The Quantification of Stream Drift

J. David Allan

Department of Zoology, University of Maryland, College Park, MD 20742, USA

Estelle Russek

Department of Animal Sciences, University of Maryland, College Park, MD 20742, USA

Allan, J. D., and E. Russek. 1985. The quantification of stream drift. Can. J. Fish. Aquat. Sci. 42: 210-215.

Although the purpose of many drift studies is to describe quantitatively the abundance of drifting invertebrates and make comparisons between seasons or sites, almost no investigations have employed replicate sampling. We analyzed drift collections from a Rocky Mountain stream in order to investigate the variability of drift sampling. The data were normalized and the variances stabilized for each taxon examined by data transformation. The fourth root transformation was favored for five taxa and the logarithmic transformation for three. Using the 95% confidence limits on 24-h drift density for an abundant mayfly (*Baetis bicaudatus*), we found that six to seven replicates are required to obtain 95% CL \pm 50% of the mean. Drift sampling appears to require fewer replicates than benthic sampling for comparable precision. Investigators may fail to replicate drift samples because they elect to sample frequently over 24 h in order to quantify the diel periodicity of drift. However, when comparison between sites or dates is the principal goal, we recommend that the effort normally put into frequent sampling over 24 h be invested instead in replicated from the remainder of the 24-h period, 60–90% of the variation in the latter was predicted from the single nighttime sample. Thus, little information appears to be lost by this recommended procedure.

Bien qu'un grand nombre d'études sur les organismes dérivants consistent à décrire quantitativement l'abondance des invertébrés qui en dérivent et à faire des comparaisons entre les saisons et les emplacements, presqu'aucun chercheur ne s'est servi de la technique d'échantillonnage répété. Nous avons analysé des organismes dérivants prélevés dans un cours d'eau des Rocheuses afin d'étudier la variabilité de l'échantillonnage de ces organismes. Nous avons normalisé les données et stabilisé les variances obtenues pour chaque taxon examiné en transformant les données. La transformation par la racine quatrième a été choisie pour cinq taxons et la transformation logarithmique pour trois. En utilisant une limite de confiance de 95 % sur une densité d'organismes dérivants calculée pendant 24 h pour une espèce d'éphémère abondante (Baetis bicaudatus), nous avons trouvé que de six à sept échantillons répétés étaient nécessaires pour obtenir une LC de 95 % ± 50 % de la moyenne. L'échantillonnage d'organismes dérivants semble exiger moins d'échantillons répétés que l'échantillonnage d'organismes benthiques pour une précision comparable. Il se peut que les chercheurs omettent de répéter les échantillons d'organismes dérivants parce qu'ils préfèrent échantillonner fréquemment sur une période de 24 h afin de quantifier la périodicité nycthémérale des organismes dérivants. Cependant, lorsque la comparaison entre les emplacements ou les dates est le principal objectif visé, nous recommandons que l'effort fourni normalement pour prélever fréquemment des échantillons sur une période de 24 h soit plutôt consacré à prélever des échantillons répétés juste après la tombée de la nuit, lorsque le nombre d'organismes dérivants est normalement à son maximum. Quand nous avons calculé la régression des organismes dérivants pour le reste de la période de 24 h, nous avons prédit 60 à 90 % de la variation de ces derniers à partir de l'unique échantillon de nuit. Par conséquent, il semble que peu d'informations soient perdues lorsqu'on utilise la procédure recommandée.

Received November 15, 1984 Accepted October 12, 1984 (J7120)

rift, the downstream transport of organisms in running waters, has been of interest to stream biologists since Müller (1954) first reported this phenomenon. Diel periodicity, the potential consequences for invertebrate distribution, possible causes of drift activity, and its quantitative reporting have resulted in many studies (see reviews by Waters 1972; Müller 1974). Drift also provides a measure of invertebrate abundance that may be used to answer questions concerning comparisons from experimental manipulations or environmental impacts. In contrast with the sampling of stream benthos, however, where the issues of appropriate statistical methods, precision, and confidence limits have Reçu le 15 novembre 1984 Accepté le 12 octobre 1984

received extensive discussion (Elliott 1977; Green 1979; Resh 1979), the quantification of drift has received little study. Only Chutter (1975) specifically investigated variability of replicate drift collections, although Elliott's (1970) treatment of sampling methodology includes statistical examples using replicated nets, and Ulfstrand (1968) compared drift collections from several depths and locations within a site. Virtually without exception, however, reports of drift lack statements of the precision of estimate.

Topics that have been addressed for benthic sampling include the need for suitable data transformation (Downing 1979) so that assumptions of parametric statistics are satisfied, how sample size affects confidence limits around the mean (Resh 1979), and recommendations as to the suitability of various statistical procedures (Elliott 1977; Green 1979). Drift methodology has been discussed from the standpoints of collecting apparatus, various ways of calculating drift, and some aspects of statistical analysis (Elliott 1970; Waters 1969). However, data transformation and confidence limits around the mean as a function of sample size have not been discussed previously, and one goal of this paper is to do so in order to provide some insight into the precision of drift sampling.

Inadequate replication in drift studies may be due to the effort usually devoted to frequent sampling over the diel cycle, in order to assess the often dramatic changes in numbers of organisms collected between day and night. If one wishes an absolute estimate of total drift per 24 h, or a detailed description of diel periodicity, it is typical to collect six to eight individual samples over 24 h at 3- to 4-h intervals. In contrast, if drift is measured in order to compare abundances between seasons or sites, it is of primary interest to have replicated estimates of some single measure of drift per site. For this reason, the second goal was to determine whether one can predict the total catch per 24 h from a single sampling interval, thus allowing the effort normally put into frequent sampling over the diel cycle to be put instead into replication of nets.

Methods

Drift was sampled over 4 yr in Cement Creek, Gunnison County, Colorado (see Allan 1975, 1982 for description of area). Nets were placed in the stream at each of three sites about 1 km apart (UC, TR, and LC of Allan 1982) on most sampling dates. On a few occasions only one site (LC) was sampled. The total of 41 collections represented 14 collecting dates ranging from early June to late September. Drift nets had a mouth area of 0.1 m^2 , a length of $\sim 2 \text{ m}$, and a mesh size of 0.3 mm. Nets widened $\sim 0.5 \text{ m}$ below the opening to minimize turbulence at the net mouth (Waters 1969). The bottom of the net always was >6 cm above the substrate, and the top of the net usually extended above the water surface.

Each drift collection consisted of eight individual samples collected at 3-h intervals over 24 h, in order to quantify diel periodicity in drift activity. The duration of the individual samples ranged from 20 min to 3 h; short sampling duration was necessary during high discharge to avoid net clogging.

Samples were preserved in the field with formalin and rose bengal to facilitate sorting. As some samples contained very large numbers of individuals, we used a plankton splitter to subsample where necessary. Subsamples of 50 or 25% were typical, but occasionally only 12.5% of a sample was counted. Inspection of replicate subsamples (unpubl. data) indicated close correspondence. Taxa abundant enough for statistical analysis included the following: *Baetis bicaudatus, Cinygmula* sp., *Epeorus longimanus, Drunella coloradensis*, and *Ephemerella infrequens*, Ephemeroptera; *Zapada haysi*, Plecoptera; *Prosimulium* spp. and the Chironomidae, Diptera. The Chironomidae were represented by an unknown number of species, as was the genus *Prosimulium*.

Stream discharge was estimated from measurement of stream width, and triplicate measures of depth and current (using a Pygmy current meter at approximately mid-depth) at each of three points across the stream. Flow through each net was estimated from triplicate readings within the net mouth at the beginning and end of each sampling period, corrected for percent of net mouth submerged.

Quantitative Expression of Invertebrate Drift

The following equations define the several estimates of drift used here, following the recommendations of Waters (1969) and Elliott (1970). First, the number of invertebrates collected per net per hour was estimated from a direct count of the sample, corrected for any deviation in sample duration or in subsampling. Then,

(1) Sample drift density =
$$\frac{\text{numbers per net-hour}}{\text{m}^3 \text{ filtered per net-hour}} \times 100$$

where m^3 filtered per net hour is estimated from area of net mouth and current at net mouth. Equation (1) provides a single estimate of numbers drifting per 100 m^3 of water filtered, for an individual hourly sample.

(2) 24-h drift rate
$$\frac{3 \times \sum_{i=1}^{8} \text{ numbers per net-hour}}{\text{proportion of flow filtered by net}}$$

where proportion of flow filtered by net = flow through net/river discharge, and there are eight sampling intervals per 24 h. Equation (2) estimates the total number of organisms drifting past the sampling point per 24 h.

(3) 24-h drift density =
$$\frac{24$$
-h drift rate
(24 HDD) $\frac{24$ -h drift rate \times 100.

Equation (3) estimates the total 24-h drift per 100 m^3 of water filtered. Waters (1972) has recommended that drift be measured as total quantity of organisms drifting past a point per 24 h, divided by total discharge as a measure of stream size. This has the units numbers $\cdot 24 \text{ h}^{-1}$ per $\text{m}^3 \cdot \text{s}^{-1}$. This differs from equation (3) only by a constant, as Waters' term divided by 864 equals 24HDD. The latter may be preferable because the units of expression are simpler.

Results

Transformation of Estimates of Drift

Many conventional statistical analyses require some transformation of the original data in order to meet the assumptions that the observations are normally distributed with sample variance independent of sample mean (Green 1979). To determine if these drift data required transformation, and which transformations were satisfactory, we used a maximum likelihood method developed by Box and Cox (1964) for locating the optimal transformation. The optimal transformation is given by

(4)
$$y = (x^{\lambda} - 1)/\lambda$$
 $\lambda \neq 0$
= log (x) $\lambda = 0$.

A value of $\lambda = 0$ would indicate a log transformation was preferred, while $\lambda = 0.25$ is the fourth root transformation and $\lambda = 0.50$ is the square root transformation. After the optimal transformation is found, it is possible to test whether each of the three standard transformations is significantly rejected by Chi-square test in comparison with the optimal λ .

We estimated 24HDD at three sites (replicates) on each of 12 collecting dates, and examined these 12 data sets for each taxon to determine the appropriate data transformations. Using the three sites as replicates allows inference about a "typical" stream section in this particular river. One could, depending on the question asked, place three nets within a single section and make statements about the variability within that section, or

TABLE 1. Need for data transformation as examined by the Taylor power law regression (equation (5)) and by the Box-Cox procedure (equation (4)). A significant r^2 indicates that sample variance depended on sample mean. The intercept (a) and slope (b) refer to equation (5). The optimal transformation by the Box-Cox procedure (λ) is compared with three standard transformations by Chi-square test. Underlined Chi-square value indicates which standard transform is closest to optimal, while asterisks indicate rejected transformations (*P < 0.05; **P < 0.01).

					Chi-square		
	r^2	а	<i>b</i> ± 95%CL	λ	$\frac{\log x}{(\lambda = 0)}$	$\frac{\sqrt[4]{x}}{(\lambda = 0.25)}$	$\frac{\sqrt[3]{x}}{(\lambda = 0.50)}$
B. bicaudatus	0.523**	0.629	1.689±1.122	0.242	2.39	0.002	2.80
Cinygmula sp.	0.821**	0.283	1.939 ± 0.631	-0.065	<u>0.63</u>	13.36**	37.43**
E. longimanus ^a	0.739**	0.435	1.425 ± 0.589	0.196	14.50**	1.04	25.05**
D. coloradensis	0.421**	0.481	1.679 ± 1.370	-0.015	0.03	6.57*	20.78**
E. infrequens ^a	0.705**	0.330	1.867 ± 0.830	0.261	3.17	0.005	2.57
Z. haysi	0.406*	0.579	1.630 ± 1.373	-0.009	0.006	3.98*	13.70**
Prosimulium spp. ^a	0.672**	0.565	1.568 ± 0.763	0.251	20.23**	0.001	13.09**
Chironomidae	0.530**	0.876	1.638 ± 1.074	0.313	2.35	<u>0.10</u>	0.89

^aData sets contained zeros, so were converted to (x + 1) prior to log transformation.

place one net in each of three nearby rivers and make inferences about rivers in that region. We suspect that our choice of scale probably is of wide application.

The estimate of the optimal transformation $(\hat{\lambda})$ is given in Table 1. The acceptability of various standard transforms is indicated by the Chi-square value (1 df) obtained by testing, via a likelihood ratio test proposed by Box and Cox (1964), each of log, fourth root, and square root against λ . This Chi-square is calculated as $-2 \log_e \gamma$, where γ is obtained by dividing the probability of the observed distributions under various null hypotheses ($\lambda = 0, 0.25, 0.50$) by the probability obtained with the maximum likelihood estimate. The optimal transformation was closest to the fourth root transformation for five taxa, closest to the log transformation for three taxa, and was never closest to the square root transformation (Table 1). However, the log transformation deviated significantly from the optimum for only two of the eight taxa examined, whereas the fourth root transformation deviated significantly from optimal for three of eight taxa.

We also used Taylor's (1961) power law method to investigate the dependency of sample variance on sample mean. For each taxon, the 12 independent estimates of sample variance and sample mean were fit to the equation

(5) $s^2 = a\bar{x}^b$

where s^2 = sample variance, \bar{x} = sample mean, and a and b are estimated by linear regression of the log-transformed variables s^2 and \bar{x} .

All regressions were significant, with values of b ranging from 1.42 to 1.94. Some 40-82% of the variation in logarithm of variance estimates was accounted for by variation in the logarithm of the means (Table 1).

Precision and 95% Confidence Limits (95% CL)

The 95% CL associated with any sample variance can be determined as a "times-divide factor" (Elliott 1977, p. 90-91) which, multiplied or divided into the back-transformed mean, is similar to a \pm CL. For example, a times-divide factor of 1.5 is equivalent to 95% CL \pm 50% of the mean. It is calculated from

(6)
$$\stackrel{\times}{+}$$
 factor = antilog $(t_{0.05} \sqrt{s^2/n})$

where s^2 = sample variance computed from log-transformed data, n = number of replicates per sample, and $t_{0.05}$ = tabulated value from Student's *t*-distribution with n - 1 degrees of freedom. In order to employ the times-divide factor, the logarithmic transformation must be used.

Sample variances were computed for log-transformed 24HDD estimates for the abundant mayfly *B. bicaudatus* on each of 12 dates (n = 3 replicates per date, data of Table 1). We then estimated 95%CL as a function of number of replicates *n*, and because true sample variance (σ^2) usually is unknown, we chose several representative values of s^2 as estimates of σ^2 . As

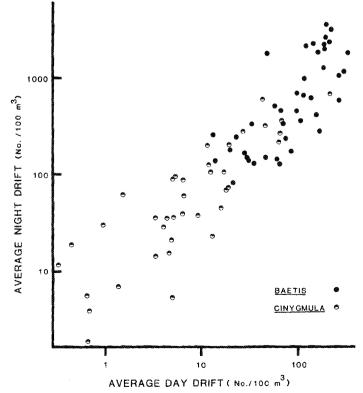


FIG. 1. Association between average day drift and average night drift for *B*. bicaudatus and Cinygmula.

TABLE 2. Proportion of variation accounted for (r^2) and coefficients (a, b) from regression analyses of drift density. The first variable listed is the dependent variable. All regressions were significant. With 10 df, $r^2 > 0.33$ is significant at 0.05, $r^2 > 0.50$ is significant at 0.01 level.

		ge night dri erage day d		First night drift sample with remaining 7 samples		
	r^2	а	b	r^2	a	Ь
B. bicaudatus ^a	0.634	0.154	1.581	0.711	-0.885	0.260
Cinygmula sp. ^b	0.729	0.670	1.044	0.724	0.579	0.156
E. longimanus ^a	0.616	0.598	1.098	0.765	0.286	0.213
E. coloradensis ^b	0.331	0.298	1.164	0.692	0.120	0.313
E. infrequens ^a	0.449	0.454	1.160	0.609	~0.186	0.237
Z. haysi ^b	0.694	0.713	0.988	0.710	0.517	0.152
Prosimuliumª	0.802	0.314	1.135	0.897	0.014	0.186
Chironomidae*	0.793	0.234	0.961	0.801	0.336	0.150

^aFourth root transformation.

^bLogarithmic transformation.

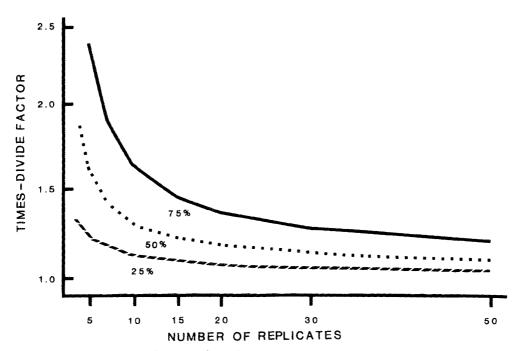


FIG. 2. The 95% CL of estimated drift density of *B. bicaudatus* expressed as a multiple of the mean (times-divide factor), as a function of number of replicate samples. Of 41 collections, 25% had variances resulting in a times-divide factor as small or smaller than the lowest curve. Half the data sets had variances that would result in times-divide factors as small or smaller than the 50% curve, while one quarter of the data sets had variances that would result in a times-divide factor greater than the 75% curve.

expected, 95% CL declined rapidly, then more slowly, with increasing n (Fig. 1). Some 25% of the 12 data sets had variances less than or equal to the value producing the lowest line in Fig. 1. The next largest 25% of variance estimates resulted in CL between the bottom and middle curves. The next largest 25% of variance estimates resulted in CL between the middle and upper curves, while 25% of the data sets result in CL larger than those indicated by the top line.

Estimating Drift from a Single Sample

For a reasonable estimate of total drift from a single sample, day and night drift should be highly correlated. We regressed average day drift, computed from four samples during the day, on average night drift, computed from three samples during the night, using sample drift density (equation (1)) and either the log or fourth root transformation, depending upon the taxon (Table 1). Some 33-80% of the variation in average night drift density was explainable by variation in average day drift density. Both *E. infrequens* and *D. coloradensis* showed low correlations, while in the other six taxa the coefficient of determination was >0.62 (Table 2).

The correspondence between night and day drift is depicted graphically for *B. bicaudatus* and *Cinygmula* (Fig. 2). There was a tendency for scatter to increase at low drift densities. The range of values was less for some taxa (e.g. the Chironomidae) than for others (e.g. *B. bicaudatus*), which may account in part for observed differences in strength of correlation (Table 2). Finally, we regressed drift density from the first night collection against drift density of the remaining seven samples from each 24-h period to determine how well a single sample predicted the information that would be lost if complete diel sampling was omitted. As 60-90% of the variation in estimates of drift from seven of the diel samples was accounted for by variation in the single sample collected just after dark (Table 2), it is clear that 24-h drift is predicted well by sampling only once in the 24-h period.

Discussion

For each of the eight taxa examined, a transformation was required to normalize the data and reduce the dependency of sample variance on sample mean. Values of b from Taylor's power law (equation (5)) suggest a fourth root transformation if b = 1.5, and a log transformation if b = 2. With the possible exception of two taxa (*Cinygmula* and *E. infrequens*), from the results of Table 1 we suggest that the fourth root should be a good general transformation for drift data. However, log transformation eliminated any significant dependence of sample variance on sample mean for each of the eight taxa (unpubl. analysis). The Box-Cox procedure gave broadly similar results (Table 1); the log transformation was significantly worse than the optimal for two taxa, while the fourth root was significantly worse than optimal for three taxa.

Reducing the dependency of variance on the mean (Taylor 1961) is one approach to meeting the assumptions of analyses of variance. However, fitting equation (5) requires several samples in order to compute an adequate regression equation. The approach of Box and Cox (1964) can be employed with a smaller number of samples (even one if replicate \times treatment degrees of freedom are adequate, approximately >20-30). Either transformation will improve the performance of subsequent analyses if data initially are skewed (Scheffé 1959), and a transformation is especially needed if multiple comparison procedures are to be employed.

Clearly, one cannot generalize about the adequacy of a single transformation to normalize data and reduce the dependence of the variance on the mean in every instance, although it may be that either the log or fourth root transformation usually will be successful. For a data set that includes a number of taxa, it may be desirable to employ the same transformation for all, using the best common or closest standard transformation, provided that whichever transformation one chooses is not strongly counterindicated for some taxon. Finally, there may be other reasons for choosing a particular transformation; for instance, the timesdivide approach to 95% CL requires log transformation of the data.

Analysis of the precision of replicate estimates of 24HDD, expressed as a times-divide factor, showed that sample variance itself is variable (see the family of curves in Fig. 1). With only one or two samples, there is little reason to believe that estimated drift density is within 100-200% of its true value. Some six to seven replicates are required, based on the data for a single, abundant species (*B. bicaudatus*), to expect 95% CL on drift to fall within $\pm 50\%$ of the mean.

This appears to be somewhat better precision than benthic samples provide. Using replicate Surber samples from the same stream, Allan (1984) estimated 95% CL using the average variance obtained for *B. bicaudatus*. Five benthic samples resulted in 95% CL of at least $\pm 130-150\%$ of the mean, and 10-15 benthic samples would be needed to achieve the same

precision obtained with 5 drift samples. This agrees with Chutter's (1975) conclusion that for the same number of replicates, drift samples are more precise than benthic samples.

Since it is common for studies of drift to utilize six or eight samples spread over a 24-h period to obtain a single summed estimate of drift density or rate, replicating the entire effort six times at perhaps two sites (for comparative purposes) becomes a formidable undertaking. This was the basis for investigating the correlation between day and night drift, in order to justify using a single sampling period as a predictor of 24-h drift. The first night sample, which typically is the largest of the 24-h series, provided most of the information (60-90%, Table 2) available in the remaining seven samples. Wherever the purpose of the investigation is to compare drift densities in stream sections exposed to various experimental or natural treatments, rather than estimate absolute numbers drifting per 24 h, we consider that effort would be better invested in some number of replicated samples collected just after nightfall, rather than an unreplicated series of collections spread over 24 h.

This recommendation assumes that one cannot simply leave a net in place continuously for 24 h. Although in some instances, one can (Waters 1969), the quantity of drifting insects and debris often precludes this possibility, and necessitates a choice between repeated sampling over 24 h and replicated sampling at a single time. Thus, for many questions concerning comparisons, we suggest it is preferable, for the same effort, to know the precision of an estimate at a single time rather than to estimate diel periodicity with unknown precision. Because most taxa exhibit peak drift just after dark, that is the best single time to sample.

The strong correlation between day and night drift may also have some bearing on the ideas of constant drift, which refers to the low level of drift usually observed during the day, and behavioral drift, which refers to the typically large increase in drift that occurs at night (Waters 1965). Based on these categorizations, one might expect that night (= behavioral) might be a very poor predictor of day (= constant) drift. Clearly, day and night drift are related, however, and the relationship was significant after the effects of benthic density and discharge were removed by partial correlation analysis (J.D. Allan, unpubl. data). This may indicate that day drift should not be viewed as an accidental event completely unrelated to the behavior that results in large peaks in night drift, but perhaps reflects incomplete suppression of the identical behaviors.

Acknowledgments

We are grateful to J. Kuhar and D. Robison for assistance in collecting and sorting samples, and L. Douglass for statistical advice. T. F. Waters, Alex Flecker, J. A. Downing, and reviewers provided helpful comments on an earlier draft. This work was supported by NSF grants to JDA and computer funds from the University of Maryland Computer Science and Sea Grant Program.

References

- ALLAN, J. D. 1975. The distributional ecology and diversity of benthic insects in Cement Creek, Colorado. Ecology 56: 1040-1053.
 - 1982. The effects of reduction in trout density on the invertebrate community of a mountain stream. Ecology 63: 1444–1455.

1984. Hypothesis testing in ecological studies of the freshwater benthos, p. 484–507. *In* V. H. Resh and D. M. Rosenburg [ed.] Ecology of aquatic insects. Praeger, New York, NY.

BOX, G. E. P., AND D. R. COX. 1964. An analysis of transformations. J. R. Stat. Soc. Ser. B 26: 211-243. CHUTTER, F. M. 1975. Variation in the day-time drift of a Natal river. Verh. Int. Ver. Limnol. 19: 1728-1735.

- DOWNING, J. A. 1979. Aggregation, transformation, and the design of benthos sampling programs. J. Fish. Res. Board Can. 36: 1454-1463.
- ELLIOTT, J. M. 1970. Methods of sampling invertebrate drift in running waters. Ann. Limnol. 6: 133-159.
 - 1977. Some methods for the statistical analysis of samples of benthic invertebrates. Sci. Publ. No. 25. 2nd ed. Freshwater Biological Association, Cumbria, England. 156 p.
- GREEN, R. H. 1979. Sampling design and statistical methods for environmental biologists. Wiley, New York, NY. 257 p.
- MULLER, K. 1954. Investigations on the organic drift in north Swedish streams. Rep. Inst. Freshwater Res. Drottningholm 35: 133-148.
 - 1974. Stream drift as a chronobiological phenomenon in running water ecosystems. Annu. Rev. Ecol. Syst. 5: 309--324.

- RESH, V. H. 1979. Sampling variability and life history features: basic considerations in the design of aquatic insect studies. J. Fish. Res. Board Can. 36: 290-311.
- SCHEFFÉ, H. 1959. The analysis of variance. Wiley, New York, NY.
- TAYLOR, L. R. 1961. Aggregation, variance and the mean. Nature (Lond.) 189: 732-735.
- ULFSTRAND, S. 1968. Benthic animal communities in lapland streams. Oikos Suppl. 10: 1-120.
- WATERS, T. F. 1965. Interpretation of invertebrate drift in streams. Ecology 46: 327–334.

1969. Invertebrate drift ecology and significance to stream fishes, p. 121-134. *In* T. G. Northcote [ed.] Symposium on salmon and trout in streams. University of British Columbia, Vancouver, B.C.

1972. The drift of stream insects. Annu. Rev. Entomol. 17: 253-272.