University of Alberta

Predicting Fish Assemblages in Small, Boreal Lakes in Alberta Using Landscape and Local Factors

by

Megan Rae MacLeod Conlon ©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment

of the requirements for the degree of Masters of Science

in

Environmental Biology and Ecology

Department of Biological Sciences

Edmonton, Alberta

Fall 2002

0-612-81384-3

University of Alberta

Library Release Form

Name of Author: Megan Rae MacLeod Conlon

Title of Thesis: Predicting Fish Assemblages in Small, Boreal Lakes in Alberta Using Landscape and Local Factors

Degree: Masters of Science

Year this Degree Granted: 2002

Permission is hereby granted to the University of Alberta Library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except as herein before provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.

Megan R Conlon

2835 Valley Dr.
Ann Arbor, MI 48103
August 08, 2002

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled *Predicting Fish Assemblages in Small, Boreal Lakes in Alberta Using Landscape and Local Factors* submitted by Megan Rae MacLeod Conlon in partial fulfillment of the requirements for the degree of Masters of Science in Environmental Biology and Ecology.

William M Tonn

Dr. William Tonn

Dr. Kevin Devito

Evelyn Merrill

Dr. Evelyn Merrill

Dr. Lee Foote

26 July 2002

ABSTRACT

I examined the effectiveness of lake order, a metric of landscape position, relative to landscape-level measures of surface connectivity and local factors at predicting the presence and types of fish assemblages and explaining variation in fish species richness in 102 small lakes in three regions of boreal Alberta. Fish presence in these lakes was best predicted in logistic regression models by isolation, a simple metric of surface connectivity. Once fish presence was established, maximum depth was most successful in predicting whether piscivore-dominated assemblages were present. Lake order was the best individual metric at explaining species richness among lakes in Poisson regression models. Thus, lake order, while not the best landscape-level metric for predicting fish assemblages, was best at explaining variation in species richness. These models can be used to predict how anthropogenic disturbance to hydrologic systems will affect fishes in boreal Alberta.

ACKNOWLEDGEMENTS

Financial support for the project was provided by an NSERC Operating Grant to W. M. Tonn, Challenge Grants in Biodiversity from Alberta Conservation Association, and a Circumpolar/Boreal Alberta Research Grant from the Canadian Circumpolar Institute. The Department of Biological Sciences, University of Alberta, and the NSERC Operating Grant to W. M. Tonn provided personal funding. Dave DeRosa of Alberta Sustainable Resource Development provided in-kind support for the fieldwork in the Utikuma Lake District.

I thank my supervisor, Dr. Bill Tonn, for helpful discussions and for patiently reading and providing excellent comments on drafts of proposals, grants, reports, my thesis, and the on-going manuscript. Hopefully, I become a better writer because of his valuable editorial skills. Members of the weekly Pasz-Tonn gathering listened to me present various iterations of my project and asked good questions at each stage. I thank Dr. Kevin Devito and members of his lab for allowing me access to research facilities in the Utikuma Lake District, for graciously sharing water chemistry data, and for discussions regarding the landscape position of lakes. Dr. Evelyn Merrill helpfully provided input on landscape ecology and modeling throughout the project. I also thank external examiner Dr. Lee Foote for being interested in my project.

I thank Mike Bonikowsky, the last of the Bonikowskys, for not dying or quitting on me through the tough field season. Charlene Popplewell, the GIS Technologist for the Department, helped me with the aspects of my project that required ArcView. Without her cheerful support, instructions, and encouragement, I would still be digitizing.

My officemates and friends, Allison Puchniak, Sarah Depoe, and Erica Heinlen, provided lots of laughter and help over the last 3 years.

Dr. Paul Webb piqued my interest in fish ecology, provided me with my first field research experience, and mentored me as an undergraduate at the University of Michigan. Dr. Jim Teeri allowed me to get my start as teaching assistant in fish ecology at the University of Michigan Biological Station.

Of course I thank my parents, Dave and Margie Conlon, for all their love and support. It is amazing how much love and good advice can be given over the telephone. I thank my sister, Erin, for visiting or talking to me when I needed to relax. Charlie, my chocolate lab, provided lots of entertainment and exercise.

And last, but by no means least, I thank Bruce McCulloch for captivating discussions of fish ecology and management, teaching me how to relax by fly fishing and hitting golf balls, and for all his love and support.

TABLE OF CONTENTS

ABSTRACT……………………………………………………………

ACKNOWLEDGEMENTS…………………………………………………

LIST OF TABLES…………………………………………………………

LIST OF FIGURES…………………………………………………………

INTRODUCTION………………………………………………………… 1
- I. Objectives and Hypotheses……………………………… 3

METHODS…………………………………………………………………… 4
- I. Study Area……………………………………………………… 4
- II. Data Sources…………………………………………………… 5
- III. Fish…………………………………………………………………… 6
- IV. Lake Morphometry………………………………………… 7
- V. Water Chemistry…………………………………………… 8
- VI. Landscape-level Metrics………………………………… 9
- VII. Data Analysis - Characteristics of Lakes…………… 10
- VIII. Data Analysis – Fish Assemblages…………………… 11
- IX. Data Analysis – Species Richness……………………… 14
- X. Model Specifications………………………………………… 14

RESULTS………………………………………………………………………… 15
- I. Characteristics of Lakes…………………………………… 15
- II. Fish Assemblages – Individual Data Sets……………… 16
- III. Fish Assemblages – Combined Data Set……………… 18
- IV. Species Richness……………………………………………… 19

DISCUSSION…………………………………………………………………… 19
- I. Fish Assemblage Type……………………………………… 19
- II. Species Richness……………………………………………… 23
- III. Lake Order: Landscape Position or Surface Connectivity?…… 24

CONCLUSION………………………………………………………………… 25

LITERATURE CITED………………………………………………………… 46

APPENDICES.. 52
- A. Data.. 52
  1. Fish Data.. 52
  2. Local Environmental Data.......................... 55
  3. Landscape Data....................................... 58
- B. National Topographic Maps List.............................. 61
- C. Factors Available for Each Data Set.......................... 64
- D. Model Specifications... 65
  1. Logistic Models of Fish Presence in UTK......... 65
  2. Logistic Models of Fish Presence in ATH......... 66
  3. Logistic Models of Piscivore Presence in ATH... 67
  4. Logistic Models of Piscivore Presence in SFM... 68
  5. Overall Logistic Models of Fish Presence......... 69
  6. Overall Logistic Models of Piscivore Presence... 71
  7. Poisson Models of Fish Species Richness......... 72

LIST OF TABLES

1. Fish Presence in UTK.. 28

2. Fish Presence in ATH.. 29

3. Piscivore Presence 30
    A. ATH
    B. SFM

4. Combined Data Set Models.. 31
    A. Fish Presence
    B. Piscivore Presence

5. Validation of Hierarchical Models.. 32

6. Fish Species Richness.. 33

LIST OF FIGURES

1. Four Factors Affect Fish Assemblages 34

2. Location of Study Sites in Alberta 35

3. Lake Order Classification 36

4. Isolation Classification 37

5. Principle Components Analysis of All Lakes 38
    - A. PCA Axis 2 by 1 with Data Set Highlighted 38
    - B. PCA Axis 3 by 2 with Data Set Highlighted 38
    - C. PCA Axis 2 by 1 by Fish Assemblage 39

6. Species Richess by Lake Order 40

7. Summary of Results 41

8. Lake Order 42
    - A. Surface Area by Lake Order 42
    - B. Maximum Depth by Lake Order 43
    - C. Total Phosphorous by Lake Order 44
    - D. pH by Lake Order 45

Introduction

Over the last 20 years, fish ecologists have identified four major factors that interact to influence the presence and composition of fish assemblages in small, temperate lakes. First, because fish must be able to immigrate into a lake, colonization routes are important (Jackson and Harvey 1989, Tonn *et al.* 1995, Snodgrass *et al.* 1996, Magnuson *et al.* 1998, Olden *et al.* 2001). Second, if there are colonization routes available, then there must be enough suitable habitat for fish to persist (Eadie and Keast 1984, Eadie *et al.* 1986, Tonn *et al.* 1990, Magnuson *et al.* 1998, Jackson *et al.* 2001). The amount of suitable habitat in lakes is influenced by local factors such as lake morphometry and water chemistry (Mathias and Barica 1980, Tonn and Magnuson 1982, Meding and Jackson 2001). Third, biotic interactions such as competition and predation can structure fish assemblages (Tonn and Magnuson 1982, Robinson and Tonn 1989, Chapleau *et al.* 1997, Findlay *et al.* 2000, MacRae and Jackson 2001). Fourth, human-mediated transfers of fish, such as accidental release, inter-basin water transfers, diversion channels, shipping canals, and stocking (Pratt *et al.* 1992, Chapleau *et al.* 1997, Ricciardi and Rasmussen 1998, Economidis *et al.* 2000, Findlay *et al.* 2000), can overcome the lack of colonization routes. This study assesses the ability of the first two factors to predict and explain the structure of fish assemblages in small, boreal lakes in Alberta.

One recent framework developed in northern Wisconsin that could potentially encompass both colonization routes and habitat suitability is the landscape position of a lake. Landscape position summarizes the interactions between a lake and the landscape in which it is embedded (Kratz *et al.* 1997, Magnuson and Kratz 1999). Landscape position is determined by the relative location of a lake in a surface and ground water flow system and by connections to neighboring lakes. Lakes lower in the landscape receive a greater proportion of their water from surface water and regional groundwater than do lakes higher in the landscape, which, in turn, influences a lake's water chemistry (Kratz *et al.* 1997, Soranno *et al.* 1999). Lakes lower in the landscape also tend to be larger and deeper and have more stream inlets and outlets then lakes higher in the landscape (Riera *et al.* 2000). Landscape position, via its relationships with surface connectivity and the local abiotic environment, should strongly influence fish assemblages in lakes (Fig. 1). As well, because it encapsulates the interactions between a lake and the landscape, landscape position is also an index of a lake's sensitivity to different kinds of disturbance (Webster *et al.* 1996), including development activities occurring on the landscape (Devito *et al.* 2000).

A newly proposed metric of landscape position, lake order, was recently developed in the Northern Highlands of Wisconsin based on the types and strengths of surface connections among lakes and the regional drainage network (Riera *et al.* 2000). Lake order is determined by geographical information easily obtained from topographic maps, making it an attractive metric. Within the Northern Highlands Lake District of Wisconsin, lake order has been correlated with lake morphometry, ion concentrations, chlorophyll *a* concentrations, crayfish abundance, fish species richness (Riera *et al.* 2000) and snail species richness (Lewis and Manguson 2000). However, its applicability outside of northern Wisconsin has not been tested.

Landscape position, because it affects the physical, chemical, and biological character of a lake, should be important in Alberta. Smaller lakes in boreal Alberta tend to be shallow and naturally eutrophic, thus, oxygen depletion during winter is extremely likely and "winterkill" is a major natural disturbance affecting fish populations (Danylchuk and Tonn, *in prep.*) and the broader lake communities (Langlois 2000, Tonn *et al. in prep.*). Small-bodied forage fishes, such as fathead minnow (*Pimephales promelas*) and brook stickleback (*Culaea inconstans*), are better able to maintain populations in lakes subject to severe oxygen depletion than are larger piscivorous fishes, such as northern pike (*Esox lucius*) and yellow perch (*Perca flavescens*) (Casselman and Harvey 1975, Klinger *et al.* 1982, Magnuson *et al.* 1985, Robinson and Tonn 1989). Where larger piscivorous fishes can persist, they can exclude small-bodied species (Robinson and Tonn 1989). Because landscape position should affect both the susceptibility of a lake to winter oxygen depletion (through its influence on the lake's physical and chemical properties) and the recovery of the fish populations following winterkill (via the stream connections of a lake), the presence and temporal stability of fish assemblages in the small lakes of boreal Alberta should be closely related to landscape position.

Objectives and hypotheses

To evaluate the applicability of lake order in boreal Alberta, I tested the usefulness of lake order in predicting the composition of fish assemblages and explaining fish species richness in 102 small (< 370 ha) lakes in three regions. For comparison, I also tested landscape-level measures of surface connectivity in combination with local factors, such as morphometry and water chemistry, in predicting types of fish assemblages and explaining fish species richness. I hypothesized that if lake order efficiently summarized surface connectivity, morphometry, and water chemistry, then lake order would predict and explain fish presence and composition as well as or better than other landscape-level metrics of surface connectivity and local factors.

Methods

Study area

To develop predictive models of fish assemblages and explanatory models of species richness in small, boreal lakes, I analyzed three data sets from three areas in north-central Alberta (Fig. 2). All three areas are located in the Boreal Mixedwood Ecoregion on the Boreal Plain, a region of low relief (Strong 1992). In this region, sedimentary bedrock is overlain by up to 200 m of glacial till containing clay (Mitchell and Prepas 1990, Devito *et al.* 2000). Soils and corresponding vegetation include well-drained brunisols with jack pine (*Pinus banksiana*), moderately drained gray luvisols with aspen (*Populus tremuloides*), and poorly drained gleysols and organics with black spruce (*Picea mariana*), willows (*Salix spp.*), and sedges (*Carex spp.*) (Strong 1992). In addition, the soils and glacial deposits tend to be rich in carbonates, which leads to well-buffered, nutrient-rich lakes (Mitchell and Prepas 1990).

The climate is continental, with a mean summer temperature of 13.5 °C and a mean winter temperature of −13.0 °C. Annual mean temperature is 0.8 °C. Summers are short (~85 frost-free days) and winter ice cover averages 6 months (Mitchell and Prepas 1990). Annual total precipitation is 400 mm, with the majority of precipitation occurring during June and July. Potential evapotranspiration often exceeds mean total precipitation (Strong 1992).

Data sources

All three of the data sets I used in my analyses included information on fish presence-absence (Appendix A-1), as well as morphometric and environmental data (Appendix A-2). I sampled 24 lakes in the Utikuma Lake District (UTK) from June through August 2000. Although concentrated geographically in the northwestern part of the study region (55.96 – 56.13°N and 115.15 – 115.70°W, Appendix A-3), lakes were chosen from four different glacial landforms: outwash plain, moraine, clay-till plain, and other. I also considered access by roads and seismic lines when selecting lakes. Lakes in UTK ranged in surface area from 2 – 362 ha (Appendix A-2). The Utikuma Lake District is largely forested, but comprises a very active region for oil and gas exploration. This project shared water chemistry data with the Hydrology, Ecology and Disturbance Research Initiative led by Drs. K. Devito, S. Bayley, and L. Foote, University of Alberta and Dr. I. Creed, University of Western Ontario. The 41 lakes in the Athabasca County Data Set (ATH) were located in the southern part of the study region (54.2 – 55.0°N, 112.5 – 113.8°W, surface area 3 – 304 ha, Appendix A-2, 3), where land clearing for agriculture and cattle grazing is the main form of landscape disturbance (Paszkowski and Tonn 2000, C. Paszkowski and W. Tonn, *unpublished data*). Finally, the Sustainable Forest Management Network Data Set (SFM, 37 lakes, 15 – 206 ha, Appendix A-2) was the most widespread geographically (54.98 – 56.82°N, 110.03 – 116.45°W, Appendix A-3), but concentrated on headwater lakes (Prepas *et al.* 2001, W. Tonn and others, *unpublished data*). The SFM landscape was least disturbed of the three, but fire and limited forest harvesting occurred in the catchments of several lakes a few years prior to sampling.

Fish

In all three data sets, fish data were collected according to the following protocol. Fish in water greater than 1 m (UTK) or 2 m (ATH and SFM) were sampled overnight (~ 16 h) with Lundgrens multi-mesh gill nets (14 3m x 1.5m panels with bar mesh sizes from 6.25 mm to 75 mm) set in a stratified random design based on water depth. Fish in water less than 1 m (UTK) or 2 m (ATH and SFM) were sampled overnight (~ 16 h) with unbaited Gee minnow traps. Fish were identified to species (Nelson and Paetz 1992), measured, and quantified using catch-per-unit-effort (CPUE); sampling effort was adjusted for lake area. However, abundance can fluctuate year-to-year due to winterkill or partial winterkill (W. Tonn, *unpublished data*), so species were classified as present or absent.

Species richness was calculated for each lake. Then, lakes were classified into one of three main categories: (1) no fish, (2) non-piscivorous, mainly smaller-bodied fish (*e.g.* brook stickleback (*Culaea inconstans*), fathead minnow (*Pimephales promelas*), finescale dace (*Phoxinus neogaeus*), northern redbelly dace (*Phoxinus eos*), pearl dace (*Margariscus margarita*), longnose sucker (*Catostomus catostomus*), white sucker (*Catostomus commersoni*)), or (3) piscivorous, with all but one of the 28 lakes dominated by northern pike (*Esox lucius*) alone or in one of three combinations: yellow perch (*Perca flavescens*) and white sucker, spottail shiner (*Notropis hudsonius*) and Iowa darter (*Etheostoma exile*), or rarely, the non-piscivorous fish listed above (*sensu* Robinson and Tonn 1989, W. Tonn, *unpublished data*, Appendix A-1). While the piscivore classification is broad (including 4 subtypes), my main interest was in determining which factors predicted the presence of piscivores, especially northern pike, which is an important game fish in Alberta (Nelson and Paetz 1992).

Lake morphometry

Surface area and perimeter of the UTK lakes were determined by digitizing the lakes on a Landsat Thematic Mapper 7 image (September 9, 1999, panchromatic band with resolution 15m x 15m, path 44, row 21, georectified in North American Datum 1983) in ArcView GIS 3.2a. ATH and SFM lakes' surface areas and perimeters were digitized from 1:50 000 National Topographic System Grid maps (NTS, Surveys and Mapping Branch, Department of Energy, Mines, and Resources, Canada, Appendix B). Shoreline development, a measure of shape, was calculated from perimeter (Wetzel 2001). Catchment areas of all lakes were estimated from 1:50 000 NTS maps (Appendix B). UTK does not have much topographic relief, so where elevation contour intervals were not fine enough to clearly delineate catchments, boundaries were drawn halfway between lakes. Estimated catchments were digitized in ArcView GIS 3.2a to measure catchment area. Slopes of the catchments were estimated by dividing the watershed into 6 to 16 equal parts, depending on the size of the catchment. The change in elevation was divided by the distance from the watershed boundary to the shoreline of the lake (*sensu* Prepas *et al.* 2001).

Maximum depths for UTK lakes were determined by measuring depth with a weighted rope along three transects that crossed the lakes. Between 12 and 20 points were located on each transect, depending on transect length. Each depth measurement location was recorded to within 10 m using a global positioning system (GPS Garmin 12 XL, Universal Transverse Mercator, North American Datum 1983). To calculate lake volume and mean depth, transect data were overlaid on the digitized polygons of the lakes in ArcView 3.2a and then the spline function of the Spatial Analyst extension was used to create continuous surfaces from the transects' points. Splines were used for local interpolation instead of kriging because the distribution of depth points was along transects instead of from more diverse locations in the lake, making kriging unfeasible (Burrough and McDonnell 1998). The resultant grid was then converted into a triangular irregular network to obtain three-dimensional characteristics of the lakes and calculate lake volume. Lake volume was divided by surface area to obtain mean depth (Wetzel 2001). For three UTK lakes for which transect data were not available, mean depth was estimated from a linear regression of mean depth on maximum depth for the other 21 lakes ($y = 0.218 + 0.315x$, $r^2 = 0.852$, $P < 0.01$ for coefficients). Mean depth was not available for ATH and SFM lakes.

Water chemistry

For ATH lakes, water chemistry data were from single samples collected in 1986, 1987, 1993, or 1994 depending on the lake Paszkowski and Tonn 2000, C. Paszkowski and W. Tonn *unpublished data*). Water chemistry data for SFM were averages of three samples collected in July, August, and September 1996 (Prepas *et al.* 2001, W. Tonn and others *unpublished data*). For UTK, I collected single water samples August 9-11, 2000 and partially processed them in the field before final analyses, using standard methods, for total phosphorus, total dissolved phosphorus, chlorophyll *a* (University of Alberta Limnology Laboratory), nitrate and ammonia (Dr. K. Devito Laboratory, University of Alberta), and calcium, magnesium, sodium, and potassium (Department of Renewable Resources, University of Alberta). Chlorophyll *a* values from August 1999 (K. Devito, *personal communication*) were substituted for missing chlorophyll *a* values for lakes U111 and U121. Total phosphorus from August 1998 was substituted for missing total phosphorus data for lake U026 (K. Devito, *personal communication*). Conductivities of water samples were measured in the laboratory using a WTW Multiline Universal Meter. Water pH was measured in the field using a Hydrolab Surveyor II (Hydrolab Corporation, Austin, TX, USA). More environmental variables were available for analysis in UTK than in ATH or SFM (Appendix C).

Landscape-level metrics

To the study lakes, I applied three landscape-level metrics. Each metric was based on the types and strengths of surface connections, which were determined from 1:50 000 NTS maps (Appendix B) and confirmed, where possible, by ground-truthing.

Lake order (*sensu* Riera *et al.* 2000; Fig. 3) - Lakes with permanent, channelized inlets and outlets were given positive lake orders based on the order of the stream that drained the lake. A headwater lake (with an outlet but no inlet) was given lake order of 0. Lakes with intermittent streams were given lake orders of -1. Lakes surrounded by wetlands were given lake order of -2. Isolated lakes (depicted without permanent or intermittent connections to streams or wetlands) were lake order of -3.

Lake order was developed in the northern Wisconsin Highlands (Riera *et al.* 2000), a region where the surface drainage networks are better defined than the Boreal Plains of Alberta, so two modifications were developed that incorporated isolation from the regional surface drainage network.

Isolation (Fig. 4) - For lakes of order 0, -1, or -2, did local surface connections actually connect to the wider drainage network or were they regionally isolated? The answer to this question produced a binary measure of surface connectivity. For a -2 order lake not to be isolated, its wetland must be connected to the drainage network by bordering on another lake with positive lake order or being drained by a permanent or intermittent stream connected to a permanent stream (based on 1:50 000 NTS maps). Some order 0 and -1 lakes were considered isolated if their outlet streams diffused into isolated wetlands. Isolated wetlands were those that did not border on a lake with positive lake order or were not drained by a permanent or intermittent stream connected to a permanent stream. Lakes could be isolated by natural factors, *i.e.,* streams and wetlands that do not connect topographically to the regional drainage network. Lakes could also be artificially isolated from the regional drainage network by roads. In the Utikuma Lake District, for example, it was not uncommon to find lakes isolated from regionally well-connected wetlands or streams (as defined above) by roads lacking culverts, with culverts blocked by debris, or with perched culverts (*personal observation*).

Connection - This system combined lake order and isolation into 3 classes: (1) those having a permanent, channelized connection (positive lake orders and lake order 0), (2) those with intermittent connections via streams or wetlands (lake orders -2 and -1), and (3) isolated lakes (lake order -3 and as defined above in *isolation*).

Data analysis – Characteristics of lakes

Because the three data sets varied in geographic location and landscape disturbances, I assessed general limnological relationships among them using principal components analysis (PCA, PC-ORD v.4.0) to determine if the data sets were similar enough to be combined or if one set could be a training set and the others could be reserved for test sets, or if they were dissimilar and should be analyzed separately. To meet assumptions of normality for PCA (McGarigal *et al.* 2000), mean depth, maximum depth, surface area, shoreline development, slope, conductivity, NH_4, total phosphorous (TP), total dissolved phosphorous (TDP), chlorophyll *a*, Ca^{2+}, Mg^{2+}, Na^{+}, and K^{+} were $\log_{10}(x+1)$ transformed. Normality was assessed via normal probability plots in SPSS (Ott 1993). Variables that were highly correlated with each other (Pearson $r > 0.70$, $P < 0.05$, SPSS v. 10.1) and did not have a strong biological reason for being included were excluded from analysis to reduce multicollinearity (Ott 1993, McGarigal *et al.* 2000). For example, mean depth was excluded as it was highly correlated with maximum depth, which is more likely related to winterkill via its function as an oxygen refuge. Variables included in the PCA were lake order, maximum depth, surface area, shoreline development, slope, pH, and TP. Correlation matrices were used in PCA. Following PCA, multiple-response permutation procedures (MRPP, Zimmerman *et al.* 1985) were used to test if the three groups of lakes from the three data sets differed in limnological characteristics. MRPP was performed using Euclidean distance on the transformed matrix (PC-ORD v.4).

Data analysis – Fish assemblages

Because results from the PCA and MRPP showed that the data sets differed in environmental characteristics and fish assemblages (see results), I modeled each data set separately, as well as the three sets combined (see below for details). For the 24 lakes in UTK, I developed logistic regression models to predict fish presence/absence only, as only 2 of the 12 fish-bearing lakes had piscivores. Thirty-three of 41 lakes in ATH had fish, of which 14 had piscivores. Therefore, I first constructed logistic regression models to predict fish presence/absence and then used the subset of lakes (n = 33) containing fish to build predictive logistic regression models of piscivore presence/absence. Since 36 of 37 SFM lakes contained fish, I used the 36 fish-bearing lakes to develop predictive logistic regression models for piscivore presence/absence. Because these individual data sets are too small (< 60) to be split into training and test sets for model validation (Olden and Jackson 2000), I used jackknife cross validation (JKV). JKV systematically excludes one observation, constructs a model, uses the model to predict the response variable of the excluded observation, and repeats the procedure for all observations.

To analyze the combined data set, I first formed a training set by grouping lakes into seven bins based on their PCA axis 1 scores and then randomly drew two-thirds of lakes from each bin. The remaining one-third of lakes from each bin formed the test set for the predictive models. This style of data-splitting for cross validation shows less model-accuracy and model-precision bias than other forms of validation (Olden and Jackson 2000).

For the training set, I first developed logistic regression models of fish presence. Then, using the subset of lakes with fish, I used logistic regression to model the presence of piscivores. Each model was assessed by applying it to the test set. To obtain an overall error rate, I applied the piscivore model to lakes in the test set that were predicted to have fish by the fish presence model. I then calculated the percentage of lakes with and without fish, with non-piscivores, and with piscivores that were correctly classified.

All modeling was done with logistic regression rather than the commonly used discriminant function analysis, because there are fewer assumptions in logistic regression, namely data do not have to be multivariate normal and the variances within groups do not have to be equal (Press and Wilson 1978, McGarigal *et al.* 2000). Also, logistic regression can accommodate data sets that have both nominal and quantitative variables (Press and Wilson 1978, ter Braak and Looman 1987, Trexler and Travis 1993), whereas mixed data sets probably should not be used in discriminant function analysis because they violate the aforementioned assumptions (Press and Wilson 1978, Williams 1983, McGarigal *et al.* 2000).

The logistic function takes the form $\pi(\eta) = \frac{e^{\eta}}{1+e^{\eta}}$, where the probability of fish (π) is a function of the linear predictor (η). The linear predictor is $\eta = \beta_0 + \beta_1 x_1 + \ldots + \beta_p x_p$, where p is the number of explanatory variables (McCullagh and Nelder 1989). I set the expected value to predict presence of either fish or piscivores (depending on the data set) as $0.5 + |(0.5 - \text{prevalence})|$ (Olden *et al.* 2002).

For each analysis, I started with two basic models, one incorporating lake order, morphometric, and chemical variables, and a second that excluded landscape position. From those starting models, several models were chosen for further examination based on a combination of χ^2 probabilities of variables (considered significant and retained if $P < 0.1$), backwards regression using Mallow's C_p (Venables and Ripley 1999), and biological interest. In the models of interest, lake order was systematically replaced with isolation and then with connection to facilitate comparison of the three metrics of landscape position. I compared models using the Aikake Information Criterion (AIC) (Hilborn and Mangel 1997, Venables and Ripley 1999), JKV (Olden and Jackson 2000), and sensitivity and specificity (MacKinnon 2000). Sensitivity is the ability to predict presence (*i.e.*, proportion of lakes with fish that were classified as having fish) and specificity is the ability to predict absence (*i.e.*, proportion of lakes without fish that were classified as fishless) (MacKinnon 2000). For the individual data set models, sensitivity and specificity were calculated using the JKV predictions. For modeling of the combined data set, sensitivity and specificity were assessed based on test set predictions.

Data analysis – Species richness

To determine which local and landscape features account for variation in fish species richness, Poisson regression models were built using the same landscape and local factors described above. Data were not transformed prior to modeling and all 3 data sets were combined for model construction. Poisson regression models, which are used for count data, take the form $\log \mu_i = \eta_i = \beta^T x_i$ where $\log \mu_i$ is the expected species richness, η_i is the linear predictor (*i.e.*, the linear combination of explanatory variables), $\boldsymbol{\beta}^T$ is the matrix of coefficients and $\mathbf{x}_i$ is the matrix of explanatory variables (McCullagh and Nelder 1989). I started with maximum depth, surface area, slope, pH, TP, and lake order and used backwards stepwise regression with Mallow's C_p (Venables and Ripley 1999) to pare the models. Subsequent to model selection, lake order was systematically replaced with isolation and then connection to examine how much explanatory power each of the three metrics of landscape position provided. In the final models, importance of each factor was evaluated by determining how much deviance it explained. The significance of the relationship between species richness and each factor was analyzed via P-values from the Poisson regression models.

Model specifications

All modeling was done in S-PLUS 2000 (Mathsoft Inc. 1999). Parameters of the linear predictors were estimated using iterative weighted least squares (Venables and

Ripley 1999). The stopping rule for convergence was $|\text{deviance}^{(i)} - \text{deviance}^{(I\text{-}1)}| < \varepsilon(\text{deviance}^{(I\text{-}1)} + \varepsilon)$ with $\varepsilon = 10^{-5}$ with maximum iterations set at 20. Dispersions of the models were also close to 1, thus, over-dispersion was not a problem (McCullagh and Nelder 1989). Partial residual plots were examined for outliers and evidence of non-linearity (McCullagh and Nelder 1989). Values of coefficients and P-values for each model are in Appendix D.

Results

Characteristics of lakes

Visual inspection of the PCA plots of the environmental characteristics of lakes suggested that ATH and UTK lakes spanned a broader range of environmental variables than the lakes in the SFM data set (Fig. 5a, b). That corresponds to the site selection criteria for the SFM lakes, which were limited to headwater lakes with enough surface area on which to land a plane (W. Tonn, University of Alberta, *personal communication*). Axis 1 was associated with maximum depth ($r = -0.813$) and axis 2 with TP ($r = 0.774$). UTK lakes tended to be shallower than lakes from ATH or SFM. The plot of axis 3 by 2 (Fig. 5b) shows SFM lakes in the lower part of the plot, an area associated with lower pH ($r = 0.863$). The result from MRPP indicated that lakes from the data sets differed significantly in their environmental characteristics ($P \ll 0.01$).

Environmental characteristics also differed among lakes when lakes were classified by fish assemblage type (Fig. 5c). Fish assemblage type corresponded to axis 1, which was strongly correlated with maximum depth. On the right hand side of the PCA are primarily shallow lakes without fish, which then grades into a region containing non-piscivore lakes, which, in turn, overlaps into the region of deep piscivore lakes (Fig.

5c). The region of high overlap between non-piscivore and piscivore lakes occurred between 3 m and 6 m maximum depth.

Further breakdown of fish composition in piscivore lakes by depth showed that below 4.5 m maximum depth, fathead minnows, brook stickleback, and dace co-occurred with northern pike in six lakes, whereas Iowa darter and spottail shiner were small-bodied species found with northern pike in lakes deeper than 9 m maximum depth. The large-bodied white sucker and yellow perch regularly co-occurred with northern pike throughout the maximum depth range of piscivore lakes.

Fish assemblages - Individual data sets

Of all landscape-level factors, lake order did the worst job predicting fish presence in UTK. Models that used lake order had fewer correct predictions, lower sensitivity and specificity, and higher AIC values than those using either isolation or connection, the landscape-level metrics of surface connectivity. Replacing lake order with isolation yielded the model with the lowest AIC value (75), highest percent correct (79%), and best sensitivity (83%) and second highest specificity (75%) (Table 1). Models using connection had values intermediate between those using isolation and lake order. Adding local factors, such as maximum depth and surface area, to models with landscape position did not improve values for percent correct, sensitivity or specificity, had higher AIC values, and used more degrees of freedom (Table 1). The best model containing only local factors (total phosphorous and surface area) did a poor job predicting presence and absence of fish correctly only 54% of the time.

Predicting fish presence in ATH produced similar results. Lake order had fewer correct predictions, lower sensitivity and specificity, and higher AIC values than models using isolation or connection. Isolation alone was as accurate, but more parsimonious than other models in terms of JKV (88%), sensitivity (88%) and specificity (87.5%) (Table 2). While connection made the same predictions as isolation, it was not as parsimonious because it had an extra degree of freedom and hence, a higher AIC value. As in UTK, local factors did not add predictive power to models containing landscape-level metrics of surface connectivity and the best model including only local factors (slope, surface area, and maximum depth) did a poor job predicting fish presence/absence (Table 2).

Models that predicted piscivore presence were not as accurate as those of simple fish presence. Additionally, landscape-level metrics did not provide the predictive power they did in models of fish presence. In ATH, the most parsimonious model of piscivore presence included only maximum depth (Table 3a). Sensitivity (71%) and specificity (84%) were well balanced and the JKV was good (79%). Adding isolation marginally improved the number of correct predictions (82%). Other measures of landscape position did not yield better results. Indeed, lake order performed worse than the simple logistic model containing only maximum depth (Table 3a).

Landscape-level metrics of surface connectivity were the first to be excluded in models predicting piscivore presence in the SFM data set. Like ATH, maximum depth was the most important explanatory variable, although JKV was lower (72%) than for ATH (Table 3b). Unlike ATH, however, sensitivity (33%) and specificity (92%) were not well balanced. It was difficult to predict the presence of piscivores when they were present, but the model gave few false presences.

Fish assemblages - Combined data set

As with the individual data sets, lake order and connection were not as parsimonious or as powerful as isolation in predicting fish presence. Isolation alone provided the most parsimonious model in the combined data set, correctly predicting fish presence/absence for 80% of the lakes in the test (Table 4a). Adding maximum depth to isolation improved the percent correct only slightly to 83%. Its improvement resulted from greater sensitivity (89%) rather than improved specificity. Similar to the individual data sets, models without landscape-level factors did a poor job predicting fish presence/absence. The model including only maximum depth correctly predicted presence/absence in 57% of the lakes.

Modeling the presence of piscivores in lakes known to have fish yielded results similar to those of the individual data sets with landscape position being excluded early on in the model building process. The most parsimonious model included only maximum depth, although the addition of surface area increased the percent correct from 71% to 79%, and also improved model sensitivity and specificity (Table 4b).

Finally, I combined the two best models for predicting fish presence (isolation; isolation plus maximum depth) with the two models for predicting piscivore presence (maximum depth; maximum depth plus surface area) to obtain overall error rates. Lakes from the test set that were predicted to have fish were subsequently entered into a piscivore presence model. Combining the two simplest models (isolation; maximum depth) led to the lowest overall percentage correct (60%), whereas the combination of the two more complicated models (isolation plus maximum depth; maximum depth plus surface area) gave the highest percentage correct (66%) (Table 5). This latter combination was equally or more accurate in predicting three of the subcategories (% no fish, % fish, % piscivore) as the other three pairings (Table 5).

Species richness

The range of species richness was 0 to 6 species, and was similar in non-piscivore (1 – 5) and piscivore lakes (1 – 6). Median species richness, however, was lower in non-piscivore lakes (1) than in piscivore lakes (3).

Overall, local and landscape factors combined to account for more variation in species richness then either type alone (Table 6). Models using lake order explained more variation than models with isolation or connection (Table 6). However, models with connection were more parsimonious than models with lake order or isolation. The model including maximum depth, surface area, and connection was the most parsimonious, but explained less than the same model with lake order 34 vs. 40%). Models with isolation explained the least amount of variation in species richness and were least parsimonious. Lake order individually accounted for the most variation (28%) with connection a close second (26%). Surface area and maximum depth accounted individually for 13% and 9%, respectively. Increasing lake order (Fig.6), surface area, and maximum depth were associated with greater species richness ($P < 0.01$, from Poisson regression models, Appendix D-7).

Discussion

Fish assemblage type

Results from the individual data set models and the combined data set models were similar. In each case, models incorporating landscape-level metrics of surface connectivity did a good job predicting presence, but not absence, of fish. Because winterkill is a major natural disturbance affecting fish populations in small, boreal lakes (Danylchuk and Tonn, *in prep.*), surface connections provide recolonization routes, and thus are important for fish presence. However, isolation alone was not sufficient to predict fish absence accurately. Five of 21 fishless lakes were predicted to have fish on the basis of their having a surface connection. These lakes probably could support fish in many years, but as they were fairly shallow (< 2.5 m), may have recently winterkilled. Indeed, lake U201, for example, was fishless in 2000 (Appendix A-1), but contained brook stickleback in 2001 (I. Norlin, University of Alberta, *personal communication*). Lake U201 was only 2.26 m deep but was connected to larger, fish-bearing lakes by an intermittent stream. Thus, the difficulty in predicting fish absence in this dynamic region was probably related to the limitations of snapshot data (Diamond 1986, Magnuson 1990).

Lake order did not predict fish presence as successfully as other landscape-level metrics of surface connectivity. The simplest metric, isolation, yielded better predictions of fish presence likely because the prediction was simple: presence or absence. At any given level of lake order, especially higher orders, there were not many lakes, which led to noise surrounding predictions. Additionally, lake order, unlike isolation or connection, did not account for isolation from the regional drainage network, and thereby overestimated availability of recolonization routes.

For lakes in which fish were present, lake order and other landscape-level metrics of surface connectivity were among the first variables to be excluded from logistic regression models predicting piscivore presence. Instead, the local factor maximum depth was the main predictor of piscivore presence (Figure 7). Its success, however, was somewhat limited. The difficulty in separating non-piscivore lakes from piscivore lakes lies in the overlap of the two assemblages between 3 m and 6 m maximum depth. There were 9 non-piscivorous lakes and 11 piscivorous lakes in this range. Greater than 6 m maximum depth, 5 lakes were non-piscivorous and 16 were piscivore lakes. The logistic regression models did not start predicting piscivore presence until 6 m maximum depth, thus, maximizing specificity at the expense of sensitivity.

What is happening between 3 and 6 m maximum depth? One possibility is that that winterkill frequency is highly variable in this depth zone. Six of ten piscivore lakes 4.5 m or shallower also had small-bodied species, such as brook stickleback and fathead minnow, that are relatively intolerant of predation (Robinson and Tonn 1989), but better able to survive low oxygen conditions due to physiological and behavioral adaptations (Klinger *et al.* 1982, Magnuson *et al.* 1985). Relatively frequent partial winterkill of larger piscivores (Casselman and Harvey 1975) might allow these small-bodied fish to persist. In lakes deeper than 6 m, northern pike are probably much less vulnerable to even partial winterkill (Mathias and Barica 1980, Meding and Jackson 2001) and can likely more effectively exclude the smaller species. Hence, the models conservatively predicted piscivore presence in lakes deeper than 6 m, even though they can be found in shallower lakes.

Although landscape factors were not important predictors of piscivores, type and strength of surface connection might influence the probability of finding northern pike in lakes between 3 and 6 m maximum depth, *i.e.*, northern pike might more likely be found in marginal lakes with better recolonization routes or nearby winter oxygen refuges (Tonn and Magnuson 1982). Indeed, 10 of 11 lakes between 3 and 6 m maximum depth with piscivores had lake order of 0 or greater. However, 8 of the 9 non-piscivore lakes in this depth range and all 5 of the non-piscivore lakes greater than 6 m also had lake order of 0 or greater. Thus, strength of surface connection, as measured by lake order, does not explain piscivore presence/absence in lakes between 3 and 6 m maximum depth. The lack of a relationship between strength of connection and northern pike presence/absence in marginal habitats may be related to the limitations of snapshot data, which do not capture the inherent temporal variability in boreal forest lakes (Diamond 1986, Magnuson 1990).

The importance of landscape-level surface connectivity and maximum depth, suggested from my models as drivers of fish assemblages in small, boreal lakes in Alberta, complements and extends previous fish community analyses of immigration (isolation) and extinction (habitat factors) (Tonn *et al.* 1995, Magnuson *et al.* 1998, Olden *et al.* 2001). There were some important differences among studies.

There were differences among data sets in types of lakes used and how immigration factors were characterized, which led to seemingly different conclusions. Tonn *et al.*'s (1995) and Magnuson *et al.*'s (1998) data sets partially overlapped, contained fish-bearing seepage (*i.e.*, lake order < 0) and drainage lakes (*i.e.*, lake order ≥ 0), measured immigration factors by local isolation, and yielded similar conclusions. Extinction factors (maximum depth and surface area) explained more variation in fish assemblages than immigration factors (related to surface connectivity). However, both argued that the presence of seepage lakes (*i.e.*, lake order < 0) in their data sets reduced the importance of immigration factors because extinction frequency was much higher than immigration frequency. Olden *et al.*'s (2001) data set included only drainage lakes (*i.e.*, lake order ≥ 0), and they measured immigration factors by metrics of local isolation. They concluded that immigration and extinction were equally important. My data set overlapped with Tonn *et al.*'s (1995) in ATH, but I measured immigration factors more regionally via landscape-level metrics of surface connectivity. Another difference between my study and the others was the inclusion of fishless lakes. This resulted in a hierarchy. Immigration factors were important for fish presence, but once fish were present, extinction factors determined the type of fish assemblage. Thus, the initial differences in conclusions were likely due to differences in data sets.

Another difference among studies was the use of different analytical methods. All four studies used different methods to assess the relative importance of immigration and extinction. Tonn *et al.* (1995) used partial canonical correspondence analysis, Magnuson *et al.* (1998) used classification and regression trees, Olden *et al.* (2001) used correspondence analysis in combination with Procrustes analysis, and I used logistic regression. Despite the differences in analytical methods, the results of the studies were similar and thus, the conclusions appear robust to analytical method.

Species richness

As in small, forested lakes in Finland (Tonn *et al.* 1990) most of the variation in species richness in boreal Alberta lakes was explained by lake order, with additional variation explained by surface area and maximum depth. Larger, deeper lakes often provide more or different habitats (Tonn and Magnuson 1982, Eadie and Keast 1984, Eadie *et al.* 1986) and are less vulnerable to oxygen depletion (Barica and Mathias 1979, Mathias and Barica 1980, Babin and Prepas 1985, Nürnberg 1995, Meding and Jackson 2001), and thus, would be expected to have more species.

Lake order: Landscape position or surface connectivity?

Landscape position, as defined by Webster *et al.* (1996) and Kratz *et al.* (1997), is a lake's relative position in a groundwater flow network. Those authors found that in northern Wisconsin, a lake's water chemistry was directly related to landscape position via groundwater flows through glacial tills, and thus, landscape position could explain a lake's chemical responses to drought. Kratz *et al.* (1997) also noted that landscape position was related to surface area and fish species richness. However, Webster *et al.* (2000) found that landscape position was not as useful for explaining lakes' chemical responses to drought in three lake districts in Ontario as in northern Wisconsin because the geology of the regions differed. Ontario has shallow (1 – 10 m), low carbonate tills over impervious granitic bedrock, so hydrologic linkages among lakes are dominated by surface water. Northern Wisconsin has thicker (~50 m), permeable, low carbonate tills and hydrologic linkages are dominated by groundwater (Webster *et al.* 2000).

To make landscape position more accessible, Riera *et al.* (2000) developed lake order, a metric based on the types and strengths of surface connections. For lakes in the Northern Highlands of Wisconsin, they found significant relationships between lake order and 13 of 25 limnological variables, including surface area, conductivity, pH, calcium, chloride, nitrogen, silica, chlorophyll *a*, and fish species richness. Because lake order was related to calcium and silica, ions found in groundwater, Riera *et al.* (2000) concluded that lake order was an effective proxy for landscape position. However, the relationships mainly reflected a dichotomy between lakes with negative lake orders and those with positive lake orders, suggesting that relationships were driven more by strength of surface connectivity than by position.

In the Boreal Mixedwood Ecoregion of Alberta, tills are thick (up to 200 m) and carbonate-rich, thus landscape position is potentially applicable to this region (see Devito *et al.* 2000). Lake order, while not useful for predicting fish assemblage type, explained more variation in species richness than did simpler landscape-level metrics of surface connectivity. Riera *et al.* (2000) posited that fish species richness was related to lake order indirectly via its correlation with surface area and chlorophyll *a*. Unlike Wisconsin, however, lake order did not appear to be strongly related to surface area, maximum depth, total phosphorous, or pH in boreal Alberta (Figs. 8a, b, c, d). I suggest that for small lakes in boreal Alberta, lake order mainly provides a more detailed measure of surface connectivity, as opposed to a measure of landscape position; species richness increases with lake order because colonization routes and oxygen refuges are more available.

Conclusion

Most of the work done to date on landscape position and lakes has focused on physical and chemical aspects of lakes (*e.g.*, Webster *et al.* 1996, Kratz *et al.* 1997, Soranno *et al.* 1999, Devito *et al.* 2000, Riera *et al.* 2000), with little work done to examine the relationship between landscape position and biota (Lewis and Magnuson 2000). However, there is increasing awareness that at the landscape level, hydrologic connectivity is important for maintaining biodiversity and biological integrity (Pringle 1998). For freshwater biota, surface connections provide colonization routes, which have been shown to influence fish assemblages and species richness (Tonn *et al.* 1990, Tonn *et al.* 1995, Magnuson *et al.* 1998, Riera *et al.* 2000, Olden *et al.* 2001, this study). In addition, the ability of organisms to recolonize local sites is especially important for long-term persistence when those sites are harsh and variable (*e.g.*, Fahrig and Merriam 1985, Pulliam 1988, Stacey and Taper 1992, Hanski 1999, Lafferty *et al.* 1999). Thus, my study indicates that fish persistence in small lakes in boreal Alberta, which are subject to frequent winterkill, depends on surface connections for recolonization following these natural disturbances. Long-term data on fish populations and assemblages are needed to determine winterkill frequency in lakes of differing depths and surface connectivities; such data should enhance predictions of fish assemblage types by incorporating the probability of winterkill and successful recolonization into models. It is also important to track the frequency of winterkill in relation to climatic factors to be able to predict the effects of climate change on fish assemblages.

Landscape position, because of its summary of lake-landscape interaction, has recently been proposed as a framework for regional management of lakes subject to anthropogenic disturbances on the landscape (Devito *et al.* 2000). In boreal Alberta, the lack of information on fish assemblages has hindered lake management (D. DeRosa, Alberta Sustainable Resource Development, *personal communication*). This study evaluated whether lake order, recently proposed as a metric of landscape position, or simpler landscape-level metrics of surface connectivity could be useful. Because of their strong relationships with landscape-level metrics of surface connectivity and maximum depth, the types and richness of fish assemblages in small, boreal lakes could be predicted with simple data derived from remote sensing, topographic maps, and field collection. This information is vital for management as disturbances stemming from forestry and oil development activities can affect surface connections and water depth via road building (Anonymous 1990, Furniss *et al.* 1991, *personal observation*, B. McCulloch, Fisheries and Oceans Canada, *personal communication*) and direct removal of water from lakes (D. DeRosa, Alberta Sustainable Resource Development, *personal communication*). Since such disturbances can have rippling effects through a drainage network, management decisions should not be made on a lake-by-lake basis. This study can contribute to an understanding and prediction of how disturbance, natural or anthropogenic, can affect fishes in boreal Alberta and thus, could play a role in regional management of the many unstudied lakes of boreal Alberta.

Table 1. Comparison of logistic models for fish presence in the Utikuma Lake District. As appropriate, variables were log transformed. $E(\pi) > 0.5$ to predict presence because 50% of the lakes had fish. MD = model number in Appendix D-1, which contains the coefficients and P-values for each model. DF = degrees of freedom used in calculating AIC, AIC = Aikake information criterion, JKV = percent correct when jackknife cross-validation used, z_{max} = maximum depth, lk ord = lake order, connect = connection factor, TP = total phosphorous.

MD	Variables	DF	Residual deviance	AIC	JKV (%)	Sensitivity (%)	Specificity (%)
1	z_{max}, lk ord	7	17.58	150	71	58	83
2	z_{max}, isolation	4	24.24	77	79	83	75
3	z_{max}, connect	5	20.92	125	75	75	75
4	area, lk ord	7	18.15	146	67	58	75
5	area, isolation	4	23.69	78	79	83	75
6	area, connect	5	19.65	125	71	67	75
7	lk ord	6	18.75	143	67	58	75
8	isolation	3	24.27	75	79	83	75
9	connect	4	21.02	122	79	83	75
10	TP, area	3	24.50	85	54	50	58

Null deviance = 33.27

Table 2. Comparison of logistic models for fish presence in Athabasca County. Variables were log transformed when appropriate. $E(\pi) > 0.8$ because 80% of the lakes had fish. MD = model number in Appendix D-2, which contains the coefficients and P-values for each model. DF = degrees of freedom used in calculating AIC, AIC = Aikake information criterion, JKV = percent correct when jackknife cross-validation used, z_{max} = maximum depth, lk ord = lake order, connect = connection factor.

MD	Variables	DF	Residual deviance	AIC	JKV (%)	Sensitivity (%)	Specificity (%)
1	area, lk ord	8	22.06	288	78	79	75
2	area, isolation	4	18.55	160	85	85	88
3	area, connect	5	18.02	216	85	85	88
4	z_{max}, lk ord	8	23.13	284	80	82	75
5	z_{max}, isolation	4	20.86	145	88	88	88
6	z_{max}, connect	5	19.91	204	88	88	88
7	lk ord	7	28.51	255	73	70	88
8	isolation	3	23.19	142	88	88	88
9	connect	4	22.56	200	88	88	88
10	slope, area, z_{max}	5	29.48	146	66	70	50

Null deviance = 40.47

Table 3a. Comparison of logistic models for piscivore presence in Athabasca County. Variables were log transformed when appropriate. $E(\pi) > 0.58$ because 43% of the lakes had piscivores and $0.5 + |0.5 - 0.43| = 0.58$. MD = model number in Appendix D-3, which contains the coefficients and P-values for each model. DF = degrees of freedom used in calculating AIC, AIC = Aikake information criterion, JKV = percent correct when jackknife cross-validation used, z_{max} = maximum depth, lk ord = lake order, TP = total phosphorous, conn = connection.

MD	Variables	DF	Residual deviance	AIC	JKV (%)	Sensitivity (%)	Specificity (%)
1	z_{max}, lk ord	8	12.26	234	76	64	84
2	z_{max}, isolation	4	22.89	117	82	71	89
3	z_{max}, conn	5	19.92	134	79	64	90
4	z_{max}, area	4	20.45	123	76	64	84
5	z_{max}	3	23.03	116	79	71	84

Null deviance = 44.98

Table 3b. Comparison of logistic models for piscivore presence in the Sustainable Forest Management lakes. Variables were log transformed when appropriate. $E(\pi) > 0.67$ because 33% of SFM lakes had piscivores and $0.5 + |0.5 - 0.33| = 0.67$. MD = model number in Appendix D-4, which contains the coefficients and P-values for each model. DF = degrees of freedom used in calculating AIC, AIC = Aikake information criterion, JKV = percent correct when jackknife cross-validation used, z_{max} = maximum depth.

MD	Variables	DF	Residual deviance	AIC	JKV (%)	Sensitivity (%)	Specificity (%)
1	slope, z_{max}	4	33.17	119	67	25	88
2	z_{max}	3	34.70	115	72	33	92

Null deviance = 45.83

Table 4a. Comparison of logistic models of fish presence for all the regions. When appropriate, variables were log transformed. E(π) > 0.8 because 80% of the lakes had fish. MD = model number in Appendix D-5, which contains the coefficients and P-values for each model. DF = degrees of freedom used in calculating AIC, AIC = Aikake information criterion, % CORR = percent correct when models used on the test set, z_{max} = maximum depth, lk ord = lake order, isol = isolation, conn = connection. Sensitivity and specificity values calculated on the test set predictions.

MD	Variables	DF	Residual deviance	AIC	%CORR	Sensitivity (%)	Specificity (%)
1	area, z_{max}, lk ord	10	22.15	444	80	89	43
2	area, z_{max}, isol	5	27.81	226	83	89	57
3	area, z_{max}, conn	6	24.32	558	77	82	57
4	z_{max}, lk ord	9	23.23	424	80	89	43
5	z_{max}, isol	4	27.89	226	83	89	57
6	z_{max}, conn	5	24.60	556	77	82	57
7	lk ord	8	31.17	374	77	86	43
8	isol	3	37.39	211	80	86	57
9	conn	4	32.38	518	80	86	57
10	area, z_{max}	4	42.30	274	71	79	43
11	z_{max}	3	45.73	280	57	68	14

Null deviance = 68.68

Table 4b. Comparison of logistic models of piscivore presence in subset of lakes from all data sets known to have fish. When appropriate, variables were log transformed. E(π) > 0.64 because 36% of lakes had piscivores and 0.5 + |0.5 – 0.36| = 0.64. MD = model number in Appendix D-6, which contains the coefficients and P-values for each model. DF = degrees of freedom used in calculating AIC, AIC = Aikake information criterion, % CORR = percent correct when models used on test set, z_{max} = maximum depth. Sensitivity and specificity values were calculated on the test set predictions.

MD	Variables	DF	Residual Deviance	AIC	% CORR	Sensitivity (%)	Specificity (%)
1	area, z_{max}	4	37.97	185	79	67	84
2	z_{max}	3	39.79	179	71	56	79

Null deviance = 69.17

Table 5. Results of validating the hierarchical logistic models of firstly, fish presence, and if fish were expected to be present, secondly, presence of a piscivore. Validation was done using the test set with the same expected values of π as when validated individually. Values given are % of the lakes in the category classified correctly by the hierarchical models. Pisc = piscivore, z_{max} = maximum depth, Fish ~ = presence of fish is modeled by the following variables, Pisc ~ = presence of piscivores is modeled by the following variables.

	Models	No fish (%)	Fish (%)	Non-piscivores (%)	Piscivores (%)	Overall (%)
1	Fish ~ isolation Pisc ~ area + z_{max}	57	86	69	56	63
2	Fish ~ isolation Pisc ~ z_{max}	57	86	69	44	60
3	Fish ~ z_{max} + isolation Pisc ~ area + z_{max}	57	89	68	67	66
4	Fish ~ z_{max} + isolation Pisc ~ z_{max}	57	89	68	56	63

Table 6. Amount of deviance explained for Poisson regression models of species richness using all lakes. MD = model number in Appendix D-7, which contains the coefficients and P-values for each model. DF = degrees of freedom used in calculating AIC, AIC = Aikake information criterion, z_{max} = maximum depth.

MD	Variables	Explained Deviance	% Deviance Explained	DF	AIC
1	z_{max}, area, lake order	54.7	40	10	299
2	z_{max}, area, isolation	40.4	29	5	305
3	z_{max}, area, connection	47.5	34	6	292
4	area, lake order	54.3	39	9	298
5	area, isolation	35.7	26	4	315
6	area, connection	45.5	33	5	294
7	lake order	39.1	28	8	313
8	isolation	26.3	19	3	308
9	connection	36.1	26	4	299
10	area	18.0	13	3	324
11	z_{max}	12.5	9	3	315

Null deviance = 137.95

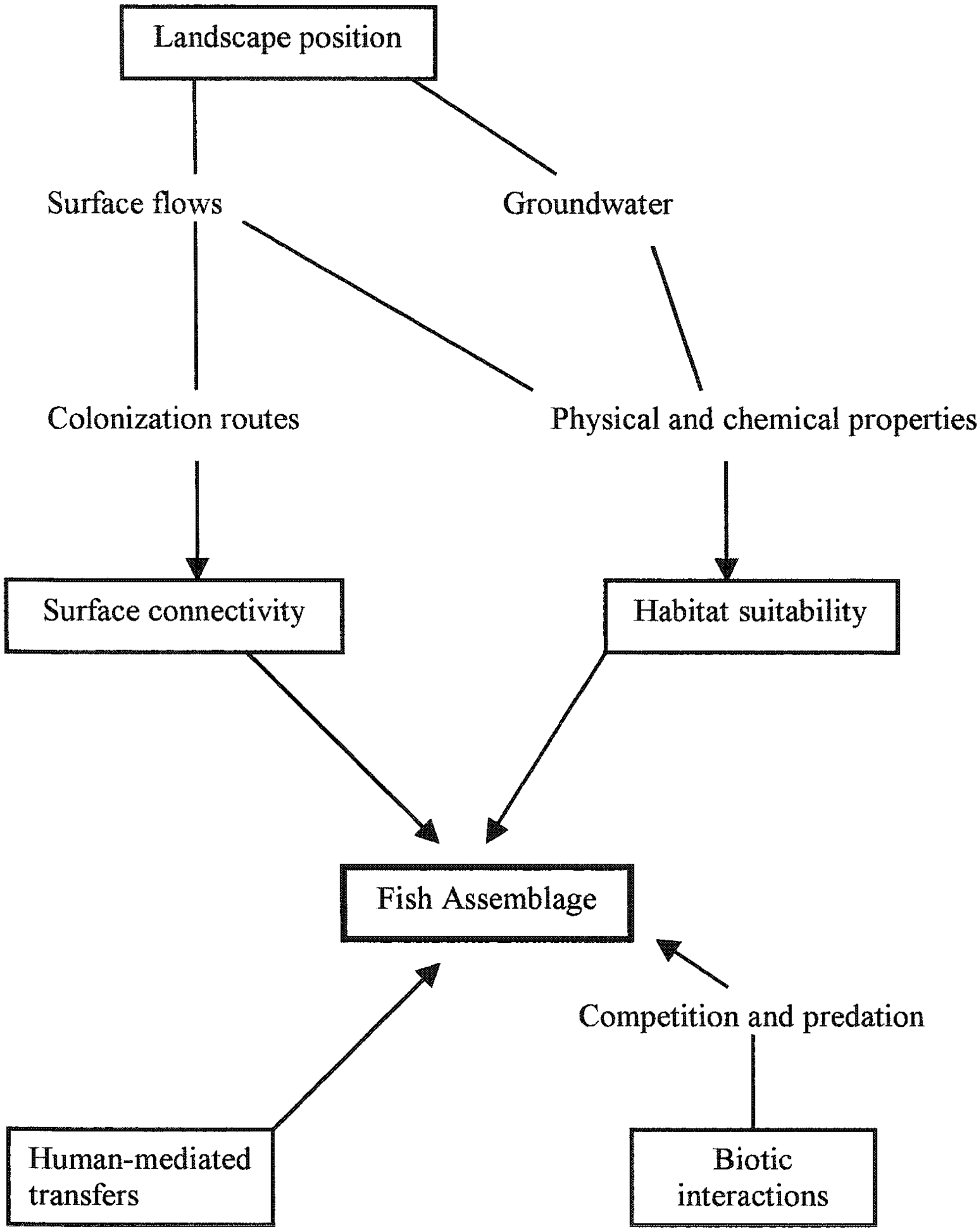

Figure 1. Conceptual proposal as to how landscape position, local factors, biotic interactions, and human-mediated transfers interact to form fish assemblages. See text for details.

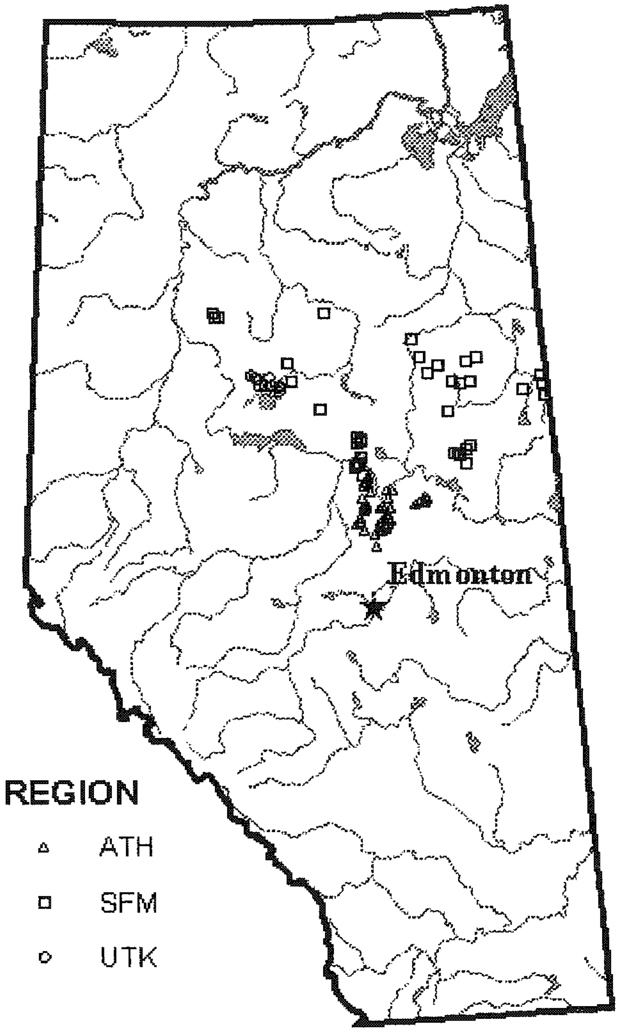

Figure 2. Location of study lakes in boreal Alberta. Data sets are as follows: ATH = Athabasca County, SFM = Sustainable Forest Management Network, UTK = Utikuma Lake District.

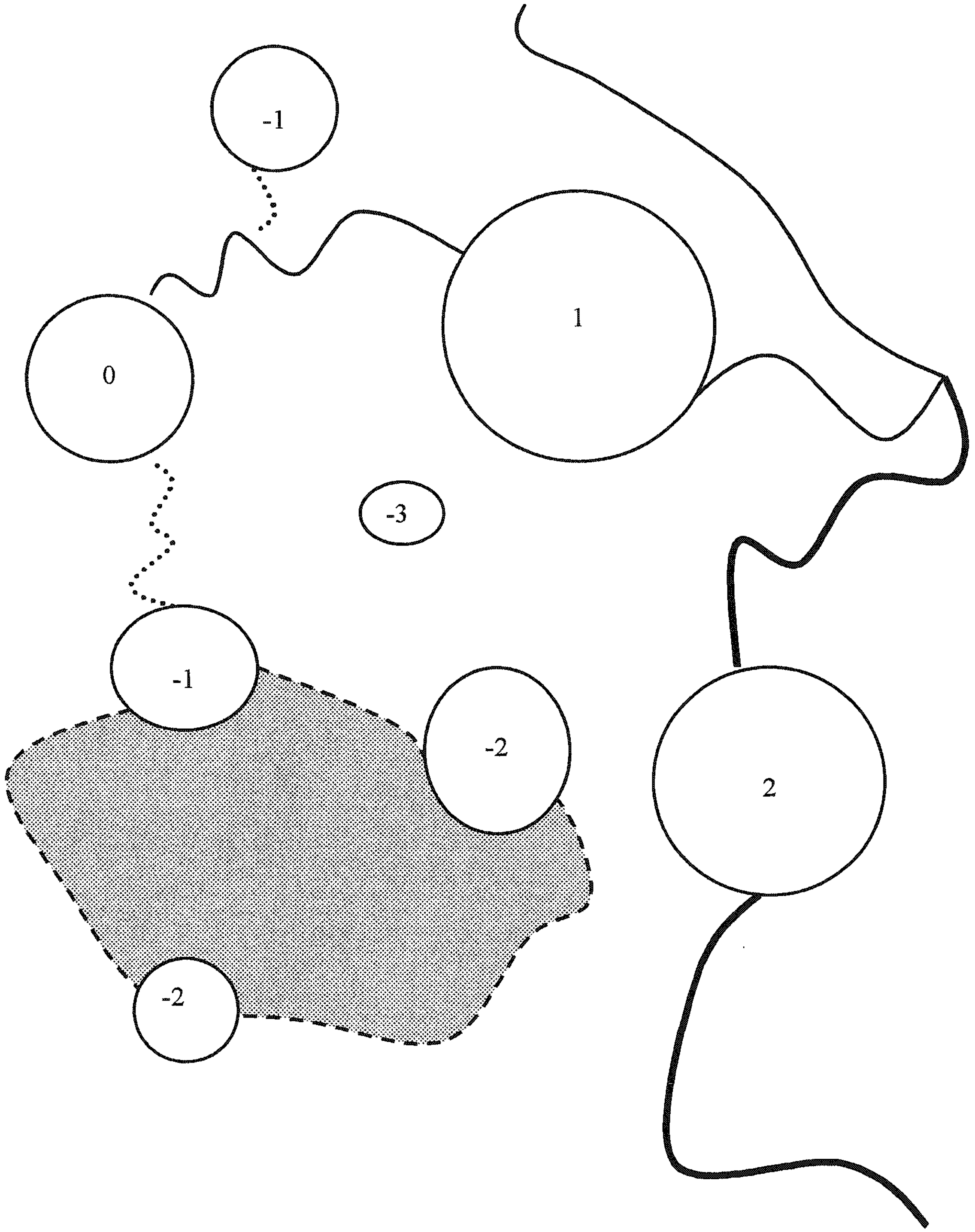

Figure 3. Classification of lakes by lake order (after Riera *et al.* 2000). Dashed lines outline a wetland; dotted lines are intermittent streams; solid lines of increasing thicknesses depict streams of increasing stream orders; circles are lakes. See text for details.

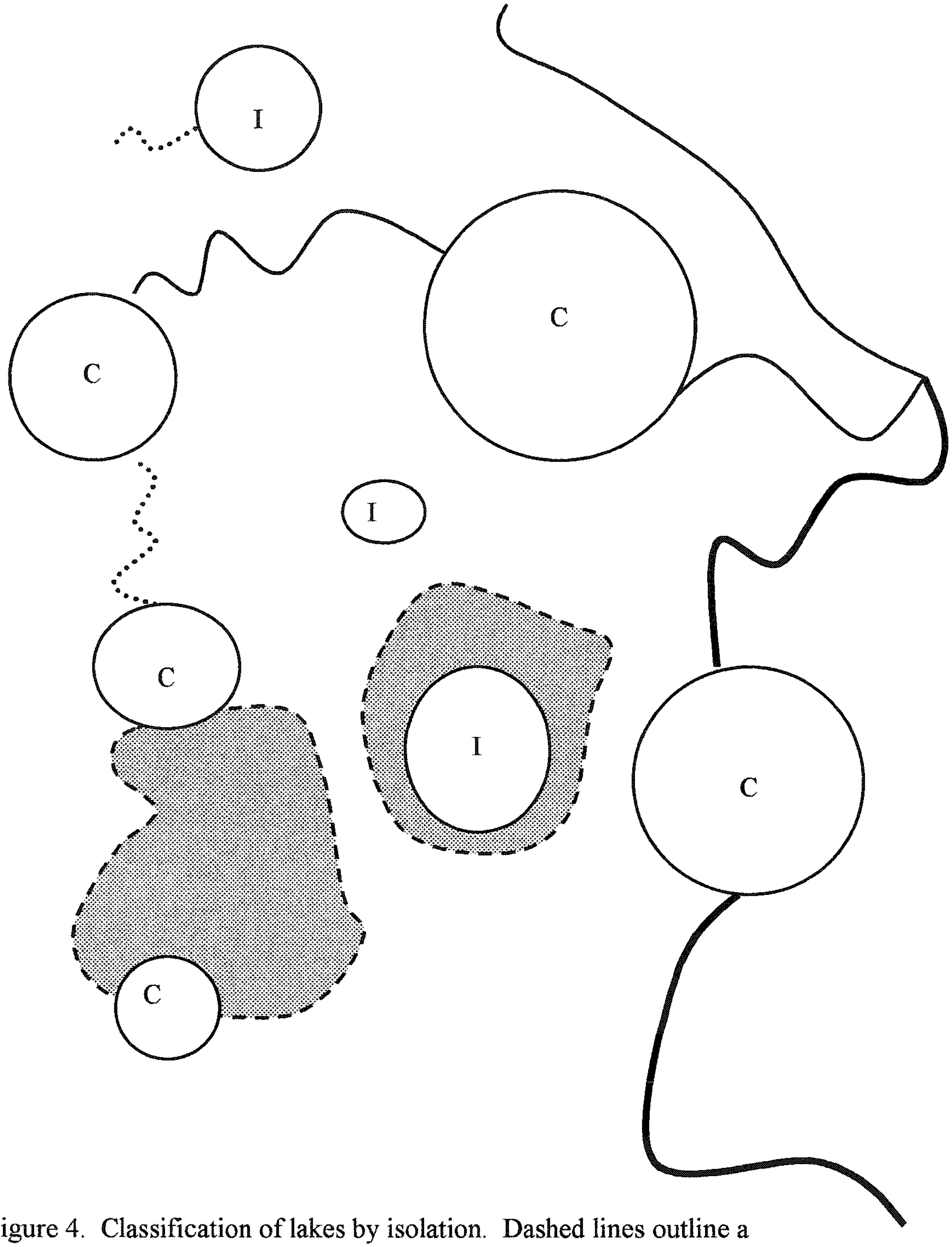

Figure 4. Classification of lakes by isolation. Dashed lines outline a wetland; dotted lines are intermittent streams; solid lines of increasing thickness depict streams of increasing stream order; circles are lakes. I – isolated, C – connected. See text for details.

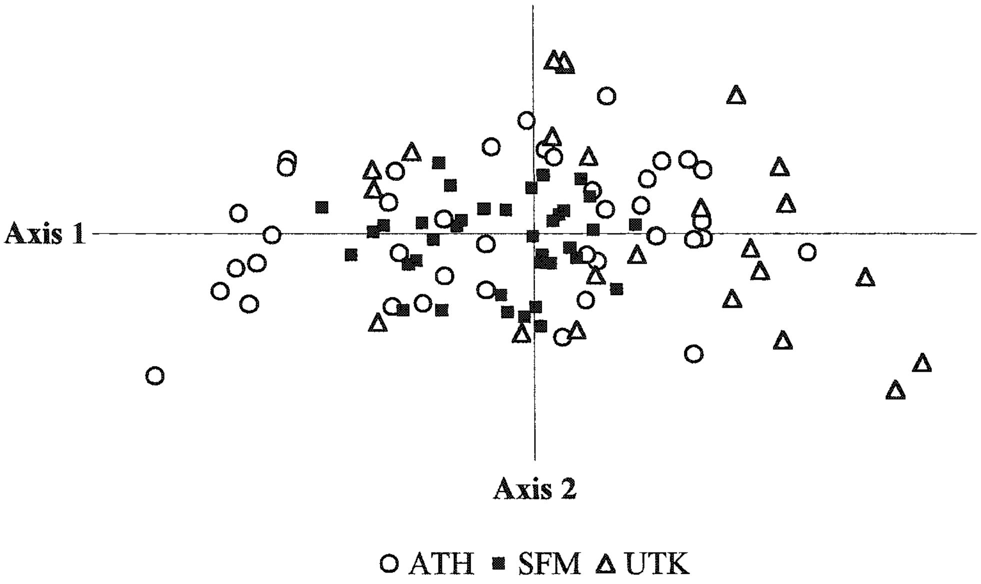

Figure 5a.

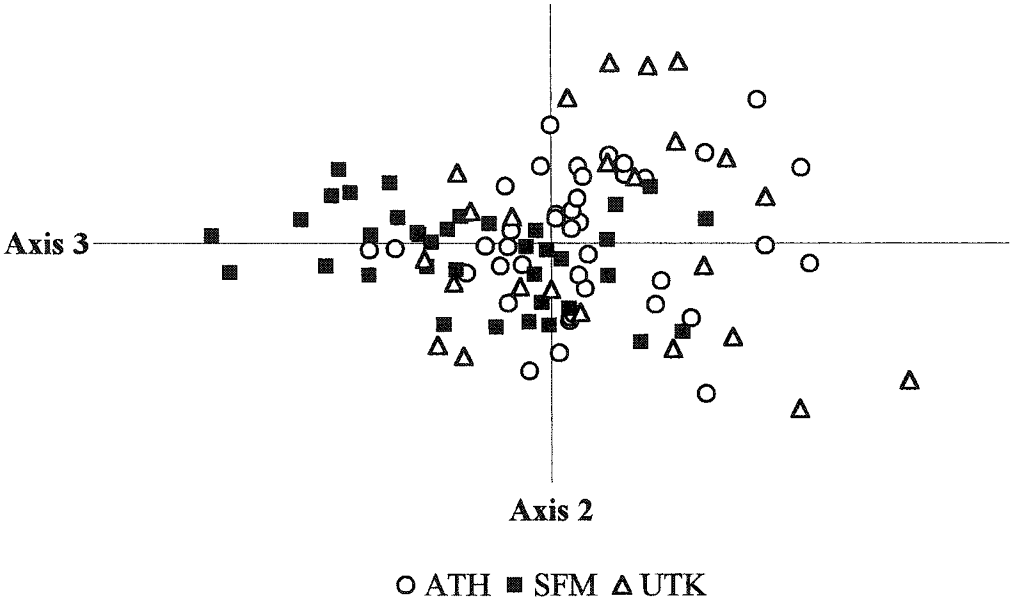

Figure 5b.

Figures 5a and 5b. PCA ordination of all lakes using local and landscape factors. ATH = Athabasca County, SFM = Sustainable Forest Management Network, UTK = Utikuma Lake District.

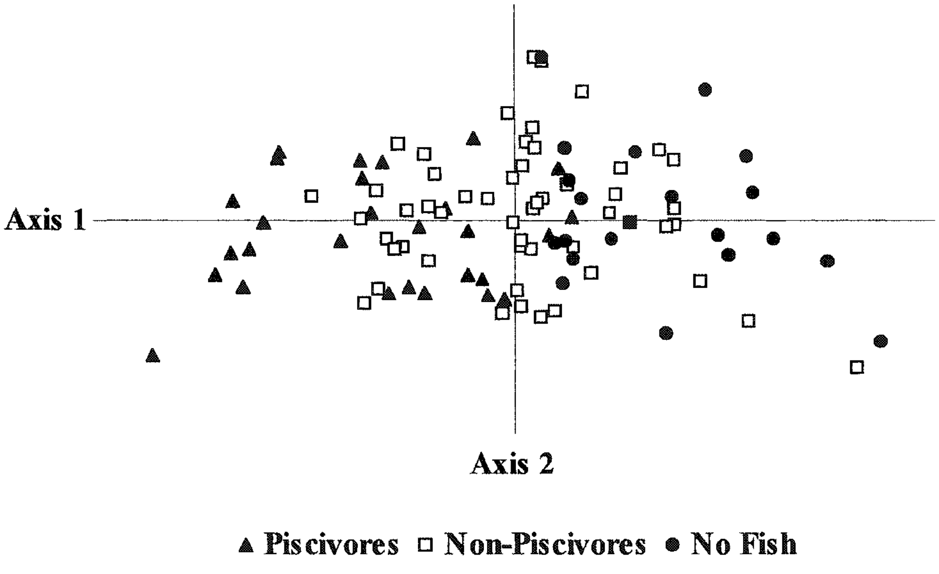

Figure 5c. The first two axes of the PCA ordination of all lakes using local and landscape factors, but classifying lakes on the basis of fish assemblage type.

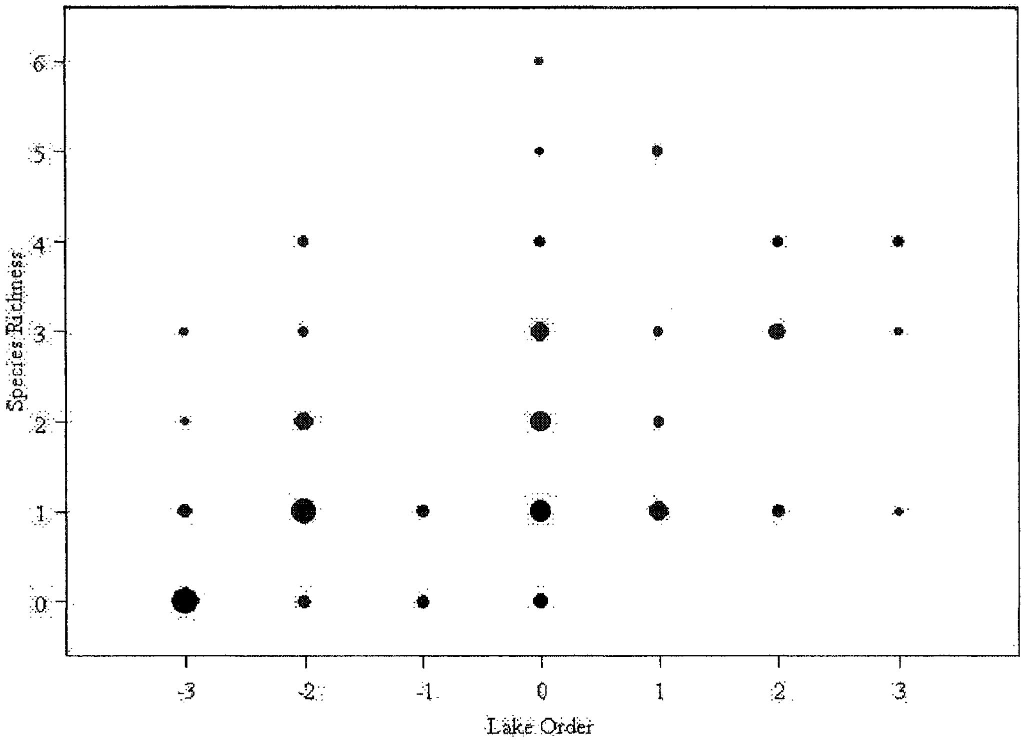

Figure 6. Bubble plot of species richness by lake order. The larger the dot, the greater the number of observations at that combination of lake order and species richness. $P < 0.001$, from Poisson regression.

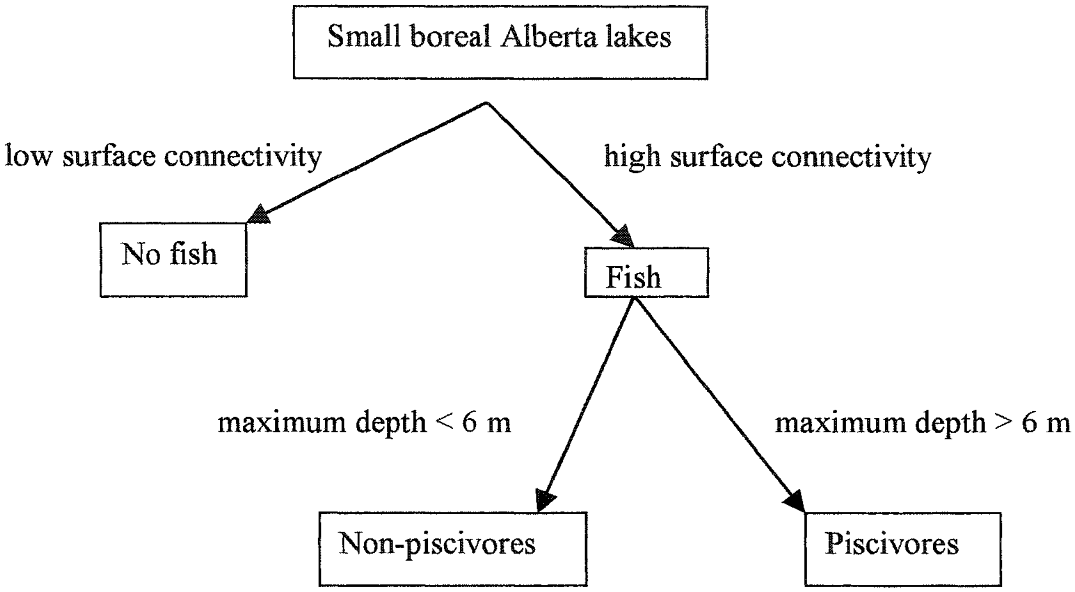

Figure 7. Summary of results from predicting fish assemblage types in small, boreal lakes in Alberta.

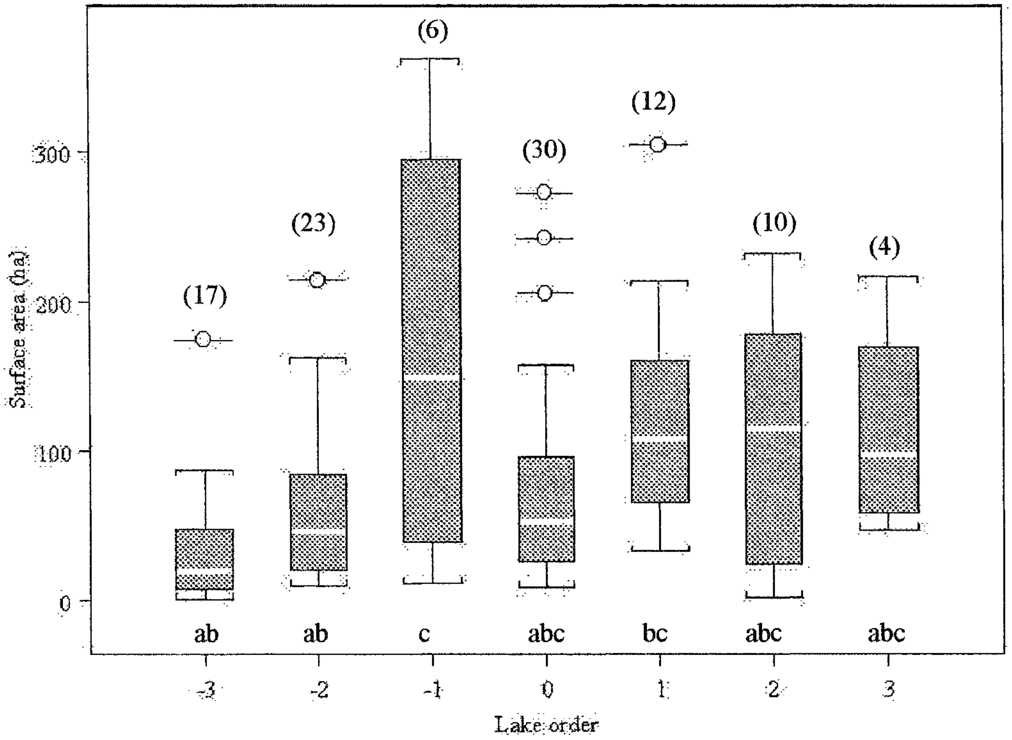

Figure 8a. Box and whisker plot of surface area by lake order. Edges of box are 1st and 3rd quantiles, middle lines are medians, whiskers represent minimums and maximums, circles are outliers. $P = 0.0009$ from one-way ANOVA. Letters indicate differences ($P < 0.05$) from Tukey's Multiple Comparisons.

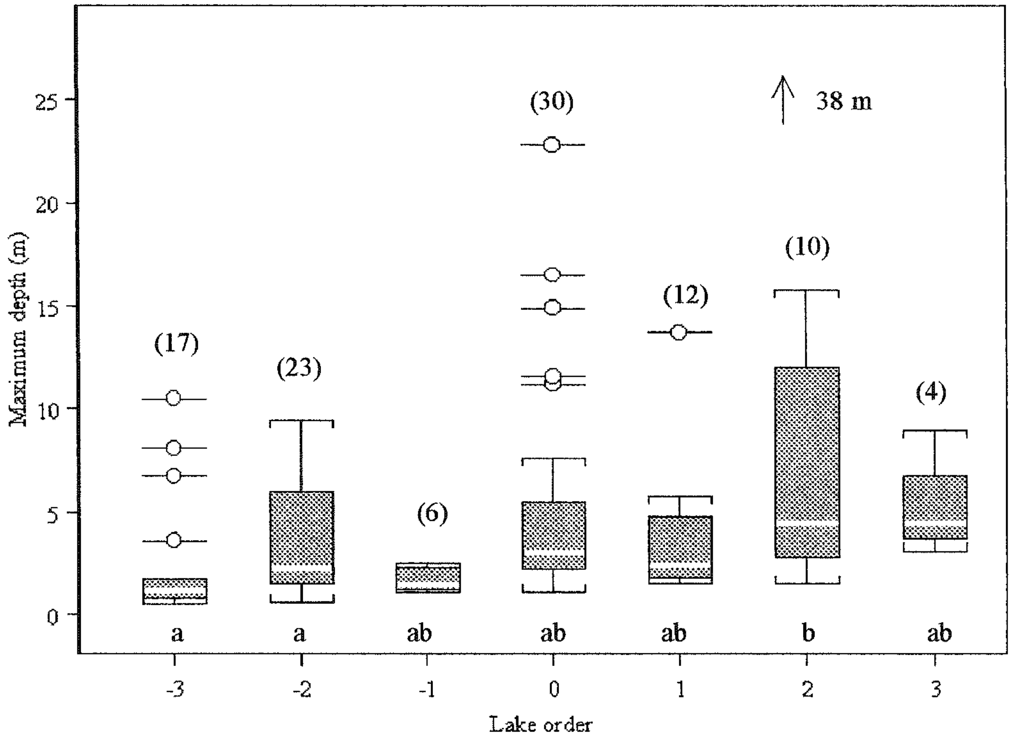

Figure 8b. Box and whisker plot of maximum depth by lake order. Edges of box are 1st and 3rd quantiles, middle lines are medians, whiskers represent minimums and maximums, circles are outliers. $P = 0.02$ from one-way ANOVA. Letters indicate differences ($P < 0.05$) from Tukey's Multiple Comparisons.

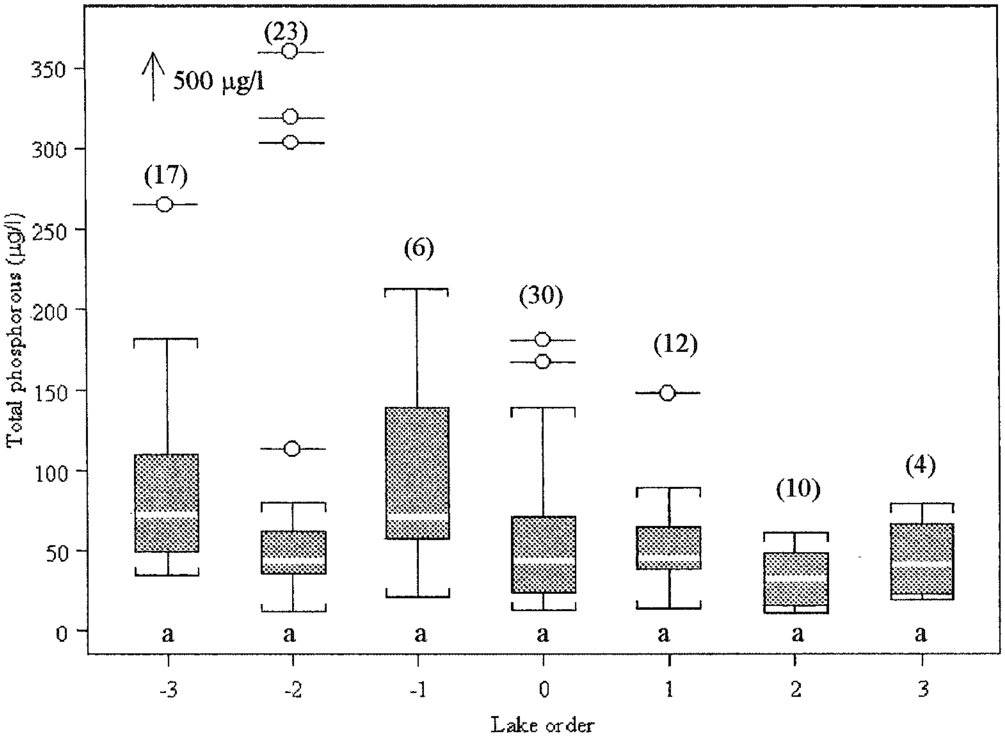

Figure 8c. Box and whisker plot of total phosphorous by lake order. Edges of box are 1st and 3rd quantiles, middle lines are medians, whiskers represent minimums and maximums, circles are outliers. $P = 0.08$ from one-way ANOVA. No differences ($P < 0.05$) from Tukey's Multiple Comparisons.

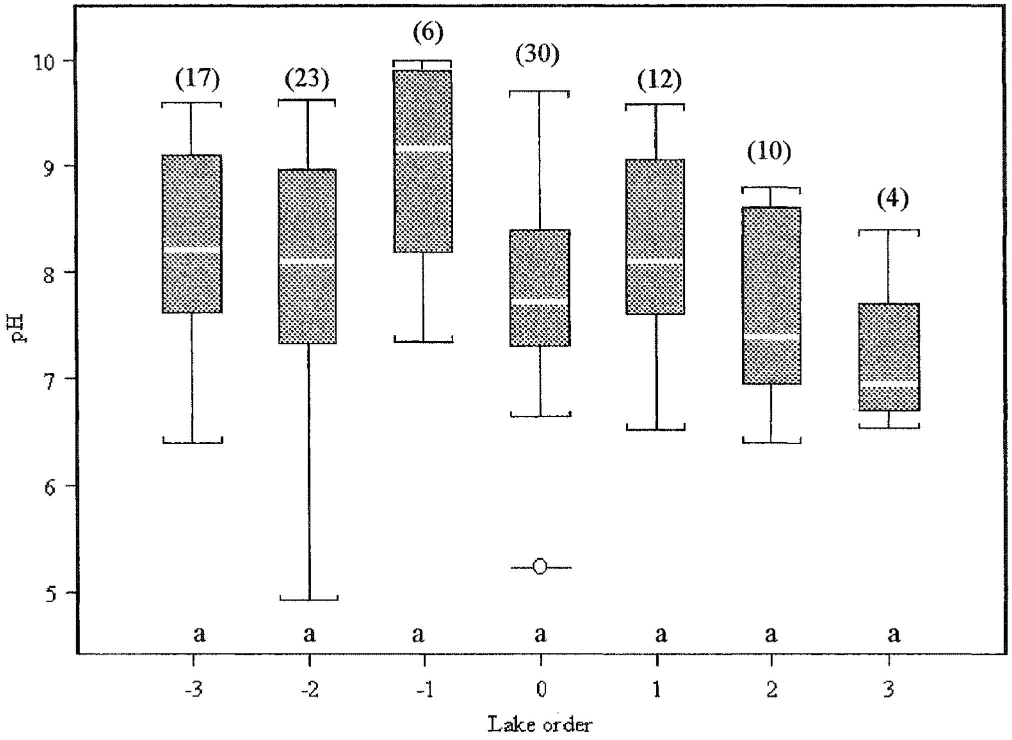

Figure 8d. Box and whisker plot of pH by lake order. Edges of box are 1st and 3rd quantiles, middle lines are medians, whiskers represent minimums and maximums, circles are outliers. $P = 0.045$ from one-way ANOVA. No differences ($P < 0.05$) were detected in Tukey's Multiple Comparisons.

Literature Cited

Anonymous. 1990. Environmental guidelines for access roads and water crossings. Ontario Ministry of Natural Resources. Toronto, Ontario. 64 pgs.

Babin, J., and E. E. Prepas. 1985. Modelling winter oxygen depletion rates in ice-covered temperate zone lakes in Canada. Canadian Journal of Fisheries and Aquatic Sciences **42**: 239-249.

Barica, J. and J. A. Mathias. 1979. Oxygen depletion and winterkill risk in small prairie lakes under extended ice cover. Journal of the Fisheries Research Board of Canada **36**: 980-986.

Burrough, P. A., and R. A. McDonnell. 1998. *Principles of Geographical Information Systems*. Oxford University Press, Oxford, UK. 333 pgs.

Casselman, J. M., and H. H. Harvey. 1975. Selective fish mortality resulting from low winter oxygen. Verhandlungen-Internationale Vereinigung für Theoreische und Angewandte Limnologie **19**: 2418-2429.

Chapleau, F., C. S. Findlay, and E. Szenasy. 1997. Impact of piscivorous fish introductions on fish species richness of small lakes in Gatineau Park, Quebec. Ecoscience **4**: 259-268.

Danylchuk, A. J., and W. M. Tonn. Natural disturbances and fish: local and regional influences on winterkill of fathead minnows in boreal lakes. *In preparation*.

Devito, K. J., I. F. Creed, R. L. Rothwell, and E. E. Prepas. 2000. Landscape controls on phosphorous loading to boreal lakes: implications for the potential impacts of forestry harvesting. Canadian Journal of Fisheries and Aquatic Sciences **57**: 1977-1984.

Diamond, J. M. 1986. Experimental methods in ecology. In *Community Ecology*. Eds. J. M. Diamond and T. J. Case. Harper and Row, New York, New York. Pgs. 3-22.

Eadie, J. M., T. A. Hurly, R. D. Montgomerie, and K. L. Teather. 1986. Lakes and rivers as islands: species-area relationships in the fish faunas of Ontario. Environmental Biology of Fishes **15**: 81-89.

Eadie, J. M., and A. Keast. 1984. Resource heterogeneity and fish species diversity in lakes. Canadian Journal of Zoology **62**: 1689-1695.

Economidis, P. S., E. Dimitriou, R. Pagoni, E. Michaloudi, and L. Natsis. 2000. Introduced and translocated fish species in the inland waters of Greece. Fisheries Management and Ecology 7: 239-250.

Fahrig, L., and G. Merriam. 1985. Habitat patch connectivity and population survival. Ecology **66**: 1762-1768.

Findlay, C. S., D. G. Bert, and L. Zheng. 2000. Effect of introduced piscivores on native minnow communities in Adirondack lakes. Canadian Journal of Fisheries and Aquatic Sciences **57**: 570-580.

Furnis, M. J., T. D. Roelofs, and C. S. Yee. 1991. Road construction and maintenance. In *Influences of Forest and Rangeland Management on Salmonid Fishes and Their Habitats*. Ed. W. R. Meehan. American Fisheries Society Special Publication 19, Bethesda, Maryland. Pgs. 297-323.

Hanski, I. 1999. Habitat connectivity, habitat continuity, and metapopulations in dynamic landscapes. Oikos **87**: 207-219.

Hilborn, R., and M. Mangel. 1997. *The Ecological Detective: Confronting Models with Data*. Monographs in Population Biology 28. Princeton University Press, Princeton, New Jersey. 315 pgs.

Jackson, D. A. and H. H. Harvey. 1989. Biogeographic associations in fish assemblages: local versus regional processes. Ecology **70**: 1472-1484.

Jackson, D. A., P. R. Peres-Neto, and J. D. Olden. 2001. What controls who is where in freshwater fish communities – the roles of biotic, abiotic, and spatial factors. Canadian Journal of Fisheries and Aquatic Sciences **58**: 157-170.

Klinger, S. A., J. J. Magnuson, and G. W. Gallepp. 1982. Survival mechanisms of the central mudminnow (*Umbra limi*), fathead minnow (*Pimephales promelas*), and brook stickleback (*Culaea inconstans*) for low oxygen in winter. Environmental Biology of Fishes 7: 113-120.

Kratz, T. K., K. E. Webster, C. J. Bowser, J. J. Magnuson, B. J. Benson. 1997. The influence of landscape position on lakes in northern Wisconsin. Freshwater Biology **37**: 209-217.

Lafferty, K. D., C. C. Swift, and R. F. Ambrose. 1999. Extirpation and recolonization in a metapopulation of an endangered fish, the tidewater goby. Conservation Biology **13**: 1447-1453.

Langlois, P. W. 2000. *The Influence of Trophic State, Thermal Structure and Winterkill on Littoral Macroinvertebrate Communities in Boreal Plains lakes (Alberta)*. M.Sc. Thesis, University of Alberta, Edmonton, Alberta, 144 pgs.

Lewis, D. B., and J. J. Magnuson. 2000. Landscape spatial patterns in freshwater snail assemblages across Northern Highland catchments. Freshwater Biology **43**: 409-420.

MacKinnon, A. 2000. A spreadsheet for the calculation of comprehensive statistics for the assessment of diagnostic tests and inter-rater agreement. Computers in Biology and Medicine **30**: 127-134.

MacRae, P. S. D., and D. A. Jackson. 2001. The influence of smallmouth bass (*Micropterus dolomieu*) predation and habitat complexity on the structure of littoral zone fish assemablages. Canadian Journal of Fisheries and Aquatic Sciences **58**: 342-351.

Magnuson, J. J. 1990. Long-term ecological research and the invisible present – uncovering the processes hidden because they occur slowly or because effects lag years behind causes. Biosciece **40**: 495-501.

Magnuson, J. J., A. L. Beckel, K. Mills, and S. B. Brandt. 1985. Surviving winter hypoxia: behavioral adaptations of fishes in a northern Wisconsin winterkill lake. Environmental Biology of Fishes **14**: 241-250.

Magnuson, J. J. and T. K. Kratz. 1999. Lakes in the lanscape: approaches to regional limnology. Verhandlungen-Internationale Vereinigung für Theoreische und Angewandte Limnologie **27**: 1-14.

Magnuson, J. J., W. M. Tonn, A. Banerjee, J. Toivonen, O. Sanchez, and M. Rask. 1998. Isolation vs. extinction in the assembly of fishes in small northern lakes. Ecology **79**: 2941-2956.

Mathias, J. A., and J. Barica. 1980. Factors controlling oxygen depletion in ice-covered lakes. Canadian Journal of Fisheries and Aquatic Sciences **37**: 185-194.

Mathsoft, Inc. 1999. S-PLUS 2000 Professional Release 2.

McCullagh, P. and J. A. Nelder. 1989. *Generalized Linear Models 2nd Ed.* Monographs on Statistics and Applied Probability 37. Chapman and Hall/CRC. New York, New York. 511 pgs.

McGarigal, K., S. Cushman, and S. Stafford. 2000. *Mutlivariate Statistics for Wildlife and Ecology Research*. Springer-Verlag, New York, New York. 283 pgs.

Meding, M. E., and L. J. Jackson. 2001. Biological implications of empirical models of winter oxygen depletion. Canadian Journal of Fisheries and Aquatic Sciences **58**: 1727-1736.

Mitchell, P. and E. Prepas. 1990. *Atlas of Alberta Lakes*. The University of Alberta Press, Edmonton, Alberta. 675 pgs.

Nelson, J. S., and M. J. Paetz. 1992. *Fishes of Alberta*. 2nd Ed. The University of Alberta Press, Edmonton, Alberta. 437 pgs.

Nürnberg, G. K. 1995. The anoxic factor, a quantitative measure of anoxia and fish species richness in central Ontario lakes. Transactions of the American Fisheries Society **124**: 677-686.

Olden, J. D., and D. A. Jackson. 2000. Torturing data for the sake of generality: how valid are our regression models? Ecoscience 7: 501-510.

Olden, J. D., D. A. Jackson, and P. R. Peres-Neto. 2001. Spatial isolation and fish communities in drainage lakes. Oecologia **127**: 572-585.

Olden, J. D., D. A. Jackson, and P. R. Peres-Neto. 2002. Predictive models of fish species distributions: a note on proper validation and chance predictions. Transactions of the American Fisheries Society **131**: 329-336.

Ott, R. L. 1993. *An Introduction to Statistical Methods and Data Analysis*. Fourth edition. Duxbury Press, Belmont, California. 1051 pgs.

Paszkowski, C. A., and W. M. Tonn. 2000. Community concordance between the fish and aquatic birds of lakes in northern Alberta, Canada: the relative importance of environmental and biotic factors. Freshwater Biology **43**: 421-437.

Pratt, D. M., W. H. Blust, J. H. Selgeby. 1992. Ruffe, *Gymnocephalus cernuus*: newly introduced in North America. Canadian Journal of Fisheries and Aquatic Sciences **49**: 1616-1618.

Prepas, E. E., D. Planas, J. J. Gibson, D. H. Vitt, T. D. Prowse, W. P. Dinsmore, L. A. Halsey, P. M. McEachern, S. Paquet, G. J. Scrimgeour, W. M. Tonn, C. A. Paszkowski, and K. Wolfstein. 2001. Landscape variables influencing nutrients and phytoplankton communities in Boreal Plain lakes of northern Alberta: a comparison of wetland- and upland-dominated catchments. Canadian Journal of Fisheries and Aquatic Sciences **58**: 1286-1299.

Press, S. J., and S. Wilson. 1978. Choosing between logistic regression and discriminant analysis. Journal of the American Statistical Association **73**: 699-705.

Pringle, C. M. 2001. Hydrologic connectivity and the management of biological reserves: a global perspective. Ecological Applications **11**: 981-998.

Pulliam, H. R. 1988. Sources, sinks, and population regulation. The American Naturalist **132**: 652-661.

Ricciardi, A., and J. B. Rasmussen. 1998. Predicting the identity and impact of future biological invaders: a priority for aquatic resource management. Canadian Journal of Fisheries and Aquatic Sciences **55**: 1759-1765.

Riera, J. L., J. J. Magnuson, T. K. Kratz, and K. E. Webster. 2000. A geomorphic template for the analysis of lake districts applied to the Northern Highland Lake District, Wisconsin, U. S. A. Freshwater Biology **43**: 301-318.

Robinson, C. L. K., and W. M. Tonn. 1989. Influence of environmental factors and piscivory in structuring fish assemblages of small Alberta lakes. Canadian Journal of Fisheries and Aquatic Sciences **46**: 81-89.

Snodgrass, J. W., A. L. Bryan, R. F. Lide, and G. M. Smith. 1996. Factors affecting the occurrence and structure of fish assemblages in isolated wetlands of the Upper Coastal Plain, USA. Canadian Journal of Fisheries and Aquatic Sciences **53**: 443-454.

Soranno, P., K. E. Webster, J. L. Riera, T. K. Kratz, J. S. Baron, P. A. Bukaveckas, G. W. Kling, D. S. White, N. Caine, R. C. Lathrop, and P. R. Leavitt. 1999. Spatial variation among lakes within landscapes: ecological organization along lake chains. Ecosystems **2**: 395-410.

Stacey, P. B., and M. Taper. 1992. Environmental variation and the persistence of small populations. Ecological Applications **2**: 18-29.

Strong, W. L. 1992. *Ecoregions and ecodistricts of Alberta, Volume 1.* Alberta, Forestry, Lands and Wildlife. Edmonton, Alberta. 77 pgs.

ter Braak, C. J. F., and C. W. N. Looman. 1995. Regression. In *Data Analysis in Community and Landscape Ecology*. Edited by R. H. G. Jongman, C. J. F. ter Braak, and O. F. R. Tongeren. Cambridge University Press. Cambridge, UK. Pgs. 29-77.

Tonn, W. M., P. W. Langlois, E. E. Prepas, and A. J. Danylchuk. Winterkill cascade: responses of littoral macroinvertebrates to natural disturbances of fish. *In preparation.*

Tonn, W. M. and J. J. Magnuson. 1982. Patterns in the species composition and richness of fish assemblages in northern Wisconsin lakes. Ecology **63**: 1149-1166.

Tonn, W. M., J. J. Magnuson, M. Rask, and J. Toivonen. 1990. Intercontinental comparison of small-lake fish assemblages: the balance between local and regional processes. American Naturalist **136**: 345-375.

Tonn, W. M., R. E. Vandenbos, and C. A. Paszkowski. 1995. Habitat on a broad scale: relative importance of immigration and extinction for small lake fish assemblages. Bulletin Français de la Pêche et de la Pisciculture **337/338/339**: 47-61.

Trexler, J. C., and J. Travis. 1993. Nontraditional regression analyses. Ecology **74**: 1629-1637.

Venables, W. N., and B. D. Ripley. 1999. *Modern Applied Statistics with S-PLUS.* 3rd Ed. Springer-Verlag, New York, New York.

Webster, K. E., T. K. Kratz, C. J. Bowser, and J. J. Magnuson. 1996. The influence of landscape position on lake chemical responses to drought in northern Wisconsin. Limnology and Oceanography **41**: 977-984.

Wetzel, R. G. 2001. *Limnology*. Third Edition. Academic Press, San Diego, California. 1006 pgs.

Williams, B. K. 1983. Some observations on the use of discriminant function analysis in ecology. Ecology **64**: 1283-1291.

Zimmerman, G. M., H. Goetz, and P. W. Mielke, Jr. 1985. Use of an improved statistical method for group comparisons to study effects of prairie fire. Ecology **66**: 606-611.

Appendix A.

Table 1. Fish species richness, composition, and assemblage type. See methods for details. BRST = brook stickleback (*Culaea inconstans*), FNDC = finescale dace (*Phoxinus neogaeus*), FTMN = fathead minnow (*Pimephales promelas*), IWDR = Iowa darter (*Etheostoma exile*), LNSC = longnose sucker (*Catostomus catostomus*), NRDC = northern redbelly dace (*Phoxinus eos*), NRPK = northern pike (*Esox lucius*), PRDC = pearl dace (*Margariscus margarita*), SPSH = spottail shiner (*Notropis hudsonius*), WHSC = white sucker (*Catostomus commersoni*), YLPR = yellow perch (*Perca flavescens*).

Lake	Species Richness	Species	Assemblage Type
		Athabasca County Data Set	
Bambi	0		Fishless
Beaver	1	FTMN	Non-piscivore
Bilsky	3	BRST, FTMN, WHSC	Non-piscivore
Bobier	1	BRST	Non-piscivore
Canoe	2	BRST, FTMN	Non-piscivore
Chump	5	IWDR, NRPK, SPSH, WHSC, YLPR	Piscivore
Cloudberry	3	BRST, FTMN, WHSC	Non-piscivore
Duggans	2	BRST, FTMN	Non-piscivore
Ghost	3	NRPK, WHSC, YLPR	Piscivore
Gilbert	1	BRST	Non-piscivore
Gods	4	BRST, FTMN, NRPK, WHSC	Piscivore
Grochowski	2	BRST, FTMN	Non-piscivore
Hope	4	IWDR, NRPK, SPSH, YLPR	Piscivore
Hutterite	0		Fishless
Jackfish	1	BRST	Non-piscivore
Jenkins	4	IWDR, NRPK, WHSC, YLPR	Piscivore
Joseph	0		Fishless
Jumping Deer	0		Fishless
Little Buck	0		Fishless
Lodge	2	NRPK, YLPR	Piscivore
Lofty	3	NRPK, SPSH, YLPR	Piscivore
Long Island	2	NRPK, YLPR	Piscivore
Low Chain	2	BRST, FTMN	Non-piscivore
Mid Chain	2	NRPK, YLPR	Piscivore
Mystic	2	BRST, FTMN	Non-piscivore
N. Crooked	3	NRPK, WHSC, YLPR	Piscivore
Narrow	4	BRST, IWDR, NRPK, YLPR	Piscivore
Nelson	2	BRST, FTMN	Non-piscivore
Orrin	2	BRST, FTMN	Non-piscivore
Pike	1	NRPK	Piscivore
Rochester	2	BRST, FTMN	Non-piscivore
S. Crooked	4	IWDR, NRPK, WHSC, YLPR	Piscivore
Schumaker	0		Fishless
Shelly	2	BRST, FTMN	Non-piscivore
Spear	0		Fishless
Tawatinaw	1	BRST	Non-piscivore

Lake	Species Richness	Species	Assemblage Type
Teen	1	NRPK	Piscivore
Tempo	0		Fishless
Tiperary	1	BRST	Non-piscivore
Turkawski	2	BRST, FTMN	Non-piscivore
Two Island	1	BRST	Non-piscivore
Utikuma Lake District Data Set			
U001	1	BRST	Non-piscivore
U002	3	BRST, NRPK, WHSC	Piscivore
U005	1	BRST	Non-piscivore
U017	1	BRST	Non-piscivore
U019	0		Fishless
U026	1	BRST	Non-piscivore
U059	0		Fishless
U060	3	NRPK, WHSC, YLPR	Piscivore
U074a	1	BRST	Non-piscivore
U081	0		Fishless
U102	0		Fishless
U111	0		Fishless
U117	0		Fishless
U121	0		Fishless
U124	0		Fishless
U127	1	BRST	Non-piscivore
U167	0		Fishless
U169	0		Fishless
U201	0		Fishless
U205	0		Fishless
U206	1	BRST	Non-piscivore
U273	1	BRST	Non-piscivore
U791	1	BRST	Non-piscivore
U792	1	BRST	Non-piscivore
Sustainable Forest Management Network Data Set			
N01	1	BRST	Non-piscivore
N02	3	BRST, NRDC, WHSC	Non-piscivore
N03	3	NRDC, NRPK, YLPR	Piscivore
N04	1	NRPK	Piscivore
N05	1	BRST	Non-piscivore
N06	1	BRST	Non-piscivore
N07	1	BRST	Non-piscivore
N08	1	BRST	Non-piscivore
N09	4	BRST, FTMN, PRDC, WHSC	Non-piscivore
N10	2	BRST, PRDC	Non-piscivore
N11	2	NRPK, YLPR	Piscivore
N12	5	BRST, FTMN, IWDR, PRDC, WHSC	Non-piscivore
N13	3	BRST, FTMN, WHSC	Non-piscivore
N14	5	BRST, FTMN, LNSC, PRDC, WHSC	Non-piscivore

Lake	Species Richness	Species	Assemblage Type
N15	6	BRST, FTMN, LNSC, NRPK, PRDC, WHSC	Piscivore
N17	1	NRPK	Piscivore
N18	2	BRST, FTMN	Non-piscivore
N19	3	BRST, FTMN, NRPK	Piscivore
N20	3	FTMN, WHSC, YLPR	Piscivore
N21	0		Fishless
N22	1	NRPK	Piscivore
N24	2	BRST, WHSC	Non-piscivore
N25	1	BRST	Non-piscivore
N26	3	BRST, FTMN, WHSC	Non-piscivore
N27	1	BRST	Non-piscivore
N30	1	BRST	Non-piscivore
N31	1	FTMN	Non-piscivore
N32	1	BRST	Non-piscivore
N33	3	NRPK, WHSC, YLPR	Piscivore
N34	1	FTMN	Non-piscivore
N35	3	BRST, FNDC, FTMN	Non-piscivore
N36	4	BRST, FNDC, FTMN, WHSC	Non-piscivore
N37	3	BRST, FNDC, FTMN	Non-piscivore
N38	3	NRPK, WHSC, YLPR	Piscivore
N39	1	BRST	Non-piscivore
N40	1	NRPK	Piscivore
N41	4	BRST, NRPK, WHSC, YLPR	Piscivore

Appendix A.

Table 2. Local environmental factors. Medians and ranges at bottom. See methods for details.

Lake	Maximum Depth (m)	Surface Area (ha)	Shoreline Development	Slope	pH	Total Phosphorous (µg/L)
Athabasca County Data Set						
Bambi	0.6	10.41	1.16	6.63	9.9	75
Beaver	4.3	2.69	1.32	4.83	7.3	26
Bilsky	9.0	8.32	1.50	2.58	8.1	37
Bobier	1.1	25.21	1.09	3.04	6.4	86
Canoe	2.4	105.04	1.55	3.97	9.3	500
Chump	13.7	304.76	1.76	4.04	8.5	22
Cloudberry	1.5	25.26	1.46	1.02	9.1	80
Duggans	2.4	123.00	1.37	3.20	8.1	37
Ghost	12.0	196.38	1.61	1.56	8.3	34
Gilbert	1.2	12.81	1.02	0.98	7.0	48
Gods	4.5	217.03	1.37	0.60	8.3	43
Grochowski	1.5	54.78	1.16	1.93	7.9	110
Hope	16.5	273.01	1.68	4.79	8.7	16
Hutterite	2.2	27.56	1.38	2.25	8.4	89
Jackfish	6.0	214.69	2.23	1.95	8.2	21
Jenkins	15.8	160.18	2.01	4.64	9.1	53
Joseph	1.5	12.05	1.13	4.88	7.8	14
JumpingDeer	1.7	66.37	1.34	1.83	8.4	21
Little Buck	1.1	51.60	1.45	5.42	9.7	181
Lodge	22.8	242.81	1.97	3.59	8.6	18
Lofty	5.5	70.38	1.29	1.93	9.6	139
Long Island	14.9	205.89	3.05	2.04	8.4	19
Low Chain	9.5	84.22	1.36	2.45	8.6	12
Mid Chain	10.5	87.96	1.28	2.48	8.6	16
Mystic	2.2	14.31	1.89	2.71	7.2	43
N. Crooked	3.1	71.03	1.56	1.82	9.1	50
Narrow	38.0	111.59	2.47	6.35	8.0	11
Nelson	2.8	9.11	1.65	3.22	7.7	360
Orrin	2.6	98.01	1.62	0.89	9.6	72
Pike	4.0	36.80	3.66	3.54	8.0	48
Rochester	1.8	14.81	1.41	2.47	8.2	86
S. Crooked	9.0	123.41	1.62	1.31	8.7	66
Schumaker	6.8	19.96	2.69	3.45	8.2	36
Shelly	1.1	27.97	1.14	1.98	8.0	39
Spear	2.0	9.83	1.35	0.56	9.0	41
Tawatinaw	4.3	47.38	2.95	1.90	8.2	46
Teen	8.4	33.76	1.92	2.26	7.8	28
Tempo	2.5	9.44	1.40	3.38	7.2	108
Tiperary	1.5	76.90	1.36	0.55	8.6	114
Turkawski	1.5	13.81	1.22	1.95	8.1	70
Two Island	1.0	57.84	1.38	1.40	9.1	96

Lake	Maximum Depth (m)	Surface Area (ha)	Shoreline Development	Slope	pH	Total Phosphorous (µg/L)
Utikuma Lake District Data Set						
U001	0.5	5.77	1.76	4.20	7.4	17
U002	3.6	232.35	1.57	4.33	9.6	50
U005	2.8	178.34	1.78	3.03	9.6	73
U017	5.0	114.73	1.88	6.62	8.8	13
U019	1.4	1.70	1.16	4.09	7.8	94
U026	2.3	55.89	1.88	2.57	9.1	167
U059	0.8	28.12	1.30	1.12	8.7	46
U060	4.5	118.14	1.52	1.60	9.0	55
U074a	0.5	3.84	1.17	10.92	8.6	36
U081	3.6	30.32	2.16	4.02	9.5	50
U102	0.5	3.37	1.61	11.20	9.6	51
U111	1.5	9.06	1.28	1.75	7.6	148
U117	1.1	11.19	1.15	1.92	7.3	57
U121	1.0	8.07	1.05	2.39	7.6	42
U124	1.1	65.32	1.11	0.82	9.6	55
U127	1.6	232.64	2.12	0.79	9.5	265
U167	1.1	48.33	1.27	0.44	9.0	319
U169	1.2	362.38	1.34	1.21	9.0	303
U201	2.3	39.76	1.15	4.72	7.1	49
U205	0.8	21.05	1.07	1.39	7.6	182
U206	2.5	12.37	1.92	2.86	7.2	13
U273	1.3	294.90	1.30	0.73	10.0	213
U791	5.1	60.00	1.73	6.31	8.4	21
U792	1.2	13.90	2.67	5.81	7.7	48
Sustainable Forest Management Network Data Set						
N01	1.8	64.59	1.31	4.60	7.3	39
N02	1.5	36.10	1.36	2.30	6.4	30
N03	11.6	63.90	1.25	0.70	7.1	32
N04	5.5	18.49	1.86	3.90	7.0	29
N05	4.6	112.72	1.16	1.40	7.1	30
N06	1.8	93.20	1.26	0.80	4.9	44
N07	3.4	105.91	1.23	0.80	5.2	26
N08	1.8	88.01	1.40	0.50	6.2	56
N09	1.8	162.69	1.10	0.30	6.6	60
N10	2.1	104.19	1.20	0.40	6.5	55
N11	4.0	59.80	1.50	8.00	7.6	44
N12	1.5	199.94	1.32	0.30	6.5	61
N13	2.6	78.22	1.65	1.00	6.9	48
N14	1.8	55.55	1.24	0.40	7.0	79
N15	4.0	157.99	1.15	0.50	6.9	27
N17	7.6	45.47	2.31	2.80	6.8	19
N18	1.8	77.64	1.47	1.10	6.5	25
N19	4.0	25.93	1.68	4.80	7.6	139
N20	1.8	33.53	1.18	3.30	8.7	106
N21	2.6	46.60	1.25	2.50	7.3	38
N22	6.1	124.03	2.35	1.20	7.4	55

Lake	Maximum Depth (m)	Surface Area (ha)	Shoreline Development	Slope	pH	Total Phosphorous (µg/L)
N24	4.0	26.55	1.70	6.60	7.7	30
N25	2.7	52.48	1.64	3.50	7.4	58
N26	5.8	214.22	2.09	1.00	7.4	35
N27	1.5	37.20	1.68	1.70	7.2	57
N30	1.8	60.22	1.16	0.70	7.4	21
N31	5.2	19.75	1.54	4.70	7.2	71
N32	3.2	39.70	1.77	3.30	7.5	87
N33	6.8	96.62	1.38	3.60	7.4	66
N34	11.2	53.26	1.37	3.50	8.1	48
N35	6.5	20.91	1.19	3.00	7.9	25
N36	2.0	113.81	1.40	4.00	9.4	42
N37	2.2	67.89	1.15	1.10	8.2	39
N38	8.1	175.25	1.30	4.50	7.6	24
N39	2.2	46.38	1.42	3.90	8.2	116
N40	8.5	39.08	1.78	3.70	7.9	36
N41	3.6	64.69	1.68	3.90	7.8	62
ATH	2.8 (0.6-38.0)	54.8 (2.7-304.8)	1.40 (1.02-3.66)	2.5 (0.6-6.6)	8.3 (6.4-9.9)	46 (11-500)
SFM	3.2 (1.5-11.6)	63.9 (18.5-214.2)	1.38 (1.10-2.35)	2.5 (0.3-8.0)	7.3 (4.9-9.4)	44 (19-139)
UTK	1.3 (0.5-5.1)	35.0 (1.7-362.4)	1.43 (1.05-2.67)	2.7 (0.4-11.2)	8.7 (7.1-9.9)	53 (13-319)

Appendix A.

Table 3. Landscape position and surface connectivity as measured by lake order, isolation, and connection. See methods for details. LAT = latitude, LONG = longitude.

Lake	LAT (°N)	LONG (°W)	Lake Order	Isolation	Connection
Athabasca County Data Set					
Bambi	54.67	-112.50	-2	1	0
Beaver	54.64	-113.59	2	0	2
Bilsky	54.68	-113.60	2	0	2
Bobier	54.55	-113.18	-2	0	1
Canoe	54.62	-113.13	0	0	2
Chump	54.66	-112.58	1	0	2
Cloudberry	54.74	-113.67	2	0	2
Duggans	54.22	-113.45	1	0	2
Ghost	54.89	-113.61	2	0	2
Gilbert	54.51	-113.17	-2	0	1
Gods	54.39	-113.64	3	0	2
Grochowski	54.47	-113.32	-2	0	1
Hope	54.66	-112.66	0	0	2
Hutterite	54.81	-113.51	0	0	2
Jackfish	54.81	-113.09	-2	0	1
Jenkins	54.92	-113.61	2	0	2
Joseph	54.56	-113.21	-3	1	0
Jumping Deer	54.83	-113.20	-3	1	0
Little Buck	54.66	-112.50	0	1	0
Lodge	54.72	-112.44	0	0	2
Lofty	54.73	-112.48	0	0	2
Long Island	54.47	-113.80	0	1	0
Low Chain	54.97	-113.51	-2	0	1
Mid Chain	54.99	-113.51	-3	1	0
Mystic	54.38	-113.35	0	0	2
N. Crooked	54.93	-113.55	3	0	2
Narrow	54.62	-113.62	2	0	2
Nelson	54.66	-113.64	0	0	2
Orrin	54.42	-113.34	-2	0	1
Pike	54.53	-113.70	1	0	2
Rochester	54.39	-113.32	0	0	2
S. Crooked	54.90	-113.53	3	0	2
Schumaker	54.64	-113.30	-3	1	0
Shelly	54.44	-113.20	-2	1	0
Spear	54.74	-113.16	0	1	0
Tawatinaw	54.34	-113.47	3	0	2
Teen	54.49	-113.72	-2	0	1
Tempo	54.62	-113.32	0	1	0
Tiperary	54.41	-113.23	1	0	2
Turkawski	54.48	-113.14	-2	0	1
Two Island	54.43	-113.22	-3	1	0

Lake	LAT (°N)	LONG (°W)	Lake Order	Isolation	Connection
Utikuma Lake District Data Set					
U001	56.11	-115.62	-3	1	0
U002	56.09	-115.60	2	0	2
U005	56.10	-115.58	2	0	2
U017	56.10	-115.54	1	0	2
U019	56.08	-115.54	-3	1	0
U026	56.08	-115.52	1	0	2
U059	56.07	-115.38	-3	1	0
U060	56.10	-115.36	2	0	2
U074a	56.02	-115.57	-3	1	0
U081	56.03	-115.56	-3	1	0
U102	56.05	-115.46	-3	1	0
U111	56.03	-115.43	-3	1	0
U117	56.05	-115.39	-3	1	0
U121	56.01	-115.35	-3	1	0
U124	56.02	-115.31	-1	0	1
U127	56.01	-115.17	-1	0	1
U167	55.97	-115.23	-3	1	0
U169	55.97	-115.19	-1	0	1
U201	56.12	-115.70	-1	0	1
U205	55.96	-115.16	-3	1	0
U206	56.13	-115.70	-1	0	1
U273	56.00	-115.15	-1	0	1
U791	56.02	-115.56	-2	0	1
U792	56.01	-115.55	-2	0	1
Sustainable Forest Management Network Data Set					
N01	55.95	-110.08	0	0	2
N02	55.94	-110.07	0	0	2
N03	55.86	-110.07	0	0	2
N04	55.73	-110.04	0	0	2
N05	55.82	-110.47	1	0	2
N06	56.21	-111.32	1	0	2
N07	56.17	-111.53	0	0	2
N08	56.46	-112.53	-2	0	1
N09	56.26	-112.40	-2	0	1
N10	56.16	-112.08	1	0	2
N11	55.15	-111.76	0	0	2
N12	56.07	-112.27	1	0	2
N13	55.97	-111.83	0	0	2
N14	55.93	-111.68	0	0	2
N15	55.94	-111.45	0	0	2
N17	55.65	-111.92	0	0	2
N18	55.24	-111.58	-2	0	1
N19	55.16	-111.82	0	0	2
N20	54.99	-113.63	1	0	2
N21	55.73	-114.40	-2	0	1
N22	56.06	-114.91	-2	1	0
N24	56.79	-114.23	0	0	2

Lake	LAT (°N)	LONG (°W)	Lake Order	Isolation	Connection
N25	56.79	-116.32	0	0	2
N26	56.82	-116.45	1	0	2
N27	56.77	-116.43	-2	0	1
N30	56.26	-114.98	2	0	2
N31	55.05	-111.65	-2	0	1
N32	55.17	-111.90	0	0	2
N33	55.20	-111.63	0	0	2
N34	55.20	-113.66	0	0	2
N35	55.09	-113.78	-2	1	0
N36	55.11	-113.73	0	0	2
N37	55.13	-113.72	-2	0	1
N38	55.42	-113.70	-3	1	0
N39	55.35	-113.72	0	0	2
N40	55.40	-113.64	-2	0	1
N41	55.37	-113.64	-2	0	1

Appendix B. List of 1:50 000 National Topographic Survey (NTS) used for each study lake. All NTS sheets were projected in North American Datum 1927.

Lake	NTS Sheet	Edition	Year
Athabasca County Data Set			
Bambi	83 I/10	2	1991
Beaver	83 I/12	3	1992
Bilsky	83 I/12	3	1992
Bobier	83 I/11	3	1991
Canoe	83 I/11	3	1991
Chump	83 I/10	2	1991
Cloudberry	83 I/12	3	1992
Duggans	83 I/03	3	1988
Ghost	83 I/13	3	1992
Gilbert	83 I/11	3	1991
God's (Armstrong)	83 I/05	3	1991
Grochowski	83 I/06	3	1991
Hope	83 I/10	2	1991
Hutterite	83 I/13	3	1992
Jackfish	83 I/14	3	1992
Jenkins	83 I/13	3	1992
Joseph	83 I/11	3	1991
Jumping Deer	83 I/14	3	1992
Little Buck	83 I/10	2	1991
Lodge	83 I/09	2	1991
Lofty	83 I/09	2	1991
Long Island Lake	83 I/05	3	1991
Lower Chain	83 I/13	3	1992
Mid Chain	83 I/13	3	1992
Mystic	83 I/06	3	1991
N. Crooked	83 I/13	3	1992
Narrow	83 I/12	3	1992
Nelson	83 I/12	3	1992
Orrin	83 I/06	3	1991
Pike	83 I/12	3	1992
Rochester	83 I/06	3	1991
S. Crooked	83 I/13	3	1992
Schumaker (Long)	83 I/11	3	1991
Shelly	83 I/06	3	1991
Spear	83 I/11	3	1991
Tawatinaw	83 I/06	3	1991
Teen	83 I/05	3	1991
Tempo	83 I/11	3	1991
Tiperary	83 I/06	3	1991
Turkawski	83 I/06	3	1991

Lake	NTS Sheet	Edition	Year
Two Island	83 I/06	3	1991
Utikuma Lake District Data Set			
U001	84 B/04	2	1989
U002	84 B/04	2	1989
U005	84 B/04	2	1989
U017	84 B/04	2	1989
U019	84 B/04	2	1989
U026	84 B/04	2	1989
U059	84 B/03	2	1989
U060	84 B/03	2	1989
U074a	84 B/04	2	1989
U081	84 B/04	2	1989
U102	84 B/03	2	1989
U111	84 B/03	2	1989
U117	84 B/03	2	1989
U121	84 B/03	2	1989
U124	84 B/03	2	1989
U127	84 B/03	2	1989
U167	83 O/14	3	1989
U169	83 O/14	3	1989
U201	84 B/04	2	1989
U205	83 O/14	3	1989
U206	84 B/04	2	1989
U273	83 O/14	3	1989
U791	84 B/04	2	1989
U792	84 B/04	2	1989
Sustainable Forest Management Network Data Set			
N01	73 M/16	1	1974
N02	73 M/16	1	1974
N03	73 M/16	1	1974
N04	73 M/09	1	1974
N05	73 M/16	1	1974
N06	74 D/03	1	1984
N07	74 D/04	1	1984
N08	84 A/07	1	1978
N09	84 A/08	1	1978
N10	84 A/01	1	1973
N11	73 M/04	1	1974
N12	84 A/01	1	1973
N13	73 M/13	1	1975
N14	73 M/13	1	1975
N15	73 M/14	1	1976
N17	73 M/12	1	1975
N18	73 M/04	1	1974
N19	73 M/04	1	1974

Lake	NTS Sheet	Edition	Year
N20	83 I/13	3	1992
N21	83 O/09	2	1981
N22	84 B/02	2	1989
N24	84 B/16	2	1989
N25	84 C/16	2	1977
N26	84 C/16	2	1977
N27	84 C/16	2	1977
N30	84 B/07	2	1989
N31	73 M/04	1	1974
N32	73 M/04	1	1974
N33	73 M/04	1	1974
N34	83 P/04	3	1979
N35	83 P/04	3	1979
N36	83 P/04	3	1979
N37	83 P/04	3	1979
N38	83 P/05	2	1974
N39	83 P/05	2	1974
N40	83 P/05	2	1974
N41	83 P/05	2	1974

Appendix C. Availability of measured variables for each data set. ATH = Athabasca County Data Set, SFM = Sustainable Forest Management Network Data Set, UTK = Utikuma Lake District Data Set.

Variable	ATH	SFM	UTK
surface area	x	x	x
perimeter	x	x	x
shoreline development	x	x	x
maximum depth	x	x	x
mean depth			x
catchment area	x	x	x
catchment slope	x	x	x
pH	x	x	x
conductivity			x
cations			x
total phosphorous	x	x	x
chlorophyll *a*			x

Appendix D - Details of the logistic and Poisson regression models.

Table 1. Logistic models of fish presence in UTK. Lake order has more than one coefficient because it is a categorical variable with multiple categories; therefore, the number of coefficients equal the number of categories minus one (number of dummy variables).

Model	Variables	Coefficient	P-value
1	intercept	10.74	
	log(maximum depth)	-6.83	0.04
	lake order	8.16	0.02
		-1.91	
		2.56	
		1.51	
2	intercept	-0.81	
	log(maximum depth)	1.58	0.04
	isolation	-1.20	0.03
3	intercept	4.31	
	log(maximum depth)	-1.20	0.04
	connection	1.10	0.02
		4.34	
4	intercept	8.73	
	log(area)	-0.99	0.10
	lake order	7.10	0.01
		-1.54	
		2.26	
		1.42	
5	intercept	1.55	
	log(area)	-1.22	0.10
	isolation	-1.92	0.01
6	intercept	6.11	
	log(area)	-1.42	0.01
	connection	1.66	0.00
		4.55	
7	intercept	7.02	
	lake order	6.85	.01
		-1.78	
		2.16	
		1.30	
8	intercept	-0.15	
	isolation	-1.35	0.00
9	intercept	3.74	
	connection	1.01	0.00
		4.23	
10	intercept	2.83	
	log(total phosphorous)	-3.48	0.12
	log(area)	2.29	0.01

Table 2. Logistic models of fish presence in ATH.

Model	Variables	Coefficient	P-value
1	intercept	1.85	
	log(area)	3.27	0.02
	lake order	1.68	0.03
		-0.01	
		2.78	
		2.17	
		1.01	
2	intercept	-2.87	
	log(area)	2.87	0.02
	isolation	-2.18	0.00
3	intercept	0.78	
	log(area)	2.75	0.02
	connection	6.44	0.00
		-0.82	
4	intercept	4.11	
	log(maximum depth)	4.60	0.02
	lake order	1.80	0.04
		-0.19	
		3.05	
		1.66	
		1.14	
5	intercept	-0.27	
	log(maximum depth)	2.74	0.02
	isolation	-1.88	0.00
6	intercept	3.11	
	log(maximum depth)	2.96	0.02
	connection	6.45	0.00
		-1.06	
7	intercept	6.52	
	lake order	1.30	0.04
		-0.07	
		2.84	
		1.71	
		1.14	
8	intercept	1.40	
	isolation	-1.96	0.00
9	intercept	4.90	
	connection	6.38	0.00
		-0.93	
10	intercept	-1.00	
	slope	-0.52	0.10
	log(area)	1.26	0.03
	log(maximum depth)	3.33	0.07

Table 3. Logistic models of piscivore presence in ATH.

Model	Variables	Coefficient	P-value
1	intercept	-15.49	
	log(maximum depth)	19.71	0.00
	lake order	-3.16	0.06
		0.85	
		0.89	
		-1.18	
		0.54	
2	intercept	-5.52	
	log(maximum depth)	7.14	0.00
	isolation	0.42	0.71
3	intercept	-7.02	
	log(maximum depth)	8.40	0.00
	connection	-1.39	0.21
		0.35	
4	intercept	-8.87	
	log(maximum depth)	6.42	0.00
	log(area)	1.99	0.11
5	intercept	-5.93	
	log(maximum depth)	7.19	0.00

Table 4. Logistic models of piscivore presence in SFM.

Model	Variable	Coefficient	P-value
1	intercept	-5.68	
	slope	0.27	0.06
	log(maximum depth)	6.10	0.00
2	intercept	-5.15	
	log(maximum depth)	6.50	0.00

Table 5. Logistic models of fish presence for all regions combined.

Model	Variable	Coefficient	P-value
1	intercept	0.04	
	log(area)	1.29	0.00
	log(maximum depth)	7.21	0.00
	lake order	2.49	0.00
		-0.20	
		0.46	
		2.32	
		1.46	
		0.84	
2	intercept	-2.32	
	log(area)	0.33	0.00
	log(maximum depth)	5.65	0.00
	isolation	-1.89	0.00
3	intercept	1.55	
	log(area)	0.61	0.00
	log(maximum depth)	4.74	0.00
	connection	1.36	0.00
		4.35	
4	intercept	1.97	
	log(maximum depth)	7.61	0.00
	lake order	2.50	0.00
		0.07	
		0.45	
		2.37	
		1.44	
		0.84	
5	intercept	-1.94	
	log(maximum depth)	5.94	0.00
	isolation	-1.94	0.00
6	intercept	2.27	
	log(maximum depth)	5.30	0.00
	connection	1.48	0.00
		4.35	
7	intercept	6.19	
	lake order	2.09	0.00
		-0.16	
		0.61	
		2.45	
		1.64	
		1.17	

Table 5. Logistic models of fish presence in all regions continued.

Model	Variable	Coefficient	P-value
8	intercept	1.15	
	isolation	-2.03	0.00
9	intercept	4.73	
	connection	1.37	0.00
		4.23	
10	intercept	-5.47	
	log(area)	1.53	0.00
	log(maximum depth)	9.50	0.00
11	intercept	-3.74	
	log(maximum depth)	11.03	0.00

Table 6. Logistic models of piscivore presence in all regions.

Model	Variable	Coefficient	P-value
	intercept	-8.87	
1	log(area)	1.37	0.01
	log(maximum depth)	8.44	0.00
2	intercept	-6.46	
	log(maximum depth)	8.60	0.00

Table 7. Poisson regression models of fish species richness in all regions.

Model	Variable	Coefficient	P-value
1	intercept	-0.09	
	maximum depth	0.01	0.00
	area	0.004	0.00
	lake order	0.55	0.00
		-0.36	
		0.27	
		0.14	
		0.12	
		0.11	
2	intercept	-0.15	
	maximum depth	0.03	0.00
	area	0.002	0.00
	isolation	-0.49	0.00
3	intercept	-0.04	
	maximum depth	0.02	0.00
	area	0.002	0.00
	connection	0.30	0.00
		0.27	
4	intercept	-0.08	
	area	0.004	0.00
	lake order	0.55	0.00
		-0.38	
		0.28	
		0.14	
		0.13	
		0.11	
5	intercept	-0.07	
	area	0.003	0.00
	isolation	-0.50	0.00
6	intercept	-0.006	
	area	0.003	0.00
	connection	0.28	0.00
		0.29	
7	intercept	0.36	
	lake order	0.60	0.00
		-0.18	
		0.26	
		0.17	
		0.14	
		0.12	

Table 7. Poisson models of species richness continued.

Model	Variable	Coefficient	P-value
8	intercept	0.13	
	isolation	-0.58	0.00
9	intercept	0.23	
	connection	0.36	0.00
		0.32	
10	intercept	0.18	
	area	0.004	0.00
11	intercept	0.32	
	maximum depth	0.04	0.00