Bycatch Estimation and Expansion in STAN

Eric J. Ward, Jason E. Jannot

2019-08-01

## Load library

library(devtools)  
devtools::install\_github("ericward-noaa/bycatch")  
library(bycatch)  
set.seed(123)

### Load data

# replace this with your own data frame  
d = data.frame("Year"= 2002:2014,   
 "Takes" = c(0, 0, 0, 0, 0, 0, 0, 0, 1, 3, 0, 0, 0),  
 "expansionRate" = c(24, 22, 14, 32, 28, 25, 30, 7, 26, 21, 22, 23, 27),  
 "Sets" = c(391, 340, 330, 660, 470, 500, 330, 287, 756, 673, 532, 351, 486))

## Simple model with constant bycatch, no covariates

We’ll start by fitting a model with constant bycatch rate,

fit = fit\_bycatch(Takes ~ 1, data=d, time="Year", effort="Sets", family="poisson",  
 time\_varying = FALSE)

If divergent transition warnings or other issues indicating lack of convergence are a problem, we can try changing some of the control arguments, e.g.

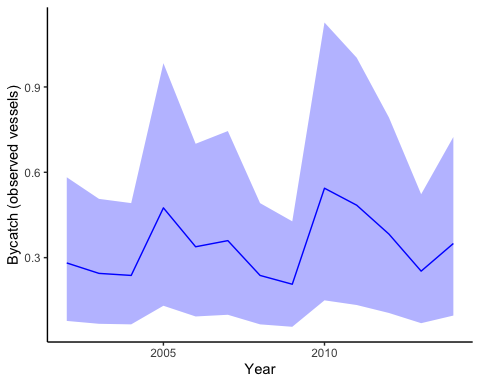
fit = fit\_bycatch(Takes ~ 1, data=d, time="Year", effort="Sets", family="poisson",  
 time\_varying = FALSE, control=list(adapt\_delta=0.99,max\_treedepth=20))

We can also increase the iterations and number of chains from the defaults (1000 and 3),

fit = fit\_bycatch(Takes ~ 1, data=d, time="Year", effort="Sets", family="poisson",  
 time\_varying = FALSE, iter=3000, chains=4)

### Make plots

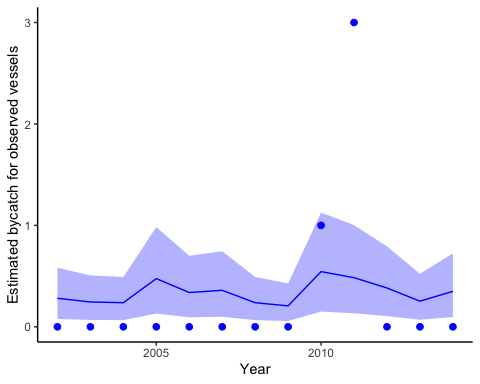
plot\_fitted(fit, xlab="Year", ylab = "Bycatch (observed vessels)")



Estimated bycatch for observed vessels (not expanded by observer coverage), but including observed takes and effort.

We can include points also,

plot\_fitted(fit, xlab="Year", ylab = "Estimated bycatch for observed vessels", include\_points = TRUE)



Observed bycatch (not expanded by observer coverage), incorporating data on observed takes and effort. Dots represent observed bycatch events.

### Extracting model selection information (LOOIC)

The loo package in R provides a nice interface for extracting leave one out information criterion (LOOIC) from stanfit objects. Like AIC, lower is better. Values of LOOIC can be used to compare models with the same response but different structure (covariates or not, time-varying bycatch or not, etc). Additional information on LOOIC can be found at [mc-stan.org](http://mc-stan.org/rstanarm/reference/loo.stanreg.html), [Vehtari et al. 2017](https://link.springer.com/article/10.1007/s11222-016-9696-4), or the vignette for the [loo package](https://cran.r-project.org/web/packages/loo/vignettes/loo2-example.html).

loo::loo(fit$fitted\_model)$estimates

## Warning: Some Pareto k diagnostic values are too high. See help('pareto-k-diagnostic') for details.

## Estimate SE  
## elpd\_loo -11.177868 6.043199  
## p\_loo 2.526928 2.162952  
## looic 22.355736 12.086398

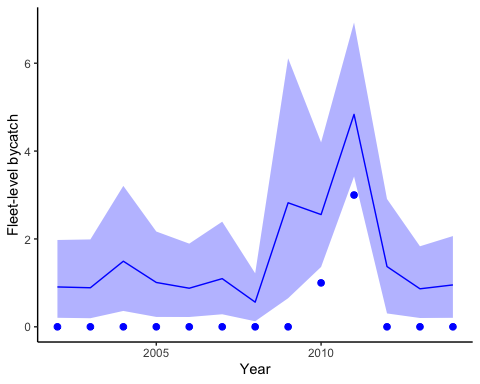
## Expanding bycatch estimates

Using our example above, the observer coverage for that dataset was less than 100% and so our estimates need to be expanded to the fleetwide level. There are some important control arguments here that are left at defaults, but should maybe be changed for huge numbers of bycatch.

expanded = expand(fit, coverage = d$expansionRate)

And we can then plot these estimates. Like the previous function we can specify whether to include the raw points or not.

plot\_expanded(fitted\_model=fit, expanded\_estimates = expanded, xlab="Year", ylab = "Fleet-level bycatch", include\_points = TRUE)



Estimated fleet-level expanded bycatch, incorporating data on takes, effort, and observer coverage. Dots represent observed bycatch events.

### Make table of expanded bycatch estimates

We can also do things like summarize the expanded estimates in table form

df = data.frame("time" = d[,"Year"],   
 "mean" = apply(expanded, 2, mean),  
 "median" = apply(expanded, 2, quantile, 0.5),  
 "lower95" = apply(expanded, 2, quantile, 0.025),  
 "upper95" = apply(expanded, 2, quantile, 0.975))  
  
write.table(df, "estimated\_bycatch.csv", row.names=F, col.names=T, sep=",")

## Negative binomial example

Using our example dataset above, we can also switch from the Poisson model to Negative Binomial.

fit = fit\_bycatch(Takes ~ 1, data=d, time="Year", effort="Sets", family="nbinom2",  
 time\_varying = FALSE)

The degree of overdispersion here is stored in the variable nb2\_phi, which we can get with

phi = rstan::extract(fit$fitted\_model)$nb2\_phi

## Example with covariates

Following [Martin et al. 2015](http://onlinelibrary.wiley.com/doi/10.1890/14-0059.1/abstract) we can include fixed or continuous covariates.

For example, we could include a julian day, and a break point (representing a regulatory change for example) in the data. We could model the first variable as a continuous predictor and the second as a factor.

d$Day = sample(seq(220,280),size=nrow(d),replace=T)  
d$Reg = ifelse(d$Year < 2008, 0, 1)

Using the formula interface makes it easy to include covariates,

fit = fit\_bycatch(Takes ~ Day + Reg, data=d, time="Year", effort="Sets", family="poisson",  
 time\_varying = FALSE)

We can get the 3 covariate effects out with the following call:

betas = rstan::extract(fit$fitted\_model)$beta

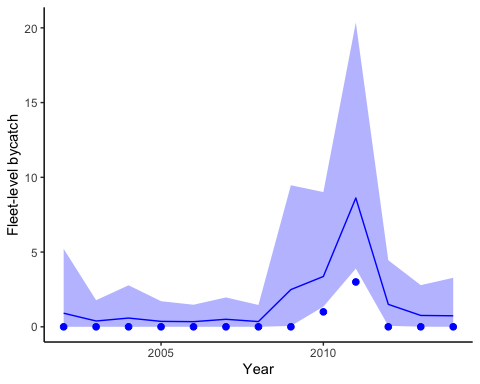
Note that ‘betas’ has 3 columns. These correspond to (1) the intercept, (2) continuous predictor, and (3) factor variable above. If we didn’t include the covariates, we’d still estimate beta[1] as the intercept.

## Fit model with time-varying effects

To incorporate potential autocorrelation, we can fit a model with time-varying random effects. This is equivalent to a dynamic linear model with time varying intercept in a Poisson GLM.

fit = fit\_bycatch(Takes ~ 1, data=d, time="Year", effort="Sets", family="poisson",  
 time\_varying = TRUE)

expanded = expand(fit, coverage=d$expansionRate)  
plot\_expanded(fitted\_model=fit, expanded\_estimates = expanded, xlab="Year", ylab = "Fleet-level bycatch", include\_points = TRUE)



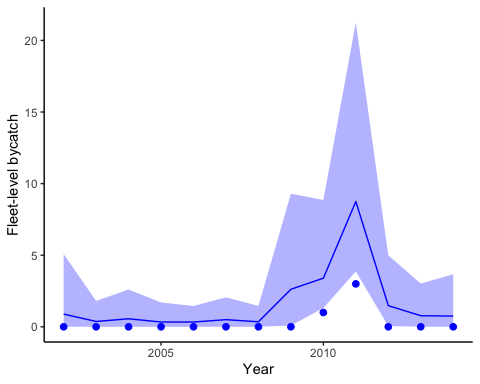
Estimated fleet-level expanded bycatch from the model with time-varying effects, incorporating data on takes, effort, and observer coverage. Dots represent observed bycatch events.

## Fit model with no bycatch events

# replace this with your own data frame  
d = data.frame("Year"= 2002:2014,   
 "Takes" = c(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0),  
 "expansionRate" = c(24, 22, 14, 32, 28, 25, 30, 7, 26, 21, 22, 23, 27),  
 "Sets" = c(391, 340, 330, 660, 470, 500, 330, 287, 756, 673, 532, 351, 486))

fit = fit\_bycatch(Takes ~ 1, data=d, time="Year", effort="Sets", family="poisson",  
 time\_varying = FALSE)

expanded = expand(fit, coverage=d$expansionRate)  
plot\_expanded(fitted\_model=fit, expanded\_estimates = expanded, xlab="Year", ylab = "Fleet-level bycatch", include\_points = TRUE)



Estimated fleet-level expanded bycatch from the dataset with no events, incorporating effort, and observer coverage. Dots represent observed bycatch events.