Survey Tutorial

Eura Nama

# Survey Tutorial Project Outline

1. Your job is to design an annual trawl survey for Dusky Scallop Shark (*Dustious maximus*; Figure 1). The survey occurs June when the stock is spawning.
2. The stock area shown has been chosen as a compromise between the stock biology, stakeholder input, and operational constraints (Figure 1).
3. In planning your survey there are 2 variables you can control
   1. You can vary the number of survey tows
      * From start to finish it takes approximately 3 hours to complete a tow and finish the scientific analyses
      * The vessel is available for the entire month of June
   2. You can decide if you want to have a random or stratified survey
      * There are 2 options for stratification
        + NAFO sub-areas (Figure 2)
        + Depth strata (Figure 3)
      * Note that the tows will be allocated proportionally to the area of each stratum.
4. You can also ‘simulate’ different biomass distributions to see if this changes your opinion on either the stratification scheme or the number of stations you need.
   1. The “Random” option distributes the biomass using a ‘random field’
      * There is a spatial pattern to the data but the pattern is randomly generated
   2. The “NAFO” option has an underlying ‘random field’ which sets the pattern for each realization
      * Then this field is augmented depending on which NAFO strata you are in
      * The biomass in the central area is elevated
      * The biomass in the nearshore and offshore areas are lowered
   3. The “Depth” option also has an underlying ‘random field’ which sets the pattern for each realization
      * Then this field is augmented by each Depth strata
      * The biomass in deeper waters are lowered

# Questions to Consider

1. How does increasing the number of tows influence the accuracy and the precision of the results?
   * Consider the trade-offs between the number of tows and accuracy and precision of the survey results with logistical constraints of running a survey
2. When you have few survey stations (e.g. ≈20) why are the stratified survey biomass estimates generally biased?
   * *Hint* Look at the number of stations in each of the NAFO stratum (Figure 5)
3. Consider Figure 1 and discuss how biological, social, economic, or political factors could influence the design of the survey of the (*D. maximus*) stock (population) in this Region.
   * *Hints*:
     + Biological versus artificial boundaries
     + Canada versus United States
     + Survey occurs during spawning
4. Consider Figures 4 - 6 and Table 2.
   * Which of the 3 sampling methods (Random, NAFO stratification, Depth stratification) would you suggest to use for the survey and why?
   * Does changing the underlying biomass distribution effect your opinion?
5. How many survey stations would you recommend?
   * Consider the ‘constraints’ on your available time in Part “C” of the Tutorial Outline above.

# The Tutorial Tutorial

1. Open the file “Survey\_tutorial.Rmd” in R-Studio
2. On line 83 of this file you can change the number of tows used in the survey. Default is 20 tows
3. On lines 86-88 of this file you can change the underlying biomass distribution
   * To do this you add a # symbol at the start of the lines you do not want, and removed it from the line you do want
4. You can change the number of realizations run on line 94
   * I suggest trying 1, 5, and 250 (this will be slow!) to see if/how your opinion changes with more realizations of your data.
5. Press the “Knit” button in the tool bar
   * Make sure that you closed (and saved if you want to keep the results) the word document.

# Survey Parameters

Just so you can keep track in the document, we return the survey parameters that were set for this simulation. You only have control over the Number of Tows and the underlying “Biomass distribution”. The other parameters are automatically set for you in the code.

Table 1: A Table of your input values for the current run of your simulation

| Parameter | Value |
| --- | --- |
| Number of Tows | 20 |
| Biomass | 100000 tonnes |
| Catchability | 0.3 |
| Area swept by a tow | 10000 m² |
| Number of Realizations | 1 |
| Biomass Distribution | Random |

## Survey Simulation

So now we can review the input data we have for our survey. First we will look at some figures. First off, lets take a look at our survey area, included in this figure are the North Atlantic Fishery Organization (NAFO) subareas that are the basis for the NAFO stratification, and the bathymetry of the region, which is used as the basis of the depth stratification (Figure 1).

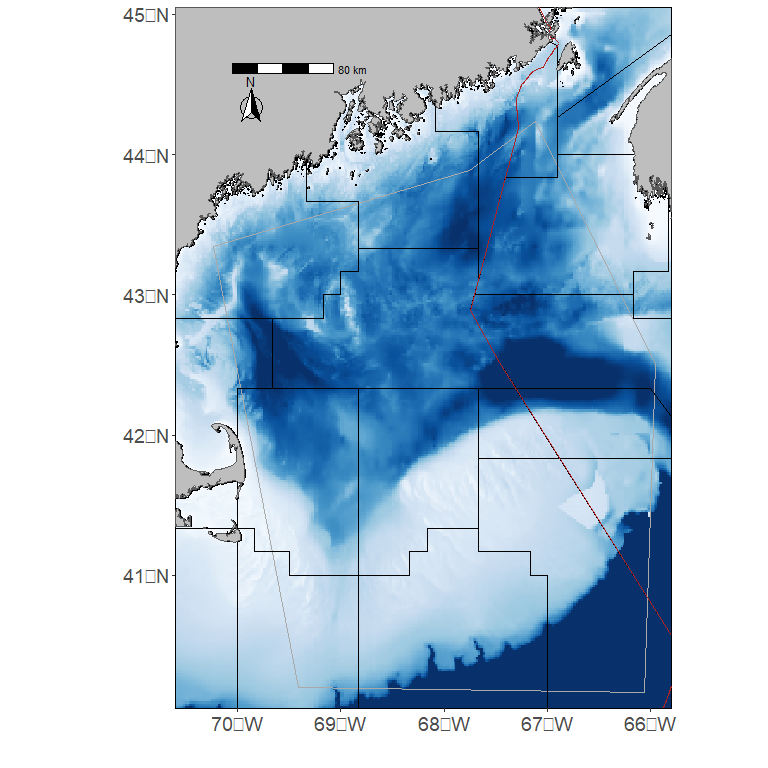


Figure 1: The assessment area for the Dusky Scalloped Shark (*Dustious maximus*) is outlined by the thick grey line. The thin black lines are the NAFO subareas in the region. The red line divides shows the division between the economic exclusive zone (EEZs) for Canada and the United States. The bathymetry in the region is also shown.

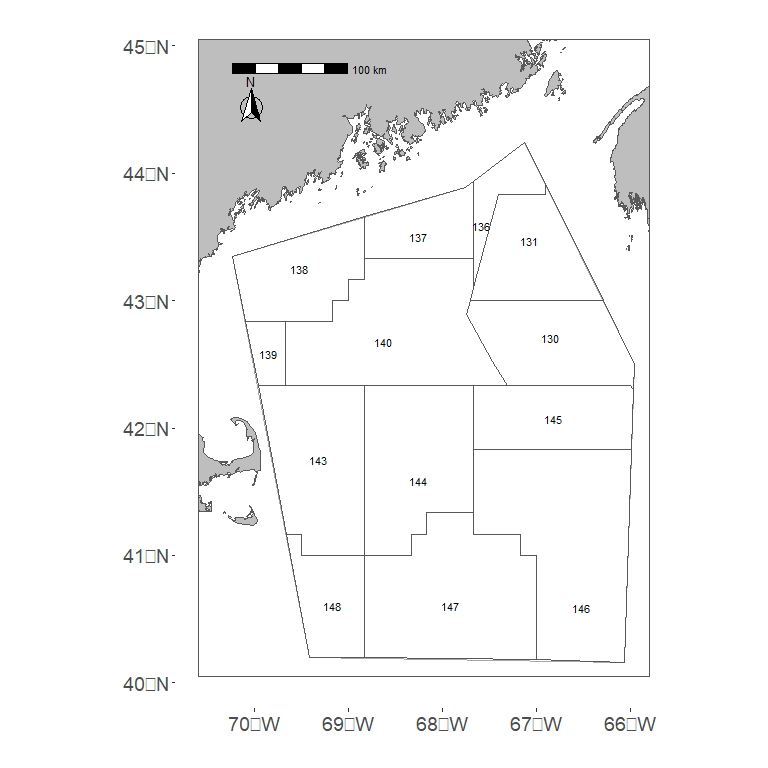


Figure 2: The NAFO stratification polygons used for stratification.

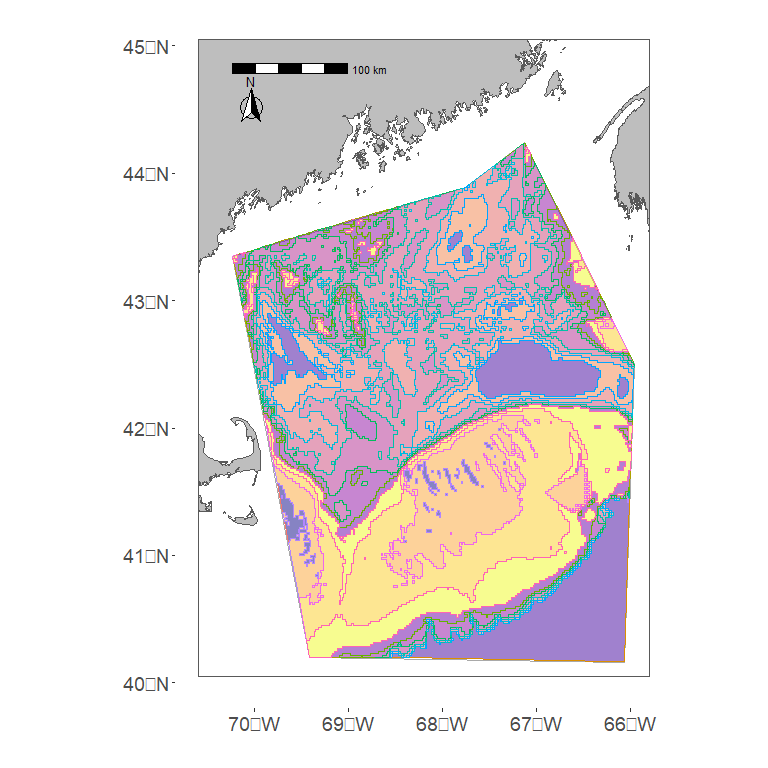


Figure 3: The Depth stratification polygons used for stratification.

Now we can also show the distribution of the biomass in the area. If 1 is greater than 1 then we’ll show two or three realizations from the models depending on how many realizations we ran. First we show the biomass distribution with the random survey stations overlain (Figure 4).

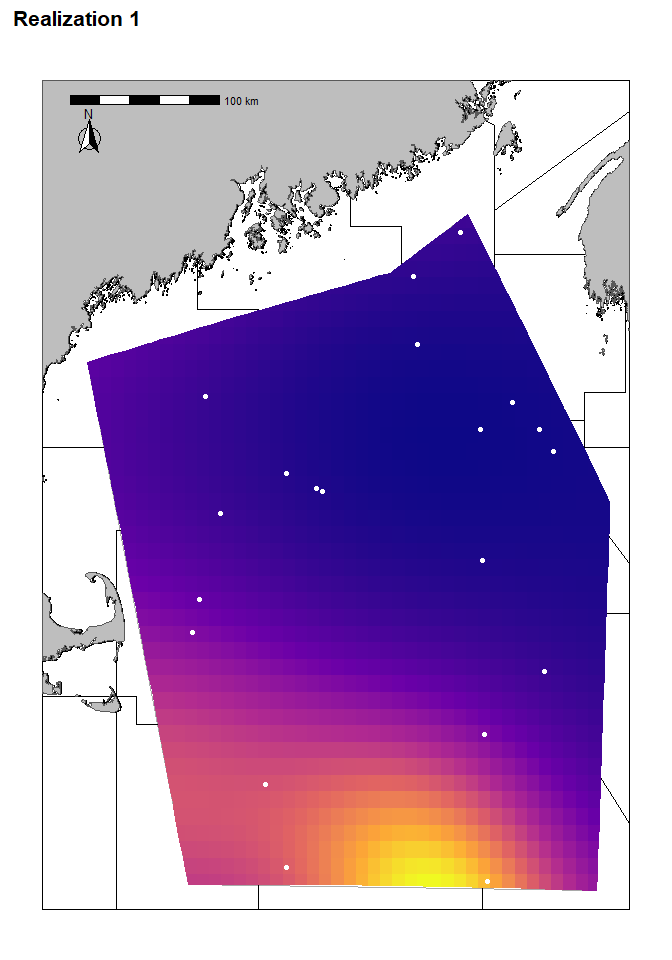


Figure 4: Biomass distribution with the random survey stations overlain

Next we show the biomass distribution with the NAFO survey stations and NAFO strata overlain (Figure 4).

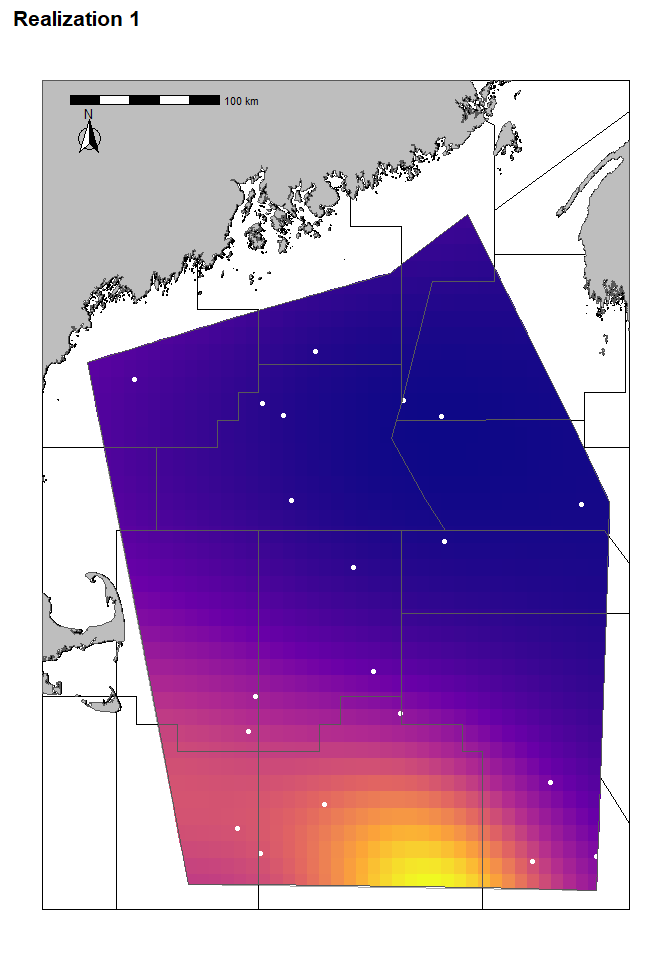


Figure 5: Biomass distribution with the NAFO survey stations and NAFO stratification polygons overlain

Finally, we show the biomass distribuiton with the Depth survey stations and Depth stratification overlain (Figure 6)

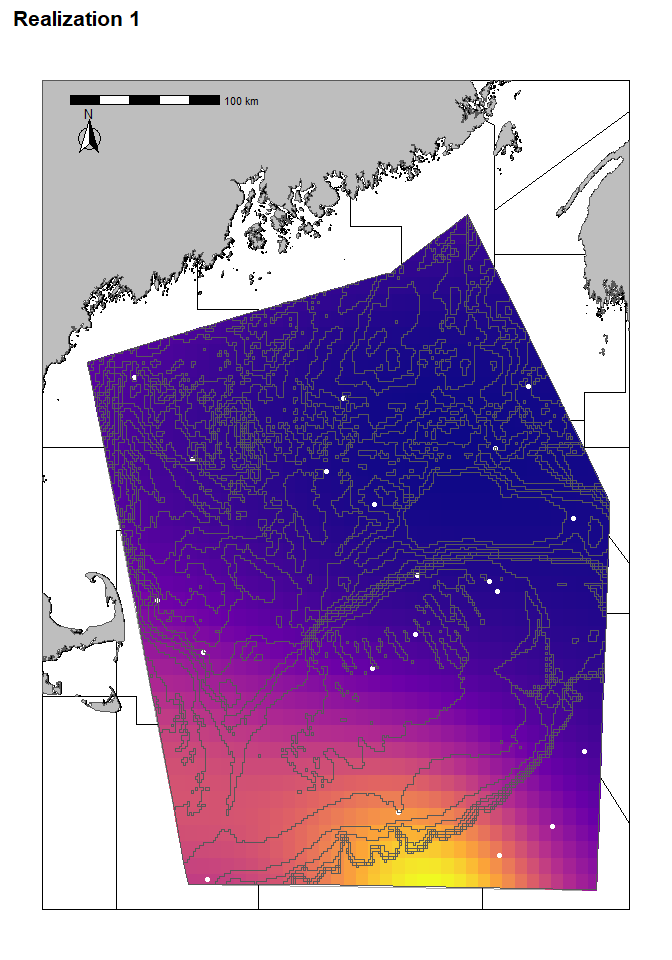


Figure 6: Biomass distribution with the Depth survey stations overlain and the Depth stratification polygons overlain.

# Now we can compare the random survey estimates to the depth and NAFO stratified surveys.

## Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.  
## ℹ Please use `linewidth` instead.

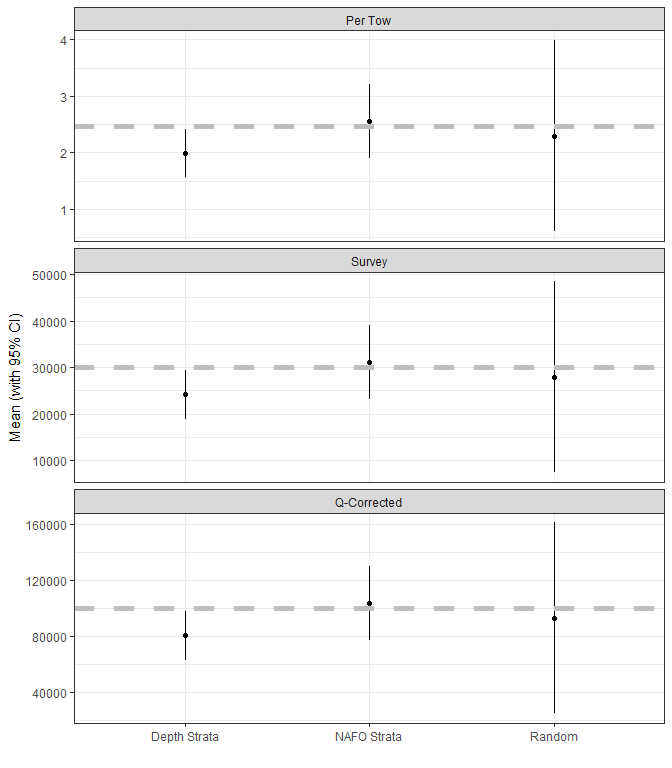


Figure 7: Biomass estimates from the 3 different survey sampling schemes. When the number of realizations run = 1 this provides the mean and 95% CI from that realization. When the number of realizations is >1 and < 10 the mean biomass for each realization is shown. When the number of realizations is 10 we show the median biomass of the realizations along with the interquartile range of the biomass from the realizations

Table 2: A Table of the results from your simulation when a single realization is performed. The Biomass is the mean, while the Upper and Lower CIs are the 95% confidence interval calculated from the survey.

| Tow Stratification | Biomass | Lower CI | Upper CI | Number of Sims | Biomass Dist |
| --- | --- | --- | --- | --- | --- |
| Random | 93070 | 24640 | 161500 | 1 | Random |
| NAFO Strata | 103600 | 76900 | 130300 | 1 | Random |
| Depth Strata | 80560 | 63140 | 97990 | 1 | Random |

# Terms and Jargon

*Important*: Pretty much all of these ‘definitions’ are simplifications that serve our purposes, but if you go deeper into Fisheries Science you will see that it is often much more complex than this.

* *Adults:* Mature individuals that are targeted by the fishery
* *Recruits:* Individuals that will mature and be targeted by the fishery next year
* *Biomass:* The Abundance of individual in a certain size/age class multiplied by the weight of those indivduals. Usually the weight is taken as some sort of average of the population
* *Growth:* The increase in size (mass) of a class of fish in a given year
* *Survey:* The process in which a stock is sampled to get an estimate of population status (e.g. abundance index) and life history parameters (e.g. growth)
* *Survey Design:* The method used to carry our a survey. For our purposes this includes the number of sampling stations and the type of stratification (if any) to use.
* *Stratification:* A method used to divide up a survey into areas with ‘similar’ characteristics. If done properly it will reduce the uncertainty of your survey indices compared to a survey in which the stations are simply randomly allocated.
* *Stock Assessment Model:* A model which uses data from a survey and/or fishery data to get metric(s) of population status (usually an estimate of stock biomass)
* *Parameter:* An input to or output from a model, generally parameters are ‘fixed’ values (e.g catchability is a parameter).
* *Variable:* An input to or output from a model, generally variables can vary (e.g. biomass is a variable).
* *Prior:* Used to inform your model what the most likely range of values are for a particular parameter. This is a “Bayesian statistics” concept.
* *Catchability:* The proportion of the individuals in the area sampled that are captured by a survey tow
* *Natural Mortality:* The proportion of the population that dies from ‘natural causes’ in a given year.
* *Exploitation:* The proportion of the population that is captured by the fishery in a given year.
* *Simulation:* Process of using certain mathematical/statistical equations and tools to develop data and test the impact of varying inputs and assumptions on the end results. Stochastic simulations incorporate statistical uncertainty in model parameters to generate multiple outcomes by accounting for various types of uncertainty.
* *Realizations:* When running a stochastic simulation a realization is one set of results, there can be many realizations.