Development of a prediction and classification system for lake (littoral, SWEPAC$_{LLI}$) and stream (riffle SWEPAC$_{SRI}$) macroinvertebrate communities

by

Richard K. Johnson and Leonard Sandin

Department of Environmental Assessment
Swedish University of Agricultural Sciences
P.O. Box 7050, SE-750 07
Uppsala, Sweden

2001 Rapport 2001:23

1.0 Introduction

In the last two decades, a number of biotic indices have been constructed to evaluate the structural and functional integrity of surface waters using macroinvertebrates (Johnson 1995). Methods range from relatively simple algorithms or biotic indices, to combinations of several indices (i.e. multimetric approaches), to relatively complex, multivariate approaches. The use of multivariate predictive algorithms, often combined with biotic metrics, show much promise as diagnostic tools in biomonitoring and impact assessment studies. Since species occur in characteristic and a limited range of habitats within their geographic range, and tend to be most abundant around their particular environmental optimum, predictive modeling of expected taxa occurrence in the absence of stress is being increasingly used to ascertain reference or ecological target conditions.

A number of modeling techniques are available that can be used to predict taxa occurrence. For example, terrestrial and aquatic ecologists have used generalized linear models (e.g. Nicholls 1989), logistic regression (Agresti 1990), Gaussian logistic regression or the more simplified weighted averaging regression (e.g. ter Braak and Looman 1986), Bayesian models (e.g. Brzeziecki et al. 1995), partial least squares regression (Wold 1982), β-functions (Austin et al. 1994), and taxon-specific (Bio et al. 1998) models to predict taxa occurrence. These generally predict taxa occurrence directly along single environmental gradients. However, species are generally being influenced simultaneously by a number of gradients. The predictive approach developed in the U.K. eloquently solves this dilemma by using discriminant function models that incorporate several environmental factors (Johnson 2000).

Since the early 1980's, discriminant function analysis has been used by ecologists for predicting community structure using sets of environmental data (e.g. Wiegleb 1981). However, Wright et al. (1984) and Moss et al. (1987) were the first to develop predictive models of stream macroinvertebrates using classification and discriminant function analysis (a.k.a. RIVPACS or River InVertebrate Prediction And Classification System). Modeling approaches in general hold much promise in assessment of biodiversity, as they de-emphasize the expertise of the individual investigator (Johnson et al. 1993). Moreover, the predicted taxa occurrence may also

be used to calculate biotic metrics such as BMWP or acidity scores. This is presently being done in the U.K., where observed BMWP scores are compared to predicted scores in national river surveys (e.g. Wright et al. 1988). Wright (1995) provides an overview of the development and application of this procedure as a means of assessing the biological condition of stream ecosystems in Great Britain. Since the advent of RIVPACS (e.g. Wright 1995), this type of predictive approach has gained widespread recognition. RIVPACS-type models have been developed for predicting profundal macroinvertebrate communities of Swedish lakes and used to determine the effects of liming on community composition (Johnson 1995). This technique has also been implemented in Australia as a standard procedure for assessing the biological condition of that nation's running waters (Simpson and Norris 2000), and it has also been evaluated as a means of detecting effects of logging practices on invertebrate assemblages in mountainous streams of California (Hawkins et al. 2000) and the impact of fish on macroinvertebrate communities (Hawkins and Carlisle in press).

Adopting the RIVPACS modeling approach developed in the U.K., the goal of this project was to develop RIVPACS-type algorithms to be used in the prediction of macroinvertebrate communities of lakes (littoral assemblages) and streams (riffle assemblages). We hope that these models will enhance our understanding of the structural and functional aspects of aquatic biodiversity and integrity, as well as supply expected reference conditions for biomonitoring. The project consists of two phases. In the first phase, predictive algorithms were developed using data collected from minimally impaired reference sites. In the second phase, as one of the goals of this study is to develop predictive models that are robust and reliable monitoring and bioassessment tools, steps were taken to test how well these models perform at detecting or discriminating impact.

2.0 METHODS

2.1 Data availability

2.1.1 Field sampling

The data sets used in the calibration of predictive models was taken from the national lake and stream survey done in the autumn of 1995 (e.g. Wilander et al. 1997; Johnson and Goedkoop 2000). A number of factors indicate that the 1995 macroinvertebrate survey of streams and lakes is a robust data set for establishing predictive algorithms of lake and stream macroinvertebrate communities. Firstly, lakes and streams were randomly selected, hence this data set represents an unbiased selection of the countries lake and stream populations. Macroinvertebrate communities of 537 lakes and 696 streams relatively evenly distributed across the country were sampled (Fig. 2.1). Secondly, a number of considerations were taken to reduce the often confounding natural or operator-induced variability of this data set: (i) macroinvertebrate sampling for RI95 was stratified temporally (autumn samples) and spatially (riffle or exposed littoral samples); (ii) samples were collected using standardized kick-sampling with a handnet with a 0.5 mm mesh size (SS EN 27 828, European Committee for Standardization 1994); five kick-samples were taken from each site and pooled to one sample for analysis.

Disclaimer or a cautionary note – As described above, a number of factors were implemented in the national survey to lower within-site variance and increase the probability of detecting change. These factors should be considered when using the SWE[invert]PAC models. Size stratification of lakes included all sizes, but only streams in catchments less than 250 km^2 were sampled. Sampling season, the habitat sampled and the method used to collect macroinvertebrates are known to affect sample composition. The data set used here consisted of *autumn* samples taken from *hard-bottom* (riffle or exposed littoral regions), using standardized *kick-sampling* (i.e. sample effort consisted of five replicate samples taken from streams [1 m x 60 sec] and lakes [1 m x 30 sec]). Further, the ranges of the environmental variables used in the respective models are given below, and we do not recommend the use of these models outside the universe of these model constraints.

2.1.2 Identification

Taxonomic identification was done to the lowest taxonomic unit possible and intercalibrations of taxonomic effort were implemented. The taxonomic resolution used was decided upon by expert opinion. Factors such as easy of identification and ecological value (discriminatory power) were two factors considered. Some 500 taxa were included in the RI95 identification protocol; see Wilander et al. (1997) for a more thorough description of the invertebrates selected and results from the intercalibration of laboratories responsible for the identifications.

Invertebrate abundance data were converted to presence/absence form before model calibration, because we decided from the outset to develop SWEPAC models for predicting only the occurrence of taxa and not their expected abundance. SWEPAC$_{LLI}$ models for lake littoral assemblages were calibrated using two-levels of taxonomic resolution (see section 3.1). Models for predicting stream riffle assemblages were only developed for "species" - level resolution (see section 4.1 below)

2.1.3 Habitat assessment

A substantial amount of environmental data is available for sites sampled in the 1995 national lake and stream survey (Table 2.1). The habitats where samples were collected were classified according to substrate types, such as the presence and proportion of inorganic (e.g. stones) and organic (detritus) substratum and vegetation. Moreover, the riparian zone just adjacent to the sampling sites as well as the whole catchments were classified according to predominant vegetation and land use. Finally, water chemistry such as indicators of acidity (pH and alkalinity) and nutrients (total phosphorus and nitrogen), as well as a number of background physico-chemical data are also available.

Table 2.1. Physico-chemical and other habitat, riparian and catchment scale metrics included in the sampling protocol for the 1995 national survey of lakes and streams.

Geographic	Habitat/Riparian/Catchment	Water chemistry (6)
latitude	substratum (2)	nutrients
longitude	vegetation (3)	pH/alkalinity/exceedence
altitude	riparian (4)	water color
ecoregion (1)	catchment (5)	metals
subecoregion (1)		

(1) 6 ecoregions according to the Nordic Council of Ministers (Anonymous 1984)
(2) habitat classification of substratum, seven classes: block/boulders, cobbles, pebbles, gravel, sand, silt/clay, fine and coarse detritus
(3) habitat classification of vegetation, 8 classes: total coverage (%), emergent, floating-leaved, isoetids, elodeids (fine and broad leaved submerged), *Fontinalis*, other mosses, filamentous algae
(4) riparian classification, 10 classes: coniferous forest, deciduous forest, mixed forest, clear-cutting, heath/grassland, arable land, wetland, alpine, urban/construction, shading
(5) catchment classification, as riparian classification, but also including the percentage of surface water in the catchment and size of the catchment
(6) other: metrics measured for streams but not lakes; slope, velocity, channel width, and depth. Sites were also recorded if affected by liming or point-source pollution.

2.1.4 Minimally disturbed or least impaired sites

The lakes and streams sampled in the 1995 national survey for macroinvertebrates are relatively well distributed across the country, and hence should provide good measures of expected community types (Fig. 2.1). However, these sites are influence not only by natural, but also by human-induced stressors. To obtain measures of communities not substantially influenced by pollution, sites affected by point-sources, acidification, liming and eutrophication or organic pollution, and sampling bias are excluded from the data set (see Johnson 1998). Removal of sites suspected to

be affected by non-natural stressors or sampling error results in a population sampled of 364 lakes (Fig. 2.1b) and 365 streams (Fig. 2.1d). Lastly, preliminary analyses of lake habitats showed that catchments classified as clear-cut (score > 2) had low taxa richness, hence these sites (n = 10 lakes) were removed. The remaining sites will be used to construct predictive models.

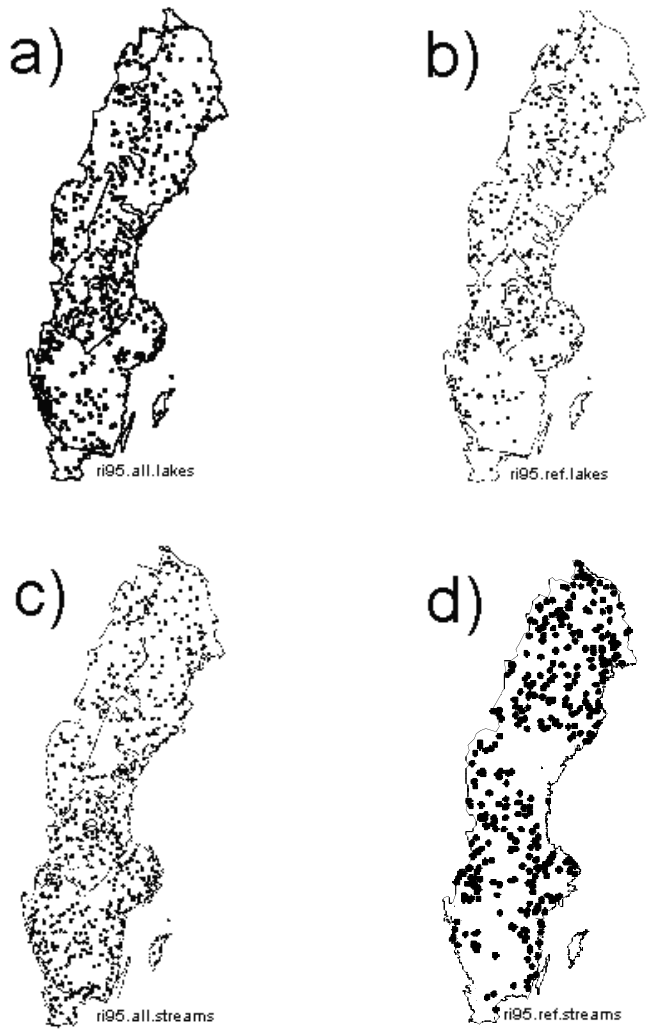

Figure 2.1. Distribution of (a) lakes (n = 537) and (c) streams (n = 696) sampled in the 1995 national survey, and reference (b) lakes (n = 364) and (d) streams (n = 365) used in model calibrations.

2.2 Analyses

RIVPACS-type predicted algorithms for lake-littoral and stream-riffle macroinvertebrate communities were developed by: (i) determination of community types using some form of clustering technique, (ii) determination of explanatory environmental variables using constrained ordination and discriminant function analysis, and (iii) prediction of the probability of taxon occurrence (Table 2.2 and Fig. 2.2). In this last step, the probability that new sites are members of established groups is established and the probability of taxa occurrence at a new site is calculated. Models will be validated using internal (cross-validation) and external (independent sites) procedures.

Table 2.2. Steps taken in the development of SWEPAC

Development and application of a RIVPACS model requires 8 main steps:

1. classification of reference sites into biologically similar groups,
2. development of a discriminant model with data collected from "reference sites" (see below for exclusion criteria to ascertain "reference") to estimate the probabilities of a new site belonging to each of the site groups defined in (1),
3. calculation of the probabilities of all taxa in the regional taxa pool occurring within each reference site group based on presence - absence data,
4. calculation of the probabilities that each taxon will occur at a new site based on (2) and (3),
5. summation of the estimated probabilities of capture of all taxa to estimate the number of taxa expected(*E*) at a new site,
6. calculation of *O/E*,
7. estimation of model error, and
8. assessing the degree of impairment of a new site given the error in the model.

2.2.1 Model calibration

The probabilities of capture of individual taxa was calculated by determining the probabilities of observing each taxon within each of the biological groups. This is simply calculated as the number of sites within a group in which a taxon was observed divided by the total number of sites within a group. Estimates of the probability of observing each taxon are 'weighted' by the probabilities that a new site will belong to each of the groups. A hypothetical example of how taxon weighted probabilities of occurrence are calculated is given in Table 2.3. To predict the macroinvertebrate community at a test site, environmental data for the site are used in the discriminant functions to calculate the probability of membership to each of the groups (here shown as four TWINSPAN groups). To calculate the probability of finding a specific taxa at the test site, the frequency with which each taxon occurs in each group is calculated as the proportion of sites where it occurs divided by the total number of sites in the group. Lastly, the probability of finding a specific taxon is simply the relative group frequency weighted by the probability of membership in each of the groups.

Table 2.3. Hypothetical example of the calculation of taxon occurrence at a test site.

TWINSPAN group	Probability that the test site belongs to group*	Frequency of *taxon a* occurrence in group (%)	Probability that *taxa a* will occur (%)
Grp. 1	0.10	0	0
Grp. 2	0	0	0
Grp. 3	0.70	70	49
Grp. 4	0.20	30	6
Weighted probability of finding taxon a at test site			*55%.*

* *Probability is calculated using a discriminant model.*

Taxon prediction based on community structure

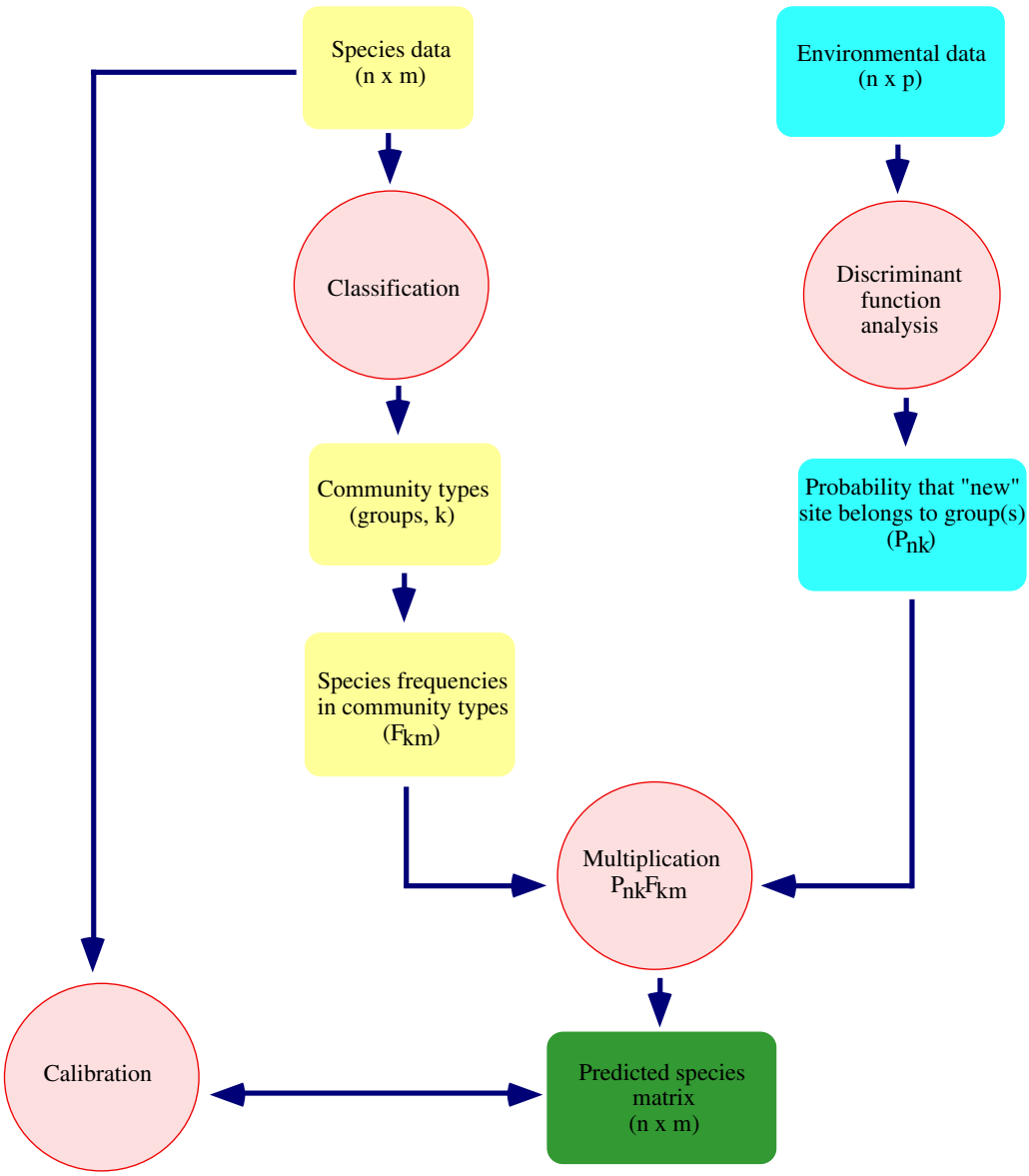

Figure 2.2. Schematic diagram showing the steps used in RIVPACS-type predictions of taxon occurrence.

Table 2.4 shows actual data taken from a site used in the calibration of region 14 models. From the regional species pool, 25 taxa are predicted to occur when an inclusion probability of 25% is used (i.e. this cut-level excludes taxa that are predicted to occur with a probability < 25%). Once a threshold value has been selected, the expected number of taxa (*E*) is expressed as the sum of the probability of occurrence values ≥ the threshold value. The observed number of taxa (*O*) is then calculated by summing the number of taxa predicted to occur at the site that were actually collected. For the example shown in Table 2.4, the expected number of taxa is12.9 at site 590. As mentioned, the observed number of taxa present is not a comparison of raw taxa richness, but is a richness value that is constrained to include only those taxa which are predicted to occur at a site or 13 taxa for site 590. This results in an *O/E* ratio of 1.01 for site 590.

Table 2.4. The number of expected (E), observed (O) and O/E for site 590 sampled in the 1995 national lake survey. An inclusion probability of 0.25 was used. Taxa are ranked in order of their predicted probabilities of occurrence (E).

Taxon	Observed (O)	Probabilities of occurrence (E)
Chironomidae	1	1.0000
Asellus aquaticus L.	1	0.9285
Oligochaeta	0	0.9233
Leptophlebia spp.	1	0.8139
Caenis spp.	1	0.7345
Heptagenia spp.	1	0.7005
Ceratopogonidae	1	0.6851
Sphaeriidae	0	0.6673
Erpobdella spp.	1	0.6224
Cyrnus spp.	0	0.5751
Mystacides spp.	0	0.5295
Limnephilus spp.	0	0.4168
Sialis spp.	0	0.3971
Cloeon spp.	0	0.3941

Ephemera spp.	1	0.3740
Oulimnius spp.	1	0.3697
Limnephilidae	1	0.3572
Turbellaria	0	0.3487
Centroptilum spp.	1	0.3431
Helobdella stagnalis (L.)	0	0.3300
Gyraulus spp.	1	0.3159
Hydracarina	0	0.3024
Polycentropus spp.	1	0.2608
Tinodes spp.	0	0.2558
Lepidostoma hirtum (Fabricius)	0	0.2551
Observed number of taxa	13.00	
Expected number of taxa		12.90
O:E	**1.01**	

Johnson and Goedkoop (2000) showed that spatial stratification of sites should result in more robust predictions. Hence models were developed for the three major biogeographic regions of Sweden. The biogeographic regions have been described by Illies (1966). We amalgamated sites in the nemoral, boreonemoral and southern boreal to fit Illies' region 14, sites in the northern and middle boreal regions were combined to constitute Illies' region 22, and the arctic/alpine complex constituted Illies' region 20 (Fig. 2.3). That ecoregion partitioning of natural variance results in more robust predictive models is also supported by work done in Australia.

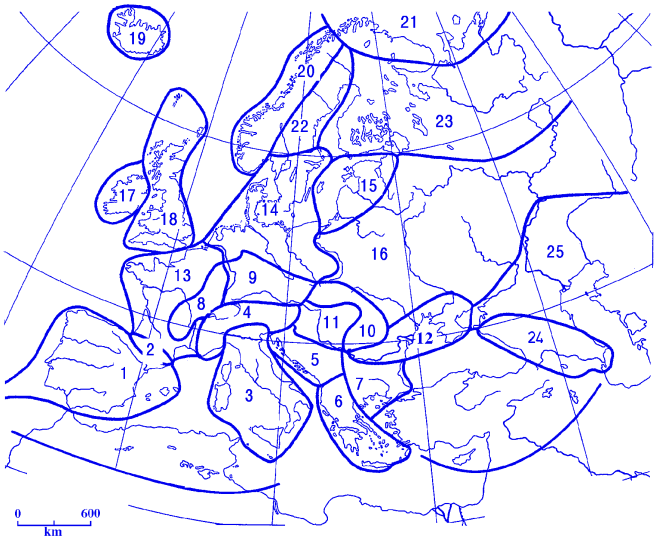

Figure 2.3.Biogeographic regions of Europe.

2.2.2 Classification

Macroinvertebrate communities were classified using both agglomerative (i.e. using Bray-Curtis Index and UPGMA, Unweighted Pair-Groups using Arithmetic Averages) and divisive techniques (using TWINSPAN (Hill 1979, modified by Dr Peter R. Minchin, Feb 1988 - June 1997). TWINSPAN was run using square-root transformed abundance and invoking the downweighting option for rare species. UPGMA was run on macroinvertebrate abundance that was transformed to presence-absence data and using Czekanowski's association measure (Czekanowski, 1913). Although the underlying algorithms of these two clustering techniques are quite different, surprisingly only slight differences were found between the use of TWINSPAN or UPGMA as a grouping technique. The discriminant models showed about the same amount of "correctly" classified sites. This finding was not too surprising since the subjective nature of group identification associated with the classification step is generally not a problem in RIVPACS-type models. Probabilities of capture are weighted by the probability of new sites belonging to each group (see below) and small errors in classification would not therefore result in large errors in predicting the expected number of taxa.

For both the lake and stream models, the classification of the "reference" sites into biologically similar groups based on the presence/absence of taxa was done using TWINSPAN. We used presence/absence data in the classification step since model calibration was focused on predicting taxon occurrence. We checked the structure of group integrity by plotting the TWINSPAN-groups in the ordination space of a Correspondence Analysis based on community composition (presence/absence) data. For stream models the computer program Twindend version 0.4 (unpublished) was also used to evaluate the homogeneity of cluster groups.

2.2.3 Development of a discriminant model

Environmental variables are used to predict site affinity and subsequently the weighted probability of taxon occurrence. A number of steps were taken to obtain a parsimonious subset of environmental variables that best discriminated among sites and between the biological groups (step 1). 1. Ordination methods were used to summarize the main structure of the two species-by-site data sets, and relate structure to environmental factors. Both indirect (e.g. PCA, Hottelling 1933) and direct (canonical correspondence analysis [CCA] and redundancy analysis [RDA], ter Braak 1986 and 1989) forms of gradient analysis were used. However, detrended correspondence analysis showed that gradient lengths were > 2.5 SD indicating that a unimodal model would best fit the species responses. Hence, canonical correspondence analysis was used to determine a subset of environmental variables that best explained the variance in the species data set. CCA ordinations were run invoking the downweighting option of species presence/absence data, forward selection of environmental variables and significance testing of the environmental variables with 999 Monte Carlo permutations. 2. This subset of environmental variables was further examined using discriminant function analysis to determine which variables also discriminated best among the biological groups. 3. An internal test was used to determine the accuracy of the models in discriminating groups. These variables were then used as predictor variables in the discriminant function models.

3.0 CALIBRATION AND VALIDATION OF LAKE (LITTORAL) MACROINVERTEBRATE MODELS

3.1 Taxonomic resolution

SWEPAC$_{LLI}$ lake models were calibrated using two-levels of taxonomic resolution (Appendix 1). "Species"-level resolution consists of 85 taxa and "family"-level resolution consists of 85 families plus other levels of resolution. Taxa included in the species-level data set were those that occurred in more than 3% of the total number of sites, whereas for family-level models all families with at least one observation were included.

3.2"Species" model calibration

Table 3.1 shows the number of biological groups and environmental variables used in the development of SWEPAC$_{LLI}$ "species"-level models.

Table 3.1. Variables used in SWEPAC$_{LLI}$ "species" models.

Region	**No. of groups**	**Environmental variables**
Region 14 (mixed forest)	4	latitude
n= 115 sites		longitude
n = 81 active taxa		log altitude
		cobble
		sand
		coarse detritus
		log chloride
		log color
Region 22 (coniferous forest)	4	latitude
n = 168 sites		longitude
n = 80 active taxa		log altitude
		cobble
		emergent vegetation
		log color

Region 20 (arctic/alpine)	4	latitude
n = 53 sites		longitude
n = 49 active taxa		log altitude
		log total nitrogen
		riparian deciduous
		log color

3.2.1 Region 14 (mixed forest region)

Eleven variables were selected in the CCA ordination of 115 sites and 81 "active" species of region 14. These 11 variables explained 18.5% of the total variance (unconstrained eigenvalues) and 37.2% of the explained variance. Of these 11 variables, eight were found to discriminate among the four TWINSPAN groups, and the discriminant model developed from these eight predictor variables was highly significant (Wilk's lamda = 0.501, F = 3.379, P < 0.001). Latitude (3.9%), longitude (5.0%), altitude (14.7%), cobble (22.7%), sand (8.2%), coarse detritus (19.6%), chloride (9.0%) and color (13.5%) discriminated among the four groups (Fig. 3.1). The first discriminant factor had an eigenvalue of 0.587, which described a substrate gradient (color and cobble substratum accounted for much of the variability). The second axis (eigenvalue = 0.223) was seemingly related to climate, with altitude and chloride strongly loading on this axis. 52.2% of the sites were classified to the "correct" group.

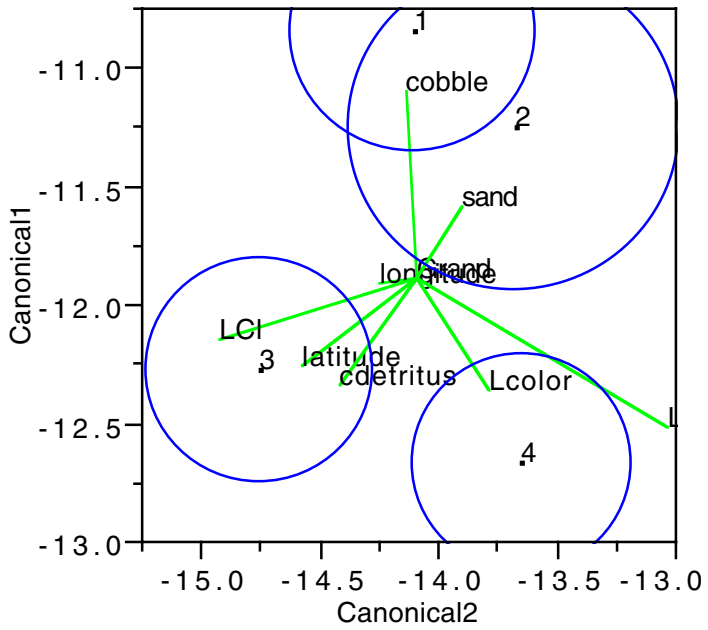

Figure 3.1. Centroid plot of TWINSPAN groups and environmental variables for the region 14 species-level model. The numbers show the centroid for each group and circles denote a 95% CI.

3.2.2 Region 22 (coniferous forest region)

Nine variables were selected in the CCA ordination of 168 sites and 80 "active" species of region 22. These nine variables explained 13.1% of the total variance (unconstrained eigenvalues) and 32.7% of the explained variance. Six variables were found to discriminate among the four TWINSPAN groups, and the discriminant model developed from these six predictor variables was highly significant (Wilk's lamda = 0.425, F = 8.8429, P < 0.001). Latitude (14.4%), longitude (10.6%), altitude (10.5%), cobble (25.1%), emergent vegetation (22.4%) and color (26.0%) discriminated among the four groups (Fig. 3.2). The first discriminant factor had an eigenvalue of 0.839, which described a substrate gradient (color and cobble substratum accounted for much of the variability). The second axis (eigenvalue = 0.171) was also related to substrate (emergent vegetation). 58.9% of the sites were

classified to the "correct" group. Group 3 had a large number of misclassifications, only 42.2% (64 sites) of were classified correctly. Group 4, on the other hand, had a high number of correct classifications (84.6% were classified correctly of n = 26 sites).

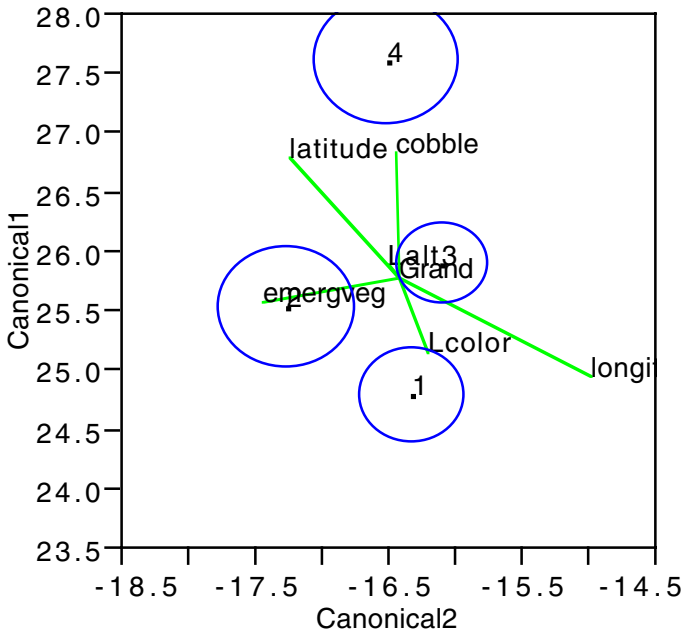

Figure 3.2. Centroid plot of TWINSPAN groups and environmental variables for the region 22 species-level model. The numbers show the centroid for each group and circles denote a 95% CI.

3.2.3 Region 20 (arctic/alpine region)

Eight variables were selected in the CCA ordination of 53 sites and 49 "active" species of region 20 (appendix 1). These eight variables explained 24.7% of the total variance and 27.7% of the explained variance. Six variables were found to discriminate among the four TWINSPAN groups, and the discriminant model

developed from these six predictor variables was highly significant (Wilk's lamda = 0.350, F = 3.1167, P < 0.001). Latitude (22.1%), longitude (18.0%), altitude (7.5%), total nitrogen (17.0%), riparian deciduous (17.4%) and color (26.9%) discriminated among the four groups (Fig. 3.3). The first discriminant factor had an eigenvalue of1.11, and was related to altitude. The second axis had an eigenvalue of 0.265; total nitrogen and riparian deciduous were strongly loaded on this axis. 56.6% of the sites were classified to the "correct" group. Groups 2 and 3 had a large number of misclassifications (only 28.6 and 45.5% of 14 and 11 sites, respectively, were classified correctly). Group 1 (n = 12 sites) had a high number of correct classifications (91.7% were classified correctly).

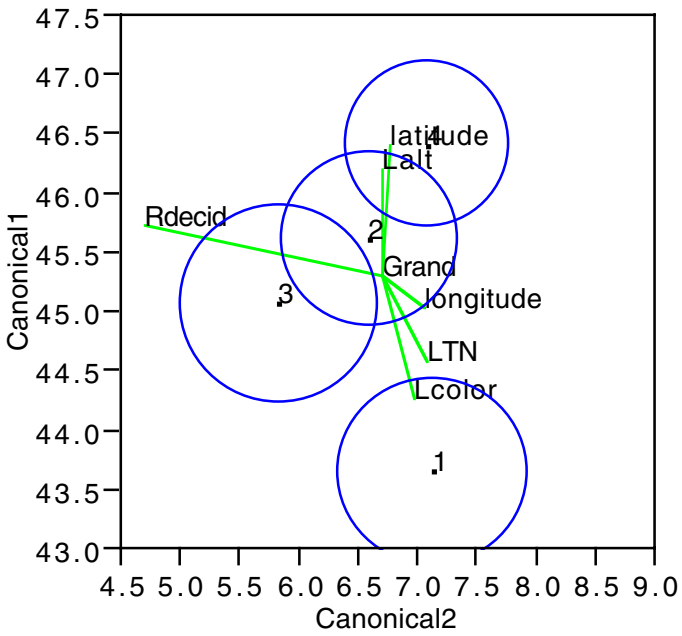

Figure 3.3. Centroid plot of TWINSPAN groups and environmental variables for the region 20 species-level model. The numbers show the centroid for each group and circles denote a 95% CI.

3.3 "Family" level model calibration

Similar to species models, three models (one for each geographic region) were developed using "family" level taxonomic resolution. Table 3.2 shows the number of biological groups and environmental variables used in the development of $SWEPAC_{LLI}$ "family"-level models.

Table 3.2. Variables used in $SWEPAC_{LLI}$ "family" models.

Region	**No. of groups**	**Environmental variables**
Region 14 (mixed forest)	4	latitude
n = 115 sites		longitude
n = 74 active families		log altitude
		coarse detritus
		sand
		log surface area
		log chloride
Region 22 (coniferous forest)	4	latitude
n = 167 sites		longitude
n = 69 active families		log altitude
		cobble
		log color
Region 20 (arctic/alpine)	3	latitude
n = 53 sites		longitude
n = 44 active families		log altitude
		emergent vegetation
		riparian deciduous

3.3.1 Region 14 (mixed forest region)

Ten variables were selected in the CCA ordination of 115 sites and 74 "active" families of region 14. These ten variables explained 17.3% of the total variance and 34.3% of the explained variance. Of these ten variables, seven were found to discriminate among the four TWINSPAN groups, and the discriminant model developed from these seven predictor variables was highly significant (Wilk's lamda = 0.333, F = 6.7055, P < 0.001). Latitude (8.7%), longitude (5.3%), altitude (17.6%), coarse detritus (27%), sand (25.7%), lake surface area (28.3%) and chloride (19.0%) discriminated among the four groups (Fig. 3.4). The first and second discriminant factors had an eigenvalues of 0.979 and 0.349, respectively. Chloride and lake surface area were strongly loaded on the first axis, while sand and altitude were strongly correlated with the second axis. 64.3% of the sites were classified to the "correct" group; classifications ranged from 57.1% for group 1 to 68.1% for group 3.

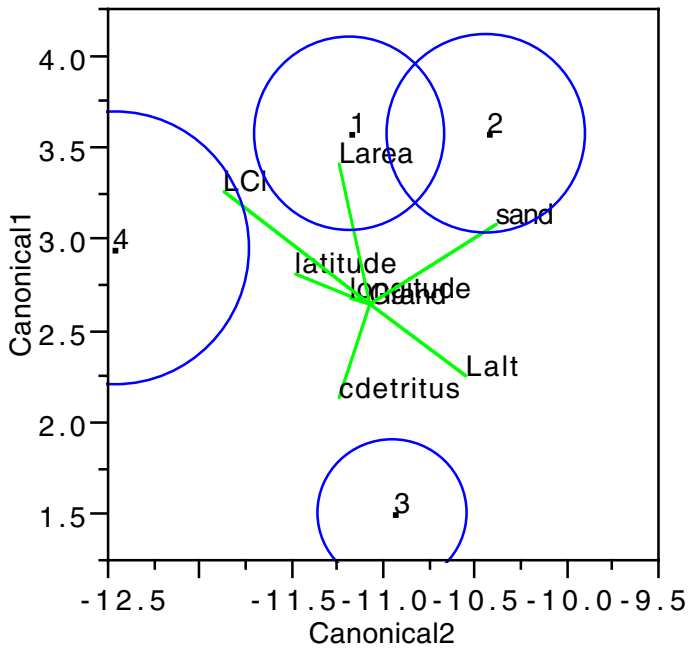

Figure 3.4. Centroid plot of TWINSPAN groups and enviromental variables for the region 14 family-level model. The numbers show the centroid for each group and circles denote a 95% CI.

3.3.2 Region 22 (coniferous forest region)

Ten variables were selected in the CCA ordination of 167 sites and 69 "active" familes of region 22. These ten variables explained 14.4% of the total variance and 37.2% of the explained variance. Five variables were found to discriminate among the four TWINSPAN groups, and the discriminant model developed from these five predictor variables was highly significant (Wilk's lamda = 0.570, F = 6.6119, P < 0.001). Latitude (8.9%), longitude (7.6%), altitude (3.2%), cobble (20.75.1%), and color (25.7%) discriminated among the four groups (Fig. 3.5). The first discriminant factor had an eigenvalue of 0.609, which described a color and substrate (cobble substratum) gradient. The second axis (eigenvalue = 0.0705) was related to altitude. 52.1% of the sites were classified to the "correct" group. Groups 2 and 3 high numbers of misclassifications; 38.8%, n = 85 sites and 42.9%, n = 9 sites, were classified correctly. Both groups 1 and 4 had correct classifications of 73.5% and 75.0%, respectively.

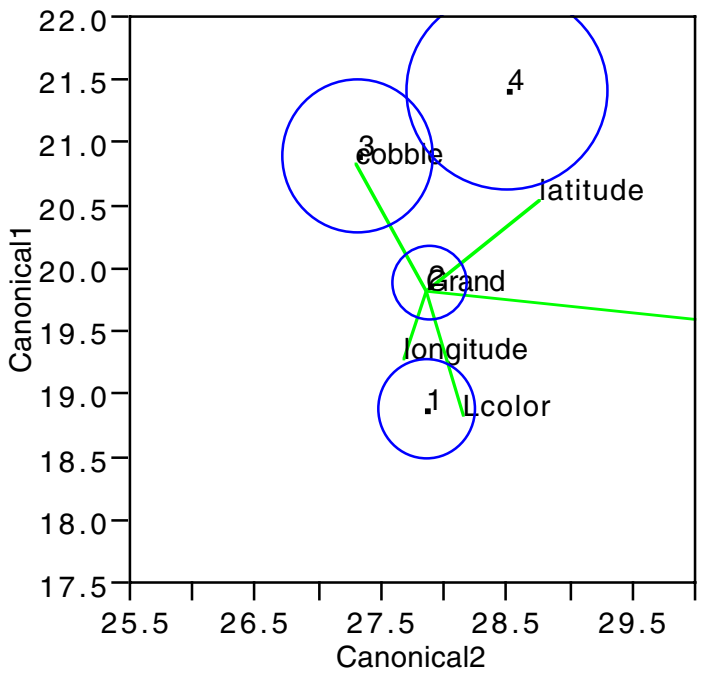

Figure 3.5. Centroid plot of TWINSPAN groups and environmental variables for the region 22 family-level model. The numbers show the centroid for each group and circles denote a 95% CI.

3.3.3 Region 20 (arctic/alpine region)

Nine variables were selected in the CCA ordination of 53 sites and 44 "active" species of region 20. These nine variables explained 28.1% of the total variance and 31.3% of the explained variance. Five variables were found to discriminate among the three TWINSPAN groups, and the discriminant model developed from these five predictor variables was highly significant (Wilk's lamda = 0.300, F = 7.5964, P < 0.001). Latitude (7.9%), longitude (3.9%), altitude (24.4%), emergent vegetation (21.3%) and riparian deciduous (24.5%) discriminated among the three groups (Fig. 3.6). The first discriminant factor had an eigenvalue of 1.67, and was related to altitude. The second axis had an eigenvalue of 0.247; emergent vegetation was strongly (positively) correlated with this axis. 79.2% of the sites were classified to the "correct" group. Group 1 had the highest percent of correct classifications (88.9%), followed by group 3 (70.6%) and group 2 (66.7%).

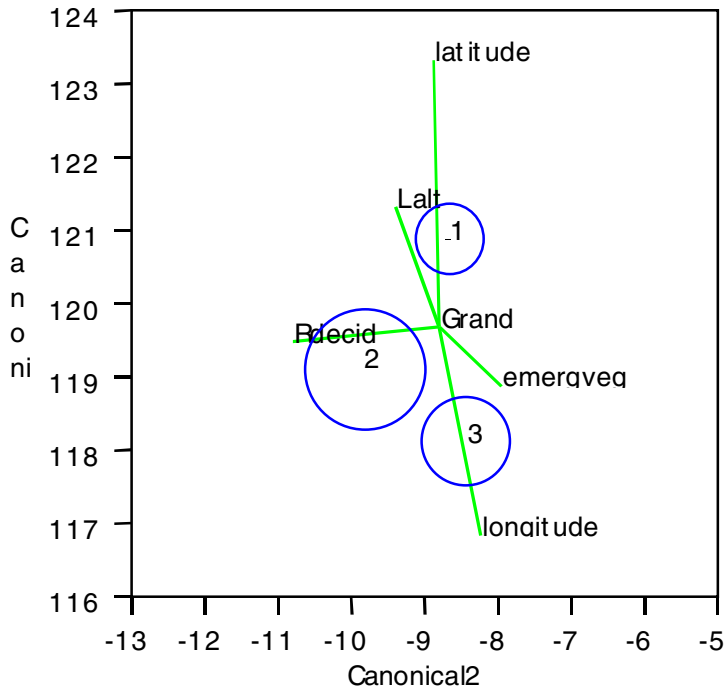

Figure 3.6. Centroid plot of TWINSPAN groups and environmental variables for the region 20 family-level model. The numbers show the centroid for each group and circles denote a 95% CI.

3.4 Inclusion probability

Before an *O/E* ratio can be calculated one must first decide on what threshold probability will be used in the calculations. RIVPACS calculates *O/E* using all probabilities > 0; based on the assumption that using "rare" taxa provides a more representative assessment. However, other RIVPACS-type models (e.g. AUSRIVAS used in Australia) calculates *O/E* using only those taxa that are predicted to occur with probabilities > 50%. Here the reasoning was that rare taxa are often "accidentals" and hence contribute no reliable information to the assessment. In other words, if these taxa are from viable populations, they cannot be modeled accurately and thus contribute mainly to model error. Two empirical analyses support this latter contention (i.e. Simpson and Norris 2000; Hawkins et al. 2000). Cao et al. (1998) and Cao and Williams (1999) contend, however, that the inclusion of rare taxa in bioassessment increases statistical power.

To determine what threshold value to use in SWEPAC$_{LLI}$ models we compared the results from a number of calibration runs. For both "species" and "family" models the number observed and expected taxa increased as the inclusion threshold decreased (Fig. 3.7). As expected, the lowest values were found in region 20 or the arctic/alpine complex presumably due to the impoverished regional species pool and/or harsher environments. Models for region 14 and 22 were similar at the species level, but family level models for region 22 had lower richness than those for region 14.

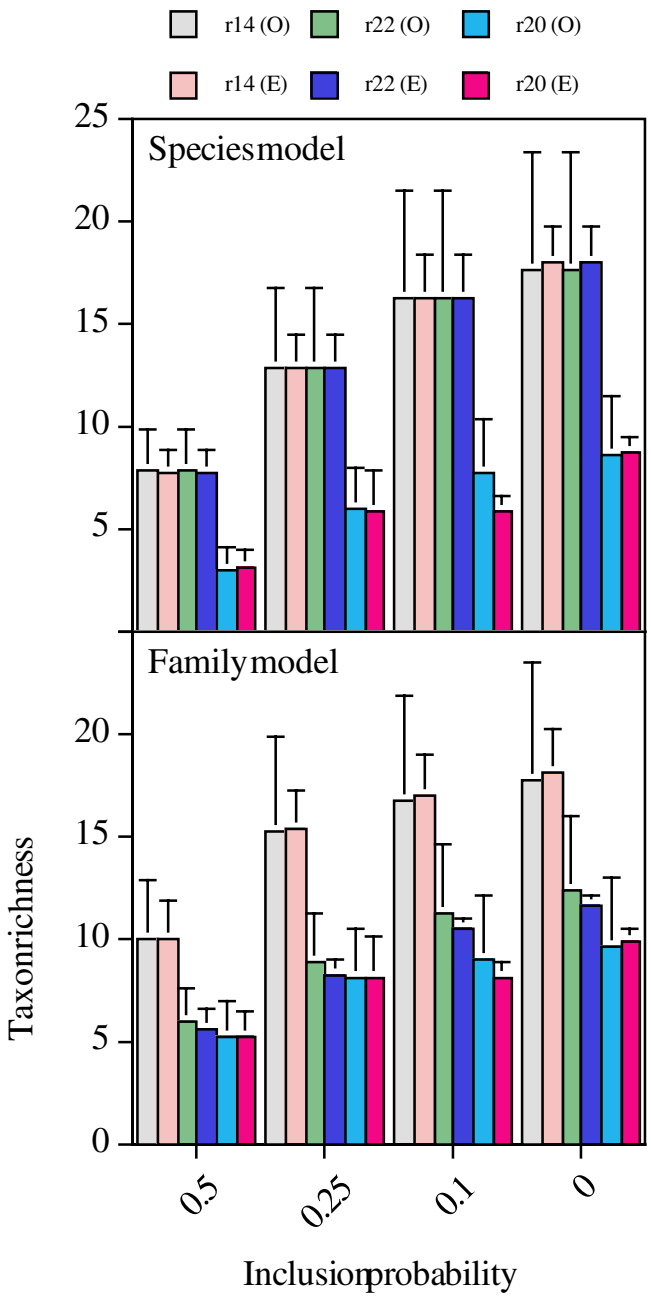

Figure3.7. Means ± standard deviations for observed (O) and expected (E) number of taxa from reference sites. Probability of occurrence thresholds used were 0.5, 0.25, 0.10 and 0.

A comparison of inclusion probability thresholds and *O/E* ratios showed only slight differences among regions and with varying threshold values (Table 3.3). Threshold levels of 50% always had lower variance (standard deviations) than other threshold levels (i.e. thresholds >50%). Species-level models had *O/E* ratios which ranged from 0.98 – 1.01 for region 14; 1.01 – 1.02 for region 22 and 0.99 – 1.01 for region 20. Somewhat greater variability was noted for family-level models in region 22.

Table 3.3. Means ± standard deviations for O/E values from reference sites. O/E_x refers to the probability of occurrence thresholds (i.e. ≥ 0.50, 0.25, 0.10 and 0).

Model	O/E_{50}	O/E_{25}	O/E_{10}	O/E_0
Region 14 species	1.01±0.21	1.00±0.25	0.99±0.28	0.98±0.28
Region 14 family	1.00±0.22	1.00±0.26	0.99±0.28	0.98±0.29
Region 22 species	1.00±0.22	1.02±0.28	1.01±0.32	1.01±0.34
Region 22 family	1.05±0.23	1.08±0.28	1.08±0.31	1.06±0.31
Region 20 species	0.99±0.27	1.01±0.34	1.01±0.33	0.99±0.31
Region 20 family	0.99±0.27	0.99±0.31	0.99±0.33	0.99±0.33

For lake models we decided to use a threshold level of 25% for $SWEPAC_{LLI}$ models. This decision was based on calibration runs comparing results of observed, expected and *O/E* ratios using thresholds of 0, 10, 25 and 50% probabilities, as well as a compromise between models that predicted fewer taxa but had lower variance associated with these predictions (e.g. models using a 50% threshold level) and models that predicted many taxa but had high variance (e.g. models using a 10% threshold level).

For the remainder of this study, model validations were only run for species-level models and using a threshold value (inclusion probability) of 25%.

3.5 Model validation

The error associated with the SWEPAC$_{LLI}$ models was estimated using two methods. First, the distribution of the reference data set (internal error validation) gives a measure of the error distribution (see Table 3.3). However, a more rigorous measure of model error is made by treating data from a series of reference sites that were not used in model construction as test data. Before model calibration, 5% of the reference sites were randomly removed from the data set. Applying the calculations in steps 3 – 6 (Table 2.2) these sites provide a second and independent measure of model error. The distribution of *O/E* values should exhibit a near-normal frequency distribution with values centered on 1, and the spread of these *O/E* values represents model error.

Internal validation showed that error estimates (here expressed as the 95% CI) for *O/E* values were for region 14; mean = 0.997, upper CL = 1.044 and lower CL = 0.950, for region 22; mean = 1.016, upper CL = 1.059 and lower CL = 0.973, and for region 20; mean = 1.011, upper CL = 1.104 and lower = 0.9185. Bootstrapping and Efron's percentile (Efron 1979) of *O/E* values by region was used to confirm these error estimates. Bootstrapped estimates obtained from 1000 iterations were similar to those obtained by conventional parametric statistics; for region 14; mean *O/E* = 0.997, upper CL = 1.037 and lower CL = 0.958, for region 22; mean *O/E* = 1.015, upper CL = 1.051 and lower CL = 0.977, and for region 20; mean *O/E* = 1.009, upper CL = 1.084 and lower = 0.9347 (Fig. 3.8). Bootstrapped error estimates will be used here after, unless mentioned otherwise.

Although the number of sites used for external validation are too few for obtaining robust estimates of model error (5% of the total number of reference sites), these sites provide nonetheless some measure of expected model error. Not surprisingly, variance estimates are much larger for external-reference than the internal-reference validations. No differences were found between internal and external validations for regions 14 and 22, but region 20 had somewhat higher values (Fig. 3.8). These differences are presumably due to the few number of sites used in this validation step. For region 14, six sites were used in the second validation step (mean *O/E* = 1.087,

upper CL = 1.259 and lower CL = 0.898). Eight sites were used in the external validation of region 22 sites (mean *O/E* = 0.946, upper CL = 1.083 and lower CL = 0.789). For region 20 only three sites were used; (mean *O/E* = 1.449, upper CL = 1.821 and lower CL = 1.179). Lastly, sites that were considered as impacted and not included in model calibration had lower *O/E* values than reference. This discriminatory power is described in greater detail below.

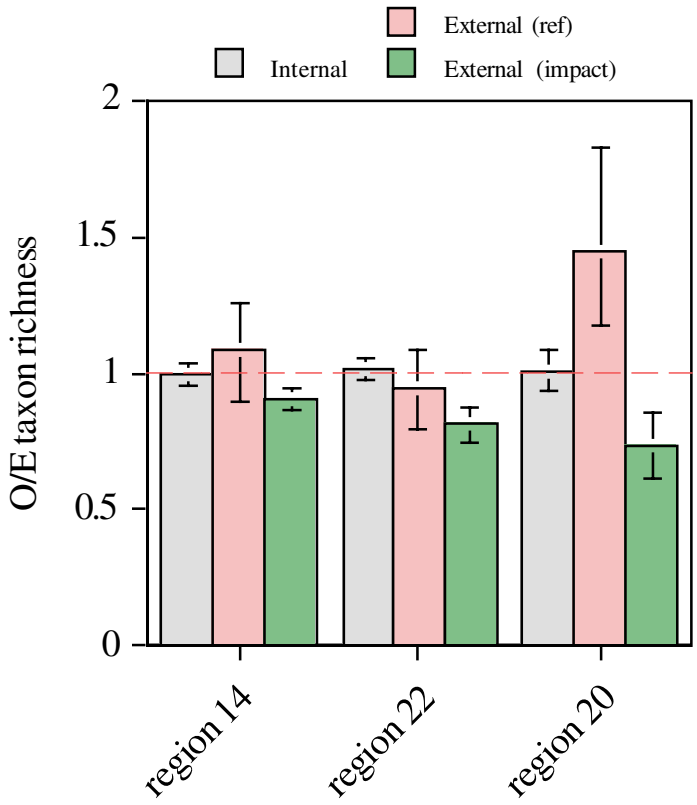

Fig. 3.8. Bootstrapped O/E values and error estimates of internal (reference or calibration data set), external (reference or 5% of original data set), and external (sites deemed to be impacted and hence removed before the model calibration) validations. Mean O/E values ± 95% confidence interval.

3.5 Assessment of ecological quality

Inferences of impairment need to be made in context of model error. Ideally threshold levels to be set with *a priori* knowledge of type I (false positive) and II (false negative) error estimates. A banding scheme is commonly used to classify degrees of impairment. The number of bands ideally should be broached by determining the error associated with estimates of expected (see Clarke et al. 1996). A number of studies have, however, shown that four bands can often be reliably distinguished for models that produce standard deviations of reference site *O/E* values of about 0.2. Further, since one of the goals of this project is to develop a classification scheme that is harmonized with the EU Water Framework Directive (EU WFD), we decided to create a banding scheme based on 4 bands that lie outside the threshold value. In other words, a total of five quality classes are recognized. Adopting the terminology of the EU WFD, these five quality classes would range from high to poor ecological quality, with high consisting of sites that are (i) within the 95% CI or (ii) within the 10 and 90 percentile distribution. Simply put, sites above the threshold value (e.g. the lower 95% CL) are deemed as not deviating significantly from reference.

Both parametric (95% CI) and non-parametric (percentile distribution) were used to established a five-band ecological quality classes. Values below the lower 95% CL or 10 percentile were used to create four impairment-bands (good, moderate, poor and bad). This was done, for example, simply by dividing the interval from *O/E* = zero to *O/E* = threshold value (lower 95% CL) into a series of equal length bands. Sites with *O/E* values above the upper 95% CL (i.e. sites with more taxa than expected) are placed in a separate band (denoted as X). On the one hand, it has been argued that large *O/E* values may occur if the effect of certain habitat features were not captured by the model (Wright et. al. 1996); these habitat features may be unique to the site and hence create sites of potentially high conservation value. Simpson and Norris (2000), on the other hand, offer an alternative interpretation for high *O/E* values. Namely, in the initial stages of eutrophication, mild nutrient enrichment may cause an increase in population densities without a loss of taxa. Although an increase in population densities should not affect models based on presence/absence data, an increase in “rare” taxa may increase the probability of their capture (i.e. sampling may have underestimated rare taxa).

To test how well model predictions are at detecting deviations or impairment, a second external validation step was performed. Sites deemed to be impacted were excluded from the calibration data set. These were sites that had exceedence of S critical load > 0, sites situated in catchments with > 20% arable, and sites classified as affected by clear-cutting. Sites that were considered as limed or affected by liming or sampling error (Västernorrland) were also omitted, but these sites will not be considered here.

3.5.1 Quality bands

Ecological quality bands were calculated for taxa richness *O/E* as described above. Error estimates were obtained by bootstrapping (1000 iterations) of *O/E* values for reference (or least impaired) sites and by using the percentile distribution (e.g. the 25^{th} or10^{th} to 90^{th} percentile = high) of reference *O/E* values. The three approaches to determining quality bands gave very different results (see Tables 3.4 to 3.9 and Fig. 3.9). Using the lower 95% CL means that 5% of the sites fall below this threshold, whereas using the 10-percentile means that 10% of the sites fall below the threshold.

Use of the 95% CI produced a banding scheme that was less prone to type II errors (i.e. not detecting changes if change has occurred). For example, for region 14 banding using the 95% CL resulted in *O/E* values < 0.958 being classified as good to poor, whereas using the 10^{th} percentile resulted in *O/E* values < 0.653. Also, the system used to "anchor" the bands affected classification. Here two approaches were used: (i) use of the upper threshold and zero as anchors and (ii) use of the upper threshold and minimum observed *O/E* ratio as anchors. In both cases, 4 equidistant bands were created by equally dividing the difference between these values. Obviously, the "corrected" banding scheme should consider both the frequency of type I and II errors. Banding systems that are conservative will err on the side of caution and allow a higher frequency of false positives.

3.4. Ecological quality classes for O/E taxon richness in region 14 (mixed forest region). Anchor-values for establishing ecological bands are upper threshold (band A) and zero.

Class and band	Threshold value by: Lower 95 CL	25th Percentile	10th Percentile
Very high, X	1.037	1.33	1.33
High, A	0.958	0.830	0.653
Good, B	0.719	0.623	0.490
Moderate, C	0.480	0.415	0.327
Poor, D	0.241	0.208	0.163
Bad, E	< 0.241	< 0.208	< 0.163

Table 3.5. Ecological quality classes for O/E taxon richness in region 22 (coniferous forest region). Anchor-values for establishing ecological bands are upper threshold (band A) and zero.

Class and band	Threshold value by: Lower 95 CL	25th Percentile	10th Percentile
Very high, X	1.05	1.37	1.37
High, A	0.978	0.827	0.651
Good, B	0.733	0.620	0.489
Moderate, C	0.489	0.414	0.326
Poor, D	0.245	0.207	0.163
Bad, E	< 0.245	< 0.207	< 0.163

Table 3.6. Ecological quality classes for O/E taxon richness in region 20 (arctic/alpine region). Anchor-values for establishing ecological bands are upper threshold (band A) and zero.

Class and band	Threshold value by: Lower 95 CL	25th Percentile	10th Percentile
Very high, X	1.08	1.52	1.52
High, A	0.935	0.803	0.511
Good, B	0.701	0.602	0.383
Moderate, C	0.467	0.402	0.256
Poor, D	0.233	0.201	0.128
Bad, E	< 0.233	< 0.201	< 0.128

Table 3.7. Ecological quality classes for O/E taxon richness in region 14 (mixed forest region). Anchor-values for establishing ecological bands are upper threshold (band A) and minimum value from reference data set.

Class and band	Threshold value by: Lower 95 CL	25th Percentile	10th Percentile
Very high, X	1.037	1.33	1.33
High, A	0.958	0.830	0.653
Good, B	0.790	0.694	0.561
Moderate, C	0.622	0.558	0.469
Poor, D	0.453	0.421	0.377
Bad, E	< 0.453	< 0.421	< 0.377

Table 3.8. Ecological quality classes for O/E taxon richness in region 22 (coniferous forest region). Anchor-values for establishing ecological bands are upper threshold (band A) and minimum value from reference data set.

Class and band	Threshold value by: Lower 95 CL	25th Percentile	10th Percentile
Very high, X	1.05	1.37	1.37
High, A	0.978	0.827	0.651
Good, B	0.782	0.669	0.537
Moderate, C	0.587	0.511	0.423
Poor, D	0.391	0.353	0.309
Bad, E	< 0.391	< 0.353	< 0.309

Table 3.9. Ecological quality classes for O/E taxon richness in region 20 (arctic/alpine region). Anchor-values for establishing ecological bands are upper threshold (band A) and minimum value from reference data set.

Class and band	Threshold value by: Lower 95 CL	25th Percentile	10th Percentile
Very high, X	1.08	1.52	1.52
High, A	0.935	0.803	0.511
Good, B	0.797	0.698	0.479
Moderate, C	0.659	0.593	0.447
Poor, D	0.520	0.487	0.414
Bad, E	< 0.520	< 0.487	< 0.414

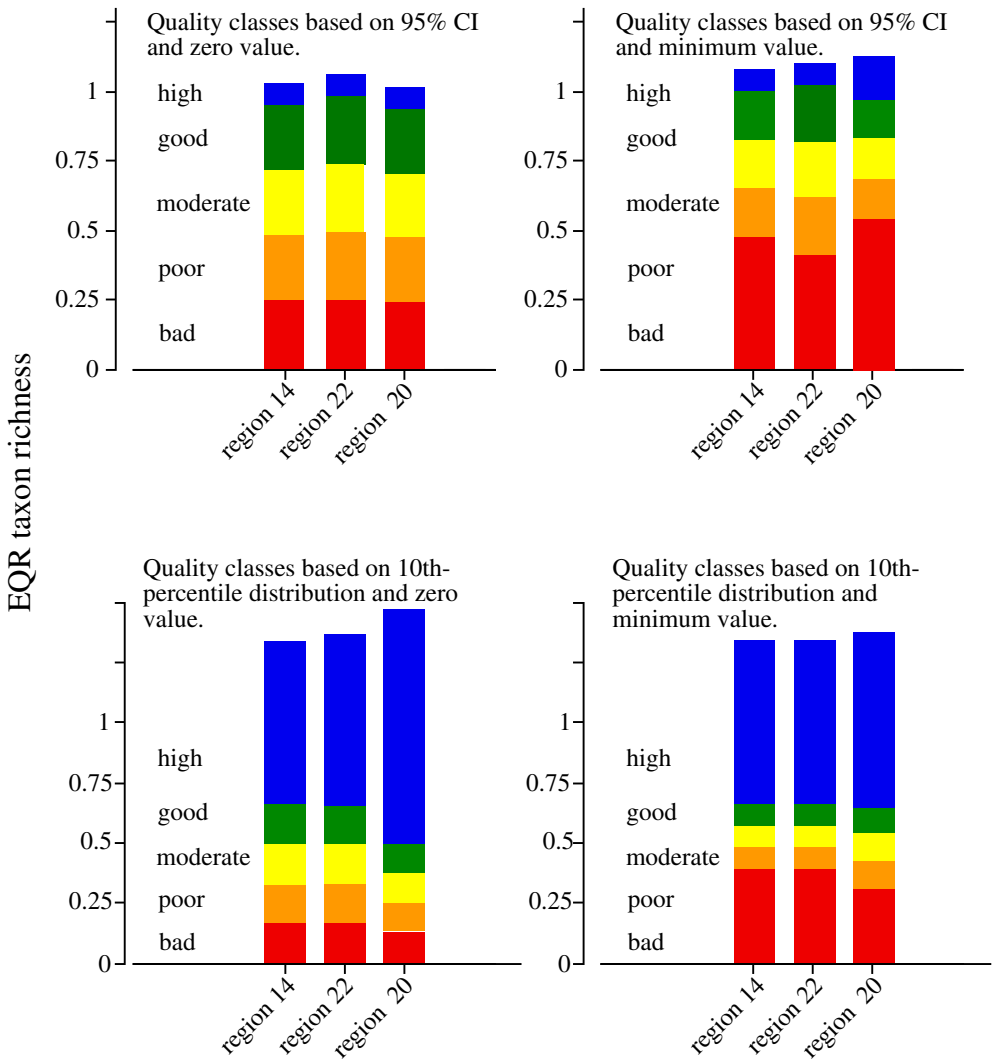

Figure 3.9. Quality bands derived using the lower 95% CL (a) and 10th percentile (b) as threshold values. Anchor-values for establishing ecological bands are upper threshold (band A) and zero or minimum values.

3.5.2 Discriminatory power

Sites in region 14 that had exceedence of S critical load > 0, but not limed or affected by agriculture (< 20% of catchments classified as agriculture) were used to evaluate how well the banding schemes classified perturbed sites (n = 23 sites) (Table 3.10 and 3.11). Of the 6 sites that had pH < 5.0 (one measure taken during the 1995 National Lake Survey), 2 were classified as poor, 2 as moderate and 2 as good using the lower 95% CL threshold value. All 6 sites were classified as moderate or better using the 10^{th} percentile threshold approach. Of the 10 sites with pH > 5.5, 4 were classified as moderate, 4 as good and 2 as very high ecological status according to the 95% banding scheme. Using the 10^{th}-percentile, 3 were classified as good and 7 as high ecological status. Using the minimum observed reference *O/E* value as the lower threshold to set the four bands (B - E) resulted in 1 site being classified as bad, 2 as poor, 1 as moderate and 2 as good for the 7 sites with pH < 5.0 (Table 3.11).

Hence, preliminary studies of comparing "impacted" sites with the two banding schemes lend support to using a more conservative threshold value. Adopting the contention that it is better to err on the side of committing false positives or type I errors, the 95% CL appears to be the more robust approach of the two tested here. However, further tests are required using "dirty" sites to test this assumption more adequately. For example, it is well relatively well established that richness metrics are good at detecting impact (e.g. Johnson 1998; Sandin and Johnson 2000), and taxon richness *O/E* ratios performed well here. However, since $SWEPAC_{LLI}$ models predict the expected taxonomic composition at a site, they also provide information on the expected presence/absence of specific taxa. If the sensitivities of taxa to different stressors are known, this information can result in diagnostics that are better able to detect human-induced change. For example, Johnson (1998) showed that pollution-specific metrics had high power to detect change.

Table 3.10. "Impacted" sites used to test the discriminatory power of the region 14 $SWEPAC_{LLI}$ model. Sites having an exceedence of S critical load are included, but limed sites and sites with >20% catchment agriculture not included. Anchor-values for establishing ecological bands are upper threshold (band A) and zero.

Site	**pH class**	**Exceedence** meq/m^2 yr	**O/E**	**Classification** 95% CL	25th percentile	10th percentile
R1042	pH < 5.0	57	0.422	poor	moderate	moderate
R1246	pH < 5.0	36	0.479	poor	moderate	moderate
R1044	pH < 5.0	47	0.655	moderate	good	high
R1029	pH < 5.0	18	0.599	moderate	moderate	good
R1122	pH < 5.0	20	0.888	good	high	high
R1123	pH < 5.0	11	0.949	good	high	high
R1152	5<pH≤5.5	27	0.693	moderate	good	high
R905	5<pH≤5.5	18	0.787	good	good	high
R662	5<pH≤5.5	34	0.826	good	good	high
R1031	5<pH≤5.5	13	1.294	very high	high	high
R246	5<pH≤5.5	33	0.713	moderate	good	high
R655	5<pH≤5.5	15	0.507	moderate	moderate	good
R959	5<pH≤5.5	44	0.441	poor	moderate	good
R1028	pH>5.5	15	0.571	moderate	moderate	good
R751	pH>5.5	22	0.993	high	high	high
R965	pH>5.5	26	0.611	moderate	moderate	good
R1171	pH>5.5	5	0.691	moderate	good	high
R749	pH>5.5	28	0.65	moderate	good	good
R628	pH>5.5	3	1.032	high	high	high
R364	pH>5.5	8	0.98	high	high	high
R1013	pH>5.5	4	1.217	very high	high	high
R907	pH>5.5	4	0.97	high	high	high
R638	pH>5.5	18	1.19	very high	high	high

Table 3.11. "Impacted" sites used to test the discriminatory power of the region 14 $SWEPAC_{LLI}$ model. Sites having an exceedence of S critical load are included, but limed sites and sites with >20% catchment agriculture not included. Anchor-values for establishing ecological bands are upper threshold (band A) and minimum value from reference data set.

Site	**pH class**	**Exceedence** meq/m^2 yr	**O/E**	**Classification** 95% CL	25th percentile	10th percentile
R1042	pH < 5.0	57	0.422	bad	poor	poor
R1246	pH < 5.0	36	0.479	poor	poor	moderate
R1044	pH < 5.0	47	0.655	moderate	moderate	high
R1029	pH < 5.0	18	0.599	poor	moderate	good
R1122	pH < 5.0	20	0.888	good	high	high
R1123	pH < 5.0	11	0.949	good	high	high
R1152	5<pH≤5.5	27	0.693	moderate	moderate	high
R905	5<pH≤5.5	18	0.787	moderate	good	high
R662	5<pH≤5.5	34	0.826	good	good	high
R1031	5<pH≤5.5	13	1.294	very high	high	high
R246	5<pH≤5.5	33	0.713	moderate	good	high
R655	5<pH≤5.5	15	0.507	poor	poor	moderate
R959	5<pH≤5.5	44	0.441	bad	poor	poor
R1028	pH>5.5	15	0.571	poor	moderate	good
R751	pH>5.5	22	0.993	high	high	high
R965	pH>5.5	26	0.611	poor	moderate	good
R1171	pH>5.5	5	0.691	moderate	moderate	high
R749	pH>5.5	28	0.65	moderate	moderate	good
R628	pH>5.5	3	1.032	high	high	high
R364	pH>5.5	8	0.98	high	high	high
R1013	pH>5.5	4	1.217	very high	high	high
R907	pH>5.5	4	0.97	high	high	high
R638	pH>5.5	18	1.19	very high	high	high

4.0 CALIBRATION AND VALIDATION OF STREAM (RIFFLE) MACROINVERTEBRATE MODELS

4.1 Taxonomic resolution

SWEPAC$_{SRI}$ models were calibrated using an adjusted taxonomic resolution for each region separately. For region 14, 174 taxa were included in the model, for region 20, a total of 70 taxa were included in the model, and for region 22, 146 taxa were included in the model (Appendix 2).

4.2 Model calibration

Table 4.1 shows the number of biological groups and environmental variables used in the development of the SWEPAC$_{SRI}$ running water models.

Table 4.1. Variables used in SWEPAC$_{SRI}$ running water models.

Region	**No. of groups**	**Environmental variables**
Region 14 *(mixed forest)*	5	latitude
n= 126 sites		longitude
n = 173 active taxa		log altitude
		stream velocity
		depth
		amount of sand
		highest coastline
Region 22*(coniferous forest)*	6	latitude
n = 186 sites		longitude
n = 146 active taxa		log altitude
		stream velocity
		emergent vegetation
		heath in riparian zone
		alpine veg. in catchment

Region 20 *(arctic/alpine)*	4	latitude
n = 30 sites		longitude
n = 72 active taxa		log altitude
		depth
		log Cl
		amount of algae in stream
		forest in catchment
		catchment area size

4.2.1 Region 14 (mixed forest region)

One TWINSPAN classification was tested including six groups with 7, 13, 18, 19, 31, and 38 sites in each group. Nine variables (i.e. latitude, longitude, altitude, stream velocity, depth, amount of sand, cobble and mosses in the stream, and whether or not the site was found above the highest coastline of the last glaciation) were chosen as candidate predictors for the classification of the six TWINSPAN groups. These were chosen from a discriminant function analysis using forward selection of environmental variables with the computer program CANOCO vs. 4.0 (ter Braak & Smilauer 1998). Using all nine environmental variables and all 126 sites to build the model, 79 out of the 126 (62.7 %) sites were classified into the correct group. Using cross-validation, where one site at a time was removed from the model-building, and only that site was predicted, 65 out of the 126 sites (51.6 %) were correctly classified. In a minimum model that only included altitude, longitude and latitude, 36 out of 126 (28.6 %) sites were correctly classified using all data for model building, whereas 32 out of 126 (25.4 %) were correctly classified using cross-validation.

After evaluating the model, we chose to use five TWINSPAN groups (i.e. adding the groups containing seven and 19 sites together). Seven environmental variables from the discriminant analysis (i.e. latitude, longitude, altitude, stream velocity, depth, amount of sand in the stream, and whether or not the site was found above the highest coastline of the last glaciation) were included in the model (Fig. 4.1). This model could predict 80 out of the 126 (63.5 %) sites into the correct groups using all

sites in model building, whereas 68 out of 126 (54.0 %) were correctly classified using cross-validation.

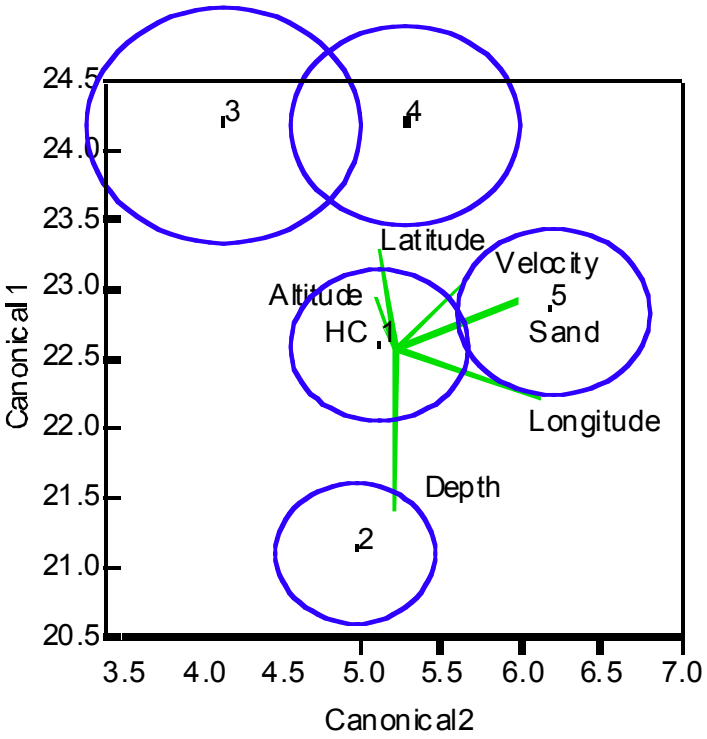

Figure 4.1. Centroid plot of five TWINSPAN groups and seven environmental variables included in the region 14 model. The numbers show the centroid for each group and circles denote a 95% CI.

4.2.2 Region 22 (coniferous forest region)

Three TWINSPAN classifications were tested. The first included five groups (with 19, 25, 39, 50 and 54 sites in the different cluster groups). The second model had eight groups (with 19, 19, 20, 23, 25, 26, 27 and 28 sites in the different cluster groups). The third classification consisted of 12 groups (with 4, 5, 6, 9, 10, 13, 16, 20, 23, 26, 27, and 28 sites in the different cluster groups). Ten variables (i.e. latitude, longitude, altitude, stream velocity, amount of algae, *Fontinalis* and gravel in the stream, the amount of heath in the riparian zone and forest and freshwater in

the catchment) were chosen as candidate predictors for the classification of five TWINSPAN groups. Using all ten environmental variables and all 187 sites to build the model 117 out of the 187 (62.6 %) of the sites were classified into the correct group. Using cross-validation, 105 out of the 187 sites (56.1 %) were correctly classified. In a minimum model that only included altitude, longitude and latitude, 71 out of 187 (38.0 %) sites were correctly classified using all data for model building, whereas 67 out of 187 (35.8 %) were correctly classified using cross-validation.

For the model including eight TWINSPAN groups, ten variables (i.e. latitude, longitude, altitude, stream velocity, amount of cobble in the stream, width of the stream, the amount of heath and the amount of alpine vegetation and freshwater in the catchment) were chosen as candidate predictor variables. Using all ten environmental variables and all 187 sites to build the model, the model predicted 94 out of the 187 (50.3 %) of the sites into the correct group, whereas 87 out of 187 (46.5 %) were correctly classified using cross-validation. Using a minimum model that only included altitude, longitude and latitude, 62 out of 187 (33.2 %) were correctly classified using all data for model building, whereas 57 out of 187 (30.5 %) were correctly classified using cross-validation.

For the model including 12 TWINSPAN groups, the same ten variables were chosen as for the model using eight groups (i.e. latitude, longitude, altitude, stream velocity, vegetation cover and amount of cobble in the stream, width of the stream, the amount of heath in the riparian zone, the amount of alpine vegetation and freshwater in the catchment). Using all ten environmental variables and all 187 sites to build the model, 82 out of the 187 (43.9 %) of the sites were classified into the correct group, whereas 59 out of 187 (31.6 %) were correctly classified using cross-validation. Using a minimum model that only included altitude, longitude and latitude, 50 out of 187 (26.7 %) were correctly classified using all data for model building, whereas 44 out of 187 (23.5 %) were correctly classified using cross-validation.

To test how many environmental variables to include in the model, all possible models using 6 – 10 environmental variables were constructed and tested with the TWINSPAN classification of four groups. Latitude, longitude, altitude and stream

depth were included in all models, and all possible combinations of the remaining six variables were tested. Models were built both by including all 30 sites and using cross-validation (see above). The best predictor model included eight environmental variables where 28 out of 30 variables were correctly predicted using all data and 23 out of 30 using cross-validation (Fig. 4.2).

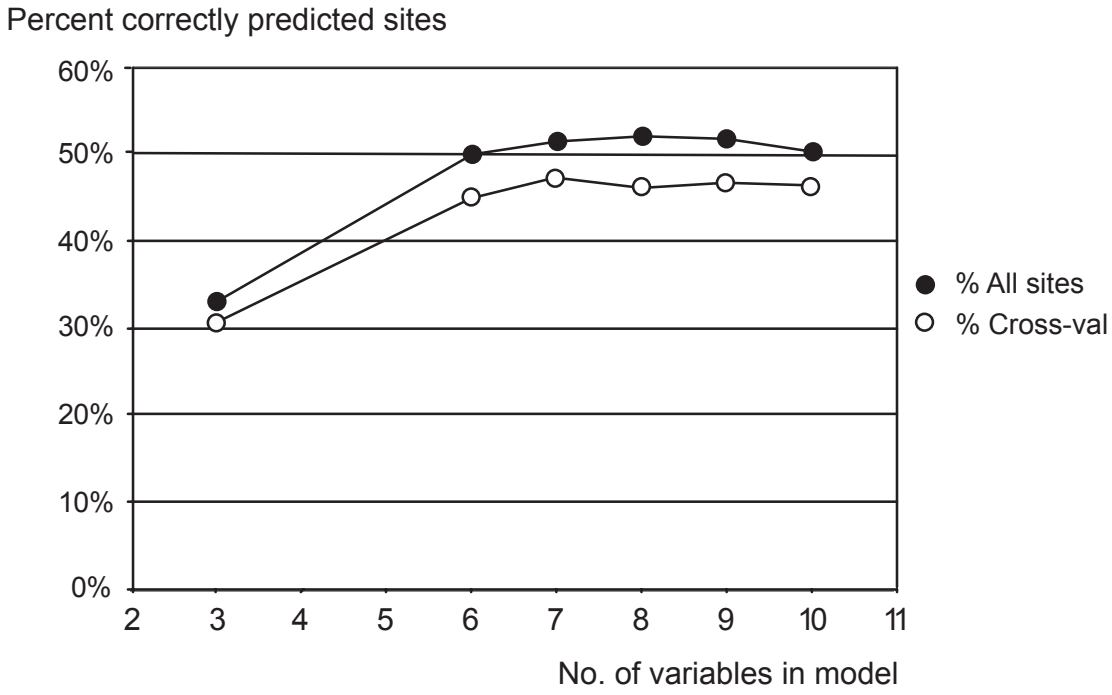

Figure 4.2. Testing a model with 8 cluster groups, using different numbers of predictor variables. Latitude, longitude, altitude and stream velocity was included in all models. The percent of correct classified sites for a model built on all data (filled circles) and cross-validation (open squares) are shown for models including different numbers of predictor variables.

The difference in number of correctly classified sites depending on the number of included environmental variables in the model was also compared between the model built on eight and 12 TWINSPAN groups, respectively. The model including eight groups, always had a higher percent correctly classified sites (i.e. it differed 6 – 9 % in a model built using all variables and 6 – 15 % using cross-validation) (Fig. 4.3).

The best predictor model included seven environmental variables (i.e. latitude, longitude, altitude, stream velocity, the amount of freshwater in the catchment, the amount of heath in the riparian zone, and the amount of vegetation cover in the stream). Using this model, 96 out of 187 (51.3 %) sites were correctly classified using all data and 88 out of 187 (47.1 %) sites using cross-validation.

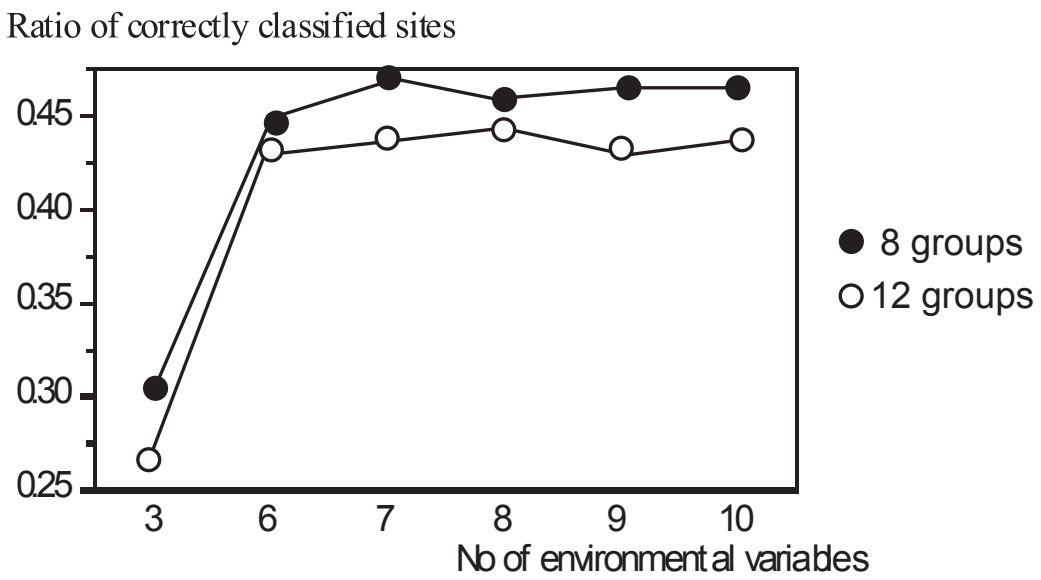

Figure 4.3. Comparison of the ratio of correctly classified sites. The model including 8 groups is the upper line, whereas the model including twelve groups is the lower line.

After evaluating these models (i.e. including five, eight and 12 TWINSPAN groups, respectively), we decided to test a model including six groups. The seven best predictor environmental variables were: latitude, longitude, altitude, stream velocity, heath in the riparian zone, alpine vegetation and freshwater in the catchment. A model including these seven environmental variables predicted 105 out of 187 (56.1%) correctly and using cross-validation 99 out of 187 (52.9 %) were correctly

classified. The number of correctly classified sites within each group varied between 35.2 % and 75.0 % for the model including six TWINSPAN groups. The discriminant model developed from these seven predictor variables (Fig. 4.4) was highly significant (Wilk's lamda = 0.165, F = 11.2865, P < 0.001).

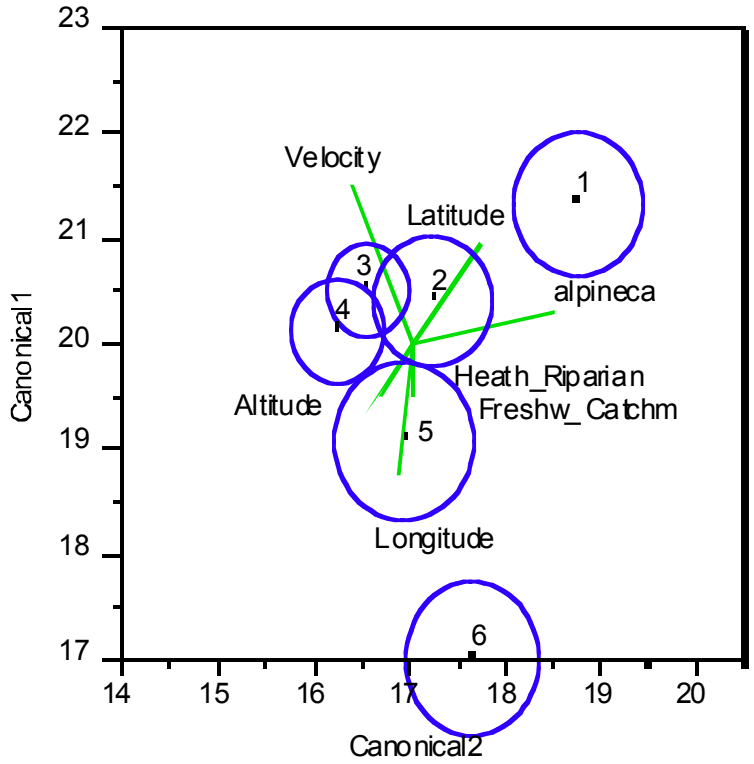

Figure 4.4. Centroid plot of six TWINSPAN groups and seven environmental variables for the region 22 model. The numbers show the centroid for each group and circles denote a 95% CI.

4.3.3 Region 20 (arctic/alpine region)

Two TWINSPAN classifications were tested, one including four groups (with 2, 7, 8, and 13 sites in the different cluster groups) and one model with 6 groups (with 2, 4, 4, 5, 7, and 8 sites). Ten variables (i.e. latitude, longitude, altitude, depth, Cl, Mg, amount of algae and *Fontinalis* in the stream, amount of forest in the catchment area, and catchment area size) were chosen as candidate predictors for the classification of

four TWINSPAN groups. These were chosen from a discriminant function analysis using forward selection of environmental variables in CANOCO. Using all ten environmental variables and all 30 sites to build the model 28 out of the 30 (93.3 %) of the sites were classified into the correct group. Using cross-validation 22 out of the 30 sites (73.3 %) were correctly classified. In a minimum model that only included altitude, longitude and latitude, 22 out of 30 (73.3 %) sites were correctly classified using all data for model building, whereas 20 out of 30 (67.7 %) were correctly classified using cross-validation.

For the model including six TWINSPAN groups, ten variables (i.e. latitude, longitude, altitude, depth, Cl, Mg, amount of algae, CPOM and *Fontinalis* in the stream, and amount of alpine vegetation in the riparian zone) were chosen as candidate predictor variables. Using all ten environmental variables and all 30 sites to build the model predicted 26 out of the 30 (86.7 %) of the sites into the correct group. Using a minimum model that only included altitude, longitude and latitude, 27 out of 30 (56.7 %) were correctly classified using all data for model building, whereas 14 out of 30 (46.7 %) were correctly classified using cross-validation.

To test how many environmental variables to include in the model, all possible models using 6 – 10 environmental variables were constructed and tested with the TWINSPAN classification of four groups. Latitude, longitude, altitude and depth of the stream were included in all models, and all possible combinations of the remaining six variables were tested. Models were built both by including all 30 sites and using cross-validation (see above). The best predictor model included eight environmental variables where 28 out of 30 variables were correctly predicted using all data and 23 out of 30 using cross-validation (Fig. 4.5).

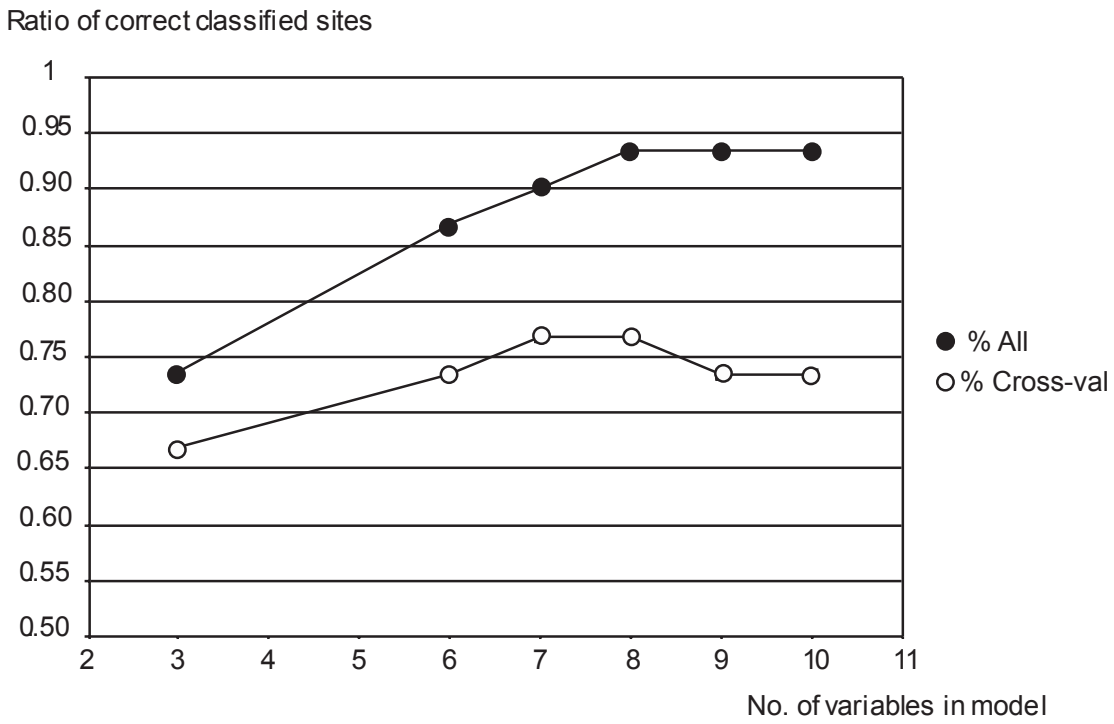

Figure 4.5. Testing all possible predictive models, using different numbers of predictor variables. Latitude, longitude, altitude and depth were included in all models. The best ratio of correct classified sites for a model built on all data (filled circles) and cross-validation (open squares) are shown for models including different numbers of predictor variables.

The best eight variables were: latitude, longitude, altitude, depth, Cl, amount of algae in the stream, amount of forest in the catchment area and size of the catchment area. The discriminant model developed from these eight predictor variables (Fig. 4.6) was highly significant (Wilk's lamda = 0.044, $F = 4.4795$, $P < 0.001$). Using this model, 23 out of 30 (76.7 %) of the sites were classified to the correct group. Group 4 had the highest percent of correct classifications (92.3%), followed by group 2 (87.5%), group 1 (50.0 %), whereas 42.9 % of the sites in group 3 were correctly classified.

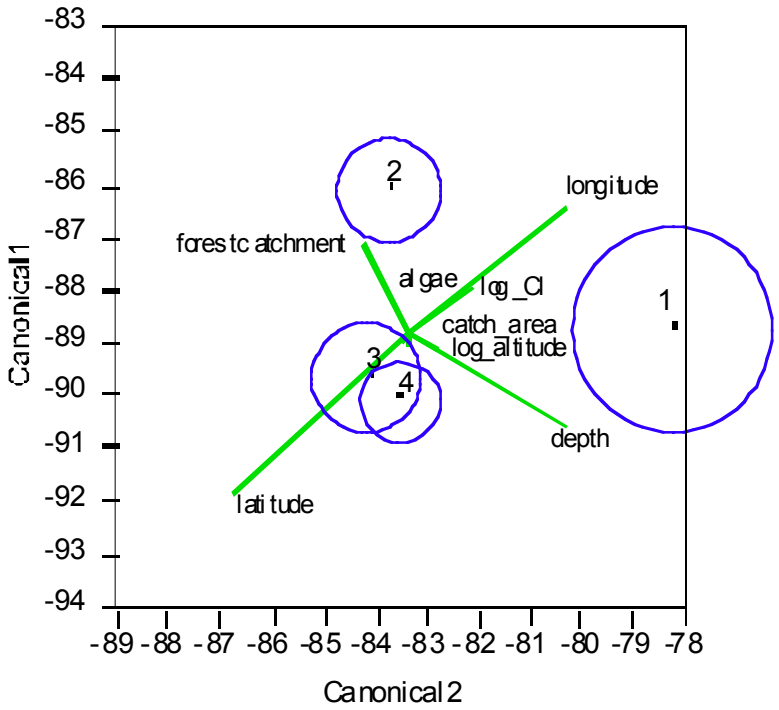

Figure 4.6. Centroid plot of four TWINSPAN groups and eight environmental variables for the region 20 model. The numbers show the centroid for each group and circles denote a 95% CL.

4.4 Inclusion probability

To determine what threshold value to use in the SWEPAC$_{SRI}$ models we compared the results from a number of calibration runs. For all models the number of observed and expected taxa increased as the inclusion threshold decreased (Fig. 4.7a-c). As expected, the lowest values were found in region 20 or the arctic/alpine complex presumably due to the impoverished regional species pool and/or harsher environments. Models for region 14 and 22 had similar taxon richness.

A comparison of inclusion probability thresholds and *O/E* ratios showed only slight differences among regions and with varying threshold values. Threshold levels of 50% always had lower variance (standard deviations) than other threshold levels (i.e. thresholds < 50%). Models had *O/E* ratios which ranged from 0.95 – 0.98 for region 14; 1.02 – 1.11 for region 22 and 1.00 – 1.01 for sites in region 20 (Table 4.2).

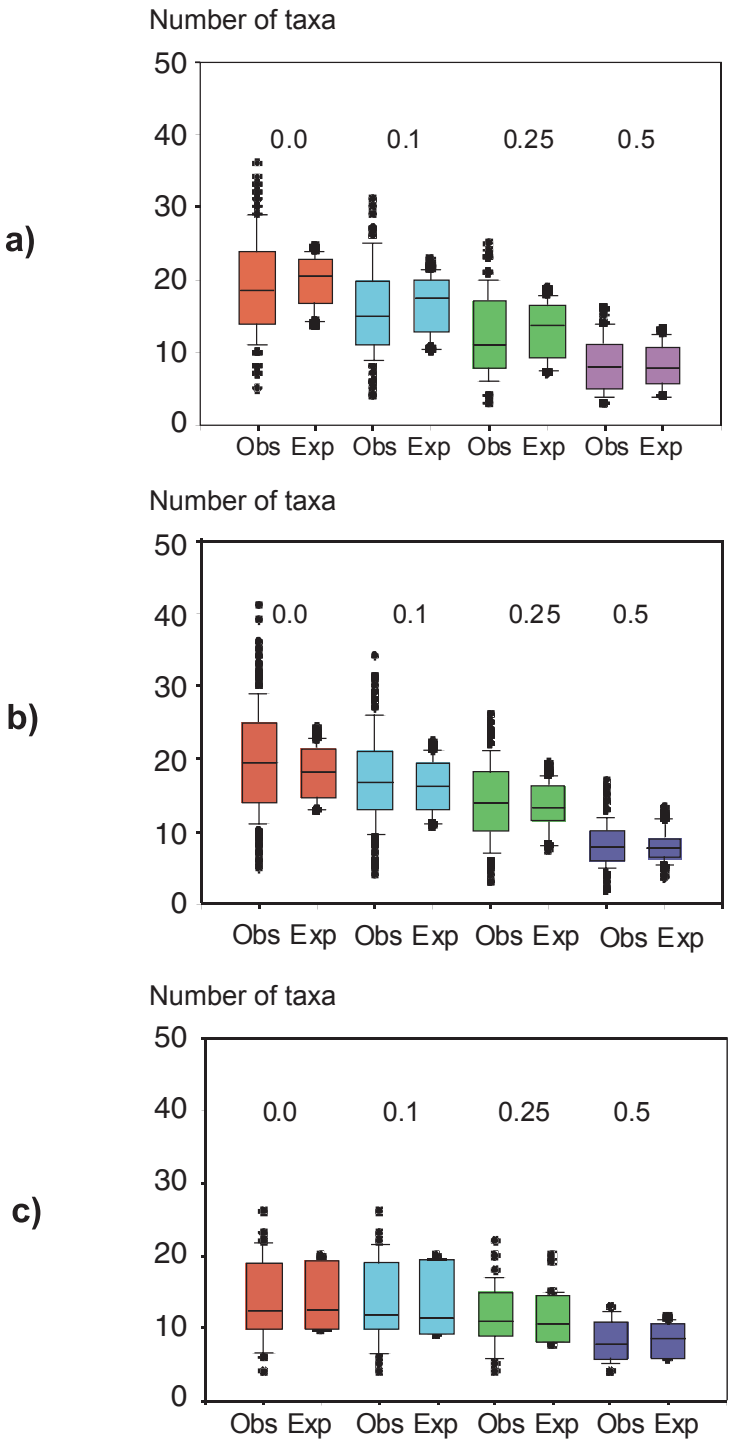

Fig 4.7a-c. Mean number of observed (left) and expected (right) taxa for different probability of occurrence thresholds in regions 14, (a), 22 (b) and 20 (c).

Table 4.2. Means ± standard deviations for O/E values from reference sites. O/E_x refers to the probability of occurrence thresholds (i.e. ≥ 0.50, 0.25, 0.10 and 0).

Model	O/E_{50}	O/E_{25}	O/E_{10}	O/E_0
Region 14	0.98±0.22	0.96±0.28	0.96±0.30	0.95±0.29
Region 22	1.02±0.25	1.05±0.30	1.07±0.32	1.11±0.37
Region 20	1.01±0.18	1.01±0.23	1.00±0.26	1.00±0.26

Another way to test which of the probability of occurrence thresholds to use is to look at the correlation between the observed number of taxa and the expected number of taxa, since analyzing only at the mean values of observed and expected values for each region might give a somewhat different picture than the actual values given by the model (Table 4.3).

Table 4.3. Results of linear regression of expected against observed number of taxa for different threshold values (or inclusion probabilities).

Region	**Threshold**	**r^2**	**p-value**
14	0.0	0.27	< 0.001
14	0.1	0.36	< 0.001
14	0.25	0.48	< 0.001
14	0.5	0.68	< 0.001
22	0.0	0.20	< 0.001
22	0.1	0.30	< 0.001
22	0.25	0.40	< 0.001
22	0.5	0.50	< 0.001
20	0.0	0.71	< 0.001
20	0.1	0.77	< 0.001
20	0.25	0.71	< 0.001
20	0.5	0.74	< 0.001

We decided to use a threshold level of 25% for $SWEPAC_{SRI}$ models. These decisions were based on calibration runs comparing results of observed, expected and *O/E* ratios using thresholds of 0, 10, 25 and 50% probabilities. This decision was based on the results discussed above, as well as a compromise between models that predicted fewer taxa but had lower variance associated with these predictions (e.g. models using a 50% threshold level) and models that predicted many taxa but had high variance (e.g. models using a 10% threshold level).

4.5 Model validation

The error associated with the $SWEPAC_{SRI}$ models (0.25 inclusion probability) was estimated using two methods. First, the distribution of the reference data set (internal error validation) gives a measure of the error distribution (see Table 4.2). However, a more rigorous measure of model error is made by treating data from a series of reference sites that were not used in model construction as test data. Before model calibration, 5% of the reference sites were randomly removed from the data set. Applying the calculations in steps 3 – 6 (Table 2.2) these sites provide a second and independent measure of model error. The distribution of *O/E* values should exhibit a near-normal frequency distribution with values centered on 1, and the spread of these *O/E* values represents model error.

Internal validation showed that error estimates (here expressed as the 95% CI) for *O/E* values were for region 14; mean = 0.962, upper CL = 1.012 and lower CL = 0.913, for region 22; mean = 1.051, upper CL = 1.095 and lower CL = 1.008, and for region 20; mean = 1.011, upper CL = 1.044 and lower = 0.978. Bootstrapping and Efron's percentile (Efron 1979) of *O/E* values by region was used to confirm these error estimates. Bootstrapped estimates obtained from 1000 iterations were similar to those obtained by conventional parametric statistics; for region 14; mean *O/E* = 0.962, upper CL = 1.000 and lower CL = 0.919, for region 22; mean *O/E* = 1.051, upper CL = 1.088 and lower CL = 1.014, and for region 20; mean *O/E* = 1.011, upper CL = 1.083 and lower = 0.935.

Although the number of sites used for external validation are too few for obtaining robust estimates of model error (5% of the total number of reference sites), these

sites provide nonetheless some measure of expected model error. Not surprisingly, variance estimates (using bootstrapping) were larger for external-reference than the internal-reference validations (Fig. 4.8). For region 14, seven sites were used in the second validation step (mean *O/E* = 0.776, upper CL = 0.958 and lower CL = 0.607). Ten sites were used in the external validation of region 22 sites (mean *O/E* = 0.829, upper CL = 0.995 and lower CL = 0.657). For region 20 only three sites were used; (mean *O/E* = 0.820, upper CL = 1.014 and lower CL = 0.634).

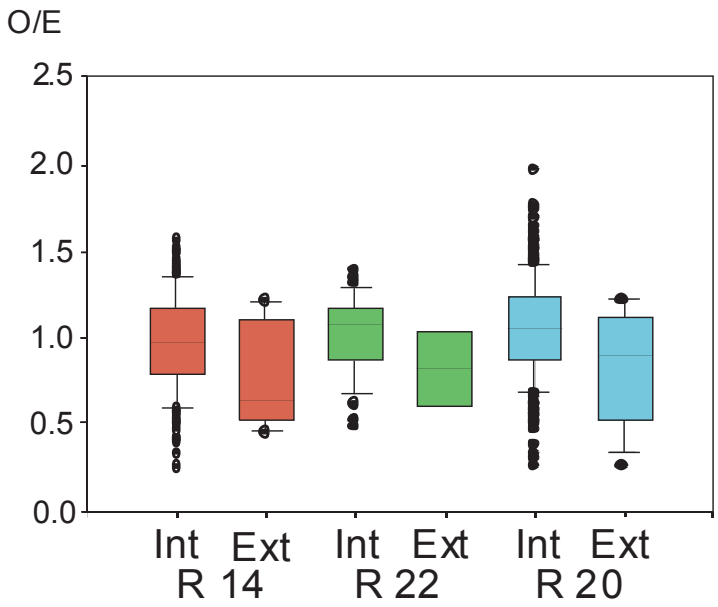

Figure 4.8. Error estimates of internal (Int) (reference or calibration data set), external (Ext) (reference or 5% of original data set), validations. Median O/E values and interquartile range (75 % and 25 %, respectively) is given within the boxes.

The region 14 model was used as an example to test how well the models could detect the effects of perturbation. Stream sites with a total phosphorous content > 100 $\mu g l^{-1}$ were deemed as impacted and used as "affected". Predicted taxa lists were calculated for these sites and compared to the internal and external references. Comparing the *O:E* ratios (number of taxa) between the internal (model) references and the affected sites, showed that there was significant differences between the two for all inclusion probabilities ($p < 0.003$). The explained variance were rather low

(r^2= 0.067-0.131) with the highest explained variance for the model with an inclusion probability of 0.5.

Comparing *O:E* ratios (number of taxa) between the external references and the affected sites, showed however, no significant differences ($p > 0.05$) for all four inclusion probabilities. This might be a results of the low power for these tests (always <0.1) and thus a larger number of external references (a new test data set) would be needed to truly evaluate the capability of the model to detect human influenced perturbation.

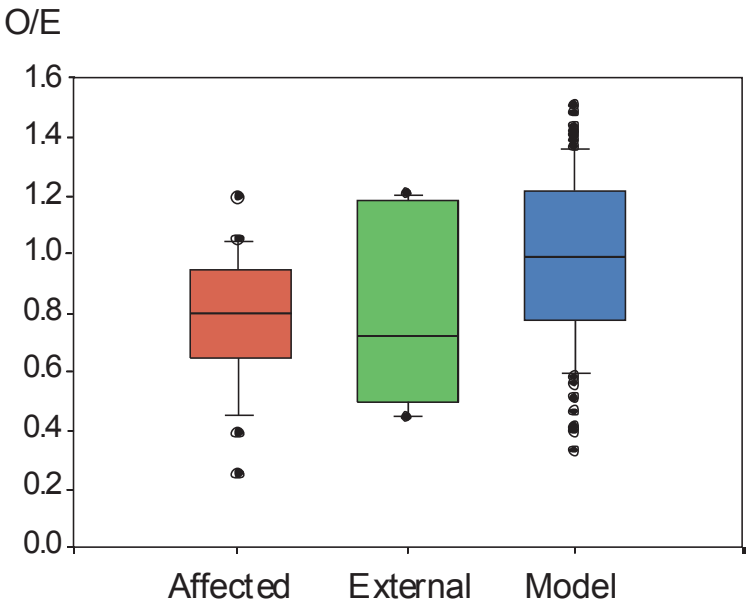

Figure 4.9. Error estimates of the O:E ratio of number of taxa. Internal (reference or calibration data set), external (reference or 5% of original data set), and affected (stream sites with a total phosphorous content > 100 µgl^{-1}) sites. Median O/E values and interquartile range (75 % and 25 %, respectively) is given within the boxes.

Another way to use the model is to calculate indices based on the observed and expected taxa lists. It is then possible to calculate the *O:E* ratio between the observed and expected index value. This was done using the region 14 model and the Average Score per Taxon (ASPT) a "clean water index" mainly indicating the presence of sensitive groups (high values) and insensitive ones (low values) (Armitage *et al.*

1983). This *O:E* ratio (only calculated for the 0.25 inclusion probability) clearly distinguished between the internal (model) references (mean = 0.945; upper CL = 0.980; lower CL = 0.910) and the affected sites (mean = 0.839; upper CL = 0.902; lower CL = 0.777) with a *t-test* ($p<0.005$). There was also a clear difference when comparing the external references (mean = 0.977; upper CL = 1.108; lower CL = 0.885) and the affected sites (mean = 0.839; upper CL = 0.896; lower CL = 0.783) with a *t-test* ($p < 0.02$) (Fig. 4.10).

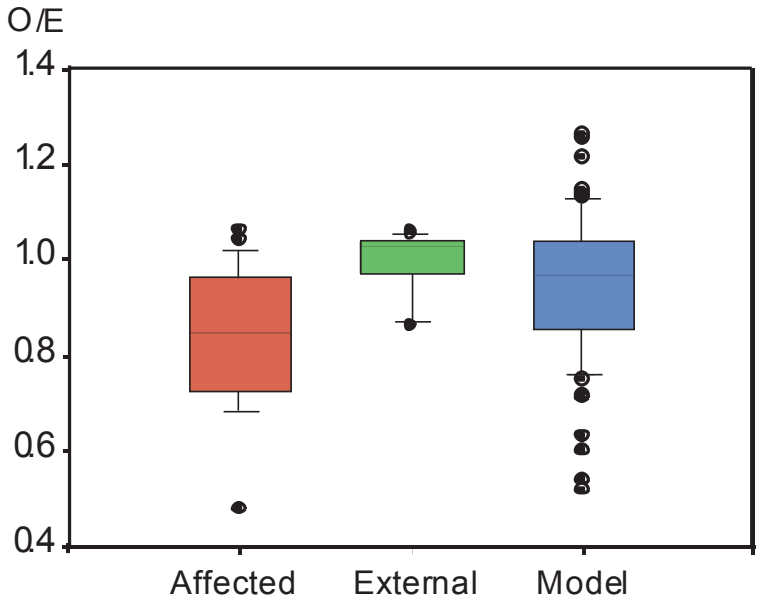

Figure 4.10. Error estimates of O:E ratios of the Average Score per Taxon (ASPT) index. Internal (reference or calibration data set), external (reference or 5% of original data set), and affected (stream sites with a total phosphorous content > 100 μgl^{-1}) sites. Median O/E values and interquartile range (75 % and 25 %, respectively) is given within the boxes.

5.0 Acknowledgements

We thank the many people who assisted in the planning, sampling and laboratory analysis of the vast number of samples collected in the 1995 national lake and stream survey. Special thanks also to SLU Miljödata who financially supported our endeavors to develop these predictive models.

6.0 References

Agresti, A. 1990. Categorical data analysis. John Wiley & Sons, New York, USA.

Armitage, P.D., Moss, D., Wright, J.F. and M.T. Furse. 1983. The perfomance of a new biological water quality score system based on macroinvertebrates over a wide range of unpolluted running-waters. Water Research 17:333-347.

Austin, M.P., Nicholls, A.O., Doherty, M.D. and J.A. Meyers. 1994. Determining species response functions to an environmental gradient by means of a β-function. Journal of Vegetation Science 5: 215–228.

Bio, A.M.F., Alkemade, R. and A. Barendregt. 1998. Determining alternative models for vegetation response analysis: a non-parametric approach. Journal of Vegetation Science 9: 5–16.

Brzeziecki, B., Kienast, F. and O. Wildi. 1995. Modeling potential impacts of climate change on the spatial distribution of zonal forest communities in Switzerland. Journal of Vegetation Science 6: 257–258.

Cao Y., Williams D.D. and Williams N.E. 1999. Rare species are important in aquatic community ecology and bioassessment? Limnology and Oceanography 43: 1403-1409.

Cao Y. and Williams D.D. 1999. Rare species are important in bioassessment (Reply to the comment by Marchant). Limnology and Oceanography 44: 1841-1842.

Czekanowski, J. 1913. Zarys method statystycznyck. Warsaw.

Efron, B. 1979. Bootstrap methods: another look at the jackknife. Annals of Statistics, 7: 1–26.

European Committee for Standardization 1994. Water quality—methods for biological sampling—guidance on handnet sampling of aquatic benthic macro-invertebrates. SS-EN-27-828. European Committee for Standardization, Brussels, Belgium.

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

Hawkins, C.P. and D.M. Carlisle. In press. Use of predictive models for assessing the biological integrity of wetlands and other aquatic habitats. In: R.B. Rader, D.P. Batzer and S. Wissinger (eds.). Biomonitoring and Management of North American Freshwater Wetlands.

Hill, M.O. 1979. TWINSPAN - A FORTRAN program for arranging multivariate data in an ordered two-way table by classification of the individuals and attributes. Cornell University, Ithaca, NY, 49 pp.

Hotelling, H. 1933. Analysis of a complex of statistical variables into principal components. Journal of Educational Psychology 24: 417–441.

Illies, J. 1966. Verbreitung der Süßwasserfauna Europas. Verhandlungen der Internationalen Vereinigung für Theoretische und Angewandte Limnologie 16: 287–296.

Johnson, R.K. 1995. The indicator concept in freshwater biomonitoring. Thienemann lecture. pp. 11 - 27, In: Chironomids - from Genes to Ecosystems. Proceedings of the 12th International Symposium on Chironomidae. (ed. P.S. Cranston), CSIRO, Melbourne.

Johnson, R.K. 1998. Classification of Swedish lakes and rivers using benthic macroinvertebrates (Bottenfauna, Bakgrundsdokument till bedömningsgrunder för sjöar och vattendrag), NV report.

Johnson R.K. 2000. RIVPACS and alternative statistical modeling techniques - accuracy and soundness of principles. Assessing the biological quality of freshwaters: RIVPACS and techniques. pp 323-332 In: J. F. Wright, D. W. Sutcliffe, and M. T. Furse (eds). Assessing the biological quality of freshwaters. RIVPACS and other techniques. Freshwater Biological Association, Ambleside, UK.

Johnson R. K., and W. Goedkoop. 2000. The 1995 national survey of Swedish lakes and streams: assessment of ecological status using macroinvertebrates. pp 229-240 In: J. F. Wright, D. W. Sutcliffe, and M. T. Furse (eds). Assessing the biological quality of freshwaters. RIVPACS and other techniques. Freshwater Biological Association, Ambleside, UK.

Johnson R.K. and W. Goedkoop. 2000. The use of biogeographical regions for partitioning variance of littoral macroinvertebrate communities. Verhandlungen der Internationale Vereinigung für Theoretische und Angewandte Limnologie 27: 333-339.

Johnson, R.K., Wiederholm, T. and D.M. Rosenberg. 1993. Freshwater biomonitoring using individual organisms, populations, and species assemblages of benthic macroinvertebrates. pp. 40-158, In: Freshwater Biomonitoring and Benthic Macroinvertebrates. (eds. D.M. Rosenberg and V.H. Resh), Chapman and Hall, New York.

Moss, D., Furse, M.T., Wright, J.F. and P.D. Armitage. 1987. The prediction of the macro-invertebrate fauna of unpolluted running-water sites in Great Britain using environmental data. Freshwater Biology 17: 41–52.

Nordic Council of Ministers. (1984). Naturgeografisk regionindelning av Norden. Nordiska ministerrådet 1984.

Nicholls, A.O. 1989. How to make biological models go further with Generalized Linear Models. Biological Conservation 50: 51–75.

Simpson, J. C. and R. H. Norris. 2000. Biological assessment of river quality: development of AUSRIVAS models and outputs. Pages 125-142 in J. F. Wright, D. W. Sutcliffe, and M. T. Furse, editors, Assessing the biological quality of freshwaters. RIVPACS and other techniques. Freshwater Biological Association, Ambleside, England.

ter Braak, C. 1986. Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. Ecology 67: 1167–1179.

ter Braak, C.J.F. 1989. CANOCO - a Fortran program for canonical community ordination by correspondence analysis, principal components analysis, and redundancy analysis. version 3.11, Agricultural Mathematics Group, Wageningen, the Netherlands.

ter Braak, C.J.F., C.W.N. Looman. 1986. Weighted averaging, logistic regression and the Gaussian response model. Vegetatio 65: 3–11.

Wiegleb, G. 1981. Application of discriminant analysis on the analysis of the correlation between macrophyte vegetation and water quality i running waters of Central Europe. Hydrobiologia 79: 91–100.

Wilander, A., Johnson, R.K., Goedkoop, W. and L. Lundin. 1997. Riksinventering 1995: En synoptisk studie av vattenkemi och bottenfauna i svenska sjöar och vattendrag. Naturvårdsverket rapport, 4813.

Wold, H. 1982. Soft modeling: the basic design and some extensions. pp. 1–54, In: Joreskog, K.G. and H. Wold (eds.), Systems under indirect observations II, North-Holland, Amsterdam.

Wright, J. F. 1995. Development and use of a system for predicting the macroinvertebrate fauna in flowing waters. Australian Journal of Ecology 20:181-197.

Wright, J.F., Armitage, P.D., Furse, M.T., and D. Moss. 1988. A new approach to the biological surveillance of river quality using macroinvertebrates. Internationale Vereinigung für Theoretische und Angewandte Limnologie Verhandlungen 23: 1548-1552.

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

Appendix 1. Two levels of taxonomic resolution used in SWEPAC$_{LLI}$ lake-model calibrations.

	"Species"	**"Families"**
1	Turbellaria	Spongillidae
2	Nematoda	Hydrozoa
3	*Bithynia tentaculata* (L.)	Turbellaria
4	*Lymnaea stagnalis* (L.)	Nematoda
5	*Radix* spp.	Neritidae
6	*Physa fontinalis* (L.)	Viviparidae
7	*Bathyomphalus contortus* (L.)	Valvatidae
8	*Gyraulus* spp.	Hydrobiidae
9	*Acroloxus lacustris* (L.)	Lymnaeidae
10	Sphaeriidae	Physidae
11	Oligochaeta	Planorbidae
12	*Glossiphonia* spp.	Ancylidae
13	*Helobdella stagnalis* (L.)	Unionidae
14	*Erpobdella* spp.	Sphaeriidae
15	Hydracarina	Oligochaeta
16	*Argyroneta aquatica* (Clerk)	Hirudidae
17	*Asellus aquaticus* L.	Piscicolidae
18	*Gammarus* spp.	Glossiphoniidae
19	*Ameletus inopinatus* Eaton	Erpobdellidae
20	*Baetis* spp.	Hydracarina
21	*Centroptilum* spp.	Araneae
22	*Cloeon* spp.	Asellidae
23	*Heptagenia* spp.	Gammaridae
24	*Leptophlebia* spp.	Collembola
25	*Ephemera* spp.	Siphlonuridae
26	*Caenis* spp.	Baetidae
27	*Nemoura* spp.	Heptageniidae
28	*Capnia* spp.	Leptophlebiidae
29	*Diura* spp.	Ephemeridae

30	Coenagriidae	Caenidae
31	*Coenagrion* spp.	Taeniopterygidae
32	*Erythromma najas* (Hansem.)	Nemouridae
33	Zygoptera	Leuctridae
34	*Aeshna* spp.	Capniidae
35	*Cordulia* spp.	Perlodidae
36	*Somatochlora* spp.	Platycnemididae
37	*Libellula* spp.	Coenagriidae
38	*Leucorrhinia* spp.	Agriidae
39	Anisoptera	Gomphidae
40	*Gerris* spp.	Cordulegasteridae
41	*Notonecta* spp.	Aeshnidae
42	Corixidae	Corduliidae
43	*Micronecta* spp.	Libellulidae
44	*Cymatia* spp.	Anisoptera
45	*Callicorixa* spp.	Gerridae
46	*Sigara* spp.	Nepidae
47	*Haliplus* spp.	Notonectidae
48	Dytiscidae	Corixidae
49	*Hygrotus* spp.	Haliplidae
50	*Hydroporus* spp.	Dytiscidae
51	*Agabus* spp.	Gyrinidae
52	*Ilybius* spp.	Hydrophilidae
53	*Potamonectes* spp.	Helodidae
54	*Nebrioporus* spp.	Dryopidae
55	*Gyrinus* spp.	Elminthidae
56	*Oulimnius* spp.	Chrysomelidae
57	*Sialis* spp.	Sialidae
58	*Plectrocnemia* spp.	Rhyacophilidae
59	*Polycentropus* spp.	Polycentropodid
60	*Holocentropus* spp.	Psychomyiidaeae
61	*Cyrnus* spp.	Hydropsychidae
62	*Ecnomus tenellus* Ramb.	Hydroptilidae
63	*Tinodes* spp.	Phryganeidae

64	*Lype* spp.	Limnephilidae
65	*Oxyethira* spp.	Molannidae
66	*Phryganea* spp.	Leptoceridae
67	*Agrypnia* spp.	Goeridae
68	Limnephilidae	Lepidostomatida
69	*Limnephilus* spp.	Brachycentridaee
70	*Glyphotaelius pellucidus* Retz.	Sericostomatida
71	*Nemotaulius punctatolineatus* Retz.	Tipulidaee
72	*Potamophylax* spp.	Dixidae
73	*Molanna* spp.	Chaoboridae
74	*Molannodes tinctus* Zett.	Culicidae
75	*Athripsodes* spp.	Ceratopogonida
76	*Mystacides* spp.	Chironomidaee
77	*Oecetis* spp.	Simuliidae
78	*Lepidostoma hirtum* (Fabricius)	Stratiomyidae
79	Tipulidae	Empididae
80	Chaoboridae	Dolichopodidae
81	Ceratopogonidae	Tabanidae
82	Chironomidae	Syrphidae
83	Empididae	Muscidae
84	Tabanidae	Cylindrotomidae
85	Limoniidae	Limoniidae

Appendix 2. Taxonomic resolution used in SWEPAC$_{SRI}$ stream-model calibrations.

Region 14

1	Acilius canaliculatus (Nicolai)	23	Caenis luctuosa-macrura
2	Acroloxus lacustris (L.)	24	Caenis rivulorum Eaton
3	Aeshna sp.	25	Callicorixa sp.
4	Agabus sp.	26	Calopteryx sp.
5	Agapetus sp.	27	Capnia sp.
6	Agrypnia sp.	28	Centroptilum sp.
7	Amphinemura borealis (Morton)	29	Ceratopogonidae
8	Amphinemura standfussi-sulcicollis	30	Chaoborus sp.
9	Ancylus fluviatilis (Müller)	31	Cheumatopsyche lepida Pictet
10	Aphelocheirus aestivalis (F.)	32	Chimarra marginata L.
11	Asellus aquaticus L.	33	Chironomidae
12	Astacus astacus (L.)	34	Cloeon sp.
13	Athripsodes sp.	35	Coenagrion sp.
14	Baetis fuscatus group	36	Collembola
15	Baetis muticus L.	37	Cordulia aenea (L.)
16	Baetis niger-digitatus	38	Cordulegaster boltonii (Donovan)
17	Baetis rhodani Pict.	39	Corixa sp.
18	Bathyomphalus contortus (L.)	40	Culicidae
19	Bithynia tentaculata (L.)	41	Cymatia sp.
20	Brachyptera sp.	42	Cyrnus flavidus McL.
21	Brachysera	43	Cyrnus insolutus McLachlan
22	Caenis horaria L.	44	Cyrnus trimaculatus Curtis
45	Dicranota sp.	71	Helobdella stagnalis (L.)
46	Dinocras cephalotes (Curtis)	72	Heptagenia dalecarlia Bengtsson
47	Diura nanseni (Kempny)	73	Heptagenia fuscogrisea Retz.
48	Dolichopodidae	74	Heptagenia sulphurea Müll.
49	Donacia sp.	75	Hesperocorixa sp.
50	Dryops sp.	76	Hexatominae
51	Dytiscus sp.	77	Holocentropus dubius Rbr.
52	Ecnomus tenellus Ramb.	78	Holocentropus picicornis
53	Elmis aenea (Müller)	79	Hydropsyche angustipennis Curtis
54	Empididae	80	Hydropsyche nevae Kolenati

55	Ephemerella aurivillii Bengtsson	81	Hydropsyche pellucidula Curtis
56	Ephemera danica Müll.	82	Hydropsyche saxonica McLachlan
57	Ephemerella mucronata Bengtsson	83	Hydropsyche siltalai Döhler
58	Ephemera vulgata L.	84	Hydropsyche silvenii Ulmer
59	Erpobdellidae	85	Hydracarina
60	Erythromma najas (Hansem.)	86	Hydraena sp.
61	Gammarus lacustris Sars	87	Hydroptilidae
62	Gammarus pulex (L.)	88	Hydroporus sp.
63	Gerris sp.	89	Hygrotus sp.
64	Glossiphonia /Batracobdella	90	Ichnura elegans (Linden)
65	Glossosoma sp.	91	Ilybius sp.
66	Gomphidae	92	Isoperla sp.
67	Gyraulus sp.	93	Ithytrichia sp.
68	Gyrinus sp.	94	Laccophilus sp.
69	Haemopsis sanguisuga (L.)	95	Lepidostoma hirtum (Fabricius)
70	Haliplidae	96	Lepidoptera
97	Leptophlebia sp.	123	Ophiogomphus sp.
98	Leuctra fusca-digitata-hippopus	124	Orectochilus villosus (Müller)
99	Libellula sp.	125	Orthetrum sp.
100	Limnius volckmari (Panzer)	126	Oulimnius sp.
101	Limnephilidae	127	Oxyethira sp.
102	Limoniidae	128	Paraleptophlebia sp.
103	Lymnaea stagnalis (L.)	129	Pasifastacus leniusculus (Dana)
104	Lype sp.	130	Pediciinae
105	Micrasema setiferum Pictet	131	Pericoma sp.
106	Micronecta sp.	132	Perlodes dispar (Rambur)
107	Molanna angustata Curtis	133	Philopotamus montanus Don.
108	Molannodes tinctus Zett.	134	Phryganea bipunctata Retz.
109	Muscidae	135	Physa fontinalis (L.)
110	Mystacides sp.	136	Piscicola geometra
111	Myxas glutinosa (Müller)	137	Planorbis sp.
112	Nematoda	138	Platambus maculatus (L.)
113	Nematomorpha	139	Platycnemis pennipes (Pallas)
114	Nemoura sp.	140	Plectrocnemia sp.
115	Nemurella pictetii Klapalek	141	Polycentropus flavomaculatus Pictet
116	Nepa cinerea L.	142	Polycentropus irroratus Curtis

117	Neureclipsis bimaculata L.	143	Polycentropus sp.
118	Notonecta sp.	144	Potamopyrgus antipodarum Quoyi & Ga.
119	Odontocerum albicorne (Scopoli)	145	Protonemura meyeri (Pictet)
120	Oecetis testacea Curtis	146	Psychomyia pusilla Fabricius
121	Oligostomis reticulata L.	147	Psychodidae
122	Oligochaeta	148	Ptychopteridae
149	Radix peregra/ovata	162	Siphonoperla burmeisteri (Pictet)
150	Rantus sp.	163	Somatochlora sp.
151	Rhithrogena sp.	164	Sphaeriidae
152	Rhyacophila fasciata	165	Stagnicola palustris (Müller)
153	Rhyacophila nubila-obliterata	166	Stenelmis canaliculata (Gyllenhal)
154	Segmentina nitida (Müller)	167	Tabanidae
155	Sericostomatidae	168	Taeniopteryx nebulosa (L.)
156	Setodes argentipunctellus McLachlan	169	Tipulidae
157	Sialis fuliginosa-group	170	Trichostegia minor Curtis
158	Sialis lutaria-group	171	Turbellaria
159	Sigara sp.	172	Valvata macrostoma Mörch
160	Silo pallipes (Fabricius)	173	Valvata piscinalis (Müller)
161	Simuliidae	174	Wormaldia subnigra McL.

Region 20

1.	Hydrozoa	36.	Dinocras cephalotes (Curtis)
2.	Turbellaria	37.	Siphonoperla burmeisteri (Pictet)
3.	Radix peregra/ovata	38.	Corixidae
4.	Physa fontinalis (L.)	39.	Hydraena sp.
5.	Gyraulus acronicus-albus-laevis	40.	Elmis aenea (Müller)
6.	Sphaeriidae	41.	Sialis fuliginosa-group
7.	Glossiphonia sp.	42.	Rhyacophila fasciata
8.	Hydracarina	43.	Glossosoma sp.
9.	Gammarus lacustris Sars	44.	Agapetus sp.
10.	Oligochaeta	45.	Philopotamus montanus Don.
11.	Gammarus pulex (L.)	46.	Polycentropus flavomaculatus Pictet
12.	Collembola	47.	Hydropsyche pellucidula Curtis
13.	Ameletus inopinatus Eaton	48.	Hydropsyche saxonica McLachlan
14.	Baetis rhodani Pict.	49.	Hydropsyche siltalai Döhler
15.	Baetis muticus L.	50.	Hydropsyche silvenii Ulmer
16.	Acentrella lapponica Bengtsson	51.	Rhyacophila Nubila/Obliterata
17.	Nigrobaetis niger-digitatus	52.	Limnephilidae
18.	Centroptilum luteolum Müll.	53.	Apatania sp.
19.	Heptagenia sulphurea Müll.	54.	Chaetopteryx-Anitella
20.	Heptagenia fuscogrisea Retz.	55.	Ceraclea sp.
21.	Heptagenia dalecarlia Bengtsson	56.	Hydroptila sp.
22.	Leptophlebiidae	57.	Lepidostoma hirtum (Fabricius)
23.	Ephemerella aurivillii Bengtsson	58.	Micrasema gelidum McL.
24.	Baetis fuscatus group	59.	Sericostoma personatum (Spence)
25.	Baetis Vernus group	60.	Tipulidae
26.	Taeniopteryx nebulosa (L.)	61.	Ecclisopteryx dalecarlica Kol.
27.	Brachyptera risi (Mort.)	62.	Psychodidae
28.	Protonemura meyeri (Pictet)	63.	Potamophylax sp
29.	Amphinemura sp.	64.	Dixidae
30.	Nemoura sp.	65.	Chaoborus sp.
31.	Leuctra fusca-digitata-hippopus	66.	Ceratopogonidae
32.	Capnia sp.	67.	Chironomidae
33.	Diura nanseni (Kempny)	68.	Simuliidae
34.	Isoperla sp.	69.	Empididae
35.	Arcynopteryx compacta (McLachlan)	70.	Limoniidae

Region 22

1.	Turbellaria	25.	Baetis muticus L.
2.	Nematoda	26.	Baetis macani Kimm.
3.	Valvata sp.	27.	Baetis fuscatus group
4.	Radix peregra/ovata	28.	Baetis Vernus-group
5.	Planorbis planorbis (L.)	29.	Baetis niger-digitatus
6.	Oligochaeta	30.	Centroptilum luteolum Müller
7.	Bathyomphalus contortus (L.)	31.	Cloeon sp.
8.	Gyraulus acronicus-albus-laevis	32.	Heptagenia sulphurea Müller
9.	Ancylus fluviatilis (Müller)	33.	Heptagenia fuscogrisea Retz.
10.	Sphaeriidae	34.	Heptagenia dalecarlia Bengtsson
11.	Glossiphonia /Batracobdella	35.	Leptophlebia sp.
12.	Helobdella stagnalis (L.)	36.	Paraleptophlebia sp.
13.	Haemopsis sanguisuga (L.)	37.	Ephemerella ignita Poda
14.	Erpobdella sp.	38.	Ephemerella aurivillii Bengtsson
15.	Hydracarina	39.	Ephemerella mucronata Bengtsson
16.	Argyroneta aquatica (Clerk)	40.	Ephemera vulgata L.
17.	Asellus aquaticus L.	41.	Ephemera danica Müller
18.	Gammarus sp.	42.	Caenis horaria L.
19.	Pallasea quadrispinosa Sars	43.	Taeniopteryx nebulosa (L.)
20.	Collembola	44.	Brachyptera risi (Mort.)
21.	Siphlonurus sp.	45.	Protonemura meyeri (Pictet)
22.	Ameletus inopinatus Eaton	46.	Amphinemura sp.
23.	Metretopus borealis Eaton	47.	Nemurella pictetii Klapalek
24.	Baetis rhodani Pict.	48.	Nemoura sp.
49.	Leuctra nigra (Olivier)	50.	Hydraena sp.
51.	Leuctra fusca-digitata-hippopus	52.	Elmis aenea (Müller)
53.	Capnia sp.	54.	Limnius volckmari (Panzer)
55.	Capnopsis schilleri (Rostock)	56.	Oulimnius sp.
57.	Diura bicaudata (L.)	58.	Chrysomelidae
59.	Diura nanseni (Kempny)	60.	Sialis lutaria-group
61.	Isoperla sp.	62.	Sialis fuliginosa-group
63.	Arcynopteryx compacta (McLachlan)	64.	Rhyacophila septentrionis McLachlan
65.	Dinocras cephalotes (Curtis)	66.	Rhyacophila fasciata
67.	Siphonoperla burmeisteri (Pictet)	68.	Rhyacophila nubila/obliterata
69.	Coenagrion sp.	70.	Glossosoma sp.

71.	Erythromma najas (Hansem.)	72.	Agapetus sp.
73.	Calopteryx virgo (L.)	74.	Philopotamus montanus Don.
75.	Onychogomphus forcipatus (L.)	76.	Wormaldia subnigra McL.
77.	Cordulegaster boltonii (Donovan)	78.	Chimarra marginata L.
79.	Aeshna sp.	80.	Neureclipsis bimaculata L.
81.	Cordulia aenea (L.)	82.	Plectrocnemia sp.
83.	Somatochlora sp.	84.	Polycentropus flavomaculatus Pictet
85.	Leucorrhinia sp.	86.	Polycentropus irroratus Curtis
87.	Gerris sp.	88.	Holocentropus sp.
89.	Notonecta lutea O.F. Müller	90.	Cyrnus trimaculatus Curtis
91.	Pleidae	92.	Cyrnus flavidus McL.
93.	Corixidae	94.	Psychomyiidae
95.	Haliplidae	96.	Hydropsyche Sp.
97.	Dytiscidae	98.	Hydropsyche pellucidula Curtis
99.	Gyrinus sp.	100.	Hydropsyche angustipennis Curtis
101.	Hydropsyche saxonica McLachlan	102.	Silo pallipes (Fabricius)
103.	Hydropsyche siltalai Döhler	104.	Lepidostoma hirtum (Fabricius)
105.	Hydropsyche silvenii Ulmer	106.	Micrasema sp.
107.	Hydropsyche nevae Kolenati	108.	Brachycentrus subnubilus Curtis
109.	Cheumatopsyche lepida Pictet	110.	Sericostoma personatum (Spence)
111.	Arctopsyche ladogensis Kolenati	112.	Lepidoptera
113.	Agraylea sp.	114.	Tipulidae
115.	Ithytrichia sp.	116.	Limonia sp.
117.	Oxyethira sp.	118.	Dicranota sp.
119.	Phryganea bipunctata Retz.	120.	Pediciinae
121.	Agrypnia sp.	122.	Eriopterinae
123.	Oligostomis reticulata L.	124.	Hexatominae
125.	Limnephilidae	126.	Psychodidae
127.	Molanna angustata Curtis	128.	Ptychoptera sp.
129.	Molanna albicans Zett.	130.	Dixidae
131.	Molannodes tinctus Zett.	132.	Ceratopogonidae
133.	Beraeodes minutus (L.)	134.	Chironomidae
135.	Athripsodes sp.	136.	Simuliidae
137.	Mystacides sp.	138.	Empididae
139.	Triaenodes bicolor	140.	Atherix sp.
141.	Oecetis sp.	142.	Tabanidae
143.	Ceraclea sp.	144.	Muscidae
145.	Goera pilosa (Fabricius)	146.	Limoniidae