Spawner Parr Regression in Stan

Scott A. Akenhead

11/10/22

```
library (rstan); library(knitr)
options(mc.cores = parallel::detectCores())
rstan_options(auto_write = TRUE)
```

Datasets

three datasets:

1. Spawners as PC-1 from AUC and PLD[1]
2. Parr_wcon as weighted mean of parr surveys during summer.
3. Parr_reg as estimates from nonlinear, weighted, survival model.

Match the spawner years (2001 to 2020) to the corresponding parr identified by smolt year: spawner year + 2.

```
print(getwd())
```

```
[1] "/Users/Scott/Documents/Projects/OK SOX 2022/OK SOX Analysis/OSO Stan"
```

```
file.exists('Spawner Parr.stan')
```

```
[1] TRUE
```

[1] AUC: area under curve, the trapezoid from linear interpolation between observations of spawner abundance; PLD: peak live plus dead is the maximum count of spawners, alive and dead.

```
file.exists('SpawnersPC.Rdata')
```

```
[1] TRUE
```

```
file.exists('parr conventional weighted.RData')
```

```
[1] TRUE
```

```
file.exists('parr regression est and stderr.RData')
```

```
[1] TRUE
```

```
spawners <- readRDS('SpawnersPC.Rdata')
dim(spawners); colnames(spawners); range(spawners$Year)
```

```
[1] 20  3
```

```
[1] "Year" "PC1"  "PC2"
```

```
[1] 2001 2020
```

```
parrWC <- readRDS('parr conventional weighted.RData')
dim(parrWC); colnames(parrWC); range(parrWC$Smolt_Year)
```

```
[1] 24 11
```

```
 [1] "Smolt_Year"     "Parr_m"         "Parr_sd"        "Parr_n"
 [5] "PreSmolt_m"     "PreSmolt_sd"    "PreSmolt_n"     "PreSmolt_lower"
 [9] "PreSmolt_upper" "Parr_lower"     "Parr_upper"
```

```
[1] 1998 2021
```

```
parrReg <- readRDS('parr regression est and stderr.RData')
dim(parrReg); colnames(parrReg); range(parrReg$Smolt_Year)
```

```
[1] 26  3
```

```
[1] "Smolt_Year" "estimate"   "stderr"
```

```
[1]    0 2021
```

```
# fix spawners
spawners = spawners[ order(spawners$Year), ]
colnames(spawners) <- c('Brood_Year','Spawners_PC1','Spawners_PC2')

# match spawners Brood_Year 2001:2020 to Smolt_Year 1998:2021
#     tricky Smolt Year 2002 is data from 2001 and early 2002, so
#     data for parr in 2001 is related to spawners year 2000
#     thus brood_year is  smolt year -1
# ParrWC is NA for smolt years 2020,2021; broods 2018, 2019.
a <- parrWC$Smolt_Year
j = (a >= 2001) & (a <= 2020)
parr1 = parrWC[j,2:3 ] *1e-6 # Parr_m,  Parr_sd
parr1$Brood_Year <- a[j]

a <- parrReg$Smolt_Year
j = (a >= 2001) & (a <= 2020)
parr2  = parrReg[j,2:3 ]   # estimate, stderr
parr2$Brood_Year <- a[j]

spawners1 = spawners[, 1:2] # Year, PC1
kable(cbind(spawners1, parr1, parr2), digits=3, row.names = FALSE)
```

Brood_Year	Spawners_PC1	Parr_m	Parr_sd	Brood_Year	estimate	stderr	Brood_Year
2001	-0.279	NA	NA	2001	1.621	0.104	2001
2002	-1.439	3.256	0.524	2002	3.777	0.086	2002
2003	-1.058	2.458	0.729	2003	2.234	0.062	2003
2004	-0.439	2.052	NA	2004	1.146	0.065	2004
2005	-0.497	1.035	NA	2005	1.054	0.066	2005
2006	-0.807	3.708	0.584	2006	3.137	0.053	2006
2007	-1.173	2.241	NA	2007	2.721	0.052	2007

Brood_Year	Spawners_PC1	Parr_m	Parr_sd	Brood_Year	estimate	stderr	Brood_Year
2008	2.121	2.736	0.944	2008	3.030	0.040	2008
2009	0.493	1.485	0.251	2009	1.421	0.052	2009
2010	3.895	8.841	0.889	2010	10.201	0.049	2010
2011	0.128	2.135	0.395	2011	1.630	0.040	2011
2012	1.047	6.657	0.372	2012	6.786	0.050	2012
2013	-0.451	4.411	0.014	2013	4.294	0.046	2013
2014	2.674	7.142	0.989	2014	7.146	0.043	2014
2015	-1.267	3.438	0.135	2015	3.069	0.049	2015
2016	0.344	12.002	2.567	2016	11.699	0.046	2016
2017	-1.315	3.012	0.072	2017	3.156	0.058	2017
2018	-0.679	7.590	0.737	2018	6.869	0.064	2018
2019	-1.131	1.802	0.042	2019	2.019	0.079	2019
2020	-0.166	NA	NA	2020	3.857	0.086	2020

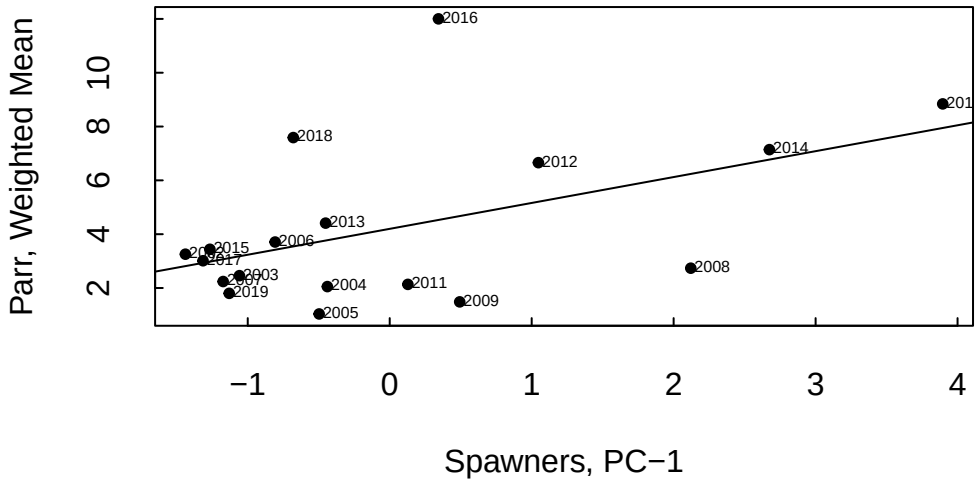
Check Plots

```
par(tcl=0.2)
x=spawners1$Spawners_PC1
y=parr1$Parr_m
r1 <- lm(y~x); summary(r1)
```

```
Call:
lm(formula = y ~ x)

Residuals:
    Min      1Q  Median      3Q     Max
-3.5030 -1.6200  0.1827  0.5995  7.4728

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   4.1986     0.6360   6.601  6.1e-06 ***
x             0.9621     0.4308   2.233   0.0402 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.698 on 16 degrees of freedom
  (2 observations deleted due to missingness)
```

```
Multiple R-squared:  0.2376,    Adjusted R-squared:   0.19
F-statistic: 4.987 on 1 and 16 DF,  p-value: 0.04017
```

```r
plot(x,y, pch=20,
     xlab= "Spawners, PC-1", ylab="Parr, Weighted Mean")
  text(x,y, labels=spawners1$Brood_Year,cex=.5,pos=4, offset=0.1)
abline(r1)
```



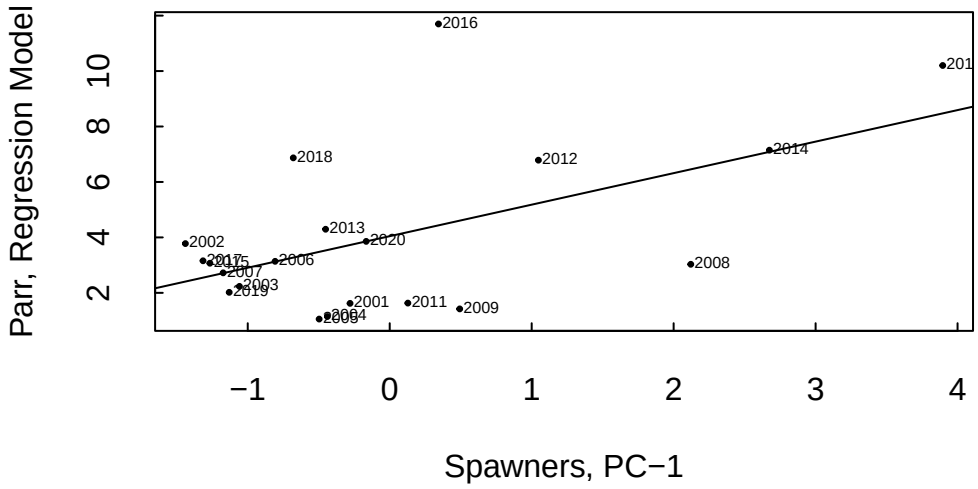

```r
# setup for Stan. eliminate NAs.
dat1 = list(
    n_years  = 18,
    spawner = x[2:19],
    parr     = y[2:19]
)
```

```r
x=spawners1$Spawners_PC1
y=parr2$estimate
r2 <- lm(y~x); summary(r2)
```

```
Call:
lm(formula = y ~ x)

Residuals:
   Min     1Q Median     3Q    Max
```

```
-3.425 -2.178  0.012  0.916  7.265

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   4.0434     0.5763   7.016 1.51e-06 ***
x             1.1372     0.4109   2.768   0.0127 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.577 on 18 degrees of freedom
Multiple R-squared:  0.2985,    Adjusted R-squared:  0.2595
F-statistic:  7.66 on 1 and 18 DF,  p-value: 0.01269
```

```r
par(tcl=0.2)
plot(x,y, pch=20,cex=.5,
     xlab= "Spawners, PC-1", ylab="Parr, Regression Model")
text(x,y, labels=spawners1$Brood_Year,cex=.5,pos=4, offset=0.1)
abline(r2)
```




```r
# setup for Stan
dat2= list(
     n_years = 20,
    spawner = x,
    parr     = y
)
```

Sampling

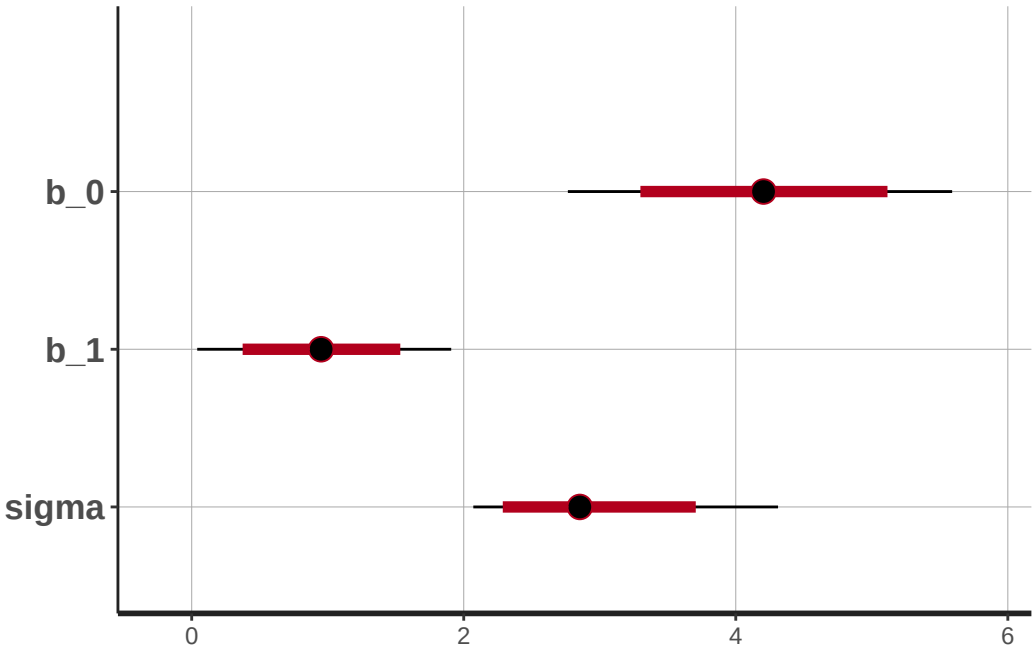
```
fit1 <- stan(
    file = "Spawner Parr.stan", # program, code
    data = dat1,            # data conforming to description in "Stan ATS.stan"
    chains = 4,             # number of Markov chains
    cores = 4,              # number of cores (one per chain)
    warmup = 1000,          # number of warmup iterations per chain
    iter = 2000             # total number of iterations per chain
)
print(summary(fit1)$summary[1:3, -2], digits=3) # drop se_mean, drop lp__
```

```
      mean    sd   2.5%   25%   50%  75% 97.5% n_eff Rhat
b_0   4.20 0.714 2.7656 3.733 4.204 4.67  5.59  3226    1
b_1   0.96 0.469 0.0406 0.663 0.951 1.25  1.91  2889    1
sigma 2.94 0.580 2.0705 2.523 2.853 3.27  4.31  2595    1
```

```
plot(fit1)
```

```
ci_level: 0.8 (80% intervals)
```

```
outer_level: 0.95 (95% intervals)
```



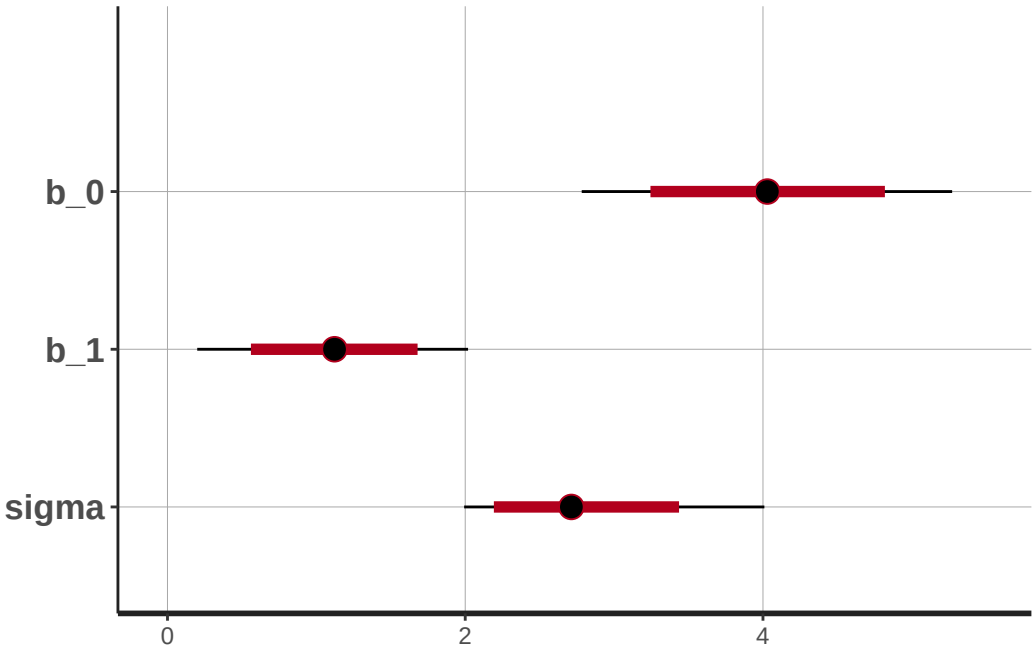

```
fit2 <- stan(
    file = "Spawner Parr.stan", # program, code
    data = dat2,            # data conforming to description in "Stan ATS.stan"
    chains = 4,             # number of Markov chains
    cores = 4,              # number of cores (one per chain)
    warmup = 1000,          # number of warmup iterations per chain
    iter = 2000             # total number of iterations per chain
)
print(summary(fit2)$summary[1:3, -2], digits=3) # drop se_mean, drop lp__
```

```
      mean    sd  2.5%   25%  50%  75% 97.5% n_eff Rhat
b_0   4.03 0.620 2.783 3.631 4.03 4.45  5.27  3162    1
b_1   1.12 0.450 0.198 0.856 1.12 1.41  2.02  2852    1
sigma 2.78 0.512 1.993 2.419 2.71 3.06  4.01  2647    1
```

```
plot(fit2)
```

```
ci_level: 0.8 (80% intervals)
```

```
outer_level: 0.95 (95% intervals)
```


b_0
b_1
sigma
0
2
4