Antarctic ice shelf vulnerability to surface-melt-induced collapse

Karen Alley, Brent Butler,
and Melissa Cushing

GIS Programming for Spatial Analysis
Final project
Spring 2014

Project background and outline:

The stability of Antarctic ice shelves, which hold back the great volume of ice contained in the Antarctic ice sheet, is a key component in understanding future sea level rise. Several ice shelf collapses on the Antarctic Peninsula in recent decades have been caused by surface meltwater forcing open crevasses and causing rapid disintegration of the shelf (Scambos et al. 2003). Surface meltwater in the summer percolates into the firn layer, where it refreezes as ice lenses and layers. When the firn becomes saturated with ice, melt ponds can be supported on the surface, where they can force open crevasses. The degree of firn saturation can be measured using winter active radar backscatter data that, when plotted against the average number of annual melt days, can reveal an ice shelf's vulnerability to surface-melt-induced collapse. The physical reasons that backscatter relates to firn saturation were described by Fahnestock et al. (1993), and the relationship on ice shelves was demonstrated by Scambos et al. (2003). Here, we confirm the Scambos et al. (2003) relationship and extend their study to include many more ice shelves using more recent, higher resolution data.

In this project, we use microwave backscatter data from Brigham Young University's Microwave Earth Remote Sensing Research Group from the Quikscat sensor (Long 2010) along with SSM/I-derived melt days data from Picard and Fily (2006). We subset these data using shape files of ice shelves derived from MODIS Mosaic of Antarctica (MOA) data products (Haran et al. 2005). Together, these data sets allow us to assess the vulnerability of ice shelves all around Antarctica to surface-melt-induced collapse.

Methodological approach:

Workflow: Our project relies on two main types of data (backscatter and melt days) that are combined in the result, and a third type (ice shelf shape files) that is used to modify the main data. Our workflow progressed as follows: 1. Convert the main data to a usable form (geotiffs);

2. Average the backscatter and melt days separately over given time periods; 3. Resample the averaged files to the same resolution; 4. Subset the resampled and averaged files using the ice shelf masks; 5. Plot the data pixel-by-pixel and by ice shelf averages; 6. Create a vulnerability index based on the backscatter/melt days relationship; 7. Map the vulnerability index across Antarctic ice shelves. Each of these steps is described below.

Data conversion: The main data came in two different formats: the backscatter data were obtained as georeferenced .sir files, while the melt days data were originally in NetCDF format, initially without a georeference but with georeferencing information provided. The conversion of these data types turned out to be the most challenging part of our project.

BYU uses a "scatterometer image reconstruction (SIR) resolution enhancement algorithm" (Long 2010) to take backscatter imagery in a difficult-to-use format and put it on a regular grid at 2.225 km resolution. However, the output format of the grid is a .sir file, which is not a standard type that many programs can work with. BYU also provided a Windows executable file that can convert one .sir file at a time into a geotiff. As imagery is available daily for Antarctica for about twenty years, this would ultimately mean converting ~2,000 files, one at a time, based on seven commands prompted by the executable. Clearly, we wanted to automate this process. We achieved this using the method "call" from the Subprocess module in Python along with a piping procedure. The call method required the path to the executable file and the path to the file to be converted as inputs. Then we used a pipe symbol ("<"), and provided a path to a text file containing all the commands required by the program. This told the method to "pipe" the contents of the text file into the command structure of the executable program. Finally, we selected "shell = True" to tell Python to make use of the shell within which the executable file ran. We were then able to use the OS module and standard Python looping routines to automatically feed the .sir files into the executable program and convert them into geoTIFFS.

NetCDF (Network Common Data Form) is a file format developed specifically for array-oriented data. The advantages offered by the netCDF format include its ability to be self-describing and scalable. The melt days data netCDF included annual cumulative daily melt data from 35 years. All 35 images were tied up into one netCDF file, making it easy to share but a challenge for us to work with. Since netCDF is becoming so commonly used, Esri developed a tool for converting the format to raster. When trying to use this tool in a scripting environment we were unable to use the output file format: "result". To work around this challenge, we opted to complete this one-time conversion in the ArcMap environment. After converting the netCDF file, we exported the output as a geoTIFF. This geoTIFF contained 35 different bands. Instead of converting each individual band into a separate geoTIFF file, we simply built functionality into our script to create a NumPy array out of each band. Since the goal was to average all 35 years of melt day data into one geoTIFF this was an appropriate technique.

After averaging, it was necessary to assign the grid a spatial reference. According to the metadata (Picard and Fily 2006), the grid was developed using the NSIDC South Polar Stereographic projection on the Hughes 1980 ellipsoid. The metadata provided us with the upper left-hand corner of the file corresponding to the NSIDC grid (though it turned out the metadata assumed a different origin than the NSIDC grid) and the cell size. The data were appearing upside-down, so we used NumPy to flip the grid vertically. Then we used ArcPy, inputting the EPSG code, lower left corner, and cell size, to assign a spatial reference. This appeared to work, except for two problems: ArcMap is unable to convert anything off of the Hughes ellipsoid to make it compatible with other data, and the data did not line up with our other datasets when projected. This second problem indicated mistakes in the metadata. Fortunately, the author had also provided two additional grids: one of latitude values and one of longitude. We ultimately used a manual procedure in ArcGIS to assign the latitude and longitude values to a grid of points, which were then converted to a raster. But since ArcMap cannot work with the Hughes 1980 ellipsoid, the data still do not quite line up with our other datasets. However, we decided it was close enough to carry out this coding project, and intend to rectify the problem in the future.

Averaging: Because the geoTIFFS containing the backscatter data are collected on a daily basis and the winters in the Antarctic are highly variable, it was necessary to average these data to reach sound correlations. To complete this, we took inputs of geoTIFF files and created loops to convert each geoTIFF file to a NumPy array. We then added each NumPy array together, and took the average of all of the arrays to make a raster that more clearly represented the data we were observing. We made this averaging script into a function so it could be called easily in any of our scripts. We also created a similar function for averaging the melt days data.

Resampling: The netCDF files of melt days were originally at a resolution of 25 km while BYU's .sir format had a resolution of 2.225 km. In order to work with the two datasets, after converting all files to geoTIFFS, it was necessary to resample the netCDF files to the size of the .sir files. We did this using Arcpy's resample function. However, even though the resample function itself ran smoothly in ArcGIS, it continually broke when implemented in Python. To combat that problem, we figured out that by using the describe function, we could determine the mean cell height and widths, and then feed those directly into the resample function instead of trying to make the resample function find those values from a given raster.

In the end, however, this code became obsolete when we learned we could resample and ensure the data was correctly lined up using the environment settings (env.cellSize and env.snapRaster) instead of using separate functions to complete the same process.

Subsetting: Both the melt days data and the backscatter data included the entire Antarctic continent. Since our analysis is concerned only with the ice shelves, we had to subset the data into individual shelf regions. This also allowed us to expedite the development of our script since we were able to test our script on smaller samples of the data.

To complete the subsetting process we utilized the ArcPy tool *Extract by mask*. Before settling on this method, we explored multiple other options that, it turned out, would not clip the raster data in a way that we could then use in NumPy arrays. This tool was perfect for our needs, since the pixels outside the shapefile were assigned a no data value. This was very important since we ultimately needed to plot the pixel values of both the backscatter and the melt day data within the ice shelf polygons, while avoiding the pixel values outside of the polygons.

Plotting: Once the backscatter and melt days data were clipped to the individual polygons, it was time to establish a technique for plotting. The intention was to have each pixel of melt days data plotted against the backscatter pixel of the same location. To do this, the subsetted raster data sets were put into NumPy arrays, the no data values were ignored, and then all of the ice shelves' NumPy arrays were simply fed into a single plot command via a for loop. After close examination of the resulting plot (fig. 1), it was clear there was too much going on for the simple purpose of visualization. To simplify the plot slightly, we opted to make an additional plot of the ice shelves' mean melt days plotted against their mean backscatter (fig. 2).

Creating and mapping a vulnerability index: Because the melt days are not yet correctly lined up, and because the results were not quite what we had expected (see discussion section), it was not yet possible or worthwhile to develop a rigorous vulnerability index. So, for the sake of simplicity and utility, we developed a rudimentary code for reclassifying the melt days raster based on the log relationship found in our plots (fig. 1). Ultimately, the vulnerability index will depend more evenly on both the melt days and the backscatter datasets, probably using some form of principal component analysis and raster math. For the moment, though, we mapped this prototype vulnerability index by simply using ArcPy to reclassify the melt days raster and choosing an appropriate color scale for the raster in ArcMap (fig. 2).

Results:

The most important result we set out to produce in this project is a pixel-by-pixel scatter plot of temporally averaged winter backscatter vs. temporally averaged annual melt days (fig. 1). Our hypothesis was that our plot would support the relationship demonstrated by Scambos et al. (2003) (appendix A fig. 1): a roughly log-linear relationship of increasing backscatter with increasing melt days, followed by an abrupt drop in backscatter indicating firn saturation and a high likelihood of ice shelf collapse. Our plot appears to support the relationship of increasing backscatter with increasing melt days, but does not show the abrupt drop-off that we expected to see.

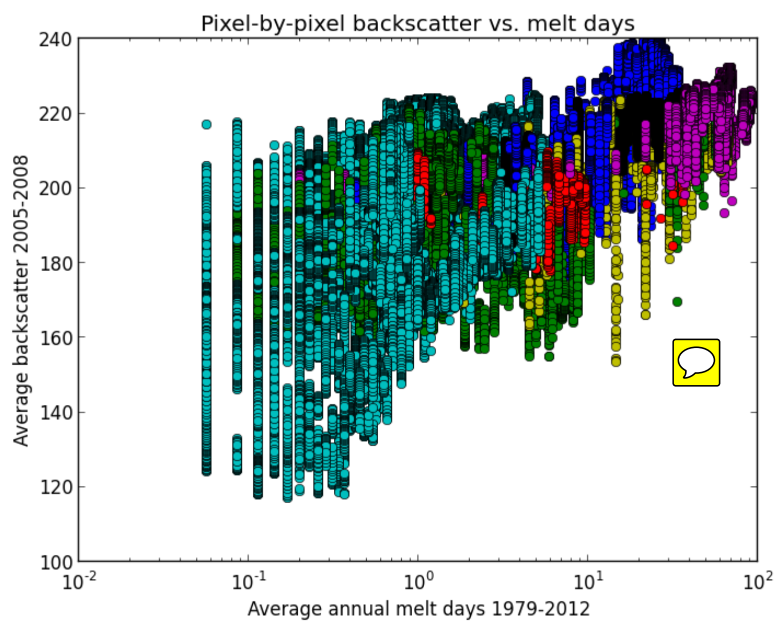


Figure 1: Pixel-by-pixel plot of melt days vs. backscatter from 19 Antarctic ice shelves. Notice the positive relationship, visible despite the large spread in data.

To further explore our data, we decided to try plotting average values for each shelf instead of plotting pixel-by-pixel (fig. 2). The result once again displayed a roughly log-linear relationship of increasing backscatter with increasing melt days, but again did not show the abrupt drop-off we expected. However, this plot helped us gain more insight into the Scambos et al. (2003) plot because it clearly identified which points belonged to which ice shelves. We could tell that the ice shelves that were highest on our relationship were the ones Scambos et al. had identified as being along the drop-off zone. This helped us realize that Scambos et al. had not used data from all of the ice shelves we had, and if we took away the points from some of the extra shelves we had used we could see the drop-off pattern start to emerge (appendix A fig. 2).

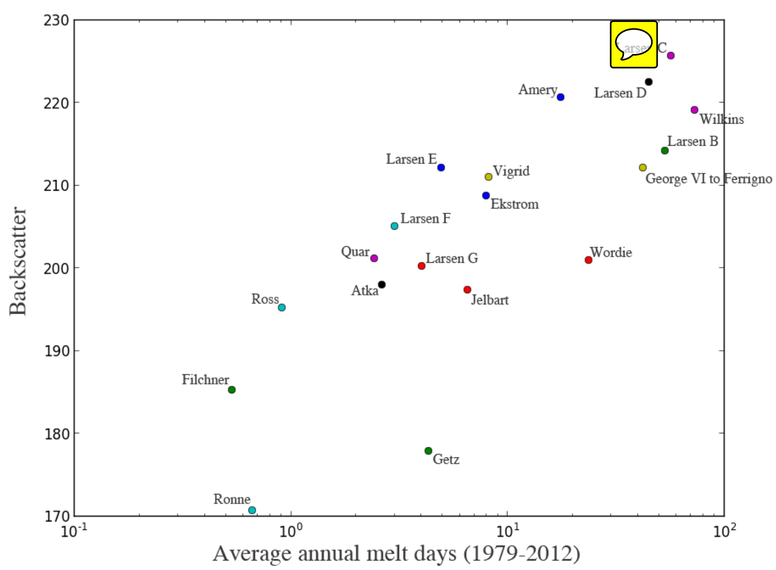


Figure 2: Shelf-averaged melt days vs. backscatter from 19 Antarctic ice shelves

Ultimately, we wanted to create a vulnerability index based on the relationship revealed in our plots, relying on both the melt days and the backscatter values. However, as discussed below, we would like to try other sampling techniques and corrections before accepting our relationship as final. The results of our rudimentary vulnerability index are mapped in appendix A fig. 3, but this index was created more for the sake of completing the coding of our workflow than for any scientific results it reveals.

Discussion:

The purpose of this project is to extend the Scambos et al. (2003) study to map the vulnerability of ice shelves to surface-melt-induced collapse on a wider scale and with higher resolution. Because we were plotting the same data over the same general areas, we expected to find the same relationship between melt days and backscatter (appendix A fig. 1). We were surprised, therefore, that our data supported the increase in backscatter with average melt days, but did not immediately support the idea of an abrupt drop-off in backscatter at high melt days, representing firn that has reached saturation. There are many limitations in our result that could explain this discrepancy.

First of all, there were differences in methodology. While Scambos et al. (2003) only used small samples of remotely sensed data for a limited number of ice shelves, we used remotely sensed data that spanned the entirety of ice shelves for a large number of shelves. The increased amount of data we used could be obscuring the physical relationship revealed by Scambos et al.'s more limited sampling. Or, alternatively, our extended data could be indicating that the relationship found by Scambos et al. is not, in reality, significant given a more representative sample. When we remove ice shelves on our average shelf value plot that were not analyzed by Scambos et al. (2003) (appendix A. fig 2), the drop-off pattern appears to emerge, indicating that data sampling has a large bearing on the relationship. We plan to try a variety of sampling techniques in the future to explore this effect.

Another limitation, due in large part to our more extensive data sampling, is the wide spread in the plotted values. Though there is clearly a strong positive relationship between melt days and backscatter, there is also a high standard deviation. This may be inherent to the nature of melt days and backscatter data. However, some of it might be due to differences in accumulation rates around Antarctica. Shelves with high accumulation rates can absorb a lot more meltwater before becoming saturated, because there is more snow full of air pockets that can be filled in. These shelves will plot below the main trend on the diagram. Shelves with low accumulation rates would experience the opposite effect and plot above the main trend. We intend to develop an empirical correction for accumulation rate to account for this bias. Theoretically, this will reduce our data spread and might even result in the emergence of a drop-off pattern.

One major problem affecting our results is the spatial referencing of the melt days data. Though the backscatter data and our ice shelf masks line up perfectly, the melt days data has an offset that is clearly visible upon inspection. We hypothesize that this has to do with ArcGIS's treatment of the Hughes 1980 ellipsoid, on which the dataset was developed. Though the plotted data is probably very close to being correct, we cannot have full confidence in the values until the referencing is corrected. Colleagues at the National Snow and Ice Data Center have offered to help us solve the problem, so we anticipate creating more reliable plots in the future.

Eventually, we hope that this project will utilize data from four different satellites: QuickScat, ASCAT, ERS-1, and ERS-2. The current project relies only on backscatter data from one satellite: QuikScat. It is unclear from the information on the SIR algorithm whether all of the data have been normalized and can be directly compared. We are waiting on a response from the dataset's author to determine whether we can cross-average and compare data from the four different satellites. In the meantime, however, we have proven the flexibility of the workflow by testing it on ERS-1 data, and found very similar results to QuikScat.

Conclusion:

Overall, we were very successful in achieving our project goal of creating a workflow to convert, average, clip, plot, and reclassify backscatter and melt days data from Antarctic ice shelves. Programming proved to be hugely valuable in this project because of the large data volume available for analysis. We were able to automate the process of data conversion and averaging, and then fed the averaged data into a script that carried out the rest of the procedures. This routine was developed using QuikScat data, but was proven flexible when we tested the same scripts successfully on ERS-1 data.

We were surprised to find that the most challenging part of the process was the data conversion. We gained experience using the Subprocess module to convert a very uncommon, proprietary data type (.sir), and became familiar with data conversion processes that can be used on netCDF files, a format growing in popularity. The spatial referencing of the melt days data proved to be the most daunting task, and despite using several approaches that seemed reliable based on the metadata, the dataset still does not line up correctly with the backscatter and shelf masks. We will continue this learning process in the future with the assistance of colleagues at NSIDC in order to align the data correctly and produce more reliable plot

Scientifically, our results supported the positive relationship between backscatter and melt days found by Scambos et al. (2003), but lacked an abrupt drop-off in backscatter representing the saturation of the firn layer. It is unclear whether our data indicate that the drop-off should not exist in reality, or if it is simply obscured by the large data spread due to our extensive sampling. We will continue the project by trying different sampling techniques and by correcting for accumulation rates, which could potentially tighten the relationship and reveal a drop-off. Once these corrections are made, we will be able to develop a vulnerability index based on the relationship using principal component analysis or other techniques. Then finally, relying on the workflow developed in this project, we will be able to create reliable maps of Antarctic ice shelf vulnerability to surface-melt-induced collapse.

Works cited:

Fahnestock, M., R. Bindschadler, R. Kwok, and K. Jezek, 1993: Greenland ice sheet surface properties and ice dynamics from ERS-1 SAR imagery. *Science*, **262**, 1530–1534.

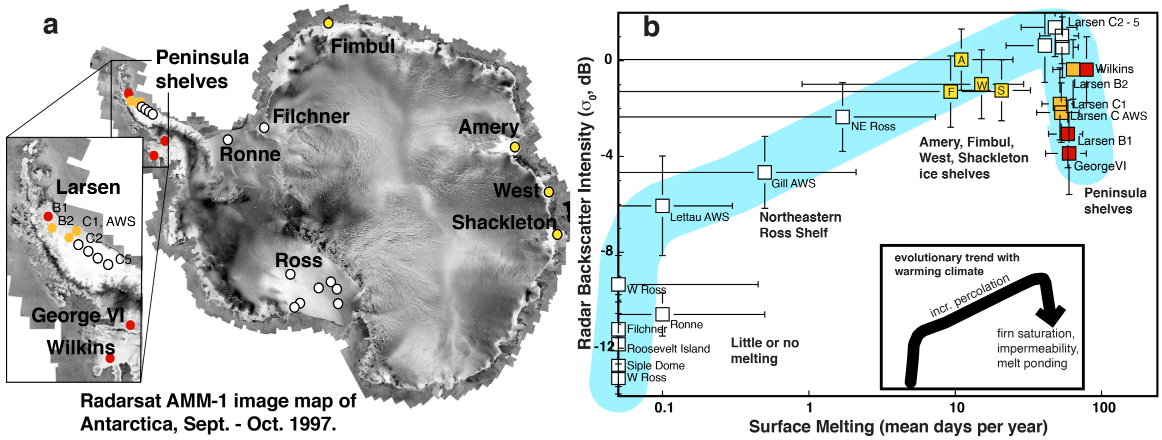
Haran, T., J. Bohlander, T. Scambos, and M. Fahnestock. 2005, updated 2013: *MODIS Mosaic of Antarctica (MOA) Image Map*. Boulder, Colorado USA: National Snow and Ice Data Center. http://dx.doi.org/10.7265/N5ZK5DM5.

Long, D. G., 2010: A QuikSCAT/SeaWinds Sigma-0 Browse Product. *BYU Microwave Earth Remote Sensing Laboratory Report*.

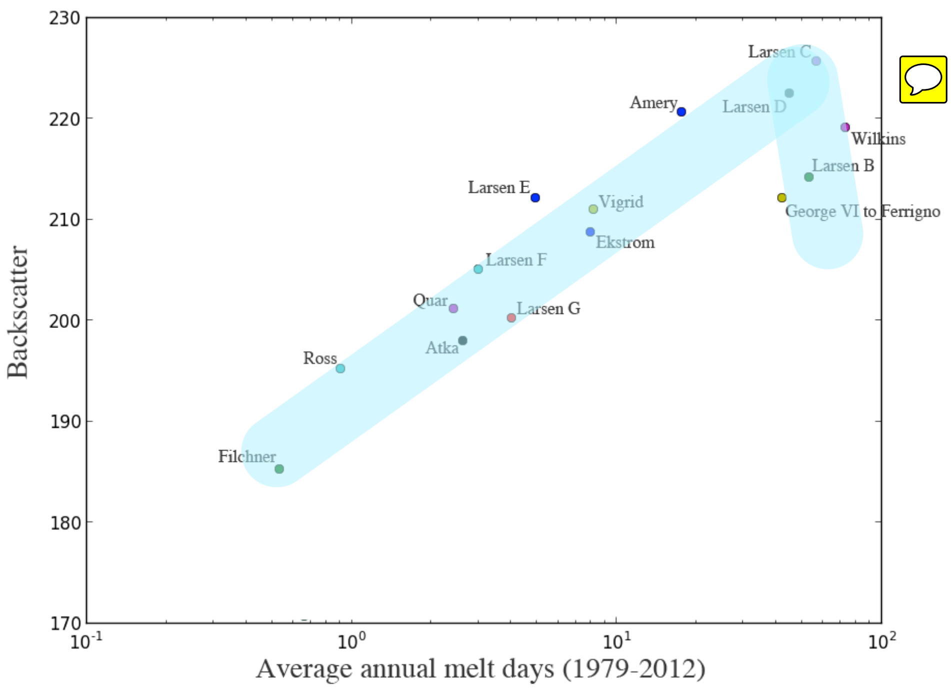
Picard, G., and M. Fily, 2006: Surface melting observations in Antarctica by microwave radiometers: Correcting 26-year time series from changes in acquisition hours. *Remote Sensing of Environment*, **104**, 325–336, doi:10.1016/,DanaInfo=dx.doi.org+j.rse. 2006.05.010.

Scambos, T., C. Hulbe, and M. Fahnestock, 2003: Climate-induced ice shelf disintegration in the Antarctic Peninsula. *Antarctic Peninsula Climate Variability: Historical and Paleoenvironmental Perspectives*, 79–92, doi:10.1029/079ARS07.

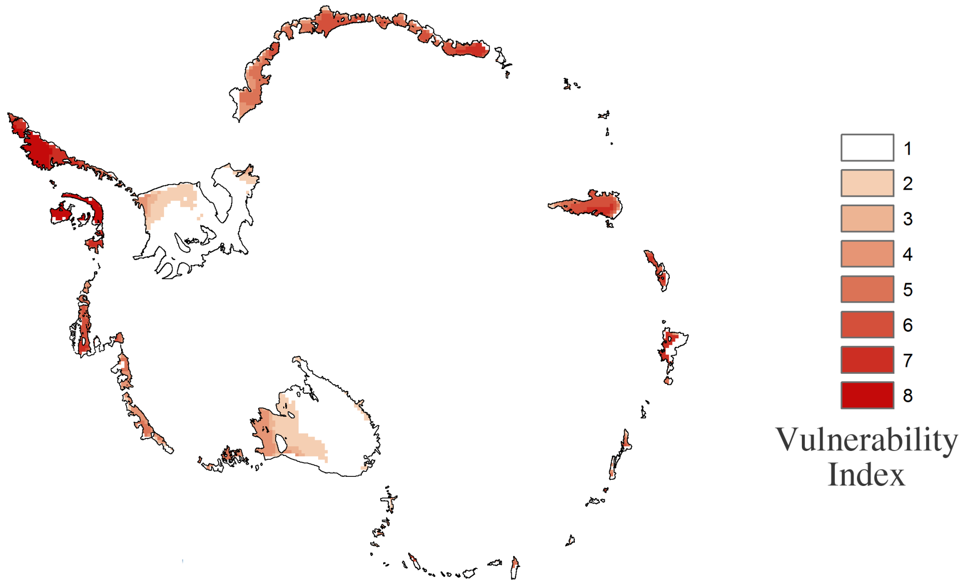
Appendix A: Additional figures



Appendix A figure 1: Figure from Scambos et al. (2003) demonstrating the relationship between melt days and backscatter. The sample sites are shown on the left, and the plotted points from the sample regions are shown on the right. Note the generally positive relationship between melt days and backscatter followed by an abrupt drop in backscatter when firn saturation takes place.



Appendix A figure 2: Shelf-averages plot from this study approximating the relationship shown by Scambos et al. (2003) by taking away ice shelves not used in Scambos et al.'s study.



Appendix A figure 3: Sample vulnerability index based on a reclassification of the melt days dataset. This is not a final vulnerability index, as it was mainly created in order to have a completely coded workflow. Ultimately, we hope to improve the relationship through better sampling techniques and corrections for accumulation and then create a more rigorous index, possibly based on a principal component analysis.

Appendix B: Code

Five scripts are used for the coding of our project: Three for data preparation, one for data analysis implementation, and one that is a module providing necessary functions for the project. They are briefly described below. Further information is included in the headers and comments of the attached script files.

Alley_sir_convert.py

Data preparation script; used to convert BYU .sir files in an efficient, automated procedure that uses a BYU executable file and a list of interactive commands that are piped into the executable. Requires user to input a path to a folder containing the script, the Windows executable, and a nested folder containing the .sir files. The name of the folder containing the .sir files is also specified.

Alley_backscatter_avg.py

Data preparation script; used to calculate and save a spatially referenced raster file of averaged backscatter rasters. Requires the user to input a path to a workspace containing any number of folders ending in “tiffs” that contain backscatter tiff files, as well as a filename for the output file.

Alley_MeltDays_avg.py

Data preparation script; used to calculate and save a non-spatially referenced raster file of averaged melt days rasters. Requires the user to input a path to a workspace containing the multiband annual melt days file, the name of the file, and a filename for the output file.

Alley_data_implementation.py

Data implementation script; used to resample, clip, plot, and reclassify the input data. Requires the user to input two main data files – backscatter and melt days – prepared using the data preparation scripts. Also requires a path to a folder containing shelf mask shape files, and a path to a folder in which reclassified rasters can be saved. Allows the user to select the output cell size and a snap raster. The script also returns pixel-by-pixel and average-value plots, and the user may wish to edit the labels on these plots.

Alley_project_module.py

Project module; contains functions utilized in the four scripts that carry out the project analysis.