Survival analysis in an experimental microbial aquatic community

11 April 2021

A single species extinction can cause a chain of secondary extinctions leading to a trophic cascade in a food web (Dunne and Williams 2009). Hence, it is important to improve our predictive understanding of the causes and consequences of species loss in an ecosystem. Knowing the reason behind species extinction can help us improve conservation efforts.

The order in which species extinct in a food web influences the consequences of the extinction event. For example, loss of larger species first, followed by smaller species, can cause a more rapid loss of ecosystem functioning in marine sediments than a random extinction order, according to simulation-based predictions (Solan 2004). Therefore, other than knowing which species might suffer extinction, it is also crucial to know when the species will go extinct.

Species in a community are linked by a complex network of interactions such as competition, predation and mutualisms. This can result in the extinction of one species caused by the loss of other species (Paine 1966). Species extinction can also be influenced by abiotic environmental factors (Cardillo 2005).

In this project, we would like to study the effect of environmental factors and interspecific interactions on the extinction rates of species in an experimental microbial aquatic community. In the experiment, four different treatment combinations had been used, varying temperature and nutrient concentration in the microcosms. We would like to investigate the combined effect of interspecific interactions and environmental factors on the determinants of extinction events.

We will use survival analysis with a Bayesian approach to estimate the mean extinction times of the species. Possible research questions of this project are:

- Using informative priors derived from empirical studies

- Using a different survival model than the current one in the preliminary study below

1 Aquatic food web

The aquatic food web consists of a 17 taxa of aquatic eukaryotic microorganisms, unknown heterotrophic nanoflagellates and an unknown bacterial flora (Fig. 1). The species in the food web (Fig. 1) are: 3 *Blepharisma japonicum*, 4 *Chilomonas paramecium*, 5 *Colpidium striatum*, 6 *Colpoda cucculus*, 7 *Cyclidium glaucoma*, 8 *Didinium nasutum*, 9 *Dileptus anser*, 10 *Entosiphon*, 11 *Euplotes patella*, 12 *Loxocephallus*, 13 *Lepadella*, 14 *Dicanophoridae*, 15 *Paramecium bursaria*, 16 *Paramecium caudatum*, 17 *Telotrochiudium*, 18 *Tetrahymena piriformis* and 19 *Vorticella*.

2 Dataset

The experiment for the current dataset was conducted by Nicholas Worsfold at the University of Sheffield (Worsfold 2007). The data is from a highly replicated microcosm experiment where the extinction times of freshwater protists species forming a small food web was recorded. Four different treatment combinations have been used, varying temperature and nutrient concentration in the microcosms.

Data from the experiment is the presence or absence of each species for each of the eight weeks of the experiment, for each of the 200 replicates.
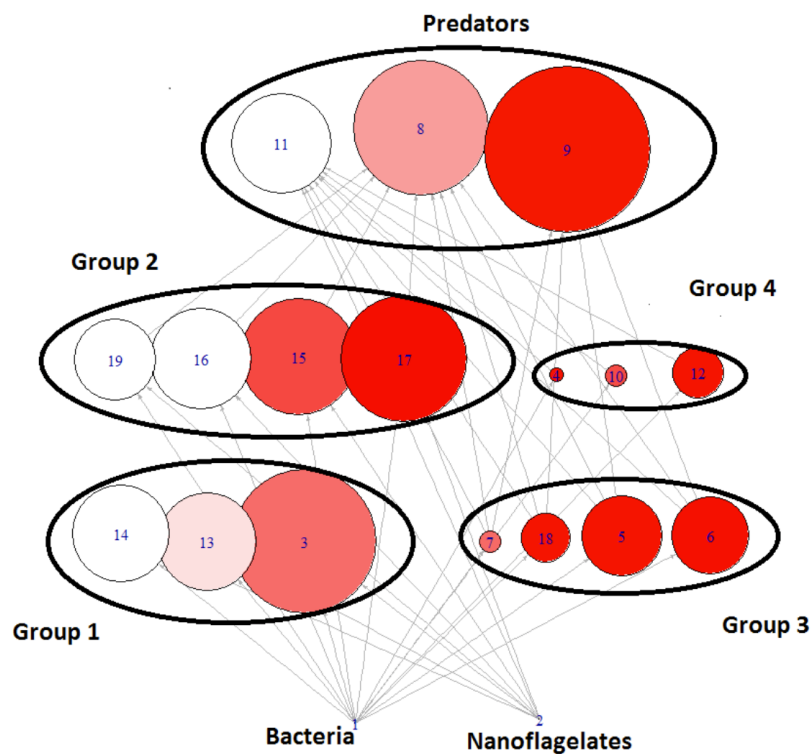
Figure 1: Food web: species are grouped according to their trophic group. The size of each node is proportional to the logarithm of the body size of the species and color indicates the mean extinction time of the species computed by maximizing likelihood across all treatments, with a darker shade of red corresponding to lower extinction times (Palamara 2015).

3 Preliminary study using a subset of the dataset

Below, we subset the dataset for the treatment with low temperature and low nutrient concentration for the species *chilomonas*. We fit the survival model with both the Bayesian approach and maximum likelihood approach to compute the mean extinction times and compare them.

```r
## Loading the packages
if ( ! require(runjags) )   { install.packages("runjags");   library(runjags) }
if ( ! require(rjags) )    { install.packages("rjags");    library(rjags) }
if ( ! require(bbmle) )    { install.packages("bbmle");    library(bbmle) }
if ( ! require(ggpubr) )   { install.packages("ggpubr");   library(ggpubr) }


## Importing the dataset
## Change the path accordingly
dd <- readRDS("main_data.Rdata")


## Some summary of the dataset
all.species <- sort(unique(dd$species))
all.temp <- sort(unique(dd$temp))
all.energy <- sort(unique(dd$energy))


## Subsetting the dataset: Low temperature; low energy and species "chilomonas"
ddll <- subset(dd, dd$temp.var==0)
ddll <- subset(ddll, ddll$energy.var==0)
ss <- 2 #corresponds to species chilomonas
spp.dd <- subset(ddll, ddll$species==all.species[ss])


## Let's look at the subsetted dataset spp.dd
head(spp.dd)
```

```
##      jar temp energy    species week.persist dd.fm temp.var energy.var day1 day2
## 201   1   15  0.275 chilomonas          0.5   0.5        0          0    0    1
## 209   9   15  0.275 chilomonas          0.5   0.5        0          0    0    1
## 215  15   15  0.275 chilomonas          0.5   0.5        0          0    0    1
## 216  16   15  0.275 chilomonas          0.5   0.5        0          0    0    1
## 226  26   15  0.275 chilomonas          0.5   0.5        0          0    0    1
## 228  28   15  0.275 chilomonas          0.5   0.5        0          0    0    1
```

```r
## day2 is the week in which the species was first not observed i.e. the species extinct by that week
day2 <- spp.dd$day2
```

3.1 Bayesian method

```r
################################# Using the Bayesian method #################################

## Define the model in JAGS:
## Considering uniform priors for the parameters

file.jags.model <- "main.jags"

cat("model {\n",
    "  shape ~ dunif(0, max_shape)\n",
    "  lambda ~ dunif(0, max_lambda)\n",
```

```
"  for ( i in 1:n ) {\n",
"    day2[i] ~ dweib(shape, lambda)\n",
"  }\n",
"}\n",
sep="",
file=file.jags.model)


## Run JAGS:
max_shape <- 100
max_lambda <- 100
jags.obj <- jags.model(file.jags.model,
                  data=list(max_shape = max_shape,
                            max_lambda = max_lambda,
                            n         = 50,
                            day2      = day2))
```
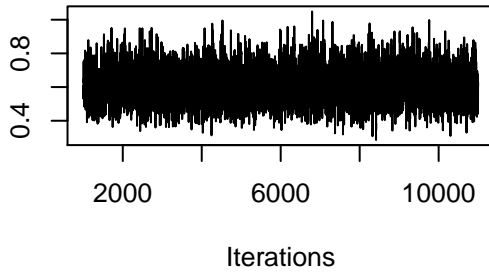
```
## Compiling model graph
##    Resolving undeclared variables
##    Allocating nodes
## Graph information:
##    Observed stochastic nodes: 50
##    Unobserved stochastic nodes: 2
##    Total graph size: 56
##
## Initializing model
```
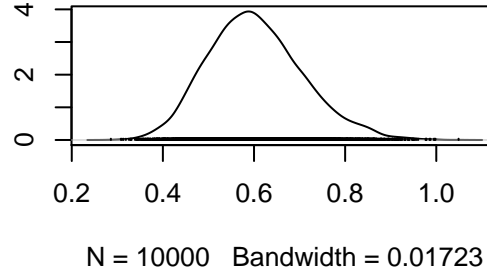
```
sampsize <- 10000
jags.res <- coda.samples(jags.obj, c("shape", "lambda"), sampsize)
plot(jags.res)
```
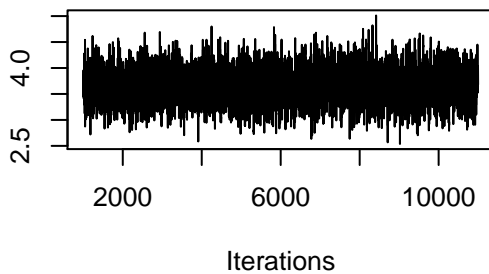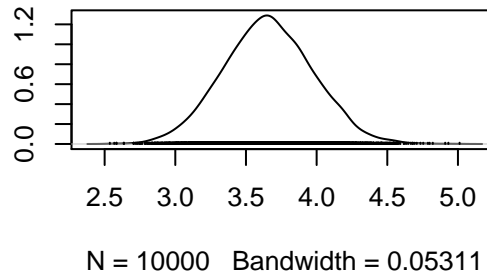
Trace of lambda	Density of lambda

Trace of shape	Density of shape


```r
## Constructing a data frame with the samples from the posterior distributions
chains <- as.matrix(jags.res)
post_par <- data.frame(lambda = chains[,1],
                       shape = chains[,2], Distribution = "Posterior")
post_par <- rbind(post_par,
                  data.frame(lambda = runif(n = 10000, 0, max_lambda),
                             shape = runif(n = 10000, 0, max_shape), Distribution = "Prior"))
```

3.2 Maximum Likelihood method

```r
################################## Using the Maximum Likelihood method ##########################

# Defining the weibull negative log likelihood here

Weibull.2x2.LF_mod <- function(lscale.0, lscale.t, lscale.e, lscale.i,
                               lshape.0, lshape.t, lshape.e, lshape.i,
                               day1, day2, temp.var, energy.var, spp.id) {
    scale <- exp(lscale.0 +
                     lscale.t*temp.var +
                     lscale.e*energy.var +
                     lscale.i*temp.var*energy.var)
    shape <- exp(lshape.0 +
                     lshape.t*temp.var +
                     lshape.e*energy.var +
                     lshape.i*temp.var*energy.var)
    -sum(log(dweibull(day2, shape, scale)))
}
```

```r
totmeansurv <- mean(spp.dd$week.persist+0.5)
startvals.SE <- list(lscale.0=log(totmeansurv), lscale.t=0, lscale.e=0, lscale.i=0,
                     lshape.0=0, lshape.t=0, lshape.e=0, lshape.i=0)

## Running the Maximum Likelihood method
m0 <- mle2(Weibull.2x2.LF_mod, start=startvals.SE, data=spp.dd,
           fixed=list(lscale.t=0,lscale.e=0,lscale.i=0,lshape.t=0,lshape.e=0,lshape.i=0))
lscale0.0<-coef(m0)[1];lscale0.t<-coef(m0)[2];lscale0.e<-coef(m0)[3];lscale0.i<-coef(m0)[4]
lshape0.0<-coef(m0)[5];lshape0.t<-coef(m0)[6];lshape0.e<-coef(m0)[7];lshape0.i<-coef(m0)[8]

scale <- exp(lscale0.0 + lscale0.t + lscale0.e + lscale0.i)
shape <- exp(lshape0.0 + lshape0.t + lshape0.e + lshape0.i)
lambda <- (1/scale)^shape

print(paste0("Estimated shape = ", shape))
```

```
## [1] "Estimated shape = 3.68740341824993"
```

```r
print(paste0("Estimated lambda = ", lambda))
```

```
## [1] "Estimated lambda = 0.583775676203063"
```

3.3   Comparing Bayesian and Maximum likelihood method

```r
## Plotting the distribution of the parameters along with the point estimates from the MLE method

x_dummy <- data.frame(x = -10, y = 1, Method = "MLE method")
cols <- c("MLE method" = "black")

plot_lambda <- ggplot(post_par) +
    geom_density(aes(x = lambda, y = ..density.., fill = Distribution), alpha = 0.5) +
    theme_classic() +
    xlim(c(0,3)) +
    geom_vline(xintercept = lambda) +
    xlab("Lambda") +
    ylab("Density") +
    geom_vline(data = x_dummy, aes(xintercept = x, color = Method)) +
    scale_color_manual(values = cols)

plot_shape <- ggplot(post_par) +
    geom_density(aes(x = shape, y = ..density.., fill = Distribution), alpha = 0.5) +
    theme_classic() +
    xlim(c(1,8)) +
    geom_vline(xintercept = shape) +
    xlab("Shape") +
    ylab("Density") +
    geom_vline(data = x_dummy, aes(xintercept = x, color = Method)) +
    scale_color_manual(values = cols)

ggarrange(plot_lambda, plot_shape, nrow = 2, ncol = 1, common.legend = TRUE, legend = "bottom")
```
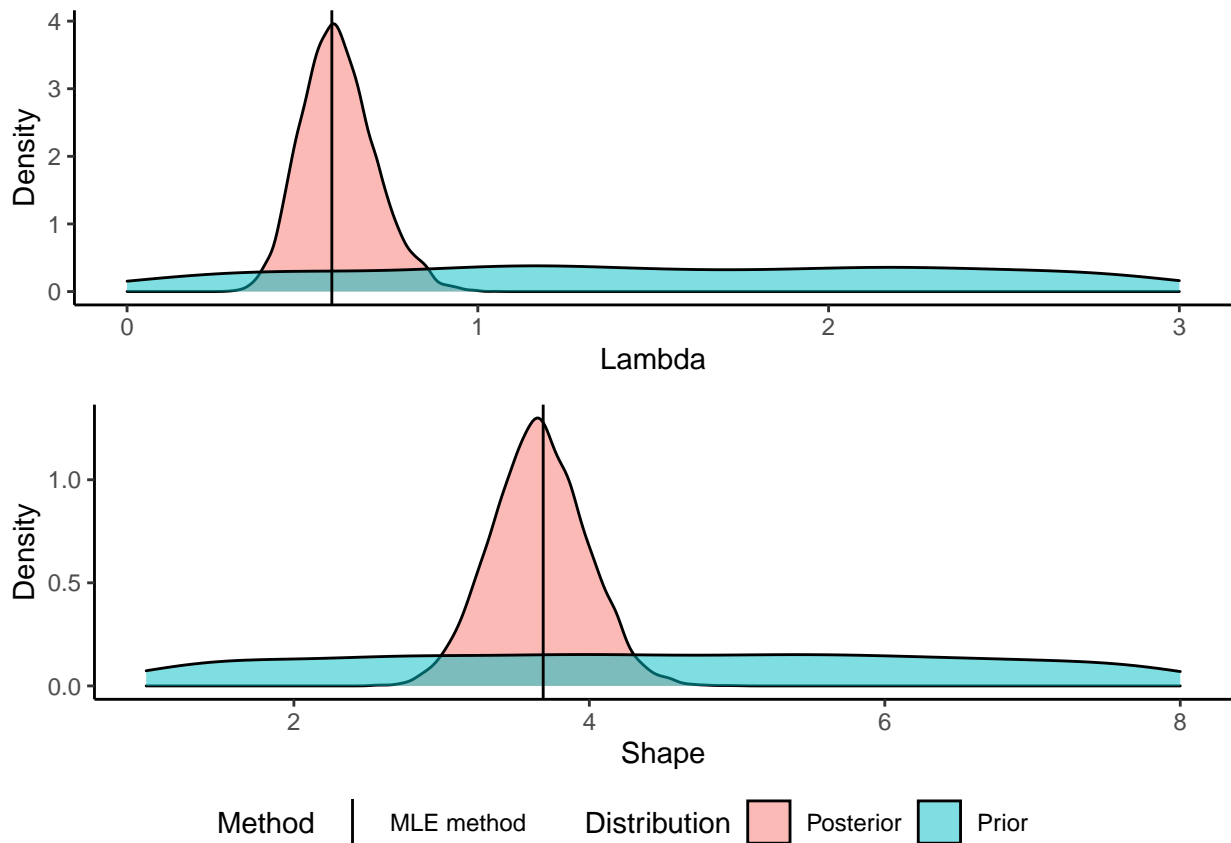
References

Cardillo, M. 2005. "Multiple Causes of High Extinction Risk in Large Mammal Species." *Science* 309 (5738): 1239–41. https://doi.org/10.1126/science.1116030.

Dunne, Jennifer A., and Richard J. Williams. 2009. "Cascading Extinctions and Community Collapse in Model Food Webs." *Philosophical Transactions of the Royal Society B: Biological Sciences* 364 (1524): 1711–23. https://doi.org/10.1098/rstb.2008.0219.

Paine, Robert T. 1966. "Food Web Complexity and Species Diversity." *The American Naturalist* 100 (910): 65–75. https://doi.org/10.1086/282400.

Palamara, Gian Marco. 2015. "Theoretical and Empirical Studies on Population Dynamics, Species Interactions and Extinctions." University of Zurich. https://doi.org/10.5167/UZH-129496.

Solan, M. 2004. "Extinction and Ecosystem Function in the Marine Benthos." *Science* 306 (5699): 1177–80. https://doi.org/10.1126/science.1103960.

Worsfold, Nicholas. 2007. "The Consequences of Extinction in Experimental Aquatic Communities." Ph.D., University of Sheffield. https://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.489745.