

A bioassessment of lakes in the Athabasca Oil Sands Region, Alberta, using benthic macroinvertebrates

Brent G. PARSONS*, Shaun A. WATMOUGH¹, Peter J. DILLON¹ and Keith M. SOMERS^{1,2}

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

¹Environmental and Resource Studies, Trent University, Peterborough, ON K9J 7B8, Canada

²Dorset Environmental Science Centre, Ontario Ministry of the Environment, Dorset, ON P0A 1E0, Canada

Present address: 1239 Winhara Road, Gravenhurst, ON P1P 1R1, Canada

*e-mail corresponding author: brentparsons@trentu.ca

ABSTRACT

Emissions of sulphur oxides, nitrogen oxides and other pollutants have increased in the Athabasca Oil Sands Region (AOSR) in Alberta, Canada. Atmospheric pollutants impact aquatic communities through a number of processes, but due to a lack of regional monitoring programs potential biological impacts have not been assessed. In this study, a bioassessment was conducted using approaches borrowed from a variety of protocols to establish a baseline dataset, determine appropriate methodologies, and to assess the current impact of emissions on benthic macroinvertebrate (BMI) communities in the AOSR. As a result, 32 lakes, including 5 test lakes located in a modelled high deposition region, were sampled for water chemistry and BMI. The Reference Condition Approach (RCA) was used because a baseline dataset does not exist and data were evaluated using three separate statistical techniques. All of the statistical methods used: One Sample T-Tests, Multivariate Analysis of Variance (MANOVA) and Test Site Analysis (TSA), showed that BMI assemblages in test lakes differed from BMI assemblages in reference lakes. Traditional statistics classified all 5 test lakes as "significantly impaired" whereas TSA identified 3 of the 5 test lakes as only potentially impaired and 2 lakes were in "reference condition". The variability in lake attributes present challenges in interpreting BMI data and establishing an accurate biomonitoring program in the AOSR which need to be addressed in future assessment studies.

Key words: bioassessment, benthic macroinvertebrates, oil sands, Test Site Analysis, Reference Condition Approach, Canada

1. INTRODUCTION

The oil sands in Alberta represent the second largest recoverable oil deposit in the world with approximately 175 billion barrels of recoverable oil (National Energy Board 2006; Reynolds 2005). By 2015, it is expected that 3 million barrels will be produced each day (National Energy Board 2006; Reynolds 2005). Recent mining operations have led to an increase in emissions of sulphur oxides (SO_x), nitrogen oxides (NO_x), heavy metals and other pollutants, and emissions are predicted to remain high for several decades (Golder 2003). Elevated levels of SO_x and NO_x may lead to the acidification of soils and surface waters in poorly buffered ecosystems. A general recovery from acidification is occurring in many other regions but these recoveries are finely balanced and additional acid inputs have the potential to disrupt the recovery process (Kowalick *et al.* 2007), making acidifying emissions both a local and global problem. Eutrophication may occur due to elevated nitrogen loading to nitrogen-limited lakes. Other pollutants such as metals (e.g., mercury, lead) are directly emitted during oil production and are transferred to waterbodies via atmospheric deposition (Moiseenko *et al.* 1995). These alterations in water chemistry cause biological damage and depending on the severity of the chemical change the entire freshwater

food web may be affected both directly and indirectly (Doka *et al.* 2003).

Mining activities in the Athabasca Oil Sands Region (AOSR) began 40 years ago but production increased dramatically when oil prices spiked in 2004 (Reynolds 2005). Since output is anticipated to grow steadily due to improving extraction techniques and increasing oil prices (Reynolds 2005; National Energy Board 2006), there are concerns that the resulting emissions may further impact the surrounding environment. Technological advances have reduced the emissions released per barrel of oil produced but the magnitude of production overshadows any such improvements as approximately 100,000 tonnes of SO₂ were emitted in 2006 (Golder 2004). Future emission limits are uncertain and will depend to a large part on estimates of the level of deposition that is likely to cause adverse effects on biota (Environment Canada 2004).

With expanding activities in the AOSR it has become increasingly important to monitor terrestrial and aquatic ecosystems for signs of adverse impacts. Aquatic bioassessments have previously been used to assess the impact of industrial emissions because they are a strong indicator of lake condition (Harman 1997). In particular, water chemistry fluctuations are often short term whereas aquatic bioassessments generally provide a more robust evaluation of water quality

(Harman 1997). For example, pronounced changes in lake chemistry during spring melt in the AOSR will not be detected during routine chemical sampling, which for the most part relies on a single chemical measurement, due in part to the remote location of the study area. Furthermore, physical and chemical characteristics of lentic environments in the AOSR are such that many widely used biomonitoring organisms such as crayfish, fish and nesting birds are often absent. Benthic macroinvertebrates (BMI) were present in all of the study lakes and BMI are good indicators of acidification, eutrophication and trace metal impacts (Moiseenko 2003; Resh & Rosenberg 1993). Benthic macroinvertebrates are also the most widely used organisms in the bioassessment of freshwater ecosystems (Bailey *et al.* 2004).

Standard methods for collecting BMI exist, but the interpretation of such data can be complex (Norris 1995), particularly in areas such as the AOSR where multiple pollutants may exist and physical and chemical lake characteristics are highly variable. Benthic macroinvertebrate data are ideally compared against a historical dataset in order to detect biological effects. Unfortunately, previous bioassessment studies in the AOSR have focussed on a small number of lakes with limited spatial distribution; a regional aquatic benthic monitoring program does not exist in the region. Accordingly, a baseline dataset was established in this study through the implementation of the Reference Condition Approach (RCA) (Reynoldson *et al.* 1997; Bailey *et al.* 1998). The RCA compares individual test sites to minimally exposed reference sites and allows for assessments to be made in the absence of historical data (Bailey *et al.* 2004). The test lake evaluations were completed using 3 different methods: One-Sample *T*-Tests, Multivariate Analysis of Variance (MANOVA) and Test Site Analysis (TSA; Bowman & Somers 2006). Multiple statistical procedures were used because a standard biomonitoring statistical procedure does not exist in Canada and we wanted to determine the most appropriate procedure for bioassessment in the AOSR. Test Site Analysis is a relatively new approach which has produced fewer false-positive and false-negative errors than other popular assessment methods such as the Benthic Assessment of Sediment (BEAST) and the multimetric approach (Bowman & Somers 2006). Test Site Analysis also identifies the metrics which are most important in distinguishing a significantly impaired test site which is an important consideration in the AOSR where multiple pollutants exist (Bowman & Somers 2006). Parsons *et al.* (2010, this issue) also determined the most appropriate metrics to monitor the impacts of atmospheric pollution in the AOSR by establishing relationships between lake water chemistry and BMI composition.

The primary objectives of this study were to establish a baseline dataset and suitable methodology for

future monitoring and assessment of effects from anticipated atmospheric emissions in the AOSR by comparing BMI communities in test lakes located in areas of high modelled deposition of SO_x (and assumed other pollutants) with BMI communities in reference lakes that are similar in available physical and chemical properties, but located outside the high deposition region.

2. METHODS

2.1. Study area

The AOSR is located approximately 400 kilometres northeast of Edmonton, Alberta and is centred around Fort McMurray, Alberta (Fig. 1). The area is part of the expansive Boreal Plains and is characterized by fens and bogs, which are dominated by black spruce (*Picea mariana*) in the lowland areas, while trembling aspen (*Populus tremuloides*) and jack pine (*Pinus banksiana*) forests typify the upland regions. The lakes in the area are generally small, naturally eutrophic and high in dissolved organic carbon (DOC).

2.2. Study design

Emissions of SO_x and NO_x are a major concern in the AOSR, and so this study focussed on lakes that are relatively sensitive to acidification (<600 µeq L⁻¹ of base cations), accessible via floatplane (>6 ha), and small enough (<1000 ha) to be characterized by 5 sampling sites (e.g., Reid *et al.* 1995). Notably, prior to sampling there was no information available on substrate or macrophytes in the lakes. There are limited deposition data for the AOSR so all test lakes were selected from an area relatively high in sulphur deposition based on A United Regional Air-quality Monitoring Program (AURAMS) (Moran *et al.* 2007) (Fig. 1) and we assumed that other pollutants follow a similar deposition pattern. Reference lakes were located well outside this zone of high deposition to account for uncertainty in deposition estimates, but close enough to the test lakes to retain environmental similarities and ensure accessibility (Fig. 1). The region where the reference lakes were located had to be sufficiently variable to ensure that every test lake had reference lakes that were similar physically and chemically. According to these parameters plus financial and logistic constraints, 5 test lakes and 27 reference lakes were selected for sampling from a dataset of 471 lakes (Fig. 1).

2.3. Benthic macroinvertebrate sampling

The bioassessment protocol of David *et al.* (1998) was used for BMI sampling. Between August 15 and August 28 2006, 5 sites in the littoral area of each lake were sampled using a travelling kick-and-sweep method. The kick-and-sweep technique was used because it collected a greater abundance of BMI than an Ekman Grab sampler in the AOSR (Parsons, unpubl.). The littoral area at each site was sampled along several

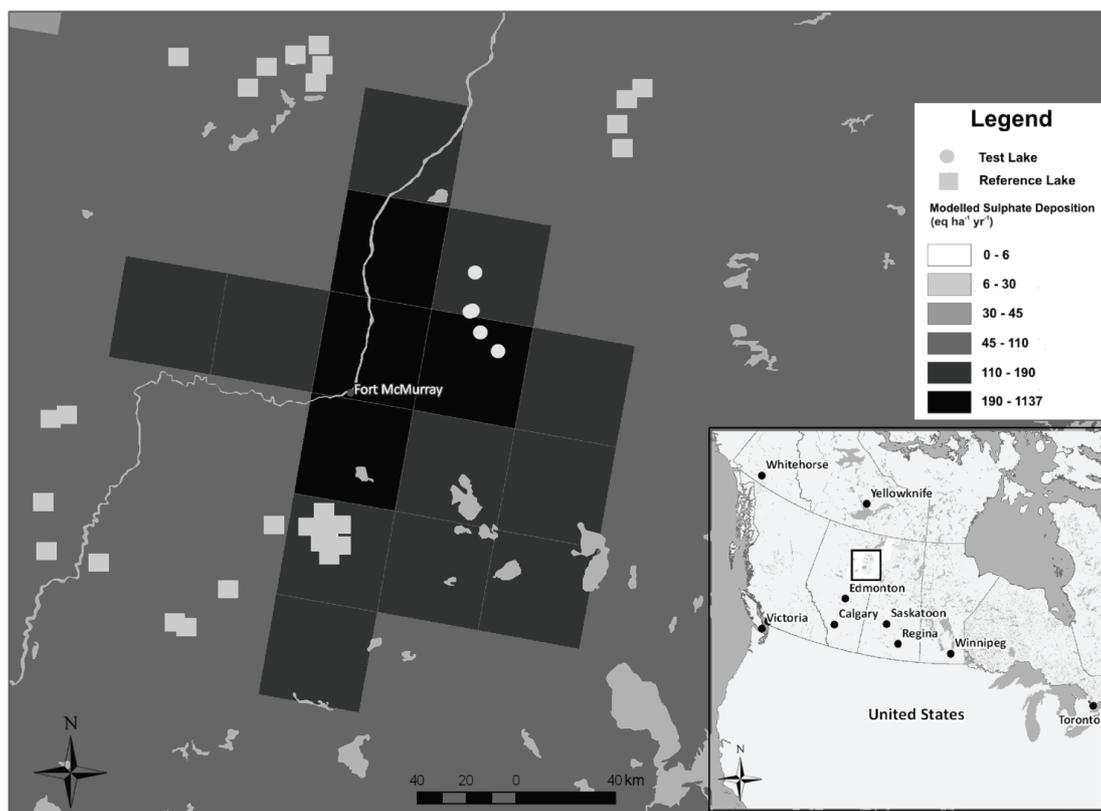


Fig. 1. Map of the study area showing modelled SO_4 deposition provided by A United Regional Air-quality Modelling System (AURAMS) (Moran *et al.* 2007) from the 2002 annual Run 3. Emissions data from the 2002 Canadian and United States national criteria air-contaminant emission inventories were utilized in the model run. The location of both test and reference lakes are also shown. Note: two test lakes were so close together that they appear as one lake.

transects extending perpendicular from shore to a maximum depth of 1 m. One person kicked and dislodged material as the other followed and collected the BMI and associated sediments with a 500 μm mesh D-net. Each site was sampled for a fixed period of 10 minutes.

The resultant sample was washed through two sieves; one contained (1 cm^2) hardware cloth to sieve out large organic debris and the other had 500 μm mesh to rinse out fine organic material, clay and silt. Glass jars (1 L) were filled with the sieved sample and preserved with 70% isopropyl alcohol. At Trent University the preserved samples were picked following David *et al.* (1998). Each sample was randomly sub-sampled and picked until 100 BMI were removed from the material or a three hour time limit was reached. A three hour time limit was implemented because BMI were sparse in many lakes and the study was based on a qualitative assessment of BMI communities so that a greater number of sites could be analyzed and spatial assessment made possible. Methods have been utilized in previous bioassessments, including time limits and sub-sampling, to make the studies logistically feasible (Rosenberg *et al.* 2003). The BMI, except for *Oligochaeta* and *Hydracarina*, were identified to the taxonomic level of family (Appendix 1). Five percent of the samples (i.e., $n = 8$,

containing 564 BMI) were re-identified by Ontario Ministry of the Environment personnel and accuracy was found to be >92%. Bailey *et al.* (2001) found that family level identification is appropriate for multivariate analysis or index calculation, both of which were completed in this study, while others have found that family identification does not provide a strikingly different description of BMI community than genus or species (Bowman & Bailey 1997).

The five 100-count BMI site abundances at each lake were combined into a single 500-count BMI lake score because the 5 sites that were selected at each lake were in proportion to nearshore habitat in an attempt to characterize each lake (David *et al.* 1998). These results allowed atmospheric pollution to be more appropriately analyzed than if each site was examined independently because deposition is a non-point source pollutant and any subsequent effects should alter BMI throughout the entire lake and not at a single sampling location.

2.4. Lake water sampling

Duplicate lake water samples were collected from the middle of each lake at a 10 cm depth. Samples were filtered through 75 μm mesh and stored in 500 mL polyethylene jars and borosilicate glass tubes. Water samples were stored on ice after collection and were

kept cool until arriving at the laboratory for analysis within two weeks after collection. Water samples were analysed for alkalinity using a PC-Titration Plus system. Nitrate (NO₃), ammonium (NH₄), and total Kjeldahl nitrogen (TKN) were determined by colourimetric measurement using a Pulse AutoAnalyzer System consisting of a reaction cell, spectrometer and data analyzer. Samples for dissolved organic carbon (DOC), dissolved carbon (DC), inorganic carbon (IC), plus anions and cations were filtered through 0.45 µm prior to analysis. Dissolved organic carbon, DC and IC were determined using a Shimadzu TOC-VCPH total organic carbon analyzer. Sulphate and chloride (Cl) were analysed using a Dionex DX-600 ion chromatograph, while sodium (Na), magnesium (Mg), potassium (K), and calcium (Ca) concentrations were determined by direct aspiration flame atomic absorption spectrometry using a Varian 240FS.

At each site, substrate assessments were conducted by manually investigating the substrate at 5 points along the sampling transect(s) and estimating the proportion of the following substrate categories (organic, organic debris, clay, silt, sand, gravel and boulder). The substrate definitions used in the Ontario Benthos Biomonitoring Network (OBBN) protocol (Jones *et al.* 2004) were employed for clay, silt, sand, and gravel, but the other categories were modified to better describe the substrates found in the lakes. These changes included an organic category consisting of heavily decomposed organic material and an organic debris category that consisted of material in an earlier state of decomposition because these substrates were very common. The OBBN protocol also identified three rock classes, but we found that the gravel category (2–65 mm diameter) and boulder category (>65 mm diameter) adequately described these substrates. The abundance of vegetation (emergent, floating and submerged) was similarly estimated by visual inspection of the transect area. The two field personnel independently estimated vegetation percentages and these estimates were averaged to provide a single value for each vegetation type at each site.

2.5. Selecting reference lakes

Rare BMI families were defined as those found at <5% of sites (e.g., Griffith *et al.* 2001). Rare families were excluded from subsequent analyses because they can obscure patterns in multivariate analyses, making interpretation difficult (Gauch 1982; Norris 1995). Data interpretation was challenged by the natural variability found throughout the study area, so rare taxa were eliminated in an attempt to clarify results.

In most bioassessment protocols, such as RIVPACS (River Invertebrate Prediction and Classification System) and BEAST (Benthic Assessment of Sediment), reference lakes for each test lake are selected which contain the most similar biota through methods such as cluster analysis (Reynoldson *et al.* 1995; Clarke *et al.*

1996). Grouping sites based on biota assumes that BMI assemblages occur in discrete groups but BMI assemblages are commonly described as continuously changing (Linke *et al.* 2005). A grouping approach based on biota was not used in this study because the significant natural variation in water chemistry, substrate types and vegetation abundances in the AOSR, especially pH and DOC, had to be taken into account when assessing biological impacts in the test lakes. By implementing a methodology that takes these variations into account it increased the probability that natural differences in BMI composition were not confused with differences caused by atmospheric pollution. A methodology similar to the Assessment by Nearest Neighbours Analysis (ANNA) (Linke *et al.* 2005) was therefore adopted to identify appropriate reference lakes for each test lake. The methodology was altered to suit application in the current study but the premise of grouping sites based on environmental predictors and assigning reference sites based on Euclidean Distances was shared. This nearest neighbours approach accounts for some of the natural variability in the study area, including pH and DOC, by comparing test lakes with reference lakes that are similar chemically and physically.

A forward stepwise multiple regression was performed in Statistica 7 (Statsoft 2004) to identify water chemistry variables that explained a significant amount of BMI variance ($p < 0.05$) and therefore should be included during reference lake selection. The stepwise multiple regression was conducted using log₁₀(x+1) transformed BMI family abundances and z-score water chemistry (pH, ALK, DOC, DC, IC, NH₄, NO₃, TKN, TP, Na, K, Mg, Ca, Cl, SO₄, ANC) variables that were standardized using:

$$X = [(x - \text{mean})/\text{SD}] \quad (1)$$

where: X is the standardized value, x is the non-standardized value, mean is the average of the values for each water chemistry parameter and SD is the standard deviation of the values for each water chemistry parameter.

Although some parameters were highly correlated (e.g., IC and DOC), it was not known how they independently affected BMI communities. The logarithmic transformation was used to reduce the effect of absolute abundances (Gauch 1982) whereas the water chemistry variables were standardized because these variables were measured in a variety of units. The water chemistry variables explaining a significant amount of BMI variance (DC, IC, TP, Ca) were used in the selection of reference lakes. In the final ordination, DOC was included although it did not explain a significant amount of variance in the multiple regression ($p = 0.2$) because DOC is found in relatively high concentrations throughout the area and it dramatically affects BMI communities in the AOSR (Parsons *et al.* 2010, this issue).

The next step in finding nearest neighbours of test lakes was to conduct a Principal Component Analysis

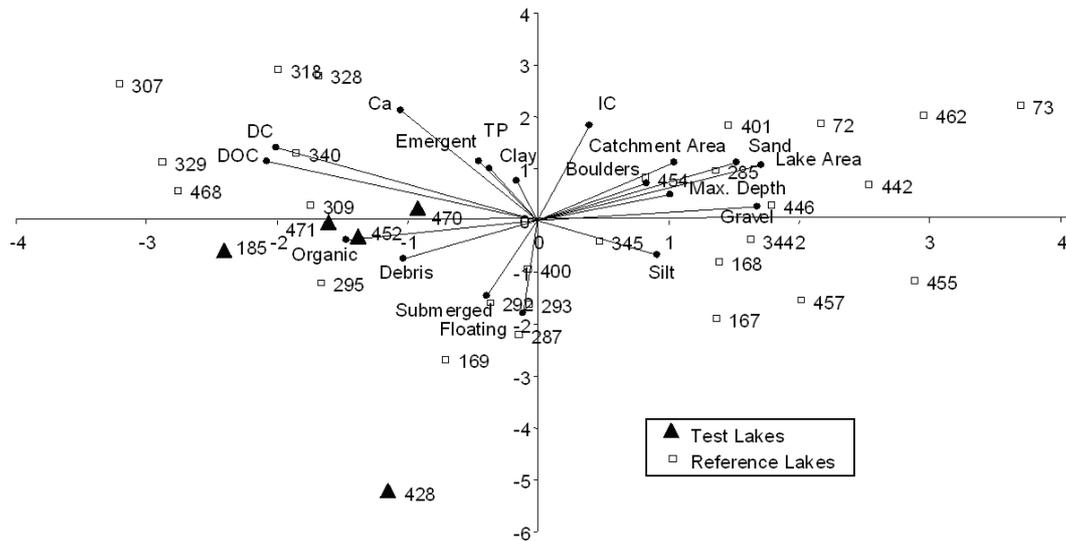


Fig. 2. Principal Component Analysis (PCA) used to find reference groups for each test lake. The environmental variables in the analysis included: water chemistry variables - DOC, DC, IC, TP, Ca; lake physical characteristics - lake area, catchment area, maximum depth; substrate percentages - organic, organic debris, silt, sand, gravel, boulders; and vegetation percentages - submerged, floating and emergent.

(PCA) from a correlation matrix in PC-Ord 4 (McCune & Mefford 1999) which included water chemistry variables (DOC, DC, IC, TP, Ca), lake physicality's (lake area, catchment area, maximum depth), substrate percentages (% Organic, % Debris, % Silt, % Sand, % Gravel, % Boulders) and vegetation percentages (% Submerged, % Emergent, % Floating) (Fig. 2). The PCA distributed the lakes along principal component axes in multivariate space so that nearest neighbours could be calculated (Fig. 2). These variables were chosen as they are not directly affected by the emissions in the AOSR and therefore were appropriate to use when selecting reference lakes. Calcium was the exception, because SO_x and NO_x emissions could directly affect Ca levels due to acidification. Acidic precipitation is often buffered within lakes by cations such as Ca and as a result of this displacement and subsequent replacement through slow weathering processes, Ca levels in lakes fluctuate (Driscoll *et al.* 2003). However, a measurement of acid sensitivity was needed during the selection process to account for the extensive natural variability in acid sensitivity observed amongst lakes, so Ca was included. Groundwater inputs of Ca to the lakes in the region are substantial (Whitfield *et al.* 2010, this issue) and any differences in current Ca concentration caused by acidification in the AOSR will be slight compared to the natural variability in Ca observed among lakes. Therefore, including Ca in an attempt to control natural variability and in turn increase assessment accuracy seemed sensible.

Principal Component Analysis axes 1 through 6 were retained after analysis of a scree plot and these scores were standardized. Euclidean Distances representing the distance between two points in multivariate space were then calculated between every test and refer-

ence lake. The reference lakes with the smallest Euclidean Distances to test lakes represented the lakes that maximized the physico-chemical similarities included in the PCA and were therefore considered the most appropriate reference lakes with which to compare biological data.

The One Sample *T*-Tests, MANOVA and TSA utilized 17 reference lakes for the assessment of each test lake. Bowman & Somers (2006) recommend 30–50 reference sites and Bowman *et al.* (2006) used 30 reference sites. Of the initial 27 reference lakes, only 17 were used here because many reference lakes were inappropriate for analysis due to their inherent differences with the test lakes under observation. Most of the test lakes contained abundant organic substrates and macrophytes, partly because the lakes in the high deposition area were all located in a confined area with limited environmental variability.

2.6. Selecting biological metrics

The following traditional metrics were chosen to evaluate the impact of atmospheric pollution on BMI in test lakes: Taxa Richness, % EPT, % Diptera and Simpson's Diversity. The metrics were selected to include both richness and compositional measures and be proven to detect changes in BMI due to acidification, eutrophication and increased trace metal concentrations. Correspondence Analysis (CA) ordinations were included along with traditional metrics because they represent the dominant trends of variation among sites and they can be more sensitive to stressors than traditional metrics (Kilgour *et al.* 2004; Bowman *et al.* 2006). Correspondence Analysis axes calculated from abundances summarize among-site patterns of abundance, while correspondence axes calculated from pres-

Tab. 1. A summary of simple benthic macroinvertebrate metrics highlighting the differences between test and reference lakes. ¹⁾ EPT = Ephemeroptera, Plecoptera and Trichoptera.

Metric	Test Lakes			Reference Lakes		
	Mean (SD)	Median	Range	Mean (SD)	Median	Range
Abundance	348 (115)	378	219–499	366 (104)	378	114–500
Richness	13.6 (3.7)	14.0	9.0–18.0	13.7 (3.0)	14.0	7.0–20.0
% Hyalellidae	22.5 (39.2)	1.6	0.0–91.4	35.0 (25.6)	40.6	0.0–74.2
% Chironomidae	30.0 (30.0)	21.5	2.2–81.1	30.7 (20.2)	26.1	6.4–78.8
% Sphaeriidae	23.0 (19.1)	22.9	0.0–47.5	7.6 (8.3)	5.2	0.0–34.7
% EPT ¹⁾	5.5 (6.2)	3.2	1.2–16.5	10.1 (11.1)	6.15	0.0–40.4

ence/absence data summarize among-site patterns in richness (Legendre & Legendre 1998). The CA axes calculated with abundance data (CA 1) and presence/absence data (ACA 1) were calculated using Pc-Ord 4 (McCune & Mefford 1999). Additional axes were not included because of the ability of the above mentioned metrics to describe BMI composition. Also, TSA assessment requires a minimum number of degrees of freedom and when the number of metrics is increased the number of reference lakes must as well; since the number of suitable reference lakes was limited, only six metrics were included.

2.7. Determining the difference in biota at test and reference lakes

The six biological metrics for individual test lakes and the mean values for the appropriate group of reference lakes were compared using One Sample *T*-Tests, MANOVA and TSA. The One-Sample *T*-Tests were performed in Statistica 7 (Statsoft 2004) to determine if differences between each test lake and the mean of their appropriate reference group were significantly different. The metrics were squared or log transformed to meet normality ($p > 0.05$), according to the Shapiro-Wilk Test prior to the *T*-Tests.

Multivariate Analysis of Variance and TSA were used to obtain multivariate assessments of difference in BMI metrics between test lakes and the means of suitable reference groups (Bowman & Somers 2006). Multivariate Analysis of Variance is calculated during TSA and it is based on simultaneously analyzing if the difference in the six metrics is significantly different (Quinn & Keough 2002). Test Site Analysis is based on a non-central test where significance is allocated if the difference in BMI metrics is greater than the normal range (95%) of reference lakes, reportedly making the results more ecologically meaningful than if a central *p* value was used (Kilgour *et al.* 1998; Bowman & Somers 2006). Test Site Analysis outputs include a non-central *p* value that measures if the difference between test lake metrics and reference group metrics is "ecologically significant" and categorizes each test site as significantly impaired ($p < 0.05$), potentially impaired ($0.05 < p < 0.95$) or in "reference condition" ($p > 0.95$) (Bowman & Somers 2006). Test Site Analysis also presents a T^2 statistic that identifies the metrics that were

most important in determining the difference between the test site and reference group (Bowman & Somers 2006). This finding can then be used to assess the causation of the impairment, which is an important component of any bioassessment study in the AOSR or other study areas where multiple pollutants exist. The multimetric approach was not solely used in the analysis because it incorporates current knowledge of relationships between structure and function of aquatic ecosystems (Barbour & Yoder 2000) and these relationships were not established in the unique aquatic environments found in the AOSR at the time of analysis, although they have been established since (Parsons *et al.* 2010, this issue). Also, the multimetric approach does not account for correlations between metrics while TSA does (Bowman & Somers 2006).

3. RESULTS

3.1. Benthic macroinvertebrates

The BMI communities were dominated by Hyalellidae, Chironomidae and Sphaeriidae, whereas EPT (Ephemeroptera, Plecoptera and Trichoptera) families were sparse (Appendix 1). There were 45 BMI families collected in total, 16 of which were considered rare as they appeared at <5% of sites (Appendix 1). These rare families were: Chrysomelidae, Elmidae, Gyridae, Haliplidae, Tipulidae, Heptageniidae, Placobdellidae, Sialidae, Hydrobiidae, Lymnaeidae, Brachycentridae, Helicopsychidae, Hydropsychidae, Hydroptilidae, Lepidostomatidae and Limnephilidae. Average abundance and richness were similar between study (348 and 13.6) and reference lakes (366 and 13.7) (Tab. 1). Percent Hyalellidae, % Chironomidae, % Sphaeriidae and % EPT varied considerably in both study and reference lakes (Tab. 1). Percent Sphaeriidae was higher in test lakes (23.0) than reference lakes (7.62), while % EPT was higher in reference lakes (10.1) than test lakes (5.53) (Tab. 1).

3.2. Chemical and physical lake attributes

The mean pH of test (6.1) and reference (6.5) lakes were quite high, but pH values vary considerably in the AOSR (4.6 to 8.0) (Tab. 2). Similarly, the mean DOC of test (25.3 mg L⁻¹) and reference (19.6 mg L⁻¹) lakes were high, while DOC values ranged between 7.1 mg

Tab.2. A summary of chemical and physical lake characteristics, substrate percentages and vegetation percentages for test and reference lakes. ¹⁾: ALK represents total alkalinity concentrations calculated using a PC-Titration Plus system; ²⁾: ANC was calculated as follows: $(K + Mg + Ca + Na) - (NO_3 + Cl + SO_4)$; ³⁾: Values from 2004 Acid Rain Assessment completed by Environment Canada (2004).

Variable	Test Lakes			Reference Lakes		
	Mean (SD)	Median	Range	Mean (SD)	Median	Range
pH	6.10 (0.08)	6.20	5.20–7.00	6.50 (0.90)	6.60	4.60–8.00
ALK ¹⁾ (mg L ⁻¹)	5.80 (5.50)	4.70	1.10–16.20	8.60 (7.60)	8.10	–0.30–28.20
ANC ²⁾ (µeq L ⁻¹)	307 (93)	288	192–469	277 (145)	260	68–551
Ca (mg L ⁻¹)	3.40 (1.00)	3.40	2.00–4.90	3.10 (1.70)	2.60	0.80–6.20
DOC (mg L ⁻¹)	25.3 (7.1)	25.0	14.6–36.8	19.6 (11.2)	15.3	7.1–52.4
SO ₄ (mg L ⁻¹)	1.24 (1.40)	0.70	0.50–3.10	1.60 (1.50)	0.90	0.00–6.30
NO ₃ (µg L ⁻¹)	19.4 (9.4)	17.9	9.9–35.8	20.6 (23.9)	13.3	5.8–120.0
NH ₄ (µg L ⁻¹)	40.8 (18.4)	44.7	16.8–66.0	119.0 (161.0)	45.7	8.3–598.0
TKN (µg L ⁻¹)	616 (256)	507	377–1080	674 (404)	557	163–1790
TP (µg L ⁻¹)	35.8 (10.7)	36.6	19.7–48.2	62.1 (45.4)	48.6	12.2–198.0
Max. Depth ³⁾ (m)	1.50 (0.60)	1.30	0.90–2.50	2.10 (1.40)	1.60	0.80–7.00
Lake Area ³⁾ (ha)	37.9 (21.0)	33.0	9.40–61.0	208.0 (165.0)	145.0	21.0–761.0
Catchment Area ³⁾ (ha)	1180 (836)	1000	400–2120	277 (3940)	1420	190–19600
Debris (%)	46.2 (28.5)	35.0	0.0–100.0	30.0 (33.0)	10.0	0.0–100.0
Organic (%)	38.2 (39.7)	25.0	0.0–100.0	23.2 (29.4)	10.0	0.0–100.0
Clay (%)	0.0 (0.0)	0.0	0.0	2.1 (7.6)	0.0	0.0–60.0
Silt (%)	0.0 (0.0)	0.0	0.0	2.8 (8.8)	0.0	0.0–60.0
Gravel (%)	5.6 (8.8)	0.0	0.0–50.0	10.4 (14.3)	0.0	0.0–80.0
Boulders (%)	0.0 (0.0)	0.0	0.0	1.4 (5.3)	0.0	0.0–40.0
Emergent (%)	5.8 (6.6)	0.0	0.0–40.0	7.3 (11.2)	0.0	0.0–80.0
Floating (%)	10.0 (17.2)	0.0	0.0–40.0	3.3 (7.3)	0.0	0.0–80.0
Submerged (%)	18.6 (30.4)	0.0	0.0–70.0	1.0 (2.8)	0.0	0.0–50.0

L⁻¹ and 52.4 mg L⁻¹ within the region (Tab. 2). The mean ANC and Ca values of test (307 µeq L⁻¹, 3.4 mg L⁻¹) lakes were similar to reference lakes (277 µeq L⁻¹, 3.1 mg L⁻¹), but as with other chemical parameters the variances were substantial for both ANC and Ca (Tab. 2). There was no evidence that SO₄ and NO₃ concentrations were higher in test lakes compared with reference lakes: the mean SO₄ concentrations in test and reference lakes were 1.2 mg L⁻¹ and 1.6 mg L⁻¹, respectively (Tab. 2), while mean NO₃ concentration in test (17.9 µg L⁻¹) and reference lakes (20.6 µg L⁻¹) exhibited the same pattern but values ranged from 5.8 µg L⁻¹ to 120 µg L⁻¹ (Tab. 2). Considering the natural variability of the area the average chemistry of the two groups were quite similar with the greatest exceptions being TP and NH₄, both of which were found in higher concentrations in reference lakes (Tab. 2).

The reference lakes were deeper (2.1 m) and larger (208 ha) than test lakes (1.5 m, 37.9 ha), respectively (Tab. 2). Substrates were predominantly organic in test lakes; very little clay, gravel and boulders were found (Tab. 2). Lastly, emergent and floating vegetation were found in relatively low abundances in the study lakes, and there was more submerged vegetation in test lakes (18.6%) than reference lakes (1%) (Tab. 2).

3.3. Selecting reference lakes

The PCA was composed of a variety of environmental variables of each study lake (Tab. 3). Six principal components were retained and together they explained 62.6% of the variance in the 32 BMI scores.

Principal Component axis 1 represented a substrate gradient where organic substrates and DOC (–0.79) had high negative loadings whereas mineral substrates, sand (0.58) and gravel (0.64) had large positive loadings (Fig. 2 and Tab. 3). Ca (0.75), TP (0.75), IC (0.64) and floating vegetation (–0.63) were the largest contributors to PC axis 2 (Fig. 2 and Tab. 3). The remaining principal components did not exhibit clear relationships with any of the water chemistry variables so they were not presented in figure 2 (Tab. 3). Except for Lake 428, the test lakes were all found in a distinct area of the multivariate space, an area associated with organic substrates and DOC (Fig. 2).

3.4. Determining the difference between BMI at test and reference lakes

One Sample *T*-Tests were performed to identify differences in BMI metrics that were composed of data from test lakes and the mean of their appropriate group of reference lakes. The number of significantly different metrics ($p < 0.05$) ranged from 3 for Lake 470 to 6 for Lake 471, while the remaining test lakes each had 5 metrics that were significantly different than the mean of their reference group (Tab. 4).

The MANOVA showed that all of the test lakes contained BMI that were significantly different than the mean of their group of reference lakes because all of the central p values were significantly different than zero (Tab. 5). The TSA results were different because according to non-central p values, no lakes were significantly impaired but 3 lakes were potentially impaired

Tab. 3 Component loadings of environmental variables upon Principal Components 1 to 6 used during the ANNA-type grouping technique.

Variable	PC 1	PC 2	PC 3	PC 4	PC 5	PC 6
DOC (mg L ⁻¹)	-0.79	0.40	0.00	0.07	-0.17	-0.09
DC (mg L ⁻¹)	-0.76	0.49	-0.07	0.08	-0.17	-0.10
IC (mg L ⁻¹)	0.15	0.64	-0.57	0.04	-0.04	-0.05
TP (µg L ⁻¹)	-0.14	0.75	-0.33	0.08	-0.05	0.11
Ca (mg L ⁻¹)	-0.40	0.75	-0.33	0.08	-0.05	0.11
Organic (%)	-0.39	-0.26	-0.59	0.41	-0.02	0.39
Clay (%)	-0.06	0.27	0.35	-0.15	0.73	0.29
Silt (%)	0.35	-0.23	0.33	0.23	-0.31	0.54
Sand (%)	0.58	0.40	0.09	0.15	0.15	-0.36
Gravel (%)	0.64	0.09	0.03	-0.38	-0.19	0.14
Boulders (%)	0.39	0.18	-0.54	0.01	-0.07	-0.02
Organic Debris (%)	-0.56	-0.14	0.39	-0.38	-0.12	-0.34
Emergent (%)	-0.17	0.40	-0.13	-0.36	0.51	0.40
Floating (%)	-0.04	-0.63	-0.45	0.10	0.35	-0.22
Submerged (%)	-0.15	-0.51	-0.35	0.30	0.35	-0.24
Lake Area (ha)	0.65	0.38	-0.16	0.31	-0.14	-0.01
Catchment Area (ha)	0.40	0.39	0.26	0.54	0.27	-0.23
Max. Depth (m)	0.32	0.25	-0.25	-0.65	0.05	-0.19

Tab. 4. The results of One-Sample *T*-Tests between each metric from each test lake and the average metric value from each reference group. The reference groups were composed of reference lakes that were nearest neighbours in the PCA. Note that some of the metric scores were transformed to meet normality but the original scores are presented here.

Test Lake Number	Metric	Observed Value	Reference Mean (SD)	<i>t</i>	<i>p</i>
185	CA 1	114	21.7 (85.8)	3.89	< 0.01
	ACA 1	-9.13	-29.2 (24.9)	-3.31	< 0.01
	Taxa Richness	13.5	9.00 (3.04)	6.14	< 0.01
	% EPT	0.11	0.05 (0.13)	1.98	0.07
	% Diptera	0.38	0.83 (0.22)	-8.26	< 0.01
	Simpson's Diversity	0.54	0.36 (0.11)	7.04	< 0.01
428	CA 1	42.2	10.2 (94.1)	6.82	< 0.01
	ACA 1	-8.94	126 (32.4)	-17.21	< 0.01
	Taxa Richness	13.4	14.0 (3.60)	-0.74	0.47
	% EPT	0.03	0.10 (0.13)	-3.58	< 0.01
	% Diptera	0.21	0.34 (0.24)	-3.54	< 0.01
	Simpson's Diversity	0.54	0.72 (0.11)	-7.11	< 0.01
452	CA 1	110	21.3 (59.0)	3.17	0.01
	ACA 1	-12.6	14.9 (27.3)	-4.16	< 0.01
	Taxa Richness	13.2	16.0 (3.03)	-3.76	< 0.01
	% EPT	0.16	0.10 (0.12)	6.77	< 0.01
	% Diptera	0.38	0.29 (0.22)	1.65	0.12
	Simpson's Diversity	0.54	0.72 (0.10)	-7.43	< 0.01
470	CA 1	102	15.9 (91.9)	1.52	0.15
	ACA 1	-7.72	37.2 (26.2)	-7.07	< 0.01
	Taxa Richness	13.5	18.0 (3.02)	-6.18	< 0.01
	% EPT	0.02	0.10 (0.13)	-5.34	< 0.01
	% Diptera	0.36	0.25 (0.24)	2.03	0.06
	Simpson's Diversity	0.54	0.58 (0.11)	-1.44	0.17
471	CA 1	-136	3.56 (89.8)	-8.09	< 0.01
	ACA 1	-9.89	88.8 (28.1)	-14.49	< 0.01
	Taxa Richness	13.3	11.0 (3.50)	2.71	0.02
	% EPT	0.01	0.08 (0.11)	-3.61	< 0.01
	% Diptera	0.36	0.02 (0.24)	5.76	< 0.01
	Simpson's Diversity	0.53	0.16 (0.11)	14.25	< 0.01

while the remaining 2 were in "reference condition" (Tab. 6). The potentially impaired category represents a level of admitted uncertainty between reference and impaired conditions (Bowman *et al.* 2006). The less impaired lakes (185, 452) contained lower D and F values than those that were more impaired (428, 470, 471) (Tabs 4 and 5).

3.5. Interpreting effects

All of the test lakes have BMI that are significantly different than the mean values for the appropriate group of reference lakes according to both the One-Sample *T*-Tests and MANOVA. However, TSA identified only 3 test lakes, 428, 470 and 471, which contained BMI

Tab. 5. Multivariate evaluation of test sites using MANOVA. MANOVA outputs include the generalized distance (D) that measures the distance between each test lake and its reference group mean, the central p value that defines impairment and the T^2 that measures the ability of the combined metrics to differentiate between the test lake and the reference group scores.

Lake	Generalized D	Central P	Difference	T^2
185	1.98	< 0.01	Significant	66
428	3.50	< 0.01	Significant	209
452	2.58	< 0.01	Significant	114
470	3.32	< 0.01	Significant	188
471	3.49	< 0.01	Significant	207

Tab. 6. Multivariate evaluation of test sites using Test Site Analysis (TSA). TSA output includes the generalized distance (D) that measures the distance between each test lake and its reference group mean, the non-central p value that defines ecological impairment, the most distinguishing metric in determining impact and the quantitative measure of this difference (partial T^2)

Lake	Generalized D	F	Non-Central p	Ecological Impairment	Most Distinguishing Metric	Partial T^2
185	1.98	7.6	1.00	Reference Condition	% Diptera	3.47
428	3.50	23.9	0.87	Potentially Impaired	ACA 1	4.01
452	2.58	13.0	1.00	Reference Condition	Simpson's Diversity	9.77
470	3.32	21.5	0.92	Potentially Impaired	CA 1	3.23
471	3.49	23.7	0.88	Potentially Impaired	Simpson's Diversity	3.23

composition that was significantly different from their reference groups, implying potential signs of ecological impairment (Tab. 6). The Correspondence Analysis axis 1 (CA 1) based on presence/absence data (partial $T^2 = 4.16$) was the most important metric in determining impact in Lake 428 (Tab. 6). Lakes 470 and 471 were both potentially impaired and the most distinguishing metrics were CA 1 (partial $T^2 = 3.23$) and Simpson's Diversity (partial $T^2 = 3.23$), respectively (Tab. 6).

4. DISCUSSION

4.1. Benthic macroinvertebrates

The BMI collected in the AOSR were dominated by Hyalellidae, Chironomidae, and Sphaeriidae. Previous lentic bioassessments in the AOSR have also reported that Chironomidae dominate the benthos whereas Hyalellidae and Sphaeriidae are common (Golder 2003b; R.L. & L. Environmental Services 1989; Tripp & Tsui 1980). Many lakes in the AOSR are shallow enough to freeze completely from the surface to the sediment in the winter, causing anoxic conditions that favour Chironomidae (Noton & Chymko 1975). Hyalellidae are probably common in the AOSR because the abundant organic material is an ideal food source and acidic lakes, which they are sensitive to, are rare (France 1992). Sphaeriidae are one of the most common and abundant BMI and they are most likely widespread in the AOSR because the lakes contain sufficient dissolved Ca for shell creation (Resh & Rosenberg 1984; Guralnick 2004). Physical lake attributes greatly influence the distribution of Hyalellidae, Chironomidae and Sphaeriidae in the AOSR but Parsons *et al.* (2010, this issue) found that lake water chemistry is an important determinant as well. Plecoptera families were absent,

whereas Ephemeroptera and Trichoptera families were sparse. This is consistent with previous lentic bioassessments which found EPT to be either absent (Syn-crude 1975) or in low abundance (Golder 2003b) in the AOSR. Ephemeroptera, Plecoptera and Trichoptera families are most likely uncommon because they are sensitive to low levels of dissolved oxygen (Resh & Rosenberg 1984; Golder 2003b), and high DOC concentrations (Parsons *et al.* 2010, this issue). Dissolved oxygen was not included in the water chemistry analysis because of temporal and within lake variability.

4.2. Water chemistry

Lake water chemistry in the AOSR is highly variable compared to many other studied areas in Canada (e.g., Dillon & Molot 1997). Previous limnological studies of lentic environments in northern Alberta reported high morphometric and chemical variability between lakes and attributed it to variable geology and drainage characteristics (Halsey *et al.* 1996; Moser *et al.* 1998). Shale is the predominant bedrock in the AOSR and its impervious nature and the low groundwater gradient often combine to form peatlands (Drake 1970). Peatland variability and varying contributions of groundwater contribute to spatially erratic water chemistry because peatlands alter pH and DOC (Dillon & Molot 1997; Halsey *et al.* 1997), while groundwater will contribute to high pH and Ca values (Siegel *et al.* 2006).

Sulphate is one of the main pollutants expected in lakes within the AOSR as a result of oil sand production. However, the SO_4 concentrations found at test lakes were similar to levels at reference lakes. This may suggest that lakes in the test area are currently not impacted or that current monitoring programs that are conducted on an annual basis are insufficient to identify

differences that may occur during periods of high flow (e.g., spring melt, storms). In addition, peatland is widespread in the AOSR and SO_4 concentrations in lakes may be low and has the potential to remain low because of the sulphur fixing capabilities of the peatlands (Siegel *et al.* 2006; Whitfield *et al.* 2010, this issue). Nitrate concentrations were very low and also similar at test and reference lakes but exhibited substantial lake-to-lake variability. Ammonium concentrations were higher at reference lakes, but again there was considerable variation among lakes. Wetlands act as sinks for NO_3 and NH_4 because of vegetation and microbial immobilization (Burns 1998), so the abundance of peatlands and macrophytes in the AOSR probably moderates NO_3 and NH_4 levels in lakes. Also, NO_3 reduction and denitrification cause NO_3 and NH_4 to be consumed rapidly within a lake, especially in eutrophic lakes like those found in many parts of the AOSR.

In the AOSR, lakes range from small, shallow depository lakes with high DOC levels to larger lakes with fewer peatlands in their watersheds, lower DOC values and more heterogeneous substrates. These differences, both physical and chemical, primarily result from underlying geology in the area as it controls drainage, nutrient input and flushing rates, in turn affecting the distribution of dissolved gases, nutrients and organisms (Moser *et al.* 1998). These variations must be controlled in bioassessment studies in the AOSR to ensure that statistical differences between BMI communities are a result of the impact under investigation and not natural variability.

4.3. Assessing the impact of atmospheric pollution

Conclusions drawn from the dataset were strongly influenced by the statistical method used for the evaluation. There were significant differences in selected BMI metrics between individual test lakes and reference lake groups according to the One-Sample *T*-Tests. The MANOVA analysis found even greater evidence of differences in BMI communities since every test lake contained BMI metrics that were significantly different from the mean value of each appropriate reference group. In contrast, the TSA analysis identified no lakes to be "significantly impaired", lakes 428, 470 and 471 were "potentially impaired" and lakes 185 and 452 were in "reference condition". Consideration of the reasons behind potential impairment should provide a better understanding as to whether differences between test and reference lakes may be due to emissions in the AOSR or simply due to the high natural variability among lakes. Lake 428 contained 100% organic substrate, 70% submerged vegetation and a high proportion of Mollusc families. Areas of abundant submerged vegetation are often dominated by Mollusc families (Resh & Rosenberg 1984), so the dominance of Mollusc families instead of Diptera families in Lake 428 most likely resulted from the difference in submerged vegetation between the test and reference lakes and not

atmospheric pollution resulting from oil production. During reference lake selection, substrate and vegetation percentages were taken into account because they are important determinants of BMI communities (Peeters *et al.* 2004), but reference lakes containing abundant organic substrates and vegetation were uncommon.

The abundance of BMI in Lake 470 was low and because of the three hour picking limit only 219 BMI were collected. The reference group for Lake 470 contained an average of 363 BMI per lake so the comparisons might have been unreasonable. Lake 471 was "potentially impaired" and Simpson's Diversity was the most important metric in determining impact because it was much lower in Lake 471 (0.16) than the reference lakes (mean = 0.51). Amphipoda were more dominant in Lake 471 (456) than in the reference lakes (mean = 149), accounting for the disparity in Simpson's Diversity. Overall, the TSA results exhibit substantial differences in BMI in test and reference lakes but it does not appear that these differences are caused by atmospheric pollution originating from industrial activities in the AOSR because BMI are not shifting cohesively as a result of a common pollutant; 2 lakes are in "reference condition"; the potentially impaired lakes contain non-central *p* values that are closer to "reference condition" (0.95) than significant impairment (0.05); and the impairment results are confounded by the natural variability in the area.

The conclusions from the 3 bioassessments are confounded by natural differences between test and reference lakes. During RCA, test lakes were selected in an area of purported impact based on modelled deposition while reference lakes were removed from this area. The test lakes were confined to a small area in the AOSR because the modelled area of high SO_x deposition is currently small. The test lakes happened to be physically similar as a result, characterized by small size, shallow depths, abundant organic substrates, abundant macrophytes and high DOC concentrations. The reference lakes were spread out over a much larger area in an initial attempt to find suitable reference lakes because at the time of lake selection the physical properties of the test lakes were unknown. As a result, some of the natural variability among test and reference lakes could not be removed from the analysis, even using the nearest neighbours approach and this is evident in figure 2. Benthic macroinvertebrates are sensitive to these natural differences in chemistry (Parsons *et al.* 2010, this issue), and especially the substantial disparity in physical lake properties between the two groups of lakes (Resh & Rosenberg 1984; Griffith *et al.* 2001; Peeters *et al.* 2004). Overall, while significant differences were observed between test and reference lakes it appears unlikely that these are due to emissions from the AOSR, although future assessments need to take into account the high variability in lake physical and chemical properties to aid interpretation of the BMI data.

Atmospheric emissions in the AOSR are expected to remain high for the foreseeable future and it is recommended that a regional biomonitoring program with a retrospective comparison to the baseline data collected in this study be employed (with potential expansion) rather than the RCA approach. The current baseline dataset covers a large geographic area, allowing for spatial assessment of impact in the future. Our findings indicate that impacts from atmospheric pollution are negligible at present and therefore the dataset is appropriate for future comparison. If data are collected over time a multimetric assessment could be used because the One-Sample *T*-Tests and MANOVA proved to be sensitive to differences in BMI assemblages. The relationships between lake water chemistry and BMI composition established by Parsons *et al.* (2010, this issue) will help multimetric assessments with metric selection and the examination of cause. If RCA is used and chemical and physical lake variability is controlled, TSA could be used successfully as it proved to be effective when appropriate reference lakes were sampled. Test Site Analysis also contains more diagnostic features than other approaches that could help in inferring cause. Initial selection of suitable reference lakes is important (including a pre-screening survey to collect sediment and macrophyte data) but a grouping technique such as the ANNA approach should also be used to further limit natural variability between test and reference lakes if RCA is used.

5. CONCLUSIONS

Significant differences in BMI between all 5 test lakes and their reference lakes were identified using One-Sample *T*-Tests and MANOVA. In contrast, the TSA analysis identified no lakes that were significantly impaired, 3 lakes (428, 470 and 471) were potentially impaired and 2 lakes (185 and 452) were in "reference condition". The BMI communities in test lakes are unquestionably different than BMI communities in their reference groups but this difference appeared to result from natural physico-chemical differences between the two groups of lakes more so than the effects of atmospheric emissions. Any assessments utilizing the RCA in the future must limit these differences to more clearly identify effects arising from industrial activities in the region. Given the lack of obvious impairment of the study lakes due to emissions from the AOSR we recommend a regional BMI monitoring program be established using the sampling approach utilized in this study to provide a more reliable assessment of impacts in the region and the baseline dataset established in this study may be considered an appropriate start in this regard.

ACKNOWLEDGEMENTS

We would like to thank Tim Seabert, Angela Adkinson, Colin Whitfield and Preston McEachern for assistance during field work and Heather Broadbent, Joe

Findeis, Scott Barker, Kyle Moorhouse and Ellen McLeod for help with laboratory work. Thanks to Ron Reid, Nicole Dmytrow and Chris Jones of the Dorset Environmental Science Centre for training and Julian Aherne and Jacob Potter for technical assistance. Alberta Environment provided logistical support during the field season and it was greatly appreciated. This work was financially supported by the Cumulative Environmental Management Association and a Collaborative Research and Development Grant from the Natural Sciences and Engineering Research Council awarded to SAW and PJD.

REFERENCES

- Bailey, R.C., M.G. Kennedy, M.Z. Dervish & R.M. Taylor. 1998. Biological assessment of freshwater ecosystems using a reference site approach: comparing predicted and actual benthic invertebrate communities in Yukon Streams. *Freshwat. Biol.*, 39: 765–774.
- Bailey, R.C., R.H. Norris & T.B. Reynoldson. 2001. Taxonomic resolution of benthic macroinvertebrate communities in bioassessments. *J. N. Am. Bentho. Soc.*, 20: 280–286.
- Bailey, R.C., R.H. Norris & T.B. Reynoldson. 2004. *Bioassessment of freshwater ecosystems using the reference condition approach*. Kluwer Academic Publishers, Norwell, MA: 167 pp.
- Barbour, M.T. & C.O. Yoder. 2000. The multimetric approach to bioassessment in the United States of America. In: Wright J.F., Sutcliffe D.W., Furse M.T. (Eds), *Assessing the biological quality of freshwaters. RIVPACS and other techniques*. Freshwater Biological Association, Ambleside, UK: 281–292.
- Bowman, M.F. & R.C. Bailey. 1997. Does taxonomic resolution affect the multivariate description of the structure of freshwater benthic macroinvertebrate communities? *Can. J. Fish. Aquat. Sci.*, 54: 1802–1807.
- Bowman, M.F. & K.M. Somers. 2006. Evaluating a novel Test Site Analysis (TSA) bioassessment approach. *J. N. Am. Benthol. Soc.*, 25: 712–727.
- Bowman, M.F., K.M. Somers, R.A. Reid & L.D. Scott. 2006. Temporal response of stream benthic macroinvertebrate communities to the synergistic effects of anthropogenic acidification and natural drought events. *Freshwat. Biol.*, 51: 768–782.
- Burns, D.A. 1998. Retention of NO₃⁻ in an upland stream environment: a mass balance approach. *Biogeochem.*, 40: 73–96.
- David, S. M., K.M. Somers, R.A. Reid, R.J. Hall & R.E. Girard. 1998. Sampling protocols for the rapid bioassessment of streams and lakes using benthic macroinvertebrates. 2nd Ed. Ontario Ministry of the Environment Data Report: 29 pp.
- Dillon, P. J. & L.M. Molot. 1997. Dissolved organic and inorganic carbon mass balances in central Ontario lakes. *Biochem.*, 36: 29–42.
- Doka, S.E., D.K. McNicol, M.L. Mallory, I. Wong, C.K. Minns & N.D. Yan. 2003. Assessing potential for recovery of biotic richness and indicator species due to changes in acidic deposition and lake pH in five areas of southeastern Canada. *J. Environ. Monitor. Assess.*, 88: 55–101.
- Drake, J.J. 1970. *The geomorphic implications of geo-hydrology of gypsum karst areas*. MSc Thesis. McMaster University, Hamilton, ON: 90 pp.
- Driscoll, C.T., K.M. Driscoll, M.J. Mitchell & R.J. Dudley. 2003. Effects of acidic deposition on forest and aquatic ecosystems in New York State. *Environ. Pollut.*, 123: 327–336.

- Environment Canada. 2004. *Canadian acid deposition science assessment*. Meteorological Service of Canada: 479 pp.
- France, R.L. 1992. Biogeographical variation in size-specific fecundity of the amphipod *Hyaletta azteca*. *Crustaceana*, 62: 240–248.
- Gauch, H.G. 1982. *Multivariate analysis in community ecology*. Cambridge University Press, New York: 435 pp.
- Golder Associates Ltd. 2003. RAMP 2003 Annual Report. Regional Aquatic Monitoring Program, Calgary, AB.
- Golder Associates Ltd. 2003b. RAMP 5 year report. Regional Aquatic Monitoring Program, Calgary, AB.
- Golder Associates Ltd. 2004. A summary of work for the management of acid emissions in the oil sands. Background information for the dynamic modelling of acid emissions. Workshop #2.
- Griffith, M.B., P.K. Kaufmann, A.T. Herlihy & B.H. Hill. 2001. Analysis of macroinvertebrate assemblages in relation to environmental gradients in Rocky Mountain streams. *Ecol. Appl.*, 11: 489–505.
- Guralnick, R.P. 2004. Life-history patterns in the brooding freshwater bivalve *Pisidium* (Sphaeriidae). *J. Mollus. Stud.*, 70: 341–351.
- Halsey, L.A., D.H. Vitt & D.O. Trew. 1997. Influence of peatlands on the acidity of lakes in northeastern Alberta, Canada. *Water Air Soil Pollut.*, 96: 17–38.
- Harman, W.M. 1997. Ostego Lake macrobenthos communities between 1968 and 1993: indicators of decreasing water quality. *Freshwat. Ecol.*, 12: 465–476.
- Jones, C., K.M. Somers, B. Craig & T. Reynoldson. 2004. Ontario Benthos Biomonitoring Network Protocol Manual, Version 1.0. Ontario Ministry of Environment.
- Kilgour, B.W., K.M. Somers & D.R. Barton. 2004. A comparison of the sensitivity of stream benthic community indices to effects associated with mines, pulp and paper mills, and urbanization. *Environ. Toxicol. Chem.*, 23: 212–221.
- Kilgour, B.W., K.M. Somers & D.E. Mathews. 1998. Using the normal range as a criterion for ecological significance in environmental monitoring and assessment. *Ecoscience*, 5: 542–550.
- Kowalik, R.A., D.M. Cooper, C.M. Evans & S.J. Ormerod. 2007. Acid episodes retard the biological recovery of upland British streams from acidification. *Global Change Biol.*, 13: 2439–2452.
- Legendre, P. & L. Legendre. 1998. *Numerical Ecology*, 2nd Ed. Elsevier, Amsterdam: 853 pp.
- Linke, S., R.N. Norris, D.P. Faith & D. Stockwell. 2005. ANNA: a new prediction method for bioassessment programs. *Freshwat. Biol.*, 50: 147–158.
- McCune, B. & M.J. Mefford. PC-Ord: Multivariate analysis of ecological data. MjM Software Design, Gleneden Beach, Oregon.
- Moiseenko, T. I. 2003. Effects of acidification on aquatic ecosystems. *Russ. J. Ecol.*, 36: 93–102.
- Moran, M.D., Q. Zheng & M. Samaali. 2007. Long-term multi-species performance evaluation of AURAMS for first 2002 annual run. Environment Canada internal report, Toronto, ON.
- Moser, K.A., J.P. Smol, D.S. Lean & G.M. MacDonald. 1998. Physical and chemical limnology of northern boreal lakes. *Can. J. Fish. Aquat. Sci.*, 377: 25–43.
- National Energy Board. 2006. *Canada's oil sands, opportunities and challenges to 2015: an overview*. Canadian Energy Assessment, Canada: 85 pp.
- Norris, R.H. 1995. Biological monitoring: the dilemma of data analysis. *J. N. Am. Benthol. Soc.*, 14: 440–450.
- Noton, L.R. & N.R. Chymko. 1978. *Water quality and aquatic resources of the Beaver Creek Diversion System*. Environmental Research Monograph, Syncrude Canada Ltd.: 350 pp.
- Parsons, B.G., S.A. Watmough, P.J. Dillon & K.M. Somers. 2010. Relationships between lake water chemistry and benthic macroinvertebrates in the Athabasca Oil Sands Region, Alberta. *J. Limnol.*, 69(Suppl. 1): 118–125. DOI: 10.3274/JL10-69-S1-12.
- Peeters, E.T.H.M., R. Gylstra & J.H. Vos. 2004. Benthic macroinvertebrate community structure in relation to food and environmental variables. *Hydrobiologia*, 519: 103–115.
- Quinn, G. P. & M.J. Keough. 2002. *Experimental design and data analysis for biologists*. Cambridge University Press, Cambridge, UK: 537 pp.
- R.L. & L. Environmental Services Ltd. 1989. *OSLO Project: Water quality and fisheries resources baseline studies*. BOVAR Environmental Services: 127 pp.
- Reid, R.A., K.M. Somers & S.M. David. 1995. Spatial and temporal variation in littoral-zone benthic invertebrates from three south-central Ontario lakes. *Can. J. Fish. Aquat. Sci.*, 52: 1406–1420.
- Resh, V.H. & D.M. Rosenberg. 1984. *The ecology of aquatic insects*. Praeger, New York: 635 pp.
- Reynolds, D.B. 2005. The economic of oil definitions: the case of Canada's oil sands. *OPEC Review*, 29: 51–73.
- Reynoldson, T.B., R.C. Bailey, K.E. Day & R.H. Norris. 1995. Biological guidelines for freshwater sediment based on Benthic Assessment of Sediment (the BEAST) using a multivariate approach for predicting biological state. *Aus. J. Ecol.*, 20: 198–219.
- Reynoldson, T. B., R.H. Norris, V.H. Resh, K.E. Day & D.M. Rosenberg. 1997. The reference condition: a comparison of multimetric and multivariate approaches to assess water-quality impairment using benthic macroinvertebrates. *J. N. Am. Benthol. Soc.*, 16: 833–852.
- Rosenberg, D.M. & V.H. Resh (Eds). 1993. *Introduction to freshwater biomonitoring and benthic macroinvertebrates*. Chapman and Hall, New York: 488 pp.
- Siegel, D.I., P.H. Glaser, J. So & D.R. Janecky. 2006. The dynamic balance between organic acids and circumneutral groundwater in a large boreal peat basin. *J. Hydrol.*, 320: 421–431.
- StatSoft Inc. 2004. STATISTICA (Data Analysis Software System), Version 7. SS Inc.
- Syncrude Ltd. 1975. Baseline environmental studies of Ruth Lake and Poplar Creek. Syncrude Environmental Research Monograph. Calgary, AB.
- Tripp D.B. & P.T.P Tsui. 1980. *Fisheries and habitat investigations of tributary streams in the southern portion of the AOSERP study area*. Volume I. Alberta Environment and Environment Canada, Edmonton, AB: 224 pp.
- Whitfield, C.J., J. Aherne, B.J. Cosby & S.A. Watmough. 2010. Modelling catchment response to acid deposition: a regional dual application of the MAGIC model to soils and lakes in the Athabasca Oil Sands Region, Alberta. *J. Limnol.*, 69(Suppl. 1): 147–160. DOI: 10.3274/JL10-69-S1-15.

