A PROTOCOL FOR MONITORING AND ASSESSMENT OF WATER QUALITY IN AGRICULTURAL STREAMS USING BENTHIC INVERTEBRATES

D.R. Barton and M.E. Farmer

Department of Biology, University of Waterloo, Waterloo, Ontario, Canada, N2L 3G1.

M.E. Dillon and D.R. Oliver

Centre for Land and Biological Research Resources,
Agriculture Canada,
Ottawa, Ontario, Canada K1A 0C6.

1.0 RATIONALE/OBJECTIVE

1.1 OBJECTIVE

The primary objective was to develop a set of techniques, a protocol, for monitoring and assessment of surface water quality in agricultural regions of southern Ontario based on benthic invertebrates.

This involved determining the best way to collect appropriate samples, how closely the animals should be identified, and how the results can be interpreted. The protocol will be used to assess the status of small streams relative to other streams draining land under the same crop, or to monitor changes in water quality due to changes in farming practices. The assessment function is intended to identify sites which are better or worse than the average for southern Ontario; monitoring will be valuable in evaluating the results of mitigative measures such as leaving riparian buffer strips or adopting conservation tillage. It was tested in a study of four pairs of streams. One of each pair drained land primarily under conservation tillage and the other under conventional tillage.

1.2 RATIONALE

Agriculture has profound local and regional effects on both the quality and quantity of surface waters. Removal of the forest which originally covered most of southern Ontario led to higher summer and lower winter water temperatures, as well as destabilization of hydrologic regimes, increased erosion and siltation. Subsequent use of this cleared land for agriculture has additional effects on the quality of surface waters, with the nature and magnitude of the effects reflecting different farming practices.

In addition to each of the basic suites of stresses which are characteristic of each major kind of crop, the relative impact of fields under the same crop may differ substantially because of individual farming practices, variations in soil conditions or chance events such as heavy rain just after application of fertilizers. Such variation can make it very difficult, and expensive, to monitor the effects of agricultural practices on surface water quality through chemical assays. Pulsed inputs of chemicals resulting from spills or heavy precipitation may easily be missed in routine surveys, but may have severe biological impacts.

The problem of rapid fluctuations in the chemistry of running waters has long been recognized with respect to discharges of wastes from industrial and municipal sources, so the alternative approach of measuring biological effects directly has been widely accepted (Rosenberg and Resh 1992). Benthic invertebrate communities are nearly ideal for this purpose because they are easily sampled and include large numbers of individuals belonging to many species, each with its own response to any particular type of stress. These animals live on the bottom of the stream or lake and have life-cycles lasting months to years, so are exposed to the conditions in the stream over fairly long periods of time.

Schindler (1987) and Gray (1989) have suggested that biological surveillance of communities, characterizing taxonomic richness and composition, is a very sensitive tool for the detection of alterations in aquatic ecosystems.

For point sources, the usual approach is to sample upstream and downstream of the point of discharge and to express the differences in terms of changes in total abundance, numbers of species, some expression of "diversity", or a biotic index. In such situations, the benthic fauna reflects average conditions as well as intermittent discharges, and biological monitoring tends to be much cheaper and more sensitive than routine chemical analyses.

The potential utility of benthic invertebrates for monitoring and assessment of the effects of agriculture on water quality should be equally great, but the situation differs in one very fundamental way. With a few exceptions such as runoff from feed lots or manure storage, agricultural effects do not originate as point sources, thus upstream-downstream comparisons are rarely practical. Furthermore, agricultural activities tend to inflict a number of stresses simultaneously so it is not easy to ascribe the composition of the fauna at any particular point to any single aspect of farming practice.

2.0 METHODOLOGY

Sampling sites were located throughout most of the major agricultural areas of southern Ontario (Fig. 1). Each of the more than 250 sites was visited at least once, most on two or more dates, during 1990-91 and 1993. Small streams were selected to isolate the effects of practices associated with specific crops such as corn, soybeans, fruit, hay, cattle and tobacco. Streams draining forested non-agricultural land were included as reference sites; all sampling sites were upstream of any urbanized areas. Four pairs of streams which were part of studies of the effects of conventional and conservation tillage were sampled intensively during 1991-92.

On each visit we noted the crop(s) being grown upstream of the site, the nature of the stream channel and substratum, and the dimensions and composition of any riparian vegetation.

Invertebrates were collected and processed from each site using a number of techniques in order to determine the most efficient sampling methods. Collecting techniques included: - qualitative dip net samples using both coarse (1 mm aperture) and fine (200 or 300 Fm) mesh, - quantitative Surber samples, - surface skimming for pupal exuviae, - sweep netting and emergence traps for aerial stages of aquatic insects. Animals were sorted from these samples either in the field (some dip net collections only) or in the laboratory under a dissecting microscope. Different numbers of animals were removed from some of the qualitative samples in order to assess the minimum number needed to estimate the number of taxa present at a site.

Considerable effort was invested in associating the immature and adult stages of aquatic insects for taxonomic study which allows more accurate and complete identification of specimens. This new information (Oliver and Dillon 1994), as well as published keys and descriptions, were used to identify all of the arthropods and molluscs in each collection to the lowest practical taxonomic level (mainly to genus, species where possible).

The animals found in each sample were used to compile lists of taxa and their percentage contributions to the site totals. Reference and agricultural sites were compared using a variety of summary measures to determine which measure best discriminated among sites.

3.0 FINDINGS

More than 475 taxa were identified from all of the material collected during this study (list available on request). Most of these were insects (43 Ephemeroptera, 32 Odonata, 23 Plecoptera, 17 Hemiptera, 65 Trichoptera, 24 Coleoptera, 141 Chironomidae and 33 other Diptera). Mollusca (28 species) and Crustacea (19 species) accounted for most of the rest. The only group of animals found in every sample containing more than 100 individuals was Chironomidae.

The Chironomidae were studied intensively because of their importance as the dominant component of the fauna in most streams. This work has resulted in the preparation of a taxonomic key for the identification of pupal exuviae to species.

3.1 EVALUATION OF SAMPLING AND PROCESSING TECHNIQUES

The ideal sampling technique should be appropriate for all types of streams, simple to use with little detailed training and yield a maximum of information with minimum cost. Using these criteria with number of taxa per sample as the simplest measure of information content, we rejected quantitative sampling (Surber sampler, etc.) as being too laborious to be cost-effective. At the other extreme, aerial sweep-netting and picking live animals in the field required the least time and effort but yielded very erratic results, suitable only for the coarsest discrimination among sites. Emergence traps yielded valuable information when deployed over the entire open-water season but require weekly visits to change the collecting bottles, are subject to disturbance by spates, animals and vandals, and collect only insects; molluscs and crustaceans are dominant groups at many sites.

The technique which yielded the most consistent results was kick-sampling using a dip net equipped with 300 Fm aperture netting. Finer mesh tends to clog too quickly and the small proportion of animals which pass through 300 Fm tend to be too immature to identify. The procedure is to disturb portions of each microhabitat in the stream (mid-channel riffles and pools, marginal vegetation, undercut banks, etc.) with a booted foot and collect the dislodged animals with the net. The contents of the net are rinsed to remove fine sediments, discarding large pieces of vegetation, wood and stones. A portion (about 400 ml) is then placed in a 500 ml jar and preserved in 10% formalin.

In the laboratory, the sample jar is emptied into a fine-meshed net and rinsed to remove the preservative. The residue is then examined under a dissecting microscope and the first 200 animals encountered are removed to 70% ethanol for identification. This minimum sample size is based upon examination of the relationship shown in Figure 2. If new species are still being found regularly, sorting should continue. Any unsorted residue which remains should be examined briefly for large rare animals which may not have been included in the first 200 individuals.

The cast skins left behind by emerging adult Chironomidae tend to accumulate in eddies along the margins of streams. These exuviae are relatively easy to handle and identify, even to species, so are increasingly being used for water quality monitoring in various parts of the world. We found pupal exuviae in 49 of 140 (35%) samples collected from May through mid-August 1991; each of the 140 kick samples taken at the same sites contained chironomid larvae or pupae. Catches of exuviae appeared to be affected by weather, the physical structure of the stream channel and operator error, so this sampling technique is not satisfactory on its own for routine bioassessment (Barton et al. in press). Pooling of kick and exuvial samples increased the estimated species richness of Chironomidae by 10 to 28%, depending on landuse. Such combined estimates cost about 40% more (in collecting and processing time), so would not seem to be cost-effective. However, several of the numerically dominant genera of Chironomidae in agricultural streams include species which have very different habitat preferences. These cannot be identified as larvae, so the information gained from species-level identification of pupal exuviae is a valuable supplement to the results of kick sampling.

3.2 INDICES AND INTERPRETATIVE TECHNIQUES

In order to determine the most appropriate way to summarize the results of benthic samples for the purposes of monitoring and assessment, we examined several indices which have been widely adopted or recently proposed (Rosenberg and Resh 1993). These included number of taxa (S), a Biotic Index (HBI) (Hilsenhoff 1987), and Percent Model Affinity (PMA) (Novak and Bode 1992). Other indices were examined, but these three illustrate the general properties of each type. Values of each index were calculated for all kick samples containing >100 animals.

The number of kinds of animals in a sample (S) is the simplest measure of diversity, and S is typically reduced by environmental stress. Therefore, we might expect S to be highest in reference streams and to decrease with increasing intensity of agricultural effort. While values of S tended to be low in streams draining areas under fruit and market vegetables, there was no clear distinction between forested and agricultural streams during the summer (e.g. Fig. 3).

Hilsenhoff's Biotic Index expresses the average tolerance of arthropods in the benthic community to organic enrichment. Tolerances for individual taxa were derived in Wisconsin so included many of the same species found in southern Ontario. Hilsenhoff's Biotic Index has been applied with some success in the assessment of agricultural streams in North Carolina (D.R. Lenat, personal communication). On average, HBI scores were higher (i.e. water quality was poorer) in agricultural streams than in reference streams (Fig. 4), but, as with S, there were no clear distinctions among different landuses. Average values of HBI tended to be slightly higher in samples collected from October through May than in samples collected in summer.

Neither of these indices clearly distinguishes between reference and agricultural streams because each allows only a single dimension of response. Reference streams should be capable of supporting large numbers of species, therefore disturbances or stresses due to agricultural practices could only be detected using S if they reduce the number of species. Hilsenhoff's Biotic Index is designed to be sensitive to inputs of organic matter or nutrients. We examined the possibility of developing an Agricultural Biotic Index by ranking cropping systems in terms of the total energy input needed for the crop. This approach was rejected because there are clear differences between the effects of various stresses, but these are not additive. Physical disturbance, organic enrichment and insecticides all cause certain kinds of animals to become more, or less, abundant but any one species may respond in opposite ways to each individual stress.

The summary technique or index which we recommend is Percent Model Affinity, recently described by Novak and Bode (1992). PMA is an expression of the degree of similarity between the animals in a sample from the stream to be assessed and the average composition of reference streams. It is not necessary to assume that any particular agricultural practice is inherently better or worse than any other in terms of effect on water quality. The reference community may be from forested non-agricultural streams, or from other streams draining the same crop, type of soil or local area. Smaller values of PMA indicate greater impacts.

The Percent Model Affinity is calculated as:

$$PMA = 3 \min(sr, su)$$

where sr and su are the percentages of each taxon in the reference community and the sample, respectively. The basic reference community is the average composition (in %) of all samples from forested reference streams in each season. Values of PMA can range from 0 to 100. Confidence limits for reference sites were calculated from the PMA's among all reference sites. Significant deviation from the reference community (i.e. a significant reduction in water quality) occurs when PMA for a site is more than two standard deviations from the mean PMA for all reference sites.

Cluster Analysis was used to examine the reference streams for regional differences which might be expected to result from differences in soils, topography or micro-climate throughout the study area. No consistent differences were found, except that streams in the Niagara area had low diversity and were dominated by Crustacea. These sites were excluded for calculation of the expected reference community.

Evaluations based on PMA are influenced by the taxonomic level to which the animals are identified, and the time of year at which the samples were collected. The proportion (in percentage) of our samples from agricultural streams which were believed to be impacted illustrate this:

Level of Identification	Species	Genus	Family
June - August	83	78	44
October - May	47	41	3

These results further confirm that summer (June through August) is the best time to sample for biomonitoring purposes. Species which are most active during the warmer months appear to be more sensitive to agricultural runoff. Identification of animals to the family level requires less time than for genus or species but is not nearly so powerful at discriminating among streams, especially during the colder months. The relatively small gain in discrimination achieved by identification to the species level is misleading: it largely reflects the present lack of taxonomic information needed to identify many chironomid larvae beyond genus.

The magnitude of the impact on water quality in streams draining agricultural land varies with crop (Fig. 5), and at least some streams draining most of the major crops surveyed were not significantly different from reference streams. All streams draining orchards/vineyards ("Fruit") and market vegetables were impacted, i.e. had values of PMA significantly different from the mean community in reference streams.

The composition of the individual samples can often give an indication of the cause of the impact. For example, the animals found in Fruit streams suggest that the effects are due to insecticides. Among Pasture streams, the lowest values of PMA occurred in mud bottomed ditches subject to heavy grazing right to the water's edge. The dominant organisms in impacted pasture streams were burrowing forms which are very tolerant of low dissolved oxygen, indicating that the important factors are physical disturbance, siltation and organic enrichment.

Other factors which appear to influence the magnitude of agricultural impacts on water quality include the dimensions and composition of riparian vegetation, and whether the stream has been channelized, i.e. turned into a drain or ditch. Significantly low values of PMA (at the generic level) were calculated for 81% of ditched sites, and 72% of natural channels. Natural channels separated from grain fields by treed buffer strips >10 m wide were not distinguishable from forested streams.

3.3 CONVENTIONAL VS. CONSERVATION TILLAGE

This part of the study was undertaken to demonstrate the sensitivity of the invertebrate communities to changes in farming practices, namely the adoption of conservation (CONS) versus conventional (mostly mouldboard ploughing - CONV) tillage. Four pairs of streams in southwestern Ontario were examined (Fig. 1): one of each pair draining a basin under conventional farming systems (CONV) or conservation farming systems (CONS). The paired drainage basins were otherwise similar to each other in size, substratum, topography, land use, soil type and hydrology. The sites had been chosen for study by the Upper Thames River Conservation Authority or as part of the Soil and Water Environmental Enhancement Program (Ontario Ministry of the Environment).

Each site was visited monthly from May 1991 through August 1991, once in the fall of 1991 (October-November) and once in the winter of 1992 (February-March). On each visit both qualitative (kick) and quantitative (Surber) samples were collected. At least the first 100 arthropods were removed from kick samples for identification. Particularly rich Surber samples were sub-sampled, otherwise all animals were removed for identification. All specimens were identified to the lowest possible taxonomic level with the exception of the Chironomidae which were identified only to subfamily.

Three indices were used to compare the arthropods collected from CONS and CONV streams: number of taxa (S), Family-level Biotic Index (FBI) (Hilsenhoff 1988) and Percent Model Affinity (PMA). The use of S and FBI as measures of impact is appropriate here because we are comparing two treatments (tillage) in matched pairs of streams. This also gave us an opportunity to compare evaluations based on PMA with these more standard methods. The expected communities for calculation of PMA were adjusted to match the level to which the animals from these samples were identified.

The Kintore streams were the only pair which did not dry up during the very dry summer of 1991. In each of the other pairs, the CONV stream stopped flowing first. Samples from CONS streams yielded significantly larger values of S (Figs. 6 and 7) and lower FBI scores (Figs 8 and 9) than found from CONV streams. Both of these measures indicate better water quality in CONS streams. The largest and most consistent differences were observed from June through October.

The Percent Model Affinity produced similar results. The animals collected from CONS streams tended to be more similar (higher PMA) to reference streams than did their CONV counterparts (Fig. 10). Again, this pattern was most consistent from June through October.

All comparisons of the benthic arthropod communities in these paired streams suggested that conservation tillage reduces the impacts of agriculture on water quality. Streams draining fields under conservation tillage supported more kinds of animals, especially those belonging to groups less tolerant of organic enrichment, and their arthropod communities were more similar to those in non-agricultural reference streams. The degree of mitigation is greatest in summer and autumn: all streams were dominated by Chironomidae (mainly Orthocladiinae) in winter, though CONS streams supported more taxa even at that time of the year.

The different tillage practices did not result in measured differences in total suspended solids or inorganic nutrients. Perhaps biologically significant inputs were missed because of the sampling schedule for these parameters was fixed, or the animals are influenced by parameters which were not measured. The streams were not perfectly paired: for example, the Pittock CONV stream flows through drainage tiles before surfacing about 150 m above the sampling location. Similarly, a small wetland forest just above the Kettle CONV sampling site probably acts as a filter for some of the deleterious components of conventional farm runoff. These potentially remedial aspects make our conclusions about biological differences between CONS and CONV streams that much more conservative.

4.0 STUDY CONCLUSIONS

This study has demonstrated that the biodiversity of benthic invertebrates in small streams in agricultural southern Ontario is very high. These diverse communities respond in distinctive ways to inputs of eroded soil, pesticides, fertilizers and other organic matter. There is a strong seasonality in the composition of the benthic fauna, and those species which grow more actively during the warmer months appear to be more sensitive to agricultural activities than those which are most active in winter.

Percent Model Affinity proved to be the most versatile and sensitive way to summarize invertebrate data for the purposes of water quality monitoring and assessment. Samples from additional streams can be compared to the expected community from forested reference streams to identify any which are significantly impacted. Inspection of the detailed composition of the fauna usually gives an indication of the nature of the problem and the kind of subsequent detailed monitoring or analysis which might be warranted. Where remedial actions or changes in farming practices have occurred, monitoring of the biological effects would involve comparing benthic samples with the average communities from both forested streams and streams draining the same crops. Changes would be seen as decreasing PMA over time relative to the same crop; increasing PMA relative to forested streams would indicate improving water quality.

5.0 NEW TECHNOLOGY AND BENEFITS

The most important result of this work is the Monitoring Protocol which has been developed for invertebrate-based monitoring and assessment of surface water quality in southern Ontario. The procedure for this protocol is as follows:

1. Field Collecting.
   a. Record location of stream (e.g. UTM), adjacent landuse and general land use in area.
   b. Take sample by disturbing portions of all microhabitats (mid-channel riffles and pools, marginal vegetation, undercut banks, etc.) in the stream with a booted foot.
   c. Capture the dislodged substrate in a dip net with 300 Fm mesh bag.
   d. Rinse the contents of the net to remove fine sediments, discarding large pieces of vegetation, wood and stones.
   e. Save 400 ml of contents in net in a jar and preserve with 10% formalin.

2. Laboratory analysis
   a. Empty sample into a fine-meshed (100 Fm aperture) net; rinse to remove the preservative and fine sediments not washed through in the field.
   b. Examine sample under a dissecting microscope and remove, at least, the first 200 animals encountered.
   c. Briefly examine unsorted sample for large rare animals which may not have been included in the first 200 individuals.
   d. Preserve animals in 70% ethanol.
   e. Identify and enumerate the animals to the lowest practical taxonomic level.
   f. Convert the counts to percentages of the total number of animals sorted from the sample.
   g. Calculate the Percent Model Affinity (PMA) between the sample and the appropriate expected reference community.
   h. Values of PMA which are more than 2 standard deviations away from the mean for the reference community are significantly different.

The cost of invertebrate biomonitoring can be estimated from the time and skill of the personnel needed for each step. Collection of the sample requires about 10-15 minutes at the site. Sorting the sample in the laboratory takes 2-3 hours. Personnel performing each of these steps require, at most, a few hours of training. The time and expertise required for the identification stage vary with the taxonomic level. An experienced benthic taxonomist should be able to identify the specimens to the family level in about half an hour, to the lowest practical level in about 1-3 hours, depending on the number of taxa present. Less experienced personnel will need more time and should have their identifications confirmed by experts in the field.

6.0 IMPLICATIONS FOR THE GREAT LAKES BASIN

Agriculture is the predominant landuse in the lower Great Lakes basin, so any inputs to surface waters enter the Lakes themselves, either directly or indirectly. Our results suggest that about 80% of the small streams in southern Ontario are negatively impacted by agriculture.

Substantial investments over the past 25 years have reduced the inputs of nutrients and organic wastes from municipal sources, and these efforts appear to be resulting in significant improvements in water quality. Some progress has also been made with respect to industrial effluents. Such point sources are easily identified so remedial efforts can be concentrated.

The Monitoring Protocol presented here can be used to identify streams subjected to significant non-point source pollution, and to monitor improvements which result from changes in landuse practices. Agricultural streams receive inputs of eroded sediment, inorganic and organic nutrients, and a variety of agro-chemicals. Broad chemically based surveys to locate problem sites are prohibitively expensive; the invertebrate fauna can provide much more information, at a much lower cost.

7.0 TECHNOLOGY TRANSFER POTENTIAL

The Monitoring Protocol is useful to any organization interested in the effects of agriculture on the water quality of streams. User groups include conservation authorities, provincial Ministries (Environment, Natural Resources, Agriculture), federal Departments (Agriculture, Environment, Fisheries), as well as private organizations (e.g. Ducks Unlimited, OFAH), and private consultants.

The taxonomic descriptions and keys will increase the capability of environmental impact workers to distinguish and name Chironomidae living in small streams.

8.0 INFORMATION GAPS AND FUTURE RESEARCH

In order to increase the sensitivity of PMA for monitoring of agricultural impacts, expected communities could be refined to reflect narrower spatial and temporal scales.

This can be accomplished through extension of the existing database by sampling at different times at some of the sites already surveyed, as well as new ones. Additional data from forested reference sites are especially needed. Similarly, sensitivity will be enhanced with improvements in our ability to identify invertebrates to the species level. This will require continuing taxonomic studies.

The database should also be expanded to include larger streams. Concentration on headwater streams during this phase, in order to isolate the effects of individual crops, has left several questions to be answered:

1. Does the fauna of larger rivers integrate all landuses upstream?
2. Are certain agricultural inputs processed (i.e. removed from circulation) within the stream?

Answers to these questions would allow more meaningful interpretations of the effects of agriculture on the Great Lakes themselves, and could increase the efficiency of biomonitoring programs by reducing the number of sites needed for initial surveys.

9.0 SELECTED REFERENCES

Gray, J.S. 1989. The effects of environmental stresses on species rich assemblages. Biological Journal of the Linnean Society 37: 19-31.

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

Hilsenhoff, W. L. 1988. Rapid field assessment of organic pollution with a family level biotic index. Journal of the North American Benthological Society 7 : 65-68.

Novak, M.A. and R.W. Bode. 1992. Percent model affinity: a new measure of macroinvertebrate community composition. Journal of the North American Benthological Society 11: 80-85.

Rosenberg, D.M. and V.H. Resh (Eds.). 1993. Freshwater Biomonitoring and Benthic Macroinvertebrates. Chapman and Hall, N.Y. 488 pp.

Schindler, D.W. 1987. Detecting ecosystem responses to anthropogenic stress. Canadian Journal of Fisheries and Aquatic Sciences 44 (Supplement): 445-455.

10.0 PUBLICATIONS

Barton, D.R., D.R. Oliver and M.E. Dillon. A comparison of pupal exuviae and larval Chironomidae for biomonitoring of the impacts of agricultural practices on surface water quality. Proceedings of the 12th International Symposium on Chironomidae. CSIRO, Australia. In press.

Oliver, D.R., D.R. Barton and M.E. Dillon. 1992. Chironomid biomonitoring of agricultural land use. Chironomus 4: 10.

Oliver, D.R. and M.E. Dillon. 1994. Systematics of some species of *Micropsectra* (Diptera: Chironomidae) living in low-order streams in southern Ontario, Canada. The Canadian Entomologist 126: 199-217.

Walker, I.R., D.R. Oliver and M.E. Dillon. 1993. The larva and habitat of *Parakiefferiella nigra* Brundin (Diptera: Chironomidae). Netherlands Journal of Aquatic Ecology 26: 527-531.