# Github Desktop

When any changes are made to any files in the github repository, these changes need to be uploaded via Github Desktop to also be reflected in Sharepoint and other areas. This includes making the survey plans, survey results, and updating the total data compendium. When making survey plans and results, the main path where data can be found or saved is:

C:\Users\herri\Documents\GitHub\HerringScience.github.io\HTML Markdown\Surveys\YEAR\SURVEY

* Once all the steps are complete for either a survey plan or survey results, load Github desktop (either on the taskbar or search for it). 
* The first page should immediately show any and all changes that occurred to the files within the github folders (such as me writing these instructions):

Graphical user interface, text, application

Description automatically generated

* These changes need to be ‘committed’ (e.g. saved). The only other step is that this commit needs to have a summary, describing the changes (e.g. “Survey Plans for SB5”, “GB3 Survey Results”, etc.). Afterwards, the ‘Commit to main’ button can be pressed.

Graphical user interface, text, application, chat or text message

Description automatically generated

* Finally, these saved changes can be uploaded by using the Push to Origin (e.g. upload to github online) button on the next screen. Graphical user interface, text, application

  Description automatically generated
* Any changes to .html files that are hosted by github pages (which are the links on Sharepoint) may take a few minutes to appear (e.g. the Total Data Compendium).

# Creating a Survey Plan

Step 1) Take screenshots of the tidal forecast for your ground (links are also in Chrome under the HSC folder):

Scots Bay – [Margaretsville station](https://tides.gc.ca/en/stations/315) | German Bank – [Yarmouth station](https://tides.gc.ca/en/stations/00365/)

The following screenshots should be taken:

1. A screenshot of the survey day (e.g. Sunday) and the proceeding day. **This needs to be saved exactly as “Daily.jpg”** and placed into the survey folder.

Table

Description automatically generated

1. A screenshot of the survey day and the proceeding day under hourly predictions. **This needs to be saved exactly as “Hourly.jpg”** and placed into the survey folder.

A picture containing text, white

Description automatically generated

Step 2) Open the Survey Lines spreadsheet in the main github.io folder.

Step 3) Form the “R Plan” and “Table” spreadsheets following the instructions in the Survey Lines document. You should end with a “survey plan.csv” and “Table.csv” with properly staggered/ordered vessel assignments saved in the survey’s github folder.

Step 4) Open the Survey Plans RMD Report in the Github folder 

Step 5) Change the options listed in the first code chunk to match the plan coordinated by the managers. Many of these options need to be formatted exactly as described.

Text

Description automatically generated

*surv* = should be “SB” for Scots Bay or “GB” for German Bank.

*surv2* = the full spelling of the ground, either “Scots Bay” or “German Bank” (can’t be Scot’s Bay or Scotts Bay, etc.)

*surv.date* = date and start time of the survey listed as YYYY-MM-DD and 24-hour clock HH:MM.

*surv.no* = the survey’s number (e.g. Scots Bay #6).

*Allocation* = the allocation of fish given to each vessel, in mt.

*SIAllocation* = relevant only to German Bank surveys, this is the allocation of fish given to each vessel for Seal Island only.

*vessels* = the number of participating vessels for the survey.

*EVessel* = relevant only to Scots Bay surveys, if a vessel was assigned the Eastern box, list their name here otherwise “NA”.

*NVessel* = relevant only to Scots Bay surveys, if a vessel was assigned the Northern box, list their name here otherwise “NA”.

*PlanktonVessel* = the vessel that the plankton tows and/or tech worked off of for the survey.

*Tagging* = a list of all vessels that had active taggers present during the survey.

Step 6) Once everything is in place, use the Knit button  to create the survey plan document. If you run into any errors, try to source the problem and troubleshoot it. Make sure you have all required files saved (Hourly and Daily tides, Survey Lines spreadsheets) with their proper format and names.

# Creating the Survey Results

After a survey, the following data will have been collected:

|  |  |  |
| --- | --- | --- |
| **Data** | **Data/Program(s) Needed** | **Where it goes** |
| Plankton Tow Data | Excel, Ruskin RBR | Source Data > planktonsamplingData.csv |
| CTD Cast | YSI Castaway CTD | Source Data > CTD\_Raw.csv  Source Data > planktonsamplingData.csv |
| Tagging Reports | Tagger Logs | Source Data > Tagging\_Raw.csv |
| SSB Estimates from Jenna  (Time delayed) | Echoview output | Main Data > SSB Estimates.csv |
| Greatest Backscatter Snip from Jenna | Save exactly as ‘Snip.jpg’ | Github/Survey/Year/Snip.jpg |
| **Final Step**: All data changes | Update RMD Data.R script | HerringScience.github.io > HTML Markdown > Update RMD Data.R |

**Plankton Tow Data**

Data collected on the Ruskin RBR Depth Probes needs to be extracted using the Ruskin Software. This can generally be done on the boat and saved ahead of time, as you need to plug the probe into the software to turn it on/off during the plankton tows anyways. Steps for doing this can be found in the appendix.

**Be sure to also take a screenshot (Win+Shift+S) of each tow’s depth profile and save it to the Github folder for your survey, as “Tow 1.jpg” and “Tow 2.jpg”, before leaving Ruskin RBR.**

Chart, line chart

Description automatically generated

Another photo that needs to be taken is of the plankton jar(s). These can then be edited together into one photo using paint or another program and **saved exactly as “Plankton.jpg”** in the survey folder.

A picture containing beverage, dirty

Description automatically generated

For the other general plankton data, fill in as much of the spreadsheet as possible. See the Appendix for a description of each column, if needed. The following columns need to be entered in a certain format:

*Ground* – “SB” or “GB”, not fully spelled such as “Scots Bay”.

*id* – same format as the plankton sampling jars, such as “GB2022-01”.

*Sample* – “Y” or “N” for yes or no, respectively.

*ExtraBox* – Answers are “No”, “Both”, “East”, or “North”, if any of the extra survey boxes were used during a Scots Bay survey.

*TideDirection* – should be “with” or “against” the tide, as written.

Other columns should simply be entered the same as previous data, such as DD/MM/YYYY for date.

**CTD Cast**

For the CTD data, export the cast data as a .csv file using the YSI Castaway CTD software (see Appendix for steps). Open this file and copy and paste the raw data (without the column headers) into the CTD\_Raw.csv file in Source Data. Be sure to add in the other columns in the datasheet manually as needed, these are: id, ground, plankton\_ID, Date, Lat, Lon, Year, Survey. Mostly just be sure to follow the same Date format as previous entries.

The general data from the CTD cast (average temperature, depth, lat/long, etc.) will also need to be added to planktonsamplingData.csv

**Tagging Reports**

Important: the columns ‘Julian’, ‘Year’, and ‘Tag\_Annual’ are added later by the R script and can be ignored. All other columns should be manually entered from each tagger’s log. Ground should be fully spelled out as “Scots Bay”, “German Bank”, or “Other”. “CTD” is the CTD id name from the .csv file name; this is if the tagger does their own CTD cast (e.g. Lisa used to have one aboard the Morning Star), and not the HSC tech’s cast from the Plankton Vessel.

**SSB Estimates**

After each survey, Jenna processes the data in Echoview and outputs a biomass estimate. This will take a few weeks after each survey to produce. Once this value is known, it can be placed into the Main Data > SSB Estimates.csv spreadsheet. \*subject to change as process may change if we can automate the process from Map+Region > Table A+B+C forward.

**Final Step: Update RMD Data.R and check for “Missing” Files**

**Update RMD Data.R**

Once all the above data is updated or needs to be changed (say after a survey some tagging logs come in late or the biomass estimate from Jenna is produced), the Update RMD Data.R script needs to be run.

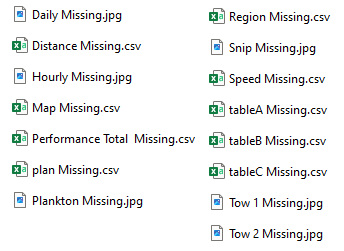
The first options in the code need to be changed to reflect the current survey, and then the script should be run in its entirety. These options need to be formatted exactly as “SB” or “GB”, and “Scots Bay” or “German Bank”.

A screenshot of a computer

Description automatically generated with medium confidence

**Checking “Missing” Files**

Every Year/Survey folder is already populated with a list of every file needed but tagged as “Missing”. Example:



After having uploaded all the data, taken and saved the correct screenshots, and run “Update RMD Data.R”, all of these missing files should have a cooresponding ‘real’ file. **Any of the “Missing” versions that have a real version should be deleted, leaving only ‘real’ files behind.** If anything *only* has the “Missing” version, it is still missing and needs to be addressed. This can also be checked within the Survey Results RMD script by running the first code chunk, as noted in the annotation:

Text

Description automatically generated

If anything labelled “missing” is in the output, that file still needs to be addressed. If the output is “character(0)”, everything has been updated correctly and “missing” files have been removed.

Finally, the Total RMD Report can now be run from the main github.io folder: 

In this markdown, just change the first code chunk option which asks for the current year, and run it (may take a few minutes).



Once all these changes have been made, the steps for Github Desktop can be followed and after a few minutes check the Sharepoint version of the Data Compendium to make sure the data has been uploaded correctly.

# Appendix

## Data Export Steps

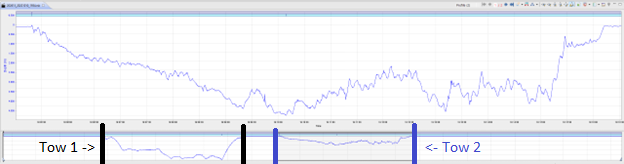
**Plankton Tow Data Export (Ruskin Depth Probe)**

Step 1) Load the Ruskin RBR software (can ‘search’ for “Ruskin” on windows). 

Step 2) In the middle taskbar on the far right there is “Toggle channel visibility” (the furthest option with a dropdown arrow); uncheck any channel that isn’t “depth”.

Graphical user interface, text, application

Description automatically generated

Step 3) At the bottom of the screen, use the left/right borders of the box to narrow in on each individual tow. For example, in the image below only the second tow is isolated (not both tows at once).

Step 4) For each tow, note the general start and end times. I personally open a notepad to write this down, as we’ll be adding a few other points soon. **Before proceeding, also take a screenshot of each tow using the Windows Snipping Tool (Win+Shift+S) and save this to the github folder for your survey. These need to be saved exactly as “Tow 1.jpg” and “Tow 2.jpg”.**

Path: C:\Users\herri\Documents\GitHub\HerringScience.github.io\HTML Markdown\Surveys\YEAR\SURVEY

Text

Description automatically generated

Step 5) Under the top left Navigator, right click the dataset you are currently using, choose Export, and select Microsoft Excel (\*.xlsx) format. This should be saved in the Github folder for your current Year and Survey. The exact path of this should be:

C:\Users\herri\Documents\GitHub\HerringScience.github.io\HTML Markdown\Surveys\YEAR\SURVEY

Graphical user interface, text, application, chat or text message

Description automatically generated

Step 6) Once extracted, open the file and navigate to the Data worksheet on the bottom left.



Step 7) Scroll down until you find the starting time of the first tow that you noted down. For example, in Step 4’s notepad I noted Tow 1 as starting at 19:00, so I would scroll to this point in the data:



Step 8) The far-right column is Depth which is what we are focusing on. Click on this first entry for Depth and drag your selection all the way down to the End time for Tow 1 (e.g., in this case 19:10). You should have all the Depth values for your time range selected (e.g., from 19:00 to 19:10).

Step 9) While selected, the bottom right of the Excel sheet will update Average and Maximum values. Record these values for both Tows 1 and 2 in the notepad.



Step 10) Both of these values should be entered in the planktonsamplingData.csv spreadsheet under the AvgTowDepth and MaxTowDepth columns.

**CTD Cast Data Export**

Step 1) Load the YSI Castaway CTD software (can ‘search’ for “CTD” on windows).

Step 2) Change from the default map page to “Show Files”

Graphical user interface, application

Description automatically generated

Step 3) Select the most recent cast (ordered chronologically with the newest on top), or find the cast you are interested in by date/time. Make sure that the cast has samples under “Number of samples” and that ‘Sample type’ is not invalid, which may indicate that you did a point sample cast instead of a regular cast (e.g. the bottom cast in this photo).

Table

Description automatically generated

Step 4) With the cast selected, choose Export Files from the top right taskbar. Make sure format is set to .csv (should be defaulted) and leave all other checkboxes selected.

Graphical user interface, application

Description automatically generated

Step 5) It is probably best to export directly into the folder associated with your survey. The export defaults to Documents, which is where the Github folder already is.

Graphical user interface

Description automatically generated with medium confidence

You will need to navigate a bit further to find the exact survey folder.

The exact path you should follow will be: Documents/GitHub/HTML Markdown/Surveys/ and then pick your Year and Survey (SB# or GB#).

# Data Dictionary

**Plankton Tow Data Columns**

Ground – “SB” or “GB” for Scots Bay or German Bank, respectively.

id – the ID tag given to the plankton tows, there should be two per survey barring weather and/or equipment issues. Format is SB2023-01 for tow 1, SB2023-02 for tow 2 from the same survey. If any tow has more than one jar, it still falls under that tow label (e.g. if tow SB2023-02 has two jars they are both under the -02 tow).

Survey.No – the survey number for this ground for this year.

Date – Date of the survey in DD/MM/YYYY format.

StartTime – the time that the survey was scheduled to start in the survey plans.

Sample – whether fishing was completed (“Y”) after the survey or not (“N”). Filling this in will be time delayed as it may take a few days after the survey for fishing to occur, but if you know that fishing happened for certain it can be added immediately.

Vessel.No – number of vessels attending the survey.

ExtraBox – for Scots Bay surveys only, whether the North or East extra boxes were assigned to any vessels for the survey. Answers are “No”, “Both”, “East”, or “North”.

EVessel – if a vessel was assigned to the East box, list the full name of the vessel here.

NVessel – if a vessel was assigned to the North box, list the full name of the vessel here.

PlanktonVessel – the vessel that the plankton tows were conducted on.

No\_jars – number of jars associated with each individual tow.

Lon1, Lat1 – the lat/lon coordinates for the start of the tow.

Lon2, Lat2 – the lat/lon coordinates for the end of the tow.

Time1 – time when the tow was started.

Time 2 – time when the tow was ended.

TowTime – total duration of the tow (or difference between the tow end and start times).

Gear – diameter of the netting used, should be defaulted to “1/500” for the 1/500um netting.

Net – diameter of the metal ring used, should be defaulted to “1” for the 1m ring.

SurfaceTemp – no longer used, previously taken from the wheelhouse sensors.

AirTemp – outside air temperature during the tow, can be taken from any weather forecast for the area.

WaterDepth1/2 – no longer used, previously taken from the wheelhouse sensors.

TowType – type of tow conducted, should be defaulted to “Surface Tow”.

Speed – speed of the vessel during the tows, in knots. Generally recorded by the captain.

Heading – heading of the vessel during the tows, in degrees. Generally recorded by the captain.

TideDirection – whether the vessel was “with” or “against” the tide during the plankton tows. Normally one tow with be with, and one will be against the tide.

AvgTowDepth – average depth of the tow. This is measured during the Ruskin export process (see Plankton Tow Data Export).

MaxTowDepth – maximum depth of the tow. This is measured during the Ruskin export process (see Plankton Tow Data Export).

CTD\_ID – file name of the CTD which can be found in the exported CTD .csv. It should be a combination of the device serial number, then the full date, then time stamp (e.g. device\_YYYYMMDD\_HHMMSS).

AvgTemp – average temperature from the CTD cast, this should be given on the CTD summary for the cast on the device itself.

AvgSalinity – average salinity from the CTD cast, this should be given on the CTD summary for the cast on the device itself.

WindDirection – a rough estimate of the direction of wind during the plankton tows. It can also be taken from any weather app.

WindSpeed – a rough estimate of the speed of wind during the plankton tows. It can also be taken from any weather app.

Swell – an estimate of the height of any swell on the ocean during the plankton tows.

Flowmeter – the brand of flowmeter that was used during the plankton tows.

FlowReading1 – the current reading on the flowmeter before conducting the plankton tow.

FlowReading2 – the reading on the flowmeter after conducting the plankton tow.

NoRevs – the difference between the two flow readings, as in the number of revolutions the flowmeter underwent during the plankton tow.

DistanceCalc – the result of a calculation ((final-initial \* 26873)/1000) to determine many meters the plankton tow covered.

Volume – the result of a calculation (DistanceCalc\*3.14159\*1m net diameter; this will change if the 1m net diameter ever changes) to determine the volume of water that the plankton tow filtered.

Observers – the tech who conducted the tow(s).

DepthDiscD – the depth at which the secchi disc disappeared.

DepthDiscA – the depth at which the secchu disc reappeared.

**CTD Cast Data Columns**

id – ID of the CTD cast, can be found in the file name of the exported CTD .csv

Pressure, Depth, Temperature, Conductivity, Specific\_conductance, Salinity, Sound\_velocity, Density – all of these columns are added from the raw CTD data (pasted in) and not added manually.

Ground – ground that the CTD cast was taken on, generally “Scots Bay” or “German Bank” spelled out fully.

Plankton\_ID – id of the associated plankton tow(s).

Date – Date of the CTD cast.

Lat/Lon – Location of the CTD cast, generally given directly by the device assuming GPS satellites are in range. If GPS can not be found in a reasonable time frame, Lat/Lon from the wheelhouse can also manually be added.

Year – Year of the CTD cast.

Survey – survey number for the current year that the CTD cast occurred.