# Replication and Improvement in Parameterizations of von Bertalanffy Growth Function through Bayesian Analysis

Abhiroop Yerramilli

STAT 3069

Assignment 2

22 December 2023

# Table of Contents

# 1. Introduction

## 1.1 Background of the Problem

Body mass is critical to all organisms, and possesses a very crucial relationship with a lot of traits [1, Introduction]. Growth (change in body size over time) is a practical way of exploring these relationships, and accurate predictions are required for most animal and plant-based industries [1, Introduction]. Traditionally the models being used have considered mass scaling components of anabolism (A), and catabolism (B) stable or not very varied, which makes the model non-scalable to various species without bad fits [1, Introduction]. A more flexible way to model these relationships would be required which does not put restrictions on variations in A and B values [1, Introduction]. The traditionally used van Bertanlanffy Growth Function (VBGF) does not allow for variation in both A and B together, which means the model is a lot less flexible and non-scalable [1, Introduction]. Different parameterizations of the VBGF can be used, with data collected from pelagic and benthic invertebrates which have shown mass scaling variation, which would provide a new framework for better and more reliable growth predictions [1, Introduction].

## 1.2 Dataset Description

The dataset used in the study was collected using the Web of Knowledge using a lot of search terms [1, The data set]. The data collected is mass-at-age data for seven pelagic and five benthic invertebrate species, as the diversity in mass scaling makes these two invertebrate species ideal for exploring variation [1, The data set]. Looking at the dataset we can conclude below [1, Data Availability]. The below conclusions were done based on Code 5.1 and [1]:

- There are columns for time, mass, individual, species, binomial, and source.
- The time mainly refers to age, while mass is the mass at that age (here at time t).
- Individual refers to the number of individuals considered for the value.
- Each species has its own separate species code, and binomial refers to the name of the species.
- Source is used to give additional information on where each species data was sourced from.

There are a total of 12 species, and each species has a different number of mass at age entries [1, Data Availability]. Every data point only has one individual considered [1, Data Availability].

Data collected only includes growth data in normal circumstances, in its natural setting [1, The data set]. Data was collected from hatch to reproductive maturity, and extracted from Graphs using WebPlotDigitizer [1, The data set].

## 1.3 Study Summary

Many growth functions have been used in the past for many groups of organisms, but some of these functions are not very good at reflecting variation growth rate [1, Abstract]. The study aims to present five different parameterizations of the von Bertalanffy Growth Function (VBGF) and apply them to twelve invertebrate species to get better and more reliable growth predictions [1, Abstract]. The study also tried to find a common best-fitting plot but failed to find the best fitting plot and instead suggested a combination of the parameterizations as a more flexible solution [1, Discussion].

## 1.4 Study Conclusions

The study compares the different parameterizations of the von Bertalanffy Growth Function (VBGF) and finds the best parameterization based on the species [1]. The study wanted to see if a parameterization can be generalized to a species or all species as a whole, but based on the plots and comparisons, there is a lack of agreement on which parameterization is the best fitting model [1, Discussion]. The study suggests a combination of parameterizations would be more flexible and reliable while capturing growth rate variations much better [1, Discussion].

Another suggestion is to ensure exceptions, like organisms increasing energy reserves with age, should be considered and evaluated differently in the current setting, and that flexibility will help with this aspect too [1, Discussion]. The study concludes by stating that the growth curve for the different parameterizations of the VBGF helps find the optimal value of the mass scaling exponent (A) which can then be used to improve current models and to model growths [1, Discussion]. The study also says that these parameterizations can also be explored in exploring growth relations with other aspects of the aquaculture like temperature and reproduction [1, Discussion].

# 2. Bayesian Model Development

## 2.1 Translation from Frequentist to Bayesian

First we look at the frequentist model and identify the components of the model. We are considering the parameterization of the exponential model and converting that model to a Bayesian model for comparison [1, Materials and methods]. Below is the model:

$$m = m_0 \exp(k(t - t_0))$$

Source:

This model has one parameter ‘k’ which is estimated iteratively [1, Materials and methods]. m is the mass being predicted. m0 is first the mass at time 0, and then this is used to fit the model [1, Materials and methods].

Before further model building there are a few points of consideration:

- There are different species at play. These species are different but also have common elements. Hence we cannot use a pooled or unpooled model. Instead a multi level partially pooled model would be the best way to implement the model.

- Growth staying proportional to mass across species would be way more probable, hence the k values can be conserved.

## 2.2 Model Specification

**Likelihood and Structure:**

For each observation i, where each observation means number of mass-at-age values we have for that species, we have

mass[i] ~ lognormal(log(mu), sigma)

mu = mu0 [species[i]] * exp(k[species[i]] * (time[i]))

**Priors:**

k ~ lognormal(log(k_mean), k_disp)

sigma ~ exponential(1)

**Hyperpriors:**

k_mean ~ exponential(1)

k_disp ~ exponential(1)

The reason we choose based on species is because the parameters mu0 and k are estimated based on species

## 2.3 Prior Justification

We can interpret and justify the priors below:

- The first prior we have, k, has a log-normal distribution. The reason is that we are working with partially pooled estimates. k is positive and non-zero for the equation to be practical, this means that the distribution of k must be more dependent on the data and less on the mean.
- The hyperprior mean is based on the fact that we are log(exp(1)) which makes k less dependent on this mean. Also, this means that smaller values are more probable which is what we want.
- The hyperprior k_disp is based on the assumption that there isn’t a lot of difference between k estimates across species. This goes in line with the points of consideration in 2.1.
- For the sigma choice, we use exponential distribution to get smaller deviations with a rate parameter of 1. This is mainly because we want to have a more general distribution as the distribution can vary a lot between different species and having a more generalized range is better.
- Another point to note is that we are using a lognormal distribution for mass as mass cannot be negative, and this also indicates positive skewness.

# 3. Interpretation and Results

## 3.1 Model Summary

First we look at the model summary. Because we have a stan model we cannot just use precis function. Instead we use the summary function to get the summary based on the below factors:

Mean and median: mean and median value of the parameter.

Rhat: Potential scale reduction factor. This refers to convergence. Values around 1 mean good convergence.

q5, q95, sd: q5 and a95 refer to the 5th and 95th quantile values, and sd refers to the standard deviation.

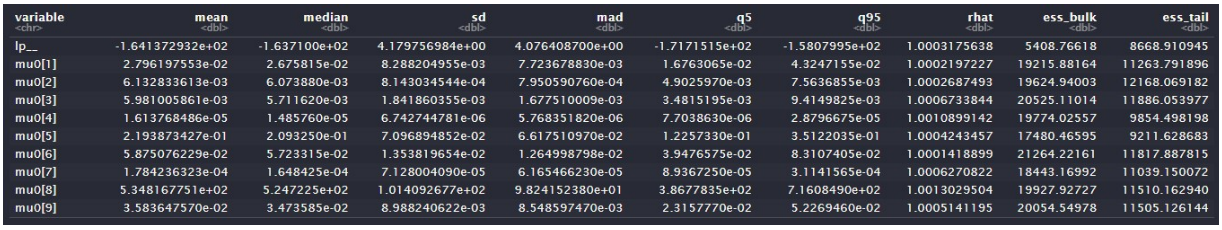
lp__: log posterior values. Mainly used to compare between models. Higher values indicate a better fit.

Ess_bulk and ess_tail: Effective sample size. Higher values desired.

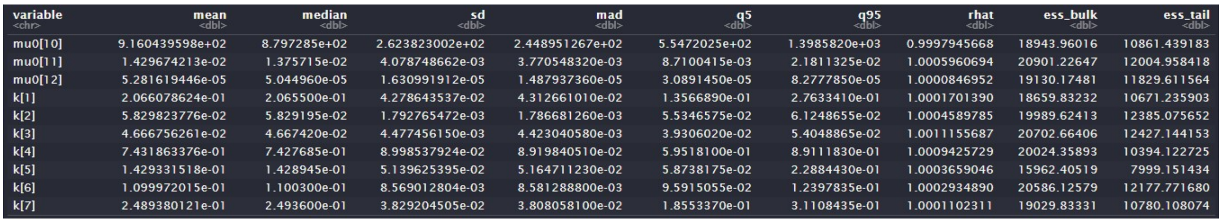
Mad: Mean absolute deviation is a measure of spread of data not influenced by outliers.

The results from running summary on the model are given below:

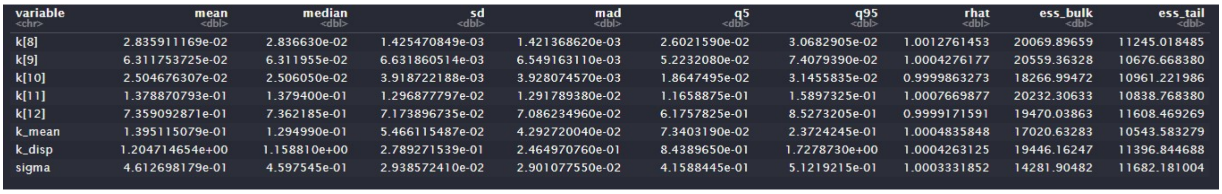
*[Generated from 5.7 after data input (5.1) and making the model (5.3, 5.4, 5.5, 5.6)]*



| variable | mean | median | sd | mad | q5 | q95 | rhat | ess_bulk | ess_tail |
|---|---|---|---|---|---|---|---|---|---|
| <chr> | <dbl> | <dbl> | <dbl> | <dbl> | <dbl> | <dbl> | <dbl> | <dbl> | <dbl> |
| lp__ | -1.641372932e+02 | -1.637100e+02 | 4.179756984e+00 | 4.076408700e+00 | -1.7171515e+02 | -1.5807995e+02 | 1.0003175638 | 5408.76618 | 8668.910945 |
| mu0[1] | 2.796197553e-02 | 2.675815e-02 | 8.288204955e-03 | 7.723678830e-03 | 1.6763065e-02 | 4.3247155e-02 | 1.0002197227 | 19215.88164 | 11263.791896 |
| mu0[2] | 6.132833613e-03 | 6.073880e-03 | 8.143034544e-04 | 7.950590760e-04 | 4.9025970e-03 | 7.5636855e-03 | 1.0002687493 | 19624.94003 | 12168.069182 |
| mu0[3] | 5.981005861e-03 | 5.711620e-03 | 1.841860355e-03 | 1.677510009e-03 | 3.4815195e-03 | 9.4149825e-03 | 1.0006733844 | 20525.11014 | 11886.053977 |
| mu0[4] | 1.613768486e-05 | 1.485760e-05 | 6.742744781e-06 | 5.768351820e-06 | 7.7038630e-06 | 2.8796675e-05 | 1.0010899142 | 19774.02557 | 9854.498198 |
| mu0[5] | 2.193873427e-01 | 2.093250e-01 | 7.096894852e-02 | 6.617510970e-02 | 1.2257330e-01 | 3.5122035e-01 | 1.0004243457 | 17480.46595 | 9211.628683 |
| mu0[6] | 5.875076229e-02 | 5.723315e-02 | 1.353819654e-02 | 1.264998798e-02 | 3.9476575e-02 | 8.3107405e-02 | 1.0001418899 | 21264.22161 | 11817.887815 |
| mu0[7] | 1.784236323e-04 | 1.648425e-04 | 7.128004090e-05 | 6.165466230e-05 | 8.9367250e-05 | 3.1141565e-04 | 1.0006270822 | 18443.16992 | 11039.150072 |
| mu0[8] | 5.348167751e+02 | 5.247225e+02 | 1.014092677e+02 | 9.824152380e+01 | 3.8677835e+02 | 7.1608490e+02 | 1.0013029504 | 19927.92727 | 11510.162940 |
| mu0[9] | 3.583647570e-02 | 3.473585e-02 | 8.988240622e-03 | 8.548597470e-03 | 2.3157770e-02 | 5.2269460e-02 | 1.0005141195 | 20054.54978 | 11505.126144 |



| variable | mean | median | sd | mad | q5 | q95 | rhat | ess_bulk | ess_tail |
|---|---|---|---|---|---|---|---|---|---|
| <chr> | <dbl> | <dbl> | <dbl> | <dbl> | <dbl> | <dbl> | <dbl> | <dbl> | <dbl> |
| mu0[10] | 9.160439598e+02 | 8.797285e+02 | 2.623823002e+02 | 2.448951267e+02 | 5.5472025e+02 | 1.3985820e+03 | 0.9997945668 | 18943.96016 | 10861.439183 |
| mu0[11] | 1.429674213e-02 | 1.375715e-02 | 4.078748662e-03 | 3.770548320e-03 | 8.7100415e-03 | 2.1811325e-02 | 1.0005960694 | 20901.22647 | 12004.958418 |
| mu0[12] | 5.281619446e-05 | 5.044960e-05 | 1.630991912e-05 | 1.487937360e-05 | 3.0891450e-05 | 8.2777850e-05 | 1.0000846952 | 19130.17481 | 11829.611564 |
| k[1] | 2.066078624e-01 | 2.065500e-01 | 4.278643537e-02 | 4.312661010e-02 | 1.3566890e-01 | 2.7633410e-01 | 1.0001701390 | 18659.83232 | 10671.235903 |
| k[2] | 5.829823776e-02 | 5.829195e-02 | 1.792765472e-03 | 1.786681260e-03 | 5.5346575e-02 | 6.1248655e-02 | 1.0004589785 | 19989.62413 | 12385.075652 |
| k[3] | 4.666756261e-02 | 4.667420e-02 | 4.477456150e-03 | 4.423040580e-03 | 3.9306020e-02 | 5.4048865e-02 | 1.0011155687 | 20702.66406 | 12427.144153 |
| k[4] | 7.431863376e-01 | 7.427685e-01 | 8.998537924e-02 | 8.919840510e-02 | 5.9518100e-01 | 8.9111830e-01 | 1.0009425729 | 20024.35893 | 10394.122725 |
| k[5] | 1.429331518e-01 | 1.428945e-01 | 5.139625395e-02 | 5.164711230e-02 | 5.8738175e-02 | 2.2884430e-01 | 1.0003659046 | 15962.40519 | 7999.151434 |
| k[6] | 1.099972015e-01 | 1.100300e-01 | 8.569012804e-03 | 8.581288800e-03 | 9.5915055e-02 | 1.2397835e-01 | 1.0002934890 | 20586.12579 | 12177.771680 |
| k[7] | 2.489380121e-01 | 2.493600e-01 | 3.829204505e-02 | 3.808058100e-02 | 1.8553370e-01 | 3.1108435e-01 | 1.0001102311 | 19029.83331 | 10780.108074 |



| variable | mean | median | sd | mad | q5 | q95 | rhat | ess_bulk | ess_tail |
|---|---|---|---|---|---|---|---|---|---|
| <chr> | <dbl> | <dbl> | <dbl> | <dbl> | <dbl> | <dbl> | <dbl> | <dbl> | <dbl> |
| k[8] | 2.835911169e-02 | 2.836630e-02 | 1.425470849e-03 | 1.421368620e-03 | 2.6021590e-02 | 3.0682905e-02 | 1.0012761453 | 20069.89659 | 11245.018485 |
| k[9] | 6.311753725e-02 | 6.311955e-02 | 6.631860514e-03 | 6.549163110e-03 | 5.2232080e-02 | 7.4079390e-02 | 1.0004276177 | 20559.36328 | 10676.668380 |
| k[10] | 2.504676307e-02 | 2.506050e-02 | 3.918722188e-03 | 3.928074570e-03 | 1.8647495e-02 | 3.1455835e-02 | 0.9999863273 | 18266.99472 | 10961.221986 |
| k[11] | 1.378870793e-01 | 1.379400e-01 | 1.296877797e-02 | 1.291789380e-02 | 1.1658875e-01 | 1.5897325e-01 | 1.0007669877 | 20232.30633 | 10838.768380 |
| k[12] | 7.359092871e-01 | 7.362185e-01 | 7.173896735e-02 | 7.086234960e-02 | 6.1757825e-01 | 8.5273205e-01 | 0.9999171591 | 19470.03863 | 11608.469269 |
| k_mean | 1.395115079e-01 | 1.294990e-01 | 5.466115487e-02 | 4.292720040e-02 | 7.3403190e-02 | 2.3724245e-01 | 1.0004835848 | 17020.63283 | 10543.583279 |
| k_disp | 1.204714654e+00 | 1.158810e+00 | 2.789271539e-01 | 2.464970760e-01 | 8.4389650e-01 | 1.7278730e+00 | 1.0004263125 | 19446.16247 | 11396.844688 |
| sigma | 4.612698179e-01 | 4.597545e-01 | 2.938572410e-02 | 2.901077550e-02 | 4.1588445e-01 | 5.1219215e-01 | 1.0003331852 | 14281.90482 | 11682.181004 |

Based on the above model summary we can make a few observations:

- Rhat values for all the variables are really low and close to 1. This means very good convergence.
- q5, q95, sd. The sd values tell us very low variation in the variables. This means we are more certain of the value. This is for all variables except mu0 for species 8 and 10.
- lp__ values are very negative. This indicates a good fit.
- Ess_bulk and ess_tail are really high between 10000-20000. This means the results are more reliable.
- Mad values are really low, showing a low spread in absence of outliers.

### 3.2 Comparing Results

We use the two given figures to compare the Bayesian model with the frequentist model. First we look at Figure 1, which shows relationship between mass and time for the seven pelagic invertebrate species for the frequentist model.

**Figure 1**

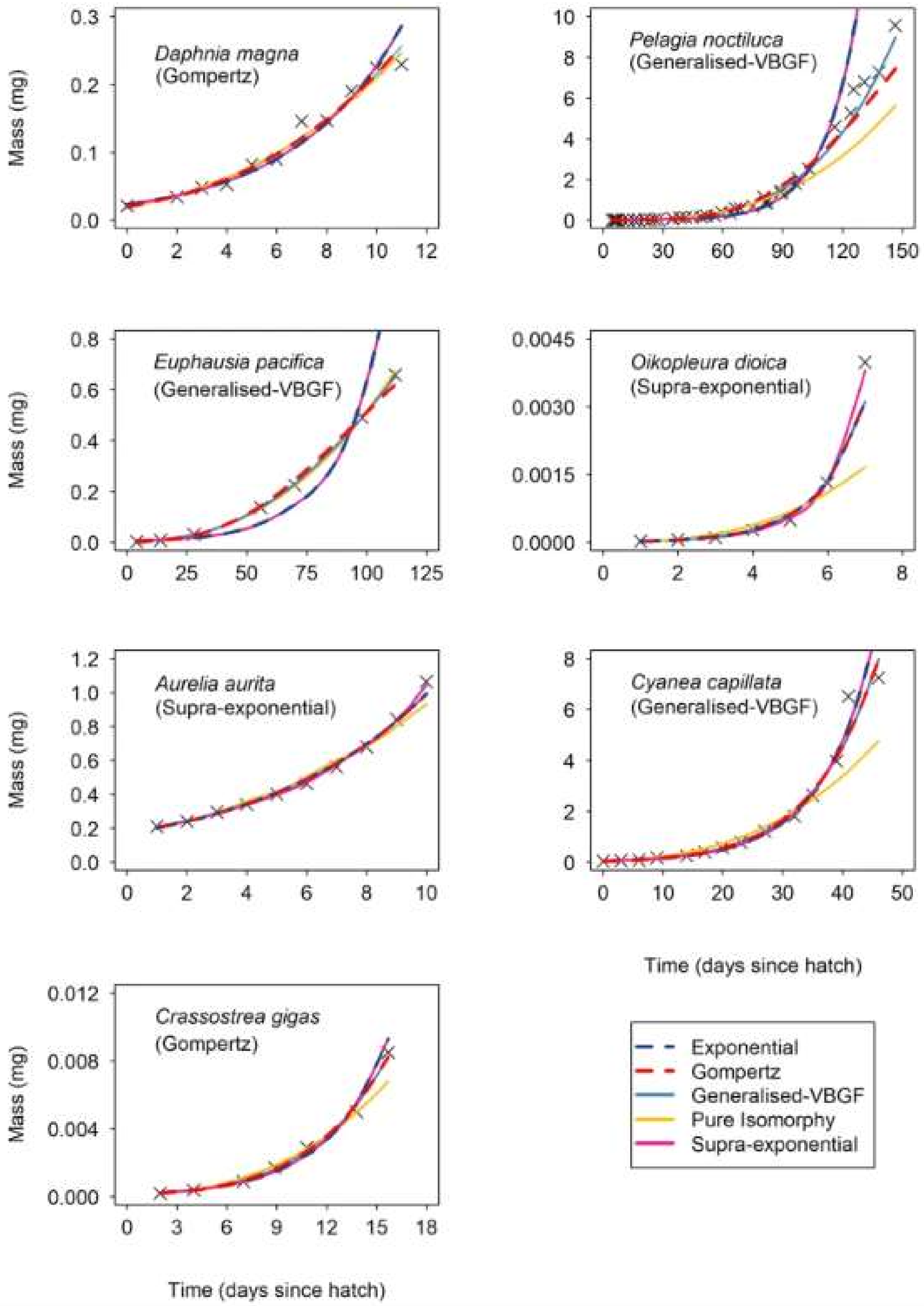
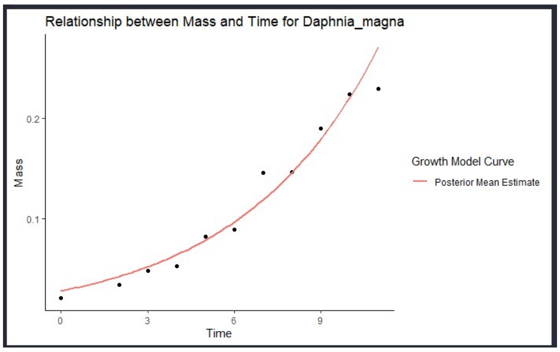


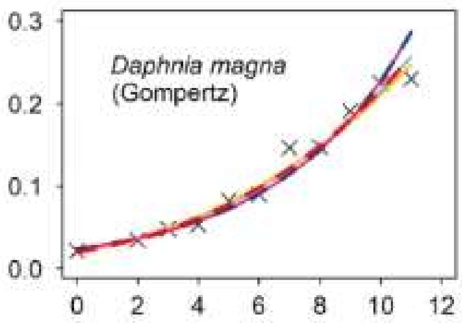
Figure obtained from [1, Results, Fig. 1]

Now we get the graphs for the same, for the seven pelagic invertebrate species for the Bayesian model we constructed. We can divide the above image based on the graphs for easier comparison. Comparing the exponential parameterization (Dotted blue line) with our Bayesian model we can conclude below:

Daphnia Magna

*[Generated from 5.8 after data input (5.1) and making the model (5.3, 5.4, 5.5, 5.6)]*

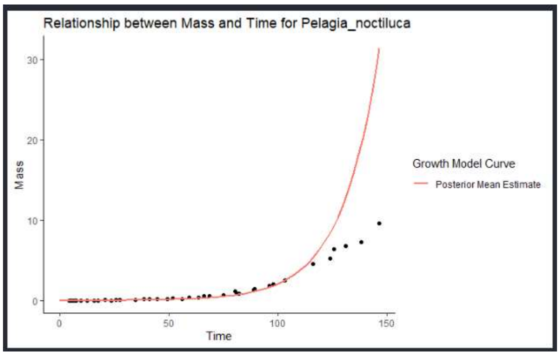


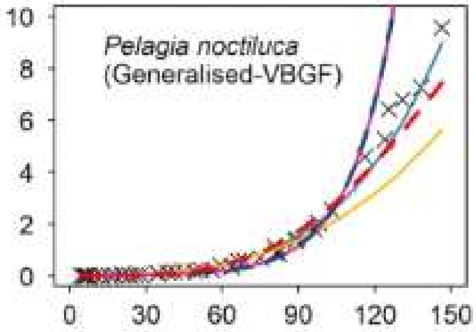


The bayesian model and frequentist model are comparable in fits. Both can be used.

Pelagia Noctiluca

*[Generated from 5.8 after data input (5.1) and making the model (5.3, 5.4, 5.5, 5.6)]*

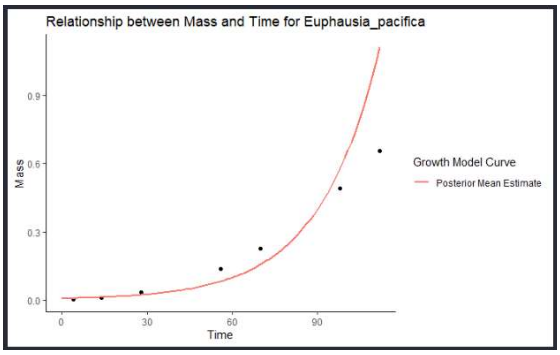


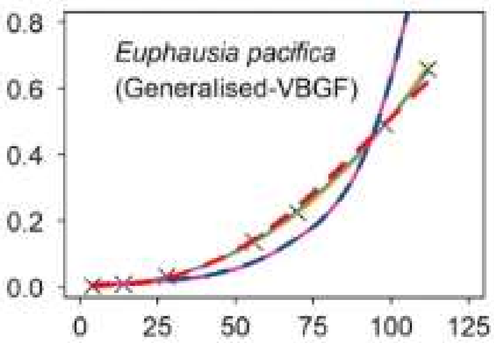


The frequentist model is better than the Bayesian model in terms of this species. The fit is better.

Euphausia Pacifica

*[Generated from 5.8 after data input (5.1) and making the model (5.3, 5.4, 5.5, 5.6)]*

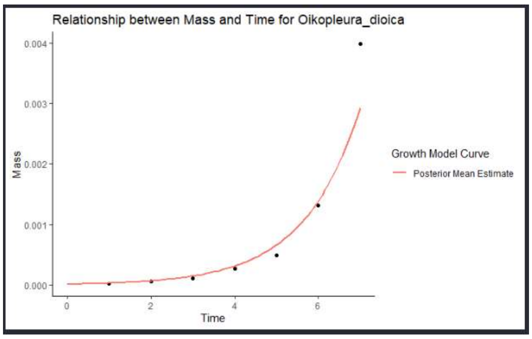


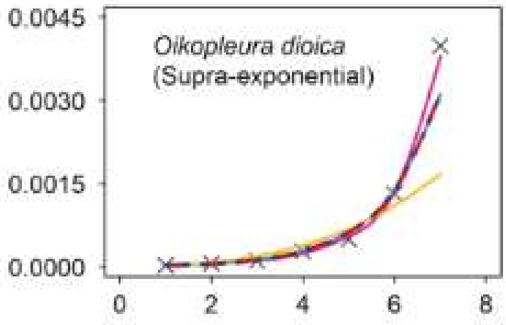


The bayesian model and frequentist model are comparable in fits. Both can be used.

Oikopleura Dioica

*[Generated from 5.8 after data input (5.1) and making the model (5.3, 5.4, 5.5, 5.6)]*

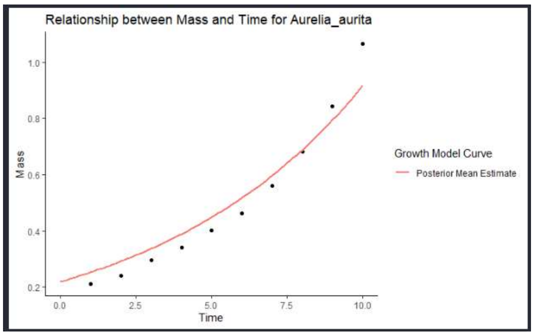


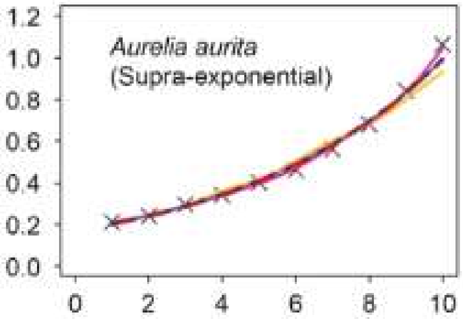


The bayesian model and frequentist model are comparable in fits. Both can be used.

Aurelia Aurita

*[Generated from 5.8 after data input (5.1) and making the model (5.3, 5.4, 5.5, 5.6)]*

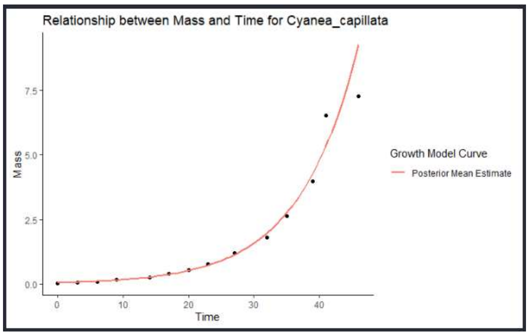


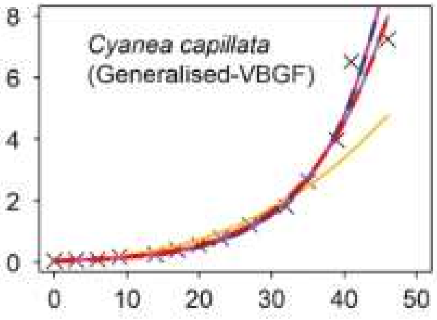


The frequentist model is better than the Bayesian model in terms of this species. The fit is better.

Cyanea Capillata

*[Generated from 5.8 after data input (5.1) and making the model (5.3, 5.4, 5.5, 5.6)]*

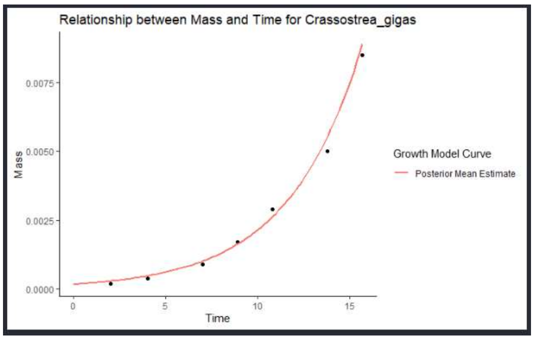


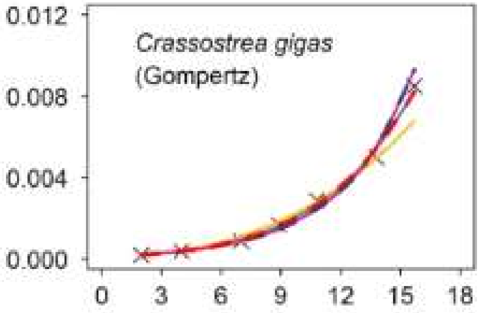


The bayesian model and frequentist model are comparable in fits. Both can be used.

Crassostrea Gigas

*[Generated from 5.8 after data input (5.1) and making the model (5.3, 5.4, 5.5, 5.6)]*





The bayesian model and frequentist model are comparable in fits. Both can be used.

Now we look at Figure 2, which shows relationship between mass and time for the five benthic invertebrate species for the frequentist model.

**Figure 2**

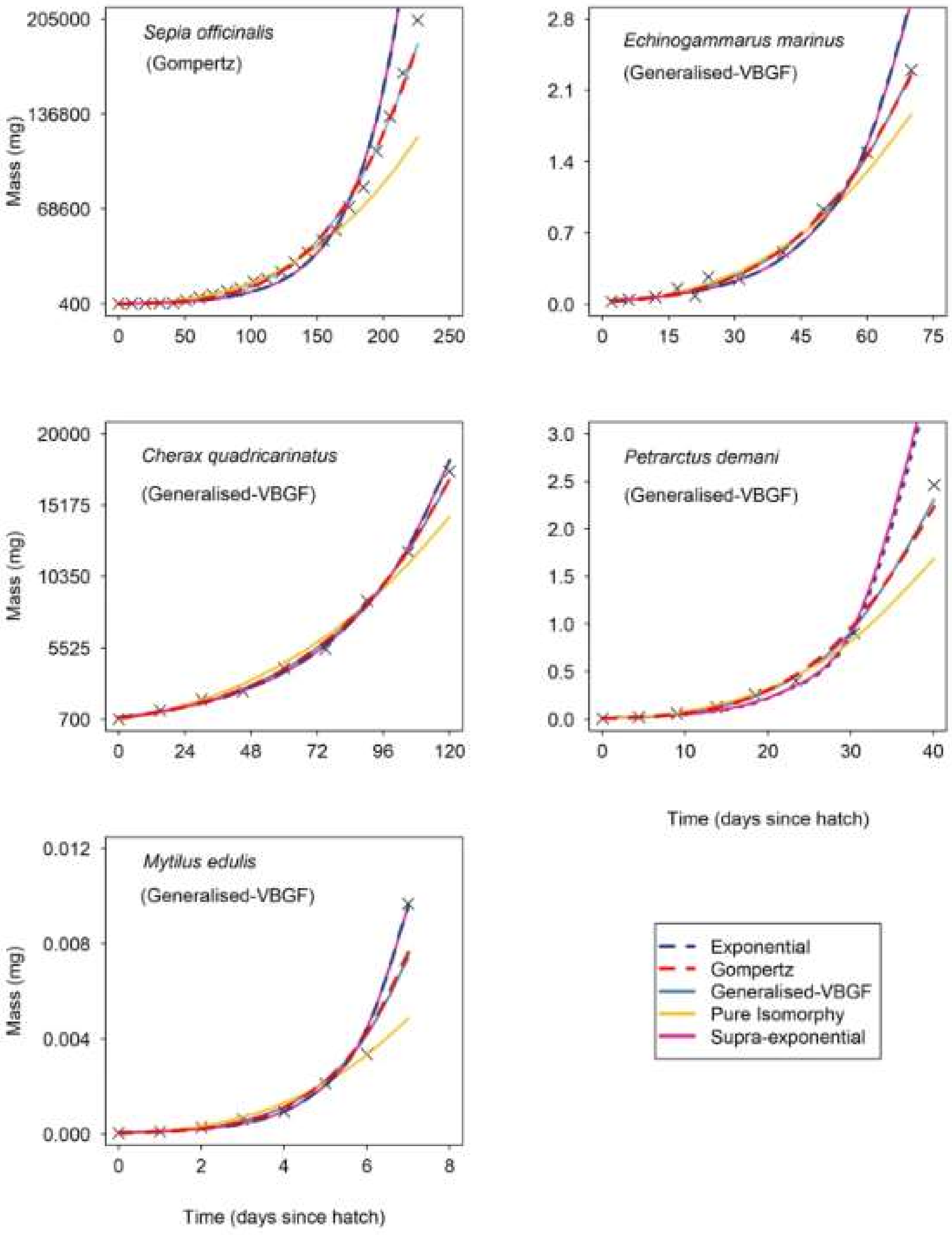
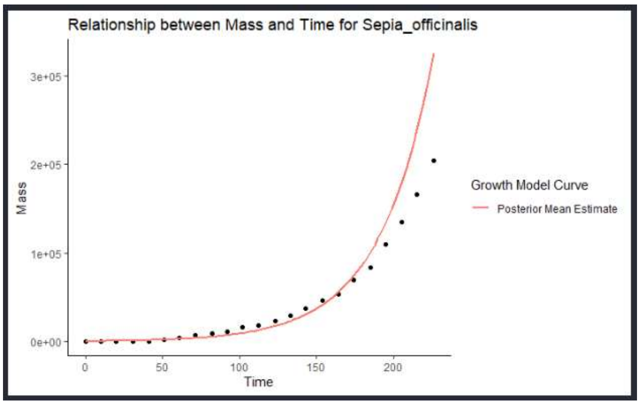


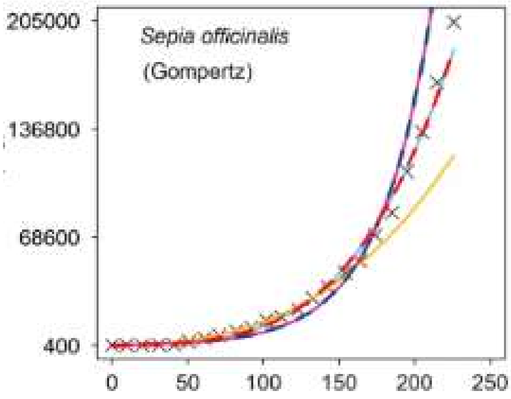
Figure obtained from [1, Results, Fig. 2]

Now we get the graphs for the same, for the seven pelagic invertebrate species for the Bayesian model we constructed. We can divide the above image based on the graphs for easier comparison. Comparing the exponential parameterization (Dotted blue line) with our Bayesian model we can conclude below:

Sepia Officinalis

*[Generated from 5.8 after data input (5.1) and making the model (5.3, 5.4, 5.5, 5.6)]*

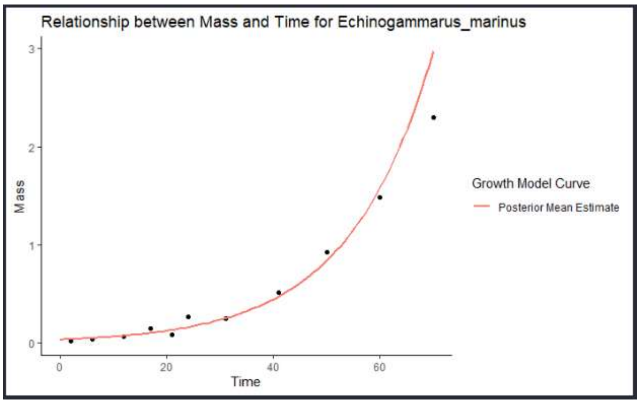


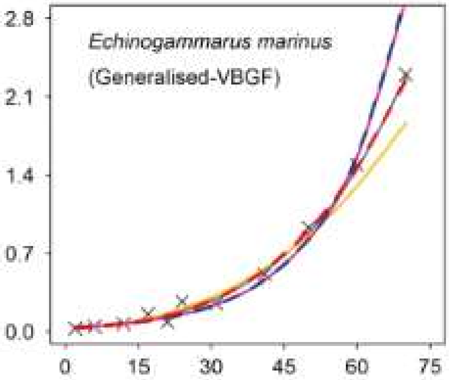


The bayesian model and frequentist model are comparable in fits. Both can be used.

Echinogammarus Marinus

*[Generated from 5.8 after data input (5.1) and making the model (5.3, 5.4, 5.5, 5.6)]*

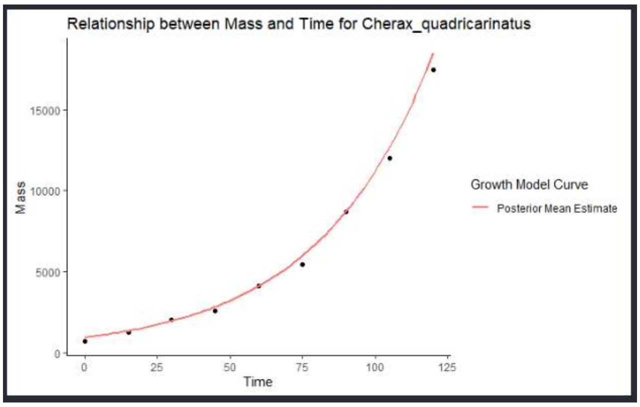


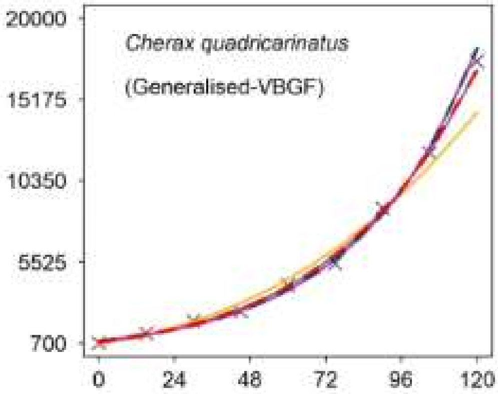


The bayesian model and frequentist model are comparable in fits. Both can be used.

Cherax Quadricarinatus

*[Generated from 5.8 after data input (5.1) and making the model (5.3, 5.4, 5.5, 5.6)]*

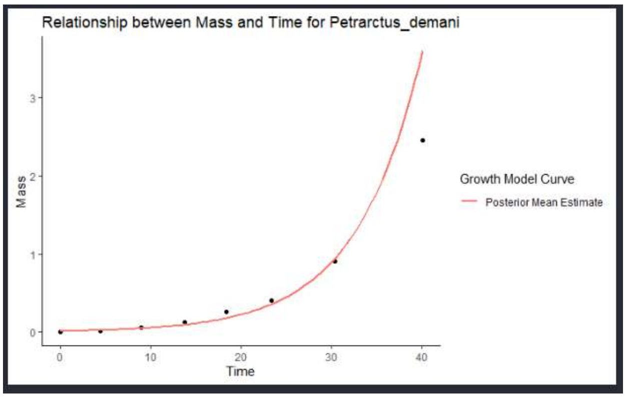


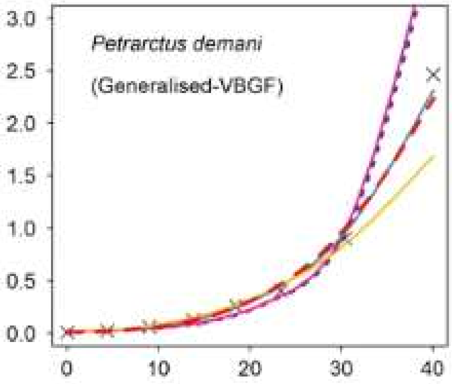


The bayesian model and frequentist model are comparable in fits. Both can be used.

Petrarctus Demani

*[Generated from 5.8 after data input (5.1) and making the model (5.3, 5.4, 5.5, 5.6)]*

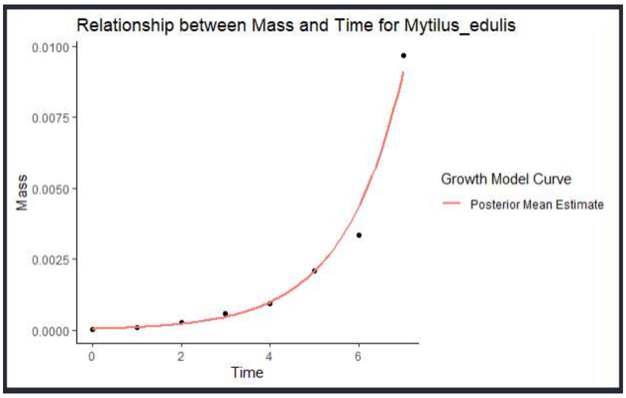


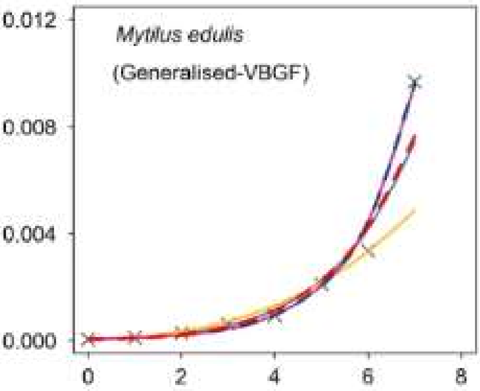


The bayesian model and frequentist model are comparable in fits. Both can be used.

Mytilus Edulis

*[Generated from 5.8 after data input (5.1) and making the model (5.3, 5.4, 5.5, 5.6)]*





The bayesian model and frequentist model are comparable in fits. Both can be used.

## 3.3 Plotting the variables

Based on the above comparisons we can make a few inferences:

- The frequentist model implemented in the study is better, if not equal to the Bayesian model.
- One of the reasons for this behavior is because the base model being used in the Bayesian model is the exponential parameterization we are using as the frequentist model.
- Both models more or less have the same parameter (k) that is being estimated, other than that all the other parameters are either derived from the data or end up the same.
- The k value we have is not that affected by its distribution that much, this means the differences between the k estimations for the frequentist and Bayesian model would not be that different.
- The k value being taken for the Bayesian model does not differ much between different species. This might be a reason why the Bayesian prediction falls behind the frequentist prediction sometimes.
- But looking at the models, the difference is so slight, that it's not very significant.

# 4. Additional Work

## 4.1 Exploring Additional Covariates

In the process of exploring additional covariates, it is important to consider the relevance and impact of each variable on the model's predictive performance. Here we have the other columns from the dataset:

- "Individual" : We can see that in all datapoints only one individual is considered for mass-at-age observations. This means individual cannot be used as a covariate as this does not vary much across datapoints.
- "Source" : This refers more to the source of the information and would not be very helpful as an additional covariate. Hence this cannot be used here.

## 4.2 Next Steps

As we do not have additional covariates in the dataset, we have to look back at the problem at hand and determine how we can make the model better. A few points we can look at are below:

- We have only considered one individual per mass-at-age data point [1, Data availability]. This would make the dataset a bit biased towards that individual. When we look at growth, we see that different individuals in the same species grow differently. Hence to get a better understanding of the species, having more individuals considered would make the data more generalized for the species.

- As the study concludes, a combination of these parameterizations can be used to provide a more flexible model that can account for mass scaling components better [1, Discussion].
- The data collected was only collected for the normal range in a natural setting [1, The data set].
- The study talks about how length data could be used if available [1, The data set].

Based on the points there are a few potential changes:

- Having more individuals for mass-at-age data points would help us get a more general model for the species. This data can be sourced from the same sources as the previous data. Another organism would help make our predictions more accurate.
- The current considered exponential parameterization considers the anabolism mass scaling component as 1, another model being used would make it better.
- Because the data collected was only done so in normal controlled circumstances, anomalies have the potential to derail the modeling. I think based on the need we can have a combination of both normal circumstances and nonideal circumstances to get a better prediction. While we will lose accuracy in modeling in normal circumstances, we will make the modeling more realistic. This data can be

collected not in controlled settings but in a wild setting and the data collected could be equally merged with controlled data in the ratio the wild circumstances happen, so we get a more realistic dataset.

- Length data can be sourced from fisheries where the original data is being sourced from. Because it is a controlled environment, it is more practical to have fewer species but more specific information like length in a controlled environment. These can be used to make a new model:

mass[i] ~ lognormal(log(mu), sigma)

mu = mu0 [species[i]] * exp((k1[species[i]] * (time[i])) + (k2[species[i]] * (length[i])))

k1 and k2 would be different priors with length being given in the dataset. This only impacts the exponent and hence would help get a more comprehensive prediction using the length.

# 5. Code

## 5.1 Importing the Dataset

***(Obtained from [1, Data Availability])***

~~~
```{r}
currentdata <- read.csv("growth_data.csv")
```
~~~

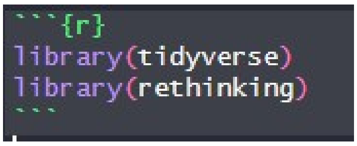
~~~
```{r}

currentdata <- read.csv("growth_data.csv")

```
~~~

## 5.2 Viewing the dataset

~~~
```{r}
currentdata
```
~~~

~~~
```{r}

currentdata

```
~~~

## 5.3 Loading libraries

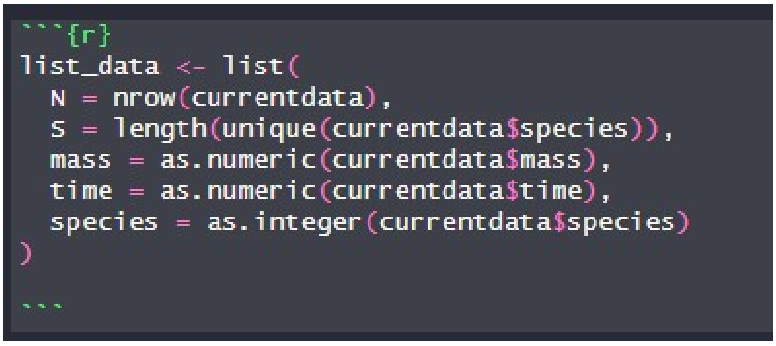


~~~
```{r}
library(tidyverse)
library(rethinking)
```
~~~

~~~
```{r}

library(tidyverse)

library(rethinking)
~~~

~~~
```
~~~

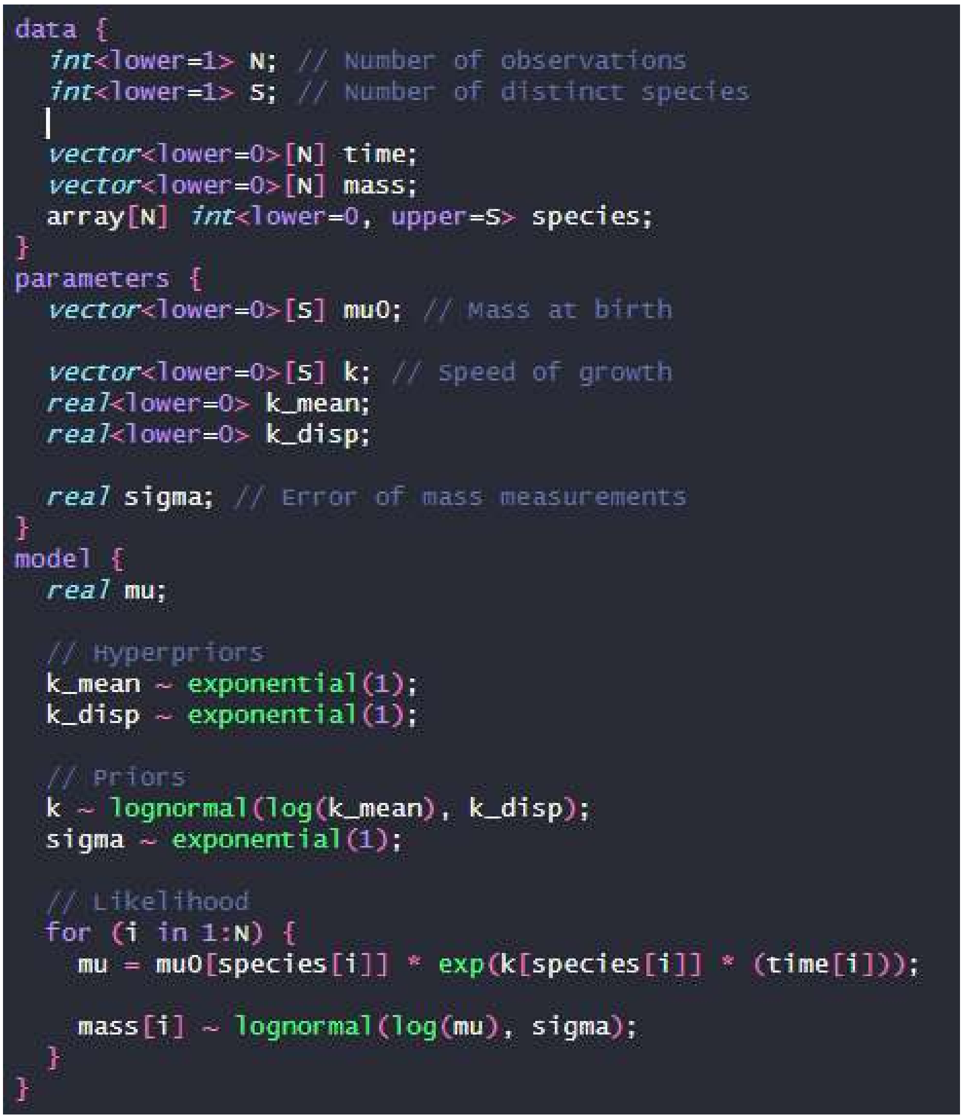
## 5.4 Converting the dataset into a usable format for the stan model



~~~
```{r}
list_data <- list(
  N = nrow(currentdata),
  S = length(unique(currentdata$species)),
  mass = as.numeric(currentdata$mass),
  time = as.numeric(currentdata$time),
  species = as.integer(currentdata$species)
)

```
~~~

~~~
```{r}
list_data <- list(
  N = nrow(currentdata),
  S = length(unique(currentdata$species)),
  mass = as.numeric(currentdata$mass),
  time = as.numeric(currentdata$time),
  species = as.integer(currentdata$species)
)

```
~~~

## 5.5 Stan model used

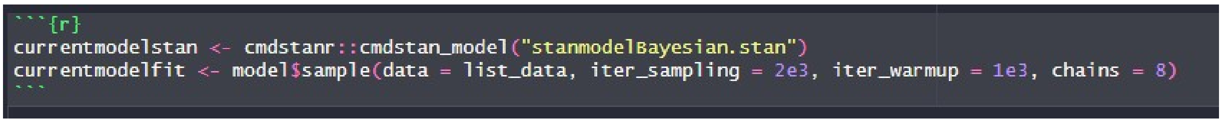


```
data {
  int<lower=1> N; // Number of observations
  int<lower=1> S; // Number of distinct species

  vector<lower=0>[N] time;
  vector<lower=0>[N] mass;
  array[N] int<lower=0, upper=S> species;
}
parameters {
  vector<lower=0>[S] mu0; // Mass at birth

  vector<lower=0>[S] k; // Speed of growth
  real<lower=0> k_mean;
  real<lower=0> k_disp;

  real sigma; // Error of mass measurements
}
model {
  real mu;

  // Hyperpriors
  k_mean ~ exponential(1);
  k_disp ~ exponential(1);

  // Priors
  k ~ lognormal(log(k_mean), k_disp);
  sigma ~ exponential(1);

  // Likelihood
  for (i in 1:N) {
    mu = mu0[species[i]] * exp(k[species[i]] * (time[i]));

    mass[i] ~ lognormal(log(mu), sigma);
  }
}
```

```
data {

  int<lower=1> N; // Number of observations

  int<lower=1> S; // Number of distinct species



  vector<lower=0>[N] time;

  vector<lower=0>[N] mass;
  array[N] int<lower=0, upper=S> species;
}
parameters {
  vector<lower=0>[S] mu0; // Mass at birth

  vector<lower=0>[S] k; // Speed of growth
  real<lower=0> k_mean;
  real<lower=0> k_disp;

  real sigma; // Error of mass measurements
}
model {
  real mu;

  // Hyperpriors
  k_mean ~ exponential(1);
  k_disp ~ exponential(1);

  // Priors
  k ~ lognormal(log(k_mean), k_disp);
  sigma ~ exponential(1);

  // Likelihood
  for (i in 1:N) {
    mu = mu0[species[i]] * exp(k[species[i]] * (time[i]));

    mass[i] ~ lognormal(log(mu), sigma);
  }
}
```

**NOTE: Saved in a different file.**

## 5.6 Calling the stan model, and fitting the model



~~~
```{r}
currentmodelstan <- cmdstanr::cmdstan_model("stanmodelBayesian.stan")
currentmodelfit <- model$sample(data = list_data, iter_sampling = 2e3, iter_warmup = 1e3, chains = 8)
```
~~~

~~~
```{r}
currentmodelstan <- cmdstanr::cmdstan_model("stanmodelBayesian.stan")
currentmodelfit <- model$sample(data = list_data, iter_sampling = 2e3, iter_warmup = 1e3, chains = 8)
~~~
```
```

## 5.7 Using summary() function to get summary statistics

~~~
```{r}
currentmodelfit$summary()
```
~~~

~~~
```{r}

currentmodelfit$summary()

```
~~~

## 5.8 Posterior mean estimate plots for all species

~~~
```{r}
currentmodeldraws <- currentmodelfit$draws(format = "df")


for (spec in unique(currentdata$species)) {
  plot_data <- currentdata |> filter(species == spec)

  mu0_mean <- currentmodeldraws[,grepl("mu0", colnames(currentmodeldraws))][, spec, drop = T] |> mean()
  k_mean <- currentmodeldraws[,grepl("k\\[", colnames(currentmodeldraws))][, spec, drop = T] |> mean()

  plot <- ggplot(plot_data, aes(time, mass)) +
    geom_point() +
    ggtitle(paste0("Relationship between Mass and Time for ", plot_data$binomial[1])) +
    xlab("Time") +
    ylab("Mass") +
    expand_limits(x = 0) +
    geom_function(fun = function(t) { mu0_mean * exp(k_mean * (t))},
                  mapping = aes(color = "Posterior Mean Estimate"), linewidth = 0.8) +
    guides(color=guide_legend(title="Growth Model Curve")) +
    theme_bw() +
    theme(panel.border = element_blank(), panel.grid.major = element_blank(),
          panel.grid.minor = element_blank(), axis.line = element_line(colour = "black"))

  print(plot)

  ggsave(filename = paste0(plot_data$binomial[1], ".png"), device = "png")
}
```
~~~

~~~
```{r}

currentmodeldraws <- currentmodelfit$draws(format = "df")
~~~

```
for (spec in unique(currentdata$species)) {
  plot_data <- currentdata |> filter(species == spec)

  mu0_mean <- currentmodeldraws[,grepl("mu0",
colnames(currentmodeldraws))][, spec, drop = T] |> mean()
  k_mean <- currentmodeldraws[,grepl("k\\[",
colnames(currentmodeldraws))][, spec, drop = T] |> mean()

  plot <- ggplot(plot_data, aes(time, mass)) +
    geom_point() +
    ggtitle(paste0("Relationship between Mass and Time for ",
plot_data$binomial[1])) +
    xlab("Time") +
    ylab("Mass") +
    expand_limits(x = 0) +
    geom_function(fun = function(t) { mu0_mean * exp(k_mean
* (t))},
                  mapping = aes(color = "Posterior Mean Estimate"),
linewidth = 0.8) +
    guides(color=guide_legend(title="Growth Model Curve")) +
    theme_bw() +
```

~~~
    theme(panel.border = element_blank(), panel.grid.major = element_blank(),
          panel.grid.minor = element_blank(), axis.line = element_line(colour = "black"))

  print(plot)

  ggsave(filename = paste0(plot_data$binomial[1], ".png"), device = "png")
}
```
~~~

# 6. References

[1] Lee, L., Atkinson, D., Hirst, A.G. et al. A new framework for growth curve fitting based on the von Bertalanffy Growth Function. Sci Rep 10, 7953 (2020). [Online]. Available: https://doi.org/10.1038/s41598-020-64839-y