8-1984

An Evaluation and Comparison of Various Diversity Indices as Illustrated by Portage Creek

Karl Josef Siebert

AN EVALUATION AND COMPARISON OF VARIOUS DIVERSITY INDICES AS ILLUSTRATED BY PORTAGE CREEK

by

Karl Josef Siebert

A Thesis
Submitted to the
Faculty of The Graduate College
in partial fulfillment of the
requirements for the
Degree of Master of Arts
Department of Biology

Western Michigan University
Kalamazoo, Michigan
August 1984

AN EVALUATION AND COMPARISON OF VARIOUS DIVERSITY INDICES AS ILLUSTRATED BY PORTAGE CREEK

Karl Josef Siebert, M.A.

Western Michigan University, 1984

The diversity index is an established method of measuring water quality. In this study various diversity indices are evaluated for their use in a rapid, accurate, and cost effective biological monitoring program. The study investigated five diversity indices: Approximate, Brillouin's, Shannon's, Simpson's and the Sequential Comparison Index. The diversity indices were tested for accuracy and sensitivity on actual macroinvertebrate samples and simulated data.

The Simpson's Index was found to provide the most accuracy and sensitivity for the indices tested. The importance of considering the evenness element of diversity indices was also supported by the study.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Clarence Goodnight, for his time, support, and assistance throughout my Master's Thesis and graduate program. I would also like to thank the other members of my thesis committee, Dr. Joseph Engemann and Dr. Leo Vander Beek, for their time and assistance in the completion of the thesis. For the background information on the Portage Creek, my thanks go to the Plainwell field office staff of the Department of Natural Resources. Finally, for putting up with long nights on the microcomputer, I would like to thank my wife, Laurie, for her patience and support.

Karl Josef Siebert

1323969

SIEBERT, KARL JOSEF

AN EVALUATION AND COMPARISON OF VARIOUS DIVERSITY INDICES AS ILLUSTRATED BY PORTAGE CREEK

WESTERN MICHIGAN UNIVERSITY M.A. 1984

University
Microfilms
International 300 N. Zeeb Road, Ann Arbor, MI 48106

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Water pollution directly and indirectly affects organisms through toxicity, oxygen depletion, eutrophication, habitat alteration, temperature change, and many other ways. Changes of organism function and distribution within the aquatic community are reflected in community and ecosystem changes. If a community changes, so may its function (Patrick, 1976). Functions such as air and water purification, food source, breeding habitat, and natural resource, are all essential to the survival of many organisms--especially humans. Too often entire ecosystems are permanently destroyed as a result of wanton disregard by society for the aquatic environment. Programs developed by federal, state and local agencies attempt to deal with both the deliberate and accidental polluter. Once the damage has been inflicted, litigation is slow and the environment is damaged. The most effective method of pollution control is prevention. The key to most prevention programs (besides education) is the use of biological monitoring.

Aquatic organisms act as natural biological monitoring systems. The assimilation of pollutants by organisms such as fish make the bioassay a useful tool in biological monitoring programs. The study of behavioral responses of organisms to effluents gives insight to the stress on biological systems resulting from pollution (Ott, 1978). Observation of community structure gives much information on pollutants otherwise not observed through measurement of chemical and

physical parameters (Cairns et al., 1973).

Aquatic community structure can be studied on two levels: qualitative and quantitative. By observing the types of organisms present, whether tolerant, intolerant, or facultative to pollutants, stressed environments can be identified (Persoone et al., 1978). The study of community structure is often time consuming, expensive, and requires expertise for taxonomic identification and interpretation. The diversity index, based on information theory, is an efficient method for summarizing the information contained within a community. Various diversity indices (indexes) have the potential to provide much information for minimal time, financial, and technical input.

The purpose of this thesis is to explore the different diversity indices and evaluate their use in a rapid, accurate, and cost effective biological monitoring program. The scope of the thesis includes the testing of five diversity indices: Approximate, Brillouin's, Sequential Comparison Index, Shannon's, and Simpson's indices. The indices are initially tested on four stream sample sites experiencing various degrees of stress. Physical, chemical, and biological data from each site serve as the basis for interpreting the accuracy of each index with respect to community structure changes (as a result of pollution). The second test of index sensitivity includes the evaluation of simulated data. Diversities are measured for two very similar data sets and conclusions are made regarding the ability of an index to distinguish different community structures.

The goals of the thesis include: selection of the "best" diversity index (if possible), collection of information and data to

supplement current knowledge of the diversity index, and to suggest changes or new directions in the study of aquatic community structure through diversity indices.

Literature Survey

Many attempts have been made to assess the impact of pollutants on aquatic communities and habitats. Chemical tests alone, as demonstrated by Patrick (1949, 1976) and others, may not indicate environmental stress (pollution) when the source is intermittent, accumulative, or synergistic/antagonistic effects have taken place. The abundance, collectability, and long life cycles of the benthic macroinvertebrates make them ideal candidates for indicating community changes due to pollution.

Many systems for evaluating benthic communities with respect to changes resulting from environmental stress have been developed. Some of the methods include: the "Saprobiensystem" (Kolkwitz & Marsson, 1908, 1909), adaptations of the same (discussion by Persoone et al., 1978; Goodnight, 1973), the "Saprobic Index" (Pantle & Buck, 1955), Beck's (1954) "Biotic Index", the "Trent Biotic Index" (Persoone et al., 1979), and the "Score System" by Chandler (1970). Most of the systems noted viewed the benthic community as a collection of indicator organisms rather than an organized structure of interrelated living organisms. The diversity indices developed out of a need to assess the overall community structure rather than the individual members.

The Diversity Concept

The majority of literature concerning diversity indices is based on information theory. Fisher, Corbet, and Williams (1943) introduced the diversity concept derived from a logseries distribution of a sample or community. MacArthur (1955) and Margalef (1958) proposed diversity indices using information theory, also developing equations for measuring evenness (distribution of individuals among species) of collections and populations. Much of the literature to follow originated with the Shannon and Weaver (1949) concept of information content. Wilhm and Doris (1968) adapted this concept to macroinvertebrate community study, known today as the Shannon-Weaver and Approximate Diversity Indices. Further discussion of these indices can be found in Pielou (1966, 1977).

Other important diversity indices evolved from information theory. Simpson's measure of concentration (1949) was adapted to biological diversity measurement and discussed by Pielou (1977). Another popular index was developed by Brillouin (1962) and supported in a study by Kaesler (1977).

The Sequential Comparison Index (SCI), although not as mathematically oriented as previous indices, presents a quicker and simpler alternative for calculating the diversity of a macroinvertebrate community. Two methods were developed by Cairns et al. (1968, 1971) using this index as a tool for technicians with little biological background. Other articles describing the use of the SCI include Persoone et al. (1979) and Goodnight (1973).

Currently there is still much controversy surrounding the use of

diversity indices. The indices have been criticized for their insensitivity to pollutants other than organic oxygen-depleting wastes (Wolf, 1980; Persoone et al., 1979). Hurlbert (1971) rejects the overuse of information theory and proposes indices of his own. Zaret (1982) questions the validity of species diversity being an indicator of the ecological health of a community. In a study of Wisconsin stream communities, Hilsenhoff (1977) found that a biotic index correlated more often (then the diversity index) with the chemical parameters of the streams. Despite criticism the diversity index is still being used extensively in evaluating aquatic communities. This includes the testing of surface waters and effluents by the Environmental Protection Agency (Weber, 1973). Pielou (1977) summarizes the merits and role of the diversity index:

> It should not be (but it is) necessary to emphasize that the object of calculating indices of diversity is to solve, not create, problems. The indices are merely numbers, useful in some circumstances, but not in all . . . . Indices should be calculated for the light (not the shadow) they cast on genuine ecological problems.

In the discussion of diversity theory, biological or species diversity will be defined as "a function of the number present (species richness or abundance) and the evenness with which the individuals are distributed among these species." (Hurlbert, 1971).

Most indices are derived from information theory; a strictly mathematical concept (Pielou, 1975, 1977). Based on this theory, diversity indices are actually reflecting bits of information in a given sample or community (Cairns et al., 1973). A community containing more information is more diverse, while the lack of information depicts a

community with low diversity. Pielou (1977) prefers the concept of "uncertainty" rather than bits of information. As diversity increases, so does the degree of uncertainty to which species a sampled individual might belong.

Also a part of the diversity theory is the concept of evenness or equitability. Determining the diversity index without calculating evenness may lead to erroneous interpretations (Hurlbert, 1971). Since a community with a few evenly represented taxa may have the same diversity as one with many unevenly represented taxa, a measure is needed to illuminate these cases. There are two measures of evenness and one of equitability. In each case, the observed diversity is compared to the maximum diversity that could be measured in that specific testing or sampling situation. There are limitations to the use of evenness measures since validity rests on the choice of censused or uncensused communities.

Once the evenness or the equitability of the sample is determined, a basis for comparing and confirming diversity indices has been established.

CHAPTER II

METHODS

Selection of Sampling Stations

The Portage Creek was chosen for the sampling stations because of previous pollution history (DNR Report 1971 & 1972). The first five miles of the creek appear clear and clean, while the lower three miles show signs of pollution such as turbidity, silt accumulations, and debris.

The Portage Creek originates in the Gourdneck State Game Area (Figure 1), passing through Hampton Lake, flowing north toward Kalamazoo, joining the west branch of Portage Creek. The creek, ranging from fifteen to about sixty feet across and one to three foot in depth, flows through residential and park areas until it reaches Cork Street. North of Cork Street, the largest single source of effluent (paper waste clarifier) enters the creek. The Portage Creek then passes through residential, commercial, and industrial properties on its route to the Kalamazoo River.

The Portage Creek offers a unique location for the study of diversity indices. Since the creek flows primarily from a single source, water quality changes are limited to pollutant inputs along the course of the creek. In this way, perturbations other than pollutants can be eliminated to allow for easier comparisons of diversity indices. The "clean" stations are selected from the lower three to five miles,

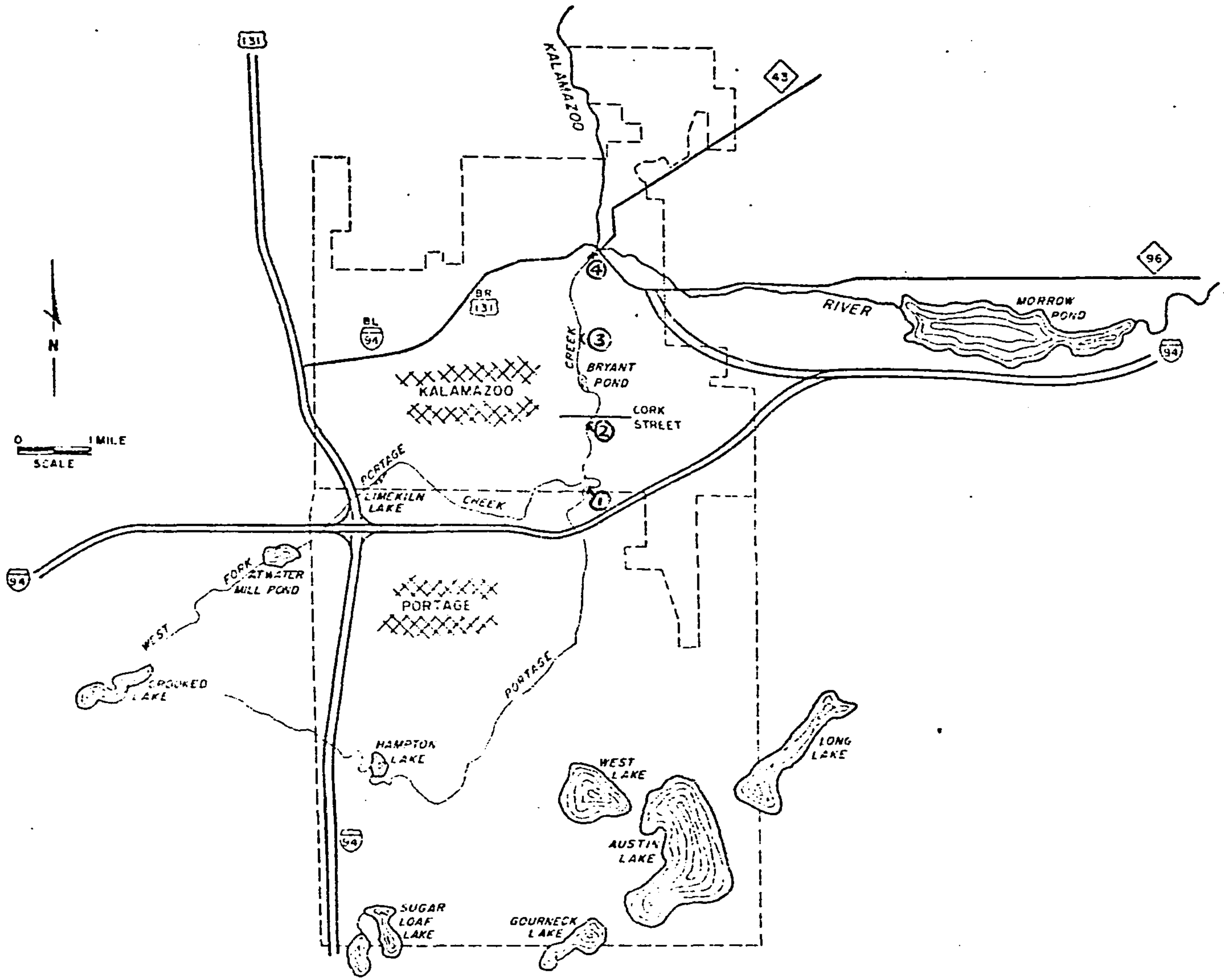

Figure 1. Portage Creek - reprinted from DNR study May 1972.

while the "dirty" stations are located in the lower three miles of the creek. The clean stations are then used as controls, theoretically reflecting higher diversity values.

Site #1 was chosen as the cleanest, or least stressed, of the sites. Before the creek reaches this station, it flows through open fields and residential areas. The station is located on the southwest corner of the Kilgore and Lovers Lane intersection. This section of the stream is characterized by clear, fast flowing water over a coarse aggregate bottom with little vegetation.

Site #2 is located just south of Cork Street and is also characterized by clear fast flowing water. The large rock and coarse aggregate bottom supports primarily algal growth with some higher plant growth in the small riffle areas. This station was also chosen as a clean sampling site, affected by perturbations similar to site #1 with the addition of the west branch of the Portage Creek.

Site #3 is located just south of Stockbridge Avenue, about one and one-quarter miles north of site #2. This station was selected as a dirth sampling site. The main sources of stress are the three million gallons per day effluent from the paper waste clarifier and runoff from storm sewers in the residential areas. The creek's flow is reduced slightly and turbidity increases. At this station the bottom is characterized by silt over sand with some debris. Vegetation is minimal, characterized by algal growth on debris.

Site #4 is located just over a mile north of site #3, by Michigan Avenue. About 150 feet below the sample station, the creek enters the Kalamazoo River. Site #4 was selected as the dirtiest site (DNR, 1972)

since it flowed through primarily industrial areas and received the already stressed waters from site #3. Site #4 is characterized by slow moving water with varied turbidity (usually greater than #3), occasionally gray in appearance. On two occasions an oily cast was observed on the creek's surface. The bottom is silt and sand with few large rocks and debris. There is no vegetation other than some algal growth on rocks.

Sampling Methods

Square meter areas were selected which best represented the creek at the sampling station. If there were both riffle and high flow areas, square meter samples were taken of each. A hand net with a 30 cm. triangular opening and two millimeter mesh was dragged through the top five centimeters of silt and sand, if present. Where bottom sediment was lacking, rocks and debris were sampled using a tweezer and eye dropper. Each square meter was sampled until five minutes lapsed without collecting an additional organism.

Samples were collected during a sixty to ninety minute period, and an attempt was made to collect at least 100 individuals when possible. Total collection sizes ranged from 400 to 675 individuals reflecting ease of collection, macroinvertebrate densities, and areas needed to obtain at least 400 individuals (refer to Table 1, Chapter III). Since one index required a minimum of 250 individuals and it took one working day to collect 400 at the site with least density, a minimum sample size of 400 macroinvertebrates was selected.

The macroinvertebrates collected were a minimum of three

millimeters in length for ease of identification. The samples were collected in one liter plastic bottles and preserved in five percent formalin and stored at approximately five degrees celcius. One of the samplings from site #4 was partially lost due to inadequate levels of formalin.

Macroinvertebrates were identified to at least family levels and genus when possible. The references used included taxonomic guides by Lehnkuhi (1979) and Meritt (1978). For the diversity index calculations, family levels were used since most families were represented by a single genus and generic levels were not always identified. Kaesler et al. (1978) suggests the use of family and genus identification for calculating diversity indices.

Index Calculations

The five methods of diversity calculation are based on mathematical properties of populations (or samples) of organisms. Each index calculation will be discussed together with conditions. Since different index values and methods will be compared, an evenness calculation will be determined for each diversity value to standardize the scale of values (0-1 units). A method for calculating a biotic index for comparison will also be discussed.

Approximate Index

The Approximate Index equation estimates the diversity of a sampled uncensused infinite community (Kaesler, 1978):

$$H'' = \sum_{i=1}^{s} N_i/N \times \ln(N_i/N)$$

where: H'' = Approximate Index
s = taxa or species number
N_i = number of individuals in the i^{th} taxon
N = total number of individuals in sample

Brillouin's Index

The Brillouin's Index measures the diversity of a fully censused community, such as a large sample:

$$H = 1/N \times \log_{10}[N!/(N_1! \times N_2! \times \ldots N_s!)]$$

where: H = Brillouin's Index
$N!$ = the factorial of the community size
$N_{1 \ldots s}!$ = factorials of the number of individuals within a taxon

Brillouin's equation accounts for the total community or sample and, therefore, is not an estimator of diversity (Pielou, 1977), but an actual diversity measurement. Brillouin's Index is dependent on sample size requiring large sample size or replicated smaller samples.

Sequential Comparison Index

The Sequential Comparison Index, or SCI was developed to simplify calculations of diversity (Cairns et al., 1968). The SCI measures diversity of a fully censused community, as with Brillouin's Index.

$$SCI = N_{runs} / N_{total}$$

where: SCI = Sequential Comparison Index
N_{runs} = number of runs of similar organisms
N_{total} = total number of individuals in sample

The original method for determining the SCI selects an individual randomly from a container and the individual is compared to the previously selected individual. If the individuals are alike, then the letter X is recorded (assuming one starts with X); if they differ, the letter 0 is recorded. The symbol of the previous individual is always recorded when the selected individual is like the previous one; if not, record the other symbol. After all the individuals are selected, the number of runs are counted as determined by symbol changes (N_{runs}). At least 200-250 specimens are necessary for this method.

Later modifications to the SCI (Cairns et al., 1971) used random number tables to reduce randomizing error. For the thesis a computer program was developed to evaluate SCI values using a random number generator. Since each randomly generated sequence of numbers is unique, so are the SCI values generated by the program. As the sample size increases the differences among values decrease. The values presented under Results are averages of five computer runs. The five values were within three percent of the average.

Shannon Index

The Shannon diversity index estimates diversity of a sampled uncensused infinite community:

$$H' = \sum_{i=1}^{s} p_i \times \log_2 (p_i)$$

where: p_i = the probability of selecting an individual belonging to the i^{th} taxon

H' = Shannon's diversity index

The Shannon equation can only estimate the diversity of a

community. To determine p_i, total number of species and population must be known. Since only small samples (less than 1000) are utilized in the study, estimates regarding total species composition and population of the community are very difficult (Slocumb et al., 1977). The taxon probabilities must be substituted by the individuals within a taxon divided by the total sample size. In this form, the Shannon Index is very similar to the Approximate Index, except for the difference in logarithmic base.

Simpson's Index

The Simpson's Index was originally designed to measure concentration or dominance of a taxon or species (Simpson, 1949). To use the index for diversity calculations two equations must be applied. The first equation determines concentration while the second utilizes a negative logarithm to transform increasing concentration to decreasing diversity.

$$C = \sum_{i=1}^{s} N_i(N_i - 1)/N(N-1)$$

$$D = -\ln C \qquad \text{also } -D = -\ln\left(\sum_{i=1}^{s} N_i(N_i-1)/N(N-1)\right)$$

Where: C = Simpson's concentration

D = Simpson's diversity index

This index can determine diversity for both censused and uncensused communities.

Biotic Index

There are many biotic indices used for evaluating water quality through macroinvertebrate sampling of lakes and streams (Persoone, 1978;

Goodnight, 1973). The "Score System" by Chandler (1970) worked best with the taxa identified from the sample stations. To obtain a "Score", a numerical value is given to each taxon according to the taxa's relative abundance. The score for the taxa in each sample are added to obtain the value. The system was modified slightly for the thesis. The original method used five minute samplings. Since the organism collection rates were relatively low (see Table 1 in Results chapter), the entire sample for each station was evaluated.

Evenness Calculations

As previously discussed, evenness or equitability measurements are essential in evaluating diversity. Hurlbert (1971) presents two methods for determining the evenness of samples or communities of macroinvertebrates:

$$V' = D/D_{max}$$

$$V = (D-D_{min})/(D_{max}-D_{min})$$

where: D = diversity as determined by various indices

D_{max} = sample diversity if all taxa were equally abundant (maximum diversity)

D_{min} = sample diversity if one taxon was represented by N-(S-1) individuals and the other taxa represented by one individual each.

V = evenness for sample of uncensused community

V' = evenness of fully censused community or sample

Both of these methods relate observed diversity to maximum diversity possible in the specific sample being measured. Care must be exercised regarding the use of V or V' since different indexes treat samples as either censused communities or samples of uncensused communities.

Evenness tends to be more sensitive than the diversity index

alone. The evenness value will fall between zero and one regardless of the specific diversity index's range being evaluated.

Chemical and Physical Measurements

In order to collect chemical and physical information for comparison of water quality and the various indices, one liter samples were taken from each site and analyzed. The analysis included temperature, dissolved oxygen, pH, turbidity, nitrites, nitrates, phosphates (ortho), and chlorides.

The Hach DR-1E colorimeter was used for the measurement of turbidity and the nitrogen and phosphorus compounds. Dissolved oxygen values were obtained using the modified Winkler method. Titration techniques were used to evaluate chlorides.

Index Sensitivity Test

To test index sensitivity or an index's ability to differentiate similar communities, two similar collections were simulated and measured for diversity and evenness.

The researcher chose collections of fifteen taxa and 400 individuals based on the actual Portage Creek collections. The collections each include five "rare" taxa (one individual per taxa) and the balance distributed almost evenly among the remaining ten taxa (see Appendix). According to the diversity definition, diversity is dependent on both species or taxa richness and the distribution of individuals among the taxa. Since taxon number and sample size are identical for the simulated collections, differences in diversity depend on

differences in distribution of the individuals. The collection with a more even distribution than the other should have the greater diversity.

Test set B was simulated to be more evenly distributed than set A. The measure of standard deviation is used to confirm the differences between the collections. The researcher wishes to note that the use of standard deviation to indicate diversity differences is unproven. The standard deviation is used statistically to distinguish distribution around the mean numbers of individuals per taxon.

CHAPTER III

RESULTS

A summary of collection data can be found in Table 1. The total number of taxa identified along with total organisms identified are listed in the table. Some organisms were lost prior to sampling as a result of insufficient preservative. Site #4 lost the most organisms; twenty-eight. There were two organism densities calculated. The first density relates organism number to area (meter^2), while the second measures organisms collected per unit time (five minutes). There were two flow measurements made; the first, volume, reflects amount of water flowing through area, and the second, rate, indicates velocity of water. The low flow volume for site #4 was not expected

Table 1

Collection Data

Station #	1	2	3	4
Taxa	14	16	16	8
Sample Size	674	588	490	372
Individuals/M^2	266	100	100	50
Individuals/5 min.	24	20	8	6
Flow Volume (ft^3/s)	44	81	90	66
Flow Rate (ft/s)	2.1	3.2	2.5	1.1

and probably a result of difficulties in measuring depth. Sites #3 and #4 should demonstrate similar flow volumes.

The organism density values (Individuals/M^2) are approximate diversities. Collection sizes were usually multiples of one hundred and sample areas were usually rounded to integer units.

Table 2 indicates the taxonomic level and the corresponding number of individuals identified in each group. The same data can be found in Table 3 represented by percent of station sample instead of numbers of individuals. This facilitates observation of changes in dominant groups among test stations. A shift from the less tolerant organisms (Ephemeroptera, Trichoptera) to the more tolerant organisms (Isopoda, Oligochaeta) can be noticed as stations become more stressed.

The diversity index values and corresponding evenness values are listed in Table 4. The table also includes a biotic index (score system) for the four stations. Generally, diversity values range from zero to three or four, except for the SCI which is zero to one. Wilhm (1968) indicated values of three-four for unpolluted streams and values less than one for polluted streams. The evenness value will fall between zero and one. Values above 0.6 represent fairly clean water quality while values less than 0.3 depict water stressed by oxygen depleting pollution (EPA, 1973).

The two bar graphs (Figures 2 and 3) illustrate the changes in diversity as the stations become more stressed. Note the low values of diversity and evenness for site #3. This station is approximately one-half mile downstream from the single registered effluent input to the stream. The effluent is of an oxygen depleting character

Table 2

Biological Data

	Site #1	Site #2	Site #3	Site #4
Amphopoda				
Gammarus	1	60	0	3
Isopoda				
Asellus	88	88	76	171
Coleoptera				
Elmidae	18	3	0	0
Diptera				
Chironomidae	4	12	324	34
Simulidae	58	17	0	0
Chaoboridae	0	0	0	1
Gastropoda				
Physa	0	0	1	0
Helisoma	0	0	1	0
Limpets	0	0	13	0
Hemiptera				
Corixidae	0	1	0	0
Naucoridae	0	1	0	0
Odonata				
Coenagrionidae	2	1	2	2
Agrionidae	0	0	1	0
Oligochaeta				
Tubifex	0	0	0	40
Lumbriculus	0	0	0	4
Trichoptera				
Hydropsychidae	11	92	4	0
Odontoceridae	0	1	0	0
Phyrganidae	1	0	1	0
Ephemeroptera				
Baetidae	289	224	27	0
Heptageniidae	145	13	4	117
Oligoneuridae	42	52	24	0
Caenidae	5	1	4	0
Arhynchobdellida				
Glossiphoniidae	0	0	1	0

Table 2, Continued

	Site #1	Site #2	Site #3	Site #4
Piscicolidae	0	4	4	0
Hydracarina	1	0	1	0
Tubellaria Planariidae	10	18	0	0
Total Number	675	588	490	372
Total Taxa	14	16	16	8

(paperwaste clarifier).

The chemical and physical data presented in Table 5 reflect single one liter samples taken from each of the four sites within a ninety minute period. Tests were performed within twenty-four hours, except for temperature, dissolved oxygen, and pH; these measurements were made within fifteen minutes of collection time.

The physical and chemical data found in Table 5 do not reveal much about the low diversity levels observed in Figures 2 and 3. According to the chemical data, the stream appears fairly healthy (except for possible heavy metals and/or organic toxins which were not measured in this study).

Sample data were simulated (Appendix) using 400 individuals distributed among 15 taxa. The Samples A and B differ slightly as demonstrated by the standard deviations. Sample A, with the highest standard deviation, should demonstrate lower diversity according to

Table 3

Biological Data - Percent Composition

	Site #1	Site #2	Site #3	Site #4
Amphopoda				
Gammarus	0	10	0	1
Isopoda				
Asellus	13	15	15	46
Coleoptera				
Elmidae	3	<1	0	0
Diptera				
Chironomidae	<1	2	65	9
Simulidae	9	3	0	0
Chaoboridae	0	0	0	<1
Gastropoda				
Physa	0	0	<1	0
Helisoma	0	0	<1	0
Limpets	0	0	3	0
Hemiptera				
Corixidae	0	<1	0	0
Naucoridae	0	<1	0	0
Odonata				
Coenagrionidae	<1	<1	<1	<1
Agrionidae	0	0	<1	0
Oligochaeta				
Tubifex	0	0	0	10
Lumbriculus	0	0	0	1
Trichoptera				
Hydropsychidae	2	15	1	0
Odontoceridae	0	<1	0	0
Phyrganidae	<1	0	<1	0
Ephemeroptera				
Baetidae	42	37	5	0
Heptageniidae	21	2	1	31
Oligoneuridae	6	9	5	0
Caenidae	1	<1	1	0

Table 3, Continued

	Site #1	Site #2	Site #3	Site #4
Arhynchobdellida				
Glossiphoniidae	0	0	<1	0
Piscicolidae	0	<1	1	0
Hydracarina	1	0	<1	0
Tubellaria				
Planariidae	2	3	0	0
	100%	100%	100%	100%

the diversity/evenness definition. This assumption was confirmed by four of the five indexes; Brillouin's index values did not agree with the other index values. Brillouin's index indicated that sample B had the highest diversity and evenness.

Of the four indexes which gave theoretically correct values for A and B, Simpson's shows the greatest difference between diversity and evenness measurement. Shannon's index also demonstrates a difference between diversity values for A and B.

Table 4

Calculation of Indices

Index	Site #1	Site #2	Site #3	Site #4
Approximate	1.68	1.88	1.24	1.31
V	0.779	0.653	0.395	0.608
Brillouin's	1.64	1.83	1.19	1.27
V	0.617	0.651	0.398	0.608
SCI	0.717	0.758	0.500	0.626
V	0.790	0.832	0.524	0.737
Shannon	2.43	2.71	1.78	1.89
V'	0.794	0.677	0.442	0.629
Simpson's	1.36	1.55	0.76	1.11
V	0.662	0.547	0.258	0.532
Biotic Index	470	450	402	178

Table 5

Chemical and Physical Data

Parameter	Site #1	Site #2	Site #3	Site #4
Temperature (°C)	19	15	19	18
O_2, dissolved	12	9.6	9.6	9.6
pH	8.3	7.9	8.2	8.0
Turbidity (FTU)	15	15	20	25
Nitrite (mg/1)	0.015	0.008	0.018	0.018
Nitrate (mg/1)	1.7	1.5	1.8	1.8
Phosphate, ortho (mg/1)	0.6	0.5	0.4	0.4
Chlorides (mg/1)	44	52	42	50

Key for bar graph identification:

Figure 2 and Figure 3

Approximate

Sequential Comparison Index

Brillouin's

Shannon's

Simpson's

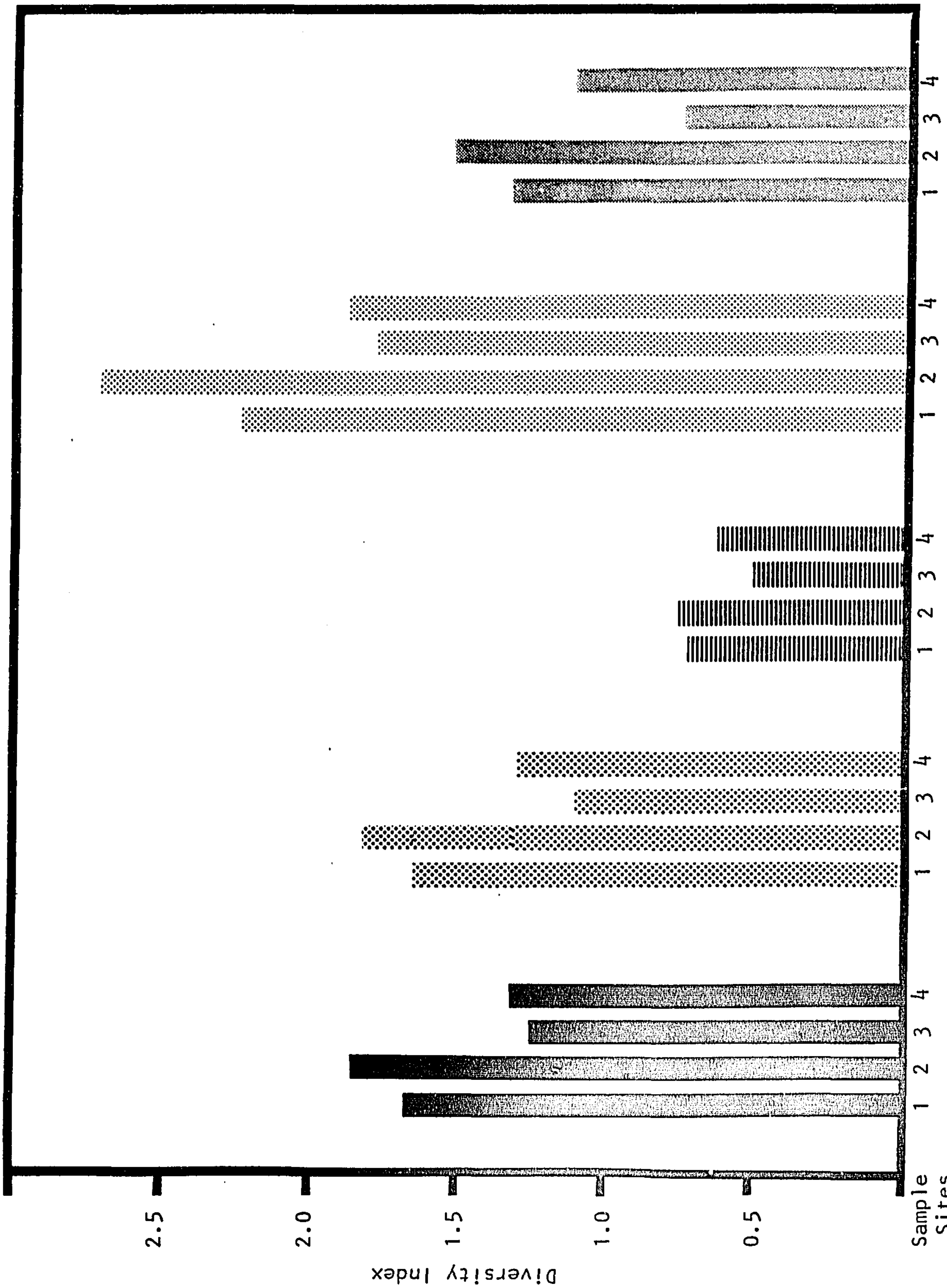

Figure 2. Comparison of Diversities

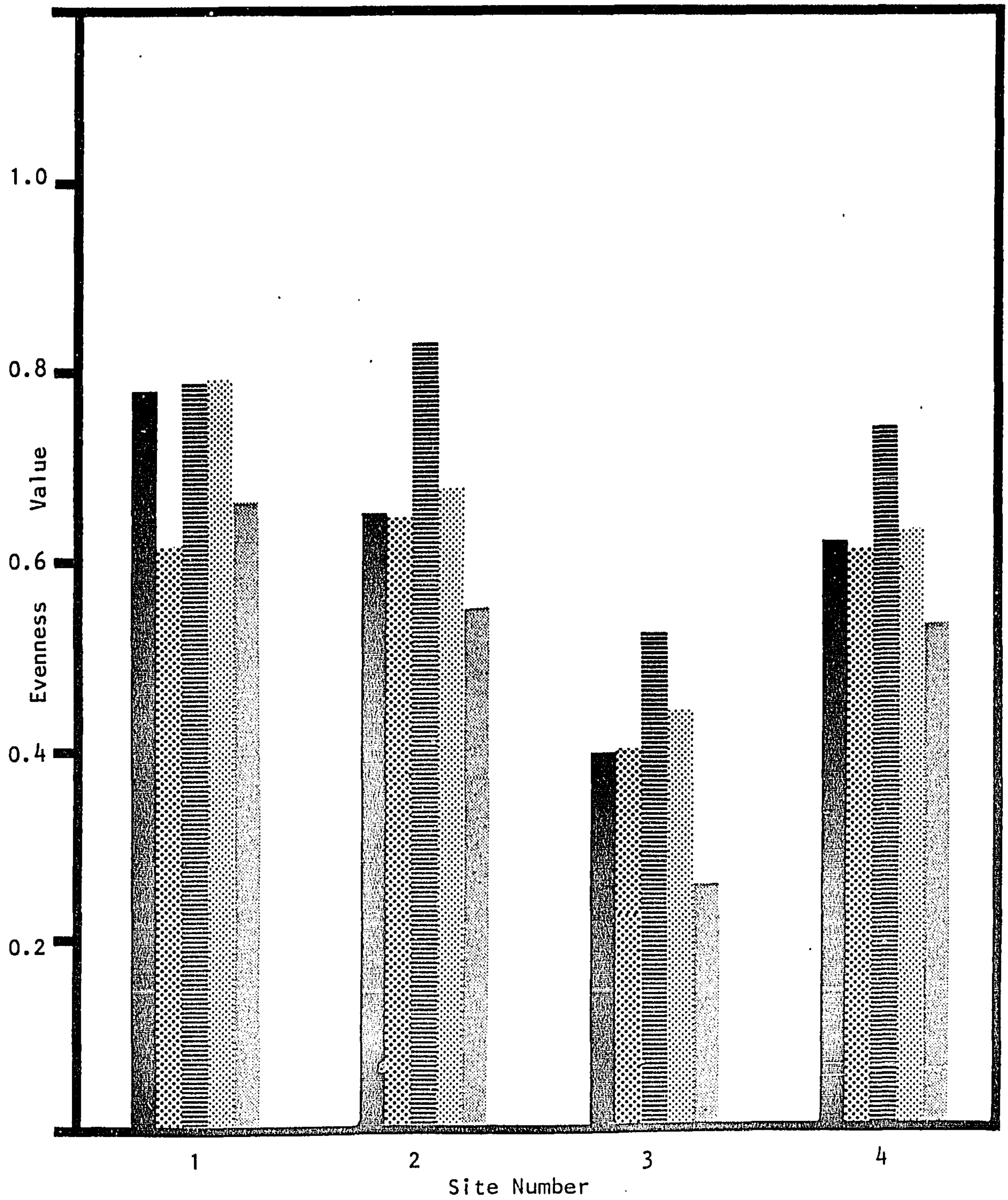

Figure 3. Comparison of Evenness

Table 6

Test Data

Test Set	A	B
Taxa	15	15
Individuals	400	400
Standard deviation	19.62	18.91
Approximate Index	2.341	2.348
*V	0.861	0.863
Brillouin's Index	2.270	2.179
V	0.855	0.818
SCI	0.845	0.848
V	0.948	0.951
Shannon's Index	3.378	3.387
**V'	0.858	0.860
Simpson's Index	2.310	2.323
V	0.846	0.851

*V = evenness of fully censused collection or community

**V' = evenness of sample from uncensused community

CHAPTER IV

DISCUSSION

Biological Data

A profile of the sampling sites along the Portage Creek will be developed utilizing different properties of the macroinvertebrate communities sampled. The properties to be discussed include tolerance, productivity and taxa richness, and a biotic index.

The ability of aquatic organisms to withstand environmental stress refers to its tolerance. There are three levels of tolerance generally accepted in water quality analysis: tolerant, intolerant, and facultative. The Environmental Protection Agency (EPA) defines these terms as:

> Tolerant - Organisms frequently associated with gross organic contamination which are generally capable of thriving under anearobic conditions.
>
> Facultative - Organisms which have a wide range of tolerance and frequently are associated with moderate levels of organic contamination.
>
> Intolerant - Organisms that are not found associated with even moderate levels of organic contaminants and are generally intolerant of even moderate reductions in dissolved oxygen. (EPA, 1973, 18-19)

All four sites (Table 1) include taxa belonging to each tolerance level (EPA, 1973; Hilsenhoff, 1977; Lehmkuhl, 1979). Relying on percent tolerant taxa, the sites rank in the order from less stressed to more stressed (by organic pollutants); #1, 2, 4, 3. The tolerant taxa are primarily numbers of the genus Asellus (Isopoda), of the

insect family Chironomidae, and the Oligechaeta family. Using percent intolerant and facultative taxa, the ranking of the sites is the same. The predominant number of organisms in this group are from the insect orders Ephemeroptera and Trichoptera, and the crustacean order Amphopoda. Based on the above data, the sites follow the sequence from #1,2,4, to 3 where; site #1 contains a macroinvertebrate community with the least stress from organic contaminants and site #3 with the most stress.

The productivity of an aquatic community can also give insight as to the effects of pollutant stress, primarily toxic and severe organic pollution. Since severe organic pollution appears to be lacking in Portage Creek (Table 5), toxic effects will be discussed. As a community experiences toxic pollution, the organism density declines (Persoone et al., 1978). According to Table 1, density is highest in site #1 and lowest in site #4. This difference in densities may result from either habitat change or toxicity.

As previously discussed, the creek's bottom changes from rock aggregate at site #1 to a silt-sand bottom with some debris for site #4. The coarse aggregate bottom of sites #1 and #2 are good substrates for different taxa of the mayflies and caddisfly, while the absence of such a habitat and an abundance of silt would not provide an ideal environment for most taxa of these two groups. One would expect to find an abundance of tubificid worms, snails, and leeches, especially at the low flow rate (Table 1). This is not the case. Not only is the productivity low, but the taxa richness (number) is reduced. Decrease in both taxa and population numbers indicate

possible effects from toxic wastes on the community(Persoone et al., 1978). Analysis by the DNR of sediment samples have identified lead, arsenic, and other heavy metals as being present. Further studies would be necessary to determine if the reduction of taxa and population were directly attributed to heavy metal toxicity. The condition of the site #4 community may also be the result from the flushing of sediment deposits during periods of high flow (after rainstorms). Visual observations of the site indicated little fluctuation in silt deposit depth and size.

Using productivity and taxa richness as indicators of community health (and water quality) indicate that site #1 is the healthiest community while site #4 is the least healthy site (possibly a result of toxicity). Sites #2 and #3 are similar, both healthier than #4, but not as healthy as site #1.

The modified "Score System" biotic index (Chandler, 1970) indicates site #1 and #2 are the least stressed (or healthiest), site #3 is slightly less, and site #4 is the most stressed. The biotic index can give much qualitative information on the community and its response to many different forms of stress (Hilsenhoff, 1977). The biotic index supports the richness and population data (Table 1) and possibly the conclusion that site #4 is stressed by toxic wastes. Since the biotic index value depends on the types and number of taxa present, one would expect the index to agree with taxa richness data.

In summary, the biological data offers two possible profiles for the water quality and community health of the sample sites. The first sequence (#1, 2, 4, 3) reflects the possible effects of organic

pollution depleting oxygen. The second sequence (#1, 2, 3, 4) reflects possible toxic pollution as a result of heavy metals in the sediment. Since the purpose of the profile is to test the accuracy of the various indices, the test profile must exhibit properties easily measured by diversity. The diversity index is most effective when measuring the effects of oxygen depleting organic wastes (Pielou, 1977; Hilsenhoff, 1977). Since toxic wastes tend to reduce both taxa richness and number of organisms, diversity indices can give false values if the individuals are distributed evenly among the remaining taxa. Therefore, the first sequence is #1, 2, 4, 3,

Chemical and Physical Data

The chemical and physical data (Table 5) indicate little difference among sample sites. The low temperature for site #2 probably results from the joining of the west fork of the Portage Creek to the main creek. Site #1 has the highest dissolved oxygen level while the other sites have identical lower values. All the values are above the 4.0 mg/1 lower limit recommended for a healthy stream by the DNR (1972). The same DNR study indicated that dissolved oxygen levels fluctuate radically with seasonal changes. The data from Table 5 came from May sampling times. The recorded values are similar to those indicated by the DNR for the time period measured.

The turbidity levels increased as the creek approached the center of the city (sequence #1, 2, 3, 4). The progressively higher turbidity may inhibit photosynthetic activity and affect community structure. The levels found are not excessive and the shallow depth of

the creek should allow sufficient light for benthic photosynthetic organisms.

Nutrient levels were adequate but not excessive since an overabundance of plant growth was not apparent (eutrophication). The nitrogen containing nutrients and phosphate appeared slightly lower in site #2. About 500 feet upstream from site #2 is located a dam and pond which temporarily slow the water flow. There is much vegetative growth in the pond which may utilize some of the nutrients. The nutrients are probably replenished with fertilizer runoff and effluent from the paper waste clarifier. The chlorides present may reflect city street runoff during rains. The salt left over from winter de-icing along with exhaust particulates (containing lead) are presumably washed into the creek.

In summary, the chemical and physical parameters measured indicate only small differences among sites. The data suggests that the water quality was good at the time of sampling. Since the chemical and physical parameters can change rapidly with time and season (DNR, 1972, 1983), adjustment of the profile to include this data would have little justification.

Background Data

The history of the Portage Creek indicates that site #4 has always been a highly polluted section of the creek (DNR, 1972, 1982). No studies were made on the upper section of the creek (site #1), but comparisons were made between sites #2 and #4. Site #2 always demonstrated better water quality. The chemical, physical, and biological

data obtained through sampling during the study tend to reflect values and collections presented by the DNR. The primary difference is that the thesis data gives a slightly healthier assessment of the water and community quality.

On several mornings during sampling periods, an oilish cast was observed on the creek's surface in the area of site #4. An employee of a firm located on the creek observed occasionally that the creek appeared milky gray in color. Both events suggest, along with the present chemical data, that the water is of generally good quality with occasional inputs of gross pollution.

Unless one can observe or chemically monitor every appearance of pollutants, chemical and physical parameters would not accurately measure perturbations caused by irregular instances of pollution.

In summary, historical data suggests site #4 is a very stressed location. The most recent data (DNR, 1983) suggests the water quality is improving chemically. The chemical testing of water quality may overlook occasional instances of pollution, especially in flowing waters. Therefore, long term monitoring is necessary to adequately detect pollutant input. Biological systems monitor continually over the life span of the organisms within the biotic community.

The biological community constantly monitors the water in which it exists, unlike the occasional testing of chemical and physical parameters. The biological data will, therefore, determine the profile against which the diversity indices will be evaluated.

Diversity Data

The water quality profile as determined by the biological data discussed earlier provides a model on which the diversity indices can be evaluated for accuracy. According to the profile, site #1 is a community experiencing the least amount of stress from organic pollutants (best water quality), followed by site #2, then site #4, and the most stressed, site #3. Although the collected data conflicts on the stressed states of sites #3 and #4, the sequence was chosen based on data more favorable to diversity indices (organic, oxygen depleting pollution).

On observing the various diversity values (Table 4, Figure 2), not one index fits the profile. Using evenness calculations for each index, unevenness within collections is compensated. According to the bar graph (Figure 3.), the Approximate, Shannon's, and Simpson's indices follow the same sequence of the proposed profile (sites #1, 2, 4, 3). Both Brillouin's and the SCI indicates the biota of site #2 is more diverse than site #1. Note that there is little difference between sites #1 and #4 using Brillouin's and SCI, which contradicts both present and historical data.

In summary, based on accuracy for detecting organic wastes, the Approximate, Shannon's and Simpson's indices are good indicators for stressed communities.

Index Sensitivity Data

The purpose of the test data sets was to evaluate the ability of each index to separate collections with similar distributions

(sensitivity) and to check for correct indications of diversity and evenness. The indices gave the expected values for each test collection except Brillouin's index. The diversity and evenness values for Brilluoin's index indicated that test set A was more diverse than B (Table 6). According to the definition of diversity and the standard deviation of each test set, Brillouin's index gave a false indication.

Of the diversity indices resulting in correct values, Simpson's demonstrated the greatest difference between test data sets using both diversity and evenness. Shannon's index produced the next greatest difference.

The ability of an index to distinguish differences among collections is very important when constructing a monitoring program. Such a program may only involve two test stations, one above the effluent input, the other below. If there is a change in community as a result of the effluent, the index should be able to distinguish any difference between the control and treated stations. A sensitive index could more rapidly identify effects from the effluent, resulting in faster action to reduce the damage to the biotic community and water quality.

CHAPTER V

CONCLUSIONS

After evaluations of index accuracy and sensitivity, Simpson's and Shannon's indices gave the best results of the indices tested on the study data. Simpson's index was determined "best" theoretically since sample and community diversity is actually measured and not estimated as with Shannon's index.

Using a microcomputer, the Simpson's index would fit into an accurate, sensitive and rapid biological monitoring program. Meaningful results were obtained with macroinvertebrate identification to the family level (Kaesler, 1977). Single samples of 350 or more individuals should be adequate for determining both diversity and evenness values. The recommended sample method uses a hand dip net, exhaustive sampling representative areas of the creek's bottom. Sample sizes should be similar to make the calculations most effective (Hurlbert, 1971).

For index calculation, the necessity of evenness calculations must be emphasized. As observed in the study, evenness values fit the stream profile while the diversity values alone probably did not. Diversity is a function of both species richness and the evenness or balance with which the individuals are distributed.

Further research is recommended using simulated and real data to investigate the validity of index calculations. The unexpected diversity and evenness value obtained using the Brillouin's index

on the test data requires more investigation.

Another unexpected result of this study is the possible use of standard deviation in the measurement of diversity. Although standard deviation may not be derived from information theory, it can give insight to the distribution of a collection around an arbitrary mean. The researcher has demonstrated the use of the standard deviation on collections with identical taxa number and size. Possibly by manipulating the standard deviation value to account for sample size (such as [SD/size] x constant), a simpler means of calculating diversity can be developed.

Diversity indices, especially the Simpson's index, are important parts of pollution monitoring programs. Diversity indices fit within a biological monitoring program (Ott, 1978) and should be complemented with other biological, chemical, and physical data (Worf, 1980) to obtain a comprehensive picture of water quality and its effect on the biotic community

APPENDIX

SENSITIVITY TEST DATA

Set#	1	2	A	B
15 species approximately 400-405 no.	27	1	1	1
	27	1	1	1
	27	1	1	1
	27	1	1	1
	27	1	1	1
	27	1	30	30
	27	1	30	30
	27	1	30	35
	27	1	40	40
	27	1	40	40
	27	1	40	40
	27	1	50	40
	27	1	50	40
	27	1	50	50
	27	386	35	50
	405	400	400	400
	-	-	19.26	18.91
	Max $\overline{D}$	Min $\overline{D}$	Test cases	

The Max and Min data sets, 1 and 2, are used for determining evenness. Test sets A & B are close in composition, except set B has a lower standard deviation than A. A good diversity index should: (1) select set B as the most diverse; (2) differentiate between A & B.

BIBLIOGRAPHY

Beck, W. M. (1954). Studies in stream pollution biology: I. A. simplified ecological classification of organisms. Quarterly Journal of Florida Academy of Science, 17, 211-227.

Branson, D. R., Armentrout, D. N., Parker, W. M., Van Hall, C., & Bone, L. I. (1981). Effluent monitoring step by step. Environmental Science and Technology, 15(5), 513-518.

Brillouin, L. (1962). Science and information theory. New York: Academic Press.

Cairns, J., Albaugh, D. W., Busey, F., & Chaney, M. D. (1968). The Sequential Comparison Index. A simplified method to estimate relative differences in biological diversity in stream pollution studies. Journal of Water Pollution and Control Federation, 40(9), 1607-1613.

Cairns, J., & Dickson, K. L. (1971). A simple method for the biological assessment of the effects of waste discharges on aquatic bottom dwelling organisms. Journal of Water Pollution and Control Federation, 43(5), 1-28.

Cairns, J., Dickson, K. L., Lake, W. (1976). Biological monitoring of water and effluent quality. Philadelphia: ASTM Publication (TP607).

Cairns, J., Dickson, K. L., & Lanea, G. (1973). Rapid biological monitoring system for determining aquatic community structure in receiving systems. Biological methods for the assessment of water quality. Philadelphia: ASTM Publication, 1973.

Chandler, J. R. (1970). A biological approach to water quality management. In G. Peroone & N. DePaun, Biological aspects of freshwater pollution. New York: Pergamon Press, 1978.

Department of Natural Resources (1972). Evaluation of the aquatic environment of the Kalamazoo River watershed, parts A, B, & C. State of Michigan.

Department of Natural Resources (1972). Water quality studies of Portage Creek in Kalamazoo County 1968 and 1980. Michigan: Bureau of Water Management.

Department of Natural Resources (1983). An intensive water quality study of the Kalamazoo River, Kalamazoo to Allegan. Michigan: Surface Water Quality Division.

Dickson, Slocumb, Cairns, Almeida, & Du (1976). A laser-based optical filtering system to analyze samples of diatom communities. Biological monitoring of water and effluent quality. Philadelphia: ASTM Publication (TP 607).

Fisher, R. A., Corbet, A. S., & Williams, C. B. (1943). The relation between the number of species and the number of individuals in a random sample of an animal population. Journal of Animal Ecology, 12, 42-58.

Gaufin, A. R. (1973). Use of aquatic invertebrates in the assessment of water quality. Biological methods for the assessment of water quality. Philadelphia: ASTM Publication.

Goodnight, C. J. (1973). The use of aquatic macroinvertebrates as indicators of stream pollution. Transactions of the American Microscopical Society, 92(1), 1-13.

Green, R. H. (1976). Some methods for hypothesis testing and analysis with biological monitoring data. Biological monitoring of water and effluent quality. Philadelphia: ASTM Publication (TP 607).

Hilsenhoff, W. L. (1977). Use of arthropods to evaluate water quality of streams. Technical Bulletin, No. 100. Wisconsin: Department of Natural Resources.

Hurlbert, S. H. (1971). The nonconcept of species diversity: A critique and alternative parameters. Ecology, 52(4), 577-586.

Kaesler, R. L., Herricks, E. E., & Crossman, J. S. Use of indices of diversity and hierarchial diversity in stream surveys. Biological data in water pollution assessment. Philadelphia: ASTM Publication (TP 652).

Kolkwitz, R., & Marsson, M. (1908). Okologie der pflanzlichen saprobien. Bericht. Deutsche Botanische Gessellschaft, 26, 505-519.

Kolkwitz, R., & Marsson, M. (1909). Okologie der tierischen saprobien beitrau zur lehre von der biologischen gewasserbeurteilung. Internationale Revue der Gesamten Hydrobiologie, 2, 126-152.

Kullberg, R. G. (1974). Distribution of aquatic macrophytes related to paper mill effluents in a Southern Michigan stream. The American Midland Naturalist, 91(2), 271-281.

Lehmkuhl, D. M. (1979). How to know the aquatic insects. Dubuque, Iowa: Wm. C. Brown Company.

MacArthur, R. H., & Wilson, E. O. (1967). The theory of island biogeography. Princeton, NJ: Princeton University Press.

Margalef, E. R. (1958). Information theory in ecology. General Systems, 3, 36-71.

Merritt, R. W., & Cummins, K. W. (Eds.) (1978). An introduction to the aquatic insects of North America. Dubuque, Iowa: Kendall/Hunt Publishing Company.

Ott, W. R. (1978). Environmental indices. Ann Arbor, MI: Ann Arbor Science.

Pantle, R., & Buck, H. (1955). Die biologische uborwachung der gewasser und die darstellung der ergenbnisse. Gas und Wasserfach, 96, 604.

Patrick, R. (1949). A proposed biological measure of stream conditions, based on a survey of the Conestaga Basin, Lancaster County, Pennsylvania. Proceedings of the Academy of Natural Sciences of Philadelphia, 101, 277-341.

Patrick, R. (1961). Proceedings of the Academy of Natural Sciences of Philadelphia, 113, 277-341.

Patrick, R. (1976). The importance of monitoring change. Biological monitoring of water and effluent quality. Philadelphia: ASTM Publication (TP 607).

Pennak, R. W. (1953). Freshwater invertebrates of the United States. New York: The Ronald Press Company.

Persoone, G., & DePaun, N. (1978). Systems of biologocal indicators for water quality assessment. Biological aspects of freshwater pollution. New York: Pergamon Press.

Phillips, D. J. H. (1980). Quantitative aquatic biological indicators. Essex, England: Applied Science Publishers.

Pielou, E. C. (1966). Species-diversity and pattern-diversity in the study of ecological succession. Journal of Theoretical Biology, 10, 370-383.

Pielou, E. C. (1975). Ecological diversity. New York: J. Wiley & Sons.

Pielou, E. C. (1975). Mathematical ecology. New York: J. Wiley & Sons.

Shannon, C. E., & Weaver, W. (1949). The mathematical theory of communication. Urbana, IL: University of Illinois Press.

Shannon, C. E., & Weaver, W. (1963). The mathematical theory of communication. Urbana, IL: University of Illinois Press.

Simpson, E.H. (1949). Measurement of diversity. Nature, 163, 688.

Slocumb, J., & Dickson, K. L. (1978). Estimating total number of species in a biological community. Biological data in water pollution assessment. Philadelphia. ASTM Publication (TP 652).

Weber, C. I. (Ed.) (1973). Biological field and laboratory methods for measuring the quality of surface waters and effluents. Environmental Protection Agency (670, 4-73-001).

Why maintain biological diversity? (1982). Environmental Science and Technology, 16(2), 94A-97A.

Wilhm, J. L., & Dorris, T. C. (1966). Species diversity of benthic macroinvertebrates in a stream receiving domestic and oil refinery effluents. The American Midland Naturalist, 76(2), 427-449.

Wilhm, J. L., & Dorris, T. C. (1968). Biological parameters for water quality criteria. Bioscience, 18, 477-481.

Woodiwiss, F. W. (1964). The biological system of stream classification used by the Trent River Board. Chemistry Industry, 14, 443-447.

Worf, D. L. (1980). Biological monitoring for environmental effects. Lexington, MA: Lexington Books.

Zaret, T. M. (1982). The stability/diversity controversy: A test of hypotheses. Ecology, 63(3), 721-731.