**Introduction**

IMPORTANT NOTE: This assignment can be completed only in the SAL classroom or on another properly equipped Windows-based computer.

Today we learn about SHIRAZ (Salmon Habitat Integrated Resource Analysis: Zowie!), a salmon life-cycle model that incorporates anthropogenic effects into fish-habitat relationships. SHIRAZ was developed under the supervision of UW Fisheries Professor Ray Hilborn (see link on the SHIRAZ lab page). The model allows for future projections of salmon population sizes as affected by habitat variables that are discussed below.

It is important to note that SHIRAZ is closely connected to DHSVM, the distributed hydrology soil-vegetation model (See flowchart below). NOAA (National Oceanic & Atmospheric Administration) is currently using both of these models to predict the effects of anthropogenic factors on the salmon life cycle for Endangered Species Act (ESA) recovery planning.

Recall that DHSVM uses such properties as land cover, soil, topography, and precipitation to predict stream flow and stream temperature. These two outputs from DHSVM are then used as inputs to drive the SHIRAZ model. SHIRAZ uses information on stream flow & temperature, sediments, and other habitat quality indicators to project future salmon populations by size of stock, life stage, and location.

SHIRAZ uses a number of inputs to run the model, but we will focus on four primary inputs:

1. Adult spawner capacity (number of fish for which suitable spawning habitat is available along the length of the creek)
2. % fine sediments (<0.65 mm diameter) in stream bottom gravel
3. Stream temperature during summer pre-spawning period (affects adults returning to spawn)
4. Stream temperature during winter incubation period (affects eggs and fry in the stream)

NOAA is currently applying SHIRAZ to recovery planning for chinook salmon in the Snohomish Watershed (WRIA 7). We will study a specialized version of SHIRAZ that has been adapted for summer chum salmon in the Big Beef Creek watershed (WRIA 15). Both of these stocks are listed under the ESA. SHIRAZ is a complicated model with numerous parameters, thus for our purposes today we will be using a simplified version to more easily discern the impacts of and relationships between the four primary input variables when predicting future salmon populations.

There are several more important differences between the SHIRAZ and DHSVM labs:

* In the previous lab, you used Excel to look at packaged output from previous runs of DHSVM. This week you will actually run SHIRAZ using Excel for input and output.
* DHSVM produces hour-by-hour time-varying results in response to time-varying input such as rainfall. This version of SHIRAZ assumes that the freshwater habitat remains constant. It uses constant values of the four primary variables over a 50-year period (assumes conditions are identical every year) and produces a prediction of what salmon populations should be under steady-state or equilibrium conditions. The early model time-varying results are generally mathematical artifacts.

**Part I: SHIRAZ Model Output: 50-Year Spawner Populations (2 points)**

We will start by looking at some of the output from SHIRAZ, and then we will go back to look at how the model was set up to produce those results.

Open the Excel file on the desktop called “ShirazBigBeefSuChum.xls.” (The file is also available for download from the SHIRAZ page on the course web site.) When the dialog boxes open, select, “Enable macros.”)

We will use three worksheets in this spreadsheet (the three leftmost bottom tabs): “Input,” “Output\_AreaAves,” and “Output\_SR.” (The rest of the worksheets contain detailed supporting information that is beyond the scope of this assignment.)

To begin, select the “Input” worksheet from the bottom tab. In tables at the upper left, you will see a box containing the input data values of the four key variables (plus some others that are less important) selected. Values are highlighted in yellow.

Four experimental model runs or “Scenarios” have already been made using different arbitrary values of these inputs. Scenario 1 (at the top) contains typical data for Big Beef Creek under present conditions. Scenarios two–four vary these data to test how salmon population predictions respond in the model.

Note that data are presented for “Section 1” and “Section 2” (in columns B & C). Big Beef Creek is about 16 km long, but chum salmon spawn only in two areas nearest the creek mouth, each 0.5 km long. These two areas are tabulated separately to account for the possibility of differing habitat conditions or alterations.

|  |  |
| --- | --- |
| **Scenario 1 (Current)** | **Location** |
| **Key habitat variables** | **Section1** | **Section2** |
| **adult spawning capacity** | **1,000** | **1,000** |
| **% fines in gravel** | **50.00** | **50.00** |
| % embeddedness | 25.00 | 25.00 |
| **temperature pre-spawning (°C)** | **12.00** | **12.00** |
| **temperature - incubation period (°C)** | **6.00** | **6.00** |
| scour depth (cm) | 10.00 | 10.00 |
| % cobble and boulder in pools | 25.00 | 25.00 |

|  |  |
| --- | --- |
|  | Northern (lower) reach of Big Beef Creek to its mouth where it enters Hood Canal. Shown are the two .5-km Sections of the creek just upstream of the mouth where summer chum salmon are observed to spawn. |

Also on this worksheet, in a box highlighted orange to the right off the values box (columns E–G), are the “Outputs,” the mean total numbers of returning salmon (per year) that the model calculates for each scenario (summing the two creek sections). These data are summaries of the SHIRAZ data output for sections 1 & 2 for each scenario. We will examine these results summaries in more detail in the next part of the assignment. First, we will look at the raw results.

|  |  |
| --- | --- |
|   | **Output** |
|   | **mean** | **sd** |
| **Sc1** | 873 | 267 |
| **Sc2** | 645 | 232 |
| **Sc3** | 1,547 | 393 |
| **Sc4** | 1,598 | 211 |

Next we will look at the kind of raw output data that the model generates. Open the “Output\_SR” worksheet from the bottom tab. The results of the four scenarios are shown in groups of columns, beginning with the baseline Scenario 1 in columns A–H (see sample below).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **A** | **B** | **C** | **D** | **E** | **F** | **G** | **H** |
| Year | Spawners | Recruits | R/Sp,Sp<=0.75alpha | Rpredicted | NLL | SPhypoth | Rpredicted |
| 2001 | 1085.114943 | 430.9280658 | 0.397126654 | 908.4100595 | -7.50618359 | 200 | 271.8914723 |
| 2002 | 697.9441101 | 571.7028225 | 0.819124073 | 702.3178707 | -6.599248031 | 300 | 380.9824882 |
| 2003 | 1741.207895 | 803.9030141 |  | 1134.554512 | -7.116883051 | 400 | 476.5944926 |
| 2004 | 1537.052235 | 1639.417462 |  | 1075.725875 | -7.965731903 | 500 | 561.080282 |
| 2005 | 652.9388649 | 410.9500746 | 0.629385225 | 672.8296582 | -6.733844517 | 600 | 636.2751791 |
| 2006 | 787.9048695 | 1023.504994 | 1.299021029 | 757.2970567 | -7.293679191 | 700 | 703.6318848 |
| 2007 | 1201.929985 | 1649.898891 | 1.372707988 | 957.6445882 | -8.246224812 | 800 | 764.3152482 |
| 2008 | 901.9710807 | 836.4082286 | 0.927311581 | 820.3013194 | -6.882592107 | 900 | 819.2701987 |
| 2009 | 922.0701627 | 845.5641996 | 0.917028046 | 830.7072442 | -6.893331506 | 1000 | 869.2713067 |
| 2010 | 1293.680991 | 1156.579526 | 0.8940222 | 993.1049882 | -7.259586244 | 1100 | 914.9595201 |

Notice that the leftmost column (Column A in Scenario 1) contains time in years, beginning in 2001. This version of the SHIRAZ model begins with an arbitrarily chosen number of spawners (around 1000) and runs for 50 years. Because of the length of the salmon life cycle, that yields 43 years of data, from 2001 through 2043. In the next column to the right (Column B in Scenario 1) are the numbers of adult salmon that returned to spawn (“spawners”) in each year, and in the next column to the right (Column C in Scenario 1) are the calculated numbers of the next year’s “recruits” (young fish that survived through the freshwater egg and fry stages to enter salt water) produced by the spawners that returned that year. This type of data stream is called a “time series.”

As we have described SHIRAZ so far, the model would run under a given set of conditions (specified as the input data for the scenario), and the number of spawners each year would gradually reach a constant value, the steady-state or equilibrium value for the conditions of that scenario. In this model, it should reach that equilibrium within about 20 years.

Notice, however, that when the year 2043 is reached in the output data, the years start over again at 2001. This version of the model is set to run each scenario three times, each time making random annual variations in three variable factors (in addition to the spawning capacity, %fines, and temperatures, which are assumed constant in this model be) that determine survival to return:

* the fishing harvest
* the survival rate of fish in the early (estuary) salt-water parts of its life cycle
* the survival rate of fish in the adult (open ocean) salt-water parts of its life cycle.

These repetitions are intended to simulate the range of fish returns that might occur as a result of natural variations in environmental conditions. Thus the number of spawners in the model time series will never reach the equilibrium value because conditions are constantly changing, as they do in the real world.

# (2 pts.) Using the graphing skills in Excel you learned in the DHSVM lab, graph the modeled changes in annual spawner returns 2001–2043 for Scenario 1.

## (0.5 pts.) Select the Year and Spawner columns for just the first repetition (2001-2043) of the model run and use the Chart Wizard. Select the XY (Scatter) type of graph, with the points connected by lines. Look at the resulting scatter of points over time, ignoring the first 20 years when the model would not be in equilibrium regardless of the changing inputs. Somewhere near the center of that range is the likely mean number of spawners over a range of conditions. Just by visual inspection, what would you estimate that mean to be?

## (0.5 pts.) Repeat step (a) two more times using the second and third repetitions (2001-2043) of the model run. You should now have three separate timelines of fish abundance on one graph. Just by visual inspection gain, what would you estimate the likely mean number of spawners to be for all three of these runs combined?

## The “Standard Deviation”  is a statistical measure of the variation above and below the mean. In theory, for a classic “bell-shaped curve” (“normal distribution”), about 95% of all occurrences should fall within 2 above and 2 below the mean (i.e., a total spread of 2).

|  |  |
| --- | --- |
| From Wikipedia.org | Dark blue is less than one standard deviation from the mean. For the [normal distribution](http://en.wikipedia.org/wiki/Normal_distribution), this accounts for 68.27 % of the set; while two standard deviations from the mean (medium and dark blue) account for 95.45%; three standard deviations (light, medium, and dark blue) account for 99.73%; and four standard deviations account for 99.994%. |

## (0.5 pts.) Look again at the resulting scatter of points over time after the first 20 years. What range of values appears to encompass about 95% of the range of spawner returns (after the first 20 years)? What would the corresponding value for  be?

## Look at the “Output\_AreaAves” worksheet. This worksheet shows the actual calculated mean (average) spawner return (Column C) and standard deviation (Column E), which you estimated visually in 1a) – 1c) above. (These means also ignore the first 20 years of model data.) For now we will look at only Scenario 1 (top set of numbers). Notice that the output shows means & standard deviations for all of the 0.5-km areas along the creek, but as noted above, only the first two areas support salmon spawning.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Scenario 1.** |  | **Means** |  | **Standard deviations** |
| **HUC code** | **Area name** | **Natural** |  | **Natural** |
|  | **All areas** | **873** |  | **267** |
| 1 | R0.5KM1 | 437 |  | 133 |
| 2 | R0.5KM2 | 436 |  | 133 |
| 3 | R0.5KM3 | 0 |  | 0 |
| 4 | R0.5KM4 | 0 |  | 0 |

## (0.5 pts.) How do your visual estimates of the mean and standard deviation for Scenario 1 compare to the calculated value? (Use the summed value for all areas = Sections 1+2.)

##

**Part II: SHIRAZ Model Input & Output: Effects of Sediments (11 points)**

This Big Beef Creek version of SHIRAZ has already been run using four scenarios incorporating different combinations of values of the four key variables. We will begin by looking at these values and the model predictions that SHIRAZ produces using them.

# (0.75 pts.) Return to the “Input” worksheet to compare the inputs and outputs from Scenarios 1 and 4. In this comparison, there is no difference between the model parameters for Sections 1 & 2 of the creek.

# (0.25 pts.) What are the differences in input values between these two scenarios?

# (0.25 pts.) What is the difference in total salmon spawner “Output” (average number of salmon returning per year for Sections 1 & 2 combined) that results from these changes in input values?

# (0.25 pts.) What can you conclude about the effect of this input value on salmon returns, at least as it is represented in the model?

# (2.25 pts.) Compare Scenario 1 to Scenarios 2 & 3.

# (0.25 pts.) What input values are changed between these Scenario 1 and the other two scenarios?

#  (1.0 pts.) In what way (increase or decrease) is each of these input values changed in Scenarios 2 & 3?

# (0.5 pts.) What is the difference in total salmon spawner “Output” (Sections 1 & 2 combined) that results from these changes in input values in each scenario (#2 & #3)?

# (0.5 pts.) What change(s) in input values do you think might have caused these differences in “Output?” Can you be sure of your answer based only on these output data?

##

# To the right of the box of model outputs on the “Inputs” worksheet are four graphs. The lower three graphs are labeled “Key functional relationships.” Each of these depicts the results of an experimental study published in the paper cited below the graph. We will use these graphs to diagnose the causes of Ouput differences between Scenarios 1, 2, &3.

# (2.5 pts.) Begin by selecting the leftmost bottom graph, labeled “% Fines effect on egg to fry survivorship.” SHIRAZ uses a formula fitted to these data to calculate the effect of fine sediments on survival.

# (0.5 pts.) Based on this graph, describe how egg survival changes as the percentage of fine sediments increases. What is the threshold value of % fines where survival begins to change rapidly?

##

# (0.5 pts.) In what way do fine sediments biologically affect salmon eggs in creeks to produce the results this graph illustrates?

##

# (0.5 pts.) What kinds of alterations in the watershed can cause changes in the percentage of fine sediments in the creek?

##

# (0.5 pts.) Compare the percentage of fine sediments used in the baseline Scenario 1 to the graph. What can you say about salmon egg survival under present conditions of fine sediments in Big Beef Creek (Scenario 1)?

##

# (0.5 pts.) Return to the comparison between Scenarios 1 & 4. Do the model results agree with what you would expect from viewing the graph of “% Fines effect on egg to fry survivorship?” Why or why not?

##

# (2.5 pts.) Return to the “Output\_AreaAves” worksheet. Recall that this worksheet shows (in column C) the mean number of salmon returning per year in each of the two river Sections modeled under each of the four scenarios.

# In #3c above, with the two Sections of the creek combined, we could not separate the effects on decreased spawning capacity from the effects of % fine sediments on egg survival. Now compare Scenario 1 to Scenario 2 and compare the magnitude of the change in mean returns caused by i) differences in adult spawner capacity and ii) differences in the percentage of fine sediments. Note that the spawning capacity was changed in only one of the two river Sections when comparing Scenario 1 to Scenario 2, and it was held constant in the other Section. But the %fine sediments was changed in both sections when comparing Scenario 1 to Scenario 2.

## (0.5 pts.) Which creek section allows you to determine the effects of changing only the percentage of fine sediments in the model without changing the spawning capacity between Scenarios 1 & 2? What effect do you observe?

## (0.5 pts.) Does this result agree with what you would expect? Why or why not?

##

## (0.5 pts.) .) Which creek section allows you to determine the effects of changing adult spawning capacity alone in the model between Scenarios 1 & 2? What effect would you expect to observe from this change?

## (0.5 pts.) What change in mean returns do we actually observe for this section between Scenarios 1 & 2 in the model? How well does this change agree with your prediction?

##

##  (0.5 pts.) Which change (spawning capacity or %fine) had the greater effect in these model scenarios?

##

# (2 pts.) Repeat the comparison of mean returns for both sections, this time between Scenarios 1 and 3.

## (0.5 pts.) Which creek section allows you to determine the effects of changing only the percentage of fine sediments in the model without changing the spawning capacity between Scenarios 1 & 3? What was the effect of changing only the percentage of fine sediments?

##

## (0.5 pts.) Use Section 2 to compare the magnitude of the change in mean returns in Scenarios 1 & 3 caused by i) differences in adult spawner capacity and ii) differences in the percentage of fine sediments. What magnitude of change would you expect from the change in adult spawning capacity alone?

## (0.5 pts.) What change in mean returns do we actually observe?

##

## (0.5 pts.) Which change (capacity or %fine) had the greater effect?

##

**Part III: Estimate Statistical Significance of Model Results (4 points)**

Recall from Part I (Question 1) that the model output for salmon returns are *means* (averages) of the three model runs shown in the “Output\_SR” worksheet. The standard deviation quantifies the degree of variation from that mean that occurs in the results, and is tabulated for each Scenario—and for each section of the creek and for sections 1 & 2 combined—in the “Output\_AreaAves” worksheet. We can use the standard deviation to consider whether differences in results from different sets of conditions in the model are statistically significant or are more likely just the result of random variations. (This is analogous to the margin of error in presidential polls and other such data.

# (2 pts.) Refer to the “Output\_AreaAves” worksheet. Note that it contains a bar graph showing the total spawners returns (creek Sections 1+2 summed) calculated in model Scenarios 1–4.

## (0.5 pts.) How can you read the total spawners returns on the graph?

##

## (0.5 pts.) How can you read the standard deviations on the graph?

##

## According to statistical procedures, two sets of data are statistically different only if their ranges of values have minimal overlap. Two sets of data are commonly considered to be statistically different if 95% or more of the values are different. We will not do a quantitative test of statistical differences, as a professional analyst would do. But recall that two standard deviations on each side of the mean (total of four standard deviations) contain about 95% of the values of a normal statistical distribution.

## (0.5 pts.) Looking visually at the chart in the “Output\_AreaAves” worksheet, which of the four scenarios are statistically indistinguishable from each other? How can you tell?

##  (0.5 pts.) Which of the four scenarios are statistically different? How can you tell?

# (1 pt.) We can make a more accurate judgment about whether the results of the various scenarios are statistically different by using simple calculations. Begin by filling in columns 3 & 4 of the table below using the data in the “Output\_AreaAves” worksheet. Fill only the rows labeled “All” (i.e, the sums of sections 1 & 2).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Scenario** | **Creek Section** | **Mean spawner return** | **Standard deviation** | **95% lower confidence limit (mean - 2** | **95% upper confidence limit (mean + 2** |
| 1 | All 1+2 |  |  |  |  |
|  | *1* |  |  |  |  |
|  | *2* |  |  |  |  |
| 2 | All 1+2 |  |  |  |  |
|  | *1* |  |  |  |  |
|  | *2* |  |  |  |  |
| 3 | All 1+2 |  |  |  |  |
|  | *1* |  |  |  |  |
|  | *2* |  |  |  |  |
| 4 | All 1+2 |  |  |  |  |
|  | *1* |  |  |  |  |
|  | *2* |  |  |  |  |

## (0.25 pts.) Next subtract two standard deviations (Column 4) from each of the means (Column 3) and enter in Column 5 to estimate the value of spawner returns which at least 95% of all values would be expected to exceed (95% lower confidence limit). Again, fill only the rows labeled “All” (i.e, the sums of sections 1 & 2).

## (0.25 pts.) Next add two standard deviations (Column 4) to each of the means (Column 3) and enter in Column 5 to estimate the value of spawner returns which at least 95% of all values would be expected to be less than (95% lower confidence limit). Again, fill only the rows labeled “All” (i.e, the sums of sections 1 & 2).

## (0.25 pts.) Now scan the pairs of lower and upper confidence limits. Do you see any pairs of confidence limits in the rows labeled “All” (i.e, the sums of sections 1 & 2) that have little or no overlap with a pair from another Scenario?

## (0.25 pts.) How would you explain these results biologically?

**Part IV: Generate New SHIRAZ Model Data (5 points)**

This part of the assignment can be completed only in the SAL classroom or on another properly equipped Windows-based computer. In this section you will actually change some if the input values and run the SHIRAZ model yourself. There are many possible variations that we could use to further explore salmon responses to environmental changes as they are modeled in SHIRAZ. However, we will take just a couple of exmples to demonstrate the model operations.

Notice that incubation (6˚C) and pre-spawning (12˚C) temperatures were the same for all the scenarios run so far.

# (1 pt.) Return to the “Input” worksheet and view the graphs of temperature effects on salmon used in this version of SHIRAZ.

# (0.5 pts.) How low would the temperature have to fall to affect egg survival during the incubation stage? How high would the temperature have to rise to affect egg survival during the incubation stage?

# (0.5 pts.) How low would the temperature have to fall to affect spawner survival during the pre-spawning stage? How high would the temperature have to rise to affect spawner survival during the pre-spawning stage?

## IMPORTANT: Make a copy of the “ShirazBigBeefSuChum.xls” spreadsheet on the computer desktop and be sure to give it a new name. Open this copy to work on the next question so that you do not lose the original values.

# (2 pts.) In the box at the top of the “Input” worksheet, change the value of “Number of Scenarios” from 4 to 1. In the input data box for Scenario 1, change the incubation temperature to **3 degrees** for both Sections 1 & 2. Go to the “SHIRAZ” menu at the top of the screen and select “Run.” This will rerun 3 repetitions of Scenario 1 and replace the output data in all the worksheets with new values for this scenario.

# (0.5 pts.) What are the new values of the mean number of spawners and the standard deviation for Scenario 1?

# (0.5 pts.) How would you explain these results biologically?

##

#  (0.5 pts.) Does the new mean number of spawners appear to be significantly different from the original number for Sections 1+2 in the original Scenario 1? How can you tell?

#  (0.5 pts.) What can you conclude about whether you have proven that the change in incubation temperature that you made has a real effect on salmon survival in the model? Why?

## IMPORTANT: Make another copy of the “ShirazBigBeefSuChum.xls” spreadsheet on the computer desktop and be sure to give it a new name. Open this copy to work on the next question so that you do not lose the original values.

# (2 pts.) In the box at the top of the worksheet, change the value of “Number of Scenarios” from 4 to 1. In the input data box for Scenario 1, change the pre-spawning temperature to **18 degrees**. Go to the “SHIRAZ” menu at the top of the screen and select “Run.” This will rerun 3 repetitions of Scenario 1 and replace the output data in all the worksheets with new values for this scenario.

# (0.5 pts.) What are the new values of the mean number of spawners and the standard deviation for Scenario 1?

# (0.5 pts.) How would you explain these results biologically?

#  (0.5 pts.) Does the new mean number of spawners appear to be significantly different from the original number for Sections 1+2 in the original Scenario 1? How can you tell?

# (0.5 pts.) What can you conclude about whether you have proven that the change in pre-spawning temperature that you made has a real effect on salmon survival in the model? Why?