Generalized growth models for aquatic species with an application to blacklip abalone (*Haliotis rubra*)

Luke R. Lloyd-Jones*, You-Gan Wang and Warwick J. Nash

Luke R Lloyd-Jones and You-Gan Wang. Centre for Applications in Natural Resource Mathematics (CARM), School of Mathematics and Physics, University of Queensland, St Lucia, Qld 4067, Australia (email: l.lloydjones@uq.edu.au, you-gan.wang@uq.edu.au).

Warwick J. Nash. Queensland Department of Agriculture, Fisheries and Forestry, Ecosciences Precinct, Joe Baker St, Dutton Park, Queensland, 4102, Australia (email: warwick.nash@daff.qld.gov.au).

* **Corresponding author**:

Luke Lloyd-Jones

Postal address: Priestley Building 67, St Lucia, Department of Mathematics, The University of Queensland, Brisbane, Qld, 4072, Australia

Ph: +61 424314627

Fax: +61 7 3365 1477

Email: luke.lloydjones@uqconnect.edu.au

Abstract: This paper presents a maximum likelihood method for estimating growth parameters for an aquatic species that incorporates growth covariates, and takes into consideration multiple tag-recapture data. Individual variability in asymptotic length, age-at-tagging, and measurement error are also considered in the model structure. Using distribution theory, the log-likelihood function is derived under a generalised framework for the von Bertalanffy and Gompertz growth models. Due to the generality of the derivation, covariate effects can be included for both models with seasonality and tagging effects investigated. Method robustness is established via comparison with the Fabens, improved Fabens, James and a non-linear mixed-effects growth models, with the maximum likelihood method performing the best. The method is illustrated further with an application to blacklip abalone (*Haliotis rubra*) for which a strong growth-retarding tagging effect that persisted3 for several months was detected.
Key Words: Aquatic species growth; Von Bertalanffy model; Gompertz model; Maximum likelihood method; Multiple tag-recapture data; Tagging effect

1. Introduction

Knowledge of how an aquatic species grows is fundamental to the stock assessment process. Growth, like other processes, is individual in nature and depends on many covariates such as seasonality, food availability, sex and many others. Growth models such as von Bertalanffy are often used to describe the mean growth of the population, however the growth process can be better understood by incorporating individuality and covariates into the model.

One of the difficulties with growth modelling is that the direct ageing of an aquatic species is often not possible. To solve this problem, fisheries rely on tag-recapture data to estimate the unknown parameters of a hypothesised representative model; for instance, the von Bertalanffy model with parameters (k, μ_∞). Determining adequate estimates of growth parameters can be a complex problem when individual variability is considered (**???**). Many methods have been developed to account for the complexity of individual variability. **?** provided a method for estimating the von Bertalanffy growth parameters via estimating functions that are unbiased and asymptotic. **?** developed a maximum likelihood method (ML method) that accounted for individual variability in L_∞ and the age at tagging using distribution theory. **?** extended this idea by developing a flexible ML method for general growth curves with less restrictive assumptions, but again only for single-recapture data.

For many aquatic species multiple recapture data can be acquired. Solving for the growth parameters of a particular growth model (such as the von Bertalanffy) using multiple-recapture data has either not been explored or is not easy to generalise mathematically. For instance, **?** explored multiple recaptures and found that the method for constructing the estimating functions for more than two recaptures, as well as unequal recaptures, was not clear. The method of **?** requires mathematically the use of a single recapture to derive the likelihood function and thus is constrained to single-recapture analyses. **?** used different growth functions under a maximum likelihood routine, but only with a single recapture. Given that many tagging studies often include multiple recaptures, it makes sense to take

advantage of these extra data. The inclusion of multiple recaptures should in theory improve growth estimates, because it allows for a better characterisation of each individuals growth trajectory by giving a greater number of growth snapshots; which is particularly important when individual variability is being modelled. More recaptures allow covariates such as seasonal growth to be modelled better because we have information across more seasons.

This paper derives a ML method to estimate growth parameters for general curves from simulated multiple recapture data; then applies it to a study of blacklip abalone (*Haliotis rubra*). The method rests on the ideas of distribution theory and requires the numerical integration of the joint distribution function for each individual. Individual variability in L_∞ and age at capture are taken into consideration, and given that the method is based on distribution theory the method should be asymptotically unbiased. To introduce other covariates into the growth model, a generalised von Bertalanffy growth model framework is used (**?**). The Gompertz model is also used and covariates are introduced via a modification of the model presented in **?**. These frameworks allow for the incorporation of explanatory variables such as seasonality into the growth model.

For abalone it has been conjectured that juvenile abalone do not follow a von Bertalanffy growth model (**???**). There are many potential solutions to this problem; for instance, to model the juvenile component with a separate (linear) model, exclude the data that do not follow a von Bertalanffy curve, or use a different growth model. To keep the model continuous the latter two approaches are investigated here, with the Gompertz model used as an alternative. The Gompertz model has some precedence in the literature when modelling growth of blacklip abalone, and data analysis also suggests it is a reasonable choice (**??**). For *Haliotis rubra* both **?**, and **?** established that growth for juvenile abalone (those < 80 mm) is best described by using a linear model. **?** also established that once adulthood is reached then the von Bertalanffy curve models growth well. The straight-line growth observed for the smaller size classes of abalone has been observed or hypothesised for haliotids in a number of other studies such as **?**, **?**, and **?**. However, **?** observed early growth of haliotids to be

non-linear.

Covariates can be included into the growth model to account for their effects. **?** concluded that for long-lived fish the von Bertalanffy model adjusted for seasonal growth is preferable to the logistic or Gompertz models. It has been shown that for juvenile blacklip abalone there is a strong seasonal effect, and its inclusion improved the model fit for juvenile blacklip abalone (**?**). It is also conjectured that blacklip abalone from the Tasmanian (Australia) fishery could experience periods of no growth in the winter months. The generalised von Bertalanffy growth model will be modified to account for the potential no-growth period of blacklip abalone.

It is hypothesised that the tagging process has an effect on the growth of blacklip abalone. **?** remarked that the most practical way to account for the effect of tagging is to quantify tagging effects concurrently with growth parameters. **?** used the generalised von Bertalanffy model framework to quantify the time taken to recover from tagging using a link function. The function represents a recuperation curve, and models a smooth transition to normal growth rather than the step-like function of **?**. If tagging leads to a period of suboptimal growth for blacklip abalone then the growth estimates are likely to be biased if this is not taken into consideration. We will investigate growth recovery after tagging of a population of blacklip abalone using the ML method under both the von Bertalanffy and Gompertz models. The principal aim of this paper is to present a ML method that uses multiple recaptures to model the effects of covariates (such as tagging) on growth for general growth curves.

2. Materials and methods

2.1. Data

To illustrate the method, multiple tag-recapture data from blacklip abalone were analysed. The tag-recapture data set for blacklip abalone, *Haliotis rubra*, was gathered from a shallow study site at George III Rock, in southern Tasmania, Australia. The mark-recapture

study included more than 1400 individuals with recapture frequency ranging from 1 to 5. Letting L_1 to L_6 be the lengths of individuals at each recapture, and T be the time at liberty, the data are summarised in Table **??**. The data appeared to exhibit significant negative growth; therefore, a cutoff of -10 mm was chosen to be reasonable to retain a sufficient error structure; all those less than this cutoff were removed from the data. The method models measurement error and thus the inclusion of significant negative growth will not lead to biased results. Only 0.5% of the data exhibited negative growth less than -4 mm and thus these data do not dominate. Capture times were also summarised to determine if seasonality could be accurately modelled. Over the ten years that the data were gathered there appears to be a sufficient distribution of times to attempt to model seasonality.

? established that for Tasmanian blacklip abalone the inverse logistic model best described growth for all age classes (adult $>$ 80 mm and juveniles $<$ 80 mm). Between 60 mm and 80 mm the data suggest that abalone begin to follow a von Bertalanffy growth curve (Fig. **??**). From other studies (e.g., **?**) it is hypothesised that blacklip abalone are particularly susceptible to the process of tagging, and that there will be a retardative effect of tagging on growth; and that the effect will be most present for those animals with short times at liberty, expressed as smaller growth rates. Those that have been at liberty longer will have had the chance to recover to normal growth. To investigate this, abalone with recapture intervals less than 4 months were compared with those at liberty for greater than 1.5 years at liberty.

(Table **??**, and Figure **??** NEAR HERE)

3. Multiple recapture ML method derivation

3.1. Derivation of the likelihood function

Methods used to estimate growth parameters for aquatic species from tag-recapture data often only make use of a single recapture. The following derivation aims to develop a ML method to incorporate multiple recaptures for each individual using distribution theory. For

each individual we would like to obtain the distribution

$$f_{\mathbf{L}}(\boldsymbol{l}) = f(l_1, l_2, \ldots, l_p), \tag{1}$$

where l_j $(j = 1, 2, \ldots, p)$ represents the length at capture j of the i-th $(i = 1, 2, \ldots, n)$ animal. The derivation of the model does not require a particular growth model to be used. In this paper we will use the generalised von Bertalanffy and Gompertz growth models. Assuming the general form from **?** for the von Bertalanffy growth function, the expected length at time t_j conditional on the asymptote being l_∞ is

$$\mu_j = l_0 + (l_\infty - l_0)\{1 - \exp[-z(t_0, t_j)]\}, \tag{2}$$

where

$$z(t_0, t) = \int_{t_0}^{t} g(\boldsymbol{\theta}, \mathbf{x}_u)\, du. \tag{3}$$

The function $g(\boldsymbol{\theta}, \mathbf{x}_u)$ represents the link function, which accommodates the effects of explanatory variables on the growth coefficient. The vector $\boldsymbol{\theta}$ contains the parameters that are related to the observed values $\mathbf{x}_u$. For example $g(\boldsymbol{\theta}, \mathbf{x}_u)$ could just be the canonical von Bertalanffy model component $-k(t - t_0)$, with k and t_0 being the parameters. Suppose t_0 is the hypothetical time when the length of the individual is 0, and $A = t_1 - t_0$ the relative age at time t_1 is a random variable with realised value a, then $l(t_0) = l_0 = 0$ and from (**??**),

$$\mu_j = l_\infty\{1 - \exp[-z(t_1 - a, t_j)]\}, \tag{4}$$

and the associated measured length is

$$l_j = \mu_j + \varepsilon_j, \tag{5}$$

where ε_j represents measurement error and is a random variable with assumed distribution $N(0, \sigma_\varepsilon^2)$. For the Gompertz model we use a length version of the model presented in **?** with an adaptation to fit the likelihood process (see **??**). The mean length is

$$\mu_j = \log(l_\infty)\{1 - \exp[-z(t_1 - a, t_j)]\}, \tag{6}$$

and the associated logged measured length is

$$\log[l_j(t)] = \log(l_\infty)\{1 - \exp[-z(t_1 - a, t_j)]\} + \log(\varepsilon_j). \tag{7}$$

The Gompertz model here is essentially equivalent to modelling von Bertalanffy on a log scale where $\log(\varepsilon)$ is still $N(0, \sigma_\varepsilon^2)$. Given these models the proposed likelihood function from conditional probability and the independence of L_∞ and A for all models with $\boldsymbol{l} = (l_1, \ldots, l_p)$ is

$$f_{\mathbf{L}}(\boldsymbol{l}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h_{L_\infty}(l_\infty) g_A(a) f_{\mathbf{L}|L_\infty, A}(\boldsymbol{l}|l_\infty, a)\, dl_\infty\, da, \tag{8}$$

where $\mathbf{L}|L_\infty, A \sim$ multivariate $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and represents measurement error. Here

$$\boldsymbol{\Sigma} = \text{diag}[\text{Var}(\varepsilon_1), \ldots, \text{Var}(\varepsilon_p)]. \tag{9}$$

The integral is performed for one individual and then the summed log represents the overall log likelihood. Given the distribution assumptions for the random variables L_∞ and A the difficulty of solving the integral of equation (**??**) can vary, and rarely can an explicit form for the integral be found. To evaluate (**??**) we first assess the distributions that we believe will model the random variables well. We choose $L_\infty \sim N(\mu_\infty, \sigma_\infty^2)$ and $A \sim \Gamma(\alpha, \beta)$ with $l_\infty \in \mathbb{R}$ and $a \in \mathbb{R}_+$. It is noted that any distributions can be chosen for the random variables, but naturally the choice of distributions must make sense with respect to the variables they

model. Given the assumed distributions equation (**??**) is now defined. Eliminating variables is of interest since we may be able to make use of other numerical techniques if the integral is over just one variable. Given $L_\infty \sim N(\mu_\infty, \sigma^2_\infty)$ (or $\log(L_\infty) \sim N(\mu_{lg\infty}, \sigma^2_{lg\infty})$ for Gompertz) the likelihood integral function can be written in terms of the random variable A

$$f(\boldsymbol{l}) = \int_0^\infty g(a) f(\boldsymbol{l}|a)\, da. \tag{10}$$

The data $\mathbf{L}$ can be expressed under the von Bertalanffy and Gompertz models as an affine transformation of the random vector $(L_\infty, \varepsilon_1, \varepsilon_2, \ldots, \varepsilon_p)^T$ or the logged equivalent for the Gompertz model. Since each of the random variables in the random vector $(L_\infty, \varepsilon_1, \varepsilon_2, \ldots, \varepsilon_p)^T$ is independent, the resultant covariance matrix is a diagonal matrix of the variance for each random variable. Using the properties of the affine transformation we can find the distribution of the random variable $\mathbf{L}|A$, which is independent of the variable L_∞. Letting

$$\mathbf{L} = \begin{pmatrix} g_1 & 1 & 0 & 0\ldots & 0 \\ g_2 & 0 & 1 & 0\ldots & 0 \\ g_3 & 0 & 0 & 1\ldots & 0 \\ \vdots & & & & \\ g_p & 0 & 0 & 0\ldots & 1 \end{pmatrix} \begin{pmatrix} L_\infty \\ \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_p \end{pmatrix}, \; G = \begin{pmatrix} f(t_1; A, \boldsymbol{\theta}) \\ f(t_1, t_2; A, \boldsymbol{\theta}) \\ f(t_1, t_3; A, \boldsymbol{\theta}) \\ \vdots \\ f(t_1, t_p; A, \boldsymbol{\theta}) \end{pmatrix},$$

$$\mathbf{D} = \begin{pmatrix} g_1 & 1 & 0 & 0\ldots & 0 \\ g_2 & 0 & 1 & 0\ldots & 0 \\ g_3 & 0 & 0 & 1\ldots & 0 \\ \vdots & & & & \\ g_p & 0 & 0 & 0\ldots & 1 \end{pmatrix}, \text{ and } \boldsymbol{\Sigma}_z = \begin{pmatrix} \sigma^2_\infty & 0 & \ldots & 0 \\ 0 & \mathrm{Var}(\varepsilon_1) & \ldots & \\ 0 & 0 & \ldots & 0 \\ \vdots & & & \\ 0 & 0 & \ldots & \mathrm{Var}(\varepsilon_p) \end{pmatrix}.$$

The resulting distribution does not depend on L_∞ or $\log(L_\infty)$, and has the form

$$\mathbf{L}|A \sim \text{multivariate } N(\mu_\infty G, \mathbf{D}\boldsymbol{\Sigma}_z\mathbf{D}^T), \tag{11}$$

for the von Bertalanffy model. For the Gompertz model the mean and covariance matrix are replaced with their log scale equivalents. The above is true because given a set of realised values from A, the matrix $\mathbf{D}$ is a constant matrix for each individual. Now the integral (**??**) is over just one variable and can be evaluated with any numerical integration technique for each individual. This is repeated for all individuals and the sum of the logged integrals forms the objective function.

3.2. Incorporation of growth effects

3.2.1. Seasonal effect

Using the general form of the above model, explanatory variables for the growth coefficient can be added into the model. To model seasonal growth for blacklip abalone, an asymmetric seasonal model as well as the canonical seasonal model is used. The asymmetric model is hypothesised to be able to model a possible period of no growth better than the canonical model. Modelling the potential no-growth period is achieved by restraining the growth link function to a positive range during the integration step of the model derivation.

To account for asymmetry in the growth season, two more parameters for the growth rate are added to the growth link function

$$g(\boldsymbol{\theta}, t) = k + \theta_3 \cos(2\pi t) + \theta_4 \sin(2\pi t) + \theta_5 \cos(4\pi t) + \theta_6 \sin(4\pi t).$$

The canonical seasonal-growth model (**?**) is the first two components of the above link function. To ensure that negative growth is not modelled we let

$$z(t_1 - A, t_j) = \int_{t_1 - A}^{t_j} \max[g(\boldsymbol{\theta}, \mathbf{x}_u), 0]\, du. \tag{12}$$

The solution for one section of the recursive cycle of the canonical model can be seen in **??**. To implement the integral in practice the other complete positive components of the trigonometric function must be added to obtain the total integral.

3.2.2. Tagging effect

A tagging effect may be of interest if a species is suspected to have a growth rate that is susceptible to the physical handling of the tagging process. Following **?**, the link function to investigate such an effect is

$$g(\boldsymbol{\theta}, t) = k_0 + \theta_1\{1 - \exp[-\theta_2(t - t_1)]\}, \quad t \geqslant t_1.$$

However, before t_1 we have $k = k_0 + \theta_1$ and as we are required to integrate from $t_1 - A$ we define the link function to be

$$g(\boldsymbol{\theta}, t) = \begin{cases} k_0 + \theta_1, & t < t_1, \\ k_0 + \theta_1\{1 - \exp[-\theta_2(t - t_1)]\}, & t \geqslant t_1, \end{cases}$$

where t_1 is the time at first capture relative to January 1. Here k_0 represents the curvature just after tagging with the true or recovered curvature being equal to $k = k_0 + \theta_1$ as time goes to infinity. To make use of this function in the method the integral of the time period is calculated. Therefore,

$$\begin{aligned} z(t_1 - A, t_j) &= z(t_1 - A, t_1) + z(t_1, t_j), \\ z(t_1 - A, t_1) &= kA, \quad \text{and} \\ z(t_1, t_j) &= \int_{t_1}^{t_j} g(\boldsymbol{\theta}, t)\, dt \\ &= kT - \frac{\theta_1}{\theta_2}\{1 - \exp[-\theta_2 T]\}, \end{aligned}$$

where $T = t_j - t_1$.

3.2.3. Combined model

To avoid possible bias we model the tagging and seasonal effect as a linear combination. Bias may arise from the fact that if only one of the effects is modelled then there is the potential for an effect to be found when there is none, arising from the other effect. For example, if only tagging is modelled there is the potential for the model to find a tagging effect when in fact the model has found the natural slowing of growth in the winter months. The link function will therefore be a combination of the canonical seasonal model and the tagging effect. The link function is

$$(13) \qquad g(\boldsymbol{\theta}, t) = \begin{cases} k_0 + \theta_1 + \theta_3 \cos(2\pi t) + \theta_4 \sin(2\pi t), & t < t_1, \\ k_0 + \theta_1\{1 - \exp[-\theta_2(t - t_1)]\} + \theta_3 \cos(2\pi t) + \theta_4 \sin(2\pi t), & t \geqslant t_1. \end{cases}$$

Again we integrate to make use of this function in the method. The integrals are

$$\begin{aligned} z(t_1 - A, t_1) &= \int_{t_1 - A}^{t_1} k + \theta_3 \cos(2\pi t) + \theta_4 \sin(2\pi t)\, dt \\ &= kA + \frac{\theta_3}{2\pi}\{\sin(2\pi t_1) - \sin[2\pi(t_1 - A)]\} - \frac{\theta_4}{2\pi}\{\cos(2\pi t_1) - \cos[2\pi(t_1 - A)]\}, \text{ and} \\ z(t_1, t_j) &= \int_{t_1}^{t_j} k - \theta_1 \exp[-\theta_2(t - t_1)] + \theta_3 \cos(2\pi t) + \theta_4 \sin(2\pi t)\, dt \\ &= kT - \frac{\theta_1}{\theta_2}\{1 - \exp[-\theta_2(t_j - t_1)]\} + \frac{\theta_3}{2\pi}[\sin(2\pi t_j) - \sin(2\pi t_1)] - \frac{\theta_4}{2\pi}[\cos(2\pi t_j) - \cos(2\pi t_1)]. \end{aligned}$$

3.3. Computational implementation

Numerical integration is computationally intensive for large data sets. Prospective slow points in the code, for example, matrix inverse calculation, unnecessary loops and determinants are optimised to improve computational speed. The structure of (**??**) allows for analytical representations of the inverse and determinant to be found for the multivariate normal distribution evaluation. The variance of each of the measurement error terms is assumed to be the same. Although all measurements in the tag-recapture study were made by an experienced field team, the assertion that the measurement error terms have equal variance across each recapture is dubious given the time period the study was made over. Unequal variances for the measurement error terms is a potential point for further generalisation. However, this assumption greatly improves the speed and implementation of the code due to the ability to solve for an analytical inverse. To solve for the inverse we notice that

$$\mathbf{\Sigma}_{L|A} = \mathbf{D}\mathbf{\Sigma}_z\mathbf{D}^T = \sigma_\varepsilon^2\mathbf{I} + \sigma_\infty^2 GG^T,$$

has known form. This is a classical matrix form and its inverse has known solution. The classical form is

$$(\mathbf{D} + uv^T)^{-1} = \mathbf{D}^{-1} - \frac{\mathbf{D}^{-1}uv^T\mathbf{D}^{-1}}{1 + v^T\mathbf{D}^{-1}u}.$$

Therefore,

$$\mathbf{\Sigma}_{L|A} = \sigma_\varepsilon^2[\frac{\sigma_\infty^2}{\sigma_\varepsilon^2}GG^T + I],$$
$$(\mathbf{\Sigma}_{L|A})^{-1} = \sigma_\varepsilon^{-2}[I - \frac{\sigma_\infty^2 GG^T}{\sigma_\varepsilon^2 + \sigma_\infty^2 G^T G}].$$

The determinant of the covariance matrix also has a known form and contributes to computational efficiency. The known form is

$$|\mathbf{D} + uv^T| = |\mathbf{D}|(1 + G^T\mathbf{D}^{-1}G),$$

which implies that for our matrix

$$|\mathbf{\Sigma}_{L|A}| = \sigma_\varepsilon^{2p}\left(1 + \frac{\sigma_\infty^2}{\sigma_\varepsilon^2}G^TG\right).$$

These two mathematical representations simplify the code and increase the speed of computation greatly.

The incorporation of a non-negative growth seasonal curve adds computational complexity to the method, as one must be able to solve for the roots of the function to perform the analytical integral (as seen in section **??**). This can be done for the yearly cycle, but is unlikely for the 6 month cycle, as it requires the solving of the roots of a quartic polynomial. This is very difficult especially with a function of this structure. Therefore, a numerical integration technique (Simpson's method) is used during the computation to evaluate the truncated integral (**??**). This adds computational complexity, but has the added benefit of being able to deal with any $g(\boldsymbol{\theta}, t)$ link function.

To use numerical integration for the whole likelihood function we state explicitly the

integral form. The function to integrate is

$$f(\boldsymbol{l}) = \int_0^\infty g(a) f(\boldsymbol{l}|a)\, da, \text{ where}$$
$$g(a) = \frac{\beta^\alpha}{\Gamma(\alpha)} a^{\alpha-1} \exp(-\beta a), \text{ and}$$
$$\mathbf{L}|A \sim MVN(\mu_\infty G, \mathbf{D}\boldsymbol{\Sigma}_\mathbf{z}\mathbf{D}^\mathbf{T}), \text{ resulting in}$$
$$f(\boldsymbol{l}|A=a) = (2\pi)^{\frac{-k}{2}} |\boldsymbol{\Sigma}_{\mathbf{L}|\mathbf{A}}|^{-1/2} \exp\left[-\frac{1}{2}(\boldsymbol{l}-\mu_\infty G)^T {\boldsymbol{\Sigma}_{\mathbf{L}|\mathbf{A}}}^{-1} (\boldsymbol{l}-\mu_\infty G)\right], \text{ and the final integral}$$
$$f(\boldsymbol{l}) = \int_0^\infty \frac{\beta^\alpha}{\Gamma(\alpha)} a^{\alpha-1} \exp(-\beta a)(2\pi)^{\frac{-k}{2}} |\boldsymbol{\Sigma}_{\boldsymbol{l}|\mathbf{A}}|^{-1/2} \exp\left[-\frac{1}{2}(\boldsymbol{l}-\mu_\infty G)^T \boldsymbol{\Sigma}^{-\mathbf{1}}_{\mathbf{L}|\mathbf{A}} (\boldsymbol{l}-\mu_\infty G)\right] da.$$

Numerical integration is used to evaluate this integral for each individual and form the log likelihood function. To integrate over the half infinite region we do the following change of variable. Let

$$a = \frac{1-u}{u}, \text{ and}$$
$$f(\boldsymbol{l}) = \int_0^1 g\left(\frac{1-u}{u}\right) f\left(\boldsymbol{l}|\frac{1-u}{u}\right) \frac{1}{u^2}\, du.$$

We can then use the integration method of our choice. Simpson's method was used as we found the difference in numerical speed between Gaussian quadrature and Simpson's method to be minimal. With these components a numerical integration routine for the multiple recapture ML method was written in the R programming language (www.r-project.org/). The integration results form the log likelihood function and then the Akaike Information Criterion (AIC) (**?**) can be calculated for each model. Once the parameters for each model are found one can also back-calculate for the realised relative ages for each individual and the asymptotic length for each individual. The mathematical approach used is outlined in **??**. Once the realised values are calculated the error structure can be evaluated and the mean squared error (MSE) and cumulative absolute percent error (CAPE) are used as goodness of fit measures.

4. Results

4.1. Simulated data

Simulated data were generated in various forms to test the effectiveness of the proposed method. For the von Bertalanffy and Gompertz models the following were used: $L_\infty \sim N(150, 15^2)$ for von Bertalanffy, $L_\infty \sim \log N(150, 15^2)$ for Gompertz, $k = 0.5$, and the times at liberty were generated from a $\Gamma(1, 1)$ distribution. The relative ages were generated from a $\Gamma(3, 5)$ distribution. Measurement error was also incorporated by setting σ_ε^2 equal to 1, 4 and 16, to model different levels of error for the von Bertalanffy model. For the Gompertz model $\log(\varepsilon) \sim N(0, \sigma_\varepsilon^2)$ with σ_ε^2 set to 0.01, 0.02, and 0.04. The simulation was performed 50 times with each run including 100 animals with 3 recaptures per animal. One thousand time points were generated for the numerical integration as this gave a satisfactory accuracy to the integral. For the simulations of the single-recapture methods the first and last captures were used.

Classic methods were used to provide a baseline comparison for the ML method. The comparison methods included those of **?**, **?**, **?**, and a non-linear mixed effect (NLME) model (**?**). The non-linear mixed effects model assumes an unknown initial age that is log-normally distributed. The NLME method takes into consideration all recaptures and therefore it is a good comparative model for the ML method; it also provides estimates for the initial age distribution and measurement error. Given this, the simulated data were generated with the above values and $A \sim$ logN(0.5,0.5) for the NLME method.

The simulations showed that the parameter estimates from the NLME and MLE methods are unbiased (Table **??**). However for the NLME and MLE methods there is an decrease in the precision as the measurement error increases. The Fabens', improved Fabens and James' methods appear to be more accurate at higher levels of measurement error, but less precise. It is noted that for the single recapture methods the first and last captures are used. However, for all levels of error the single-recapture methods do not exhibit the same accuracy and precision as the NLME and MLE methods. The simulation study has shown

both the NLME and MLE methods to be very reliable, and capable of arriving at estimates for all parameters with excellent accuracy and precision.

(Table **??** near here.)

4.2. Blacklip abalone tag-recapture data

Tag-recapture data from blacklip abalone (*Haliotis rubra*) were analysed to further demonstrate the method. The von Bertalanffy model was run on all the data for comparison with the Gompertz model, and for those abalone with length < 70 mm removed. To establish the cutoff of 70 mm, a generalised additive model (GAM) was fitted to a plot of growth rate versus length at tagging. This was done to determine the overall growth shape of the data and to choose a cutoff for those abalone that do not follow the von Bertalanffy curve (Fig. **??**). A cutoff of 70 mm was chosen since this was the point at which the plot visually turns linear and the best tradeoff point between not misrepresenting the von Bertalanffy curve and keeping as many data as possible. When the Gompertz curve is used all samples are included. Given that the growth rate appears to increase and then decrease linearly, the use of the Gompertz model for the whole population was a logical choice (Fig. **??**).

As a first investigation, the method was used to fit the von Bertalanffy and Gompertz models without effects (Table **??**). This was done to provide a baseline for further investigation. With the basic model parameters estimated, the complexity was increased by incorporating effects. Firstly, seasonal effects for a yearly cycle were estimated with no truncation to 0. This provided both a computational and parameter baseline to compare the more complex seasonal models. This model was then truncated to attempt to model the effects of a potential no-growth period during the winter months. The AIC of the truncated model did not improve when compared to the original model (no-truncation); when the extra computational time required to fit this model is also taken into account, the no-truncation model is preferred. In a further attempt to quantify the potential for a no-growth period, an asymmetric form of the growth model was assumed. No improvement in the objective function was found and both the resultant plot and parameters did not make sense. There-

fore, seasonal growth is always positive according to this model and the chosen final model was the canonical seasonal model. The results of the seasonal assessment show that growth for blacklip abalone peaks in March and is at its lowest in September (Fig. **??**) for the von Bertalanffy model ($<$ 70 mm removed), and peaks on the in February and is at its lowest in August for the Gompertz model (Fig. **??**). We refrain from specific dates as would be misleadingling precise given the time period the data were collected over. The basic seasonal model did not improve the AIC over the basic no-effect model for the von Bertalanffy but did for the Gompertz model (Table **??**). The AIC score is negative for the Gompertz model due to the narrow nature of the $\log(L_\infty)$ distribution each likelihood value is greater than 1, which when logged leads to a positive value and thus a negative negative log-likelihood. The MSE and CAPE showed for the von Bertalanffy model that goodness of fit does not improve, but it did for the Gompertz model (Table **??** and Fig. **??**).

The data were analysed to determine if a tagging effect was present. It was hypothesised that those abalone that have been at liberty longer will have had the chance to recover to normal growth. To investigate this, abalone with recapture intervals less than 4 months were compared with those at liberty greater than 1.5 years. There was a substantial difference in the average growth rate (Fig. **??**). To investigate this further the recaptures were divided into 4 length classes, as a proxy for age, and the differences in the growth rates were compared quantitatively (Table **??**). A one sided t-test (assumptions tested) was also run with the null hypothesis being that the means (growth rate) for each length class are the same. The alternative hypothesis was that the mean growth rate for those with shorter times at liberty would be less than those with longer times at liberty. Interestingly this was the case for all age classes. The larger length classes should be interpreted with caution because when an individual nears its asymptotic length the likelihood of measurement error influencing the growth rate for small time at liberty increases. Therefore, there may be a bias towards extreme values in the higher length classes due to the nature of the error, and a biased weighting towards negative growth may result. However, this is unlikely to be the case for

smaller animals because their growth rate is high enough such that the growth increment will be greater than measurement error. Therefore, the difference between the growth rates for younger animals points towards the influence of a tagging effect. The tagging effect was thus quantified using the MLE method.

The results from the MLE method show a strong tagging effect, with recuperation to normal growth after approximately 22-24 months for both the von Bertalanffy and Gompertz models (Fig. **??**). It is noted that these recoveries are modelled without the effect of seasonality. The tagging-effect model shows a substantial increase in k and a small decrease in the asymptotic length (Table **??**). This drastically changes the shape of the mean Gompertz and von Bertalanffy curves (Fig. **??**). The inclusion of the tagging effect also lowers the AIC relative to both the no-effect and seasonal model. When the MSE is compared between the normal and tagging models there is an improvement for both the Gompertz and von Bertalanffy models. However, the seasonal model performs better under the MSE for the Gompertz model.

The combined (tagging + season) model shows a concurrent influence from both the seasonal and tagging effects. Recuperation to normal growth takes less time, with normality resuming after 8 to 9 months for both the von Bertalanffy and Gompertz models (Fig. **??**). The recuperation curve represents a theoretical individual that has been tagged on January 1. Given that not all individuals are tagged on this date, the curves in Fig. **??** must be interpreted with caution. Again there is an increase in k relative to the other models, and a decrease in the asymptotic length (Table **??**). Based on both the MSE and AIC the combined model is either the second best or best model for both the von Bertalanffy and Gompertz models.

Comparing the performance of the von Bertalanffy and the Gompertz models is difficult when using only the AIC because they use different log-likelihoods. However, the MSE results allow comparison between the two models. The von Bertalanffy model was run on the whole data set so that comparison with the Gompertz model could be made. The

von Bertalanffy model appears to perform better under the MSE criteria and also when the cumulative absolute percent error is plotted (Table **??** and Fig. **??**). The cumulative absolute percent error shows that the von Bertalanffy models perform less well on smaller animals; they accrue more error up to 130 mm but then the Gompertz model accrues error after this length (Fig. **??**). The combined von Bertalanffy model has the lowest cumulative absolute percent sum and the best MSE when compared with the Gompertz model (on all the data). The normal von Bertalanffy model has the lowest MSE when only the subset of data is analysed, which may indicate that the juvenile data contribute significantly to explaining the influence of the covariates.

(Tables **??** and **??**, and Figures **??**, **??**, **??**, **??**, **??**, **??** and **??** near here.)

5. Discussion

The results suggest that the proposed multiple-recapture ML method is a powerful technique for the investigation of growth in aquatic species. The most advantageous component of this method is its generality. Generality is achieved through the ability to choose any initial distribution and growth link function, and to take advantage of extra data and the ML method. The ability to use any link function is due to the optional numerical integration of the link function; this option negates the need to present a link function that is analytically integrable. There is also the possibility to use different growth curves such as Gompertz, provided they fit the methodological framework. The method is also general in the sense that it does not require a set number of recaptures: one can use a different number of recaptures for each individual, as done in the abalone application. The problem of solving for multiple and uneven recaptures has been a weakness of many other tag-recapture methods (**?**). This method provides a solution that is statistically rigorous, incorporates individual variability, models k as a time-dependent variable to incorporate effects, produces unbiased estimates, and incorporates multiple-recapture data that have varying frequency of recaptures. The method also allows one to back calculate for each individual the asymptotic lengths and

relative ages.

The important assumptions are that k is a constant, A and L_∞ are independent, the random L_∞ is adequately described by the normal distribution, and the random variable A is adequately described by a gamma distribution. **?** concluded that applying a model that incorporates variability in just one of the growth parameters (μ_∞, k), even if the variability comes from the other parameter, still results in smaller biases than if one did not include individual variability at all. Given this conclusion and the known high correlation between μ_∞ and k, which points to a potential model over-parameterisation if variability in both μ_∞ and k is modelled, we conclude that the choice to model variability solely in L_∞ is a sound one. As discussed by **?** the assumption that L_∞ and A are independent is valid if we assume that the distribution of L_∞ is identical over all age classes, and that the distribution of L_∞ is the same for the length or age range considered for the population. The choice of the normal distribution to model L_∞ makes reasonable intuitive sense, and as seen in the methods section simplifies the mathematics and increases the speed of computation considerably. The precedence of **?** using the gamma distribution is the predominant reason for the choice of this distribution to model A. Given this historical precedent and the positive distribution property of the gamma distribution, we believe that the assumption of a gamma distribution to model A is reasonable.

? used linear models to fit a dataset consisting of two cohorts of juvenile size classes of blacklip abalone (*Haliotis rubra*). They found a persistent seasonal signal through the juvenile size range, with slow growth in winter and fast growth during summer. They used both statistical and biological model-selection criteria to judge the best model; the seasonal growth model had the best fit. In the present study of mature blacklip abalone we did not arrive at the same result as **?**: for the von Bertalanffy we found no improvement in the AIC when seasonality was included but we did for the Gompertz model. The same picture was also seen for the MSE. Growth was found to be at its fastest in the summer months and conversely at its slowest in winter. The lack of improvement may be attributed to the

need to model a tagging effect concurrently. Although the AIC score did not improve for the seasonal von Bertalanffy, the information gathered on the seasonal nature of blacklip abalone growth must also be taken into consideration. This could also be said for the inclusion of the tagging effect, which improves the AIC for the Gompertz and von Bertalanffy models. Although the tagging effect lowers the AIC value, compared with that of the seasonal model, the information gathered is different and explains a different biological process. Therefore, it is important, as **?** stated, that fit not be assessed solely from a quantitative perspective.

Xiao (1994) remarked that the introduction of a tagging effect into growth models is essential to reduce or eliminate biases introduced by the tagging process. The investigation of the tagging effect is particularly important for blacklip abalone growth, which have been shown to be vulnerable to handling during the tagging process. **?** and **?** have shown that New Zealand abalone, *Haliotis iris*, exhibited physiological stress in response to handling that caused the heart to stop pumping for up to an hour, and stop feeding for up to three weeks. Although a tagging effect for blacklip abalone growth was found here, **?** found no tagging effect for blacklip abalone. The results have shown that the tagging procedure leads to a substantial time of suboptimal growth. This study demonstrates that adjusting for a tagging effect leads to a marked increase in the estimate of k. This is important because if the model concluded that there was a tagging effect but the estimate of k was not much different from the base (Fabens, MLE) estimate of k, then the results would be of less practical application to fisheries assessment. It is noted that k can only be compared within model because k has a different meaning for the Gompertz model than it does for the von Bertalanffy model.

It is interesting that both the von Bertalanffy and Gompertz models predicted a very similar time and seasonal shape to recovery for the combined model. This concordance is also encouraging. However, due to the shapes of the two curves the growth trajectories are very different as seen in the growth curve plots. For both models the inclusion of seasonality (combined model) resulted in a small period of negative growth in the tagging recovery curve

(less so for the Gompertz model). Due to the shelled nature of abalone negative growth is unlikely. The period of negative growth shown in the recovery curve figures may be an accumulation of error in the code, or a reflection of the negative growth present in the data due to measurement error. Although measurement error was modelled, the method may not have been powerful enough to eliminate its effect. Nevertheless, the combined model provides the best fit for all three models in one of the selections criteria and thus is concluded to best explain the growth process for blacklip abalone.

From the results gathered the combined tagging and seasonal von Bertalanffy model is the best. The Gompertz model explains the data better for smaller lengths, but due to the nature of the Gompertz model the fit at larger lengths suffers. The von Bertalanffy model was hypothesised to model the growth of small animals poorly and the CAPE shows this. However, the von Bertalanffy appears to out perform the Gompertz model in terms of fit for the longer lengths. This analysis is an exposition of the well known strengths of each model. The ease of interpretation and precedence in the literature of the von Bertalanffy model may also be desirable to many researchers and thus contributes to it being the best model. Its ubiquity in the literature also allows for better comparison with other research. Although the Gompertz model has not performed well on this data set the problem may be the higher number of data at larger lengths, which leads to a large accruement of error. If there were many more juvenile data than adult data perhaps the Gompertz model would be more applicable. The process followed in this paper shows the level of consideration that must be taken when choosing to use not only a methodology (MLM), but the underlying growth model as well.

Blacklip abalone were recaptured between one and five times, therefore, it is possible that a similar growth-retarding effect may be caused by recapture. Thus, each recapture event may cause several months of slower growth. The effect of recapture on abalone growth will be investigated in future studies. This weakness of mark-recapture methods to investigate growth and ageing is avoided with direct-ageing techniques such as oxygen isotope analysis

of shells post mortem (**?**). Although the isotope method avoids the bias associated with a tagging effect, the ability to quantify the tagging effect, as described here, is another solution.

In a simulation study of error structures, **?** noted that while covariates can be included in the growth model to better represent other factors that may influence growth, the information content in single-recapture tagging data will generally be insufficient to support such complexity. This point establishes a further need to take advantage of the extra data that multiple recaptures provide. Although in fisheries the primary role of growth estimates is for stock assessment, modelling environmental effects on growth gives an interesting avenue to investigate biological concerns unique to a species; for example, in the present study the effect of tagging on the growth of an individual, or the change in growth due to migration as in **?**. Temperature is also an important biological factor affecting numerous biological processes and their associate phenomena, including the distribution, growth and reproduction of plants and animals (**?**). In a growth-specific model **?** estimated the effect of water temperature on the growth coefficient k. Given the generality of the method described here, a possible next step would be to incorporate such an effect into the current model framework.

The multiple-recapture ML method is an excellent tool for analysing growth of an aquatic species. We have shown how it can be used to incorporate effects on growth while providing unbiased estimates of key growth parameters. This allows for biological investigations into the species, instead of solely focusing on growth parameters for stock assessment. The results have shown that when tag-recapture data are being used to investigate growth one should be mindful of the influence of both seasonality and tagging together.

Appendix A. Derivation of Gompertz growth curve with covariates

For the inclusion of time-dependent growth covariates into the Gompertz growth model we will use the model version from **?**. Using a length interpretation

$$L(t) = L_\infty \left(\frac{L_0}{L_\infty}\right)^{\exp[-z(b_0,t)]} \varepsilon,$$

where $L(b_0) = L_0$, the other parameters are as usual and $\varepsilon \sim \log N(\mu_{lg\varepsilon}, \sigma^2_{lg\varepsilon})$. Taking the natural log of $L(t)$ we have

$$\log[L(t)] = \log(L_\infty) + \log(L_0)\exp[-z(b_0,t)] - \log(L_\infty)\exp[-z(b_0,t)] + \log(\varepsilon).$$

If we assume that L_0 is equal to 1 and assuming $\log(\varepsilon) \sim N(0, \sigma^2_\varepsilon)$ then the mean is

$$\log[L(t)] = \log(L_\infty)\{1 - \exp[-z(b_0,t)]\},$$

and the log of the measured length is

$$\log[L(t)] = \log(L_\infty)\{1 - \exp[-z(b_0,t)]\} + \log(\varepsilon).$$

This is a similar structure to the von Bertalanffy model but we are estimating the length on a log scale. If we assume that $L_\infty \sim \log N(\mu_\infty, \sigma^2_\infty)$ is distributed log-normally then the $\log(L_\infty)$ has a normal distribution $\log(L_\infty) \sim N(\mu_{lg\infty}, \sigma^2_{lg})$ thus we can use the affine transformation seen in the derivation for the von Bertalanffy terms with the error term on the log scale also being normally distributed. Therefore $\log(\mathbf{L})|A \sim MVN(\mu_{lg\infty}G, \mathbf{D\Sigma}_{lgz}\mathbf{D}^T)$ where $\mathbf{L}$ has multivariate log-normal distribution. The link functions used for the Gompertz model remain the same and thus the parameter results have the same meaning as the von Bertalanffy. The conversions between the means and variances of the normal random variables

and their log-normal counterparts are given. If

$$L_\infty \sim \log N(\mu_\infty, \sigma^2_\infty), \text{ and } \log(L_\infty) \sim N(\mu_{lg\infty}, \sigma^2_{lg}),$$

then estimating $\mu_{lg\infty}$ and σ^2_{lg} we can calculate the lognormal means and variances

$$\mu_\infty = \exp(\mu_{lg\infty} + \frac{\sigma^2_{lg}}{2}),$$

$$\sigma^2_\infty = \exp(2\mu_{lg\infty} + \sigma^2_{lg})[\exp(\sigma^2_{lg}) - 1],$$

and vice versa

$$\mu_{lg\infty} = 2\log(\mu_\infty) - \frac{1}{2}\log(\sigma^2_\infty + \mu^2_\infty),$$

$$\sigma_{lg} = -2\log(\mu_\infty) + \log(\sigma^2_\infty + \mu^2_\infty).$$

Assuming that $\log(\varepsilon) \sim N(0, \sigma^2_\varepsilon)$ then

$$\mu_{lg\varepsilon} = \exp(\frac{\sigma^2_\varepsilon}{2}), \text{ and}$$

$$\sigma^2_{lg\varepsilon} = \exp(\sigma^2_\varepsilon)[\exp(\sigma^2_\varepsilon) - 1].$$

For the Gompertz model the result is $\sigma_\varepsilon = 0.0195$. Therefore, the resultant multiplicative error term has distribution $\log N(\mu_{lg\varepsilon}, \sigma^2_{lg\varepsilon})$ with

$$\mu_{lg\varepsilon} = 1.00019, \text{ and}$$

$$\sigma^2_{lg\varepsilon} = 0.00038.$$

Appendix B. Integration of the link function for the seasonal model

Depending on the choice of link function this integral has varying difficulty. The difficulty rests on the ability to solve for the roots of the polynomial formed by the trigonometric functions. Due to the inability to solve for the roots of the quartic polynomial for the asymmetric model, only the analytical integral for the canonical seasonal model is derived. We would like to find when $g = g(\boldsymbol{\theta}, \mathbf{x}_t) = 0$, which occurs only if $\theta_3^2 + \theta_4^2 > k^2$ or the magnitude of the wave is greater than the positive offset k. Let $g = k+\theta_3 \cos(2\pi t)+\theta_4 \sin(2\pi t)$ and $G = \max(g, 0) = (g + |g|)/2$. To solve for the roots of g let

$$0 = k + \theta_3 \cos(2\pi t) + \theta_4 \sin(2\pi t).$$

Let $R = 2\pi t$

$$k + \theta_3 \cos(R) + \theta_4 \sqrt{1 - \cos^2(R)} = 0,$$

and $\cos(R) = u$

$$\begin{aligned}
&k + \theta_3 u + \theta_4 \sqrt{1 - u^2} = 0,\\
&\theta_4^2(1 - u^2) = (k + \theta_4 u)^2,\\
&\theta_4^2 - \theta_4^2 u^2 = k^2 + \theta_3^2 u^2 + 2k\theta_3 u,\\
&0 = u^2(\theta_4^2 + \theta_3^2) + u2k\theta_3 - \theta_4^2 + k^2.
\end{aligned}$$

Letting $a = \theta_4^2 + \theta_3^2, b = 2k\theta_3,$ and $c = -\theta_4^2 + k^2$. The roots are,

$$u_1 = \frac{-b + \sqrt{b^2 - 4ac}}{2a}, u_2 = \frac{-b - \sqrt{b^2 - 4ac}}{2a},$$
$$u = \cos(2\pi t),$$
$$r_1 = \frac{\arccos(u_1)}{2\pi},\ r_2 = 1 - \frac{\arccos(u_2)}{2\pi}.$$

Given these roots, for the *ith* individual we would like to find

$$\int_{t_1}^{t_2} G\ dt.$$

We are only required to solve for

$$\int_{t_1}^{t_2} |g|\ dt,$$

because we can easily calculate the analytical integral of the whole function $\xi(t_2, t_1) = \int_{t_1}^{t_2} g\ dt$. The integral depends on the ordering of the time points relative to roots of the function. If we analyse the case where one section of the integral of g is positive between r_1 and r_2, then we cover all orderings of the roots and time points by letting w_1, w_2, w_3, w_4 be the ordered values of t_1, t_2, r_1, r_2 and

$$\int_{t_1}^{t_2} |g|\ dt = \left| \int_{w_1}^{w_2} g\ dt - \int_{w_2}^{w_3} g\ dt + \int_{w_3}^{w_4} g\ dt \right| = |\xi(w_2, w_1)| + |\xi(w_3, w_2)| + |\xi(w_4, w_3)|,$$

conditioning on the fact that if $t_2 < r_2$ then $\int_{t_2}^{r_2} g\ dt = 0$, if $r_1 < t_1$ then $\int_{r_1}^{t_1} g\ dt = 0$, and if $t_2 < r_1$ then $\int_{t_2}^{r_1} g\ dt = 0$ and $\int_{r_1}^{r_2} g\ dt = 0$.

Appendix C. Estimation of the realised values of of A and L_∞

We follow the ideas of **?** to attempt to find the realised values of A and L_∞ for each individual. **?** suggest calculating the estimated posterior distribution using Bayes' theorem. For an individual fish

$$g(A = a|\boldsymbol{l}) = \frac{f(\boldsymbol{l}|a)g(a)}{f(\boldsymbol{l})},$$

and then the expected value corresponds to the integral

$$\mathbb{E}(A|\boldsymbol{l}) = \int_0^\infty a\frac{f(\boldsymbol{l}|a)g(a)}{f(\boldsymbol{l})}da.$$

This back calculation was performed for each model using Simpson's method over one variable. A similar calculation was done for the expected asymptotic lengths for each individual with the following formula:

$$\mathbb{E}(L_\infty|\boldsymbol{l}) = \frac{\int_0^\infty \mathbb{E}(L_\infty|\boldsymbol{l}, a)f(\boldsymbol{l}|a)g(a)da}{f(\boldsymbol{l})}, \text{ where}$$

$$\mathbb{E}(L_\infty|\boldsymbol{l}, a) = \mu_\infty + \frac{\sigma^2_\infty}{\sigma^2_\varepsilon + \sigma^2_\infty(\sum_{j=1}^p f_j^2)}\left[\sum_{j=1}^p f_j(l_j - \mu_\infty f_j)\right],$$

with $f_j = 1 - \exp[z_j(t)]$ and $j = 1, \ldots, p$. Again the numerical integral is calculated for each individual using Simpson's method. These integrals are performed for each model structure and with the ML estimates substituted in.

Acknowledgements

The authors would like to thank the Queensland Department of Agriculture, Fisheries and Forestry and the Australian Research Council for the scholarship that funded this research. We also thank the Tasmanian government, the Australian Fisheries Research and Development Corporation and the Tasmania abalone research team for supporting the original study.

References

Akaike, H., 1974. A new look at the statistical model identification. Automatic Control, IEEE Transactions. 19, 716–723.

Catchpole, E.A., Freeman, S.N., Morgan, B.J.T., Nash, W.J., 2001. Abalone I: Analyzing mark-recapture-recovery data incorporating growth and delayed recovery. Biometrics. 57, 469–477.

Chen, H.C., 1984. Recent innovations in cultivation of edible molluscs in Taiwan, with special reference to the small abalone *Haliotis diversicolor* and the hard clam *Meretrix lusoria*. Aquaculture. 39, 11–27.

Cox, K.W., 1962. California abalones, family Haliotidae. California Fish and Game Fish Bulletin 1, 8–133.

Day, R.W., Fleming, A.E., 1992. The determinants and measurement of abalone growth. Abalone of the world: biology, fisheries and culture. 1, 141–168.

Eveson, J.P., Polacheck, T., Laslett, G.M., 2007. Consequences of assuming an incorrect error structure in von Bertalanffy growth models: a simulation study. Canadian Journal of Fisheries and Aquatic Sciences. 64, 602–617.

Fabens, A.J., 1965. Properties and fitting of the von Bertalanffy growth curve. Growth. 29, 265–289.

Forster, G.R., 1967. The growth of *Haliotis tuberculata*: results of tagging experiments in Guernsey 1963-65. Journal of the Marine Biological Association of the United Kingdom. 47, 287–300.

Gamito, S., 1998. Growth models and their use in ecological modelling: an application to a fish population. Ecological Modelling. 113, 83–94.

Garcia-March, J.R., Marquez-Aliaga, A., Wang, Y.G., Surge, D., Kersting, D.K., 2011. Study of *Pinna nobilis* growth from inner record: How biased are posterior adductor muscle scars estimates? Journal of Experimental Marine Biology and Ecology. 407, 337–344.

Harrison, A.J., Grant, J.F., 1971. Progress in abalone research. Tasmanian Fisheries Research 5, 1–10.

Helidoniotis, F., Haddon, M., 2012. Growth model selection for juvenile blacklip abalone (*Haliotis rubra*): assessing statistical and biological validity. Marine and Freshwater Research. 63, 23–33.

Helidoniotis, F., Haddon, M., Tuck, G., Tarbath, D., 2011. The relative suitability of the von Bertalanffy, Gompertz and inverse logistic models for describing growth in blacklip abalone populations (*Haliotis rubra*) in Tasmania, Australia. Fisheries Research. 112, 13–21.

Jackson, C.J., Wang, Y.G., 1998. Modelling growth rate of *Penaeus monodon* Fabricius in intensively managed ponds: effects of temperature, pond age and stocking density. Aquaculture Research. 29, 27–36.

James, I.R., 1991. Estimation of von Bertalanffy growth curve parameters from recapture data. Biometrics. 47, 1519–1530.

Koike, Y., Flassch, J., Mazurier, J., 1979. Biological and ecological studies on the propagation of the ormer, *Haliotis tuberculata* Linnaeus. II. Influence of food and density on the growth of juveniles. La Mer. 17, 43–52.

Laslett, G.M., Eveson, J.P., Polacheck, T., 2002. A flexible maximum likelihood approach for fitting growth curves to tag-recapture data. Canadian Journal of Fisheries and Aquatic Sciences. 59, 976–986.

Lloyd-Jones, L.R., Wang, Y.G., Courtney, A.J., Prosser, A.J., Montgomery, S.S., 2012. Latitudinal and seasonal effects on growth of the Australian eastern king prawn (*Melicertus plebejus*). Canadian Journal of Fisheries and Aquatic Sciences. 69, 1525–1538.

Mayfield, S., Mundy, C., Gorfine, H., Hart, A.M., Worthington, D., 2012. Fifty years of sustained production from the Australian abalone fisheries. Reviews in Fisheries Science. 20, 220–250.

McShane, P.E., Smith, M.G., Beinssen, K.H.H., 1988. Growth and morphometry in abalone (*Haliotis rubra* Leach) from Victoria. Marine and Freshwater Research. 39, 161–166.

Momma, H., 1980. Studies on the variation of the abalone I. On the growth of the different aged young abalone (in Japanese). Aquaculture. 28, 142–55.

Morse, D.E., 1984. Biochemical and genetic engineering for improved production of abalones and other valuable molluscs. Aquaculture. 39, 263–282.

Peck, L.S., 1989. Feeding, growth and temperature in the ormer, *Haliotis tuberculata* l. Progress in Underwater Science. 14, 95–107.

Pitcher, T.J., MacDonald, P.D.M., 1973. Two models for seasonal growth in fishes. Journal of Applied Ecology. 10, 599–606.

Poore, G.C., 1972. Ecology of New Zealand abalones, *Haliotis species* (Mollusca: Gastropoda) 3. growth. New Zealand Journal of Marine and Freshwater Research. 6, 534–559.

Prince, J.D., 1991. A new technique for tagging abalone. Marine and Freshwater Research. 42, 101–106.

Prince, J.D., 2004. The decline of global abalone (genus *Haliotis*) production in the late twentieth century: is there a future. Stock Enhancement and Sea Ranching. 1, 427–443.

Prince, J.D., Sellers, T.L., Ford, W.B., Talbot, S.R., 1988. A method for ageing the abalone *Haliotis rubra* (Mollusca: Gastropoda). Marine and Freshwater Research. 39, 167–175.

Ragg, N.L.C., Taylor, H.H., 2006. Heterogeneous perfusion of the paired gills of the abalone *Haliotis iris* Martyn 1784: an unusual mechanism for respiratory control. Journal of Experimental Biology. 209, 475–483.

Ragg, N.L.C., Taylor, H.H., Behrens, J., 2000. Stress and weight loss associated with handling in the blackfoot abalone, *Haliotis iris*. Journal of Shellfish Research. 19, 528–529.

Sakai, S., 1962. Ecological studies on the abalone *Haliotis discus hannai* Ino-I. Experimental studies on the food habit. Bulletin of the Japanese Society of Scientific Fisheries 28, 766–779.

Shepherd, S.A., Hearn, W.S., 1983. Studies on southern Australian abalone (genus *Haliotis*). IV. Growth of *H. laevigata* and *H. ruber*. Marine and Freshwater Research. 34, 461–475.

Troynikov, V.S., Day, R.W., Leorke, A.M., 1998. Fisheries-estimation of seasonal growth parameters using a stochastic Gompertz model for tagging data. Journal of Shellfish Research. 17, 833–838.

Wang, Y.G., 1998a. Growth curves with explanatory variables and estimation of the effect of tagging. Australian and New Zealand Journal of Statistics. 40, 299–304.

Wang, Y.G., 1998b. An improved Fabens method for estimation of growth parameters in the von Bertalanffy model with individual asymptotes. Canadian Journal of Fisheries and Aquatic Sciences. 55, 397–400.

Wang, Y.G., 1999. Estimating equations for parameters in stochastic growth models from tag–recapture data. Biometrics. **55**, 900–903.

Wang, Y.G., Die, D., 1996. Stock-recruitment relationships of the tiger prawns (*Penaeus esculentus* and *Penaeus semisulcatus*) in the Australian northern prawn fishery. Marine and Freshwater Research. 47, 87–95.

Wang, Y.G., Thomas, M.R., Somers, I.F., 1995. A maximum likelihood approach for estimating growth from tag-recapture data. Canadian Journal of Fisheries and Aquatic Sciences. 52, 252–259.

Xiao, Y., 1994. Growth models with corrections for the retardative effects of tagging. Canadian Journal of Fisheries and Aquatic Sciences. 51, 263–267.

Table C.1: Summary of tag-recapture data for blacklip abalone. The variable L_j represents length at recapture j (mm), and T_j time at liberty between measurements L_{j-1} and L_j (years). The abbreviation SD represents standard deviation.

	Mean	SD	Range	Frequency
L_1	129	23.9	(26, 169)	1400
L_2	137	20.6	(34, 180)	1400
L_3	138	19.9	(40, 177)	383
L_4	142	17.1	(73, 165)	51
L_5	146	4.68	(139, 151)	9
L_6	140		(140, 140)	1
T_1	1.28	1.22	(0.00274, 6.99)	1400
T_2	1.91	1.41	(0.148, 7.45)	383
T_3	2.62	1.53	(0.441, 6.25)	51
T_4	3.73	1.81	(1.40, 6.31)	9
T_5	6.53		(6.53, 6.53)	1

Table C.2: Von Bertalanffy model parameter estimates from simulated data with various methods. The simulated data consist of 100 individuals run 50 times. The distributions used are $L_\infty \sim N(150, 225)$, $A \sim \Gamma(3, 5)$ ($\Gamma(\alpha, \beta)$) for MLE method and $A \sim$ log N(0.5, 0.5) for NLME, times at liberty are taken from $\Gamma(1, 1)$, and with $p = 3$ recaptures. For the ML method (MLM) 1000 integral steps are used. For the Gompertz model the parameters stay the same but $L_\infty \sim \log N(150, 225)$ and $\varepsilon \sim N(0, \sigma_\varepsilon)$ where σ_ε takes values 0.01, 0.02 and 0.04. The first and last measurements are used for the single-recapture methods.

Method	μ_∞	k	β_{IF}	σ_∞	α	β	σ_ε
True value	150	0.500		15	0.5/3	0.5/5	1
Fabens	154.0 (4.7)	0.469 (0.031)					
Improved Fabens	151.0 (5.5)	0.495 (0.038)	0.159 (0.11)				
James	152.2 (7.8)	0.486 (0.072)					
NLME	149.5 (1.5)	0.503 (0.007)		15.1 (1.14)	0.50 (0.05)	0.50 (0.03)	1.0 (0.1)
MLM	149.9 (1.4)	0.500 (0.004)		15.2 (0.80)	3.1 (0.5)	5.1 (0.8)	1.0 (0.1)
True value	150	0.500		15	0.5/3	0.5/5	2
Fabens	153.4 (4.6)	0.473 (0.029)					
Improved Fabens	150.8 (5.1)	0.497 (0.039)	0.152 (0.12)				
James	153.1 (8.7)	0.476 (0.064)					
NLME	149.8 (1.8)	0.511 (0.011)		14.8 (0.87)	0.48 (0.05)	0.49 (0.04)	2.0 (0.1)
MLM	150.2 (1.8)	0.498 (0.008)		15.0 (1.0)	3.2 (0.5)	5.2 (0.9)	2.0 (0.1)
True value	150	0.500		15	0.5/3	0.5/5	4
Fabens	152.7 (6.0)	0.479 (0.041)					
Improved Fabens	150.1 (6.4)	0.503 (0.047)	0.139 (0.09)				
James	151.7 (8.4)	0.489 (0.073)					
NLME	148.2 (2.3)	0.526 (0.024)		14.9 (1.1)	0.46 (0.06)	0.50 (0.04)	4.0 (0.2)
MLM	150.4 (2.3)	0.499 (0.015)		15.0 (1.1)	3.1 (0.5)	5.3 (0.9)	4.0 (0.2)
True value Gom	150	0.500		15	3	5	0.01/0.02/0.04
MLM Gom	150.6 (1.3)	0.500 (0.001)		15.3 (1.0)	3.2 (0.5)	5.6 (0.6)	0.01 (0.0006)
MLM Gom	150.0 (2.8)	0.500 (0.002)		14.6 (0.96)	3.2 (0.6)	5.4 (0.8)	0.02 (0.0009)
MLM Gom	149.7 (3.6)	0.501 (0.004)		14.7 (1.4)	3.1 (0.6)	5.2 (1.0)	0.04 (0.0020)

Table C.3: Model parameter estimates for blacklip abalone from the von Bertalanffy and Gompertz models. The Gompertz models used all recaptures and lengths. The von Bertalanffy models were run with animals <70 mm length at first recapture removed and all the data (ALL) for comparison with the Gompertz model. Gom=Gompertz; VB=von Bertalanffy; seas=seasonal effect; tag=tagging effect; seastag=combined seasonal and tagging effect.

Method	μ_∞	k	σ_∞	α	β	σ_ε	θ_1	θ_2	θ_3	θ_4	-loglk	AIC	$\widehat{\text{MSE}}$
MLM (VB)	151.3	0.448	11	4.0	0.69	1.8					10917.9	21847.8	**3.04**
MLM (VBseas)	149.6	0.457	10	2.5	0.39	1.8			0.136	0.352	11005.4	22026.8	4.35
MLM (VBtag)	151.0	0.549	9.7	2.5	0.49	1.8	0.341	2.87			10875.6	21767.2	4.07
MLM (VBseastag)	150.5	0.613	10	3.0	0.65	1.7	0.722	4.35	0.0643	-0.0838	10818.5	**21657.0**	3.69
MLM (VBAll)	153.2	0.376	11.8	3.5	0.55	2.1					11490.5	22993.0	4.41
MLM (VBseasAll)	153.2	0.376	11.5	3.4	0.52	2.1			0.0309	0.0513	11487.3	22990.6	4.45
MLM (VBtagAll)	151.0	0.659	10.1	2.9	0.73	1.9	0.866	3.55			11295.6	**22607.2**	4.28
MLM (VBseastagAll)	148.1	1.80	12.7	4.1	2.9	1.8	1.98	0.948	0.151	-0.164	11344.0	22708.0	**3.56**
MLM (Gom)	153.1	0.414	12.0	4.6	0.45	0.0179					-4009.6	-8007.2	6.10
MLM (Gomseas)	154.1	0.388	12.9	5.7	0.53	0.0165			0.123	0.125	-4077.7	-8139.3	5.22
MLM (Gomtag)	148.9	0.769	9.8	3.1	0.52	0.0171	0.698	1.83			-4157.3	-8298.6	5.38
MLM (Gomseastag)	151.5	0.632	10.3	4.7	0.66	0.0161	0.660	3.04	0.128	-0.0390	-4273.1	**-8526.4**	**4.43**

Table C.4: Summary of growth rates (mm/yr) for blacklip abalone. Data are separated into 4 length classes, and differing times at liberty to analyse the possible influence of tagging. The p-value represents the one-sided Welch t-test between the mean growth rates (mm/yr) of time at liberty T < 4 months and T > 1.5 years.

Previous Length (mm)	T <4 mths	T >1.5 yrs	All	P-value
25-90 mm				
Mean	6.51	22.9	13.9	0.00160
Sd	11	6.8	8.6	
Range	(-7.8, 21.3)	(10.1, 31.7)	(-22.6, 31.7)	
No.	8	8	96	
90-120 mm				
Mean	4.40	11.1	12.6	0.00030
Sd	7.6	3.7	5.8	
Range	(-5.82,19.4)	(-3.41, 19.9)	(-5.82, 36.5)	
No.	21	189	453	
120-150 mm				
Mean	1.63	4.24	3.83	0.00002
Sd	7.5	2.8	4.9	
Range	(-26.1, 36.5)	(-2.72, 13.7)	(-26.1, 36.5)	
No.	156	371	973	
150-180 mm				
Mean	-1.16	0.951	-0.282	0.00700
Sd	6.5	1.7	4.8	
Range	(-23.9,11.1)	(-2.01, 4.6)	(-23.9, 26.2)	
No.	65	50	258	

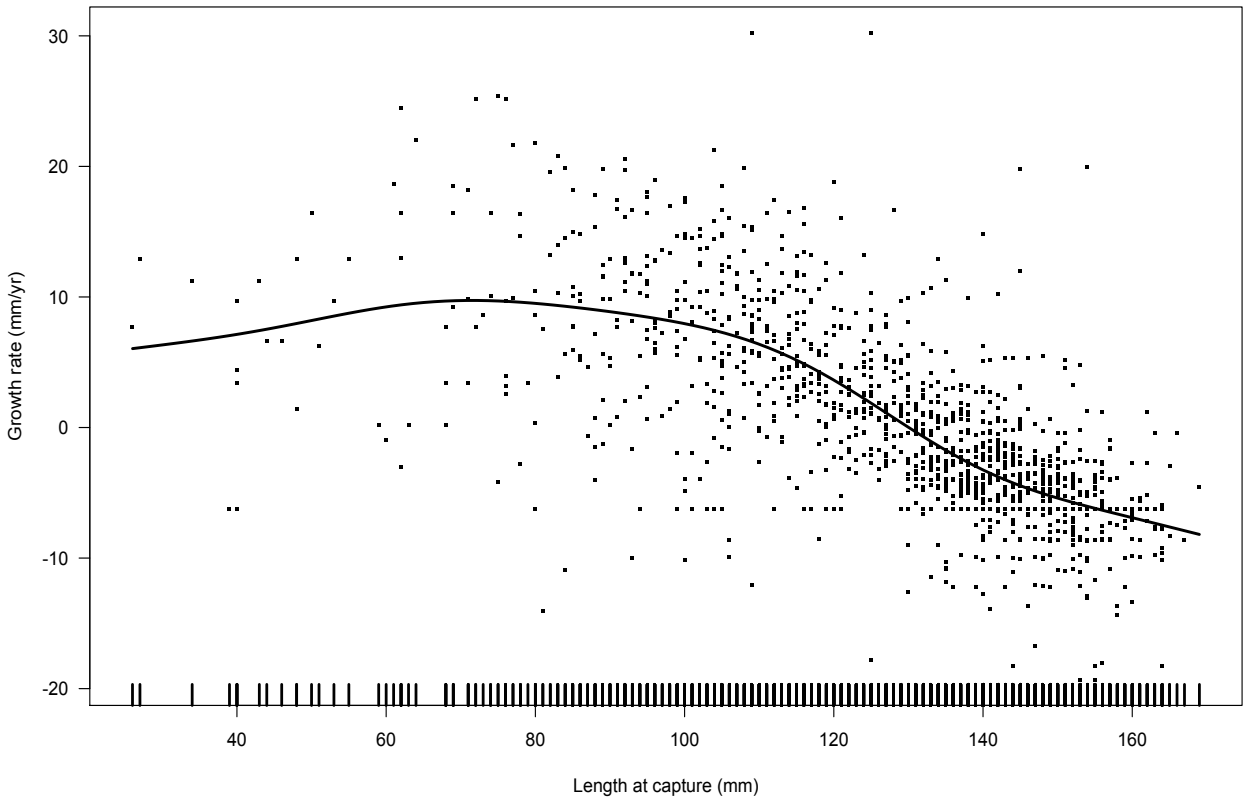

Figure C.1: Length at capture vs growth rate plot to investigate when blacklip abalone begin to follow a von Bertalanffy growth curve. The solid line is the fitted GAM model.

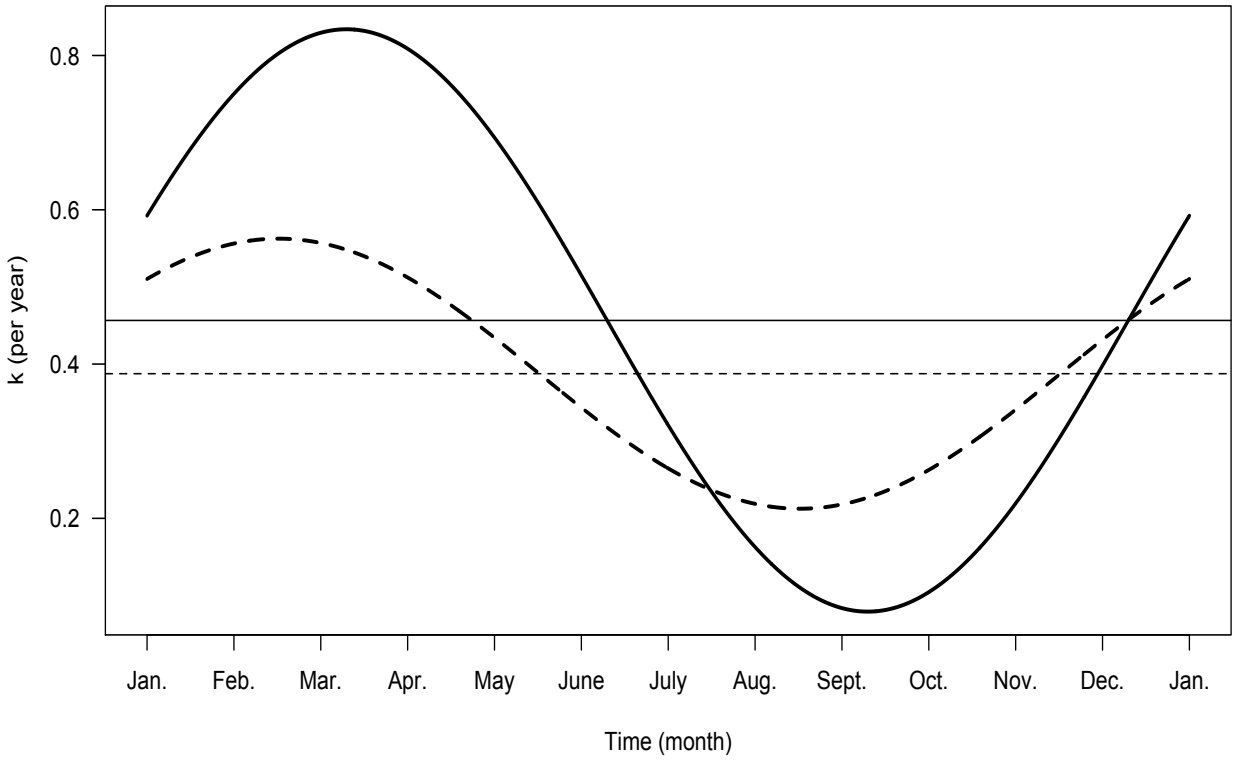

Figure C.2: Seasonal curves for blacklip abalone with canonical yearly seasonal model $k + \theta_1 \cos(2\pi t) + \theta_2 \sin(2\pi t)$ for the von Bertalanffy model (purple) and Gompertz (blue). Dotted lines indicate k for this model.

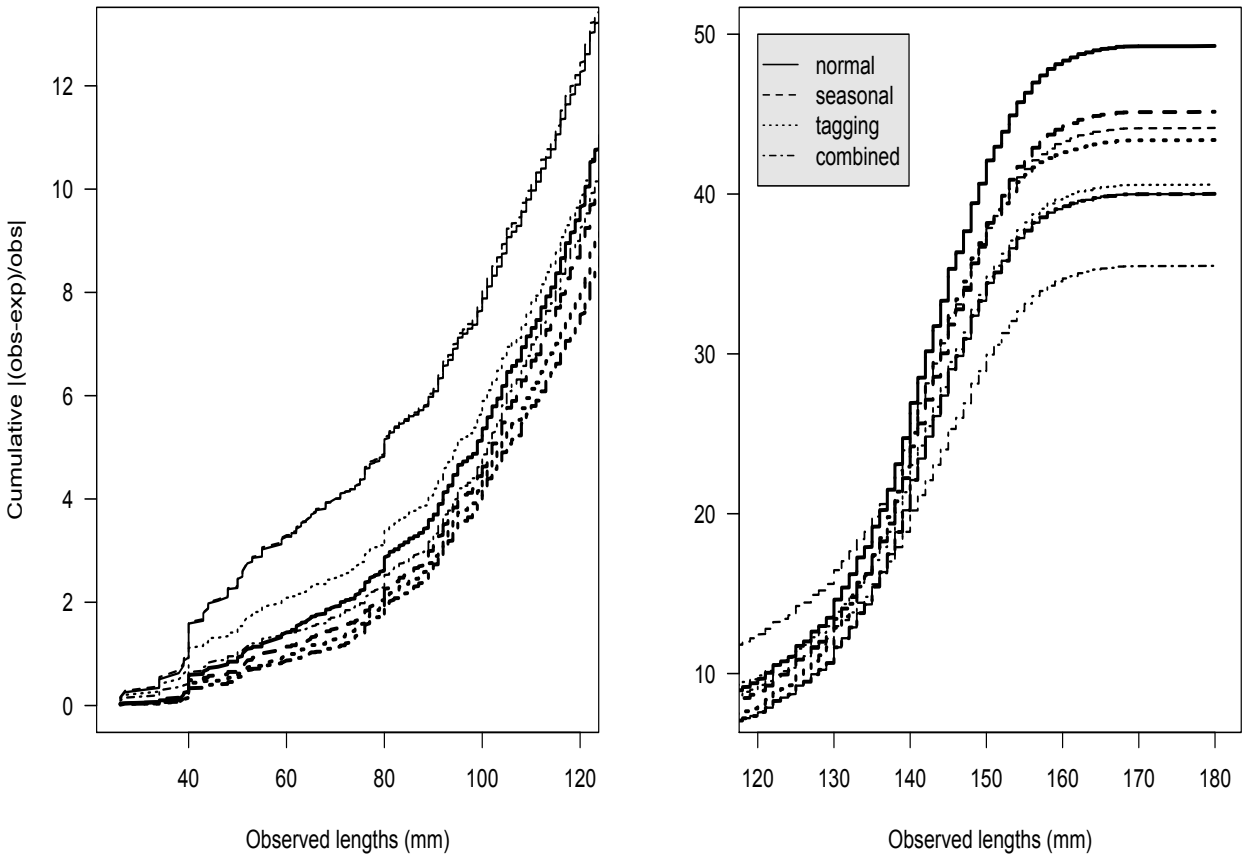

Figure C.3: Cumulative absolute percent error plots for each model. The solid lines represent the von Bertalanffy model and the dashed lines the Gompertz model. Colours indicate each model with blue the normal model, red seasonal, green tagging and purple combined.

includegraphics[width=17cm,height=12cm]growthBWt

Figure C.4: Previous length vs growth rate plot to investigate difference between growth rates for those with less than 4 months at liberty and those with greater than 1.5 years at liberty. Green line represents the spline of all points, red points and spline those with < 4 months at liberty, and blue points and spline those with > 1.5 years at liberty.

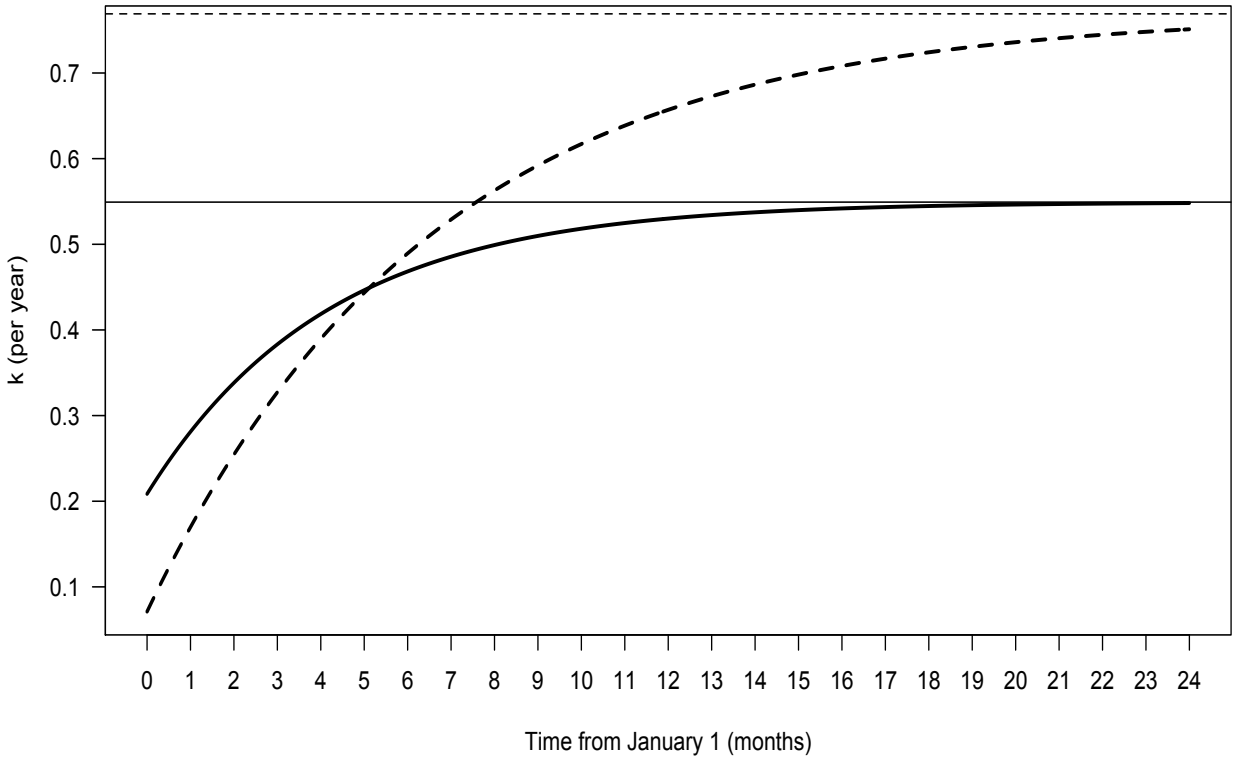

Figure C.5: Tagging effect recuperation curves for blacklip abalone for the von Bertalanffy model (purple) and Gompertz model (blue). Dotted lines indicate k for this model.

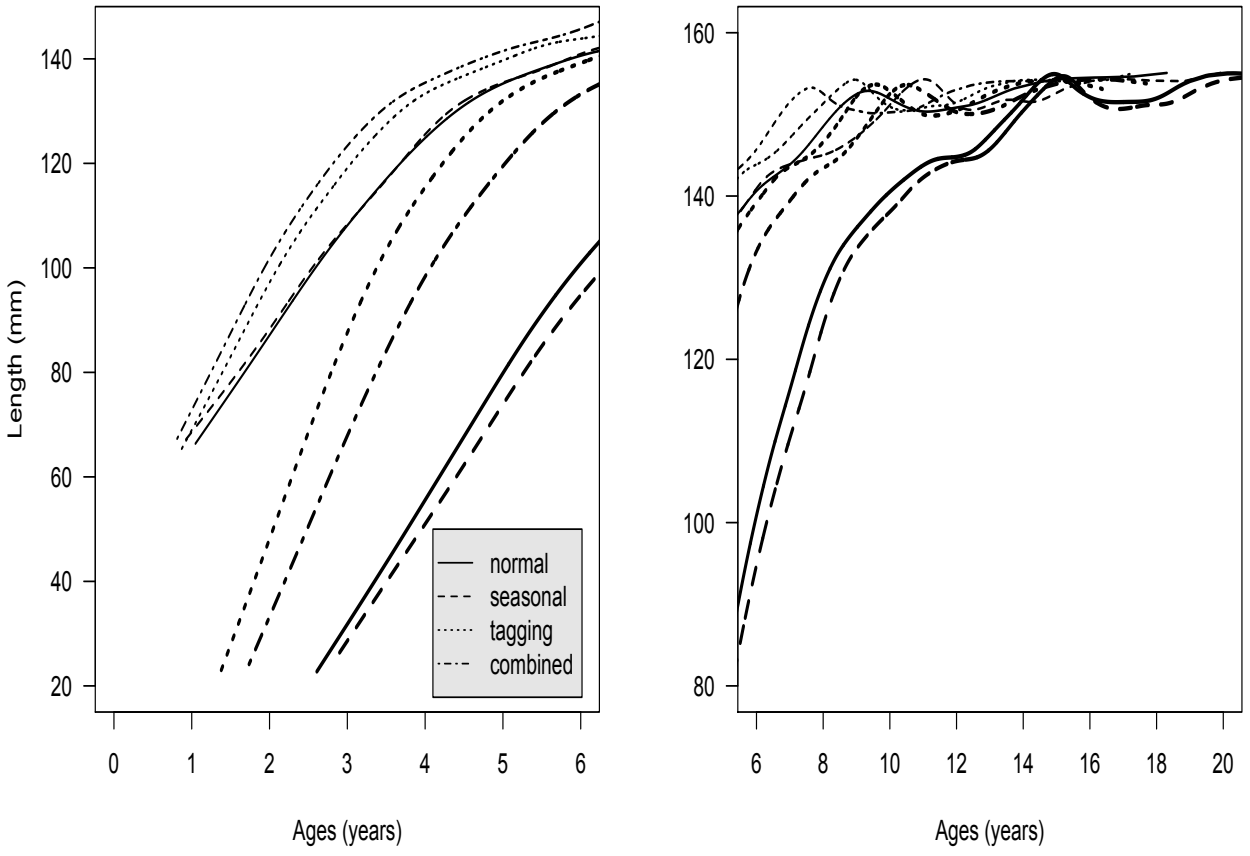

Figure C.6: Varying mean von Bertalanffy curves (solid) Gompertz curves (dashed) for the four different models: blue curve represents the no effect model, red seasonal, green tagging, and purple combined.

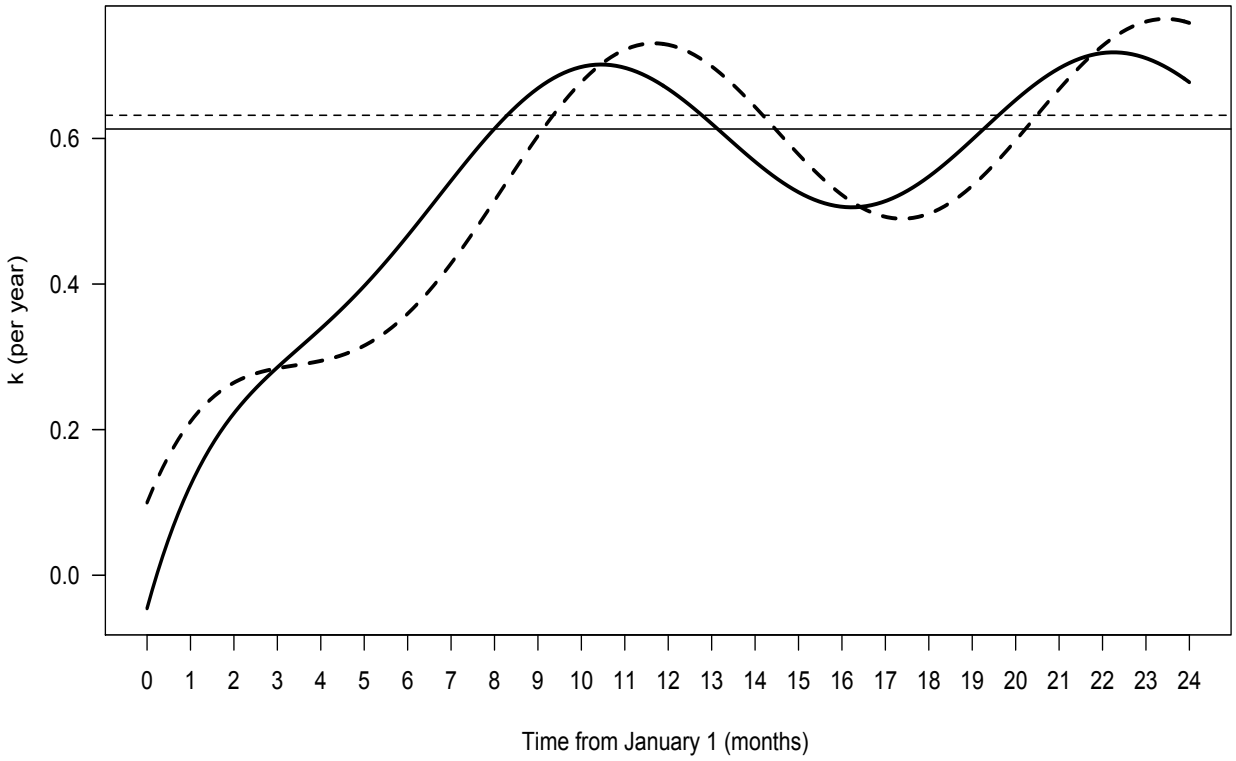

Figure C.7: Theoretical combined tagging and seasonal recuperation curve for an individual tagged on January 1 using the von Bertalanffy growth model (purple) and Gompertz model (blue). Dotted lines indicate k for this model.