Computation of confidence intervals for the size-frequency (Hynes) method of estimating secondary production

Abstract—An estimator of variance is described for production values calculated by the size-frequency (Hynes) method. The number of sampling dates is identified as having an important influence on the variance estimate. A geometric mean is proposed to estimate the average weight of individuals when lost between size classes. In an example calculation, Ephemerella dorothea (Ephemeroptera) from North Branch Creek, Minnesota, had an annual production value of 8.5 g·m⁻² wet and an approximate 95% confidence interval of 7.0–10.0. Similar calculations for Baetis sp. (Ephemeroptera) from the same river yielded an annual production value of 4.7 g·m⁻² wet with a confidence interval of 3.6–5.8. Significant differences in annual production thus exist between these mayflies.

The size-frequency (Hynes) method of estimating secondary production, as originally proposed, facilitated calculations involving entire benthic faunas, permitting the combination of data on all benthic taxa present (Hynes 1961). An important advantage of the method was that single cohorts within the data did not need to be identified to calculate production. Hynes at the time claimed only an accuracy of “an order of magnitude” for the production estimates. The method was later modified due to various conceptual errors and certain criteria were established for the data (Hynes and Coleman 1968; Hamilton 1969). These

1 Paper 10,664, Scientific Journal Series, Minnesota Agricultural Experiment Station, St. Paul 55108. This research was financed in part by a research grant from the National Science Foundation.

© 1980, by the American Society of Limnology and Oceanography, Inc.


Submitted: 24 July 1979
Accepted: 6 January 1980

The restrictions imposed by these criteria required a more complex categorization of data for production estimation than initially proposed by Hynes (1961). Theoretically, it remains possible to combine data from various taxa, provided that they all meet the criteria. In practice, however, the consequence of satisfying the criteria usually is that the benthic data must be taxonomically organized at a generic or species level in order to maintain the uniformity of life history required. As a result, empirical applications of the method have been primarily to single species (e.g. Waters and Crawford 1973; Eckblad 1973; Winterbourne 1974; McClure and Stewart 1976; Martien and Benke 1977) or to combinations of similar species (Neves 1979; Krueger 1979). The most notable exception is the work of Fisher and Likens (1973) in which benthic taxa were combined regardless of life cycle or trophic level. The method does however retain the original advantage of being independent of the need to identify single cohorts.

The precision of the size-frequency
method has also been examined in relation to its present taxonomically specific use. Size-frequency production estimates have been consistent with estimates made by other calculation methods either with empirically derived data (Waters and Crawford 1973; Resh 1977; Welton 1979) or those from computer simulations (Cushman et al. 1978). The accuracy of the size-frequency method in its present application appears to be nearly equal to that of other methods and much greater than Hynes’ original suggestion of an order of magnitude. Recent modifications of the technique (Benke 1979) further increase its accuracy. It is well accepted in secondary production research (Cushman et al. 1978; Benke 1979; Waters 1979).

It now seems appropriate to determine the effect of sampling variability on the size-frequency production estimates. An estimate of variance would provide a way to evaluate the precision of production calculations. We here define an estimator of variance and compute confidence intervals for size-frequency production estimates. The use of the variance estimate is demonstrated through a comparison of production estimates for two mayflies.

Raw data for size-frequency production estimates include observations of the numbers of animals within size classes per unit area at intervals over a certain period of time (frequently 1 year). In the following equations, sampling dates are indexed by $i$ and proceed from 1 through $n$, where $n$ represents the total number of sampling dates ($i = 1, 2, \ldots, n$). Size classes are indexed by $j$ ($j = 1, 2, \ldots, a$) from the smallest to the largest. Samples within date $i$ are ordered by $k$ ($k = 1, 2, \ldots, b_i$) where $b_i$ represents the total number of samples taken on date $i$.

The mean number of individuals by size class $j$ observed on the $i$th date is

$$\hat{Y}_{ij} = \frac{\sum_{k=1}^{b_i} Y_{ijk}}{b_i} \quad (1)$$

where $Y_{ijk}$ is the number of individuals observed on the $i$th date in the $j$th size class from the $k$th sample drawn. The sampling variance for the mean number of individuals per size class by sampling date ($\hat{Y}_{ij}$) is estimated by

$$v(\hat{Y}_{ij}) = \frac{\sum_{k=1}^{b_i} (Y_{ijk} - \hat{Y}_{ij})^2}{b_i(b_i - 1)}. \quad (2)$$

The estimate of the number of individuals on the average during the sampling year for each size class ($\tilde{Y}_{ij}$) is weighted by sampling interval length. This calculation of $\tilde{Y}_{ij}$ can easily be thought of as the prevalence of the $j$th size class in the population. $\tilde{Y}_{ij}$ is expressed as

$$\tilde{Y}_{ij} = \frac{\sum_{i=1}^{n-1} (D_{i+1} - D_i) \hat{Y}_{ij} + \hat{Y}_{i+1\cdot j}}{2D_n - D_1} \quad (3)$$

where $D_i$ is the number of days between the first sampling day and the $i$th sampling date. The variance for the estimate of the prevalence of the $j$th size class during the sampling year ($\tilde{Y}_{ij}$) is calculated according to

$$v(\tilde{Y}_{ij}) = \left\{ (D_2 - D_1)^2 \frac{v(\hat{Y}_{i+1\cdot j})}{4} + \sum_{i=2}^{n-1} (D_{i+1} - D_{i-1})^2 \frac{v(\hat{Y}_{ij})}{4} + (D_n - D_{n-1})^2 \frac{v(\hat{Y}_{nj})}{4} \right\} / (D_n - D_1)^2. \quad (4)$$

In the above variance derivation $\text{cov}(\hat{Y}_{ij}, \hat{Y}_{i+1\cdot j}) = 0$ since sampling is independent from one date to another. If a new and independent randomization of the sampling procedure occurs on each sampling date, it follows that the means on two successive sample dates are statistically independent, and their covariance is zero.

In the production calculation given here, the geometric mean $(W_j W_{j+1})^{0.5}$ of the mean weights per individual of size classes $j$ and $j + 1$ (see Sokal and Rohlf 1969) is used to estimate the weights of individuals lost. The geometric mean gives a value always less than the arith-
metric mean (Simpson et al. 1960) and is more accurate for logarithmic growth patterns. This modification of the size-frequency production calculation can be written as

\[
Pw = a \left[ \sum_{j=1}^{a-1} (\bar{Y}_{j, \cdot} - \bar{Y}_{j+1, \cdot})(W_jW_{j+1})^{0.5} \right] + (\bar{Y}_{\cdot, a})(W_a)
\]

(5)

where \(W_j\) is the mean weight per individual in size class \(j\). An earlier version of Eq. 5 is given by Hamilton (1969). The variance of this estimate, \(v(Pw)\), is estimated by

\[
v(Pw) = a^2 \left[ G_{1,1}^2 v(\bar{Y}_{1, \cdot}) \right. \\
\left. + \sum_{j=2}^{a-1} (G_j - G_{j-1})^2 v(\bar{Y}_{j, \cdot}) \right] + (W_a - G_{a-1})^2 v(\bar{Y}_{\cdot, a})
\]

(6)

where \(G_j\) is \((W_jW_{j+1})^{0.5}\). This calculation of \(v(Pw)\) is based on the fact that \(Pw\) is the linear combination of random variables \(\bar{Y}_{j, \cdot}\) which are assumed to be uncorrelated. If the variables \(\bar{Y}_{j, \cdot}\) are correlated, then we suspect that a negative relationship is likely. Such a correlation could be due to differences in environmental requirements between life stages affecting, through microdistributional patterns, the capture success of size classes within samples (Cummins 1964; Minshall 1967; Lehmkohl and Anderson 1972; Minshall and Minshall 1977). If \(\text{cov}(\bar{Y}_{j, \cdot}, \bar{Y}_{j+1, \cdot})\) is negative then Eq. 6 becomes a conservative estimate of \(v(Pw)\). Mean weights by size class \((W_j)\) are treated as constants in the derivation of Eq. 6 since they can be determined relatively precisely. An approximate 95% confidence interval which represents 2 SE on either side of the production estimate \((Pw)\) can be calculated:

\[
Pw \pm 2[v(Pw)^{0.5}].
\]

(7)

The calculations described above can be used only for univoltine species where pupal, adult, or egg stages do not comprise a significant portion of the total generation time. When the cohort production interval (CPI), in days from hatching to the attainment of the largest aquatic size class, is less than 1 year for a univoltine species, or if multivoltine or semivoltine species are involved, the production estimate \((Pw)\) must be multiplied by 365/CPI (Benke 1979). Variance for this new estimate would then become

\[
v(Pw) \cdot (365/\text{CPI})^2.
\]

(8)

The confidence interval follows as described above (Eq. 7) except that the variance for this new estimate (Eq. 8) is used. Our computer program, which performs these calculations, is available at cost.

We present an empirical example for *Ephemerella dorothea* from North Branch Creek, Fillmore Co., Minnesota. A description of the sampling procedures and life cycle determinations was given by Krueger (1979). The CPI for this mayfly was about 9 months (275 days). The basic data, means \(\bar{Y}_{\cdot, j}\), and their variances \(v(\bar{Y}_{\cdot, j})\) from Eq. 2 are given in Table 1. The estimate of annual production \((Pw)\) for *E. dorothea* was 8.5 \(\text{g} \cdot \text{m}^{-2} \cdot \text{yr}^{-1}\) wet and had a variance of 0.54 (Table 2). This variance is based on data from 13 dates and 10 samples from each date. A close look at Eq. 4 for the \(v(\bar{Y}_{\cdot, j})\) shows that the number of sampling dates is the most important influence on controlling variance and that the variance is minimum when sampling dates are equally spaced. The confidence interval derived from the variance surrounding this estimate indicates that the production value is much greater than 0.0 and with good probability occurs between 7.0 and 10.0 \(\text{g} \cdot \text{m}^{-2} \cdot \text{yr}^{-1}\) wet.

A similar calculation was performed for another mayfly, *Baetis* (species unknown), from the same location in North Branch Creek (Krueger 1979). Annual production was 4.7 \(\text{g} \cdot \text{m}^{-2} \cdot \text{yr}^{-1}\) wet. For this estimate a confidence interval of 3.6–5.8 was calculated. The lack of interval overlap between the production estimates for *E. dorothea* and *Baetis* sp. suggests that
Table 1. Means $\bar{Y}_b$ and (variances) $v(\bar{Y}_b)$ of numbers per 0.1 m$^2$ of *Ephemera dorothea* captured in bottom samples (North Branch Creek, Minnesota), March 1977–February 1978.

<table>
<thead>
<tr>
<th>Date</th>
<th>Size classes (mm)</th>
<th>1.0-2.5</th>
<th>2.5-4.0</th>
<th>4.0-5.5</th>
<th>5.5-7.0</th>
<th>7.0-8.5</th>
<th>8.5-10.0</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$b_i$=10</td>
<td>$\bar{Y}_b$</td>
<td>$\bar{Y}_b$</td>
<td>$\bar{Y}_b$</td>
<td>$\bar{Y}_b$</td>
<td>$\bar{Y}_b$</td>
<td>$\bar{Y}_b$</td>
</tr>
<tr>
<td>1 Mar</td>
<td>$b_1$=10</td>
<td>86.0</td>
<td>111.6</td>
<td>41.3</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(604.3)</td>
<td>(1,017.1)</td>
<td>(139.3)</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>23 Mar</td>
<td>$b_2$=10</td>
<td>25.4</td>
<td>63.6</td>
<td>32.9</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(28.4)</td>
<td>(175.9)</td>
<td>(42.2)</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>24 Apr</td>
<td>$b_3$=10</td>
<td>3.4</td>
<td>23.1</td>
<td>73.1</td>
<td>66.0</td>
<td>18.6</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.9)</td>
<td>(37.7)</td>
<td>(368.2)</td>
<td>(301.8)</td>
<td>(24.3)</td>
<td>—</td>
</tr>
<tr>
<td>26 May</td>
<td>$b_4$=10</td>
<td>1.8</td>
<td>6.9</td>
<td>5.4</td>
<td>3.3</td>
<td>3.4</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.2)</td>
<td>(2.3)</td>
<td>(1.4)</td>
<td>(1.1)</td>
<td>(1.9)</td>
<td>(0.1)</td>
</tr>
<tr>
<td>24 Jun</td>
<td>$b_5$=10</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>22 Jul</td>
<td>$b_6$=10</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>2 Aug</td>
<td>$b_7$=10</td>
<td>0.2</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.02)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>21 Sep</td>
<td>$b_8$=10</td>
<td>2.5</td>
<td>0.0</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.2)</td>
<td>(0.01)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>20 Oct</td>
<td>$b_9$=10</td>
<td>22.4</td>
<td>3.3</td>
<td>0.2</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(136.3)</td>
<td>(2.0)</td>
<td>(0.02)</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>29 Nov</td>
<td>$b_{10}$=10</td>
<td>4.0</td>
<td>13.2</td>
<td>2.9</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2.5)</td>
<td>(23.1)</td>
<td>(0.5)</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>24 Dec</td>
<td>$b_{11}$=10</td>
<td>3.9</td>
<td>13.4</td>
<td>14.1</td>
<td>1.2</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.4)</td>
<td>(16.4)</td>
<td>(18.2)</td>
<td>(0.2)</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>22 Jan</td>
<td>$b_{12}$=10</td>
<td>2.5</td>
<td>16.8</td>
<td>17.7</td>
<td>4.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.9)</td>
<td>(39.9)</td>
<td>(43.6)</td>
<td>(2.0)</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>28 Feb</td>
<td>$b_{13}$=10</td>
<td>2.0</td>
<td>11.4</td>
<td>30.2</td>
<td>7.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.3)</td>
<td>(9.7)</td>
<td>(65.8)</td>
<td>(3.5)</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

Table 2. Calculation of annual production (g·m$^{-2}$ wet) of *Ephemera dorothea* by size-frequency method.

<table>
<thead>
<tr>
<th>Size class (mm)</th>
<th>$\bar{Y}_b$ (m$^{-2}$)</th>
<th>(2) $v(\bar{Y}_b)$</th>
<th>(3) $\bar{Y}<em>b - \bar{Y}</em>{b+1}$</th>
<th>(4) $W_i$ (mg)</th>
<th>(5) $C_i$</th>
<th>(6) $(3 \times 5)$</th>
<th>(7) Production$^*$ (g·m$^{-2}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0-2.5</td>
<td>82.3</td>
<td>197.5</td>
<td>-70.6</td>
<td>0.40</td>
<td>0.67</td>
<td>-0.047</td>
<td>-0.28</td>
</tr>
<tr>
<td>2.5-4.0</td>
<td>152.9</td>
<td>284.5</td>
<td>2.2</td>
<td>1.12</td>
<td>1.91</td>
<td>0.004</td>
<td>0.02</td>
</tr>
<tr>
<td>4.0-5.5</td>
<td>150.7</td>
<td>387.5</td>
<td>81.9</td>
<td>3.27</td>
<td>4.95</td>
<td>0.405</td>
<td>2.43</td>
</tr>
<tr>
<td>5.5-7.0</td>
<td>68.8</td>
<td>235.5</td>
<td>49.6</td>
<td>7.51</td>
<td>9.17</td>
<td>0.454</td>
<td>2.72</td>
</tr>
<tr>
<td>7.0-8.5</td>
<td>19.2</td>
<td>20.0</td>
<td>18.5</td>
<td>11.20</td>
<td>13.00</td>
<td>0.240</td>
<td>1.44</td>
</tr>
<tr>
<td>8.5-9.5</td>
<td>0.7</td>
<td>0.1</td>
<td>0.7</td>
<td>15.00</td>
<td>15.00</td>
<td>0.010</td>
<td>0.06</td>
</tr>
</tbody>
</table>

\[ P_w = \text{total} \times 365/275 = 8.5 \]
\[ v(P_w) = 0.54 \]
\[ 7.0 \leq P_w \leq 10.0 \]

* Values are 6 (-a) × values in col. 6.
significant differences in annual production exist between these mayflies. The calculation of the variance of size-frequency production estimates as described allows an interpretation of the reliability of the estimate and provides a means of comparison with other production values.

Charles C. Krueger
Department of Entomology,
Fisheries and Wildlife
University of Minnesota
St. Paul 55108

Frank B. Martin
Department of Applied Statistics
University of Minnesota
St. Paul 55108

References


WINTERBOURNE, M. J. 1974. The life histories, trophic relations and production of Stenopera prasina (Plecoptera) and Delatidium sp. (Ephemeroptera) in a New Zealand river. Freshwater Biol. 4: 507–524.

Submitted: 29 January 1979
Accepted: 28 February 1980