Environment Canada   Environnement Canada

National Agri-Environmental Standards Initiative (NAESI)

Report No. 4-7

Assessing Biotic Integrity and Local Habitat Quality in Agricultural Landscapes in Support of Developing Standards

Technical Series 2008

Photos:
Bottom Left- clockwise

Fraser Valley near Abbotsford, B.C.: Wayne Belzer, Pacific Yukon Region, Environment Canada
Crop spraying: Corel CD photo # 95C2840
Elk Creek, BC: Joseph Culp, National Water Research Institute, Environment Canada
Prairie smoke and bee: Emily Wallace, Prairie Northern Region, Environment Canada

This report can be cited as follows:

Kirk, D.A. 2008. Assessing Biotic Integrity and Local Habitat Quality in Agricultural Landscapes in Support of Developing Standards. National Agri-Environmental Standards Initiative Technical Series Report No. 4-7. 104 p.

Prepared and published by
Environment Canada
Gatineau, QC

March 2008

NATIONAL AGRI-ENVIRONMENTAL STANDARDS INITIATIVE TECHNICAL SERIES

ASSESSING BIOTIC INTEGRITY AND LOCAL HABITAT QUALITY IN AGRICULTURAL LANDSCAPES IN SUPPORT OF DEVELOPING STANDARDS

REPORT NO. 4-7

© Her majesty the Queen in Right of Canada, represented by the Minister of the Environment, 2008. All rights reserved. Reproduction authorized if source is acknowledged. The reproduction must be presented within its proper context and must not be used for profit.

NOTE TO READERS

The National Agri-Environmental Standards Initiative (NAESI) is a four-year (2004-2008) project between Environment Canada (EC) and Agriculture and Agri-Food Canada (AAFC) and is one of many initiatives under AAFC's Agriculture Policy Framework (APF). The goals of the National Agri-Environmental Standards Initiative include:

- Establishing non-regulatory national environmental performance standards (with regional application) that support common EC and AAFC goals for the environment
- Evaluating standards attainable by environmentally-beneficial agricultural production and management practices; and
- Increasing understanding of relationships between agriculture and the environment.

Under NAESI, agri-environmental performance standards (i.e., outcome-based standards) will be established that identify both desired levels of environmental condition and levels considered achievable based on available technology and practice. These standards will be integrated by AAFC into beneficial agricultural management systems and practices to help reduce environmental risks. Additionally, these will provide benefits to the health and supply of water, health of soils, health of air and the atmosphere; and ensure compatibility between biodiversity and agriculture. Standards are being developed in four thematic areas: Air, Biodiversity, Pesticides, and Water. Outcomes from NAESI will contribute to the APF goals of improved stewardship by agricultural producers of land, water, air and biodiversity and increased Canadian and international confidence that food from the Canadian agriculture and food sector is being produced in a safe and environmentally sound manner.

The development of agri-environmental performance standards involves science-based assessments of relative risk and the determination of desired environmental quality. As such, the National Agri-Environmental Standards Initiative (NAESI) Technical Series is dedicated to the consolidation and dissemination of the scientific knowledge, information, and tools produced through this program that will be used by Environment Canada as the scientific basis for the development and delivery of environmental performance standards. Reports in the Technical Series are available in the language (English or French) in which they were originally prepared and represent theme-specific deliverables. As the intention of this series is to provide an easily navigable and consolidated means of reporting on NAESI's yearly activities and progress, the detailed findings summarized in this series may, in fact, be published elsewhere, for example, as scientific papers in peer-reviewed journals.

This report provides scientific information to partially fulfill deliverables under the Biodiversity Theme of NAESI. This report was written by D.A. Kirk of Aquila Conservation & Environment Consulting. The report was edited and formatted by Denise Davy to meet the criteria of the NAESI Technical Series. The information in this document is current as of when the document was originally prepared. For additional information regarding this publication, please contact:

Environment Canada
National Agri-Environmental Standards
Initiative Secretariat
351 St. Joseph Blvd. 8th floor
Gatineau, QC
K1A 0H3
Phone: (819) 997-1029
Fax: (819) 953-0461

NOTE À L’INTENTION DES LECTEURS

L’Initiative nationale d’élaboration de normes agroenvironnementales (INENA) est un projet de quatre ans (2004-2008) mené conjointement par Environnement Canada (EC) et Agriculture et Agroalimentaire Canada (AAC) et l’une des nombreuses initiatives qui s’inscrit dans le Cadre stratégique pour l’agriculture (CSA) d’AAC. Elle a notamment comme objectifs :

- d’établir des normes nationales de rendement environnemental non réglementaires (applicables dans les régions) qui soutiennent les objectifs communs d’EC et d’AAC en ce qui concerne l’environnement;
- d’évaluer des normes qui sont réalisables par des pratiques de production et de gestion agricoles avantageuses pour l’environnement;
- de faire mieux comprendre les liens entre l’agriculture et l’environnement.

Dans le cadre de l’INENA, des normes de rendement agroenvironnementales (c.-à-d. des normes axées sur les résultats) seront établies pour déterminer les niveaux de qualité environnementale souhaités et les niveaux considérés comme réalisables au moyen des meilleures technologies et pratiques disponibles. AAC intégrera ces normes dans des systèmes et pratiques de gestion bénéfiques en agriculture afin d’aider à réduire les risques pour l’environnement. De plus, elles amélioreront l’approvisionnement en eau et la qualité de celle-ci, la qualité des sols et celle de l’air et de l’atmosphère, et assureront la compatibilité entre la biodiversité et l’agriculture. Des normes sont en voie d’être élaborées dans quatre domaines thématiques : l’air, la biodiversité, les pesticides et l’eau. Les résultats de l’INENA contribueront aux objectifs du CSA, soit d’améliorer la gérance des terres, de l’eau, de l’air et de la biodiversité par les producteurs agricoles et d’accroître la confiance du Canada et d’autres pays dans le fait que les aliments produits par les agriculteurs et le secteur de l’alimentation du Canada le sont d’une manière sécuritaire et soucieuse de l’environnement.

L’élaboration de normes de rendement agroenvironnementales comporte des évaluations scientifiques des risques relatifs et la détermination de la qualité environnementale souhaitée. Comme telle, la Série technique de l’INENA vise à regrouper et diffuser les connaissances, les informations et les outils scientifiques qui sont produits grâce à ce programme et dont Environnement Canada se servira comme fondement scientifique afin d’élaborer et de transmettre des normes de rendement environnemental. Les rapports compris dans la Série technique sont disponibles dans la langue (français ou anglais) dans laquelle ils ont été rédigés au départ et constituent des réalisations attendues propres à un thème en particulier. Comme cette série a pour objectif de fournir un moyen intégré et facile à consulter de faire rapport sur les activités et les progrès réalisés durant l’année dans le cadre de l’INENA, les conclusions détaillées qui sont résumées dans la série peuvent, en fait, être publiées ailleurs comme sous forme d’articles scientifiques de journaux soumis à l’évaluation par les pairs.

Le présent rapport fournit des données scientifiques afin de produire en partie les réalisations attendues pour le thème de la biodiversité dans le cadre de l’INENA. Ce rapport a été rédigé par D.A. Kirk d'Aquila Conservation & Environment Consulting. De plus, il a été révisé et formaté par Denise Davy selon les critères établis pour la Série technique de l’INENA. L’information contenue dans ce document était à jour au moment de sa rédaction. Pour plus de renseignements sur cette publication, veuillez communiquer avec l’organisme suivant :

Secrétariat de l’Initiative nationale
d’élaboration de normes
agroenvironnementales
Environnement Canada

351, boul. St-Joseph, 8eétage
Gatineau (Québec) K1A 0H3
Téléphone : (819) 997-1029
Télécopieur : (819) 953-0461

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

EXECUTIVE SUMMARY

Assessing ecological or habitat quality is a formidable challenge that needs to be met so that sites/landscapes altered by human disturbance can be restored towards their former unimpaired state. While much effort has been devoted towards selection of biotic and environmental indicators, it is being increasingly recognized by some that an ecosystem or community approach is needed, and that selecting individual species or environmental variables alone as surrogates for biodiversity is fraught with difficulties and pitfalls. This necessitates the monitoring of multi-species taxa groups as recognized by leading terrestrial monitoring programs such as the Alberta Biodiversity Monitoring Institute, the United States National Parks Service and the USDA Forest Service and aquatic monitoring programs such as RIVPACS (UK), AUSRIVAS (Australia), CABIN (Canada) and EUWFD (Europe). Once multi-species taxa data are collected as part of a bioassessment monitoring program, the challenge is to seek an approach that will not only provide a sensitive and accurate measure of ecological quality, but can also be used to set standards and guidelines for management.

In this report, we first provide a critique of four reports commissioned by the National Agri-Environmental Standards Initiative (NAESI) of Environment Canada to review methods for assessing local habitat quality, at the level of the individual farm. These reports dealt separately with potential frameworks and approaches for four ecosystems – woodlands, grasslands, riparian areas and wetlands. All four reports lacked emphasis on an empirical (data-driven) base from which to assess site quality. While they referred to multimetric indices, they did not address the advantages and disadvantages of their use or how they could be applied in terrestrial and aquatic habitats. None specifically mentioned the use of multivariate techniques, though one report referred to similarity indices. Moreover, the reports lacked a holistic approach to the effects of

scale; habitat and stressors need to be considered at multiple scales for site level assessments of biotic and habitat quality.

Second, we review two of the main methods used in the toolbox of ecological quality assessment: Multimetric indices and multivariate ordination. While both have strengths and weaknesses and can be used in conjunction with each other, our preference is for multivariate ordination. The reason for this is that this approach is probably more objective than multimetric indices, as well as having less redundancy and additive effects. Moreover, species abundance or occurrence data can be modeled with natural environmental factors and stressors within a framework of flexibility and inclusiveness not shown in multimetric indices. We also believe that ordination results are easier to visualize (with guidance and interpretation by experts) – the concept of viewing a graphical 2 or 3D plot where sites close together are more similar in species composition and abundance than those far apart is an intuitive concept easy to grasp by managers and laypersons alike. One previously used approach that we follow is the Reference Condition Approach (RCA) but our application of RCA differs in that we did not attempt to identify pristine sites, but only those that were relatively unimpaired; "pristine" sites do not occur in most agricultural landscapes. However, we emphasize the importance of including sites as least disturbed as possible in the spectrum of candidate sites.

In the RCA approach a range of minimally impacted sites is used to characterize the biological condition of a region and to explain as much of the natural variability that occurs in these sites as possible; following this a test site is compared to a subset of the reference sites (with similar environmental characteristics not influenced by humans) or to all the reference sites using probability weightings.

Third, we present a case study to demonstrate our approach for one of the ecozones (Prairie

Potholes ecozone); we used pre-existing survey data on plant communities collected as part of the Saskatchewan Correlates of Biodiversity study. While these data were not specifically collected using a RCA approach, they include sites that could be considered relatively unimpaired ("Wild" sites), as well as sites with no pesticide use (organically-farmed sites), and sites with pesticide use (nonorganic farms using chemical pesticides and minimum tillage). The stressor gradient available was invasive plant species richness which we hypothesized would be higher in impaired than unimpaired sites. We also included a range of environmental variables measured at different scales, including sites (quadrats), ponds (intermediate) and quarter sections (64 ha). We caution that the case study modelling was done for demonstration purposes only and that there was no evidence to support our hypothesis that invasive species richness was lowest on unimpaired sites. We also emphasize that because the case study was done *a posteriori* and was not set up with an RCA approach in mind, the designation of minimally impacted (reference) sites chosen is subjective; usually specific criteria are developed *a priori* to select reference sites. However, some support for our *a priori* classification of sites was that of the 262 plant species found around farm ponds, 51 were unique to wild sites, 24 to organic sites, 23 to minimum tillage sites and 13 to conventional sites.

We use three types of multivariate methods to compare results: a) a standard RCA approach – cluster analysis and distance-based ordination (non-metric multidimensional scaling, nMDS); b) and a modified, simpler RCA type approach using weighted-averaging (eigenanalysis) ordination (canonical correspondence analysis, CCA); and c) a wetland plant index (WPI) based on the species tolerances and standard deviations from a CCA on all sites and habitat variables. We followed a series of steps as in the RCA approach (a). These were to: 1) Run a cluster analysis on "relatively unimpaired" sites (wild and organic) to derive groups of sites that represent a community assemblage; 3) Deploy Analysis of Similarity (ANOSIM) to determine if groups of sites established *a priori* are significantly different and linear discriminant function analysis to determine which biophysical variables not influenced by human disturbance were the best predictors for the site cluster groups; 4) Use regression models (principal axis correlation in nMDS) to determine the relationship with ordination (nMDS) axes; 5) Substitute test sites in an ordination model to predict the expected species composition and compare this to that observed; and 6) draw probability ellipses around reference sites and see where the test sites lie relative to reference sites. The ellipses can be used to set a measure of difference or similarity between reference and test.

For the weighted-averaging ordination approach (b) we: 1) performed a detrended correspondence analysis (DCA) to determine whether a linear or unimodal model was appropriate; 2) Conducted a canonical correspondence analysis (CCA) on "unimpaired" sites to determine the significant environmental gradients using forward selection of biophysical variables; 3) Conducted a partial CCA (*p*CCA), using the number of invasive species as the environmental (stressor) variable while controlling for significant biophysical variables as covariates; 4) Included test sites as supplementary in the ordination (i.e., passive sites that did not affect the analysis) and 5) Included the centroids for wild and organic sites as supplementary and then determined the distance and trajectory from the test site to the centroid in the direction of decreased richness of introduced species. We also tested a third approach c) which was to conduct a CCA to determine the main environmental gradients in the data and then develop a Wetland Plant Index (WPI) based on the tolerance and niche breadth of individual species.

Techniques such as the RCA can be used to set standards and targets; this would involve several steps including: 1) determining the magnitude and deviation from reference condition (based on vector distances between sites and directionality on ordination axes in response to stressor gradients). For example, a distance in ordination space of 2 standard deviations of species turnover can be used as a measure of change and be set as a standard for assessing changes in the ecological quality of sites; 2) investigating why the site failed to meet the Reference Condition (using information on biota composition and abundance, as well as relationship with stressors and other predictors); and 3) building scenarios to predict the effect of restoration or increasing stressor values using Generalized Linear Models (GLM) for individual stressors for example. The CCA modification of the RCA approach could also help set standards; for example, the length of vector between the impaired (test) site (s) and centroid of unimpaired sites provides information on the change in community composition to attain reference condition. If surveys are repeated over time (post management) then this type of analysis provides information on the rate of change in communities; setting a target (e.g., 5 years to attain a reference condition following mitigative management) is another way of setting guidelines.

In our case study, specific management would be to 1) actively reduce the number of invasive species (or manage another stressor) using targeted management and then 2) re-survey plant species at the site and 3) re-run the analyses to test whether site quality had improved in terms of a reduced number of invasive species. This uses a combination of a benchmark or baseline with time zero, since sites need to be monitored over time to evaluate improvement in their condition.

These types of multivariate frameworks and approach can be used for any one of the four ecosystem types. We recommend as a next step that these approaches be tested by specifically collecting data using an RCA approach (particularly for terrestrial ecosystems where the approach has not been used) and to see how effective the framework is for setting standards and guidelines.

1 INTRODUCTION

Landscapes are increasingly fragmented by anthropogenic disturbances and remnant biotic communities may vary in quality from highly modified to relatively undegraded.

Much debate has centred on how best to evaluate habitat quality[1] and biodiversity value at these sites (e.g., Karr and Chu, 1999; Taft et al., 2006), and a wide variety of indices have been developed to do so. These indices of local ecological quality need to fulfill several requirements. First, they must be spatially explicit so that they can be tied to habitat and stressors at multiple scales (though bioassessment is at the level of individual sites). Second, they need to be sensitive to adaptive management such as restoration efforts implemented to reduce the impact of environmental stressors. Third, they need to be monitored over time so that changes in biotic response to stressors can be evaluated; and fourth, they need to be sensitive to changes in vegetation succession. Among several obstacles is the challenge of identifying relatively unimpaired ("reference[2]") sites and allowing for patterns of natural dynamic vegetation succession. In the Reference Condition Approach (RCA), a range of sites relatively unimpacted by human disturbance are used to characterize the biological condition of a region; following this a test site is compared to a subset of the reference sites (with similar environmental characteristics) or to all the reference sites using probability weightings (Reynoldson et al., 1997; Bailey et al., 2004). Establishing benchmarks or reference sites is important so that sites which are relatively impaired habitat quality because of human stressors can be compared with them and restored to a similar state.

Comparing relatively impaired sites to relatively unimpaired ones can be done: 1) spatially – by

[1] Note that here we do not use "Ecological Integrity" since this is a nebulous concept and hard to define
[2] Although we use the term "Reference Condition Approach" this is meant to describe relatively unimpaired sites, and not pristine sites, since no sites can be considered truly natural – in agricultural landscapes the term "least disturbed" may be more appropriate

comparing biological condition between sites that are least disturbed (or relatively "pristine" in some landscapes) and those that are impaired by anthropogenic disturbance; 2) temporally – by comparing changes in site habitat quality over time (time zero); and 3) by comparing sites to known thresholds for maintaining composition, ecological processes and function. For 1) a way of comparing relatively pristine sites to impaired sites is to match sites for biophysical characteristics at different ends of an impairment spectrum with the restoration goal of moving impaired sites to an unimpaired condition. Assessing site habitat quality over time is an integral part of this process. Although generally insufficient information is available about ecosystems to deploy the third option, in some cases there are data available such as in possible threshold effects shown by forest birds to forest loss and fragmentation in urban-agricultural landscapes. For some species it has been demonstrated that landscape configuration is important below a critical amount of suitable habitat, often between 10-30% (nonlinear fragmentation hypothesis; e.g., Andrèn, 1994; Fahrig, 1998, Flather and Bevers, 2002). However, it has been suggested that rather than being sharp thresholds, there is a continuum of response (e.g., Villard et al., 1999).

Before such an assessment of assemblage composition can be tackled for agricultural landscapes, we need to determine what type of framework or approach is best suited for this purpose. Some work has already been done on evaluating single species populations and an extremely promising technique is the approach to assess relative intactness deploying logistic regression models for single species (Nielsen et al., 2007) and extrapolating the model to the condition in the absence of a human stressor (such as road density). Other measures of intactness, such as the biodiversity intactness index (BII; Scholes and Biggs, 2005), have been criticized because scores can be inflated (underestimating degradation) and they may not represent the range of variation in living biodiversity (quantity can be overemphasized at the expense of variability; see Faith et al., 2008).

A multi-species or community-level assessment is essential to complement single species statistical modelling and to provide an index of biodiversity[3] which can be used to determine the main ecosystem drivers and to factor out human anthropogenic change from natural disturbance. This is partly because adequate models can only be derived for a limited number of species – and selecting the "best" species as indicators can be challenging and sometimes arbitrary. Another advantage of a multispecies approach is that it can help drive the identification and selection of individual species indicators.

Many programs have explored the use of multimetric indices (MMIs), such as the index of biotic integrity (IBI – Karr and Chu, 1999), to evaluate site condition, particularly in aquatic systems. For example, the United States Protection Agency (US EPA) has invested heavily in the MMI approach and this is used to evaluate stream condition using benthic macroinvertebrates and fish as indicators of stream condition (Klemm et al., 2002). The focus has been on aquatic rather than terrestrial systems for four major reasons: 1) aquatic systems are more closed than terrestrial ones, and the origin of reference conditions and impairment came from point-source pollution studies; 2) Clean water acts in various parts of the world have been instrumental in the development of standards in relation to water quality, which in turn has forced bioassessments of aquatic biological condition; 3) Aquatic bioassessments have focused more on a multi-species approach whereas terrestrial ones have tended towards using single species; and 4) to some extent the inputs and relationships in aquatic systems are simpler to understand than those in terrestrial ones (see Andreasan et al., 2001; Browder et al., 2002; Bryce et al., 2002).

[3] Although it has been defined as ''the variety of living organisms, the ecological complexes in which they occur, and the ways in which they interact with each other and the physical environment'' (Redford and Richter, 1999), ''biodiversity'' is most commonly used to refer to elements of biota (species and communities) which is how we refer to it here.

A complementary approach to use of MMIs is to use multivariate ordination and other multivariate techniques such as classification (e.g., cluster analysis and classification and regressions trees, CART). Multivariate techniques (cluster analysis, linear discriminant analysis and non-metric multidimensional scaling) have been used in aquatic bioassessment programs first in Europe (**R**iver **I**n**v**ertebrate **P**rediction **a**nd **C**lassification **S**ystem; RIVPACs - Wright et al., 1984; Wright et al., 1993; Simpson and Norris, 2000), and later in Australia (the **Aus**tralian **Riv**er **A**ssessment **S**ystem – AUSRIVAS; Smith et al., 1999), Canada (**C**anadian **A**quatic **Bi**omonitoring **N**etwork - CABIN; Reynoldson et al., 1999) and the United States (Hawkins et al., 2000). For example, in CABIN, reference sites are first classified based on organism assemblages (benthic invertebrates), then the community types are related to a set of habitat features that are independent of human disturbance. This relationship then allows a new site to be probabilistically assigned to a reference group so that a comparison can be made between observed and expected conditions (Reynoldson et al., 1997; Bailey et al., 2004). Because this involves categorizing sites into groups using classification methods (e.g., cluster analysis or techniques such as Classification and Regression Tree CART analysis – see Cao et al., 2007, Lougheed et al., 2007), and this can be artificial, it has been argued that it is better to examine a continuum along a stressor gradient rather than clustering sites into groups (Linke et al., 2005; T. Reynoldson pers. comm.). This can be done using ordination combined directly or *a posteriori* with multiple regression analyses (Linke et al., 2005).

Ordination techniques reduce complex multi-dimensional parameters to visual 2-3 dimensional graphical interpretation, where sites close together have more similar species composition than those far apart (ter Braak, 1995). They are believed by some to be more objective than MMIs and have less redundancy and additive effects. However, they have also been used in conjunction with MMIs (see Reynoldson et al., 1997, Sylvestre et al., 2005) and can be viewed as just one tool in a toolbox of modelling and metric development.

1.1 Why is this project needed?

Previous studies conducted under the auspices of the National Agri-Environment Standards Initiative (NAESI) of Environment Canada have focused on modelling habitat required by focal and other species groups at the landscape scale (patch size) using Population Viability Analysis (PVA) and optimization models, as well as habitat amount needed at specific scales to support viable populations (e.g., Baldwin et al., 2006). It was recognized that a multi-scale approach was needed and that these landscape level analyses using remote sensing need to be supplemented by habitat quality indices at the level of individual farms. The goal of this project is to develop a framework and approach to assessing the habitat quality and biodiversity value of sites in semi-natural habitats in agricultural landscapes. While the bioassessment framework presented is at the site level, it is critically important to recognize that multi-scale environmental features are incorporated into the approach and modelling.

Specifically, we:

1) Review and compare four reports (focusing on woodlands, grasslands, riparian areas and wetlands, respectively) describing approaches to measuring local scale biotic and habitat integrity in agricultural landscapes in Canada;
2) Evaluate the pros and cons of two or more main methods for assessing habitat quality; these are multimetrics such as the Index of Biotic Integrity (IBI) widely used to assess water quality in the United States and parts of Canada; and multivariate ordination, a more

flexible modelling approach that has also been widely deployed in Europe, parts of Canada, the United States and Australia;

3) Synthesize the information and recommendations made to produce an overall framework/approach that can be applied to assess local habitat quality and set performance standards at the farm scale. This will be done through an over-arching framework for all ecosystems;
4) As a case study, we apply three analytical frameworks to a study of plant assemblages at pond edges in the Prairie Potholes region of Saskatchewan. The first uses a Reference Condition Approach (RCA) which has been widely tried and tested in aquatic ecosystems (Bailey et al., 2004). The second approach is a modification of the RCA and a simpler integrated analytical framework and the third a modification of the Wetland Macrophyte Index (WMI) used by P. Chow-Fraser and co-workers.

2 METHODS

2.1 Comparing evaluations of biotic/habitat quality

We reviewed four reports describing assessment protocols for wetlands, grasslands, woodlands and riparian areas. As a first step in this review process we tabulated differences and similarities between the reports. We identified major weaknesses or gaps in the reports and proposed assessment frameworks. All reports attempted to address two levels of biotic and habitat assessment. First, a professional assessment based on compound metrics of biological “integrity” and, second, a layperson’s rapid environmental assessment (REA) that can be carried out by individual farmers/ landowners or other volunteers. This report will focus on level 1 (the professional assessment) for two reasons: 1) while the REAs carried out by farmers/landowners are extremely important for education and outreach they are unlikely to provide the quality of data necessary to assess habitat quality empirically using a modelling approach; 2) they may provide subjective assessments of local habitat quality and monitoring responses to management over time, but it will be hard to integrate them with level 1 assessments (unless using standardized protocols and training of personnel to the same standards, which seems unlikely).

2.2 Evaluations of multimetrics and multivariate ordination

As background to the applicability of MMIs and/or multivariate techniques for local site quality assessment we reviewed programs conducted by the US EPA, as well as water quality assessment programs such as RIVPACs, AUSRIVAs and CABIN. We then prepared a table with a series of questions about MMIs and multivariate techniques which we attempted to answer. We sent drafts of the table to eight reviewers (see acknowledgements) and solicited comments which were then incorporated into the final version. Note that we did not specifically ask these reviewers to choose between multimetric indices and multivariate techniques. The views of reviewers were clearly reflected in their backgrounds and experiences with the different techniques and relatively few people who conduct MMIs are familiar with multivariate techniques and vice versa.

However, for a number of reasons, we advocate the use of multivariate techniques, particularly because of model flexibility and visual representation in an ordination diagram (see ter Braak, 1995; Lêps and Smilauer; 2007; Legendre and Legendre, 1998; Urban, 2006). We propose to outline a potential framework application using three main approaches – 1) a reference condition approach using CABIN protocols, including non-metric multidimensional scaling and cluster analysis (Reynoldson et al., 1999, Sylvestre et al., 2005); 2) an RCA approach using weighted averaging ordination (canonical correspondence analysis, CCA) but omitting the clustering step; and 3) an assessment of site quality using CCA and following the Wetland “indices” approach of Chow-Fraser (2006) and colleagues (Seilheimer and Chow-Fraser, 2006; Croft and Chow-Fraser, 2007). Conceptually, all of these approaches compare matrices – a species matrix and an environmental matrix, which is a better option than looking at either matrix separately.

2.3 Case study using 2 "reference" condition approaches, and Wetland Plant Index (WPI) approach

Of the 10 candidate case studies of the distribution and abundance of various taxa in farming landscapes selected, we had direct or indirect involvement in nine (Appendix A). Four were conducted as part of the Saskatchewan Correlates of Biodiversity (SCB) study and include inventories of plants, and invertebrates (aquatic and terrestrial) on four farm types – wild, organic, minimum tillage and conventional (chemical) farms. Two projects involved comparisons of organic and conventional farming systems, one an analysis of bird use of croplands in southern Ontario, and another a combination of two investigations comparing bird species composition and abundance between woodland patches of different size in eastern Ontario (see Appendix A).

We chose the SCB plant study for two main reasons; a) it involved a taxa group that has been shown to be highly sensitive to anthropogenic stressors (e.g., Albert and Minc, 2004; Fuller et al., 2005; Croft and Chow-Fraser, 2007); b) it included relatively "wild" areas that could be used as relatively unimpaired sites, together with "restored" sites (organic farms) and relatively impaired sites (nonorganic farms and minimum tillage). Normally, reference sites would not be identified in this manner (i.e., *a priori* based on farming types) but according to specific criteria; it is important to acknowledge that identifying such sites is difficult and can involve some subjectivity (e.g., expert opinion from local/regional management staff). The main general recommended criteria for choosing reference sites is that they have *minimal exposure to the stressors of concern* (see Bailey et al., 2004) or that they are in the "best condition that could be expected"

(Reynoldson et al., 1997). For terrestrial situations, and especially in agricultural landscapes, near pristine sites are likely not available and relevant “least disturbed” sites could be selected and compared against those with more substantial land use stressors (See Discussion for detailed criteria for selection of reference sites; Davies, 1994).

The data consisted of 42 upland sites (plots), with 146 recorded native plant species. In addition, there were 30 introduced species and 23 extrinsic variables (some include descriptions of general vegetation type). Note that among the extrinsic variables are trend surface variables (used to incorporate variability attributable to spatial autocorrelation) and habitat features measured at multiple scales (e.g., landscape scale - variables included wetland area, soil types, precipitation; local scale; local scale - variables included % shoreline treed and % shoreline in willow/shrub). We used three types of multivariate approaches to demonstrate how our framework could be applied to semi-natural habitats within agricultural landscapes in Canada. These were: 1) an RCA approach using cluster analysis and hybrid non-metric multidimensional scaling (hMDS using the Bray-Curtis distance; see Reynoldson et al., 1997, 1999; T. Reynoldson, pers. comm.); 2) a variation on the RCA approach using canonical correspondence analysis (CCA); and 3) an adaptation of the Wetland Quality Index (WQI) or Wetland Macrophyte Index (WMI) developed by P. Chow-Fraser and her students (Lougheed and Chow-Fraser, 2002; Seilheimer and Chow-Fraser, 2006; Croft and Chow-Fraser, 2007).

Approach 2) is closest to the ANNA model used by Linke et al. (2005), except that we used weighted-averaging ordination rather than distance-based methods. In the ANNA framework, Linke et al. (2005): 1) weighted the predictor variables using a multivariate approach analogous to principal axis correlations; 2) calculated the weighted Euclidian distance from a test site to reference sites based on the environmental predictors; 3) predicted the faunal composition based on the nearest reference sites; and 4) calculated an observed/expected (O/E) analogous to RIVPAC/AUSRIVAS. We also omitted the clustering step that is used in RIVPACs and AUSRIVAs. Usually sites are clustered into different groups using two-way-indicator-species-analysis TWINSPAN or cluster analysis (single-linkage); following this discriminant function analysis (DFA) is used to identify the environmental variable that best discriminates among the groups (as in approach 1 above).

We used the two main ordination methods for the case studies: weighted-averaging (or eigenanalysis) ordination using the software CANOCO version 4.5 (ter Braak and Smilaeur, 2002), PATN vs 3.1 for windows (Belbin, 2008) and Primer (Clarke and Gorley, 2001) for the distance-based ordination. The reason we chose CANOCO was because it includes various options not available in some other packages (e.g., use of covariates – which means that the effects of background natural features can be controlled for while examining the effect of stressor variables). We chose PATN because it deploys vector analysis (principal axis correlation PCC) and hence at some level is comparable to CCA. Moreover, PATN and Primer were the softwares used in developing the BEAST (Benthic Assessment of Sediment) assessment method as part of the CABIN protocols (Canadian Aquatic Biomonitoring Network; Reynoldson et al., 1999). We did not use PC-Ord for the nMDS analysis because it has been suggested that the solution is unstable (P. Minchin, pers. comm., J.L. Pearce pers. comm.). For the standard RCA approach we used SYSTAT for the DFA and ANOVA analyses (SYSTAT version 9.0; SPSS Inc., 1999).

We acknowledge that other ordination methods are available – another possible candidate is Fuzzy set ordination (FSO) which is the constrained form of nMDS (Boyce and Ellison, 2001); however, this analysis is in development in the R-language (R. Boyce, pers. comm.) and so was not considered for our case study. Note that all of the above analyses can be implemented in the

R-language which is extremely flexible and versatile (R-Core Development Team, 2007).

1) CABIN RCA approach (cluster analysis and nMDS)

In both of the RCA approaches (1 and 2 below), wild or organic sites were treated as the least disturbed (reference sites). It should be noted that the usual method in RCA for defining reference sites is to have an independent set of criteria, and to sample all possible habitat types within a geographic area (see Discussion). For example, reference criteria could be: 1) no pesticide use in the last 5 years; or 2) no agricultural activity over the past 3 years. One criteria for choosing wild and organic sites as reference was the number of species unique to each farm type. Of the 262 species found at farm ponds, 51 were unique to wild sites, 24 to organic sites, 23 to minimum tillage sites and 13 to conventional sites. This suggested that wild and organic sites had different plant communities to farms that use chemicals. Reference sites can also be selected and modified) later using multivariate models, provided that a gradient of unimpaired to impaired sites is surveyed.

Model Building

- Step 1: Because of the relatively large number of species for the sites, the data were reduced by excluding non-native species; these introduced species were treated as the stressor variable. Analyses were performed on species that occurred at >5% of sites and that had % occurrence values of ≤ 5%. However, although Linum spp. only occurred at 2.3% of the plots, in one plot it represented 30.8% of the species found, and was therefore retained. The exclusion of rare taxa resulted in a final species data matrix of 83 species.
- Step 2: We then classified species at relatively unimpaired (“reference”) sites (wild and organic sites). Classification was done using cluster analysis (unweighted pair group means

average clustering), setting the beta coefficient at -0.1 (in PATN, Belbin, 2008). Prior to clustering, the Bray-Curtis similarity coefficient was used to construct the similarity matrix and the data were not transformed. A maximum of four groups was examined.

- Step 3: A combination of Analysis of Similarity (ANOSIM) and nMDS ordination was used to decide between a few different cluster solutions. Then a discriminant function analysis (DFA) was performed on 2-3 of alternative models to determine the optimal model. To identify candidate habitat variables, several approaches were used:
- Univariate ANOVA and pairwise testing with the groups formed by the species data as the factor. This determined which habitat attributes show the greatest difference among the groups, and used *post hoc* tests (not required with 2 groups) to determine which variables discriminated best between specific groups.
- Principle components (axis) correlation (PCC) which uses multiple linear regression (MLR) to 'add' habitat variables into the species 'ordination space'.
- Principle components analysis (PCA) of the habitat data was used to examine the structure of the habitat data, and the relationship between the habitat and groups created from the species data.
- Stepwise DFA indicated which variables best discriminated among the biological groups. The predictive model itself was constructed using DFA.
- Several different methods were used to relate habitat to species pattern. Sixteen habitat variables were examined, however the variable "introduced species" was not considered as a potential predictor variable in the model.
- Because habitat data were in different values and units, the data were standardized before analysis; range standardization was used (the default recommended analysis – Belbin,

2008).

- Step 4: We then decided on the potential list of habitat predictor variables, and ran stepwise DFA (forward and backward) with the groups from Step 1 as the categorical variable. The best set of variables was adjusted using PCC and ANOVA to determine the best set of variables (determined by the model with the lowest error rate using cross validation in DFA).

Assessment

- Step 1: Using the model above we ran DFA with a weighting variable set to 1 for reference sites used in building the model and 0 for the test set. Then the prior probabilities were saved for the test sites.
- Step 2: Created an excel file for each reference group containing the species data and appended the test sites predicted (highest probability) to that group.
- Step 3: We then performed an ordination of reference sites with test sites included (this is usually done one at a time to reduce the effect of the test site on the ordination, and is especially important when few reference sites are available, as in this case study). PCC was used (PATN) to determine which taxa and habitat attributes differentiated the test site from the reference sites (if they were different – i.e., , interpretation of the response).

Sites were then plotted and ellipses built around the reference sites only.

Based on probability ellipses a test site can be assigned to one of four categories:

- Band 1 – inside the 90% ellipse – in reference condition
- Band 2 – between the 90 and 99% ellipse – possibly disturbed

- Band 3 – between the 99 and 99.9% ellipse – disturbed
- Band 4 – outside 99.9% ellipse – very disturbed

2) RCA approach using canonical correspondence analysis (CCA)

Model building

- Step 1: We performed a detrended correspondence analysis (DCA) to test whether a unimodal or linear response model was appropriate (length of axes > 2 and data matrix sparsely populated – ter Braak and Smilaeur, 2002). These analyses were done on all 146 species, including rare ones (partly to circumvent the criticism of Karr and Chu (1999) that ordination does not consider rare species). We used the downweighting option for rare species so that they did not unduly influence analyses; rare species are often outliers, can skew ordination diagrams and obscure patterns. Another way of dealing with rare species is to set criteria for inclusion such as (occurs in a minimum number of sites, 5% or 10%, as in approach 1 above).
- Step 2: Having determined that a unimodal model was suitable, we conducted a canonical correspondence analysis (CCA) on reference sites (wild and organic) - minus invasive species. We tested which habitat features were significantly related to ordination axes using stepwise forward selection in CCA ($P < 0.05$); this was required because of the relatively large number of variables compared to the number of sites. We used Hill's scaling and focused scaling on the inter-sample distances.
- Step 3: We then performed a partial CCA (pCCA) using a stressor (invasive plant species richness) and controlling for significant natural habitat variables as covariates

Assessment

- Step 1: We re-ran the *p*CCA model (Step 3) on the “reference” sites and inserted test sites (relatively impaired sites) as supplementary (passive) in the ordination. We ensured that these were positively associated with stressor gradient.
- Step 2: We then ran the *p*CCA model (above) and included dummy variables for reference sites (wild and organic) as supplementary environmental variables. We calculated the distance from the reference site centroid (s) to the test site to determine the level of impairment. As in the RCA approach (1), the response can be interpreted by examining why the test site differs from the reference condition.

3) Wetland Plant Index (WPI) – based on WQI and WMI

Model building

- Step 1: We first performed a DCA to test whether a unimodal or linear response model was appropriate (length of axes > 2 and data matrix sparsely populated – ter Braak and Smilaeur, 2002)
- Step 2: We then performed a CCA or *p*CCA to determine the main environmental gradients and specifically identified a stressor axis (invasive species richness) and how much the latter contributed to the total variation explained by the environmental variables. We used stepwise forward selection to evaluate the significance of the variables. We used the biplot scaling and focused the scaling on interspecies distances (ter Braak and Smilaeur, 2002)
- Step 3: We then used general formula below as in Lougheed and Chow-Fraser, 2002; Seilheimer and Chow-Fraser, 2006; Croft and Chow-Fraser, 2007):

Equation 1:

$$\text{WFI} = \frac{\sum_{i=1}^{n} Y_i T_i U_i}{\sum_{i=1}^{n} Y_i T_i}$$

Where: Y*i* = if the species is present, this value is 1; if absent, it is 0

T*i* = value from 1-3 or niche breadth of species *i*

U*i* = value from 1-5, tolerance of species *i* to impairment (invasive species in this case)

We deployed the position of species along CCA axis 2 (this was the invasive species stressor gradient) to determine the U value for that species. The U-value is an index of species' tolerances of (or sensitivity to), invasive species; a value of 1 is indicative of highest tolerance, whereas a value if 5 indicates least tolerance. Thus, species given a U-value of 1 were associated with sites having high numbers of invasive species and/or extensive bare ground cover (located at the top of the ordination – see Results), whereas species with high negative values (located to at the bottom of the ordination).

- Step 4: We then determined the positions of species' centroids along the stressor gradient (in this case axis 2 of the CCA) to calculate U-value for that species. Species centroids were first sorted in descending order and then ranked into five U groups (T. Seilheimer pers. comm.) based on the CCA axis 2 range;

>0.75	U =1
0.25 to 0.75	U =2
0.25 to -0.25	U = 3
-0.25 to -0.75	U = 4

<-0.75	U = 5

- Step 5: This step was to derive the T-values for each species and is based on the standard deviation (SD) of the species scores from the CCA output file. As for U-values these standard deviations were sorted in descending order; species with a broad niche (large SD) were assigned a T-value of 1, whereas those with a narrow niche (small SD) were assigned a value of 3. The ranks used were:

0	T = 3
0-0.66	T = 2
>0.67	T = 1

- Step 6: We identified plant species that are indicative of good conditions (U-value of 4 or 5), those of degraded conditions (U value of 1) and those that are tolerant of a wide range of conditions. Part of this step is to provide a discussion of the ecology of different species and biological relevance.
- Step 7: We then used equation 1 to calculate a wetland plant index (WPI). We compared this index with other indices of site quality (e.g., using linear regression).
- Step 8: The WPI index can be validated by: a) examining before and after impact; b) or before and after adaptive management; c) or by comparing historical WPIs to current conditions; d) or by comparing WPI from reference sites with study sites; e) comparing the index with other indices of condition (e.g., Simpson's Index, IBI scores). Compared WPI index between farm types using Generalized Linear Model (PROC GLIMMIX, SAS Institute, 2000).

3 RESULTS

3.1 Comparing evaluations of biotic/habitat quality

Generally, the four reports focused on three aspects of assessment: 1) the selection of indicators for assessing local habitat quality with a view to setting standards and 2) the development of a biotic and habitat integrity index and 3) the use of reference conditions (derived from protected areas) or time zero (using the first assessment as a baseline from which to assess future change). Detailed comparisons of the reports are presented in Table 1.

3.1.1 Woodlands

The report by Summit Environmental Consultants Ltd. (2008) is the most in depth treatment of approaches to assessing local habitat quality and biotic integrity. Although some extremely useful techniques and frameworks are suggested (e.g., a general framework for monitoring and evaluation, numbers of reference sites needed etc.) the final recommendations seem to be largely based on compound indices (MMIs). A major weakness is the extensive focus on richness indices, which are not currently recommended for biodiversity assessments (e.g., Nebbia and Zalba, 2007). Yet elsewhere in the report, the authors acknowledge that evaluations should be based on species composition and abundance data (as in the Alberta Biodiversity Monitoring Program).

Table 1: Detailed assessment and critique of studies

Study	Wetlands in agricultural landscapes	Indicators of riparian diversity in agricultural areas	Biodiversity standards in grasslands	Habitat and biotic indices for woodland
Consultant	• Del Degan, Massé and Associé Inc. (2008)	Iris Environmental Systems Inc. (2008).	AMEC Earth & Environmental (2008)	Summit Environmental Consultants Inc. (2008)
Commonalties	All studies examined 2 levels: • Level 1 – detailed professional assessment • Level 2 – rapid assessment of farm woodland biodiversity monitoring			
Goals	*Primary goals* • To develop a habitat and biotic index for wetlands • To summarize information required for development of ecological integrity assessment methods	• To develop indices of biotic and habitat integrity for riparian areas through review, assessment and synthesis and existing riparian health assessment protocols • To develop one protocol for a monitoring program and one for rapid environmental assessment (REA) by landowner/ farmer	• Objectives are (still) not clear in final report but presumably to identify surrogates and indicators of biotic and habitat integrity in grasslands with a view to developing performance standards for this ecosystem type	• To support development biodiversity monitoring goals that define habitat quality at fine scale (individual farm) • To develop assessment protocols for forest/woodland habitats in agricultural landscapes
Methods – what did they do?	*General* • Examined bogs, fens, swamps, marshes, shallow water • Reviewed relevant information – functions of wetlands, wetland classification • Summarized existing EI assessment protocols for wetlands • Identified factors	• Reviewed and synthesized ecosystem health protocols • Classified survey types: as rapid or inventory application to rivers/streams Species/clades/guilds/groups of interest Method of measurement Associated program Applicability to 7 ecozones in agricultural regions of Canada	*General* • Reviewed NAESI themes, performance standards, relevance to other initiatives • Reviewed background material for biodiversity theme • Provided brief literature review of IBI and IHI in the context of NAESI project • Reviewed surrogates and indicators • Suggested surrogates for grasslands • Provided list of criteria for surrogates (from earlier NAESI reports)	• Reviewed all background for NAESI and biodiversity theme • First reviewed stressors – habitat loss and fragmentation, agricultural inputs and uptake, biotic associations with agricultural inputs, large scale climate change • Reviewed literature related to developing indices of habitat integrity (IHI) and biotic integrity (IBI) – also included monitoring methods and protocols - created access database

Study	Wetlands in agricultural landscapes	Indicators of riparian diversity in agricultural areas	Biodiversity standards in grasslands	Habitat and biotic indices for woodland
	that should be considered • Summarized and selected relevant tools for wetland health assessment • Applied to case study • Made recommendations and identified gaps *Specifics* • First carried out screening and consultation with experts • Then defined and selected indicators, evaluated indices, constructed scoreboard – reviewed EPA WRAP (Ohio, Massachussetts most relevant), EMAN • Then Stakeholders guide to assessing wetland health (EPA Wetland Walk Manual) • Selected indicators using coarse filter from 100 candidate indicators – based on literature survey. Then used fine filter approach based on 10 criteria to select specific indicators (criteria were – accessibility, efficiency, feasibility, measurability, cost, adaptability, sensitivity, rigour, aggregation, uniformity)	Included following land uses: native pasture, tame pasture, irrigated and non-irrigated perennial crops, annual crops, agricultural warehouses, all agricultural use • Conducted literature review • Consulted experts (including industry, academics) to assess utility of indicators and Cows and Fish Program • Case study was Alberta's Cow and Fish Riparian Health Assessment • Only the Rapid Assessment part of the Cows and Fish Program was used as the inventory requires technical expertise and training and there is a disclaimer on the web site (also time constraints)	• Identified potential agricultural stressors • Discussed use of indicator rating methods – similarity/dissimilarity matrices, reference conditions *Specifics* • Provided Steps for ecosystem health assessment, these are: Determine ecosystem goals and associated questions Map project area Choose indicators to be measured Site selection Timing Site assessment Reference conditions Scoring and analysis • Compared results of ecosystem health assessment qualitative and quantitative approaches • Provided details on rangeland evaluation assessment protocols	• Identified, reviewed and synthesized existing ecosystem health protocols • Recommended elements of IHI or IBI for forest/woodland with thresholds or relation to reference condition • Evaluated indicators according to whether they were scientifically defensible, logistically feasible, would be diagnostic, spatially/temporally relevant to farm woodlands, and whether they were communicable to the public/scientists • Completed case study southern Okanagan

Study	Wetlands in agricultural landscapes	Indicators of riparian diversity in agricultural areas	Biodiversity standards in grasslands	Habitat and biotic indices for woodland
Outcome and results	• Selected 10 indicators based on coarse and fine filter approaches outlined: IBI • Plant life – distribution of plant communities, invasive species • Animals – diversity animal species (birds, amphibians), occurrence of species at risk IHI (water) • Physico-chemical properties of water (turbidity, algae cover, water levels) • Hydrology of wetland • Size of wetland IHI (Landscape) • Connectivity (hydro-connectivity of wetland, distance between nearest road and wetland) • Adjacent land use (human activities in buffer zone, size of buffer zone, pressures inside buffer zone)	• During literature review analyzed 76 references – found 620 indicators • After removing duplicates, 372 indicators left remaining • Number of indicators exclusive to rapid assessment and inventory was 304 (total 335) and 104 (total 135) respectively. • Initially contacted 293 expert individuals/organizations – 43 responses; in total contacted 420 individuals/organizations – 54 responded • Of 372 indicators, 243 had been used by at least 1 expert • Summarized indicators for different ecozones (e.g., 113 for Pacific Maritimes, 8 Montane Cordillera, 271 Prairies, 167 Boreal Plains, 5 Boreal Shield, 0 Mixedwood Plains, 47 Atlantic Maritime) • Of indicators assessed 5/5 stream habitat, 12/14 hydrology, 3/3 geomorphology, 22/22 disturbance, 6/6 pollutants, 14/14 land use, 12/12 health trend, 7/7 biodiversity, 20/49 birds, 5/10 herptiles, 8/26 invertebrates, 14/16 wildlife, 14/14 patch characteristics, 11/11 soils, 15/16 ground cover, 92/109 vegetation,	• IBI and IHI require reference conditions but these can be hard to find (most ecosystems have some level of human influence) • Hard to provide a complete list of species and this is subject to observer bias (depends on how skilled observer is – this applies to most/all types of biological monitoring) • Also identifications vary by taxa; taxonomic resolution and skill of observer may affect assessment • Suggested choice of indicators and assessment of ecosystem health are subjective • IBI indices assume that animal and plant populations are proportional to amount of habitat (may not be true in disturbed landscapes where source-sink effects are common) • Measures need to be selected for IBI and IHI that best measure ecosystem health – these include species and habitat characteristics • Suggested approach that is logistically feasible to develop IBI for each natural grassland type within each ecozone is to: ➢ List species present at reference conditions (abundance/evenness) ➢ Relate habitat requirements of these species to the IHI parameters • IBI and IHI should summarize	• Found little agreement in literature about types of indicators to use or on surrogates for level 1 (professional assessment) e.g., surrogate species may not be representative of other taxa, taxonomic solution may be an issue (family or species?). • Recommended using species composition indicators as representative of biotic integrity (IBI) and structure, function as indicators of habitat integrity (IHI) • Also recommended scoring IBI and IHI as deviations from reference conditions (dividing observed difference from mean of conditions by deviation of all reference sites); also Test Site analysis may be useful • Concluded that selection of reference sites is critical and recommend at least 20 sites; if reference sites are not possible then use time zero • Because project is outcome-based, monitoring is essential to measure effectiveness of policies and management actions and to provide data on EI in

Study	Wetlands in agricultural landscapes	Indicators of riparian diversity in agricultural areas	Biodiversity standards in grasslands	Habitat and biotic indices for woodland
	Applied to St. Lawrence Lowlands as case study • Assessed 6 wetlands – had to make some modifications to protocols • Attained first objective – consistency of protocols and relevance of findings • Applied protocols for various parameters • Found indicators performed successfully in terms of assessment criteria • Calculated time required to complete spreadsheet/guide • Identified obstacles and difficulties experienced in adjustments • Found consistent results despite change in observers • Of wetlands examined, in terms of IBI; 5 stable 1 concern; in terms of IHI, 1 Excellent, 2 good, 2 stable, 1 of concern; according to guide 4	13/13 bank stability, 23/25 channel characteristics • Found diminishing return in terms of new indicators with addition of more references or interviews with more experts	condition of ecosystem • By tracking changes over time can use as tool to support management decisions • In summary concluded that IBI is simply using living organisms as measure of biotic integrity of system in order to evaluate consequences human actions • IHIs are actual measurement or estimate of habitat available and sensitivity of habitat to disturbance – e.g., in grassland this is number of features • All IBIs assume plant and animal populations are proportional to quality and quantity of available habitat • Did not recommend diversity indices (Shannon-Weaver, or Simpson) as these contain information on invasive species and do not measure processes/ecosystem function; also changes over time cannot be easily related to stressors/environmental changes (does not help management/restoration) • Noted that indices of habitat integrity can suggest that ecosystem is intact but this may not mean a species is found there	agricultural landscapes – 4-5 monitoring cycles needed to detect trends (6 mentioned in summary) – suggested that this time-lag may be a problem for stakeholders • However, did not recommend using reference condition approach of ABMP because it involved monitoring "indicators" at range of sites: Stated that this was not the most cost efficient method for assessing EI in ag landscapes for NAESI program. For example, the BI cannot be used for individual farms. • Recommend multivariate approach – a variety are superficially reviewed, many are missing • Multiple metrics may be best approach to measure response human-induced stressors and variance around metric

Study	Wetlands in agricultural landscapes	Indicators of riparian diversity in agricultural areas	Biodiversity standards in grasslands	Habitat and biotic indices for woodland
	healthy, 2 in decline; in terms of IEI 4 acceptable, 2 endangered • Some inconsistencies in designation between IBI, IHI and guide.			
	• No conclusions or discussion were presented in the final report – unfinished draft	• Stated that database can be used with specific project goal in mind, level of assessment and land use/ecozone specified • Recommended creating user interface for indicator database • Further analyses are necessary for screening of indicators and identify ones exclusive to each ecozone and assessment level • Refine indicator list to exclude those with no measurement protocols • Provide French translation of database • In collaboration with AAFC and Cows and Fish, modify program protocols so that they are specific to each province • Supplying the database in a public domain portal was also recommended • Solicit data sharing agreement to elicit feedback	• Recommended for surrogates that there is one indicator per agricultural stressor (sensitive to known stressor), 1 per ecotype and 1 per life cycle • For biotic integrity: ➢ Suggested using species list for community – plant, animal, insect and to use plants and animals as surrogates – then to measure occurrence, abundance, richness, evenness for each species ➢ Recommended multiple site visits to increase chance of detection ➢ Species indicator groups – songbirds, waterfowl, raptors, (mammalian) ungulates, (mammalian) carnivores, small mammals, reptiles, amphibians, ➢ Rare species ➢ Invasive species (presence/absence) ➢ Population viability of indicator species ➢ Groups of species/guilds	• Review of sampling designs indicated that random tessellation or stratified systematic were best • Hierarchical Stratification of woodlands necessary (down to agricultural ecumenes) • Use classification of forest types • Recommended hierarchical reporting and monitoring as in ABMP (at level of agricultural ecumenes) • For Level 2 suggested using existing Forest Range Health Assessment from Alberta

Study	Wetlands in agricultural landscapes	Indicators of riparian diversity in agricultural areas	Biodiversity standards in grasslands	Habitat and biotic indices for woodland
		from users in relation to relevance to ecozones • Also suggested creating a parallel system for wetland and lake indicators Knowledge gaps • Province-specific modifications to Cows and Fish methods • Need to correlate indicator with riparian function/benefit • Need to identify habitat-based or biota-based relevance to each indicator	➢ Abundance ➢ Species richness ➢ Evenness ➢ Predation levels • For habitat integrity ➢ Structure: all vegetation layers (tall shrub, medium grass, forbs), site stability, soil moisture regime, soil nutrient regime (nutrient recycling), soil texture, salinity ➢ Function: domestic/native ungulates; human influence (pollutants, habitat fragmentation); fire periodicity, natural or human-caused; decomposition cycle • Recommended that criteria be developed for selection of indicators and that definitive list of indicators be developed for different grassland types • For criteria for selection of indicators recommended: ➢ Defining goal (i.e., , level of biodiversity relative to optimal reference conditions) ➢ Sensitivity of indicator to change ➢ Reliability/accuracy of indicator in measuring ecosystem processes	

Study	Wetlands in agricultural landscapes	Indicators of riparian diversity in agricultural areas	Biodiversity standards in grasslands	Habitat and biotic indices for woodland
			➢ Biological characteristics of indicator (e.g., migrant or resident) ➢ Is species keystone or umbrella? ➢ Species habitat requirements/sensitivities ➢ Can measurements be done using existing programs? (e.g., monitoring programs) ➢ Are they indicators across range of temporal/spatial scales?	
Conclusions and criticism	• Conceptual problem in relation to ecosystem integrity: "An ecosystem has integrity when it is deemed characteristic of natural region in composition, abundance and richness of native species and rates of change and supporting processes" - habitat in most agricultural landscapes will not be considered intact; however, possible need to set biodiversity standards bar higher than possible to achieve • Not clear why non-specialist approach needed (educational, involvement, reducing	• Assessment of indicators from literature review and expert opinion may be of some use and application. • Many of the biophysical indicators suggested for measurement could be useful in multi-species, multivariate modelling techniques • No mention of level 1 professional assessment – Cows and Fish Program is really appropriate as a REA and not inventory and is more related to the landscape function of riparian areas than specifically related to setting standards/targets for biodiversity • Measures of health using Cows and Fish Program are subjective and not based on empirical data (except inventory	• Most of report is a review of existing information – no clear objectives • Some suggested goals may be unrealistic – for example "conserve species composition typical for region" (many/most natural communities in agricultural regions have already been lost) – but may need to set biodiversity standard bar high • Not sure that biodiversity standards can only be implemented by farmers and ranchers – more professional guidance may be needed • General comment: Artificial separation of site level and landscape level factors – studies should be multi-scale • Also priority goal of biodiversity theme is to ensure habitat quantity and quality at multiple scales but for some habitats this does not consider the fact	• General comments – seem to be some conflicting comments/recommendations. E.g., one point recommend multi-species composition and abundance, then for case study use lots of richness/diversity indices • Talks about "assessment protocols" and "monitoring protocols" as if these are interchangeable terms – I think they are quite different • List of indicators (Table) may need modification: e.g., large carnivores poor indicator for isolated and fragmented woodlots in ag. landscapes (most large carnivores are extinct from intensively modified landscapes); large herbivores may be management problem (e.g., White-tailed

Study	Wetlands in agricultural landscapes	Indicators of riparian diversity in agricultural areas	Biodiversity standards in grasslands	Habitat and biotic indices for woodland
	cost?) and/or how this can be used to achieve empirical standards/targets (general criticism of REA in all 4 reports). • Apparent bias in the wetland experts listed/consulted – noticeable absence of wetland scientists who probably have the most knowledge of monitoring and ecological assessment (could include experts from US as well, other wetland experts in EC, aquatic invertebrate experts LRTAP CWS) • Appears to be disproportionate number of experts represented from QC compared to other provinces – this is fine, but perhaps need to balance more nationally? • Also appears to have been little done in the way of a literature review – for example, there are all sorts of peer-reviewed papers	component of Cows and Fish which was not used in this report) – therefore it is only level 2 assessment	that 1) many species use the intervening agricultural matrix (food, shelter etc.) – e.g., carabid beetles use hedges overwinter, fields for foraging; 2) farming practices on the intervening matrix affect species in non-crop habitat (e.g., pesticide spray drift); 3) other habitats are managed for agricultural use (e.g., in this case grassland – pasture, hay, grazing etc.) • May be best to refer to insects as "arthropods" throughout report • Also statement "the Shannon Index takes in account the number of species and the evenness of **the** species" may be better rephrased "evenness of species" otherwise it sounds as if the authors are talking about individual species whereas obviously the SI is based on all species • Recommends that reference conditions should be taken from guidebooks, empirical data, literature or expert opinion – but reference conditions should only really be determined from empirical data • Problems over conflicting recommendations for individual indicators (page 63) could be avoided by choosing multi-species groups for grassland (plants, arthropods, herptiles, birds, small mammals) • Section criticizing IBI/IHI is very weak – observer skills, not being able	Deer) rather than indicators of habitat quality (effects on regeneration). Purple Martin and Eastern Bluebird both poor indicators for woodland. Diversity or species richness generally are not good indicators. • Agree with following: Birds, small/medium-sized mammals (including bats), amphibians, Arthropods (including spiders, ants, moths, also butterflies good volunteer potential), soil micro-arthropods, vascular plants (trees, shrubs, dicots, monocots), bryophytes, lichens, fungi. • Structural features – Canopy composition by tree species and cover, tree size, tree age (old growth listed), snags, snag dbh, special features (e.g., density size cavity trees), downed woody debris, Specific comments • In relation to the multivariate approaches reviewed these are quite narrow and relate to applications rather than the general concepts – ordination (constrained and

Study	Wetlands in agricultural landscapes	Indicators of riparian diversity in agricultural areas	Biodiversity standards in grasslands	Habitat and biotic indices for woodland
	impacts of changes in farming practices on biodiversity in the literature (Milko, 1998 are gray literature and not good references on impact of agriculture on wetlands) • Sentence "Wetlands should be located at least 100 m away from cultivated land in order to insulate them, insofar as possible, from the effects of human activities" should be re-phrased – Land should not be cultivated closer than 100 m of wetlands. This is an arbitrary distance and not based on empirical data.		to record all species, detectability issues apply to many different types of monitoring/ biodiversity assessments (e.g., any data on species distribution and abundance) – there are other, more specific criticisms of multimetric indices that may be more relevant here • Ecosystem health protocol examples in Appendix seem superfluous in report – moreover, many are more applicable to agronomic rangeland health than ecological rangeland habitat quality • Also extensive RIC monitoring protocols (e.g., for waterfowl, raptors etc.) are really outside the scope of report which should be focused on evaluating habitat quality/ecosystem integrity	unconstrained, eigenvalue versus distance based etc.) and cluster analysis or other techniques. So generally this section lacks depth. • In relation to specific indicators listed for case study – richness and compound indices have numerous deficiencies • An important omission is not to include woodland in western provinces. For example, Boreal transition zone in Saskatchewan and Alberta have extensive woodland (agriculture is moving northward); also aspen parkland, natural aspen bluffs. Also authors refer to riparian areas in prairie regions but these are excluded (perhaps because covered in other reports?) • Monitoring protocols listed in Appendix omit some major and well-tested protocols using multi-species groups in US (US National Parks Service and USDA Forest Service – though these are referred to elsewhere and in the references cited) • Indices of richness, diversity and evenness are

Study	Wetlands in agricultural landscapes	Indicators of riparian diversity in agricultural areas	Biodiversity standards in grasslands	Habitat and biotic indices for woodland
				reviewed in Appendix but these are not recommended by current literature and notably lacking are indices using species composition and abundance – • Little mentioned about disturbance patterns – important to include forest patches of different successional stages – this involves active management (landscape planning) • Moreover, conflicting management options for different species at site level can be resolved by examining regional/ larger scale population targets – this highlights a disadvantage of site by site approach • Note Bird Studies Canada is not responsible for Breeding Bird Survey (BBS); this is managed by USGS (Patuxent Wildlife Center) and EC, CWS (National Wildlife Research Centre) • Not sure why winter tracking data for mammals could not be collected at individual farm level??
Application to determining	• Could use wetland classification system to	• Cows and Fish is probably more relevant to other NAESI	• Other than recommendation to use IBI and IHI little substantial relevant	• Useful suggestions for reference sites (20 per ecumene),

Study	Wetlands in agricultural landscapes	Indicators of riparian diversity in agricultural areas	Biodiversity standards in grasslands	Habitat and biotic indices for woodland
local habitat quality and setting standards	stratify sampling design for monitoring program • Discusses thresholds of health for wetland (derived from IBI/IHI)	themes – water, pesticides but not biodiversity • Perhaps application is of greatest use from an agronomic perspective, education and communication with farmers and landowners rather than empirically determining local habitat quality. • No rigorous comparison with reference conditions and threshold levels chosen are probably much too general (healthy, healthy with problems, unhealthy) • However, general assessment of improvement of riparian areas since time zero may be possible	information additional to that in Summit Environmental report	stratified sampling of woodlands • Idea of multi-species groups is mentioned but not pursued sufficiently • Species richness measures finally recommended are suspect and may not be sensitive enough to measure local habitat quality and set standards or deviation from reference conditions

3.1.2 Grasslands

Most of the report on grasslands (AMEC Earth & Environmental, 2008) reviews the general NAESI background and recently completed studies for NAESI. While it reviews the use of biotic and habitat indices of integrity, as well as diversity indices, there is little mention of multivariate techniques other than similarity indices. As in the other reports scale is not considered adequately; models of biotic-habitat relationships need to incorporate local and landscape scale variables and indicators of landscape scale habitat should not be considered separately from the local scale. Moreover, much of the report focuses on protocols for field surveys rather than an analytical framework and study design for bioassessment (see Table 1).

3.1.3 Riparian

The approach used by IRIS Environmental Systems Inc. (2008) was to survey the literature for lists of the different indicators used in riparian habitat, evaluate the frequency with which these were used in riparian areas and then investigate their usefulness by asking experts to rank them. They compiled a database of these indicators for riparian areas within the different ecozones of Canada. While the indicator database may be useful, the emphasis of the report is on the Cows and Fish Health Assessment method which (at least at level 2) appears to be more suited to assessing the utilitarian and functional value of riparian areas, rather than local habitat quality from a biodiversity perspective (though clearly there is overlap in the two goals).

3.1.4 Wetlands

This report focused on wetlands and selected 100 coarse filter indicators which were narrowed down to 10 – among these were plants, animals, physical and chemical properties of water, landscape factors (wetland size, connectivity), and adjacent land use (buffers). These indicators were then tested on six wetlands in the St. Lawrence Lowlands. Some of these indicators may be useful as a component in developing multivariate models.

3.1.5 General comments

The main criticisms of all reports are the following;

1) That the process by which indicators are selected is complex and perhaps unnecessary. For example, despite a carefully thought out protocol for indicator selection, these may be the wrong variables, measured at the wrong scale and they may not respond well to current or future environmental stressors. A better alternative is to use survey data on multi-species groups (which can also be considered indicators). The multi-species approach has been tested extensively by the United States Forest Service, the US National Parks Service and the Alberta Biodiversity Monitoring Institute. The basic concept of such an approach is to survey as many species groups as possible with the fewest sampling protocols and thus for the minimum feasible cost (see Manley et al., 2004).

2) There is considerable emphasis on the use of MMIs for bioassessment. While there are pros and cons of MMIs and multivariate ordination, both can be used within the toolbox of bioassessment techniques (see Cao et al., 2007). There are some disadvantages of MMIs. First, calculation of metrics may be initially cumbersome (though this can be done effectively nowadays using Microsoft Access databases). Second, there may be some redundancy among metrics – though with careful metric selection this can be avoided. The problem of "Eclipsing" (when one component scores offset each other) can be avoided by choosing non-redundant metrics. Third, a new IBI has to be calculated for each new area surveyed, or at least recalibrated for new areas. Fourth, some criticism has been made of IBIs in terms of their insensitivity to change – large changes can occur before they are detected and this could be critical. Fifth, they

may be habitat-specific and do not incorporate spatial or temporal variability (Herbers and Schieck, 2004). Sixth, they have largely been developed for aquatic systems (so could feasibly be used for riparian and wetland ecosystems) but have only recently been adapted for terrestrial ones (grasslands and woodland). One reason for this is the large number of confounding variables in terrestrial ecosystems. Finally, IBIs do not easily incorporate a hypothesis-testing paradigm or information theoretic techniques.

Karr and Chu (1999) criticized multivariate ordination because of the assumption of normality, the exclusion of rare species and the assumption that there is biological significance to the maximum variance in the data (Seilheimer and Chow-Fraser, 2006). However, not all ordination techniques assume normality (nMDS for example, makes no assumptions about the data) and rare species can be included in any ordination technique (though they are often not well described by ordination – or for that matter any other model). In fact, an advantage of ordination is that extremely rare species can be included, though their niche space is not well represented by ordination (Elith and Burgman, 2003). Another advantage of ordination techniques is that they can be used for prediction. Prediction is very important – particularly when the areas to be inventoried are very large and species data are often lacking.

(Note that the following two points were not within the scope of the current contract but are included for information purposes only).

3) There is an almost complete lack of reference in the reports to the possible use of existing data in agricultural landscapes – or the critical need to collect new data. Although the North American Breeding Bird Survey (NABBS) is mentioned, there is little discussion of its usefulness for using birds as an indicator of the condition of agricultural landscapes. In grasslands, extensive data have been collected on birds through the NABBS and Grassland Bird Survey and models developed to

model bird species presence absence and environmental variables (using Generalized Linear Models and an Information Theoretic approach; Franken et al., 2003). In the case of wetlands, although the Marsh Monitoring Program is briefly mentioned, more than 10 years of data are available from this program on marsh bird and amphibian populations – models have already been developed using canonical correspondence analysis (CCA) linking local habitat quality to avian assemblages in wetlands for eastern Canada (Kirk et al., 2001b). Such analyses provide an empirical base from which to develop standards (based on species-habitat models). Other studies do exist but most are in preparation or in review in scientific journals and thus not available.

4) None of the reports refer to two frameworks and protocols being used for aquatic bioassessment in Canada – the CABIN RCA approach and the Wetland Quality Index or Wetland Macrophyte Index approach (Chow-Fraser, 2006). These frameworks can be applied in terrestrial as well as aquatic environments but this is not mentioned.

Finally, 5) None of the reports address the issue of declining biodiversity in agricultural landscapes – and that stemming or reversing declines could be a standard or goal. While cause and effect has not been demonstrated (in Canada) using biodiversity data, retrospective analyses can be done using counts from the NABBS and spatially overlaying these with agricultural statistics data (e.g., Mineau et al., 2005). Trend analyses can also be conducted to determine which species are declining in agricultural regions. One target could be to reverse or stem declines in these species of concern (e.g., Vesper Sparrow, Horned Lark). Another could be based on population targets set by Partners-in-Flight regional conservation plans (these are mentioned in one report). However, empirical models need to be developed to isolate the specific local and landscape environmental factors contributing to these declines, so that best practices can be implemented to alleviate these.

3.2 Evaluations of multimetrics and multivariate ordination

Both MMIs and multivariate methods have strengths and weaknesses and these are highlighted in Table 2. Note that this table was conceived to provide enlightenment on the similarities and differences of the two techniques. Although the table was initiated to

compare and contrast MMIs and multivariate techniques, it was viewed by some as a false dichotomy since in a sense these techniques are achieving the same goal (e.g., scores on ordination axes are essentially a biodiversity index). It was not intended to be dichotomous or to suggest that one approach was optimal. Both approaches could be used in the toolbox of bioassessment, either separately or in conjunction with each other.

Moreover, the column describing multivariate techniques includes methods such as cluster analysis but the real focus is ordination; some reviewers found this confusing. Nevertheless, we include the table for information purposes in this report (Table 2).

3.3 General framework application

First, it should be emphasized that a model-based framework is required using an empirical (data-driven) approach to assess habitat quality, establish reference/benchmark conditions and set restoration targets (i.e., no short cuts, no subjectively-derived indicators or environmental variables). The reason it is required is to stem or reverse declines in biodiversity in agricultural landscapes and to assess efforts to improve habitat quality and enhance species' populations. While the method does not necessarily have to be simple and comprehensible to all laypersons (if so then this could seriously compromise value and validity), nevertheless, it may need interpretation and demonstration (at different levels).

One advantage of ordination techniques is that they reduce complex multi-species multidimensional relationships into a (usually) 2 dimensional graph. This framework can guide restoration efforts by comparing sites impaired by stressors with those in a less disturbed state (reference) within a multivariate framework, such as that used by the Reference Condition Approach. Its strength lies in explaining as much of the natural variability in species composition and abundance at reference sites as possible by relating these to extrinsic environmental factors and then predicting what species assemblages should look like at test sites were stressors mitigated by management. An objective standard to assess habitat quality or health is the magnitude of the deviation of a test site (s) affected by stressors from that expected based on minimally disturbed sites. Specifically, the distance in ordination space that is required for the impacted site (s) to move towards that expected if the sites were minimally affected by stressors can be implemented as a standard. Were sites to be monitored over time, the length of the vector in ordination space can give a measure of the speed at which it is achieving the desired condition. Targets for the time taken to reach the desired state can also be set as a standard (e.g., 5 years for an impacted site to reach a reference condition with adaptive management). This is a powerful and objective tool for setting standards and guidelines.

A general framework application is shown in Figure 1. These are intended as generic and all inclusive steps – specific steps for assessing site quality using 3 multivariate approaches are included in the Methods (case study). Suggested general steps are:

- Assessment - Conduct meta-analysis on pre-existing studies in agricultural landscapes in Canada – birds, invertebrates, plants (stimulate completion of publication and conduct meta-analysis of these studies)
- Collect extensive new data on species abundance and composition in agricultural landscapes (e.g., taxa and known stressor gradients could be chosen; i. e. collection of data can be targeted with specific targeted goals/objectives informed by existing models). For

example, declines in biodiversity have been attributed to the intensity of farming – stressors indicative of farming intensity include fertilizer and pesticide applications, reduced crop diversity (lower heterogeneity in farming systems) and a lower percentage of semi-natural habitat (Billeter et al., 2007). Taxa known to respond to these stressors include birds, carabid beetles (Carabidae), bees (Apoidea), true bugs (Heteroptera), spiders (Araneae), hoverflies (Syrphidae), vascular plants and others. This is needed to: 1) establish baseline/reference conditions (assessment) and 2) to evaluate restoration goals and set standards (restoring sites to baseline/reference condition using adaptive management).

Table 2: Comparison of multimetric and multivariate assessment methods to determine relationships between local habitat quality and biological condition

Multimetric indices	Multivariate techniques
1. GENERAL CONSIDERATIONS • Index of Biotic Integrity (IBI) measures biotic condition using trophic levels, species richness, community composition, disease prevalence, abundance of pollution-tolerant species etc.; Index of Habitat Integrity (IHI) is index of habitat integrity using weighted score, for example, for different stream reaches in environmental measures • IBI has been mostly used in aquatic systems for fish, benthic invertebrates, plants (macrophytes, algae) • Can help identify descriptors of structure and function • Index based on additive metrics (community descriptors) – potential for redundancy in metrics (i.e., #of clingers and # of mayflies – many mayflies are clingers) but a good IBI should avoid this and scenario should be uncommon • Then rank systems from poor to excellent based on variable number of criteria (for rivers or entire watersheds) • Provides intuitive ranking of sites based on health of communities - ranking usually based on quartile distributions; depending on index, quantitative information, including variances can be provided • Can provide quantitative visual (graphical) representation of natural range of variability (e.g., using boxplots) – but not as effectively as multivariate techniques • Many different methods are used and many different metrics are incorporated into the index (modifications in different states in US); If a robust index development process is used, the metrics may depend on the geographical region, the type of water body (i.e., , lake, stream, river, blackwater stream, etc), and the predominant types of stressors in that region. • Enables quick comparisons among sites within calibrated area • May be flexible, but calculation of metrics somewhat cumbersome initially (however, calculation of only a handful of metrics is simple within a database program like Access); regional calibration or completely different sets of metrics should be examined for different regions.	• Many approaches exist including classification or ordination based on species composition and abundance or presence/absence data –e.g., ordination uses species data to investigate environmental gradients either indirectly or directly • Highly complex multidimensional relationships can be shown in simplified ordination space • Provides quantitative visual (graphical) representation of natural range of variability • Wide variety of techniques (linear discriminant analysis – RIVPACS, non-metric multidimensional scaling used in RCA, also weighted averaging such as canonical correspondence analysis CCA, fuzzy set ordination constrained form of nMDS) • Highly flexible • Standardized framework – in the case of ordination can be applied anywhere (but some specific applications, e.g., RIVPACs are very region-specific); certain methods have restrictive data requirements • Can be used in conjunction with many other techniques • Can be used to rank sites in terms of site quality and compare with IBI results • Have been criticized by proponents of multimetric indices for substituting statistical models for expert judgment (but note any model is subject to misapplication/misuse), and the complexity of statistics involved (it has been argued that some of this complexity can be hidden using “canned” software, portals but users need to have an understanding of the programs so that they can recognize errors and make informed decisions during analyses) • Unlike multimetric indices produces a statistical model of relationships between biota and stressor/biophysical factors – does not necessarily claim to be an evaluation of ecosystem health (though CABIN uses RCA to evaluate ecosystem health) – this depends on which multivariate technique is used (e.g., RIVPAC-type models use only physical site characteristics or other features not affected by disturbance to create discriminant models – i.e., does not directly result in any relationships between O/E and stressors. • Have higher data requirements than multimetric indices (gradient analysis) • Widely used in Australia, EU and UK. Less so in Canada (Great Lakes Region,

Table 2: Comparison of multimetric and multivariate assessment methods to determine relationships between local habitat quality and biological condition

Multimetric indices	Multivariate techniques
• For regular use, may be simpler to relay to general public than multivariate techniques (however, this is not necessarily of scientific benefit) • Has been criticized for arbitrary mathematical statistics, unitless scores produced from metric standardization so that they can be combined across metrics, eclipsing (see later), also claims to evaluate "health" (hard to define) – redundancy in some metrics • Useful as investigative tool • Widely used in US for benthic and fish communities, used locally in regional/municipal programs in Canada, Greater Vancouver Regional District, northern BC)	Fraser River, Northern BC)
2. SAMPLING DESIGN • IBIs calibrated by sampling along a gradient of stress (i.e., urban) to determine which metrics best correlated with that gradient then sampling sites were targeted to evaluate the degree of impact (e.g., a specific example is urbanization) urbanization). • (Ideally) requires probabilistic sampling design (as painstakingly set up by EPA for United States). • Some programs are now being implemented using non-probabilistic sampling design (which may or may not be a good idea – depending on the program goals) because it is being argued that local level site requirements are not being met by probabilistically chosen sites. • For example, if goal is assessment of condition over an extensive region then a probabilistic design is necessary, however, if the goal is to locate and mitigate problems, a targeted approach may be more effective. (Goal in this case is to assess condition over an extensive region)	• (Ideally) requires probabilistic sampling so that statistical inference can be made at other sites not surveyed outside study area • Initial development of something like a RIVPACS model actually requires a relatively robust set of reference sites. Assuming that this type of model is already developed then probabilistic sampling is ok.
3. WHAT TYPES OF DATA? • 2 phases – 1) collection of stressor and natural habitat data (though may not be used in B-IBI) and 2) collection of biological taxa data • In some IBIs previous development of set of regionally-defined B-IBI scores and development of e.g., 5 levels of stream condition (excellent, good, fair, poor, very poor) is carried out – however development of these	• Species composition and abundance data and biophysical/anthropogenic stressor data can be used for wide gamut of models as well as ordination (e.g., Generalized Linear Models, GLMs for single species, Neural Networks etc.); these are the same data requirements as IBI. • Taxon data can be abundance values, ranks (e.g., Braun-Blanquet scale or presence/absence)

Table 2: Comparison of multimetric and multivariate assessment methods to determine relationships between local habitat quality and biological condition

Multimetric indices	Multivariate techniques
levels of condition may not be necessary to relate an IBI to habitat quality	• Predictor/explanatory variables can be measured on different scales, quantitative or qualitative (e.g., ranked, binary); some transformation of variables may be necessary (though in CCA universal transformation of variables, e.g., using log-transformations is not advised)
4. HOW ARE SPECIES AND ENVIRONMENTAL DATA HANDLED? • IBI and Index of Habitat Integrity (IHI) must be calculated separately and interpreted *a posteriori* • Environmental data are not used to develop relationships between indices and habitat; rather it is assumed that IBIs should be similar within a region regardless of the variation of the natural physical site characteristics not related to disturbance features.	• In constrained ordination, explanatory variables are incorporated directly into ordination (e.g., in CCA, ordination axes are linear combinations of environmental variables; in Fuzzy Set Ordination FSO hypotheses about species distributions are developed *a priori* and axes are tested for relationship with explanatory variables - FSO is constrained form of Bray Curtis Ordination); also Distance-based Redundancy Analysis (db-RDA assumes linear model) • Unconstrained ordination techniques can be used for species and habitat data separately – matrices can be compared using Mantel tests • Typically, different weighted-averaging ordination techniques would be used for species compared to environmental data – species data tend to have many zeros, low abundance scores (unimodal model such as CA, DCA, also nMDS), whereas habitat data are more complete/full matrix – linear model such as Principal Components Analysis (PCA) are appropriate • Note that using PCA on habitat variables assumes one is not as interested in individual variables (but this may not be the case)

Table 2: Comparison of multimetric and multivariate assessment methods to determine relationships between local habitat quality and biological condition

Multimetric indices	Multivariate techniques
5. HOW ARE REFERENCE/BASELINE CONDITIONS CHOSEN? • Reference sites identified based on absence of exposure to stressors can indicate reference conditions • For biological metrics may need to know relationship between environmental condition (stressor) and biological response to choose reference condition • Reference or baseline conditions chosen according to criteria (e.g., for aquatic systems – total N, total P, chloride, mean RBP score) • Problem with IBI is that two reference sites may have very different scores/ranking due to natural variability (e.g., hydrological regime) – even two physically similar reference sites may have different scores because of natural temporal or spatial variability	• Reference or baseline conditions can be determined from position in ordination space – environmental envelopes can be drawn around sites using probability ellipses and/or using centroid of cluster of sites as desired condition – assumes abiotic variables included so that good and poor conditions can be determined with ordination • Can use linear discriminant analysis (canonical variates analysis CVA) to identify clusters and the environmental variables which characterize clusters (are significant) • Can also incorporate other types of analyses into ordination –could use cluster analysis, TWINSPAN, discriminant analysis or regression trees on biotic data to identify reference, intermediate and impaired conditions. These sites are then identified in the ordination diagram and the centroid in ordination space represents the condition for reference, intermediate and impaired. This approach is using biotic communities to inform observer about condition of sites – therefore it is objective. • In eigenanalysis ordination can separate out contribution of abiotic factors as well as temporal and spatial variability using variance partitioning

Table 2: Comparison of multimetric and multivariate assessment methods to determine relationships between local habitat quality and biological condition

Multimetric indices	Multivariate techniques
6. WHAT TYPES OF ECOSYSTEMS? • IBIs have been mostly tested and applied in aquatic environments (but applicable in variety of situations including human health and terrestrial ecology) – more recently attempts have been made in terrestrial ecosystems or mixed terrestrial and aquatic ecosystems, but not well developed or tested. • For many multimetric indices, application beyond the region in which they were developed has been considered difficult. Applicability of an IBI approach is to developing an indicator of condition – not using actual indicator itself in a new type of system (assumed that actual measures of interest depend on type of system of interest) • IBI began in aquatic systems – this should not mean that it is limited to them. However may be inappropriate for terrestrial ones (trophic patterns may not differ greatly among disturbance types plus health of individual animals hard to assess) • Some recent IBIs for birds have been developed and appear successful in woodland and grassland ecosystems	• Ordination techniques have been widely applied in aquatic and terrestrial ecosystems – framework is applicable to all ecosystems where data on species and sites are available • Well tested in both aquatic and terrestrial systems
• IS THE FRAMEWORK/TECHNIQUE APPLICABLE IN DIFFERENT GEOGRAPHIC REGIONS? • New IBI must be developed for each new geographical area or at least an existing IBI must be evaluated and possibly recalibrated for each new area • Also possible to use historical data with an existing IBI as long as the geographical area is the same or at least an existing IBI must be evaluated and possibly recalibrated for each new area (depending on quality of the older data). If they are much better conditions, the existing IBI would not adequately describe conditions.	• New sites can be entered into initial (constrained ordination) passively to see where they lie – for example, species data from a different geographical area can be included passively in CCA to determine where they are positioned in relation to other sites • However, would need to consider how data from new area (with new/different natural characteristics) fits in with original data • Can also enter historical data passively in ordination

Table 2: Comparison of multimetric and multivariate assessment methods to determine relationships between local habitat quality and biological condition

Multimetric indices	**Multivariate techniques**
8. TYPES OF METHODS • Many different methods available (Karr’s Index of Biotic Integrity not really the most widely accepted anymore). Old approach of using professional judgment to choose metrics and set scoring has generally fallen out of favour. • No one way to develop an index, but it is much more common to develop an index empirically and to use more continuous scoring approaches (rather than Karr's 1-3-5 approach). EPA pushes this more empirical approach based on actual data in a few different documents (used this approach in its Wadeable Streams Assessment in 2004).	• Somewhat bewildering variety of methods requires level of understanding – but richness of analytical techniques allows models to be compared and contrasted – to obtain optimal or best solution • Appropriate approach (for either multimetric or multivariate may depend on type and amount of existing data – so only a few approaches end up being relevant/applicable. • distance-based methods (nMDS, MDS, FSO) vs weighted averaging (eigenanalysis methods, CA, DCA, CCA, RDA) • unconstrained (CA, DCA, nMDS) vs constrained (CCA, RDA, FSO); • linear (PCA, RDA) vs non-linear, unimodal (DCA, CCA) • Main difference in techniques is in how to evaluate/best represent the distance between sites/species in ordination space – distance-based techniques algorithm may more truly represent ecological distance compared to eigenanalysis ordination (less distortion, and can portray discontinuities in data) but with large datasets still can be analytically slow • Two main schools of ordination approaches – weighted averaging vs distance-based. Both require large numbers of decisions by investigator but as a generalization weighted-averaging has been more widely adopted because of easy to use, “canned” software • Ordination method to use can be chosen objectively – e.g., Shepard Diagram to test sample separation in ordination space from ecological dissimilarity • Can use unconstrained ordination to *a posteriori* interpret gradients (relate species gradients to environmental variables) and substantiate results from constrained ordination (e.g., test that important variables not missed, optimal model produced, no model error/distortion) – this can be done for weighted averaging techniques: detrended correspondence analysis DCA vs CCA or non-metric dimensional scaling (nMDS) vs fuzzy set ordination (constrained form of nMDS);
9. COMPARABILITY OF ASSESSMENT • May be challenges in comparing IBIs with other models but have been used to compare with ordination methods (site ranking)	• Models can be compared and contrasted within wide framework of ordination techniques using model fit parameters (e.g., variance explained by environmental variables in eigenanalysis ordination compared to unconstrained model or stress in distance-based ordination)
10. SENSITIVITY OF METHOD	• With regular monitoring ordination techniques are very sensitive to modeling

Table 2: Comparison of multimetric and multivariate assessment methods to determine relationships between local habitat quality and biological condition

Multimetric indices	Multivariate techniques
• IBI may not provide early enough warning of environmental change – tend to require large changes to occur before substantial change in IBI occurs. • Indices may oversimplify effects on communities (especially when organisms grouped into higher taxonomic categories)	change in site habitat quality • With ordination techniques determining what a change means may be difficult – a shift in community composition along a gradient of increasing or decreasing stressor value could be measured as site movements in ordination space over time • In RIVPACS-type model it is clear what a drop in O/E means.
11. FLEXIBILITY AND APPLICATIONS • Combining species data into guilds or groups (e.g., tolerant species etc) could obscure relationships - can also illuminate relationships that are not obvious at species level because they affect an entire group of organisms • However original species and biophysical data can be used for other analyses (e.g., multivariate ordination)	• Species composition and abundance data can be used for many other applications • There is a critical need for flexibility between multi-species and single species models (both coarse filter and fine filter approach is needed for biodiversity conservation) • Single species management is still important (e.g., species at risk legislation) - – e.g., Generalized Linear Models (GLMs) of Generalized Additive Models (GAMs) models of single species. Single species may be indicators for part of communities and specific environmental conditions. • Also GLM and GAMs single species models are very powerful, can be used for prediction and within model selection framework • Also potential value in aggregating across similar groups for ordination (i.e., , ordinating by trophic groups has been examined for diatom data).
12. IS METHOD OBJECTIVE? • Subjectivity of metrics selection and some metric definitions is a concern • Current most common approach to selecting metrics involves identifying metrics that respond to specific disturbance gradients. Selecting metrics based purely on judgment has been a problem in the past but is not current practice. However, the set of candidate metrics examined is largely based on judgment.	• Higher degree of objectivity in ordination framework • But have to make many decisions during analyses - however, these are based on experience/expert knowledge, and can be substantiated using different techniques/tests - could be argued that this is not any different from multimetric approaches where “expert judgement” is used for choosing cutoffs or selecting metrics based on a series of metric evaluations.
13. HOW TRANSPARENT ARE THE ASSESSMENTS? • Changes in species composition/group membership are “hidden” in index making it ambiguous • Could suffer from “Eclipsing” - component scores offset each other when combined; however, if not using redundant metrics, not a real problem. Each metric is intended to measure some different aspect of the assemblage, so some aspects will look better than others.	• Changes in species composition and abundance are directly interpretable in ordination diagram (changes over time or comparisons between paired low quality and reference/baseline sites) • Could be argued that interpretation for the lay person and for some managers can be less than obvious with ordination (therefore, need ordination experts to interpret). Site shifts within ordination may not be obvious to those without intimate knowledge of the data. However, RIVPACS models are fairly easy to interpret because lower O/E score has a very specific meaning, and the actual taxa

Table 2: Comparison of multimetric and multivariate assessment methods to determine relationships between local habitat quality and biological condition

Multimetric indices	Multivariate techniques
• Multimetric index can lose information – but single metrics can also be used independently (related more to data analysis than data type) and can inform the user in identifying potential stressors	lost can be directly assessed. • Graphical techniques allow site by site summary of species composition and abundance (diversity indices etc.) • Techniques are transparent provided one can understand the methods and graphs
14. STATISTICAL INDEPENDENCE OF VARIABLES • Large number of confounding variables (e.g., in terrestrial ecosystems) • If two metrics are perfectly correlated, both are not needed • However, if completely uncorrelated may cause "eclipsing" – zero net change in index because some metrics increasing, other decreasing. Scoring of "negative" metrics is usually adjusted so that increasing scores are associated with lower values, whereas increasing scores of "positive" metrics are associated with higher values. In this way, eclipsing should not occur. • Often redundancy in metrics used to develop the index – likely to be relationships among metrics in an index; however, idea is to ensure that each metric provides new information (i.e., is not too highly correlated with another metric	• Fuzzy Set Ordination (FSO) is constrained form of nMDS; Can use vector analysis – e.g., Principal axis correlation (PCC) can be used to relate habitat variables to nMDS ordination axes (uses multiple regression) • Constrained ordination methods like CCA or RDA can handle large numbers of explanatory variables (chemical and physical stressors), are robust to collinearity (highly correlated variables), and include methods for reducing number of variables – analogous process in nMDS is to use PCC and stepwise linear discriminant analysis; provided enough appropriately replicated data are collected • Completely collinear variables are dropped in CCA and RDA (Variance Inflation Factor statistic provides information on variable collinearity).
• DOES THE METHOD CONSIDER SPATIAL AUTOCORRELATION? • No real way of incorporating spatial autocorrelation (sites close to each other are more similar to each other in species composition/environmental variables than sites far apart) into IBI or examining effects of different scales i.e., IBIs developed at site/local level are not applicable to regions or landscapes - IBIs should not even be developed based on local scale (K. Blocksom, pers. comm..) • However, tests can be done on raw data prior to development of index • Scores from IBI can be normalized (e.g., if relationship with watershed demonstrated)	• Can test for spatial autocorrelation in species and environmental data matrices using spatial eigenvector maps, Mantel tests for multispecies data; single species can use spatial lag distances; Dutilleul's test can be used to correct for SA. • Can incorporate spatial autocorrelation into models using trend surface variables (factor out as covariates or include in model if these are significant)
16. HOW DOES THE METHOD FIT INTO A MODELLING PARADIGM (HYPOTHESIS TESTING OR MODEL SELECTION APPROACH)? • Statistical modeling paradigm and testing (application of model) are disconnected - does not fit into traditional hypothesis testing paradigm or	• Can be used in model selection framework (determining best model fit to data) • Hypotheses can be developed *a priori* and then tested (e.g., FSO)

Table 2: Comparison of multimetric and multivariate assessment methods to determine relationships between local habitat quality and biological condition

Multimetric indices	**Multivariate techniques**
more model selection approach (choosing model that best fits the data) • Requires *a priori* identification of specific anthropogenic impacts and effects on biological communities	
17. CAN ANALYSIS BE INCORPORATED INTO GEOGRAPHIC INFORMATION SYSTEMS (GIS)? • No mechanism to incorporate into GIS • Scale effects cannot be incorporated – need to develop different set of metrics for different scales	• Recent developments have been made to incorporate ordination into GIS – this means that the data can used for prediction and also incorporated into conservation planning algorithms at different scales • Can investigate species/environmental matrix over range of spatial scales
18. HOW IS THE ASSESSMENT VALIDATED? ARE DIAGNOSTICS PRODUCED? • Observed effect (metric value) = diagnostic of exposure (human disturbance) • IBI scores can be correlated with disturbance factors (e.g., simple correlations of fish or invertebrate indices with stressors) • Can develop biota-habitat-stressor relationships with "training" set of data, then compare observed metric values with those predicted from the relationships	• Model diagnostics are produced with ordination models (how well model performs, total variation in species matrix explained by model) • Model can be validated statistically; e.g., in CCA the relationship between species matrix and environmental variables can be tested for significance; similarly in NMDS vector analysis or linear discriminant analysis can be used to test significance of environmental variables on axes or can use FSO to do this directly • In constrained methods (e.g., CCA) relative importance of different environmental/explanatory variables can be determined using forward selection (though this suffers from same criticisms of stepwise selection in multiple regression models) • Can use cross validation (e.g., leave one out) – as in RCA
19. HOW ARE CHANGING SITE CONDITIONS MONITORED? • By changes in rank of sites according to IBI scores	• Sampling design can be set up so that sites with similar habitat or in the same geographical area (watershed) but with different anthropogenic stressors are paired (the distance between these pairs in ordination space provides a measure of desired change – see below) • Vector distance (trajectory analysis forthcoming) can be measured between sites in ordination space to indicate change in communities – desired/target change is in direction of unimpaired site (s)
20. CAN METHOD BE USED TO DIFFERENTIATE NATURAL DISTURBANCE FROM ANTHROPOGENIC FACTORS? • Differentiating anthropogenic stressors from natural disturbances is a major concern in IBIs – this can give false/misleading IBI indices • **Can factor out effects of e.g., elevation or watershed which are**	• Differentiating natural disturbance from anthropogenic disturbance patterns is critically important (e.g., in Great Lakes natural disturbance can have greater effect than area of developed land in catchment) • Ordination can include other variables such as climate which influence species distribution and abundance

Table 2: Comparison of multimetric and multivariate assessment methods to determine relationships between local habitat quality and biological condition

Multimetric indices	**Multivariate techniques**
correlated with e.g., fish IBI	• In CCA can use variance partitioning to determine how much variation in species data is due to natural disturbance factors compared to anthropogenic stressors
21. DOES METHOD DIRECTLY MODEL STRESSOR AND RESPONSE? • No – statistical relationship between environmental condition (stressor) and biological response can be established *a priori* or *a posteriori* (using correlation analysis – metrics vs stressors – can be done on metrics already corrected for natural variation too)	• Ordination is based on the premise that species are responding to environmental condition/gradient • This relationship is explored in ordination diagram • Because CCA is a combination of ordination and multiple regression can control for variables not of interest as covariates, while examining effects of stressor variables
22. HOW DOES METHOD INCORPORATE VEGETATION SUCCESSION/DISTURBANCE? • Need to develop different metrics for early successional compared to later successional vegetation types	• Same ordination framework used for early succession and late succession - data from early succession to later seral stages can be incorporated in same ordination model (gradient)
23. HOW ARE ADAPTIVE MANAGEMENT ACTIVITIES TESTED FOR EFFECTIVENESS? • Cannot incorporate management as a vector but IBI score can be correlated with intensity/level of management	• Can readily incorporate management experiments and natural disturbance as vectors – can test trajectory and direction of change on site by site basis • Directionality of variables can be hard to determine in some cases
24. RECOMMENDATIONS • Could use combination of MM and MV assessments • No one approach will give all of the information obtainable with a combination of both	

- For bioassessment monitoring purposes, a statistical sampling design (e.g., stratified random) to select range of sites along environmental gradients within each ecosystem type; could use paired design of high quality/impaired sites in same geographical area (i.e., , reference vs impaired sites)
- Taxa groups could include as many of the following as possible: mammals, birds, reptiles, amphibians, fish, arthropods (terrestrial invertebrates such as Carabid beetles, spiders, hoverflies, bees, aquatic macroinvertebrates), diatoms, vascular plants, fungi, lichens, bryophytes. Use minimum number of sampling protocols for maximum number of species (Manley et al., 2004); collect abundance and composition data rather than simple presence/absence
- Simultaneously collect (multi-scale) biophysical data specific to each ecosystem type at multiple scales (though analysis is based on site level response)
- (tentatively –e.g., based on recommendations in NAESI reports - these could include);
- Woodland – woodland patch size, tree species composition, canopy height, tree species dbh and density, structural characteristics, snags, dead downed woody material, herbaceous layer
- Grassland – patch size, edge vegetation, soil type, hydrology, slope, sward composition and height
- Riparian – open water area, width of edge vegetation, plant species composition, cover and height of adjacent vegetation (e.g., buffer zone), stream flow etc.
- Wetland – area of open water, water depth, water chemistry (calcium, magnesium, pH), turbidity, dissolved organic carbon
- Collect quantitative information on all potential human activities affecting biodiversity

(e.g., for grasslands - grazing, tillage, pesticide applications, timing of harvest – AMEC Earth & Environmental, 2008)

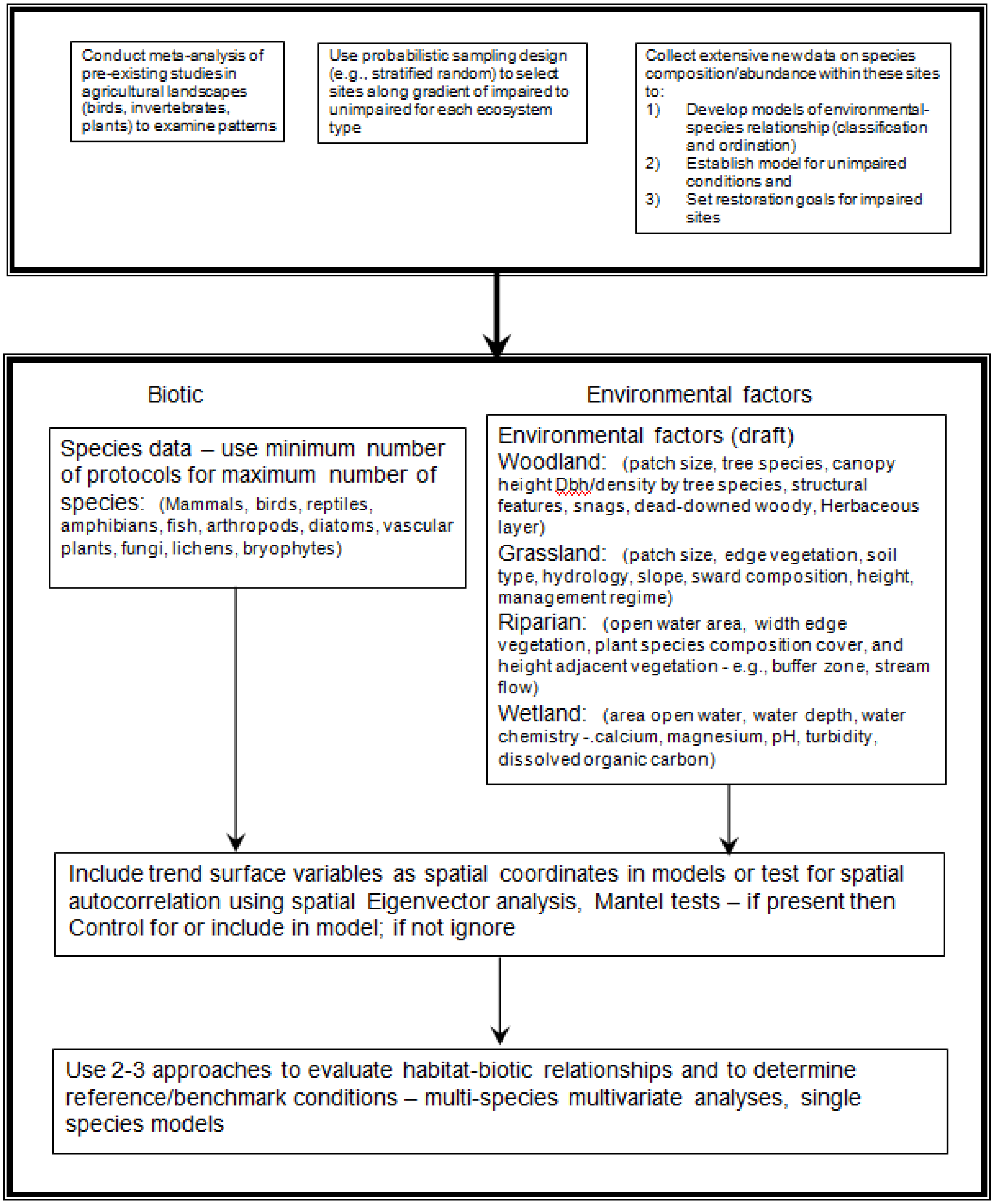
- Use 3-4 approaches to determine reference conditions/baseline
- Multi-species multivariate ordination (see later)
- Single-species models (GLM, regression trees)
- Simulation/scenario modelling
- *For ordination approach* – determine whether to best use distance-based methods (e.g., non-metric multidimensional scaling nMDS, fuzzy set ordination FSO) and/or weighted averaging (e.g., canonical correspondence analysis CCA) or unconstrained/constrained (e.g., detrended correspondence analysis DCA versus CCA) or linear/nonlinear
- Use series of different analytical techniques to provide confirmation of patterns (e.g., use nMDS, FSO and CCA simultaneously)
- From ordination plots determine main environmental gradients (e.g., agricultural stressors such as phosphorus concentration in water) on ordination axes
- Use best estimate of position of species and sites in ordination space from nMDS; use vector analysis (King et al., 2004) to determine the main environmental gradients. And/or use FSO – constrained form of Bray Curtis Ordination BCO. Simultaneously use CCA to confirm patterns (species ordination axes are constrained by linear combinations of environmental variables; do variance partitioning to determine unique contribution of different groups of extrinsic variables; for example, to differentiate human activity from "natural" changes). Note that although the bioassessment is at the site level, it is important to include extrinsic environmental variables at multiple scales (e.g., landscape and local) as these can all have impact on species composition and abundance.

- Determine effect of spatial autocorrelation – there are two schools of thought here: 1) that spatial autocorrelation is a natural component of natural variation and should be included in models, or 2) that it is a hindrance to understanding natural communities and its effects should be tested for and controlled for, or removed, in statistical modelling (see Legendre, 1993). Thus (following 1), spatial coordinates can be used within models to increase the amount of natural variability explained (e.g., in the RCA approach case study, latitude and longitude were included as variables) or spatial autocorrelation can be tested for using eigenvector maps in R- (Griffith and Peres-Neto, 2006) or using Mantel tests (Mantel r correlation between data matrices); if significant then include in models, if not then ignore
- Use cluster analysis (or TWINSPAN), discriminant analysis (canonical variates analysis) or regression trees (CART, boosted regression trees) to classify sites (latter would also provide information on environmental variables characterizing sites and used for classification). Highlight impaired, intermediate and reference sites in ordination space (use different symbols)
- Use vector or trajectory analysis (latter being developed by Peter Minchin, University of Illinois) to determine change in position of sites following restoration efforts (adaptive management). These tests determine whether communities are changing in specified manner; e.g., goal could be to shift sites along a specified environmental gradient, a vector of maximum correlation between ordination scores and a variable that expresses position on the gradient would be a suitable target direction. Can re-project data from sites in ordination space over different points in time; current conditions compare distances between sites in ordination space (pairs of impaired versus reference conditions). Also can use ordination for large number of sites (e.g., “sustainable” or reference/baseline sites) and

insert new sites that need to be assessed passively (as in CCA, RCA case study)

- Indicators should be developed from empirical data and models, not selected subjectively *a priori* (as in all 4 NAESI reports); there are several types of software available to do this most notably "indicator species analysis" Dufrêne and Legendre (1997) and also TWINSPAN (two-way-indicator-species-analysis). Indicators can also be derived using multivariate ordination – species located close to environmental vectors can be used (Kremen, 1992) and modeled using single species models (see below). Species known to respond to specific stressors (and perhaps not included in multivariate models) could be included in single species models as well.
- *For single-species models (e.g., in RCA this is part of the scenario building)*; Based on multi-species patterns in ordination conduct single species models using GLM (negative-binomial or zero-inflated errors) – could use approach of Nielsen et al. (2007) to identify reference/baseline conditions.
- *For simulation models*: Need to incorporate concepts of vegetation dynamics/succession in modelling approach (e.g., fire models, hydrological models SWAT)
- Final step is to evaluate uncertainty in models and test/validate models

Figure 1: Framework for evaluating local site quality in agro-ecosystems using empirical data



3.4 Example of multivariate techniques – application to case study

For the SCB plant study, vegetation surveys were conducted around the margins of wetlands between the last week of June and first week of August. We defined the margin as all native vegetation that occurred around ponds within the cropland matrix. This varied from a ring of willow or aspen around the edge of the basin to areas of native upland vegetation that included grasses, forbs and shrubby species. For wild sites, the demarcation between wetland margin and “upland” (field) habitats was the point at which grassland vegetation or planted duck nesting cover became relatively homogeneous.

Surveys were timed to occur at least two weeks after herbicide use by conventional and minimum tillage farmers (Leeson et al., 2000) because herbicide applications can have an important effect on plant species cover and composition (Moreby and Southway, 1999). Five transects were randomly located at each pond edge. Along each of these transects, we positioned three 1 m^2 quadrats, one in the wetland (1 m from the water), the second at a midpoint (midway between the water and the field edge) and the third in the field (1m into the margin from the edge of the field). Within each quadrat, we recorded presence and cover of all submergent and emergent plant species. We ranked vegetation cover using the domin scale (Shimwell, 1971).

Because cover of dicotyledonous and monocotyledonous plants was assessed differently (dicotyledonous by stem counts and monocotyledonous by % cover), and thus not strictly comparable, for statistical analyses we used the % occurrence of each plant species in the five quadrats at each of the three moisture locations. For the purposes of the case study we used data from the field sites only for several reasons: 1) because invasive species tend to be associated with crops it would be most likely that they would seed themselves into the adjacent few metres into the pond margin; supporting this contention, Leeson et al. (2000) found that the highest number of weed (invasive) species occurred at the edge between the crop and pond margin; 2) the number of invasive species would be low near the water's edge because only species adapted to very wet conditions would survive there; 3) at the midpoint there would likely be fewer introduced species because of competition from native species and woody species (D. Forsyth, pers. comm.).

The original goal of the SCB plant study was to examine the effect of management and moisture regimes on plant species frequency of occurrence and not specifically on evaluating site quality. To account for relationships at multiple scales, we included not only site level environmental factors, but also some at an intermediate scale (the entire wetland) and also at the quarter section level (approximately 64 ha) [4]. At the site level, we measured biophysical factors that could affect plant biodiversity, including aspect, slope, percentage of bare ground and dead vegetation in each quadrat. However, only the percentage of bare ground and dead vegetation could be used in analyses, because averaging aspect and slope from quadrat level information was not meaningful. We included two variables at an intermediate scale (% shoreline treed and % shoreline willow) from Shutler et al.'s (2000) study; variables at the landscape scale were soil types derived from soil classification maps, precipitation and area of wetlands. Soil types and precipitation levels were determined by overlaying the latitude and longitude centroids of each pond on soil classification maps to derive soil classes, and on maps of precipitation levels. We also included corrected latitude and longitude as variables to incorporate measures of space. We did not examine the effect of specific farming practices like herbicide or fertilizer use in adjacent fields, although these can affect field boundary vegetation through spray drift (Kleijn and van der Voort, 1997; Kleijn and Spoejing, 1997; Kleijn and Verbeek, 2000).

[4] Note that although we assessed the importance of variables at different scales we did not carry out an in depth assessment of scale effects as this would require different types of analyses (e.g., variance partitioning in CCA or GLM for single species) to see which scale contributed most to variation in species composition and abundance.

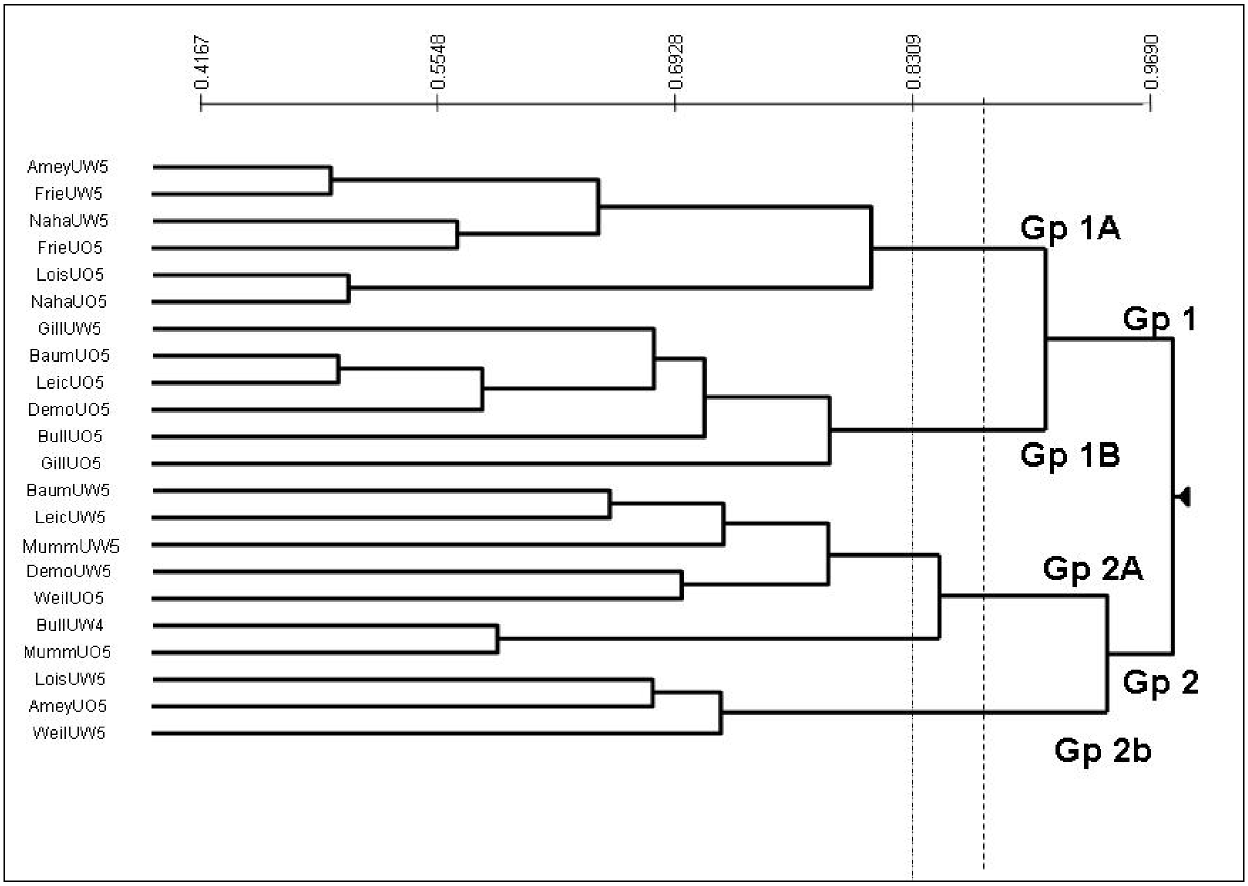
1) CABIN RCA approach (cluster analysis and nMDS)

The dendrogram for the 22 "reference" sites shows the results of the classification analysis (Step 1; Figure 2). Although 22 sites are shown, we discuss only the solutions for four groups. Two groups of 12 sites (Group 1) and 10 sites (Group 2) were created at the first division. Group 2 was then divided into three sites (Group 2B) and seven sites (Group 2A); Group 1 was further split into two groups of six sites. For the purposes of assessment a minimum group size of 10 sites has been recommended (Reynoldson and Wright, 2000), so to build models we have not gone further than characterizing these two groups (Table 3).

Table 3: Biological characteristics of two groups (see Table 11 for full species names).

Characteristic	Group 1	Group 2
Mean no taxa (SD)	8.3 (4.3)	10.5 (3.9)
Top 5 Species characterizing groups	*Cirsium arvense*	*Cirsium arvense*
	Agropyron trachycaulum	*Aster hesperius*
	Vicia americana	*Aster falcatus*
	Aster hesperius	*Potentilla anserine*
	Salix lucida	*Achillea millefolium*
No. of species accounting for 90% similarity within group	7	13
Taxa contributing > 4% to differences between groups	Average % occurrence per plot	
Cirsium arvense	55.0	32.0
Agropyron trachycaulum	35.0	5.0
Aster falcatus	0	25.0
Aster hesperius	15.0	27.0
Populus tremuloides	13.3	15.5
Potentilla anserine	3.3	20.5

Figure 2: UPGMA dendrogram of 22 "reference sites'.



In general Group 1 was less diverse, and plot coverage tends to be dominated by two taxa, *Cirsium arvense* and *Agropyron trachycaulum*. Group 2 tends to be more diverse and although *Cirsium arvense* is also the dominant taxa, its coverage is lower and other species contribute to a greater degree (*Aster hesperius, Aster falcatus, Potentilla anserine*). Two species were absent or occurred rarely in group 1 (*Aster falcatus* and *Potentilla anserine*). Of the 42 species accounting for 90% of the difference between the two groups, 12 are unique to Group 2 and ten to Group 1, with 20 species common to both groups (Table 3, Figure 2). Figure 3 shows the sites plotted in nMDS ordination space (PRIMER) with the four groups and demonstrates good discrimination between Groups 1 and 2.

Figure 3: Two dimensional nMDS ordination of 86 plant taxa from 22 reference sites, showing four groups identified from cluster analysis

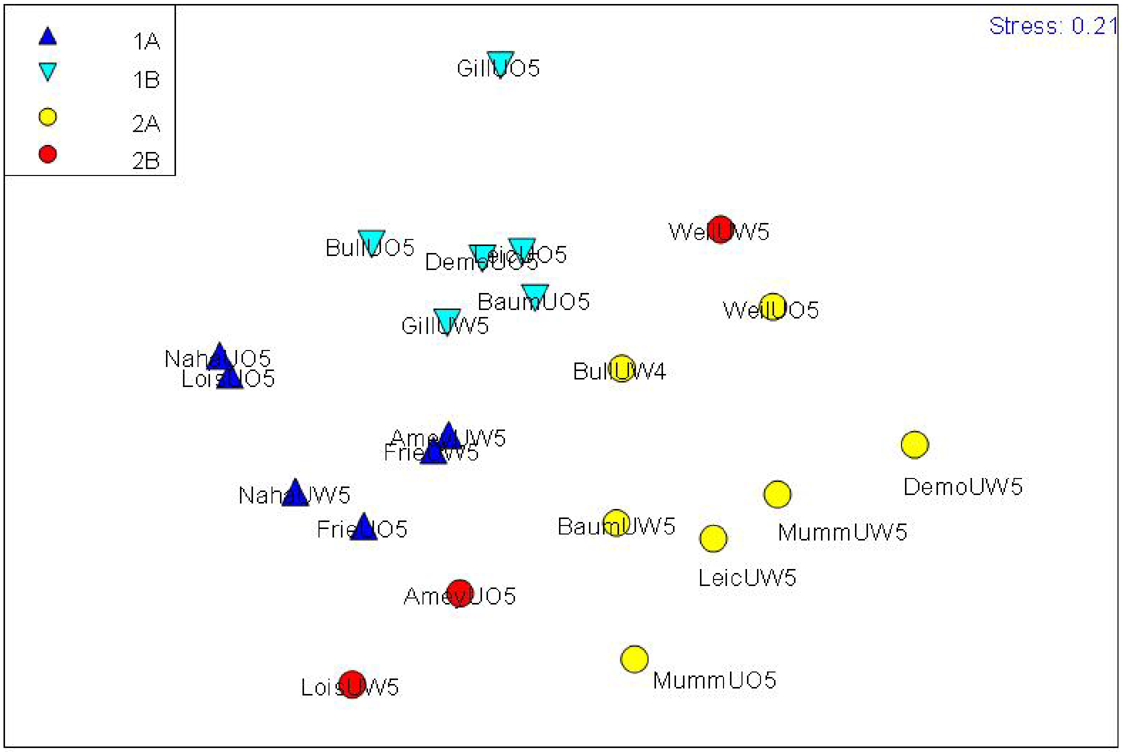
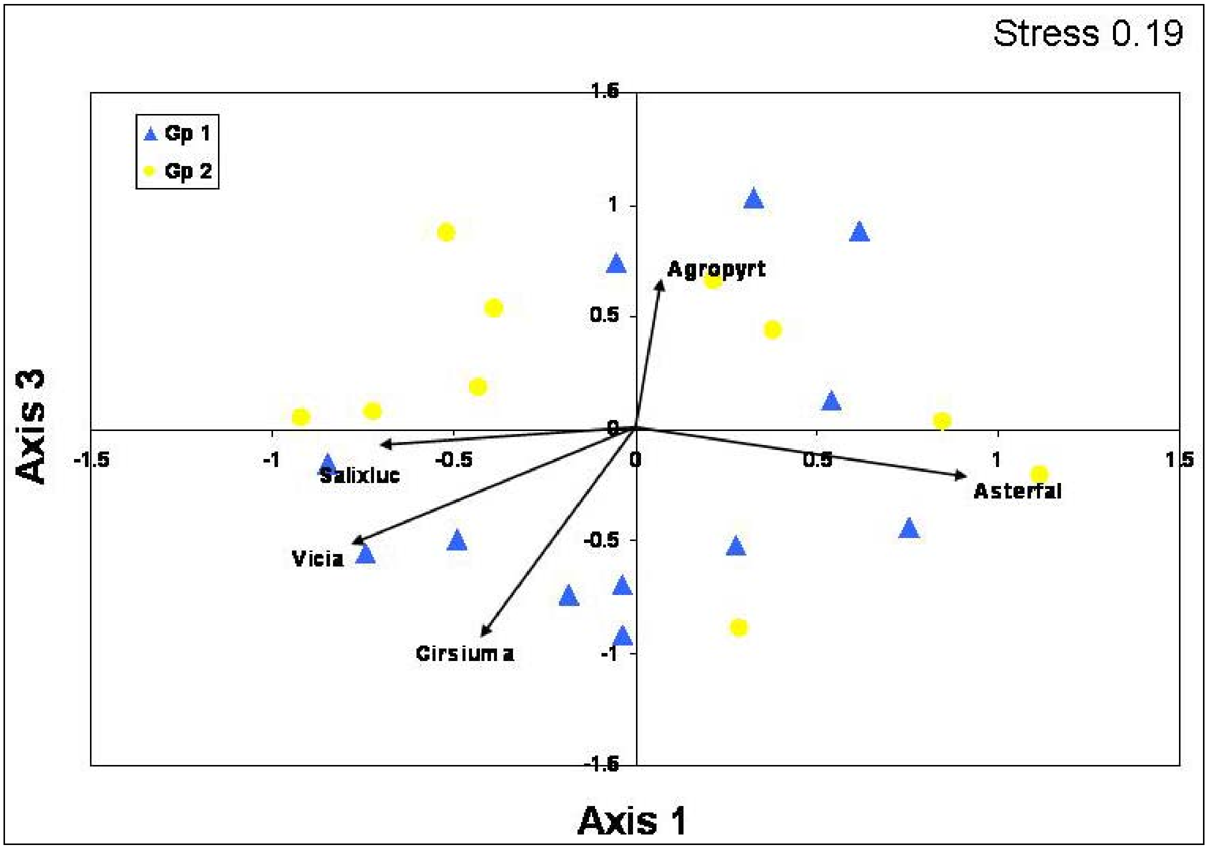


Figure 4 is a plot of axes 1 and 3 from hybrid nMDS; taxa that discriminate the two groups are shown (Table 3). There is a negative association between *Aster falcatus* with *Vicia americana* and *Salix lucida*. The species *Agropyron trachycaulum* is independent of these but there is some negative association with the widely distributed *Cirsium arvense*.

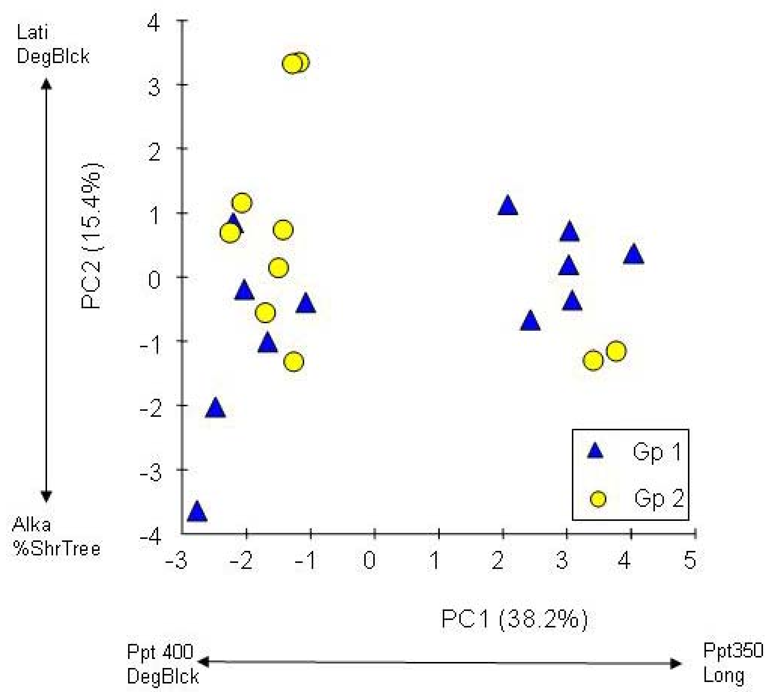
Figure 4: Three dimensional hybrid nMDS ordination (PATN) for 86 plant taxa showing 2 axes and the vectors associated with the ordination.



Relationship with Habitat and Model building

No difference (ANOVA, P > 0.05) was found between any of the environmental variables in the two groups formed by plant species composition. The highest F scores were for Year, Annual mean precipitation (ppt350-400 and ppt400-450); all with F = 2.134 and P = 0.160). A PCA demonstrated that the environmental data do appear to show some difference between the two groups (Figure 5).

Figure 5: Principal components analysis (PCA) of 16 habitat attributes from reference sites, and with groups formed by plant species illustrated



Sites in Group 2 are largely clustered on the left of the graph; these are sites with higher Annual Mean Precipitation (of ppt400-450) and Degraded black soils, and there is clear separation of extrinsic variables on the first component. The summarized results of the PCA are shown in Table 4.

Table 4: Summary of PCA of standardized environmental data.

Variable	PC1	PC2	PC3	PC4	PC5
%Variation	38.2	15.4	11.9	9.3	7.1
Cumulative % variation	38.2	53.6	65.4	74.7	81.8
Year	-0.394	0.043	-0.005	-0.075	-0.046
Julian	0.366	0.008	-0.155	0.079	0.057
Latitude	0.104	0.522	-0.028	-0.255	0.062

Table 4: Summary of PCA of standardized environmental data.

Variable	PC1	PC2	PC3	PC4	PC5
Longitude	0.345	0.258	-0.018	-0.002	0.068
Area wetland (km^2)	0.235	0.044	-0.216	-0.047	-0.189
% Shoreline treed	0.300	-0.273	0.170	0.031	-0.158
% Shoreline willow/shrub	-0.036	-0.470	0.054	-0.444	0.057
Annual Mean Precipitation	0.394	-0.043	0.005	0.075	0.046
(ppt350-400)	-0.394	0.043	-0.005	-0.075	-0.046
Annual Mean Precipitation	0.213	-0.173	0.455	0.084	0.109
(ppt400-450)	0.064	-0.091	-0.650	-0.223	-0.118
Dark Brown (soil)	-0.133	0.426	0.281	0.071	-0.376
Black (soil)	-0.074	-0.018	0.156	-0.459	0.544
Degraded Black (soil)	-0.082	0.063	-0.027	0.500	0.579
Gray (Soil)	-0.101	-0.331	0.156	0.221	-0.324
Alluvium (Soil)	-0.184	-0.156	-0.374	0.374	0.122
Alkali (Soil)					
Introduced species					

The first two components explained only 53.6% of the variation; five components are generally required to explain more than 80% of the variation. This indicates that there is not a strong pattern or structure to the environmental data. Nevertheless, this does provide an indication of which variables may best discriminate the biological groups.

Principle axis correlation (e.g., Figure 3) uses multiple linear regression to “add” extrinsic environmental variables into the biological ‘ordination space’. The extrinsic (or taxa) variables become vectors embedded into the space defined by the ordination axes. The tail of the vectors are always fixed at the centroid of all of the sites in this space and the direction of the vector correlates (literally) maximally with the values of the variables ascribed to the objects. An r^2

value provides an indication of the correlation of the environmental variable with the biological ordination and Monte Carlo analysis gives and indication of significance.

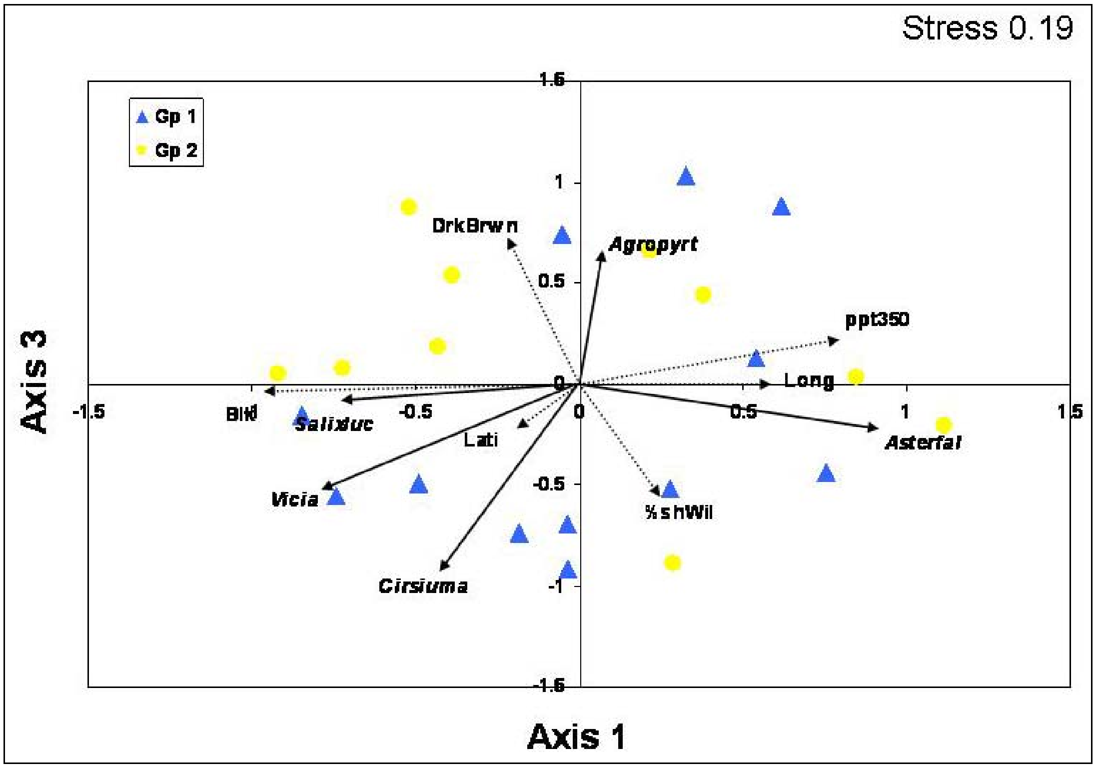
The variables contributing most to the first three components were plotted in nMDS ordination space (axes 1 and 3; Figure 6). These show a positive association between longitude (Long) and Annual Mean Precipitation (ppt350-400) with *Aster falcatus* and these are negatively associated with Black (soils), similarly Dark brown (soils) and % shoreline willow/shrub are negatively associated. However, the relationship between the two matrices is weak (Table 5).

Table 5: Summary of relationship of habitat data with biological nMDS ordination

Variable	**R^2**	**MC analysis -100 random permutations (number more significant than actual)**
Longitude	0.344	6
Evapotranspiration index (ppt350-400)	0.224	14
Annual Mean Precipitation (ppt400-450)	0.224	14
Latitude	0.203	21
Julian	0.198	27
Dark Brown (Soils)	0.189	25
Gray (Soils)	0.185	12
Introduced species	0.146	41
% Shoreline treed	0.141	43
Alluvium (soils)	0.133	57
Alkali (soils)	0.103	83
Area wetland (km^2)	0.082	74
Black (soils)	0.063	72
% Shoreline willow/shrub	0.059	74
Degraded black (soils)	0.040	83

Only longitude showed a significant relationship ($P < 0.10$), and several of the variables contributing to the first three principle components (Black, Dark Brown soils) show little relationship with the species ordination.

Figure 6: Three dimensional hybrid nMDS ordination (PATN) for 86 plant taxa showing 2 axes and the plant species and habitat vectors associated with the ordination.



From these analyses a set of candidate environmental variables were selected for use in DFA that may best separate the two groups formed by the native plant species. Variables were added iteratively beginning with those that seem to best differentiate the two groups in PCA (e.g., latitude and Annual Mean Precipitation ppt350-400).

The number and percentage of sites classified (predicted) to the correct group (by crossvalidation, not resubstitution) are shown in Table 6. Cross validation is a more rigorous test of model performance than resubstitution as the classified site is removed from the model. In addition the between groups F score and the variables used are shown. The best of these models based on how well sites are predicted and the F score is the one using three variables: Annual Mean Precipitation (ppt350-400), longitude and Dark Brown soil. This correctly predicts 80% of the sites in Group 2 and 66% of the sites in Group 1. This model was used to match the test sites to one of the two groups.

Table 6: Results of DFA models tested

Variables in model	Predicted to Group 1 (n 12)	Predicted to Group 2 (n 10)	% correct	F score
Annual Mean Precipitation (ppt350-0400)	6	8	64%	2.134
Annual Mean Precipitation (ppt350-0400), Longitude	9	7	73%	2.380
Annual Mean Precipitation, Longitude, Latitude	7	7	64%	1.688
Annual Mean Precipitation (ppt350-0400) Longitude, Degraded Black (soils)	8	7	68%	1.544
Annual Mean Precipitation (ppt350-0400) Longitude, % Shoreline willow/shrub	7	6	59%	1.923
Annual Mean Precipitation (ppt350-0400), Longitude, Black (soils)	7	5	55%	2.065
Annual Mean Precipitation (ppt350-0400), Longitude, Dark Brown (soils)	8	8	73%	4.613

Test site assessment

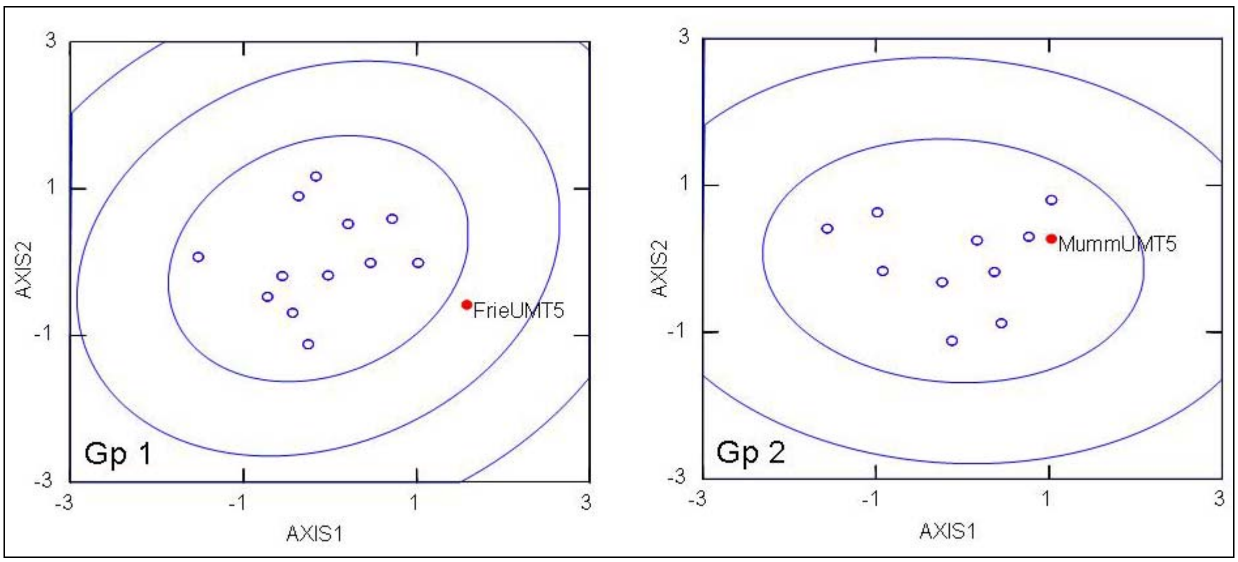
The selected model (Annual Mean Precipitation ppt350-400, Longitude, Dark Brown soils) was used to determine a probability of the test sites belonging to one of the two groups. This was done using DFA with a weighting variable with reference sites weighted as 1 and test sites weighted as 0. The model uses reference sites only to calculate coefficients and then runs the model on test sites. The results are shown in Table 7 for each test site the group to which it is assigned and the probability.

Table 7: Results of DFA in assigning test sites to a reference group

Site	Assigned Group	Probability of being Group 1	Probability of being Group 2
Friesen MT	1	0.986	0.014
Friesen C	1	0.981	0.019
Naharney MT	1	0.924	0.076
Naharney C	1	0.920	0.080
Amey MT	1	0.914	0.086
Amey C	1	0.913	0.087
Gillis C	1	0.900	0.100
Gillis MT	1	0.795	0.205
Leicht C	1	0.748	0.252
Leicht MT	1	0.722	0.278
Mumm MT	2	0.021	0.979
Mumm C	2	0.021	0.979
DeMong C	2	0.139	0.861
DeMongMT	2	0.143	0.857
Bauml MT	2	0.236	0.764
WeilandMT	2	0.282	0.718
Weiland C	2	0.328	0.672
Loiselle MT	2	0.350	0.650
Bull C	2	0.350	0.650
Bull MT	2	0.382	0.618
Loiselle C	2	0.433	0.567

Ten of the sites were assigned to Group 1 and 11 to Group 2. The model predicted sites to the groups with probabilities ranging form 0.99 to 0.72 for Group 1 and 0.98 to 0.57 for Group 2. Site assessment is conducted by comparing the test sites to the reference sites with which it is assigned, so for example site Friesen MT would be compared with the twelve reference sites comprising Group 1, on the basis that it has the same environmental characteristics and if it is undisturbed (in reference condition) it will have the same assemblage of native plants as the reference sites. Determination of the similarity of the native plant assemblage is done by ordination. Reference sites and the test site are ordinated together. Because of the small number of reference sites, this is done one test site at a time. To determine if the test site is the same as the reference sites, probability ellipses are constructed around the reference sites only. In aquatic assessment four quality bands are used to establish a categorized response using 90, 99 and 99.9% ellipses (Reynoldson et al., 2000). In essence these probability ellipses set theoretical Type 1 error rates, that is, using a 90% ellipse around reference sites there is a 10% chance that a reference site will be designated as non-reference.

Figure 7: Assessment of two test sites using ordination (BEAST)



To illustrate the assessment process results of assessing two sites (Friesen MT and Mumm MT) are shown in Figure 7. The left hand plot shows the Group 1 reference sites (open circles) and the test site Friesen MT (red circle). This site is outside the 90% ellipse constructed around the reference sites and will be assessed as non reference. The right hand plot shows test site Mumm MT compared with Group 2 reference sites (open circles) to which it was assigned (P = 0.979), this site would be assessed as in reference condition. In the left hand figure, all three ellipses are shown and a test site can be ranked to one of four categories:

Band 1 – inside the 90% ellipse – in reference condition

Band 2 – between the 90 and 99% ellipse – possibly disturbed

Band 3 – between 99 and 99.9% ellipse – disturbed

Band 4 – outside 99.9% ellipse – very disturbed

In these two examples, site Friesen MT would be assessed in Band 2, and Mumm MT in Band 1.

Summary

The section above describes the basic approach of classification, model building and site assessment that is the reference condition approach (RCA). It should be noted that this is a very minimal data set on which to construct models, and the environmental variables selected may not have been the most appropriate to relate to native plant distribution and abundance. These are both key elements of RCA study designs (Bailey et al., 2004). Therefore, this analysis must be seen as illustrative only.

2) RCA approach using canonical correspondence analysis (CCA)

The first step in this approach was to use DCA to determine whether a unimodal model was appropriate. Gradient lengths >4 in a DCA (Step 1) demonstrate that at least some species in the data matrix show a unimodal response; in our case the gradient length on axis 1 was 3.46 which indicates a modest level of unimodality (ter Braak and Smilaeur, 2002). Because of this finding and the high number of zeros in the data matrix we used CCA rather than redundancy analysis (RDA).

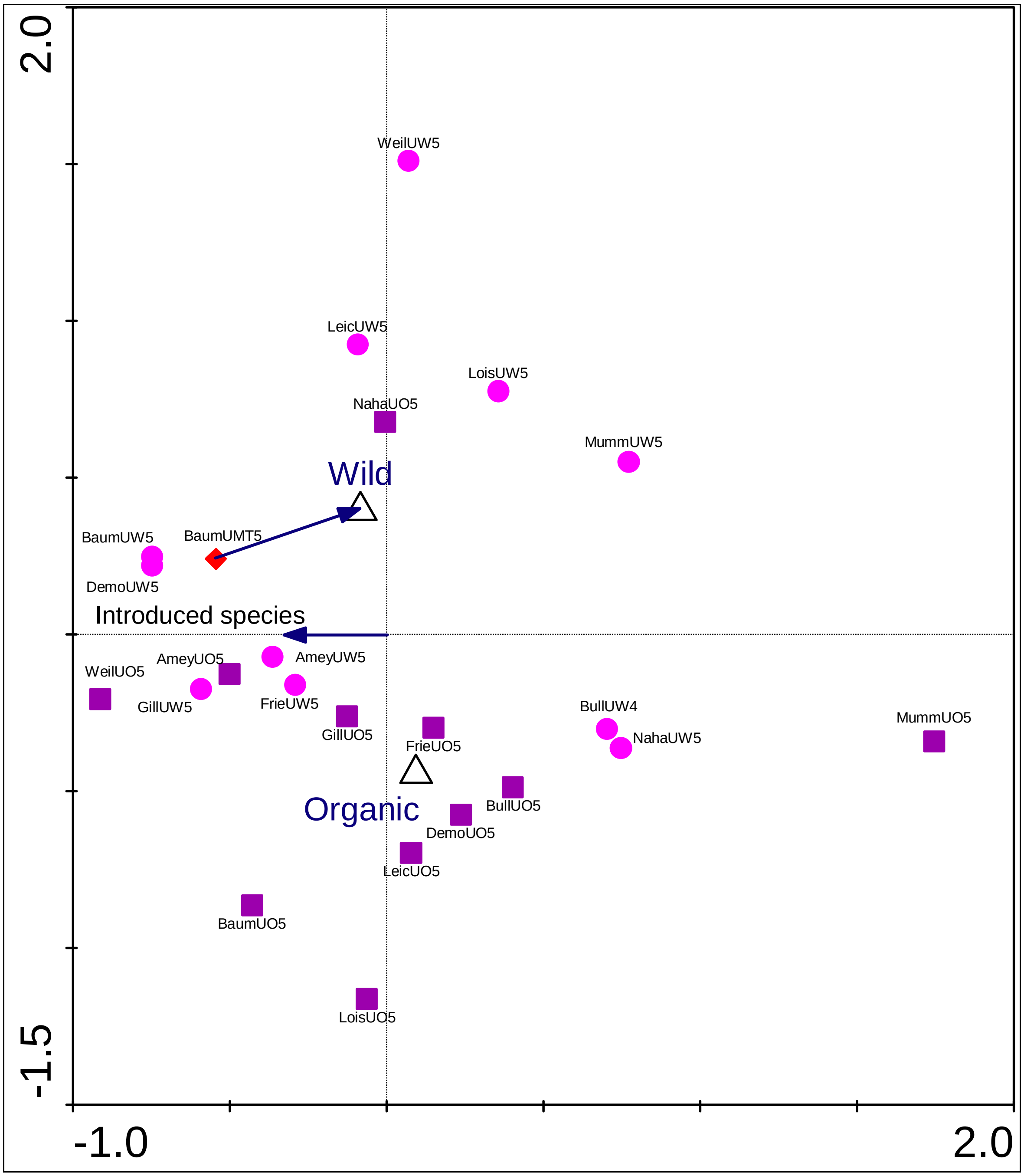
In step 2, stepwise forward selection in CCA indicated that in order of importance the variables explaining significant variation in the species matrix were; bare ground, dark brown soils and percent shoreline willow/shrub (Table 8). We then performed a *p*CCA, where introduced species richness was deployed as the environmental variable and the significant habitat features as covariables.

The resulting ordination plot (Figure 8) shows the *p*CCA results with the introduced species vector on the left side of the graph. The vector linking the centroid of the wild or organic sites (or the centroid of all of these farm types) to the test site can be used to examine the directionality and magnitude of the differences. Note that because these centroids are entered as supplementary in the ordination diagram they do not influence the analysis. We deliberately chose a “test” site which was on the negative end of the axis and thus the trajectory of site quality improvement is in the direction of the wild centroid (i.e., lower introduced species richness).

Table 8: Importance of variables derived from stepwise forward selection in canonical correspondence analysis (for “reference” sites – wild and organic)

	Conditional Effects			
Variable	SR	TVE	P	F
Bare ground	1	0.59	0.001	2.21
Dark brown (soil)	2	0.47	0.009	1.82
% shoreline willow	3	0.39	0.004	1.57
Alluvium (soil)	4	0.35	0.122	1.45
Gray (soil)	5	0.35	0.128	1.46
Moisture (ppt350-4)	6	0.32	0.067	1.39
% shoreline treed	7	0.26	0.29	1.13
Dead vegetation	8	0.24	0.373	1.09
Alkali (soil)	9	0.24	0.454	1.01
Area wetland (km^2)	10	0.23	0.432	1.02
Black (soil)	11	0.16	0.748	0.7

Figure 8: Biplot of site scores from partial canonical correspondence analysis (*p*CCA axes 1 and 2). In this analysis, significant habitat variables were controlled for as covariates and introduced species richness is the explanatory variable). Wild sites are magenta-filled circles, organic sites are purple-filled squares. The centroids for Wild and Organic sites (large open upright triangles) are inserted as supplementary in the ordination (i.e., they do not affect the analysis). The test site Bauml MT is shown as a red-filled diamond. The vector between the test site and the wild centroid shows the desired directionality and distance for change (fewer introduced species)



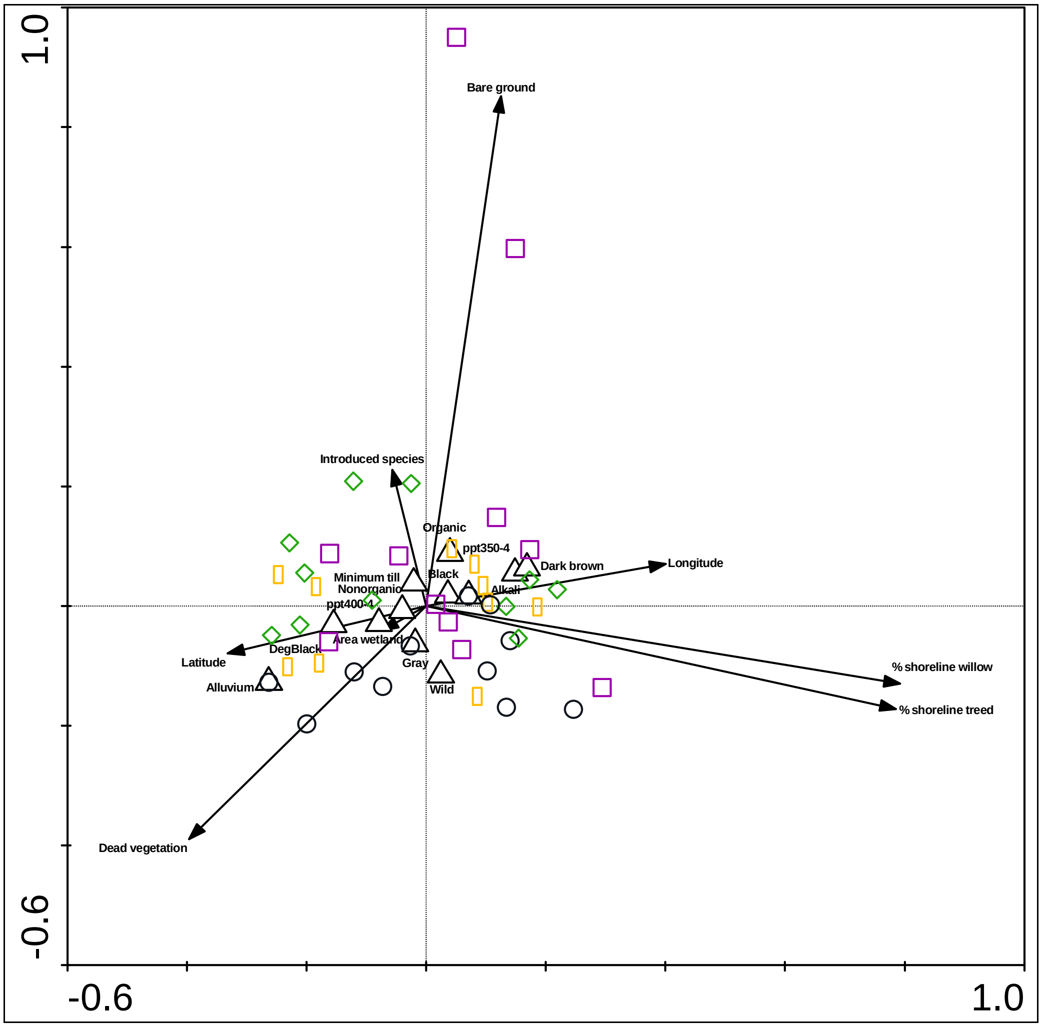
3) Wetland Plant Index (WPI) – based on the Wetland Macrophyte Index (WMI)

In Step 2, we first conducted a CCA on plant species of field sites and examined the relative importance of the environmental variables using forward selection (the variables are listed in Table 9). In order of importance these were % shoreline willow, bare ground, moisture regime, wild farm type, introduced species richness and longitude (Figure 9). Introduced species richness (the stressor variable of interest), accounted for 6.8% of the variation explained by all measured variables (total inertia explained by measured variables was 2.211, or 50.4% of the total inertia in a DCA) and about the half the variation of the most important variable (% shoreline willow explained 14% of the variance from measured variables).

Table 9: Wetland indices approach: Importance of variables derived from stepwise forward selection in canonical correspondence analysis (CCA)

Variable	SR	TVE	F	P
% shoreline willow	1	0.31	3.08	0.001
Bare ground	2	0.25	2.65	0.001
Annual Mean Precipitation (ppt350-4)	3	0.23	2.49	0.001
Wild	4	0.16	1.76	0.004
Invasive species	5	0.15	1.73	0.002
Longitude	6	0.14	1.53	0.011
Latitude	7	0.11	1.26	0.123
% shoreline treed	8	0.1	1.2	0.172
Dead vegetation	9	0.1	1.19	0.205
Conventional	10	0.1	1.13	0.257
Dark Brown soils	11	0.1	1.19	0.225
Gray Soils	12	0.09	1.05	0.38
Alkali soils	13	0.08	1.01	0.377
Minimum tillage	14	0.08	0.91	0.61
Alluvium soils	15	0.08	0.88	0.602
Black Soils	16	0.06	0.77	0.818
Area of wetland (km^2)	17	0.07	0.78	0.745

Figure 9: Biplot of site scores from canonical correspondence analysis (CCA axes 1 and 2). Vectors for the 7 continuous environmental variables are shown as lines with arrows. Strengths of correlations with the axes are indicated by the length of lines and their proximity to the axes. Centroids of nominal environmental variables (farm type, soil types, moisture regime) are show as large open triangles. Farm types are: wild = black open circles; organic = purple open squares; minimum tillage = green open diamonds; nonorganic = ochre open rectangles



Having categorized species by their U and T values (Step 3, 4, 5; see Table 10), we then used equation 1 to rank each site in relation to habitat quality (degree of impairment from introduced species). If a WMI of ≥ 3 is considered indicative of relatively unimpaired sites and a <3 indicative of impaired status, then 15 of the 43 sites can be considered impaired (relative to their invasive species richness) compared to 26 that were unimpaired. Interestingly, of the wild sites (considered reference in the other two approaches, only 1 was considered impaired, compared to 10 that were unimpaired. Of the other farm types, approximately half of each was in the impaired and unimpaired category (Table 10).

Axis 1 was a gradient from more heavily treed shorelines (willow and other shrubs) on the right to sites with low shoreline shrub cover, at higher latitudes (Figure 10). Axis 2 was a gradient from sites with a high percentage of bare ground and introduced species to sites with less bare ground, fewer introduced species and more sites in the wild farming type category. Species that were positively associated with the sites with higher mean introduced species richness (i.e., , indicators of impaired sites – see Step 6) included *Crepis runcinata, Hordeum jubatum, Cryptantha fendleri, Atriplex patula*, *Artemisia biennis* and *Juncus bufonius* (close to the vector for introduced species in Figure 10). All of the latter species were positioned in the upper part of the ordination diagram (close to the vector for introduced species richness) and thus had low U values (1; Table 10). By contrast, species that were negatively associated with introduced species richness (i.e., , indicators of relatively unimpaired sites were *Anemone riparia*, *Psoralea esculenta*, *Antennaria parvifolia, Ribes lacustre,* and *Achillea millefolium.* All of the latter species were positioned in the lower part of the ordination diagram (opposite the vector for introduced species richness) and thus had high U values (5; Table 10).

Figure 10: Biplot of species scores from canonical correspondence analysis (CCA axes 1 and 2). See Table 12 for species abbreviations (only some species shown as examples to prevent name crowding)

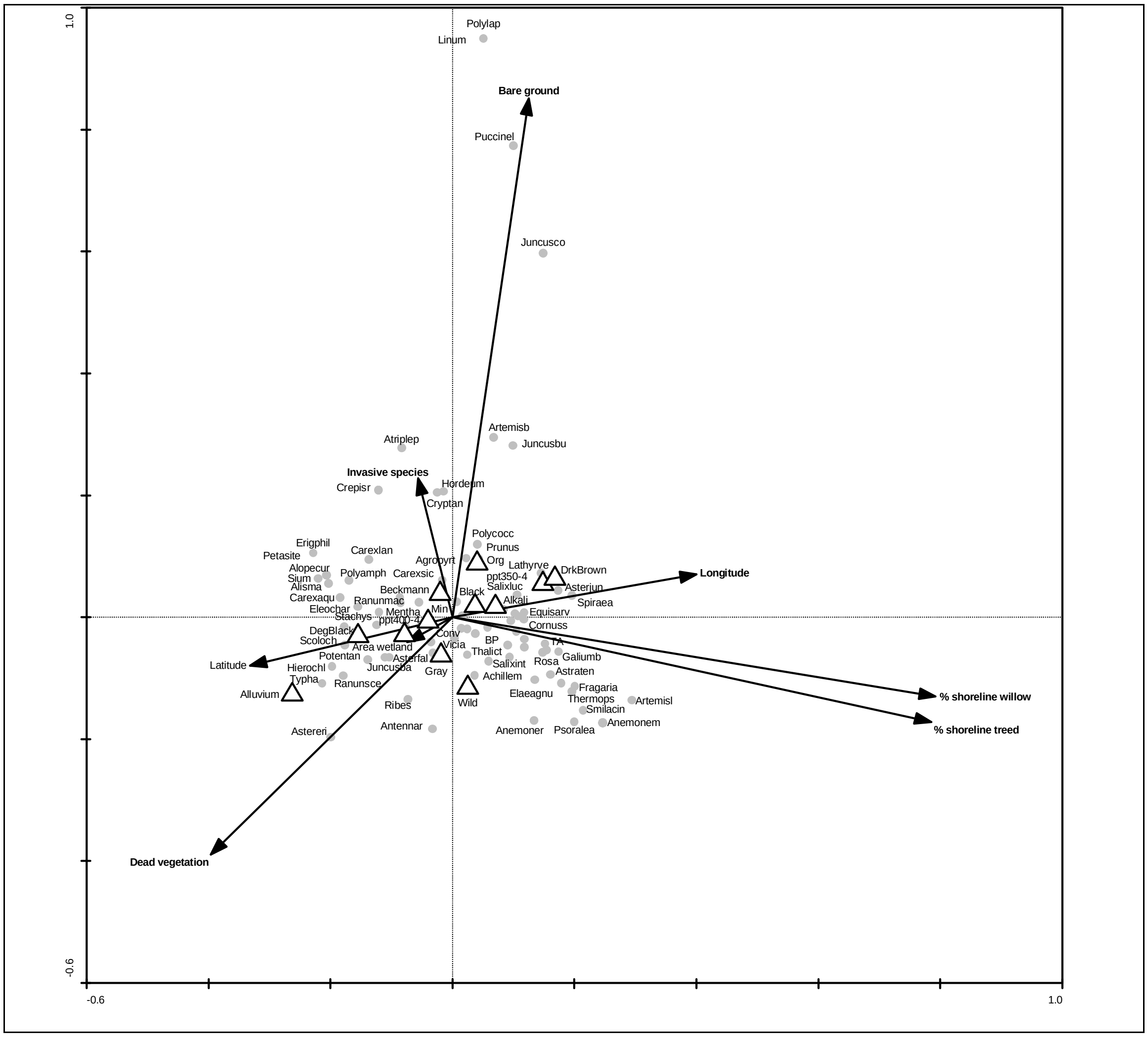


Table 10: List of plant species from SCB case study for wetland indices approach showing acronyms, % occurrence at sites and U and T values

Abbreviation	Scientific name	% occurrence	T	U
Achillem	*Achillea millefolium*	18.6	2	4
Agropyrs	*Agropyron smithii*	14.0	2	3
Agropyrt	*Agropyron trachycaulum*	51.2	1	2
Alisma	*Alisma plantago*	2.3	3	2
Alopecur	*Alopecurus aequalis*	4.7	2	2
Amelanch	*Amelanchier alnifolia*	9.3	2	4

Table 10: List of plant species from SCB case study for wetland indices approach showing acronyms, % occurrence at sites and U and T values

Abbreviation	Scientific name	% occurrence	T	U
Anemonec	*Anemone canadensis*	20.9	2	3
Anemonem	*Anemone multifida*	2.3	3	5
Anemoner	*Anemone riparia*	2.3	3	5
Antennar	*Antennaria parvifolia*	4.7	2	5
Artemisb	*Artemisia biennis*	16.3	1	1
Artemisc	*Artemisia campestris*	2.3	3	5
Artemisl	*Artemisia ludoviciana*	2.3	3	5
Asclepia	*Asclepias ovalifolia*	2.3	3	5
Astercil	*Aster ciliolatus*	2.3	3	4
Astereri	*Aster ericoides*	2.3	3	5
Asterfal	*Aster falcatus*	18.6	1	4
Asterhes	*Aster hesperius*	67.4	2	3
Asterjun	*Aster junciformis*	2.3	3	2
Asterlae	*Aster laevis*	2.3	3	5
Asterpta	*Aster ptarmicoides*	2.3	3	5
Astradan	*Astragalus danicus*	9.3	2	4
Astraten	*Astragalus tenellus*	7.0	2	4
Atriplep	*Atriplex patula*	4.7	1	1
Beckmann	*Beckmannia syzigachne*	16.3	2	3
BP	*Populus balsamifera*	16.3	2	4
Campanap	*Campanula aparinoides*	2.3	3	5
Carexaqu	*Carex aquatilis*	2.3	3	3
Carexatd	*Carex atherodes*	20.9	2	4
Carexats	*Carex athrostachya*	2.3	3	5
Carexlan	*Carex lanuginosa*	4.7	2	2
Carexpra	*Carex praegracilis*	2.3	3	4
Carexsic	*Carex siccata*	7.0	2	2
Castille	*Castilleja miniata*	2.3	3	4
Cirsiuma	*Cirsium arvense*	83.7	1	3
Cornuss	*Cornus stolonifera*	9.3	2	3
Crepisr	*Crepis runcinata*	2.3	3	1
Cryptan	*Cryptantha fendleri*	2.3	3	1
Elaeagnu	*Elaeagnus commutata*	7.0	2	4
Eleochar	*Eleocharis spp.*	18.6	1	3
Epiloban	*Epilobium angustifolium*	2.3	3	3
Equisarv	*Equisetum arvense*	4.7	2	3

Table 10: List of plant species from SCB case study for wetland indices approach showing acronyms, % occurrence at sites and U and T values

Abbreviation	Scientific name	% occurrence	T	U
Equispra	*Equisetum pratense*	23.3	2	3
Erigcana	*Erigeron canadensis*	2.3	3	3
Erigphil	*Erigeron philadelphicus*	2.3	3	2
Fragaria	*Fragaria virginiana*	4.7	2	4
Galiumb	*Galium boreale*	20.9	2	4
Galiumtd	*Galium trifidum*	9.3	2	3
Galiumtr	*Galium triflorum*	2.3	3	2
Geumtrif	*Geum triflorum*	2.3	3	5
Hedysaru	*Hedysarum boreale*	2.3	3	5
Hierochl	*Hierochloe odorata*	4.7	2	4
Hordeum	*Hordeum jubatum*	14.0	1	1
Iva	*Iva axillaries*	2.3	3	1
Juncusba	*Juncus balticus*	23.3	2	4
Juncusbu	*Juncus bufonius*	7.0	1	1
Juncusco	*Juncus confuses*	2.3	3	1
Lactucap	*Lactuca pulchella*	2.3	3	5
Lathyroc	*Lathyrus ochroleucus*	4.7	2	3
Lathyrve	*Lathyrus venosus*	7.0	2	2
Linum	*Linum lewisii*	2.3	3	1
Lysimcil	*Lysimachia ciliata*	2.3	3	4
Mentha	*Mentha arvensis*	34.9	2	3
Monarda	*Monarda fistulosa*	2.3	3	5
Moss	*Moss*	2.3	3	2
Muhlen	*Muhlenbergia cuspidata*	2.3	3	1
Penstemp	*Penstemon procerus*	2.3	3	5
Petalost	*Petalostemon purpureus*	2.3	3	5
Petasite	*Petasites sagittatus*	4.7	2	2
Phalaris	*Phalaris arundinacea*	9.3	2	3
Poapalu	*Poa palustris*	23.3	2	3
Polyamph	*Polygonum amphibium*	9.3	2	2
Polycocc	*Polygonum coccineum*	4.7	2	2
Polylap	*Polygonum lapathifolium*	2.3	3	1
Potentan	*Potentilla anserine*	27.9	2	4
Potentar	*Potentilla arguta*	2.3	3	5
Prunus	*Prunus virginiana*	2.3	3	2
Psoralea	*Psoralea esculenta*	4.7	2	5

Table 10: List of plant species from SCB case study for wetland indices approach showing acronyms, % occurrence at sites and U and T values

Abbreviation	Scientific name	% occurrence	T	U
Puccinel	*Puccinellia nuttalliana*	4.7	1	1
Ranuncym	*Ranunculus cymbalaria*	4.7	1	4
Ranunmac	*Ranunculus macounii*	9.3	2	3
Ranunsce	*Ranunculus sceleratus*	2.3	3	4
Ribes	*Ribes lacustre*	2.3	3	5
Rorippa	*Rorippa islandica*	2.3	3	2
Rosa	*Rosa woodsii*	23.3	2	4
Rubus	*Rubus idaeus*	2.3	3	2
Rumexmar	*Rumex maritimus*	2.3	3	2
Salixbeb	*Salix bebbiana*	7.0	2	3
Salixint	*Salix interior*	2.3	3	4
Salixluc	*Salix lucida*	27.9	1	3
Salixlut	*Salix lutea*	2.3	3	3
Salixser	*Salix serissima*	9.3	2	3
Scoloch	*Scolochloa festucacea*	16.3	2	4
Sium	*Sium suave*	7.0	2	2
Smilacin	*Smilacina stellata*	4.7	2	5
Solidago	*Solidago canadensis*	11.6	2	3
Spiraea	*Spiraea alba*	4.7	2	3
Stachys	*Stachys palustris*	25.6	2	3
Stellaril	*Stellaria longipes*	4.7	2	3
Stipa	*Stipa viridula*	2.3	3	5
Syoc	*Symphoricarpos occidentalis*	23.3	2	4
TA	*Populus tremuloides*	30.2	2	4
Thalict	*Thalictrum venulosum*	11.6	2	4
Thermops	*Thermopsis rhombifolia*	7.0	2	5
Trifohyb	*Trifolium hybridum*	2.3	3	4
Trisetum	*Trisetum wolfii*	2.3	3	2
Typha	*Typha latifolia*	4.7	2	4
Urtica	*Urtica gracilis*	2.3	3	3
Vicia	*Vicia americana*	27.9	2	3
Violanut	*Viola nuttallii*	2.3	3	5
Zizia	*Zizia aptera*	2.3	3	3
Zygaden	*Zygadenus elegans*	2.3	3	5

We compared the WPI scores derived for each site for the different farm types (Table 11). As predicted, wild sites had the highest mean score indicating that they were the least impaired and had significantly higher scores than all other farm types (Generalized Linear Model; PROC GLIMMIX in SAS; SAS, 2000; F = 4.93, P = 0.032). Wild WPI scores were significantly different from minimum tillage (F = 5.18, P = 0.0284) and organic scores (F = 4.33, P = 0.044), but not nonorganic farm scores (F = 1.17, P =0.287). The latter seems counterintuitive as organic and minimum tillage would be expected to be more similar to wild scores than nonorganic; however, no difference was found in contrasts between all other farm pairs which suggested that the most important finding was that wild sites were most distinctive.

While organic sites had a slightly higher mean WPI score than minimum tillage or nonorganic sites, this difference was not significant. We were not able to validate the scores since we did not have another set of data from the same area (temporal comparison) or from other more pristine sites (Steps 7 and 8). This type of comparison can be done by conducting a linear regression model or paired t-tests.

Table 11: WMI scores from approach 3 showing quarter section farm owners and farm type

Quarter section	Farm type	WMI
Friesen	Wild	2.83
Gillis	Wild	3.10
Naharney	Wild	3.17
Amey	Wild	3.30
Bull	Wild	3.31
DeMong	Wild	3.33
Mumm	Wild	3.57
Leicht	Wild	3.70
Loiselle	Wild	3.74
Weiland	Wild	3.81
Bauml	Wild	3.87
	Mean	3.43

Table 11: WMI scores from approach 3 showing quarter section farm owners and farm type

Quarter section	Farm type	WMI
	SD	0.33
Loiselle	Organic	1.71
Naharney	Organic	2.00
Gillis	Organic	2.75
DeMong	Organic	2.91
Leicht	Organic	2.94
Bauml	Organic	3.08
Weiland	Organic	3.25
Bull	Organic	3.33
Mumm	Organic	3.36
Friesen	Organic	3.40
Amey	Organic	4.16
	Mean	2.99
	SD	0.67
Bauml	Mintill	1.57
Loiselle	Mintill	2.58
Amey	Mintill	2.80
Bull	Mintill	2.93
Leicht	Mintill	2.94
Weiland	Mintill	3.00
Naharney	Mintill	3.05
DeMong	Mintill	3.22
Gillis	Mintill	3.29
Mumm	Mintill	3.33
Friesen	Mintill	3.73
	Mean	2.95
	SD	0.55
Gillis	Nonorganic	2.80
Naharney	Nonorganic	2.89
Bull	Nonorganic	2.94
Leicht	Nonorganic	2.95
Loiselle	Nonorganic	3.00
DeMong	Nonorganic	3.36
Mumm	Nonorganic	3.40
Amey	Nonorganic	3.41
Weiland	Nonorganic	3.50

Table 11: WMI scores from approach 3 showing quarter section farm owners and farm type

Quarter section	Farm type	WMI
Friesen	Nonorganic	3.71
	Mean	2.96
	SD	0.85

5 DISCUSSION

We have outlined several different approaches to assessing habitat quality within a framework of multivariate analyses. Using such a modelling approach based on empirical data (species composition and abundance) as the response variable and environmental data (biophysical features, ecological processes and agricultural stressors) as the predictor variables provides a rigorous and scientifically defensible framework for bioassessment. Such models can be developed to compare between sites at one point in time (e.g., unimpaired versus impaired sites), or over time (e.g., comparing the same site between years, so that trends in its condition can be assessed). These models should be evidenced-based and thus benefit from much accumulated existing scientific and expert knowledge on different taxa and their responses to environmental features and anthropogenic stressors.

It has been recommended that diverse taxa and levels of organization be used because indicator groups or individual species may fail to reflect human disturbance. The idea of including groups representing structure, function and composition is useful, and all or many of these parameters would be encompassed by straightforward monitoring of as many taxa as possible for the minimum possible cost. For example, numerous studies have demonstrated the taxa groups that can be monitored in farmland, and the environmental and stressor variables to which they respond (i.e., that determine their composition and abundance; Canada, Freemark and Kirk, 2001; Europe, Billeter et al., 2007).

While it has been argued that collating information on multi-species taxa groups may contain redundancy, and that this is not cost effective, having some redundancy is useful because not all environmental problems can be easily detected or foreseen and individual taxonomic groups vary in their response to different stressors. Several successful terrestrial and aquatic monitoring programs have already been set up and tested which use multi-species groups for bioassessment (e.g., terrestrial, Manley et al., 2004, ABMI; aquatic RIVPACS, AUSRIVAS, CABIN).

Species richness is not appropriate as a measure biodiversity conservation measure for three reasons. First, trends in counts are subject to bias because of changes in detectability over time. Second, richness alone can be misleading to evaluate conservation value because shifts in community composition which are significant in terms of biodiversity value (e.g., replacement of rare or specialist species by generalist or invasive species) are not reflected in a simple species richness index. This is particularly important in agricultural landscapes, where species populations may survive in source sink habitat, where levels of edge/generalist species are high, and various processes such as predation and brood parasitism by Brown-headed Cowbirds *Molothrus ater* is elevated. To take a specific example, although avian species richness in shelterbelts may be high (but not as high as woodlands) often the species in shelterbelts are generalists that can adapt well to anthropogenic landscapes and not the specialist species that cannot. Third, recent studies have suggested a lack of congruence from species richness data from different taxa; i.e., the species richness of one taxa group is not necessarily a good predictor for other groups.

A second important component of any framework is that environmental variables be integrated into a model that relates to the distribution and abundance of these multi-species taxa groups.

The benefit here is that relationships and links between variables and species composition and abundance can be demonstrated empirically. This does require some prior biological knowledge about the relevant environmental variables to measure as well as the farming practices and other stressors that are important to different taxa. However, without identifying these relationships standards could be somewhat arbitrary. Some guidance is available in the scientific literature but it is important to consider that taxa responses may vary by geographic location – however, meta-analyses can help in generalizations. The general recommendation is that a wide range of environmental descriptors are measured at a variety of scales (see Bailey et al., 2004).

The multivariate techniques used for the case study are ideally suited to examination of site quality. However, each has advantages and disadvantages, and they all have some commonalties: 1) they all use a species and environmental data matrix; 2) each method attempts to identify the most important habitat/extrinsic features determining species composition and abundance; 3) each attempts to identify a stressor gradient; and 4) at least the first 2 approaches (RCA using CABIN and RCA using CCA) uses a set of sites as reference (relatively unimpaired) and then examines the deviation of test sites from a goal (within reference condition probability ellipses or distance from a centroid).

Generally, the steps required for approach 1 are by far the most complex since it involves exploratory analyses of which species and environmental factors are indicators for different reference site cluster groups, how these separate in nMDS ordination space and then inserting test sites to see if they lie inside or outside probability ellipses. By contrast, the steps in approach 2 are simpler in that environmental variables are incorporated directly into the ordination model – the ordination axes are constrained to be linear combinations of environmental variables. The final analysis is based on residuals since the effect of significant habitat variables (not influenced by human disturbance) are factored out, while examining the effect of the stressor variable of interest (introduced species richness).

There are several general limitations of our case study analysis. First, as previously stated, the study was designed as a comparison of the effects of farming regime and moisture content on the distribution and abundance of plant species. While the farms chosen incorporated a range of farming practices and sites that could potentially act as reference sites, as well as potentially impaired and restored sites, this was not tested independently from farm type. In the case study, we chose wild and organic farms as reference sites for demonstration purpose only and based on the fact that wild sites were either in native grassland or permanent cover program sites (Ducks Unlimited Canada), and organic sites had no pesticide use in the years prior to the study and were managed in accordance with organic farming standards (see http://www.organicagcentre.ca/std_canadian.html). However, both wild and organic farm types could be problematic as reference sites. Except for the Loiselle cluster which was native prairie, most wild sites were cropland seeded to perennial cover, including wheatgrass (*Agropyron* spp.), bromegrass (*Bromus* spp.), alfalfa (*Medicago* spp.) and crested wheatgrass (*Agropyron desertorum*). Thus, they included introduced species – which was the stressor selected for the case study. If a larger sample size had been collected, native prairie grassland sites could have been selected as reference sites, and would presumably have had a very low incidence of introduced species. Organic sites, because of their lack of herbicide use for weed control, may also have more introduced species than conventionally farmed areas.

Second, we had a relatively small sample size to test the results and the groups of four farm types were close to the minimum cluster used for some approaches (the CABIN minimum group size for clusters is 10 – Reynoldson and Wright, 2000). Because of small sample size we had to combine both wild and organic sites as reference. Small sample size also limited the number of variables that could be included in models. Third, the stressor gradient we chose may not have been the most appropriate to use with these data, but this was the only variable available.

In relation to the main ordination methods used, some controversy exists over whether to use eigenanalysis (e.g., CA, CCA) or distance-based ordination (e.g., nMDS) methods and there are two schools of thought. Generally speaking, proponents of distance-based methods are in Australia, the United States and Canada, whereas eigenanalysis ordination is used more in Europe and to a lesser extent in Canada and the United States. There are few completely objective examinations of the merits of both types of ordination methods, partly because there is a vested interest in experts expounding their particular softwares. The best approach for comparing methods may be to use simpacted (simulated impact) data. This procedure is completely objective and is currently being used to compare Type 1 and 2 error rates with the BEAST, RIVPACS and regression modelling (T. Reynoldson, pers. comm.).

One of the main criticisms of eigenanalysis ordination relates to how the distances between sites are calculated (they are based on the weighted abundance of each species – the chi-squared distance) and the "double-zero" problem. The latter is the case when a species is absent at two sites; the sites may both be above or below the optimal niche value for the species in question, or one of the sites may above and the other below the value. It is impossible to tell which of these assumptions are true (Legendre and Legendre, 1998). In this situation, the best recourse may be to not make any ecological assumptions on the significance of the species' absence from the two sites (Legendre and Legendre, 1998). The two types of ordinations (eigenanalysis and distance-based techniques) differ in how they treat these zero values. In CCA, absences are counted as indications of resemblance. However, some criticisms of eigenanalysis ordination, particularly

CCA, are based on an incomplete understanding of the technique (e.g., Austin, 2002). Moreover, many criticisms of CCA tend to be applicable to a whole array of regression modelling techniques and it is therefore not appropriate to single out CCA (Palmer, 2008; M. Palmer pers. comm.).

In distance-based techniques, the distance between sites in ordination space is closer to ecological differences (Urban, 2006). Moreover, in nMDS no assumptions at all are made about the data (and double zeros are skipped altogether when computing similarity coefficients). Thus it could be argued distance-based ordination methods may be more appropriate for an RCA type approach where distances between sites in ordination space either spatially or over time is critical. For example, we need to assess first whether adaptive management is moving sites in the direction of desired change and second, how fast these sites are "moving" in ordination space (i.e., , the length of the environmental vector of interest). Both of these are implicit in setting standards for site quality.

Retaining the species information is extremely useful, both for ordination interpretations and to identify potential indicators and is done in eigenanalysis ordination; CA and CCA have been used extensively in both aquatic and terrestrial ecosystems.

Specific comments on Approaches 1-3 (case study)

Approach 1

We have described in some detail how bioassessments can be carried out using the BEAST method (approach 1) which deploys nMDS ordination and other techniques to determine if the species assemblage at a test site matches the reference site to which it is assigned. The method allows a test site to be classified to one of four quality bands. One advantage of this method is that it makes no prior assumptions as to which attributes of a community are important and it uses all of the species composition and abundance data. The major disadvantage is that it uses the probability and in some cases the group to which a site belongs can be equivocal (e.g., Loiselle C, in Table 5). However, this is a relatively rare occurrence. Moreover, this approach has been used and tested extensively in on aquatic invertebrate communities in the Great Lakes, Fraser River in British Columbia and the Yukon. It is the basis of national monitoring programs in the UK and Australia and has been extensively tested and reported on in the primary literature.

There are alternative assessment methods using RCA and the probabilistic output. Perhaps the best known of these is the RIVPACS method, which deploys presence-absence data and uses the expected to observed taxon occurrence information. The method uses the weighted probabilities of the site group membership and the percent occurrence of taxa within the group to calculate an expected taxa richness for a site. It also provides the probability of any taxon's occurrence. The advantages of the method are that it uses all the probabilities, so error associated with mis-assignment are eliminated and it also gives an expected taxon occurrence list (i.e., predicts what taxa will occur). The disadvantages are that assessment is based on richness only and that it only uses presence absence data. Other approaches that can use the method are to use various metrics and measures univariately (e.g., species diversity) but use the RCA model to select an appropriate group of reference sites for univariate comparison.

Approach 2

Using stepwise forward selection in CCA to identify significant habitat variables suffers from the same criticism as it does in regression modelling. Some of the variables entered may be arbitrary, especially if they are highly correlated with other variables. Thus a careful evaluation of the relationships among variables needs to be done to see which variables are correlated. The approach also assumes that the best drivers of species distribution and abundance have been selected and that a sufficient range of samples are available to determine gradients (note that this also applies more generically to other models). Moreover, it is important to determine whether stressor variables are co-correlated with other (unmeasured) variables. Examining the effect of one stressor at a time may also be problematic, unless they are along the same impairment gradient. The latter would often be the case (for example in the case of wetlands, those with high phosphorus concentrations would also be likely to have high nitrate levels). We used the centroids of reference conditions (wild and organic sites) as a measure to compare positions of test sites in ordination space. This could be done with probability ellipses as in Approach 1. However, it is important to note that the analysis would be done on the residuals of the species data (after extracting the effect of significant habitat covariates, not influenced by human activities). If the test sites fell outside the ellipse, it would be relatively simple to interpret, namely, the test sites are less likely affected by the stressors under study. However, if test sites fell within the ellipse, there could be more than one explanation. The following are two alternatives: 1) The stressors have an impact on the test sites; 2) The test sites may be affected by other factors that are confounded with the stressors. So it is important that the effects of the most relevant variables have been taken into accounts and partitioned out.

Approach 3

In this method, sites are ranked according to the positions of species along the ordination axis of interest. The approach has been tried and tested extensively for fish, macroinvertebrates and aquatic plant species in the Great Lakes Region. As in most approaches there is an element of investigator skill and experience in approaching the bioassessment (moreover, plant identification skills during field surveys could also affect results and is the subject of current research). Moreover, some degree of trial and error is involved in running the ordinations (as with the other two approaches) and it is important to conduct analyses that make ecological sense. For example, ubiquitous species and rare species can skew the data. It may be advisable to experiment with different runs of the ordinations to see how different options affect analyses (e.g., including rare species but downweighting them, or excluding rare species by setting a minimum threshold of occurrence or abundance). To some extent dividing up the U and T scores of species along ordination axes can be arbitrary. We chose equal categories to assign scores to avoid bias (A. Wei, pers. comm.).

Establishing reference conditions

Establishing reference conditions in agricultural landscapes is a considerable challenge, especially deciding which unimpaired ecosystems to use as a baseline. Some may argue that the entire concept of finding pristine sites is flawed since today no ecosystems escape human influence (e.g., climate change). Moreover, the definition of "healthy" is problematic. For example, consider the simple example of a riparian system which has a narrow fringe of woodland cover containing both native and exotic tree and shrub species. In terms of stream bank quality and role in intercepting nutrients and pollutants, and reducing sediment run-off, the riparian area may be considered healthy. However, from a biodiversity perspective, it may provide a corridor for predators that prey on birds in adjacent grassland landscapes. Moreover, while it may have higher species richness than adjacent habitats, the species may be composed largely of generalist or edge species, rather than the specialist species that are declining in agricultural landscapes. At a larger scale this could result in regional reductions in populations of habitat specialist species and increases in generalist or edge species.

Bailey et al. (2004) summarized the approach used for aquatic conditions by Davies (1994) in Australia and recommended that reference sites: 1) encompass a wide range of physical and chemical conditions and include rivers varying in size, water discharge and altitude within the study region; 2) that they are minimally disturbed – with the caveat that lowland rivers would be affected by land use practices; 3) they are representative of location or reach and not impacted by flow regulations; 4) they should be easily accessible and safety considerations for field workers considered.

Benchmarks are important and can be measured in four ways (Nielsen et al., 2007): 1) protected areas; 2) time-zero; 3) desired goals or targets; and 4) using empirical estimates of the relationship between species' occurrences/abundance and stressors (human footprint) one can estimate reference conditions under pristine conditions (see Nielsen et al., 2007). The problem with protected areas is that most have been selected arbitrarily based on their remoteness, inaccessibility and low resource potential, rather than by a systematic conservation planning procedure. It is important to note that the concept of a reference needs to be flexible and appropriate to the context, and partially based on societal value judgments. For example, in agricultural landscapes reference sites could be habitats on farms using best management practices. While these could be compared with protected area reference sites, this would answer a slightly different question.

Using time zero (selecting a point in time to compare against current conditions) to evaluate changes in local habitat quality in response to adaptive management is problematic because usually an insufficient period of time is available to provide sufficient data to inform models. Such an approach also requires an ongoing monitoring program, because time-series data are needed on species, habitat and processes to model responses. While this approach should be implemented in agricultural landscapes, at the same time it is important to survey a range of sites along an anthropogenic disturbance gradient. Not only is this important to provide data that would improve understanding and prediction, but it could also provide the instrument to categorize sites based on their condition, and thus set standards in human-disturbed landscapes.

The last technique has used generalized linear models to estimate species occurrence and abundance in the absence of human impact. For example, Nielsen et al. (2007) used this approach to model mammalian snowtracking data in relation to road density. Because CCA is a regression technique it could also be used to empirically identify reference conditions, but for a multi-species taxa group rather than single species one at a time. Moreover, multiple regression is used in PCC to examine the importance of vectors in nMDS so this could also be used.

How can these approaches be used to set standards and targets?

It is important to point out that approaches outlined here are only a first step and that the management decisions made following the evaluation are the key to restoration of habitat quality. The multivariate techniques used in the case study could be used in a variety of ways to set standards and target goals for biodiversity conservation in agricultural landscapes. For example, in the RCA approach, the probability ellipses are a standard by which to measure impairment of test sites. Test sites that do not occur within the different probability ellipses can be considered to have failed the Reference Condition. However, a integral part of the process to setting standards is to determine the degree of deviation from reference. This can be determined from a site's position in ordination space (vector distance) or by categorizing sites by the degree of damage. These categories of damage are a way of simplifying assessments made from graph or map presentations and are used in RIVPACS (Clarke, 2000), AUSRIVAS (Simpson and Norris, 2000), BEAST (Reynoldson et al., 2000) and the IBI (Kerans and Karr, 1994). In AUSRIVAS, these bands are based on the observed to expected ratio of taxa (Ball et al., 2001). Band definition is extremely critical because incorrect assignment could lead to similar sites being allocated to different band categories and receiving very different management treatments. Similarly, in approach 3, based on the WPI scores, sites can be divided into those that are relatively impaired compared to those that are relatively unimpaired. Thresholds could be used to determine what WPI value to use for sites to attain a standard. These can be derived from comparisons of the WPI score from before and after impact studies (BACI), before or after adaptive management, comparing historical WPIs with current conditions, comparing reference sites with test sites or comparing the index with other condition assessments (e.g., RCA). Within the modelling environment of CCA it is possible to set standards for environmental variables by examining directionality and length of vectors. Both CCA and nMDS can be used to inform selection of species for single species regression models. Beyond assessing the severity of failure of an impaired site we also need to know specifically why it failed. This can be done by detailed site assessments – biotic composition and ecology, sensitivity or tolerance of species to stressors and habitat condition and position of the site in ordination space. Finally, scenarios can be built to determine the possible outcomes of restoration or increased exposure to stressors. These could be Generalized Linear Models (GLMs) and both reference sites and test sites would be included together in the model. An important distinction is that in the assessments for the RCA alone the predictor variables are those which describe natural variability in communities (i.e., not influenced by human activity), whereas in the scenario models predictors include stressor variables such as those responsible for damaging ecosystems (see Bailey et al., 2004).

Conclusions and recommendations

In conclusion, it is strongly recommended that habitat quality bioassessments for different ecosystems within farmland be made:

1) Using a data-driven approach that uses a modelling framework;

2) Deploying pre-existing data (e.g., collaboration and coordination with Environment Canada supported programs, non-government supported programs);
3) An immediate goal should be to carry out a pilot study and collect data along an environmental stressor gradient for any of the above taxa to demonstrate the techniques used in the case study.
4) Over the longer term, a broadly implemented multi-species monitoring program should be set up in agricultural landscapes. The latter should include birds, small mammals, herptiles, arthropods (terrestrial and aquatic), vascular plants, lichens, bryophytes and fungi (perhaps using a similar design and set up to the ABMI).
5) It is recommended that a multivariate multi-species modelling procedure be used on both existing and future data that can include both species and environmental (habitat and agricultural stressor) variables within the same model. This will allow the main drivers of species composition and abundance to be identified; avoid having to subjectively set standards in separately measured indicators; allow sites of differing habitat quality to be identified and provide a mechanism by which to measure improvements in site quality brought about by adaptive management (by examining changes over time in the position of sites in ordination space).
6) Finally it is recommended that a combination of the techniques be used to substantiate the findings from a multi-species multivariate modelling approach; these could be compared using an objective method such as impacted data.

6 ACKNOWLEDGEMENTS

We would like to thank Environment Canada, National Agri-Environmental Standards Initiative (NAESI) for funding the preparation of this report, and particularly Cathy Nielsen for her help and guidance at all stages of the project. We are particularly grateful to Doug Forsyth (Canadian Wildlife Service, Prairie and Northern Region, Saskatoon) for kindly letting us use his data on plants for the case study and to Trefor Reynoldson who helped with and ran the analyses for the RCA approach (case study approach 1). We would like to extend special thanks to Titus Seilheimer, Melanie Croft and Pat Chow-Fraser who helped with the WMI index, and Rod Brook who helped compute the index. We would also like to thank Karen Blocksom (US EPA, Ohio), Stephanie Strachan (Stephanie Sylvestre, Environment Canada, British Columbia), Lee Grapentine (Environment Canada, Burlington), Tim Pascoe (Environment Canada, Burlington), Chuck Hawkins (University of Utah), and Trefor Reynoldson (Acadia University) for comments and ideas on Table 1. Stephanie Strachan was particularly helpful in ordering the questions and Karen Blocksom provided extensive comments on the multimetric column. For sending manuscripts in preparation and pdfs of published articles we would like to thank Pat Chow-Fraser (McMaster University), Ian Dodkins (University of Dublin), Chuck Hawkins (University of Utah), Vanessa Lougheed (University of Austin, Texas), Trefor Reynoldson (Acadia University) and Adam Yates (University of Waterloo). For help with software and discussions we would like to thank Lee Belbin (Blatant Fabrications Ltd.), Peter Minchin (University of Illinois), Jennie Pearce (Pearce & Associates Ecological Research), Grant Wardell-Johnson (Curtin University of Technology, Perth, Australia), Phil Rundel (University of California), Dean Urban (Duke University) and Anhua Wei (National Resources Canada). Finally we would like to thank Cathy Nielsen and Trefor Reynoldson for their extremely valuable comments on the report and Cheryl

Widdifield for help with formatting.

7 REFERENCES

Albert, D.A. and L.D. Minc. 2004. Plants as regional indicators of Great Lakes coastal wetland health. Aquatic Ecosystems Health Management. 7: 233-247.

AMEC Earth & Environmental. 2008. Development of Habitat and Biotic Indices for Grasslands in Agricultural Regions. National Agri-Environmental Standards Initiative Technical Series Report No. 4-5. 137 p.

Andrén, H. 1994. Effects of habitat fragmentation on birds and mammals in landscapes with different proportions of suitable habitat: a review. Oikos. 71: 355-366.

Andreasen, J.K., R.V. O'Neill, R. Noss, and N.C. Slosser. 2001. Considerations for the development of a terrestrial index of ecological integrity. Ecological Indicators. 1: 21–35.

Austin, M.P. 2002. Spatial prediction of species distribution: An interface between ecological theory and statistical modeling. Ecological Modeling. 157: 101-118.

Bailey, R.C., R.H. Norris, and T.B. Reynoldson. 2004. Bioassessment of Freshwater Ecosystems: Using the Reference Condition Approach. Kluwer Academic Publishers, Boston, Massachusetts.

Baldwin, D., E. Neave, C. Nielsen, J. Pearce, M. Rowsell, F. Schnekenburger, J. Tews, L. Tischendorf, L. Venier, and R. Waldick. 2006. Eastern Ontario Pilot - Interim Progress Report. National Agri-Environmental Standards Initiative Technical Series Report No. 2-28. 828 p.

Ball, J., L. Donnelly, P. Erlanger, R. Evans, A. Kollmorgen, B. Neal, and M. Shirley. 2001. Inland waters, Australia State of the Environment Report 2001 (Theme report), Commonwealth Scientific and Industrial Research Organizaion Publishing on behalf of the Department of the Environment and Heritage Canberra. http://www.ea.gov.au/soe/2001/index

Belbin, L. 2008. PATN vs 3.0. Blatant Fabrications Pty. Ltd, Carlton, Tasmania, Australia 7173.

Billeter, R., J. Liira, D. Bailey, R. Bugter, P. Arens, I. Augenstein, S. Aviron, J. Baudry, R. Bukachek, F. Burel, M. Cerny, G. De Blust, R. De Cock, T. Diekötter, H. Dietz, J. Dirksen, C. Dormann, W. Durka, M. Frenzel, R. Hamersky, F. Hendrickx, F. Herzog, S. Klotz, B. Koolstra, A. Lausch, D. Le Coeur, J.P. Maelfait, P. Opdam, M. Roubalova, A. Schermann, N. Schermann, T. Schmidt, O. Schweiger, M.J.M. Smulders, M. Speelmans, P. Simova, J. Verboom, W.K.R.E. van Wingerden, M. Zobel, and P.J. Edwards. 2007. Indicators for biodiversity in agricultural landscapes: a pan-European study. Journal of Applied Ecology. 45 (1): 141-150

Boutin, C., K.E. Freemark, and D.A. Kirk. 1999a. Farmland birds in southern Ontario: field use, activity patterns and vulnerability to pesticide use. Agriculture, Ecosystems and Environment. 72: 239-254.

Boutin, C., K.E. Freemark, and D.A. Kirk. 1999b. Spatial and temporal patterns of bird use of farmland in southern Ontario. Canadian Field-Naturalist. 113: 430-460.

Boyce, R.L. and P.C. Ellison. 2001. Choosing the best similarity index when performing fuzzy set ordination on binary data. Journal of Vegetation Science. 12: 711-720.

Browder, S.F., D.H. Johnson, and I.J. Ball. 2002. Assemblages of breeding birds as indicators of grassland condition. Ecological Indicators. 2: 257–270

Bryce, S.A., R.M. Hughes, and P.R. Kaufman. 2002. Development of a bird integrity index: Using bird assemblages as indicators of riparian condition. Environmental Management. 30: 294-310.

Cao, Y., C.P. Hawkins, and J. Olson. 2007. Modeling natural environmental gradients improves the accuracy and precision of diatom-based indicators. North American Benthological Society. 26: 566-585.

Chow-Fraser, P. 2006. Development of the Wetland Quality Index (WQI) to assess effects of basin-wide land-use alteration on coastal marshes of the Laurentian Great Lakes. In: T.P.

Simon and P. M. Stewart (eds.). Coastal wetlands of the Laurentian Great Lakes: Health, habitat and indicators. Authorhouse. Bloomington, IN. pp. 137–166.

Clarke, R. 2000. Uncertainty in estimates of biological quality based on RIVPACS. In: J.F. Wright, D.W. Sutcliffe, and M.T. Furse. (eds.). Assessing the biological quality of fresh waters: RIVPACS and other techniques. Freshwater Biological Association, Ambleside, UK. pp. 39-54.

Clarke, K.R. and R.N. Gorley. 2001. Primer v5: User manual/tutorial. Primer-E, Plymouth, UK.

Croft, M.V. and P. Chow-Fraser. 2007. Use and development of the Wetland Macrophyte Index to detect water quality impairment in fish habitat of Great Lakes coastal marshes. Journal of Great Lakes Research. 33 (SI3): 172-197.

Davies, P.E. (ed.). 1994. Monitoring River Health Initiative. River Bioassessment Manual. National River Processes and Management Program. Freshwater Systems, Tasmania.

Del Degan, Massé et Associés Inc. 2008. Élaboration de l’indice de l’habitat et de l’indice biotique pour les terres humides des régions agricoles. Série technique de l'Initiative nationale d'élaboration de normes agroenvironnementales rapport no 4-5. 136 p.

Dufrene, M. and P. Legendre. 1997. Species assemblages and indicator species: the need for a flexible asymmetrical approach. Ecological Monographs. 67: 345-366.

Dunford, W. and K.E. Freemark. 2005. Matrix matters: Effects of surrounding land uses on forest birds near Ottawa, Canada. Landscape Ecology. 20: 497-511.

Elith, J. and M.A. Burgman. 2003. Habitat models for population viability analysis. In: A. Brigham and M. W. Schwartz (eds.). Population viability in plants. Springer-Verlag, Heidelberg, Germany. pp. 203–235.

Faith, D.P., S. Ferrier, and K.J. Williams. 2008. Getting biodiversity intactness indices right: ensuring that ‘biodiversity’ reflects ‘diversity’. Global Change Biology. 14: 207-217.

Fahrig, L. 1998. When does fragmentation of breeding habitat affect population survival?

Ecological Modelling. 105: 273-292.

Flather, C.H. and M. Bevers. 2002. Patchy reaction-diffusion and population abundance: the relative importance of habitat amount and arrangement. American Naturalist. 159: 40-56.

Franken, R., T. Wellicome, B. Dale, M. Schmoll, D. Kirk, R. Quinlan, and R. Scobie, 2003. Identifying species-habitat linkages for priority landbirds, shorebirds and species at risk: Final report. Report to Alberta North American Waterfowl Management Plan. Biodiversity Fund, Prairie and Northern Region, Canadian Wildlife Service, Environment Canada, Edmonton, Alberta.

Freemark, K. and B. Collins, 1992. Landscape ecology of birds breeding in temperate forest fragments. In: J.M. Hagan and D.W. Johnson (eds.). Ecology and Conservation of Neotropical Migrant Landbirds. Smithsonian Institution Press, Washington, D.C. pp. 443-454.

Freemark, K.E. and D.A. Kirk. 2001. Birds breeding on organic and nonorganic farms in Ontario: Partitioning effects of habitat and practices on species composition and abundance. Biological Conservation. 101: 337-350.

Fuller, R.J., L.R. Norton, R.E. Febee, P.J. Johnson, D.E. Chamberlain, A.C. Joys, F. Mathews, R.C. Stuart, M.C. Townsend, W.J. Manley, M.S. Wolfe, D.W. Macdonald, and L.G. Firbank. 2005. Benefits of organic farming to biodiversity vary among taxa. Biology Letters. 1(4): 431–434.

Griffith, D.A. and P.R. Peres-Neto. 2006. Spatial modeling in ecology: The flexibility of eigenfunction spatial analysis. Ecology. 87: 2603-2613.

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

Herbers, J. and J. Schieck. 2004. A review of national and international biodiversity indices. Appendix 2, Alberta Biodiversity Monitoring Program, Edmonton, Alberta. Unpublished report available at http://www.abmp.arc.ab.ca

IRIS Environmental Systems Inc. 2008. Indicators of Riparian Biodiversity within the Agricultural Areas of Canada. National Agri-Environmental Standards Initiative Technical Series Report No. 4-3. 196 p.

Karr, J.R. and E.W. Chu. 1999. Restoring life in running waters: better biological monitoring. Island Press, Washington, DC.

Kerans, B.L. and J.R. Karr. 1994. A benthic index of biotic integrity (B-IBI) for rivers of the Tennessee Valley. Ecological Applications. 4: 768-785.

King, R.S., C.J. Richardson, D.L. Urban, and E.A. Romanowicz. 2004. Spatial dependency of vegetation-environment linkages in an anthropogenically influenced wetland ecosystem. Ecosystems. 7: 75-97.

Kirk, D.A., M. Csizy, C. Francis, and J.D. McCracken. 2001a. Habitat associations of marsh-nesting birds in the Great Lakes Region: Implications for local conservation and management. Report to Wildlife Habitat Canada, Bird Studies Canada.

Kirk, D.A., C. Boutin, and K.E. Freemark. 2001b. A multivariate analysis of bird species composition and abundance between crop types and seasons in southern Ontario, Canada. ÉcoScience. 8: 173-184.

Kleijn, D. and L.A.C. van der Voort. 1997. Conservation headlands for rare arable weeds: the effects of fertilizer application and light penetration on plant growth. Biological Conservation. 81: 57-67.

Kleijn, D. and I.J. Spoejing. 1997. Field boundary vegetation and the effects of agrochemical drift: botanical change caused by low levels of herbicide and fertilizer. Journal of Applied Ecology. 34: 1413-1425.

Kleijn, D. and M. Verbeek. 2000. Factors affecting the species composition of arable field boundary vegetation. Journal of Applied Ecology. 37: 256-266.

Klemm, D.J., K.A. Blocksom, W.T. Thoeny, F.A. Fulk, A.T. Herlihy, P.R. Kaufmann, and S.M. Cormier. 2002. Methods development and use of macroinvertebrates as indicators of ecological conditions for streams in the Mid-Atlantic Highlands Region. Environmental Monitoring and Assessment. 78: 169–212.

Kremen, C. 1992. Assessing the indicator properties of species assemblages for natural areas monitoring. Ecological Applications. 2: 203–217.

Leeson, J.Y., J.W. Sheard, and A.G. Thomas. 2000. Weed communities associated with arable Saskatchewan farm management systems. Canadian Journal of Plant Science. 80: 177-185.

Legendre, P. 1993. Spatial autocorrelation: Trouble or new paradign? Ecology. 74: 1659-1673.

Legendre, P. and L. Legendre. 1998. Numerical ecology. Second edition. Elsevier, Amsterdam.

Lêps, J. and P. Smilauer. 2007. Multivariate analysis of ecological data using CANOCO. Cambridge University Press, New York, NY, USA.

Linke, S., R.H. Norris, D. Faith, and D. Stockwell. 2005. ANNA: A new prediction method for bioassessment programs. Freshwater Biology. 50 (1): 147-158.

Lougheed, V.L. and P. Chow-Fraser. 2002. Development and use of a zooplankton index of wetland quality in the Laurentian Great Lakes Basin. Ecological Applications. 12: 474-486.

Lougheed, V.L., C.A. Parker, and R.J. Stevenson. 2007. Using non-linear responses of multiple taxonomic groups to establish criteria indicative of wetland biological condition. Wetlands. 27: 96-109.

Manley, P.N., W.J. Zielinski, M.D. Schledsinger, and S.R. Mori. 2004. Evaluation of a multiple-species approach to monitoring species at the ecoregional scale. Ecological Applications 14: 296-310. Online at: http://www.fs.fed.us/research/Monitoring/multiple_species.pdf

Milko, R. 1998. Directive pour les évaluations environnementales relatives aux milieux humides, Direction de la protection de la biodiversité, Service canadien de la faune, Environnement Canada.

Mineau, P., C.M. Downes, D.A. Kirk, E. Bayne, and M. Csizy. 2005. Patterns of bird species abundance in relation to granular insecticide use in the Canadian prairies. EcoScience. 12: 267-278.

Moreby, S.J. and S.E. Southway. 1999. Influence of autumn applied herbicides in summer and autumn food available to birds in winter wheat fields in southern England. Agriculture, Ecosystems & Environment. 72(3): 285-297.

Nebbia, A.J. and S.M. Zalba. 2007. Designing nature reserves: traditional criteria may act as misleading indicators of quality. Biodiversity and Conservation. 16: 223-233.

Nielsen, S.E., E.M. Bayne, J. Schieck, J. Herbers, and S. Boutin. 2007. A new method to estimate species and biodiversity intactness using empirically derived reference conditions. Biological Conservation. 137: 403-414.

Palmer, M. 2008. Ordination web page. Accessed March 2008. URL http://ordination.okstate.edu/

R-Core Development Team. 2007. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. http://www.R-project.org.

Redford, K.H. and B.D. Richter. 1999. Conservation of biodiversity in a world of use. Conservation Biology. 13: 1246-1256.

Reynoldson, T.B. and J.F. Wright. 2000. The reference condition: problems and solutions. 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. Chapter 20. pp. 293-303.

Reynoldson, T.B., K.E. Day, and T. Pascoe. 2000. The development of the BEAST: a predictive approach for assessing sediment quality in the North American Great Lakes. 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. Chapter 11. pp. 165-180.

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

Reynoldson, T.B., M. Bombardier, D. Donald, H.J. O'Neill, D.M. Rosenberg, H. Shear, T. Tuominen, and H. Vaughan. 1999. Strategy for a Canadian Aquatic Biomonitoring Network. National Water Research Institute Contribution No. 99-248.

SAS Institute. 2000. SAS/STAT Users Guide: Statistics. Release 8, SAS Institute, Inc. Cary, North Carolina.

Scholes, R.J. and R. Biggs. 2005. A biodiversity intactness index. Nature. 434: 45–49.

Seilheimer, T.S. and P. Chow-Fraser. 2006. Development and validation of the WFI to assess the quality of coastal wetlands in the Laurentian Great Lakes. Canadian Journal of Fisheries and Aquatic Sciences. 63: 354-366.

Shimwell, D.W. 1971. Description and classification of vegetation. Billing and Sons Ltd., London, UK.

Shutler, D., A. Mullie, and R.G. Clark. 2000. Bird assemblages of prairie uplands and wetlands in relation to farming practices in Saskatchewan. Conservation Biology. 14: 1441-1451.

Simpson, J. and R.H. Norris. 2000. Biological assessment of water quality: development of AUSRIVAS models and outputs. 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 and Environment Agency, Ableside, Cumbria, U.K. pp. 125–142.

Smith, M.J. W.R. Kay, D.H.D. Edward, P.J. Papas, K. St. J. Richardson, J.C. Simpson, A.M.

Pinder, D.J. Cale, P.H.J. Horwitz, J.A. Davis, F.H. Yung, R.H. Norris, and S.A. Halse. 1999. AusRivAS: using macroinvertebrates to assess ecological condition of rivers in Western Australia. Freshwater Biology. 41(2): 269–282.

Summit Environmental Consultants Ltd. 2008. Habitat and Biotic Indices for Forests and Woodlands in Agricultural Regions. National Agri-Environmental Standards Initiative Technical Series Report No. 4-4. 155 p.

Sylvestre, S., M. Fluegel, and T. Tuominen. 2005. Benthic invertebrate assessment of streams in the Georgia Basin using the reference condition approach: Expansion of the Fraser River invertebrate Monitoring Program 1982-2002. Environment Canada, Environmental Conservation Branch, Vancouver.

SPSS Inc. 1999. Systat 9.0: Statistical Software, Chicago, USA.

Taft, J.B., C. Hauser, and K.R. Robertson. 2006. Estimating floristic integrity in tallgrass prairie. Biological Conservation. 131: 42-51.

ter Braak, C.J.F. 1995. Ordination. In: R.H.G. Jongman, C.J.F. ter Braak, and O.F.R. van Tongeren (eds.). Data analysis in community and landscape ecology. Cambridge University Press, Cambridge, UK. pp. 91-173.

ter Braak, C.J.F. and P. Smilauer. 2002. CANOCO reference manual and CanoDraw for Windows User's guide: Software for canonical community ordination (version 4.5). Microcomputer Power, Ithaca, New York, USA.

Urban, D.L. 2006. A modeling framework for restoration ecology. In: D.A. Falk, M.A. Palmer, and J.B. Zedler (eds.) Foundations of restoration ecology. Island Press, Washington DC, USA. pp. 238-256.

Villard, M-A, M.K. Trzcinski, and G. Merriam. 1999. Fragmentation effects on forest birds: relative influence of woodland cover and configuration on landscape occupancy. Conservation Biology. 13: 774-783.

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

Wright, J.F., M.T. Furse, and P.D. Armitage. 1993. RIVPACS- a technique for evaluating the biological quality of rivers in the U.K. European Water Pollution Control. 3: 15–25.

8 APPENDICES

APPENDIX A: Potential candidate research projects for case study

Study	Lead organization	Province	Taxa	Habitat/ecosystems	Year	Comparison	Lead contact/researcher	Status
1) Crop use by birds	Canadian Wildlife Service Environment Canada	Ontario	Birds	Mixed small grain, rowcrop, pasture Mixedwood Plains	1988 1989	Apple orchards, corn	Drs. C. Boutin and K. Freemark Lindsay	Kirk et al., 2001a; Boutin et al., 1999a,b
2) Habitat and farming practices on birds of organic and conventional farms	Canadian Wildlife Service Environment Canada	Ontario	Birds	Mixed small grain, rowcrop, pasture Mixedwood Plains	1988	Organic vs conventional	Dr. K. Freemark Lindsay	Freemark and Kirk, 2001; Freemark-Lindsay et al. in preparation
3) Habitat and farming practices and birds of organic and conventional farms	Canadian Wildlife Service Environment Canada	Saskatchewan	Birds	Prairie grassland, cropland Prairie Potholes	1990	Organic vs conventional	Dr. K. Freemark Lindsay	Freemark Lindsay and Kirk in review; Freemark Lindsay et al. in preparation

Study	Lead organization	Province	Taxa	Habitat/ecosystems	Year	Comparison	Lead contact/researcher	Status
4) Bird communities in watersheds affected by different levels of agricultural intensification	Canadian Wildlife Service Environment Canada	Iowa	Birds	Rowcrops, pasture, woodland	1996	Watershed comparison	Dr. K. Freemark Lindsay	Freemark Lindsay et al. in preparation
5) Habitat and farming practices on invertebrates of organic and conventional farms (pitfall, sticky traps, sweep nets)	Canadian Wildlife Service Environment Canada	Ontario	Terrestrial arthropods	Hedges, field margins and field interior in mixed small grain, rowcrop, pasture Mixedwood Plains	1999	Organic vs conventional	Pamela Martin and Dr. C. Boutin	Martin et al. in preparation (2 mss pitfall traps and sweep nets/sticky traps – not submitted to journals) Boutin et al. in review
6) Influence of farm management and moisture regime on plant communities of wetland field margins	Canadian Wildlife Service Environment Canada	Saskatchewan	Plants	Wetland edges Prairie Potholes	1996	Wild, organic, minimum tillage, conventional	Dr. D. Forsyth	Forsyth et al. in preparation

Study	Lead organization	Province	Taxa	Habitat/ecosystems	Year	Comparison	Lead contact/researcher	Status
7) Influence of farm management on terrestrial invertebrates of wetland field margins	Canadian Wildlife Service Environment Canada	Saskatchewan	Terrestrial invertebrates (Carabid beetles, spiders)	Wetland edges Prairie Potholes	1996	Wild, organic, minimum tillage, conventional	Dr. D. Forsyth	Forsyth et al. in preparation
8) Effects of tillage on invertebrates	Canadian Wildlife Service Environment Canada	Saskatchewan	Terrestrial invertebrates (Carabid beetles, spiders)	Mixed arable land Prairie Potholes	1996	Compared different tillage systems, organic vs conventional	Dr. D. Forsyth	Forsyth et al. in preparation
9) Influence of farm management on aquatic invertebrates of wetlands	Canadian Wildlife Service Environment Canada	Saskatchewan	Terrestrial invertebrates (Carabid beetles, spiders)	Wetland edges Prairie Potholes	1996	Wild, organic, minimum tillage, conventional	Dr. D. Donald	Donald et al. (submitted and rejected by Wetlands) – plan to revamp
10)	Canadian Wildlife Service Environment Canada	Ontario	Birds	Mixedwood Plains	1982, 2001?	Compared bird species composition and abundance between blocks	Dr K. Freemark Lindsay and W. Dunford	Freemark and Collins, 1992; Dunford and Freemark, 2005

APPENDIX B: Acronyms

AUSRIVAS **Aus**tralian **Riv**er **A**ssessment **S**ystem
BEAST Benthic Assessment of Sediment
B-IBI Benthic index of biotic integrity
CA Correspondence analysis (unconstrained ordination, weighted averaging, unimodal model)
CABIN **C**anadian **A**quatic **Bi**omonitoring **N**etwork
CCA canonical correspondence analysis (constrained ordination, weighted averaging, unimodal model)
CVA canonical variates analysis (constrained CCA ordination form of linear discriminant analysis)
DCA detrended correspondence analysis (unconstrained ordination, weighted averaging, unimodal model, CA corrected for arch effect)
db-RDA distance-based redundancy analysis (constrained ordination, weighted averaging, linear model)
EUWFD European Union Water Framework Directive
FSO fuzzy set ordination (constrained form of non-metric multidimensional scaling nMDS)
IBI index of biotic integrity
IHI index of habitat integrity
GLM Generalized Linear Model
hMDS hybrid metric multi-dimensional scaling (unconstrained distance-based ordination, seeks best solution from non-metric/metric)
nMDS non-metric multi-dimensional scaling (unconstrained distance-based ordination, different similarity measures – Bray Curtis is usually considered best distance measure)
PCA Principal components analysis (weighted-averaging, linear model)
PCC Principle axis correlation (vector analysis – multiple regression of environmental variables on ordination axes derived from nMDS)
RCA Reference Condition Approach
RDA redundancy analysis (constrained form of PCA, weighted averaging, linear model)
RIVPACS **R**iver **I**n**v**ertebrate **P**rediction **a**nd **C**lassification **S**ystem
TWINSPAN two-way-indicator-species-analysis (divisive cluster analysis)
WMI Wetland Macrophyte Index
WPI Wetland Plant Index
WQI Wetland Quality Index