**Tag-recapture Antarctic toothfish random effects environment**

Authors: D’Arcy N. Webber1, James T. Thorson2

1Quantifish

1 Saint Michaels Crescent

Kelburn

Wellington 6012

New Zealand

2Fisheries Resource Assessment and Monitoring Division

Northwest Fisheries Science Center

National Marine Fisheries Service

National Oceanic and Atmospheric Administration

2725 Montlake Blvd. East

Seattle, WA 98112

**Abstract**

**Keywords:** Antarctic toothfish; time-varying growth; random effect; von Bertalanffy growth function

## 1 Introduction

Growth in fishes arises from complex behavioral trade-offs for each individual in a population. Growth (measured as an increase in standard length) arises from the allocation of energy between increased size (length) and increased condition (i.e., weight at length). Individuals can also increase their rate of energy acquisition via increased foraging activity, but increased foraging may also cause changes in the risk of predation. Finally, environmental conditions modulate the potential for growth, as well as the trade-offs faced by individual fish. For example, increases in water temperature are associated with increased activity levels but also increased energetic demands for maintaining existing body tissue, such that the net effect of changing temperature on growth may vary among individuals. Similarly, changes in survival rates (via fishery harvest) may favor earlier maturation, thus affecting the relative allocation of energy between growth and reproduction.

Many recent studies of captive or wild populations have demonstrated persistent differences in behavioral or phenotypic traits among individuals (termed differences in “personality”). Persistent differences in activity level or tolerance of predation risk (i.e., when foraging in high vs. low-quality habitat) will likely lead to persistent differences in growth rates. Variation in growth rates has been discussed for over 100 years for fish populations. For example, persistent differences in growth rate, combined with size-selective harvest, will result in older individuals being composed primarily of slow-growing individuals (termed “Rosa Lee’s Phenomenon”). Failure to account for persistent differences in growth rate can therefore lead to biased estimation of average growth rates in wild populations, and population dynamics models are increasingly developed to account for these effects (Taylor and Methot 2013).

In addition to persistent differences in growth rate among individuals, populations may experience synchronous variation in growth rates in a given year. For example, annual variation in water temperature may drive changes in metabolic costs as well as food availability in upwelling systems such as the California Current (Black et al. 2010), and this may cause all individuals to have elevated or depressed growth rates in a given year. Recent meta-analytic work has supported the hypothesis that time-variation in average growth rates is the rule rather than the exception for marine populations (Thorson and Minte-Vera In press). Similarly, many studies have illustrated persistent spatial differences in growth rates (Gertseva et al. 2010), e.g., where populations often have a smaller maximum size in warmer waters.

We therefore hypothesize that most marine populations will have multiple sources of variability in individual growth rates, including:

1. Persistent variation among individuals (i.e., some individuals grow faster or to a larger size than others)
2. Spatial variation (i.e., some areas have different size at age on average than others)
3. Temporal variation (i.e., all individuals having elevated or depressed growth rates in a given year)
4. Residual variation (i.e., individuals will have spurts and drops in growth rates).

Mixed-effects models are generally used to decompose variance in a data set into multiple sources of variability (Thorson and Minto In press). Mixed-effects models accomplish this by estimating true size at age as a latent variability, which is integrated across during estimation of growth parameters. Hierarchical models are increasingly advocated as a strategy to partition variability into multiple sources, and hence to interpret which source of variance is worth further study (Larsen et al. 2001, Gelman 2005).

Growth rates in natural populations are generally estimated using capture-mark-recapture (CMR) models. In CMR studies for marine fishes, thousands of individuals are typically captured and marked with a persistent and unique tag, and then measured and released. If a tagged individual is recaptured, it is often killed and its age at recapture identified via analysis of hard parts. Given the known time between capture and recapture, its size and age at both times can be calculated. This study design allows a model for growth increments to be fitted to size and age data for two times for each individual in a data set.

Many studies have examined the impact of persistent or residual variation on estimation of growth patterns using CMR data for wild marine populations (Sainsbury 1980, Francis 1988). However, few studies have explicitly modeled multiple sources of variability in growth rates. One exception is Shelton et al. (2013), which modeled persistent, temporal, and residual variation in growth rates in the analysis of data from a growth experiment for steelhead trout. In this study, we additionally incorporate sex-specific differences and spatial variation in growth rates, and implement the model using maximum marginal likelihood estimation. The code is distributed as a publicly available package in the R statistical environment, and is illustrated using CMR data for Antarctic toothfish (*Dissostichus mawsoni*). Using these data, we show that …. [1-2 sentence summary of most important result for toothfish].

**2 Methods**

We start with the specialized von Bertalanffy growth function:

(1)

where *dL/dt* is change in length as a function of time, *a* scales with energy acquisition, and *k* represents metabolic upkeep costs. Variation among individuals can be estimated by treating each individual's demographic parameters as a random effect that arises from a population-level distribution (Thorson and Minto In press). Here, we treat metabolic upkeep costs as varying among individuals, due to differences in ambient temperature, intrinsic variation in activity levels, and other factors:

(2)

where *ki* is the upkeep costs for individual *i* and *μlog(k)*  is the average log(*k*) among individuals in the population.

Individuals that are more highly active may obtain more food (increased *a*) and simultaneously have greater upkeep costs (increased *k*). Following Shelton et al. (2013), we include this association via the following equation:

(2)

where *Ψ* determines whether upkeep costs (*ki*) and energy acquisition rate (*ai*) are highly correlated (*Ψ=1*) or independent among individuals (*Ψ=0*), and *γ* controls the average energy acquisition rate among individuals.

We also assume that environmental conditions experienced by each individual will vary over time, such that each individual will also have variation over time in their upkeep costs (*ki*), in turn affecting their access to food (*ai*). In particular, we assume that upkeep costs for a given individual *i* for a short time interval *t* can be approximated by a constant value (*ki,t*), which will vary around the average value for this individual (*ki*). In the proceeding, we use a time interval of one week (7 days), while confirming that results are similar for other small time intervals. Integration then yields:

(3)

where

where *Δt* is the number of time-periods elapsed between length intervals and *nΔ* is the number of time-periods per year, such that *k* is the Brody growth coefficient (with units yr-1). The average asymptotic length for individuals in the population can be found by setting Eqn. 1 equal to zero, substituting in Eqn. 2, and rearranging:

(3)

Readers are referred to Shelton et al. (2013) for an expanded model that also incorporates variability in *γ* over time.

## 2.2 Estimation

We implement this model using Template Model Builder (TMB;available from: <https://github.com/kaskr/adcomp>) software called from the R statistical environment (R Core Development Team 2013) and all code necessary for replicating the case study and simulation experiment are publicly available (Appendix A). Estimation was conducted using maximum marginal likelihood (Thorson and Minto In press), while marginalizing across the probability of all random variables representing variation among individuals and over time.

**2.0 Case study application**

We fit the model to data collected in the Antarctic toothfish tagging programme wherein each individual is uniquely tagged and aged at recapture.

Initial model exploration using the Antarctic toothfish data set illustrated that the model had little ability to identify the value of asymptotic maximum length because most individuals were captured during ages experiencing fast growth. We therefore include a penalty on the value of average axymptotic maximum length.

Set at mu=and sig= (Dunn et al. 2006).

* Data grooming. Removed unsexed individuals. Linked aged fish that were tagged to the tag-recapture data set. 315 individuals. 166 female, 149 male (check this)
* Did fits with no random effects, k, z, k and z…
* Calendar year effect incorporated as a white-noise random effect. Need to estimate a year effect for years before the data begins. I.e. from the year the fish were born.
* Fixed psi=0
* Priors?
* Seasonality could not be incorporated due to timing of fishery

## 2.1 Simulation

A simulation study was done to validate the model. Four different scenarios were simulated including: no random-effects, random-effects for *k* only, random-effects for *z* only, and random-effects for *k* and *z*. A power analysis was done for each of the four scenarios where 50, 100, 250 and 500 individuals were sampled. We did 200 replicates for each of the scenario power combinations.

The model parameters *t0*, *k*, Linf and the coefficient of variation of growth were taken from Dunn et al. (2006) and used to derive the values of all parameters in the simulation study (Table 1). Where required parameters were converted to have weekly units (i.e. *k*, *t0*, age). The parameters *L0* and gamma can be calculated as:

(2)

and

(2)

Initially psi was fixed at a value. Later fixed at zero. Table 2.

* Set values for sdk, sdz, sdobs that resulted in reasonable variation in indiv trajectories. These pars were sex specific.
* Sex sampled with replacement from the obs sex. Age1, age2 and liberty all sampled with replacement from obs, then one of the three is chosen at random to be calculated from the remaining two. Round to the nearest integer.
* Fit the model to the data.
* No prior on L0, no prior on Linf

## 3 Results

## 3.1 Simulation

* Found that psi could not be estimated reliably but when psi was fixed at 0 the model could recover the pars well.
* Do table showing number of pdh fits for each of the 4 scenarios. Also show true par values, and mean/median of estimates.

## 3.2 Estimation

## 4 Discussion

* Mention how when the model was fit to annual rather than daily growth increments that the estimated parameters resulted in biased high growth schedules.
* Growth shock or growth retardation due to tagging. Tag shock has been estimated before as about 0.7 year lost in growth. Confounded by live/dead.
* Snapper may also have some data like this.
* Seasonality

## 5 Conclusions

**Bibliography**

Black, B.A., Schroeder, I.D., Sydeman, W.J., Bograd, S.J., and Lawson, P.W. 2010. Wintertime ocean conditions synchronize rockfish growth and seabird reproduction in the central California Current ecosystem. Can. J. Fish. Aquat. Sci. **67**(7): 1149–1158.

Francis, R. 1988. Maximum likelihood estimation of growth and growth variability from tagging data. N. Z. J. Mar. Freshw. Res. **22**(1): 43–51. [accessed 19 December 2014].

Gelman, A. 2005. Analysis of variance—why it is more important than ever. Ann. Stat. **33**(1): 1–53.

Gertseva, V.V., Cope, J.M., and Matson, S.E. 2010. Growth variability in the splitnose rockfish Sebastes diploproa of the northeast Pacific Ocean: pattern revisited. Mar. Ecol. Prog. Ser. **413**: 125–136. [accessed 3 May 2013].

Larsen, D.P., Kincaid, T.M., Jacobs, S.E., and Urquhart, N.S. 2001. Designs for Evaluating Local and Regional Scale Trends. BioScience **51**(12): 1069–1078. doi: 10.1641/0006-3568(2001)051[1069:DFELAR]2.0.CO;2.

R Core Development Team. 2013. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. Available from http://www.R-project.org/.

Sainsbury, K.J. 1980. Effect of individual variability on the von Bertalanffy growth equation. Can. J. Fish. Aquat. Sci. **37**(2): 241–247. [accessed 19 December 2014].

Shelton, A.O., Satterthwaite, W.H., Beakes, M.P., Munch, S.B., Sogard, S.M., and Mangel, M. 2013. Separating intrinsic and environmental contributions to growth and their population consequences. Am. Nat. **181**(6): 799–814. [accessed 8 October 2013].

Taylor, I.G., and Methot, R.D. 2013. Hiding or dead? A computationally efficient model of selective fisheries mortality. Fish. Res. **142**: 75–85. [accessed 20 April 2013].

Thorson, J.T., and Minte-Vera, C. In press. Relative magnitude of cohort, age, and year effects on size at age of exploited marine fishes. Fish. Res.

Thorson, J.T., and Minto, C. In press. Mixed effects: a unifying framework for modelling in aquatic ecology. ICES J. Mar. Sci.

Table 1 – Parameter values from Dunn et al. 2008. Coefficient of variation.

|  |  |  |  |
| --- | --- | --- | --- |
| Parameter | Units | Female | Male |
| *t0* | years | 0.021 | -0.256 |
| *k* | years-1 | 0.090 | 0.093 |
| Linf | cm | 180.20 | 169.07 |
| c.v. | – | 0.102 | 0.102 |

Table 2a – Parameter values used for all scenarios in the simulation experiment.

|  |  |  |
| --- | --- | --- |
| **Parameter** | **Female** | **Male** |
| *L0* | -0.3409 | 3.9777 |
| *k* | 0.00173 | 0.00179 |
| gamma | 0.3119 | 0.3024 |
| Psi | 0.0 | 0.0 |

Table 2a – Parameter values that vary among scenarios in the simulation experiment.

|  |  |  |  |
| --- | --- | --- | --- |
| Scenario | Parameter | Female | Male |
| No variation | Sd\_obs | 0.102 | 0.102 |
|  | Sd\_k | - | - |
|  | Sd\_z | - | - |
| Variation in k | Sd\_obs | 0.05 | 0.05 |
| Sd\_k | 0.1 | 0.2 |
|  | Sd\_z | - | - |
| Variation in z | Sd\_obs | 0.05 | 0.05 |
| Sd\_k | - | - |
|  | Sd\_z | 0.2 | 0.4 |
| Variation in k and z | Sd\_obs | 0.05 | 0.05 |
| Sd\_k | 0.1 | 0.2 |
| Sd\_z | 0.2 | 0.4 |

Table 2 – List and definition of symbols used in the text and equations.

|  |  |  |
| --- | --- | --- |
| **Parameter name** | **Symbol** | **Units** |
| The hypothetical length at |  | cm |
| Anabolic rate |  |  |
| Energy expenditure rate (“Brody growth coefficient”) |  |  |
| Anabolism coefficient |  |  |
| Allometric scaling of anabolism and energy expenditure |  |  |
| Time variation |  |  |
| Asymptotic maximum length |  | cm |
| The hypothetical time at which |  | days |
| Observation error standard deviation |  |  |
| Catabolic rate standard deviation |  |  |
|  |  |  |
|  |  |  |
| **Data** | **Symbol** | **Units** |
| Time interval |  | days |
| Length |  | cm |
| **Indices** | **Symbol** | **Units** |
| Individual |  | - |
| Time |  | days |
| Sex |  | - |
| Area |  |  |

Table 2 – List of simulation runs.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Random effects | pdH | L0 | k | gamma | Sd obs |  |  |  |
|  |  |  |  |  |  |  |  |  |
| None |  |  |  |  |  |  |  |  |
| k |  |  |  |  |  |  |  |  |
| z |  |  |  |  |  |  |  |  |
| k and z |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |

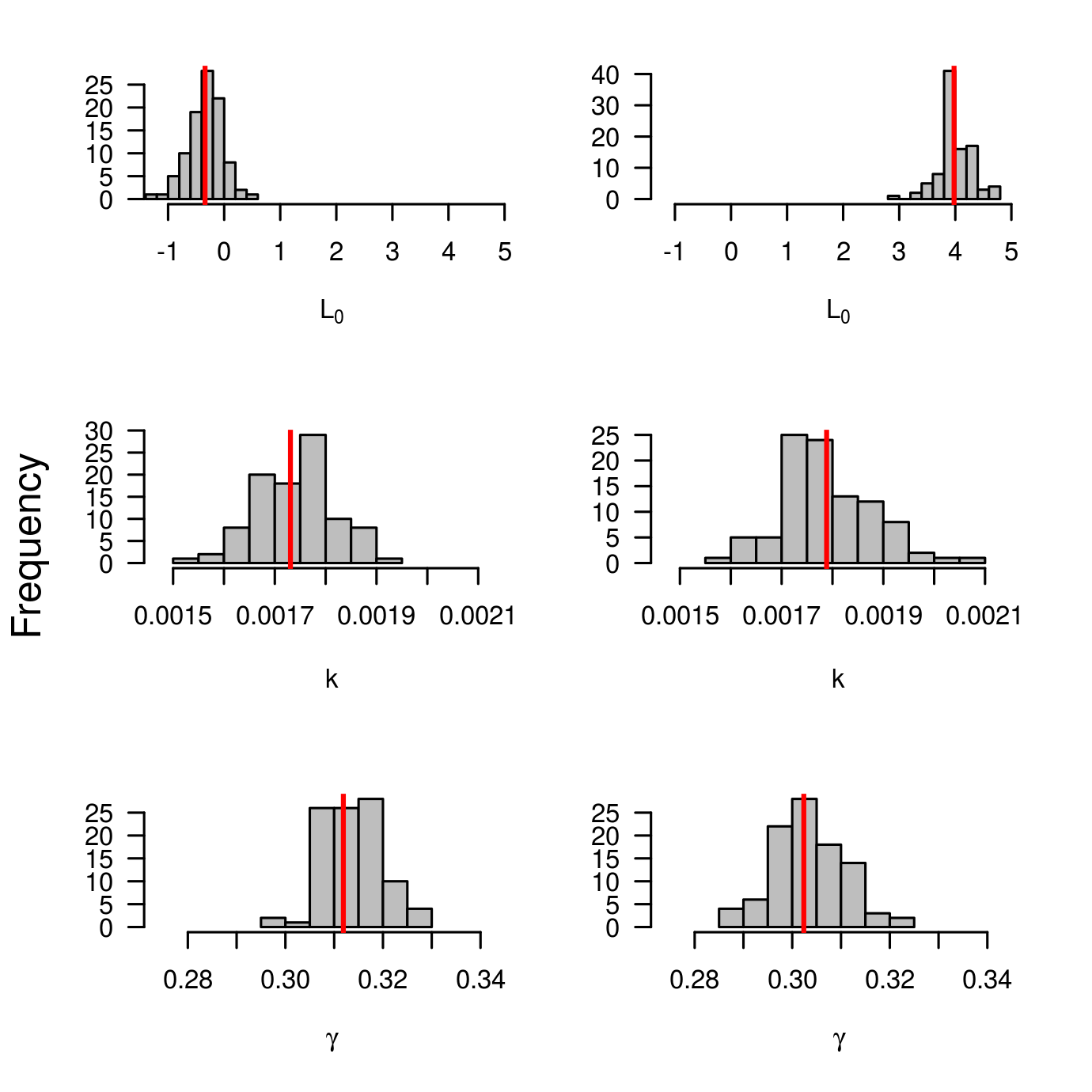


Figure 1 - estimates

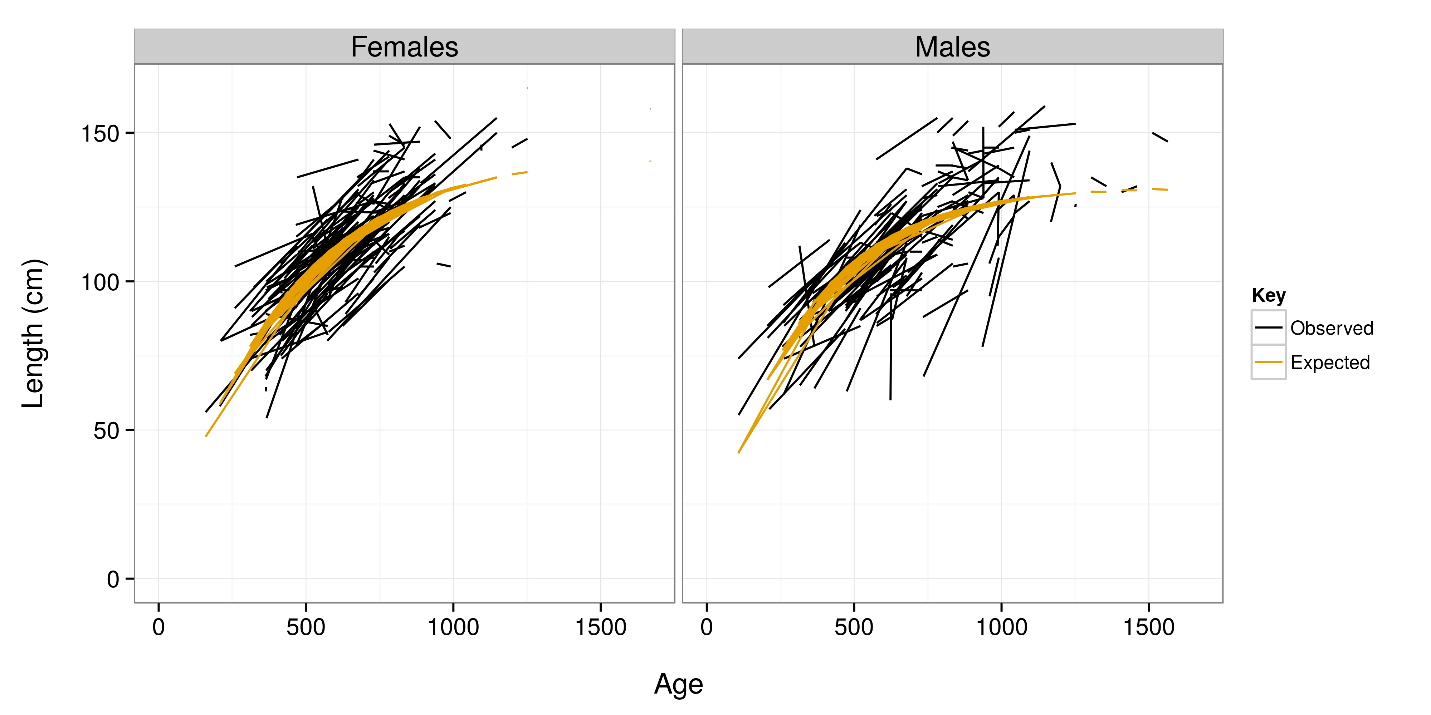


Figure X – no random effects

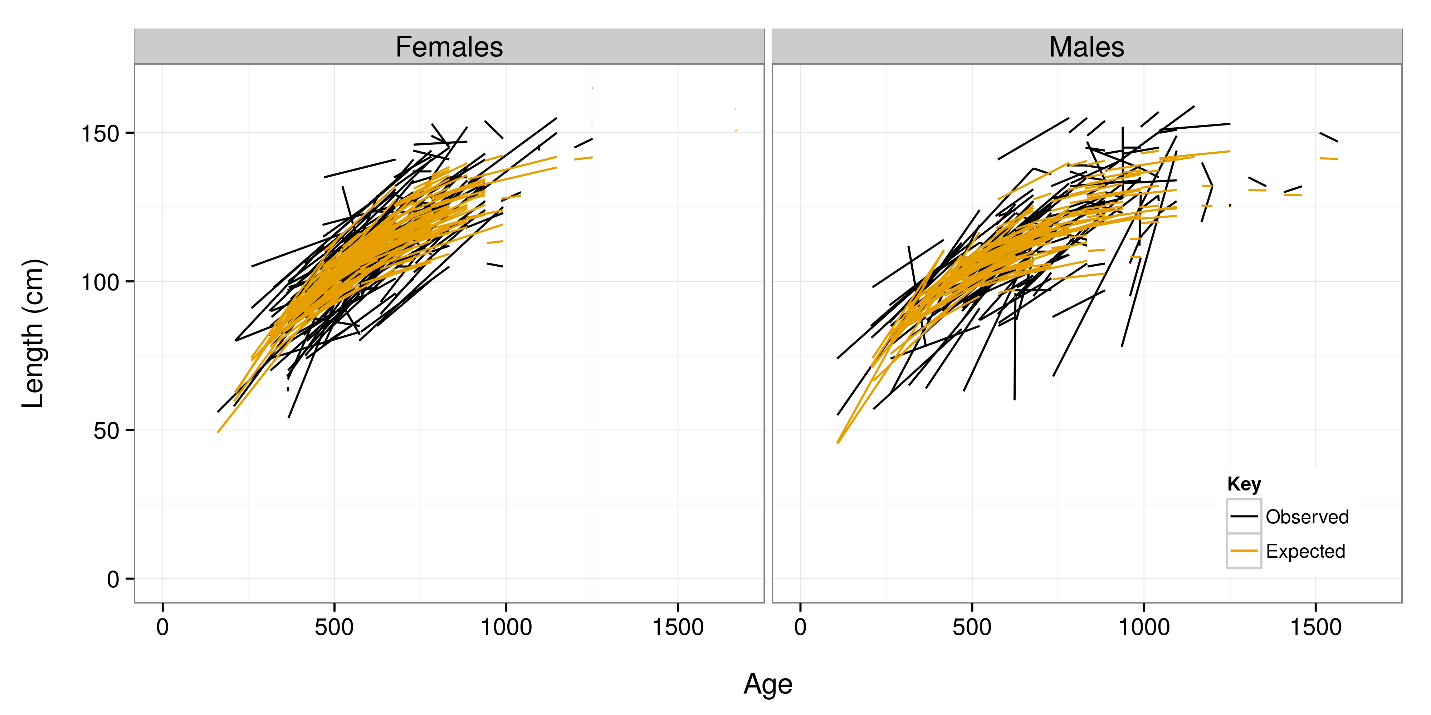


Figure X – random effects of k

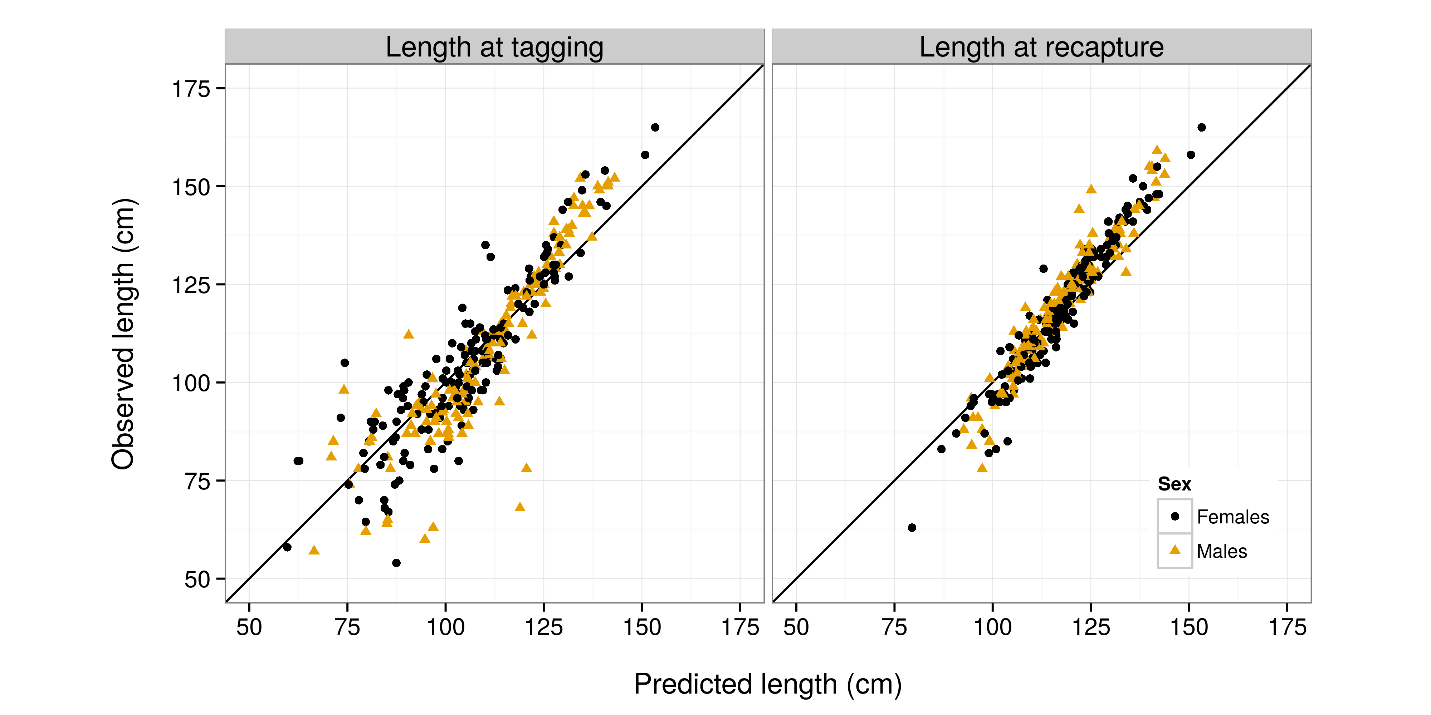


Figure X

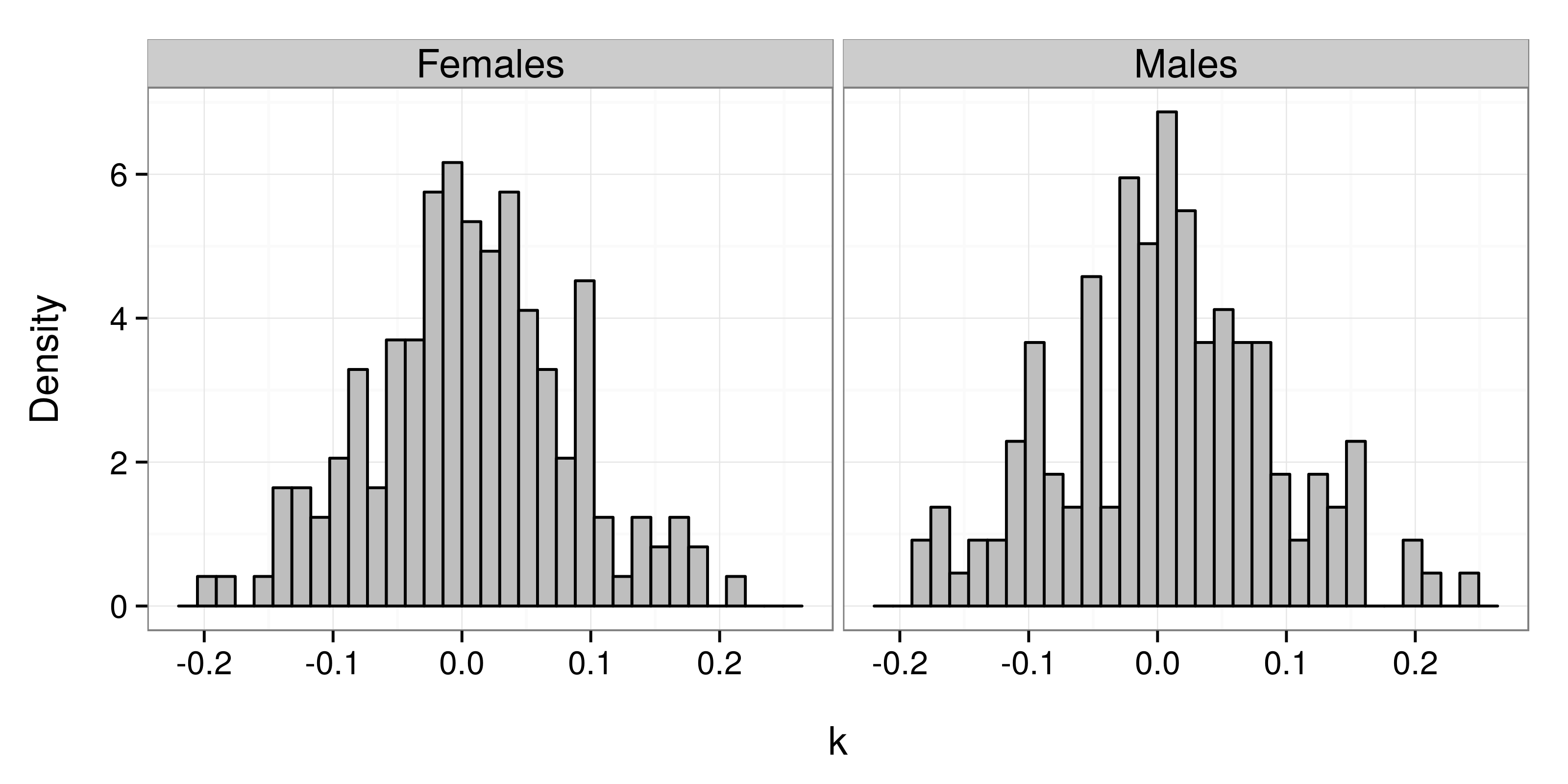


Figure X