**Ricker Estimation Lab Instructions**

**I. Simulation Model**

*Green cells: SRR Parameters- USER INPUT*

*Blue cells: SRR Parameters- Calculated with formulae to be preserved*

*Yellow cells: SRR Estimates (^)-Calculated with formulae to be preserved*

*All instructions that require direct action are underlined, with cell formulas to enter in blue**font.*

1. Go to sheet "Simulator". This sheet simulates the year-to-year functioning of a salmon fishery with two relationships, the Spawner-Recruit Relationship or “SRR” and the Recruit-Spawner Relationship or “RSR”. The SRR should be familiar- it models recruitment as a function of spawners using a Ricker model, which you’ve been hearing about in the lectures. For the sake of simplicity, it assumes all fish mature in one year. The RSR is new to you- it models spawners as a function of recruitment by accounting for harvest. The RSR model in this lab mimics simple escapement goal management, where harvest rate is responsive to run size but the goal is represented as a threshold (Parameter **Sgoal**). Parameter **F1** is the instantaneous fishery mortality during years when spawning escapement falls below the threshold, whereas parameter **F2** is the (presumably higher) instantaneous fishing mortality when escapement exceeds the threshold. The final RSR parameter is **F**, which controls variability around the intended harvest rate (the difference between the actual harvest rate and the intended harvest rate). It is roughly equivalent to the CV of implementation error, so a fishery with **F** = 0.1 would be precisely managed, whereas a fishery with **F** = 0.4 would be hard to control.

In this lab we will look at how fishing affects estimation and the goal setting process. You will be assigned to one of two groups; “high harvest rate” (**F1** = 0.4 and **F2** = 1.6) or “low harvest rate” (**F1** = 0.1 and **F2** =0.4). Think of **F1** and **F2** as the mortality that results when, respectively, the run is weak and fisheries are being restricted, versus when the run is strong and fisheries are being liberalized.

2) Enter your assigned values for **F1** and **F2** and press the F9 key to simulate a new dataset.

3) Locate the plot labeled “SRR”. Blue squares represent true median recruitment from the SRR we used to simulate data. Empty circles represent simulated S and R data. Red diamonds represent median recruitment as estimated from the data. Press the F9 key repeatedly. Inspect the SRR plot to get a feel for the range of escapements and recruitments observed under your assigned harvest rates. Also look at differences between the true and estimated SRR curves.

Do the estimated and true recruitments agree over the entire range of data? If not, where are the largest differences?

4) Locate the estimated SRR parameter values and the associated errors relative to the true values. Press the F9 key repeatedly. Inspect the estimated parameter values and the associated errors for each generated dataset.

Is there a Ricker parameter (**α** and **β**) that tends to have the largest relative error?

Is there a management reference point (SMAX, SEQ, SMSY, UMSY) that tends to have smaller relative error? If so, why?

5) Locate the SRR plot. Press the F9 key repeatedly. Look at the how the simulated data is distributed around the true values.

What happens to the simulated data in the SRR plot when you change **w** (Try values of 0.1. 0.4 and 0.8)?

6) Set **w** back to the default (0.4). Locate the RSR plot. Press the F9 key repeatedly. Look at the how the simulated data is distributed.

What happens to the simulated data in the RSR plot when you change **F**(Try values of 0.1 0.4 and 0.8)?

Set **F** back to the default (0.4) and change **F1** and **F2** to match the other groups settings. How does the relationship between simulated spawning escapements and simulated recruitments in the RSR plot change?

7) Return all parameter values back to their default settings.

|  |  |  |
| --- | --- | --- |
| **Default parameter values** | | |
| Log productivity | **ln()** | 1.5 |
| Density dependence | **** | 0.001 |
| Process error | **w** | 0.4 |
| Auto-correlation | **** | 0 |
| Observation error | **s** | 0.1 |
| Low fishing mortality | **F1** | 0.1 or 0.4 |
| Escapement threshold | **Sgoal** | 500 |
| High fishing mortality | **F2** | 0.4 or 1.6 |
| Outcome error | **F** | 0.4 |

**II. Point Estimates**

1) Go to sheet "Ricker Sim". The S and R data are linked to the “Simulator” sheet. Break the links and fix the S and R values by copying (green) cells B9-C28, then "Paste Special, Values” to the same location. The scatter plots of R vs S should now remain static when you press F9.

2) Calculate ln(R/S) in cell D9 and copy down through D28. The scatter plot centered near K33 now displays ln(R/S) vs S. Notice that (the log of) productivity (R/S) declines with increasing S. It may dip below 0 at large S, signifying escapements that did not replace themselves. These are the same points as the ones below the R=S replacement line in the R vs S scatter plots.

3) Conduct the simple linear regression of ln(R/S) on S and use the intercept as an estimate of lnalpha, the negative of the slope is an estimate of beta. Go to cell K3 and type “=INTERCEPT(d9:d58,b9:b58)” and to cell K4 and type “=-SLOPE(d9:d58,b9:b58)” (don’t miss the minus sign). The estimate of sigma is already completed in cell K5. The estimated Ricker curve will now appear in the top R vs S scatter plot.

4) In cells K7-K11, calculate estimates of Smax, Seq, Smsy, Umsy, and MSY as functions of the 3 basic parameters. See Equations sheet. Note that Seq, Smsy, and Umsy are all functions of the corrected lnalpha.p, not the raw lnalpha.

5) Calculate fitted ln^R/S = ln(alpha)^ - beta^ S in column E. That is, in cell E9, type in

“=$K$3-$K$4\*B9” and copy down. These fitted values define the estimated regression line.

6) Calculate log residuals = ln(R/S) - ln^R/S in column F. That is, in cell F9, type in

“=D9-E9” and copy down. These are plotted versus year at the bottom of the home screen (home screen = the part of the spreadsheet visible with cell A1 at the upper left corner). Do you see signs of serially correlated productivity in this stock? The Durbin-Watson test statistic (DW) for AR(1) serial correlation is calculated in cell AA41.

For n=20 years of data:

If DW <1.20 conclude that positive AR(1) serial correlation exists.

If DW>1.41 conclude that there is no AR(1) serial correlation.

If 1.20 <DW< 1.41, the test is inconclusive.

Note that the critical values differ for different n (see any stat book, or hidden sheet DW). Does your stock have serially correlated productivity?

7) Calculate fitted R^ = S exp(lnalpha – beta S) in column G. That is, in cell g9, type in

“=B9\*exp($K$3-$K$4\*B9)” and copy down.

8) You now have point estimates of the Ricker spawner-recruit parameters, and the associated curve. Write down estimates for the following:

ln()^ ^ SMSY^\_\_\_\_\_\_\_\_ UMSY^\_\_\_\_\_\_\_\_

Compare your estimates to the parameter values used to simulate the data. Write the simulation parameters below:

ln()  SMSY \_\_\_\_\_\_\_\_ UMSY \_\_\_\_\_\_\_\_

**III. Quantifying Uncertainty**

*Our point estimates are not equal to the true values. In this section we will quantify how far off they could be.*

1) Look at rows 62-64 on sheet “Ricker Sim”, starting in column K. The values for S, fitted ln(R/S), and residuals appear here, transposed. We will resample ("bootstrap") the regression residuals to quantify the uncertainty around our parameter estimates. For each observation, we take the original fitted ln(R/S) value (the point on the regression line) and add to it a residual randomly selected from the n=20 we have observed.

2) Use the INDEX function to randomly resample residuals. Go to cell K65 and copy it to the right, through cell AD65. You now have 20 residuals in random order, selected with replacement.

3) Go to cell BJ65 and note that this adds the new resampled residual to the fitted ln(R/S) value. Copy it to the right, through cell CC65. These numbers are the new pseudo-data (ln(R/S) values) representing one of the many bootstrap samples we will create. We will regress each of these bootstrap samples against the original S values in the following sections.

4) Notice that cells A65-I65 contain functions and formulae equivalent to those you entered in column K, above, but reference the row of pseudo-data to the right. These are the Ricker parameter estimates for bootstrap replicate number 1.

5) Select cells A65-CC65 and copy down through row 1064. You now have up to 1000 bootstrap replicates of your data, each with an associated set of parameter values. *Consider these to be possible versions of the true parameter values.* The IF() function is used in columns C-I to censor bootstrap reps that result in negative values of lnalpha or beta. Neither lnalpha nor beta could be negative in nature.

6) Use the FREQUENCY() array function to tally and plot distributions of the bootstrap estimates. This has already been done for lnalpha, beta, S.msy and U.msy. To repeat for S.eq, select cells BJ7:BJ47, type “=FREQUENCY(E65:E1064,BI7:BI46)” and hit ctrl-shift enter. The bootstrap distributions are plotted on the home screen. In this lab the data is simulated and the true parameter estimates are known. These values are represented in the plots as light blue vertical lines.

7) Using PERCENTILE(), calculate the 10th and 90th percentiles of bootstrap distributions in the yellow cells of columns J and L. This has already been done for the first several quantities in column I. These are lower and upper bounds of 80% confidence intervals for each quantity. Look at the formulae in J5 and L5 for sigma and modify them to calculate the appropriate percentiles for Seq and Smsy. Note that you cannot just copy and paste the formulae down- you have to choose the right data ranges for the percentile function in each case. What is an 80% interval for Smsy for your stock? (\_\_\_\_\_\_ , \_\_\_\_\_\_). How would you get a 90% interval?

8) It can be useful to standardize the uncertainty for comparison with other stocks. Calculate the relative uncertainty by dividing the range of the bootstrap interval by the median. Further dividing by the constant 2.56 makes this a nonparametric (NP) equivalent of the familiar coefficient of variation (CV). Calculate NPCV for beta, Seq, Smsy, and MSY in column M. This has already been carried out for lnalpha.

9) Create a “horsetail” plot. Go to cell DI65 and copy it to the right, through cell EC65, and then down to cell EC114. *The resulting 50 Ricker curves in the horsetail plot at the bottom of the home screen are examples of true curves that could have resulted in the observed data*. Recalculate (key F9) to see a new set of 50[[1]](#footnote-2). The horsetail plot is not used to calculate anything further, but provides a visual describing variability in our estimation of the Ricker curve with the data at hand.

10) The bootstrap distributions completely summarize what you know about the lnalpha, beta, Smsy, and Seq. *When documenting a stock-recruit analysis, it is better to report interval estimates derived from those distributions, rather than single point estimates*. Similarly, *it is more illuminating to include a “horsetail plot”, rather than a plot of a single fitted curve*.

**IV. Graphical Tools for Evaluating Escapement Goals**

If we knew the true Ricker curve (alpha, beta) exactly, choosing an escapement goal would be easy. For instance, we could pick an EG range which would result in sustained yield (SY=R-S) greater than some high percentage, say 90%, of maximum sustained yield (MSY = Rmsy – Smsy, see equations). But we are uncertain about the true Ricker curve, and we want a goal that performs best across the entire family of Ricker curves that could have plausibly led to our data, not just the single estimated curve. We will create two graphical tools to help choose an escapement goal. Both are based on the collection of bootstrap estimates.

1) The first tool, called an ***Optimal Yield probability profile***, we plot the probability that different values of S will result in yield exceeding 90% of MSY.

a) An incremental series of prospective escapements S\* has already been calculated in cells FA61:GE61.

b) For the first bootstrap dataset, across all prospective escapements S\*, determine whether or not is greater than 90% of the bootstrap MSY value. Record 0 if not or 1 if so. This has been done, using the IF() function, in cell FA65. To repeat for other S\* values, copy cell FA65 to the right through cell GE65.

c) Do the same for the remaining bootstrap datasets: select FA65:GE65 and copy it down through row 1064. Now we have, for each bootstrap dataset, a '1' if the value of SY at each S is >90%MSY and a "0" if not. Scroll through the values and notice that the "1"s tend to concentrate in a certain “zone” of the S\* series. These are the values of S which are most likely to result in near-maximal yields.

d) For each value of S\*, estimate the probability that SY>90%MSY by calculating the average of the 0s and 1s over all bootstrap reps[[2]](#footnote-3). In cell FA63, type “=AVERAGE(FA65:FA1064)” and copy it to the right through cell GE63. Each one of these is an estimate of the probability of achieving 90% of MSY for a prospective value of S.

e) Return to your home screen and look at the plot centered on Q14. *Optimal yield* probabilities are plotted vs S\* on a plot labelled “Optimal Yield Probability Profile”. They reach a maximum near the point estimate of Smsy. You can choose an escapement goal that maximizes the probability of achieving near-maximal sustained yield by choosing a horizontal gridline, finding where it intersects with the profile, and projecting it down to the S axis using the red vertical lines provided. Note that higher grid lines give narrow goals and lower grid lines wider goals.

2) The second tool is a plot of ***Expected Sustained Yield*** (= R - S) versus spawning escapement, with confidence intervals that reflect our uncertainty about expected sustained yield.

1. Prospective values of S\* are in EE64:EY64. Go to cell EE65, where expected sustained yield has already been calculated for the first bootstrap dataset. Copy cell EE65 to the right through cell EY65, Then copy the entire row (EE65:EY65) down to cell EY1064. This calculates expected sustained yield for all 21 values of S\* and all 1000 bootstrap datasets.
2. Calculate point and interval estimates of expected sustained yield for each prospective value of S. Go to cells EE58:EE62 and copy the formulae right to column EY.
3. Now return to your home screen and look at the Expected Sustained Yield graph. Notice that expected sustained yield is a relatively flat function for a broad range of S. The bands represent 80% and 90% interval estimates. Generally, yields are predictably low at small escapements but can be unpredictable at higher escapements. There can sometimes be an important trade-off between high yield and predictability.

3) Let’s use the tools we have developed to think about the process of picking a goal with a range of the form X–2X. That is, one in which the maximum escapement is about twice the minimum escapement. There is nothing special about the X-2X form, except that many existing escapement goals will take a form close to this one. In this lab we are using the X-2X form to set an initial goal, which we will refine. Thus, the lab mirrors the actual escapement goal review process.

1. Using your optimum yield profile, find an approximate EG range with the X-2X property that also maximizes the probability of achieving 90% of maximum sustained yield (\_\_\_\_\_\_ , \_\_\_\_\_\_).
2. Note that the optimum yield profile and the expected sustained yield plot have common X-axes that are lined up with one another. Move the vertical red lines left and right until they line up with your tentative goal from (III 3 a) above. How does your tentative goal perform with respect to expected sustained yield?
3. Once you have selected an *escapement goal range, the expected sustained yield plot and the optimal yield probability profile can be used to qualify your proposed goal, in the same way that confidence intervals qualify your estimates of ln(), , and SMSY.* An optimal yield profile with steep sides and that peaks near 100% probability indicates that you can be fairly certain that your escapement goal will result in near maximal sustained yield. Conversely, a profile that peaks at only 50% probability indicates less certainty of achieving optimal yields. For your stock, complete the following description of your proposed escapement goal.

***“If we maintain the spawning escapement between \_\_\_\_\_\_\_\_ and \_\_\_\_\_\_\_\_ fish, we can be \_\_\_\_% to \_\_\_\_% certain that we will achieve at least 90% of the maximum sustained yield for the stock. At these escapements, we can expect sustained yields between \_\_\_\_ and \_\_\_\_\_\_.***

The first percentage is the height of the horizontal gridline chosen to define the EG range, and the second percentage is the maximum height of the optimum yield profile. The remaining two values can be read from the expected sustained yield plot. The first number would be the minimum value of the lower 90% confidence interval over the range of the proposed escapement goal, and the second number would be the maximum value of the upper 90% confidence interval for the goal range. *We recommend that a statement similar to the one above accompany any proposed escapement goal range that is produced using a stock-recruit analysis.* Ideally, the diagnostic graphics (optimum yield profile and expected sustained yield plot) should also be presented.

4) In the real world, you would consider other factors that bear on selection of a goal. These might include: the range of the current goal, the power of the fishery and in-season management capability, the range of recent escapements, the quality of stock assessment, status of nearby stocks, subsistence considerations, etc. Think about the harvest rates you were assigned.

Does an EG range with a X-2X property make sense? Do you think a narrower or wider goal range is appropriate with the harvest rates associated with your data?

*Low harvest rates suggest ADF&G lacks the ability to control escapement within a narrow goal.*

By attempting to maximize the probability of achieving 90% of maximum sustained yield you are likely choose a symmetric goal (one with the same probability of achieving 90% of maximum sustained yield at both ends). Non-symmetric goals sacrifice yield at one end of the goal range (relative to the other end of the goal range). Is a symmetric goal ideal in this case?

Which end would you change and how do you justify the yield you sacrificed?

*Yield will be sacrificed by increasing the upper bound of a symmetric escapement goal.*

**V. Percentile Goals**

In the really real world, you may not have the data to construct the brood table in the first place, and all of these tools will not be available to you. In this situation percentile goals are used as a proxy for the full analysis conducted above.

|  |  |  |
| --- | --- | --- |
|  | Situation | Percentiles |
| Tier 1 | High contrast (>8), high measurement error, low-moderate harvest rate (<40%) | 20th-60th |
| Tier 2 | High contrast (>8), low measurement error, low-moderate harvest rate (<40%) | 15th-65th |
| Tier 3 | Low contrast (<8), low-moderate harvest rate (<40%) | 5th-65th |
|  | High harvest rate (>40%) | **Not recommended, improve stock assessment** |

Select the appropriate tier and enter the percentile values in cells Q56:R56. The goal will be calculated and plotted on the optimum yield profile for you (Note: Don’t use this workbook to calculate percentile goals outside of the lab. When calculating a percentile goal you want to use all available escapements, not just those for which you have the associated recruitments.).

How does this goal compare to the goal you created when SR information was available?

**Notes**

Use the completed (and error-checked!) version of the spreadsheet (“Ricker Data”) when analyzing other datasets. This sheet is like the one you just completed but removes references to the simulation data and parameters. Right click any sheet tab and select Unhide then “Ricker Data”. This worksheet can handle up to 50 years of data with little or no modification.

**Equations**

Ricker ln() corrected for log transformation bias

|  |  |
| --- | --- |
|  | (1) |
| Spawning escapement where recruitment is maximum |  |
|  | (2) |
| Spawning escapement at equilibrium under no harvest (= carrying capacity) |  |
|  | (3) |
| Spawning escapement at which sustained yield is maximized |  |
|  | (4) |
| Sustainable harvest rate at Smsy |  |
|  | (5) |
| Recruitment at Smsy |  |
|  | (6) |
| Maximum sustained yield |  |
|  | (7) |

1. Percentile estimates, frequency distributions, the horsetail plot, and the graphics of section III are all subject to some volatility, because each time you recalculate (key F9) the bootstrap replicates are re-randomized. [↑](#footnote-ref-2)
2. This calculation is the same type you would make when you estimate the probability of heads on a coin toss; you toss the coin lots of times, add up the heads(1’s) and divide by the number of tosses. [↑](#footnote-ref-3)