EVS ENVIRONMENT CONSULTANTS

FINAL REPORT

REVIEW AND ANALYSIS OF ANC RIVER MONITORING STUDIES ON THE ATHABASCA RIVER

Prepared for: ALBERTA NEWSPRINT COMPANY
P.O. Box 9000
Whitecourt, Alberta
T7S 1P9

Prepared by: EVS CONSULTANTS
195 Pemberton Avenue
North Vancouver, B.C.
V7P 2R4

EVS Project No: 3/561-01.2

July 1992

Environmental Services for
Industry and Government

Vancouver, Canada
Seattle, U.S.A.

EVS ENVIRONMENT
CONSULTANTS

Our File: 3/561-01.2

July 13, 1991

Mr. Brian Steinback
Alberta Newsprint Company
Postal Bag 9000
Whitecourt, Alberta
T7S 1P9

Dear Mr. Steinback:

Re: Review of ANC River Monitoring Studies on the Athabasca River

Eight (8) copies of the final report for the above project are enclosed. The review and analysis was divided into two tasks. Task I was the initial review and comparison with EEM protocols; Task II was an in-depth analysis of the benthos data, comparing pre- and post-operational benthic communities in the Athabasca River.

We found the Beak/Sentar studies to be technically defensible, generally conforming with requirements of the EEM protocols and with a rational benthic monitoring study design. Collection of pre-operational baseline data is very important for impact assessment monitoring studies, and the availability of such data is a strong positive aspect of the Beak/Sentar studies. Some minor areas in which improvements could be made in the future are noted in the Task I report.

Our initial review (Task I) indicated that oxygen sags below the mill were not of concern, as oxygen saturation remained above 90% at all sites. There was evidence of elevated phosphorous levels in the first year of operation, but subsequent reduction of phosphorous loads from the mill has reduced phosphorous levels in the river downstream of the mill.

Our second task was to summarize and analyze statistically the differences within and between stations and between years for spring and fall sampling. The analyses indicated some changes identified as impacts which were statistically significant relative to natural changes. These were:

- a post-operational increase in total abundance downstream of the ANC mill

.../2

195 Pemberton Avenue
North Vancouver, B.C.
Canada V7P 2R4
Tel: (604) 986-4331
Fax: (604) 662-8548

2517 Eastlake Ave. East
Suite 200
Seattle, WA
USA 98102
Tel: (206) 328-4188
Fax: (206) 328-4291

- a post-operational decrease in adjusted richness in sites downstream of the mill was observed with a partial or complete recovery at near-field sites in 1991. Irrespective of the cumulative inputs of phosphorous by Millar Western and the Town of Whitecourt, 1991 data exhibited a complete recovery at far-field sites.

- a substantial post-operational increase in abundance of the Chironomidae at near-field sites, with very little change in the abundance of the more pollution-sensitive EPT (may flies, stone flies and caddis flies) taxa.

For your information, adjusted richness represents a way of expressing the number of taxa present relative to the total abundance. The 1991 recovery in terms of adjusted richness was coincident with the 50-75% reduction in phosphorous loads from the ANC mill.

The Task II report identifies a serious general flaw in the fundamental design of all benthic monitoring studies. Basically, natural temporal changes are often statistically significant, so that it is difficult to distinguish other temporal changes such as post-operational impacts from "background noise". Means of resolving this problem without altering the study design and abandoning past data are suggested in the Task II report.

In conclusion, release of nutrients in the ANC effluent has resulted in elevated phosphorous levels. A coincident increase in the abundance of some invertebrate taxa tolerant of organic material was also observed. The reduction of phosphorous loads in 1991 is coincident with a biological recovery which should become more evident as more post-operational data become available. The effluent discharge has not affected taxa that are sensitive to organic wastes, consistent with the fact that dissolved oxygen levels have remained high. These impacts were identified from the fall data; no impacts were evident from the spring data. We therefore recommend that the spring sampling be dropped, subject to approval from Alberta Environment. It is evident from our data analysis that changes represent a modest localized enrichment for part of the year. These effects are reversible as evidenced by the 1991 data, and hence are not cumulative.

.../3

EVS ENVIRONMENT CONSULTANTS

We trust that this report satisfactorily concludes our reiew. We have enjoyed working with you and Alberta Newsprint Co., and would welcome further opportunities to continue our association.

Yours truly,

EVS CONSULTANTS

M. Paine

Michael D. Paine, Ph.D.
Environmental Scientist

MDP/ubl

TASK I
Review of ANC River Monitoring Studies on the Athabasca River

EXECUTIVE SUMMARY

From 1989 to 1991, Beak and Associates / Sentar Consultants conducted annual benthic monitoring studies on the Athabasca River in the vicinity of the Alberta Newsprint Company (ANC) paper mill (Luoma and Shelast, 1990, 1991, 1992). As requested by ANC, a two faceted critique of these studies was preformed by EVS Consultants. Firstly, we evaluated the interpretations of chemistry and benthos data provided in the reports. The strongest positive aspect of the Beak/Sentar studies was the availability of baseline data. Secondly, we compared the study design and methods to draft Environmental Effects Monitoring (EEM) protocols for pulp and paper mills (Environment Canada, 1991). EVS Consultants realized that all three studies were conducted well before the draft EEM protocols were available. Therefore, Beak/Sentar could not be expected to meet the protocol requirements in every detail. However, it was important that the Beak/Sentar studies conform with the general requirements of the protocols, and with good benthic monitoring study design.

Our interpretation of the data resulted in three main conclusions regarding phosphorous (P), dissolved oxygen (DO) and benthos. In the spring sampling season, the addition of P from the ANC discharge has not markedly increased total P in the Athabasca River downstream of the mill. Total P concentrations during the fall have increased in the Athabasca River immediately below the mill, but these increases have been substantially reduced as a result of treatment improvements. A sag and subsequent recovery of DO occurred downstream of ANC but % saturation never fell below 90% in the spring or 100% in the fall. Therefore, DO was not significantly affected by ANC inputs. In terms of benthic structure, mill discharge has increased standing crop, particularly in near-field areas, without eliminating any taxa. As a result, diversity has decreased downstream of the discharge relative to upstream. These changes would be indicative of mild enrichment in addition to that already present prior to mill operation.

The evaluation of chemistry and benthos data provided three recommendations. Firstly, a more complete chemical analyses is warranted for all seven sites rather than just sites 2 and 3. Secondly, the spring sampling period could be discarded from future studies. However, this issue should be discussed with both Sentar and government representatives. Finally, the benthos data over all three years should be analyzed. Temporal and spatial changes over all three years should be examined using the methods suggested by EVS Consultants.

The comparison of draft EEM protocols with the study design and methods determined that only Quality Assurance / Quality Control (QA/QC) data were lacking. More detailed QA/QC information for chemical analyses and benthos sample processing should be included as Appendices in future reports. In addition, QA/QC protocols, such as external verification and re-sorting, should be included in any future benthos study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

The Task I report was written by Dr. Michael Paine and Clark Gunter and reviewed by Dr. Gary Vigers. Figures were produced by Mr. Gunter. The report was produced by Arlene Booth, Ursula Lowinger, and Leanne Cook. We thank Maire Luoma of Sentar Consultants for providing chemistry QA/QC, and for answering other questions on procedures. Brian Steinback of Alberta Newsprint Company provided comments on an earlier draft.

Citation

Paine, M.D. and C.P. Gunter. 1992. Review and analysis of ANC river monitoring studies on the Athabasca River. Task I: Review of ANC river monitoring studies on the Athabasca River. Prepared for Alberta Newsprint Co., Whitecourt, Alberta by EVS Consultants, North Vancouver, B.C. 12 pp. + Tables + Figures + App.

1.0 INTRODUCTION

This report provides a brief review and critique of three Beak/Sentar reports on benthic monitoring studies conducted on the Athabasca River in the vicinity of the Alberta Newsprint Company (ANC) paper mill (Luoma and Shelast, 1990, 1991, 1992). Our critique consists of:

- an evaluation of the chemistry and benthos data provided in the reports.
- a comparison of the study design and methods with draft Environmental Effects Monitoring (EEM) protocols for pulp and paper mills (Environment Canada, 1991).

In our critique, we recognized that:

- Collection of baseline (pre-operational) data is very important in monitoring studies, but is often not done. The availability of baseline data is the strongest positive aspect of the Beak/Sentar studies.

- The studies began well before the draft EEM protocols were available, and thus cannot be expected to meet the protocol requirements in every detail. Many of these details may be subject to negotiations between ANC and local regulators. However, it is important that the Beak/Sentar studies conform with the general requirements of the protocols, and with good benthic monitoring study design.

2.0 CHEMICAL ANALYSES

The principal potential impacts of the ANC discharge on the Athabasca River are related to increases in total phosphorous and decreases in dissolved oxygen at downstream sites.

2.1 Phosphorous

Three identified point sources of total phosphorous (P) exist between Site 1 and Site 7 on the Athabasca River. The ANC paper mill began production prior to the fall 1990 sampling while Miller Western (MW) Pulp mill and the Whitecourt sewage effluent were in operation throughout 1989-1991. For comparison purposes, means of Site 1 and 2 were computed and subtracted from P levels at other sites.

Spring (Figure 1)

In 1989, total P increased downstream, with a dip at Site 5. In 1990, before ANC was in production, total P at all downstream sites was below the mean concentration at Sites 1 and 2. However, total P did increase slightly downstream of Site 3. In both 1989 and 1990, increased total P at Sites 6 and/or 7 reflected discharges from the MW mill and the Whitecourt sewage plant. In 1991, total P was relatively constant at 0.025 mg/L at all sites, except Site 6. The high level of total P at Site 6 (0.07 mg/L, which exceeded the ASWQO of 0.05 mg/L) reflected other inputs (e.g., MW mill). Addition of total P from the ANC discharge has not markedly increased total P in the Athabasca River downstream of the mill.

Fall (Figure 2)

In fall 1989, total P concentrations at all sites downstream of ANC, except Site 6, was similar to concentrations at Sites 1 and 2. Again, the high level at Site 6 probably reflected inputs from the WM mill and/or the Whitecourt sewage treatment plant. However, in 1990 and 1991, after the ANC mill became operational, total P concentrations immediately below the ANC mill (Sites 3 and 4) were greater than concentrations above the mill (Sites 1 and 2). These downstream increases were much greater in 1990 than in 1991, reflecting the effects of the 1991 implementation of waste treatment processes that reduced mass loading of total P 50 - 75% despite increases in mill production rate. Again, the increases in total P at Site 7 in both 1990 and 1991 suggest other inputs. The ASWQO of 0.05 mg/L was not exceeded in any 1991 samples. It appears that the ANC discharge of total P has increased concentrations in the Athabasca River immediately below the mill, but that these increases have been substantially reduced as a result of treatment improvements.

2.2 Dissolved Oxygen

Dissolved Oxygen (DO) was measured both as concentration (ppm) and as percent saturation. Expressing DO as percent saturation removes the effects of temperature on the solubility of oxygen in water.

Spring (Figures 3 and 4)

In 1989, no percent saturation data were recorded due to mechanical problems. The 1990 data (Figure 3) were collected prior to ANC operations. However, construction was under way and may have had an impact on DO. The patterns for 1990 and 1991 were similar - a decrease in DO at Sites 3 and 4 relative to Sites 1 and 2 (Figure 3). In both years, there was a recovery at Site 5 followed by another sag at Sites 6 and 7. The input of the Miller Western Pulp Mill and the Whitecourt sewage lines may have contributed to this sag further downstream. Although the data suggest that construction and operation reduced DO downstream of ANC, it is important to note that all values were above 90% saturation. DO concentrations also indicated a DO sag below the ANC mill in both 1990 and 1991, but no recovery at Site 5, and no or little decrease further downstream at Sites 6 and 7. The differences between Figure 3 and 4 reflect temperature differences, as the downstream sites (Sites 5-7) tended to be cooler than Sites 3 and 4. The 1989 concentration data do not indicate any trend, but the method of measurement employed that year had very low precision (±1 ppm).

Fall (Figures 5 and 6)

In the fall of 1989-91, DO levels at Sites 3-7 were similar to or greater than levels above the ANC mill (Figure 5), providing no evidence of a DO sag either before or after the ANC mill began operation. All values were close to 100% saturation. The concentration data indicated that DO below the mill was higher than above the mill in 1990, but lower in 1991 (Figure 6). There were only minimal differences among sites in 1989. The magnitude of differences among stations in all years was low (<1 ppm) relative to the levels themselves (< 10 ppm). Therefore, even though DO concentrations decreased immediately downstream of ANC in 1991, there appears to be no reason for concern about fall DO levels.

3.0 BENTHIC COMMUNITIES

3.1 Statistical Analyses

The statistical methods used in the three Beak/Sentar reports followed the draft EEM protocols (Environment Canada, 1991, Chapter 10). Environment Canada sponsored a Benthic Monitoring Workshop in February, 1992, to develop a manual of standardized methods for benthic monitoring studies. The written Workshop summary and first draft of the manual are currently in preparation (Gibbons et al, 1992). The working group on data analysis was critical of the methods given in the EEM document, and suggested alternatives, which will appear in the draft manual from the Workshop. The major criticism of the methods in the draft EEM protocols was that they did not specifically address and test hypotheses related to potential impacts. That is also a criticism of the methods used in the Beak/Sentar reports.

A critique of, and some suggested alternatives to the Beak/Sentar methods of analyses are presented below, with the following caveats:

- Although we strongly agree with the alternative methods suggested in the Benthic Monitoring Workshop, and discussed below, there is no guarantee that these methods will be adopted for the federal pulp and paper EEM program. It can be argued that ANC has fulfilled the requirements for statistical analyses given in the most recent EEM document, and that inadequacies in the EEM protocols should be addressed by Environment Canada.

- The alternative methods test different hypotheses than the methods used by Beak/Sentar, but obviously the data do not change. Thus, these alternative methods may provide the same conclusions, but in a more defensible and rigorous manner.

- Our criticisms of the analysis have no bearing on the quality of the data and study design.

3.2 Critique

Minor criticisms

Standing crop, and probably number of taxa (richness), should be log-transformed. Abundance or standing crop values are almost invariably log-normally distributed. If both variables are log-transformed, then evenness/diversity can be analyzed by examining regressions of richness on standing crop. This regression procedure is more objective than using diversity indices, which assume a relationship between the two variables which usually does not exist (see Washington, 1984, for criticisms of diversity indices).

Use of all-possible pairwise comparisons between means (SNK tests)

Post-hoc comparison procedures such as SNK are very inefficient, and do not take advantage of the selection of reference, near-field, and far-field sites. These post-hoc methods should only be used when the investigator cannot develop *a priori* comparisons which would indicate impacts. In most cases, it is possible to make a restricted set of *a priori* comparisons which test for differences among sites which are consistent with impacts from a discharge.

The method of choice for *a priori* comparisons is the use of orthogonal contrasts (Sokal and Rohlf, 1981, p. 232-242). Contrasts are simply differences between groups of sites (or levels of other factors such as sample years). We note that a reviewer of the Beak/Sentar reports intuitively grasped the use of contrasts. In the copy of the 1992 report we received, someone had calculated the difference in numbers of taxa in fall samples between the average of the two post-operational years (1990-91) and the pre-operational year (1989) (Table 6 of Luoma and Shelast, 1992). That difference represents a contrast, one which represents an obvious test of impacts. The use of such contrasts for temporal changes is discussed in more detail below.

Spatial contrasts are also of interest and follow from the logic of site selection. Two obvious and independent contrasts of interest are:

- difference between average of "impact" or downstream sites (3-7) and the average of the reference sites (1 and 2)

- difference between the average of the near-field sites (3-5) and the average of the far-field sites (6 and 7)

The first contrast or difference examines the average impact (change from reference) over all five downstream sites; the second contrast tests for far-field effects. Values for these contrasts have been calculated for standing crop (on a log scale), richness, and diversity for the three study years (Table 1). Over the three years, standing crop was consistently greater below the discharge than above, and greater at the near-field sites than at the far-field sites. No consistent patterns were apparent for richness and diversity. The significance of these differences can be tested statistically. Note that by using contrasts, we have only two comparisons (more could be tested, up to a maximum of 6), rather than the 21 pairwise comparisons made by the SNK tests.

Reciprocal averaging (RA) versus MANOVA approach

RA is a technique for sorting samples and taxa along gradients; presumably these gradients correspond to environmental gradients. RA is probably superior to alternatives such as principal components analysis (PCA). However, methods such as RA, PCA, and any clustering methods all have the same failing - they identify only the major gradients, groupings, or sources of variance present, and do not explicitly examine potential impacts. Thus, RA might identify a gradient which reflects the effects of a natural factor such as current velocity, and gradients corresponding to pollution or impacts such as nutrient enrichment may be difficult to distinguish from this background gradient. More often, the axes derived from RA may combine natural and impact gradients, so that the two cannot easily be separated. Methods such as RA ask the question:

- what gradients exist?

rather than the more specific:

- what gradients exist which are consistent with the presence of impacts?

The specific question above is best answered by hypothesis-testing methods such as MANOVA (multivariate analyses of variance). In a MANOVA, abundances of dominant taxa are used as variables, and sites (and years, if data from several years are available) as levels of a factor, in the same way that standing crop and richness are analyzed in ANOVA. The general test of the site factor may not be of interest, but it is possible to use contrasts in a MANOVA in the same way as in ANOVA. In a test of a spatial contrast, MANOVA will generate the multivariate vector which maximizes the difference specified by the contrast. For example, if a contrast comparing sites upstream *versus* downstream of the discharge is tested, the vector will maximize the separation between the two sets of sites.

The vectors produced by MANOVA are simply linear combinations of the original variables (taxon abundances), with each variable multiplied by a coefficient or weight. Note that many indices used in benthic monitoring studies are also constructed by multiplying taxon abundances by weights, except that these weights are arbitrarily chosen. We have found that the vectors generated by MANOVA usually resemble commonly used indices, and we use the index values to present the results. For example, the vectors produced by contrasts testing impacts often assign weights of opposite signs to sensitive and tolerant taxa, and are thus similar to Biotic or Tolerance Indices. We prefer to use the MANOVA as an objective means of identifying appropriate indices, and to help us select only a few indices from the many available.

The vectors generated by specific impact contrasts can be quite different from vectors generated from the overall test of site differences. For example, if sites differed in current velocity, the vector from the overall test of site differences might assign weights of opposite sign to sprawling or crawling taxa *versus* swimming or climbing taxa. The positioning of the sites along this vector might be very similar to the positioning of sites along the primary gradient identified by a procedure such as RA, provided that variance among replicates is low. However, the vector from the impact contrast might assign weights of opposite signs to tolerant *versus* intolerant taxa, identifying a gradient consistent with impacts and unrelated to the background natural gradient. Note that RA, like SNK tests, is an inefficient, scatter-gun approach whereas MANOVA and contrasts are directed towards specific impact-related hypotheses.

Time by site interaction not tested

The real test of whether or not the ANC discharge has affected benthic communities would come from a comparison of spatial differences or patterns before and after the mill became operational. As noted previously, a reviewer has already made some preliminary comparisons. However, examination of spatial differences in each individual year is of little benefit in assessing impacts, as there may be natural differences among the sites or pre-existing differences from the impacts of other dischargers on the river. In Beak/Sentar's defence, it should be noted that the draft EEM protocols make no mention of testing temporal changes in spatial patterns in the chapter on benthos (Chapter 10), although the chapter on study design and site selection (Chapter 3) does address this issue.

The proper method for testing for impacts involves using combined temporal and spatial contrasts. The spatial contrasts have already been discussed; the only temporal contrast of interest is the difference between post-and pre-operational years. We have calculated the difference represented for that temporal contrast for standing crop, richness, and diversity for each site for the fall samples (Table 2). Note that the post-operational increase in abundance at the near-field sites (2-5) has generally been greater than at the reference sites (1 and 2) and at the far-field sites (6 and 7). Temporal patterns in richness show no obvious pattern. As a result, analysis of diversity is roughly equivalent to analyzing the inverse of standing crop (i.e., $1/N$ where N is standing crop), and the results are the opposite of those for standing crop - post-operational reductions in diversity at the near-field sites have been greater than at either the reference or far-field sites (Table 2).

Finally, we can combine the temporal contrast with the spatial contrasts to provide the most appropriate test for impacts. The post-operational changes in the differences represented by the two spatial contrasts are also provided in Table 1. With respect to standing crop, both spatial differences increased after operation indicating that abundance relative to other sites increased downstream of the discharge, especially at near-field sites. There were some changes in the spatial differences for richness after operation, but these were small relative to the differences between the two post-operational years (i.e., the difference in 1989 was intermediate between the two post-operational years). Thus, there was little evidence of impacts on richness. As a result, diversity is again roughly equivalent to $1/N$, and decreased at the downstream sites relative to the reference sites, and at the near-field sites relative to the far-field sites, after the mill became operational.

3.3 Interpretation

Our simple analysis indicates that the mill discharge has increased standing crop, particularly in near-field areas, without eliminating any taxa. As a result, diversity has decreased downstream of the discharge relative to upstream. These changes would be indicative of mild enrichment *in addition to that already present prior to mill operation.* This interpretation appears to coincide with that given in the 1992 report, although the conclusion was not very clearly stated in the report. Two obvious questions remain:

- are the changes observed after operation statistically significant?
- did the abundance of all taxa, tolerant and intolerant, increase or was there a shift to more tolerant taxa?

These two questions can be answered by the statistical analyses discussed in Section 3.2. The answer to the first question would provide a formal test of the interpretation offered. The answer to the second question would indicate whether the impact has been negative or neutral to positive. If both tolerant and intolerant taxa have increased in abundance as a result of enrichment (or other impacts), then the impact should be regarded as neutral to positive. If there has been a shift to more tolerant taxa, then the impact should be considered negative. Based on a superficial look at the community data, we suspect the impact has been neutral to positive, and would not expect a negative impact unless there was a major nutrient addition to the river and/or dissolved oxygen levels declined in conjunction with enrichment. The water quality data indicate that increases in nutrient levels have been minor, and that dissolved oxygen levels have not decreased, after operation.

4.0 COMPARISON WITH EEM PROTOCOLS

4.1 Chemical Analysis

The principal deficiencies of the Beak/Sentar studies in comparison with EEM protocols were:

- Methods for collection and storage of samples were not provided.
- QA/QC (Quality Assurance/Quality Control) information was not provided.
- The most important parameters, dissolved oxygen (DO) and phosphorous (P), were measured at all seven sites, but more complete chemical analyses including metals and resin acids were done at only two sites.

The draft and final EEM protocols now or will contain extensive QA/QC requirements. The current draft devotes a complete chapter to QA/QC; the chapter on chemical analyses is largely devoted to QA/QC; and most of the other chapters contain a section on QA/QC. The protocols are also very specific about sample handling and collection methods. Most chemical laboratories will provide QA/QC data such as Standard Operating Procedures (SOP), precision and accuracy of measurements, and recoveries from spiked samples. This material can be included in an Appendix, with the data reports sent by the analytical laboratory.

Upon request, Sentar supplied us with more detailed QA/QC information. The QA/QC for chemical analyses conformed to that proposed in the EEM guidelines. The only noticeable variant was that percent recovery for abietic acid in 1991 was below that recommended in the EEM guidelines. QA/QC data for 1989 and 1990 were not supplied.

The draft EEM protocols require sampling of at least four areas (these are equivalent to sites in the Beak/Sentar reports) for most receiving water studies. The sites should include reference (unimpacted), near-field, and far-field areas. The Beak/Sentar studies included two reference sites above the mill discharge (Sites 1 and 2), three near-field sites between the mill and the first major tributary (McLeod River) (Sites 2 to 5), and two far-field sites (Sites 6 and 7). These sites were well chosen, based on predicted plume dispersion, and provide a dilution gradient, with effects expected to decrease at successive sites downstream of the mill. However, water chemistry analyses for metals, resin acids and fatty acids were only conducted at two sites (2 and 3). The lack of these analyses at all seven sites is probably a minor issue because concentrations of contaminants just below the mill were generally low, and effects on invertebrates minimal.

4.2 Benthic Communities

As with chemical analyses, the EEM draft and final regulations have and will place substantial emphases on QA/QC for benthic structure. All relevant information should be included in an Appendix. From QA/QC information supplied by Beak/Sentar, the following EEM QA/QC protocols were not complied with:

- re-sorting of sample residues to verify that a minimum of 95% of the organisms were recovered
- verifying identifications with an external expert
- retaining a voucher and reference collections (Sentar did reference all chironomids and archive all specimens)

These points should be included in any future benthic structure study.

The taxonomic level of identifications met or exceeded the requirements in the draft EEM protocols, except for Oligochaeta. The protocols require identification of Oligochaeta to genus or species, but they were identified only to family in the Beak/Sentar reports. Oligochaetes were not as abundant as many insect taxa, and identifying Oligochaeta to genus or species is not possible for immatures and would double the cost of identifications. For these reasons, we disagree with the draft protocol requirements, and do not regard the failure to identify Oligochaeta to genus or species as a serious deficiency. Government representatives may not share this opinion, so we recommend archiving the Oligochaeta in case additional identifications are required.

Only one sampling season per year is required in the draft EEM protocols. Late summer/fall is the preferred time for sampling, as many of the insects are more mature and easier to identify than they are in spring. Spring samples tend to be more variable from year to year than are fall samples, and this was true for the Beak/Sentar studies (e.g., compare fluctuations in standing crop in Table 7 of Luoma and Shelast, 1992). We have also experienced difficulty with high water in the spring. Eliminating the spring sampling season is one option available to ANC to reduce the costs of EEM, while still meeting requirements.

5.0 RECOMMENDATIONS

1. More complete chemical analyses should be done at all seven sites rather than just Sites 2 and 3.

2. The spring sampling period could be dropped from future studies, although this issue should be discussed with both Sentar and government representatives.

3. The benthos data should be reanalysed, examining temporal and spatial changes over all three years using the methods suggested above. Ideally, these analyses should be conducted prior to meetings with regulatory authorities in order to provide support for conclusions and interpretations and to narrow the focus of presentations to statistically significant changes. We appreciate that time constraints may not allow for this analysis prior to Ottawa presentations. The analyses will require raw data in electronic format.

4. More detailed QA/QC information for chemical analyses and benthos sample processing should be included in future reports as a Appendices.

5. Additional QA/QC protocol, such as re-sorting and external verification, should be included in any future benthic community study.

6.0 REFERENCES

Environment Canada. 1991. Technical guidance manual for aquatic environmental effects monitoring at pulp and paper mills. Volume 2: methodology. Draft rept. 199 pp.

Gibbons, W.G., M.D. Munn and M.D. Paine. 1992. Guidelines for monitoring benthos in freshwater environments. Draft rept. prepared for Environment Canada, North Vancouver, by EVS Consultants, North Vancouver. 81 pp.

Luoma, M.E. and B.M. Shelast. 1990. A benthic invertebrate monitoring study and fish habitat assessment on the Athabasca River, Whitecourt, Alberta, 1989. Report prepared for Alberta Newsprint Co., Whitecourt, by Beak Associates Consulting, Calgary, Alberta.

Luoma, M.E. and B.M. Shelast. 1991. A benthic invertebrate monitoring study on the Athabasca River, Whitecourt, Alberta, 1990. Report prepared for Alberta Newsprint Co., Whitecourt, by Beak Associates Consulting, Calgary, Alberta. 62 pp. + App.

Luoma, M.E. and B.M. Shelast. 1992. A benthic invertebrate monitoring study and fish habitat assessment on the Athabasca River, Whitecourt, Alberta, 1991. Report prepared for Alberta Newsprint Co., Whitecourt, by Sentar Consultants, Calgary, Alberta. 64 pp. + App.

Sokal, R.R. and F.J. Rohlf. 1981. Biometry: the principles and practice of statistics in biological research. 2nd ed. W.H. Freeman and Co., New York. 859 pp.

Washington, H.G. 1984. Diversity, biotic and similarity incides. A review with special reference to aquatic ecosystems. Water Res. 18: 655-694.

TABLES

Table 1. Differences between downstream and reference sites (DS-R) and between near-field and far-field sites (NF-FF) for standing crop, richness and diversity (H') in fall samples.

Variable	Year	DS-R	NF-FF
Standing Crop (log no./m^2, with antilog in parentheses)	1989	0.08 (1.19)[1]	0.02 (1.04)
	1990	0.60 (3.95)	0.23 (1.71)
	1991	0.18 (1.52)	0.19 (1.54)
	post minus pre[2]	0.31 (2.05)	0.19 (1.55)
Richness (number of taxa)	1989	-1.0	-2.5
	1990	-1.6	-0.2
	1991	3.4	-11.0
	post minus pre[2]	1.9	-3.1
Diversity	1989	-0.06	0.05
	1990	-0.99	-0.27
	1991	0.08	-0.36
	post minus pre[2]	-0.39	-0.37

[1] the difference in logs was 0.08, which means that standing crop at the downstream sites was 1.19 times (19% greater than) standing crop at the downstream sites.

[2] the average of the post-operational years (1990-91) minus the pre-operational year (1989).

Table 2. Post-operational (average of 1990-91) minus pre-operational (1989) values for standing crop, richness, and diversity (H') in fall samples.

Site	Post-minus pre-operational		
	Standing crop (no./m²)[1]	Richness (no. taxa)	Diversity
1	-0.07 (0.86)[2]	-4.0	-0.34
2	0.41 (2.56)	3.0	-0.22
3	0.28 (1.90)	-4.5	-0.91
4	0.84 (6.90)	2.0	-0.85
5	0.62 (4.17)	3.0	-0.72
6	0.35 (2.25)	2.5	-0.49
7	0.33 (2.12)	4.0	-0.42

[1] - log, with antilog in parentheses.

[2] - the difference in logs was -0.07, which means that standing crop after operation was 0.86 times (14% less than) standing crop before operation.

FIGURES

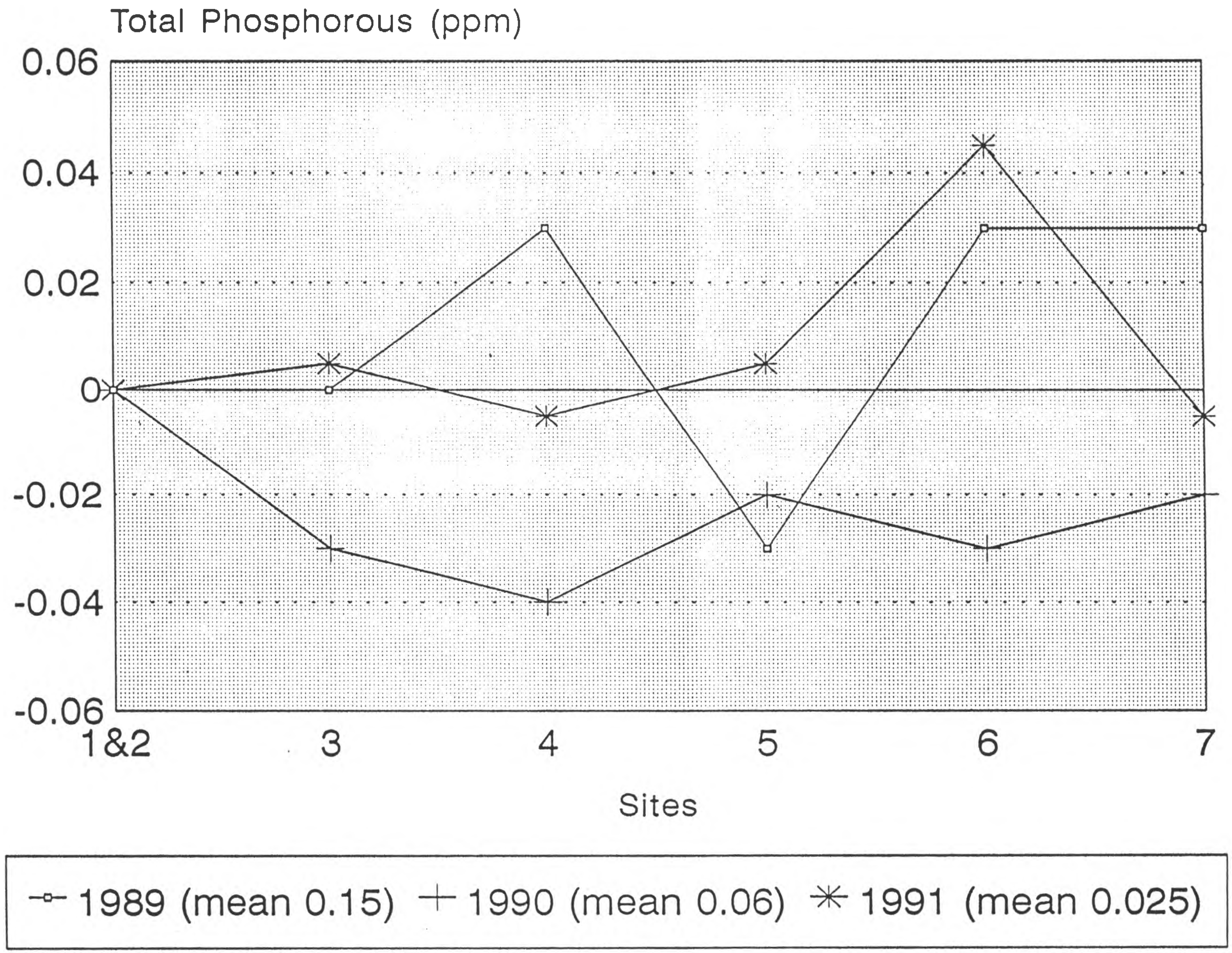

Figure 1. Spring total phosphorous concentrations for selected sites on the Athabasca River relative to mean of sites 1 and 2

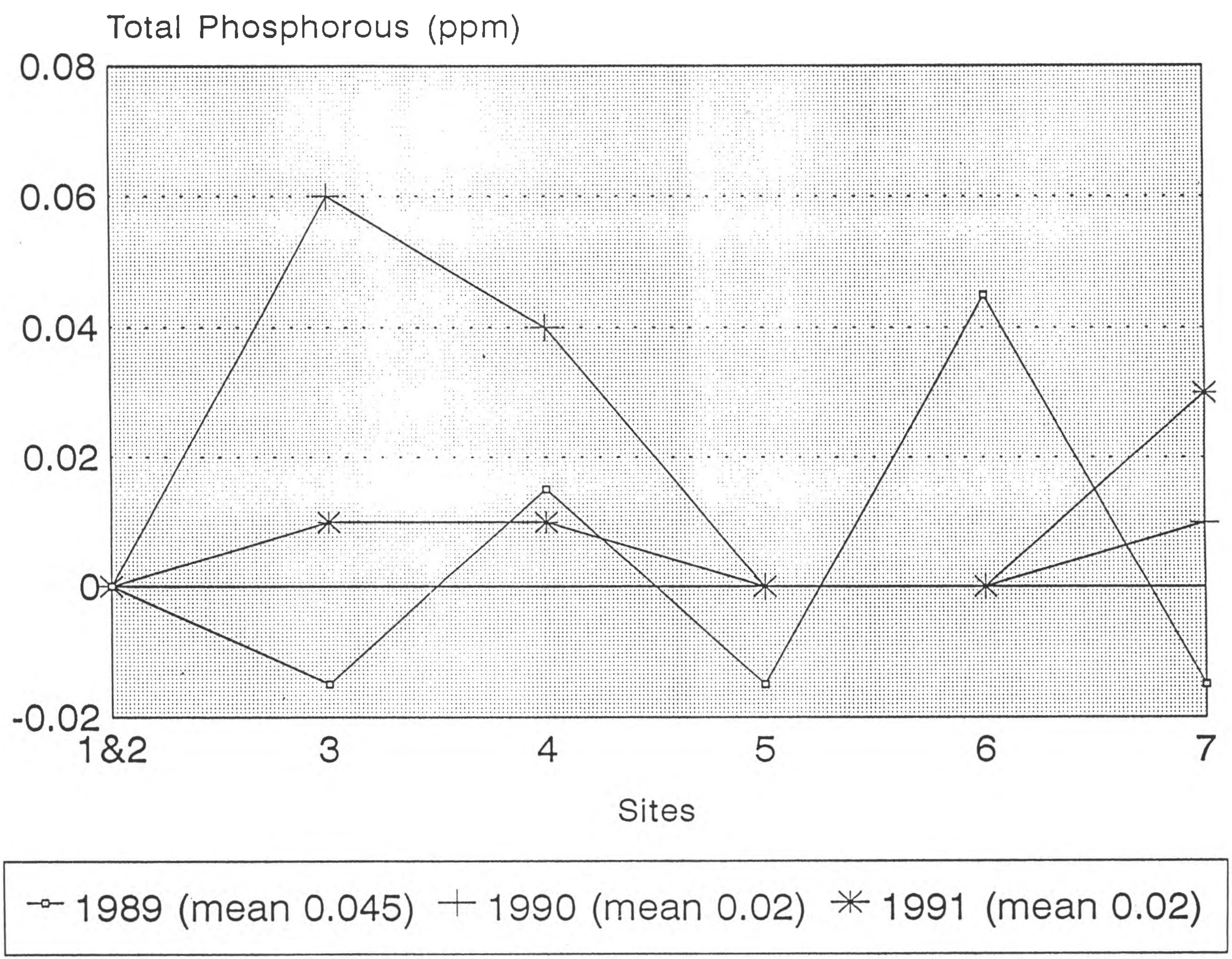

Figure 2. Fall total phosphorous concentrations for selected sites on the Athabasca River relative to mean of sites 1 and 2

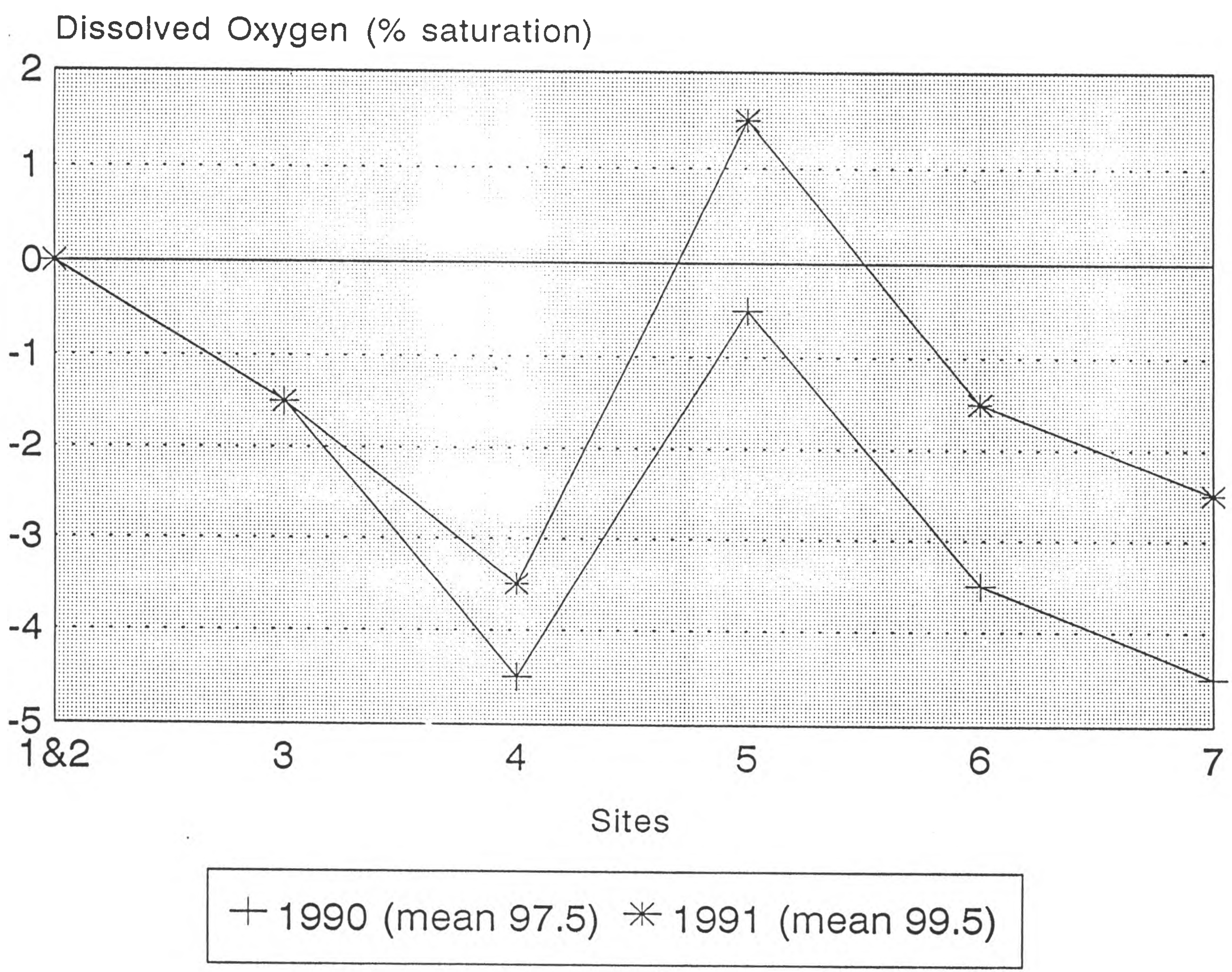

Figure 3. **Spring dissolved oxygen saturations for selected sites on the Athabasca River relative to mean of sites 1 and 2**

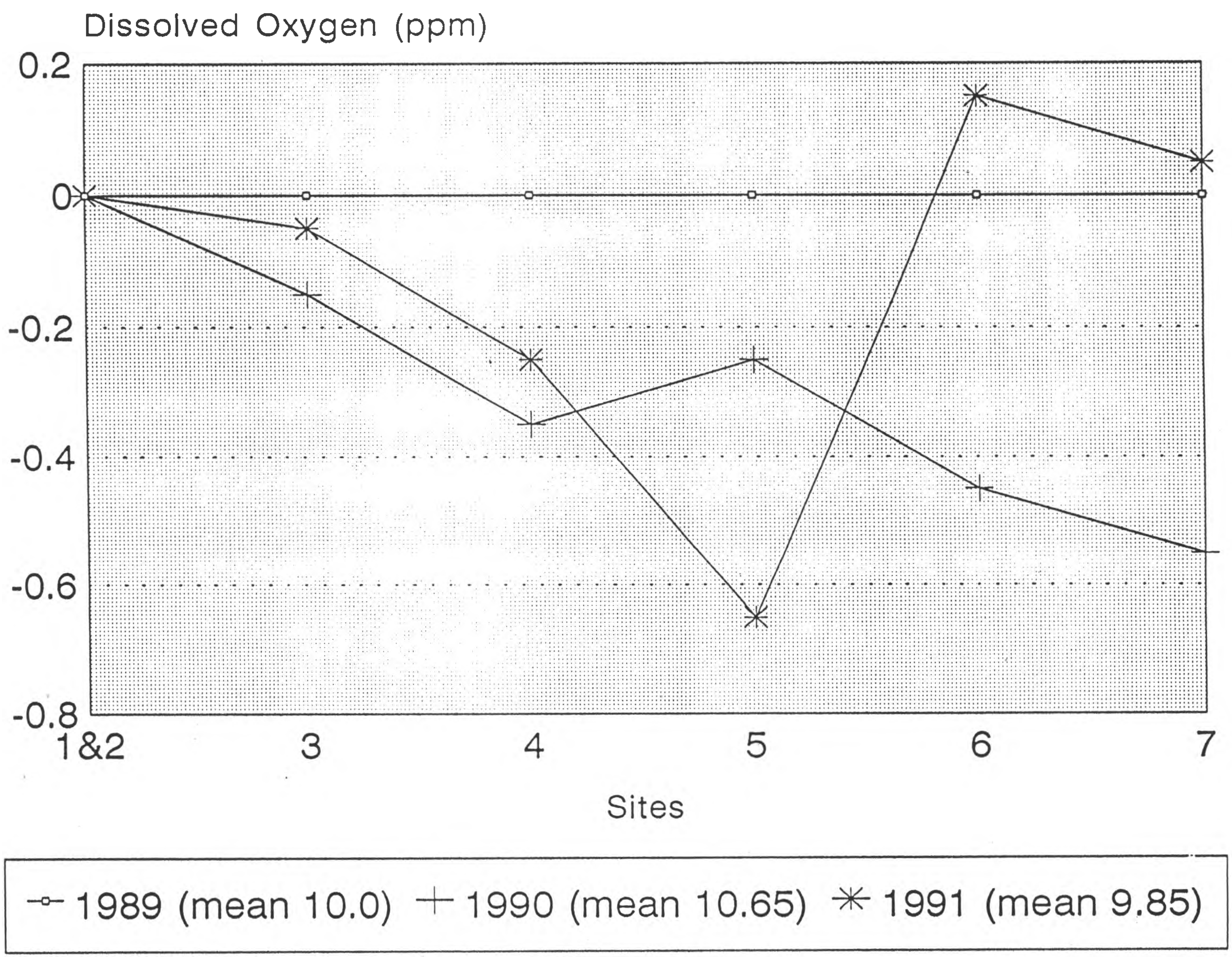

Figure 4. Spring dissolved oxygen concentrations for selected sites on the Athabasca River relative to mean of sites 1 and 2

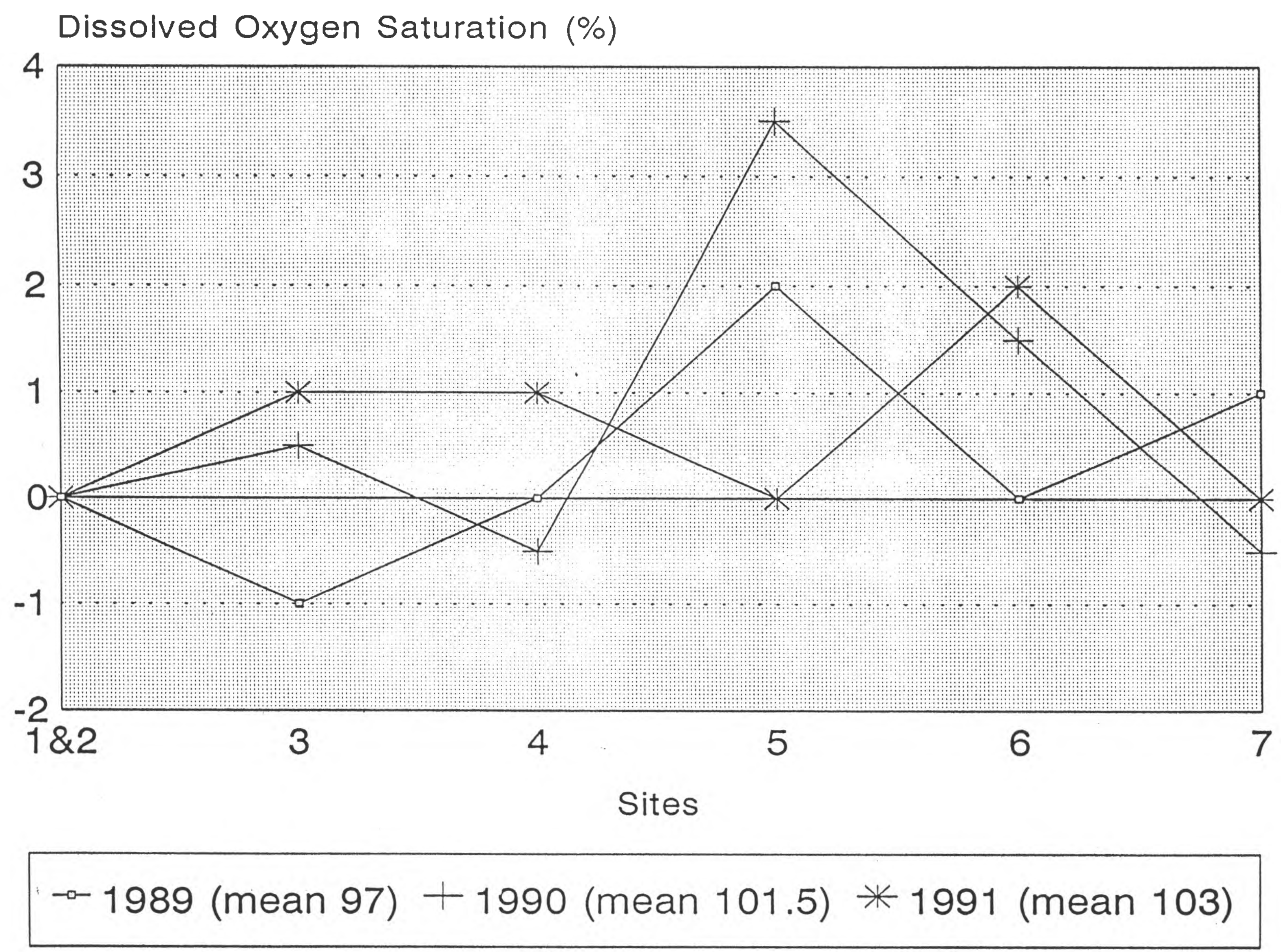

Figure 5. Fall dissolved oxygen saturations for selected sites on the Athabasca River relative to mean of sites 1 and 2

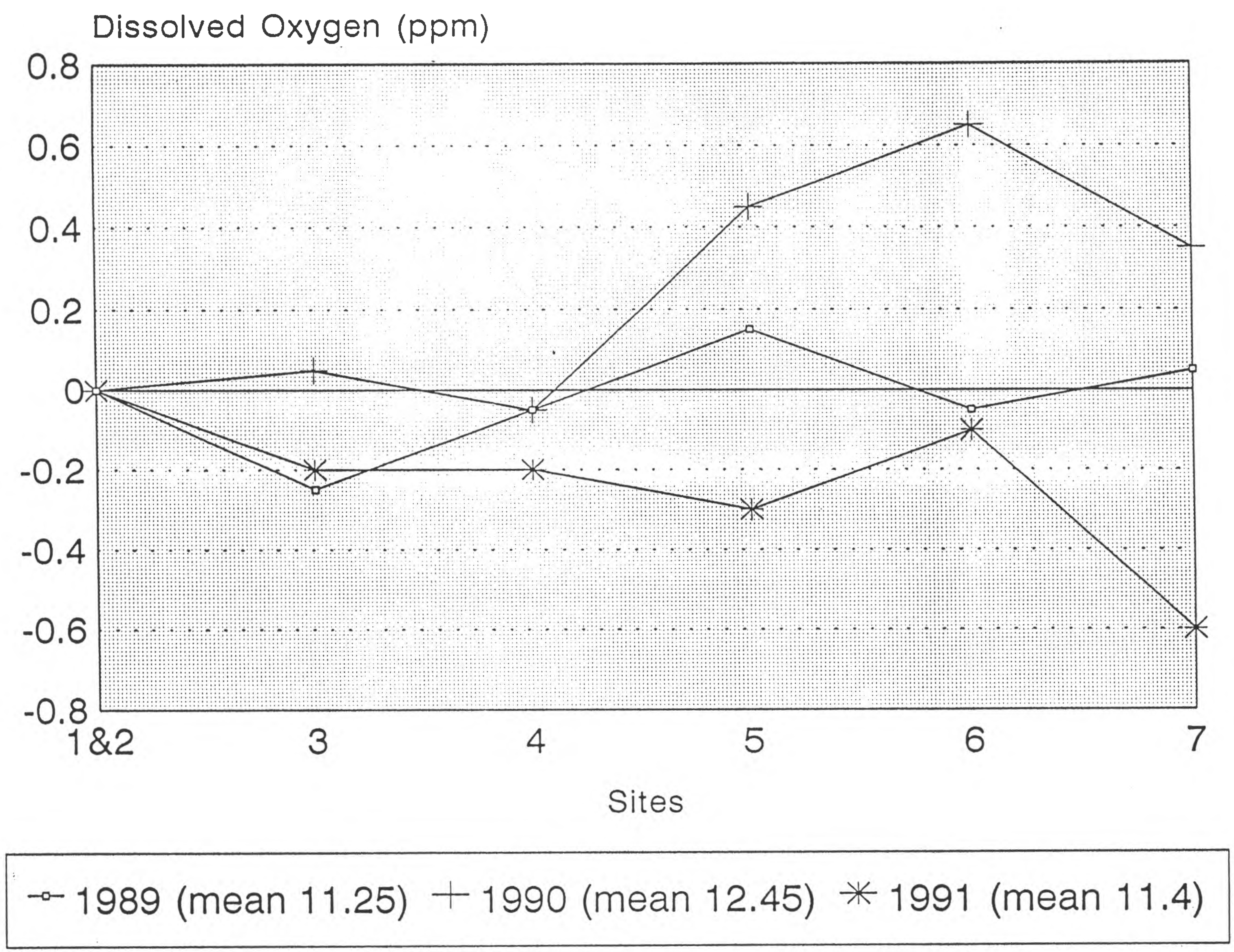

Figure 6. Fall dissolved oxygen concentrations for selected sites on the Athabasca River relative to mean of sites 1 and 2

APPENDIX A
QA/QC FOR CHEMICAL AND BENTHOS ANALYSES

SAMPLING PROCEDURES FOR WATER QUALITY

The following are the sampling procedures used for collecting water quality samples in the field:

- All bottles used for water samples are received from the analytical laboratory.
- All bottles are rinsed three times with river water at the site, prior to filling the bottles.
- The bottles are filled with river sample and preserved, where required with standard preservatives, as instructed by the analytical laboratory.
- All samples are kept cool in coolers with wet ice.
- All BOD samples are shipped on ice to the analytical laboratory within 24 hours.
- All other samples are delivered to the analytical laboratory, as soon as the field survey is completed.
- A chain-of-custody form is completed and delivered with the samples. An example is attached.

The following are the detection limits for the chemical parameters which were analyzed for by the analytical laboratory:

• True Color	5 units
• Total Phosphorus	0.01 mg/L
• TKN	0.1 mg/L
• TSS	1 mg/L
• BOD	1 mg/L
• Total Phenols	0.001 mg/L
• Total Organic Carbon	1 mg/L
• Total Cadmium	0.001 mg/L
• Total Copper	0.001 mg/L
• Total Nickel	0.001 mg/L
• Total Lead	0.001 mg/L
• Total Arsenic	0.001 mg/L
• Total Mercury	0.0005 mg/L
• Total Manganese	0.5 mg/L
• Total Cobalt	0.001 mg/L
• Total Chromium	0.001 mg/L
• Total Iron	0.001 mg/L
• Total Selenium	0.001 mg/L
• Total Silver	0.001 mg/L
• Total Vanadium	0.001 mg/L
• Total Molybdenum	0.001 mg/L
• Resin and Fatty Acids	10 µg/L (May)
	0.63 µg/L (October)

Enclosed also is information on the QA/QC provided by the analytical laboratory on resin and fatty acids, as well as a lab spike which was conducted by the laboratory.

All instrument operation manuals are kept in files with the instrument. Laboratory personnel initial inside cover of the manual after having read the manual, completed training and demonstrated competence in the operation of the instrument. All maintenance records are kept with this file.

4. QUALITY CONTROL

Listed are specific groups of parameters and the quality control methods used.

F. Quality Control Measures

Quality Control Measures

i) Proficiency Samples*
- Lab Duplicate
- Blind Duplicate (submitted by clients)
- Blind Samples (submitted by Lab Manager)
- External Standards (supplied by APG or ERA)
- Matrix Spikes

ii) Control charting of results of Proficiency sample data

iii) Interlaboratory Round Robins
- AWAC
- Alberta Environment

iv) External blind samples sent twice per year by APG and ERA

* Proficiency samples are run at the following frequency:

- **Lab Duplicate - every 10 samples; Independent aliquots tested**
- **Blind Duplicate - one or more submitted per group of samples (at the client's discretion)**
- **External Standards:**
  - **1 per batch of samples for each parameter**
  - **1 for each type of sample matrix**
- **Matrix Spikes - analysed to determine percent recovery of the parameter being analysed in the matrix present in a particular batch of samples**

SAMPLING PROCEDURES FOR WATER QUALITY

The following are the sampling procedures used for collecting water quality samples in the field:

- All bottles used for water samples are received from the analytical laboratory.
- All bottles are rinsed three times with river water at the site, prior to filling the bottles.
- The bottles are filled with river sample and preserved, where required with standard preservatives, as instructed by the analytical laboratory.
- All samples are kept cool in coolers with wet ice.
- All BOD samples are shipped on ice to the analytical laboratory within 24 hours.
- All other samples are delivered to the analytical laboratory, as soon as the field survey is completed.
- A chain-of-custody form is completed and delivered with the samples. An example is attached.

The following are the detection limits for the chemical parameters which were analyzed for by the analytical laboratory:

• True Color	5 units
• Total Phosphorus	0.01 mg/L
• TKN	0.1 mg/L
• TSS	1 mg/L
• BOD	1 mg/L
• Total Phenols	0.001 mg/L
• Total Organic Carbon	1 mg/L
• Total Cadmium	0.001 mg/L
• Total Copper	0.001 mg/L
• Total Nickel	0.001 mg/L
• Total Lead	0.001 mg/L
• Total Arsenic	0.001 mg/L
• Total Mercury	0.0005 mg/L
• Total Manganese	0.5 mg/L
• Total Cobalt	0.001 mg/L
• Total Chromium	0.001 mg/L
• Total Iron	0.001 mg/L
• Total Selenium	0.001 mg/L
• Total Silver	0.001 mg/L
• Total Vanadium	0.001 mg/L
• Total Molybdenum	0.001 mg/L
• Resin and Fatty Acids	10 μg/L (May)
	0.63 μg/L (October)

Enclosed also is information on the QA/QC provided by the analytical laboratory on resin and fatty acids, as well as a lab spike which was conducted by the laboratory.

SAMPLING PROCEDURES FOR BENTHIC INVERTEBRATE SAMPLES

The following are the sampling procedures used for collecting the benthic invertebrate samples:

- All samples are collected in the field, as described in the methods section of the reports and preserved with 10% formaldehyde.
- The samples are delivered to the sorter, once the field survey is completed.
- The sorter follows a set of protocols for sorting the samples and all subsamples are counted for total numbers of organisms.
- The sorted samples are then delivered to the taxonomist for the identification of the samples.
- The taxonomist identifies and enumerates all organisms in the samples.
- The same taxonomist has been used for all survey years to provide consistency in identifications.

Sampler: (Signature) Maire Luoma

Phone: 291-5080

Date Shipped: 7 Oct. 1991

Carrier:

Weigh Bill No.:

SHIP TO: Envirotest Laboratories
9936 - 67th Avenue
Edmonton, Alta. T6E 0P5
ATTENTION: Dieb Birkholz

SEND RESULTS TO: Maire Luoma ~~Dr. Stella Swanson~~
Beak Associates Consulting Ltd.
#155, 2635-37th Avenue NE
Calgary, AB T1Y 5Z6

** Please quote Beak project number on results **

Project Name: Alberta Newsprint, Whitecourt

Project No.: 09-063-01-01

P.O. No.:

Relinquished by: (Signature)	Received by: (Signature)	Date	Time
Maire Luoma	S Prendergast	Oct 7/91	10:00 am
Relinquished by: (Signature)	Received by: (Signature)	Date	Time
Relinquished by: (Signature)	Received at lab by: (Signature)	Date	Time
Discarded at lab by: (Signature)	Discard approved by: (Signature)	Date	Time

ANALYSIS REQUEST

Sample ID No.	Sample Description	Date/Time Sampled	Analysis Requested	Sample Condition Upon Receipt
ANC Site 2	water sample	3 Oct. 1991	Resin Acids*	
ANC Site 3	water sample	3 Oct. 1991	Resin Acids*	

NOTE: DO NOT DISCARD SAMPLES UNTIL DISCARD APPROVED BY BEAK

Special Instructions/Comments: * See list provided for this project in the spring of 1991.

Expected lab turn-around time: Rush (surcharge): ______ Standard: ✓

*** PLEASE RETURN WHITE COPY TO BEAK WITH FINAL RESULTS ***

ENVIROTEST

RESIN AND FATTY ACIDS RESULTS

LAB SAMPLE # 91-D1330-2 IN-MATRIX SPIKE
CLIENT ID ANC SITE 3
OCTOBER 3, 1991

COMPOUND	% RECOVERY
Oleic acid	66
Linoleic acid	82
Pimaric acid	66
Sandaracopimaric acid	66
Palustric/levopimaric acid	46
Isopimaric acid	118
Dehydroabietic acid	68
Abietic acid	26
Neoabietic acid	NA
14-Chlorodehydroabietic acid	70
12-Chlorodehydroabietic acid	71
12,14-Dichlorodehydroabietic acid	73

Detection Limit - 0.63 ppb for all target compounds
ND - Not detected, below detection limit.
NA - Not analyzed

ENVIROTEST

QA/QC

i. To ensure resin acid extraction efficiency, the effluent was fortified with a surrogate compound prior to extraction. Its recovery reflects the extraction efficiency:

o-methylpodocarpic acid	66% (n=3)

To ensure resin acid derivatization efficiency, the final extracts were fortified with tricosanoic acid prior to methylation with diazomethane. Its recovery reflects the derivatization efficiency:

tricosanoic acid	63% (n=3)

CERTIFIED BY: ____________________
Don Boothe, Residue Analyst

APPROVED BY: ____________________
Detlef [Deib] Birkholz, MSc., PhD.
Manager, Environmental Services

ALL SAMPLES WILL BE DISPOSED OF AFTER 30 DAYS FOLLOWING ANALYSIS. PLEASE CONTACT THE LAB IF YOU REQUIRE ADDITIONAL SAMPLE STORAGE TIME.

ACCREDITED BY THE: AMERICAN INDUSTRIAL HYGIENE ASSOCIATION (AIHA) - Industrial Hygiene analysis
STANDARDS COUNCIL OF CANADA - Organic & Industrial Hygiene analysis

1. OBJECTIVE

The objective of Alpha Laboratory Services Ltd.'s QA/QC Program is to ensure that data generated for our clients is of known accuracy to some stated quantitative degree of probability.

2. ORGANIZATION

A. Personnel

i) Laboratory Manager: Bob Lickacz, B.Sc., P.Biol.
ii) Supervisor, Analytical Services: Garry K. Ogletree
Quality Assurance Officer
iii) Laboratory Personnel:
Nadia Li, B.Sc.
Bill Durnford
Binh Luu
Phuc Truong
Angela Hollinger

B. Responsibilities

i) Laboratory Manager:
- communicates commitment to and delegates responsibility for quality assurance.
- allocates funds and resources for effective quality assurance.

ii) Quality Assurance Officer:
- plans and evaluates QA/QC program.
- reports any plans or problems of QA/QC to management.

Supervisor:
- supervises compliance to QA/QC program.
- trains employees in QA/QC operations.
- helps establish Standard Operating Procedures (S.O.P.)

iii) Laboratory Personnel:
- have appropriate education and experience for job.
- input in establishing Standard Operating Procedures.
- follows approved analytical procedures and reporting of data as per S.O.P.

3. QUALITY ASSURANCE DOCUMENTATION AND RECORD KEEPING

A. Laboratory Note Books

- Use bound consecutively numbered pages.
- All entries of raw analytical results are recorded in ink, and are dated and signed.
- Errors crossed out with a single line.
- All blank spaces are cancelled.

- All lab books are periodically inspected and signed by supervisor.
- New books are issued when the old books are turned in.
- Old books are archived in secure area and kept for ten years.

B. Sample Log In

- Chain of Custody forms (in triplicate) sent with sample bottles
- Chain of Custody signed by lab staff when samples received. Sample containers sealed, dated and stored in sample reception shelves.
- Samples are manually logged in by Supervisor.
- Are assigned a lab job number prior to any analysis being conducted. No samples brought into lab without being labelled with a job number.
- Sample log book contains:
  - lab assigned job number
  - name of client
  - name of contact
  - sample description
  - date of receipt.
- Chain of custody form is kept in job file for each job number.

C. Analytical Methods

All analysis of samples are conducted in accordance with the following:

a) the Latest Edition of "Standard Methods for the Examination of Water and Waste Water,"
b) Alberta Environment Methods Manual for Chemical Analysis of Water and Wastes,
c) latest EPA protocols.

D. Calibration Standards

Anions - Standards supplied by ERA (Environmental Resource Associates)
Metals and Cations - Fisher Scientific

All standards supplied with Certificates of Traceability. This documentation is stored in the QA/QC files.

A series of standards are analysed to establish a calibration curve for each parameter to be tested. Calibration curves are determined for each batch of samples tested.

E. Instrument Calibration Results

All calibration results are kept in the analysts' notebooks with the raw data with the exception of atomic absorption, ion chromatograph and total organic carbon analyses. These calibration results are included in the printout of analytical results and as such are stored with the individual job files.

TASK II
Comparison of Pre- and Post-Operational Benthic Communities

EXECUTIVE SUMMARY

Task II consisted of a comparison of pre- and post-operational data on benthic communities collected from 1989-91 in the Athabasca River in the vicinity of the Alberta Newsprint Company (ANC) paper mill. Analyses followed the procedures recommended in Task I. Temporal changes in differences among sites were examined using contrasts which corresponded with impact and non-impact (i.e., natural) changes. The general hypothesis was that if impacts from the ANC mill were present, then the differences between sites just below the mill (near-field) and other sites further downstream (far-field) or upstream of the mill (reference) would change after start-up. Other contrasts tested for more subtle changes within the near- and far-field zones. The variables examined were total abundance (per sample), richness (number of taxa), adjusted richness (residuals from regression of richness on total abundance), number of EPT (Ephemeroptera, Plecoptera, and Trichoptera) taxa, EPT abundance, and Chironomidae abundance. EPT taxa are considered intolerant of organic wastes; Chironomidae are considered tolerant. Spring and fall samples were analyzed separately.

The analyses indicated that almost every contrast tested was statistically significant, regardless of whether it reflected potential impacts or natural changes. Thus, the issue was whether the changes in difference among sites associated with impact contrasts were greater than natural changes. The following changes were identified as impacts which were large relative to natural changes:

- a post-operational increase in total abundance downstream of the ANC mill
- a post-operational decrease in adjusted richness in sites downstream of the mill with a partial or complete recovery at near-field sites, and a complete recovery at far-field sites, in 1991
- a large post-operational increase in abundance of the tolerant Chironomidae at near-field sites, with very little change in the abundance of the more sensitive EPT taxa

All of these impacts were identified from the fall data; no impacts were evident from the spring data. These impacts are consistent with the chemical analyses discussed in Task I, which indicated post-operational nutrient enrichment downstream of the ANC mill, with little or no decrease in dissolved oxygen. The maintenance of high dissolved oxygen levels and the low level of nutrients released were probably responsible for the absence of any decreases in abundance of the sensitive EPT taxa. The 1991 recovery from impacts on adjusted richness may have resulted from 50-75% reductions in phosphorous loads from the ANC mill. There was also some evidence of improvements (reduced impacts) in far-field sites which were possibly the results of treatment improvements by other downstream dischargers (Millar Western Pulp, Whitecourt sewage plant).

This study, and another conducted by EVS Consultants on the Lesser Slave River in Alberta, indicated that natural temporal changes in spatial differences in benthic communities are almost invariably statistically significant. Because temporal changes in spatial patterns are also used in most studies as evidence of impacts, the natural occurrence of such changes will make it difficult to identify impacts. This is a serious flaw in the design of this and other benthos monitoring studies. Another type of design, repeated measures, is recommended for any monitoring studies, including EEM studies, which sample sites over several time periods. We recommend that the ANC study continue with the sampling procedures and study sites used in the past, but suggest a method of analyses which would treat the study as a repeated measures design.

Finally, we repeat our recommendation from Task I that the spring sampling be dropped, subject to approval from Alberta Environment.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

The Task II report was written by Dr. Michael Paine, and reviewed by Dr. Gary Vigers (EVS) and Brian Steinback (ANC). The report was produced by Ursula Lowinger, Arlene Booth and Leanne Cook. We thank Sentar Consultants for permission to reproduce Figure 1, and Maire Luoma of Sentar for providing raw data files.

Citation:

Paine, M.D. 1992. Review and analysis of ANC river monitoring studies on the Athabasca River. Task II: Comparison of pre- and post-operational benthic communities. Prepared for Alberta Newsprint Co., Whitecourt, Alberta by EVS Consultants, North Vancouver, B.C. 18 pp. + Tables + Figures + App.

1.0 INTRODUCTION

Task II consisted of a statistical comparison of pre- and post-operational benthic communities of the Athabasca River in the vicinity of the Alberta Newsprint Company (ANC) paper mill, as recommended in the Initial Review (Task I). The data came from benthos studies conducted by Beak and Associates/Sentar Consultants in spring and fall of 1989-1991 (Luoma and Shelast, 1990, 1991, 1992). These investigators sampled seven sites on the Athabasca, shown in Figure 1; the original reports should be consulted for a complete description of methods.

The objective of these analyses was to determine if spatial patterns of the macroinvertebrate communities changed in a manner consistent with the presence of impacts from the ANC mill discharge. The basic approach involved comparing spatial differences before and after the mill became operational. The statistical analyses followed the protocols recommended by the Benthic Monitoring Workshop (February, 1992) sponsored by Environment Canada (Gibbons et al., 1992a).

2.0 METHODS

All analyses were conducted using SYSTAT statistical packages (Wilkinson, 1990).

2.1 Data

The data set consisted of the abundances of 122 taxa in samples taken during six sampling times (spring and fall of each year from 1989-1991). At each sampling time, five replicate samples were taken from each of the seven sites shown in Figure 1. The data files were provided by Maire Luoma of Beak/Sentar. The variables analyzed were total abundance (per 0.0892 m^2 sample), richness (number of taxa per sample), EPT taxa (number of Ephemeroptera, Plecoptera and Trichoptera taxa per sample), and the abundances of dominant individual taxa or higher level groupings (e.g., EPT, Chironomidae). The EPT taxa are considered sensitive to or intolerant of organic wastes, whereas other taxa such as Chironomidae are considered less sensitive or more tolerant (Klemm et al., 1990).

2.2 Basic Statistical Design

The basic design for all analyses was a two-factor analysis of variance (uni- or multivariate), with Year and Site as factors. The two seasons, spring and fall, were analyzed separately, because there were substantial seasonal differences in benthic communities. The term of interest was the interaction between Year and Site (Year X Site), which measures temporal changes in spatial patterns (i.e., differences among sites). If there were impacts from the mill discharge, all sites would not be affected equally. Sites upstream of the mill would not be affected at all, and far-field sites would be affected less than near-field sites. Therefore, the differences among sites should change between pre- and post-operational years if there is an impact.

The specific approach used was to test the significance of contrasts associated with the Year X Site interaction. Contrasts are differences between combinations of site and/or year means. For example, a simple spatial contrast testing for impacts from the ANC mill would be the difference between the sites upstream and downstream of the mill. However, that difference may have been significant even prior to mill operation if there were some natural downstream gradient or abrupt change in habitat in the vicinity of the mill. The true test of impact would be a comparison of the upstream minus downstream contrasts before and after the mill began operation. That comparison is another contrast combining a spatial contrast and a temporal contrast, and would be associated with the Year X Site interaction. Other interaction contrasts can also be constructed from spatial and temporal contrasts of interest. Several contrasts can be pooled and tested simultaneously. Appendix A provides more details on contrasts.

The Year X Site interaction was broken down into a number of independent contrasts, following rules provided in Sokal and Rohlf (1981). The number of independent contrasts is equal to the number of degrees of freedom (d.f.) for the interaction or 12 (= [3-1] • [7-1]; with 3 years and 7 sites). Tables 1 and 2 provide the contrasts used in this study to analyze the fall and spring data. Interaction contrasts were formed by multiplying coefficients for temporal and spatial contrasts (see Appendix A). Temporal contrasts were divided into pre-*versus* post-operational years, and differences between post-operational (fall) or pre-operational years (spring; there were two post-impact years for the fall and one for the spring because the mill began discharging in fall of 1990). There where six spatial contrasts. The differences between sites downstream (DS) and upstream (US) of the mill, and between near- and far-field sites, are standard comparisons in impact assessment studies. The NF-linear contrast tests for a linear increase or decrease in some variable from site 3 to 5, and would reflect dilution or plume dispersion effects. The NF-quadratic contrast tests for a curvilinear relationship over the same three sites, and would not reflect impacts unless those impacts were localized at Site 3 only, or at Sites 3 and 4 only. The Between FF contrast compares the two far-field (FF) sites and would not reflect impacts unless those impacts extended to Site 6 but not Site 7 (or unless other discharges below the ANC mill also had impacts). The Between US contrast tests for differences between the two sites upstream of the mill, which do not reflect impacts.

The temporal and spatial contrasts were combined and the resulting interaction contrasts divided into "Impact" and "Non-Impact" contrasts. The Impact contrasts were ranked roughly according to their expected magnitude if impacts existed (Tables 1 and 2). For example, any comparisons between the upstream sites and between pre-operational years are obviously Non-Impact contrasts. Differences between pre- and post-operational years are likely to be greater than differences between post-operational years, unless impacts are increasing exponentially or process changes markedly alter impacts. Finally, the expected magnitude of Impact contrasts involving pre- *versus* post-operational years can be ranked as shown in Tables 1 and 2, as the higher ranked contrasts involve more sites and/or reflect greater differences in effluent concentrations. Thus, if an impact were present, differences would be expected to decrease, and probability (P) values increase, from top to bottom in Tables 1 and 2.

If the interaction between Year and Site is broken down into contrasts as shown in Tables 1 and 2, then the sums-of-squares of deviations from the mean (SS) for the individual contrasts should add to the total interaction SS. The SS are a measure of variance, as a standard deviation (SD) is the square root of (SS/[n-1]). The SS are also a measure of the magnitude of the contrast difference or effect size. In the absence of any impacts, the SS for the 12 possible contrasts would be equal to each other and to 1/12 or 8.3% of the total interaction SS. If several contrasts are pooled and tested simultaneously, then the expected SS for the pooled contrasts in the absence of impacts would be equal to the total SS times the number of contrasts (e.g., the number of contrast d.f.) divided by 12. For example if the 7 Non-Impact contrasts for the spring analyses were pooled, and no impacts were present, the total SS for those 7 contrasts would be 7/12 of the total SS. The additivity of contrast SS is a useful property if non-impact changes tend to be statistically significant, which they often are. By comparing the observed magnitude of Impact and Non-Impact contrasts with their expected SS in the absence of impacts, we can determine whether statistically significant impact-related contrasts are greater than statistically significant non-impact contrasts.

2.3 Analysis of Abundance and Richness

Total abundance per sample and richness (all taxa, EPT taxa) were analyzed in ANOVA, followed by the contrasts listed in Tables 1 and 2. Total abundance (both seasons) and richness (fall data) were log-transformed; EPT taxa (fall data) was transformed using log $(x+1)$; richness and EPT taxa (spring data) were not transformed because the log $(x+1)$ transformation was an over-transformation. For the fall data, richness was also analyzed in analyses of covariance (ANCOVA) with total abundance as a covariate (X). The use of total abundance as a covariate removed the effects of total abundance on richness, adjusting all richness values to a common abundance. We refer to the adjusted means (i.e, means of residuals or differences from the regression line) as adjusted richness, which has some relationship with concepts such as evenness (number of individuals per species) and diversity. The ANCOVA approach was not used on the spring data, because preliminary analyses indicated that slopes of log-log relationships between richness and total abundance differed significantly ($P<0.0001$) among year-site combinations.

2.4 Analysis of Dominant Taxa

The ten dominant taxa in each season were identified based on their contribution to the total number of invertebrates collected over the three-year study. The percentage of samples in which each taxon occurred was also calculated, and any taxa not occurring in >90% of the samples were eliminated from further analysis. This last step ensured that most values for abundance were non-zero, but also eliminated all but one taxon for the spring data. Therefore, taxa in the spring samples were pooled to higher levels: EPT and Chironomidae. To be consistent, the same pooling was used for the fall data, and analyses of those data were presented in the main text of the report (Section 3.0). However, abundances of the fall dominants at the lowest taxonomic level was also performed and the results are provided in Appendix B.

The approach used for analyses of the dominant taxa was multivariate analysis of variance (MANOVA) followed by the contrasts in Tables 1 and 2, as recommended in Task I. Section 3.0 of Task I provides more details on MANOVA. The MANOVA conducted were similar to the ANOVA discussed above, with Year and Site as factors, except that more than one variable was analyzed. Contrasts can be used to test impact-related hypotheses. MANOVA provides multivariate tests of significance for factors (e.g., sites) or contrasts. There are a number of test statistics available, and probability values will differ slightly among these statistics, except for single degree-of-freedom contrasts (Wilkinson, 1990); we used Wilks' λ. We also used F-values for Wilks' λ as a measure of effect size or magnitude of impacts, as multivariate differences are difficult to quantify in any other way. These F-values were compared between Impact and Non-Impact contrasts, and among Impact contrasts to determine where the greatest effects were.

MANOVA also provides canonical vectors or discriminant functions; the number of vectors is equal to the degrees of freedom for the factor or contrast, or to the number of variables, whichever is smaller. These vectors are linear combinations of the original variables, with each variable multiplied by a coefficient or weight. The vectors provide the maximum possible separation of factor levels in multivariate space, or maximize the difference associated with a contrast. The vectors provided from MANOVA are similar to some indices, in that taxon abundances are assigned weights for both indices and vectors. In many cases, vectors closely resemble one or more indices, and the index values can be used for presentation. Thus, MANOVA can be useful in objectively identifying which of the many available indices are useful in describing impact-related changes. The weights for individual variables are usually adjusted by dividing by the standard deviation within groups; these adjusted weights are called standardized coefficients (SC). Correlations between the variables and vector scores are also provided by the SYSTAT program used, and these correlations are referred to as loadings.

3.0 RESULTS

3.1 Fall

Table 3 lists the 10 dominant taxa in the fall samples over the three year study. Cumulatively, these taxa accounted for 87% of the total number of invertebrates collected in fall samples. All ten dominants were detritivores and/or herbivores. As Tolerance Values from Hilsenhoff (1987, 1988) indicate, the Chironomidae present were generally tolerant of organic wastes, whereas the EPT taxa, with the possible exception of *Baetis*, were intolerant. Thus, our pooling of EPT taxa, and pooling of Chironomidae, was justified in terms of tolerance to organic wastes. However, Hilsenhoff's Tolerance Values are based on Wisconsin species (EPT taxa) or genera (Chironomidae) and may not apply if species not native to Wisconsin were present in the Athabasca River. Tolerance Values from the U.S. EPA, based on a broader survey of North American species (Klemm et al., 1990), indicate a greater range of tolerance within the taxa listed, and probably a lesser difference between EPT and Chironomidae. The EPA values can be biased as well, because the lower end of the range given is often dominated by relatively rare and localized species which probably do not occur in the Athabasca.

3.1.1 Abundance and Richness

Table 4 provides probability (*P*) values for univariate analyses of changes in abundance and richness. Only one contrast for total abundance was not significant, and the remainder were significant at $P<0.001$, indicating that highly significant changes in abundance were almost universal regardless of whether or not they were associated with impacts from the ANC discharge. Only one of the contrasts for richness which involved comparing pre- and post-operational years (1989 vs. 1990,91) was significant, but the pooled contrast examining the two post-operational years was highly significant. These results suggest that if any impacts occurred, their magnitude varied considerably between the two post-operational years. Results for adjusted richness were similar to those for richness. However, there were some alterations in the relative significance of the pre- vs. post-operational contrasts, and pooling the between post-operational years contrasts may have obscured other similar shifts (see discussion of means below). None of the contrasts comparing number of EPT taxa between pre- and post-operational years was significant, but the pooled contrast involving the two post-operational years was highly significant. These results also indicate that if any impacts occurred they varied considerably between the post-operational years. Finally, note that for all four variables, temporal changes in the differences between the two sites upstream of the mill were significant, often at lower *P*-values than for the Impact contrasts. These results suggest that "impacts" may have been no larger than "natural" changes.

Table 5 provides the breakdown of sums-of-squares (SS) for the interaction between Year and Site for abundance and richness variables. These breakdowns confirm the results from Table 4: non-impact differences or changes were often as large as impact-related differences or changes. For total abundance, the pre-versus post-operational DS-US contrast was the only Impact contrast accounting for substantially more than the expected SS. The SS for the three richness variables indicated that the greatest changes were usually between the two post-operational years and that the difference between the pre-and post-operational years were often less than the natural changes occurring between the two upstream sites.

Table 6 provides means for the four variables for each site in each year. There was a general tendency for total abundance to increase from 1989 to 1991 at all sites, but the increases differed in magnitude among the sites, and did not always occur in the same year. For example, the greatest increase (almost 10-fold) from 1989 to 1991 occurred at Site 4; the second smallest increase (≈50%) occurred at Site 3, immediately upstream, and abundance actually decreased at this site between 1990 and 1991. There was also a general tendency for the post-operational increases in abundance to be greater below the mill.

In 1989, richness did not vary much among Sites (range 21-26), but the variation among sites increased slightly in 1990 (range 21-29), and substantially in 1991 (range 18-36). For all sites, the differences between 1989 and 1990 were ±0-15%, indicating minimal change at any site. Changes from 1990 to 1991 were greatest at the two downstream (far-field) sites (+40-50%). These results suggest that downstream discharges from Millar Western and/or the sewage treatment plant may have improved from 1990 to 1991, resulting in a recovery (increased number of taxa) from previous impacts (see Figure 1 for location of these discharges). The number of EPT taxa was 9-11 at all sites in all years, with the exception of an increase at sites 4, 6, and 7 in 1991. As with richness, these results suggest very little impact from the ANC discharge, but a possible improvement in the two other downstream discharges between 1990 and 1991.

Adjusted richness values showed patterns which were more consistent with an impact from the ANC discharge, confirming our suspicion that pooling the contrasts comparing 1990 with 1991 may have obscured some difference between results for raw and adjusted richness in Tables 4 and 5. In 1989, adjusted richness was similar at all sites, although there was a slight downstream decrease. In 1990, adjusted richness remained at 1989 levels at the two upstream sites but decreased at all sites below the ANC mill. In 1991, adjusted richness decreased at Sites 1 and 2, remained relatively unchanged at Sites 3-5, and increased markedly at Sites 6 and 7. The increases at the two downstream sites would again suggest some improvement in the Millar Western and/or sewage plant discharges. The difference between Sites 3-5 vs. Sites 1 and 2 in 1991 was similar in magnitude to the difference in 1989 suggesting that Sites 3-5 had largely recovered from any impacts which occurred in 1990.

3.1.2 EPT and Chironomidae

Table 7 provides the results of uni- and multivariate analyses of EPT and Chironomidae abundances. All of the multivariate contrasts were significant, usually at very low P values, suggesting that both impact and non-impact changes were large. The F-values for Wilks' λ followed a pattern consistent with the presence of impacts largely restricted to near-field sites. The highest F-values occurred for the three most obvious Impact contrasts (the first three listed). These were the only F-values for Impact contrasts substantially greater than those for the Non-Impact contrasts, suggesting that they were the only impacts greater than "background noise". The pattern of P-values (and SS in Table 5) for Chironomidae followed a pattern similar to that for the multivariate P. EPT abundance made almost no contribution to the first three impact-related contrasts, as those contrasts were not significant (Table 7) and accounted for very little of the SS (Table 5). Instead, some of the lesser Impact contrasts and the Non-Impact contrasts were the most significant and accounted for most of the interaction SS, suggesting some temporal changes unrelated to the ANC discharge.

Standardized coefficients (SC) and loadings for EPT and Chironomidae abundances for the multivariate vectors generated by the contrasts are given in Table 8. For the first three Impact contrasts, and for the dominant vector from the pooled 1990 vs. 1991 contrasts, SC were negative for EPT and positive for Chironomidae. These SC suggest impact-related changes in the relative abundance of these two groups. For these contrasts, the loadings for Chironomidae were positive and high, whereas the loadings for EPT were close to 0. These loadings indicate that the raw abundance of EPT remained relatively constant, and that any changes in relative abundance were due to marked increases in Chironomidae abundance. The Between FF contrast showed almost the opposite pattern - an increase in the raw and relative abundance of EPT with very little change in Chironomidae raw abundance. The other two Impact vectors reflected increases in both taxa as SC and loadings were positive. Finally, the two Non-Impact vectors had high SC and/or loadings for one of the groups, and SC and/or loadings near 0 for the other, suggesting that abundances for the two groups were changing relatively independently.

Table 9 provides mean abundances of EPT taxa and Chironomidae. First, the raw abundance of both groups tended to increase from 1989 to 1991 at most sites, consistent with increases in total abundance (Table 6). We consider these to be partly natural and universal changes, but the increases were much greater for Chironomidae, especially at some of the sites downstream of the ANC mill. The patterns of changes in raw abundance tend to be obscure or difficult to interpret with respect to impacts. However, the patterns in relative (%) abundance of EPT were clear. In 1989, the relative abundance of EPT was reasonably similar at Sites 1-5, and greater than at Sites 6 and 7. In 1990, there was a large decrease in relative abundance of EPT taxa at Sites 3-6 (and a lesser decrease at Site 2), not because the raw abundance of EPT decreased, but because the raw abundance of Chironomidae increased by approximately an order of magnitude. In 1991, the relative abundance of EPT at Sites 3-6 increased, because the raw abundance of Chironomidae remained constant or decreased, whereas the raw abundance of EPT increased. The increased EPT abundance in 1991 was presumably natural, as the largest change (3-fold increase) occurred at Site 1. The changes in relative EPT abundance are consistent with an impact primarily affecting (increasing) Chironomidae abundance in 1990, with some recovery (lesser impact) in 1991. This interpretation assumes some large natural increases in the abundance of both EPT and Chironomidae over the study period, and leaves the large temporal changes in the differences in relative abundance of EPT between Sites 1 and 2 unexplained.

3.2 Spring

Table 10 lists the ten dominant taxa in the spring samples. As in the fall, chironomids in the *Cricotopus/Orthocladius* complex accounted for approximately half of the total number of invertebrates collected. The community was dominated by detritivores, herbivores, and omnivores, but the carnivorous *Isoperla* was also abundant. Tolerance Values from Hilsenhoff (1987, 1988) indicated that there was a mix of sensitive EPT taxa and tolerant Chironomidae, with *Baetis* and *Hydropsyche* intermediate. As for the fall dominants, Tolerance Values from the U.S. EPA (Klemm et al., 1990) indicated less distinction between EPT taxa and Chironomidae, and considerable variation within genera. Only *Cricotopus/Orthocladius* occurred in >90% of the samples; seven taxa occurred in <80% of the samples. As a result, there were many zeroes in the abundance data set, especially in 1991 samples, and pooling to higher taxonomic levels (EPT and Chironomidae) was necessary for statistical analyses.

3.2.1 Abundance and Richness

Table 11 provides *P*-values for differences in abundance, richness, and number of EPT taxa associated with main effects (Year, Site) and their interaction, as well as interaction contrasts. For all three variables, Non-Impact contrasts were more significant (lower *P*) than Impact contrasts, suggesting that any effects from the ANC discharge were weaker than effects from "natural" causes. That conclusion was verified by the breakdown of sums-of-squares (SS), as the bulk of the SS were associated with Non-Impact contrasts (Table 12).

Table 13 provides mean abundance and richness values for each site in each year. Values of all three variables decreased substantially from 1989 to 1991, as sampling occurred earlier in the year in 1990-91, and water levels were high in 1991. Invertebrate abundance is expected to increase over the summer (compare Table 13 and Table 6), and sampling efficiency decreases and drift or displacement increase in high water. The major changes in spatial patterns of abundance occurred between 1989 and 1990, prior to start-up of the ANC mill. In 1989, abundance at Sites 6 and especially 7 was greater than at other sites, but in 1990 abundance was greatest at Sites 2-4. Changes in richness and number of EPT taxa followed a similar patterns, because of the strong correlations between these variables and abundance. The decreased abundance (relative to other sites) at Sites 6 and 7 in 1990 may suggest some reduction of nutrient loads from downstream sources such as the sewage treatment plant and/or Millar Western.

3.2.2 Abundance of Chironomidae and EPT Taxa

Table 14 provides results of uni- and multivariate tests of changes in Chironomidae and EPT abundances. All multivariate Impact contrasts were significant, but with *F*-values less than for the pooled Non-Impact contrasts. Only 4 (of 10) univariate Impact contrasts were significant; the pooled Non-Impact contrast was highly significant for both EPT and Chironomidae. As with total abundance and richness, these results suggest that any impacts from the ANC discharge were small relative to "natural" changes in spatial patterns. This conclusion was verified by the breakdown of SS (Table 12), as most of the SS were associated with the Non-

Impact contrasts. The NF-quadratic contrast for EPT was the only Impact contrast which accounted for more than the expected 8% of SS, and this is probably the most difficult Impact contrast to relate to impacts.

Table 15 provides standardized coefficients and loadings for vectors generated by the various contrasts. For the Impact contrasts (except NF-linear), SC for EPT and Chironomidae had opposite signs, indicating an increase in the relative abundance of one group and a decrease in the relative abundance of the other. In fact, loadings for the two groups had opposite signs for three of the vectors, indicating that there may have been opposite effects on raw, as well as relative abundance. SC and loadings for both groups were positive for the first vector for the pooled Non-Impact contrasts, indicating that changes in the raw and/or relative abundance of both groups were in the same direction. The most obvious "natural" changes in abundance of these two groups would be those described for total abundance in Section 3.2.1.

Table 16 provides mean raw abundances of EPT taxa and Chironomidae and mean relative abundance of EPT. The decrease in abundance of both groups throughout the study parallelled the decrease in total abundance. The greatest reductions occurred at Sites 6 and 7 between 1989 and 1990. Changes in relative abundance of EPT were not informative. In 1989, there was a slight tendency for the relative abundance of EPT taxa to decrease in a downstream direction, but in 1990, this tendency was not obvious, and the differences among sites were less than in 1989. Note that there was also an overall decrease in the relative abundance of EPT from 1989 to 1990. The same decrease was evident in the fall samples, except at Site 1 (Table 9). We interpreted the fall changes as indicative of an impact from the ANC discharge, but the spring data suggest that the fall changes may have been unrelated to impacts. There were some changes in relative abundance of EPT from 1990 to 1991 (i.e., after start-up), specifically decreases at Sites 1 and 2, and increases at other sites. These changes could be interpreted as a positive impact from the mill discharge but the numbers of both EPT and Chironomidae were too low to permit any meaningful conclusions.

3.3 Summary of Impacts

The fall data provided some evidence of impacts from the ANC discharge; the spring data provided no evidence of impacts. One of the factors confounding interpretations was the presence of highly significant, apparently natural temporal changes in spatial patterns, which were as large or larger than any changes associated with potential impacts. Therefore, we constructed "pseudo-*F*" tests, comparing the magnitude of Impact contrasts with Non-Impact contrasts. For both univariate and multivariate tests, pseudo-*F* values were *F*-values for the Impact contrasts divided by the *F*-values for the pooled Non-Impact contrasts. For univariate contrasts, pseudo-*F* values can also be obtained by dividing the % SS for Impact contrasts by % SS for Non-

Impact contrasts, and then multiplying by the Non-Impact d.f. divided by the Impact d.f.. Pseudo-*F* values greater than 1 indicate Impact effects larger than Non-Impact effects.

All pseudo-*F* values >2 have been listed in Table 17; an arbitrary cut-off of 2 was chosen to restrict this summary and the subsequent discussion (Section 4.0) to the largest and most obvious impacts. There were no pseudo-*F* values >2 for the spring data, and in fact there was only one value >1. Thus, the spring data set provided no evidence of impacts from the ANC discharge. For the fall data, the largest impacts were (all increases or decreases relative to natural temporal changes occurring at all sites):

- a post-operational increase in total abundance downstream of the ANC mill
- a post-operational decrease in adjusted richness in sites downstream of the mill with a partial or complete recovery at near-field sites, and a complete recovery at far-field sites, in 1991
- a large post-operational increase in abundance of Chironomidae at near-field sites, with very little change in EPT abundance (the multivariate pseudo-*F* values were not much larger than those for Chironomidae alone)

This summary ignores changes in richness, which were considered partially confounded by changes in abundance.

4.0 DISCUSSION

4.1 Evaluation of ANC Impacts

This section considers only those impacts identified in Section 3.3 as exceeding background or natural changes. An obvious question is whether even those impacts were real, as they were only apparent in the fall data. We would argue that the impacts were real, but were not observed in the spring data for a number of reasons. First, there were large, apparently natural year-to-year changes in abundance in the spring samples at all sites. If factors such as high water (1992), and earlier sampling (1991, 1992) affected all sites equally, their effects would be referred to as strictly additive, and our ability to detect impacts by examining the interactions between year and site would not be impaired. However, it is very unlikely that such large changes would be additive, because we are unlikely to choose the appropriate additive scale (i.e., transformation) to express these changes, and because the variance among sites in the magnitude of any effect is likely to increase with the magnitude of the effect. Thus, large natural temporal changes are likely to be non-additive, contributing to the Non-Impact component of the Year X Site interaction, and/or obscuring any impacts. Second, the spring samples may reflect conditions in the previous winter, rather than conditions in the previous growing season (summer). If impacts are largely restricted to enrichment, rather than reductions in dissolved oxygen, as argued in the Task I report and in Luoma and Shelast (1992), these impacts may only be evident in fall samples, collected soon after the period of maximum growth.

The impacts identified were also noted by Luoma and Shelast (1992). Their explanation for the observed changes appears reasonable - nutrient (phosphorous) addition was sufficient to cause an increase in the abundance of tolerant taxa such as Chironomidae below the mill, but oxygen depletion was not sufficient to cause a decrease in the more sensitive EPT taxa. Because abundance increased but no new taxa were added (i.e., some existing taxa increased in abundance), changes in richness were limited but adjusted richness decreased immediately below the mill. The changes in adjusted richness correspond roughly with the changes in diversity noted in the Task I report. There was also some evidence suggesting that the reductions in phosphorous loads resulting from treatment improvements between the 1990 and 1991 samples led to a recovery from impacts on adjusted richness.

EVS Consultants has conducted a similar study on the Lesser Slave River in Alberta, collecting benthos samples before and after the Slave Lake Pulp Corporation (SLPC) pulp mill began operation (Gibbons et al., 1992b). The SLPC mill does not use.chlorine for bleaching, so nutrient enrichment and oxygen depletion were identified as the most likely impacts, as they also were for the ANC mill. The Lesser Slave River benthos study indicated that mild enrichment from the SLPC mill had led to increased numbers of all taxa, including EPT, and also increased adjusted richness. No depletion of oxygen levels was observed. Therefore, the effects of the SLPC and ANC mill were similar, in that enrichment was mild and the sensitive EPT taxa were not negatively affected. Both studies suggest that enrichment will not negatively affect EPT taxa unless oxygen levels also decrease. However, the impacts from the ANC mill were stronger (i.e., more negative) than those from the SLPC mill in that adjusted richness was negatively affected (decreased) and only Chironomidae increased in abundance. Assuming that the treatment improvements made at the ANC mill after fall 1990 will be or were effective in reducing impacts, then future impacts on benthic communities in the Athabasca River are predicted to decrease and more closely resemble those in the Lesser Slave River. Specifically, adjusted richness below the mill should increase relative to upstream sites (which may already have occurred), abundance of Chironomidae should decrease, and abundance of EPT taxa should remain the same or even increase.

4.2 Potential Impacts from Millar Western

Millar Western also discharges effluent enriched with nutrients into the Athabasca River and might be expected to have impacts on the benthic community similar to those from the ANC discharge. Earlier studies conducted by Beak Associates indicated that the dominance of tolerant taxa increased downstream of the Millar Western mill after start-up in 1988 (Luoma and Shelast, 1988, 1989). Since that time, Millar Western has made a number of treatment improvements and process changes which might be expected to reduce any impacts. Several instances in which impact reductions might have occurred between 1989 and 1991 were noted in Section 3.0. The contrasts used for the analysis of the 1989-91 data were not designed to identify potential impacts (or reduction in impacts) from Millar Western, but those impacts may have affected contrasts comparing far-field sites with near-field or upstream sites, and contrasts comparing the two far-field sites. For the spring data, any reductions in impacts from the Millar Western discharge prior to 1991 would be included with Non-Impact contrasts, and might cause us to overestimate the magnitude of "natural" effects. For the spring and fall, any reductions in impacts from the Millar Western discharge after the ANC mill became operational would be included with Impact contrasts, and would cause us to overestimate near-field effects and underestimate far-field effects.

4.3 Adequacy of Study Design

The Beak/Sentar studies on the Athabasca River are examples of a time by site factorial design. Analyses of that type of design assume that impacts are specific types of interactions between site and time; these types of interactions are identified by the Impact contrasts. The statistical tests of these contrasts use the variance among replicate samples collected at the same time as an error term. Specifically, the tests indicate whether the changes associated with the contrasts are larger than expected, based on the variation observed among the replicates. If the contrasts are large relative to the variance among replicates, we assume that the contrasts were unlikely to have occurred by chance, and instead represent real impacts. The problem with this approach is that it assumes that variance among replicates is the major source of variance other than impacts influencing the magnitude of interactions between time and site. Obviously this is not true, because in both this report, and in the study of benthic communities in the Lesser Slave River, we identified numerous statistically significant non-impact interactions. Thus there is some additional natural and presumably random variance of temporal changes in differences among sites.

In this report, the problem posed by the additional natural temporal variance of spatial differences was addressed by comparing the magnitude of Impact contrasts with the magnitude of Non-Impact contrasts. This approach is complex and lengthy, and cannot provide statistical tests of the hypothesis that impact-related changes are, or are not, greater than those expected by chance. The pseudo-*F* values in Table 17 are the test statistics required to test such an hypothesis, but there were not enough degrees of freedom for the Non-Impact contrasts to provide any meaningful statistical tests. For example, for the fall data, there were only two d.f. for the Non-Impact contrasts. The critical value of *F* (i.e., for $P=0.05$) for a single d.f. contrast, with 2 d.f. for the numerator, is 18.5. Therefore, none of the pseudo-*F* values would be statistically significant, and any impact would have to be enormous to be declared significant. The same considerations apply to the spring data, although with 7 d.f. for the Non-Impact contrasts, the critical *F* is only 5.59 for a single d.f. contrast. Critical *F* for multiple d.f. Impact contrasts will be lower than those for single d.f. contrasts, but not so low that any of the pseudo-*F* for multiple d.f. Impact contrasts in Table 17 would be significant.

There is a design which would more directly test the variance associated with impact interactions or contrasts against non-impact interaction variance. That design is referred to as a repeated measures design (e.g., Winer, 1971; Green, 1989). In a repeated measures design, replicate sampling locations (sub-sites) would be chosen at random at each site. In subsequent years, these sub-sites would be re-sampled (hence repeated measures), as opposed to the current Sentar study (and many others), in which a new set of randomly selected sub-sites is sampled (referred to as re-randomization). The problem with re-randomization is that it cannot provide a measure of the natural variance of temporal changes because individual replicates or sub-sites are not followed over time. In a repeated measures design, individual sub-sites are monitored over time, and the temporal trends can be considered -plots of the variable of interest against time (time trajectories). The variance of these trajectories among sub-sites within sites is then used as an error term to test the significance of differences in time trajectories among sites (i.e., time X site interactions).

Green (1989; pers. comm.) argues that only repeated measures designs are appropriate for impact assessments and that re-randomization designs should be avoided. His argument has obvious implications for the ANC study on the Athabasca River, for other studies conducted by other consultants including EVS Consultants, and for the federal EEM program for monitoring pulp and paper mills (Environment Canada, 1991). Re-randomization designs are almost universal in environmental studies, although repeated measures designs are becoming more common in basic ecological studies. There are problems with repeated measures. First, natural differences in time trajectories among sites may still be significant relative to differences in time trajectories within sites. If that is the case, then our pseudo-*F* approach is actually superior to a repeated measures design because we are testing impact-related differences in time trajectories among sites against natural differences in time trajectories *among sites rather than within sites* (i.e, the pseudo-*F* tests account for additional variance among sites; repeated measures does not). Second, in a repeated measures design, sub-sites must be located far enough apart that they can be easily relocated and re-sampled, and so that differences among sub-sites are meaningful, yet the distances among sub-sites must be small relative to the distances among sites. The appropriate spatial distribution of sites and sub-sites can be especially difficult to determine in a river, where there are natural longitudinal changes in physical conditions and effluent dilution. A repeated measures design might force artificial divisions onto a continuous longitudinal gradient. Third, switching to a repeated measures design means abandoning existing data. In the case of the Athabasca study, that would mean abandoning baseline (pre-operational) data which can never be collected again.

We would recommend repeated measures designs for most environmental studies, including those to be undertaken for the federal EEM programs. Even though tests similar to those used in repeated measures designs such as our pseudo-F test can be constructed from contrasts, a repeated measures analysis is simpler, more direct, and more elegant. We stand by the approach used in this report, but recognise that it is jury-rigged, unconventional, and not easy to apply to all studies. However, we would not recommend that ANC abandon the present design and data for the Athabasca River benthos studies, because it would mean abandoning baseline data. The data from the first few post-operational years are also valuable, because they represent another baseline if treatment improvements, process changes, expansion, or other changes are planned in the future. Analyses in the future can be strengthened by:

1. Pooling weak Impact contrasts with Non-Impact contrasts to increase d.f.

   The error d.f. for pseudo-F tests could be increased by pooling all impact contrasts with pseudo-F values less than, for example, 1.5 with the Non-Impact contrasts. As data from post-operational years increase, the error d.f. could be great enough to provide meaningful tests.

2. Subdivide the existing sites into two or three groups, and use the sites as replicates in a repeated measures design

   The existing seven sites could be grouped into upstream (Sites 1 and 2), near-field (Sites 3-5), and far-field (Sites 6 and 7) areas, and the site means used as replicates within the areas in a repeated measures design. This approach assumes that differences in time trajectories within these areas are random with respect to impacts. This assumption may be difficult to accept, considering that concentrations of effluent may vary considerably within the near-field area, and the concentration of other effluents may vary between Sites 6 and 7. However, the results of our analyses actually support the assumption. Basically, this approach is the same as the one discussed above, except that an *a priori* decision is made about which Impact contrasts are to be pooled with the Non-Impact contrasts. Specifically, the NF-linear and quadratic, and Between FF, Impact contrasts would be pooled with the Between US contrasts. Note that only the NF-linear contrast appears in Table 17, suggesting that most differences in time trajectories within areas are no greater than random natural differences.

We prefer the second approach suggested, if only to reduce the number of tables included in reports. Appendix C provides the analyses associated with that approach; conclusions did not differ from those presented elsewhere in this report. To be safe, a global pseudo-*F* test could be performed for the Impact contrasts involving within-area comparisons prior to pooling.

5.0 RECOMMENDATIONS

1. The present study design should be continued, but the analyses should proceed as if the design were a repeated measures, with sites as individual replicates within broader areas.

2. As suggested in the Recommendations for Task I, the spring sampling should be dropped if that is acceptable to Alberta Environment. The spring samples were not informative about impacts.

6.0 REFERENCES

Environment Canada. 1991. Technical guidance manual for aquatic environmental effects monitoring at pulp and paper mills. Volume 2: methodology. Draft rept. 199 pp.

Gibbons, W.N., M.D. Munn and M.D. Paine. 1992a. Guidelines for monitoring benthos in freshwater environments. Draft rept. prepared for Environment Canada, North Vancouver by EVS Consultants, North Vancouver. 81 pp.

Gibbons, W.N., M.D. Paine, W.M. Gibson, G.A. Vigers and K. Kingston. 1992b. Operational monitoring of the Lesser Slave River. Volume III - 1991. Draft rept. prepared for Slave Lake Pulp Corporation, Edmonton, Alberta, by EVS Consultants, North Vancouver, BC. 63 pp. + Tables + Fig. + App.

Green, R.H. 1989. Inference from observational data in environmental impact studies: what is valid? what is possible? Proc. 47th Session, Intl. Statist. Inst., Invited Papers, Paris (Aug. 29 - Sept. 6, 1989).

Hilsenhoff, W.L. 1987. An improved biotic index of organic stream pollution. Great Lakes Entomol. 20: 31-39.

Hilsenhoff, W.L. 1988. Rapid field assessment of organic pollution with a family-level biotic index. J. North Amer. Benthol. Soc. 7: 65-68.

Klemm, D.J., P.A. Lewis, F. Fulk and J.M. Lazorchak. 1990. Macroinvertebrate field and laboratory methods for evaluating the biological integrity of surface waters. U.S. Environmental Protection Agency, Rept No. EPA/600-4-90/030. Environmental Monitoring Systems Laboratory, Office of Research and Development, Cincinnati. 256 pp.

Luoma, M.E. and B.M. Shelast. 1988. A baseline benthic invertebrate monitoring study on the Athabasca and McLeod Rivers, near Whitecourt, Alberta, 1987. Report prepared for Millar Western Pulp Ltd., Edmonton, Alberta, by Beak Associates Consulting, Calgary, Alberta. 44 pp. + App.

Luoma, M.E. and B.M. Shelast. 1989. A benthic invertebrate monitoring study on the Athabasca and McLeod Rivers, Whitecourt, Alberta, 1988. Report prepared for Millar Western Pulp Ltd., Edmonton, Alberta, by Beak Associates Consulting, Calgary, Alberta. 47 pp. + App.

Luoma, M.E. and B.M. Shelast. 1990. A benthic invertebrate monitoring study and fish habitat assessment on the Athabasca River, Whitecourt, Alberta, 1989. Report prepared for Alberta Newsprint Co., Whitecourt, by Beak Associates Consulting, Calgary, Alberta.

Luoma, M.E. and B.M. Shelast. 1991. A benthic invertebrate monitoring study on the Athabasca River, Whitecourt, Alberta, 1990. Report prepared for Alberta Newsprint Co., Whitecourt, by Beak Associates Consulting, Calgary, Alberta. 62 pp. + App.

Luoma, M.E. and B.M. Shelast. 1992. A benthic invertebrate monitoring study and fish habitat assessment on the Athabasca River, Whitecourt, Alberta, 1991. Report prepared for Alberta Newsprint Co., Whitecourt, by Sentar Consultants, Calgary, Alberta. 64 pp. + App.

Merritt, R.W. and K.W. Cummins (eds.). 1984. An introduction to the aquatic insects of North America. 2nd ed. Kendall Publ. Co., Dubuque, Iowa. 441 pp.

Sokal, R.R. and F.J. Rohlf. 1981. Biometry: the principles and practice of statistics in biological research. 2nd ed. W.H. Freeman and Co., New York. 859 pp.

Wilkinson, L. 1990. SYSTAT: the system for statistics. SYSTAT, Inc., Evanston, Illinois. 677 pp.

Winer, B.J. Statistical principles in experimental design. 2nd ed. McGraw-Hill, New York.

TABLES

Table 1. Fall. Contrasts used for analysis of temporal changes in spatial patterns. Contrasts for Year X Site interactions are formed by multiplying coefficients for temporal contrasts by coefficients for spatial contrasts (see Appendix A).

Contrast	Temporal (Year)	Spatial (Site)[1]	Comments
Impact			
1989 vs. 1990, 91	1/2 [1990 + 1991] - [1989)		Contrasts pre- and post-operationai years
DS - US		1/5 [3 + 4...+7] - 1/2 [1 + 2]	Contrasts sites downstream (DS) vs. upstream (US) of mill
NF - FF		1/3 [3 + 4 + 5] - 1/2 [6 + 7]	Contrasts near (NF) - and far-field (FF) sites
NF - linear		[3] - [5]	Tests for linear trend among near-field sites
NF - quadratic		[3] - 2 [4] + [5]	Tests for quadratic (curvilinear) trend among near-field sites
Between FF		[6] - [7]	Contrasts far-field sites
1990 vs. 1991	[1990] - [1991]	same 5 contrasts as above	Contrasts two post-operational years
Non-Impact			
Between US	1/2 [1990 + 1991] - [1989]	[1] - [2]	Tests for temporal changes in difference between two US sites
	[1990] - [1991]	[1] - [2]	Tests for temporal changes in difference between two US sites

1 Numbers in square brackets are site numbers.

3/561-01.2
92/07/13

Table 2. Spring. Contrasts used for analysis of temporal changes in spatial patterns. Contrasts for Year X Site interactions are formed by multiplying coefficients for temporal contrasts by coefficients for spatial contrasts (see Appendix A).

Contrast	Temporal (Year)	Spatial (Site)[1]	Comments
Impact			
1989 vs. 1990, 91	[1990] - 1/2 [1989 + 1990]		Contrasts pre- and post-operational years
DS - US		1/5 [3 + 4...+7] - 1/2 [1 + 2]	Contrasts sites downstream (DS) vs. upstream (US) of mill
NF - FF		1/3 [3 + 4 + 5] - 1/2 [6 + 7]	Contrasts near (NF) - and far-field (FF) sites
NF - linear		[3] - [5]	Tests for linear trend among near-field sites
NF - quadratic		[3] - 2 [4] + [5]	Tests for quadratic (curvilinear) trend among near-field sites
Between FF		[6] - [7]	Contrasts far-field sites
Non-Impact	[1990] - [1989]	same 5 contrasts as above	Contrasts two pre-operational years
	[1990] - [1989]	[1] - [2]	Tests for temporal changes in differences between two US sites
	[1991] - 1/2 [1989 + 1990]	[1] - [2]	Tests for temporal changes in differences between two US sites

1 Numbers in square brackets are site numbers.

Table 3. Fall. Dominant taxa, ranked according to % of total abundance in all samples from 1989-91.

Taxon[1]	% of Total Abundance	% Occurrence	Feeding Group[2]	Tolerance Values	
				Klemm et al.[3] (1990)	Hilsenhoff[4] (1987, 1988)
Cricotopus/Orthocladius (C)	50.0	100	DH	0 - 3	6 - 7
Ephemerella inermis (E)	8.7	100	DH	0	1
Rheotanytarsus (C)	7.1	97	D	3	6
Capniidae (P)	3.9	99	D	-	1
Taenionema (P)	3.9	93	H	-	2[5]
Micropsectra (C)	3.7	69	D	0 - 3	7
Baetis (E)	3.4	93	DH	0 - 3	2 - 6
Rhithrogena (E)	2.9	100	DH	0 - 2	0
Tvetinia (C)	2.1	97	D	-	5
Naididae (O)	1.7	83	D	1 - 5	-

1 Higher taxonomic levels given in parentheses: C = Chironomidae; E = Ephemeroptera; O = Oligochaeta; P = Plecoptera

2 D = detritivore; H = herbivore (from Merritt and Cummins, 1984 as adapted by Luoma and Shelast, 1992)

3 Scale: 0 (intolerant) to 5 (tolerant)

4 Scale: 0 (intolerant) to 10 (tolerant)

5 Tolerance Value for family Taeniopterygidae

Table 4. Fall. Results of univariate (ANOVA, ANCOVA) analyses of changes in total abundance and taxon richness.

Source	d.f.	*P*			
		Total *N*	Richness	Adj. richness[1]	EPT taxa
Year	2	****	*	***	***
Site	6	****	****	****	**
Year X Site	12	****	****	****	***
Contrasts					
Impact					
1989 vs. 1990, 91					
DS - US	1	****	NS	NS	NS
NF - FF	1	***	NS	**	NS
NF - linear	1	***	***	*	NS
NF - quadratic	1	****	NS	NS	NS
Between FF	1	NS	NS	NS	NS
1990 vs. 1991 (pooled)	5	****	****	****	***
Non-Impact					
Between US	2	****	**	**	*

Note: * - $P \leq 0.05$
** - $P \leq 0.01$
*** - $P \leq 0.001$
**** - $P \leq 0.0001$
NS - Not significant ($P > 0.05$)

1 From ANCOVA with total abundance as covariate.

3/561-01.2
92/07/13

EVS

Table 5. Fall. Sums-of-squares for Impact and Non-Impact contrasts.

Contrast	d.f.	Sums-of-squares (%)						
		Expected	Total *N*	Richness	Adj. Richness	EPT taxa	EPT (*N*)	Chironomidae (*N*)
Impact								
1989 vs 1990-91								
DS - US	1	8.3	20.0	3.2	1.0	1.4	6.1	18.4
NF - FF	1	8.3	7.1	2.6	10.7	1.3	0.0	15.0
NF - linear	1	8.3	8.4	15.8	4.0	0.9	3.0	13.4
NF - quadratic	1	8.3	12.8	2.6	0.5	1.5	20.1	7.0
Between FF	1	8.3	0.0	0.8	0.9	0.4	13.9	0.0
Total	5	41.7	48.3	24.9	17.1	5.4	43.0	53.8
1990 vs. 91 (pooled)	5	41.7	33.7	61.3	72.3	70.3	25.1	32.1
Non-Impact Between US	2	16.7	18.0	13.9	10.4	24.3	31.9	14.1
TOTAL	12	100.0	100.0	100.0	99.8	100.0	100.0	100.0

Table 6. Fall. Mean abundance, richness and number of EPT taxa.

Variable	Site	Year		
		1989	1990	1991
Total Abundance (per sample)	1	335	151	529
	2	460	873	1359
	3	452	958	701
	4	294	1866	2209
	5	305	1857	905
	6	654	1352	1639
	7	791	1617	1986
Richness	1	24	21	19
	2	26	29	29
	3	25	23	18
	4	23	23	27
	5	21	24	24
	6	24	21	32
	7	26	26	36
Adjusted Richness	1	29	29	20
	2	29	28	26
	3	28	22	18
	4	28	19	22
	5	25	21	24
	6	25	19	28
	7	27	20	30
EPT taxa	1	11	9	9
	2	9	11	10
	3	9	10	10
	4	10	11	13
	5	10	11	11
	6	11	9	13
	7	11	10	15

Note: Means for total abundance, richness, adjusted richness are geometric means; means for EPT taxa are back-transformed from log $(x + 1)$.

Table 7. Fall. Results of MANOVA and ANOVA examining changes in abundance of EPT taxa and Chironomidae.

Source	MANOVA (λ)[1]		ANOVA	
	F	*P*	EPT taxa (*P*)	Chironomidae (*P*)
Year	118.8	****	****	****
Site	29.5	****	****	****
Year X Site	14.5	****	****	****
Contrasts				
Impact				
1989 vs 1990,91				
DS - US	34.0	****	NS	****
NF - FF	34.2	****	NS	****
NF - linear	25.6	****	NS	****
NF - quadratic	11.5	****	***	****
Between FF	5.8	**	**	NS
1990 vs 1991 (pooled)	14.4	****	*	****
Non-Impact				
Between US	12.7	****	***	****

Note:
* - $P \leq 0.05$
** - $P \leq 0.01$
*** - $P \leq 0.001$
**** - $P \leq 0.0001$
NS - Not significant ($P > 0.05$)

[1] Wilks' λ is test statistic for multivariate differences

Table 8. Fall. Standardized coefficients (SC) and loadings (correlations, *r*) for EPT and Chironomidae abundances for vectors generated by multivariate contrasts.

Contrast	SC		*r*	
	EPT	Chir	EPT	Chir
Impact				
1989 vs. 1990-91				
DS - US	-44	118	22	93
NF - FF	-65	121	2	84
NF - linear	-49	119	18	91
NF - quadratic	20	88	69	99
Between FF	120	-72	80	-5
1990 vs. 1991 (pooled)				
Vector 1	-79	120	-12	75
Vector 2	91	15	99	66
Non-Impact				
Between US				
Vector 1	-11	106	48	100
Vector 2	120	-58	88	9

Note: SC and *r* are multiplied by 100.

Table 9. Fall. Mean raw abundance (no./sample) of EPT taxa and Chironomidae, and relative abundance (% of sample total) of EPT taxa.

Variable	Site	Year		
		1989	1990	1991
EPT (*N*)	1	192	91	288
	2	212	259	257
	3	204	131	257
	4	176	336	444
	5	210	184	369
	6	238	129	374
	7	225	404	437
Chironomidae (*N*)	1	99	46	218
	2	188	577	954
	3	194	814	422
	4	85	1474	1637
	5	60	1651	483
	6	360	1197	1103
	7	408	1164	1353
EPT (%)	1	58	61	55
	2	46	30	22
	3	46	14	37
	4	60	18	20
	5	69	10	41
	6	37	10	23
	7	29	25	22

Note: Raw abundances are geometric means.

Table 10. Spring. Dominant taxa, ranked according to % of total abundance in all samples from 1989 -1991.

Taxon[1]	% of Total Abundance	% Occurrence	Feeding Group[2]	Tolerance Values	
				Klemm et al.[3] (1990)	Hilsenhoff[4] (1987, 1988)
Cricotopus/Orthocladius (C)	43.6	94	DH	0 - 3	6 - 7
Eukiefferiella (C)	8.5	50	DH	0[5]	8
Hydropsyche (T)	5.0	44	O	0 - 3	0 - 7
Naididae (O)	4.4	68	O	1 - 5	-
Baetis (E)	4.4	82	DH	0 - 3	2 - 6
Rhithrogena (E)	3.8	81	DH	0 - 2	0
Brachycentrus (T)	3.2	51	O	0 - 1	0 - 1
Enchytraeidae (O)	3.0	67	O	-	-
Ephemerella inermis (E)	2.7	77	DH	0	1
Isoperla (P)	2.6	51	C	0 - 2	0 - 5

1 Higher taxonomic levels in parentheses: C = Chironomidae; E = Ephemeroptera; O = Oligochaeta; P = Plecoptera; T = Trichoptera
2 C = carnivore; D = detritivore; H = herbivore; O = omnivore (from Merritt and Cummins, 1984 as adapted by Luoma and Shelast, 1992)
3 Scale: 0 (intolerant) to 5 (tolerant)
4 Scale: 0 (intolerant) to 10 (tolerant)
5 Value for one species only provided; there are >20 species in the genus.

Table 11. Spring. Results of univariate (ANOVA) analyses of changes in total abundance and taxon richness.

Source	d.f.	*P*		
		Total *N*	Richness	EPT taxa
Year	2	****	****	****
Site	6	****	****	***
Year X Site	12	****	****	****
Contrasts				
Impact				
1989-90 vs. 1991				
DS - US	1	***	NS	NS
NF - FF	1	NS	NS	NS
NF - linear	1	***	***	NS
NF - quadratic	1	****	****	***
Between FF	1	NS	NS	NS
Non-Impact (pooled)	7	****	****	****

Note: * - $P \leq 0.05$
** - $P \leq 0.01$
*** - $P \leq 0.001$
**** - $P \leq 0.0001$
NS - Not significant ($P > 0.05$)

Table 12. Spring. Sums-of-squares for Impact and Non-Impact contrasts.

Contrast	d.f.	Sums-of-squares (%)					
		Expected	Total *N*	Richness	EPT taxa	EPT (*N*)	Chironomidae (*N*)
Impact							
1989-90 vs. 1991							
DS - US	1	8.3	4.4	0.9	0.4	1.0	4.6
NF - FF	1	8.3	0.0	0.2	0.0	0.5	0.9
NF - linear	1	8.3	4.1	6.2	4.8	2.5	3.1
NF - quadratic	1	8.3	7.3	14.2	15.2	10.1	0.4
Between FF	1	8.3	0.1	0.2	1.7	2.1	0.5
Total Impact	5	41.7	15.8	21.6	22.2	16.3	9.6
Non-Impact	7	58.3	84.2	78.4	77.8	83.7	90.4
TOTAL	12	100.0	100.0	100.0	100.0	100.0	100.0

Table 13. Spring. Mean abundance, richness, and number of EPT taxa.

Variable	Site	Year		
		1989	1990	1991
Total Abundance (per sample)	1	173	54	20
	2	277	106	11
	3	143	177	8
	4	115	453	6
	5	437	60	26
	6	601	62	12
	7	1988	52	18
Richness	1	21	13	9
	2	25	19	6
	3	22	18	5
	4	22	30	4
	5	20	11	9
	6	31	12	7
	7	32	13	10
EPT taxa	1	9	4	2
	2	10	6	1
	3	9	7	2
	4	11	10	2
	5	10	5	4
	6	14	4	2
	7	14	5	4

Note: Means for abundance are geometric means.

Table 14. Spring. Results of MANOVA and ANOVA examining changes in abundance of EPT taxa and Chironomidae.

Source	MANOVA (λ)[1]		ANOVA	
	F	*P*	EPT taxa (*P*)	Chironomidae (*P*)
Year	164.7	****	****	****
Site	10.0	****	****	****
Year X Site	16.1	****	****	****
Contrasts				
Impact				
1989-90 vs. 1991				
DS - US	13.4	****	NS	****
NF - FF	3.3	*	NS	NS
NF - linear	5.9	**	*	***
NF - quadratic	9.2	***	****	NS
Between FF	4.5	*	NS	NS
Non-Impact (pooled)	21.1	****	****	****

Note: * - $P \leq 0.05$
** - $P \leq 0.01$
*** - $P \leq 0.001$
**** - $P \leq 0.0001$
NS - Not significant ($P > 0.05$)

[1] Wilks' λ is test statistic for multivariate differences.

Table 15. Spring. Standardized coefficients (SC) and loadings (correlation, *r*) for EPT and Chironomidae abundances for vectors generated by multivariate contrasts.

Contrast	SC		*r*	
	EPT	Chir	EPT	Chir
Impact				
1989-90 vs. 1991				
DS - US	-67	105	-26	78
NF - FF	-78	100	-39	69
NF - linear	27	86	61	97
NF - quadratic	104	-12	99	29
Between FF	-97	84	-64	45
Non-Impact (pooled)				
Vector 1	26	87	60	97
Vector 2	106	-65	80	-24

Note: SC and *r* multiplied by 100.

Table 16. Spring. Mean raw abundance (no./sample) of EPT taxa and Chironomidae and relative abundance (% of sample total) of EPT taxa.

Variable	Site	Year		
		1989	1990	1991
EPT (N)	1	85	11	2
	2	80	25	2
	3	65	34	3
	4	80	101	2
	5	113	23	8
	6	221	7	2
	7	345	12	8
Chironomidae (N)	1	58	33	13
	2	116	60	7
	3	46	128	3
	4	10	287	2
	5	281	27	15
	6	270	51	7
	7	1431	30	7
EPT (%)	1	51	21	13
	2	29	25	14
	3	46	20	43
	4	70	23	50
	5	27	39	33
	6	38	12	18
	7	18	24	48

Note: Raw abundance means back-transformed from log $(x + 1)$.

Table 17. Impact contrasts with pseudo-F > 2 (see text for derivation of pseudo-F values). These values indicate the magnitude of impact-related changes relative to non-impact or natural changes.

Variable	Contrast		Pseudo-F
	Year	Site	
Total Abundance	1989 vs. 1990, 91	DS-US	2.2
Richness	1989 vs. 1990, 91	NF-linear	2.3
Adjusted Richness	1989 vs. 1990, 91 1990 vs. 1991	NF-FF pooled	2.1 2.8
Chironomidae (N)	1989 vs. 1990, 91	DS-US NF-FF	2.6 2.1
Multivariate	1989 vs. 1990, 91	DS-US NF-FF NF-linear	2.7 2.7 2.0

Note: All values are for fall data; no pseudo-F values for spring impact contrasts were >2.

FIGURES

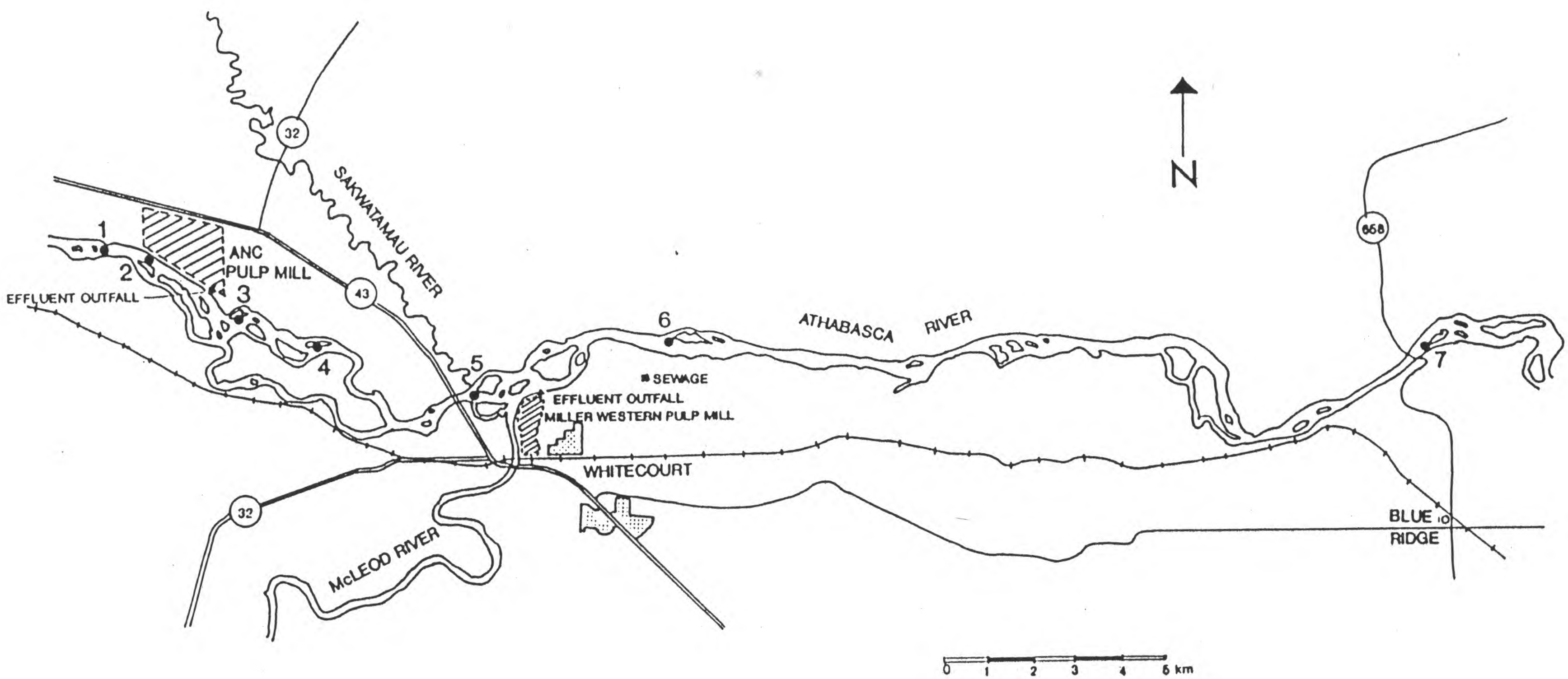

Figure 1. Benthic invertebrate sampling sites on the Athabasca River (reproduced with permission of Sentar Consultants Ltd.)

APPENDIX A
CONSTRUCTION OF INTERACTION CONTRASTS

This Appendix briefly describes how contrasts for the interaction between Year and Site were constructed. In a contrast, Year and/or Site means are assigned coefficients; the sum of the coefficients equals 0. For example, a pre- *versus* post-operational contrast for Year is shown in boldface in Table A.1. Contrasts can also be constructed for Site, and the upstream minus downstream (US-DS) spatial contrast is also shown in boldface in Table A.1. To construct a contrast which compares the difference between sites upstream and downstream of the ANC mill before and after start-up, the coefficients for the Year contrast are multiplied by the coefficients for the Site contrast, as shown in Table A.1. Additional contrasts can be constructed in the same manner.

Statistical texts provide rules to ensure that all contrasts constructed are independent or orthogonal (e.g., Sokal and Rohlf, 1981). Basically, a table is set up, with the coefficients for each contrast as rows, and the cells (combinations of Year and Site) as columns. The coefficients in each column are multiplied together, and the sum of their products will sum to 0 if the contrasts are orthogonal. In practice, it is simpler to construct sets of orthogonal contrasts for Year and for Site separately, and to multiply the coefficients of all possible Year and Site contrast combinations to produce a set of orthogonal interaction contrasts, than to attempt to directly construct an entire set of orthogonal contrasts for the interaction.

Table A.1. Construction of interaction contrasts. The temporal (Year) and spatial (Site) coefficients in **bold face are multiplied to provide the interaction coefficients.**

		Site						
Year		1	2	3	4	5	6	7
		(-1/2)	**(-1/2)**	**(1/5)**	**(1/5)**	**(1/5)**	**(1/5)**	**(1/5)**
1989	**(-1)**	1/2	1/2	-1/5	-1/5	-1/5	-1/5	-1/5
1990	**(1/2)**	-1/4	-1/4	1/10	1/10	1/10	1/10	1/10
1991	**(1/2)**	-1/4	-1/4	1/10	1/10	1/10	1/10	1/10

Note: Contrast is pre-versus post-operational change in difference between stations upstream versus downstream of the ANC mill.

APPENDIX B
ANALYSIS OF DOMINANT TAXA IN FALL SAMPLES

This Appendix provides the results of analyses of abundances of eight dominant taxa in the fall samples. The taxa included were the taxa listed in Table 3, with the two taxa (*Micropsectra* and Naididae) occurring in <90% of the samples excluded. The objective of the analyses were to determine if analyzing the eight taxa separately produced results similar to those obtained when taxa were pooled into EPT and Chironomidae. The methods were the same as those used to analyze EPT and Chironomidae abundances, except that there were eight variables, rather than two.

Table B.1 provides the results of multivariate contrasts testing for various impact and non-impact differences. All contrasts were significant at $P<0.05$; all but one were significant at $P<0.0001$. F-values for all but one of the Impact contrasts were greater than the F-value for the Non-Impact contrasts, indicating that there were impacts greater in magnitude than natural temporal changes in spatial differences. Two of the Impact contrasts (1989 vs. 1990,91 DS-US and NF-linear) had pseudo-F values (F divided by Non-Impact F) >2, suggesting some overall changes downstream of the mill with a linear decrease (or increase) in these changes among the near-field sites. These results were roughly similar to those obtained using EPT and Chironomidae abundance (Table 7), except that the rankings of the Impact contrasts differed somewhat between the two analyses.

Table B.2 provides loadings for the eight taxa abundances for the vectors generated by the contrasts. There are two things to look for in these loadings:

- the same sign for loadings for all taxa, which would indicate a general increase or decrease in abundance associated with the contrast

- absolute value of loadings for all EPT taxa (first five listed) near 0 and lower than the absolute values of loadings for Chironomidae (last three taxa), which would indicate little change in abundance of EPT and large changes in abundance of Chironomidae (the dominant trend observed in the analyses in Section 3.2)

If a general abundance vector is defined as one for which at least six taxa, including *Cricotopus/Orthocladius*, have the same sign, there were only two:

- 1989 vs. 1990,91 NF-quadratic

- Non-Impact Vector 1

The first contrast listed also produced a vector with high positive loadings when only EPT and Chironomidae were analyzed (Table 8). In general, the analyses of the eight dominant taxa did not identify the same trends as were identified by the analyses of total abundance (see Section 3.1.1). Total abundance depended primarily on the abundance of *Cricotopus/Orthocladius*, which accounted for 50% of the invertebrates collected, so any vector correlated with that taxon is likely to reflect total abundance regardless of loadings for other taxa. Thus, for example, Vector 1 from the pooled 1990 vs. 1991 contrast was strongly correlated with *Cricotopus/Orthocladius* and no other taxa, and would probably be strongly correlated with total abundance.

There were no vectors for which absolute values of loadings for EPT were low and those for Chironomidae high, except perhaps for:

- 1989 vs. 1990,91 DS-US

- 1990 vs. 1991 pooled (Vector 1)

- 1990 vs. 1991 pooled (Vector 3)

Loadings for *Cricotopus/Orthocladius* were high for the first two vectors listed above, suggesting that changes in abundance of this taxon may have accounted for most of the impact-related increases in Chironomidae abundance. Loadings for all or most EPT taxa were rarely of the same sign, except for the general abundance vectors, partly because many of the loadings were weak and distributed around 0. However, there were some instances in which loadings for a pair of EPT taxa were strong, but of opposite signs. For example, loadings for the two Plecopterans, Capniidae and *Taenionema*, were either of opposite signs or both near 0 for all of the Impact vectors. Those results suggest that the abundance of one of the taxa may have been reduced by impacts, but that the abundance of the other increased, possibly because of decreased interspecific competition. Thus, it is possible that no strong impacts were observed on abundance of EPT as a whole, because reductions in abundance of some taxa were compensated for by increases in abundance of their competitors.

Table B.1. Fall. Results of MANOVA examining changes in abundances of 8 dominant taxa.

Source	$F(\lambda)$[1]	P
Year	94.7	****
Site	17.2	****
Year X Site	7.7	****
Contrasts		
Impact		
1989 vs. 1990, 91		
DS - US	24.1	****
NF - FF	9.8	****
NF - linear	15.3	****
NF - quadratic	3.2	**
Between FF	9.6	****
1990 vs. 1991 (pooled)	7.1	****
Non-Impact Between US	6.4	****

Note: * - $P \leq 0.05$
** - $P \leq 0.01$
*** - $P \leq 0.001$
**** - $P \leq 0.0001$

[1] F - value for Wilks' λ

Table B.2. Fall. Loadings (correlation, r) for dominant taxa abundances for vectors generated by multivariate contrasts.

Contrast	Ephem	Rhith	Cap	Taen	Baetis	Rheo	C/O	Tvet
Impact								
1989 vs. 1990-91								
DS - US	1	15	33	-37	-22	-33	-67	-33
NF - FF	26	18	-33	-16	-10	44	-40	3
NF - linear	40	-1	35	-36	-47	17	37	18
NF - quadratic	55	51	5	14	18	55	44	80
Between FF	30	19	-22	33	68	-2	-13	12
1990 vs. 1991 (pooled)								
Vector 1	10	-2	-9	-6	-15	27	76	1
Vector 2	45	-2	9	-8	-46	47	-24	10
Vector 3	51	30	40	-7	26	-30	-7	-58
Non-Impact								
Between US								
Vector 1	32	56	14	41	64	47	85	41
Vector 2	10	-30	-24	-46	-47	40	24	-18

Note: r multiplied by 100.

APPENDIX C
REPEATED MEASURES ANALYSIS

This Appendix provides the results of analysis of the ANC benthic community data using a repeated measures design.

1.0 Methods

To conform with the requirements for a repeated measures design, the study sites were grouped into three areas:

Upstream (Reference)	Sites 1,2
Near-field	Sites 3,4,5
Far-field	Sites 6,7

The sites were considered replicates within areas, re-sampled each year. Site means were used as individual observations, providing a total of 21 observations for any analysis. The ANOVA table for the analysis was:

Source	d.f.
Area	2
Site{Area}	4
Year	2
Year*Area	4
Year*Site{Area} = Error	8

The Area effect is tested against the Site{Area} effect; all other effects are tested against the Year*Site{Area} effect or the Error. Note that the Year*Site{Area} effect represents variation in time trajectories among sites within areas; the Year*Area effect represents variations in time trajectories among areas. The Year*Area interaction can be broken down into contrasts, corresponding to those used in the EVS report.

2.0 Results

Tables 1 and 2 provide the results of the analyses. Note that a liberal *P* of 0.10, rather than 0.05, was used as a significance level because the limited Error d.f. reduced the power of the tests.

For the fall data, only the contrasts comparing the two post-operational years were significant. These results indicate that the impacts noted (increase in total abundance, chironomid abundance; decrease in adjusted richness) occurred largely between 1989 and 1990 (first operational year), and that there was a subsequent recovery. The one exception was the number of EPT taxa, which increased downstream of the mill in 1991, after very little change between 1989 and 1990. This could be considered a positive impact, although it may have more to do with decreased impacts from Millar Western or the Whitecourt sewage treatment plant than decreased impacts from the ANC mill.

For the spring data, only contrasts comparing the difference between near- and far-field sites between the two pre-operational years were significant. These results strongly suggest reduced impacts from either Millar Western or the Whitecourt sewage treatment plant, with no evidence of effects from the ANC discharge.

3.0 Conclusions

The repeated measures analysis identified the same impacts for the fall data set as did the EVS report did, except that recovery in 1991 was more apparent. The analysis of the spring data set strongly suggested some reductions in impacts from downstream sources.

Table C.1. Fall samples. Results of repeated measures analysis of benthic community variables. Values are probabilities (P); P ≤0.10 are given in boldface.

Contrast		Variable					
Temporal	Spatial	Total *N*	Richness	Adj. Richness	EPT Taxa	Chironomidae (*N*)	EPT (*N*)
1989 vs. 1990,91	DS - US	0.12	0.45	0.73	0.64	0.13	0.50
	NF - FF	0.30	0.16	0.16	0.61	0.13	0.95
1990 vs. 1991	DS - US	**0.10**	0.21	**0.03**	**0.04**	**0.09**	0.98
	NF - FF	0.38	**0.04**	**0.06**	**0.02**	0.42	0.97
Year * Area Interaction		0.14	0.17	**0.06**	**0.07**	**0.10**	0.97

Table C.2. Spring samples. Results of repeated measures analysis of benthic community variables. Values are probabilities (P); P ≤0.10 are given in boldface.

Contrast		Variable				
Temporal	Spatial	Total *N*	Richness	EPT Taxa	Chironomidae (*N*)	EPT (*N*)
1989,90 vs. 1991	DS - US	0.42	0.73	0.76	0.42	0.60
	NF - FF	1.00	0.85	0.94	0.73	0.72
1989 vs. 1990	DS - US	0.62	0.52	0.29	0.76	0.50
	NF - FF	**0.03**	**0.02**	**<0.01**	**0.05**	**<0.01**
Year * Area Interaction		0.18	0.17	**0.05**	0.28	**0.04**