North American Journal of Fisheries Management 30:191–208, 2010

DOI: 10.1577/M08-237.1

[Article]

Performance of Three Alternative Estimators of Stream Residence Time Based on Live and Dead Counts of Salmonids

John R. Skalski*
School of Aquatic and Fishery Sciences, University of Washington,
1325 Fourth Avenue, Suite 1820, Seattle, Washington 98101-2509, USA

Shiquan Liao
King County Department of Judicial Administration,
516 Third Avenue, Room 609, Seattle, Washington 98104, USA

Rebecca A. Buchanan
School of Aquatic and Fishery Sciences, University of Washington,
1325 Fourth Avenue, Suite 1820, Seattle, Washington 98101-2509, USA

Abstract.—Data from live and dead fish counts at salmonid spawning areas have been used for decades to estimate stream residence time. Simulation studies indicate that the peak-to-peak and median-to-median estimates of stream residence time will have at least a 50% negative bias under the best of circumstances. A new estimation technique called the "expectation-to-expectation" method is shown to be approximately unbiased over a range of arrival and lifetime distributions. The method also allows corrections for incomplete detection probabilities and carcass retention rates less than 1.0. An important byproduct of the approach is a concurrent estimate of total escapement.

Spawner surveys are an integral part of salmonid management in natural systems (Knudsen 2000:251–254). A common method of studying stream escapement is the area-under-the-curve (AUC) method, in which an estimate of total fish days is divided by an estimate of stream residence time to obtain an abundance estimate (English et al. 1992; Hill 1997; Lady and Skalski 1998). Stream residence time has also been called "survey life" (Ames 1984; Perrin and Irvine 1990) or "stream life" (Bue et al. 1998). Of six techniques reviewed by Perrin and Irvine (1990) for estimating stream residence time in natural streams, none was found to be without bias.

Survey data suggest that stream residence time for salmonid species is not a constant but varies from year to year and from stream to stream (Thomason and Jones 1984). Killick (1955) found stream residence time to be correlated with run timing, and there is substantial evidence that residence time changes over the course of the spawning season (Thomason and Jones 1984; Fukushima and Smoker 1997; Bue et al. 1998). This natural variability dictates that estimates of stream residence time used in estimation of spawner escapement must be seasonwide, site specific, and year specific in order to yield unbiased abundance estimates.

Knudsen (2000:247) summarized escapement methods used on U.S. salmonid populations. He found that out of 3,688 populations for which escapement surveys were performed, 241 (6.5%) populations were evaluated with foot surveys of live and/or dead counts. Foot surveys were most commonly applied to populations of coho salmon *Oncorhynchus kisutch*, followed roughly by Chinook salmon *O. tshawytscha*, chum salmon *O. keta*, and pink salmon *O. gorbuscha*. Knudsen (2000) rated the foot surveys as generally of poor quality and recommended additional research on AUC techniques. The AUC method has also been used in conjunction with boat or snorkel surveys (Cousens et al. 1982) and aerial counts (Hill 1997; Bue et al. 1998) to estimate escapement.

Although mark–recapture methods have been specifically developed to estimate stream residence time (Lady and Skalski 1998; Manske and Schwarz 2000), the need for site- and time-specific estimates calls for cost-effective methods that can be readily implemented. Two methods based on periodic live and dead counts of spawners—the peak-to-peak (Perrin and Irvine 1990) and mean-to-mean methods (Lewis 1987; Perrin and Irvine 1990)—have been used extensively. This paper examines the bias of these two estimators and introduces a new approach that uses live and dead counts and that is nearly unbiased over a range of arrival and death time distributions.

---

* Corresponding author: jrs@cbr.washington.edu

Received November 18, 2008; accepted October 13, 2009

Study Area

Few escapement studies have formally attempted to estimate detection probabilities of carcasses or their retention rate in streams. Chum salmon escapement was studied at Perry Creek, Thurston County, Washington, in winter 1992–1993. The creek was located in a semiurban area approximately 6.6 km west of Olympia, Washington. The stream was 7.5 km in length and was blocked to anadromous fish use by a waterfall at river kilometer 2.5. During the chum spawning season (November and December), average flow was 0.70–0.85 m^3/s and average stream width was 7.6 m. The stream flows directly into the marine waters of Hood Canal, Washington. Annual chum salmon escapement to the stream has averaged 4,750 fish, with a range of 1,650–11,850 (1968–1982; Washington Department of Fish and Game, unpublished data).

Methods

Using live and dead counts, three different estimators of stream residence time were investigated. Two methods have been used for decades (Perrin and Irvine 1990): one is based on the difference in times of the peak live and peak dead counts (i.e., peak-to-peak method), and the other is based on the difference in the median dates (i.e., 50%) of the cumulative live and dead counts (i.e., median-to-median method). The third method, which we call the expectation-to-expectation method, will be developed below. Simulation studies were used to examine the sampling error and bias of the three alternative methods under a variety of escapement patterns and survey sampling designs.

The AUC method estimates total escapement (N) by dividing an estimate of total fish days ($\hat{F}$) by an estimate of stream residence time ($\hat{T}$), where

$$\hat{N} = \frac{\hat{F}}{\hat{T}}.$$

In the case where $\hat{F}$ and $\hat{T}$ are independent, the variance of the escapement can be estimated by

$$\widehat{\mathrm{Var}}(\hat{N}) = \hat{N}^2\left[\frac{\widehat{\mathrm{Var}}(\hat{F})}{\hat{F}^2} + \frac{\widehat{\mathrm{Var}}(\hat{T})}{\hat{T}^2}\right]$$

or

$$\widehat{\mathrm{Var}}(\hat{N}) = \hat{N}^2[\mathrm{CV}(\hat{F})^2 + \mathrm{CV}(\hat{T})^2],$$

based on the delta method (Seber 1982:7–9). Hence, the variance of $\hat{N}$ is a function of the precision of the separate estimates of F and T expressed in terms of their coefficients of variation (CV). From inspection of the above variance formula, $\hat{F}$ and $\hat{T}$ contribute equally to the precision of $\hat{N}$.

Peak-to-Peak Estimator

The peak-to-peak estimator ($\hat{T}_1$) provides an estimate of stream residence time based on the difference between the time (d) of maximum (peak) daily live counts and the time of maximum daily carcass counts. Assume that m stream surveys are conducted during a spawning season of M days and that the time for the ith survey is t_i ($i = 1, \ldots, m$). The observed numbers of live fish and carcasses are l_{t_i} and c_{t_i} ($i = 1, \ldots, m$), respectively. The peak-to-peak estimator can then be mathematically written as

$$\hat{T}_1 = t_{c,\max} - t_{l,\max}, \tag{1}$$

where $t_{c,\max} = t|\max(c_{t_i})$ is the time (d) the peak (i.e., maximum) carcass count occurs over the season. Analogously, $t_{l,\max} = t \mid \max(l_{t_i})$ is the time (d) of the maximum live count over the season.

During a season, it is possible that the number of live fish or carcasses observed during the season may be multimodal. When either of the live or dead counts has multiple modes, $t_{c,\max}$ and/or $t_{l,\max}$ are not well defined. As such, the associated estimator ($\hat{T}_1$) of stream residence time will also fail to be well defined. In this study, the peaks with the highest counts were used when multimodal distributions occurred.

Median-to-Median Estimator

The second existing estimator based on the live and dead counts is the median-to-median estimator ($\hat{T}_2$). This estimator computes stream residence time based on the difference in time between the 50th percentiles of the dead counts and live counts. Let F_l and F_d denote the cumulative live count curve and the cumulative dead count curve, respectively. The median-to-median estimator can then be expressed as

$$\hat{T}_2 = t_{c,0.5} - t_{l,0.5}, \tag{2}$$

where $t_{c,0.5} = t \mid F_c^{-1}(0.5)$ is the time (d) when 50% of the total counted carcasses have been observed or is the time (t) such that

$$F_d(t) = \frac{\sum_{j=1}^{t} c_{t_j}}{\sum_{j=1}^{m} c_{t_j}} = 0.5.$$

Similarly, $t_{l,0.5} = t \mid F_l^{-1}(0.5)$ is the time when 50% of the total counted live fish have been observed. The estimator $\hat{T}_2$ is illustrated in Figure 1. It is worth noting that the time when the 50% live or 50% dead count occurs may not be unique as shown in the lower panel of Figure 1. In such a case, a decision will be required to

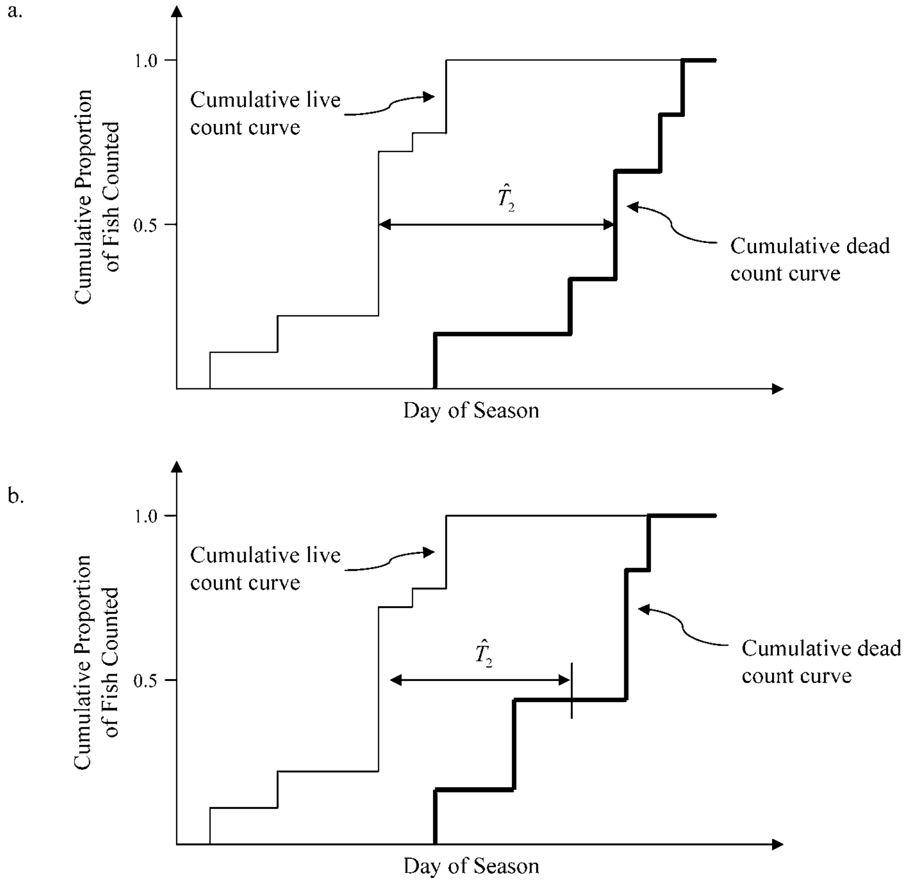

FIGURE 1.—Illustration of the median-to-median estimator of average stream residence time for all salmonid spawners ($\hat{T}_2$). The average stream residence time is estimated by the difference in time (d) between 50% death counts and 50% live counts. In (**a**), the time for 50% death counts and the time for 50% live counts are uniquely determined; in (**b**), the time for 50% deaths is not uniquely determined.

determine the time when the 50% carcass counts or live counts occur. In this study, the middle point on the cumulative curves is used if the time of the median is not uniquely determined.

Expectation-to-Expectation Estimator

The existing estimators, $\hat{T}_1$ and $\hat{T}_2$, compute estimates of stream residence time (T) using the numbers of observed live fish and carcasses in the stream. However, T is the average time that the spawners spend in the stream between their arrivals and deaths. Therefore, a methodologically correct estimate should be constructed based on these arrivals and death times. Though the actual arrivals and mortalities in the stream are not directly observable during stream surveys, these quantities can be estimated from the live and dead counts.

Letting T_{ai} denote the arrival time of the ith spawner and letting T_{di} denote the time of the ith spawner's death, then the stream residence time for the ith spawner is

$$T_i = T_{di} - T_{ai}$$

for $i = 1, \ldots, N$, where N is the total escapement during the season. Hence, the average stream residence time for all spawners is

$$T = \frac{1}{N}\sum_{i=1}^{N} T_i = \frac{1}{N}\sum_{i=1}^{N}(T_{di} - T_{ai}). \tag{3}$$

An estimator for T can be constructed by rewriting equation (3) as

$$\begin{aligned} T &= \frac{1}{N}\sum_{i=1}^{N} T_{di} - \frac{1}{N}\sum_{i=1}^{N} T_{ai} \\ &= \frac{1}{N}\sum_{j=1}^{M} jD_j - \frac{1}{N}\sum_{j=1}^{M} jA_j, \end{aligned} \tag{4}$$

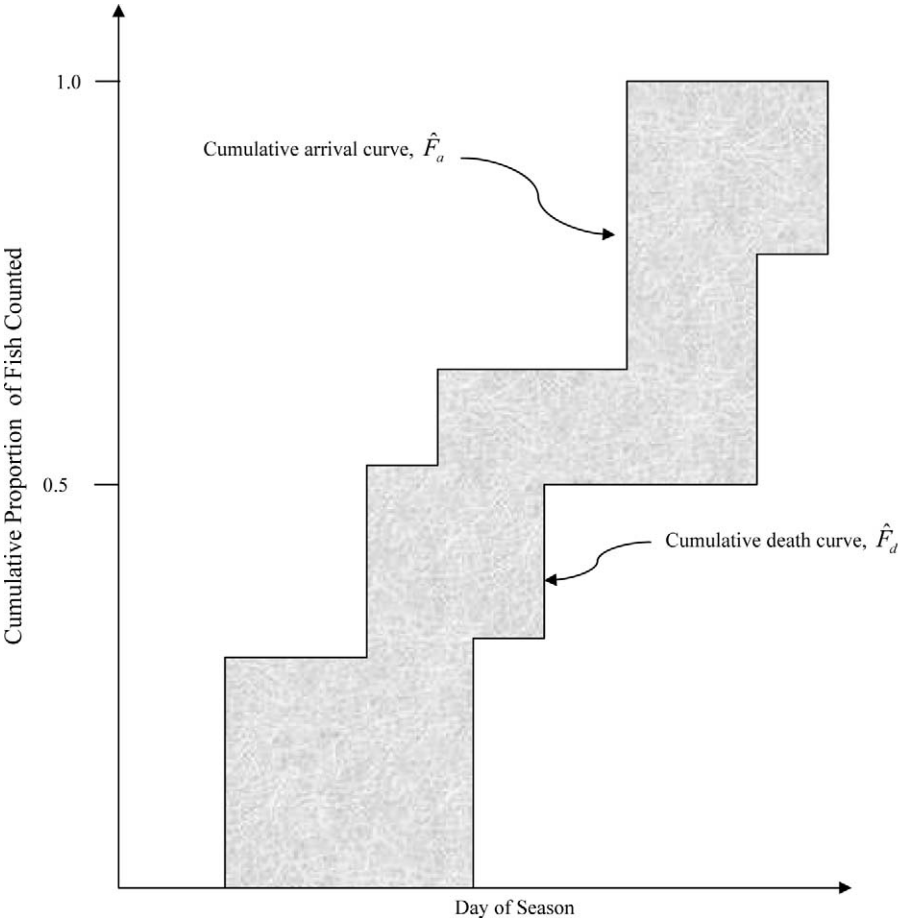

FIGURE 2.—Illustration of the expectation-to-expectation estimator of average stream residence time for all salmonid spawners ($\hat{T}_3$). The average stream residence time is estimated by the area between the estimated cumulative arrival curve ($\hat{F}_a$) and the estimated cumulative death curve ($\hat{F}_d$).

where M = length of the spawning season, D_j = number of mortalities occurring on day j ($j = 1, \ldots, M$), and A_j = number of arrivals on day j ($j = 1, \ldots, M$).

If $f(j)$ is a probability mass function with support 1, …, M and $F(j)$ is its associated cumulative distribution function, then (Mood et al. 1974:65):

$$\sum_{j=1}^{M} jf(j) - 1 = \sum_{j=1}^{M}[1 - F(j)].$$

Using this relationship, equation (4) simplifies to

$$\begin{aligned} T &= \sum_{j=1}^{M}[F_a(j) - F_d(j)] \\ &= \sum_{j=1}^{M}[j - (j-1)][F_a(j) - F_d(j)], \qquad (5) \end{aligned}$$

where F_a and F_d are the cumulative distribution functions for arrivals and dead fish, respectively. Starting off with a straightforward definition of stream residence time (equation 3), it was then re-expressed in terms of relative frequency distributions for arrival and death times, which in turn are re-expressed in terms of cumulative distributions. Equation (5) shows that the average stream residence time is the area between the cumulative arrival curve and the cumulative death curve (Figure 2).

In the case of only m days of stream surveys ($m \leq M$), equation (5) can be estimated by noting that the quantity $[j - (j - 1)]F_d(j)$ is approximated by

$$(t_j - t_{j-1})\frac{[\hat{F}_d(t_j) + \hat{F}_d(t_{j-1})]}{2}.$$

Similarly, $[j - (j - 1)]F_a(j)$ is approximated by

$$(t_j - t_{j-1})\frac{[\hat{F}_a(t_j) + \hat{F}_a(t_{j-1})]}{2}.$$

Hence, an estimator of the average stream residence time can be constructed as

$$\hat{T}_3 = \frac{1}{2}\sum_{j=1}^{m}(t_j - t_{j-1})\{[\hat{F}_a(t_j) + \hat{F}_a(t_{j-1})] - [\hat{F}_d(t_j) + \hat{F}_d(t_{j-1})]\}, \qquad (6)$$

where the cumulative dead curve is estimated by

$$\hat{F}_d(t_j) = \sum_{k=1}^{j}\frac{\hat{D}_{(t_k)}}{\hat{N}}, \qquad (7)$$

with $D_{(t_k)}$ representing the number of mortalities that occurred after survey day t_{k-1} and before day t_k:

$$D_{(t_k)} = \sum_{j=t_{k-1}+1}^{t_k} D_j.$$

Similarly, the cumulative arrival curve is given by

$$\hat{F}_a(t_j) = \sum_{k=1}^{j}\frac{\hat{A}_{(t_k)}}{\hat{N}}, \qquad (8)$$

with $A_{(t_k)}$ representing the number of arrivals between survey days t_{k-1} and t_k:

$$A_{(t_k)} = \sum_{j=t_{k-1}+1}^{t_k} A_j.$$

The estimates of arrivals ($\hat{A}_{(t_k)}$) and deaths ($\hat{D}_{(t_k)}$) during the survey period will be explained in the following two sections of the paper.

Estimating arrivals.—In practice, the $D_{(t_k)}$ and $A_{(t_k)}$ are not directly observable; only the number alive on day j (i.e., L_{t_j}) and the carcass count (i.e., C_{t_j}) are observable. The $\hat{A}_{(t_k)}$ can be derived as a function of the live and dead counts based on the fundamental relationship among the arrivals, mortalities, and live fish, where

$$N_j = \sum_{i=1}^{j} A_i = L_j + \sum_{i=1}^{j} D_i; \quad \forall j = 1, \ldots, M, \qquad (9)$$

with N_j representing the total number of fish in the stream on day j ($j = 1, \ldots, M$) and L_j representing the number of live fish in the stream on day j ($j = 1, \ldots, M$).

Equation (9) implies that escapement (defined as the cumulative arrivals) to any specified day during the season must be equal to the cumulative deaths plus the number of live fish still present in the stream on that day. From equation (9), we can write the daily arrivals over the season in terms of the mortalities and live fish as

$$A_j = D_j + L_j - L_{j-1}; \quad \forall j = 1, \ldots, M. \qquad (10)$$

Equation (10) simply implies that the number of arrivals on day j is equal to the mortalities occurring on the same day (D_j) plus the difference between the numbers of live fish on days j and $j - 1$ (i.e., $L_j - L_{j-1}$). For example, assume that there were 10 live fish present on day 5 ($L_5 = 10$) and that the number of live fish in the stream on day 6 was 20 ($L_6 = 20$). Also assume that on day 6, five spawners died ($D_6 = 5$). Hence, the number of arrivals on day 6 must be 15 fish ($A_6 = 5 + 20 - 10 = 15$).

A matrix expression of equation (10) can be written as

$$\underset{\sim}{A} = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_M \end{bmatrix} = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix}\begin{bmatrix} D_1 \\ D_2 \\ \vdots \\ D_M \end{bmatrix} + \begin{bmatrix} 1 & & & \\ -1 & 1 & & \\ & \ddots & \ddots & \\ & & -1 & 1 \end{bmatrix}\begin{bmatrix} L_1 \\ L_2 \\ \vdots \\ L_M \end{bmatrix}. \qquad (11)$$

The matrices on the right-hand side of equation (11) are of size $M \times M$ with zeroes for the blank entries.

When there are only m surveys ($m < M$), equation (10) can be modified as

$$A_{(t_k)} = D_{(t_k)} + L_{t_k} - L_{t_{k-1}}; \quad \forall k = 1, \ldots, m, \qquad (12)$$

where $t_1, t_2, \ldots, t_m$ are the times of the m surveys and L_{t_k} is the number of live fish in the stream on survey day t_k. Instead of expressing the daily arrivals on the M days of the season, equation (12) expresses the arrivals for a time period as a function of the deaths occurring during the same time period and the number of live fish on the two survey occasions. In matrix form, equation (12) can be written as

$$\begin{bmatrix} A_{(t_1)} \\ A_{(t_2)} \\ \vdots \\ A_{(t_m)} \end{bmatrix} = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix}\begin{bmatrix} D_{(t_1)} \\ D_{(t_2)} \\ \vdots \\ D_{(t_m)} \end{bmatrix} + \begin{bmatrix} 1 & & & \\ -1 & 1 & & \\ & \ddots & \ddots & \\ & & -1 & 1 \end{bmatrix} \times \begin{bmatrix} p_{lt_1} & & & \\ & p_{lt_2} & & \\ & & \ddots & \\ & & & p_{lt_m} \end{bmatrix}^{-1}\begin{bmatrix} l_{t_1} \\ l_{t_2} \\ \vdots \\ l_{t_m} \end{bmatrix}, \qquad (13)$$

where p_{lt_k} is the probability of detecting a live fish present in the stream on survey occasion t_k and l_{t_k} is the observed number of live fish on day t_k ($k = 1, \ldots, m$). The matrices in equation (13) have dimensions of $m \times m$.

Estimating deaths.—In order to estimate the above $\underset{\sim}{A}_{(t_k)}$, estimates of the number of fish that died during survey period t_k [i.e., $\underset{\sim}{D}_{(t_k)}$] must in turn be estimated from the observed carcass counts (i.e., $\underset{\sim}{c}$). This can be accomplished by modeling the expected values of the observed carcass counts as a function of the mortalities on the survey days ($\underset{\sim}{D}_{t_k}$) where

$$E(\underset{\sim}{c}_{m \times 1}) = \mathbf{P}_{c_{m \times m}} \boldsymbol{\cdot} \boldsymbol{\Delta}_{m \times M} \boldsymbol{\cdot} \boldsymbol{\Xi}_{M \times m} \boldsymbol{\cdot} \underset{\sim}{D}_{t_{m \times 1}}. \tag{14}$$

The detection matrix $\mathbf{P}_c$ is of size $m \times m$ and is a diagonal matrix with the probabilities of detecting and removing a carcass (p_{ct_k}) from the stream at survey times t_k ($k = 1, \ldots, m$) down the diagonal. The other terms in equation (14) are the carcass retention matrix ($\boldsymbol{\Delta}$) and the daily interpolation matrix ($\boldsymbol{\Xi}$), as described below.

Matrix $\boldsymbol{\Delta}$ is the carcass retention matrix of size $m \times M$ that expresses the probabilities of a fish carcass persisting in the stream from death to the time of detection and removal. Let $R_{t_k,j}$ be the probability a fish that died on day j ($j = 1, \ldots, M$) is retained to the survey time t_k ($k = 1, \ldots, m$). The retention matrix is of the form

$$\boldsymbol{\Delta}_{m \times M} = \begin{bmatrix} \Delta_{11} & \Delta_{12} & \Delta_{13} & \ldots & \Delta_{1t_1} & 0 & \ldots & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ \Delta_{21} & \Delta_{22} & \Delta_{23} & \ldots & \Delta_{2t_1} & \Delta_{2t_1+1} & \ldots & \Delta_{2t_2} & 0 & 0 & 0 & \ldots & 0 & 0 & & 0 \\ \Delta_{31} & \Delta_{32} & & \ldots & \Delta_{3t_1} & \Delta_{3t_1+1} & \ldots & \Delta_{3t_2} & \ldots & \Delta_{3t_3} & 0 & \ldots & 0 & 0 & & 0 \\ \vdots & & & & & & & & & & & & & & & \vdots \\ \Delta_{m1} & \Delta_{m2} & & \ldots & \Delta_{mt_1} & \ldots & & \Delta_{mt_2} & \ldots & \Delta_{mt_3} & \ldots & & \Delta_{mt_m} & 0 & & 0 \end{bmatrix}, \tag{15}$$

where the final $M - t_m$ columns consist entirely of zeroes (if $t_m < M$). The nonzero elements of the first row are $\Delta_{11} = R_{t_1,1}$, $\Delta_{12} = R_{t_1,2}, \ldots, \Delta_{1t_1} = R_{t_1,t_1} = 1$. The second row must account for carcass retention and the probability that a carcass was neither detected nor removed in the first survey period. The nonzero elements of the second row are therefore $\Delta_{21} = R_{t_2,1}(1 - p_{ct_1})$, $\Delta_{22} = R_{t_2,2}(1 - p_{ct_1}), \ldots, \Delta_{2,t_1} = R_{t_2,t_1}(1 - p_{ct_1})$, and $\Delta_{2,t_1+1} = R_{t_2,t_1+1}$, $\Delta_{2,t_1+2} = R_{t_2,t_1+2}, \ldots, \Delta_{2t_1} = R_{t_2,t_2} = 1$.

The third row of $\boldsymbol{\Delta}$ takes into account carcass retention rates and the probabilities that a carcass was not detected and removed during survey periods 1 and 2. For example, $\Delta_{31} = R_{t_3,1}(1 - p_{ct_1})(1 - p_{ct_2}), \ldots, \Delta_{3t_1} = R_{t_3,t_1}(1 - p_{ct_1})(1 - p_{ct_2})$, $\Delta_{3,t_1+1} = R_{t_3,t_1+1}(1 - p_{ct_2}), \ldots, \Delta_{3,t_2} = R_{t_3,t_2}\ (1 - p_{ct_2})$, and $\Delta_{3,t_2+1} = R_{t_3,t_2+1}, \ldots, \Delta_{3t_3} = R_{t_3,t_3} = 1$. If retention rates and detection probabilities are assumed equal to 1, then $\boldsymbol{\Delta}$ simplifies to a matrix of the form

$$\Delta_{m \times M} = \begin{bmatrix} 1 & 1 & \ldots & 1 & 0 & \ldots & 0 & 0 & \ldots & 0 & \ldots & 0 & \ldots & 0 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 & 1 & \ldots & 1 & 0 & \ldots & 0 & \ldots & 0 & & 0 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 & 1 & \ldots & 1 & \ldots & 0 & & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & & & & \ldots & & & & & & & & & & & \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 & 0 & \ldots & 0 & \ldots & 1 & \ldots & 1 & 0 & \ldots & 0 \end{bmatrix}.$$

The third term in equation (14), $\boldsymbol{\Xi}$, is the daily interpolation model. This $M \times m$ matrix estimates by interpolation the daily deaths between the survey dates across the season. Here, a key assumption is that the mortalities for the days between consecutive survey occasions can be approximated as a linear interpolation of the deaths on the two survey occasions (i.e., t_{k-1} and t_k) for $k = 1, \ldots, m$. In this case,

$$\hat{D}_j = \frac{t_k - j}{t_k - t_{k-1}} D_{t_{k-1}} + \frac{j - t_{k-1}}{t_k - t_{k-1}} D_{t_k}$$

for days j when $t_{k-1} \le j \le t_k$. The linear interpolation is all that is available in the absence of direct observations, and the assumption will likely degrade as the duration between surveys (i.e., $t_k - t_{k-1}$) and increases. Subsequent computer simulations will assess the effectiveness of the linear interpolation out to 7 d between survey occasions. If the spawning season continues past the final survey day ($t_m < M$), we assumed zero deaths on the last season day ($D_M = 0$).

Matrix $\boldsymbol{\Xi}$ is of the form

$$\boldsymbol{\Xi}_{M\times m} =$$

$$\begin{bmatrix}
\Xi_{11} & 0 & \cdots & 0 & 0 \\
\Xi_{21} & 0 & & \vdots & \vdots \\
\vdots & \vdots & & & \\
\Xi_{t_1 1} & 0 & & & \\
\Xi_{t_1+1,1} & \Xi_{t_1+1,2} & & & \\
\vdots & \vdots & & & \\
\Xi_{t_2 1} & \Xi_{t_2 2} & & & \\
0 & \Xi_{t_2+1,2} & & & \\
\vdots & \vdots & & & \\
 & \Xi_{t_3 2} & & & \\
 & 0 & & & \\
 & \vdots & & \vdots & \vdots \\
 & & & \Xi_{t_{m-1},m-1} & 0 \\
 & & & \Xi_{t_{m-1}+1,m-1} & \Xi_{t_{m-1}+1,m} \\
 & & & \vdots & \vdots \\
 & & & \Xi_{t_m,m-1} & \Xi_{t_m,m} \\
 & & \cdots & 0 & \Xi_{t_m+1,m} \\
\vdots & \vdots & & & \vdots \\
0 & 0 & \cdots & 0 & \Xi_{Mm}
\end{bmatrix}. \qquad (16)$$

The elements of the first column of $\boldsymbol{\Xi}$ are

$$\Xi_{11} = \frac{1}{t_1}, \quad \Xi_{21} = \frac{2}{t_1}, \ \ldots, \ \Xi_{t_1 1} = 1,$$

$$\Xi_{t_1+1,1} = \frac{t_2 - (t_1+1)}{t_2 - t_1},$$

$$\Xi_{t_1+2,1} = \frac{t_2 - (t_1+2)}{t_2 - t_1}, \ \ldots, \ \Xi_{t_2 1} = 0, \quad \text{and}$$

$$\Xi_{j1} = 0 \quad \text{for } j > t_2.$$

The elements of the second column are computed as

$$\Xi_{j2} = 0 \quad \text{for } j \le t_1 \quad \text{or} \quad j > t_3,$$

$$\Xi_{t_1+1,2} = \frac{1}{t_2 - t_1}, \quad \Xi_{t_1+2,2} = \frac{2}{t_2 - t_1}, \ \ldots,$$

$$\Xi_{t_2 2} = 1, \quad \Xi_{t_2+1,2} = \frac{t_3 - (t_2+1)}{t_3 - t_2},$$

$$\Xi_{t_2+2,2} = \frac{t_3 - (t_2+2)}{t_3 - t_2}, \ \ldots, \ \Xi_{t_3 2} = 0.$$

The entries of column 3 to column $m-1$ are computed analogously to that of column 2. The last column (column m) has nonzero elements:

$$\Xi_{t_{m-1}+1,m} = \frac{1}{t_m - t_{m-1}},$$

$$\Xi_{t_{m-1}+2,m} = \frac{2}{t_m - t_{m-1}}, \ \ldots, \ \Xi_{t_m m} = 1,$$

and if $t_m < M$,

$$\Xi_{t_m+1,m} = \frac{M - (t_m+1)}{M - t_m},$$

$$\Xi_{t_{m+2},m} = \frac{M - (t_m+2)}{M - t_m}, \ \ldots, \ \Xi_{Mm} = 0.$$

From equation (14), the estimate of the number of fish that died on the m survey days can be calculated as

$$\hat{\underset{\sim}{D}}_{t_k} = [\mathbf{P}_c \boldsymbol{\cdot} \boldsymbol{\Delta} \boldsymbol{\cdot} \boldsymbol{\Xi}]^{-1}_{m\times m} \underset{\sim}{c}_{m\times 1} \qquad (17)$$

using matrix inversion. Finally, estimates of the daily mortalities can be calculated using, once again, the interpolation matrix, where

$$\hat{\underset{\sim}{D}}_j = \boldsymbol{\Xi}_{M\times m} \boldsymbol{\cdot} \hat{\underset{\sim}{D}}_{t_k}. \qquad (18)$$

The estimates of mortalities during the survey periods, $\hat{\underset{\sim}{D}}_{(t_k)}$, come from the estimated daily mortalities, $\hat{\underset{\sim}{D}}_j$.

The variance estimator for $\hat{T}_3$ (equation 6) is presented in Appendix 1. Computer code written in R (R Development Core Team 2008) to estimate T and its variance is given in Appendix 2 and is available in the online version of this article at afs.allenpress.com.

Finally, it should be noted that escapement can be estimated from the total number of spawners that died in the stream. Hence, an estimate of total escapement ($\hat{N}$) can be readily calculated as

$$\hat{N} = 1'\hat{\underset{\sim}{D}}_j = 1'\boldsymbol{\Xi}[\mathbf{P}_c \boldsymbol{\cdot} \boldsymbol{\Delta} \boldsymbol{\cdot} \boldsymbol{\Xi}]^{-1} \underset{\sim}{c}, \qquad (19)$$

based on equations (17) and (18), assuming the number of daily mortalities can be linearly interpolated between survey occasions. The R code to compute $\hat{N}$ and its variance estimator are also given in Appendix 2.

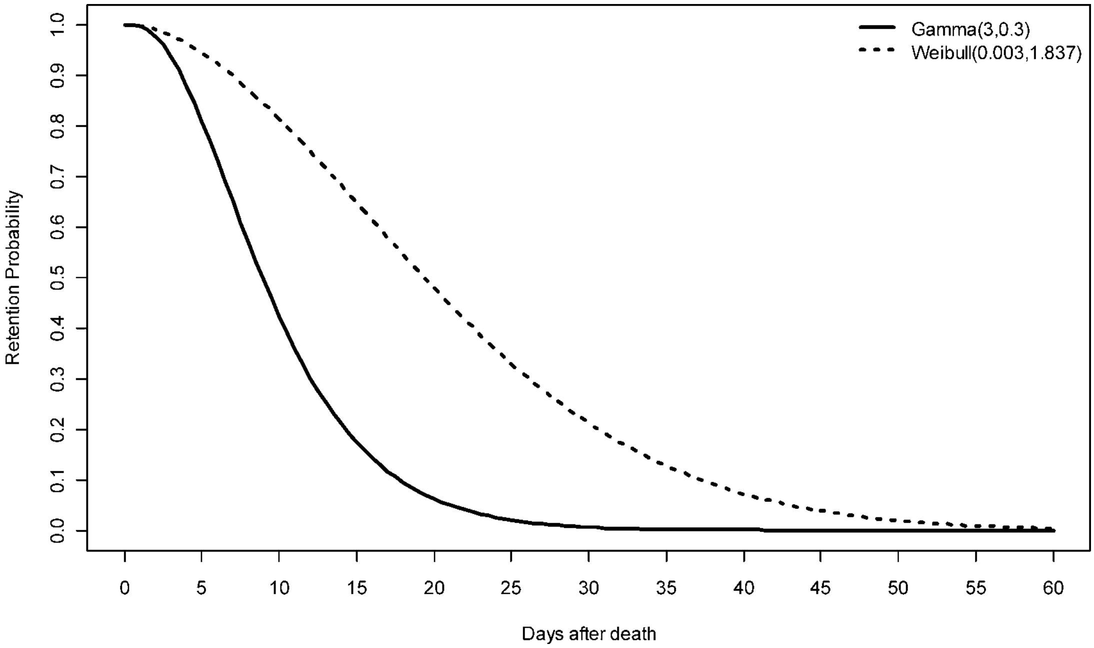

FIGURE 3.—Retention curves used to simulate the retention rates of salmonid spawner carcasses in a stream based on a gamma (3, 0.3) distribution or a Weibull (0.003, 1.837) distribution. The Weibull distribution describes chum salmon in Perry Creek, Washington (Liao 1994).

Simulation Studies

An initial set of simulations was performed to compare estimation methods under ideal circumstances of 100% detection and carcass retention. These simulations were performed to first assess the alterative estimators of stream residence time independent of the problems of imperfect detection. Subsequent simulations were conducted with less-than-perfect detection probabilities and carcass retention rates to more realistically represent actual survey conditions. Carcass retention was simulated with a Weibull (location parameter = 0.003, scale parameter = 1.837) distribution fit to data from chum salmon at Perry Creek. This distribution produced an expected carcass retention time of 21 d. Using a gamma (shape parameter = 3, scale parameter = 0.3) distribution, carcass retention was also modeled with a shorter expected retention time of 10 d (Figure 3). For the chum salmon population at Perry Creek, Liao (1994) estimated an average carcass detection probability of 0.937 ($\widehat{\text{SE}}$ = 0.019). To more rigorously assess the behavior of the expectation-to-expectation method, detection probabilities were reduced to $P_l = 0.85$ for live fish and $P_l = 0.80$ for carcasses.

Two alternative arrival distributions of fish were considered. A left-skewed arrival distribution was generated using a beta ($\alpha = 4$, $\beta = 2$) distribution (Figure 4a), while a right-skewed arrival distribution was generated using a beta (3, 7) distribution (Figure 4b). The actual number of arrivals on each of the M days of the season was stochastically determined for each computer run. In addition, two alternative survival functions were used to generate carcass counts. A gamma (5, 0.5) distribution was used to generate values of stream residence time with an average life expectancy of 10 d, while a gamma (6, 0.4) distribution generated stream residence times with a life expectancy of 15 d (Figure 5).

Three alternative sampling schemes were considered. For daily surveys, 100% of the spawning season was canvassed. When systematic sampling was used, either every fourth day (25%) or every seventh day (14%) was canvassed. Start times for the systematic sampling were randomly selected during either the first 4 d or the first 7 d of escapement, respectively.

For all scenarios, 1,000 simulations were performed in order to estimate sampling error and bias. Simulations were performed with total escapement fixed at 1,000 and spawning seasons of 30 d and more. Performance of estimators was based on average value and mean squared error, which accounts for both sampling variance and bias. Sampling error estimated by the model was compared with the empirical variance among the replicate Monte Carlo simulations.

Field Sampling

A total of 82 recently dead chum salmon were tagged, and their positions were flagged. Their retentions were monitored daily during 21 November–3 December 1992 and thereafter were monitored every 2 d through 6 January 1993. Fate times (i.e., loss times) were fit to a two-parameter Weibull distribution using maximum likelihood estimation.

Two independent survey crews were used to estimate detection probabilities of carcasses. The first crew tagged all observed carcasses (90) with a spaghetti tag, followed 1 h later by a second survey crew that recorded all observed tagged (82) and nontagged (3) carcasses. A two-sample Lincoln–Petersen model (Seber 1982:59–114) was used to establish detection rates and overall carcass abundance. Maximum likelihood estimation and likelihood ratio tests were used to estimate a common detection rate across survey crews.

Results

The results begin with an annotated example of the calculations for the expectation-to-expectation estimator of T. Simulation results comparing the statistical performances of the three alternative estimators of T follow.

Numerical Example

For simplicity of illustration, a 6-d season (i.e., $M = 6$) will be assumed with stream surveys on days $t_1 = 2$, $t_2 = 4$, and $t_3 = 6$ ($m = 3$). The number of carcasses observed during the three surveys were $\underset{\sim}{c}' = (10, 20, 50)$ and live counts were $\underset{\sim}{l}' = (48, 45, 0)$. The detection probabilities for carcasses and live fish will be assumed constant at $p_{lt_k} = p_{ct_k} = 0.80$ for all $k = 1, 2, 3$. The carcass retention rates are assumed to be $R_{t_k,t_k} = 1.0$, $R_{t_k,t_k-1} = 0.80$, $R_{t_k,t_k-2} = 0.60$, $R_{t_k,t_k-3} = 0.50$, $R_{t_k,t_k-4} = 0.40$, and $R_{t_k,t_k-5} = 0.30$ for $k \leq 3$.

The first task was to estimate the numbers of deaths. The carcass detection matrix was of the form

$$\mathbf{P}_c = \begin{bmatrix} 0.8 & 0 & 0 \\ 0 & 0.8 & 0 \\ 0 & 0 & 0.8 \end{bmatrix},$$

and from equation (15), the retention matrix was

$$\mathbf{\Delta} = \begin{bmatrix} 0.8 & 1 & 0 & 0 & 0 & 0 \\ 0.1 & 0.12 & 0.8 & 1 & 0 & 0 \\ 0.012 & 0.016 & 0.10 & 0.12 & 0.8 & 1 \end{bmatrix}.$$

Because the surveys were conducted every other day, the linear interpolation matrix was simply

$$\mathbf{\Xi} = \begin{bmatrix} 0.5 & 0 & 0 \\ 1 & 0 & 0 \\ 0.5 & 0.5 & 0 \\ 0 & 1 & 0 \\ 0 & 0.5 & 0.5 \\ 0 & 0 & 1 \end{bmatrix}.$$

The estimate of mortalities on the three survey days was then

$$\hat{\underset{\sim}{D}}_{t_k} = [\mathbf{P}_c\mathbf{\Delta}\Xi]^{-1}\underset{\sim}{c} = \begin{bmatrix} 1.12 & 0 & 0 \\ 0.456 & 1.12 & 0 \\ 0.058 & 0.456 & 1.12 \end{bmatrix}^{-1} \begin{bmatrix} 10 \\ 20 \\ 50 \end{bmatrix} = \begin{bmatrix} 8.9 \\ 14.2 \\ 38.4 \end{bmatrix}.$$

Interpolating for the days not surveyed, the daily estimates of mortality were as follows:

$$\Xi\hat{\underset{\sim}{D}}_{t_k} = \hat{\underset{\sim}{D}}_j = \begin{bmatrix} 0.5 & 0 & 0 \\ 1 & 0 & 0 \\ 0.5 & 0.5 & 0 \\ 0 & 1 & 0 \\ 0 & 0.5 & 0.5 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 8.9 \\ 14.2 \\ 38.4 \end{bmatrix} = \begin{bmatrix} 4.45 \\ 8.90 \\ 11.55 \\ 14.20 \\ 26.30 \\ 38.40 \end{bmatrix}.$$

The cumulative death curve, based on the daily $\hat{D}_{t_k}$ for the days of the actual survey, was then estimated from equation (7) as follows:

$$\hat{F}_d = \left(\frac{4.45 + 8.90}{103.80} = 0.1286, \quad \frac{4.45 + 8.90 + 11.55 + 14.20}{103.80} = 0.3767, 1.0\right). \tag{20}$$

Because the detection probability was 0.80 for the live fish, the estimated number of live fish present on the three survey occurrences was estimated to be

$$\hat{\underset{\sim}{\mathrm{L}}}_{(t_k)} = \begin{bmatrix} 0.8 & & \\ & 0.8 & \\ & & 0.8 \end{bmatrix}^{-1} \begin{bmatrix} 48 \\ 45 \\ 0 \end{bmatrix} = \begin{bmatrix} 60.0 \\ 56.3 \\ 0 \end{bmatrix}.$$

Using equation (13), the number of new arrivals between the survey occurrences was then estimated as follows:

$$\hat{\underset{\sim}{\mathrm{A}}}_{(t_k)} = \begin{bmatrix} 13.35 \\ 25.75 \\ 64.70 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 60.0 \\ 56.3 \\ 0 \end{bmatrix} = \begin{bmatrix} 73.35 \\ 22.05 \\ 8.40 \end{bmatrix},$$

where $13.35 = 4.45 + 8.90$; $25.75 = 11.55 + 14.20$; and so on. These arrival estimates were then used to produce the estimated cumulative arrival distribution, where

$$\hat{F}_a = \left(\frac{73.35}{103.8} = 0.7066, \frac{73.35 + 22.05}{103.8} = 0.9191, 1.0\right). \tag{21}$$

Using the two estimated cumulative distributions (i.e.,

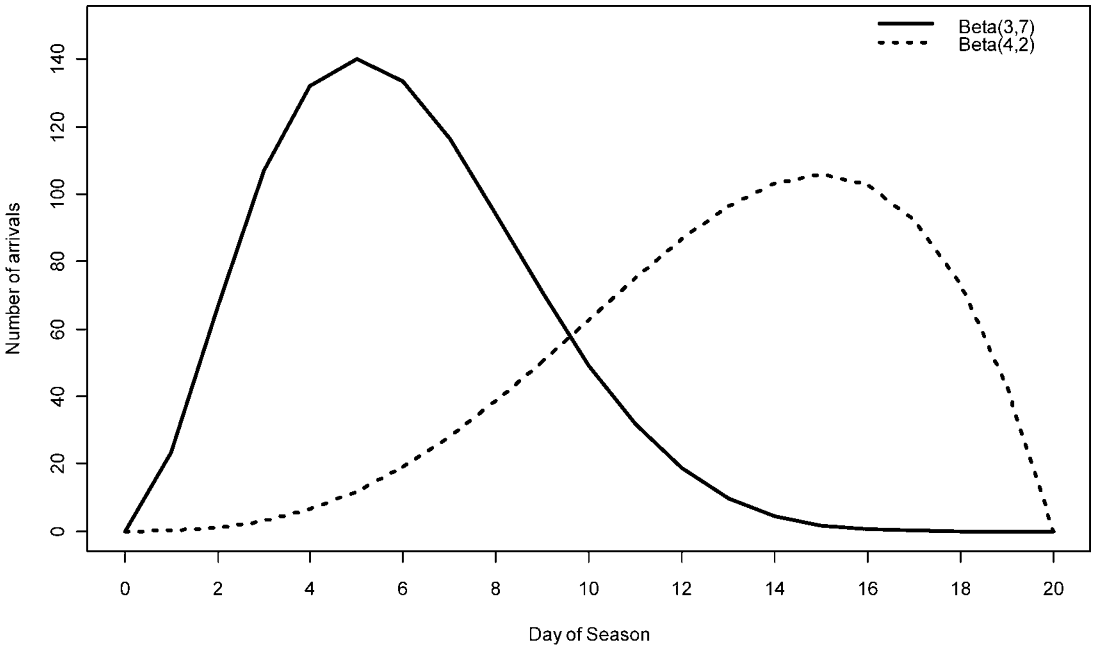

FIGURE 4.—Beta distributions used to simulate the arrival process of salmonid spawners based on either a beta (3, 7) distribution or a beta (4, 2) distribution.

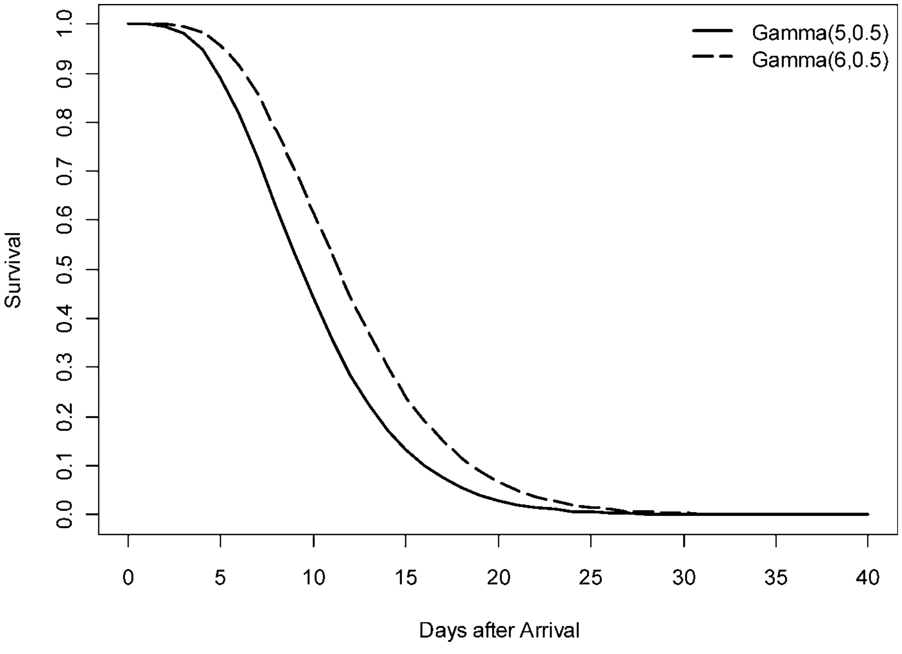

FIGURE 5.—Gamma survivorship functions used to simulate salmonid stream residence times based on either a gamma (6, 0.5) distribution or a gamma (5, 0.5) distribution.

TABLE 1.—Results of Monte Carlo simulations used to evaluate the performance of peak-to-peak, median-to-median, and expectation-to-expectation estimators of salmonid stream residence time (d; $\hat{T}_1$, $\hat{T}_2$, $\hat{T}_3$, respectively) under ideal conditions with no carcass loss during the season, 100% detection rates of all live and dead fish, and daily surveys (MSE = mean squared error).

	Stream residence time = gamma (5, 0.5)[a]						Stream residence time = gamma (6, 0.4)[b]					
	Peak-to-peak		Median-to-median		Expectation-to-expectation		Peak-to-peak		Median-to-median		Expectation-to-expectation	
Arrival model and survey scheme	$\hat{T}_1$	MSE	$\hat{T}_2$	MSE	$\hat{T}_3$	MSE	$\hat{T}_1$	MSE	$\hat{T}_2$	MSE	$\hat{T}_3$	MSE
Arrival = beta (4, 2) distribution; daily sampling	5.17	27.11	4.61	29.35	10.07	0.21	7.92	52.61	6.99	64.90	15.01	0.019
Arrival = beta (3, 7) distribution; daily sampling	4.67	29.97	4.61	29.41	9.91	0.034	9.33	35.69	6.99	64.78	14.94	0.020

[a] The expected stream residence time for the gamma (5, 0.5) distribution model is 10 d.

[b] The expected stream residence time for the gamma (6, 0.4) distribution model is 15 d.

equations 20 and 21), stream life was estimated to be

$$
\begin{aligned}
\hat{T}_3 = {} & \frac{1}{2}\{(2-0)[(0.7066+0)-(0.1286+0)]\} \\
& + \frac{1}{2}\{(4-2)[(0.9191+0.7066) \\
& \qquad -(0.3767+0.1286)]\} \\
& + \frac{1}{2}\{(6-4)[(1+0.9191)-(1+0.3767)]\} \\
= {} & 2.24 \text{ (days)}.
\end{aligned}
$$

The estimated SE for $\hat{T}_3$ was 0.16 (Appendices 1, 2). Finally, an estimate of total escapement is the sum of the estimated daily deaths,

$$\hat{N} = 1'\underset{\sim}{\hat{D}}_j = 103.8$$

with an estimated SE of 5.48.

Simulation Results

In an ideal situation consisting of 100% carcass retention, 100% detection rates for both live and dead fish, and daily surveys throughout the spawning season, the peak-to-peak and median-to-median estimators have about a 50% negative bias (Table 1). This magnitude of bias was observed for all arrival and lifetime distributions simulated. In practice, this would mean estimates of total escapement based on the AUC method could be overestimated by a factor of 2. The expectation-to-expectation method was unbiased as anticipated from its development (Table 1).

Arrival distributions in this study had little or no appreciable effect on estimator performance. Consequently, only results for the beta (4, 2) arrival distribution (Figure 4a) will be subsequently reported. For a complete analysis, see Liao (1994). For the peak-to-peak method, as the sampling fraction for systematic sampling decreased, the estimates of stream residence time (T) became increasingly more negatively biased (Table 2). With systematic sampling of one stream visit every 7 d, the peak-to-peak method underestimated T by approximately 90% (Table 2). For the median-to-median method, accuracy interestingly increased as sampling fraction declined for systematic sampling. However, the degree of negative bias was still substantial (i.e., $\geq$ 25%; Table 2). In the case of the expectation-to-expectation method, the estimator remained essentially unbiased regardless of the sampling fractions simulated (Table 2).

Under less-than-perfect circumstances (Table 3) with capture and retention rates less than 1, the expectation-to-expectation mortality remained relatively unbiased, along with its variance estimator. Some degradation in performance was observed with systematic sampling of 1 d/week when the arrival distribution [i.e., beta (3,7)] was essentially only 2 weeks long. In this circumstance, performance was limited by the small number of survey events across the spawning season.

Discussion

Knowledge of stream residence time is critical to the application of the AUC method for estimating escapement. However, the natural variability in average stream residence time (T) requires that estimates of T be both time specific and site specific (Perrin and Irvine 1990). While mark–recapture methods can be used to estimate T (Lady and Skalski 1998), their complexity and costs preclude their wide application. Instead, data from live and dead counts of fish can be collected at numerous streams annually with reasonable costs.

The two early methods for estimating T from live and dead counts (methods summarized by Perrin and Irvine 1990) have substantial negative bias even under the best of circumstances (Tables 1, 2). Computation of the expectation-to-expectation estimator is more complex that the earlier methods but provides a means of unbiasedly estimating salmonid stream residence time.

TABLE 2.—Results of Monte Carlo simulations used to evaluate the performance of peak-to-peak, median-to-median, and expectation-to-expectation estimators of salmonid stream residence time (d; $\hat{T}_1$, $\hat{T}_2$, $\hat{T}_3$, respectively) based on live and dead fish counts, assuming two alternative gamma distributions for stream residence time and detection probabilities of 1 (MSE = mean squared error). Only the results for the beta (4, 2) arrival model are shown.

	Peak-to-peak		Median-to-median		Expectation-to-expectation	
Sampling scheme	$\hat{T}_1$	MSE	$\hat{T}_2$	MSE	$\hat{T}_3$	MSE
Gamma (5, 0.5) distribution[a]						
Daily	3.83	40.40	4.56	29.90	10.10	0.020
Systematic: every fourth day	1.60	94.47	6.09	20.44	9.97	0.035
Systematic: every seventh day	1.08	92.82	7.54	7.09	9.98	0.065
Gamma (6, 0.4) distribution[b]						
Daily	4.75	29.71	6.94	65.68	15.03	0.021
Systematic: every fourth day	2.38	212.72	8.51	56.24	14.94	0.058
Systematic: every seventh day	1.51	212.54	9.96	29.78	14.85	0.106

[a] The expected stream residence time for this model is 10 d.

[b] The expected stream residence time for this model is 15 d.

This new method also accounts for imperfect detection probabilities and carcass retention rates less than 1.

One means of estimating carcass retention has been to periodically revisit dead fish that were flagged at the time of death. This information on retention and loss is then analyzed using a failure time curve or survivorship curve to estimate the probability of loss as a function of time (Hosmer and Lemeshow 1999). This approach works best when the primary sources of loss are decay and scavengers. When there is the possibility that carcasses may be washed out of the spawning area, a mark–recapture method such as the single release–recapture model of Skalski (1998), as a special case of the Cormack (1964) model, can be used to estimate carcass retention probabilities as well as detection probabilities. One rapid way of estimating detection probabilities has been to have one crew mark carcasses, followed shortly behind by another survey crew that records the numbers of fish marked. A binomial sampling model is then used to estimate the probability of detection assuming a closed population. Liao (1994) used a two-sample Lincoln–Petersen method to estimate detection probabilities and found it to be more robust than the binomial model.

Acknowledgments

Funding for this work came from the Pacific Northwest region's electrical ratepayers through the Columbia River Fish and Wildlife Program administered by the Bonneville Power Administration through Project Number 1989-107-00.

TABLE 3.—Results of Monte Carlo simulations examining the performance of the expectation-to-expectation estimator of salmonid stream residence time ($\hat{T}_3$), where the carcass detection probability (p_c) is 0.80, the live fish detection probability (p_l) is 0.85, and carcass retention follows a gamma (3, 0.3) distribution with an expectation of 10 d or a Weibull distribution (0.003, 1.831) with an expectation of 21 d. Salmon arrival distributions were allowed to be either beta (4, 2) or beta (3, 7) (see Figure 3). Stream residence time was gamma (6, 0.4) distributed with an expectation (E[SRT]) of 15 d or gamma (5, 0.5) distributed with an expectation of 10 d (see Figure 4).

			Carcass retention = gamma (3, 0.3)			Carcass retention = Weibull (0.003, 1.837)		
Arrival model	Survey scheme	E(SRT)	$\overline{\hat{T}}_3$	$s^2_{\hat{T}_3}$	$\widehat{\text{Var}}(\hat{T}_3)$	$\overline{\hat{T}}_3$	$s^2_{\hat{T}_3}$	$\widehat{\text{Var}}(\hat{T}_3)$
Beta (4, 2)	Daily	15 d	14.96	0.0037	0.0040	14.98	0.0041	0.0039
	Systematic: every fourth day	15 d	14.80	0.0302	0.0286	14.95	0.0222	0.0196
	Systematic: every seventh day	15 d	14.57	0.0867	0.0656	14.83	0.0604	0.0394
Beta (3, 7)	Daily	15 d	15.05	0.0038	0.0039	14.89	0.0036	0.0039
	Systematic: every fourth day	15 d	14.57	0.0768	0.0280	14.70	0.0601	0.0191
	Systematic: every seventh day	15 d	13.63	0.5812	0.0613	13.86	0.5808	0.0356
Beta (4, 2)	Daily	10 d	9.99	0.0025	0.0026	10.01	0.0026	0.0025
	Systematic: every fourth day	10 d	9.86	0.0195	0.0159	9.54	0.0146	0.0120
	Systematic: every seventh day	10 d	9.69	0.0602	0.0345	9.88	0.0455	0.0232
Beta (3, 7)	Daily	10 d	9.99	0.0024	0.0026	10.01	0.0026	0.0026
	Systematic: every fourth day	10 d	9.65	0.0584	0.0155	9.73	0.0537	0.0115
	Systematic: every seventh day	10 d	8.76	0.5354	0.0307	8.92	0.5496	0.0195

References

Ames, J. J. 1984. Puget Sound chum salmon escapement estimates using spawner curve methodology. Canadian Technical Report of Fisheries and Aquatic Sciences 1326:135–148.

Bue, B. G., S. M. Fried, S. Sharr, D. G. Sharp, J. Wilcock, and H. J. Geiger. 1998. Estimating salmon escapement using area-under-the-curve, aerial observer efficiency, and streamlife estimates: the Prince William Sound pink salmon example. North Pacific Anadromous Fisheries Commission Bulletin 1:240–250.

Cormack, R. M. 1964. Estimates of survival from the sighting of marked animals. Biometrika 51:429–438.

Cousens, N., G. Thomas, C. Swann, and M. Healey. 1982. A review of salmon escapement estimation techniques. Canadian Technical Report of Fisheries and Aquatic Sciences 1101:122.

English, K. K., R. C. Bocking, and J. R. Irvine. 1992 A robust procedure for estimating salmon escapement based on the area-under-the-curve method. Canadian Journal of Fisheries and Aquatic Sciences 49:1982–1989.

Fukushima, M., and W. W. Smoker. 1997. Determinants of stream life, spawning efficiency, and spawning habitat in pink salmon in the Auke Lake system, Alaska. Canadian Journal of Fisheries and Aquatic Sciences 54:96–104.

Hill, R. A. 1997. Optimizing aerial count frequency for the area-under-the-curve method of estimating escapement. North American Journal of Fisheries Management 17:461–466.

Hosmer, D. W. Jr., and S. Lemeshow. 1999. Applied survival analysis: regression modeling of time to event data. Wiley, New York.

Killick, S. R. 1955. The chronological order of Fraser River salmon during migration, spawning, and death. International Pacific Salmon Fisheries Commission Bulletin VII.

Knudsen, E. E. 2000. Managing Pacific salmon escapements: the gaps between theory and reality. Pages 237–272 *in* E. E. Knudsen, C. R. Steward, D. D. MacDonald, J. E. Williams, and D. W. Reiser, editors. Sustainable fisheries management. Lewis Publishers, Boca Raton, Florida.

Lady, J., and J. R. Skalski. 1998. Estimators of stream residence time of Pacific salmon (*Oncorhynchus* spp.) based on release-recapture data. Canadian Journal of Fisheries and Aquatic Sciences 55:2580–2587.

Lewis, A. F. J. 1987. The effects of breeding completion and abiotic factors on the behaviour and reproductive success of female chum salmon (*Oncorhynchus keta*). Master's thesis. Simon Fraser University, Burnaby, British Columbia.

Liao, S. 1994. Statistical models for estimating salmon escapement and stream residence time based on stream survey data. Doctoral dissertation. University of Washington, Seattle.

Manske, M., and C. J. Schwarz. 2000. Estimates of stream residence time and escapement based on capture-recapture data. Canadian Journal of Fisheries and Aquatic Sciences 57:241–246.

Mood, A. M., F. A. Graybill, and D. C. Boes. 1974. Introduction to the theory of statistics, 3rd edition. McGraw-Hill, New York.

Perrin, C. J., and J. R. Irvine. 1990. A review of survey life estimates as they apply to the area under the curve method of estimating spawning escapement of Pacific salmon. Canadian Technical Report of Fisheries and Aquatic Sciences 1733.

R Development Core Team. 2008. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available: www.R-project.org. (November 2008).

Seber, G. A. F. 1982. The estimation of animal abundance. MacMillan, New York.

Skalski, J. R. 1998. Estimating season-wide survival rates of outmigrating smolt in the Snake River, Washington. Canadian Journal of Fisheries and Aquatic Sciences 55:761–769.

Thomason, G. J., and J. D. Jones. 1984. Southeastern Alaska pink salmon (*Oncorhynchus gorbuscha*) stream life studies, 1983. Alaska Department of Fish and Game, Information Leaflet Number 236, Juneau.

Appendix 1: Variance Estimator for Stream Residence Time Estimator $\hat{T}_3$

The variance estimator for $\hat{T}_3$ is a function of the survey times ($\underset{\sim}{t}_k$), the observed numbers of live fish ($\underset{\sim}{l}_{t_k}$), the estimated numbers of dead fish on each day of the season ($\hat{\underset{\sim}{D}}_j$), the detection probabilities for live fish ($\underset{\sim}{p}_{lt_k}$) and for carcasses ($\underset{\sim}{p}_{ct_k}$), and the carcass retention rates ($\underset{\sim}{R}_{t_k,j}$). If detection of live fish is independent of detection of dead fish, then the variance is estimated as follows:

$$\widehat{\mathrm{Var}}(\hat{T}_3) = \hat{T}_3^2(V_1 + V_2),$$

where

$$V_1 = \frac{\displaystyle\sum_{k=1}^{m}\left[\frac{(t_{k+1} - t_{k-1})^2 l_{t_k}(1 - p_{lt_k})}{p_{lt_k}^2}\right]}{\left\{\dfrac{\displaystyle\sum_{k=1}^{m}[(t_{k+1} - t_{k-1})l_{t_k}]}{p_{lt_k}}\right\}^2}$$

(with $t_0 = 0$ and $t_{m+1} = M$) and

$$V_2 = \frac{W[\mathrm{diag}(\mathbf{P}_c \boldsymbol{\Delta} \underset{\sim}{\hat{D}}_j) - \mathbf{P}_c \boldsymbol{\Delta} \mathrm{diag}(\underset{\sim}{\hat{D}}_j)(\mathbf{P}_c \boldsymbol{\Delta})']W'}{(1'\underset{\sim}{\hat{D}}_j)^2}$$

with

$$W = 1'\boldsymbol{\Xi}[\boldsymbol{P_c \Delta \Xi}]^{-1}.$$

The matrices $\mathbf{P}_c$, $\boldsymbol{\Delta}$, and $\boldsymbol{\Xi}$ are defined in the body of the article. The notation diag($\underset{\sim}{x}$) indicates the diagonal matrix for which the diagonal is the vector $\underset{\sim}{x}$.

The variance estimator for estimated escapement $\hat{N}$ is

$$\widehat{\mathrm{Var}}(\hat{N}) = (1'\underset{\sim}{\hat{D}}_j)^2 V_2 = W[\mathrm{diag}(\mathbf{P}_c \boldsymbol{\Delta} \underset{\sim}{\hat{D}}_j) - \mathbf{P}_c \boldsymbol{\Delta} \mathrm{diag}(\underset{\sim}{\hat{D}}_j)(\mathbf{P}_c \boldsymbol{\Delta})']W'.$$

Appendix 2: R Code

```
## This file contains R code to implement the expectation-to-expectation estimator (T3) of stream
## residence time from Skalski, Liao, and Buchanan (2009), “Performance of three alternative estimators
## of stream residence time based on live and dead counts of salmonids", North American Journal of
## Fisheries Management.

## Code is presented for both the point estimator and the variance estimator for both stream residence
## time (T3) and escapement (N).

## Functions and an example of usage are given in this file.

##########################################################################################
## Functions:
##     1. res.time = calculates point estimate and standard error estimate of stream residence time
##                   and escapement.
##     2. res.time.pt = calculates point estimate of stream residence time and escapement
##                      (used in res.time)

res.time=function(times,carcass.count,live.count,R.matrix,P.carcass,P.live,se=T)
{
# Returns point estimate and standard error (or variance) estimate of stream residence time (T3)
# and escapement (N).
#
# Arguments:
# times = vector of survey times (in days); length = m
# carcass.count = vector of observed carcass counts at survey times; length = m
# live.count = vector of observed live counts at survey times; length = m
# R.matrix = matrix of carcass retention probabilities R[tk,j]; m rows, M columns
#     - rows = survey days; columns = season days
#     - entry R[tk,j] = probability that the carcass of a fish that died on day j remains in the stream
#                 on day tk, in the absence of detection and removal by surveyors;
#                 = 1 for j = tk; = 0 for j > tk
# P.carcass = vector of carcass detection probabilities; length = m
# P.live = vector of live detection probabilities; length = m
# se = logical; if T (default), return standard error; if F, return variance
#
# Uses formula: var.T=T.hat^2*(var.F/F.hat^2 + var.N/N.hat^2)
# where
```

```
# F.hat = estimated fish days
# var.F = estimated variance of F.hat
# var.N = estimated variance of escapement estimate (N.hat)
# Equivalent to: var.T=T.hat^2*(V1 + V2), where V1 = var.F/F.hat^2 and V2 = var.N/N.hat^2

M=ncol(R.matrix)   # length of season (in days)
m=length(times)    # number of survey days (m <= M)

# Define matrices and estimates from point estimator
res.pt=res.time.pt(times,carcass.count,live.count,R.matrix,P.carcass,P.live) # point estimate
Delta=res.pt$Delta          # matrix of carcass retention rates and detection rates
Xi=res.pt$Xi                # transition matrix from D.survey to D.season
D.season=res.pt$D.season # estimated number dying on season days
N.hat=res.pt$N.hat          # estimated escapement
T.hat=res.pt$T.hat          # estimated stream residence time

## Get point estimate and variance estimate of fish days (F)
t.tmp=c(times[-1],M)-c(0,times[-m]) # vector = times[i+1]-times[i-1]
F.hat=0.5*sum(t.tmp*live.count/P.live)
var.F=0.25*sum(t.tmp^2*live.count*(1-P.live)/P.live^2)
V1=var.F/F.hat^2

## Get variance estimate of N.hat, using var.N=W%*%var.c%*%t(W)
P=diag(P.carcass)
W=rep(1,M)%*%Xi%*%solve(P%*%Delta%*%Xi) # matrix with 1 row, m columns
var.c=diag(c(P%*%Delta%*%D.season)) - P%*%Delta%*%diag(c(D.season))%*%t(P%*%Delta)
var.N=W%*%var.c%*%t(W)
V2=var.N/N.hat^2

var.T=T.hat^2*(V1+V2)

if(se) {
 out=c(T.hat,sqrt(var.T),N.hat,sqrt(var.N)); names(out)=c("T","se.T","N","se.N")
 } else {
   out=c(T.hat,var.T,N.hat,var.N); names(out)=c("T","var.T","N","var.N")
   }
 return(out)
}


res.time.pt=function(t.vec,c.vec,l.vec,R.mat,P.dead.vec,P.live.vec)
{
# Estimates stream residence time, T3 (point estimate).
#
# Arguments:
# t.vec = vector of survey times (in days); length = m
# c.vec = vector of observed carcass counts at survey times; length = m
# l.vec = vector of observed live counts at survey times; length = m
# R.mat = matrix of carcass retention probabilities R[tk,j]; m rows, M columns
#    - rows = survey days; columns = season days
#    - entry R[tk,j] = probability that the carcass of a fish that died on day j remains in the stream
#                    on day tk, in the absence of detection and removal by surveyors;
#                    = 1 for j = tk; = 0 for j > tk
# P.dead.vec = vector of carcass detection probabilities; length = m
```

```
# P.live.vec = vector of live detection probabilities; length = m
#
# Returns list containing defined parameters:
# Delta = function of R.mat and P.dead.vec; matrix, m rows, M columns
# Xi = function of t.vec; matrix, M rows, m columns
# D.survey = estimated number dying on survey days
#       = function of P.dead.vec, Delta, Xi, c.vec; vector, length m
# D.season = estimated number dying on season days
#       = function of D.survey, Xi; vector, length M
# D.period = estimated number dying during survey periods
#       = function of D.season; length m
# L.survey = estimated number alive on survey days
#       = function of l.vec, P.live.vec; vector, length m
# A.period = estimated number arriving in survey periods
#       = function of D.period, l.vec, P.live.vec; vector, length m
# N.hat = estimated escapement; function of D.season; scalar
# D.distn = estimated cumulative distribution of dead over survey periods
#       = function of D.period, N.hat; vector, length m
# A.distn = estimated cumulative distribution of arrivals over survey periods
#       = function of A.period, N.hat; vector, length m
# T.hat = estimated stream residence time; function of D.distn, A.distn, t.vec; scalar
#       = the point estimate

M=ncol(R.mat) # length of season (in days)
m=length(t.vec) # number of survey days (m <= M)

###### Define Matrices
### Constant matrices (for estimating A.distn)
I.mat=diag(rep(1,m))

tmp.mat=matrix(0,nrow=m,ncol=m)
for(i in 2:m) tmp.mat[i,]=c(rep(0,i-2),-1,rep(0,m-i+1))
J.mat=I.mat+tmp.mat
rm(tmp.mat,i)

### Variable Matrices
 # Delta = function of P.dead.vec, R.mat
 del=matrix(0,nrow=m,ncol=M)
 for(i in 1:m)
 {
  tmp=rep(0,m)
  tmp[i]=1
  if(i<m) {for(j in (i+1):m) tmp[j]=prod((1-P.dead.vec)[i:(j-1)])}
  if(i==1) col.1=1 else col.1=t.vec[i-1]+1
  col.2=t.vec[i]
  del[,c(col.1:col.2)]=tmp
 }
 Delta=R.mat*del
 rm(i,j,tmp,col.1,col.2,del)

 # Xi = function of t.vec
 Xi=matrix(0,nrow=M,ncol=m)
 Xi[1:t.vec[1]]=c(1:t.vec[1])/t.vec[1]
 for(j in 2:m) Xi[(t.vec[j-1]+1):t.vec[j],j]=c(1:(t.vec[j]-t.vec[j-1]))/(t.vec[j]-t.vec[j-1])
```

```
 for(j in 1:(m-1))Xi[(t.vec[j]+1):t.vec[j+1],j]=(t.vec[j+1]-t.vec[j]-c(1:(t.vec[j+1]-t.vec[j])))/(t.vec[j+1]-t.vec[j])
 if(t.vec[m]<M) {for(j in (t.vec[m]+1):M) Xi[j,m]=(M-j)/(M-t.vec[m])}
 rm(j)

###### Estimate death and arrival distributions
 D.survey=as.vector(solve(diag(P.dead.vec)%*%Delta%*%Xi)%*%c.vec)
 D.season=Xi%*%D.survey
 t.diff=c(t.vec[1],t.vec[2:m]-t.vec[1:(m-1)])
 D.period=unlist(lapply(split(D.season[1:t.vec[m]],rep(t.vec,times=t.diff)),sum))
 N.hat=sum(D.season)
 L.survey=as.vector(solve(diag(P.live.vec))%*%l.vec)
 A.period=as.vector(I.mat%*%D.period + J.mat%*%L.survey)

 f1=function(i,x) return(sum(x[1:i]))
 D.distn=sapply(c(1:m),f1,x=D.period/N.hat)
 A.distn=sapply(c(1:m),f1,x=A.period/N.hat)

##### Estimate stream residence time (T.hat)
 A.tmp=c(A.distn[1],A.distn[2:m]+A.distn[1:(m-1)])
 D.tmp=c(D.distn[1],D.distn[2:m]+D.distn[1:(m-1)])
 T.hat=0.5*sum(t.diff*(A.tmp-D.tmp))

 out=list(Delta,Xi,D.survey,D.season,D.period,L.survey,A.period,N.hat,D.distn,A.distn,T.hat)
 names(out)=c("Delta","Xi","D.survey","D.season","D.period","L.survey","A.period","N.hat",
              "D.distn","A.distn","T.hat")
 return(out)
}

#####################################################################################
## Example: from Skalski, Liao, and Buchanan (2009)
# M=6; m=3
t1=c(2,4,6)
c1=c(10,20,50)
l1=c(48,45,0)
P1=c(0.8,0.8,0.8)
R1=matrix(c(c(0.8,1,0,0,0,0),c(0.5,0.6,0.8,1,0,0),c(0.3,0.4,0.5,0.6,0.8,1)),nrow=3,byrow=T)

# Run res.time to estimate T, N, and their standard errors:

res.time(t1,c1,l1,R1,P1,P1)
# returns:
#           T        se.T           N        se.N
# 2.2379227   0.1571949 103.8909894   5.4809977



# If the reader wants to separately calculate the matrices and estimates of the cumulative distribution
# of deaths and arrivals that are used in estimating T3 and N, then run res.time.pt:

res.time.pt(t1,c1,l1,R1,P1,P1)
# returns:
# $Delta
#      [,1] [,2] [,3] [,4] [,5] [,6]
# [1,] 0.800 1.000   0.0 0.00   0.0    0
```

```
# [2,] 0.100 0.120   0.8 1.00   0.0   0
# [3,] 0.012 0.016   0.1 0.12   0.8   1
#
# $Xi
#      [,1] [,2] [,3]
# [1,]   0.5   0.0   0.0
# [2,]   1.0   0.0   0.0
# [3,]   0.5   0.5   0.0
# [4,]   0.0   1.0   0.0
# [5,]   0.0   0.5   0.5
# [6,]   0.0   0.0   1.0
#
# $D.survey
# [1] 8.928571 14.221939 38.393313
#
# $D.season
#           [,1]
# [1,] 4.464286
# [2,] 8.928571
# [3,] 11.575255
# [4,] 14.221939
# [5,] 26.307626
# [6,] 38.393313
#
# $D.period
#        2        4    6
# 13.39286 25.79719 64.70094
#
# $L.survey
# [1] 60.00 56.25   0.00
#
# $A.period
# [1] 73.392857 22.047194   8.450938
#
# $N.hat
# [1] 103.8910
#
# $D.distn
# [1] 0.1289126 0.3772228 1.0000000
#
# $A.distn
# [1] 0.7064410 0.9186557 1.0000000
#
# $T.hat
# [1] 2.237923
```