**How OGMAP is Currently Used for Shrimp Assessments:**

Currently Ogmap is utilized to generate survey based indices for the Shrimp Assessments. These indices include biomass and abundance based on different criteria of the Shrimp (total, female, fishable, recruitment sized, etc.) and also abundance (or biomass) at length (i.e. Logmap).

Two space delimited data sets (one for numbers/weights and the other for length frequencies) are generated for running through Ogmap. If there is a data from different seasons, then a separate data set is generated for each season. Also a separate data set is created for differing species (i.e. Northern Shrimp and Striped Shrimp). You could combine the data such that there are columns, with meaningful header names, for both species in the same data set. These data are generated/formatted separately from the Ogmap program (I use SAS).

The numbers and weights file includes columns year, lat, long, depth, NAFO, sfa, stratum, fisharea, temp, light, survey, n\_preds, w\_preds, n\_cape, w\_cape, n\_totals, w\_totals, n\_fishable, w\_fishable, n\_females, w\_females, n\_males, w\_males, n\_recruit, w\_recruit, n\_trans, w\_trans, n\_nonovig, w\_nonovig, n\_ovig, w\_ovig, n\_primi, w\_primi, n\_multi, and w\_multi. Column headers should make sense to the user of the data as they will likely differ for each species! At a bare minimum, you will need year, lat, long, depth, area and weight/number where area could be a NAFO division, a stratum, a management area, etc. You can include multiple area columns as well.

I had edited my version of bottom\_survey.py and it required that all variable headers began with either n or w. The reason for doing so was to call upon the same bandwidth file for either weights or numbers in the same area. Geoff’s most recent version won’t care so much about the header names given that it grabs the same bandwidth file for all analysis on any given species and area.

The length frequency input file has the same beginning columns as the file of weights and numbers; however it is followed by columns containing both numbers and weights at length. I’m uncertain if this data set can be combined with the weights and numbers files, but likely so. Some examples of the header names for shrimp are n\_total1…n\_total80, w\_male1…w\_male80, n\_female1…n\_female80 where the number (i.e. 1 to 80) are the carapace length multiplied by 2. The length groupings will differ for all species of course.

There is an arena file (LNshelf.arena) which contains many previously surveyed points of known lat, long and depth. This file, for shrimp, has headers lat, long, rootdepth, stratum, depth, NAFO, SFA, fisharea and area. The lat, long, rootdepth, depth and area columns should not be modified and are a necessity in any arena file. The stratum, NAFO, SFA, fisharea columns should be modified to match the headers within your input data set. So if your input data set has headers NAFO, SFA, CMA, stratum then the arena file should have the same headers within it. For species such as cucumber and scallop, where strata are not groundfish strata, the strata numbers in the arena file must be edited to reflect those strata and match the strata numbers within the input data set.

Now to review the steps involved in running Ogmap. Directory structure is an important factor in using Ogmap. You will want to have a main directory containing, at a minimum, the input data, the arena file, ogmap.py, bottom\_survey.py and gte\_utils.py. The steps required to run Ogmap will differ depending on if Ogmap was previously (successfully) executed on the particular area and species. For completeness, the steps listed will assume that it is a brand new analysis.

1. Upon opening python, navigate to your Ogmap folder. Run gte\_utils.py, Ogmap.py and bottom\_survey.py in that order. The command syntax is *run –i bottom\_survey.py*.
2. Create a subdirectory in which you will run Ogmap. This will use the general arena file from the parent directory to create a local arena file within the subdirectory. Hence that arena file should match strat, area (i.e. SFA or CMA), etc. The syntax of the command is *make\_arena\_directory(‘arena directory and name’, ‘subdirectory name’, NAFO=’div’, strata=[stratnum])*. This can specify numerous areas, such as NAFO, SFA and stratum. An example from Shrimp, used to run for *Pandalus borealis* in SFA 5 is *make\_arena\_directory('C:\Shrimp\Ogmap\LNshelf.arena', 'SFA5 PB', SFA=5)* where SFA5 PB will be the name of the new directory. This step is to be omitted if the analysis was successfully run on this area and species before. If you attempt to recreate an existing directory, a message will appear that reads “arena exists.”
3. na*('C:\Shrimp\Ogmap\PB\_fall\_new.dat',24)* if the number of number/weight columns is 24. This used to create a file called *tiebreak* in the parent directory however Geoff has made edits to modify this. NOTE: In the previous version a “magic number” was hidden within the bottom survey file and represented the swept area of the survey gear (of course given ideal trawl geometry!). Geoff has modified the code to make this the default, but does give the user to input a swept area for whatever gear they are using (i.e. it would differ for dredges).
4. Navigate to the directory you created in step 3. Syntax: *cd SFA5 PB*. The only file within this directory should be *local.arena.*
5. Now we will run Ogmap for the first time. Type in *use\_ogmap()*. Prompts will appear on the screen to guide you through the execution, the user must enter the info. Each step below displays the prompt in bold, with examples of what I could type following.
   1. **Name of variable to analyse:** *w\_totals* (one of the header names in the original input data)
   2. **Years (‘last’, ‘all’ or array):**  *‘all’* (an array is necessary if there are missing data years in the input data, or if you wish to only run on the last 5 years, but this example uses all years in the data set).
   3. **Bandwidth options:**
      1. **Specify the bandwidths**
      2. **Take the bandwidths from the .median\_bw file**
      3. **Create the .median\_bw file and then use the result. If a new analysis, run w\_totals first.**

I added in the statement at the end of #3 above, and 3 is selected here since we are assuming that it is a brand new analysis. A bandwidth file (i.e. .median\_bw) file is required to run Ogmap. A bandwidth file is required for the analysis. On the first run of a particular area and species, this file will need to be created. Elsewhere in this document I mentioned that manually copy the bandwidth file from the run on w\_totals to the other subdirectories in an area (i.e. in each area I will use the same for SSB, males, totals, logmap, etc). This ensures that the same bandwidths are used, and are comparable, for that area and species. After the first run, I select 2 so that the same bandwidths are used annually. I rarely use 1, but you can manually enter the bandwidths if you wish.

* 1. **Ref probs:** *0.025,0.975* – These are the confidence intervals, so this example is 95% CIs.
  2. **Ref values:** *27* – This could be something like the LRP and will provide a probability that the estimate is below that value. Keep in mind that the estimate is in thousands, so if your LRP is 27000 then you enter 27. You can leave this blank and enter through this command.
  3. **Sanity check: (y/[n]):** I’m not entirely certain what this does, although I should know.

1. The initial run will always be the longest one. Once it is ran once, then every year you can return to the same directory and run on just the last year, or other years if there had been data edits. I like to pull the data off the archive, re-generate the data, and run on all years for each assessment. For me, it keeps all the estimates in one place and makes it easier for plotting.
2. The steps for running logmap are quite similar. You will need to visit the parent directory and pre\_process the data for logmap (example *preprocess\_data('C:\Shrimp\Ogmap\PBlfd\_fall\_new.dat',480)* then re-enter the same directory. The command to run logmap is simply *logmap(‘all’, 480)*. You could replace ‘all’ with ‘n\_totals’ or whatever variable you wish.

**Things I Know OGMAP Can Do but Not Sure How:**

I know that there are ways to look at maps of probability distributions and influence while running Ogmap on a particular area, but I don’t fully understand how to do so.

I know there is a way to generate gini coefficients, so far I have not seen this work.

**Things I Would Like to Do (*Easily*) Using OGMAP:**

For Shrimp, I manually copy the bandwidth file into each shrimp “type” (females, males, total, fishable, etc). Ideally this file might live in the directory for the area being analysed, so it might be in the directory SFA6PB rather than in each subdirectory. DONE; thanks Geoff!

I would like to have generated the “republic” plots using Ogmap rather than strap. At present, when strap is used to generate these plots, there are instances where the biomass index might go up using strap, but down using Ogmap or vice versa. While this makes sense to us scientists, it causes confusion for many meeting participants who would prefer to see an index higher than lower regardless of our explanation. Essentially I would like to have the option to select to run Ogmap for all strata and years in an SFA. Currently I believe I would have to create separate directories, with separate local arenas and run Ogmap on each stratum separately. This can get time consuming and leave room for omissions/human error. IN PROGRESS; thanks Geoff!

This can be a work in progress that I will get to someday, but it would be ideal to have the creation of graphs automated. This is an ideal project for somebody learning to graph in python, or can be done in R or whatever program. ☺

This is likely related to my previous point given the way that I have been plotting the abundance at length indices and addresses the logmap output. The logmap output is a very long file. IF plotting in excel (which will undoubtedly be done by some future users of the data), inputting the data into excel can be time consuming particularly for the very first run of logmap that covers all years. I’m uncertain if possible, but an output file utilizing columns rather than rows would be great. For example, headers would be carapacelength, numtot1995, numtot1996, numtot1997, … This might not be feasible, and certainly not necessary if Python (or R) is utilized to plot the data without the tedious methods of plotting in Excel. THIS ONLY APPLIES IF PLOTTING IN EXCEL, WHICH WON’T BE DONE BY MOST USERS OF OGMAP.

**Things I Have Modified from the Original OGMAP program:**

I prefer to have all the output formatted such that there is one line for each year of survey data. I can’t accurately recall if these are things I changed from the original, but I know some things in the output data were! The log file is named, and has information within it, to avoid mixing it up with another variable Ogmap output and to provide clarity to a wider range of potential users. I cannot recall if the original bottom\_survey.py file named the log file in a meaningful way.

For example, I have a log file called w\_totals SFA6 AllStrata.log and the top few lines appear as follows:

w\_totals c:\Shrimp\Ogmap\SFA6 AllStrata

Biomass Indices x 1000 tonnes

Using bandwidths [ 3.87, 43, 2.33 ]

Ref probs (i.e. Confidence Limits) [ 0.025 0.975 ]

Year Estimate LowerCI UpperCI

1995 293.824 220 345

1996 498.738 419 622

I modified the code so that w\_totals.median.bw is used for both abundance and biomass runs and that file is used for each variable for that particular species.