Received: 6 October, 2011
IOTC-2011-WPTT13-30Rev_1

PRELIMINARY APPLICATION OF THE BROWNIE-PETERSEN METHOD TO SKIPJACK TAG-RECAPTURE DATA

J. Paige Eveson

CSIRO Marine & Atmospheric Research
Hobart, Tasmania, Australia
paige.eveson@csiro.au

Prepared for the 13th Session of the IOTC Working Party on Tropical Tuna, 16-23 October 2011, Maldives

Table of Contents

Abstract 1
Introduction 1
Methods 1
Population dynamics model 1
Estimation Model 4
Data and assumptions 5
Results 9
Discussion 12
References 13
Appendix: Updated growth estimates 15

Abstract

Results from applying the Brownie-Petersen method for estimating mortality rates and abundance to skipjack tag-recapture data and catch data from years 2005 to 2009 (using data corresponding to cohorts 2001 to 2005) are presented. The analysis used an annual time-step and a single fishery (i.e., tag returns and catches were aggregated within each year across fisheries). Several alternative scenarios were considered; however, overall, the results suggest natural mortality rate at ages 1 to 4+ is U-shaped (highest at ages 1 and 4+ and lower at ages 2 and 3). Fishing mortality rates vary significantly between years and ages, but were higher in 2006 and 2007 than in 2008 and 2009. When interpreting the results, it is important to note that a large number of uncertainties exist in the data and the model assumptions. The results presented can only be considered preliminary until some of these issues are resolved and/or further sensitivity runs are conducted.

Introduction

The Brownie-Petersen method, presented in Polacheck et al. (2006), is a method for estimating natural mortality rates, fishing mortality rates and abundance from multi-year tagging data integrated with catch data. The inclusion of catch data not only improves estimation of mortality rates (especially fishing mortality) but also allows for direct estimation of cohort size at the time of tagging. This method provides a potentially powerful alternative to CPUE and fishery-independent surveys for augmenting traditional stock assessments.

In this paper, we apply the Brownie-Petersen method to Indian Ocean skipjack tuna (*Katsuwona pelamis*) tag-recapture and catch data. Stock assessment of skipjack is difficult due to a number of reasons including: they have rapid (but hard to quantify) growth, relatively short life spans, spawn almost continuously, and have questionable relative abundance indices due to the nature of the fishery (mostly purse seine and pole and line). A large-scale conventional tagging program, referred to as the Regional Tuna Tagging Project - Indian Ocean (RTTP-IO), was started in 2005 and since then a large amount of tag release and recapture data has accumulated for skipjack (as well as yellowfin and bigeye). These data, either stand-alone or integrated with other data sources, have the potential to provide valuable information for assessing the stock. At the 2010 IOTC Scientific Committee (SC) meeting, the SC marked conducting an assessment of skipjack as a priority, and they recommended the estimation of natural mortality rates, as well as abundance and exploitation rates, using the Brownie-Petersen method applied to latest RTTP-IO tag-recapture data.

Methods

The Brownie-Petersen method is presented in detail in Polacheck et al. (2006), but the relevant information is reproduced here for convenient reference. Modifications that were used in the application to skipjack data are also described.

Population dynamics model

The basic model underlying the analyses of the multi-year tagging experiments used here is the general population dynamic equations commonly used in fisheries. These equations involve exponential and competing natural and fishing mortality rates. Thus for a cohort of animals of a given age, the number that survive one time step is

$$P_{i,\,t+1} = P_{i,\,t}\,\exp\{-F_{i,t} - M_{i,t}\} \tag{1}$$

$$C_{i,t} = \frac{F_{i,t}}{F_{i,t} + M_{i,t}} P_{i,t}\left(1 - \exp\{-F_{i,t} - M_{i,t}\}\right) \tag{2}$$

where:

$P_{i,t}$ = the number of individuals of age *i* at time *t*
$C_{i,t}$ = the catch of individuals of age *i* at time *t*
$F_{i,t}$ = the instantaneous fishing mortality rate for individuals of age *i* at time *t*
$M_{i,t}$ = the instantaneous natural mortality rate for individuals of age *i* at time *t*.

In most fisheries contexts, $M_{i,t}$ will be assumed to be constant with time, although multi-year and multi-cohort tagging programs can provide year and age specific natural mortality rates. Here, we focus on a multi-year tagging experiment involving a single cohort. As such, we will drop the *t* subscript and express everything in terms of age.

Note that the model and equations are presented in terms of a single cohort as this is the minimum required by the model and makes the notation simpler. In practice, it is likely that several cohorts (age-classes) would be tagged in each time period of tagging. To include multiple cohorts in the model, one simply needs to develop the likelihood for each cohort as described in the next section, and then multiply them together to form a joint likelihood. Note that if all parameters being estimated vary with both year and age, then maximizing the likelihood for each cohort separately is equivalent to maximizing the joint likelihood (i.e., will yield the same parameter estimates). More likely, however, some parameters will be shared; for example, if natural mortality varies with age but not with year, then all fish recaptured at a given age will have a common M parameter regardless of the year.

In the context of a tagging experiment, the above equations provide the basis for predicting the expected number of returns assuming that the tagged fish constitute a representative sample of the population. Following Brownie et al. (1985), the expected number of tags recaptured and returned from a particular cohort at age *i* from releases at age *a* ($R_{a,\,i}$) are given by the expressions in Table 1.

Table 1. Expressions for the expected number of tag returns by age corresponding to releases at a particular age, for a tagging experiment in which a cohort of fish is tagged at ages 1 to 3 and recaptured at ages 1 to 5.

Release	#	Expected # returns from age class *i*				
Age	Releases	1	2	3	4	5
1	N_1	$\lambda_1 N_1 f_1$	$\lambda_2 N_1 S_1 f_2$	$\lambda_3 N_1 S_1 S_2 f_3$	$\lambda_4 N_1 S_1 S_2 S_3 f_4$	$\lambda_5 N_1 S_1 S_2 S_3 S_4 f_5$
2	N_2		$\lambda_2 N_2 f_2$	$\lambda_3 N_2 S_2 f_3$	$\lambda_4 N_2 S_2 S_3 f_4$	$\lambda_5 N_2 S_2 S_3 S_4 f_5$
3	N_3			$\lambda_3 N_3 f_3$	$\lambda_4 N_3 S_3 f_4$	$\lambda_5 N_3 S_3 S_4 f_5$

where:

N_a= the number of tag releases of age *a* fish from a specific cohort

$f_i = F_i/(M_i+F_i)*[1 - \exp\{-(M_i+F_i)\}]$

$S_i = \exp\{-(M_i+F_i)\}$

λ_i= tag reporting rate for fish captured at age *i*.

The above expressions for the expected number of returns assume complete and instantaneous mixing of tagged fish and no tagging mortality or loss. In our application to skipjack data, we modify the equations to incorporate tag shedding as follows:

$$f_i = \alpha F_i/(M_i+F_i+\Omega)*[1 - \exp\{-(M_i+F_i+\Omega)\}]$$
$$S_i = \exp\{-(M_i+F_i+\Omega)\}$$

where α is the instantaneous retention rate (i.e., the proportion of tags that are not shed immediately after tagging) and Ω is the continuous shedding rate (i.e., the rate at which tags shed over time).

Note that these modified equations pertain to single-tagged fish, but for simplicity we assume here that they hold for double-tagged fish as well. In actuality, the probability of a fish retaining (at least) one tag will be greater for a double-tagged fish than a single-tagged fish; however, for skipjack, the shedding rate estimates are low enough (see ‘Data and assumptions’ section below) that we assume the difference can be ignored.

We also want to account for the fact that newly tagged fish will not be fully mixed with the untagged population immediately after tagging, and for the fact that tagging generally occurs during the fishing season so tagged fish are only vulnerable for part of the season. To do so, we allow the F parameters to differ between tagged fish in the time period they were tagged and untagged fish in that same time period (see application to southern bluefin tuna in Polacheck et al. 2006). We assume that tagged and untagged fish are fully mixed by the time period following release, where a time period is one year in applications using an annual time-step, and one quarter in applications using a quarterly time step.

Equations (1) and (2) can also be used to provide analogous expressions for the expected catches of age *i* fish from a particular cohort, conditional on the size of the cohort at the age of first tagging, assumed here to be age 1 and denoted by P_1 (Table 2).

Table 2. Expressions for the expected number of fish caught at ages 1 to 5 from a cohort which had an age 1 abundance of P_1.

Size of	Expected catch from age class *i*				
cohort	1	2	3	4	5
P_1	P_1f_1	$P_1S_1f_2$	$P_1S_1S_2f_3$	$P_1S_1S_2S_3f_4$	$P_1S_1S_2S_3S_4f_5$

Essentially, the catch data can be viewed as a tagging experiment in which the number of releases (P_1) is unknown and is a parameter to be estimated. However, unlike a tagging experiment where there is little uncertainty in the numbers of tags returned[1], the numbers of fish caught at each age will be estimated quantities. These quantities are usually derived from a multi-stage sampling of catches for length combined with age-length keys derived from otoliths, or obtained via cohort slicing. Because P_1 is unknown, it is not possible to derive estimates of mortality rates from the catch at age data alone[2]. However, combining the catch at age data with the multi-year tagging data allows P_1 to be estimated and additional information on F and M contained in the catch data to be extracted.

Estimation Model

As developed in Brownie et al. (1985), if each tag recapture is assumed to be independent, then the numbers of returns at age corresponding to a given release event are expected to be multinomial. The likelihood function for the observed numbers of returns from all release events is the product of multinomials:

$$L_R = \prod_a \left(\frac{N_a\,!}{\prod_{i \ge a} R_{a,i}\,!\,(N_a - R_{a,\bullet})!} \prod_{i \ge a} p_{a,i}^{R_{a,i}} (1 - p_{a,\bullet})^{N_a - R_{a,\bullet}} \right) \qquad (3)$$

where a indexes release age, i indexes recapture age, and $p_{a,i}$ is the probability of a tag being returned from an age i fish released at age a. An expression for $p_{a,i}$ can be obtained from the expected number of returns in Table 1 by dividing by N_i. Explicitly,

$$p_{a,i} = \begin{cases} \lambda_i f_i & i = a \\ \lambda_i S_a \cdots S_{i-1} f_i & i > a \end{cases}$$

Note that in equation (3) and in subsequent equations, a dot in the subscript denotes summation over the index it replaces.

Variance in the tag return numbers may be greater than a multinomial distribution predicts (due to factors such as tagged fish remaining in schools). Overdispersion (i.e., extra variability) can be accounted for by using a Dirichlet-multinomial distribution, but this requires specifying the level of overdispersion since it cannot be reliably estimated within the model. Assuming a multinomial distribution should not bias the parameter estimates; however, it means that their estimated standard errors will be too small if that return data are in fact overdispersed (Polacheck et al. 2006).

---

[1] The numbers of tags *recaptured* can have high uncertainty due to uncertain reporting rates, but the numbers of tags actually returned (i.e., the data that enters the model) are usually known accurately.
[2] Even if M is assumed known as in many stock assessments, there are still too many parameters and this is the reason that catch at age stock assessment models require additional sources of data.

Similar to the tag-return data, if we assume that all fish in a cohort are independent, then the catch at age data can be modelled as random multinomial, where each fish has a probability of being captured at age *i*. Expressions for the catch probabilities can be obtained by dividing the expected catches in Table 2 by the initial cohort size (P_1).

The age distribution of the catch is usually determined by taking a sample of the catch, estimating the ages of fish in the sample (either from lengths or from direct aging of hard parts), and then scaling up the estimated age frequencies of the sample by the ratio of the catch size to the sample size. We have chosen to represent the error in the catch at age data that results from this estimation procedure as Gaussian with a common coefficient of variation (CV), υ, across all age classes. To fit a model with both multinomial "process" error and Gaussian sampling/measurement error would require a relatively sophisticated approach, such as a Kalman filter. However, in most fisheries the number of fish in the cohort from which catches are being taken will be very large such that the multinomial error will be negligible compared to the Gaussian sampling error (see Polacheck et al. 2006). In such cases, only the latter needs to be considered. Thus, the likelihood for the catch at age data can be expressed as

$$L_C = \prod_i \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{1}{2}\left(\frac{C_i - E(C_i)}{\sigma_i} \right)^2 \right) \tag{4}$$

where the expected catch at age *i*, $E(C_i)$, is given in Table 2 and $\sigma_i = \upsilon E(C_i)$.

The overall likelihood for the combined recapture and catch data can be obtained by multiplying likelihoods (3) and (4) together:

$$L = L_R \times L_C \tag{5}$$

Estimates of the *F*, *M* and *P* parameters can be obtained by maximizing the likelihood in (5) (or, equivalently, by minimizing the negative log of this likelihood). The parameter υ cannot be estimated from the data when a separate *F* is estimated for each year of recapture, thus we assume that it is known.

The information for estimating M_i comes from the differential between the expected returns at age $i+1$ of fish released at age *i* and those released at age $i+1$. Thus, in an experiment with *n* release events, estimates can only be obtained for M_i to M_{n-1} because subsequent *M*'s are not separable from the corresponding *F* parameters. In an experiment with three release events, as illustrated in Table 1, only M_1 and M_2 are estimable. Therefore we assume that $M_i = M_{n-1}$ for $i \geq n$.

Data and assumptions

The data used in the analysis presented here are the RTTP-IO skipjack (SKJ) tag-recapture data, and the SKJ catch data used in the stock assessment being conducted by D. Kolody (for which the data are compiled by year, quarter and fishery, where 4 fisheries have been defined). In each year-quarter-fishery combination that had length information, the sample length-frequency data was scaled up to the total catch in that year, quarter and fishery. For a year-quarter-fishery combination that did not have any length information, the length-frequency data for that fishery from adjacent years (i.e., years y-1, y and y+1) were combined to calculate a length-frequency distribution, which was then scaled up to the total catch for that year, quarter and fishery.

The release data and the catch data were aged based on length and an assumed growth curve, using a simple ‘cohort slicing’ method (i.e., fish with lengths between $L(a_1)$ and $L(a_2)$ are considered to be age a_1, where $L(a)$ is the expected length at age a calculated from the growth curve). The growth curve used was a von Bertalanffy (VB) curve estimated from applying the method of Laslett et al. (2002) to the SKJ tag-recapture data. This method was applied to YFT, BET and SKJ tag-recapture data in 2008 (Eveson and Million 2008), and was updated using the most recent SKJ recapture data for the current analysis (see Appendix). Even the updated tag-recapture data does not have much information on large fish (over 65cm), so when no constraints are put on the parameters, the asymptotic length parameter (Linf) is estimated to be lower than seems believable (considering the maximum sizes seen in the catches). As such, the model was fit fixing Linf at 75cm (based on the largest fish in the catch data used in this analysis) and 83cm (based on length data from longline fisheries; pers. comm. D. Kolody). These two growth curves are referred to as VB75 and VB83 respectively (Figure 1). VB75 was used as the default in this analysis.

For the analyses presented here, a single fishery was assumed and an annual time-step was used; thus, releases, recaptures and catches were aggregated by year over all fisheries, and integer ages were calculated.

Cohorts 2001 to 2005, release ages 1 to 5 and recapture ages 1 to 6, were included in the analysis, since these had sufficient data. Note, however, that each cohort was not tagged at all ages (see Table 3). The Brownie-Petersen method requires a cohort to be tagged in multiple consecutive time periods in order to separate M from F; thus, even though a large number of fish from cohort 2006 were tagged at age 1, they are not informative with regard to M. Table 3 gives a summary of the data used in the base analysis, in the format required for the Brownie-Petersen model. In particular, the release and recapture data are broken down by cohort and age of release, and age at recapture. The catch data are broken down by cohort and age. Release age and age of the catch were estimated using the VB75 growth curve (note that recapture age was calculated from the estimated age at release and the time at liberty). Cohort is calculated as year (of tagging or catch) minus estimated age.

Table 3. Data used in base analysis. (a) Number of tag releases by (estimated) cohort and release age, and corresponding number of tag returns by age. (b) Catch numbers (in millions) by cohort and age. Ages were estimated using VB75 growth curve.

(a) Release-recapture data

	Release	Number	Number recaptures by age					
Cohort	age	releases	1	2	3	4	5	6
2001	1	0	0	0	0	0	0	0
	2	0	0	0	0	0	0	0
	3	0	0	0	0	0	0	0
	4	678	0	0	0	15	18	2
	5	386	0	0	0	0	18	4
2002	1	0	0	0	0	0	0	0
	2	0	0	0	0	0	0	0
	3	3918	0	0	51	297	26	0
	4	1278	0	0	0	101	15	0
	5	538	0	0	0	0	17	2
2003	1	0	0	0	0	0	0	0
	2	9027	0	75	1219	217	12	0
	3	12409	0	0	1001	808	31	0
	4	1056	0	0	0	42	5	1
	5	0	0	0	0	0	0	0
2004	1	247	0	22	10	1	0	0
	2	25011	0	1567	3031	175	12	2
	3	5431	0	0	562	101	5	0
	4	0	0	0	0	0	0	0
	5	0	0	0	0	0	0	0
2005	1	2125	22	351	21	1	0	0
	2	11857	0	1032	534	69	2	0
	3	0	0	0	0	0	0	0
	4	0	0	0	0	0	0	0
	5	0	0	0	0	0	0	0

(b) Catch data (in millions)

	Age					
Cohort	1	2	3	4	5	6
2001				18.6	21.7	10.9
2002			47.2	21.4	11.4	8.0
2003		46.3	51.8	9.2	7.8	2.4
2004	9.5	43.2	29.4	5.6	3.0	
2005	11.2	46.0	18.2	4.1		

In fitting the Brownie-Petersen model, we allowed M to vary with age but not across years, and F to vary with both age and year (with no assumptions about selectivity patterns). Because we are assuming a single fishery, this means we estimate a total F for each year and age (i.e., we do not estimate fishery-specific Fs).

Reporting rates are required for each year and age of tag returns being included in the model. Although reporting of a tag is not expected to depend on the age of a fish, it is still necessary to estimate age-specific reporting rates in situations where there are multiple fisheries with different selectivities, implying different age-structures in the catches (Hearn et al. 1999). An average reporting rate across all fisheries is calculated for year *y* and age *a* by taking a weighted average of the fishery-specific reporting rates, where the weights are the proportion of the catches in year *y* belonging to age class *a* in each fishery. Reporting rates have been estimated from tag seeding data for the Seychelles purse seine catches (pertaining to both log set (LS) and free school (FS) catches). Two sets of estimates exist: those used in the assessment and those by Hillary et al. (2008). The assessment estimates were used here because they exist for most of the years required and therefore do not require as much extrapolation. Reporting rates for the "at sea" purse seine catches (LS and FS) are assumed to be 100%; they are assumed to be 0% for all other fisheries, and any tag returns from these fisheries are excluded from analysis[3]. The reporting rates used in the analysis are given in Table 4. Note that the small values for some years and ages are due to the fact that the fisheries assigned a 0% reporting rate catch a large portion of the total catch for that year and age.

Table 4. Reporting rate estimates (averaged over all fisheries) by cohort and age of recapture. See text for details.

	Age					
Cohort	1	2	3	4	5	6
2001				0.22	0.24	0.06
2002			0.34	0.38	0.16	0.07
2003		0.36	0.42	0.40	0.14	0.28
2004	0.08	0.54	0.46	0.30	0.53	
2005	0.34	0.55	0.35	0.71		

Tag shedding estimates were assumed to remain the same over time, and were taken from Table 3 of Gaertner and Hallier (2009). In particular, the values used were α=0.987 and Ω=0.015 (per year).

Recall that the CV of the catch data needs to be specified. A value of 0.3 was used here; although this value was chosen rather arbitrarily, previous investigations have shown the results to be fairly insensitive to the value used (Eveson et al. 2007).

In addition to fitting the model according to the above specifications, a few alternatives were also considered.

[3] The Maldives pole and line fishery does return some tags; however, in the absence of any information on reporting rates, it is simplest to omit these returns from the analysis and assume a 0% reporting rate for this fishery.

Alternative 1: Catch data omitted.

Alternative 2: VB83 growth curve used to age data.

Alternative 3: VB83 growth curve used to age data, and catch data omitted.

Results

The results obtained from the base model run and alternatives 1 to 3 are presented in Table 5. For the base run, M at age is estimated to be U-shaped, with relatively large values at ages 1 and 4+ and smaller values at ages 2 and 3; however, note that the CVs are quite large at ages 3 and 4. The F estimates for a given age can differ a lot between cohorts (i.e., years); for example, the age 4 F estimate was ~0.45 for cohorts 2002 and 2003 (years 2006 and 2007), but was only 0.15 and 0.04 for cohorts 2004 and 2005 (years 2008 and 2009) respectively. When considering the F estimates, it is important to keep the CVs in mind since they can be very large for years and ages where the numbers of tag returns were low, such as age 6. The cohort size estimates must also be interpreted carefully because they correspond to the size of the cohort at the age when it was first tagged. Thus, the estimate of 0.56 billion for cohort 2001 is an estimate of the number of age 4 fish in year 2005; whereas the estimate of 4.2 billion for cohort 2004 is an estimate of the number of age 1 fish in 2005.

Comparing the base run results with those from the three alternative runs, we can make the following general statements. When catch data are omitted from the analysis, the F estimates all increase and some of them quite substantially. Correspondingly, M estimates also increase (since the F and M estimates are positively correlated). Population size estimates cannot be obtained without catch data, nor can meaningful estimates of F in the year of tagging (since F in the year of tagging is allowed to differ from the general population to allow for a period of non-mixing). Using the VB83 growth curve as opposed to the VB75 growth curve results in M estimates that are again U-shaped over ages 1 to 4+, but with the age 1 estimate being larger than the age 4 estimate, and the age 3 estimate significantly smaller than the age 2 estimate. The F estimates are quite similar, especially when CVs are taken into account. The population size estimates are on the same order, but are quite a bit larger for cohort 2003, age 2 (year 2005) and cohort 2005, age 1 (year 2006) for the VB83 results than the VB75 results.

Table 5. Natural mortality rate (M) estimates by age, fishing mortality rate (F) estimates by cohort and age, and abundance at age of first release (P) by cohort (in billions). Results are given for the default model run and the 3 alternatives specified in the methods. Note that in the alternatives that do not include catch data, F in the year of release is not estimated (because it is assumed to differ from the general population to allow for a period non-mixing). When catch data are included, F in the year of release can be estimated from the catch data. The numbers in parenthesis below the point estimates give the CVs of the estimates.

Base run (VB75):

M		Age 1	2	3	4+		
		0.585	0.376	0.210	0.707		
		(0.12)	(0.10)	(0.74)	(0.44)		
F	Cohort	Age 1	2	3	4	5	6
	2001				0.148	0.466	0.882
					(0.35)	(0.54)	(1.30)
	2002			0.132	0.448	0.278	0.263
				(0.41)	(0.22)	(0.56)	(1.01)
	2003		0.022	0.791	0.444	0.131	0.028
			(0.52)	(0.08)	(0.23)	(0.55)	(0.96)
	2004	0.003	0.412	0.707	0.155	0.016	
		(0.43)	(0.26)	(0.08)	(0.22)	(0.52)	
	2005	0.010	1.091	0.309	0.039		
		(0.39)	(0.13)	(0.08)	(0.21)		
P	Cohort:	2001	2002	2003	2004	2005	
	Age:	4	3	2	1	1	
		0.56	1.02	4.28	4.20	1.38	
		(0.29)	(0.33)	(0.45)	(0.35)	(0.30)	

Alternative 1 (VB75 and omit catch data):

M		Age 1	2	3	4+		
		0.618	0.367	0.290	1.094		
		(0.11)	(0.10)	(0.45)	(0.17)		
F	Cohort	Age 1	2	3	4	5	6
	2001					0.902	5.000*
					NA	(0.41)	(NA)
	2002				0.609	0.765	0.206
				NA	(0.19)	(0.51)	(0.99)
	2003			0.820	0.601	0.329	0.013
			NA	(0.07)	(0.20	(0.45)	(1.21)
	2004		0.478	0.732	0.205	0.025	
		NA	(0.25)	(0.07)	(0.19)	(0.41)	
	2005		1.159	0.323	0.047		
		NA	(0.12)	(0.07)	(0.19)		

* upper bound

Alternative 2 (VB83):

M		Age 1	2	3	4+		
		0.634	0.498	0.115	0.492		
		(0.16)	(0.07)	(0.87)	(0.51)		
F	Cohort	Age 1	2	3	4	5	6
					0.314	0.430	0.465
	2001				(0.32)	(0.46)	(0.89)
				0.140	0.464	0.385	0.169
	2002			(0.39)	(0.16)	(0.48)	(0.85)
			0.009	0.780	0.376	0.106	0.010
	2003		(0.57)	(0.06)	(0.16)	(0.43)	(0.81)
		0.003	0.404	0.781	0.208	0.020	
	2004	(0.47)	(0.56)	(0.06)	(0.16)	(0.42)	
		0.003	1.303	0.286	0.062		
	2005	(0.44)	(0.18)	(0.07)	(0.17)		
P	Cohort:	2001	2002	2003	2004	2005	
	Age:	4	3	2	1	1	
		0.61	1.07	6.98	2.97	2.66	
		(0.25)	(0.31)	(0.51)	(0.40)	(0.38)	

Alternative 3 (VB83 and omit catch data):

M		Age 1	2	3	4+		
		0.677	0.501	0.131	0.816		
		(0.15)	(0.07)	(0.71)	(0.27)		
F	Cohort	Age 1	2	3	4	5	6
	2001					0.831	1.094
					NA	(0.47)	(0.97)
	2002				0.563	0.817	0.228
				NA	(0.14)	(0.48)	(0.90)
	2003			0.792	0.453	0.199	0.006
			NA	(0.06)	(0.14)	(0.38)	(0.94)
	2004		0.728	0.794	0.251	0.027	
		NA	(0.55)	(0.06)	(0.14)	(0.37)	
	2005		1.438	0.293	0.070		
		NA	(0.18)	(0.07)	(0.16)		

Discussion

The results presented here provide a useful first step in estimating mortality rates and abundance from the growing amount of skipjack tag-recapture data. The estimates of natural mortality may prove particularly useful given the lack of alternative methods for estimating this parameter. Nevertheless, a large number of uncertainties exist in the data inputs and assumptions of the model, and the results must be considered preliminary. For example, a reliable growth curve for skipjack has yet to be established so the age estimates are highly uncertain. Also, an annual time-step is arguably too coarse given the rapid population dynamics of skipjack, and the fact that different components of the fishery operate at different times and can have highly variable exploitation rates by quarter. As such, a number of sensitivity analyses should be conducted before drawing conclusions from the Brownie-Petersen results. For example, further model runs that would be informative include:

- use quarterly time step
- check sensitivity to which cohorts and release/recapture ages are included in the analysis
- estimate fishery-specific Fs
- use the alternative reporting rates by Hillary et al. (2008) once they have been updated to include more years
- test sensitivity to larger range of growth models
- include overdispersion in the tag return data

Lack of time has prevented these runs from being conducted for inclusion in this report. If time permits, it is hoped that some of these runs will be run prior to the meeting so that the results can be presented and discussed at that time.

References

Brownie, C., D. R. Anderson, K. P. Burnham, and D. S. Robson, 1985. Statistical inference from band recovery data: a handbook. U.S. Fish and Wildlife Resource Publication 156.

Eveson, J.P., Polacheck, T., and Laslett, G.M. 2007. Incorporating fishery observer data into an integrated catch-at-age and multiyear tagging model for estimating mortality rates and abundance. Fish. Bull. 105: 493-508.

Eveson, P. and Million, J. 2008. Growth of tropical tunas using the Laslett, Polacheck and Eveson (LEP) method. IOTC-2008-WPTT-09.

Gaertner, D. and J.P. Hallier. 2009. An updated analysis of tag-shedding by tropical tunas in the Indian Ocean. IOTC-2009-WPTT-34.

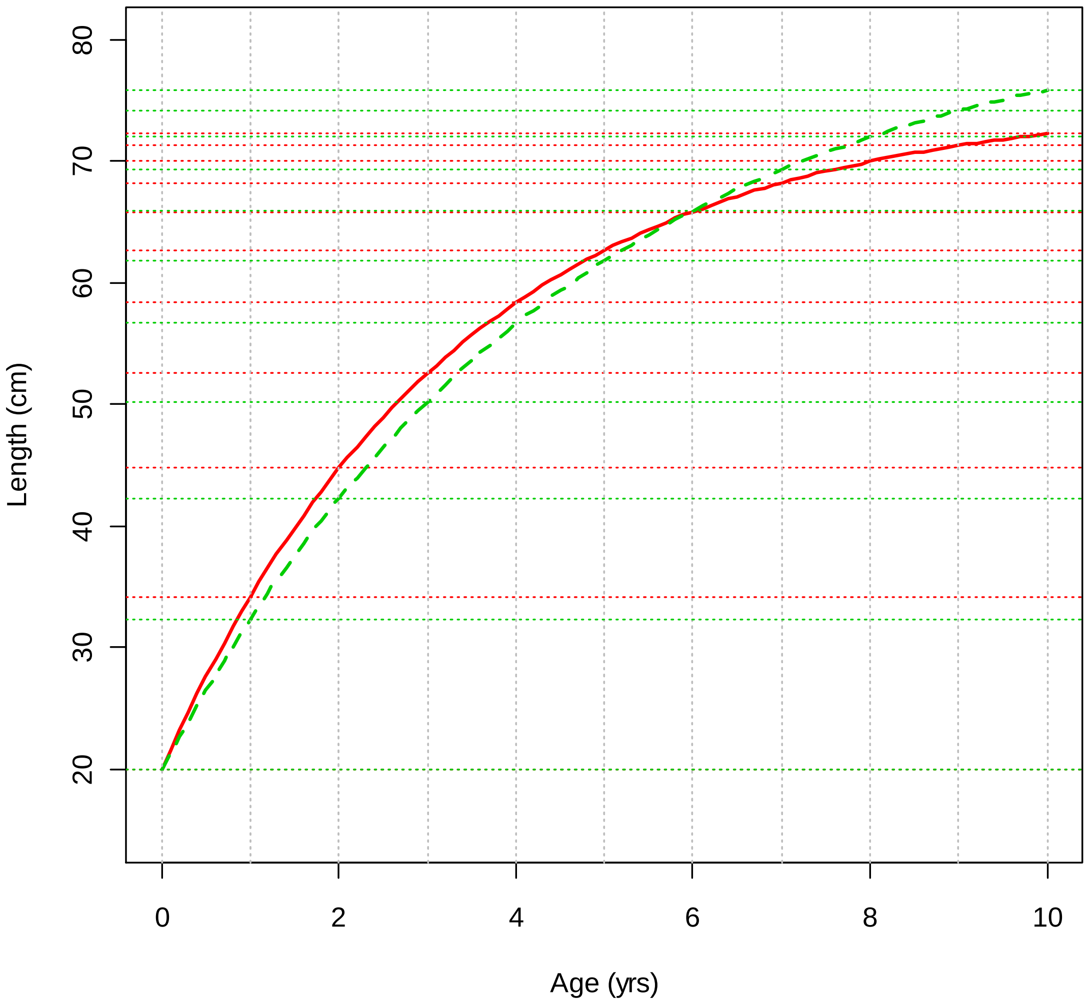
Hearn, W. S., T. Polacheck, K. H. Pollock, and W. Whitelaw. 1999. Estimation of tag reporting rates in age-structured multicomponent fisheries where one component has observers. Canadian Journal of Fisheries and Aquatic Sciences 56: 1255-1265.

Hillary, R.M., IOTC Secretariat, Areso, J.J. 2008. Reporting rate analyses for recaptures from Seychelles port for yellowfin, bigeye and skipjack tuna. IOTC-2008-WPTT-18.

Laslett G.M., Eveson, J.P., and Polacheck, T. 2002. A flexible maximum likelihood approach for fitting growth curves to tag-recapture data. Can. J. Fish. Aquat. Sci. 59: 976-986.

Polacheck, T., Eveson, J.P., Laslett, G.M., Pollock, K.H., and Hearn, W.S. 2006. Integrating catch-at-age and multiyear tagging data: a combined Brownie and Petersen estimation approach in a fishery context. Can. J. Fish. Aquat. Sci. 63: 534-548.

Figure 1. Growth curves used in estimating age from length. Solid red line = VB curve with Linf fixed at 75cm (VB75); dashed green line = VB curve with Linf fixed at 83cm (VB83). Dotted red/green horizontal lines indicate upper and lower length cut-offs for each age class corresponding to the VB curve of the same colour (e.g., using the red VB curve, fish between 20 and 34cm are estimated to be age 0 and fish between 34 and 45cm are estimated to be age 1; using the green VB curve, fish between 20 and 32cm are estimated to be age 0 and fish between 32 and 42cm are estimated to be age 1).



<u>Appendix:</u> Updated growth estimates for skipjack obtained from the LEP method applied to the most recent RTTP-IO tag-recapture data

J. Paige Eveson
CSIRO Marine and Atmospheric Research

Introduction

Growth information for skipjack tuna is limited. As part of the Regional Tuna Tagging Project - Indian Ocean (RTTP-IO), large numbers of yellowfin, bigeye and skipjack tuna were tagged in 2005 to 2007, and have subsequently been recaptured in fisheries. Length measurements of all fish were taken upon release. When a tagged fish is recaptured, the fisherman is asked to record the tag number, species name, date and location of catch, and fork length of the fish (i.e., length from the nose to the middle of the tail), and to return the tag along with this information to the RTTP-IO headquarters in the Seychelles.

The change in length of a tagged fish between the time it was released and the time it was recaptured provides useful information for modelling growth. Because the age of a fish at release is unknown, the traditional approach has been to model the incremental change in length of the fish over the time it was at liberty (Fabens 1965; Francis 1988; James 1991). Such methods can lead to biased parameter estimates when individual variability in the growth parameters exists (Sainsbury 1980; Maller and deBoer 1988; Eveson et al. 2007). To overcome this problem, maximum likelihood approaches have been developed that model the joint density of the release and recapture lengths as opposed to modelling the length increment (Palmer et al. 1991; Wang et al. 1995; Laslett et al. 2002). In these cases, the age at release is modelled as a random variable.

In IOTC-2008-WPTDA-07, the method of Laslett et al. (2002), referred to as the LEP method, was applied to the tag-recapture data for yellowfin, bigeye and skipjack. Since that time, a large number of additional recaptures have occurred. Here, the analysis for skipjack is updated using the most recent data.

Data

The skipjack tag-recapture dataset used in this report was a “cleaned” dataset, obtained using the criteria outlined in Appendix V of the 2008 Report of the First Session of the IOTC Working Party on Tagging Data Analysis. The cleaned data set contains 4345 recaptures.

Histograms of the release lengths, recaptures lengths, and times at liberty for all fish in the cleaned dataset are shown in Figure A1. The median release length is 46 cm (interquartile range 46-52 cm), the median recapture length is 53.6 cm (IQ range 50.7-57.1 cm), and the median time at liberty is 256 days (IQ range 129-345 days). A number of fish had recapture lengths less than their release lengths, implying negative growth (Figures A2 and A3); this suggests that measurement error may be a significant component of the large observed variability in growth.

Figure A2 shows an obvious decline in the empirical growth rate (change in length divided by time at liberty) with release length up to ~40-45 cm, after which it appears to be roughly constant at 0.02 cm/day. If growth occurs according to a VB model, then the growth rate should decline linearly with release length. This suggests that perhaps a two-stage growth model may be appropriate for skipjack; however, the degree of variability in the data is high and the amount of data for fish under 40 cm is low, so it is difficult to be conclusive.

Methods

Given our initial explorations of the data, two growth functions were considered for skipjack: (1) von Bertalanffy (VB) and (2) VB with a logistic growth rate parameter (VB log k) (see Laslett et al. 2002). Both functions can be expressed as

$$l(a) = L_\infty f(a - a_0; \theta)$$

where L_∞ is asymptotic length and *f* is a monotone increasing function with parameter set $\{a_0, \theta\}$ that equals 0 when $a = a_0$. The parameter a_0 can be thought of as the theoretical age at which a fish would have had length 0 if we were to project its growth curve backwards. This parameter cannot be estimated from tag-recapture data alone, so in order to fully define the growth curve, it must be determined from other sources.

For the VB curve, $\theta = \{k\}$ and $f(a - a_0; \theta) = 1 - \exp(-k(a - a_0))$.

For the VB log k curve, $\theta = \{k_1, k_2, \alpha, \beta\}$ and

$$f(a - a_0; \theta) = 1 - \exp(-k_2(a - a_0)) \left\{ \frac{1 + \exp(-\beta(a - a_0 - \alpha))}{1 + \exp(\alpha\beta)} \right\}^{-(k_2 - k_1)/\beta}$$

The VB log k function allows for a change in growth from a VB curve with growth rate parameter k_1 to a VB curve with growth rate parameter k_2, with a smooth transition between the two stages that occurs according to a logistic function. The parameter α governs the age at which the midpoint of the transition occurs, and β governs the rate of the transition (being sharper for larger values).

These growth curves were fitted to the data using the LEP method, details of which can be found in Laslett et al. (2002). The key feature of this method is that it models the release and recapture lengths as functions of age by treating age at tagging, *A*, as a random variable[4]. *A* is assumed to follow a specified distribution, and the parameters of this distribution are estimated within the model. In applying the LEP method to the three tuna data sets, a lognormal distribution (with mean and standard deviation on the

[4] Note that *A* actually represents the age at tagging, a_1, relative to a_0 (i.e., $A = a_1 - a_0$), because it is not possible to estimate a_0 directly from the tagging data.

log scale of μ_A and σ_A) was chosen for *A*; Laslett et al. (2002) showed that the results were fairly robust to the distribution used for *A* so long as it provided a reasonable approximation. Another feature of the LEP method is that it allows for individual variability in growth by modelling the asymptotic length parameter as a random effect. For all species, L_∞ was assumed to follow a normal distribution with mean $\mu_{L\infty}$ and standard deviation $\sigma_{L\infty}$. Additional random normal variability in length due to measurement error was also included with mean 0 and standard deviation σ_ε.

The LEP method is based on maximum likelihood, so the likelihood values could be used to compare model fits. Residual plots were also used to evaluate the fits. Note that to calculate the fitted recapture values (and thus the residuals) for the LEP method requires a realized value of *A* and L_∞ for each fish. These were estimated using the procedures described in Laslett et al. (2002). Briefly, for each fish, the mean of the posterior distribution for *A* and for L_∞ was calculated given the fish's release length and recapture length. Also, in order to plot the growth curve in terms of age instead of 'age relative to a_0', we assume that a fish has length 20 cm at age 0 (i.e., $l(0) = 20$) and calculate a_0 such that it meets this criteria. This value seems reasonable given the smallest sizes in the catches, but cannot be verified without direct ageing data.

Results and discussion

The results from fitting a VB model to the data without any parameter constraints are given in Table A1 and Figure A4; similarly the results for the VB log k model are given in Table A1 and Figure A5.

Both models, the VB model in particular, suggest a lower asymptotic length than seems probable for skipjack considering the maximum sizes seen in the catches. This is due, at least in part, to the fact that the tag-recapture data has very little information on large fish (over 65cm). As such, we refit the VB model fixing the mean asymptotic length parameter ($\mu_{L\infty}$) at 75 cm (based on the largest fish in the catch data used in this analysis) and 83 cm (based on length data from longline fisheries; pers. comm. D. Kolody). We will refer to these two growth models as VB75 and VB83 respectively. The results for these two models are given in Table A1 and Figures A6 and A7.

We also tried fitting the VB log k model with the mean asymptotic length parameter fixed at 75 cm, but encountered convergence problems and the model appeared to want to converge to a simple VB model (with the change in growth occurring around age 5 and the growth rate prior to that age being the same as for the VB75 model).

The mean growth curves estimated from the VB, VB75, VB83 and VB log k growth models are compared in Figure 8. There is a marked difference between them, with the VB and VB log k curves suggesting much more rapid initial growth compared to the VB75 and VB83 curves. Rapid initial growth is consistent with the empirical growth rate estimates (Figure A3).

Table A1. Parameter estimates for the four growth models fit to the skipjack tag-recapture data. The a_0 parameter was calculated post hoc such that the expected length of a fish at age 0 is 20 cm. (NLL = negative log likelihood.)

Model	$\mu_{L\infty}$	$\sigma_{L\infty}$	k_1	k_2	α	β	μ_A	σ_A	σ_ε	a_0	NLL
VB	61.6	4.27	0.76	–	–	–	0.75	0.16	1.68	-0.52	23279.7
VB75	75	4.57	0.30	–	–	–	1.27	0.12	1.94	-1.04	23367.9
VB83	83	4.93	0.23	–	–	–	1.42	0.10	1.99	-1.27	23405.1
VB log k	67.6	4.69	1.06	0.43	1.12	28.3	0.35	0.23	1.62	-0.33	23100.8

The negative log likelihood value is much smaller for the VB log k model than any of the VB models (Table A1), and suggests it provides a significantly better fit even when the extra number of parameters is taken into account. Also, the residuals look better for the VB log k model (Figure 5) compared to the VB models (Figures A4, A6 and A7). For instance, all of the VB models show a lack of fit to the smallest lengths, and the unconstrained VB model also underestimates the largest fish. Note that the range of the y-axis differs between the residual plots. Thus, the pattern of the residuals for the VB75 and VB83 models is better than for the unconstrained VB model, but the absolute values are larger (hence the significantly larger negative log likelihoods).

The estimated release age distribution for the VB log k model (Figure A5) has a fairly narrow mode at age 1, which is perhaps too young to be realistic. The VB, VB75 and VB 83 models estimate the mode of the release ages to be around 2, 2.5 and 3 years respectively (Figures A4, A6 and A7 respectively).

When the data have a high degree of variability, as for skipjack, the LEP method has a lot of flexibility as to how it makes the data best fit the assumed growth model – it can assume that individuals have a high degree of variability in their asymptotic lengths, it can attribute the variability to measurement error, or it can “adjust” the estimates of release ages to provide the best fit. Without information to ground truth at least one of these components (e.g., more information on older fish, estimates of measurement error, otolith data to verify realistic release ages), it makes determining the most appropriate growth curve difficult. The apparently better fit achieved by the VB log k model compared to the VB model may be due to the model being too flexible in the face of uncertain data, as opposed to it being more biologically realistic.

References

Eveson, J.P., Polacheck, T., and Laslett, G.M. 2007. Consequences of assuming an incorrect error structure in von Bertalanffy growth models: a simulation study. Can. J. Fish. Aquat. Sci. 64: 602-617.

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

Francis, R.I.C.C. 1988. Maximum likelihood estimation of growth and growth variability from tagging data. N.Z. J. of Mar. Freshwater Res. 22: 42-51.

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

Laslett, G.M., Eveson, J.P., and Polacheck, T. 2002. A flexible maximum likelihood approach for fitting growth curves to tag-recapture data. Can. J. Fish. Aquat. Sci. 59: 976-986.

Laslett G.M., Eveson, J.P., and Polacheck, T. 2004. Estimating the age at capture in capture-recapture studies of fish growth. Aust. N.Z. J. Stat. 46: 59-66.

Maller, R.A., and deBoer, E.S. 1988. An analysis of two methods of fitting the von Bertalanffy curve to capture-recapture data. Aust. J. Marine Freshwater Res. 39: 459-466.

Palmer, M.J., Phillips, B.F., and Smith, G.T. 1991. Application of nonlinear models with random coefficients to growth data. Biometrics 47: 623-635.

Sainsbury, K.J. 1980. Effect of individual variability on the von Bertalanffy growth equation. Can. J. Fish. Aquat. Sci. 37: 241-247.

Wang, Y.-G., Thomas, M.R., and Somers, I.F. 1995. A maximum likelihood approach for estimating growth from tag-recapture data. Can. J. Fish. Aquat. Sci. 52: 252-259.

Figure A1. Distribution of (a) release lengths; (b) recapture lengths; (c) times at liberty (in days) for recaptured skipjack.

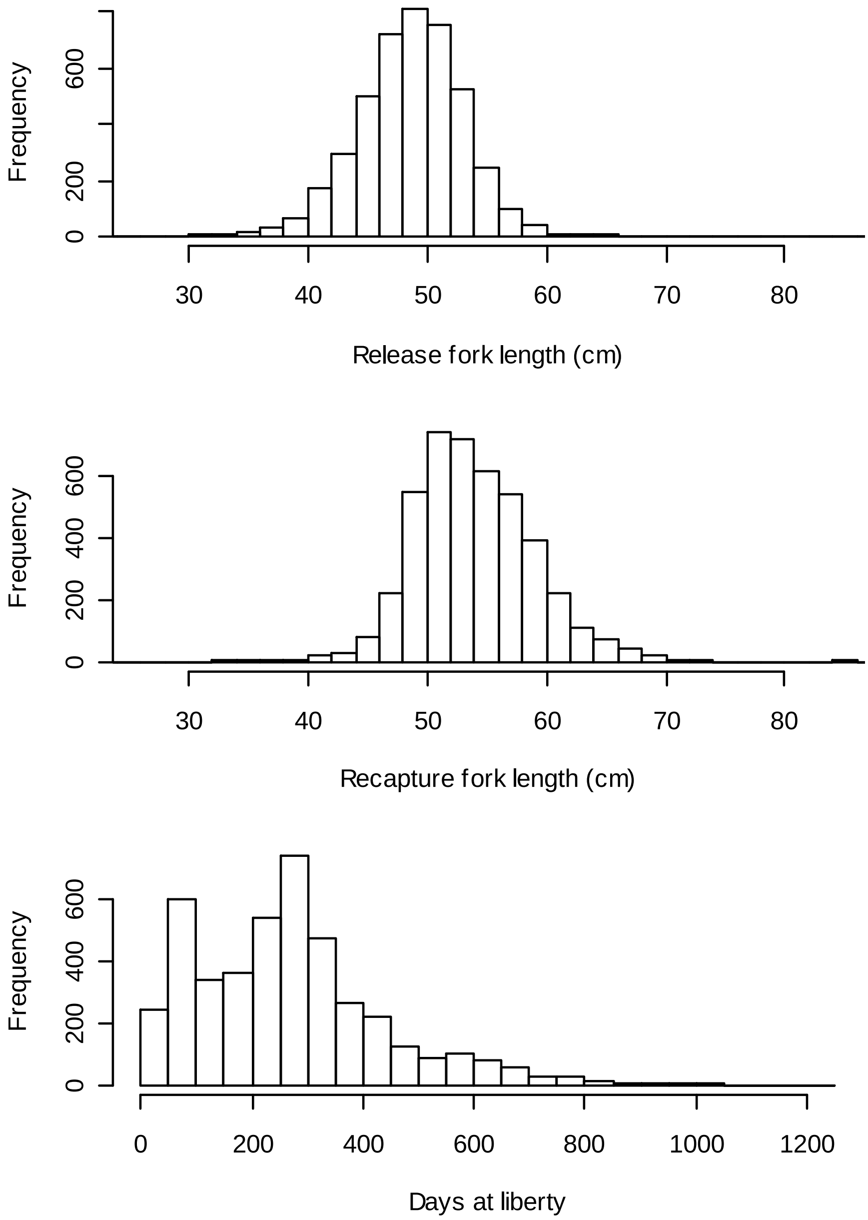


Figure A2. Time at liberty (days) versus length increment (cm) between release and recapture. The negative growth increments are presumably due to measurement error.

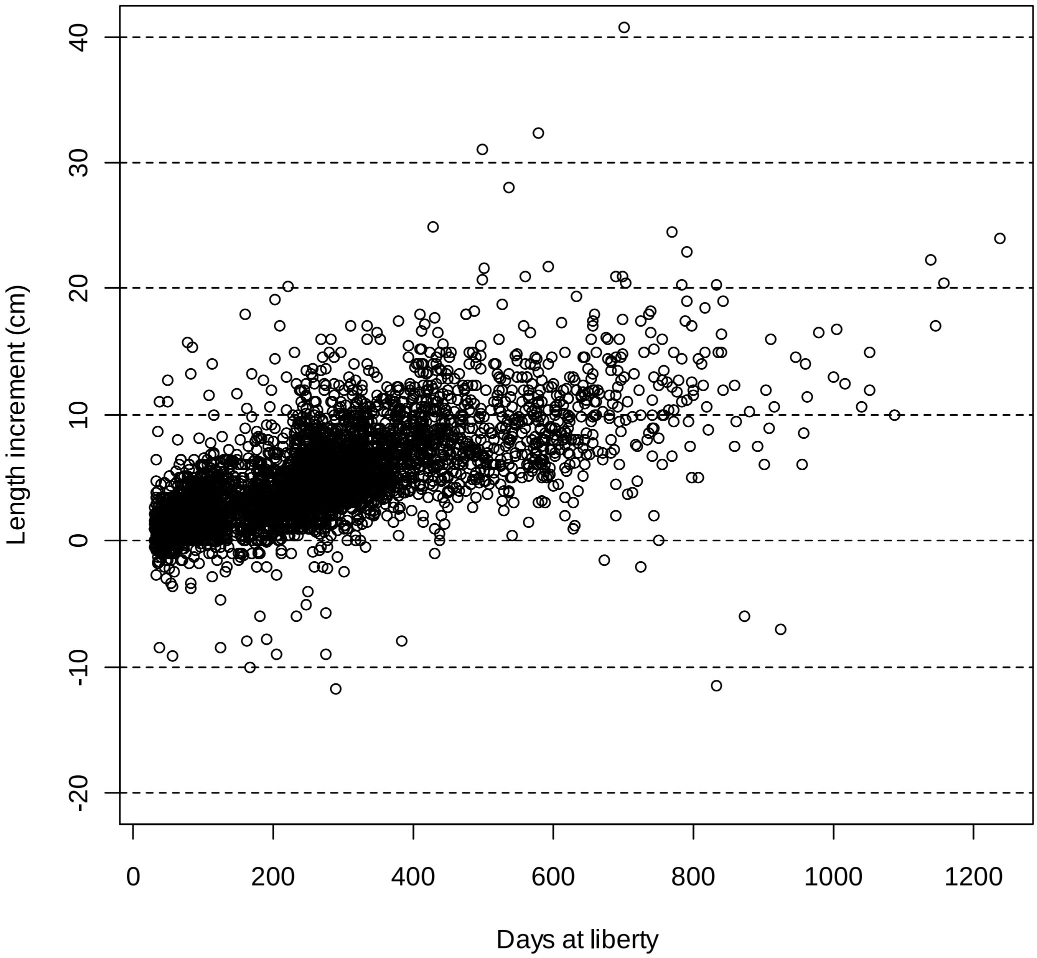


Figure A3. Growth rate (cm/day) versus release length (cm). The solid pink line is a smooth of the data, and the dashed blue line indicates a growth rate of 0.02 cm/day.

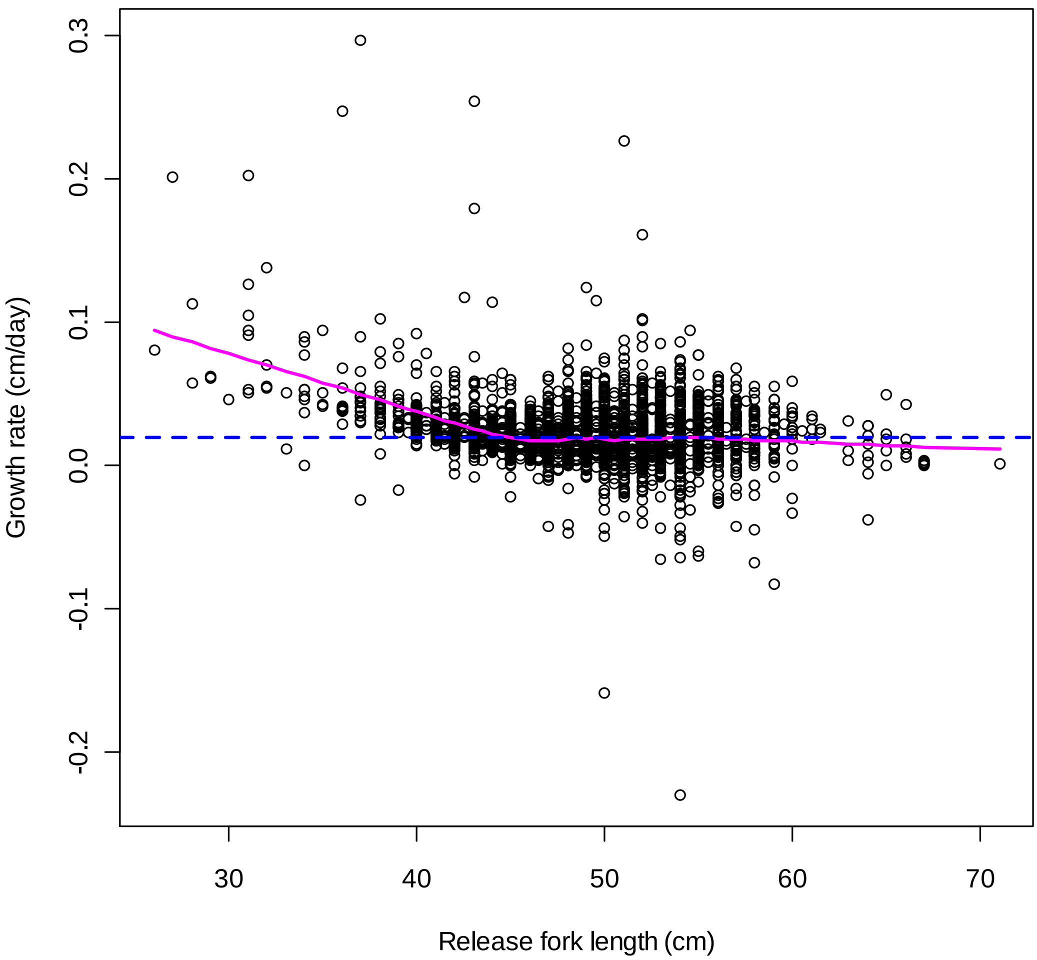
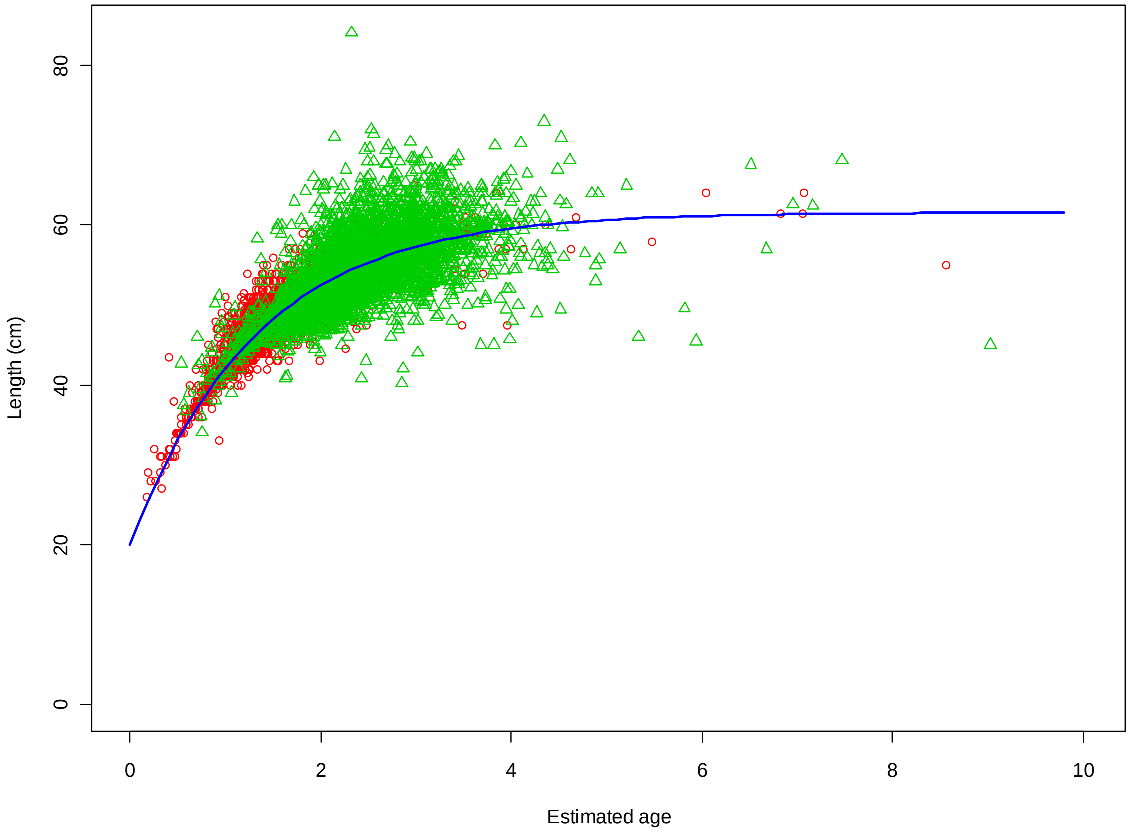
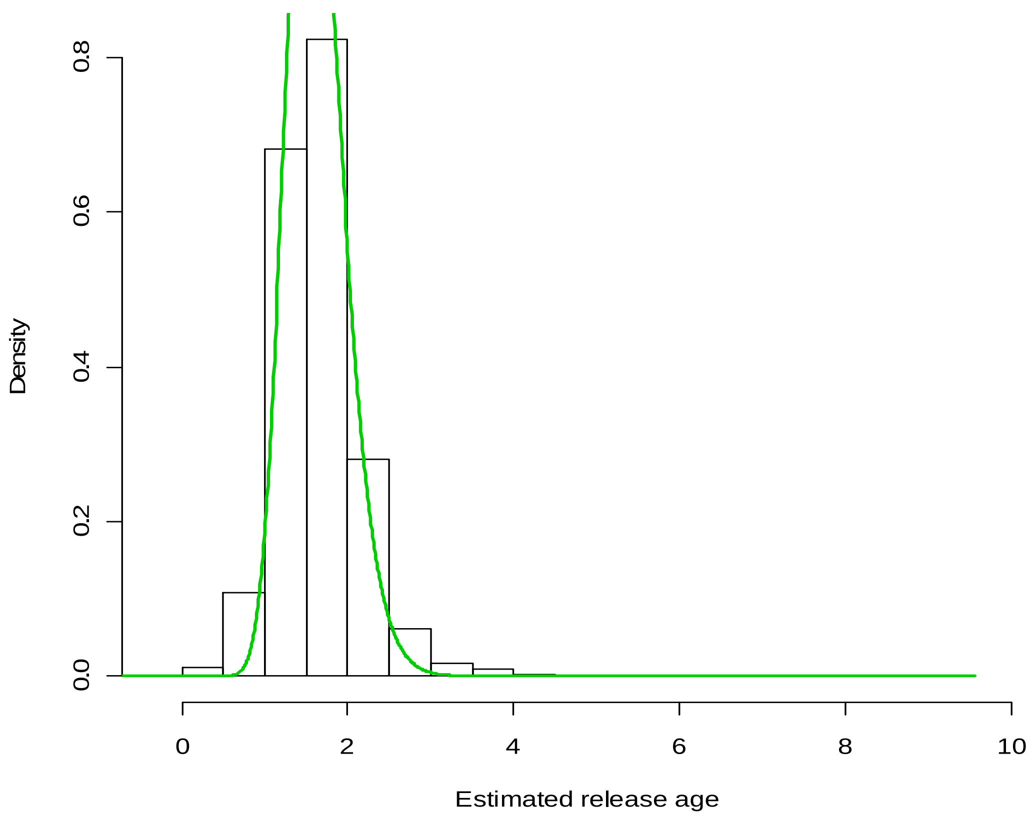
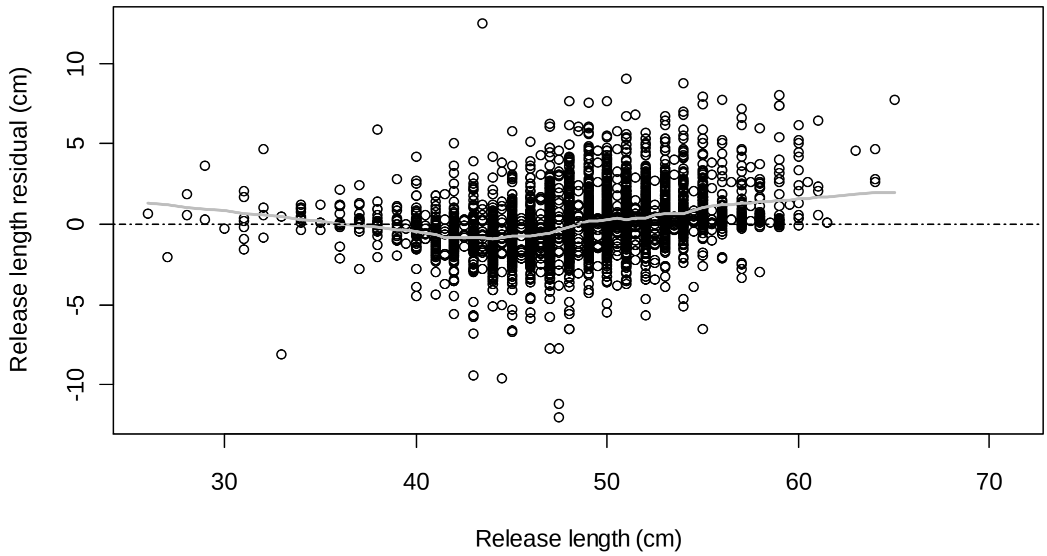


Figure A4. Fitted curve and diagnostic plots for VB model fitted to all skipjack tag-recapture data using the LEP method with no parameter constraints. The age axis was set assuming L(0) = 20cm.






Release length residual (cm)
-10
-5
0
5
10
30
40
50
60
70
Release length (cm)

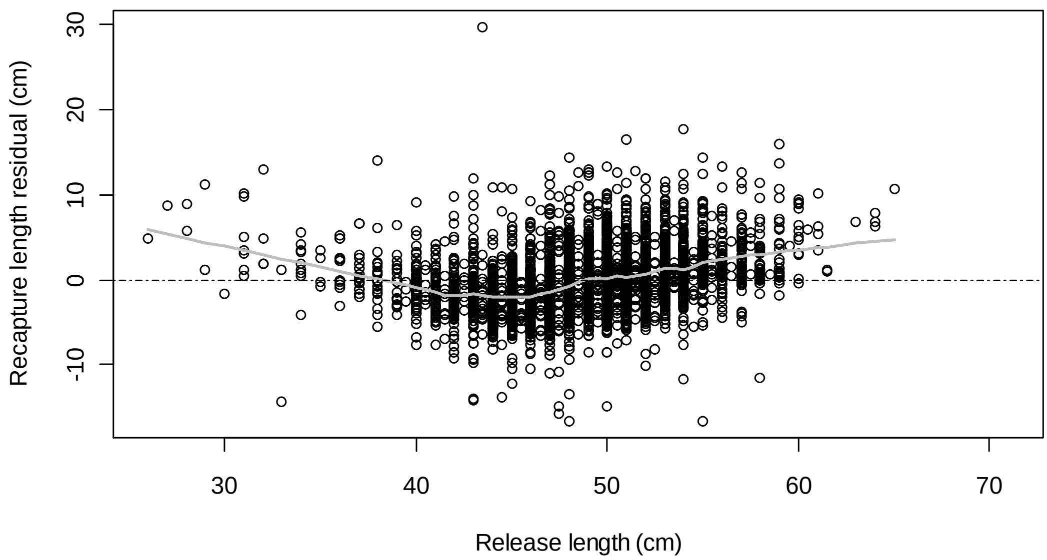
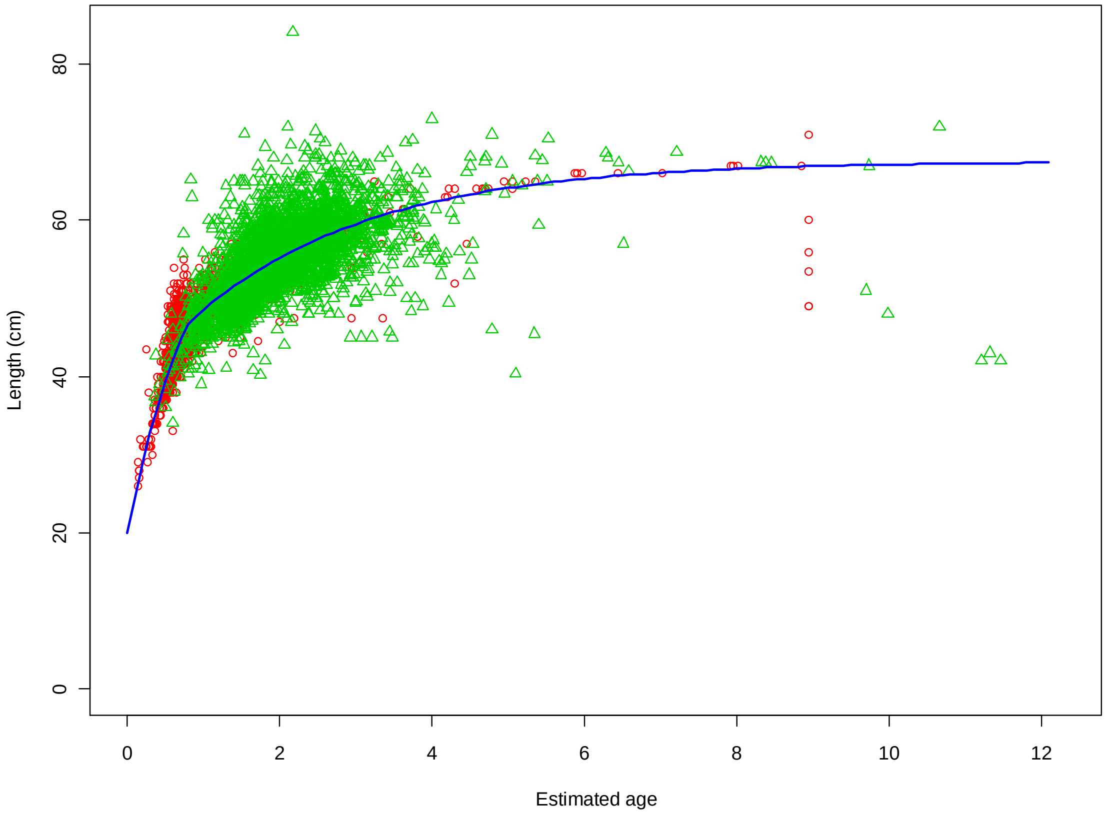
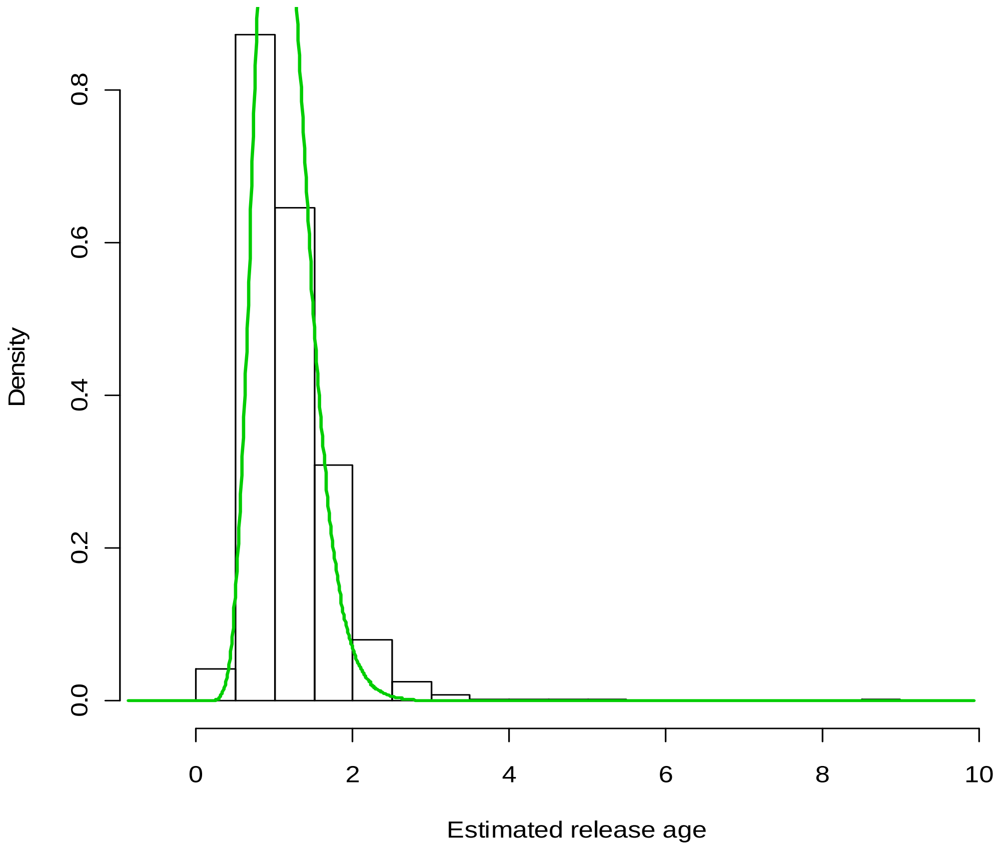
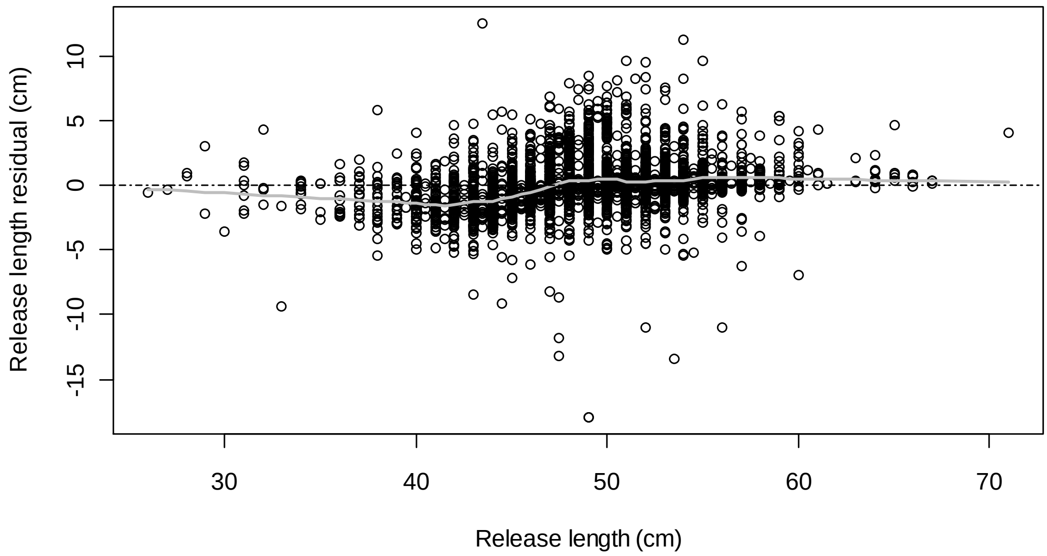

Recapture length residual (cm)
-10
0
10
20
30
30
40
50
60
70
Release length (cm)

Figure A5. Fitted curve and diagnostic plots for VB log k model fitted to all skipjack tag-recapture data using the LEP method with no parameter constraints. The age axis was set assuming L(0) = 20cm.






Release length residual (cm)
10
5
0
-5
-10
-15
30
40
50
60
70
Release length (cm)

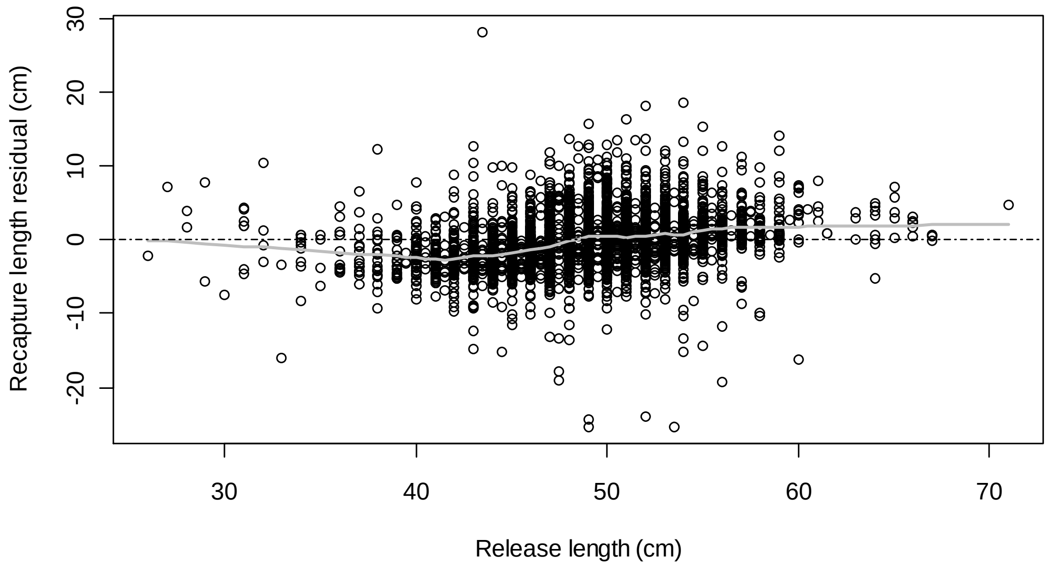
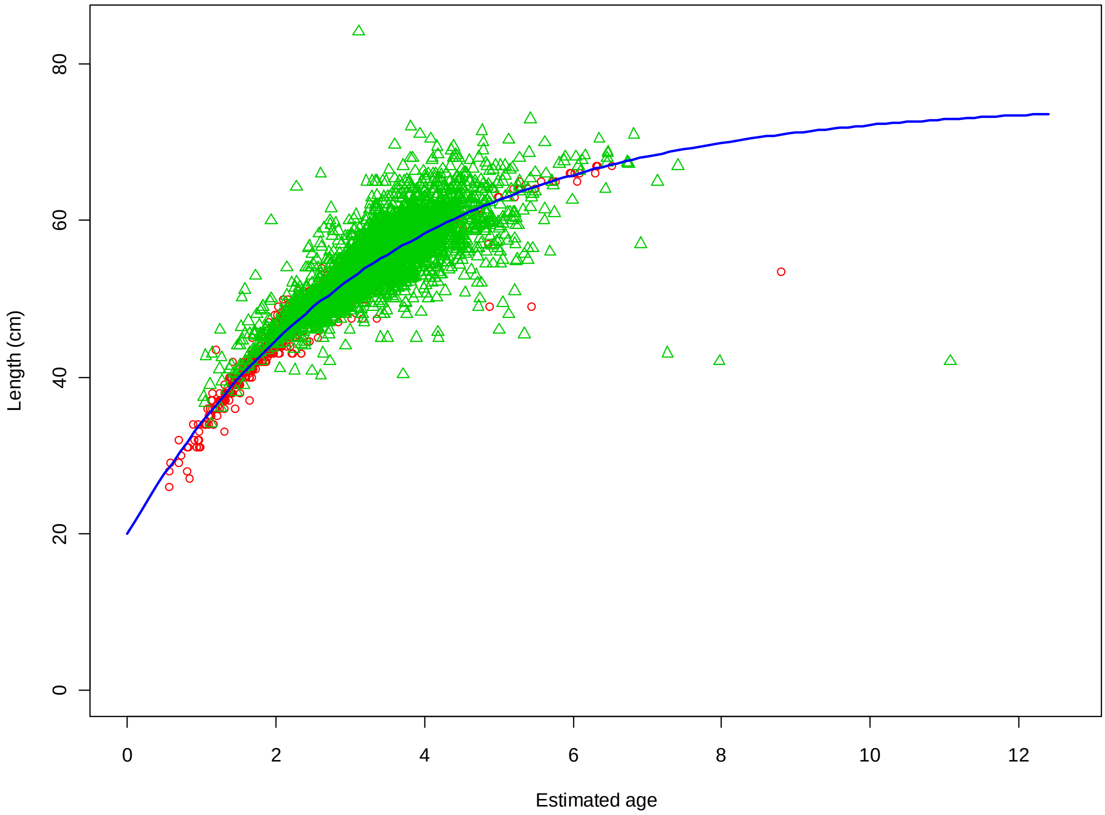
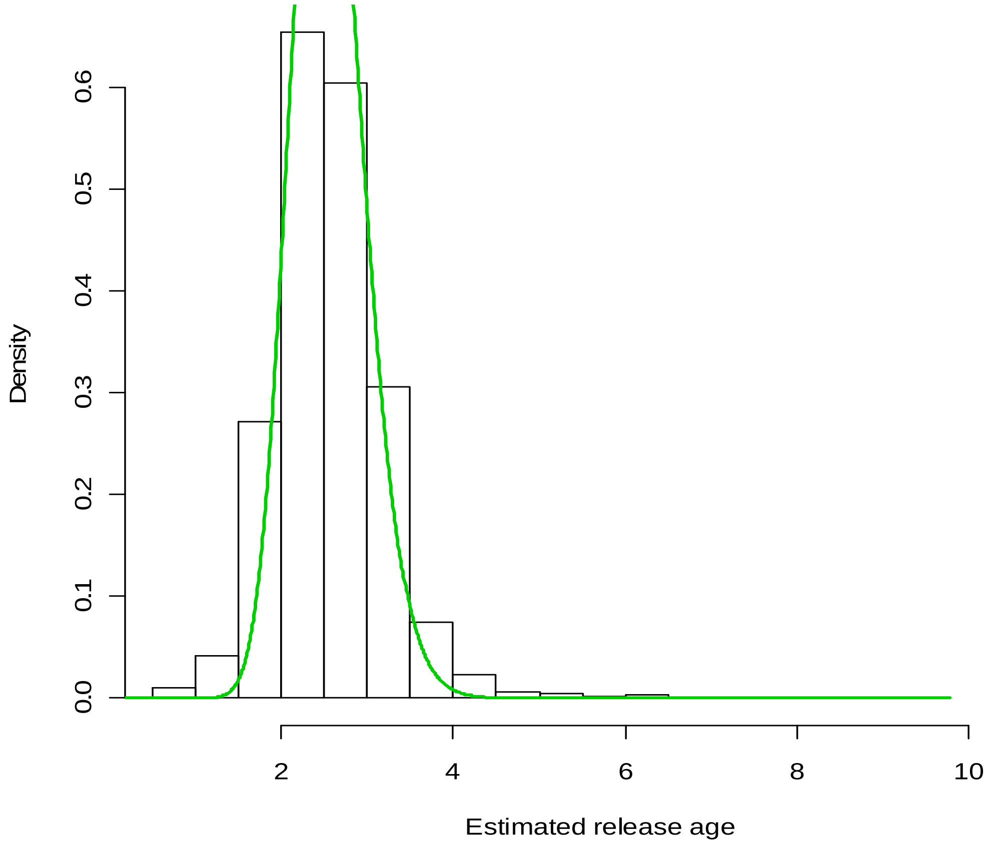
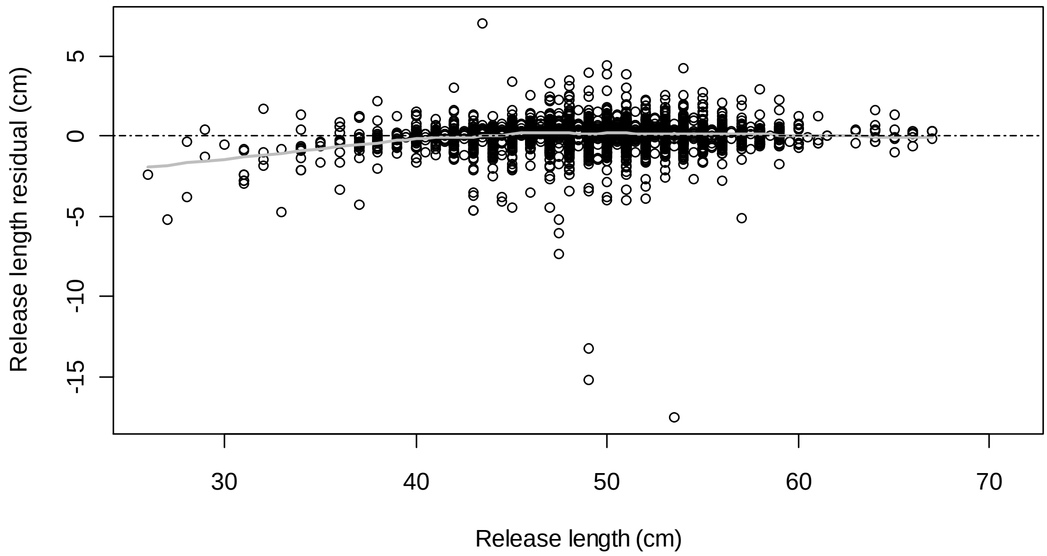

Recapture length residual (cm)
30
20
10
0
-10
-20
30
40
50
60
70
Release length (cm)

Figure A6. Fitted curve and diagnostic plots for VB model fitted to all skipjack tag-recapture data using the LEP method with mean Linf fixed at 75cm (VB75). The age axis was set assuming L(0) = 20cm.






Release length residual (cm)
5
0
-5
-10
-15
30
40
50
60
70
Release length (cm)

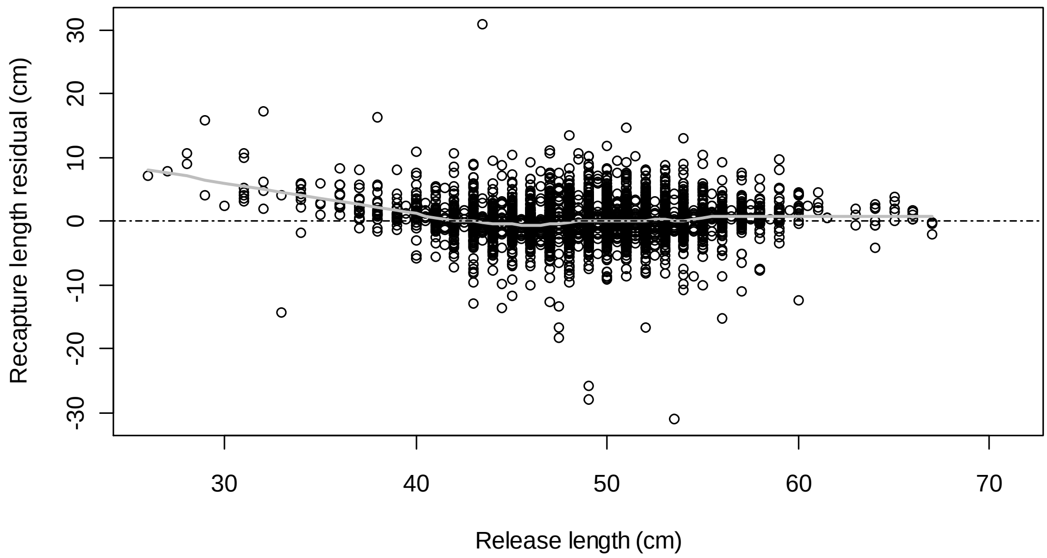
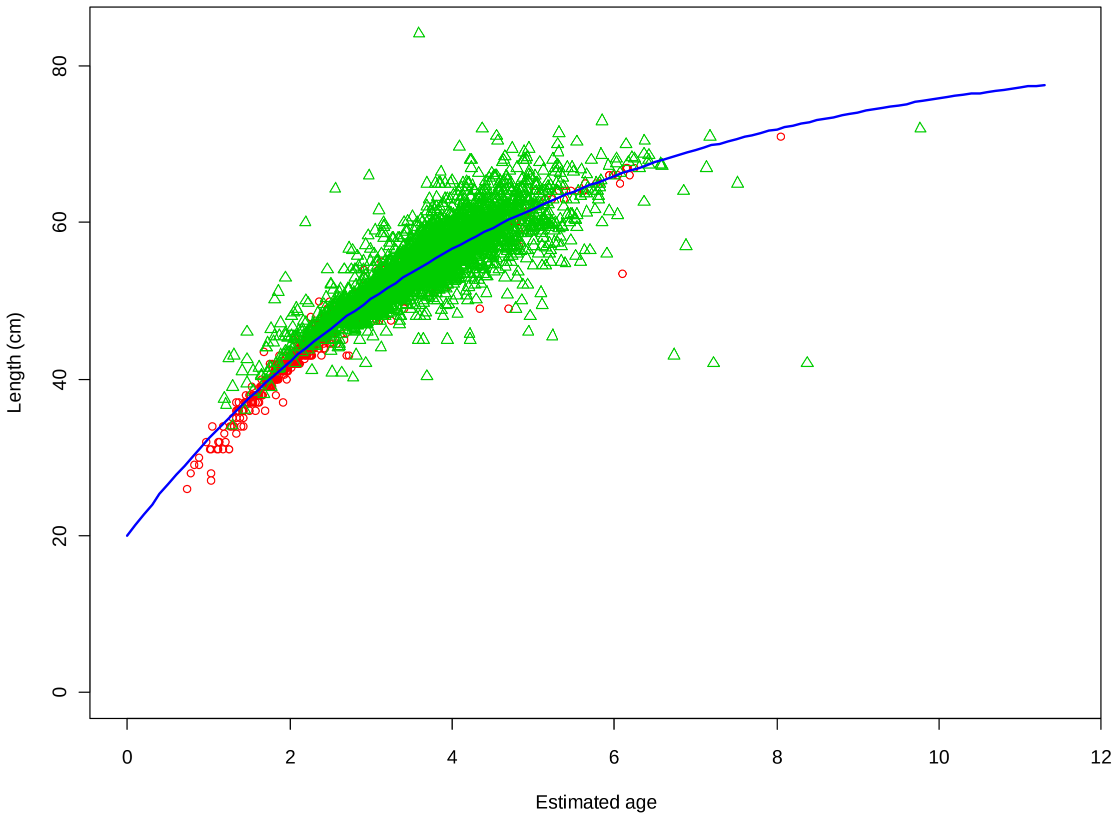
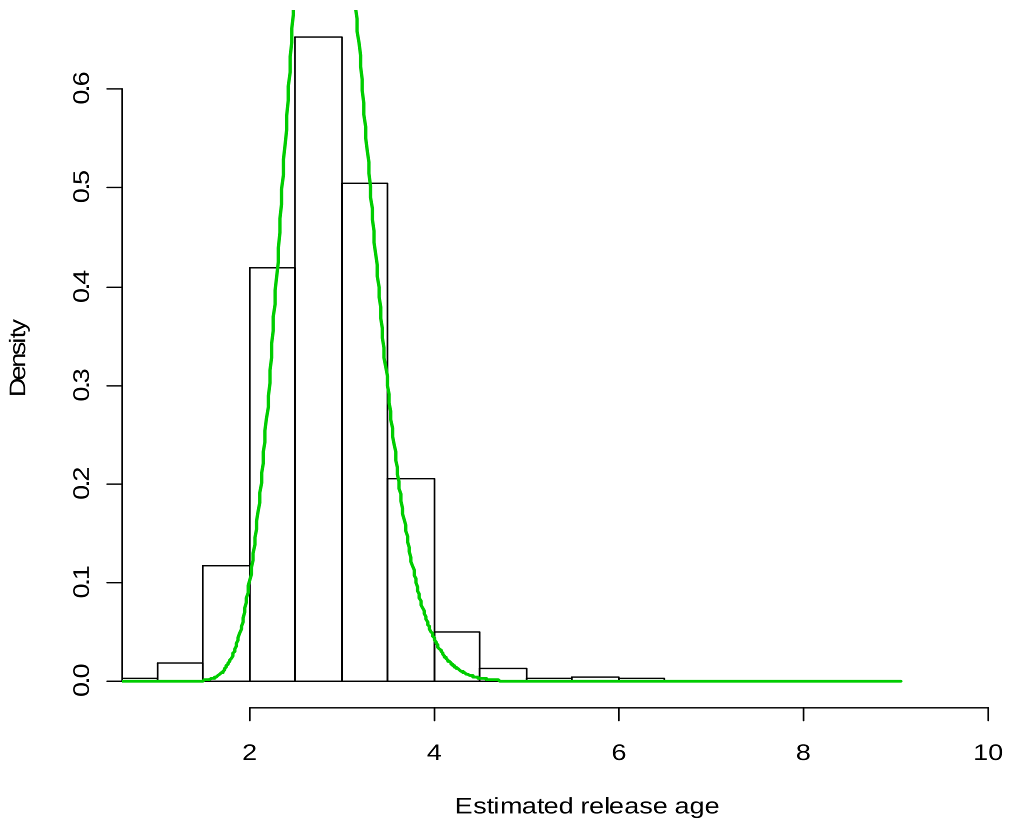
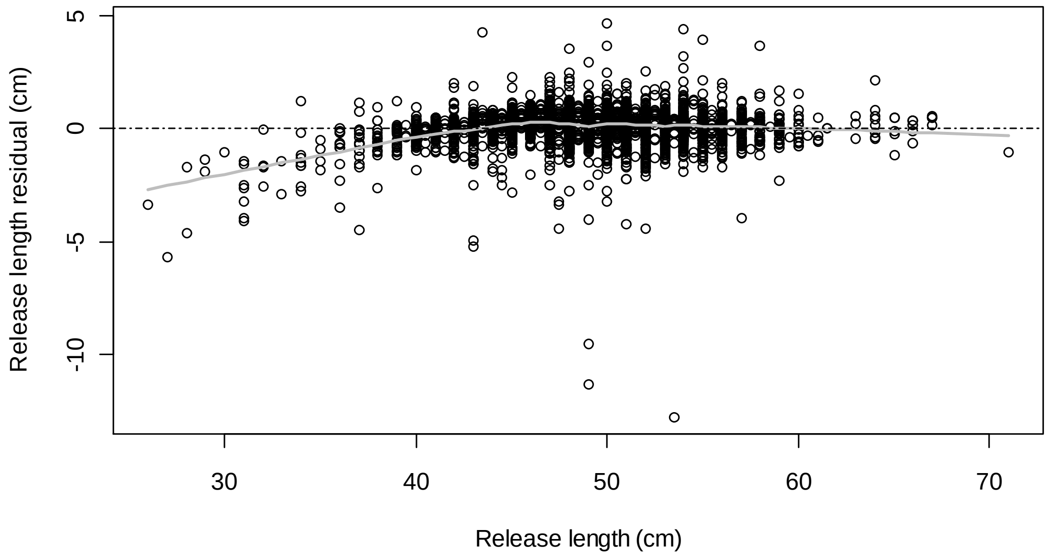

Recapture length residual (cm)
30
20
10
0
-10
-20
-30
30
40
50
60
70
Release length (cm)

Figure A7. Fitted curve and diagnostic plots for VB model fitted to all skipjack tag-recapture data using the LEP method with mean Linf fixed at 83cm (VB83). The age axis was set assuming L(0) = 20cm.






Release length residual (cm)
5
0
-5
-10
30
40
50
60
70
Release length (cm)

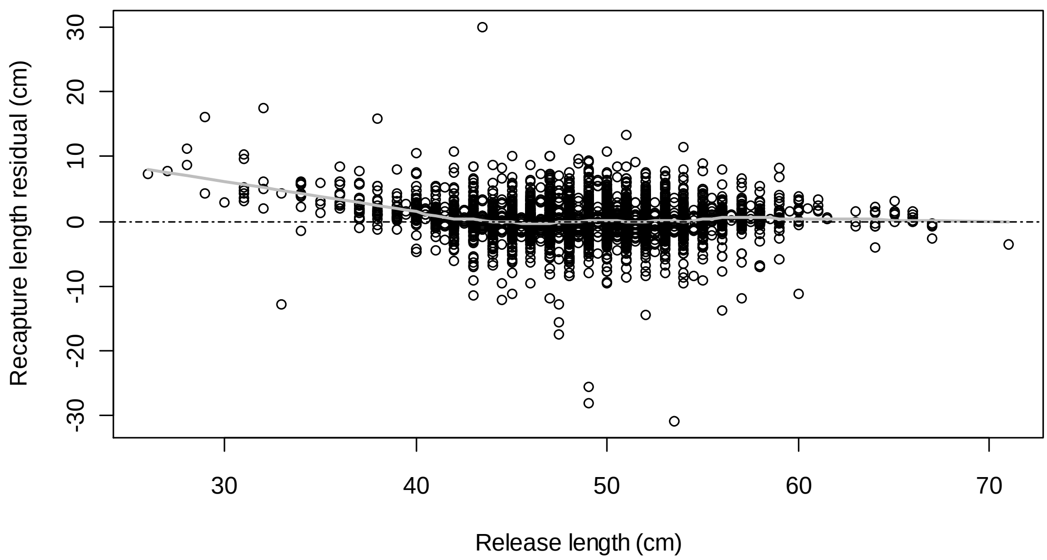

Recapture length residual (cm)
30
20
10
0
-10
-20
-30
30
40
50
60
70
Release length (cm)

Figure A8. Comparison of mean growth curves, all plotted such that L(0) = 20cm.