Stochastic growth models for analyzing crustacean data

Chuan Hui Foo and You-Gan Wang

Centre for Applications in Natural Resource Mathematics (CARM), School of Mathematics and Physics, The University of Queensland, Brisbane, Australia
Email: c.foo@uq.edu.au

Abstract: The contemporary methodology for growth models of organisms is based on continuous trajectories and thus it hinders us from modelling stepwise growth in crustacean populations. Growth models for fish are normally assumed to follow a continuous function, but a different type of model is needed for crustacean growth. Crustaceans must moult in order for them to grow. The growth of crustaceans is a discontinuous process due to the periodical shedding of the exoskeleton in moulting. The stepwise growth of crustaceans through the moulting process makes the growth estimation more complex. Stochastic approaches can be used to model discontinuous growth or what are commonly known as “jumps” (Figure 1). However, in stochastic growth model we need to ensure that the stochastic growth model results in only positive jumps. In view of this, we will introduce a subordinator that is a special case of a Levy process.

A subordinator is a non-decreasing Levy process, that will assist in modelling crustacean growth for better understanding of the individual variability and stochasticity in moulting periods and increments. We develop the estimation methods for parameter estimation and illustrate them with the help of a dataset from laboratory experiments. The motivational dataset is from the ornate rock lobster, Panulirus ornatus, which can be found between Australia and Papua New Guinea.

Due to the presence of sex effects on the growth (Munday et al., 2004), we estimate the growth parameters separately for each sex. Since all hard parts are shed too often, the exact age determination of a lobster can be challenging. However, the growth parameters for the aforementioned moult processes from tank data being able to estimate through: (i) inter-moult periods, and (ii) moult increment.

We will attempt to derive a joint density, which is made up of two functions: one for moult increments and the other for time intervals between moults. We claim these functions are conditionally independent given pre-moult length and the inter-moult periods. The variables moult increments and inter-moult periods are said to be independent because of the Markov property or conditional probability. Hence, the parameters in each function can be estimated separately. Subsequently, we integrate both of the functions through a Monte Carlo method. We can therefore obtain a population mean for crustacean growth (e.g. red curve in Figure 1).

Figure 1: Growth trajectories of crustaceans

Keywords: Subordinator, discontinuous growth trajectories, moulting
1. INTRODUCTION

Crustacean fisheries, such as crabs, lobsters and prawns, play a vital role in global markets. A number of studies have been implemented to study the growth of crustaceans (Hoenig and Restrepo, 1989; Hartnoll, 2001; Wahlle and Fogarty, 2006). As Hoenig and Restrepo (1989) discussed estimation of inter-moult parameters from tag-recapture data, Hartnoll (2001) reviewed some improvements in methods to model growth of crustaceans.

We consider two general problems in growth estimation, as a result of the moulting process. Firstly, the growth of crustaceans has a stepwise fashion that has hindered us from understanding the growth thoroughly. Secondly, the growth of crustaceans is determined by inter-moult periods and moult increments concurrently and therefore both processes should be integrated to describe the individual growth paths.

The purpose of this paper is to introduce a stochastic growth component into a growth model that accounts for stepwise growth of crustaceans. A Levy process has right-continuous with left-limits paths, implying that the process has a countable number of jumps on a finite interval. Hence, an individual's growth can be modelled by a Levy process.

A Levy process, $X = (X_t)_{t \geq 0}$ is a stochastic process with the four properties,

(i) $P(X_0 = 0) = 1$.  
(ii) $X_t$ is almost surely right-continuous with left-limits.  
(iii) For any $s \leq t$, increments $X_t - X_s$ have the same distribution as $X_t - s$.  
(iv) For any $0 \leq t_1 < t_2 < \cdots < t_n < \infty$, $X_{t_2} - X_{t_1}$, $X_{t_3} - X_{t_2}$, $\ldots$, $X_{t_n} - X_{t_{n-1}}$ are independent.

A number of papers reformulated growth models as the stochastic differential equations, but some models may predict negative size increments (Garcia, 1983). Therefore, we introduce a subordinator, which is a non-decreasing Levy process that can be used to model nondecreasing measurement in growth, e.g. length of crustaceans. We consider two examples from a class of subordinator for fitting the moult increment from tank data: the Gamma process and the inverse Gaussian process. By estimating the population mean and variance of the size increment, we can then average over the discontinuous growth paths to yield a population growth trend.

2. STOCHASTIC GROWTH MODELS

2.1 Subordinator

A subordinator is a non-decreasing Levy process. It has no diffusion component, non-negative drift which means the individual never moves backward and non-negative jumps (Cont and Tankov, 2004). Let $\{Z_t, t \geq 0\}$ be a subordinator starting from the origin such that $Z_0 = 0$, the process is right continuous, and takes the value on $[0, \infty)$. It has independent and stationary increments on $[0, \infty)$. For $t \geq 0$, $Z_t \geq 0$ and $Z_{t_1} \leq Z_{t_2}$ at $t_1 < t_2$, it is more convenient to look at the Laplace transform of $Z_t$

$$M_t(\theta) = E(e^{-\theta Z_t}), \quad \theta \geq 0,$$

where $l(\theta)$ is the Laplace exponent.

The Laplace transform in equation (1) will be able to estimate the mean and the variance of the subordinators easily. If $Z_t$ has a finite mean $\mu$ and finite variance $\sigma^2$, then the first moment (mean) of the subordinator can be written as

$$E(Z_t) = \mu t,$$

and

$$Var(Z_t) = \sigma^2 t,$$

where $\mu$ and $\sigma^2$ are the mean and variance of $Z_t$ at $t = 1$. A subordinator can also take into consideration the individual and environmental variation.

Russo et al. (2009) applied both Gamma and inverse Gaussian processes as examples of subordinators to the Atlantic herring population growth. We can make use of the subordinators to model the growth of crustaceans based on their statistical framework.

A Gamma process is a subordinator following a Gamma distribution at time $t$ with density in the form of

$$f(x) = x^{\alpha-1} e^{-\lambda x} / \Gamma(\alpha, \lambda), \quad x > 0,$$

where

$$\Gamma(\alpha, \lambda) = \int_0^\infty x^{\alpha-1} e^{-\lambda x} dx.$$
for $\alpha, \lambda > 0$ and $f(x) = 0$ elsewhere. The parameter $\alpha$ is the average rate of jumps and the scaling parameter $\lambda$ controls the distribution of jump sizes.

The Laplace transform of (4) is

$$M_r(\theta) = E e^{-\lambda \theta} = \exp \{-\lambda t \log \lambda/\lambda + \theta\}$$

$$= (\lambda/\lambda + \theta)^{-\alpha}.$$  

(5)

Hence, the first and second derivatives of (5) yield the mean and variance of gamma process.

$$E(Z_t) = \alpha t / \lambda.$$

and

$$Var(Z_t) = \alpha t / \lambda^2.$$

Next, we will look into the second example of the subordinator which is the inverse Gaussian process. While a Gaussian describes a Brownian motion's level at a fixed time, an inverse Gaussian process describes how long a Brownian motion with positive drift takes to reach a fixed positive level.

Here, the subordinator given by $Z_t = \sqrt{a} B_t + bt$; $t \geq 0$ where $(B_t)$ is a standard Brownian motion with drift $b > 0$ and $a > 0$. Then the first passage time of $Z_t$ for a fixed level $\alpha > 0$ is

$$T_\alpha = \inf\{t \geq 0 | Z_t = 0, Z_t \geq \alpha\}.$$  

Assuming $T_\alpha$ follows an inverse Gaussian distribution, where $\text{IG}(\mu, \lambda)$, with the mean and scale parameters

$$\mu = a / b, \quad \lambda = a^2 / a.$$  

The first passage time $T_\alpha$ for fixed level $\alpha$ has probability density function

$$f_\alpha(x) = \exp \left[\left(-b^2 x^3 / (2a) - (a^2 / 2ax)\right)\right] a e^{b/2} / \sqrt{2\pi a x}.$$  

(6)

The Laplace transform of eqn (6) is

$$M_\alpha(\theta) = \exp \{-\alpha \sqrt{b^2/\alpha^2} + (2a/\alpha b)\}.$$  

The corresponding subordinator has a mean $E(Z_\alpha) = \mu = a / b$, and variance $= \mu^2 / \lambda = a\alpha / b^2$, where $a, b > 0$.

3. MODELLING THE STEPWISE GROWTH PROCESS

The growth of crustaceans can be described by two processes, which are moult time interval (inter-moult) and moult length increment (moult increment). We need to account for the discontinuities or jumps due to the moulting process in individuals. In the past, researchers mainly focused on the study of continuous growth rather than on discontinuous growth processes (Chang et al., 2012). In this context, we will take into consideration two of the stochastic processes (inter-moult periods and moult increment) that characterize the stepwise trajectories of each individual. Subsequently, we can produce the corresponding population growth curve.

We focus on two types of data: (i) tank data from rearing experiments and (ii) tagging data from tag-recapture experiments. Tank data consist of exact moulting times as well as increments from each moult throughout the study. For tank data, it is relatively straightforward to estimate the size-at-age of animals since we observe their growth throughout the moulting process. We can also use tag-recapture data to estimate the growth parameters.

First of all, we will construct a likelihood approach to accommodate the aforementioned growth parameter estimation. Suppose individual $i$ consists of $n_i$ repeated measures $\{L_{ij}, T_{ij}, I_{ij}\}$ from $j = 1$ to $n_i$ where $L_{ij}^{-1}$ is the $j$th pre-moult length, $T_{ij}$ is the $j$th inter-moult time, and $I_{ij}$ is the $j$th increment length. Since each individual is measured at different times, a general density function is required to describe a sample of different lobsters. We define $f(L_{ij})$ as an initial length function for an animal. We assume $g(\cdot)$ is a function of inter-moult periods and $h(\cdot)$ is a moult increment function, that both condition on pre-moult length and inter-moult periods. For example, if lobster $i$ has moulted three times, we can derive the joint density function of inter-moult periods and moult increment as

$$f((L_{i1}, T_{i1}, I_{i1}), (L_{i1}, T_{i2}, I_{i2}), (L_{i1}, T_{i3}, I_{i3}))$$

$$= f(L_{i1}, T_{i1}, I_{i1}) f(L_{i1}, T_{i1}, I_{i1}) f(L_{i1}, T_{i1}, I_{i1})$$

$$= f(L_{i1}, T_{i1}, I_{i1}) f(L_{i1}, T_{i2}, I_{i2}) g(T_{i2}, L_{i1}, T_{i1}, I_{i1}) h(I_{i2}, L_{i1}, T_{i1}, I_{i1}) g(T_{i2}, L_{i1}, T_{i1}, I_{i1})$$

$$= f(L_{i1}, T_{i1}, I_{i1}) f(L_{i1}, T_{i2}, I_{i2}) f(L_{i1}, T_{i3}, I_{i3}) h(I_{i3}, L_{i1}, T_{i1}, I_{i1}) g(T_{i3}, L_{i1}, T_{i1}, I_{i1})$$

$$= f(L_{i1}, T_{i1}, I_{i1}) f(L_{i1}, T_{i2}, I_{i2}) f(L_{i1}, T_{i3}, I_{i3}) x h(I_{i3}, L_{i1}, T_{i1}) g(T_{i3}, L_{i1}, T_{i1}) f(L_{i1}).$$  

(7)
Because $L_{i1} = L_{i0} + I_{i1}$, and $L_{i2} = L_{i1} + I_{i2}$, the function (7) can be written as

$$f(L_{i2}, T_{i3}, L_{i3}) = f(L_{i1}, T_{i2}, L_{i2}) g(T_{i2} | L_{i1}, T_{i1}) h(I_{i2} | L_{i1}, T_{i1}) g(T_{i1} | L_{i0}, T_{i0}) f(L_{i0})$$

$$= g(T_{i3}, L_{i3}) h(I_{i3} | L_{i2}, T_{i3}) g(T_{i3}, L_{i3}) L_{i2} T_{i3} h(I_{i2} | L_{i1}, T_{i1}) g(T_{i1} | L_{i0}, T_{i0}) f(L_{i0})$$

$$= g(L_{i2} | L_{i1}, T_{i2}) g(T_{i2} | L_{i1}, T_{i1}) h(I_{i2} | L_{i1}, T_{i1}) g(T_{i1} | L_{i0}, T_{i0}) f(L_{i0})$$

$$= g(T_{i1} | L_{i2}, T_{i2}) g(T_{i2} | L_{i1}, T_{i1}) h(I_{i2} | L_{i1}, T_{i1}) g(T_{i1} | L_{i0}, T_{i0}) f(L_{i0})$$

$$= g(T_{i0} | L_{i2}, T_{i2}) g(T_{i2} | L_{i1}, T_{i1}) h(I_{i2} | L_{i1}, T_{i1}) g(T_{i1} | L_{i0}, T_{i0}) f(L_{i0})$$

In general, the joint density function for the $i$th individual can be written as

$$\prod_{j=1}^{n} [g(T_{ij} | L_{ij-1}, T_{ij-1}) \prod_{j=1}^{n} h(I_{ij} | L_{ij-1}, T_{ij}) f(L_{ij})].$$

Suppose there are $N$ individuals in our study. The likelihood function for all individuals is

$$L(\theta) = \prod_{i=1}^{N} \prod_{j=1}^{n} g(T_{ij} | L_{ij-1}, T_{ij-1}) \prod_{j=1}^{n} h(I_{ij} | L_{ij-1}, T_{ij}) f(L_{ij})$$

$$= \prod_{i=1}^{N} \prod_{j=1}^{n} g(T_{ij} | L_{ij-1}, T_{ij-1}) \prod_{j=1}^{n} h(I_{ij} | L_{ij-1}, T_{ij}) f(L_{ij})$$

(8)

where the first term is for the inter-moult periods function, and the second term is for the moult increments. Parameters in $g(\cdot)$ and $h(\cdot)$ can be estimated by maximising the two parts separately. Therefore, the first and second term are asymptotically independent given pre-moult length and inter-moult periods. The next section will discuss the models of interest with respect to the inter-moult periods function $g(\cdot)$ and moult increment function $h(\cdot)$ separately.

### 3.1 Modelling inter-moult periods from tank data

Here, we know exactly the moult ing time $T$ of the animals. An estimator of the inter-moult periods follows a log-normal distribution will be considered in the growth model. Suppose log(T) is distributed $N(\mu_t$, $\sigma^2$), the mean of inter-moult period $T$ is based on pre-moult length $L^-$ (Restrepo, 1989), therefore, we have

$$\mu_t = \log(\alpha) - \sigma^2/2 + \beta L^-,$$

where $\alpha > 0$ and $\beta \geq 0$ are constants. There are three parameters to be estimated: $\alpha$, $\beta$ and $\sigma$. The likelihood for inter-moult periods is

$$L(T; \alpha, \beta, \sigma) = \prod_{i=1}^{N} \prod_{j=1}^{n} [1/(\sigma T_{ij}^{\sqrt{2\pi}})] \exp \left(-[\log(T_{ij}) - \log(\alpha) - \sigma^2/2 + \beta L_{ij}]^2 / 2\sigma^2 \right).$$

For simplicity we assume all the observations in $j$ are independent for each sample $i$. Parameter estimation can be carried out by maximising the log $L(T; \alpha, \beta, \sigma)$.

### 3.2 Modelling moult increment from tank data

Let $\{Z_i, t \geq 0\}$ be a subordinator, the time changing Levy process $L_t = Z_{t_0}$, with the size-at-age of an animal denoted by $L_t$ at time $t$, and the time transformation is followed a von Bertalanffy $V_t = 1 - e^{-k(t-t_0)}$ where $k$ is the growth speed and $t_0$ is the age of fish at zero length (Russo et al., 2009). As discussed earlier, the size-at-age can be estimated as long as $Z_t$ has mean and variance based on equation (2) and (3) such that

$$E(L_t) = \mu V_t,$$

and

$$Var(L_t) = \sigma^2 V_t,$$

(9)

(10)

where $\mu$ and $\sigma^2$ are real numbers. Here, $\mu$ plays the role of $L_{t_0}$, thus the estimates follow von Bertalanffy growth function (VBGF) and possess monotonic increment behavior. However, we do not have information on age $t_0$ of the individuals. The Fabens’ method (1965) for estimating the size of animals will be used under the assumption that the ages are all unknown. The rearranged form of VBGF with its increment size of $I$ is

$$I = (l_e - L^-)(1 - e^{-kt})$$

where $l_e$ are the parameters.

Firstly, a Gamma process has mean $\mu = \alpha(1 - e^{-kt})\lambda$, and variance $\sigma^2 = \alpha(1 - e^{-kt})\lambda^2$. Thus $\alpha = \lambda(l_e - L^-)$, the
Foo and Wang, Stochastic growth models for analyzing crustacean data

 Estimated increment length after time interval $T$ from eqn (9) is

$$E(I|T, L^-) = (l_\infty - L^-) (1 - e^{-kT}) ,$$

and the variance of the increment length after time interval $T$ from eqn (10) is

$$Var(I|T, L^-) = (l_\infty - L^-) (1 - e^{-kT}) / \lambda .$$

Subsequently, the likelihood of the increment lengths from the $n$ animals is

$$L(I|T, L^-; l_\infty, k, \lambda) = \prod_{i=1}^{N} \prod_{j=1}^{n} \left[ \frac{\lambda^{(l_\infty - L^-)(1 - e^{-kT})}}{\Gamma(\lambda(l_\infty - L^-)(1 - e^{-kT}))} I_y^{\lambda(l_\infty - L^-)(1 - e^{-kT}) - 1} e^{-\lambda I_y} \right] ,$$

where $i$ is the individual, and $j$ is the number of measurements in each individual.

Secondly, we assume an inverse Gaussian distribution as in eqn (6), the mean and variance of a subordinator following an inverse Gaussian process are $\mu = (1 - e^{-kT}) / b$, and variance $\sigma^2 = a(1 - e^{-kT}) b^2$, where $1/b$ plays the role of $(l_\infty - L^-)$. Assuming the increment length to be $\{I|T, L^-\}$, $Z_{ij}$ follows an inverse Gaussian process $IG(1 - e^{-kT} / \sqrt{a}, b / \sqrt{a})$, where $L^-$ is the pre-moult length of the individuals, and $T$ is the moult time interval. The estimated increment length after time interval $T$ can be written as in equation (9)

$$E(I|T, L^-) = (l_\infty - L^-) (1 - e^{-kT}) ,$$

while the variance of the increment size after time interval $T$ is

$$Var(I|T, L^-) = a (l_\infty - L^-)^3 (1 - e^{-kT}) .$$

We could write the joint likelihood function on increment length at the time of $n$ individuals as follows

$$L(I|T, L^-; l_\infty, k, \lambda) = \prod_{i=1}^{N} \prod_{j=1}^{n} \left[ \frac{\lambda^{(l_\infty - L^-)(1 - e^{-kT})}}{2\pi a I_y^3} \exp \left( -\frac{I_y}{2a(l_\infty - L^-)} - \frac{(1 - e^{-kT})^2}{2al_y} \right) \right] ,$$

where the estimated parameters are $a, l_\infty$ and $k$.

Application and Numerical Studies

A total of 75 lobsters (39 females, 36 males) were reared in tanks over a time span. The dataset was collected from 1995 to 1999 and included observations of the carapace length ranging from 6.3 mm to 158.3 mm. The growth of the lobster was recorded via moult time intervals and moult increments. The individual lengths increased monotonically for both females and males throughout the growing process (Figure 2).

Figure 2: Growth length over years 1995 to 1999

Figure 3 indicates that males and females grow similarly before premoult length of 100 mm and Table 1 shows the parameter estimates through the maximum likelihood estimation. Females spent a longer time to moult than males as they grow older. Males appeared to spend shorter moulting periods to reach the maximal
length in less than 300 days, whilst females took more than 300 days to obtain the same length as the males. However, females have a greater maximal length of over 150 mm more than the males.

<table>
<thead>
<tr>
<th>SEX</th>
<th>α</th>
<th>β</th>
<th>σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>37.18</td>
<td>0.0150</td>
<td>0.6367</td>
</tr>
<tr>
<td>Male</td>
<td>39.30</td>
<td>0.0139</td>
<td>0.5829</td>
</tr>
</tbody>
</table>

Table 1: Parameter estimates of inter-moult periods using log normal model

Figure 4(a) shows the carapace lengths of males and females increased but did not have significant discrepancy until the males have overtaken the females from the intermediate stage of the moulting process, thus the growth of female slowed down gradually. This implies that males have larger moult increments than females, indicating that males possess greater asymptotic length compared to females. As for Table 2 and Table 3, the growth rate of females is relatively faster than male lobsters, resulting in a lower length in $L_\infty$ for females. From biological perspective, once the females have reached maturity, their growth is said to be slowed down faster than males as more energy will be converted to reproduction purposes. From the plots, we notice both the Gamma processes and inverse Gaussian processes models share similar increment lengths.

Figure 4(a): Moult increment estimates for: (i) Gamma process and (ii) inverse Gaussian process

Figure 4(b) displays the variance for both processes respectively. However, those figures do not account for the fluctuation of the environment variability. Based on pre-moult length, variances of males were greater than females. However, the variability of the inverse Gaussian process is relatively higher than that of the Gamma process, specifically during the first 50 days. Therefore, some diagnostic tests ought to be implemented to determine which model is better in the overall context.

<table>
<thead>
<tr>
<th>SEX</th>
<th>$L_\infty$</th>
<th>λ</th>
<th>k</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>175.38</td>
<td>0.6199</td>
<td>0.00083</td>
</tr>
<tr>
<td>Male</td>
<td>258.45</td>
<td>0.6179</td>
<td>0.00056</td>
</tr>
</tbody>
</table>

Table 2: Parameter estimates for Gamma process

<table>
<thead>
<tr>
<th>SEX</th>
<th>$L_\infty$</th>
<th>α</th>
<th>k</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>178.20</td>
<td>0.00017</td>
<td>0.00087</td>
</tr>
<tr>
<td>Male</td>
<td>196.34</td>
<td>0.00016</td>
<td>0.00086</td>
</tr>
</tbody>
</table>

Table 3: Parameter estimates for inverse Gaussian

Figure 4(b): Expected variance of (i) Gamma process versus (ii) inverse Gaussian process

Once the parameters for the moult increments and intermoult periods processes are obtained, it is easy to integrate both processes as stated in eqn (8) via Monte Carlo approach. The step-like growth trajectories are made up of moult increments associated with time intervals between moults. The purpose of this study is to simulate a large sample of observations to estimate the population mean growth curve. The simulation studies were implemented in R. Eventually, we find out some interesting features from the plots as shown in Figure
Foo and Wang, Stochastic growth models for analyzing crustacean data

5, the growth pattern shows a monotonically increasing function. In this context, it implies that the aforementioned model and approach is plausible and sensible in the real data analysis. Comparatively from the plots, the males surpass the females to have a bigger size (length, mm). The drawback of the previous subordinated stochastic process is where the assessment of the size-at-age only considered for the growth length (increment) at a given time without considering moult interval. We could make some improvement of the underlying model by deriving a joint density function between moult increment and moult interval by applying the Monte Carlo approach.

![Simulation of growth for 20 tank spiny lobsters by Gamma process](image)

**Figure 5:** Simulation of growth for 20 tank spiny lobsters by Gamma process

The most commonly used approach to describe the relationship between length and age in fisheries is known as the von Bertalanffy growth function (VBGF). However, this deterministic function does not account for the individual variability. In view of that, Russo et al., (2009) provided a new solution to curb the above problems where a subordinator is included in VBGF for stochastic models. The model considers variability of individuals as well as environmental sources.

Numerous approaches are available to estimate the growth parameters. In addition to tank data, tag-recapture experiments are commonly used for growth parameter estimation of the individuals. For animals reared in the tank, the moulting times can be observed directly, however, the exact moulting time is unknown during the periods from when the tagged animals are released to the sea until they are recaptured.

**ACKNOWLEDGEMENTS**

The authors acknowledge Darren Dennis (CSIRO) who has provided us the lobster dataset and and Dr Liya Fu for interesting discussion on this work.

**REFERENCES**


