

# Larkin Ricker in Salmon population

*Rui Zhang*

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Data</b>	<b>2</b>
<b>3</b>	<b>Model</b>	<b>8</b>
<b>4</b>	<b>Model comparison and analysis</b>	<b>9</b>
<b>5</b>	<b>Conclusion</b>	<b>23</b>
<b>6</b>	<b>Supplement</b>	<b>24</b>
6.1	Significance of $b$ . . . . .	24
6.2	result on the Early Stuart stock . . . . .	24
<b>7</b>	<b>References</b>	<b>25</b>

---

## Objective

- Finding a model for cyclic dominance in sockeye salmon population.
  - Statistically determining whether the Larkin Ricker model is needed.
- 

## 1 Introduction

It is a common fact that the population of salmon fluctuates periodically, with a cohort that hugely outnumbers the others. This phenomenon is usually called cyclic dominance. Some researches have been done and claim that cyclic dominance is possibly caused by environmental variabilities when the population of fish is in some kinds of situations, such as narrow spawning age distribution and low persistence (White et al.[1], Botsford et al.[2]).

The populations of the sockeye salmon in Fraser River periodically fluctuates with period four, which is the same as generation time. A cycle with period four can have many appearances, except the cyclic dominance. In addition to random environmental noise, is there any further mechanism that cause this pattern – cyclic dominance? My goal is to analyze the interaction between cohorts, if any.

## 2 Data

My data is from the supplemental material of White et al.[1]. We can take a glance at the Seymour data, which is one of the 19 Fraser River sockeye stocks.

```
library(RCurl)
Url=getURL("https://raw.githubusercontent.com/ruizhang-ray/salmon_larkin/master/data/Seymour.csv")
seymour=read.table(text=Url,header = FALSE,sep=',')
colnames(seymour)=c("year","eff","rec3","rec4","rec5")
summary(seymour)
```

```
##      year      eff      rec3      rec4
##  Min.   :1948   Min.    :   311   Min.    :    0.0   Min.    :  1944
## 1st Qu.:1963   1st Qu.:  2869   1st Qu.:    0.0   1st Qu.: 18347
## Median :1978   Median :   6585   Median :   136.0   Median : 55150
## Mean   :1978   Mean    : 14854   Mean    :   911.6   Mean    :124433
## 3rd Qu.:1994   3rd Qu.: 18984   3rd Qu.:   564.0   3rd Qu.:175000
## Max.   :2009   Max.    :108000   Max.    :23882.0   Max.    :812000
##
##      NA's      :3      NA's      :4
##
##      rec5
##  Min.    :    0
## 1st Qu.:   191
## Median :  1934
## Mean    :  5083
## 3rd Qu.:  7081
## Max.    :42537
## NA's    :    5
```

We can see there are about 62-year data, and in each year there are effective female spawner (eff) and adult recruits which returns to the spawning ground at age three, four and five (rec3, rec4, rec5).

According to the run-timing, the 8 most abundant sockeye stocks can be divided into four groups (Peterman et al[4]). Followings are the plots of Early Stuart, Seymour and Late Stuart, which return before all other Fraser stocks, at early-summer and at mid-summer, respectively.

```
Url=getURL("https://raw.githubusercontent.com/ruizhang-ray/salmon_larkin/master/data/Seymour.csv")
seymour=read.table(text=Url,header = FALSE,sep=',')
row.aval=which(!is.nan(apply(seymour,1,sum)))
seymour=as.matrix(seymour[row.aval,])
seymour=cbind(seymour,apply(seymour[,3:5],1,sum))
for (i in 1:3){
  seymour=cbind(seymour,seymour[,2+i]/seymour[,6])
}
colnames(seymour)=c('year','eff','rec3','rec4','rec5','rec.total','r3.per','r4.per','r5.per')

Url=getURL("https://raw.githubusercontent.com/ruizhang-ray/salmon_larkin/master/data/EarlyStuart.csv")
earlystuart=read.table(text=Url,header = FALSE,sep=',')
row.aval=which(!is.nan(apply(earlystuart,1,sum)))
earlystuart=as.matrix(earlystuart[row.aval,])
earlystuart=cbind(earlystuart,apply(earlystuart[,3:5],1,sum))
for (i in 1:3){
  earlystuart=cbind(earlystuart,earlystuart[,2+i]/earlystuart[,6])
}
```

```

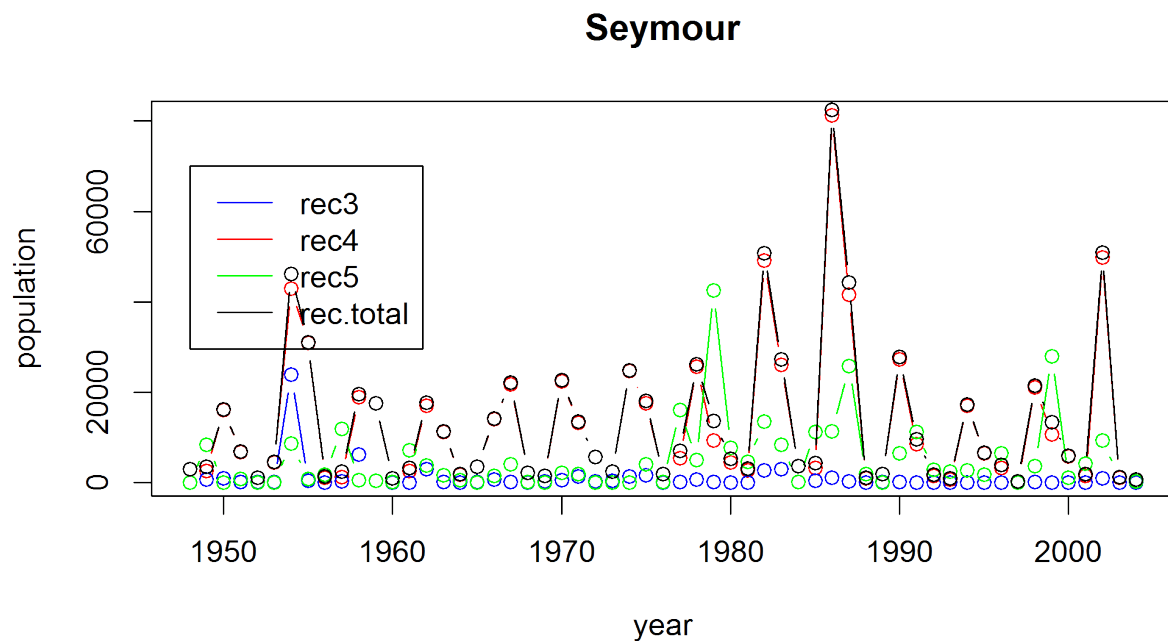
colnames(earlystuart)=c('year','eff','rec3','rec4','rec5','rec.total','r3.per','r4.per','r5.per')

Url=getURL("https://raw.githubusercontent.com/ruizhang-ray/salmon_larkin/master/data/LateStuart.csv")
latestuart=read.table(text=Url,header = FALSE,sep=',')
row.aval=which(!is.na(apply(latestuart,1,sum)))
latestuart=as.matrix(latestuart[row.aval,])
latestuart=cbind(latestuart,apply(latestuart[,3:5],1,sum))
for (i in 1:3){
  latestuart=cbind(latestuart,latestuart[,2+i]/latestuart[,6])
}
colnames(latestuart)=c('year','eff','rec3','rec4','rec5','rec.total','r3.per','r4.per','r5.per')

###-----data display-----

##Seymour
par(mfrow=c(1,1))
plot(seymour[, 'year'],seymour[, 'rec4']/10,type='b',col='red',main='Seymour',xlab='year',ylab='population')
lines(seymour[, 'year'],seymour[, 'rec3'],type='b',col='blue')
lines(seymour[, 'year'],seymour[, 'rec5'],type='b',col='green')
lines(seymour[, 'year'],seymour[, 'rec.total']/10,type='b',col='black')
legend(seymour[1, 'year'],70000,c('rec3','rec4','rec5','rec.total'),col=c('blue','red','green','black'),)

```



```

max(seymour[, 'r3.per'])

```

```

## [1] 0.05174147

```

```
mean(seymour[, 'r3.per'])
```

```
## [1] 0.005044865
```

```
max(seymour[, 'r5.per'])
```

```
## [1] 0.4846439
```

```
mean(seymour[, 'r5.per'])
```

```
## [1] 0.07176551
```

```
## EarlyStuart
```

```
par(mfrow=c(1,1))
```

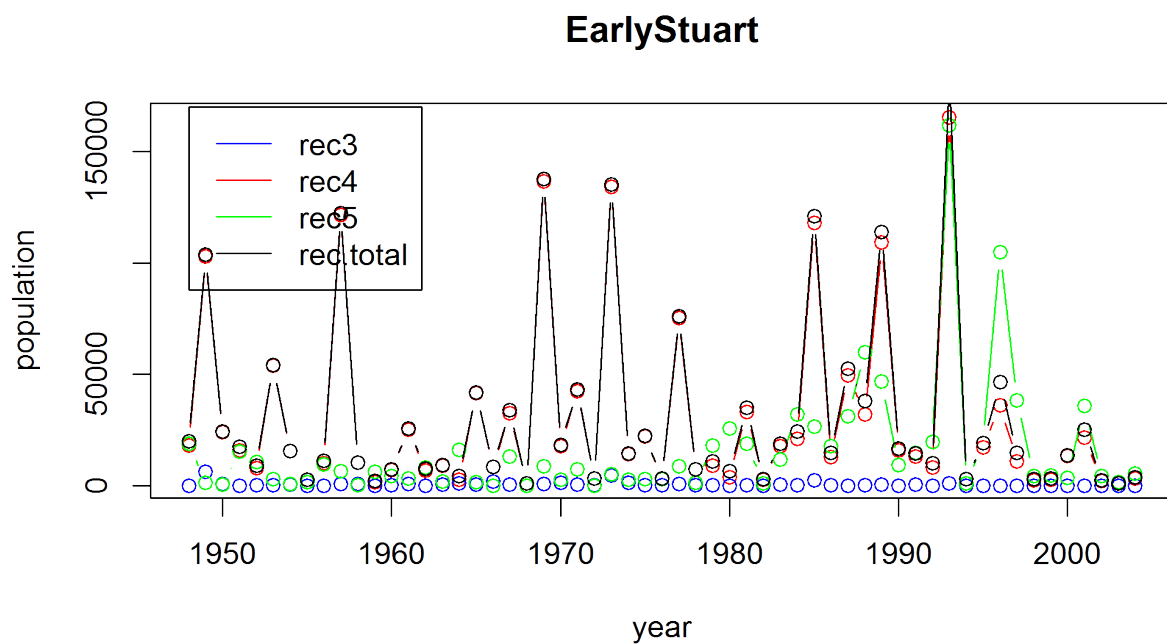
```
plot(earlystuart[, 'year'], earlystuart[, 'rec4']/10, type='b', col='red', main='EarlyStuart', xlab='year', ylab='population')
```

```
lines(earlystuart[, 'year'], earlystuart[, 'rec3'], type='b', col='blue')
```

```
lines(earlystuart[, 'year'], earlystuart[, 'rec5'], type='b', col='green')
```

```
lines(earlystuart[, 'year'], earlystuart[, 'rec.total']/10, type='b', col='black')
```

```
legend(earlystuart[1, 'year'], 170000, c('rec3', 'rec4', 'rec5', 'rec.total'), col=c('blue', 'red', 'green', 'black'))
```



```
max(earlystuart[, 'r3.per'])
```

```
## [1] 0.02394665
```

```
mean(earlystuart[, 'r3.per'])
```

```
## [1] 0.001962232
```

```
max(earlystuart[, 'r5.per'])
```

```
## [1] 0.4022141
```

```
mean(earlystuart[, 'r5.per'])
```

```
## [1] 0.08478021
```

```
## LateStuart
```

```
par(mfrow=c(1,1))
```

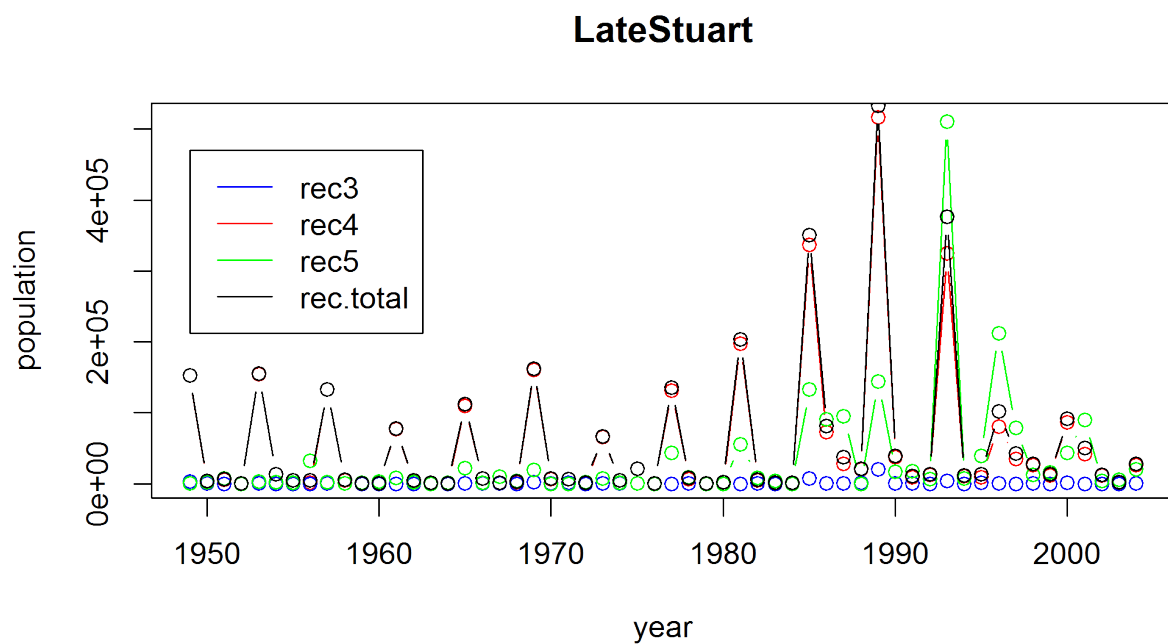
```
plot(lateststuart[, 'year'], lateststuart[, 'rec4']/10, type='b', col='red', main='LateStuart', xlab='year', ylab='population')
```

```
lines(lateststuart[, 'year'], lateststuart[, 'rec3'], type='b', col='blue')
```

```
lines(lateststuart[, 'year'], lateststuart[, 'rec5'], type='b', col='green')
```

```
lines(lateststuart[, 'year'], lateststuart[, 'rec.total']/10, type='b', col='black')
```

```
legend(lateststuart[1, 'year'], 470000, c('rec3', 'rec4', 'rec5', 'rec.total'), col=c('blue', 'red', 'green', 'black'))
```



```
max(lateststuart[, 'r3.per'])
```

```
## [1] 0.02164611
```

```
mean(lateststuart[, 'r3.per'])
```

```
## [1] 0.00384186
```

```
max(lateststuart[, 'r5.per'])
```

```
## [1] 0.7077784
```

```
mean(lateststuart[, 'r5.per'])
```

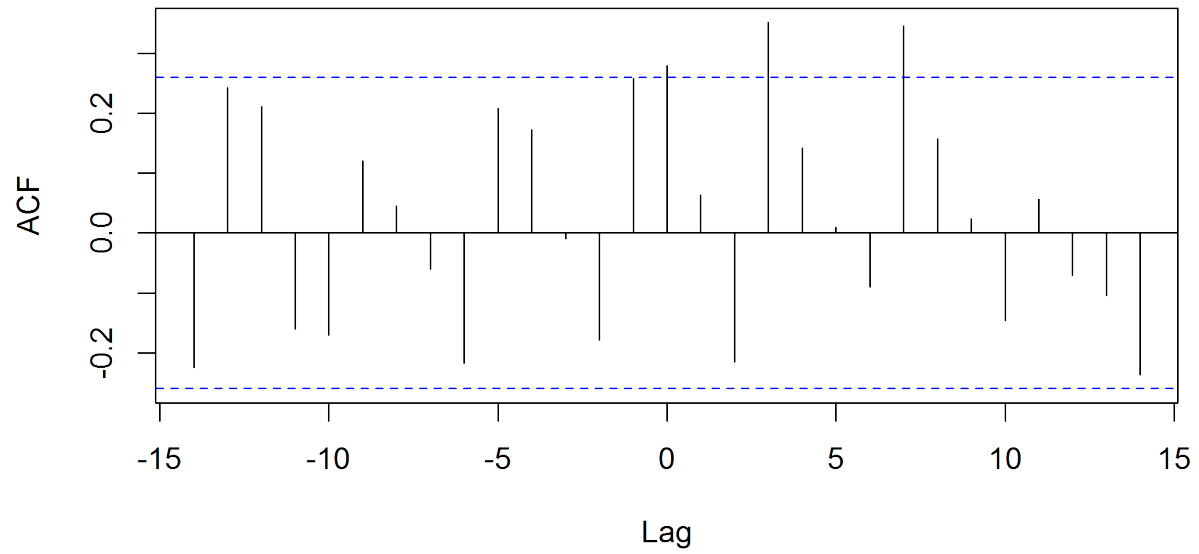
```
## [1] 0.1019538
```

1. To make the plots comfortable to be read, the number of four-year-old adult returns and the total number of the adult returns are divided by 10.
2. Since the trajectory of the red dots and black dots are very the same, we can tell that the populations of these stocks are mainly composed by the four-year-old adult returns .
3. From the plot of Seymour stocks, we can see the five-year-old adult returns are abundant right after the years of the appearances of dominant cohorts. This is predictable because both of them are the off-springs of the last dominant cohorts. However, from the plots of Early Stuart and Late Stuart, we can see most of the five-year-old adult returns are abundant at the same year as the abundance of whole population. There are two possible accounts for this. First, the number of five-year-old adult returns are miscounted. The age of fish can be determined by its size. As we can see, the total population of the fish and the percentage of five-year-old adult returns of Early Stuart and Late Stuart are higher than the Seymour. It is possible that people miscount the large four-year-old fish as the five-year-old fish and this error become severer when the whole population is larger. Second, it is also possible that the fish that return at the same year shares the same living environment, which cause this simultaneous abundance of them.
4. We can notice that in 1967, the dominant cohort of Seymour delayed one year, in 1996 the dominant cohorts in both Early Stuart and Late Stuart shifted to one year early. Phase switching and shifting are common in the population that fluctuates periodically (White et al.[1]). An example of phase switching can be found in the cyclic population of pink salmon (Krkosed et al.[7])

Furthermore, we can investigate the correlation plots.

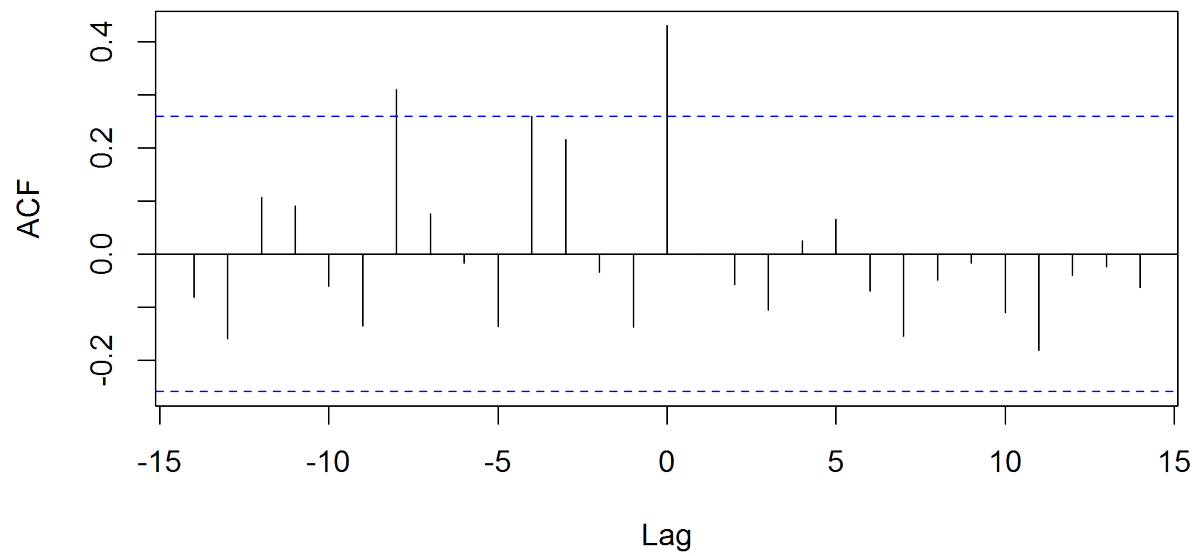
```
##cross correlation of rec4 & rec5  
ccf(seymour[, 'rec4'], seymour[, 'rec5'])
```

### seymour[, "rec4"] & seymour[, "rec5"]

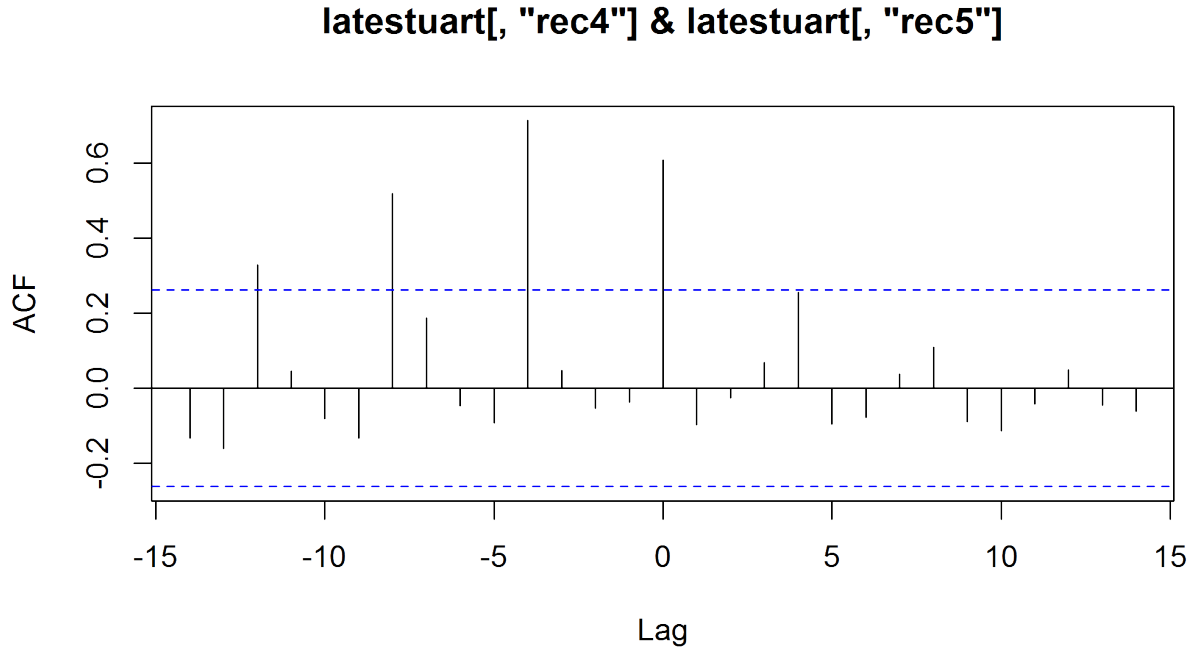


```
ccf(earlystuart[, 'rec4'], earlystuart[, 'rec5'])
```

### earlystuart[, "rec4"] & earlystuart[, "rec5"]



```
ccf(latestuart[, 'rec4'], latestuart[, 'rec5'])
```



*# evidences that there is strong positive correlation between the recruits with same  
# returning year instead of the same birth year*

1. In these three stocks, the populations of five-year-old adult returns and four-year-old adult returns at the same years are highly positive correlated.
2. In the plots of Early Stuart and Late Stuart, we can see the population of five-year-old adult returns are highly positive correlated with the population of four-year-old adult returns at the four years before.

### 3 Model

Process model:

$$\begin{aligned}
 F_t &= \alpha(R_{t,4} + R_{t,5}) \exp(\gamma_0 R_{t,4} + \omega_0 R_{t,5} + \\
 &\quad \gamma_1 R_{t-1,4} + \omega_1 R_{t-1,5} + \\
 &\quad \gamma_2 R_{t-2,4} + \omega_2 R_{t-2,5} + \\
 &\quad \gamma_3 R_{t-3,4} + \omega_3 R_{t-3,5} + \epsilon_t) \quad [LarkinRicker] \\
 \epsilon_t &\sim N(0, \sigma^2) \\
 p_t &= 1 / (1 + \exp(a + b R_{t,4} + c R_{t,5})) \\
 X_{t,4} &= p_t F_t \\
 X_{t,5} &= (1 - p_t) F_t \\
 R_{t,4} &= X_{t-4,4} \\
 R_{t,5} &= s X_{t-5,5}
 \end{aligned}$$



Measurement model:

$$\begin{aligned} Rec_4 &\sim \text{LogNormal}(\log(\phi_1 R_{t,4}), \psi_1) \\ Rec_5 &\sim \text{LogNormal}(\log(\phi_2 R_{t,5}), \psi_2) \end{aligned}$$

1. In order to reduce the redundancy, effective female spawner (eff) is not included. Moreover, the three-year-old adult return (rec3) is not included as well because of its small amount.
2.  $F_t$  is the newborn generation which is born at year  $t$ .  $R_{t,k}$  is the  $k$ -year-old adult return that reaches spawning ground at year  $t$ .  $X_{t,k}$  is part of  $F_t$ , which will return spawning ground  $k$  years later.  $s$  is the survival rate from four years old to five years old.
3.  $p_t$  is the proportion that  $X_{t,4}$  takes among  $F_t$ . Mathematically, it is a logistic regression, of which the outcome is between 0 to 1. Biologically, it is assumed that the proportion that the fry will return four years later is related to the number of their parents of different ages.
4.  $\alpha$  is often referred as productive rate (Peterman et al.[6], Myers et al.[5]). However here, it is the product of productive rate and the accumulative survival rate till four years old.
5. This is an age-structure model with only explicitly specifying age-4 and age-5 population. Moreover, this is a state space model with non-linearity, which makes the iterated filtering a good algorithm for analysis.

## 4 Model comparison and analysis

In this part, a few tests will be done to analyze the significance of the parameters. Furthermore, the Larkin Ricker model will be compared with the standard Ricker model, Beverton-holt model, Beverton-holt model with delayed density dependence (Larkin version) and AMAR model.

Beverton-holt with delayed density dependence:

$$F_t = \frac{\alpha(R_{t,4} + R_{t,5})}{\gamma_0 R_{t,4} + \gamma_1 R_{t-1,4} + \gamma_2 R_{t-2,4} + \gamma_3 R_{t-3,4} + \omega_0 R_{t,5} + \omega_1 R_{t-1,5} + \omega_2 R_{t-2,5} + \omega_3 R_{t-3,5}} \exp(\epsilon_t)$$

Figure 1 is the results (b:Beverton-holt):

1. Model 1 is the model of Null Hypothesis that the number of five-year-old adult returns are not significant and only two-year delayed density dependence is needed.
2. By comparing 2 with 1, we can know that the number of five-year-old adult returns are not significant for computing the proportion of four-year-old adult return.
3. By comparing 3 with 1, we can know that the delayed density dependence on the population of five-year-old adult returns are not significant.
4. We can see all the Ricker and Beverton-holt model outperform the ARMA model. Choosing by AIC, the best model is model 11, which includes three-year delay density dependence. This model has the log likelihood much higher than the others and is the only model that the Beverton-holt version are not close to the Ricker version.

Following is the R codes that building pomp object for the model 11:

Model	log Likelihood	Parameters	AIC	constraint
1	301.36	18	-566.72	$\gamma_3=0, \omega_3=0$
1b	301.36	18	-566.72	$\gamma_3=0, \omega_3=0$
2	301.32	17	-568.64	$\gamma_3=0, \omega_3=0, c=0$
2b	301.12	17	-568.24	$\gamma_3=0, \omega_3=0, c=0$
3	301.7	21	-561.4	$\gamma_3=0, \omega_3=0$
3b	301.68	21	-561.36	$\gamma_3=0, \omega_3=0$
5	299.95	16	-567.9	$\gamma_1=0, \gamma_3=0, \omega_3=0, c=0$
5b	299.88	16	-567.76	$\gamma_1=0, \gamma_3=0, \omega_3=0, c=0$
6 (Ricker)	296.46	15	-562.92	$\gamma_1=0, \omega_3=0, c=0$
6b (Beveton-holt)	296.45	15	-562.9	$\gamma_1=0, \omega_3=0, c=0$
7	295.79	15	-561.58	$\gamma_1 < 0, \gamma_1=0, \omega_3=0, c=0$
7b	295.8	15	-561.6	$\gamma_1 > 0, \gamma_1=0, \omega_3=0, c=0$
8	298.98	16	-565.96	$\gamma_2=0, \omega_3=0, c=0$
8b	298.82	16	-565.64	$\gamma_2=0, \omega_3=0, c=0$
9	295.38	16	-558.76	$\gamma_1 < 0, \gamma_1=0, \gamma_2 < 0, \gamma_3=0, \omega_3=0, c=0$
9b	295.47	16	-558.94	$\gamma_1 > 0, \gamma_1=0, \gamma_2 > 0, \gamma_3=0, \omega_3=0, c=0$
10	299.06	16	-566.12	$\gamma_1 < 0, \gamma_1=0, \gamma_3=0, \omega_3=0, c=0$
10b	298.29	16	-564.58	$\gamma_1 > 0, \gamma_1=0, \gamma_3=0, \omega_3=0, c=0$
11	307.81	18	-579.62	$\omega_3=0, c=0$
11b	302	18	-568	$\omega_3=0, c=0$
12	301.6	17	-569.2	$\gamma_1=\omega_1, \gamma_1=\omega_1, \gamma_2=\omega_2, \gamma_3=\omega_3, c=0$
12b	301.2	17	-568.4	$\gamma_1=\omega_1, \gamma_1=\omega_1, \gamma_2=\omega_2, \gamma_3=\omega_3, c=0$
ARMA (5, 1)+ARMA (0, 0)	195.055	8	-374.11	rec4:ARMA (5, 1), rec5:ARMA (0, 0)
SARMA (p=1, q=2, P=1, Q=2)+SARMA (0, 0, 0, 2)	200.4295	10	-380.8592	
SARMA (p=1, q=2, P=1, Q=2, period=2)+SARMA (2, 2, 0, 5, 5)	207.8715	17	-381.7426	

Figure 1: AIC Table

```
library(RCurl)
Url=getURL("https://raw.githubusercontent.com/ruizhang-ray/salmon_larkin/master/data/Seymour.csv")
seymour=read.table(text=Url,header = FALSE,sep=',')
row.aval=which(!is.nan(apply(seymour,1,sum)))
seymour=as.matrix(seymour[row.aval,])
seymour=cbind(seymour,apply(seymour[,3:5],1,sum))
for (i in 1:3){
  seymour=cbind(seymour,seymour[,2+i]/seymour[,6])
}
colnames(seymour)=c('year','eff','rec3','rec4','rec5','rec.total','r3.per','r4.per','r5.per')
##-----pomp-----
library(pomp)
library(ggplot2)
data=data.frame(seymour)
##normalizing

#data$eff=data$eff/sd(data$eff)
data$rec3=data$rec3/max(data$rec.total)
data$rec4=data$rec4/max(data$rec.total)
data$rec5=data$rec5/max(data$rec.total)
data$rec.total=data$rec.total/max(data$rec.total)
```

Notice that I normalize adult returns data by dividing them by the max total adult return.

```
salmon_statenames <-c("F0","F1","F2","F3","F4","F5",
                      "R04","R14","R24","R34","R44","R54",
                      "R05","R15","R25","R35","R45","R55")
salmon_rp_names <- c("alpha",
                     "gamma0","gamma1","gamma2","gamma3",
```

```

        "a","b","c","s5","sigma","phi1","phi2","psi1","psi2")
salmon_ivp_names <- c("F0_0","F1_0","F2_0","F3_0","F4_0","F5_0",
        "R04_0","R14_0","R24_0","R34_0","R44_0","R54_0",
        "R05_0","R15_0","R25_0","R35_0","R45_0","R55_0")
salmon_paramnames <- c(salmon_rp_names, salmon_ivp_names)
salmon_rproc <- '
double  epsilon=rnorm(0,sigma);
double  y=0;
double  X4=0;
double  X5=0;

F5=F4;
F4=F3;
F3=F2;
F2=F1;
F1=F0;

R54=R44;
R44=R34;
R34=R24;
R24=R14;
R14=R04;

R55=R45;
R45=R35;
R35=R25;
R25=R15;
R15=R05;

X4=(1/(1+exp(a+b*R44+c*R45)))*F4;
X5=(1-1/(1+exp(a+b*R54+c*R55)))*F5;

R04=X4;
R05=s5*X5;
F0=alpha*(R04+R05)*exp(gamma0*R04+gamma1*R14+gamma2*R24+gamma3*R34+epsilon);
'
salmon_initializer <- "
R04=R04_0;
R14=R14_0;
R24=R24_0;
R34=R34_0;
R44=R44_0;
R54=R54_0;
R05=R05_0;
R15=R15_0;
R25=R25_0;
R35=R35_0;
R45=R45_0;
R55=R55_0;
F0=F0_0;
F1=F1_0;
F2=F2_0;

```

```

F3=F3_0;
F4=F4_0;
F5=F5_0;
"
salmon_rmeasure <-"
rec4=rlnorm(log(phi1*R04),psi1);
rec5=rlnorm(log(phi2*R05),psi2);
"
salmon_dmeasure <-'
double tol = 0.00001;
lik=dlnorm(rec4+tol,log(phi1*R04+tol),psi1,1);
lik+= dlnorm(rec5+tol,log(phi2*R05+tol),psi2,1);
if (!(lik>-999)&(lik<999)) lik=-9999;
lik=(give_log) ? lik : exp(lik);
'

salmon_toEst <-"
Tsigma=log(sigma);
Ts5=logit(s5);
Talpha=log(alpha);
TF0_0=log(F0_0);
TF1_0=log(F1_0);
TF2_0=log(F2_0);
TF3_0=log(F3_0);
TF4_0=log(F4_0);
TF5_0=log(F5_0);
Tphi1=log(phi1);
Tphi2=log(phi2);
Tpsi1=log(psi1);
Tpsi2=log(psi2);

"
salmon_fromEst <-"
Tsigma=exp(sigma);
Ts5=expit(s5);
Talpha=exp(alpha);
TF0_0=exp(F0_0);
TF1_0=exp(F1_0);
TF2_0=exp(F2_0);
TF3_0=exp(F3_0);
TF4_0=exp(F4_0);
TF5_0=exp(F5_0);
Tphi1=exp(phi1);
Tphi2=exp(phi2);
Tpsi1=exp(psi1);
Tpsi2=exp(psi2);

"

n=dim(data)[1]
salmon<- pomp(data=data.frame(rec4=data$rec4[6:n]+data$rec3[6:n],

```

```

        rec5=data$rec5[6:n],year=data[6:n,'year']),
    statenames=salmon_statenames,
    paramnames=salmon_paramnames,
    times="year",
    t0=data$year[5],
    rmeasure=Csnippet(salmon_rmeasure),
    dmeasure=Csnippet(salmon_dmeasure),
    rprocess=discrete.time.sim(step.fun=Csnippet(salmon_rproc),
                                delta.t=1),
    initializer=Csnippet(salmon_initializer),
    toEstimationScale=Csnippet(salmon_toEst),
    fromEstimationScale=Csnippet(salmon_fromEst)
)
salmon_ivp_fix=c(0,0,data$rec5[1:5],0,data$rec3[1:5]+data$rec4[1:5])
names(salmon_ivp_fix)=c("F5_0","R55_0","R45_0","R35_0","R25_0","R15_0","R05_0",
                        "R54_0","R44_0","R34_0","R24_0","R14_0","R04_0")

```

We can use iterated filtering to search the MLEs. Following is the code with parallel computing.

```

run_level=3
CORES<-10 ##update for flux
salmon_Np=c(100,5000,60000)
salmon_Nmif=c(10,200,350)
salmon_Nglobal=c(10,20,120)
salmon_Neval=c(5,20,120)
JOBS=salmon_Nglobal[run_level]
nodefile <- Sys.getenv("PBS_NODEFILE")
CLUSTER <- nchar(nodefile)>0
CLUSTER
if (CLUSTER) {
  require(doSNOW)
  hostlist <- tail(scan(nodefile,what=character(0)),-1)
  cl <- makeCluster(hostlist,type='SOCK')
  cl
  registerDoSNOW(cl)
} else {
  require(doParallel)
  registerDoParallel(CORES)
}
library(pomp)
tic <- Sys.time()
#cat("session1")
mpar <- foreach(
  i=1:JOBS,
  .packages=c('pomp'),
  .inorder=FALSE,.combine=c) %dopar% {
  Sys.sleep(i*.1)
  setwd("~/Larkin/June_27/model11")
  NMIF<-salmon_Nmif[run_level]
  NP<-salmon_Np[run_level]
  source("salmon.r")
  salmon_box <-rbind(
    alpha=c(0.4,6),

```

```

    gamma0=c(-4,4),
    gamma1=c(-4,4),
    gamma2=c(-4,4),
    gamma3=c(-4,4),
    a=c(-6,-2),
    b=c(-8,-3),
    s5=c(0.2,0.6),
    sigma=c(0.05,0.8),
    phi1=c(0.8,1.6),
    phi2=c(0.2,3),
    psi1=c(0.55,0.7),
    psi2=c(1.8,1.9),
    F4_0=c(0.01,0.06),
    F3_0=c(0.01,0.22),
    F2_0=c(0.4,1),
    F1_0=c(0.2,1),
    F0_0=c(0.04,0.08)
  )
  salmon_fix=c(salmon_ivp_fix,c=0)
  salmon.rw.sd <- c(rep(0.02,13),rep(0.1,5))
  names(salmon.rw.sd) <- row.names(salmon_box)
  salmon_cooling.fraction.50 <- 0.5

  set.seed(1129+i)
  Sys.sleep(i*0.1)
  th.draw=c(apply(salmon_box,1,function(x) runif(1,x[1],x[2])),salmon_fix)
  cat("session2")
  m<-try(mif2(salmon,
             start=th.draw,
             Np=NP,
             Nmif=NMIF,
             cooling.type="geometric",
             cooling.fraction.50=salmon_cooling.fraction.50,
             transform=TRUE,
             rw.sd=salmon.rw.sd
           ))
  m
}
L.if <- foreach(i=1:JOBS,.packages='pomp',
               .combine=rbind,.inorder=FALSE) %dopar%
{
  Sys.sleep(i*.1)
  setwd("~/Larkin/June_27/model11")
  source("salmon.r")
  set.seed(1993+i)
  Sys.sleep(i*.1)
  logmeanexp(
    replicate(salmon_Neval[run_level],
              logLik(pfilter(salmon,params=coef(mpar[[i]]),Np=salmon_Np[run_level]))
    ),
    se=TRUE)
}
setwd("~/Larkin/June_27/model11") ##turn on for flux

```

```

source("salmon.r")

toc <- Sys.time()
print(toc-tic)

r.if=data.frame(logLik=L.if[,1],logLik_se=L.if[,2],t(sapply(mpar,coef)))
summary(r.if$logLik,digits=5)
#18 parameters
pairs(~logLik+s5+alpha+gamma0+gamma1+gamma2+gamma3,
      data=subset(r.if,logLik>max(logLik)-20))
pairs(~logLik+a+b+c,data=subset(r.if,logLik>max(logLik)-20))
pairs(~logLik+alpha+gamma0+gamma1+gamma2+gamma3+sigma+phi1+phi2+psi1+psi2+s5,
      data=subset(r.if,logLik>max(logLik)-20))
plot(mpar)

if (run_level>1)
  write.table(r.if,file="salmon_global_model11_params.csv",
             append=TRUE,col.names=TRUE,row.names=FALSE,sep=',')
save.image(file="result.rda")
if (CLUSTER) stopCluster(cl)

```

It takes over 4 hours for 60 cores to compute the following results.

```

r.if=read.csv("salmon_global_model11_params.csv")
summary(r.if$logLik,digits=5)

```

```

##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 297.82  304.32  306.15  304.96  307.08  307.81

```

```

r.if[which.max(r.if$logLik),]

```

```

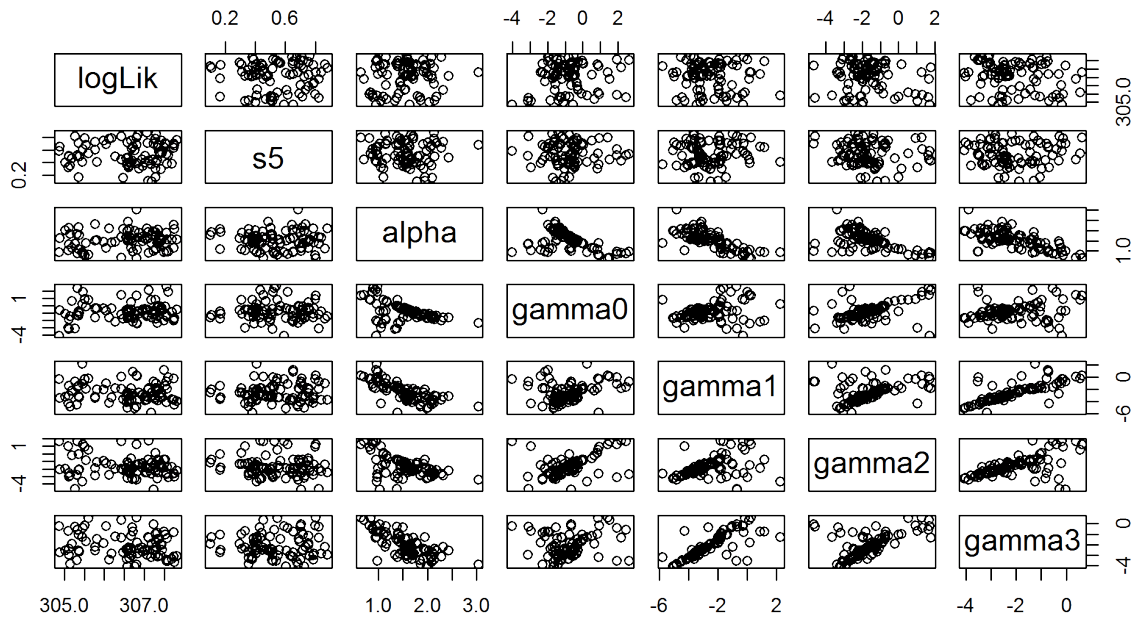
##      logLik  logLik_se  alpha  gamma0  gamma1  gamma2  gamma3
## 23 307.8107 0.0003138325 1.572604 -0.5334972 -3.876305 -2.189927 -3.07504
##           a           b           s5          sigma          phi1          phi2          psi1
## 23 -2.581353 -8.611213 0.7625507 0.01418107 2.725208 1.515285 0.6294496
##           psi2          F4_0          F3_0          F2_0          F1_0          F0_0 F5_0
## 23 2.487886 6.066661e-05 0.02172296 0.08065268 0.07572891 0.005889906 0
##      R55_0 R45_0          R35_0 R25_0          R15_0 R05_0 R54_0          R44_0
## 23      0      0 0.01013398      0 0.001035837      0      0 0.0359729
##           R34_0          R24_0          R14_0          R04_0 c
## 23 0.03196055 0.1963833 0.08258678 0.01364418 0

```

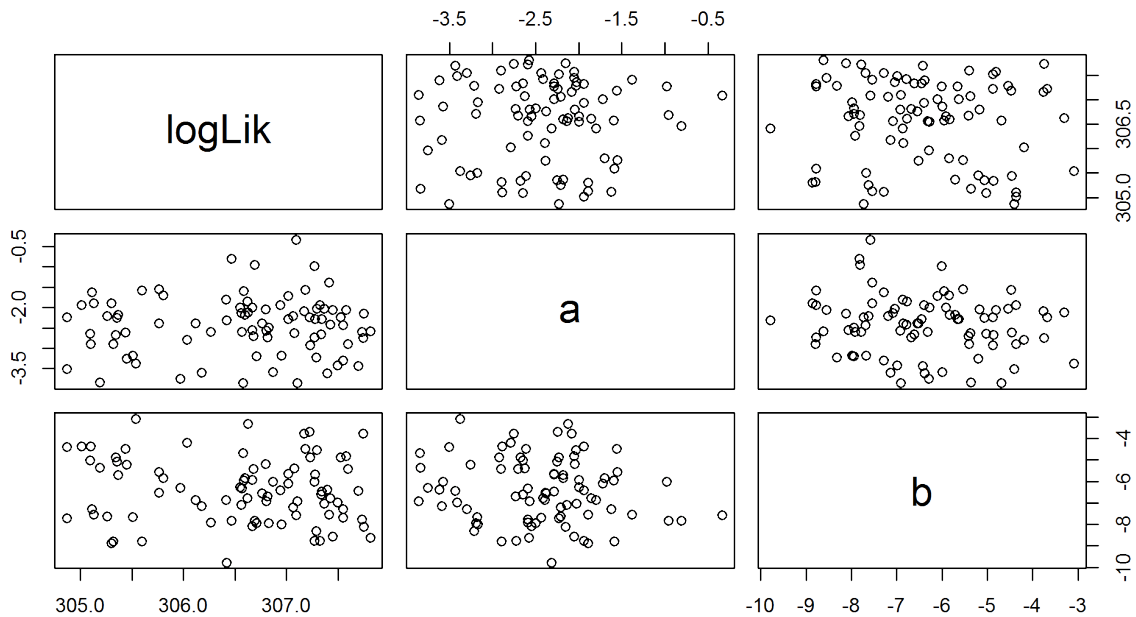
```

pairs(~logLik+s5+alpha+gamma0+gamma1+gamma2+gamma3,
      data=subset(r.if,logLik>max(logLik)-3))

```

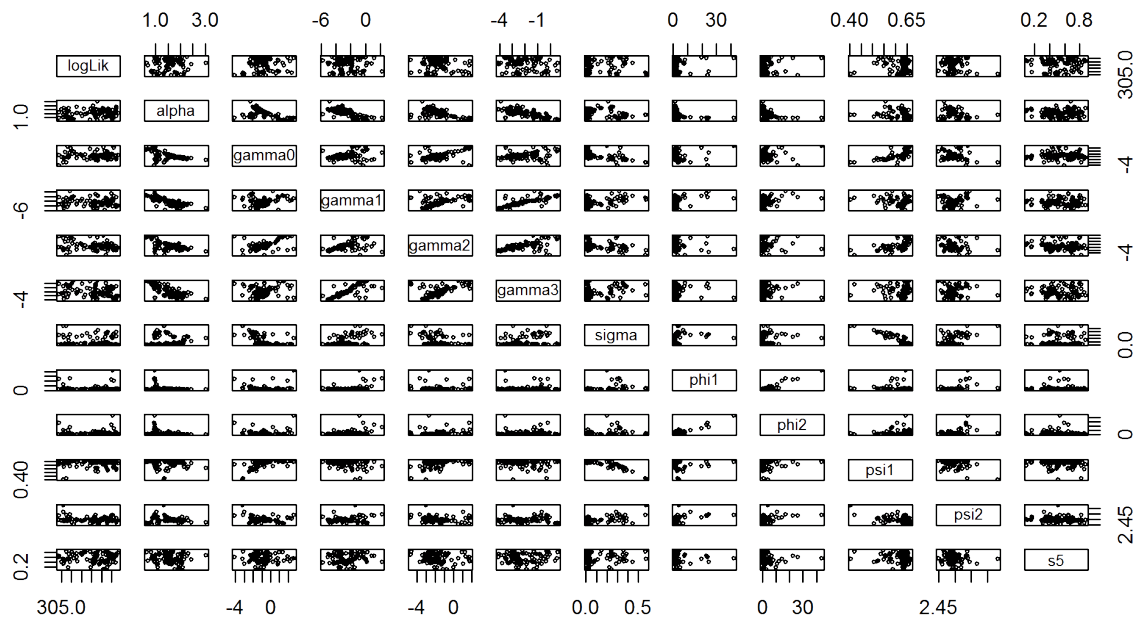


```
pairs(~logLik+a+b,data=subset(r.if,logLik>max(logLik)-3))
```

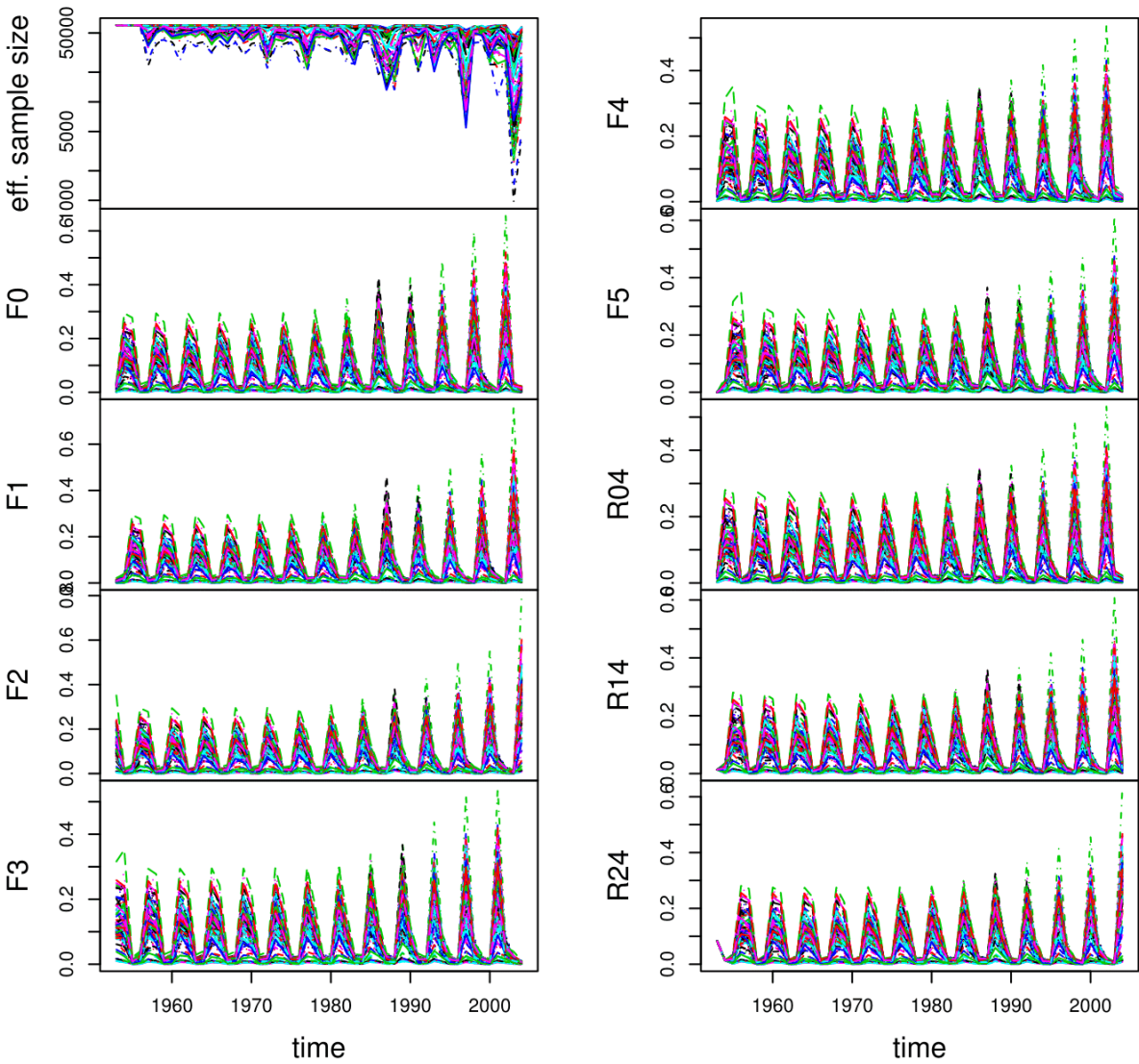


```
pairs(~logLik+alpha+gamma0+gamma1+gamma2+gamma3+sigma+phi1+phi2+psi1+psi2+s5,
      data=subset(r.if,logLik>max(logLik)-3),cex=0.4)
```

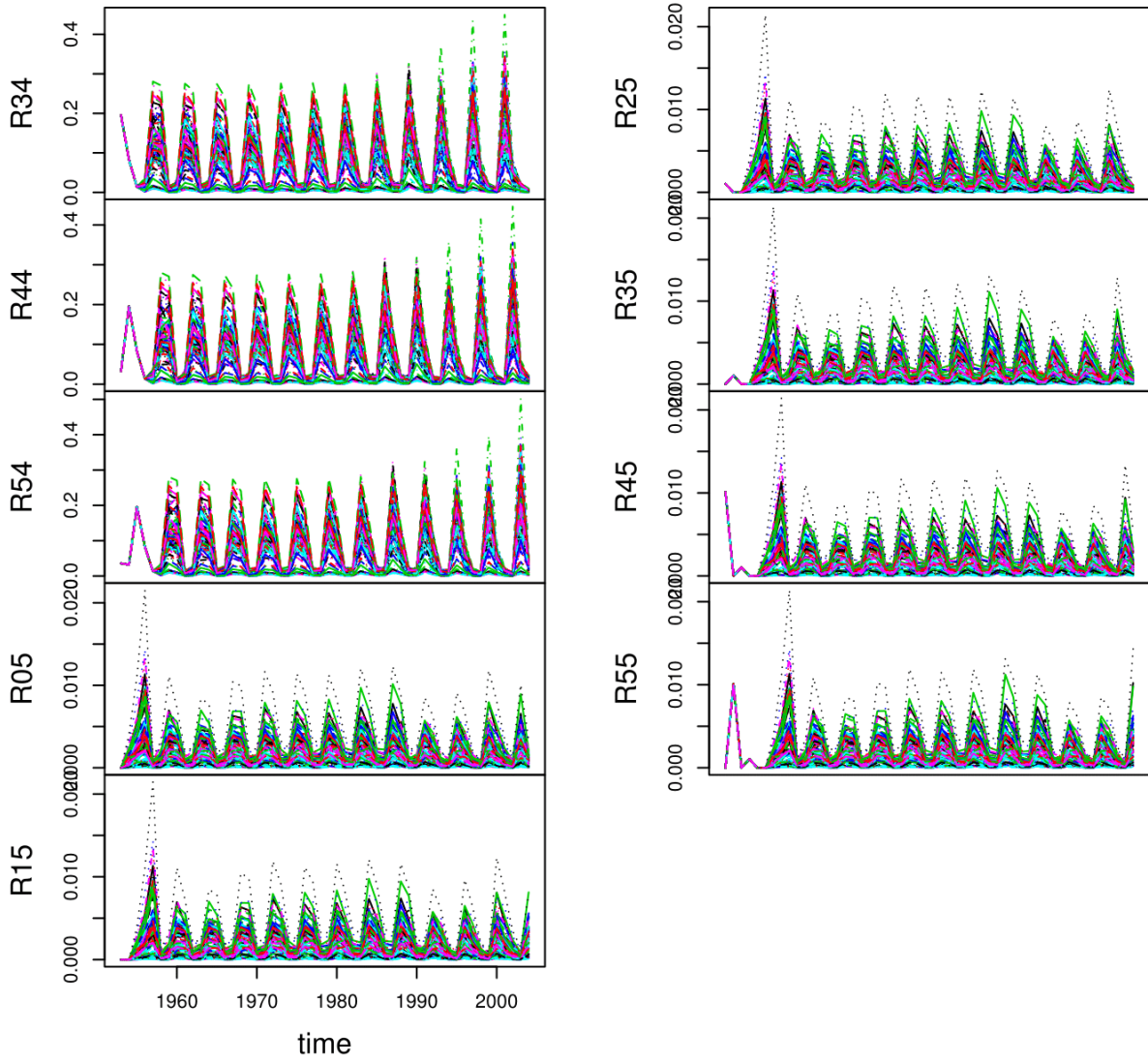




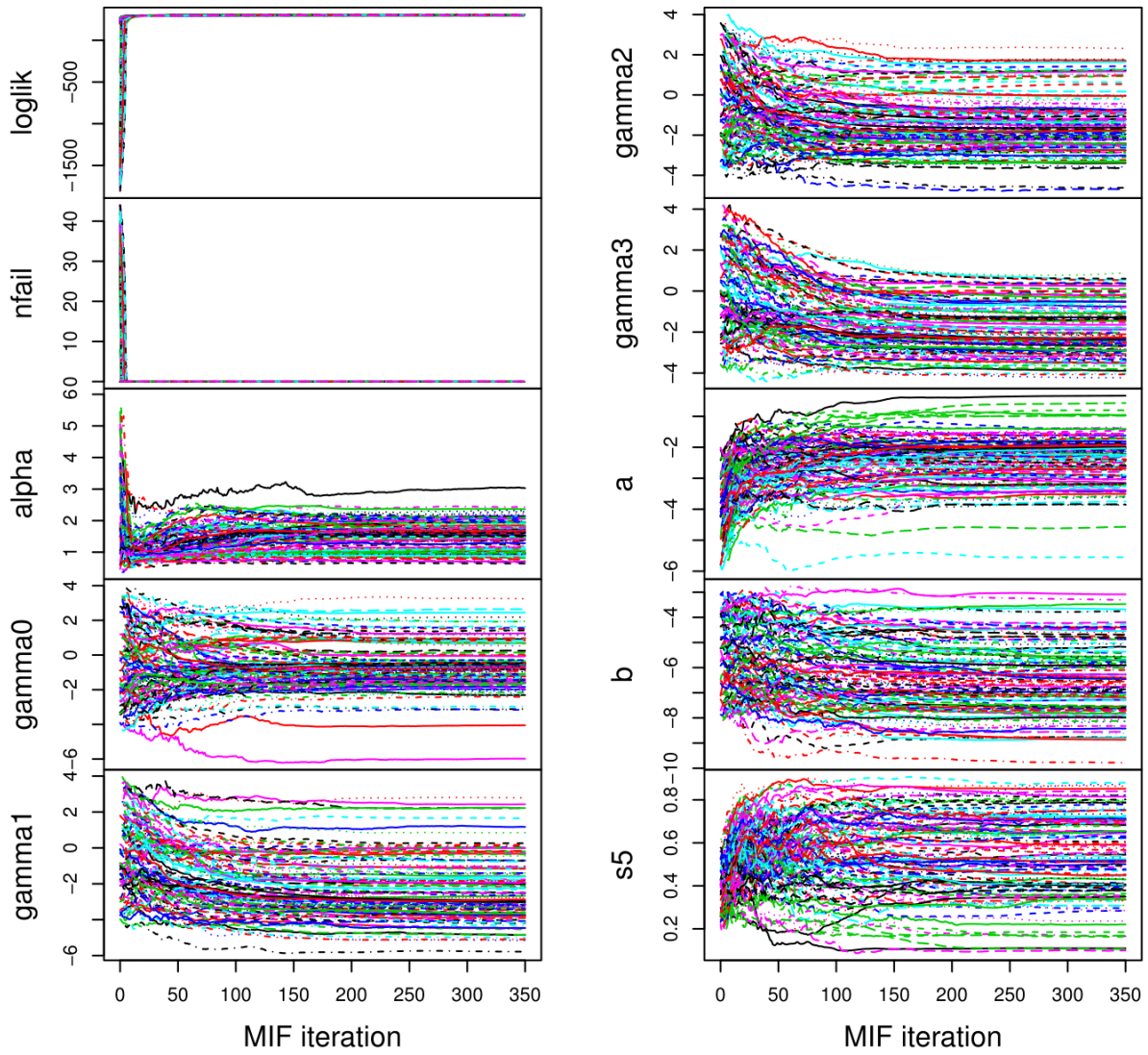
Filter diagnostics (last iteration)



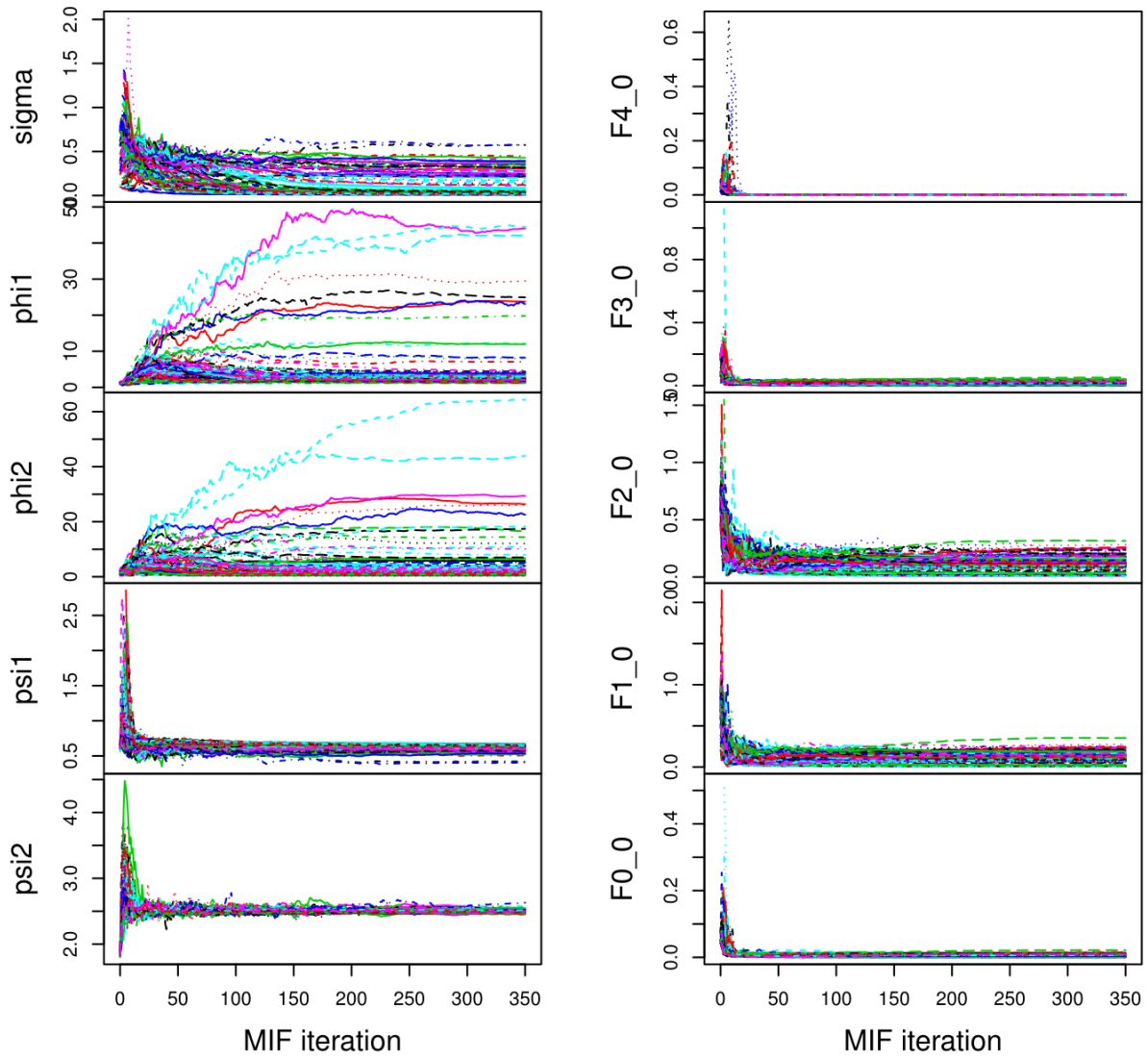
# Filter diagnostics (last iteration)



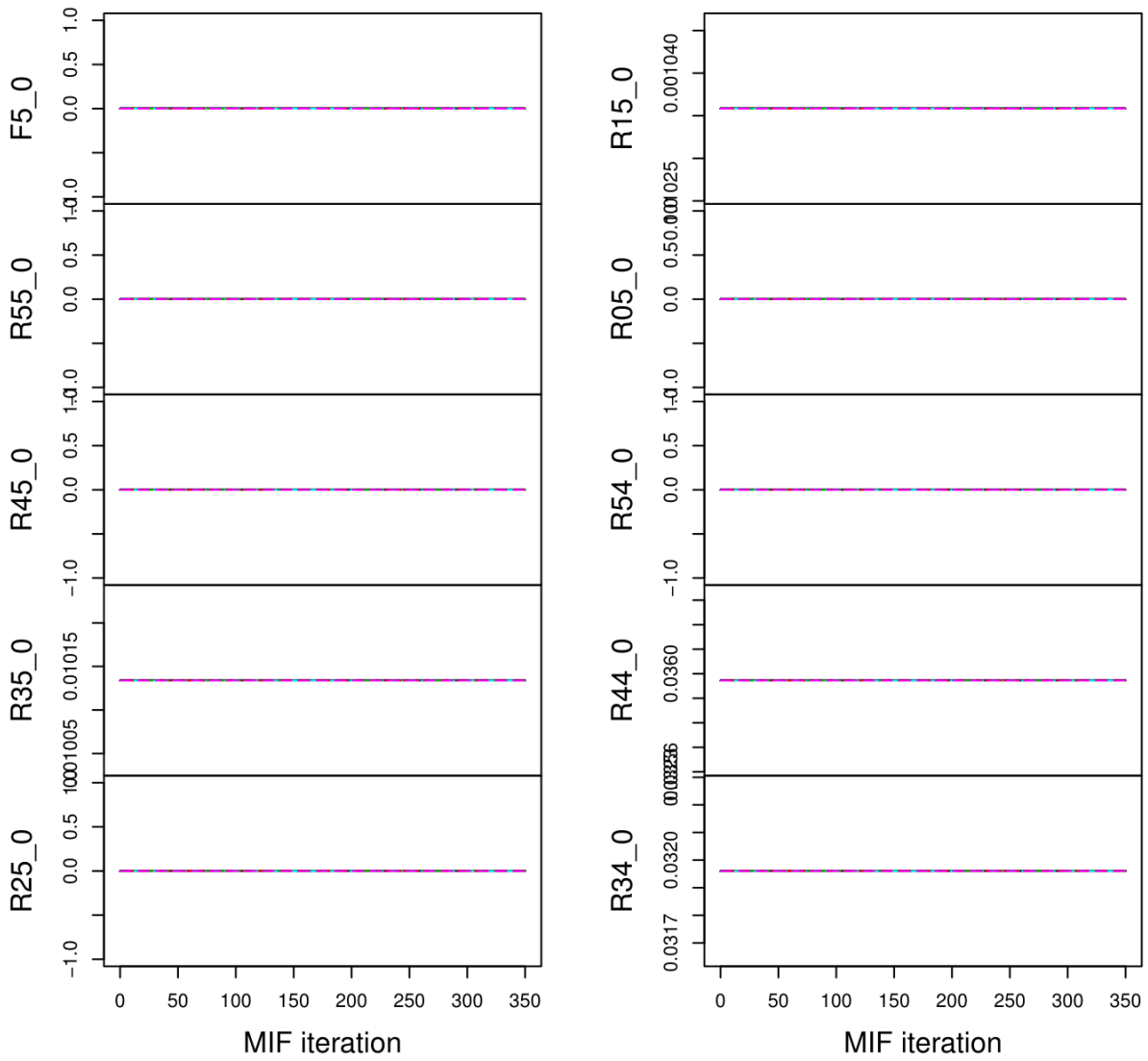
# MIF2 convergence diagnostics



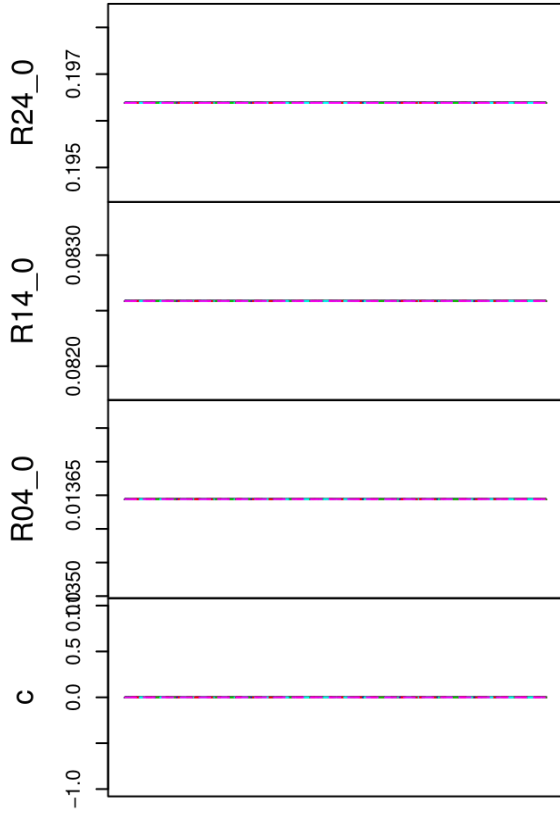
# MIF2 convergence diagnostics



# MIF2 convergence diagnostics



## MIF2 convergence diagnostics



## 5 Conclusion

1. Evidence shows that the data favors the model that with negative delayed density dependence of 3 years. Delayed density dependence can be caused by cannibalism by the older cohorts and the transimission of disease and salmon louse from the olders (Bjornstad et al.[8]). The same research has been done for pink salmon, fluatuated population with period two year (Krkosek et al.[7]).
2. The existence of interaction between cohorts are supported by the data. It seems the cohort resonance induced by the environmental unstability is not the only reason that the population of salmon fish shows cohort dominance. With negative delayed density dependence of 3 years, one extremely abundance cohort will suppress all other cohorts in the following cycle, which leads to the cohort dominance.
3. We can notice that  $\psi_2$  is large, which means model can not find any mechanism for the pattern of

five-year-old adult returns since there are lots of noise.

4. The rationalities of the parameters with specific biological meanings, such as  $s$  and  $\alpha$ , need further discussion.
5. So far, the model still fails to consider the effects of living condition shared by the fish which return at the same year. The effects of delayed density dependence for the salmon with the same birth year are the same. However, even though the salmon are born at the same year, the fish returns at four years old and the one returns at five years old are actually living in different living condition. For an individual fish, it shares living environment with the fish of same birth year at the early stage of life and shares living environment with the fish of same return year at the later stage of the life. Only the former one is considered in my model.

## 6 Supplement

### 6.1 Significance of $b$

By roughly computation of the MLE, we can know,  $p_t$  are actually very closed to 1. Model 13 is run to test the significance of  $b$ , to see whether a fixed  $p_t$  is enough.

13	306.33	17	-578.66	$\omega_3=0, c=0, b=0$
13b	294.76	17	-555.52	$\omega_3=0, c=0, b=0$

Figure 2: model13

We can see that,  $b$  is significant in Beveton-holt but not significant in Larkin Ricker.

### 6.2 result on the Early Stuart stock

Now we can apply the model to other stocks, to see whether we can get the similar result. Due to the time limit, only some models are done for the Early Stuart stock.

Model	log Likelihood	Parameters	AIC	constraint
1	274.13	18	-512.26	$\gamma_3=0, \omega_3=0$
1b	273.89	18	-511.78	$\gamma_3=0, \omega_3=0$
2	274.76	17	-515.52	$\gamma_3=0, \omega_3=0, c=0$
2b	274.03	17	-514.06	$\gamma_3=0, \omega_3=0, c=0$
11	272.87	18	-509.74	$\omega_3=0, c=0$
11b	276.5	18	-517	$\omega_3=0, c=0$
ARMA (3, 3) + ARMA (3, 3)	171.67	14	-315.34	
SARMA (p=0, q=0, P=5, Q=4, period=5) + SARMA (0, 3, 0, 2, 3)	199.735	16	-367.47	

Figure 3: AIC in Early Stuart stock

Still, the Ricker model and Beveton-Holt still outperform the ARMA model. The best model for Seymour stock is no longer the best in Early Stuart. The best one is the Beveton-holt with four-year delay density dependence. Although the AIC of the best one are just little higher than the others, we can still conclude that the interactions between cohorts exist. More computation is needed since there are optimization errors (the small model has larger likelihood than the big one).



## 7 References

1. *Stochastic models reveal conditions for cyclic dominance in sockeye salmon populations*, J. Wilson White et al.
2. *Cohort resonance: a significant component of fluctuations in recruitment, egg production, and catch of fished populations*, Louis W. Botsford et al.
3. *Equation-free mechanistic ecosystem forecasting using empirical dynamic modeling*, Hao Ye et al.
4. *Patterns of covariation in survival rates of British Columbian and Alaskan sockeye salmon (*Oncorhynchus nerka*) stocks*, Randall M. Peterman et al.
5. *Simple dynamics underlie sockeye salmon (*Oncorhynchus nerka*) cycles*, Ransom A. Myers et al.
6. *A widespread decrease in productivity of sockeye salmon (*Oncorhynchus nerka*) populations in western North America*, Randall M. Peterman et al.
7. *Cycles, stochasticity and density dependence in pink salmon population dynamics*, Martin Krkosek et al.
8. *Trends and cohort resonant effects in age-structured populations*, Ottar N. Bjornstad et al.