# 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\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 title (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) **Scots Bay ONLY:** Take screenshots of the tidal forecast for your ground:

[Margaretsville station](https://tides.gc.ca/en/stations/315)

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 Plans RMD Report in the Github folder 

Step 3) 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.

**Example German Bank:**

A computer screen shot of a program

Description automatically generated

*Example notes*: EVessel and NVessel are left “NA” as this only pertains to Scots Bay. As “vessels = 4”, there are four vessels named below under V1-V4. The PlanktonVessel (often Lady Janice II) should be the furthest east vessel (and last listed in V#) for German Bank, to line up with the CTD/tow box better.

**Example Scots Bay (with both Boxes):**

A computer screen shot of a program

Description automatically generated

*Example notes*: In this case, ‘vessels’ is increased to 6 but only 4 are listed in the V1-V4 section as this is for the main box only. The other two vessels are listed as the EVessel (East Box) or NVessel (North Box). For Scots Bay, the main box V# vessels + EVessel + NVessel should equal the total ‘vessels’. SIAllocation is ignored as this pertains only to Seal Island during German Bank surveys. Finally, the PlanktonVessel is often assigned V#2-3 as this will line up with the CTD and tow box.

Step 4) 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 for Scots Bay) with their proper format and names.

# Survey Plan Coding Dictionary

*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.

*Tagging* = a string of vessel names for boats that are participating that have known taggers aboard.

*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.

*V#*: A list of all main box vessels. These should be randomized between surveys, although PlanktonVessel is often either lines 2-3 for Scots Bay or the last vessel (furthest east) for German Bank.

# 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 | Fill out PlanktonData.csv in Survey Folder |
| CTD Cast | YSI Castaway CTD | Export raw data .csv to Survey Folder |
| Depth Probe | Ruskin RBR | Export raw data .xlsx to Survey Folder as “Ruskin.xlsx” |
| Tagging Reports | Tagger Logs | HerringScience.github.io > HTML Markdown > Tagging Data.R |
| Map and Region .csv’s from Jenna’s Echoview | Echoview output | Github/Survey/Year as Map.csv and Region.csv |
| Survey Box Polygons from Jenna | polygon\_SBEastern, polygon\_SBNorthern, polygon\_SB, and any GB changes | Github/Survey/Year as:  polygon\_SBEastern.csv polygon\_SBNorthern.csv  polygon\_SB.csv |
| Greatest Backscatter Snip from Jenna | Save exactly as ‘Snip.jpg’ | Github/Survey/Year/ |
| **Final Step**: All data changes | Update Data.R script | HerringScience.github.io > HTML Markdown > Update Data.R |

**Full Survey Results Checklist**

*All the below files are made at some point in the results process, but this checklist can help if you find one is missing.* **All files must be named exactly as below, and the same .jpg or .csv format!**

|  |  |
| --- | --- |
| **Daily.jpg** | .jpg snip of the daily tides forecast for the survey day and the next day |
| **Hourly.jpg** | .jpg snip of the hourly tides forecast for the survey day and the next day |
| **Distance.csv** | Analysis of vessel transect distances made after Update RMD Data, if tableA and plan.csv exist |
| **Map.csv** | Output from Jenna’s Echoview work, should be manually saved into the survey folder |
| **Performance Total.csv** | Analysis of vessel performance made after Update Data, if tableA and plan.csv exist |
| **survey plan.csv** |  |
| **Plankton.jpg** | Photograph of any jars of plankton from the tow(s), needs to be saved as .jpg |
| **Region.csv** | Output from Jenna’s Echoview work, should be manually saved into the survey folder |
| **Snip.jpg** | .jpg snip of the single-point with the greatest backscatter in Echoview, should be taken by Jenna and saved in the survey/year folder |
| **Speed.csv** | Analysis of vessel speed made after Update Data, if tableA and plan.csv exist |
| **adhoc.csv** | Should be saved directly to the survey/year folder if any vessels conducted an adhoc school survey after the scheduled survey |
| **tableA.csv, tableB.csv, tableC.csv** | Tables created in the Update Data script, if Map and Region exist. Also requires any “adhoc.csv” files to be present, if there was an adhoc survey conducted |
| **Tow 1.jpg, Tow 2.jpg** | Created in Update Data.r through ggplot and the raw Ruskin data |
| **polygon\_X.csv** | .csv files of any polygon changes sent by Jenna. If they didn’t change, need to copy the same set from the previous survey’s folder, or the default boxes are in the “Box Coordinates” folder in the root github directory. |
| **CTD Data: e.g., 10K100766\_20230730\_224751.csv** | Raw .csv export of the CTD data via the CTD software. No name change necessary, but the CTD\_ID (file name) needs to be recorded in the PlanktonData.csv sheet |
| **Ruskin.xlsx** | Raw .xlsx export of the Ruskin RBR Depth Probe data via their software. |

# Plankton Tow Data

Each Github survey folder (Surveys/Year/Survey, e.g. Surveys/2023/SB5/) will have a blank “PlanktonData.csv” sheet that needs to be filled out as entirely as possible. There will be **one row per plankton tow completed**.



# Plankton Data Columns

*Set\_Number:* the set number written on the plankton jar, which is cumulative throughout the year (e.g. SB2023-01, with SB2023-02 being written in the second row).

*No\_jars*: how many jars were needed to used to capture plankton from the tow, previously more than 1 jar was used but now we only sub-sample the plankton and tend towards using only 1 jar per tow.

*Tow\_No*: the tow number, written in numerical format, e.g. 1 or 2.

*Lon1, Lat1, Lon2, Lat2*: coordinates of the tow written by the plankton vessel captain. Need to be entered in the following format: 44 60.28 or 65 20.35. No negative is used on the longitude, and it is required to have one space and then a decimal for the formatting to work properly.

*Time 1, Time2*: The start (time 1) and end (time 2) of the plankton tow, generally 10:00min apart. Written as MM:SS.

*TideDirection*: ‘with’ or ‘against’ written all lower case. Whether the plankton tow was conducted with or against the tide.

*Speed*: the speed of the plankton vessel during the tow in knots, written on the captain’s sheet.

*Heading*: the heading of the plankton vessel during the tow in degrees, written on the captain's sheet.

*Observers*: name of the technician who conducted the plankton tows.

*WindDirection*: direction of wind during the tow, generally gathered from any Weather Network website, using Margaretsville (SB) or Yarmouth (GB) stations.

*WindSpeed*: speed of wind during the tow, gathered as above.

*Swell*: an estimation of the swell height during the tows, in 0.25m increments.

*AirTemp*: air temperature during the tow, gathered the same as WindDirection and WindSpeed.

*FlowReading1 and 2*: Flow readings recorded off the General Oceanics flowmeter at the start and end of the tows. Not that FlowReading2 for tow 1 and FlowReading1 for tow 2 should be nearly identical.

*DiscDepthD and DiscDepthA*: for the secchi diskl, the depth in meters at which it ‘D’isappeared and ‘A’ppeared. Can be recorded in 0.5m increments as needed.

*CTD\_ID*: the ID of the raw CTD data file that is exported after conducting a cast (e.g. ‘10K100766\_20230730\_224751’). Can be copy+pasted from the file name directly while renaming it (but don’t rename it). 

*CTD\_Lat, CTD\_Lon*: the Lat and Lon recorded on the CTD device, recorded in the format the device presents (e.g. 45.000, -65.000).

# Ruskin Depth Probe 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.

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

Step 2) 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. **It needs to be named as “Ruskin.xlsx”.** The exact path of this should be:

C:\Users\herri\Documents\GitHub\HerringScience.github.io\Surveys\YEAR\SURVEY\Ruskin.xlsx

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

Description automatically generated

**Plankton Tow Jars**

A photo of the plankton tow jars needs to be taken and saved as **“Plankton.jpg” in the survey folder.** It is best to take one photo of both jars, but if more detail is needed of the captured plankton then two separate photos can be taken and stitched together later (in Paint or other programs).

A picture containing beverage, dirty

Description automatically generated

# CTD Cast Data

For the CTD data, export the cast data as a .csv file using the YSI Castaway CTD software:

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/Surveys/ and then pick your Year and Survey (SB# or GB#).

# Tagging Reports

Data used to be manually added to TaggingEvents.csv but now can be added by using the Tagging Data.R script. Only add information to the first fields as marked and run the script, no other data needs to be added or modified normally. Lat/Lon is added as degree-min-sec (e.g. 44°16’23) but **written with only numbers and a space between (e.g. “44 16 23”)**; the script will convert it to decimal degrees.

A picture containing text, screenshot, font, electric blue

Description automatically generated

If Vessel names or associated Taggers change, the below portion can be modified:

A screenshot of a computer

Description automatically generated with low confidence

# Map and Region Data Files

After each survey, Jenna processes the acoustic data in Echoview and outputs a “Map.csv” and “Region.csv” file. These need to go into the Github/Survey/Year file specific to the survey. Once these are in place, running the Update RMD Data script (details below) will process this data into tableA+B+C.csv files. Furthermore, **if there was an adhoc school survey conducted** the .csv for it needs to be saved as “adhoc.csv” in the same github/survey/year folder.

**Final Step: Update Data.R**

# Update Data.R

Once all the above data is updated 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. ‘ids’ need to be set to the boat initials for only Scots Bay and only Main Box vessels. Area values need to be updated to match Jenna’s values in tableC.csv., and then the script should be run in its entirety.

**If for whatever reason you need to re-run this (**e.g., you receive some tagger logs late and they need to be added or are troubleshooting errors**) there are three locations you need to manually go in and delete the entry for the current survey, that would have been entered from each time you’ve run the script.**

1. **Main Data/SSB Estimates.csv** and delete the entry for the current survey and any NA blanks.
2. **Source Data/CTD\_Raw.csv** and delete any overlapping CTD entries from the same survey. You can also select the entire dataframe and press Data > Remove Duplicates
3. **Main Data/CTD Full.csv** and delete any overlapping CTD entries from the same survey. You can also select the entire dataframe and press Data > Remove Duplicates

A screenshot of a computer

Description automatically generated

A computer screen shot of a blue screen with text

Description automatically generated

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

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.

# 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.

Fishing – whether the survey was a fishing (“Y”) or non-fishing (“N”; typically German Bank is non-fishing).

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.