Code for producing results and figures of “Estimating sperm whale cue rates to inform passive acoustic density estimation via cue counting”

Marques, T. A.

25 October 2024

Table of Contents

# Introduction

This document includes the code required to reproduce the analysis and resulting figures for the manuscript “Estimating sperm whale cue rates to inform passive acoustic density estimation via cue counting”, submitted to The Journal of the Acoustical Society of America, by **add all names here**.

In the paper we report on estimates of cue rates for sperm whales at a selection of time and year combinations that were sampled, but the main interest is in estimating a mean cue rate, and corresponding precision, for a time and location that might not have been observed. We also look at how a number of additional covariates can impact cue rates.

This is hosted in a dedicated github repository at

<https://github.com/TiagoAMarques/SpermWhaleCueRates/>

# Related repositories and files

## Related repositories

This document was created by simplification of another document describing lots of dead ends and open questions related to Gamma GLMMs and their implementation both in lme4::glmer and in a Bayesian context.

These analysis and the data to make it fully reproducible are hosted at

<https://github.com/TiagoAMarques/Report4BB>

in case that helps the readers.

Nonetheless, the document presented here is self contained.

# Related Files

For this code to run the file clall.txt is required, but that file is over 400Gb. I will archive it as a file somewhere but in the mean time you can download it from here.

“<https://www.dropbox.com/scl/fi/eihokxcimaz15y83c4v37/clall.txt?rlkey=uyz2l2l8zrzbx6b73dnio3mji&dl=0>”

# About the data used for analysis

The main dataset lies within ddata1, the single object in data\_4\_article\_clickrates\_deep\_dive.rda that gets front loaded below. Under

<https://github.com/TiagoAMarques/Report4BB>

there are additional details about how this object is created, but for the purpose of this paper, this R data file should be treated as data.

The data in ddata1 consists of summaries of numbers of regular echolocation clicks per deep dive cycle, for each of the sperm whale tags considered on the manuscript. Note that despite having been recorded at the deep dive cycle level, a level which was also considered for comparisons in the manuscript

Marques, T. A., Marques, C. S. & and Gkikopoulou, K. C. (2023) A sperm whale cautionary tale about estimating acoustic cue rates for deep divers. : 1577-1584

the first step in the data pre-processing below is to pool data for each tag record, as the tag is the fundamental (and more importantly independent) sampling unit we consider here.

# Reading the data

We begin by reading the deep dive cycle data in:

# file created in Cue\_Rates\_For\_Sperm\_Whales.Rmd  
# Reading the data that contain the information per deep dive cycle - object ddata1  
load("data/data\_4\_article\_clickrates\_deep\_dive.rda")  
ddata1$location[ddata1$location=="Gulf of Mexico"]<-"GoM"  
ddata1$location[ddata1$location=="North Atlantic Delaware"]<-"Delaware"  
ddata1$location[ddata1$location=="DOMINICA"]<-"Dominica"  
ddata1$location[ddata1$location=="Norway Andenes"]<-"Norway"

# Data pre-processing

We actually have data from whales which have been subjected to SONAR exposure under controlled exposure experiments. We discard the data from those tags here, since evaluating the effect of SONAR is the focus of other research programs. Understanding the effects of sonar exposure on whale behaviour, and in particular cue rate production, is a separate research thread which requires information to be analysed at a much finer resolution, and for which context would have to be considered. For the majority of the tags the context information is unavailable to us within the ACCURATE project.

#removing the tags for animals we know were exposed to sonar  
DDCs<-ddata1[ddata1$sonar!="sonar",]

Since we will treat tags, not deep dive cycles, as the independent sampling units, we aggregate the deep dive cycle data for each tag into a single tag record (object tags)

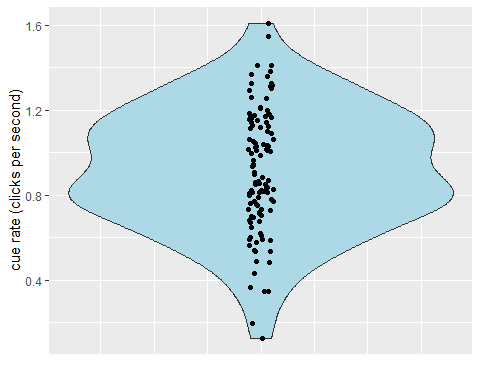
# Creating the data per tag  
tags <- DDCs %>%  
 group\_by(tag) %>%   
 summarise(location=unique(location), year=unique(year), sex=unique(sex),  
 duration= sum(durations,na.rm=T),nclicks=sum(nclick,na.rm=T),  
 crate=sum(nclick,na.rm=T)/sum(durations,na.rm=T),ddc=max(absdives+1,na.rm = T))  
#recoding "Norway Andenes" as just Norway  
tags$location[tags$location=="Norway Andenes"] <- "Norway"

# Exploratory data analysis

We have a total of 118 whales tagged which were not, knowingly, exposed to sonar.

Tag recording duration ranged from 0.144 to 25.781 hours. The observed cue rates per tag varied between 0.126 and 1.61 clicks per second, with a mean value of 0.921 and median value of 0.904 clicks per second. The individual cue rates per tag record, pooled across years and locations, are shown in the following figure:

ggplot(tags,aes(x=1,y=crate),fill="lightblue")+  
theme(axis.ticks.x=element\_blank(),axis.text.x=element\_blank())+geom\_violin(fill="lightblue")+  
geom\_jitter(width=0.03)+ylab("cue rate (clicks per second)")+xlab("")



This corresponds to Figure 1 in the main paper

Below we present a table with the locations and years covered by these tags:

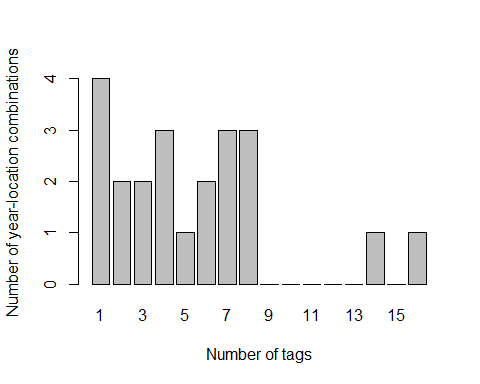
kable(table(tags$location,tags$year))

|  | 2001 | 2002 | 2003 | 2005 | 2009 | 2010 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 | 2020 | 2021 |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Azores | 0 | 0 | 0 | 0 | 0 | 8 | 0 | 0 | 0 | 0 | 0 | 3 | 5 | 6 | 8 |
| DOMINICA | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 7 | 16 | 4 | 1 | 4 | 0 | 0 | 0 |
| Gulf of Mexico | 4 | 14 | 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Kaikoura | 0 | 0 | 0 | 0 | 0 | 0 | 6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Mediterranean | 1 | 0 | 7 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| North Atlantic Delaware | 0 | 0 | 7 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Norway | 0 | 0 | 0 | 1 | 1 | 3 | 0 | 0 | 0 | 2 | 0 | 0 | 2 | 0 | 0 |

The number of tags per year-location combination varies considerably. Out of a total of 7 different locations and 15 different years, leading therefore to 105 possible year-location combinations, tags are not available for the majority of these possible combinations (83 combinations), with only 22 combinations having any tags associated with.

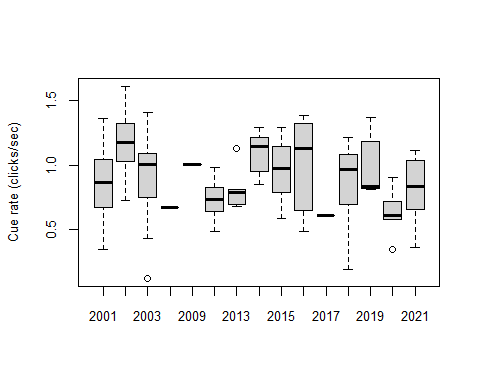
What might be the number of tag records required to obtain a reliable year-location cue rate estimate remains hard to evaluate, but several year-location combinations are certainly below that minimum, namely for those with less than a handful of tags. The distribution of the number of tags per year-location combination for which we have tags is represented in the image below:

#note the need to remove the first count  
#corresponding to un-interesting year-location combinations  
#with 0 tags  
counts<-table(tags$location,tags$year)  
maxtags<-max(table(tags$location,tags$year))  
counts2<-numeric(maxtags)  
for (i in 1:maxtags){  
 counts2[i]<-sum(counts==i)  
}  
barplot(counts2,names.arg=1:maxtags,ylab="Number of year-location combinations",xlab="Number of tags")



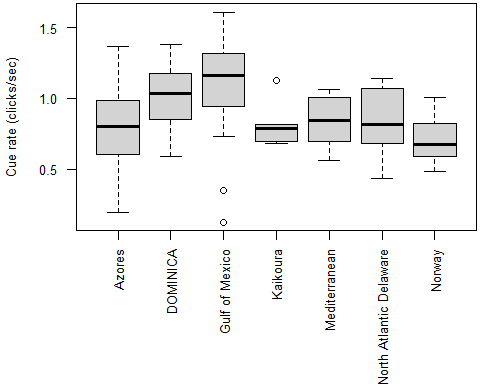
We can take a look at the cue rates (pooled across locations) per year

with(tags,boxplot(crate~year,ylab="Cue rate (clicks/sec)",xlab="",cex.lab=0.8,cex.axis=0.8))



and those (pooled across years) per location

par(mar=c(8,4,0.2,0.2))  
with(tags,boxplot(crate~location,las=2,ylab="Cue rate (clicks/sec)",xlab="",cex.lab=0.8,cex.axis=0.8))



We are hoping to explain variability in cue rates as a function of year and location. To do so, year-location combinations with a small number of tags are difficult to deal with. In particular, for those with a single tag, a year-location effect would be hard to estimate: with a model with an interaction, the effect would be strictly unidentifiable. We therefore removed years and locations for which only a single tag existed, effectively removing the single tag from Norway Andenes, which was also the only tag from 2005. Additionally, we pooled tags from single-tag year-location combinations into an adjacent year for the same same location. Hence, we pooled

* the single tag from Norway in 2009 with the 3 Norway tags from 2010;
* the single tag from Dominica in 2017 with the remaining 4 Dominica tags from 2016;
* the single tag from the Mediterranean in 2001 with the 7 tags from Mediterranean from 2003.

# removing the location with a single tag  
# tags<-tags[tags$location!="Norway Andenes",]  
# grouping single tag per year into adjacent years  
tags$year[tags$location=="Norway" & tags$year==2009]<- 2010  
tags$year[tags$location=="Dominica" & tags$year==2017]<- 2016  
tags$year[tags$location=="Mediterranean" & tags$year==2001]<- 2003

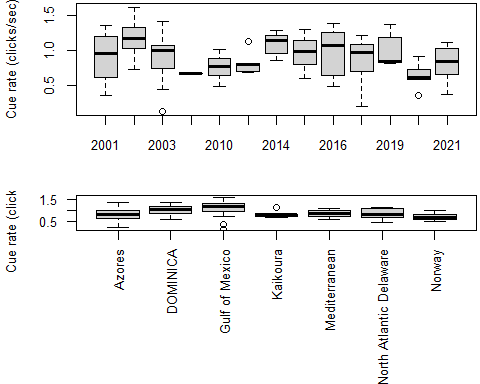
After this assignment, we created a couple of factor covariates inside tags:

* fyear, representing year as a factor
* locyear, representing each year-location as a factor

#building relevant factor covariates given this new assignment  
#making a new variable, year as factor  
tags$fyear<-as.factor(tags$year)  
#adding a variable that represents each location-year combination  
tags$locyear<-as.factor(paste0(tags$location,tags$fyear))  
#and also explicitly code location as a factor  
tags$location<-factor(tags$location)

Given this data pooling, the above plots become:

par(mfrow=c(2,1),mar=c(4,4,0.2,0.2))  
with(tags,boxplot(crate~year,ylab="Cue rate (clicks/sec)",xlab="",cex.lab=0.8,cex.axis=0.8))  
par(mar=c(8,4,0.2,0.2))  
with(tags,boxplot(crate~location,las=2,ylab="Cue rate (clicks/sec)",xlab="",cex.lab=0.8,cex.axis=0.8))



# Estimating cue rates

Our objective is to obtain a cue rate estimate, and its desired precision to include in a cue counting density estimator. We assume cue rate might depend on a number of covariates, for which location and year might act as relevant proxies.

After this we will look at whether some animal specific covariates might be relevant to explain cue rates.

## Considering Year and Location

Ignoring what required predictions might be to begin with, from first principles, location might be a sensible fixed effects covariate, since different locations will present different depths and prey distributions, and hence foraging at different depths might occur, and consequently different cue rates per location (across years). On the other hand, it seems like the variability from year might be not driven by year itself, but as random fluctuations over time, perhaps more sensibly accounted for as a factor (or a random effect). If one believes this to be the case, to predict cue rates for:

1. a sampled location and a new year, one could consider to use a model with location as a fixed effect, propagating the variability of year as a random effect;
2. for a new location and year, one could would use a model with both location and year as random effects.

In other words, intuitively year seems more sensibly modeled as a random effect, while location could be modeled either as as a fixed effect or as a random effect.

We can distinguish different levels of difficulty in terms of cue rate estimation with regards to the available information to do so. From easiest to hardest, we might want to estimate a cue rate for:

* a year-location combination we have data for;
* a location we have data for at a year we do not have data for;
* a year we have data for at a location we do not have data for;
* a completely new year-location combination.

From a conceptual point of view, if we had enough data across years and locations, one might want to:

1. treat as fixed effects those covariates for which we observed the level for which predictions are desired, but
2. treat as random effects those covariates for which the level at which we would like to predict were not observed

When considering a model with random effect(s) care must be had to propagate into predictions the variability associated with predicting for a new, previously unobserved, level of the random effect. As will be described below, this is not necessarily straightforward, and different alternatives are available to do so.

Note we will not be able to separate the effect of year and location for Kaikoura, since this location was only sampled in 2013, and no other location was sampled in that year. Additionally, we cannot separate effects for years 2020 and 2021 as we only have tags from a single site (the Azores) in those years.

A precautionary approach, when predicting a cue rate for new year and location combinations, might be to use both year and location as random effects in a random effects model and then propagate the uncertainty associated with the overall mean with respect to both random effects, year and location, that would be unobserved.

We can think of a few different ways to conceptualize the random effects. These include:

* Independent random effects for year and location
* A single random effect associated with each location-year combination
* A random effect associated with location and a nested-within-location random effect associated with year. This means that effect of location is shared across years, but the effect of year is not shared across locations

Below we will consider Bayesian implementations of [Nimble](#ilinkGLMMlyreNimble) to fit the models with random effects and propagate the uncertainty of unobserved random effect levels to predictions, but in <https://github.com/TiagoAMarques/Report4BB> there are other implementations, including in a frequentist framework.

### Model fitting

We will assume that cue rates, a strictly positive quantity, follow a Gamma distribution, and a log-link function will be considered within a generalized linear model (GLM) or generalized linear mixed model (GLMM) to ensure only admissible, i.e. positive, predictions.

#### GLMs

We begin by looking at GLM models where both location and year are treated as fixed effects including or not interaction terms between year and location. While for preliminary analysis we considered year as a numerical covariate, we then decided that there was no temporal trend expected, and hence year was more sensibly included as a factor. The use of year as a factor seems more sensible , as noted above. On the other hand, even if it were a real effect, the interaction would be almost impossible to evaluate/support with the available data, given there are very few year-location combinations for which such an interaction term could be estimated.

#run models  
CRglm0<-glm(crate~1,data=tags,family=Gamma(link="log"))  
CRglmL<-glm(crate~location,data=tags,family=Gamma(link="log"))  
CRglmfY<-glm(crate~fyear,data=tags,family=Gamma(link="log"))  
CRglm3<-glm(crate~location+fyear,data=tags,family=Gamma(link="log"))  
CRglm4<-glm(crate~location+fyear+fyear:location,data=tags,family=Gamma(link="log"))

AICtable <- cbind(data.frame(terms=c("~1","location","fyear","location+fyear","location\*fyear")),AIC(CRglm0,CRglmL,CRglmfY,CRglm3,CRglm4))  
AICtable$deltaAIC <- AICtable$AIC-min(AICtable$AIC)  
AICtable <- AICtable[order(AICtable$deltaAIC),]  
kable(AICtable,caption="Diferent GLMs considered to model the cue rate data")

Diferent GLMs considered to model the cue rate data

|  | terms | df | AIC | deltaAIC |
| --- | --- | --- | --- | --- |
| CRglmL | location | 8 | 49.53301 | 0.000000 |
| CRglm3 | location+fyear | 18 | 54.27206 | 4.739056 |
| CRglm4 | location\*fyear | 20 | 57.84576 | 8.312746 |
| CRglmfY | fyear | 14 | 58.67131 | 9.138305 |
| CRglm0 | ~1 | 2 | 60.37225 | 10.839238 |

According to AIC the best model considers only location as a factor, while year and the interaction between these factors is not deemed relevant. We can look at the summary of said model

summary(CRglmL)

##   
## Call:  
## glm(formula = crate ~ location, family = Gamma(link = "log"),   
## data = tags)  
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -0.22889 0.05152 -4.443 2.11e-05 \*\*\*  
## locationDOMINICA 0.24856 0.07172 3.466 0.000753 \*\*\*  
## locationGulf of Mexico 0.31390 0.07562 4.151 6.51e-05 \*\*\*  
## locationKaikoura 0.02654 0.12621 0.210 0.833812   
## locationMediterranean 0.05480 0.11229 0.488 0.626500   
## locationNorth Atlantic Delaware 0.05424 0.11846 0.458 0.647938   
## locationNorway -0.11955 0.10725 -1.115 0.267422   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for Gamma family taken to be 0.0796389)  
##   
## Null deviance: 14.561 on 117 degrees of freedom  
## Residual deviance: 12.041 on 111 degrees of freedom  
## AIC: 49.533  
##   
## Number of Fisher Scoring iterations: 4

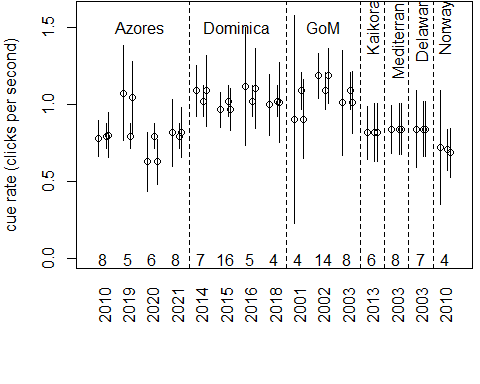
We illustrate how to estimate the cue rate, and its precision, for all year-location combinations for which we have more than 3 tags. The choice of 3 is arbitrary, but we considered that 3 or less tags would be unreliable.

We present empirical estimates (i.e. cue rate means for tags from a given year and site combinations) and estimates based on the GLM model. Note that these are not necessarily the same as empirical estimates use the data for a given year site combination, since the GLM estimates only use the predictions based on a fixed effect location. The predictions would only be the same if we used a model with a location-year interaction term.

# create all possible unique location-year combinations  
all.comb<-expand.grid(location=sort(unique(tags$location)),year=sort(unique(tags$year)))  
  
# select only those year-location combinations for which we have tags  
index.min<-which(table(tags$location,tags$year)!=0)  
# create an object to hold all possible location-year's independent of the number of tags (1 is enough to be in)  
byly.all <- all.comb[index.min,]  
  
# define minimum number of tags required to produce an estimate  
nmin<-3  
# select only those  
index.min<-which(table(tags$location,tags$year)>nmin)  
# create an object to hold results by location-year  
# all.comb was created way above!  
byly <- all.comb[index.min,]  
  
#now, for each combination  
for(i in 1:nrow(byly)){  
 index <- tags$location==byly$location[i] & tags$year==byly$year[i]  
 #-----------------------------------------------------------------  
 #get the empirical average cue rate estimate  
 #-----------------------------------------------------------------  
 #get the actual number of tags  
 byly$ntags[i] <- sum(index)  
 #calculate empirical cue rate  
 byly$ecr[i] <- mean(tags$crate[index])  
 #calculate empirical standard deviation  
 byly$ecrsd[i] <- sd(tags$crate[index])  
 #and the margin for a confidence interval  
 byly$margin.ecr[i] <- qt(0.975,byly$ntags[i]-1)\*byly$ecrsd[i]/sqrt(byly$ntags[i])  
 # get the empirical lower and upper 95% CI  
 byly$lcl.ecr <- with(byly,ecr - margin.ecr)  
 byly$ucl.ecr <- with(byly,ecr + margin.ecr)  
}  
# get the glm estimate  
temp1 <- predict(CRglm3,newdata = data.frame(location=byly$location,fyear=as.character(byly$year)),type="response",se=TRUE)  
temp2 <- predict(CRglmL,newdata = data.frame(location=byly$location),type="response",se=TRUE)  
# get the point estimates  
byly$glm.cr1 <- temp1$fit  
byly$glm.cr2 <- temp2$fit  
# get the ci's  
byly$glm.lci.cr1 <- temp1$fit - qt(0.975,summary(CRglm3)$df.residual)\*temp1$se  
byly$glm.uci.cr1 <- temp1$fit + qt(0.975,summary(CRglm3)$df.residual)\*temp1$se  
# get the ci's  
byly$glm.lci.cr2 <- temp2$fit - qt(0.975,summary(CRglmL)$df.residual)\*temp2$se  
byly$glm.uci.cr2 <- temp2$fit + qt(0.975,summary(CRglmL)$df.residual)\*temp2$se

We compare below standard averages (left) against values obtained from the glms (right, just location and left, location + year) for year-site combinations with more than 3 tags in terms of both point estimates (black points) and confidence intervals (solid vertical black lines). In between each such pair of estimates we added the number of tags available for each year-location combination.

#sort by location, then date  
byly <- byly[order(byly$location,byly$year),]  
#standard average  
par(mfrow=c(1,1),mar=c(6,4,0.1,0.4))  
plot(x=(1:nrow(byly))-0.2,y=byly$ecr,xaxt="n",ylim=c(0,1.6),xlim=c(0.5,nrow(byly)+0.5),xlab="",ylab="cue rate (clicks per second)")  
with(byly,segments(x0=(1:nrow(byly))-0.2,x1=(1:nrow(byly))-0.2,y0=lcl.ecr,y1=ucl.ecr))  
#GLMs  
points(x=(1:nrow(byly))+0.2,y=byly$glm.cr1)  
points(x=(1:nrow(byly))+0.1,y=byly$glm.cr2)  
with(byly,segments(x0=(1:nrow(byly))+0.2,x1=(1:nrow(byly))+0.2,y0=glm.lci.cr1,y1=glm.uci.cr1))  
with(byly,segments(x0=(1:nrow(byly))+0.1,x1=(1:nrow(byly))+0.1,y0=glm.lci.cr2,y1=glm.uci.cr2))  
#draw axis and annotations  
#axis(1, at=(1:nrow(byly))-0.2,byly$location,cex.axis=0.6,las=2)  
axis(1, at=(1:nrow(byly)),byly$year,tick=FALSE,cex.axis=1,las=2,line=0)  
text(x=(1:nrow(byly)),y=0,labels=byly$ntags,cex=1)  
abline(v=c(4.5,8.5,11.5,12.5,13.5,14.5),lty=2)  
myat <- 1.50  
text(2.5,myat,"Azores")  
text(6.5,myat,"Dominica")  
text(10,myat,"GoM")  
text(12,myat,"Kaikora",srt=90)  
text(13,myat,"Mediterranean",srt=90)  
text(14,myat,"Delaware",srt=90)  
text(15,myat,"Norway",srt=90)



This corresponds to figure 2 in the main paper

For easier reading and for possible use elsewhere, we also present these values in a table:

byly$crcv<-100\*byly$ecrsd/byly$ecr  
kable(byly,digits = 3)

|  | location | year | ntags | ecr | ecrsd | margin.ecr | lcl.ecr | ucl.ecr | glm.cr1 | glm.cr2 | glm.lci.cr1 | glm.uci.cr1 | glm.lci.cr2 | glm.uci.cr2 | crcv |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 29 | Azores | 2010 | 8 | 0.781 | 0.138 | 0.116 | 0.665 | 0.897 | 0.802 | 0.795 | 0.655 | 0.949 | 0.714 | 0.877 | 17.716 |
| 71 | Azores | 2019 | 5 | 1.074 | 0.247 | 0.307 | 0.767 | 1.381 | 1.044 | 0.795 | 0.810 | 1.278 | 0.714 | 0.877 | 23.013 |
| 78 | Azores | 2020 | 6 | 0.629 | 0.183 | 0.192 | 0.437 | 0.821 | 0.629 | 0.795 | 0.485 | 0.773 | 0.714 | 0.877 | 29.130 |
| 85 | Azores | 2021 | 8 | 0.816 | 0.258 | 0.216 | 0.600 | 1.032 | 0.816 | 0.795 | 0.654 | 0.978 | 0.714 | 0.877 | 31.655 |
| 44 | DOMINICA | 2014 | 7 | 1.089 | 0.178 | 0.165 | 0.924 | 1.254 | 1.089 | 1.020 | 0.858 | 1.320 | 0.919 | 1.121 | 16.352 |
| 51 | DOMINICA | 2015 | 16 | 0.965 | 0.213 | 0.114 | 0.852 | 1.079 | 0.965 | 1.020 | 0.830 | 1.101 | 0.919 | 1.121 | 22.096 |
| 58 | DOMINICA | 2016 | 5 | 1.116 | 0.307 | 0.381 | 0.735 | 1.497 | 1.101 | 1.020 | 0.844 | 1.359 | 0.919 | 1.121 | 27.483 |
| 65 | DOMINICA | 2018 | 4 | 0.997 | 0.122 | 0.195 | 0.802 | 1.192 | 1.014 | 1.020 | 0.754 | 1.274 | 0.919 | 1.121 | 12.286 |
| 3 | Gulf of Mexico | 2001 | 4 | 0.905 | 0.423 | 0.673 | 0.232 | 1.577 | 0.905 | 1.089 | 0.651 | 1.158 | 0.969 | 1.208 | 46.730 |
| 10 | Gulf of Mexico | 2002 | 14 | 1.185 | 0.252 | 0.145 | 1.040 | 1.330 | 1.185 | 1.089 | 1.007 | 1.363 | 0.969 | 1.208 | 21.250 |
| 17 | Gulf of Mexico | 2003 | 8 | 1.012 | 0.407 | 0.340 | 0.672 | 1.353 | 1.012 | 1.089 | 0.811 | 1.213 | 0.969 | 1.208 | 40.192 |
| 39 | Kaikoura | 2013 | 6 | 0.817 | 0.162 | 0.170 | 0.647 | 0.987 | 0.817 | 0.817 | 0.630 | 1.004 | 0.630 | 1.003 | 19.826 |
| 19 | Mediterranean | 2003 | 8 | 0.840 | 0.187 | 0.156 | 0.684 | 0.996 | 0.840 | 0.840 | 0.673 | 1.007 | 0.674 | 1.006 | 22.217 |
| 20 | North Atlantic Delaware | 2003 | 7 | 0.840 | 0.268 | 0.248 | 0.592 | 1.087 | 0.840 | 0.840 | 0.662 | 1.018 | 0.662 | 1.017 | 31.882 |
| 35 | Norway | 2010 | 4 | 0.722 | 0.232 | 0.370 | 0.353 | 1.092 | 0.686 | 0.706 | 0.525 | 0.847 | 0.574 | 0.837 | 32.152 |

These estimates illustrate that point estimates are not that different from each other using either means or the GLM, with considerable overlap in confidence intervals across year site combinations. In terms of cue rates, most location-year combinations are relatively consistent, with notable exceptions for the GoM in 2002, with higher estimated cue rate, and the Azores in 2020, with a considerably lower estimated cue rate.

Interestingly, the estimates from the GLM are similar or even slightly less precise then those from standard means, with notable exceptions being the most variable year-location combinations, where the glm returned narrower confidence intervals, namely for the GoM in 2001, Norway in 2010 and Dominica in 2016, where the strength borrowed from all the tags analysis means the model-based estimate is considerably more precise.

The GLM vs. empirical average comparison also illustrates how a naive cue rate confidence interval drawn from a reduced number of tags could result in inadmissible estimates for the latter (the naive 95% CI for the GoM in 2001 approaches 0, while negative values for the cue rate are not possible). From that perspective, estimates from the fitted model might be better, as these will avoid negative values by construction, induced by the log link.

#### GLMMs

##### Using lme4

Here we implemented 3 different GLMM models, with a different random effect structure:

1. Location and Year as independent random effects
2. Year random effect nested in location random effect
3. Year-location combo single random effect

This is implemented here:

# a model that considers each of year and location to be fully crossed and   
#independent random effects  
crglmer<-glmer(crate~(1|location)+(1|fyear),data=tags,family=Gamma(link="log"))  
# a Pooled model not separating the effect of year and location  
# so each year location combination is its own random effect level  
# equates to saying there's no site effect not year effect  
crglmerP<-glmer(crate~(1|locyear),data=tags,family=Gamma(link="log"))  
# Random effect of year nested in location  
crglmerN<-glmer(crate~(1|location/fyear),data=tags,family=Gamma(link="log"))

Interestingly, AIC - which might not be the best way to choose from different random effect stuctures - seems to favour the model where each year-site combination is a different level of a single year-site random effect.

kable(AIC(crglmer,crglmerP,crglmerN))

|  | df | AIC |
| --- | --- | --- |
| crglmer | 4 | 50.94656 |
| crglmerP | 3 | 50.63517 |
| crglmerN | 4 | 50.47458 |

It becomes an open question for now whether, to predict for a new location-year combination, one should consider the precision of this overall mean, or the precision associated with a new location-year mean (the classical difference between prediction intervals, for a new observation, and confidence intervals, for the mean). Here I am working on the latter, harder to estimate, although one might argue the former, much easier to estimate, might be appropriate.

To do so we need to propagate the variability associated with the random effects into the point estimate of the intercept of the model. This is not necessarily a straightforward thing to do.

Here we implement the GLMM model within a Bayesian context, choosing NIMBLE for implementation (<https://r-nimble.org/>). NIMBLE adopts and extends BUGS as a modeling language and lets you program with the models you create.

##### Location and Year as independent random effects

First we define the Nimble model:

## define the model  
GLMMcode <- nimbleCode({  
 # define priors  
 # the overall intercept  
 beta0 ~ dnorm(0, sd = 10)  
 # random effect standard deviation associated with location, a uniform, might change this to be something else latter  
 sigmal\_RE ~ dunif(0,10)  
 # random effect standard deviation associated with year, a uniform, might change this to be something else latter  
 sigmay\_RE ~ dunif(0, 10)  
 # the gamma dispersion (or variance - see commented parametrization 1) parameter, a uniform, might change this to be something else latter  
 #dispersion ~ dunif(0, 10)  
 disp ~ dunif(0, 10)  
 ## sd ~ dhalfflat()  
 #get year random effects  
 for(yy in 1:nyears){  
 #REy[yy] ~ dnorm(0, sd = sigmay\_RE)  
 REy[yy] ~ dnorm(0, sd = sigmay\_RE)  
 }   
 #get location random effects  
 for(ll in 1:nlocs){  
 #REl[ll] ~ dnorm(0, sd = sigmal\_RE)  
 REl[ll] ~ dnorm(0, sd = sigmal\_RE)  
 }   
 for (i in 1:N){  
 #get the linear predictor, consider a log link function  
 log(mean[i]) <- beta0 + REy[year[i]] + REl[loc[i]]  
 # now Using decentered parametrization, a suggestion by Ben Augustine  
 # log(mean[i]) <- beta0 + REy[year[i]]\*sigmay\_RE + REl[loc[i]]\*sigmal\_RE  
 # parametrization 1 - now I know that to not be what is to be used  
 #left here for future reference  
 # crate[i] ~ dgamma(shape=(mean[i]^2)/disp,scale=disp/mean[i])  
 #parametrization 2  
 crate[i] ~ dgamma(shape=1/disp, scale=mean[i]\*disp)  
 }  
})

And then we define the required constants, data, and initial values, as well as the nodes to monitor in the MCMC

## constants, data, and initial values  
  
#constant, sample size and number of levels for each of the random effects  
#number of rows in tags  
N<-nrow(tags)  
#number of different years  
nyears <- length(unique(tags$year))  
#number of different locations  
nlocs <- length(unique(tags$location))  
#the year covariate is passed as a constant  
year <- as.numeric(tags$fyear)  
#the location covaraite is passed as a constant  
loc <- as.numeric(tags$location)  
#bundle all in a suitable object  
constants <- list(N = N,nyears=nyears,nlocs=nlocs,year = year,loc = loc)  
  
#data  
data <- list(crate = tags$crate)  
  
#initial values  
#inits <- list(beta0 = 0, sigmal\_RE = 1, sigmay\_RE = 1, dispersion = 0.8,REy = rep(0,nyears),REl = rep(0,nlocs))  
inits <- list(beta0 = 0, sigmal\_RE = 1, sigmay\_RE = 1, disp = 1,REy = rep(0,nyears),REl = rep(0,nlocs))

## create the model object  
myGLMMModel <- nimbleModel(code = GLMMcode, constants = constants, data = data,   
 inits = inits, check = FALSE, buildDerivs = TRUE) ## Add buildDerivs = TRUE for AD

## Defining model

## Building model

## Setting data and initial values

## Running calculate on model  
## [Note] Any error reports that follow may simply reflect missing values in model variables.

## Checking model sizes and dimensions

cmyGLMMModel <- compileNimble(myGLMMModel)

## Compiling  
## [Note] This may take a minute.  
## [Note] Use 'showCompilerOutput = TRUE' to see C++ compilation details.

#things to monitor  
#tomon<-c("beta0","dispersion","sigmay\_RE","sigmal\_RE","crate","REy","REl")  
tomon<-c("beta0","disp","sigmay\_RE","sigmal\_RE","REy","REl")

Then we run the code, considering 50000 iterations with a 10000 iterations burnin period, leaving 40000 iterations for inference.

test<-nimbleMCMC(myGLMMModel,monitors=tomon,niter=50000,nburnin=10000,progressBar=TRUE,summary=TRUE)

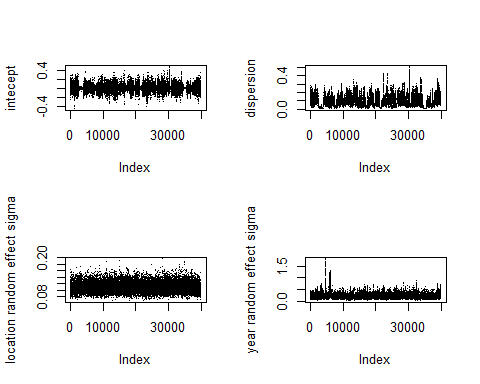
## Compiling  
## [Note] This may take a minute.  
## [Note] Use 'showCompilerOutput = TRUE' to see C++ compilation details.

## running chain 1...

## |-------------|-------------|-------------|-------------|  
## |-------------------------------------------------------|

There were no apparent issues with convergence/mixing of MCMC chains

par(mfrow=c(2,2))  
#trace plot intercept  
plot(test$samples[,20],pch=".",ylab="intecept")  
#trace plot dispersion  
plot(test$samples[,nrow(test$summary)],pch=".",ylab="dispersion")  
#trace plot year random effect standard deviation  
plot(test$samples[,nrow(test$summary)-2],pch=".",ylab="location random effect sigma")  
#trace plot location random effect standard deviation  
plot(test$samples[,nrow(test$summary)-1],pch=".",ylab="year random effect sigma")



We can look at the main results, i.e. all top-level stochastic nodes of the model, namely the intercept, the dispersion parameter for the Gamma and the standard deviations of both random effects

#look at main results  
kable(test$summary,digits=2)

|  | Mean | Median | St.Dev. | 95%CI\_low | 95%CI\_upp |
| --- | --- | --- | --- | --- | --- |
| REl[1] | -0.09 | -0.08 | 0.12 | -0.35 | 0.12 |
| REl[2] | 0.13 | 0.13 | 0.12 | -0.09 | 0.38 |
| REl[3] | 0.15 | 0.15 | 0.12 | -0.07 | 0.38 |
| REl[4] | -0.04 | -0.03 | 0.14 | -0.34 | 0.23 |
| REl[5] | -0.03 | -0.02 | 0.13 | -0.30 | 0.22 |
| REl[6] | -0.03 | -0.02 | 0.13 | -0.30 | 0.22 |
| REl[7] | -0.15 | -0.14 | 0.14 | -0.47 | 0.07 |
| REy[1] | -0.02 | -0.01 | 0.08 | -0.21 | 0.14 |
| REy[2] | 0.07 | 0.04 | 0.09 | -0.06 | 0.28 |
| REy[3] | 0.00 | 0.00 | 0.07 | -0.15 | 0.15 |
| REy[4] | -0.01 | 0.00 | 0.09 | -0.21 | 0.18 |
| REy[5] | -0.01 | 0.00 | 0.07 | -0.16 | 0.13 |
| REy[6] | -0.01 | 0.00 | 0.08 | -0.21 | 0.16 |
| REy[7] | 0.02 | 0.01 | 0.08 | -0.12 | 0.20 |
| REy[8] | -0.02 | -0.01 | 0.07 | -0.19 | 0.12 |
| REy[9] | 0.00 | 0.00 | 0.07 | -0.15 | 0.16 |
| REy[10] | -0.03 | -0.01 | 0.07 | -0.20 | 0.11 |
| REy[11] | 0.07 | 0.05 | 0.09 | -0.05 | 0.31 |
| REy[12] | -0.06 | -0.03 | 0.09 | -0.29 | 0.06 |
| REy[13] | 0.01 | 0.00 | 0.07 | -0.14 | 0.17 |
| beta0 | -0.12 | -0.13 | 0.11 | -0.33 | 0.10 |
| disp | 0.11 | 0.11 | 0.01 | 0.08 | 0.14 |
| sigmal\_RE | 0.19 | 0.16 | 0.12 | 0.05 | 0.46 |
| sigmay\_RE | 0.08 | 0.07 | 0.06 | 0.01 | 0.22 |

# getting the overall estimate for a new year and location and respective CI's we are looking for  
# heuristically appealing but... is this kosher?  
nsamples<-nrow(test$samples)  
myquants95CI.Bayes<-quantile(exp(test$samples[,which(rownames(test$summary)=="beta0")]+rnorm(nsamples,mean=0,sd=test$samples[,which(rownames(test$summary)=="sigmal\_RE")])+rnorm(nsamples,mean=0,sd=test$samples[,which(rownames(test$summary)=="sigmay\_RE")])),probs=c(0.025,0.5,0.975))  
myquants95CI.Bayes

## 2.5% 50% 97.5%   
## 0.5285296 0.8826588 1.5026216

##### Year random effect nested in location random effect

First we define the Nimble model:

## define the model  
GLMMcodeYnL <- nimbleCode({  
 # the overall intercept  
 beta0 ~ dnorm(0, sd = 10)  
 # random effect standard deviation associated with location, a uniform, might change this to be something else latter  
 sigmal\_RE ~ dunif(0,10)  
 # random effect standard deviation associated with year, nested in location, a uniform, might change this to be something else latter  
 sigmaynl\_RE ~ dunif(0, 10)  
 # the gamma dispersion (or variance - see commented parametrization 1) parameter, a uniform, might change this to be something else latter  
 #dispersion ~ dunif(0, 10)  
 disp ~ dunif(0, 10)  
 #get year random effects, nested in locations (i.e. more levels than just years)  
 for(yy in 1:nyearsnl){  
 REynl[yy] ~ dnorm(0, sd = sigmaynl\_RE)  
 }   
 #get location random effects  
 for(ll in 1:nlocs){  
 REl[ll] ~ dnorm(0, sd = sigmal\_RE)  
 }   
 for (i in 1:N){  
 #get the linear predictor, consider a log link function  
 log(mean[i]) <- beta0 + REynl[yearnl[i]] + REl[loc[i]]  
 #on the response scale  
 crate[i] ~ dgamma(shape=1/disp, scale=mean[i]\*disp)  
 }  
})

And then we define the required constants, data, and initial values, as well as the nodes to monitor in the MCMC

## constants, data, and initial values  
  
#constant, sample size and number of levels for each of the random effects  
#number of rows in tags  
N<-nrow(tags)  
#number of different years nested in locations (= number of location year combinations)  
nyearsnl <- length(unique(tags$locyear))  
#number of different locations  
nlocs <- length(unique(tags$location))  
#the year within loc covariate is passed as a constant  
yearnl <- as.numeric(tags$locyear)  
#the location covaraite is passed as a constant  
loc <- as.numeric(tags$location)  
#bundle all in a suitable object  
constants <- list(N = N,nyearsnl=nyearsnl,nlocs=nlocs,yearnl = yearnl,loc = loc)  
  
#data  
data <- list(crate = tags$crate)  
  
#initial values  
#inits <- list(beta0 = 0, sigmal\_RE = 1, sigmay\_RE = 1, dispersion = 0.8,REy = rep(0,nyears),REl = rep(0,nlocs))  
inits <- list(beta0 = 0, sigmal\_RE = 1, sigmaynl\_RE = 1, disp = 1,REynl = rep(0,nyearsnl),REl = rep(0,nlocs))

## create the model object  
myGLMMModelynl <- nimbleModel(code = GLMMcodeYnL, constants = constants, data = data,   
 inits = inits, check = FALSE, buildDerivs = TRUE) ## Add buildDerivs = TRUE for AD

## Defining model

## Building model

## Setting data and initial values

## Running calculate on model  
## [Note] Any error reports that follow may simply reflect missing values in model variables.

## Checking model sizes and dimensions

cmyGLMMModelynl <- compileNimble(myGLMMModelynl)

## Compiling  
## [Note] This may take a minute.  
## [Note] Use 'showCompilerOutput = TRUE' to see C++ compilation details.

#things to monitor  
tomonynl<-c("beta0","disp","sigmaynl\_RE","sigmal\_RE","REynl","REl")

Then we run the code, considering 55000 iterations with a 15000 iterations burnin period, leaving 40000 iterations for inference.

testynl<-nimbleMCMC(myGLMMModelynl,monitors=tomonynl,niter=55000,nburnin=15000,progressBar=TRUE,summary=TRUE)

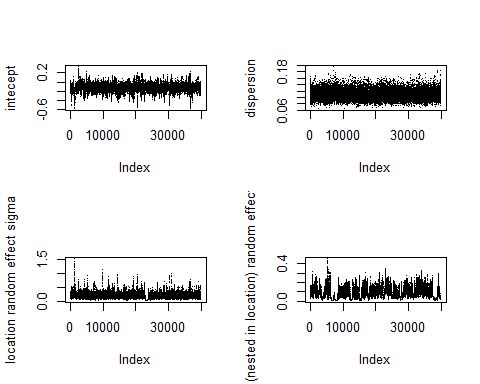
## Compiling  
## [Note] This may take a minute.  
## [Note] Use 'showCompilerOutput = TRUE' to see C++ compilation details.

## running chain 1...

## |-------------|-------------|-------------|-------------|  
## |-------------------------------------------------------|

There were no apparent issues with convergence/mixing of MCMC chains

par(mfrow=c(2,2))  
#trace plot intercept  
plot(testynl$samples[,which(rownames(testynl$summary)=="beta0")],pch=".",ylab="intecept")  
#trace plot dispersion  
plot(testynl$samples[,which(rownames(testynl$summary)=="disp")],pch=".",ylab="dispersion")  
#trace plot location random effect standard deviation  
plot(testynl$samples[,which(rownames(testynl$summary)=="sigmal\_RE")],pch=".",ylab="location random effect sigma")  
#trace plot year random effect standard deviation  
plot(testynl$samples[,which(rownames(testynl$summary)=="sigmaynl\_RE")],pch=".",ylab="year (nested in location) random effect sigma")



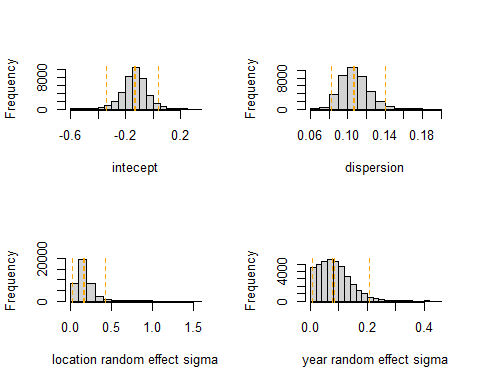
We can look at the main results, i.e. all top-level stochastic nodes of the model, namely the intercept, the dispersion parameter and the standard deviations of the random effects

#look at main results -   
testynl$summary

## Mean Median St.Dev. 95%CI\_low 95%CI\_upp  
## REl[1] -0.072144082 -7.044076e-02 0.10275122 -0.281247092 0.14016827  
## REl[2] 0.126264187 1.177502e-01 0.10895097 -0.060447039 0.36818456  
## REl[3] 0.150190805 1.421835e-01 0.11924370 -0.052231671 0.41178956  
## REl[4] -0.025756970 -2.009531e-02 0.12777711 -0.293816500 0.22907197  
## REl[5] -0.013665103 -1.155181e-02 0.12155911 -0.257561418 0.24433610  
## REl[6] -0.014672373 -1.386718e-02 0.12471542 -0.265641162 0.25137841  
## REl[7] -0.118895907 -1.090222e-01 0.12253482 -0.387580692 0.09842536  
## REynl[1] -0.014293492 -6.770309e-03 0.07326253 -0.180899894 0.13213129  
## REynl[2] -0.039399845 -1.993823e-02 0.08908105 -0.257603592 0.10897281  
## REynl[3] 0.074358683 5.118288e-02 0.09389634 -0.056935113 0.30426426  
## REynl[4] -0.066828662 -4.465719e-02 0.09344072 -0.294625518 0.06714446  
## REynl[5] 0.001151951 -7.959793e-06 0.07222661 -0.154060519 0.15580310  
## REynl[6] 0.031839708 1.575316e-02 0.08166994 -0.110662672 0.22927918  
## REynl[7] -0.008088292 -5.553818e-03 0.07021762 -0.156646619 0.14513314  
## REynl[8] 0.034197660 1.842759e-02 0.08420263 -0.112540635 0.23591984  
## REynl[9] 0.003316526 5.051348e-04 0.08153766 -0.166165119 0.18588344  
## REynl[10] -0.018804156 -8.885366e-03 0.08359983 -0.206711630 0.14628055  
## REynl[11] 0.073940862 5.203301e-02 0.09155221 -0.054790798 0.29229705  
## REynl[12] 0.007908553 2.756750e-03 0.07712511 -0.149452150 0.18253202  
## REynl[13] -0.009303619 -4.843556e-03 0.08362682 -0.190151564 0.16028157  
## REynl[14] -0.007401313 -2.645527e-03 0.08199379 -0.190879483 0.15990765  
## REynl[15] -0.006348170 -3.061123e-03 0.08164564 -0.185935687 0.16598861  
## REynl[16] -0.012763961 -4.827715e-03 0.09275923 -0.227291959 0.17387950  
## REynl[17] -0.015532761 -6.424084e-03 0.08581621 -0.212388984 0.15201940  
## REynl[18] -0.039769873 -1.828266e-02 0.09696039 -0.282912695 0.12396548  
## REynl[19] 0.008146330 3.803156e-03 0.08827318 -0.175825915 0.20430058  
## beta0 -0.134792613 -1.310947e-01 0.09069314 -0.339880218 0.04196548  
## disp 0.107535126 1.063333e-01 0.01503944 0.082261941 0.14066472  
## sigmal\_RE 0.172921015 1.549770e-01 0.10434110 0.022775773 0.42233620  
## sigmaynl\_RE 0.083868856 7.762735e-02 0.05407698 0.006734904 0.20731270

We can see posterior distributions

par(mfrow=c(2,2))  
#posterior plot intercept  
intercepts<-testynl$samples[,which(rownames(testynl$summary)=="beta0")]  
hist(intercepts,pch=".",xlab="intecept",main="")  
abline(v=quantile(intercepts,probs=c(0.025,0.5,0.975)),col="orange",lty=2)  
abline(v=mean(intercepts),col="orange",lty=2)  
#posterior plot dispersion  
dispersions<-testynl$samples[,which(rownames(testynl$summary)=="disp")]  
hist(dispersions,pch=".",xlab="dispersion",main="")  
abline(v=quantile(dispersions,probs=c(0.025,0.5,0.975)),col="orange",lty=2)  
abline(v=mean(dispersions),col="orange",lty=2)  
#posterior plot location random effect standard deviation  
lsigmas<-testynl$samples[,which(rownames(testynl$summary)=="sigmal\_RE")]  
hist(lsigmas,pch=".",xlab="location random effect sigma",main="")  
abline(v=quantile(lsigmas,probs=c(0.025,0.5,0.975)),col="orange",lty=2)  
abline(v=mean(lsigmas),col="orange",lty=2)  
#posterior plot year random effect standard deviation  
ynlsigmas<-testynl$samples[,which(rownames(testynl$summary)=="sigmaynl\_RE")]  
hist(ynlsigmas,pch=".",xlab="year random effect sigma",main="")  
abline(v=quantile(ynlsigmas,probs=c(0.025,0.5,0.975)),col="orange",lty=2)  
abline(v=mean(ynlsigmas),col="orange",lty=2)



Assuming that the MCMC approach, which does not use any approximations, is the one most reliable, we would finally obtain the variability around the mean by simulating from the relevant posterior distributions

# getting the overall estimate for a new year and location and respective CI's we are looking for  
# heuristically appealing but... is this kosher?  
nsample.ynls<-nrow(testynl$samples)  
myquants95CI.Bayes.ynl<-quantile(exp(testynl$samples[,which(rownames(testynl$summary)=="beta0")]+rnorm(nsamples,mean=0,sd=testynl$samples[,which(rownames(testynl$summary)=="sigmal\_RE")])+rnorm(nsamples,mean=0,sd=testynl$samples[,which(rownames(testynl$summary)=="sigmaynl\_RE")])),probs=c(0.025,0.5,0.975))

##### Year-location combo single random effect

First we define the Nimble model:

## define the model  
GLMMcode1re <- nimbleCode({  
 # the overall intercept  
 beta0 ~ dnorm(0, sd = 10)  
 # random effect standard deviation associated with locatio-year, a uniform, might change this to be something else latter  
 sigma1re\_RE ~ dunif(0,10)  
 # the gamma dispersion (or variance - see commented parametrization 1) parameter, a uniform, might change this to be something else latter  
 #dispersion ~ dunif(0, 10)  
 disp ~ dunif(0, 10)  
 #get year random effects  
 for(yy in 1:nyearsnl){  
 RE1re[yy] ~ dnorm(0, sd = sigma1re\_RE)  
 }   
 for (i in 1:N){  
 #get the linear predictor, consider a log link function  
 log(mean[i]) <- beta0 + RE1re[yearnl[i]]  
 #on the response scale  
 crate[i] ~ dgamma(shape=1/disp, scale=mean[i]\*disp)  
 }  
})

And then we define the required constants, data, and initial values, as well as the nodes to monitor in the MCMC

## constants, data, and initial values  
  
#constant, sample size and number of levels for each of the random effects  
#number of rows in tags  
N<-nrow(tags)  
#number of different years nested in locations (= number of location year combinations)  
nyearsnl <- length(unique(tags$locyear))  
#the year within loc covariate is passed as a constant  
yearnl <- as.numeric(tags$locyear)  
#bundle all in a suitable object  
constants1re <- list(N = N,nyearsnl=nyearsnl,yearnl = yearnl)  
#data  
data <- list(crate = tags$crate)  
  
#initial values  
inits1re <- list(beta0 = 0, sigma1re\_RE = 1, disp = 1,RE1re = rep(0,nyearsnl))

## create the model object  
myGLMMModel1re <- nimbleModel(code = GLMMcode1re, constants = constants1re, data = data,   
 inits = inits1re, check = FALSE, buildDerivs = TRUE) ## Add buildDerivs = TRUE for AD

## Defining model

## Building model

## Setting data and initial values

## Running calculate on model  
## [Note] Any error reports that follow may simply reflect missing values in model variables.

## Checking model sizes and dimensions

cmyGLMMModel1re <- compileNimble(myGLMMModel1re)

## Compiling  
## [Note] This may take a minute.  
## [Note] Use 'showCompilerOutput = TRUE' to see C++ compilation details.

#things to monitor  
tomon1re<-c("beta0","disp","sigma1re\_RE","RE1re")

Then we run the code, considering 50000 iterations with a 15000 iterations burnin period, leaving 40000 iterations for inference.

test1re<-nimbleMCMC(myGLMMModel1re,monitors=tomon1re,niter=55000,nburnin=15000,progressBar=TRUE,summary=TRUE)

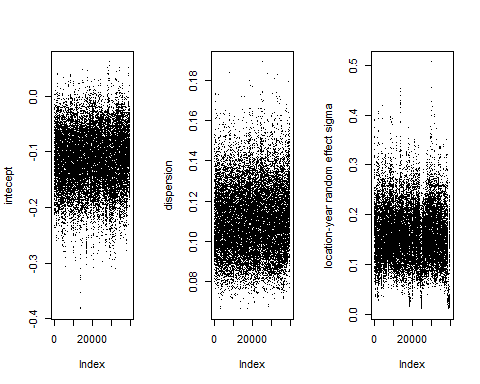
## Compiling  
## [Note] This may take a minute.  
## [Note] Use 'showCompilerOutput = TRUE' to see C++ compilation details.

## running chain 1...

## |-------------|-------------|-------------|-------------|  
## |-------------------------------------------------------|

There were no apparent issues with convergence/mixing of MCMC chains

par(mfrow=c(1,3))  
#trace plot intercept  
plot(test1re$samples[,which(rownames(test1re$summary)=="beta0")],pch=".",ylab="intecept")  
#trace plot dispersion  
plot(test1re$samples[,which(rownames(test1re$summary)=="disp")],pch=".",ylab="dispersion")  
#trace plot location random effect standard deviation  
plot(test1re$samples[,which(rownames(test1re$summary)=="sigma1re\_RE")],pch=".",ylab="location-year random effect sigma")



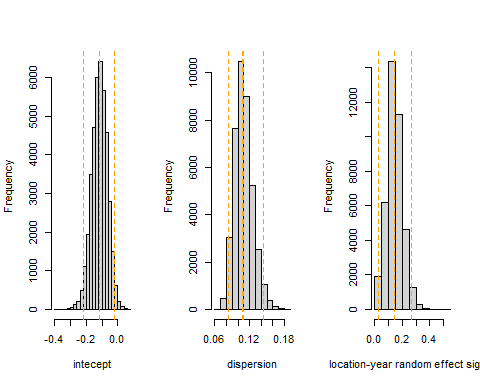
We can look at the main results, i.e. all top-level stochastic nodes of the model, namely the intercept, the dispersion parameter and the standard deviations of the random effects

#look at main results -   
test1re$summary

## Mean Median St.Dev. 95%CI\_low 95%CI\_upp  
## RE1re[1] -0.07060900 -0.065984337 0.09603578 -0.270394451 0.11127654  
## RE1re[2] -0.10528768 -0.093927249 0.12804696 -0.387462444 0.12100718  
## RE1re[3] 0.09351435 0.083762643 0.10999069 -0.097725417 0.32900256  
## RE1re[4] -0.16326105 -0.156551669 0.11983692 -0.419199469 0.03919658  
## RE1re[5] -0.04315879 -0.039505794 0.09461470 -0.234939722 0.14120330  
## RE1re[6] 0.11449006 0.107284770 0.10405989 -0.064375975 0.33660314  
## RE1re[7] 0.05947461 0.055279258 0.07923230 -0.087424218 0.22363087  
## RE1re[8] 0.11164515 0.103682899 0.11214979 -0.086040200 0.35174014  
## RE1re[9] 0.05426878 0.044040842 0.10924222 -0.143454646 0.29101189  
## RE1re[10] 0.01174959 0.008609002 0.10853102 -0.202039127 0.23784308  
## RE1re[11] 0.19902253 0.196875489 0.10129571 0.008480719 0.40230032  
## RE1re[12] 0.07785059 0.070369458 0.09653246 -0.096055902 0.28053908  
## RE1re[13] -0.03820086 -0.034997745 0.09893053 -0.240255251 0.15962057  
## RE1re[14] -0.02728512 -0.025727450 0.09290512 -0.210571895 0.15892551  
## RE1re[15] -0.02844707 -0.024261410 0.09688776 -0.227235370 0.16536284  
## RE1re[16] -0.04200486 -0.032188126 0.13947259 -0.341897776 0.22546792  
## RE1re[17] -0.08149677 -0.073088790 0.11504800 -0.324635199 0.13093098  
## RE1re[18] -0.11147383 -0.096108149 0.14200505 -0.429796108 0.13004369  
## RE1re[19] -0.01553352 -0.011691847 0.12390754 -0.268021427 0.23289110  
## beta0 -0.11567392 -0.114336506 0.05039011 -0.217442018 -0.02150094  
## disp 0.10970925 0.108247095 0.01556429 0.083341411 0.14396535  
## sigma1re\_RE 0.14441806 0.141488298 0.05764713 0.025696789 0.26574811

We can see posterior distributions

par(mfrow=c(1,3))  
#posterior plot intercept  
intercepts<-test1re$samples[,which(rownames(test1re$summary)=="beta0")]  
hist(intercepts,pch=".",xlab="intecept",main="")  
abline(v=quantile(intercepts,probs=c(0.025,0.5,0.975)),col="orange",lty=2)  
abline(v=mean(intercepts),col="orange",lty=2)  
#posterior plot dispersion  
dispersions<-test1re$samples[,which(rownames(test1re$summary)=="disp")]  
hist(dispersions,pch=".",xlab="dispersion",main="")  
abline(v=quantile(dispersions,probs=c(0.025,0.5,0.975)),col="orange",lty=2)  
abline(v=mean(dispersions),col="orange",lty=2)  
#posterior plot location random effect standard deviation  
lysigmas<-test1re$samples[,which(rownames(test1re$summary)=="sigma1re\_RE")]  
hist(lysigmas,pch=".",xlab="location-year random effect sigma",main="")  
abline(v=quantile(lysigmas,probs=c(0.025,0.5,0.975)),col="orange",lty=2)  
abline(v=mean(lysigmas),col="orange",lty=2)



Assuming that the MCMC approach, which does not use any approximations, is the one most reliable, we would finally obtain the variability around the mean by simulating from the relevant posterior distributions

# getting the overall estimate for a new year and location and respective CI's we are looking for  
# heuristically appealing but... is this kosher?  
nsample.1re<-nrow(test1re$samples)  
myquants95CI.Bayes.1re<-quantile(exp(test1re$samples[,which(rownames(test1re$summary)=="beta0")]+rnorm(nsample.1re,mean=0,sd=test1re$samples[,which(rownames(test1re$summary)=="sigma1re\_RE")])),probs=c(0.025,0.5,0.975))

##### Comparison of random effect stuctures

Considering two independent random effects for year and location, the mean cue rate would be estimated at 0.89 (median: 0.88, 95% credible interval 0.55-1.47). This corresponds to a CV of 50.76.

Considering year nested within location for the random effect structure, the mean cue rate would be estimated at 0.88 (median: 0.88, 95% credible interval 0.55-1.38). This corresponds to a CV of 47.09.

Considering a single year-location combination as random effect structure, the mean cue rate would be estimated at 0.89 (median: 0.9, 95% credible interval 0.65-1.22). This corresponds to a CV of 33.07.

## Animal specific covariates

### Inter-Click-Interval

Here we look at whether the Inter-Click-Interval (ICI) data might be used to predict cue rate.

Read the click data

#reading the clicks data  
#clall<-read.table(file="data/clall.txt")  
clall<-fread(file="data/clall.txt")

## Warning in fread(file = "data/clall.txt"): Detected 3 column names but the data  
## has 4 columns (i.e. invalid file). Added 1 extra default column name for the  
## first column which is guessed to be row names or an index. Use setnames()  
## afterwards if this guess is not correct, or fix the file write command that  
## created the file to create a valid file.

nclall<-nrow(clall)

do a bit of data processing

#get the ici, not last ici of each whale does not exist, so here  
#add ici of NA for the last click of the last whale  
clall$ici <- c(clall$sss[2:nclall]-clall$sss[1:(nclall-1)],NA)  
  
#index of where you change whale, so last click of a whale  
clall$cw <- c(clall$tag[2:nclall]==clall$tag[1:(nclall-1)],NA)  
#and remove those ici's, which are not interpretable as they are calculated across whales  
  
clall <- clall[clall$cw,]  
clall <- clall[-nrow(clall),]  
# calculating per whale  
# mean/median/mode ici's, with trimmed data  
  
#first, only get regular click ICI, which implies  
#between 0.2 seconds and say 5 seconds  
clall <- clall[clall$ici>0.2 & clall$ici<5,]  
  
#a bespoke function to get a mode  
get.mode <- function(x){  
xdens = density(x)  
modex = xdens$x[which.max(xdens$y)]  
return(modex)  
}

Calculate summary statistics per tag

#summarize data  
icitags<-tapply(X=clall$tag,INDEX=clall$tag,FUN=head,n=1)  
icimean<-tapply(X=clall$ici,INDEX=clall$tag,FUN=mean)  
icimean0.05<-tapply(X=clall$ici,INDEX=clall$tag,FUN=mean,trim = 0.05)  
icimean0.1<-tapply(X=clall$ici,INDEX=clall$tag,FUN=mean,trim = 0.1)  
icimedian<-tapply(X=clall$ici,INDEX=clall$tag,FUN=median)  
icimode<-tapply(X=clall$ici,INDEX=clall$tag,FUN=get.mode)  
# make a single suitable object  
ici.sums<-data.frame(tag=icitags,mean=icimean,mean0.05=icimean0.05,mean0.1=icimean0.1,median=icimedian,mode=icimode)  
#pairs(ici.sums[,-1])

And finally merge this information with the click rates, to be able to model click rates as a function of ICI

tagsici<-merge(tags,ici.sums,by="tag",all.x=TRUE,all.y=FALSE)

We look below at linear models of click rate as a function of several statistics based on ICI’s, namely:

* mean ICI
* mean ICI, trimed 0.05
* mean ICI, trimed 0.1
* median ICI
* mode ICI

and at the end we present a GLM model of the statistic that seems to better correlate with cue rate, which is the mean.

par(mfrow=c(3,2))  
with(tagsici,plot(crate~mean))  
lmmean<-lm(crate~mean,data=tagsici)  
summary(lmmean)

##   
## Call:  
## lm(formula = crate ~ mean, data = tagsici)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.87238 -0.14217 0.00855 0.17907 0.54976   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1.7318 0.1196 14.482 < 2e-16 \*\*\*  
## mean -1.4243 0.2063 -6.903 2.88e-10 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.2419 on 116 degrees of freedom  
## Multiple R-squared: 0.2912, Adjusted R-squared: 0.285   
## F-statistic: 47.65 on 1 and 116 DF, p-value: 2.884e-10

abline(lmmean)  
#  
with(tagsici,plot(crate~mean0.05))  
lmmean0.05<-lm(crate~mean0.05,data=tagsici)  
summary(lmmean0.05)

##   
## Call:  
## lm(formula = crate ~ mean0.05, data = tagsici)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.86357 -0.14401 0.01012 0.17221 0.56835   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1.7043 0.1194 14.276 < 2e-16 \*\*\*  
## mean0.05 -1.4205 0.2126 -6.682 8.61e-10 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.2442 on 116 degrees of freedom  
## Multiple R-squared: 0.2779, Adjusted R-squared: 0.2717   
## F-statistic: 44.65 on 1 and 116 DF, p-value: 8.61e-10

abline(lmmean)  
#  
with(tagsici,plot(crate~mean0.1))  
lmmean0.1<-lm(crate~mean0.1,data=tagsici)  
summary(lmmean0.1)

##   
## Call:  
## lm(formula = crate ~ mean0.1, data = tagsici)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.86439 -0.13920 0.01618 0.16824 0.57306   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1.6940 0.1188 14.26 < 2e-16 \*\*\*  
## mean0.1 -1.4141 0.2133 -6.63 1.11e-09 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.2447 on 116 degrees of freedom  
## Multiple R-squared: 0.2748, Adjusted R-squared: 0.2685   
## F-statistic: 43.95 on 1 and 116 DF, p-value: 1.114e-09

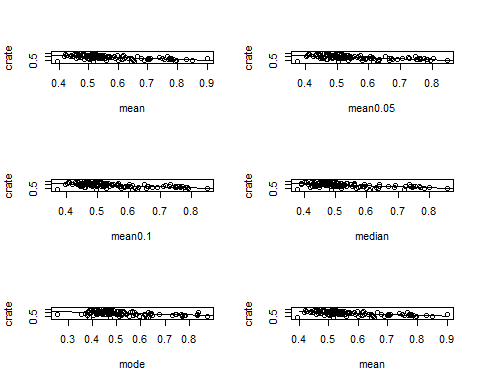
abline(lmmean)  
#  
with(tagsici,plot(crate~median))  
lmmedian<-lm(crate~median,data=tagsici)  
summary(lmmedian)

##   
## Call:  
## lm(formula = crate ~ median, data = tagsici)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.86358 -0.14132 0.00838 0.16879 0.58454   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1.6540 0.1176 14.063 < 2e-16 \*\*\*  
## median -1.3640 0.2147 -6.354 4.26e-09 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.2475 on 116 degrees of freedom  
## Multiple R-squared: 0.2582, Adjusted R-squared: 0.2518   
## F-statistic: 40.38 on 1 and 116 DF, p-value: 4.261e-09

abline(lmmedian)  
#  
with(tagsici,plot(crate~mode))  
lmmode<-lm(crate~mode,data=tagsici)  
summary(lmmode)

##   
## Call:  
## lm(formula = crate ~ mode, data = tagsici)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.8379 -0.1607 0.0074 0.1860 0.6172   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1.4335 0.1052 13.621 < 2e-16 \*\*\*  
## mode -0.9913 0.1981 -5.003 2.03e-06 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.2606 on 116 degrees of freedom  
## Multiple R-squared: 0.1775, Adjusted R-squared: 0.1704   
## F-statistic: 25.03 on 1 and 116 DF, p-value: 2.029e-06

abline(lmmode)  
  
with(tagsici,plot(crate~mean))  
glmmean<-glm(crate~mean,data=tagsici,family=Gamma(link="log"))  
newicis<-seq(0.4,0.9,length=100)  
preds <- predict(glmmean,newdata=data.frame(mean=newicis),type="response")  
lines(newicis,preds)

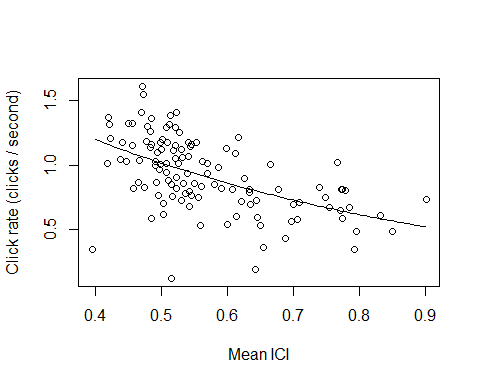


summary(glmmean)

##   
## Call:  
## glm(formula = crate ~ mean, family = Gamma(link = "log"), data = tagsici)  
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.8450 0.1279 6.605 1.26e-09 \*\*\*  
## mean -1.6559 0.2208 -7.501 1.39e-11 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for Gamma family taken to be 0.06700011)  
##   
## Null deviance: 14.561 on 117 degrees of freedom  
## Residual deviance: 10.922 on 116 degrees of freedom  
## AIC: 27.829  
##   
## Number of Fisher Scoring iterations: 4

This image might be used in the paper

par(mfrow=c(1,1))  
with(tagsici,plot(crate~mean,ylab="Click rate (clicks / second)",xlab="Mean ICI"))  
glmmean<-glm(crate~mean,data=tagsici,family=Gamma(link="log"))  
newicis<-seq(0.4,0.9,length=100)  
preds <- predict(glmmean,newdata=data.frame(mean=newicis),type="response")  
lines(newicis,preds)



summary(glmmean)

##   
## Call:  
## glm(formula = crate ~ mean, family = Gamma(link = "log"), data = tagsici)  
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.8450 0.1279 6.605 1.26e-09 \*\*\*  
## mean -1.6559 0.2208 -7.501 1.39e-11 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for Gamma family taken to be 0.06700011)  
##   
## Null deviance: 14.561 on 117 degrees of freedom  
## Residual deviance: 10.922 on 116 degrees of freedom  
## AIC: 27.829  
##   
## Number of Fisher Scoring iterations: 4

#save file for further processing  
write.table(tagsici,file="tagsici.txt")

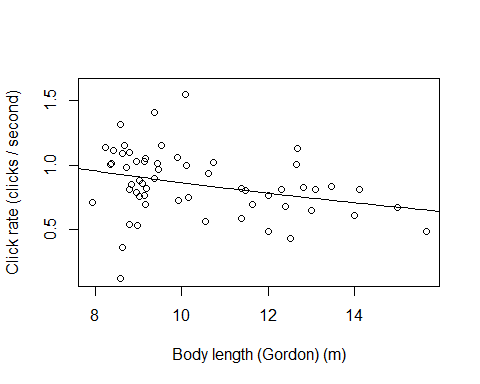
### Inter-Pulse-Interval

Here we look at the Inter-Pulse-Interval (IPI) data

#reading the IPI data  
IPIs <- read\_excel("data/ACCURATE\_spermwhale\_sizefromIPIs.xlsx")  
#merging with the cue rate and ICI data  
tagsipi<-merge(tagsici,IPIs,by="tag",all.x=TRUE,all.y=FALSE)  
  
#not sure what this is not read as a number, forcing it so  
tagsipi$'IPI (ms)' <- as.numeric(tagsipi$'IPI (ms)')  
  
#cleaning up the data.frame which is, by now, full of garbage and repeated measurements from the multiple merges above  
tagsipi <- tagsipi[,c("tag","location.x","sex.x","year.x","duration.x","nclicks.x","crate.x","fyear","locyear","mean","mean0.05","mean0.1","median","mode",'IPI (ms)','Body length (Gordon) (m)','Body length (Growcott) (m)')]  
names(tagsipi) <- c("tag","location","sex","year","duration","nclicks","crate","fyear","locyear","mean","mean0.05","mean0.1","median","mode",'medianIPI','BLGor','BLGro')

and, sure enough, there also seems to be an indication that click rate decreases with size

par(mfrow=c(1,1))  
with(tagsipi,plot(crate~BLGor,ylab="Click rate (clicks / second)",xlab="Body length (Gordon) (m)"))  
glmmeanipi<-glm(crate~BLGor,data=tagsipi,family=Gamma(link="log"))  
newBLGor<-seq(7,16,length=100)  
preds <- predict(glmmeanipi,newdata=data.frame(BLGor=newBLGor),type="response")  
lines(newBLGor,preds)

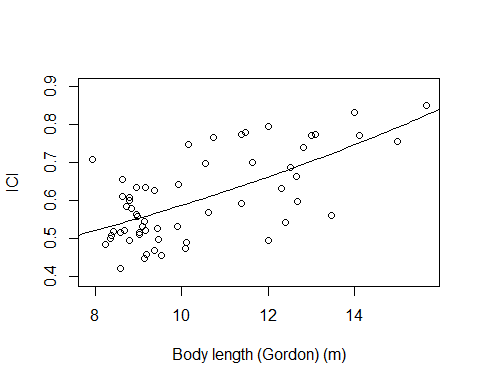


summary(glmmeanipi)

##   
## Call:  
## glm(formula = crate ~ BLGor, family = Gamma(link = "log"), data = tagsipi)  
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.35136 0.20363 1.725 0.0900 .  
## BLGor -0.04950 0.01932 -2.562 0.0131 \*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for Gamma family taken to be 0.07941491)  
##   
## Null deviance: 6.6337 on 57 degrees of freedom  
## Residual deviance: 6.1444 on 56 degrees of freedom  
## (60 observations deleted due to missingness)  
## AIC: 16.521  
##   
## Number of Fisher Scoring iterations: 4

How do ICI and body length relate?

par(mfrow=c(1,1))  
with(tagsipi,plot(mean~BLGor,ylab="ICI",xlab="Body length (Gordon) (m)"))  
glmmeanipi<-glm(mean~BLGor,data=tagsipi,family=Gamma(link="log"))  
newBLGor<-seq(7,16,length=100)  
preds <- predict(glmmeanipi,newdata=data.frame(BLGor=newBLGor),type="response")  
lines(newBLGor,preds)



summary(glmmeanipi)

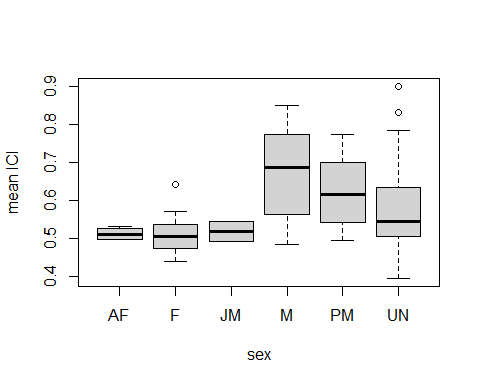
##   
## Call:  
## glm(formula = mean ~ BLGor, family = Gamma(link = "log"), data = tagsipi)  
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -1.133379 0.102947 -11.009 1.25e-15 \*\*\*  
## BLGor 0.060118 0.009769 6.154 8.53e-08 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for Gamma family taken to be 0.02029688)  
##   
## Null deviance: 1.9313 on 57 degrees of freedom  
## Residual deviance: 1.1324 on 56 degrees of freedom  
## (60 observations deleted due to missingness)  
## AIC: -117.84  
##   
## Number of Fisher Scoring iterations: 4

They most certainly do, as expected. The larger the size, the larger the ICI.

### Sex

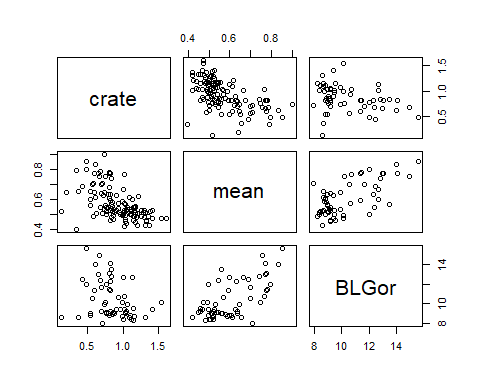
Naturally, if one looks at click rate per sex there is also a corresponding pattern

with(tagsipi,boxplot(mean~sex,ylab="mean ICI"))



which is not surprising since cue rate, ICI and IPI/size are all related

pairs(tagsipi[,c("crate","mean","BLGor")])



### All together now

par(mfrow=c(1,1))  
glmall<-glm(crate~location+fyear+mean+sex,data=tagsipi,family=Gamma(link="log"))  
summary(glmall)

##   
## Call:  
## glm(formula = crate ~ location + fyear + mean + sex, family = Gamma(link = "log"),   
## data = tagsipi)  
##   
## Coefficients: (3 not defined because of singularities)  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.22219 0.39212 0.567 0.572281   
## locationDOMINICA 0.54441 0.26915 2.023 0.045887 \*   
## locationGulf of Mexico 0.45092 0.21127 2.134 0.035363 \*   
## locationKaikoura 0.39422 0.37540 1.050 0.296298   
## locationMediterranean 0.44864 0.25122 1.786 0.077275 .   
## locationNorth Atlantic Delaware 1.34230 0.49088 2.734 0.007442 \*\*   
## locationNorway 0.97001 0.34794 2.788 0.006396 \*\*   
## fyear2002 -0.06958 0.18135 -0.384 0.702062   
## fyear2003 -0.28323 0.21155 -1.339 0.183784   
## fyear2005 0.09089 0.30311 0.300 0.764939   
## fyear2010 0.10274 0.12796 0.803 0.424013   
## fyear2013 NA NA NA NA   
## fyear2014 -0.04908 0.28667 -0.171 0.864425   
## fyear2015 -0.17231 0.27894 -0.618 0.538219   
## fyear2016 -0.01186 0.24518 -0.048 0.961516   
## fyear2018 -0.10275 0.18190 -0.565 0.573485   
## fyear2019 0.16371 0.13991 1.170 0.244844   
## fyear2020 -0.23226 0.13918 -1.669 0.098426 .   
## fyear2021 NA NA NA NA   
## mean -1.30100 0.32699 -3.979 0.000134 \*\*\*  
## sexF 0.02820 0.31892 0.088 0.929717   
## sexJM 0.17638 0.37371 0.472 0.638017   
## sexM -0.69823 0.44783 -1.559 0.122253   
## sexPM NA NA NA NA   
## sexUN 0.29352 0.34675 0.846 0.399384   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for Gamma family taken to be 0.06641151)  
##   
## Null deviance: 14.5610 on 117 degrees of freedom  
## Residual deviance: 8.6302 on 96 degrees of freedom  
## AIC: 39.661  
##   
## Number of Fisher Scoring iterations: 7

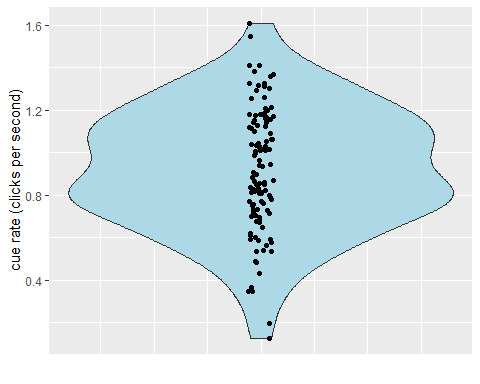
# Reproducing the paper tables and figures

While some of the paper images were already reproduced above, in this section we show the code used to recreate all the figures in the manuscript for easier reproduction.

Images are also (silently, i.e. ) exported as pdfs to be pasted into Word at production stage.

### Figure 1

ggplot(tags,aes(x=1,y=crate),fill="lightblue")+  
theme(axis.ticks.x=element\_blank(),axis.text.x=element\_blank())+geom\_violin(fill="lightblue")+  
geom\_jitter(width=0.03)+ylab("cue rate (clicks per second)")+xlab("")

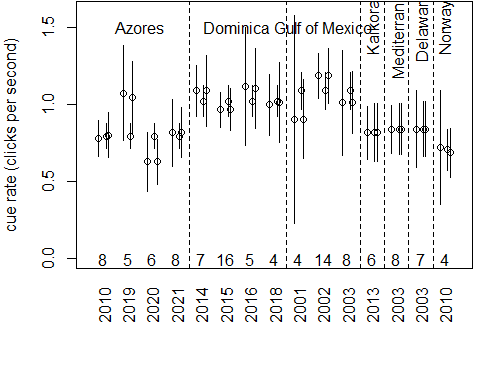


This corresponds to Figure 1 in the main paper

## png   
## 2

### Figure 2

#sort by location, then date  
byly <- byly[order(byly$location,byly$year),]  
#standard average  
par(mfrow=c(1,1),mar=c(6,4,0.1,0.4))  
plot(x=(1:nrow(byly))-0.2,y=byly$ecr,xaxt="n",ylim=c(0,1.6),xlim=c(0.5,nrow(byly)+0.5),xlab="",ylab="cue rate (clicks per second)")  
with(byly,segments(x0=(1:nrow(byly))-0.2,x1=(1:nrow(byly))-0.2,y0=lcl.ecr,y1=ucl.ecr))  
#GLMs  
points(x=(1:nrow(byly))+0.2,y=byly$glm.cr1)  
points(x=(1:nrow(byly))+0.1,y=byly$glm.cr2)  
with(byly,segments(x0=(1:nrow(byly))+0.2,x1=(1:nrow(byly))+0.2,y0=glm.lci.cr1,y1=glm.uci.cr1))  
with(byly,segments(x0=(1:nrow(byly))+0.1,x1=(1:nrow(byly))+0.1,y0=glm.lci.cr2,y1=glm.uci.cr2))  
#draw axis and annotations  
#axis(1, at=(1:nrow(byly))-0.2,byly$location,cex.axis=0.6,las=2)  
axis(1, at=(1:nrow(byly)),byly$year,tick=FALSE,cex.axis=1,las=2,line=0)  
text(x=(1:nrow(byly)),y=0,labels=byly$ntags,cex=1)  
abline(v=c(4.5,8.5,11.5,12.5,13.5,14.5),lty=2)  
myat <- 1.50  
text(2.5,myat,"Azores")  
text(6.5,myat,"Dominica")  
text(10,myat,"Gulf of Mexico")  
text(12,myat,"Kaikora",srt=90)  
text(13,myat,"Mediterranean",srt=90)  
text(14,myat,"Delaware",srt=90)  
text(15,myat,"Norway",srt=90)

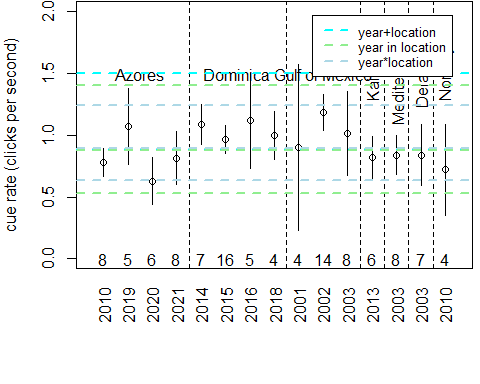


This corresponds to figure 2 in the main paper

## png   
## 2

### Figure 3

#sort by location, then date  
byly <- byly[order(byly$location,byly$year),]  
#standard average  
par(mfrow=c(1,1),mar=c(6,4,0.1,0.4))  
plot(x=(1:nrow(byly)),y=byly$ecr,xaxt="n",ylim=c(0,2),xlim=c(0.5,nrow(byly)+0.5),xlab="",ylab="cue rate (clicks per second)")  
with(byly,segments(x0=(1:nrow(byly)),x1=(1:nrow(byly)),y0=lcl.ecr,y1=ucl.ecr))  
#draw axis and annotations  
#axis(1, at=(1:nrow(byly))-0.2,byly$location,cex.axis=0.6,las=2)  
axis(1, at=(1:nrow(byly)),byly$year,tick=FALSE,cex.axis=1,las=2,line=0)  
text(x=(1:nrow(byly)),y=0,labels=byly$ntags,cex=1)  
abline(v=c(4.5,8.5,11.5,12.5,13.5,14.5),lty=2)  
myat <- 1.50  
text(2.5,myat,"Azores")  
text(6.5,myat,"Dominica")  
text(10,myat,"Gulf of Mexico")  
text(12,myat,"Kaikora",srt=90)  
text(13,myat,"Mediterranean",srt=90)  
text(14,myat,"Delaware",srt=90)  
text(15,myat,"Norway",srt=90)  
abline(h=myquants95CI.Bayes,col="cyan",lty=2,lwd=2)  
abline(h=myquants95CI.Bayes.ynl,col="lightgreen",lty=2,lwd=2)  
abline(h=myquants95CI.Bayes.1re,col="lightblue",lty=2,lwd=2)  
legend("topright",legend=c("year+location","year in location","year\*location"),lty=2,lwd=c(2,2,2),col=c("cyan","lightgreen","lightblue"),inset=0.05,cex=0.8,bg="white")



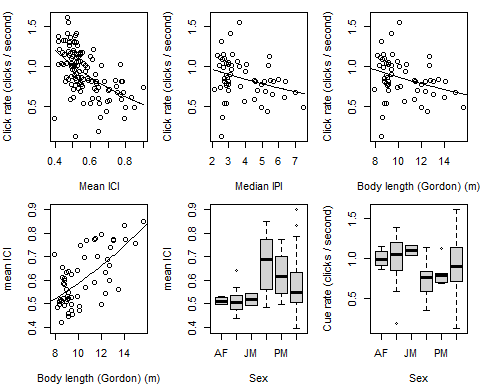
This corresponds to figure 3 in the main paper

## png   
## 2

### Figure 4

Below we reproduce figure 4 in the manuscript

par(mfrow=c(2,3),mar=c(4,4,1,1))  
# click rate as a function of mean ICI  
with(tagsipi,plot(crate~mean,ylab="Click rate (clicks / second)",xlab="Mean ICI"))  
glmmean<-glm(crate~mean,data=tagsipi,family=Gamma(link="log"))  
newicis<-seq(0.4,0.9,length=100)  
preds <- predict(glmmean,newdata=data.frame(mean=newicis),type="response")  
lines(newicis,preds)  
  
# click rate as a function of median IPI  
with(tagsipi,plot(crate~medianIPI,ylab="Click rate (clicks / second)",xlab="Median IPI"))  
glmmedian<-glm(crate~medianIPI,data=tagsipi,family=Gamma(link="log"))  
newipis<-seq(2.1,7.5,length=100)  
preds <- predict(glmmedian,newdata=data.frame(medianIPI=newipis),type="response")  
lines(newipis,preds)  
# click rate as a function of animal size  
with(tagsipi,plot(crate~BLGor,ylab="Click rate (clicks / second)",xlab="Body length (Gordon) (m)"))  
glmBLGor<-glm(crate~BLGor,data=tagsipi,family=Gamma(link="log"))  
newBLGor<-seq(7,16,length=100)  
preds <- predict(glmBLGor,newdata=data.frame(BLGor=newBLGor),type="response")  
lines(newBLGor,preds)  
# relation between ICI and IPI (i.e. size)  
with(tagsipi,plot(mean~BLGor,ylab="mean ICI",xlab="Body length (Gordon) (m)"))  
glmmeansize<-glm(mean~BLGor,data=tagsipi,family=Gamma(link="log"))  
newBLGor<-seq(7,16,length=100)  
preds <- predict(glmmeansize,newdata=data.frame(BLGor=newBLGor),type="response")  
lines(newBLGor,preds)  
# ICI by sex  
with(tagsipi,boxplot(mean~sex,ylab="mean ICI",xlab="Sex"))  
# cue rate by sex  
with(tagsipi,boxplot(crate~sex,ylab="Cue rate (clicks / second)",xlab="Sex"))



This corresponds to figure 4 in the main paper

## png   
## 2

### Summary tables

The following tables can be used to reproduce all the paper’s analysis without the need for all the data wrangling that went into producing it from multiple data sources:

Cue rates per tag

kable(tagsipi[,c(1:8)],digits=2,caption="A summary table with the cue rate information")

A summary table with the cue rate information

| tag | location | sex | year | duration | nclicks | crate | fyear |
| --- | --- | --- | --- | --- | --- | --- | --- |
| pm10\_211b | Azores | UN | 2010 | 47819 | 43037 | 0.90 | 2010 |
| pm10\_222a | Azores | UN | 2010 | 22249 | 18037 | 0.81 | 2010 |
| pm10\_222b | Azores | UN | 2010 | 53744 | 28762 | 0.54 | 2010 |
| pm10\_226a | Azores | UN | 2010 | 61415 | 52345 | 0.85 | 2010 |
| pm10\_227b | Azores | UN | 2010 | 1399 | 1027 | 0.73 | 2010 |
| pm10\_228a | Azores | UN | 2010 | 71632 | 50026 | 0.70 | 2010 |
| pm10\_228b | Azores | UN | 2010 | 52105 | 38085 | 0.73 | 2010 |
| pm10\_230a | Azores | UN | 2010 | 47131 | 46455 | 0.99 | 2010 |
| pm13\_050a | Kaikoura | PM | 2013 | 5874 | 4771 | 0.81 | 2013 |
| pm13\_051a | Kaikoura | PM | 2013 | 10972 | 8940 | 0.81 | 2013 |
| pm13\_056a | Kaikoura | PM | 2013 | 28837 | 22162 | 0.77 | 2013 |
| pm13\_057a | Kaikoura | PM | 2013 | 6775 | 7634 | 1.13 | 2013 |
| pm13\_061a | Kaikoura | PM | 2013 | 22007 | 15348 | 0.70 | 2013 |
| pm13\_061b | Kaikoura | PM | 2013 | 17519 | 11932 | 0.68 | 2013 |
| sw01\_200a | Gulf of Mexico | F | 2001 | 30131 | 26214 | 0.87 | 2001 |
| sw01\_204a | Gulf of Mexico | F | 2001 | 18773 | 19558 | 1.04 | 2001 |
| sw01\_208b | Gulf of Mexico | M | 2001 | 8022 | 2791 | 0.35 | 2001 |
| sw01\_209c | Gulf of Mexico | F | 2001 | 10568 | 14358 | 1.36 | 2001 |
| sw01\_265a | Mediterranean | UN | 2003 | 11857 | 8016 | 0.68 | 2003 |
| sw02\_235c | Gulf of Mexico | F | 2002 | 6540 | 6181 | 0.95 | 2002 |
| sw02\_237a | Gulf of Mexico | UN | 2002 | 10383 | 7584 | 0.73 | 2002 |
| sw02\_238a | Gulf of Mexico | F | 2002 | 13462 | 17838 | 1.33 | 2002 |
| sw02\_238b | Gulf of Mexico | UN | 2002 | 7550 | 8853 | 1.17 | 2002 |
| sw02\_239a | Gulf of Mexico | F | 2002 | 44469 | 52441 | 1.18 | 2002 |
| sw02\_239b | Gulf of Mexico | UN | 2002 | 2539 | 3929 | 1.55 | 2002 |
| sw02\_240a | Gulf of Mexico | UN | 2002 | 2177 | 3069 | 1.41 | 2002 |
| sw02\_240c | Gulf of Mexico | UN | 2002 | 17847 | 20945 | 1.17 | 2002 |
| sw02\_248a | Gulf of Mexico | UN | 2002 | 2737 | 4407 | 1.61 | 2002 |
| sw02\_249a | Gulf of Mexico | UN | 2002 | 5904 | 7758 | 1.31 | 2002 |
| sw02\_253a | Gulf of Mexico | UN | 2002 | 12125 | 14366 | 1.18 | 2002 |
| sw02\_254a | Gulf of Mexico | F | 2002 | 41142 | 47351 | 1.15 | 2002 |
| sw02\_254b | Gulf of Mexico | F | 2002 | 42275 | 34540 | 0.82 | 2002 |
| sw02\_254c | Gulf of Mexico | UN | 2002 | 42180 | 43407 | 1.03 | 2002 |
| sw03\_156a | Gulf of Mexico | UN | 2003 | 17329 | 22542 | 1.30 | 2003 |
| sw03\_162a | Gulf of Mexico | UN | 2003 | 2970 | 4190 | 1.41 | 2003 |
| sw03\_163a | Gulf of Mexico | UN | 2003 | 22802 | 2867 | 0.13 | 2003 |
| sw03\_164a | Gulf of Mexico | UN | 2003 | 48144 | 48702 | 1.01 | 2003 |
| sw03\_165a | Gulf of Mexico | UN | 2003 | 47260 | 48091 | 1.02 | 2003 |
| sw03\_165b | Gulf of Mexico | UN | 2003 | 58773 | 77272 | 1.31 | 2003 |
| sw03\_167a | Gulf of Mexico | UN | 2003 | 9203 | 7629 | 0.83 | 2003 |
| sw03\_173b | Gulf of Mexico | UN | 2003 | 3642 | 3966 | 1.09 | 2003 |
| sw03\_197a | North Atlantic Delaware | M | 2003 | 6825 | 5561 | 0.81 | 2003 |
| sw03\_197b | North Atlantic Delaware | M | 2003 | 8447 | 3650 | 0.43 | 2003 |
| sw03\_201b | North Atlantic Delaware | M | 2003 | 9967 | 5922 | 0.59 | 2003 |
| sw03\_202a | North Atlantic Delaware | M | 2003 | 3093 | 2355 | 0.76 | 2003 |
| sw03\_206a | North Atlantic Delaware | M | 2003 | 13426 | 15323 | 1.14 | 2003 |
| sw03\_206c | North Atlantic Delaware | M | 2003 | 8846 | 9134 | 1.03 | 2003 |
| sw03\_207a | North Atlantic Delaware | M | 2003 | 14327 | 15787 | 1.10 | 2003 |
| sw03\_247a | Mediterranean | UN | 2003 | 13106 | 7371 | 0.56 | 2003 |
| sw03\_249a | Mediterranean | UN | 2003 | 8496 | 6411 | 0.75 | 2003 |
| sw03\_249b | Mediterranean | UN | 2003 | 1006 | 1031 | 1.02 | 2003 |
| sw03\_249c | Mediterranean | UN | 2003 | 39195 | 27938 | 0.71 | 2003 |
| sw03\_253a | Mediterranean | UN | 2003 | 21840 | 21729 | 0.99 | 2003 |
| sw03\_253b | Mediterranean | UN | 2003 | 42393 | 39568 | 0.93 | 2003 |
| sw03\_253c | Mediterranean | UN | 2003 | 17782 | 18898 | 1.06 | 2003 |
| sw05\_196a | Norway | M | 2005 | 77806 | 52271 | 0.67 | 2005 |
| sw09\_153a | Norway | M | 2010 | 31208 | 31423 | 1.01 | 2010 |
| sw10\_147a | Norway | M | 2010 | 56725 | 27495 | 0.48 | 2010 |
| sw10\_149a | Norway | M | 2010 | 57324 | 46243 | 0.81 | 2010 |
| sw10\_150a | Norway | M | 2010 | 53019 | 31333 | 0.59 | 2010 |
| sw14\_092a | DOMINICA | F | 2014 | 6941 | 7929 | 1.14 | 2014 |
| sw14\_106a | DOMINICA | F | 2014 | 18030 | 22651 | 1.26 | 2014 |
| sw14\_106b | DOMINICA | F | 2014 | 32045 | 27223 | 0.85 | 2014 |
| sw14\_114a | DOMINICA | F | 2014 | 24849 | 25834 | 1.04 | 2014 |
| sw14\_114b | DOMINICA | F | 2014 | 23424 | 30254 | 1.29 | 2014 |
| sw14\_119a | DOMINICA | F | 2014 | 29112 | 34328 | 1.18 | 2014 |
| sw14\_119b | DOMINICA | F | 2014 | 26550 | 22948 | 0.86 | 2014 |
| sw15\_061a | DOMINICA | F | 2015 | 14804 | 17392 | 1.17 | 2015 |
| sw15\_061b | DOMINICA | F | 2015 | 26392 | 22588 | 0.86 | 2015 |
| sw15\_061c | DOMINICA | F | 2015 | 21238 | 23865 | 1.12 | 2015 |
| sw15\_078a | DOMINICA | F | 2015 | 26280 | 33166 | 1.26 | 2015 |
| sw15\_078b | DOMINICA | F | 2015 | 9353 | 12109 | 1.29 | 2015 |
| sw15\_078c | DOMINICA | JM | 2015 | 28730 | 29605 | 1.03 | 2015 |
| sw15\_078d | DOMINICA | JM | 2015 | 83449 | 97145 | 1.16 | 2015 |
| sw15\_085a | DOMINICA | F | 2015 | 31513 | 22288 | 0.71 | 2015 |
| sw15\_090a | DOMINICA | F | 2015 | 34775 | 27744 | 0.80 | 2015 |
| sw15\_090b | DOMINICA | F | 2015 | 32316 | 25286 | 0.78 | 2015 |
| sw15\_090c | DOMINICA | F | 2015 | 26702 | 20133 | 0.75 | 2015 |
| sw15\_091a | DOMINICA | F | 2015 | 27180 | 25514 | 0.94 | 2015 |
| sw15\_091b | DOMINICA | F | 2015 | 23770 | 24167 | 1.02 | 2015 |
| sw15\_097a | DOMINICA | F | 2015 | 22696 | 25448 | 1.12 | 2015 |
| sw15\_100a | DOMINICA | F | 2015 | 18575 | 10916 | 0.59 | 2015 |
| sw15\_100b | DOMINICA | F | 2015 | 33932 | 28206 | 0.83 | 2015 |
| sw16\_103a | DOMINICA | F | 2016 | 21807 | 23217 | 1.06 | 2016 |
| sw16\_115b | DOMINICA | F | 2016 | 3331 | 4602 | 1.38 | 2016 |
| sw16\_131a | Norway | M | 2016 | 42992 | 27885 | 0.65 | 2016 |
| sw16\_133a | DOMINICA | F | 2016 | 22144 | 29327 | 1.32 | 2016 |
| sw16\_134a | Norway | M | 2016 | 20379 | 9912 | 0.49 | 2016 |
| sw16\_134ad | DOMINICA | F | 2016 | 13210 | 15835 | 1.20 | 2016 |
| sw17\_119c | DOMINICA | UN | 2016 | 17794 | 10905 | 0.61 | 2016 |
| sw18\_083a | DOMINICA | AF | 2018 | 28873 | 29147 | 1.01 | 2018 |
| sw18\_140a | DOMINICA | AF | 2018 | 33231 | 38359 | 1.15 | 2018 |
| sw18\_140b | DOMINICA | AF | 2018 | 36336 | 35061 | 0.96 | 2018 |
| sw18\_143a | DOMINICA | AF | 2018 | 45105 | 38741 | 0.86 | 2018 |
| sw18\_267a | Azores | F | 2018 | 22063 | 4291 | 0.19 | 2018 |
| sw18\_281a | Azores | UN | 2018 | 2693 | 3262 | 1.21 | 2018 |
| sw18\_292a | Azores | UN | 2018 | 23684 | 12670 | 0.53 | 2018 |
| sw19\_088a | Azores | UN | 2019 | 83995 | 68203 | 0.81 | 2019 |
| sw19\_137a | Azores | UN | 2019 | 34559 | 28394 | 0.82 | 2019 |
| sw19\_158a | Azores | UN | 2019 | 27438 | 37537 | 1.37 | 2019 |
| sw19\_160a | Azores | UN | 2019 | 92812 | 107607 | 1.16 | 2019 |
| sw19\_163a | Azores | UN | 2019 | 50040 | 60501 | 1.21 | 2019 |
| sw19\_253b | Norway | M | 2019 | 33659 | 28126 | 0.84 | 2019 |
| sw19\_259a | Norway | M | 2019 | 21520 | 17655 | 0.82 | 2019 |
| sw20\_161a | Azores | UN | 2020 | 70360 | 63822 | 0.91 | 2020 |
| sw20\_184a | Azores | UN | 2020 | 42912 | 24825 | 0.58 | 2020 |
| sw20\_189a | Azores | UN | 2020 | 19411 | 6753 | 0.35 | 2020 |
| sw20\_191a | Azores | UN | 2020 | 517 | 372 | 0.72 | 2020 |
| sw20\_192a | Azores | UN | 2020 | 50980 | 31509 | 0.62 | 2020 |
| sw20\_225a | Azores | UN | 2020 | 63339 | 38204 | 0.60 | 2020 |
| sw21\_196a | Azores | UN | 2021 | 47905 | 53284 | 1.11 | 2021 |
| sw21\_209a | Azores | UN | 2021 | 39267 | 41351 | 1.05 | 2021 |
| sw21\_211a | Azores | UN | 2021 | 60044 | 47354 | 0.79 | 2021 |
| sw21\_215a | Azores | UN | 2021 | 48634 | 37430 | 0.77 | 2021 |
| sw21\_215b | Azores | UN | 2021 | 30478 | 26953 | 0.88 | 2021 |
| sw21\_221a | Azores | UN | 2021 | 32993 | 33452 | 1.01 | 2021 |
| sw21\_230a | Azores | UN | 2021 | 48092 | 25953 | 0.54 | 2021 |
| sw21\_242a | Azores | UN | 2021 | 31680 | 11674 | 0.37 | 2021 |

ICI per tag (mean is the mean, mean0.05 is the trimmed mean for 5% of the data, mean0.1 is the trimmed mean for 10% of the data, median and mode)

kable(tagsipi[,c(1,2,3,4,10:14)],digits=2,caption="A summary table with the ICI information")

A summary table with the ICI information

| tag | location | sex | year | mean | mean0.05 | mean0.1 | median | mode |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| pm10\_211b | Azores | UN | 2010 | 0.63 | 0.61 | 0.61 | 0.61 | 0.64 |
| pm10\_222a | Azores | UN | 2010 | 0.61 | 0.59 | 0.59 | 0.58 | 0.58 |
| pm10\_222b | Azores | UN | 2010 | 0.56 | 0.54 | 0.53 | 0.52 | 0.51 |
| pm10\_226a | Azores | UN | 2010 | 0.58 | 0.57 | 0.56 | 0.56 | 0.53 |
| pm10\_227b | Azores | UN | 2010 | 0.90 | 0.80 | 0.75 | 0.58 | 0.54 |
| pm10\_228a | Azores | UN | 2010 | 0.64 | 0.61 | 0.61 | 0.59 | 0.58 |
| pm10\_228b | Azores | UN | 2010 | 0.64 | 0.61 | 0.60 | 0.59 | 0.62 |
| pm10\_230a | Azores | UN | 2010 | 0.59 | 0.57 | 0.56 | 0.56 | 0.47 |
| pm13\_050a | Kaikoura | PM | 2013 | 0.77 | 0.76 | 0.76 | 0.76 | 0.74 |
| pm13\_051a | Kaikoura | PM | 2013 | 0.63 | 0.60 | 0.59 | 0.55 | 0.44 |
| pm13\_056a | Kaikoura | PM | 2013 | 0.50 | 0.46 | 0.45 | 0.43 | 0.36 |
| pm13\_057a | Kaikoura | PM | 2013 | 0.60 | 0.57 | 0.57 | 0.54 | 0.46 |
| pm13\_061a | Kaikoura | PM | 2013 | 0.70 | 0.67 | 0.66 | 0.64 | 0.64 |
| pm13\_061b | Kaikoura | PM | 2013 | 0.54 | 0.51 | 0.48 | 0.43 | 0.26 |
| sw01\_200a | Gulf of Mexico | F | 2001 | 0.49 | 0.48 | 0.48 | 0.47 | 0.44 |
| sw01\_204a | Gulf of Mexico | F | 2001 | 0.44 | 0.42 | 0.42 | 0.41 | 0.39 |
| sw01\_208b | Gulf of Mexico | M | 2001 | 0.79 | 0.78 | 0.78 | 0.72 | 0.63 |
| sw01\_209c | Gulf of Mexico | F | 2001 | 0.49 | 0.48 | 0.47 | 0.47 | 0.46 |
| sw01\_265a | Mediterranean | UN | 2003 | 0.79 | 0.78 | 0.79 | 0.80 | 0.83 |
| sw02\_235c | Gulf of Mexico | F | 2002 | 0.51 | 0.50 | 0.49 | 0.48 | 0.47 |
| sw02\_237a | Gulf of Mexico | UN | 2002 | 0.53 | 0.52 | 0.51 | 0.51 | 0.51 |
| sw02\_238a | Gulf of Mexico | F | 2002 | 0.45 | 0.44 | 0.43 | 0.42 | 0.41 |
| sw02\_238b | Gulf of Mexico | UN | 2002 | 0.54 | 0.54 | 0.53 | 0.53 | 0.53 |
| sw02\_239a | Gulf of Mexico | F | 2002 | 0.44 | 0.43 | 0.43 | 0.43 | 0.42 |
| sw02\_239b | Gulf of Mexico | UN | 2002 | 0.47 | 0.47 | 0.47 | 0.47 | 0.46 |
| sw02\_240a | Gulf of Mexico | UN | 2002 | 0.52 | 0.51 | 0.51 | 0.50 | 0.47 |
| sw02\_240c | Gulf of Mexico | UN | 2002 | 0.55 | 0.54 | 0.53 | 0.51 | 0.46 |
| sw02\_248a | Gulf of Mexico | UN | 2002 | 0.47 | 0.47 | 0.46 | 0.46 | 0.44 |
| sw02\_249a | Gulf of Mexico | UN | 2002 | 0.51 | 0.50 | 0.50 | 0.49 | 0.48 |
| sw02\_253a | Gulf of Mexico | UN | 2002 | 0.48 | 0.47 | 0.47 | 0.46 | 0.45 |
| sw02\_254a | Gulf of Mexico | F | 2002 | 0.46 | 0.45 | 0.45 | 0.44 | 0.43 |
| sw02\_254b | Gulf of Mexico | F | 2002 | 0.46 | 0.45 | 0.44 | 0.44 | 0.42 |
| sw02\_254c | Gulf of Mexico | UN | 2002 | 0.45 | 0.44 | 0.44 | 0.43 | 0.42 |
| sw03\_156a | Gulf of Mexico | UN | 2003 | 0.48 | 0.47 | 0.47 | 0.46 | 0.45 |
| sw03\_162a | Gulf of Mexico | UN | 2003 | 0.47 | 0.46 | 0.46 | 0.45 | 0.45 |
| sw03\_163a | Gulf of Mexico | UN | 2003 | 0.52 | 0.50 | 0.50 | 0.49 | 0.47 |
| sw03\_164a | Gulf of Mexico | UN | 2003 | 0.51 | 0.50 | 0.49 | 0.49 | 0.48 |
| sw03\_165a | Gulf of Mexico | UN | 2003 | 0.42 | 0.40 | 0.40 | 0.39 | 0.39 |
| sw03\_165b | Gulf of Mexico | UN | 2003 | 0.42 | 0.41 | 0.41 | 0.41 | 0.40 |
| sw03\_167a | Gulf of Mexico | UN | 2003 | 0.74 | 0.72 | 0.72 | 0.69 | 0.64 |
| sw03\_173b | Gulf of Mexico | UN | 2003 | 0.61 | 0.61 | 0.60 | 0.59 | 0.58 |
| sw03\_197a | North Atlantic Delaware | M | 2003 | 0.77 | 0.76 | 0.75 | 0.75 | 0.75 |
| sw03\_197b | North Atlantic Delaware | M | 2003 | 0.69 | 0.67 | 0.66 | 0.64 | 0.60 |
| sw03\_201b | North Atlantic Delaware | M | 2003 | 0.65 | 0.63 | 0.63 | 0.63 | 0.65 |
| sw03\_202a | North Atlantic Delaware | M | 2003 | 0.52 | 0.50 | 0.50 | 0.50 | 0.45 |
| sw03\_206a | North Atlantic Delaware | M | 2003 | 0.48 | 0.47 | 0.47 | 0.47 | 0.53 |
| sw03\_206c | North Atlantic Delaware | M | 2003 | 0.56 | 0.54 | 0.54 | 0.55 | 0.57 |
| sw03\_207a | North Atlantic Delaware | M | 2003 | 0.49 | 0.48 | 0.48 | 0.48 | 0.51 |
| sw03\_247a | Mediterranean | UN | 2003 | 0.70 | 0.68 | 0.68 | 0.69 | 0.78 |
| sw03\_249a | Mediterranean | UN | 2003 | 0.75 | 0.73 | 0.73 | 0.74 | 0.83 |
| sw03\_249b | Mediterranean | UN | 2003 | 0.77 | 0.75 | 0.73 | 0.74 | 0.84 |
| sw03\_249c | Mediterranean | UN | 2003 | 0.71 | 0.70 | 0.71 | 0.72 | 0.77 |
| sw03\_253a | Mediterranean | UN | 2003 | 0.49 | 0.46 | 0.46 | 0.45 | 0.44 |
| sw03\_253b | Mediterranean | UN | 2003 | 0.57 | 0.55 | 0.55 | 0.53 | 0.48 |
| sw03\_253c | Mediterranean | UN | 2003 | 0.53 | 0.50 | 0.49 | 0.47 | 0.44 |
| sw05\_196a | Norway | M | 2005 | 0.76 | 0.75 | 0.75 | 0.75 | 0.73 |
| sw09\_153a | Norway | M | 2010 | 0.66 | 0.65 | 0.65 | 0.64 | 0.63 |
| sw10\_147a | Norway | M | 2010 | 0.80 | 0.79 | 0.79 | 0.78 | 0.77 |
| sw10\_149a | Norway | M | 2010 | 0.78 | 0.77 | 0.77 | 0.77 | 0.77 |
| sw10\_150a | Norway | M | 2010 | 0.77 | 0.76 | 0.76 | 0.77 | 0.78 |
| sw14\_092a | DOMINICA | F | 2014 | 0.54 | 0.52 | 0.52 | 0.52 | 0.52 |
| sw14\_106a | DOMINICA | F | 2014 | 0.53 | 0.51 | 0.51 | 0.50 | 0.49 |
| sw14\_106b | DOMINICA | F | 2014 | 0.52 | 0.50 | 0.50 | 0.50 | 0.50 |
| sw14\_114a | DOMINICA | F | 2014 | 0.47 | 0.45 | 0.45 | 0.44 | 0.40 |
| sw14\_114b | DOMINICA | F | 2014 | 0.51 | 0.49 | 0.49 | 0.48 | 0.46 |
| sw14\_119a | DOMINICA | F | 2014 | 0.51 | 0.49 | 0.49 | 0.47 | 0.46 |
| sw14\_119b | DOMINICA | F | 2014 | 0.47 | 0.45 | 0.44 | 0.43 | 0.42 |
| sw15\_061a | DOMINICA | F | 2015 | 0.50 | 0.48 | 0.47 | 0.46 | 0.45 |
| sw15\_061b | DOMINICA | F | 2015 | 0.55 | 0.53 | 0.52 | 0.52 | 0.49 |
| sw15\_061c | DOMINICA | F | 2015 | 0.50 | 0.48 | 0.47 | 0.47 | 0.46 |
| sw15\_078a | DOMINICA | F | 2015 | 0.48 | 0.47 | 0.46 | 0.45 | 0.42 |
| sw15\_078b | DOMINICA | F | 2015 | 0.52 | 0.51 | 0.51 | 0.50 | 0.47 |
| sw15\_078c | DOMINICA | JM | 2015 | 0.49 | 0.48 | 0.47 | 0.46 | 0.43 |
| sw15\_078d | DOMINICA | JM | 2015 | 0.55 | 0.53 | 0.53 | 0.52 | 0.47 |
| sw15\_085a | DOMINICA | F | 2015 | 0.50 | 0.49 | 0.48 | 0.48 | 0.42 |
| sw15\_090a | DOMINICA | F | 2015 | 0.54 | 0.53 | 0.52 | 0.52 | 0.53 |
| sw15\_090b | DOMINICA | F | 2015 | 0.54 | 0.52 | 0.52 | 0.52 | 0.51 |
| sw15\_090c | DOMINICA | F | 2015 | 0.56 | 0.54 | 0.54 | 0.54 | 0.52 |
| sw15\_091a | DOMINICA | F | 2015 | 0.54 | 0.53 | 0.52 | 0.52 | 0.52 |
| sw15\_091b | DOMINICA | F | 2015 | 0.57 | 0.55 | 0.55 | 0.55 | 0.56 |
| sw15\_097a | DOMINICA | F | 2015 | 0.53 | 0.51 | 0.51 | 0.51 | 0.50 |
| sw15\_100a | DOMINICA | F | 2015 | 0.49 | 0.47 | 0.46 | 0.45 | 0.41 |
| sw15\_100b | DOMINICA | F | 2015 | 0.47 | 0.46 | 0.46 | 0.45 | 0.39 |
| sw16\_103a | DOMINICA | F | 2016 | 0.54 | 0.53 | 0.53 | 0.53 | 0.53 |
| sw16\_115b | DOMINICA | F | 2016 | 0.51 | 0.50 | 0.49 | 0.48 | 0.48 |
| sw16\_131a | Norway | M | 2016 | 0.77 | 0.75 | 0.75 | 0.75 | 0.57 |
| sw16\_133a | DOMINICA | F | 2016 | 0.46 | 0.44 | 0.44 | 0.43 | 0.42 |
| sw16\_134a | Norway | M | 2016 | 0.85 | 0.85 | 0.85 | 0.86 | 0.87 |
| sw16\_134ad | DOMINICA | F | 2016 | 0.50 | 0.48 | 0.48 | 0.47 | 0.43 |
| sw17\_119c | DOMINICA | UN | 2016 | 0.83 | 0.79 | 0.78 | 0.77 | 0.77 |
| sw18\_083a | DOMINICA | AF | 2018 | 0.50 | 0.48 | 0.47 | 0.47 | 0.46 |
| sw18\_140a | DOMINICA | AF | 2018 | 0.52 | 0.50 | 0.49 | 0.49 | 0.48 |
| sw18\_140b | DOMINICA | AF | 2018 | 0.50 | 0.48 | 0.47 | 0.46 | 0.42 |
| sw18\_143a | DOMINICA | AF | 2018 | 0.53 | 0.52 | 0.51 | 0.50 | 0.49 |
| sw18\_267a | Azores | F | 2018 | 0.64 | 0.61 | 0.61 | 0.61 | 0.62 |
| sw18\_281a | Azores | UN | 2018 | 0.62 | 0.58 | 0.57 | 0.56 | 0.57 |
| sw18\_292a | Azores | UN | 2018 | 0.65 | 0.63 | 0.62 | 0.60 | 0.60 |
| sw19\_088a | Azores | UN | 2019 | 0.68 | 0.66 | 0.66 | 0.67 | 0.68 |
| sw19\_137a | Azores | UN | 2019 | 0.52 | 0.50 | 0.50 | 0.49 | 0.42 |
| sw19\_158a | Azores | UN | 2019 | 0.42 | 0.41 | 0.41 | 0.40 | 0.39 |
| sw19\_160a | Azores | UN | 2019 | 0.49 | 0.47 | 0.47 | 0.45 | 0.43 |
| sw19\_163a | Azores | UN | 2019 | 0.42 | 0.41 | 0.40 | 0.39 | 0.38 |
| sw19\_253b | Norway | M | 2019 | 0.56 | 0.54 | 0.53 | 0.50 | 0.42 |
| sw19\_259a | Norway | M | 2019 | 0.59 | 0.57 | 0.57 | 0.55 | 0.50 |
| sw20\_161a | Azores | UN | 2020 | 0.52 | 0.50 | 0.50 | 0.49 | 0.37 |
| sw20\_184a | Azores | UN | 2020 | 0.71 | 0.69 | 0.69 | 0.69 | 0.62 |
| sw20\_189a | Azores | UN | 2020 | 0.40 | 0.38 | 0.37 | 0.38 | 0.39 |
| sw20\_191a | Azores | UN | 2020 | 0.62 | 0.58 | 0.58 | 0.58 | 0.46 |
| sw20\_192a | Azores | UN | 2020 | 0.50 | 0.47 | 0.46 | 0.42 | 0.38 |
| sw20\_225a | Azores | UN | 2020 | 0.61 | 0.59 | 0.58 | 0.57 | 0.52 |
| sw21\_196a | Azores | UN | 2021 | 0.52 | 0.50 | 0.49 | 0.48 | 0.45 |
| sw21\_209a | Azores | UN | 2021 | 0.52 | 0.50 | 0.49 | 0.46 | 0.44 |
| sw21\_211a | Azores | UN | 2021 | 0.63 | 0.61 | 0.60 | 0.59 | 0.58 |
| sw21\_215a | Azores | UN | 2021 | 0.55 | 0.52 | 0.52 | 0.51 | 0.47 |
| sw21\_215b | Azores | UN | 2021 | 0.51 | 0.49 | 0.49 | 0.47 | 0.43 |
| sw21\_221a | Azores | UN | 2021 | 0.53 | 0.50 | 0.50 | 0.49 | 0.38 |
| sw21\_230a | Azores | UN | 2021 | 0.60 | 0.58 | 0.58 | 0.57 | 0.55 |
| sw21\_242a | Azores | UN | 2021 | 0.65 | 0.63 | 0.63 | 0.62 | 0.61 |

IPI and body sizes, per tag

kable(tagsipi[,c(1,2,3,4,10:14)],digits=2,caption="A summary table with the cue rate IPI and animal size information")

A summary table with the cue rate IPI and animal size information

| tag | location | sex | year | mean | mean0.05 | mean0.1 | median | mode |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| pm10\_211b | Azores | UN | 2010 | 0.63 | 0.61 | 0.61 | 0.61 | 0.64 |
| pm10\_222a | Azores | UN | 2010 | 0.61 | 0.59 | 0.59 | 0.58 | 0.58 |
| pm10\_222b | Azores | UN | 2010 | 0.56 | 0.54 | 0.53 | 0.52 | 0.51 |
| pm10\_226a | Azores | UN | 2010 | 0.58 | 0.57 | 0.56 | 0.56 | 0.53 |
| pm10\_227b | Azores | UN | 2010 | 0.90 | 0.80 | 0.75 | 0.58 | 0.54 |
| pm10\_228a | Azores | UN | 2010 | 0.64 | 0.61 | 0.61 | 0.59 | 0.58 |
| pm10\_228b | Azores | UN | 2010 | 0.64 | 0.61 | 0.60 | 0.59 | 0.62 |
| pm10\_230a | Azores | UN | 2010 | 0.59 | 0.57 | 0.56 | 0.56 | 0.47 |
| pm13\_050a | Kaikoura | PM | 2013 | 0.77 | 0.76 | 0.76 | 0.76 | 0.74 |
| pm13\_051a | Kaikoura | PM | 2013 | 0.63 | 0.60 | 0.59 | 0.55 | 0.44 |
| pm13\_056a | Kaikoura | PM | 2013 | 0.50 | 0.46 | 0.45 | 0.43 | 0.36 |
| pm13\_057a | Kaikoura | PM | 2013 | 0.60 | 0.57 | 0.57 | 0.54 | 0.46 |
| pm13\_061a | Kaikoura | PM | 2013 | 0.70 | 0.67 | 0.66 | 0.64 | 0.64 |
| pm13\_061b | Kaikoura | PM | 2013 | 0.54 | 0.51 | 0.48 | 0.43 | 0.26 |
| sw01\_200a | Gulf of Mexico | F | 2001 | 0.49 | 0.48 | 0.48 | 0.47 | 0.44 |
| sw01\_204a | Gulf of Mexico | F | 2001 | 0.44 | 0.42 | 0.42 | 0.41 | 0.39 |
| sw01\_208b | Gulf of Mexico | M | 2001 | 0.79 | 0.78 | 0.78 | 0.72 | 0.63 |
| sw01\_209c | Gulf of Mexico | F | 2001 | 0.49 | 0.48 | 0.47 | 0.47 | 0.46 |
| sw01\_265a | Mediterranean | UN | 2003 | 0.79 | 0.78 | 0.79 | 0.80 | 0.83 |
| sw02\_235c | Gulf of Mexico | F | 2002 | 0.51 | 0.50 | 0.49 | 0.48 | 0.47 |
| sw02\_237a | Gulf of Mexico | UN | 2002 | 0.53 | 0.52 | 0.51 | 0.51 | 0.51 |
| sw02\_238a | Gulf of Mexico | F | 2002 | 0.45 | 0.44 | 0.43 | 0.42 | 0.41 |
| sw02\_238b | Gulf of Mexico | UN | 2002 | 0.54 | 0.54 | 0.53 | 0.53 | 0.53 |
| sw02\_239a | Gulf of Mexico | F | 2002 | 0.44 | 0.43 | 0.43 | 0.43 | 0.42 |
| sw02\_239b | Gulf of Mexico | UN | 2002 | 0.47 | 0.47 | 0.47 | 0.47 | 0.46 |
| sw02\_240a | Gulf of Mexico | UN | 2002 | 0.52 | 0.51 | 0.51 | 0.50 | 0.47 |
| sw02\_240c | Gulf of Mexico | UN | 2002 | 0.55 | 0.54 | 0.53 | 0.51 | 0.46 |
| sw02\_248a | Gulf of Mexico | UN | 2002 | 0.47 | 0.47 | 0.46 | 0.46 | 0.44 |
| sw02\_249a | Gulf of Mexico | UN | 2002 | 0.51 | 0.50 | 0.50 | 0.49 | 0.48 |
| sw02\_253a | Gulf of Mexico | UN | 2002 | 0.48 | 0.47 | 0.47 | 0.46 | 0.45 |
| sw02\_254a | Gulf of Mexico | F | 2002 | 0.46 | 0.45 | 0.45 | 0.44 | 0.43 |
| sw02\_254b | Gulf of Mexico | F | 2002 | 0.46 | 0.45 | 0.44 | 0.44 | 0.42 |
| sw02\_254c | Gulf of Mexico | UN | 2002 | 0.45 | 0.44 | 0.44 | 0.43 | 0.42 |
| sw03\_156a | Gulf of Mexico | UN | 2003 | 0.48 | 0.47 | 0.47 | 0.46 | 0.45 |
| sw03\_162a | Gulf of Mexico | UN | 2003 | 0.47 | 0.46 | 0.46 | 0.45 | 0.45 |
| sw03\_163a | Gulf of Mexico | UN | 2003 | 0.52 | 0.50 | 0.50 | 0.49 | 0.47 |
| sw03\_164a | Gulf of Mexico | UN | 2003 | 0.51 | 0.50 | 0.49 | 0.49 | 0.48 |
| sw03\_165a | Gulf of Mexico | UN | 2003 | 0.42 | 0.40 | 0.40 | 0.39 | 0.39 |
| sw03\_165b | Gulf of Mexico | UN | 2003 | 0.42 | 0.41 | 0.41 | 0.41 | 0.40 |
| sw03\_167a | Gulf of Mexico | UN | 2003 | 0.74 | 0.72 | 0.72 | 0.69 | 0.64 |
| sw03\_173b | Gulf of Mexico | UN | 2003 | 0.61 | 0.61 | 0.60 | 0.59 | 0.58 |
| sw03\_197a | North Atlantic Delaware | M | 2003 | 0.77 | 0.76 | 0.75 | 0.75 | 0.75 |
| sw03\_197b | North Atlantic Delaware | M | 2003 | 0.69 | 0.67 | 0.66 | 0.64 | 0.60 |
| sw03\_201b | North Atlantic Delaware | M | 2003 | 0.65 | 0.63 | 0.63 | 0.63 | 0.65 |
| sw03\_202a | North Atlantic Delaware | M | 2003 | 0.52 | 0.50 | 0.50 | 0.50 | 0.45 |
| sw03\_206a | North Atlantic Delaware | M | 2003 | 0.48 | 0.47 | 0.47 | 0.47 | 0.53 |
| sw03\_206c | North Atlantic Delaware | M | 2003 | 0.56 | 0.54 | 0.54 | 0.55 | 0.57 |
| sw03\_207a | North Atlantic Delaware | M | 2003 | 0.49 | 0.48 | 0.48 | 0.48 | 0.51 |
| sw03\_247a | Mediterranean | UN | 2003 | 0.70 | 0.68 | 0.68 | 0.69 | 0.78 |
| sw03\_249a | Mediterranean | UN | 2003 | 0.75 | 0.73 | 0.73 | 0.74 | 0.83 |
| sw03\_249b | Mediterranean | UN | 2003 | 0.77 | 0.75 | 0.73 | 0.74 | 0.84 |
| sw03\_249c | Mediterranean | UN | 2003 | 0.71 | 0.70 | 0.71 | 0.72 | 0.77 |
| sw03\_253a | Mediterranean | UN | 2003 | 0.49 | 0.46 | 0.46 | 0.45 | 0.44 |
| sw03\_253b | Mediterranean | UN | 2003 | 0.57 | 0.55 | 0.55 | 0.53 | 0.48 |
| sw03\_253c | Mediterranean | UN | 2003 | 0.53 | 0.50 | 0.49 | 0.47 | 0.44 |
| sw05\_196a | Norway | M | 2005 | 0.76 | 0.75 | 0.75 | 0.75 | 0.73 |
| sw09\_153a | Norway | M | 2010 | 0.66 | 0.65 | 0.65 | 0.64 | 0.63 |
| sw10\_147a | Norway | M | 2010 | 0.80 | 0.79 | 0.79 | 0.78 | 0.77 |
| sw10\_149a | Norway | M | 2010 | 0.78 | 0.77 | 0.77 | 0.77 | 0.77 |
| sw10\_150a | Norway | M | 2010 | 0.77 | 0.76 | 0.76 | 0.77 | 0.78 |
| sw14\_092a | DOMINICA | F | 2014 | 0.54 | 0.52 | 0.52 | 0.52 | 0.52 |
| sw14\_106a | DOMINICA | F | 2014 | 0.53 | 0.51 | 0.51 | 0.50 | 0.49 |
| sw14\_106b | DOMINICA | F | 2014 | 0.52 | 0.50 | 0.50 | 0.50 | 0.50 |
| sw14\_114a | DOMINICA | F | 2014 | 0.47 | 0.45 | 0.45 | 0.44 | 0.40 |
| sw14\_114b | DOMINICA | F | 2014 | 0.51 | 0.49 | 0.49 | 0.48 | 0.46 |
| sw14\_119a | DOMINICA | F | 2014 | 0.51 | 0.49 | 0.49 | 0.47 | 0.46 |
| sw14\_119b | DOMINICA | F | 2014 | 0.47 | 0.45 | 0.44 | 0.43 | 0.42 |
| sw15\_061a | DOMINICA | F | 2015 | 0.50 | 0.48 | 0.47 | 0.46 | 0.45 |
| sw15\_061b | DOMINICA | F | 2015 | 0.55 | 0.53 | 0.52 | 0.52 | 0.49 |
| sw15\_061c | DOMINICA | F | 2015 | 0.50 | 0.48 | 0.47 | 0.47 | 0.46 |
| sw15\_078a | DOMINICA | F | 2015 | 0.48 | 0.47 | 0.46 | 0.45 | 0.42 |
| sw15\_078b | DOMINICA | F | 2015 | 0.52 | 0.51 | 0.51 | 0.50 | 0.47 |
| sw15\_078c | DOMINICA | JM | 2015 | 0.49 | 0.48 | 0.47 | 0.46 | 0.43 |
| sw15\_078d | DOMINICA | JM | 2015 | 0.55 | 0.53 | 0.53 | 0.52 | 0.47 |
| sw15\_085a | DOMINICA | F | 2015 | 0.50 | 0.49 | 0.48 | 0.48 | 0.42 |
| sw15\_090a | DOMINICA | F | 2015 | 0.54 | 0.53 | 0.52 | 0.52 | 0.53 |
| sw15\_090b | DOMINICA | F | 2015 | 0.54 | 0.52 | 0.52 | 0.52 | 0.51 |
| sw15\_090c | DOMINICA | F | 2015 | 0.56 | 0.54 | 0.54 | 0.54 | 0.52 |
| sw15\_091a | DOMINICA | F | 2015 | 0.54 | 0.53 | 0.52 | 0.52 | 0.52 |
| sw15\_091b | DOMINICA | F | 2015 | 0.57 | 0.55 | 0.55 | 0.55 | 0.56 |
| sw15\_097a | DOMINICA | F | 2015 | 0.53 | 0.51 | 0.51 | 0.51 | 0.50 |
| sw15\_100a | DOMINICA | F | 2015 | 0.49 | 0.47 | 0.46 | 0.45 | 0.41 |
| sw15\_100b | DOMINICA | F | 2015 | 0.47 | 0.46 | 0.46 | 0.45 | 0.39 |
| sw16\_103a | DOMINICA | F | 2016 | 0.54 | 0.53 | 0.53 | 0.53 | 0.53 |
| sw16\_115b | DOMINICA | F | 2016 | 0.51 | 0.50 | 0.49 | 0.48 | 0.48 |
| sw16\_131a | Norway | M | 2016 | 0.77 | 0.75 | 0.75 | 0.75 | 0.57 |
| sw16\_133a | DOMINICA | F | 2016 | 0.46 | 0.44 | 0.44 | 0.43 | 0.42 |
| sw16\_134a | Norway | M | 2016 | 0.85 | 0.85 | 0.85 | 0.86 | 0.87 |
| sw16\_134ad | DOMINICA | F | 2016 | 0.50 | 0.48 | 0.48 | 0.47 | 0.43 |
| sw17\_119c | DOMINICA | UN | 2016 | 0.83 | 0.79 | 0.78 | 0.77 | 0.77 |
| sw18\_083a | DOMINICA | AF | 2018 | 0.50 | 0.48 | 0.47 | 0.47 | 0.46 |
| sw18\_140a | DOMINICA | AF | 2018 | 0.52 | 0.50 | 0.49 | 0.49 | 0.48 |
| sw18\_140b | DOMINICA | AF | 2018 | 0.50 | 0.48 | 0.47 | 0.46 | 0.42 |
| sw18\_143a | DOMINICA | AF | 2018 | 0.53 | 0.52 | 0.51 | 0.50 | 0.49 |
| sw18\_267a | Azores | F | 2018 | 0.64 | 0.61 | 0.61 | 0.61 | 0.62 |
| sw18\_281a | Azores | UN | 2018 | 0.62 | 0.58 | 0.57 | 0.56 | 0.57 |
| sw18\_292a | Azores | UN | 2018 | 0.65 | 0.63 | 0.62 | 0.60 | 0.60 |
| sw19\_088a | Azores | UN | 2019 | 0.68 | 0.66 | 0.66 | 0.67 | 0.68 |
| sw19\_137a | Azores | UN | 2019 | 0.52 | 0.50 | 0.50 | 0.49 | 0.42 |
| sw19\_158a | Azores | UN | 2019 | 0.42 | 0.41 | 0.41 | 0.40 | 0.39 |
| sw19\_160a | Azores | UN | 2019 | 0.49 | 0.47 | 0.47 | 0.45 | 0.43 |
| sw19\_163a | Azores | UN | 2019 | 0.42 | 0.41 | 0.40 | 0.39 | 0.38 |
| sw19\_253b | Norway | M | 2019 | 0.56 | 0.54 | 0.53 | 0.50 | 0.42 |
| sw19\_259a | Norway | M | 2019 | 0.59 | 0.57 | 0.57 | 0.55 | 0.50 |
| sw20\_161a | Azores | UN | 2020 | 0.52 | 0.50 | 0.50 | 0.49 | 0.37 |
| sw20\_184a | Azores | UN | 2020 | 0.71 | 0.69 | 0.69 | 0.69 | 0.62 |
| sw20\_189a | Azores | UN | 2020 | 0.40 | 0.38 | 0.37 | 0.38 | 0.39 |
| sw20\_191a | Azores | UN | 2020 | 0.62 | 0.58 | 0.58 | 0.58 | 0.46 |
| sw20\_192a | Azores | UN | 2020 | 0.50 | 0.47 | 0.46 | 0.42 | 0.38 |
| sw20\_225a | Azores | UN | 2020 | 0.61 | 0.59 | 0.58 | 0.57 | 0.52 |
| sw21\_196a | Azores | UN | 2021 | 0.52 | 0.50 | 0.49 | 0.48 | 0.45 |
| sw21\_209a | Azores | UN | 2021 | 0.52 | 0.50 | 0.49 | 0.46 | 0.44 |
| sw21\_211a | Azores | UN | 2021 | 0.63 | 0.61 | 0.60 | 0.59 | 0.58 |
| sw21\_215a | Azores | UN | 2021 | 0.55 | 0.52 | 0.52 | 0.51 | 0.47 |
| sw21\_215b | Azores | UN | 2021 | 0.51 | 0.49 | 0.49 | 0.47 | 0.43 |
| sw21\_221a | Azores | UN | 2021 | 0.53 | 0.50 | 0.50 | 0.49 | 0.38 |
| sw21\_230a | Azores | UN | 2021 | 0.60 | 0.58 | 0.58 | 0.57 | 0.55 |
| sw21\_242a | Azores | UN | 2021 | 0.65 | 0.63 | 0.63 | 0.62 | 0.61 |

All these tables merged are also exported as a single table:

write.table(tagsipi,file="tagsipi.txt")

# Aknowledgements

We thank Ben Bolker for helpful advice via the answer to our question on “Stack Exchange” at <https://stats.stackexchange.com/questions/616697/how-to-estimate-precision-on-a-prediction-for-a-glmm-for-an-unobserved-level-of/>. Also thanks to Ben Augustine and Perry du Valpine over email, regarding the comparison between Bayesian inferences in Nimble and lme4. Luca (in private) and Wei Zhang (both in private and publicly) provided helpful advice following the question on the nimble user group at <https://groups.google.com/g/nimble-users/c/7UHlKdCC8B4> about the same comparison. Len Thomas provided useful discussion and advice.

All errors remaining are the sole responsibility of the authors.