns about multiple comparisons. In the EEM program a significance level of \(\alpha = 0.05\) is used at the mill or mine level for each comparison. There were several safeguards put into place: progression to a new tier of monitoring only occurs if an endpoint responds above a CES, that effect is confirmed in a subsequent cycle of monitoring (three years later) in the same direction, and that related endpoints are responding together in an interpretable pattern (eutrophication, food limitation, metabolic disruption). Furthermore, analyses of responses have shown that responses are usually confirmed in the opposite sex, or a second species (two fish species are required for monitoring in the EEM program), and that responses seldom occur individually (<8% of the time).

There has been some debate over the levels of \(\alpha\) (the probability of committing a type I error) and \(\beta\) (the probability of committing a type II error) in EEM studies. In earlier cycles of EEM, \(\alpha\) was set at 0.05 and \(\beta\) was set to 0.20 (equivalent to a power level of \(1 - \beta = 0.80\)). The EEM program now recommends that \(\alpha\) and \(\beta\) equal one another, letting the risk to the environment (probability of committing a type II error) equal the risk to industry (probability of committing a type I error). If values are set at \(\alpha = \beta = 0.10\), the sample sizes required to detect the same effect are approximately the same as when \(\alpha = 0.05\) and \(\beta = 0.20\). Where possible it is encouraged to reduce \(\alpha = \beta = 0.05\) (the traditional level for \(\alpha\)). These recommendations are to help ensure that studies are designed to provide a reasonably high probability of statistically detecting a predetermined effect size (ES) if it has occurred (i.e., the power of the test \((1 - \beta)\) should be high).

Briefly, pseudoreplication (Hurlbert 1984) arises because it is not possible to randomly assign field sites, and the exposure treatment (exposure site) cannot be replicated. In a river system, it has been shown that, where possible, upstream and downstream data are essential for interpreting the effects of the effluent (Munkittrick et al. 2000). But in these situations, the reference site must be upstream and the exposed site must be downstream (except in unique cases). Since the sites cannot be randomly assigned, there may be factors that have not been controlled (i.e., factors other than timing, sample gear, habitat type etc.) that may account for the differences between sites. Also if a significant difference is detected, this provides evidence that there is a difference between two locations (a site downstream of the effluent and a site upstream). This difference cannot be interpreted as a difference resulting from the stressor of interest, but only as a difference between two specific locations.

EEM studies focus on trying to develop an understanding of that specific reach, so that site-specific concerns can be identified, and necessary changes can be made on a local basis. Monitoring is conducted over time and consistent significant responses over time are used to direct more focused monitoring to uncover the course of the impact. The ability to extrapolate the results in most studies is not important. Unless there is another site with identical stressors and similar habitats, extrapolation will seldom be possible. It is not possible with data from a single site to make wide conclusions. It must be restricted to comment on one situation and design hypotheses to test at others. The best approach is based on developing an iterative understanding (Hodson et al. 1996). The conclusion of a round of monitoring should only be used to design subsequent monitoring steps which ultimately lead to uncovering cause, thus the design must be iterative. The repeated observation of consistent changes over a longer time period will increase confidence that the changes are real and interpretable as differences between sites. An observed change of decreased food availability at a site can equally be interpreted as increased food availability at the reference site. Iterative studies must be designed and used to focus studies to test the validity of the conclusion.

Wide conclusions about the impacts require analysis of data from multiple sites (e.g., Lowell et al. 2005). With pulp and paper impacts, the studies were initiated in Canada at a single mill (Munkittrick et al. 1991), expanded to eight mills (Munkittrick et al. 1994), and then went to data from 65 mills (Munkittrick et al. 2002; Lowell et al. 2005). Meaningful information and industry- and country-wide conclusions can be generated with data from 65 mills, but many monitoring programs will need to make decisions based on a minimal number of reference sites.

Potential Genetic Influences

Potential genetics issues include whether the absence of a difference between sites represents genetic adaptation, whether differences between populations are genetically-based and natural, and whether reductions in genetic variability are an important response on their own. If there is no detectable difference between sites, then for the purposes of evaluating sustainability, the conclusion is that the situation is sustainable, regardless of the mechanism that led to it. If differences exist that are genetically based, the iterative sampling program will eventually lead to a conclusion that the differences between sites are independent of the stressor, and that the situation for the population is sustainable.

The issue of reducing the genetic variability of exposed populations is real (e.g., Bell and Collins 2008) and is thought to represent increased vulnerability of the population. The issue is whether the situation may be close to an unidentified threshold that is only evident through biochemical or chemical measurements, and this issue of potentially more significant lower thresholds is currently beyond the scope of EEM.

Selection of Species and Timing of Sampling

The most important factors when selecting fish species for assessing the health of fish populations are exposure,
abundance, relevance to the study area (Munkittrick et al. 2000), and sensitivity to the effluent. The recommended method for carrying out a fish population survey is to monitor adults (sexually mature) of relatively sedentary finfish that have been exposed to effluent over a long period of time. In selecting a species, the species selected for previous population health studies should be considered if possible, and preference should be given to: resident (nonmigratory) fish species identified in site characterization, sexually mature female and male fish species that are abundant in both the exposure and reference areas, fish species for which fishing or sampling permits can be obtained, and fish species that have the highest exposure to effluent.

Some of the challenges related to species selection relate to the attempt to design a single program for multiple purposes. It is very difficult to accomplish this. Concerns about contamination of fishery resources for human consumption would direct the study design to collect a species that is long-lived (so that contaminants can accumulate over a longer timeframe), is piscivorous (so that biomagnification is higher), matures late (to increase concentration), preferably focuses on male fish or species that do not spawn every year (so that elimination of contaminants through egg deposition is lessened), and are of importance for local consumption. These characteristics are exact features that decrease the sensitivity for detecting environmental impacts, where the preference is for species that are benthic (because generally they are less mobile), are not commercially or recreationally important (because it obscures determining cause), mature early, contribute a lot of energy to reproduction (so that energy demands are high), and are short-lived (so impacts are recent), and the focus is on female fish (environmental impacts are often more serious on female egg producers since female gametes are typically limited relative to male gametes).

There are a number of other factors that need to be considered when selecting a sentinel species (see Munkittrick and McMaster 2000; Munkittrick et al. 2000), including ensuring that the species are active participants in the local aquatic food web. Other life history characteristics, like spawning time and migration, need to be evaluated site-specifically because the interaction between discharge site, spawning habitats, seasonal changes in flow, and dilution all play a role in influencing the importance of the characteristics and how they potentially impact the sensitivity of the monitoring program.

The trend towards the increasing use of forage fish species (Munkittrick et al. 2002) has continued in the EEM pulp and paper program, rising from their use in 10% of surveys in Cycle 1, to 26% in Cycle 2, and 34% in Cycle 3. Their use has several advantages and disadvantages. On a practical level, small-bodied fish species are usually more abundant, easy to capture, and more sedentary than larger-bodied fish species. Shuter (1990) noted that small-bodied fish species were more sensitive to acidification than larger species, or at least responded more quickly to changes in pH. In freshwater fish, home range size has been positively correlated with body size (Minns 1995), and many small-bodied species integrate local conditions very well. On the other hand, they require more sensitive analytical balances, more careful measurement, and are more sensitive to microhabitat differences because they integrate the local habitat so well. They are also more sensitive to differences in timing of sampling (Barrett and Munkittrick 2010).

The suitability of fish as sentinels depends on a number of factors, but the timing of sampling is very important. A variety of factors needs to be considered, including potential migratory behaviour of the sentinel species, water conditions (flow, turbidity, wave action), accessibility, and the cycle of gonadal development for the sentinel species. If historical data exists, it would be useful to examine that data and, if appropriate, conduct the survey during similar periods so that the surveys can be compared. Similarly, subsequent monitoring should be conducted during similar periods of the year to be comparable if that sampling time makes sense for the study. Barrett and Munkittrick (2010) have developed species-specific sampling times for the more than sixty different species that have been used in EEM studies in Canada. The sampling times are based on the reproductive strategies and cycle of gonadal development for each species and are provided in Table 2 along with some species-specific aspects that may affect study designs.

Nonlethal Sampling Methods

Nonlethal sampling of fish can also be used to evaluate the effects of stressors on fish populations. Sample size requirements for these studies are recommended to be a minimum of 100 fish older than young-of-the-year (YOY), and an additional 100 YOY fish from each study site. It is usually possible to separate YOY from older age classes by size distributions; however, this may not be possible for species with extended spawning periods. If YOY abundance is extremely high (>80 to 90%), the proportion of fish that are YOY should be estimated from the first 100 fish collected, and then the collection can continue concentrating only on collecting 100 larger individuals for calculating size distributions of older fish. The fish collected that are older than YOY should represent the whole range of fish sizes and be representative of the population (mature and immature). When YOY are abundant and constitute a high relative proportion of the population, there will not be sufficient information collected on all other size/age classes. In the latter two situations, the collection of the additional non-YOY fish allows for a higher discrimination of the older fish classes to be achieved. The use of this recommended sample size in each area will give a good idea of the population distribution when plotting endpoints such as the length-weight frequency. As well, when examining differences between the relative abundance of young
TABLE 2. Commonly used fish species in past environmental effects monitoring studies; aspects to consider during study design and recommended sampling times

<table>
<thead>
<tr>
<th>Family</th>
<th>Species</th>
<th>Scientific name</th>
<th>Considerations</th>
<th>Sample time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Salmonids</td>
<td>Lake trout</td>
<td>Salvelinus namaycush</td>
<td>F M *</td>
<td>4–6</td>
</tr>
<tr>
<td></td>
<td>Brook trout</td>
<td>Salvelinus fontinalis</td>
<td>F M A</td>
<td>4–6</td>
</tr>
<tr>
<td></td>
<td>Arctic char</td>
<td>Salvelinus alpinus</td>
<td>F M D</td>
<td>4–6</td>
</tr>
<tr>
<td></td>
<td>Dolly varden</td>
<td>Salvelinus malma</td>
<td>F M?</td>
<td>4–6</td>
</tr>
<tr>
<td></td>
<td>Bull trout</td>
<td>Salvelinus confluentus</td>
<td>F M?</td>
<td>4–6</td>
</tr>
<tr>
<td></td>
<td>Cutthroat trout</td>
<td>Salmo clarki</td>
<td>W M * A</td>
<td>LF</td>
</tr>
<tr>
<td></td>
<td>Rainbow trout</td>
<td>Oncorhynchus mykiss</td>
<td>S/M A</td>
<td>LF</td>
</tr>
<tr>
<td></td>
<td>Arctic grayling</td>
<td>Thymallus arcticus</td>
<td>S M A</td>
<td>LF</td>
</tr>
<tr>
<td></td>
<td>Mountain whitefish</td>
<td>Prosopium williamsoni</td>
<td>F M?</td>
<td>4–6</td>
</tr>
<tr>
<td></td>
<td>Round whitefish</td>
<td>Prosopium cylindraceum</td>
<td>F/W M?</td>
<td>4–6</td>
</tr>
<tr>
<td></td>
<td>Lake whitefish</td>
<td>Coregonus clupeaformis</td>
<td>F M *</td>
<td>4–6</td>
</tr>
<tr>
<td></td>
<td>Cisco</td>
<td>Coregonus artedii</td>
<td>F</td>
<td>4–6</td>
</tr>
<tr>
<td>Hiodontids</td>
<td>Goldeye</td>
<td>Hiodon alosoides</td>
<td>S M P H</td>
<td>LF</td>
</tr>
<tr>
<td></td>
<td>Mooney</td>
<td>Hiodon tergisus</td>
<td>S M?</td>
<td>LF</td>
</tr>
<tr>
<td>Esocids</td>
<td>Northern pike</td>
<td>Esox lucius</td>
<td>S M</td>
<td>LF</td>
</tr>
<tr>
<td>Cyprinids</td>
<td>Carp</td>
<td>Cyprinus carpio</td>
<td>X M *</td>
<td>4–6</td>
</tr>
<tr>
<td></td>
<td>Fallfish</td>
<td>Semotilus corporalis</td>
<td>S/X?/G</td>
<td>4–6</td>
</tr>
<tr>
<td></td>
<td>Creek chub</td>
<td>Semotilus atromaculatus</td>
<td>S/X *</td>
<td>4–6</td>
</tr>
<tr>
<td></td>
<td>Peamouth chub</td>
<td>Mylocheilus caurinus</td>
<td>S</td>
<td>4–6</td>
</tr>
<tr>
<td></td>
<td>Lake chub</td>
<td>Cottus plumbeus</td>
<td>S M</td>
<td>4–6</td>
</tr>
<tr>
<td></td>
<td>Longnose dace</td>
<td>Rhinichthys cataractae</td>
<td>X</td>
<td>4–6</td>
</tr>
<tr>
<td></td>
<td>Blacknose dace</td>
<td>Rhinichthys atratulus</td>
<td>X</td>
<td>4–6</td>
</tr>
<tr>
<td></td>
<td>Pearl dace</td>
<td>Margariscus margarita</td>
<td>X?</td>
<td>4–6</td>
</tr>
<tr>
<td></td>
<td>Redbelly dace</td>
<td>Phoxinus espinus</td>
<td>A</td>
<td>SPAW</td>
</tr>
<tr>
<td></td>
<td>Spottail shiner</td>
<td>Notropis hudsonius</td>
<td>X? M?</td>
<td>4–6</td>
</tr>
<tr>
<td></td>
<td>Mimic shiner</td>
<td>Notropis volucellus</td>
<td>S? 2-yr</td>
<td>4–6</td>
</tr>
<tr>
<td></td>
<td>Emerald shiner</td>
<td>Notropis atherinoides</td>
<td>X? M?</td>
<td>4–6</td>
</tr>
<tr>
<td></td>
<td>Blacknose shiner</td>
<td>Notropis heterolepis</td>
<td>X</td>
<td>4–6</td>
</tr>
<tr>
<td></td>
<td>Common shiner</td>
<td>Luciulus cornutus</td>
<td>S?</td>
<td>4–6</td>
</tr>
<tr>
<td></td>
<td>Golden shiner</td>
<td>Notemigomus crysoleucus</td>
<td>X?</td>
<td>4–6</td>
</tr>
<tr>
<td></td>
<td>Redside shiner</td>
<td>Richardsonius balteatus</td>
<td>X?</td>
<td>4–6</td>
</tr>
<tr>
<td></td>
<td>Bluntnose minnow</td>
<td>Pimephales notatus</td>
<td>A/G</td>
<td>SPAW</td>
</tr>
<tr>
<td></td>
<td>Fathead minnow</td>
<td>Pimephales promelas</td>
<td>A/G</td>
<td>SPAW</td>
</tr>
<tr>
<td>Catostomids</td>
<td>White sucker</td>
<td>Catostomus commersoni</td>
<td>S M</td>
<td>LF</td>
</tr>
<tr>
<td></td>
<td>Longnose sucker</td>
<td>Catostomus catostomus</td>
<td>S M</td>
<td>LF</td>
</tr>
<tr>
<td></td>
<td>Largescal sucker</td>
<td>Catostomus macrolechilus</td>
<td>S M?</td>
<td>LF</td>
</tr>
<tr>
<td></td>
<td>Bridgelip sucker</td>
<td>Catostomus columbiae</td>
<td>S M?</td>
<td>LF</td>
</tr>
<tr>
<td></td>
<td>Shorthead redhorse sucker</td>
<td>Mozostoma macrolepidotum</td>
<td>S M</td>
<td>LF</td>
</tr>
<tr>
<td></td>
<td>Silver redhorse sucker</td>
<td>Mozostoma anisurum</td>
<td>S M?</td>
<td>LF</td>
</tr>
<tr>
<td>Ictalurids</td>
<td>Brown bullhead</td>
<td>Icterus nebulossus</td>
<td>S/G</td>
<td>4–6</td>
</tr>
<tr>
<td></td>
<td>Channel catfish</td>
<td>Icterus punctatus</td>
<td>S/G D</td>
<td>4–6</td>
</tr>
<tr>
<td>Fundulids</td>
<td>Mummichog</td>
<td>Fundulus heteroclitus</td>
<td>A R</td>
<td>SPAW</td>
</tr>
<tr>
<td>Gadids</td>
<td>Burbot</td>
<td>Lota lota</td>
<td>W M? H</td>
<td>LF</td>
</tr>
<tr>
<td>Atherinids</td>
<td>Atlantic silverside</td>
<td>Menidia menidia</td>
<td>S 1-yr</td>
<td>SPAW</td>
</tr>
<tr>
<td>Gasterosteids</td>
<td>Brook stickleback</td>
<td>Gasterosteus acustans</td>
<td>X</td>
<td>SPAW</td>
</tr>
<tr>
<td></td>
<td>Threespine stickleback</td>
<td>Gasterosteus aculeatus</td>
<td>X M *</td>
<td>SPAW</td>
</tr>
<tr>
<td></td>
<td>Ninespine stickleback</td>
<td>Pungitius pungitius</td>
<td>X M *</td>
<td>SPAW</td>
</tr>
</tbody>
</table>

TABLE 2 continued on next page
versus mature fish, fairly good resolution is achieved (Gray et al. 2002).

Fish for nonlethal sampling should be measured for length and weight, and external sex determination should be made if possible (Gray and Munkittrick 2005). If only adults are used, the priority should be to sample prior to or at the start of the spawning season. However, if YOY are to be collected, the timing should move to the late fall when it will be easier to measure YOY for most species (spring spawners).

The size distribution should be examined as a surrogate for differences in age. If a site difference is present, subsequent monitoring should focus on understanding the difference and possible causes. Size distributions can be analyzed by the Kolmogorov-Smirnov test, although this test is very conservative. There are challenges to using age information on many short-lived species of fish. If a fish only lives two or three years, it will not be possible to measure a 25% difference in mean age (CES).

### TABLE 2 continued

<table>
<thead>
<tr>
<th>Percopsidae</th>
<th>Trout-perch</th>
<th>Percopsis omiscomaycus</th>
<th>X M *</th>
<th>4–6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Centrarchidae</td>
<td>Rock bass</td>
<td>Ambloplites rupestris</td>
<td>S/G</td>
<td>4–6</td>
</tr>
<tr>
<td>Pumpkinseed sunfish</td>
<td>Lepomis gibbosus</td>
<td>S/G</td>
<td>4–6</td>
<td></td>
</tr>
<tr>
<td>Smallmouth bass</td>
<td>Micropterus salmoides</td>
<td>S/G</td>
<td>4–6</td>
<td></td>
</tr>
<tr>
<td>Percidae</td>
<td>Walleye</td>
<td>Sander vitreus</td>
<td>S M</td>
<td>LF</td>
</tr>
<tr>
<td>Yellow perch</td>
<td>Percina flavescens</td>
<td>S M * H D</td>
<td>LF</td>
<td></td>
</tr>
<tr>
<td>Iowa darter</td>
<td>Etheostoma exile</td>
<td>S/G M *</td>
<td>4–6</td>
<td></td>
</tr>
<tr>
<td>Johnny darter</td>
<td>Etheostoma nigrum</td>
<td>S/G M *</td>
<td>4–6</td>
<td></td>
</tr>
<tr>
<td>Logperch</td>
<td>Percina caprodes</td>
<td>S M *</td>
<td>4–6</td>
<td></td>
</tr>
<tr>
<td>Cottidae</td>
<td>Mottled sculpin</td>
<td>Cottus bairdii</td>
<td>S/G LM</td>
<td>4–6</td>
</tr>
<tr>
<td>Slimy sculpin</td>
<td>Cottus cognatus</td>
<td>S/G R LM</td>
<td>4–6</td>
<td></td>
</tr>
<tr>
<td>Torrent sculpin</td>
<td>Cottus rhothous</td>
<td>S/G LM</td>
<td>4–6</td>
<td></td>
</tr>
<tr>
<td>Spoonhead sculpin</td>
<td>Cottus ricei</td>
<td>S/G LM</td>
<td>4–6</td>
<td></td>
</tr>
<tr>
<td>Shortnose sculpin</td>
<td>Myxocyprhus scorpius</td>
<td>W/G</td>
<td>LF</td>
<td></td>
</tr>
</tbody>
</table>

*First letter or characters separated by a slash(es) /) represent “SPAWNING,” and the subsequent characters represent “MIGRATION” and “OTHER” considerations as follows:

**SPAWNING**

S = spring or early summer spawners, usually a single spawning event per year, usually migrate to spawn.

F = fall spawners, usually single spawners.

W = winter spawners, usually single spawners.

X = Multiple spawners – usually summer spawners, and need to worry about whether small fish are reproducing less (or more) than larger fish (view GSI versus length), variability will increase as approach spawning season.

G = males are guarders – reproductive investment in males will be low, but are likely not very mobile.

A = asynchronous spawners.

? = some doubt about number of spawns.

**MIGRATION**

M = Mobile – are known to migrate considerable distances (or change habitats significantly) for spawning at some sites.

* = may be site-specific.

M? = probably migrate.

A = anadromous forms may be present, and may be much larger and much more mobile.

C = Catadromous, migrate out as adults.

R = research available to show that they are usually resident.

**OTHER**

P = plankton feeder – migratory feeder.

1-yr = 1-year life cycle.

2-yr = 2-year life cycle.

H = juvenile and mature fish may utilize very different habitats.

LM = females may start to mature very late in the year, fish may spawn during freshet – can be difficult to get mature female gonads.

D = diet changes within the size of sexual maturity to piscivorous at larger sizes (when available)

*Sampling times adapted from Barrett and Munkittrick (2010); 4–6 = 4 to 6 weeks prior to spawning; LF = late fall (before ice cover); SPAW = as close to the first spawning event of the season as possible.
It should be possible at most sites to get an estimate of growth and reproductive success using nonlethal methods. Growth can be evaluated by the size of YOY at the end of the growing season, and by the size of fish older than YOY. A comparison of the size of YOY fish gives a good indicator of growth since it is a direct indicator, in comparison with size-at-age which is indirect. Differences between sites in spawning times will be integrated into this endpoint. It is also possible to get a growth estimate by a shift in size distributions over time (i.e., repeating measurements two months apart at the same sites), or differences in average size. If the fish species chosen is externally sexually dimorphic, it is possible to examine whether there are gender-specific differences in growth rate.

Reproductive success can be assessed using relative age class strength or by the relative abundance of YOY individuals (Gray et al. 2002). A length frequency distribution may be plotted as a surrogate of an age frequency distribution. Size frequency analysis can be used to examine age distributions, size-at-age data, and condition factors for the fish (Gray et al. 2002). It is recommended that, if possible, aging structures be sampled from a subsample of each size class for situations where age may need to be verified. In slimy sculpin (Cottus cognatus), it has been found that there is rapid growth of YOY fish in the spring, which can cause some overlap with fish older than YOY, making resolution difficult (Gray et al. 2002). Thus, length frequency distributions may be easier to make on late summer and early fall data.

Condition factor can also be evaluated by the relationship 
\[ K = 100,000 \times \left( \frac{\text{weight}}{\text{length}^3} \right) \] 
for the fish examined (when weight is in grams and length is in millimetres). A large number of areas can typically be sampled using this approach, and it is encouraged to sample multiple exposure and reference areas.

### Sample Size Requirements, Effect Sizes, and Power

Munkittrick (1992) suggested that 15 to 25 female fish of a properly selected species and size range will yield sufficient information to characterize a population. Data from more than 300 fish population surveys from the EEM pulp and paper program have been examined to determine the statistical power to detect a 25% CES in the EEM effect endpoints (10% difference for the condition endpoint) when sample sizes were between 15 to 25. The distributions of observed power levels by endpoint are provided in Fig. 2. In general, analyses of condition, relative liver weight, and weight-at-age had high power levels to detect an effect equal to the respective CES. The relative gonad weight and age comparisons are typically more variable than the other endpoints and thus these comparisons have lower power to detect effects. The proportions of comparisons with power levels greater than 0.8 by endpoint are 78, 58, 49, 61, and 42% for condition, relative liver weight, relative gonad weight, weight-at-age, and age, respectively. Based on these power levels, a sample size of 15 to 25 per sex and species seems reasonable for a preliminary evaluation, and variability from preliminary studies can be used in power analyses to determine sample sizes required in follow-up studies.

An extensive literature review has shown that CESs which have been defined in other programs are often consistent with a CES of around 25% or two standard deviations for many biological or ecological monitoring endpoints, and this value appears to be reasonable for use in a wide variety of monitoring programs and with a wide variety of endpoints (Munkittrick et al. 2009). Barnthouse et al. (1989) argue that a 10% change in variables would be societally and ecologically significant, although they were concerned primarily with laboratory toxicity tests and not field surveys. Their proposed CES was deliberately conservative (small) because of concerns about the uncertainty in extrapolating laboratory results to the field (Environment Canada 1998). The EEM program currently uses CESs of 25% for all fish population endpoints except for condition which has been set at 10%. These values were chosen primarily on data distributions of observed effects from within the program.

When preliminary analyses show that power will be insufficient given reasonable sample sizes, the assessments should be redesigned. Studies are designed site-specifically and the priority should be given to reducing variability rather than increasing sample size. CESs in the EEM program are defined as percentages of the reference mean and are not represented in the measurement units of the response variable, as these CESs would vary for different studies; therefore the coefficient of variation (COV), expressed as a percentage of the reference mean, is used as a measure of variability in sample size calculations (COV = standard deviation / reference mean × 100%). For a basic control/impact ANOVA (analysis of variance) design with untransformed data (e.g., as used for the age endpoint in EEM studies), the estimated sample size required to detect a given ES (effect size) at a given power level can be calculated as follows (Green 1989):

![Boxplot of distributions of observed power in EEM pulp and paper fish surveys by endpoint.](https://iwaponline.com/wqrj/article-pdf/45/2/209/229720/wqrjc0450209.pdf)
\[ n = 2(t_a + t_p) \left( \frac{\text{COV}}{\text{ES}} \right)^2 \]  

(1)

where \( n \) is the sample size required at each site, \( t_a \) and \( t_p \) are the values of Student’s \( t \) statistic (two-tailed for \( t_a \) and one-tailed for \( t_p \)) with \( n - 1 \) degrees of freedom at a significance level of \( \alpha \) and \( \beta \) respectively. COV and ES are each expressed as a percentage. The sample size equation is solved iteratively by choosing an approximate value of \( n \) to start with (usually 20).

For a basic control/impact ANCOVA (analysis of covariance) design using log transformed data (e.g., as used for the relative gonad weight endpoint), the estimated sample size required to detect a given ES at a given power level can also be calculated by using a different version of equation 1. This equation is as follows (Green 1989):

\[ n = 2(t_a + t_p) \left( S_f / \Delta_x \right)^2 \]  

(2)

where \( S_f \) is the standard deviation of the residuals using log transformed data and \( \Delta_x = \log(f + 1) \), where \( f = \text{ES} \) represented as a fraction of the reference mean (e.g., ES of 25% \( \Rightarrow f = 0.25 \)).

Unusual Observations, Statistical Assumptions, and Data Analysis

Barrett et al. (2010) discuss unusual observations in the EEM fish survey data. These observations can occur for a number of different reasons and should only be removed under certain circumstances. Outliers can occur when a test subject that does not belong to the population of interest is included in the study; outliers should be removed if they can be correctly identified. Obvious data entry errors should be corrected if possible and naturally occurring large or small observations should not be removed. Comments associated with unusual observations should be recorded in the field upon sampling to help determine whether data should be removed from an analysis. It is recommended to perform two separate analyses: one containing all data, and another with any outliers removed to determine the consequences of removing the data and to allow the reader to decide whether the observations should be removed or not.

The inclusion of immature fish data in statistical analyses can provide misleading results. Immature fish devote proportionally more energy towards growth, and have varying body size and gonadal growth relationships. For data analysis, fish identified as immature should be removed. The gonadosomatic index (GSI = gonad weight / body weight \times 100\%) can be useful in identifying immature fish. For many fish species, immature fish can be identified as having a GSI of <1\%. A plot of gonad weight versus body weight can also be useful in identifying immature fish. Comments from field observations may also assist in identifying unusual data that are suspected to be from immature fish. If immature fish are being used in any data analyses, it should be in comparing immature fish to immature fish between sites.

The standard statistical assumptions required for many parametric statistical tests are those of independence, normality, and homogeneity of variances. The two parametric analyses used to analyze the EEM fish population endpoints are ANOVA and ANCOVA. The assumptions for these tests are discussed below.

Independence (Pseudoreplication).

When designing experiments, it is desirable to ensure that replicates are randomly allocated to different treatment levels such that the response of each replicate is independent of other replicates. This element of randomness provides some assurance that observed differences in responses among treatments result from treatment effects and not for some other reason. Randomly allocating replicates to different treatment levels is a relatively easy procedure when conducting manipulative experiments (e.g., controlled laboratory tests), but is less obvious for observational field studies. Observational studies, such as environmental impact studies (e.g., single stressor EEM studies) or environmental assessments (i.e., multiple stressors), test hypotheses about the presence and magnitude of effects. However, the strength of inferences from these types of experiments is limited for two reasons (Paine 1990): the stressor (e.g., mill outfall, hydroelectric dam) cannot be replicated, and stressors cannot be applied randomly to replicates. What this means is that the stressor or treatment is always partly or wholly confounded with space or time and that the observed effects may or may not be caused by the stressor of interest. When significant differences are observed between reference and exposed fish populations, it is only possible to conclude there are differences between these two populations, not that the differences were caused by effluent exposure. Interpreting significant differences as treatment effects when either treatment is not replicated or replicates are not independent is referred to as pseudoreplication (Hurlbert 1984), as discussed previously.

It is critical that observations be confirmed, through replication over time, and that some effort be expended to confirm that the stressors of interest are involved in the responses before attributing cause to any specific stressor.

Normality and Homogeneity of Variances

The assumptions of normality and homogeneity of variance should be assessed before applying parametric procedures. However, most univariate normal distribution-based statistical methods are quite robust and can support moderate violations of the assumptions. Transformation of original data will help normalize the data or homogenize the variances. Logarithmic transformations are often preferred because most biological measures are considered to operate on a log or exponential scale (Peters 1983), and such a transformation
is biologically meaningful. If the transformations are unable to produce data that meet the assumptions, then a plot of the residuals may reveal problematic data points that may warrant investigation. Most of the univariate statistical methods are robust under moderate violations of assumptions, with some exceptions, including analyses with small and unequal samples. For serious violations, nonparametric statistics can be considered.

Additional Assumptions

An assumption of ANCOVA is that the independent variable (covariate) is fixed and measured without error. This assumption is frequently violated and Draper and Smith (1981) discuss the consequences of this violation. This is likely to prove problematic only in situations where the range of the independent variable is very small. The assumption of a linear relationship can be tested for samples with multiple observations at different values of the independent variable. This may be possible for discrete variables such as age, but not for continuous independent variables such as body weight (Environment Canada 1998).

An additional assumption of ANCOVA is that the regression slopes of each treatment (site) are parallel. When this assumption is not met, the ANCOVA procedure cannot proceed since there is a covariate by treatment interaction, and differences in the response variable among treatments vary at different values of the covariate. There are a few options for dealing with nonparallel regression slopes in ANCOVA that are useful in analyzing fish population survey data. Barrett et al. (2010) provide two different methods for dealing with data sets with nonparallel regression slopes so that they can be analyzed using the parallel slope ANCOVA model. The first method identifies data sets where the slopes are forced nonparallel by high-influence observations, which can be removed to provide parallel slopes. The second method identifies data sets where a model with nonparallel slopes is statistically, but not practically, significant and ANCOVA can proceed using the parallel slope regression model. Lowell and Kilgour (2008) provide another method for analyzing data when regression slopes are not parallel. They estimate the effects for smaller (or younger) and bigger (or older) fish by calculating the difference in the response variable at the values of covariate where the ranges for each site overlap. These estimates can then be compared to CESs.

It has been suggested that the range of the covariate in ANCOVA should be approximately the same for each site. This will be difficult to assure in practice, but the violation of this should be considered when interpreting results from such cases. If there is reason to believe that there are issues with the overlap of the range of covariate values, then a single factor ANOVA can be performed on the covariate values between sites. If the covariate means do not significantly differ between sites, then the results of the ANCOVA will probably be reliable (Quinn and Keough 2002). A significant difference in the mean covariate values between sites is on its own a significant effect. Things to consider would be the consistency of sampling gear between sampling sites and selection of samples in interpreting differences in the covariate means or ranges observed. It may be appropriate to provide an analysis on a subset of the data, omitting unusually high or low covariate values to provide a reliable analysis. Also, for several small-bodied fish species the range of the covariate (age) might only be 2 and 3 or 2, 3, and 4. An ANCOVA with only two or three values of the covariate can provide misleading results. In these cases it is appropriate to perform a one-factor ANOVA on body weight using site as the factor for each age group.

ANOVA and ANCOVA are robust to violations of the assumptions of the test when sample sizes are equal (Huitema 1980; Hamilton 1977) but quite often unequal sample sizes are prone. When assumptions are seriously violated and sample sizes are unequal, nonparametric alternatives to these tests could be considered. The Kruskal-Wallis test on the equality of medians is an appropriate alternative to ANOVA, and several different nonparametric techniques have been proposed for ANCOVA using ranks. Conover and Iman (1982) proposed a nonparametric alternative in which the response and the covariate are replaced by their ranks. The analysis is the same as the parametric ANCOVA using the ranks as data and is the simplest nonparametric alternative. Some other nonparametric alternatives to ANCOVA are discussed in Shirley (1981) and Quade (1967).

Conclusions

Monitoring fish at the population level gives a balance between the ecological relevance of changes at the community level and the sensitivity of responses at the biochemical level. Changes at the population level are important to monitor for the protection of fish species. Since the implementation of the pulp and paper EEM program in 1992, there have been many advancements along with new research available to improve fish population health assessments. The guidance provided in this paper can be used to improve study designs and the interpretation of fish surveys.

It is important that monitoring of fish populations is conducted over time so that observations can be confirmed and follow-up studies can be designed to address issues of concern. Life history characteristics of sentinel species should be studied and species should be selected to best answer the questions of the monitoring program. Monitoring should be conducted at an appropriate time to assess the questions of the monitoring program. Usually 15 to 25 fish of each species and sex are sufficient for preliminary assessments, and power analyses can be used to determine sample sizes for subsequent monitoring. Nonlethal sampling methods have been developed and should be considered if the methods can be used to answer the questions of concern. Careful data quality assurance / quality control and data analysis are necessary to obtain reliable results and design subsequent monitoring.
References


Barrett TJ, Munkittrick KR. 2010. Seasonal reproductive patterns and recommended sampling times for sentinel fish species used in environmental effects monitoring programs in Canada. Environ. Rev. 18:115–135.


Received: 1 September 2009; accepted: 30 December 2010.