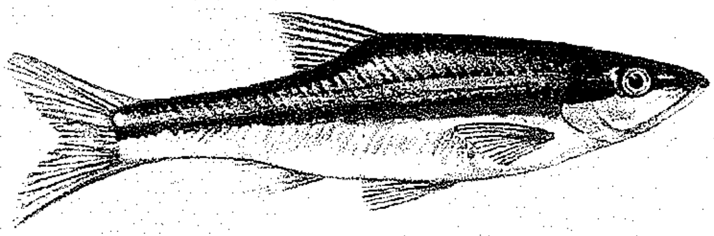
Guidelines for Designing and Interpreting Stream Surveys in Ontario



A compendium manual to the Ontario Stream Assessment Protocol

Version: 1.1

2003

Table of Contents

Introduction[1]

The materials provided within this manual are intended to complement the procedures described in the Ontario Stream Assessment Protocol by providing advise on study design and interpretation of the results. This manual should not be considered as being all that is required. In fact it is simply a holding place for some of the information that project managers might find useful in the challenges of project management. This compendium manual will evolve considerably over time as new interpretation techniques and models develop and as we develop more experience in using stream data to answer complex questions. We advise users to also treat this as a holding place for information that helps you with your job.

1.0 ORGANIZATION

As with OSAP the manual is organised into sections and modules within each section. This enables new material to be added to the manual without significant disruption to the overall organisation. The current contents are as follows:

Section 1: Designing a stream survey

This section includes several reports that were written to provide guidance to project managers in defining study questions and approaches to be applied to answer these questions (module 1). Module 2 offers a stepped process for developing a study design to answer specific questions. Field forms are provided to assist project managers and this data can (and should) be stored in the HabProgs application and database (see section 5 of OSAP). Recent work in Ontario has focused on utilizing landscape data to stratify field surveys, The third module describes a tool that was developed for delineating, attributing and classifying rivers into segments. This protocol has been developed into an application (Aquatic Landscape Inventory System) and it has been applied to large parts of the Great Lakes basin by the geomatics section of OMNR (Peterborough, contact Scott Christilaw).

Section 2: Interpreting Water Quality Evaluations

The initial focus for this section is to provide guidance on the interpretation of benthic invertebrate data (modules 1 and 4) and water temperature data (modules 2 and 3). All but the fourth module provides interpretation advice on rapid assessment techniques. The fourth is a new manual being developed for a new initiative of the Ministry of the Environment and describes both study design and interpretation guidelines for data collected using several protocols and varying levels of effort/diagnostic power.

Section 3: Fish Community Sampling

The current list of modules in this section focus mainly on articulating the influence of varying effort of field surveys on catches. The first module describes how data collected using single pass surveys (with varying effort) are adjusted to reflect 3 pass population estimates. The second module is a copy of a manuscript by Jones and Stockwell that

[1] Reference this document as: Stanfield, L. W. (Ed.). 2003. Guidelines for Designing and interpreting stream surveys: A compendium manual to the Ontario Stream Assessment Protocol. Ontario Ministry of Natural Resources, Aquatic Research and Development Section, Picton. internal document.

describes the original research that supports the rationale for using single pass electrofishing surveys.

Section 4: Physical Habitat and Geomorphology

This section describes the rationale for and basis of the queries that are available from within the HabProgs application. Further, users are advised how to access the data and the summary reports from within the application.

Section 5: Models generated from OSAP data

A number of researchers are currently working to develop models to relate various measures of habitat and biota currently found within HabProgs. As these models are generated and where appropriate, the algorithms will be included in the application. To date models exist that relate the physical attributes of channel structure, temperature, Hilsenhoff index and geomorphology to fish abundance. These models are summarized in the first module and are available from within HabProgs.

Section 1: HabProgs database structure and Management Protocols

To date only HabProgs is available for storing and interpreting the habitat data collected with OSAP. The database generates many tables and in this section we will periodically update the list provided to enable users to easily locate data from within the application.

Table of Contents for Companion manual:
Guidelines for Designing and interpreting Stream Surveys

Sections:

Section 1:

Designing a stream survey

Section 1 Module 1

A framework for incorporating the science of impact, risk and state of resource assessments into management decisions for flowing waters[1]

Table of Contents

List of Appendices

List of Tables

[1] Authors: B. Kilgour and L. W. Stanfield

List of Figures

1.0 INTRODUCTION

A number of federal, provincial, tribal and municipal governments share in the responsibility for managing habitat in the province of Ontario. The Department of Fisheries and Oceans (DFO), has a mandate under the Fisheries Act (Section 35(2)) to ensure that fish habitat is not harmfully altered, disrupted or destroyed (HADD). The Ontario Ministry of Natural Resources (OMNR) has a mandate to manage the aquatic ecosystems and the fisheries they support for long term sustainability. Municipalities and tribal governments manage local landuses under the Planning Act to ensure their compliance with Federal and Provincial Policies and laws. Each agency contributes to habitat management differently.

As part of its mandate, DFO evaluates "projects" submitted to it for approval. Projects may be as simple as a culvert installation on a small intermittent stream, or as complex as a new hydro dam on a large river. DFO examines information about the biophysical condition of the site, as well as information about proposed developments to determine the likelihood that a HADD will result. Where a HADD is likely to occur, DFO will either recommend mitigation or deny the project. In some cases, DFO's decisions require site-specific biological, physical or chemical data to assist in the estimating the likelihood of a HADD. DFO has also recently identified culverts, stream realignments and shoreline hardening as three types of developments for which there are information gaps, that are difficult to screen the environmental risks of, and that thus require science to fill those gaps (Gillespie et al., 2002). DFO, therefore, periodically requires field studies to address the information needs of referrals, either generically (as in culvert assessments) or site-specifically. The studies that DFO requires can vary widely in the attributes measured and statistical design (Gillespie et al., 2002).

Municipalities and Conservation Authorities carry out subwatershed studies that require the characterization of biophysical conditions in stream systems. On the basis of existing or historical conditions, subwatershed studies (and plans) determine the future developments that can occur in a system while minimizing environmental damage. Subwatershed studies often rely on historical data (aquatic resource inventories), but can incorporate new data from field studies.

Conservation Authorities also carry out routine biomonitoring that may be part of larger environmental monitoring programs (e.g., CVC, TRCA). The design of those monitoring programs is challenging because of fiscal restraints and an underlying desire to sample as much as possible in as many places as possible. Sampling is, however, limited to specific biophysical or chemical variables at specific key locations. Though monitoring is carried out annually, sampling locations may be visited only once in every three to five years.

Finally, industry and Municipalities are often required to carry out site-specific monitoring to address permit requirements or certificates of approval to operate. Monitoring is usually intended to determine the effects of point-source discharges on aquatic receiving environment biology, chemistry or physical attributes. Pulp mill effluents, mines and municipal wastewaters are example discharges that may require monitoring. Monitoring requirements often dictate rigorous study designs to document effects, including high levels of statistical replication and the characterization of attributes (e.g., sentinel fish populations) that may require considerable effort in the field.

Each of the above examples are assessments. Those that characterize existing conditions can be termed state of the resource assessments are may or may not require new data. Studies that examine the effects of existing developments can be called impact assessments, while those that predict future impacts are typically called risk assessments or environmental assessments. Though the goals and objectives of each of these types of assessment differ, they should all follow a generalized thought process beginning with problem formulation, followed by identification and characterization of the stressor being evaluated (or not), definition of physical and temporal boundaries on the assessment, existing information assessment, and design of field studies if required. Where field studies are carried out, the challenge is in selecting measurement variables, protocols for measurement, timing, locations, etc., i.e., all those factors that comprise the study design. Those designing the studies need assistance to determine if and when data need to be collected; if so, what type, where, when, and how intensively.

The principal objective of this document is to provide a generalized but flexible framework for assessing the biophysical condition of stream systems. The framework will provide general guidance on how to determine when new field studies are required, the design of those studies, and follow-up actions. Further, this framework formalizes a flexible process for developing study designs that satisfy individual objectives. It is anticipated that this document will be a useful frame of reference for DFO Habitat managers, referral biologists, and Conservation Authority biologists, among others, in assessing existing conditions, or to justify studies that are designed to predict or document effects associated with development.

This document describes a generic framework that summarizes a thought process for carrying out assessments and designing studies. The underlying principles of the document follow from approaches used in site-specific risk assessments (EPA, 1993; CCME, 1996; MOE, 1996). While the original purpose for this document is to provide assistance to those working on streams, the concepts and process are applicable to any environmental question. The document cannot stand alone, and does require companion documents that (1) describe individual measurement protocols (e.g., Ontario Stream Assessment Protocol), (2) provide guidance on statistical aspects of study design (e.g., Gillespie et al., 2002), and (3) recommend measurement precision.

Apart from this introduction, this document has two main sections. Section 2.0 overviews the framework and provides general background. Section 3.0 provides examples of how the framework might be applied.

2.0 FRAMEWORK

In support of a workshop on riverine management held in 1998, Boyd et. al. (2000) developed a management framework for ecosystems (Figure 1). In this framework, overall ecosystem goals and objectives are specified on the basis of human values. Where development activities threaten those values, issues emerge. The political environment determines which issues become priority and require management. Where there is sufficient knowledge on which to immediately base a management decision, management can be carried out. The success of a management action may or may not be monitored over time (as part of Adaptive Experimental Management) to feed back to Issue Analysis and provide additional knowledge against which to propose management actions for future issues. Where there is not enough knowledge on which to propose a management action, the framework anticipates that science will be carried out to address the knowledge gap. Where science addresses the gaps, that new knowledge can be used to inform the Issue Analysis stage.

The 1998 workshop confirmed the need for an overall framework that would guide managers through the science analysis (McGuinness et. al., 2000). The assessment framework presented here (Figure 2) is intended to address that gap and to assist in the development of study designs. This framework broadens the Issues Analysis Stage into six discrete steps, the Science Analysis Stage into two discrete steps and Management into two discrete steps. As such, the framework described here provides further guidance towards the issue assessment.

The proposed framework (Figure 2) for carrying out stream assessments has the following 11 fundamental steps:

1. **Question Definition:** Development of a clear question that the assessment is intended to answer. The construction and interpretation of Hypothesis of Effect diagrams or their equivalent should aid this step.

2. **Assessment Typing**: Clear articulation of the type of assessment that is intended. That is, is the study intended to characterize existing conditions as part of a state of the resource assessment, predict future environmental effects (i.e., risk assessment) or evaluate the present-day effects of specific stressors (i.e., impact assessment).

3. **Stressor characterization:** Description of the development pressure, proposed project, chemical stressor, etc., that is under evaluation.

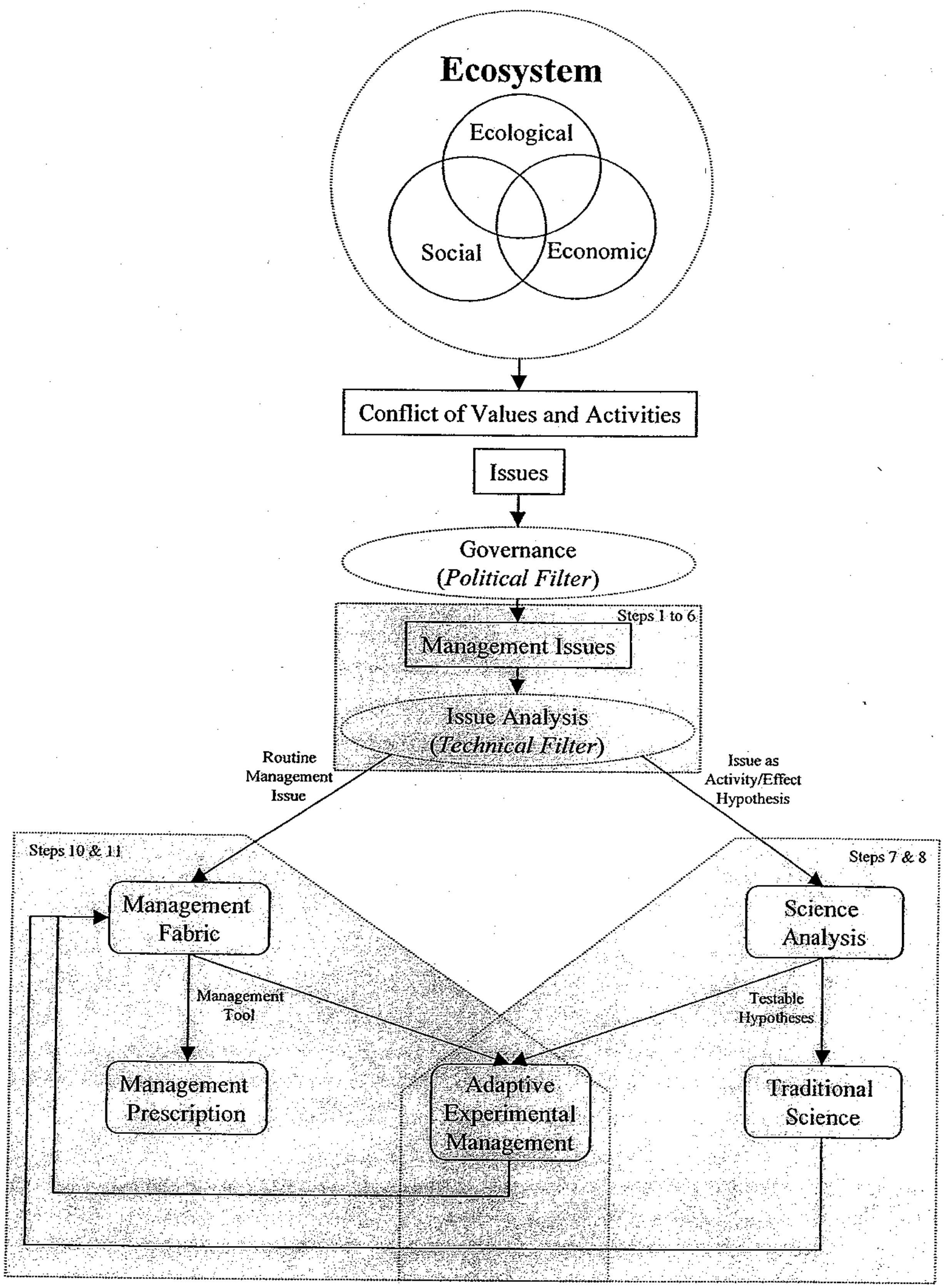


Figure 1. Management framework for stream systems (from Boyd et al., 1998).

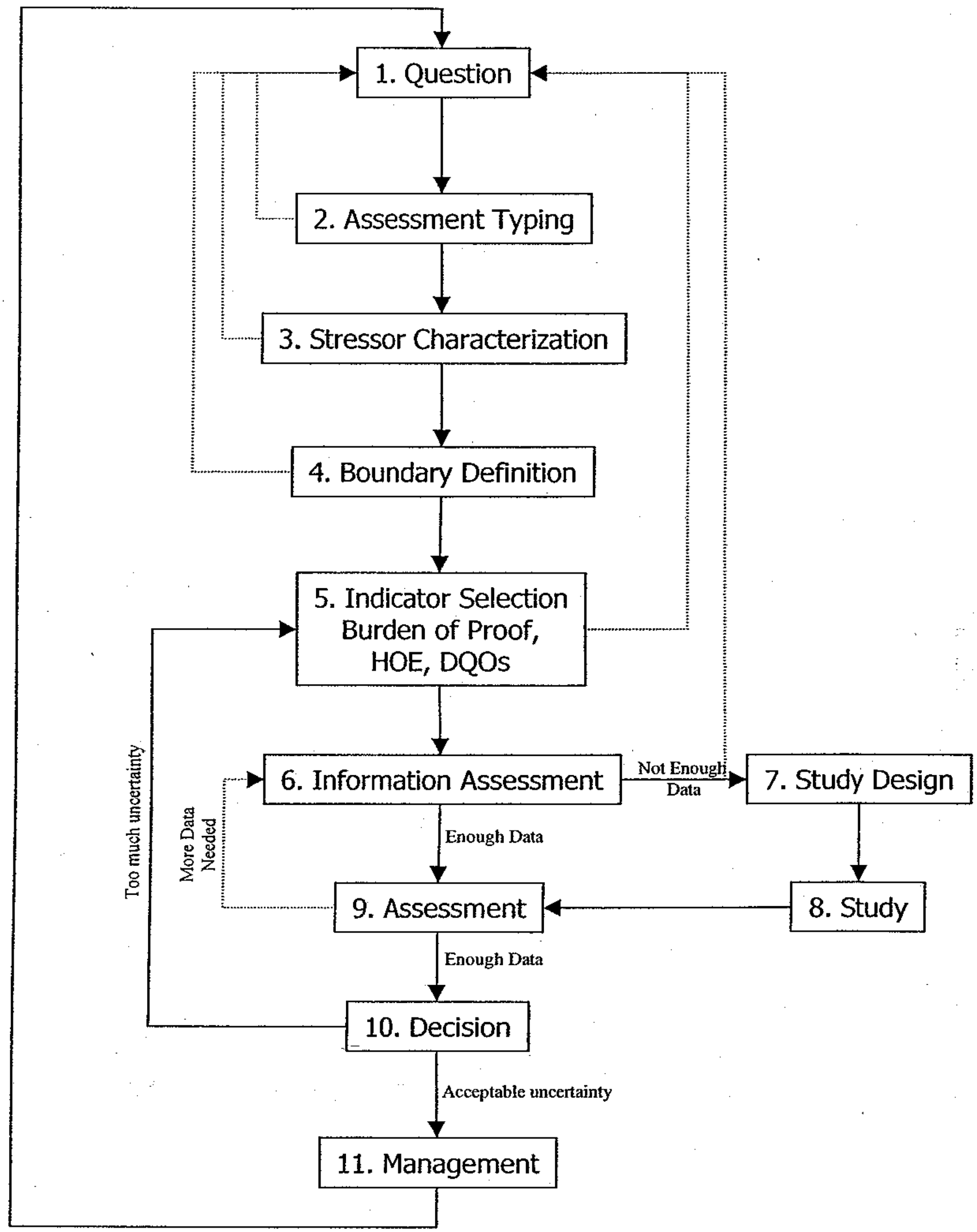


Figure 2. Schematic illustration of the process for carrying out a stream assessment.

4. **Boundary definition**: Description of the spatial and temporal bounds of the assessment.

5. **Indicator selection**: Justified selection of environmental indicators that will be used in the assessment.

6. **Information Assessment**: This step determines whether the existing available information and knowledge are sufficient for making the assessment.

7. **Study Design (if necessary)**: This step details the attributes of fields studies. Measurement variables and protocols, timing, locations are incorporated into a studies' design.

8. **Study (if necessary)**: Conduct of the field study.

9. **Assessment**: Carry out the assessment with existing or new field data.

10. **Decision**: On the basis of the assessment, decide a course of follow-up action. Regardless of the nature of the assessment, each assessment should lead to a decision regarding follow up activity.

11. **Management Action**: Carry out the recommended management action

Each of these steps is described fully below. Management Issues or questions "trigger" the use of the framework. The framework then guides the assessment to answer the question either through the use of existing information and knowledge, or through the development of new knowledge obtained through studies specifically designed to address knowledge gaps. This framework builds upon some of the basic "risk" assessment frameworks developed in Canada (CCME, 1996; OMOE, 1996; FCEMS, 2000), the US (USEPA, 1989), and elsewhere (e.g., Denmark, Pedersen et al., 1995). The process of evaluating effects is intended to be logical, and to initially incorporate a screening-level exercise that may build to larger, more comprehensive assessments where there are more gaps in the understanding of the relationship between the development pressure (stressor) and valued environmental components.

2.1 Level of Detail

Following from CCME (1996), a tiered approach to assessment is recommended. Here, it is recommended that screening and comprehensive assessments be carried out iteratively. If an initial screening-level assessment cannot adequately address the question posed with an acceptable degree of uncertainty, an additional level of detail is required for the assessment (see loop from Step 10 to Step 5, Figure 2). Thus, assessments may be carried out iteratively until uncertainties are at an acceptable level. The impetus for carrying out additional iterations is the uncertainty associated with a decision (Stage 10). Sources of uncertainty include:

- normal variation in environmental conditions;
- imperfect or incomplete knowledge; and,
- human error in carrying out the assessment.

CCME (1996) recommends that the uncertainty considered acceptable for terminating an assessment must be determined by the individual assessor, while financial and regulatory considerations may also come into play. According to CCME (1996), we should also recognize that there is no single best approach to an assessment, and that multiple lines of evidence are better than single approaches.

2.2 Step 1: Question Definition

The most important task in any assessment is articulating the question that the assessment is to address. It is recommended that the underlying objective of the assessment be articulated in a few sentences. See Appendix 1 for example questions. Questions need not be detailed in this step since Steps 2 through 5 will increase the detail of the overall question. A well articulated question will assist in defining the information requirements of the next 4 steps. When carrying out Steps 2 through 5, it may become evident that the initial question did not adequately capture the underlying objectives, or was not a question that could be feasibly answered and should be revised. Thus the feedback loops from Steps 2 and 5 to Step 1.

2.3 Step 2: Assessment Typing

The framework is considered appropriate for a variety of assessment types including:

1. **Impact**;
2. **State of the Resource**; and,
3. **Risk Assessment**.

Project objectives often combine aspects of each of these assessments. In these cases, assessments should be carried out for each specific question, with the results combined at the end where warranted. Assessments with combined objectives run the risk of compromising one or all of the multiple objectives, though they can be done with careful planning (Section 3.4).

Definitions for each of these categories vary. Here, impact assessments are considered studies carried out after an activity (development or natural event) has occurred with the objective of determining the nature of the effect on biophysical conditions. These impact assessments are also called environmental effects monitoring (EEM) studies (e.g., Environment Canada, 1998). Impact assessments are generally site-specific and document the effects of a particular development on environmental characteristics. See Environment Canada (1998, 2001) and Gillespie et al. (2002) and references therein for considerable guidance on the design of these types of studies for aquatic environments. These studies generally compare biological responses in "affected" locations with biological responses from "unaffected" locations. Differences in response between affected and unaffected locations are used as evidence of effects. Guidance on optimal and practical statistical designs is provided in Gillespie et al. (2002)

Conservation Authorities, among others, often conduct studies that are designed to characterize existing biophysical conditions, and may or may not test for spatial or temporal trends. Here, these types of studies are termed state-of-the-resource assessments. Data from state-of-the-resource assessments can be used as baseline studies against which to judge future development. In those cases, state-of-the-resource assessments become the initial stages of comprehensive impact assessments. If there is no stressor[1] or modifier (see definition below), the assessment is a state-of-the-resource assessment. Stressors should not be confused with strata[2], which are incorporated into studies to account for natural sources of variation in biophysical responses.

Classically, risk assessments are defined as those studies that evaluate the likelihood that adverse effects will occur, are occurring, or have occurred as the result of development (U.S. EPA, 1992). Risk assessments are typically carried out to determine the likelihood that chemical contaminants will have adverse effects on terrestrial or aquatic animals (CCME, 1996). In the context of fish habitat assessment, however, risk assessments represent evaluations of the likelihood that habitat alterations will affect the productive

[1] *A stressor is an activity (natural or anthropogenic) that is predicted to have an effect on a valued ecosystem component.*
[2] *Strata are natural features such as catchment area, slope, and surficial geology that cause variation in biological condition. From a statistical point of view, they can be used to stratify study designs and account for natural variation in biological endpoints.*

potential of fish and/or fish habitat (i.e., cause HADD). There is considerable overlap in the classic definition of risk assessment and the definition of impact assessment presented here. Where there is a new development, risk assessment will rely on existing data from other facilities to derive the burden of proof that an environmental impact (HADD) may result. Where the development is existing, the risk assessment will require at least some site-specific data to characterize physical and chemical condition, which would be used to estimate the likelihood of a biological impact. Where biological responses are measured as part of a site-specific risk assessment, the assessment becomes an impact assessment as described above.

2.4 Step 3: Stressor Characterization

The stressor characterization stage should identify the major stressors that have potential to affect selected responses (Step 5). Variables that reflect the degree of human activity are modifiers, because they reflect the degree of modification imposed on the natural biophysical environment (Figure 3). Modifiers generally refer to the human activity or development (e.g., mill, mine, urban area, and culvert). It may be possible to quantify the modifier, for example the area of a catchment that is under urban area.

2.5 Step 4: Boundary Definition

In this step, the physical/spatial and temporal bounds for the assessment are identified. Spatial bounds require consideration of the physical extent of potential effects and/or the scale at which effects might be manifest. Boundary definition includes an understanding of the primary landscape variables that might influence the assessment endpoint (Figure 3).

Primary variables are those characteristics of the landscape that provide the underlying foundation for biophysical conditions. Primary variables are also difficult to modify by human activity. Kilgour and Stanfield (2001) listed the following general features that can be considered primary variables, and that are particularly important to the distributions of fish:

- **Upstream drainage area**, because it reflects/determines flows and discharge, and is a strong determinant for fish community composition and species richness.
- **Position** which is a measure of how close to a large river a stream site is. Being in close proximity to a large river influences species richness.
- **Connectivity** which reflects connections of a stream site to other types of water bodies including inland lakes, great lakes, salt water and wetlands.

- **Bedrock geology** because it determines deep acquifer flow patterns and basic water chemistry.

- **Surficial geology** because it determines soil permeability and is associated with the likelihood of there being significant groundwater resources.

- **Slope** because it is associated with flow velocities, substrate particle size, and thus the kinds of animals found at a site.

- **Climate**

Primary variables represent the template against which natural variations in biological responses (compliance indicators) are often set (Kilgour and Stanfield, 2001). Thus, their potential influence on biological responses should be considered in any assessment. For example, biological responses in stream systems vary greatly with catchment area. Small streams with small upstream catchments tend to be cooler with a limited diversity of cold-water fauna. In contrast, larger streams/rivers with larger upstream catchment areas tend to be warmer with larger diversity of fauna (Vannote et al., 1987). Catchment area may be a very important factor to consider in the design for state-of-the-resource assessments covering catchments or watersheds, but might not be an important factor for the assessment of a point-source discharge on a large river. Where these factors are expected to be important, they would be used to stratify study designs.

2.6 Step 5: Indicator Selection

This step involves the selection of indicators on which to make the assessment. The assessment endpoints should be selected on the basis of Hypothesis of Effects models, and the approach that will be taken to develop the burden-of-proof in the assessment. Data quality objectives determine the protocols under which data should have been or should be collected. In selecting indicators, it is assumed that managers have already established valued ecosystem components (VECs) through other activities (Figure 1). In classical risk assessment frameworks, identification of VECs is a critical component of the overall assessment. The indicator should provide a direct measure of the VEC, or if surrogates are selected, the managers must ensure that the relationship between the VEC and the surrogate are well understood. The best indicator should be capable of being measured to the same degree of precision necessary to answer the question emerging from the issue analysis.

In stream assessments, several possible indicators are possible. At the site level, Cairns et al. (1993) recognized three general kinds of indicators: (1) compliance (i.e., VECs); (2) surrogate; and (3) diagnostic. Early-warning indicators can be a fourth type of indicator (Kilgour and Barton, 1999). Each of these kinds of indicators can be used to address one of four general questions relating to ecosystem management (Cairns et al., 1993): (1) are stated ecosystem objectives being met, i.e., are the VECs in an acceptable condition? (2) if VECs are not in an acceptable condition, why not? (3) how can impending impacts on VECs be predicted before they are actually observed? (4) are there indications that VECs may be in jeopardy of being impacted in the future.

The first part of this section describes some types of indicators that are typically used in environmental assessments, and their uses and limitations (Section 2.6.1). Then, the importance of defining criteria to be used for decision making and the level of precision required for a state-of-the-environment assessment is described (Section 2.6.2). The third part of this section introduces the use of hypothesis of effect diagrams (Section 2.6.3), while burden-of-proof approaches are described in Section 2.6.4. Both are integral in risk and impact assessments.

2.6.1 Indicator Types

2.6.1.1 Compliance Indicators

As defined by Cairns et al. (1993), compliance indicators are those environmental attributes (like valued ecosystem components, VECs) that we are trying to protect, enhance, and otherwise prevent impacts to. Thus, if we know what it is we are protecting, defining the compliance indicator for any assessment should be relatively easy. That is rarely the case. Cairns et al. (1993) suggest that compliance indicators (VECs) should be the most obvious part of any monitoring effort, and thus their significance should be communicable to the public and policy makers. Compliance indicators should have biological and social significance, be measurable with standard approaches, be interpretable, and have historical data against which to judge trends and be measurable at scales that are relevant to management issues. Compliance indicators are typically aspects of ecosystem structure or composition (Figure 3) that are chosen because they integrate conditions within the ecosystem. Because they are integrators, they can take longer to respond to changes in chemical and physical habitat conditions. Thus the detection of effects on compliance endpoints is often a signal that the whole ecosystem has been altered and that the effects may be irreversible. For routine monitoring, compliance indicators are not preferred because they do not detect effects until after it may be too late to manage the problem.

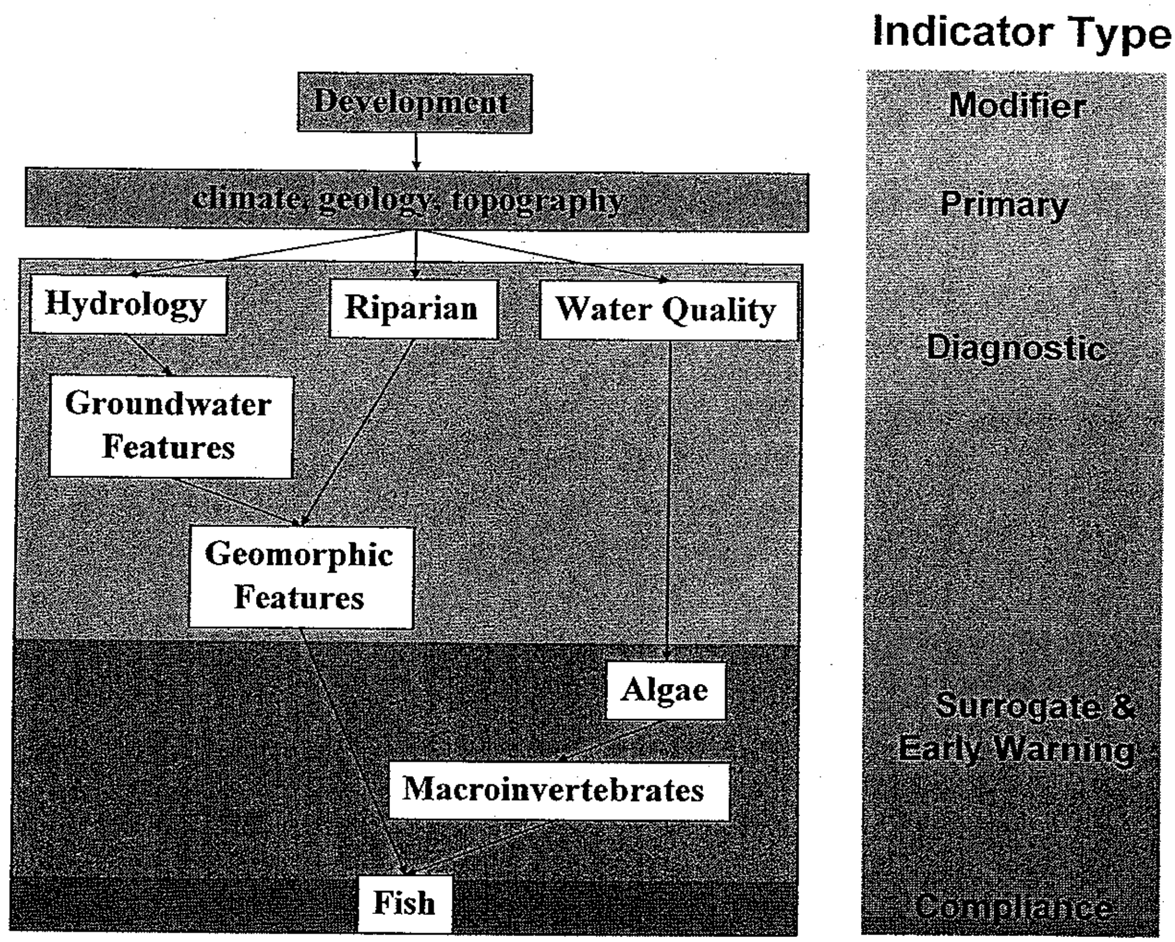


Figure 3. Schematic showing the relationship between a hypothesized effect and measurable assessment endpoints.

The Fisheries Act has the underlying objective of having *no net loss of fish and/or fish habitat*. In the context of that Act, a logical compliance indicator would be some measure of actual or potential fish production. Not only is production difficult to measure (quantify), but it also does not tend to respond to environmental stress as rapidly as do changes in species assemblages (Schindler, 1987). Thus, though the Act may imply one or more measures worth assessing, common sense should also be used when making the final selection. In Canada, there is no common VEC used in the assessment of stream resources, but species assemblages and the biomass of key species have been put forward in numerous cases.

2.6.1.2 Surrogate Indicators

Surrogate indicators are those indicators that are measured, not because they have a direct ecosystem value, but because they actually or theoretically reflect the condition of a compliance indicator (VEC). Some reasons for incorporating surrogates into an assessment include:

- they are easier to sample in a quantitative fashion;
- when sampling of the compliance indicator would be too destructive;
- when there is a defined relationship with the compliance indicator;
- when the surrogate indicator is measurable at a more appropriate temporal or spatial scale; and,
- when the surrogate indicator has lower natural variability and can be measured more precisely
- when data for the surrogate indicator can be produced in a more timely fashion.

In Canada, the Pulp and Paper Effluent Regulations (PPER) require proponents (mills) to monitor the condition of sentinel fish species (typically white sucker in freshwater environments) and benthic macroinvertebrates. Sentinel-species surveys are proposed because pulp mills discharge into large receiving environments where it is difficult to characterize effects on a fish community or assemblage. The sentinel-species surveys are thus a surrogate for potential effects on fish communities. Similarly, benthic community surveys are conducted, not because they are considered a VEC (though some agencies treat them as such), but because they are considered a surrogate measure of potential effects on fish habitat (Environment Canada, 1998). Kilgour and Barton (1998) have demonstrated that the composition of benthic invertebrate communities can to some degree predict the composition of fish communities. In part, that demonstration supports the idea that benthic communities could be used as surrogate indicators of potential effects on fish communities.

2.6.1.3 Early-Warning Indicators

Early warning indicators are used to detect trends in environmental conditions that may affect the compliance indicator at some point in the future (Cairns et al., 1993). They should have a well-defined relationship with the compliance indicator, have standardized measurement protocols that are cost effective and result in defensible data, and should be able to integrate effects of various kinds of stressors in a manner that is similar to the compliance indicator. Sampling should be non-destructive to the compliance indicator, and there should be potential for continuity in measurement over time so that trends can be detected. Finally, it should be possible to obtain the data relatively quickly without time lags so that effects, if present, can be managed in a timely fashion.

In aquatic environments, early-warning indicators have included the following kinds of indicators:

- behavioural;
- physiological; and,
- morphological.

Early-warning indicators are generally biological because VECs are typically biological endpoints, and they are a direct measure of a VECs initial responses to stress. Indicators of physical and chemical habitat features are not considered early-warning indicators because changes in the chemical or physical makeup of a stream do not guarantee that biological effects will occur.

2.6.1.4 Diagnostic Indicators (Drivers)

Diagnostic indicators are those indicators that provide insight into the cause of non-compliance. Diagnostic indicators should have known linkages to the biological attributes of interest (i.e., compliance indicator or VEC), and may be identified through hypotheses-of-effects models (Section 2.6.3). As with other indicators, there should be standardized measurement protocols, with the potential to make measurements at appropriate spatial and temporal scales. Physical and chemical habitat descriptors measured in the field can also be thought of as *drivers* because they are the fundamental features that drive biological responses (Figure 3).

2.6.2 Criteria for Assessment Endpoints

Regardless of the assessment endpoint, criteria must be established *a priori* to define acceptable and unacceptable conditions for the compliance, surrogate, early-warning and diagnostic indicators. Criteria for compliance or VEC endpoints crystallize the anticipated or expected biological condition. Where the VEC has a condition other than the expected condition, one should conclude that an undesirable impact or HADD has occurred. Criteria for surrogate, early-warning and diagnostic indicators are values that correspond with unacceptable effects on the compliance indicator (Figure 4). Thus, when surrogate, early-warning or diagnostic indicators have conditions outside of the establish criteria, there is good evidence of existing or potential future impacts to the VEC (which may also be considered a HADD).

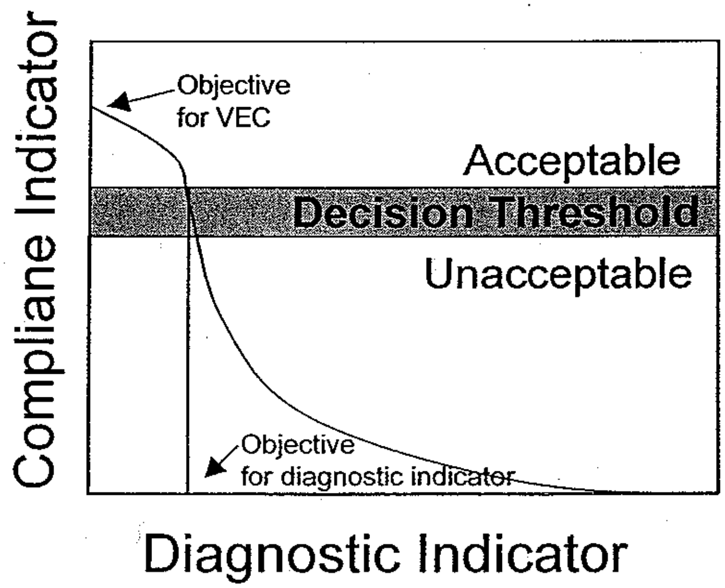


Figure 4. Possible relationship between a compliance indicator (VEC) and a diagnostic indicator showing potential criteria/objectives for both.

Criteria are also required in order to specify data quality standards. In screening-level assessments and depending on the assessment endpoint, criteria may already be established in published literature. Generic water and sediment quality objectives (CCME, 1999) define critical concentrations above or below which there may be a risk of biological impairment. In more comprehensive assessments, criteria may have to be developed through field programs that characterize acceptable natural background conditions, and or through consultative processes (Environment Canada, 2001). Criteria may be as simple as "*brook trout should be present*", or as complex as "*indices of fish community composition shall not exceed natural background conditions for these specific kinds of habitats, where natural background is defined as the mean plus 2 standard deviations*".

The intended uses of data (e.g., hypotheses to be tested, summary statistics involved and total uncertainty that can be tolerated) influence the required quality of the data. Data quality objectives should specify sensitivity, accuracy and precision. These characteristics will in turn affect the selection of specific assessment endpoints. Take for example a proposed impact assessment with benthic invertebrates as one assessment endpoint. If the criterion for declaring an undue impact was a loss of a major taxonomic group such as mayflies, stoneflies, gastropods, etc., then coarse taxonomy might be all that would be required. Generally, the criteria against which judgements are to made often dictate the quality of data, and thus the measurement protocols.

These two important steps are combined in an iterative process to define the assessment endpoints. An integral step in this process is an evaluation of existing data that could be used to answer the question. Note that different data sets are likely to be used depending on whether the assessment is a screening or comprehensive assessment. From the above example, a manager may choose to simply use numbers of water quality infractions in an area for the screening tool. For the more comprehensive assessment they may choose the rapid bioassessment methods of the kick and sweep technique would be appropriate to answer this question (Module 3, Stanfield et. al., 2000).

2.6.3 Hypothesis of Effects Models

It is highly recommended that hypothesis-of-effect (HOE) models be developed as part of the early stages of an assessment to clarify the pathways through which a stressor will interact with the proposed VECs. These models identify modes of action of stressors on assessment endpoints. HOE models should be structured to clearly illustrate the current understanding of how management or development activities are linked to VECs through important physical, chemical, biological, economic and social processes (Greig et al., 1992; Gillespie et al., 2002).

On the basis that VECs have been identified, that the stressor has been well characterized, and the spatial and temporal boundaries have been established, HOE models effectively summarize in a visual format, the expected mode of action of the stressor. By doing so, HOE models clearly identify measurable indicators.

2.6.4 Burden of Proof

On the basis of HOE diagrams, an approach to a burden-of-proof as to whether an impact is likely (i.e., risk assessment) or whether an impact can be attributed to a specific stressor can be developed. For example, the burden-of-proof that mining effluents cause effects on aquatic organisms is derived from demonstrating concordance between environmental chemistry, bioavailability (i.e., elevated tissue contaminant concentrations or toxicity) and biological effects (Chapman, 1991; Green et al., 1993; ESG International, 1998). The specific burden-of-evidence will depend on the question, but will be key to selecting indicators.

2.7 Step 6: Information Assessment

On the basis of the hypotheses to be tested, the burden-of-proof approach taken, and the data quality objectives, the data required to carry out the assessment should be apparent. If the data are available, the assessment can be carried out. If the data are not available, the data need to be developed (Steps 7 and 8). For impact assessments, site-specific data that are up to date are usually required. Where the data are outdated or were collected using inappropriate methods, or did not demonstrate a cause-effect relationship, more data may be required. At the screening stage, risk assessments usually rely on data from other sites or studies. If the screening-level assessment does not produce a decision with a high degree of confidence, more data may be required. In risk assessments, the requirement to have more data usually infers that more site-specific data are necessary to take into account site-specific factors. For state-of-the-resource studies, there are an increasing number of datasets (or data layers) available that may enable a screening-level assessment to be carried out or with minimal effort conduct more rigorous assessments.

2.8 Steps 7 and 8: Study Design and Study

Where there are inadequate existing data to carry out an assessment with sufficient confidence, it is necessary to obtain more data that will produce a more confident assessment. The specific study design employed will vary with the hypothesis, and resources available. Study designs generally specify the following attributes of a study:

- statistical design;
- study area boundaries (temporal and spatial); and,
- sampling methods

Study area boundaries should be evident from Step 4 (Boundary Definition). Common and powerful statistical designs are well documented elsewhere (e.g., Gillespie et al., 2001, and references therein). Impact and risk assessments often incorporate elements of what are referred to as Before-After, Control-Impact (BACI) designs that have been well described by Green (1979), Underwood (1991, 1993, 1994) and others. BACI designs compare data from impaired (exposed) sites to data from reference or control sites. Recognizing that two locations can naturally differ in biological responses at any given time, Green (1979) and others have recommended that the principal evidence of an effect is a difference in changes from before to after development, between control and impacted sites.

Specific sampling methodologies need to be identified (Appendix 1) and defended. Typically, impact and risk assessments rely on the most rigorous methods available. For example, the environmental effects monitoring program for the pulp and paper sector incorporates benthic invertebrate surveys with identifications to the lowest practical level, principally genus and species (Environment Canada, 1998). The mining EEM program is proposing family-level identifications on the basis that they are as statistically rigorous, but much less costly to produce (Environment Canada, 2001).

Finally, requiring more data does not necessarily imply that more field work is required. Existing data sets often contain data that are useful to an assessment. Examples include:

- Ontario Ministry of Natural Resources (OMNR) FISHNET database that is used to store fisheries-related data;
- OMNRs NRVIS databases that is used to store natural resource information such as topography, forest cover, wetlands, and fish and wildlife habitats;
- OMNR/ROM inventory database;
- OMNR's HabProgs which warehouses data collected under the Ontario Stream Assessment Protocol (Stanfield et al., 2001); and,
- Environment Canada's BEAST database that warehouses benthic invertebrate and physical habitat data major systems in Ontario.

2.9 Step 9: Assessment

The assessment step has one primary objective, that is to determine whether there is enough information to make a management decision. For state-of-the-resource assessments, it is necessary to know whether enough information is available to characterize environmental conditions. As part of the characterization in a state-of-the-resource assessment, new questions may emerge, forcing the assessment back to Step 1 (Question).

For risk and impact assessment this Step requires:

- Understanding the relationship between the development (stressor) and the assessment endpoint (often termed Hazard Assessment, e.g., CCME, 1996); and,

- On the basis of expected relationships between the development (stressor) and endpoint, predict the extent and nature of impacts; and

- Evaluate the level of uncertainty that the development has had or will have effects (often termed risk characterization, CCME, 1996).

By their nature, screening-level risk assessments can rely on published data and models. Where data and models do not exist for a specific question, new data and models would have to be developed through Steps 7 and 8.

For impact and risk assessments, CCME (1996) describe both qualitative and quotient methods for characterizing the likelihood that effects are associated with a given stressor. Qualitative methods rely on professional judgement of the likelihood of effects in terms of high, moderate, or low (similar to approaches used in most habitat suitability models). Quantitative methods can be used in situations where the stressor is measurable, and are typically applied where concentrations of chemicals are of concern. Where environment concentrations of chemicals exceed objectives, there is a potential risk of biological impacts.

To quantify risk associated with environmental impacts (due to chemicals), CCME (1996) recommend calculating quotients, which are the ratio between the expected environmental condition (concentration), and the chemical objective. Quotients < 1 imply a slight risk with little or no required action (Burns, 1991). Quotients > 1 imply that risk is greater and that management action may be required. Quotients are commonly used in ecotoxicological assessments involving chemicals, but could easily be applied to other stressors for which there are quantitative relationships with relevant biological endpoints (Figure 5). Where there are multiple stressors, individual quotients can be summed to produce an overall hazard index (CCME, 1996).

2.10 Step 10: Decision Matrix

Every assessment should lead to a decision and recommendation regarding management action. The decision matrix (Table 2) should be visited at the end of each assessment iteration, whether it is a screening assessment or a comprehensive assessment. Where the risk that the stressor(s) has or will affect the assessment endpoint is high, and that assessment has a high confidence, then the logical resulting decision should be a recommendation to manage the stressor. If risk is high but confidence is low, more data may be required as part of a more comprehensive assessment. Where there is high confidence of low risk, then the defensible decision may be to recommend maintenance of that level of risk. Clearly, many of these decisions are subjective and require professional judgement.

Table 2. Decision matrix

Confidence	Recommended Action	
	High Risk of Biological Impact or HADD	Low Risk of Biological Impact or HADD
High	Manage or improve existing development. Disallow or re-design new or proposed development.	Maintain existing development. Allow new or proposed development.
Low	Increase comprehensiveness of assessment	Increase comprehensiveness of assessment

With state-of-the-resource assessments, this step should be used to evaluate whether the assessment was sufficient and no additional questions emerge, or whether the assessment creates new questions that initiate a new assessment.

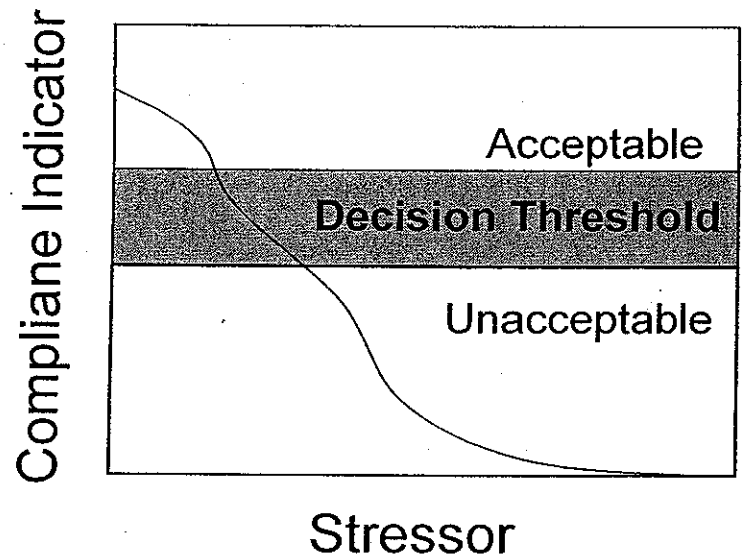


Figure 5. Potential relationship between a stressor and compliance indicator (VEC).

2.11 Step 11: Management Action

All decisions should lead to action following from the decisions in Step 10. Where the burden of evidence indicates existing effects (impact assessment) or potential for effects (risk assessment), some action to mitigate or prevent those effects should be taken.

Management of a stressor may lead to different questions or issues arising. The framework may, therefore, loop back to Step 1 (Question Definition).

Action from state-of-the-resource assessments may include evaluation of new hypothesis or *no action.*

3.0 APPLICATION OF THE FRAMEWORK

In this section, four examples are worked through representing the three assessment types: impact, state of the resource, and risk. A fourth study that combines elements of all three is also provided as an example.

3.1 Impact Assessment

Impact Assessment	
Step	**Approach/Result**
1. Question Definition	Have brook trout declined in the Humber R. watershed with development, since 1970?
2. Assessment Type	Impact
3. Stressor Characterization	Converted lands (lands characterized in non-natural states within the watershed)
4. Boundaries	Spatial: Humber R. watershed Temporal: Assessment should compare the present (2002) condition to an expected pre-1970 condition.
5. Indicator Selection	Compliance Indicator (VEC): brook trout biomass Surrogate Indicator: fish community composition, benthic community Early-Warning Indicator: not applicable Diagnostic Indicator: physical in-stream habitat, temperature Modifier Variable: % impervious area because it summarizes the principle mode of action of urban areas on stream hydraulics and a large component of the development in this watershed is classed as urban. Burden of evidence: The best burden of evidence that development has had an effect on brook trout biomass requires data on brook trout biomass both before and after urbanization in both the Humber R. watershed and a non-urbanized watershed. Those data are not achievable.

Impact Assessment	
Step	**Approach/Result**
	Therefore, a secondary burden of evidence would involve spatial comparisons of developed and less developed streams. Where compliance, surrogate and diagnostic indicators are correlated in their responses to urbanization, there would be good evidence of a development-related effect. For the screening level assessment, data could be taken from other watersheds.
6. Information Assessment	It is perceived that there is adequate information on effects of development on fish communities from other catchments.
7. Study Design	Not necessary at the screening stage.
8. Study	Not necessary at the screening stage.
9. Assessment	Data from the U.S. show the relationship between % imperviousness and impacts on fish communities. Data on imperviousness within the Humber R. catchment indicate that imperviousness is > level that would cause impairment of fish communities.
10. Decision	Conclusion is that urbanization in the Humber River watershed is at a level that would likely cause impairments to fish communities. Further studies may be warranted to quantify the observed effect at which point the assessment proceeds back to step 5 for re-evaluation of indicators, information assessment, and possible study design.
11. Management	This assessment has already concluded that the Humber R. is at high risk to loss of brook trout production as a result of past landuse activities. One management action would be to develop an action plan to reverse the impacts from these stressors. Another management action might be to initiate a study to quantify the effect that landuse has had on this species and to protect remaining habitats. After 1 or more iterations of the assessment process, it will be determined whether development in the Humber R. catchment has significantly affected brook trout biomass. If it has, then mitigation would be recommended. If not, maintenance would be recommended.

3.2 State of the Resource Assessment

State of the Resource Assessment	
Step	**Approach**
1. Question Definition	What is the condition of brook trout in the Humber R. watershed in 2002?
2. Assessment Type	State of the resource
3. Stressor Characterization	none identified
4. Boundaries	Spatial: Humber River watershed Temporal: present day (2002) condition
5. Indicator Selection	Compliance: brook trout biomass Surrogate: not required Early warning: not required Diagnostic: water temperature might be measured because it is so critical to determining the presence/absence of brook trout.
6. Information Assessment	Because there are no present-day data, new data would be required. Study design is then required.
7. Study Design	Study design would potentially stratify for major landscape features including catchment area, slope and surficial geology.
8. Study	Would be carried out.
9. Assessment	Would characterize spatial variations in brook trout biomass and the diagnostic indicator.
10. Decision	The spatial characterization of brook trout biomass may produce new questions of interest to the project manager. Otherwise, the assessment has fulfilled its objective.
11. Management	The management action here would involve the evaluation of new hypotheses, or may require no action at all.

3.3 Risk Assessment

Risk Assessment	
Step	**Approach**
1.	Will brook trout biomass be affected by further development in the Humber R. watershed?
2. Assessment Type	Risk
3. Stressor Characterization	Urbanization
4. Boundaries	Spatial: Humber R. watershed Temporal: Assessment should compare the present (2002) condition to an expected pre-development condition.
5. Indicator Selection	Compliance: brook trout biomass Surrogate: fish community, benthic community Early-warning: not applicable Diagnostic: physical in-stream habitat Modifier Variable: % impervious area because it summarizes the principle mode of action of urban areas on stream hydraulics. Burden of evidence: The best burden of evidence that urbanization could have an effect on brook trout biomass requires data on brook trout biomass both before and after urbanization in both a developed and non-urbanized watershed. Those data are not available. Therefore, a secondary burden of evidence would involve spatial comparisons of urbanized and non-urbanized streams. Where compliance, surrogate and diagnostic indicators are correlated in their responses to urbanization, there would be good evidence of an urbanization-related effect. For the screening level assessment, data could be taken from other watersheds.
6. Information Assessment	It is perceived that there is adequate information

Risk Assessment	
Step	**Approach**
	on effects of urbanization on fish communities from other catchments to carry out a screening-level assessment.
7. Study Design	Not necessary at the screening stage.
8. Study	Not necessary at the screening stage
9. Assessment	Data from the U.S. show the relationship between % imperviousness and impacts on fish communities. Our intention is to develop models that could relate the percent of imperviousness in Lake Ontario tributaries with species biomass. Once these are developed we could use the relationship to predict the degree of change expected from a prescribed increase in imperviousness.
10. Decision	Conclusion is that further urbanization in the Humber River watershed would negatively affect fish communities and brook trout biomass by x%. It may be recommended that the growth of urban areas be minimized or that mitigative measures be incorporated into future growth.
11. Management	Management would either ensure limited growth, or ensure that mitigative measures are put in place to eliminate effects.

3.4 Combining Study Designs

The following has been extracted and modified from (Stanfield et. al., 2000). As is common with most field surveys, there are often competing objectives among the partners/managers. The Bowmanville - Soper's watershed study was no exception. In the summer of 1998, the project team began to design a study to assess the current state of the stream resources in these two watersheds. The team chose to use the stream protocol as one of the tools for collecting field data and began by working through a process similar to the framework presented here. As the team articulated the study design, the project steering committee identified three seemingly conflicting study designs.

- a state of the resource survey stratified by physical features such as physiography and adjacent landuse;

- a state of the resource survey stratified by management zones (catchments with varying long term planning strategies); and,

- an impact assessment study testing study intended to identify changes in conditions over time.

The challenge was to design a study that could satisfy all of these objectives while recognizing that funds were limited. A balanced stratified random design was applied to satisfy the project requirements. The watershed was divided into 4 physiographic zones. Each would receive 25 % of the sites. Next, each of 3 landuse types (forested, agricultural and settlement) would be divided among the 4 physiographic zones, allocating equal numbers of sites to each landuse type. Finally, the eight management zones were not treated as strata per se, but acted as modifiers for the randomization. Site selection was done such that the sites were approximately equally distributed among the management zones. This design permitted the study team to ask a number of questions relating to the various strata.

4.0 LITERATURE CITED

Barbour, M.T., J. Gerritsen, B.D. Snyder and J.B. Stribling. 1999. Rapid bioassessment protocols for use in streams and wadeable rivers: periphyton, benthic macroinvertebrates, and fish. Second Edition. EPA 841-B-99-002.

Boyd, A., Fraser and B. Pond. 1998.

Cairns, J.Jr., P.V. McCormick and B.R. Niederlehner. 1993. A proposed framework for developing indicators of ecosystem health. Hydrobiologia, 263:1-44.

Canadian Council for Ministers of the Environment (CCME). 1996. A framework for ecological risk assessment: general guidance. Canadian Council for Ministers of the Environment, Winnipeg.

David, S.M., K.M. Somers, R.A. Reid, R.J. Hall and R.E. Girard. 1998. Sampling protocols for the rapid bioassessment of streams and lakes using benthic macroinvertebrates. 2nd edition. Dorset Environmental Sciences Centre, Aquatic Science Section, Ontario Ministry of the Environment, Dorset, Ontario.

Dissmeyer, G.E. 1994. Evaluating the effectiveness of forestry best management practices in meeting water quality goals or standards. United States Department of Agriculture, Southern Region, Atlanta, GA.

Doloff, C.A., D.G. Hankin and G.H. Reeves 1993. Basinwide estimation of habitat and fish populations in streams. General Technical Report SE-83, Asheville, NC: United States Department of Agriculture, Forest Service, Southeastern Forest Experimental Station. 25 pp.

Environment Canada. 1998. Pulp and Paper Technical Guidance for Aquatic Environmental Effects Monitoring. EEM/1998/1, April 1998.

Environment Canada. 2001. Metal mining guidance document for aquatic environmental effects monitoring. June 2001.

FCEMS (Federal Committee on Environmental Management Systems). 2000. An approach for assessing and managing wastewater effluent quality for federal facilities – final report, June 1, 2000. FCEMS Wastewater Working Group, Technical Secretariat, National Guidelines and Standards Office, Environmental Quality Branch, Environment Canada.

Green, R.H. 1979. Sampling design and statistical methods for environmental biologists. John Wiley, New York. 257 p.

Green, R.H., J.M. Boyd and J.S. MacDonald. 1993. Relating sets of variables in environmental studies: the sediment quality triad as a paradigm. Environmetrics, 4:439-457.

Greig, L.A., J.K. Pawley, C.H.R. Wedeles, P. Bunnell and M.J. Rose. 1992. Hypotheses of effects of development in the Moose River Basin, workshop summary. Prepared for the Department of Fisheries and Oceans, Burlington, by ESSA Environmental and Social Systems Analysts Ltd., Richmond Hill.

Karr, J. R. and E. W. Chu. 1999. Restoring Life in running waters: Better Biological Monitoring. Island Press. 206 pp. Kilgour, B.W. 1999. Using benthic macroinvertebrate communities to monitor stream quality with reference to approaches in Ontario. Report prepared by Water Systems Analysts for the Regional Municipality of Ottawa-Carleton.

Kilgour, B.W. and L. Stanfield. 2001. Protocols for delineating, characterizing and classifying valley segments. Report prepared by Jacques Whitford Environment Limited for the Ontario Ministry of Natural Resources and the City of Ottawa.

Klemm, D.J., P.A. Lewis, F. Fulk and J.M. Lazorchak. 1990. Macroinvertebrate field and laboratory methods for evaluating the biological integrity of surface waters. EPA 600/4-90/030.

McGuiness, F., L. Carl, R. Mackareth, A. Boyd, J. Fraser and B. Pond. 2000. Proceedings of the Ontario Workshop on Riverine Science Requirements, February 21 – 24, 1999 at the Leslie M. Frost Natural Resources Centre, Dorset. A Watershed Sciences Centre Report.

Ministry of Environment and Energy. 1996. Guidance on site specific risk assessment for use at contaminated sites in Ontario. Ontario Ministry of Environment and Energy, Standards Development Branch.

Newbury R. W., and M. N. Gaboury. 1993. Stream analysis and fish habitat design: A Field Manual. Co-Published by Newbury Hydraulics Ltd., the Manitoba Habitat Heritage Corporation and Manitoba Natural Resources, Gibsons, British Columbia.

Pedersen, F., A. Damborg and P. Kristensen. 1995. Guidance document for risk assessment of industrial waste water. Danish Environmental Protection Agency, Ministry of Environment and Energy, Denmark, Miljρprojekt nr. 298l.

Simonson T. D., J. Lyons, and P. D. Kanehl. 1994a. Guidelines for evaluating fish habitat in Wisconsin streams. U.S. Department of Agriculture, Forest Service, General Technical Report NC-164, St. Paul, Minnesota.

Stanfield, L., M. Jones, M. Stoneman, B. Kilgour, J. Parish and G. Wichert. 2000. Stream Assessment Protocol for southern Ontario. V 4.1. Ministry of Natural Resources training manual.

Suter, G. and L. Barnthouse. 1993. Assessment Concepts. In, G.W. Suter (ed), Ecological Risk Assessment. Lewis Publishers, Boca Raton.

U.S. EPA. 1992. Framework for ecological risk assessment. Risk Assessment Forum. February 1992. Washington, DC. EPA/630/R-92/001.

U.S. EPA. 1992. Framework for ecological risk assessment. Risk Assessment Forum. February 1992. Washington, DC. EPA/630/R-92/001.

Underwood, A.J. 1993. The mechanics of spatially replicated sampling programmes to detect environmental impacts in a variable world. Aust. J. Ecol. 18:99-116.

Underwood, A.J. 1994. On beyond BACI: sampling designs that might reliably detect environmental disturbances. Ecological Applications, 4:3-15.

Underwood, A.J. 1991. Beyond BACI: experimental designs for detecting human environmental impacts on temporal variations in natural populations. Australian Journal of Marine and Freshwater Research 42: 569-587.

United States Environmental Protection Agency (USEPA). 1999. Screening level ecological risk assessment protocol for hazardous waste combustion facilities, volume one. USEPA, Solid Waste and Emergency Respose. EPA530-D-99-001A.

WRC. 2000. Watershed Report Card: Manuals for community involvement in watershed management. Published By Watershed Report Card Inc., Wiarton, On.

Appendix 1

Partial Listing of Available Stream Sampling Protocols

Protocol Element	Protocol	Source
Fish Community	1-pass electrofishing (biomass)	Stanfield et al. (2000)*
	3-pass removal electrofishing (population estimate)	Stanfield et al. (2000)* Zippen 1958
	qualitative electrofishing	Stoneman MTO
Benthic Community	rapid bioassessment[3]	Stanfield et al. (2000)*
	rapid bioassessment[3]	Barbour et al. (1999)
	rapid bioassessment[3]	David et al. (1998)*
	Point-source EEM programs	Environment Canada (1998, 2001)
	Travelling kick methodologies	Reynoldson et al. (1999)*
	misc	Klemm et al. (1990)
	Point transect kick and sweep	
Habitat	point-transect	Stanfield et al. (2000)* Simonson et. al., (1994)*
	Visual	Dodge et al. (1984)
		Barbour et al. (1999)
	Calibrated visual based	Stanfield et al. (2000)* Dolloff et. al., (1993) Watershed Report Card (1999)
Geomorphic assessments	Natural Channel Design Manual	Jack has reference?
	Diagnostic indicators of channel stability	Stanfield et al. (2000)
	Stream habitat and fish habitat design	Newbury and Gaboury (1993)

Note: Protocols with an * are accompanied by a data management system

[3] These protocols are designed to evaluate the same indicator and are virtually the same except for minor differences in specific habitats sampled, and taxonomic effort.

Section 1 Module 2

Project Design and Site Selection[1]

TABLE OF CONTENTS

APPENDICES

[1] Authors: L. W. Stanfield and B. Kilgour

1.0 INTRODUCTION

The most important question a stream surveyor needs to answer before you begin a habitat survey is "What is the question that I want to answer from the data"? The answer to this question is integral to a successful study design and will guide you in deciding where to go to collect the data, determining how many sites you should visit and how often. The most precise habitat data you can possibly collect at a particular site will be of little value if it cannot be shown to meaningfully represent some larger area of interest. In this module, we discuss the issue of project design and its relation to site selection for two common study types; hypothesis testing and impact/risk assessment.

The previous module (S1. M2) provides a framework to guide managers in selecting what type of study should be conducted to meet the objectives of the project and in selecting appropriate indicators. In this module we provide minimal guidance to managers on this process. Instead we focus on developing the study design, once the question has been defined.

1.1 Background

Habitat surveys can be done for any number of reasons including:

- baseline studies (state of some resource)
- impact assessments (to evaluate an effect from some action (i.e., landuse etc.,)
- Risk assessments (to predict the effect from some action)
- Or experiments to diagnose relationships between attributes, functions or processes

Without exception, the first issue that you should address at the start of any habitat survey is to *clearly* define the purpose of the survey. It is impossible to make wise choices about where to locate your sites without a clearly defined purpose. We recommend you write down the study purpose in sufficiently clear and unambiguous terms that someone could pick up your statement of purpose and from it design the same survey. If you find this task difficult, you are not ready to go out into the field. If you have been contracted to do the survey for someone else, make sure this statement of purpose is in the contract's terms of reference or make sure that the client agrees to the purpose before you start to choose sites. There is essentially no way to correct your data after the fact by reanalysing the data in a different way to accommodate for a poorly designed survey.

1.2 Evolution of the Module

This module evolved from a number of efforts to develop guidelines for risk or impact assessment assessments (EPA, 1993; CCME, 1996; MOE, 1996; Gillespie 2002) and experiences of the authors in supporting project managers attempting to develop stream surveys in Great Lakes tributaries.

1.3 Potential Uses

This module can remind you of the importance of carefully considering the issue of site selection in your study design. It can provide some background information on the kind of variation that you might expect in the habitat attributes that you will be measuring. This variation is critical to determining how many sites you need to survey. Finally, this module can provide you with a process for designing your study.

This module is intended to provide project managers with the tools to have some confidence in the design of the study before the crews go out and to prevent the default study design of "sample as many sites as possible wherever access is good".

1.4 Issues to be considered

1.4.1 Effort to be Expended

Sampling effort depends on the magnitude of difference that is considered important to detect for experiments and for parameter estimation studies, it depends on the precision in estimation that is considered relevant. For example, in an experiment, we might wish to determine if restored streams had better habitat for brook trout. Our null hypothesis for the study is: "H_0: there is no difference in the mean habitat suitability of restored and unrestored streams. However, we will want to survey enough streams to ensure that we have a reasonable chance of detecting a difference (i.e., rejecting the null hypothesis) when the true difference is of some magnitude. For example, we might specify that we want to survey enough streams to have greater than a 90% chance of detecting a difference in brook trout HSI scores when the true difference is 25%.

In contrast to experiments, the number of streams required in a study is dependent on the precision we hope to obtain for our estimate of a stream habitat attribute. Precision refers to how narrow the confidence limits on the attribute will be. For example, we might wish to "estimate the potential parr production of Atlantic salmon from north shore tributary streams within ±25%". **This is perhaps the hardest part of any study design, with the possible exception of addressing scale.**

1.4.2 The appropriate scale for a study

A tremendous challenge for a project manager is deciding the temporal and spatial scale of any study. That is over what geographic area do I need to sample to answer the question and over what time frame. Spatial scale questions must consider both the landscape component (i.e., geographic boundaries of the survey) and the area of the stream to be sampled.

In the OSAP manual an attempt has been made to standardize the definition of site length. The advantage of this is that data collected across sites and times can be compared without the additional variance that is introduced by varying site length. These moderate length sites represent a compromise that enables local conditions to be quantified, while at the same time enabling variance between sites to be measured. This design is appropriate for measuring some micro-habitat features (i.e., physical habitat) but not others (i.e., microhabitat use by fish). For specific studies, managers may find that the benefit of having longer or shorter sites outweighs the lost comparability to other studies.

1.4.3 What diagnostic power do I need to apply to answer the question

Managers should attempt to balance the cost and benefit of selecting any protocol for conducting field surveys with the desire to sample the site(s) in an efficient manner. This means matching the diagnostic potential of a protocol to the study question. It may be helpful to think of questions in terms of whether you are using the data for screening, assessment or diagnostics (see S1.M1), with intensity (cost) increasing through the continuum.

While some protocols are appropriate at all three levels of intensity, many are not. Therefor the project manager must select the appropriate standard sampling protocol to meet both the short and long-term needs of the project. This requires a clearly articulated objective for each study. Within each section of the OSAP manual an effort has been made to evaluate the suitability of each protocol for meeting each of these three major study types. Note that many more protocols are available for diagnostic purposes than are described in the OSAP manual.

Definition: Microhabitat Features

Attributes of streams that are measured at small scales such as depth profiles, cover and vegetation distributions.

1.5 Limitations

This module does not address study designs associated with regression studies. These are popular research tools designed to demonstrate the relationship between one or more variables from a data set. If your study fits this definition we advise you to seek out one of the more popular statistical design books (i.e., Krebs 1989,

1.6 Time and Resources Required

In most impact assessment studies, the location of the sites will be predetermined. In which case filling out the appropriate forms associated with this module should take no more than 1 hour. In other situations, developing a sound study design could take as long as a week or more, but again the filling out of the forms should not take more than an hour. A GIS or maps (stream network, geology, landuse etc.,) of the study area are invaluable in helping design a study.

All the information is recorded on the three linked forms: Project Description form; Project Objectives; or the Project stratification and site numbers determination. This information can be entered and stored within the Habprogs application as a permanent record of the rationale and biases associated with the sites selected for sampling with each project. We encourage all users to make use of this tool.

2.0 DESCRIBING THE PROJECT

Each time a new project is carried out you will need to clearly describe it. This means giving it a name, code, and identifying the project manager and organization responsible for the project.

On the Project Description Form, record the project name. This is entered into a memo field but ideally should not take more than 1 typed line. The title should include some information about the sample area and purpose of the study, i.e., Assessment of Waring Creek Restoration Efforts or Assessment of Potential Atlantic salmon Parr Production from North Shore Streams of Lake Ontario.

Next, record the name of the project leader and the name of the organization responsible for the project. By responsible, we refer to the group who initiated and is responsible for the quality of the study design. This may or may not be the same group who carried out the project.

Provide a brief description of the project in the description box. Record the name of the upper tier level of government responsible for the study area in the field called "geographic boundaries". This field is intended to provide future users of the data with a guide to available data sets. In most areas, this level will be the county or municipality. For studies which occur over extensive areas, list all of the areas included in the study area (i.e., Durham, Peel, York, and Metro Toronto). In unorganized areas, record the Ministry of Natural Resources district.

Record the final location of the Project Description Forms and field data sheets in the box titled "location of data sheets/etc". As manuscripts, data summaries, or technical reports are written which utilize these data, record this information in the box titled "References associated with this project".

Hint: Don't forget to update these forms if storage places change or new reports are created. As the originator of the data, you have a vested interest in ensuring that it is available to future generations.

2.1 Defining Project Objectives

There are several decisions required to ensure the project objectives are clearly articulated. First, you need to decide which **study type** you are carrying out. The next step is to **list the parameters** you intend to use to make your assessment and the **level of precision you desire in the results**. The final step involves defining and **limiting the number of modifying factors** and then carefully considering whether your study would benefit from **stratification**. For example, you need to **define the boundaries** for the study building on what you already started in Step 1. Each of these steps requires serious consideration, as they will have an enormous impact on your study results.

2.1.1 Determining the study type

The first task is to determine whether the study you are proposing to carry out fits either of our definitions:

- **hypothesis testing:** involves comparison between treatments. Whether the comparison involves sampling the same streams before and after an intervention or sampling streams at the same time that have been treated differently doesn't really affect the number of streams that are required for the study.
- **parameter estimation** involves studies intended to make some inference about a habitat or biotic variable over an extensive area. Examples include a watershed study (or inventory) to estimate the overall productive capacity of the streams or a study to assess the potential production of Atlantic salmon in Lake Ontario tributaries.

If your study design does not fit either of these scenarios, we strongly advise you to obtain the help of a professional statistician, as this module will not meet all your needs.

Based on the study description, determine whether you are carrying out a hypothesis testing or
a parameter estimation survey and mark off the appropriate box on the Project Objectives Form (see definitions below).

Definition:

Hypothesis Testing:
A study that will result in a comparison of one or more parameters between sites i.e.,

impact assessment	channel diversion
restoration work	species introductions
experimental manipulations	channel modifications
comparative projects	differences between basins or reaches

Parameter Estimation:
A study to derive an estimate of some attribute over an extensive area i.e.,

production estimate	fish production
sediment supply	food supply (invertebrates)
habitat availability	habitat suitability across broad range

Note: It is possible for a study to have multiple objectives such that both a hypothesis test and a parameter estimation study are carried out. In this case we suggest you mark both off on the Project Objectives Form but use the hypothesis test formula to determine the number of sites (Step 4).

2.1.2 Documenting the Parameters of Interest an Protocols to be used

Make a list of the parameters to be assessed in the study. Next, review your list and record the parameters that can adequately be assessed either directly or through a surrogate (see below) using the OSAP in the box titled "study parameters assessed by protocol". This step is necessary to ensure that users do not have false expectations about what the procedures in this manual can do. This step should be done by or with someone who has had the training course.

If, after this review, you conclude there are parameters not measured with this protocol that you believe are important to your study design, you will need to list these now in the box "study parameters not assessed by protocol". We also provide an area for you to describe methods/protocols you intend to use to address these additional parameters.

Before listing all the parameters of interest, you need to determine the most appropriate parameter on which to base the study. Selecting parameters as a basis for sample size calculations depends on the variable that you wish to use for your critical decisions. This decision will require balancing the need for precision in your estimate and the attribute or species of interest. The coefficient of variation is the best indicator of how precisely you can measure a given variable. This may relate to an attribute from an HSI model for an individual species such as rainbow darter channel homogeneity, or it may reflect a more generic variable such as the substrate composition (D50).

As a general rule and regardless of the habitat attribute, you can detect a difference that is about twice as great as the CV with as little as five samples from each stratum. This is about twice as great as the normal range of variability of the habitat attribute within each treatment (Environment Canada, 1998). Of course, the caveat is that this will only happen about 90% of the time.

2.1.3 Determining the Desired Level of Precision

Based on the study objectives and experience, you must decide the desired level of precision to allow the appropriate conclusions from the study. The precision is expressed as a percent with lower precision values indicating a higher degree of confidence in the estimate.

Again, as a general rule, the level of precision can be set at levels that would be considered meaningful to the target organism. The suitability curves provided in section 5 provide guidance for making these decisions. For example, the curves for temperature and habitat stability for some species tend to have steeper slopes on the ascending and descending sides of the suitability curves. These variables also tend to have low Coefficients of Variation (CV's) (Table 1.2)[1]. This information suggests that you should be able to set the precision level fairly low for these parameters (i.e., 20 - 40 %).

For parameters which have higher variation in the estimates (higher CV's) or which have lower slopes (contrast in the fishes' response to these variables), you

[1] For the most up to date and complete listing check HABPROGS. Click on the add or edit project definitions box, then click the "CV's" box.

will need to set your precision level higher to reflect the lower power of the data (i.e., 50 -100 %).

It is always useful to know the level of precision you can expect from a study before the data is collected so that you know whether you will be able to meet your objectives[2].

Table 2: Coefficients of variation for selected variables found within the Stream Assessment database

	Species Grouping						
	Generic	creek chub	rainbow darter	mottled sculpin	brook trout	blacknose dace	longnose dace
Channel structure SI		19.68	33.14	22.58	2.43	0.17	3.25
Channel stability SI		31.91	28.02	21.1	29.8	33.21	21.95
Substrate quality SI		23.98	35.1	30.86	42.21	32.78	34.23
Thermal stability SI		9.74	12.69	0.34	2.8	7.38	9.73
Homogeneity SI		50.76	5.03	65.9	72.08	97.52	73.2
Thermal stability	8.78						
Hilsenhoff score	5.08						
Homogeneity index	9.19						
% pools	14.32						
%depth < 10 cm	20.57						
% rock cover	35.67						
% wood cover	47.22						
D50 point particle	112.56						
D50 max particle	20.03						
Sediment sorting index	28.54						

Definition: Surrogates

There are many parameters of interest to ecologists for which direct measures are difficult to obtain. Biologists tend to substitute measures which are more directly and often easier to measure as surrogates for the parameter of interest. Often, you will need to look at several surrogates before you can begin to draw conclusions about your target parameter. For example, the width of a stream is a pretty good surrogate for the size of the catchment upstream of the site. Conclusions about the hydrology of a stream can be made based on the width/depth alone, but are greatly strengthened by a number of other geomorphic measurements.

---

[2] Note: Although it is not a requirement of this module to record the rationale for your decisions, it is advisable to do so as part of your overall project design

2.1.4 Defining the Modifying factors to be considered

There are many factors (too many to list here) which cause variability in study endpoints or variables. The intention of this section is to provide some guidance to users not experienced in identifying these modifying variables and using these to define the scope of a study. More detailed information about modifying factors (variable(s) that addvariation to the measured parameter) and methods to use these to stratify your study are provided in S1.M3.

This step is optional because as long as you follow the steps to choosing a random sampling design, your study will be sound. This section is only intended to help reduce variability in the data set so as to increase the likelihood that real differences will be found with a given study design, or to reduce sample size requirements for parameter estimation studies.

With that, the next question that the project manager should answer is "*Are there modifying factors which are likely to increase variability in the results to such an extent that the scope of the study should be reduced?*"

If the answer to this question is yes, list the modifying factors that you will use to bound your study as well as the criteria to be used to identify boundaries. Boundary criteria are the definitions that will be used to split sample areas. For example: if geology is used as a modifying factor, the boundary might be shield versus non-shield. There are many possibilities for you to consider. We provide a partial list to of some of these in (Table 1.3).

Remember that each variable you identify reduces the scope of your study, so beware of choosing too many.

Table 1.3: Partial list of Potential Modifying Variables for Study Scoping or Stratification

Modifying Variable	Description
Soils	influence drainage into stream, nutrient passage (i.e., pass quickly through sand) and channel type
Bedrock	sets the overall hardness and baseline productivity for the system; its proximity to the surface effects channel dimensions
Physiography	is important to any study comparing the hydrology (particularly the baseflow conditions), thermal stability, energy dynamics, substrate composition, channel type, valley profile between sites or reaches
Topography	has a tremendous bearing on the channel and its energy dynamics, particularly slope and

	valley configuration.
Climate	affects the potential thermal stability of the stream, periodicity and intensity of high flow events, and the degree-days of growth.
Drainage Area	directly influences the size of the stream and therefore amount of available habitat and number of modifying conditions acting on the stream from upstream; some species of fish are restricted by size of stream
Tablelands	Landuse on the lands which lie higher than the 100 year flood level influence has its greatest effect on hydrology by altering peak and low flow conditions but also affects contaminant levels and sediment loading.
Valleylands	Activities in the lands which lie within the 100 year flood plain can influence many functions such as disconnecting the active channel from the flood plain, altering channel stability and the amount of cover, nutrient inputs, reduced shade, etc.
Channel Roughness	is used to describe the barriers to flow within a stream; includes the meander shape, and amount of wood or roughness of the channel bottom; can influence both the geomorphology and habitat suitability of a stream
Species Distribution	projects which have a biological focus must be aware of issues related to the distribution of species, whether it be natural or influenced by barriers

Action: Record the scope of the study, including a list of factors used to reduce the scope of the study on the Project Objectives Form.

2.2 Determining the Need for Stratification and Number of Sites to Sample

The next step is to determine whether your study would benefit from stratification. Stratification (see definition) is used to reduce variation between sites within a grouping. Review the study parameters and the potential modifying factors and determine whether each factor will become a stratum (with categories > 1) or simply a modifying factor (1 category).

Stratification

To stratify means to divide the study design into equal or representative groupings of various factors. Stratification is usually used for one of two reasons. First, in a hypothesis testing study, stratification might be used to allow a number of questions to be answered. In effect they become independent tests which share a common data set. In this instance, stratification increases the number of sites to be sampled because each stratum is independent. Stratification in a parameter estimation survey is designed to increase precision of an estimate within a grouping.

List each of the strata in your design on the Requirements for Stratification and Boundary Determination Form. This step tightens the study design and ensures that you are able to achieve your objectives given the logistics of how many sites you can sample. Note that this is an iterative process: once you have described the ideal scenario, the next step is to calculate the number of sites required meeting the study design. Depending on the results of that exercise, you may be back to this step to redefine the scope of the study.

Warning: For hypothesis testing studies, stratify with caution (i.e., > 1 category per factor) as it can greatly increase sample sizes.

2.3 Step 4: Determining How Many Sites are Required for the Study

The next step is to decide how many sites to include in your survey or how many samples to take at one or several sites over time. Different procedures are used depending on whether the study is to carry out a hypothesis test or parameter estimation. For this reason, we have split this section.

In both instances, the formula used to calculate the number of sites relies on some estimate of the Coefficient of Variation (CV) for the parameter of interest. Fortunately, in 1996, we carried out surveys at 3 sites along a number of stream segments. These surveys provided us with the necessary data to calculate the natural ranges in variability (CV's) for both the biotic and abiotic habitat parameters summarized using this protocol. A subset of the results of this analysis is provided in Table 1.2. A much more complete table of the CV's for each variable measured at these sites is located in the HABPROGS database, in the project description section (see CV's).

To use either of these formulae, you will need the CV's for the **most appropriate test parameter** in your study design. Choose your decision parameter carefully: the CV of this variable will control the precision of your study and the number of sites you will need to sample.

2.4 Determining the number of sites per strata

2.4.1 Parameter Estimation Study

The formula to be applied for parameter estimation surveys follows from Mendenhall (1983):

$$n = \frac{Z_\alpha^2 \, CV^2}{\%d^2}$$

where:

n is the estimated sample size,

Z_α is the standard normal deviation for a probability of α

%d is the desired precision expressed relative to the mean (i.e., we hope to estimate the true mean within 25 %)

CV is an estimate of the coefficient of variation

This equation essentially estimates the number of samples required to get within some percentage of the true population mean. If we set $\alpha = 0.05$ for a 95% confidence limit, Z becomes 1.96. If you wish to set a different level of confidence, you will need to consult a Z table in a statistics text.

To complete the equation, fill in the precision estimate you desire for your study from the project objective section. Then, look up the CV for the most sensitive parameter in your study and include this in the equation and you will have an estimate of the number of sites required to have sufficient precision in your study to meet the project objectives.

For the Atlantic salmon parr assessment example, the variables are as follows:
$Z = 1.96$
%d = 25 (i.e., we had set the level of significant change at 25 %)
CV = 30.86 (the surrogate CV for substrate quality for mottled sculpin)

With these numbers, $n = \frac{1.96^2(30.86^2)}{25^2} = 6$

The project should carry out a minimum of 6 samples in each sample stratum (i.e., above and below barriers, etc.).

We have used this example to illustrate:

1. the effect of the desired level of precision
2. the effect of using a parameter with a relatively low CV on the number of sites required
3. how useful this process is for checking feasibility on the project objectives

Note: This equation assumes that the original estimate of the coefficient of variation was based on a large sample (> 30 observations). The results presented in Table 1.2 were determined from 65 sites.

Another way to use this formula is to estimate the precision of your estimate for a given sample size (i.e., your study design is constrained by the number of sites which can practically be sampled). The formula is rearranged to provide the following:

$$\%d = \sqrt{\frac{1.96^2 CV^2}{n}}$$

This can be very useful to managers forced to allocate limited resources and still come up with an answer. You can use this formula as a reference to see whether the scope of your study is in the right ballpark. For example, if you knew you could only sample 10 sites, you can run the CV for the various habitat features through the formula and see whether this level of sampling will achieve the level of precision directed by the project manager.

Reminder: The number of sites estimated using this formula represents the number of sites required in each stratum. This is another reason to be very cautious in adding strata to your design. The Rapid Assessment Module can help by providing a means to obtain many sites worth of habitat data (albeit with higher CV's) at a fraction of the effort (Module 11).

Action: Fill in the variables used to calculate the number of sites. Then, using the appropriate formula, calculate the number of sites to be used in each stratum and record this number in the box marked "n =".

2.4.2 Hypothesis Testing Study

For hypothesis testing studies, the approximate sample size required is determined from the formula provided by Alldredge (1987):

$$n = \frac{2(Z_\alpha + Z_\beta)^2 CV^2}{\%d^2} + 1$$

where:

n is the sample size
Z_α is the standard normal deviate for a two-tailed test with significance of α
Z_β is the standard normal deviate for probability β
CV is the coefficient of variation of the measured variable
%d is the percent change that is perceived to be important to detect

This equation estimates the number of samples or stream sites needed in each stratum to correctly conclude that an effect has occurred when it has (i.e., 1-β) or to correctly conclude that an effect has not occurred when in fact it has not (i.e., 1-α).

For most applications of this protocol, the confidence limit is likely to be set at 95%. If you do this, the equation reduces to:

$$n = \frac{26(CV^2)}{\%d^2} + 1$$

This happens because 1-α becomes 1.96 and 1-β becomes 1.645. If you decide to use a different detection level, look up α and β values in a statistics book and solve for ($Z_\alpha + Z_\beta$).

For a s study designed to measure a 40 % change in the primary output variable channel stability, the formula would be as follows:

$$\frac{26(29.8^2)}{40^2} + 1 = 15 \text{ sites per stratum}$$

The CV for channel stability is 29.8 and a moderate precision rate of 40%. Still this represents a significant amount of sampling on a small stream in order to detect a meaningful difference

Table 1.2 can be used to evaluate which parameter might provide useful surrogates with sufficiently low CV's to permit detection of change with reasonable sampling effort. The suitability curves provided in section 5 should provide some guidance to project managers in selecting which parameters provide the greatest explanatory power for fish species of interest.

Finally, the manager also has the option of using data from the individual transect or point level surveys, but this would require additional work to determine the CV's for the parameter of choice. For example, a survey could test that the stream width has changed and interpret the data on a transect to transect basis (10 per site) instead of the average of a site.

Alternatively, you could reconsider your level of precision required or the detection level. Note that changing the detection level will have the least effect on the sample size estimate. Should you decide to change the objectives, enter the new criteria into the equation and recalculate.

The formula can also be rearranged to determine the detectable change when a sample size is fixed:

$$\%d = \sqrt{\frac{(Z_\alpha + Z_\beta)^2 CV^2}{n - 1}}$$

or for 95 % C.I.

$$\%d = \sqrt{\frac{26CV^2}{n - 1}}$$

A word of caution

The CV is not a measure of the accuracy of the data collected, although it is certainly part of it. More appropriately, this number represents an estimate of how variable the observations are between sites within sections of stream which have relatively similar adjacent landuses. In all instances the estimated CV between sites is much larger than that found when we compared visits at the same site, implying that most of this variation represents true differences between the sites.

2.4.3 Locating the Sites within the Study Area

Now that you know how many sites you need, the next step is to decide where they should be. We strongly recommend you adopt random sampling principles (see Advantages of a Random Sampling Design; Appendix 3).

First, determine how many potential sites you have within your sample area. Measure the total length of stream in each stratum. This can be done in many ways, ranging from using a digitizer to simply laying a ruler or piece of string on a topographic map. A number of GIS applications have been developed to assist with this task (ALIS, 2002, ORSECT 2002 and WRIP 2003) and in some instances (i.e., ALIS) the stream lengths are already available for segments that were determined using the rules described in the next module (S1.M3).

Next divide the total length of stream available to sampling in each strata by the expected average site length[3] and subtract 1 to obtain the search list for the set of sites ((i.e., total length/45) - 1). Finally, use random numbers from a table or function within a spreadsheet database to pick which of the potential list will become the sample sites. The number of sites to be surveyed in each stratum dictate how many random numbers you will choose. The procedure is the same for the parameter estimation procedure. The last step is to mark on a map the location of each site (see Appendix 2).

2.5 Tips for applying this module

Do this module at the early stages of any project.

Have someone else review your criteria to ensure you have not missed anything. One of the most important aspects of this module is that it provides a mechanism to articulate study designs so that they are very clear to everyone. Therefore, this makes the data you are collecting even more valuable to future users. Take the time to ensure that there are no ambiguities in the project design.

Remember: the most important step is that the sites chosen are true to the objectives of the study design. In most instances, this will be random selection.

[3] use the approximate relationship of 7* average stream width for streams wider than 6 m or the default length of 45 for those narrower

3.0 DATA MANAGEMENT

1. Enter the data from the three forms into the HABPROGS database (see OSAP manual: S5.M1), and save backup copies that are stored in a separate location from the master copy.

2. (Optional) Provide a copy of the database to the OMNR for inclusion into the master database.

By storing the data digitally in HABPROGS, the data can be shared with a large number of users province-wide. Data sharing will facilitate the refinement and development of habitat suitability models, which will lead to improved habitat management practices and policies.

4.0 LITERATURE CITED

ALIS, 2002. Aquatic landscape Inventory System. Ontario Ministry of Natural Resources, Geomatics Section, Peterborough Ontario.

Canadian Council for Ministers of the Environment (CCME). 1996. A framework for ecological risk assessment: general guidance. Canadian Council for Ministers of the Environment, Winnipeg.

Chang, C., F. Ashenhurst, S. Damaia and W. Mann. 2002. Ontario Flow Assessment Techniques, Version 1.0. OMNR, Northeast Science & Information, South Porcupine, Ontario. NESI Technical, Manual-011 122p. + App.

Krebs, C. J. 1989. Ecological methodology. Harper Collins.

Ministry of Environment and Energy. 1996. Guidance on site specific risk assessment for use at contaminated sites in Ontario. Ontario Ministry of Environment and Energy, Standards Development Branch.

WRIP, 2003. Water Resources Information Project: GIS Toolbox. Ontario Ministry of Natural Resources, Geomatics Section, Peterborough Ontario.

Appendix I

Examples of Parameter Estimation, Hypothesis Testing and a balanced stratified random design survey

To assist in the explanation of how to use this module, we will describe 3 projects. The first two are projects that they have yet to be carried out, therefore there are no data or reports. The first is a hypothesis test that we will refer to as the Waring Creek study (Example 1). The second is a parameter estimation study where our goal is to determine the overall suitability of Lake Ontario tributaries for Atlantic salmon restoration (Example 2). As the various steps are described for these projects, refer to the example sheets at the end of this module. The last project described is a real project that shows how a slight variation on the process described here was necessary to meet all the project needs without compromising the overall study design.

Example1: <u>Hypothesis Testing</u>

Waring Creek is a small (10 km long) stream, which is located in Prince Edward County. It has its origins in an esker and drains through an intensive agricultural area, which has extensive zones of reclaimed forest. A community group has been carrying out extensive restoration work along the main stem of the stream and is interested in assessing whether this work has made the system suitable for brook trout to be reintroduced.

The parameters to be assessed were chosen to be channel structure, thermal stability, substrate quality, percentage deep pools, and habitat stability. The habitat stability parameter for brook trout was selected as the output variable for the study. The project leader came to this conclusion based on the information that the group has been concerned about channel stability because of the sediment loading and transport in the stream, this parameter has a reasonable CV (at 29.3), and the suitability curve for both age classes of brook trout show a high R^2 for this variable. The group decided to set the precision level at 40 %, meaning that the null hypothesis would be: that habitat stability scores in restored areas of Waring Creek have not improved by more than 40 % from unrestored areas.

This study is being conducted in an area with intensive agriculture and riparian forest areas that are in a recovery stage. This study could benefit by reducing the study scope to stream sections where adjacent landuses and the riparian zone vegetative community are in a similar state. These criteria were used to identify the study boundaries from a map of the area.

This hypothesis testing project should only be stratified by the restored and unrestored sections of streams (2 categories) and the adjacent landuse and riparian vegetation should be used as modifiers to control site selection (1 category i.e., the same). This would prevent samples being taken in the control area where the modifying effects might overwhelm data interpretation.

For this experiment, we conclude that the project can realistically only assess 5 sites in each stratum. We can use the rearranged formula to determine that this sampling strategy will be sufficient to detect a difference of only 68 %.

The available stream length for Waring Creek was 2800 m of restored habitat and 1200 m of control area. Of these totals, the group determined that 1200 and 800 m of stream from each stratum met the criteria of having similar landuse and riparian vegetation.

The potential number of sites for each stratum are 27 (1200/45) and 18 (800/45) so that the random numbers search list for both would be from 0 -26 and 0 - 17 (you must subtract 1 in order to search on the right number of sites). The random number generator picked the location for the five sites in each stratum to be: 2, 9, 21, 6, 23 and 1, 8, 15, 4, 7.

The first site of the treatment area would begin at the closest crossover, occurring 90m (2 x 45) up from the downstream end of the study area. The second site begins at the closest crossover which occurs at 450 m (9 x 45) upstream of the downstream end of the study area etc.

Example 2: Parameter Estimation

This study will develop a quantitative estimate of the potential production of Atlantic salmon parr (fall fingerlings), in both accessible and presently inaccessible north shore tributaries of Lake Ontario. The survey will utilize parr survival data from experimental stocking studies in combination with both site level and basin wide (GIS) habitat data collected across the north shore tributaries of Lake Ontario.

This study proposed to use the substrate quality for mottled sculpin as a surrogate for Atlantic salmon since no curves are available for this species yet. Substrate quality was chosen because there is high contrast in scores within the parameter, Atlantic salmon densities are known to be related to the amount and quality of coarse substrate, and the CV for this parameter is reasonable (30.9). The precision level for the study was chosen as 25 %.

There are many potential modifying factors. The project manager chose four. Access from Lake Ontario was chosen to reflect the evidence showing the effect of pacific salmonids on Atlantic salmon production. Thermal stability was chosen because these fish do not thrive in warm water streams. Gradient was chosen as a surrogate for the preferred habitat of coarse substrate and riffle habitat. Finally, stream order was chosen as a surrogate of width. Each of these modifiers was chosen because of evidence that they are likely to have a substantial effect on data variability.

The parameter estimation project should stratify by access (2 factors), well vs. poorly drained catchments (2 factors), and steep vs. gentle stream gradient (2 factors). In addition, a decision was made to avoid 1st order streams because of the large amount of effort they would require to sample (it would double the number of sites required) as well as a belief that the main stems were likely to be the main parr producers. Basically, this decision postpones work on these smaller sites until a later date. This design should minimize variance within strata to permit the overall confidence limit to be even better than the ± 25 % predicted from this design.

The Atlantic salmon study would first require that catchments be divided into well and poorly drained sections, accessible and inaccessible reaches, and low and high gradient areas. The next step would be to highlight all areas that are 2nd order and greater in size. This would be greatly aided by a GIS. Once completed, the total available sites for each of the 8 strata could be determined.

Example 3: Bowmanville - Sopers Watershed Studies:

As is common with most field surveys, there are often competing objectives among the partners/managers. The Bowmanville - Soper's watershed study was no exception. In the summer of 1998, the project team began to design a study to assess the current state of the stream resources in these two watersheds. The team chose to use the stream protocol as one of the tools for collecting field data and began by working through this module to articulate the study design. The project steering committee soon identified three seemingly conflicting study designs.

a parameter estimation survey stratified by physical features such as physiography and adjacent landuse.
a parameter estimation survey stratified by management zones (catchments with varying long term planning strategies).
a hypothesis testing study intended to identify changes in conditions over time.

The challenge was to design a study that could satisfy all of these objectives while recognizing that funds were limited.

The hard part was done, the objectives had been articulated. The solution was quite simple, although it was a modification of what is written in this section. A balanced stratified random design was applied to satisfy the project requirements.

The watershed was divided into 4 physiographic zones. Each would receive 25 % of the sites. Next, each of 3 landuse types (forested, agricultural and settlement) would be divided among the 4 physiographic zones, allocating equal numbers of sites to each landuse type. Finally, the eight management zones were not treated as strata per se, but acted as modifiers for the randomization. Site selection was done such that the sites were approximately equally distributed among the management zones. This design permits the study team to ask a number of questions relating to the various strata. Finally, the sites can still contribute to a long term monitoring strategy, permitting comparisons to be made about changes across the various strata.

Site selection required several steps, including creating an overlay with the physiographic zones, landuse category, and management zone and river. Next, the available length of river was measured for each stratum (i.e., Oak Ridge Moraine, forested). A random numbers generator was used to determine exactly where the site should be selected within each category. Once the target number of sites for each category (including management zone) was reached, that category was closed. Random selection would continue to fill the remaining categories.

This design permits long term sampling to test a number of hypothesis relating to how both landuse and physiography influence stream habitat and biota. The project team discovered how important it is to spend the time on study design, which took over 2 weeks. The end result was well worth the effort according to Samantha Mason, project biologist with CLOCA.

Appendix II

Advantages of a Random Sampling Design

Random sampling poses a major logistic challenge to the design of a habitat survey. It is much easier to survey sites nearby road crossings or other access points than to hike half a mile into dense bush because of random numbers, but there is no defensible way of avoiding this. If your purpose is to say something about habitat conditions in a sub-watershed or a stream section several kilometres long, you cannot simply assess sites that are readily accessible. We have tried to develop a protocol that is relatively quick and not equipment-intensive so getting to the sites is not additionally onerous. Besides, in southern Ontario, access points are sufficiently abundant that random sampling is not out of the question.

Random sampling comes in a variety of flavours. We will not go into the details here. See the monograph by C. J. Krebs - *Ecological Methodology* (Krebs, 1989) for an excellent and readable discussion of survey sampling design. It is likely that you will want to use stratified random sampling, particularly for surveys where you expect big differences among sub-areas in habitat conditions (i.e., watersheds with a mixture of forested and open riparian areas or big changes in topography). One stratification possibility you might wish to consider for pragmatic reasons would be dividing your total area into easily and not-so-easily accessible zones. You could then sample more intensively in the easily accessible zone and test the null hypothesis that the easily accessible zone is not significantly different from the less accessible zone. If this hypothesis is accepted, you can justify relatively few samples from the less accessible zone.

Section 1 Module 3

Guidelines for delineating, characterizing and classifying valley segments[1]

TABLE OF CONTENTS

INDEX OF FIGURES

[1] Authors:B. Kilgour, L. W. Stanfield and R. Kuyvenhoven

INDEX OF TABLES

APPENDIX 1: Classification of riverine habitats using temperature and nutrient status.

1.0 INTRODUCTION

1.1 Background

Frissell et al. (1996) described the relationships between habitat features measured at different scales within a catchment. Valley segments represent the 2nd level in the stream-habitat hierarchy with several segments making up a catchment, 1+ reaches comprising a segment, 1+ riffle/pool systems comprising a reach, etc., (Figure 1.1).

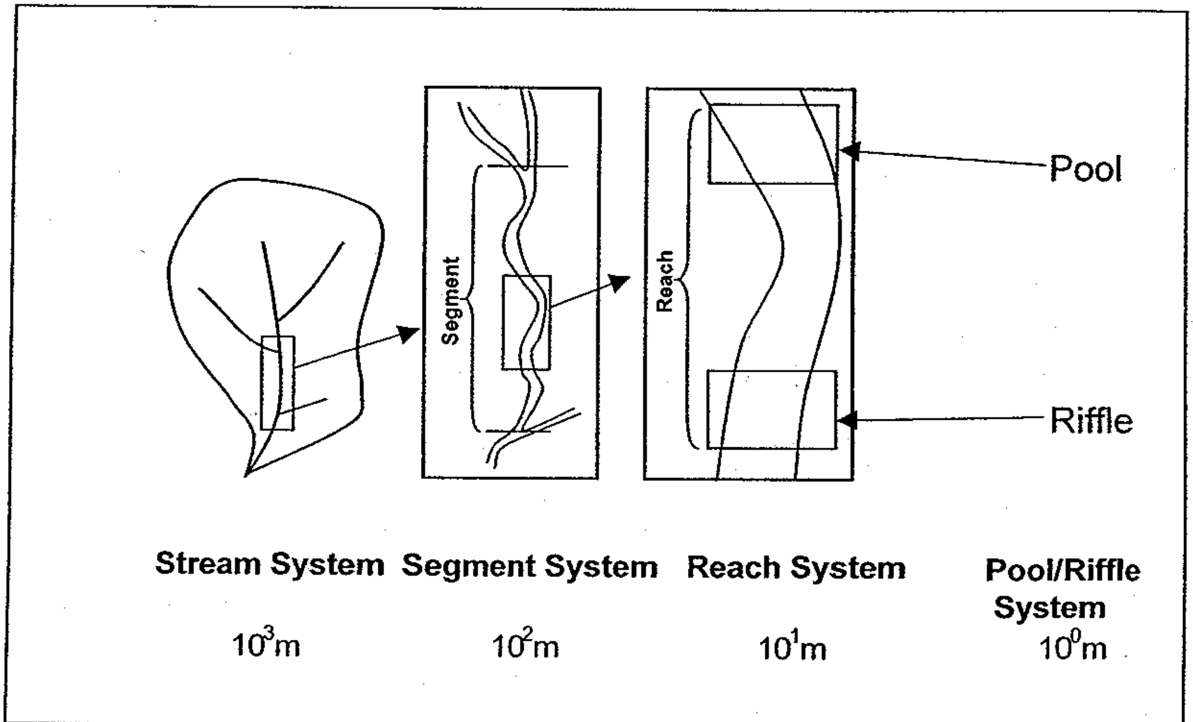


Figure 1.1. Relationships among stream habitats.
Redrawn from Frissell et al. (1986). Pools and riffles, etc., comprise reach systems. One or more reaches comprise a river segment, and one or more segments comprise a stream/river system.

Seelbach et al. (1997) proposed that valley segments could be used as the fundamental unit for managing river systems because:

- They are at the scale (size) at which rivers react to local slope, geologic materials and land cover/use (Maxwell et al., 1995).
- They cover the typical home range of riverine fish species. A single valley segment is large enough to contain the multiple habitats required by most stream fish during their life cycle (Schloesser, 1991).

- They are fairly homogeneous. Therefore, they have internal organization of smaller-scale units (reaches), and further-nested channel habitats (pools, riffles, and substrates) that are used by organisms during their life stages and seasons. Therefore, segment attributes can include the description of patterns in local habitat structure and provide a framework for smaller-scale classification where desired.

- They are the smallest habitat unit that can be interpreted from low-resolution maps and analyzed across large geographic areas (Seelbach et al., 1997).

The use of valley segments as the fundamental unit for managing riverine resources has considerable potential. For example, in subwatershed or other large-scale studies, valley segments are a potentially useful stratification tool. The Sydenham River Recovery Team is using valley segments as a basic unit on which to compile biophysical data (Kilgour et al., 2000). The Nature Conservancy Canada (TNCC) has also recently proposed classification of river habitats in Ontario using valley segments as the basic habitat unit. In the U.S. Michigan DNR and The Nature Conservancy have collaborated to inventory valley segments of river systems draining the upper Great Lakes. TNC has also identified rare and sensitive aquatic habitats, using valley segments as the basic habitat unit.

Since valley segments are a proposed habitat unit for applying management decisions, it is important that approaches for defining and delineating valley segments be similar. It is also important that the attributes of valley segments, on which management decisions are based, be characterized in the same way across jurisdictions. Presently, this is not the case.

1.2 Objective

The objective of this series of protocols is to standardize methods for delineating, characterizing and classifying segments. Much of the material in these protocols is taken from either Seelbach et al. (1997) and/or Higgins et al. (1998). Many of the attributes on which segments are delineated, characterized and classified are measured at the landscape scale, either from digital or hard copy maps. The protocols presented not only attempt to standardize data collection, but also how the data may be stored in electronic media.

1.3 Organization

This document has several modules that are inter-linked, but that can also be used as stand-alone documents. Delineation is the process of dividing up stream and river habitats into relatively homogeneous sections. That process is described in Module A1. The naming of valley segments requires organization so that data obtained from other exercises (i.e., field collections) can be linked to the correct segments. That process is described in Module A2. Module A3 describes protocols for characterizing features of catchments that, *a priori*, are expected to be predictive of physical, and chemical characteristics of streams, and of the kinds of animals found in reaches and sites within valley segments. Because biophysical attributes are often measured at the site-scale (e.g., Stanfield et al., 2000), some comparison between what is predicted and what is actually observed can be used for assessment purposes. That and other uses of valley segments are described in Module A4.

1.4 Literature Cited

Ashenhurst, F., K. Strilchuk and P. Tran. 2000. A GIS tool to automate watershed information extraction. Final Report to the Ontario Ministry of Natural Resources by students at Sir Sandford Fleming College, GIS Applications Specialist program.

Frissell, C.A., W.J. Liss, C.E. Warren and M.D. Hurley. 1986. A hierarchical framework for stream habitat classification: viewing streams in a watershed context. Environmental Management, 10:199-214.

Higgins, J., M. Lammert, M. Bryer, M. DePhilip and D. Grossman. 1998. Freshwater conservation in the Great Lakes Basin: development and application of an aquatic community classification network. Report to The George Gund Foundation, Cleveland, Ohio and U.S. Environmental Protection Agency, Great Lakes National Program Office, by The Nature Conservancy, Chicago, Il..

Kilgour, B.W., G.L. Mackie, J. DiMaio, J. Smith, S. Staton, J. Parish, M. Nelson. 2000. Synthesis report for the Sydenham River Recovery Project. Prepared for the Ontario Ministry of Natural Resources and the St. Clair Region Conservation Authority.

Maxwell, J.R., C.J. Edwards, M.E. Jensen, S.J. Paustian, H. Parrott and D.M. Hill. 1995. A hierarchical framework of aquatic ecological units in North America (Nearctic Zone). United States Department of Agriculture, Forest Service, North Central Forest Experimental Station, General Technical Report NC-176.

Schloesser, I.J. 1991. Stream fish ecology: a landscape perspective. BioScience, 41:704-712.

Seelbach, P.W., M.J. Wiley, J.C. Kotanchik, and M. E. Baker. 1997. A Landscaped-Based Ecological Classification System for River Valley Segments in Lower Michigan (MI-VSEC Version 1.0). Michigan Department of Natural Resources, Fisheries Research Report No. 2036.

The Don Watershed Council, The Metropolitan Toronto and Region Conservation Authority and the Ontario Ministry of Natural Resources (TDWC et al.). 1997. Don watershed fish community and habitat management plan.

2.0 MODULE A1: DELINEATING VALLEY SEGMENTS

2.1 Background

Valley segment boundaries can be placed using the system of priorities defined by Seelbach et al. (1997) and Higgins et al. (1998). Delineation is based on natural features including hydrography, surficial geology and slope. Delineation can also be based on the presence of dams. Mapping requirements for delineation are described in Section 2.3.

2.2 Natural Segment Boundaries

2.2.1 Hydrography

Based on hydrography, valley segments can be placed on a stream where:

- The adjoining stream is of equal or higher order in the tributary.
- The upstream catchment area of the adjoining stream doubles the upstream catchment area of the mainstem[1].
- A stream/river joins with a wetland or lake (i.e., at the downstream end of inlets and the upstream end of outlets).

2.2.2 Surficial Geology

Based on surficial geology, additional valley segments can be placed on a stream where there is a change in surficial geology, and where surficial geology can be classified on the basis of texture or hydraulic conductivity as in Table 2.1:

- well drained
- moderately well drained
- poorly drained

[1] Here, the criterion that upstream catchment area has to double before a new node is selected, is based on the same concept that a doubling of Strahler stream orders must occur. It is recommended that research be conducted to determine if criteria (1) and (2) are redundant, and to determine if different criteria based on upstream catchment area, may give more useful delineation. Until this research is conducted, exclude this criteria

Table 2.1. Classification of soil textures.
Hydraulic conductivity's are taken from Freeze and Cherry (1979).

Classification	**Soil Texture**	**Saturated Hydraulic Conductivity, *K* (m/yr)**
Well Drained	• Gravel • Clean Sand • Fractured Metamorphic Rock • Permeable basalt	$1x10^{1}$ to $1x10^{7}$
Moderately Well Drained	• Silty sand • Silt	$1x10^{-2}$ to $1x10^{1}$
Poorly Drained	• Glacial Till • Clay • Shale • Sandstone • Unfractured metamorphic and igneous Rocks • Limestone and dolomite	$1x10^{-5}$ to $1x10^{-2}$

Valley segment nodes should be placed at the point where the surficial geology boundary crosses the stream.

2.2.3 Slope

Based on slope, additional valley segments boundaries can be placed on a stream where there is a change in longitudinal gradient classification, and where gradient is classified following Higgins et al. (1998) as:

- low (0 to 2 m/km)
- medium (2 to 20 m/km)
- high (>20 m/km)

Use ~~whichever procedure~~ most accurately and efficiently identifies the location on the landscape where slope changes between the three criteria[2] [2]

---

[2] Two methods have been proposed. The first develops contours from the DEM at predetermined intervals (1-10 m intervals) and then slopes were generated between the contours. Summarizing these slopes could identify the approximate location of nick points or changes in elevation. The second, derived slope from a direct interpretation from the DEM using a moving average as determined from a predetermined number of pixels (4 were proposed).

Using this approach, it is conceivable that nodes will be placed at both the top and bottom of transition areas like waterfalls. Transition areas will therefore become valley segments. It is anticipated that identification of transition areas as distinct valley segments will only assist in understanding the form and function of a riverine system, and is therefore considered a good result.

2.3 Unnatural Boundaries

2.3.1 Dams[3]

Based on the presence of dams, additional valley segment nodes can be placed on the stream.

2.4 Literature Cited

Freeze, R.A. and J.A. Cherry. 1979. Groundwater. Prentice Hall Inc., Englewood Cliffs, N.J.

Higgins, J., M. Lammert, M. Bryer, M. DePhilip and D. Grossman. 1998. Freshwater conservation in the Great Lakes Basin: development and application of an aquatic community classification network. Report to The George Gund Foundation, Cleveland, Ohio and U.S. Environmental Protection Agency, Great Lakes National Program Office, by The Nature Conservancy, Chicago, Il..

Seelbach, P.W., M.J. Wiley, J.C. Kotanchik, and M. E. Baker. 1997. A Landscaped-Based Ecological Classification System for River Valley Segments in Lower Michigan (MI-VSEC Version 1.0). Michigan Department of Natural Resources, Fisheries Research Report No. 2036.

[3] Re: Dams. There should be an option to use or not use dams as a criterion for delineation.

3.0 CHARACTERIZING VALLEY SEGMENTS

3.1 Background

In Section 1, a protocol for delineating valley segments on the basis of size, slope and surficial geology was presented. Here, a set of rules for characterizing these attributes (and others) for each valley segment is provided. For each attribute, an attempt is made to provide a protocol that would result in a number describing the attribute (i.e., as continuous variables). For example, stream size, position, and elevation can all be expressed numerically. Sites can be classified into various geoclimatic settings (Maxwell *et al.*, 1995), but such classification ignores the reality of gradients in environmental conditions. Numeric data provide more information about a location, while classification (if desired) can be imposed at some later time based on the numeric value. The only type of attributes that are not continuous variables are connectivity and zoogeography. Both are strictly categorical variables.

Some of the variables (e.g., size) will vary as one moves downstream in the valley segment, especially if the valley segment is quite large (e.g., 50 to 60 km long). To address this concern data will calculated for both the up and downstream nodes for the valley segment. In all but the last segment in a first order tributary (headwater) data will already be available for the upstream node, as this will also be the downstream end of the upstream segment). These data will be duplicated in column headings indicating that it represents the upstream node. For those segments where there are no upstream nodes the results for the upstream node will be blank.

3.2 Primary Variables

Primary landscape variables are attributes of a basin that have a controlling effect on the biochemical processes within the catchment. These variables are considered to be the basic template within which a stream ecosystem evolves. While we have a basic understanding of which attributes control these processes (Appendix I), it is recognized that others may also be important.

The proposed primary variables for characterizing valley segment conditions include:

1. **UPSTREAM DRAINAGE AREA** – because it reflects/determines flows and discharge, and is a strong determinant of fish community composition and species richness.

2. **POSITION** – which is a measure of how close to a large river a valley segment is. Being in close proximity to a large river has been associated with higher species richness.

3. **CONNECTIVITY** – which reflects connections of a valley segment to other types of water bodies including inland lakes, great lakes, salt water and wetlands.

4. **SURFICIAL GEOLOGY** – because it determines soil permeability and is associated with the likelihood of there being significant groundwater resources.

5. **SLOPE** – because it is associated with flow velocities, substrate particle size distributions and thus the kinds of animals that are found at a site.

6. **CLIMATE** – i.e., total annual precipitation and average summer temperature

3.2.1 Drainage Area

Drainage area is the most robust method of characterizing stream size. It is measured as the total area within the upslope catchment boundary. Estimation of drainage area requires delineation of the watershed boundary from a DEM or hard-copy map with elevations. This procedure has been automated as part of the Ontario Flow Assessment Techniques (OFAT).

In addition to simply measuring the drainage area we should also summarize the area in upstream catchment consisting of both lakes and wetlands.

3.2.2 Position

Here it is recommended that the upstream catchment area of the downstream segment be used as an indicator of position within the catchment. The greater the difference in area between the target valley segment and the downstream segment, the closer the segment is to a larger waterbody. This attribute is measured from the bottom of the segment.

Areas are measured from the bottom of the segment in either ha or ac depending on jurisdiction. Data are reported into a column called "*upstream catchment area of downstream segment*".

3.2.3 Connectivity

Connectivity affects the dispersal of organisms and the flow of water and nutrients among aquatic ecosystems. This attribute accounts for local zoogeography by distinguishing stream macrohabitats that are linked (uninterrupted) to the Great Lakes, to inland lakes (2 ha or 5 ac), wetlands, and those that have no obvious surface connection.

Connections to the attributes listed below will be recorded as a logical field (y/n). If there are connections to the listed attributes, then the distance (km) between the valley segment and the feature is determined.

Connectivity should be determined from hydrographic maps. For each site, determine whether the site is connected, in both the upstream or downstream directions, to the following habitat classes. Record the minimum distance of travel between any location within a segment and the object.

- Lakes: (2 ha or 5 ac)
- Wetlands (2 ha or 5 ac)
- Great Lakes
- Salt water environments

Finally there are unconnected valley segments, that occur when a valley segment is not connected to any other waterbody, including a river. For example, when a segment of a headwater stream flows into an ephemeral reach.

This protocol recognizes that intermittent streams cannot be identified from maps (digital or hard copy).

Wetlands that are adjacent to or interrupt a stream/river are conceivably equally important. Therefore, streams flowing through or adjacent to a wetland should be considered to be connected to a wetland.

If the valley segment is connected to any of the features listed below (in either up or downstream directions), *yes* will appear in the appropriate column, and a distance will be recorded. Where dams or other obstructions create a year-round barrier to fish migration, that obstruction disconnects the valley segment from habitats on the other side.

Connectivity will require an external database, listing dams and whether they are passable, partially passable (during certain times of the year) or impassable. In Ontario, NRVIS has a layer that identifies dams. Other barriers would have to be identified by local experts. In addition, designation of a barrier as being passable, partially passable and impassable would also have to be made by local experts. While connectivity is being assessed both upstream and downstream of a valley segment, it is anticipated that the dams database would be cross-referenced.

3.2.4 Surficial Geology

For each valley segment, determine the area of the upstream catchment that is covered in each of the four surficial geology classes (i.e., organic matter/peatland, well drained, moderately well drained, and poorly drained; Table 3.1).

Also determine the proximate surficial geology class, i.e., the material that the valley segment flows across. Classify as well drained, moderately well drained, poorly drained or organic matter as in Table 3.1.

Determine the surficial geology class from the immediate segment upslope of the valley segment being assessed.

The accuracy of these determination will vary with map resolution. There is currently no guidance on the map resolution that is required to reflect distributions of fishes[4].

3.2.5 Slope

Gradient of both the valley segment and the upstream catchment can be measured as the change in elevation divided by the measured length of the segment (or catchment).

The coefficient of variation of elevation along the river channel measures the variability in slope within a segment and should provide some insight into channel structure. It may provide an alternate tool for delineating valley segments in the future. Therefor we should also measure the coefficient of variation for each valley segment

3.2.6 Climate

Regardless of the map source, it is recommended that the following climate data be obtained:

- mean daily maximum temperature during summer (July)
- average annual precipitation

It is recognized that these are the minimum data requirement[5].

[4] Research Need: Determine map resolution needed to reflect distribution of fishes.
[5] Research Need: There are many climate variables not considered above. Therefore, research is needed to establish which climate variables influence fish distributions.

3.3 Modifier Variables

3.3.1 Landcover

Even landcover data that are summarized in the coarsest fashion (i.e., 4 categories) provide support for the premise that landcover influences the biophysical makeup of valley segments (Appendix I). Landcover data have been summarized in Ontario (all) and the U.S., with high resolution, but with variations in data standards, data legends (rules) and map scales.

The recommended approach for summarizing landcover data will involve extraction of the existing information (regardless of the source) by extracting the area of upstream catchment that is covered by each of the existing land cover types. There is no clear understanding of the type of landcover that has the most influence on valley segment attributes[6]. There are therefore two reasons to extract all of the landcover data from an available source. First, it does not assume which landcover classes are the most important. This will allow for research to be done examining the importance of the finer classifications. Second, the data can be agglomerated at some later time.

To accommodate the differences among databases, it is recommended that data be agglomerated to the following minimum landcover classes as initially proposed by Anderson et al. (1976) (Table 4.1). In Ontario, LANDSAT satellite imagery is interpreted into 28 landcover classifications. Table 4.2 recommends how those classifications can be agglomerated to fit with Anderson et al.'s (1976) original landcover classification.

To accomplish this, 3 steps are required. First, the database should have a default list of fine-resolution landcover classes and their definitions. Second, each class should be assigned to a major landcover class. Third, the database should have provision for new classes that do not match those in the default categories.

Table 4.1. Major land cover/use classes for remote sensor data from Anderson et al. (1976), and Ontario landcover classification from White (2000).

Major Landcover Group	Ontario Landcover classification
Urban or built up land	SETTLEMENT AND DEVELOPED LAND: Clearings for human settlement and economic activity; major transportation routes.
Agricultural	CROPLAND: Row crops mapped in Southern Ontario; hay or open soil in areas of agricultural land use.
Rangeland	PASTURE AND ABANDONED FIELDS: Open grassland with sparse shrubs mapped in

[6] Research Need: Determine the association of landcover type with valley segment characteristics.

[7] Need to determine how accurately CV can be measured using various map scales

Major Landcover Group	Ontario Landcover classification
	agricultural areas of Southern Ontario; includes orchard lands. ALVAR: Homogeneous areas of dry grassland growing on thin soils over a limestone substrate, mapped only where they occur in clusters in the central and eastern portions of Southern Ontario
Forest	DENSE DECIDUOUS FOREST: Largely continuous forest canopy composed of at least 80 percent of deciduous species; includes deciduous shrub cover on old burns and alder thicket swamps in the Hudson Bay-James Bay Lowlands. DENSE CONIFEROUS FOREST: Largely continuous forest canopy composed at least 80 percent of coniferous species; includes dense conifer swamp in the Hudson Bay-James Bay Lowlands. CONIFEROUS PLANTATION: Mature conifer plantations, mostly pine, occurring in evenly spaced rows, mainly in Southern Ontario. This class does not include artificially regenerated cutovers or burns in Northern Ontario. MIXED FOREST, MAINLY DECIDUOUS: Largely continuous forest canopy composed of coniferous and deciduous species, with deciduous species dominant (i.e., comprising more than 50 percent of the canopy). MIXED FOREST, MAINLY CONIFEROUS: Largely continuous forest canopy composed of coniferous and deciduous species, with coniferous species dominant (i.e., comprising more than 50 percent of the canopy). SPARSE CONIFEROUS FOREST: Patchy or sparse forest canopy (i.e. approximately 30 to 40 percent canopy closure) composed approximately 80 percent of coniferous species.

Major Landcover Group	Ontario Landcover classification
	SPARSE DECIDUOUS FOREST: Patchy or sparse forest canopy (i.e., approximately 30 to 40 percent canopy closure) composed approximately 80 percent of deciduous species. RECENT CUTOVERS: Forest clear-cuts estimated at less than 10 years of age. RECENT BURNS: Forest burns estimated at less than 10 years of age. OLD CUTS AND BURNS: Forest clear-cuts and burns estimated at more than 10 years of age.
Water	WATER: All waterbodies, both deep/clear and shallow/sedimented.
Wetland	COASTAL MUDFLATS: Unvegetated coastal areas of the Hudson Bay-James Bay Lowlands, partly submerged at high tide. INTERTIDAL MARSH: Coastal marshes of the Hudson Bay-James Bay Lowland lying between the coastal mudflats and the supertidal zone. SUPERTIDAL MARSH: Coastal marshes of the Hudson Bay-James Bay Lowland lying inland of the Coastal Mudflat and Intertidal Marsh classes and subject to only exceptionally high tides. FRESHWATER COASTAL MARSH/ INLAND MARSH: Coastal marshes of the Hudson Bay-James Bay Lowland lying beyond the area of saltwater influence; marshes occurring along lakeshores; Southern Ontario inland marshes characterized by a range of moisture conditions: seasonal marshes, flooded in spring but often dry by fall, that may appear flooded more deeply

Major Landcover Group	Ontario Landcover classification
Barren Land	MINE TAILINGS, QUARRIES, AND BEDROCK OUTCROPS: Clearings for mining activity scattered in all parts of the province; aggregate quarries occurring mainly in Southern Ontario; bedrock outcrops.
Tundra	OPEN BOG: Non-treed bog that may have a partial cover of stunted trees occurring generally in the province but most extensively in the Hudson Bay-James Bay Lowlands, where it also includes lichen-rich peat plateau. TREED BOG: Bog with a low to high density of tree cover. There is expected to be some degree of overlap between densely treed bog and sparse conifer forest in more northerly parts of the province and especially in the Hudson Bay-James Bay Lowlands. TUNDRA HEATH: Areas of dense ericaceous vegetation occurring on better-drained areas only in the Hudson Bay coastal zone.

3.4 Literature Cited

Anderson, J.R., E.E. Hardy, J.T. Roach and R.E. Witmer. 1976. A land use and land cover classification system for use with remote sensor data. Geological Survey Professional Paper 964. U.S. Government Printing Office, Washington, D.C. 28 pp.

Maxwell, J.R., C.J. Edwards, M.E. Jensen, S.J. Paustian, H. Parrott and D.M. Hill. 1995. A hierarchical framework of aquatic ecological units in North America (Nearctic Zone). United States Department of Agriculture, Forest Service, North Central Forest Experimental Station, General Technical Report NC-176.

White, D. 2000. Provincial landcover database derived from LANDSAT satellite imagery. Ontario Ministry of Natural Resources.

4.0 MODULE A4: CLASSIFYING VALLEY SEGMENTS

4.1 Background

The objective of this module is to describe a process of valley segment classification. Classification would assist (1) in the identification of similar (like) valley segments, (2) in generalizing within classes, and (3) in comparison across classes. In the Great Lakes Basin, classification schemes have been used or proposed for a variety of reasons, including to:

- identify rare environments (Higgins et al. 1998)
- identify sensitive habitats in need of protection (i.e., cold-water valley segments, Higgins et al., 1998)
- stratify study areas in subwatershed studies (CVC et al. 1997; Cover, pers. comm.)
- stratify study areas for research needs
    - development of habitat suitability indices for aquatic fauna (Stanfield et al., 2000)
    - evaluate the effect of barriers on fish communities
    - quantify production of migratory salmonids (Zorn et al., 2000)
    - describe background variations in fish and benthic communities in the Moose River Basin (Kilgour et al., 2000).

For each of these studies/proposals, the classification schemes relied on a relatively common suite of attributes (i.e., stream order, slope, surficial geology, etc.,). The proponents of these projects agree that a common classification system would be beneficial, in that results from one study area could be applied across study areas. But, no common scheme has thus far been proposed. Therefore, this module describes one proposal that has been developed through a peer-review process, and that standardizes a classification scheme based on stream order, slope and surficial geology.

It is recognized that this proposed classification scheme is coarse and may not provide the detail required for all managers and researchers.

4.2 Recommended Classification Scheme

The intent of this recommended scheme is to provide a default classification. Where classification is automated, it is recommended that databases be written in such a way as to be flexible in allowing for more or less variables, more or less classes of each variable, and varying ranges within a class. Whether this is done in a dichotomous key format or a series of drop-down windows, will depend on the software being used. This flexibility will ensure that data can be examined in a number of ways.

The recommended default classification is given below.

On the basis of upstream catchment area:

- small (upstream catchment area < 10 km^2),
- medium (10 to 200 km^2) and
- large (>200 km^2).

On the basis of surficial geology:

- highly porous (>25% coarse soils in the upstream drainage area)
- moderately porous (10 to 25% coarse soils in the upstream drainage area)
- poor porosity (< 10% coarse soils in the upstream drainage area)

On the basis of slope:

- low slope (0 to 3 m/km)
- moderate slope (4 to 10 m/km)
- high slope (11 to 50 m/km)

4.3 Literature Cited

Credit Valley Conservation, Environmental Water Resource Group, Ortech Corporation, Terraqua Investigations Ltd. and Water Systems Analysts (CVC et al.). 1997. Caledon Creek and Credit River subwatershed study - characterization report. Prepared for the Region of Peel and Town of Caledon.

Environment Canada, Ontario Ministry of Natural Resources and Ontario Ministry of Environment and Energy (EC, OMNR and OMOEE). 1998. A framework for guiding habitat rehabilitation in Great Lakes Areas of Concern.

Higgins, J., M. Lammert, M. Bryer, M. DePhilip and D. Grossman. 1998. Freshwater conservation in the Great Lakes Basin: development and application of an aquatic community classification network. Report to The George Gund Foundation, Cleveland, Ohio and U.S. Environmental Protection Agency, Great Lakes national Program Office, by The Nature Conservancy, Chicago, Il..

Kilgour, B.W., D.G. Dixon, R.C. Bailey and T.B. Reynoldson. 2000. Development of a reference condition approach (RCA) to assess fish and benthic communities and in-stream habitat attributes in the Moose River Basin. Reported submitted to the Forest Ecosystem Science Co-Operative Inc.

Stanfield, L., M. Jones, M. Stoneman, B. Kilgour, J. Parish and G. Wichert. 1999. Stream Assessment Protocol for southern Ontario. V 3.1. Ministry of Natural Resources training manual.

The Don Watershed Council, The Metropolitan Toronto and Region Conservation Authority and the Ontario Ministry of Natural Resources (TDWC et al.). 1997. Don watershed fish community and habitat management plan.

Zorn, T, P. Seelbach, M. Wiley, L. Stanfield. 2000. Lake Huron Aquatic Habitat, Geographic Information System

5.0 FRAMEWORK FOR USING VALLEY SEGMENTS IN AQUATIC ENVIRONMENT ASSESSMENT

6.1 Background

The objective of this section is to describe how valley segments can be used in the three basic stages of aquatic environment assessment, i.e., baseline characterization, assessment of current conditions relative to a standard, and prediction of future condition under assumed scenarios of human development. Focus here is on assessments within subwatersheds, because planning and implementation of environmental programs typically take place at that scale.

6.2 Baseline Characterization

Baseline studies are conducted for three reasons:

- as a general inventory
- to characterize conditions prior to human intervention
- to understand the normal causes of animal distributions

Baseline studies are typically conducted in areas that lack a significant amount of human development. The actual environmental condition is characterized using surveys at the site scale. However, since it is not possible to characterize all locations (reaches/sites) within a watershed, some means of reducing the number of sampling locations must be used. Therefore, the following approach can be used to reduce the number of sites for sampling.

- Define the limits of the study area (subwatershed?).
- Delineate valley segments within the study area.
- Characterize valley segments
- Classify those valley segments and determine the number of segments of each valley segment type.
- If there are multiple valley segments for each segment type, randomly select one or more segments for study.
- Within selected valley segments, randomly select one or more sites at which field surveys will be conducted. In practice, sites may be selected on the basis of accessibility.
- Conduct field sampling at those sites.
- Analyze the data and establish a relationship between environmental conditions at the reach or site scale and landscape attributes (i.e., a landscape-based biophysical model).
- Alternatively, there may already exist, landscape models that can be used to predict the biophysical conditions within valley segments. Landscape models may either be empirically based, or mechanistic (Appendix 1).
- Use landscape models to predict biophysical condition elsewhere in the system.

6.3 Assessment of Current Conditions Relative to Standards

Environmental assessments are often conducted to establish current status of a subwatershed relative to standards. Standards are usually numeric values of indices of biological composition derived from locations that are representative of an acceptable condition (Section 6.4). As with baseline studies, there is no advantage in sampling every possible location within a subwatershed. Therefore, some means of stratfiying the subwatershed would be useful for minimizing the number of sampling locations. A recommended procedure for site selection for assessment purposes is:

- Define the physical limits of the study area (subwatershed?).
- Delineate valley segments within the study area.
- Characterize valley segments.
- Classify valley segments and determine the number of segments of each valley-segment type.
- If there are land uses that could potentially alter biological potential, further stratify valley segments on that basis, and sample each strata (if possible). Where land uses alter the natural landscape, it is unlikely that multiple segment types would have similar land cover in the upstream catchment. In those cases, it may be necessary to sample each valley segment in order to establish the current condition of each segment.
- If there are multiple valley segments for each segment type, randomly select one or more segments for study. If all valley segments are unique, then each segment should be sampled.
- Within selected valley segments, randomly select one or more sites at which field surveys will be conducted. In practice, sites are often selected on the basis of accessibility.
- Conduct field sampling at those sites.
- Compare observed site conditions with expected site conditions established at reference streams (Section 6.4).

6.4 Developing Standards in Least-Impaired Systems

Numeric criteria (standards) are typically used to evaluate the biological health of test sites. Where specific criteria are lacking (usually the case), the use of regional-reference locations provides an objective way to define the limits of acceptable and unacceptable conditions (Hughes, 1995). For example, biologists with the State of Ohio sample fish and benthos from hundreds of relatively unimpacted streams to establish the biological characteristics of healthy fish and invertebrate communities (Yoder and Rankin, 1995). Various numerical indices are calculated for each of these regional-reference locations and the observed distribution of values for each index is often used to define the "normal range" of variation for unimpacted communities (or acceptable conditions; Kilgour et al., 1998; Figure 6.1). Frequently, the lower 5^{th} or upper 95^{th} percentile of the distribution (depending on the direction of the index) is used as a biological criterion. Thus, any test location falling outside of the range of 95% of the reference locations is considered potentially degraded.

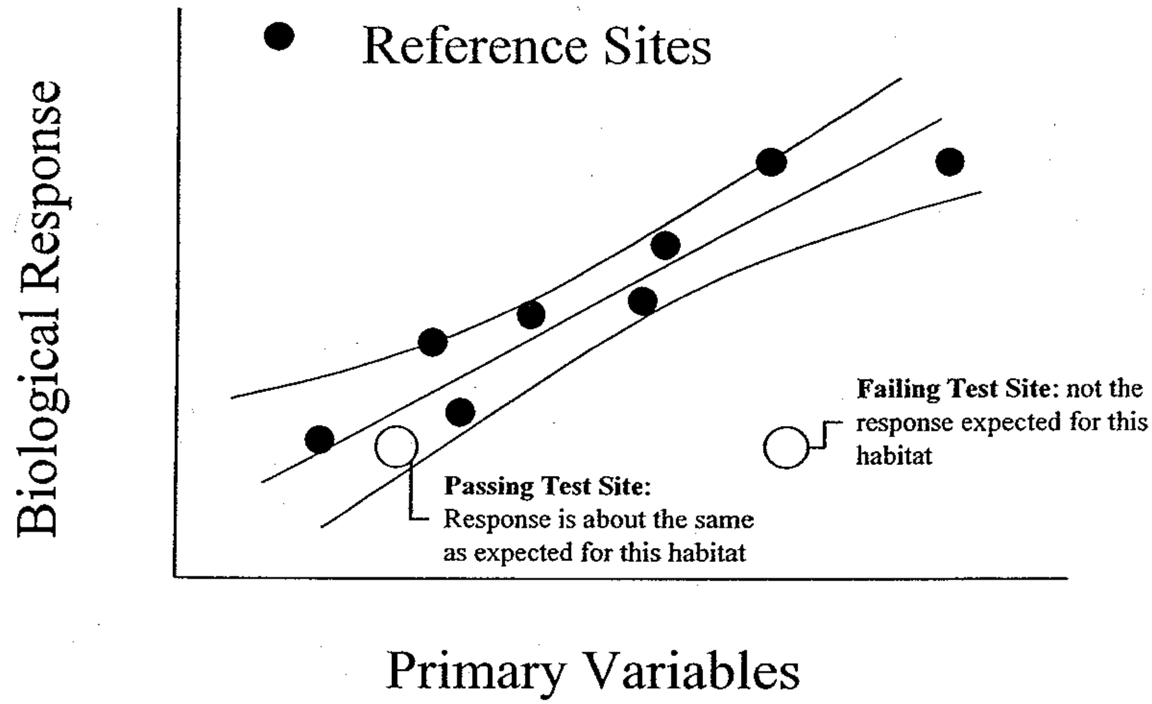


Figure 6.1. Relationship between Biological Response and Primary Habitat Possible relationship between a biological response (Compliance Indicator) and a Primary habitat variable that can be used to predict the expected biological condition.

Defining the reference condition is challenging. For 1st and 2nd order streams in southern Ontario, Barton (1996) defined reference streams as those with > 50% forest cover in the upstream catchment. This definition is satisfactory because there are always small catchments that are still highly forested. However, this definition cannot be applied to larger 3rd and 4th order rivers because rivers of that size nearly always have significant development in the upstream catchment (agricultural or urban).

The objective of this section is to describe a procedure for developing reference-condition models that can be used for assessment. Much of the following material is from Hughes (1995). Hughes (1995) cautions that the selection of reference sites is not a trivial undertaking, and that it can be a more time-consuming process than the actual sampling. The focus here is on deriving numeric biological criteria for surveys of aquatic fauna. However, the principles applied here can also be applied to physical and chemical attributes as well.

- Define the spatial limits of the study area,
- Define the waterbody types, sizes and classes of interest (i.e., low-order streams, high-order rivers),
- Identify candidate reference catchments. Sites can be rejected for the following reasons:

- known pollutant discharges
- landfills
- mines
- intense land use in the upstream catchment (i.e., logging, fire, agriculture, urbanization, oil fields, poultry farms, fish hatcheries,
- fish stocking practices (especially with exotic species)
- proximity of major roadways

- Conduct Field Reconnaissance to screen potential sites. Hughes (1995) recommends that field reconnaissance be used to ensure that reference sites have the following characteristics:
  - roads distant
  - riparian vegetation extensive and old
  - riparian structure complex
  - bank complex
  - channel complex
  - bank modification minimal
  - habitat structure complex
  - chemical stressors minimal
  - channel/flow manipulation minimal
  - odors, films, scums and slicks minimal
  - pipes, drains, ditches and tiles absent
  - wildlife and benthos evident
  - human and livestock activity minimal

- Determine the number of reference sites needed. There is no adequate guidance on the number of appropriate reference sites. Hughes (1995) recognized 3 sites as minimum for any class and size of stream within any ecoregion. Walters et al. (1988) recommended 1 reference site for every 4 to 5 test sites, whereas David et al. (1998) recommended 10 reference sites for every 1 test sites.

  In part, the number of reference sites required depends on the underlying model that will be used to partition natural variation from anthropogenically induced variation. If numeric criteria are calculated for like groups of sites, then a large number of sites per group (about 10 to 20) will be required to provide adequate statistical power for detecting effects when they occur (Kilgour et al., 1998). If, however, a linear regression model is used (Figure 6.1), a total of 30+ sites across all stream types may be adequate, depending on the number of variables needed to account for variation in the biological indices. The larger the number of reference sites, the greater the potential for detecting effects when they occur. If too few sites are used, then effects must be very large before they can be detected (Kilgour et al., 1998).

- Sample reference sites, and generate numeric biological criteria. As before, criteria are represented by the background variability in indices of composition. Where background variation can be accounted for by natural landscape features, those models should be developed (e.g., Figure 6.1) to optimize the ability to detect effects when they occur. A variety of approaches for building reference condition models exists including the use of discriminant models (Reynoldson et al., 1997), regression models (Bailey et al., 1998; Wiley et al., 2000) and combinations of the two (Kilgour, 2000).

6.5 Prediction of Future Conditions

Impact and risk assessments have the general task of determining the likely outcome of a human intervention (disturbance) on the biophysical condition of a study area. These predictions can be done one of two ways: (1) using mechanistic models; and (2) using empirical models. Mechanistic models tend to be complex, requiring data on a number of physical and chemical processes that may or may not be available for the study area. Empirical models are often simpler, relying on observational data from other study areas. However, they apply only to the study area and conditions for which they were calibrated. Caution should be used when applying empirical models to new study areas or new conditions.

In this section, a procedure for developing empirical models is provided. Assuming that the objective is to develop a model that predicts effects caused by changes in land cover/use, the procedure is:

- Define the limits of the study area (subwatershed?).
- Delineate valley segments within the study area.
- Characterize valley segments.
- Overlay land use/cover to further stratify the study area.
- Classify those valley segments and determine the number of segments of each valley-segment type. The classification will be based on primary, diagnostic and modifier variables.
- If there are multiple valley segments for each segment type, randomly select one or more segments for study.
- Within selected valley segments, randomly select one or more sites at which field surveys will be conducted. In practice, sites may be selected on the basis of accessibility.
- Conduct field sampling at those sites.
- Using data from sites considered acceptable (Section 6.4), develop a biophysical model predicting biological conditions, given primary variables (Figure 6.1).
- Determine the deviation from expected for sites not considered acceptable (Figure 6.1)

- Relate deviations to modifier variables that reflect the degree of development within the upstream catchment (Figure 6.2). This is the model that can then be used to predict change in biological condition. Variables such as % urban area, % agricultural, etc., would be good candidate variables for predictive models.

Variables that reflect the degree of human activity in a catchment are "Modifying" variables because they reflect the degree of human modification of the natural biophysical environment. Ideally, measures of human modification would reflect the full history of development within a catchment. However, these histories are not easily quantified or standardized. Therefore, human modification of the landscape is typically measured using current conditions with variables such as urban area, percent forest cover, percent rural area, and human population density (Lenat and Crawford, 1994; Townsend et al., 1997; Kennen, 1999). The major assumption is that these conditions reasonably reflect the degree of human disturbance over time.

To establish the degree of impairment in biological indicators, models relating deviation from the expected biological condition to modifier variables are often used. Models that relate the existing biological condition (not the deviation from expected) to modifier variables cannot be confidently used to predict the future consequence of human activity because they do not take into account the underlying effect of primary variables. For example, Barton et al. (1985) demonstrated a strong relationship between percent forest cover and fish communities in southern Ontario. However, their model cannot be used to infer the effect of reducing forest cover in a catchment because forest cover may have been autocorrelated with underlying physiography (Portt et al., 1989). A more useful approach would relate forest cover to fish communities within each physiographic region or other stratifying variable (e.g., climatic zone).

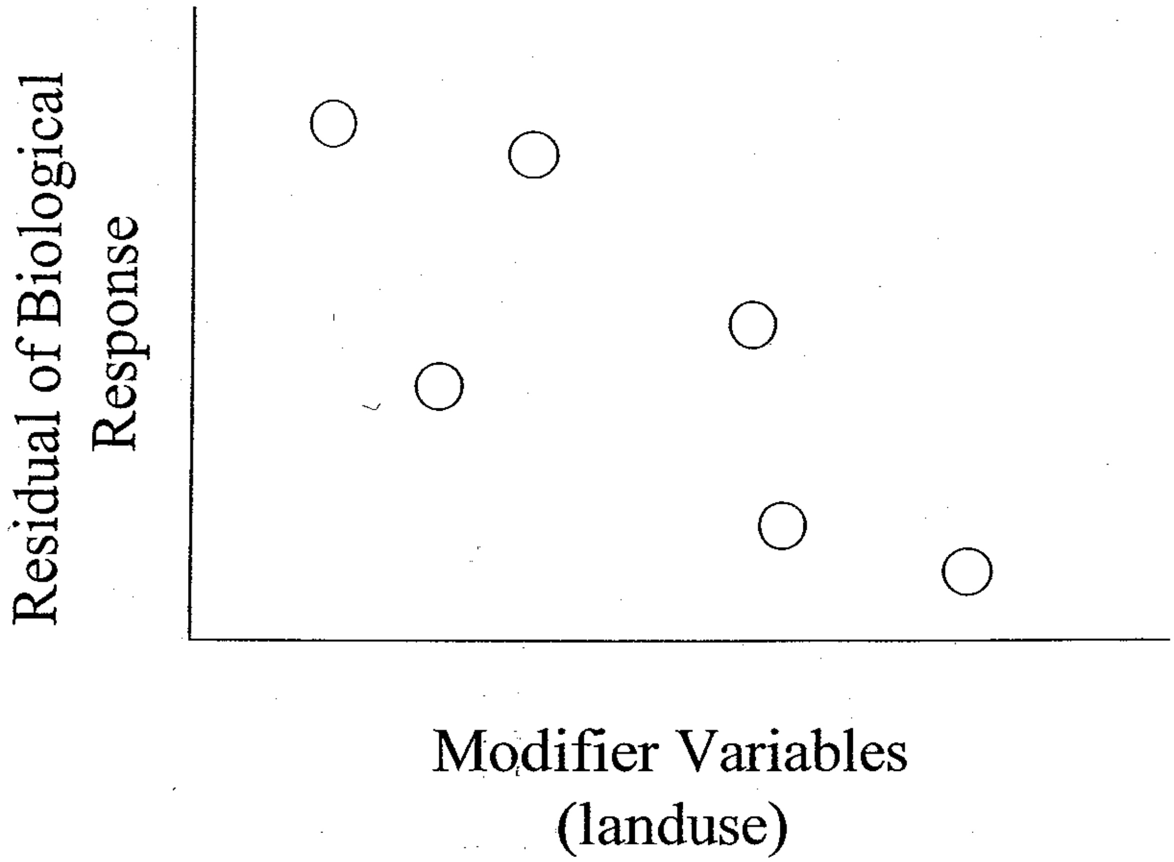


Figure 6.2. Relationship between Biological Deviation and Modifier Variables Relationship between biological deviation and some set of modifier variables. Models like this can be used to predict future biological impact given changes in land cover/use.

5.1 Literature Cited

Bailey, R.C., M.G. Kennedy, M.Z. Dervish and R.M. Taylor. 1998. Biological assessment of freshwater ecosystems using a reference condition approach: comparing predicted and actual benthic invertebrate communities in Yukon streams. Freshwater Biology, 39:765-774.

Barton, D.R. 1996. The use of percent model affinity to assess the effects of agriculture on benthic invertebrate communities in headwater streams of southern Ontario, Canada. Freshwater Biology, 36:397-410.

Barton, D.R., W.D. Taylor and R.M. Biette. 1985. Dimensions of riparian buffer strips required to maintain trout habitat in southern Ontario streams. North American Journal of Fisheries Management, 5:364-378.

David, S.M., K.M. Somers, R.A. Reid, R.J. Hall and R.E. Girard. 1998. Sampling protocols for the rapid bioassessment of streams and lakes using benthic macroinvertebrates. 2nd edition.

Dorset Environmental Sciences Centre, Aquatic Science Section, Ontario Ministry of the Environment, Dorset, Ontario.

Hughes, R.M. 1995. Defining acceptable biological status by comparing with reference conditions. In W.S. Davis and T.P. Simon (eds), Biological Assessment and Criteria, Tools for Water Resource Planning and Decision Making. Lewis Publishers, Boca Raton.

Kennen, J.G. 1999. Relation of macroinvertebrate community impairment to catchment characteristics in New Jersey streams. Journal of the American Water Resources Association, 35:939-955.

Kilgour, B.W. 2000. Assessing Impairment of Benthic Macroinvertebrates in the Humber River Watershed Using a Reference-Condition-Approach, 1999. Report prepared by Water Systems Analysts Inc. for the Toronto and Region Conservation Authority.

Kilgour, B.W., K.M. Somers, and D.E. Matthews. 1998. Using the normal range as a criterion for ecological significance in environmental monitoring and assessment. Écoscience, 5:542-550.

Lenat, D.R. and J.K.Crawford. 1994. Effects of land use on water quality and aquatic biota of three North Carolina Piedmont streams. Hydrobiologia, 294:185-199.

Portt, C.B., S.W. King and H.B.N. Hynes. 1989. A review and evaluation of stream habitat classification systems with recommendations for the development of a system for use in southern Ontario. Prepared by C. Portt and Associates for the Ontario Ministry of Natural Resources, Fish Community and Habitat Section, Fisheries Branch, Toronto.

Reynoldson, T.B., R.H. Norris, V.H. Resh, K.E. Day and D.M. Rosenberg. 1997. The reference condition: a comparison of multimetric and multivariate approaches to assess water-quality impairment using benthic macroinvertebrates. Journal of the North American Benthological Society, 16:833-852.

Townsend, C.R., C.J. Arbuckle, T.A. Crowl and M.R. Scarsbrook. 1997. The relationship between land use and physicochemistry, food resources and macroinvertebrate communities in tributaries of the Taieri River, New Zealand: a hierarchically scaled approach. Freshwater Biology, 37:177-191.

Walters, C.J., J.S. Collie and T. Webb. 1988. Experimental designs for estimating transient responses to management disturbances. Canadian Journal of Fisheries and Aquatic Sciences, 45:530-538.

Wiley, M.J., P.W. Seelbach and P. Rentschler. 2000. Feasibility of explicitly modeling ecological reference conditions for Michigan streams using landscape data. Final Report to U.S. E.P.A., Water Quality Criteria Program.

Yoder, C.O., and E.T. Rankin. 1995. Biological criteria program development and implementation in Ohio. In, W.S. Davis and T.P. Simon (eds), Biological Assessment and Criteria, Tools for Water Resource Planning and Decision Making. Lewis Publishers, Boca Raton, FL.

APPENDIX 1

CLASSIFICATION OF RIVERINE HABITATS USING TEMPERATURE AND NUTRIENT STATUS

BACKGROUND

In impact and risk assessment, and in subwatershed studies where the emphasis is on determining the likely impact of human activity on stream systems, it may be important to predict thermal and nutrient status after development. Other than physical habitat features, thermal and nutrient status are two attributes of stream systems that have significant influence on the kinds of animals found. Therefore, where there are significant changes in thermal and nutrient status, significant changes in the fauna of a system can be expected. A variety of models are available for predicting changes in thermal and nutrient status.

Modeling may not need to be excessively precise. In aquatic environments, major differences in faunal assemblages are associated with cold, cool and warm thermal regimes, and with oligotrophic, mesotrophic and eutrophic systems (Table I-1). Stream and river sites that are classified as cold water can support a cold-water fauna including brook and brown trout, and sculpins. Cool-water sites will support rainbow trout, while warm-water sites tend not to support any salmonids during peak summer temperatures, but will support a variety of cyprinids, depending on zoogreography. Eutrophic systems tend to have assemblages of fish and other animals that can tolerate oxygen sags at night (associated with decaying and respiring vegetation). Therefore, as long as a model can classify a site into basic categories, it should be reasonably useful at identifying biologically significant changes. At times however, more detailed classification may be required if the biological endpoint is something other than a gross change in species composition.

Models can either be process based, empirical, or mixtures. Process-based models replicate the key physical and chemical processes in a system. Empirical models reflect the relationships between response variables (i.e., those things we hope to predict effects in) and predictor variables. The objective of this Appendix is to provide a brief overview of process and empirical models that can be used to predict physico-chemical characteristics of streams.

PROCESS MODELS

In Southern Ontario, at least three different stream temperature modeling exercises have taken place (Meisner, 1990; LeBlanc et al., 1997; and Weatherbee et al., 1999). Typically, these models require some estimate of groundwater discharge to a site, as well as canopy and a variety of climatological data (sun angle, air temperature). Water temperatures are usually calculated hourly in the process-driven models, so day-to-night fluctuations can be predicted. None of the models for southern Ontario have been incorporated into user-friendly interfaces.

The more detailed, process-driven models have been reviewed extensively. Readers are referred to Marchant and Heathcote (1998) for a southern-Ontario perspective and to Donigian and Huber (1991) for a U.S. perspective. A list of commonly used process models for modeling stream temperatures and nutrient status are listed in Table I-2. QUAL2E, AGNPS and WASP5 are suitable for modeling nutrients and temperatures in branched river systems. Of these three, QUAL2E has the simplest interface, in that it uses a single plain text input file (under the Windows version). It is also extensively used. HEC-2, HEC-RAS and SWMM are free, validated, and relatively simple to interface.

Table I.1. Classification of riverine environments on the basis of thermal and nutrient status.

Source	Class	Criteria
Stoneman and Jones (1996)	Cold-Water Fauna	Maximum summer temperature < 14°C
	Cool-Water Fauna	Maximum summer temperature between 14°C and 23°C
	Warm-Water Fauna	Maximum summer temperature > 23°C
Wehrly et al. (1997)	Cold-Water Fauna	Mean July temperature < 19°C
	Cool-Water Fauna	Mean July temperatures between 19 and 22°C
	Warm-Water Fauna	Mean July temperature > 22°C
MOE (1994)	Eutrophic (taxa tolerant of high nutrients and low dissolved oxygen levels present)	Total Phosphorus > 30 μg/L
Seelbach et al. (1997)	Oligotrophic (taxa sensitive to nutrients and low dissolved oxygen)	< 15 g SRP/L and < 100 g N/L[7]
	Mesotrophic	15 to 30 g SRP/L and 100 to 700 g N/L
	Eutrophic (taxa tolerant of high nutrients and low dissolved oxygen levels present)	>30 g SRP/L and/or >700 g N/L

In Southern Ontario, at least three different stream temperature modeling exercises have taken place (Meisner, 1990; LeBlanc et al., 1997; and Weatherbee et al., 1999). Typically, these models require some estimate of groundwater discharge to a site, as well as canopy and a variety of climatological data (sun angle, temperature). Temperatures are usually calculated hourly in the process-driven models, so day-to-night fluctuations can be predicted. None of the models for southern Ontario have been incorporated into user-friendly interfaces.

[7] N is NO_2+NO_3

The more detailed, process-driven models have been reviewed extensively. Readers are referred to Marchant and Heathcote (1998) for a southern-Ontario perspective and to Donigian and Huber (1991) for a U.S. perspective.

Table I-2. List of commonly used mechanistic models for predicting physical and chemical attributes at stream sites.

Subject	Name	Type/Sys Req	Source	Characteristics	Availability
Water Quality	QUAL2E Sep 95	PC-WIN DNS	US EPA Watershed Modeling Section 401 M Street, S.W. Washington, DC 20460	Simulates the major reactions of nutrient cycles, algal production, benthic and carbonaceous demand, atmospheric reaeration and their effects on the dissolved oxygen balance; includes a sophisticated heat budget calculation.	Free - download at http://www.epa.gov
	AGNPS 98	PC-WIN DPM	ARS National Sedimentation Laboratory Oxford, Mississippi	Simulates sediment and nutrient transport from agricultural watersheds; the basic components of the model are hydrology, erosion, sediment transport, nutrient transport, chemical oxygen demand and temperature simulation.	Free - download at http://www.sedlab.olemis.edu/AGNPS98.html
	CORMIX 3.2 Dec 96	PC-DOS DPM	US EPA-CEAM960 College Stn Rd Athens, GA 30605-2700	Analysis, prediction and design of aqueous toxic conventional pollutant discharge into diverse water bodies. Includes temperature simulation.	Available at no charge via CEAM disk exchange or download at http://ftp.epa.gov/epa_ceam/wwwhtml
	SWAT 98.1	PC-WIN95	Grassland, Soil and Water Research Laboratory Temple, TX.	SWAT is a continuous time model (daily time step) that is required to look at long-term impacts of management (i.e., reservoir sedimentation over 50-100 years) and also timing of agricultural practices within a year (i.e., crop rotations, planting and harvest dates, irrigation, fertilizer, and pesticide application rates and timing).	Free – download at http://www.brc.tamus.edu/swat/swat/swatdoc.html
	WASP5 93	PC-DOS DPM	US EPA	Temperature and contaminant fate and transport model in 1-D, 2-D, 3-D.	Available at no charge via CEAM disk exchange or download at http://ftp.epa.gov/epa_ceam/wwwhtml
	CH2D	2-D finite difference flow model	WES	Uses primitive shallow water eqn.	Public domain

Notes: 1. DPM = Digital Process Model
2. DNS = Deterministic Numerical Simulation 3. General approach: find common, validated, inexpensive modelling tools running on a PC.

EMPIRICAL MODELS

Empirical models have two major advantages over process models in that they:

- Are usually simpler, and require less input before giving a result, and,
- Often do a better job of predicting the parameter of interest.

However, there are two major disadvantages of empirical models:

- They are not valid outside of the geographic area in which they were calibrated, and
- They cannot be easily modified to reflect changes in underlying processes. For example, if a relationship between %urban area in a catchment and water quality is developed, that model may not be appropriate for predicting future changes in water quality if the runoff from urban areas is managed in a new manner.

In the Great Lakes drainage area, there are no generally recognized empirical models for deriving estimates of temperatures and nutrient status. However, there have been a few models derived for specific applications. Those are described below in the event that the approaches used may have application elsewhere.

Temperature Modelling

Wehrly et al. (1997) used empirical relationships to predict the average of the three weekly maximum and minimum temperatures. These were called the maximum and minimum July temperatures. The average of the minimum and maximum temperatures was used to estimate the mean July temperature, while difference between them was taken to estimate the temperature fluctuation. Maximum and minimum July temperatures can be predicted from stream width, cross-sectional area, forest cover in the catchment, gradient of the reach, maximum air temperature and groundwater. Groundwater contributions were estimated using Darcy's Law which estimated groundwater velocity to a site.

Seelbach et al. (1997) demonstrated that estimated mean July temperatures and temperature fluctuation are related to baseflow yield (i.e., the 90 percent exceedance flow) and drainage area (Seelbach et al., 1997; Figure 4.7). These graphs can therefore be used in Michigan to classify the thermal condition of sites. Such extrapolations should be used cautiously because the relationships between thermal conditions and baseflow yield and drainage area are weaker than the more detailed models derived in part using the estimate of groundwater velocity (P. Seelbach, pers. comm.). What these graphs do illustrate is that for streams of equal size, the stream with the greatest baseflow will have lower temperatures. Local groundwater contributions can override this general observation.

To estimate the potential input of ground water to a stream network, Wehrly et al. (1997) developed a spatial model (GIS map layer) that predicts maximum potential ground water velocity. That map is based on Darcy's Law which states that ground water velocity is proportional to local hydraulic head (or slope) times the hydraulic conductivity of the underlying materials (Dunne and Leopold, 1978). Slope was calculated from a DEM (1:250,000). Conductivity values for surficial geology classes were taken from published tables (Rahn, 1980; Todd, 1980; Bedient and Huber, 1988). Surficial geology data were digital at 1:500,000. This representation of the potential ground water velocity was summarized as the mean velocity within a catchment.

The use of the Darcy Flux has been proposed because of the general difficulty in obtaining estimates of ground-water discharge (m^3/s) to a site. However, the use of the Darcy Flux as a surrogate for stream temperature is still in the developmental stage, and supporting documents from colleagues in Michigan were unavailable at the time of printing this report.

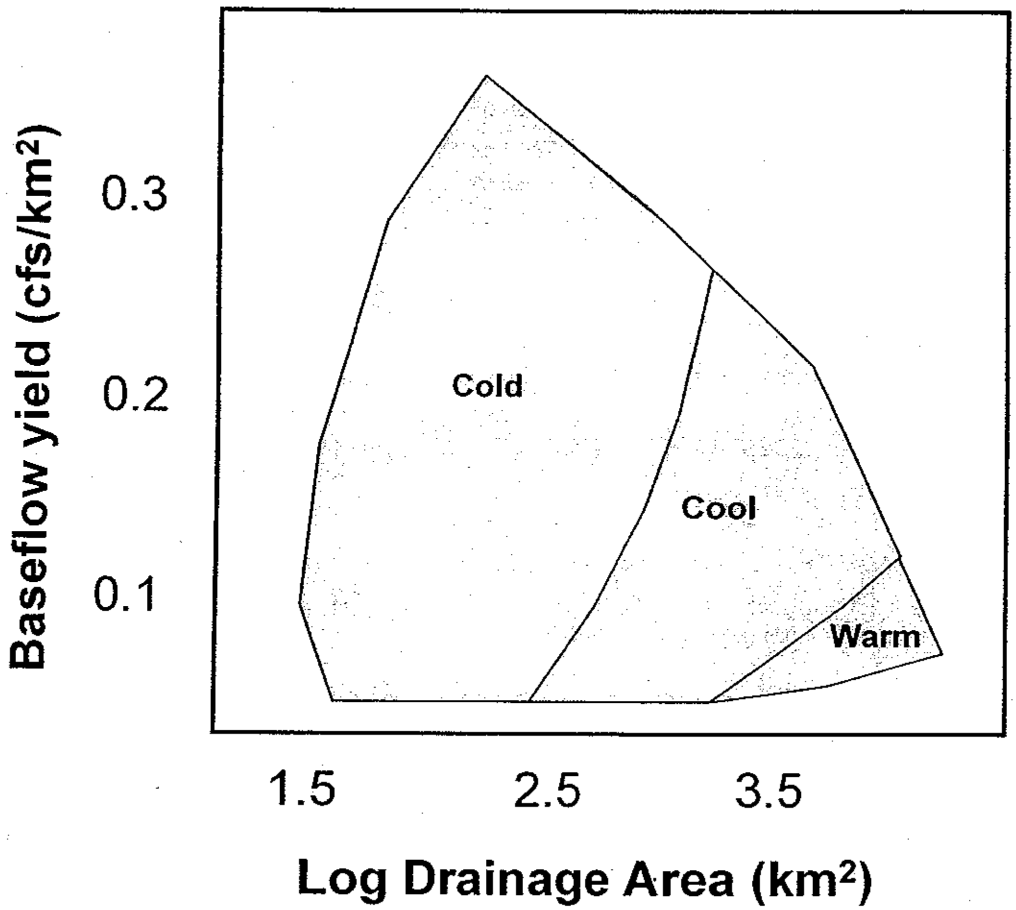


Figure I.1. Patterns of estimated July mean temperatures in lower Michigan streams plotted against catchment area and baseflow yield (90% exceedance flow per km^2 of the catchment).

Perhaps the simplest model that can be used to predict thermal status, or at lease the potential for a site to have cold, cool or warm-water fauna was a model constructed by Toronto and Region

Conservation Authority (EC, OMNR and OMOEE, 1998). TRCA found that when > 25% of the upstream catchment has well-drained soils, the receiving site is likely to have cold water and support a cold-water fish fauna. When the upstream catchment has between 10 and 25% well-drained soils, the site will likely have cool-water fauna, and when the site has < 10% well-drained soils, the site is likely to have warm-water fauna. This model is useful for predicting what should be found at a site, given the underlying soil porosity. Urbanization and other landuses will alter the original porosity, and thus the actual observed thermal status and biological community.

Nutrient Status Modelling

For the continental U.S., Omernik (1977) developed landscape-nutrient models explaining between 20 and 60% of the variance in nutrient concentrations. In the Salt Fork watershed (Illinois), Osborne and Wiley (1988) were generally able to explain between 90 and 97% of the variance in nitrate and soluble reactive phosphorus concentrations using landscape attributes. In the Saginaw Bay catchment of central Michigan, Johnson et al. (1997) were able to explain between about 60 and 70% of the variability in summer water quality parameters using percent agriculture, urban and forest cover. Finally, as cited in Seelbach et al. (1997), Kleiman (1995) was also able to relate surficial geology, soils and land cover/use to water chemistry. Kleiman's (1995) models have been used by Seelbach et al. (1997) to classify the nutrient status of streams in Michigan. They may be useful in Ontario but need to be verified/tested.

LITERATURE CITED

Bedient, P.B. and W.C.Huber. 1988. Hydrology and Floodplain Analysis. Addison-Wesley Publishing Company.

Cumming-Coburn. 2000. A technical reference manual for water budget analysis on a watershed basis. Prepared for the Ontario Ministry of Natural Resources, Ontario Ministry of the Environment, and Environment Canada.

Donigian, A.S. Jr. and W.C. Huber. 1991. Modeling of nonpoint source water quality in urban and non-urban areas. EPA/600/3-91/039. U.S. Environmental Protection Agency, Athens, GA.

Dunne, T. and L.B. Leopold. 1978. Water in Environment Planning. W.H. Freeman and Company, N.Y.

Environment Canada, Ontario Ministry of Natural Resources and Ontario Ministry Environment and Energy (EC, OMNR and OMOEE). 1998. A framework for guiding habitat rehabilitation in the Great Lakes Areas of Concern.

Johnson, L.B., C. Richards, G.E. Host and J.W. Arthur. 1997. Landscape influences on water chemistry in midwestern stream ecosystems. Freshwater Biology, 37:193-208.

Kleiman, R. 1995. Modeling water quality in Michigan rivers from landscape variables. M.S. Thesis, University of Michigan, Ann Arbor.

LeBlanc, R.T., R.D. Brown and J.E. Fitzgibbon. 1997. Modeling the effects of land use change on the water temperature in unregulated urban streams. Journal of Environmental Management, 49:445-469.

Marchant, D. and I.W. Heathcote. 1998. A review of available models for the prediction of pollutant loadings from rural non-point sources. Prepared for the Grand River Conservation Authority.

Meisner, J.D. 1990. Potential loss of thermal habitat for brook trout, due to climatic warming, in two southern Ontario streams. Transactions of the American Fisheries Society, 119:282-291.

Omernik, J.M. 1977. Nonpoint source-stream nutrient level relationships: a nationwide study. Environmental Research Laboratory, Office of Research and Development, U.S. Environmental Protection Agency, EPA-500/3-77-105.

Ontario Ministry of the Environment and Energy. 1994. Water Management, Policies, Guidelines, Provincial Water Quality Objectives of the Ministry of Environment and Energy. Queen's Printer for Ontario, 1994.

Osborne, L.L. and M.J. Wiley. 1988. Empirical relationships between land use/cover and stream water quality in an agricultural watershed. Journal of Environmental Management, 26:9-27.

Rahn, P.H. 1980. Engineering Geology: An Environmental Approach. Elsevier Science Publishing Co., New York, N.Y.

Seelbach, P.W., M.J. Wiley, J.C. Kotanchik, and M. E. Baker. 1997. A Landscaped-Based Ecological Classification System for River Valley Segments in Lower Michigan (MI-VSEC Version 1.0). Michigan Department of Natural Resources, Fisheries Research Report No. 2036.

Stoneman, C.L. and M.L. Jones. 1996. A simple methodology to evaluate the thermal stability of trout streams. North American Journal of Fisheries Management, 16:728-737.

Todd, D.K. 1980. Groundwater Hydrology. Wiley. New York, N.Y.

Weatherbee, D.G., D.L. Charlton, R. Tufgar, G. Rungis and J. Etienne. 1999. Temperature model application to evaluate stream naturalization. Proceedings of the Second International Conference on Natural Channel Systems, Niagara Falls, March 1 to 4, 1999.

Wehrly, K.E., M.J. Wiley and P.W. Seelbach. 1997. A thermal habitat classification for Lower Michigan rivers.. State of Michigan, Department of Natural Resources, Fisheries Division, Research Report Number 2038.

Section 2:

Interpreting Water Quality Evaluations

Section 2 Module 1

Developing and interpreting Hilsenhoff index scores[1]

Table of Contents

APPENDICES

[1] Authors: Kilgour, B. and L. W. Stanfield

1.0 INTRODUCTION

Both the screening level and kick and sweep surveys for benthic invertebrates provide insight into the overall suitability of a stream for invertebrates and thus other biota as well. While it may be difficult to diagnose causes it is possible to flag the site(s) as being potentially impacted, particularly if compared to reference sites. This section provides guidance about the situations when you should consider the possibility that a site is impacted.

1.1 Background

There is a direct link between the productivity of a stream and its nutrient loading regime, particularly for phosphorus and nitrogen. Streams that have an overabundance of nutrients are eutrophic and sometimes have problems with excessive growths of algae, followed by oxygen (O_2) depletion resulting from algal die off. Where O_2 depletion occurs, species that require highly oxygenated water (e.g., trout, and stoneflies) may become locally affected and replaced by more tolerant species such as stickleback or tubificid worms.

The nutrient status of a stream is usually determined by measuring the water-borne concentrations of nutrients such as nitrogen and phosphorus. In streams, however, nutrient concentrations vary daily with rain and runoff events. Characterization of the nutrient status of a stream requires at least monthly water chemistry sampling, as well as sampling during both dry and wet weather periods, and is therefore costly and time consuming. Rather than measure nutrients, the composition of animal and plant communities can also be used to classify the nutrient status of a site. In addition, the macroinvertebrates residing within a site reflect a variety of features including substrate texture, water quality, and temperature (which are also important factors for determining fish habitat suitability).

1.2 Evolution of the Module

There are recognized relationships between the relative abundances of various benthic taxa and the average nutrient concentrations at a site. For example, oligochaete worms and midges are more abundant at nutrient-enriched sites than at nutrient-poor sites. In contrast, mayflies and stoneflies tend to be more abundant in nutrient-poor sites and less abundant in nutrient-rich sites. Hilsenhoff (1987) was one of the first to develop a simplified index of stream health based on the tolerances of benthic families to nutrient status. In Hilsenhoff's model, taxa are rated on a scale of 0 (least tolerant to nutrient enrichment) to 10 (most tolerant). These data can be summarized for taxa identified using the diagnostic tool (Section 2: module 6). Kilgour (1998) averaged the biotic index tolerance values for major groups such as mayflies, stoneflies, chironomids, etc., which allows the nutrient index to be calculated for sites using data from rapid bioassessments (Table 1).

Table 1. Tolerance values for benthic orders in Southern Ontario.

Taxon	Modified Tolerance Value
Acarina (mites)	6
Oligochaeta (worms)	8
Hirudinea (leeches)	8
Amphipoda (scuds)	6
Isopoda (sowbugs)	8
Chironomidae (midges)	7
Simuliidae (blackflies)	6
Other Diptera (2-winged flies)	5
Tipulidae (craneflies)	3
Ephemeroptera (mayflies)	5
Plecoptera (stoneflies)	1
Hemiptera (true bugs)	5
Coleoptera (beetles)	4
Megaloptera (helgrammites)	4
Anisoptera (dragonflies)	5
Zygoptera (damselflies)	7
Trichoptera (caddisflies)	4
Gastropoda (snails)	7
Pelecypoda (clams)	8
Decapoda (crayfish)	6
Ostracoda (seed shrimp)	7

Kilgour (1998) used data collected from 55 streams in southern Ontario to demonstrate that the modified or major-groups biotic index can be used to distinguish between nutrient-poor (i.e., average phosphorus concentrations of less than 0.03 mg P/L) and nutrient-rich (i.e., average phosphorus concentrations greater than 0.03 mg/L) streams with about 70% accuracy.

1.3 Potential Uses

1.4 Screening-Level Impact Assessment

Data from screening level surveys are useful to determine whether additional information should be collected. The best evidence that this is the case is when data from a "test" site to data from reference sites that are known to have benthic communities in good condition. With this tool there would be a difference in the numbers of taxa between them. That is major taxonomic groups are absent from one or more sites that are present at others.

Further, these assessments could compare any number of indices (including the modified BI). For assistance in the interpretation of benthic data, see for example Rosenberg and Resh (1993), various articles in Davis and Simon (1995), Reynoldson *et al.* (1997), David *et al.* (1998), and Barbour *et al.* (1999).

1.5 3.0 Calculating the Biotic Index

As stated in Section 1 above, the data from the two replicate collections can be used to estimate a modified Hilsenhoff's biotic index (BI) for the site, using the following equation:

$$BI = \frac{\sum t_i n_i}{N} \qquad [1]$$

where

- n_i is the number of animals in taxon i;
- t_i is the tolerance value for taxon i; *and,*
- N is the total number of animals in the sample (calculated here as the average number of animals collected from the two replicates).

The HABPROGS database (Module 8) is capable of calculating the modified Hilsenhoff's BI, or the calculation can be easily executed in standard spreadsheet software.

1.6 Determining Nutrient Status

With animals identified to the level of major group as in Table 1, the relationship between the BI and nutrient status (average phosphorus concentrations) is depicted in Figure 5. Communities that produce BI values < 4 are generally in streams with < 0.03 mg/L total phosphorus and would be classified as oligotrophic, while communities that produce a BI> 6 to 6.5 are generally in streams with TP > 0.03 mg/L and that would be classified as potentially eutrophied. This BI can be used to correctly classify the nutrient status of sites about 70% of the time.

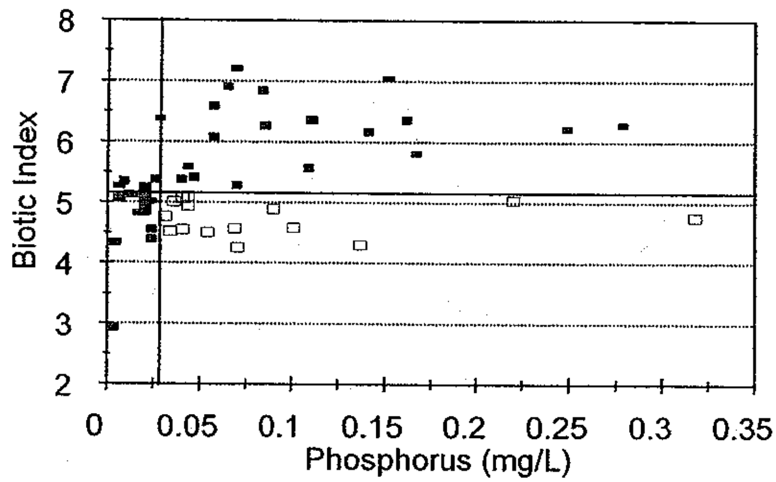


Figure 5. Relationship between total phosphorus concentrations and biotic index values in streams in southern Ontario. From Kilgour (1998).

1.7 Relating the Biotic Index to Fish Communities

Some preliminary work has established good relationships between the major-groups BI and biomasses of key stream fish species (Appendix 3).

In general, benthic community composition is related to fish community composition across both natural (e.g., river continuum) and pollutant-induced gradients (Kilgour and Barton, 1999). For example, cold-water streams will have cold-water benthos (e.g., Plecoptera) and fish (e.g., brook trout) while warm-water streams will have warm-water benthos (e.g., Odonata) and warm-water fish (e.g., common shiner). Credit Valley Conservation (1997) showed that the BI was related to fish community classifications. BI values < 5 corresponded with unimpaired cold-water fish communities, while BI values > 7.5 corresponded with impaired fish communities.

6.0 LITERATURE CITED

Barbour, M.T., J. Gerritsen, B.D. Snyder and J.B. Stribling. 1999. Rapid bioassessment protocols for use in streams and wadeable rivers: periphyton, benthic macroinvertebrates, and fish. Second Edition. EPA 841-B-99-002.

Credit Valley Conservation. 1997. Caledon Creek & Credit River Subwatershed Study: Draft Characterization Report (November 1997). Prepared by Credit Valley Conservation for the Region of Peel and the Town of Caledon.

David, S.M., K.M. Somers, R.A. Reid, R.J. Hall and R.E. Girard. 1998. Sampling protocols for the rapid bioassessment of streams and lakes using benthic macroinvertebrates. Ontario Ministry of the Environment, Aquatic Sciences Section.

Hilsenhoff, W. L. 1987. An improved biotic index of organic stream pollution. Great Lakes Entomologist. 20:31-39.

Kilgour, B.W. 1998. Developing an index of nutrient status based on rapid assessment methodology for collecting benthic macroinvertebrates. Prepared by Water Systems Analysts for the Ontario Ministry of Natural Resources.

Kilgour, B.W. and D.R. Barton .1999. Associations between stream fish and benthos across environmental gradients in southern Ontario, Canada. Freshwater Biology, 41:553-566.

Reynoldson, T.B., R.H. Norris, V.H. Resh, K.E. Day and D.M. Rosenberg. 1997. The reference condition: a comparison of multimetric and multivariate approaches to assess water-quality impairment using benthic macroinvertebrates. Journal of the North American Benthological Society, 16:833-852.

Section 2 Module 2

Interpreting point in time Stream Temperature Measures of summer maximums[1]

Table of Contents

APPENDICES

[1] Authors: L. W. Stanfield and M. L. Jones

1.0 INTRODUCTION

Stream water temperatures collected during periods of the summer that approximate the maximum temperatures observed at a site are very useful for understanding the types of biota expected to be found at a site as well as giving a measure of groundwater contributions to the system. We offer a few tools for helping interpret the temperature data in this section.

1.1 Calculating Thermal Stability

Stoneman and Jones (1996) found a relationship between the daily maximum stream temperatures and air temperatures at a number of sites with contrasting temperature conditions (i.e., from low to high variation in daily temperatures). These streams ranged from cold spring-fed to warm surface-discharge dominated systems. For streams that are cold and spring fed (i.e., are considered to have stable or relatively constant temperatures over an annual cycle), the study developed a good linear relationship between measured maximum water temperatures and maximum air temperatures on hot summer days.

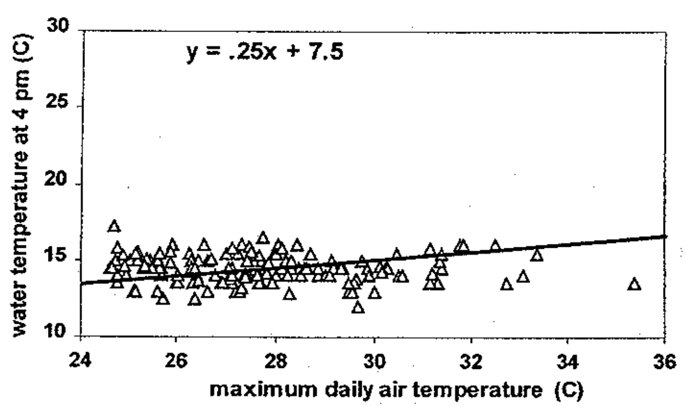


Figure 1: Water temperature at 4 pm and Maximum Daily Air Temperature at two cold reference sites.

The relationship is defined by the equation:

$$y = 0.2513x + 7.5131, \qquad [1]$$

where,

- y is the predicted temperature of a spring-fed cold-water stream; and,
- x is the maximum air temperature (Figure 1).

The relationship is specific to water temperatures measured at 4 p.m. on hot days. Deviation from the predicted relationship is a measure of thermal stability. A temperature close to the predicted relationship implies a stable temperature regime, while an observed temperature much higher than the predicted relationship implies a non-stable temperature regime (i.e., variable from day to night and from winter to summer).

1.2 Classification of Streams based on temperature

A number of organisations have used the nomograms created from the Stoneman and Jones (1995) manuscript to classify streams into cold, cool and warmwater groups. To do this users would simply apply the observations collected using the results of the rapid assessment of temperature (S2.M1 of the OSAP manual) to the nomogram in Figure 2 to determine which of the three classes the site falls into.

We caution users that this nomogram and the classifications that it implies should not be used as a means of classifying fish communities in streams as these criteria have not been related to fish communities. In fact the cool water criteria may be optimal for salmonid production in the province of Ontario (see section 6).

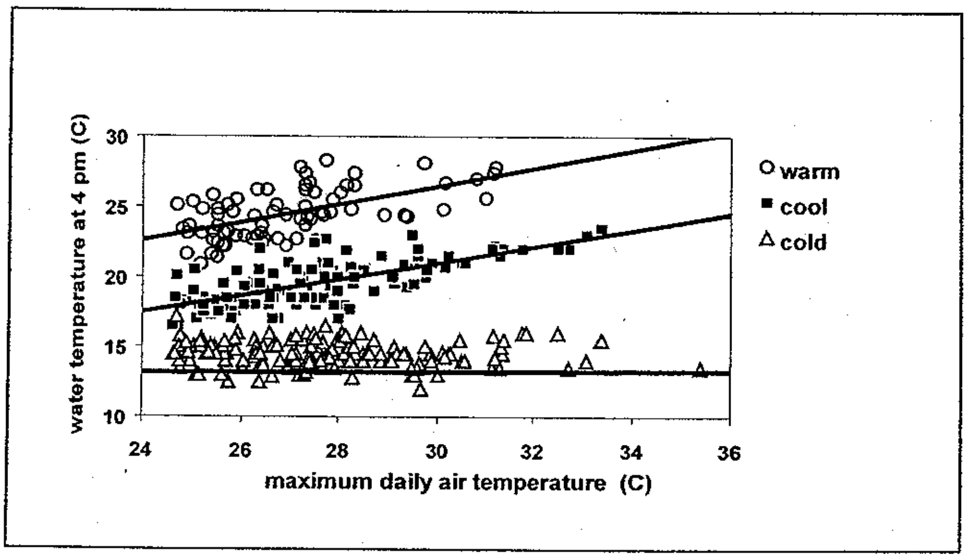


Figure 2. Nomogram showing regressions between air and water temperatures for three reference stream types.

1.3 Estimating Summer Maximum Temperature

Ideally summer maximum temperature would be obtained from all sites in a study on the same day. This rarely happens and as a result air temperatures can vary substantially between sites. As noted in Figures 1 and 2 differences in air temperature influence the water temperatures that are observed. Here we describe a procedure for standardizing stream temperatures between sites, thus improving the capabilities of making comparisons between sites.

The first step is to determine which relationship most closely describes the theraml properties for each site (i.e., which regression line should be used of the three available): Plot each observation in your dataset on the nomogram (Figure 2) and pick the regression line that each observation is closest to. Next use the appropriate regression line to estimate the water temperature for a standard air temperature that approximates the maximum for the summer (i.e., 32 C). The next step is to compare the degree that the site deviated from the reference stream. For example if a site was measured at 22 C and the predicted temperature for the same air temperature was 19 C, then this site would be 3 C warmer than the reference streams. This difference is added to the predicted summer maximum. The equations to be used for each regression equations are:

Coldwater class $y = 0.2513x + 7.5131 - (a - b)$
Coolwater class $y = 0.5831x + 3.5007 - (a - b)$
Warmwater class $y = 0.6414x + 7.206 - (a - b)$

Where y = standardized water temperature; x equals the air temperature you wish to standardize to; a = the predicted water temperature at the observed air temperature and b= the observed water temperature at an observed air temperature.

The final number then represents the standardized maximum water temperature for each site. This methodology assumes that the slope of the temperature relationships will not vary significantly. An example of this procedure is provided in Box 1.

Box 1: Example of calculation of standardized thermal stability measures:

At site A, we measured water temperature to be 25 °C and obtained an air temperature of 26°C. Based on the reference stream regression lines this would be classified as a warm water stream. We would then use the warm water regression line to predict the water temperature using the following formula:

Predicted water temp at 32 C = air temp (32) X 0.6414(slope of warm water regression line)
+ 7.206 (intercept of warm water regression line) - (23 – 25.9)
= 27.7 +1.1 °C
= 28.8 °C

We would predict that this site would be 28.8 °C at a standardized air temperature of 32 °C

1.4 Relating Stream temperature to fish species composition

There is a tremendous wealth of literature describing the suitability of streams to support various fish species based on their thermal tolerances. In section 5 we show suitability curves for many Ontario species to thermal stability data collected using the procedures described in the OSAP manual. Recently several agencies have adopted classificiation schemes for describing the potential fish communities that might be found at a site based on summer maximum temperatures. For example the OMNR's "Let's Fish Ontario!: Fish Biology and Identification" web page (http://www.mnr.gov.on.ca/MNR/fishing/bio.html) classifies species of Ontario sportfish according to the following three thermal regimes:

- Cold: maximum water temperature 10 to 18°C
- Cool: maximum water temperature 18 to 25°C
- Warm: maximum water temperature over 25°C

Table 1 presents some typical fish species for thermal regimes based on July weekly mean stream temperatures (Wehrly *et al.*, 1999).

Table 1: Typical Fish Species for each Thermal Regime

Thermal Regime	Typical Fish Species
Cold (10 to 18°C)	Slimy sculpin, mottled sculpin, brook trout, brown trout.
Cool (19 to 21°C)	Northern pike, white sucker, creek chub, longnose dace, blacknose dace, central mudminnow.
Warm (22 to 26°C)	Green sunfish, channel catfish, common carp, smallmouth bass, rock bass.

Regardless of which classification scheme is used, some overlap of these groups may occur during extreme summer conditions as fish are exposed to temperatures warmer than their preferred thermal regime, or where thermal refuge areas are available. Different species have different tolerance levels for temperature variation, so some species may be more adaptable to rising temperatures than others (e.g., northern pike, creek chub, central mudminnow).

Remember that warmwater streams provide critical habitats for cold water species for the majority of the year (i.e., feeding and spawning areas in the fall and spring, overwintering habitat and staging areas prior to spawning or smolting).

2.0 LITERATURE CITED

Stoneman, C.L., and M.L. Jones. 1996. A simple method to evaluate the stability of trout streams. North American Journal of Fisheries Management. 2:728-737.

Wehrly, K.E., M.J. Wiley and P.W. Seelbach. 1999. A thermal habitat classification for Lower Michigan Rivers. State of Michigan, Department of Natural Resources, Fisheries Division, Research Report Number 2038.

North American Journal of Fisheries Management 16:728–737, 1996

A Simple Method to Classify Stream Thermal Stability with Single Observations of Daily Maximum Water and Air Temperatures[1]

CHRISTINE L. STONEMAN[2]

Watershed Ecosystems Graduate Program, Trent University
Peterborough, Ontario K9J 7B8, Canada

MICHAEL L. JONES

Ontario Ministry of Natural Resources, Aquatic Ecosystems Research Section
Rural Route 4, Picton, Ontario K0K 2T0, Canada

Abstract.—The relationship between instream water temperature and ambient air temperature at six stream sites in southern Ontario was examined. At two sites, maximum summer water temperatures never exceeded 17°C; at two others, temperatures remained below 23°C; and at the remaining two, temperatures reached 28°C. The relationship that best distinguished the three pairs of sites was the regression of water temperature measured at 1600 hours on maximum air temperature. Analysis of covariance indicated that the regression slopes for the first (cold) and second (cool) pairs of sites were nonhomogeneous; those for the second (cool) and third (warm) pairs were homogeneous, and the adjusted means were significantly different. Where data were available, analysis of covariance indicated that the relationship did not differ between years. Graphical analysis of the data indicated little overlap of 95% confidence intervals at air temperatures greater than 25°C. The regression results were used to develop a nomogram to determine the thermal stability of stream sites from a single observation of water temperature at 1600 hours on a warm summer day and a maximum air temperature estimate for the same day.

Biologists and fishery managers have long recognized that different streams and different locations within a stream can exhibit very different thermal regimes and that these regimes have a critical influence on the biota that the site can support. Some streams remain very cold throughout hot summer months; others can reach temperatures close to ambient air temperatures. The maintenance of cool summer water temperatures in thermally stable sites minimizes heat stress on stenothermal fish. These same sites may exhibit relatively warm water temperatures in the winter, thereby preventing the formation of anchor ice, another important determinant of habitat suitability. Because water temperatures are closely linked to biota, it is important to both managers and researchers to be able to accurately and efficiently evaluate the thermal regime of a site.

Thermal regime is often assessed by using direct measurements of temperatures within the stream over an appropriate period of time (i.e., spanning the normally observed range of meteorological conditions for an area). This usually requires continuous monitoring of stream temperatures throughout the summer with maximum–minimum thermometers (e.g., Bowlby and Roff 1986) or continuous recording devices. While this method does provide adequate information to describe the thermal regime of a stream site, it is both labor intensive and expensive, particularly for projects that cover a large geographical area or comprise numerous sites.

Models have been built to predict thermal regime by including several factors that may affect water temperatures. They are useful in describing the effects of changes in the watershed on the thermal regime (e.g., Brown 1969; Raphael 1962). However, the data requirements for these models can be excessive. Kothandaraman (1972) found that air temperatures may be used to predict steam temperatures, provided that data exist for a stream with similar attributes (e.g., depth, shading). Stefan and Preud'homme (1993) determined that stream temperature estimation from air temperature exhibited smaller standard deviations within a stream than across streams. Also, small, shallow streams exhibited less deviation than large, deep streams. Stefan and Preud'homme (1993) also noticed a time lag between maximum air temperature

[1] Contribution 96-05 of the Ontario Ministry of Natural Resources Aquatic Ecosystems Research Section.

[2] Present address: Department of Fisheries and Oceans, Fisheries and Habitat Management, Bayfield Institute, 867 Lakeshore Road, Post Office Box 5050, Burlington, Ontario L7R 4A6, Canada.

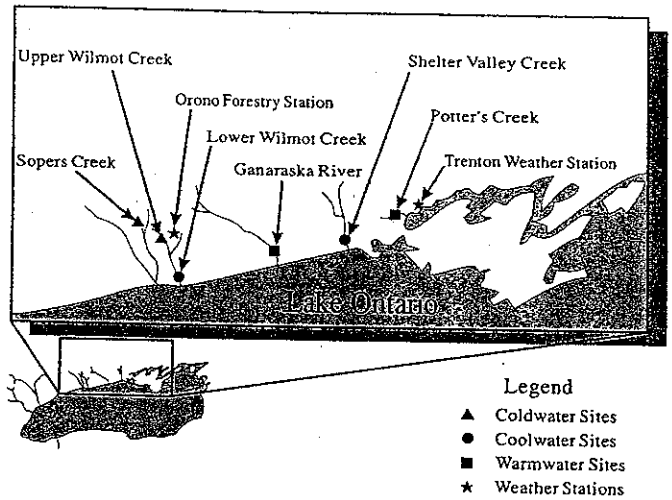


FIGURE 1.—Location of streams, sites, and weather stations used in the study.

and maximum stream temperature and that this time lag increased with the depth of the stream. These studies suggest that at a fixed air temperature, the water temperature of streams with differing thermal stability will respond differently, but in a consistent way. Sites dominated by groundwater inputs or with extensive shading will tend to remain cold on very hot days, whereas open sites with relatively little groundwater contribution can attain water temperatures approaching ambient air temperatures.

In this study, an investigation of the relationships between air and water temperature on streams of contrasting thermal regimes was conducted. It was decided a priori that the sites selected would represent coldwater, coolwater, and warmwater streams so that the differences between these three categories of sites could be assessed and quantified.

Methods

Data Collection

Continuous water temperature data were collected from six sites on Lake Ontario tributary streams (Figure 1) during the summer months of July, August, and September. Two sites were chosen to represent each of three thermal regimes—coldwater sites that support large populations of coldwater species (mainly Salmonidae and Cottidae), coolwater sites that support a mix of coldwater and coolwater–warmwater species (Cyprinidae, Percidae, Centrarchidae), and warmwater sites that support few, if any, coldwater species. The definitions of thermal regime based upon species composition were compared to actual water temperature data to verify the classification.

Between 1990 and 1992, data were collected at a coldwater and a coolwater site on Wilmot Creek (coldwater #1 and coolwater #1; Figure 1). At each site a Ryan type J paper thermograph recorder was placed in a cement block and submerged in a shaded pool. The paper tapes were replaced at roughly 2.5-month intervals. Between placements, the thermographs were periodically checked and the date and time were marked on the tape. From these markings, the thermographs were checked to determine the accuracy of their internal clocks. All paper thermographs used in this study functioned correctly.

In the summer of 1994, data were collected from the Ganaraska River (warmwater #1), Shelter Valley Creek (coolwater #2), Sopers Creek (coldwater #2), and Potter's Creek (warmwater #2; Figure 1) with Onset Technologie's Stowaway model CT16 temperature recorders. These recorders provided a digital temperature reading (set to read every 10 min). Stowaways were placed at each stream site in early July. As with the Ryan J thermographs, the Stowaways were secured to a cement block and then placed in a shaded pool. The data were downloaded monthly and the Stowaways were removed by mid-September.

Regional air temperature data were obtained from the weather stations at Orono and Trenton (Figure 1) for the required time periods. These weather stations record the maximum and minimum daily air temperature, as well as precipitation data. Species composition data were collected by

TABLE 1.—Summary of characteristics of the six stream sites studied.

Characteristic or measurement	Coldwater 1	Coldwater 2	Coolwater 1	Coolwater 2	Warmwater 1	Warmwater 2
Location	Upper Wilmot Creek	Sopers Creek tributary	Lower Wilmot Creek	Mid–lower Shelter Valley Creek	Lower Ganaraska River	Potters Creek
Maximum temperature						
Water	16°C	17.2°C	23.5°C	22.7°C	28°C	28°C
Air	35.4°C	32°C	33.4°C	32°C	32°C	32°C
Average daily maximum temperature						
Water	14°C	14°C	18°C	18.9°C	23.4°C	22.4°C
Air	27°C	25°C	25°C	25°C	25°C	25°C
Fish species composition[a]	Brown trout, rainbow trout, cottids	Brown trout, brook trout, rainbow trout, cottids	Some brown trout, rainbow trout, cyprinids, percids, catostomids	Some brown trout, rainbow trout, cyprinids, percids, catostomids	Very few rainbow trout; cyprinids, percids dominate	No salmonids
Anchor ice formation	No	No	Some	Some	Yes	Yes
Period of temperature data collection	May–Nov 1990, 1991	Jul–Sep 1994	May–Nov 1991, 1992	Jul–Sep 1994	Jul–Sep 1994	Jul–Sep 1994

[a] Brown trout *Salmo trutta*, rainbow trout *Oncorhynchus mykiss*, brook trout *Savelinus fontinalis*.

electrofishing, with the exception of the Potters Creek site, where Moira River Conservation Authority personnel reported no occurrence of salmonids in the tributary. Data on anchor ice formation were collected by visually assessing the relative amount of anchor ice in winter. Site characteristics are summarized in Table 1.

Data Analysis

With the data collected from Ryan thermographs, water temperature values were read from tapes at 900, 1400, 1500, and 1600 hours for all available data (all years, May–November). These four times were chosen in order to examine the change in water temperatures between the morning and afternoon and to pinpoint the approximate time of maximum daily water temperature. Daily changes in air temperature (maximum air temperature minus minimum air temperature) were plotted against the corresponding change in water temperature (maximum water temperature minus minimum water temperature) for one site in each thermal regime category (coldwater #1, coolwater #1, and warmwater #1). Water temperatures at 900, 1400, 1500, and 1600 hours were also plotted against the regional maximum air temperature.

The plotted data were examined visually and with analysis of covariance (ANCOVA) and regression methods to determine which relationship provided the greatest amount of separation between the three categories. The ANCOVA was used to determine if the slopes of the different sites were homogeneous and to compare between years for the coldwater #1 and coolwater #1 sites.

With the data from all sites, regressions of maximum air temperature against water temperature at 1600 hours were completed, and ANCOVA methods were used to examine relationships between sites and categories of sites for the midsummer months. These same analyses were repeated for only those days on which the maximum air temperature was above 24.5°C.

The residuals from the regressions of maximum air temperature against water temperatures at 1600 hours were plotted against the following: the previous day's maximum water temperature, the previous day's maximum air temperature, an average of the maximum air temperature on the three previous days, the time at which the day's maximum air temperature occurred, the minimum air temperature, the previous day's minimum air temperature, and the amount of precipitation for that day. A simple correlation analysis was completed on these data.

By determining the equivalence points (the set of *x*, *y* points equally likely to be from either of two thermal regime categories, as determined by the confidence intervals around the regression lines) for the coldwater versus coolwater and the coolwater versus warmwater thermal regime categories, a nomogram was developed to assign other streams to a category based on a single obser-

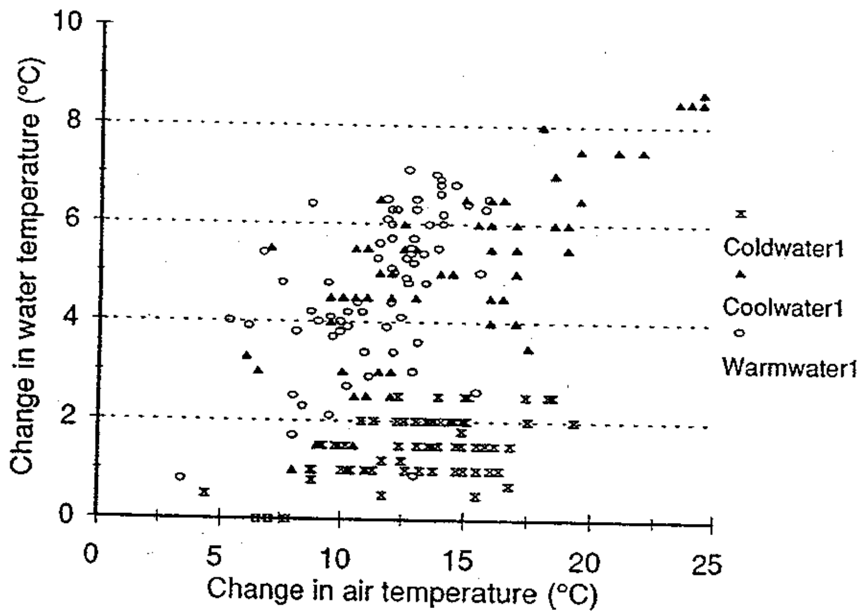


FIGURE 2.—The daily change in air temperature plotted against the corresponding change in water temperature for site coldwater #1 in 1990, site coolwater #1 in 1991, and site warmwater #1 in 1994.

vation of daily maximum air temperature and water temperature at 1600 hours.

Results

The relationship between the daily change in air temperature and the corresponding change in stream temperatures did not differentiate well between the three thermal regime categories, particularly in the coolwater and warmwater sites (Figure 2). Closer examination of the data showed that the coolwater and warmwater sites exhibited similar diurnal fluctuations in water temperature but that on warm days the warmwater sites tended to remain at high temperatures throughout the night and following day. This suggests that actual water temperature at a particular time rather than daily changes in water temperature are more likely to differentiate between the site categories.

Plotting maximum daily air temperatures against the water temperature at 1600 hours effectively differentiated the three stream categories. When all six sites were plotted on the same graph (Figure 3), the sites clustered together by category: coldwater #1 and #2, coolwater #1 and #2, and warmwater #1 and #2. At all sites, water temperature at 1600 hours never deviated by more than 0.5°C from the daily maximum water temperature. At the sites on smaller streams, maximum daily temperature began to decline shortly after 1600 hours; in larger systems, the water temperatures had just peaked by 1600 hours. This same phenomenon was reported by Stefan and Preud'homme (1993), who stated that the time lag between maximum air and maximum water temperatures was directly related to water depth.

The separation between the three categories of sites was most evident for temperatures above 25°C obtained during July, August, and the first week of September. Before and after these dates, average maximum air temperatures usually did not reach high enough values to allow adequate separation of sites. During the summer months, the categories of sites showed sufficient differentiation, such that the 95% confidence intervals did not overlap at air temperatures above 25.5°C for coolwater and coldwater sites and showed little overlap between coolwater and warmwater sites (Figure 4).

When the regressions for the three categories were compared graphically and with ANCOVA, the nonhomogeniety of slopes between the coldwater sites and both the coolwater and warmwater sites was evident (Table 2; Figure 4). The lower (smaller) slope for the coldwater sites caused the coldwater sites to diverge from the others as air temperature increased. This result indicates that the coldwater sites react differently to changes in ambient air temperatures than do the other sites. The lower slope of the coldwater sites indicates that they are more thermally stable than the cool-

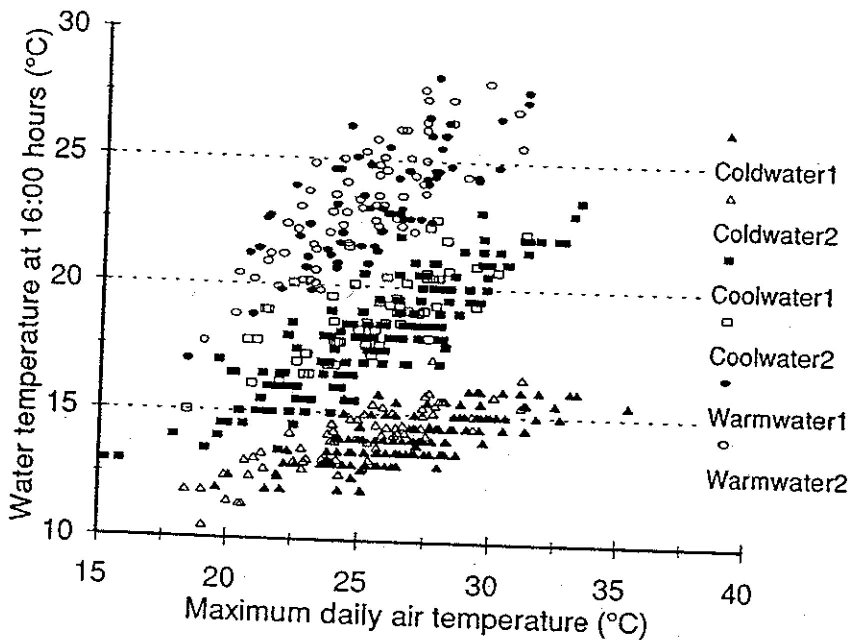


FIGURE 3.—Maximum daily air temperature data plotted against the corresponding water temperature at 1600 hours for all sites and all years' data. The sites cluster by thermal category: coldwater, coolwater, and warmwater.

water and warmwater sites because they are less affected by increasing air temperatures. The steeper slopes of the coolwater and warmwater sites indicates that the water temperatures at these sites are more closely linked to the ambient air temperatures.

The ANCOVA was subsequently used to investigate the similarity of the sites that had been selected to represent streams with similar thermal regimes. The slopes of the two coldwater sites are homogeneous, but the adjusted means are significantly different (Table 2). This result is impossible to distinguish graphically and there is sufficient overlap between the sites so that a separation based upon mean water temperature is not possible (Table 2). The same is true for the cool-

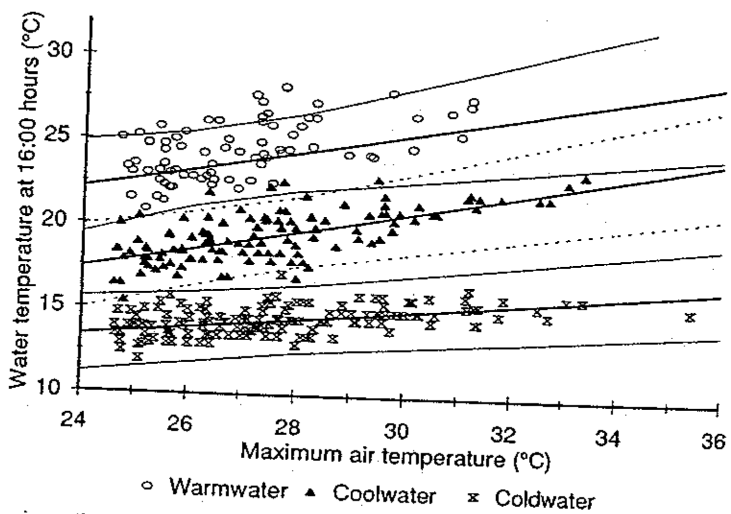


FIGURE 4.—Regressions (heavy solid lines) of maximum air temperature plotted against water temperature at 1600 hours for each thermal category. Also shown are the 95% confidence limits (light solid lines or broken line) for each of the three regression lines. The slopes of the coolwater and warmwater sites are homogeneous, and the coldwater slope is significantly different from the other two slopes.

TABLE
and wate
on days
corrected,

Covari
Coldwater
coldwater
Coolwater
coolwater
Coldwater
coldwater
Coolwater
coolwater
Warmwater
warmwate
Coldwater si
coolwater
warmwate
Coldwater si
coolwater
Coolwater si
warmwater
Coldwater si
warmwater

water site
are also h
not differ
Compar
location in
The result
for site co
mogeneous
coolwater
2). At both

FIGURE 5.—
dashed lines)
lines) and 19

TABLE 2.—Results of analysis of covariance. For all tests, maximum daily air temperature is the independent variable and water temperature at 1600 hours is the dependent variable. Only data collected between July 1 and September 9 on days where maximum air temperatures exceeded 24.5°C were used. Comparisons between means are Bonferroni corrected, meaning $P > 0.005$ is considered significant; NA = not applicable.

Covariates	N	Slopes (b_1, b_2)	P ($b_1 = b_2$)	Adjusted means		P ($x_1 = x_2$)
				x_1	x_2	
Coldwater 1, 1990; coldwater 1, 1991	105	Homogeneous	0.522	14.529	14.179	0.009
Coolwater 1, 1991; coolwater 1, 1992	85	Homogeneous	0.998	19.792	18.495	0.000
Coldwater 1; coldwater 2	145	Homogeneous	0.750	14.258	15.044	0.000
Coolwater 1; coolwater 2	118	Homogeneous	0.698	19.996	19.318	0.005
Warmwater 1; warmwater 2	74	Homogeneous	0.694	24.126	24.879	0.025
Coldwater sites; coolwater sites; warmwater sites	337	Nonhomogeneous	0.000	NA	NA	NA
Coldwater sites; coolwater sites	263	Nonhomogeneous	0.000	NA	NA	NA
Coolwater sites; warmwater sites	192	Homogeneous	0.586	19.381	24.668	0.000
Coldwater sites; warmwater sites	219	Nonhomogeneous	0.000	NA	NA	NA

water sites. The slopes of the two warmwater sites are also homogeneous and the adjusted means do not differ significantly.

Comparisons of data obtained from the same location in different years yielded similar results. The results of ANCOVA indicate that the slopes for site coldwater #1 for 1990 and 1991 are homogeneous. The same result was obtained for site coolwater #1 for the years 1991 and 1992 (Table 2). At both of these sites, the mean water temperature, when adjusted for the covariate, was not significantly different between years. These results are also evident graphically, as the amount of overlap between years for each site is substantial (Figures 5, 6).

When the residuals (actual water temperature at 1600 hours minus the predicted water temperature) for each thermal regime category were correlated against other physical factors, several factors were significantly correlated with the coldwater resid-

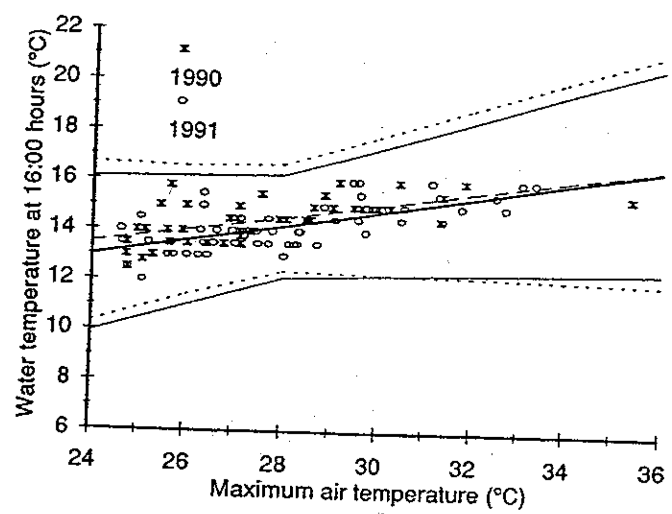


FIGURE 5.—Regressions (heavy solid and long dashed lines) and 95% confidence intervals (light solid and short dashed lines) of maximum air temperature against water temperature at 1600 hours for data from 1990 (dashed lines) and 1991 (solid lines) at the coldwater 1 site.

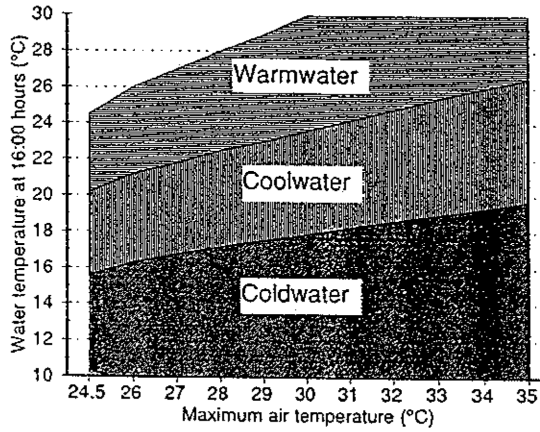


FIGURE 7.—A nomogram of maximum air temperature and water temperature at 1600 hours. Upon obtaining a measure of the maximum daily air temperature and the water temperature at 1600 hours, the nomogram can be used to determine whether the site is most like a coldwater, coolwater, or warmwater site.

served maximum daily water temperature (which typically occurs at or near 1600 hours) on warm days, corrected for the actual maximum air temperature on the same day, appears to be diagnostic of the type of stream being sampled.

These measurements provide a simple, practical method for classifying streams according to their thermal regime. The regression lines depicted in Figure 4 for the three site categories can be used to develop a nomogram (Figure 7) that allows users to classify new sites by using the water temperature at 1600 hours and maximum air temperature measurements on warm summer days. The nomogram works by defining a "most probable" stream type for any combination of maximum air temperature and water temperature at 1600 hours. Classification is based on the confidence intervals surrounding each of the three regression lines. The boundary lines between the site categories are defined by the set of points at which it is equally likely that an observation represented by those points came from either of two stream types. The further a point lies from the boundary line between two stream types, the less likely it is that the site will be classified incorrectly.

The development of this nomogram implies that the three sets of streams used in this study constitute "reference" streams against which data collected from other streams can be compared. In reality, the range of thermal conditions (groundwater importance, shading, etc.) exhibited by streams in a region of varying physiographic and land use features will represent a continuum from well-shaded streams almost exclusively fed by groundwater to open streams dominated by surface runoff. Discrete categories of stream thermal types do not exist. Nevertheless it is often useful to be able to classify systems into approximate categories for purposes such as regional assessment and watershed planning. The nomogram provides an easily applied, objective tool for this purpose.

Some of the residual variation in maximum water temperatures (after the effect of maximum air temperature was accounted for) is explained by antecedent conditions (Table 3). The previous day's minimum air temperature shows the greatest correlation for the cool- and warmwater sites and also was significantly correlated for the coldwater sites. However, in all cases, the correlations are not strong, suggesting that the addition of easily collected data on antecedent conditions will not substantially improve the discriminatory power of the regression models presented earlier. Evidently, the current day's air temperature explains most of the day-to-day variation in maximum water temperatures among streams of a given type.

It is likely that the circumstances under which antecedent conditions will cause misleading deviations from the expected air–water temperature relationship will be those in which the weather on preceding days has been quite different from that day's weather. This is illustrated in Figure 8, in which a set of three high outliers, one from each of the regression lines (cold, cool, warm) are highlighted. These points are associated with a relatively cool day which followed a period of unusually warm weather. The practical ramification of this and similar observations is that data collection for the purposes of applying the nomogram (Figure 7) should be limited to days on which no major changes in weather have occurred relative to the preceding 2–3 d.

It should be emphasized that the method of stream thermal classification presented in this paper is an approximate one. As the preceding discussion illustrates, circumstances under which a single day's water temperature observation may lead to a misleading interpretation of a stream's thermal stability are not difficult to imagine. The most obvious way to reduce the risk associated with a single observation is to increase the number of observations. If several observations are collected from a single site they can all be plotted on the nomogram and their collective position (i.e., centroid) used to classify the stream. Further, when resources are available, continuous temperature data collected over a period of weeks or months

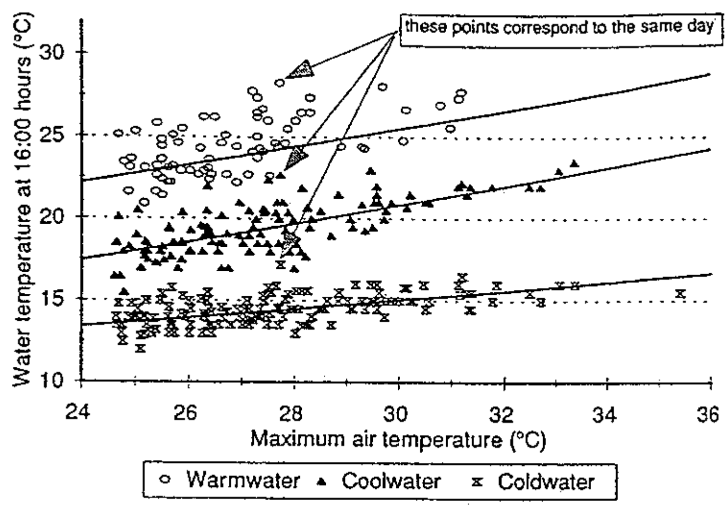


FIGURE 8.—Regressions (heavy solid lines) of maximum air temperature plotted against water temperature at 1600 hours for each thermal category. The arrows highlight data that were collected on a 27.7°C day that had been preceded by a relatively hot day with a maximum air temperature of 31.2°C and a minimum air temperature of 19°C. It appears that these data were influenced by the previous day's high air temperatures because the water temperatures were higher than expected.

will invariably provide a more accurate description of the stream's thermal regime. However, for many applications, such as watershed planning and stream assessment, such intensive data collection may be both impractical and unnecessary. The rapid assessment technique described here offers a far more cost-effective and manageable solution for describing thermal habitat.

The data collected in this study do not shed light on the mechanisms that give rise to the differences among stream types in air–water temperature relationships. In southern Ontario, where these data were collected, groundwater inputs often play an important role in determining the thermal regime found in a stream (Meisner 1990) and thereby influence the suitability of the stream for coldwater species, such as salmonids (Bowlby and Roff 1986). A simple two-box mixing model applied to our data, together with the assumption of a fixed groundwater temperature of 8.5°C (Meisner 1990), implied that the fractional groundwater contributions to our cold-, cool-, and warmwater sites were approximately 0.7, 0.4, and 0.15 respectively (M. L. Jones, unpublished data). Other factors that influence the degree to which stream temperatures will respond to changes in air temperature include shading by riparian vegetation and stream size (i.e., the volume of water in the stream).

Finally, it is important to note that the thermal stability of a stream, as defined by the method described above, does not necessarily determine the fish species that one will find at a site in any given year. For example, in relatively cool summers, coolwater streams may support coldwater species. Summer water temperature averag 19.6°C at the lower Wilmot Creek (coolwater #. site during 1991 but was only 16.4°C in 1992. Despite these differences in average temperatures, the regression line slopes for the 2 years were not significantly different. Data collected from this site in either year would have resulted in the same classification. In 1992, however, thermal conditions at this site would have been significantly more favorable for stenothermal species than one would typically expect in a coolwater stream. The biological community found in a particular stream should reflect the interaction between the thermal stability of the stream and the recent meteorological history of the area. The lag between meteorological change and community change will depend on the thermal sensitivity and turnover rates of the species present. However, general conclusions can be drawn about the types of fish communities that could be supported by a site in most years. A warmwater site is very unlikely to support large numbers or biomass of salmonids. In contrast, a coldwater site will provide optimal thermal habitat for brook trout and brown trout.

Acknowledgments

We thank the many people involved in the project, Mike Bozak, Leon Carl, Gina Cunningham, Jack Imhof, Serge Metikosh, Cam Portt, Norm Smith, Les Stanfield, and Mike Stoneman reviewed this and earlier drafts of the manuscript. We also thank the Department of Fisheries and Oceans' Action Plan on Fish Habitat and the Ontario Ministry of Natural Resources for providing funding for this project.

References

Bowlby, J., and Roff, J. 1986. Biomass and habitat relationships in southern Ontario streams. Transactions of the American Fisheries Society 115:503–514.

Brown, G. W. 1969. Predicting temperatures of small streams. Water Resources Bulletin 5:68–75.

Kothandaraman, V. 1972. Air–water temperature relationship in Illinois river. Water Resources Bulletin 8:38–45.

Meisner, J. D. 1990. Effect of climatic warming on the southern margins of the native range of Brook trout. *Salvelinus fontinalis*. Canadian Journal of Fisheries and Aquatic Sciences 47:1065–1070.

Raphael, J. M. 1962. Prediction of temperature in rivers and reservoirs. Journal of the Power Division. Proceedings of the American Society of Civil Engineers 88:157–181.

Stefan, H. G., and E. B. Preud'homme. 1993. Stream temperature estimation from air temperature. Water Resources Bulletin 29:27–45.

Section 3:

Fish Community Sampling

Section 3: Module 1

Criteria for interpreting electrofishing survey results[1]

Table of Contents

[1] Authors: L. W. Stanfield, M. L. Jones, Kilgour, B. and N. Lester

1.0 INTRODUCTION

This section provides a summary of some of the background to the relationships between single pass and multiple pass catches and between single pass low intensity (i.e., screening) surveys and the standard higher effort single pass surveys. This section will be useful to users as they interpret the results of electrofishing surveys.

1.1 Relating Species Abundance from single to multiple pass surveys

While analyzing some of the Great Lakes Salmonid Unit's long term data sets, a relationship was observed between the total population estimates of salmonids derived from removal method electrofishing and the catch of salmonids in the first pass. These data had all been collected using three-pass removal methods with a slow, thorough approach. The Salmonid Unit tested whether this relationship would hold up in a more rigorous assessment. Surveys were carried out using single- and three-pass techniques several weeks apart. A strong correlation was found between the single- and multi-pass population estimates (Figure 1). This provided a means of meeting both the needs of a rigorous monitoring program and the desire of managers to increase the number of stations that could be sampled in a season.

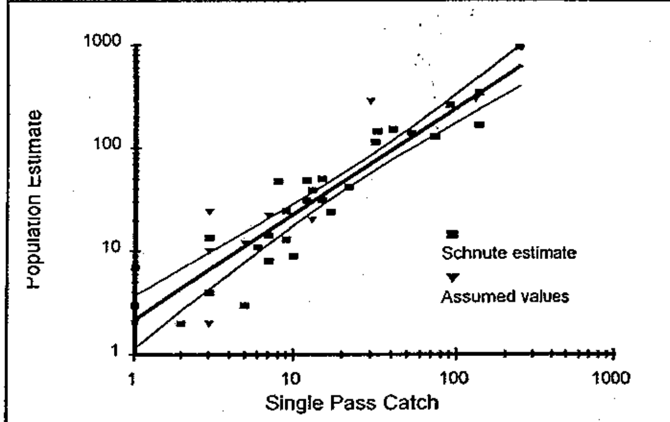


Figure 1 Relationship between population estimates and independent single-pass catches for salmonids at 12 Wilmot Creek Sites. From Jones and Stockwell (1995).

The single-pass methodology (as described in Section 3.1.2) has been tested and is known to provide reliable estimates of the abundances of juvenile salmonids (R^2 = 0.86, Jones and Stockwell, 1995), when the shocking effort is ∃15s/m^2. The single-pass methodology also provides an objective and repeatable means of estimating relative abundances of other non-salmonid taxa.

This method typically uses 15 to 21 seconds of effort per m^2, and is intended to allow crews to produce a thorough and repeatable assessment of fish communities at a site, while minimizing the potential for bias due to varying skill levels and sampling experience among crews. If the effort applied is between 15 and 21 s/m^2, the data will be comparable to data collected by others using this protocol. Further, the data can be used in the further refinement or development of Overall Habitat Suitability Index Models (OHSIM) that relate habitat attributes to stream fish species abundances (section 5).

The formula that is used to estimate total population of salmonids (TPE) is:

$$TPE = 1.02 \times SPC + 0.78 \qquad [1]$$

where,

- *TPE* is the total population estimate; and,
- *SPC* is the single-pass catch.

This calculation is carried out automatically as part of HABPROGS, but can be carried out manually for all or individual species of salmonids.

HABPROGS has not yet been set up to carry out the equations generally associated with removal methods. Where multi-pass removal methods were used to generate data, users should consult Zippen (1958), Mahon (1980), Schnute (1983), and others for guidance on carrying out calculations of total population estimates. The Schnute regression formula, which is used to calculate total abundance's of a species at a site based on multi-pass data, will be incorporated in future versions of the HABPROGS database.

1.2 Relating Screening survey results to Standard assessments

In 1999, Stanfield *et al.* (2002) examined the possibility of reducing the total shocking effort to 5s/m^2. They repeated the earlier study on Wilmot Creek varying the shocking effort by 5, 10 and 15 s/m^2 and comparing the catches at the various intensities. As shown in Figure 2, catches at 5 s were equal to catches at 15 s for some species (e.g., longnose dace) while for other species catches at 5 s were lower but predictive of catches at 15 s (rainbow trout juveniles). From this work we were able to determine which of the catches of the 12 species in our dataset were effected by effort and were able to generate correction factors for these species. Results of this work are currently under review and once peer reviewed will be incorporated into the HabProgs database. This will enable users to standardize catch estimates to a single effort criteria.

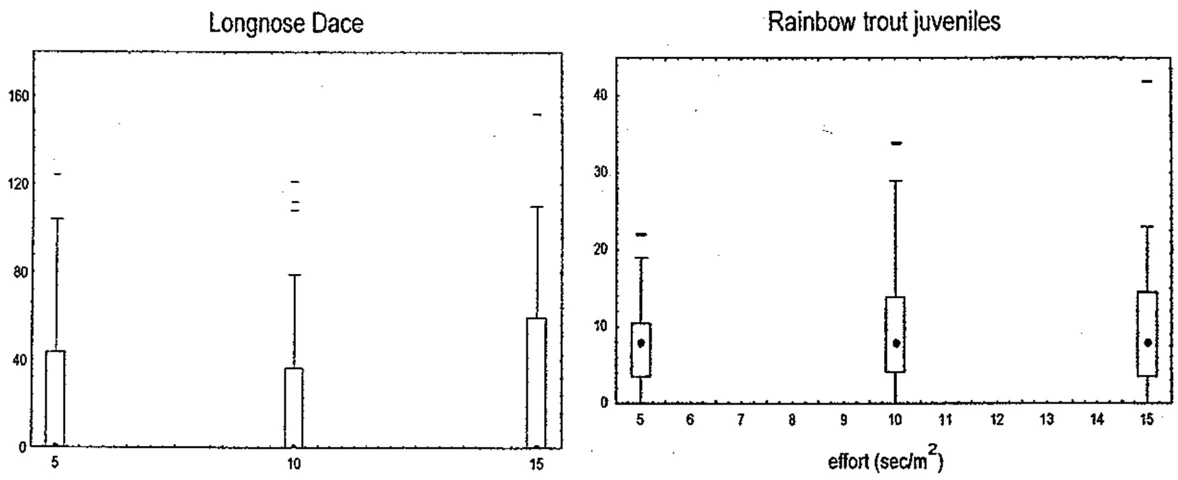


Figure 2: Total catch of rainbow trout and longnose dace with varying effort.

The possibility of sampling many more, but smaller sections of streams as a means of increasing accuracy and efficiency in calculating population estimates for a study area is being investigated by the Ontario Ministry of Natural Resources, Picton (L. Stanfield, pers. comm., OMNR, 613-476-8777).

2.0 LITERATURE CITED

De Lury, D.B. 1958. The estimation of population size by a marking and recapture procedure. Journal of the Fisheries Research Board of Canada, 15:19-25.

Jones, M.L. and J.D. Stockwell. 1995. A rapid assessment procedure for enumeration of Salmonine populations in streams. North American Journal of Fisheries Management, 15:551-562.

Mahon, R. 1980. Accuracy of catch-effort methods for estimating fish density and biomass in streams. Environmental Biology of Fishes, 5:343-360.

Schnute, J. 1983. A new approach to estimating removal by the removal method. Canadian Journal of Fisheries and Aquatic Sciences, 40:2153-2169.

Stanfield, L., Lester N, and B. Kilgour, 2002. The influence of electrofishing effort, season and sampling area on fish catches in wadable streams. In review

Zippen, C. 1958. The removal method of population estimation. Journal of Wildlife Management, 22:82-90.

North American Journal of Fisheries Management 15:551–562, 1995

A Rapid Assessment Procedure for the Enumeration of Salmonine Populations in Streams[1]

MICHAEL L. JONES

Research Science, and Technology Branch, Ontario Ministry of Natural Resources Glenora Fisheries Station, Rural Route 4, Picton, Ontario, K0K 2T0, Canada

JASON D. STOCKWELL

Department of Zoology, University of Toronto, Erindale College Mississauga, Ontario, L5L 1C6, Canada

Abstract.—Population enumeration is a key component of fisheries investigations for riverine salmonines. We examined the reliability of a rapid population assessment technique for stream salmonines that depends on a single episode of electrofishing rather than traditional multiple fishing, removal, or mark–recapture methods. We show that for 12 sites sampled in 1992 on Wilmot Creek, a small, coldwater tributary to Lake Ontario, the catch of salmonines from a single electrofishing episode predicted the population estimate obtained with a more time-consuming multiple-pass removal method. When we collected similar data from eight additional Lake Ontario tributary sites in 1994, the relationship was not significantly different from that obtained in 1992. We also conducted nonparametric analyses of covariance on subsets of these data and found no significant differences in the above relationship for rainbow trout *Oncorhynchus mykiss* versus brown trout *Salmo trutta* or for age-0 versus older fish, although the statistical power of the species comparison was low because of the small number of samples containing brown trout. The regression model developed from the pooled 1992 and 1994 data predicts population sizes that only slightly exceed estimates obtained with the removal method at 12 other sites throughout southern Ontario. We use the rapid assessment technique to obtain rainbow and brown trout population estimates with measurable precision for an entire catchment and show that the error in our site-specific estimator of population size is small relative to among-site sampling error, even though our sampling fraction is relatively large compared to that of typical surveys. Finally, we provide suggestions to potential users of the proposed methodology regarding data collection and analytical techniques.

Stock or population assessment is an integral part of virtually all fisheries management activities, and developing efficient methods for reliable stock assessment poses a continuing challenge to the fisheries management profession. Much of the attention devoted to stock assessment techniques has focused on lake and ocean-dwelling fish stocks of commercial and recreational significance (e.g., Hilborn and Walters 1992). The challenge is equally great for stream fish populations, even though methods may be quite different. In this paper we present the results of a study to evaluate a rapid technique for assessing stream salmonid population size, based on electrofishing methods.

Stream salmonid populations have most commonly been quantitatively assessed with one of three techniques: mark–recapture, removal methods, or direct counts via snorkeling. Several authors have compared these methods (Heggberget and Hesthagen 1979; Peterson and Cederholm 1984; Slaney and Martin 1987; Zubik and Fraley 1988; Heggenes et al. 1990; Rodgers et al. 1992) without any clear resolution emerging as to which is most reliable or accurate. As a result, logistical considerations and the nature of the questions asked usually determine which method is used. For example, snorkeling may be the preferred method when details of microhabitat distribution are of interest, whereas mark–recapture is perhaps the best approach when a population estimate is required for a relatively large area containing a relatively sparse population.

In smaller streams where isolating a section of the stream with block nets is practical, the removal method has been the most commonly used technique. With this method, successive fishings of a closed population without replacement should result in a pattern of declining catches. Several models have been developed that allow estimation of the population at the site from this pattern (e.g., DeLury 1947; Zippen 1956; Carle and Strub 1978; Schnute 1983). For these models to be applied, it is desirable, if not necessary, to have at least three successive fishings during which a fairly large

[1] This is contribution 95-05 of the Ontario Ministry of Natural Resources, Aquatic Ecosystems Research Section, Maple, Ontario.

fraction of the population is caught. The validity of these models depends on certain assumptions, most notably that all fish in the population are equally catchable and, with the exception of Schnute's (1983) model, that catchability does not vary from one fishing to the next. As a rule, these requirements (minimum of three fishings, high average probability of capture, equal catchability) necessitate a fishing procedure that is slow, methodical, and consistent; they are reflected in the electrofishing procedures guidelines issued by the Ontario Ministry of Natural Resources. Consequently, for routine quantitative electrofishing surveys in southern Ontario streams, a crew can rarely sample more than one site per day.

Often, population estimates for salmonids in streams are obtained to provide an indication (index) of salmonid production or habitat quality in an area with data from a "representative" sample of sites. In such cases, the degree to which the information gathered accurately reflects the conditions of the area as a whole depends upon the extent to which the sites sampled constitute a statistically valid (i.e., random) sample of the area. Further, the precision of index depends on the number of samples collected, together with the precision of the estimates from the individual samples. Generally, the practical limitations of the sampling methodology (one site per day) combined with limited resources lead investigators to nonrandomly select a small number of "index" sites and assume they represent the overall area (e.g., watershed, basin, stream reach). If it were possible to obtain population estimates with measurable precision more rapidly, more sites could be sampled with the same resources, and investigators might be prepared to adopt a more statistically valid sampling scheme.

Hankin (1984) showed that this type of stream electrofishing survey is a two-stage sampling procedure. The first stage involves choosing the sites (or sampling units) to visit from a population of possible sampling units. The variance at this stage is a consequence of the sampling fraction (i.e., the proportion of all possible sampling units that are actually visited) and the variability among individual sampling units. The second stage involves estimating the statistic of interest (in this case population size) at each site. Here the variance is determined by the precision of the estimation method. Hankin pointed out that in most cases the first-stage variance is likely to be far greater than the second-stage variance. Thus, there is a greater payoff in reducing first-stage variance (by increasing the sampling fraction) than second-stage variance (by increasing the precision of the estimator). Hankin and Reeves (1988) used these ideas to develop a sampling procedure for stream population estimation that involved applying a relatively fast, if imprecise, site-specific estimation method (diver observations) at a relatively large number of sites, thereby increasing the precision of the catchment-level population estimate compared to that obtained from a more precise site-specific estimator (e.g., removal method) applied to fewer sampling units.

In this paper we consider another approach to optimizing sampling strategy by comparing the catch obtained from a single electrofishing episode to the population estimate obtained from the traditional removal method. The purpose of the successive fishings used in the removal method is to estimate the catchability, or probability of capture, of an individual fish from the population at that site. All of the models developed for the removal method are statistical variations on a basic theme of using the total number of fish caught and the estimate of catchability to estimate the total population of fish at the site. But if catchability does not vary greatly among sites, then it could be estimated from a subset of sites and applied to a larger group of sites where only a single fishing episode would be required. More specifically, if a repeatable relationship exists between the single-pass catch and the multiple-pass population estimate for a variety of locations, then this relationship, together with single-pass catch, can be used to estimate the population at other locations more rapidly.

There is little doubt that electrofishing catchability does vary considerably among sites because of differences in species composition, physical characteristics of the site, and specific details of the sampling methods used. We were interested, however, in determining whether conditions could be controlled sufficiently in a particular survey, involving perhaps only a restricted range of stream types, certain target species, and a crew or crews trained to operate in a consistent, methodical fashion, such that variations in catchability among sites would be small enough to allow development of a single-pass, rapid assessment technique.

Methods

First, we examined existing data sets from southern Ontario streams, for which removal method estimates of salmonid abundance or biomass were available, to determine whether the first-pass

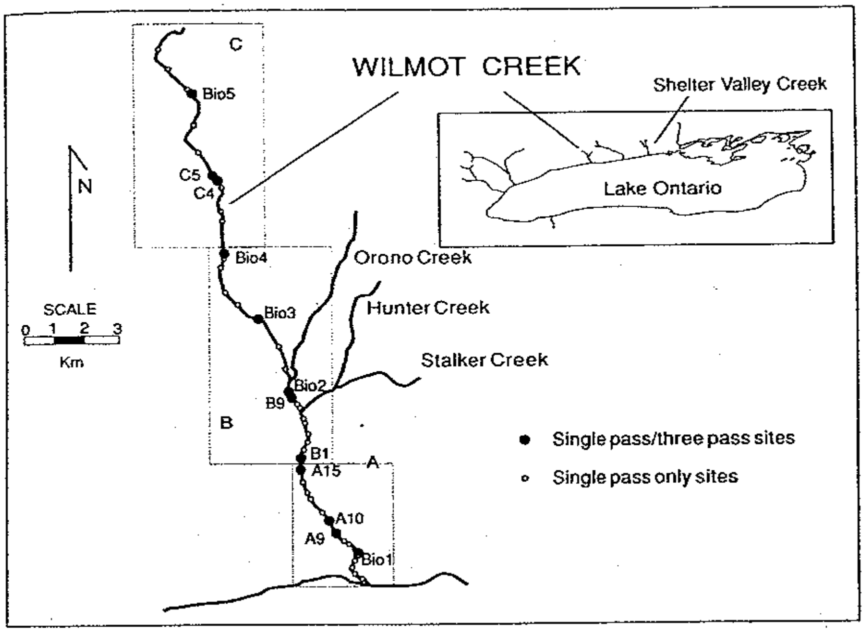


FIGURE 1.—Map showing Wilmot and Shelter Valley creeks (inset) and the single-pass sampling sites in Wilmot Creek. Zones A, B, and C correspond to the zones used to develop the whole-stream population estimates. The 12 sites used to develop the rapid assessment regression model are individually labeled for cross-reference to Table 1.

catch from multiple-pass sampling provided a consistent indicator of population size for a range of sites and species. We obtained biomass estimates (population × mean weight) for brook trout *Salvelinus fontinalis* from 48 sites throughout southern Ontario from Bowlby and Roff (1986). Although biomass estimates incorporate an additional source of error (i.e., variance in fish weights) the pattern we sought should not have been qualitatively affected by this factor. We also examined our own data (1988–1992) for rainbow trout *Oncorhynchus mykiss* and brown trout *Salmo trutta* from six sites at Wilmot Creek (Figure 1), a small (base flow, 0.5 m^3/s), coldwater stream draining into the northwest end of Lake Ontario (M. L. Jones, unpublished data).

Existing data sets such as those cited above do not allow a statistically valid comparison of the catch from a single electrofishing episode to the population estimate for the same site, because the latter is not measured independently of the former. To examine this relationship further, we sampled 12 sites (each approximately 50 m long) on Wilmot Creek during summer 1992. These 12 sites were chosen to cover the range of habitat characteristics (e.g., headwaters to lower reaches, limited instream cover to extensive cover) seen in this stream. The top and bottom of each site were chosen to correspond with a crossover point (riffle). Each site was first fished during August 1992 with only a single pass, without block nets. Three to five weeks later each site was sampled again, this time with block nets and the three-pass method. Block nets consisted of 12.5-mm seine netting attached to steel bars pounded into the substrate and weighted at the bottom with small boulders. For both the single-pass and three-pass samples, fish were captured with Smith-Root model VII 12-V backpack electrofishing units. A Smith-Root model GPP shore-based electrofishing unit was used at the most downstream site because of the size and depth of the river at this location.

A crew consisted of three persons, one operating the anode and two netting fish. The crew was instructed to move slowly upstream from the lower end of the site in a sinuous pattern, repeatedly crossing from one side of the stream to the other. All habitat types except deep (>1.5 m) pools were fished; pools too deep to fish were rare in Wilmot Creek. Typically, a crew spent 60–90 min to complete a single pass at a 250-m^2 site.

All captured fish were held live for subsequent processing and release. At the end of each pass, the fish were identified to species. Rainbow trout and brown trout juveniles were divided into age-0 and older groups, based on length-frequency analyses, which clearly indicated an age-group separation at about 100 mm total length (Jones, unpublished data). All trout less than 100 mm were classified as age-0 fish.

To determine the repeatability of the results from the 1992 data, eight additional sites were surveyed during August 1994. Three sites were on Orono Creek, a tributary of Wilmot Creek, and the other five were on Shelter Valley Creek, a physiographically similar stream approximately 70 km east of Wilmot Creek (Figure 1). The methods used during the single-pass and three-pass surveys in 1992 were used in 1994.

We analyzed the three-pass catch data using two removal models (Carle and Strub 1978; Schnute 1983). Both models use maximum-likelihood methods to estimate population size and catchability. The Schnute model allows catchability to vary among passes and provides a method for testing whether such variation is significant. The model also provides an objective criterion for rejecting population estimates with unacceptably large standard errors. However, because the Carle and Strub model (and the related model of Zippen 1956) has been more widely used in the past, we chose to use both models (Carle and Strub, Schnute) and compare the estimates.

We used a type I (predictive) regression model to examine the relationship between single-pass catch and the population estimate obtained using the three-pass method. We judged this model to be appropriate because our objective was to develop a predictive tool, not to explore the underlying relationship between two variables. We used a $\log_e$ transformation to stabilize the variances in both single-pass catches and multiple-pass population estimates. Small sample sizes and evidence of heteroscedasticity in the data led us to use a numerical method, the percentile bootstrap method (Efron and Tibshirani 1993), to estimate confidence limits and for tests of significance.

To test if the slope and intercept of the regression for the 1992 data were greater than zero, the data were randomly sampled with replacement $B = 10{,}000$ times. The regression coefficients were calculated for each of the B bootstrap samples. The probability that a regression coefficient was equal to zero is P = (number of bootstrap samples with coefficient ≤ 0)/10,000. The tests for the slope and the intercept were done independently (i.e., B = 10,000 random samples were drawn to test for a significant slope, then another B = 10,000 random samples were drawn to test for a significant intercept). The same procedure was used to test the regression coefficients of the 1994 data set.

We estimated the 95% confidence limits for the regression by first generating B = 10,000 bootstrap estimates of the regression coefficients. From each pair of regression coefficients we computed the bootstrap-predicted population estimate (Y^*) for each observed value (single-pass catch, X). This yielded a distribution (B = 10,000) of Y^* values for each X observation, from which we obtained the 2.5% and 97.5% quantiles for Y. We used a spline curve to connect the individual quantiles for each observed X to arrive at an approximate 95% confidence region around the regression line.

We also used the bootstrap method to conduct nonparametric analysis of covariance (ANCOVA) to compare the regressions obtained for different years, species, and age-groups. For each data set of interest (e.g., 1992 versus 1994), we generated B = 10,000 bootstrap estimates of the regression coefficients. We then calculated the difference in slopes and intercepts for each of the B samples, which created a distribution of B = 10,000 slope and intercept differences between the two data sets. Comparing years, we used a two-tailed test at $\alpha = 0.05$. If the range of differences from the 2.5th to the 97.5th percentile (i.e., the 95% confidence interval) included a difference of zero, we concluded that the null hypothesis of no difference should be accepted. We found no significant difference between years (see Results); thus, we pooled the data for both years to test for species and age differences. Because these two comparisons used the same data set, we applied a Bonferroni correction for multiple comparisons (α = 0.05/2 = 0.025).

To investigate whether our failure to reject the null hypotheses of common slopes and intercepts for age and species comparisons resulted from small sample sizes, we used the bootstrap method to estimate statistical power for each comparison. To estimate power, we first had to determine an appropriate contrast size. We were interested in detecting a difference that would be of practical (ecological) rather than simply statistical significance. Bowlby and Roff (1986) repeatedly sampled 10 of their sites eight times during a single season and computed the variance in the population estimates (their Table 5). The average coefficient of variation (SE/mean) for the 10 sites was 37%. This is an indication of the variation (measurement error and natural variability combined) associated with repeated measurements of trout abundance at a single site. We used this result to define our contrast size of practical significance by determining slope and intercept differences that would lead to differences in predicted population estimates (over the range of likely values of single-pass catch) that exceeded 74% (equivalent to an approximate 90% confidence interval or 2× coefficient of variation). Essentially we assumed that differences in predicted abundance that were less than the inherent variability associated with measuring trout abundance in streams were not of practical significance. This definition of an appropriate contrast size is admittedly open to debate; there are no completely objective criteria for making such a determination.

To compute power, we repeated the bootstrap ANCOVA procedure 100 times as described earlier. Each time we generated a distribution of slope (or intercept) differences using B = 1,000 bootstrap samples instead of 10,000 to reduce computing time. If the prespecified difference lay beyond the appropriate percentile values (1.25 and 98.75 for $\alpha/2 = 0.025$), we considered the result to indicate a significant difference. According to Efron and Tibshirani (1993), the power of the test for the prespecified difference at the given sample sizes and α value is given by power = (number of significant differences)/100. If the power of a particular comparison was well below 0.8, we determined, by trial and error, the sample size that would have sufficed to test for the prespecified difference with power of approximately 0.8.

During the same period (August–September 1992) we sampled 33 additional sites (approximately 50 m each) in three sections of Wilmot Creek using the single-pass method only. We selected these sites at random from the total population of 50-m sections of Wilmot Creek in which juvenile rainbow or brown trout are found (Figure 1). We used this random sample, together with the regression estimators described above, to estimate rainbow and brown trout population sizes for the entire stream length and to compare the relative contribution to the overall variance of our population estimates made by sampling (stage 1) error and errors in regression estimators (stage 2 error). Again, we used bootstrap procedures to estimate standard errors. To compute a bootstrap estimate for a section of the stream (A, B, or C; Figure 1), we generated a bootstrap sample of single-pass catches from the n sites within that section, and also generated an independent bootstrap estimate of the regression coefficients obtained from the 12 intensively surveyed sites discussed earlier. We then calculated the bootstrap population estimate (Y^*) from:

$$Y^* = N\frac{\sum_{i=1}^{n} f^*(x_i^*)}{n};$$

N is the total number of possible sites in the section, f^* is the bootstrap regression, and x_i^* is the bootstrap sample of single-pass catch at site i. We repeated this procedure B = 250 times and calculated the standard error for that section from:

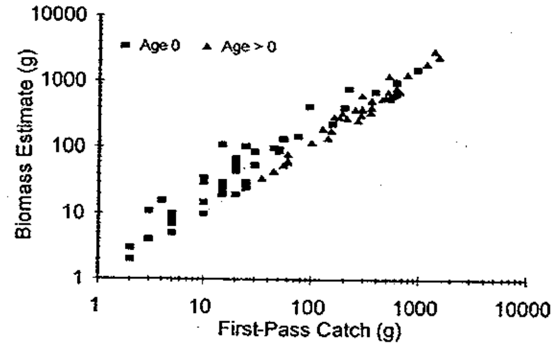
$$\mathrm{SE}_{s+m} = \left[\frac{\sum_{b=1}^{B}(Y_b^* - \bar{Y}^*)^2}{B-1}\right]^{1/2};$$

the ($s+m$) subscript refers to the combined contribution of sampling and model (regression) error to the estimate of the standard error, Y_b^* is the bth bootstrap population estimate, and $\bar{Y}^*$ is the mean bootstrap population estimate. We also calculated a second bootstrap standard error estimate (SE_s) using a fixed (least-squares estimates) set of regression coefficients (i.e., in effect assuming that these coefficients are known without error), again using a bootstrap sample size of 250. The ratio $\mathrm{SE}_s/\mathrm{SE}_{s+m}$ indicates the contribution of stage 1 (sampling) error to the overall uncertainty in the population estimate for each section.

Results

We plotted the brook trout biomass estimate against the first-pass catch for all 48 sites from the Bowlby and Roff (1986) study on a single graph (Figure 2). The result was a remarkably strong relationship between these two values for both age-0 and older brook trout, implying that first-pass catch was a very good predictor of the population estimate or, in effect, that catchability did not vary sufficiently among sites to mask the contribution of the first-pass catch to the biomass (or population) estimate. Data for rainbow and brown trout from six sites on Wilmot Creek sampled over a 5-year period showed a similar pattern (Figure 2).

In the analyses to follow, each sample consists of the catch data for an individual species and age-group at a particular site. We assume that the relationship between single-pass catch and three-pass catch for a particular species and age-group is not significantly affected by the abundance of fish in other species and age-groups and therefore, for this analysis, the samples are independent. We cannot be certain of this independence, but when



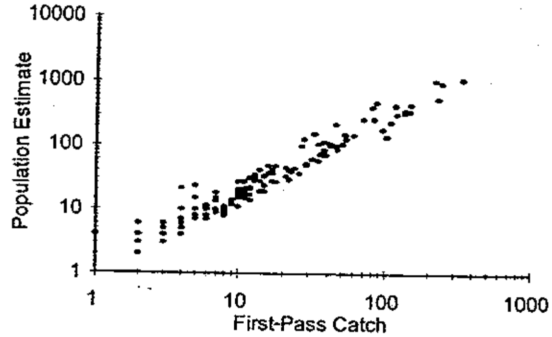


FIGURE 2.—Population estimates plotted against first-pass catch for two southern Ontario data sets: **(upper)** brook trout biomass estimates from 48 sites sampled by Bowlby and Roff (1986); **(lower)** abundance estimates for all ages of rainbow trout and brown trout from six sites in Wilmot Creek sampled during 1988–1992.

we examined the correlation between population estimates for species and age-group i (e.g., age-0 rainbow trout) and the catchability estimate for species and age-group j ($j \neq i$; e.g., brown trout of age > 0), we found no significant correlations for any of the 12 possible cases. The data comprise 40 samples from the 12 sites sampled with both the single-pass and three-pass methods (Table 1). Of the 40 estimates, the Schnute method rejected 9 because of unacceptably large standard errors (i.e., $F(\infty) < F' + 1.92$; Schnute 1983); the total numbers caught over three passes were no greater than 12 in each of the nine cases. We rejected two other samples for which the Schnute method yielded implausible population estimates (9,499 and 53,541) because of nondecreasing, relatively high catches; in both of these cases the numerical algorithm used to obtain the population estimate failed to converge. Of the 29 remaining samples, three had significantly different catchabilities after the first pass. We compared these 29 Schnute population estimates with their corresponding Carle–

TABLE 1.—Catch data summary for rainbow and brown trout caught at the 12 rapid assessment sites on Wilmot Creek.

Site	Age	Catch by pass			Carle–Strub population estimate	Schnute population estimate	Catchability constant?
		First	Second	Third			
Rainbow trout							
A15	>0	17	10	7	40	42.1	Yes
A9	0	17	12	8	47	50.0	Yes
A10	0	25	10	8	48	48.5	Yes
B1	0	69	43	23	164	166.0	Yes
B1	>0	6	2	1	9	9.0	Yes
Bio1	0	38	20	20	108	114.0	Yes
Bio1	>0	13	10	4	30	31.5	Yes
Bio2	0	250	166	142	927	942.0	Yes
Bio2	>0	17	5	6	30	31.0	Yes
Bio3	0	71	48	40	250	260.0	Yes
Bio3	>0	5	6	0	11	11.0	No
Bio4	0	150	70	56	340	342.0	Yes
Bio4	>0	19	11	8	46	47.8	Yes
Bio5	0	61	33	23	144	146.1	Yes
C4	0	54	31	25	147	152.0	Yes
C5	0	41	28	20	123	128.0	Yes
C5	>0	6	5	2	13	14.3	Yes
Brown trout							
A9	>0	2	1	0	3	3.0	Yes
Bio1	>0	2	1	0	3	3.0	
Bio2	0	4	3	0	7	7.0	
Bio2	>0	2	0	0	2	2.0	Yes
Bio3	0	7	4	2	13	13.6	Yes
Bio4	0	17	10	6	38	39.0	Yes
Bio4	>0	14	5	4	24	24.5	Yes
Bio5	0	103	22	12	140	139.7	Yes
Bio5	>0	14	10	0	25	24.0	No
C4	0	3	1	0	4	4.0	Yes
C4	>0	4	4	0	8	8.0	No
C5	>0	11	1	1	13	13.0	Yes
Rejected estimates: rainbow trout							
A9	>0	4	3	4	11	23.0	Yes
A10	0	0	0	2	2	259.0	Yes
Bio5	>0	1	0	0	1	1.0	Yes
C4	>0	6	4	2	12	12.8	Yes
A15	0	42	57	43	534	9,499.0	Yes
B9	0	82	26	39	191	53,541.0	No
B9	>0	2	4	4	10	2.067.0	Yes
Rejected estimates: brown trout							
A10	0	0	1	0	1	1.0	Yes
A15	>0	2	2	1	5	5.0	Yes
A9	0	1	0	0	1	1.0	Yes
B9	>0	3	2	1	6	6.0	Yes

Strub population estimates and found only minor differences (mean difference, 2.0%; range, 4.0–9.5%). We therefore included all 29 estimates in subsequent analyses. When the 29 Schnute estimates were plotted against their respective first-pass catches (Figure 3), the pattern was similar to that seen in the earlier data sets (Figure 2

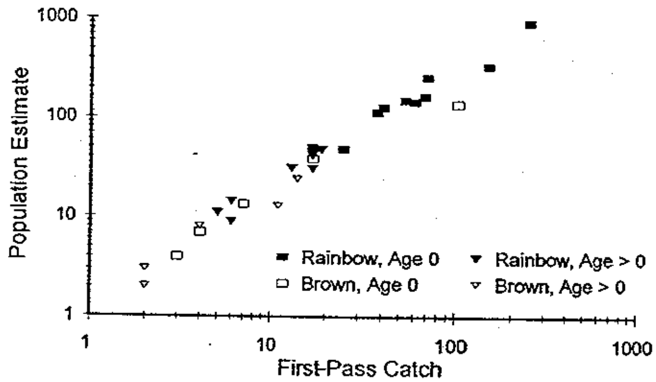


FIGURE 3.—Relationship between the Schnute population estimates for brown trout and rainbow trout and the catches from the first of three passes at each of 12 sites on Wilmot Creek sampled during 1992.

strong qualitative relationship existed between the first-pass catch and the population estimate.

The regression of the Schnute population estimates on the independent single-pass catch (both variables $\log_e$-transformed) was highly significant (Figure 4; $r^2 = 0.86$; slope = 1.02, 95% confidence interval [CI] = 0.86–1.22, $P(\text{slope} = 0) < 0.001$; intercept = 0.78, 95% CI = 0.14–1.31, $P(\text{intercept} = 0) = 0.009$). The nine rejected samples are also plotted in Figure 4; the population estimate for each rejected sample was approximated as 2× total three-pass catch.

The data collected in 1994 yielded 21 additional samples, 16 of which produced acceptable Schnute population estimates. Again the regression of population estimate on independent single-pass catch was highly significant (Figure 5; $r^2 = 0.76$; slope = 1.10, 95% CI = 0.89–1.61, $P(\text{slope} = 0) < 0.0001$), although in this case the intercept did not differ significantly from zero ($P = 0.2581$). Nonparametric ANCOVA revealed no significant differences in either the slopes or intercepts of the regression lines obtained from the 2 years' data: $P(\text{slope}_{92} = \text{slope}_{94}) = 0.6587$; $P(\text{intercept}_{92} = \text{intercept}_{94}) = 0.6688$.

Because we found the 1992 and 1994 data had common slopes and intercepts, we pooled the data to examine differences between age-groups (age-0 versus older fish) and between species (rainbow versus brown trout), again using nonparametric ANCOVA. We did not test for an age × species interaction effect because of small sample sizes. Allowing for the Bonferroni correction, we did not observe any significant differences in slopes or

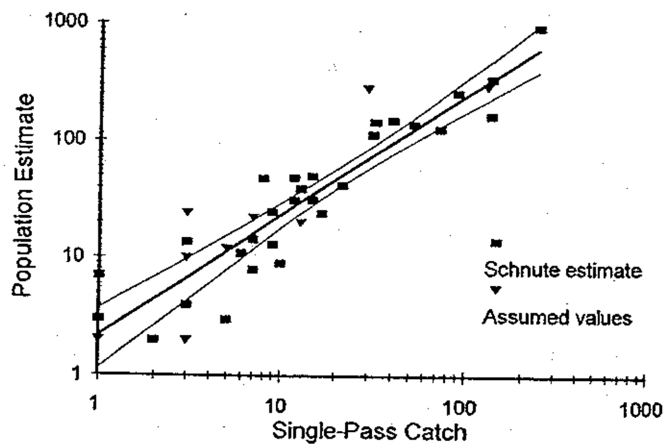


FIGURE 4.—Relationship between the population estimates obtained from a three-pass survey and independent single-pass catches obtained at 12 sites in Wilmot Creek sampled during 1992. Also shown are the 95% confidence intervals (thin curves) for the regression line, and the assumed population estimates (2× total three-pass catch) for sites not included in the regression analysis because the Schnute method led to unacceptable population estimates.

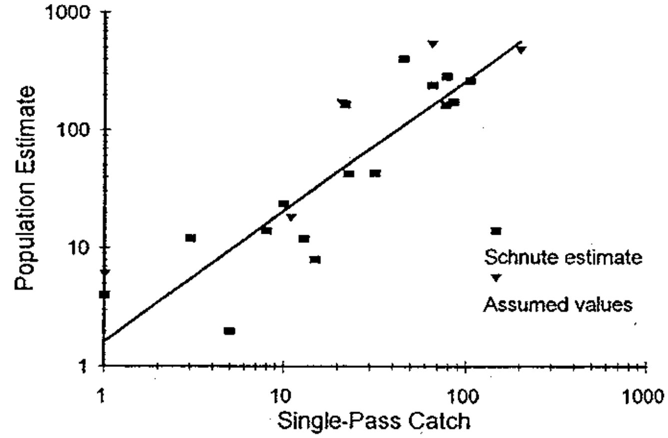


FIGURE 5.—Relationship between the population estimates from a three-pass survey and independent single-pass catches obtained at eight sites in Orono and Shelter Valley creeks in 1994. Also shown are the assumed population estimates (2× total three-pass catch) for sites not included in the regression analysis because the Schnute model led to unacceptable population estimates.

intercepts, although the P value for the comparison of intercepts between age-0 and older fish was 0.0375, marginally nonsignificant (Table 2).

The age-0 versus older fish comparison exhibited high power for the prespecified contrast size and the sample sizes available from this study (Table 2). The species comparison exhibited lower power, particularly for the slope comparison. This is largely a consequence of a small number (15) of samples for brown trout. To achieve a power of approximately 0.8 for the slope comparison, we would have needed a sample size of $N = 18$ for brown trout, given $N = 30$ for rainbow trout (Table 2).

The precision of the population estimates we obtained by applying the regression model derived from the 1992 Wilmot Creek data to the extensive Wilmot Creek survey of 33 randomly chosen sites

TABLE 2.—Results of the nonparametric analysis of covariance tests and power analysis for the species and age-group contrast (a versus b) with combined 1992 and 1994 data.

Parameter	Sample size a	Sample size b	P (a = b)	Contrast size	Power (a − b > contrast size)	Sample sizes for power ≈ 0.8 a	Sample sizes for power ≈ 0.8 b
Age-0 (a) versus age > 0 (b)							
Slope	21	24	0.4443	0.38	0.96		
Intercept	21	24	0.0375	1.31	>0.99		
Rainbow trout (a) versus brown trout (b)							
Slope	30	15	0.1524	0.38	<0.01	30	18
Intercept	30	15	0.3438	1.31	0.30	30	27

varied from 16 to 46% (coefficients of variation SE/mean; Table 3). As expected, most of the variance (69–94%) was due to sampling error, the contribution of uncertainty in the regression parameters was relatively small.

Discussion

Plots of the numbers (or weight) of fish caught in the first pass of a multiple-pass fishing survey versus the resulting population (or biomass) estimate (Figures 2, 3) suggest that the former provides a very strong indication of the magnitude of the latter. This does not necessarily imply that catchability is highly conservative among the sites surveyed. It does suggest, however, that among-site variation in catch overwhelms among-site variation in catchability in influencing the resulting population estimate. If the among-site variation in catch were small but factors influencing catchability varied greatly, one would expect a much weaker relationship than is suggested by Figure 2. In our experience with trout streams in southern Ontario, the among-site variation in catch rates (i.e., first-pass catch) consistently dominates the influence of catchability variation on population estimates derived with the removal method.

Our comparison of three-pass population estimates with independent single-pass catch data from 12 Wilmot Creek sites surveyed in 1992 (Figure 4) confirms this pattern. These data suggest that the catch of rainbow and brown trout from a single pass at a site provides a reli-

TABLE 3.—Population estimates and bootstrap standard errors for three sections of Wilmot Creek surveyed during 1992. The section-level population estimate (Y), the bootstrap standard error of the estimate accounting for both stage 1 and stage 2 error (SE_{s-m}), and the bootstrap standard error due to stage 1 error alone (SE_s) is given for each species, age-group, and stream section. Figure 1 shows the division of Wilmot Creek into the three sections.

Variable	Section A	Section B	Section C
Number of sites surveyed	12	13	8
Area of section (m^2)	50,272	48,414	24,344
	Rainbow trout		
Age 0			
Y	6,294	29,311	10,344
SE_{s-m}	1,882	5,886	1,837
SE_s	1,683	4,163	1,262
Age > 0			
Y	2,689	3,849	816
SE_{s-m}	699	767	275
SE_s	635	628	234
	Brown trout		
Age 0			
Y		1,406	2,461
SE_{s-m}		345	1,007
SE_s		261	950
Age > 0			
Y	265	2,027	1,968
SE_{s-m}	123	497	322
SE_s	101	443	243

and consistent predictor of the trout population as estimated from the traditional three-pass removal method. However, our study does not shed light on the accuracy of either approach as an estimate of the true trout population at a site. Other authors have suggested that the removal method tends to underestimate actual population sizes (Heggberget and Hesthagen 1979; Rodgers et al. 1992). When this is true, our regression model would lead to similar underestimates. Nevertheless, in situations where the removal method is judged to be the most appropriate of the traditional methods, the single-pass approach described earlier offers an efficient and reliable, if somewhat less precise, alternative.

When we repeated the single-pass–three-pass comparison in 1994 at eight new sites, five of which were in a different catchment, the results were remarkably similar. The nonparametric ANCOVA showed no significant differences in either slopes or intercepts, and our power calculations suggested that we were unlikely to have committed a type II error for differences in slopes or intercepts that would be of practical significance. This result further confirms that for streams characteristic of this area of southern Ontario, and provided the electrofishing methods described earlier are followed, the catch of trout from a single-pass survey provides a repeatable predictor of the population present at the site surveyed. The low power associated with our comparison of brown trout versus rainbow trout data suggests, however, that additional brown trout samples should be gathered before we can confidently conclude that a single regression model applies well to both species.

In 1992 and 1994 some of the samples could not be included in our regression analyses because the three-pass catch data did not allow for reasonable estimates of population size with the Schnute model. In all cases this resulted from a nondeclining pattern of catches, suggesting that catchability was relatively low at these sites. As a first approximation, we assumed that the total population for these samples was equal to twice the total catch from all three passes, which is equivalent to assuming a constant catchability of approximately 0.2, well below the mean catchability estimated for the samples that were included in the analysis (0.55). When we plotted these estimates against their corresponding single-pass catches (Figures 4, 5), they did not appear as obvious outliers, suggesting that the regression model may apply acceptably to these sites as well. The failure of catches to decline on successive passes is a fundamental difficulty with removal method estimation models which can be properly resolved only by increasing the number of passes when practical, thereby significantly increasing the time required to complete the survey. Alternatively, one can accept the premise, as suggested by our analysis of these data, that most of the variation in abundance among sites is explained by first-pass catch rather than catchability, and use a calibrated single-pass method.

The data used in this study were collected from only two catchments. To examine whether our results applied more generally to the region within which these catchments are found, we considered data from 12 additional sites throughout southern Ontario and collected as part of an extensive electrofishing survey in 1992. The survey data comprised three-pass removal method estimates of trout populations obtained with the methods described earlier and included 25 samples. The trout population estimates predicted for each site with independent first-pass catch and the regression from our pooled 1992 and 1994 data were very similar to those derived from the Schnute model in all but one case (Figure 6). In general the regression model predicted a slightly higher popu-

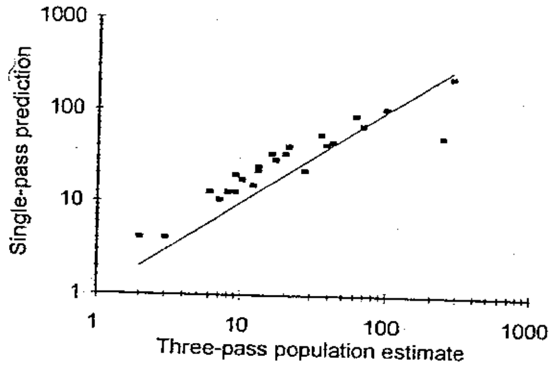


FIGURE 6.—Relationship of trout population estimates predicted from our pooled regression (1992 plus 1994 data) with first-pass catch data to Schnute population estimates derived from all three passes for 25 samples at 12 sites throughout southern Ontario.

lation than the three-pass data would suggest. This might be expected because the regression model was developed from sites where the single-pass catch was obtained without block nets. In contrast, the first-pass catch used in this comparison involved use of block nets. The first-pass catch (and therefore the resulting population estimate) was likely somewhat higher than would have been expected without the nets.

The primary reason for using a rapid assessment procedure as described in this paper is to increase the number of sites one can include in a survey in which a fixed amount of time is available for data collection. When one uses a less time-consuming method for enumerating the population at a site, it will usually be at the expense of precision. This was certainly the case for the single-pass method. Instead of estimating catchability by repeatedly sampling the same site without replacement, we assumed an "average" catchability from a calibration study conducted at other sites or at other times. The critical question was whether this loss of precision at the site level could be justified by a gain in precision for the overall survey. Our example (Table 3) suggested that this was the case; most (69–94%) of the imprecision of our catchment-level population estimates for Wilmot Creek was from sampling error, even though we sampled what would normally be considered a relatively large number of sites (33 sites, ≈13% of the total stream length being surveyed). Clearly, the precision gained by conducting multiple-pass surveys would not be sufficient to justify reducing the number of sites surveyed by two-thirds. Hankin and Reeves (1988), following Hankin (1984), used a similar conclusion to argue for adopting a visual estimation method to expedite stream population surveys. Because resources for such surveys are nearly always severely limited, it is important that rapid assessment techniques such as the one described herein or that of Hankin and Reeves be used wherever possible to increase the coverage (and thus the overall precision) of the survey.

To obtain our catchment-level population estimates for Wilmot Creek, we used fixed-length stream sections as our sampling unit. The single-pass method does not require that fixed-length sites be used, however. Hankin (1984) argued that the accuracy of catchment-level population estimates can be substantially increased by using variable-length sampling units that correspond to natural habitat units. These gains in accuracy are possible when habitat units can be selected such that the variation in fish abundance among units is strongly related to the units' size. Within a particular habitat type (e.g., riffles) this is a reasonable assumption. The single-pass method described herein, properly calibrated for the habitat units being sampled and with the appropriate two-stage estimators for unequal-sized sampling units (Hankin, 1984, 1986), could be used as an alternative to visual methods (Hankin and Reeves 1988) to greatly increase efficiency and effectiveness of a population assessment.

The importance of conducting a statistically valid survey to obtain catchment-level population estimates can be illustrated by comparing the results in Table 3 with an earlier, nonstatistical extrapolation of index station density estimates to the whole Wilmot Creek catchment (Jones, unpublished data; Figure 7). The statistically valid estimates are considerably lower, at least for age-0 fish. This is not surprising, as index stations are probably chosen to represent relatively good sites for the species of interest. Thus, the index stations on which the extrapolations were based contained higher-than-average trout densities, in our case particularly for age-0 fish.

Although we believe the rapid assessment approach described herein is worth considering for other stream salmonine assessments, we strongly discourage users from simply applying our regression model. Other electrofishing crews, using different equipment and investigating streams with different physical characteristics, could well experience significantly different average catchabilities. Thus, population estimates obtained with our regression equation would be in error. Instead we suggest that individual crews calibrate their pro-

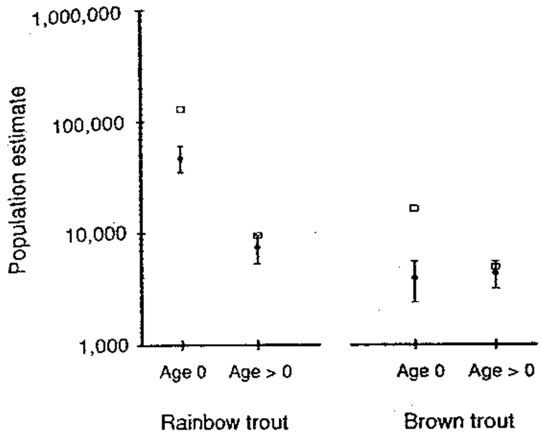


FIGURE 7.—Comparison of the estimated population size for age-0 and older rainbow and brown trout as derived from the current study (solid circles, standard error bars shown) and from an earlier, subjective extrapolation from five index stations on Wilmot Creek (open squares). The index stations are Bio1-Bio5 and shown in Figure 1. We obtained the latter population estimates by dividing the entire length of the creek into sections presumed to be represented by each of the index stations and then scaling up the population estimate for each station according to the ratio of the area of the section to the area of the station.

cedures by following the methods described in this study to develop a predictive regression based on independent single-pass catch data, for either equal-sized stream sections or unequal-sized sections whose boundaries are defined according to natural habitat units.

As a general rule, the greater the catchability, the more reliable and precise the single-pass method will be. For the 45 samples we used (1992 and 1994 data combined), the average catchability estimated with the Schnute method was 0.55 (i.e., more than 50% of the estimated population was caught in the first pass). Therefore, electrofishing crews should be encouraged to move slowly and systematically through the site, making every effort to capture as many fish as possible. Moreover, we suspect the single-pass method will not be as effective a quantitative assessment tool for species which are more difficult to catch (i.e., have lower catchability), such as sculpins *Cottus* sp. or darters *Etheostoma* spp. For stream salmonine assessments, however, we expect the rapid assessment technique described in this paper could provide a useful and reliable assessment tool, provided careful attention is given to consistency in methods.

Acknowledgments

Several individuals assisted with the collection of the data for this study, most notably M. Stoneman, L. Stanfield, J. Tyminski, S. Good, and S. Croke. We thank Jim Bowlby for permission to use the data from Bowlby and Roff (1986) for Figure 1. R. Tibshirani and J. Brunner provided valuable statistical advice. G. Sprules, N. Collins, J. Imhof, N. Lester, and three anonymous reviewers all provided helpful comments on an earlier draft of this paper.

References

Bowlby, J. N., and J. C. Roff. 1986. Trout biomass and habitat relationships in southern Ontario streams. Transactions of the American Fisheries Society 115: 503–514.

Carle, F. L., and M. R. Strub. 1978. A new method for estimating population size from removal data. Biometrics 34:621–630.

DeLury, D. B. 1947. On the estimation of biological populations. Biometrics 3:145–167.

Efron, B., and R. Tibshirani. 1993. An introduction to the bootstrap. Chapman and Hall, New York.

Hankin, D. G. 1984. Multistage sampling designs in fisheries research: applications in small streams. Canadian Journal of Fisheries and Aquatic Sciences 41:1575–1591.

Hankin, D. G. 1986. Sampling designs for estimating the total number of fish in small streams. U.S. Forest Service Research Paper PNW-360.

Hankin, D. G., and G. H. Reeves. 1988. Estimating total fish abundance and total habitat area in small streams based on visual estimation methods. Canadian Journal of Fisheries and Aquatic Sciences 45:834–844.

Heggberget, T. G., and T. Hesthagen. 1979. Population estimates of young Atlantic salmon, *Salmo salar* L., and brown trout, *S. trutta* L., by electrofishing in two small streams in north Norway. Institute of Freshwater Research Drottningholm Report 58:27–33.

Heggenes, J. A., Å. Braband, and S. J. Saltveit. 1990. Comparison of three methods for studies of stream habitat use by young brown trout and Atlantic salmon. Transactions of the American Fisheries Society 119:101–111.

Hilborn, R., and C. J. Walters. 1992. Quantitative fisheries stock assessment: choice, dynamics and uncertainty. Routledge, Chapman and Hall, New York.

Peterson, N. P., and C. J. Cederholm. 1984. A comparison of the removal and mark–recapture methods of population estimation for juvenile coho salmon in a small stream. North American Journal of Fisheries Management 4:99–102.

Rodgers, J. D., M. F. Solazzi, S. L. Johnson, and M. A. Buckman. 1992. Comparison of three techniques to estimate juvenile coho salmon populations in small streams. North American Journal of Fisheries Management 12:79–96.

Schnute, J. 1983. A new approach to estimating populations by the removal method. Canadian Journal of Fisheries and Aquatic Sciences 40:2153–2169.

Slaney, P. A., and A. D. Martin. 1987. Accuracy of underwater census of trout populations in a large stream in British Columbia. North American Journal of Fisheries Management 7:117–122.

Zippen, C. 1956. An evaluation of the removal method of estimating animal populations. Biometrics 12: 163–189.

Zubik, R. J., and J. J. Fraley. 1988. Comparison of snorkel and mark–recapture estimates for trout populations in large streams. North American Journal of Fisheries Management 8:58–62.

Section 4:

Physical Habitat and Geomorphology

Section 4, Module 1

Summary procedures and descriptions used for classifying physical habitat[1]

Table of Contents

[1] Authors: L. W. Stanfield, M. Stoneman, B. Kilgour and J. Parish
Summary procedures and descriptions used for classifying physical habitat

1.0 INTRODUCTION

The methods described in section 4 of the OSAP manual produce raw data on many attributes of streams that are useful for characterizing the physical dimensions of habitat. The raw data can be summarized and interpreted in many ways that vary with each study design. There are a number of routine ways that biologists use for these purposes and queries have been developed within HABPROGS for these purposes. This section provides a synopsis of various summary procedures provided Habprogs. Many of these are used in the modelling initiatives described in section 5. These features are accessed through the data summary screen from within HabProgs (Figure 1).

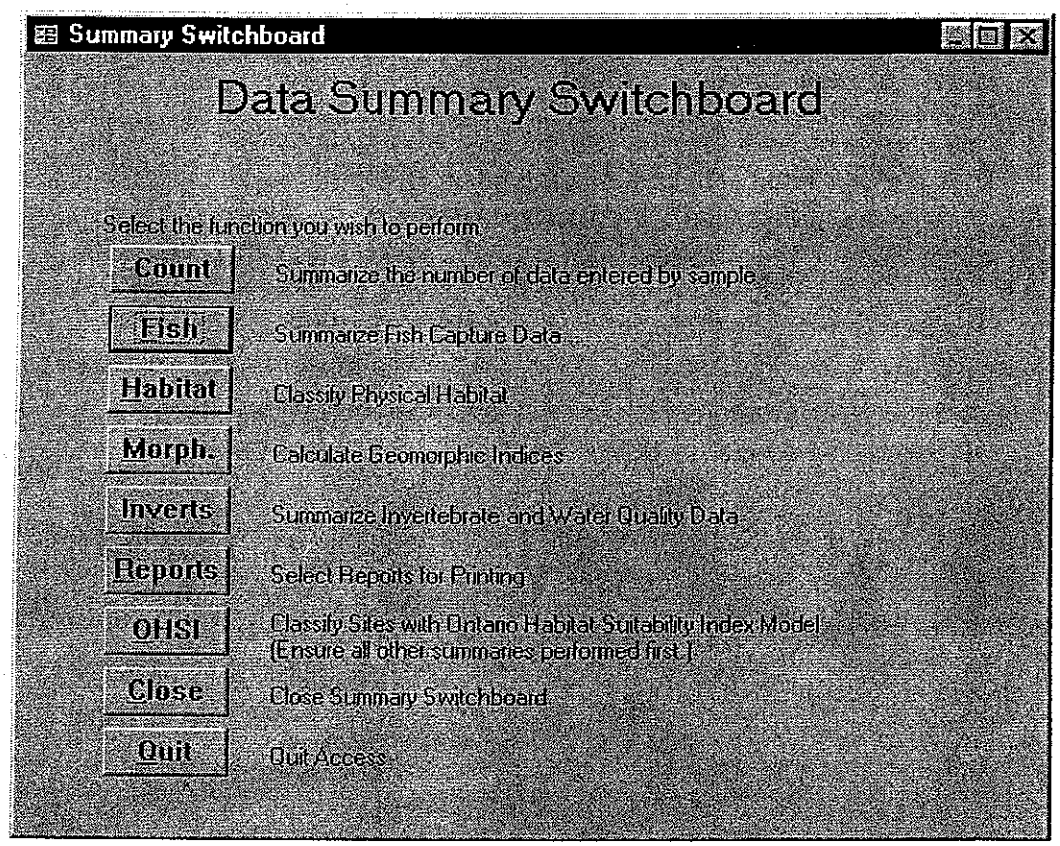


Figure 1: Data Summary Switchboard for database system

2.0 Summary Procedures for data collected using the Rapid Assessment methods (S4.M1 of OSAP)

The data collected using RAM provides summary statistics for channel structure, substrate, cover and bank stability using methods that while being biased are comparable to those used in S4.M2, i.e., the point transect methodologies. Within HabProgs, a summary report has been created to provide a hard copy of the results of these surveys. Access this report by opening the reports window from within the data summary switchboard (Figure 1). Identify the site sample you wish to generate the reports on and either click on the "RAM" report or the "All" button, if you wish to print off a complete set of reports available for that site.

3.0 Summary Procedures for data collected using Point Transect methods (S4.M2 of OSAP)

Data described in this section can be printed off by accessing the Reports window and printing off the channel structure reports for each site sampled. The summary data is stored in a number of tables within HabProgs (see section 6)

3.1 Channel Structure

The term channel structure refers to the makeup of the various flow-depth and cover, substrate of the stream. From the raw data we determine the proportion of the site occupied by various habitat types (as defined by hydraulic head (water velocity) and depth), and the amount and types of cover. To obtain these data, enter the summary switchboard and run the Habitat (classify physical habitat) module.

This function summarizes the physical dimensions of the site (i.e., area, length, width), the break down of points on islands versus stream, and the maximum depth observed. It also summarizes the point level data collected using Section 4 Module 2 into a series of categories of channel structure (i.e., riffles and pools). Summaries of the amount of cover, its distribution, and the types observed are also provided in table format by depth and hydraulic head categories. Finally, a summary of the type of instream and riparian vegetation is produced.

Once you click on 'Habitat' a form will open containing a dichotomous key to habitat type (Figure 1) which is used for summarizing the data using the default classes (i.e., the same features used to develop the habitat suitability models (section 5). Click on the 'RUN' button. After the summary finishes, a notification will appear. Click 'Ok' to close the notification. A graph of the frequency distribution of the habitat types in your database can be viewed or printed by clicking on the graph button. Should you wish to classify the habitat using different criteria simply change the depth categories in each box and then rerun the dichotomous key.

Warning: Do not under any circumstances change these criteria in your master copy of the database. To do so means that you will not be able to apply the OHSIM to your data set without first converting the criteria back, which you are likely to forget to do. If you wish to apply this option, we suggest you make a copy of the database and rename it before proceeding.

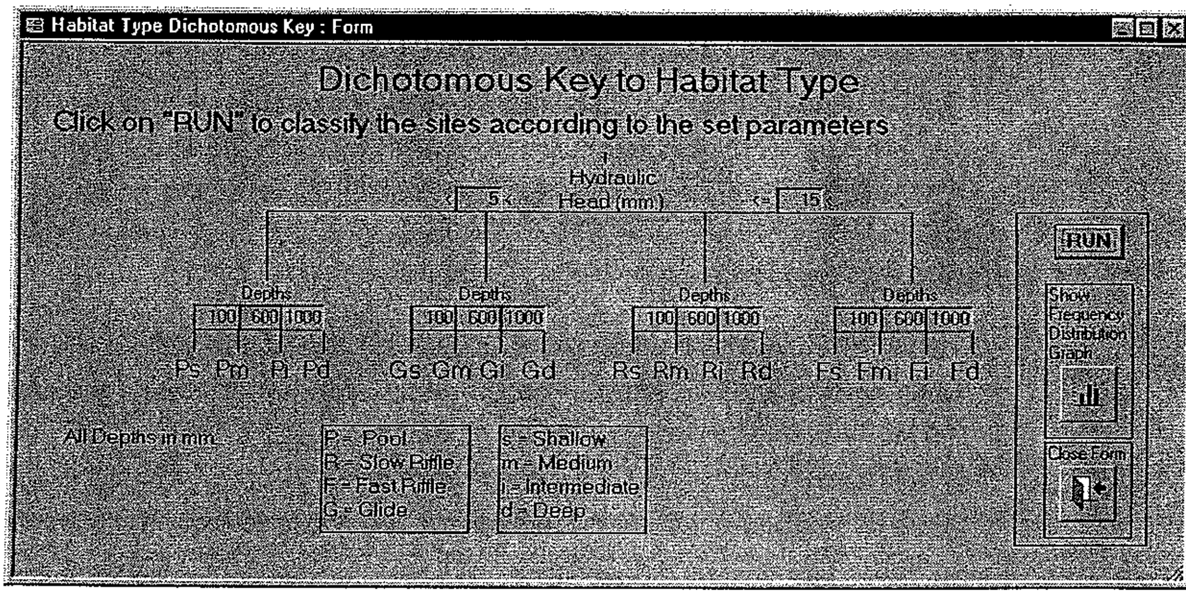


Figure 2: Dichotomous Key for Habitat Type

3.2 Percent Cover

Each observation point is classified by presence and type of cover. These are then totaled for the entire site and are converted to a percentage of the usable area of the site. The percentage of point observations in each habitat type are determined.

3.3 Homogeneity

The homogeneity score of a site reflects the proportion that is represented by relatively similar habitat conditions. To determine the homogeneity score, the most abundant habitat type (as determined above) is chosen. If there are two types that are equally abundant, it does not matter which is used for the calculation. Add up the percentages of all the habitat types that are only one branch away from the dominant type on the habitat hierarchical key (Figure 3). The overall homogeneity score is the sum of the most abundant habitat type(s) and all habitat types within one branch on the dichotomous key. A site with high homogeneity has one major habitat type that dominates and normally means low channel diversity.

Calculating Homogeneity

Example: A station consists of 55% shallow slow riffle, 30% deep glide, 10% medium slow riffle and 5% deep pool.

The dominant type in this case is shallow slow riffle (55%). The adjacent habitat types to shallow slow riffle are deep glide (30%) and medium slow riffle (10%). Thus, the overall homogeneity score is 55+30+10, or 95% or 0.95.

3.4 Substrate

Substrate provides spawning habitat, refuge and cover for fish and their prey. Fish are adapted to spawn in particular ways and to eat particular food; thus they tend to select substrate which provides conditions to suit these requirements. In other words, some species of fish are most abundant in areas of coarse substrate and others are predominantly found in streams with sandy bottoms.

In HABPROGS, stream sites are grouped into three major substrate classes based on the dominant particle size: cobble, gravel and sand. Each of these three classes is subdivided into three classes reflecting varying amounts of fines: low, medium and high. Thus, each stream site is classified into 1 of 9 discrete categories as shown in Table 6. Table 7 summarizes the specific criteria used in the classification:

Table 1: Substrate Classification Categories

Summary procedures and descriptions used for classifying physical habitat

	Substrate Quality			
	Pt. Particle / Max. Particle	High	Medium	Low
Sand D50	≤ 40	Q1	Q2	Q3
Gravel D50	> 40 ≤ 100	Q4	Q5	Q6
Cobble D50	> 100 ≤ 1000	Q7	Q8	Q9
*Bedrock D50	> 1000	Q10	Q11	Q12

*Insufficient samples have been obtained from bedrock-based streams to allow testing of the association of these substrate types to fish community abundance. There are, therefore, no suitability scores for bedrock.

Table 7: Criteria for Rating Substrate Quality

Quality Rating	Criteria
Q1	Max. Part. D50 ≤ 40 and Pt Part. ≥ 10 mm ≥ 40 %
Q2	Max. Part. D50 ≤ 40 and Pt. Part ≥ 10 mm < 40 ≥ 20 %
Q3	Max. Part. D50 ≤ 40 and Pt. Part ≥ 10 mm <20 %
Q4	Max. Part. D50 > 40 ≤ 100 and Pt. Part. ≤ 2 mm = ≤ 25 %
Q5	Max. Part. D50 > 40 ≤ 100 and Pt. Part. ≤ 2 mm = >25 ≤ 50 %
Q6	Max. Part. D50 > 40 ≤ 100 and Pt. Part ≤ 2 mm = >50 %
Q7	Max. Part. D50 >100 and Pt. Part .≤ 2 mm is ≤ 20 % and Pt. Part. ≥ 10 ≤ 64 are ≥10 < 60 %
Q8	Max. Part. D50 >100 and Pt. Part. ≤ 2 mm is >20 ≤ 50 or Pt. Part. ≤ 10 ≤ 64 are > 60 ≤ 80 %
Q9	Max. Part. D50 >100 and Pt. Part. ≤2 mm is >50 or Pt. Part ≤ 10 ≤ 64 are >80 or < 10 %
Q10	Max. Part. D50 >1000 and Pt. Part .≤ 2 mm is ≤ 20 %
Q11	Max. Part. D50 >1000 and Pt. Part. ≤ 2 mm is >20 ≤ 50
Q12	Max. Part. D50 >100 and Pt. Part. ≤2 mm is >50

In cobble bed streams, this classification process is intended to identify an excess of fines (i.e., embedded) or a lack of fines (i.e., armoured).

Assign each site to one of the nine categories.

3.5 Classification of Geomorphic properties of stream using data from section 4 module 2:

From the Data Summary Switchboard, click 'Morph.' to open the Bank Stability Dichotomous Key. Click 'RUN' to run the summary function. This function summarizes a number of channel morphology attributes for the site, including the bank stability rating, the low flow width/depth ratio, and the substrate particle size summaries for both the maximum and point particle sizes.

Definition: Particle Size Distribution Variables

The d16, d50 and d84 particle size measures represent the bed particle size corresponding to the various percents in the particle size distributions. They roughly correspond to the distribution of fines (d16), median (d50), and coarse (d84) materials.

The sorting index is defined as (0.5(d84/d50 +d50/d16)) and provides an indicator of how well sorted the site is. The higher the number, the greater the spread in particle sizes. In geomorphic terminology, this represents a poorly sorted site.

The bank stability ratings are determined by first summarizing the data for each attribute and then applying the dichotomous key shown in Figure 9.3.

3.6 Channel Stability

To quantify dynamic channel stability, a scoring mechanism has been developed incorporating the results of the four parameters (width/depth ratio, bank stability, and sediment sorting and transport potential). Each of these parameters contains several variables that are used to calculate the summary score.

3.6.1 Width/Depth Ratio

Generally speaking, channels with small width/depth ratios move water more efficiently. Flow resistance is reduced in deep narrow channels because they have less bank surface area relative to volume.

Normally the width/depth ratio is measured at the bankfull height of discharge, or channel defining flow (Channel Morphology, Module 12). However, there is a strong relationship between the low flow and the bankfull width/depth ratios.

For each site, calculate the width/depth ratio and apply the scoring mechanism in Table 8:

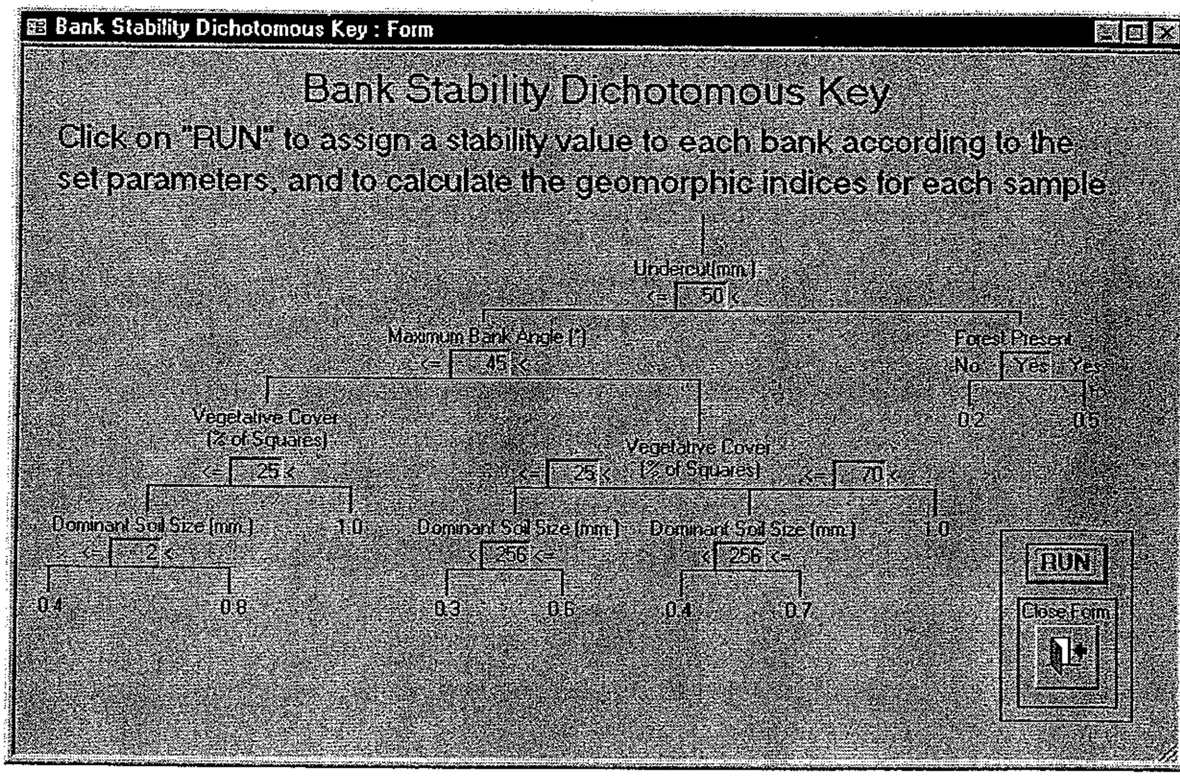


Figure 3: Bank Stability Dichotomous key

Table 8: Scoring the Width/Depth Ratio

Width/Depth Ratio	Variable Score
< 20	0.3
> 20 < 40	0.2
> 40 < 60	0.1
> 60	0

3.6.2 Bank Stability

The amount of erosion on stream banks is a useful indicator of the dynamic stability of each river. The more banks that are in an erosive state the more dynamic is the channel. Within Habprogs a dichotomous key (Figure 3) is used to determine the percentage of the banks within a site that are vulnerable to or are actively eroding. The criteria used for these key include: the composition of the bank (from pebble counts), bank angle (from height to tape measurements), the amount of vegetation on each bank and whether trees are present. Each bank is classified first and then these are summarized for the entire site.

3.6.3 Sediment Sorting

This variable measures the range in particle sizes present at a site. A poorly sorted stream has a diverse mixture of sizes of substrate. This indicates a system that has balance between substrate transport areas (high energy areas) and storage areas (low energy areas). A well-sorted site is one that has particles of similar sizes.

The criterion used to calculate the sediment sorting index is the coefficient of variation for the maximum particle sizes as follows:

$$\text{Sediment Sorting Index} = \text{S.D.}_{MP} / \text{Mean}_{MP} * 100$$

where:

- S.D. = standard deviation of maximum particle sizes;
- Mean = average size for all of the maximum particles, or

$$\text{Mean} = \frac{\sum (\text{ClassSize} \bullet \text{No.Points})}{\text{TotalPoints}};$$

- Class Size is the average particle size for fines, gravel, cobble and bedrock;
- No. Points is the number of dots tallied for any given substrate class; and,
- Total points is the total number of dot tallies, usually 100.

The scoring criteria are presented in Table 9 below:

Table 9: Scoring the Sediment Sorting Index

Sediment Sorting Index	Variable Score
> 120	0.2
> 70 < 120	0.1
< 70	0

3.6.4 Sediment Transport

Sediment transport potential is characterized with the following ratio:

$$\text{Sediment Transport Potential} = \frac{\text{D50 maximum particle}}{\text{D50 point particle}}$$

where,

- D50 maximum particle is the median size of maximum particles at a site; and,
- D50 point particle is the median size of point particles at a site.

Sites for which there is a low ratio are deemed to be in balance with respect to sediment transport. The D50's for each of the maximum and point observations are calculated according to the procedures outlined in Leopold *et al.* (1964). The scoring mechanism is shown in Table 10 below:

Table 10: Scoring the Sediment Transport Potential

Sediment Transport Potential	Variable Score
≤ 12	0.2
$> 12 \leq 36$	0.1
> 36	0

4.0 Summary Procedures for Bankfull Profiles and Entrenchment

Summary reports for data collected using the S4. M3 section of OSAP are accessed through the "Bankfull indices" summary report button. The query determines the cross-sectional area of stream and reports the bankfull width and mean depth. These data are used to report the bankfull width to depth ratio. If velocities are collected along with the bankfull profiles, discharge summaries will also be available, see next section.

4.1 Cross sectional area and mean depth at Bankfull height:

Within Habprogs a number of queries are used to generate summary statistics for the cross-sectional profiles that are created from section 4 module 3 of OSAP. First the cross-sectional area of the river at the bankfull height is determined for each profile. Data from each observation point is assessed as being in the middle of each panel. Next the area within each panel is calculated and added together to create a total cross-sectional area for each profile. These are then averaged for each site to create an average cross-sectional area at bankfull.

5.0 Summary statistics for discharge measurements

Discharge data can be obtained from several related methods and as a result the data is stored in several locations within HabProgs. Data for each can be summarized and printed off as a report. This is accessed from the Main Switchboard through "Summ.", then selecting "Reports" from the Data Summary Switchboard, then selecting the site and sample of interest and choosing the "Discharge" option. The report summarizes discharge in five sections:

(1) *Historic Discharge Data* reports data collected using the Area times Velocity method, however the raw data that generated these observations are not available within HabProgs. For data collected following the National Water Survey of Canada protocol there is a measure of the associated error (i.e., a + or – estimate to relate to the 95 % confidence limits for each sample is available. If you wish more information on the source of the data (check the project description table for the data set.

(2) *Measured Area X Measured Velocity Discharge Method.* This is the standard formula for estimating discharge (Q). It is calculated as a sum of discharge per panel, where discharge per panel is a product of depth, width and velocity[1]. Ideally panel width would be held constant for all observations. However, it has been our experience that this is not always the case. Therefor, our summary procedure has been adapted to accommodate slight differences in panel width.

Discharge (Q) is calculated by weighting each panel observation by the proportion of the total width covered by each panel. Note: If the panels are equal in width across the transect this equation will still derive discharge correctly.

$$Q = \sum_{i=1}^{n} \left[d_i \times \left(w_i \times \frac{w_i}{W_t} \right) \times v_i \right]$$

[1] The assumptions of this technique are that panel widths are similar and that the depth at the midpoint of the panel provides a reasonable measure of the average depth across the panel.

Q = discharge
n = number of observations along the transect
d_i = water depth at centre of panel i
w_i = panel width of panel i
v_i = velocity measurement of panel i
W_t = Active Channel Width (i.e., right active channel location – left active channel location)

(3) *Estimated Area X Estimated Velocity Method* calculates discharge (Q) as a product of stream width, mean depth across the transect, and mean velocity.

$$Q = StreamWidth \times EstimatedMEanDepth \times EstimatedMeanVelocity$$

(4) *Comparative Discharge Estimate Method* discharge data is an estimate of discharge that is assigned by the field crew and is based on their knowledge of discharge values known to exist at other sites with similar conditions.

(5) Volume/Time Method calculates discharge by dividing the volume of the container used to collect water by the time it takes to fill the container. The number of replicates are reported and the discharge is averaged and reported as measured discharge. During these measurements the container of the volume is assumed to remain the same. The amount of each sample that is not sampled (i.e., misses the bucket) is obtained and this estimate (in percent) is used to calculate a "corrected discharge estimate" using the following formula: meaured discharge (m^3/s) * (1+ percent not sampled). Only one value is measured per sample (In theory, since the container volume does not change between replicates, the estimated loss should not change either).

6.0 Literature Cited

Leopold L. B., M. G. Wolman and J.P. Miller. 1964. Fluvial processes in geomorphology. Freeman, San Francisco, CA: 522 pp.

North American Journal of Fisheries Management 18:657–675, 1998

A Comparison of Full-Station Visual and Transect-Based Methods of Conducting Habitat Surveys in Support of Habitat Suitability Index Models for Southern Ontario

LES W. STANFIELD* AND MICHAEL L. JONES[1]

Aquatic Ecosystems Science Section, Ontario Ministry of Natural Resources Rural Route 4, Picton, Ontario, K0K 2T0, Canada

Abstract.—Over the years, ecologists have developed several models that reliably relate stream habitat and fish biomass. Under today's development pressures, managers are demanding that inputs to these models be objectively and repeatably applied by minimally trained crews and at low cost. In this study we compared how repeatable three techniques were at measuring the surficial area of morphological habitat features of wadable streams in southern Ontario. Two crews surveyed 20 sites twice with two survey designs (20 sites × 2 crews × 2 visits × 2 survey designs = 160 visits). For the first two visits, the crews surveyed the entire site (minimum length = 40 m, beginning and ending at a crossover) using a technique that entailed visually determining and then mapping the spatial extent occupied by morphological habitat units. For the next two visits, the crews applied the same visual classification system to a point–transect design (evenly spaced observation points located along evenly spaced transects placed at right angles to the flow. At each point, the crews made measurements of six stream variables: depth, velocity, velocity range, water surface slope, vertical roughness, and substrate. These observations were used to characterize each point as a morphological unit by means of a hierarchical key. We compared the percentage of each habitat type and an overall habitat suitability score observed at each site between visits and crews for each of the techniques. We found that the measured point–transect technique was more repeatable than either of the visual techniques for within-crew and especially between-crew comparisons. We speculate that the lack of repeatability for the visual techniques are largely due to difficulties encountered by observers in consistently identifying transition habitats, such as flats and runs, and to differences among observers in the scale at which they recognize individual habitat units. We estimate that as few as 50–60 point observations within a typical third-order stream site (100 m^2) should be sufficient to characterize the morphological features of the stream for the purposes of determining habitat suitability with acceptable precision (i.e., a coefficient of variation [100·SD/mean] of <10%).

Today's challenges in freshwater fisheries management have heightened the need for valid, practical models that predict the effects of habitat modification on fish populations and communities. Many of these models are intended to be applied at the site level (30–150-m-long sections). Development and validation of these models in turn depends on the existence of consistent and reliable methods for quantitatively measuring the habitat features that constitute the inputs to these models. Stream researchers have found the relative abundance of morphological habitat units (pools, riffles, etc.) to be an important component of fish habitat models (e.g., Binns and Eiserman 1979; Bowlby and Roff 1986). Several methods have been developed over the years to measure morphological habitat units, yet relatively little attention has been paid to the precision or repeatability of these methods. In this paper we evaluate the repeatability of three different methods for measuring these habitat components.

Historically, habitat assessment protocols have tended to use a visual approach to estimate the proportion and location of habitat units throughout the entire site (e.g., Ontario Ministry of Natural Resources 1989). These protocols rely on surveyors to categorize habitat features, which are based on semiquantitative, textual definitions (i.e., Table 1), to estimate the proportion of each habitat type at the site. Platts (1983) cautioned that these techniques were prone to biases and poor repeatability. This warning led

* Corresponding author: lstanfil@gov.on.ca

[1] Present address: Department of Fisheries and Wildlife, 13 Natural Resources Building, Michigan State University, East Lansing, Michigan 48824-1222, USA.

TABLE 1.—Habitat type definitions for the visual technique.

Habitat type	Description
Marginal flat	Little or no flow, located along the margins of streams, usually in backwater areas; very shallow (<10 cm) and tend to have fine substrate, such as silt, detritus or fine sand
Riffle	Relatively shallow (≈1–20 cm), swift-flowing section of river where the water surface is broken
Flat (glide)	Relatively shallow, with visible flow but mostly laminar in nature; minimal turbulence observable; bottom is relatively featureless, but coarser than marginal flats
Run	Deep, swift-flowing sections with turbulent (helical) flow; surface is generally not broken
P1 pools	Slow-moving, deep section of water (>50 cm) with sufficient cover to hide at least one juvenile salmonid from predators
P2 pools	Slow-moving deep section of water (>50 cm); good circulation but no cover present; generally considered good habitat for only very young fish
P3 pools	Slow-moving, moderate depth (10–50 cm) sections of water; no cover, poor circulation; usually removed from the main flow or located in disturbed areas
Plunge pools	Created by water dropping over some object; current usually turns back upstream and fish cover is provided by turbulence and often by undercutting; separated from P1 pools by differing hydrology

to many improvements in field procedures to reduce this risk. Some suggestions have included applying a hierarchical classification system (Hawkins et al. 1993), measuring habitat area and using aids (rulers, twigs, etc.) to assist in determining flow regime (Stoneman et al. 1996), or limiting the number of habitat types classified (Dolloff et al. 1993). Although the adequacy of visual surveys has been challenged from time to time (Binns and Eiserman 1979; Orth 1983), their use has continued, largely because (1) they can be applied fairly efficiently; (2) they have been shown to relate to fish abundance (Binns and Eiserman 1979; Bowlby and Roff 1986, Jones and Stanfield 1993); and (3) it has been assumed that with well-trained and careful crews, accurate and precise site-level descriptions of the habitat features can be obtained (Hawkins et al. 1993; Simonson et al. 1994a).

In contrast, some recent methodologies have used a transect-based approach (as proposed by Platts et al. 1983), which is presumed to provide greater accuracy and precision in the measures of habitat features. For example, Jowett (1993) developed criteria that related to visually determined morphological habitat units based on measurements of depth, velocity, water surface slope, and average substrate size at points along a transect. The instream flow methodologies (Bovee 1982) also rely on objectively measured habitat features to characterize a particular feature of a stream. These methods are generally preferred for collecting data to be incorporated into habitat models. However the intensity of the collection methods generally make these methods cost prohibitive for studies involving large spatial or temporal comparisons.

Some authors have proposed striking a balance between the contrasting approaches by making visual observations within a structured framework, such as along a transect (Simonson et al. 1994a, 1994b) or in quadrants (Wright 1981).

For the past several years, we have been working to develop habitat suitability models for southern Ontario stream salmonines. In the pas (Stoneman et al. 1996), we had applied a site-based visual assessment methodology comparable to that described by Hawkins et al. (1993) as a technique for meeting the data requirements of the model. In this study, we wished to compare the repeatability, or precision, of this method with those obtained with two versions of a transect-based approach, which adapted components of both Simonson et al. (1994a, visually based) and Jowett (1993, measurement based). The three methods will be referred to as the visual site (VS), visual point–transect (VPT) and measured point–transect (MPT) methods for the remainder of this paper.

Methods

We surveyed 20 sites that provided contrast in both morphological habitat features (i.e., riffles versus pools) and the amount and type of cover (i.e., woody cover versus boulders). All sites were located on tributaries to Lake Ontario that drain from the Oak Ridges Moraine through glacial tills and lacustrine deposits (Figure 1). Each site had a minimum length of 40 m and at least one riffle–pool sequence that began and ended where the deepest part of the channel (thalweg) crossed the stream's center line (crossovers). Typically, cross

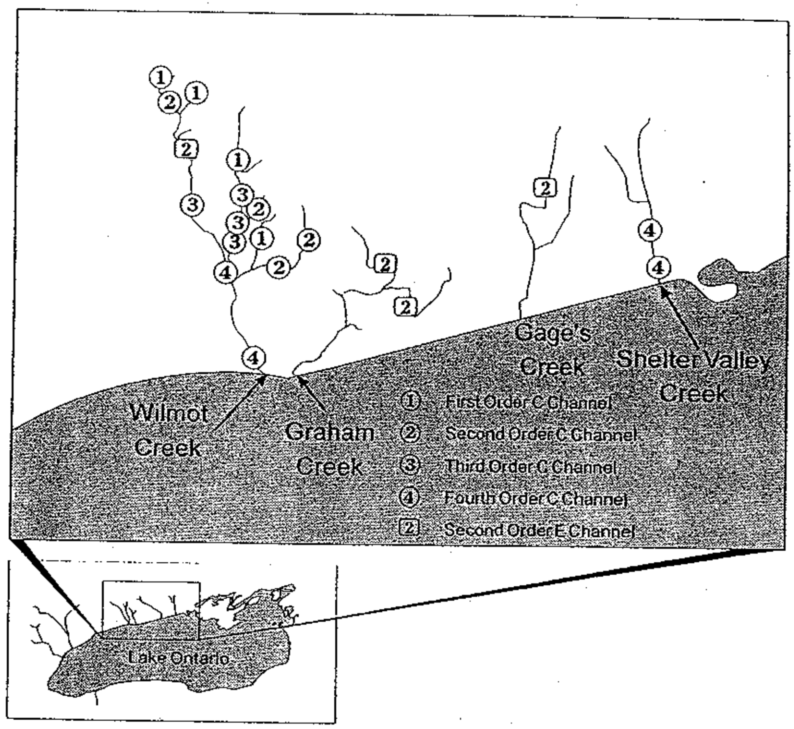


FIGURE 1.—Location of the twenty sites where the habitat data were collected. All sites were on streams draining into Lake Ontario along the north-central shoreline. The C- and E-type channel designations are determined by width : depth ratios, sinuosity, and channel confinement.

overs are associated with shallow areas of uniform depth, such as riffles. On the first visit to each site, the survey crew marked both banks at the upper and lower boundaries.

To assess the influence of stream size on the repeatability of the three methods we sampled four sites each in first-, second-, third-, and fourth-order streams, (as determined by interpretation from a 1:50,000 topographical map), all of which had C-type channels (sensu Rosgen 1996). The C-type channels are characterized by moderate width : depth ratios, sinuosity, and confinement and are by far the most common channel form in southern Ontario. To assess whether there were differences in repeatability between different channel forms, we chose four additional second-order sites with E-type channels, which had low width : depth ratios, high sinuosity, and unconfined flood plains. We visually assessed channel types during site selection following procedures described by Rosgen (1996).

We used two, two-person crews to conduct the habitat assessments. Each crew visited each site twice for both the visual and point–transect approaches, for a total of 160 site visits (20 sites × 2 methods × 2 crews × 2 visits). The crews, beginning at different locations, visited all 20 sites in sequence. All surveys were conducted between May 18 and August 5, 1994. The average number of days between visits by crews were 10 and 18 for the visual and point–transect methods, respectively. This design was chosen to minimize the effect that seasonal changes in base flow would have on physical habitat, at least within one strata. To address daily changes that base flow might have, crews were instructed to carry out all surveys under low-flow conditions. In practice this meant that if the site was turbid or the discharges were noticeably higher than on previous visits, it was not to be surveyed on that day.

Our objective was to mimic a management agency survey. Students with varied educational and practical backgrounds were selected and provided with basic training. The training consisted of two 1-d courses of both lecture and practice, which were supplemented with a written protocol. Each crew assessed all 20 sites twice using the VS method before they were

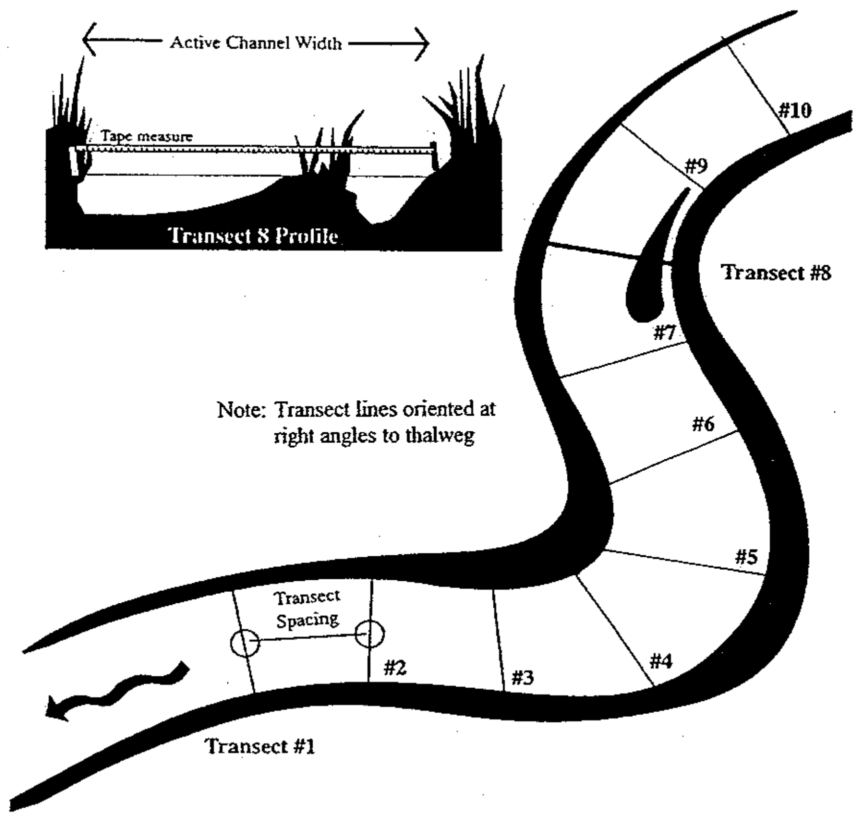


FIGURE 2.—Survey design for the measured point–transect method.

trained in the use of the MPT method. Although it might have been preferable to have one crew use the MPT method first, this was not possible because the VPT and MPT methods were applied simultaneously (for the rationale for this, see below), and thus the crews needed to be familiar with the visual categories before applying the point–transect method.

The visual method of categorizing habitats involved determining morphological habitat units based on a description of the distinguishing features of each habitat type (Table 1). Crews also recorded the abundance of plunge pools, although this habitat type was very rare at our sites and has been left out of the analysis. The two crew members would discuss and agree on a habitat type and its spatial extent and then either transfer the information to a scaled sketch of the site (VS method) or record the habitat type for each point observation (VPT method). To ensure accuracy of scale for the VS method, crews measured wetted-stream widths to the nearest 0.1 m every 2 m along the length of the site and used compass bearings to map around corners. We encouraged crews to use aids such as dropping silt into the stream to help identify flow types (i.e., corkscrew versus laminar) or rulers to verify depths, velocities, and areal extent of habitat types.

The MPT survey involved measuring six variables and the visually determined habitat type at 2–10 uniformly spaced observation points along 10–30 evenly spaced transects that were placed at right angles to the flow (Figure 2). The point–transect survey design varied with stream width to provide 100 observation points for each site. The six variables measured were (1) average and (2) range in velocity over a 10 s interval, (3) water depth, (4) local water surface slope, (5) substrate particle size, and (6) vertical roughness. The crews made vertical measurements (depth, roughness) from the subpavement (defined as the lower part of the active flowing channel) and operationally where substrate particles formed a fairly uniform layer across the stream bottom. Velocity was measured at 0.4 times the depth of the water from th

TABLE 2.—Summary of the correspondence between the point-level habitat type assignment for the hierarchical ke (MPT) and for the visual (VS) interpretation. Data for both crews and visits are included. The hierarchical key assigr ments that are exact are shown in bold italic (matrix diagonal) and those that differ from the visual assignment by onl a single criterion in the key are followed by an asterisk.

Visual assignment or statistic	Hierarchical key assignment: Flat	Marginal flat	P3 pool	P2 pool	P1 pool	Plunge pool	Run	Riffle	Point bar	Total
Flat	***529***	127	244	245*	64	2	60*	207*	6	972
Marginal flat	17	***113***	120*	5*			1	58	8	197
P3 pool	10	156*	***311***	14*	26*			30	9	360
P2 pool	172*	148*	917*	***330***	206*		37	93	13	473
P1 pool	18*	12	69	20*	***657***		26	3	14	995
Plunge pool	1	1	11	1	10	***3***				27
Run	513*	24	41	69	41		***258***	247*	1	434
Riffle	148*	8	42	17	4	3	22*	***309***		383
Point bar		4							***77***	81
Total	1,408	593	1,755	701	1,015	8	404	947	128	6,959
Percent correct	38	19	18	47	65	38	64	33	60	37
Percent within one one key position	98	70	77	87	89	75	88	81	73	85

subpavement (Newbury and Gaboury 1993). We measured water surface slope with a water level (tube length of 0.6 m and 1.25-cm precision) that was placed parallel to the flow. Substrate particle sizes were measured to the nearest millimeter along the median axis (Newbury and Gaboury 1993). Similarly to Jowett (1993), vertical roughness was defined as the vertical distance occupied by any stationary object within the cross-sectional profile that extended from the subpavement to the surface.

We used data from a subset of four sites ranging from first to fourth order to develop a hierarchical key that defined the habitat type present at each point, based on the values measured for each of the six variables in the MPT survey. We divided the 400 observation points from these four sites into groups corresponding to the habitat types judged by the field crews to be present at each point (see criteria in Table 1). Then we examined the frequency distributions of each of the six measured variables, as well as three compound variables (velocity : depth ratio, roughness : depth ratio, and Froude number [see below]) to seek distinguishing attributes of the individual habitat types. Results of this were used to create a hierarchical key to classify points into one of the habitat types listed in Table 1 or as point bars. This key is essentially tautological and therefore cannot be empirically validated. Nevertheless, we examined the extent to which the habitat types assigned to each point by means of the key corresponded to the visual assignment made at the time of the survey. Observation points for which variables could no be measured because of insufficient depth (velocity) or the presence of obstructions (slope were not included in this or subsequent analyses

We assessed the repeatability of the three sampling methods by comparing the absolute dif ferences in the percentage of each habitat typ measured at each site, both between visits an between crews. Unless otherwise stated, data i graphs include all results. To evaluate the repeatability of the MPT data we summarized th point-level habitat type assignments obtaine from the hierarchical key into tables of percen area (one point = one percent of the site area by habitat type. We used the average absolut percent differences (p) for all sites (or a strat ified subset such as first-order sites) between vis its or crews as our indicator of repeatability. Fo example,

$$R_{i,m} = \frac{1}{K} \sum_{k=1}^{K} |P_{i,1,k,m} - P_{i,2,k,m}|, \qquad (1$$

where $R_{i,m}$ is the nonrepeatability index for crev i and habitat type m between visit one and visi two, averaged over K sites. The higher the valu the less repeatable the measurement. We used similar calculation to evaluate the repeatabilit between crews. We did not include data for habitat types that were less than 1.5% of the tota area of the site for all of the observations for eac methodology.

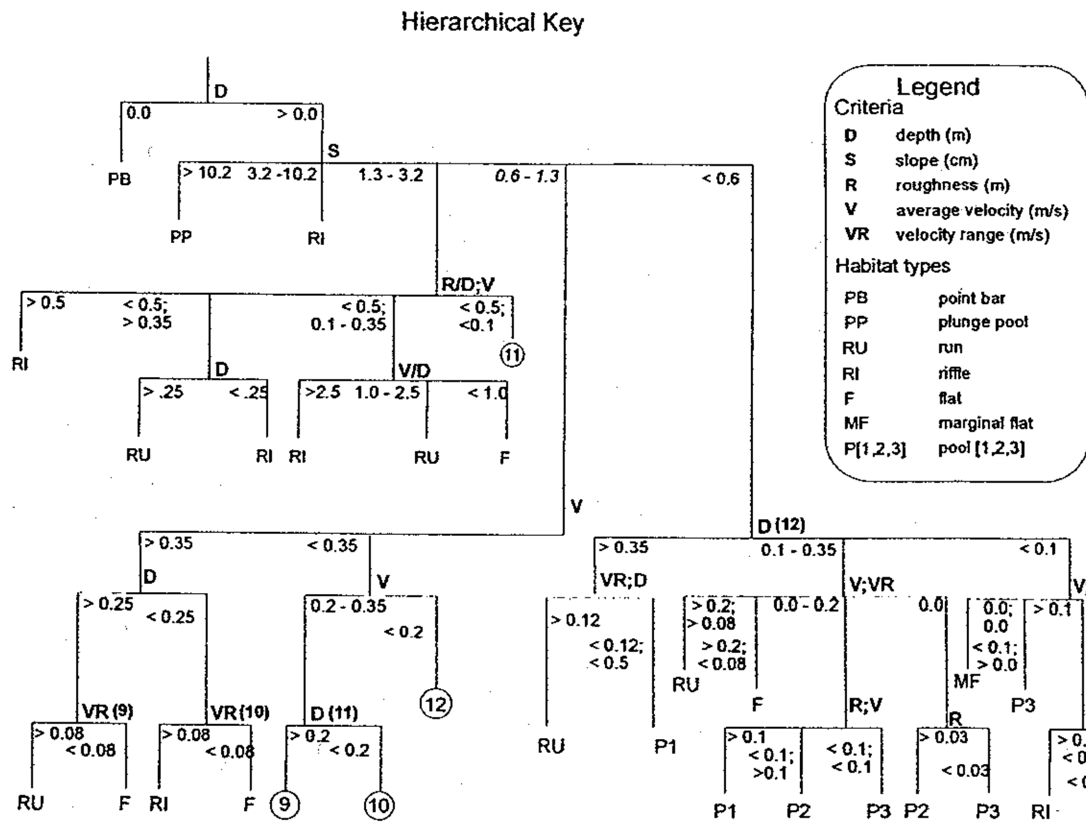


FIGURE 3.—The key used for the measured point–transect method to classify each point observed into one of eight habitat types. At each branch of the key a choice is made about which direction to follow depending upon the value of one or more physical habitat variables for that point. For example, points with water surface slopes greater than 10.2 cm are classified as plunge pools (second level from top, leftmost branch), while those with slopes between 3.2 and 10.2 cm are classified as riffles (second branch from left, same level). All other points at this level require further physical information to be classified (three branches to the right). Circled numbers move classification to a new branch as identified by the number in parentheses.

Differences among crews or visits in measured habitat type frequencies are of greater significance if they can be shown to translate into differences in the measured suitability of the habitat at each site for fish. To further test the repeatability of the three methods in this context, we used the habitat type data to compute a habitat suitability score for each site. We assigned each habitat type a value (referred to as a suitability weighting) between 0 and 1 that we assumed reflected the quality of that habitat type as a nursery area for age-0 rainbow trout *Oncorhynchus mykiss*. We then calculated the habitat suitability index (HSI) score for each site and visit from

$$\text{score}_{i,j,k} = \sum_{m=1}^{M} p_{i,j,k,m} \cdot w_m, \qquad (2)$$

where $p_{i,j,k,m}$ is the percent (by area) of habitat type m observed by crew i on visit j at site k and w_m is the suitability weighting for habitat type m. The weightings for each habitat type were: P1 pool = 1.0; P2 pool = 1.0; P3 pool = 0.5; marginal flat = 0.2; flat = 0.5; riffle = 1.0; run = 0.0. We compared calculated scores among crews and among visits.

We also evaluated whether increased repeatability for the point–transect survey could be obtained by either collapsing the number of habitat types to three (riffles, runs, and pools) or by using a simpler classification system, such as the Froude number. Comparisons were made of the VPT and MPT surveys with these new classification systems for only one crew to remove the crew effect. The Froude number is calculated from the following formulae: $-\ V/(gD)^{1/2}$, where V = velocity (m/s), D = depth (m), and g = acceleration due to gravity (9.81 m/s^{-2}). For this comparison, marginal flats and flats were considered to be pools.

To determine whether fewer (<100) MPT observations would be sufficient to characterize the physical habitat at a site with acceptable precision, we selected five sites that varied from first to fourth order and included both C-type and E-type channels. We computed the standard error of the mean habitat suitability score among the points sampled with the suitability weightings we defined for age-0 rainbow trout (equation 2). We used both the standard textbook formula and a nonparametric bootstrap method (Efron and Tibshirani 1993) with 400 bootstrap samples to compute the standard error. The latter allowed us to reestimate the standard errors with only 50 observations (every second point along a transect) or 25 observations (every fourth point) to see how the standard errors increased as sample size became smaller. Finally, we computed coefficients of variation (CV = 100·SD/mean) for the mean score for each site and for 100, 50, and 25 point observations.

Results

The correspondence on a point-by-point basis between the habitat type assigned for all sites, crews, and visits for the VPT and MPT methods was relatively low (37%; Table 2). When we summarized the frequency with which the key assignment differed from the visual assignment by only one physical attribute (i.e., a single branch on the hierarchy in Figure 3), the correspondence improved considerably (85%; Table 2). This suggests that the visual criteria and the key are approximately similar.

Although the visual and key-based habitat classifications differ on a point-by-point basis, when they were applied to the entire site they produced reasonably similar overall distributions of habitat types for all the sites combined (Figure 4). The visual techniques tended to classify more points as P2 pools (depth greater than 50 cm and good circulation), and the measured point–transect method yielded more P3 pools (depth between 10 and 50 cm and poor circulation). There was general agreement about the relative proportion of fast-water habitats, but the VPT procedure reversed the ordering of riffle and run habitats classified by the other two methods.

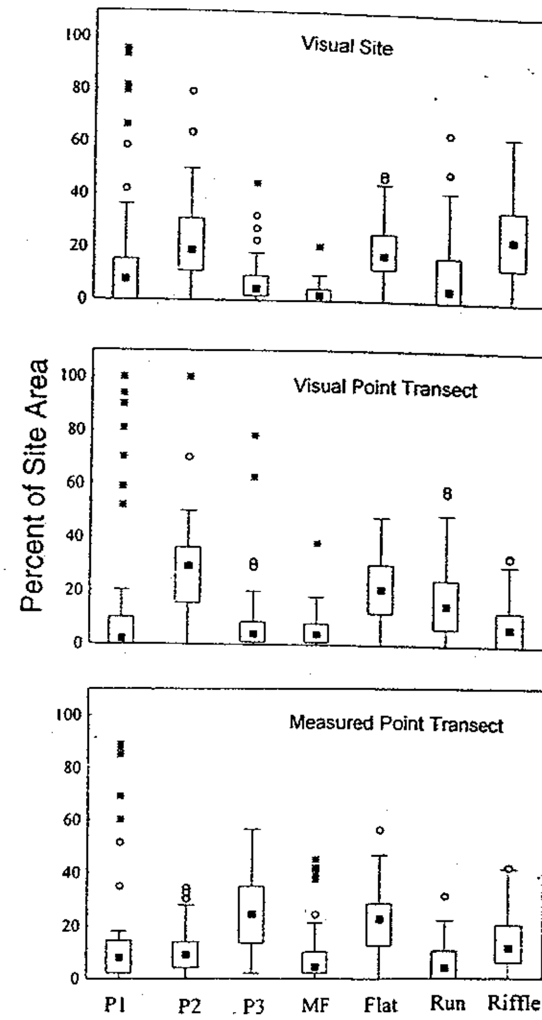


FIGURE 4.—A comparison among all 20 sites of the observed distribution of habitat type abundances (see Table 1 for definitions) between the visual site, visual point-transect, and measured point–transect methods. Data are given as means (solid squares), 25–75 percentiles (bars), and 10–90 percentiles (vertical lines); data outside the 90th percentiles are plotted as circles or asterisks.

For data collected by the same crews, the MPT method provided the lowest index of nonrepeatability for 8 of 14 cases, never exceeding 10% differences, compared with 5 and 6 times each for both the VS and VPT methods (Table 3). The scatter plots (Figures 5,6,7) also show that the MPT method provided a more repeatable means of measuring habitat than either of the visual methods for the individual habitat types.

Patterns in the differences in repeatability of habitat type assessments between the three methods were even more pronounced for the between-

TABLE 3.—Indices of nonrepeatability of the individual habitat-type assessments between visits but within crews, all strata combined. The index is the average difference between visits of the percent of the site area represented by that habitat type, averaged across all sites for which that habitat type was more than 1.5% of the total site area on at least one visit. Note that a higher value for the index implies poorer repeatability.

Variable	P1 pool	P2 pool	P3 pool	Marginal flat	Flat	Run	Riffle
			Visual site method				
Crew 1 index	7.6	8.3	3.7	2.0	10.4	8.9	7.7
Crew 2 index	10.7	16.2	6.7	3.3	7.7	11.9	10.5
Number of sites	19	20	20	20	20	19	18
			Visual point–transect method				
Crew 1 index	21.0	10.2	6.8	8.9	13.0	16.1	9.0
Crew 2 index	6.3	12.2	5.2	3.6	8.5	18.1	5.4
Number of sites	16	19	18	19	19	19	19
			Measured point–transect method				
Crew 1 index	3.6	4.1	5.9	4.6	6.7	4.1	7.3
Crew 2 index	5.7	4.6	6.0	5.3	8.4	4.8	5.8
Number of sites	16	19	19	19	17	15	17

crew comparison (Table 4; Figures 8, 9, 10). The index of nonrepeatability was lower for the MPT method in 10 of 14 categories (Table 4), exceeding the 10% difference in only one category; both of the visual surveys exceeded this value six times. One unexpected result was the between-crew bias that occurred between the percentage of flats and riffles with the MPT method. These habitats are mainly differentiated by their velocity ranges (see Figure 3).

The calculated age-0 rainbow trout HSI scores provide a more integrated (and ecologically relevant) measure of the physical habitat conditions at the site. Again, the repeatability of the MPT method was clearly greater than that of either visual method, both between visits and especially between crews (Figure 11).

There were no major differences in the repeatability of any of the techniques that were related to stream order. On the other hand, the repeatability of the visual technique was lower for E-type channels than for C-type channels (Figure 12), although there was considerable scatter for both channel types. The MPT technique performed equally well on both types of channel.

When we reduced the number of habitat types classified visually by a crew and compared the classifications obtained by using either the hierarchical key or the Froude number, we found that both methods produced very good agreement (83–90%) for all types of pools combined. The hierarchical key (79.5%) performed considerably better than the Froude number (12.9%) at classifying riffles, and there was little difference between the hierarchical key (49.8%) and Froude number (56.7%) for runs (Table 5). Many of the points we subjectively classified as riffles had lower velocities (relative to depth) than would lead to a riffle classification based on the Froude number or were too shallow to measure velocity with the velocity meter, as reflected by the higher percentage of unclassified points for this category.

The CVs of the mean habitat suitability scores for the five sites we chose for sample size assessment with the MPT method were consistently low (range, 5.8%–8.2%). The bootstrap CV estimates for $N = 100$ were very similar to the textbook estimates (5.2–7.6%). For $N = 50$, the bootstrap CV estimates increased to 7.1–10.9%. At $N = 25$, the CV estimates exceeded 10% for all five sites. This analysis suggests that a sample size of 50–60 observations should be sufficient to describe the physical habitat with a precision of no worse than 10%.

Discussion

Our results suggest that the MPT method is considerably more repeatable in assessing the habitat types present at a site than either the VS or VPT method, particularly for between-crew comparisons. This increased repeatability appeared to be consistent across different stream sizes and types (C-type versus E-type channels). There are a variety of reasons why the visual methods might suffer from a lack of repeatability. These include (1)

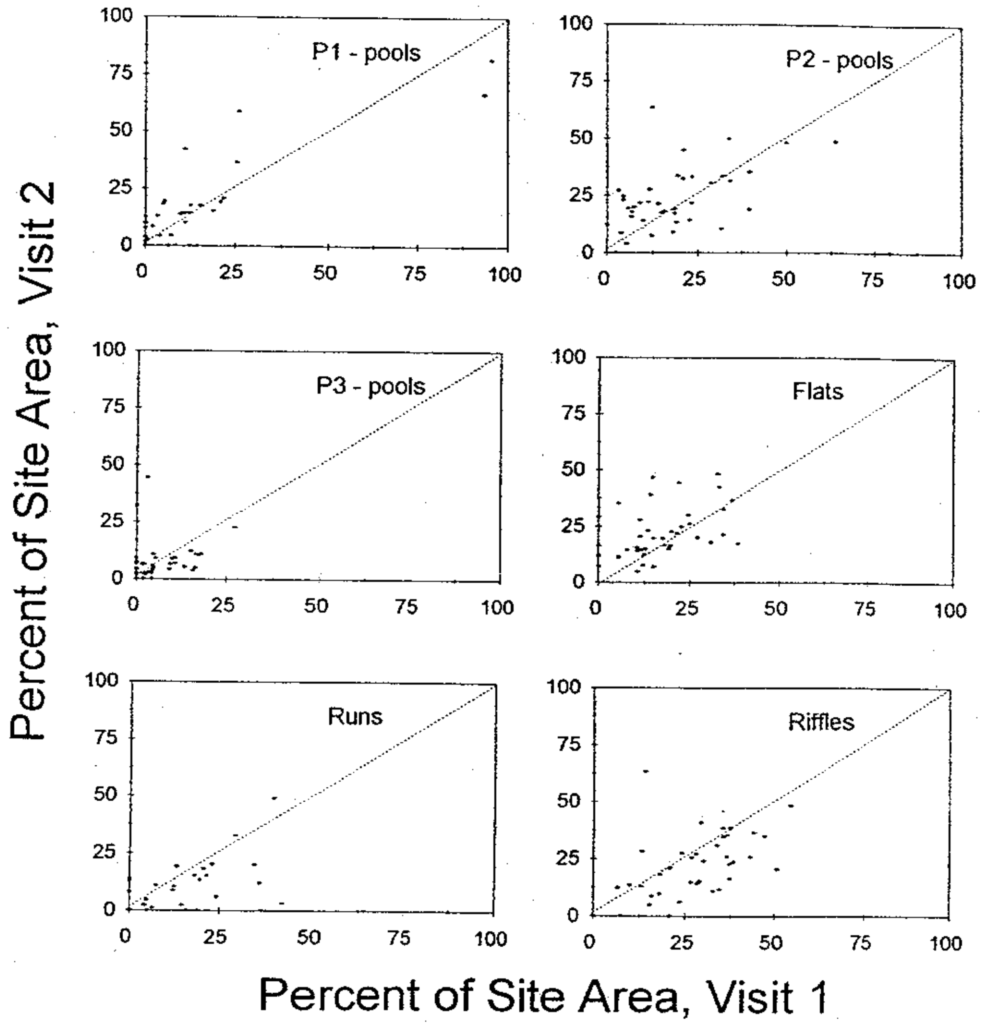


FIGURE 5.—Between-visit comparisons of the observed relative abundances of six habitat types classified by the visual site method. Each point corresponds to the two visits to an individual site by one crew. The graphs include data from both crews.

difficulties in visually identifying transitional habitats and their boundaries, (2) difficulties in maintaining consistent interpretations of habitat types across a range of stream sizes and types, and (3) errors derived from creating and interpreting drawings made at the site.

The unexpected between-crew bias that was observed for surveys conducted with the MPT method may have resulted from the crews using velocity meters of differing makes, with differing sensitivities for recording velocity range.

Reliance on a general, semiquantitative description of the distinction between one habitat unit (e.g., riffle) and another (e.g., run; Table 1) allows differences in perception to influence the consistency of habitat descriptions among observers. Identifying transitional habitat types (e.g., runs, flats) and their boundaries can be particularly difficult. It was apparent that our crews had difficulty consistently identifying the beginning and end of these habitat units. For example, the site depicted in Figure 13 has a long section of low-gradient, moderately fast, relatively shallow habitat. This type of stream section would typically include riffles, runs, and flats. Both crews defined the boundaries between the individual habitat units quite differently on the two visits, and the second crew produced a

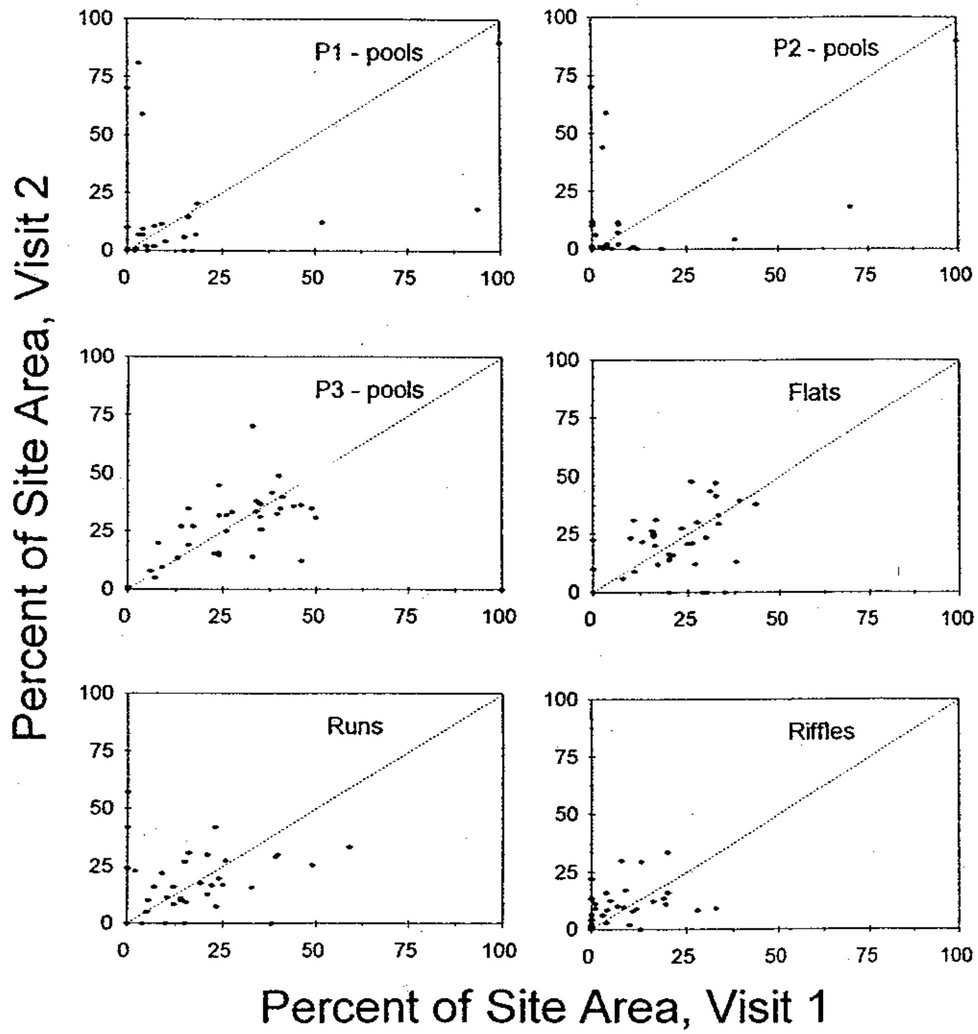


FIGURE 6.—Between-visit comparisons of the observed relative abundances of six habitat types classified by the visual point–transect method. Each point corresponds to the two visits to an individual site by one crew. The graphs include data from both crews.

noticeably less complex description of the habitat on their second visit (Figure 13).

Roper and Scarnecchia (1995) applied a hierarchical classification system to categorize habitats by means of a visual method and also found high variation between observers, particularly for secondary categories and transitional habitats, such as glides. Hawkins et al. (1993) summarized the debate on the issue of habitat classification and also concluded that transitional areas (which they referred to as glides) were particularly difficult to classify.

The failure to improve repeatability by applying the visual method to the point–transect design was, in addition, probably the result of the surveyors being influenced by the surrounding habitat when making their subjective determination of the habitat present at a particular point. With the visual methods, the observers must continually decide issues of scale, for example, whether an individual habitat type is "large" or "important" enough to be considered a distinct unit (i.e., split) or simply judged an unimportant anomaly within a larger habitat unit such as small pools within a riffle complex (i.e., lumped). In a similar study, Jowett (1993) compared the habitat classification he obtained based on point-level physical measurements to a sub-

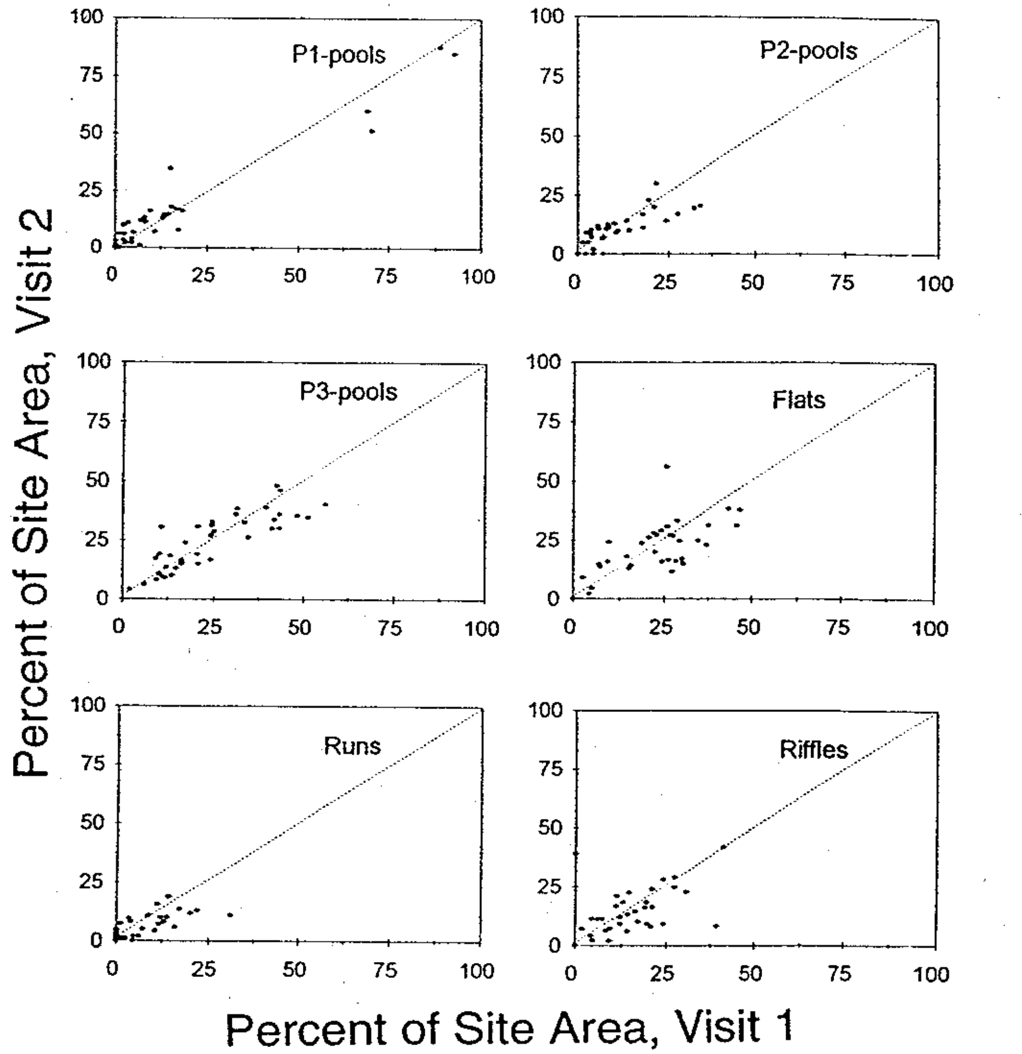


FIGURE 7.—Between-visit comparisons of the observed relative abundances of six habitat types classified by th measured point–transect method. Each point corresponds to the two visits to an individual site by one crew. The graph include data from both crews.

jective (visual) interpretation of habitat type present at each point. At one-third of the points the classification based on physical measures differed from the visual assessment. He concluded that the primary reason for these discrepancies was the influence of surrounding conditions on the visual interpretation of the habitat type present at a particular point. Simonson et al. (1994a) recommend a grid be used at each point to more clearly define the boundaries of the visual point observations.

Finally, the process of accurately mapping the habitat features at a site can be quite challenging, especially at sites with complex habitat fea tures and high sinuosity. As a result, the poten tial exists for a substantial degree of variatio among observers and, thus, low precision i quantitative habitat descriptions obtained wit this approach.

The MPT method, by concentrating on sam pling the physical attributes of a stream that ar guably determine the habitat types that are pres ent, avoids the pitfalls of the visual method. In stead of providing a complete, but subjective description of the habitat present at the site, th MPT technique provides a statistical descriptio

TABLE 4.—Indices of nonrepeatability of the individual habitat-type assessments, between crews but within visits, all strata combined. The index is the average difference between crews of the percent of the site area represented by that habitat type, averaged across all sites for which that habitat type was more than 1.5% of the total site area for at least one visit. Note that a higher value for the index implies poorer repeatability.

Variable	P1 pool	P2 pool	P3 pool	Marginal flat	Flat	Run	Riffle
			Visual site method				
Visit 1 index	6.0	16.2	4.4	4.1	8.1	10.8	16.4
Visit 2 index	9.4	12.7	6.7	2.1	7.3	11.4	13.6
Number of sites	19	20	20	20	20	19	18
			Visual point–transect method				
Visit 1 index	7.8	13.7	9.2	4.8	7.2	12.2	8.4
Visit 2 index	21.7	14.2	9.3	7.0	21.8	15.6	9.5
Number of sites	16	19	18	19	19	19	19
			Measured point–transect method				
Visit 1 index	3.1	3.6	8.3	4.5	10.5	4.9	8.8
Visit 2 index	5.3	4.6	6.1	7.3	7.0	2.7	6.8
Number of sites	16	19	19	19	17	15	17

(sample) of the joint distribution of the physical conditions at the site. The interpretation of these physical measures by means of a consistent classification system (i.e., the hierarchical key) is therefore scale- and context-independent in that it depends only on the physical attributes of the individual point. Also, because the points are geo-referenced, the data can be used to reconstruct a habitat map. The bootstrap results suggest that 60 point observations are more than adequate to provide a reasonably precise description (CV < 10%) of the distribution of habitat features present at a site for second- to fourth-order streams and that even fewer are required (we recommend 50) for first-order streams.

Another potentially valuable feature of the MPT method is that it is robust to post survey changes in the habitat type classification system. In this study, we developed a hierarchical key to classify a point into one of seven habitat types; some, such as riffles and flats, were only subtly different. The MPT technique provides the flexibility of redefining the hierarchical key to reflect improvements in our knowledge of the habitat requirements of fish species or as the species of interest shift. This would not be possible for the visual method, except through the lumping of the predefined categories into a smaller number of aggregated groups. Related to this is the value of the MPT method to habitat supply modeling. For example, if an instreamflow model is developed that predicts changes in stream depth and velocity profiles, the model predictions can be combined with the hierarchical key to yield estimates of changes in habitat composition or suitability with changes in flow. This is in effect how the instream flow incremental methodology process works (e.g., Bovee 1982).

The MPT approach is not a novel idea. More than a decade ago Platts et al. (1983) presented a strong case for the use of transects to standardize habitat data collection. Simonson et al. (1994b) extended the transect approach for application to a reach-level habitat assessment manual and demonstrated how systematic quantitative data of the sort obtained from transect measurements can be used to determine the required sampling intensity. More recently, Wang et al. (1996) have documented the observed precision and accuracy of the various habitat measures described in the manual.

By reducing the overall number of habitat types to three we were able to greatly improve the overall agreement between the visually determined habitat types and those determined by the hierarchical key (83.9%). Our use of the Froude number produced similar overall agreement (63.6%) with the visually determined habitat types, as Jowett (1993) found (66%). The hierarchical key performed considerably better at characterizing the riffle habitats. The poorer performance of the Froude number could be related to several factors. First, there could be subtle differences in the criteria used in the two studies to define riffles. Secondly, the Froude number may not be an appropriate formula for defining microhabitats because it is intended to

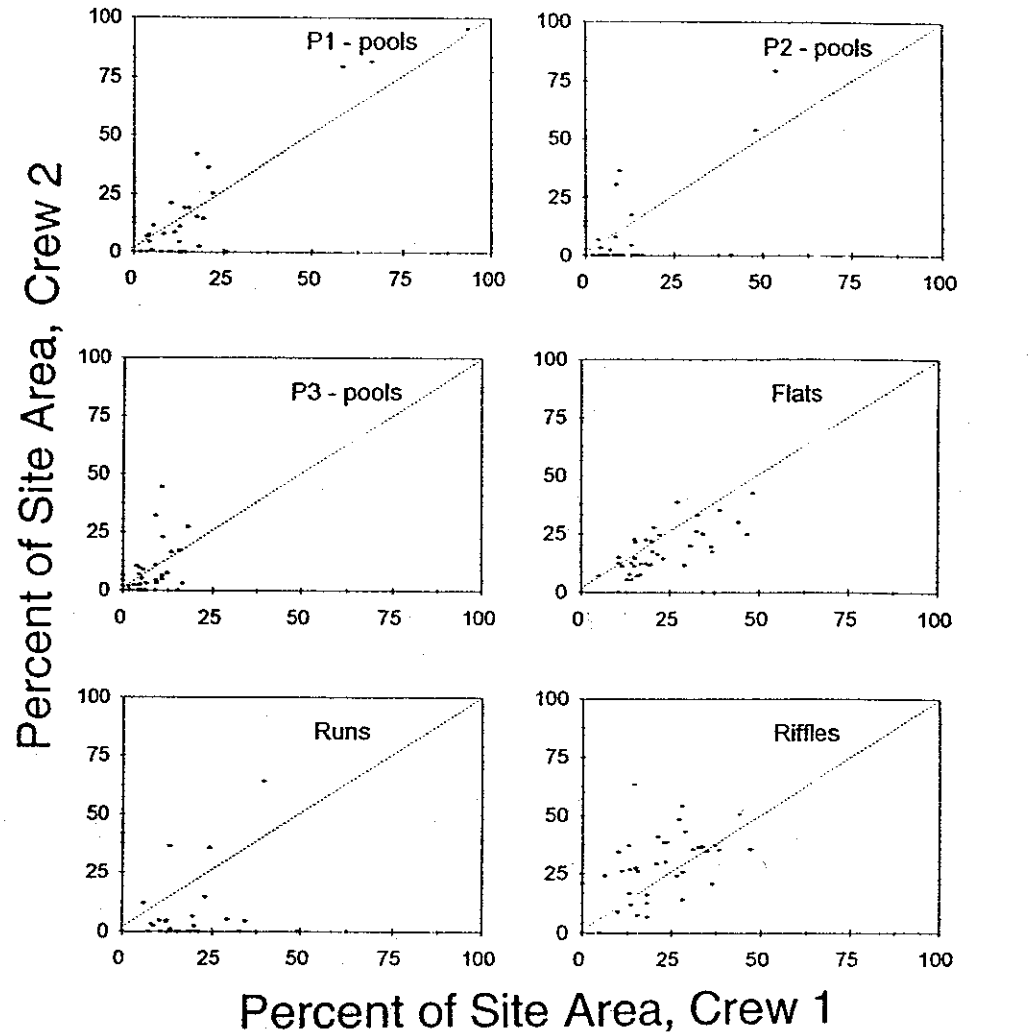


FIGURE 8.—Between-crew comparisons of the observed relative abundances of six habitat types classified by th visual site method. Each point corresponds to the two visits to an individual site by each crew.

characterize flows over a larger section of the river (Newbury and Gaboury 1993). We found that our crews were classifying many microhabitats as riffles, particularly in smaller streams, that would not meet the Froude definition because the velocity relative to depth was too low. We believe these habitats should be called riffles because their characteristics (shallow, relatively fast flow, high degree of roughness) are those associated with the habitat values typically attributed to riffles (invertebrate production, spawning substrate, partial cover for small fish) and, thus, that the Froude number (based only on velocity and depth) is an insufficient criterion for habitat classification across a broad range o stream sizes. Regardless, the likelihood that th poor performance of Jowett's method on our da taset is due to subtle differences in habitat typ definitions lends further support to the value o a more objective system, such as the MPT meth od.

Some of the differences we observed betwee crews and visits for the methods may have bee the result of real changes in the habitat at eac site rather than observation error. The time be tween visits was at least 5 d. For the MPT meth od in particular, relatively small changes in dis charge (and thus both depth and velocity) ma

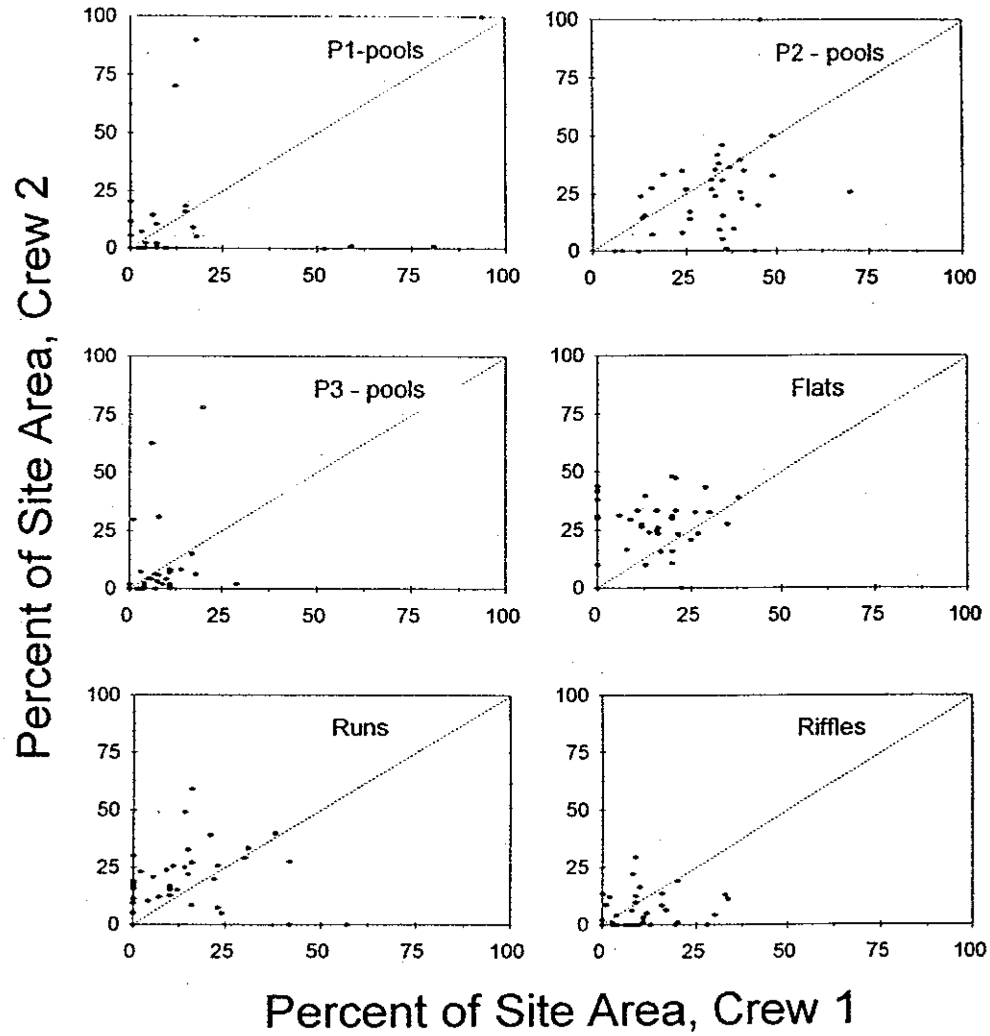


FIGURE 9.—Between-crew comparisons of the observed relative abundances of six habitat types classified by the visual point-transect method. Each point corresponds to the two visits to an individual site by each crew.

have led to changes in the habitat type present at some points. This leads us to conclude that the true repeatability of the methods, and especially the MPT method, is actually somewhat greater than the estimates in Tables 3 and 4 would suggest.

The real test of the value of a habitat assessment and classification system is whether it can be used to explain variations among sites in fish biomass or production that are probably a result of variations in habitat quality. In the past, we have found habitat descriptions based on the visual method to be useful in explaining differences in survival and density of juvenile Atlantic salmon *Salmo salar* (Jones and Stanfield 1993) and the biomass of brook trout *Salvelinus fontinalis*, brown trout *Salmo trutta*, and rainbow trout in southern Ontario streams (Stoneman et al. 1996). It has yet to be shown how much explanatory power of habitat models is lost by lumping habitat types. This is a significant challenge to future habitat suitability modeling.

To conclude, this work was stimulated by the challenges laid down by Hawkins et al. (1993) in their valuable discussion of stream habitat classification issues. We have shown that visual methods have significant problems with repeatability, or precision, especially when different crews are compared. We ad

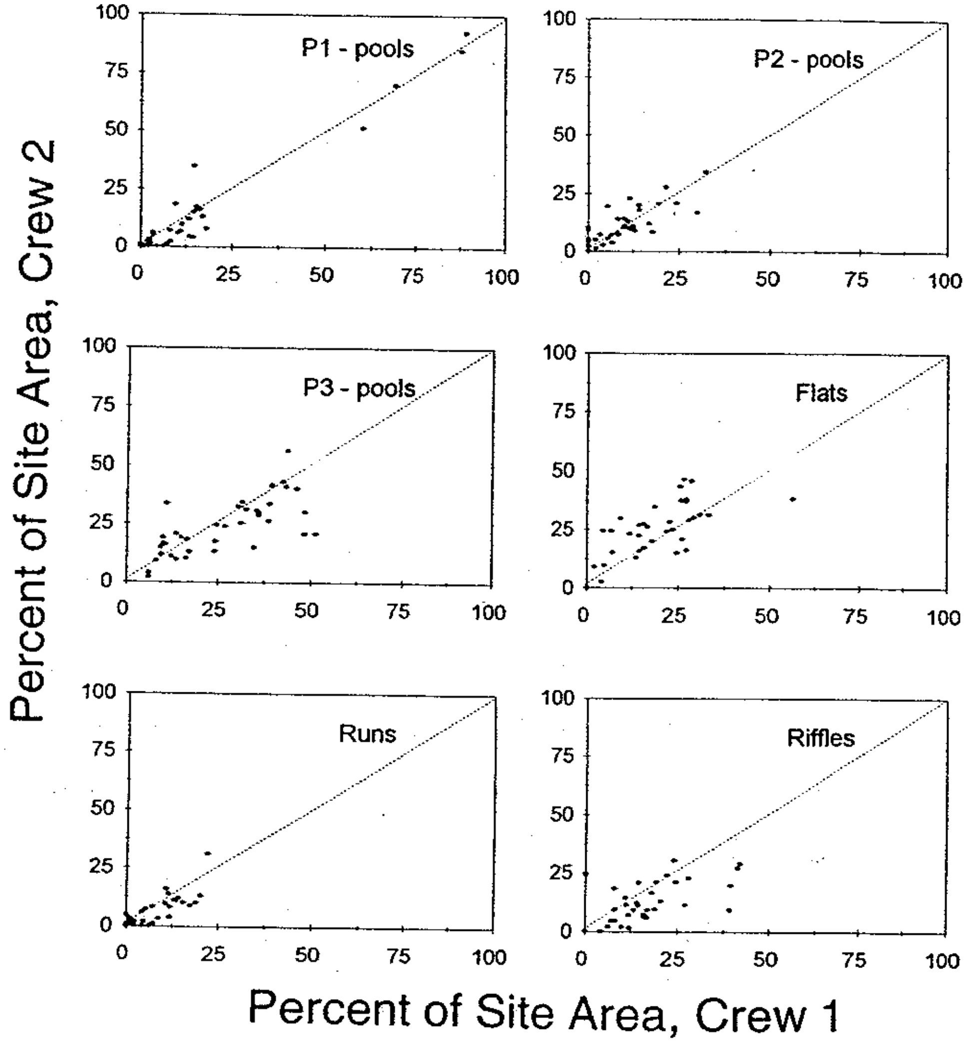


FIGURE 10.—Between-crew comparisons of the observed relative abundances of six habitat types classified by the measured point–transect method. Each point corresponds to the two visits to an individual site by each crew.

vise any users of a visual survey to ensure that the questions being asked of the data can be adequately addressed by the accuracy and precision of the data being collected.

We applied a transect-based system that is more repeatable for habitat type classification and that provides greater flexibility for post-survey interpretations. We have also shown that the measured point–transect method appears to be transferrable across a range of stream sizes, from first to fourth order. It took between 2 and 3 h/site to apply the MPT survey. This intensity of survey may not be cost-effective for large-scale watershed studies. Dolloff et al. (1993) propose a survey design that uses an objective survey to calibrate a visual survey. We are currently exploring this design and encourage other investigators to respond to the challenge of Hawkins et al. (1993) by developing and testing stream assessment methods that contribute to the common goal of an objective, repeatable, yet efficient (i.e., practical) methodology for stream habitat classification.

Acknowledgments

The field crew of A. Kauppinen, C. Byrne, K. Ryall, M. Kary, A. Sluiter embraced the design of this study and endured the monotony of sur-

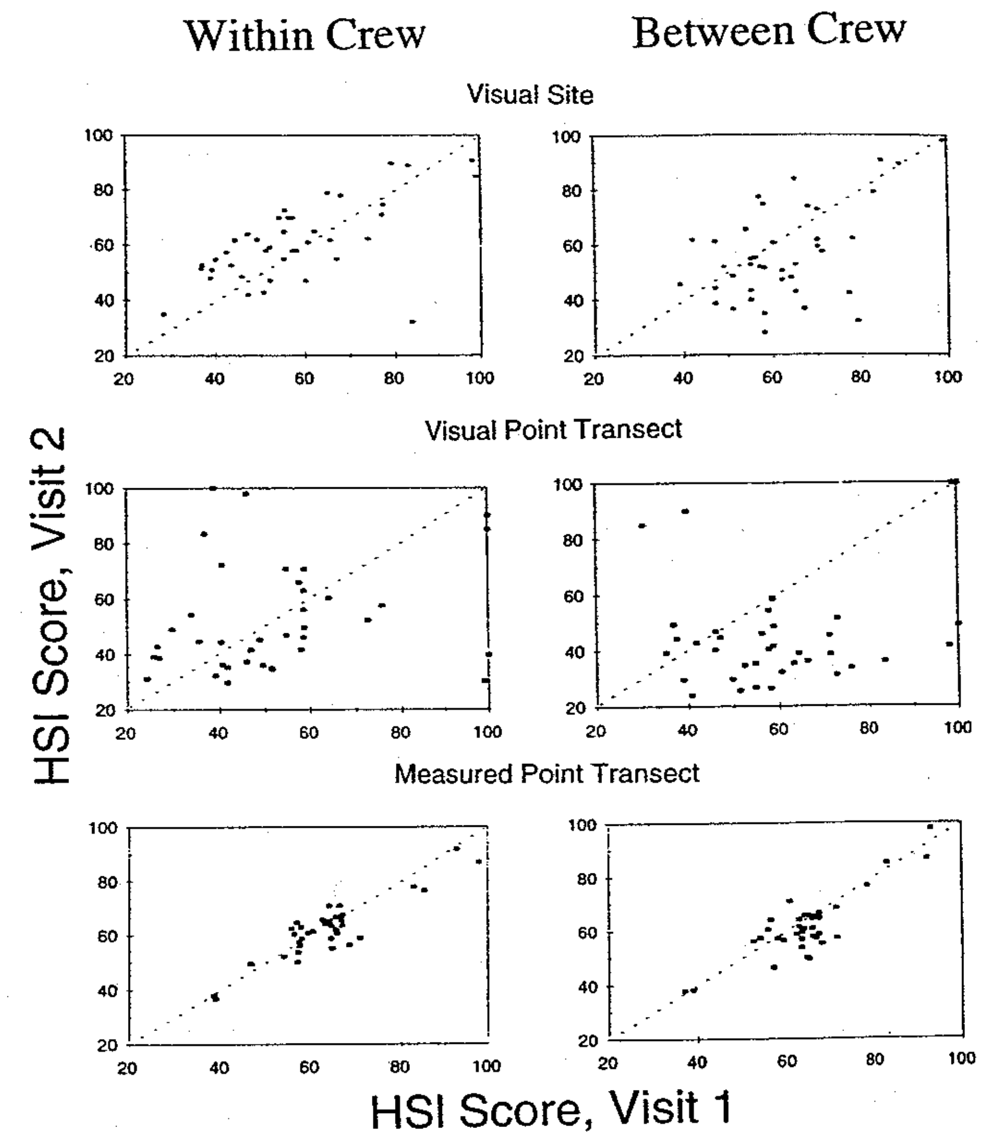


FIGURE 11.—Between-visit (left panels) and between-crew (right panels) comparisons of the age-0 rainbow trout habitat suitability index (HSI) scores for the visual site, visual point-transect, and the measured point-transect methods. The HSI score is a sum of the habitat types present at the site, weighted according to the presumed suitability of the habitat type for age-0 rainbow trout. It is expressed as a percentage, with 100% indicating an ideal site (with respect to morphology) for age-0 rainbow trout.

veying the same site over and over. Mike Stoneman provided invaluable insight into the development of the procedures and supervised the field crews. We particularly thank Ken Minns for his insightful (if off the cuff) question of our habitat suitability report which predated this work, "Are your results repeatable?" The manuscript was improved by helpful comments from our colleagues Nigel Lester and Leon Carl and from Tim Simonson and two unnamed reviewers. Finally, we thank the Greg Clarke Chapter of Trout Unlimited and the Ontario Environmental Youth Corp. for providing and administering the field crews for this study. This is Ontario Ministry of Natural Resources contribution 97-08.

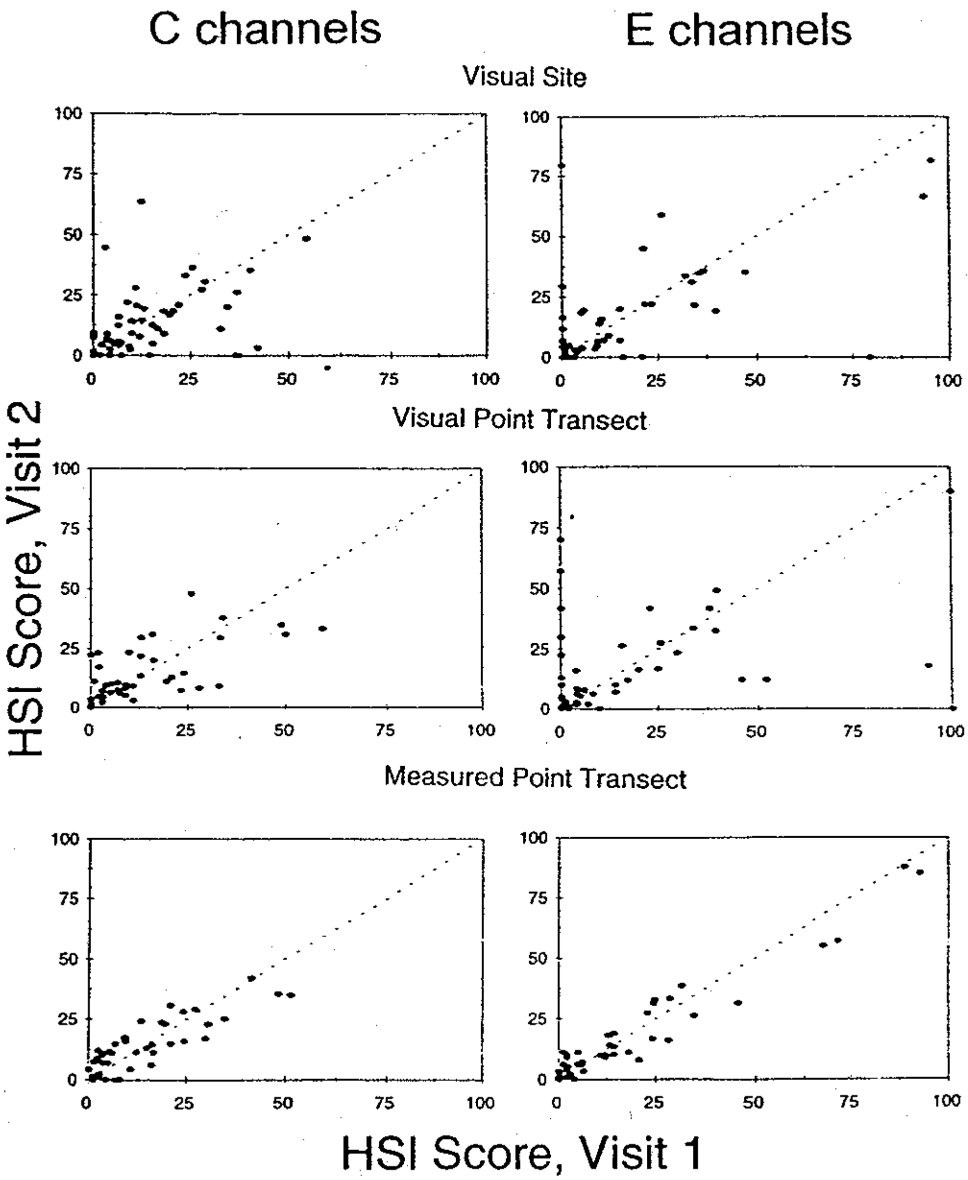


FIGURE 12.—Between-visit comparisons of the observed relative abundances of all habitat types for C-type channels versus E-type channels and comparing the visual site, visual point–transect and measured point–transect methods. Each point corresponds to the two visits to an individual site by one crew. The graphs include data for both crews.

TABLE 5.—Comparison of the agreement between the hierarchical key and Froude number classification methods against the visual technique, using one crew's data from both visits. Observation points classified as flats by the visual method were not included in the calculations. Numbers in parentheses are the percent of points not classified because of missing data.

Technique	Pools	Riffles	Runs	Overall
Hierarchical key	90.0% (12.6%)	79.5% (19.9%)	49.8% (6.4%)	83.9% (15.0%)
Froude number	83.6% (13.4%)	12.9% (32.0%)	56.7% (5.1%)	63.6% (14.4%)
Number of visual observations	2,249	322	328	2,899

References

Binns, N. A., and F. M. Eiserman 1979. Quantification of fluvial trout habitat in Wyoming. Transactions of the American Fisheries Society 108:215–228.

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

Bowlby, J. N., and J. C. Roff 1986. Trout biomass and habitat relationships in southern Ontario streams. Transactions of the American Fisheries Society 115: 503–514.

Dolloff, C. A., D. G. Hankin, and G. H. Reeves 1993. Basinwide estimation of habitat and fish populations in streams. U.S. Forest Service General Technical Report SE-83.

Efron, B., and R. J. Tibshirani 1993. An introduction to the bootstrap. Chapman and Hall, New York.

Hawkins, C. P., and ten coauthors 1993. A hierarchical approach to classifying stream habitat features. Fisheries 18(6):3–11.

Jones, M. L., and L. W. Stanfield 1993. Effects of exotic juvenile salmonines on growth and survival of juvenile Atlantic salmon (*Salmo salar*) in a Lake Ontario tributary. Canadian Special Publication of Fisheries and Aquatic Sciences 118:71–79.

Jowett, I. G. 1993. A method for objectively identifying pool, run, and riffle habitats from physical measurements. New Zealand Journal of Marine and Freshwater Research 27:41–248.

Newbury, R. W., and M. N. Gaboury 1993. Stream analysis and fish habitat design: a field manual. Copublished by Newbury Hydraulics Ltd., the Manitoba Habitat Heritage Corporation and Manitoba Natural Resources, Gibsons, British Columbia.

Ontario Ministry of Natural Resources 1989. Manual of instructions: aquatic habitat inventory surveys, 8th edition. Ontario Ministry of Natural Resource Fisheries Branch, Toronto.

Orth, D. J. 1983. Aquatic habitat measurements. 61–8 *in* L. A. Nielson and D. L. Johnson, editors. Fish eries techniques. American Fisheries Society, Be thesda, Maryland.

Platts, W. S., W. F. Megahan, and G. W. Minshall 198 Methods for evaluating stream, riparian and bioti conditions. U.S. Forest Service General Technic Report INT-38.

Roper, B. B., and D. L. Scarnecchia 1995. Observ variability in classifying habitat types in stream su veys. North American Journal of Fisheries Man agement 15:49–53.

Rosgen, D. 1996. Applied river morphology. Wildlan Hydrology, Pagosa Springs, Colorado.

Simonson, T. D., J. Lyons, and P. D. Kanehl 1994 Guidelines for evaluating fish habitat in Wisconsi streams. U.S. Forest Service General Technical Re port NC-164.

Simonson, T. D., J. Lyons, and P. D. Kanehl 1994 Quantifying fish habitat in streams: transect spac ing, sample size and a proposed framework. Nort American Journal of Fisheries Management 1 607–615.

Stoneman, C. L., M. L. Jones, and L. W. Stanfield 199 Habitat suitability assessment models for souther Ontario trout streams. Canadian Manuscript Repo of Fisheries and Aquatic Sciences 2345.

Wang L., T. D. Simonson, and J. Lyons 1996. Ac and precision of selected stream habitat esti North American Journal of Fisheries Manageme 16:340–347.

Wright, J. F., P. D. Hiley, S. F. Ham, and A. D. Berr 1981. Comparison of three mapping procedures de veloped for river macrophytes. Freshwater Biolog 11:369–379.

Received May 27, 19
Accepted February 6. 19

Section 5:

Models generated from OSAP data

Section 5 Module 1

Habitat Suitability indices for selected Ontario fish species[1]

Table of Contents

APPENDICES

[1] Authors: L. W. Stanfield, M. Stoneman, B. Kilgour and J. Parish

1.0 INTRODUCTION

Details of the approaches used, inherent assumptions, and results of the development for each of the 14 models are provided in this appendix. The models described here are incorporated into the summary procedures available through HABPROGS.

2.0 DEVELOPING SUITABILITY INDICES FOR FISH

In 1999 research staff at OMNR developed a number of suitability indices for common fish for which both density and habitat data was available. Data used for this initiative followed the earlier versions of the OSAP manual (Stanfield et al. 1999). The first step in developing the SI's was to randomly select 2/3 of the available sites for use in developing the suitability curves. This translated to 159 sites, although only 110 had complete data, while the remainder lacked thermal stability data.

For each of these habitat attributes and for each species/age group, the following procedures were used to fit the data to a Poisson curve.

General steps to develop suitability curves:

Habitat scores for each variable were divided into 9-15 categories depending on the variable of interest. The boundaries for each category were based on the range of scores observed for each variable.

For categories with a sample size of 4 or more, we calculated the 90th percentile value of the biomass for each category (i.e., 1 C thermal unit). We judged that the 90th percentile approximates the maximum potential biomass for sites with given habitat characteristics (Hubert et al., 1996). If a category had less than 4 observations, it was not included.

Adjacent cell averaging was used to calculate a smoothed average maximum biomass for each category. The data was smoothed to reduce the variation in 90th percentile values. This was also done to fill in missing data.

Finally, we standardized the distribution to a suitability score to 1, by dividing the smoothed averages for biomass for each category by the highest 90th percentile of biomass observed. This ensured that each suitability curve would range from 0 to 1.

For each of the habitat parameters, curves or plots through the standardized data were fitted using non-linear Poisson equations. Figure 1 provides an example of curve fitting of thermal stability and creek chub biomass.

Initial scanning of the curves produced some unusual shapes. To test whether these curves were true representations of the fishes' response to that variable or simply were a result of random chance, we chose to replicate the process and compare the two curves. A chi-square goodness of fit test were used to determine if the curves were different. In some instances, visual comparisons were used when high r^2 values precluded statistical significance. Where curves were considered the same, the first was used. Where they were different, data from model A and B were combined to generate a third model.

This process generated the SI curves and criteria for each of the 11 species. The parameters and criteria for each habitat variable are available in the database in the Table: Fish Species Habitat Associations. The Poisson equation is as follows:

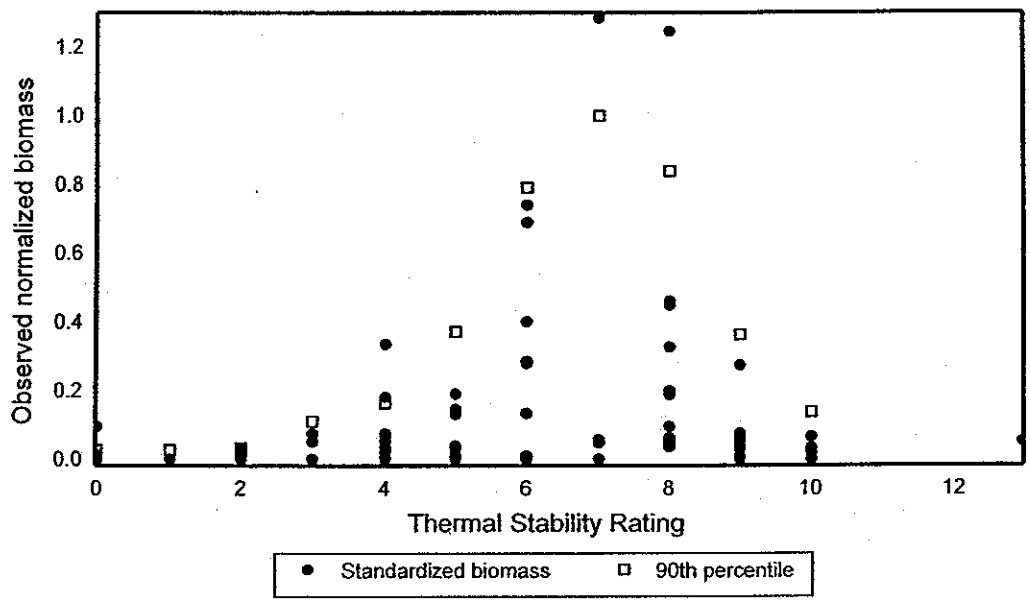
SI = (((b - x)/(b - a))^c) * e^((c/d) * (1 - ((b - x)/(b - a))^d))

where:
a = value of x where f(x) = 1.0
b = value of x where f(x) = 0, x < b
c = shape parameter for the part of the curve to the right of x = a
d = shape parameter for the part of the curve to the left of x = a
e = base of the natural logarithm
x = the habitat variable

1.1 Suitability Indices for Thermal Stability

Thermal stability ratings are calculated following the procedures described in Section 2 in order to obtain a deviation in degrees C from a reference cold water stream. Data were summarized to the nearest degree for this curve.

Creek chub: thermal stability raw data and 90th percentiles



Creek chub: thermal stability curve through 90th percentiles

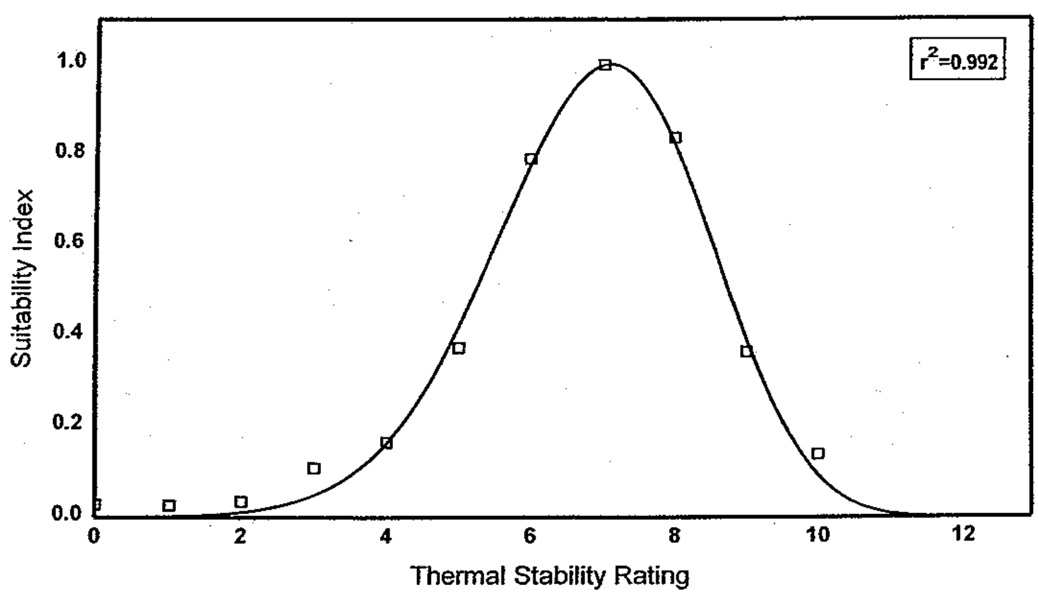


Figure VI.1: Top: Scatter plot of normalized biomass for creek chub and thermal stability (90^{th} percentiles shown as solid circles). Bottom: Fitted Poisson regression line through the 90^{th} percentiles of normalized fish biomass.

1.2 Suitability Indices for Hilsenhoff Index of Biotic Integrity

Some preliminary work had established good relationships between the major-groups of taxa and biomass of key stream fish species see (Kilgour and Barton, 1999). Development of SI curves for this attribute required no additional changes to the data, once the BI was developed. Results can only be interpreted within the range of scores available from the dataset at that time (i.e., 3.75-7.25).

1.3 Suitability Indices for Substrate Quality

Substrate categories were discrete, so regressions were not used. For this variable, biomass in all 9 substrate categories was standardized by dividing by the largest observed biomass in any substrate class. The resulting standardized biomass values were considered equivalent to HSI values. Adjacent cell averaging was not done unless there were no sites in a category as for Q9. The category value then became the average of the adjacent substrate classes (i.e., average of Q5, Q6, and Q8). Standardized values for brook trout biomass in various substrate categories are shown in Figure IV.1.

1.4 Suitability Indices for Channel Stability Rating

The composite scores for channel stability (as defined in Section 4 Module 1) were used to generate suitability ratings. These data produce ratings between .3 and 1 with .3 representing highly unstable (dynamic) channels.

1.5 Suitability Indices for Channel Structure

We have developed a way to preserve spatial links between the point observations and the associated habitat and cover characteristics when the OHSIM is run, but have yet to find a way to present this information.

Each individual species was classified by the suitability of each morphological feature and cover categories. Each habitat and cover category was classified by suitability as a "home" location for a fish species. Points were classified as either "0" as unsuitable; "+1" for a preferred habitat; "+0.5" for marginally suitable habitat or "-1" for habitats that are generally avoided.

In order to associate the habitat attributes to individual fish species, we created a table which included each of the habitat categories (16 morphological features, 3 cover categories, and 5 cover types) and all the stream fishes of Ontario. Each category was rated by the suitability of that habitat attribute for each species following the criteria described above. Criteria for assigning scores to cells used adjacent cell averaging wherever possible. For example, if we determined that shallow pools without cover were preferred by a species, then all shallow flats and medium depth pools would be scored with a .5. This meant that any habitat type which was only one category away from the preferred would be rated at 0.5. Finally, the avoided category occurs in cover types where, for example, a species is reported to avoid vegetation. When present, this overrides the possible suitability of the habitat type and sets the point score to 0.

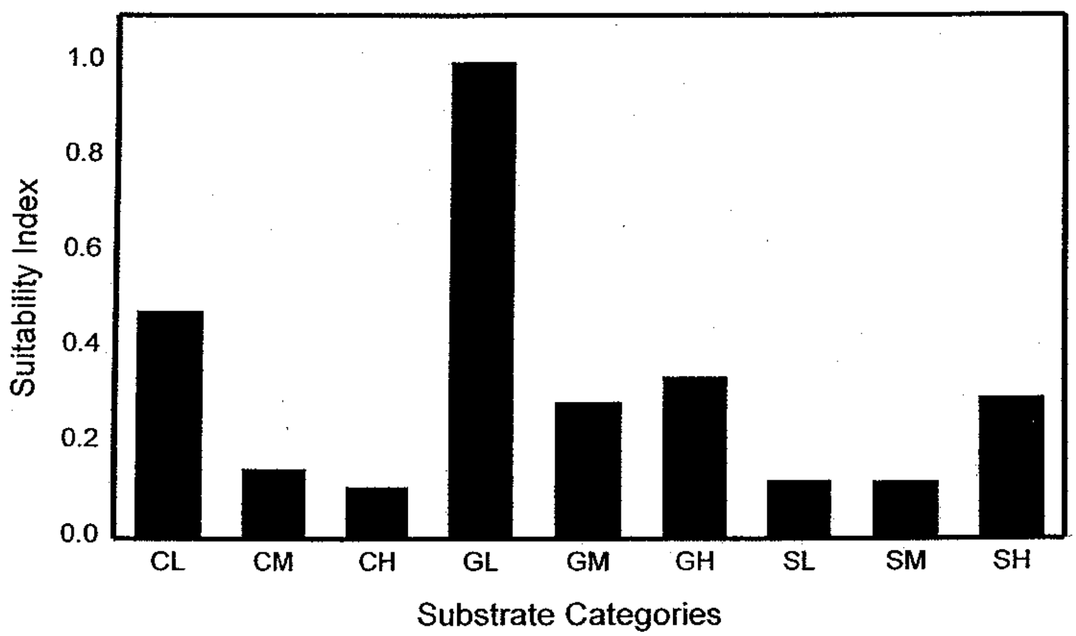


Figure IV.1: Histogram of normalized biomass for brown trout > 70 mm. Each of 193 development sites were scored for substrate quality and grouped into 1 of 9 categories. For each category with sample size of more than 4, the 90th percentile of the mean observed biomass estimate was standardized and plotted.

The classifications of each habitat attribute were based on extracts from the primary literature or technical documents (e.g. Aadlands et al., 1991). For species habitat relationships not reported in the primary literature, we consulted publications containing general accounts for species. Examples of such sources include *Freshwater Fishes of Canada* (Scott and Crossman, 1973) and *Fishes of Ohio* (Trautman, 1981). If habitat information for a species was not found in either of the two types of sources listed above, we used the associations for a species from a similar guild which would have similar habitat requirements. Finally, Drs. Ed Crossman and Erling Holm of the Royal Ontario Museum were consulted to verify the species habitat associations.

In most instances, interpretations were required to fit the described habitat requirements from the literature to our definitions. For example, we would fit the description "prefers shallow margins of streams" to shallow pools and we interpreted "often in association with submergent vegetation" as shallow pools with macrophyte vegetation. As more research describing species habitat relationships becomes available, the information used to develop this habitat suitability information model will be improved (Next Steps).

A query first categorizes the point observation into a habitat type, then evaluates the suitability of both the habitat and cover type for a particular species. If cover is present but not suitable, the score for the observation point reverts to a score reflecting no suitable cover present. This process occurs for all of the point observations at a site and is converted to a percent score for the site. This percent score for the site is then normalized to the total amount of usable area.

Output from this query provided a channel morphology rating or score for each site for particular species. The ratings were then categorized (10 groupings, based on the range observed for each species) to facilitate the creation of the suitability index. These results were plotted against the biomass for the target species in order to permit the determination of the 90 th percentile of the biomass for each category. Once determined, this data was smoothed using adjacent cell averaging. Finally, the data was fit to a regression line. The rationale to fitting this variable to a regression line was that it was assumed that there would be a linear relationship between available habitat and the biomass of each species. The regression line was: $SI = ax + b$

where:

a = the slope co-efficient

b = the y - intercept

x = the habitat variable

2.0 SUITABILITY INDICES (SI'S) FOR ONTARIO RIVERINE SPECIES

The following figures are the suitability indices for the 14 species/age groups of fish for which models have been developed in Ontario. These figures can be applied individually to evaluate the suitability of a stream segment for a target fish group for an individual habitat attribute. For example, if only a subset of data was collected (i.e., thermal stability), the manager could evaluate the system for that attribute alone to determine suitability for a species. Collectively these SI's, along with stream width, form the input to the OHSI model.

List of Figures Page

1.0 INTERPRETING THE Y AXIS

We define habitat suitability in terms of biomass: the higher the biomass of a given species at a site, the greater the suitability of that site for the species of interest. By dividing the observed biomass at each site by the maximum biomass observed at any site, we have

Figure IV.4 Suitability Indices for Brook trout (< 70 mm)

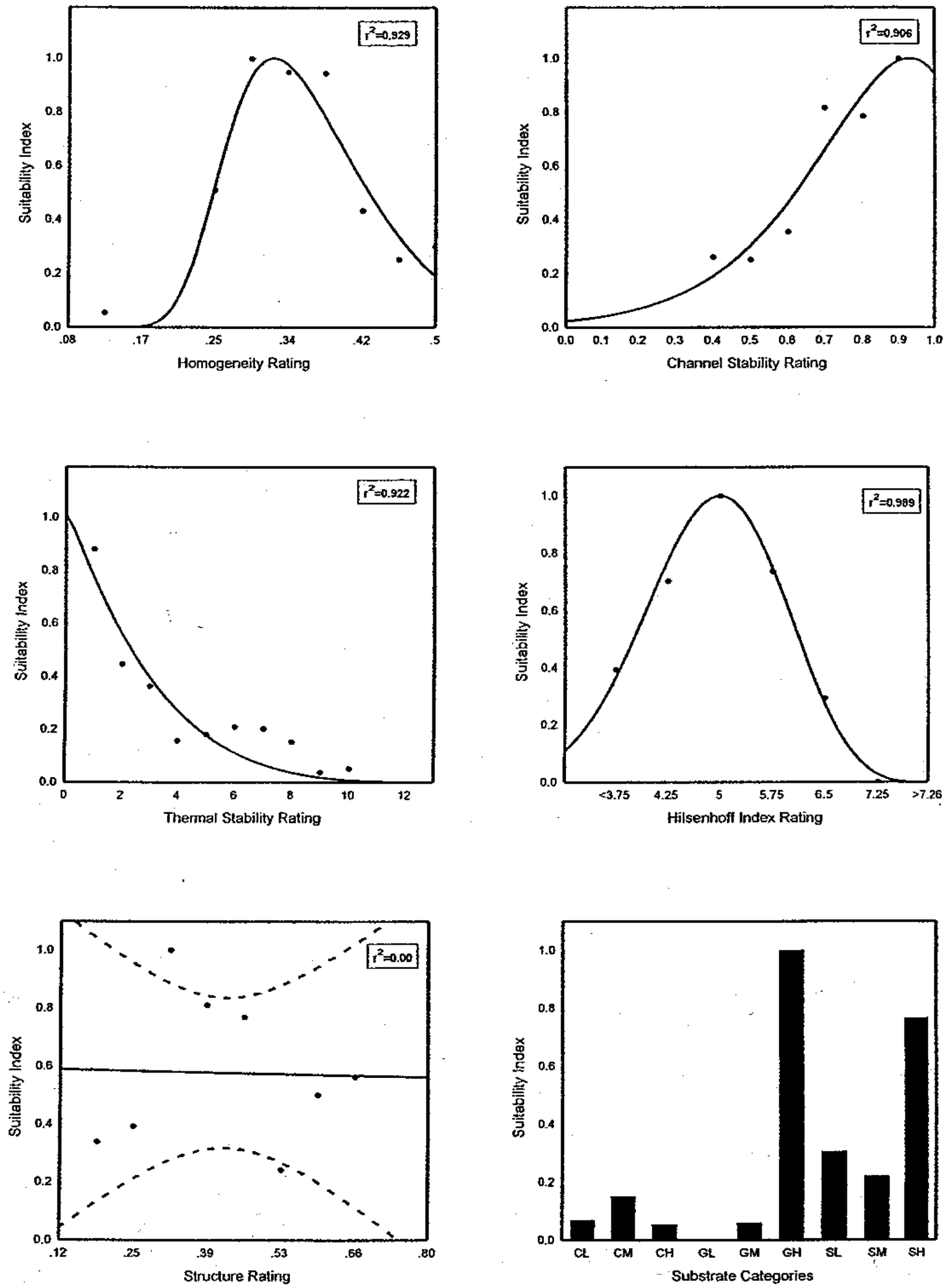


Figure IV.5 Suitability Indices for Brook trout (> 70 mm)

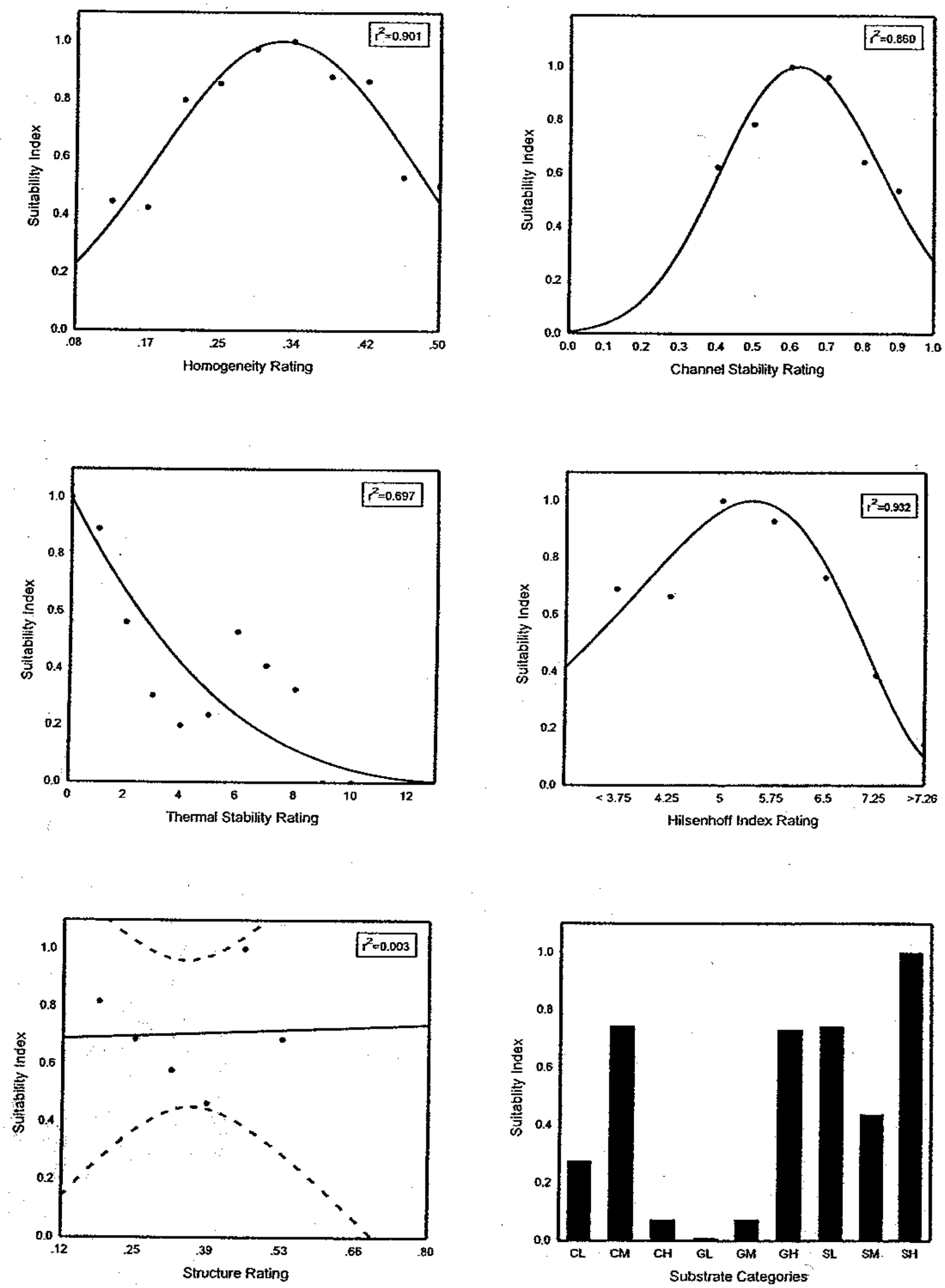


Figure IV.6 Suitability Indices for Blacknose dace

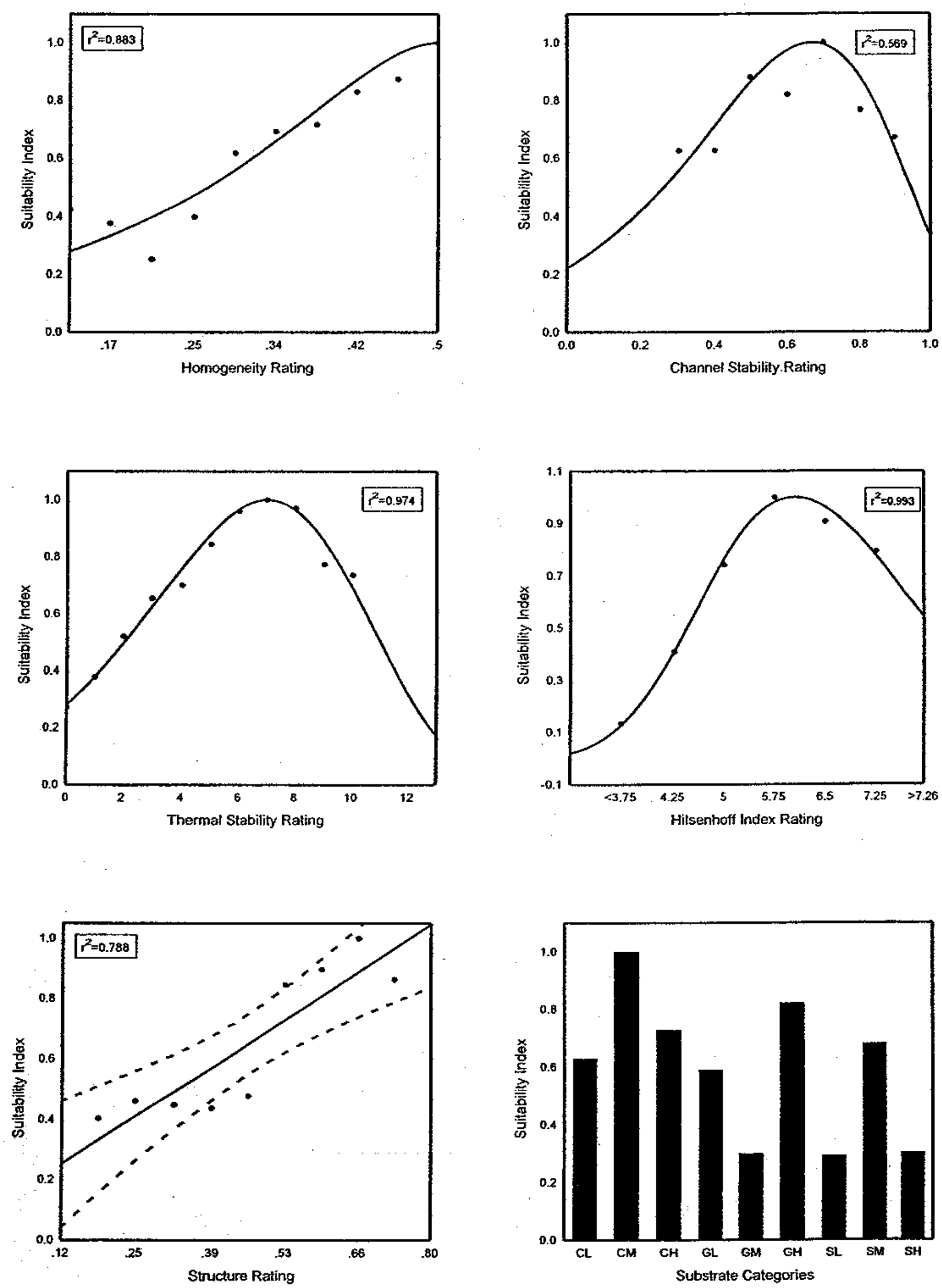


Figure IV.7 Suitability Indices for Brown trout (< 70 mm)

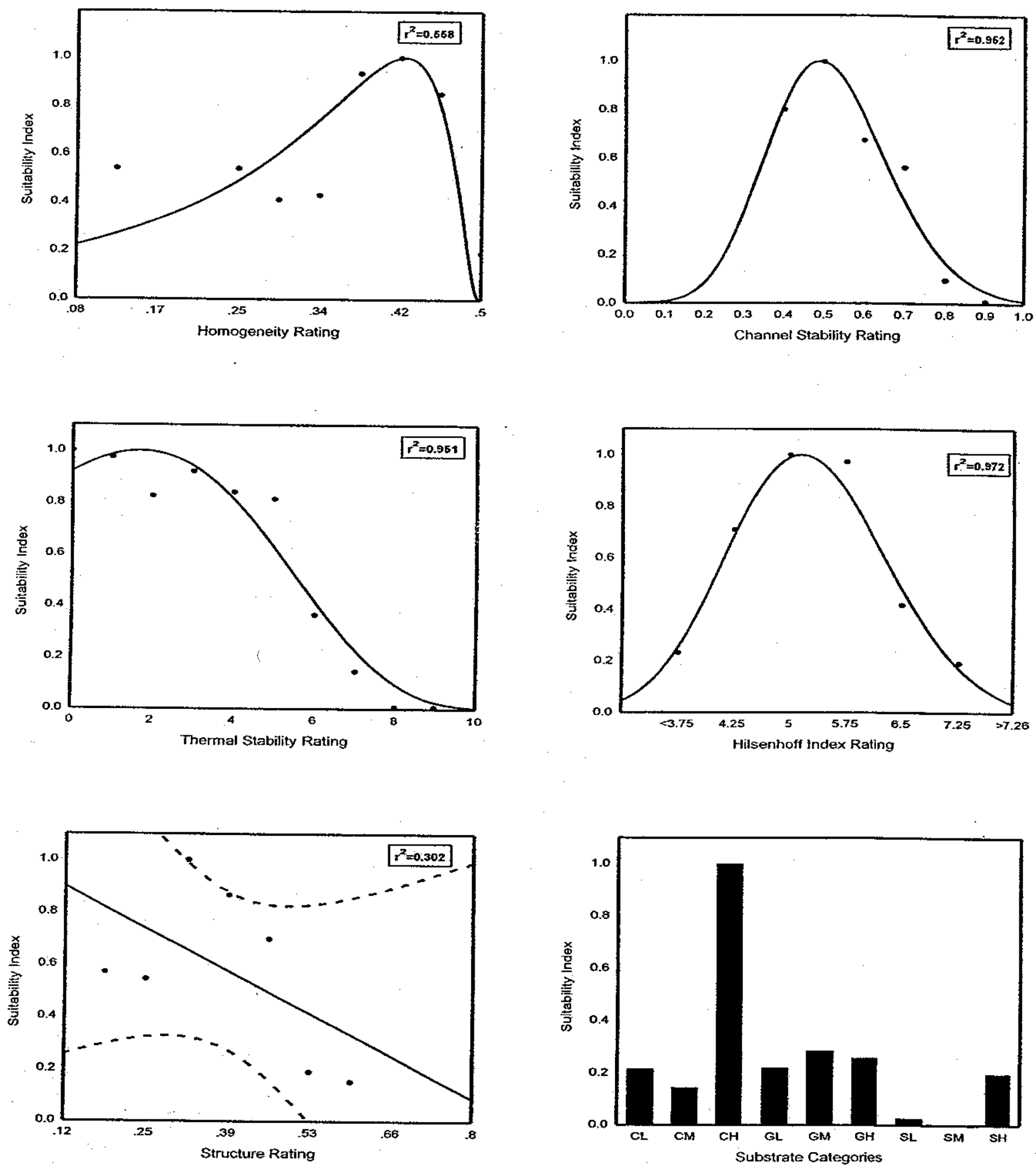


Figure IV.8 Suitability Indices for Brown trout (> 70 mm)

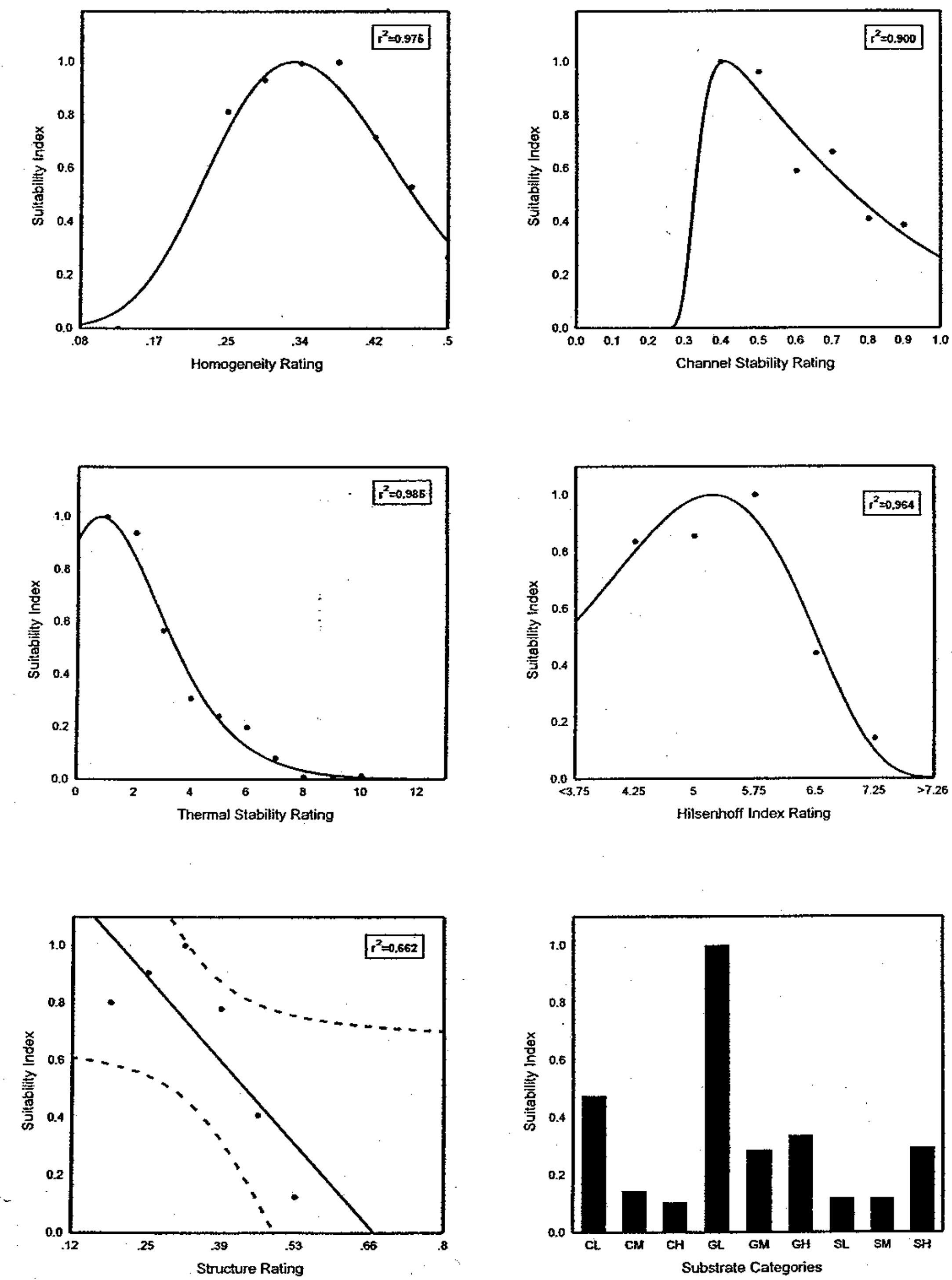


Figure IV.9 Suitability Indices for Creek chub

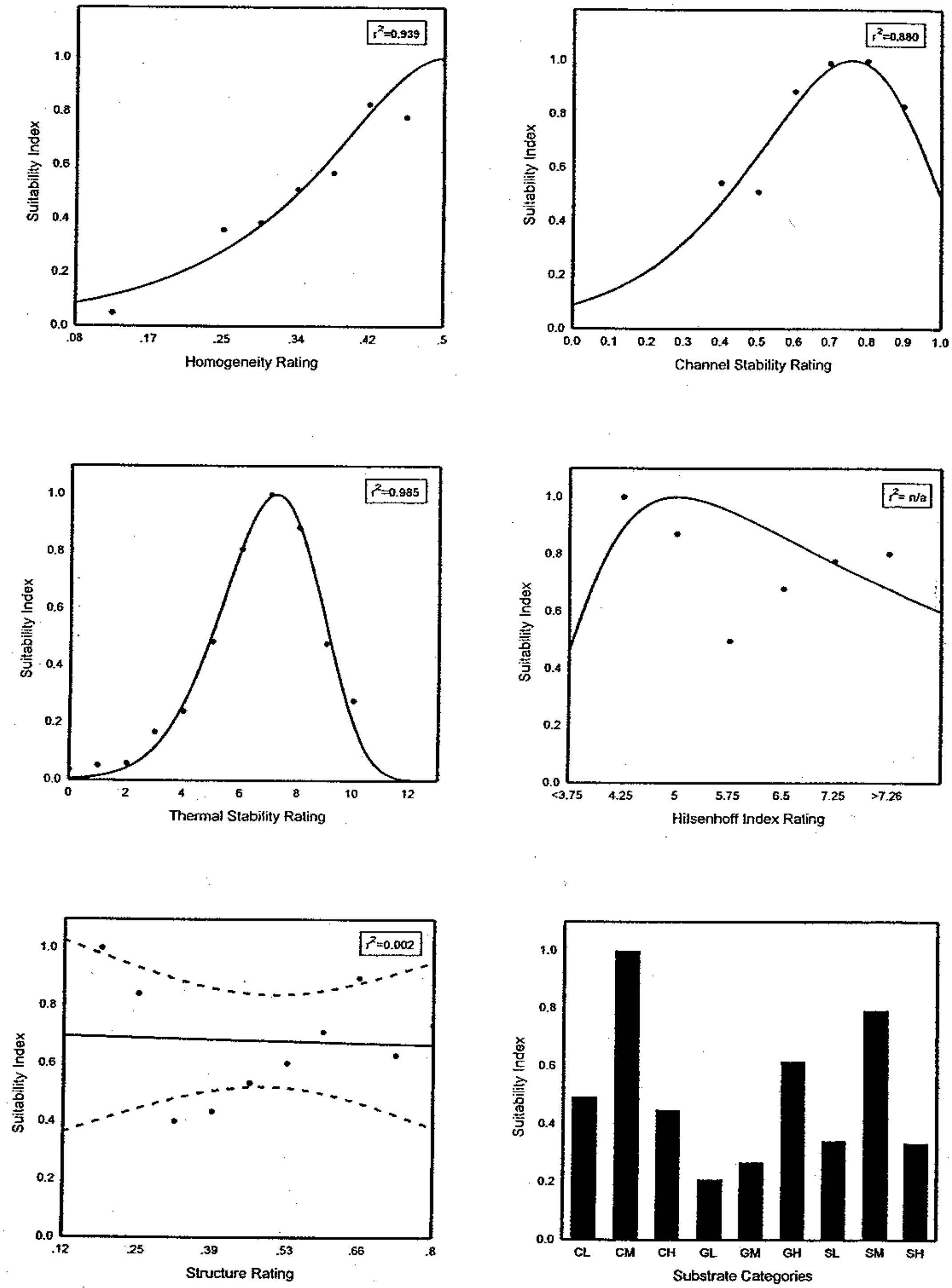
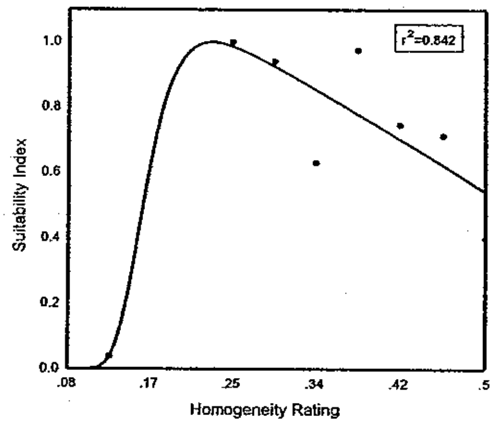
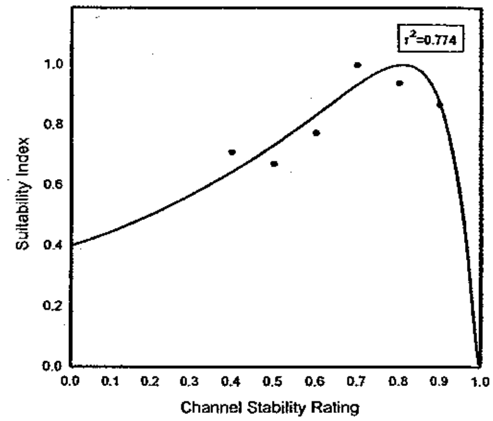
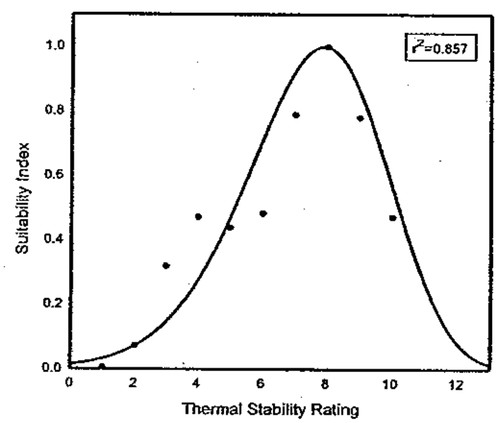
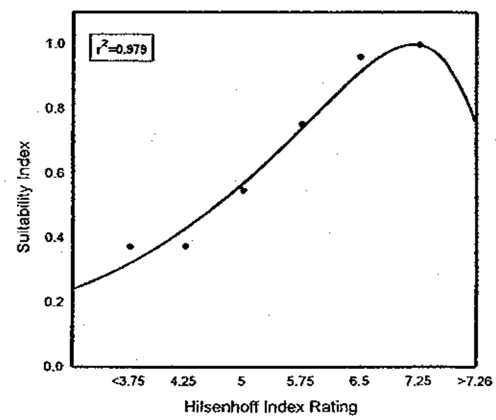


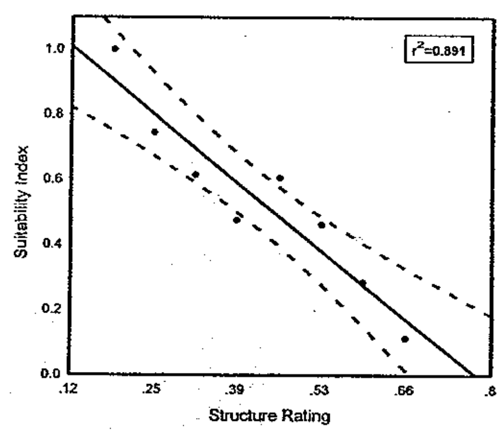
Figure IV.10 Suitability Indices for Common shiner


r²=0.842
Suitability Index
Homogeneity Rating
1.0
0.8
0.6
0.4
0.2
0.0
.08
.17
.25
.34
.42
.5


r²=0.774
Suitability Index
Channel Stability Rating
1.0
0.8
0.6
0.4
0.2
0.0
0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0


r²=0.857
Suitability Index
Thermal Stability Rating
1.0
0.8
0.6
0.4
0.2
0.0
0 2 4 6 8 10 12


r²=0.979
Suitability Index
Hilsenhoff Index Rating
1.0
0.8
0.6
0.4
0.2
0.0
<3.75 4.25 5 5.75 6.5 7.25 >7.26


r²=0.891
Suitability Index
Structure Rating
1.0
0.8
0.6
0.4
0.2
0.0
.12 .25 .39 .53 .66 .8

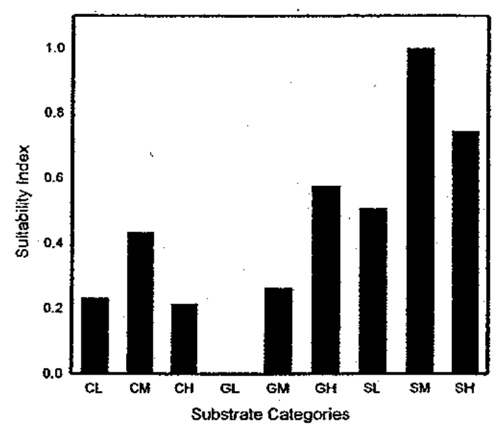

Suitability Index
Substrate Categories
1.0
0.8
0.6
0.4
0.2
0.0
CL CM CH GL GM GH SL SM SH

Figure IV.11 Suitability Indices for Johnny darter

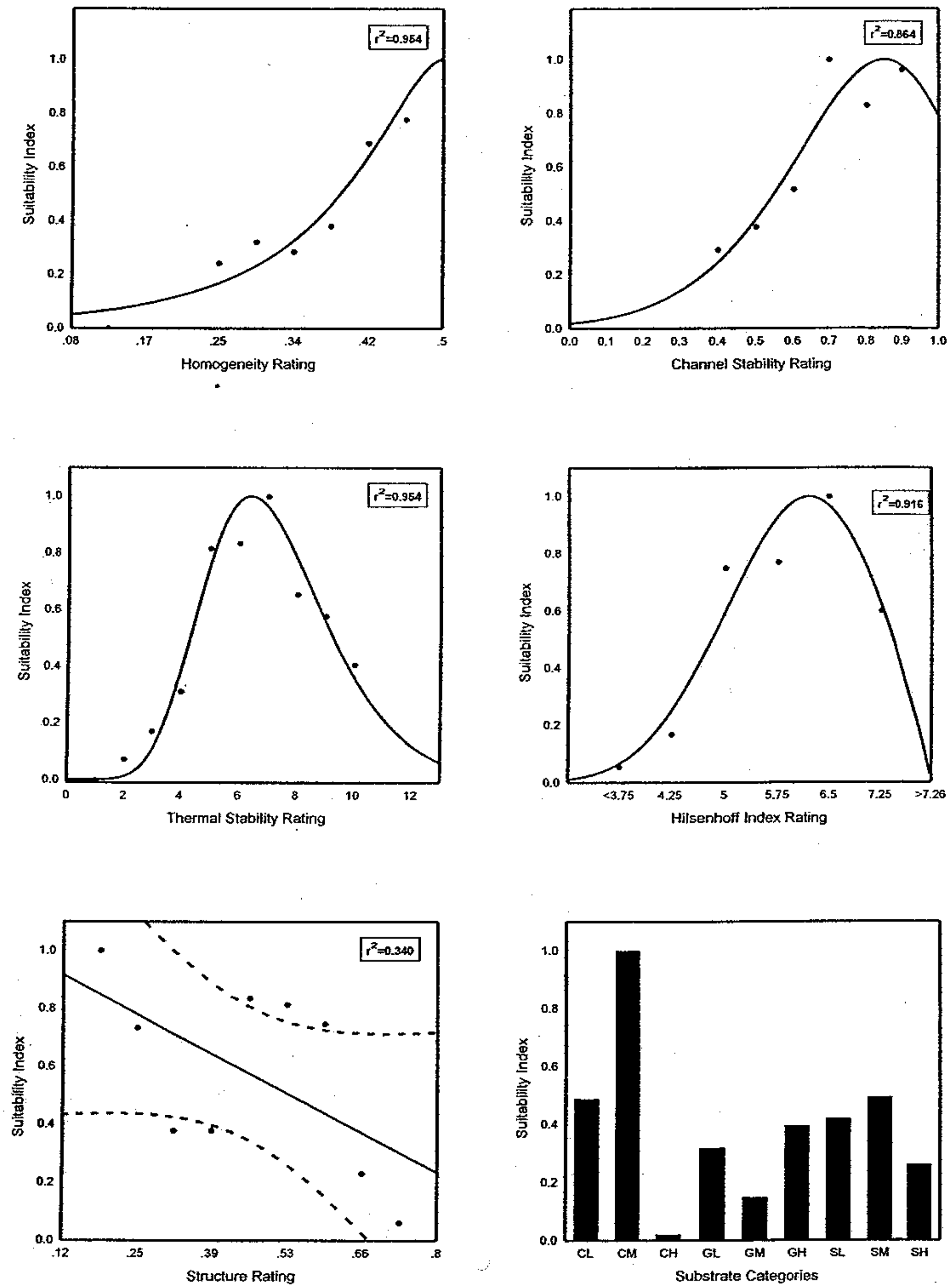


Figure IV.12 Suitability Indices for Longnose dace

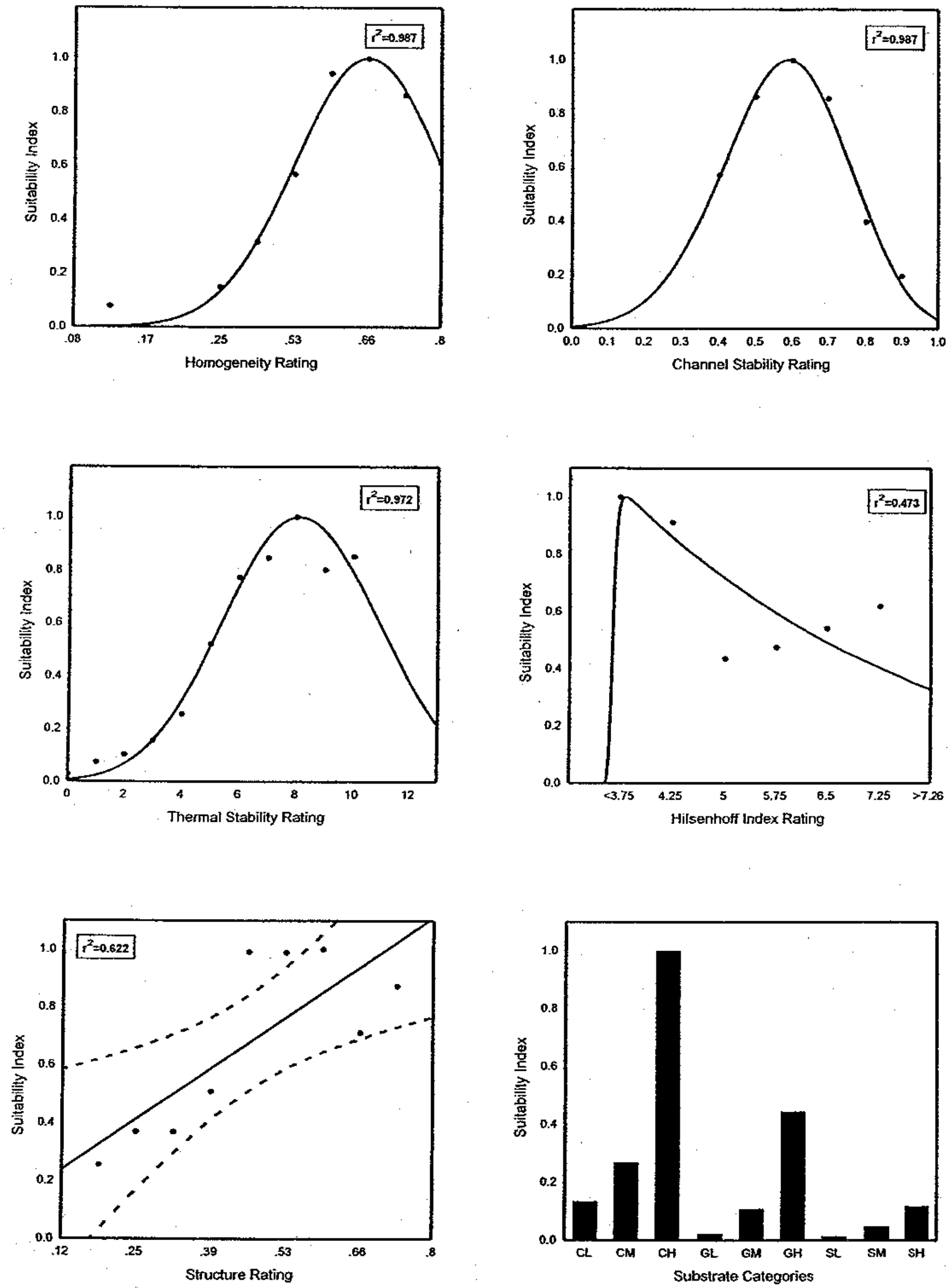


Figure IV.13 Suitability Indices for Mottled sculpin

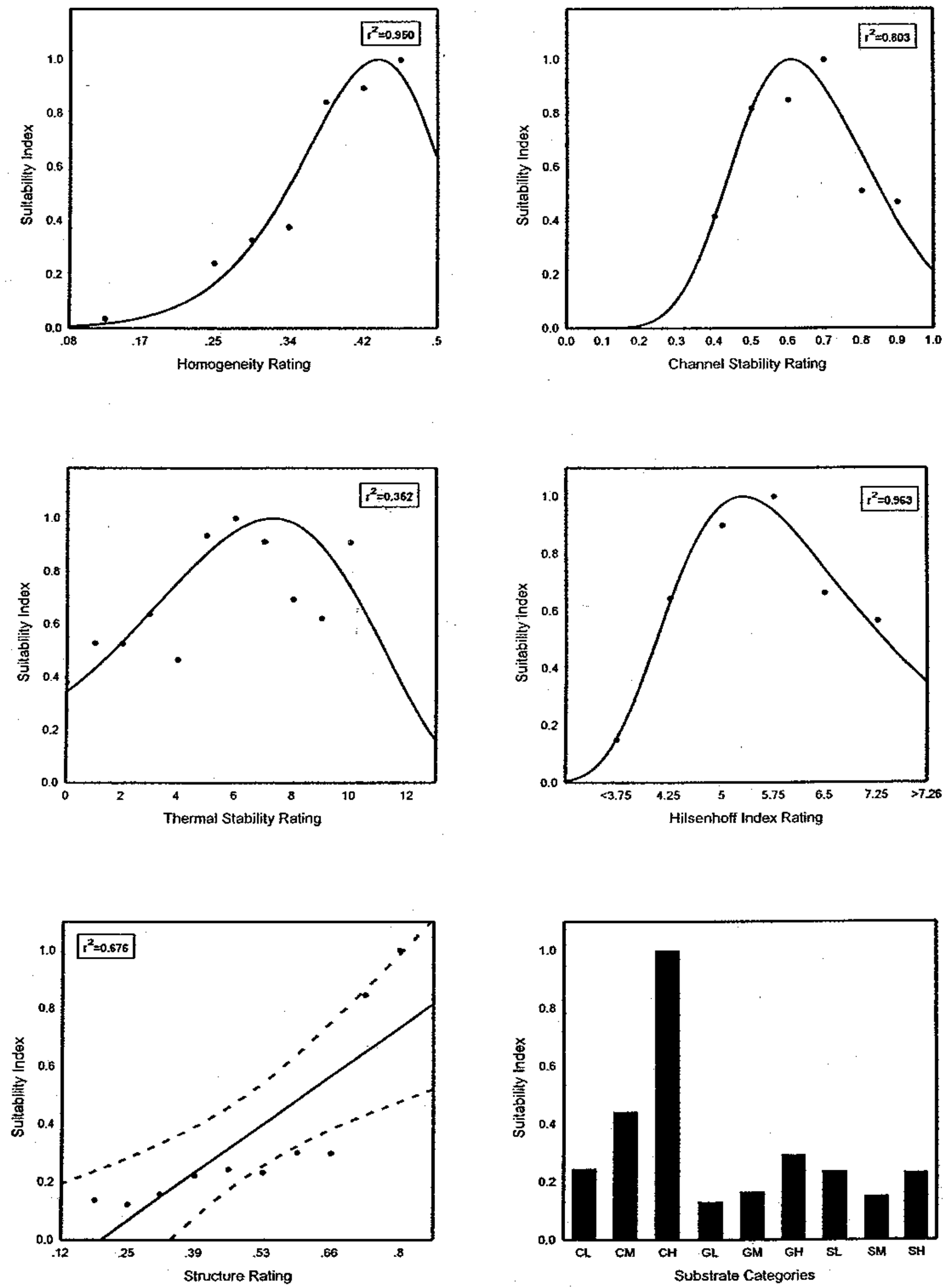


Figure IV.14 Suitability Indices for Rainbow darter

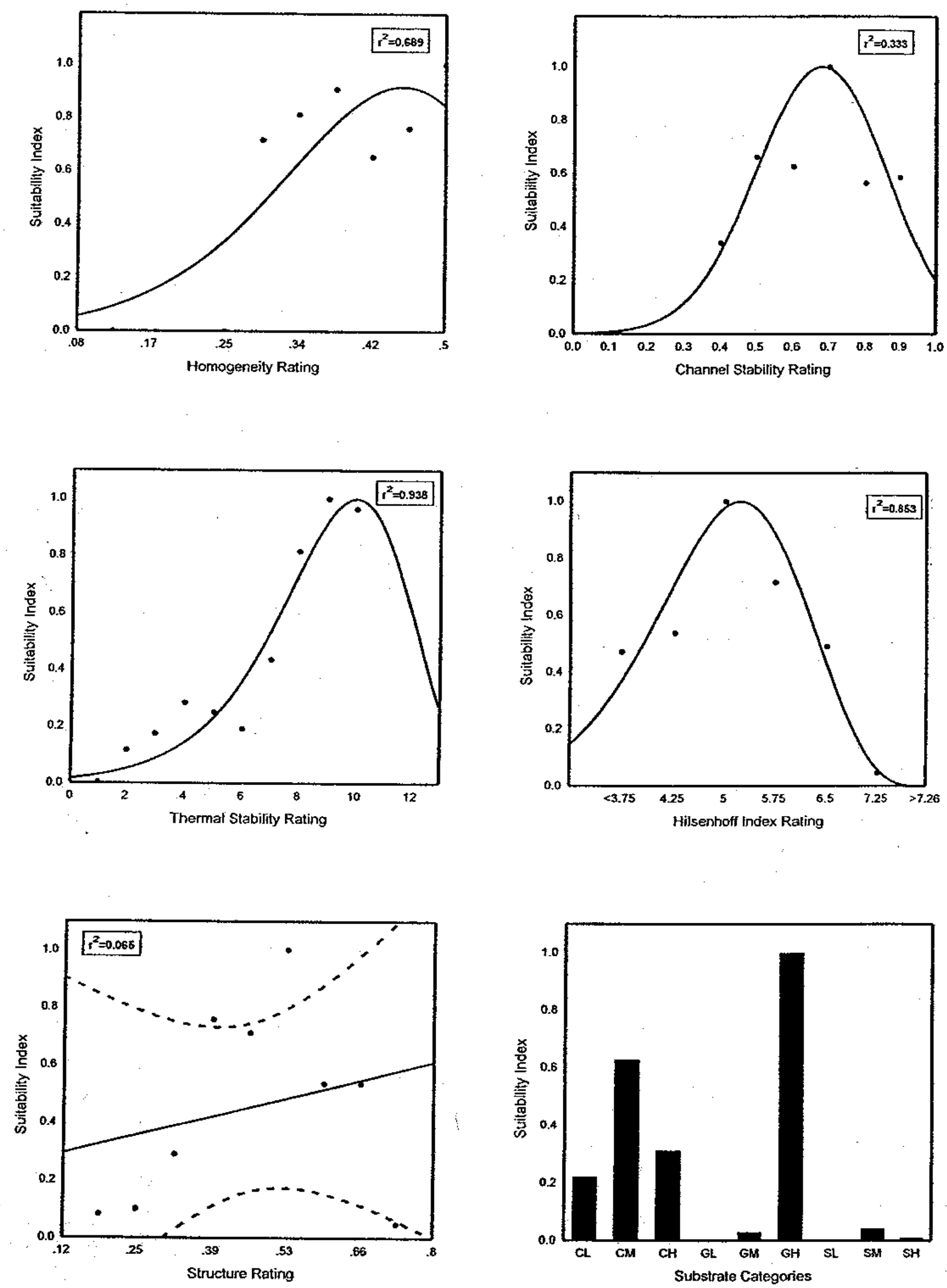


Figure IV.15 Suitability Indices for Rainbow trout (< 70 mm)

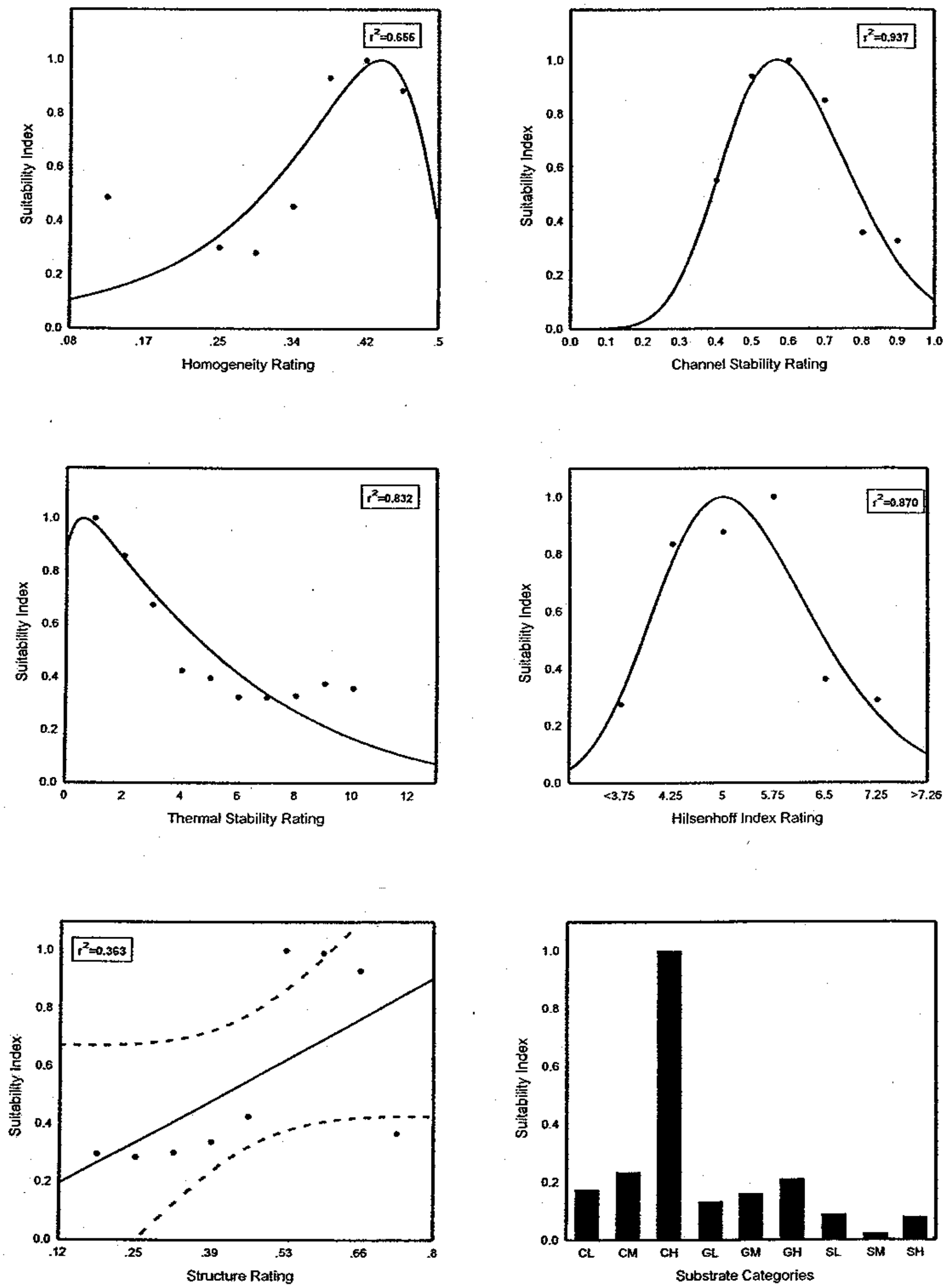


Figure IV.16 Suitability Indices for Rainbow trout (> 70 mm)

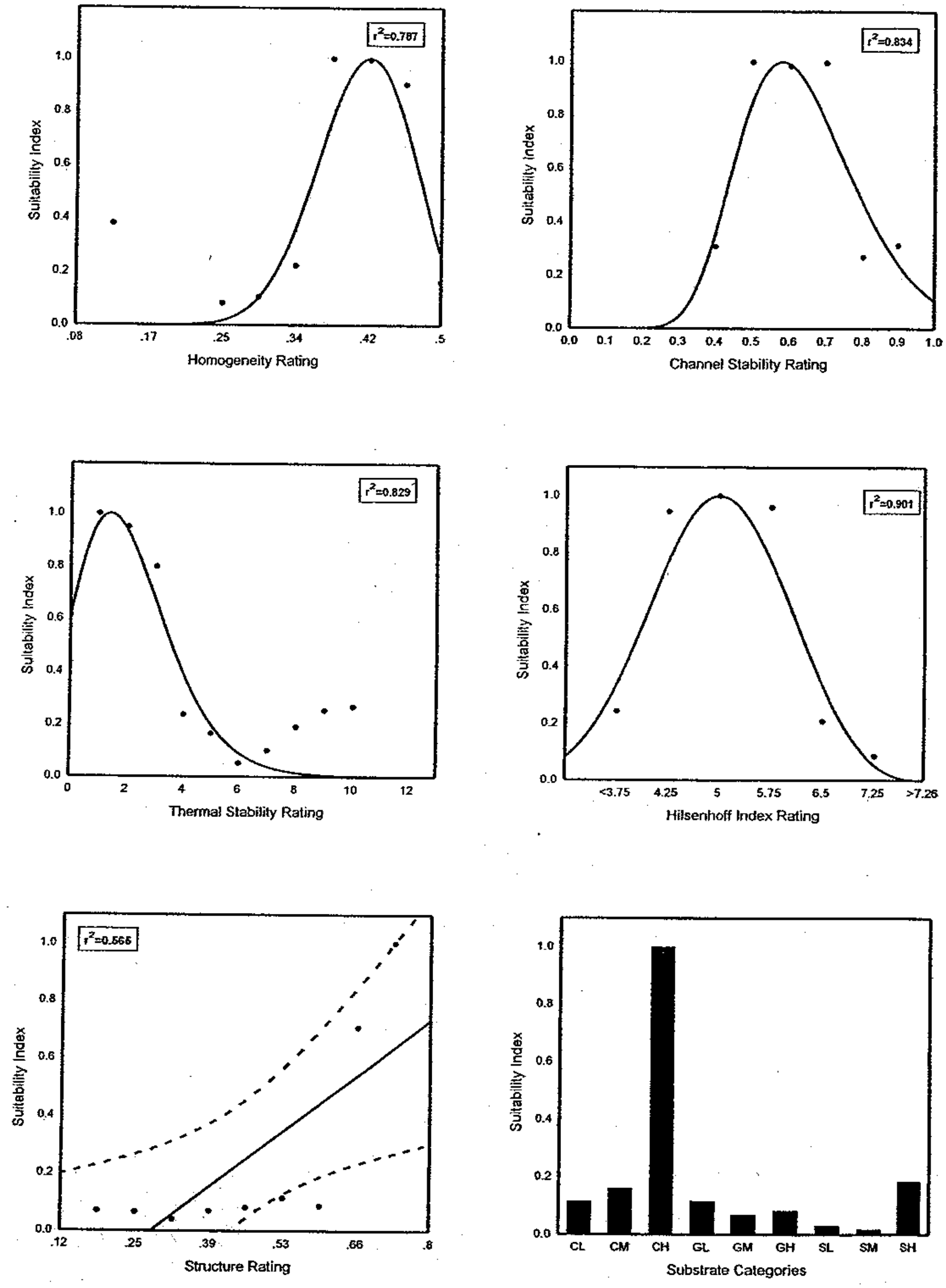
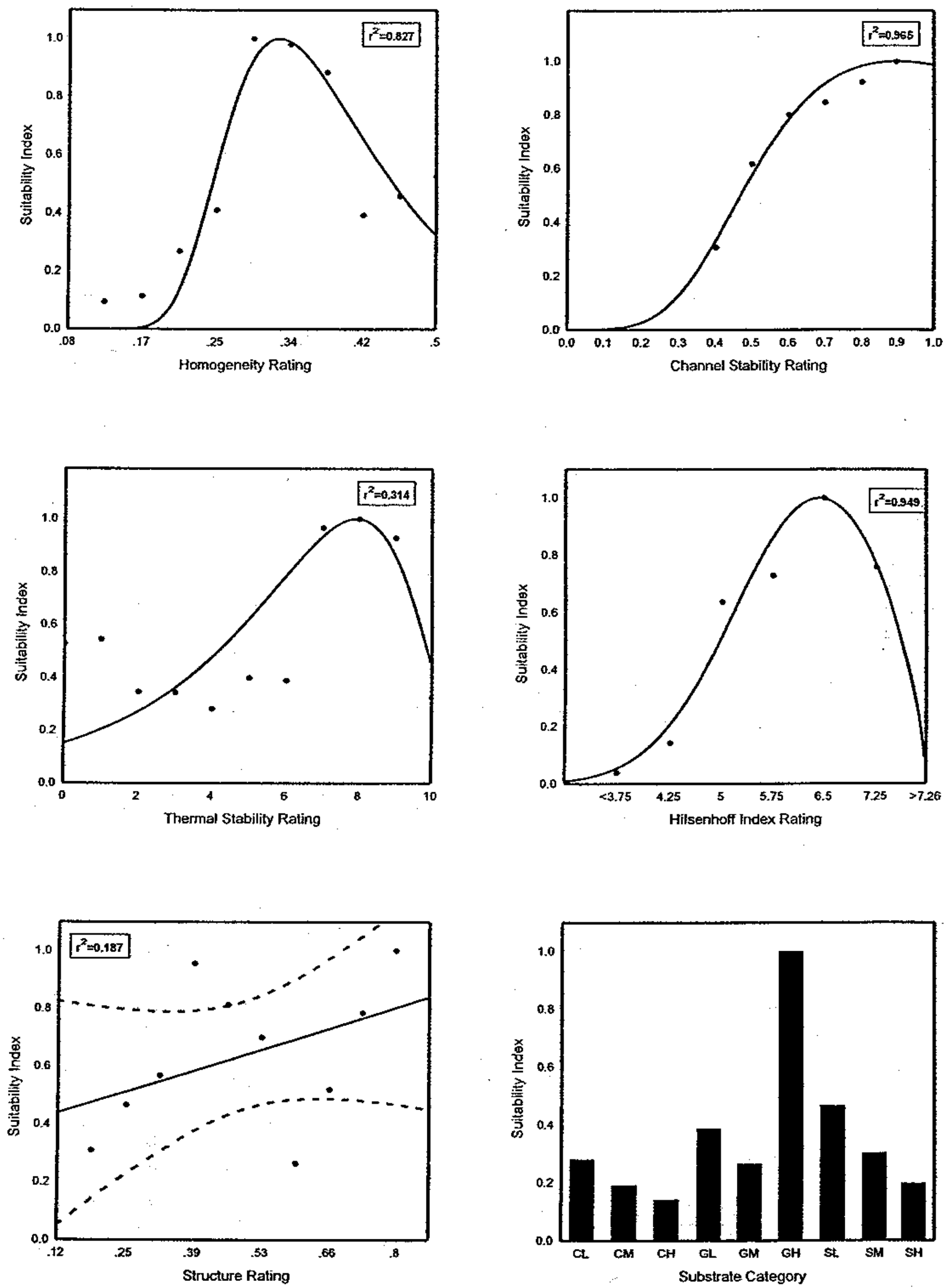


Figure IV.17 Suitability Indices for White sucker



Appendix I: Criteria for each habitat type used to create Channel Structure score.

	Rainbow Trout < 70 mm			
	Velocity Category			
Depth	Pool	Glide	Riffle	Chute
≤100	1.5	0.5	1.5	0
> 100 ≤ 600	1.5	1.5	1.5	0.5
> 600 ≤ 1000	2	1.5	2	0.5
> 1000	3	1.5	2	0.5

	Rainbow Trout >70 mm			
	Velocity Category			
Depth	Pool	Glide	Riffle	Chute
≤100	0.5	0	1	0.5
> 100 ≤ 600	1	1	1	1
> 600 ≤ 1000	1	1	1	0.5
> 1000	3	1	1	0.5

	Brown Trout < 70 mm			
	Velocity Category			
Depth	Pool	Glide	Riffle	Ch
≤100	0.5	0.5	1	
> 100 ≤ 600	1.5	0.5	0.5	
> 600 ≤ 1000	1.5	1	0.5	
> 1000	2	1	1	

	Brown Trout > 70 mm			
	Velocity Category			
Depth	Pool	Glide	Riffle	Chute
≤100	0	0	0	0
> 100 ≤ 600	1.5	0.5	0.5	0
> 600 ≤ 1000	1.5	2	0.5	0
> 1000	2	2	0.5	0

	Brook Trout < 70 mm			
	Velocity Category			
Depth	Pool	Glide	Riffle	Chute
≤100	0.5	0.5	0.5	0
> 100 ≤ 600	1.5	2	1	0
> 600 ≤ 1000	2	1.5	1	0
> 1000	2	1.5	1.5	0

	Brook Trout > 70 mm			
	Velocity Category			
Depth	Pool	Glide	Riffle	Ch
≤100	0	0	0	
> 100 ≤ 600	1.5	0.5	0.5	
> 600 ≤ 1000	1.5	0.5	0.5	
> 1000	2	2	0.5	

	Common White Sucker			
	Velocity Category			
Depth	Pool	Glide	Riffle	Chute
≤100	0	0	0	0
> 100 ≤ 600	1	1.5	0	0
> 600 ≤ 1000	1.5	1.5	2	0
> 1000	3	1.5	2	0

	Common Shiner			
	Velocity Category			
Depth	Pool	Glide	Riffle	Chute
≤100	0.5	0.5	0	0
> 100 ≤ 600	1	1	0.5	0
> 600 ≤ 1000	2	1.5	0.5	0
> 1000	3	1	0	0

	Blacknose Dace			
	Velocity Category			
Depth	Pool	Glide	Riffle	Ch
≤100	1	2	0.5	
> 100 ≤ 600	2	2	1	
> 600 ≤ 1000	2	2	1	
> 1000	1.5	1.5	0.5	

	Longnose Dace			
	Velocity Category			
Depth	Pool	Glide	Riffle	Chute
≤100	0	1.5	1.5	0.5
> 100 ≤ 600	0	2	1.5	0.5
> 600 ≤ 1000	0	2	2	0
> 1000	0	1.5	1.5	0

	Rainbow Darter			
	Velocity Category			
Depth	Pool	Glide	Riffle	Chute
≤100	0	1.5	2	0.5
> 100 ≤ 600	0	2.5	1	0.5
> 600 ≤ 1000	0	2.5	1	0.5
> 1000	0	0.5	0.5	0

	Johnny Darter			
	Velocity Category			
Depth	Pool	Glide	Riffle	Ch
≤100	0	2	2	
> 100 ≤ 600	1	2.5	2	
> 600 ≤ 1000	1	1.5	2.5	
> 1000	1	1	1	

	Mottled Sculpin			
	Velocity Category			
Depth	Pool	Glide	Riffle	Chute
≤100	0.5	0.5	2	0.5
> 100 ≤ 600	0.5	1.5	2.5	1.5
> 600 ≤ 1000	0.5	2	2.5	0.5
> 1000	1	1	2	0.5

HabProgs database table structure

Section 6: Module 1[1]

Table of Contents

[1] Authors: M. Stoneman and L. W. Stanfield

Section 5 Module 1

Habitat Suitability indices for selected Ontario fish species[1]

Table of Contents

APPENDICES

[1] Authors: L. W. Stanfield, M. Stoneman, G. Wichert, M. L. Jones and F. McGuiness

Section 6:

Habprogs database structure and management protocols

HabProgs database table structure

Section 6: Module 1

Table of Contents

1.0 Introduction

A number of tables are contained or are generated as part of the database system. All of these tables are available for extraction to other spreadsheet or database programs. They are also available to be reinterpreted using the query procedures within the program. The following is a partial list of the tables produced by this program and a brief summary of the information contained within them. It is incomplete because as new modules and options are built into the system, new tables are created. This appendix will be updated periodically to reflect these changes. The relationships of these tables within the database are shown in Figure 1.

1.1 Tables Created for Field Data by the Database System

Data: Stream ID Table

The Stream ID Table (Data: Stream Identification) is designed to ensure that each stream has only one official, properly spelled name. It contains the Stream Name, Stream Code, and a general stream location description.

Data: Site Identification Table

There can be many sites on any given stream. Each of the sites on a particular stream must have a unique code that is entered into the Site ID Table (Data: Site Identification).

Data: Sites - Years Sampled

A list each site and the years for which data are available in the database.

Data: Sample Identification

In some situations, a site may be sampled more than once in a year. Therefore, the database is linked to the sample number. Each sample has a corresponding record in the Sample Identification Table (Data: Sample Identification). This table also contains fields used to track the entry, verification, and correction process for each sample.

Data: Project Description Table

This contains the name and code for the project.

Data: Sample Project Codes

This is the list of the name and code for each project available in the database.

Data: Project Assessed Parameters

This is the list of all the parameters to be assessed by the project.

Data: Project Unassessed Parameters

This is the list of parameters not assessed by the project.

Data: Project Modifying Factors

This is the list of parameters that will be considered as modifying factors within the project design.

Data: Project Stratifying Factors

This is the list of criteria that were used to stratify a project.

Data: Project Reference

This lists any reports or manuscripts that were produced in association with a project.

Data: Site Description

This table (Data: Site Description) contains most of the information that is recorded on the Site Identification Form (Module 2). It also contains information on the stream and air temperature, the time they were taken, the daily maximum air temperature, and the source of the information (Module 6).

Data: Transect Level

The Transect Level Table (Data: Transect Level) is where all records that are recorded once per transect are recorded. This includes data on the width, compass bearing, and point spacing.

Data: Transect Banks

The Transect Banks Table (Data: Transect Banks) contains all the data that is recorded at each of the two banks associated with each transect, including the undercut depth, heights of the banks (bank angle), soil type, and the amount and type of vegetation.

Data: Pavement

This is an optional table for surveyors wishing to collect additional data on the pavement material size.

Data: Cover Type

This lists the type and quality of cover found at each point observation.

Data: Extra Point Pebble Counts

This is an optional table for surveyors wishing to collect additional pebble count data.

Data: Extra Point Velocity

This is an optional table for surveyors wishing to collect additional velocity data.

Data: Extra Wood

This is an optional table for surveyors wishing to collect additional wood concentration information.

Data: Invertebrate Sample

The Invertebrate Table (Data: Inverts) contains the depth and hydraulic head measured at the invertebrate sample site as well as a count of each of the taxonomic categories of invertebrates sampled.

Data: Invert Substrate

This table contains the median sizes of the ten substrate particles sampled at the invertebrate collection site.

Data: Efish Sample

This table contains the description of the electrofishing effort used to sample the fish community. Included are the start and stop times, the number of runs, the counter reading, and the names of the crew.

Fish Community Individual Specimens Table

All fish for which individual lengths or weights were measured are entered in this table (Data: Efish Individual Weights). Also included are fields to indicate if scale or otolith samples were taken and if the fish was preserved.

Fish Family Codes

This lists the family name and code for all Ontario fish.

Fish Species Codes

This lists the common and scientific name and code for all Ontario fish.

Data: EFish habitat Measurements

This contains the length and width data collected while electrofishing.

Data: EFish Individual Weights

Individually measured and weighed fish are recorded in this table.

Data: EFish Bulk Weights

Fish that were weighed in batches rather than individually are recorded in this table (Data: Efish Bulk Weights). The number of fish in the batch, the bulk weight, and the number of fish preserved are all stored in this table. The bag number assigned to the preserved sample is also stored here.

Data: Unique Features

This table contains the data from the Site Features Form that describes the presence of features at the site which could influence the habitat or the fish community. This includes chemical or nutrient sources, historical alterations, and the riparian composition.

1.2 Summary Tables derived from raw data available from within HabProgs

Most summary procedures created within HabProgs store the data in summary tables that are available for export by users of this application. Below we summarize a partial list of the summary tables that are currently available for export. As more methods are developed to summarize these data this list will be updated.

summary:geomorphic indices

Contains data on the average width, width/depth ratio, d16, d50 and d84 for the point observation and maximum particles in the ring as well as the standard deviation for the point and maximum particle size observations.

aaPHSthermal:

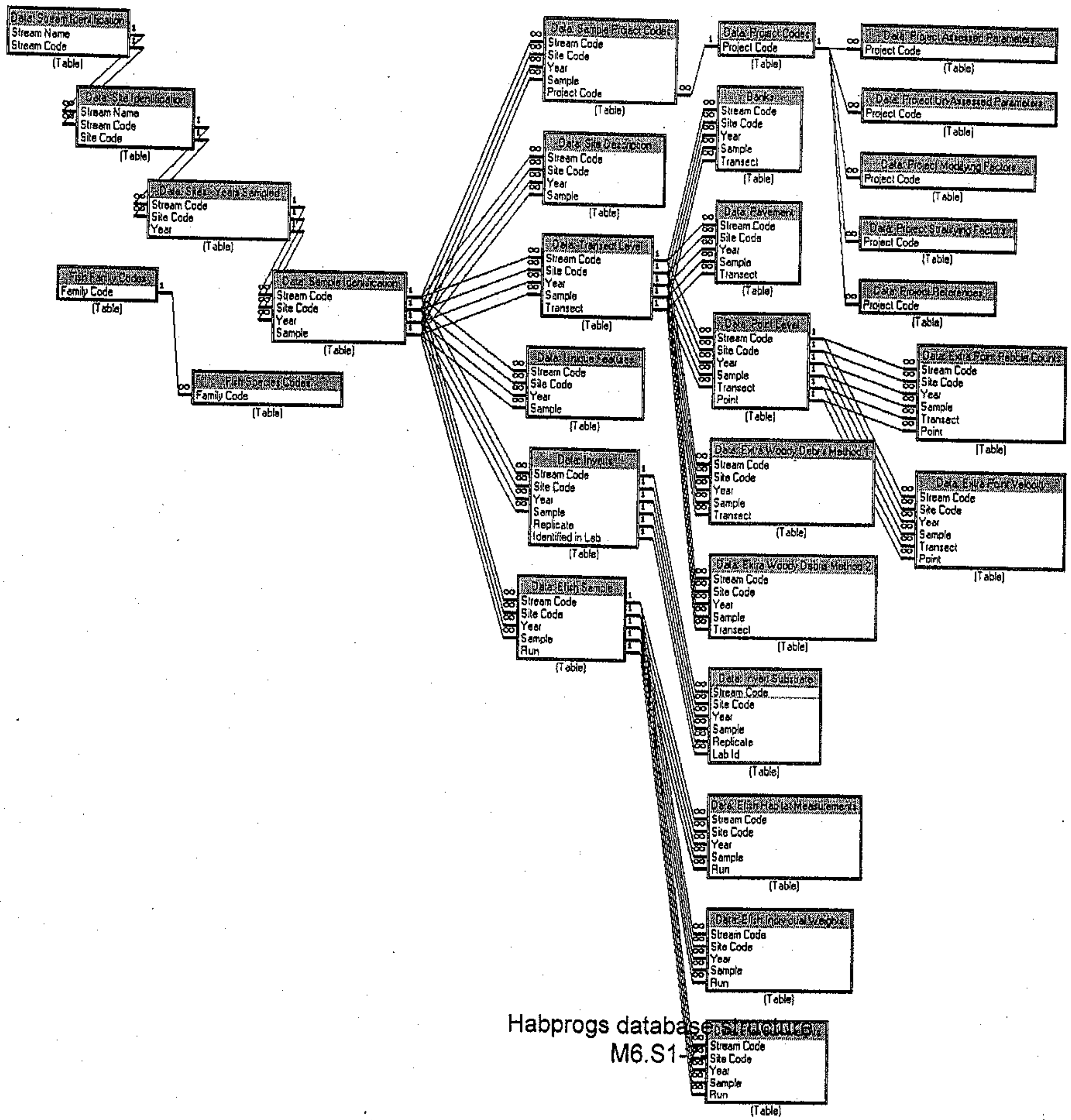
The thermal deviation from a reference cold water stream for each site sampled

inverts: percent taxa table

Contains the number and percent of the sample for each invertebrate group identified

fish: summary by area no size classes

Summary of the abundance and biomass of species caught at each site by area of available habitat. This data combines size groupings for each species



Habprogs database structure
M6.S1-