Vignette for the Bayesian Surplus Production Model with Catch-Reslience Method C_{msy}
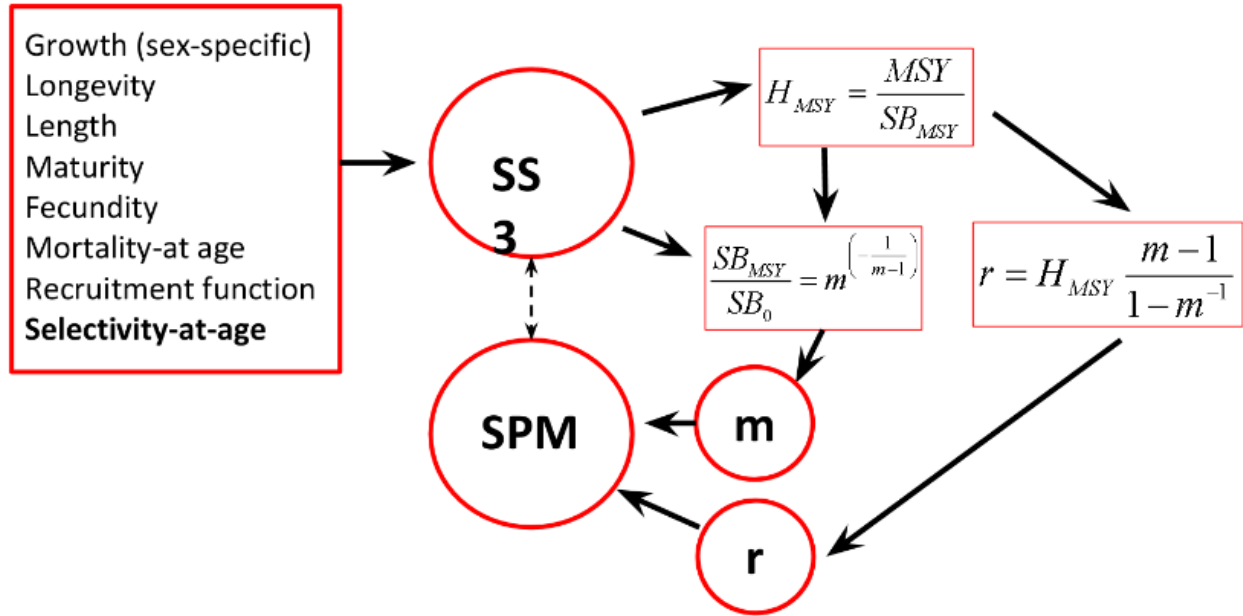
Part II - BSP Model Structure

Henning Winker, Felipe Carvalho

Contents

This document provides an overview of the structure & function of the Bayesian Surplus Production model. We illustrate a test case using the Shortfin Mako shark *Isurus oxyrinchus*, with reproducible code and explanatory formulae. The following provides a schematic of the model structure.

BSPSP Model Development

We seek to improve the estimation properties of Bayesian state-space surplus production models (SPMs) by building on previous formulations by Pella and Tomlinson (1969), Gilbert (Gilbert 1992, Wang et al. 2014) and Fletcher (1978, c.f. Thorson et al. 2012a). An advantage of the proposed generalization is that it links surplus production models more directly to conventional age-structured model formulations (e.g. SS3; Methot and Wetzel 2013) . First we focus on the surplus production function of the generalized three parameter SPM by Pella and Tomlinson (1969):

$$SP_t = \frac{r}{m-1} B_{t-1}(1 - (\frac{B_{t-1}}{K}^{m-1}))$$

where r is the intrinsic rate of population increase at time t, K is the unfished biomass and m is a shape parameter that determines at which B/K ratio maximum surplus production is attained. If the shape parameter is $m = 2$, the model reduces to the Schafer form, with the surplus production g(Bt) attaining MSY at exactly $K/2$. If $0 < m < 2$, (Bt) attains MSY at depletion levels smaller than $K/2$ and vice versa. The Pella-Tomlinson model reduces to a Fox model if m approaches one (m=1) resulting in maximum surplus production at $\sim 0.37K$, but there is no solution for the exact Fox SP with $m = 1$. $B_M SY$ is given by:

$$B_{MSY} = K m^{\frac{-1}{m-1}}$$

and the corresponding harvest rate at MSY H_{MSY} is:

$$H_{MSY} = \frac{r}{m-1}(1 - \frac{1}{m})$$

Where the harvest rate H is defined here as the ratio of:

$$H = \frac{C}{B}$$

Where C denotes the catch. Correspondingly, H_{MSY} can be expressed by:

$$H_{MSY} = \frac{MSY}{SB_{MSY}}$$

Combing and re-arranging equation (3) and (5), it follows that r in equation (1) can be expressed as:

$$r = \frac{MSY}{SB_{MSY}} \frac{m-1}{1 - m^{-1}}$$

or

$$r = H_{MSY} \frac{m-1}{1 - m^{-1}}$$

This allows re-formulating the production function of the Pella-Tomlinson equation as a function of HMSY, such that:

$$SP_t = \frac{H_{MSY}}{1 - m^{-1}} SB_{t-1}(1 - (\frac{B_t - 1}{K})^{m-1})$$

where m can be directly translated into BMSY/K and thus determines the biomasss depletion level where MSY is achieved (Thorson et al. 2012b), using the following relationship:

$$\frac{B_{MSY}}{K} = m^{\frac{-1}{m-1}}$$

Part II: Code for the BSPSP Model

This is the underlying assessment model which is invoked via `source()` in Part I.

Bayesian State-Space Formulation

The BSPSP Model which is sourced by the front-end command was built upon the Bayesian state-space estimation framework proposed by Meyer and Millar (1999). The biomass B_y in year y is expressed as proportion of K (i.e. $P_y = B_y/K$) to improve the efficiency of the estimation algorithm.

The model is formulated to accommodate multiple CPUE for fisheries f. The initial biomass in the first year of the time series was scaled by introducing model parameter to estimate the ratio of the spawning biomass in the first year to K (Carvalho et al. 2014). The stochastic form of the process equation is given by:

$$P_y = \frac{\psi e^{\eta_y}}{(P_{y-1} + \frac{H_{MSY}}{1-m^{-1}} P_{y-1}(1 - P_{y-1}^m) - \frac{\sum_f C_{f,y-1}}{K})e^{\eta_y}} \quad y = 2, 3, ...n$$

where η_y is the process error, with $\eta_y \sim N(0, \sigma_\eta^2)$, $C_{f,y-1}$ s the catch in year y by fishery f. The corresponding biomass for year y is:

$$B_y = P_y K$$

The observation equation is given by:

$$I_{f,y} = q_f B_{f,y} e^{\epsilon_y} \quad y = 1, 2, ...n$$

where q_f is the estimable catchability coefficient associated with the abundance index for fishery f and ϵ_y is the observation error, with $\epsilon_{f,y} \sim N(0, \sigma_{\epsilon,f,y}^2)$, where is the observation variance for fishery f in year y.

To incorporate available standard errors of the year-effect estimated from the standardization models, we modified adopted an additional variance approach for the observation error variance (Booth and Quinn 2006, Carvalho et al. 2014), such that:

$$\sigma_{\epsilon,y,f}^2 = \hat{\sigma}_{SE,y,f}^2 + \sigma_{add,f}^2 \quad \epsilon_{y,f} \sim N(0, \sigma_{\epsilon,y,f}^2)$$

where $\hat{\sigma}_{SE,y,f}^2$ is the externally estimated standard error for year y and abundance index f and $\sigma_{add,f}^2$ is the estimable additional variance.

The full BSPSP model projected over n years requires a joint probability distribution over all unobservable hyper-parameters $\theta = K, H_{MSY}, \psi, \sigma_\eta^2, q_f, \sigma_{\epsilon,y,f}^2$ and the n process errors relating to the vector of unobserved states $\eta = \eta_1, \eta_2...\eta_y$, together with all observable data in the form of the relative abundance indices for fisheries f, $I_f, = I_{f,1}, I_{f,2}...I_{f,y}$ (Meyer and Millar, 1999). According to Bayes' theorem, it follows that joint posterior distribution over all unobservable parameters, given the data and unknown states, can be formulated as:

$$p(\theta|\eta, I) = p(K)p(H_{MSY})p(\psi)p(\sigma_\eta^2)p(q_f)p(\sigma_\epsilon^2)p(P_1|\psi, \sigma_\eta^2) \times \prod_{y=1}^{n} p(P_y|P_{y-1}, K, \psi, \sigma_\eta^2) \times \prod_{y=1}^{n} p(I_{f,y}|P_t, q_f, \eta_t, \sigma_{\eta,y,f}^2)$$

Format Input Data

```r
indices = names(cpue)[2:ncol(cpue)]
n.indices = max(length(indices),1)
catches = names(catch)[2:ncol(catch)]
n.catches = length(catches)

years=catch[,1]
styr = min(years)
endyr = max(years)
n.years = length(years)
styr.cpue = min(cpue[,1])
styr.I = styr.cpue-styr+1


# Convert input data to matrices
conv.cpue = as.numeric(rbind(matrix(rep(NA,(styr.I-1)*n.indices),styr.I-1,n.indices),as.matrix
CPUE=matrix(conv.cpue,nrow=n.years,ncol=n.indices)

conv.se = as.numeric(rbind(matrix(rep(NA,(styr.I-1)*n.indices),styr.I-1,n.indices),as.matrix(se
se2 = matrix(ifelse(is.na(conv.se),0.2,conv.se)^2,n.years,n.indices)#/2

conv.catch = as.numeric(rbind(matrix(rep(NA,(styr.I-1)*n.catches),styr.I-1,n.catches),as.matri
Catch=matrix(conv.catch,nrow=n.years,ncol=n.catches)
Catch[is.na(Catch)] = 0 # Replace any NA by zero

# Total Catch
TC = apply(Catch,1,sum)

set.seed(123)

# head(conv.cpue)
```
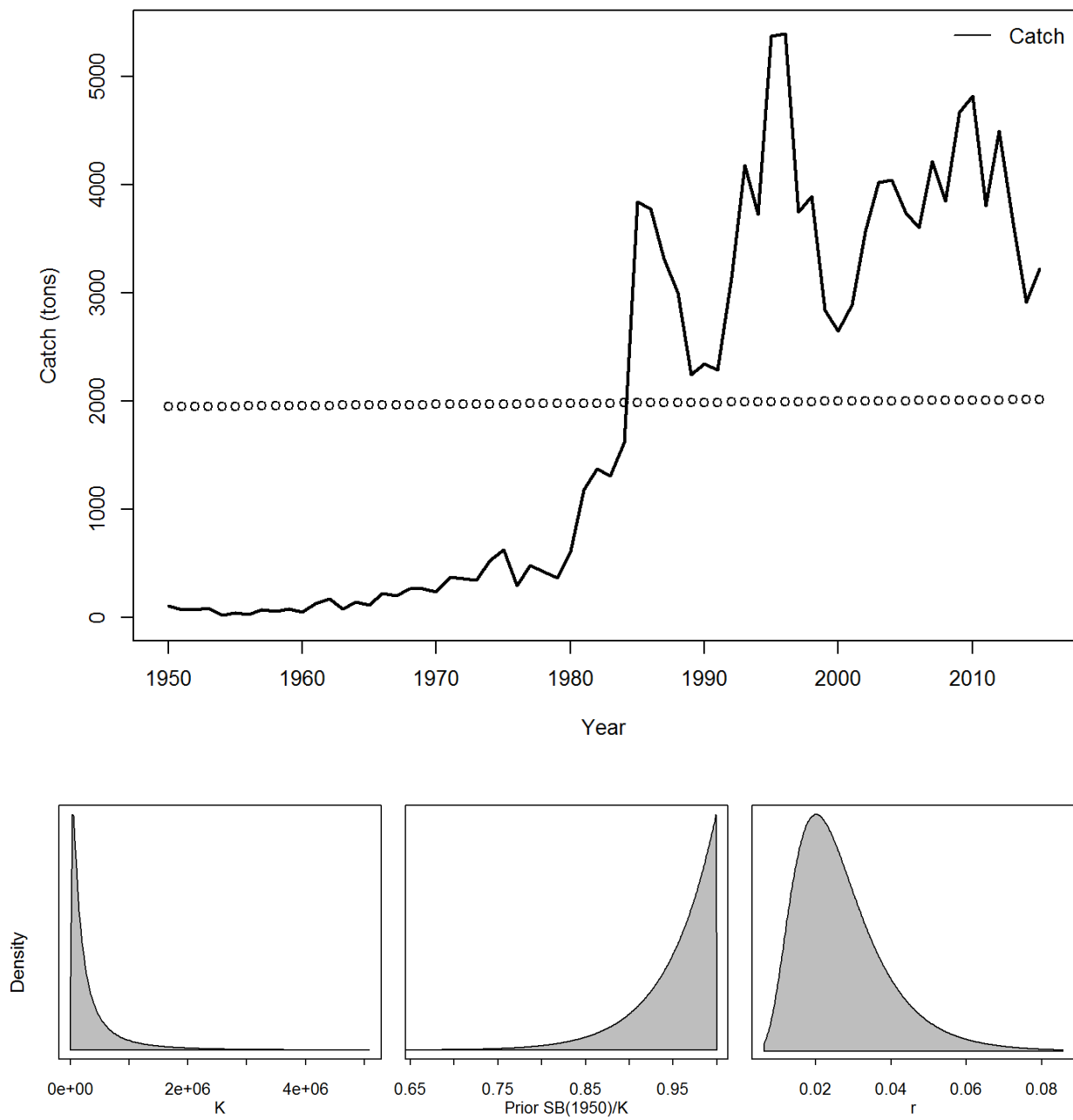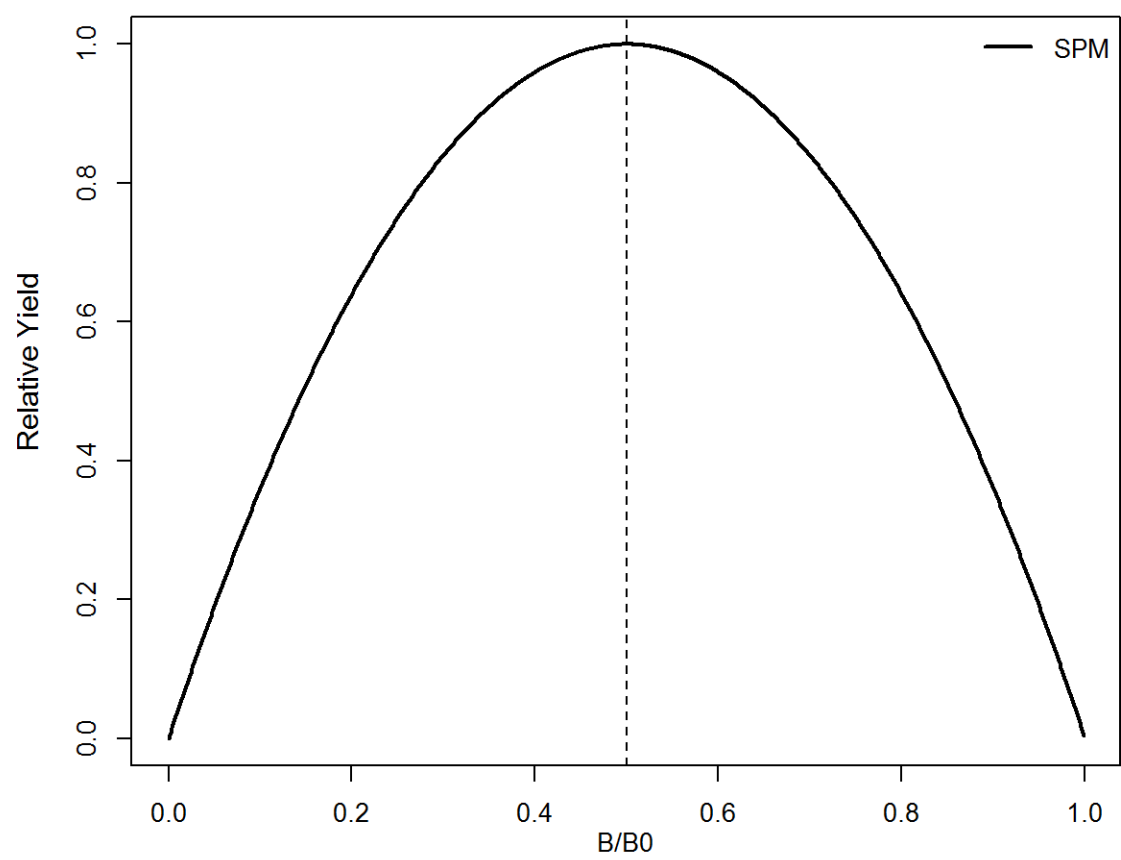
Plot Catch Input Data The plots generated in this step are saved to the Input file in the user's directory.

BSPSP prior formulations for Shortfin Mako

For this example vignette, we assumed a vaguely informative lognormal prior for K = 200,000 metric tons with a CV of 200% for both stocks. In contrast to the 2012 ICCAT SMA assessment, where the same lognormal (mean = log(0.058), sd = 0.12) was assumed for the base-case scenario, we worked with different ranges of r values for the North Atlantic and South Atlantic, which were suggested based on updated life history analysis (SCRS/2017/126). These ranges were rmin = 0.01 and rmax = 0.06 for the North Atlantic and rmin = 0.03 and rmax = 0.09, which were then converted in log-normal priors by:

$$log(r) = \frac{r_{min} + r_{max}}{2}$$

$$\sigma_r = \frac{\sqrt{log(r) - log(r_{min})^2}}{2}$$

where σ_r is the approximated lognormal standard deviation for the assumed range of r values. The corresponding lognormal priors are:

$$LN \sim (log(0.0254), 0.434) \quad North \quad Atlantic$$
$$LN \sim (log(0.052), 0.151) \quad South \quad Atlantic$$

The prior means for r were as translated into $H_{MSY} = r/2$.

```r
# Function to get beta prior parameters
get_beta <- function(mu,CV,Min=0,Prior="x"){
  a = seq(0.0001,1000,0.001)
  b= (a-mu*a)/mu
  s2 = a*b/((a+b)^2*(a+b+1))
  sdev = sqrt(s2)
  # find beta )parameter a
  CV.check = (sdev/mu-CV)^2
  a = a[CV.check==min(CV.check)]
  #find beta parameter b
  b = (a-mu*a)/mu
  x = seq(Min,1,0.001)
  pdf = dbeta(x,a,b)
  plot(x,pdf,type="l",xlim=range(x[pdf>0.01]),xlab=paste(Prior),ylab="",yaxt="n")
  polygon(c(x,rev(x)),c(rep(0,length(x)),rev(ifelse(pdf==Inf,100000,pdf))),col="grey")
  return(c(a,b))
}

get_gamma <- function(mu,CV,Prior="x"){
  a = seq(0.0001,1000,0.0001)
  b = a/mu
  s2 = (a/b^2)
  sdev = sqrt(s2)
  # find beta )parameter a
  CV.check = (sdev/mu-CV)^2
  a = a[CV.check==min(CV.check)]
  #find beta parameter b
  b = a/mu
  x = sort(rgamma(1000,a,b))
```

```
  pdf = dgamma(x,a,b)
  plot(x,pdf,type="l",xlim=range(x[pdf>0.01]),xlab=paste(Prior),ylab="",yaxt="n")
  polygon(c(x,rev(x)),c(rep(0,length(x)),rev(ifelse(pdf==Inf,100000,pdf))),col="grey")
  return(c(a,b))
}
```

```
plot_lnorm <- function(mu,CV,Prior="x"){
  sdev= sqrt(log(CV^2+1))
  rand.pr = rlnorm(1000,log(mu),sdev)
  x = seq(min(rand.pr),quantile(rand.pr,0.995),max(rand.pr/500))
  pdf = dlnorm(x,log(mu),sdev)
  plot(x,pdf,type="l",xlim=range(x),xlab=paste(Prior),ylab="",yaxt="n")
  polygon(c(x,rev(x)),c(rep(0,length(x)),rev(ifelse(pdf==Inf,100000,pdf))),col="grey")
  return(c(mu,sdev))
  }
```

Models I - IV

As in previous assessments, we assumed a Schaefer model for the base-case scenario by fixing input values for the shape parameter to $m = 2$. Initial depletion priors $\psi = \frac{B_1}{K}$ are inputted as means and CVs and then converted into corresponding beta priors. For our base-case scenarios, we assumed that the North Atlantic and South Atlantic were fairly unexploited at the start of the catch time series in 1950 and 1971, respectively. To reflect this we assumed the same beta prior for both stocks with a mean of 0.9 and a CV of 10%. All catchability parameters were formulated as uninformative uniform priors, while the process variance and observation variance were implement by assuming inverse-gamma priors.

```
# Determine initial ranges for r (see CMSY)

# initial range of r based on resilience (FishBase.org)
 if(length(r.prior)>1){ start.r = r.prior} else
 if(res == "High") {
   start.r <- c(0.6,1.5)} else if(res == "Medium") {
     start.r <- c(0.2,0.8)}    else if(res == "Low") {
       start.r <- c(0.05,0.5)}  else { # i.e. res== "Very low"
         start.r <- c(0.015,0.1)}

 log.r = mean(log(start.r))
 sd.r = abs(log.r - log(start.r[1]))/2
 r.prior = c(exp(log.r),sd.r)
 CV.r = sqrt(exp(sd.r^2)-1)

 # Translate Schaefer r prior to Hmsy
 Hmsy.pr = c(r.prior[1]/2,r.prior[2])
```

```r
# Set shape m for Fox and Schaefer: Fox m ~1; Schaefer m =2
if(shape==FALSE){
if(Model < 3){m=2} else {m = 1.001}}else{m=shape}


# Critical B/K at recruitment impairment
if(Model == 1 | Model == 3) {P.crit = 0.0001} else {P.crit = 0.25}


Par = list(mfrow=c(1,3),mai=c(0.5,0.1,0,.1),omi = c(0.1,0.2,0.1,0) + 0.1,mgp=c(2,1,0), tck =
png(file = paste0(assessment,"/",Scenario,"_",Mod.names,"/Input/Priors_",assessment,".png"),
    res = 200, units = "in")
par(Par)
plot_lnorm(mu.K,CV.K,Prior="K")
psi.pr = get_beta(mu=mu.psi,CV=CV.psi,Min=0,Prior=paste0("Prior SB(",years[1],")/K"))
r.pr = plot_lnorm(mu=r.prior[1],CV=CV.r,Prior="r")
mtext(paste("Density"), side=2, outer=T, at=0.5,line=1,cex=0.9)
dev.off()


# Plot MSY
Par = list(mfrow=c(1,1),mai=c(0.6,0.3,0,.15),omi = c(0.1,0.2,0.2,0) + 0.1,mgp=c(2,1,0), tck =
png(file = paste0(assessment,"/",Scenario,"_",Mod.names,"/Input/Prodution",assessment,".png")
    res = 200, units = "in")
par(Par)

# Get Bmsy/B0 as a fucntion of M
Bmsy=(m)^(-1/(m-1))
P = seq(0.0001,1,0.001)
SP = ifelse(P>P.crit,Hmsy.pr[1]/(1-1/m)*P*(1-P^(m-1)),Hmsy.pr[1]/(1-1/m)*P*(1-P^(m-1))*4*P)
#if(is.null(refBmsy)==TRUE) refBmsy = Bmsy
plot(P,SP/max(SP),type="l",ylab="Relative Yield",xlab="B/B0",lwd=2)
mtext(paste("Relative Yield"), side=2, outer=T, at=0.6,line=1,cex=0.9)
legend("topright",c("SPM"),col=c(1),lwd=2,bty="n")
abline(v=Bmsy,lty=2)
dev.off()
```

Set up BSPSP

The model selection made by the user in Part I dictates which of the following models is invoked.

```r
# remove scientific numbers
#options(scipen=999)

# starting values
nq = length(unique(sets.q))
nvar = length(unique(sets.var))
if(length(min.obsE)==1) min.obsE = rep(min.obsE,ncol(CPUE))
```

```r
# BSPSP Schaefer/Fox Models 1-4

  nSel = 1
  nI = ncol(CPUE)
  inits <- function(){list(K= rlnorm(1,log(mu.K),0.3),q = runif(nq,0.005,0.5), isigma2.est=runi
  surplus.dat = list(N=n.years, TC = TC,I=CPUE,SE2=se2,m=m,Hmsy.pr=Hmsy.pr,psi.pr=psi.pr,K.pr =

  BSPSP = "BSPSP.jags"

  # PARAMETERS TO MONITOR
  params <- c("K","r", "q", "psi","sigma2", "tau2","m","Hmsy","SBmsy", "MSY", "BtoBmsy","HtoHms
  # monitored quantaties (sorted alphabetically)
```

Execute the JAGS call

Note: the JAGS text file is not shown, but is called here by the BSPSP object.

```
mod <- jags(surplus.dat,
            inits,
            params,
            paste(BSPSP),
            n.chains = nc,
            n.thin = nt,
            n.iter = ni,
            n.burnin = nb)   # adapt is burn-in


cat(paste0("\n","- Produce results output","\n"))


# if run with library(rjags)
posteriors = mod$BUGSoutput$sims.list
```