The Use of Cluster Analysis in the Assessment of Spills of Hazardous Materials

JOHN S. CROSSMAN[1]
ROGER L. KAESLER[2]
and
JOHN CAIRNS, JR.
Department of Biology and Center for Environmental Studies, Virginia Polytechnic Institute and State University, Blacksburg 24061
Department of Geology, University of Kansas, Lawrence 66045

ABSTRACT: The macrobenthic community of Clinch River, near Carbo, Virginia, has twice been subjected to acute stress caused by major industrial spills from a power plant. The first spill, which resulted in a high *p*H shock, was from a fly-ash retaining pond in 1967. The second was an acid spill in 1970 with consequent low *p*H shock. Stream surveys were made in 1969, 1970 and 1971. This paper reports the results of Q-mode cluster analysis of presence-absence data on the total aquatic insect fauna, several orders of insects considered separately, and Gastropoda from those surveys.

Recovery from the effects of the fly-ash spill by all elements of the fauna studied except the Gastropoda was well under way by the summer of 1969. Nevertheless, the insect fauna in samples from the area affected by the spill was still different from that in unpolluted reaches of the stream, although it was not possible to discriminate between remnant effects of the spills and chronic stress due to the day-to-day operation of the power plant. The spill of acid in 1970 eliminated many elements of the fauna from about 30 km of the river. Again, by the end of the summer recovery was well under way for all groups except the Gastropoda.

Cluster analysis was particularly useful in determining the effects of type of substrate, time of sampling, longitudinal succession, and flooding on the composition of the macrobenthic community. It is suggested that one effect of flooding may be to make the fauna more homogeneous so there is a more nearly equal distribution of macrobenthic organisms among the stations from which samples were collected.

INTRODUCTION

Since 1967, the Clinch River in southwestern Virginia has been subjected to two major industrial spills and several episodes of flooding that have had serious deleterious effects on the fish population and on the community of macrobenthic invertebrates (Table 1). These spills and floods and their effects on density and diversity of the aquatic community have been discussed previously (Cairns *et al.*, 1971; Cairns *et al.*, 1972; Crossman *et al.*, 1973). The paper by Crossman *et al.* also presented cluster analysis of some of the data on aquatic insects. The purpose of this presentation is to make more extensive use of

---

[1] Present address: Teledyne Brown Engineering, Huntsville, Alabama.
[2] Work was done while author was Visiting Professor in Department of Biology, Virginia Polytechnic Institute and State University, Blacksburg 24061.

cluster analysis in order to further assess the responses of the aquatic community to the two spills.

The Clinch River is one of the headwater tributaries of the Tennessee River (Fig. 1). It flows through steep mountainous topography of the valley and ridge physiographic province, and flooding is common in the winter and spring. In addition, hurricanes moving up the Atlantic coast may cause heavy flooding in the late summer and early autumn (Table 1). The drainage basin is largely rural. More than half the basin is forested; the remainder is cropland, pastureland and rock quarries, with appreciable amounts of coal mining on the NW flank of the basin. Acid drainage from the coal mines is usually neutralized by the carbonate terrain through which the river flows, but coal-washing operations may have deleterious effects on the macrobenthic community, particularly along some of the Clinch River's tributaries. Crossman *et al.* (1973) have presented a detailed description of the study area.

In the vicinity of the Appalachian Power Company's 700-Mw steam power-generating plant at Carbo, Va., the Clinch River has been exposed to acute stress from the two industrial spills mentioned previously and possibly to chronic stress as a result of the day-to-day operation of the power plant. In addition, other reaches of the stream

TABLE 1.—Events with a major effect on the macrobenthic community of the Clinch River

Date	Event
June 1967	Fly-ash pond spill, high *p*H
August 1969	Hurricane Camille, flooding
19 June 1970	Acid spill, low *p*H
7 -11 May 1971	Flooding
13-16 May 1971	Flooding

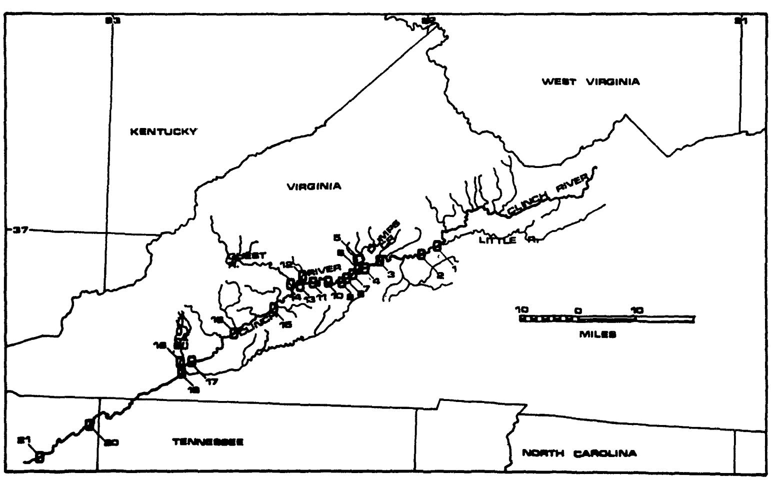

Fig. 1.—Map of the Clinch River in Virginia and Tennessee showing tributaries and collecting stations

were exposed to untreated sewage, acid mine-drainage and possibly other contaminants.

The first of the two spills that were the basis of this study occurred in June 1967 when a dike surrounding a fly-ash retaining pond at the power station failed. Within less than an hour, 4.9×10^5 m^3 of caustic waste (*p*H 12.0 to 12.7) poured into Dump's Creek and into the Clinch River less than 1 km downstream (Anonymous, 1967b). This caustic slug equalled 40% of the Clinch River's daily flow at the time of the spill, and it blocked the normal flow for several minutes. In addition, due to the raised water level, some of the waste flowed upstream for nearly 1 km (Cairns *et al.,* 1972; Crossman *et al.,* 1973). The alkaline slug moved downstream at about 1.5 km per hr killing virtually all the fish in its path for about 4½ days. During this period more than 215,000 fish were killed in Virginia and Tennessee before

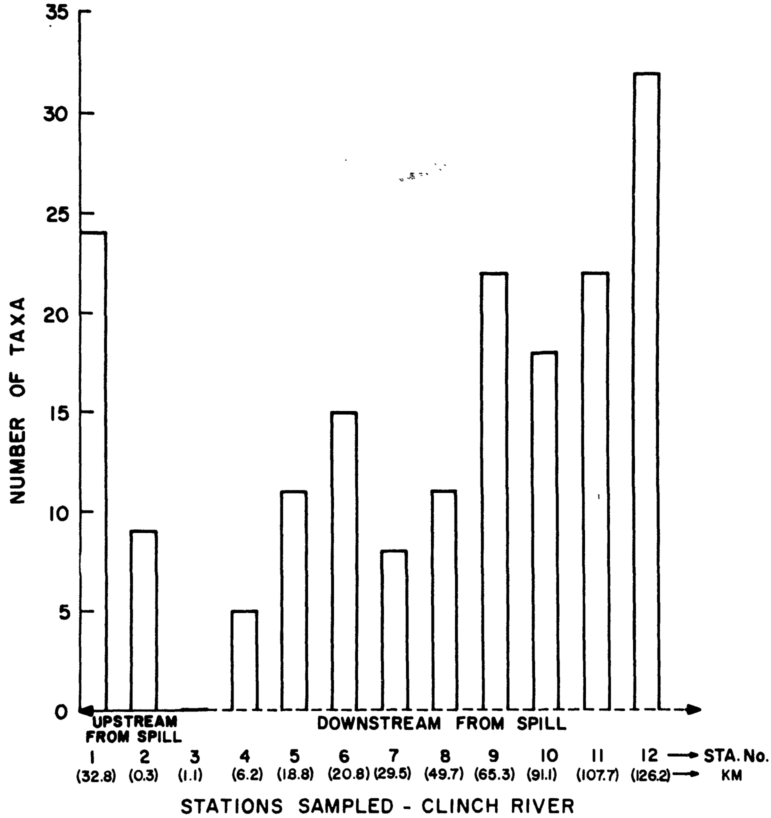

Fig. 2.—The number of taxa of benthic macroinvertebrates found at stations above and below the power plant 10 days after the fly-ash pond spill. Sample numbers do not coincide with those from later surveys; spill occurred between Stations 2 and 3

the high *p*H was neutralized.

As a result of the high *p*H and possibly a depression in dissolved oxygen, bottom-dwelling fish-food organisms were completely eliminated for 5 to 6 km below the site of the spill. The number and kinds of fish-food organisms were drastically reduced for about 125 km below the spill (Figs. 2, 3), and snails and mussels were eliminated for 18 km

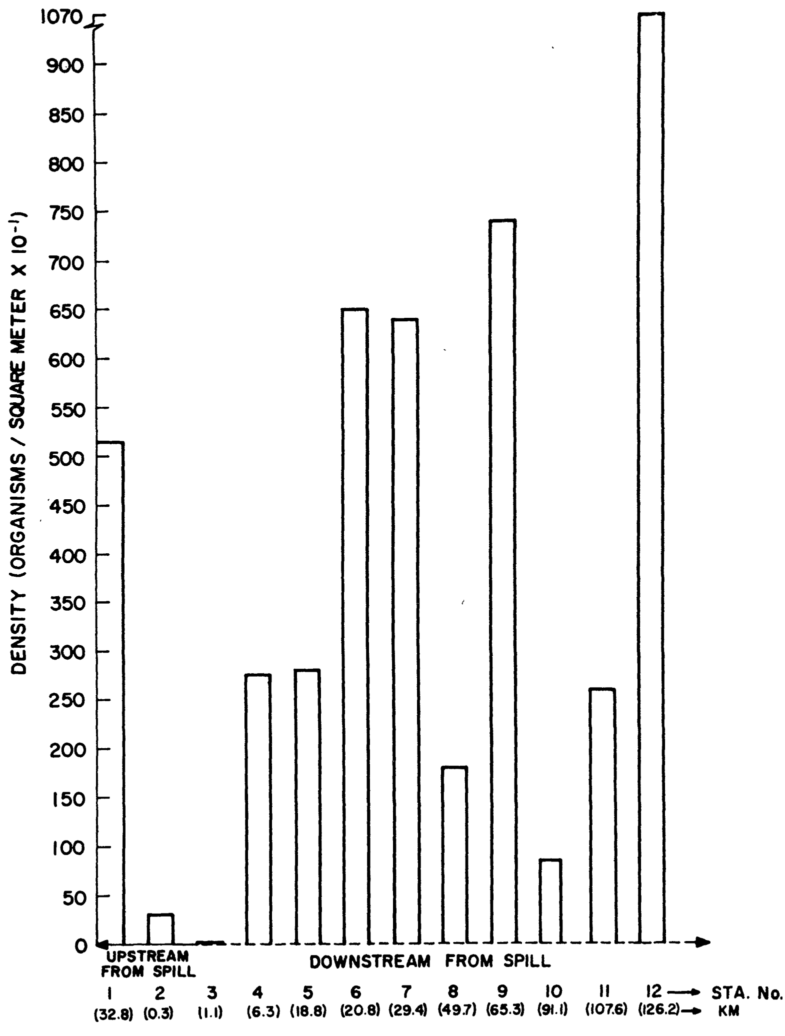

Fig. 3.—Density of benthic macroinvertebrates at stations above and below the power plant 10 days after the fly-ash pond spill. See Figure 2 for explanation of station numbers

(Crossman *et al.*, 1973). The Virginia State Water Control Board predicted that, with respect to total weight per unit area, the biota of the Clinch River would recover within 3 months of the spill, and both the Virginia and the Tennessee Game Commissions believed that restocking would be successful in the autumn of 1967 (Anonymous, 1967a).

The second spill at the power plant occurred in June 1970 just after our 2nd year's survey of the bottom fauna had been completed. This spill involved the release of an undetermined amount of sulfuric acid that killed 5300 fish. Damage to the aquatic community began about 1.5 km below the power plant and extended downstream for 22 km.

It is important to distinguish between chronic and acute stresses of the aquatic environment. Chronic stress may result from a long-term addition of an effluent to a stream. If it produces severe biological effects, one expects to see a linear pattern of change in the aquatic community downstream from the outfall. This linear pattern represents a gradual *accommodation* or restoration (as evidenced by increased diversity) of the community to improved water quality as one progresses downstream. This is achieved by means of both dilution and transformation of the waste. Acute stresses resulting from industrial spills such as the ones described here have a different net effect. The composition of the community may change gradually with distance downstream, but when the slug from the spill has passed, the community may recover from the stress in a way that is not possible with chronically stressed environments. Measures of diversity are well-suited for the study of accommodation to chronic stress, but where acute stress has occurred, diversity values may be misleading simply because stream recovery presupposes an upstream or tributary source of organisms for recolonization. Since our research was focused on recovery from massive, acute pollutional events, we have chosen to augment the use of diversity with cluster analysis, which deals with specific organisms in the zone of recovery and not merely the diversity of the community.

MATERIALS AND METHODS

Information on the location of the stations at which samples were collected is given in Table 2. A more complete discussion has been given by Crossman *et al.* (1973). The stations were selected to include similar habitats so that comparisons could be made between them. Care was taken to make an equal effort in collecting at each station. In all cases, stations were divided into three substations, the left bank, mid-channel and right bank facing upstream, in order to assess lateral differences that might result from channelization of discharged waste. Four stations on the Clinch River were located above the site of the power plant, and 12 stations were located below it. Five stations were established on tributaries of the Clinch River with pollution sources possibly deleterious to the main stream.

Samples of the macrobenthic community were taken with Turtox 20 × 25 × 45 cm rectangular bottom nets and a Surber 0.1 m^2

sampler. For purposes of the research discussed here, only the presence or absence of species in each sample was noted, although data on absolute abundances were also recorded (Crossman *et al.,* 1973). Allen and Koonce (in press) have demonstrated the usefulness of presence-absence data in comparisons of ecosystems. In our study, presence-absence data may carry most of the useful information because we are dealing with tolerances of organisms to pollutants.

Samples from three limnological surveys were used in this portion of the study (Table 3). In the summer of 1969, the macrobenthic community was surveyed in order to determine the extent of the biological recovery from effects of the 1967 fly-ash spill. All stations were sampled except those located on tributaries. Sampling was done according to the schedule in Table 3. Long intervals between times of sampling were necessitated by the disruptive effects of high water due to flooding at several times during the summer. At times of high flow, sampling the macrobenthic community was physically impossible, and the composition of the community itself was likely to be adversely affected by the swiftly moving water.

Just after a portion of a follow-up survey had been completed during the summer of 1970, the spill of sulfuric acid occurred. A series of samples was collected during the remainder of the summer according to the schedule shown in Table 3. Note the lettering scheme used. The letters A through F refer to the time at which the sample was collected. The first time a station was sampled in 1970 it was assigned the letter A, the second time the letter B, and so on to the sixth time (F). Not all stations were sampled during each sampling period, nor were they all sampled the same number of times. Thus the third time Stations 7, 8, 9 and 10 were sampled (7-9 July) was before the second time Stations 4, 11 and 13 were sampled (21-24 July). It is important to note, then, that samples with the same letter were not necessarily sampled at the same time. This lettering scheme has been used in all references to the samples collected during 1970. In addition, all sample numbers are

TABLE 2.—Descriptions of stations sampled on the Clinch River and its tributaries (from Crossman *et al.,* 1973)

Station number	Distance from mouth (km)	Stream	Station number	Distance from mouth (km)	Stream
1	477.3	Clinch River	12	......	Bull Run Creek
2	464.1	" "	13	396.2	Clinch River
3	439.0	" "	14	......	Guest River
4	433.6	" "	15	381.4	Clinch River
5	......	Dump's Creek	16	366.0	" "
6	......	" "	17	350.7	" "
7	430.0	Clinch River	18	......	Stock Creek
8	426.5	" "	19	343.0	Clinch River
9	423.4	" "	20	304.8	" "
10	412.5	" "	21	286.8	" "
11	401.9	" "			

TABLE 3.—Times of sampling in 1969-71 and alphabetical notation for the different sampling periods. Letters A to F refer to the sequence of sampling at a particular station during the 1970 survey

Station number	1969 Mid-June	Late July	Late Aug. Early Sept.	1970 6/9-6/12	6/18-6-19	6/23-6/26	6/29-7/2	7/7-7/9	7/21-7/24	8/5-8/10	8/15-8/18	8/24-8/28	1971 6/14 6/11-
1			X	A								B	X
2			X	A								B	X
3	X	X		A								B	X
4				A					B		C		X
5T				A								B	X
6T				A								B	X
7	X			A		B		C	D	E	F		X
8	X				A	B		C	D	E	F		X
9		X			A	B		C	D	E	F		X
10		X			A	B		C	D	E	F		X
11		X				A			B		C		X
12T						A						B	X
13		X				A			B		C		X
14T						A						B	X
15			X				A					B	X
16			X				A					B	X
17			X				A					B	X
18T							A					B	X
19			X				A					B	X
20			X				A					B	X
21			X				A					B	X

prefixed by the station number and suffixed by an R, MC or L to indicate right bank, mid-channel or left bank facing upstream.

During June 1971, each substation was sampled a single time (Table 3). Previous to the time of sampling, on 7-11 May and 13-16 May, the community of macrobenthic organisms was disrupted by high water conditions. Recovery seems not to have been complete at most stations by the time they were sampled.

Cluster analysis has been widely used in ecological work. Williams (1972) has provided an excellent discussion of many aspects of its application. We have chosen to base our measurements of similarity on Jaccard's coefficient (Jaccard, 1908) and to cluster using the unweighted pair-group method with arithmetic averages (Sokal and Sneath, 1963). These methods have been discussed by Kaesler and Cairns (1972) and by Crossman *et al.* (1973).

Since the surveys extended over 3 summers, and because of the nature of the pollution events that occurred, each year's work has been considered separately. The large size of the data matrices precluded simultaneously clustering of all the data collected each year. Therefore, the samples were separated into insect and noninsect fractions. After separation, cluster analysis was done on each year's total insect fauna as well as each taxonomic order (Table 3). Other noninsect taxa were also analyzed as total other macroinvertebrates and Gastropoda. A total of 26 dendrograms were formed (Table 4), although not all of them have been discussed here. Those dendrograms providing little information over and above the ones discussed have been omitted. The results of cluster analysis supported the results obtained by study of community structure (density and diversity) and in several instances suggested relationships that had not been recognized previously.

RESULTS AND DISCUSSION — 1969 SURVEY

Total aquatic-insect fauna.—Figure 4A shows the results of Q-mode cluster analysis (Sokal and Sneath, 1963) of the total aquatic-insect fauna from 48 benthic samples collected from the substations surveyed during the summer of 1969. When 0.4 is chosen as the limit of simi-

TABLE 4.—Cluster analyses computed, dendrogram presented here, XX; cluster analysis computed only, X; cluster analysis not computed, ...

Subset of fauna	Year 1969	1970	1971
Total insects	XX	XX	XX
Ephemeroptera	X	XX	X
Coleoptera	X	X	X
Odonata	X	X	......
Trichoptera	X	X	X
Diptera	X	X	X
Plecoptera	X	X	X
Total other macroinvertebrates	X	X	X
Gastropoda	XX	X	X

larity, eight clusters are formed. Cluster A at the top of the dendrogram indicates similarity among samples from five of the six substations at Stations 1 and 2. Three factors probably contributed to the high similarity of these samples. First, Stations 1 and 2 were sampled much later in the year than the other streams, possibly resulting in faunal differences due to seasonal variation or succession. Second, although these samples were collected more than 6 weeks after the flooding associated with Hurricane Camille, the macrobenthic insect populations may not have had adequate opportunity to recover from the flushing action of the high stream flow. Third, the stream gradients at Stations 1 and 2 are higher than at any other stations with the exception of Station 13. A distinct benthic macroinvertebrate fauna could be expected at these stations due to any of the three factors above acting alone or in conjunction.

Cluster B is the largest, comprising 18 samples. Most of the samples in this cluster were from Stations 15 through 21, all of which were collected 10-14 September, 30 to 34 days after Hurricane Camille's flooding. The only samples from substations not in this interval are

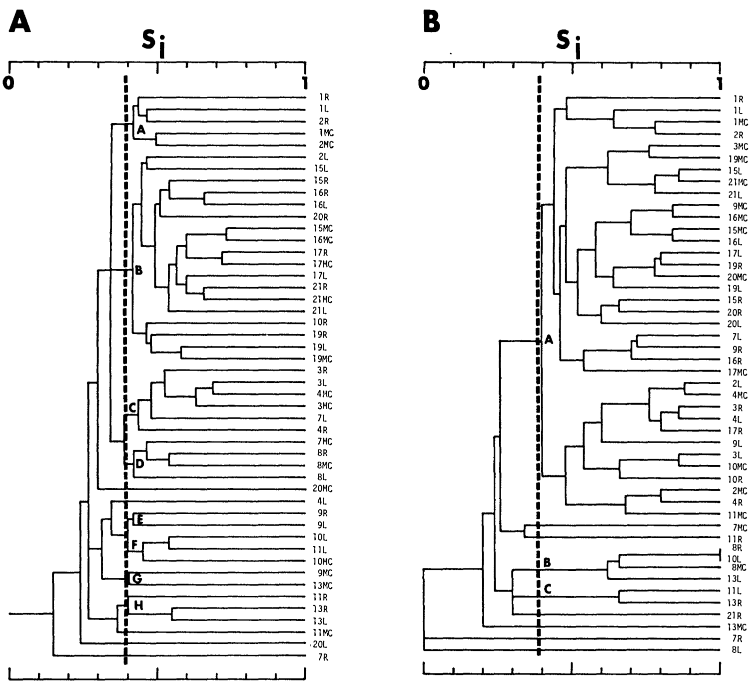

Fig. 4.—A. Dendrogram from Q-mode cluster analysis of presence-absence data on total insect fauna from 1969 survey; cophenetic correlation = 0.793; B. dendrogram from Q-mode cluster analysis of presence-absence data on Gastropoda from 1969 survey; cophenetic correlation = 0.825

samples 20L and 20MC. Their lack of similarity to other samples in this cluster can almost certainly be attributed to the rock substrate that was present at Station 20. This cluster may also reflect the effects of the time of year of sampling, the relatively low stream gradient in this section of the stream, substrate dissimilarities or, possibly, the effects of flooding.

Since all of the substations in clusters C and D except 4R and 4MC were sampled during the same time period and had similar substrates and stream gradients, a cause other than these must be sought to account for the two clusters and their separation. Five of the six samples in cluster C were collected at stations above the power plant. The other, 7L, is opposite the power plant and has been determined by measures of diversity to have been unaffected by discharges from the power plant. All samples in cluster D were collected downstream from the power plant. The separation of these two clusters supports the view of Cairns *et al.* (1971) that a combination of interrelated factors affected Stations 7 and 8: the 1967 fly-ash spill, the severity of initial damage from the spill, differing rates of recolonization by different species, and the existence of a mild, chronic environmental stress upon a limited portion of the stream caused by the power station. Note the position of Sample 7R in Figure 4A relative to the other samples. This substation is dissimilar to other samples because it was located immediately downstream from the power plant's discharges.

The remaining four small clusters (E, F, G and H) are each made up almost entirely of samples from the same station or substations of adjacent stations. Faunal similarity of adjacent substations suggests a longitudinal accommodation to the effluents from the power station. Samples from Station 13 may also be affected by the rock-slab substrate and the high stream gradient at this station, both of which might act as factors limiting colonization by macrobenthic organisms.

Cluster analyses of data on distributions of Ephemeroptera, Trichoptera, and Plecoptera were similar to that of the total insect fauna.

Gastropoda.—Figure 4B shows the results of cluster analysis of data on Gastropoda. This dendrogram was included for two reasons. First, the cluster analysis of data on gastropod distribution showed almost exactly the same results as the analysis of the total noninsect macroinvertebrate fauna. Second, recolonization by the molluscs was known from previous study to be very slow in damaged reaches of streams. The cluster analysis of the Gastropoda shows the effects of the slow rate of recolonization as well as giving a good impression of the results for all noninsect macroinvertebrates.

When a level of similarity just less than 0.4 is chosen, three clusters are formed with six samples left unclustered. Two of the unclustered samples were barren of gastropods. We have chosen not to subdivide cluster A because, left intact, it shows the effect on the snail population of the 1967 spill from the fly-ash pond, of chronic stress from the power plant, or a combination of these. For the most part, samples that are not part of the large cluster are from stations affected by the 1967 spill and also in a position to be chronically stressed by the day-to-day operations of the power station. These include Substations 7R, 7MC, 8R, 8MC, 8L, and 10L. Other samples not included in the large cluster include those from Substations 11R, 11L, 13R, 13MC, 13L and 21R. Samples from 11R, 11L, and 21R were all collected at substations with gradients lower than 1.1 m per km, although not all stations with low gradients were excluded from the large cluster. The exclusion of Substations 13R, 13MC, and 13L was attributed to the scouring effects of the swift current at this station as well as unfavorable substrate conditions.

RESULTS AND DISCUSSION — 1970 SURVEY

Total aquatic-insect fauna.—Figure 5 shows the dendrogram computed from data on the total aquatic-insect fauna collected from 163 substations during the 1970 survey. If a limit of similarity of 0.45 is chosen, 19 clusters are formed with several samples left unclustered, particularly samples collected from tributaries of the Clinch River at Stations 5, 6, 12, 14 and 18. Cluster A consists of the six samples from Substations 1AR through 2BL. These samples were collected from the substations that are the farthest upstream, and all of them were collected in late August 1970 (Table 3). All substations represented in the cluster were characterized by high stream gradients and velocity, similar substrates, and lack of pollution. That these samples cluster together underscores the importance of seasonal succession and time of sampling, as do clusters B and C.

Clusters D, E and F are large and closely similar. It is unwarranted to make too much of their differences. Nevertheless, the small differences that exist seem to be attributable to different times of sampling and differences in stream gradient. Most of the samples in these clusters, especially in E and F, are from Stations 15 through 21, the stretch of stream that is farthest downstream and has the lowest stream gradients, highest discharge, lowest velocity and a drainage area typified by farmland rather than mining areas found in the upper reaches of the stream.

The composition of cluster H is strongly influenced by the time samples were taken in relation to the time of the acid spill on 19 June. The following observations were made from Table 5 where the cluster is summarized:

(1) Except for sample 13AL, all A-suffix samples from Stations 4, 11 and 13 are present in the cluster. As shown in Table 3, these samples were collected during June 1970.

(2) One or more samples from each of Stations 4, 11 and 13 collected after 26 June 1970 are present in the cluster. These samples have a suffix starting with B or C.

(3) Samples collected at Stations 8, 9 and 10 on 18 and 19 June (A-suffix) are all in this cluster, along with samples from Station 10 collected in August (E- or F-suffix).

(4) Samples from the mid-channel and left-bank substations at

Station 7 are usually found in this cluster irrespective of the time of collection. However, samples from the right-bank substations are not found in the cluster but are scattered throughout the remainder of the

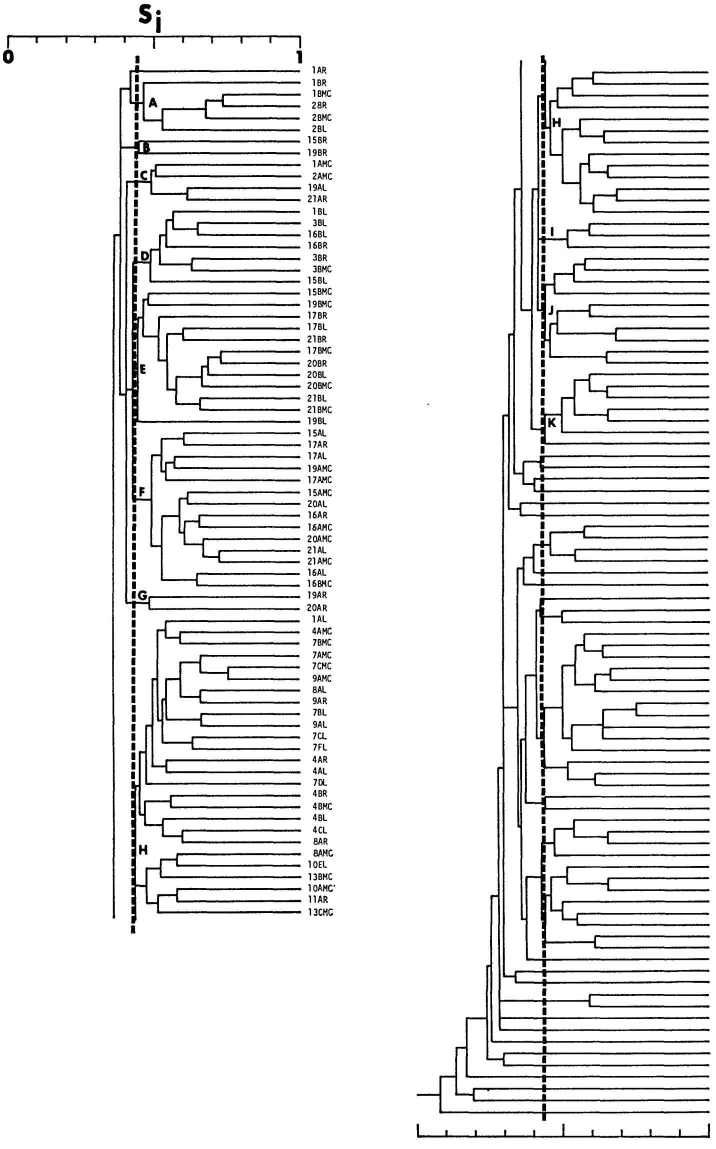

Fig. 5.—Dendrogram from Q-mode cluster analysis of presence-absence data on total insect fauna from 1970 survey; cophenetic correlation = 0.753

dendrogram and have very low levels of similarity with other samples.

(5) None of the samples collected from Stations 8 and 9 after the 19 June acid spill are in the cluster.

(6) Of the samples collected at Station 10 from 19 June-5 August (suffixes B, C, and D), only 10BR and 10DR are in the cluster.

Cluster H is clearly related to the effects of the acid spill. Because of the channel location, damage at Station 7 was restricted to the right bank adjacent to the power station. Mixing occurred rapidly, and damage was widespread at Stations 8, 9 and 10, with the exception of Substation 10R. By the time the acid reached Substation 10R, its toxicity had been nearly neutralized by natural physicochemical reactions in the stream, and organisms along the right bank of Station 10 apparently were able either to tolerate the acidic slug as it passed or to avoid it altogether.

Three reference stations were sampled to furnish a basis for com-

TABLE 5.—Samples from the largest cluster of Figure 5, total aquatic insect fauna, 1970 survey

Left bank	Mid-channel	Right bank
1 AL	–	–
4 AL	4 AMC	4 AR
4 BL	4 BMC	4 BR
4 CL	–	–
–	7 AMC	–
7 BL	7 BMC	–
7 CL	7 CMC	–
7 DL	–	–
–	–	–
7 FL	–	–
8 AL	8 AMC	8 AR
–	–	–
–	–	–
–	–	–
–	–	–
–	–	–
9 AL	9 AMC	9 AR
–	–	–
–	–	–
–	–	–
–	–	–
–	–	–
10 AL	10 AMC	10 AR
–	–	10 BR
–	–	–
–	–	10 DR
10 EL	10 EMC	10 ER
–	–	10 FR
11 AL	11 AMC	11 AR
11 BL	–	–
–	11 CMC	–
–	13 AMC	13 AR
–	13 BMC	–
–	13 CMC	–

parison. These were Station 4 above the power plant and Stations 11 and 13 located 28.1 and 33.8 km below the power plant, respectively. Stations 11 and 13 were sampled after spill, and it was important to determine their degree of similarity with Station 4 and other stations sampled before the spill. Since samples from Stations 11 and 13 clustered with the stations sampled before the spill, it is possible to conclude that damage was limited to the section of the river from Substation 7R to Station 10 (*see also* Cairns *et al.,* 1971).

Cluster J consists of 10 samples, most of which were collected 4 or more weeks after the acid spill, presumably after biological recovery from the spill was well under way. Samples 8CL and 10CMC, collected 19 and 20 days after the spill, are in this cluster and suggest that biological recovery may have occurred more rapidly at these main-channel substations than at other substations in the cluster, due to increased drift. Note that samples in this cluster are, on the average, not very different from samples of the previous two clusters and the succeeding one.

Examination of the remaining clusters indicates a possible recovery sequence. Samples from Stations 8 and 9 tend to cluster with each other depending on the elapsed time after the spill when they were collected. In several instances, samples collected from substations of one of these stations are closely similar to the corresponding samples from the other station. This result is not surprising since the two stations are only 3.2 km apart.

Several samples at the bottom of the dendrogram (Fig. 5) did not join any cluster at the limit of similarity of 0.45. With only a few exceptions, these samples were from tributaries, stations numbered 5, 6, 12, 14 and 18. The low similarities of samples from tributaries to each other and to samples from the Clinch River are indicative of the paucity of the macrobenthic fauna in the tributaries, presumably because the aquatic environment in most of the tributaries had been stressed by the wastes from coal-washing operations. In addition, Station 14 received a nutrient load from the town of Coeburn, Va.; and Station 18 on Stock Creek was affected by seepage from stored underground alkaline wastes from the Foote Mineral Company (Crossman *et al.,* 1973).

In the cluster analysis of data from the 1969 survey, stream discharge was important in explaining the results. During the summer of 1970, no flooding occurred. This lack of flooding was fortuitous since it could have seriously complicated both sampling and the interpretation of our results due to its unmeasurable effect on the stream biota.

Ephemeroptera.—The dendrogram showing relationships among samples on the basis of Ephemeroptera is shown in Figure 6A. It is included here because of the proclivity of mayfly larvae to drift in streams, suggesting that at some seasons of the year they may be among the first to recover from acute pollution events.

If a level of similarity of 0.6 is chosen, 17 clusters result, one of which is a barren cluster of 12 samples containing no mayfly nymphs.

Only six samples remained unclustered. Cluster C, summarized in

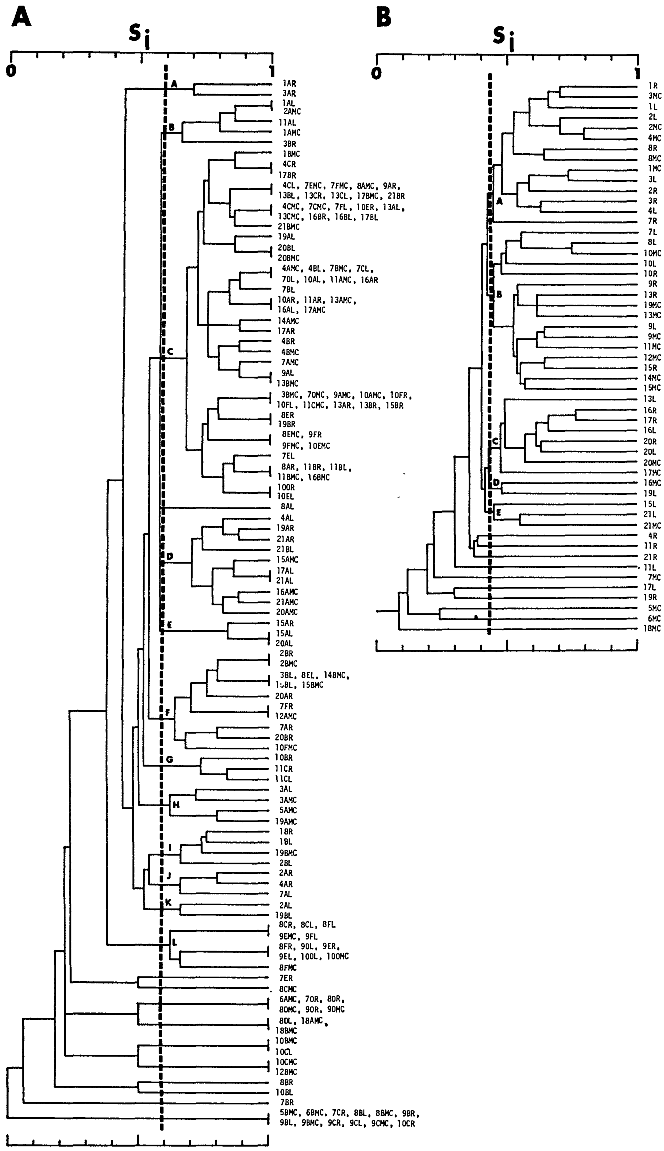

Fig. 6.—A. Dendrogram from Q-mode cluster analysis of presence-absence data on Ephemeroptera from 1970 survey; cophenetic correlation = 0.917; B. dendrogram from Q-mode cluster analysis of presence-absence data on insect fauna from 1971 survey; cophenetic correlation = 0.868

Table 6, is the only one discussed here. It is of interest largely because of the samples it does not contain. Samples from three parts of the

TABLE 6.—Samples from Cluster C of Figure 6, Ephemeroptera, 1970 survey

Left bank	Mid-channel	Right bank
–	–	–
–	1 BMC	–
–	–	–
–	–	–
–	–	–
–	3 BMC	–
–	4 AMC	–
4 BL	4 BMC	4 BR
4 CL	4 CMC	4 CR
–	7 AMC	–
7 BL	7 BMC	–
7 CL	7 CMC	–
7 DL	7 DMC	–
7 EL	7 EMC	–
7 FL	7 FMC	–
–	8 AMC	8 AR
–	–	–
–	–	–
–	–	–
–	8 EMC	8 ER
–	–	–
9 AL	9 AMC	9 AR
–	–	–
–	–	–
–	–	–
–	–	–
–	9 FCM	9 FR
10 AL	10 AMC	10 AR
–	–	–
–	–	–
–	–	10 DR
10 EL	10 EMC	10 ER
10 FL	–	10 FR
–	11 AMC	11 AR
11 BL	11 BMC	11 BR
–	11 CMC	–
13 AL	13 AMC	13 AR
13 BL	13 BMC	13 BR
13 CL	13 CMC	13 CR
–	14 AMC	–
–	–	–
–	–	15 BR
16 AL	–	16 AR
16 BL	16 BMC	16 BR
–	17 AMC	17 AR
17 BL	17 BMC	17 BR
19 AL	–	–
–	–	19 BR
–	–	–
20 BL	20 BMC	–
–	–	–
–	21 BMC	21 BR

stream are excluded from the large cluster. Samples from only two substations above Station 4 have been included in the cluster. At these upstream stations, the stream gradient is quite steep. Mayfly diversity is reduced in this area, thus accounting for the low similarity of samples from those substations to samples from the region between Stations 4 and 13. Similarly, samples from below Station 13, but especially from Stations 19, 20 and 21, have clustered elsewhere in the dendrogram. This area was one of low stream gradient, which favored burrowing or sprawling forms but not the ephemeropterans living at Stations 4 to 13.

The cluster summarized in Table 6 and the barren cluster demonstrate the pronounced effect of the acid spill on the ephemeropterans. Samples collected before the acid spill (Table 3) had a relatively high diversity except for the sample from Substation 7R. All were similar enough to each other to be in this cluster. Samples collected at Stations 8, 9 and 10 after the spill had low diversities and were too dissimilar to the samples collected before the spill to cluster with them. By late July at Substation 10R and early August at the other stations, recovery of the ephemeropteran portion of the macrobenthic community was well under way, as evidenced by samples collected from the zone of pollution clustering with samples collected before the spill. Note that samples from Substation 7R, while rarely barren, were always of low diversity and never joined this cluster of samples from presumably healthy stations.

Other fauna.—Cluster analysis of other groups of organisms showed similar results and a few interesting exceptions. The Tricoptera appear to have been affected very little by the acid spill or to have made a rapid recovery from its effects.

The Odonata yielded a large barren cluster comprising 48 samples, 36 of which were from mid-channel substations. Throughout the study it had been noted that the Odonata favored habitats near the banks of the stream rather than in mid-channel. The Plecoptera produced three clusters of interest. Two of these, a barren cluster of 32 samples and a cluster of 41 samples that contained only *Acroneura,* included almost all samples from the area affected by the spill. They also included all samples from the tributaries, attesting to the scarcity of stoneflies in the tributaries. The third cluster indicated perfect similarity of samples 7AR, 7BR and 10CL. These samples contained only *Perlesta placida,* which was emerging at the time of sampling and was thus very much involved in stream drift. *P. placida* was not a permanent inhabitant of these stations, particularly the chronically stressed Substation 7R. In general, the small numbers of Plecoptera in the summer months and their proclivity to drifting led to indistinct clusters. As was noted previously, the snail population was slow to recover from the 1967 fly-ash spill so that recovery was by no means complete by the summer of 1970.

RESULTS AND DISCUSSION — 1971 SURVEY

Total aquatic insect fauna.—Interpretation of the dendrogram relating the 53 samples collected during the summer of 1971 (Fig. 6B) will be facilitated by reference to data on stream discharge (Crossman *et al.*, 1973). For our purposes, information on times of flooding will suffice (Table 1). Two gauging stations were considered: Cleveland, Va., the farthest upstream gauging station, and Speers Ferry, the farthest downstream gauging station in Virginia. Before samples were collected in June, flood conditions—that is, discharges greater than 0.5×10^5 liters per sec—occurred on 7-10 and 13-16 May at Cleveland and 7-11 and 13-16 May at Speers Ferry.

Previous investigators have noted the detrimental effects of rapid changes in discharge on the structure of macrobenthic communities. Patrick (1970) pointed out that populations of all species are not affected equally: some species may be eliminated entirely, while others are able to adapt to, avoid or otherwise withstand high-velocity conditions. Viewed in this way, flooding provides a selective, short-term stress which reduces diversity and population size, with each species being affected in proportion to its resistance to flooding. Depending on the rate of recovery, the effects of flooding may be evident for a varying length of time. Among the most important mechanisms for biological recovery is stream drift, which is also highly dependent on variations in discharge (Minshall and Winger, 1968). In certain situations drift may be regarded as an equalizing process, since organisms that do not normally inhabit an area may become temporary residents even if the habitat is unsuitable for permanent residence. Therefore, it is possible that one major effect of flooding is to equalize benthic samples from habitats that may be ecologically dissimilar before flooding. We believe our data show that one of the effects of flooding and stream drift may be to cause benthic organisms to be distributed more uniformly throughout a stream, thereby, resulting in greater similarity among samples.

The equalizing effect of flooding is a possible explanation for the composition of clusters in the dendrogram of Figure 6B. If a level of similarity just higher than 0.435 is chosen, five clusters are formed. Cluster A comprises samples from Stations 1, 2, 3, 4, 7 and 8. Cluster B includes samples from Stations 7-15 in the stream. Cluster C is smaller and includes samples from Stations 13, 16, 17 and 20. The remaining two clusters are very small and are closely linked to the third cluster. If they are added to the third cluster, disregarding the limit of similarity, the third cluster contains samples from Stations 13, 15, 16, 17, 20 and 21.

In the analyses of the 1969 and 1970 surveys, the mainstream stations were repeatedly divided into three groups: Stations 1-4 above the power plant; Stations 7-13 below the power plant, parts of which were affected by the 1970 acid spill; and Stations 15-21 in the area where the gradient of the stream was lowest. The samples from the 1971 survey did not fit this pattern as well as did those from previous years. For example, samples from Stations 1-4 were clustered with samples from Stations 7 and 8. The second cluster includes not only samples from Stations 7-13 but also samples from Stations 14 and 15. In the third cluster, one sample from Station 13 has been added to Stations 15-21. In effect, the boundaries previously noted between groups of stations seem to have been moved downstream.

In addition to this downstream shift of similarity, the inclusion of sample 7R in the first cluster and 14MC in the second cluster requires some explanation. In previous surveys Substation 7R was noted to be sparsely inhabited because of effluents from the power station and an unsuitable substrate. Substation 14MC was similarly sparsely inhabited because it was located in a nutrient-enriched, fifth-order tributary of the Clinch River where the substrate was unsuitable for most benthic organisms because of coal-washings. A possible explanation for these differences from previous patterns is that the communities at both healthy and unhealthy stations have been equalized by flooding and stream drift. This hypothesis is supported by data which show Substation 7R to have greater generic richness after the flooding in 1971 than at any other time. Moreover, the normally healthy Substations 1R, 4R and 13R had lower generic richness after flooding than previously. The number of genera found at Substation 8R after flooding in 1971 was as low as the number found in 1970 immediately after the acid spill. It seems, then, that substations with rich faunas were depressed by flooding, whereas sparsely inhabited substations were less likely to lose species by flooding but instead gained species because of increased drift associated with high water conditions. It is clear that periodic flooding of the Clinch River has a pronounced effect on the macrobenthic insect fauna and that at the time of the 1971 survey the insects had not had enough time to recover from the adverse effects of flooding.

Redundancy in data from stream surveys.—Kaesler and Cairns (1972), on the basis of study of data from many elements of the biota of a portion of the upper Potomac River, concluded that a great deal of redundancy resided in their data. They suggested that the study of only aquatic insects and diatoms may yield most of the ecological information from limnological surveys at greatly reduced cost and effort. The fact that cluster analyses of data from several groups of organisms from the macrobenthic community tell much the same story suggests that appreciable redundancy may reside within data from the macrobenthic community as well. We plan to investigate this hypothesis in a later study.

Conclusions

From cluster analysis of these data, supported by previous studies of density and diversity, it is possible to draw the following conclusions:

(1) By 1969, the macrobenthic invertebrate community had recovered in large part from the effects of the 1967 spill of material from the fly-ash retaining pond. This was especially true of the insects but not for the molluscs, which recovered very slowly.

(2) In 1969, in spite of being well on the way to recovery, the insect fauna of the affected area was still dissimilar to other areas not

affected by the spill. It was not possible to separate the acute effects of the spill from the chronic effects of the day-to-day operation of the power station.

(3) Cluster analysis of data on the total insect fauna from the 1970 survey has revealed that the acid spill had deleterious effects on Stations 8, 9 and 10. As was shown by analysis of density and diversity, Substation 7R was also affected, but chronic stress of the community at that locality precluded the detection of acute stress by means of cluster analysis. Recovery of the total macrobenthic insect community was well under way by the time the survey ended, but recovery of other elements of the fauna, especially the gastropods, had not progressed as far.

(4) The Ephemeroptera, which had been completely eliminated by the acid spill, had nearly recovered by the end of the survey at all of the stations affected. Several taxa of gastropods, on the other hand, remained drastically depleted, suggesting characteristically slow recovery.

(5) Cluster analysis was particularly useful in detecting the effects of type of substrate, time of sampling, longitudinal succession, and flooding on the macrobenthic community. The 1971 survey indicated that flooding may have a "homogenizing" effect on a stream's macrobenthic community (*i.e.*, make the distribution more homogeneous). For the Clinch River, at least, it appears that stations normally rich in taxa were more likely to lose representatives of some taxa than sparsely inhabited stations. The sparsely inhabited stations, on the other hand, seemed more likely to gain taxa due to increased stream drift associated with higher rates of flow. Thus *at some stations* the net effect of flooding can be as great as the net effect of an industrial spill.

Acknowledgments.—The authors wish to express their appreciation to Dr. T. F. H. Allen, who read an early version of our manuscript and contributed substantially to its improvement. We are also indebted to Mr. T. A. Miskimen, American Electric Power, and to Messrs. T. A. Abolin, R. E. Senter and R. G. McComas, Appalachian Power Company, for their cooperation and assistance in all phases of this project.

The financial support for the cluster analysis was provided by Biomedical Sciences Support Grant 4869-5706-6 from The University of Kansas and by Grant A-054-Kan from the Office of Water Resources Research through the Water Resources Institute of The University of Kansas. Principal support for other aspects of the research was from the Appalachian Power Company and the Department of Biology, Virginia Polytechnic Institute and State University.

Particular thanks are due to Dr. Kenneth L. Dickson who helped initiate the study and to the following persons who helped collect specimens, process samples, or type the paper: Mike Masnik, Jerry Smrchek, Zane Grey Conner, Donald Ball, David DeHuff, Terry Lawrence, Gary Webb, Gary E. Wheeler, Russell Donlan, Tom Muir, Mrs. Mary Look, Greg Lewis, Mrs. Brenda Armstrong, and Mrs. Jerilee Crossman.

REFERENCES

ALLEN, T. F. H. AND J. F. KOONCE. In press. Multivariate approaches to algal strategies and tactics in the systems analysis of phytoplankton. *Ecol. Monogr.*

ANONYMOUS. 1967a. Clinch River fish kill, June 1967. Federal Water Pollution Administration, Middle Atlantic Region, United States Department of Interior. 1-20 p.

———. 1967b. Fish-kill on Clinch River below the steam-electric power plant of Appalachian Power Company, Carbo, Virginia, June 10-14, 1967. Tennessee Valley Authority. 29 p.

CAIRNS, J. JR., J. S. CROSSMAN AND K. L. DICKSON. 1972. The biological recovery of the Clinch River following a fly ash pond spill. *Proc. 25th Ind. Waste Conf. Purdue Univ.*, p. 182-192.

———, ———, ——— AND E. E. HERRICKS. 1971. The recovery of damaged stream. *Ass. Southeast. Biol. Bull.,* **18**:72-106.

CROSSMAN, J. S., J. CAIRNS, JR. AND R. L. KAESLER. 1973. Aquatic invertebrate recovery in the Clinch River following pollutional stress. *Bull. Water Resour. Res. Center,* Virginia Polytechnic Institute and State Univ.

JACCARD, P. 1908. Nouvelles recherches sur la distribution florale. *Soc. Vaudoise Sci. Natur. Bull.,* **44**:233-270.

KAESLER, R. L. AND J. CAIRNS, JR. 1972. Cluster analysis of data from limnological surveys of the upper Potomac River. *Amer. Midl. Natur.,* **88**: 56-67.

MINSHALL, G. W. AND P. V. WINGER. 1968. The effect of reduction in stream flow on invertebrate drift. *Ecology,* **49**:580-582.

PATRICK, R. 1970. Benthic stream communities. *Amer. Sci.,* **5**:546-549.

SOKAL, R. R. AND P. H. A. SNEATH. 1963. Principles of numerical taxonomy. Freeman and Co., San Francisco. 359 p.

WILLIAMS, W. T. 1971. Principles of clustering. *Annu. Rev. Ecol. System.,* **2**: 303-326.

SUBMITTED 23 APRIL 1973 ACCEPTED 8 NOVEMBER 1973