IS BIOMONITORING REALLY EFFECTIVE?

P. Baisyet
CPG New Zealand Limited, Auckland, New Zealand

ABSTRACT

The purpose of freshwater biomonitoring is to provide an indication of a stream's, river's, or lake's health, typically in order to make a connection between a human activity and resulting effects. CPG has been undertaking biomonitoring annually within the Waitangi Stream since 2001 to assess possible impacts of landfill leachate and wastewater discharge originating from the Waiouru Military Camp and settlement. This paper discusses preliminary results from analysis of a seven-year longitudinal biomonitoring study to assess the impacts to Waitangi Stream from these discharge activities. Data was analysed using two approaches including one which examines variations and differences in each individual quantity measured and one which instead analyses the entire data set at once using principal components analysis (PCA). The two approaches provide different perspectives, however with consistent outcomes. As expected, the integrated approach (PCA) provided a broader perspective on the factors most influential in description of site disposition. Landfill control and impact sites were clearly differentiable in terms of habitat, but this did not manifest in large differences in macroinvertebrate health, whereas wastewater control and impact sites were not greatly different in terms of habitat, but the impact site did show evidence of compromised macroinvertebrate health. In conclusion, despite the inherent variability of biological data, biomonitoring is effective in understanding the inter-yearly variability between control versus impacted sites. The approach described herein, and related approaches, will help ensure that biomonitoring can be used to maximum utility.

KEYWORDS

Biomonitoring, landfill, wastewater, discharge, macroinvertebrates, biological indices, principal components analysis, multivariate analysis.

1 INTRODUCTION

Resource consent monitoring is required by a number of activities, and discharging effluent to natural waters often requires monitoring upstream and downstream of the discharge as a condition of consent. Biomonitoring is undertaken to detect the impacts of anthropogenic inputs such as discharge on macroinvertebrate and/or periphyton communities.

Transfield Services, Waiouru, holds two main resource consents that require biomonitoring, one to operate a landfill for Waiouru Military Camp and settlement and a second to discharge 3,200 m^3/day of treated wastewater from Waiouru Military Camp's wastewater treatment plant to Waitangi Stream via an unnamed drain. Discharges from both sites ultimately flow into the Waitangi Stream, and biomonitoring has been undertaken annually upstream and downstream at both sites to comply with resource consent conditions set out by Horizons Regional Council (HRC). Transfield Services has engaged CPG to undertake this freshwater biomonitoring within the Waitangi Stream since 2001.

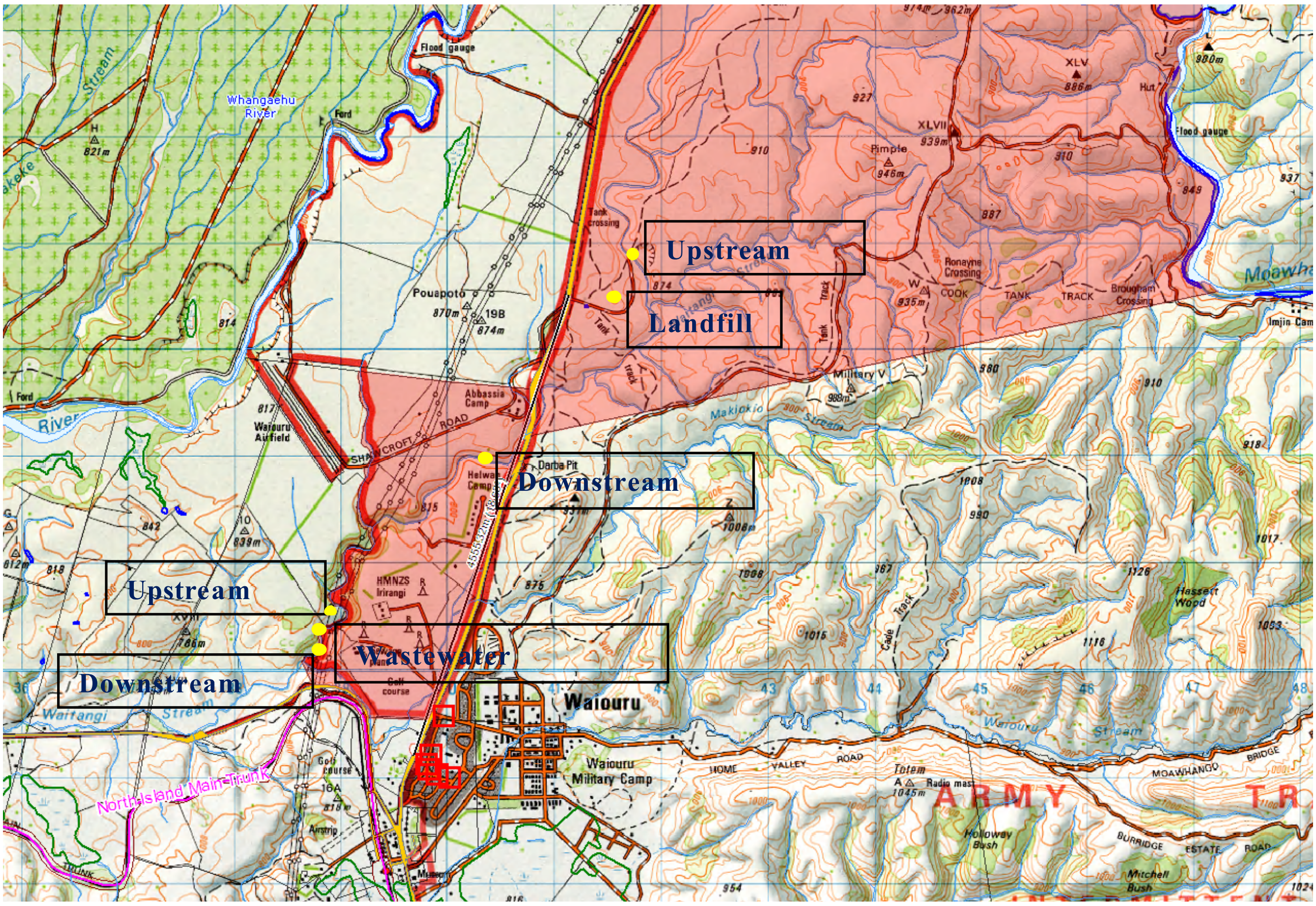
This paper presents results of the seven years of biomonitoring data undertaken from 2005-2011, the monitoring period for which the same types of data were collected each year, hence establishing a longitudinal record. Results are analysed by two different approaches, the first of which examines variations and differences in each individual quantity measured (univariate analysis), and the second of which analyses the entire data set at once (using a multivariate approach).

2 METHODS

2.1 STUDY AREA AND SAMPLING SITES

The study sites referred to in this paper are located on the Central Plateau of the North Island, New Zealand. Waiouru is home to a major New Zealand Army base, Waiouru Military Camp. A map of the study area is depicted in Figure 1. The Waiouru landfill site is located on military land, approximately 4.5 km north east of Waiouru Township. The landfill itself is located within the army training area along the Desert Road, North of Waiouru. The wastewater treatment plant and discharge is located on military land to the west and immediately to the northwest of the golf course and State Highway 1 at Waiouru. The wastewater from the treatment plant is discharged into an unnamed open drain before discharging into the Waitangi Stream. Waitangi Stream is a tributary of the Whangaehu River. Locations of both control (upstream) and impact (downstream) sampling sites are shown in Figure 1, wherein shaded land represents defence land.

Figure 1: Location of Study Area and Sampling Sites



2.2 SAMPLE COLLECTION AND HABITAT ASSESSMENT

Macroinvertebrate samples were collected using protocols for sampling macroinvertebrates in wadeable streams described by Stark et al. (2001). Samples were collected within riffle habitat using a handheld D-net with a 0.5 mm mesh and from an area totaling 1 m^2, as recommended by Stark et al (2001). Samples were preserved in 70% ethanol prior to subsequent laboratory analysis for the identification and enumeration of macroinvertebrates. Sample processing followed the Protocol P1 (coded-abundance) as outlined in Stark et al (2001). There are alternative methods for hard bottom streams that are increasingly used, however, the original Protocol P1 has been continued to retain internal consistency of the data set. By this protocol, each taxon is assigned to one of five coded abundance categories: R = Rare (1-4 individuals); C = Common (5-19 individuals); A = Abundant (20-99 individuals); VA = Very abundant (100-499 individuals); or VVA = Very, very abundant (500+ individuals) per sample. Macroinvertebrate data were analysed for: Macroinvertebrate Community Index

(MCI), Semi-Quantitative Macroinvertebrate Community Index (SQMCI) and Ephemeroptera (mayflies), Plecoptera (stoneflies) and Trichoptera (caddisflies) (EPT).

Habitat assessments, which evaluate the quality of the stream and riparian habitat, were completed at each site. Assessment, for both left and right stream banks, consisted of evaluating 20 metrics that are divided into five groups. Each individual metric is scored on a scale of 1 to 10, and the sum for each group is determined (i.e., group sum maxima are 40). Then the group sums are multiplied by a factor and the final value is reported. The factors are specified by the Regional Council and are commonly used by HRC and Hawkes Bay Regional Council. Groups and factors are as follow: Group A – adjacent land use (factor = 4); Group B – vegetation (width, structure, type, age, shading, buffer, periphyton and macrophyte, factor = 0.5); Group C – stability (bank and channel stability, factor = 2); Group D – disturbances caused by stock management (stock access and damage) or other external disturbances (inputs of sedimentation, contaminations, and artificial drainage, natural drainage, factor = 0.67); and Group E – bottom habitat diversity (hydraulic indicators, substrate availability and embeddedness of dominant coarse substrate, factor = 1.34).

2.3 DATA ANALYSIS

Data analysis included descriptive statistics/univariate analysis (mean, standard deviation, *t*-test) and multivariate analysis via Principal Components Analysis (PCA). The primary purpose of these analyses is to determine if impact sites are different from controls, and, if so, the extent to which they differ.

PCA is a multivariate statistical data analysis technique which reduces a set of raw data into a number of principal components (PCs) that can be thought of as revealing which aspects of a data set, when analysed all together, best explain variance within the original data in order to identify possible patterns among variables. For this work, PCA was performed by arranging data into a matrix using ten variables: macroinvertebrate metrics (total number of taxa richness, MCI, SQMCI, number of EPT, and percentage of EPT abundance) and habitat scores (adjacent land use, vegetation, stability, disturbances, and bottom habitat diversity). The columns in the matrix defined the variables and the rows defined upstream and downstream sites by year. The inputs are used to calculate a covariance matrix, whose eigenvectors (i.e. from linear algebraic operations) contain coefficients, or loadings, that relate the original data to the resulting PCs. Eigenvalues of the covariance matrix relate to the amount of the variability in the data set explained by each PC, and these were analysed by examination of variance and cumulative variance graphs. Using the so-called elbow rule (where the elbow bends), scree plots of cumulative variance suggested which PCs contained the most useful information. Once PCs of interest are identified, PCA factor loadings greater than or equal to 0.5 were considered significant.

3 RESULTS AND DISCUSSION

3.1 SUMMARY OF RESULTS FOR MACROINVERTEBRATES AND HABITAT ASSESSMENT

Means and standard deviations of macroinvertebrate metrics and habitat scores are shown in Table 1. The highest number of macroinvertebrates was recorded at the upstream landfill site and the lowest number downstream of the wastewater site. Percentage of taxa present that were mayflies, stoneflies and caddisflies (EPT) was greater at the landfill site compared to the wastewater site. The mayfly (*Deleatidium*) and caddisflies (*Aoteapsyche, Hydrobiosis, Pycnocentria*) were always present each year in rare, common or in abundance quantities.

Table 1: Summary statistics (mean ± standard deviation) for landfill and wastewater sites from 2005 – 2011[a]

Variables	Landfill		Wastewater	
	Upstream	Downstream	Upstream	Downstream
Total No. of Taxa Richness	15 ± 5	13 ± 3	17 ± 4	10 ± 4
MCI	119 ± 12	114 ± 10	113 ± 9	87 ± 13
SQMCI	6 ± 1	6 ± 1	5 ± 1	3 ± 1
No of EPT	7 ± 4	7 ± 3	8 ± 3	3 ± 2
% EPT Abundance	47 ± 15	48 ± 17	48 ± 7	24 ± 13

Adjacent land use	240 ± 0	91 ± 30	169 ± 46	163 ± 42
Vegetation	61 ± 22	34 ± 9	44 ± 17	41 ± 24
Stability	115 ± 47	57 ± 17	75 ± 15	80 ± 14
Disturbances	57 ± 7	39 ± 5	40 ± 6	41 ± 18
Bottom habitat diversity	42 ± 10	52 ± 9	54 ± 11	45 ± 16

[a] Total Habitat Scores: Landfill upstream = 3606; Landfill downstream = 1916; Wastewater upstream = 2669; Wastewater downstream = 2593.

Adjacent land use upstream of the landfill site has always scored high due to tussock grassland within the army training area, while extensive farmland surrounds the downstream site. As a consequence the upstream site has stable natural banks (stability score = 115), whereas stock access at the downstream site has resulted in increased bank erosion (disturbance score = 39). Undulating arid tussock land predominates upstream of the wastewater discharge. Within the immediate vicinity of the wastewater discharge, land use includes extensive pastoral farming on the true left bank and production forestry on the true right bank. Lower habitat scores downstream are predominately due to vegetation (vegetation scores = 44 upstream and 41 downstream).

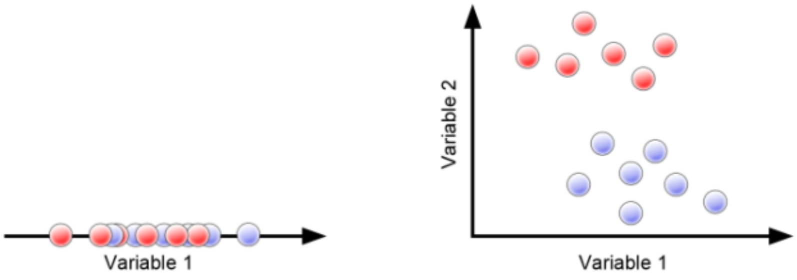
Differences between the landfill upstream and downstream sites are mostly due to habitat scores, which do not correspond with a decline in the macroinvertebrate community at the downstream site. Macroinvertebrate metrics have been consistently lower at the site downstream of the wastewater discharge, whereas habitat scores for this site are lower for only three of five habitat groups with averages that are reasonably close in all cases. The decline in water quality between upstream and downstream sites is likely a response to the discharge of treated wastewater from Waiouru Military Camp's wastewater treatment plant.

3.2 UNIVARIATE VS MULTIVARIATE APPROACHES TO ANALYSIS OF RESULTS

The Student's *t*-test was used for the univariate data analysis method. The null hypothesis for this investigation is that there is no difference between the control versus the impacted sites. Analysis of results indicates there is habitat scores that are different from the upstream and downstream landfill sites with high to very high levels of confidence, whereas macroinvertebrate metrics were not different to any reasonable level of confidence upstream vs downstream. The wastewater sites showed the opposite trend: habitat scores were not significantly upstream vs downstream whereas macroinvertebrate metrics were. Based on the results of univariate analysis then, we can see clear difference between upstream and downstream sites, and in this case, we can say that biomonitoring has been and is effective. Unfortunately, the case is not always so simple due to the inherently high nature of variability in biological data.

Using univariate *t*-testing is a 1-dimensional approach, i.e. each variable is analysed separately and independently from the others. Increased dimensionality through multivariate analysis gives the freedom of looking at data from a new viewpoint, and, more importantly, when differences between two sites are not statistically significant by univariate analysis, multivariate methods are often more sensitive. Figure 2 below uses a data representation in one and two dimensions that reveals how simultaneous analysis in two dimensions can reveal more than simple univariate analysis in one. However, there are several problems with multivariate analysis. One is that for more than three dimensions it is difficult to find a useful way to visualise results, a

Figure 2: A Schematic Diagram Showing the Effects of Increased Dimensionality on Data Patterns.



Two groups of data
1-D View

Same two groups of data
2-D View

phenomenon that is referred to as the curse of dimensionality. A second important problem is that multivariate methods typically less familiar to environmental scientists and engineers, hence often neglected or even distrusted through lack of familiarity. Thus this data set, the results for which can be very clearly interpreted by conventional univariate analysis, provides an ideal case for demonstration of the links between univariate and multivariate analysis. PCA was the method chosen for multivariate data analysis trials in this study because it is a technique with widely demonstrated power in pattern recognition, however the ability to extract a small number of PCs for final analysis reduces the curse of dimensionality.

For PCA analysis using the PCs for ten variables, the first two PCs accounted for 84.5% of the total variation in the data set, with 73.3% of the variability in the data set being expressed in first principal component. The first component arises due to adjacent land use while the second component arises due to stability (based on loadings ≥ 0.5). An example of how results can be simply visualised is given in Figure 3. In Figure 3A, the red dashed line shows a linear decision boundary indicating a clear separation between upstream and downstream landfill sites. The predominant grouping inside the red circle suggests that for most years there is a uniform habitat condition. Clusters for the two sites are distributed similarly along the PC2 axis, and hence are primarily distinguishable as different based on PC1. Figure 3B, in contrast, shows that for the wastewater site there is no clear decision boundary between the control and impact sites, indicating that the most variable habitat contributors are the same, to within variability, for the control and impact sites.

Figure 3: Bivariate Plots of PC1 and PC2 for the Landfill (3A, Squares) and Wastewater (3B, Triangles) Sites. Upstream and Downstream Sites are Denoted by Open and Filled Symbols, Respectively. Years are Color Coded: Red = 2005, Green = 2006, Blue = 2007; Black = 2008; Yellow = 2009, Magenta = 2010 and Cyan = 2011.

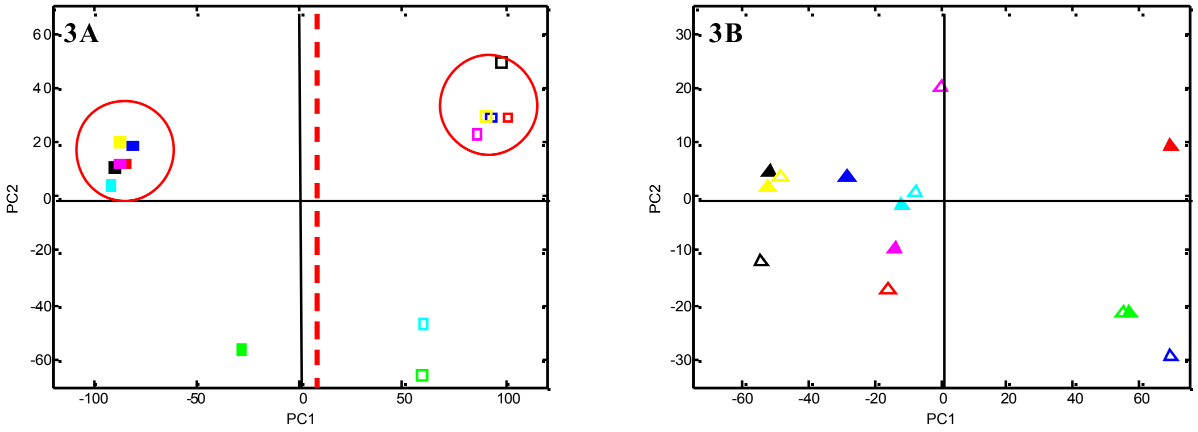


Table 1 shows that adjacent land use is the habitat score that is by far most different for the upstream and downstream landfill sites, so this input was deleted from the data and the PCA analysis was performed using nine variables.

For the new nine variable data set, the first two PCs accounted for 76% of the total variation in the data set, with 55% of the variation being expressed in PC1. The first component arises due to stability, while the second component arises due to MCI and %EPT abundance, i.e., for PC2 all factor loadings with absolute magnitude greater than 0.5 relate to variability in macroinvertebrate metrics. Comparing PC1 (habitat influenced) to PC2 for the landfill site, there is a difference in the central tendency of the data based on PC1 however, based on PC2, the two groups do not separate much. Since PC1 is constituted of habitat factors, this suggests that variability in macroinvertebrate results are not much of a differentiating factor overall for the landfill site, whereas the landfill control and impact sites are distinctly different in terms of habitat. For the wastewater sites, there is a clear decision boundary near PC2 = 0, and the control and impact sites are separable based on PC2. For these sites, PC2 is rather stable from year to year, however, there is a lot of inter-year variability injected from habitat.

4 CONCLUSIONS

The results of macroinvertebrate and habitat assessments presented here suggest that the downstream landfill site is not affected by leachate, however, the downstream wastewater site is affected by the wastewater discharge. For the landfill sites, there is a clear difference between upstream and downstream sites based on habitat contributors, with only occasional inter-year variability. For the wastewater sites, though health at the downstream site is substantially less than upstream, there was no discernible difference in habitat contributors. The results of multivariate analysis are consistent with those from univariate analysis, while simultaneously offering greater insight into covariability of different metrics and the ability to confidently assess an entire complex data set in an integrated fashion. Based on this preliminary investigation, we can conclude that biomonitoring is effective. In cases where the high variability in biological data reduces confidence in assessing differences in control vs impact results, multivariate approaches can be employed to ensure that the efficacy of biomonitoring is ensured.

ACKNOWLEDGEMENTS

I would like to acknowledge Colin Nightingale at Transfield Services, Waiouru for his assistance annually out in the field. In addition, I am greatly appreciative of the technical support, guidance and critical advice provided by Dr Mona Wells and other staff at CPG NZ Ltd.

REFERENCES

Stark, J. D. 1985. A macroinvertebrate community index of water quality for stony streams. New Zealand Ministry of Works and Development.

Stark, J. D. 1993. Performance of the Macroinvertebrate Community Index: Effects of sampling method, sample replication, water depth, current velocity, and substratum on index values. *New Zealand Journal of Marine and Freshwater Research* 27: 463-478.

Stark, J. D. 1998. SQMCI: A biotic index for freshwater macroinvertebrate coded abundance data. *New Zealand Journal of Marine and Freshwater Research* 32: 55-66.

Stark, J. D; Boothroyd, I. K. G; Harding, J.S; Maxted, J.R; Scarsbrook, M.R. 2001. Protocols for sampling macroinvertebrates in wadeable streams. New Zealand Macroinvertebrate Working Group Report No. 1. Prepared for Ministry for the Environment, Sustainable Management Fund Contract No. 5103.