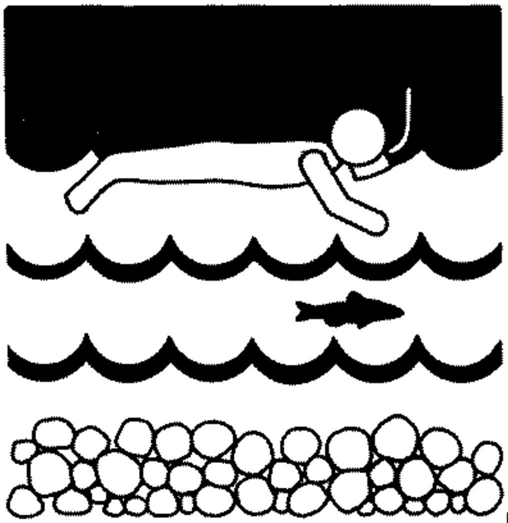
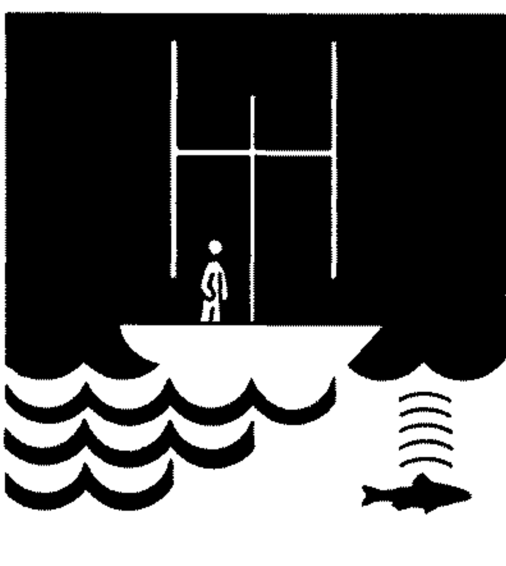
c.2



DEVELOPMENT AND EVALUATION OF HABITAT SUITABILITY CRITERIA FOR USE IN THE INSTREAM FLOW INCREMENTAL METHODOLOGY

INSTREAM
FLOW
INFORMATION
PAPER: NO. 21

Biological Report 86(7)
September 1986



INSTREAM FLOW
AND AQUATIC
SYSTEMS GROUP

Biological Report 86(7)
September 1986

DEVELOPMENT AND EVALUATION OF HABITAT SUITABILITY CRITERIA FOR USE IN THE INSTREAM FLOW INCREMENTAL METHODOLOGY

Instream Flow Information Paper No. 21

by

Ken D. Bovee
Instream Flow and Aquatic Systems Group
National Ecology Center
U.S. Fish and Wildlife Service
2627 Redwing Road
Fort Collins, Colorado 80526-2899

National Ecology Center
Division of Wildlife and Contaminant Research
Fish and Wildlife Service
U.S. Department of the Interior
Washington, DC 20240

DISCLAIMER

This report should be cited as:

Bovee, K. D. 1986. Development and evaluation of habitat suitability criteria for use in the Instream Flow Incremental Methodology. Instream Flow Information Paper 21. U.S. Fish Wildl. Serv. Biol. Rep. 86(7). 235 pp.

PREFACE

Past users of the Instream Flow Incremental Methodology (IFIM) have occasionally complained about the lack of biology in routine applications of the method. Using biological information in the IFIM is a relatively simple process. In fact, this simplicity may have trivialized the importance of the biological component. Experienced users realize that the important decisions related to biological data are made outside the mechanical operation of the models, and that the outcome of the analysis hinges on assumptions and decisions made long before the models are run. Many of these determinations are made by the user when planning an IFIM study, but many others were made by the researcher who developed the original biological information. All of these judgements may have some bearing on the outcome of an IFIM analysis.

This information paper concentrates on only one of the biological components of the IFIM, the development and evaluation of microhabitat criteria. Other equally important biological components include information related to water quality, temperature, trophic relationships, and population dynamics. These subjects are beyond the scope of this paper, which is fortunate because the topic of microhabitat criteria is large enough by itself. These other items will be discussed only to the extent that they may affect microhabitat usage.

This book has two intended audiences, but essentially only one goal. The first audience is the research community, the people involved in collecting data for the expressed purpose of developing habitat suitability criteria. The operational community, which uses the IFIM to analyze specific instream flow problems on specific streams, is the second audience. The ultimate goal of criteria research is to provide the user group with high quality, bias-free biological information that can be used confidently in any application of the IFIM. The achievement of this goal requires certain obligations from both researchers and users. The research community is responsible for developing criteria that are precise, accurate, unbiased, and sufficiently detailed as to be biologically meaningful. Users are charged with the responsibility of ensuring that the criteria used in an instream flow analysis are appropriate to the application. This requires a recognition by both groups, of the factors that can lead to errors and biases in the determination of habitat suitability.

Developing microhabitat suitability criteria is not like determining a species' tolerance to different concentrations of dissolved chemicals. The most obvious distinction is that nature is the laboratory in which microhabitat criteria are developed. Where there are very few options in conducting a bioassay study, there are literally dozens of pathways that could be followed in a microhabitat study. No single pathway is inherently better than any alternative route. The best approach is often determined by the environment (both natural and political) in which the criteria are developed. If the investigator is not careful, he or she may become part of the experiment, regardless of the scope of the exercise. The potential for these problems exists whether the criteria are developed from professional judgement or from field data.

SUMMARY

The Instream Flow Incremental Methodology (IFIM) is a habitat-based tool used to evaluate the environmental consequences of various water and land use practices. As such, knowledge about the conditions that provide favorable habitat for a species, and those that do not, is necessary for successful implementation of the methodology. In the context of the IFIM, this knowledge is defined as habitat suitability criteria: characteristic behavioral traits of a species that are established as standards for comparison in the decision-making process.

Habitat suitability criteria may be expressed in a variety of types and formats. The type, or category, refers to the procedure used to develop the criteria. Category I criteria are based on professional judgement, with little or no empirical data. Category II criteria have as their source, microhabitat data collected at locations where target organisms are observed or collected. These are called "utilization" functions because they are based on observed locations that were used by the target organism. These functions tend to be biased by the environmental conditions that were available to the fish or invertebrates at the time they were observed. Correction of the utilization function for environmental availability creates category III, or "preference" criteria, which tend to be much less site specific than category II criteria.

There are also several ways to express habitat suitability in graphical form. The binary format establishes a suitable range for each variable as it pertains to a life stage of interest, and is represented graphically as a step function. The quality rating for a variable is 1.0 if it falls within the range of the criteria, and 0.0 if it falls outside the range. The univariate curve format establishes both the usable range and the optimum range for each variable, with conditions of intermediate usability expressed along the portion between the tails and the peak of the curve. Multivariate probability density functions, which can be used to compute suitability for several variables simultaneously, are conveyed as three dimensional figures with suitability on the z-axis, and two independent variables on the x-y plane. These functions are useful for incorporating interactive terms between two or more variables. Such interactions can also be demonstrated using conditional criteria, which are stratified by cover type or substrate size. Conditional criteria may be of any category or format, but are distinguishable by two or more sets of functional relationships for each life stage.

One of the most important aspects of developing criteria is the formulation of a study plan that addresses the goals of the study and the intended use of the results. The study plan should anticipate sampling strategies and methods, and potential sources of error or bias so that the results will meet the perceived needs of the study. Regardless of the goal, the study plan should include:

(1) a statement of purpose and objectives,

(2) a list of target species and their selection criteria,

(3) a description of data stratification procedures, and

(4) a list of variables to be measured or described and how they will be expressed.

The above articles are required for all study plans. In addition, studies designed to develop empirical criteria must also include:

(1) stream locations where the data will be collected,

(2) identification of sampling strategies and methods,

(3) an estimate of sample size requirements, and

(4) a list of necessary equipment and supplies.

The statement of purpose and objectives establishes the orientation of the study. Studies designed to produce criteria for restricted use will have very different objectives from those intended for wide transferability.

The selection of target species is often influenced by the intended audience for the criteria. Some studies will concentrate on only one or two species of particular importance to a specific instream flow determination. Others may include many species, or guilds of species, to expand the biological data base as much as possible. The decision to study many species or only a few is important. It is generally more efficient to collect data on several species at the same time, but if some restrictions are not applied, the effort can be diluted among species of lesser interest. Investigators should consider the information content of each criteria set when selecting target species. This is determined by the management importance of the species, its adaptations to riverine environments, and the amount of information already available for a particular area or stream type.

Data stratification refers to the subdivision of criteria for a species to reflect spatial or temporal changes in microhabitat utilization patterns. Common divisions include size classes or age groups, diurnal or seasonal changes in habitat usage, different activity patterns, and variations in tolerable hydraulic conditions as a function of cover or substrate type. Understratification of data can be a serious problem, either resulting in overly-broadened criteria or bimodal frequency distributions.

The sampling protocol is a formalized description of the variables to be measured or described, and procedures for measuring, describing, and recording the data. The purposes for establishing a sampling protocol are to enhance consistency and reduce ambiguity. Many investigators use coding systems or abbreviations to record the species, size class, activity, substrate, and cover. An important aspect of the sampling protocol is to cross reference these codes to a written definition for each variable. The sampling protocol also defines how certain variables are to be measured, such as measuring the mean column velocity or the nose velocity at each location. Units of measurement should also be defined under this component of the study plan.

One of the most important elements for the design of category II and III criteria is the selection of appropriate study areas. Habitat availability can be a major source of error in the development of these criteria and is particularly serious in category II criteria. The ideal study site would contain all conceivable combinations of microhabitat conditions in equal abundance. Fish observed in such a stream would reflect the true preference and avoidance behavior of the species, because the fish would have free and equal access to all microhabitat conditions. Although this ideal situation is virtually impossible to find in nature, the closer the study stream approximates this condition, the smaller the bias in the resulting criteria. Other important considerations in the selection of the source stream are factors that may alter a species' selection of microhabitats, such as water quality, temperature, and the presence or absence of competitors or predators.

A coherent sampling strategy is necessary to avoid biases due to disproportionate sampling effort. Investigators who emphasize the quantity of observations rather than the quality, tend to sample more intensively where they expect to find fish (or macroinvertebrates). Consequently, the resulting criteria become self-fulfilling prophecies. This is an especially serious problem, because it is almost impossible to detect this type of bias. Selection of a particular sampling strategy is contingent on the intended sampling method, because certain strategies are compatible only with particular types of gear or data collection techniques.

Obtaining an adequate sample size is not only necessary to preserve accuracy in the criteria, but also to facilitate fitting a function to the observed frequency distribution. Typically, 150 to 200 observations are necessary to construct a reasonably smooth histogram. An observation refers to a single location where microhabitat utilization is observed, regardless of the number of fish found at the location. The actual sample requirement may need to be adjusted up or down, depending on the variance of the samples. Sample size estimates of less than 150, however, may be symptomatic of restricted microhabitat availability in the source stream, suggesting that the study should be moved to another area.

Habitat suitability criteria are not always developed from field studies. There are numerous situations that can dictate the formulation of category I criteria, which are largely based on literature sources and professional judgement. Of the literature sources, reports of previously conducted criteria development studies are much more useful than the more common life history or distribution and abundance studies. The habitat descriptions of the latter are usually not quantitative enough for the formulation of criteria. Unfortunately, many of the better criteria studies are found in relatively obscure journals and reports or are unpublished.

Development of category I criteria by professional judgement is a common solution when data for higher categories are unavailable. Three techniques have evoled to this end: roundtable discussions, the Delphi technique, and habitat recognition. The roundtable is an informal, face-to-face discussion among group participants. The success or failure of such group interactions depends on the composition of the group and the leadership abilities of the moderator. The advantages of the roundtable approach are that all participants have equal access to information exchanged by the group, and feedback is instantaneous. The disadvantages of this approach include scheduling problems, repetitive meetings, a tendency to discount minority opinions, and potential domination of the group by strong personalities.

The Delphi technique was devised to overcome many of the disadvantages of face-to-face discussions. The most common Delphi exercise uses a questionnaire, developed by a small monitor team and sent to a larger respondent group. The use of the questionnaire surmounts two of the major problems of the roundtable approach. Respondents can participate at their convenience, so specific times do not need to be scheduled for meetings. The anonymous nature of the questionnaire also prevents the "bandwagon" effect of a group dominated by a strong personality. Whereas feedback is instantaneous in roundtable discussions, it is delayed in a Delphi exercise. This places a greater responsibility on the monitor team to be absolutely clear in the definitions of terms, and in communications in general. It may also be more difficult to prevent the introduction of tangential subjects, although this problem occurs with roundtable discussions as well.

Habitat recognition is founded on the premise that although the most qualified experts may not be able to quantify usable and unusable habitat, they can recognize it when they see it. This approach involves field data collection, but relies on the opinions of the experts rather than sampling of fish. Each participant is provided with a secret ballot and, at specific locations in the river, indicates whether or not the specified target organism would be likely to use that location. Microhabitat measurements are then made at the location. A frequency distribution of all the responses is then assembled. Each "yes" vote is assigned a frequency of one and each "no" vote is assigned a frequency of zero. Functional relationships are then fit to the frequency distributions using the same techniques that would be used for empirical data.

Many research biologists are critical of category I criteria because of their lack of an empirical data base. When time or resources precludes the collection of empirical data, however, category I criteria are much better than no criteria at all. Furthermore, verification studies comparing category I criteria with subsequently developed category II criteria have shown good agreement between the two.

Criteria in categories II and III are based on data collected specifically for determining the habitat characteristics of a species. Two types of data bases may be required. The first is a utilization data base, data collected at locations where the target organism was observed or captured. Utilization data are required for both categories of criteria. Additionally, an availability data base, a frequency distribution of the environmental conditions in the stream at the time of sampling, is needed for the development of category III criteria.

The development of a valid utilization function requires the unbiased measurement of microhabitat variables at specific fish (or macroinvertebrate) locations. This means that the focal point or centroid of the home range of the organism must be determined as accurately as possible, further implying that the observer and the collection technique must not interfere with the observation. These conditions place limitations on sampling techniques and gear that can be used successfully in criteria studies. Methods designed to capture fish in transit, or which attract fish to the capture device, cannot be used in this type of study. This eliminates most passive capture techniques, such as stationary gill nets or baited traps. The technique must also be capable of sampling specific locations in the stream. Chemicals cannot be used in criteria studies, because they are nonspecific.

There are still many techniques that are applicable to criteria studies, although procedural or equipment modifications are often required to overcome gear biases and limitations. The favored approaches are those involving direct observation of the fish: surface observation, snorkeling, and SCUBA. The advantages of these methods are that they allow precise definition of the activity and focal point of the subject with a minimum of disturbance. Factors influencing visibility, such as turbidity, surface glare, and turbulence are the primary limitations to observational techniques, although the effect of these factors varies according to the method used. Divers are limited in the kinds of conditions under which they can work in safety and comfort. There will be some situations where the water is too fast or too cold, covered by surface ice, or is otherwise too dangerous for a diver.

Remote observation techniques, such as underwater video or radiotelemetry, are useful in situations where direct observations cannot be made. Limited visibility causes the same kinds of problems for underwater video as it does for divers, but radiotelemetry can be used in any water. Both systems allow the investigator to monitor the movements or activities of the fish, and can be used in places where unsafe conditions prevent diving. It is much more difficult, however, to identify the exact position of the fish with these techniques. A gridwork of underwater monuments must be used with video systems to identify focal points, and special procedures must be used with radiotelemetry systems to reduce triangulation errors.

Several active capture techniques can be used to sample specific locations in streams to determine their usage by fish and macroinvertebrates. Electrofishing, explosives, and small area samplers can all be used in criteria studies, but can be very disruptive and potentially biasing unless certain modifications are made. The prepositioned area shocker is a relatively new innovation that may overcome many of the problems or traditional electrofishing procedures. This technique uses a large anode that is placed at a predetermined location in the stream and left in place for a period of time to allow fish to resume their normal behavior patterns. After this waiting period, the electrode is energized, and fish immobilized within the area bounded by the electrode are captured, identified, and counted. A similar procedure involves the use of preset explosive charges, generally primacord. Explosives are not as size selective as electrofishing, but the fish are killed rather than being immobilized. They also tend to sink to the bottom, which makes their recovery more difficult. Certain procedures can be used to enhance the capture efficiency of both techniques to overcome some of their limitations.

Small area samplers are most effective for capturing relatively immobile organisms, such as macroinivertebrates and very small fish. These include a variety of bottom samplers, nets, seines, and traps. In nearly every case, special procedures must be used to obtain localized, homogeneous samples with area samplers. These devices are usually designed to collect a particular type of organism, which can be an advantage when other techniques are ineffective. They are also limited in the types of environments in which they can be used and in the sizes or types of animals they can capture. Therefore, area samplers are usually considered to be specialty devices to supplement or replace other techniques, when these become ineffective.

All procedures for observing or collecting fish and macroinvertebrates have some inherent limitations. If the conditions of the sampling environment do not exceed the limitations of the technique, however, the data will probably not be biased by differential efficiency. The important aspect of developing the utilization data base is to understand the limitations of each approach, determine the characteristics of the sampling environment and the subject organism(s), and then select the most effective technique for that particular application. This undoubtedly means that a wide variety of methods must be used to develop seasonally stratified criteria for all life history phases of a species.

Determining the distribution of microhabitat availability is much easier than determining utilization. Despite this difference in effort, habitat availability data is as important to a criteria study as fish observations are. Without such data, the investigator risks a severe enviromental bias in the criteria. There are two basic approaches for determining habitat availability: random and proportional sampling. The method used to determine availability is somewhat determined by the method used to determine utilization. Random sampling is most compatible with methods that involve fish collections, rather than observations, to develop the utilization function. This approach uses randomly selected locations for fish sampling and microhabitat measurements are made at each location, regardless of whether fish were captured or not. Proportional sampling is most compatible with direct and remote observation techniques. Using this approach, the study area is mapped and appropriate measurements made to conduct a PHABSIM simulation. The relative frequency distribution of various microhabitat combinations is then determined on the basis of surface areas.

Combining data from different sources, whether they be different streams, different sites, or the same sites at different flows can result in data pooling bias. When habitat utilization data are taken from more than one site, it is important to recognize that the frequency of observations or collections is a function of several factors besides habitat suitability. The surface area sampled, the time spent sampling, and the efficiency of the technique used also influence the frequencies recorded in the data base. The easiest way to eliminate data pooling error from category II data is to use study sites of equal area, sample each of them the same number of times, and use the same observation or collection technique each time. The aforementioned random and proportional sampling designs will internally correct for unequal effort. However, certain other sampling approaches, such as basing the number of random samples of the environment on the catch rate or using certain types of systematic sampling, can lead to data pooling bias.

Once the data have been collected, they must be reduced to an easily interpretable graphical display. This involves fitting univariate or multivariate curves or functions to the data. Three basic approaches have evolved for the processing of habitat utilization and preference data: histogram analysis, nonparametric tolerance limits, and function fitting.

Histogram analysis is conceptually simple but, because of the irregular nature of utilization and availability histograms, may actually be more difficult to use than the other techniques. The basic approach is to fit a curve, by eye, to the frequency distribution. This is often fairly imprecise, because different investigators will draw different curves. One way to improve precision is by smoothing the histogram through the grouping of intervals, but this may result in a decrease in accuracy. Another technique is to compute the residual sum of squares for each curve and use the curve that minimizes this statistic. This is a haphazard and tedious approach, however, and is rarely employed.

Nonparametric tolerance limits are used to determine a range of an independent variable within which a certain percentage of the population will be found. Suitability for a given interval is computed as:

$$SI = 2(1-P)$$

where P is the proportion of the population under the curve. Thus, the central 50% is assigned a suitability of 1.0, whereas the range including the central 90% has a suitability of 0.2. This approach has many desirable attributes. It is easy to use, it can be used with small sample sizes, it is insensitive to irregularities of the frequency distribution, and it does not require the presumption of any particular distribution or curve shape. Because the resultant suitability curve represents cumulative frequencies, however, the relative frequency distribution must be estimated in order to calculate the preference function.

Nonlinear regression techniques involve many of the same concepts as histogram analysis, except that a mathematical equation is used to draw the curve. Once an appropriate function has been chosen, a series of trials is made to determine the equation coefficients that will minimize the residual sum of squares. Many nonlinear regression programs contain solution algorithms that solve for the roots of an equation. Others rely on a "brute force" approach, where the user is required to change the coefficients manually, in search of this solution. These programs require much less memory than those with automatic search routines and can be used on smaller computers, but they are not very efficient. Nonlinear regression techniques can be used to fit either univariate curves or multivariate probability density functions. Exponential polynomial equations are commonly used for multivariate analysis, and the logistic regression approach has recently been suggested as an alternative.

The primary advantage of using a multivariate function is that it can incorporate interactive terms between independent variables in the calculation of habitat suitability. The use of univariate curves assumes that the selection of certain environmental conditions is not significantly affected by variable interactions. The importance of this assumption has been a serious source of confusion and misunderstanding because some interactions have biological importance, and some do not. The error of attributing biological meaning to variable interactions when they are spurious is as serious as assuming independence when they are not. The most common types of biologically important interactions are related to hydraulics and cover types. Fish may use shallow water in the presence of overhead cover and deep water in its absence, but will not use shallow water without cover, for example. This type of interactive behavior is best described by developing conditional criteria. Interactions between depth and velocity have been assumed to be biologically important, but are usually artifacts of the sampling environment that are eliminated when the utilization function is corrected for availability. Criteria developers should test their data for interactive terms and determine whether such interactions are biologically induced or merely artifacts of the environment. Univariate curves are much more flexible and are easier to use in PHABSIM than are multivariate functions. In many cases, they are more accurate than multivariate functions, which can represent only a few types of distributions. If it is determined that the interactive terms have biological significance, however, the user may be required to use the multivariate format.

The criteria used in an IFIM application will often originate from streams other than those being evaluated with IFIM, because of the time and expense of developing an empirical data base. Furthermore, the stream under investigation may not meet the criteria of a good source stream for criteria development. Before offsite criteria are used in an operational IFIM study, they must be evaluated to determine if they are adequate for the needs of the study. This evaluation consists of two parts: a review of comprehensiveness and a determination of accuracy.

The review of comprehensiveness is concerned with the data stratification procedures and sampling protocol followed in the criteria study. The purpose of this evaluation step is to determine whether the level of detail exhibited by the criteria is compatible with the perceived needs of the IFIM study.

This process will reveal information gaps, such as missing criteria for a particular life stage, activity, or season. The review is also useful in determining the adequacy of the criteria for certain variables with respect to the river in which they will be applied. In particular, it is important to determine whether nose velocities or mean column velocities were measured, and whether the velocity criteria are appropriate to the study stream. The level of detail in substrate descriptions and the stratification of criteria by cover type are also important determinants of the adequacy of the criteria. Often, it will be found that the existing criteria are satisfactory, but that certain critical information is missing. Additional criteria may need to be acquired, or existing information supplemented.

Evaluations of accuracy and precision can take two mutually exclusive pathways. The easiest, but least definitive, is a screening level review of the study plan and implementation. The other approach is to implement one of several field verification studies. These are more costly in terms of time and money, but the results can provide a solid basis for acceptance or rejection of the criteria.

Factors to be considered in a screening level evaluation include the diversity of the source stream, potential biases associated with the sampling design, and errors associated with data collection. A general rule is that criteria may be transferred from highly diverse streams to those with lower diversity, but not the opposite direction. An easy method of determining the relative diversity of the source stream is to compare the utilization function with the preference function for the same data stratum. If the two are very similar, they were likely derived in a highly diverse environment. If they are radically different from one another, they probably originated from a very simple or restricted environment. In this case, it is likely that neither function is very accurate, and neither should be used.

The investigator should also evaluate any potential biases inherent in the sampling design used in the criteria study. Some sampling designs may be theoretically better than others, especially when data are pooled from several sources. In the context of a criteria review, however, the description of a sampling design at least indicates that the original researcher recognized its importance. Whether the best strategy was used is often less important than knowing that the field crew did not confine their sampling to places where they expected to find fish.

Three types of error are often associated with data collection: precision error, disturbance, and gear bias. Precision error refers to the ability of the observer to determine the focal point, or home range centroid. Precision errors are generally lowest when direct observation techniques are used, although prepositioned electrodes and preset explosives are also capable of high-precision samples. Area samplers, unless they are very small, generally exhibit the largest amount of precision error. Underwater video and radiotelemetry are intermediate, with the amount of error affected and controlled by the skill of the observer.

Disturbance is a serious precision error, because the fish are displaced from their original focal points, and then observed. This source of error can often be evaluated simply by determining the observation technique used. Snorkeling, surface observation, radiotelemetry, and underwater video are all considered to be low-disturbance techniques. SCUBA can be moderately disruptive, depending on the reaction of the fish to the diver and bubbles vented from the regulator. The amount of disturbance can be controlled to some extent by the diver, however, or at least documented so that it can be easily interpreted. Active capture techniques, especially electrofishing and explosives, can be extremely disruptive unless specific procedures are followed to reduce the disturbance.

Gear bias is a generic term related to the effects of the environment on the effectiveness of the sampling or observation technique. All observation and capture methods have some inherent limitations. For example, it is easier to see or collect fish in shallow water than in deep water with some methods. Other methods may be affected by velocity or turbulence, turbidity, temperature, or other factors. If the limitations of the sampling method can be considered constant over the range of measured conditions, however, the technique may be presumed to be unbiased.

A verification study is an empirical test of the accuracy and reproducibility of offsite criteria. These studies all require the collection of certain types of data in the subject stream. The confidence that can be placed in the results of a verification study is directly related to the amount of effort invested in the study. In order of increasing sophistication, the three approaches to field verification are: the abbreviated convergence approach, habitat suitability overlay, and Monte Carlo simulation.

The abbreviated convergence approach is essentially a miniature criteria study, wherein a small number of fish is collected or observed, and histograms constructed for each data stratum and variable. These histograms are then superimposed on the corresponding suitability curve. If the peaks and tails of the curve and the histogram are approximately the same, the test is considered positive. The test is negative if they are noticeably different from each other. Negative results may not be very meaningful, however, because of the small sample size used to develop the test histogram.

The habitat suitability overlay method is based on the premise that areas of the stream computed to be high quality habitat should contain more fish than those estimated to be low quality. This approach compares the computed suitability of each stream cell with the observed fish distribution, and is a more definitive test of the criteria than the previous method. There is only one test result that confirms the accuracy of the criteria. All other results are considered negative, but the nature of the results can be used to diagnose the general problem (such as the curves being too narrow, too wide, or simply incorrect).

The most rigorous verification test is the Monte Carlo simulation. The field application is virtually identical to the overlay method, but this method uses a random process to predict the actual locations of the fish.

These predicted locations are then correlated with observed locations. The correlation coefficient can be used to evaluate how well the criteria agree with the behavior of the fish in the subject stream.

As a result of the review and evaluation phase, it may become apparent that some of the curves or functions should be modified before they are applied to the subject stream. The most common form of modification is extension beyond the limits of the existing criteria. This is essentially a matter of letting professional judgement take over where the data leave off. Actual modification involves changing the shape or the intercepts of the original functions. Legitimate reasons for modifying criteria include:

(1) addition of information not contained in the original data,

(2) resolution of differences between two or more models,

(3) incorporation of professional opinion in the final model, and

(4) formulation of a mixed model.

The purpose of these changes should be to improve the accuracy of microhabitat predictions in PHABSIM. It is not legitimate to change criteria simply to alter the results of PHABSIM. This constitutes deliberate manipulation of the model to justify a preconceived outcome, a practice that can undermine the credibility of the user and the model.

The most definitive test of habitat suitability criteria is mathematical convergence, where several investigators working in different areas derive the same functional relationships. This requires several replicate studies to be conducted on the same species, using the same data stratifications and sampling protocol in all the studies. Any deviations from one study to another invite divergence in the resulting criteria. It is unreasonable to expect reproducibility when the same procedures are not followed in any experiment. A goal of these studies should be to develop regionalized criteria that are applicable for a species in a specified geographical area. The applicable regions should be determined on the basis of convergence, however, and not assigned by arbitrary boundaries. Until such criteria are available, researchers must strive to develop comprehensive, accurate, and transferable criteria, and users must continue to evaluate and test it.

CONTENTS

CONTENTS (Concluded)

FIGURES

FIGURES (Continued)

FIGURES (Continued)

FIGURES (Continued)

FIGURES (Concluded)

TABLES

ACKNOWLEDGMENTS

This has been a very rewarding and interesting project partly because I find the subject so intriguing, and partly because I had so much to learn. Although I have had some practical experience in the development of habitat suitability criteria, it did not take long before a topic arose that I knew virtually nothing about. Fortunately, I was able to draw on a wealth of practical experience and advice from friends and colleagues who had struggled with many of the problems I was writing about. Dr. Jeff Gosse, whose several innovations are documented in this book, also provided valuable advice regarding safety precautions when using SCUBA observation techniques. Mr. Johnie Crance provided many useful comments on the Delphi technique, and also provided the introductory Delphi package in Appendix A. I am indebted to Dr. Harold Tyus, USFWS, Vernal, Utah, for his review of my discussion on radiotelemetry, and his suggestions for increasing the accuracy of this technique in criteria studies. Dr. Mark B. Bain, Argonne National Laboratory, was very helpful in his discussion of sampling strategies and techniques. Dr. Bain and his associates are the developers of the Prepositioned Area Shocker; and researchers who work in streams with zero visibility may be very appreciative of this innovation in the future. Dr. Stacy Li, Holton Associates, Berkeley, California, provided a comprehensive review of the entire paper. His suggestions relating to snorkeling will make this technique easier and safer for future divers. Mr. Allan Locke, Alberta Fish and Wildlife Division, asked exactly the right questions to make my discussion of data pooling problems more relevant and understandable. Mr. David Wegner, USBR, Flagstaff, Arizona, commented on the overall content of the paper. His experiences with techniques ranging from roundtable discussions to underwater video and explosives add substance to my explanations of data acquisition techniques. Ms. Catherine Sabaton, Department Environmental Aquatique et Atmospherique, Electricite de France, Chatou, France, was instrumental in clarifying the chapter on data analysis. Her insightful questions regarding the nature of the preference index were especially important as they disclosed many of the solutions to the data pooling dilemma. I am also indebted to Mr. Lynn Kaeding, USFWS, Grand Junction, Colorado, and Mr. Larry Moore, Midwest Divers Supply, Ft. Collins, Colorado, for their help in assembling the equipment list in Appendix B.

I am grateful to several authors and publishers who kindly gave their permission to reprint some of the tables and figures found in this book. These include:

Elsevier Scientific Publishing Company, Amsterdam, The Netherlands,

The Institute of Mathematical Statistics, Hayward, California, Ms. Jean Bladridge, Entrix, Anchorage, Alaska,

Dr. George Innis, Colorado State University,

Dr. Kenneth A. Voos, Woodward-Clyde Consultants, San Francisco, California.

Their contributions have added significantly to my explanations of concepts and ideas expressed herein.

In my writing, I have attempted to be thorough and accurate. Unfortunately, the quest for comprehensiveness is often an invitation to redundancy and verbosity. Two people are to be commended for their attempts to keep me to the subject (the reader may judge how successful they were). Ms. Marje Blaine, Alamogordo, New Mexico, provided technical and compositional comments that resulted in a major revision of the first draft. Mr. James Zuboy is the NEC editor, and the readability of this manuscript is largely due to his efforts. Ms. Susan Strawn translated my hand sketches into legible illustrations. Ms. Madeline Sieverine and Ms. Elizabeth Barstow handled the typing duties, and I thank them for their efficiency and accuracy.

Finally, I wish to especially thank Mr. Erwin G. Dieter, for teaching me the most important things I know about fish.

1. INTRODUCTION

The Instream Flow Incremental Methodology (IFIM) is a habitat-based tool for evaluating environmental changes in streams and rivers. A prerequisite of any habitat-based methodology is knowledge about those conditions that constitute "habitat" and those that do not. The fact that different species of fish and macroinvertebrates occupy different habitat types in streams is intuitive to anyone who has spent any time observing the animals in the wild. There is a difference, however, between this intuitive knowledge and the ability to quantify the microhabitat characteristics selected by the organism. The quantification of these characteristics is what distinguishes microhabitat suitability criteria from anecdotal descriptions.

Webster's New Collegiate Dictionary (1976) gives two definitions to the term "criteria": (1) a characterizing mark or trait, and (2) a standard on which a judgement or decision may be based. If criteria were restricted to the first definition, they would be descriptive only. Standards for making judgements, however, require quantification. Taken in this context, the microhabitat criteria used with the IFIM employ both definitions. They are, in essence, characteristic behavioral traits of a species that are established as standards for comparison in the decision-making process.

The subject of this information paper is the development and evaluation of physical microhabitat criteria used in the IFIM. Two separate, but related, audiences are addressed by this paper: people who develop criteria, and people who use criteria. Therefore, the paper has two goals. For those investigators developing criteria in a research mode, the purpose of the paper is to describe data acquisition and analysis procedures, with the ultimate goal of improving the quality, flexibility, and transferability of the criteria data base. The purpose with respect to the user community is to describe the variety of methods by which microhabitat criteria can be developed, and methods by which criteria can be evaluated, modified, and tested.

1.1 HABITAT CRITERIA CONCEPTS

1.1.1 Use of Habitat Criteria in the IFIM

The Instream Flow Incremental Methodology contains a series of mathematical, empirical, and conceptual models that are used to compute habitat availability under various water management alternatives. Figure 1 is a simplified flow diagram of the IFIM that illustrates how the various models are linked together. Biological information enters the IFIM at three places:

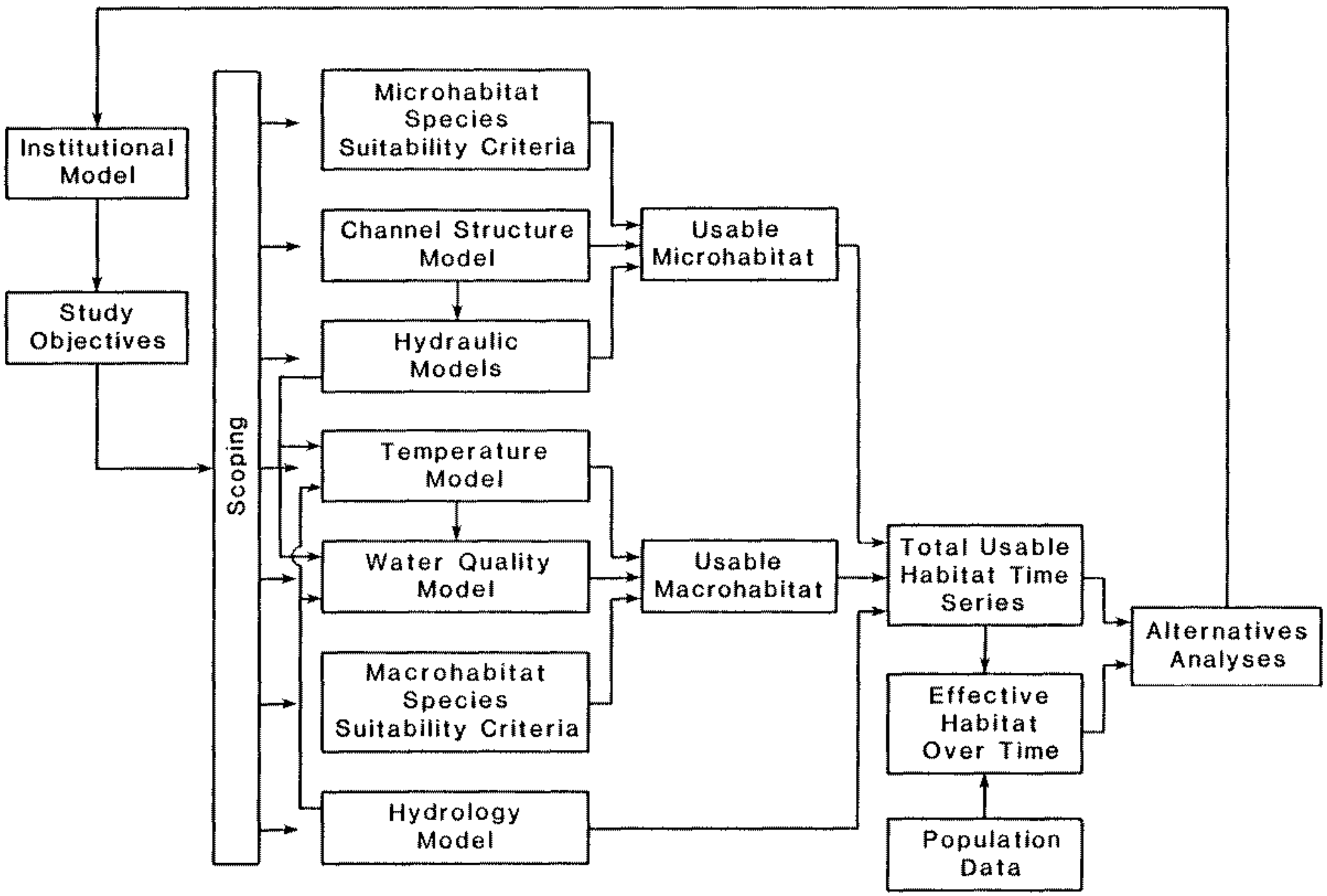


Figure 1. Schematic representation of model linkages in the IFIM showing input points for biological criteria.

as input to the Physical Habitat Simulation System (PHABSIM), in the integration of water quality and temperature to compute total habitat, and in the "population model" shown as the last step in Figure 1. This information paper will deal exclusively with the first type of biological information, but more advanced applications of the methodology will inevitably increase the importance of the latter two.

The Physical Habitat Simulation System (PHABSIM) has been described extensively in the literature (Stalnaker 1979, 1982; Bovee 1982; Milhous et al. 1984), so it will be discussed only briefly here. Measurements at a stream study site divide the stream into a large number of rectangular or trapezoidal cells, as shown in Figure 2. Each cell is considered to have a unique combination of depth, velocity, substrate, and cover at any particular discharge. Cells near the edge of the stream may also have a surface area that varies with discharge; however, those in the center of the channel generally have a fixed surface area. Substrate and cover characteristics are also fixed for each cell. Changes in depth and velocity at unmeasured discharges are predicted by the use of hydraulic simulation models described by Bovee and Milhous (1978) and Milhous et al. (1984). Depths and velocities can also be entered directly into PHABSIM if the stream cannot be accurately simulated by the hydraulic models. Thus, the physical models within PHABSIM describe how the environment changes with respect to streamflow.

The translation from flow to habitat occurs when each cell is evaluated against the microhabitat criteria for a species. The depth, velocity, and substrate or cover codes for each cell are compared with the criteria for each variable to determine the overall habitat quality of the cell. This quality is a composite based on the combination of hydraulic and structural (i.e., substrate and cover) characteristics of the cell. A composite quality value of 1.0 means that the cell is entirely satisfactory as microhabitat; 0.0 equals unsuitable habitat. The composite quality value is then multiplied by the surface area of the cell to obtain an index of microhabitat, called weighted usable area (WUA). This process is continued for all the cells. All the individual cell weighted WUA's are then summed to give one weighted usable area for the whole study site at one discharge. The sequence is then repeated for other discharges.

Several different types of criteria can be used in PHABSIM, and are distinguished according to format and category. The format refers to what they look like; the category refers to how they were developed. It is important for the PHABSIM user to understand this nomenclature, because criteria that look similar may have been developed by different techniques, and may imply different behavioral characteristics. Alternatively, criteria developed following the same basic procedures may be expressed in different functional forms. The format may have important implications in terms of the behavioral characteristics ascribed to the organism. By the time weighted usable area is calculated, it is virtually impossible to determine which kinds of criteria were used.

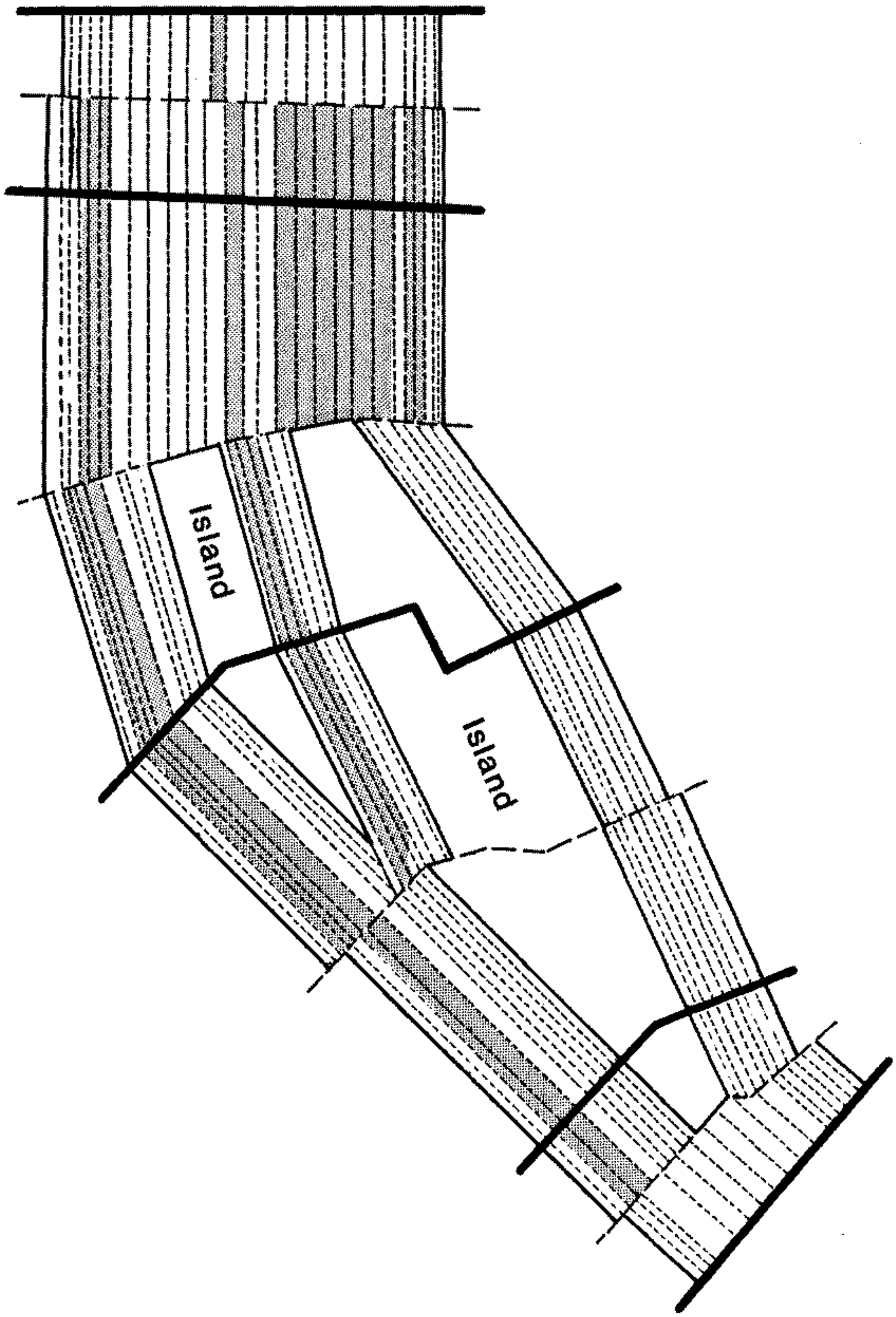


Figure 2. Conceptualization of a stream reach divided into cells for microhabitat analysis in PHABSIM.

1.1.2 Criteria Formats

The easiest and least theoretical approach of distinguishing among the different types of microhabitat criteria is by the formats in which they are expressed. Three formats can be used with PHABSIM: binary criteria, univariate curves, or multivariate response surfaces. The differences between these formats are illustrated in Figure 3.

The binary format (Figure 3a) establishes a suitable range of conditions for each variable as it pertains to a life stage of a species. The quality rating for a variable is 1.0 if it falls within the range established by the criteria. Any variable outside the criteria range is given a value of 0.0, which renders the cell unusable regardless of the quality assigned to the other variables. Therefore, a cell can be considered to be suitable habitat only if all the variables fall within their respective suitable ranges. The range considered to be usable is typically quite broad, often encompassing the conditions which 80% to 100% of the individuals are likely to inhabit.

Waters (1976) suggested the use of weighting factors between 0.0 and 1.0 to define habitat suitability for fish. He argued that, within the range of conditions considered suitable, there is a narrower range that fish select as preferred or optimal for that variable. This format expresses the behavioral characteristics of an animal as a series of univariate curves, rather than the block or step functions expressed by binary criteria. The univariate curve format is shown in Figure 3b. The peak of the curve represents the most suitable, most used, or most preferred range for each variable (the differences among these terms is discussed under categories of criteria, Section 1.1.3). The tails of the curve represent the bounds of suitability for each variable. Conditions of intermediate suitability are expressed along the portion between the tails and peak of each curve. The preferred technique of determining values between 0.0 and 1.0 is to fit a curve to a frequency distribution of empirically-derived data. Sometimes, only the optimal range and the locations of the tails are known, and intermediate values are estimated by straight line connections between 0.0 and 1.0 on the curve.

An example of a three-dimensional orthogonal response surface is shown in Figure 3c. The axis of the response surface appears twisted as the correlation increases between two variables. This interaction is most easily visualized by looking down on the response surface as though it were a topographic map. The primary advantage of the multivariate response surface is the ability to express interactions among the variables. There is no difference between an orthogonal response surface (e.g., Figure 3c) and two or more univariate curves multiplied together. Various considerations of using univariate versus multivariate functions are discussed in Chapters 5 and 6.

An alternate way to describe behavior-induced interactions is to group intervals of a continuous variable and treat them as categorical variables. A continuous variable is one that can theoretically assume any value between two given values; a discrete variable is one in which intermediate values between two given values do not exist (or are assumed not to exist). As more, or finer, discrete intervals are defined for a variable, the series more closely

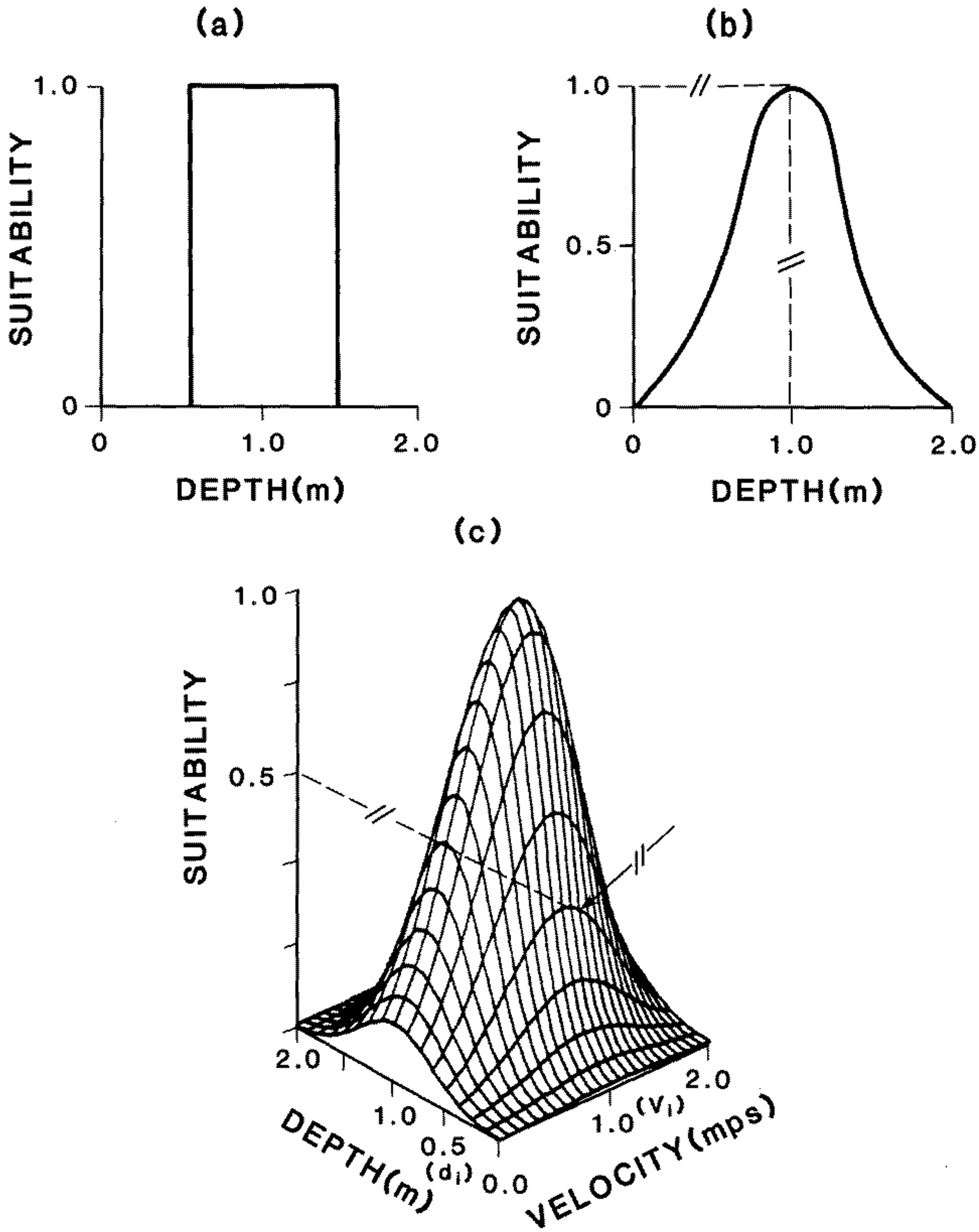


Figure 3. Examples of the three formats of habitat criteria that can be used in PHABSIM: (a) binary, (b) univariate curve, and (c) multivariate response surface.

approximates a continuous variable. Some variables, such as size, time of day, or season, are continuous but can be stratified into categories, whereas other variables, such as cover type, are truly discrete. The use of discrete variables is the basis for the development of conditional criteria.

Conditional criteria employ a separate set of criteria for each category of a discrete variable. A common example of conditional criteria is the development of separate criteria for fry, juveniles, and adults because it is typical for each of these sizes of fish to use different types of habitat. Conditional criteria are especially useful in describing behavioral interactions with respect to cover and substrate. Many species exhibit cover-conditional behavior, utilizing shallow water in the presence of overhead cover, fast water in the presence of large substrate, or deep water in the absence of overhead cover. Conditional criteria are in somewhat of a class by themselves. They may be expressed in any format: binary, curve, or response surface. The distinguishing format characteristic of this type of criteria is the appearance in sets of two or more.

1.1.3 Criteria Categories

The criteria format is largely a description of how the habitat suitability for one or more variables is expressed. Criteria categories refer to the kind of data used to generate the criteria and how those data have been processed. For simplicity, three categories have been identified.

Category I means that the habitat criteria are based on information other than field observations made specifically for the purpose of criteria development. Category I criteria are derived from life history studies in the literature or from professional experience and judgement. Category II criteria are based on frequency analysis of microhabitat conditions utilized by different life stages and species. These criteria are termed "utilization functions" because they depict the conditions that were being used when the observations were made. Utilization functions may not always accurately describe a species' preferences because the preferred conditions might be absent or in short supply. Category III criteria attempt to correct this bias by factoring out the influence of limited habitat availability. The purpose of this correction is to increase the transferability of the criteria to streams that differ from those where the criteria were originally developed, or in the same stream at different flows.

1.2 DISCUSSION

There are numerous meanings of the term "microhabitat suitability criteria." This term has a range of definitions from binary - category I criteria, to conditional - category III criteria. There are also several connotations associated with each of these types that can affect the adequacy of the criteria in any problem-solving setting. These connotations have created difficulties for past users of the IFIM. Many of these problems have been created by limitations of available criteria, but nearly as many have been self-inflicted by the inappropriate use of the information. The goals of this information paper are to encourage the development of more and better criteria and to eliminate, if possible, their inappropriate use.

The adequacy of the criteria is usually one of the first areas of criticism when an instream flow study is challenged. A variety of potential problem areas have emerged from these criticisms, some directed at the criteria, some at their use. Most of these problems can be classified into one of the following groups: credibility, quality control, transferability, comprehensiveness, or cost. It should come as no surprise to learn that these factors are highly interrelated.

Credibility may be the paramount issue in the development and use of habitat suitability criteria. Although problems associated with credibility can often be linked to such topics as accuracy, precision, and transferability, there are certain perceptions about criteria that may be more difficult to overcome. One such perception is that criteria based on field observation are inherently more accurate, more precise, and more transferable than criteria based on professional judgement. While this is often the case, it is not always true. One study (Baldridge 1981) showed that a group of experienced fisheries biologists could derive category I criteria that were essentially identical to category II curves that were developed independently. The accuracy of the category I curves is intimately related to the expertise of the people who developed them. Unfortunately, the level of expertise may not become evident until higher-level categories are developed.

Users of the IFIM must critically evaluate the credibility of the criteria to be applied to a problem, and determine the best way to maintain it. It might seem that the collection of site specific data would always be the best way to maintain credibility, but there are many exceptions to this idea. There is a time and a place for every type of criteria, and it is the user's responsibility to ensure that the type is appropriate to the situation. Category III criteria should be more accurate than category I, but sometimes the perception of accuracy is more important than accuracy itself. This is particularly true under arbitrator-dominated decision processes. When instream flow issues are subject to negotiation, the best way to maintain credibility may be to include all interested parties in the criteria development process. This can often be done simply by assembling a panel of experts from the various interest groups and developing category I criteria. Sometimes, the best approach is the joint development of category II or III criteria, or to engage in some type of joint verification study.

There is also a distinction between development of criteria for a particular application and the broader context of criteria research. While category I criteria might suffice for the former, the development of highly transferable category III data is the implied goal of the latter. Such data are usually expensive and must be of the highest quality and transferability in order for their development to be cost effective. At a time when research dollars are at a premium, it is important for criteria to be as broadly applicable as possible, without compromising the accuracy implied by site-specific data. Realizing this goal requires close attention to study planning, sampling design and tactics, analytical techniques, and, most of all, documentation. These are the dominant topics of this information paper, although the research community is not the only intended audience. It is equally important for the user community to understand these concepts so that rational evaluations of non-site-specific (off-site) criteria can be made.

Chapter 2 discusses the most important phase in the development of any type of criteria, the plan of study. It is at this time that decisions are made regarding the selection of the target species and life stages, the category of criteria to be developed, the stratification of data, selection of study streams, sampling strategies and methods, variables to be measured, and analytical techniques to be applied. All of these factors can influence the utility of criteria outside the stream in which they were developed. Virtually all problems associated with the comprehensiveness of the criteria are created or solved at this stage. Transferability is at least partially controlled by decisions made during the study design phase, and many quality control problems can be anticipated and avoided at this time. Finally, a good study plan will also include an estimation of costs, both in terms of time and money.

The development of category I criteria is the subject of Chapter 3. Although such criteria might seem less desirable that those of higher levels, there may be occasions when these are the most accurate criteria obtainable. A variety of options exist for the formulation of criteria at this level, ranging from informal roundtable discussions to limited field reconnaissance studies. There are numerous advantages to such criteria, in addition to minimal cost and time commitments. One of their uses is in the role of interim criteria, allowing quick analysis of a problem. Many times, such interim criteria reveal obvious solutions without resorting to the development or evaluation of more expensive criteria.

Chapter 4 examines field techniques useful in the development of category II and III criteria. While direct observation techniques are usually preferred, there are many situations where these are either biased or infeasible. Several observation and capture techniques are described in this chapter, identifying the strengths, weaknesses, and potential sources of data bias associated with each method. Techniques for quantifying habitat availability are also described in this chapter.

Analytical techniques for fitting curves to data are discussed in Chapter 5. These techniques include histogram analysis, the use of nonparametric statistics, and nonlinear regression. This chapter also explores the ramifications of the various methods of computing habitat preference.

Chapter 6 is devoted to the subjects of criteria evaluation, verification, modification, and testing. This may be the most important chapter for the IFIM user, as it describes methods by which criteria from outside sources may be tested for use in a study stream. These methods range from relatively cursory methods of examining criteria, to more sophisticated techniques involving computerized predictions of fish locations and abundances in streams.

2. DEVELOPING THE PLAN OF STUDY

A study plan is an essential component of any habitat suitability study. Study plans may be simple and rather informal when the study is of limited scope and the criteria are not intended to be used in other geographic areas. If the intent of the study is to produce transferable, comprehensive criteria, however, the study plan must anticipate as many potential sampling problems as possible. Although it may be impossible to foresee and avoid all problems, failure to develop a study plan nearly guarantees that some difficulties or uncertainties will arise. If nothing else, the study plan serves to remind the investigator what the goals and objectives of the study are, preventing the waste of time and resources on interesting, but unproductive tangents.

There are several common components to any plan of study, regardless of the category of criteria to be developed. These include:

(1) a statement of purpose and objectives;

(2) a list of target species or groups of species, and some rationale for their selection;

(3) a description of the various data stratifications to be used in the study; and

(4) a list of the variables to be measured, or for which criteria will be developed, and an explanation of how they will be expressed or recorded (termed the sampling protocol).

Several additional items must be included in the study plans for category II or category III criteria:

(1) selection of the study stream(s);

(2) identification of appropriate sampling or observation methods, and compatible sampling strategies and tactics;

(3) an estimate of sample size requirements; and

(4) determination of equipment needs and potential suppliers.

2.1 PURPOSE AND OBJECTIVES

The statements of purpose and objectives are, word for word, the most important parts of the study plan. Although usually the shortest sections, they set the stage for everything that follows. The statement of purpose should state the reason for conducting a study and the intended use of the criteria. A study to develop in-house criteria for a specific application may be quite different from one intended to produce widely usable and transferable data, applicable to many streams and instream flow studies. Although difficult to elucidate, the expectations and acceptable limitations of the criteria should be expressed as clearly as possible in the statement of purpose.

The objectives may be stated in only a sentence or two. As a rule, objectives should be succinct to the point of bluntness and should concentrate on the ultimate goal of the study. The statement of objectives should include, at least, the proposed category of the criteria and its intended use. It may contain a list of target species and life stages, although this will be included in the text of the study plan and might be redundant. Any reference to how, where, or when the study is to be done is inappropriate in the statement of objectives.

2.2 SELECTING THE TARGET SPECIES

Several strategies can be used in the selection of target species. The most elementary decision is whether the study is to concentrate on one or two species or whether data will be collected for a large number of species. There are advantages and liabilities with either approach, the latter being more generally efficient. Unless some limits are placed on the number of species, however, the effort expended in obtaining data on the species of primary interest will be diluted among those of lesser interest. There are also some logistical limitations to the number of species that can be observed, sampled, or monitored, depending on the method used.

The central issue in the selection of a few target species is what constitutes a primary interest. This definition is highly influenced by the intended audience. A limited audience with a limited interest might require focusing the study on a single target species. A different target species or group of species might be needed to satisfy an audience with broader (or, at least, different) interests. A typical example of this divergence of interest is the managerial dichotomy between endangered species and sport fisheries in some river systems.

Unless a target species has been identified by a narrowly defined interest group, it is most desirable to select species that will provide the maximum usable information to the largest possible audience. There are at least two classification systems that can be used to make this selection: one based on fisheries management interests and the other on the species' adaptation to riverine environments. The ultimate selection may be based on both systems.

2.2.1 Target Species Classifications

Target species can be classified by management goals of the intended audience. When evaluating the potential utility of criteria, it is important to recognize both the breadth and the intensity of the audience. Sport species and endangered species both rank very high in the hierarchy of management goals. Endangered species may be of interest to a relatively small audience, but the intensity of interest is very high. Game species often enjoy a larger following than endangered species, and frequently are supported by an interest group as intense as those supporting endangered species. There is rarely any conflict in the allocation of study resources between the two groups, but when there is, the statutory support for endangered species usually gives them precedence over game species.

Two classes of organisms occupy a lower position in the management hierarchy, even though they may actually be more important from an ecological perspective. One is the support group, species upon which the management species are dependent at one or more phases of their life history. Forage fish, such as small cyprinids and catostomids, and various groups of macroinvertebrates typically fall into this category. The other group is composed of species that are competitors or predators of the primary management species. Because many sport fish are often the top predators in a given river, the predator/competitor group is often overlooked in the selection of target species.

Several factors may be involved in the lesser popularity of the support and predator/competitor groups as target species. The most compelling reason may be simple name recognition; it may be easier to justify funding a study on rainbow trout than one on rainbow darters. Related to the name recognition factor is the "weak link" syndrome. Because the primary species of interest are most commonly the game species, the importance of the support and predator/competitor groups to the game species must be demonstrated. This linkage may not always be intuitively obvious and the ecological importance of a particular group might be overlooked. There may also be a problem of species dilution among the support groups. It may be difficult to demonstrate that a game species' survival is dependent on only one or two forage species. So it may become necessary to study the whole assemblage of support species, either individually or as guilds of species having similar habitat requirements. Finally, there may often be an added barrier imposed by difficulty in taxonomic identification, especially in the support group. Many forage species, such as young cyprinids, look very much alike and may require identification under a dissecting microscope. This dictates the collection, preservation, and labeling of samples rather than simple field identification.

Name recognition, linkage, and taxonomy, are special problems in the selection of target species, but linkage is really the only legitimate reason to include or exclude a group of potential target species. This linkage can often be strengthened by evaluating the adaptations to riverine environments made by the species of primary management interest.

The three behavioral categories by which river-dwelling organisms may be classified include: obligate riverine, facultative riverine, and facultative lacustrine. An obligate riverine species is one that requires a river environment for the completion of part or all its life cycle. The requirement of running water, such as by trout and salmon for successful reproduction, is characteristic of obligate riverine species. A facultative riverine species can live in either lakes or streams, but makes behavioral adaptations to take advantage of the microhabitat types available in rivers. This group will often utilize areas in rivers that are quite different from conditions they would use in lakes. Completely different behavior is demonstrated by facultative lacustrine species, which seek out conditions in rivers that are as close to a lake as possible. The smallmouth bass often fits the description of a facultative riverine species, whereas the spotted and largemouth basses are more typically facultative lacustrine.

2.2.2 Research Priorities by Species

The relevance of the previous categorization to the selection of target species lies in the species' sensitivity to streamflow. Obligate and facultative riverine species are more dependent on streamflow and, therefore, more sensitive to changes in riverine microhabitat. Lacustrine species tend to reside in robust habitat types: deep pools, backwaters, and slough channels. These habitat types are not very sensitive to flow changes. It is for this group, in particular, that the support and competitor/predator groups become increasingly important. The habitat vs. discharge relationships for lacustrine species nearly always peak at a very low (often zero) flow. While low flows usually do provide a maximum usable area for a lacustrine species, these are rarely the most beneficial flows for the species. Such low flows can reduce the habitat potential for one or more support groups to the point that food or water quality becomes more limiting than space for the management species.

A hierarchy of potential target species emerges when both classification systems are considered simultaneously. High-priority management species that are obligate riverine will usually provide the most broadly useful information. Lacustrine species, regardless of their management priority, provide the least. Obligate or facultative riverine species in the support category are generally high in this hierarchy. They should always rank higher than lacustrine species and may outrank riverine sport species on occasion.

Another aspect of target species selection is the amount of information currently available for a species. Conducting a duplicate study on a much-studied species may be a waste of time and resources, unless the goal of the study is to test the quality of the existing data or increase its comprehensiveness. While it is desirable to conduct replicate studies to update and improve existing criteria, there are literally hundreds of species for which no criteria are available.

Table 1 contains the elements of a ranking system for the selection of target species. This system is based on three ranking criteria: fisheries management objectives (A), riverine adaptation (B), and existing information (C). Each potential target species is assigned one score for each of the

Table 1. Ranking criteria and scores for assigning research priorities to potential target species.

Ranking criteria	Score
A. Importance to fisheries management	
1. Threatened or endangered species	10
2. Sport or commercial species	
a. Highest management priority	10
b. High management priority	8
c. Moderate management priority	5
d. Low management priority	2
e. Disinterest	0
3. Support or predator/competitor species	
a. Intimately linked to well-being of management species	10
b. Strongly linked to well-being of management species	8
c. Moderately linked to well-being of management species	2
e. Insignificant to well-being of management species	0
B. River adaptation	
1. Obligate riverine	10
2. Facultative riverine	
a. Tends toward riverine habits most of the time	8
b. Tends toward riverine habits some of the time	6
c. Tends toward lacustrine habits most of the time	4
3. Facultative lacustrine	2
C. Status of existing information	
1. No information available on species	10
2. Existing criteria are incomplete or were developed in another geographic location	8
3. Existing criteria are mostly complete and fewer than three studies have been conducted in same geographic region	6
4. Existing criteria are mostly complete, and three to five studies have been conducted in same geographic region	4
5. Existing criteria are complete, and more than five studies have been conducted in same geographic region	0

three categories and a total score is calculated by adding the category scores. The highest possible score is 30. (Conceivably, the score could be 40 by double-counting category A. Such a double count would be justified for the greenback cutthroat trout, a highly valued game fish that is also endangered. In practice, such double counting is rarely necessary, but may be useful in tie-breaking situations.)

2.2.3 Guilds and Indicator Species

The system for assigning research priorities discussed in Section 2.2.2 will probably be most useful for habitat criteria studies that are oriented to individual species. The value of such a system lies in the ability to determine individual species for which habitat information is most needed or which would provide the greatest geographic coverage. In certain situations, however, it is impractical to conduct a separate habitat suitability study on every species. This is particularly true when dealing with forage fishes and macroinvertebrates. Unless the primary management species feeds on only one or two forage species, some strategy is desirable to avoid studying every species in the river. The most common approach is to group species that behave similarly into guilds.

The term "guild" has been variously defined and interpreted. One of the more modern definitions is attributed to Root (1967), who defined a guild as a "...group of species that exploit the same class of environmental resources in a similar way." Balon (1975) interprets this definition to mean that the primary characteristic for classification is diet. Balon built on the earlier works of Kryzhanovsky (1949) and developed a system of reproductive guilds for fishes, based on the type of habitat used for spawning, parental behavior, fecundity, and physiological adaptations of the eggs and larvae.

Grouping animals into habitat utilization guilds is an attractive idea, but one that may confront the investigator with a curious paradox. The reason for identifying a guild is to develop criteria for the group as a whole, rather than studying each species individually. Without some knowledge about an individual species' habitat preferences, however, it is impossible to determine the guild it belongs in. Furthermore, guild membership may change during various portions of the life histories of the collective species. Species that share a common habitat type as fry may share different habitat types with other species as adults, for example.

Procedures for determining guild membership are variable, depending on how rigorously the investigator wishes to define the guild. The simplest of all the guild strategies applies to aquatic macroinvertebrates. Habitat suitability criteria can be developed on the basis of biomass density (mean dry weight per unit area). In this case, the guild is composed of all species of aquatic macroinvertebrates and the only sorting done to the samples involves the removal of debris, stones, and terrestrial invertebrates. Although this technique may seem overly simplified, the resulting criteria will show the types of microhabitats that support the highest densities of macroinvertebrates. Similarly, repeated sampling of similar areas can yield an estimate of benthic invertebrate production, which may be a better overall index. This simplified guilding strategy produces criteria with two implied assumptions. The first is that microhabitat areas that produce or support the maximum biomass of macroinvertebrates supply the bulk of the aquatic food supply to the fish. This is probably not a bad assumption. The second assumption is that the availability of invertebrates is the same, on average, from all habitat types. This assumption may be violated quite frequently, but if the first assumption is valid, it may not make much difference.

One way to circumvent the assumption of equal availability is to divide the macroinvertebrate community into accessible vs. nonaccessible species. The easiest way to do this is to take a series of drift samples and determine which species are most common in the drift. Then, criteria are based on the density or production of these species, only. This approach would be most valid in fish communities where drift-feeding is the primary food acquisition mechanism. If the fish actively forage on the bed, an analysis of stomach contents may be required to determine which species contribute most to the diet. Investigators using this approach must be careful not to bias the guild membership toward species that are hard to digest, as these may appear to be more abundant in the gut samples.

The functional feeding group classification described by Cummins (1979) is a more traditional method of assigning guild memberships among macroinvertebrate species. This classification is based on morphological and behavioral adaptations for food acquisition. Functional groups are convenient guild classifications for the development of habitat suitability criteria, but they can be cumbersome and confusing when used in a PHABSIM analysis. The reason is that a user, when evaluating potential impacts of food-producing habitat on fish, is generally more interested in total (or at least, most substantial) food-producing habitat and not its components. If the components are mutually exclusive (e.g., collectors cannot co-occur with scrapers), then the predicted habitats are additive and there is no problem. If two or more functional groups partially overlap in their habitat utilization, however, the predicted habitats must be weighted before they can be added. Since the weighting factors between functional groups are usually unknown, a user can be placed in the unenviable position of trading off one functional group for another without any knowledge about overall food supply implications of doing so. Segregation of macroinvertebrates into functional groups may ultimately prove to be the best guilding strategy. But, until the relative contributions of each group to the total fish diet (i.e., the weighting factors) are determined, such a classification cannot be used effectively. This does not mean that criteria by functional group should not be developed. It does mean that such a classification should be accompanied by criteria based on relative density or production of all species, or those most common in the diet.

Forage fishes may not exhibit the degree of feeding and microhabitat specialization that characterizes the aquatic macroinvertebrate community. Nevertheless, it may be possible to develop forage fish guilding strategies similar to those described by Balon (1975). The class boundaries for microhabitat utilization by forage species are likely to be less distinct than those for reproductive or functional feeding guilds, however, so it may be necessary to assign guild membership somewhat more arbitrarily.

The most common approach to the development of microhabitat utilization guilds is to subdivide the river environment into distinct microhabitat types (e.g., main channel riffles, side channels, connected sloughs and oxbows, eddies, backwaters) and determine the species most commonly associated with each subdivision. There are numerous examples of this type of habitat association in the literature. (see, for example, Thompson and Hunt 1930; Martin and Campbell 1953; Minckley 1963). The descriptions provided in the literature may often suffice to assign species to their respective guilds, without conducting a field survey, but it may be necessary to do some preliminary sampling of predetermined habitat classes in order to determine guild membership. The study by Martin and Campbell (1953) provides an excellent prototype for this type of survey. They used current velocity as the primary demarcation of four major habitat types (backwaters, pools, riffle margin, and riffle channel) and then subdivided the species according to their vertical distribution (benthic vs. upper water column) and cover orientation (open water vs. cover).

The distribution of species among the various habitat groupings may often be as distinct as Martin and Campbell (1953) indicate. More commonly, some species will occur in more than one habitat type, especially if the subdivisions between types are subtle. There are at least two ways of handling these species. The easiest is to develop a guild for ubiquitous species. This may be the most satisfactory solution if there are only a few species (i.e., three or less) that could be considered ubiquitous, and they are relatively unimportant to the principal management species. The species, in this case, could be dismissed with little or no consequence. A different scenario develops if many species can be considered ubiquitous and they are very important to the management species of interest. The implication of a species that is present everywhere is that it is cosmopolitan in its habitat usage. Therefore, it can presumably find suitable habitat somewhere in the river, provided that the streambed is not totally dry. The wisdom of attempting to assemble habitat suitability criteria for a ubiquitous guild should be seriously questioned. However, before dismissing an entire guild, it is important to be certain that its members are really ubiquitous, and not simply represented in small numbers in several of the habitat types.

Species may occur in several habitat types, but may be most abundant in only one or two of them. It is also likely that the environments in two or more habitat types are very similar if the same animals occur in them at about the same level of abundance. Where this happens, the best approach is to combine the two habitat types into a single guilding unit. If a species is truly ubiquitous and of major importance to the principal management species, there may be little choice but to develop criteria for that species, individually.

2.3 DATA STRATIFICATION

Data stratification refers to the subdivision of criteria for a species to reflect changes in habitat utilization patterns. The most common stratification is to divide a species into size classes, age groups, or life stages. Other strata include seasonal or diurnal changes in habitat utilization, activity patterns, species associations, water quality characteristics, or sampling gear. Cover-conditional criteria, introduced in Section 1.1.2, is merely a different way of stratifying data.

The degree of data stratification is one of the most important considerations of a criteria study. Each stratum essentially represents a separate study, so the more of them there are, the greater the total data (or information) requirements to complete the study. Overstratification, if there is such a thing, affects only the time and money required to develop a complete set of criteria. This is important, but not as serious as understratification. The risk of overlooking a critical habitat type for a life stage, activity, or time of year increases as fewer stratifications are made. Conversely, compensatory habitat utilization mechanisms are ignored if data are not stratified by cover type. In any event, the degree of data stratification should be dictated by the behavior of the fish, not by the study's budget. Some species require many disparate habitats during the course of their lives and, consequently, can only be described by highly stratified criteria. Other species may use the same type of habitat throughout their lives and would need little or no subdivision. If money is an overriding concern, it is better to reduce the number of target species than to understratify the data.

There are no standardized procedures or nomenclature for stratifying criteria, although some standardization seems ultimately desirable. At this time, many researchers tend to use different terms rather synonymously. The following sections describe some of the more common data stratification units, synonymous terminology, and recommended nomenclature.

2.3.1 Stratification Based on Size

This is, by far, the most common unit of criteria stratification. It is common knowledge that many (if not most) species of fish exhibit changing patterns of habitat utilization as they grow. Synonyms include size class, life stage, and age group. Of these, size class is the preferred unit of stratification, because it is more precise than life stage or age group. Once the size classes have been determined, however, they may be cross-referenced to terms such as post-larvae, fry, or age 0+. It may be desirable to increase the number of strata for younger fish, because growth is rapid during the first year or two. More shifts in habitat utilization are likely to occur prior to adulthood than afterward.

2.3.2 Stratification Based on Time or Activity

The activity patterns of fish frequently vary by season and by time of day. Examples of seasonal variations include migrations, use of staging habitats, spawning, "rearing," and hibernation. With the exception of those activities relating to migration and spawning, most criteria can be easily stratified by season, the most important division being between winter and summer. Basing the stratification on the season is usually more exact than attempting to distinguish specific activities, such as rearing or hibernation.

There may be several important stratifications relating to reproduction for many species: migration, staging, egg deposition, and post-spawning behavior (e.g., nest guarding and incubation). Migration usually refers to the mass movement of a large number of fish prior to spawning. Typical criteria associated with migration include minimum passage depth and maximum sustained or burst swimming speed. However, other types of information relating to migration cues, such as a rise in stage or temperature, are also very important. Staging refers to the use of refuge areas, often pools or backwaters, immediately preceeding spawning or migration. The suitability of some spawning sites appears to be related to the proximity to a suitable staging area. Therefore, in addition to traditional PHABSIM criteria, it may be necessary to determine a threshold distance between staging areas and spawning sites. Spawning or egg deposition are either the easiest or the most difficult activities to develop criteria for, depending on whether or not the fish can be observed. If they can only be sampled or monitored by radiotelemetry it is very difficult to determine the actual spawning act. Moreover, the selection of a spawning site may be influenced by variables not typically used in PHABSIM, such as groundwater upwelling. (Upwelling can be included in PHABSIM analyses, but only if the criteria include it as an important feature of the stream.)

Changes in habitat utilization on a daily basis are also quite common. Many species move from one area to another between day and night, presumably from a "resting" habitat to a "feeding" habitat. Sometimes, only the movement is detectable and not the activity. Stratification according to time of day can be substituted for the actual activity, however, if the shift occurs on a relatively predictable diurnal pattern.

Gosse (1982) reported one type of habitat division that appears to be time independent. He found the same life stages of the same species utilizing two distinct habitat types, depending on their mode of feeding. Stationary swimming was defined as "maintaining a stationary position by actively swimming against the current." Random swimming was defined as "swimming without orientation toward a current (found only in low velocity water) that did not produce a net change in location." Fish engaged in one mode of habitat utilization did not switch to another habitat type during the course of a day. Like spawning, this kind of habitat division would be very difficult to detect, unless the fish could be observed directly. One indication that a set of criteria data contains observations of more than one activity is a bimodal frequency distribution. Whenever a bimodal or polymodal distribution is encountered, the investigator should immediately suspect understratification.

2.4 SAMPLING PROTOCOL

The sampling protocol is a formal description of the variables for which criteria will be developed or data collected, procedures for measuring or describing each variable, and procedures for abbreviating, codifying and recording data. The purposes for developing a sampling protocol are to enhance consistency and to avoid ambiguity. One of the most frustrating problems in conducting an instream flow study is to find that whoever developed the criteria used ambiguous terminology, incompatible units of measurement, undecipherable codes, or measured the "wrong" variables entirely. The only thing worse is to learn that such errors were self-inflicted.

The minimum amount of information that can be used effectively to construct habitat suitability criteria includes: a species descriptor, a life stage or size descriptor, frequency (number of fish per observation), depth, mean column velocity, and a substrate or cover descriptor. This fundamental data base can easily be doubled, providing information that can be input directly to PHABSIM or used by potential users of the criteria to evaluate the transportability of the criteria to other stream systems. It is especially for this latter group that criteria investigators must be absolutely clear in their description of the sampling protocol. Please note that "sampling protocol" is equally applicable to category I criteria, even though the term implies measurement of variables.

There are at least three areas of sampling protocol that should be of concern to researchers. The first is the use of abbreviations or numerical codes to denote species, life stage, activity, substrate composition, or cover types. The second area of concern is the consistent use of units of measurement. Finally, there are many optional variables, in addition to the required ones, that can be used in conjunction with PHABSIM. Researchers and users alike should be aware of how such ancillary information is used, and why and when it is needed.

2.4.1 Abbreviations and Coding Systems

a. Species. The first entry for a line of data describing a fish sighting is usually an abbreviation or code for the species. Researchers can help themselves, as well as others, by recording the translation between the abbreviations and the respective species' or guild names. Many investigators prefer the use of common names and abbreviations, such as RBT for rainbow trout or SMB for smallmouth bass. Others prefer the use of scientific names. It does not make much difference, as long as the translation is written down. It is confusing, however, when data recorders switch back and forth between common and scientific names for the same organism. The use of either the accepted common or scientific name as published by the American Fisheries Society (Bailey et al. 1970) is highly recommended. Avoid colloquialisms that can be confused with those of other species (e.g., bullhead, rockfish) and abbreviations that could mean any of several species (e.g., CS = coho salmon, chum salmon, chinook salmon, common shiner, or common sucker).

b. Life stage. As mentioned in Section 2.3, the best way to denote different sizes, life stages, or ages of fish is to record the length of each individual. Normally, a length frequency histogram of all fish of a species can be used to delineate age groups and life stages. This concept works well in all but two situations. The first is where the fish are observed, but not collected. The second is where several hundred post-larval fish are scooped up in a single sample. In either case, it is impractical, if not impossible, to measure individual fish. The best alternative is to categorize the individuals into predetermined size classes. These size classes should conform to age-length groupings found in the literature or (preferably) localized age and growth studies. Size classes should not be systematic (uniform increment) or geometric unless supported by some form of length-frequency information. Arbitrary size classes should be discouraged.

Size classes are also preferable to the use of life stage descriptors. The simple reason is that fish of the same age and life stage can vary dramatically in size. Although such size differences can be found quite often among immature life stages, the most significant discrepancies occur among the adults. For example, the average brown trout adult in the Colorado Rockies might be 30 cm long and weigh 400-500 grams. Those in the Green River in Utah might be twice that length and more than twice the weight. To assume they both select the same habitat would be a mistake.

c. Activity. A fish's activity may play an important role in its habitat selection during different times of the day or year. If direct observations are made, the activity should be monitored and noted. A simple code can be used to denote activity, such as:

1 - resting, holding

2 - feeding, foraging

3 - stationary swimming

4 - random swimming

5 - staging

6 - spawning, nest guarding

7 - hibernating

8 - migrating

9 - escape

Unless fish observations are made directly, it is nearly impossible to tell what the fish were doing at the time of sampling. Activity may not always be detectable or fit a preassigned category, but many shifts in activity occur on diurnal and seasonal bases. Therefore, recording the dates and times of samples will often suffice to distinguish some of the changes in activity. Time and date may not be enough to distinguish all activities, such as pelagic spawning. Considerations of special problems related to activity will be discussed in Chapter 4.

d. Substrate codes. The procedure for describing substrate is especially important in a study design because there are numerous classification and coding systems, similar in appearance, but very different in information content. There are actually three components to the substrate coding procedure to be addressed in the study plan: the dimensions of the sample, the particle size classification system, and the coding system.

Sampling dimension refers to the area around an organism that represents the substrate being utilized. Defining the horizontal dimension is usually not much of a problem for those organisms or life stages that are highly influenced by substrate, such as macroinvertebrates or demersal spawners. The most common source of confusion is illustrated by an adult fish hiding behind a boulder, over a gravel substrate. The substrate that the fish is actually using is the boulder, and it may be oblivious to the substrate immediately underneath. The preferred approach, in this case, would be to record the substrate as gravel, but to record the boulder as a large instream cover object. There would be nothing wrong, however, in describing the same observation as a boulder embedded in gravel. The point is to discuss such situations ahead of time and decide how to handle them in a consistent manner.

The vertical dimension of the substrate can be more troublesome than the horizontal, especially when dealing with benthic macroinvertebrates. Many investigators record the undisturbed surface materials, prior to sampling, as the utilized substrate. As long as the matrix is uniformly mixed, this is no problem, but if the substrate is armored, there may be more invertebrates below the surface than on the surface. Conversely, there may be very few invertebrates deep in the substrate, if it is covered with fine sediment. Since it is not always easy to predict which aspect of the substrate is more important, it may be necessary to record both the surface description and the at-depth-of-sample description. This paradox will also become apparent to researchers describing salmonid spawning substrate, where the substrate in the redd will have been disturbed and most of the fines removed. This is not the same substrate originally excavated by the fish. Thus, the redd materials would represent the "at-depth" sample. The surface description might need to be made adjacent to the redd, or prior to spawning.

The size range (or type of vegetation) assigned to each substrate descriptor is the second item that must be defined. Platts et al. (1983) suggest the use of substrate terminology and size classes accepted by the American Geophysical Union (AGU). Since one of the biggest problems in substrate analysis is lack of standardization, this recommendation has merit. The AGU classification of substrate materials is a geometric scale, which means that the size range for each class is based on a doubling of the size range for the previous class. This process inevitably leads to very fine subdivisions of small sediment sizes and broad size ranges for larger size classes. For most biological studies, it is usually sufficient to be able to distinguish between clay and silt or silt and sand without resorting to fine distinctions within these classes.

The AGU substrate classification system is well suited to many biological applications, but it does not include two very important substrate types: vegetation and bedrock. It has been repeatedly demonstrated that vegetation and detritus are extremely important substrates for the production of aquatic macroinvertebrates (Percival and Whitehead 1929; Egglishaw 1964; Cummins 1966; Gosse 1982). Furthermore, macroinvertebrate density may vary depending on the species of vegetation or the type of detritus, so it may be necessary to subdivide these substrate types into specific components.

Bedrock cannot be dismissed as a substrate type, either. It may form most of the streambed in many eastern U.S. streams. As a substrate type, bedrock may be a valuable form of cover for fish; it may be teaming with aquatic macroinvertebrates; or it may be a lifeless slab of rock. The value of bedrock as a substrate type depends on the degree of fracturing, whether its bedding plane is tilted or flat, and the orientation of the bedding plane to the flow of the river. Highly fractured bedrock contains numerous cracks and crevices which may be colonized by macroinvertebrates. Tilted slabs of bedrock may provide velocity shelters for fish if the slabs are perpendicular to the flow, but may be nearly worthless if the slabs are parallel to the flow.

Table 2 contains a generalized guide to substrate classification, following the AGU particle size classification and including organic materials and bedrock as substrate types. The size classes for clay, silt, and sand in Table 2 represent the combination of finer divisions from the AGU classification. It may be appropriate in some studies to combine these three classes and simply call them "fines." These small size classes should be kept separate in studies involving benthic macroinvertebrates because clay is a more productive substrate than sand (Percival and Whitehead 1929; Sprules 1947).

The final process to be determined is how mixtures of different substrate sizes are to be described. The bed materials in many streams are not of uniform size, but consist of various mixtures of different size classes. The most detailed method of describing the substrate is by sieving it to determine percent composition of each size class by weight. However, this amount of detail is unwieldy to use in PHABSIM, may be superfluous in biological significance, and would require the same level of detail in routine data collection for operational IFIM analyses. Therefore, sieve samples are probably impractical for this type of study.

Brusven (1977) developed a system by which the essential elements of the substrate can be reduced to a simple numerical code. The Brusven substrate index consists of three parts: an index for the dominant particle size, an index for the subdominant particle size, and an index describing the relative percentage of the subdominant with respect to the dominant, termed embeddedness. A three-digit code is used to describe a substrate mixture, based on these components. The index for the dominant particle size is recorded in the ten's place, the subdominant index in the one's place, and embeddedness is expressed as a decimal. If small cobble has an index of 5 and medium gravel an index of 3, a mixture containing 40% small cobble and 60% medium gravel would be expressed in code as 53.6. The original index, based on a three-digit code, only has room for particle size codes from zero to nine. By using a five digit code, however, up to 100 different size classes could be described. Using the classification scale from Table 2, a combination of large cobble with attached *Cladophora* could be described as 1303.0 to 1303.9, depending on the amount of algae, where 13 is the code for large cobble and 03 is the code for attached algae.

Table 2. Generalized substrate classes for use in studies to determine substrate utilization and preference.

Class names (optional subdivision)		Size range mm	Size range inches
01	Organic detritus		
	(logs, branches)		
	(pine needles)		
	(leaf detritus)		
02	Vascular plants		
	(*Potamogeton*)		
	(*Zanichellia*)		
	(*Ranunculus*)		
03	Attached algae		
	(*Cladophora*)		
	(*Chara*)		
	(*Nitella*)		
	Inorganic substrates		
04	Clay	.00024 - .004	9.5 E-6 - 1.58 E-4
05	Silt	.004 - .062	1.58 E-4 - 2.44 E-3
06	Sand	.062 - 2.0	2.44 E-4 - 7.87 E-2
07	Very fine gravel	2 - 4	.08 - .16
08	Fine gravel	4 - 8	.16 - 0.3
09	Medium gravel	8 - 16	0.3 - 0.6
10	Coarse gravel	16 - 32	0.6 - 1.3
11	Very coarse gravel	32 - 64	1.3 - 2.5
12	Small cobbles	64 - 128	2.5 - 5
13	Large cobbles	128 - 256	5 - 10
14	Small boulders	256 - 512	10 - 20
15	Medium boulders	512 - 1,024	20 - 40
16	Large boulders	1,024 - 2,048	40 - 80
17	Very large boulders	>2,048	>80
	Bedrock		
18	Plain, unfractured		
19	Plain, jointed		
20	Tilted, perpendicular, unfractured		
21	Tilted, parallel, unfractured		
22	Tilted, perpendicular, jointed		
23	Tilted, parallel, jointed		

Bovee (1982) and Platts et al. (1983) modified the Brusven index slightly by substituting an estimated percentage of fines for embeddedness in the decimal's place. The reasoning was that the amount of fine material filling the interstices between the larger particles is an important factor determining substrate suitability for macroinvertebrates (Percival and Whitehead 1929; Wene and Wickliff 1940; Sprules 1947) and embryo survival for benthic spawning species (Terhune 1958; Coble 1961; Johnson 1961; Condore and Kelly 1961). The example code of 53.6, shown above, would refer to a mixture of small cobble and medium gravel containing 60% sand or smaller material. Some practitioners ignore the subdominant particle size, which results in an even simpler code containing only dominant particle size and percent fines.

Although there are many ways of expressing substrate composition, the Brusven index combines descriptive flexibility with ease of use. These attributes have led to the popularity of the index among fisheries biologists. Modifications to the Brusven index are certainly possible, but some coding system like it is highly recommended. The extent to which the original or a modified version is used depends somewhat on the organism for which the criteria are being developed. Large fish, if they use or select substrate at all, often use substrate in the large cobble or boulder category. It may be more convenient to include these categories as cover descriptors rather than as substrate. Small benthic fishes, such as darters and madtoms, are often found in the pore spaces between moderate sized bed materials. The size of these pore spaces is a function of any smaller material filling the voids (e.g., a mixture of medium cobble and small gravel). The original Brusven index is well suited for describing this substrate combination. The abundance and survival of macroinvertebrates and buried embryos is also a function of the pore spaces between the larger particles, but the pore spaces may be much smaller. The percentage of sand and small materials in the pore spaces may be more important than the percent of small gravel mixed in with large gravel, so the modified Brusven index would be more appropriate for describing this mixture.

e. Cover codes. Cover is easier to codify than it is to quantify. There are two principal differences between cover and substrate, with regard to codification. First, combinations of cover types tend to be uncomplicated, compared to combinations of particle sizes in the substrate. The dominant cover type utilized by a fish is quite apparent, in most cases. Second, cover is a discrete variable; substrate is continuous. It is possible to interpolate a particle size between sand and medium gravel, but not possible to interpolate a cover type between a log and an undercut bank. Therefore, the most significant aspect of codifying cover is the grouping of individual cover features according to type or function. An example of such a cover grouping is illustrated in Table 3.

Table 3 is not comprehensive, but it should not be too difficult to determine the classification of unlisted cover objects. A major type or function of cover omitted from Table 3 is the "edge effect." Edge cover is defined as a feature of the stream providing a distinct interface between two different habitat types. There are numerous examples of such interfaces in a stream:

Table 3. Numerical cover codes, type and function groups, and typical examples of cover features suggested for cover codification in habitat suitability investigations.

Code	Type	Function	Examples
1	No cover	No cover	Open water, deep pools*
2	Instream object	Velocity shelter	Large rocks, partially buried logs, bedrock ledges
3	Instream overhead	Visual isolation (direct)	Undercut banks, floating vegetation, open log jams, surface turbulence, deep pools*
4	Offstream overhead	Visual isolation (indirect)	Overhanging canopy, shadows
5	Combination object + overhead	Combination velocity shelter + visual isolation	Root wads, brush piles, emergent vegetation, log jams, any superimposed object with overhead cover

*Deep pools can serve as overhead cover, but only in conjunction with cover conditional criteria. Refer to Section 2.4.3 for more detail.

(1) the use of shadowed areas adjacent to open sunlight;

(2) fish hiding in the wake of a boulder, next to an area of high velocity;

(3) the use by some species of shallow ledges, affording no cover, but immediately adjacent to a steep drop-off; and

(4) the use of areas of emergent vegetation at the interface with open water.

This edge effect can be extremely important, because some species and life stages use only the edge created by the cover object. For these species, more cover in a reach is not synonymous with better habitat, but more edge is. Fortunately, the use of edges is usually quite easy to detect and to incorporate into the cover code. Simply add a one in the ten's place, in front of the type code, to signify an edge effect (i.e., 4 = offstream overhead, 14 = offstream overhead with edge effect).

The cover code for a criteria study is not necessarily the same as the code used in a PHABSIM analysis, at least not at this stage of development. The cover code used in PHABSIM must account for the cover dimensions; these must be determined in the criteria study, but do not need to be incorporated in the code (see Section 2.4.3).

2.4.2 Units of Measurement

PHABSIM is based on the English decimal system of measurement; depths and velocities are measured in feet and tenths of feet. The criteria for depth and velocity should, therefore, be expressed in the same system of measurement. It is perfectly acceptable to measure these variables in either cgs or mks metric units. The resulting criteria, however, should be converted to English units prior to use in PHABSIM. An investigator should be cognizant of two guidelines relating to units. First, do not mix units. It is especially important not to mix English and mks metric, as the numbers can be easily confused. Second, be sure to document the measurement system, both in the study plan and in the field notes.

The protocol for units of measurement applies mainly to depth and velocity. It is far less important for other habitat-related variables, at least with respect to PHABSIM. There are, however, preferred or conventional units for many of the required and optional variables included in a habitat suitability study (refer to Section 2.4.3). Temperature, for example, is typically measured in degrees Celsius. Substrate sizes are often classified by centimeters, millimeters, or inches (see Table 2, for example). Cover dimensions are generally recorded in feet, inches, or centimeters. Chemical concentrations should be expressed as mg/L. Turbidity is the only variable that deviates from standard units. The preferred expression is Secchi disc visibility, in feet or meters, rather than Jackson (JTU) or nephelometric (NTU) turbidity units. As with depth and velocity, the units of measurement are less critical than documenting what they are.

2.4.3 Variables and Measurement Techniques

There are relatively few variables required for a habitat suitability study. These include a species descriptor, a life stage descriptor, an observation frequency (number of fish associated with the observation), depth, mean column velocity, and a substrate or cover descriptor, or both. Some of these items have been discussed previously, in the context of nomenclature and coding systems. In addition to the standard variables, there exist many optional variables that might be measured or described in a criteria study. These include: nose depth, nose velocity, adjacent velocity, interspersion distance, temperature, dissolved oxygen, alkalinity, conductivity, and visibility. This section discusses the measurement of these variables, situations under which certain optional measurements are desirable, and other information that should be included in a criteria data base.

a. Depth. Depth is the thickness of the column of water between the streambed and the atmosphere, usually measured with a wading rod, sounding line, acoustical sounding, or pressure sensitive depth gage. It may also be necessary to record "fish depth," in addition to total depth at each fish position. Fish depth refers to the distance above the bed occupied by the fish and is used in conjunction with "nose velocity" discussed below. The only convention in regard to fish depth is that it is typically measured as the distance of the fish from the streambed.

b. Velocity. Several different types of velocity criteria can be used with PHABSIM. The most common is the mean column velocity, which represents the arithmetic average of the velocity distribution between the bed and the surface. The second type is termed "nose velocity," measured at the vertical position in the water column actually occupied by the fish. The difference between mean column and nose velocity is the location in the water column where the measurement is made.

A standard procedure for the measurement of mean column velocity has evolved from stream gaging techniques. A single measurement at 60% of the total depth, measured down from the surface, gives a good estimate of the mean column velocity in water less than 75 cm deep. In deeper water, two measurements are taken, one at 20% of the depth and one at 80%. These two measurements are then averaged to obtain mean column velocity. Velocity profiles in turbulent water often necessitate all three measurements to obtain a reliable estimate of the mean:

$$V = \frac{V_{.8} + V_{.2} + 2V_{.6}}{4} \tag{1}$$

where V = the mean column velocity,

$V_{.8}$ = the velocity measured at 80% of the distance from the surface,

$V_{.6}$ = the velocity measured at 60% of the distance from the surface, and

$V_{.2}$ = the velocity measured at 20% of the distance from the surface.

In its current configuration, PHABSIM is a two-dimensional model. This means that the nose velocity can be simulated at any position in the water column, but the position is fixed. Given sufficient depth, some species of fish will move vertically in the water column to find their preferred nose velocity (Gosse 1982). This behavior cannot currently be simulated in PHABSIM, although future modifications may allow such simulation (producing weighted usable volume rather than weighted usable area). This limitation means that nose velocities should be measured at a constant distance from the streambed. A standard distance of 12 cm (0.4 ft) above the bed has frequently been used for the nose velocity of spawning salmon. In other cases, the standard distance will not be known, requiring the concurrent measurement of "fish depth" as discussed previously. Nose velocity can also be assigned to a vertical stratum (e.g., upper third, middle third, lower third) if the fish tend to be confined to a particular portion of the water column. If the fish are found to use the entire water column, however, mean column velocity would be the preferred variable.

A third velocity option has recently become available for use with PHABSIM. This version is designed to simulate the sheer zone, or velocity edge, effect. Some species of fish prefer low nose velocities, but only if such areas are adjacent to higher velocity water. Drift feeding by salmonids is a common example of sheer zone utilization. This option requires three types of information. The first is the focal point velocity for the fish, which may be either the mean column or nose velocity, measured at the fish's location. The second type is the adjacent velocity, measured at locations lateral to the fish's position. These may also be mean column or nose velocity measurements, but since the role of the adjacent velocity is usually food delivery, it is more logical to measure mean column velocities. The third piece of information needed is a measurement of the distance between the low velocity focal point and the high velocity zone. This information will ultimately be used to establish threshold distances, which define the effective sizes of sheer zones for different species.

The use of adjacent velocities may also be exhibited in the vertical dimension by some species. A typical example is the use of a low velocity area behind a rock, where the overstory velocity is quite high. The phenomenon is essentially the same as lateral adjacent velocities, but would require a three-dimensional version of PHABSIM. This technology is not currently available, but may be developed in the future. Therefore, it might be desirable to develop criteria for vertical adjacent velocities, even though this information cannot be used effectively at this time.

c. Substrate. As mentioned in Section 2.4.1, there are usually two or three aspects of the substrate that should be described. Depending on the coding system, the investigator will need to describe a dominant particle size, percent fines or embeddedness, and, possibly, the subdominant particle size. One aspect that must be addressed prior to describing substrate, by any method, is a definition of dominance. This definition has caused a great deal of confusion because dominance can be related to size, abundance, or both. Unless some consistent procedure is used, the substrate descriptions will inevitably contain a mixture of definitions.

There have been many attempts to develop quick and easy field methods for quantifying particle size distributions in the substrate. Unfortunately, the only completely quantitative method is to remove a sample of the substrate material, sieve it, and weigh each particle size fraction. This method will give a good estimate of the dominant particle size if it is defined as the particle size fraction that contributes the largest percentage of the total weight. The estimate of the percentage of fine materials in the matrix can range from good to poor, depending on the success of removing the sample from the river without washing all the fines out of it. This method is not quick or easy. Furthermore, most sampling procedures, with the exception of the freeze-core method (Platts and Penton 1980), destroy the arrangement of particles, which precludes any post-sampling estimation of embeddedness. The most serious drawback, however, is that criteria developed by this method would require the same level of effort in routine PHABSIM applications. Such effort might be justifiable in the development of criteria, but is impractical in operational use.

Visual estimation of the particle size distribution lies at the opposite extreme of the methodological spectrum. This method is quick and easy, but lacks the precision of the sieve method. There are several ways of improving the reproducibility of substrate descriptions, however, without resorting to sampling and sieving. Bain et al. (1985a) present a method for estimating the coarseness and heterogeneity of the substrate. They used a 2-meter, lead core rope, divided into twenty 10-cm sections alternately painted orange and white. The rope was laid on the bed, perpendicular to the flow, and the most common (i.e., dominant) particle size recorded under each 10-cm segment. The mean of the 20 values was used as an index of coarseness and the standard deviation as an index of heterogeneity. This approach will probably improve reproducibility of the substrate description, but may be incompatible with a Brusven-format coding system, because the mean may not equate with dominance and heterogeneity may not be equivalent to embeddedness.

Another technique for estimating the dominant particle size is the use of a geometric wire grid (Figure 4). The apparatus consists of a square metal frame, subdivided into grids corresponding to the particle size classifications described in Table 2. The largest dimension of the frame is 256 mm and the smallest, 8 mm. This encompasses particle size classifications from medium gravel to large cobbles, the size range that has proven to be the most difficult to judge by eye. The grid is used by laying it on the bed and comparing the grid sizes to the particle sizes. The dominant particle size is defined as the grid size that is most closely "filled" by individual particles. If most of the wires for a particular grid size fall inside the outer perimeter of individual particles, the dominant particle size is assigned to the next largest grid size, where most of the particles are contained between the wires.

Visual estimation of embeddedness or percent fines may prove to be the most reproducible and practical technique. The reason for this is that the arrangement of the particles with respect to one another may be more important than their weight composition. For example, consider a clean gravel lying beneath a thin veneer of silt. Although the silt contributes very little to the weight of the sample, it may plug all the interstices among the larger materials, rendering the substrate useless for some organisms. If the same amount of silt were completely mixed among the other particles, it would probably be inconsequential.

The danger of visually estimating the amount of fine materials in or on the bed is the potential lack of reproducibility. This problem is exacerbated as more people are involved in a study. There are several procedures, however, that can be used to increase the precision of the estimates. First, limit the accuracy to which the estimate is made. It is virtually impossible to achieve

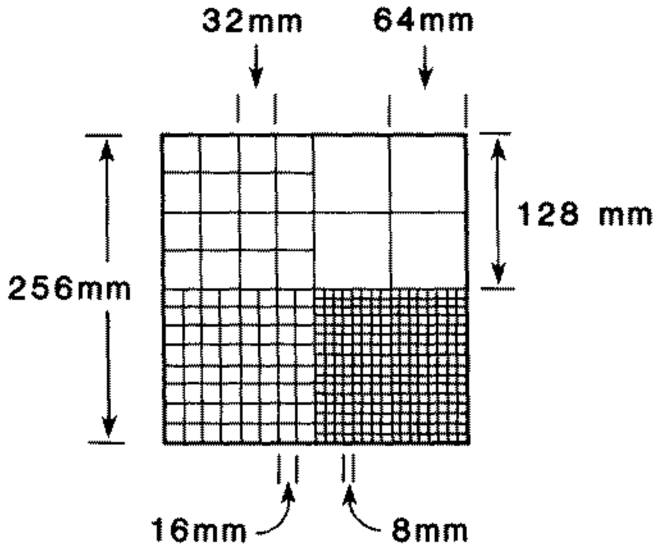


Figure 4. Geometric wire grid for estimating dominant particle size.

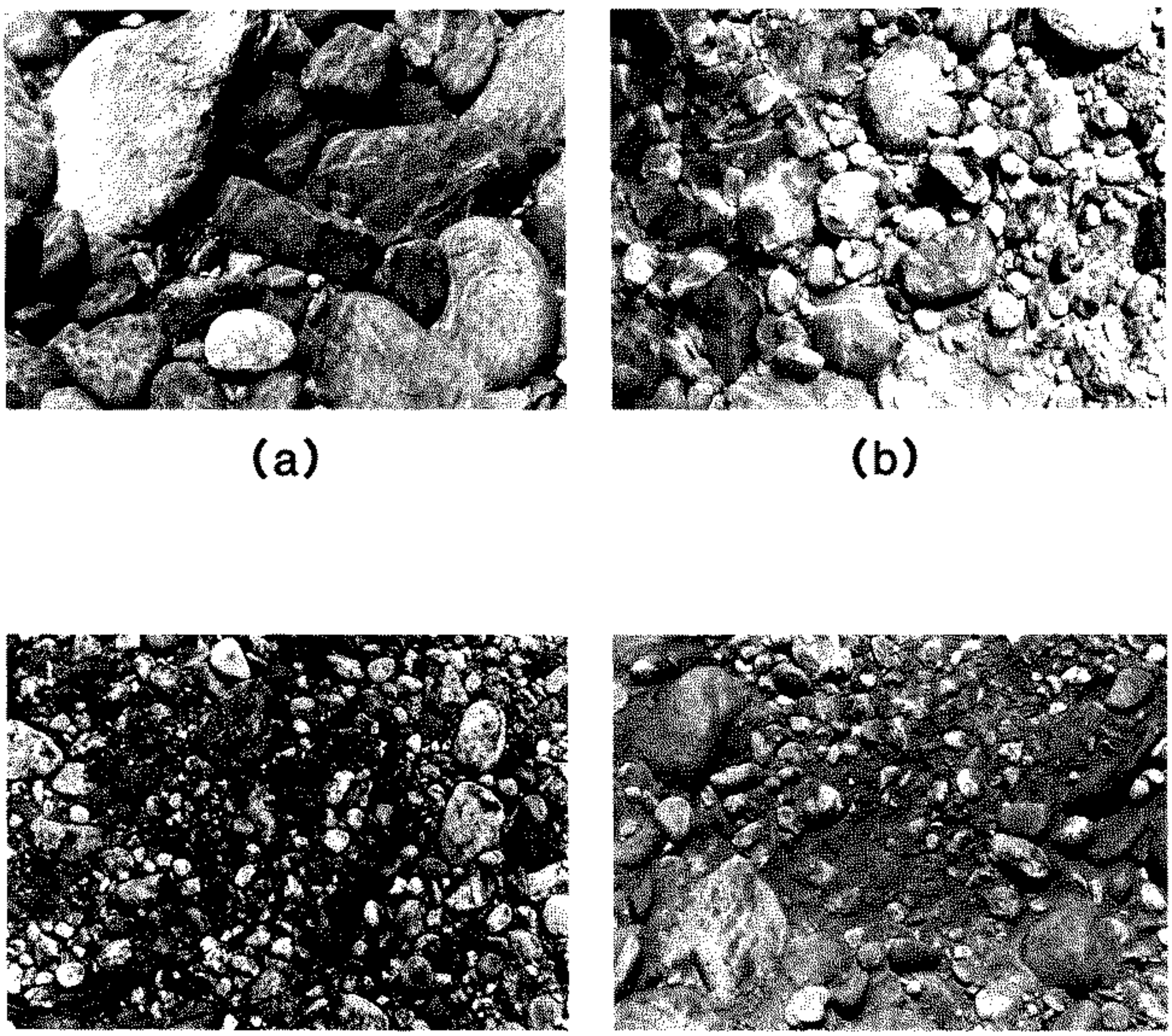
reproducibility if embeddedness is estimated to the nearest 10%, however, Platts et al. (1983) found a high level of precision among different investigators when they estimated quartile embeddedness (i.e., 0-25%, 25%-50%, 50%-75%, 75%-100%). Second, provide written criteria that describe each of the embeddedness quartiles. Table 4 contains such a description based on photographs in Platts et al. (1983). Third, whenever possible, provide several photographs of different substrate mixtures representing each of the four quartiles, as illustrated in Figure 5.

The assumption of the foregoing discussion is that the investigator can see the streambed well enough to distinguish grid sizes and estimate embeddedness. There are instances in clear water where the streambed cannot be seen this well and many others where the streambed will be totally obscured by turbidity. Visual estimations of the substrate composition can usually be conducted when the visibility is at least 15 cm, although special equipment may be needed. Different methods may be necessary when the visibility is less than that.

Surface turbulence and suspended bubbles, which prevent an unobstructed view of the streambed, are common problems encountered in clear water. The solution is simple and inexpensive: a viewbox. One of the most durable and easily-used viewboxes consists of a bucket with the bottom cut out and replaced with a piece of plexiglass. The glass is normally attached to the bucket with

Table 4. Descriptions of substrate materials by percentages of embeddedness, to the nearest quartile.

Quartile	Percent fines	Description
1	0-25	Openings between dominant sized particles appear dark and are ¼ to ½ the size of the particles. Materials seen through the openings are about the same size as the dominant particles. Edges of particles clearly discernible.
2	25-50	At least half the openings between dominant sized particles appear dark. Openings are apparent, but less than ¼ the size of the particles. Most particle edges are clearly discernible, but up to half the edges are obscured by fine materials.
3	50-75	Openings between dominant sized materials appear to be completely filled with finer materials. Less than half the edges of dominant particles are clearly discernible, but the size of the larger materials can be determined without removing them from the bed.
4	75-100	All openings between larger materials are obscured. Bed appears to consist of fine materials, but is solid to the touch. Only one or two edges of dominant particles may be visible. Size of dominant particles cannot be determined without removing them from the bed.



(a) (b) (c) (d)

Figure 5. Gravel mixtures representing differing degrees of fine particle embeddedness: (a) 0-25%; (b) 25-50%; (c) 50-75%; (d) 75-100%.

silicone caulking as shown in Figure 6. Perhaps the most important design feature of the view box is that it is round. Square viewboxes may provide a slightly larger field of view, but they have two serious deficiencies. First, a square box tends to be caught in the current and can act like a parachute to the person holding it. Second, the leading edge of the box can cause air entrainment, resulting in bubble formation on the underside of the viewplate. Round viewboxes do not seem to have either of these problems, at least not to the extent of square ones. It is extremely important to use some type of scale, such as the geometric wire grid, when using any type of viewbox. Objects viewed through any water-glass-air interface are magnified by approximately 25%. Therefore, there is a potential to overestimate bed material sizes unless some type of scale is used. (The scale must also be in the water, not in the air.)

A different type of viewbox will be needed when the depth exceeds the visibility. A prototype design of a view-tube is shown in Figure 7. The bucket used for the viewbox is replaced with a length of 15-cm (6-inch) PVC pipe. The pipe should exceed the maximum depth to be viewed by about 60 cm. Other important adaptations include an external light source, counterweight, handles, and a viewing mask at the top of the tube.

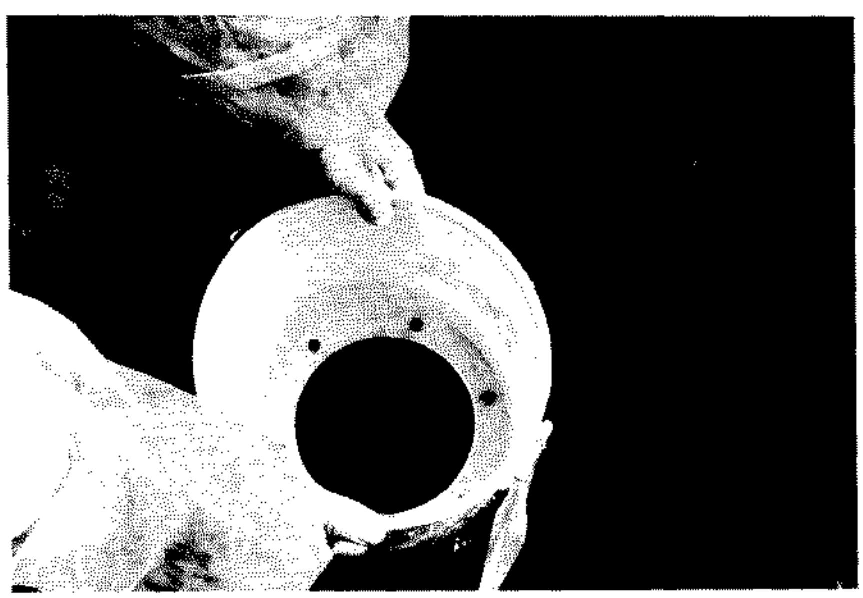


Figure 6. Bucket viewbox to determine substrate composition in relatively shallow, clear water.

Theoretically, the view-tube can be used when the visibility is as low as 5 cm, by placing the lens nearly on the bottom. However, use of this device is discouraged when the visibility is less than about 15 cm, because repeated contact with the bed will scratch the lense or cause damage to the external lights. Furthermore, the presence of suspended fine particulates may make it difficult to estimate the amount of fines in the substrate. The preferred approach in "zero visibility" water is to remove a sample with a dredge or other device. Platts et al. (1983) advocate freeze-core sampling or the use of a McNeil core sampler, because both methods preserve most of the fines in the sample. Both techniques, as well as a description of sample processing (sieving) techniques, are presented in Platts et al. (1983). These methods should be considered as a last resort because of the time involved in using them.

d. Cover. There are three different approaches to the development of cover suitability criteria. The first is to assign a numerical code and a weighting factor implying relative utilization or preference to each cover type or function. This is done by using a (normalized) histogram as illustrated in Figure 8, which shows highest usage of submerged logs, undercut banks, and rootwads, and lowest usage of areas having no cover or shadows from overhead canopy cover. Cover criteria of the form shown in Figure 8 would require the least amount of field data because the histogram for cover utilization would be developed separately from those for depth, velocity, and substate. This type of cover criteria implies that the fish exhibits the same velocity, depth, and substrate preferences regardless of what type of cover it is using.

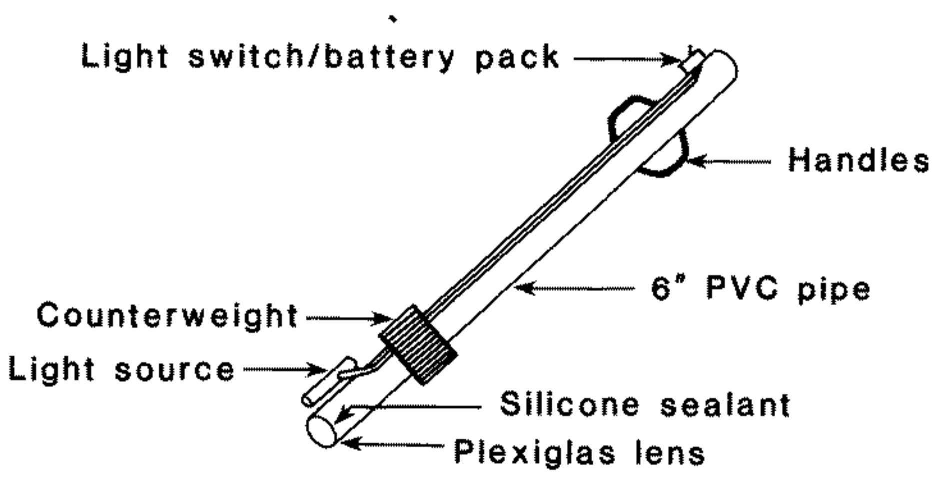


Figure 7. Prototype design of a view-tube used to examine bed materials in turbid water.

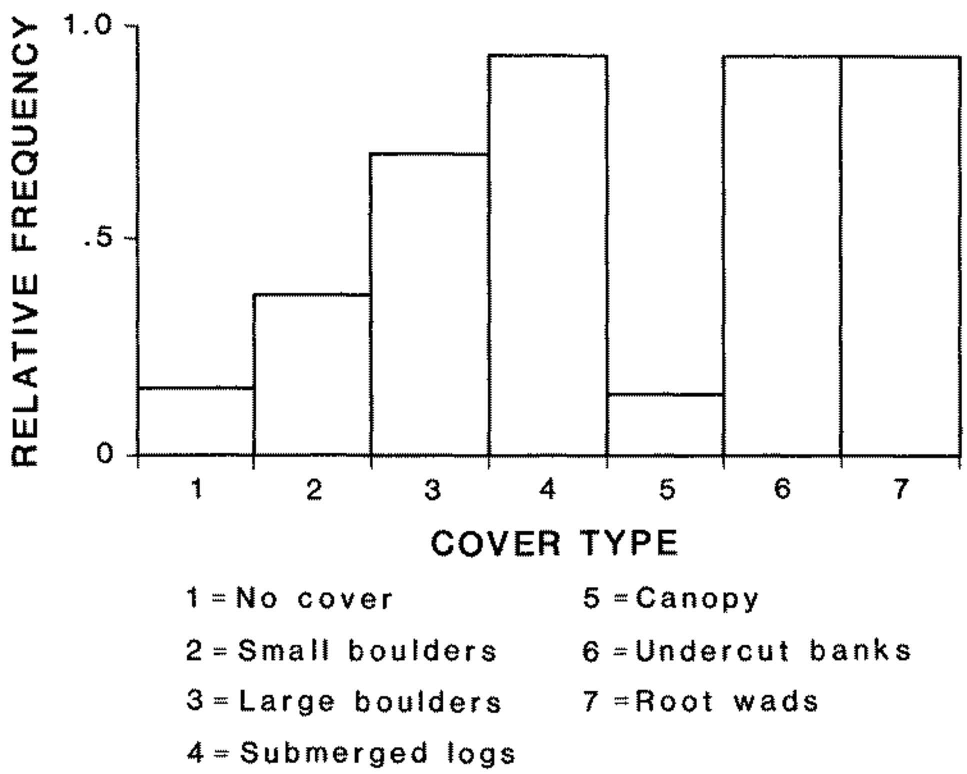


Figure 8. Example of a histogram showing relative frequency of various cover types utilized by a species of fish.

Conditional criteria, illustrated in Figure 9, allow the use of different depth, velocity, or substrate criteria as functions of the type of cover present. This figure shows that the presence of undercut banks and root wads is accompanied by high utilization of shallow, slow water. The optimum depth is slightly greater and the velocity optimum broader with submerged logs. Greater depths and velocities are used in association with boulders, whereas only deep, very slow water is used when cover is absent. Utilization of depth as a form of cover is simulated using this approach. Ordinary conditional criteria assume that the fish show no particular preference for any cover type, but alter their use of other microhabitat variables based on the cover type present. Conditional criteria require more data than those based on independent histogram analysis because depth, velocity, and perhaps, substrate criteria are developed for each cover type. This may create a problem in completing the data set if the fish tend to favor microhabitats containing cover. It may be quite easy to complete the "with cover" part of the data base, but very time consuming to obtain the necessary observations in the absence of cover.

Weighted conditional criteria combine the attributes of the two forms discussed previously by including conditional use of depth and velocity (and substrate, if appropriate) as a function of cover type <u>and</u> a relative

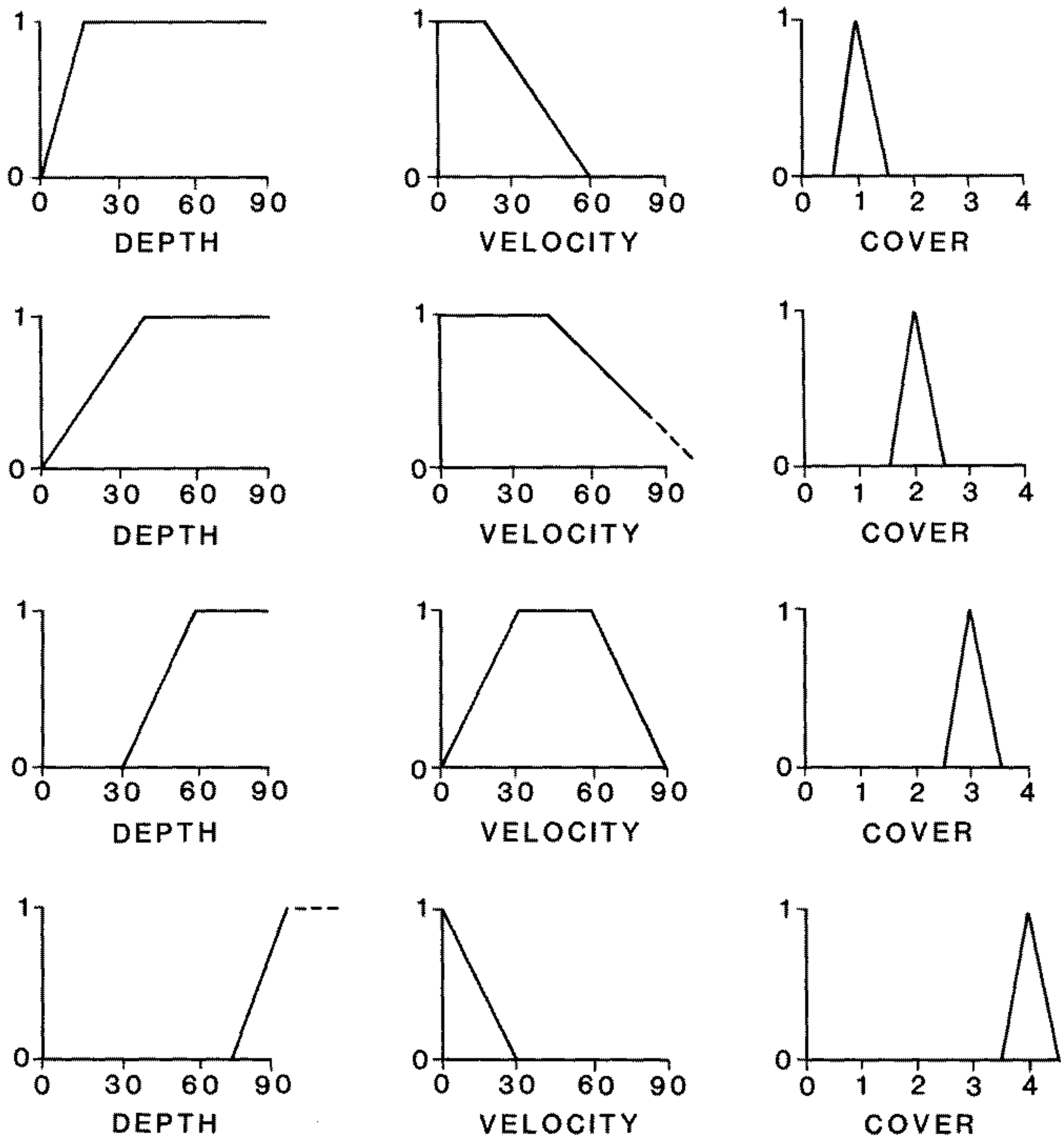


Figure 9. Example of ordinary conditional criteria for various cover types.

preference for different cover types. The use of this type of criteria in PHABSIM is very easy, but will require a different approach in data collection. Data collected for ordinary conditional criteria can not be used to detect preferential cover selection because the same number of observations are typically made in each cover type. Conversely, data collected for simple histogram analysis (Figure 8) will almost certainly be inadequate for developing conditional criteria.

The most expedient approach for developing this type of criteria is to conduct the criteria study in two parts. The first phase consists of a simple cover utilization study without collecting any depth, velocity, or substrate data. The simple histogram approach is used to determine relative cover preferences. The second phase would consist of observation and data collection in the same manner as ordinary conditional criteria. The results are then combined, as illustrated in Figure 10.

It is obvious that the manner of describing cover utilization will have a large influence on the amount of data to be collected. One way of reducing the data requirements for cover-related criteria is to use as few cover classifications as possible, but at the same time avoiding understratification problems with respect to form or function. At this stage, methods for determining the dimensions of utilized cover objects must also be considered.

Cover dimensions define the threshold size of a channel feature in order for it to provide cover for a specific organism. This information is crucial to field crews conducting instream flow studies because they need to know which features to consider as cover and which to ignore during data collection. Cover dimensions will probably be assigned arbitrarily by the field crew collecting data for PHABSIM, if this information is not forthcoming in the criteria.

Examples of some common cover dimensions are shown in Figure 11. The important dimensions for an instream velocity shelter, such as a boulder, are the perpendicular component of the object with respect to the current and its height above the streambed. A long, narrow object, such as a log, has no effective length if lying parallel to the flow. The important component in this case would be the diameter of the log. Other important dimensions include: the height and width of undercut banks, the distance from the streambed to the undersides of root wads, the aggregate width and height of clusters of small instream objects, and the distance from the water surface to the undersides of overhanging vegetation.

e. Optional variables. The variables discussed previously (items a through d) should be considered essential to any microhabitat utilization or preference study. The measurement of additional variables, such as water temperature, dissolved oxygen, alkalinity, visibility, and conductivity, may also be useful in an IFIM analysis, but are most important in evaluating the transferability and possible biases of a set of criteria.

Temperature is easy to measure and is among the most valuable of the optional variables. If a criteria study is conducted over several seasons, temperature data may provide important clues regarding spawning periodicity,

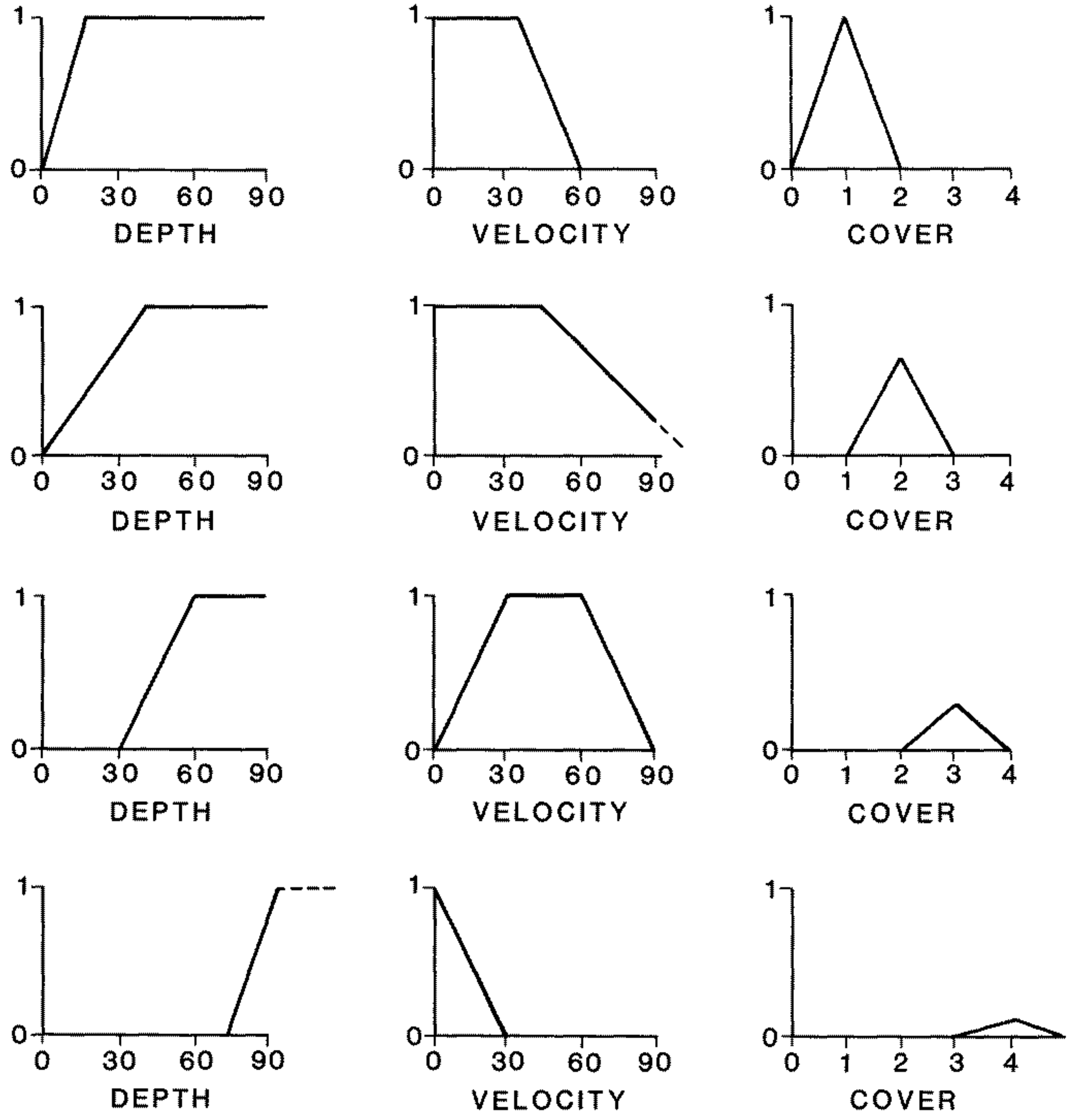


Figure 10. Example of weighted conditional criteria, combining conditional depth and velocity with behavioral selection of various types of cover.

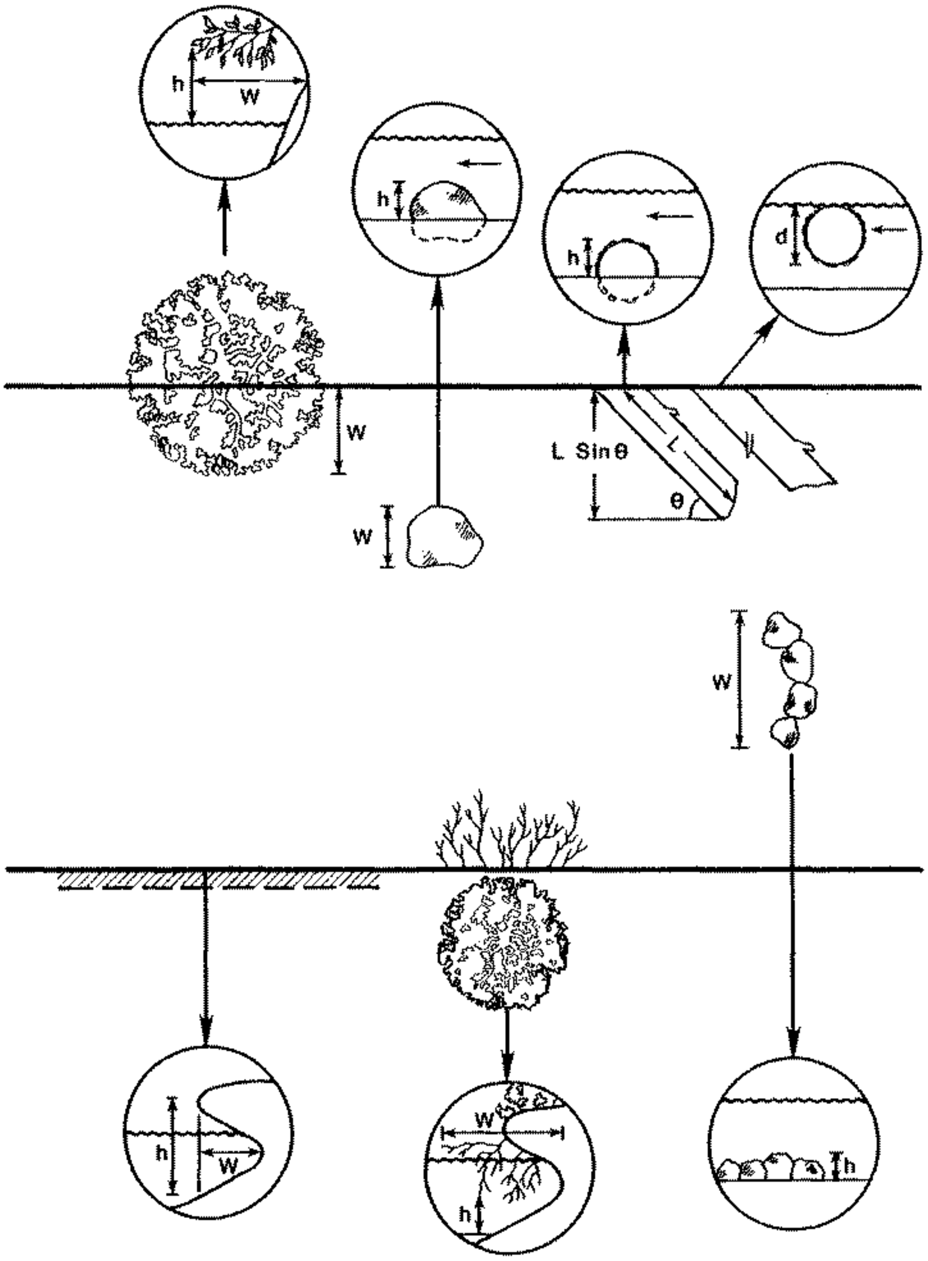


Figure 11. Examples of important dimensions associated with common types of instream and bankside cover.

incubation time, and seasonal or diurnal changes in behavior as reflected in microhabitat utilization. It may also be possible to ascertain tolerable or preferred temperature ranges for a species if the study sites exhibit longitudinal or lateral thermal gradations during the sampling period.

The measurement of "traditional" water quality variables, such as dissolved oxygen, is of limited value in a microhabitat criteria study. Water quality characteristics should seldom become unfavorable in a study site, if the study site has been selected properly. Measurements of certain water quality constituents may be warranted at some sites to ensure that water quality is not influencing microhabitat utilization. It may be possible to develop interactive microhabitat and water quality criteria concurrently, because changes in water quality might be accompanied by changes in microhabitat utilization. If this approach is taken, category III microhabitat criteria should be developed, because longitudinal changes in water quality often parallel changes in microhabitat availability, which can affect the utilization function. Although this approach has some intuitive appeal, it would require a very large sample size. It may be easier simply to categorize water quality as being completely acceptable or marginal, and then develop conditional criteria for the two categories.

Alkalinity, or some other index of primary productivity, is a useful variable to measure during microhabitat criteria studies. It seems logical, at least conceptually, that habitat utilization by fish might be influenced by the availability of food in a stream. Preferred microhabitat sites for many species are actually highly efficient feeding stations. As food supply increases, the need for the most efficient feeding stations may decrease, so fish in a highly productive stream may be found in more marginal microhabitats than those in less productive streams. An index of the productivity would be helpful in comparing the criteria source stream with other streams to determine similarities in the food supply. As described for water quality, it may also be desirable to categorize streams as productive, moderately productive, and unproductive for the development of conditional criteria.

A measurement of conductivity may be necessary in studies using electrofishing as the primary collection method. This information is used to calculate the size and strength of the electric field around the probe. Field size and current density should be determined so the investigator knows the area sampled when the probe is activated. This is especially important in water with low visibility where the fish's original position cannot be determined precisely.

One or two columns in the field book should be reserved for comments. Dr. Paul Turner (Department of Fisheries, New Mexico State University, Las Cruces; pers. comm.), suggests that a numerical rating for the quality of each observation should be recorded under this category. The idea is that "in-situ" fish observations are better than those where the fish are not seen until they are brought to the surface. If a "1.0" means that a fish was collected where observed and a "5.0" means that it was sampled from an unknown location within the general area, an average of all the ratings gives a good overall evaluation of the quality of the observations. The comments column should also cross-reference frame and roll numbers for any photographs taken at fish locations.

This is absolutely essential if photographic analysis will be used to determine substrate composition or fish positions.

Figure 12 is a sample data form showing the types of data that might be collected in a criteria study. Items marked by an asterisk are considered mandatory, including:

1. the name of the stream,
2. a site identifier (cross-reference to a topographic map),
3. the discharge (cfs) on the date of sampling,
4. the date,
5. the time or time period of sampling,
6. members of the data collection crew,
7. method of sampling or observation,

12. species,
13. size (total length in mm),
14. activity code (if activity can be observed),
15. frequency (number of fish observed/collected at each location or catch per unit effort),
16. total depth (feet or cm),

18. mean column velocity (feet or cm per second), and

21-25. substrate or cover codes.

The choice to determine substrate composition, cover type, or both is optional. Substrate should be considered a required element in studies of macroinvertebrates or spawning salmonids. Studies of adult and juvenile fish should include cover as a required element and substrate optional. Other optional data include:

8. Secchi disk visibility (ft or cm),
9. conductivity (µmho/cm),
10. alkalinity or calcium hardness (mg/l),
11. other water quality constituents (mg/l),

MICROHABITAT UTILIZATION DATA

(1) *Stream: ______________ (4) * Date: ______________ (7) * Method/Gear: ______________ (10) Alkalinity (mg/)

(2) *Site ID: ______________ (5) * Time: ______________ (8) Visibility (ft) (11) Other W.Q. ______________

(3) *Discharge: ______________ (6) * Crew: ______________ (Secchi disk or) ______________ ______________

Conductivity (mmho) ______________ ______________

(12) * Species	(13) * Size -TL- (mm)	(14) * Activity (code)	(15) * Freq. (#obs) or CPUE)	(16) Depth Total* (ft)	(17) To Fish (ft)	(18) Velocity Mean* (fps)	(19) Nose (fps)	(20) Adjacent (fps)	(21) Substrate Dom. (code)	(22) Sub. (code)	(23) Embed. (%)	(24) Cover Type/ Form code	Func. (code)	(25) Dim. (ft)	(26) Temp. (°C)

*Required element

Either is optional, but one is required.

Figure 12. Sample field data form showing required (*) and optional measurements and information recorded for a microhabitat suitability study.

17. fish depth (in feet or cm above streambed),

19. nose velocity (feet or cm per second),

20. adjacent velocity (feet or cm per second), and

26. temperature (°C or °F).

In addition, a site description including stream width, gradient, elevation, pool:riffle ratio, sinuosity, and the range of depths, velocities, substrates, and cover types within the reach are invaluable aids in evaluating transferability of criteria. These should be mandatory accouterments with category II criteria. They are optional, but welcome, additions to category III criteria.

2.5 SELECTING STUDY AREAS

Most category I criteria are not based on actual field data, so study areas are usually not needed for their development. Selecting the appropriate setting for data collection, however, can be one of the most critical aspects in developing category II and category III criteria. As mentioned in Section 1.1.3, category II criteria are based on observations of fish or macroinvertebrates under the environmental conditions that were available to them at the time of observation. If great care is not exercised in the selection of the time and place of observation, the resulting criteria may only represent tolerance of marginal habitat, rather than true elective behavior on the part of the organism. Category III criteria are designed to factor out this environmental bias, but habitat availability can also create problems in the development of these criteria.

The easiest way to envision the ideal situation for determining habitat preferences is to consider a stream in which all conceivable combinations of depth, velocity, substrate, and cover are equally distributed. Fish observed in such a stream would reflect the preferred microhabitats, as well as those the species avoids, because the fish would have free and equal access to all other microhabitat combinations. In this situation, category II data would be equivalent to category III, because the term for habitat availability would be a constant.

Unfortunately, streams do not have equal distributions of all combinations of microhabitats. In fact, most streams do not even have all microhabitat combinations available, regardless of their distribution. Given these facts what characteristics are desirable in a "source" stream for habitat criteria, and which types of streams should be avoided? The most important feature of a good source stream is habitat diversity. Other important considerations include stream dimensions, water quality, and characteristics of the biological community. These factors may be irrelevant, however, (or, at least, secondary) if the stream does not display a wide variety of habitat types.

Habitat diversity is often considered only on a spatial scale, but it also has a temporal component. Channel structure plays an important role in determining the distribution of microhabitat features in a stream, but so does discharge. It is fairly easy to distinguish between channels having simple and complex structures, but more difficult to distinguish discharges having simpler or more complex hydraulic patterns in the same channel. At the risk of overgeneralizing, it seems logical that the widest range of conditions would occur at a moderate flow level, for example, at the mean annual flow. Habitat diversity is probably lower at extreme high or low flows in most channels. The important concept, though, is the selection of stream reaches and discharge levels that provide a variety of conditions: deep-slow, deep-fast, shallow-slow, and shallow-fast, with various combinations of substrate and cover. The closer these conditions approach equal proportions, the better.

Stream size is also important in setting the dimensions of the criteria. Criteria developed in a small stream with a maximum depth of 90 cm may not be directly applicable in a large stream with a maximum depth of 900 cm. A complicating factor in transferring criteria from small streams to large ones is that the mean column velocity in a small stream is about the same as the nose velocity of the fish. This is usually not the case in a large stream. The most important dimensional consideration in selecting a stream, however, is to make sure that at least some of the available habitat is within the preferred range of the target species. Otherwise, the peak of the utilization function will not reflect the optimum condition. Even worse, the optimum will be estimated by extrapolation if a preference function is developed. This can lead to a larger error than using the utilization function alone (Bovee 1982). Furthermore, the utilization and preference functions may look totally different, leaving the investigator the unenviable task of deciding which one is correct. (In this case, it is very likely that both are wrong.)

It is also desirable for the source stream to contain some conditions outside the estimated tolerance range of the target species. This is important because a criteria curve, whether utilization or preference, intersects zero when fish are no longer found over an interval of a variable. If fish are present over all intervals, the curve never goes to zero and the end points of the distribution are undefined or must be estimated. The consequence of these requisites is that the investigator needs some a priori knowledge about the habitat characteristics selected by the target species. This may require the construction of interim criteria by one of the category I methods discussed in Chapter 3.

The influence of temperature and water quality on habitat selection is another characteristic of the source stream that should be considered. Macrohabitat is analyzed in parallel to microhabitat in the IFIM, an approach that assumes suitable water quality and temperature at the time the microhabitat criteria were generated. Evidence suggests that fish (and presumably macroinvertebrates) will abandon otherwise suitable microhabitats under marginal water quality or temperature conditions. Elevated water temperatures often drive fish into deeper, cooler water (Reynolds 1983), whereas low dissolved oxygen levels may force a retreat into shallower water, backwaters, or mouths of tributaries (Thompson 1925; Simpson 1978). It also seems likely that fish would move into slower water to reduce their metabolic requirements, under conditions of high temperature or low dissolved oxygen. Microhabitat criteria, conditional on water quality and temperature, could be incorporated into an IFIM analysis, but it would be difficult and might not gain much in terms of accuracy. Therefore, water quality and temperature in the source stream should be well within the acceptable range of the target species to avoid these complications.

A final consideration in the selection of a source stream is the relative abundance of the target species and its competitors. At very low densities, locating enough specimens to conduct a meaningful frequency analysis becomes difficult and time consuming. Furthermore, the fish will not be under the competitive pressure forcing them to use suitable, but suboptimal habitat areas. Consequently, the preferred habitat may be well defined, but the end points defined poorly. Conversely, very high densities (i.e., above saturation) may result in unusual behavior of both the fish and the data collector. Fish may be forced into areas that are only tolerable for brief periods, whereas the data collector may be overwhelmed by the abundance of fish, and experience great difficulty in identifying or enumerating members of the target species. The presence or absence of closely related competitors can also change the frequency distribution of microhabitat usage. Competitive release often occurs in allopatric populations, so fish inhabit a wider range of conditions than in sympatric populations. Whether the source stream contains an allopatric or sympatric population of the target species has little bearing on how the criteria are developed. It may have a large influence, however, on where the criteria can be applied. Therefore, the ideal source stream should be near saturation of the target species and have roughly the same species composition as the streams to which the criteria are to be applied.

2.6 SELECTING A SAMPLING STRATEGY

Some of the largest sources of bias in habitat utilization or preference criteria can be traced to poor (or no) sampling strategy. One distinction between criteria studies and other fisheries studies is the emphasis on the quality of observations, rather than the quantity. Investigators who emphasize quantity tend to look for fish where they expect to find them. Consequently, the resultant criteria become self-fulfilling prophecies. Disproportionate sampling effort is one of the most serious problems encountered in criteria studies, because it is a form of bias that may be totally undetectable (even to the person who collected the data).

Johnson and Nielsen (1983) list four basic sampling strategies: (1) simple random, (2) stratified random, (3) clustered, and (4) systematic. Bain et al. (1982) have developed a fifth sampling strategy called proportional sampling. The best sampling strategy for a criteria study often depends on the organism under study and the observation technique used to study it.

2.6.1 Random Sampling

Simple random sampling is one of the most basic and least biased of all sampling designs. This approach requires only a moderate amount of preparation before sampling. A plan map of the study area, drawn to scale, with a numbered grid system superimposed, is the basic tool for random sampling (Figure 13). Grid numbers are randomly selected from numbered slips of paper or from a random number table. A variation of this technique is a double random sample, in which the grids are also numbered randomly, as in Figure 14. The size of the grids may be dependent on the area sampled by the collection device.

2.6.2 Stratified Random Sampling

In order to increase the precision of random sampling and to ensure that certain habitat types will be sampled, the population of potential sampling sites can be divided into homogeneous groups. Each homogeneous group can then be gridded and randomly sampled. This process is called stratified random sampling. The advantage of stratified random sampling is that differences among sampling locations or times are recognized and accounted for in the sampling procedure.

Stratified random sampling requires slightly more preparation than simple random sampling. One obvious difference is that the plan map of the study area requires considerably more detail because discrete habitat types must be delineated on the map. Examples of habitat strata include pools, riffles, runs, different substrate types, and areas with and without cover.

Unless caution is exercised, stratified random sampling can result in the inadvertent biasing of the data. For example, suppose a stream is stratified simply into riffles and pools, but 75% of the stream area is riffle and only 25% pool. If both strata are sampled with the same effort, this approach would give too much weight to the pool samples and not enough to the riffle samples. One way to avoid this type of bias is to use proportional sampling.

2.6.3 Proportional Sampling

Bain et al. (1982) proposed the use of a proportional sampling design in conjunction with use of a grid shocker (see Section 4.1.4). Proportional sampling employs a very detailed habitat map, similar to the type used in the Physical Habitat Simulation (PHABSIM) system (Trihey and Wegner 1981; Bovee 1982). Transects are established to describe the longitudinal variation in habitat, while verticals across each transect define the lateral variation. The cells defined by the transects and verticals are, at least theoretically, completely homogeneous. The cells will be small where habitat characteristics change rapidly, but large where the habitat is gradually varied or homogeneous. Since each cell has a surface area, the ratio between cell area and total area provides an index to the amount of sampling effort expended in each habitat (cell) type. Site preparation and sampling effort preparations are illustrated in Figure 15.

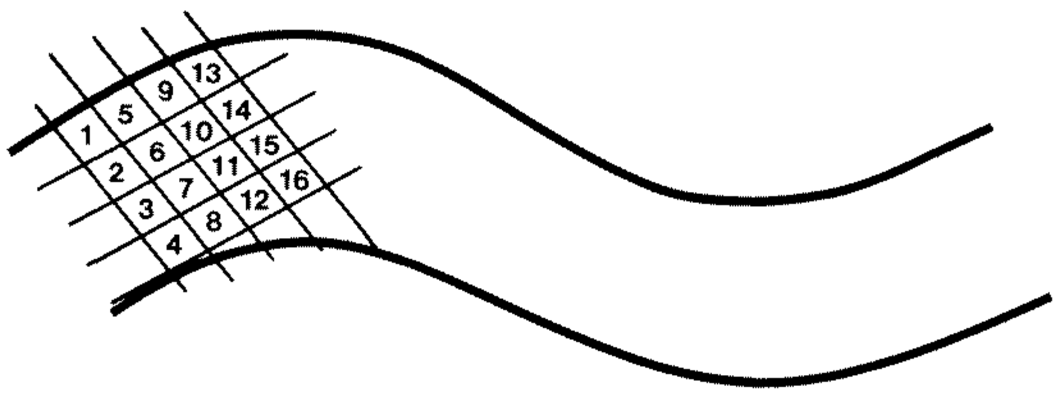


Figure 13. Gridded plan map of a river used to select sampling locations from a simple random sampling design.

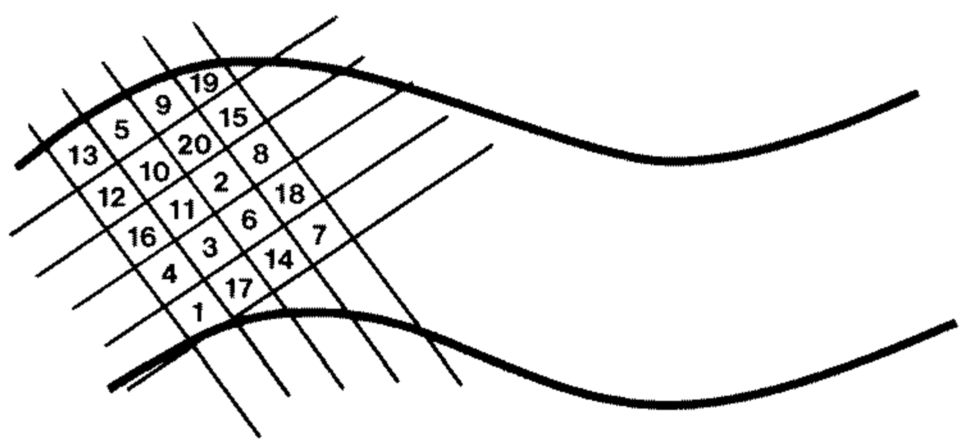


Figure 14. Gridded plan map of a river used to select sampling locations from a double random sampling design.

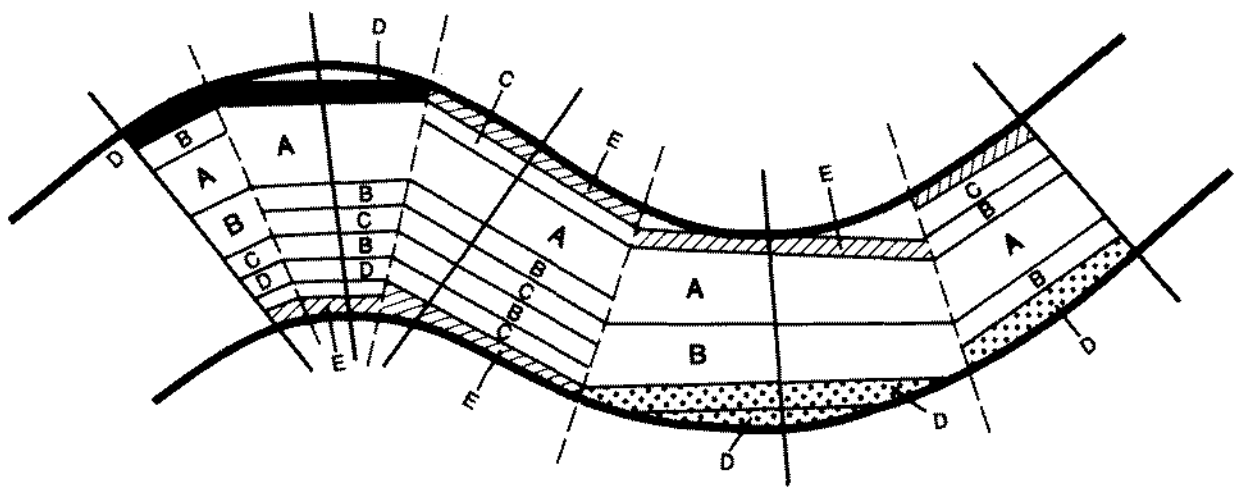


Figure 15. Site preparation for proportional sampling.

Each cell in Figure 15 has a discrete combination of microhabitat variables and, for simplicity, have been distinguished by the letters A through E. The total surface area of the reach is 3,300 m^2. Type A habitat represents about 45% of the total, whereas Type E represents only 6%. If 200 samples were to be collected in this reach, 90 of them should be collected in the cells designated by A and 12 in cells marked E.

2.6.4 Modified Cluster Sampling

The modified cluster sampling design uses a much simpler stratification procedure than described above for proportional sampling. The stratification elements suggested for stratified random sampling are probably sufficient, but the proportion of the study area represented by each stratum would be determined in the same manner as proportional sampling. Each stratum is gridded and randomly sampled to determine sampling locations, but samples are weighted. The number of samples taken from each stratum is proportional to the abundance of that stratum in the study area. Using the example of a stream with a 75/25 riffle/pool ratio, for every 100 samples taken, 75 would be randomly selected from the riffle stratum. This approach may also be compatible with field techniques that do not involve the sampling of a discrete area. For example, if the data are to be collected by a diver, 75% of the observation time should be spent in the riffles and 25% in the pools.

2.6.5 Systematic Sampling

Systematic or uniform sampling is probably not applicable to criteria studies because it is based on a uniform spacing of sampling locations in one dimension, usually along a riverbank or beach. Microhabitat sampling is distinctly two-dimensional because of the lateral variability across rivers. Modifications of systematic sampling, however, may be applicable to this type of study.

Blanket sampling is a systematic technique applicable in those studies using divers to collect the data on fish observations. Enough divers are deployed that the entire width of stream is covered. Each diver is responsible for enumerating and locating fish within a uniform distance (usually an arm's length) from the next diver. The entire row of divers then drifts through the reach in a line, essentially covering the total surface area of the study site in one sweep. Obviously, this technique will only be practical in small streams with good visibility.

The systematic random walk is a more generally applicable technique, with many of the desirable attributes of simple random sampling. As its name suggests, the distance between successive sampling locations is uniform, but the bearing from one location to the next is selected at random, as illustrated in Figure 16. One problem that can arise with the random walk method is that potential sampling locations may be located on dry ground or behind a previous location. This problem can be avoided by limiting the sampling of potential bearings to a range of ±90° from the current location, which will confine the next location to the river. This technique is probably most applicable in large rivers where there is ample room to work, and with unidirectional sampling gear, such as drift boats.

Bain (1985) developed a stratified-systematic approach that, in his opinion, was more efficient than random sampling, but less subjective than proportional sampling. Representative transects were chosen to include major types of habitat. Then, sampling locations were systematically chosen on each transect: one adjacent to each bank, one at one-third the distance across the stream, and one at two-thirds the distance. This approach should give unbiased results as long as the areas represented by each transect are about the same. When certain habitat types are more extensive than others, however, it is advisable to use more transects in the more abundant habitat types. Otherwise, a proportionately larger number of samples would be taken in the less extensive habitats. Of course, if no fish are found there, the bias would be negligible.

2.6.6 Sampling Design Considerations

Selection of the sampling design for a criteria study cannot be made solely on the theoretical properties of one design versus another. Other equally important considerations include the amount of site preparation needed, efficiency of the design, and compatability with sampling techniques. Regardless of the sampling design selected, sampling strategy should be determined before the actual data collection begins. Stealth is crucial in making good observations for habitat criteria and one factor that can contribute to unobtrusive sampling is the preselection of sites. The crew

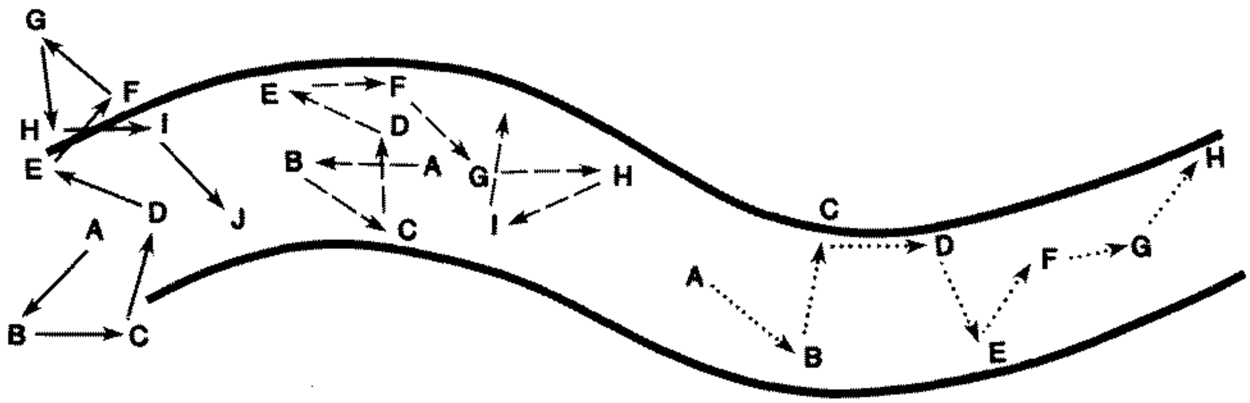


Figure 16. Examples of systematic random walk sampling designs: unlimited random bearings (solid line); bearings limited to river boundaries (dashed line); bearings limited to river boundaries with downstream orientation (dotted line).

should be provided with a good map of the study site, with the sampling locations already marked. Then, it is up to the crew to determine the best way to approach the locations undetected.

Preselection of sampling locations means that some site preparation is needed for virtually all sampling designs. Even an approach as simple as blanket sampling requires knowledge about the width, depth, and morphometry of the stream so the required number of divers and spacing between them can be determined. Some sampling designs require more site preparation than others; the proportional sampling technique suggested by Bain et al. (1982) requires the most. If an intensively surveyed study site can serve double duty, either for determining habitat availability or as an analytical site for an instream flow study, then a large investment in site preparation might be warranted. Stratified random and modified cluster sampling both require a moderate amount of site preparation, including mapping the area to scale, delineating strata, determining the area of each stratum, and randomly sampling locations. Simple random sampling requires only the preparation of a plan map, but translating sampling locations from a map to a place in the river can be difficult, especially if the map is sketchy. The systematic random walk is probably the least intensive sampling design, requiring only a plan map and a means of picking random bearings (a spinner from a child's game or a small "wheel of fortune" will often suffice).

The efficiency of the sample design is another important consideration. The biggest drawback to any form of purely random sampling is that many randomly sampled sites in a river will not contain any fish. An example of this problem is presented by Bain et al. (1982) from a study conducted in a small Massachusettts stream to develop preference criteria for smallmouth bass. Of the 214 sites that were selected for sampling, only 39 contained smallmouth bass. Furthermore, the 39 sites that did contain smallmouth bass included both juveniles and adults. The combined efficiency in the Deerfield River for all smallmouth bass was 18% and would have been lower for juveniles or adults alone. At this rate, 850 locations would need to be sampled to assemble a typical data base of 200 observations for each life stage. This means that the total number of samples would be around 1,700 for the development of separate data bases for juveniles and adults. Bain (pers. comm.) suggests that this is probably an extreme case, but agrees that efficiency is a major concern with random sampling methods.

Sampling efficiency improves if data for several species are collected concurrently. From the 214 sampling locations sampled by Bain et al. (1982), 15 longnose and 13 blacknose dace were taken, along with 261 fallfish, 32 American eels, and a scattering of darters and suckers. Random or proportional sampling designs are inefficient with respect to a single species, but the efficiency increases dramatically when criteria data are simultaneously collected for numerous species. Efficiency of random or proportional sampling is also influenced by the organism being sampled. Aquatic macroinvertebrates, for example, tend to be much more ubiquitous in a stream than fish. Efficiency for this group would be fairly high. Conversely, a highly mobile top predator, such as a walleye, might be nearly impossible to find using a random sampling approach.

A third consideration in choosing a sampling design is its compatability with the proposed sampling technique. For example, electrofishing can be very disruptive to fishes in a localized area. Simple or stratified random sampling often results in two or more adjacent sampling locations. Electrofishing in one location may disturb fish in an adjacent location, thereby creating a potential data bias. The systematic random walk, especially if unidirectional, would not create this problem because sampling locations would be far enough apart that one sample would not interfere with another. The random walk pattern, on the other hand, would be fairly incompatible with SCUBA observations. The diver would spend more time swimming from site to site than observing fish. Modified cluster sampling would be better, because the diver would spend a proportional amount of time in each area, locating and enumerating as many fish as possible. Table 5 gives some broad guidelines to the applicability of different sampling designs, depending on the circumstances of a particular study.

Table 5. Considerations for selecting sampling designs depending on physical, biological, and logistical characteristics.

Study characteristics	Sampling design: Simple random	Stratified random	Proportional	Cluster	Random walk
Site characteristics					
Large river	/	/	X	/	X
Small river	X	X	X	X	/
Species characteristics					
Multiple species	X	X	X	0	X
Few species	0	/	/	X	/
Ubiquitous	X	X	X	/	X
Rare	0	0	/	X	/
Sedentary	X	X	X	X	X
Active	0	0	/	X	/
Sampling technique					
Surface observation	0	0	X	X	0
Subsurface observation	0	0	X	X	0
Electrofishing	X	X	X	0	X
Area Sampling	X	X	/	/	X
Biotelemetry	0	0	X[1/]	0	0

X - mostly applies
/ - somewhat applies
0 - probably not applicable
1/ - proportional or systematic sampling of frequencies.

2.7 ESTIMATING SAMPLE REQUIREMENTS

It is desirable to know how many observations or collections are needed to obtain a convergent frequency distribution, at the outset of a study. Convergence means that the frequency distribution would not change significantly if a larger data base were used. The number of observations for each life stage and species is a nearly universal specification in statements of work when criteria data are collected by a contractor.

There is a distinction between a sample, an observation, and a frequency in the context of a criteria study. A sample is an attempt to capture or observe fish or aquatic macroinvertebrates, whereas an observation refers only to samples containing one or more target organisms. The frequency refers to the number of target organisms captured or observed under each "observation." Many samples have a frequency of zero, but one observation might have a frequency of 1, 10, or 100.

The reason that this distinction is important is that experience indicates that 150-200 observations are typically needed to construct a reasonably smooth frequency histogram. This does not mean 150-200 samples, nor does it mean a total of 150-200 fish. It means 150-200 data points representing places where one or more fish were captured or observed. This number should be considered to be a minimum data requirement, although a larger number of observations may be needed.

There are numerous statistical equations that can be used to estimate "sample" sizes (assuming each sample counts as an observation). An example of this type of equation is given in Platts et al. (1983):

$$n = \frac{t^2 s^2}{E^2} \tag{2}$$

where n = the required sample size,

t = Student's t for an assumed sample size, n at a specified probability level,

s^2 = the sample variance, and

E = the acceptable error between the sample mean and the population mean.

Other sample size estimators are similar in form to equation 2. These equations, in general, are designed to provide a sample size large enough to ensure that the sample mean is within a specified percentage error of the population mean. If sample sizes are estimated using this approach, it is important to base the estimate on the variable with the largest variance over the frequency distribution. This variable is usually, but not always, depth.

Armour et al. (1983) illustrate the use of another sample size estimator that can be used without an estimate of variance:

$$n = 4/\left(\frac{cv}{\delta}\right)^2 \tag{3}$$

where cv = the coefficient of variation, and

δ = the relative half width of the confidence interval about the true mean.

Armour et al. (1983) state that for planning purposes, the coefficient of variation can be assumed to be 100% (1.0) and a "good" estimate of the population mean can be obtained if the relative half width, $\delta = 0.1$. This reduces equation 3 to:

$$n = 4/(0.1)^2 = 400 \tag{4}$$

The final sample size requirement ultimately depends on the variance of the frequency distribution over the range of utilized conditions. This, in turn, is influenced both by the range of conditions available and the subset of conditions used by the organism. Equation 2 can be used to determine if more than 400 observations are needed, but if equation 2 indicates a sample size of less than 150, the investigator should re-examine the suitability of the study stream. It may not contain a wide enough variety and range of conditions for the development of transferable criteria. Fewer than 150 observations are probably acceptable if the stream appears to have a wide variety of conditions and the organism is utilizing a very narrow subset of those conditions.

2.8 DISCUSSION

It is no accident that this chapter is one of the longest in this manuscript. The limitations of existing habitat suitability criteria are nearly universally linked to an oversight or misjudgement in the original data collection. Many of these limitations could have been circumvented if the investigators had spent more time designing the studies. In fairness to previous researchers, it must be stated that most of these studies were intended for site-specific analyses. Furthermore, some of the limitations of the criteria could not have been foreseen as recently as five years ago. Such limitations became obvious only with increased knowledge about fish behavior and habitat selection. Most of the study plan components suggested in this chapter are direct results of the collective experiences of previous research efforts.

If there is one common misconception by developers of habitat suitability criteria, it is that their criteria will not be used in any other study. It may be fair to say that many researchers report the results of habitat utilization studies in the context of academic enlightenment, and are chagrined to find the fruits of their efforts sullied in the decision-making arena. In reality, there is a strong tendency for IFIM users to rely on criteria developed by someone else. The reason is that most operational applications of IFIM have such limited budgets and deadlines that the user cannot afford to develop site-specific criteria. Given that habitat suitability criteria will be transferred from stream to stream, the research community has certain obligations to the user community. Foremost among these is to develop criteria that are comprehensive, accurate, and precise.

Comprehensiveness is primarily related to data stratification and sampling protocol. Data stratification refers to the subdivision of the criteria according to factors that may affect habitat selection by a species. Typical strata include size, activity, time of day, cover availability, and season. Sampling protocol establishes the variables to be measured and how they will be described. There is an apparent link between comprehensiveness and precision. By definition, if a functional relationship is theoretically robust, it should be possible for two researchers, measuring the same thing to arrive at the same result. The fact that some curve sets do not converge can often be traced to differences in data stratification and sampling protocol. Reproducibility should not be expected if researchers are not measuring the same thing, the same way.

The primary factors influencing accuracy are sample size, characteristics of the study stream, sampling design, and limitations of sampling gear. A large enough sample size is needed to ensure that the sample mean is close to the population mean. The bigger question, however, is whether the population has been sampled in an unbiased manner. The principal concern of anyone developing criteria is to eliminate sampling bias as much as possible. While most investigators are very conscientious about bias introduced by gear limitations (discussed in Chapter 4), the potential biases due to sampling location and design are often overlooked. The causes of these bias problems are either limitations of habitat availability (site limitations) or disproportionate sampling effort (sampling design). These sources of bias were discussed briefly in Sections 2.5 and 2.6, in the context of data collected at one site. The problem becomes more serious and difficult to deal with when data are pooled from several sources.

The study stream and sampling design can also affect the precision of the criteria. The profound influence of the study stream on category II criteria was discussed in Section 2.5. There, it was suggested that one of the most desirable characteristics of a "source" stream is a high degree of habitat diversity. The goal of developing category III (preference) criteria is to increase precision by eliminating or reducing the bias of habitat availability. Curiously, however, data pooling problems become more severe when developing category III criteria, and can potentially reduce precision rather than increase it. Data pooling biases and possible solutions are discussed in Chapter 4.

3. DEVELOPMENT OF CATEGORY I CRITERIA

Many scientists question the value of category I criteria because they are not based on field data. Most people feel more confident using category II or III criteria, but category I criteria must be used in many instream flow studies for the foreseeable future. Data intensive criteria have not been developed for very many species, to date, and data gaps are common for certain life stages or physical variables within criteria that are available. The ultimate goal of the research community should be the development of category III criteria, but in the interim, category I criteria must be relied on to fill the voids. Decisions regarding water management will proceed regardless of the quality of the available biological information, and may be made with no input from the biological community, at all. In view of this reality, category I criteria are vastly superior to no criteria.

The two most common sources of information for category I criteria are the professional literature and the collective judgements of experts. While it might seem easy to construct habitat suitability criteria by conducting a literature search or by convening a panel of experts, both methods can be frustrating and difficult. The greatest obstacles to the development of category I criteria are insufficient information and poor communication. The following sections illustrate techniques to minimize these problems. Unfortunately, they can not be totally eliminated.

3.1 LITERATURE SOURCES

Most of the habitat related information found in the literature is contained in life history and distribution and abundance studies that were conducted from the 1930's to the 1950's. Studies, such as Thompson and Hunt's (1930) habitat typing of warmwater fishes in Illinois, or Minckley's (1963) ecological study of Doe Run in Kentucky, can provide valuable insights into the types of habitats utilized by different species and life stages. These studies come tantalizingly close to providing enough information on which to formulate habitat suitability criteria. Bovee and Cochnauer (1977) devised a variety of techniques to transform anecdotal descriptions of habitat usage into somewhat quantifiable terms. Subsequent experience has shown, however, that these methods are only marginally useful in the development of criteria. The basic problem is that life history studies (at least the older ones) lack the kind of quantification necessary to formulate habitat suitability curves. For example, current velocity has proven to be a dominant variable in microhabitat selection for nearly every river-dwelling organism, but measurements of velocity are nearly universally absent from most life history studies.

Another problem with this type of study is that quantified descriptions are often sufficiently vague as to be meaningless. A description of adult channel catfish being most commonly found in pools with depths "ranging from two to 20 feet" does little to determine the preferred depth of the animal. This discussion is not intended as a criticism of these studies. Rather, the problem is one of relying too much on the judgement and interpretation of the person attempting to use such information to develop criteria.

Actual criteria development studies are much more valuable sources of information than life history studies. Unfortunately, most studies of this type are found only in the "gray" literature. Typical examples include Master's theses, State fisheries investigations (Dingall-Johnson or "DJ" reports), documentation for operating licenses from the Federal Energy Regulatory Commission (FERC studies), and environmental impact studies involving instream flow issues. Some States and Canadian Provinces (notably Washington, California, and Alberta) are actively developing criteria for use within their areas of jurisdiction.

The emerging interest in developing habitat suitability data bases is a welcome indication of a commitment to improving the accuracy and confidence in instream flow studies. It will, however, place greater responsibility on the user to evaluate, verify, and/or modify available criteria. When criteria are modified, even if only slightly, they become category I criteria regardless of their original classification. One of the most important, and often, most difficult decisions a user must make is whether to use criteria directly from the literature or to adapt the criteria to local conditions. In either case, professional judgement is required. For this reason, it is as important to the user to understand how criteria are developed as it is to the researcher. Although the development of new criteria is accelerating, one problem remains that may force a user to develop or modify criteria: lack of standardization. As explained in Chapter 2, there are literally dozens of ways to stratify and collect data. The only way a single criteria study could satisfy all potential users is by stratifying to the ultimate level of detail and by measuring every possible variable. Even this level of detail might not guarantee criteria that would exactly meet the needs of any particular user. Therefore, it is almost inevitable that the second type of category I development, through collective expert judgement, will be employed sometime during an instream flow study.

3.2 PROFESSIONAL JUDGEMENT

The development or modification of habitat criteria by a group of fishery scientists is often a wise substitute or supplement to the use of literature-based criteria. The convening of such a group may be standard practice when an instream flow study is highly controversial, even if the criteria are derived empirically. The benefit of criteria-by-committee is that if consensus can be achieved, the derivation and validity of the criteria should not arise as issues in subsequent reviews and interpretations of the results. The disadvantage is that the principal science involved in the derivation is frequently psychology, not biology. Obtaining a consensus is time-consuming and may be impossible. Participation in such a committee may be an imposition to some of the members and as the size of the committee grows, so do scheduling problems. Despite these problems, it is wise to attempt to obtain consensus before using criteria of any kind in an instream flow study.

Group development or modification of criteria can take several forms. The most informal of these are face-to-face discussions of the various participants in the group, termed roundtable discussions. A more formal approach, called the Delphi technique, is conducted such that the participants remain anonymous to one another. A third technique involves data collection, but substitutes professional opinion for fish observations. This method is termed "pattern recognition" because it is based on the idea that professionals can recognize usable and unusable habitat, although they might not be able to quantify it.

3.2.1 Roundtable Discussions

A roundtable is the least structured and most informal of the professional committee approaches. The success of such a group interaction depends on the composition of the committee and on the ability of the moderator to guide the group towards the intended objective. Failures can often be traced to an ineffective or nonexistent moderator. Someone must have the authority and the responsibility to be the group's leader and spokesperson, but must not be so heavy handed that group interactions are stifled.

Scheele (1975) states that a mix of three types of participants is desirable in any type of professional committee. First are the "stakeholders," those who are or will be directly affected by the outcome of the criteria development effort. In face-to-face discussions it is important to ensure that no single agency or interest group is overrepresented by stakeholders. Many can empathize with the lonely feeling of being the sole representative for an agency or group, facing a whole cadre of opposition across the table. The group chairman or organizer should be careful that each group has the same number of spokespersons (preferably one), no matter how many people show up for the meeting. The idea is to encourage a diversity of opinions, not to outvote the other side.

The second group of people that should be included in such a committee are the "experts," who have applicable specialties or relevant experiences. People associated with instream flow studies are often stakeholders as well as experts. The difference is that experts, in Scheele's context, are further removed from the operational aspects of a study and more objective than stakeholders. Experts are also much harder to motivate than stakeholders. From their viewpoint, engaging in "abstract speculation" with people they do not know, about a subject not central to their interest, may be seen as a waste of their time. Scheele suggests that prestige is one of the strongest motivators for inducing quality participation among members of this group. If potential panelists tend to be disinterested, try to find a prestigious sponsor for the effort or try to staff the panel with a group of stimulating peers. Above all, reinforce the idea that the goals of the discussion are important and not simply an exercise in mental gymnastics. A token payment or honorarium may also stimulate interested responses. At the very least, travel expenses should be provided. A certificate of appreciation from a prestigious source (e.g., the governor's office, the Department of Interior, the president of the company) may be a more powerful incentive than money.

The third type of person Scheele recommends as a panelist is the "facilitator." These people are skilled at organizing, clarifying, and synthesizing information. They often act as mediators, a role that can be most important in face-to-face discussions. It may be necessary to hire a professional mediator. This can be a wise investment even (especially) if the mediator knows nothing about fish habitat.

Once a panel has been assembled, the second design consideration falls into the general classification of orienting and directing the flow of information. The responsibility of the chairperson or mediator is to prevent the introduction of tangential subjects and to keep the information flow focused on the objective. A foremost consideration is to make the committee interaction "context specific." Most inquiries of this type are stated in terms of general propositions, but individuals work within a context defined by their experiences. This can become an obstacle in the development of habitat suitability criteria. Individuals with a holistic philosophy may tend to view streams in the context that "as long as it's wet, it's habitat for something." Such a viewpoint might result in the understratification problems discussed in Chapter 2. Although all of these habitat strata may be equally important to a target species, they must be addressed one at a time. Several narrow sets of criteria imply that habitat types A, B, <u>and</u> C are needed, whereas one broad set implies that A, B, <u>or</u> C are needed. There is a large difference between these two meanings. Therefore, the general proposition needs to be defined succinctly enough that everyone knows exactly what it is. If resting criteria for adult smallmouth bass are desired, define what is meant by "resting" and "adult."

Criteria development by committee involves a sequential process. It is unlikely that consensus will be achieved on the first round. In fact, if this happens, the results should be viewed suspiciously because there may be too many like-minded individuals on the panel. The chairperson of a diverse panel must be able to orchestrate group interactions, documenting areas of divergence and consensus. Above all, the direction in which the group is moving must be expressed and compared against the goals of the exercise with regularity. The chairperson should recognize tangential discussions early and stipulate constraints focusing on the goal, when necessary.

There are positive and negative sides to any face-to-face group interaction. Among the positive aspects are such factors as rapid feedback, equal information flow to all participants, and a short response time in obtaining results. Among the disadvantages of committee meetings are such factors as scheduling problems, repetitive meetings, a tendency to discount minority or dissenting opinions, and the potential domination of the group by strong personalities. Because of these shortcomings, it may be necessary to advance beyond roundtable discussions and use one of the more anonymous group techniques.

3.2.2 Delphi

The Delphi technique, at least theoretically, overcomes many of the disadvantages of the traditional committee meeting. Delphi was first used in military strategic planning, and has been applied in areas such as transportation, health care, the environment, and fisheries management (Zuboy 1981). The primary characteristics of the Delphi are: (1) anonymity of the participants, (2) controlled feedback, and (3) an estimator of group consensus. The most common Delphi exercise uses a questionnaire, designed by a small monitor team and sent to a larger respondent group. The use of a questionnaire overcomes two of the major problems of a conventional committee meeting. First, the respondents can participate at their convenience, so specific times do not need to be scheduled for meetings. Second, the anonymous nature of the questionnaire prevents the "bandwagon effect" of a group dominated by a strong personality. After the questionnaire is returned to the monitor team, an estimator of group opinion is summarized. This usually consists of the median (50%) and interquartile (25% and 75%) ranges of the initial responses. The monitor team then provides this information to each respondent, who is asked to answer the questionnaire again, in light of the new information. If a respondent's second response is outside the interquartile range from the previous round, the respondent is asked to provide a brief explanation in support of this estimate. These explanations are then provided to all the respondents in the next round, along with the revised median and interquartile responses. This process is repeated until consensus has been achieved (Linstone and Turoff 1975; Zuboy 1981). Consensus does not necessarily mean that all respondents agree or provide the same answers. Stability of the distribution of the responses over successive rounds is a more significant measure for a stopping criterion than degree of convergence (Linstone and Turoff 1975).

Crance (1985) has documented a complete Delphi exercise he conducted to derive category I criteria for inland stocks of striped bass. This article makes interesting reading for anyone contemplating a Delphi exercise because Crance not only describes his own participation and preparation, but also summarizes the responses of the Delphi panelists during each iteration. Several important concepts regarding the conduct of a Delphi exercise emerge from the article. First, a problem with definitions became immediately apparent, even though Crance was fairly explicit at the outset. Some of the comments vividly demonstrate this problem:

1. "It was not clear what was meant by spawning runs."

2. "You should define larval and juvenile life stages since there is some variation in the use of these terms when dealing with striped bass."

3. "In assigning values to turbidity, I placed importance on total color, rather than just turbidity from sedimentation."

This illustrates a major difference between roundtable and Delphi exercises. Feedback is instantaneous with the former, but may take a month with the

Delphi. Thus, great care must be taken to be totally explicit with as many definitions as can be anticipated. Of course, it is probably impossible to anticipate all potential ambiguities arising from definitions.

A second interesting managerial concept that Crance dealt with was keeping the panelists' responses "context specific." Some of the comments from the first round were related to food availability:

1. "Foraging availability obviously does not ... fall into the categories of variables that you listed, but some assessment of food resources must be made to accurately evaluate habitat suitability for the species."

2. "Striped bass are pelagic nomads that travel great distances in search of food ... I feel the success of a striped bass year class is mostly dependent upon the availability of the proper size and type of plankton during the first week after hatching. Turbidity is important to adult striped bass because shad seek the warm turbid water. Adult striped bass follow the shad into this turbid zone."

While the goal of the exercise was to develop SI curves for the striped bass, these experts identified another habitat type important to striped bass that was not included in the original study, namely feeding habitat. An obvious solution, but one that would have increased the time and effort of both the panelists and moderators, would have been to expand the questionnaire to include habitat types for major forage species. This type of dilemma may occur in every Delphi exercise, and if not controlled, may cause the study to expand to the point where a final result is impossible. For example, parallel exercises would be needed to describe habitat requirements for plankton and all stages of shad to address the two comments cited. Perhaps the best solution would be to query the panelists regarding the most important food items of the striped bass, and then conduct follow-up Delphi exercises to construct suitability criteria for the forage organisms.

Linstone and Turoff (1975) suggest that Delphi forecasts are improved by using a "blank questionnaire" in the first round. For a criteria study, this would amount to a suitability graph with the axes labeled, but no curve on it. Crance (1985) provided several hypothetical curves for each variable during the first round. He later decided, however, that an intermediate form of a blank questionnaire is better than asking panelists to respond to a hypothetical curve. The reason is that a curve of any kind tends to predispose the panelists to a particular outcome, either biasing the result or creating an inertia that must be overcome before a final curve can be completed.

Experience has shown that the best way to start a Delphi inquiry is by asking panelists to respond to a series of short answer questions, such as:

1. What is the minimum depth used (if any)?

2. What is the maximum depth used (if any)?

3. What is the minimum depth considered to be optimal?

4. What is the maximum depth considered to be optimal?

Because the first round is so important in orienting and directing a Delphi panel, a sample first round packet has been included in Appendix A.

3.2.3 Habitat Recognition

The best qualified experts on a given species may not have the degree of expertise to directly develop habitat suitability curves. Most experts are sufficiently familiar with a species to distinguish which habitats it will or will not use, but if they have never measured those habitats, the experts are hard pressed to quantify habitat suitability. Habitat recognition is an intermediate technique useful in those situations where there are insufficient resources or time to collect fish observation data and insufficient expertise to assemble criteria by conventional roundtable or Delphi techniques. Habitat recognition involves moving the committee to the river. The principle behind this technique is simple. At specified locations in the river, each committee member is polled regarding the potential utilization of each location by the target species. The characteristics of the location are then measured. The results of numerous polls then constitute a frequency distribution of opinions rather than of fish. This frequency distribution is analyzed using the same techniques one would use for actual fish observations, discussed in Chapter 5.

Because habitat recognition is intermediate to category I and category II methods, it is important to observe the guidelines and cautions inherent with both. For example, the need to be specific about definitions is as important here as it is in designing a Delphi questionnaire. Anonymity is virtually impossible with this approach, but a secret ballot, similar to the one shown in Figure 17, can be used to minimize this problem. Each sampling location is denoted by a location identification number and the panelists have only two choices regarding the utilization of the sites by the target organism: yes or no. Each "yes" counts as a frequency of one and each "no" as a frequency of zero. Experience suggests that if "maybe" is a choice, it will nearly always be selected by everyone on the committee. The summed frequency of all the votes would then be entered as conventional data and analyzed as though they were actual observations of fish. Although this technique may seem rather subjective, it has some definite strengths. First, it is expedient, especially for species that are difficult to observe in the field (such as rare and endangered species, with very low distribution density). Second, it requires very little specialized equipment. Third, because fish capture and observation are not required, habitat recognition is free of many (but not all) of the potential biases of these techniques.

3.3 DISCUSSION

Critics of category I approaches are numerous, vociferous, and often, derogatory. Phrases such as "crystal ball" and "voodoo science" are often associated with criticisms of these techniques. To be sure, there are numerous pitfalls to be avoided in category I inquiries. Communications (or lack thereof) are a constant worry to the project officer. Experts often do not know as much as they think they do. The project leader may be guilty of a

Species:	Brown trout		
Life stage:	Juvenile	Size limits:	5 - 15 cm
Activity:	Resting	Season:	Summer

Location ID	Suitability of location for organism specified above: Yes	No
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		
13		
14		
15		
16		
17		
18		
19		
20		

Figure 17. Sample secret ballot for use with habitat recognition for developing habitat suitability criteria.

superficial analysis of responses and respondents may become impatient to get the job over with. One particular concern with Delphi is its potential for deceptive, manipulative purposes. The anonymity of the process may facilitate the potential for deception.

In defense of these methods, several points must be made. First, none of the methods should be viewed as "crystal ball" or "voodoo science." To place these approaches in the same category as fortune telling is to discount the years of experience and knowledge accumulated by professional fisheries scientists. When any category I method is used, a conscious decision has been made to substitute expert opinion for direct knowledge (data). Therefore, these approaches cannot legitimately be called science, voodoo or otherwise. They are, however, rational and organized approaches for providing information to the decision-making process. As Zuboy (1981) concluded, "In the final analysis, it is the basic integrity of the investigator that determines the results of an experiment, not the inherent characteristics of a particular methodology." One can only add that Zuboy's statement is true of real science, as well.

The accuracy of suitability criteria varies directly with the experience and knowledge of the panelists. Naturally, people who have spent large amounts of time observing the target species and measuring components of its habitat are more able to estimate tolerances and preferences than those who have not. Baldridge (1981) conducted a two-step criteria development study on several species of salmon in the Terror River, Kodiak, Alaska. The first step consisted of a traditional committee meeting of fisheries biologists to develop criteria through professional judgement. The second step consisted of criteria developed using empirically-derived data. The results of Baldridge's comparison are illustrated for spawning pink salmon in Figure 18. These results are typical of the criteria for most life stages and species documented in her study.

The amount of agreement between the interim and final curves for the Terror River study may come as a surprise to some. It should be recognized, however, that fisheries biologists in Alaska have been measuring microhabitat conditions over salmon redds for years. The real surprise would be to obtain this degree of convergence for a lesser known species. The point of this example is to illustrate that true experts can assemble highly accurate habitat criteria using only their experience and intuition.

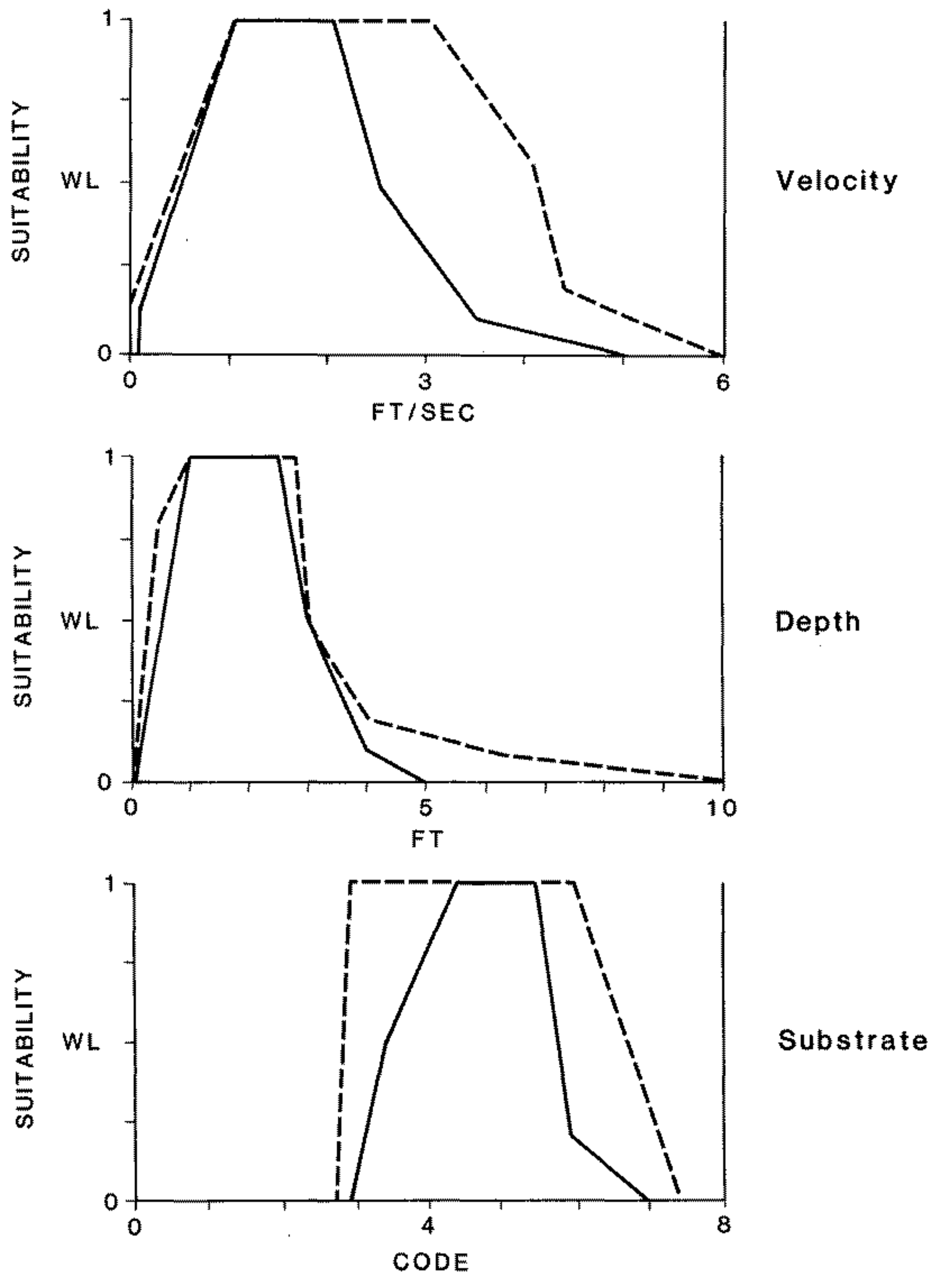


Figure 18. Comparison of habitat criteria for spawning pink salmon generated by professional judgement (dashed curve) and by data analysis (solid curve). From Baldridge (1981).

4. COLLECTION OF HABITAT UTILIZATION AND PREFERENCE DATA

Habitat utilization and preference criteria are based on data collected and analyzed specifically for determining the microhabitat characteristics of a species. The increased accuracy and precision of category II or III criteria requires an investment of time and money and a commitment to quality control during development. Whereas category I criteria may be constrained by a lack of information, the major concern with categories II and III is biased information. An investigator will usually recognize the general nature and potential inaccuracies of criteria developed by a committee. Category II and III criteria carry with them the strength of empirical data, and because data biases are often well disguised, the investigator may not recognize these limitations as readily. Consequently, it is imperative to understand the advantages, disadvantages, and limitations of various sampling strategies and methods in order to select the most appropriate techniques for a particular river and target species.

As previously defined, a utilization (category II) function is based on a frequency analysis of microhabitat characteristics at specific locations in a stream. A datum in a utilization function represents the presence of one or more individuals of the target species at a measured location, either utilizing it as a focal point or as a central portion of the species' home range. Because data are only collected where target organisms are found, the function defined is $P[E|F]$, the probability of occurrence of a particular environment, given the presence of a fish or macroinvertebrate. $P[E|F]$ represents, in part, the habitat characteristics preferred and selected by a species. It also represents, in part, the habitat characteristics available to the species at the time of sampling. This means that utilization functions may be applicable only in the streams in which they were developed, or in others very similar to them.

The development of new habitat criteria every time a different project or river came under study would be very inefficient, time consuming, and costly. Furthermore, it is difficult to evaluate trade-offs among different rivers or sites if different criteria are applied to each. These limitations to the utilization function led to the development of the preference function (Voos 1981). The concept underlying the habitat preference function is the same as that discussed by Ivlev (1961) for determining food electivity. The ratio between the proportions of food items in the ration and in the food complex is used as an index of selective feeding. A food item that occurs in a higher proportion in a fish's stomach than it does in the food complex means that the fish has actively selected that item in preference to others. By analogy, an organism found in a higher proportion in a particular environment, compared to the availability of those conditions, has actively selected that set of conditions. If the organism's frequency distribution is identical to the distribution of environmental conditions, then it is randomly distributed; all measured conditions are equally suitable for the species. The relative preference of an organism (henceforth called P_r) is defined as the ratio between utilization and availability (Voos 1981):

$$P_r = \frac{P[E|F]}{P[E]} \tag{5}$$

where P_r = the relative preference index of a species for a specific set of environmental conditions;

$P[E|F]$ = the probability of occurrence of a specific set of environmental conditions, given the presence of one or more individuals; and

$P[E]$ = the probability of occurrence of that set of environmental conditions in the stream at the time the organism was sampled.

The advantage of the preference function is that it is more independent of the sampling environment than the utilization function. Although this attribute greatly expands the transferability of preference functions, they are not universal, nor are they completely free of environmental bias. The primary disadvantage of category III criteria is that the data requirements to develop them are larger than those to develop category II criteria. A complete category II data base is needed to describe the numerator in equation 5. Additionally, the availability of environmental conditions at the time of sampling, P[E], must also be determined each time data are collected for P[E|F]. Depending on how the investigator chooses to describe P[E], the additional data requirement can increase field time from 20% to 100% over typical category II studies. These measurements are generally easier to obtain than those related to microhabitat utilization, but they can represent a fairly substantial data collection effort.

As mentioned in Section 2.6, the sampling design for category II and III criteria is very important in minimizing bias due to disproportionate sampling. This concern is valid when collecting category II data under a single set of environmental conditions, and it becomes increasingly important when data from several sources are pooled together. The potential for this type of bias is greatest when developing category III criteria from pooled data. Since category III criteria are generally more desirable from a user's perspective, precautions against bias must be taken at each step by the researcher. The following sections, therefore, discuss the development of the category II data base, the habitat availability data base, and strategies to avoid data pooling problems.

4.1 DEVELOPING THE CATEGORY II DATA BASE

The development of a valid utilization function requires the unbiased measurement of microhabitat variables at specific fish (or macroinvertebrate) locations. This means that the actual point location of each fish must be determined as accurately as possible, further implying that the observer and the collection technique must not interfere with the observation. Wickham (1967) defined the focal point as a central point in a fish's home range. For a stationary swimming trout, the home range is small and the point location well defined. The home range for a smallmouth bass is larger, but the focal point can still be determined quite accurately (Gerking 1953). Migrating or transient fish, however, do not exhibit a home range behavior so they must not be included in the sample, unless migration criteria are being developed. This requirement automatically disqualifies the use of techniques or gear designed to capture fish in transit, such as stationary gill nets or Fyke nets. It is equally important not to lure the fish away from its focal point, and then capture it. This eliminates the use of trotlines, baited traps, and angling as collection techniques. Finally, the collection or observation technique should not be so disruptive, or sample such a wide range of habitat conditions, that the point location cannot be accurately determined. The use of explosives sometimes fits the first category, while beach seines and poisons fit the latter.

Techniques that may be applicable to criteria studies include direct observation, biotelemetry, electrofishing, small-area samplers (including explosives), and certain laboratory techniques. The appropriateness of any technique depends on its inherent limitations, the behavior of the animal, and the environmental conditions under which it is applied. Once these factors have been evaluated, the efficiency, safety, and relative cost of each technique must also be considered.

4.1.1 Direct Observation

One of the most effective ways to learn about microhabitat utilization by fish is to quietly and unobtrusively watch them. The basic distinction among direct observation techniques is whether the vantage point of the observer is above or below the surface.

a. Surface observation. Surface observation of fish can be an extremely efficient method of collecting habitat utilization data, but only if its several limitations can be overcome. The concept of surface observation is nearly too simple to document. The observer simply scans a section of stream documenting the location, species, approximate size, frequency, and activity of every member of the target species in the section scanned. The most practical way of documenting these locations is to mark them on a scale planimetric map. Then, after the scanning period is over, microhabitat measurements are taken at each marked location.

The advantages of surface observation are high efficiency and minimal disturbance to the fish. It is possible to observe many fish quickly, easily, and at low cost. The primary limitation of surface observation is the potential bias caused by differences in visibility from one location to another.

The low incident angle between water level and eye level often results in a large amount of glare. The main source of bias caused by glare is that fish nearest the vantage point are easier to see and identify than those farther away. This problem can be partially corrected by wearing polarized sunglasses and using binoculars. Another remedy is to increase the angle of incidence by selecting an elevated vantage point. Bachman (1983) describes the use of observation towers for monitoring microhabitat usage by brown trout in a small stream in Pennsylvania. Bachman enjoyed such good visibility from his towers that he could identify individual fish by the spots on their backs. This situation might be the exception rather than the rule, but it does demonstrate the effectiveness of surface observation in small, clear streams. The construction of permanent observation towers is probably impractical for most criteria studies, but most streams have built-in observation towers, called trees. A portable tree-stand, such as used in bow hunting, can provide a stable and comfortable platform from which to make observations. The tree-stand is versatile, economical, and much safer than sitting on a branch. A major limitation of surface observation techniques is the glare occurring under "flat light" conditions associated with gray, overcast days. Polarized glasses and elevated vantage points seem to provide little improvement of visibility under these conditions.

Surface visibility can also be affected by a broken water surface and entrained air bubbles associated with wave action and turbulence. This creates a bias toward more observations being made in pools and other flat water areas than in riffles or runs. Sometimes, the use of a viewbox from a boat can alleviate this problem, but it is usually better to use another method in turbulent water.

Surface observation is obviously influenced by water clarity and incident light. Its effectiveness is reduced in water with less than about 6 meters (20 feet) visibility. The potential for missed fish and misidentification increases (probably geometrically) as turbidity increases or incident light decreases, so surface observation should probably not be attempted at night or if visibility is less than two or three meters. The technique is also limited to relatively small streams, unless a boat or some other flotation device is used. The use of observation towers or tree stands can increase the allowable stream size somewhat, but even with these, streams over 30-40 meters wide are difficult to scan without introducing bias. Finally, there may be a bias toward larger fish because they are easier to see, and against some species such as rock bass, madtoms, and darters, which may be virtually invisible from the surface.

b. Underwater observation. Fishes that flee from a person walking the bank, wading, or even floating in a canoe are often nearly oblivious to a diver. To some extent, underwater observations are affected by visibility problems similar to those of surface observations. A diver, however, can usually approach fish much closer than a surface observer could, and can often sight fish under conditions of higher turbidity. Divers are not affected by the problems of glare and wave action that plague surface observers, and visibility in turbulent water is often good, even with air bubbles entrained in the water. The proximity of the diver to the fish alleviates problems of misidentification and size estimation, and allows the accurate determination of the fish's vertical position in the water column for nose velocity measurements. Also, fish positions can be physically marked by the diver, rather than estimated from a map. These features of direct underwater observation make this one of the most valuable of all data collection techniques. It does require special preparation and procedures, however, and is not without disadvantages and limitations.

The simplest, least expensive, and least disruptive method of observing fish in the field is snorkeling (Helfman 1983). The basic equipment for a skin diver includes a mask, fins, snorkel, and a wetsuit (in all but the warmest waters) or a drysuit (in cold water). Drysuits are recommended in all but the warmest water, when the diver will be immersed for more than an hour. The experience of many divers suggests that discomfort can lead to erroneous, if not biased, data. The drysuit's primary function is protection against the cold, but it also provides bouyancy, an important safety feature, and can protect against sunburn and abrasion. Skin divers not wearing a wetsuit should wear a life vest or an inflatable bouyancy compensator (commonly referred to as a "bc"). The bc is preferred because it allows underwater excursions, whereas the life vest gives constant flotation. Skin diving in high gradient streams, often resembles rock climbing more than swimming, so divers may wish to wear tennis shoes instead of fins. Knee pads and gloves are not necessary, but will often be greatly appreciated after a day of crawling over rocks. A dive knife should be worn by each diver, especially if the site has received any substantial fishing pressure. Monofilament line becomes invisible under water and can be a considerable hazard to a diver who becomes entangled in it. The knife should be strapped to the inside of the thigh or calf to prevent it from becoming entangled and ensure accessibility by either hand (Helfman 1983). If the stream has any boat traffic, a surface tender should accompany the diver to alert boaters. Many boaters do not recognize or respect dive flags, so these should not be relied on for protection.

The preferred sampling strategy with snorkel or SCUBA is to work in an upstream direction. Nearly all species of fish face into the current (positive rheotaxis), which allows the diver to approach the fish from behind. Although moving upstream is more tiring than drifting with the current, it has several other advantages besides a stealthy approach. The first is that a more constant rate of movement can be maintained in the upstream direction. Drifting with the current can introduce a slow-water bias because the diver will spend proportionately less time in high velocity areas and more observation time in slow areas. The second advantage of upstream movement is that it is safer. No one has ever been swept upstream over a waterfall!

It is difficult for humans to swim upstream, even against a weak current. This poses a special problem for skin divers on the surface where the current is strongest. In shallow water, the diver can pull him or herself along the bottom or move around the wakes of boulders. This is impossible in depths over about 90 cm unless the diver resorts to free diving, which can also be very tiring. The most common solution is to stretch a static line, such as 1/8 inch aircraft cable, across the stream and suspend one or more lengths of

1/2 - 3/4 inch polypropylene rope (usually about 100 meters long) from the cable. Divers can then pull themselves upstream along the rope. Some investigators tie overhand knots at intervals along the rope, while others use mountaineering ascenders to reduce diver fatigue. If ascenders are used, they should have a quick release mechanism, or each diver should carry a knife capable of cutting the rope very quickly. Although this technique has a good safety record, it is potentially dangerous if the diver cannot disengage rapidly from the rope.

Underwater observation is effective for locating, identifying, and enumerating fish, but it is difficult for a diver to carry any equipment needed to measure microhabitat variables. The most popular solution to this limitation is to mark fish locations with small weights attached to colored surveyor's flagging and measure microhabitat variables at each marker after the observations have been completed. Flagging can be color coded to represent different species or size classes and the appropriate colored flag dropped at each fish location. Helfman (1983) suggests the use of large nails (10-penny or larger) as weights, and describes how to make a carrier for them out of a bicycle inner tube. Lead fishing weights are equally acceptable and can be carried in a diver's "goody bag," a closable mesh purse. It is important that the weights are heavy enough that the markers will not drift downstream, but not so heavy that they encumber the diver. Surveyor's flagging is available in a variety of colors, but, for this use, the best colors are white, yellow, hot pink, and international orange. Avoid using blue or green as these colors will disappear instantly under water. If four colors are not enough, various combinations can be used to specify a life stage or species. It is also a good idea to number the flags and drop them sequentially. This helps the measurement crew determine when all the fish sighting locations have not been measured. An alternative to using underwater markers is to deploy marker buoys at each fish sighting. These consist of an anchor weight attached to a small float, such as a styrofoam ball, by a length of string. The floats can also be color coded or numbered. Buoys are less likely to be overlooked by the measurement crew, but are far more difficult for a diver to handle. Therefore, underwater markers are generally preferable, unless the measurement crew has great difficulty in relocating them. The surveyed area and the number of different color codes should be small when using the "sight and mark" technique. The area covered per diver during a sighting session should probably be no larger than about 1,000 square meters. This is ample room in which to lose markers. Use of more than six or eight different color combinations is also discouraged. It becomes difficult to remember what colors represent which fish, so dropping the wrong flag is increasingly likely if there are too many combinations. In these situations, recording the marker number, species, and related information, is the preferred technique.

An alternative is to use numbered markers. This requires the diver to record the pertinent information for each sighting (species, size, frequency, and activity) corresponding to the number on the marker. A dive cuff (Figure 19) is probably the most useful underwater data recording system for a skin diver. Although very little data can be recorded on a dive cuff, very little needs to be recorded using this system. The best feature of the cuff is that it leaves both hands free except when recording data.

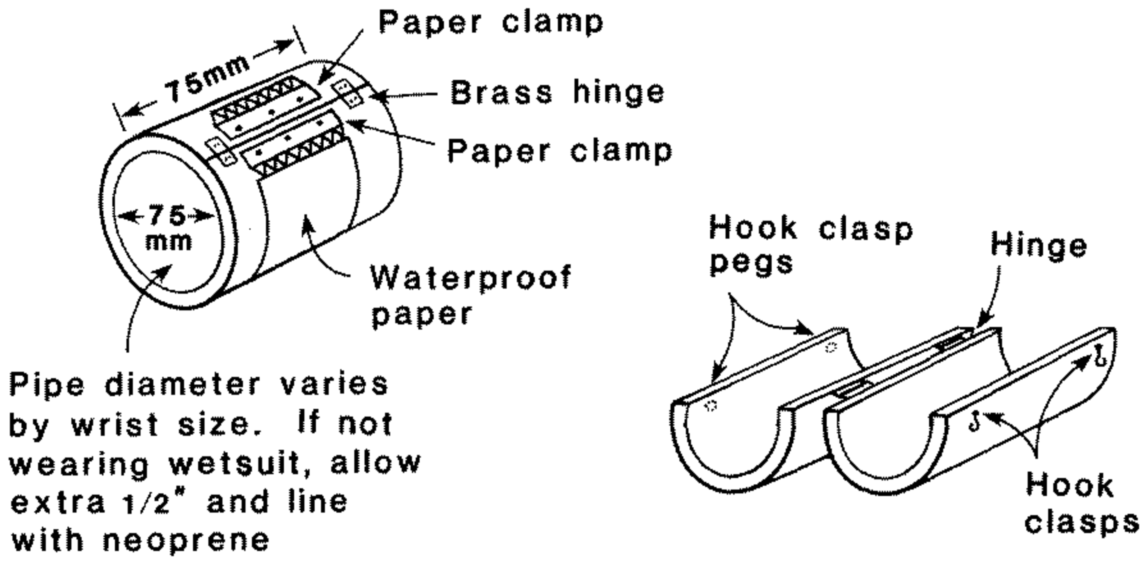


Figure 19. Dive cuff for underwater data recording.

SCUBA allows the diver to spend extended periods of time on the bottom of the stream rather than on the surface. Frictional resistance of the streambed causes water velocities to be much lower at the bottom, so the SCUBA diver is not exposed to the high velocities encountered by the skin diver. Furthermore, the streambed offers many good handholds, by which the diver can maneuver from place to place. The superior maneuverability and increased bottom time of SCUBA allows the diver to carry a limited amount of the equipment needed for microhabitat measurements. This can allow microhabitat measurements to be taken at each fish sighting, although the "sight and mark" technique can be used effectively with SCUBA, as well. SCUBA divers often have a much better view of the bottom than surface divers, so their substrate descriptions are likely to be more accurate in deep or turbid water.

Depth can be measured directly from the diver's depth gage, rather than by sounding. Many of the cheaper depth gages, however, are only accurate to a few decimeters, so it is important to obtain one that has higher accuracy. Furthermore, the depth gage should be adjustable for elevation and barometric pressure as these will affect the reading. Velocity measurements may require some minor deviation from the way they are made by wading or from a boat. The fact that the measurer will be submerged essentially eliminates the use of electronic current meters, unless the circuit box is encased in a water-tight container or left aboard a tender (boat), with the diver carrying only the probe. This all seems rather clumsy when compared to having the diver carry a Price AA or other horizontal vane current meter. In theory, all the diver really needs to carry is the meter itself, but mounting the meter on a short wading rod will give both the meter and the diver added stability. Obviously, the standard headset used with this type of meter cannot be used under water. The solution is to paint one of the cups of the meter with a bright color. Revolutions are counted each time the painted cup passes through the yoke of the meter. Use of this meter requires either a waterproof stopwatch (usually needs to be waterproof to only about 15 meters) or a good dive watch with a sweep second hand. Digital underwater dive watches with stopwatch, interval time, alarm, and light have recently become available and may be well suited for this work. It may also be desirable to carry the wire or string grid described in Section 2.4.3, for determining substrate composition. Other useful equipment includes: a half-meter stick with thickened lines at 10 cm intervals (for estimating fish sizes), a thermometer, a small dip net for collecting small fish, and a camera. Fish sizes may also be estimated from the wading rod, so the measuring stick may be unnecessary.

Collecting microhabitat data under water is relatively simple; recording it is not. Unless the "sight and mark" technique is used, the amount of data that might be recorded is potentially quite large. Hence, only a few data recording systems are feasible with SCUBA observations. Clipboards with underwater paper, or continuous scrolls appear to be the most useful for this type of application. The former consists of a standard clipboard, fitted with a surgical tubing loop in one corner for a handle. The loop can also be worn on the elbow or wrist to keep one hand free. A number 2 pencil, sharpened at both ends is attached by a string or a length of tubing at another corner. Two large rubber bands are slipped around the bottom of the board to hold the loose end of the writing material in place (Helfman 1983). It is probably advisable to use one at the top, as well, in case the clip is inadvertently opened or broken off. The actual writing material is a plasticized, waterproof paper, such as Polypaper (See Appendix B). Waterproof field books used in surveying might also be appropriate and are available at most surveying and engineering supply houses.

The original design of the continuous scroll is described by Ogden (1977). Instead of individual writing sheets, the scroll consists of a long strip of polyester drafting film that is fed from one roller to another across a flat writing surface. The best feature of the scroll is that it eliminates the need for changing writing sheets or flipping pages in a field book, both of which can be frustrating in currents and while wearing diving gloves. Losing a page of data in a river is a serious accident because the sheet will be gone with the current with little or no chance of recovery.

Some researchers have used verbal data recording techniques to overcome the difficulties inherent with writing tablets and scrolls. One approach has been the use of underwater tape recorders. (Suppliers listed in Helfman 1983.) The biggest problem with these is the difficulty of speaking through a standard SCUBA regulator, which usually results in rather garbled messages. Gosse (1982) used a full face mask with a sonic transceiver to communicate with surface personnel who recorded data as it was relayed from below the surface. Gosse gives no indication of communications difficulties, and it is likely that his conversations were easier to understand than those made into a standard regulator. The difficulties in recording data discussed above may have been overstated for many applications. The majority of criteria studies will be conducted in streams that are shallow enough that the "sight and mark" technique will be feasible. This is probably a more efficient and safer method than direct measurement and recording. The latter should be used only in large systems where "sight and mark" cannot be used effectively.

One of the advantages of underwater observation techniques (either snorkeling or SCUBA) is that observations can be made at night, at least in some streams and under certain conditions. There are several precautions that must be considered prior to night diving, however. One of these is the reaction the fish might exhibit in response to a dive light. Most fish are alarmed by bright lights, especially if they are suddenly exposed to them when a diver sweeps the light back and forth. In some situations, it may be possible to turn off the dive light and approach the fish using a Cyalume light stick. These produce a diffuse green light that illuminates only a small area around the diver. Fish spotted near the margins of the range of the dive light may be approached under near blackout with the light stick.

Helfman (1983) suggests that the objectives of night dives should be reduced, compared to day dives. One way to accomplish this in a criteria study is to follow a modified "sight and mark" procedure. Rather than attempting to measure microhabitat variables in the dark, fish positions should simply be marked and measurements taken the next morning. It may be more convenient to use color coded markers than numbered ones, when diving at night. A staff gage should be installed near the study site to monitor flow changes that might occur between the observation and measurement periods. Sampling periods when the flow will be relatively stable are preferred to those exhibiting large diurnal changes. Measurements can be taken immediately if the staff gage reading is the same in the morning as it was the preceding night. If the flow has changed, wait until the same staff gage reading is reached before measuring microhabitat variables at the markers. It is also a good idea to firmly anchor the marker to the bottom so it will not move downstream during the interim period.

Perhaps the largest drawback to SCUBA as an observational technique is the potential disturbance created by the venting of bubbles. The reaction of fish to exhaled bubbles seems to vary not only from species to species, but from population to population. The same species might ignore a diver in one area and exhibit a fright reaction in another. The possibility that vented bubbles could frighten the fish might encourage the diver to hold his breath, a practice called skip-breathing. This procedure is strongly discouraged, first, because it can be dangerous and second, because it is counter-productive. After holding one's breath, the exhalation tends to be explosive. If the fish were skittish around a normally breathing diver, they are certain to be frightened by one who is skip-breathing. The best solution is for the diver to maintain a nonthreatening distance to the fish if they appear frightened by the bubbles. Another approach would be to make several "get acquainted" dives to let the fish become accustomed to the presence of the diver before making observational dives. Underwater blinds might also be used to conceal the diver. Although this will not solve the bubble problem, it may be the combined sight of diver and bubbles that frightens the fish. Finally, the diver should avoid wearing bright or shiny equipment. Avoid red or yellow wetsuits (wear black, blue, green, or camouflage) and shiny tanks or other gear.

SCUBA diving is considered to be more hazardous than most of the other data collection techniques. These hazards are compounded in criteria studies because they are conducted in rivers rather than flat water. Therefore, safety is paramount, and several steps must be taken prior to using this technique. First, it is mandatory that all divers are certified by a licensed instructor. Reputable dive shops will refuse to fill tanks unless a diver's certification card is presented. In fact, many dive shops insist that a current logbook also be presented. Divers who were certified several years ago and have not dived since, may be required to take a refresher course. In addition, most government agencies have a safety officer or committee that must be consulted prior to diving.

Beyond satisfying safety regulations for diving, one should also prepare for the special hazards of river diving. Nearly everyone learns to dive in still water: swimming pools, ponds, quarries, or ocean bays. Before engaging in a river dive, additional instruction by a licensed instructor is highly recommended. It is advisable to begin river diving in the least threatening situation possible: snorkeling in a stream shallow enough to wade. Experience is the most important aspect of safe diving, so it is best to start in the safest situation and move to more difficult situations at a comfortable rate. It may be necessary to allocate 10 to 20 dives to gaining the experience needed for safe river diving.

The obvious distinction of river diving is the problem posed by the current. The most common solution for a SCUBA diver is to wear additional weight. This, in turn, may require supplemental bouyancy control. The current can present several dangers to a diver. Simple exhaustion from fighting the current is one, but, perhaps the greatest danger is in becoming wedged in bottom obstructions. The headgates of diversions are particular hazards in the West. For this reason, above all others, no one should ever dive alone, not even snorkeling. Diving anywhere near diversions or intake structures should be strictly avoided.

Despite the dangers presented by currents, river diving may actually be safer than other types of diving in some respects. Turbidity frequently causes problems in diving, but since the goal of a river dive is to observe fish, dives should not be conducted at all if the visibility is less than a meter and a half (Platts et al. 1983). This is much greater visibility than one might encounter in a rescue or recovery operation. Although ice diving can be conducted in lakes or ponds, it should never be attempted in a river unless the diver is highly qualified for such dives.

Most river dives will be no-decompression dives because depths rarely exceed 30 feet. The diver should be aware, however, that decompression tables are calibrated at sea level. At higher elevations, the atmospheric pressure is reduced, so the no-decompression limits and depths are shorter and shallower. Consult a local diver or diving instructor to determine safe limits and depths at altitude.

Safety considerations for river diving are considerable but not insurmountable. Under proper circumstances, SCUBA or skin diving are possibly the best techniques that can be employed to determine habitat utilization of fishes. The key is to use these techniques only where it is safe to do so. In some respects, a person's instinct for self-preservation is the best indicator. A person who is not comfortable with a particular dive is not likely to collect very good data. Therefore, if the diver is more concerned with survival than with data collection, some other, safer, technique should be used.

4.1.2 Underwater Video

Divers are limited in the kinds of conditions under which they can work in safety and in comfort. While diving is superior in many situations, there will be others where the water is too fast or too cold, covered by surface ice, or otherwise too dangerous for a diver. Obtaining microhabitat utilization data under these circumstances may be accomplished using a remote observational technique, rather than direct observation.

Much of the literature discussing the use of underwater video refers to its use in monitoring the behavior of fish with respect to capture devices. The use of video in making quantitative microhabitat measurements is a new application that may require different procedures and equipment. Furthermore, because this technique is so new, its feasibility for criteria studies has not been tested or evaluated.

Essential features of a remote underwater video system consist of an underwater camera, a remote monitor or recorder, and a support system. The support system includes some sort of base on which the camera can be mounted, auxilary lighting system, power supply, and cables. Similar equipment is needed in support of fixed-base still photography systems, the primary difference being the type of camera mounted on the base.

Use of a remote system implies that conditions are too dangerous to attempt direct observation. Given this implication, an immediate problem is how to lower a very expensive piece of equipment into the stream and make it stationary on the streambed, without damaging or losing the camera. Two approaches seem plausible: have a diver carry the camera down and mount it to a preset base, or attach the camera to the base aboard a boat, and lower the entire apparatus to the bottom using a winch. In either case, the base must be heavy enough that it will be stable on the streambed. It must also be designed in such a way that the torque exerted on the camera will not tip the base over. A dome-shaped concrete base may be more stable than other designs because of its wide bottom and streamlined contour. A concrete dome 60 cm in diameter by 30 cm tall will weigh nearly 140 kg, so the necessity of a winch is obvious. A slightly smaller dome, 24 cm tall and 48 cm diameter, would weigh about 70 kg.

A single-unit base and camera may create some problems in recovery. The simplest means of recovery would be to leave the winch cable attached to the base. This will be satisfactory most of the time, but drifting debris could snag the cable and pose a threat to the camera and crew. Therefore, the cable system should be equipped with some type of breakaway or quick release device aboard the boat so the winch end of the cable can be jettisoned in emergencies. (Fixing a small buoy to the free end of the cable will aid in its recovery once the crisis has passed.) This makes it desirable to have a remotely detachable swivel and camera with a separate recovery line. In the event that the base becomes hopelessly caught on the bottom, at least the camera and swivel can be recovered.

Observing fish with video equipment is relatively easy, but pinpointing their position in the stream is not. Use of this equipment almost certainly dictates some type of site preparation. A detailed site map, including cross sectional profiles and underwater landmarks will usually be needed. Sonar profiles should be made if surveyed transects are not used. At the very least, these should help in camera placement to prevent the base from being set up over a ledge or on the side of a drop-off. If underwater landmarks are large and distinctive, it may be possible to locate fish positions with respect to the landmark and the camera position. Where landmarks are absent or non-descript, it will be necessary to place some sort of reference monuments on the bed. One approach is to use rocks painted a contrasting color from the rest of the substrate. Each rock should be numbered prominently and placed on the bed in a grid pattern, with the numbers pointed toward the camera. As the camera scans around, fish can be located, identified, their size estimated, and their position with respect to the monuments determined. As an alternative to using monuments, Wegner (U.S. Bureau of Reclamation, Flagstaff, AZ; memo dated Feb. 19, 1985) suggests lowering a weighted grid to the streambed to outline an area of interest. Either way, this information is then transferred to a plan map showing the grid pattern, and the appropriate locations measured at the end of the observation session.

A good monument needs to have two qualities: it must be obvious on film or monitor, but ignorable to the fish. The type of monument used depends on the existing substrate and whether the camera is color or black and white. Figure 20 shows several monuments, made from natural and man-made materials. The best color pattern to use with a black and white camera depends on the background color of the substrate. Do not use black monuments on basalt, for example. With a color camera, any contrasting background color can be used, with hot pink, international orange, or yellow numerals (or black and white). The second attribute of a monument is that it must not act as a fish attractor. Natural substrate materials are preferred monuments for this reason, provided they are large enough to be seen by the camera. The concrete monument shown in Figure 20 might be needed in sand, mud, or small-gravel substrates. Before mapping any fish locations, it is advisable to watch the fish for a while to make sure they are not attracted by the color patches that identify the monuments. The same is especially true when using artificial monuments in streams lacking cover or large substrate, because fish may use the monuments for shelter.

A video system or camera can be used anywhere it is feasible to dive and many places where it is not. This may be a good alternative for observing fish that are disturbed by bubbles emitted from a SCUBA diver. (If the camera emits bubbles, there is a more serious problem than frightened fish.) Video systems are not limited to bottom time, they do not get cold and tired, they can withstand fairly strong currents, and they can be used under ice. If a video system is destroyed, it is a serious matter, but not nearly as serious as losing a diver.

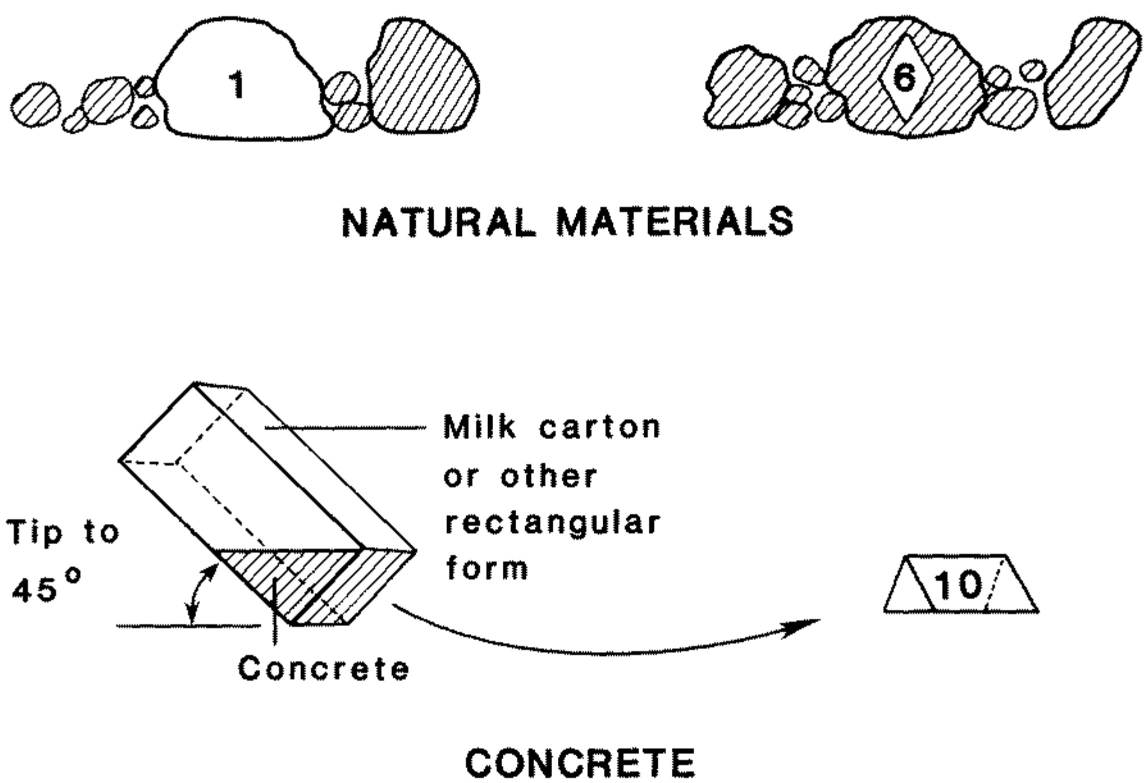


Figure 20. Natural and manmade grid monuments used to locate fish positions with underwater photography and video systems.

Cameras are obviously limited by the same visibility problems that hamper divers. Because this equipment is less maneuverable than a diver, it may be more restricted by turbidity. It may not be worth the effort to set the camera up if only a two-meter radius can be scanned. Furthermore, it may be unwise to observe fish in close proximity to the camera because they may be using the base as cover. Conversely, the use of camera equipment used at night might be more effective than a diver, through the use of remote-controlled auxillary lights. To overcome the problems of light avoidance, the video camera is aimed at a specific area of the grid, activated, and then the lights turned on. Nighttime operations essentially use video systems as very expensive still cameras, and require video recorders in addition to monitors. The fish will be photographed just for an instant as the lights are switched on and they will undoubtedly be alarmed soon thereafter. The fact that the fish have been disturbed is immaterial if their original position has been determined. After each shot, the lights should be turned off for several minutes before another is taken.

Underwater video systems are expensive, costing about 10 times more than a good still camera, but are preferred to still photography despite the large cost differential. The most important reason is the instantaneous monitoring capacity of video. The locations of fish can be determined instantly with a video system. The quickest determination using still photography requires a polaroid-type camera in a watertight case (cases cost about $200). Standard underwater still cameras would be largely infeasible because the film would need to be developed before fish locations could be established. Even a polaroid-type camera is not very efficient, because the camera would need to be retrieved after each roll of film was shot. The second advantage of the video system is that it is generally more compatible with remote operation. Camera orientation, lighting, and focus can all be adjusted electronically, rather than relying on a more cumbersome mechanical system. Finally, video offers the ability to observe the fish over a period of time instead of taking a snapshot. This allows the observer to describe the activity of the fish and to distinguish between stationary or cruising fish and those in transit through the reach.

Remote photography and video are generally safe techniques. The greatest danger occurs when the base is being lowered or recovered. It is extremely important to use a large enough boat to accomodate a winch, boom, and the camera base. Long winch booms should be avoided, especially with smaller boats. Long-armed booms create rotational torque when lifting the base, which can lead to capsizing. Finally, the gears and drum of the winch should be shielded to prevent entanglement of clothing or fingers.

4.1.3 Biotelemetry

Biotelemetry involves the attachment (either externally or internally) of a device that emits a signal and the use of a receiver to locate the source of the signal. This technology has been used extensively in biological studies since the early 1960's, primarily in the determination of the movement of animals or to monitor physiological parameters, such as the body temperature of hibernating mammals. The use of biotelemetry to determine habitat preferences of fish (with the precision required by IFIM) is more recent, dating to the early 1980's (Tyus 1982; Gerardi 1983; Tyus et al. 1984). The relative youth of this use of the technology means that the state of the art is still subject to change and evolution. Biotelemetry requires a variety of skills on the part of the investigator, ranging from surgical procedures to repair of electronic equipment. There are also numerous pitfalls to be avoided in its use, and it may not be applicable everywhere. Nevertheless, the effectiveness of telemetry has been demonstrated in several instances (particularly in large, turbid rivers) where other methods were either biased or ineffective.

Two types of telemetry systems are commonly used in the study of fishes. Ultrasonic systems transmit and receive a high-frequency sound wave, whereas radio frequency systems send and receive relatively low-frequency radio waves. Because sound will not effectively cross the interface between water and air, the receiver (hydrophone) used with ultrasonic systems must be submerged. Radio waves are rapidly attenuated in water, more so in highly conductive water, but those signals that reach the surface at an incident angle of less than six degrees are transmitted through air. Attenuation of radio signals in air is negligible and signals can often be detected at considerable distances. It is beyond the scope of this article to describe the equipment and theory underlying either system in much detail. A good description of ultrasonic telemetry systems can be found in Stasko and Pincock (1977). A complete and comprehensive guide to radiotelemetry equipment and techniques is provided by Amlaner and MacDonald (1980). This reference provides information on circuitry, battery performance, and antenna design for transmitters and receivers, as well as numerous case studies of applications. For a variety of reasons, radiotelemetry has proven to be superior to ultrasonic systems in riverine environments. Therefore, the following discussion will concentrate on radio systems, although some concepts apply to ultrasound, as well.

a. Transmitter implantation. There are three methods of attaching transmitters to fish: external attachment, insertion into the stomach, and surgical implantation. The latter technique is generally considered the most satisfactory of the three. Externally mounted transmitters create drag, can become snagged, cause abrasions, or create balance problems for the fish (Ross and McCormick 1981). Any of these factors can cause a change in the behavior of the animal. Stomach inserted transmitters can injure the esophagus or stomach of the implanted fish, either resulting in altered feeding behavior or regurgitation of the implant. The former problem can result in tracking and monitoring an abnormally behaving fish, whereas the latter can result in time wasted tracking a naked transmitter.

These consequences are not as likely with surgically implanted transmitters. Once the incision has started to heal, fish exhibit normal activity and behavior patterns and the loss of transmitters is minimal (Chamberlain 1979). Surgical implantation may result in some mortality, and is not suitable for gravid females or very small fish. The biggest disadvantage of surgical implantation is that the time required to implant transmitters and the rate of mortality among implanted fish are both related to the experience of the surgeon. Although there are several good references on surgical techniques (Hart and Summerfelt 1975; Bidgood 1980; Winter 1983), newcomers to the field would be well advised to have an experienced person show them how to do it, and then practice the techniques before attempting to conduct a field study.

There is general agreement that subject fish should be held for a day or two prior to surgery, to ensure that their capture has not overly stressed them. Immediately prior to surgery, it is standard practice to anesthetize the fish, most commonly with a quinaldine or tricaine methanesulfonate (MS222) solution. Appropriate concentrations for sedation and for anesthesia with these and other chemicals can be found in Stickney (1983). Once the fish have attained the desired level of relaxation, surgery should be initiated as quickly as possible to avoid overdosage.

A common field surgery technique is to place the fish in a V-shaped trough, bathing the gill cavity with water or anesthetic via a siphon tube or squeeze bottle (Winter 1983). An incision just large enough to admit the transmitter is then made in the deepest part of the body. Some practitioners suggest making the incision slightly to one side of the mid-ventral line (Hart and Summerfelt 1975; Winter 1983), while others prefer a more lateral incision, anterior and dorsal to the pelvic fin (Tyus 1982; Otis and Weber 1982). The best location may vary from species to species, but the important point is to avoid vital organs.

After making the incision, the transmitter capsule is implanted; those with whip antennas by extending the whip inside the body cavity first, using a hemostat forceps. If the transmitter has a loop antenna, the loop should be inserted more or less vertically to maximize signal propagation toward the surface. Otis and Weber (1982) had to disassemble their transmitters, rotate the antenna 90 degrees, and reassemble them prior to implantation; however, by so doing, the range was increased by over 100%. It is also important to remember that some transmitters are equipped with magnetically activated leaf switches. If so, do not forget to turn them on before implantation.

Following implantation, the incision is sutured shut, possibly the trickiest part of the surgery. Winter (1983) suggests practising suturing on a piece of cloth stretched over the open end of a can before trying it on a live specimen. Poor suturing is time consuming and can increase the time it takes for the incision to heal. More critically, the transmitter may work its way out of the incision if the sutures come untied.

Some investigators finish the surgical procedure by treating the wound with an antiseptic solution, such as benzalkonium chloride. This can sometimes irritate the incision, however, and is considered unnecessary if the instruments are kept sterile and water is kept out of the incision (Hart and Summerfelt 1975; Winter 1983). Most investigators suggest observing the fish for several days after surgery to ensure that the healing process is underway, and to prevent the release of unhealthy fish.

b. <u>Signal reception and triangulation</u>. Once a transmitter-implanted fish is released into its natural environment, it is assumed that the fish will behave the same as it did prior to implantation. In most documented cases, this is a valid assumption. By locating the source of the radio signal, it is possible to monitor movement, activity patterns, and microhabitat utilization of the fish. This process typically proceeds in several steps: determining the general location of an instrumented fish, ascertaining its identity by the frequency and pulse rate of the signal, pinpointing its specific location in the stream, monitoring its location to determine activity and movement, and measuring microhabitat variables at its focal point.

Two different types of receiving gear are necessary to complete this sequence. The first type is a search receiver, or scanner, that is equipped with an omnidirectional whip antenna. These are designed to scan the range of transmitted frequencies (preferably all frequencies in use) to detect the presence of transmitters within range of the receiver. This range typically varies from a kilometer to less than a few hundred meters. A scanner may not be necessary if the subject species is sedentary, but is essential in the study of highly migratory species, which must sometimes be located from low flying aircraft.

The second type of receiver is a tracking or pinpointing receiver, used to determine the specific location of individual transmitters. These rely on a directional antenna to determine the direction of the signal source. There are several types of directional antennas: the T-shaped dipole, the H-shaped Adcock, the Yagi, which looks like a television antenna, and the loop, which may take the form of a circle, square, or diamond. The two most commonly used antennas in fisheries work have been the Yagi and the loop.

The most desirable attributes of the Yagi are its high gain (a ratio between signal strength received at the antenna to that delivered to the amplifier) and excellent directionality. Amlaner (1980) states that directional accuracy is ±0.5 degrees when the signal source is more than 20 wavelengths away. The most serious detractions of the Yagi are its size and relatively fragile construction. The elements of a half-wavelength Yagi, operating in the 30 mHz range, would be five meters long and one to two meters apart. Because of the vertical polarization of radio waves transmitted underwater, the elements of the Yagi would also be vertically mounted. Still, maneuvering such an apparatus through bankside vegetation, or rafting one through a white-water rapid, would be formidable. Thus, Yagi antennas are best suited for use on large, tranquil rivers where they can be boat mounted in such a way that their size and shape do not interfere with their use.

Loop antennas are more useful on smaller rivers, where the distance between transmitter and receiver is small, or where the receiver is likely to be shore-based. Loops are bidirectional antennas, receiving the maximum signal when either edge of the loop is pointed at the transmitter. The weakest signal, or null, occurs when the plane of the loop is perpendicular to the signal. It is generally agreed that the null is easier to detect and distinguish than is the peak signal (i.e., no noise vs. more noise). Therefore, the null is usually used to determine bearings to the signal source. These antennas do not attain the high directional accuracy of the Yagi. Amlaner (1980) lists a typical directional accuracy of ±5 degrees for loops, although Minor (1981) reported a mean angular error of about ±1 degree for a quarter wavelength loop antenna. Loops have several distinct advantages over multi-element antennas. Because the loop is not as efficient as the Yagi, the observer can get much closer to the transmitter without overloading the antenna (receiving such a strong signal that the null cannot be detected). In this respect, it is very desirable to have a tracking receiver on which the gain can be turned down (regardless of the antenna). The other desirable attributes of the loop are compactness and durability. Loops encased in plastic frames are virtually indestructible and ideal for applications in heavy streamside vegetation or other rough terrain.

Two or more directional antennas and receivers are used to pinpoint the location of a transmitter, using the principle of triangulation. Figure 21 shows a simplified representation of this principle as it might be applied in a microhabitat study. Two observers on shore (A and B) determine the direction to the signal source, from their respective positions. They direct a third crewmember, either in a boat or on foot, to the point of intersection between the two bearings. Once the position has been determined with some certainty, the crewmember in the boat either measures the microhabitat variables immediately, or drops a marker buoy for subsequent measurement.

c. Triangulation error. Unfortunately, the actual location of radio tagged fish is seldom as simple as the preceeding description of triangulation indicates. Several complications can arise that reduce the accuracy of the fix, and thus, the certainty that the indicated point of intersection is the actual location of the fish. Among the most important sources of triangulation error are: system error, movement error, and topographic error (MacDonald and Amlaner 1980).

System error refers to the precision with which the null bearing can be detected. This type of error is inherent in the equipment, but can be intensified if the equipment is not functioning at peak efficiency. Twisted or deformed antenna elements, pinched cables, and loose connections can all contribute to increased system error. Another common problem is overloading the antenna by getting too close to the transmitter, and not being able to sufficiently turn down the gain on the receiver. Some system error is present, however, even when the equipment is operating under optimal conditions.

The following example illustrates the system error associated with the use of loop antennas, both operating within the ±5 degree angular error cited by Amlaner (1980). The displacement error from each bearing increases with distance from the source. The displacement errors from two or more bearings form what is called an error polygon, as illustrated in Figure 22. In this case, both receivers are 50 m away from the signal source and the resulting error polygon has an area of about 46 m^2. The most likely location of the fish is where the two bearings intersect, but in reality, the fish could be anywhere within the area defined by the polygon. In streams exhibiting gradual change from one microhabitat to another, this may be close enough. Where the difference of a meter or two makes a big difference in the microhabitat conditions, however, this is a serious problem.

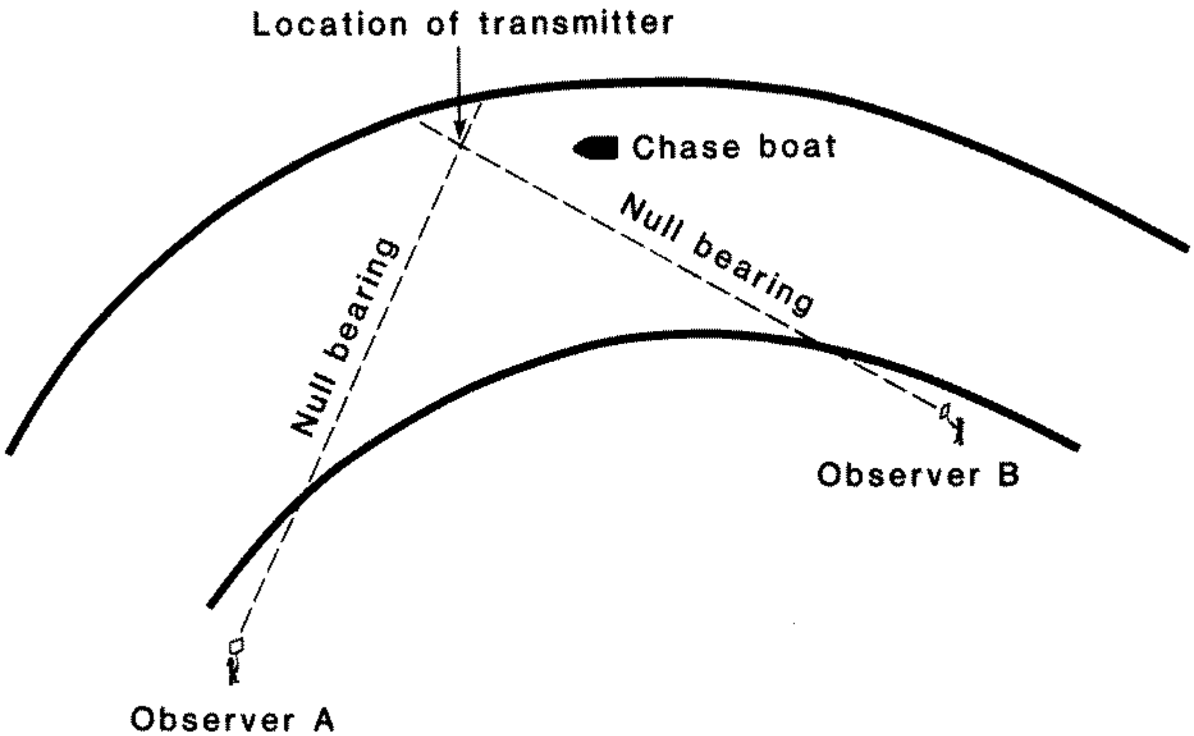


Figure 21. Principle of triangulation used to locate the position of a transmitter using two directional antennas.

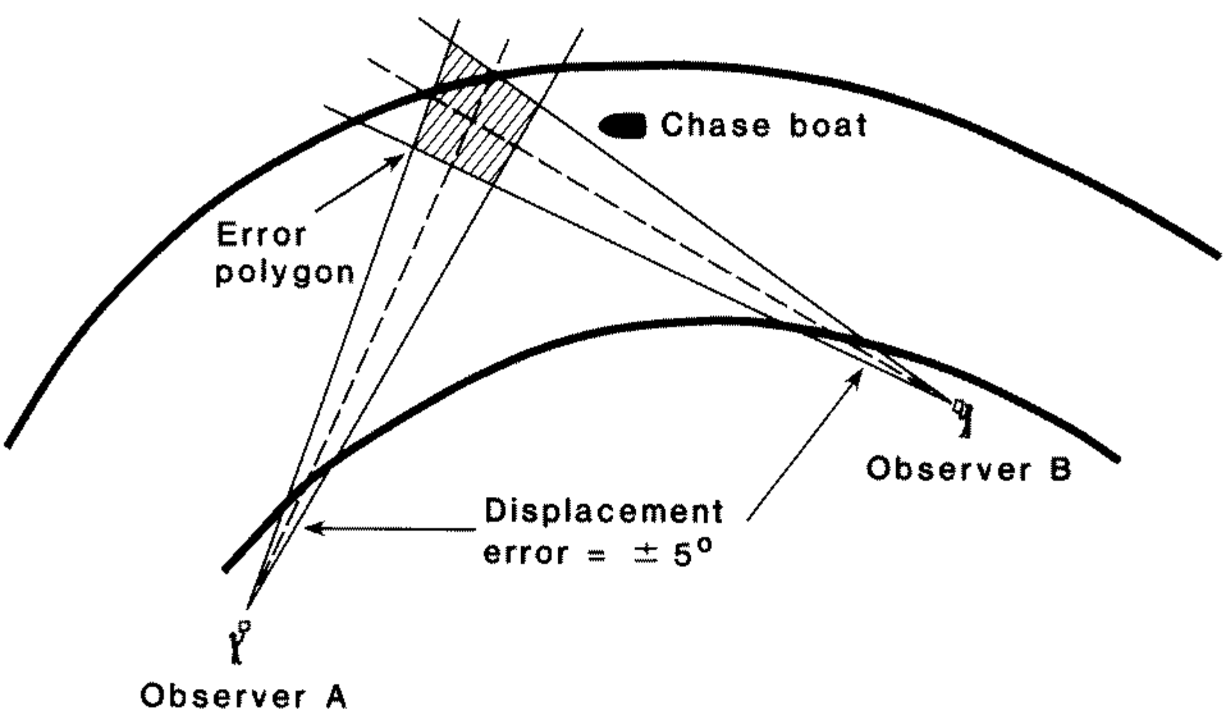


Figure 22. Triangulation error polygon resulting from ± 5° angular error in obtaining "true" null position.

There are several steps that the observers can take to reduce (but not eliminate) the size of the error polygon. The first is to try to approximate a 90 degree angle between pairs of bearings. As the intersecting bearings deviate from a right angle, the resulting polygon more resembles a parallelogram than a square, increasing the displacement error along the long axis. A second technique is to take more than two simultaneous bearings. The addition of a third observer (D) between A and B, as shown in Figure 23, reduces the area of the error polygon by about two-thirds. Third, try to get as close to the subject as possible without overloading the antenna. This will require some experimentation with a naked receiver (in water) to determine the best combination of distance and receiver gain setting.

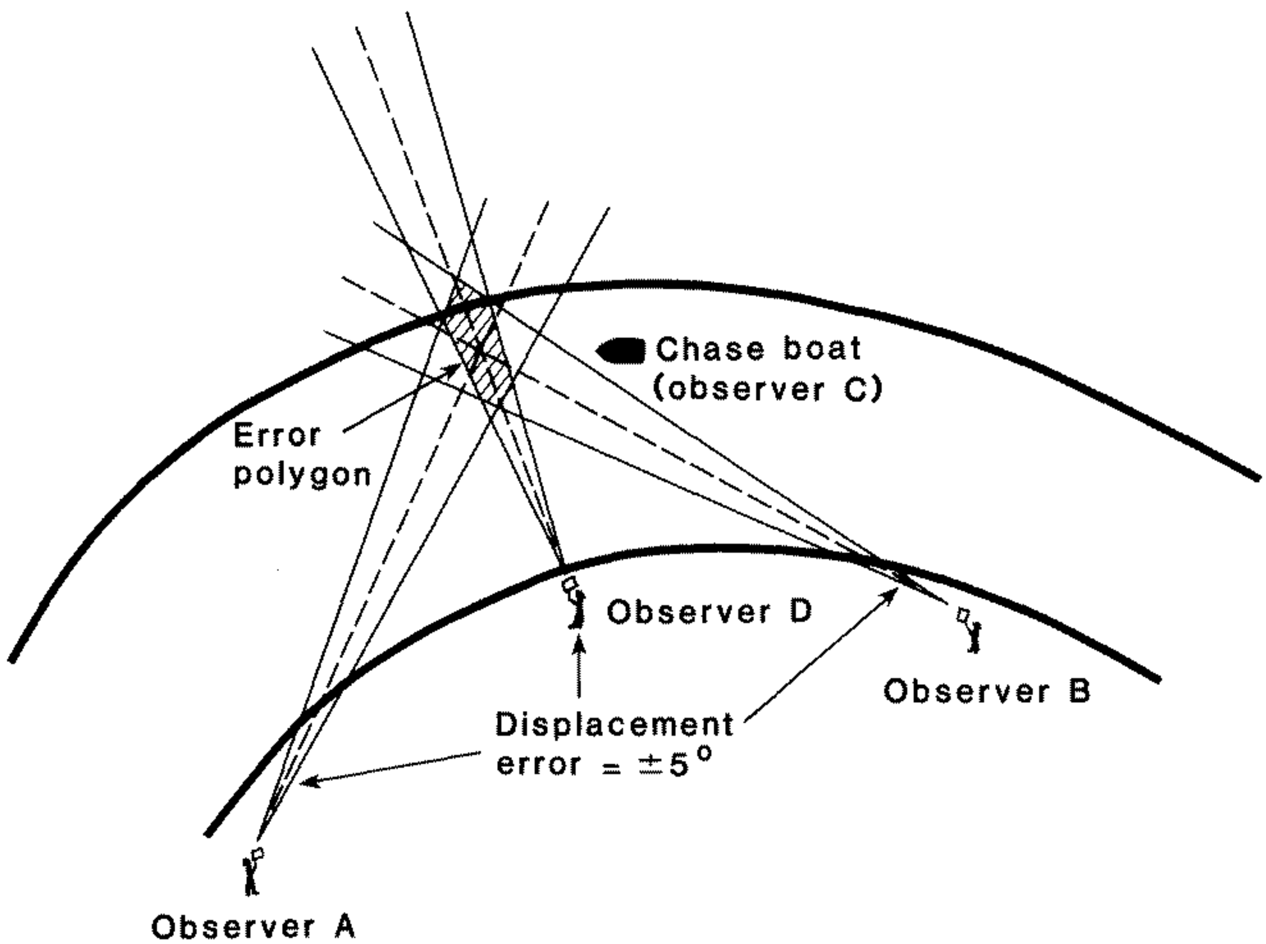


Figure 23. Reduction of triangulation error by taking additional bearings.

It should be obvious that triangulation can only determine the approximate location of the fish, not its absolute position. Sometimes this location can be determined within a meter or two of the absolute position, and sometimes only to the nearest 10 m. The adequacy of either level of accuracy is strictly a matter of the spatial heterogeneity of the environment. In terrestrial applications, triangulation is used to determine the general location of the animal, and actual sightings or other confirmatory evidence are generally considered necessary for an absolute fix. The same philosophy is appropriate for underwater applications. If close is good enough, then triangulation is good enough. If an absolute location is needed, do not rely solely on triangulation. In clear water, triangulation should provide a good enough location that an observer can make visual contact. This is the best form of confirmation, but not very likely since most radio telemetry studies are conducted in turbid water. One approach that may be useful in turbid water is the use of a separate receiver, either with no antenna or a very inefficient one (such as a length of coaxial cable or a paper clip), in the observer boat. The only time such a receiver will pick up a signal is when it is directly over the transmitter. This technique has been used to recover lost transmitters, and should work with implanted transmitters, provided the fish are not frightened by the presence of the boat overhead. A second approach is the use of an underwater aerial. Minor (1981) and Solomon (1982) describe the use of such aerials and state that they have been used to pinpoint transmitter locations within a few centimeters. To date, such devices have only been used to find lost transmitters, but they should also be appropriate for finding fish.

A second type of error mentioned by MacDonald and Amlaner (1980) is caused by movement of the subject between the time of its original fix and its final confirmation. This is probably a more serious problem in tracking migrating animals than it is in a criteria study, where the subject animal is monitored at a focal point. In fact, one of the advantages of radiotelemetry is the ability to distinguish between stationary and moving fish in turbid water. No other method allows this distinction.

Movement problems of a different sort may occur in a criteria study, as a result of random swimming within a home range or vertical movement of the fish through the water column. The cruising patterns of Sacramento squawfish described by Dettman (1977) and the random movements of trout in eddies (Gosse 1982) are two typical examples. Such movements, in themselves, are not much of a problem in a criteria study, because a series of measurements can be made to determine the average conditions utilized within the area. The larger problem is trying to distinguish between system error and movement. The best solution is to minimize the system error as much as possible, thereby ensuring the best definition of the home range area. It is also a good idea to monitor the position of successive bearings, repeated over a period of time. Tyus et al. (1984) suggest that highly mobile fish (in this case, Colorado squawfish) should be monitored for 30 minutes. Significant movement, such as from one pool to another, may indicate transitional behavior, whereas movement within the same pool would indicate some form of home range.

Vertical motion of the fish through the water column can present a different type of problem. Radio signal attenuation is directly related to the distance and conductivity of the water column through which the signal is propagated. This means that the strongest signals will occur when the fish is near the surface and the weakest when the fish is near the bottom in deep water. Tyus (1982) found an apparent reception limit at about three meters in water with a conductivity of about 800 µmhos. The complete loss of signal in deep water can become a source of bias, but only if the species shows a consistent preference for a benthic position in very deep water. Otherwise, the loss of signal becomes more of a nuisance than a real detriment. One problem that can result from such vertical movement is an apparent null, detected when the antenna is not correctly oriented. Receiver operators can check for a false null by rotating their antennas to sweep the signal area. If the amount of rotation is more than about 10 degrees before a signal is detected, either the gain on the receiver should be turned up or the observer should try to get closer.

The third source of triangulation error is due to the effects of local topography on the signal. A receiver can only detect the direction from which a radio signal reaches the antenna, and this is not always the direction from which it originated. Radio signals are reflected from canyon walls and vegetation, or reradiated from barbed wire fences and telephone lines. Any of these conditions can lead the observer into tracking a ricochet instead of the subject animal. MacDonald and Amlaner (1980) suggest several precautions that can minimize triangulation errors induced by topography. First, be aware of local topographic features and how they might affect radio signals. This refers to underwater topography as well as terrestrial features. Radio waves can reflect off submerged boulders as well as canyon walls, with the result that at least part of the signal may emerge from the water at some distance from the fish's location. Presumably, this false signal should not be as strong as the one emitted from the fish and should be detected as such. The second suggestion is to take several alternative bearings from different positions. Rank these bearings and derived signal locations according to the width of the null points and the variation of successive bearings that do not intersect at the same spot. Because the only signals that leave the water are those aimed upward, total signal loss can be expected if the subject moves under a ledge or undercut streambank.

In addition to these suggestions, it is usually worth the trouble to attempt confirmation of the absolute location. It may be possible to determine the location visually if the water is clear enough. Otherwise, it may be necessary to use some of the techniques described earlier for turbid water confirmations. Even assuming success with these techniques, it may be necessary to recover the subject from time to time simply to confirm that it is the original animal in its original condition. Gerardi (1983) reported tracking a snapping turtle that had eaten one of his tagged fish and swallowed the transmitter intact.

As mentioned in the introduction to this section, radiotelemetry for the development of habitat suitability criteria is still in its infancy. Gerardi (1983) found that the combination of movement error, topographical error, and background noise (mostly from CB radios) contributed to such large triangulation errors, that such data could not reliably be used to develop criteria. On the other hand, Tyus et al. (1984) used radiotelemetry to develop suitability criteria for Colorado squawfish in the turbid and isolated Green

River system of Utah and Colorado, concluding that this technique provided more representative and less biased data than other techniques they used.

It appears that the quality of data obtained by radiotelemetry depends on the transmitting/receiving system chosen, the terrain and topography of the environment, the behavior of the subject animal, and the skill and experience of the observer crew. It is therefore advisable to select the equipment that will perform best under the conditions anticipated to exist in the field. Then, the crew should practice with the equipment to learn how best to operate it and what its limitations are. Finally, it is important to set appropriate expectations on precision. If absolute locations are required, but cannot be obtained accurately enough by triangulation, then some confirmatory method will need to be employed.

4.1.4 Electrofishing

Electricity has been used extensively to capture fish since the early 1950's. Largely due to the efficiency and ease of use of electrofishing equipment, the technique has gained considerable popularity for applications such as population estimation, age and growth studies, and food habits studies. Electricity has also been used in microhabitat utilization studies, and has been variously praised or condemned for this application. The major complaint about electrofishing is that it can be a very disruptive method of determining fish positions in a stream. This complaint, however, is as often the result of misapplication or poor technique as it is of the equipment itself.

Complete discussions of the principles of electrical fields and the strengths and limitations of various types of gear are somewhat tangential to the subject of criteria development. Such knowledge is essential, however, for safe and effective use of electrofishing gear. Reynolds (1983), and citations therein, provide a comprehensive discussion of electrofishing theory and techniques. Such articles are highly recommended reading for anyone considering the use of electricity in a criteria study. Factors influencing the effectiveness of electrofishing gear and sampling tactics with this equipment are more germane topics for this paper, but a thorough understanding of the theory is also required.

a. Factors influencing the effectiveness of electrofishing. Data collected for fish frightened by or herded ahead of an electroshocker are worthless. This means that electrofishing will be more effective for some criteria studies than for others. Factors influencing this effectiveness include: characteristics of the equipment, characteristics of the species, skill of the sampling crew, and operating conditions.

The susceptibility of different species and sizes of fish plays a major role in determining the effectiveness of electrofishing in criteria studies. It is fairly common knowledge that electrofishing is size selective, being more effective on larger fish. Small fish do not have the body mass, and therefore, the electrical resistance, to create much of a voltage drop through the body. Consequently, smaller species (e.g., darters, small cyprinids, and sculpins) can often swim through an electric field with impunity, while their larger relatives experience a rather different effect. The same selectivity occurs among fry, juveniles, and adults of the same species.

Species characteristics other than size can also influence capture efficiency. Bony fishes are more susceptible to electricity than are cartilaginous species (Reynolds 1983). This factor is somewhat offset for paddlefish and sturgeons due to their large size. (Their large size may create other problems, however, such as increased sensitivity to the escape field and the logistics of capturing them with a dip-net.) Other species are vulnerable to electrical fields, but are difficult to capture. Ictalurids do not rise to the surface when shocked, but roll or tumble along the bottom of the stream. Because of this tendency, large members of the catfish family (channel, flathead, and blue catfish) are usually very difficult to capture with electrofishing gear. Trout may escape the electrical field by leaping out of the water and their capture is often a function of the netter's reflexes. Conversely, electrofishing may be the most effective method for reclusive or highly camoflauged species that are difficult for divers or surface observers to see.

Certain species, smallmouth and spotted bass in particular, seem to be especially sensitive to the presence of people and electrofishing gear. These fish tend to maintain a distance between themselves and the sampling crew and are virtually impossible to approach in clear water. Their avoidance behavior is apparently optical as they can be captured easily in murky water. One way to estimate the fish reaction to the approach of a sampling crew is to attach a bobber or needle float to a previously caught fish, and then release it. The effectiveness of the approach and subsequent sampling by an electrofishing crew can be evaluated by watching the movement of the float. If the float moves away from an approaching sampler, it is very likely that untagged fish are doing the same thing.

There are several modifications to gear or deployment that can reduce avoidance behavior. Sensitization to an approaching electrical field can be reduced by equipping the electrode circuitry with a "deadman switch." Originally developed as a safety feature, the deadman switch energizes the electrode only when the handler squeezes the switch contacts together. By selecting a prespecified sampling location, and energizing the electrode only after it has been placed in that location, fish do not detect the approach of the edge of the field and are more likely to be captured in place.

Some researchers have extended this concept to a fairly extreme, but effective, equipment modification: a prepositioned area shocker. Rather than moving about with an electrode and periodically energizing it, the electrode is positioned at a location to be sampled and left in place for a period of time. This allows the fish to resume their normal behavior following the disturbance of deploying the electrode. After a specified waiting period the electrode is energized with a shore-based generator. Specific details of this equipment are provided by Bain et al. (1985b).

Major drawbacks of prepositioned electrodes are the time required to complete a sample and the relative inefficiency of the system in some applications. It may take 5 to 15 minutes to position the electrode and another 10 to 15 minutes for the waiting period. Bain et al. (1985b) concluded that waiting more than 10 minutes did not affect the catch rate of their sampler, so until evidence suggests otherwise, this is the recommended time interval. In another study, Bain et al. (1982) found that only 18% of the set-ups produced captures of their target species. Such low efficiency, combined with a fairly large time per sample requirement, might discourage the use of this technique. Overall time efficiency can be improved significantly, however, by deploying several units in sequence.

Starting at the lower end of a stream reach, one crew would begin setting out electrodes and running the lead wires to the bank. After the specified waiting period, a second crew would follow, energizing electrodes, identifying and counting captured fish, and measuring the habitat characteristics. As soon as the first crew had finished laying out the last electrode, they could retrieve the first, following the data crew upstream. By leap-frogging upstream, it is likely that between 10 to 15 samples could be taken per hour. The most likely bottleneck with this procedure will probably be moving the generator from place to place. The process can be speeded up by using a small backpack generator or by floating the generator in a small boat. Battery powered units might also work, but will require numerous back-up batteries to complete a full day of sampling. Bain (Argonne National Laboratory, Chicago, Illinois; letter dated May 1985) advises that by using only two electrodes and a three-person crew, it was possible to sample three or four locations per hour, following this procedure.

The characteristics of the sampling environment can affect the performance of electrofishing gear and the crew using it. Some of the more important environmental factors include: depth, visibility, conductivity, temperature, velocity, and turbulence. Some of these variables simply affect the efficiency of capture, whereas others have a differential effect that can lead to data bias.

The efficiency of many electrofishing units is inversely proportional to the depth. Boom shockers are most affected because the electrodes are always near the surface. Therefore, current density is highest near the surface. To counteract this problem, fisheries workers have resorted to using more powerful generators and a fascinating variety of electrode designs. These have not overcome the basic problem, however. No matter how big the generator or the electrode, current density will always be strongest near the surface.

One solution may be a mobile electrode, which can be lowered to the bottom, energized, and retrieved slowly. Fish in the stun field exhibit galvanotaxis and swim toward the electrode as it is being pulled up. It is possible to draw fish to the surface from considerable depths using this technique. Mobile electrodes also exhibit a depth efficiency bias, because the farther fish must be drawn to the surface, the greater the chances of their escaping the field before they can be netted. This can be caused by retrieving the probe too rapidly or by the fish exhibiting narcosis instead of electrotaxis. The prepositioned area shocker is not designed to be used as a mobile electrode. Therefore, it probably shares many of the depth bias problems of boom shockers. Bain et al. (1985b) suggest that their area shocker should not be used in water exceeding wader depth (about 90 cm), although this restriction may apply more to the netter than to the field strength of the shocker. It might be possible to sample deeper water by netting from a boat, or to design a more rigid prepositioned area shocker that could be retrieved to the surface following energization. In any event, such sampling devices are probably limited to effective depths of two meters or less, with present technology.

Water clarity can also have an effect on electrofishing success. Most electrofishing devotees seem to prefer transluscent water with visibility of one to two meters. Opaque water creates sighting problems for the netters, who spot the fish only as they break the water surface. In very transparent water, the fish can detect the approach of the sampling crew too well and can avoid capture. Unless the fish learn to recognize a prepositioned electrode, water transparency is not a limitation with this device. When using any other electrofishing technique, however, it is probably better to work in opaque water than in clear. If the netter can see the fish, it is likely that the fish have already seen the netter.

Water conductivity also influences electrofishing effectiveness. If the conductivity is less than 100 µmho/cm^3, the fish may be unaffected by the electrical field until they touch an electrode, whereupon they are promptly electrocuted. At the other extreme, when conductivity exceeds about 500 µmho/cm^3, the current density may be too diffuse to affect the fish (Reynolds 1983). One technique sometimes used to overcome low conductivity is to place several salt blocks at the upstream end of the sampling section. This practice has had mixed results, but it would seem to work better in small streams. It is conceivable that it could also introduce a bias if the conductivity increased with proximity to the salt blocks, because habitats closer to the salt blocks would be sampled more effectively than those farther away.

Temperature is probably a greater concern in terms of overall effectiveness, and not as a source of bias. The principal influence of temperature on electrofishing effectiveness is its control over the metabolic rate of the fish. Fish are more active in warm water (if not too warm), may be more sensitive to the escape field, and are generally more elusive. Conversely, fish may become so lethargic in very cold water that they are not drawn to the electrode or simply sink to the bottom when shocked.

A common problem of sampling in swift and turbulent water is that these conditions are more hazardous than slow, flat water. A crew that feels unsafe in any situation will be concentrating more on staying alive than on collecting quality data. Several precautions should be taken to enhance crew performance under dangerous conditions. The first is to make sure the equipment is in top working condition and equipped with good safety accessories. Reynolds (1983) provides an excellent discussion of safety procedures and equipment. Another precaution is to thoroughly scout the area before sampling. This should be done anyway, to prepare the sampling design and select the sampling locations. Additionally, places where special care will be needed should be identified and inspected before attempting a sampling run. In this context, moving upstream under power is usually safer than drifting downstream. A third precaution is to allow the crew time to gain experience with the area and conditions to be sampled. This may mean devoting time for one or two dry runs up or down the reach without sampling. Dry runs may initially seem like a waste of time, but they will ultimately pay off in increased efficiency, better data, and, most importantly, a safer operation.

High velocity, in areas where it does not pose a safety hazard, can still cause difficulties in data collection. The most serious problem occurs when the current carries a mobile electrode away from the boat, so the crew cannot determine the original positions of captured fish. This problem is less serious when the boat is drifted downstream because the electrode will drift at about the same rate as the boat. Mounting the electrode on a long, rigid pole, such as a dip-net handle, can alleviate the problem of electrode drift. Reynolds (1983) discourages this practice for safety reasons, mainly because of the possibility of hitting someone with the pole or electrode. Stationing the operator of a long-handled electrode on the foredeck with the netter and inserting the electrode vertically into the water, however, should be no more cumbersome than two netters.

Water velocity and turbulence require good reflexes on the part of the netters to avoid data bias. Fish are much easier to capture in slow, flat water; they are easier to see, and they are not carried out of the electrical field by water movement. A netter has only a split second to capture a stunned fish in rapidly flowing water. If not captured immediately, the stunned fish will drift out of the electrical field, recover, and dart away in a matter of seconds. This tendency can bias the data toward lower velocities, where capture efficiency is higher. Evaluation of potential bias can be made by calculating an efficiency rating based on the ratio of captures to sightings (e.g., 50 fish sighted, 30 captured, efficiency = 60%), stratified between slow and fast water samples. Unfortunately, the same determination cannot be made for depth-influenced efficiency losses because actual sightings are needed, and these may not be possible in deep water.

b. Sampling tactics. The use of electrofishing gear in a criteria study requires several procedural modifications from its use in general fisheries investigations. The most fundamental of these is a modification of sampling objectives. The objective of a population study is to capture sufficient numbers of fish to estimate the size of the population. How they are captured is of little significance. The objective of a criteria study is to sample specific microhabitat areas to learn which species are using them. How fish are captured is paramount. Although quantity is important for obtaining the necessary sample size, it is not nearly as important as the quality of each capture. The difference between these two objectives is clear, but old habits are sometimes hard to break.

The second significant procedural step is to select an appropriate sampling design. Random walk designs are most compatible with mobile electrodes, whereas proportional, random, or stratified systematic schemes are best with prepositioned electrodes. Selected locations should be marked on a plan map of the reach, and new locations should be selected for each day of repetitive sampling in the same area. The number of samples (i.e., each time the electrode is energized) that can be made in a day depends on the size of the river, proportion of samples containing the target species, and whether or not each sample location is measured to determine availability of microhabitats (see Section 4.2). Generally, 100 to 200 individual locations can be sampled in a day with a mobile electrode and 25 to 50 with prepositioned electrodes.

Once the locations have been determined for a day's sampling, the tactics for approaching each location should be planned by the crew. Certain locations, such as undercut banks and debris jams, will require a special plan of attack, but there are a few general guidelines that can contribute to a stealthy approach:

1. Energize the electrode only at the sample location. This helps prevent sensitization to the field, and also preserves battery life.

2. When approaching a sampling location, shut off the boat motor and avoid making sudden, loud noises. Sound moves through the water about four times faster than it does in air and can carry for considerable distances. Certain noises, such as creaking oarlocks or objects striking the hull or gunwales of a boat seem to transmit exceptionally well under water.

3. Electrofishing, in a criteria study, should probably be confined to those situations where visibility is too poor to use an observational technique. If the water is clear enough to see the fish, they can also see the approaching sampling crew.

4. When possible, sample in an upstream direction so the fish can be approached from the rear. When wading, dip-netters should stay abreast of, or lag slightly behind, the electrode. Do not lag too far behind or the fish will recover before they can be netted. It is also possible to approach from downstream in a boat, but there is a trade-off between disturbing the fish with motor noise and disturbing them visually by drifting downstream. If visibility is poor, drifting downstream is better, provided it can be done safely. An electric trolling motor may allow a silent upstream approach if the current is not too strong.

5. Avoid consecutive sampling of locations in close proximity to one another. Fish in the escape zone of one sample will often flee into or through a nearby location, triggering a brief underwater stampede. Flight distance probably varies by species, but if sampling follows the guidelines listed above, it will probably be less than 50 meters. The advantage of the systematic random walk sampling design is that it avoids this problem.

6. Allow at least two hours between samples taken at the same or adjacent locations. Cross and Stott (1975) found that learned avoidance of electrical fields is most pervasive if the same area is shocked two or more times at intervals of less than two hours. Allowing a day between samples at the same location would be better.

4.1.5. Area Samplers

An area sampler is a device designed to collect fish or macroinvertebrates from a specified area, rather than a single, well-defined location. One of the best known is the Surber sampler, used to collect quantitative samples of benthic macroinvertebrates. All area samplers should be operated following the same principles as a Surber sampler. That is the area sampled should be homogeneous so a microhabitat measurement taken at the center of the area is representative of the entire area. This concept seems reasonable when the area sampled is only one-tenth of a square meter. The same concept, however, should hold for much larger samplers in a criteria study. Among the larger samplers that can meet this operating criterion are a variety of seines, nets, active traps, and benthic samplers, as well as certain types of explosives.

The use of some area samplers requires a considerable amount of site preparation. A complete habitat map should be prepared, showing lateral and longitudinal distributions of different habitat features and underwater hazards that may snag a net. Cross sections should also be surveyed or sounded acoustically to make sure that homogeneity is maintained in each of the samples. The cross-sectional information can be incorporated as contours on the plan map as shown in Figure 24.

It is important to minimize habitat variability within each sample because only one depth, velocity, substrate type, and cover index can be entered into the data base for each sample location. The suggested approach for data collection is to take six to eight measurements within the area sampled and compute the mean and standard deviation for each of the variables. If the coefficient of variation exceeds 25%, the sample should probably not be used in the data base. This may result in discarding a considerable amount of data, and a decision must be made regarding the acceptable level of precision. The alternatives are to confine sampling to areas of known homogeneity or to sample smaller areas.

Most area samplers should be considered as specialty equipment for sampling fish or stream areas that cannot be effectively fished using other methods. The limitations of these devices are such that their use alone is bound to introduce bias: they cover relatively large areas, increase the potential for internal variance for each sample, and cannot be used very successfully around cover or in moderate currents. The strength of these samplers is that they are fairly effective for schooling fish and in deep water.

a. Seines and haul nets. The use of these devices to develop habitat suitability criteria has been criticized for many of the same reasons as electrofishing. The same characteristics that make seines and haul nets (e.g., drifted trammel or gill nets) effective fish catchers, make them less desirable as equipment used in a criteria study. Nets and seines are known to be size selective, a desirable attribute only if certain sizes of fish are not wanted in a sample. This is rarely the case in a criteria study, but a problem that can be partially overcome by using a variety of mesh sizes. A more serious problem is the traditional method of net deployment. It is virtually impossible to obtain a homogeneous sample using the standard sweep or seine haul procedure. Whether both ends of the seine are drifted, or one end made stationary, it is customary to finish the drift by hauling the net toward and onto the shore. This practice inevitably includes shallow water areas in the sample zone, whether intended or not. The only way to avoid this bias is to alter the deployment procedure.

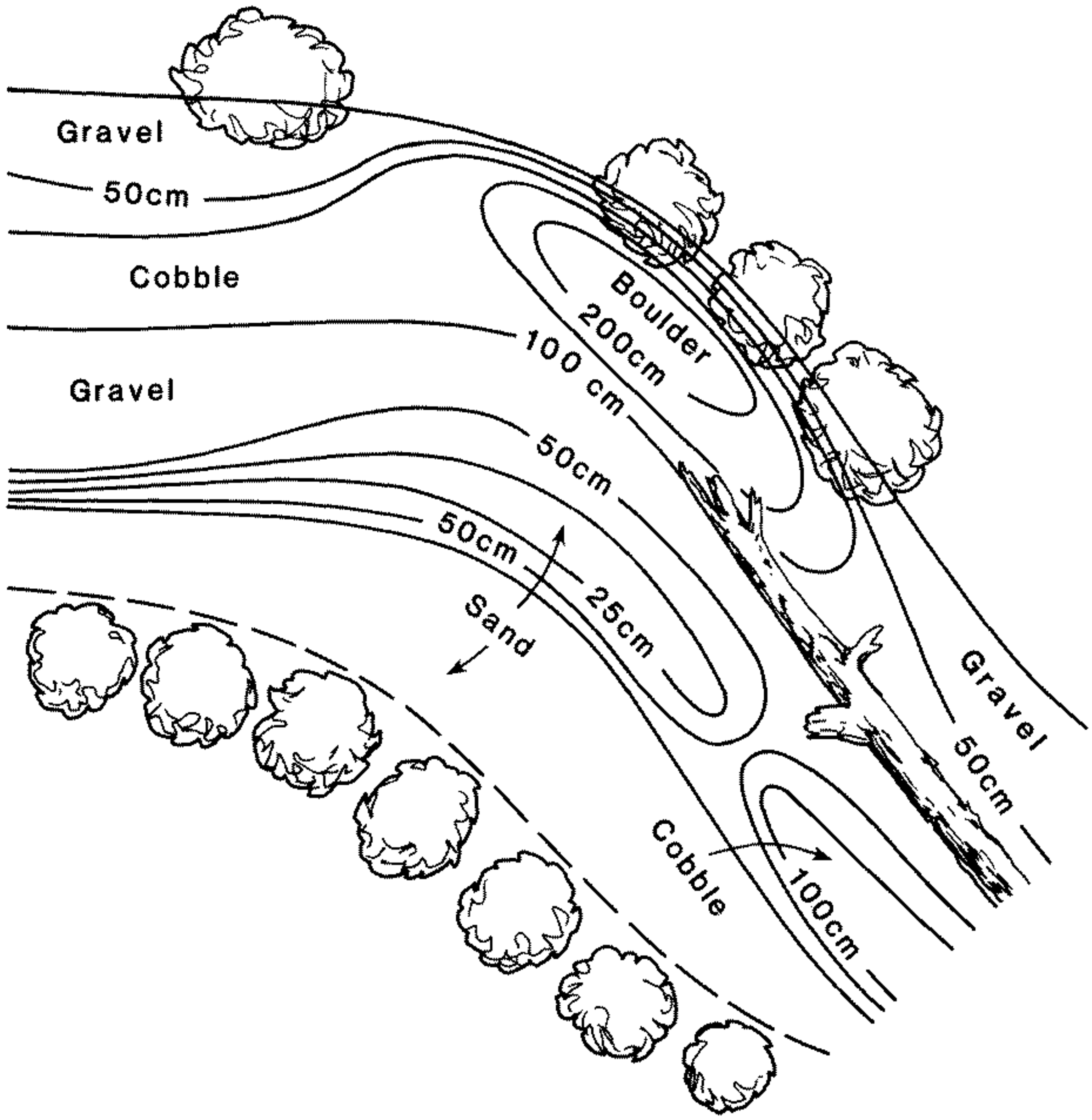


Figure 24. Example of a contour map used to maintain sample homogeneity using seines, cast nets, or lift nets.

The most promising strategy may be to use a purse seine deployment. The seine is first set out in a straight line, then one or both ends are pulled in an arc to encircle an area. The encircling maneuver can be used to isolate a particular area, but retrieval of fish is difficult with a standard seine. One solution is to use a purse seine to begin with, since the maneuver was originally devised for this gear. Understandably, a commercial purse seine is much too large for this application. Hunter et al. (1966) give design specifications for a miniature purse seine to be used in research applications. Their design may still be too large for habitat suitability investigations, but could be scaled down.

An alternative to using a purse seine might be an entanglement net, such as a trammel net. Following the encirclement, both ends of the net would be hauled in simultaneously. As the area inside the net becomes ever smaller, fish darting away from the center would entangle themselves in the pockets of the trammel net. Because this approach will only work if the fish are caught in the netting, standard beach seines probably cannot be used. Gill nets might work, but these can cause considerable injury to the fish. To prevent fish from escaping through the opening between the two free ends, it may be possible to use an electrofishing electrode to keep them in the net.

It is undoubtedly fortunate that criteria studies require small seines, because the encircling maneuver described above is not easy in a current. One method of deployment involves extending the seine perpendicular to the current and moving both ends upstream until they are close enough to be pursed or hauled in (Figure 25a). This approach presents two problems. First, the open seine acts as a parachute (more so as it fills with fish or detached algae) as it is deployed. Second, the current will tend to collapse the upstream side of the enclosure as the net is being closed. This approach is obviously limited by current velocity and relies on power. A second approach, illustrated in Figure 25b, relies on speed. Instead of deploying the seine perpendicular to the current, it is deployed parallel to it. The downstream end remains fixed, while the upstream end is towed in a downstream direction around the area to be sampled. The critical procedure is to close off the circle and purse the seine before the upstream side collapses significantly (i.e., before the float line drifts over half the distance to the downstream end). It is likely that power boats will be needed wherever there is an appreciable current, using either approach.

Seines are best applied to pelagic or mid-column schooling species, but ineffective for benthic species. Their use is most effective in deep, slow pools and backwaters where the efficiencies of other techniques are reduced by excessive depth or turbidity. Seining, like electrofishing, becomes more effective in turbid water because it is easier to avoid detection.

Unlike electrofishing, seines are not very useful around cover objects, such as boulders or tree snags. Tangling or snagging a seine can be extremely dangerous, because there is a good chance of swamping a boat or entangling a crew member. There is also a good chance that the seine will be damaged or lost, so it is simply good policy not to deploy one near these objects. As a safety procedure, the ends of the floatline should be looped around boat cleats and belayed by hand, rather than tied. This way, the boat can be freed from the net immediately, if necessary.

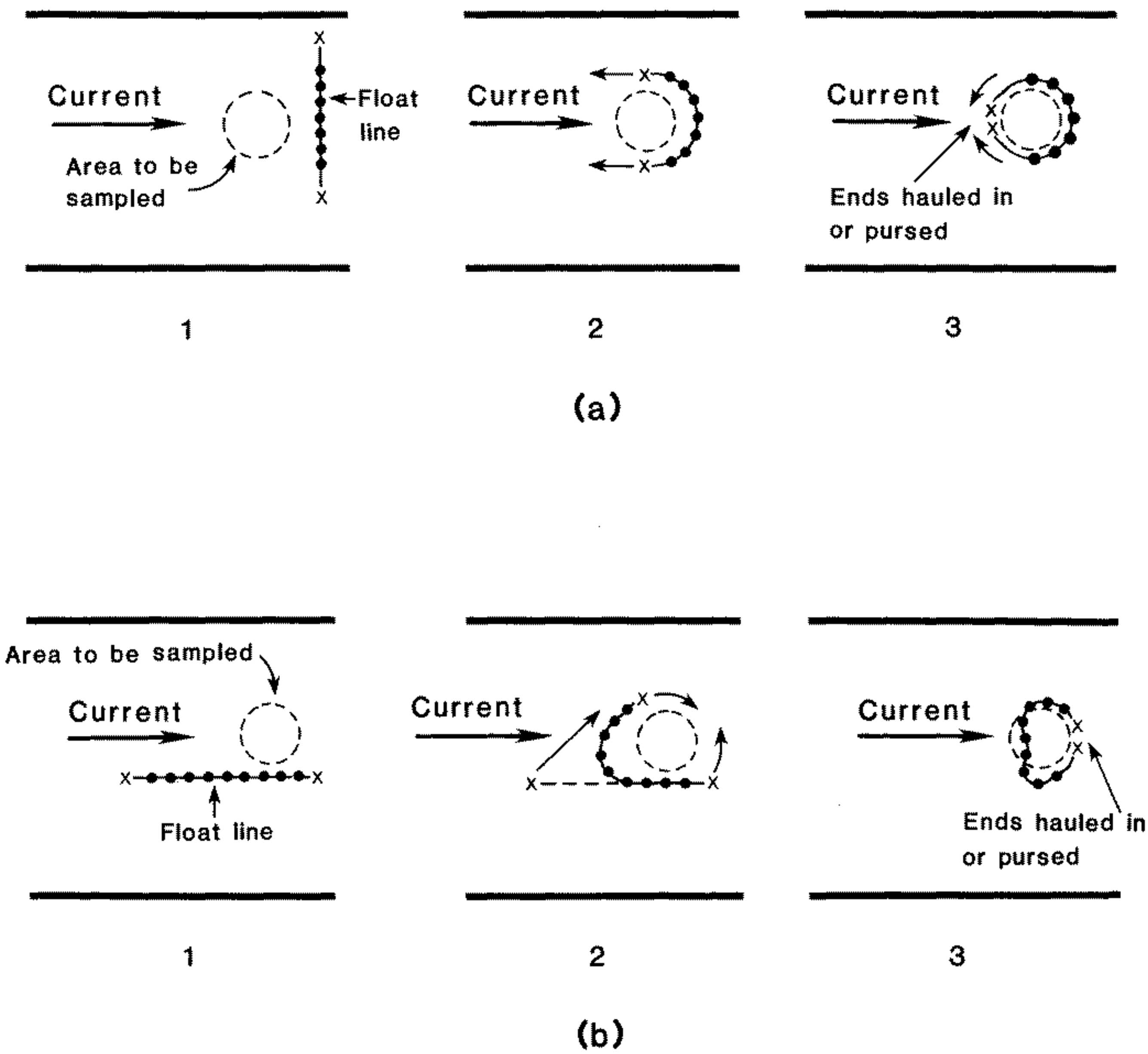


Figure 25. Seine deployment strategies to encircle a relatively small, homogenous area of stream: (a) upstream deployment, (b) downstream deployment.

b. Throw/lift nets. Throw nets and lift nets may have wider application than the limited habitats that can be seined. Both are relatively small, typically with an opening of less than two or three meters in diameter. Throw nets are conical shaped nets, cast or dropped over an area to be sampled. The open end of the net is larger than the leadline end. When retrieved, fish are trapped in mesh pockets formed between the lead line and open end. Throw nets can also be designed with a purse-type closure (Hayes 1983). Considerable skill is required to throw a cast net so that the net opens properly as it sinks to the bottom. Making such precise casts takes a considerable amount of practice.

The lift net works on nearly the opposite principle as the cast net and requires comparatively little skill. The lift net is simply a small net attached to a rigid frame that holds it open. The net is lowered to the bottom, left for a period of time to allow the fish to restore normal activity, and retrieved through the water column. The efficiency of lift nets is affected by visibility and the rate of retrieval: generally, the faster they are retrieved, the better. In clear water, fish are alarmed by the movement the instant the net comes off the streambed. Furthermore, they seem to be as wary of slow, steady motion as they are of sudden movement. Conceivably, a large net may have a better capture efficiency, but the retrieval rate may be reduced because of the increased resistance. In relatively shallow streams, a tripod and counterweight can be used to assist in lifting the net (Figure 26). In deep water, it may be necessary to use a boat-mounted, high speed winch to lift the net. The net may also be equipped with a remotely activated flotation collar, using buoyancy to provide the lift.

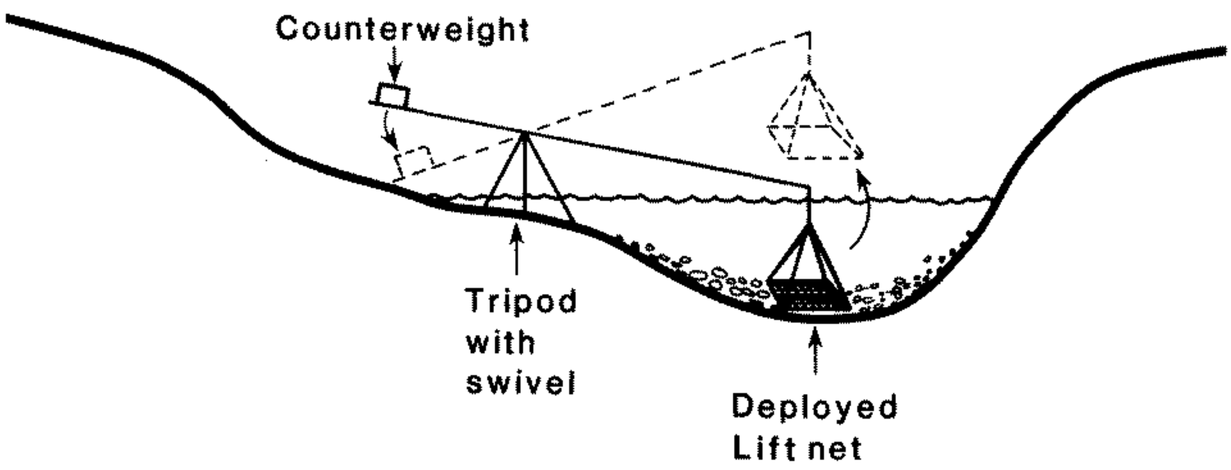


Figure 26. Tripod and counterweight to assist manual raising of a lift net in a shallow stream.

Another important consideration of lift nets is to make them as neutral to the fish as possible. For example, nets that have come in contact with salmon roe will be attractive to small salmon; nets that were previously used in a rotenone application will be repulsive to nearly everything. Dirty or camoflauged netting material is also more neutral than clean, white netting.

d. Explosives. The use of explosives for criteria studies provokes as much controversy as electrofishing. Anyone who has spent more than a year or two in fisheries science has heard at least one story about someone using a big enough charge to blast all the water out of a channel. True or not, such legends damage the credibility of using explosives to determine habitat utilization. Explosives can be used effectively to this end, however, provided that their limitations are understood, as is true of other techniques.

One advantage of explosives is that they can be deployed at predetermined locations in much the same way as the prepositioned area shocker (Section 4.1.4). This allows discrete areas to be sampled with the element of surprise entirely on the side of the investigator. The size of the sampling area can also be varied by altering the size of the charge. Explosives are effective in much deeper water than electrofishing and are equally effective for all sizes and species of fish. Many of the biases associated with depth and size selectivity, common to electrofishing and seines, are not inherent problems with explosives.

Despite these desirable attributes, there are several negative aspects of explosives that may detract from their usefulness in criteria studies. First, explosives are potentially very dangerous. Primacord is generally considered to be one of the safest types of explosives available. It is relatively insensitive to heat, impact, or static electricity, so premature or accidental detonation is unlikely. Platts et al. (1983) provide a brief description of primacord use, and recommend "Primacord Detonating Fuse - What It Is and How To Use It" by Ensign-Bickford Co., Simsbury, Connecticut, as required reading prior to using this method. The company has recently updated this manuscript, which is now available as the "Primacord Handbook" (Ensign-Bickford Co., 660 Hopmeadow Street, Simsbury, CT 06070). Many agencies may also require specialized training and certification in the use of explosives to comply with OSHA regulations. Whether such training is required or not, it is still advisable, especially for novices.

Unlike electrofishing, the majority of fish affected by an explosive device are not stunned; they are killed. Platts et al. (1983) state that primacord can achieve 100% mortality of fish within three to five meters of the cord. This fact will preclude the use of explosives in many criteria studies: those involving rare and endangered species or important sport and game species, in particular. There may also be somewhat of a philosophical paradox of ostensibly saving fish by blowing them up. Philosophy aside, the concussion from an explosive device ruptures the air bladders of the fish, which causes them to sink to the streambed. This can make recovery and accurate enumeration difficult, if not impossible, under conditions of reduced visibility.

If explosives are to be used in a criteria study, several procedures should be followed to increase the effectiveness of the method. Sampling should be confined to relatively small areas, generally in the range of 2.5 to 10 square meters. Remember that sample homogeneity is important, and the smaller the sample area, the more homogeneous it is likely to be. The most effective deployment patterns for sampling small areas are either circular or parallel primacord sets. These patterns focus the concussion toward the center of the circle or between two parallel lines of cord.

The appropriate length and number of strands of cord are determined and laid out to outline the sample area, and all microhabitat variables within the sample area measured. It is important to make the microhabitat measurements prior to detonation, as the explosion will probably alter the streambed configuration and local hydraulic characteristics. A detonating device is then attached to the end of the cord. Platts et al. (1983) recommend electric blasting caps as detonators, but these should not be attached until after the measurements have been made. It is unlikely that the cord would detonate while measurements were being made, but there is no point in taking chances.

Depending on the depth and turbidity, it may be necessary to set up a block net downstream of the sample area. If the water is shallow and clear enough to recover the fish with a dip net, the block net will not be needed. The mesh size for all nets should be small enough to recover young of the year fish. Block nets should be set up no closer than two meters of the sample area to prevent damage to the net (Platts et al. 1983). A block net frame shaped like a hockey goal (Figure 27) may facilitate quick set-ups, although it may be somewhat cumbersome to maneuver through shoreline vegetation and will take up more space in a boat.

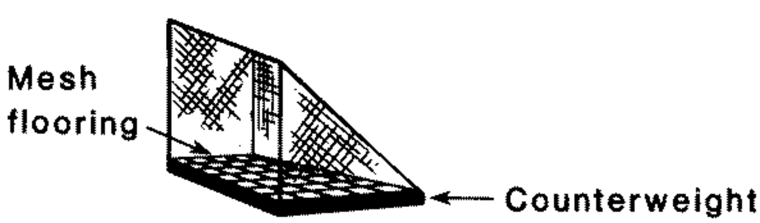


Figure 27. Prototype block net used to recover fish from an area sampled with primacord.

After the microhabitat measurements have been completed and the block net set up (if needed), the blasting cap can be attached to the end of the cord with electrician's tape. The long electrical wires from the cap are run to a safe location downstream or on the bank. The sampling area should then be left undisturbed for at least 10 minutes to allow the fish to resume normal activity. At the end of the waiting period, the charge is detonated. If water clarity allows, dead fish should be recovered immediately by dip net. In turbid water, the block net should not be retrieved until the water that was in the sample area at the time of the explosion has passed through the net. In fact, several water exchanges may be desirable.

d. Benthic samplers. These samplers were originally designed for collecting benthic macroinvertebrates, but are also quite effective in collecting small fish inhabiting relatively swift water, such as darters, longnose dace, and young channel catfish. The Hess or Water's round bottom sampler consists of a cylindrical metal frame, surrounded by a fine mesh. A sock-shaped bunt extends from the approximate middle of one side of the sampler (Figure 28). These samplers are operated by jamming the metal base into the streambed and stirring the enclosed substrate. A polyurethane foam cuff can be attached to the base for sampling coarse substrates. Organisms dislodged by the mechanical disturbance of the bed are swept (sometimes chased) into the bunt. Because the samplers are only about 50 cm tall, they have typically been confined to shallow water, but Gosse (1982) used a similar device with a mesh closure over the top to sample macroinvertebrates in deep water. The device was operated by a diver in much the same manner as the original Hess sampler would be operated by a person wading.

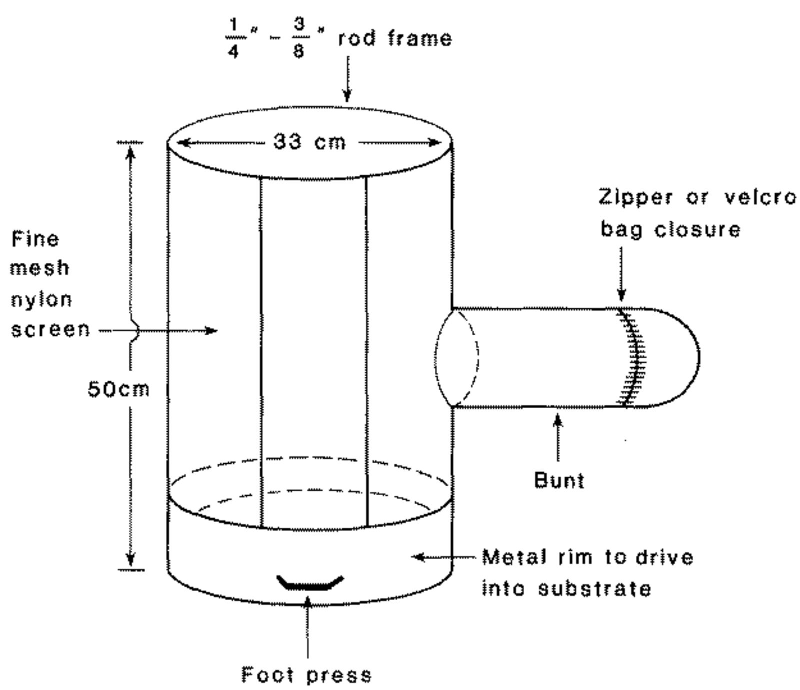


Figure 28. Design features of the Hess or Water's round bottom sampler.

The advantages of these samplers are that they can be used in fast water and are effective in capturing small, benthic fish that easily evade electrofishing and large mesh nets. Their principal disadvantage is that they are somewhat cumbersome, particularly when operated by a diver. When used in the specialized environments for which they were intended, however, this is more of an inconvenience than a true limitation. The only major limitation is that they cannot be used very effectively to capture fish larger than about 10 cm.

4.1.6 Laboratory Streams

Laboratory streams and flumes are rarely large enough to conduct meaningful habitat studies on adult and juvenile fish. They are, however, very useful in determining environmental influences on incubation and hatching success. Incubation is not an elective behavior, so the controlled conditions of the laboratory stream make it possible to determine the effects of individual variables on the survival and maturation of fish eggs. Hatching success is a function of the history of the habitat between the time the eggs are laid and when they hatch. This may be impossible to control in a natural stream, so grab samples measured at a particular time during the incubation period may be totally unrelated to the survival of the eggs. Therefore, the development of incubation criteria from laboratory data may be superior to studies in natural streams.

The basic elements of a laboratory stream are a trough of some sort, through which water can flow, and a water supply and control system. The trough may be either fixed gradient or variable gradient. Both can be used in incubation studies, but the variable gradient type is preferred because of the wider range of velocities that can be evaluated. Water supply and control systems are categorized as flow-through or recirculating. Flow-through systems require a constant source of good quality water, but this aspect makes them somewhat more similar to natural stream conditions. Eggs are exposed to the same temperature, dissolved oxygen, and general water quality conditions they would experience in the wild. Temperature and water quality can be controlled and experimentally modified in recirculating systems.

Experiments are performed by introducing various substrate mixtures into the trough, altering the bed configuration by creating small pools and riffles, and adjusting flow rates. Viable eggs are buried or broadcast over the substrate in approximately the same manner that they would have been in the wild. Then, a particular set of environmental conditions is held constant until fry emergence, at which time the percent of surviving eggs can be calculated.

Details for constructing artificial streams, or components thereof, can be found in Burrows and Combs (1968), McCrimmon and Berst (1966), Scherer et al. (1977), and Scott and Allard (1985) and in references cited in their articles. Before attempting to build a laboratory stream, it is wise to try to find an existing facility that could be used instead. Raceways at fish hatcheries are not always in use, and may be vacant long enough to conduct an incubation experiment. The advantages of using hatchery facilities is that equipment for dechlorinating and purifying recirculated water are usually already in place, so these components do not need to be retrofitted into the facility. The limitation of raceways is that they have fixed gradients. About the only way to alter the velocity in a raceway is by changing the discharge. This aspect will definitely limit the range of velocities that can be examined.

Another possibility exists for researchers near universities with large civil engineering colleges. Most of these institutions have one or more flumes that are used for hydraulics research. It may be possible to conduct incubation experiments at these facilities when they are not in use. Many of these flumes are equipped with hydraulic lifts, so the gradient can be altered. Hydraulic flumes are nearly always closed, recirculating systems with no temperature, dechlorination, or water purification capability so these components would need to be incorporated before a flume could be used in a biological study. Scott and Allard (1985) give details for a four-tank recirculation system with a cycling rate of 130 L/minute (34 gal/min). This is probably too small to be used with a three-meter flume, but the size could be scaled up to handle larger flow rates.

4.1.7 Special Sampling Problems: Spawning and Larval Fish

a. Spawning. The collection of data related to spawning is either very easy or very difficult, with relatively little middle ground between these extremes. Spawning is a specific, short-duration activity that is best detected by direct observation. The determination of spawning habitat must be based on circumstantial evidence, if the act cannot be observed. When the fish must be sampled, rather than observed, their activity must be inferred from deviations from normal behavior, and additional evidence is often necessary.

The presence of large aggregations of fish during the spawning season may indicate imminent or ongoing spawning activity. Such aggregations may be misleading, however, if only one sex is present in the group. A large group of males may congregate in portions of streams that may be nothing more than convenient places to await the arrival of females. The absence of females serves as circumstantial evidence that spawning is not happening at the time of sampling. While this type of evidence is useful, it is little help in developing criteria. More conclusive evidence of spawning would be the capture of a female and one or more males in the same, small sample. The evidence is even more conclusive if both are running milt and eggs at the time.

The behavior patterns and movements of fish can also serve as indicators of spawning activity, if these movements can be detected. For example, spawning carp create a visible, noisy commotion on the surface, often in rather large aggregations. Few species are as boisterous as carp, but many exhibit deviations from normal behavior during spawning. Tyus (U.S. Fish and Wildlife Service, Vernal, Utah; pers. comm.) used radiotelemetry to monitor the movements of Colorado squawfish during the spawning season. The squawfish tended to hold in large pools adjacent to deep, swift riffles. At random intervals, the squawfish would exhibit quick and erratic movement into the riffle, where they remained for a short period of time (usually less than 15 minutes). Following this interlude, the fish returned to the pool. Although other explanations of this behavior are possible, similar behavior has been observed of other squawfish species during spawning. Therefore, there is strong, but not conclusive, evidence that the squawfish were indeed spawning.

Uncharacteristic behavior is not confined to local movement or surface displays. Long-range movements and large aggregations of fish, especially if they are normally solitary, can indicate potential spawning activity. Some species exhibit both traits. A large group of paddlefish (usually solitary) collecting over a gravel bar in a river (they normally reside in slow water or lakes) is a strong indication that spawning is impending, if not ongoing. Unfortunately, these examples of behavioral deviation are not conclusive evidence of spawning. Although some are stronger indicators than others, they all serve mainly to alert the investigator to the likelihood of spawning. More evidence is usually needed to confirm these suspicions.

The most conclusive evidence is undoubtedly the collection of fertilized eggs at suspected spawning locations. Adhesive eggs of demersal spawners will usually be found very near the original spawning location. The eggs of pelagic spawners, even if they sink to the bottom, may be found at considerable distances downstream from the spawning area. In either case, the combination of suspected spawning activity followed by the immediate collection of eggs is valuable for two reasons. First, the spawning area itself can be confirmed quite precisely. Second, identification is easier because the adults can be identified rather than the eggs. Identification of fish eggs can be difficult and requires the services of an embryologist. Sometimes, the eggs must be reared until hatching to determine their identity.

Discussions of the sampling gear and techniques for collecting fish eggs and young fish can be found in Bagenal and Nellen (1980) and in Snyder (1983). The simplest method of collecting adhesive eggs from a benthic spawner is to remove small samples of the substrate and inspect them for eggs. If the eggs are incorporated within the substrate matrix, they can often be separated by placing the substrate sample in a 7% solution of hydrogen peroxide. The peroxide causes all organic matter in the mixture to float to the surface, where it can be skimmed off (Guy 1969). If the eggs are strongly attached to the rocks, the peroxide may not remove them, but the rest of the organic matter can be removed. Hydrogen peroxide is a fairly strong oxidizing agent and will eventually destroy the eggs. Therefore, mixtures should be examined within an hour or less of treatment. This method of separation cannot be used if the identity of the eggs is unknown, because the eggs will decompose before they can be identified. Snyder (1983) recommends that eggs requiring identification should be fixed and preserved in the field, transported to the laboratory, and sorted by hand. A 5% to 10% formalin solution is recommended to fix the sample (killing and stabilizing tissues quickly to maintain anatomical form and structure) and a 3% to 5% formalin solution to preserve it. Alcohol should not be used because it dehydrates and distorts the specimens.

Sampling for nonadhesive eggs, broadcast by pelagic or demersal spawners, can be more complicated. The problem is not so much one of sampling the eggs as it is of determining where they originated. This problem is reduced appreciably if the investigator can reasonably isolate the probable spawning area. Suspended fish eggs can be sampled by a variety of methods, but the most effective is the use of stationary drift samplers as described by Dovel (1964) or Snyder (1983). These are banks of plankton nets distributed laterally across the river (or vertically through the water column), suspended from an anchor rope. The advantages of these systems are that they allow the collection of control samples, they can be deployed fairly quickly, and they allow the simultaneous sampling of several areas.

Control samples are helpful in the isolation of probable spawning locations. These are collected by setting a line of drift samplers upstream from the suspected spawning area. The "experimental" samplers are deployed downstream. Control samples help distinguish locally produced eggs from background drift. If the spawning area is between the control and experimental samplers, the concentration of eggs should be higher in the experimental samplers.

Further confirmation of the precise location of spawning is provided by the lateral distribution of the plankton nets. For a relatively short distance downstream of the spawning area, the eggs should be concentrated (i.e., they will not be completely mixed). Therefore, if the experimental samplers are deployed fairly close to the spawning area, those nets directly downstream from the actual spawning locations should have higher concentrations of eggs than those on either side. (Be careful to account for the volume of water filtered by each net in making this comparison.)

The inherent limitation of plankton nets is their fine mesh. These nets can clog and fill with debris at an alarming rate. Therefore, they must be removed and cleaned frequently. The drag created by the nets is fairly strong, even when they are clean. When clogged with debris, the drag may be sufficient to break the anchor rope, pull out the anchors, or put so much tension on the anchor rope that it is difficult to remove from the anchor. For this reason, loops should be tied in the anchor rope, and the plankton nets individually attached with an easily untied knot (such as a bow knot or bowline). Carabiners can also be used to attach the nets to the loops, provided the drag is not so strong that they cannot be released.

Eggs cannot be separated from the debris in a drift sample using hydrogen peroxide, because virtually everything in the net will be organic. One simple way to separate the eggs is to wet-sieve the mixture. The sample is washed through a stack of sieves having different mesh openings (usually corresponding to the geometric size distributions listed in Table 2, between .062 mm and 16 mm). It is probably not necessary to use a complete sieve set, but make sure that the smallest sieve will filter out the eggs. The sample may be wasted if the eggs pass completely through the whole stack. Wet sieving is not as destructive as using peroxide, but there is some potential for injury to the egg. It is a good idea to preserve a subsample prior to wet-sieving in the event that the eggs must be identified. Wet sieving will not totally separate eggs from debris, but it will remove a large portion of the debris. The final stage of separating eggs from a sample may involve hand sorting, which is facilitated if most of the debris has been removed by wet sieving.

b. Larvae. Larval fish can often be sampled with the same types of gear used to collect macroinvertebrates or eggs. Young fish, however, are considerably more adept at avoiding capture than are fish eggs and aquatic insects. The Hess sampler shown in Figure 28 is often quite useful for collecting larval fish in riffles and shallow backwaters. A desirable feature of this sampler is that the area sampled is known precisely, allowing a standardized unit of effort. The current is usually strong enough in riffles to carry dislodged fry into the bunt of the sampler. They must be herded into the bunt when the sampler is used in very slow water. They must also be convinced to stay there until sampling is terminated, which may require modification of the bunt. Installing a mesh holding chamber, similar to a minnow trap or other pot gear, within the bunt may solve the problem of fish swimming out before the sample can be removed (Figure 29).

Larger fry may simply dart around inside the sampler and avoid final capture. A more active capture system may be needed for these fish, the simplest of which is a hand seine. This consists of two sticks with a small (usually 1 m^2) piece of screen between them. The hand seine is pushed along the bottom of the stream, with the top of the screen slightly above the water surface. A slight bow should be allowed to develop in the center of the screen (i.e., do not stretch it taut) to form a bunt. After the screen has been moved a specified distance, the base is raised slowly. This action captures fish on top of the screen as it is raised. When using a hand seine, avoid submerging the top of the net at any time, as fish can escape over the top. Also do not raise the bottom of the net too fast because water will run off the screen rather than through it. This can result in fish being washed off the lower lip of the screen. It is difficult not to sample heterogeneous microhabitats with a hand seine. The most common operation is to start in the middle of a pool or backwater and push the seine toward the bank. Although this is undeniably effective in capturing fish, it invariably results in the sampling of a wide range of habitat types.

Active collection gear, such as the hand seine, generally requires a considerable amount of operating room. Furthermore, it is difficult to standardize the area sampled and the unit of effort. The Hess or Water's samplers are superior in this respect. The biggest drawback to the Hess sampler is its open top and side-mounted bunt. A Hess sampler can be converted to a drop trap, however, by equipping it with a mesh lid and a diaphragm or pursing closure on its base (Figure 30). The sampler is dropped through the water column and, upon hitting bottom, the lower diaphragm closes. The diaphragm could be spring loaded so it would close automatically on contact with the bed, or manually operated so it could be closed anywhere in the water column.

Pump samplers may be very effective for developing criteria for young fish. The pump sampler can be used effectively in nearly all types of habitats. Its greatest value is in sampling around cover objects, which usually cannot be sampled with nets or traps. The pump sampler also has a few negative features. The equipment is bulky and awkward if not boat mounted (Burch 1983) and specimen damage is higher in pump samples than with other types of collection gear. Manz (1961) and Allen and Hudson (1977) describe modifications such as prefilters and bypasses that can reduce specimen damage. There is also an apparent tendency for larger fry to avoid the sampler. Cada and Loar (1982) compared the effectiveness of the pump sampler and towed plankton nets for collecting shad larvae. Both samplers were about equal at night, but the pump sampler collected significantly smaller larvae during daylight samples.

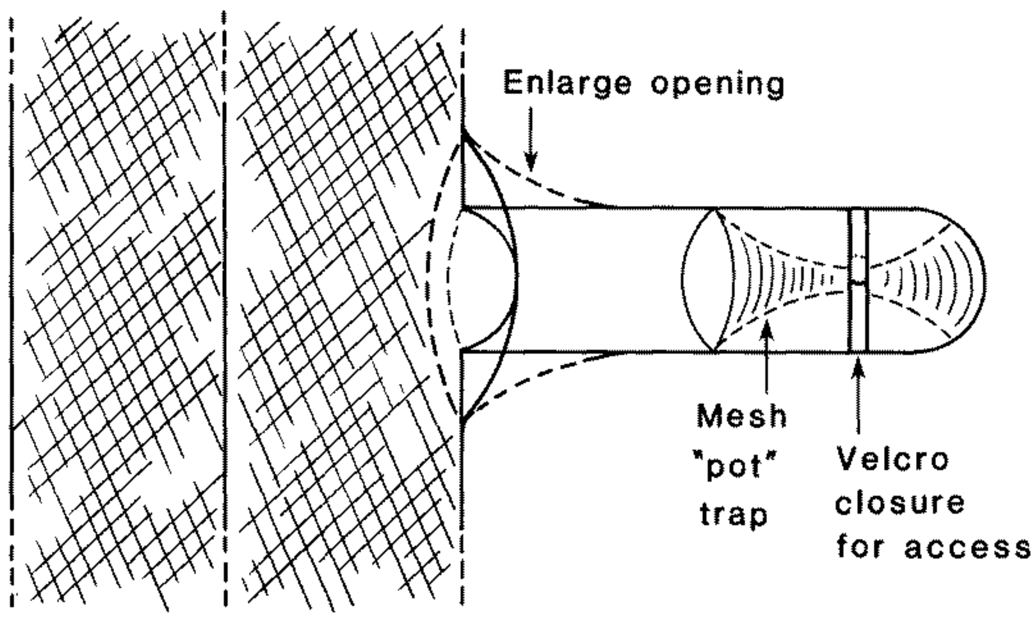


Figure 29. Modification to the bunt of a Hess or Water's round sampler for capturing young fish.

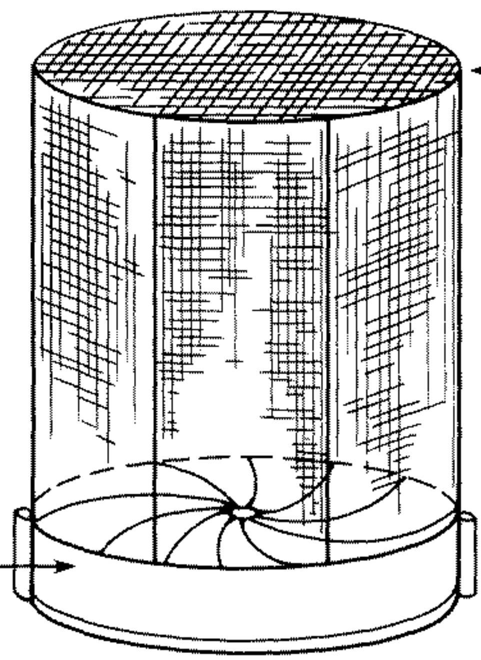


Figure 30. Prototype design of a drop trap converted from a Hess sampler, for sampling larval fish.

As in sampling for adults, a predetermined sampling design should be used to select sampling locations for larval fish. This will undoubtedly reveal the need to use a variety of equipment to sample different types of habitats. It is extremely important to be equipped to sample virtually all available habitats. Otherwise, the resulting criteria will reflect a scale for relative gear efficiency rather than habitat requirements.

The two areas where larval sampling deviates the most from adult sampling are in the gear used and in data recording. With the exceptions of direct observation and explosives, sampling equipment that is effective for juveniles and adults may be useless for fry, and vice versa. Therefore, sampling should be confined to one or the other for two reasons. First, concurrent sampling for fry, juveniles, and adults requires a large amount of equipment, creating a logistical nightmare. Second, the sampling tactics for adults may interfere with larval sampling and vice versa. The main consideration, however, is not to load the boat with so much gear that it interferes with crew performance and safety.

Many larval fish cannot be identified in the field, so it will be necessary to preserve samples for later identification and enumeration. Samples must be numbered and cross referenced in the field notes along with the measured habitat variables. If numerical codes have been used to denote species or life stages in the field book, be careful not to confuse these with sample numbers.

4.2 DETERMINATION OF HABITAT AVAILABILITY

Determining habitat availability is so much simpler than determining habitat utilization, that devotion of a separate section to this aspect of a criteria study may seen unwarranted. Habitat availability may not be very difficult to determine, but is equal in importance to utilization. Habitat availability is used to calculate preference; category III criteria cannot be developed without it. Furthermore, availability functions are indispensible for evaluating criteria quality and transferability.

There are two basic approaches to the determination of habitat availability during a criteria study: random or proportional sampling. Regardless of which method is used, habitat availability should be determined for each discharge, stream reach, and day that data on habitat utilization are collected. This may result in a data base for availability that is considerably larger than the one for utilization, but an incomplete or poorly defined availability function can produce some strange results when used to calculate the preference function (see discussion in Chapter 5).

4.2.1 Random Sampling

The basic random sampling approach combines the collection of habitat availability data with that for habitat utilization. Under this approach, sampling locations are randomly selected and sampled for fish utilization, but habitat measurements are made at every sampling location regardless of whether fish were observed or not. This is probably the statistically purest form of random sampling. It reduces data pooling problems (Section 4.3) almost entirely and inevitably results in an availability data base five to ten times larger than the utilization data base. The detraction of this approach is that it can increase field time two- to threefold. The near guarantee of a very good availability function and the elimination of data pooling problems, however, may make the extra investment worthwhile. This method is most compatible with utilization data based on fish collections rather than observations: data from electrofishing, area samplers, and explosives. This method is not very compatible with radiotelemetry or observation techniques, because every "sample" results in an observation. It is possible to take a number of random samples in the area of observation or around a telemetered fish position to determine availability, but care must be taken when pooling such data.

4.2.2 Proportional Sampling

Proportional sampling uses the same habitat mapping approach as the PHABSIM system. Details of the measurement procedures are given in Trihey and Wegner (1981). Although this technique is initially more time consuming than random sampling, most measurements only need to be made once. PHABSIM is then used for mathematical simulation of the reach for discharges other than the ones measured. This is particularly advantageous if the same reach is sampled several times, but the discharge is different each time. This method is compatible with virtually all sampling designs and techniques used to assemble the utilization data base. Obviously, proportional sampling is less desirable if each reach is only sampled once, as it might be in a telemetry study of a nomadic species. The drawback to its use with telemetry is that all the measurements would have to be repeated at a new reach every time a tagged fish wandered out of the surveyed reach. In some instances, this problem can be solved by first determining the home range of the fish, then surveying one large reach that completely overlays the home range. This is feasible as long as the home range is less than about 5 km. Unfortunately, some species wander this far in an afternoon. Taking a multitude (100 or more) of random measurements within the area being utilized by such species is probably a better approach.

4.3 POOLING DATA FROM DIFFERENT SOURCES

Pooling data refers to the practice of combining data sets collected from different reaches or at different times into a common data base. Obviously, data pooling is no concern if all the data are collected in the same reach under the same flow conditions. It is common, however, for data to be obtained from several reaches in the same stream, from different streams, under different streamflow conditions, or with different gear. It is under these circumstances that data pooling problems are created.

The crux of the data pooling dilemma is to avoid overrepresentation of data from one source, with respect to others. When habitat utilization data are taken from more than one site, it is important to recognize that the frequency of observations or collections is a function of several factors besides habitat suitability. The surface area sampled, the time spent sampling, and the efficiency of the technique used also influence the frequencies recorded in the data base. In developing a category II data base, these secondary factors are assumed to be constant. They will <u>be</u> constant, however, only if the investigator makes them so.

The easiest way to eliminate data pooling error from category II data is to use study sites of equal areas, sample each of them the same number of times, and use the same observation or collection technique each time. This is the only way to guarantee that the frequencies of fish observation are not influenced by unequal sampling effort. The alternative to equalizing effort is to record catch per unit effort (CPUE) rather than raw frequency. This will be easier to do with capture techniques (e.g., electrofishing, explosives, or area samplers) than with observation techniques (e.g., snorkeling or radio-telemetry). It is difficult, if not impossible, to define a unit of effort for many observational methods.

The aforementioned data pooling problems and solutions are particularly appropriate in the development of category II criteria. Some of the sampling strategies used to develop category III criteria will automatically correct for differential sample areas. Some techniques used to describe habitat availability, however, can misrepresent the area sampled, and actually reverse the bias of unequal sample areas. (Reversing the bias does not mean neutralizing it, but that the data will be biased in the opposite direction).

The two types of sampling designs that internally correct for differential sample areas and unequal effort in each area include:

(1) active capture techniques with a standardized unit of effort, used to describe utilization ($P[E|F]$ from equation 5) and availability ($P[E]$ from equation 5) at the same time; and

(2) observational techniques utilizing a proportional sampling design to determine availability.

The first case is illustrated by the use of a prepositioned area shocker at randomly selected locations in three streams, and measuring the environmental variables at each location, regardless of whether or not fish were taken. Thirty fish are taken in stream A, with 90 set-ups; 40 fish are taken in stream B, with 150 set-ups; 10 fish are taken in stream C with 20 set-ups. Based on raw frequencies only, the environmental conditions in stream B would appear to be the best because the most fish were caught there. On a catch per unit effort basis, however, stream C is obviously better. One way of standardizing $P[E|F]$ would be to use catch per unit effort instead of raw frequencies, but this is unnecessary. A total of 260 samples were taken and used to describe $P[E]$. Of these, the proportion taken in stream C was only about 8% of the total. When the environmental conditions represented by this small fraction are added into the denominator, the effect is approximately the same as multiplying all the fish observations in stream C by 12.5 (1/.08) and assuming the same number of samples in all streams. In essence, the equation is standardized because each sample represents a constant unit of effort, and each raw frequency is additive. This is equivalent to stating that if as many samples had been taken in stream C as in stream B, the investigator would have caught 75 fish instead of 10. The sure way to ensure pooling compatibility with this method would be to standardize the sampling areas so that this is a fact and not an assumption.

The second case, proportional sampling, is represented by a team of divers observing fish in three stream reaches, and determining P[E] with PHABSIM habitat mapping. Reach A encompasses 25,000 m^2, reach B, 15,000 m^2, and reach C, 40,000 m^2. In this case, P[E] is determined for each increment or combination of environmental variables on the basis of the total area having that combination in all three reaches, divided by the total surface area (80,000 m^2). In essence, this approach would state that the conditions in stream reach C are 2.67 times more available than those in reach B. Again, the reason that such data can be pooled directly is that the units of availability are additive. (Note: PHABSIM output expressed in units of area per unit length must be corrected to reflect true reach length). Observations should be confined to the actual area encompassed by the PHABSIM site to avoid the occurrence of fish in conditions that appear to be unavailable from the environmental data.

Two methods of determining habitat availability are certain to create a data pooling bias when calculating category III criteria:

(1) taking a standard number of random samples of the environment based on the number of fish observed; and

(2) systematic sampling of the environment where different intervals between samples are used in different reaches.

To illustrate the first case, suppose that 10 random samples of the environment are taken each time a fish was observed. Study reach A and reach B are the same size, but 40 fish are observed in A and 20 in B. With this sampling design, 400 measurements of the environment would be taken in A and only 200 in B, implying that the conditions in A are twice as available as those in B. When using any type of random sampling, where P[E] will be defined by raw frequencies, the number of random samples in each reach must be in proportion to the total area sampled. Clearly, in this case, the same number of random samples should have been taken at both sites. If A is twice as large as B, then A should have twice the number of random samples.

A similar bias can occur when systematic sampling designs are used. One popular sampling design uses a zig-zag pattern, where a transect is laid diagonally across the channel. A diver follows the transect, counting fish found within a meter on either side of the line. Measurements of environmental conditions are made 1/4, 1/2, and 3/4 the way across each transect, as well as at each edge. Mathematically, the problem with this sampling design is identical to the previous example; each sample of the environment enters P[E] as a frequency, but the frequencies do not represent the same areas. The two solutions are to use a constant spacing between measurement points, regardless of the size of the stream, or select study sites that all have the same width.

4.4 DISCUSSION

Several important concepts are introduced in this chapter. First, there are many techniques that can be used to develop the habitat utilization data base. Second, each technique has certain strengths and limitations that must be evaluated with respect to the organism under study and the environment in which the study is conducted. One certain way to bias a data base is to use a technique in a situation where the limitations of the method are exceeded. Third, there are many factors, besides gear limitations, that can introduce error into a data base. How the data are collected has a much broader implication than the physical act of observing or collecting fish.

Most practitioners agree that the best overall technique for collecting habitat utilization data is by snorkeling. This method has been field-proven to be unobtrusive, efficient, and, with proper gear and training, relatively safe. Its major limitations are that the field of vision must exceed the maximum depth of the stream, the observer must be able to discern the size and species without handling the fish, and the quality of the data may be influenced by the safety and comfort of the observer. Observation efficiency may decline in very fast or very cold water due to observer fatigue or discomfort. Under these conditions, and where beneath-ice observations are necessary, underwater video or periscopes operated from the surface may provide a suitable alternative. When using either of these, the reaction of the fish to the lighting system should be evaluated before collecting any data.

SCUBA observations can generally be made under more marginal conditions than by snorkeling. A SCUBA diver's observations are not biased by water depth, for example, whereas a skin diver has more difficulty observing fish in deep water. SCUBA observations can often be made under conditions of poorer visibility than skin diving, and a SCUBA diver is more mobile than an underwater video camera. The two dominating concerns in using SCUBA to obtain observations are safety and fish disturbance. SCUBA is generally considered to be more dangerous than skin diving, although an experienced diver will rarely place himself in an overly hazardous position. The key word is experience. Inexperienced divers should start by making observations under the safest of conditions (such as a small stream where a diver could walk out, if necessary) and work up slowly to more difficult conditions.

Another potential problem with SCUBA is that the behavior of the fish may be altered by bubbles exhaled by a diver. Some investigators have attached long hoses to the regulator to vent their bubbles downstream. This can be a very dangerous practice and is discouraged for all but the most experienced divers, especially in water over about three meters deep. The problem occurs during surfacing, when the air in the hose expands and creates a back pressure in the regulator, preventing exhalation. As a consequence, the diver must perform a free ascent or risk an embolism. Rather than using an exhaust hose, it is better to make several trial dives to determine how close the diver can approach the fish without disturbing them. The diver should then maintain that distance throughout the observations, or until the fish become accustomed to his or her presence.

Collection methods may be needed where the visibility is too poor to make observations. Of these, the prepositioned area shocker designed by Bain et al. (1985b) appears to be the most promising, although small, prepositioned explosive charges have potential for some applications. The strength of either technique is that the sampling device is left undisturbed long enough to allow the fish to resume normal activity. Both methods are also very compatible with random sampling designs, which reduce data pooling problems. Electrofishing, in general, is size selective and prone to depth bias, traits not shared by primacord. The major problem with explosives is that they kill fish, rather than stunning them. Recovery and enumeration of fish killed by an explosive charge may also be more difficult than capturing fish in an electrical field. A serious concern with both methods is that they are potentially very dangerous. Training in safety procedures, first aid, and CPR should be required of all members of field crews employing these techniques. Both techniques also suffer somewhat from inefficiency. Many of the samples will contain no fish, which means that a large number of samples must be taken to obtain a sufficient data base. The most promising deployment strategy for either is to set out several sampling devices at once, then activate them sequentially, moving in an upstream direction. When using either method, it is advisable to measure the habitat characteristics of each sample location, regardless of whether fish were collected or not. This will produce utilization and availability functions that are unbiased by differential units of effort or sampling areas.

Radiotelemetry has been demonstrated to be an effective method of locating fish in streams. The most appropriate applications of this technique are in large, turbid rivers and with highly migratory species having relatively low population densities (Colorado squawfish, paddlefish, sturgeons, walleye, blue suckers, and flathead catfish, for example). Radiotelemetry is one of the few methods that can be used to locate fish under ice cover. The principal limitations of this technology are triangulation error, the inability to monitor a large number of fish, and the inability to monitor small species or life stages. Triangulation error can be reduced by a number of procedures, but pinpointing the precise location of a transmitter usually requires some confirmatory technique. Unfortunately, there is not much that can be done to overcome the second and third limitations. It is quite easy to build up a data base containing many observations, but it must be recognized that most of these are the same fish, observed repeatedly. This is less of a problem if the fish are highly mobile because they will be exposed to a wide variety of habitat types and will have the opportunity to select their preferred habitats from a larger universe. When this technique is applied to sedentary species, however, the exact habitat locations may be measured over and over. Part of the solution to this problem is to install as many transmitters as can feasibly be monitored. Winter (1983) suggests that this number is between 20 and 40 animals. Another suggestion is to try to make about the same number of observations on each transmitter, rather than 10 observations on one and one on another. Finally, individual observations on a particular transmitter should be spread out temporally (e.g., a week or so) to allow the fish time to select alternative microhabitat sites. This time period is obviously a function of the amount of movement exhibited by the fish.

The aforementioned field techniques have all been used successfully in habitat utilization studies. Their limitations are fairly well known and methods of circumventing or lessening these limitations have been devised. The same cannot be said for nets, seines, and traps used in an active capture mode. Although seines have been used on occasion, the way that they have been operated (by sweeping them up the bank) practically guarantees erroneous data. The encircling maneuver discussed for miniature purse seines and trammel nets may alleviate this problem, but may also prove to be very difficult to implement in a river. Lift nets may provide a viable technique for sampling deep areas that cannot be reached by electrofishing and are too turbid for underwater observation. Drop traps, such as the modified Hess sampler, could be useful for sampling small schooling fish. Unfortunately, none of these tools has been used extensively enough in criteria studies to establish much of a track record. The performance of their predecessors was not very good, and it is too early to determine if the modifications suggested in this chapter will help.

The same techniques will not work equally well for all life stages or species, under all environmental conditions. This is a critical concept in developing a category II data base, because once a technique has been chosen, it should be used consistently for that particular data stratification. This approach is suggested to avoid problems with differential gear efficiency influencing the frequencies of fish recorded in the data base. This does not mean that the same technique must be used throughout the study. The best method of determining adult brown trout habitat utilization during the winter might be radiotelemetry. During the summer, it might be skin diving and during the spring, electrofishing. What this means is that it may be necessary to utilize a wide variety of equipment, for a study with only modest objectives. A list of suggested equipment for each sampling technique is given in Appendix B, to assist investigators in preparation for a field study.

Where samples are taken is as important in a criteria study as how they are taken. There are several important considerations when establishing study areas, including: habitat diversity, size of the study area, and method of quantifying habitat availability.

The importance of habitat diversity was emphasized in Chapter 2. Perhaps the most important study design consideration is that a very long reach, or several disjunct reaches, may be required to obtain a sufficiently diverse universe in which to observe fish. Use of a single, highly diverse, and relatively large (i.e., length = 20-30 times the channel width) study reach is probably the best way to avoid data pooling problems. The use of several smaller reaches will be more likely, however, in most studies. When selecting this latter option, it is highly recommended that the study sites all encompass the same surface areas, even when using a pure random or proportional sampling design. This is the only certain way of avoiding bias due to differential sampling effort.

The two most applicable methods of measuring habitat availability are pure random sampling and habitat mapping (proportional sampling). Other methods include stratified random and systematic sampling schemes. When using any technique that uses a raw frequency to describe habitat availability, it is essential that the number of samples are proportional to the area sampled. This is especially important with stratified random and systematic sampling designs.

The most important aspect of using a proportional sampling strategy is to confine the fish observations to the area within the upper and lower transects of the study area. Otherwise, the area sampled for fish is not the same as the area sampled for availability. This may not be a problem if only one reach is sampled, but if data are pooled, the potential exists to bias either the utilization or the availability data base, or both. Some investigators have mapped a small representative reach to describe habitat availability, and made fish observations over a much larger reach of stream. There are two hazards with this approach. First, fish might be observed utilizing conditions not described by the representative reach. Consequently, preference for that set of conditions is undefined because the denominator in equation 5 is zero. Second, the surface areas sampled may not be the same, leading to a bias in the utilization data base, or the availability functions for the various reaches are incorrect, resulting in an error in the category III criteria. Again, the easiest solution is to establish study areas of equal size.

5. DATA PROCESSING AND DOCUMENTATION

The goal of data processing is to reduce raw frequency data down to an easily interpreted graphical display that represents the behavioral response of a species with respect to one or more environmental variables. There are numerous smoothing techniques, some more sophisticated than others, that can be employed to this end. Most of these techniques attempt, in one way or another, to do two things: (1) to fit a "line" through the data points that best represents the shape of the distribution, and (2) to minimize the amount of scatter of the data points about the chosen "line." (The term "line" is in quotes because it is sometimes multidimensional.) Three basic approaches have evolved for the processing of habitat utilization and preference data: histogram analysis, nonparametric tolerance limits, and function fitting.

5.1 HISTOGRAM ANALYSIS

Histogram analysis is the most elementary, although not necessarily easiest of the three curve-fitting methods. A histogram is little more than a bar graph showing frequency on the ordinate (Y-axis) and the range of some variable on the abscissa. A frequency polygon is created by plotting a point at the middle of each bar and connecting the points. If the data base is large and smoothly distributed, such as the one shown in Figure 31, the frequency polygon will closely approximate the suitability curve that would be obtained by virtually any other technique. In this case, the relative frequencies of the polygon could simply be normalized (i.e., the largest frequency given a value of 1.0, and the other frequencies proportioned to a maximum value of 1.0) and used directly as a utilization (or availability) function.

5.1.1 Derivation of Habitat Utilization and Availability Curves

Data are seldom distributed as regularly as those shown in Figure 31. The frequency distribution shown in Figure 32 is more typical of data collected for utilization or preference criteria. The basic shape of the curve is evident from the histogram, but the unevenness of the distribution makes it difficult to determine exactly where the line should be placed.

There are many reasons for the irregular shape of the histogram in Figure 32, few of which have anything to do with fish behavior. The most common reason is that microhabitat measurements, especially depth and velocity, are precise to about 3 cm (or cm/sec) and are, initially, not grouped. This can result in irregular histograms from measurement and roundoff error. As more intervals are grouped together, the effects of these errors are lessened, and the histogram is smoother. The occurrence of schools of fish in some of

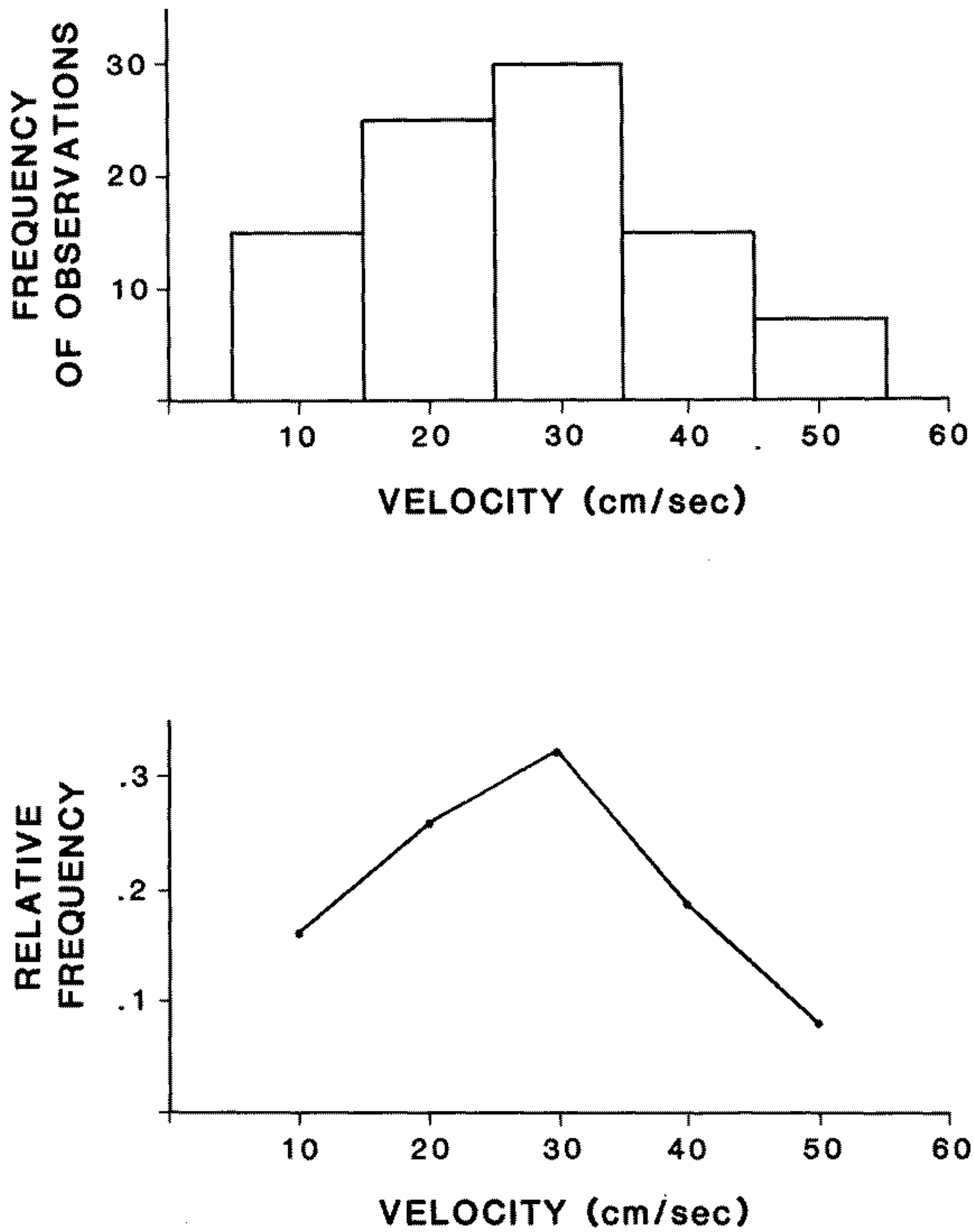


Figure 31. Example of smooth histogram and resulting frequency polygon.

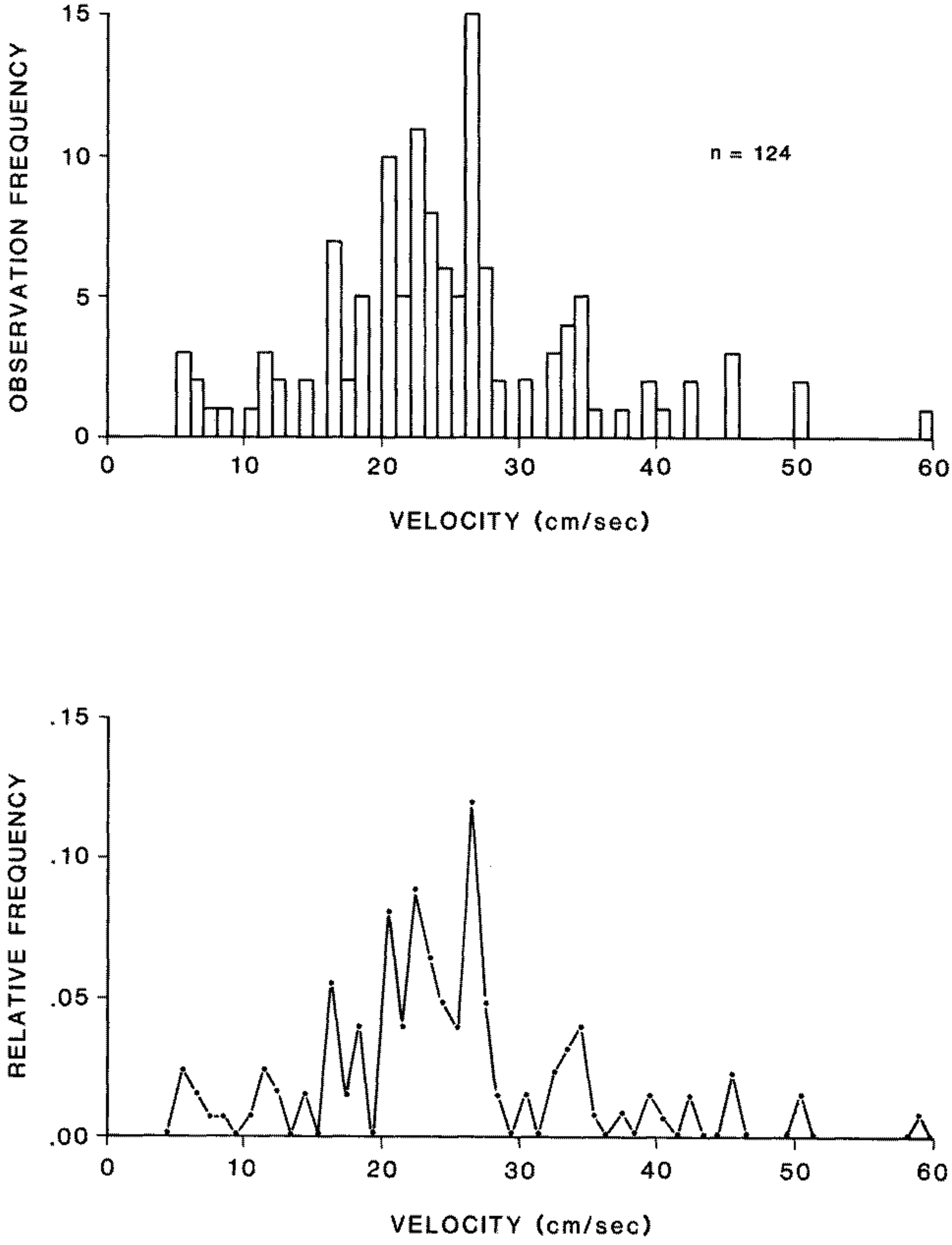


Figure 32. Histogram and frequency polygon typical of habitat utilization data.

the samples is another frequent cause of irregular histograms. Where a normal sample might turn up one or two fish, a location with 35 of them is bound to cause irregularities in a histogram. This problem is partially controlled by obtaining a larger data base and by expressing frequency as CPUE. The best solution for highly clumped data may be to use a binary system where each observation is assigned a frequency of 1.0, regardless of the number of fish found at the location.

It should be readily apparent that the frequency polygon shown in Figure 32 should not be used as a utilization curve. This frequency polygon, as a reflection of the behavior of an animal, simply does not make any sense. For example, 10 animals were found at 20 cm/sec and none at 19 cm/sec; 15 were found at 26 cm/sec and only 3 at 27 cm/sec. Although fish can sense changes in velocity as small as one cm/sec, this type of distribution is more often the result of schooling, inadequate sample size, or measurement (or roundoff) error than from discriminatory behavior by the fish.

The frequency polygon in Figure 32 needs to be smoothed before it can be used as a criteria curve. One way to accomplish this is to cluster the data into larger increments, as shown in Figure 31. Although this technique will result in a smoother curve, it may be at the expense of a less accurate histogram. Each time two or more adjacent increments are grouped together, the relative frequency is assigned to the midpoint of the interval. It is never very obvious how large the increment should be. Larger intervals result in smoother histograms but accuracy may be lost at each grouping. Furthermore, by grouping the data, all that has been accomplished is a better fit of the curve to the histogram. The process is counterproductive if smoothing the histogram distorts the original distribution.

A more accurate technique of fitting a curve to a frequency distribution like the one in Figure 32 is to sketch in a curve and compute the residual sum of squares. An observed frequency on a histogram can be described as a single point on a frequency polygon, having the coordinates (X_i, Y_i), where Xi is the distance along the abscissa and Y_i is the observed frequency. When a line is drawn through a collection of data points, there is a predicted value of Y_i (denoted as $\hat{Y}_i$) for each interval on the x-axis. The vertical deviation of each data point from the line is described as $(Y_i - \hat{Y}_i)$. The residual sum of squares (SS_r) is computed as:

$$SS_r = \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2 \qquad (6)$$

where Y_i = the observed frequency at interval X_i, and

$\hat{Y}_i$ = the predicted frequency at interval X_i (Zar 1974).

The line that best fits the distribution is the one that minimizes the residual sum of squares. This technique is also known by the more familiar term, least squares fit. The main difference between this approach and more standard statistical curve fitting procedures is that each curve is fitted by eye, the residual sum of squares manually calculated, and the process repeated until the least squares curve is found or the investigator is exhausted. Because the distributions do not always fit into neat categories for which a function has been derived, this process is largely one of manual trial and error. The time consuming and rather haphazard aspects of this approach are the main reasons it is seldom used. A more systematic approach to the same principle will be discussed in Section 5.3.

The advantages of histogram analysis are the relative simplicity of concept and the freedom of not being confined to predetermined mathematical functions. Some distributions, easily and obviously drawn by hand, may be nearly impossible to approximate in equation form. Another advantage is that both utilization and availability are derived in terms of relative frequency, facilitating the transition to preference criteria. The major drawback to the approach is the difficulty posed in determining precisely the right line to draw through the data. This is usually less of a problem with very large amounts of data, but can be quite serious with small or medium-sized data bases.

5.1.2 Calculation of the Availability Distribution

The frequency distribution of habitat availability is determined by one of two methods, depending on the sampling approach. Using the random sampling approach, each sample (i.e., measurement of depth, velocity, substrate, and cover) has a frequency of one. All samples having the same habitat measurements are combined to determine the total frequency of samples having that unique combination of characteristics. The probability of encountering that particular combination is then computed by dividing the frequency of occurrence by the total number of samples in the data base.

An entirely different approach is used with proportional sampling, but the result is the same. Each cell in a PHABSIM site is represented by a depth, velocity, substrate, cover type, and surface area. The surface areas of all cells having the same combination of habitat variables are added together and the probability of occurrence for that combination is calculated by dividing these summed surface areas by the total surface area of the reach. The advantage of this sytem is that once the original set of data has been calibrated, the entire process can be completed on a computer, with no further data entry.

5.1.3 Derivation of the Preference Function

Preference is simply computed as the ratio between utilization and availability:

$$P_i = U_i/A_i \tag{7}$$

where P_i = an unnormalized index of preference at x_i,

U_i = the relative frequency of fish observations at x_i,

A_i = the relative frequency of x_i available during the observation period, and

x_i = the interval of the variable (x).

Two techniques can be used to derive a preference curve from habitat utilization and availability. The first is based on curves previously fit to both the utilization and availability histograms. For each increment across the range of a measured microhabitat variable (X_i), there will be a predicted relative frequency ($\hat{Y}_i$) for either the utilization or availability of that increment. The unnormalized preference for the increment is computed as the ratio between these two predicted values. The second method is to divide the observed relative frequencies from the utilization and availability histograms and then fit a curve to the resulting "preference histogram." This latter technique is somewhat easier if both histograms are fairly smooth, because a curve is fit to only one (the preference) histogram. If either the utilization or availability histogram is uneven or has zero frequencies within the distribution, however, the uneveness may be accentuated in the preference histogram. This makes curve fitting more difficult, so in this case, it would be better to compute preferences by the first technique.

Preference ratios must be normalized before they can be used in PHABSIM, because the maximum value of any weighting factor used in the program is 1.0. All suitability curves are normalized for the following reasons. First, the weighting factor is used to convert an actual area (e.g., in square feet) to an equivalent area of preferred or most highly utilized habitat. Therefore, the total area of a cell having these preferred conditions should be fully counted as suitable habitat. Normalization allows the direct comparison of these weighted areas, whereas relative frequencies do not. Second, the only type of criteria that would be compatible with the use of empirical probabilities are unnormalized category II functions. These tend to be the most biased and site specific of all types of criteria. Empirical probabilities are unknown when developing category I criteria. Category III functions are indexes developed by the division of relative frequencies. These indexes, if unnormalized, can result in suitability values greater than unity. Values for suitability cannot exceed unity, because given enough area of preferred conditions, such weighting factors could result in more habitat than wetted area. The logic of this upper limit seems fairly obvious.

The computation of preference, starting with relative frequency of utilization and availability and ending with normalized preference criteria, is illustrated in Table 6. Corresponding utilization, availability, and preference curves are shown in Figure 33.

Table 6. Calculation and normalization of preference criteria from smoothed utilization and availability relative frequencies.

Depth (cm)	Utilization (rel. freq.)	Availability (rel. freq.)	Ratio	Preference (normalized)
15	.05	.30	.167	0.042
30	.10	.20	.500	0.125
45	.15	.20	.750	0.188
60	.20	.15	1.333	0.333
75	.30	.10	3.000	0.75
90	.20	.05	4.000	1.0

5.2 NONPARAMETRIC TOLERANCE LIMITS

The general concepts of nonparametric tolerance limits are attributed to Wilks (1941), who showed that for continuous populations, the proportion of the population between two order statistics is independent of the population sampled and a function only of the order statistics chosen (Wilks 1941; Somerville 1958). The concept underlying nonparametric tolerance limits is illustrated in Figure 34, where F is a theoretical distribution with respect to x.

At each couple (a,b) there is a corresponding probability, $F_{a,b}$, that a certain proportion of the population is between a and b:

$$F_{a,b} = P[a<X<b] \tag{8}$$

Consider a sample of n observations, and let x_a and x_b be the a^{th} and b^{th} order values in the sample. The values associated with x_a and x_b are approximations of the true values for the random variables x_a and x_b. At each of these values for x_a and x_b, there is also a corresponding value of Fx_a, x_b, which is itself an approximation of the true proportion of the population that is between x_a and x_b. The true proportion lying between the a^{th} and b^{th} value of a sample will be at least equal to Fx_a,x_b, with a probability equal to α

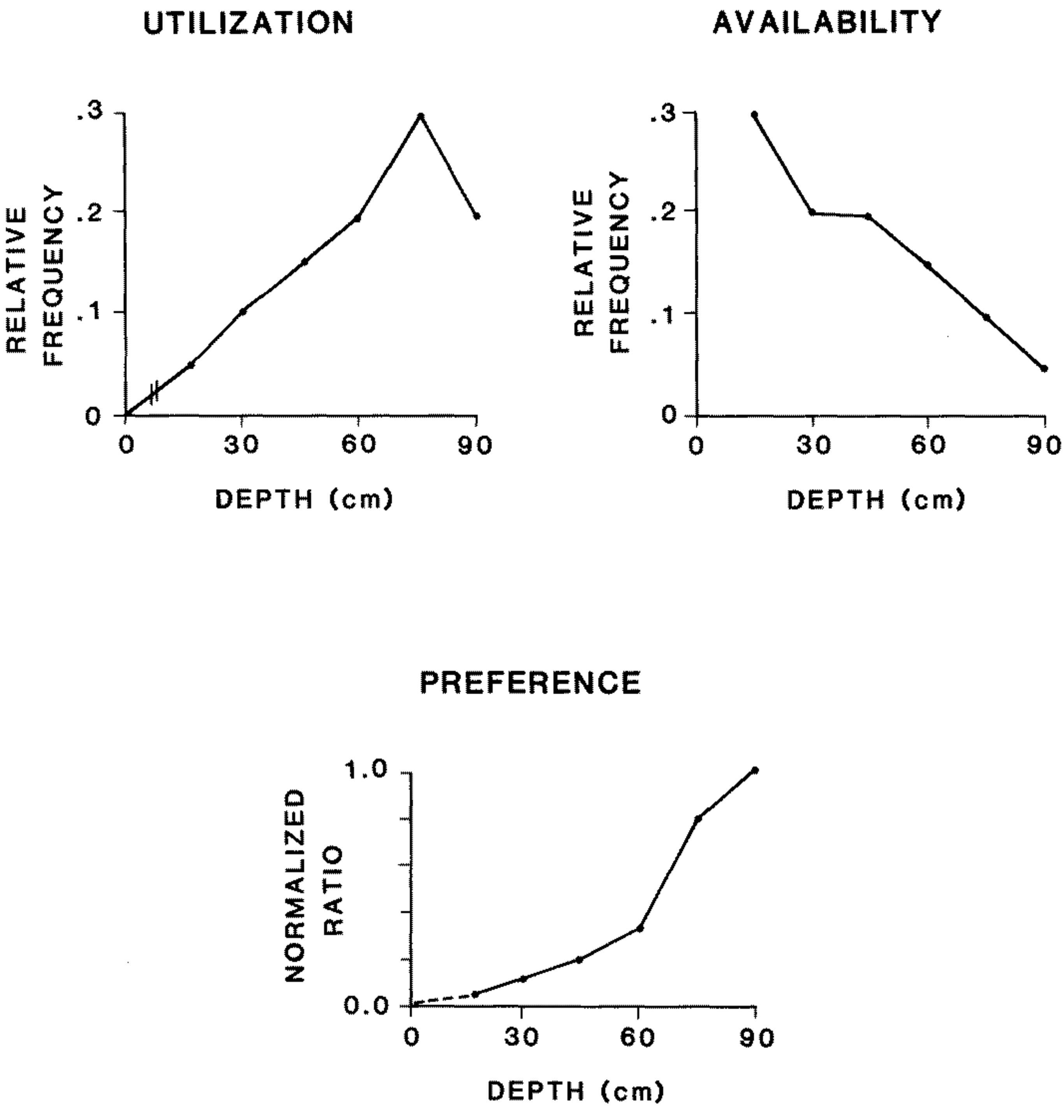


Figure 33. Comparison of utilization, availability, and preference curves derived from histogram analysis.

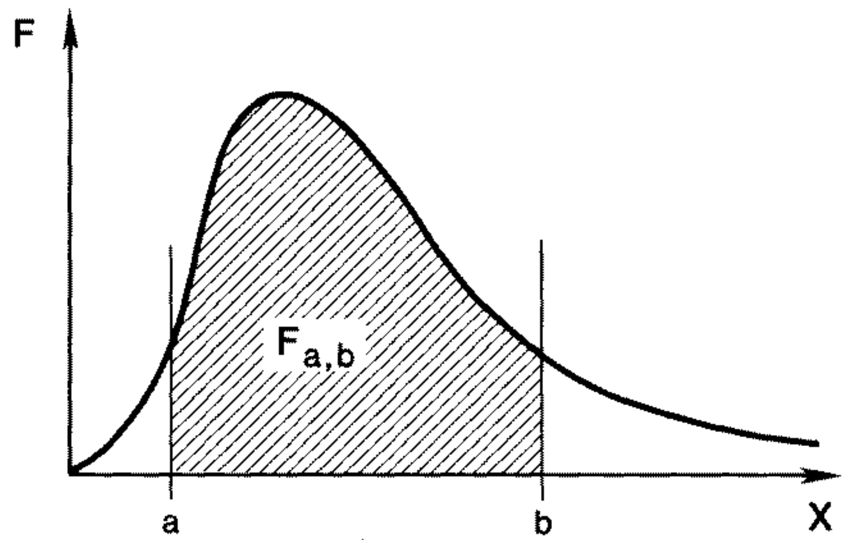


Figure 34. Theoretical distribution of F with respect to x, illustrating the concept of nonparametric tolerance limits.

(i.e., the confidence level, α). By choosing the desired level of confidence (α) and the proportion of the population to be included between x_a and x_b, it is possible to find the limits of x_a and x_b that will contain the specified proportion. The larger the sample size, n, the "tighter" these limits can be determined at a given confidence level. This means that if individual sample data are ordered (e.g., from lowest to highest) it is possible to determine, with a given level of confidence, the range of the variable that will contain a specified proportion of the population. This is true regardless of the distribution of the data (i.e., it does not need to be a normal or other predetermined distribution).

5.2.1 Development of Utilization Curves

Gosse (1982) was the first instream flow researcher to advocate the use of nonparametric tolerance limits to derive habitat utilization curves. The use of this technique is made easier by the availability of contingency tables containing the order values for a given sample size, proportion, and confidence level. These tables generally contain pairs of numbers, one of which is the smallest, the other the largest, ordered values in a sample of size n. These ordered values define the limits of ordered data such that a specified proportion of the population will be between a and b, with a given confidence level (Remington and Schork 1970).

The use of nonparametric tolerance limits is introduced using the depth utilization data in Table 7 and the tolerance limits listed in Remington and Schork (1970). The data in Table 7 have been ordered by increasing depth and a frequency has been entered opposite each recorded depth. This sample has a total of 128 observations. For a sample of this size, Remington and Schork (1970) give order values of (13, 13) as the upper and lower limits within which 75% of the observations will occur, at the 90% confidence level. The 13th smallest value in Table 7 occurs at a depth of 35 cm: the 13th largest value occurs at 70 cm. Therefore, 75% of the population can be found between 35 and 70 cm depth, at the 90% confidence level. At the same level of confidence, 90% of the population should be found between order values of (5, 4) according to Remington and Schork (1970). The 5th smallest value occurs at 25 cm and the 4th largest at 75 cm. Similarly, 95% of the population should occur between order values (2, 1) corresponding to a lower depth limit of 20 cm and an upper limit of 85 cm.

Gosse (1982) extended this approach to curve development by assigning a normalized weighting factor of 1.0 to the range of the variable that encompasses 50% of the observations. A weighting factor of 0.5 was assigned to the range encompassing 75% of the observations, 0.10 to the range encompassing 90%, and 0.05 to the 95% range. One improvement to the approach suggested by Gosse is to compute the normalization factor as:

$$NSI = 2(1-P) \tag{9}$$

where NSI is the normalized suitability index and P is the proportion of the population under the curve (i.e., the 50%, 75%, 90%, and 95% ranges).

This will keep each of the normalized values in proportion to the original areas under the curve. This really means that the 90% range should have a weighting factor of 0.2, and the 95% range, a weighting of 0.1. Weighting them at 0.1 and 0.05, respectively, makes these values out of proportion (i.e., inconsistent) with the weights given to the 50% and 75% ranges. Figure 35 shows a histogram of the data presented in Table 7, superimposed with the utilization curve derived according to Gosse's (1982) suggested approach.

Somerville (1958) has developed a table (Table 8) that contains order values for the 50%, 75%, 90%, 95%, and 99% proportions. This table contains only one number beneath each proportion column, which is the sum of the highest and lowest order values, corresponding to those found in Remington and Schork (1970). Somerville leaves the division of the sum between the upper and lower components to the discretion of the user, but simply dividing by two to determine upper and lower order values is consistent with Remington and Schork (1970).

Whereas histogram and function-fitting approaches attempt to construct a line between data points, the use of tolerance limits results in a curve forming an umbrella over the data points. The only sure way to make the

Table 7. Ordered depth utilization data to illustrate curve construction by nonparametric tolerance limits.

Depth (cm)	Frequency
10	0
15	1
20	1
25	3
30	5
35	9
40	12
45	18
50	15
55	22
60	19
65	19
70	9
75	7
80	2
85	1
90	0
	n = 128

Table 8. Nonparametric tolerance limits. From Somerville (1958). Reprinted with permission of publisher.

n	γ = 0.50					γ = 0.75					γ = 0.90					γ = 0.95					γ = 0.99				
P	.50	.75	.90	.95	.99	.50	.75	.90	.95	.99	.50	.75	.90	.95	.99	.50	.75	.90	.95	.99	.50	.75	.90	.95	.99
50	25	12	5	2	0	22	10	3	1	--	20	9	2	1	--	19	8	2	--	--	16	6	1	--	--
55	28	14	5	3	0	25	12	4	2	--	23	10	3	1	--	21	9	2	--	--	19	7	1	--	--
60	30	15	6	3	0	27	13	4	2	--	25	11	3	1	--	24	10	2	1	--	21	8	1	--	--
65	33	16	6	3	0	30	14	5	2	--	27	12	4	1	--	26	11	3	1	--	23	9	2	--	--
70	35	17	7	3	1	32	15	5	2	--	30	13	4	1	--	28	12	3	1	--	25	10	2	--	--
75	38	19	7	4	1	35	16	6	2	--	32	14	4	1	--	30	13	3	1	--	27	10	2	--	--
80	40	20	8	4	1	37	17	6	3	--	34	15	5	2	--	33	14	4	1	--	30	11	2	--	--
85	43	21	8	4	1	39	19	7	3	--	37	16	5	2	--	35	15	4	1	--	32	12	3	--	--
90	45	22	9	4	1	42	20	7	3	--	39	17	5	2	--	37	16	5	1	--	34	13	3	1	--
95	48	24	9	5	1	44	21	7	3	--	41	18	6	2	--	39	17	5	2	--	36	14	3	1	--
100	50	25	10	5	1	47	22	8	3	--	44	20	6	2	--	42	18	5	2	--	38	15	4	1	--
110	55	27	11	5	1	51	24	9	4	--	48	22	7	3	--	46	20	6	2	--	43	17	4	1	--
120	60	30	12	6	1	56	27	10	4	--	53	24	8	3	--	51	22	7	2	--	47	19	5	1	--
130	65	32	13	6	1	61	29	11	5	--	58	26	9	3	--	56	25	8	3	--	52	21	6	2	--
140	70	35	14	7	1	66	31	12	5	1	62	28	10	4	--	60	27	8	3	--	56	23	6	2	--
150	75	37	15	7	1	71	34	12	6	1	67	31	10	4	--	65	29	9	3	--	61	26	7	2	--
170	85	42	17	8	2	81	39	14	7	1	77	35	12	5	--	74	33	11	4	--	70	30	9	3	--
200	100	50	20	10	2	95	46	17	8	1	91	42	15	6	--	88	40	13	5	--	84	36	11	4	--
300	150	75	30	15	3	144	70	26	12	2	139	65	23	10	1	136	63	22	9	1	130	58	19	7	--
400	200	100	40	20	4	193	94	36	17	3	187	89	32	15	2	184	86	30	13	1	177	80	27	11	--
500	250	125	50	25	5	242	118	45	22	3	236	113	41	19	2	232	109	39	17	2	224	103	35	14	1
600	300	150	60	30	6	292	143	55	26	4	284	136	51	23	3	280	133	48	21	2	272	126	44	18	1
700	350	175	70	35	7	341	167	65	31	5	333	160	60	28	4	328	156	57	26	3	319	149	52	22	2
800	400	200	80	40	8	390	192	74	36	6	382	184	69	32	5	377	180	66	30	4	367	172	61	26	2
900	450	225	90	45	9	440	216	84	41	7	431	208	79	37	5	425	204	75	35	4	415	195	70	30	3
1000	500	250	100	50	10	489	241	94	45	8	480	233	88	41	6	474	228	85	39	5	463	219	79	35	3

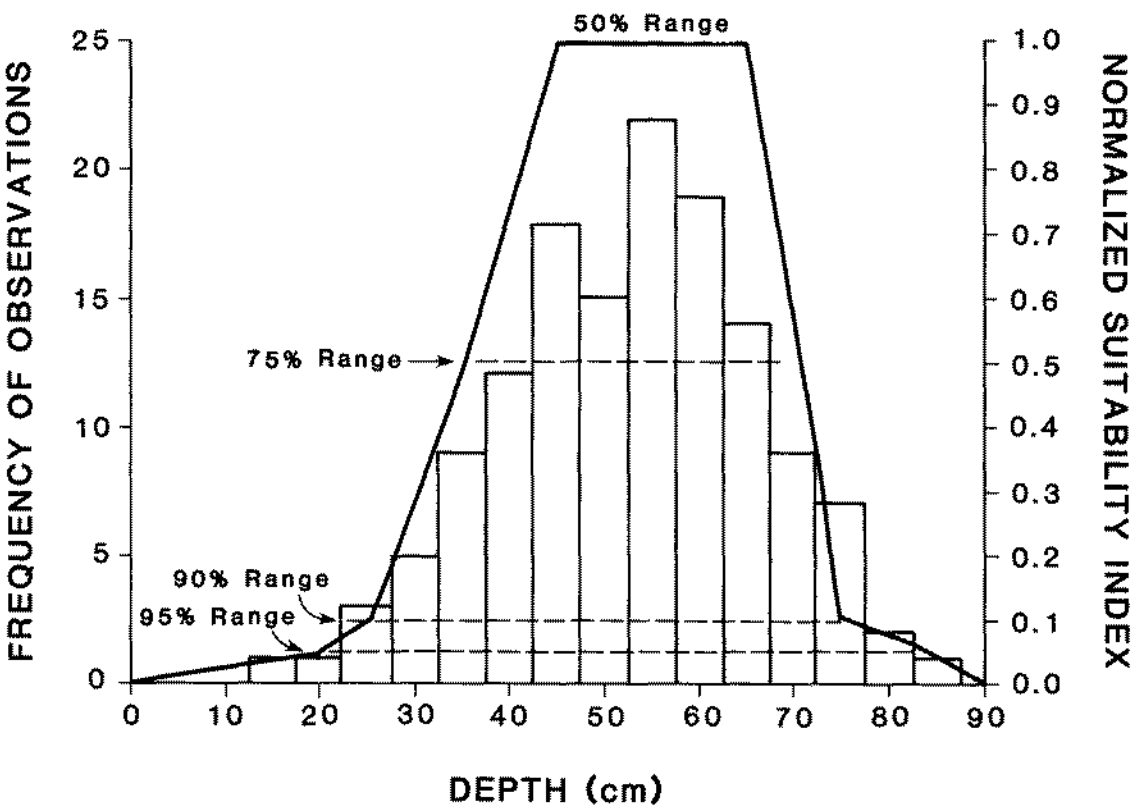


Figure 35. Frequency histogram and corresponding utilization curve derived from nonparametric tolerance limits. Curve represents tolerance limits at 90% confidence level.

umbrella smaller is to lower the confidence level (e.g., from 90% to 75%). Using a larger sample size may result in a narrower range of tolerance limits, but this is not guaranteed.

This approach has several appealing features, not the least of which is its simplicity of application. The tolerance interval is not influenced by irregularities in the frequency histogram. Therefore, one does not need to worry about clumped data or whether the bumps and dips in a histogram reflect species behavior or are simply artifacts of data collection. There is no need to compute the residual sum of squares because the technique does not involve function-fitting. Like histogram analysis, this method does not presume any particular distribution or curve shape; assymetrical curves can be produced as easily as symmetrical ones.

5.2.2 <u>Development of Preference Criteria</u>

Gosse (1982) originally suggested the tolerance limits approach for the development of utilization curves. There is no reason that availability curves could not be likewise constructed. The complicating factor with regard to preference criteria, is that tolerance limits represent integrated areas under the curve rather than relative frequencies. Computation of preference requires the use of some sort of relative frequency. Therefore, it is necessary to approximate the relative frequency distribution from the tolerance limits. The conceptual aspects of this process are somewhat complicated, but procedurally, it is quite straightforward. The basic concept is to divide the utilization and availability curves into intervals representing theoretical distributions of the population under various parts of the curve. The easiest interval to identify is the portion of the curve with a suitability (or availability) index of unity. By definition, 50% of the population falls between the upper and lower bounds of the variable within this range. The expected relative frequency at any interval within that range is 0.50 divided by the number of intervals (i.e., the average relative frequency). The number of intervals can vary, depending on the range and class width, but initially, the class width should be set equal to the precision of the instrumentation. This precision is usually 1 cm (or cm/sec) for metric instruments or one-tenth foot (or ft/sec) for English instruments. Therefore, if 50% of the population falls within the range of 1 to 2 feet, there would be 10 intervals, with an average relative frequency of 0.05 for each.

A variation of the same concept is used for tolerance limits containing larger portions of the population. The 75% tolerance limits also contain the central 50% of the population, so only 25% of the population has been added to the total. Assuming that this 25% is equally distributed outside the limits of the central (50%) block, two new blocks are added, each containing 12.5% of the population. The first of these extends from the lower limit of the 75% interval to the lower limit of the 50% interval. The second extends from the respective upper limits of these two intervals.

This process is continued for each subsequent set of tolerance limits. The 90% tolerance limit represents a net addition of 15% of the population beyond the 75% interval, or 7.5% of the population in each of the blocks between the lower and upper 90% to 75% tolerance limits, respectively. The

95% tolerance limits represent the addition of two 2.5% blocks, and so forth. It is assumed that the remainder of the population not accounted for within the tolerance limits occurs within the tails of the distribution between the last interval and the end of the range of the observations.

An example of this partitioning is shown for larval stages of the loach minnow (*Tiaroga cobitus*) in Figure 36. Expected frequencies are computed for each interval across the entire measured range of the variable, again using the precision specified above. The expected relative frequency for each interval is assumed to be the average for the entire block (i.e., the total population percentage within the block divided by the number of intervals). These expected relative frequencies are then arrayed, by interval, across the utilized range of the variable for both utilization and availability. The preference ratio for each interval is then computed, normalized, and preference indexes plotted at the midpoints of intervals having the same ratio. This computational procedure is shown in Table 9 for larval *Tiaroga cobitus*, and the resulting preference curve plot, as well as the utilization and availability curves on which it is based, are shown in Figure 37.

The assumption of equal proportioning of the 75%, 90%, and 95% tolerance limits about the central block could be questioned. The underlying principle of the tolerance limit approach, however, is that these limits delineate the central portion of the population at the specified percentage. Therefore, this assumption is consistent with the basic concept of tolerance limits. It is also common for the 75% and 90% (and even 99%) tolerance limits to fall on the same interval in asymmetrical distributions. A block where this happens would have an expected frequency equal to the area beneath both tolerance limits. For example, if the 75% and 90% tolerance limits fall at the same interval, the resulting block would extend from the lower 90% interval to the lower 50% interval (or upper 50% to upper 90%), and the expected frequency within the block would be 12.5% plus 7.5%, or 20%. This characteristic of tolerance limits somewhat corrects for asymmetrical distributions about the central block.

5.3 NONLINEAR REGRESSION

5.3.1 Univariate Analysis

Nonlinear regression is, in some respects, similar to hand-fitting curves to histograms as described in Section 5.1. The major difference is that an equation is used to fit the curve, rather than fitting it by eye. The analogy between nonlinear regression and freehand curve construction is that selecting the appropriate equation in the former is comparable to choosing the correct French curve for the latter. Figure 38 shows some of the more typical curve shapes encountered with habitat utilization frequency distribution, and their associated mathematical functions.

Once an appropriate function has been selected, a series of trials is made to determine the equation coefficients that will minimize the residual sum of square. This process is also similar to the one described for histogram analysis, but is much more orderly. Most nonlinear regression techniques use

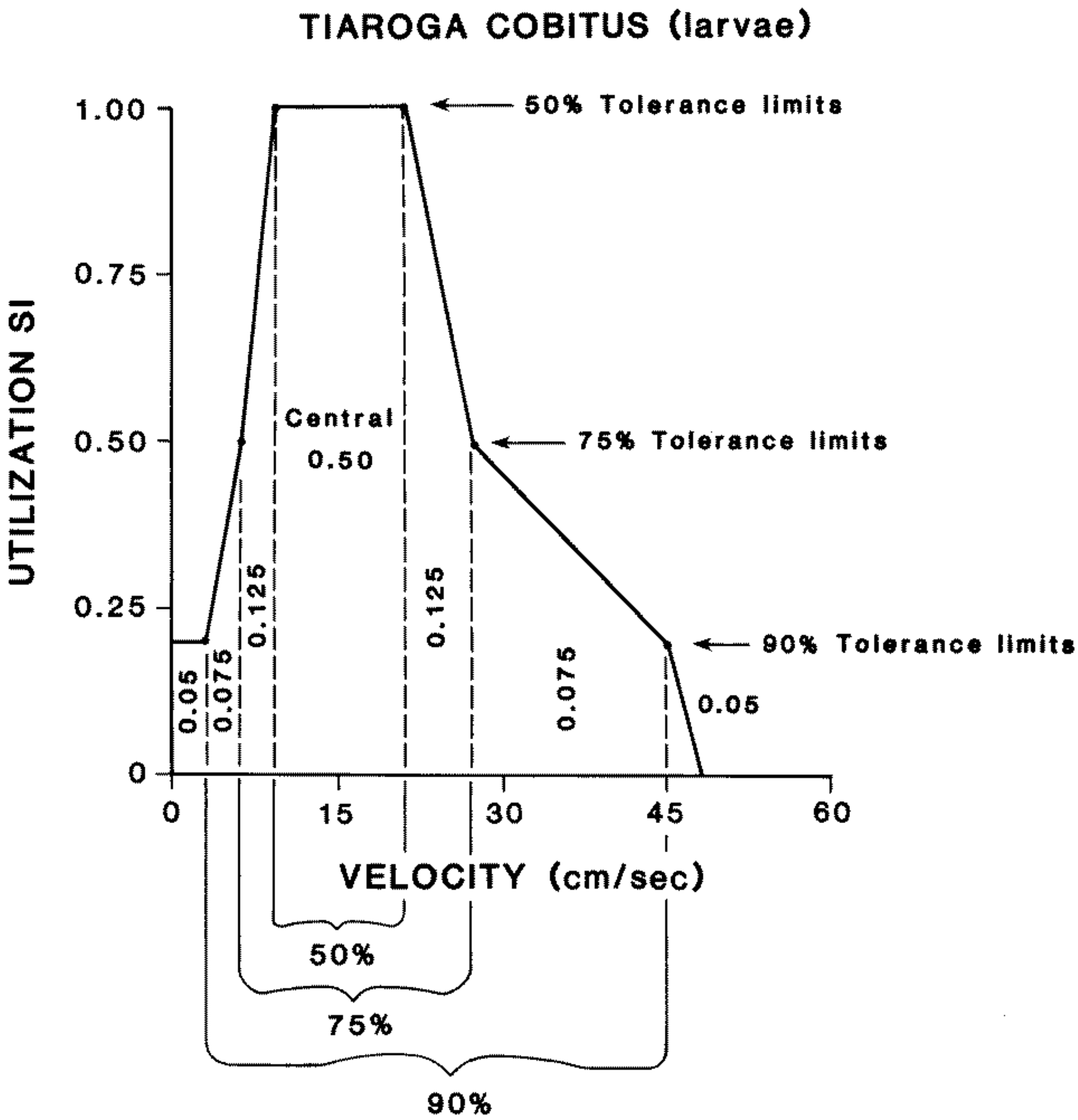


Figure 36. Partitioning of a velocity utilization curve developed from tolerance limits. Expected relative frequencies for each interval shown for each added increment of population.

Table 9. Preference ratios computed from estimated relative frequencies of utilization and availability as reconstructed from tolerance limits.

Velocity interval (cm/sec)	Utilization (expected F)	Availability (expected)	Preference ratios	Preference SI (normalized)
0-3	.05	.06125	0.82	.25
3-6	.075	.06125	1.22	.37
6-9	.125	.06125	2.04	.62
9-12	.125	.06125	2.04	
12-15	.125	.038	3.29	
15-18	.125	.038	3.29	1.0
18-21	.125	.038	3.29	
21-24	.0625	.038	1.65	0.5
24-27	.0625	.038	1.65	
27-30	.0125	.038	0.33	
30-33	.0125	.038	0.33	
33-36	.0125	.038	0.33	0.1
36-39	.0125	.038	0.33	
39-42	.0125	.038	0.33	
42-45	.0125	.038	0.33	
45-48	.05	.038	1.32*	
48-51	0	.038	0	0

*It is not uncommon for a tail to turn up with real frequency distributions either. This is usually an artifact of the data and requires judgement by the investigator regarding its validity. Usually, such upturned tails are ignored and the curve smoothed to its end point.

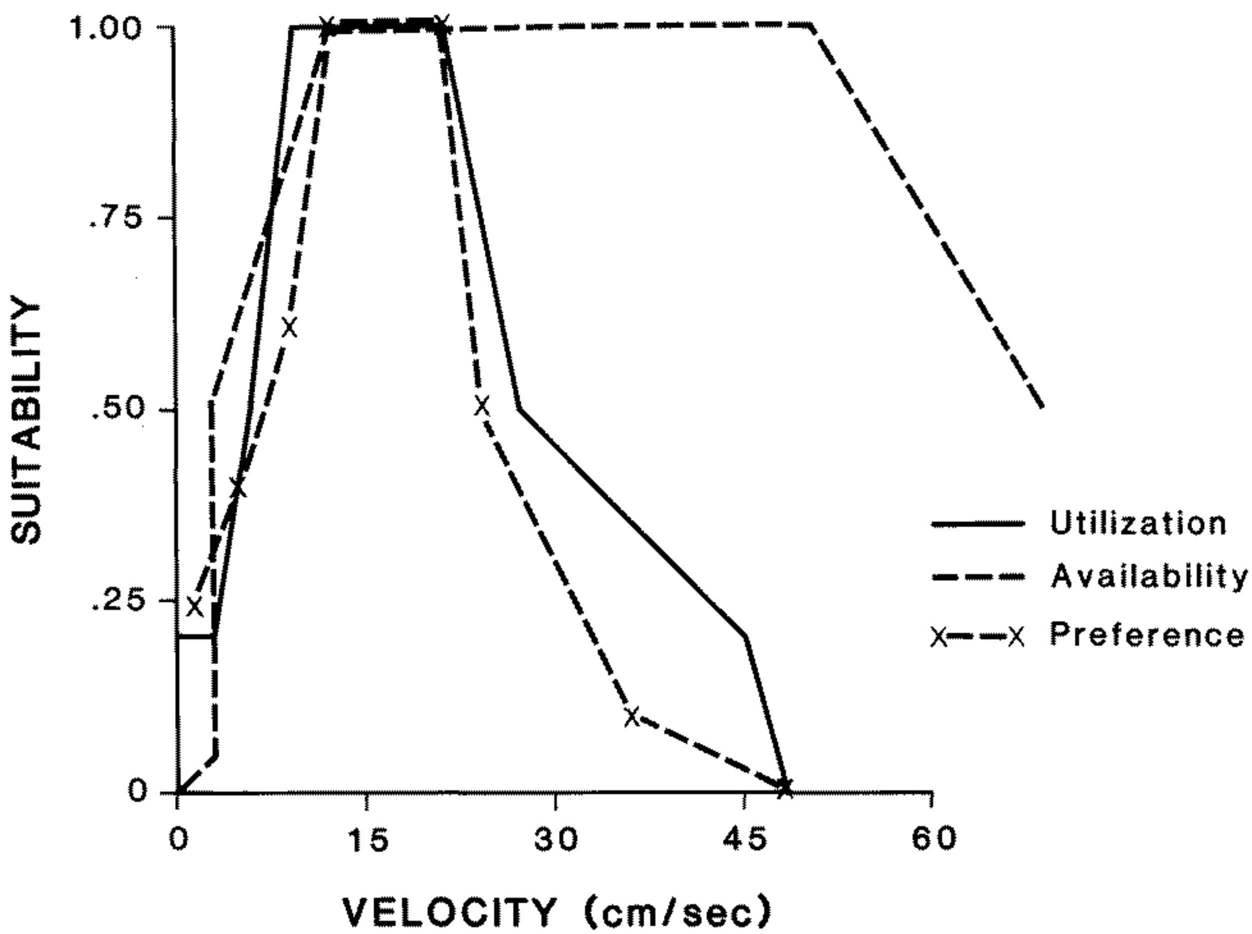


Figure 37. Velocity preference curve for Tiaroga cobitus, derived from tolerance limit curves for utilization and availability.

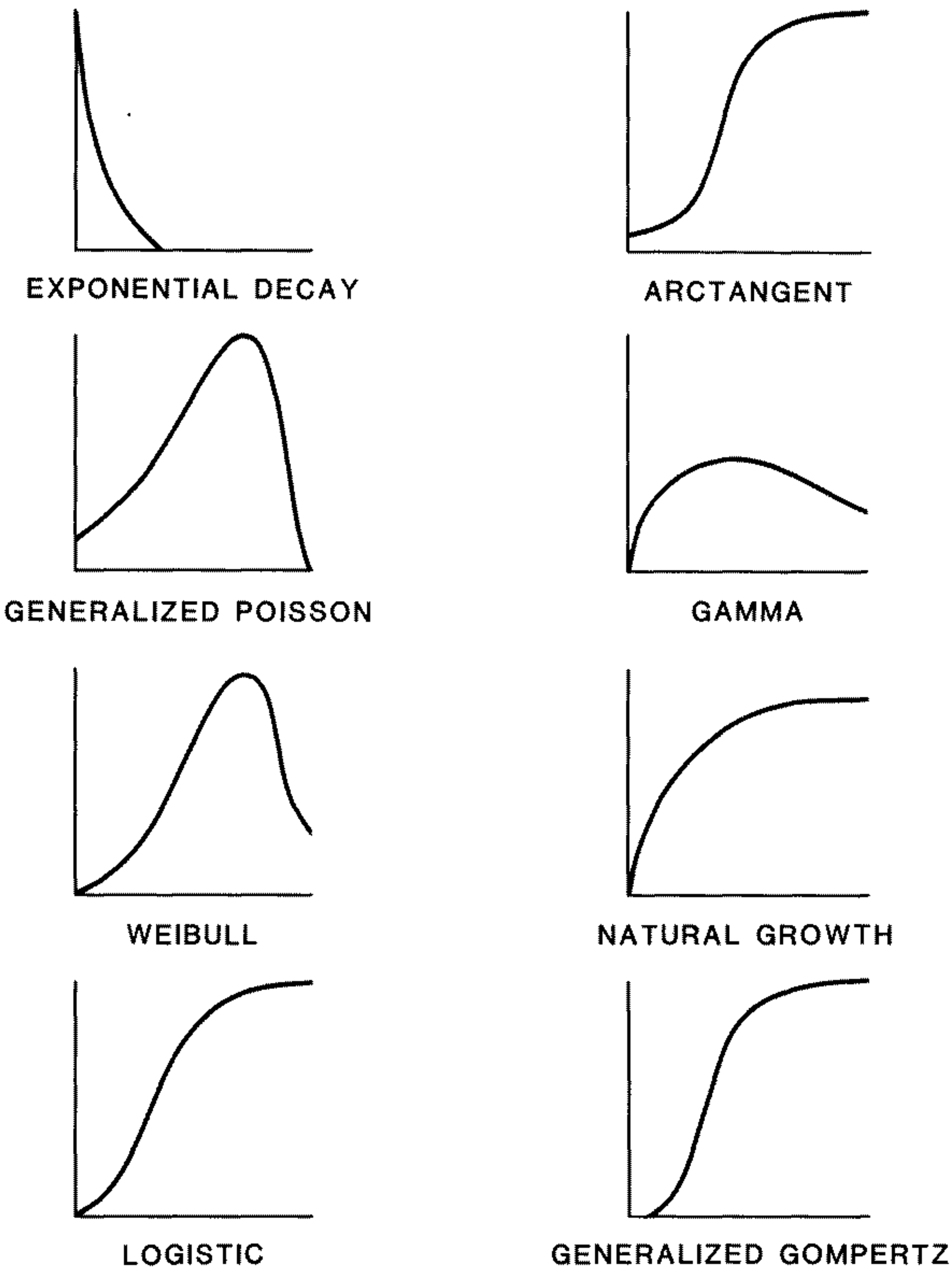


Figure 38. Some common curves and functional forms of frequency distributions encountered in habitat utilization and preference studies.

what are called gradient solution techniques to solve for the roots of an equation. Gradient techniques are based on the Newton recursion method, shown in Figure 39. Starting with a given approximation of one or more equation coefficient(s), the equation is solved over the range of the variable of interest, resulting in a curve through the data. At each interval within the range, the difference between the observed value (Y_i) and the predicted value ($\hat{Y}_i$) is computed and the sum of squares (or sometimes the mean square root of the sum of squares) is computed. A second trial is then made, by changing the values of one or more of the coefficients, solving the new equation, and recomputing the sum of squares. If the sum of squares from the second run is larger than for the first run, the direction of change of one or more of the coefficients is reversed. The direction of change is maintained if the second iteration produces a smaller sum of squares than the first run. This search procedure is continued until a minimum sum of squares is found, or is convergent (i.e., minor changes in the coefficients result in the same sum of squares).

There are numerous computer programs with nonlinear regression capabilities. One of several search procedures may be employed to find the least squares coefficients for an equation. Some of the less sophisticated programs use what is called the "brute force" technique. This method requires the user to specify an upper and lower limit for each of the coefficients. Each linear coefficient is varied, one at a time, holding all others constant for each trial. Then, both linear and exponential coefficients are varied, based on the best estimates of the linear coefficients. This is not a very efficient

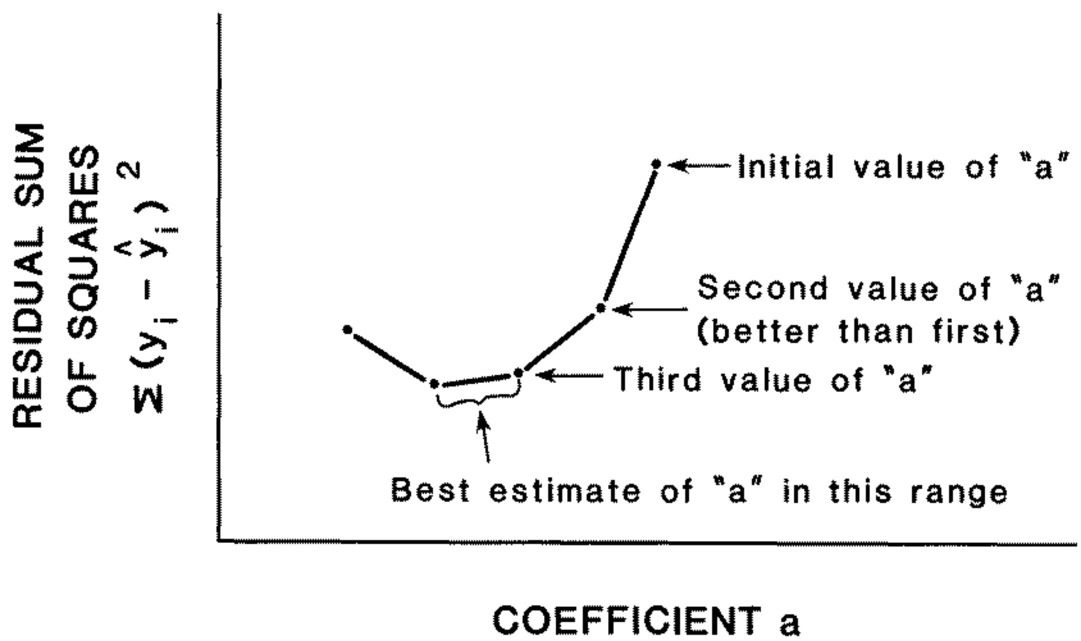


Figure 39. Estimation of coefficients to minimize residual sums of squares with nonlinear equations, by Newton's recursion method.

search technique, and becomes less efficient as the number of coefficients and the specified ranges increase. Obviously, the more the user knows about the effect of each coefficient on the shape of the curve, the more precisely the range can be specified, and the more efficient brute force becomes. The effect of a coefficient on the shape of a function may be illustrated in some statistics texts for some functions. One of the more useful sources for such illustrations is a technical report prepared for the U.S. International Biological Program by Parton and Innis (1972). This report contains many of the functional forms shown in Figure 38 (with the exception of the gamma and Weibull distributions). Among the functional forms illustrated by Parton and Innis (1972), the arctangent, generalized Poisson, logistic, and generalized Gompertz functions are commonly encountered in habitat utilization frequency distributions. Descriptions of these functions, parameter descriptions, and the effects of changing coefficient values are reproduced in Appendix C. Similar descriptions are available for the gamma and Weibull distributions in Walpole and Meyers (1972).

Some of the more sophisticated nonlinear regression programs use a search procedure known as the Newton-Raphson method. The Newton-Raphson technique allows the modification of several coefficients at once, using simultaneous solution of equations. Aside from the added complexity of simultaneous solution, Newton-Raphson is still a gradient method, following essentially the same concepts as the Newton recursion method. A complete description of the Newton-Raphson method is given by Matz (1978). Another technique, described by Marquart (1963), employs an interpolation of parameter estimates derived from a Taylor series expansion and from a gradient method. Marquart's technique is sometimes used as an alternative to the Newton-Raphson when convergence of the roots of the equation becomes painstakingly slow or appears to be unobtainable. A third search technique, called the Simplex algorithm, is described by Caceci and Cacheris (1984). The Simplex method is based on a geometric analysis of a multidimensional response surface, where the dependent variable is the residual sum of squares and the independent variables are the values assigned to the model coefficients. When two parameters are used the resulting response surface is three-dimensional and looks somewhat like a blanket suspended loosely at all four corners. The best estimates of both parameters are those where the response surface achieves the maximum amount of sag.

Zar (1974) categorizes computer programs with nonlinear regression capabilities into two groups. First are programs written especially to fit a particular function, or a group of functions. Second are general programs that can fit data to any of a wide variety of models. The former are usually much easier to use because the data processing sequences are more formalized and, consequently, more adequately documented. Also, because the functions are "canned," many of the intricacies of the statistics are internalized. This can be a blessing if the user understands the nature of the function used in the program, and uses the appropriate function for the frequency distribution. It can be a curse if the user treats the program as a black box, attempting to fit any or all frequency distributions to a limited number (or type) of functions, without understanding the nature of the function. The second type of program requires considerably more knowledge regarding each function to which data might be fit. It is often necessary to supply the model with the partial derivatives of the regression function with respect to each parameter. This attribute not only requires a stronger background in mathematics, but also a better understanding of computer language. Since the use of any particular function is not routine for these programs, documentation will often be found to be less than adequate. If the first type of program can be termed "user friendly," then the latter type can often be described as "user hostile."

A computer package, whether of the first or second type, should have certain capabilities and features besides being able to fit data to a nonlinear equation. One of the most important aspects is the ability to build, modify, and sort data files. Habitat utilization data should be rather voluminous. Therefore, the easier it is to input data and to manage data files, the better. Data entry systems allowing a user-specified format and interactive data entry are among the easiest and most flexible. These systems usually contain a subroutine allowing the user to specify the number, order, and location of data to be contained in a record. Data are then entered in "free format" from a terminal. (Free format means that each piece of information is separated by a space or comma, rather than being placed in specific columns.) Systems based on keypunched, rigidly-formatted data are nearly intolerable once a user has become accustomed to free format.

A second attribute of a good data management system is the ability to sort records according to the variables the user wishes to analyze. A typical record for habitat utilization data might contain information on species, size, time of day, date, frequency, depth, velocity, nose velocity, substrate code, cover code, and temperature. A good data management system would allow the user to select records for one species and one size range, for any particular time of day or year, and to conduct a frequency analysis against any or all of the microhabitat variables in the record. Some systems require the user to build separate files for each life stage of each species, for only one or two variables at a time. Such systems are not too inconvenient if they are interactive, but are certainly not as effective as those allowing records to be sorted.

Other considerations in selecting nonlinear regression programs include the types of functions to which data can be fitted and the search routine used by the program to find the least squares fit. It is quite possible to obtain a best fit of data to the wrong function. The flexibility to obtain best fits for several equations allows the user to select the one giving the overall best fit to the data. Zar (1974) states that the most important output from the program includes: estimates of the parameters in the model; the standard error of each of these estimates; and an analysis of variance summary including at least the regression and residual sum of squares and degrees of freedom. Many programs also require solutions of the first and second derivatives of the input equation, or partial derivatives for each of the coefficients. The better statistical programs have internal subroutines that solve for these derivatives. This is a desirable characteristic in curve-fitting packages, especally if the user does not have a strong background in calculus.

The search procedure used by the program is an important determinant of its ease of use. Many of the nonlinear regression programs available for microcomputers rely on a brute force method of optimizing parameters. The brute force method requires much less internal memory than a more sophisticated method, such as Newton-Raphson. It also requires a firmer hand at the controls on the part of the user. Programs employing Newton-Raphson or Marquart-type search procedures are essentially automatic, requiring very little user interaction. These, however, are usually only available on mainframe computers, although some of the more powerful microcomputers can also handle gradient solution techniques.

5.3.2 Multivariate Analysis

Fitting data to a multivariate function follows many of the same steps as univariate nonlinear regression. The principle difference is that the function selected to describe the frequency distribution contains terms for more than one variable and terms describing the interactions among the variables. One of the most common functions used to describe a multivariate response surface is the exponential polynomial of the form:

$$P_{(d,v)} = \frac{1}{N} e^{-(a_1 d + a_2 v + a_3 d^2 + a_4 v^2 + a_5 dv)} \tag{10}$$

where $P_{(d,v)}$ = a joint probability of utilization for a combination of depth and velocity;

a_i = least squares parameters for the terms v, d, and dv; and

N = a normalizing term reducing the area beneath the response surface to unity.

The term $a_5 dv$ in equation 10 describes the interaction between depth and velocity, determined from the frequency distribution.

The use of exponential polynomials as multivariate functions has been favored by many researchers because of several attributes:

1. They are flexible, generalized functions capable of fitting a wide variety of frequency distributions.

2. Transformation of the function to yield an integral equal to one is not needed (as it frequently is for linear polynomial functions).

3. The value of P (equation 10) can never be less than zero (another undesirable attribute of linear polynomials) (Voos 1981; Prewitt 1982).

Most biological data can be fit to one of four polynomial orders (Figure 40). The polynomial order is determined by the number of complete peaks defined by the histogram. A first order exponential polynomial is an exponential decay function; it has only half a peak, which occurs at the origin. A second order function has the outline of a bell shaped curve, with one complete peak. A third order function has half a peak at the origin and a complete peak somewhere in the middle of the distribution. Fourth order functions have two complete peaks. Third and fourth order exponential polynomials invariably represent bimodal distributions resulting from inadequate data stratification. Subdivision of the data base should be strongly considered when third or fourth order polynomials give the best apparent fit.

The response surface generated by an exponential polynomial can theoretically contain terms for any number of variables. Inclusion of more than four variables is discouraged, however, because of the massive amounts of computer time consumed.

One of the primary drawbacks to the exponential polynomial is that it produces a symmetrical response surface with a single maximum value. This tendency appears to hold regardless of which order polynomial is used. This function is fairly limited in the variety of shapes that the marginals (Figure 40) can take. Skewed distributions are difficult, if not impossible, to derive. The single maximum value of the exponential polynomial also makes derivation of functions with plateaus (such as arctangent, logistic, or natural growth functions) impossible. Since many depth suitability curves take this shape, exponential polynomials can usually represent only the shallow depth portion of the utilization curve, up to the peak of the curve. Beyond that, a suitability of 1.0 must be assigned, manually, to greater depths. Extensions of criteria functions are discussed in Section 6.3.1.

A second multivariate function, the logistic regression, has been proposed by Thielke (1985), as a method of deriving a suitability index. This approach has some features similar to the exponential polynomial. The logistic equation takes the form:

$$E\left(\frac{s}{n}\right) = \frac{e^{Bx}}{1 + e^{Bx}} \tag{11}$$

where $E\left(\frac{s}{n}\right)$ = suitability,

B = a vector of regression coefficients,

x = a vector of independent variables,

s = number of binary positive results, and

n = the number of trials or samples.

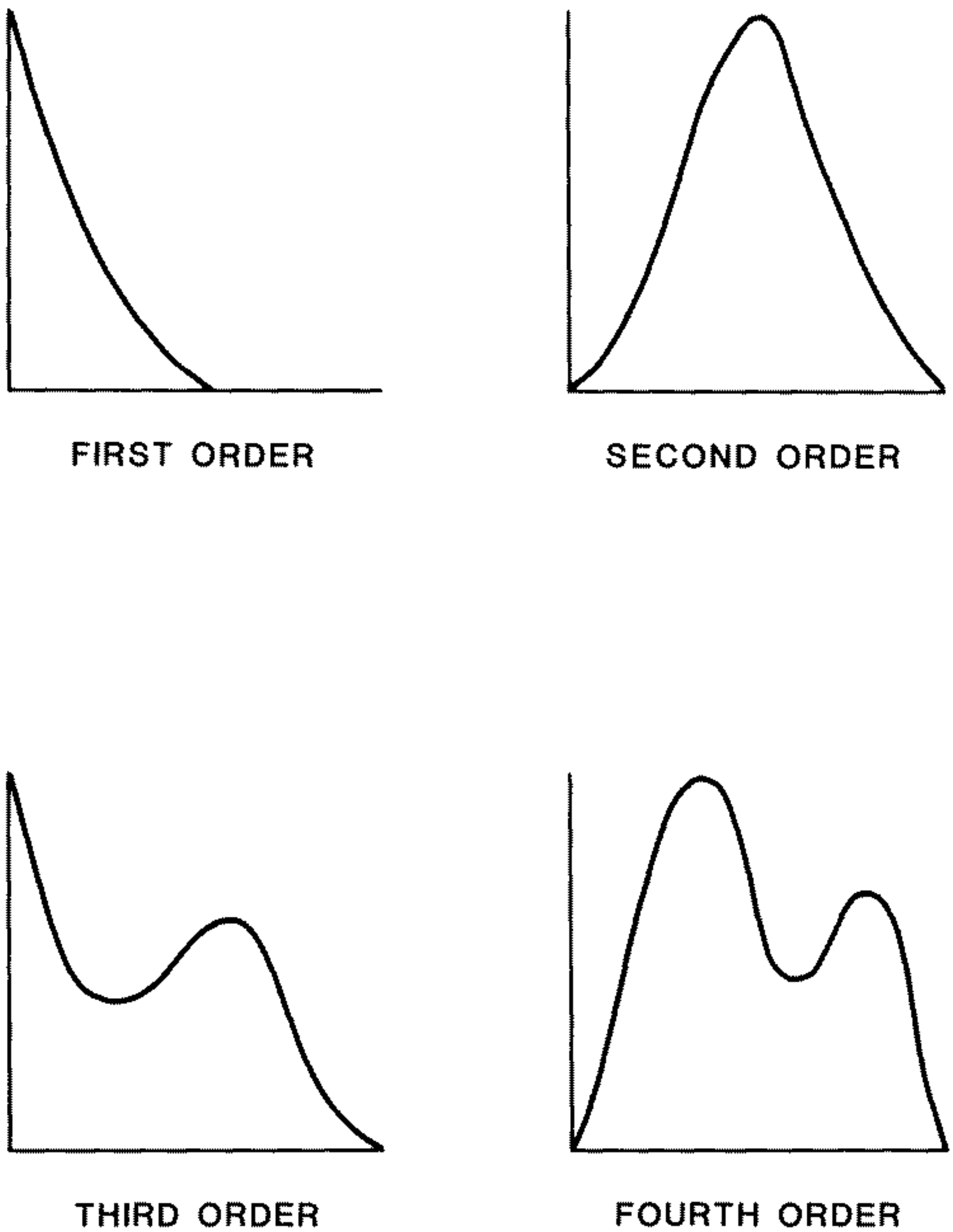


Figure 40. Shapes of multivariate response surfaces associated with various exponential polynomial orders.

The dependent variable is based on binary events, such as success or failure and response or no response. Since presence or absence (without considering frequency) can be used as a binary variable, $E\left(\frac{S}{n}\right)$ can be considered a suitability index. In a stepwise manner, logistic regression estimates the vector of parameters (B_i) where x represents one or more independent variables (Hill 1984). In other words, the term e^{Bx} is an exponential polynomial.

The principal difference between this formulation, and equation 10 is the form of the probability density function. Negative parameters (B_i) in equation 11 result in values of $E\left(\frac{S}{n}\right)$ that approach zero, and large positive parameters (B_i) cause the values of $E\left(\frac{S}{n}\right)$ to approach unity. Therefore, it is the shape of the response variable, $E\left(\frac{S}{n}\right)$, that takes the form of a logistic curve. In other words, the comparison between a probability density function from an exponential polynomial and $E\left(\frac{S}{n}\right)$ is analogous to a comparison between a histogram and a cumulative frequency distribution.

Equation 11 has many of the characteristics desirable of a multivariate suitability index: it produces values between zero and one, and negatives cannot be derived, several variables can be used simultaneously, and cross product terms can be included in the calculation of $E\left(\frac{S}{n}\right)$. It is not clear, however, exactly what kind of a suitability index $E\left(\frac{S}{n}\right)$ represents. Because the equation is solved using data from places fish were and were not found, $E\left(\frac{S}{n}\right)$ might be a preference function. $E\left(\frac{S}{n}\right)$ does not explicitly solve equation 7 (i.e., $P = U/A$), however, and hence is not a preference function in the context of previous definitions used with the IFIM. The context of $E\left(\frac{S}{n}\right)$ in terms of utilization, preference, or suitability may become more apparent through additional use and comparison with more traditional indexes. At this time, more research is needed to determine what $E\left(\frac{S}{n}\right)$ actually represents.

5.3.3 Computation of Preference Functions

Section 5.1 presented two options for developing a preference curve from utilization and availability frequency distributions. One way is to develop a smoothed curve for both distributions and divide the predicted frequencies from the respective (unnormalized) curves to obtain the preference curve. The second way is to develop a preference histogram by dividing the observed utilization frequencies by the observed availability frequencies, and then fit a curve to the resulting histogram. A third technique can sometimes be used when utilization and availability equations are derived: the equation for utilization can be divided by the availability equation, resulting in an equation describing preference. This can be plotted as a preference curve or response surface. This technique is limited, however, to equations of the same type or order. Dividing an arctangent utilization function by a Poisson availability function may have totally unpredictable results. Furthermore, coefficients with the same symbols might have totally different meanings in two different functions and cannot logically be divided. In these cases, the recommended approach is to obtain the best fit function for both utilization and availability, draw both curves, and then divide the respective predicted frequencies to obtain the preference curve. Unfortunately, division of equations must be used with multivariate equations. This may cause mathematical instabilities in the preference response surface. An investigator using this approach should always obtain a graphical representation of the response surface for visual examination.

5.4 DISCUSSION

This chapter has outlined several approaches by which a graphical representation of a frequency distribution can be developed. The selection of any particular approach should be based on several conditions and considerations: foremost among these are accuracy and precision. Other considerations include availability and access of statistical software, time, cost, required expertise, and ease of use of the final product. The variety and supply of statistical programs has been expanding rapidly over the last several years, parallel to the expansion of the market for microcomputers. A good way to determine the availability of appropriate statistical packages is by referring to reviews and advertisement indexes in computer-oriented magazines. Also, various user groups may have knowledge of both commercial and public-domain software with nonlinear regression capability.

Regardless of the curve-generating technique selected, transferring data from field forms to computer files is highly recommended. This reduces the space occupied by the data and makes retrieval, sorting, and management of data easier and more efficient. At a minimum, a computer package should:

(1) allow interactive, free-format file building,

(2) allow sorting of data by record and by variable, and

(3) produce an ordered frequency array of the data (e.g., frequency of observations vs. each increment of each variable).

These characteristics will at least allow the development of curves through histogram analysis or tolerance limits.

Nonlinear regression is superior to histogram analysis in terms of accuracy and mathematical acceptability, provided the investigator has chosen the correct function to fit the data. A function should not be used, however, simply because it is the only one available with the software at hand. A desirable attribute of a statistical software package is the ability to evaluate the goodness-of-fit to several nonlinear functions. Provided the investigator has access to the appropriate programs, nonlinear regression is a relatively easy method of constructing curves.

The tolerance limit approach is a quick and easy method for developing curves and, in some cases, may actually be better than nonlinear regression techniques. This approach is especially valuable when developing curves from small data bases. The conservative nature of the tolerance limits is designed to include the specified percentage of the population (not the sample) at a given confidence level. This means that the interval under which 50% of the population will occur is a conservative estimate (i.e., covers a broader range of conditions). The smaller the sample size, the more conservative the estimate.

This characteristic of tolerance limits tends to produce broader curves than regression or histogram techniques, but this may sometimes be advantageous. It is quite common for utilization data, collected in different streams or under different conditions, to exhibit dissimilar frequency distributions. Sometimes, the dissimilarity can be traced to different habitat availability, different gear, or different effort. The disparity, however, can often be traced to small sample sizes. Tolerance limits are not highly influenced by internal variations in the frequency distribution, so the disparity in utilization curves developed in different streams is often reduced appreciably. Because tolerance limits result in a higher degree of interstream agreement and more robust curves, such curves may be easier to defend in adversarial situations.

Tolerance limits may be the preferred approach to use with small sample sizes, but they should not be used to fit data with bimodal distributions. As previously mentioned, a bimodal distribution may signal some sort of biologically induced interaction between variables. Such data should usually be partitioned and conditional criteria developed. It would be a mistake to apply tolerance limits to a bimodal distribution, because the resulting curve will encompass both peaks, and assign a suitability index value of 1.0 to the trough between them. Bimodal distributions should not occur if the data are stratified correctly. If one occurs, it may be an indication that the sample size is too small or that there is another intervening variable within the stratified data base. The cause of the bimodal distribution should be resolved before proceeding with curve development.

The use of univariate curves in PHABSIM assumes the absence of significant interactions among the variables represented by each of the curves. Critics of the methodology (Mathur et al. 1985) point out that multiplication of univariate probabilities is only valid when the two probabilities are independent. Unfortunately, the nature of the independence assumption is nearly universally misunderstood and misinterpreted.

Three factors can lead to apparent correlations between variables in multivariate regression analyses. The first is the collection of data in hydraulically simple channels, where a physical correlation exists between depth and velocity. A utilization function developed in such a channel will contain the same degree of correlation found in the environment from which the data were collected. The second is failure to correct for environmental availability (i.e., developing category II data instead of category III). When correlations appear between depth and velocity in the utilization function, the same cross product invariably appears in the availability function. The two cross products cancel out when the preference function is calculated (K. A. Voos, Woodward-Clyde Consultants, San Francisco, CA; pers. comm.). Correlations between depth and velocity are physical phenomena, not biological ones, and should not be attributed biological significance when they are merely artifacts of the data. The third is that the interaction may be the result of behavioral selection on the part of the fish. One conceivable instance where a correlation between depth and velocity might be important is the case of fish utilizing surface turbulence as a form of overhead cover. In such a case, it could be expected that fish would be found in shallow, fast water, but not in shallow, slow water.

There are two important considerations regarding depth-velocity interactions. First, are the interactions real (i.e., do the fish select microhabitats on the basis of the interaction) or are they spurious? Second, if they are real, will treating the two variables independently introduce an unacceptable amount of error into the results?

The first consideration can be evaluated by two different approaches, both of which test essentially the same hypothesis. If the correlation between depth and velocity is spurious, then the same relationship should be apparent in both the utilization data base and the availability data base, i.e., both sets of data are from the same population. If the data were originally fit to exponential polynomials, the coefficients on the depth-velocity cross products can be used as indicators of this phenomenon. Both coefficients will have about the same magnitude and the same sign, if the correlation is spurious. In such cases, the two terms would cancel one another out when the preference function was computed. An important interaction is suggested when the two cross product coefficients have different signs. This event is conceivable, but based on data sets analyzed to date, such an occurrence would be unusual. Should an investigator find reversed signs on the cross product terms, the multivariate preference function should probably be used instead of univariate curves. The investigator, however, should also attempt to discover the underlying cause of such a phenomenon. If no biological basis can be explained, then a corroborative study should be conducted to determine whether or not the phenomenon is intrinsic to the species. Until the true nature of the resultant preference cross product is understood, the function should probably not be used for general applications of the IFIM.

Sometimes, the two cross product coefficients (i.e., utilization and availability) will have different numerical values, although they have the same sign. Generally, a spurious correlation is indicated when the cross product term in the numerator (utilization) is larger than the cross product term in the denominator (availability). A potentially significant correlation is indicated when the cross product for availability exceeds the one for utilization by more than 50%.

The significance of depth-velocity interactions can also be tested by comparison of the simple linear regressions between depth and velocity as they appear in the utilization and availability data bases. If both sets of data represent the same phenomenon (i.e., the interaction is due to physical correlation and not behavioral selection), then a regression between depths and velocities in the utilization data base will have the same (or nearly the same) slope and intercept as a depth-velocity regression of the availability data base. Student's t statistic can be used to test the hypotheses that:

(1) the two slopes are not significantly different, and

(2) the two intercepts are not significantly different.

Details of this test are given in Zar (1974:228-230) and other statistics textbooks.

Although PHABSIM will accept biological criteria in either univariate or multivariate format, the former are much easier to use. Therefore, a user may opt for univariate curves, even though there appears to be a significant correlation between depth and velocity in the preference or utilization function. Sometimes, this is acceptable, and sometimes it is not. The primary questions that determine the acceptability of violating the assumption of independence are the following:

(1) What category (II or III) of criteria does the function represent?

(2) How much error will be introduced into the results?

A depth-velocity correlation in a utilization function may be a harmless artifact of the sampling environment, but without availability data, the significance or insignificance of the cross product cannot be determined. This leaves the user in a quandry, because the error involved in assuming dependence (i.e., incorporating the cross product in the suitability index) when none really exists is as bad or worse than assuming independence when the correlations are real. Orth and Maughan (1982) found that inclusion of the cross product improved the fit of their data, as judged by the mean square error. They developed exponential polynomial utilization functions with and without the cross product terms, and found that the mean square error decreased by 3% to 19% for various species, when the cross product was included. A better fit to data collected in one stream, however, does not necessarily mean that a better model has been developed for the species in general. A utilization pdf developed in a hydraulically simple channel is likely to contain a rather large depth-velocity cross product term. Applying the same utilization function to a diverse channel may result in an underestimation of WUA, because the habitat will not exhibit the same degree of depth-velocity correlation. Therefore, utilization functions containing large cross product terms should <u>not</u> be transferred to streams having a more diverse habitat than the stream from which the criteria data were obtained. There is also some question regarding the applicability of such criteria in the same stream in which they were developed.

Voos (1981) concluded that the development of the preference function resolved many of the potential problems of variable interactions apparent in utilization data. Experience to date appears to confirm Voos' conclusion that the preference pdf is usually orthogonal (or nearly orthogonal), regardless of the cross product term in the utilization pdf. Nevertheless, the possibility exists that the preference pdf might contain an interactive depth-velocity term. The question then arises: "If variable interactions do affect the computation of WUA, at what levels of correlation are results likely to become unreliable?" Prewitt (1982) conducted an extensive sensitivity analysis of PHABSIM, using five guilds of hypothetical preference functions with varying degrees of depth-velocity correlation (r values ranging from 0 to 0.8). He found that the effect of the correlation was most pronounced in small, simple channels using criteria for rather cosmopolitan guilds (i.e., species with broad habitat preferences). Under these circumstances, a depth-velocity correlation of 0.2 in the preference function was sufficient to reduce the reliability of the results beyond acceptable limits (± 20%). Conversely, in medium to large, complex rivers, the depth-velocity correlation could be as high as 0.6 before the results deviated beyond these limits. Also, species having very restrictive curves appeared to give reliable results across a very broad range of stream sizes and channel configurations. These guidelines are useful, but users should be provided with enough information (i.e., multivariate functions with and without the cross product) to conduct their own sensitivity analysis. The use of sensitivity analysis to evaluate criteria is discussed in Chapter 6.

The third factor leading to cross variable interactions is a genuine behavioral response by the fish. Most typically, these interactive terms are strongest between mean column velocity and substrate size or instream cover objects, and between depth and overhead cover. Interactive terms between substrate size and velocity can be treated either as continuous variables in a bivariate model (such as an exponential polynomial) or by treating substrate size as a categorical variable, used in the form of conditional univariate criteria. If the use of certain depths or velocities appears to be influenced by cover type, conditional criteria must be developed, because cover is not a continuous variable.

Although it is sometimes possible to detect interactive behavior by a bimodal distribution, this is an unreliable indicator. There are two problems with a posteriori stratification. First, unless the data base is large, bimodal distributions may be deceptive. In one case, a bimodal distribution may not be apparent, even though the data should have been stratified. In another case, a bimodal distribution might be apparent, but is not caused by interactive behavior. Second, only one of the strata might have enough samples on which to base a curve. In this case, the stratification would be made and the research team would have to return to the field to collect data for the under-represented stratum.

A disguised bimodal distribution can occur when the preferred range of a variable, when associated with one cover type, is overlapped by the preferred range for that variable when associated with another cover type. An example would be a preferred depth range of one-half to two meters with cover type A and one and a half to four meters with cover type B. The resultant histogram may show nearly equal frequencies across all depths from one-half to four meters.

Apparent bimodal distributions are often the result of too few samples taken of a schooling species of fish. Large aggregations of fish, though they may be indicative of good habitat, can nonetheless create a very irregular histogram that gives the illusion of a bimodal or polymodal distribution. Although certain histogram smoothing techniques can be attempted, the best way to smooth it may simply be to collect more data and fill in the gaps.

Most of these problems can be avoided by judicious data stratification at the study design phase. It is generally advisable to assume that certain types of conditional behavior will be exhibited by the target species and to anticipate these by selecting strata before going to the field. If it is found that there is no significant difference between depth or velocity curves developed in association with different cover or substrate types, the worst connotation is that the researcher collected more data than were absolutely necessary.

In conclusion, the data analysis and documentation phase of criteria development should follow this general outline:

1. The category of the criteria must be specified; if category II, the axis representing relative suitability must be labeled "utilization" on all graphical output; if category III, the suitability axis should be labeled "preference."

2. If the criteria are category II, provide the best written description of the study area possible. Include in the documentation all photographs taken at the study area. If the criteria are category III, provide the above and include graphical representations (histograms, smoothed curves, and/or multivariate response surfaces) of the availability distribution(s).

3. Build separate data files for each data stratum: species, size class, season, activity, time of day, cover type, species association, or other, as appropriate.

4. Evaluate the degree and significance of depth-velocity correlations in the utilization data base by comparing them with the availability data base for each stratum. This can be done by developing utilization and availability pdf's and subtracting the availability cross product term from the utilization cross product, or by comparing the slopes and intercepts of the depth-velocity regressions using the Student's t test (Zar 1974). If availability data were not collected, determine the correlation coefficient and significance level of the correlations between depth and velocity in each utilization data base.

5. If it is concluded that depth-velocity correlations in the utilization data base are artifacts of availability, proceed with univariate curve construction. For small sample sizes ($n < 100$), nonparametric tolerance limits are preferred over nonlinear regression techniques. For large samples sizes ($n > 300$), nonlinear regression or histogram analysis is preferred. Both techniques should be used for intermediate sample sizes, and the resulting curves compared. Some judgment will be required regarding the quality of the two curves, and sometimes they will be averaged. Both curves should be similar, however, with the nonlinear regression curve nested inside the one developed by tolerance limits. If the regression curve appears as a very small subset of the tolerance limits curve, it is an indication of an inadequate sample size.

6. Compute the preference function by division of expected relative frequencies from smoothed utilization and availability curves. With large sample sizes and smooth histograms, development of a preference histogram is also satisfactory. Division of nonlinear regression equations is discouraged unless the same type of equation is used to describe utilization and availability (division of exponential polynomials, for example).

7. If it appears (from step 4) that the preference function contains a significant depth-velocity cross product, evaluate possible causes of the phenomenon (such as the use of surface turbulence as a form of overhead cover). The data should be fitted to a multivariate function if the cross product appears to have a behavioral cause. At this time, the characteristics of the exponential polynomial as a suitability model are better known than the logistic regression approach, although the latter may also be applicable. Both multivariate response surfaces and marginal projections (two dimensional profiles of the response surface) or independently derived univariate curves should be produced if the nature of the cross product is inexplicable. In this case, a corroborative study in another stream system is strongly recommended. The purpose of such a study would be to substantiate the biological validity of the cross product term.

8. Provide complete study documentation, including all pertinent information from the plan of study (e.g., study purpose and objectives, data stratifications and sampling protocol, coding systems, sampling methods, sample sizes, and analytical techniques). If the implementation of the study did not deviate substantially from the study plan, the plan may be substituted for formal documentation (but note any exceptions between the study plan and the implemented study). Provide criteria graphics, including histograms, univariate curves, or bivariate response surfaces. Truncate curves or response surfaces at the maximum and minimum values of habitat availability or, if curves are extended by professional judgement beyond the range of the data, indicate the extended portion with a dashed line. This will allow future users to evaluate the range of a variable for which data were collected versus the range that must be based on judgement and interpretation.

6. CRITERIA EVALUATION

The criteria used in an IFIM application will often originate from streams other than the one(s) being evaluated. Some investigators conducting instream flow studies have suggested that site-specific criteria must be developed for correct application of the methodology. While this suggestion appears logical, site-specific criteria are not always the most appropriate for a particular stream.

One obvious consideration in the use of offsite criteria is cost, especially if no site-specific information is available and the data base must be completely assembled from empirical observations. It is not unusual for empirically derived criteria to cost as much as $10,000 per species, and more if the species is particularly difficult to study (e.g., highly migratory or rare species, difficult access to study sites, expensive tracking and locating technologies). The less that is known about the species at the outset, the higher the cost, because it often becomes necessary to expend large amounts of effort to determine elementary information regarding life histories, movement, and basic life requisites.

The second consideration is that it takes time to develop criteria. Applications of the IFIM generally fall into the category of operations and management, meaning that results are expected nearly immediately. Criteria development often becomes the responsibility of the user, but this task more appropriately belongs with research. Operational studies are nearly always constrained by tight deadlines, many being completed in a matter of weeks or months. Habitat preference studies can rarely be completed in less than a year, because it takes at least that long to be able to observe all life history phases and seasonal habitat usage by juveniles and adults. Unusually high or low water years may alter the periodicity and the normal habitat utilization patterns of the fish, thereby upsetting the schedule of the investigator. The need to collect additional data, to further partition data, or to alter sampling strategies often becomes apparent only after a certain amount of effort has been expended. The results of a study can be put off a year simply by missing the spawning run by a week. One of the biggest mistakes made by operations-oriented biologists has been to underestimate the time it will take to assemble criteria, even by category I methods. Consequently, large amounts of resources may be devoted to other aspects of the instream flow study, only to find that these data are virtually worthless without adequate criteria. Addressing the problem of the availability and adequacy of the criteria should be one of the very first activities of a project scoping exercise. If criteria are not available or are inadequate, the project developer needs to know that the instream flow study will not be completed very quickly. A preproject negotiation may follow to determine the type and quality of criteria that can be developed under different deadlines and budgets.

A final consideration is the adequacy of criteria data, for use in the streams being analyzed with the IFIM. Chapter 2 provided a lengthy discussion regarding the desirable attributes of streams from which habitat criteria are developed. These included such factors as high habitat diversity, good water quality and temperature conditions, high standing crop of the target species, presence of competitors, good food supply, and gradually varying streamflow conditions. Instream flow studies are often initiated on a stream because one or more of these conditions is not being met under current water management practises. A habitat preference study should not be initiated in such streams because the resulting criteria will probably be meaningless, if they can be developed at all. The influence of habitat availability on both the utilization and preference functions has been demonstrated repeatedly in previous chapters. It should be easy to see that criteria developed in a low quality habitat could lead to erroneous conclusions regarding the streamflow requirements of a species. Therefore, criteria should not be developed in the stream that is under an operational IFIM study unless it meets the conditions specified in Chapter 2.

The three limitations of cost, time, and adequacy mean that most of the criteria used for instream flow analyses will not (perhaps, should not) be site specific. Indeed, the IFIM would soon become a tool for only financially well-endowed natural resource agencies and interest groups if site-specific criteria were needed for every study. Developing these criteria should not be the responsibility of the IFIM user; but acquisition, evaluation, and, sometimes, modification of existing criteria are in the user's jurisdiction. The user may be assigned development responsibilities if criteria are not available or judged inadequate, but this must be determined early in the IFIM study design phase.

Evaluations of offsite criteria can take several forms, ranging in complexity from a simple review process to an empirical verification study. The goal of this evaluation process is to ensure that the criteria are appropriate for the stream and the water management problem being investigated. The foremost decision factors for evaluating criteria are comprehensiveness and accuracy. The adequacy of the former can be determined by a quick review, but the latter may require more detailed analysis.

6.1 REVIEW FOR COMPREHENSIVENESS

During the IFIM study planning process, decisions are made regarding the target species and life stages of concern, what variables to include in the habitat analysis, how these variables should be described, and how criteria should be stratified by season or activity. These decisions should guide the selection and evaluation of criteria, but unfortunately, the study planning process often operates in reverse, i.e., the habitat analysis is predicated on what criteria are available rather than on what criteria are needed.

Sometimes, the criteria evaluation process will uncover gaps in the information base stage for which criteria must be developed, such as a missing life stage. Therefore, one of the first steps in a comprehensiveness review is an evaluation of the data stratifications used in the criteria study.

The second category of comprehensiveness is related to the sampling protocol. Perhaps criteria exist for all the appropriate size classes and activities, but certain variables were not measured in a manner consistent with the needs of a particular instream flow study. The most important variables to evaluate in terms of sampling protocol are velocity, substrate, and cover.

A protocol evaluation should first determine whether the velocity criteria represent mean column or nose velocities and, if the latter, at what distance from the streambed they were measured. Nose velocity criteria are generally more transferable to streams of different sizes than are mean column velocities. When using nose velocity criteria, it is also important to evaluate whether the nose depth is appropriate for the target species and life stages. For example, the nose depth for chinook salmon appears to increase from California to Alaska, possibly because the average size of the adults increases over that range. Stream size is a much more important consideration when using mean column velocity criteria. It is generally not a good idea to use mean column velocities in streams that differ significantly in size from the criteria source stream. The greatest errors occur when such data are transferred from a small source stream to a large study stream, but problems may arise in the reverse situation, as well. Obviously, the most important size comparison between the two streams is the average depth, because the depth determines the point(s) in the water column where the velocity is measured. If the average depth of the study stream is appreciably greater than the source stream, the velocity criteria may not be applicable there. Two additional factors should be considered, however, before rejecting the criteria on this basis. First, if the life stage of interest is confined to shallow-water areas, mean column velocities may be acceptable despite the size difference of the two streams. Second, mean column velocities measured in a small stream may be equivalent to nose velocities in a large one. The average "nose depth" can be estimated by multiplying the average depth of the fish observations by 0.4. (This is the approximate distance from the streambed at which the measurements would have been made.) An investigator should determine the amount of variance in the depth measurements, however, before accepting this transformation. If the coefficient of variation of the depth observations exceeds about 25%, this transformation is not advisable.

There is an alternative transformation that might be acceptable if the coefficient of variation exceeds 25% and no other criteria are available. Find the central 50% of the depth range, using nonparametric tolerance limits as discussed in Section 5.2. The average depth is then computed for this portion of the utilized depth range, and nose depth estimated by the 0.4 multiplier. This may result in an increased distance between the streambed and the assumed nose depth, compared to using all the depths in the average. Consequently, the computed WUA may be an underestimate of the actual usable area, but the error is less than if mean column velocities were used directly in PHABSIM.

The second important evaluation element is the substrate code employed in the development of the criteria. Complex substrate codes can always be simplified if the level of detail in the criteria exceeds the needs for a particular analysis. Transformation of a simple code into a more complex one is more difficult, but sometimes possible. The main points to consider are how the investigator described the dominant particle size (e.g., by size, abundance, or some combination) whether the code incorporates an index of embeddedness or a percentage of fines in the substrate matrix. The acceptability of a substrate code is contingent on its intended application. A study of aquatic macroinvertebrates would normally suggest that the substrate code contain a descriptor for percent fines, for example. If the stream in which the criteria are to be applied has a uniformly clean substrate, however, the absence of such a descriptor would not pose much of a problem. The need for a complex substrate code is greatest in study streams having a wide variety of substrate types, when the analysis includes a species that is highly affected by this variability. It may be necessary to modify a set of simple criteria by professional judgement or by supplemental data collection if it does not include the desired level of detail. Such criteria should not be rejected outright, however, because the information on depth, velocity, and dominant particle size may be entirely appropriate.

The evaluation of cover-related criteria is similar to that of substrate codes. The main distinction is that an inappropriate system for analyzing cover may be impossible to rectify and, therefore, may potentially invalidate a complete set of criteria. The most important consideration is whether the criteria were stratified by conditional cover types when this type of data is necessary. A species that is suspected of changing its microhabitat utilization as a function of cover type should be analyzed with cover conditional criteria. If conditional criteria are not available, a verification study to determine the accuracy of the criteria is advisable. PHABSIM results can be very sensitive to the way that cover is treated in the model, so it behooves the investigator to apply the form that is most representative of the target species.

6.2 EVALUATIONS OF ACCURACY AND PRECISION

Two types of evaluations can be conducted to assess the applicability of a criteria set to a particular stream. The first is a screening-level review of the study plan and results to determine the overall quality and transferability of the criteria. The second is one of several types of verification studies that attempt to replicate the results of a criteria study in the stream under investigation. In essence, these two types of evaluations are mutually exclusive. The verification approach is much more definitive, so if this approach is to be used, there is little to be gained by a screening level review. Verification studies are also more costly and time-consuming, and may not be feasible for some IFIM applications. All applications of the methodology, however, should include some type of criteria quality control. If the study is highly controversial or subject to litigation, a verification analysis is strongly recommended.

6.2.1 Quality Review

The quality of criteria can be affected by any of the factors of study design, sampling strategy, sampling methods, or data analysis that have been discussed in previous chapters. The evaluation is facilitated if the investigator has thoroughly documented each step in the design, implementation, and completion of the study.

The first factor a potential criteria user should evaluate is the diversity of the source stream from which the data were derived. This is one of the most important sources of bias in a criteria data base and is especially critical with category II data. One of several diagnostic indicators can be used for this evaluation, depending on the amount of information provided by the original researcher. The easiest is a comparison of the utilization, availability, and preference curves, all plotted on the same graph. The preference curve should appear as a subset of the utilization curve; the more similar they are, the more likely the data were derived from a highly diverse environment. Both curves should be considered suspect if they are radically different from one another. The second diagnostic test is to examine the raw frequencies of the availability curve. The availability histogram from a highly diverse stream should have a fairly uniform frequency distribution for any particular variable. Furthermore, there should be little correlation between depth and velocity. There is probably no stream on earth that will absolutely meet this standard, but the closer the source stream approximates the ideal, the better. Prewitt (1982) suggested that the coefficient of variation for either depth or velocity was a good indicator of channel complexity. Complex streams, as Prewitt defined them, always had a coefficient of variation greater than 50%.

Another indicator of habitat heterogeneity is a habitat diversity index. Gorman and Karr (1978), in a study relating habitat diversity to species diversity, used the Shannon-Weiner index:

$$H' = \quad P(i) \log P(i) \tag{12}$$

where H' = the index of diversity, and

$P(i)$ = the probability of occurrence of a particular interval of a variable or combination or variables.

Based on the descriptions provided by Gorman and Karr, complex (heterogeneous) environments usually had a value of H' greater than 3.0 when $P(i)$ contains terms for three variables, or 1.0 to 1.5 for $P(i)$ representing any single variable. These guidelines apply only to the Shannon-Weiner index and should not be used with other diversity equations.

Photographs of the study sites are also useful in the evaluation of habitat diversity, although they cannot be used to estimate a quantitative index. The level of discrimination is too low in most streamside photographs to make much of a distinction beyond pools and riffles. It is often impossible to judge the depth or velocity of the water from a still photograph, and the bed is often not visible.

The absence of availability information does not necessarily indicate low-quality criteria. Rather, it simply means that the criteria cannot be evaluated at this level. A verification study may be needed, in place of further screening-level evaluation.

The other two sources of bias to be assessed during a screening-level evaluation are problems with study design and sampling error. Limitations of study design may be more serious than those associated with observational or collection techniques, because the former are more difficult to identify than the latter. Field investigators nearly always record the type of gear used to collect the data, but may omit a description of how sampling locations were selected. In a sense, the presence or absence of a sampling design description is itself an indicator of the quality of the criteria. If the investigator has described the sampling design, it is at least an indication that its importance has not been overlooked. Whether the best sampling design or strategy was used is less important than knowing that the field crew did not confine their sampling to places where they expected to capture fish.

The sampling designs discussed in Chapter 2 are all intended to sample habitat areas within a study area in relative proportion to their occurrence. Although some, such as random or proportional designs, might be rated higher in the context of sampling theory, the use of any of the techniques should be rated about equally in criteria evaluations. The sampling design, however, does have important implications with respect to pooling data from several streams. The discussion of pooled data, in Section 4.3, outlines preferred sampling strategies when multiple study areas are used. The simplest technique for avoiding data pooling biases is to standardize the effort among all the study areas. If the investigator has done this, or if the sampling design is obviously unbiased, the user should not suspect disproportionate sampling bias. Conversely, one should be suspicious of random sampling designs, unless the investigator has demonstrated that the number of sample attempts in each stream was proportional to the size of the areas sampled. The same number of samples taken from two different-sized streams may tend to overemphasize the observations made in the smaller one. This tendency can be counteracted by measuring the habitat variables at every sampling location, whether fish were observed or not. This approach allows the precise determination of CPUE, as well as providing an excellent availability function.

Errors and biases associated with data collection can generally be categorized into three groups: precision error, disturbance error, and gear bias. Precision error refers to the ability of the observer to identify the focal points of observed or captured fish. Disturbance error refers to the observation or collection of fish at locations to which they were frightened or displaced due to the activities of the observer. Gear bias is introduced by variations in the efficiency of different observation or collection techniques imposed by the environments in which they are used. All three types of error are often interrelated.

Precision errors are often related to visibility and gear limitations. The reason that collection techniques such as seining and poisoning are not recommended for criteria studies is that they are imprecise. That is, the original focal point of the fish can seldom be identified with these methods. Precision error is also important when evaluating telemetric or photographic data. Neither technique is particularly disruptive or biased, but unless the investigator has used procedures to obtain a precise fix, the true location of the fish may be in error by several meters. Such positional errors are acceptable in streams exhibiting gradual microhabitat changes, but not where conditions change abruptly. Precision errors can also be introduced when the activity of the fish is unobservable or undifferentiated. The capture of transient fish moving through a section of stream is one such source of error, although usually a relatively minor one. More common is the capture of fish engaged in random swimming. These fish usually exhibit focal point orientation, but the focal point is poorly defined. Several habitat measurements representing the average conditions utilized by the fish are preferable to a single measurement taken where each was captured. Unless the fish were observable, however, it is generally not known whether they were engaged in stationary or random swimming patterns. This is a precision problem generally associated with low visibility. As is typical with precision errors, this is not a serious problem in gradually varying environments, but very serious in abruptly changing ones. One of the reasons that skin diving and SCUBA observations are so highly recommended is that they both allow precise identification of focal points and activity patterns.

Disturbance is the most serious source of error because fish are captured or observed in locations where they have fled, and not at their original focal points. Fortunately, the potential effects of disturbance on the quality of a criteria data base can easily be evaluated by determining the methods used to collect the data. Skin diving, surface observation, radiotelemetry, and underwater photography are generally considered to be low-disturbance observational techniques. SCUBA observation may also cause little disruption, provided the fish are not alarmed or attracted to the bubbles vented by a diver. Fish exhibit a variety of reactions to divers, depending on the direction of approach, color of the wet suit, and other factors. Since a diver can also observe the behavior of the fish, however, the observation procedure can either be changed to eliminate the reaction, or documented so possible errors can be detected.

Electrofishing, explosives, and other active capture techniques are moderately to highly disruptive methods, depending on how they are employed. One or more of these methods may be necessary, however, when the visibility is low. The amount of disturbance can be controlled through the use of prepositioned electrodes, preset charges, lift nets, or other sampling techniques designed to allow fish to recover from the disturbance of placing the sampling device. Even with these precautions, these methods can still be fairly disruptive compared to skin diving or radiotelemetry. The major redeeming feature of capture techniques is that they can be more effective in turbid water than in clear water. High-quality data can be obtained, but only if the technique is applied as unobtrusively as possible. Therefore, in evaluating criteria developed from capture data, it is important to consider factors such as sampling strategy, technique, and gear effectiveness.

All types of observation or capture methods have some inherent limitation, but if the limitation can be considered constant over the range of measured conditions, the technique may be presumed to be unbiased. Conversely, if a technique is more effective in shallow water than in deep water, in slow water than in fast water, or in open water as opposed to areas of heavy cover, then it has the potential to introduce bias. The best way to evaluate a data set for potential bias is to fully understand the limitations of each data collection technique, as they were outlined in Chapter 4.

The second aspect of the review process is an evaluation of transferability, to determine whether or not criteria developed in one geographic area should be applied in another. Transferability is a concept that applies to all categories of criteria, but is particularly critical in the evaluation of category II data. The reason is that environmental availability exerts the greatest amount of influence on utilization criteria.

There are only two simple rules governing screening-level evaluations. First, and perhaps most important, is that criteria may be transferred from higher diversity streams to those with lower diversity. They may or may not be transferable from streams of lower diversity to higher. Second, the more similar the source and study streams, the greater the probability that the criteria are transferable to the study stream. Stream size is important only if it affects the relative frequency distribution of available habitat conditions. If the maximum depth of the source stream is considerably less than that of the study stream, the criteria may need to be extended by the judgement of the user. Conversely, criteria developed in a larger stream lacking the shallower depths or slower velocities of a smaller study stream may not be directly transferable. Such differences are not important where the utilization curve is enveloped by the availability curves for both streams. If the total range of utilized velocities for a species is from 0.25 to 0.75 m/sec, it would not matter if the source stream had a maximum velocity of 1.0 m/sec and the study stream, 2.0 m/sec, because both are outside the usable range. Evaluating differences between source and study streams is not a question of whether the two streams are virtually identical. Rather, it is a question of whether the differences between the two streams are large enough to cause changes in the observed fish behavior. The answer to this question may lie in one of the verification studies discussed in Section 6.2.2.

One of the most useful review techniques for evaluating the transferability of criteria is convergence. As it applies to criteria evaluations, convergence means that several investigators, studying different streams, all develop similar habitat preference (or, less likely, utilization) functions for a species. This tends to confirm that the preference function represents the species throughout a particular geographic range. For instance, if similar preference functions are developed for smallmouth bass in Missouri, Arkansas, and Kentucky, it is likely that these criteria would also be valid in Tennessee. Caution is advised, however, in extending this type of logic too far. Convergent criteria from California and New York are not necessarily applicable in Michigan.

A good example of convergent criteria is given in Shirvell and Dungey (1983), who investigated microhabitat utilization by brown trout in six different streams in New Zealand. Although the streams varied considerably in both size and morphometry, and the researchers developed utilization rather than preference functions, Shirvell and Dungey found remarkable similarity in functions for a specific activity. This type of replication reinforces the concept that the information could be used in practically all the streams in New Zealand, because it represents the species, and not the streams in which they were studied.

Unfortunately, convergence requires the availability of multiple criteria sets developed by similar techniques. Microhabitat criteria have not yet been developed for many species, let alone replicated. Existing criteria are almost exclusively utilization functions and, therefore, fairly divergent. There may be some reluctance among researchers to replicate criteria that have already been developed elsewhere. The importance of such replicative studies cannot be overstated, to achieve the goal of regionalized habitat preference criteria. It may be necessary, however, to design studies similar to the one by Shirvell and Dungey (1983) to realize this end.

6.2.2 Verification Studies

A verification study is an empirical test of the accuracy and reproducibility of offsite criteria in a stream being analyzed with the IFIM. All verification studies require the collection of certain types of data in the subject stream, although the type and amount of data varies according to the test. Not unexpectedly, the confidence that can be placed in the results of such a study is directly related to the effort invested. None of the verification studies described in this section, however, is as data intensive as a complete criteria development study. When selecting one of these techniques, the investigator should also evaluate the degree of certainty needed for a particular IFIM application. Some noncontroversial applications need only an approximate confirmation, while others will require a more sophisticated test. Three techniques have been developed for criteria verification. In order of increasing complexity, they are: the abbreviated convergence approach, habitat suitability overlay, and Monte Carlo simulation.

The abbreviated convergence method is essentially a miniature criteria study. A small number of fish is observed or collected in the stream being analyzed with the IFIM, and relative frequency histograms developed for each data stratum and variable. Each histogram is then superimposed with the suitability curve for that stratum and variable. If the peaks and tails of the histogram agree with those of the curve, the test is considered positive. Minor disagreement may also be considered as confirmation of the criteria, but if the histogram is radically different from the curve, something is wrong. Unfortunately, at this level of analysis, it is impossible to determine which one is incorrect (or if both are incorrect, for that matter).

One way to improve this test is to develop a preference histogram rather than one for utilization. This would require the measurement of 25 to 50 randomly selected points in the test reach. These data would then be used to construct a relative frequency histogram for habitat availability. The preference histogram is developed following the same approach described in Section 5.1.2. It is more appropriate to superimpose the suitability curve and the preference histogram, because the utilization histogram is affected by the stream environment the same way it affects category II criteria. The small sample size (usually 25 to 50 observations per data stratum) used to build the utilization histogram can also contribute to disagreement, but the effect should be more noticeable at the tails of the distribution. There should be fairly good agreement between the optimum ranges of the preference histogram and the suitability curve.

When conducting an abbreviated convergence study, it is important to follow the same data stratification and sampling protocol as the original criteria developer. It is not necessary to use the same sampling techniques as the original investigator, but if different techniques are used, the potential biases of the sampling procedure should be recognized. It is difficult to evaluate the precision of a set of criteria if the two investigators have not conducted the same experiment. Any deviations from the original study design invite "apples and oranges" comparisons.

The second criteria testing procedure is called habitat suitability overlay (HSO). The underlying premise of this approach is that there should be good agreement between the computed suitabilities of stream cells and the observed distribution of fish, if the criteria are valid for a particular stream. To apply this method, a test reach of stream is mapped and measured for analysis with PHABSIM. The composite suitability index (CSI) for each cell is then computed by running the HABTAT program (a component of PHABSIM) at one or more test flows, using the criteria set under investigation. The cell-by-cell suitability information is obtained by invoking an option in the program. A planimetric map of the test reach is then prepared, with all the stream cells overlaid on the map. One such map is prepared for each test flow and each criteria set to be evaluated. Then, the investigator visits the stream site during a period of time that one of the test flows is present in the stream. The locations of fish, corresponding to each data stratum, are determined and marked on the appropriate map.

Finally, the suitability for each of the cells containing fish of a certain life stage or data stratum is determined. These cells are grouped by quartile (i.e., combined suitabilities from 0.0-0.25, 0.26-0.50, 0.51-0.75, 0.76-1.0). The total number of fish and the total stream area in each quartile is then determined. The number of fish for each quartile is divided by the total surface area for the quartile, and a bar graph, such as shown in Figure 41, is developed.

If the results of a habitat suitability overlay are similar to Figure 41a, then the criteria are likely to be valid for the stream reach in question. The most fish per unit area are associated with the cells having the highest computed suitability, and the fewest fish with the lowest suitability. Any different results (Figure 41b is but one of many possibilities) would indicate a fundamental inaccuracy in one or more of the curves used to compute the CSI. Although this test will not reveal which curves are incorrect, the nature of the problem can sometimes be diagnosed. Table 10 lists some of the possible outcomes from an overlay test, and the probable causes.

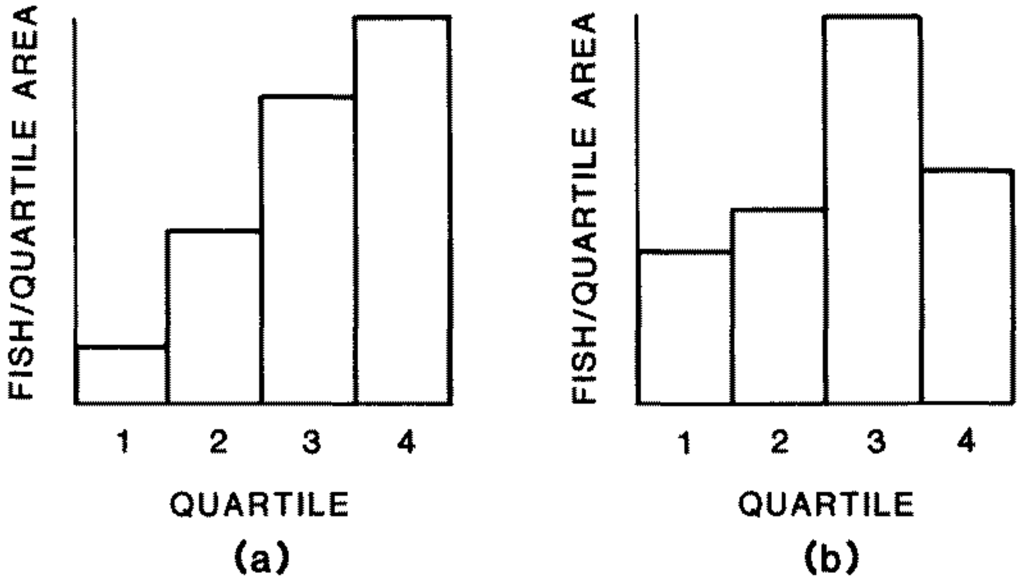


Figure 41. Bar graphs illustrating results of a habitat suitability overlay test for evaluating the accuracy of criteria.

Several precautions should be followed when implementing HSO evaluations. First, the internal homogeneity of each cell must be maintained in the habitat mapping phase. Although this is the goal of transect and vertical placements in standard PHABSIM analysis, it is much more critical in HSO studies. Internal homogeneity is important because the cell descriptors are taken at the verticals on each transect, but the fish are observed throughout the cell. If there is very much heterogeneity in the cells, the most typical problem will be the observation of fish in cells that, from their vertical measurements, would have a low suitability. The opposite can also occur (finding few fish in cells with high suitability), but is not as likely. Typical PHABSIM analyses are not particularly sensitive to cell heterogeneity, because WUA is computed from a large sample of cells. An HSO evaluation is much more sensitive to this factor because each cell is evaluated individually. A site mapped for an HSO study will probably require more transects than a standard PHABSIM site because of the necessity for homogeneity.

Table 10. Diagnosis of the probable causes for negative results from habitat suitability overlay tests of criteria accuracy.

Definitions

Quartile 1 - CSI = 0 - .25
Quartile 2 - CSI = .26 - .50
Quartile 3 - CSI = .51 - .75
Quartile 4 - CSI = .75 - 1.0

Error Type	Symptom	Diagnosis
A	Same fish/area in Quartiles 3 and 4	1. Curve(s) too narrow under optimum range.
B	All fish in Quartile 4	1. Curve(s) too broad under optimum range or 2. Population in test stream far below saturation.
C	More fish/area in Quartile 3 than 4	1. Peak(s) of curve(s) do not represent true optimum or 2. The wrong algorithm was used to compute CSI.
D	More fish/area in Quartile 2 than 3 or	1. Rising or descending limbs of curves wrong shape (concave and should be convex) or
D	More fish/area in Quartile 1 than 2	2. End points of curves too restrictive.
E	More fish/area in Quartile 1 than 4	1. Curve(s) have serious errors and should be refined or rejected.

The second precaution is to make sure that the test flow at which the fish were observed is the same as the discharge simulated in PHABSIM. The easiest way to ensure this is to measure the discharge during the fish observation phase and, subsequently, enter the test flow into PHABSIM. This is easier and less risky than simulating the flows first, and then waiting for the stream to cooperate.

It should be fairly obvious that the quality of the test is affected by the investigator's ability to translate fish positions in the stream to the correct cell on the map. This ability can be enhanced several ways. First, the map should be as accurate as possible. This means that the map should be drawn to scale and not sketched. Obvious landmarks, such as large boulders, tree trunks, submerged logs, or other features that can serve as reference points should be accurately displayed on the map. The second suggestion is to use cells that are as large as possible, while still maintaining internal homogeneity. It is easier to locate (place) a fish somewhere in a big cell than in one of several small cells. It may also be practical in some studies to grid the cells directly over the stream. Parachute cord, marking the ends and edges of each cell can be suspended over the stream to mark cell locations. This is the most accurate method of translating fish locations to the map, but also the most tedious. If this technique is used, fish locations should be determined and marked in the stream with a monument or marker buoy first, and the gridwork laid out afterward. Otherwise, the fish will be disturbed by the activity and the grid lines can be a nuisance to a sampling crew in the stream. Once the grid is in place, however, an investigator might choose to leave it up for other test flows. This would depend on the chances of the grid being destroyed by high flows, wind, canoeists and rafters, or other typical hazards encountered in stream investigations.

Finally, the investigator should attempt to observe <u>all</u> of the fish of interest within the reach. The reason is that the evaluation variable is based on a density term, so the number of fish accounted for is very important. This is especially true when there are few fish in the reach. The smaller the sample size, the greater the effect of missed fish on the results. It may be advisable to obtain a population estimate for the reach after the observation phase has been completed. The adequacy of the test results can then be judged on the basis of the proportion of the population that was observed. The results should be trustworthy if 75% or more of the population was observed. If less than 50% was observed, the results of the test may be unreliable, especially if type A, C, or D errors (Table 10) occur. This problem is also a concern with the next verification approach, Monte Carlo simulation, but may not be as serious as it is in HSO evaluations.

The Monte Carlo approach is similar to HSO in concept, but is a more rigorous test of the criteria. As with the overlay method, PHABSIM is used to predict a composite suitability index (CSI) for each cell at a test flow, and fish are observed and assigned to cells on a plan map. To this point, implementation of both techniques is identical. The principal difference is that a Monte Carlo simulation predicts where the fish will be found. These predictions are then tested statistically against the observed distribution.

A Monte Carlo simulation recognizes that there is an element of chance associated with the use or nonuse of an area of stream, regardless of its suitability. That is, areas of "perfect" habitat may be unused while areas of less-than-perfect habitat are occupied, even if the original criteria are completely adequate for the stream. The underlying concept of a Monte Carlo simulation is illustrated by the two roulette wheels shown in Figure 42. Stream cells predicted to have a combined suitability of 1.0 are represented by the upper wheel, on which two-thirds of the slots are labled "no" and one-third, "yes." This division of the wheel is based on the relative frequency distribution of the original criteria data base, which showed that the maximum probability of finding a fish in "perfect" habitat was only 33%. The second roulette wheel in Figure 42 represents cells with a combined suitability of 0.1. This wheel is similar to the first one, except that the chances of finding a fish are only 3%. Consequently, this wheel has three slots marked "yes" and 97 slots marked "no." When a cell with a combined suitability of 1.0 is encountered, the upper wheel is spun and a fish predicted to occupy the cell only if the roulette ball lands on a "yes." Likewise, the lower roulette wheel would be spun each time a cell with a combined suitability of 0.1 was encountered. To continue this analogy, there are 101 possible suitability indexes between, and including, 0.00 and 1.00. Therefore, to conduct this type of simulation, one would need a very large room with 101 roulette wheels, each with a face marked with yes and no slots corresponding to the probability of finding a fish in cells ranging in suitability from 0.00 to 1.00. For each cell, the combined suitability would be determined and the correct roulette wheel spun. Depending on the outcome of the spin, a fish would be predicted to occupy or not occupy the cell. Then, proceeding to the next cell, the appropriate roulette wheel would be spun, and another prediction made. This process continues, cell by cell, until the number of predicted fish locations equals the number of fish actually observed in the reach.

In actual applications, a computer generates the predicted utilization of each cell. There are two basic classes of programs capable of Monte Carlo simulations. First, are generalized programs available commercially or in the public domain. These are usually easy to use and well documented, but may not be totally compatible with this type of application. The second type of program is developed specifically with criteria evaluation in mind. These are generally compatible with this use but, because they are usually written for similar research, the programs may be too specific. This means that the original program may need to be modified or rewritten to fit a particular application. A listing of the program GETFISH is located in Appendix D as a general guide for persons wishing to conduct a Monte Carlo simulation. This program was written by Dr. Kenneth A. Voos during his tenure at the Instream Flow Group, and was provided with the understanding that the program was developed in a research mode and no attempt has been made to make it operational. Potential users are thus cautioned to have the program translated by a computer programmer before trying to use it.

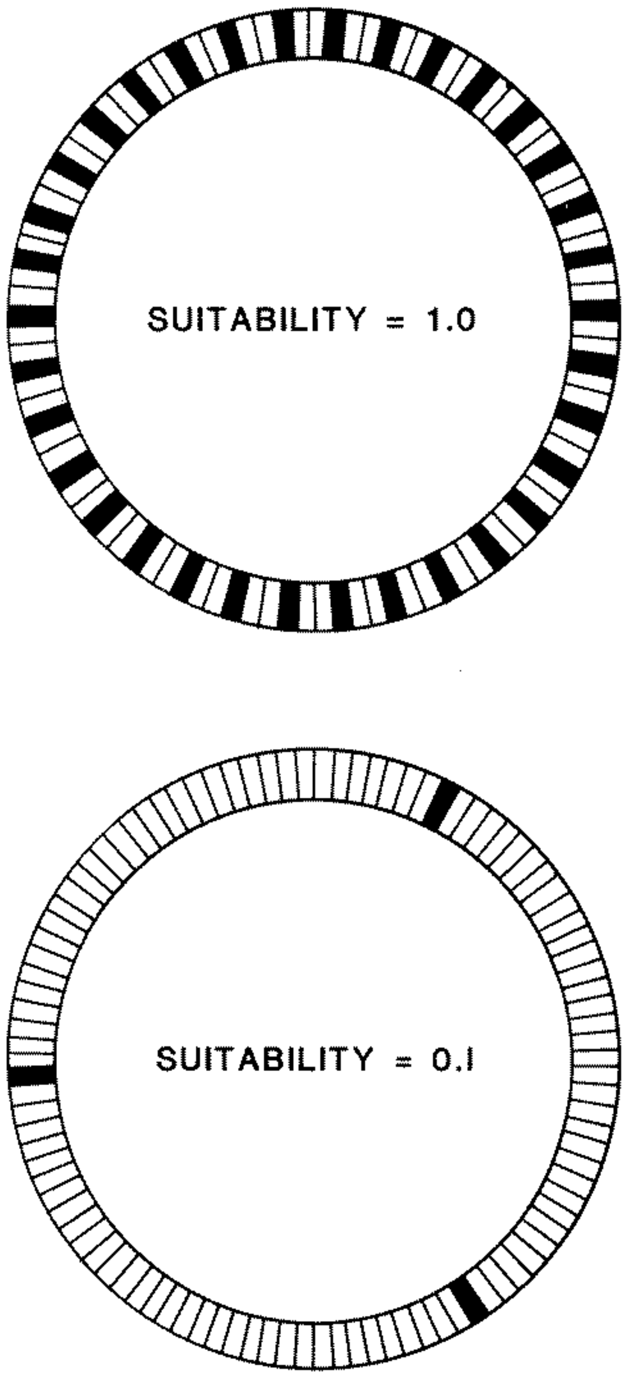
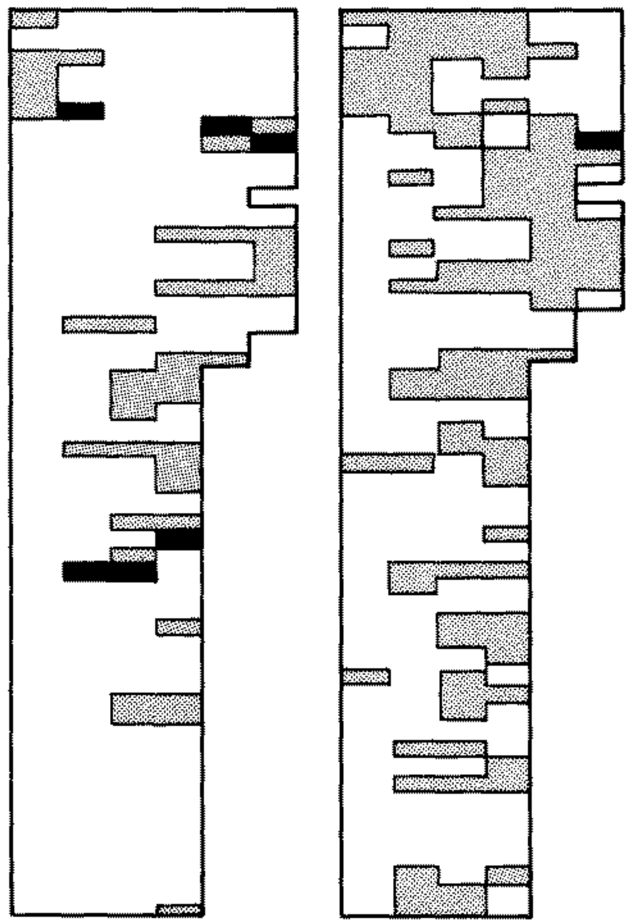


Figure 42. Roulette wheel analogy illustrating Monte Carlo prediction of cell utilization to test transferability of habitat suitability criteria.

The first application of a computerized Monte Carlo simulation in the evaluation of habitat preference criteria was conducted by Hardy et al. (1982), for a small desert stream in Nevada. The species under investigation in this study included several that exhibited rather narrow habitat preferences, Gila robusta, Rhinichthys osculus, and Gambusia affinis, and a generalist species, Poecilia mexicana. Figures 43 through 46 show the observed vs. predicted densities for each of these species. These figures illustrate that PHABSIM was able to identify the more restricted habitats of Gila, Rhinichthys, and Gambusia and the more ubiquitous distribution of Poecilia. The real strength of this technique, however, is that the results can be tested with correlation analysis. Hardy et al. (1982) noted that the correlations between observed and predicted densities of Gila and Rhinichthys were low (r=0.47 and r=0.39, respectively), although there was fairly good visual agreement between the type of cell actually utilized and that simulated. They suspected that this conflicting information might be due to the high association with cover exhibited by these species. In this case, cover might be considered to be an intervening variable influencing fish distributions. One strong point in favor of Monte Carlo simulation is that it can identify the presence of such intervening variables. Unless the habitat preferences of the animal are fairly well known, however, the actual identity of the variable may be a mystery. Failure to recognize the potential influence of some variable not included in the predictive model might result in the rejection of the criteria, even though they may be accurate. Low correlations between observed and predicted densities might be the result of using invalid criteria, but they can also be the result of leaving an important variable out of the simulation. This distinction should not be overlooked in an evaluation of transferability.

It was stated earlier that an investigator should attempt to observe all the fish in a reach when using HSO, but that this constraint is not as critical with Monte Carlo simulations. The reason is that the Monte Carlo approach simulates the same number of fish that were observed in the reach. The fish are more likely to be observed in high-quality habitat, and this method is likely to place the fish in the highest quality areas first. Therefore, it is possible to obtain an estimate of the quality of the criteria with relatively few observations. The quality of the test, however, is affected by sample size in much the same way that it influences the results of the overlay method. Very small samples will result in an analysis that concentrates on the highest quality habitat, because these areas will be filled first. The areas of marginal habitat are not evaluated very well unless a fairly large sample size is used. A small sample can also result in a poor correlation between predicted and observed fish locations, especially if there is an abundance of good habitat. The reason is that there would be many cells with a high CSI, but only a few cells in which fish were observed. This increases the likelihood that the computer would predict fish in cells where they were not observed.



Observed Simulated

G. robusta

r = 0.66

Figure 43. Simulated vs. observed distributions of Gila robusta using a Monte Carlo simluation. From Hardy et al. (1982). Reprinted with permission of the publisher.

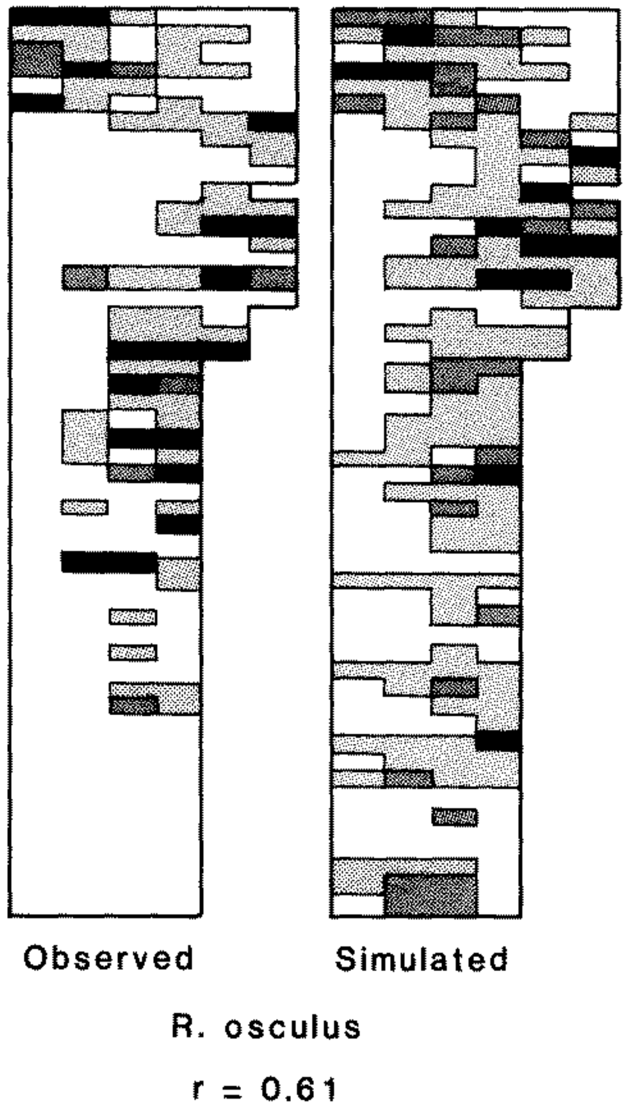


Figure 44. Simulated vs. observed distributions of Rhinichthys osculus using a Monte Carlo simulation. From Hardy et al. (1982). Reprinted with permission of the publisher.

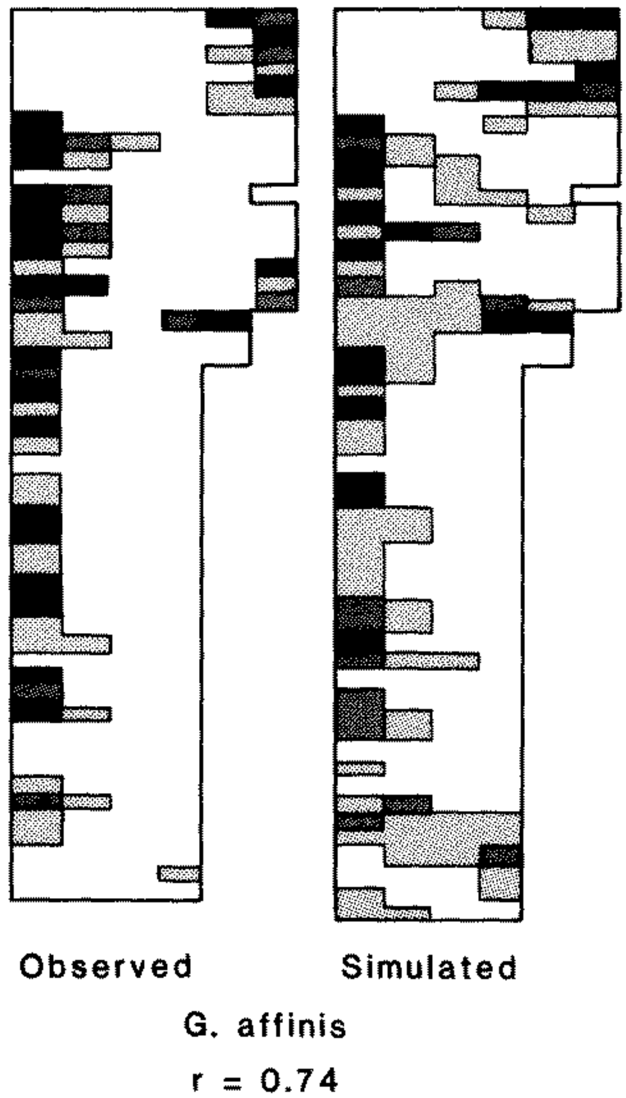


Figure 45. Simulated vs. observed distributions of Gambusia affinis using a Monte Carlo simulation. From Hardy et al. (1982). Reprinted with permission of the publisher.

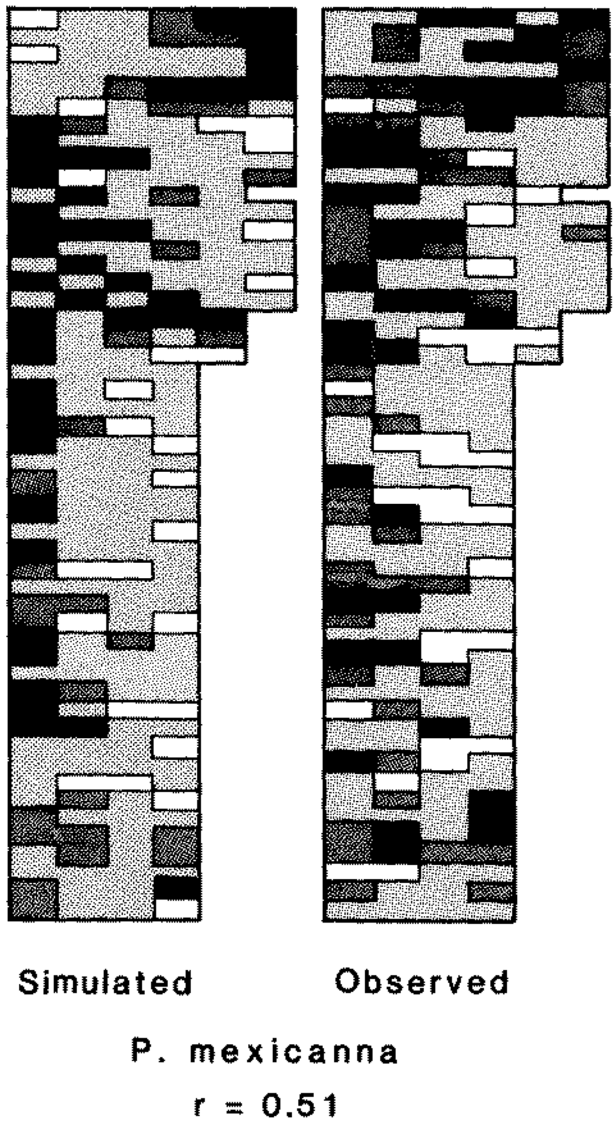


Figure 46. Simulated vs. observed distributions of *Poecilia mexicanna* using a Monte Carlo simulation. From Hardy et al. (1982). Reprinted with permission of the publisher.

6.3 EXTENSION AND MODIFICATION OF CRITERIA

Sometime in the future, convergent criteria sets for specified geographical regions may be available to IFIM users. This should be the goal of future criteria research, because convergence will eliminate much of the subjectivity involved with evaluating criteria. Until that time, IFIM users will need to judge the applicability of criteria for each stream under investigation. Sometimes, the user will be required to extend or modify the criteria to fit the application. Many scientists are reluctant to do either, because such actions imply "tinkering" with someone else's data. It is worth remembering that criteria are not data, but someone's interpretation of data. It is also important to recognize that the criteria are usually based on data, so indiscriminant modifications are equivalent to outright rejection of the criteria.

6.3.1 Criteria Extension

Extension is one of the most common forms of criteria modification. The reason that criteria must often be extended is because the range of a variable encompassed by a curve is limited by the conditions available in the source stream. This can result in a bell-shaped curve or truncated logistic function that represents the species' behavior for one side of the curve, but not the other. This restriction is most commonly associated with depth curves, but can occur with any microhabitat variable. Examples of restricted criteria and some typical extensions are shown in Figure 47.

The determination of suitability values beyond the end points of the data is mandatory when using PHABSIM. If univariate curves are used, suitability is computed by interpolation. This means that the range of the criteria must exceed the range of each variable simulated in the system. For example, if the maximum depth of the criteria curve is 2 meters, and a depth of 2.1 meters is encountered somewhere in the reach, the simulation will be aborted. It is equally important to specify upper (and, less often, lower) limits when using multivariate equations. Some functions, especially higher order polynomials, are mathematically unstable when extrapolated beyond the end points of the data. They may appear as bell-shaped functions, but at some value beyond the apparent upper tail of the function (where SI = 0), the suitability suddenly increases. This occurrence is unpredictable and may result in suitability indexes greater than 1.0 for some intervals of a variable. Therefore, it is important to place an upper threshold on these functions to prevent this from happening.

The easiest functions to extend are depth curves that take the form of a truncated logistic growth curve (e.g., Figure 47b). This is essentially a one-tailed curve that reaches a maximum value at the greatest depth encountered. It is usually a fairly safe assumption that greater depths will be equally suitable, at least for the range of depths normally encountered in rivers.

Bell-shaped curves are somewhat more challenging, because they imply reduced suitability as the value of a variable increases beyond the peak of the curve. This is logical for a velocity curve because some velocities will inevitably exceed the tolerances of an organism. It is not so easy to draw

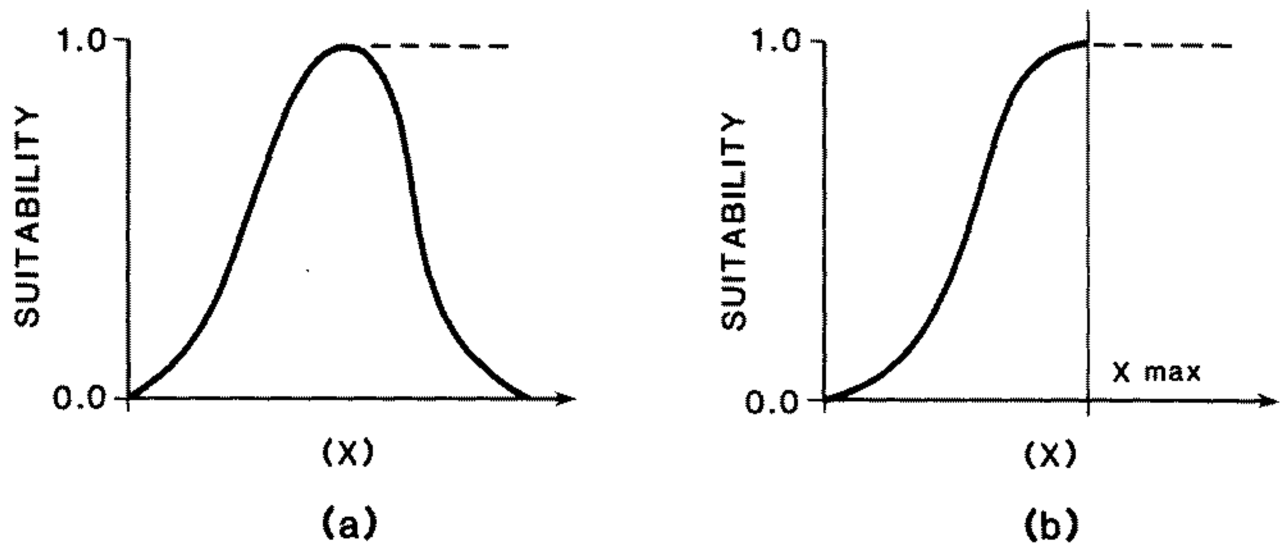


Figure 47. Examples of truncated criteria, and common methods of extension.

the same conclusion for depth and substrate curves. A bell-shaped substrate curve might be logical for a nest-builder or group of macroinvertebrates, but perhaps not for an adult fish. Similarly, a bell-shaped depth curve may represent the best condition for young fish that avoid predation by staying in shallow water. Whether the curve should retain the shape suggested by the data, or should be extended, requires an interpretation of the mode of usage by the organism.

6.3.2 Criteria Modification

Although extension is a form of criteria modification, it is really a matter of letting professional judgement take over where the criteria leave off. Actual modification of the criteria involves changing the shape or the intercepts of the original function. There may be many reasons for modifying a set of criteria; some are legitimate and some are not. The legitimate modifications include such things as:

(1) the addition of information not contained in the original data,

(2) the resolution of differences between two or more models,

(3) the injection of professional opinion into the criteria, and

(4) the development of "mixed" models.

The purpose for making any of these changes is to improve the accuracy of microhabitat predictions in PHABSIM, in essence, to improve the model for a particular situation. It is not legitimate to change the criteria to alter the results of PHABSIM. This constitutes deliberate manipulation of the model to justify a preconceived outcome. This is not a valid use of PHABSIM, and the credibility of both the user and the model can suffer from the practice.

The addition or consolidation of information is one of the easiest types of modifications that can be made to a set of criteria. This type of change usually affects the substrate criteria. A common alteration involves the addition of a term for embeddedness or percent fines to a criteria set that contains information on the dominant particle size only. This addition is illustrated in Figure 48, which shows the original criteria as a solid line and the added information as a dashed line.

The original information was retained in Figure 48, but suitabilities for various percentages of fine materials in the substrate matrix were added. In this case, the original suitability was used if the percent fines was zero (i.e., the decimal value of the code was zero), but a zero value was assigned each time the decimal in the code was 9, implying 90%-100% fines. It is important to stress that by modifying Figure 48, part of the curve is data based (category II or III) and part is category I. Such additions, and the judgemental portions of the resultant curve, must be documented by the user.

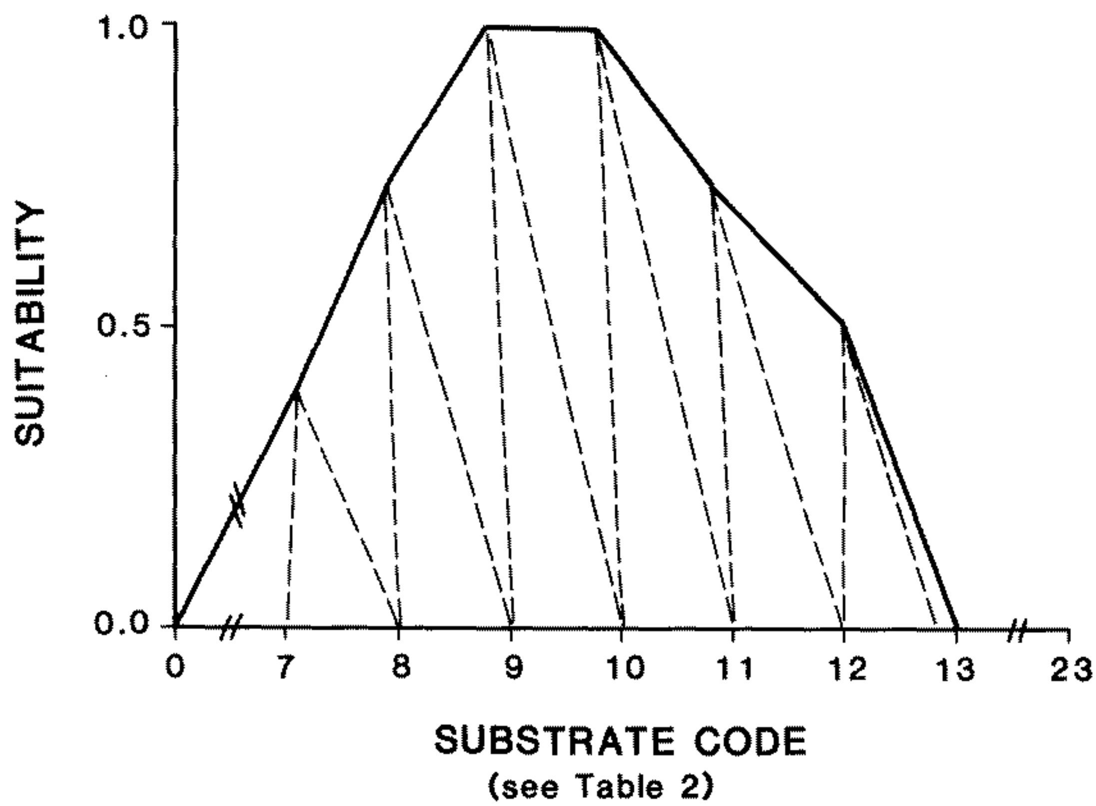


Figure 48. Addition of suitability indexes for percent fines to a spawning substrate curve developed for dominant particle size only.

The desirability of changing one or more curves in a criteria set becomes apparent as the result of either of two circumstances. The first is the availability of two or more divergent criteria sets, with no clear indication that one is any better than the others. The second situation occurs when the results of a verification analysis, especially HSO and Monte Carlo simulations, indicate low reliability of the available criteria. In either case, the solution may lie in a "mixed" model.

A mixed model follows the basic concept of Bayes theorem in statistical analysis. Bayes decision theory is used for decision-making when there is uncertainty about the true state of nature. Sometimes, the best way to reduce uncertainty is to obtain more information, but when evaluating or modifying criteria, this may not be possible, nor will more information necessarily gain anything. Bayes theorem enables the combination of new information, which may be in the form of a few samples or professional opinion, with old information to create a new model (Bostock and Davis 1975).

Bayesian decision-making is used fairly extensively in groundwater hydrology and petroleum geology for estimating the optimal design of well fields. It may not be directly applicable to criteria modification because empirical probabilities, rather than normalized indexes, are required. Conceivably, category II frequency distributions could be used with this approach, but not category I, III, or normalized category II criteria. The mixed model approach, however, adopts many of the same concepts as Bayes theorem, and allows the use of indexes instead of empirical probabilities.

The development of a mixed model depends on the availability of at least two sets of criteria. Both sets may be of any category or mixture of categories. The simplest explanation of a mixed model is illustrated in Figure 49, which shows two depth curves from different sources. In this case, there is no reason to favor one curve over the other, so the mixed model is simply the average of the two curves.

One of the principles of Bayes theorem is that an investigator can estimate the "correctness" of a particular model. The Bayesian approach combines sample information with other available prior information that may appear to be pertinent. The probabilities associated with this prior information are called subjective probabilities in that they measure a person's degree of belief in a proposition (Walpole and Meyers 1972). Essentially, the "*a posteriori* probability" is a function of the weight assigned to a model, with respect to other models.

The same concept can be applied to the combination of two or more sets of criteria. In the previous example, both curves were given equal weight in the average because there was no reason to believe that one was any better than the other. Suppose, however, that the investigator believes that curve 2 is more accurate than curve 1, based on past experience. Yet, the study supporting curve 1 was well done and the results cannot be rejected. The mixed model, in this case, would be a weighted curve. The weights assigned to the

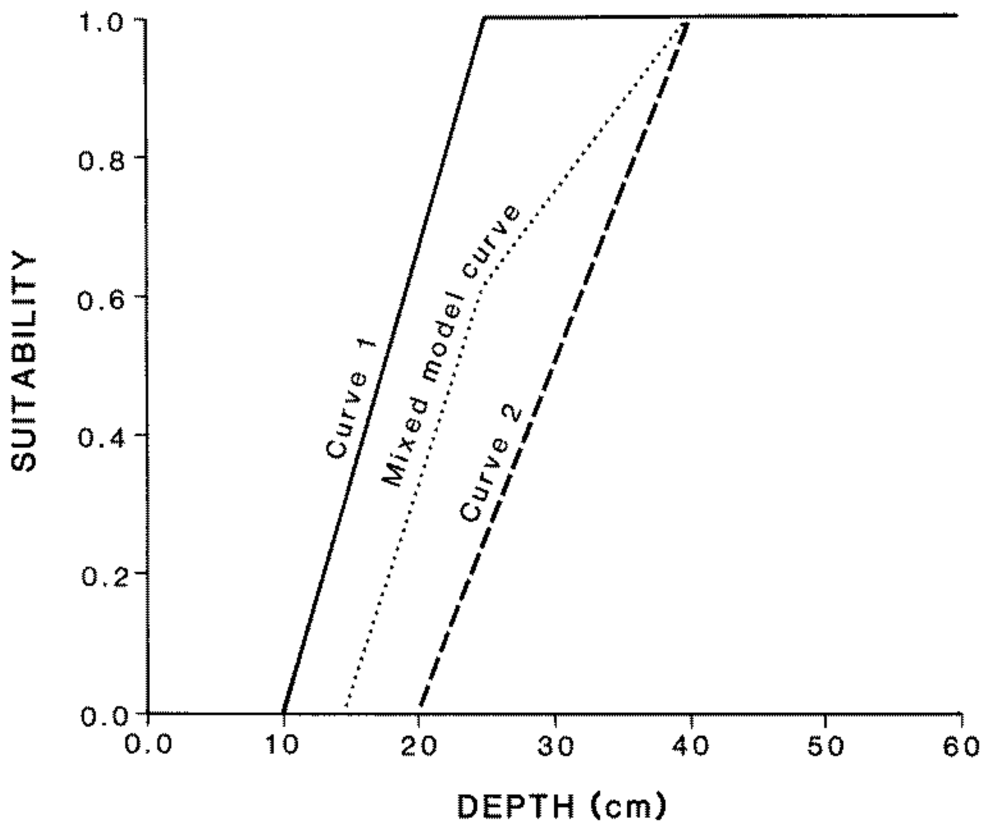


Figure 49. Mixed model criteria from two equally weighted depth curves.

two curves reflect the degree to which the investigator believes that curve 2 is more accurate than curve 1. Construction of the mixed curve must be done point-by-point, generally following the equation:

$$SI'_{(i)} = \frac{a(SI_{(i,1)}) + b\ (SI_{(i,2)}) + \ldots z\ (SI_{(i,n)})}{a + b + \ldots\ z} \tag{13}$$

where $SI'_{(i)}$ = the mixed model suitability index for interval (i) of a variable,

$SI_{(i,1)}$ = the suitability index for interval (i) of a variable from curve 1,

a = the weight assigned to curve 1,

$SI_{(i,2)}$ = the suitability index for interval (i) of a variable from curve 2,

b = the weight assigned to curve 2,

$SI_{(i,n)}$ = the suitability index for interval (i) of a variable from the nth curve, and

z = the weight assigned to the nth curve.

Figure 50 shows a mixed model for the two depth curves in Figure 49, where curve 2 was weighted 3:2 over curve 1.

Any time a mixed model is created, it is not immediately known if it is an improvement over the previous models. Unless the modifications are relatively minor, an empirical verification study to test the accuracy of the modified curves is strongly recommended. The Monte Carlo technique is well adapted to this analysis, because it gives a quantifiable indication of whether or not the mixed model improves the predictive capability of PHABSIM. Habitat suitability overlay can also be used to test old or new curves, but it lacks the ability to discriminate the better of two similar, but nonidentical, criteria sets.

A sensitivity analysis can also be conducted in conjunction with a verification test to improve the overall quality of the model. Sensitivity analysis is used to investigate the effects of different estimates of the input variables on the output variables in a mathematical model. One quick analysis to determine the importance of differences between two sets of criteria is to use both in PHABSIM to calculate WUA's over the flow range of interest. If both flow vs. WUA curves are the same, it makes little difference which set is used for subsequent analyses. As long as both curves have about the same intercepts and peak at the same flow(s), it is likely that the same management decisions would be made from a habitat time series (Bovee 1982). If decisions will be made by an optimization technique, however, or using an effective habitat time series, the magnitude of the curves is important. Differences in the shapes of the flow-WUA relationships are important regardless of the analysis procedure. It is in these instances that selecting the correct set of criteria is most critical.

Sensitivity analysis is also a useful technique for fine-tuning individual curves. After conducting an HSO or Monte Carlo verification, an investigator will have a good idea about the overall accuracy of the criteria set. Neither of these techniques, however, can identify which curves are causing problems when the results of a test indicate poor predictive capability. One solution is to enter modified curves into the PHABSIM analysis, one at a time, and then to repeat the verification analysis. (Since the fish observation data are already available, this takes relatively little additional time.) The test results should improve when the correct curve or curves are identified. Progress in this direction can be best gauged by examining the correlation coefficients produced through Monte Carlo simulations, so this technique is preferred.

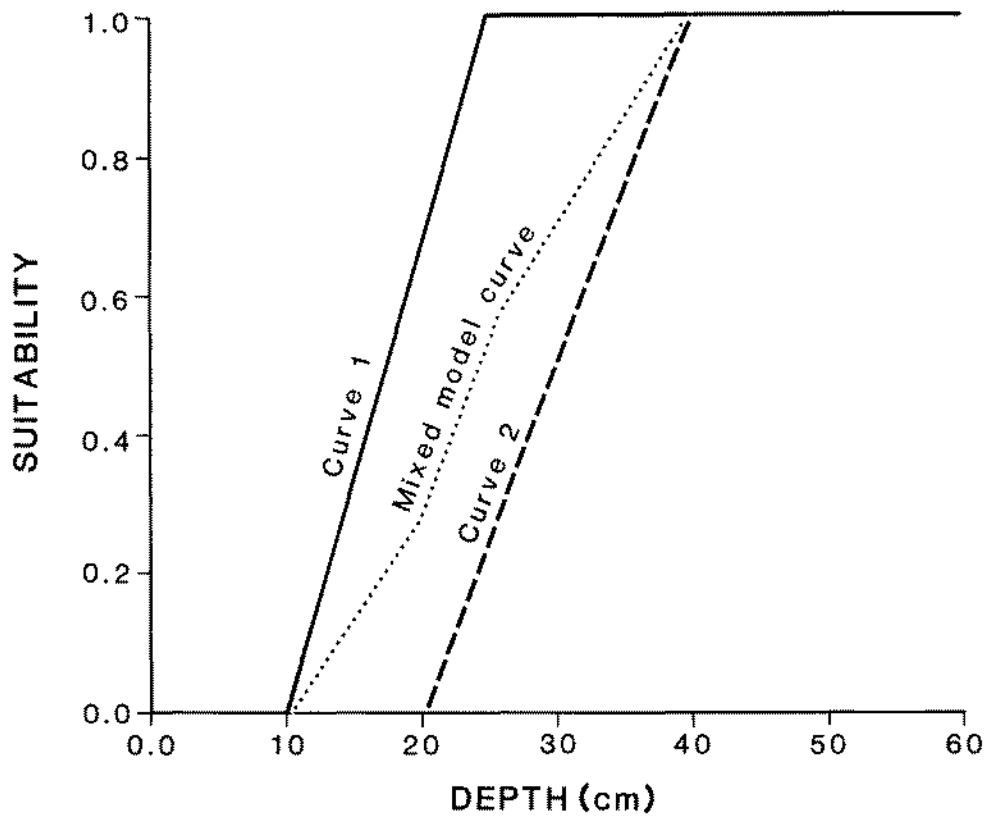


Figure 50. Mixed model for two depth curves, weighting curve 2 over curve 1 by a factor of 3:2.

6.4 DISCUSSION

The procedures outlined in this chapter should allow a systematic and flexible process for reviewing, evaluating, modifying, and testing criteria. Some situations will dictate very thorough evaluation and testing procedures, whereas a brief review will often suffice in other situations. It is important for a potential user of the IFIM to evaluate the criteria needs of an IFIM analysis during the study planning phase. It is equally important to judge the legal and institutional setting of the decision process, and anticipate the level of criteria evaluation and testing that will be required. As the intensity of conflict increases between water users and instream flow interests, the investigator will increasingly be required to justify the criteria used in a study. This means that the review and evaluation process must be more comprehensive and testing procedures more definitive.

The initial review of the criteria accomplishes several things, in addition to identifying data gaps and potential incompatibilities between the criteria and the needs of the study. One of the most practical by-products of the criteria review, is that the investigator will learn how data should be collected for the PHABSIM analysis. One of the worst mistakes a PHABSIM field crew can make is to collect data that are not compatible with the criteria. The easiest way to prevent this from happening is to review the data stratification and sampling protocol of the criteria study, and use the same procedures when collecting data for PHABSIM. The second important aspect of the review is that the investigator will gain an appreciation of the quality and relative transferability of the criteria. In some cases, this is all the evaluation needed. Such information is also valuable when modifying or combining criteria. If various criteria sets are weighted for a mixed model, it helps to know how much faith to place in each set.

The best way to evaluate a set of criteria is to verify it in the field. Where this is not necessary or feasible, an evaluation review should be conducted to determine the relative quality and transferability of the criteria. If it is determined at the study planning phase that a field verification is necessary, however, it should never become infeasible to conduct one later in the study implementation. Sufficient time and resources to conduct a verification study should be incorporated in the plan of study and reserved for that purpose. A field verification may add approximately one year to the completion of an IFIM analysis. This is because the field lay-out for the PHABSIM sites should not be done until the criteria are finalized. It can take as much as a year to complete a verification study because not all life stages and activities are present at the same time.

Of the three verification methods discussed in this chapter, the habitat suitability overlay technique may be the most generally useful. The abbreviated convergence method has the advantages that it does not require the establishment of a PHABSIM site, nor does it require much data. The small sample sizes associated with this method also contribute to its major disadvantage. That is, the quality of the utilization or preference histograms developed by this approach may be so poor that divergence from the criteria is meaningless. It is certainly uninterpretable as a means of judging the criteria.

The habitat suitability overlay method requires slightly more data than the abbreviated convergence approach, but the results from HSO are far more definitive. Positive results are positive and negative results, negative. Furthermore, this method allows some diagnosis of the nature of the problem when negative results are obtained (see Table 10). Although HSO results do not indicate which curve or curves are in error, the method can be used with a sensitivity analysis to test the influence of individual curves. One advantage of HSO over the Monte Carlo approach is that it does not require any specialized computer programs to conduct the analysis. The results from the overlay method are not as quantitative as those from a Monte Carlo simulation, but for many studies they are totally adequate.

The Monte Carlo approach is by far the most quantitative of the verification tests. This method has the advantage of producing a test statistic, the correlation coefficient between observed and predicted fish positions. This allows a rigorous test of individual curves or groups of curves, and is a powerful tool for "fine tuning" criteria when used in conjunction with a sensitivity analysis. This approach requires a specialized simulation program, but once the investigator has access to the software, the field application is virtually identical to the overlay method.

Both the HSO method and Monte Carlo simulation can be used to test more than just the accuracy of individual criteria functions. There are at least six different algorithms that can be used to compute WUA in PHABSIM. These techniques can also be used to determine the best algorithm to use. This can be an extremely valuable test when questions arise regarding the use of univariate curves vs. multivariate response surfaces. If the investigator can show that the former are as predictive as the latter, further justification of univariate curves is unnecessary. Conversely, if one algorithm is clearly superior to the others, it should be used in subsequent analyses.

HSO and Monte Carlo are both more data intensive than the abbreviated convergence method. A potential disadvantage of both methods is that they require the establishment of at least one PHABSIM site before the criteria are finalized. This is not a penalty, however, if the site can be used later in the PHABSIM analysis. The most serious drawbacks to both methods may be the requirements of observing a large proportion of the population at the study site and assigning individual fish to the correct cells. The observation of a large proportion of the local population is necessary when using either method, but may be more critical with HSO. One should be suspicious of tests where less than half of the local population was observed. The results of both tests can be affected by the ability of the investigator to observe the fish. When it is necessary to capture fish by electrofishing or other sampling techniques, the investigator must be very careful not to displace them during sampling.

Verification testing should not be confined to the evaluation of offsite criteria, alone. Several techniques for extending and modifying criteria were discussed in Section 6.3. Each time a curve is changed, a new criteria set has been created. Field verifications of these modified criteria sets are strongly encouraged, unless the modifications are trivial. Furthermore, the results of a field verification, when used with sensitivity analysis, may provide the basis for the modified curve. If the investigator is not careful, a feedback loop will be established where the verification test essentially becomes a criteria development study. This is an acceptable means of modifying criteria, but unless this feedback loop is broken, the modified criteria are not really tested. The best way to break this pattern of self-reinforcement is to repeat the test at another flow or at another site, where the habitat distribution is significantly different from the original test conditions. Upon completion of such a replicate test, it may be determined that the modifications essentially made the criteria site-specific, and that the unmodified criteria are superior overall.

Perhaps it will some day be unnecessary to review, evaluate, modify, and test criteria to the extent described in this chapter. That time will come when criteria can be regionalized, based on mathematical convergence. Many IFIM users have suggested the development of criteria that are applicable to specified geographical areas. This is an admirable goal, but will require a long term commitment of the research community to achieve. One problem with the regionalization concept is that the regions are specified first, and then the criteria developed. Ideally, the process should be reversed. Criteria should be developed and replicated several times, and the regions defined by families of convergent criteria. This aspect of biology is so new, that it will be a long time before the goal is reached (and given the evolutionary nature of the subject, it could be a very long time). Until then, researchers must do the best they can to provide comprehensive, accurate, transferable criteria, and users must continue to evaluate and test it.

REFERENCES

Allen, D. M., and J. H. Hudson. 1970. A sled-mounted suction sampler for benthic organisms. U.S. Fish. Wildl. Serv. Spec. Sci. Rept., Fisheries 614, Washington, DC.

Amlaner, C. J., Jr. 1980. The design of antennas for use in radio telemetry. Pages 251-262 in C. J. Amlaner and D. W. MacDonald, eds. A handbook on biotelemetry and radio tracking. Pergamon Press, New York.

Amlaner, C. J., and D. W. MacDonald, eds. 1980. A handbook on biotelemetry and radio tracking. Pergamon Press, New York. 804 pp.

Armour, C. L., K. P. Burnham, and W. S. Platts. 1983. Field methods and statistical analyses for monitoring small salmonid streams. U.S. Fish. Wildl. Serv. FWS/OBS-83/33. 200 pp.

Bachman, R. A. 1983. Foraging behavior of free-ranging brown trout (Salmo trutta) in a stream. Ph.D. Dissertation. Pennsylvania State University, University Park. 158 pp.

Bagenal, T. B., and W. Nellen. 1980. Sampling eggs, larvae and juvenile fish. Pages 13-36 in T. Backiel and R. L. Welcomme, eds. Guidelines for sampling fish in inland waters. Food and Agriculture Organization of the United Nations, European Inland Fisheries Advisory Commission. Tech. Paper 33. Rome, Italy.

Bailey, R. M., J. E. Fitch, E. S. Herald, E. A. Lachner, C. C. Lindsey, C. R. Robins, and W. B. Scott. 1970. A list of common and scientific names of fishes from the United States and Canada. Am. Fish. Soc. Spec. Publ. 6. 150 pp.

Bain, M. B. 1985. Fish community structure in rivers with natural and modified daily flow regimes. Ph.D. Dissertation. University of Massachusetts, Amherst. 152 pp.

Bain, M. B., J. T. Finn, L. J. Gerardi, Jr., M. R. Ross, and W. P. Saunders, Jr. 1982. An evaluation of methodologies for assessing the effects of flow fluctuations on stream fish. U.S. Fish. Wildl. Serv. FWS/OBS-82/63. 199 pp.

Bain, M. B., J. T. Finn, and H. E. Booke. 1985a. Quantifying stream substrate for habitat analysis studies. N. Am. J. Fish. Man. 5(3B): 499-500.

Bain, M. B., J. T. Finn, and H. E. Booke. 1985b. A quantitative method for sampling riverine microhabitats by electrofishing. N. Am. J. Fish. Man. 5(3B):489-493.

Baldridge, J. E. 1981. Development of habitat suitability criteria: Appendix 3. Pages 357-419 in W. J. Wilson, E. W. Trihey, J. E. Baldridge, C. D. Evans, J. G. Thiele, and D. E. Trudgen, eds. An assessment of environmental effects of construction and operation of the proposed Terror Lake hydroelectric facility, Kodiac, Alaska. Arctic Environmental Information and Data Center. University of Alaska, Anchorage.

Balon, E. K. 1975. Reproductive guilds of fishes: a proposal and definition. J. Fish . Res. Bd. Can. 32(6):821-864.

Bidgood, B. F. 1980. Fish surgical procedure for implantation of radio tags in fish. Fish. Res. Rept. 20, Alberta Div. Fish and Game, Edmonton, Alberta, Canada.

Bostock, C. A., and D. R. Davis. 1975. Application of Bayesian decision theory in well field design. Pages 113-124 in D. L. Chery, Jr., ed. Hydrology and water resources in Arizona and the Southwest. Proc. Ariz. Acad. Sci.5.

Bovee, K. D., and T. Cochnauer. 1977. Development and evaluation of weighted criteria, probability-of-use curves for instream flow assessments: fisheries. Instream Flow Information Paper 3. U.S. Fish. Wildl. Serv. FWS/OBS-77/63. 38 pp.

Bovee, K. D., and R. T. Milhous. 1978. Hydraulic simulation in instream flow studies: theory and technique. Instream Flow Information Paper 5. U.S. Fish. Wildl. Serv. FWS/OBS-78/33. 130 pp.

Bovee, K. D. 1982. A guide to stream habitat analysis using the instream flow incremental methodology. Instream Flow Information Paper 12. U.S. Fish. Wildl. Serv. FWS/OBS-82/26. 248 pp.

Brusven, M. A. 1977. Effects of sediments on insects. Page 43 in D. L. Kibbee, ed. Transport of granitic sediments in streams and its effects on insects and fish. U.S. Forest Service. Forest, Wildl., and Range Expt. Sta. Bull. 17. University of Idaho, Moscow.

Burch, O. 1983. New device for sampling larval fish in shallow water. Prog. Fish-Cult. 45(1):33-35.

Burrows, R. E., and B. D. Combs. 1968. Controlled environment for salmon propagation. Prog. Fish-Cult. 30:123-136.

Caceci, M. S., and W. P. Cacheris. 1984. Fitting curves to data. Byte 5:340-362.

Cada, G. F., and J. M. Loar. 1982. Relative effectiveness of two ichthyoplankton sampling techniques. Can. J. Fish. Aquat. Sci. 39:811-814.

Chamberlain, A. 1979. Effects of tagging on equilibrium and feeding. Underwater Telemetry Newsletter 9(1):1-3.

Coble, D. W. 1961. Influence of water exchange and dissolved oxygen in redds on survival of steelhead trout embryos. Trans. Am. Fish. Soc. 90(4):469-474.

Cordone, A. J., and D. W. Kelley. 1961. The influence of inorganic sediment on the aquatic life of streams. California Fish Game 47:189-228.

Crance, J. H. 1985. Delphi technique procedures used to develop habitat suitability index models and instream flow suitability curves for inland stocks of striped bass. U.S. Fish. Wildl. Serv. WELUT-85/W07. 58 pp.

Cross, D. G., and B. Stott. 1975. The effect of electrical fishing on the subsequent capture of fish. J. Fish. Biol. 7:349-357.

Cummins, K. W. 1966. A review of stream ecology with special emphasis on organism-substrate relationships. Pages 2-51 in Organism-substrate relationships. Spec. Pub. No. 4, Pymatuning Lab. Ecol., University of Pittsburgh, Pittsburgh, PA.

Cummins, K. W. 1979. The natural stream ecosystem. Pages 7-24 in J. V. Ward and J. A. Stanford, eds. The ecology of regulated streams. Plenum Press, New York. 398 pp.

Dettman, D. H. 1977. Habitat selection, daytime behavior, and factors influencing distribution and abundance of rainbow trout (Salmo gairdneri) and Sacramento squawfish (Ptychocheilus grandis) in Deer Creek, California. M.S. Thesis. University of California, Davis. 47 pp.

Dovel, W. L. 1964. An approach to sampling estuarine macroplankton. Chesapeake Sci. 55:77-90.

Egglishaw, H. J. 1964. The distributional relationship between the bottom fauna and plant detritus in streams. J. Anim. Ecol. 33:463-476.

Gerardi, L. J., Jr. 1983. Radio telemetry investigations of short-term habitat use by smallmouth bass in the Deerfield River, Massachusetts. M.S. Thesis. University of Massachusetts, Amherst. 142 pp.

Gerking, S. D. 1953. Evidence for the concepts of home range and territory in stream fishes. Ecology 34:347-365.

Gorman, O. T., and J. R. Karr. 1978. Habitat structure and stream fish communities. Ecology 59(3):507-515.

Gosse, J. C. 1982. Microhabitat of rainbow and cutthroat trout in the Green River below Flaming Gorge Dam. Final report, contract 81 5049. Utah Division of Wildlife Resources, Salt Lake City. 114 pp.

Guy, H. P. 1969. Laboratory theory and methods for sediment analysis. USGS Techniques for Water-Resources Investigations, Chapter C1, Book 5. 58 pp.

Hardy, T. B., C. G. Prewitt, and K. A. Voos. 1982. Application of a physical habitat usability model to the fish community in a spring-fed desert stream. Pages 391-397 in W. K. Lauenroth, G. V. Skogerboe, and M. Flug, eds. Analysis of ecological systems: state of the art in ecological modelling. Elsevier Scientific Pub. Co., New York.

Hart, L. G., and R. C. Summerfelt. 1975. Surgical procedures for implanting ultrasonic transmitters into flathead catfish (Pylodictus olivaris). Trans. Am. Fish. Soc. 104:56-59.

Hayes, M. L. 1983. Active fish capture techniques. Pages 123-145 in L. A. Nielsen and D. L. Johnson, eds. Fisheries techniques. Am. Fish. Soc., Bethesda, MD.

Helfman, G. S. 1983. Underwater methods. Pages 349-369 in L. A. Nielsen and D. L. Johnson, eds. Fisheries techniques. Am. Fish. Soc., Bethesda, MD.

Hill, M. 1984. BMDP user's digest. BMDP Statistical Software, Inc., Los Angeles, CA. pp. 111-116.

Hunter, J. R., D. C. Aasted, and C. T. Mitchell. 1966. Design and use of a miniature purse seine. Prog. Fish-Cult. 28(3):175-179.

Ivlev, V. S. 1961. Experimental ecology of the feeding of fishes. Yale University Press, New Haven, CT. 302 pp.

Johnson, D. L., and L. A. Nielsen. 1983. Sampling considerations. Pages 1-21 in L. A. Nielsen and D. L. Johnson, eds. Fisheries techniques. Am. Fish. Soc., Bethesda, MD.

Johnson, F. H. 1961. Walleye egg survival during incubation on several types of bottom in Lake Winnibigoshish, Minnesota, and connecting waters. Trans. Am. Fish. Soc. 90(3):312-322.

Kryzhanovsky, S. G. 1949. Eco-morphological principles of development among carps, loaches, and catfishes. Tr. Inst. Morph. Zhiv. Severtsova. 1:5-332. [Translated from Russian by Fish. Res. Bd. Can. Transl. Ser. No. 2945, 1974.]

Linstone, H. A., and M. Turoff. 1975. The Delphi method. Addison-Wesley, Reading, MA. 620 pp.

MacDonald, D. W., and C. J. Amlaner, Jr. 1980. A practical guide to radio tracking. Pages 143-160 in C. J. Amlaner and D. W. MacDonald, eds. A handbook on biotelemetry and radio tracking. Pergamon Press, New York.

Manz, J. V. 1961. A pumping device used to collect walleye eggs from offshore spawning areas in western Lake Erie. Trans. Am. Fish. Soc. 93:201-206.

Marquart, D. W. 1963. An algorithm for least-squares estimation of nonlinear parameters. J. Soc. Indust. Appl. Math. 11(2):431-441.

Martin, R. G., and R. S. Campbell. 1953. The small fishes of Black River and Clearwater Lake, Missouri. Univ. Missouri Studies 26(2):45-65.

Mathur, D., W. H. Bason, E. J. Purdy, Jr., and C. A. Silver. 1985. A critique of the Instream Flow Incremental Methodology. Can. J. Fish. Aquat. Sci. 42:825-831.

Matz, A. W. 1978. Maximum likelihood parameter estimation for the quartic exponential distribution. Technometrics 20(4):475-484.

McCrimmon, H. R., and A. H. Berst. 1966. A water recirculation unit for use in fishery laboratories. Prog. Fish-Cult. 28:165-170.

Milhous, R. T., D. L. Wegner, and T. Waddle. 1984. User's guide to the Physical Habitat Simulation System (PHABSIM). Instream Flow Information Paper 11. U.S. Fish. Wildl. Serv. FWS/OBS-81/43 (revised). 475 pp.

Minckley, W. L. 1963. The ecology of Doe Run, Meade County, Kentucky. Wildl. Monog. 11:124 pp.

Minor, J. D. 1981. Aquatic applications of radiotelemetry and corresponding methods of data analysis in studying the behavior of fish in their natural environment. Pages 183-196 in E.F. Long, ed. Third interim conference on wildlife biotelemetry. University of Wyoming, Laramie.

Ogden, J. C. 1977. A scroll apparatus for the recording of notes and observations underwater. Marine Tech. Soc. J. 11:13-14.

Otis, K. J., and J. J. Weber. 1982. Movement of carp in the Lake Winnebago system as determined by radiotelemetry. Tech. Bull. 134. Wisc. Dept. Nat. Res., Madison, WI. 16 pp.

Orth, D. J., and O. E. Maughan. 1982. Evaluation of the incremental methodology for recommending instream flows for fishes. Trans. Am. Fish. Soc. 111(4): 413-445.

Parton, W. J., and G. S. Innis. 1972. Some graphs and their functional forms. Tech. Rept. 153. U.S. International Biological Program. Nat. Res. Ecol. Lab., Colorado State University, Ft. Collins. 41 pp.

Percival, E., and H. Whitehead. 1929. A quantitative study of the fauna of some types of stream-bed. J. Ecol. 17:282-314.

Platts, W. S., W. F. Megahan, and G. W. Minshall. 1983. Methods for evaluating stream, riparian, and biotic conditions. USDA Forest Service, General Tech. Rept. INT-138. Intermountain Forest and Range Expt. Sta., Ogden, UT. 70 pp.

Platts, W. S., and V. E. Penton. 1980. A new freezing technique for sampling salmonid redds. USDA Forest Service Res. Paper INT-248. Intermountain Forests and Range Expt. Sta. Ogden, UT. 22 pp.

Prewitt, C. G. 1982. The effect of depth-velocity correlations on aquatic physical habitat usability estimates. Ph.D. Dissertation. Colorado State University, Ft. Collins. 83 pp.

Remington, R. D., and M. A. Schork. 1970. Statistics with applications to the biological and health sciences. Prentice-Hall, Englewood Cliffs, NJ. 418 pp.

Reynolds, J. B. 1983. Electrofishing. Pages 147-163 in L. A. Nielsen and D. L. Johnson, eds. Fisheries techniques. Am. Fish. Soc., Bethesda, MD.

Root, R. B. 1967. The niche exploitation pattern of the blue-gray gnatcatcher. Ecol. Monogr. 37:317-350.

Ross, M. J., and J. H. McCormick. 1981. Effects of external radio transmitters on fish. Prog. Fish-Cult. 43(2):67-73.

Scheele, D. S. 1975. Reality construction as a product of Delphi interaction. Pages 37-71 in H. A. Linstone and M. Turoff, eds. The Delphi method: techniques and applications. Addison-Wesley, Reading, MA.

Scherer, E., K. R. Scott, and S. H. Nowak. 1977. A modular large scale laboratory system to acclimate and test aquatic organisms. Can. Fish. Mar. Serv. Tech. Rep. 728.

Scott, K. R., and L. Allard. 1985. A four-tank recirculation system with a hydrocyclone prefilter and a single water reconditioning unit. Prog. Fish-Cult. 46(4):254-261.

Shirvell, C. S., and R. G. Dungey. 1983. Microhabitats chosen by brown trout for feeding and spawning in rivers. Trans. Am. Fish. Soc. 112(3):355-367.

Simpson, D. 1978. Evaluation of electrofishing efficiency for largemouth bass and bluegill populations. M.S. Thesis. University of Missouri, Columbia.

Snyder, D. E. 1983. Fish eggs and larvae. Pages 165-197 in L. A. Nielsen and D. L. Johnson, eds. Fisheries techniques. Am. Fish. Soc., Bethesda, MD.

Solomon, D. J. 1982. Tracking fish with radio tags. Symp. Zool. Soc. Lond. 49:95-105.

Somerville, P. N. 1958. Tables for obtaining non-parametric tolerance limits. An. Math. Stat. 29:599-601.

Sprules, W. M. 1947. An ecological investigation of stream insects in Algonquin Park, Ontario. Publ. Ontario Fish. Res. Lab, Univ. Toronto, 69:1-81.

Stalnaker, C. B. 1979. The use of habitat preferenda for establishing flow regimes necessary for maintenance of fish habitat. Pages 321-337 in J. V. Ward and J. A. Stanford, eds. The ecology of regulated streams. Plenum Press, New York.

Stalnaker, C. B. 1982. Instream flow assessments come of age in the decade of the 1970's. Pages 119-142 in W. T. Mason, ed. Research of fish and wildlife habitat. EPA-600/8-82-022. U.S. Environmental Protection Agency, Washington, DC.

Stasko, A. B., and D. G. Pincock. 1977. Review of underwater biotelemetry with emphasis on ultrasonic techniques. J. Fish. Res. Bd. Can. 34:1261-1285.

Stickney, R. R. 1983. Care and handling of live fish. Pages 85-94 in L. A. Nielsen and D. L. Johnson, eds. Fisheries techniques. Am. Fish Soc., Bethesda, MD.

Terhune, L. D. B. 1958. The Mark IV groundwater standpipe for measuring seepage through salmon spawning gravel. J. Fish. Res. Bd. Can. 15(5):1027-1063.

Thielke, J. 1985. A logistic regression approach for developing suitability-of-use functions for fish habitat. Pages 32-38 in F. W. Olson, R. G. White, and R. H. Hamre, eds. Proc. Symp. on small hydropower and fisheries. Am. Fish. Soc., Bethesda, MD.

Thompson, D. H. 1925. Some observations on the oxygen requirements of fishes in the Illinois River. Ill. Nat. Hist. Surv. 25(7):423-437.

Thompson, D. H., and F. D. Hunt. 1930. The fishes of Champaign County. Ill. Nat. Hist. Sur. Bull. 19(1):101 pp.

Trihey, E. W., and D. L. Wegner. 1981. Field data collection procedures for use with the physical habitat simulation system of the Instream Flow Group. U.S. Fish. Wildl. Serv., Cooperative Instream Flow Service Group, Ft. Collins, CO. 151 pp.

Tyus, H. M. 1982. Fish radiotelemetry: theory and application for high conductivity rivers. U.S. Fish. Wildl. Serv. FWS/OBS-82/38. 26 pp.

Tyus, H. M., B. D. Burdick, and C. W. McAda. 1984. Use of radiotelemetry for obtaining habitat preference data on Colorado squawfish. N. Am. J. Fish. Man. 4:177-180.

Voos, K. A. 1981. Simulated use of the exponential polynomial/maximum likelihood technique in developing suitability of use functions for fish habitat. Ph.D. Dissertation. Utah State University, Logan. 85 pp.

Walpole, R. E., and R. H. Meyers. 1972. Probability and statistics for engineers and scientists. Macmillan, New York. 506 pp.

Waters, B. F. 1976. A methodology for evaluating the effects of different stream flows on salmonid habitat. Pages 254-266 in J. F. Orsborn and C. H. Allman, eds. Instream flow needs. Am. Fish. Soc., Bethesda, MD.

Wene, G., and E. L. Wickliff. 1940. Modification of a stream bottom and its effect on the insect fauna. Can. Ent. 72:131-135.

Wickham, G. M. 1967. Physical microhabitat of trout. M.S. Thesis. Colorado State University, Ft. Collins. 42 pp.

Wilks, S. S. 1941. Determination of sample sizes for setting tolerance limits. Ann. Math. Stat. 12:91-96.

Winter, J. D. 1983. Underwater biotelemetry. Pages 371-395 in L. A. Nielsen and D. L. Johnson, eds. Fisheries techniques. Am. Fish. Soc., Bethesda, MD.

Zar, J. H. 1974. Biostatistical analysis. Prentice-Hall, Englewood Cliffs, NJ. 620 pp.

Zuboy, J. R. 1981. A new tool for fisheries managers: the Delph technique. N. Am. J. Fish. Man. 1:55-59.

APPENDIX A

First Round Sample Packet For a Delphi Inquiry

by

Johnie H. Crance
Instream Flow and Aquatic Systems Group
U.S. Fish and Wildlife Service
Fort Collins, Colorado

Dear

Thank you for agreeing to serve as a panelist for the sockeye salmon Delphi exercise.

The purpose of the exercise is to develop Suitability Index (SI) curves for use with the Instream Flow Incremental Methodology (IFIM) in the assessment of riverine habitat of sockeye. The Delphi technique is being used to develop SI curves for sockeye because data available in the literature on habitat suitability for the species is inadequate. Published and unpublished reports on sockeye will be used in developing the curves but opinions of the 12 Delphi panelists, including yourself, will be the primary basis for the resultant curves.

General information about the Delphi technique and SI curve development, and instructions and materials for completing the first round of the exercise are enclosed. A few hours of your time will be required to complete the first and subsequent rounds of the Delphi. You, no doubt, have many demands on your time but please respond to each round promptly. We should complete the exercise in about 6 to 8 months, assuming that four or five rounds will be required and that all panelists respond to each round within 10 days after receipt of material. You may wish to get an associate to serve as panelist in your behalf when you are unable to respond within 10 days.

I will serve as evaluator of the exercise. This means that I will prepare the material for each round, summarize responses, and prepare a final report, including rationale for the curves developed. Anonymity among panelists will be maintained until the exercise is completed. A representative of Region 7, U.S. Fish and Wildlife Service, Anchorage, and a representative of the Alaska Department of Game and Fish, Anchorage, will serve as monitors of the Delphi exercise. The monitors will receive all material sent to panelists, including a summary of each round, and will assist in clarifying any issue that may arise during the exercise.

Thank you again for consenting to be a panelist. I look forward to receipt of your input.

Sincerely,

Johnie H. Crance
Fishery Biologist

Enclosures

SOCKEYE SALMON

Definition of Terms

The sockeye salmon Delphi exercise will be concerned with the riverine (lotic) habitat used by the various life stages of the species. A definition of some terms to be used during the Delphi exercise has been assumed. If you disagree with any general definitions listed below, please give your definition of the term and/or any other terms that you feel need clarification.

Inmigrant. Mature adult during migration upstream from the estuary to the riverine spawning site.

Spawning. Nest building, egg and sperm release, and fertilization in lotic habitat.

Incubation. Egg, from fertilization to hatching.

Preemergent larval. Larval stage, from hatching to emergence from redd.

Postemergent larval. Larval or (juvenile?) stage during migration from riverine hatching site to lacustrine nursery habitat.

Outmigrant. Juvenile during migration downstream to estuary.

INSTRUCTIONS

SOCKEYE SALMON DELPHI - ROUND 1

1. Consider the relationships between riverine habitat suitability for sockeye salmon for each of the variables -- velocity, depth, substrate, cover and temperature. What is the relationship between each variable and habitat suitability for the various life stages/activities (e.g., inmigration, spawning, incubation, preemergent larval, postemergent larval, outmigrant, or other life stage or activity).

2. Next, fill in the columns of each of the tables (attached). List references, data sources, or any information available which you may use as the basis of your curve. It is important that you use your gut feeling or opinion if no data are available. You may choose to ignore all available data or information and use only your gut feeling or opinion as the basis of your curve. If you mention a reference to data, please give the complete citation or send the evaluator a copy of the report. If the reference has been published in a popular journal or has been widely circulated and is likely available in libraries, you need not send it.

3. Write comments, ideas, logic, reference, etc., at the bottom of each table or on the reverse of the page.

4. If you feel that a variable or a life stage other than those listed in a table is important and should be considered for an SI curve, please clearly define the variable, how the variable is quantified, and what the specific size-group, season or unique life stage/activity the variable applies to.

5. If you have questions you may call Johnie Crance (303)226-9318 or FTS 323-5318.

6. Please return your response within 10 days to:

Johnie H. Crance
U.S. Fish and Wildlife Service
National Ecology Center
Division of Wildlife and Contaminant Research
Creekside One Building
2627 Redwing Road
Fort Collins, CO 80526-2899

SOCKEYE SALMON-DELPHI EXERCISE ROUND 1.

Cover

Date ____________

Panelist ____________

Describe what you consider to be important cover (i.e., type of cover important to the well-being of the species) for any life stage/activity of sockeye salmon. Please describe what the cover is, how it is used, how it may be quantified in relation to habitat suitability, what happens if there is more cover, less cover, no cover, etc. Sketch your version of any cover SI considered to be important. Use the space below and reserve side of page, if needed.

SOCKEYE SALMON - DELPHI EXERCISE ROUND 1 - % FINES IN REDD

Complete this table by filling in each column with the percent fines[a] (< 0.3 mm) in redd appropriate for the life stage/activity of the species.

% fines condition	% fines < 0.3 mm diameter		
	spawning	incubation	preemergent fry
1. Lowest % fines considered optimal[b]			
2. Highest % fines considered optimal[b]			
3. % fines for SI=0.5			
4. Highest % fines considered useable			
5. % fines must increase to for SI=0[c]			

[a] What do you mean by % fines. Please explain below or on reserve of page.

[b] SI=1.0

[c] % fines is totally unsuitable.

SOCKEYE SALMON - DELPHI ROUND 1 - WATER VELOCITY

Date ____________________

Panelist ________________

Complete this table by filling in each column with the water velocity (ft/sec) appropriate for the life stage/activity of the species.

	Velocity (feet/second)						
Velocity Condition	Immigrant	Spawning	Incubation	Preemergent larval	Postemergent larval	Outmigrant	Other
1. Minimum velocity used.							
2. Maximum velocity used.							
3. Lowest velocity considered to be optimal.							
4. Highest velocity considered to be optimal.							
5. Level velocity must decrease to for SI=0 (use N if never occurs)							
6. Level velocity must increase to for SI=0 (use N if never occurs).							
7. Velocity level(s) where SI=0.5 (use N is never occurs).							

Generally the mean column velocity (velocity at 0.6 of depth measured from water surface). However, more specific measurements are used sometimes. What do you mean by velocity relative to the values you will give in this table? Underline the following phrase that most closely describes your use of velocity: Velocity at surface of water. Velocity within 6 inches of stream bottom. Velocity at site of fish/activity (e.g., nose velocity). Other (please define) ____________________.

Specify any other riverine life stage/activity that you consider to be important and fill in column.

SI=1.

Velocity level is totally unsuitable.

SOCKEYE SALMON - DELPHI ROUND 1 - WATER DEPTH

Date ______________________________

Panelist _________________________

Complete this table by filling in each column with the water depth (feet) appropriate for the life stage/activity of the species.

Depth Condition	Water Depth (feet)						
	Inmigrant	Spawning	Incubation	Preemergent larval	Postemergent larval	Outmigrant	Other
1. Minimum depth used.							
2. Maximum depth used.							
3. Minimum depth considered optimal.							
4. Maximum depth considered optimal.							
5. Depth water must decrease to for SI=0.							
6. Depth water must increase to for SI=0 (use N if never occurs).							
7. Depth(s) where SI=0.5.							

Indicate what you mean by depth in the context of the values you will use in this table by underlining the following phrase that most clearly describes your use depth: Minimum water depth. Average water depth. Nose depth or depth at fish/egg/activity. Other (please define) ______________________________.

Specify any other riverine life stage/activity you consider to be important and fill in column.

SI=1.

Depth is totally unsuitable.

SOCKEYE SALMON - DELPHI EXERCISE ROUND 1 - Water Temperature

Date ______________________

Panelist ____________________

Complete this table by filling in each column with the water temperature (°F) appropriate for the life stage/activity of the species.

Temperature Condition	Water condition (°F)						
	Inmigrant	Spawning	Incubation	Preemergent larval	Postemergent larval	Outmigrant	Other
1. Minimum temperature used.							
2. Maximum temperature used.							
3. Lowest temperature considered to be optimal.							
4. Highest temperature considered to be optimal.							
5. Temperature water must decrease to for SI=0.							
6. Temperature water must increase to for SI=0.							
7. Temperature(s) where SI=0.5.							

Average mean daily water temperature at warmest time of day, usually mid-afternoon.

Specify any other riverine life stage/activity that you consider to be important and fill in column.

SI=1.

Temperature is totally unsuitable.

SOCKEYE SALMON - DELPHI EXERCISE ROUND 1. <u>Substrate</u>

Date ______________________

Panelist ______________________

Complete this table by filling in each column with the appropriate SI value (0.0-1.0) for the substrate -life stage/activity.

Substrate Type Code Particle Size	Suitability Index (0.0-1.0)						
	Immigrant	Spawning	Incubation	Preemergent larval	Postemergent larval	Outmigrant	Other
1. Organic material							
2. Mud/soft clay							
3. Silt, 0.062 mm							
4. Sand, 0.062-2 mm							
5. Gravel, 2-64 mm							
6. Cobble, 64-250 mm							
7. Boulder, 250-4000 mm							
8. Bedrock							

Substrate is totally unsuitable when SI=0. If substrate is optimal, SI=1.

Indicate what you mean by substrate in the context of how you will use it for this table. Underline the following phrase that most closely describes your meaning: Dominant substrate particles observed on surface of substrate. Material comprising highest percentage (by weight) of grab sample. Other (please define) ______________________

Specify other riverine life stage/activity that you consider to be important and fill in column.

APPENDIX B

Generalized Equipment List for Field Studies With Partial List of Suppliers

APPENDIX B. EQUIPMENT FOR VARIOUS CAPTURE, OBSERVATION, AND MEASUREMENT TECHNIQUES

Items marked by star (*) are recommended

Item Description	Approximate Unit Cost $
GENERAL	
* Topsetting wading rod	
4'; 6'; 8'; and 10' length	200.00 - 400.00
Current Meter	
Marsh-McBirney Electromagnetic	1700.00 - 3000.00
* Price AA	500.00 - 600.00
* Pigmy	300.00 - 400.00
Headset[1]	30.00 - 40.00
Stopwatch[1]	30.00 - 50.00
* Diver's stopwatch[1]	150.00 - 300.00
Substrate Analysis wire grid	
wire twisted	100.00 - 150.00
wire brazed	200.00 - 250.00
Dredges	
Emery	190.00 - 200.00
Ekman	175.00 - 200.00
Ponar	335.00 - 950.00[2]
Dredge accessories	
Cable and hoist	335.00 - 675.00
Ekman dredge handle (for depths <5')	75.00 - 80.00
Freeze core sampler (support equipment + 32 probes)	5200.00
Wet sieves	
.062 mm - 16 mm stack (plastic-stainless)	55.00 - 300.00
Selected sieves	
Brass	35.00 - 50.00
Stainless steel	45.00 - 70.00

[1]Applied only to Price AA and Pigmy meters
[2]Stainless steel

Item Description	Approximate Unit Cost $
Boats and boat equipment	
Boat, 14'-16', flat bottom or tri-hull, with trailer	1000.00 - 2000.00
25 hp outboard motor	1600.00 - 2000.00
Electric trolling motor	800.00 - 1100.00
Raft, 4-5 person inflatable, wood transome and floor	1500.00 - 3500.00
Gasoline powered compressor	500.00 - 800.00
Extra oars	15.00 - 30.00
Life vests, USCG approved, type III	30.00 - 50.00
Chest waders, insulated, felt sole	150.00 - 250.00
* Chest waders, neoprene, wading boots (except electrofishing)	250.00 - 350.00
Thermometer	
Hand held	25.00 - 50.00
* Immersible max/min	75.00
Portable Water Chemistry Kits	
Dissolved oxygen	130.00
Alkalinity	100.00
* Multiparameter	180.00 - 1600.00
Portable Water Chemistry Meters	
Conductivity	250.00 - 550.00
Dissolved oxygen	200.00 - 1100.00
pH	185.00 - 550.00
* Multiparameter	800.00 - 1600.00
* Secchi disk (limnological type) with calibrated sounding line	65.00 - 210.00
Field Books and Data Recording	
Spiral, Rag Paper	1.50 - 3.00
* Bound, Rite in Rain	7.00 - 15.00
* Bound, Waterproof	10.00 - 20.00
* First Aid Kit	5.00 - 40.00
SURFACE OBSERVATION	
* Polarized sunglasses	7.00 - 100.00
* Tree stand kit	
Tree stand platform	70.00 - 105.00
Climbing safety harness (or)	18.00
Tree steps (twist-in) per pair	4.00
Tree seat	18.00

Item Description	Approximate Unit Cost $
Ladder tree stand	85.00
Poision ivy cream	4.00 - 6.00
Wasp/hornet spray	4.00 - 6.00
Binoculars	40.00 - 500.00
* with polarizing filter	140.00

UNDERWATER OBSERVATION

* Dive mask	25.00 - 85.00
* Snorkel	12.00 - 40.00
* Bouyancy compensator	
Horse collar	100.00 - 200.00
Stabilization jacket	150.00 - 550.00
* Wet suit (with hood) for temperatures above 10 °C	450.00 - 1100.00
* Dry suit for temperatures below 10 °C	
* Rubber galoshes (if wearing booties)	15.00 - 25.00
* Canvas tennis shoes (if not wearing booties)	20.00 - 40.00
Swim fins	40.00 - 60.00
* Goody bag	10.00 - 20.00
* 1-2 oz. lead weights (for markers)	0.15 - 0.30
Surveyor's flagging tape	0.90 - 1.75/ 300 ft roll
Indelible marking pens	1.50 - 2.00
* Dive cuff	20.00 - 30.00
SCUBA tank	140.00 - 260.00
Regulator	200.00 - 400.00
Dive weights and belt	10.00 - 30.00 + $3/lb.
Depth gauge	35.00 - 120.00
Dive light	9.00 - 400.00
Cyalume light sticks	3.00/each
Polypaper (8½ x 11 sheets)	25.50/pkg of 100
Writing scroll, polyester drafting film	74.50/30" x 2 yd. roll
Static line (1/8" aircraft cable)	0.75/ft
Wire rope grips (cable pullers)	20.00 - 30.00
Cable real	200.00 - 700.00
Hand winch (come-along)	15.00 - 20.00
1/2" - 3/4" polypropylene rope	0.20 - 0.50/ft
Mountaineering ascenders (Gibbs, or CMI shorty)	25.00 - 60.00
Velcro straps (for closing wetsuit openings)	3.00 - 5.00

REMOTE UNDERWATER OBSERVATION

Submersible periscope or view tube (homemade)	200.00 - 300.00
Underwater video camera with monitor-recorder	2200.00 - 3000.00
Remote control swivel	?
Base unit	?
Crane and winch	375.00 - 700.00
Monuments	0.50

Item Description	Approximate Unit Cost $
RADIOTELEMETRY	
Search receiver	2000.00
Pinpointing receivers (minimum of 2)	2000.00
Directional antennas (minimum of 2)	
Yagi	200.00 - 300.00
Square loop	100.00
Implant transmitters and batteries	200.00
Surgical equipment/kit	50.00
Anesthetic	75.00
ELECTROFISHING	
Prepositioned area shocker	
12-gauge solid strand copper wire[3] (rubber insulated)	.05/ft
Anchors (4 per electrode - 4 kg steel bars)	0.45/kg
Extension cords (10 amp rating, 100 ft)	12.00 - 15.00
Generator (1500 - 3000 watt)	600.00 - 1000.00
Remote control box	10.00 - 25.00
DC rectifier	1500.00 - 2500.00
Dip nets	30.00 - 50.00
Live boxes	50.00
Battery powered backpack shocker	1200.00 - 1800.00
SEINES, NETS, TRAPS	
Miniature purse seine (materials only)	150.00 - 250.00
Trammel nets	70.00 - 250.00
Cast nets	15.00 - 45.00
Lift nets	35.00 - 90.00
Surber samplers	100.00 - 125.00
Hess samplers	300.00
Modified Hess drop traps	400.00 - 500.00
Plankton nets	35.00 - 50.00
Drift nets	150.00 - 250.00
Pump samplers	400.00 - 800.00
Slurp guns (build from volume pump with foot valve)	20.00 - 100.00
EXPLOSIVES	
Primacord®	Contact sales representative[4]
Detonators	
Electric blasting caps	
Nonel® shock tube (nonelectric)	Contact sales representative[4]
Nonel® starter	Contact sales representative[5]

[3]According to Bain et al. (1985) approximately 35-70 ft. of wire are needed per electrode.
[4]Ensign-Bickford Co. (see list of suppliers)
[5]Powder Horn Supply, P.O. Box 230, Adams Center, NY 13606. (315) 583-5654.

PARTIAL LIST OF SUPPLIERS

The following list of suppliers is provided as a service to researchers engaged in habitat criteria research. It is representative, but by no means comprehensive. There are, undoubtedly, additional sources for much of this equipment, so the wise investigator will shop around before purchasing major equipment. Mention of specific companies or trademarks in this list does not constitute endorsement.

Stream Gaging Equipment

U.S. Geological Survey
Hydrologic Instrumentation Service
Gulf Coast Hydroscience Center
Building 2101
NSTL Station, MS 39529
FTS 494-2108
1-601-688-2108

Scientific Instruments
518 Cherry Street, West
Milwaukee, WI 53212
Comm. 1-414-263-1600

Montedoro-Whitney
2741 E. McMillan Road
San Luis Obispo, CA 93401
1-800-235-4104

McMaster-Carr Supply Co.
P.O. Box 4365
Chicago, IL 60680
1-312-833-0300

Teledyne Gurley
514 Fulton Street
Troy, NY 12181
1-710-443-8156

Marsh-McBirney, Inc.
8596 Grovemont Circle
Gaithersburg, MD 20877
1-301-869-4700

WESCO
3895 Joliet
Denver, CO
1-800-625-0266

McMaster-Carr Supply Co.
P.O. Box 54960
Los Angeles, CA 90064
1-213-945-2811

Superex Electronics Corp.
161 Ludlow Street
Yonkers, NY 10706
1-914-965-6906

Dredges, Grabs, Sieves, Winches

Scientific Instruments (see above)
U.S. Geological Survey (see above)

Wildco Wildlife Supply Company
301 Cass Street
Saginar, MI 48602
1-517-799-8100

McMaster-Carr (see above)

Kahl Scientific Instrument Corp.
P.O. Box 1166
El Cajon, CA 92022
1-714-444-2158 and
1-714-444-5944

Freeze Core Samplers

(See Platts and Penton 1980 for specifications and suppliers)

Boats, Rafts, and Related Equipment

Refer to yellow pages under Boat Dealers

Note: Most Federal agencies must purchase boats and rubber rafts, motors, and boating accessories from a list of suppliers available from GSA Office of Federal Supply Services. Ask for:

Federal Supply Schedule (FSC) No. 19
Small craft, marine equipment, and
floating marine barriers.

Non-Federal users may also wish to obtain this schedule.

Portable Water Chemistry Kits/Meters

Hach Company
P.O. Box 389
Loveland, CO 80539
1-800-525-5940

Forestry Suppliers, Inc.
205 W. Rankin Street
P.O. Box 8397
Jackson, MS 39204-0397
1-601-354-3565

Ben Meadows Company
3589 Broad Street
P.O. Box 80549
Atlanta (Chamblee) GA 30366
1-800-241-6401
1-800-241-3161 (in Georgia)

Cole-Parmer Instrument Co.
7425 N. Oak Park Avenue
Chicago, IL 60648
1-800-323-4340

Thomas Scientific
P.O. Box 99
Swedesboro, NJ 08085-0099
Main Office: 1-215-574-4500
(call for Regional toll free
numbers)

Ben Meadows Company
2601-B West 5th Ave.
P.O. Box 2781
Eugene, OR 97402
1-800-547-8813
1-800-452-9010 (in Oregon)

Diving Gear

See yellow pages under Sporting Goods

Radiotelemetry Equipment

Smith-Root, Inc.
14014 N.E. Salmon Creek Ave.
Vancouver, WA 98665
1-206-573-0202

Advanced Telemetry Systems
23859 N.E. Highway 65
Bethel, MN 55005
1-612-434-5040

Custom Telemetry and Consulting
185 Longview Drive
Athens, GA 30605
404-548-1024

AVM
6575 Trinity Court
Dublin, CA 94566
1-415-449-2286

Electrofishing Equipment

Coffelt Electronics
3910 S. Windermere St.
Englewood, CO 80110
1-303-761-3505

Dirigo Electronics Engineering
1307 NW Buchanon
Corvallis, OR 97330
1-503-752-5337

Smith-Root, Inc.
(see address under Radiotelemetry)

Generators, wire, extension cords - see yellow pages under Hardware, Electrical Supplies, Farm and Ranch Supplies.

Seines and Nets

Memphis Net and Twine Co., Inc.
2481 Matthews Ave.
P.O. Box 8331
Memphis, TN 38108
1-901-458-2656

Wildco Wildlife Supply
(see address under
Dredges, Grabs)

Kahl Scientific Instruments
(see address under
Dredges, Grabs)

Explosives

Primacord® - Ensign-Bickford Sales Offices -

660 Hopmeadow St.
Simsbury, CT 06070
1-23-658-4411

P.O. Box 97
Louviers, CO 80131
1-303-798-8625

5011 Washington Ave.
Evansville, IN 47715
1-812-476-1329

P.O. Box 322
Wexford, PA 15090
1-412-935-5712

5036 Snapfingers Drive
Decatur, GA 30035
1-404-987-1000

1325 Airmotive Way
Reno, NV 89502
1-702-786-7822

Nonel® starter - Powder Horn Supply
P.O. Box 230
Adams Center, NY 13606
1-315-583-5654

Specialty Items

Wire rope grips - Electrical suppliers such as
Graybar Electrical (National chain)

D&M Wire Rope, Inc.
682 W. Gunnison
Grand Junction, CO
1-303-342-1144

M. Klein & Sons, Inc. (manufacturer)
Chicago, IL
(Part # 1613-30)

Tree stand kits - Cabela's
812 13th Avenue
Sidney, NE 69160
1-800-237-4444

Forestry Suppliers, Inc.
(see address under Portable Water Chemistry Kits)

Ascenders - See yellow pages under Sporting Goods:
Mountaineering

Holubar Mountaineering
3500 South College Avenue
Ft. Collins, CO 80526
1-303-226-3683

Polypaper/Polyester Drafting Film

Ben Meadows Company
(see address under Portable Water Chemistry Kits)

Under Water Video Systems - See yellow pages under Sporting Goods: Dive Shops. Many dive shops can offer specific advice regarding the performance of various systems. They will often allow or demonstrate field tests of the equipment. There are many brand names to select from, but it is a good idea to try it before purchase.

APPENDIX C

Some Graphs and Their Functions

Excerpts reprinted with permission from: W.J. Parton and G.S. Innis. 1972. Some graphs and their functional forms. U.S. International Biological Program, Tech. Rep. 153, Colorado State University, Ft. Collins, CO. 41 pp.

ARCTANGENT FUNCTION

Functional Form

$$f(x,a,b,c,d) = b + \frac{c}{\pi} \arctan [\pi d(x - a)]$$

Derivative

$$f'(x,a,b,c,d) = cd \left[\frac{1}{1 + [\pi d(x - a)]^2}\right]$$

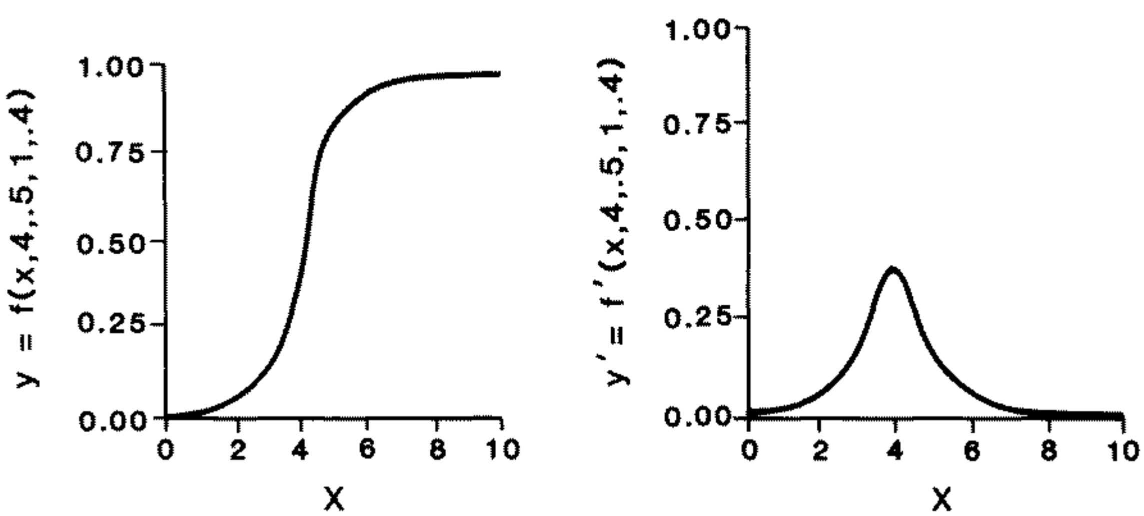
Parameter Definitions

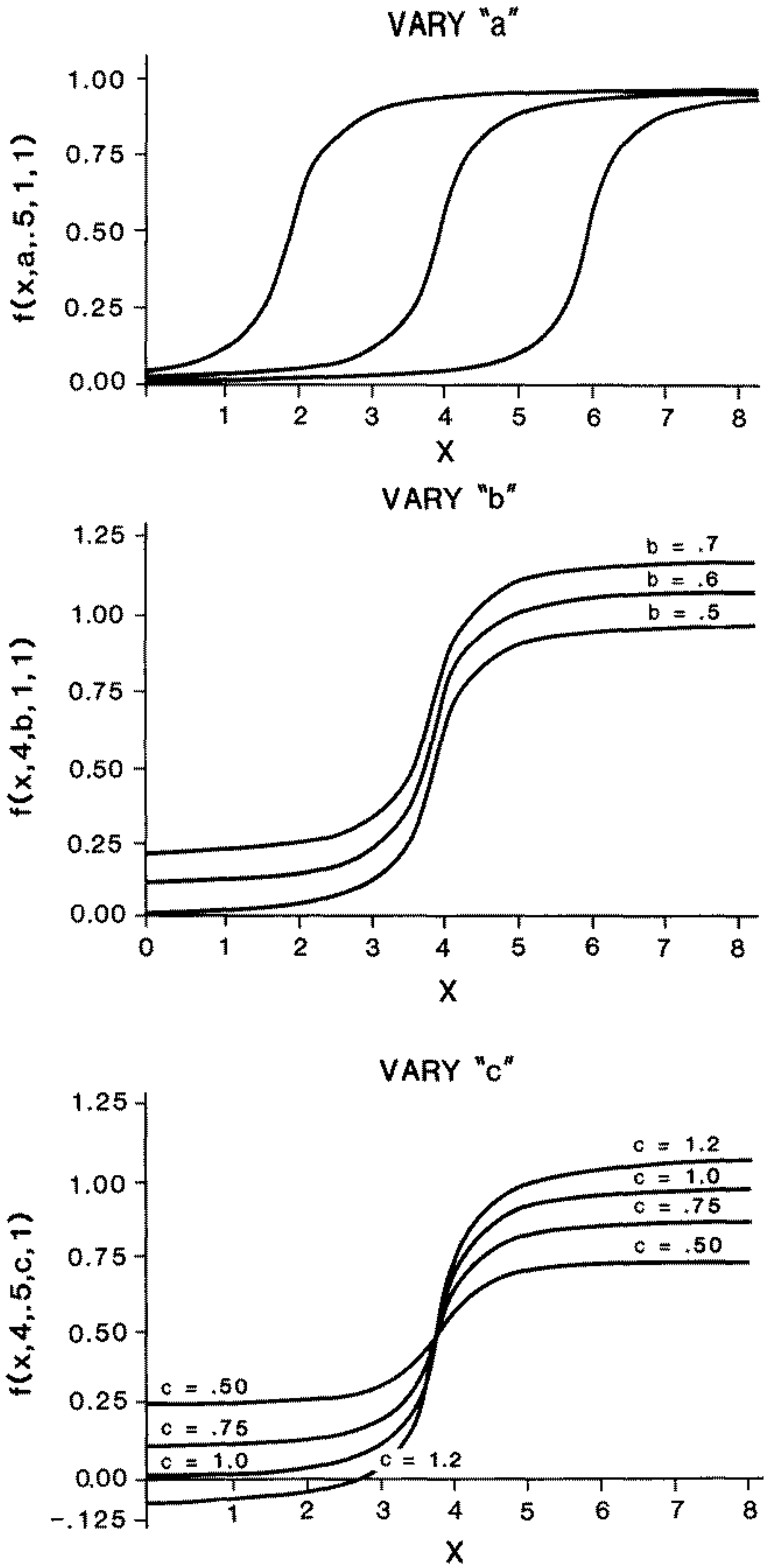
a = "x" location of inflection point

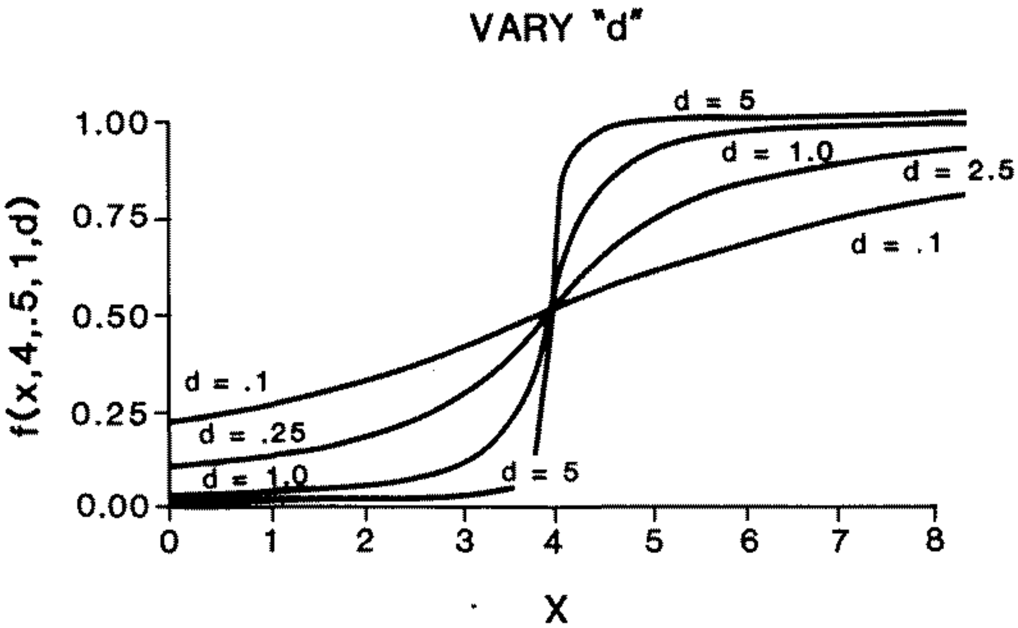
b = "y" location of inflection point

c = step size (distance from the maximum point to the minimum point)

d = slope of line at inflection point




VARY "a"
1.00
0.75
0.50
0.25
0.00
f(x,a,.5,1,1)
1
2
3
4
5
6
7
8
X
VARY "b"
1.25
1.00
0.75
0.50
0.25
0.00
f(x,4,b,1,1)
b = .7
b = .6
b = .5
0
1
2
3
4
5
6
7
8
X
VARY "c"
1.25
1.00
0.75
0.50
0.25
0.00
-.125
f(x,4,.5,c,1)
c = 1.2
c = 1.0
c = .75
c = .50
c = .50
c = .75
c = 1.0
c = 1.2
1
2
3
4
5
6
7
8
X



FORTRAN CODE

```
      FUNCTION ATANF (X,A,B,C,D)
      ATANF=B+(c/3.14159)*ATAN(3.14159*D*(X-A))
      RETURN
      END
C...  NOTE THAT 3.14159 IS AN APPROXIMATION TO PI
C...  AND THAT ATAN IS A SYSTEM SUPPLIED ROUTINE
C...  THAT COMPUTES THE PRINCIPAL BRANCH (PI/2>
C...  ATAN (x)>-PI/2)
```

GENERALIZED POISSON DENSITY FUNCTION

Functional Form

$$f(x,a,b,c,d) = \left[\frac{b-x}{b-a}\right]^{c} \cdot e^{\left(\frac{c}{d}\right)\cdot\left[1-\left(\frac{b-x}{b-a}\right)^{d}\right]}$$

Derivative

$$f'(x,a,b,c,d) = e^{\left[\left(\frac{c}{d}\right)\left[1-\left(\frac{b-x}{b-a}\right)^{d}\right]\right]} \cdot \left(\frac{b-x}{b-a}\right)^{c-1} \cdot \frac{c}{b-a} \cdot \left[\left(\frac{b-x}{b-a}\right)^{d} -1\right]$$

Parameter Definitions

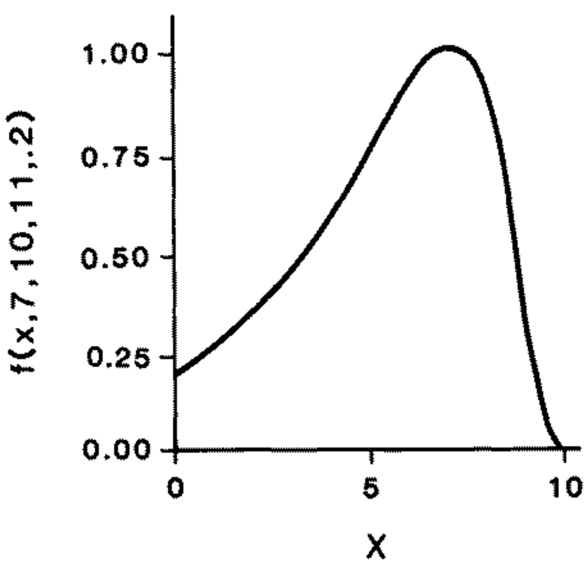
a = value of "x" where f(x) = 1.0

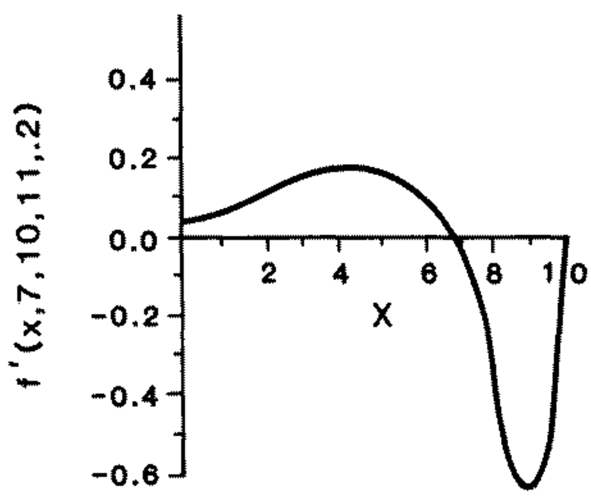
b = value of "x" where f(x) = 0.0 (x < b)

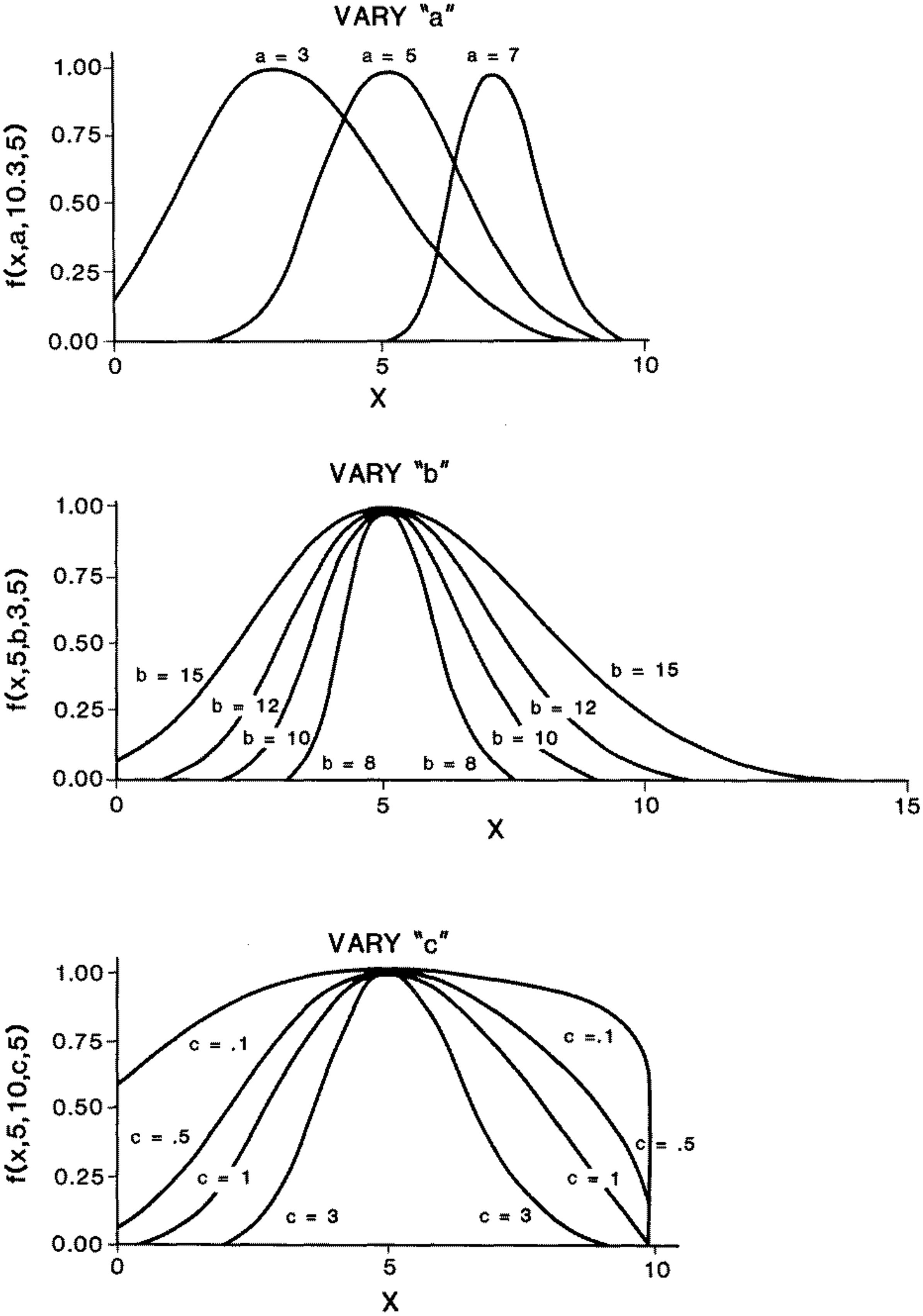
c = shape parameter for part of the curve to the right of x = a

d = shape parameter for part of the curve to the left of x = a

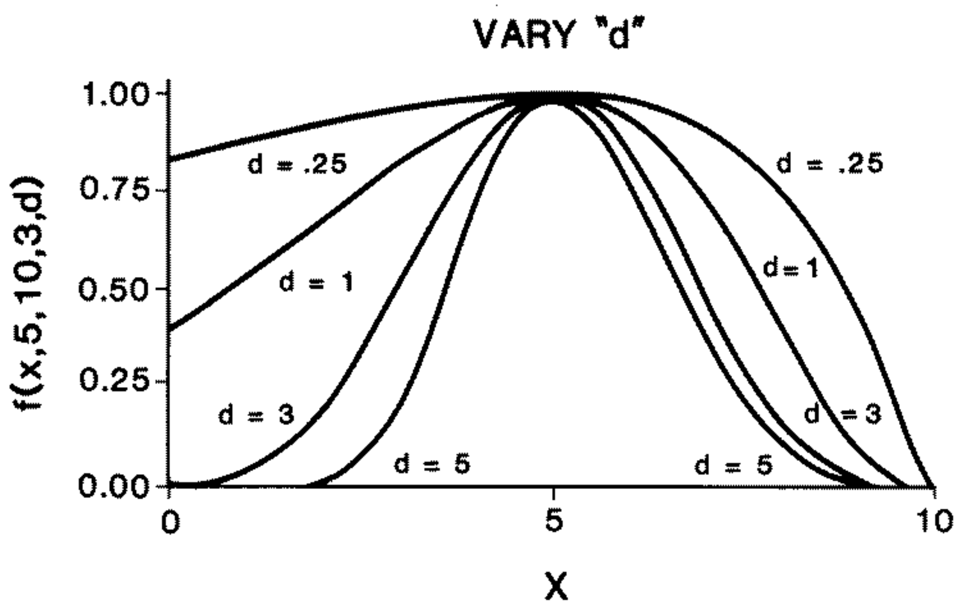
e = base of the natural logarithm ≈ 2.71828







VARY "a"
a = 3
a = 5
a = 7
f(x,a,10,3,5)
1.00
0.75
0.50
0.25
0.00
0
5
10
X
VARY "b"
f(x,5,b,3,5)
b = 15
b = 12
b = 10
b = 8
b = 8
b = 10
b = 12
b = 15
0
5
10
15
X
VARY "c"
f(x,5,10,c,5)
c = .1
c = .5
c = 1
c = 3
c = 3
c = 1
c = .5
c =.1
0
5
10
X



FORTRAN CODE

```
      FUNCTION GPDF (X,A,B,C,D)
      FRAC=(B-X)/(B-A)
      IF (FRAC.LE.0.) GO TO 1
      GPDF=EXP(C/D*(1.-FRAC**D))
      GPDF=(FRAC**C)*GPDF
      RETURN
    1 GPDF=0
      RETURN
      END
C... EXP IS A SYSTEM SUPPLIED FUNCTION WHICH
C... EXPONENTIATES (BASE E) THE ARGUMENT.
C... THE IF STATEMENT IS INCLUDED TO ASSURE
C... THAT ONE DOES NOT ATTEMPT TO EXPONENTIATE
C... A NEGATIVE NUMBER. N.B. IF X>B. GPDF=0
C... ALTHOUGH THE FUNCTION IS REALLY NOT DEFINED
C... THERE.
```

NATURAL GROWTH FUNCTION

Functional Form

$$f(x,a,b) = a(1 - e^{-bx})$$

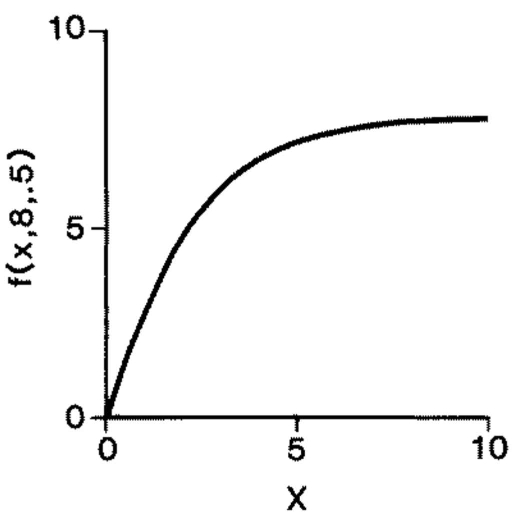
Derivative

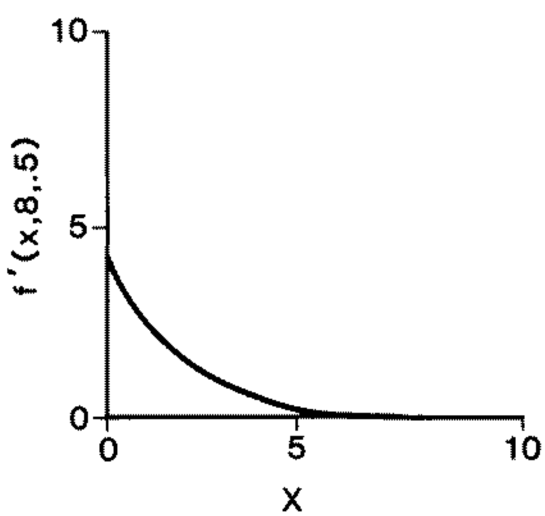
$$f'(x,a,b) = abe^{-bx}$$

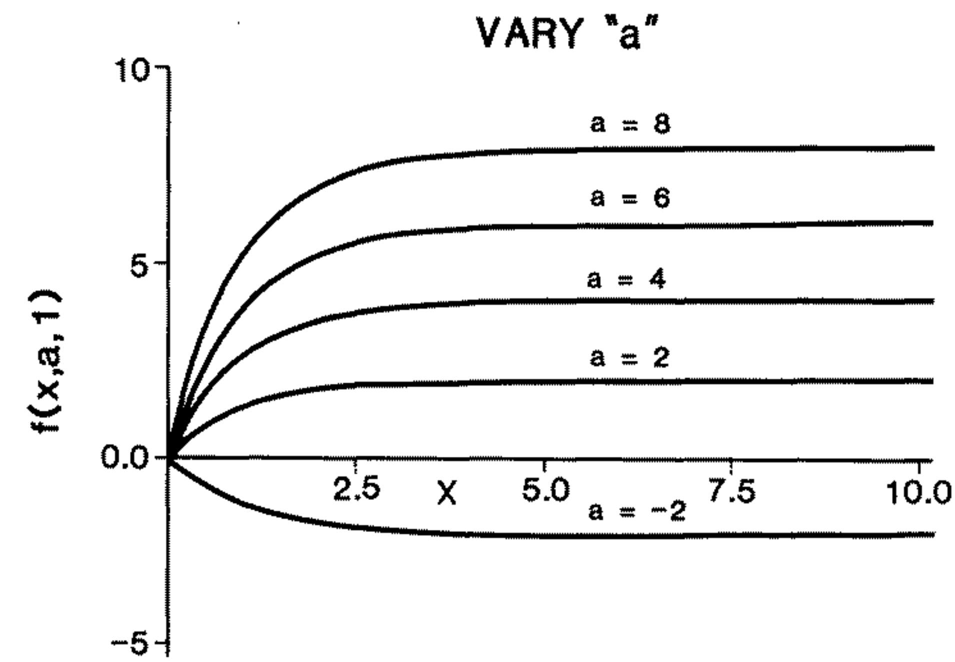
Parameter Definitions

a = the maximum or minimum value of $f(x)$

b = parameter that controls the rate which $f(x)$ approaches "a"

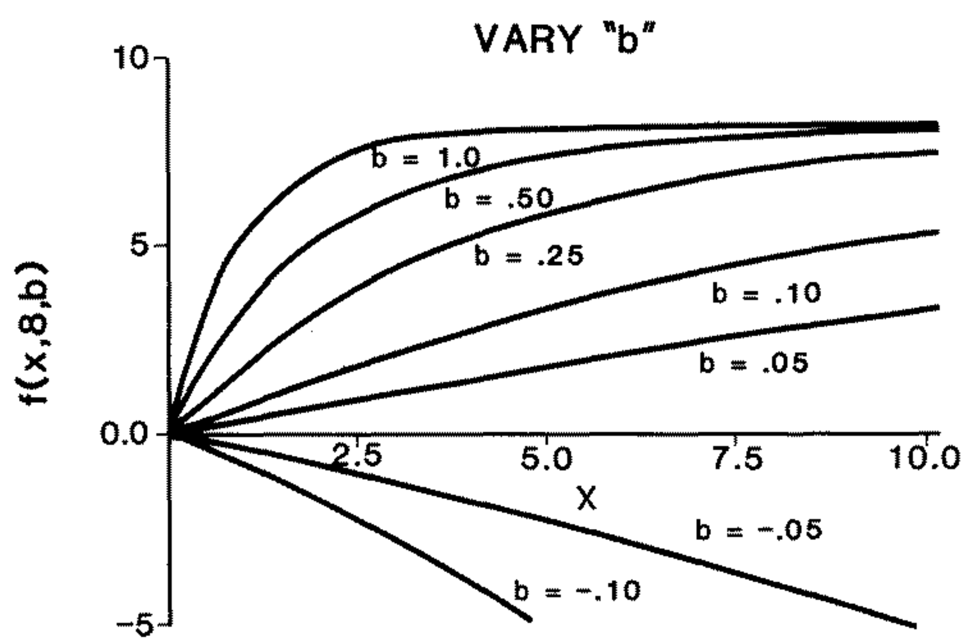






FORTRAN CODE

```
      FUNCTION GFN(X,A,B)
      GFN = A*(1.-EXP(-B*X))
      RETURN
      END
```



LOGISTIC FUNCTION

Functional Form

$$f(x,a,b,c) = \frac{a}{1 + be^{-cx}}$$

Derivative

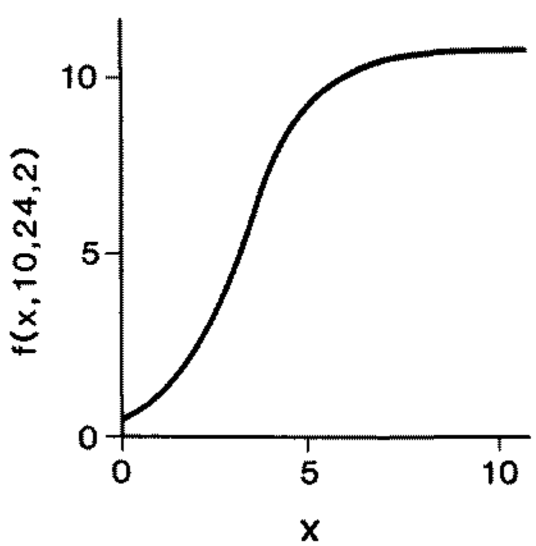
$$f'(x,a,b,c) = \frac{abce^{-cx}}{(1 + be^{-cx})^2}$$

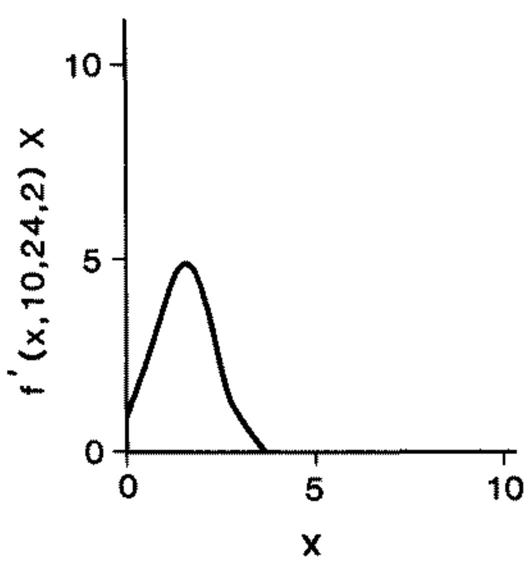
Parameter Definitions

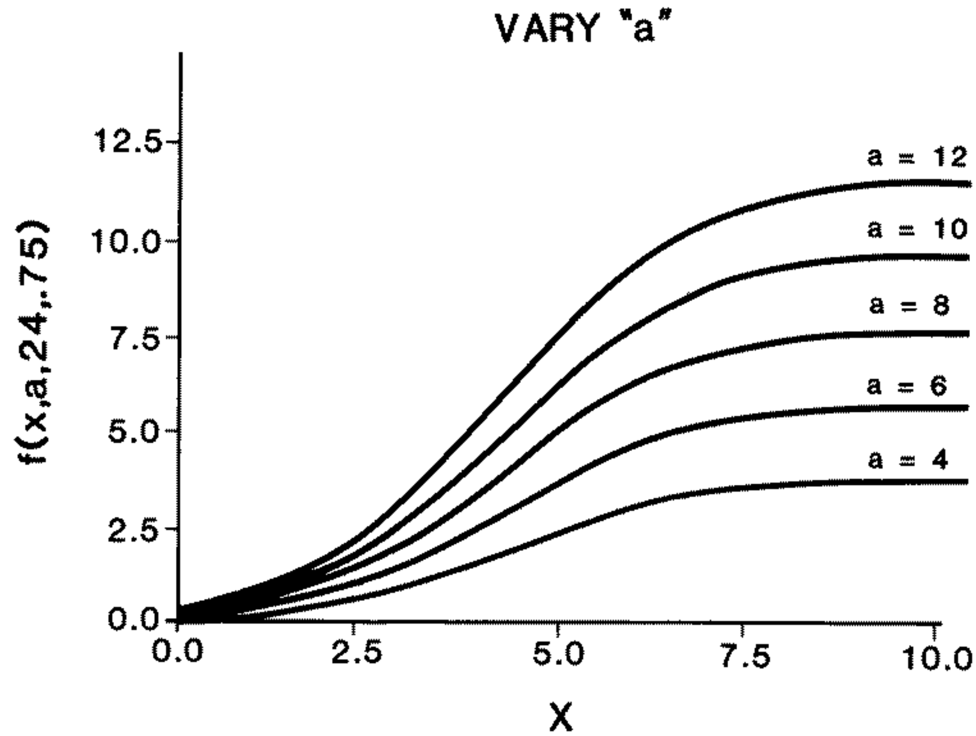
a = the maximum value of f(x) [f(x) equals $\frac{a}{2}$ at the inflection point of the curve]

b = control parameter for value of f(x) when x = 0.0

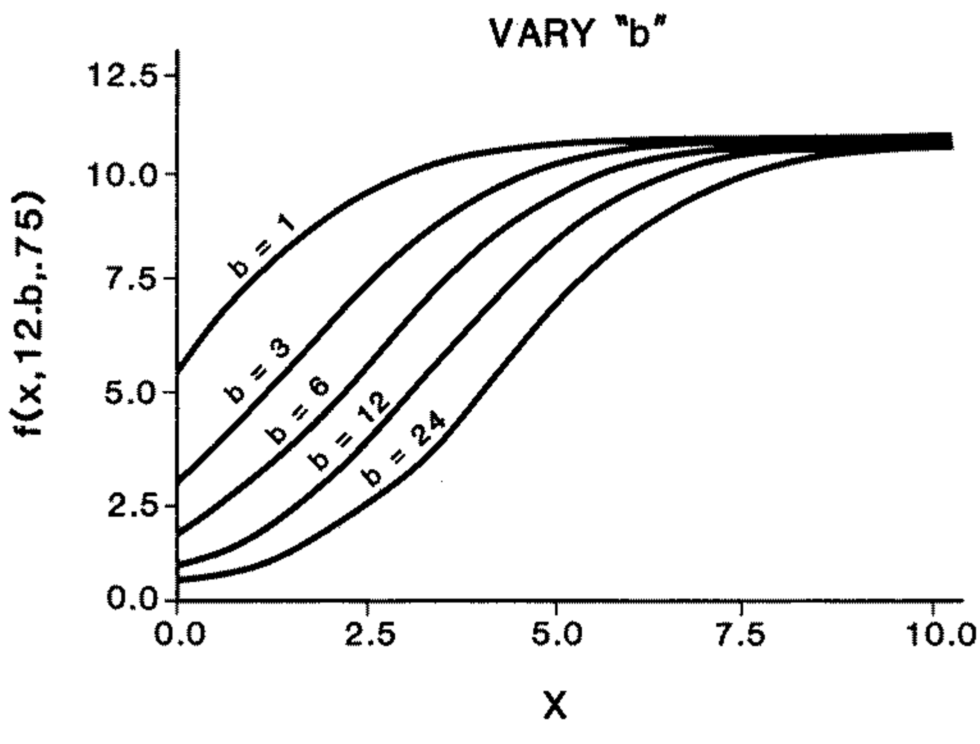
c = control parameter for the value of "x" at the inflection point of the curve



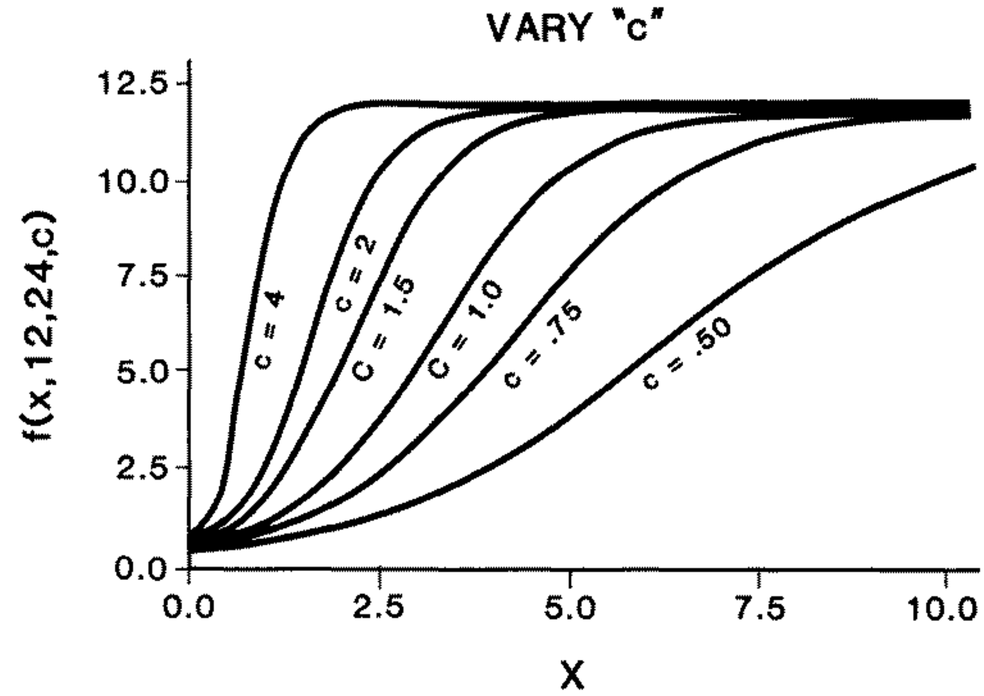




VARY "a"
f(x,a,24,.75)
12.5
10.0
7.5
5.0
2.5
0.0
0.0
2.5
5.0
7.5
10.0
X
a = 12
a = 10
a = 8
a = 6
a = 4



VARY "b"
f(x,12,b,.75)
12.5
10.0
7.5
5.0
2.5
0.0
0.0
2.5
5.0
7.5
10.0
X
b = 1
b = 3
b = 6
b = 12
b = 24



FORTRAN CODE

```
      FUNCTION FL(X,A,B,C)
      FL = A/(1.+B*EXP(-C*X))
      RETURN
      END
```

GENERALIZED GOMPERTZ EQUATION

Functional Form

$$f(x,a,b,c,d) = (ab)^{(-cb)^{(-dx)}}$$

Derivative

$$f'(x,a,b,c,d) = (acdb)^{(-cb)^{(-dx)}} \bullet (\ell n_e b)^2 \bullet (b)^{(-dx)}$$

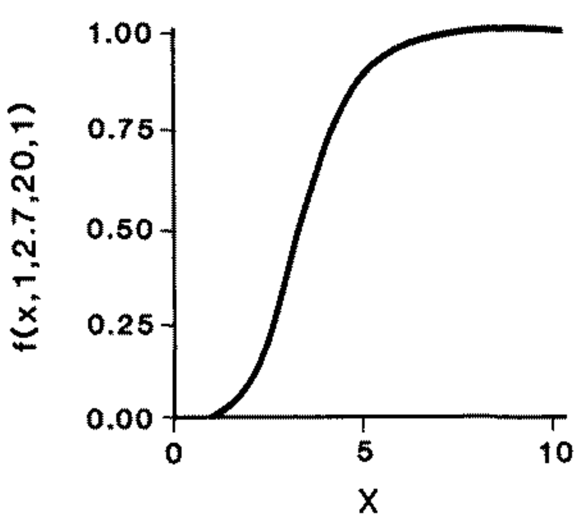
Parameter Definitions

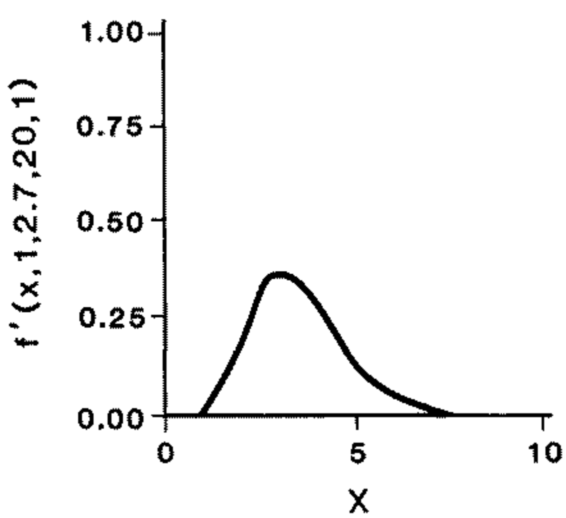
a = the maximum value of f(x)

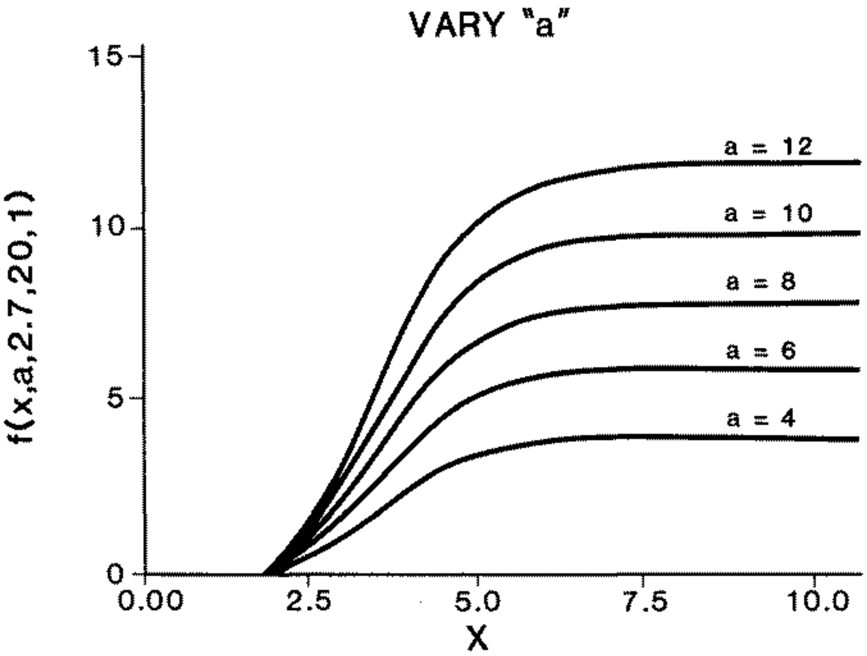
b = control parameter that changes value of f(x) where the inflection point is located [f(x) ≈ a/b at the inflection point for values of "b" between 2 and 6]

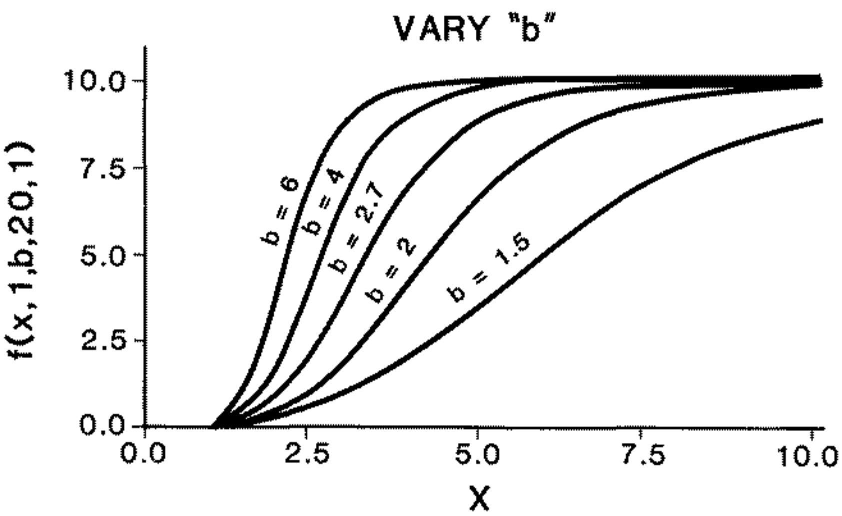
c = control parameter that moves the "x" location of the inflection point

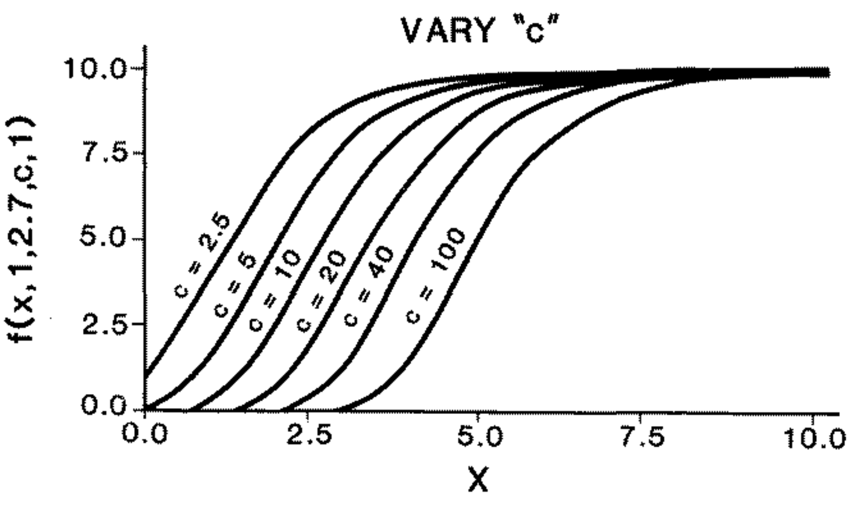
d = control parameter that changes the slope of the curve at inflection point

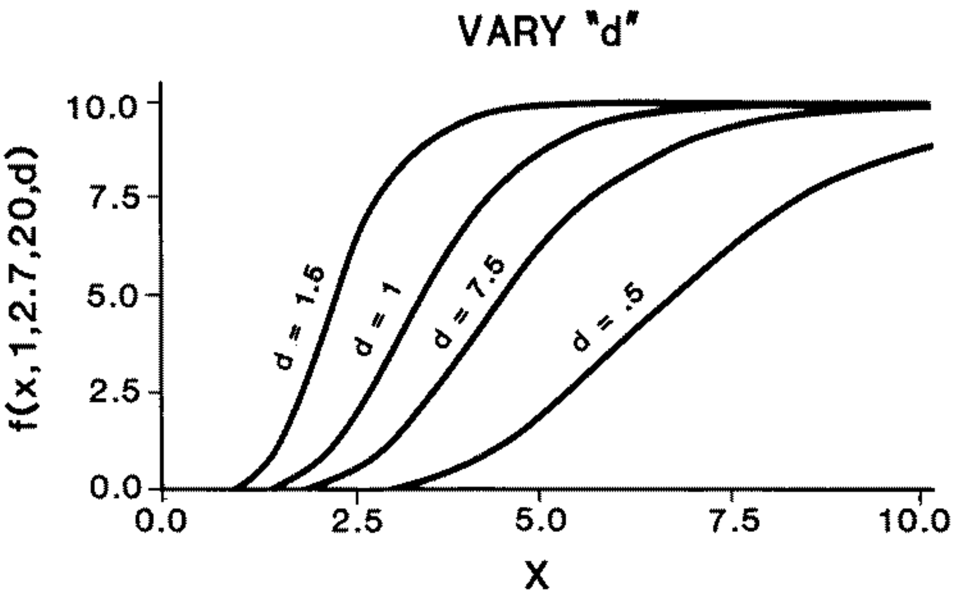





VARY "a"
f(x,a,2.7,20,1)
X
0
5
10
15
0.00
2.5
5.0
7.5
10.0
a = 12
a = 10
a = 8
a = 6
a = 4


VARY "b"
f(x,1,b,20,1)
X
0.0
2.5
5.0
7.5
10.0
b = 6
b = 4
b = 2.7
b = 2
b = 1.5


VARY "c"
f(x,1,2.7,c,1)
X
0.0
2.5
5.0
7.5
10.0
c = 2.5
c = 5
c = 10
c = 20
c = 40
c = 100



FORTRAN CODE

```
      FUNCTION GGEF (X,A,B,C,D)
      IF (B.LE.0.) GO TO 1
      GGEF=C*B**(-D*X)
      GGEF=A*B**(-GGEF)
      RETURN
   1  GGEF=0.
      RETURN
      END.
C...  THIS FUNCTION IS GENERALLY MULTIPLE VALUED AND COMPLEX
C...  IF B≤0.  THE ROUTINE RETURNS THE VALUE 0.
```

APPENDIX D

Program listing for GETFISH, a computer program for conducting Monte Carlo tests of habitat suitability criteria. Reprinted with permission of the author, Dr. K. A. Voos.

APPENDIX D

```
 PROGRAM GETFSH     73/74     OPT=2          FTN 4.8+498     82/09/09 . 08.40.13

              PROGRAM GETFSH (TAPE5, TAPE6, INPUT, OUTPUT)
        C
              DIMENSION P(2000), V(2000), D(2000), S(2000)
              DIMENSION IPNT (2000)
        C
              REWIND 5
              REWIND 6
        C
              MXIT = 1000000
              PMX = 0.0
              IFSH = 0
              ITCNT = 0
              SEED = 0.0
              ATOT = 0.0
              CALL SECOND (SEED)
              CALL RANSET (SEED)
        C
              I=0
         10   I = I + 1
              READ (5,5000) A, D(I), V(I), S(I)
              IF (EOF(5) .NE. 0.0) GO TO 20
              ATOT = ATOT + A
              PR = FGF (V(I),D(I),S(I)
              P(I) = PR * A
              IF (PR .GT. PMX) PMX = PR
              GO TO 10
        C
         20   N = I - 1
        C
              PRINT*, "ENTER NUMBER OF FISH DESIRED",
              READ*,NFSH
        C
        C----------MAIN LOOP
         30   ITCNT = ITCNT + 1
              IF (ITCNT .GE. MXIT) GO TO 200
              FCHECK = PMX * RANF (NRAND)
              I = N * RANF(NRAND)
        C---------SELECT RANDOM POINT
              IF (FCHECK .GT. P(I) / ATOT) GO TO 30
```

```
            IFSH = IFSH + 1
            IPTN(I) = IPNT(I) + 1
            WRITE (6,6000) I, 1.0, D(I), V(I), S(I)
            IF (IFSH .LT. NFSH) GO TO 30
C
C
  200       PRINT*, "TOTAL NUMBER OF ITERATIONS = ,ITCNT
            PRINT *, "TOTAL NUMBER OF FISH = ",IFSH
            PRINT*, " "
            DO 222 I=1,N
  222       IF(IPNT(I) .NE. 0)PRINT*,I," ",IPNT(I)
C
  5000      FORMAT (4F8.3)
  6000      FORMAT (I8,F8.1,3F8.3)
            END
```

INDEX

50272-101

REPORT DOCUMENTATION PAGE

1. REPORT NO. Biological Report 86(7)

2.

3. Recipient's Accession No.

4. Title and Subtitle
Development and Evaluation of Habitat Suitability Criteria for Use in the Instream Flow Incremental Methodology
Instream Flow Information Paper No. 21

5. Report Date
September 1986

6.

7. Author(s)
Ken D. Bovee

8. Performing Organization Rept. No.

9. Performing Organization Name and Address
National Ecology Center
Division of Wildlife & Contaminant Research
U.S. Fish and Wildlife Service
2627 Redwing Road
Fort Collins, CO 80526-2899

10. Project/Task/Work Unit No.

11. Contract(C) or Grant(G) No.
(C)
(G)

12. Sponsoring Organization Name and Address
National Ecology Center
Division of Wildlife & Contaminant Research
U.S. Fish and Wildlife Service
U.S. Department of the Interior
Washington, DC 20240

13. Type of Report & Period Covered

14.

15. Supplementary Notes

16. Abstract (Limit: 200 words)

Accurate and comprehensive suitability criteria are critical to the effective use of the Instream Flow Incremental Methodology. Five major topic areas relating to the development and evaluation of microhabitat suitability criteria are discussed in this paper:
(1) study planning and design,
(2) development of criteria by professional judgment and consensus,
(3) field methods for the acquisition of microhabitat utilization data,
(4) statistical methods for fitting data to curves or mathematical functions, and
(5) methods for evaluating criteria accuracy and transferability.
The discussion of each technique includes a brief summary of the advantages, limitations, and potential sources of error and bias. The paper also provides a foundation for the development of regionalized microhabitat suitability criteria by a strategy of complementary study planning and mathematical convergence.

17. Document Analysis a. Descriptors
Habitability
Mathematical models
Fishes
Water resources
Stream flow

b. Identifiers/Open-Ended Terms
Habitat suitability criteria

c. COSATI Field/Group

18. Availability Statement
Release unlimited

19. Security Class (This Report)
Unclassified

20. Security Class (This Page)
Unclassified

21. No. of Pages
235

22. Price

(See ANSI–Z39.18) See Instructions on Reverse OPTIONAL FORM 272 (4–77) (Formerly NTIS–35) Department of Commerce

*U.S. GOVERNMENT PRINTING OFFICE: 1986-0-672-848/45301

Take Pride in America

Use Natural Resources Wisely

DEPARTMENT OF THE INTERIOR
U.S. FISH AND WILDLIFE SERVICE

As the Nation's principal conservation agency, the Department of the Interior has responsibility for most of our nationally owned public lands and natural resources. This includes fostering the wisest use of our land and water resources, protecting our fish and wildlife, preserving the environmental and cultural values of our national parks and historical places, and providing for the enjoyment of life through outdoor recreation. The Department assesses our energy and mineral resources and works to assure that their development is in the best interests of all our people. The Department also has a major responsibility for American Indian reservation communities and for people who live in island territories under U.S. administration.