CPUE_predict_example

Markus Min

3/8/2021

Load libraries

```
library(here)
library(tidyverse)
library(pscl)
```

Import data

```
dat <- read.csv(here("hook_and_line_data", "CPUE_data_for_model.csv"), row.names = 1)
dat$month <- as.character(dat$month)
```

Model survey in the detection process

```
# Survey on the detection side
yelloweye_zip_effort_depth_basin_survey <- zeroinfl(yelloweye_catch ~
                            offset(angler_hours)|
                            # Predictors of detection
                            mean_depth + basin + survey, data = dat)
summary(yelloweye_zip_effort_depth_basin_survey)
```

```
##
## Call:
## zeroinfl(formula = yelloweye_catch ~ offset(angler_hours) | mean_depth +
##     basin + survey, data = dat)
##
## Pearson residuals:
##        Min         1Q     Median         3Q        Max
## -2.733e+00 -1.233e-02 -2.880e-03 -9.032e-04  2.993e+03
##
## Count model coefficients (poisson with log link):
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -14.486      0.106  -136.7   <2e-16 ***
##
## Zero-inflation model coefficients (binomial with logit link):
##                               Estimate Std. Error z value Pr(>|z|)
```

```
## (Intercept)                     3.19678    1.16475   2.745 0.006058 **
## mean_depth                     -0.03339    0.01305  -2.558 0.010537 *
## basinHood Canal               -12.47619  117.44141  -0.106 0.915397
## basinSan Juan Islands          -3.40881    0.99382  -3.430 0.000604 ***
## basinSouth Sound               14.59825  620.09492   0.024 0.981218
## basinStraits of Juan de Fuca   14.58340 1105.44554   0.013 0.989474
## basinWhidbey Island             1.70732    1.11469   1.532 0.125608
## surveylingcod_bycatch          -2.61817    1.42686  -1.835 0.066518 .
## surveyPercy_Washington         -0.47352    0.85331  -0.555 0.578946
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Number of iterations in BFGS optimization: 22
## Log-likelihood: -735.8 on 10 Df
```

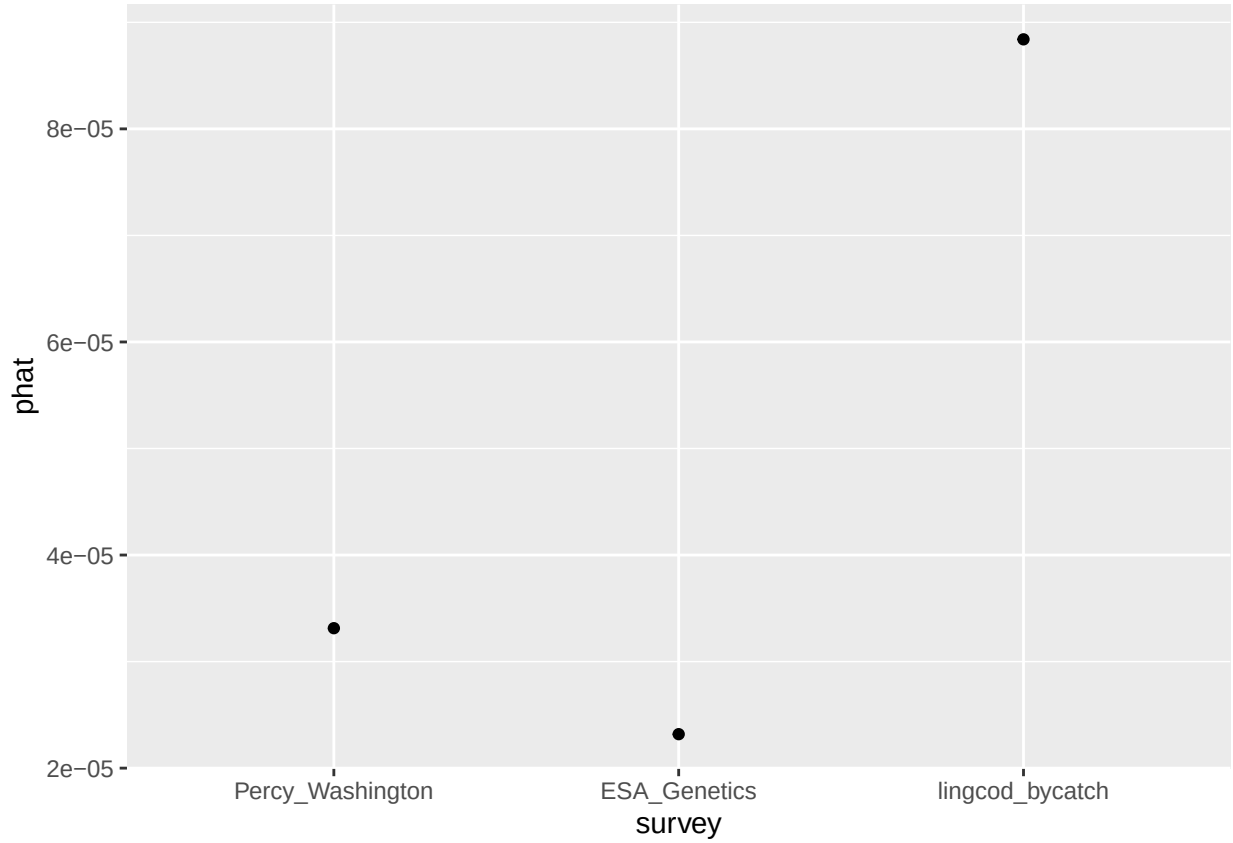
```
# Create df for predicted values
angler_hours = rep(mean(dat$angler_hours), 3)
mean_depth = rep(mean(dat$mean_depth, na.rm = TRUE), 3)
basin = factor(rep("Central Sound", 3), levels = c("Central Sound", "Whidbey Island", "South Sound", "S
survey = factor(c("Percy_Washington", "ESA_Genetics", "lingcod_bycatch"), levels = c("Percy_Washington"
newdata <- data.frame(angler_hours, mean_depth, basin, survey)

# Get predicted values
newdata$phat <- predict(yelloweye_zip_effort_depth_basin_survey, newdata = newdata)

newdata
```

```
##   angler_hours mean_depth         basin           survey         phat
## 1     5.402264   55.01862 Central Sound Percy_Washington 3.313309e-05
## 2     5.402264   55.01862 Central Sound     ESA_Genetics 2.318827e-05
## 3     5.402264   55.01862 Central Sound  lingcod_bycatch 8.840316e-05
```

```
ggplot(newdata, aes(x = survey, y = phat)) +
  geom_point()
```



Model survey in the count process

```
# Survey on the count side

yelloweye_zip_effort_depth_basin_survey <- zeroinfl(yelloweye_catch ~
                            offset(angler_hours) + survey |
                            # Predictors of detection
                            mean_depth + basin, data = dat)
summary(yelloweye_zip_effort_depth_basin_survey)
```

```
##
## Call:
## zeroinfl(formula = yelloweye_catch ~ offset(angler_hours) + survey |
##     mean_depth + basin, data = dat)
##
## Pearson residuals:
##        Min         1Q     Median         3Q        Max
## -2.143e+00 -7.260e-02 -4.361e-03 -9.007e-04  3.926e+03
##
## Count model coefficients (poisson with log link):
##                         Estimate Std. Error  z value Pr(>|z|)
## (Intercept)             -15.0282     0.1434 -104.825   <2e-16 ***
## surveylingcod_bycatch     7.5497     0.3835   19.688   <2e-16 ***
```

```
## surveyPercy_Washington   2.2676     0.2341    9.687   <2e-16 ***
##
## Zero-inflation model coefficients (binomial with logit link):
##                                Estimate Std. Error z value Pr(>|z|)
## (Intercept)                     3.59600    0.88496   4.063 4.83e-05 ***
## mean_depth                     -0.04051    0.01288  -3.146  0.00166 **
## basinHood Canal               -12.91192  145.57515  -0.089  0.92932
## basinSan Juan Islands          -2.93110    0.74684  -3.925 8.68e-05 ***
## basinSouth Sound               14.41951  416.19273   0.035  0.97236
## basinStraits of Juan de Fuca   14.33100 1072.65538   0.013  0.98934
## basinWhidbey Island             1.92355    1.13244   1.699  0.08940 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Number of iterations in BFGS optimization: 22
## Log-likelihood: -639.9 on 10 Df
```

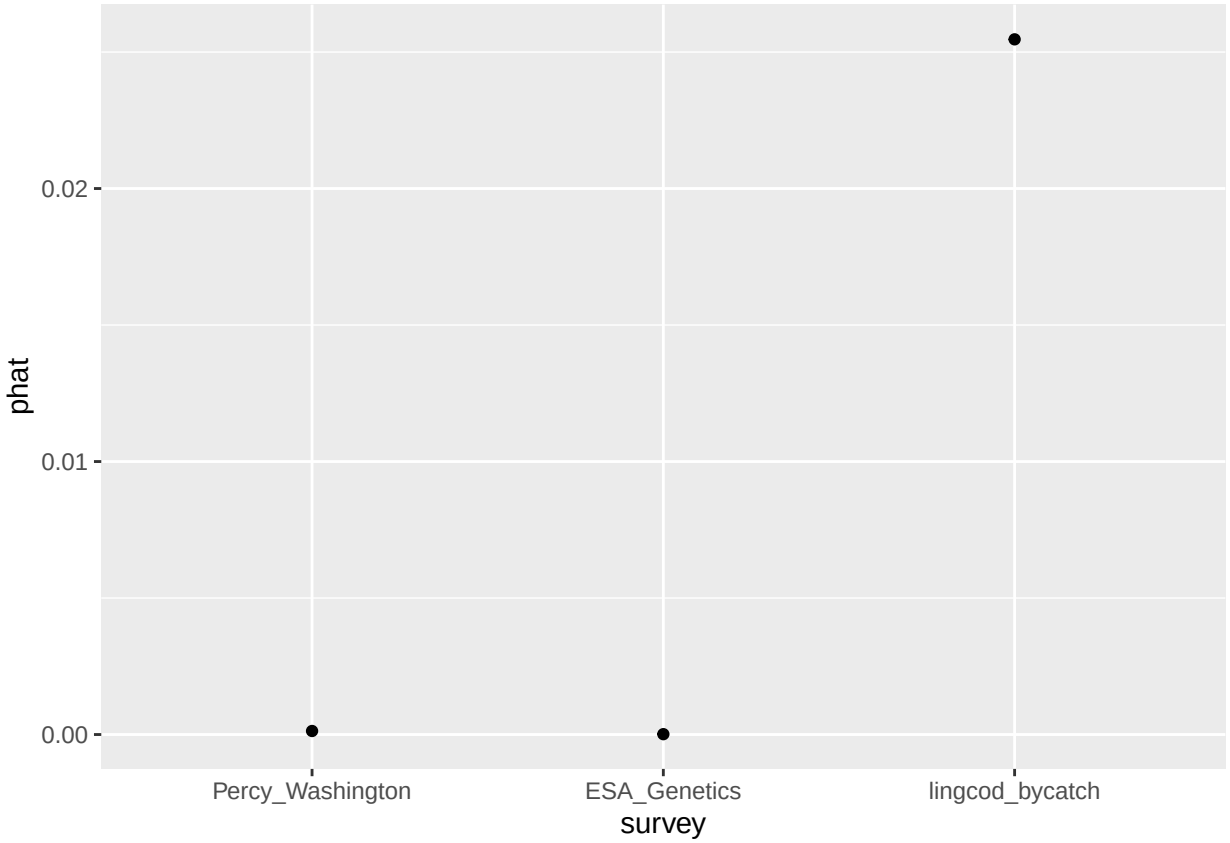
```
# Create df for predicted values
angler_hours = rep(mean(dat$angler_hours), 3)
mean_depth = rep(mean(dat$mean_depth, na.rm = TRUE), 3)
basin = factor(rep("Central Sound", 3), levels = c("Central Sound", "Whidbey Island", "South Sound", "S
survey = factor(c("Percy_Washington", "ESA_Genetics", "lingcod_bycatch"), levels = c("Percy_Washington"
newdata <- data.frame(angler_hours, mean_depth, basin, survey)

# Get predicted values
newdata$phat <- predict(yelloweye_zip_effort_depth_basin_survey, newdata = newdata)

newdata
```

```
##   angler_hours mean_depth         basin           survey         phat
## 1     5.402264   55.01862 Central Sound Percy_Washington 1.293983e-04
## 2     5.402264   55.01862 Central Sound     ESA_Genetics 1.340091e-05
## 3     5.402264   55.01862 Central Sound  lingcod_bycatch 2.546508e-02
```

```
ggplot(newdata, aes(x = survey, y = phat)) +
  geom_point()
```


0.02
0.01
0.00
phat
Percy_Washington
ESA_Genetics
lingcod_bycatch
survey