Estimating the Internmol Periods in Asynchronously Molting Crustacean Populations

John M. Hoening
University of Miami, Rosenstiel School of Marine and Atmospheric Science, Cooperative Institute for Marine and Atmospheric Studies, 4600 Rickenbacker Causeway, Miami, Florida 33149, U.S.A.

and

Victor R. Restrepo
University of Miami, Rosenstiel School of Marine and Atmospheric Science, Division of Biology and Living Resources, 4600 Rickenbacker Causeway, Miami, Florida 33149, U.S.A.

SUMMARY

In some crustacean populations, including many tropical lobsters and crabs, molting occurs asynchronously throughout the year. That is, at any time of year, the animals are uniformly distributed throughout the internmolt cycle. We present methods for finding maximum likelihood estimates of the size-specific internmolt period for this situation when mark-recapture data are available consisting of initial size, time at liberty, and whether the animal molted while at liberty. Alternatively, the internmolt period can be determined by the same estimation procedure when wild animals are brought into captivity and observations are made on which animals molt within some fixed time period or before a fixed date. The method is applied to data on spiny lobster (Panulirus argus) to estimate internmolt period as an exponential function of length.

1. Introduction

Crustacean growth is a discontinuous process that is often described in terms of two components: the change in size at molt and the duration of the internmolt period (Botsford, 1985). Information on these processes is used to construct a model of growth in size over time for use in fishery assessment models. Both growth components vary with the size of the animal. These can be studied by observing animals held in captivity though the results may not be representative of growth in wild populations. An alternative is to estimate the components of growth from mark-recapture data.

In this paper, we focus on the problem of estimating the internmolt period under the assumption that the individuals in a population at any time are uniformly distributed throughout their internmolt period. We use the method of maximum likelihood to derive a procedure for estimating the internmolt periods of tropical and subtropical decapods from mark-recapture data when seasonal changes in growth are negligible. The procedures can also be used to analyze laboratory data on the proportion of animals molting when animals are held for periods of time sufficiently short to minimize the effects of captivity.

In Section 2, we consider mark-recapture studies and develop the estimators for internmolt period as a function of body size and other covariates. The growth of spiny lobsters

Key words: Growth; Mark-recapture; Maximum likelihood estimation; Spiny lobster; Tropical decapod crustacean.

7
2. Models for Mark–Recapture Data

For the models which follow, we assume that the intermolt period is some nondecreasing function $g$, of known form, of the animal’s size $L$. The objective is to estimate the coefficients of the function $g(L)$.

We also assume a lack of molting synchronicity in the population. This means that, if an animal is selected at random from the population, the elapsed time since the last molt, $U$, is distributed as a uniform random variable with parameter $g(L)$. Thus,

$$U \sim U[0, g(L)].$$  

Finally, it is necessary to assume that tagged animals do not have an increased probability of dying or losing the tag at the time of molting (relative to other times in the growth cycle).

2.1 The Molt–No-Molt Model

Consider an animal of length $L$ that is marked, released, and recaptured $\Delta t$ time units later. Suppose first that $\Delta t$ is less than $g(L)$, i.e., that the time at liberty is less than the intermolt period of time. It follows that the probability the animal molted once while at liberty is given by (Fig. 1, line labelled type 1)

$$P(L, \Delta t) = Pr[U > g(L) - \Delta t] = 1 - \int_{g(L) - \Delta t}^{g(L)} g(L)^{-1} du = \Delta t / g(L).$$  

![Figure 1](image.png)

Figure 1. Probability of molting as a function of the time at liberty under the molt–no-molt model. Type 0 records are animals that did not molt; type 1 molted once by chance; type 2 necessarily molted (at least) once.
Thus, the probability of molting for this case is proportional to the time at liberty. The probability that an animal did not molt, given that $\Delta t < g(L)$, is

$$P(D, \Delta t) = 1 - P(L, \Delta t).$$  \hspace{1cm} (5)$$

Now, if $\Delta t$ is greater than or equal to $g(L)$, the probability of molting one or more times is simply

$$P(D, \Delta t) = 1.$$  \hspace{1cm} (4)$$

That is, the animal must have molted.

The above leads us to define three kinds of mark-recapture records:

- type 0: no molt, $g(L) > \Delta t$,
- type 1: molt, $g(L) > \Delta t$,
- type 2: molt, $\Delta t \geq g(L)$.

In a typical mark-recapture experiment the data would consist of $n$ records, each comprised of an initial length $L$, the time at liberty $\Delta t$, and an indication of whether the animal molted. We will treat the initial lengths and the times at liberty as being fixed rather than as random variables. Let the number of records of type 0, 1, and 2 be $n_0$, $n_1$, and $n_2$, respectively. Then, if we could assign each record to the proper type, the likelihood function would be

$$L = \prod_{j=1}^{n_0} P(D_j, \Delta t_j) \prod_{j=1}^{n_1} P(D_j, \Delta t_j) \prod_{j=1}^{n_2} P(D_j, \Delta t_j).$$  \hspace{1cm} (5)$$

The downhill simplex search algorithm of Nelder and Mead (1965) can be used to maximize the log-likelihood function. A subroutine in R is suitable for the computations is presented in Press et al. (1986). For each set of coefficients tried, it is necessary to compute the estimated $g(L)$ for each animal that molted. If the $\Delta t$ value is greater than the estimated $g(L)$ in the current iteration then the animal is classified as type 2; otherwise, it is treated as type 1.

Since the usual regularity conditions are not met by the model, one cannot assume most of the desirable asymptotic properties of maximum likelihood estimation. However, the estimators are consistent and asymptotically unique (Bury, 1975). Variance-covariance estimation can be accomplished using a sample reuse technique such as the jackknife or the bootstrap (Efron, 1982).

Note that if none of the animals in the sample molted, the values of $g(L)$ that maximize $L$ would be infinite. One should then redesign (extend) the experiment to obtain longer times at liberty. Conversely, if all of the animals molted, the $g(L)$ values would be predicted to be 0. In this case, the multiple-molt model of Section 2.2 would be preferable.

2.2 The Multiple-Molt Model

In this section we assume that each recapture record can be classified into one of $K + 1$ categories according to the animal molted 0, 1, ..., $K + 1$ times, where "$K + 1$" means "$K$ or more." We also assume the size-specific molt increments are known. The molt size increments can be determined from mark-recapture data using some variation of the Hiai growth diagram (Hiai, 1948), in which the size at recapture is plotted against the size at
marking. Animals that did not molt form a cluster of points along the 45° line. Those that molted once form another cluster slightly above the 45° line, and so forth. Generally, it is not possible to distinguish between animals that molted two or more times. Recently, Shetton and Chagottin (1987) discovered that, if a piece of cuticle with attached epidermis is implanted in the abdomen of a crustacean, then the cuticle will develop alternating layers of cuticle and epidermis in synchrony with the molting cycle of the animal. This method may provide an accurate method for determining the number of times an animal has molted and should be especially useful for large animals in which the molt increment may be only slightly larger than the measurement error.

For convenience, we will illustrate the method by assuming that (i) animals molt either 0, 1, or 2+ times, and (ii) molt increments are given by the linear relationship

\[ L_i = c + dL_i, \]

where \( L_i \) and \( L_0 \) are the sizes before and after molting, respectively. The values of \( c \) and \( d \) can be treated as known or can be estimated simultaneously with the coefficients of \( g(L) \). Extensions to higher values of \( K \), and to other formulations for the molt increment, are straightforward.

Recapture records can be thought of as comprising six types when \( K = 2 \) (3 × \( K \) types in general). These are defined in Table 1. Again, for any set of guesses of the coefficients, the values of the internal molt period \( g(L_0) \) and the subsequent internal molt period \( g(L) = g(c + dL_0) \) can be computed in order to separate animals that molted once into types 1, 2, and 3, and to separate animals molting two or more times into types 4 and 5 (Fig. 2).

For each set of guesses of the parameter values, the likelihood function would be constructed as

\[ L = \prod_{j=1}^{n} P_i(L_0, \Delta t) \prod_{j=1}^{n} P_j(L_i, \Delta t_i) \prod_{j=1}^{n} P_k(L_i, \Delta t_i) \prod_{j=1}^{n} P_m(L_0, \Delta t), \]

where the subscript \( i \)'s refer to the number of records assigned to each type, and the form of the probability functions \( P_i(L_i, \Delta t) \) (\( i = 0, 1, \ldots, 5 \)) are as given in Table 1. Of course, the products over \( j \) and \( m \) are unity and can thus be neglected.

**Table 1**

<table>
<thead>
<tr>
<th>Type of record, ( x )</th>
<th>Number of molts</th>
<th>Relationship between ( \Delta t ) and ( g(L) )</th>
<th>Probability, ( P_i )</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>9</td>
<td>( \Delta t &lt; g(L_0) )</td>
<td>( P = 1 - \Delta t/g(L_0) )</td>
<td>Did not molt, by chance</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>( \Delta t &lt; g(L_0) )</td>
<td>( P = 1 - \Delta t/g(L_0) )</td>
<td>Molted once (as opposed to zero times, by chance)</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>( g(L_0) &lt; \Delta t &lt; g(L_0) )</td>
<td>( P = 1 )</td>
<td>Never molted once</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>( g(L_0) &lt; \Delta t &lt; g(L) )</td>
<td>( P = \Delta t/g(L) )</td>
<td>Molted once (as opposed to twice, by chance)</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>( g(L_0) &lt; \Delta t &lt; g(L_0) + g(L) )</td>
<td>( P = \Delta t/g(L) )</td>
<td>Molted twice (as opposed to once, by chance)</td>
</tr>
<tr>
<td>5</td>
<td>2+</td>
<td>( g(L_0) + g(L) &lt; \Delta t )</td>
<td>( P = 1 )</td>
<td>Necessarily molted 2 or more times</td>
</tr>
</tbody>
</table>
2.3 Effect of Covariates

Growth of decapod crustacea is believed to vary according to factors such as sex and whether the animal has been injured (Davis, 1981; Hunt and Lyons, 1986). These can be incorporated into the analysis using indicator variables. For example, if the intermolt time model is given by

\[ g(L) = ae^{\beta L} \]

then the covariate \( Y \) indicating sex could be added as

\[ g(L) = \left( a + \alpha Y \right)e^{\beta L} \]

The \( Y \) values are considered fixed. Likelihood-ratio tests can be used to decide among models.

2.4 Diagnostics

An obvious diagnostic procedure for detecting lack of fit would be to compare the observed number of animals molting, \( M \), with the expected number, \( E(M) \), estimated by summing the estimated probabilities of molting over all animals. This can be computed separately for different length and time intervals. Another statistic to consider is the number of animals, \( FM \), that failed to molt even though the estimated probability of molting was unity (because the time at liberty was greater than the estimated intermolt period). This statistic should be equal or close to \( 0 \).

3. Example: Intermolt Periods of Spiny Lobster

We use for illustrative purposes the tagging data of Little (Florida Department of Natural Resources Marine Research Laboratory, Special Scientific Report No. 31, 1972) to estimate
the intermolt periods of spiny lobster \((Panulirus argus)\) from South Florida. We assume that the intermolt period \(g(L)\) is an exponential function of the carapace length, \(L\), i.e.,

\[
g(L) = ae^{-bL}.
\]

where \(a > 0\) and \(b > 0\). In addition to being flexible, this formulation for \(g(L)\) is consistent with generally accepted beliefs that intermolt period increases with size, and is consistent with observations on the species reported by Hunt and Lyons (1986).

We assume that at any time of the year the animals are uniformly distributed throughout the intermolt period (equation (1)). This means that animals do not molt in synchrony, e.g., in response to environmental cues such as lunar cycle or rising spring water temperatures. Empirical support for the assumption consists of observations that molting occurs at all times of the year with minor, if any, seasonal trend (Lefcious, 1985).

3.1 Parameter Estimation

In the unpublished scientific report previously cited, Little presented 68 mark-recapture records consisting of carapace length at tagging, length at recapture, and number of weeks at liberty. Little considered an animal to have molted if its size at recapture exceeded the size at tagging by more than 2 mm. This criterion was also adopted by Hunt and Lyons (1986) and was used in our analysis. We used the computer routine in Press et al. (1986) to obtain the estimates \(\hat{a} = 1.15\) and \(\hat{b} = 0.033\). Using the jackknife procedure, we estimated the variances and covariances to be

\[
\text{var}(\hat{a}) = 0.410;
\]

\[
\text{var}(\hat{b}) = 6.29 \times 10^{-4};
\]

\[
\text{cov}(\hat{a}, \hat{b}) = -5.08 \times 10^{-1}.
\]

A confidence band around the estimated function \(g(L)\) was constructed using the Monte Carlo procedure described in the next section (Fig. 3).

![Figure 3](image)

**Figure 3.** Estimated relationship between intermolt time and carapace length for spiny lobsters \((Panulirus argus)\) from South Florida. The dashed lines represent 95% confidence limits obtained by Monte Carlo simulation.
3.2 Monte Carlo Studies

In order to study the performance of the maximum likelihood procedure under realistic conditions, we conducted Monte Carlo simulation studies based on the example in Section 3.1. The times at liberty and initial sizes were fixed at the values given in the scientific report by Litjens. The values of \( a \) and \( b \) were assumed to be 10, with 0.5% and 0.1% respectively for data sets size 68 and 272. Jackknifed estimates of \( a \) and \( b \) were used rather than the non-jackknifed estimates and consequently are not considered further.

Table 2

<table>
<thead>
<tr>
<th>Parameter (true value)</th>
<th>Mean estimate (% bias) from simulations when</th>
</tr>
</thead>
<tbody>
<tr>
<td>( a ) (1.15)</td>
<td>100 (4.02)</td>
</tr>
<tr>
<td>( b ) (0.033)</td>
<td>100 (3.02)</td>
</tr>
<tr>
<td>( \sigma^2 ) (1.812)</td>
<td>100 (0.83)</td>
</tr>
<tr>
<td>( \sigma^2 (b) \times 10^{-3} )</td>
<td>100 (5.02)</td>
</tr>
<tr>
<td>( \sigma^2 (a) \times 10^{-3} )</td>
<td>100 (5.02)</td>
</tr>
<tr>
<td>( \sigma^2 (a) \times 10^{-3} )</td>
<td>100 (5.02)</td>
</tr>
</tbody>
</table>

Parameter estimates for the coefficient \( a \) were somewhat skewed when the size of the data sets was 68 (Fig. 6A). However, for data sets of size 272, the parameter estimates appeared approximately normally distributed (Fig. 6B). Estimates for the coefficient \( b \) appeared approximately normally distributed for both samples of size 68 and 272 (Fig. 5).

The actual variances and the covariance of the parameter estimates were approximated by computing the sample estimates for the 1,000 simulations for data sets of size 68. These results were compared to the mean of 500 estimates of the covariance matrix obtained by jackknifing (Table 2). The jackknife procedure provided conservative estimates.

In a second simulation study, we investigated the effect of individual variability in the interval by assuming a given length, the interval period is given by

\[ g(L) = \exp^{\frac{L}{M}}, \]
where $e_j$ is a log-normally distributed error,

$$e_j \sim \text{LN}(0, \sigma^2),$$

and

$$\gamma e^{\gamma/2} = a = 1.15$$

such that

$$E[\log(L)] = a e^{\mu} = 1.15 e^{a \text{var}(\mu)}.$$

Values of $\sigma^2$ were chosen such that the coefficient of variation, CV,

$$CV = 100(\sigma^2 - 1)^{1/2},$$

took on the values 0, 25, 50, 75, and 100%. Using the original data set of size 68, we ran 500 simulations for each value of the coefficient of variation, using random number
Table 3 Results of Monte Carlo studies of the effects of individual variability in intermolt periods on estimates obtained with the most-no-molt model. Estimates are based on 500 simulations for each coefficient of variation (CV).

<table>
<thead>
<tr>
<th>CV (%)</th>
<th>Mean estimate (% bias)</th>
<th>Mean estimate (%)</th>
<th>var(a)</th>
<th>var(b)</th>
<th>cov(a, b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.23 (6.9)</td>
<td>0.05 (4.5)</td>
<td>0.748</td>
<td>0.828</td>
<td>0.699</td>
</tr>
<tr>
<td>25</td>
<td>1.57 (36.3)</td>
<td>0.03 (0.0)</td>
<td>2.319</td>
<td>1.013</td>
<td>-1.237</td>
</tr>
<tr>
<td>50</td>
<td>2.09 (81.8)</td>
<td>0.03 (2.7)</td>
<td>4.762</td>
<td>1.677</td>
<td>-2.172</td>
</tr>
<tr>
<td>75</td>
<td>2.60 (126.3)</td>
<td>0.03 (-8.3)</td>
<td>8.299</td>
<td>1.849</td>
<td>-2.847</td>
</tr>
<tr>
<td>100</td>
<td>3.18 (176.4)</td>
<td>0.029 (-30.3)</td>
<td>11.135</td>
<td>2.577</td>
<td>-4.699</td>
</tr>
</tbody>
</table>

The results indicate that the bias in the estimates of the coefficient \( a \) became progressively greater with increased variability in intermolt periods, being roughly 7% for a CV of 0% and increasing up to 176.4% for a CV of 100% (Table 3). The bias in \( b \) was slightly positive (4.9%) for a CV of 0% and became negative for increased variability in \( g(L) \) (−10.3% for CV = 100%). The variance and covariance of the parameters, as approximated from the observed variability in the estimates in the simulations, also increased in absolute magnitude with increasing variability in intermolt periods (Table 3).

We examined the bias in estimated length-specific intermolt periods caused by individual variability in intermolt period. For each of the 500 pairs of parameter estimates generated in the simulations, we calculated \( g(L) \) values and compared the means to the true values (calculated with \( a = 1.15, b = 0.033 \)). For the range of lengths represented in the data, bias in the intermolt period was slightly negative for a CV of 0% and became increasingly positive with larger CV (greater than 25%) (Fig. 6).

![Figure 6](image)

Figure 6. Percentage bias in the estimated length-specific intermolt periods obtained by Monte Carlo simulation for different magnitudes of individual variability in intermolt periods. CV = coefficients of variation of the length-specific intermolt period. The results are based on the average intermolt period calculated from 500 simulated estimates of the parameters \( a \) and \( b \) for each value of the CV.
4. Discussion

The principal assumption of the methods presented here is encoded in equation (1), which implies that the intermolt period $g(L)$ does not vary over the course of the study and that there is no synchronicity in molting. This is likely to be true only in the tropics where molting can occur at all times of the year with little seasonal modulation. In certain deepwater shrimp populations, the effects of seasonality on growth may also be minimal. In contrast, crustacean populations in temperate waters tend to have depressed molting in winter so that $g(L)$ varies seasonally. Also, if reproduction is restricted to a narrow period of time the youngest life stages may exhibit a large degree of molting synchronicity. This would invalidate the use of our methods for the smallest animals.

Monte Carlo studies indicate that biases in estimates of $a$, $b$, and $g(L)$ are small (under 10%) where there is no individual variability in intermolt periods. The biases in estimates of $g(L)$ were under 5% when the coefficient of variation (CV) of the length-specific intermolt period was under 25%, and biases were less than 20% when the CV was under 50%. We could not find any information in the literature on likely values of the CV. However, we estimated a value of 34% from data on another large decapod, the stone crab (Menippe mercenaria) in captivity (Restrepo, unpublished results). Thus, it appears that the molt-to-molt model of Section 2.1 is reasonably robust to failure of the assumption of no variability in intermolt periods.

4.1 Other Methods

Munro (1974, 1985) suggested that if animals are uniformly distributed throughout their intermolt periods at all times, then the time it takes for half of a cohort of tagged animals (of similar size) to molt is an estimate of half the intermolt period. This follows from equations (2)-(4) since, for a suitably small length interval, the expected proportion of animals in a small time interval that have molted will be proportional to the time at liberty until the time at liberty exceeds the intermolt period (Fig. 1, lines labeled types 1 and 2). Thus, Munro divided the recapture records into groups according to the time at liberty and noted the time interval in which approximately half the animals had molted. The grouping of animals by length and by time makes Munro's method data-intensive.

Departures from the expected linear relationship in a Munro plot could be caused by several factors, such as individual variability in intermolt times, and acute tag loss or mortality at the time of molting. Munro's method appears to be quite robust to these failures of assumption (Keister and Hoening, 1988). The Munro plot may also have some value as a diagnostic procedure. However, it should be noted that nonlinearity can also be an artifact of using length and time intervals that are not sufficiently small. The diagnostic procedure of comparing observed and predicted number of animals molting (Eq. 4) is preferable particularly when the number of recapture records is small.

4.2 Laboratory Studies

It is common practice to estimate intermolt times by bringing wild animals into captivity, waiting for them to molt, and then recording the times until the next molt. This is a time-consuming procedure (one has to wait two intermolt cycles to obtain the results) and the results may not be indicative of growth in the wild.

If one can assume that at any time of the year the animals are uniformly distributed throughout their intermolt cycle [equation (1)], then the times to first molt in captivity can be used to estimate the intermolt period. For example, for animals in a narrow size class, the maximum likelihood estimator of the intermolt period would simply be the largest order statistic (Freund, 1971). This procedure is obviously sensitive to outliers. Thus, a
wounded as diseased animal that takes a long time to molt would greatly influence the estimate. A more robust alternative would be to use the moment estimator, i.e., \( \hat{g}(L) \) for a narrow site class would be twice the mean time to first molt. This still requires keeping animals in captivity for up to one full intermolt period. Another alternative is to use twice the mean time to first molt as the estimate of \( g(L) \). This requires less work and is more robust as animals need be held in captivity only until half have molted (Childress and Price, 1978). The procedure is structurally equivalent to Muroy’s (1974) estimator based on mark-recapture data. Of course, the choice of the median as the basis for the estimator is arbitrary and, in general, one can estimate the intermolt period by

\[ \hat{g}(L) = u \]

where \( t \) is the duration of the observation period in captivity and \( p \) is the proportion of animals molting (unpublished doctoral dissertation, University of Califorina, Santa Barbara, 1977). Either \( t \) or \( p \) can be fixed in advance. If \( t \) is large, the total observation time can be divided into intervals of equal duration and the cumulative proportion molting can be plotted against the rank of the interval. This method can be used to check graphically on the goodness of fit since the expected relationship is linear.

All of the above methods deal with length-specific variability in intermolt times by requiring animals to be grouped into narrow length intervals which are studied separately. However, this inefficient grouping of data can be avoided if one employs the molt-so-molt model of Section 2.1. For example, one can collect animals over a period of time and make observations of molting in captivity until a fixed date. Then for each animal one will have information on the initial size, the time in captivity, and whether the animal molted. The times in captivity and lengths are considered fixed. If the exact times of molting are also recorded, then one can make a series of estimates by “fixing” the termination date at several points. One can then look for a trend over experiment time in the estimates of \( g(L) \) in order to check for failure of the assumption embodied in equation (1).

ACKNOWLEDGMENTS

We thank Adrian Wenner, Dennis Heisey, Kenneth Follock, and two anonymous reviewers for their helpful comments and encouragement. Support for this study was provided through the Cooperative Institute for Marine and Atmospheric studies by National Oceanic and Atmospheric Administration Cooperative Agreement No. NA85-WCH-06134.

RESUME

Dans quelques populations de crevettes, notamment dans le genre de Crustacea, le" molt" est bien synchronisé sur l'an. La moitié des animaux a pu être observée a différentes époques de la même année. Les animaux sont distribués uniformément sur le cycle interne. Nous prêtons de la mort d'une partie des animaux à des variations du rapport de naissance. Le" molt" intervient quand les animaux sont maintenus en captivité et qu'on observe les animaux qui sont en phase de mue. On applique la méthode à des données sur la longueur (Leroy, 1974) pour estimer la période intermolt comme une fonction exponentielle de la longueur.

REFERENCES

