Fall 2006

Ecological patterns, community classification, and a comparison of approaches for predicting biological and habitat reference conditions in New Hampshire streams

Brian R. Frappier
University of New Hampshire, Durham

ECOLOGICAL PATTERNS, COMMUNITY CLASSIFICATION, AND A COMPARISON OF

APPROACHES FOR PREDICTING BIOLOGICAL AND HABITAT REFERENCE

CONDITIONS IN NEW HAMPSHIRE STREAMS

BY

BRIAN R. FRAPPIER

B.S. University of Maine, 1999

M.S. University of New Hampshire, 2001

DISSERTATION

Submitted to the University of New Hampshire

in Partial Fulfillment of

the Requirements for the Degree of

Doctor of Philosophy

in

Natural Resources and Environmental Studies

September, 2006

UMI Number: 3231350

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI®
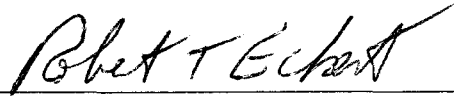
UMI Microform 3231350

Copyright 2006 by ProQuest Information and Learning Company.

All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.
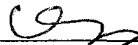
ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

This dissertation has been examined and approved.
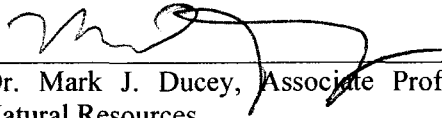
_____
Dissertation Director, Dr. Robert T. Eckert,
Professor of Environmental Conservation and
Forest Resources
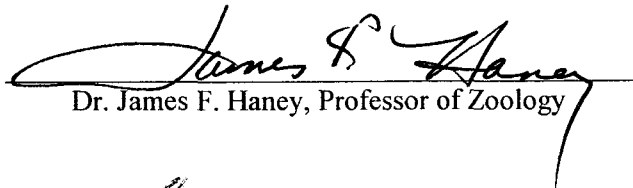
_____
Dr. Kimberly J. Babbitt, Associate Professor of
Wildlife Ecology
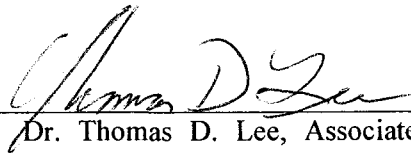
_____
Dr. Mark J. Ducey, Associate Professor of
Natural Resources

_____
Dr. James F. Haney, Professor of Zoology

_____
Dr. Thomas D. Lee, Associate Professor of
Forest Ecology

_____
Date

ACKNOWLEDGEMENTS

I would like to thank all of the wonderful people who helped me through grueling long days in the field and many boring hours in the laboratory: Brian Topping, Sarah Mikulak, Danielle Adams, Jon Miller, Kirsten Nelson, Liz Durfee, Lindsey Scott, and Ethan Gyles. My advisor, Bob Eckert, gave me many years of support and advice; I owe most of what I know about science to him and my doctoral committee. The most special thanks is reserved for Amy Frappier for filling in on a moments notice and patiently listening to my esoteric grumblings with a genuine interest. You all really made this an enjoyable experience and I hope you are proud of what has resulted from all of your hard work.

iii

TABLE OF CONTENTS

CHAPTER   PAGE

v

vi

x

ABSTRACT

STREAM COMMUNITY PATTERNS AND CLASSIFICATION OF MINIMALLY

IMPACTED NEW HAMPSHIRE STREAMS AND A COMPARISON OF NOVEL

APPROACHES FOR PREDICTING BIOLOGICAL AND PHYSICAL HABITAT

REFERENCE CONDITIONS

by

Brian R. Frappier

University of New Hampshire, September, 2006

Reference conditions play a vital role in many challenges facing both conservation and natural resources management. This dissertation sought to establish minimally-impacted reference conditions for stream biota and habitat in New Hampshire and explore alternative statistical methodologies to predict reference conditions for biological and habitat assessments. The fish, stream-dwelling salamander, macroinvertebrate and periphyton assemblages as well as the co-occurring physical habitat and riparian conditions of 76 minimally –impacted first to fourth order streams in New Hampshire were estimated using USEPA Environmental Monitoring and Assessment Program protocols over a four year period. Several statistical approaches and data standardizations for classifying multi-taxonomic assemblages were investigated for the strength of the classification they produced; log transformed abundances classified using TWINSPAN produced the best classification as measured using specific criteria. Seven natural biotic community types primarily arranged along the longitudinal stream profile were classified. Geographic classifications based on ecoregions and watersheds poorly explained organism distributions and abundances. Organism distributions were primarily associated with substrate characteristics, elevation, latitude, and the proportion of mesohabitat types (e.g. pool, riffle, etc.).

xvi

A new approach to constructing a biological assessment index that is based on the Bray-Curtis percent similarity between the observed and predicted communities was developed to allow taxa density information into the multivariate predictive assessments. Separate linear regression models to predict the densities of each taxon resulted in the most accurate predictions of expected community structure. Multivariate predictive models that included classification steps were not in general less accurate than approaches based on continuous prediction of taxon densities such as nearest-neighbor or ordination-based analyses. Including abundance information into the predictive models did not increase relative prediction error compared to an AusRivAS-style assessment index based solely on predicted taxon occurrences. Habitat prediction followed similar results. Inter-annual variation in three streams sampled every year of the study was highest in the vertebrates and lowest in the macroinvertebrates. In contrast, vertebrate assemblages were more resistant to a summer spate than the macroinvertebrates. Greater sampling intensity in the field and laboratory are probably the only remaining avenues for increasing assessment accuracies and reducing unexplained variation in reference conditions..

INTRODUCTION

We rely on ecosystem goods and services (Costanza et al. 1997). On a human time-scale of hundreds of years, many ecosystems are naturally relatively stable compared to the dramatic changes in ecosystem characteristic that can accompany longer time frames (thousands to millions of years), such as glacial cycles and long-term changes in precipitation patterns. Thus, we have come to expect certain levels of ecosystem goods and services in our planning and natural resources management. While the exact nature of the relationship is far from clear, many ecosystem functions appear to be affected by changes in species composition and richness (Loreau et al. 2002, Schmid 2002). Changes in community structure and composition due to human-induced stresses or alterations to ecosystems have the potential to negatively impact our resource planning and conservation. This perspective on the interactions between humans and ecosystems has partly given rise to the concept of ecosystem management (Grumbine 1994, Christensen et al. 1996).

The concern for species and ecosystems goes beyond our need for goods and services. Many see an inherent value in natural ecosystems little altered by humans and believe that all species have an intrinsic right to exist. Understanding and appreciating the natural character of ecosystems is an important part of this concern. In addition, efforts to protect species depend on a firm understanding of the distributions of species and their interactions with each other and the physical environment. Ensuring that multiple representatives of each species, community, or ecosystem in a region are located in conservation areas has become a key goal of biological diversity conservation (Noss 1987). As many regions contain far more species than can be directly managed, assessment of representation has often been focused at the higher community and ecosystem level (Noss 1990, Franklin 1995, Grossman et al.1998).

1

It is clear that conservation biology and natural resources management for our needs are often intertwined. In both of those pursuits, reference conditions play a vital role as they are the basis for answering many of the questions and challenges facing both conservation and natural resources management:

- Are All Terrain Vehicles affecting this stream?

- Is this a good example of a natural community for conservation representation?

- Has fishing reduced the average size of the mature cod population?

- What should our large woody debris density targets be for restoring this stream's physical habitat?

- How do salamanders respond to human disturbance of forest cover?

- Are these newly instituted pollution controls reducing the impacts on this stream's biota?

Understanding the natural conditions and functioning of ecosystems provides a baseline for assessing the resource impacts of management decision and changes wrought by humans on ecosystems. Yet, very little is known about the natural ecological patterns in New Hampshire streams, providing very little guidance as to the reference conditions needed to adequately answer the common questions facing lotic conservation and management in the State.

The collected papers that make up this dissertation represent a first attempt at providing vitally needed reference conditions, both biological and physical, for New Hampshire streams. The goals of this dissertation were to:

2

1. Classify the natural stream communities in New Hampshire using the major taxonomic groups periphyton, macroinvertebrates, fish, and stream dwelling salamanders (Chapter 3)

2. Explore the ecological patterns in New Hampshire streams (Chapter 3)

3. Produce a statistical model to map community locations using GIS (Chapter 3)

4. Predict organism distributions for constructing the most accurate theoretical biotic reference conditions possible for biological monitoring (Chapter 4)

5. Similarly predict theoretical habitat conditions for habitat assessment (Chapter 5)

6. Explore the temporal dynamics in stream ecosystem composition and structure (Chapter 6)

The chief mechanism for accomplishing these goals was a database of the biotic, physical, and chemical characteristics of minimally impacted streams in New Hampshire built over a four year period from 2002 to 2005. Chapter 1 describes the field and laboratory methods for collecting this information. Along the way, it became clear that the conventional methods for estimating periphyton abundance were woefully inadequate. Hence, Chapter 2 describes an investigation seeking to improve estimation of periphyton biovolume by comparing estimation error resulting from the standard methods compared to a line-intercept technique.

3

CHAPTER I

FIELD AND LABORATORY METHODS

Reference Stream Selection

Locating minimally impacted sites is the most important step in establishing reference conditions for biological criteria (Hughes 1995) and identifying natural communities. Stream segments, defined as the length of stream between two tributaries, were randomly selected using the GRANIT hydrography GIS layer (Complex Systems Research Center 2001). To ensure adequate representation of the full range in stream types across New Hampshire, the selection of segments was stratified to be proportional to the total permanent stream length in each aquatic ecoregion (Omernik 1987).

Each randomly selected segment was evaluated for anthropogenic impact using GIS layers identifying known point and non-point source pollution, land-use, right-of-ways, dams, public and private water extraction, groundwater hazards, clear-cuts, underground storage tanks, and digitized recent aerial photos available from New Hampshire Department of Environmental Services or New Hampshire GRANIT. Segments identified as having any upstream water quality threats using those layers were discarded. Once a stream segment was selected and visited for field sampling, an additional assessment was made for any potential impacts not detected using the remotely collated GIS data.

Field Data Collection

Stream Reach Location

Field methods, with only a few exceptions, followed the wade-able streams techniques developed by the United States Environmental Protection Agency's Environmental Monitoring and Assessment Program (hereby known as USEPA-EMAP-SW; Lazorchak et al. 1998). What follows is a synopsis of the USEPA-EMAP-SW field methods along with a description of the slight modifications made to them. For a full description accompanied by explanatory figures and summary tables, consult the USEPA-EMAP-SW manual. Departures from the USEPA-EMAP-SW methods are typeset in italics to ease identification.

All sampling was performed during the base flow period of 15 June to August 30 every year between 2002 and 2005, inclusive. Sites that experienced a recent spate were not sampled for 6 weeks following the disturbance. In 2002, a regional drought caused small, normally permanent, streams to dry; consequently, field sampling was ended early. In 2003, a series of heavy summer rainstorms caused many streams in the White Mountains regions to reach flood-stage and sampling was abandoned early that year as well. Although the order in which segments were sampled in each base-flow period was randomized, for logistical reasons sampling in each year was clustered into the major regions of New Hampshire. Southern New Hampshire as far west as U.S. route 93 was sampled in 2002, the White mountain region in 2003, southwestern New Hampshire in 2004, and region north of the White Mountains in 2005. A few additional sites in the Seacoast region were sampled in 2005 to maintain the stratification proportions of the sampling design.

The basic sampling unit was the stream reach, defined as a length of stream 40 times the wetted width, with a minimum length of 150 m and a maximum width of 500 m. The reach center was randomly located along the selected stream segment at least 100 m upstream or

5

downstream of the bounding tributaries and upstream of any roads or trails during the stream segment selection and screening process. The wetted width was determined by measuring the width at a few representative locations near the reach center and averaging the result to the nearest meter. Reaches were divided into 10 subsections delineated by 11 transects oriented perpendicular to the flow and spanning the width of the stream. The first downstream transect was placed at the lower boundary of the reach and the each successive transect was equally spaced along the reach (i.e. 4 wetted widths apart). Transects were placed and marked endeavoring to disturb the stream as little as possible. If an unknown stream confluence or impoundment (e.g. beaver pond) was found while setting the transects, the stream reach was moved upstream or downstream to accommodate the 100 m proximity limit to those features.

Water Chemistry

Water chemistry samples were collected at the reach center (middle transect) before any other sampling activities. Water was collected from the main flow using a 40 ml plastic syringe and filtered through a 0.7 μm pre-combusted (5+ hours at 450 C) glass fiber filter into an acid-washed 60 ml HDPE bottle. The filled bottled was capped and kept in a shady part of the stream to keep the sample cool. The water sample was frozen upon return to the laboratory at the end of the day. Water pH, temperature, and conductivity were measured in the same area the water chemistry sample was taken using an Oakton 35630 portable pH/conductivity/temperature meter.

Physical Habitat Measurements and Biological Sampling

Biological sampling was performed before most habitat measurements. A randomized, systematic spatial sampling design was used to locate biological sampling points on the transects. Starting at the downstream transect, a sampling point (left, center, or right) on the first transect

was randomly chosen. Subsequent sample locations were assigned to each upstream transect, alternating in order as left, center, or right. The type of habitat present at the biological sampling point (i.e. pool, riffle, glide, or rapid) and the dominant substrate type (i.e. fine/sand, gravel, coarse, or macrophytes) were recorded.

Periphyton were collected at the transect sample points before macroinvertebrate sampling. In erosional habitats, a sample of rock or wood substrate was removed from the stream and a 12 cm^2 area on the upper surface of the substrate brushed with a stiff-bristled toothbrush for 30 seconds to dislodge periphyton. Dislodged periphyton were washed into a 500 ml bottle using stream water. In depositional habitats, the top 1 cm from a 12 cm^2 area of soft sediment was vacuumed into a 60 ml syringe and placed in the 500 ml bottle. All periphyton samples were composited into the same 500 ml bottle and preserved with 10% formalin to a final concentration of 2%.

Macroinvertebrates were collected at each of the transect sample points using a 500 μm d-net (net). In flowing water habitats, the net was placed securely on the stream bottom. Large and small rocks in a 0.09 m^2 sample area in front of the net were rubbed to dislodge organisms and large obvious organisms were hand-picked and placed into the net. A 20 second kick sample of the 0.09 m^2 sample area was then taken. In wade-able pool habitats, large organisms in a 0.09 m^2 sample area were hand-picked and placed into a net. The same 0.09 m^2 area of substrate was disturbed by vigorous kicking and a 20-second sample was collected by dragging the net repeatedly through the disturbed area just above the bottom while kicking. If the water was too shallow to use the net, the 0.09 m^2 area of substrate was stirred by hand and a US Standard #30 sieve used instead of the net. Net or sieve contents were rinsed into a bucket half filled with water that contained all of the samples as a single composite. The composite sample was elutriated to remove inorganic sediment, placed in large plastic containers, and preserved with 95% ethanol.

7

Latitude and longitude were measured using a Garmin portable GPS unit at the reach center. Discharge was measured using a small portable flow meter at 20 equally spaced points along a transect placed perpendicular to the flow. There was a minimum spacing of 10 cm between measurement points. This, some flow estimations were made using fewer than 20 measurements. Depth at the flow-measurement points and the wetted width were recorded to calculate the cross-section. The discharge estimation transect was located in an area with flow characteristics that would yield the most accurate discharge estimation (e.g. confined weir-like flow). If there was no section of the segment that would yield a good estimate because the flow was too slow, then the neutral buoyant object approach was used.

The measurement of other habitat parameters can be broadly divided into those parameters measured along the transects, those measured using the stream area 5 m on either side of the transect, and those measured in the riparian area 5 m on either side of the transect to a distance of 10 m from the stream edge. The transect methods and between-transect area methods will be described for the first downstream transect; the methods were then repeated at each successive upstream transect and between-transect area.

The following physical habitat variables were measured along each transect: wetted width, angle of each bank, undercut length, bankfull channel width and height above water surface, and canopy cover in the stream center and edges. The embeddedness, size substrate particle size (in 11 categories), and water depth were measured at 5 equally spaced points along the transect starting at one bank. *Additional transects labeled as 'supplementary transects' to take additional substrate and depth measurements in the USEPA-EMAP-SW methods were not taken* after the first sampling year due to time constraints.

In the area of stream 5 m on both sides of the transect, the following habitat measurements were made: Tally of large woody debris in the bankfull cross-section greater than 10 cm in diameter at the large end into 4 classes based on large end diameter (0.1-0.3 m and >0.3 m diameter) and length (1.5 – 5 m and >5 m). Additionally, large woody debris above the

8

bankfull cross-section but over the stream area were tallied into the same diameter/length classes. *The larger diameter/length classes described in the USEPA-EMAP-SW methods were not utilized* as very few pieces of woody debris reach those sizes in the northeastern Unites States. The areal cover of several categories of fish cover habitat (small brush, large woody debris, undercut banks, overhanging vegetation [< 1 m from water surface], filamentous algae, macrophyte, boulders, and in-channel live trees or roots) were measured using four ordinal classes (0 = absent, 1 = 1-10%, 2 = 11-40%, 3 = 41-75%, 4 = 76-100%). In a departure from the EPA-EMAP-SW methods, *the areal cover of live-trees or roots, overhanging vegetation, and undercut banks were measured as the percentage of the bank on both sides of the stream covered by those habitat features* as the cover of those habitat features over the entire water surface area was never greater than 10%. The water surface gradient and directional bearing (aspect) were measured between pairs of transects for a total of 10 measurements. *The thalweg profile was not measured.*

For each side of the stream, the riparian zone bounded by an area 5 m upstream and downstream of the transect and a distance of 10 m from the stream edge was visually inspected for riparian vegetation characteristics in 3 layers: canopy (>5 m), under-story (0.5-5 m), and ground cover (0-0.5 m). The dominant vegetation type of the canopy (none, conifer, deciduous, evergreen, or mixed) was recorded and the areal coverage of vegetation in the separate layers measured using four ordinal categories (0 = absent, 1 = 1-11%, 2 = 11-40%, 3 = 41-75%, 4 = 76-100%).

Abundances of vertebrates (fish and amphibians) and crayfish (Decapoda: Cambaridae) were estimated after all other sampling activities using a backpack electroshocker. *The USEPA-EMAP-SW methods did not estimate amphibian or crayfish abundances.* A single-pass electro-fishing method attempting to fish all available cover in the entire reach was used starting at the downstream limit of the reach. An anode net was used and one additional person with a net followed behind the person electroshocking to catch organisms that bypassed the electroshocker in the flow. In contrast to the variable time-limit for electro-fishing in the USEPA-EMAP-SW

methods, *a consistent effort was applied throughout the reach for 50 minutes by fishing for 5 minutes between each transect.* Block nets were placed at the downstream and upstream limits of the sampling reach when the sample reach was a large continuous pool. Collected fish were placed in a bucket of water, identified, and returned to the stream; however, fish that were not caught but could be confidently identified by the shocker or netter were also tallied. An additional person was responsible for the bucket and recording the identifications. Fish of questionable taxonomy were killed and preserved in 70% ethanol for lab identification.

Lab Analyses

Physical Habitat

Basin- (watershed-) level physical factors were estimated using ArcView and the hydrography and digitized 1:24,000 USGS topographic quads datasets from New Hampshire GRANIT. The following basin-scale parameters were measured:

- Drainage area

- Elevation

- Maximum elevation

- Cumulative perennial stream length

- Cumulative intermittent stream length

- Distance from source (furthest point in network)

- Distance from nearest impoundment (wetland or lake in the hydrography layer)

- Level IV aquatic ecoregion membership (Omernik 1987)

- Stream order

- Percent of watershed as lakes

10

- Percent of watershed as wetland,

- Length of roads per watershed area

- Percent watershed area in three bedrock types (metamorphic, plutonic, and volcanic)

Additionally, an index of the relative magnitude of the annual (spring) flood was calculated as the base-flow wetted cross-section divided by the bankfull cross-section. A listing of the habitat variables taken at all scales can be found in Table 1-1.

Vertebrate and Cambaridae Data Processing

Vertebrate and Cambaridae densities were calculated by dividing the number of organisms by taxon from the 50-minute electro-fishing by the length of stream sampled to yield a density per 100 m of stream length with a sampling effort of 50 minutes.

Macroinvertebrate Sample Processing

A fixed-count sub-sample procedure based on the USGS-NAWQA protocols (Moulton et al. 2000) was used to estimate abundances of aquatic macroinvertebrates. Each sample was rinsed and sieved using a 500 μm sieve. The sample was uniformly distributed in a sub-sampling frame (stage-1 sub-sampling frame) with 15 grids. The grids to be sorted were randomly selected from the stage-1 sub-sampling frame. An estimate of the average number of organisms per stage-1 grid was obtained using two grids and the number of grids needed to achieve a 500-organism fixed count calculated. Doberstein et al. (2000) found no significant differences in several taxa measurements between whole sample processing and 1000 count sample processing; lower count sub-samples showed decreased power to detect differences between sites. However, resource constraints necessitated sub-sampling to 500 organisms, which improved on the very poor

11

precision of 100-300 fixed-count sub-samples with much less effort than doing a 1000-count sub-sampling. *A second stage was not used as the stage 1 grid was large enough (30 cm by 18 cm) to consistently achieve the critical density required to sub-sample 500 organisms with at least 3 grids.* Grids were sorted using a dissecting microscope at 8x power. Large, rare organisms were collected from any remaining unsorted portions of the sample in a 15 minute search.

Most macroinvertebrates were identified to the family level, using keys by Peckarsky et al. (1990), Thorp and Covitch (2001), and Merritt and Cummins (1996), though some were only identified to higher taxonomic level (e.g. Oligochaeta, Hydrachnida, Turbellaria, Nematomorpha). There has been a rich conversation regarding the role of taxonomic resolution in biological monitoring (Lenat and Resh 2001). Several studies have found that identification to genus-species level (also known as lowest-practical level) either classifies sites into biotic groups or detects suspected impacts better than family level (Barton 1996, Hawkins and Norris 2000, Hawkins et al. 2000, Marchant and Hehir 2002). In contrast other studies have found family-level as sensitive, or more so, than genus-species level identifications (Olsgard et al 1997, Hewlett 2000).

Bowman and Bailey (1997) found a very high correlation between family-level communities and genus-level communities. They suggest that a possible explanation for the differences in impact sensitivity with taxonomic level found in other studies may be related to whether the data is qualitative (usually presence-absence) or quantitative; Bowman and Bailey found a high correspondence between family and genus-species communities only when they conducted their analyses with abundance data, not binary data. McCreadie et al. (1997) found that black fly species co-occurrences within a group of very similar streams were no different from random. The distributions of species within higher taxonomic levels in families other than blackflies may also be random with only the distributions of higher taxonomic levels following tightly defined habitat templates. Indeed, Reynoldson et al. (1997) found that AUSRIVAS, which uses family-level presence absence data, more reliably detected known pollution than

12

multimetric approaches that use genus-species identifications. They suggest that the way in which reference sites are aggregated for comparison with potentially polluted sites plays a greater role in determining impact sensitivity than taxonomic resolution.

Other problems present themselves with the use of lower taxonomic levels. Volunteers almost certainly cannot identify to lower than family level. There is often much wasted effort in identifications to lower taxonomic levels. Ordination analyses usually drop rare taxa, which results in many more individuals dropped from analysis at lower taxonomic levels than higher. Additionally, damaged or very small individuals often preclude confident identifications to lower than family level that require aggregation to higher taxonomic levels for all of the individuals to be used in statistical analysis. Resource constraints, even in areas with high-quality genus-species level prediction approaches such as RIVPACS in the United Kingdom, often force local users to use the family-level versions in common practice.

The balance of evidence suggests that there is only a small decrease in impact sensitivity due to identifications at the family level compared to genus-species identifications, and that that decrease may not exist when quantitative abundance data is available. Doberstein et al. (2000), found that sub-sampling significantly increased the variance in component metrics. It appears that the use of high quality quantitative data, powerful approaches for explaining variation in the reference sites, and a high sub-sampling fixed-count target are much more important considerations than taxonomic level for increasing impact sensitivity of a bioassessment approach. Thus, the level of effort spent in sub-sampling was increased to 500 organisms from the usual 100-200 organisms, rather than increase the effort spent in taxonomic identification, as that resource trade-off was more likely to result in higher sensitivity to potential impacts.

Because exactly 1 m^2 of stream benthos was sampled for macroinvertebrates, the density of macroinvertebrates (#/m^2) in the stream reach and the estimated number of macroinvertebrates in the composite sample were equivalent and were calculated as:

13

$$Density = total\ \#\ organisms\ in\ sub\text{-}sample \cdot \frac{\#\ grids\ sorted}{total\ \#\ of\ grids}$$

The same equation was used to calculate the densities of individual taxa as well. Macroinvertebrates from the large-rare search were added to the final density estimate only if no individuals of a large-rare taxon were collected in the 500-organism sub-sample.

Periphyton Sample Processing

Periphyton were processed in a blender at low speed for 10 seconds to break up filamentous algae and Lugol's iodine added to help distinguish Chlorophyta from Cyanophyta. A line-intercept method using ten 50 mm transects in a 20 mm X 50 mm X 0.4 mm deep counting chamber (0.4 ml) was used to estimate periphyton biovolume. Transects were split between two counting chambers to reduce bias associated with using a single 0.4 ml aliquot. Periphyton were identified to the genus level for soft algae and morphometric type for the Bacillariophyta (naviculoid, cymbelloid, centroid, and a few genera that were easily distinguished such as *Gyrosigma*) using the key by Prescott (1978). Details of the development and testing of this novel application of the line-intercept method can be found in Chapter 2. Periphyton abundance was calculated in biovolume units of mm^3 per m^2 of stream benthos using the equations in Chapter 2.

Chemical Analyses

Water chemistry samples were analyzed for a full range of common constituents. Major anions were analyzed using ion chromatography with suppressed conductivity; major cations using ion chromatography; Ammonia using alkaline phenol method (Berthelot reaction) with discrete colorimetric analysis (based on EPA 350.1); Phosphate using molybdate blue method and

14

discrete colorimetric analysis (based on EPA 365.2); Silica using discrete colorimetric analysis (based on EPA 370.1); dissolved organic Carbon as non-purgeable organic Carbon using high-temperature catalytic oxidation (HTCO); total dissolved nitrogen using HTCO with chemiluminescent detection; and total alkalinity according to the standard color change titration method (Clesceri et al. 1989). Dissolved organic Nitrogen was calculated as the fraction of total dissolved Nitrogen remaining after subtracting Ammonia and Nitrate (all measured as mg N/L). A complete list of the individual compounds and ions analyzed along with their respective units can be found in Table 1-1.

Table 1-1. Habitat descriptors estimated for minimally impacted stream reaches in New Hampshire.

Brief variable name	Description and Units
Granit_ID	GRANIT hydrography ArcView layer ID number
Ecoregion4	Level IV EPA Aquatic Ecoregion
Ecoregion3	Level III EPA Aquatic Ecoregion
Lat	Latitude in decimal degrees
Long	Longitude in decimal degrees
Order	Stream order
Wetlands	Proportion of watershed as wetlands
Lakes	Proportion of watershed as lakes
Permanent_stream	Total length of permanent streams (m)
Intermittent_stream	Total length of intermittent streams (m)
Highest_point	Maximum elevation in watershed (m)
Distance_impound	Distance to nearest upstream impoundment (km)
Distance_source	Distance to furthest point in the upstream stream network (km)
Area	Watershed area (ha)
Metamor	Proportion area of metamorphic bedrock
Volcan	Proportion area of volcanic bedrock
Pluton	Proportion area of plutonic bedrock
Length	Reach length (m)
Elevation	M
Discharge	m^3/sec
Slope	Degrees
Aspect	Degree deviation from south
Fine	Proportion of 11 macroinvertebrate and periphyton sample sites classified as "fine" substrate
Gravel	Proportion of 11 macroinvertebrate and periphyton sample sites classified as "gravel" substrate
Coarse	Proportion of 11 macroinvertebrate and periphyton sample sites classified as "coarse" substrate
Macro	Proportion of 11 macroinvertebrate and periphyton sample sites classified as "macrophyte" substrate
Pool	Proportion of 11 macroinvertebrate and periphyton sample sites classified as "pool" habitat
Glide	Proportion of 11 macroinvertebrate and periphyton sample sites classified as "glide" habitat
Riffle	Proportion of 11 macroinvertebrate and periphyton sample sites classified as "riffle" habitat
Rapid	Proportion of 11 macroinvertebrate and periphyton sample sites classified as "rapids" habitat
pH	Water Ph
Temp	Water temperature (°C)
Cond	Specific conductivity (uS)
Embedded	Mean % embeddedness of substrate
Phi	Mean phi value of substrate particles
SDphi	Standard deviation of phi values for substrate particles
Fine_sub	Proportion of 55 substrate measurements classed as fine

Table 1-1. Continued.

Brief variable name	Description and Units
Sand	Proportion of 55 substrate measurements classed as sand
Fine_gravel	Proportion of 55 substrate measurements classed as fine gravel
Coarse_gravel	Proportion of 55 substrate measurements classed as coarse gravel
Cobble	Proportion of 55 substrate measurements classed as cobble
Boulder_sub	Proportion of 55 substrate measurements classed as boulder
Bedrock	Proportion of 55 substrate measurements classed as bedrock
Depth	Mean water depth (m)
Max_depth	Maximum water depth (m)
Width	Mean channel width (m)
Cross_section	Mean channel width x mean water depth (m^2)
Bankfull_width	Mean bankfull channel width (m)
Bankfull_height	Mean bankfull channel height (m)
Bankfull_crossection	Mean bankfull channel width x mean bankfull channel height (m2)
Flood_mag	Annual flood magnitude indexed by the bankfull cross-section divided by wetted cross-section (unit-less)
Undercut	Mean undercut length (m)
Angle	Mean Bank angle (degrees)
Filamentous	Mean cover index of filamentous algae in the channel (1-4)
Macrophytes	Mean cover index of macrophytes in the channel (1-4)
Woody_debris	Mean cover index of large woody debris in the channel (1-4)
Brushwood	Mean cover index of brush in the channel (1-4)
Live_trees	Mean cover index of live trees along the bank (1-4)
Overhanging	Mean cover index of overhanging vegetation along the bank (1-4)
Undercut_banks	Mean cover index of undercut along the bank (1-4)
Boulders	Mean cover index of boulders in the channel (1-4)
Canopy_cover	Mean % canopy cover
Big_trees	Mean cover index of large trees (>0.3 m DBH) in the canopy riparian vegetation (1-4)
Small_trees	Mean cover index of small trees (<0.3 m DBH) in the canopy riparian vegetation (1-4)
Shrubs	Mean cover index of shrubs in the under-story riparian vegetation (1-4)
Tall_herbs	Mean cover index of tall herbs (>0.5 m) in the under-story riparian vegetation (1-4)
Short_herbs	Mean cover index of short herbs (<0.5 m) in the ground cover riparian vegetation (1-4)
Bare	Mean cover index of bare ground in the under-story riparian vegetation (1-4)
Deciduous	Proportion of riparian canopy vegetation classified as "deciduous"
Coniferous	Proportion of riparian canopy vegetation classified as "coniferous"
Mixed	Proportion of riparian canopy vegetation classified as "mixed" (>20% other)
None	Proportion of riparian canopy vegetation classified as "none present"
IN1.5-0.1	Density (per m of reach length) of large woody debris 0.1-0.3 m in diameter and 1.5 -5 m in length in the bankfull channel
IN1.5-0.3	Density (per m of reach length) of large woody debris >0.3 m in diameter and 1.5 -5 m in length in the bankfull channel

17

Table 1-1. Continued.

Brief variable name	Description and Units
IN5-0.1	Density (per m of reach length) of large woody debris 0.1-0.3 m in diameter and 5-15 m in length in the bankfull channel
IN5-0.3	Density (per m of reach length) of large woody debris 0.1-0.3 m in diameter and 5-15 m in length in the bankfull channel
IN15-0.3	Density (per m of reach length) of large woody debris >0.3 m in diameter and >15 m in length in the bankfull channel
OUT1.5-0.1	Density (per m of reach length) of large woody debris 0.1-0.3 m in diameter and 1.5 -5 m in length above the bankfull channel
OUT1.5-0.3	Density (per m of reach length) of large woody debris >0.3 m in diameter and 1.5 -5 m in length above the bankfull channel
OUT5-0.1	Density (per m of reach length) of large woody debris 0.1-0.3 m in diameter and 5-15 m in length above the bankfull channel
OUT5-0.3	Density (per m of reach length) of large woody debris >0.3 m in diameter and 5-15 m in length above the bankfull channel
Total_LWD	Total tally of all large woody debris
IN_LWD	Density of all large woody debris within the bankfull channel (per m of reach length)
OUT_LWD	Density of all large woody debris above the bankfull channel (per m of reach length)
Chloride	Chloride (mg Cl/L)
Nitrate	Nitrate (mg N/L)
Sulfate	Sulfate (mg S/L)
Sodium	Sodium (mg Na/L)
Potassium	Potassium (mg K/L)
Magnesium	Magnesium (mg Mg/L)
Calcium	Calcium (mg Ca/L)
Ammonia	Ammonia (ug N/L)
Phosphate	Phosphate (ug P/L)
Silica	Silica (mg SiO_2/L)
DOC	Dissolved organic Carbon (mg C/L)
TDN	Total dissolved Nitrogen (mg N/L)
DON	Dissolved organic Nitrogen (mg N/L)
Total_alkalinity	Total alkalinity (mg $CaCO_3$/L)
Vertebrate taxa	Number per 100 m of reach captured in 50 minutes of electro-fishing
Macroinvertebrate taxa	Number per m^2 of benthos
Periphyton taxa	Biovolume (mm^3/m^2 benthos)

18

CHAPTER II

LINE-INTERCEPT LABORATORY SUB-SAMPLING MORE EFFICIENTLY ESTIMATES PERIPHYTON BIOVOLUME

Summary

Algal communities have been included as a biological monitoring assemblage in a number of national, large-scale, and long-term biological monitoring and assessment programs of streams, lakes, and wetlands. However, there has been no substantial investigation of alternative enumeration strategies. We compared the sampling efficiency of the common periphyton sub-sampling strategies of counting 300-cells by strip counts or random fields and measuring 10 cells per taxon for biovolume with line-intercept sampling (LIS) using samples collected from four streams differing in canopy cover and discharge. We also simulated the optimum ratio of cells tallied to cells measured for biovolume for all approaches. LIS was 1.6 to 5.5 times more efficient at estimating algal taxa biovolume than the 300 cell-count methods. LIS performed best relative to conventional 300-cell count methods in streams with high biovolume, though efficiency gains were still noticeable in low biovolume streams. In addition, LIS tended to detect more taxa than the conventional 300-cell target biological assessment methods.

In simulations, LIS was more efficient than the conventional methods for nearly all combinations of target cell-count and cells measured for biovolume. No single optimal ratio of cells-counted to cells-measured was found. The target cell-count that resulted in the

19

greatest efficiency was higher in assemblages with large variation in cell density. Likewise, assemblages with high variance in taxa biovolume were most efficiently sampled by measuring more cells.

An increase in sampling efficiency due to greater adoption of LIS sampling may result in more accurate estimation of the relationships between periphyton community composition and environmental factors and more sensitive detection of impacts using stream periphyton. Additionally, LIS may improve estimates in other ecological research fields employing microscopic counting chambers.

Introduction

Although algae have a long history of use as a potential biological indicator of ecosystem health and anthropogenic impact to aquatic systems (Patrick 1949), they have been gaining increasing attention in recent years (see recent reviews by McCormick & Cairns 1994; Lowe & Pan 1996; Danielson 1998; Stevenson & Smol 2002). To date, algae have been used to assess impacts in a variety of ecosystems (see for example Porter et al. 1993; Adamus 1996; Pan et al. 1996; Jameson et al. 1998; Hill et al. 2001) , with indications that they may respond differently to impacts and thus be complementary with other aquatic biological assemblages in those same systems. As a result, algal communities have been included as a biological monitoring assemblage in a number of national, large-scale, and long-term biological monitoring and assessment programs of streams, lakes, and wetlands (Porter *et al.* 1993; Lazorchak, Klemm & Peck 1998; Barbour *et al.* 1999; Biggs & Kilroy 2000). Despite the burgeoning application of algae to assessing aquatic ecosystem integrity, there have been few investigations of laboratory sub-sampling error (see Alverson *et al.* 2003) or alternative methodology. Although sampling efficiency is essential to a large-scale high-throughput monitoring program, there have been no

20

attempts to evaluate alternative laboratory sub-sampling approaches for increased efficiency. An assessment of the potential contribution of measurement error to correlation strengths is necessary to gaining a complete understanding of the relationships between algal communities and the environment.

Currently, there is one basic template for the enumeration and biomass estimation of algal taxa for community composition in biomonitoring programs (Porter *et al.* 1993; Lowe and Pan 1996; Gerritsen *et al.* 1998; Barbour *et al.* 1999; Biggs & Kilroy 2000; Hill *et al.* 2001). Preserved algae, which could be phytoplankton, or periphyton scraped from an erosional substrate or siphoned from a depositional substrate, or periphyton scraped from colonized artificial substrates, are blended with the liquid they were stored in and a precise amount placed onto a counting chamber suitable for a high magnification compound microscope (e.g. Palmer-Maloney chamber). Cells are then counted and identified to a fixed-count of 300 cells either in random fields or in strips placed along the long axis of a counting chamber. High-density samples usually are counted with random fields to ensure that the 300-cell target is not surpassed. Algal community ecology investigations also use this basic template but with a wider range in cell-count targets (Alverson *et al.* 2003)

Because algae vary widely in size, some measurement of average biomass per cell of each taxon must be made to more accurately reflect dominance and energy transfer within the assemblage. The approach taken by most biomonitoring studies is to scan the sample after enumeration is complete to record measurements on 5 –15 cells per taxa identified in the enumeration for use in biovolume estimation. The target time allotment for completing the entire laboratory sub-sampling effort on a sample is 1-2 hours, inclusive of the measurement of biovolume (Porter *et al.* 1993; Biggs & Kilroy 2000).

Line-intercept sampling (LIS) is one of the predominant sampling strategies for terrestrial plant ecology (Knapp 1984; Thompson 1992), with a long history of use (e.g. Canfield 1941). It may offer substantial benefits in sampling efficiency and accuracy compared to the plot-based

21

approach of random microscope fields or strips. LIS samples particles in a region if they intersect a line segment, here called a 'transect.' Particles are sampled in proportion to the perimeter of their convex hull, i.e. the smallest convex shape that encompasses the particle (Kendall and Moran 1963, p. 58). This method has been reviewed and a unified theory of its use constructed by Kaiser (1983). LIS has been found to be more efficient than plot-based sampling, of which random fields and strip counts are analogous, in many studies of macroscopic items, predominantly terrestrial vegetation (as cited by Kaiser 1983: Canfield 1941; Bauer 1943; Warren & Olsen 1964; Bailey 1970). Any number of characteristics could be recorded for each cell that intersects a transect line, including taxon identity, whether the cell was alive or dead at capture, and cell measurements needed for biovolume calculation. Although Nedoma *et al.* (2001) demonstrated that LIS can be used to reliably calculate total filamentous algae length in a counting chamber, no comparison was made with other approaches to assess the relative efficiency of LIS.

In this paper, we present a comparison of the relative efficiency of three methods of estimating algal biovolume in a counting chamber: LIS, a 300-cell random fields count, and a 300-cell strip count. Next, we evaluated the impact of the number of cells measured and number of cells tallied for fields, strips, and LIS using simulations based on observed cell densities and sizes.

Methods

Periphyton Biovolume Estimation

To compare the sampling efficiency of the traditional two fixed-count methods and a line intercept sampling, we selected four streams representing the full range of discharges and canopy

22

cover available in the sampled stream population due to the large effects of those habitat

parameters on algal communities (Biggs 1996).   The four streams selected were: a high

discharge - low canopy cover stream; a medium discharge - high canopy cover stream; a low

discharge - low canopy cover stream; a low discharge -high canopy cover stream (Table 2-2-3,

see Supplementary Material).

We blended the samples in a high-speed kitchen blender for approximately 10 seconds to

break apart filamentous algae and placed 0.4 ml in a rectangular counting chamber measuring 20

mm x 50 mm and 0.4 mm deep that we designed. We selected a rectangular counting chamber

because a precise estimate of the length of line sampled is required for LIS. A round counting

chamber would introduce additional measurement error by requiring the measurement of transect

length . Dried Lugol's iodine was present in the counting chamber before addition of the 0.4 ml

sub-sample.

For each sample, three methods were applied to make separate estimates of periphyton

biovolume: a 300 fixed-count random field approach; a 300 fixed-count strip count approach; a

line-intercept approach using 7 transect lines. As in the usual biomonitoring approach to

estimating periphyton biovolume, high density samples were estimated using random fields and

low density samples using strip counts. We maintained this strategy so we could accurately

assess the sampling efficiency of the current protocols that call for operator interpretation of

which method to use. For the LIS approach, we randomly located transects along the short axis

of the counting chamber and ran them parallel to the long axis, a transect length of 50 mm in our

counting chamber, identifying only cells that intersected the transect. Transects were designated

on the field of view using an eyepiece micrometer. We selected 7 transect lines based on initial

estimates of density and mean cell size in the high discharge – low cover sample such that we

would expect to tally 300 cells using the LIS approach. However, in order to compare the

efficiency of all methods, we also performed the alternate fixed-count method on each sample

23

with limits of five strip counts on the high density samples and 40 random fields on the low density samples. Any more than 5 strip counts in high density samples would have required that we count thousands of cells, which we did not feel was necessary for accurate variance estimation. We counted only 40 random fields in the alternate method because attaining a 300-cell count using random fields would have required many more random fields than was necessary for variance estimation. We tallied each random field, counting strip, and transect separately for variance estimation and timed each approach.

Using the key by Prescott (1978), we identified soft algae to genus and diatoms to morphometric type (naviculoid, other pennate, cymbelloid, or centric; Porter *et al.* 1983; Lowe & Pan 1996) where genus could not be conclusively identified without acid treatment. We used the equations of Hillebrand *et al.* (1999) to calculate biovolume for all taxa tallied in each method. In the strip count and random field approaches, the biovolume of all encountered taxa was estimated by scanning the sample after counting and identification was complete, consistent with the methodology of the United States Geological Survey's National Water Quality Assessment Program (Porter *et al.* 1993), the New Zealand National Institute of Water and Atmospheric Research (Biggs & Kilroy 2000), and the U.S. EPA Rapid Bioassessment Protocols (Barbour *et al.* 1999),. A maximum of 10 cells of each taxa encountered were measured, but not greater than the number of cells encountered for that taxa if less than 10 were tallied.

Because it was already necessary to take measurements of cell size and shape for biovolume estimation, we were able to calculate the perimeter of a cell's convex hull and, consequently, estimate the probability that a cell would be tallied using LIS (cf. Kaiser 1983, equations 2.5 and 3.3). Thus, when using LIS, we took the necessary measurements for biovolume and perimeter of the convex hull as each taxon was encountered, to a limit of 10 cells per taxa. For example, *Closterium* (Prescott 1978), a sickle-shaped cell, normally would be measured for biovolume using equations for a double cone requiring only width at the median and total length; we added a

24

measurement of the width of the of the total sickle including the indentation so as to calculate the perimeter of the convex hull using an equation for half the perimeter of an ellipse. The measurements and equations we used to estimate biovolume for all taxa recorded by LIS can be compared with the measurements and equations used to calculate the perimeter of the convex hull in Table 2-2-4 (see Supplmentaty Material)..

Data Analysis

We calculated the mean biovolume per mm^2 of the counting chamber for each taxon in the random field and strip count approaches for each sample by multiplying the mean cell density mm^{-2} of counting chamber for a taxon by that taxon's mean biovolume in μm^3. Suppose that a total of n fields or strips have been evaluated, and that each field or strip has area A (mm^{-2}). Let X_{ik} be the number of cells of taxon i that are tallied in the kth field or strip, let Y_{ij} be the biovolume (μm^3) of the jth measured cell of taxon i, and let J_i be the number of cells of taxon i that have been measured. Then the corresponding estimators of biovolume density are

$$V_i = \left(\frac{1}{nA}\sum_k X_{ik}\right) \bullet \left(\frac{1}{J_i}\sum_j Y_{ij}\right) = \frac{1}{A}\bar{X}_{ik} \bullet \bar{Y}_{ij}$$
$$V_{total} = \sum_i V_i$$

(1)

Each of these estimates is associated with a sampling variance, which describes the reliability of inferences about biovolume by taxon or in total. Under the assumption that homogenization of samples makes the density and size distribution of each taxon independent, the variances for mean cell count per field or transect, and mean biovolume per cell, can be propagated by the chain rule (Goodman 1960), yielding

25

$$s^2(V_i) = \left(\frac{\overline{X}_{ik}}{A}\right)^2 s^2(\overline{Y}_{ij}) + (\overline{Y}_{ij})^2 \left(\frac{1}{A}\right)^2 s^2(\overline{X}_{ik})$$

$$s^2(V_{total}) = \sum_i s^2(V_i)$$

(2)

A higher-order term in the exact variance for individual taxa is nearly always negligible and can be omitted (Goodman 1960).

LIS is an unequal probability method, sampling cells with probability proportional to the perimeter of their convex hull; computation of biovolume density differs slightly. Suppose that a total of n lines across the counting chamber have been evaluated, and that each line has length L (mm^{-2}). As before, let X_{ik} be the number of cells of taxon i that are tallied on the kth line, and let J_i be the number of cells of taxon i that have been measured. Let Z_{ij} be the biovolume (μm^3) of the jth cell of taxon i, divided by the perimeter of that cell's convex hull (μm^2). Now the corresponding estimators of biovolume density are

$$V_i = \left(\frac{\pi}{L} \overline{X}_{ik}\right) \bullet (\overline{Z}_{ij})$$

$$V_{total} = \sum_i V_i$$

(3)

and the corresponding variance estimators are

$$s^2(V_i) = \left(\frac{\pi}{L} \overline{X}_{ik}\right)^2 s^2(\overline{Z}_{ij}) + (\overline{Z}_{ij})^2 \left(\frac{\pi}{L}\right)^2 s^2(\overline{X}_{ik})$$

$$s^2(V_{total}) = \sum_i s^2(V_i)$$

(4)

Equation (3) represents the usual unconditional estimator of density for LIS (Kaiser 1983), modified to account for sub-sampling for Z_{ij}.

The efficiency of a method depends both on the time required to process a sample, and on the resulting sample variance. We calculated relative efficiency of an estimation method (i.e. fields or strips) compared to LIS as:

26

$$R.E. = \frac{CV_i^2 * t_i}{CV_{LIS}^2 * t_{LIS}}$$ (5)

where CV_i is the coefficient of variation of a fixed-count method, t_i is the time in minutes of a fixed count method, CV_{LIS} coefficient of variation of the LIS estimation for that same sample, and t_{LIS} is the time in minutes of performing the LIS for that sample. Coefficients of variation can be calculated simply as

$$CV = \sqrt{s^2(V)/V^2} .$$

Efficiency gains found for LIS could be attributed to measuring biovolume as cells are encountered, as opposed to a second scan to measure the required number of individuals of each taxon. Thus, we estimated the time to measure each cell independent of tally time in LIS by regressing the total time to count against the number of cells counted and number of cells measured for biovolume for all four LIS samples. We used the unstandardized coefficients for the number of cells measured after controlling for the number of cells counted as the time to measure each cell for biovolume as encountered. To predict the time it would take to sample the fixed count methods as if we had measured cells for biovolume as encountered, we multiplied the time to measure for biovolume as encountered per cell by the number of measured cells in a fixed-count sample and adding the time it took to tally the unmeasured cells. We then re-calculated relative efficiencies between the fixed count methods and LIS based on the new times to compare any efficiency loss or gain of measuring cells as encountered.

To assess the variance impact of counting cells *versus* measuring cells for biovolume, we simulated the efficiency of sampling the stream periphyton using a range of target cell-count and cells measured for biovolume on all three methods. To do so, we randomly sampled the sampling units associated with each method (i.e. a field, strip, or 50 mm line-intercept line) from the dataset

27

collected to compare relative efficiency above. We simulated target cell-counts ranging from 50 to 1000. Random sampling of the parent dataset was completed when adding a sampling unit resulted in the target cell-count. We also altered the number of cells measured for each taxon in each simulated sample from 2 to 50 cells. We estimated the biovolume and variance in the biovolume estimate for each simulated sample using the methodology appropriate for each sampling approach. We also calculated the time to sample using the sampling time to count cells, measure cells, and examine each sampling unit derived from the time cost regressions described above. We calculated sampling efficiency for each simulated sample, using eqn 5, in which efficiency was relative to the most efficient simulated sample across all three sampling methods for a stream.

Results

Periphyton community structure differed greatly between streams, possibly reflecting the widely differing canopy covers and summer discharges; the three sub-sampling methods generally estimated similar assemblages (Table 2-1 and Figure 2-1). All three methods yielded similar biovolume estimates within streams and the relative ranking of streams by biovolume was the same for all three methods (Table 2-1). The high discharge – low cover stream had the highest biovolume estimates, while the low discharge – high cover stream consistently yielded the lowest periphyton biovolume(Table 2-1).

Taxa richness was slightly higher under LIS compared to those estimates made using the strict 300-count limit. In the high discharge – low cover 300-count random field sample, LIS detected 3 more taxa, and in the low discharge – low cover random field sample, LIS detected 5 additional taxa. Taxa richnesses for the two 300-count strip count samples were identical to LIS

28

(Table 2-1). It does not appear that any taxa are being missed by not counting 300 cells using LIS; to the contrary, LIS appears to detect more taxa.

LIS was more efficient than either of the fixed count methods for all streams. Relative efficiency ratios ranged from 1.6486 to 5.4564. LIS performed best relative to conventional 300-cell count methods in streams with high biovolume (Table 2-1). The efficiency gains were due to a combination of lowered time to sample in LIS, possibly because less time was spent searching empty fields or strips, and to lower coefficients of variation (C.V.) (Table 2-1). C.V.s for the LIS samples were generally half that of strip counts or random fields where the target 300-cells were counted. The only exception to this pattern was the low discharge – low cover stream, which had an equivalent C.V. in all three methods. The LIS sample with the highest C.V was the stream with the lowest cell densities and total biovolume (low discharge-high cover). . Not surprisingly, this stream also displayed the lowest efficiency gains for LIS relative to the other methods in which the target 300-cells were counted, though LIS was still 2.807 times more efficient (Table 2-1).

The mean time to both tally and measure cells for biovolume on an LIS transect was 10.50 +/- 4.89 min. Multiple regression of the sampling times in LIS resulted in a time to measure independent of the time to count a cell of 0.71 min. After adjusting the sampling times for the fixed-count methods to reflect an estimated time to sample based on measuring cells as they were encountered, the relative efficiencies did not change substantially (Table 2-1). The relative efficiencies of LIS compared to all methods were still above 1.0 after adjusting the sampling times, potentially indicating that the relative efficiency gain of LIS *versus* the fixed-count methods were due primarily to LIS rather than to measuring cells for biovolume as encountered.

The relative sampling efficiency of the simulated samples varied widely (Figure 2-2). LIS was more efficient than the conventional methods for nearly all combinations of target cell-count

and cells measured for biovolume (Figure 2-2) and resulted in the most efficient combination of target cell-count and cells measured for biovolume for all streams (Table 2-2). No single optimal ratio of cells-counted to cells-measured was found (Table 2-2). In general, the least efficient combination involved counting few cells while measuring many cells, although the high discharge-low cover stream and the conventional methods on the low discharge-low cover stream did not follow that pattern (Figure 2-2). The target cell-count that resulted in the greatest efficiency was higher in assemblages with large variation in cell density. Likewise, assemblages with high variance in taxa biovolume were most efficiently sampled by measuring more cells.

Discussion

Line-intercept sampling was consistently more efficient at measuring periphyton biovolume than fixed-count random fields or strip count approaches as sampling error was often lower even with less time spent sampling. Increased efficiency of LIS compared to fixed-area sampling, in this case random fields or strip counts, has been demonstrated in many studies sampling a variety of objects (as cited in Kaiser 1983: Canfield 1941; Bauer 1943; Warren and Olsen 1964; Bailey 1970). There are a number of theoretical reasons why we would expect sampling proportional to some characteristic of the target object to be more efficient than fixed-area sampling. First, there is no need to spend time finding random fields on the microscope stage. Secondly, because algae vary widely in size and with the large cells contributing the largest amount of biovolume usually rare in number, sampling proportional to the convex perimeter of the cell results in detection of large cells more reliably than in plot-based or frequency-proportional sampling methods. In this study, taxa richness using LIS was either equal to or greater than random fields or strip counts in

30

all samples despite tallying far fewer cells (Table 2-1). Lastly, the LIS estimate of biovolume for a taxon is calculated in part using the ratio of biovolume to perimeter (eqn 2). We would expect that the biovolume to perimeter ratio (which scales as the square of the characteristic length scale of the cell) would exhibit a smaller variance over a wider range of cell sizes than mean biovolume, which increases with the third power of the characteristic length scale of a cell.

In our view, the coefficients of variation obtained for biovolume using seven transects are still high (Table 2-1). Fortunately, the seven transects we utilized for LIS took substantially less time to sample than either 300-cell count method. We would, therefore, suggest increasing the number of transects to match the suggested 1-2 hour sampling target that the national programs are planning (Porter *et al.* 1993; Barbour *et al.* 1999; Biggs & Kilroy 2000). The additional transect length should continue to decrease error in the estimates, but with no additional resource cost compared to the fixed-count approaches. Because 10.5 minutes on average were required to perform one 50 mm transect in our counting chamber, a total of 500 mm, would fit into current program budgets. The increase in relative efficiency for LIS did not take into account the additional calculations required; we assume that the available algal biovolume computer programs could be easily modified to compute biovolume using LIS eliminating the need for additional computation time.

Although the simulations did not provide a single optimal ratio of target cell-count to number of cells measured for biovolume, they nonetheless are informative for designing an efficient sampling approach under varying expected assemblage conditions. They confirm that LIS remains more efficient than the conventional methods across a wide range of combinations of target cell-counts and cells measured. The simulations also confirm the intuitive expectation that more cells should be measured if one expects large within-taxa variation in cell biovolumes. Conversely, more cells should be counted if large variation in taxa cell densities is expected. More filamentous or colonial forms could result in higher variation in cells densities because the

31

attached cells would be less likely to be randomly distributed in the sample. However, the most efficient combination may require each research or monitoring program to evaluate its own most efficient combination tailored to the characteristics of the algal community it is investigating. Algal community ecology investigations with fewer budgetary constraints than monitoring and assessment programs may downplay sampling efficiency in favor of increased accuracy; however, given equal time spent sampling, LIS should provide greater accuracy than conventional high cell-count methods.

In theory these findings should apply to sub-sampling any algal assemblages stored in a liquid medium. However, given the different dominance of life-forms in marine *versus* freshwater or benthic *versus* pelagic taxa (Graham & Wilcox 2000), that assumption should be tested. Additionally, Alverson *et al.* (2003) found that unlike for the soft algae, valves in diatom mounts are often clustered towards the center of a counting chamber. The utility of LIS for diatom mounts, while promising, also needs further examination. Perhaps most importantly for the large-scale efforts to use periphyton in long-term biological monitoring and assessment programs, we found error in total biovolume as high as 34.0% of the mean associated with the traditional 300-cell fixed-count methods (Table 2-1). Furthermore, those high coefficients of variation were not offset by increased detection of taxa (Table 2-1). The U.S. EPA has forgone use of individual taxon biovolume estimation in its recent assessment of the Mid-Atlantic Highlands area in favor of chlorophyll *a* concentration (Hill *et al.* 2000), presumably in response to seemingly low correlations between periphyton community composition and environmental parameters or potential impacts. However, that decision may be premature, as we believe that the true strength of correlation between periphyton community composition and potential impacts may not have been accurately estimated using the traditional biomonitoring sub-sampling method.

Long-term datasets produced using consistent methodology that is accurate and resource efficient is a key goal of monitoring and adaptive ecosystem-based management (Ringold *et al.* 1999).

Given that several large-scale programs are still developing their laboratory sub-sampling

procedures for algae (Lazorchak, Klemm & Peck 1998; Jameson *et al.* 1998; Gibson *et al.* 2000;

U.S. EPA 2002) continued research into the statistical sampling aspects of the measurement of

algal biovolume such as we have presented here is needed to avoid costly mistakes and increase

monitoring and assessment program efficiency.

Table 2-1. A comparison of random field and strip fixed-count methods for estimating periphyton biovolume with line-intercept sampling. R.E. refers to relative efficiency where the most efficient approach for each stream was scaled to equal 1.

Stream	Method	Biovolume $(mm^3\ m^{-2})$	C.V.	Time (min.)	Taxa*	R.E.: biovolume post counting**	R.E.: biovolume as encountered***
High discharge – low cover†	Fields	8,851	0.194	118	8	5.023	4.347
Low discharge – high cover	Fields	33	1.054	22	-	2.719	2.89
Med. discharge – high cover	Fields	300	0.348	70	-	2.305	1.339
Low discharge – low cover†	Fields	2,991	0.296	167	6	5.456	3.028
High discharge – low cover	LIS	6,122	0.086	119	11	1.0	1.0
Low discharge – high cover	LIS	138	0.321	82	6	1.0	1.0
Med. discharge – high cover	LIS	902	0.191	106	4	1.0	1.0
Low discharge – low cover	LIS	1,262	0.186	78	11	1.0	1.0
High discharge – low cover	Strips	5,636	0.074	266	-	1.649	1.387
Low discharge – high cover†	Strips	19	0.307	267	6	2.807	2.724
Med. discharge – high cover†	Strips	274	0.34	123	4	3.852	2.438
Low discharge – low cover	Strips	932	0.188	182	-	2.405	1.606

* Taxa richness was only presented for the fixed-count method that would have been chosen, based on overall cell density, by the conventional sub-sampling estimation techniques. Taxa richness was compared with LIS only for these method applications to control for the sample size effect. See methods for details.

** Relative efficiency of the fixed-count sample method to the LIS sample method where the time to measure was based on measuring for biovolume after the cells are tallied in the fixed-count method as per typical biomonitoring program protocol.

*** Relative efficiency of the fixed-count sample method to the LIS sample method where the time to measure was based on the estimated time for measuring biovolume as cells are encountered.

Table 2-2. Summary of simulations to calculate the most efficient target cell-count and cells measured for biovolume for three sub-sampling approaches. The efficiency of the most efficient combination of target cell-count and cells measured for biovolume for each stream irrespective of method was scaled to equal 1.

Stream	Estimation method	Target cell count	Cells measured for biovolume	Relative efficiency
Low discharge-high cover	Random fields	75	40	2.954
	Random strips	200	40	2.0176
	Line-intercept	50	50	1
High discharge-low cover	Random fields	1000	4	6.4957
	Random strips	800	3	3.3879
	Line-intercept	200	4	1
Low discharge-low cover	Random fields	900	5	3.5734
	Random strips	300	10	2.6543
	Line-intercept	400	50	1
Med. discharge-high cover	Random fields	75	40	2.2555
	Random strips	100	15	3.1588
	Line-intercept	450	15	1

Table 2-3. Descriptive physical habitat information for four streams used to compare the sampling efficiency of three methods for estimating periphyton biovolume: fixed-count random fields or strip counts and line-intercept sampling. Mean Φ-value refers to the average particle size on the phi scale of Krumbein (1934).

Stream identifier	Latitude	Longitude	Discharge $(m^3\ sec^{-1})$	Canopy Cover (%)	Mean Φ-value	Depth (m)	Width (m)
Low discharge – high cover	43.35558	71.31392	< 0.01	99.43	-4.36	0.03	1.94
High discharge- low cover	44.14039	71.26849	1.43	39.75	-5.47	0.13	9.32
Low discharge – low cover	42.76310	71.29010	< 0.01	4.01	4.27	0.12	0.94
Medium discharge – high cover	43.19658	71.19557	0.09	90.87	4.06	0.11	2.25

Table 2-4. Equations used to calculate biovolume and perimeter of the convex hull for all taxa recorded by the line-intercept sampling method to illustrate the utilization of biovolume measurements for calculating the perimeter of the convex hull. The equations' dimension parameters follow the conventions of Hillebrand *et al.* (1999), though generally *a* refers to length, *b* to width, and *c* to depth. Most equations are self-explanatory and follow the generally accepted formulas for calculating the perimeter of common shapes such as rectangles and circles. *Closterium*, cymbelloid diatoms, *Gyrosigma*, and naviculoid diatoms use a simplified formula for estimating the perimeter of an ellipse. See Prescott (1978) for genera references.

	Biovolume	Perimeter of the convex hull
Anabaena	$\dfrac{\pi}{6} \bullet b^2$	$\pi \bullet b$
Bulbochaete	$\dfrac{\pi}{4} \bullet b^2 a$	$2a + 2b$
Closterium	$\dfrac{\pi}{12} \bullet b^2 a$	$\dfrac{\pi}{2}\sqrt{2 \bullet \left[\left(\dfrac{a}{2}\right)^2 + (b+d)^2\right] - \dfrac{1}{2} \bullet \left(\dfrac{a}{2} - b + d\right)^2}$
Cymbelloid diatom	$\dfrac{2}{3} ac^2 \bullet \arcsin\left(\dfrac{b}{2c}\right)$	$\dfrac{\pi}{2}\sqrt{2 \bullet \left[\left(\dfrac{a}{2}\right)^2 + \left(\dfrac{b}{2}\right)^2\right] - \dfrac{1}{2} \bullet \left(\dfrac{a}{2} - \dfrac{b}{2}\right)^2}$
Eunotia	$\dfrac{\pi}{4} abc$	$2a + 2b$
Gyrosigma	$\dfrac{\pi}{4} abc$	$\pi\sqrt{2 \bullet \left[\left(\dfrac{a}{2}\right)^2 + \left(\dfrac{b}{2}\right)^2\right] - \dfrac{1}{2} \bullet \left(\dfrac{a}{2} - \dfrac{b}{2}\right)^2}$
Lyngbya	$\dfrac{\pi}{4} \bullet b^2 a$	$2a + 2b$
Mougeotia	$\dfrac{\pi}{4} \bullet b^2 a$	$2a + 2b$
Naviculoid diatom	$\dfrac{\pi}{4} abc$	$\pi\sqrt{2 \bullet \left[\left(\dfrac{a}{2}\right)^2 + \left(\dfrac{b}{2}\right)^2\right] - \dfrac{1}{2} \bullet \left(\dfrac{a}{2} - \dfrac{b}{2}\right)^2}$
Oscillatoria	$\dfrac{\pi}{4} \bullet b^2 a$	$2a + 2b$
Pennate diatom	$\dfrac{\pi}{4} abc$	$2a + 2b$
Rhizoclonium	$\dfrac{\pi}{4} \bullet b^2 a$	$2a + 2b$
Staurastrum	$\dfrac{\pi}{12} \bullet b^2 a$	$\sqrt{\left(\dfrac{b}{2}\right)^2 + a^2} + a + b$
Tabellaria	abc	$2a + 2b$

37

Table 2-4. Continued.

	Biovolume	Perimeter of the convex hull
Terpsinoe	$\dfrac{\pi}{4}abc$	$2a + 2b$
Ulothrix	$\dfrac{\pi}{4} \bullet b^2 a$	$2a + 2b$

38

Figure 2-1. Summary biovolumes for periphyton measured in four stream samples. The samples were sub-sampled using a 0.4 ml rectangular counting chamber and three different estimation methods: random fields, random strips, and line-intercept sampling.

(a) Low discharge-high cover

(b) High discharge-low cover

(c) Low discharge-low cover

(d) Medium discharge-high cover
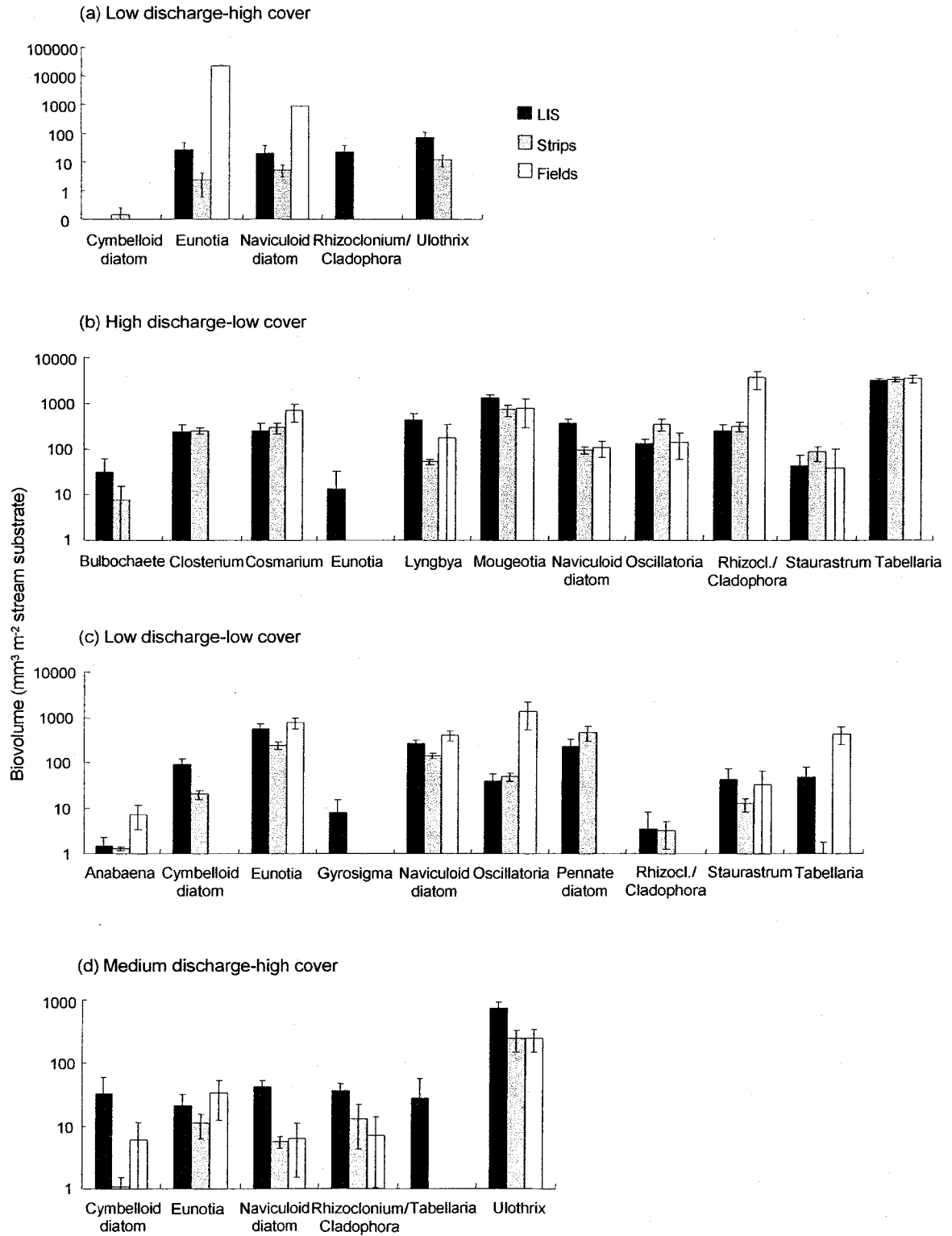
Figure 2-2. Simulated relative sub-sampling efficiency of varying target cell-counts and number of cells measured for periphyton samples from four streams using three different sub-sampling methods. Each combination was sampled from the original sampling distributions and bootstrapped 1000 times. Efficiency was calculated as relative to the most efficient combination within each stream across sub-sampling method; note that z-axes are on different scales.

(a) Low discharge – high cover

(b) High discharge – low cover
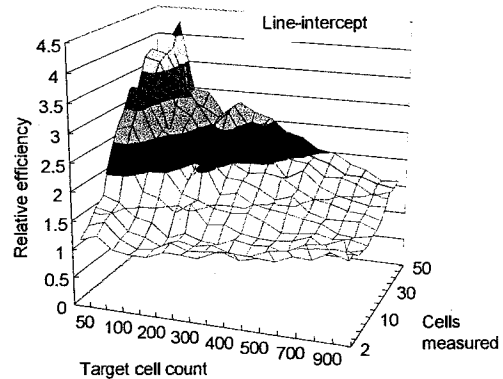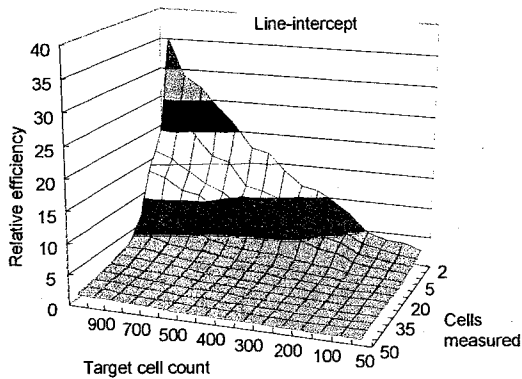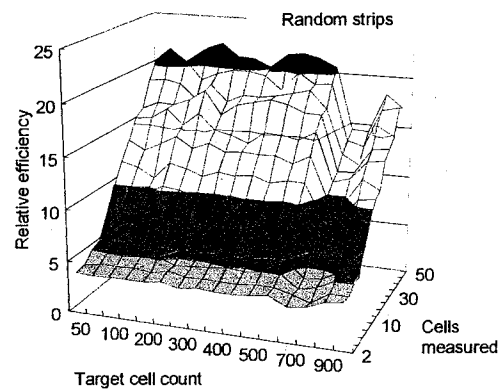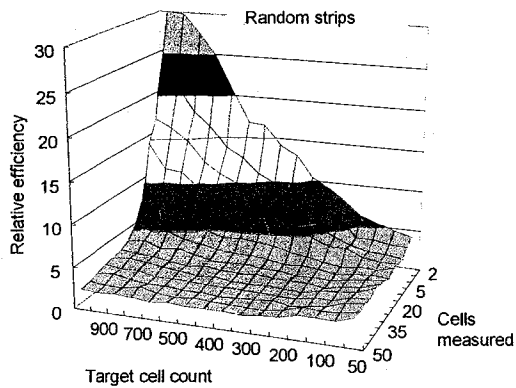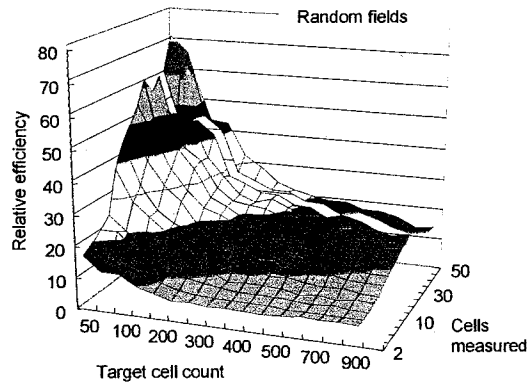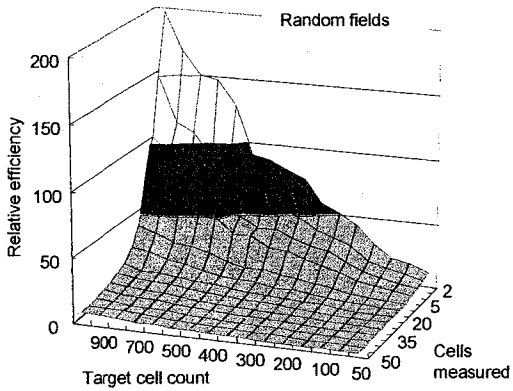
Figure 2-2 coninued.

(c) Low discharge – low cover
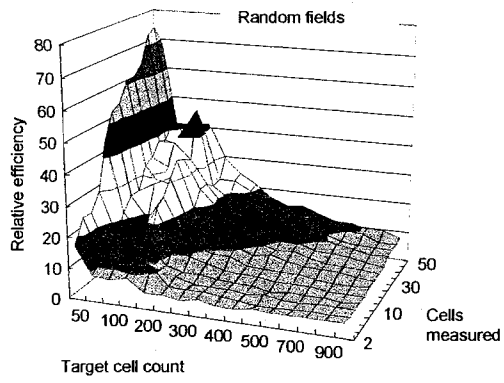
(d) Med. discharge – high cover
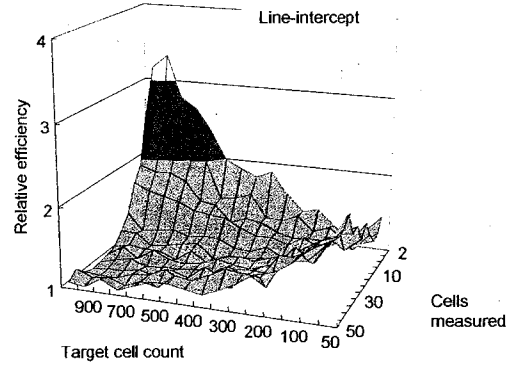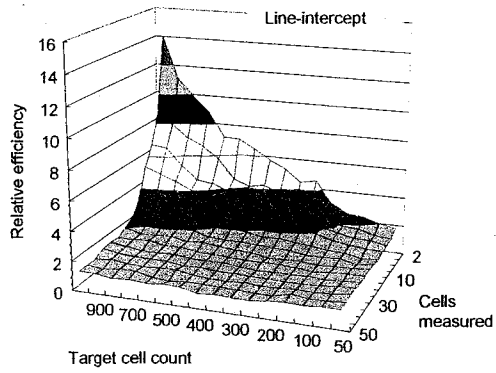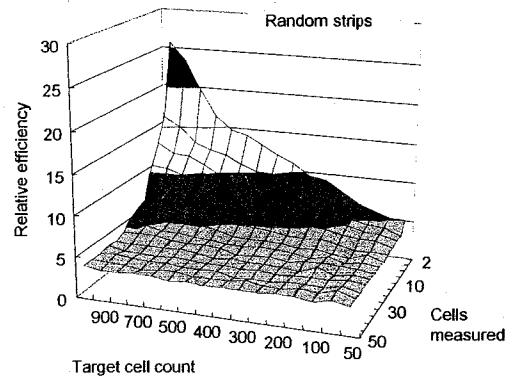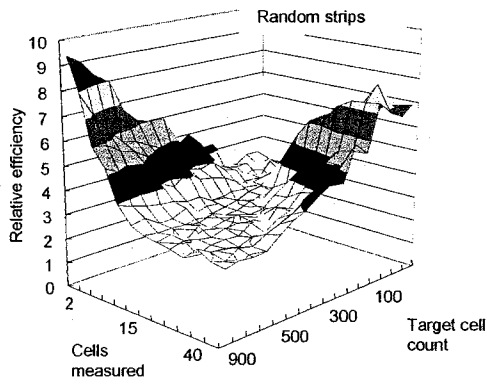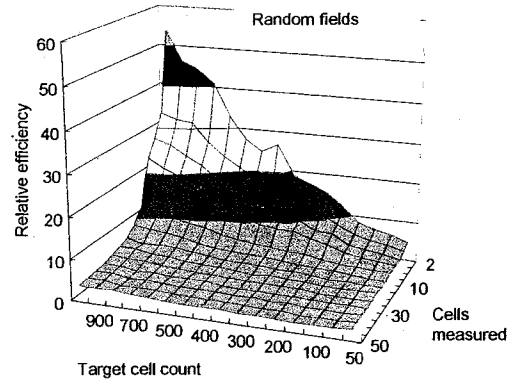
CHAPTER III

MULTI-TAXONOMIC STREAM ASSEMBLAGES: METHODS OF ANALYSIS FOR COMMUNITY CLASSIFICATION, ECOLOGICAL PATTERNS, AND A TEST OF GEOGRAPHIC CLASSIFICATIONS

<u>Summary</u>

Many of the ecosystem types in which biological assessment is being pursued contain a variety of taxonomic groups that require diverse sampling and density estimation approaches which results in very different scales of measurement. Clustering procedures, a key step in many biological assessment analyses, are greatly influenced by differing measurement scales. This study investigated the classification strengths of minimally-impacted lotic community types in New Hampshire produced using several approaches to data standardization and clustering on a dataset that incorporated vertebrate, macroinvertebrate, and periphyton assemblages. Log transformation of densities resulted in higher classification strengths than relative abundance, species maximum, or two inherently scale-independent distance measures. TWINSPAN and SPSS's Two-Stage Clustering displayed higher classification strengths than UPGMA or furthest-neighbor, without a high number of very small groups composed of single outliers. The classification produced using all of the taxonomic groups also appeared to adequately classify the taxonomic groups separately with no loss in classification strength for the macroinvertebrates and periphyton; the vertebrates were better classified using a classification produced using that taxonomic group separately. Seven community types were delineated that closely followed the

42

longitudinal stream profile based primarily on position in the watershed. As a result of the strong influence of the longitudinal stream profile, a geographic classification of New Hampshire watersheds and a classification of aquatic ecoregions poorly explained organism distributions. The sensitivity of biological assessment using the reference condition approach is greatly affected by the ability to explain natural variation in the reference sites. These results demonstrate that differing taxonomic groups can be adequately combined to produce one classification for biological assessment and natural resource management, rather than a complex set of separate classifications for each group, when appropriate statistical approaches are taken. Similar investigations are needed in other regions to test the generality of these findings.

Introduction

Biological assessment is being pursued in a variety of ecosystem types, particularly wetlands, coral reefs, streams, and lakes (Gerritsen et al. 1998, Danielson 1998, Jameson et al. 1998, Gibson et al. 2000). All of these ecosystems contain very different taxonomic groups such as macroinvertebrates, fish, amphibians, algae, plants, and bacteria. Bioassessment has been pursued because organisms integrate impacts, make detection of multiple impacts easier, and indicate the ability of an ecosystem to provide the goods and services on which humans rely. As ecosystem goods and services and organism interactions span taxonomic groups, it seems artificial, and an unnecessary complication for resource management, to construct separate community classifications for each taxonomic component. Additionally, a general goal of conservation biology is to conserve and manage all species present in an ecosystem.

Collecting organisms from multiple sampling seasons improves the classification upon which RIVPACS-style predictive bioassessment models are reliant (Clarke et al. 2002). Sampling from multiple habitat types also improves the impact sensitivity of a bioassessment

43

approach (Kerans et al. 1992) because habitats are differentially sensitive to anthropogenic impacts and contain different species (e.g. Bradley and Ormerod 2002). Similarly, taxonomic groups differ in their response to various impacts (see reviews by Davis and Simon 1995, Karr and Chu 1999). For example, Joy and Death (2002) found that a fish biotic index was not correlated to a macroinvertebrate biotic index in New Zealand, possibly because each group was responding differently to impacts. Thus, including all taxonomic groups present in an ecosystem increases the ability of a biological assessment to detect a wide range of impacts (Karr 1991, Metcalfe-Smith 1996). Additionally, differences in assemblage in one taxonomic group may propagate to other trophic levels resulting in a more discreet community classification when additional taxonomic groups are included. For example, the preferred macroinvertebrate prey species for an invertivorous fish will usually co-occur; by expanding the taxonomic groups included, a stream community that may not be distinguishable when only the fish are classified may become apparent when the macroinvertebrate prey are included.

However, different taxonomic groups often require very different sampling approaches, which results in different scales of measurement for abundance estimates. Most clustering algorithms are based on distance measures that are sensitive to the differing scales of measurement. Unless a binary presence-absence approach is taken, Taxa that are measured in larger units exert artificially increased influence on distance measures and, therefore, the classification. Precise classification of communities is important for a variety of reasons. Resource management and species conservation are simplified and enhanced by focusing on community types rather than individual species, especially for less well known taxonomic groups such as macroinvertebrates and periphyton. It is a critical step in the process of RIVPACS-style predictive modeling for biological assessment (Moss et al. 1999). A classification that explains more natural variation results in greater sensitivity of a predictive bioassessment model to small deviations in expected community structure.

A related issue is the use of geographic classifications to explain species abundances, i.e. classifications of land area or watersheds on the basis of physical landscape factors thought to influence the distribution and abundance of species (Omernik 1987, Higgins et al. 2005). Geographic classifications contrast with biotic classifications in which biological assemblages are classified into community types using either organism abundances or occurrences. The classification requirements for bioassessment and assessing representation of organisms or communities in conservation areas are similar; the classification framework should precisely explain organism distributions such that we can use the classification groups as a proxy for organism distributions and abundances (Hughes 1995, Omernik 1995). However, several studies have found that while geographic classifications such as ecoregions or stream order explain variation in organism distributions in stream ecosystems, they do not explain distributions as powerfully as biotic classifications (Marchant et al. 1999, Hawkins et al. 2000, Hawkins and Vinson 2000, Sandin and Johnson 2000, Waite et al. 2000).

This chapter examines the use of multi-taxonomic data for classifying biotic communities in minimally-impacted stream ecosystems in New Hampshire. Specifically, several clustering approaches and methods of standardizing abundance data to minimize measurement scale differences were investigated to assess which combination of standardization approach and clustering algorithm achieved the highest classification strength. The effect of including all taxonomic groups on classification strength was also investigated by comparing a classification produced using all taxonomic groups with separate classifications for each taxonomic group. Lastly, the strength of the biotic classifications was compared with geographic classifications.

Methods

Field and laboratory data were collected from minimally-impacted first to fourth order streams using the USEPA-EMAP-SW methods described in Chapter 1.

Standardization and Clustering Approaches

Taxa that occurred in fewer than three reaches were not included in any statistical analyses. Five standardization approaches for combining the broad taxonomic groups measured on different scales were investigated for their effectiveness in classification:

1. Taxa maximum (Max.): Taxa abundances were standardized to taxon maximum by dividing each taxon's abundance at a reach by the maximum abundance for that taxon across reaches.

2. Relative abundance (Rel.): The relative abundance within the three broad taxonomic groups, vertebrates, macroinvertebrates, and periphyton. Thus, a reach's total relative abundance totaled to 3.

3. Log abundance (Log): Taxa abundances were equal to $\log_{10}$ of abundance plus 1.

In addition, two distance measures that are scale independent, requiring no standardization, were investigated:

4. Mahalanobis distance (Digby and Kempton 1987).

5. Gower's General Similarity Coefficient (Gower 1971).

Taxa were classified using four techniques: TWINSPAN (Jongman et al. 1995) using PC-Ord (McCune and Mefford 1997), furthest-neighbor clustering, UPGMA clustering, and Two-Stage clustering developed by SPSS (algorithm included in SPSS version 14.0; SPSS Inc. 2005). SPSS's two stage clustering was included as it is a convenient pre-packaged algorithm for

46

clustering data measured on different scales by standardizing variables to Z-scores. For furthest-neighbor and UPGMA clustering, all five standardization approaches were used. For the three linear transformation approaches to standardization (taxa maximum, relative abundance, and log+1), Bray-Curtis (Sorenson's) percent dissimilarity was used as the distance metric for clustering (Faith et al. 1987). The distance metric standardization approaches (Gower's and Mahalnobis) obviously did not require an additional distance measure and the values were directly used in the clustering approaches. However, TWINSPAN could not be performed using the scale-independent distance measures (Mahalanobis and Gower's) as it is a weighted averaging approach that does not require a distance matrix. Reaches were classified into 2 through 12 groups/clusters for all classification approaches.

Classification approaches were compared using the LR-IND as an index of classification strength (Warton and Hudson 2004). LR-IND is defined as:

$$LR - IND = \sum_{j=1}^{p} N \cdot \log_e \left(\frac{SS_{j,0}}{SS_{j,1}} \right)$$
equation 1

where, p is the total number of taxa and $SS_{j,0}$ and $SS_{j,1}$ denote the residual sum of squares of the jth taxon under H_0 and H_1 respectively. Higher LR-IND thus indicates increased group homogeneity and difference between groups. LR-IND has been shown to be as powerful as distance-based approaches at detecting differences in organism abundances (Warton and Hudson 2004). Although this may be an unorthodox measure of classification strength, Analysis of Similarity (ANOSIM) and Multi-Repsonse Permutation Procedure (MRPP) would have biased the comparison of classification approaches as both require the choice of a distance measure. Classification approaches that resulted in very small, if not single case, groups were to be avoided on the premise that they are sensitive to outliers and offer much less information than classifications with more even groups. Thus, Pielou's J was calculated for each classification approach in which the abundances were the number of reaches within each group/cluster. LR-IND and evenness were both used to assess the standardization and classification approaches.

47

The clustering performance of a classification built using the combined taxonomic groups was compared with classifications constructed using the same clustering method but on the separate taxonomic groups. However, as there would have been no need to standardize the abundances of taxa within taxonomic groups, no standardization was applied in the separate taxonomic group classifications. LR-IND was used as the measure of classification strength and groups/clusters between 2 and 12 were produced.

The classification strength of three geographic classifications was assessed for comparison with biotic classifications:

1. Stream order: Horton-Strahler stream order.

2. Ecoregion: Level IV USEPA aquatic ecoregions (Omernik 1987).

3. HUC10 Watershed Classification: A classification of New Hampshire's USGS 10-digit HUC scale watersheds into homogenous classes produced by The Nature Conservancy. The watersheds were clustered based on physical characteristics such as elevation, landform, bedrock, and stream network characteristics. In contrast to the generalized approach of the Nature Conservancy to freshwater classification (Higgins et al. 2005, TNC-SWI 2006) the HUC10 classifications were not nested within their Ecological Drainage Units. Two levels were available: a 14-group fine-scale classification and a coarser 7-group classification.

Again, LR-IND was used as the measure of classification strength.

Ecological Patterns

To investigate basic ecological relationships among taxa as well as among taxa and physical and chemical habitat, taxa densities were ordinated using Detrended Correspondence Analysis (DCA) in PC-Ord (McCune and Mefford 1997). The correlations between the DCA axes and environmental parameters were calculated. Environmental parameters were first

48

screened to remove variables that were correlated with a Spearman's rank R greater than 0.7; one environmental variable was chosen within each set of highly co-correlated variables.

Stepwise Discriminant Function Analyses was performed on the final community classification to investigate which environmental variables best separated community types and to construct predictive models for their occurrence. Two models were constructed. One model used all non-highly correlated environmental variables to assess congruence between the ecological patterns found in the DCA ordination and provide a template for predicting community occurrence in stream reaches where a site visit is feasible. The second model was designed to assess the accuracy of mapping the community types across New Hampshire without a site visit. Thus, it was built using only large-scale variables commonly available for GIS analysis. The predictive accuracy of the two models would also inform the relative importance of local-scale habitat and watershed-scale environmental descriptors at separating community types. Both analyses used a p-to-enter of 0.5 and a p-to-remove of 1.0 based on Wilks' Lambda.

Lastly, the congruence of taxa abundances and distributions with the predictions of the trophic cascades hypothesis, which predicts negative correlations between adjacent trophic levels, was investigated. Taxa were grouped into five trophic guilds. The vertebrates were grouped into either fish or stream-dwelling amphibians. Macroinvertebrate functional feeding groups (e.g. predator, scraper) derived from Merritt and Cummings (1996) were used to group macroinvertebrate families into predator and non-predator macroinvertebrates. Periphyton biovolumes were summed for the primary producer level. Simple bivariate Spearman's correlations were calculated between the trophic groups. The significance and sign of the rank correlation coefficients were used to assess the agreement of the relationships with the trophic cascades hypothesis. Additionally, the periphyton were grouped by growth form (e.g. filamentous, colonial) nested within the phylum level (Prescott 1978) to investigate the relationships between the macroinvertebrate functional feeding groups and periphyton growth form using bivariate Spearman's correlations.

49

Matlab 7.1 was used for statistical tests and computations not attributed to another computer program.

Results

A total of 76 minimally impacted stream reaches were sampled. A total of 92 environmental descriptors were measured; 30 were removed from further analysis for the multivariate statistical analysis because they were highly correlated. The full list of environmental descriptors and those excluded can be found in Table 3-1.

Standardization and Clustering Approaches

Two-Stage clustering and TWINSPAN tended to produce classifications with the highest classification strength, as measured by high positive values of LR-IND (Table 3-2). Furthest neighbor produced slightly better classifications than UPGMA clustering, though the trend was weak (Figure 3-1). Both TWINSPAN and Two-stage clustering produced very even groups, as denoted by a high Pielou's J (Figure 3-1). Despite common suggestions that Furthest Neighbor tends to find tighter, smaller groups, it tended to produce more even group sizes than UPGMA (Figure 3-1).

Log and species maximum standardizations resulted in the best classifications compared to the other standardization approaches. There was a weak trend towards log transformation detecting more groups than species max as assessed by the number of groups where the peak in LR-IND was reached; this trend was more pronounced in the TWINSPAN and Furthest Neighbor clustering (Table 3-2). Relative abundance and Gower's dissimilarity performed intermediate to

50

the scale-independent distance measures and Log or species maximum. Mahalanobis dissimilarity measure resulted in very poor classifications (Table 3-2).

Gower's similarity measure produced classifications with very unequal group sizes (Figure 3-2). Some groups in the Gower's classifications had only one member. Species maximum, relative abundance within taxonomic group, and log standardizations displayed similar evenness values that were somewhat higher than Mahalanobis distance. Z-scores used within Two-Stage clustering produced highly even groups.

Overall, Two-Stage clustering and log-transformed TWNSPAN produced the most homogenous and even-sized groups. The combination of log transformation and TWINSPAN clustering was selected for community classification for two reasons. Their properties are better known (Jongman et al. 1995) and more widely used than SPSS's Two-Stage clustering. Logarithmic transformation, unlike species maximum, maintains the relative differences in total abundance between sites and therefore more information available for subsequent ordinations (Digby and Kempton 1987).

Comparison with Geographic Classifications

USEPA Level IV Aquatic Ecoregions resulted in the most homogenous groups with an LR-IND of 45.2. The 7-group and 14-group HUC10 classifications of the Nature Conservancy had values of 41.6 and 21.8, respectively. Biotic classifications produced with all distance measures and standardizations except Mahalanobis distance resulted in classifications that explained more variation in taxa distributions and abundances than the geographic classifications investigated (Table 3-2). For example the Log-transformed TWINSPAN 7-group classification had an LR-IND value of 99.2; because the LR-IND contains a natural log term, these differences roughly translate into a 4 to 6 times larger multivariate (MANOVA) F-ratio describing group differences than the geographic classifications (see equation 1).

51

Cross-Taxonomic Group Classification Comparisons

A general scan of LR-IND values across all cluster sizes indicated than classifications built with one taxonomic group resulted in better classifications of that same group than classifications based on other taxonomic groups. For example, a TWINSPAN classification of vertebrates resulted in generally higher LR-IND values across all numbers of clusters on the vertebrate group, and therefore more homogenous groups, than a macroinvertebrate-based TWINSPAN applied to the vertebrates (Table 3-3). A TWINSPAN classification based on all taxa included together ('all taxa') resulted in as effective a classification when applied to the macroinvertebrate and periphyton groups separately as classifications based on each taxonomic group separately. Only the vertebrate classification resulted in a better classification when applied to its respective taxonomic group compared to the all taxa combined classification (Table 3-3).

Community Descriptions

A community classification containing seven groups, produced using the log transformed abundances and TWINSPAN, was selected. Organism densities for the seven community types can be found in Table 3-3-4 and select environmental descriptors in Table 3-3-5. The first four communities were cold-water streams that, on average, were at higher elevations. These communities contained higher densities of the Siphlonuridae, Baetidae, and Rhyacophilidae. They also had higher concentrations of sulfate.

1. Very steep scour streams (n= 6), usually at high elevation, with very coarse substrate and frequent bedrock exposure. The increase in water volume at the annual flood is very

52

large relative to the size of the base-flow stream. Water temperatures are very cold with low pH and calcium. Fish are largely absent, with the vertebrates dominated by the stream dwelling amphibians *Eurycea bislineata* and *Gyrinophilus porphyriticus*. The macroinvertebrate Nemouridae, Collembola, Simuliidae, Asellidae, and Acari occur in much higher densities than other communities. This community also had greater abundance of the Uenoidae, Leuctridae, *Cladophora*, and the periphyton *Microspora*. Hydropsychidae were largely absent.

2. <u>High gradient, very cold streams (n= 21)</u>, higher average elevation with many rapids, but with less bedrock than type 1 and a smaller relative increase in flood volume. Vertebrates were almost exclusively *Salvelinus fontinalis*, though this type contained the second highest densities of *Eurycea bislineata*. The Perlodidae, Peltoperlidae, and Gomphidae occurred in relatively high densities. There were fewer Hydropsychidae than in types 3 and 4.

3. <u>Lower gradient cold-water streams (n= 17)</u>. This community is transitional between types 2 and 4. Vertebrates are dominated by *Salvelinus fontinalis*, though *Rhinicthys atratulus* can be found as well. In addition to the relatively high densities Perlodidae and Peltoperlidae, Perlidae also have higher densities. Ephemeroptera can be found in higher densities than type 2, especially the Leptophlebiidae and Leuctridae. Elmidae and Hydropsychidae reach much higher densities in this community compared to type 2.

4. <u>Very large, shallow, low gradient cold-water rivers (n = 9)</u>. Although these are the largest rivers in the classification, they are still wade-able rivers with coarse substrate. The canopy is more open due to the higher stream width. The fish assemblage was more mixed, usually with several species of cold-water fish: *Salvelinus fontinalis*, *Rhinicthys atratulus*, *Rhinicthys cataractae*, and *Cottus cognatus*. There were higher densities of

53

Leptophlebiidae, Perlidae, Corydalidae, Elmidae, Ceratopogonidae, as well as the

periphyton *Lyngbya* and *Cladophora*.

The last three communities were lower elevation, warm-water streams. They had much

higher densities of the Sphaeridae and their watersheds had greater percent cover of wetlands and

lakes. Conductivity, DOC, TDN, and DON concentrations were higher. These communities

were:

5. Warm-water, riffle streams (n = 13) with dense riparian tree cover. This community had

    higher densities of *Ictalurus nebulosus* and *Luxilus cornutus* with very few cold-water

    fish species. The Cambaridae were higher in density. The macroinvertebrate community

    was similar to the types 3 and 4, with higher densities of Hydropshychidae, Corydalidae,

    Simuliidae. However, this community was largely missing the families present only in

    the four cold-water types (Siphlonuridae and Rhyacophilidae), and had lower densities

    of the Chloroperlidae and Lepidostomatidae than the cold-water communities.

6. Sandy glide streams (n = 6). These streams tended to have a high amount of

    overhanging vegetation and more open riparian tree cover. Fish that prefer low flow

    velocity, such as species of *Esox* and *Lepomis,* dominated the vertebrates. Gammaridae

    attained far higher densities in this community. The Psephenidae, Tabanidae,

    Molannidae, and Turbellaria also had somewhat higher densities in this community as

    well. The Heptageniidae were absent.

7. Low gradient fine-muck streams (n= 4) with very high proportion of fine mucky

    substrate, pool habitat and macrophytes. More open canopy on average with fewer trees

    in riparian zone, though this was very variable. Some streams had dense forest cover

while others were completely devoid of trees in the riparian zone. The water had low Nitrate but higher Ammonia. The Asellidae were very prominent component of the macroinvertebrate assemblage; there were much higher densities of the Glossiphonidae and Dytiscidae as well. This community also had increased densities of the Hirudinidae, Phryganeidae, Sialidae, Ceratopogonidae, Libellulidae, Planorbidae, and Physidae. The diatoms (naviculoid diatoms, cymbelloid diatoms, and *Gyrosigma*) had much higher densities and there were increased densities of *Closterium, Cosmarium, Ulothrix,* and *Mougeotia. Notophthalmus viridescens* was the most common vertebrate. All Ephemeroptera densities were low compared to the other community types.

Ecological Patterns

Because log transformation of species abundances appeared to standardize species abundances best for classification, the combined taxa were log transformed for ordination to ensure that taxa measured on larger scales did not artificially influence the ordination. Three DCA axes were computed that explained 41% of the variation in taxa densities. The first axis explained 18.3% of the variance in taxa densities, only slightly more than the 11.2% of variance explained by the second axis (Figure 3-3).

The first axis was most highly correlated with substrate, elevation, and bankfull height (Figure 3-3). It represented a complex gradient of taxa preferring low-elevation slow velocity streams at one extreme and those taxa preferring high elevation streams with coarse substrate and rapids at the other. Taxa were distributed along this substrate/stream gradient/elevation gradient in ways that would be expected from their known habitat preferences (Merritt and Cummings 1996). The Turbellaria, leech families, Gammaridae, and Asellidae were all associated with slow flowing streams. In addition, all of the Odonates except for Gomphidae had highest abundances

55

in finer substrate streams. In contrast, the Rhyacophilidae, Plecoptera, *Salvelinus fontinalis* and *Rhinichthys* spp. were associated with coarse substrate and high elevation. *Eurycea bislineata*, *Gyrinophilus porphyriticus*, Nemouridae, Siphlonuridae, and Collembola were found in higher densities in small, low pH streams and their optimums clustered distant to those of the fish species (Figures 3-3 and 3-4).

The second DCA axis was primarily a temperature and pH axis (Figure 3-3). All of the fish species density optima were above the overall mean pH. In contrast, all of the stream-dwelling salamanders attained highest densities below the overall mean pH. Axis three was best correlated with measures of stream size such as mean width, bankfull height, and canopy cover. The amount of visible filamentous algae cover was also associated with this axis. Not surprisingly, the periphyton were associated with higher filamentous algae cover and a more open canopy.

Some of the community types were not very distinct. The stream communities overlapped with each other in ordination space (Figures 3 and 4). Reaches within a community varied substantially in their community composition as measured by the standard deviation in taxa densities (Table 3-4). In addition, the range in weighted-average reach scores was not as broad as the range in taxa optimums (Figures 3 and 4) indicating that many taxa optimums were highly influenced by their densities in only a few sites.

Large woody debris and purported measures of fish cover habitat (e.g. brush, undercut banks) were not highly correlated with assemblage structure. Additionally, DOC was the only chemical constituent correlated with an axis above an R^2 of 0.2.

A stepwise discriminant model to predict community memberships using watershed-scale physical descriptors available for mapping community locations correctly classified 63.2% of the reaches. The model retained 6 variables (Table 3-6). The first axis explained 64% of the total variance explained by the model (Table 3-6) and was most highly correlated with elevation (R = 0.484), membership in Omernik's (1987) aquatic ecoregion 591 (R = 0.335), order (R = 0.326),

56

and latitude (R = 0.323). Aquatic ecoregion 591 is transitional between the coastal plain and the White Mountains uplands. A stepwise discriminant model to predict community memberships that included local-scale physical habitat descriptors correctly classified 86.8% of the reaches. The model retained 15 variables, 11 of which were local habitat descriptions (Table 3-7). The first axis explained 43% of the total variance explained by the model and was most highly correlated with bankfull height (R = 0.503), elevation (R = 0.373) and order (R = 0.366). Stream order, percent of watershed as lakes, and elevation were retained as important predictors in both models.

There was very little support for trophic abundance alternations predicted by the trophic cascade hypothesis (Table 3-8). There was a weak negative relationship between salamanders and fish followed by a highly significant positive relationship between fish and predator macroinvertebrates. However, in direct opposition to the predictions of the trophic cascade hypothesis, there was a strong positive relationship between predator macroinvertebrates and non-predators. The relationship between non-predator macroinvertebrates and periphyton densities also strongly deviated from the predictions of the trophic cascade hypothesis.

Discussion

Standardization and Clustering Approaches

Log transformation of species abundances appeared to result in stronger classification of multi-taxonomic group communities (Table 3-2 and Figure 3-2). Digby and Kempton (1987) argued that on theoretical grounds log transformation of species abundances was the best . standardization approach as it minimizes differences in measurement scale while, unlike species maximum or relative abundance, preserving differences in total abundance between sites. Warton

57

and Hudson (2004) also found that log transformation of taxa abundances increases the power of multivariate analyses to detect assemblage differences. The results presented in this dissertation further extend the conclusion that log transformation may be the most appropriate standardization approach for multivariate analysis to broadly mixed taxonomic assemblages that differ in measurement scale.

TWINSPAN and SPSS's Two-Stage clustering appear to achieve the greatest classification strength compared to other common methods for classification such as UPGMA and Furthest-Neighbor clustering (Table 3-2 and Figure 3-1). Additionally, those two methods tended to discriminate more community types. Moss et al. (1999) also found that TWINPSAN performed as well or better than the alternative classification approaches they studied for the classification step in the RIVPACS biological assessment predictive model. In the RIVPACS assessment of TWINSPAN's classification strength, presence-absence data was used. This study expands the classification strength of TWINSPAN to quantitative abundance data as well. TWINSPAN has been used widely, been studied in detail, and it properties are well known. Before a general recommendation for SPSS's Two-Stage clustering can be made for ecological data involving different taxonomic groups, it should be studied in greater detail.

Classifying communities using all taxonomic groups did not achieve stronger classification than separate classifications for each taxonomic group. The all taxonomic groups classification classified macroinvertebrates and periphyton as strongly as classifications using those groups separately. While the vertebrates classified themselves more strongly than the all taxonomic groups classification, the vertebrate classification did a poor job of classifying the other taxonomic groups (Table 3-3). Additionally, 10 of the 76 stream reaches did not contain vertebrates and were not classified by the vertebrate classification except as a 'no-vertebrate' group. Thus, a classification produced using all taxonomic groups is more flexible than separate classification and with only a small loss in classification strength solely for the vertebrates.

58

There was high variation in group sizes between the classifications (Figures 3-1 or 3-2). Moss et al. (1999) also found high variation in group sizes between different classification approaches on the same dataset. Although the communities delineated using TWINSPAN differed significantly on the combined taxa (MANOVA $F = 9.268$, $p < 0.001$), the groups appeared to overlap to a great degree in ordination space (Figures 3-3 and 3-4). As has been often postulated for lotic ecosystems, taxa abundances appear to change more continuously rather than discreetly over environmental gradients (Vannote et al. 1980, Gauch 1982, Minshall and Robinson 1998). These results support the contention of Linke et al. (2005) that predictive modeling for biocriteria (Reynoldson et al. 1997) might be improved by using statistical techniques that do not require a classification step.

Ecological patterns

Stream assemblages in New Hampshire appear to be shaped primarily by substrate and water velocity (Figures 3-3 and 3-4). The community types delineated in this study arrange along the typical longitudinal stream profile. The community types ranged from the steep bedrock streams found at the upper end of the longitudinal profile to the sluggish wetland channels typically, though not always, found in the lower slopes of watersheds. The other community types were intermediate along the longitudinal profile.

It was surprising that measures of fish cover and large woody debris were so poorly correlated with assemblage structure. One potential reason for this lack of influence may be that the range in large woody debris densities may have been narrow, the ecological legacy of reduced region-wide forest clearance of the 19th century (Foster et al. 2003). The range in large woody debris density present today may not be wide enough to influence community structure; streams may require higher densities of large woody debris than currently found to influence stream habitat and community structure. As forests continue to mature in New Hampshire, the range in

59

large woody debris densities found in streams may increase and its influence should be re-assessed.

Other alterations to stream ecosystems that could not be accounted for in this classification of minimally impacted stream communities were the presence of exotic species, anthropogenically increased regional acid and nitrogen deposition, and downstream dams and artificial impoundments as barriers to dispersal. Very few anadromous species were found in this survey. If dam removal becomes wide-spread, the possible re-introduction of anadromous vertebrates may alter community structure and classes. Because pH was an important correlate with stream assemblages, reductions in acid deposition may also change composition of communities in New Hampshire. Changes in any of the regional factors would require a re-examination of the minimally impacted community types in New Hampshire.

Geographic classifications

Most of the biotic classification approaches, particularly those that did not use Mahalanobis and Gower's distances, produced stronger classifications of taxa distributions than geographic classifications. These results agree with a growing body of literature that has reached the same conclusion (Marchant et al. 1999, Hawkins et al. 2000, Hawkins and Vinson 2000, Sandin and Johnson 2000, Waite et al. 2000). The Level IV aquatic ecoregions of the USEPA had slightly higher classification strength than the HUC10 watershed classification The Nature Conservancy produced for New Hampshire. Mykra et al. (2004) also found that watersheds explained less variation in assemblage attributes than other geographical classifications.

The strong influence of stream gradient on substrate and, subsequently, on community patterns may be the reason that geographic classifications of watersheds, such as the HUC10 classification examined here, did not explain variation as well as biotic classifications. Larger watersheds, like the HUC10 size, contain within them the gradient along which differences in

60

taxa distributions are most apparent in New Hampshire – the longitudinal profile. Because watersheds bound and contain the entire stream profile behind a drainage point, they cannot separate and classify along lines that will explain natural variation in taxa abundances as well as classifications that can split and classify along that gradient. Additionally, the factors watersheds are best at delineating, water chemistry (e.g. Vitousek 1977, Martin 1979), do not appear to play a large role in structuring stream assemblages in New Hampshire (Tables 3-5 and 3-6, Figures 3-3 and 3-4). Ecoregions may have had a slightly stronger classification than the HUC10 classification because they ignore watershed boundaries and were able to delineate boundaries that better followed profile changes.

In areas where biological data are scarce, physical classifications may be the only available proxy of taxonomic distributions for conservation planning or bioassessment. Higgins et al. (2005) proposed a hierarchical classification framework for aquatic resources in which aquatic ecosystems would be classified at finer scales within ecologically determined large drainages. They argue that increasingly finer-scale geographic classifications nested within higher levels may improve explanatory power. Indeed, in this study, a discriminant model accurately predicted stream reach community memberships using a combination of local and large-scale physical factors. However, the predictive model that used only large-scale factors was less accurate. Weigel et al. (2003) similarly found that reach-scale habitat was a more important influence on lotic species composition than catchment-scale factors. Clearly, local habitat exerts a strong influence on taxa distributions and any attempt to accurately predict distributions will require information on local habitat.

The ecological patterns seen in minimally-impacted New Hampshire streams suggest that a classification unit and scale sensitive to the local habitat effects of the longitudinal stream profile is required at the smaller scales described by Higgins et al. (2005) for New Hampshire and similar areas. The analyses in this paper indicate that watersheds as large as HUC10 are not a fine enough split of the ecoregion scale. Where local habitat information is just as limited as

61

biological data, classifying individual stream segments, instead of watersheds, on the basis of physical coarse-scale characteristics such as elevation, slope, aspect, and stream order may better discriminate biological variation. Very small watersheds would be functionally similar for upper parts of the longitudinal profile, but might contain too large an area to adequately classify large stream order stream segments. Ecoregions (e.g. World Wildlife Federation ecoregions or USEPA Aquatic Ecoregions) and Ecological Drainage Units (Higgins et al. 2005) might form the higher-level stratification within which to refine and constrain the stream-segment classifications or predictive models and account for evolutionary processes in unglaciated areas.

Conclusions

In this study, log transformation with TWINSPAN divisive clustering was the best approach for clustering stream communities where biological abundances are measured on widely different scales. However, there was little support for the notion that including multiple taxonomic groups improved classification strength, though a classification based on multiple taxonomic groups did not substantially reduce classification strength on the separate taxa. The suggestion that including multiple taxonomic groups into a combined bioassessment would improve impact detection still remains to be examined. In addition, geographical classifications based on HUC10 watersheds and aquatic ecoregions poorly explained organism distributions and abundances compared to biotic classifications. The pattern of factors associated with stream organism distributions and community classification were primarily factors that vary within a catchment along the longitudinal stream profile, making it unlikely that watershed-based classifications would explain ecological variation well in New Hampshire. While a stream geographic classification based on classifying individual stream segments was proposed as an alternative for areas without adequate biological data, the assumption that this approach would better explain variation in lotic organism distributions and abundances than larger scale

62

geographic classifications clearly needs to be rigorously tested using quality biological data. Lastly, the conclusions drawn in this paper were based on only one dataset and should be tempered. Warton and Hudson (2004) demonstrate that testing analysis methods on multiple datasets can alter the inferences drawn. As bioassessment datasets expand beyond streams to lakes, coral reefs, and wetlands and come to increasingly include multiple taxonomic groups, the conclusions reached in this paper should be re-examined with a greater number of datasets.

63

Table 3-1. Environmental descriptors measured for minimally impacted streams in New Hampshire. Variables in bold were excluded from multivariate analysis because they had a bivariate Spearman's R greater than 0.7 with another included variable.

Brief variable name	Description and Units
†Ecoregion4	Level IV EPA Aquatic Ecoregion
†Lat	Latitude in decimal degrees
†Long	Longitude in decimal degrees
†Order	Stream order
†Wetlands	Percent of watershed as wetlands
†Lakes	Percent of watershed as lakes
Permanent_stream	**Total length of permanent streams (m)**
Intermittent_stream	**Total length of intermittent streams (m)**
†Highest_point	Maximum elevation in watershed (m)
Distance_impound	**Distance to nearest upstream impoundment (km)**
†Distance_source	Distance to furthest point in the upstream stream network (km)
Area	**Watershed area (ha)**
†Metamor	Percent area of metamorphic bedrock
†Volcan	Percent area of volcanic bedrock
†Pluton	Percent area of plutonic bedrock
Length	**Reach length (m)**
†Elevation	M
Discharge	**m³/sec**
†Slope	Degrees
†Aspect	Degree deviation from south
Fine	**Percent of 11 macroinvertebrate and periphyton sample sites classified as "fine" substrate**
Gravel	Percent of 11 macroinvertebrate and periphyton sample sites classified as "gravel" substrate
Coarse	Percent of 11 macroinvertebrate and periphyton sample sites classified as "coarse" substrate
Macro	**Percent of 11 macroinvertebrate and periphyton sample sites classified as "macrophyte" substrate**
Pool	Percent of 11 macroinvertebrate and periphyton sample sites classified as "pool" habitat
Glide	Percent of 11 macroinvertebrate and periphyton sample sites classified as "glide" habitat
Riffle	Percent of 11 macroinvertebrate and periphyton sample sites classified as "riffle" habitat
Rapid	Percent of 11 macroinvertebrate and periphyton sample sites classified as "rapids" habitat
pH	Water pH
Temp	Water temperature (°C)
Cond	**Specific conductivity (uS)**
Embedded	Mean % embeddedness of substrate
Phi	**Mean phi value of substrate particles**
SDphi	Standard deviation of phi values for substrate particles
Fine_sub	Percent of 55 substrate measurements classed as fine

64

Table 3-1. Continued.

Brief variable name	Description and Units
Sand	Percent of 55 substrate measurements classed as sand
Fine_gravel	Percent of 55 substrate measurements classed as fine gravel
Coarse_gravel	Percent of 55 substrate measurements classed as coarse gravel
Cobble	Percent of 55 substrate measurements classed as cobble
Boulder_sub	Percent of 55 substrate measurements classed as boulder
Bedrock	Percent of 55 substrate measurements classed as bedrock
Depth	Mean water depth (m)
Max_depth	**Maximum water depth (m)**
Width	Mean channel width (m)
Cross_section	**Mean channel width x mean water depth (m^2)**
Bankfull_width	**Mean bankfull channel width (m)**
Bankfull_height	Mean bankfull channel height (m)
Bankfull_crossection	**Mean bankfull channel width x mean bankfull channel height (m2)**
Flood_mag	Annual flood magnitude indexed by the bankfull cross-section divided by wetted cross-section (unit-less)
Undercut	Mean undercut length (m)
Angle	Mean Bank angle (degrees)
Filamentous	Mean cover index of filamentous algae in the channel (1-4)
Macrophytes	Mean cover index of macrophytes in the channel (1-4)
Woody_debris	**Mean cover index of large woody debris in the channel (1-4)**
Brushwood	Mean cover index of brush in the channel (1-4)
Live_trees	Mean cover index of live trees along the bank (1-4)
Overhanging	Mean cover index of overhanging vegetation along the bank (1-4)
Undercut_banks	Mean cover index of undercut along the bank (1-4)
Boulders	Mean cover index of boulders in the channel (1-4)
Canopy_cover	Mean % canopy cover
Big_trees	Mean cover index of large trees (>0.3 m DBH) in the canopy riparian vegetation (1-4)
Small_trees	Mean cover index of small trees (<0.3 m DBH) in the canopy riparian vegetation (1-4)
Shrubs	Mean cover index of shrubs in the under-story riparian vegetation (1-4)
Tall_herbs	Mean cover index of tall herbs (>0.5 m) in the under-story riparian vegetation (1-4)
Short_herbs	Mean cover index of short herbs (<0.5 m) in the ground cover riparian vegetation (1-4)
Bare	**Mean cover index of bare ground in the under-story riparian vegetation (1-4)**
Deciduous	Percent of riparian canopy vegetation classified as "deciduous"
Coniferous	Percent of riparian canopy vegetation classified as "coniferous"
Mixed	Percent of riparian canopy vegetation classified as "evergreen"
None	Percent of riparian canopy vegetation classified as "mixed" (>20% other)
IN1.5-0.1	**Percent of riparian canopy vegetation classified as "none present"**

65

Table 3-1. Continued.

Brief variable name	Description and Units
IN1.5-0.3	**Density (per m of reach length) of large woody debris 0.1-0.3 m in diameter and 1.5 -5 m in length in the bankfull channel**
IN5-0.1	**Density (per m of reach length) of large woody debris >0.3 m in diameter and 1.5 -5 m in length in the bankfull channel**
IN5-0.3	**Density (per m of reach length) of large woody debris 0.1-0.3 m in diameter and 5-15 m in length in the bankfull channel**
IN15-0.3	**Density (per m of reach length) of large woody debris >0.3 m in diameter and 5-15 m in length in the bankfull channel**
OUT1.5-0.1	**Density (per m of reach length) of large woody debris 0.1-0.3 m in diameter and 1.5 -5 m in length above the bankfull channel**
OUT1.5-0.3	**Density (per m of reach length) of large woody debris >0.3 m in diameter and 1.5 -5 m in length above the bankfull channel**
OUT5-0.1	**Density (per m of reach length) of large woody debris 0.1-0.3 m in diameter and 5-15 m in length above the bankfull channel**
OUT5-0.3	**Density (per m of reach length) of large woody debris >0.3 m in diameter and 5-15 m in length above the bankfull channel**
Total_LWD	**Total tally of all large woody debris**
IN_LWD	Density of all large woody debris within the bankfull channel (per m of reach length)
OUT_LWD	Density of all large woody debris above the bankfull channel (per m of reach length)
Chloride	**Chloride (mg Cl/L)**
Nitrate	Nitrate (mg N/L)
Sulfate	Sulfate (mg S/L)
Sodium	Sodium (mg Na/L)
Potassium	Potassium (mg K/L)
Magnesium	**Magnesium (mg Mg/L)**
Calcium	Calcium (mg Ca/L)
Ammonia	Ammonia (ug N/L)
Phosphate	Phosphate (ug P/L)
Silica	**Silica (mg SiO$_2$/L)**
DOC	Dissolved organic Carbon (mg C/L)
TDN	**Total dissolved Nitrogen (mg N/L)**
DON	**Dissolved organic Nitrogen (mg N/L)**
Total_alkalinity	Total alkalinity (mg CaCO$_3$/L)

† Variables considered routinely available for GIS analysis to map community types.

Table 3-2. Classification strength of differing data standardization and clustering approaches on combined abundance data of vertebrates, macroinvertebrates, and periphyton from minimally impacted streams in New Hampshire. The values are the sum of the natural log sums of squares ratios by taxa (LR-IND of Warton and Hudson 2004).

Clustering Algorithm	Standardization	Number of groups										
		2	3	4	5	6	7	8	9	10	11	12
TWINSPAN	Log+1	63.9	73.3	103.6	93.7	92.9	99.2	92.0	92.1	86.6	83.7	83.8
	Taxa max.	85.6	71.8	100.1	94.3	81.5	89.4	87.1	93.5	89.3	82.3	78.1
	Rel. abund.	60.7	71.1	99.2	90.2	92.7	89.4	81.9	74.8	73.4	74.7	80.4
Two-stage Clustering	Z-score	22.9	85.2	97.0	89.0	85.5	92.6	91.3	94.9	96.6	101.2	97.6
UPGMA	Mahalanobis	-142.1	-67.2	-46.0	-34.9	-28.9	-16.9	-10.6	-6.0	-4.8	-3.5	-1.8
	Gower's	-110.8	-44.8	-27.6	0.2	19.3	29.9	42.6	49.8	53.9	60.2	57.6
	Log+1	-98.3	-71.1	36.6	26.5	51.9	41.0	51.3	56.3	57.5	62.8	65.4
	Taxa max.	-116.3	-19.3	-21.6	-13.0	40.1	31.4	46.4	44.3	60.2	60.8	57.5
	Rel. abund.	-189.8	-43.1	-13.4	32.9	27.7	43.9	37.8	37.6	35.4	39.4	38.9
Furthest Neighbor	Gower's	-110.8	-16.5	53.6	58.0	73.2	73.6	72.3	72.5	71.0	66.9	77.3
	Mahalanobis	-117.2	-49.7	-37.3	-20.7	-17.2	-12.4	-10.0	-7.8	-9.6	0.0	0.0
	Log+1	67.3	45.3	54.4	76.1	79.7	77.9	87.9	86.1	85.0	84.4	81.5
	Taxa max.	11.8	83.0	69.1	103.2	88.6	85.9	82.8	80.0	81.8	87.1	83.9
	Rel. abund.	17.0	34.4	51.6	53.0	59.0	48.0	55.9	56.8	57.5	59.3	58.6

Table 3-3. Strengths of TWINSPAN classifications built using either all taxa together ('All Taxa') or the macroinvertebrates, vertebrates, and periphyton separately. The four classifications were then applied to the other taxonomic groups (taxonomic group applied to) to assess how well classifications built using one taxonomic group could explain variation in other stream taxonomic groups. The groups used to build the four classifications were also included for comparison purposes. The values are the sum of the natural log sums of squares ratios by taxa (LR-IND of Warton and Hudson 2004). Thus, comparison among taxonomic groups tested (e.g. classification strengths on Macroinvertebrates *versus* Periphyton) is inappropriate as taxonomic groups differed in the number of taxa within them.

TWINSPAN Classification	Taxonomic group applied to:	Number of groups										
		2	3	4	5	6	7	8	9	10	11	12
All Taxa	All Taxa	63.9	73.3	103.6	93.7	92.9	99.2	92.0	92.1	86.6	83.7	83.8
Vertebrates	All Taxa	-43.9	20.1	42.2	52.1	48.8	54.8	59.6	58.3	52.9	60.1	55.8
Macroinvertebrates	All Taxa	86.6	81.1	98.1	86.2	79.3	79.3	72.9	64.5	61.7	64.6	61.5
Periphyton	All Taxa	-68.6	-47.6	1.8	5.8	12.6	10.9	11.7	19.5	23.1	15.4	13.5
All Taxa	Macroinvertebrates	55.5	50.8	69.0	65.7	63.5	63.9	60.8	60.8	58.6	51.0	49.7
Vertebrates	Macroinvertebrates	30.1	31.6	36.8	39.8	35.5	38.8	40.3	36.9	34.9	30.7	28.1
Macroinvertebrates	Macroinvertebrates	62.5	68.2	78.1	72.1	64.8	63.3	60.1	53.2	45.7	50.4	46.7
Periphyton	Macroinvertebrates	-37.0	-36.0	-14.6	-14.9	-9.3	-7.7	-6.7	-4.7	-2.8	-0.9	-1.5
All Taxa	Vertebrates	16.5	10.4	15.3	13.5	12.4	14.9	15.0	13.8	12.7	8.3	6.8
Vertebrates	Vertebrates	12.8	19.6	24.3	25.8	24.9	24.5	24.5	23.7	22.2	23.2	22.3
Macroinvertebrates	Vertebrates	17.5	12.2	15.3	11.8	13.0	12.4	11.4	9.7	13.2	-1.3	-2.5
Periphyton	Vertebrates	-15.4	-16.8	-11.2	-8.0	-7.9	-7.4	-5.9	-6.1	-1.6	-2.7	-2.8
All Taxa	Periphyton	-8.2	12.1	19.3	14.5	17.1	19.2	16.2	17.6	15.3	15.9	13.7
Vertebrates	Periphyton	-26.3	-10.2	-11.1	-7.3	-3.7	-1.0	1.0	4.1	2.8	-2.7	-3.4
Macroinvertebrates	Periphyton	6.5	0.7	4.6	2.3	1.5	3.6	1.4	1.6	2.7	-1.4	0.6
Periphyton	Periphyton	-11.4	-11.2	14.3	15.0	17.1	15.2	14.7	21.6	21.0	17.7	16.4

Table 3-4. Mean densities for all taxa by community type for minimally-impacted streams in New Hampshire. Vertebrates and Cambaridae are in number per 100 m of stream length captured in 50 min. of electro-fishing. Macroinvertebrates are in number per m^2 of stream benthos. Periphyton densities are in biovolume (mm^3/m^2).

Taxon	Very steep scour streams		Lower gradient cold-water streams		Low gradient wetland streams		Warm-water riffle streams		Sandy glide streams		High gradient cold-water streams		Large, low gradient cold-water rivers	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Eurycea bislineata	28.4	30.4	0.9	1.0	0	0	1.0	2.9	7.6	14.3	8.1	25.1	0.2	0.4
Ictalurus nebulosus	0	0	0.1	0.5	0	0	0.8	1.9	0.1	0.3	0	0	0.0	0.1
Catostomus commersoni	0	0	0	0	0	0	0.3	1.1	0	0	0	0	0.4	0.6
Rhinichthys atratulus	0	0	3.1	7.2	0	0	2.7	9.8	0	0	0	0	7.3	7.9
Semotilus corporalis	0	0	0	0	0	0	0.1	0.4	0	0	0	0	0	0
Esox niger	0	0	0	0	0	0	0.2	0.6	0.3	0.6	0	0	0	0
Esox americanus	0	0	0	0	0	0	0.1	0.2	0	0	0	0	0	0
Perca flavescens	0	0	0	0	0	0	0.1	0.2	0	0	0	0	0	0
Decapoda: Cambaridae	0	0	0.2	0.6	0	0	3.2	10.9	0.9	2.2	0	0	0.4	1.0
Salvelinus fontinalis	1.1	2.8	14.2	16.6	2.0	4.0	0.1	0.4	0	0	13.6	15.5	6.4	7.3
Cottus cognatus	0	0	1.0	3.7	0	0	0	0	0	0	0.8	2.4	6.6	9.0
Gyrinophilus porphyriticus	0.8	1.6	0.0	0.1	0	0	0	0	0	0	0.1	0.5	0.2	0.4
Lepomis macrochirus	0	0	0	0	0	0	0.1	0.4	0.3	0.6	0	0	0	0
Notophthalmus viridescens	0	0	0.0	0.2	3.2	3.9	1.0	2.1	0.3	0.8	0	0	0.1	0.4
Luxilus cornutus	0	0	0	0	0.3	0.4	1.5	4.8	0	0	0	0	0.1	0.4
Lepomis gibbosus	0	0	0	0	0	0	0.1	0.4	0	0	0	0	0	0
Phoxinus eos	0	0	0	0	0	0	0	0	0	0	0	0	0.1	0.2
Rhinichthys cataractae	0	0	0.2	0.9	0	0	0	0	0	0	0	0	2.6	5.2
Ambloplites rupestris	0	0	0	0	0	0	0	0	0.2	0.5	0	0	0	0
Salmo gairdneri	0	0	0.4	1.5	0	0	0.5	1.7	0	0	0	0	3.8	5.6
Desmognathus fuscus	0	0	0	0	0.5	1.0	0	0	0	0	0	0	0	0
Semotilus atromaculatus	0	0	0.2	0.6	0	0	0	0	0	0	0	0	0.9	2.5
Enneacanthus obesus	0	0	0	0	0	0	1.7	6.1	0	0	0	0	0	0
Anguilla rostrata	0	0	0	0	0	0	2.8	10.0	0	0	0	0	0	0
Notropis hudsonius	0	0	0	0	0	0	0	0	0	0	0	0	0.1	0.4

Table 3-4. Continued.

Taxon	Very steep scour streams		Lower gradient cold-water streams		Low gradient wetland streams		Warm-water riffle streams		Sandy glide streams		High gradient cold-water streams		Large, low gradient cold-water rivers	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Oligochaeta	26.8	39.5	20.3	26.3	44.4	37.8	36.2	23.8	12.9	11.7	12.5	19.0	26.0	27.6
Hirudinidae	0	0	0.1	0.2	5.6	11.3	0.2	0.8	0	0	0	0	0	0
Glossiphonidae	0	0	0.2	0.5	2.8	3.6	0.6	2.1	2.1	3.3	0.2	0.9	0	0
Gammaridae	1.3	3.1	2.8	6.7	25.6	36.8	52.8	120.1	547.0	358.8	0.3	1.4	0	0
Heptageniidae	13.4	19.9	20.3	24.2	7.5	15.0	20.5	36.5	0	0	37.6	47.7	29.3	26.0
Siphlonuridae	16.6	23.5	0.1	0.2	0	0	0.3	1.0	0	0	2.2	4.9	3.9	6.9
Caenidae	0	0	0	0	0	0	0.3	1.0	0	0	0	0	0	0
Baetiscidae	0	0	0	0	0	0	0	0	0	0	0.1	0.3	0.2	0.5
Baetidae	5.0	12.2	21.1	17.1	1.3	2.5	0.6	2.1	0	0	32.6	55.6	58.6	57.8
Leptophlebiidae	0.7	1.8	97.5	107.9	20.3	8.7	0.5	1.7	1.5	2.4	6.1	9.5	55.9	69.4
Unknown Ephemeroptera	0	0	0	0	0	0	0	0	0	0	2.5	7.5	0	0
Ephemerellidae	26.7	19.4	116.6	109.6	3.8	7.5	6.8	21.5	5.2	10.0	90.7	273.1	48.8	33.4
Nemouridae	42.9	56.9	4.7	10.2	0	0	0.5	1.7	0.2	0.4	9.2	22.6	0	0
Perlodidae	0.9	1.7	1.6	3.8	0	0	0	0	0	0	0.7	1.7	0	0
Perlidae	0	0	3.8	5.9	0.3	0.5	3.0	10.8	0	0	0.1	0.2	2.3	2.3
Peltoperlidae	0	0	2.8	4.4	0	0	0	0	0	0	1.1	3.1	0	0
Capniidae	0	0	0	0	0	0	0	0	0	0	0	0	0.1	0.3
Leuctridae	72.6	65.6	83.4	87.5	6.3	12.5	18.4	27.9	20.6	25.0	34.5	34.8	11.4	12.6
Rhyacophilidae	8.1	6.8	5.0	5.0	0	0	0	0	2.7	6.1	8.6	11.2	4.6	5.5
Chloroperlidae	22.4	35.2	34.2	34.7	1.3	2.5	0	0	2.9	7.0	42.8	84.2	50.5	50.9
Hydropsychidae	0	0	26.9	28.4	6.3	7.5	46.6	55.1	6.5	13.5	4.2	10.1	47.2	65.6
Hydroptilidae	0	0	1.9	4.5	0	0	2.5	8.3	0	0	1.3	4.2	9.2	18.3
Glossosomatidae	0	0	5.0	9.4	0	0	0	0	0	0	3.0	9.8	12.3	24.2
Brachycentridae	0.4	1.0	6.4	16.8	0	0	0.2	0.8	0	0	5.1	22.9	4.6	12.8
Phryganeidae	0	0	2.2	2.7	10.3	20.6	1.5	5.6	3.3	5.4	1.1	3.3	1.4	2.8
Molannidae	0	0	0	0	0	0	0.5	1.3	1.6	2.2	0	0	0.6	1.7
Leptoceridae	0	0	0.3	0.9	0	0	2.3	4.8	0	0	0	0	8.5	19.9
Philopotamidae	1.9	4.6	20.9	38.2	2.5	5.0	10.8	29.1	1.3	3.1	7.5	16.5	26.5	46.9

Table 3-4. Continued.

Taxon	Very steep scour streams		Lower gradient cold-water streams		Low gradient wetland streams		Warm-water riffle streams		Sandy glide streams		High gradient cold-water streams		Large, low gradient cold-water rivers	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Odontoceridae	1.9	4.6	3.6	5.0	2.8	5.6	15.2	29.6	3.7	8.7	3.5	6.6	5.2	7.5
Lepidostomatidae	22.7	20.3	18.9	20.6	0	0	4.1	10.4	33.8	29.5	15.2	14.0	16.4	30.6
Polycentropidae	0	0	11.6	12.3	5.6	7.2	4.0	6.7	0.2	0.4	1.1	2.4	3.9	7.6
Limnephilidae	15.8	22.4	1.9	3.0	1.3	2.5	3.5	4.9	9.2	11.6	6.6	8.0	8.2	22.3
Notonectidae	0	0	0	0	0	0	0	0	0.8	2.0	0	0	0	0
Belastomatidae	0	0	0	0	0.3	0.5	0	0	0.8	2.0	0	0	0	0
Corixidae	0	0	2.2	9.1	1.9	3.8	0.3	1.0	1.3	3.1	0	0	0	0
Sialidae	1.9	3.1	3.1	7.4	5.9	4.7	4.0	4.7	2.3	3.0	0	0	0.2	0.7
Corydalidae	0	0	5.3	7.8	1.3	2.5	12.8	19.0	0	0	0	0	7.4	6.1
Dytiscidae	8.0	10.9	0.7	2.0	26.3	33.3	6.0	17.1	6.3	7.3	3.1	6.8	0.7	1.5
Chrysolmelidae	0	0	0	0	0	0	0	0	1.0	2.0	0.1	0.2	0	0
Elmidae	1.6	3.0	51.7	37.1	8.1	11.8	32.6	35.4	9.6	10.8	6.3	16.7	29.8	25.7
Circulionidae	0.6	1.5	0	0	0	0	0	0	0	0	0.1	0.2	0	0
Haliplidae	0	0	0.6	1.9	0	0	0	0	0	0	0	0	0	0
Psephenidae	0	0	4.6	8.1	0	0	5.3	15.2	75.4	160.4	0.4	1.0	0	0
Empididae	2.6	4.6	1.3	2.6	2.5	5.0	2.0	4.5	0	0	2.7	6.3	0.8	2.5
Tipulidae	1.9	1.9	12.8	14.8	1.9	3.8	5.0	4.8	4.3	6.9	16.6	19.1	12.0	11.3
Tabanidae	0	0	3.2	7.5	0.9	1.9	1.0	2.6	5.5	5.3	0.8	2.0	0.6	1.4
Dixidae	0	0	0.4	1.1	0	0	0	0	0	0	0.5	1.5	0	0
Psychodidae	0	0	0	0	0	0	0	0	0.2	0.4	0.1	0.5	0.8	2.5
Culcidae	0	0	0.1	0.2	0	0	0	0	0	0	0	0	0	0
Simuliidae	142.4	135.6	28.4	79.6	7.5	15.0	50.7	93.9	2.3	4.1	15.9	23.7	7.0	13.3
Chironomidae	1238.7	1389.4	868.8	736.5	1065.6	629.4	881.8	877.7	698.9	650.1	534.0	528.0	1167.2	859.6
Ceratopogonidae	1.3	3.1	16.8	17.9	24.7	18.0	13.1	21.9	4.0	3.7	12.5	14.6	27.2	19.4
Blepharicidae	0	0	0.2	0.9	0	0	0	0	0	0	0.1	0.3	1.7	5.0
Libellulidae	0	0	0	0	3.1	3.8	0	0	5.8	14.3	0	0	0	0
Ancylidae	0	0	0.4	1.8	0	0	0.3	1.0	0	0	0	0	0	0
Physidae	0	0	0.1	0.2	33.8	67.5	0	0	0	0	0	0	0	0
Planorbidae	0	0	0.1	0.3	35.6	61.7	0.8	2.2	0.6	1.5	0	0	0	0

Table 3-4. Continued.

Taxon	Very steep scour streams		Lower gradient cold-water streams		Low gradient wetland streams		Warm-water riffle streams		Sandy glide streams		High gradient cold-water streams		Large, low gradient cold-water rivers	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Sphaeridae	18.8	45.9	9.9	15.5	285.3	549.8	122.8	173.4	36.1	37.2	0.2	0.9	1.4	2.2
Collembola	12.2	17.2	1.1	2.5	0	0	1.0	1.9	0	0	0.9	2.5	0.2	0.5
Cordulegastridae	0	0	2.7	3.5	0	0	2.5	5.0	3.0	3.3	0.3	0.9	3.8	9.9
Corduliidae	0	0	0.9	3.6	0	0	0.1	0.3	0	0	0	0	0	0
Gomphidae	0	0	5.4	6.6	0	0	0.9	2.3	0	0	1.7	2.9	0.6	0.9
Aeshnidae	2.5	6.1	17.5	53.8	5.0	3.5	4.4	4.7	8.1	11.3	0.1	0.7	2.4	5.0
Calopterygidae	0	0	3.1	11.0	0	0	2.1	5.1	0	0	0.1	0.2	0.4	0.9
Ptychopteridae	0	0	0	0	3.8	7.5	0	0	0.2	0.4	0	0	0	0
Turbellaria	0.4	0.9	0	0	0	0	1.4	4.2	2.3	2.6	0	0	0	0
Nematomorpha	0	0	0.1	0.5	0	0	0.8	1.5	0	0	1.1	4.4	2.1	4.9
Hydrachnida	24.2	22.0	3.3	3.0	4.7	9.4	1.8	6.2	2.0	2.5	5.4	6.2	9.2	14.5
Pyralidae	1.8	2.9	0	0	0	0	0.2	0.8	0.2	0.4	0.4	0.9	0.7	1.7
Pteronarcyidae	0	0	0.4	1.1	0	0	0	0	0	0	0.3	1.2	1.7	3.3
Nymphomyiidae	0	0	0.1	0.4	0	0	0	0	0	0	0.1	0.2	0	0
Athericidae	0.6	1.5	1.8	5.6	0	0	0	0	0	0	0.5	1.4	0.9	1.8
Strationyidae	0	0	0	0	0	0	0	0	0	0	0.5	1.7	0.8	2.5
Uenoidae	6.3	8.4	1.5	5.4	0	0	0	0	0	0	0.3	1.0	0.6	1.7
Unknown Trichoptera	2.3	2.3	12.8	18.3	2.5	5.0	2.3	8.3	0	0	2.7	8.6	6.3	14.7
Sciomyzidae	0	0	0.1	0.4	0	0	0	0	0	0	0.1	0.2	0	0
Gerridae	0	0	0	0	0	0	0	0	0	0	0.2	0.8	0	0
Unknown Other	0	0	0	0	0	0	0	0	0	0	0	0	0.5	1.4
Asellidae	1.3	3.1	0	0	633.8	1267.5	5.8	18.7	0.4	1.0	0.1	0.5	0	0
Unknown Diptera	0	0	0	0	0	0	0	0	0	0	0.1	0.4	0	0
Psychomyiidae	0	0	0.1	0.4	0	0	0	0	0	0	0	0	0	0
Ephemeridae	0	0	3.0	10.1	0	0	0	0	0	0	0.1	0.2	3.0	9.1
Lymnaeidae	0	0	0.1	0.2	0	0	0	0	0	0	0	0	0	0
Lestidae	0	0	0	0	0	0	0	0	0	0	0	0	0.2	0.5
Anabaena	0	0	0	0	0.0	0.1	0.0	0.0	0.0	0.0	0	0	0	0
Calothrix	0	0	0	0	0	0	0	0	0	0	0.6	2.9	0	0

Table 3-4. Continued.

Taxon	Very steep scour streams		Lower gradient cold-water streams		Low gradient wetland streams		Warm-water riffle streams		Sandy glide streams		High gradient cold-water streams		Large, low gradient cold-water rivers	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Gleocapsa	0	0	0	0	0.1	0.1	0	0	0	0	0	0	0	0
Lyngbya	0	0	3.0	8.6	0	0	0.0	0.0	0	0	0.4	0.9	10.2	21.1
Oscillatoria	0.3	0.5	1.2	2.6	1.1	1.5	0.0	0.1	0.3	0.6	2.3	8.8	5.3	13.8
Microcystis	0	0	0	0	5.8	11.5	0	0	0	0	0	0	0	0
Spirulina	0	0	0	0	0	0	0.1	0.3	0	0	0.0	0.1	0	0
Centroid diatom	0	0	0	0	0	0	0.3	0.7	0	0	0	0	0	0
Asterionella	0	0	0	0	2.1	4.2	0.1	0.3	0	0	0	0	0	0
Gleocapsa	0	0	0	0	0.1	0.1	0	0	0	0	0	0	0	0
Cymbelloid diatom	0.8	1.6	0.2	0.6	2.7	5.4	0.4	0.7	0.3	0.7	0	0	0.1	0.3
Eunotia	19.3	19.6	171.7	474.8	6.9	11.4	4.3	5.6	3.2	6.4	0.7	1.7	4.7	11.1
Fragillaria	0	0	1.5	5.6	0	0	0	0	0.2	0.6	0	0	0.5	1.2
Gomphonema	1.2	2.3	5.0	14.3	2.4	4.8	0	0	2.0	3.4	0.1	0.2	6.6	8.4
Melosira	0	0	0	0	11.7	23.5	0	0	0	0	0	0	0	0
Naviculoid diatom	10.1	14.2	14.8	26.0	83.5	59.4	11.4	10.0	7.1	10.0	2.0	3.7	22.4	19.8
Tabellaria	278.2	522.8	7.5	20.1	84.8	112.3	10.7	25.5	7.4	10.4	1.4	3.2	106.5	216.8
Terpsinoe	0	0	0	0	0	0	0	0	0.4	0.9	0	0	0	0
Gyrosigma	0	0	0.1	0.3	8.3	10.4	0.0	0.1	0	0	0	0	0.2	0.6
Batrachospermum	0	0	0	0	0	0	0	0	0	0	0	0	9.5	28.5
Bulbochaete	0	0	2.9	12.1	1.3	2.6	0	0	0	0	1.3	5.9	4.3	13.0
Closterium	6.9	17.0	5.5	15.3	12.8	15.8	2.0	3.2	0	0	0	0	4.2	9.8
Cosmarium	1.2	2.9	1.8	4.3	4.6	5.4	0.2	0.6	1.0	2.3	0.8	2.8	10.4	25.1
Cylindrocapsa	0	0	0	0	4.2	8.4	0	0	0	0	0	0	0	0
Cylindrocystis	0	0	0	0	0	0	0	0	0	0	0.8	3.7	6.8	20.5
Desmidium	0	0	0	0	3.7	7.4	0	0	0	0	0	0	0	0
Draparnaldia	0	0	0	0	157.2	314.5	0	0	0	0	0	0	0	0
Euastrum	0	0	0	0	2.0	3.9	1.0	3.6	0	0	0	0	0	0
Gloeocystis	0	0	0	0	0.1	0.1	0	0	0	0	0	0	0	0
Hyalotheca	0	0	0.3	1.3	0	0	0	0	0	0	0	0	0	0
Microspora	225.4	359.0	0	0	1.2	1.9	0.1	0.2	0	0	1.0	4.2	8.2	24.5

73

Table 3-4. Continued.

Taxon	Very steep scour streams		Lower gradient cold-water streams		Low gradient wetland streams		Warm-water riffle streams		Sandy glide streams		High gradient cold-water streams		Large, low gradient cold-water rivers	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean
Mougeotia	58.6	138.5	0.7	1.8	195.0	253.3	8.8	27.6	0	0	0.3	1.3	143.1	313.7
Oedigonium	0	0	0	0	5.6	11.1	0	0	0	0	0	0	0	0
Pediastrum	0	0	0	0	0	0	0.5	1.8	0	0	0	0	0	0
Plectonema	0	0	0.1	0.5	0	0	0	0	0	0	0	0	0	0
Rhizoclonium/Cladophora	2.8	2.3	0.3	0.7	0.5	1.1	1.6	3.3	1.2	1.9	2.2	4.4	4.8	6.2
Scenedesmus	0	0	0	0	0.2	0.4	0.0	0.1	0.1	0.3	0	0	0.5	1.3
Selenastrum	0	0	0	0	0	0	0	0	0	0	0	0	0.0	0.1
Arthrodesmus	0	0	0	0	7.8	15.5	0	0	0	0	0	0	0.9	2.1
Staurastrum	0	0	0	0	51.0	102.0	0	0	1.1	1.7	0	0	1.1	1.7
Ulothrix	57.3	103.5	27.3	89.4	307.6	541.4	134.3	396.0	3.6	6.0	2.0	7.0	27.2	79.8

74

Table 3-5. Select environmental descriptors for seven stream community types classified using minimally-impacted streams in New Hampshire. The environmental descriptors included in the table were mentioned in the community descriptions. See Table 3-1 for definitions.

	Very steep scour streams		Lower gradient cold-water streams		Low gradient wetland streams		Warm-water riffle streams		Sandy glide streams		High gradient cold-water streams		Very large, shallow, low gradient cold-water rivers	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Order	2.0	0.9	2.2	0.8	1.8	1.0	1.2	0.4	1.0	0.0	1.8	0.7	2.8	0.8
Wetlands	0.0	0.0	2.1	3.6	7.1	5.9	4.3	4.4	9.6	12.7	0.5	1.3	0.6	1.0
Lakes	0.5	1.0	0.8	1.0	2.9	3.9	1.9	2.1	0.7	1.1	0.0	0.0	0.3	0.5
Highest_Point	700	440	594	420	517	414	484	325	253	143	879	411	1005	426
Elevation	432	133	333	143	211	189	110	71	94	89	393	145	301	93
Width	4.72	5.25	4.13	1.66	3.14	1.34	2.54	1.11	1.48	0.88	4.07	2.99	6.63	2.60
Slope	8.62	3.66	3.76	2.54	2.00	1.59	3.76	1.84	2.20	2.65	6.01	3.55	2.47	1.27
pH	5.56	0.48	6.26	0.60	6.10	0.60	5.94	0.87	6.04	0.67	6.29	0.61	6.67	0.29
Temp	14.15	1.23	17.11	1.65	21.65	1.17	18.79	3.64	18.17	2.87	14.76	2.21	18.44	2.94
Cond	20.29	5.06	28.40	9.76	66.69	80.87	54.83	50.34	74.92	40.58	29.04	14.39	27.47	6.55
Phi	-6.85	1.60	-3.43	3.10	3.39	4.89	-2.25	3.20	2.39	3.23	-4.13	2.09	-3.69	2.02
Pool	37.9	10.6	23.0	17.6	79.6	15.5	42.2	31.8	43.8	38.4	27.3	26.6	39.1	24.2
Glide	1.5	3.7	8.0	19.0	11.4	17.2	11.3	17.9	42.5	35.4	6.2	10.1	10.1	27.1
Riffle	15.2	9.4	49.7	26.5	9.0	10.5	34.7	23.1	3.0	4.7	37.8	25.5	40.2	27.3
Rapid	45.5	8.1	17.7	18.0	0.0	0.0	4.9	10.9	12.1	29.7	29.6	21.8	10.6	11.4
Fine_sub	2.1	5.2	7.1	15.9	48.5	49.8	9.1	15.2	17.8	13.7	4.1	9.2	4.4	6.0
Sand	2.1	1.8	8.3	8.0	20.3	23.7	21.8	20.1	52.0	29.6	7.9	7.7	12.4	10.7
Fine_gravel	6.7	5.4	9.5	6.4	6.3	7.9	5.0	4.9	4.4	5.2	10.2	5.5	11.0	6.7
Coarse_gravel	18.5	6.0	27.2	14.4	4.8	7.8	14.9	10.3	7.6	13.8	26.0	8.9	26.7	13.3
Cobble	16.7	8.1	23.4	10.0	4.7	5.4	24.9	12.3	6.3	9.1	23.2	9.3	19.3	13.7
Boulder	27.0	18.2	17.9	11.9	11.3	15.5	18.4	13.2	7.0	12.7	21.5	11.8	18.9	13.0
Bedrock	25.5	20.5	1.7	6.1	0.5	1.0	1.8	4.4	0.0	0.0	3.7	10.6	5.5	9.0
Bankfull_Height	1.01	0.47	0.62	0.25	0.30	0.16	0.33	0.15	0.25	0.06	0.64	0.25	1.06	0.28
Flood_ratio	5.03	2.48	3.16	1.56	1.60	0.89	3.20	1.32	1.55	0.60	3.63	2.24	3.23	1.55
Canopy Cover	75.43	24.53	91.26	4.87	63.23	37.53	91.56	17.18	80.75	37.69	90.59	6.56	60.73	28.32

75

Table 3-6. A discriminant model to predict membership of minimally impacted stream reaches in New Hampshire into seven community types using predictors available for mapping community locations. The model correctly classified 63.2% of the stream reaches.

Predictors	_Unstandardized Coefficients for Each Function_					
	_1_	_2_	_3_	_4_	_5_	_6_
Order	0.746	0.436	-0.472	0.910	-0.375	0.030
Lakes	-33.698	20.702	25.106	24.837	-14.843	-43.781
Highest point	0.001	0.001	0.000	-0.001	0.001	-0.002
Area	0.000	0.001	0.000	-0.001	0.000	0.000
Elevation	0.002	-0.007	-0.002	0.001	-0.001	-0.003
Slope	0.097	-.114	0.138	-0.184	-0.238	0.006
Ecoregion 580	2.891	3.034	3.661	1.287	1.197	2.424
Ecoregion 581	3.042	2.109	3.756	0.707	2.302	1.445
Ecoregion 591	0.998	2.243	4.080	0.353	0.605	1.505
Constant	-4.795	-1.779	-2.973	-1.096	0.727	0.615

Communities	_Functions at Group Centroids_					
	_1_	_2_	_3_	_4_	_5_	_6_
Very steep scour streams	1.284	-1.065	0.463	-0.262	-1.083	-0.011
Lower gradient cold-water streams	0.529	0.093	-0.071	0.796	0.028	0.234
Low gradient wetland streams	-2.293	0.349	-0.149	1.078	-0.154	-0.761
Warm-water riffle streams	-2.419	0.570	1.038	-0.336	0.064	0.083
Sandy glide streams	-3.398	-1.079	-1.732	-0.411	-0.021	0.123
High gradient cold-water streams	1.426	-0.718	0.137	-0.246	0.338	-0.108
Very large, shallow, low gradient cold-water rivers	1.596	1.949	-0.773	-0.474	-0.129	-0.047
Eigenvalue	3.575	0.923	0.568	0.313	0.141	0.054
Percent of variance	64.1	16.6	10.2	5.6	2.5	1

Table 3-7. A discriminant model to predict membership of minimally impacted stream reaches in New Hampshire into seven community types using all available local- and watershed-scale predictors for local site evaluation of the probability of stream reach membership in each community type. The model correctly classified 86.8% of the stream reaches.

Predictors	Unstandardized Coefficients for Each Function					
	1	2	3	4	5	6
Order	0.796	0.239	-0.026	-0.249	0.08	-0.597
Lakes	-5.921	60.169	28.48	-9.864	25.378	3.966
Distance from source	-0.008	-0.063	-0.169	-0.195	0.026	0.261
Elevation	0.003	-0.005	-0.001	0.002	0.004	-0.002
Rapids	-2.71	-3.46	-2.514	2.231	-0.444	-1.397
Fine	0.373	5.538	-3.475	2.166	5.404	2.632
Sand	-3.366	2.02	-3.817	4.599	-1.051	1.249
Fine gravel	7.025	7.138	-8.014	-2.647	3.793	2.261
Boulder	0.174	5.885	1.427	3.036	2.891	5.334
Bedrock	7.305	7.912	5.784	8.446	1.39	2.591
Bankfull height	3.505	3.503	0.998	-0.55	-2.193	-0.083
Flood ratio	-0.398	-0.34	0.196	0.089	0.055	0.079
Brush/wood cover	-0.339	-0.475	1.636	0.035	-0.312	-0.772
Mixed canopy	-0.528	-1.648	1.346	-0.643	0.169	0.439
Log DOC	-0.614	0.282	2.883	-1.08	-0.918	-2.303
Constant	-1.835	-1.054	-3.712	-0.266	-0.809	1.13

Communities	Functions at Group Centroids					
	1	2	3	4	5	6
Very steep scour streams	2.397	0.544	1.782	2.446	-0.045	0.07
Lower gradient cold-water streams	0.502	-0.601	0.487	-0.602	0.095	-0.767
Low gradient wetland streams	-0.8	3.821	-0.728	0.009	2.278	0.008
Warm-water riffle streams	-2.57	0.397	1.292	-0.605	-0.278	0.448
Sandy glide streams	-3.341	1.109	-1.776	1.339	-0.978	-0.5
High gradient cold-water streams	0.255	-1.783	-0.747	0.143	0.307	0.333
Very large, shallow, low gradient cold-water rivers	3.154	1.922	-0.724	-0.85	-0.825	0.307
Eigenvalue	4.131	2.547	1.193	0.935	0.518	0.251
Percent of variance	43.1	26.6	12.5	9.8	5.4	2.6

Table 3-8. Spearman correlation coefficients between trophic level densities in minimally-impacted streams in New Hampshire.

	Fish	Salamanders	Predator Macroinvertebrates	Non-predator Macroinvertebrates
Salamanders	-0.191			
Predator Macroinvertebrates	0.394*	-0.055		
Non-predator Macroinvertebrates	0.196	0.297*	0.523*	
Periphyton Total	-0.096	0.115	0.294*	0.409*

*$P<0.01$ (2-tailed)

78

Figure 3-1. The classification strengths of four clustering approaches were compared to the evenness of the group sizes. UPGMA and furthest neighbor clustering were performed using five different standardization approaches. TWINSPAN clustering was performed using three standardization approaches. Two-stage clustering was performed directly on the raw organism densities. Each clustering approach was computed to each give a range in number of groups of between 2 and 12. See methods for more explanation. Classification strength was measured using the sum of the univariate F-ratios by taxon (Warton and Hudson 2004).

79
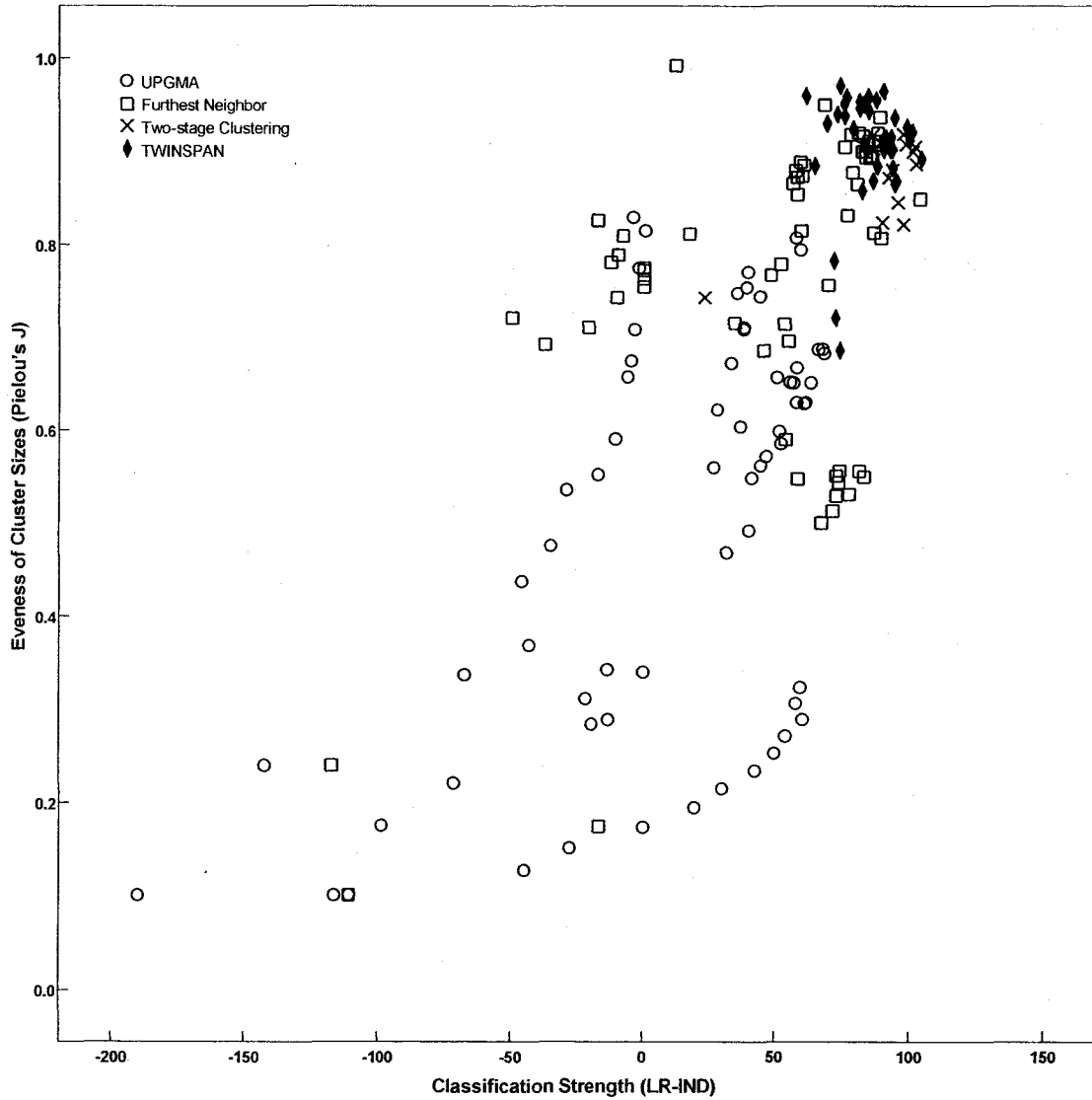
Figure 3-2. The classification strengths of six approaches to standardizing densities from multiple taxonomic groups measured on different scales were compared to the evenness of the group sizes. Four clustering approaches were used for taxa maximum, log transformation, and relative abundance by taxonomic group. Gower's and Mahalanobis each were clustered using only two clustering approaches. Two-stage clustering groups were computed directly from the raw organism densities. See methods for more explanation. Classification strength was measured using the sum of the univariate F-ratios by taxon (Warton and Hudson 2004).
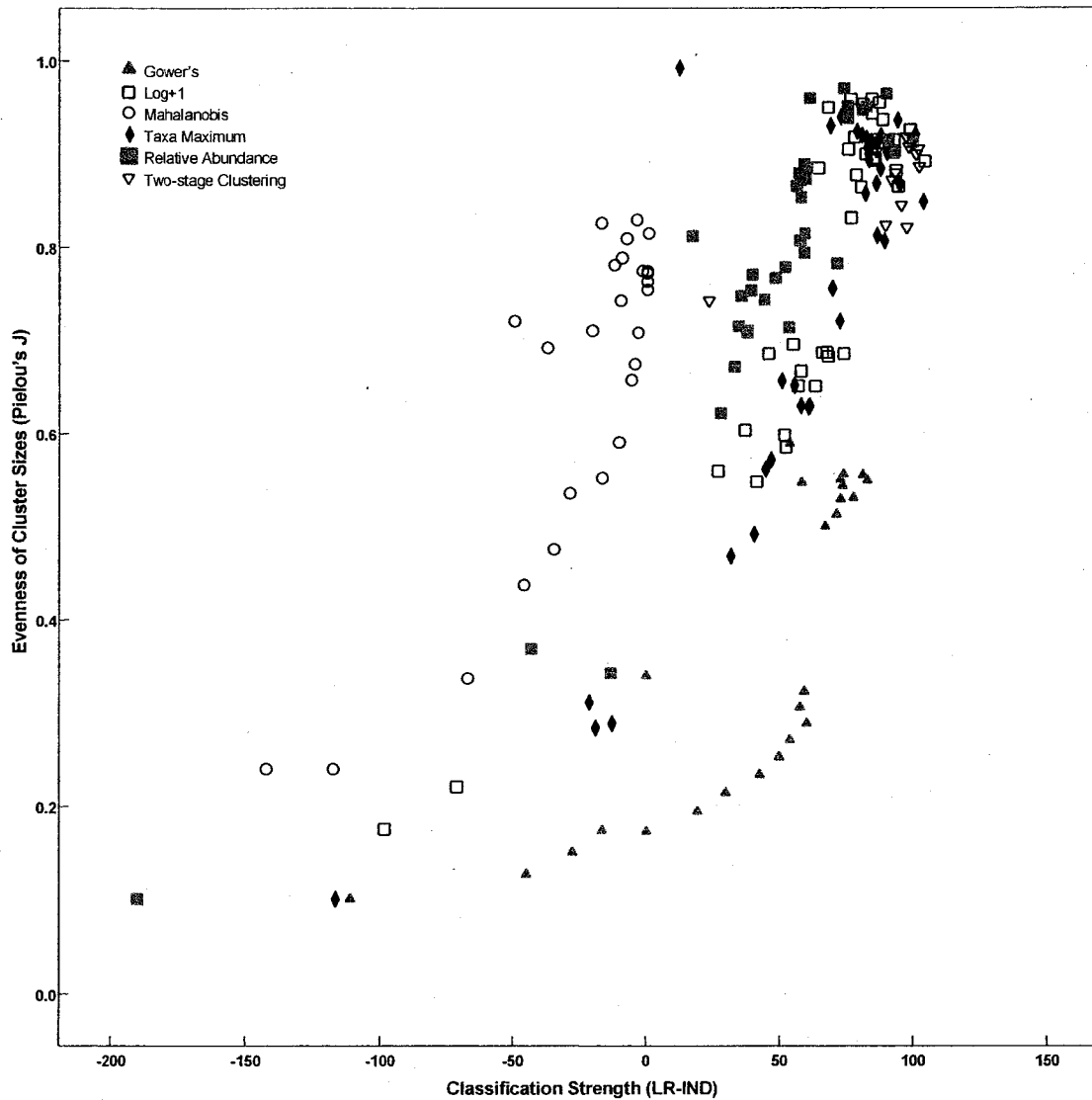
Figure 3-3. Axes 1 and 2 of a DCA ordination of minimally-impacted stream reaches in New Hampshire. Reaches are coded by the community type to which they were classified. The plotted environmental variables had correlations with the axes greater than 0.2 R^2.

Communities

△	Very steep scour streams
▲	Lower gradient cold-water streams
▽	Low gradient wetland streams
▼	Warm-water riffle streams
◇	Sandy glide streams
◆	High gradient cold-water streams
○	Very large, shallow, low gradient cold-water rivers

81
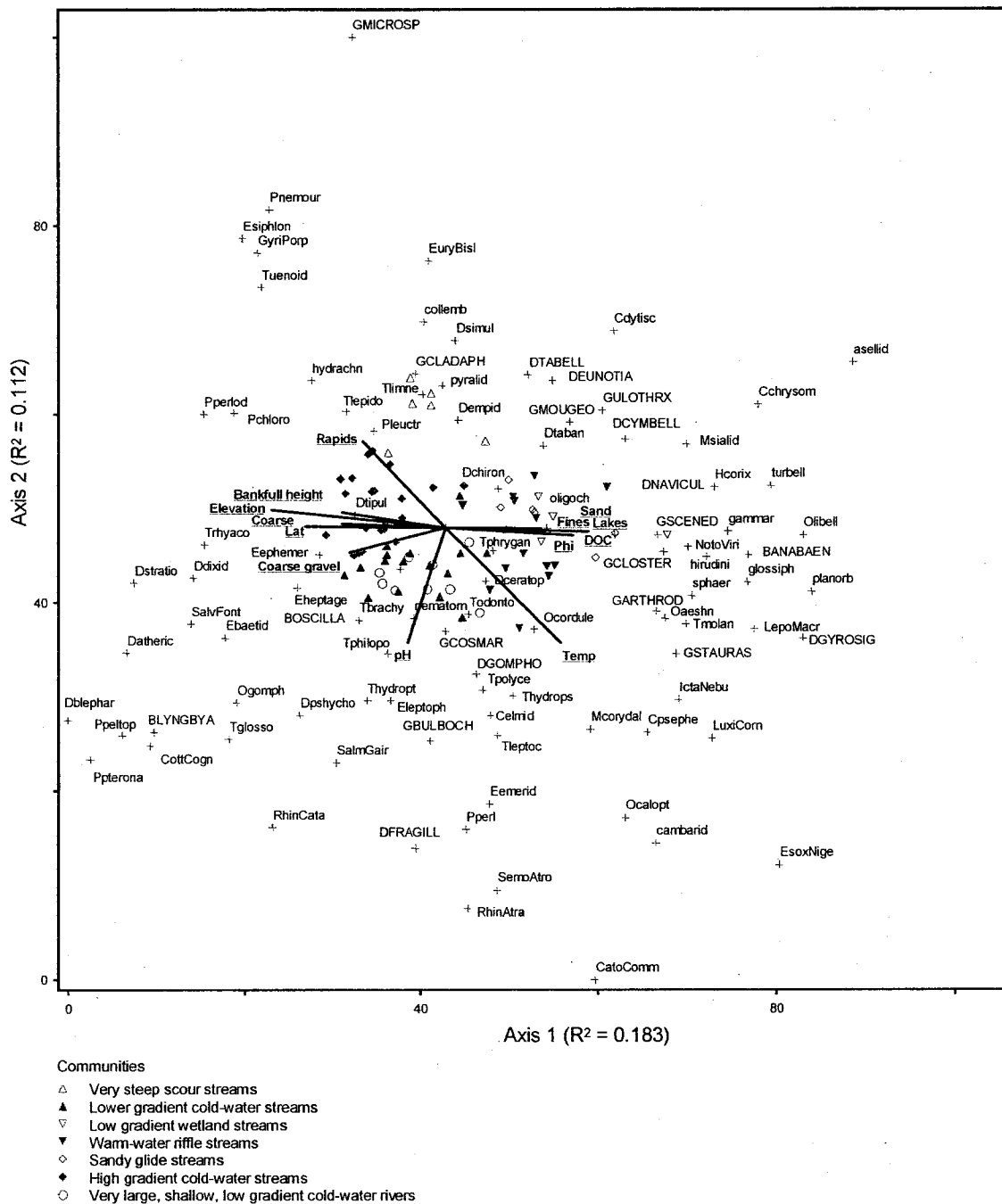
Figure 3-4. Axes 1 and 3 of a DCA ordination of minimally-impacted stream reaches in New Hampshire. Reaches are coded by the community type to which they were classified. The plotted environmental variables had correlations with the axes greater than 0.2 R^2.

Axis 3 (R^2 = 0.113)

Axis 1 (R^2 = 0.183)

Communities

△ Very steep scour streams
▲ Lower gradient cold-water streams
▽ Low gradient wetland streams
▼ Warm-water riffle streams
◇ Sandy glide streams
◆ High gradient cold-water streams
○ Very large, shallow, low gradient cold-water rivers

82
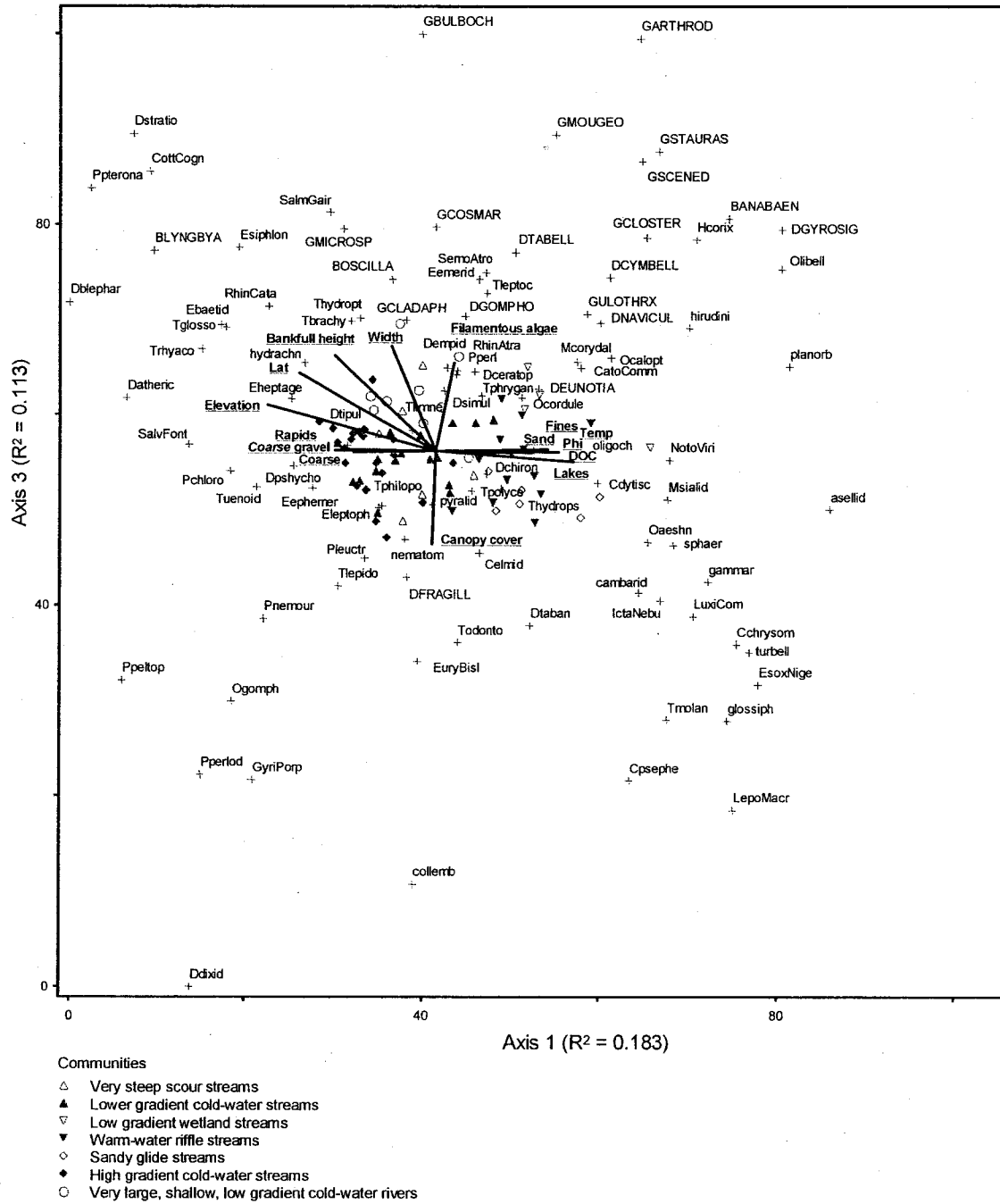
CHAPTER IV

EVALUATION OF SEVERAL METHODS TO PREDICT STREAM BIOLOGICAL REFERENCE CONDITIONS

Summary

Sensitive biological assessment to detect and measure the level of human impacts to an ecosystem requires that natural variation in community structure can be constrained. As part of a program to construct biological monitoring reference conditions for New Hampshire streams, 76 minimally-impacted streams were sampled for their fish, salamander, macroinvertebrate, and periphyton densities along with the physical habitat conditions of the streams and their watersheds. The predictive accuracies of eleven statistical analysis methods to predict macroinvertebrate, fish, and periphyton densities was investigated to identify those approaches that best constrain natural spatial variation for constructing biological assessment reference conditions. A new approach to constructing a biological assessment index that is based on the Bray-Curtis percent similarity between the observed and predicted communities was developed to allow taxa density information into the multivariate predictive assessments. Separate linear regression models to predict the densities of each taxon resulted in the most accurate predictions of expected community structure. Multivariate predictive models that included classification steps were not in general less accurate than approaches based on continuous prediction of taxon densities such as nearest-neighbor or ordination-based analyses. Including abundance information into the predictive models did not increase relative prediction error compared to an

83

AusRivAS-style assessment index based solely on predicted taxon occurrences. While some approaches were better than others, most of the multivariate prediction approaches investigated in this study differed little in predictive accuracy. And although neural networks and Bayesian approaches may improve predictive accuracy, it seems that current biological assessment sensitivity may be more limited by the intensiveness and effort involved in field and laboratory sampling than statistical analysis techniques.

Introduction

The condition of the biological component of ecosystems is a powerful indicator of the level of human influence on ecosystems (Rosenberg and Resh 1993, Karr and Chu 1999) and the ability of ecosystems to sustainably provide the goods and services on which humans rely (Karr 1999). To successfully carry out biological assessment and monitoring, the expected condition of an ecosystem and biological criteria must be set (Davis and Simons 1995). The definitions of biological diversity (Noss 1990) and biological integrity (Karr and Dudley 1981) both contain measures of function, composition, and the processes that sustain them (Davis and Simon 1995). Biological assessment and criteria are strongly linked to the community composition of minimally-impacted ecosystems (Hughes and Noss 1992). Given broad recognition of this link, the concept of reference conditions has risen to the forefront of the current conception of what constitutes appropriate biological criteria (Cairns et al. 1993, Hughes 1995). In the reference condition framework, biological criteria have been defined as expressions that describe the least impacted (reference) biological integrity of communities within a region (Cairns et al. 1993, Hughes 1995, Reynoldson et al. 1997) such that the acceptable range in biological integrity is the variation among unimpaired reference sites (Linke et al. 1999).

Hence, the central goal of biological assessment is detecting a change or difference (differences for spatial comparisons and changes for temporal) in community structure compared

84

to minimally-impacted conditions. Understanding background variation in community structure is important to determining whether observed differences or changes are due to human impacts or natural differences or changes (Philippi et al. 1998, Micheli et al. 1999, Schaefer et al. 2005). Bioassessment has most commonly been employed in stream ecosystems, often understood to be patchy and variable in space and time (Tonn 1990). The community structure of streams varies primarily along the physical gradients controlled by the longitudinal stream profile (Vannote et al. 1980). They are also affected by differences in regional climate (Winterbourn et al. 1981), tributary inflows (Bruns et al. 1984), and hydrologic conditions (Corkum 1989). This background variation must be explained before any variation in species composition detected as part of a biological assessment can be concluded to be due to human influence.

There have been two main approaches to constraining and explaining natural variation within the reference conditions for biological assessment of streams: the multimetric approach and predictive modeling (both reviewed by Reynoldson et al. 1997). In the multimetric approach, potentially disturbed streams are compared with a batch of reference minimally impacted streams chosen because they are most similar to the potentially impacted stream (Barbour et al. 1995). Usually the choice of which streams to use for the comparison is made along ecoregion lines, stream order, or both, which primarily represents catchment-level constraint of physical habitat variation. Metrics describing community composition and structure are chosen by selecting those that are most related to the presumed pollution found in non-reference sites.

A number of criticisms have been leveled at the mutltimetric approach. Local habitat factors have been found to be more important for determining lotic species composition (Weigel et al. 2003). As such, geographic classifications, e.g. ecoregions (Omernik 1987) and stream order, do not constrain biological variation as well as biotic classifications (Marchant et al. 1999, Hawkins et al. 2000, Hawkins and Vinson 2000, Sandin and Johnson 2000, Waite et al. 2000). Finally, the process of selecting metrics as a measure of biological change and human impact is inherently circular as the metrics best correlated with presumed pollution (measured by discharge

85

permits, amount of agriculture, and others) are selected over those that are not highly correlated (Reynoldson et al. 1997). As such, in the multimetric approach a given level of presumed pollution in a watershed is expected and assumed to have the same level of impact all streams. By eliminating metrics that do not correlate with the presumed pollution in the majority of streams in a region, no possibility is allowed that a metric may be indicating biological change in some streams and not others due to an usual characteristic of the streams or their assemblages. The multimetric system thus is a proxy for the easily measured presumed pollution indicators included in the metric selection process and not necessarily a measure of community change in relation to reference conditions.

The multivariate predictive modeling approach attempts to predict and model the assemblage expected for a stream using the environmental conditions of the potentially impacted test stream that are minimally influenced by humans (Reynoldson et al. 1997). There are many examples of predictive models for bioassessment (Wright et al. 1984, Reynoldson et al. 1995, Marchant et al. 1997, Chessman 1999, Hawkins et al. 2000, Linke et al. 2005, Van Sickle et al. 2005). The majority of predictive multivariate models (e.g. RIVPACS, AusRivAS, and BEAST in references above) constrain unexplained variation in species composition by classifying sites into relatively homogenous community types and then predicting which community type a test stream should belong to using environmental conditions minimally influenced by humans. By constraining more variation in assemblage attributes using biotic classifications than the geographic classification used in the multimetric approach and tailoring the biological criteria to better match the specific local habitat conditions of a site, predictive modeling should be more sensitive at detecting community alterations due to human impact. Comparisons of multimetric and predictive modeling approaches found that predictive modeling resulted in fewer misclassifications of reference sites and apparently more robust detection of impacts in potentially impacted sites (Reynoldson et al. 1997, Hawkins et al. 2000).

RIVPACS and AusRivAS procedures have been criticized because they require classification of sites into communities as an intermediate step (Chessman 1999, Linke et al. 2005). These researchers argue that classification is an artificial step in ecosystems, such as streams, where organisms display a continuum of variation and little evidence for discreet community types. Indeed, the RIVPACS procedures recognize the continuous nature of lotic organism distributions by weighting the probability an organism should be expected to occur in a test site by the probability of that test site's membership within every community and the probability of that organism's membership within every community type. This weighted probability of occurrence attempts to remove the artificial boundaries step up during classification and can thus be seen as an unnecessary complication that potentially introduces an additional source of prediction error; a direct means of predicting organism abundances directly from the reference sites might yield more accurate predictions (Chessman 1999). A nearest-neighbor approach to predicting reference organism abundances for a test site, ANNA, was developed in which the reference sites most closely matching the test site in environmental conditions were chosen for comparison with the observed test site assemblage (Linke et al. 2005). The developers compared the sensitivity and accuracy of ANNA with AusRivAS, but found no substantial improvement in correct classification of reference sites as reference or in detecting relatively well known impacts.

However, the largest criticism of the RIVPACS and AusRivAS approaches that ANNA did not respond to was that neither approach included taxa abundances in the assessment process. The indices they use for assessment are based on the presence of expected organisms; *a taxon must be locally extirpated for a change in condition to be measured. Any changes in relative abundance in response to an impact cannot be detected.* Thus, the RIVPCAS O/E approach may decrease the sensitivity of impact compared to an assessment that took abundances into account (Reynoldson *et al.* 1997, Chessman 1999). Indeed, recently Schaefer et al. (2005) found that

87

community change was detected more reliably in several datasets when abundance information was used compared to binary indices such as Jaccard's metric.

The Benthic Assessment of Sediment (BEAST) predictive bioassessment approach takes taxa abundances into account during the assessment process (Reynoldson et al. 1995). However there is still a classification step where test sites are compared with a community type it is most likely to belong to based on physical conditions. The test site is then ordinated with the reference sites from the most likely community; a test site is impaired if it is outside a pre-defined confidence interval around the reference sites. The authors of BEAST admit that there is a circularity problem as the test site is included in the ordination by which it is assessed. Additionally, there is no weighted averaging or method for allowing continuous change in organism abundances.

Chessman (1999) developed the first predictive bioassessment approach that did not involve classification, allowed for continuous change in species distributions, and assessed sites using abundance data. Chessman regressed differences in environmental conditions between pairs of reference sites against Bray-Curtis biological dissimilarity for each pair of reference sites to predict the expected dissimilarity of a test site. There are a few drawbacks to this approach. First, including differences, biological and environmental, between each pair of reference sites violates the assumption independence in regression to a large degree; X reference sites becomes $X!$ cases in the analysis. Second, if a test site differs from the expected dissimilarity, there is no explicit indication as to what is different in the test site assemblage from the expected reference conditions assemblage.

This chapter investigates several alternative methods of multivariate predictive modeling for bioassessment. The goals of all the methods developed were to include abundance data in the assessment. However, both classification and classification-less approaches were developed. The approaches were compared for the accuracy of their predictions in the reference sites, measured by percent similarity between the observed and predicted assemblages, and for the

88

precision of the prediction, measured as the standard deviation in the predicted percent similarity. The AusRivAS approach was included for comparison. In addition, the metrics usually used in the mutltimetric approach were also included and parsed by ecoregion and stream order to assess their relative variation with the multivariate predictions. These measures of accuracy and precision indicate which approach explained the greatest amount of variation in reference biological condition, a key step in determining the overall sensitivity of a bioassessment approach to detecting human influence.

Methods

The dataset of biological and physical attributes from minimally-impacted first to fourth order streams in New Hampshire (see Chapter 1) was used to compare multivariate predictive bioassessment approaches in their ability to accurately and precisely predict reference conditions in New Hampshire reference streams. Only minimally impacted streams were sampled. Any appeal to how often a bioassessment approach detects impact inherently has some circular logic involved (Reynoldson et al. 1997) as it is rarely ever truly known if a level of potential human impacts should or should not be altering organisms abundances in a particular site. Indeed, there are always two conclusions that may be drawn when a site is not found to be impacted by a bioassessment: either it was not impacted or the bioassessment approach was not sensitive enough to detect the changes in assemblage. This ambiguity limits the inferences that can be drawn from such an investigation. The only scientific basis for comparing the ability of bioassessment approaches to detect impacts would be to experimentally manipulate ecosystems with varying levels of specific impacts in a carefully controlled situation akin to estimating the minimum detection level for chemical analytic procedures (APHA 1992). What we are left is a comparison of the amount of unexplained variation in the reference conditions resulting from a

89

bioassessment approach as the most logically secure basis for comparing bioassessment approaches.

Ten alternative approaches were investigated for the ability to constrain unexplained variation in the reference stream reach assemblages:

1. Weighted Averaging MDA of Communities: This was essentially the same process as the AusRivAS (Marchant et al. 1997) approach, except that organism densities were included. Streams were classified using log transformed taxa densities and TWINSPAN. The number of communities with the largest classification strength was chosen; classification strength was measured by the sum of the natural log sums of squares ratios by taxa among the groups (LR-IND of Warton and Hudson 2004). See Chapter 3 for a complete description of the communities and classification process. A stepwise multiple discriminant analysis (MDA) was run to predict community membership using environmental variables relatively robust to human influence (Table 4-1). Each reference stream was then compared with an expected assemblage that was produced by weighting the average density for each taxon in each community by the probability of occurrence in that community. Taxa that were predicted to have a base 10 log+1 abundance below 0.3 (less than 1 predicted) were scored as zero to reduce the number of marginally predicted species occurrences in the predicted community alone.

2. MDA of Communities: This was similar to the above approach except that the predicted abundances were not weighted by probability of group membership. Each taxon density in the predicted assemblage was equal to the mean density of that taxon in the community with the highest probability of occurrence in the discriminant model. It was expected that a lack of weighted averaging would increase prediction error because it would not control for continuous organism distributions (Chessman et al. 1999, Linke et al. 2005).

90

3. Weighted Averaging CART Prediction of Communities: This was similar to approach 1 except that a CRT classification and regression tree (CART) model was used to predict the probabilities of community membership for each reference stream instead of discriminant analysis. Classification and regression trees models do not require the normal distribution assumption that discriminant analysis requires. The CRT model was pruned with a maximum difference of risk of 1 standard error.

4. CART Prediction of Communities: Similar to approach 3, except each taxon's density in the predicted assemblage was equal to the mean density of that taxon in the community with the highest probability of occurrence in the discriminant model. Just as in approach 2, it was expected that a lack of weighted averaging would increase prediction error because it would not control for continuous organism distributions (Chessman et al. 1999, Linke et al. 2005). The CRT model was pruned with a maximum difference of risk of 1 standard error.

5. Linear Regression Habitat Models: In the simplest approach to a classification-less prediction, a habitat model for each taxon was constructed using least-squares stepwise multiple linear regression. Taxon density was the dependent variable and the same environmental variables robust to human influence used in the previous 4 approaches were the independent variables (Table 4-1). Taxon densities were log-transformed to reduce skew, though little could be done to reduce the occurrence of zero densities. The predicted assemblage was the combined densities predicted by the separate regression models for all taxa.

6. CART Habitat Models: This approach was similar to approach 5 except that classification and regression tree models were used to construct habitat models for each taxon instead of least-squares linear regression. CART models do not have the distributional assumptions of least-squares regression and might be expected to produce more accurate prediction models for non-linear relationships between taxa and

91

environmental conditions. The CRT model was pruned with a maximum difference of risk of 1 standard error.

7. PCA Prediction: A principal components analysis (PCA) was used to ordinate the streams using the log transformed taxa densities. Scores on the first three axes were predicted for each site using three stepwise multiple linear least-squares regression with site scores on each axis as the dependent variables and environmental variables relatively robust to human influence as independent variables (Table 4-1). The predicted site scores were used to predict taxa densities for each reference stream.

8. DCA Gaussian prediction: This was similar to approach 7 except Detrended Correspondence Analysis (DCA) and quadratic log-linear regressions were used to approximate a Gaussian response of taxa densities to environmental conditions. Taxa densities were ordinated using DCA. To approximate a Gaussian ordination using DCA, two quadratic least-squares linear regressions were performed for each taxon with the log-transformed taxa densities as the dependent variables and site scores on each of the first two DCA axes as the independent variables in the two separate regressions (Jongman et al. 1995, p. 113). Scores for each reference reach on the first two axes were then predicted using two stepwise multiple linear least-squares regressions with site scores on each axis as the two separate dependent variables and environmental variables relatively robust to human influence as independent variables (Table 4-1). The predicted reach scores and the quadratic regression models were then used to predict taxa densities for each reference reach.

9. FA Nearest-Neighbor: Stream reaches were ordinated on the basis of environmental variables relatively robust to human influence (Table 4-1) using Factor Analysis. Environmental variables were standardized by dividing each value by the maximum value for each variable to remove measurement scale differences between the variables that could artificially influence the ordination. The nearest k neighbors were identified

92

using Euclidean distance between the site scores on all axes. The predicted community was equal to the mean densities of each taxon in the k nearest reaches. Values of k from 2 to 10 were investigated to see which produced the most accurate and precise predictions. Because there was no stepwise selection process involving the environmental variables, this approach was expected to require the largest number of environmental predictors.

10. <u>DCA Nearest-Neighbor:</u> This approach was similar to approach 9, except that reaches were ordinated on the basis of log transformed taxa densities rather than environmental variables. As such, it is the approach most similar to the ANNA approach of Linke et al. (2005) with the only differences being that reaches were ordinated using density data rather than presence-absence and DCA was used for ordination instead of non=metric multidimensional scaling (NMDS). It was thought that this approach to nearest-neighbor prediction would result in more accurate predictions than the Factor Analysis nearest-neighbor because the nearest reaches would be as similar as possible on the basis of community composition. In the Factor Analysis approach above, reaches were nearby because they had similar environmental conditions. Environmental conditions are thus only a proxy for similar community composition. However, this approach, like ANNA, required the extra step of predicting the DCA axis scores. Multiple linear regressions were used to predict reach scores on the first three axes of the DCA ordination using environmental variables relatively robust to human influence. One advantage of this is that this approach would require fewer measured environmental predictors than approach 9. The predicted scores for each reach were used to identify the k nearest-neighbors on the first three DCA axes' reach scores using Euclidean distance. The predicted community was equal to the mean densities of each taxon in the k nearest reaches. Values of k from 2 to 10 were investigated to see which produced the most accurate and precise predictions.

93

The observed reference community and the predicted community log densities were compared using Bray-Curtis percent similarity (Faith et al. 1987), hereafter referred to as O/E similarity in this paper. However, taxa that were predicted to have a base 10 log+1 abundance below 0.3 (less than 1 organism predicted) were scored as zero to reduce the number of marginally predicted species occurrences in the predicted community. Cao et al. (2002) used mean and S.D. in similarity as basis for assessing sample representativeness for biological assemblages. As assessing the ability of multivariate models to predict community composition is conceptually similar, this study also used the mean and standard deviation in the Bray-Curtis similarity between the predicted and observed assemblages as the basis for comparing the prediction approaches.

An AusRivAS-style assessment based on classification and subsequent comparison of observed and expected organism occurrences was constructed to compare the precision of the reference predictions with the multivariate approaches that included organism abundances. The simplest predictive model would be to simply identify the aquatic ecoregion (Omernik 1987) to which test reaches belong and compare the observed assemblage with the average assemblage of reference reaches within an ecoregion. To quantify the ability of such a system to predict reference conditions, the similarities between the reaches and the mean taxon densities in the ecoregions to which they belong were calculated.

In the multimetric approach, the first steps are to place reference streams into classes that explain variation in assemblages and then assess variability in component metrics in reference reaches. Part of the selection process for component metrics looks at how much variability a metric displays in the reference reaches (Barbour et al. 1992). Thus, the coefficients of variation for some common component metrics describing community structure within two common classification schemes - ecoregion and stream order nested within ecoregion - were calculated to compare the precision of two of the usual multimetric approach to defining reference conditions

94

(Barbour et al. 1995, Barbour et al. 1999) with multivariate predictive modeling. The metrics which were investigated were: vertebrate total density, macroinvertebrate total density, periphyton total density, vertebrate species richness, macroinvertebrate taxa richness, periphyton taxa richness, total richness, vertebrate diversity, vertebrate Simpson's diversity, percent Ephemeroptera/Plecoptera/Trichoptera (EPT), percent Epehemeroptera (E), percent Plecoptera (P), percent Trichoptera (T), EPT taxa, EPT taxa, E taxa, P taxa, T taxa, percent Diptera, percent Chironomidae, percent Oligochaeta, percent gatherers, percent filterers, percent scrapers, percent shredders (Barbour et al. 1999).

Prediction of taxa densities using a Canonical Correspondence Analysis approach was not investigated because there were enough environmental predictor variables available for prediction; CCA is best used when there is some doubt that most of the major explanatory variables have been measured for an assemblage (Jongman et al. 1995). Chessman's (1999) regression of community distances approach was not investigated for the reasons given in the introduction and because Linke et al. (2002) found that this approach required excessive computing power to achieve accuracies that were equivalent to AusRivAS. The ANNA nearest-neighbor approach (Linke et al. 2005) was not investigated because it did not take taxa abundances into account in the assessment process.

The classification and regression tree models were produced using SPSS version 14.0 (SPSS Inc. 2005). DCA ordinations were performed using PC-ORD (McCune and Mefford 1997). All other statistical analyses were performed in Matlab 7.1. The FATHOM toolbox of Jones (2002) was used to calculate Bray-Curtis percent similarity in Matlab.

95

Results

The multivariate approaches investigated for predicting reference conditions varied in their accuracy and precision (Table 4-2). Mean Bray-Curtis similarity between observed community composition and predicted composition (O/E similarity) ranged from 0.440 to 0.656. The coefficient of variation was used to assess precision instead of the standard deviation so that models with higher mean O/E similarities were not penalized. The precision of the predictions, as measured using the coefficient of variation in the O/E similarities, ranged from 0.095 to 0.283.

The linear regression habitat models for each taxon predicted reference conditions with the highest accuracy and precision, achieving a mean O/E similarity of 0.656 with a low coefficient of variation of 0.095 (Table 4-2). While the CART habitat models for each taxon had higher accuracy, it also displayed the second highest range in prediction accuracy (C.V. = 0.217) with very poor predictions in some cases (min. = 0.19, Table 4-2).

The PCA prediction approach had the lowest accuracy (O/E similarity of 0.440, Table 4-2). Three PCA ordination axes were used for prediction in this approach; however, the ordination required more than 16 axes to explain more than half of the variation in taxa densities. While using an approximation of Gaussian ordination improved the prediction accuracy (O/E similarity of 0.518, Table 4-2), it still displayed a much lower accuracy than the linear regressions approach and resulted the worst prediction of all the approaches (min. = 0.17). Thus, a Gaussian response model on an ordination did not substantially improve prediction ability in this dataset. The length of the gradients in DCA axes scores were low enough (Axis 1 = 2.638, Axis 2 = 1.943) that a linear model response model may just as accurately describe the taxon responses to environmental gradients (Jongman et al. 1995).

The majority of the approaches without a classification step did not predict community composition more accurately or with greater precision than procedures involving classifications (Table 4-2). The four approaches involving a classification step (Weighted Averaging MDA of

96

Communities, MDA of Communities, Weighted Averaging CART Prediction of Communities, and CART Prediction of Communities) achieved prediction accuracies and precisions above the methods that did not require discreet classifications (Table 4-2). Additionally, weighting the taxon abundances by the probability of community membership as a means of approximating continuous change in taxon densities did not improve prediction accuracy or precision at all compared to weighted-average approaches (Table 4-2).

The nearest-neighbor approaches all performed poorly (Table 4-2). The nearest-neighbor approach that used an ordination based on taxon densities (DCA nearest-neighbor) had somewhat higher prediction accuracy than the nearest-neighbor approach based on an ordination of environmental variables (FA nearest-neighbor). This improvement in accuracy was achieved primarily through better minimum predictions, with minimums of 0.41 and 0.27, respectively.

While a full description of the model details and the environmental variables involved for each approach would not fit into the constraints of this paper, some general observations can be made. The most common environmental parameters in all approaches were stream substrate descriptors (e.g. percent bedrock, percent fine, etc). Habitat type descriptors (e.g. percent rapids) and elevation were also frequent predictor variables. Although the PCA ordination approach included all the environmental variables robust to human influence, the factor loadings for the previous variable types tended to be higher.

Using aquatic ecoregion membership to predict taxa composition in the reference reaches resulted in inaccurate and imprecise predictions. The mean percent similarity between reaches and the mean ecoregion community to which they belonged was 0.296 with a standard deviation of 0.104. Variation in component metrics were not well explained by aquatic ecoregions (mean C.V. = 0.837, s = 0.439). Stream order nested within ecoregion did not explain much more variation (mean C.V. = 0.761, s = 0.457). The multimetric approach would normally select only the metrics with low variation in reference reaches. However, only two metrics, total taxa

richness and macroinvertebrate taxa richness, had average coefficients of variation below 0.3, the C.V. in observed to expected similarity attained by predictive modeling.

Discussion

The highest Bray-Curtis percent similarity between an observed minimally-impacted stream assemblage and the assemblage predicted using multivariate techniques (O/E similarity) was 0.672. A perfect prediction would yield an O/E similarity very close to 1. Thus, all of the prediction approaches for taxon densities fell short (Table 4-2). In contrast, the AusRivAS analysis predicting taxon occurrence (rather than taxon densities) displayed a very high accuracy (O/E ratio = 0.981, Table 4-2). However, the mean prediction accuracy does not completely describe prediction performance. The coefficient of variation around the mean prediction accuracy for the best performing multivariate prediction of taxon densities (linear regression habitat models) was about half of that of the AusRivAS model (O/E of 0.095 versus 0.179, Table 4-2).

This last observation is important because in any assessment approach, the range of the indicator or index in reference conditions needs to be accounted for when assessing whether there was a departure from reference conditions (Davis and Simon 1995, Barbour et al. 1995, Hughes 1995, Linke et al. 1999). Typically, some description of variance in the index value in the reference sites, such as standard deviation or inter-quartile range, is chosen beyond which test sites are considered to be different from reference conditions (Davis and Simon 1995, Wright et al. 2000). For the index investigated in this study, O/E similarity, a decrease would indicate pollution. Thus, the number of standard deviation categories between the mean O/E similarity and zero range describes the upper limit on how sensitive an index might be at detecting impacts. For example, if an index that decreases in response to pollution had a mean of 1 and a standard

deviation in the reference sites of 0.5 and two standard deviations is the assessment range, then only 1 category of impact below reference can be resolved. This analysis is similar to the one done by Doberstein et al. (2000) on measuring the impact on assessment sensitivity of variance due to sub-sampling. Clearly, a high percent similarity and low variance around that prediction are both desirable because those conditions are likely to increase sensitivity at detecting impacts.

The number of standard deviation categories below the mean predicted O/E similarity can be quantified by dividing the mean prediction accuracies for each approach by their standard deviations. On this metric, the AusRivAS binary assessment approach might discriminate a maximum of 5.6 categories (Table 4-2). However, the classification approaches most similar to AusRivAS, but predicting taxon densities rather than occurrence, might discriminate a maximum of between 7.5 and 9.0 categories of impairment. The best performing multivariate prediction approach, linear regression habitat models, might discriminate 10.6 categories and thus was the prediction approach most likely to be the most sensitive at detecting impact.

The nearest-neighbor approaches were not among the best prediction approaches by any measure (Table 4-2). Linke et al. (2005) found that Assessment by Nearest Neighbor Analysis (ANNA) had roughly equivalent prediction accuracy compared to AusRivAS, which is based on classification and discriminant analysis. The same conclusion can also be reached when taxon abundances are included. The nearest-neighbor prediction approach most closely similar to ANNA, DCA Nearest-Neighbor, had similar, though slightly lower, prediction accuracies than the taxon abundance equivalent of AusRivAS, Weighted Averaging MDA of Communities and MDA of Communities (Table 4-2). Nearest neighbor approaches have the potential to include cases/sites that are very different from the test site because a constant number of cases/sites is always used for comparison. A nearest-neighbor approach that contained a cut-off on distance at which a nearby site would be included for analysis might improve prediction accuracy, but would take a fair amount of calculation and may result in an over-fit model.

Weighting the predicted taxon densities in the approaches with a classification step did not substantially improve predictions of expected assemblages. From what is known about the lotic community types in New Hampshire (see Chapter 3), some communities that differ greatly from each other in taxa composition do not differ greatly on many of the environmental variables that discriminate the communities in a discriminant analysis. For example, a low-elevation warm-water riffle community type has a very different taxa composition than a cold-water riffle community. However, of the habitat predictors in the discriminant model, only elevation contributed substantially to discriminating the two communities. Thus, a reach that is higher in elevation than the average for the warm-water community type may be given a large enough probability of belonging to the cold-water community that weighted averaging of the taxon densities introduces some prediction error that would not be present when taxon densities are simply equal to the average for the most likely community.

Variation in stream assemblages was not adequately explained by ecoregions or ecoregions nested within stream order. The utility of these classifications for reducing unexplained variation in reference conditions is low. Only two metrics achieved coefficients of variation as low as those in observed to expected similarity attained by the multivariate predictive models. A study examining the explanatory power of ecoregions compared to biotic classifications also found that ecoregions displayed relatively poor power to explain lotic organism distributions and abundances (Chapter 3).

Although predictive modeling explained organism distributions and abundances in these reference streams better than ecoregions, the percent similarities between the observed and predicted communities did not reach 100% in any predictive modeling approach (Table 4-2). There are a number of reasons the expected communities derived from the predictive models might not closely match observed communities. The prediction errors that result from predictive modeling for bioassessment are actually a combination of all of the error inherent in every step of the model building process, from initial data collection to statistical model construction.

Sampling error in the field for both habitat conditions and biota, error due to sub-sampling the macroinvertebrates and periphyton in the laboratory, unexplained spatial variation due to inadequate predictor variables, unexplained temporal variability in assemblages, and error due to inappropriate assumptions in the statistical models all combine to produce the final prediction accuracy and error. Given the large number of quantitative habitat variables collected in this study, it is unlikely that the predictive models are limited by the availability of predictors of spatial variability. As of now, many of the potential statistical approaches to prediction have been explored and compared for prediction accuracy (Chessman 1999, Moss et al. 1999, Davies 2000, Reynoldson et al. 2000, Linke et al. 2005, Van Sickle et al. 2005, Van Sickle et al. 2006). The only additional avenues for improvement of statistical techniques foreseeable at this point lie in artificial neural networks or Bayesian approaches (Johnson 2000, Walley and Fontama 2000), though those approaches are very complex compared to the approaches explored thus far.

Cao et al. (2002) have demonstrated that sampling effort greatly affects the multivariate comparison of lotic fish and macroinvertebrate assemblages. They found that classification and group separation only stabilize when replicate samples achieve a percent community similarity of 100. Because Cao et al. analyzed the effects of sampling on Bray-Curtis distances, their results have broad implications for all multivariate approaches involving distances, including many of the approaches presented in this paper. In seeming contrast, Ostermiller and Hawkins (2004) found that predictions in a RIVPACS-style model in the Pacific Northwest of the United States were not greatly affected by sampling approach. They found that 50% of the prediction error was unexplained by either sub-sampling target (to 450 organisms) or combinations of sampling methods and collectors. However, additional samples or more accurate sampling methods are limited by the amount of sub-sampling of the collected samples in the laboratory when greater than 500 organisms are counted; sub-samples targets may need to be as high as 1000 for macroinvertebrates before the sub-samples resemble the whole sample and sub-samples below 500 did not differ in the amount of error produced (Doberstein et al. 2000). Indeed, Wolda

101

(1981) found that percent similarity of sub-samples compared to the whole sample in several datasets stayed between 0.6 and 0.7 for sub-samples of 500 organisms, which was used as the sub-sampling target in this study; only sub-sampling targets approaching the whole sample can achieve percent similarities near 1. Thus, the relatively low predicted percent similarity for all of the approaches except AusRivAS (Table 4-2), which was not based on community similarity, may be in part due to sub-sampling error. Greater sampling effort in the field and laboratory than has been conventionally exerted may be needed for accurate bioassessment.

Lastly, recent studies have found that temporal variation in lotic assemblages results in inter-annual Bray-Curtis similarities that are very close (mean similarity ranging between 0.49 to 0.70) to those achieved by the predictive models presented here (Milner et al. 2006, Chapter 5). This suggests that the accuracies of statistical models may be limited by the amount of unexplained temporal variation in assemblages. However, every measure of temporal variability is based on replicate samples through time; thus it also includes field and laboratory sub-sampling error. The relative contributions to model error of all the possible sources of error from field to statistical model construction need to be fully decomposed and separated to identify where the largest contributions to error are arising and identify strategic changes to reduce the total prediction error (Clarke 2000).

Conclusions

Separate linear regressions for each taxon present in this survey of 76 minimally impacted stream reaches in New Hampshire using environmental variables relatively robust to human influence yielded the most accurate prediction of stream reach assemblages compared to nine other multivariate prediction approaches. The accuracy was not much greater than the traditional AusRivAS approach involving discriminant analysis prediction of community types. Nearest neighbor and ordination approaches generated relatively poor predictions. The low

102

coefficient of variation in O/E similarity and high number of potential categories of impairment for the two best prediction approaches (MDA of Communities and Linear Regression Habitat Models) compared to the presence-absence AusRivAS approach indicate that variation or noise in lotic assemblages due to field sampling, sub-sampling, and natural variation is not great enough to reduce potential impact sensitivity. Thus, presence-absence does not appear to improve prediction of reference communities by reducing unexplained variation or noise in taxon abundances in these analyses (also see Chapter 5). A similar comparison of multivariate techniques for predicting organism abundances in other regions might yield different results, primarily related to how variable organism abundances are in the region and how discreetly they classify into community types. While these results suggest that organism abundances can be accurately predicted under reference conditions, the ability of predictive models that include taxon abundances to improve sensitivity at detecting impacts should be further investigated. Lastly, given the imperfect prediction of reference communities for bioassessment despite many varied attempts to improve statistical analysis (e.g. Moss et al. 1999, Hawkins et al. 2000, Linke et al. 2002), the most fruitful directions for improving the accuracy and sensitivity of bioassessments may be greater sampling effort and accuracy in the field and laboratory (Doberstein et al. 2000, Cao et al. 2002, Ostermiller and Hawkins 2004) and better understanding of temporal variability (Linke et al. 1999, Milner et al. 2006).

Table 4-1. Environmental descriptors that were determined to be robust enough to human influence to use for predictive modeling of expected reference assemblages for minimally impacted streams in New Hampshire. Variables in bold were excluded from multivariate analysis because they had a bivariate Spearman's R greater than 0.7 with another included robust habitat descriptor.

Robust Habitat Descriptors

Latitude in decimal degrees

Longitude in decimal degrees

Stream order

Percent of watershed as wetlands

Percent of watershed as lakes

Total length of permanent streams (m)

Total length of intermittent streams (m)

Maximum elevation in watershed (m)

Distance to nearest upstream impoundment (km)

Distance to furthest point in the upstream stream network (km)

Watershed area (ha)

Reach length (m)

Percent area of metamorphic bedrock

Percent area of volcanic bedrock

Percent area of plutonic bedrock

M

m^3/sec

Degrees

Degree deviation from south

Percent of 11 macroinvertebrate and periphyton sample sites classified as "fine" substrate

Percent of 11 macroinvertebrate and periphyton sample sites classified as "gravel" substrate

Percent of 11 macroinvertebrate and periphyton sample sites classified as "coarse" substrate

Percent of 11 macroinvertebrate and periphyton sample sites classified as "pool" habitat

Percent of 11 macroinvertebrate and periphyton sample sites classified as "glide" habitat

Percent of 11 macroinvertebrate and periphyton sample sites classified as "riffle" habitat

Percent of 11 macroinvertebrate and periphyton sample sites classified as "rapids" habitat

Mean phi value of substrate particles

Standard deviation of phi values for substrate particles

Percent of 55 substrate measurements classed as fine

Percent of 55 substrate measurements classed as sand

Percent of 55 substrate measurements classed as fine gravel

Percent of 55 substrate measurements classed as coarse gravel

Percent of 55 substrate measurements classed as cobble

Percent of 55 substrate measurements classed as boulder

Percent of 55 substrate measurements classed as bedrock

Mean water depth (m)

Maximum water depth (m)

Mean channel width (m)

Mean channel width x mean water depth (m^2)

Mean bankfull channel width (m)

Mean bankfull channel height (m)

Mean bankfull channel width x mean bankfull channel height (m2)

Annual flood magnitude indexed by the ratio of the bankfull cross-section to the wetted cross-section (unit-less)

104

Table 4-2. The summary results presented compare the accuracy and precision of eleven multivariate approaches to predicting reference community composition in minimally-impacted New Hampshire streams. O/E similarity was the Bray-Curtis percent similarity between the predicted community for a minimally impacted stream and the observed community composition.

	Mean O/E similarity	Min.	Max.	S^1	$C.V.^2$	Categories of s between mean O/E and zero
Weighted Averaging MDA of Communities	0.617	0.44	0.78	0.071	0.115	8.69
MDA of Communities	0.605	0.42	0.74	0.067	0.112	9.02
Weighted Averaging CART Prediction of Communities	0.612	0.38	0.74	0.077	0.125	7.95
CART Prediction of Communities	0.591	0.35	0.74	0.079	0.134	7.48
Linear Regression Habitat Models	0.656	0.51	0.79	0.062	0.095	10.58
CART Habitat Models	0.672	0.19	0.92	0.146	0.217	4.60
PCA Prediction	0.440	0.18	0.61	0.089	0.203	4.94
DCA Gaussian prediction	0.518	0.17	0.77	0.147	0.283	3.52
FA Nearest-Neighbor	0.536	0.27	0.74	0.081	0.151	6.62
DCA Nearest-Neighbor	0.561	0.41	0.68	0.068	0.120	8.25
AusRivAS Binary O/E	0.981	0.49	1.36	0.176	0.179	5.57

1 s = standard deviation
2 $C.V.$ = coefficient of variation

105

CHAPTER V

A COMPARISON OF APPROACHES TO PREDICT AND ASSESS STREAM HABITAT

Summary

Stream habitat quality assessment is complementary to biological assessment by

providing a mechanism for ruling out habitat degradation as a potential stressor and as reference

targets for the physical aspects of stream restoration projects. This chapter analyzed several

multivariate statistical approaches for predicting reference stream habitat conditions in habitat

assessments of potentially degraded streams based on discriminant analysis, linear regressions,

ordination, and nearest neighbor analyses. Quantitative physical and chemical habitat and

riparian conditions in minimally-impacted streams in New Hampshire were estimated using

USEPA Environmental Monitoring and Assessment Program protocols. A new assessment index

comparing and summarizing the degree of correspondence between predicted and observed

habitat based on Euclidean distance between the standardized habitat factors is described; the new

index avoids the erroneous prediction of multiple mutually exclusive habitat conditions that

confused previous habitat assessment approaches. Separate linear regression models for each

habitat descriptor yielded the most accurate and precise prediction of reference conditions with a

coefficient of variation (C.V.) between predictions and observations for all reference sites of

0.269. However, for a unified implementation in regions where a classification-based approach

has already been taken for biological assessment, a discriminant analysis approach that predicted

membership in biotic communities and compared the mean habitat features in the biotic

communities with the observed habitat features was a close second in prediction accuracy and precision. (C.V. = 0.293) As the best model still returned an error 27% of the mean index value for the reference sites, there still remains substantial room for improvement in habitat assessment techniques. Additional basin characteristics than were available for this analysis, such as average rainfall or winter snow-pack, surficial geological characteristics, or past land-use history, may improve the precision of the predicted habitat features in the reference streams. Land-use history in New Hampshire and regional environmental impacts have greatly impacted stream habitat conditions even in streams considered minimally-impacted today; thus as regional environmental impacts change and riparian forests mature, reference habitat conditions should be re-evaluated.

Introduction

Considerable effort has been expended recently on measuring the quality of lotic systems using resident biota. The premise for biological monitoring is that biota respond to and temporally integrate a wide range of physical and chemical parameters (Rosenberg and Resh 1993). Thus, they are the best indicator of the condition of an aquatic system (Reice and Wohlenberg 1993, Karr and Chu 1999). While measurement of river form and structure has a long history, the relative neglect recently of directly assessing the physical habitat of streams and rivers (hereafter 'streams') in ecosystem health assessments may stem from the view by some that physical (and chemical) measures may indicate causes of degradation rather than ecosystem condition or 'health' (Karr 1991, Chapman 1992). However, exactly because physical parameters indicate potential causes of degradation, this inattention has hampered assessment of a critical diagnostic of total stream 'health' (Maddock 1999, Norris and Thoms 1999).

Physical conditions, as separate from water chemistry, form the most central aspects of the habitat (the 'habitat templet' of Southwood [1977]) in which lotic organisms evolve, adapt,

and interact to form the assemblages (Hildrew and Giller 1994) that biological monitoring uses for assessment of stream condition. That organisms adapt and respond to physical habitat is well known (Hynes 1970, Allan 1995, Cushing et al. 1995, Petts and Calow 1996) and given the wide variety of adaptations and strategies, a through review is beyond the scope of this question. However, I believe an understanding of spatial and temporal scale in stream ecosystems is necessary to understanding that habitat templet, particularly the hierarchical organization of stream systems. In this hierarchical system, higher scales impose constraints on the lower scales nested within those higher scales (Frissell et al. 1986, Minshall 1988, Hildrew and Giller 1994). The processes operating at each scale differ in the spatiotemporal scale at which they work. Higher scale processes act over longer time frames and larger spatial areas. Minshall (1988) noted that streams operate on spatial and temporal scales that spread over 16 orders of magnitude. At the highest scales, stream ecology merges with terrestrial ecology such that, ultimately, "the valley rules the stream" (Hynes 1975). This view of scale and connection with the surrounding landscape is the theoretical framework within which any system to assess physical stream condition must operate. Because the end point of stream health is biological (Reice and Wohlenberg 1993, Karr and Chu 1999), the measurement of physical habitat needs to include factors that influence biotic communities at scales relevant to the organism of interest (Weins 1989; Cooper et al. 1998).

There are several avenues upon which stream habitat assessment has been pursued. Within a geomorphological perspective, measurement of habitat tends to relate fluvial processes to channel structure and form at scales that reflect the hierarchical organization of stream systems (Frissell *et al.* 1986; Harper and Everard 1998, Maddock, 1999). In this view, physical habitat is the result of predictable geomorphic processes (Harper and Everard 1998, Maddock, 1999). Because the stream habitat that forms is a result of these processes, the link between the biological perspective of habitat and the geomorphological perspective is that both consider a

'healthy' habitat necessary for a 'healthy' stream ecosystem (Maddock 1999, Norris and Thoms 1999). Examples of this hydrogeomorphic approach include the Geomorphic River Styles (Brierly et al. 1996), the Index of Stream Condition (Ladson et al. 1999), and the detailed, rigorous, hydrogeomorphic approach of Fitzpatrick and Knox (2000). They require very intensive field and historical surveys to complete their assessments and are, as such, not readily applicable to an extensive survey. In addition, the expected conditions at a site in the Geomorphic River Styles and Index of Stream Condition assessments are somewhat subjective.

The physical habitat component of the United States Environmental Protection Agency's (USEPA) Rapid Bioassessment Protocols (RBP) attempts to evaluate the structure of the surrounding physical habitat that influences the quality of the stream and the condition of the biological community (Barbour et al. 1999). The method is visually based and involves no quantitative sampling. Instead, the values for 10 components of stream habitat are assessed by consensus between 2 or more biologists after a walk of the 100 m stream segment. Each component metric is rated on a scale of 1-20, with 20 being optimal, based on narrative descriptions. Pictures are provided to aid scoring and each operator is trained to ensure some modicum of consistency. Because high and low gradient stream differ greatly in expected habitat conditions, there are separate narrative descriptions for each type. Scores are summed to make the composite index (Barbour et al. 1999).

There are a number of drawbacks to the rapid bioassessment approach. The reference conditions used are still based on highly subjective ideas of what an "ideal" stream looks like. Hannaford and Resh (1995) found very high inter-operability variance in metric scores even after training. Although low and high gradient streams are separated, there are still only two types of ideal stream conditions. However, the largest drawback of the habitat measurement of the rapid bioassessment protocol is that the measurements are not quantitative. Kaufmann and Hughes (2003) found that the RBP's qualitative measures of habitat had a signal to noise ratio in

109

reference sites that was only a tenth of the ratio for quantitative measurements of stream habitat. As a result, the scores of reference, minimally impacted streams in one study varied tremendously (Rabeni 2000), which reduces sensitivity at detecting alterations in lotic habitat.

There are several methods for describing stream habitat that use objective and quantitative measures. The USEPA's Environmental Monitoring and Assessment Program (EMAP) has developed detailed methods for collecting mostly quantitative measurements of stream habitat in a single site visit (Lazorchak et al. 1998). Along a similar avenue, the nationally standardized system to measure, classify and report on the physical structure of rivers in the United Kingdom, the River Habitat Survey (RHS), also measures many quantitative and presence/absence descriptors of stream habitat (Raven *et al.*, 1998). The end product of the RHS will be a prediction of the physical features of a stream that would occur under unmodified conditions for use in assessing potentially impacted streams for habitat degradation (Raven *et al.*, 1998). Kaufmann and Hughes (2003) found that the EMAP measurements explained double the amount of variance in biological metrics as the RBP methods. Clearly, quantitative measurements of habitat are required to represent habitat in a biologically meaningful manner and to produce an assessment process that is sensitive to habitat alterations that are not profound.

However, a method of assessing test streams in either program, EMAP or RHS, is still being developed. An approach similar to RIVPACS (Wright et al. 2000) for biological assessment may be taken in which the reference streams will be grouped into similar types and their occurrence predicted using predictor variables robust to human influence. New streams will be compared to the appropriate reference groups using those same predictor variables (Raven et al. 1998).

A RIVPACS-style approach is taken in the most developed predictive model for assessing stream habitat in large-scale and extensive regional surveys, the Habitat Predictions Modeling (HPM; Davies et al. 2000). In HPM, local scale habitat features and catchment scale variables are measured, usually on a continuous ratio scale. The habitat descriptors were

110

categorized into discreet categories and classified using UPGMA into groups of similar sites. The group membership of sites was then predicted using the large scale variables, thought to be robust to human influence, as predictors in a multiple discriminant analysis. However, the authors of HPM recognize that categorizing the ratio and interval scale data produces results that are sometimes problematic. When expected discreet habitat conditions are predicted, there was often prediction of multiple levels of a particular habitat condition. For example, categories of less than 20% sand and 40-50% sand could both be predicted to occur for a site. Davies et al. (2000) show that O/E ratios can be adversely affected by such results and suggest that statistical analyses be explored to predict continuous variables and correct these anomalies in prediction.

This chapter reports an investigation of five multivariate prediction methods for predicting habitat variables in minimally-impacted stream reaches in New Hampshire. All five methods avoid categorizing interval-scale habitat into discreet categories. The goal was to evaluate alternative methods for their accuracy at predicting habitat using environmental descriptors that are very difficult for humans to alter without leaving obvious evidence of alteration.

Methods

The dataset of physical attributes from minimally-impacted first to fourth order streams in New Hampshire (see Chapter 1) was used to compare multivariate predictive approaches for their ability to accurately and precisely predict reference habitat conditions. In all models, the environmental variables used to predict habitat were highly robust to human influence. They represented a mix of catchment-scale and local-scale variables (Table 5-1). To predict habitat for the purposes of assessing the degree of potential habitat alteration by humans, it is important that

111

the predictors are difficult to alter (Davies et al. 2000). Because there were four Level IV Aquatic Ecoregions in New Hampshire (Omernik 1987), ecoregion membership was transformed into three dummy-coded variables (Ecoregions 580, 581, 591) such that a 1 was assigned if a reach belonged to an ecoregion and a 0 if it did not. Measures of stream width and flood magnitude were not used to predictor habitat variables, despite their potential to be strong predictors, because they are often altered by flow regulation. The local-scale physical and chemical habitat features that were measured for prediction and assessment of habitat quality (Table 5-2) were thought to represent habitat features that are relevant to stream organisms (Weins 1989, Cooper et al. 1998).

In current bioassessment research, there is discussion about the role of classification in prediction of unimpaired or reference conditions. Streams are generally seen as transitioning continuously in physical and biological conditions from headwaters to mouth (Vannote et al. 1980, Gauch 1982, Statzner and Higler 1986). Thus, classification into discreet types may not be a powerful method for constraining variation in reference conditions. Multivariate approaches that require classification of the stream reaches into homogenous groups and those that do not require a classification step were investigated. Five approaches were examined to predict local-scale physical and chemical habitat features:

1. <u>MDA of Habitat Types:</u> The reaches were classified on the basis of the habitat descriptors using Ward's agglomerative clustering and Euclidean distance. To remove artificial effects on the classification of descriptors measured in units with different magnitudes, the habitat descriptors were transformed to a maximum magnitude of 1 by dividing each value by the maximum value for that descriptor. Membership in between 2 and 14 clusters were calculated. Classification strengths of the possible clusters were calculated using the sum of the univariate F-ratios for each habitat descriptor (Warton and Hudson 2004). A

112

stepwise multiple discriminant analysis was used to predict stream reach membership in the habitat types. Environmental predictors highly robust to human influence were the predictor variables. An F-to-enter of 0.1 and F-to-remove of 0.15 were used for stepwise selection. The predicted habitat descriptors were equal to the mean for each descriptor in the group to which membership was predicted. This approach is very similar to HPM (Davies et al. 2000); the key difference is that habitat descriptors are predicted on a continuous basis, not in discreet categories.

2. <u>MDA of Biotic Communities:</u> This approach was similar to the first except that a stepwise multiple discriminant analysis was used to predict stream reach membership in the biotic community types delineated for New Hampshire streams (Chapter 3). Biotic groupings were used as the grouping variable, instead of classifying the reaches on the basis of habitat as in HPM and the first approach above, because biological assessments and resource management decisions are often made using biotic community classifications. Using the biotic classifications would thus be simpler than basing the habitat assessment on a separate, parallel classification for stream reaches for habitat. Seven community types were predicted using the robust environmental descriptors as the predictor variables. An F-to-enter of 0.1 and F-to-remove of 0.15 were used for stepwise selection. The predicted habitat descriptors were equal to the mean for each descriptor in the group to which membership was predicted.

3. <u>Separate Linear Regressions:</u> A least-square stepwise linear regression model was constructed to separately predict each habitat variable. The robust environmental predictors were the independent variables and the habitat

descriptors were the separate dependent variables. The chemical concentrations were log transformed to improve their distributional characteristics, though the other variables were not substantially improved by log or square root transformations and were left in their original units. An F-to-enter of 0.1 and F-to-remove of 0.15 were used for stepwise selection.

4. Factor Analysis: The stream reaches were ordinated using principal axis factoring on the local-scale habitat descriptors to be predicted. The site scores on the first three axes were then predicted using separate stepwise least-squares linear regression models for each axis. The robust environmental predictors were the independent variables and the site scores were the dependent variables. An F-to-enter of 0.1 and F-to-remove of 0.15 were used for stepwise selection. The regression models were then used to predict the site scores for each reach and to calculate the predicted habitat descriptors based on the predicted site scores and their axis correlations.

5. Nearest-Neighbor: This approach is similar to the Analysis of Nearest Neighbor Assessment (ANNA) bioassessment approach (Linke et al. 2005) except that habitat is being predicted instead of organism occurrences. The reaches were ordinated using principle axis factoring, however, in contrast to the above ordination approach, the reaches were ordinated on the basis of the robust environmental predictors. The Euclidean distance between the reaches was calculated as the distance between the site scores. The predicted habitat was equal to the means for each habitat descriptor on the k nearest reaches. Values of k from 2 to 10 were investigated.

114

In all approaches, the predicted habitat for a reach was compared with the observed habitat using Euclidean distance. However, there were large differences in measurement units between habitat descriptors that could artificially influence the distance metric (Jongman et al. 1995). The predicted and observed habitat descriptors were thus transformed to range between 0 and 1 by dividing each value by the maximum observed value for each descriptor. As there is no meaning to the relative relationships between habitat descriptors, the loss of the absolute scale for each descriptor by this transformation was not as relevant as when organism abundances in an assemblage are transformed by the maximum abundance in each taxon. The use of a distance metric eliminates the need for categorizing the habitat descriptors into discreet categories to be predicted on a binary basis as in HPM (Davies et al. 2000).

The mean and standard deviation of the distances between the predicted and observed habitat for each reach were calculated for each of the four prediction approaches. Cao et al. (2002) used mean and standard deviation in similarity as basis for assessing sample representativeness for biological assemblages. As assessing the ability of multivariate models to predict multiple habitat descriptors is conceptually similar, this study also used the mean and standard deviation in the Euclidean distance between the predicted and observed habitat as the basis for comparing the accuracy and precision of prediction approaches. All statistical analyses were performed in Matlab 7.1.

Results

The multivariate approaches investigated for predicting reference conditions varied in their accuracy and precision (Table 5-3). Mean Euclidean distances between observed and predicted habitat descriptors ranged from 1.370 to 1.798. The coefficient of variation was used to assess precision instead of the standard deviation to separate the effects of higher mean distance

115

from error around the prediction. The precision of the predictions, as measured using the coefficient of variation in the Euclidean distances, ranged from 0.269 to 0.341 (Table 5-3).

Separate linear regressions for each habitat descriptor resulted in the lowest distances between the observed habitat and the predicted habitat for each minimally-impacted stream reach (Table 5-3). The coefficient of variation around the mean predicted distance and maximum predicted distance were also lowest. However, the approaches that achieved the next highest prediction accuracies and precisions were the classification-based approaches MDA of habitat types and MDA of biotic communities (Table 5-3). Classifying the sites on the basis of habitat or biota did not seem to affect prediction ability much as both approaches displayed very similar prediction accuracies.

Nearest-neighbor clustering had an optimal number of neighbors of 11. However, the classification-based approaches and separate linear regressions yielded better predictions. Factor analysis resulted in the worst predictions (Table 5-3).

While a full description of the model details and the environmental predictor variables and their statistical parameters for each approach would not fit into the constraints of this chapter, some general observations can be made. Stream order, elevation, watershed area, and aquatic ecoregion membership were consistently powerful predictors of habitat in the stepwise models. Overall, stream order was the most powerful predictor variable across all of the models as measured by either p-values or correlation coefficients. Percent of watershed as lakes appeared solely in the discriminant function to predict biotic communities and was otherwise absent from the models or a weak predictor.

Discussion

The Euclidean distance between the predicted habitat and the observed habitat at a test stream is the index upon which an assessment of habitat alteration would be based. As such, a sensitive method for predicting and assessing habitat quality or degree of alteration will display a low distance between predicted and observed habitat conditions in the reference reaches. Additionally, the coefficient of variation of the predicted distances in the reference reaches would ideally be low because it is the range of an index in reference conditions that partly determines how likely an assessment approach will detect alteration from expected reference conditions (Davis and Simon 1995, Barbour et al. 1995, Hughes 1995, Linke et al. 1999). Typically, some description of variance in the index value in the reference sites, such as standard deviation or inter-quartile range, is chosen beyond which test sites are considered to be different from reference conditions (Davis and Simon 1995, Wright et al. 2000).

Using these mechanics of assessment as a guide, separate linear regression models for each habitat descriptor yielded the most accurate and precise prediction of reference conditions (Table 5-3). However, in regions where a classification-based approach has already been taken for biological assessment, such as Australia's AusRivAS and the United Kingdom's RIVPACS, the discriminant analysis approach that predicted biotic communities and compared the mean habitat features in the biotic communities with the observed habitat features was a close second in prediction accuracy and precision (Table 5-3). The relative ranking of the prediction approaches on their accuracy is very similar to that found for predicting biological reference conditions; the separate linear regressions approach yielded the best prediction of reference assemblages for bioassessment (Chapter 4). Just as the discriminant analysis models resulted in the next best predictions habitat, they also yielded the next best predictions of organism densities (Chapter 4). The ordination approaches, i.e. factor analysis, produced the worst predictions. However, the habitat predictions had much higher coefficients of variation in the reference streams than even

117

the worst multivariate prediction approaches for biological assemblages. Additional large-scale predictor variables, such as additional basin characteristics beyond size, may improve the precision of the predicted habitat features in the reference streams.

Unlike for organisms, the presence/absence of a habitat feature does not always make sense. For example, sand nearly always occurs in a stream, the real question is in what quantity or spatial arrangement. Thus, unlike biological assessment where the relative power of detecting impacts using presence/absence of organisms versus density information is a viable topic of discussion, habitat descriptors for streams should be largely be quantitatively predicted. The approach to predicting discreet categories of quantitative habitat features of Davies et al. (2000) sometimes resulted in contradictory and confusing predictions of habitat; the index that resulted from the predictions were adversely affected by that approach.

Using Euclidean distance to compare the predicted and observed habitat features eliminates the need for discreetly categorizing the habitat features. There is no possibility of contradictory predictions as in the HPM approach (Davies et al. 2000). As the distance increases, there is more deviation from the expected minimally-impacted habitat conditions and thus this index of habitat quality increases with the degree of habitat alteration. A test stream with a Euclidean distance between its predicted and observed habitat beyond the range of distances in the reference streams would be beyond the natural variation in habitat.

However, much information is lost in reducing the assessment index to a single value. With 62 habitat descriptors in this study, there are many possible permutations of which features were different from reference condition. A single index does not elucidate what were the specific deviations. One mechanism for assessing which features deviated from the expected range would be to calculate the magnitude of deviations between predicted values and observed values for every habitat descriptor for every reference stream. The standard deviation in the deviations between predicted and observed values for each descriptor would provide a metric to compare the magnitude of the deviation of each habitat feature in a test stream. If the magnitude of the

observed deviation in a test stream was larger than the range of deviations found in the reference streams, then that specific habitat feature would be considered altered in the test stream.

Both the Euclidean distance index and the deviations in the habitat features would provide information for diagnosing the cause or stressor in streams that were assessed to be biologically impaired. In the absence of water quality degradation, physical habitat will greatly influence biotic communities. As such, a parallel assessment of habitat can discriminate whether biotic impairment at a site is related to poor habitat quality or to water quality degradation by a process of elimination akin to medical diagnosis. Sites that are biologically impaired but do not show habitat impairment can be concluded to be impacted by chemical water quality degradation (Davies et al., 2000). The modeled habitat features also have the potential to form restoration targets for a stream that are not based on subjective judgment but on empirical relationships in minimally impacted streams (Davies et al. 2000).

Conclusions

Assessment of habitat provides powerful tools for ecosystem management, restoration, assessment, and monitoring. Considering the large amount of laboratory time that goes into each biological assessment of a stream using macroinvertebrates, making additional measurements of habitat features that takes on average four hours is not a substantial increase in the effort already being expended in stream ecosystem assessment. The benefits are potential stressor identification and restoration targets for impaired streams. However, probably even more than for stream biota, stream habitat has been greatly influenced by historical land management (Foster et al. 2003). Following land clearance, particularly in forested regions such as New England, streams change their morphology through bed and bank erosion and large woody debris is often decreased compared to forested catchments; often streams have coarser substrate and fewer pools. These

119

changes persist and are detectable more than 100 years after prior land clearance (Foster et al. 2003). As such, the reference habitat conditions and the restoration targets (e.g. number of large woody debris pieces or stream substrate characteristics) inferred using currently minimally impacted streams may be substantially different than before forest clearance. Additionally, changes in hydrology due to climate change may alter the habitat conditions to which streams in a region can attain even in the absence of direct human alteration. As lotic ecosystems and their watersheds mature following the abandonment of large-scale land clearance, stream habitat reference conditions will need to be re-assessed on a periodic basis and should in no way be taken as static or equivalent to pre-human conditions.

120

Table 5-1. Environmental descriptors to predict local-scale physical and chemical habitat of stream reaches in New Hampshire that were considered to be very robust to human influence. Variables in bold were excluded from multivariate analysis due to high univariate correlations (R > 0.7) with other predictors.

Predictor Environmental Descriptors
Latitude
Longitude
Strahler stream order
Percent of watershed as wetlands
Percent of watershed as lakes
Total length of permanent streams (m)
Total length of intermittent streams (m)
Maximum elevation in watershed (m)
Distance to nearest upstream impoundment (km)
Distance from source (km)
Watershed area (ha)
Reach length (m)
Percent metamorphic bedrock
Percent volcanic bedrock
Percent plutonic bedrock
Elevation (m)
Stream slope (%)
Degree deviation from south
Aquatic ecoregion

121

Table 5-2. Local-scale physical and chemical habitat conditions for stream reaches in New Hampshire that were predicted using robust environmental variables.

Description and Units

Discharge (m³/sec)

Percent of 11 macroinvertebrate and periphyton sample sites classified as "pool" habitat

Percent of 11 macroinvertebrate and periphyton sample sites classified as "glide" habitat

Percent of 11 macroinvertebrate and periphyton sample sites classified as "riffle" habitat

Percent of 11 macroinvertebrate and periphyton sample sites classified as "rapids" habitat

Water pH

Water temperature (°C)

Specific conductivity (uS)

Mean percent embeddedness of substrate

Mean phi value of substrate particles

Standard deviation of phi values for substrate particles

Percent of 55 substrate measurements classed as fine

Percent of 55 substrate measurements classed as sand

Percent of 55 substrate measurements classed as fine gravel

Percent of 55 substrate measurements classed as coarse gravel

Percent of 55 substrate measurements classed as cobble

Percent of 55 substrate measurements classed as boulder

Percent of 55 substrate measurements classed as bedrock

Mean water depth (m)

Maximum water depth (m)

Mean channel width (m)

Mean channel width x mean water depth (m²)

Mean bankfull channel width (m)

Mean bankfull channel height (m)

Mean bankfull channel width x mean bankfull channel height (m2)

Annual flood magnitude indexed by the bankfull cross-section divided by wetted cross-section (unit-less)

Mean undercut length (m)

Mean Bank angle (degrees)

Mean cover index of filamentous algae in the channel (1-4)

Mean cover index of macrophytes in the channel (1-4)

Mean cover index of large woody debris in the channel (1-4)

Mean cover index of brush in the channel (1-4)

Mean cover index of live trees along the bank (1-4)

Mean cover index of overhanging vegetation along the bank (1-4)

Mean cover index of undercut along the bank (1-4)

Mean cover index of boulders in the channel (1-4)

Mean % canopy cover

Mean cover index of large trees (>0.3 m DBH) in the canopy riparian vegetation (1-4)

Mean cover index of small trees (<0.3 m DBH) in the canopy riparian vegetation (1-4)

Mean cover index of shrubs in the under-story riparian vegetation (1-4)

Mean cover index of tall herbs (>0.5 m) in the under-story riparian vegetation (1-4)

Mean cover index of short herbs (<0.5 m) in the ground cover riparian vegetation (1-4)

Mean cover index of bare ground in the under-story riparian vegetation (1-4)

Percent of riparian canopy vegetation classified as "deciduous"

Percent of riparian canopy vegetation classified as "coniferous"

Percent of riparian canopy vegetation classified as "evergreen"

Percent of riparian canopy vegetation classified as "mixed" (>20% other)

Percent of riparian canopy vegetation classified as "none present"

122

Table 5-2. Continued.

Description and Units
Density (per m of reach length) of large woody debris 0.1-0.3 m in diameter and 1.5 -5 m in length in the bankfull channel
Density (per m of reach length) of large woody debris >0.3 m in diameter and 1.5 -5 m in length in the bankfull channel
Density (per m of reach length) of large woody debris 0.1-0.3 m in diameter and 5-15 m in length in the bankfull channel
Density (per m of reach length) of large woody debris >0.3 m in diameter and 5-15 m in length in the bankfull channel
Density (per m of reach length) of large woody debris 0.1-0.3 m in diameter and 1.5 -5 m in length above the bankfull channel
Density (per m of reach length) of large woody debris >0.3 m in diameter and 1.5 -5 m in length above the bankfull channel
Density (per m of reach length) of large woody debris 0.1-0.3 m in diameter and 5-15 m in length above the bankfull channel
Density (per m of reach length) of large woody debris >0.3 m in diameter and 5-15 m in length above the bankfull channel
Density of all large woody debris within the bankfull channel (per m of reach length)
Density of all large woody debris above the bankfull channel (per m of reach length)
Chloride (mg Cl/L)
Nitrate (mg N/L)
Sulfate (mg S/L)
Sodium (mg Na/L)
Potassium (mg K/L)
Magnesium (mg Mg/L)
Calcium (mg Ca/L)
Ammonia (ug N/L)
Phosphate (ug P/L)
Silica (mg SiO_2/L)
Dissolved organic Carbon (mg C/L)
Total dissolved Nitrogen (mg N/L)
Dissolved organic Nitrogen (mg N/L)
Total alkalinity (mg $CaCO_3$/L)

123

Table 5-3. Summary statistics for the Euclidean distance between the observed and predicted habitat are presented for each multivariate prediction approach.

	Mean Distance	Min.	Max.	s	C.V.
Nearest-neighbor	1.700	0.873	3.555	0.552	0.325
Separate linear regressions	1.370	0.737	2.406	0.369	0.269
MDA of biotic communities	1.567	0.799	2.995	0.459	0.293
MDA of habitat types	1.617	0.861	3.306	0.552	0.341
Factor analysis	1.798	1.062	3.582	0.596	0.331

124

CHAPTER VI

ANNUAL VARIATION AND SPATE RESPONSE OF STREAM ASSEMBLAGES IN NEW HAMPSHIRE

Summary

The biological condition of ecosystems is increasingly being used to indicate the level of anthropogenic alteration to the ecosystems. However, the sensitivity of assessments to anthropogenic alteration is in part constrained by the degree of natural temporal variation in the reference ecosystems. This study quantified temporal variability in taxonomic composition and abundances for vertebrate (fish and stream-dwelling salamanders), macroinvertebrate, and periphyton assemblages in minimally-impacted streams in New Hampshire. Two types of temporal variability were investigated: inter-annual variability in base-flow summer assemblages over a four-year period and variability due to a natural flow disturbance resulting from unusually high rainfall during the summer base-flow period. The macroinvertebrate assemblages were the most stable taxonomic group on an inter-annual time scale, inter-annual Bray-Curtis percent similarities ranged between 0.133 and 0.400 in the vertebrates, 0.666 to 0.703 in the macroinvertebrates, and 0.122 and 0.616 in the periphyton. Sorenson's similarities based on taxon occurrences were very similar indicating that most of the inter-annual variation was due to changes in taxon presence rather than abundance. In contrast, the vertebrates displayed the highest Bray-Curtis similarities (0.786 to 0.0.974) before and after a spate, periphyton the lowest (0 to 0.351), and macroinvertebrates intermediate resistance (0.493 to 0.725). A model to explain

these taxonomic differences is described wherein lagged response of vertebrates to disturbance results in high inter-annual variability while relatively quick recovery by the periphyton to disturbance results in relatively higher inter-annual similarities. However, temporal variability as measured in this and other studies is a actually mix of sampling error in the field, error due to sub-sampling in the laboratory, and true temporal variability in assemblages. The relative contributions of those factors to variation need to be decomposed to quantify true temporal variability. Incorporating predictors of temporal variability, such as past rainfall, snow-pack, and other weather parameters, might increase our ability to explain temporal variation in reference streams and increase the sensitivity of biological assessments at detecting differences in biological condition due solely to anthropogenic impacts.

Introduction

The study of biological community persistence, the stability of organism occurrences and abundances over time, has a long history in ecology (Connell and Sousa 1983). The persistence and stability of ecological communities has broad level implications for our understanding of ecosystems, conservation biology, biological community classification, and more recently, ecosystem management and biological assessment (Weatherly and Ormerod 1990). The biological assemblages of ecosystems are increasingly being used to indicate the level of anthropogenic alteration to the ecosystem, and in the case of streams, also to the watershed surrounding it (Wright et al. 1984, Rosenberg and Resh 1993, Davis and Simon 1995, Karr and Chu 1999, Marchant et al. 1997, Yoder and Rankin 1998, Wright et al. 2000). The reference condition approach to biological assessment (Cairns et al. 1993, Hughes 1995) has seen widespread adoption. In the reference condition framework, biological criteria have been defined as expressions that describe the least impacted (reference) biological integrity of communities

126

within a region (Cairns et al. 1993, Hughes 1995, Reynoldson et al. 1997) such that the acceptable range in biological integrity is the variation unimpaired reference sites (Linke et al. 1999). Thus, the level of persistence communities display is a central aspect of the reference condition approach (Weatherly and Ormerod 1990, Humphrey et al. 2000).

Three conditions are required for accuracy and sensitive bioassessment using the reference condition approach: (1) low unexplained spatial variability of community composition (taxa occurrences) and community structure (taxa abundances) in the reference sites (Cairns et al. 1993, Hughes 1995), (2) low unexplained temporal variability in community composition and structure (Weatherly and Ormerod 1990), and (3) large response of the taxa to human-caused stressors relative to the unexplained variability (Rosenberg and Resh 1993, Karr and Chu 1999). Many studies have examined ways to best explain spatial variation in the reference site assemblages through either predictive modeling or multimetric approaches (see Reynoldson et al. 1997 for a review as well as Chessman 1999, Linke et al. 2005, and Chapter 4). Gauging the sensitivity of taxa to human induced stressors is a difficult task that requires experimentation and carefully controlled observation if circular logic is to be avoided (Reynoldson et al. 1997).

There have been many studies of temporal variability in streams in response to flow disturbance and variability such as droughts and floods (Townsend et al. 1987, Meffe and Minkley 1987, Richards and Minshall 1992, Bradt et al. 1999). To the extent that these disturbances are similar to those produced by human disturbance and alteration, they provide some measure of the response of taxa to anthropogenic stress. For example, Richards and Minshall (1992) found that communities in anthropogenically disturbed streams were more similar to one another than to reference streams and *vice versa* over a five year period and concluded that the response to catchment-wide disturbance was larger than temporal variability, a requisite for successful bioassessment. Recent emphasis is being placed on developing a greater understanding of temporal variability in the structure and composition of minimally disturbed communities (Humprhey et al. 2000, Robinson et al. 2000, Metzeling et al. 2002, Milner et al.

127

2006) as low temporal variability in the biological reference conditions is also a requisite for sensitive bioassessment (Weatherly and Ormerod 1990) that has seen less attention (Metzeling et al. 2002).

However, the few studies of temporal variability in minimally-impacted stream ecosystems have focused on only one taxonomic group at a time. Limited resources in government agencies responsible for stream monitoring have often forced a decision as to which of the several major taxonomic groups in streams (e.g. fish, periphyton, and macroinvertebrates) to focus on for bioassessment. The taxonomic groups have been found to differ in their response to human impacts (Joy and Death 2002). Including all taxonomic groups present in an ecosystem increases the ability of a biological assessment to detect impacts (Karr 1991, Metcalfe-Smith 1996). However, they may differ in their temporal stability. An understanding of the relative amount of temporal variability of the major taxonomic groups in reference streams within the same region would enhance reasonable decision making as to which groups to include when constructing predictive models for bioassessment.

The goal of this study was to quantify temporal variability in taxonomic composition and abundances for vertebrate (fish and stream-dwelling salamanders), macroinvertebrate, and periphyton assemblages in minimally-impacted streams in New Hampshire. Two types of temporal variability were investigated: inter-annual variability in base-flow, summer assemblages over a four-year period and variability due to a natural flow disturbance resulting from unusually high rainfall during the summer base-flow period. The temporal variability of the taxonomic groups were compared to provide a sound basis for selecting which of the taxonomic groups to use in a biological assessment and monitoring system in New Hampshire and to inform the general question of whether aquatic assemblages in New Hampshire are stable enough to use for bioassessment of human impacts.

128

Methods

All streams were sampled for biota and habitat descriptors using the methods described in Chapter 1.

Annual Variation

Three minimally-impacted stream reaches in New Hampshire were sampled annually for vertebrate and macroinvertebrate densities and periphyton biovolume in the summers between 2002 and 2005 (see Chapter 1 for detailed field methods). The three sites were part of the initial batch of reaches sampled in 2002. They were selected to represent all three Level IV Aquatic Ecoregions (Omernik 1987) sampled in 2002 and a broad range in habitat types available in the first year of sampling (Table 6-1). Nelson Brook reach is a cold-water, high gradient brook with coarse substrate. The Oyster River reach is a small low-gradient, headwaters stream of a coastal river; water it drained a large wetland complex and the substrate varied between angular boulders and fine substrate. The Lovell River is a large river with alternating deep sandy pools and gravel riffles.

The annual stability of stream assemblages was quantified in several ways. Because the reaches differed greatly in the taxon composition, all calculations of annual variability were done within reaches. To examine annual variation in taxon occurrences (community composition), the number of years in which a taxon was present in the annual samples within a reach, hereafter termed co-occurrences, were calculated to quantify annual variation in community composition; as each reach was sampled for four consecutive years, four was the maximum co-occurrence value. The mean co-occurrence of taxa within a reach was calculated only for taxa that were detected at least once in a reach. Sorenson's coefficients of community (Sorenson's CC) were calculated on taxon occurrences between each year-pair within reaches. Bray-Curtis similarity

129

reduces to Sorenson's CC when presence/absence data is used (Jongman et al. 1995) and is thus the most similar binary similarity metric to compare with Bray-Curtis similarity. Similarities were calculated for all taxa combined and for the vertebrates, macroinvertebrates, and periphyton separately to investigate taxonomic differences in annual variation in community composition.

To examine annual variation organism abundances, the coefficients of variation in densities were calculated for each taxon within reaches (as was done by Robinson et al. 2000). Bray-Curtis percent similarities were calculated between each year-pair within reaches as it is the often seen as the best distance metric for ecological data (Faith et al. 1987) and to facilitate comparison with one of the few other studies of annual variation in stream assemblages for bioassessment (Metzeling et al. 2002, Milner et al. 2006). Bray-Curtis distance is also the bioassessment index in predictive multivariate bioassessment models developed for wade-able New Hampshire streams (Chapter 3). Bray-Curtis distances were calculated for all taxa combined and for the vertebrates, macroinvertebrates, and periphyton separately to investigate taxonomic differences in annual variation in community structure. To facilitate comparison with similar studies (Metzeling et al. 2002), Spearman rank correlation coefficients were calculated between taxon densities in year-pairs by reach (see Weatherly and Ormerod 1990).

Spate Response

During August of 2003, a series of large rainfall events raised many streams in the White Mountains region of New Hampshire above flood stage on August 10; stream flows during the usual summer base-flow period exceeded even the spring snowmelt (Figure 6-1). Five streams that were sampled for lotic assemblages no earlier than three weeks prior to the regional flooding were re-sampled no more than two weeks after the stream flow receded to the normal range

130

discharges (after August 16) using the same sampling techniques. Habitat descriptions for each reach can be found in Table 6-1.

To test for effects of the spate on taxon densities before and after the spate, paired t-tests were performed on the pre- and post-spate log-transformed densities for each taxon. The additional descriptors of community composition and structure total vertebrate density and Shannon diversity, total macroinvertebrate density and Shannon diversity, total periphyton density and Shannon diversity, as well as density and family richness of macroinvertebrate orders were tested for spate effects using paired t-tests. Bray-Curtis similarities were also calculated between pre-spate and post-spate samples for each reach. To investigate taxonomic differences in spate response, similarities were calculated for all taxa combined and for the vertebrates, macroinvertebrates, and periphyton separately.

Results

Annual Variation

Community composition varied greatly between sampling years within a reach (Table 6-2). The average number of times a taxon was detected within each reach over the four years of annual re-sampling (co-occurrences) ranged between 2.029 and 2.304. Thus, on average a taxon was detected in a reach in only half of the annual samples. The majority of taxa were detected only once within a reach (Table 6-2). Mean Bray-Curtis similarities (Sorenson's index) between the year-pairs varied between 0 and 0.713 (Table 6-3). Community composition stability was highest for macroinvertebrates and lowest for vertebrates. Only three vertebrate individuals were detected in the four years of sampling the Oyster River (Table 6-2); the mean similarity for that reach for vertebrates is therefore misleadingly low. Nonetheless, community composition

131

stability in the other two reaches, which had greater vertebrate densities, was low compared to the other taxonomic groups (Table 6-3).

Taxon densities varied substantially from year to year. The mean coefficients of variation in taxon densities in Nelson Brook, the Oyster River, and the Lovell River were 1.437, 1.469, and 1.249, respectively (Table 6-2). Even relatively common and abundant taxa such as Chironomidae and *Salvelinus fontinalis* had inter-annual variation in densities greater than 80% of the mean (Table 6-2). Common taxa contributed as much to total inter-annual variation as rare taxa (Figure 6-2).

As measured by inter-annual Bray-Curtis similarities, the three streams annually re-sampled were as stable in community structure as in community composition. Bray-Curtis similarities ranged between 0 and 0.703 (Table 6-3). There was very close agreement between community composition and community structure in relative taxonomic group stability; vertebrates were least stable in community structure and macroinvertebrates the most stable. The same caveat that the Oyster River reach had very few vertebrates (Table 6-2) applies to interpreting inter-annual stability in vertebrate densities. Spearman R^2 between taxon densities in year-pairs followed the Bray-Curtis similarity results very closely (Table 6-4).

Spate Response

The only significant effects of the spate were on total periphyton abundance ($p = 0.017$, n = 5); periphyton biovolume density averaged 145 mm^3/m^2 prior to the spate and 8 mm^3/m^2 after. All other paired t-tests were not significant at the $p < 0.05$ level. Bray-Curtis similarities between pre- and post-spate samples for all taxa combined ranged between 0.500 and 0.689 (Table 6-5). In contrast to the taxonomic differences in annual variation in stream community composition and structure, the vertebrate similarities were very high compared to the other taxa. Periphyton

community structure changed substantially in response to the spate, while macroinvertebrate community structure responses were intermediate between the taxonomic groups (Table 6-5).

Discussion

These patterns in temporal variability of lotic organisms provide valuable insights for interpreting ecological change and bioassessment in lotic ecosystems. The findings in this chapter are consistent in many respects with other studies of inter-annual variation in minimally-impacted stream macroinvertebrate communities. However, this was the first study of annual variation in stream assemblages that examined the major taxonomic groups together. There were stark differences between the response of the three major taxonomic groups to the summer spate and in temporal variability that have broad implications for bioassessment using those taxonomic groups.

Milner et al. (2006) found that the mean inter-annual variability in genus-level community composition (taxon occurrences), as measured using Jaccard's similarity, in six sites sampled over a nine-year period ranged between 0.49 and 0.70. The similarities in minimally-impacted New Hampshire streams as measured using Sorenson's CC were similar, though somewhat higher, to the Alaskan streams Milner et al. (2006) examined; mean similarity in community composition ranged between 0.59 and 0.71 (Table 6-3). An increase in stability at the family-level is consistent with other examinations of community stability in minimally impacted streams. For instance, family-level macroinvertebrate community structure was found to be more stable over time than genus-species community composition in Australian streams (Metzeling et al. 2002). As genus-level identifications were not available, no direct comparisons of stability between macroinvertebrate taxonomic identification levels can be made in this study.

There was also some agreement with similar studies on inter-annual variation in minimally-impacted stream macroinvertebrate community structure. Metzeling et al. (2002) reported mean stability in family-level community structure, as measured by Spearman's correlations, of 0.59. Spearman correlations between annual estimates of macroinvertebrate community structure in this study were very similar (Table 6-4). Robinson et al. (2000) found that coefficients of variation in individual taxon annual abundances in minimally-impacted Idaho streams were high and also greater than variation in community measures. This study also found high inter-annual variation in taxon densities. This study also found that community measures, such as Bray-Curtis similarity and Spearman correlations, varied less than individual taxon densities, which exhibited very high annual variation (Tables 2 and 3). Thus, it appears that community similarity measures for macroinvertebrates tend to be less variable than individual taxon abundances, though the number of studies is small.

However, there were some differences with previous studies with regard to the relative stabilities of community composition and structure. In previous comparisons of community structure and composition, community composition was found to be more stable over time than community structure. Milner et al. (2006) found that mean inter-annual variability in community structure (taxon abundances), as measured using Bray-Curtis similarities, was higher than for community composition and ranged between 0.28 and 0.44 (compare with community composition similarities above). In contrast, this study found that macroinvertebrate community structure, measured by Bray-Curtis similarity, was as stable as community composition measured using Sorenson's CC (Table 6-3). In contrast to Robinson et al. (2000), rare taxa did not contribute a disproportionately high share of the total variation in community structure in the three minimally-impacted streams examined (Figure 6-2). Common taxa displayed very high coefficients of variation in annual samples (Table 6-2).

It was somewhat of a surprise that annual variation in vertebrate (primarily fish species) composition and structure seems to have been much higher than annual variation in

134

macroinvertebrate families (Tables 3 and 4). Other studies have found high stability of fish community composition and structure (Ross et al. 1985, Matthews et al. 1988). Hoeinghaus et al. (2003) reported fish community structure stability between two samples years between 0.67 and 0.87 (Morisita-Horn index). Fausch and Bramblett (1991) noted a difference in stability between streams of differing morphology; they concluded that fish species composition structure remained relatively constant in streams with deep pools and diverse habitats but more variable at sites with shallow and/or less diverse habitats. In contrast to Fausch and Bramblett's conclusions, the site with the greatest number of pools in this study, the Lovell River, displayed the lowest inter-annual stability in fish community structure. However, the stream with the greatest diversity of habitat types and substrate, Nelson Brook, did indeed display the highest stability in fish community structure (Tables 1, 3, and 4).

The vertebrates were hardly affected by the summer spate (Table 6-5), yet as already discussed, displayed high annual variation in community composition and structure at two of the three reaches sampled annually (Tables 3 and 4). The vertebrates in a species-poor region such as New Hampshire already faced challenges in their use for accurate bioassessment of lotic ecosystems (e.g. Lyons et al. 1996). Whether the low stability of vertebrate assemblages in this study is due to the low species richness is not clear from these data. It is unlikely that the sampling approach for the vertebrates was responsible for the low temporal community stability as very high similarities were attained between the pre- and post-spate sample estimates. Whatever the cause, the large amount of unexplained temporal variability in the vertebrates assemblages makes them a poor choice for bioassessment.

In contrast, the macroinvertebrate community appeared to be the most stable on an annual scale (Tables 3 and 4) yet were substantially affected by the summer spate (Table 6-5). Macroinvertebrates have long been argued as the ideal choice for bioassessment as, unlike the vertebrates, they are present in all streams and are diverse enough in most regions to potentially respond in unique ways to different stressors, aiding stressor identification (Rosenberg and Resh

135

1993). When temporal stability is also taken into account, it appears they may also be the best taxonomic group for bioassessment.

Inter-annual stability in periphyton community composition and structure differed by stream. It was surprising that the periphyton were as stable as the macroinvertebrates in two of the three streams (Tables 3 and 4). Periphyton in the Lovell River and Oyster River were as stable as macroinvertebrate assemblages. Though, the periphyton communities at Nelson Brook varied greatly from year to year (Tables 3 and 4). Nelson Brook had a higher gradient than the other streams studied. Spates dramatically reduce periphyton abundances and alter community structure over the scale of weeks (Table 6-5). As scouring is more likely in Nelson Brook during high-flows, stochastic high-flows may be responsible for high variation in periphyton community at Nelson Brook. The effects of normal variation in stream flow interacting with stream habitat on periphyton communities need more attention if they are to be used for stream bioassessment.

The differences in community stability by taxa in response to the summer spate and annual re-samples suggests a theoretical model for how these taxonomic groups respond to perturbations in New Hampshire (Figure 6-3). The periphyton's large response to the spate but relatively stable assemblage in the annual re-samples suggests they are sensitive to flow perturbations, but recover quickly enough to present roughly the same assemblage in subsequent years. The macroinvertebrates displayed an intermediate response to the summer spate, but high stability from year to year. It is postulated that the macroinvertebrate community is substantially impacted by summer spates and continues to decline for a short while following the spate in response to lower periphyton biovolume and leaf-litter that has not yet been colonized by decomposing bacteria and fungi. However, the macroinvertebrate assemblage recovers to pre-spate structure through high fecundity and predominantly univoltine life-histories such that the community comes to resemble the pre-spate community after one year. In contrast, it is hypothesized that the vertebrates are not immediately affected by spate conditions due to their large body size. Instead, over a longer time frame, they experience population decline due to

136

lowered food resources, primarily macroinvertebrates, following a summer spate. The lower fecundity and longer life-cycles of fish in cold, nutrient poor streams seem preclude a rapid recovery to pre-spate conditions. Thus, environmental variability causes the vertebrate assemblage to cycle more erratically on an annual scale, which would explain the high annual variability in vertebrate assemblages and low initial spate response (Tables 3 and 5). Clearly, this tentative hypothetical model for temporal change in New Hampshire streams will need to be tested with long-term monitoring and careful experiments (Franklin 1989).

The overall question for this paper is whether temporal variability in stream assemblages is stable enough for bioassessment? The high coefficients of variation in all individual taxa in this study (Table 6-2) suggest that interpretation of individual taxa to infer that a stress is present or infer the type of stressor would be highly suspect. Similarly, Milner et al. (2006) found that the usual metrics used in multimetric approaches to bioassessment varied over their entire range over nine years in reference sites in Alaska. However, the interpretation of community-level measures such as distance indices and other multivariate statistical approaches may be stable enough over time to detect alteration to community structure. Several studies have found that despite substantial temporal variability in macroinvertebrate community structure and composition, there is enough of a range in similarity metrics to allow for a response above and beyond temporal variability (Metzeling et al. 2002, Milner et al. 2006). The Bray-Curtis similarity between assemblages predicted for reference streams using several multivariate predictive models in New Hampshire streams and observed assemblages achieved a similarity of 0.656 (S.D. = 0.062) in the most accurate approach (Chapter 4). The predicted similarities and temporal similarities in the macroinvertebrates (Table 6-3) are very similar and may indicate that the predictive models are limited in their prediction accuracy primarily by the amount of temporal variability and that the particular level of temporal variability observed in this study may not be enough to seriously impair multivariate predictions.

137

What has been termed temporal variability in this study and previous studies like it is actually a mix of sampling error in the field, error due to sub—sampling the macroinvertebrates and periphyton in the laboratory, and unexplained temporal variability in assemblages. The amount of temporal variation may be over-estimated if substantial variation is also introduced by field sampling (Karr and Chu 1999) and laboratory sub-sampling (Doberstein et al. 2000). Thus, the relative contributions to variation of temporal change and sampling methods needs to be assessed before the true nature of temporal variability can be evaluated. The same is true for predictive multivariate model accuracies (Chapter 4). Most of the studies that have examined long-term trends in lotic assemblages have used relatively small sub-sampling targets and have sometime differed in sampling strategies over time, which may have over-estimated temporal variability in macroinvertebrates. Increasing sub-sampling targets for macroinvertebrates to 1000 may eliminate the contribution of sub-sampling to measures of temporal variability (Doberstein et al. 2000).

If temporal variability in community structure remains as substantial as it currently seems with better sampling methods, then explanatory predictors of temporal variability will need to be incorporated into multivariate predictive modeling for bioassessment if prediction accuracies are to be improved. The substantial differences in temporal stability and spate responses between the streams (Tables 3, 4, and 5) argues that a greater understanding of the intersection of temporal variability and site conditions is necessary if temporal variability is to be accounted for in bioassessment. Poff and Ward (1989) offer a framework for investigating those interactions. However, incorporating predictors of temporal effects on assemblages into multivariate predictive models would require sampling reference sites for many consecutive years along with a variety of potential predictors, such as rainfall or winter snow pack (Bradt et al. 1999), to ascertain which are important to explaining hitherto unexplained temporal variability in lotic community structure. That task is enormous for any region, especially given that reference sites are usually remotes and difficult to access (Milner et al. 2006). It may be more efficient to focus on

138

improving field and laboratory sampling and then use some long-term data collected using

improved sampling methods to assess if temporal variability and the feasibility of controlling for

it in predictive bioassessment models.

139

Table 6-1. Habitat Description for the minimally-impacted New Hampshire stream reaches sampled on an annual basis and those sampled before and after a summer spate.

	Annual Re-sample Reaches			Spate Response Reaches				
	Nelson Brook	*Oyster River*	*Lovell River*	*Moriah*	*Mad*	*Smarts*	*Stony*	*Wonalancet*
Order	2	1	3	3	2	2	2	2
Elevation (m)	239	73	155	338	513	430	268	358
Discharge (m³/sec)	0.036	1.905	0.702	1.026	0.969	0.427	0.063	0.516
Slope (%)	3.6	4.5	1.0	2.8	2.7	4.2	6.0	6.5
Pools	36%	9%	73%	0%	9%	18%	45%	18%
Glide	9%	18%	0%	0%	0%	0%	0%	18%
Riffle	55%	45%	27%	82%	73%	64%	27%	36%
Rapids	0%	36%	0%	18%	18%	18%	27%	27%
pH	7.02	4.42	6.42	7.08	6.84	6.64	6.01	6.95
Temp. (°C)	17.7	13.4	16.8	19.5	13.6	16.5	15.8	12.8
Fine	0%	14%	0%	0%	0%	0%	5%	4%
Sand	1%	4%	26%	0%	2%	4%	11%	25%
Fine gravel	6%	4%	1%	16%	16%	7%	13%	7%
Coarse gravel	37%	2%	55%	24%	25%	29%	31%	13%
Cobble	33%	17%	17%	25%	40%	31%	22%	22%
Boulder	23%	41%	1%	35%	16%	29%	16%	24%
Bedrock	0%	0%	0%	0%	0%	0%	0%	4%
Depth (m)	0.06	0.20	0.28	0.11	0.09	0.11	0.07	0.13
Width (m)	3.10	4.77	9.20	9.10	8.72	4.10	1.79	5.64
Bankfull width (m)	6.96	4.87	20.79	17.82	13.51	7.79	4.55	9.92
Bankfull Height (m)	0.45	0.47	0.97	1.46	1.22	1.24	0.73	0.63

140

Table 6-2. Statistics describing inter-annual variability in taxon densities and occurrence over four years in minimally-impacted New Hampshire streams. Co-occurrences refers to the number annual samples within a stream reach in which a taxon was detected; as there were four annual samples, the maximum number of co-occurrences if the taxa was always present was four. The overall mean co-occurrences was based only on taxa that were detected at least once in a stream reach. C.V. refers to the inter-annual coefficient of variation in densities.

Taxon	Nelson Brook			Oyster River			Lovell River		
	Mean	C.V.	Co-occurrences	Mean	C.V.	Co-occurrences	Mean	C.V.	Co-occurrences
Ictalurus nebulosus	0.0	-	0	0.1	2.000	1	0.1	2.000	1
Catostomus commersoni	0.0	-	0	0.0	-	0	0.6	1.641	2
Rhinichthys atratulus	2.3	2.000	1	0.0	-	0	3.8	1.245	2
Salvelinus fontinalis	35.7	0.800	3	0.1	2.000	1	1.3	1.350	2
Cottus cognatus	7.7	0.877	3	0.0	-	0	1.5	1.701	2
Phoxinus eos	0.0	-	0	0.0	-	0	0.2	2.000	1
Rhinichthys cataractae	0.0	-	0	0.0	-	0	0.1	2.000	1
Salmo gairdneri	0.0	-	0	0.0	-	0	0.1	2.000	1
Anguilla rostrata	0.0	-	0	0.1	2.000	1	0.0	-	0
Oligochaeta	4.2	0.791	3	64.7	1.312	4	32.0	1.277	4
Hirudinidae	0.3	2.000	1	0.0	-	0	0.0	-	0
Glossiphonidae	0.3	2.000	1	0.0	-	0	0.0	-	0
Gammaridae	0.0	-	0	440.6	0.387	4	0.0	-	0
Heptageniidae	4.2	1.330	2	0.0	-	0	6.5	0.399	4
Siphlonuridae	2.5	2.000	1	0.0	-	0	5.6	0.685	3
Baetidae	16.3	0.444	4	0.0	-	0	8.5	0.724	4
Leptophlebiidae	110.0	0.802	4	0.0	-	0	15.6	1.081	4
Ephemerellidae	41.5	0.522	4	0.0	-	0	5.0	0.771	4
Nemouridae	4.1	1.461	2	0.0	-	0	0.0	-	0
Perlodidae	5.7	0.745	3	0.0	-	0	0.0	-	0
Perlidae	6.3	1.128	3	0.0	-	0	2.0	1.322	3
Leuctridae	14.2	0.624	4	0.0	-	0	12.2	0.965	4
Rhyacophilidae	4.4	0.729	3	0.0	-	0	0.3	2.000	1
Chloroperlidae	55.8	0.673	4	0.9	2.000	1	12.9	0.615	4
Hydropsychidae	4.5	0.825	3	4.7	1.200	2	2.7	1.190	2
Hydroptilidae	3.3	1.788	2	0.0	-	0	0.0	-	0
Glossosomatidae	1.7	1.169	2	0.0	-	0	2.4	1.468	2
Brachycentridae	17.5	1.095	4	0.0	-	0	11.5	1.596	2

Table 6-2. Continued

Taxon	Nelson Brook			Oyster River			Lovell River		
	Mean	C.V.	Co-occurrences	Mean	C.V.	Co-occurrences	Mean	C.V.	Co-occurrences
Phryganeidae	0.8	2.000	1	0.0	-	0	0.0	-	0
Leptoceridae	0.0	-	0	0.9	2.000	1	0.3	2.000	1
Philopotamidae	5.3	0.724	3	0.0	-	0	2.4	1.169	2
Odontoceridae	5.5	0.855	4	0.0	-	0	0.0	-	0
Lepidostomatidae	20.0	1.360	3	0.0	-	0	0.0	-	0
Polycentropidae	8.0	1.827	2	0.0	-	1	1.7	1.066	3
Limnephilidae	9.2	1.507	4	0.9	2.000	1	2.1	1.414	2
Sialidae	2.5	2.000	1	1.9	2.000	3	0.5	2.000	1
Corydalidae	0.0	-	0	6.6	0.719	1	14.2	1.142	4
Dytiscidae	0.3	2.000	1	0.9	2.000	0	0.6	1.187	2
Elmidae	28.6	0.724	4	0.0	-	0	3.4	0.595	4
Haliplidae	0.8	2.000	1	0.0	-	0	0.0	-	0
Psephenidae	1.3	1.187	2	0.0	-	0	0.0	-	0
Empididae	1.3	2.000	1	0.0	-	2	0.5	2.000	1
Tipulidae	23.6	0.658	4	1.9	1.155	4	12.1	1.247	4
Simuliidae	17.0	0.786	4	222.2	0.517	4	2.5	0.773	3
Chironomidae	727.7	0.824	4	264.4	0.448	2	186.2	0.289	4
Ceratopogonidae	8.3	0.673	3	4.7	1.200	1	45.0	0.537	4
Ancylidae	0.0	-	0	0.9	2.000	0	0.0	-	0
Physidae	0.3	2.000	1	0.0	-	0	0.0	-	0
Planorbidae	0.3	2.000	1	0.0	-	2	0.3	2.000	1
Sphaeridae	0.0	-	0	7.5	1.683	1	1.0	1.390	2
Collembola	0.0	-	0	0.9	2.000	0	0.0	-	0
Cordulegastridae	0.8	2.000	1	0.0	-	0	1.3	1.186	2
Cordulidae	0.0	-	0	0.0	-	0	0.0	-	0
Gomphidae	1.3	2.000	1	0.0	-	2	0.0	-	0
Aeshnidae	0.0	-	0	2.8	1.277	0	0.4	2.000	1
Calopterygidae	0.0	-	0	0.0	-	2	0.5	2.000	1
Turbellaria	0.0	-	0	6.6	1.641	1	0.0	-	0
Nematomorpha	0.0	-	0	2.8	2.000	0	0.0	-	0
Hydrachnida	7.0	1.278	3	0.0	-	0	0.0	-	0
Pyralidae	0.3	2.000	1	0.0	-	0	0.4	2.000	1
Pteronarcyidae	0.3	2.000	1	0.0	-	0	0.0	-	0

142

Table 6-2. Continued.

Taxon	Nelson Brook			Oyster River			Lovell River		
	Mean	C.V.	Co-occurrences	Mean	C.V.	Co-occurrences	Mean	C.V.	Co-occurrences
Nymphomyiidae	0.0	-	0	0.9	2.000	1	0.3	2.000	1
Strationyidae	0.5	2.000	1	0.0	-	0	0.0	-	0
Uenoidae	1.3	2.000	1	0.0	-	0	0.0	-	0
UnknownTrichoptera	6.1	1.057	3	0.0	-	0	0.5	2.000	1
UnknownOther	1.3	2.000	1	0.0	-	0	0.0	-	0
Asellidae	0.0	-	0	336.6	0.601	4	0.0	-	0
Anabaena	0.0	-	0	0.0	2.000	1	0.0	-	0
Lyngbya	8.1	2.000	1	0.0	-	0	0.0	-	0
Oscillatoria	19.3	1.930	2	0.0	-	0	3.0	1.572	4
Cymbelloid diatom	0.0	-	0	0.1	2.000	1	0.0	-	0
Eunotia	4.7	1.944	2	14.7	1.104	3	0.0	-	0
Gomphonema	0.7	2.000	1	0.0	-	0	0.0	-	0
Naviculoid diatoms	1.6	2.000	1	12.3	0.812	3	26.8	0.937	4
Tabellaria	3.2	0.877	3	0.2	2.000	1	99.4	0.981	4
Gyrosigma	0.0	-	0	0.0	-	0	0.4	2.000	1
Batrachospermum	0.0	-	0	0.0	-	0	21.4	2.000	1
Bulbochaete	0.0	-	0	0.0	-	0	6.1	0.936	3
Closterium	0.0	-	0	3.1	0.262	4	0.0	-	0
Cosmarium	0.0	-	0	0.0	-	0	4.4	1.988	2
Cylindrocapsa	0.0	-	0	0.0	-	0	1.6	2.000	1
Cylindrocystis	0.0	-	0	0.0	-	0	15.4	2.000	1
Hyalotheca	12.8	2.000	1	0.0	-	0	0.0	-	0
Mougeotia	5.0	1.630	2	52.3	0.746	4	295.7	1.509	3
Plectonema	0.5	2.000	1	0.0	-	0	0.0	-	0
Rhizoclonium/Cladaphora	22.5	1.724	2	4.5	2.000	1	4.9	0.870	4
Scenedesmus	0.0	-	0	0.0	-	0	0.9	2.000	1
Arthrodesmus	0.0	-	0	0.0	-	0	1.6	2.000	1
Ulothrix	0.0	-	0	50.2	0.897	3	3.9	2.000	1
Overall mean	**14.4**	**1.437**	**2.241**	**16.8**	**1.469**	**2.029**	**9.9**	**1.443**	**2.304**
Overall standard deviation	**77.4**	**0.567**	**1.189**	**68.0**	**0.615**	**1.218**	**38.1**	**0.533**	**1.249**

143

Table 6-3. Mean inter-annual Bray-Curtis similarities of taxon abundances and Sorenson's coefficient of community of taxon occurrences for each taxonomic group by stream reach.

		Bray-Curtis Similarity	Sorenson's CC
All Taxa	Nelson Brook	0.571	0.622
	Oyster River	0.679	0.580
	Lovell River	0.587	0.652
Vertebrates	Nelson Brook	0.400	0.433
	Oyster River	0.000	0.000
	Lovell River	0.133	0.153
Macroinvertebrates	Nelson Brook	0.651	0.671
	Oyster River	0.703	0.586
	Lovell River	0.666	0.713
Periphyton	Nelson Brook	0.122	0.256
	Oyster River	0.616	0.643
	Lovell River	0.479	0.676

144

Table 6-4. Spearman rank correlations and standard deviation (s) in the squared coefficients by taxonomic group between annual samples of three minimally-impacted streams in New Hampshire.

		Mean R^2	s
All Taxa	Nelson Brook	0.510	0.088
	Oyster River	0.567	0.128
	Lovell River	0.558	0.078
Vertebrates	Nelson Brook	0.840	0.139
	Oyster River	-0.189	0
	Lovell River	-0.070	0.445
Macroinvertebrates	Nelson Brook	0.581	0.133
	Oyster River	0.590	0.125
	Lovell River	0.660	0.035
Periphyton	Nelson Brook	0.117	0.259
	Oyster River	0.591	0.286
	Lovell River	0.530	0.222

145

Table 6-5. Bray-Curtis similarities between samples before and after a summer spate by taxonomic group in five minimally-impacted streams in New Hampshire.

	All Taxa	Vertebrates	Macroinvertebrates	Periphyton
Wonalancet	0.660	0.974	0.692	0.000
Moriah	0.689	0.950	0.725	0.245
Mad	0.582	0.953	0.643	0.071
Stony	0.500	0.786	0.493	0.322
Smarts	0.541	0.791	0.563	0.351

146

Figure 6-1. Stream flow in the Pemigewasset River of central New Hampshire showing the large spate during August of 2003 due to a period of heavy rainfall provided by A U.S. Geological Survey stream gauge in Lincoln, New Hampshire. The spate (see arrow) resulted in stream flows higher than the spring snowmelt. The previous August (2002) is more typical of August flow in the region.

USGS 01074520 EAST BRANCH PEMIGEWASSET RIVER AT LINCOLN, NH

EXPLANATION

—— DAILY MEAN STREAMFLOW   —— ESTIMATED STREAMFLOW

147
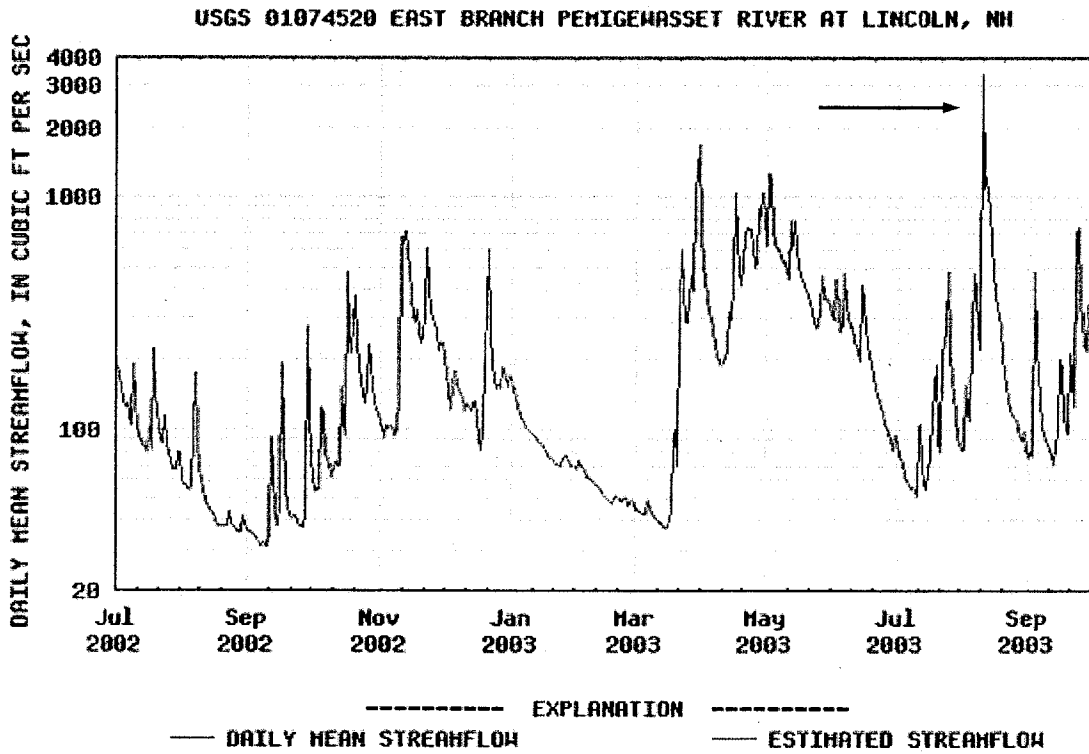
Figure 6-2. Cumulative contribution of taxa, ranked by relative abundance within each taxonomic group, to the total of the coefficients of variation (C.V.) within three minimally-impacted stream reach in New Hampshire. The coefficients of variation are based on the inter-annual variation in taxon abundances over four years of sampling.

148
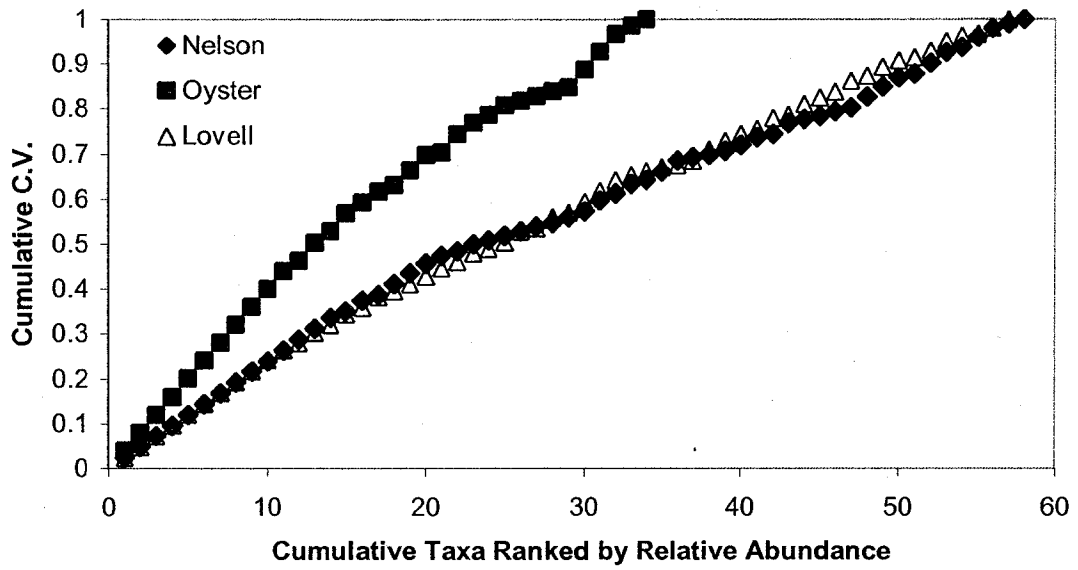
Figure 6-3. This figure presents a conceptual model of the responses of minimally-impacted stream communities to summer base-flow disturbance that may account for the observed differences in community stability for each taxonomic group in response to short-term flow disturbance and inter-annual sampling. Relative abundance refers to the abundance before a spate, though the vertical axis could also be interpreted as community similarity as well.

149
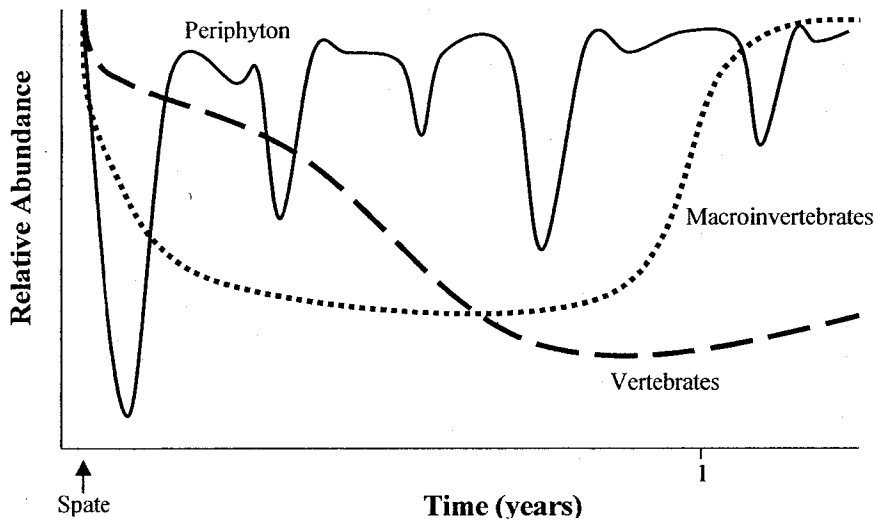
CONCLUSIONS

Minimally-impacted stream assemblages, encompassing periphyton, macroinvertebrates, stream-dwelling salamanders, and fish, in New Hampshire are primarily influenced by local-scale physical habitat conditions, especially benthic substrate characteristics and mesohabitat (e.g. pool, riffle, etc.; Figures 3-3 and 3-4). Stream benthic substrate and mesohabitats are primarily controlled by water flow (Minshal 1984, Naiman and Decamps 1997, Skinner 2003). In turn, hydrological conditions primarily vary along the longitudinal stream profile (Vannote et al. 1980). Thus, it seems that the longitudinal stream profile has a central influence on stream organism distributions and abundances in New Hampshire absent major human influence. Many other studies have found hydrological conditions and substrate to heavily influence stream organism ecology (see reviews by Allan 1975, Winget and Magnum 1991, Allan 1995).

Seven community types best described the pattern of variation in stream organism distributions in New Hampshire. As stream organisms individually varied in response to the longitudinal stream profile, so too did the community types arrange along a gradient from high elevation, steep, coarse substrate scour streams to flat, fine sediment, wetland streams (Chapter 3). The strong organizing influence of the stream longitudinal profile places constraints on the ability of geographic classifications of stream to explain organism distributions. Neither the aquatic ecoregions of Omernik (1987) nor a classification of New Hampshire's USGS HUC10 watersheds by The Nature Conservancy explained organism distributions even half as well as the biotic classification of streams (Chapter 3). By encompassing much of the range in variation of the primary controlling process in New Hampshire stream ecology, the changes inherent in the longitudinal stream profile, watersheds will fail to separate biological variation in stream unless

150

they are very small. Indeed, aquatic ecoregions explained stream organisms distributions and abundances better than the HUC10 watershed classification (Chapter 3). Thus, resource management and the assessment of representation of stream biological diversity on conservation lands should be performed within the framework of the biotic community types and the longitudinal stream profile.

Both individual stream organisms and the general community types were predictable using local-scale and basin-scale environmental descriptors (Tables 3-6, 3-7, and 4-2). In addition, the taxon densities of the minimally-impacted streams sampled for this dissertation could also be predicted using environmental descriptors that were determined to be robust to human influence (Table 4-2). Thus, reference conditions for biological assessment and monitoring of other streams in New Hampshire can be accurately constructed and tailored to each stream segment that is to be assessed for the level of human alteration to the biotic assemblage relatively independent of unexplained natural variation in community structure (Chapter 4). While several statistical approaches could be used to predict reference conditions for a particular stream segment, separate habitat regression models for each taxon most accurately predicted stream community structure in the reference streams sampled in this dissertation.

Physical stream and riparian habitat in reference streams was also predictable using large-scale physical environment descriptors such as watershed area and elevation (Chapter 5; Table 5-??). In addition, a new index for comparing observed physical stream habitat and expected minimally-impacted habitat conditions based on the Euclidean distance between the standardized expected and observed habitat descriptors. This new index improved on previous attempts to index the degree of habitat alteration by eliminating the need to construct discreet pseudo-classes of habitats to predict that often resulted in multiple mutually exclusive habitat conditions predicted for a stream in the absence of human influence (Davies et al. 2000). An assessment of stream habitat condition parallel to a biological assessment would provide powerful information in identifying the stressor responsible for an altered community structure. A lack of habitat

alteration as measured by the habitat assessment index would rule out habitat alteration

potentially shift emphasis to chemical sources.

Cautions

The accuracies of both the biotic and habitat prediction approaches were not verified on

an independent dataset of minimally-impacted streams and likely overestimate the accuracy as a

result of potential model over-fitting to the training dataset. The accuracies of all prediction

models presented in this dissertation should be tested on an independent set of minimally-

impacted New Hampshire streams. However, finding additional reference streams will be

difficult in the Seacoast area of the State as widespread development has greatly limited the

number of streams that fall within the definition of minimally-impacted conditions used in this

study.

However, the accurate prediction of reference conditions is only part of the conditions

that produce a sensitive biological assessment system. The degree to which an assessment

approach is sensitive to perturbations is also important and the ability of the biological and habitat

assessment approaches outlined in this dissertation should be tested. However, the usual and easy

approaches to measuring the sensitivity to impact of an assessment system are fraught with

circular logic as the only metric for assessing the degree of alteration to the biotic community is

often the prediction approach itself (Reynoldson et al. 1997). For example, one assessment

system may find a stream to be unimpacted and another assessment system may find it impacted.

Is the latter system more sensitive? If streams can differ in their response to similar levels of

anthropogenic stress, then it cannot be assumed that any individual stream should be impacted by

a particular level of human impact. To answer that question then, the prior community structure

of a stream would need to be known. Something akin to a controlled experiment manipulating

152

the level of alterations to a stream would need to be conducted on streams with known reference biological community structure; a sensitive system would detect impacts in this experiment at lower impact levels. However, assuming that stresses vary in their impact on stream biota, an additional treatment consisting of different impact types would also need to be included; multiplying the number of potential types of human impacts by a number of impact levels yields a very large experiment. That large experiment may also need to be run for years as some impacts may not show effects for multiple seasons. Also, three taxonomic groups were included in the predictive models generated in this dissertation; the taxonomic groups may differ in their sensitivity to various impacts. Clearly, the task of assessing the sensitivity to impact of a biological assessment approach without recourse to circular logic or tenuous assumptions is daunting.

Because the predictions of stream community structure and physical habitat derive from minimally-impacted streams, they can be used to establish restoration targets. However, there are a number of caveats attached to the use of these reference conditions for restoration targets. First, the impact of past land use on current stream condition even in stream that currently minimally-impacted cannot be overstated; land-use practices, primarily forest clearance of the majority of New Hampshire 150+ years ago, has left a legacy of alteration to stream habitat (Foster et al. ??) and possibly to stream assemblages, though the latter point is much less well known. Therefore, these reference conditions probably do not represent a stream's biotic assemblage or physical habitat completely absent of any human influence. For example, predictions of large woody debris (LWD) densities for a restoration target would probably underestimate LWD densities as past forest clearance often increased river flows, washed out LWD, and removed downed LWD that might have contributed to in stream LWD (??). Additionally, the minimally-impacted streams sampled were potentially influenced by regional anthropogenic environmental changes such as acid precipitation, nitrate deposition, exotic species, dams, and climate change (Carpenter et al. 1992). Future policy modifications that reduce those large-scale, ongoing environmental

153

changes by humans may require a re-assessment of stream reference conditions; otherwise, the reference conditions predicted for stream assemblages and habitat for assessment or restoration targets would no longer reflect the least altered stream ecosystems.

Another potential hindrance to interpreting predicted biotic reference conditions is annual temporal variation in minimally-impacted stream assemblages. When there is a deviation between the predicted and observed assemblages, annual temporal variation in reference conditions represents an alternative hypothesis to the inference that anthropogenic impacts have caused those deviations from predicted biotic assemblages. To the extent that annual variation in assemblages is typically greater than the prediction error of any predictive model, then annual variation represents the greatest constraint on the sensitivity of detecting alterations to stream assemblages. Inter-annual variation in three stream segments over four years, as measured by Bray-Curtis percent similarity, was high for fish, stream-dwelling salamanders, and periphyton (Table 6-3), but inter-annual variation in the macroinvertebrates were within the ranges of the Bray-Curtis percent similarity prediction errors displayed by the multivariate predictive models of biotic reference conditions (Table 4-2).

At this point, it does not seem that natural inter-annual variation in macroinvertebrate assemblages is limiting the sensitivity of any biological assessment using the predictive models described in Chapter 4; the vertebrates and periphyton may vary too much over an annual time scale in first to fourth order New Hampshire streams for use in bioassessment. However, only three stream segments were observed and over a short period of time, making this conclusion tentative. Additionally, as any analysis of temporal variation involves repeat sampling, sampling error is also a part of any estimate of temporal variation. The relative contributions of sampling error and natural temporal variability to the observed annual variability need to be decomposed before it is concluded that annual differences in assemblage are causing the observed temporal variability and not poor sampling methods; though, it is unlikely that the electroshocking

154

sampling for the fish was responsible for the high inter-annual variability because samples taken immediately before and after a spate showed very high community similarity (Table ??).

The water chemistry parameters that the habitat predictions report are based on only one summer measurement. Stream water chemistry of streams of all size in the northeastern United States is highly variable over very short time spans (Chetelat and Pick 2001). Therefore, the individual predictions of chemical concentrations for a particular stream may be inaccurate for a specific time period. Collectively, as the collection of measurements was made over the course of four summers, the dataset may indicate general concentration ranges for all minimally-impacted streams in New Hampshire; those ranges are simply not tailored to any particular stream's conditions and therefore may display higher unexplained spatial variability than other habitat parameters. More intensive temporal sampling of stream chemistry in the minimally-impacted stream segments sampled as part of this dissertation might improve prediction of the expected reference range in chemical concentrations for other streams to be assessed for human impacts.

Future Work

The data and analyses presented in the previous two chapters represent a step forward in understanding the natural ecology of streams in New Hampshire. This hopefully will translate into a better understanding of anthropogenic alterations to stream ecology and ultimately ease resource management and conservation decisions. As always, avenues for additional work remain.

Although seven community types were chosen as the best description of stream organism distributions in New Hampshire, classifications that split the stream segments into additional groups also found substantial distinct biological differences between the groups (Table 3-2). Additional distinct natural community types may still exist. In particular, there may be rare

155

community types nested within the described seven community types (Chapter 3) that were not encountered often enough to present a detectably distinct group in classification. Targeted sampling of additional minimally-impacted streams within the community types may find rare sub-types. The warm-water riffle community type showed substantial variation in taxon densities (Table 3-3) and as a result is a good candidate to investigate sub-types; sites along the coast that contain American eels (*Anguilla rostrata*) seemed particularly distinct from other stream segments within that community type (pers. obs.). At the time of this dissertation, several dam removal projects are scheduled in the Seacoast region of the State that would allow greater access to streams by anadromous vertebrates; their effects on the warm-water riffle communities that dominate the southeastern portion of New Hampshire should be investigated and tracked using a BACI (before-after control impact) design.

Ease of implementation is always a concern in applied research. Several statistical models were presented including models to construct biotic reference conditions, construct habitat reference conditions, and map community locations using available GIS data. Those models involved either elaborate or tedious calculations. Automated computer algorithms that could translate the required parameters for each model into predictions for biological or habitat assessment would make it much more likely that these models would be used by State agencies, environmental consultants, and water quality volunteer groups. Detailed and relevant information output and summary would also be crucial to ensuring that the model results are understood and placed in the proper context.

Volunteers are playing an increasing role in water quality assessment and monitoring, particularly for basic chemical parameters and macroinvertebrate bioassessments (USEPA 1997). Making the prediction of reference conditions easily accessible and automated is particularly important for volunteers who are unlikely to be skilled in statistical computing. There are several additional important considerations involved with translating this work into a volunteer program. The predictive models generated in this dissertation were based on 500-organism fixed-counts for

156

the macroinvertebrates and identification to family-level in the laboratory. Live-sorting of fewer organisms by volunteers many not accurately compare with the model predictions or be sensitive to human impacts (Doberstein et al. 2000). Additionally, some families are difficult to distinguish in the field. Whether the predictions can be collapse up to order level and still be sensitive at detecting human impacts is uncertain; many studies indicate little information is lost at the family level versus genus-species, but higher taxonomic levels (order, class) often display a substantial loss of information (Chapter 1). It also takes substantial resources to train volunteers to accurately identify macroinvertebrates. It may actually be less costly overall to send volunteer collected samples to contract biological laboratories to perform similar counts to those performed in this dissertation as annually training a rotating cadre of volunteers to identify organisms if program staff time and salaries were included in the calculation. Additionally, the use of contract labs makes quality assurance requirements easier as most contract labs follow documented and detailed quality assurance and control procedures. For the predictive models presented here, then, volunteers would collect the habitat information needed to predict reference conditions for biological or habitat predictions and assessments; volunteers may find that habitat is easier to measure and more enjoyable, thereby increasing volunteer retention and further reducing program training needs. The periphyton are particularly ill suited to identification by volunteers as microscopes, dyes, blenders, and detailed biovolume calculations are needed beyond identification skills. Again however, volunteers can easily collect periphyton samples for which the identification and quantification for biovolume can be sub-contracted to a commercial biological lab.

When a volunteer program might lead to enforcement action, a period of parallel assessments by professionals to measure the degree of correspondence between the volunteer and professional assessments would be required (USEPA 1997). However, volunteers might be included as a screening step before any decision on punitive or remedial action by a State agency. Thus, although it is unlikely that a volunteer macroinvertebrate bioassessment program based on

157

higher taxonomic levels or fewer organisms counted would be as sensitive as the 500-count version presented in this dissertation (Chapter 4), a reduced count or higher taxonomic level identification, may be sensitive enough to detect gross impairments in a low-cost screening program. A volunteer-based screening effort would target professional assessment efforts using the higher 500 fixed-count family level predictive model bioassessment and make monitoring of stream biotic integrity more efficient.

Biological assessments are being actively pursued in other ecosystem types (Gerritsen et al. 1998, Danielson 1998, Jameson et al. 1998, Gibson et al. 2000). As biological assessment has been best developed in streams, there are several findings that may be translated to these efforts that might reduce development time. First, it is clear that every assessment of a site is a small scientific project subject to the usual rules of scientific inference; alternative hypotheses as to the causes of any deviation in expected condition need to be ruled out before human alteration can be inferred to be the cause (Karr and Chu 1999). The reference condition concept (Hughes 1995) is central to eliminating the hypothesis that natural spatial or temporal variation is the cause of any observed alterations to biological condition.

Second, the ability of the reference conditions to explain natural spatial and temporal variation seems primarily constrained by the precision and accuracy of the sampling methods (Karr and Chu 1999, Doberstein et al. 2000, Cao et al. 2002, Ostermiller and Hawkins 2004) and the ability of statistical models to explain and control for natural variation (Reynoldson et al. 1997). Many attempts to improve the prediction of reference conditions suing sophisticated alternative multivariate statistical analyses have often resulted in little improvement for stream ecosystems compared to those developed 20 years ago such as RIVPACS (Moss et al. 1999, Hawkins et al. 2000, Linke et al. 2002, Linke et al. 2005). Yet, studies of the impact of sub-sampling on macroinvertebrate assessments consistently find major improvements in assessment accuracy due to higher sampling intensity when an adequate range in sampling intensity is

158

examined (Doberstein et al. 2000). Thus, it seems that substantial attention should be paid to the precision and accuracy of sampling results

Third, the taxonomic groups found in streams often respond differently to various impacts (Karr 1991, Metcalfe-Smith 1996, Karr and Chu 1999). As most ecosystems contain multiple taxonomic groups, the widest possible range in taxonomic groups should be included in initial explorations into biological assessments of other ecosystem types to determine which groups are complementary and sensitive to detecting impacts. Similarly, sampling from multiple habitat types has also improved the impact sensitivity of stream bioassessments (Kerans et al. 1992) because habitats are differentially sensitive to anthropogenic impacts and contain different species (e.g. Bradley and Ormerod 2002); thus a wide range of within-ecosystem habitat types should also be investigated early in the development process.

The variety and potential number of anthropogenic impacts to ecosystems seem to be increasing, particularly from relatively new chemical sources such as nanotechnology and pharmaceuticals. Additionally, the increasing global population is increasing resource utilization and stress on ecosystems. Much of human welfare ultimately derives from ecosystem good and services (Costanza et al. 1997) and those services may be linked to the biological condition of those systems (Loreau et al. 2002). Thus, close attention should be paid to the biological integrity of ecosystems to ensure their sustainability. Biological assessment and monitoring will play a substantial role in ensuring the sustainability of our natural resources.

159

LITERATURE CITED

Introduction

Christensen, N.L., et al. 1996. The report of the Ecological Society of America Committee on the Scientific Basis for Ecosystem Management. *Ecological Applications* 6:665-691.

Costanza, R. *et al.* 1997. The value of the World's ecosystem services and natural capital. *Nature* 387: 253-260.

Franklin, J. 1995. Predictive vegetation mapping: Geographic modelling of biospatial patterns in relation to environmental gradients. *Progress in Physical Geography* 19:474–499.

Grossman, D. H., D. Faber-Langendoen, A. S. Weakley, M. Anderson, P. Bourgeron, R. Crawford, K. Goodin, S. Landaal, K. Metzler, K. D. Patterson, M. Pyne, M. Reid, and L. Sneddon. 1998. International classification of ecological communities: terrestrial vegetation of the United States. Volume I. The National Vegetation Classification System: development, status, and applications. The Nature Conservancy: Arlington, Virginia, USA.

Grumbine, R.E. 1994. What is ecosystem management? *Conservation Biology* 8:2738.

Loreau, M., S. Naeem,. and P. Inchausti (Eds.). 2002. Biodiversity and Ecosystem Functioning: Synthesis and Perspectives. Oxford University Press: Oxford, United Kingdom.

Noss, R.F. 1987. From plant communities to landscapes in conservation inventories: A look at The Nature Conservancy (USA). *Biological Conservation* 41:11-37.

Noss, R.F. 1990. Indicators for monitoring biodiversity: A hierarchical approach. *Conservation Biology* 4: 355-364.

Schmid, B. 2002. The species richness – productivity controversy. *Trends in Ecology and Evolution* 17: 113-114

Chapter I

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.

Bowman, M.F., and R.C. Bailey. 1997. Does taxonomic resolution affect the multivariate description of the structure of freshwater benthic macroinvertebrate communities? *Canadian Journal of Fisheries and Aquatic Sciences* 54:1902-1807.

Clesceri, L.S., A. E. Greenberg, R. R. and Trussell, (eds.), 1989. Standard Methods for the Examination and Treatment of Water and Wastewater: 17th Edition. Washington, D.C.

160

Complex Systems Research Center. 2001. Hydrography Dataset. University of New Hampshire, Durham, New Hampshire. Accessible at: http://www.granit.sr.unh.edu/cgibin/nhsearch?dset=nhhd/nh.

Doberstein, C. P., J. R. Karr, and L. L. Conquest. 2000. The effect of fixed-count subsampling on macroinvertebrate biomonitoring in small streams. *Freshwater Biology* **44**: 355-371.

Hawkins, C.P., and R.H. Norris. 2000. Effects of taxonomic resolution and use of subsets of the fauna on the performance of RIVPACS-type models. Pages 217-228 *in* J.F. Wight, D.W. Sutcliffe, and M.T. Furse, editors. Assessing the Biological Quality of Fresh Waters: RIVPACS and other Techniques. Freshwater Biological Association: Ambleside, Cumbria, United Kingdom.

Hawkins, C.P., R.H. Norris, J. Gerritsen, R.M. Hughes, S.K. Jackson, R.K. Johnson, and R.J. Stevenson. 2000. Evaluation of the use of landscape classifications for the prediction of freshwater biota: synthesis and recommendations. *Journal of the North American Benthological Society* **19**: 541-556.

Hughes, R. M. 1995. Defining acceptable biological status by comparing with reference conditions. Pages 31-48 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, Florida, USA.

Lazorchak, J.M., D. J. Klemm, and D.V. Peck (Eds.). 1998. Environmental Monitoring and Assessment Program -Surface Waters: Field Operations and Methods for Measuring the Ecological Condition of Wadeable Streams. EPA/620/R-94/004F. U.S. Environmental Protection Agency, Washington, D.C.

Lenat, D. R. and V. H. Resh. 2001. Taxonomy and stream ecology – the benefits of genus and species-level identifications. *Journal of the North American Benthological Society* **20**:287-298.

Marchant, R., and G. Hehir. 2002. The use of AUSRIVAS predictive models to assess the response of lotic macroinvertebrates to dams in south-east Australia. *Freshwater Biology* **47**:1033-1050.

McCreadie, J.W., P.H. Adler, and J.F. Burger 1997. Species assemblages oflarval black flies (Diptera: Simuliidae): random or predictable? *Journal of North American Benthological Society* **16**, 760–770.

Merritt, R. W. and K. W. Cummins (eds.). 1996. An Introduction to the Aquatic Insects of North America: 3rd Edition. Kendall/Hunt Publ. Co., Dubuque, Iowa.

Moulton, S. R. II, J. L. Carter, S. A. Grotheer, T. F. Cuffney, and T. M. Short. 2000. Methods of Analysis by the U.S. Geological Survey National Water Quality Laboratory —Processing, Taxonomy, and Quality Control of Benthic Macroinvertebrate Samples. U.S. Geological Survey Open-File Report 00-212.

Olsgard F., P. J. Somerfield, and M. R. Carr. 1997. Relationships between taxonomic resolution and data transformations in analyses of a macrobenthic community along an established pollution gradient. *Marine Ecology Progress Series* **149**:173–181.

Omernik, J. M. 1987. Ecoregions of the conterminous United States. *Annals of the Association of American Geographers* **77**: 118-125.

Peckarsky, B. L., P. Fraissinet, M. A. Penton, and D. J. Conklin, Jr. 1990. Freshwater Macroinvertebrates of Northeastern North America. Cornell University Press, Ithaca, New York.

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.

Thorp, J.H. and A.P Covitch (eds.) 2001. Ecology and Classification of North American Freshwater Invertebrates: 2nd Edition. Academic Press, Burlington, Massachusetts.

Chapter II

Adamus, P.R. (1996) *Bioindicators for Assessing Ecological Integrity of Prairie Wetlands.* National Health and Environmental Effects Laboratory, U.S. Environmental Protection Agency, Corvallis, Oregon, EPA/600/R-96/082.

Alverson, A.J., Manoylov, K.M., & Stevenson, R.J. (2003) Laboratory sources of error for algal community attributes during sample preparation and counting. *Journal of Applied Phycology,* **15**, 357-369.

Barbour, M.T., Gerritsen, J., Snyder, B.D. & Stribling, J.B. (1999) *Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates and Fish,* 2nd edn. United States Environmental Protection Agency, Office of Water, Washington, DC, EPA 841-B-99-002.

Biggs, B.J.F. (1996) Benthic algal communities as biological monitors. *Algal Ecology: Freshwater Benthic Ecosystems* (eds. R.J. Stevenson, M.L. Bothwell, & R.L. Lowe), pp. 31-56, Academic Press, New York.

Biggs, B.J.F. & Kilroy, C. (2000) *Stream Periphyton Monitoring Manual.* NIWA, New Zealand Ministry for the Environment, Christchurch, New Zealand.

Canfield, R. H. (1941) Application of the line interception method in sampling range vegetation. *Journal of Forestry,* **39**, 388-394.

Danielson, T.J. (1998) *Indicators for monitoring and assessing biological integrity of inland, freshwater wetlands: a review of the technical literature (1989-1996).* United States Environmental Protection Agency, Washington, DC, EPA43-R-98-002.

Gerritsen, J., Carlson, R.E., Dycus, D.L., Faulkner, C., Gibson, G.R., Harcum, J. & Markowitz, S.A. (1998) *Lake and Reservoir Bioassessment and Biocriteria: Technical Guidance Document.* United States Environmental Protection Agency, Office of Water, Washington, DC, EPA 841-B-98-007

Gibson, G.R., Bowman, M.L., Gerritsen, J. & Snyder B.D. (2000) *Estuarine and Coastal Marine Waters: Bioassessment and Biocriteria Technical Guidance.* United States Environmental Protection Agency, Washington, DC, EPA 822-B-00-024.

Goodman, L.A. (1960) On the exact variance of products. *Journal of the American Statistical Association,* **55**, 708-713.

Graham, L.E. & Wilcox, W. (2000) *Algae.* Prentice Hall, Upper Saddle River, New Jersey.

Hill, B.H., Herlihy, A.T., Kaufmann, P.R., Stevenson, R.J., Mccormick, F.H. & Johnson, C.B. (2000) Use of periphyton assemblage data as an index of biotic integrity. *Journal of the North American Benthological Society,* **19**, 50-67.

Hillebrand, H., Dürlsen, C.-D, Kirschtel, D., Pollingher, U. & Zohary, T. (1999) Biovolume calculation for pelagic and benthic microalgae. *Journal of Phycology,* **35**, 403-424.

Horvitz, D.G. & Thompson, D.J.. (1952) A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association,* **77**, 89-96.

Jameson, S.C., Erdmann, M.V., Gibson Jr, G.R. & Potts, K.W. (1998) *Development of Biological Criteria for Coral Reef Ecosystem Assessment.* United States Environmental Protection Agency, Office of Science and Technology, Health and Ecological Criteria Division, Washington, D.C.

Kaiser, L. (1983) Unbiased estimation in line-intercept sampling. *Biometrics,* **39**, 965-976.

Kendall, M.G. & Moran, P.A.P. (1963) *Geometrical Probability*. Griffin, London.

Knapp, R. (1984) *Sampling methods and taxon analysis in vegetation science. Handbook of Vegetation Science 1*. Part 4. Junk, The Hague.

Krumbein, W.C. (1934) Size frequency distributions of sediments. *Journal of Sedimentary Petrology*, **4**, 65-77.

Lazorchak, J.M., Klemm, D.J. & Peck, D.V. (1998) *Environmental Monitoring and Assessment Program - Surface Waters: Field Operations and Methods for Measuring the Ecological Condition of Wadeable Streams*. United States Environmental Protection Agency, Washington, DC, EPA/620/R-94/004F.

Lowe, R.L. & Pan, Y. (1996) Benthic algal communities as biological monitors. *Algal Ecology: Freshwater Benthic Ecosystems* (eds. R.J. Stevenson, M.L. Bothwell, & R.L. Lowe), pp. 705-740, Academic Press, New York.

McCormick, P.V. & Cairns Jr., J. (1994) Algae as indicators of environmental change. *Journal of Applied Phycology*, **6**, 509-526.

Nedoma, J., Vrba, J., Hanzl, T. & Nedbalova, L. (2001) Quantification of pelagic filamentous microorganisms in aquatic environments using the line-intercept method. *Microbiology Ecology*, **38**, 81-85.

Pan, Y., Stevenson, R.J., Hill, B.H., Herlihy, A.T. & Collins, G.B. (1996) Using diatoms as indicators of ecological conditions in lotic systems: a regional assessment. *Journal of the North American Benthological Society*, 15, 481-495.

Porter, S.D., Cuffney, T.F., Gurtz M.E., Meador, M.R. (1993) *Methods for Collecting Algal Samples as Part of the National Water-Quality Assessment Program*. United States Geological Survey, Raleigh, North Carolina, Report 93-409.

Prescott, G.W. (1978) *How to Know the Freshwater Algae*. McGraw-Hill, Boston.

Ringold, P.L., Alegria, J., Czaplewski, R.L., Mulder, B., Tolle, T. & Burnett, K. (1996) Adaptive monitoring for ecosystem management. *Ecological Applications*, **6**, 745-747.

Stevenson, R.J. & Smol, J.P. (2002) Use of algae in environmental assessments. *Freshwater Algae in North America: Classification and Ecology* (eds J.D. Wehr & R.G. Sheath), pp. 775-804, Academic Press, San Diego, California.

Thompson, S.K. (1992) *Sampling*. Wiley-Interscience, New York.

Chapter III

Bradley, D.C. and S.J. Ormerod. 2002. Evaluating the precision of kick-sampling in upland streams for assessments of long-term change: the effects of sampling effort, habitat and rarity. *Archiv Fur Hydrobiologie* **155**: 199-221.

Clarke, R.T., M.T. Furse , R.J.M. Gunn , J.M. Winder , J.F. Wright. 2002. Sampling variation in macroinvertebrate data and implications for river quality indices. *Freshwater Biology* **47**: 1735-1751.

Danielson, T.J. 1998. Indicators For Monitoring And Assessing Biological Integrity of Inland, Freshwater Wetlands: A Review of the Technical Literature (1989-1996). United States Environmental Protection Agency, Washington, DC, EPA43-R-98-002.

163

Davis, W. S. and T. P Simon (Eds.). 1995. Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making. Lewis Publishers: Boca Raton, Florida.

Digby, P. G.N and R.A. Kempton. 1987. Multivariate Analysis of Ecological Communities. Chapman and Hall: New York, New York.

Faith, D.P., P.R. Minchin, and L. Belbin. 1987. Compositional dissimilarity as a robust measure of ecological distance. *Vegetation* **69**:57-68.

Gauch, H.G., Jr. 1982. Multivariate Analysis and Community Structure. Cambridge University Press, Cambridge, United Kingdom.

Gerritsen, J., R.E. Carlson., D.L. Dycus, C. Faulkner, G.R. Gibson, J. Harcum, and S.A. Markowitz. 1998. Lake and Reservoir Bioassessment and Biocriteria: Technical Guidance Document. United States Environmental Protection Agency, Office of Water, Washington, DC, EPA 841-B-98-007

Gibson, G.R., M. L. Bowman, J. Gerritsen, and B.D. Snyder. 2000. Estuarine and Coastal Marine Waters: Bioassessment and Biocriteria Technical Guidance. United States Environmental Protection Agency, Washington, DC, EPA 822-B-00-024.

Gower, J.C. 1971. A general coefficient of similarity and some of its properties. *Biometrics* **27**:857-872.

Hawkins, C.P. and M.R. Vinson. 2000. Weak correspondence between landscape classifications and stream invertebrate assemblages: implications for bioassessment. *Journal of the North American Benthological Society* **19**: 501-517.

Hawkins, C.P., R.H. Norris, J. Gerritsen, R.M. Hughes, S.K. Jackson, R.K. Johnson, and R.J. Stevenson. 2000. Evaluation of the use of landscape classifications for the prediction of freshwater biota: synthesis and recommendations. *Journal of the North American Benthological Society* **19**: 541-556.

Higgins, J.V., M.T. Bryer, M.L. Khoury, and T.W. Fitzhugh. 2005 A freshwater classification approach for biodiversity conservation planning. *Conservation Biology* **19**:432–445.

Hughes, R.M. 1995. Defining acceptable biological status by comparing with reference conditions. Pages 31-48 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, Florida, USA.

Jameson, S.C., M.V. Erdmann, G.R. Gibson, Jr, and K.W. Potts. 1998. Development of Biological Criteria for Coral Reef Ecosystem Assessment. United States Environmental Protection Agency, Office of Science and Technology, Health and Ecological Criteria Division, Washington, D.C.

Jongman, R. H. G., C. J. F. ter Braak, and O. F. R. Van Tongeren. 1995. Data Analysis in Community and Landscape Ecology. Cambridge University Press: Cambridge, United Kingdom.

Joy, M. K., and R.G. Death. 2002. Predictive modeling of freshwater fish as a biomonitoring tool in New Zealand. *Freshwater Biology* **47**:2261-2275.

Karr, J.R. 1991. Biological integrity: a long-neglected aspect of water resource management. *Ecological Applications* **1**: 66-84.

Karr, J.R. and E.W. Chu. 1999. Restoring Life in Running Waters - Better Biological Monitoring. Island Press, Covelo, California.

Kerans, B.L., J.R. Karr, and S.A. Ahlstedt. 1992. Aquatic invertebrate assemblages - spatial and temporal differences among sampling protocols. *Journal of the North American Benthological Society* **11**: 377-390.

164

Linke S, R.H. Norris, D.P. Faith, D. Stockwell. 2005. ANNA: A new prediction method for bioassessment programs. *Freshwater Biology* **50**: 147-158.

Marchant, R., A. Hirst, R. Norris, and L. Metzelling 1999. Classification of macroinvertebrate communities across drainage basins in Victoria, Australia: consequences of sampling on a broad spatial scale for predictive modelling. *Freshwater Biology* **41**: 253-268.

Martin, C.W. 1979. Precipitation and streamwater chemistry in an undisturbed forested watershed in New Hampshire. *Ecology* 60:36-42.

McCune, B. and M.J. Mefford. 1997. Multivariate Analysis of Ecological Data (PC-Ord for Windows, V. 3.06). MJM Software Design.

Merritt, R. W. and K. W. Cummins (eds.). 1996. An Introduction to the Aquatic Insects of North America: 3rd Edition. Kendall/Hunt Publ. Co., Dubuque, Iowa.

Metcalfe-Smith J.L. 1996. Biological water-quality assessment of rivers: use of macroinvertebrate communities. Pages 17-43 in Petts G.E., Calow P. (Eds.). River Restoration. Blackwell Science: Oxford, United Kingdom.

Minshall, G.W. and C.T. Robinson. 1998. Macroinvertebrate community structure in relation to measures of lotic habitat heterogeneity. *Archiv für Hydrobiologie* **141**:129-151.

Moss D., J.F. Wright, M.T. Furse, R.T. Clarke. 1999. A comparison of alternative techniques for prediction of the fauna of running-water sites in Great Britain. *Freshwater Biology* 41:167-181.

Mykra H, J. Heino, T. Muotka. 2004. Variability of lotic macroinvertebrate assemblages and stream habitat characteristics across hierarchical landscape classifications. *Environmental Management* **34**: 341-352.

Omernik, J.M. 1987. Ecoregions of the conterminous United States. *Annals of the Association of American Geographers* **77**: 118-125.

Prescott, G. W. 1978. How to Know the Freshwater algae, 3rd Edition. Wm. C. Brown Publishers: Dubuque, Iowa.

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.

Sandin, L. and R.K. Johnson. 2000. Ecoregions and benthic macroinvertebrate assemblages of Swedish streams. *Journal of the North American Benthological Society* **19**: 462-474.

SPSS for Windows, Rel. 14.0.1. 2005. Chicago: SPSS Inc.

TNC-SWI (The Nature Conservancy—Sustainable Waters Initiative). 2000. Tools for GIS analysis. The Nature Conservancy, Arlington, Virginia. Available from http://www.freshwaters.org/info/large/. (accessed February 2006).

Vannote, R. L., G. W. Minshall, K. W. Cummins, J. R. Sedell, and C. E. Cushing. 1980. The river continuum concept. *Canadian Journal of Fisheries and Aquatic Science* **37**: 130-137.

Vitousek, P.M. 1977. The regulation of element concentrations in mountain streams in the northeastern United States. *Ecological Monographs* 47:65-87.

Waite, I.R., A.T. Herlihy, D.P. Larsen, and D.J. Klemm. 2000 Comparing strengths of geographic and nongeographic classifications of stream benthic macroinvertebrates in the Mid-Atlantic Highlands, USA. *Journal of the North American Benthological Society* **19**: 429-441.

Warton, D.I. and H.M. Hudson. 2004. A MANOVA statistic is just as powerful as distance-based statistics for multivariate abundances. *Ecology* **85**:858–874.

Weigel, B.M., L. Wang, P.W. Rasmussen, J.T. Butcher, P.M. Stewart, T.P. Simon, M.J. Wiley. 2003. Relative influence of variables at multiple spatial scales on stream macroinvertebrates in the Northern Lakes and Forest ecoregion, U.S.A. *Freshwater Biology* **48**:1440-1461.

Chapter IV

APHA. 1992. Standard Methods for the Evaluation of Water and Wastewater, 18th ed. American Public Health Association: Washington, D.C.

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. U.S. Environmental Protection Agency; Office of Water: Washington, D.C.

Barbour, M.T., J.L. Plafkin, B.P. Bradley, C.G. Graves, and R.W. Wisseman. 1992. Evaluation of EPA's rapid bioassessment benthic metrics: Metric redundancy and variability among reference stream sites. *Environmental Toxicology and Chemistry* **11**:437-449.

Barbour, M.T., J.B. Stribling, and J.R. Karr. 1995. Multimetric approach for establishing biocriteria and measuring biological condition. Pages 63-77 *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, Florida.

Bruns D.A., G.W. Minshall, C.E. Cushing, K.W. Cummins, J.T. Brock, and R.L. Vannote. 1984. Tributaries as modifiers of the river continuum concept: analysis by polar ordination and regression models. *Archiv für Hydrobiologie*, **99**:208-220.

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

Cao, Y., D.D. Williams, and D.P. Larsen. 2002. Comparison of biological communities: the problem of sample representativeness. *Ecological Monographs* **72**:41-56.

Chessman, B.C. 1999. Predicting the macroinvertebrate faunas of rivers by multiple regression of biological and environmental differences. *Freshwater Biology*, **41**:747-757.

Clarke, R. 2000. Uncertainty in estimates of biological quality based on RIVPACS. Pages 39–54 in J.F. Wright, D.W. Sutcliffe, and M.T. Furse, eds. Assessing the Biological Quality of Fresh waters: RIVPACS and Other Techniques. Freshwater Biological Association: Ambleside, UK.

Corkum L.D. 1989. Patterns of benthic invertebrate assemblages in rivers of northwestern North America. *Freshwater Biology* **21**:191-205.

Davies, P.E. 2000. Development of a national river bioassessment system (AUSRIVAS) in Australia. Pages 113-124 in J.F. Wright, D.W. Sutcliffe, and M.T. Furse, eds. *Assessing the Biological Quality of Fresh waters: RIVPACS and Other Techniques.* Freshwater Biological Association: Ambleside, UK.

Davis, W.S. and T.P. Simons (Eds.). 1995. Biological Assessment and Criteria: Tools for Resource Planning and Decision Making. Lewis Publishers: Boca Raton, Florida.

Doberstein, C. P., J. R. Karr, and L. L. Conquest. 2000. The effect of fixed-count subsampling on macroinvertebrate biomonitoring in small streams. *Freshwater Biology* **44**: 355-371.

Faith, D.P., P.R. Minchin, and L. Belbin. 1987. Compositional dissimilarity as a robust measure of ecological distance. *Vegetation* **69**:57-68.

Hawkins, C.P. and M.R. Vinson. 2000. Weak correspondence between landscape classifications and stream invertebrate assemblages: implications for bioassessment. *Journal of the North American Benthological Society* **19**: 501-517.

Hawkins, C.P., R.H. Norris, J. Gerritsen, R.M. Hughes, S.K. Jackson, R.K. Johnson, and R.J. Stevenson. 2000. Evaluation of the use of landscape classifications for the prediction of freshwater biota: synthesis and recommendations. *Journal of the North American Benthological Society* **19**: 541-556.

Hawkins, C.P., R.H. Norris, J.N. Hogue, and J.W. Feminella. 2000. Development and evaluation of predictive models for measuring the biological integrity of streams. *Ecological Applications* **10**:1456-1477.

Hughes, R.M. 1995. Defining acceptable biological status by comparing with reference conditions. Pages 31-48 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, Florida, USA.

Hughes, R.M. and R.F. Noss. 1992. Biological diversity and biological integrity: Current concerns for lakes and streams. *Fisheries* **17**: 11-19.

Johnson, R.K. 2000. RIVPACS and alternative statistical modeling techniques: accuracy and soundness of principles. Pages 323-332 in J.F. Wright, D.W. Sutcliffe, and M.T. Furse, eds. Assessing the Biological Quality of Fresh waters: RIVPACS and Other Techniques. Freshwater Biological Association: Ambleside, UK.

Jones, D.L. 2002. Users manual for FATHOM: a MATLAB toolbox for Multivariate Ecological & Oceanographic Data Analysis. Available from: http://www.rsmas.miami.edu/personal/djones/

Jongman, R. H. G., C. J. F. ter Braak, and O. F. R. Van Tongeren. 1995. Data Analysis in Community and Landscape Ecology. Cambridge University Press: Cambridge, United Kingdom.

Karr, J. R. and M. Chu. 1999. Restoring Life in Running Waters: Better Biological Monitoring. Island Press: Washington, D. C., USA.

Karr, J.R. 1999. Defining and measuring river health. *Freshwater Biology* **41**:221-234.

Karr, J.R. and D.R. Dudley. 1981. Ecological perspective on water quality goals. *Environmental Management* **5**: 55-68.

Larsen, D.P., R.M. Hughes, P.L. Angermeier, T.M. Patton. 2002. Sampling effort affects multivariate comparisons of stream assemblages. *Journal of the North American Benthological Society* **21**:701-714.

Linke S., R.C. Bailey, and J. Schwindt. 1999. Temporal variability of stream bioassessments using benthic macroinvertebrates. *Freshwater Biology* **42**: 575-584.

Linke, S., R. Norris, and D. Faith. 2002. Australian River Assessment System: Improving AusRivAS Analytical Methods DDRAM and E-Ball (Phase I Final Report), Monitoring River Heath Initiative Technical Report no 27. Australian Museum, Sydney, Commonwealth of Australia, Canberra and University of Canberra, Canberra.

Linke, S., R.H. Norris, D.P. Faith, and D. Stockwell. 2005. ANNA: A new prediction method for bioassessment programs. *Freshwater Biology* **50**:147-158.

Marchant, R., A. Hirst, R. Norris, and L. Metzelling 1999. Classification of macroinvertebrate communities across drainage basins in Victoria, Australia: consequences of sampling on a broad spatial scale for predictive modelling. *Freshwater Biology* **41**: 253-268.

Marchant, R., A. Hirst, R.H. Norris, R. Butcher, L. Metzeling, and D. Tiller. 1997. Classification and prediction of macroinvertebrate communities from running waters in Victoria, Australia. *Journal of the North American Benthological Society* **16**:664-681.

McCune, B. and M.J. Mefford. 1997. Multivariate Analysis of Ecological Data (PC-Ord for Windows, V. 3.06). MJM Software Design.

Micheli F., K. Cottingham, J. Bascompte, O. Bjornstad, J. Eckert, J. Fischer, T. Keitt, B. Kendall, J. Klug, and J. Rusak. 1999. The dual nature of community variability. *Oikos* **85**:161-169.

Milner A.M., S.C. Conn, and L.E. Brown. 2006. Persistence and stability of macroinvertebrate communities in streams of Denali National Park, Alaska: implications for biological monitoring. *Freshwater Biology* **51**:373-387.

Noss, R.F. 1990. Indicators for monitoring biodiversity: A hierarchical approach. *Conservation Biology* **4**: 355-364.

Omernik, J.M. 1987. Ecoregions of the conterminous United States. *Annals of the Association of American Geographers* **77**: 118-125.

Ostermiller, J.D. and C.P. Hawkins. 2004. Effects of sampling error on bioassessments of stream ecosystems: application to RIVPACS-type models. Journal of the North American Benthological Society **23**:363-382

Philippi, T., P. Dixon, and B.L. Taylor. 1998 Detecting change and trends in species composition. *Ecological Applications* **8**:300-308.

Reynoldson, T.B., R.C. Bailey, K.E. Day, and R.H. Norris. 1995. Biological guidelines for freshwater sediment based on BEnthic Assessment of SedimenT (the BEAST) using a multivariate approach for predicting biological state. *Australian Journal of Ecology.* **20**:198-219.

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.

Rosenberg, D.M. and V.H. Resh (Eds.). 1993. Freshwater Biomonitoring and Benthic Macroinvertebrates. Chapman and Hall: New York.

Sandin, L. and R.K. Johnson. 2000. Ecoregions and benthic macroinvertebrate assemblages of Swedish streams. *Journal of the North American Benthological Society* **19**: 462-474.

Schaefer, J., K. Gido, M. Smith. 2005. A test for community change using a null model approach. *Ecological Applications* **15**:1761-1771.

SPSS for Windows, Rel. 14.0.1. 2005. Chicago: SPSS Inc.

Tonn, W.M. 1990. Climate change and fish communities: a conceptual framework. *Transactions of the American Fisheries Society* **119**: 337-352.

Van Sickle J., D.D. Huff, and C.P. Hawkins. 2006. Selecting discriminant function models for predicting the expected richness of aquatic macroinvertebrates. *Freshwater Biology* **51**:359-372.

Van Sickle, J., C.P. Hawkins, D.P. Larsen, and A.T. Herlihy. 2005.A null model for the expected macroinvertebrate assemblage in streams. *Journal of the North American Benthological Society*: **24**:178–191.

Vannote, R.L., G.W. Minshall, K.W. Cummins, J.R. Sedell, and C.E. Cushing. 1980. The river continuum concept. *Canadian Journal of Fisheries and Aquatic Science* 37:130-137.

Waite, I.R., A.T. Herlihy, D.P. Larsen, and D.J. Klemm. 2000 Comparing strengths of geographic and nongeographic classifications of stream benthic macroinvertebrates in the Mid-Atlantic Highlands, USA. *Journal of the North American Benthological Society* **19**: 429-441.

Walley, W.J. and V.N. Fontama. 2000. New approaches to river quality classification based upon Artificial Intelligence. Pages 263-280 in J.F. Wright, D.W. Sutcliffe, and M.T. Furse, eds. Assessing the Biological Quality of Fresh waters: RIVPACS and Other Techniques. Freshwater Biological Association: Ambleside, UK.

Warton, D.I. and H.M. Hudson. 2004. A MANOVA statistic is just as powerful as distance-based statistics for multivariate abundances. *Ecology* 85:858–874.

Weigel, B.M., L. Wang, P.W. Rasmussen, J.T. Butcher, P.M. Stewart, T.P. Simon, and M.J. Wiley. 2003. Relative influence of variables at multiple spatial scales on stream macroinvertebrates in the Northern Lakes and Forest ecoregion, U.S.A. *Freshwater Biology* 48:1440 – 1461.

Winterbourn, M.J., J.S. Rounick and B. Cowie. 1981. Are New Zealand ecosystems really different? New Zealand Journal of Marine and Freshwater Research 15:321-328.

Wolda, H. 1981. Similarity indices, sample size, and diversity. *Oecologia* 50:296-302.

Wright, J., D. Moss, P. Armitage, and M. Furse. 1984. A preliminary classification of running water sites in Great Britain based on macroinvertebrates and the prediction of community types using environmental data. *Freshwater Biology* 14: 221-256.

Wright, J.F., D.W. Sutcliffe, and M.T. Furse (Eds.). 2000. Assessing the Biological Quality of Fresh waters: RIVPACS and Other Techniques. Freshwater Biological Association: Ambleside, UK.

Chapter V

Allan, J.D. 1995. Stream Ecology : Structure and Function of Running Waters. Kluwer Academic Press: Boston, USA.

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. U.S. Environmental Protection Agency; Office of Water: Washington, D.C.

Barbour, M.T., J.B. Stribling, J.R. Karr. 1995. Multimetric approach for establishing biocriteria and measuring biological condition, p. 63-77 *in* Davis, W.S. and T.P. Simon, eds. Biological assessment and criteria--Tools for water resources planning and decision making: Lewis Publishers: Boca Raton, Florida.

Brierley, G., K. Fryirs, and T. Cohen. .1996. Development of a generic geomorphic framework to assess catchment character. Part 1. A geomorphic approach to catchment characterisation. Working Paper 9603, Macquarie University, Graduate School of the Environment.

Cao, Y., D.D. Williams, and D.P. Larsen. 2002. Comparison of biological communities: the problem of sample representativeness. *Ecological Monographs* 72:41-56.

169

Chapman, D. (Ed.). 1992. Water Quality Assessments. A Guide to the Use of Biota, Sediments and Water in Environmental Monitoring. Chapman and Hall: London, United Kingdom.

Cooper, S.D., Diehl, S., Kratz, K. and Sarnelle, O. 1998. Implications of scale for patterns and processes in stream ecology. *Australian Journal of Ecology* 23:27-40.

Cushing, C., C. McIntyre, J. Sedell, K. Cummins, G. W. Minshall, R. Peterson, and R. Vannote. 1995. Comparative study of physico-chemical variables in streams using multivariate analyses. *Archiv fur Hydrobiologie* 89: 343-352.

Davis, W. S. and T. P Simon (Eds.). 1995. Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making. Lewis Publishers: Boca Raton, Florida.

Fitzpatrick, F.A. and J.C. Knox. 2000. Spatial and Temporal Sensitivity of Hydrogeomorphic Response and Recovery to Deforestation, Agriculture, and Floods. *Physical Geography* 21:89-108.

Foster D., F. Swanson, J. Aber, I. Burke, N. Brokaw, D. Tilman, and A. Knapp. 2003 The importance of land-use legacies to ecology and conservation. *Bioscience* 53:77-88.

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

Gauch, H.G., Jr. 1982. Multivariate Analysis and Community Structure. Cambridge University Press, Cambridge, United Kingdom.

Hannaford, M.J. and V.H. Resh. 1995. Variability in macroinvertebrate rapid-bioassessment surveys and habitat assessment in a northern California stream. *Journal of the North American Benthological Society* 14: 430-439.

Harper, D. and Everard, M. 1998. Why should the habitat-level approach underpin holistic river survey management? *Aquatic Conservation: Marine and Freshwater Ecosystems* 8:395-413.

Hildrew, A. G. and P. S. Giller. 1994. Patchiness, species interactions, and disturbance in the stream benthos. Pages 21-62 in P. S. Giller, A. G. Hildrew, and D. G. Rafaelli, eds. Aquatic Ecology: Scale, Pattern, and Process. 34th Symposium of the British Ecological Society, Blackwell, Oxford, UK.

Hughes, R. M. 1995. Defining acceptable biological status by comparing with reference conditions. Pages 31-48 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, Florida, USA.

Hynes, H. B. 1970. The Biology of Polluted Rivers. Liverpool University Press: Liverpool, UK.

Hynes, H.B.N. 1975. The stream and its valley. *Internationale Vereinigung fur Theoretische und Angewandte Limnologie* 19: 1-15.

Jongman, R. H. G., C. J. F. ter Braak, and O. F. R. Van Tongeren. 1995. Data Analysis in Community and Landscape Ecology. Cambridge University Press: Cambridge, United Kingdom.

Karr, J.R. 1991. Biological integrity-A long-neglected aspect of water resource management: *Ecological Applications* 1:66-84.

Karr, J.R. and E.W. Chu. 1999. Restoring Life in Running Waters - Better Biological Monitoring. Island Press, Covelo, California.

170

Kaufmann, P. and B. Hughes. 2003. W-EMAP Quantitative River Physical Habitat Assessment. Presenatation to the National Biological Criteria and Assessment Workshop, 31 March, 2003.

Ladson, A.R., White, L.J., Doolan, J.A., Finlayson, B.L., Hart, B.T., Lake, P.S. and Tilleard, J.W. 1999. Development and testing of an Index of Stream Condition for waterway management in Australia. *Freshwater Biology* **41**: 453-468.

Lazorchak, J.M., D. J. Klemm, and D.V. Peck (Eds.). 1998. Environmental Monitoring and Assessment Program -Surface Waters: Field Operations and Methods for Measuring the Ecological Condition of Wadeable Streams. EPA/620/R-94/004F. U.S. Environmental Protection Agency, Washington, D.C.

Linke S, R.H. Norris, D.P. Faith, D. Stockwell. 2005. ANNA: A new prediction method for bioassessment programs. *Freshwater Biology* **50**: 147-158.

Linke S., R.C. Bailey, and J. Schwindt. 1999. Temporal variability of stream bioassessments using benthic macroinvertebrates. *Freshwater Biology 42:* 575-584.

Maddock, I. 1999. The importance of physical habitat assessment for evaluating river health. *Freshwater Biology* 41: 373-391.

Minshall, G.W. 1988. Stream ecosystem theory: a global perspective. *Journal of the North American Benthological Society* **7**: 263-288.

Norris, R.H. and Thoms, M.C. 1999. What is river health? *Freshwater Biology,* 41: 197-209.

Omernik, J.M. 1987. Ecoregions of the conterminous United States. *Annals of the Association of American Geographers* **77:** 118-125.

Petts and P. Calow (Eds.) 1996. River biota, diversity, and dynamics. Blackwell: Oxford, United Kingdom.

Rabeni, C.F. 2000. Evaluating physical habitat integrity in relation to the biological potential of streams. *Hydrobiologia* **422/423**:245-256.

Raven, P.J., N.T.H. Holmes, F.H. Dawson, and M. Everard. 1998. Quality assessment using River Habitat Survey data. *Aquatic Conservation: Marine and Freshwater Ecosystems* **8**:477-499.

Southwood, T. R. E. 1977. Habitat, the templet for ecological strategies? *Journal of Animal Ecology* **46**: 337-465.

Statzner, B. and B. Higler. 1986. Stream hydraulics as a major determinant of benthic invertebrate zonation patterns. *Freshwater Biology* **16**:127-139.

Vannote, R. L., G. W. Minshall, K. W. Cummins, J. R. Sedell, and C. E. Cushing. 1980. The river continuum concept. *Canadian Journal of Fisheries and Aquatic Science* **37**: 130-137.

Warton, D.I. and H.M. Hudson. 2004. A MANOVA statistic is just as powerful as distance-based statistics for multivariate abundances. *Ecology* **85**:858–874.

Weins, J.A. 1989. Spatial scaling in ecology. *Functional Ecology* **3**:385-397.

Wright, J.F., D.W. Sutcliffe, and M.T. Furse (Eds.). 2000. Assessing the Biological Quality of Fresh waters: RIVPACS and Other Techniques. Freshwater Biological Association: Ambleside, UK.

Bradt, P., M. Urban, N. Goodman, S. Bissell, and I. Spiegel. 1999. Stability and resilience in benthic macroinvertebrate assemblages. *Hydrobiologia*. 403:123-133

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

Chessman, B.C. 1999. Predicting the macroinvertebrate faunas of rivers by multiple regression of biological and environmental differences. *Freshwater Biology*, 41:747-757.

Connell, J.H., and W.P. Sousa. 1983. On the evidence needed to judge ecological stability or persistence. *American Naturalist* 121:789-824.

Davis, W. S. and T. P Simon (Eds.). 1995. Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making. Lewis Publishers: Boca Raton, Florida.

Doberstein, C. P., J. R. Karr, and L. L. Conquest. 2000. The effect of fixed-count subsampling on macroinvertebrate biomonitoring in small streams. *Freshwater Biology* 44: 355-371.

Faith, D.P., P.R. Minchin, and L. Belbin. 1987. Compositional dissimilarity as a robust measure of ecological distance. *Vegetation* 69:57-68.

Fausch, K. D., and R. G. Bramblett. 1991. Disturbance and fish communities in intermittent tributaries of a western Great Plains river. *Copeia* 1991:659-674.

Franklin, J. F. 1989. Importance and justification of long-term studies in ecology. Pages 3-19 in G. E.

Hoeinghaus, D.J., C.A. Layman, D.A. Arrington, and K.O. Winemiller. 2003. Spatiotemporal variation in fish assemblage structure in tropical floodplain creeks. *Environmental Biology of Fishes* 67:379-387.

Hughes, R. M. 1995. Defining acceptable biological status by comparing with reference conditions. Pages 31-48 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, Florida, USA.

Humphrey, C.L., A.W. Storey, and L. Thurtell. 2000. AUSRIVAS: operator sample processing errors and temporal variability - implications for model sensitivity. Pages 143-163 *in* Wright, J.F., D.W. Sutcliffe and M.T. Furse (eds.). Assessing the Biological Quality of Freshwaters. RIVPACS and Similar Techniques Freshwater Biological Association: Ambleside, United Kingdom.

Jongman, R. H. G., C. J. F. ter Braak, and O. F. R. Van Tongeren. 1995. Data Analysis in Community and Landscape Ecology. Cambridge University Press: Cambridge, United Kingdom.

Joy, M. K., and R. G. Death. 2002. Predictive modelling of freshwater fish as a biomonitoring tool in New Zealand. *Freshwater Biology* 47:2261-2275.

Karr, J.R. 1991. Biological integrity: a long-neglected aspect of water resource management. *Ecological Applications* 1: 66-84.

Karr, J.R. and E.W. Chu. 1999. Restoring Life in Running Waters - Better Biological Monitoring. Island Press, Covelo, California.

Linke S, R.H. Norris, D.P. Faith, D. Stockwell. 2005. ANNA: A new prediction method for bioassessment programs. *Freshwater Biology* 50: 147-158.

172

APPENDIX

UNIVERSITY OF NEW HAMPSHIRE INSTITUTIONAL ANIMAL CARE AND USE

COMMITTEE AGREEMENT

UNIVERSITY of NEW HAMPSHIRE

May 19, 2005

Eckert, Robert
Natural Resources
James Hall
Durham, NH 03824

IACUC #: 050402
Approval Date: 04/22/2005
Review Level: C
Project: A Classification of New Hampshire's Freshwater Stream Communities

The Institutional Animal Care and Use Committee (IACUC) reviewed and approved the protocol submitted for this study under Category C on Page 4 of the Application for Review of Vertebrate Animal Use in Research or Instruction - *the research potentially involves minor short-term pain, discomfort or distress which will be treated with appropriate anesthetics/analgesics or other assessments.*

Approval is granted for a period of three years from the approval date above. Continued approval throughout the three year period is contingent upon completion of annual reports on the use of animals. At the end of the three year approval period you may submit a new application and request for extension to continue this study. Requests for extension must be filed prior to the expiration of the original approval.

Please Note:
1. All cage, pen, or other animal identification records must include your IACUC # listed above.
2. Use of animals in research and instruction is approved contingent upon participation in the UNH Occupational Health Program for persons handling animals. Participation is mandatory for all principal investigators and their affiliated personnel, employees of the University and students alike. A Medical History Questionnaire accompanies this approval; please copy and distribute to all listed project staff who have not completed this form already. Completed questionnaires should be sent to Dr. Gladi Porsche, UNH Health Services.

If you have any questions, please contact either Van Gould at 862-4629 or Julie Simpson at 862-2003.

For the IACUC,

Robert G. Mair, Ph.D.
Chair

cc: File

Research Conduct and Compliance Services, Office of Sponsored Research, Service Building, 51 College Road, Durham, NH 03824-3585 * Fax: 603-862-3564

176