
6 Analysis of Ephemeroptera from Hester-Dendy samplers: A bootstrap rarefaction method

Michael D. Hubbard¹ and Kerrie M. Swadling²

¹Entomology/Center for Water Quality, Florida A&M University, Tallahassee, FL 32307-4100 USA and ²Department of Biological Science, Florida State University, Tallahassee, FL USA (Present address: Department of Zoology, University of Tasmania, Hobart, Tasmania, Australia)

A bootstrap rarefaction method is presented to allow testing of faunal differences between mayfly samples collected by artificial samplers. A complete randomization procedure is used to determine the number of species present on all possible combinations of samplers. A method to determine 95 per cent confidence limits is given to allow differentiation between populations. Simulation experiments illustrate the method. Temporal variation in mayfly populations in Burnt Mill Creek in North Florida is demonstrated by this method using results from modified Hester-Dendy samplers.

Introduction

A number of the aquatic macroinvertebrates, in particular many of the Ephemeroptera (Hubbard and Peters 1978), are generally thought to be extremely sensitive to various types of pollution and are commonly used as bioindicators of environmental conditions. Therefore, effective methods for estimating the population parameters of the macroinvertebrate fauna (in particular, the composition of the fauna and the density of each taxon) are often crucial in biological monitoring of water quality.

The use of artificial substrates has several advantages over standard collecting methods for sampling freshwater benthic macroinvertebrates: (i) they are useful when collecting data from habitats difficult to sample by other means; (ii) they permit non-destructive sampling of the environment; (iii) they often are quite cost-effective; and (iv) they permit flexibility in and standardization of sampling programs (Mason et al. 1973; Rosenberg and Resh 1982).

The use of modified Hester-Dendy samplers for monitoring and sampling freshwater macroinvertebrates is one of the more common methods used in the

North Florida area by researchers, including governmental, regulatory, university and private researchers (e.g., Florida Department of Environmental Regulation 1985, 1987; Mason 1991; Pescador and Rasmussen this volume). The number of samplers used and the method of their deployment varies among studies because no standard number or method has been agreed on among the various research efforts. Unfortunately, the efficacy of this method to determine what taxa are present and their population parameters is little known and has not been investigated for North Florida streams. It is therefore important to determine the accuracy and repeatability of this sampling method.

Hester-Dendy samplers mimic snags (logs, limbs, twigs and other woody inclusions), a frequently occurring and important habitat for benthic invertebrates in the sandy-bottomed streams of the southeastern coastal plain. Because of the unstable shifting sand bottoms, and lack of cobble and gravel, in these streams, snags are often one of the only stable substrates for colonization by macroinvertebrates (Benke et al. 1984). The Hester-Dendy samplers do not sample the entire mayfly fauna, only that suite of species that inhabits snags. The use of these artificial samplers allows comparable standardized sampling of this important suite of species between locations with different substrate types and permits comparison of results between the differing habitats, an important consideration in water quality studies (Rosenberg and Resh 1982).

Variability of abundance is much greater than variability of number of taxa (Minshall and Petersen 1985; Clements 1991), which suggests that comparison of numbers of species might be much more useful than comparison of abundances for biomonitoring benthic macroinvertebrates. Therefore, we considered species richness, rather than individuals. Traditional methods (e.g., Elliott 1977; Downing 1979; Morin 1985) for determining the number of samples required to estimate reliably faunal numbers consider numbers of individuals present in the samples, not just taxa present. This creates complications when clumping of individuals is common, as it is in the use of Hester-Dendy samplers in North Florida streams. This problem is discussed in greater detail below in regard to artificial samplers. The method proposed here overcomes some of these problems in regard to clumping of individuals.

The aim of our study was to develop and illustrate a new method to estimate reliably the taxonomic species richness of the mayfly fauna collected with modified Hester-Dendy samplers. In this paper we present a simple bootstrap method designed to accomplish this task. The number of taxa present on all possible combinations of replicate samplers is determined. The number of samplers in each combination is plotted against the mean number of species in each combination. This plot is then bracketed by a (95 per cent) confidence interval. This method allows comparison of separate samples to determine whether differences between them are statistically significant. We also present here results of simulation experiments designed to illustrate the method and results of field samples to show the method in actual practice.

Table 1. Values of parameters for simulation experiments.

Simulation Number	Number of Pseudospecies	No. of individuals in each pseudospecies
1, 2, 3	4	20, 15, 10, 5
4, 5, 6	4	15, 12, 9, 6
7, 8, 9	12	25, 24, 22, 21, 19, 18, 16, 15, 13, 12, 5, 4
10, 11, 12	3	15, 10, 5

Materials and Methods

Simulations

Simulations, mimicking collection of species by artificial samplers, were conducted to illustrate a bootstrap method to estimate reliably the taxonomic species richness of the mayfly fauna using artificial samplers. Individuals of pseudospecies were randomly assigned to each of 20 samplers to simulate the data from a 20-sampler collection (n_i individuals of each of s_j species were assigned at random to $x=20$ samplers). The values of the simulation population parameters are listed in Table 1. Although the actual simulation parameters chosen to illustrate this bootstrap method do not purport to represent natural population parameters (these parameters are undetermined for North Florida streams) we think that they are not unreasonable. They appear to be in the range found for actual collections. Each set of parameters was used in three simulations, resulting in three sets of data for each population. Twelve simulations in total were performed.

For each experiment, after pseudospecies were assigned, each possible combination of the 20 samplers taken x at a time, where $x = 1$ to 20, was determined, and the number of species present in each combination was ascertained.

Because the number of combinations of samplers can be quite large (Table 2) a computer program was written in Microsoft QBasic and run on an MS-DOS personal computer to perform the simulations and calculate the combinations. There are 190 combinations of 20 samplers taken two at a time, 1140 combinations

Table 2. Number of combinations of 20 samplers taken x at a time.

Value of x	No. of Combin.	Value of x	No. of Combin.	Value of x	No. of Combin.
1	20	8	125,970	15	15,540
2	190	9	167,960	16	4845
3	1140	10	184,756	17	1140
4	4845	11	167,960	18	190
5	15,540	12	125,970	19	20
6	38,760	13	77,520	20	1
7	77,520	14	38,760		

taken three at a time, up to 184,756 combinations taken ten at a time. The number of combinations then decreases to the trivial one combination of 20 taken 20 at a time.

Field Collections

The field portion of the study was conducted in Burnt Mill Creek, Jefferson County, Florida. Burnt Mill Creek is a typical small, sandy-bottomed, blackwater North Florida stream with well defined banks. At the experimental site it is 3-4 m wide and 0.5 m to 1.5 m deep. The pH typically ranges from 5.5 to 6.4 and the dissolved oxygen from 7.3 to 9.6 mg L⁻¹. The heavy forest overstory (especially *Pinus* spp., *Quercus* spp. and *Magnolia* spp.) results in a profusion of snags in the creek. Five species of mayflies are known from this area of Burnt Mill Creek: *Acerpenna pygmaea* (Hagen), *Eurylophella temporalis* (McDunnough), *Habrophlebiodes brunneipennis* Berner, *Paraleptophlebia volitans* (McDunnough), and *Stenonema smithae* Traver.

The Hester-Dendy samplers were modified from the original design (Hester and Dendy 1962). They were constructed of 12 fibreboard discs (6.4 mm thick x 70 mm diameter, separated by eight 6.4 mm and three 12.8 mm x 20 mm spacers of the same material; Fig. 1). The total surface area of each sampler available for

colonization by the aquatic fauna is approximately 0.09 m². Concrete blocks that held four samplers each with screw anchors were used to hold the samplers in place on the stream bottom.

During each collection period, five anchoring blocks were placed on the stream bottom for colonization by the aquatic macroinvertebrate fauna. The 20 samplers all were placed parallel to the direction of water flow, thus avoiding possible effects of orientation in relation to water flow (Hill and Matter 1991). The anchoring blocks were placed contiguously, suggesting that placement effects in the habitat were nonexistent or minimal.

The samplers were collected after one month. Preliminary experiments by us and published reports (Tsui and Breedlove 1972; Rosenberg and Resh 1982) indicate that this is sufficient time for the fauna to reach a colonization "equilibrium" (MacArthur and Wilson 1967) on the samplers. Each sampler was removed from the anchoring block and placed in a separate polyethylene bag half-filled with stream water. The bags were stored on ice and returned immediately to the laboratory. The samplers were carefully disassembled, then the macroinvertebrates were removed, screened through a 0.71 mm sieve and preserved in ethanol. The mayflies were separated from the rest of the macroinvertebrates and identified to species.

Three trials of these field collections were conducted (January-February, May-June and June-July 1992). The third trial was subjected to a strong spate, which buried two of the anchoring blocks (containing eight of the Hester-Dendy samplers) in the sand. Therefore, only 12 samplers were recovered and treated in the analysis. The data were analyzed in the same way as the simulations.

Results

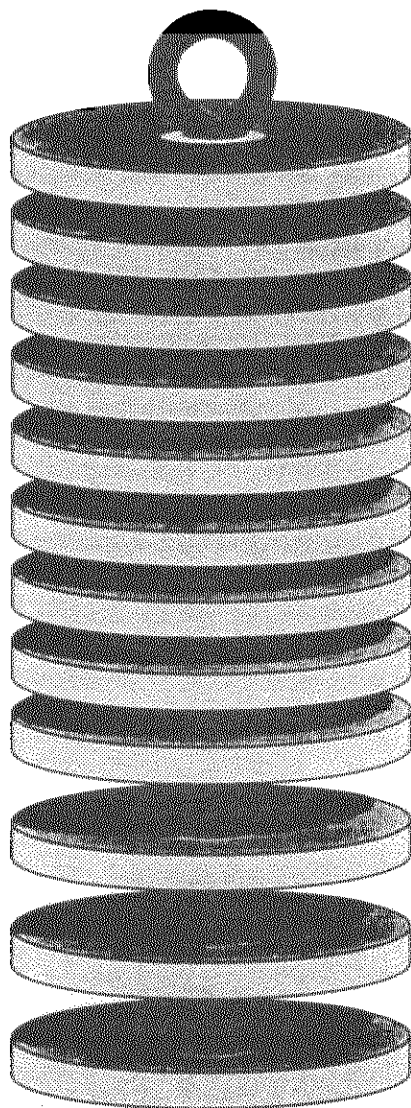
Assumptions

There are several assumptions underlying this analysis. It may be possible to relax these to varying degrees, with appropriate caution, when interpreting the results.

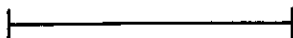
One assumption is that placement effects are negligible. The exact placement of the modified Hester-Dendy samplers in the stream should make little difference in the fauna collected, because the entire suite of snag fauna mayfly species appears to be quite vagile. Such vagility is required for colonizing the scattered and often temporary snag habitats.

A second assumption is that the results are inter-sampler independent: i.e., the presence of more than one sampler in a location does not affect the result obtained by the other samplers.

Figure 1. **Diagram of a modified Hester-Dendy sampler: full description in text.**



5 cm



We have also assumed that the presence of one organism on the samplers should not affect the presence of other organisms. This is not strictly true, because intraspecific clumped distribution is common and it should have little practical effect. The presence of individuals on a sampler can be compared to an equilibrium process, considering the samplers as substrate islands in the stream (c.f., MacArthur and Wilson 1967; Hubbard 1973; Boothroyd and Dickie 1991). Individuals encounter samplers at a given frequency, they remain on an encountered sampler at a given frequency, and then they leave a sampler at a given frequency. These frequencies are obviously species-specific. The number of individuals on any sampler is a function of the interaction of these frequencies. In the case of clumped distribution, the frequency of leaving is lessened by the presence of conspecifics. This should have little effect on the frequency of arrival and, of course, the presence of multiple individuals of the same species obviously does not, in itself, change the number of taxa on the sampler.

We also have assumed that, in the field collections, the total of 20 samplers (12 in collection 3) collected all of the mayfly species present in the suite of snag fauna species at that time. This assumption appears to be both valid, except for the possibility of extremely rare species, and of little practical import in these analyses.

Because of various problems encountered in preliminary collections, in particular the common occurrence of clumped distribution in some species, which violates the assumption that organisms should not affect the presence of others (see discussion above), numbers of individuals were not considered in this study.

Simulations

Curves of the mean number of species for all possible combinations of each number of samplers for each of the 12 simulation experiments are presented in Figs. 2-3. These curves represent the limit that would be approached by infinite resampling with replacement for the different combinations of samplers.

This method, a randomization test, belongs to the general category of Monte Carlo tests (Sokal and Rohlf 1981) although it is more accurate than traditional Monte Carlo tests because it results in the limit (i.e., the actual number) rather than in an approximation. This is somewhat analogous to rarefaction methods, outlined and discussed by Simberloff (1978), although obviously somewhat different from those traditional rarefaction methods that deal with hierarchical taxa.

Unfortunately, this method shares a drawback with most rarefaction techniques (and indeed, with all species richness methods). It is a fairly weak statistical test because it does not distinguish between different sets of taxa. It is entirely possible that two sets of taxa are indistinguishable by this method even though they have no species in common.

Figure 2. Means (\pm 95% C.I.) for simulations 1 to 6, see Table 1 for simulation parameters.

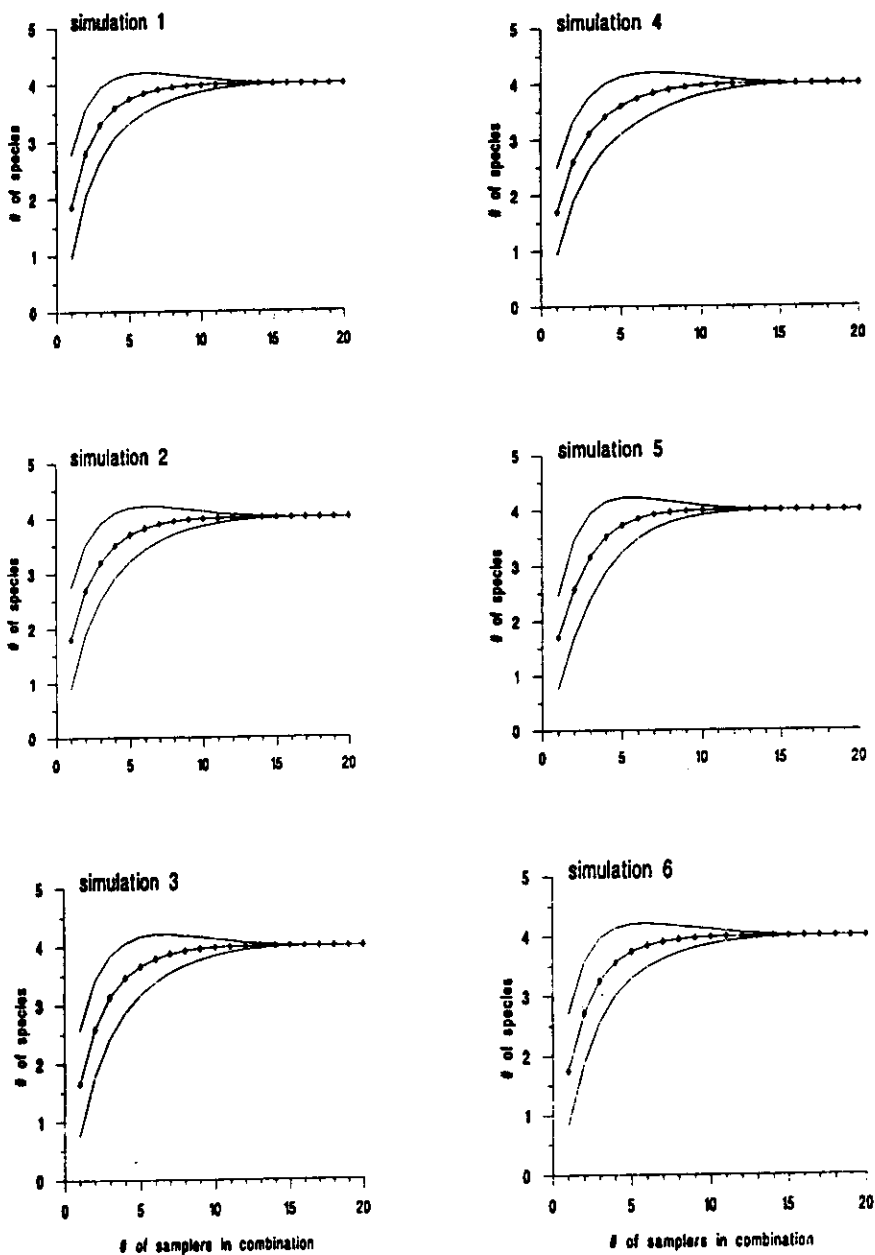
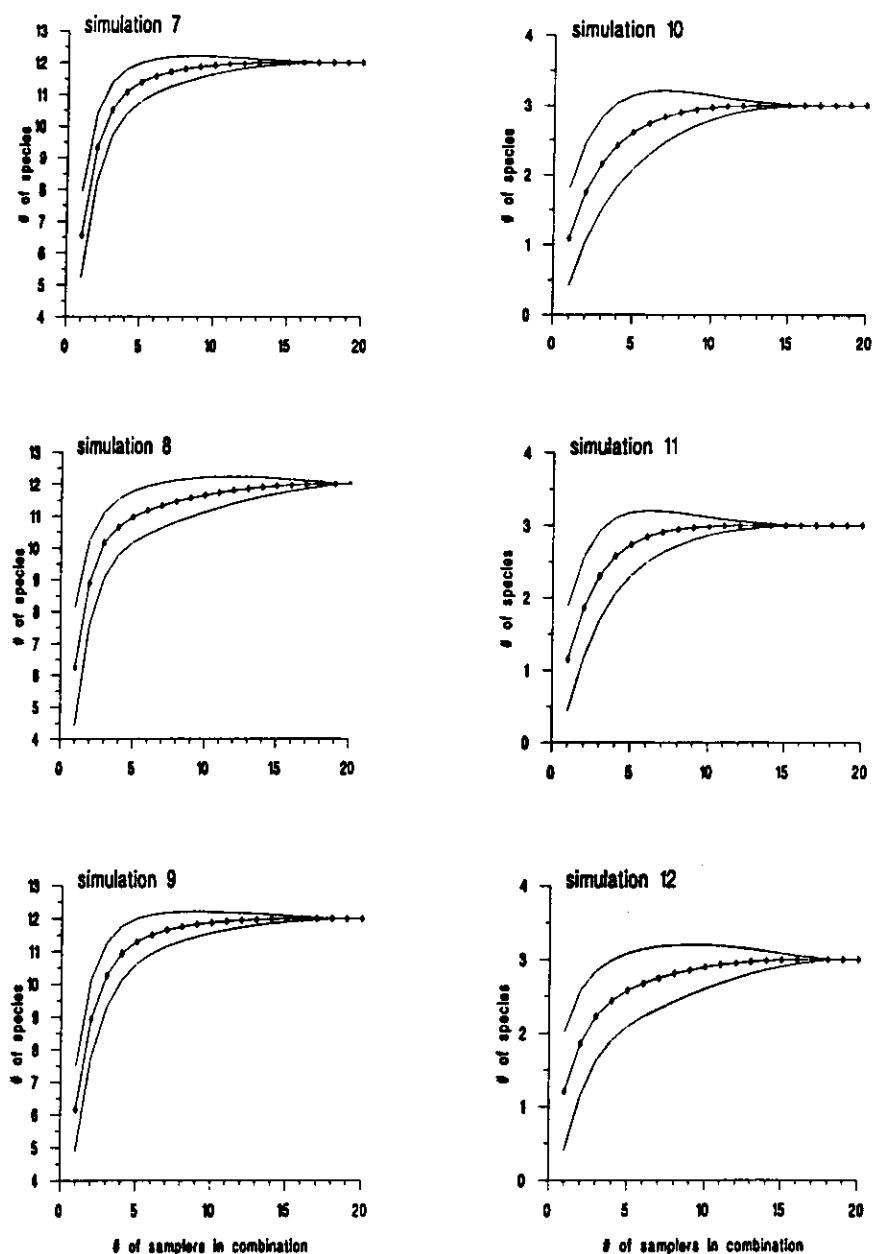


Figure 3. Means ($\pm 95\%$ C.I.) for simulations 7 to 12, see Table 1 for simulation parameters.



The mean species numbers curves are bracketed by a band 1.96 standard deviations wide, which can be taken to represent approximate 95 per cent confidence limits for the mean number of species for each set of combinations. In other words, for any two sets of samples, if the mean for one set of samples falls outside the 95 per cent confidence limits for the other set of samples, the hypothesis that the samples were taken from the same population can be rejected with a probability of $P < 0.05$. A 1.96 standard deviation band was used to approximate 95 per cent confidence levels because calculation of exact confidence levels is quite tedious and yields results very close to the approximation (Simberloff 1978).

Each set of simulations with the same population parameters yielded results consistent with the fact that they were drawn from the same population. That is, the means of each were within the 95 per cent confidence limits of the others, and there was no reason to believe that they were from different populations. In fact, the replicates within each simulation experiment are virtually indistinguishable.

Area graphs showing the frequency of occurrence of species numbers for each of the combinations of samplers for simulation experiments numbers 1, 4, 7 and 10 (one representative example from each set of three replicates) are given in Fig. 4. These graphs show the frequency at which the different species numbers appeared in each set of combinations of samplers. As is readily apparent from these graphs, frequency of the maximum species numbers increases quite rapidly, much faster than is apparent from inspecting the curves of the means. The maximum species number usually appears as the median by about six to seven samplers per combination.

Empirical evidence from both the simulation and field experiments reveals that the confidence interval curves can be used as a good approximation to derive estimated frequency curves by extending the confidence intervals to include the upper and lower species numbers, which are approached by the confidence intervals. For example, if the confidence intervals go from 2.3 to 4.6 for a particular number of samplers, then approximately 95 per cent of the sampler combinations of that number will yield from two to five species.

Field Collections

The data from the field collections in Burnt Mill Creek were treated in the same manner as the data from the simulation experiments; the number of species on all possible combinations of samplers was calculated. Four species of mayflies were collected in the January-February and the June-July collections. Three species were collected in the May-June collection.

Because the samples were taken from the same location at different times, it was possible to use the results of the field experiments to test the hypothesis that the population parameters of the snag fauna in Burnt Mill Creek remained unchanged temporally throughout the three sampling periods.

Figure 4. Frequency of occurrence of species numbers for combinations of samples in simulations 1, 4, 7 and 10: numbers on graphs refer to number of species.

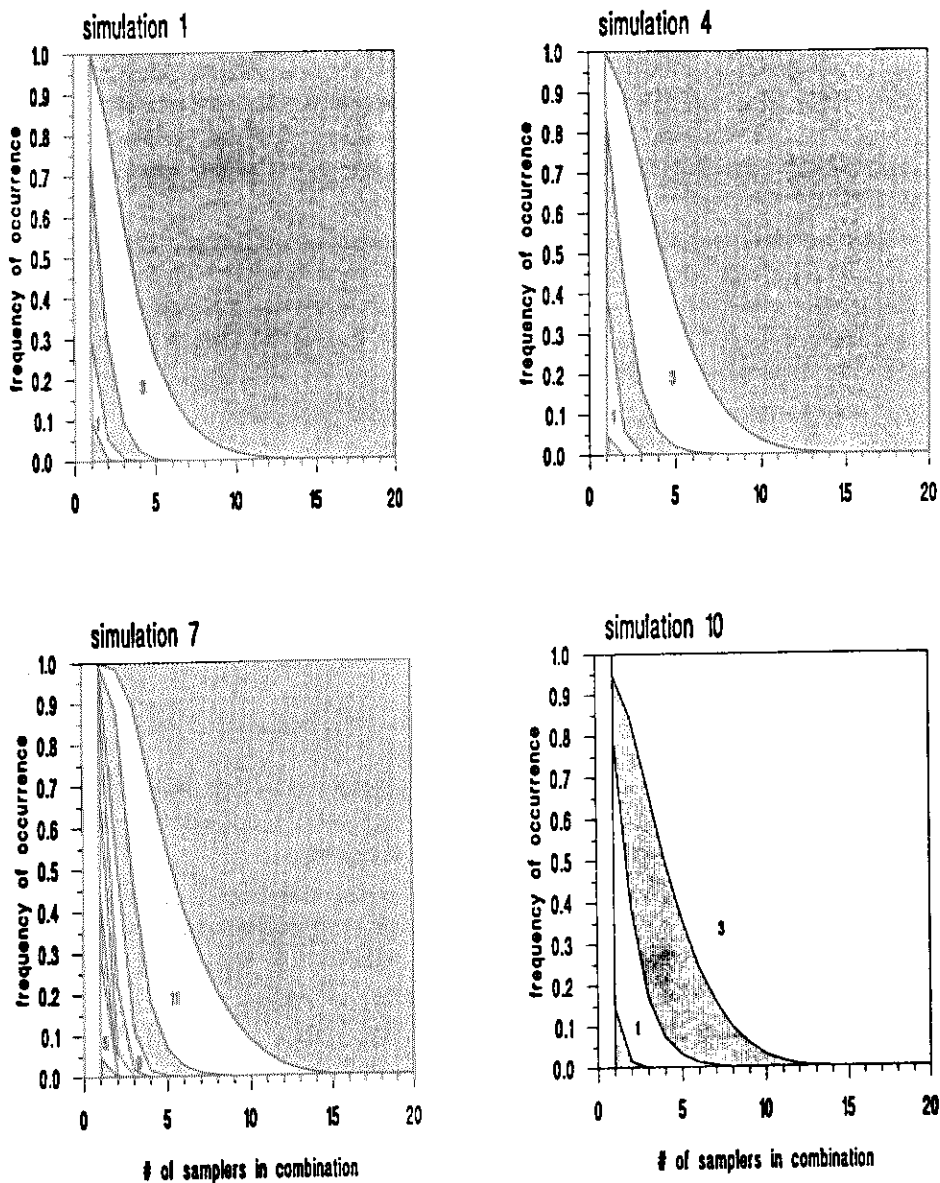


Fig. 5 presents a comparison of the means (± 95 per cent confidence intervals) of numbers of species on each of the combinations for the field data from Burnt Mill Creek. Comparisons of the January-February vs. May-June and May-June vs. June-July sampling periods are shown.

As can be seen clearly by inspection of Fig. 5, the means of species number for each combination of samplers fall outside the 95 per cent confidence limits for the other collecting dates. The means of the January-February and the May-June collection combinations do not overlap with the 95 per cent confidence intervals of each other at all. The means of the May-June and June-July collection combinations separate from the confidence intervals of the distributions of each other at about the eight sampler combination level.

Clearly the hypothesis that the mayfly population parameters remained unchanged over time between the sampling periods can be rejected with $P < 0.05$ even at combinations of small numbers of samplers. Thus we can conclude that the faunas were different during each of the sampling periods and the population parameters changed over time.

If the total number of samplers included in the randomization test had been smaller, it is quite likely that the probable higher variances would have masked these differences at some level.

One particular phenomenon illustrated by this study is the concept that if only one sampler is deployed the concept of variance is meaningless and it is impossible to tell if the population parameters of the sampling site or time differ from any other sampling site or date. With no concept of sample variation, no conclusion at all can be drawn.

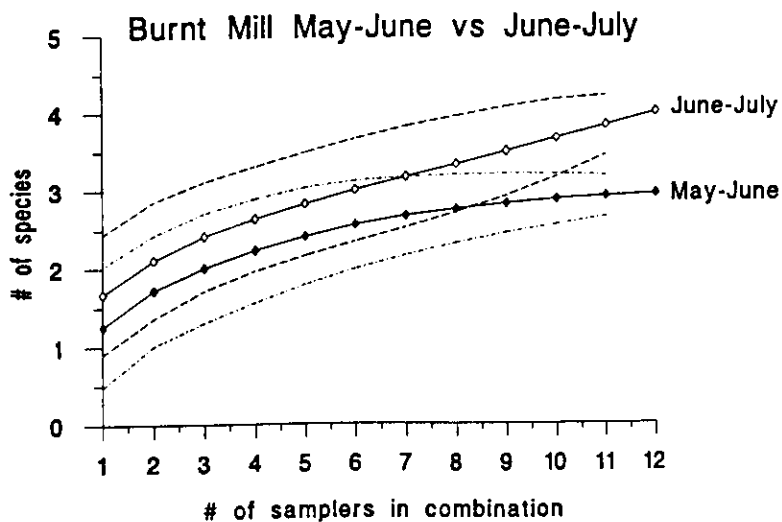
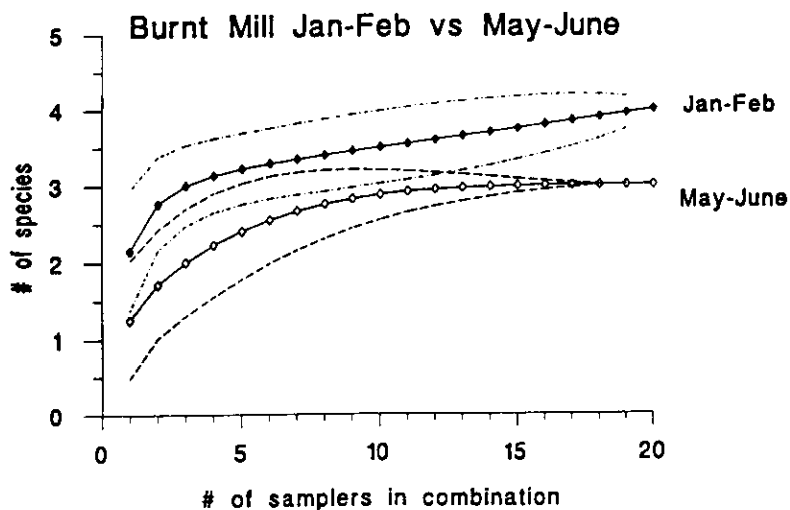
Discussion

It is obvious from an examination of the simulation results that at least four factors are extremely important in determining the number of species collected by these artificial samplers and for comparing results of collections.

The first factor is the number of species present in the snag fauna species pool. The more species present in the pool, the easier it is to discriminate between samples. The fewer species present in the pool, the faster the expected number approaches an asymptote.

The second factor is the density of the individuals of the species present. Because sampling variability is inversely related to density in benthic macroinvertebrate communities (Downing 1979; Morin 1985), an increase in density reduces the variance and allows use of fewer samplers to discriminate between samples.

Figure 5. Comparison of means (\pm 95% C.I.) for samples collected from Burnt Mill Creek.



The third factor is the number of samplers deployed. The larger the number of samplers, the closer the expected number approaches the true number in the species pool. A larger number of samplers also reduces the variances and allows greater discrimination between samples.

A fourth factor is the interaction between the frequency at which individuals of each species encounter the samplers and the frequency at which they may depart instead of remaining on the samplers (see discussion above).

In North Florida streams, where the suite of snag fauna mayfly species is quite low compared with many areas, the accuracy of modified Hester-Dendy samplers in determining the number of taxa is also somewhat low. In those areas with a much larger suite of species, the accuracy should increase correspondingly.

Because one of the reasons for using artificial samplers is to reduce the amount of sampling effort (e.g., cost and time), it is obvious that the lowest number of samplers needed to produce an acceptable result is optimum. With a suite of at least four species, it appears that for many purposes four samplers might be an adequate number for many monitoring and sampling programs. The curves of mean species number begin to level towards an asymptote at about four samplers. Each investigator must decide on an acceptable compromise between accuracy and cost.

One caveat in the use of this method is that less reliance should be placed on the ends of the curves where x_i (the number of combinations) is small. The test is much more robust in the middle areas of the curves where x is greater.

Among the principal reasons for monitoring the fauna of aquatic habitats is to determine if the population parameters are changing or to determine if two different habitats differ in these parameters. Causes for the changes or differences can then be profitably investigated. There is little value in striving to determine causes for effects that cannot be demonstrated.

We feel that with the application of this method it will be possible to show, in many instances, whether two sampling locations have different population parameters. In that case, further investigation of the causes of the difference might be profitable. If the hypothesis that the two sets of samples were drawn from the same set of population parameters cannot be rejected, one cannot know that the sampled habitats are faunistically different.

There is little reason that this method could not be generalized to other sampling methods. It should apply equally well to other types of artificial samplers as well as the traditional collecting methods such as kick sampling, pond nets, Surber samplers and Ponar grabs.

Although this method does require extensive computation, it appears to be the first method that allows statistically valid comparison of the results of artificial sampling collections at the taxon level, with the obvious benefits that accrue.

Acknowledgements

This study grew out of the Florida A&M University Water Quality / Entomology Research Program, and the assistance of the members of our research team (M. L. Pescador, J. Jones, R.W. Flowers and A.K. Rasmussen) is gratefully acknowledged. Mary Baggett and Duane Meeter of the Florida State University Statistical Consulting Center were most helpful in discussion of the problem. Comments by Jan Ciborowski of the University of Windsor improved an earlier draft of this paper. Jane Jimeian of the National High Magnetic Field Laboratory at Florida State University helped with the figures. Partial support for this research was furnished by a grant from the U. S. Department of Agriculture.

References

- Benke, A.C., T.C. Van Arsdall, Jr., D.M. Gillespie and F.K. Parrish. 1984. Invertebrate productivity in a subtropical blackwater river: the importance of habitat and life history. *Ecol. Monogr.* 54: 25-63.
- Boothroyd, I.K.G. and B.N. Dickie. 1991. Macroinvertebrate drift and community colonisation on Perspex artificial substrates in the Ohinemuri River, New Zealand. *N.Z. J. Mar. Freshwat. Res.* 25: 167-176.
- Clements, W.H. 1991. Characterization of stream benthic communities using substrate-filled trays: colonization, variability, and sampling selectivity. *J. Freshwat. Ecol.* 6: 209-221.
- Downing, J.A. 1979. Aggregation, transformation, and the design of benthos sampling programs. *J. Fish. Res. Bd. Can.* 36: 1454-1463.
- Elliott, J.M. 1977. Some methods for the statistical analysis of samples of benthic invertebrates. 2nd Ed. *Spec. Publ. Freshwat. Biol. Assoc.* No. 25.
- Florida Department of Environmental Regulation. 1985. *Limnology of the Suwannee River, Florida*. Div. Environ. Prog., Biol. Sect., Tallahassee.
- Florida Department of Environmental Regulation. 1987. *An Investigation of the Water Quality of the Ochlockonee River*. Div. Environ. Prog., Biol. Sect., Tallahassee.
- Hester, F.E. and J.S. Dendy. 1962. A multiple-plate sampler for aquatic macroinvertebrates. *Trans. Am. Fish. Soc.* 91: 420-421.
- Hill, J.P. and W.J. Matter. 1991. Macroinvertebrate colonization of Hester-Dendy samplers in different orientations to water flow. *Calif. Fish Game* 77: 94-97.
- Hubbard, M.D. 1973. Experimental insular biogeography: ponds as islands. *Florida Sci.* 36: 132-141.
- Hubbard, M.D. and W.L. Peters. 1978. *Environmental Requirements and Pollution Tolerance of Ephemeroptera*. Cincinnati: U.S. Environmental Protection Agency (EPA-600/4-78-061).

- MacArthur, R.H. and E.O. Wilson. 1967. *The Theory of Island Biogeography*. Princeton: Princeton University Press.
- Mason, W.T., Jr. 1991. A survey of benthic invertebrates in the Suwannee River, Florida. *Environ. Monit. Assess.* 16: 163-187.
- Mason, W.T., Jr., C.I. Weber, P.A. Lewis and E.C. Julian. 1973. Factors affecting the performance of basket and multiplate samplers. *Freshwat. Biol.* 3: 409-436.
- Minshall, G.W. and R.C. Petersen, Jr. 1985. Towards a theory of macroinvertebrate community structure in stream ecosystems. *Arch. Hydrobiol.* 104: 49-76.
- Morin, A. 1985. Variability of density estimates and the optimization of sampling programs for stream benthos. *Can. J. Fish. Aquat. Sci.* 42: 1530-1534.
- Pescador, M.L. and A.K. Rasmussen. This volume. Nymphal abnormalities in *Stenacron interpunctatum* (Ephemeroptera: Heptageniidae) from the Fenholloway River, Florida. In L.D. Corkum and J.J.H. Ciborowski (Eds.), *Current Directions in Research on Ephemeroptera*. Toronto: Canadian Scholars' Press Inc.
- Rosenberg, D.M. and V.H. Resh. 1982. The use of artificial substrates in the study of freshwater benthic macroinvertebrates. P. 175-235 in J. Cairns, Jr. (Ed.), *Artificial Substrates*. Ann Arbor: Ann Arbor Science Publishers Inc.
- Simberloff, D. 1978. Use of rarefaction and related methods in ecology. P. 150-165 in K.L. Dickson, J. Cairns, Jr. and R.J. Livingston (Eds.), *Biological Data in Water Pollution Assessment: Quantitative and Statistical Analyses*. ASTM STP 652, American Society for Testing and Materials.
- Sokal, R.R., and F.J. Rohlf. 1981. *Biometry*. 2nd Ed. New York: W. H. Freeman and Co.
- Tsui, P.T.P. and B.W. Breedlove. 1978. Use of the multiple-plate sampler in biological monitoring of the aquatic environment. *Florida Sci.* 41: 110-116.