Mixture of Time-Dependent Growth Models with an Application to Blue Swimmer Crab Length-Frequency Data


1Centre for Neurogenetics and Statistical Genomics, Queensland Brain Institute, University of Queensland, St Lucia, Brisbane, Queensland 4072, Australia
2School of Mathematics and Physics, University of Queensland, St Lucia, Brisbane, Queensland 4072, Australia
3Queensland Department of Agriculture and Fisheries, Ecosciences Precinct, Joe Baker St, Dutton Park, Brisbane, Queensland 4102, Australia
4School of Mathematical Sciences, Queensland University of Technology, Brisbane, Queensland 4000, Australia

*email: luke.lloydjones@uqconnect.edu.au
**email: h.nguyen7@uq.edu.au,
***email: g.mclachlan@uq.edu.au,
****email: Wayne.Sumpton@daf.qld.gov.au
*****email: yougan.wang@qut.edu.au

SUMMARY. Understanding how aquatic species grow is fundamental in fisheries because stock assessment often relies on growth dependent statistical models. Length-frequency-based methods become important when more applicable data for growth model estimation are either not available or very expensive. In this article, we develop a new framework for growth estimation from length-frequency data using a generalized von Bertalanffy growth model (VBGM) framework that allows for time-dependent covariates to be incorporated. A finite mixture of normal distributions is used to model the length-frequency cohorts of each month with the means constrained to follow a VBGM. The variances of the finite mixture components are constrained to be a function of mean length, reducing the number of parameters and allowing for an estimate of the variance at any length. To optimize the likelihood, we use a minorization–maximization (MM) algorithm with a Nelder–Mead sub-step. This work was motivated by the decline in catches of the blue swimmer crab (BSC) (Portunus armatus) off the east coast of Queensland, Australia. We test the method with a simulation study and then apply it to the BSC fishery data.

Key Words: Blue swimmer crab; Growth model estimation; Length-frequency data; Minorization–maximization algorithm; Mixture modeling.

1. Introduction

Understanding how aquatic species grow is fundamental in fisheries because stock assessment often relies on growth dependent statistical models. The statistical model used to make inference about growth depends on the data available; three common data types used for growth model estimation are (i) individual length and time-at-liberty data from tag-recapture experiments; (ii) length-frequency data gathered from fishers and/or research surveys; and (iii) length data that are directly aged via otolith readings, scales, and or calcified structures such as spines. It is common to use tag-recapture data to estimate the parameters of a potential growth model. However, tagging studies can be costly and ineffective for some species, especially crustaceans (McPherson, 2002; González-Vicente et al., 2012). Length-frequency data are recordings of individual lengths at a time point and can be interpreted as a mixture distribution; this interpretation allows for the identification of separate modes, which may be attributed to different age cohorts. Following the modes of these cohorts through time provides a natural way to model growth of an aquatic species (Pauly and Morgan, 1987; Montgomery et al., 2011).

The von Bertalanffy growth model (VBGM) is commonly used in fisheries research. The VBGM has a biological basis, with its derivation stemming from a differential equation that models growth rate as the difference between the rate of anabolism versus the rate of catabolism. Since its inception (Von Bertalanffy, 1938), the VBGM has remained one of the models of choice for representing growth in aquatic species largely due to its biological interpretability and success in outperforming competing multi-parameter functions (Pauly, 1979; Chen et al., 1992; Essington et al., 2001). In its length form, the solution to von Bertalanffy’s differential equation with relevant assumptions is

\[ L(t) = L_\infty \left[ 1 - \exp\left( -K(t - T_0) \right) \right], \]

where \( L_\infty \) represents the asymptotic length, \( K \) controls the curvature, and \( T_0 \) is defined to be the time when an individual would have had length 0 if its post-juvenile growth