**FISH 458/558 – Fisheries Population Dynamics**

Take-home Midterm Exam

***Student Name: Carter Adamson***

**Instructions**: This take-home part of your midterm exam is worth a total of 45 points. This exam is for both 458 and 558 students. You may draw upon your notes or any other written supporting material. **You may NOT consult with other students (past or present) or other people.** Please provide sufficient depth in your answers that convey your grasp of the topic. **This exam is NOT to be shared with any other individuals or students (current or future)**. Submit your exam as a Word document via Canvas by the due date and time. I am available for answering any questions you may have! Good luck!

**Formatting guidelines:**

* Overall
  + Include an appropriate header for your document
  + CLEARLY identify your answers to the questions and include the correct question number
  + Write in a manner that would be suitable for a scientific article (i.e., write professionally!)
  + Avoid grammatical errors 🡪 Proofread your answers!
  + Include your R code as an appendix to the exam (i.e., at the end), instead of having separate chunks of code for each question/section.
* For all figures:
  + make sure that the X and Y limits are appropriate,
  + make sure the size and scale of your plots are appropriate (i.e., no squashed or squeezed figures),
  + label axes and include units,
  + include a legend where appropriate (e.g., when plotting multiple lines), and don’t have the legend cover any part of your plot,
  + include an informative figure caption below the figure. (Note: multi-panel figures with a single caption will likely be useful for the diagnostic plots)
* For all tables:
  + include a table caption above the table

**QUESTIONS (45 points)**

1. **Below are some results for an occupancy analysis of a terrestrial salamander (Table 1, Figure 1). Use the results to answer the questions. (8 pts)**
   1. **What conclusions can you make from Table 1 regarding the results of the study and the variables that are important? Explain your reasoning. (3 pts)**

Table 1 summarizes the occupancy analysis for this salamander, and looking at it can tell us about the effects of various covariates on Psi, the probability of occupancy, and p, the probability of detection. The model indicated as the “best” (dAIC = 0) includes disturbed habitat as a covariate affecting occupancy and detection, as well as time affecting detection. All models reasonably close to the best model (using some “rule of thumb” cutoffs: dAIC < 2, AIC weights < 10%) also include disturbance as a covariate for Psi, indicating the importance of this variable. Models including disturbance for occupancy universally perform better than models that don’t include a covariate for occupancy. The covariates for detection make relatively less of an impact, since all variations are still within < 2 dAIC.

* 1. **How do the AICc, ΔAICc, and Weight columns in Table 1 relate to one another generally? (2 pts)**

The Akaike Information Criterion (corrected, in this case) can be used to compare the relative fit of a set of candidate models, with a lower AIC indicating a better model fit. dAIC is calculated as the difference between a model and the best fit of the candidate models. Thus, the best fit model with have dAIC = 0. dAIC values between 0 and 2 generally indicate strong models. Another way to see howstrong these models are, comparatively, is by looking at the Akaike weight, which can be thought of as a probability of that model being the best, given its AIC score and the AIC scores of the other candidate models. Taken together, these three statistics can help narrow down a set of models to find the most important covariates.

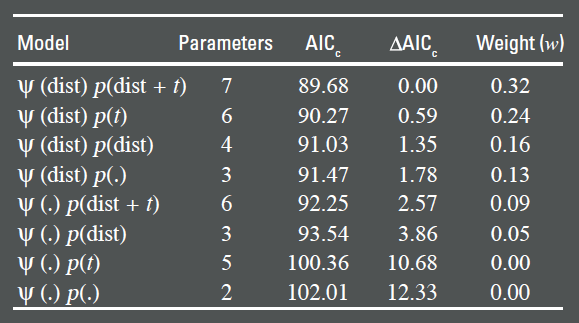
* 1. **What main conclusions can you make from the Figure and how does this relate to the Table? Explain your reasoning. (3 pts)**

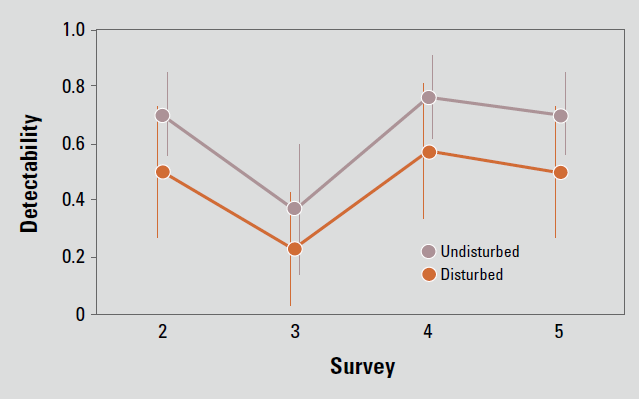
The figure shows the impacts of time (x-axis) and disturbance (lines) on the detectability (y-axis) or probability of detection p. From the table, we saw that the dAIC = 0 best model included both disturbance and time as covariates for detectability. On the graph, we can indeed see that detectability varies with time (survey 3, for example, seems to have lower detectability) and with habitat disturbance (undisturbed having consistently higher detectability). However, the AIC table shows some uncertainty; among the four best-scoring models, the AIC weights indicate no clear consensus on which is the best. All possibilities of distance, time, or neither affecting detectability fall within weights of 0.13 – 0.32. As such, it’s difficult to be certain about drawing stark conclusions from the table. Looking at the graph, we can see large error bars that often overlap, which should likewise give us pause about our conclusions.

* 1. **Do these results (i.e., Table 1 and Figure 1) indicate how the occupancy changes with disturbance history? Why? (1 pt)**

Not exactly. Table 1 does show us that disturbance is important for modeling occupancy. However, Figure 1 shows disturbance’s impact on detectability, not occupancy, so we cannot use it as evidence. The table indicates that it is likely that disturbance has significant impacts on occupancy, since the best models all include it. Beyond that, however, the table and figure give no information on the nature or direction of the relationship. We would have to look at other information to assess that.

**Table 1**. Results of occupancy model selection for salamanders in a National Park. The models considered assume occupancy was either constant across sites ψ (.), or varied according to the sites’ previous disturbance history, ψ (dist). Detectability was modeled in four ways: it was constant across all surveys and sites p(.), it varied among surveys p(t), it varied across sites according to previous disturbance history p(dist), or it varied by both disturbance history and survey p(dist+t). Model selection was based on Akaike’s Information Criteria for small sample size (AICc), and Akaike weight (w).





**Figure 1.** Estimates of detectability across surveys and among sites with different disturbance histories. The estimates are averaged across all possible models.

1. **Develop predictive stock-recruitment relationships for the hypothetical Pacific sardine (*Sardinops sagax*) data contained in the file “Exam\_2\_SRdata.csv” (units are millions of fish for the recruits and kmt for the spawners). Use nonlinear least squares (the nls() function in R) to fit your models. For your candidate set of models, fit the density-independent, Beverton-Holt, and Ricker models to the data, and assume that the models have multiplicative error. Do the following (15 points):**
   1. **Generate a single plot that includes the observed data and the 3 predicted lines for each model (in different colors). (4 pts)**

A graph with different colored lines and dots

Description automatically generated

**Figure 2.** Stock-recruitment plot showing the number of recruits versus the number of spawners for a population of Pacific sardine. Black points show annual observed numbers of spawners and recruits for years between 1966 and 2010. Lines show model fits obtained through three alternative stock-recruitment modeling methodologies: a density-independent model, Beverton-Holt model, and a Ricker model.

* 1. **Create a table that includes the estimated *a* and *b* parameters (and their standard errors) for each of the 3 models. Include also the estimated AIC and AIC differences (i.e., Δ or dAIC) for each model. [Note, you do not have to make the table in R if you don’t want to]. (4 pts)**

**Table 2**. Summary table for stock recruitment analysis for the Pacific sardine population. Data were analyzed with three competing model methodologies. Table compares parameter estimates between models for a (the productivity parameter) and b (the density-dependence parameter), as well as AIC and dAIC values.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Model** | **a** | **b** | **AIC** | **dAIC** |
| Density Independent | 0.380 | N/A | 80.59 | 23.77 |
| Beverton-Holt | 0.952 | 0.010 | 58.62 | 1.81 |
| Ricker | 0.725 | 0.004 | 56.81 | 0 |

* 1. **Present diagnostic plots for each model for assessing the normality and homogeneity of variance assumptions. Provide a written description of your conclusions and reasoning regarding the assumptions based on you plots. (5 pts)**

A group of graphs showing different types of data

Description automatically generated

**Figure 3.** Diagnostic plots for the three stock-recruitment models presented in Figure 2 and Table 2. Each plot is identified with the name of the model methodology it refers to, organized row-wise.

These model diagnostics give me some concerns regarding the ability of these models. The first column, the histograms of residuals for each of the three models, look off in all cases, showing a distinct left skew in all cases rather than the expected bell shape. Looking at the residual vs. fitted plots in the second column, the desntiy independent model clearly fails the assumption of homogeneity of variance, showing positive residuals at low fitted values and negative residuals at high fitted values. The other two models are harder to assess, due to being relatively data-poor at low fitted values, but may also exhibit the ”cone” shape associated with violations of the assumption of homogeneity of variance.

* 1. **Based on your analysis, describe what the best SR model is for the dataset and why. (2 pts)**

Table 2 clearly indicates that the Ricker model performs the best of the three stock-recruitment models attempted (dAIC = 0). However, Figure 3 shows that this model and the others may violate the assumptions of normality and homogeneity of variance. This indicates that perhaps none of the models describe the dataset well enough.

1. **Conduct yield per recruit (YPR) and spawning stock biomass per recruit (SBPR) analyses for a population of rainbow trout in Blue Lake, Sitka, Alaska. Your goal is to determine potential biological reference points (BRPs) to be used in management decisions and to describe the tradeoffs among the different BRP options. Natural mortality is estimated to be M=0.26 for the population. You should assume that 50% of fishing and natural mortality occur before spawning, and that there is no plus group. Using the “Exam2\_trout data.csv” data (also in Table 2), answer the following questions. (20 pts)**
2. **Use a YPR analysis to determine the estimates for Fmax and F0.1, and the associated YPR at each of those F values. Report the estimates. You will use function ypr() function from the *fishmethods* package for this. (2 pts)**

**Table 3.** Results of a yield per recruit analysis for the Blue Lake rainbow trout population. The analysis reports fishing mortality and associated yield per recruit (YPR, in kg) at two reference points: fishing mortality at the highest yield per recruit (Fmax), and mortality when the YPR curve has reached 10% of its original slope (F0.1).

|  |  |  |
| --- | --- | --- |
|  | **F** | **YPR** |
| **Fmax** | 0.940 | 0.167 |
| **F0.1** | 0.538 | 0.156 |

1. **Use a SBPR analysis to determine the estimates for F30% and F40%, and the associated spawning stock biomass per recruit at each of those F values. Report the estimates. You will use function sbpr() function from the *fishmethods* package for this. (2 pts)**

**Table 4.** Results of spawning stock biomass per recruit analysis for the Blue Lake rainbow trout population. The analysis reports fishing mortality and spawning stock biomass per recruit (SBPR, kg) at two reference points: fishing mortality at 30% and 40% of the max spawning potential ratio (i.e. the spawning potential when there is no fishing pressure).

|  |  |  |
| --- | --- | --- |
|  | **F** | **SBPR** |
| **F\_30%** | 0.581 | 0.242 |
| **F\_40%** | 0.427 | 0.322 |

1. **Make a graph of *YPR vs. F* with all four reference points (Fmax, F0.1, F30%, F40%) identified using vertical lines. (4 pts)**

A graph with different colored lines

Description automatically generated

**Figure 4.** Yield-per-recruit curve for Blue Lake Salmon population. Plots yield-per-recruit (kg) against fishing mortality (F). Vertical lines represent fishing mortality at various reference points obtained in YPR and SBPR analyses.

1. **Make a graph of *% of Max SPR vs. F* with all four reference points (Fmax, F0.1, F30%, F40%) identified using vertical lines. Recall that the “*% of Max SPR”* should range from 0-100% (4 pts)**

A graph with different colored lines

Description automatically generated

**Figure 5.** Spawning potential ratio curve for Blue Lake Salmon population. Plots the percentage of the maximum possible spawning potential ratio (achieved under no fishing pressure) against fishing mortality (F). Vertical lines represent fishing mortality at various reference points obtained in YPR and SBPR analyses.

1. **Create a single table to summarize the results of your YPR and SPR analyses. For each of the four BRPs (Fmax, F0.1, F30%, F40%), you should present: the associated F value, the corresponding annual exploitation rate (u; calculate this using the equation from lecture 14, assuming this is a Type I, pulse fishery), YPR (in kg), percent of max YPR (i.e., YPR expressed as a percentage of the maximum YPR), the spawning stock biomass per recruit (in kg), and the spawning stock biomass per recruit as a percentage of the maximum spawning potential. (4 pts)**

**Table 5.** Summary table for the results of yield-per-recruit and spawning-stock-biomass-per-recuit analyses for the Blue Lake rainbow trout population. Results are reported according to four reference points associated with values of fishing mortality (F). Also reported for each reference point are fishing mortality re-expressed as an annual exploitation rate (u), yield per recruit (YPR, in kg), yield per recruit re-expressed as a percentage of the maximum possible yield per recruit, spawning stock biomass per recruit (SBPR, in kg), and the corresponding percentage of the maximum possible spawning potential ratio.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Reference Point** | **F** | **u** | **YPR** | **% max YPR** | **SBPR** | **% max SPR** |
| Fmax | 0.940 | 0.609 | 0.166 | 100% | 0.135 | 17% |
| F0.1 | 0.538 | 0.416 | 0.156 | 94% | 0.260 | 32% |
| F\_30% | 0.581 | 0.441 | 0.159 | 95% | 0.242 | 30% |
| F\_40% | 0.426 | 0.347 | 0.146 | 88% | 0.320 | 40% |

1. **Based on your plots and table with your calculated values, discuss the tradeoffs in yield and spawning potential ratio that result from the four different BRP options (Fmax, F0.1, F30%, F40%). Which reference point would you recommend to a manager and why? (4 pts)**

The central point to take home form Figures 4 and 5, and Table 5, concerns the tradeoff between fishery yield and the spawning potential of the population. The reference point results in the highest yield (Fmax, 100% max YPR) also results in the lowest % max SPR of the four reference points. Likewise, the scenario that conserves the highest % max SPR (F\_40%) results in the lowest yields (88% max YPR). A fishery manager focused on generating the maximum profit for the fishery would likely prefer the former scenario, but a more conservation-minded manager would prioritize preserving the spawning potential of the population to ensure resilience. The other two reference points are similar to each other and are found in between the other two. Of these “in-between” reference points, F0.1 slightly favors spawning potential and F\_30% slightly favors yield. Both could serve as compromise points for a fishery manager. Since my values bias me towards a heavily conservation-focused management strategy, I would personally recommend F\_40% as a reference point to benefit the population of trout as much as possible, though this decision may not be popular with commercial fishermen.

1. **Re-read the instructions and formatting guidelines. State whether you followed them.**

Yep!

1. **How many hours did you spend on this exam? (1 pt)**

6 hours

**CODE:**

library(ggplot2)

library(here)

library(fishmethods)

# QUESTION 2 ####

sardine <- read.csv(here("Exam 2","Exam\_2\_SRdata.csv"))

## a - stock-recruitment plot with three models ####

#make density independent model

sr.ind <- nls(log(R)~log(a\*S), data=sardine, start=c(a=.2), trace=T)

SEE.sr.ind <- summary(sr.ind)$sigma

pred.sr.ind <- exp(predict(sr.ind))\*exp((SEE.sr.ind^2)/2)

#make B-H model

sr.bh <- nls(log(R)~log(a\*S/(1+b\*S)), data=sardine,start=c(a=1.0,b=5e-5), trace=T)

SEE.sr.bh <- summary(sr.bh)$sigma

pred.sr.bh <- exp(predict(sr.bh))\*exp((SEE.sr.bh^2)/2)

#make Ricker model

# log(R)=log(a)+log(S)-b\*S+e

sr.ric <- nls(log(R)~log(a)+log(S)-(b\*S), start=c(a = 1.0, b=5e-5), data=sardine,trace=T)

SEE.sr.ric <- summary(sr.ric)$sigma

pred.sr.ric <- exp(predict(sr.ric))\*exp((SEE.sr.ric^2)/2)

#plot

plot.sr <- ggplot(data=sardine, aes(x=S, y=R))+geom\_point()+

theme\_bw()+labs(x="Spawners", y="Recruits")+

geom\_line(aes(y=pred.sr.ind,color="Density Independent"))+

geom\_line(aes(y=pred.sr.bh,color="Beverton-Holt"))+

geom\_line(aes(y=pred.sr.ric,color="Ricker"))

plot.sr

## b - table of a and b parameters, AIC, dAIC ####

#assemble table

m <- c("Density Independent", "Beverton-Holt", "Ricker")

a <- c(0.3797, 0.9522, 0.7251)

b <- c("N/A", 0.0101, 0.0039)

AIC <- c(AIC(sr.ind),AIC(sr.bh),AIC(sr.ric))

dAIC <- AIC - min(AIC)

table.2b <- data.frame(Model=m, a=a, b=b, AIC=AIC, dAIC=dAIC)

table.2b

## c - diagnostic plots (normality, HOV) ####

#plot 3x2 grid of diagnosic plots

par(mfrow=c(3,2))

hist(resid(sr.ind), xlab="Residual", main = "Density Independent")

plot(resid(sr.ind)~fitted(sr.ind), main = "Density Independent")

abline(h=0, lty=2)

hist(resid(sr.bh), xlab="Residual", main = "Beverton-Holt")

plot(resid(sr.bh)~fitted(sr.bh), main = "Beverton-Holt")

abline(h=0, lty=2)

hist(resid(sr.ric), xlab="Residual", main = "Ricker")

plot(resid(sr.ric)~fitted(sr.ric), main = "Ricker")

abline(h=0, lty=2)

# QUESTION 3 ####

trout <- read.csv(here("Exam 2", "Exam\_2\_trout data.csv"))

par(mfrow=c(1,1))

## a - YPR analysis to find Fmax and F0.1 ####

#YPR analysis

trout.ypr <- ypr(age=trout$age,wgt=trout$weight,partial=trout$selectivity, M=0.26,

plus=FALSE,maxF=2,incrF=0.01, graph=TRUE)

trout.ypr$Reference\_Points

## b - SBPR analysis for F30 and F40 ####

#SBPR analyses for each reference point

trout.sbpr.30 <- sbpr(age=trout$age, ssbwgt=trout$weight, partial=trout$selectivity,

pmat=trout$maturity, M=0.26, pF=0.5, pM=0.5, MSP=30,

plus=FALSE, maxF=2, incrF=0.01, graph=TRUE)

trout.sbpr.30$Reference\_Point

trout.sbpr.40 <- sbpr(age=trout$age, ssbwgt=trout$weight, partial=trout$selectivity,

pmat=trout$maturity, M=0.26, pF=0.5, pM=0.5, MSP=40,

plus=FALSE, maxF=2, incrF=0.01, graph=TRUE)

trout.sbpr.40$Reference\_Point

## c - graph YPR vs F, including all four reference points ####

plot.3.data<-data.frame(F=trout.ypr$F\_vs\_YPR$F, YPR=trout.ypr$F\_vs\_YPR$YPR,

SPR=trout.sbpr.30$F\_vs\_SPR$PSPR)

#basic graph

plot.3c <- ggplot(data=plot.3.data, aes(x=F, y=YPR))+geom\_line()+

theme\_bw()+ labs(colour="Reference Point")+

geom\_vline(aes(xintercept=0.94, color="Fmax"), linetype="dashed")+

geom\_vline(aes(xintercept=0.5383887, color="F0.1"), linetype="dashed")+

geom\_vline(aes(xintercept=0.5808088, color="F\_30%"), linetype="dashed")+

geom\_vline(aes(xintercept=0.4266088, color="F\_40%"), linetype="dashed")

#add points

F <- c(0.94, 0.5383887, 0.5808088, 0.426088)

YPR <- c(0.1668414, 0.1564271, 0.159070255, 0.146347072) #kg

color <- c("Fmax", "F0.1","F\_30%","F\_40%")

points.df <- data.frame(F=F, YPR = YPR, color = color)

plot.3c <- plot.3c + geom\_point(data=points.df, aes(color=color))

plot.3c

## d - graph % max SPR with all four reference points ####

plot.3d <- ggplot(data=plot.3.data, aes(x=F, y=SPR))+geom\_line()+

theme\_bw()+ labs(y="% Max SPR", colour="Reference Point")+

geom\_vline(aes(xintercept=0.94, color="Fmax"), linetype="dashed")+

geom\_vline(aes(xintercept=0.5383887, color="F0.1"), linetype="dashed")+

geom\_vline(aes(xintercept=0.5808088, color="F\_30%"), linetype="dashed")+

geom\_vline(aes(xintercept=0.4266088, color="F\_40%"), linetype="dashed")

#add points

F <- c(0.94, 0.5383887, 0.5808088, 0.426088)

SPR <- c(16.772855, 32.299491, 30.043468, 39.737697) #%

color <- c("Fmax", "F0.1","F\_30%","F\_40%")

points.df <- data.frame(F=F.vector, YPR = YPR.vector, color = color.vector)

plot.3d <- plot.3d + geom\_point(data=points.df, aes(color=color))

plot.3d

## e - table to summarize results for all reference points ####

ref <- c("Fmax", "F0.1", "F\_30%", "F\_40%")

u <- 1-exp(-F)

YPR <- c(0.1668414, 0.1564271, 0.159070255, 0.146347072) #kg

Percent\_max\_YPR <- (YPR/0.1668414)\*100

SPR <- c(0.13506823,0.26010092,0.24193365,0.31999921) #kg

Percent\_max\_SPR <- c(16.772855, 32.299491, 30.043468, 39.737697)

table.3e <- data.frame(Reference\_point=ref, F=F, u=u, YPR=YPR,

Percent\_max\_YPR=Percent\_max\_YPR, SPR=SPR,

Percent\_max\_SPR=Percent\_max\_SPR)

table.3e