ESRM433/SEFS533
Lab 9

Objectives:

- *Introduction to full waveform lidar*
- *Introduction to Global Ecosystem Dynamics Investigation (GEDI) Lidar*
  - https://gedi.umd.edu/
- *Introduction to The Ice, Cloud and land Elevation Satellite-2, or ICESat-2*
  - https://www.nasa.gov/content/goddard/about-icesat-2
- *Downloading and using GEDI data*
- *Characterizing forested landscapes with full waveform lidar*

Data and Software:

- *https://openaltimetry.org/data/gedi/*
- *https://search.earthdata.nasa.gov/search*
- *ArcGIS Pro*

What you will turn in:

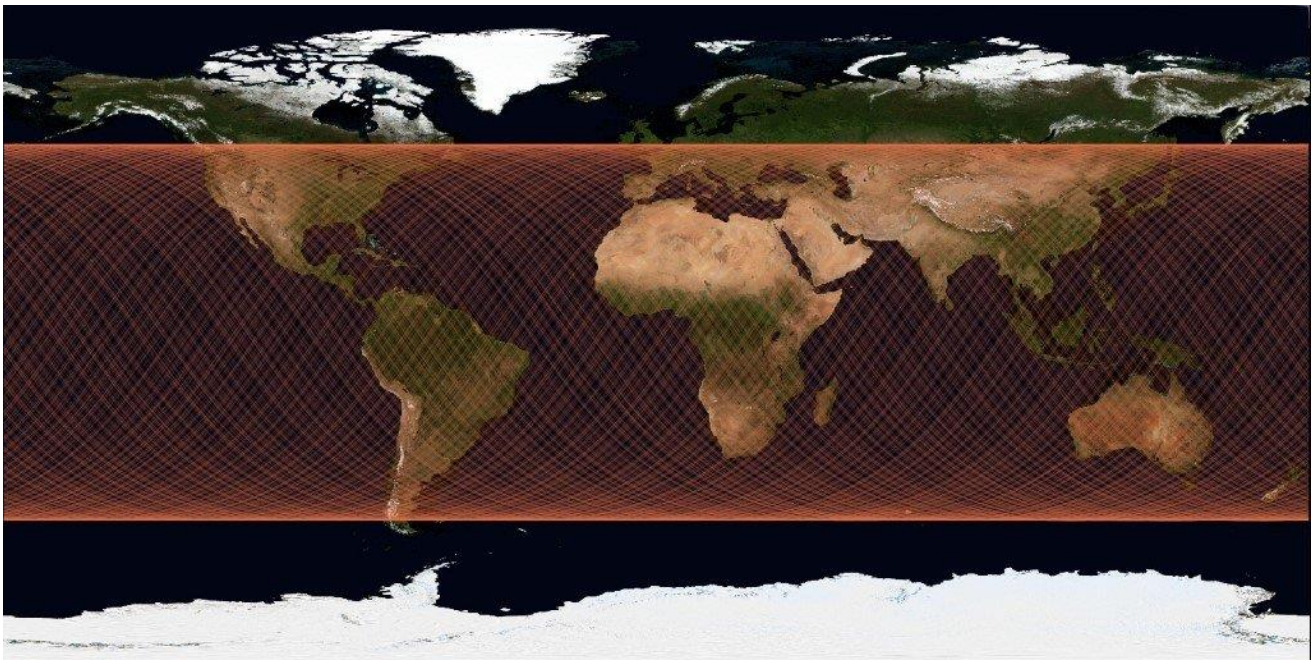
- *Submission Template as a PDF*

Welcome to Lab 9 for ESRM433/SEFS533

This lab is all about Spaceborn lidar. We have covered aerial lidar, and terrestrial lidar, now we will talk about spaceborn lidar scanning (SLS).
Two prominent spaceborn lidar sensors are ICESat and GEDI.
A lot of this lab will be you reading about the sensors and searching online resources for answers.

PART 1a GEDI background

For this lab, we are going to focus on GEDI data.
The GEDI sensor was launched to the International Space Station (ISS) in December of 2018 and is slated for a two year mission to scan terrestrial ecosystems. The data collected is limited to the orbit of the ISS so data isn't available for latitudes greater than 51.6^{o}.

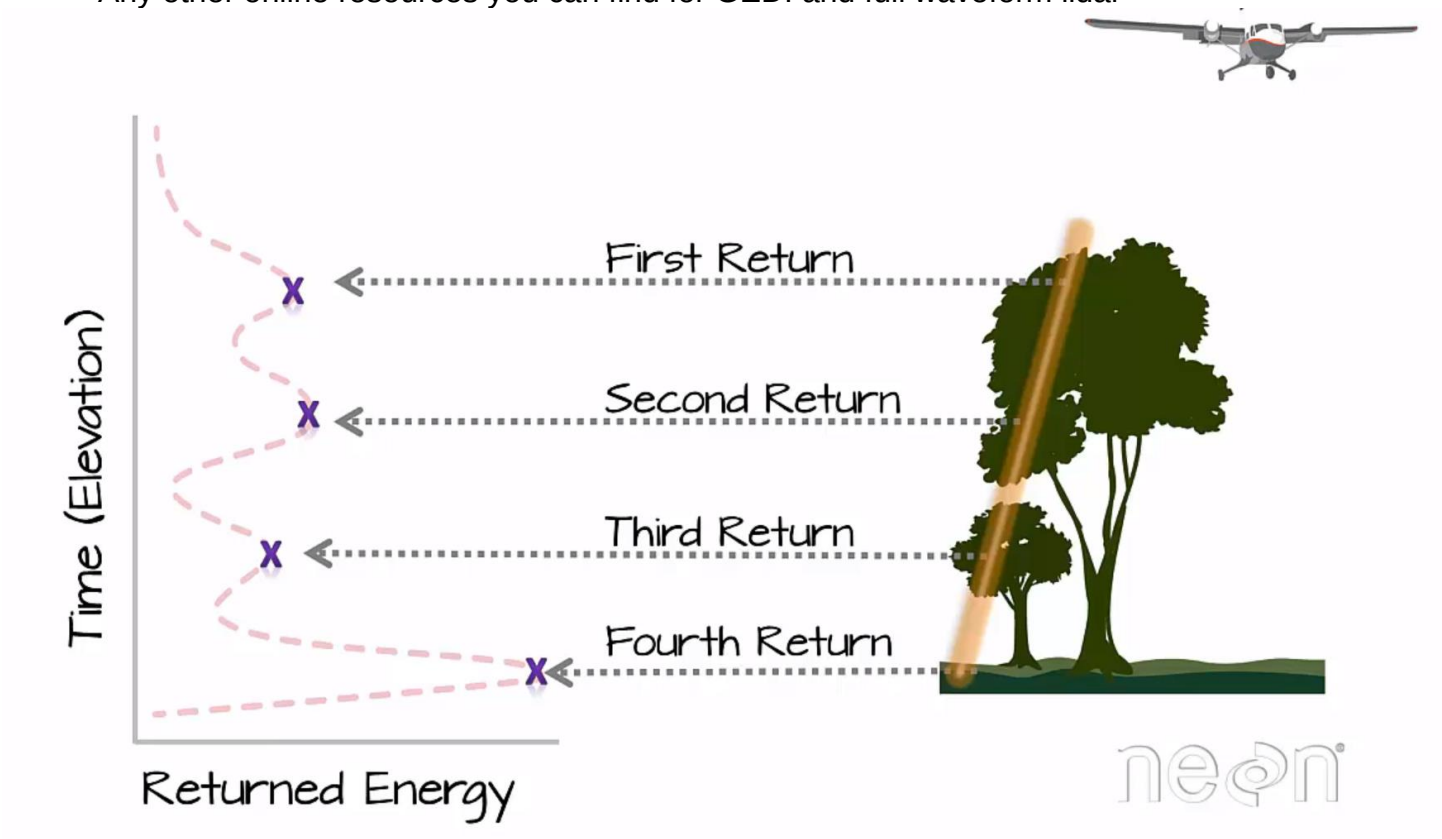


ESRM433/SEFS533

Lab 9

For the following questions refer to:

- GEDI's website https://gedi.umd.edu/
- GEDI videos
  - https://www.youtube.com/watch?v=SSdDPFfUVIo
  - https://www.youtube.com/watch?v=wxgrxvAKpTo
  - https://www.youtube.com/watch?v=XFlm-TmhvjM
- Any other online resources you can find for GEDI and full waveform lidar



The lidar data from GEDI retains it's full waveform. ALS data is also based on the amount of energy returned from a laser pulse, but the full waveform is discarded and only the peaks in the waveform are preserved. These peaks are what we call returns when using terms like "discreate return lidar" vs "full wavefrom lidar"

The beam divergence of lasers is extremely important to keep in mind when discussion SLS. Beam divergence is in direct relationship to the distance the laser pulse travels. For TLS, objects are much closer to the scanner so beam divergence is likely to be a few mm to a few cm depending on the distance. For ALS, beam divergence can be ~10cm to ~30cm. This divergence is what allows for multiple returns from a single pulse.

QUESTION 1: Is GEDI data free to the public? Where can you get GEDI data?

QUESTION 2: What is the footprint (resolution) of GEDI laser pulses? This is the same as beam divergence.

QUESITON 3: How many beams of data is GEDI collecting? What is their spacing?

QUESTION 4: Is every square meter of the earths surface getting sampled by GEDI?

QUESTION 5: What are the three science questions for GEDI?

QUESTION 6: For Data Products, there are multiple Algorithm Theoretical Basis Documents (ATBDs) for GEDI data. We are most interested in L1B, L2A, & L2B. What is the name of the data product for those 3 ATBDs?

PART 1b ICESat background

ICESat is like GEDI in many respects. An orbiting lidar sensor in space. A major difference is that GEDI is on the international space station while ICESat has its own dedicated satellite in a polar orbit. Take a moment to review the information here about ICESat:
https://www.nasa.gov/content/goddard/about-icesat-2

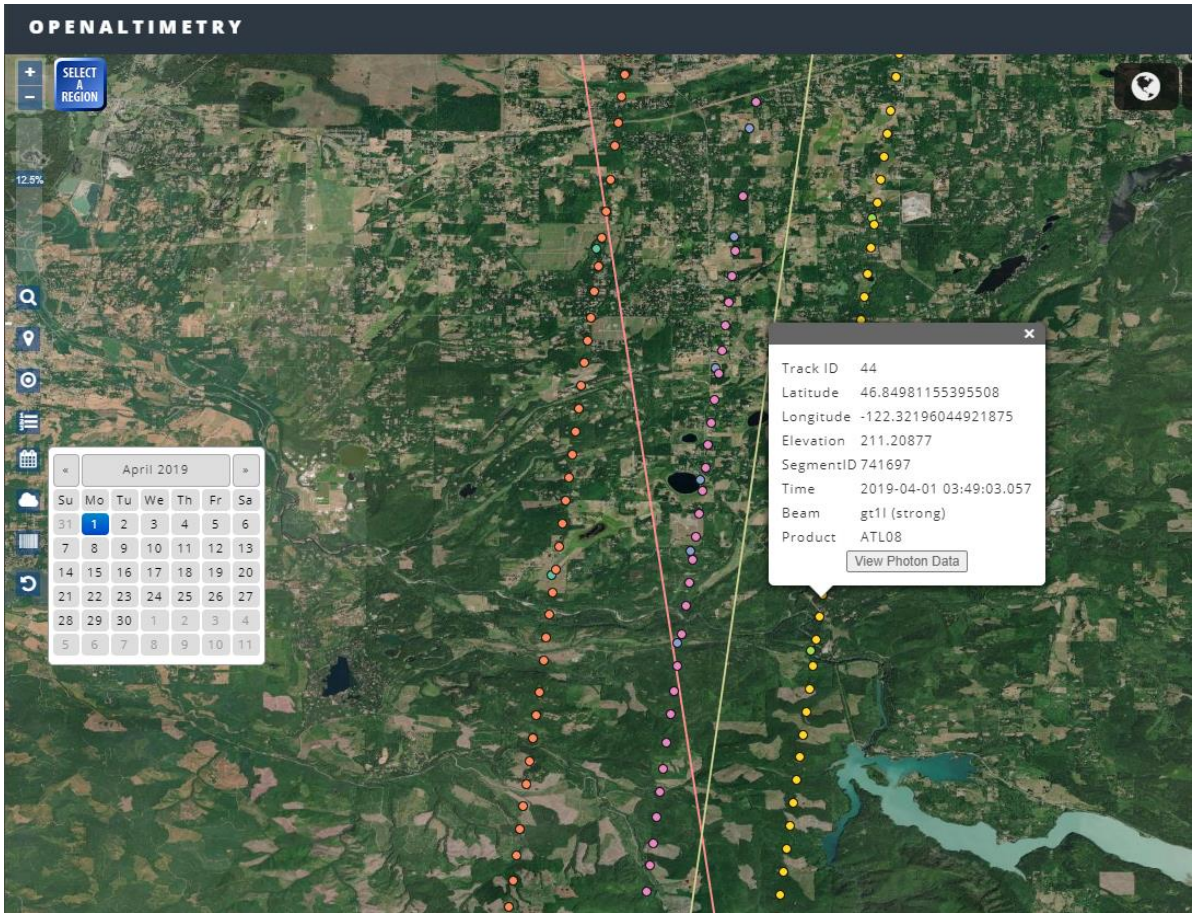
QUESTION 7: What are some similarities and some differences between GEDI and ICESat2?

QUESITON 8: What wavelength and "color" is the ICESat2 laser? What wavelength and "color" is the GEDI laser?

QUESTION 9: What is the spacing of the ICESat2 laser pulses?

The best way to manipulate and use ICESat data is by using python. Currently there is not an easy way to use R for ICESat2 data, but this will likely change soon. You can look at the locations of the track lines, and get photon data from ICESat2 here:
https://openaltimetry.org/data/icesat2/ (you can also see some GEDI data here)

You can zoom into the pack forest area and select the path nearest one of our areas of interest and select the different dates until you find data that has pulses near the forest that we have been looking at in the past labs.
If you click on the individual pulses, you can get the photon data.



QUESTION 10: Find the pulse closest to our pack forest area of interest. Take a screen shot of the photon data and discuss what the data represents.

QUESTION 11: Where can you download ICESat2 data?

PART 2: Getting and processing GEDI data

There is an R package for downloading and processing GEDI data. Unfortunately, it is currently not working with the new GEDI data volume. Check it out, but for this lab we are going to be using data we can acquire directly from GEDI data websites.
https://github.com/carlos-alberto-silva/rGEDI

An important part of this class is to demonstrate how accessible lidar data is for researchers. This is certainly true about GEDI data, however, GEDI data downloads are huge. The bundles are about 6GB with subsets of bundles only recently becoming available for download. However, even the subsets are over 2GB. This is in the highly compressed .h5 format. H5 is a Hierarchical Data format and if you want more info check out: https://www.neonscience.org/about-hdf5

ESRM433/SEFS533

Lab 9

Because of the size constraints, and that the R package rGEDI is currently not working, we will be using online and GIS resources to complete this lab. I encourage you all to keep an eye on the rGEID page and to run though the tutorials there if it gets updated.

Go to https://openaltimetry.org/data/gedi/ to view the online mapping of a small selection of GEDI data. There is only a small sample of GEDI data available to be viewed on this site, but it is sorted and visualized well. GEDI data is collected in swaths that cross the majority of the planet. There is only one pass of GEDI data across Washington state visualized on the Openatimetry site.

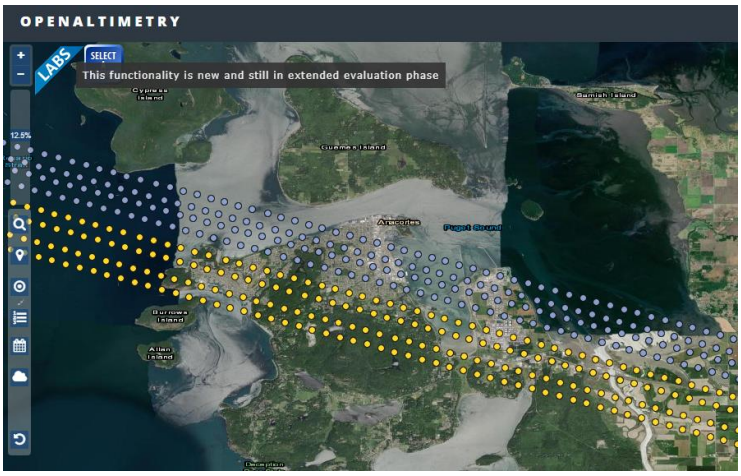
You are going to look at the data for 6 GEDI footprints. Which 6 are your choice. You are going to use these same footprints for lab 10 so the criteria are:

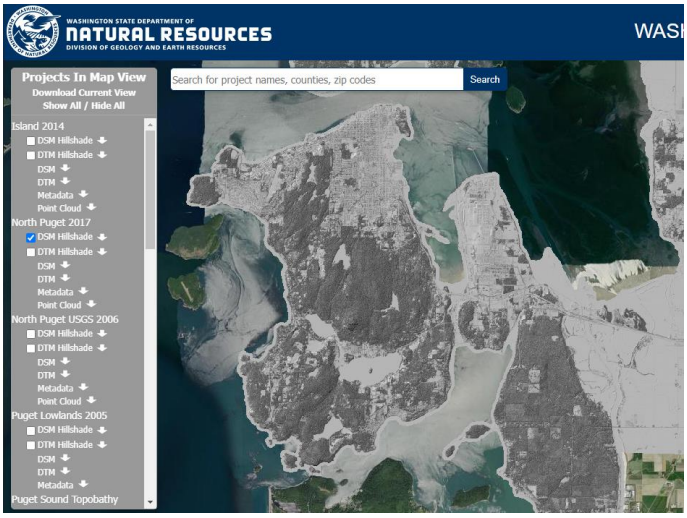
- 2 samples from sites with little to no vegetation.
  - Urban, plowed field, beach, desert, meadow, etc…
- 2 samples from sites with medium to light vegetation.
  - Young forest, park, agricultural, etc…
- 2 samples from sites with heavy vegetation
  - Mature forest

Make sure that there is ALS data for the sites. You can refer the WA DNR lidar portal to make sure your GEDI data area has ALS data. You won't be using ALS data for this lab, but you will be doing a comparison between the lidar data sets for lab 10. You should review lab 10 before selecting your sites.

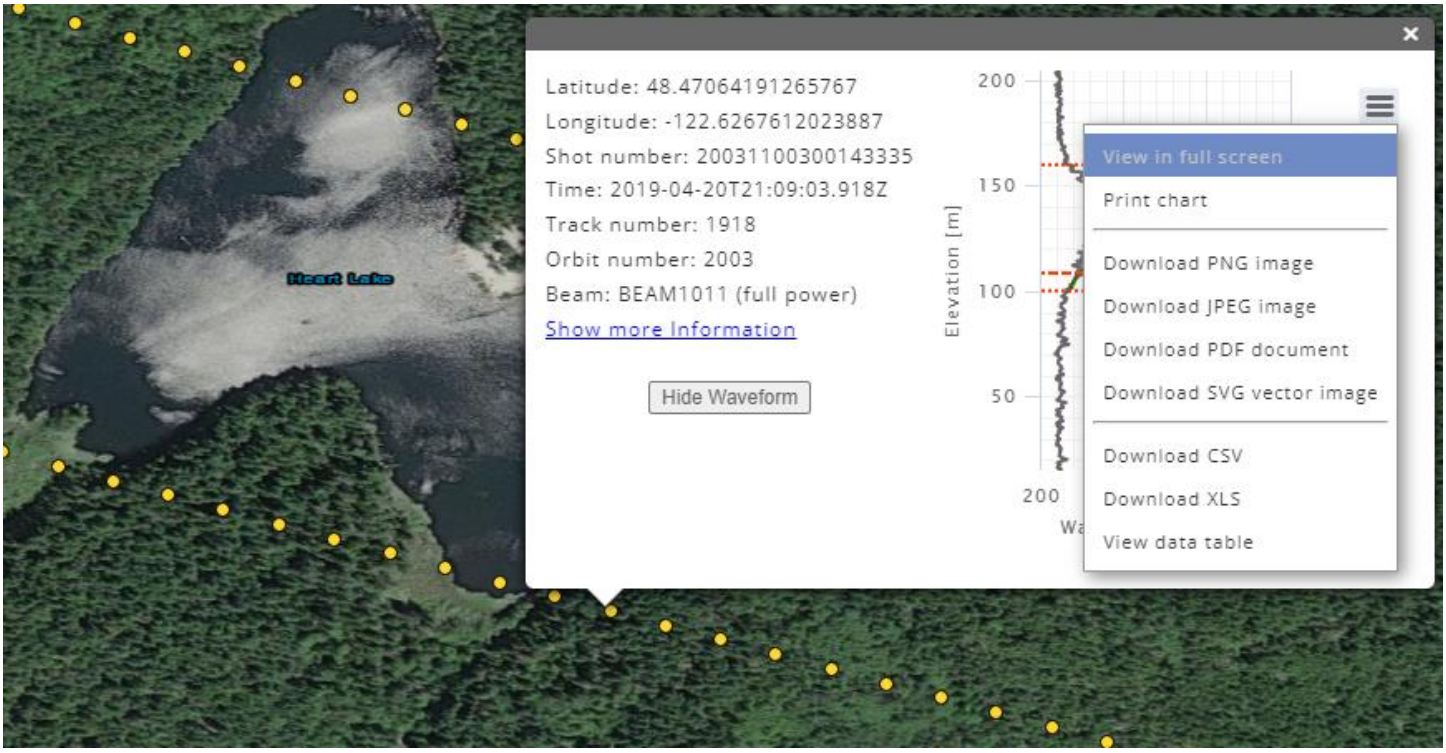
Fidalgo island and the city of Anacortes is a good location to work with, but you can choose anywhere that meets the above criteria. Don't sample points that are directly next to each other. Space them out so we can get a good sampling variability.

ESRM433/SEFS533
Lab 9

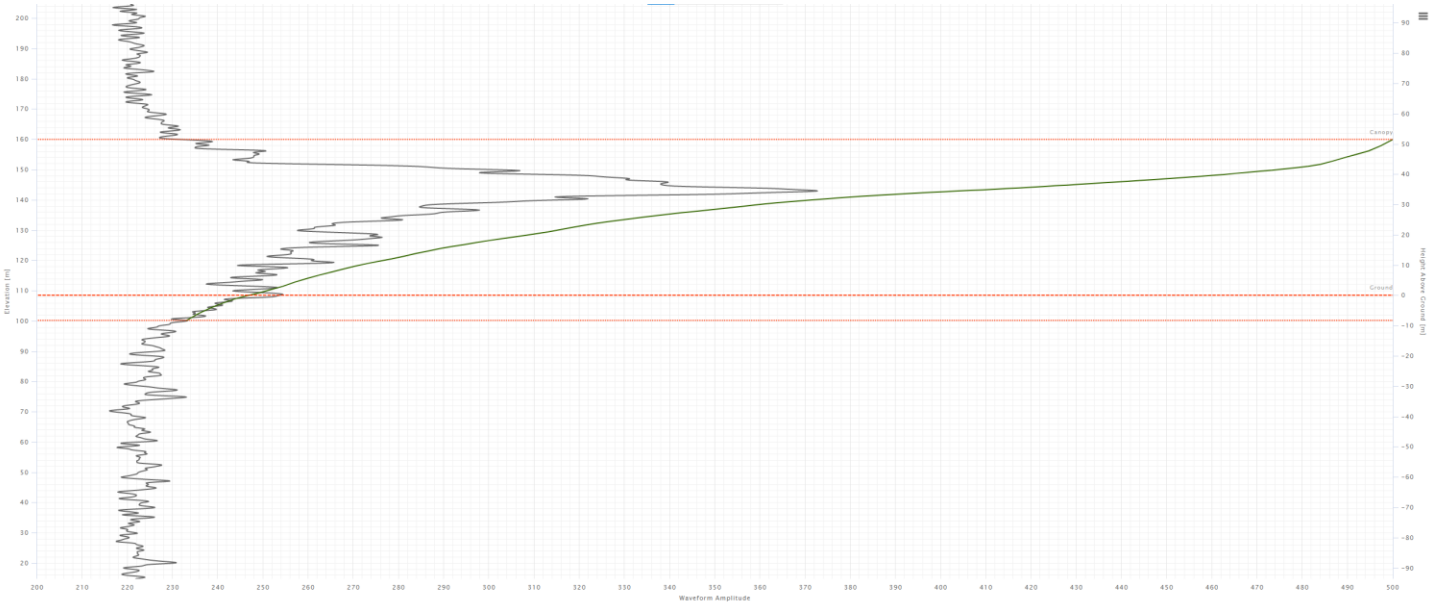




Zoom in and click on one of the footprints locations. Click on the “Show more information” to see metrics derived from the pulse. You can view the waveform and then click on the menu icon to “view in full screen”. Use one of the full power beams for your samples.
<u>Remember that there is a 7 to 10m potential error in the geolocation of each pulse.</u>



This provides a view of the full waveform data with an interactive slider.



PAI

rGEDI uses a metric called Plant Area Index (PAI). This is similar to leaf area index but it doesn't distinguish between live green leaves and woody tree stems.
https://en.wikipedia.org/wiki/Leaf_area_index
https://lpdaac.usgs.gov/documents/588/GEDI_FCCVPM_ATBD_v1.0.pdf

PAI differs from height metrics as it relies on the amount of energy that filters through the different layers of forest canopy. A forest with shorter trees but the trees have very full crowns and there is a full midstory layer of vegetation could have a much higher PAI than a forest with very tall trees, but a small crown ratio and little to no midstory vegetation.

Tang, Hao, et al. "Retrieval of vertical LAI profiles over tropical rain forests using waveform lidar at La Selva, Costa Rica." *Remote Sensing of Environment* 124 (2012): 242-250.

Zhao, Feng, et al. "Measuring effective leaf area index, foliage profile, and stand height in New England forest stands using a full-waveform ground-based lidar." *Remote Sensing of Environment* 115.11 (2011): 2954-2964.

LAI and PAI can be expressed as ratios with a value of 1 indicating that the area of plant matter surfaces is equal to the spatial planar area of the sample location. A value less than 1 will have less vegetation, while a value greater than 1 will have more vegetation. There is no upper limit to the lai values but a value approaching 10 would indicate a very dense forest. A value of 0 would be bare ground.

QUESTION 12: Fill out the table below in the submission template for your 6 sample locations. You are going to use this table again in lab 10. Note that most online web mapping applications default to WGS 84 web Mercator

Sample	Lat	Long	RH (100)	Canopy Cover	PAI	FHD
Low Veg 1						
Low Veg 2						

Med Veg 1						
Med Veg 2						
High Veg 1						
High Veg 2						

QUESTION 13: Include screenshots of the full waveform for each of your six sample locations AND a screenshot of the point location from the online map. TOTAL of 12 small screenshots. You will be using these again for lab 10. Make sure to label them with what Sample they are. Describe the similarities and differences between the waveforms and amount of vegetation present.

QUESTION 14: What exactly is the RH metric measuring? Check out: https://gedi.umd.edu/mission/technology/

The GEDI on the open altimetry site is from one week in 2019. GEDI data has been actively collected since that date with ever increasing overlap between the scan lines. Most users of GEDI data won't be doing anything with the single pulse footprint data, but rather the aggregate raster output of GEDI metrics that form the level 3 and 4 GEDI products.

You are going to need to create a free account with https://search.earthdata.nasa.gov/ to be able to download GEDI level 3 and 4 products.

Follow the prompts to make a new account. You will need to provide some basic information. This website is one of the best sources for global imagery data beyond just GEDI and ICESAT products.

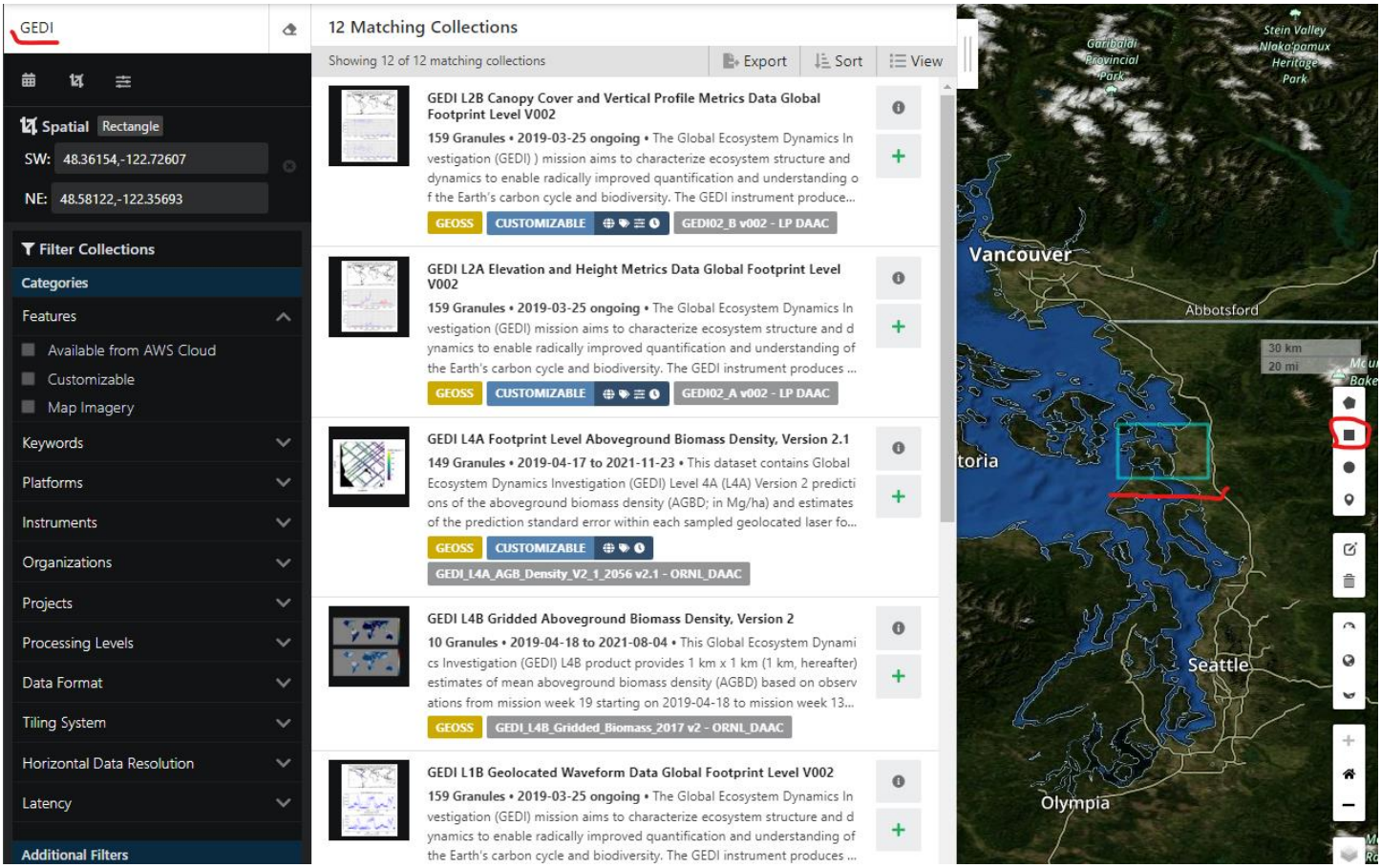
You can't zoom in that far, but you can use the polygon or square tool to highlight an area of interest and search for the term GEDI to pull up all of the available data. Remember, there is now 2 versions of GEDI data. Version 1 (V001) and version 2 (V002). A list of the improvements to version 2 can be found here:
https://lpdaac.usgs.gov/news/release-gedi-version-2-data-products/

QUESTION 15: Summarize the improvements of version 2 over version 1 for GEDI data.

When you search for GEDI data you can download L1B, L2A, L2B, L3, and L4 products. The L1 and L2 data are H5 files while the L3 and L4 products are GeoTIFFs.

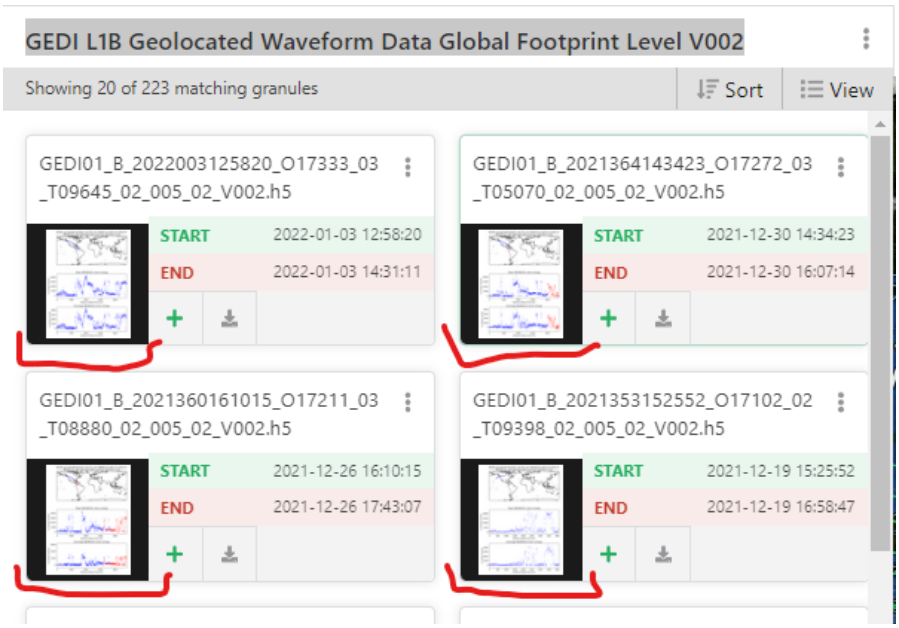
ESRM433/SEFS533

Lab 9



Lets check out the L1B data first. Click on the “GEDI L1B Geolocated Waveform Data Global Footprint Level V002” item and it will take you to all of the L1B data that cover our area of interest. There are summaries for all of the sections of data provided as image files.

We are not going to download the H5 files, but we can look at the data swath and summary point data for the L1B, L2A, and L2B data.



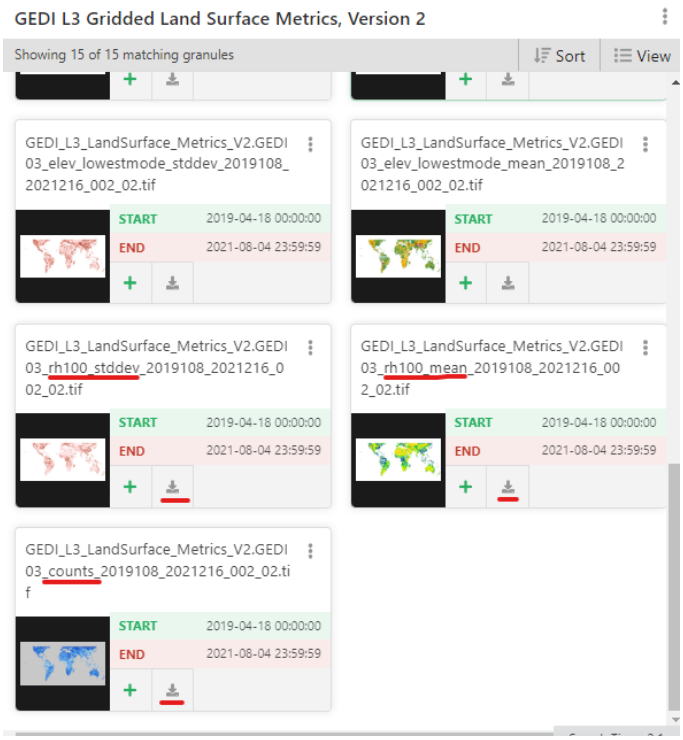
QUESTION 16: Include a screenshot of one of the summary image files for either a L1B, L2A, or L2B data product. Discuss the location of the swath and what the two scatter plots represent. Full descriptions are needed and not just repeating the title.

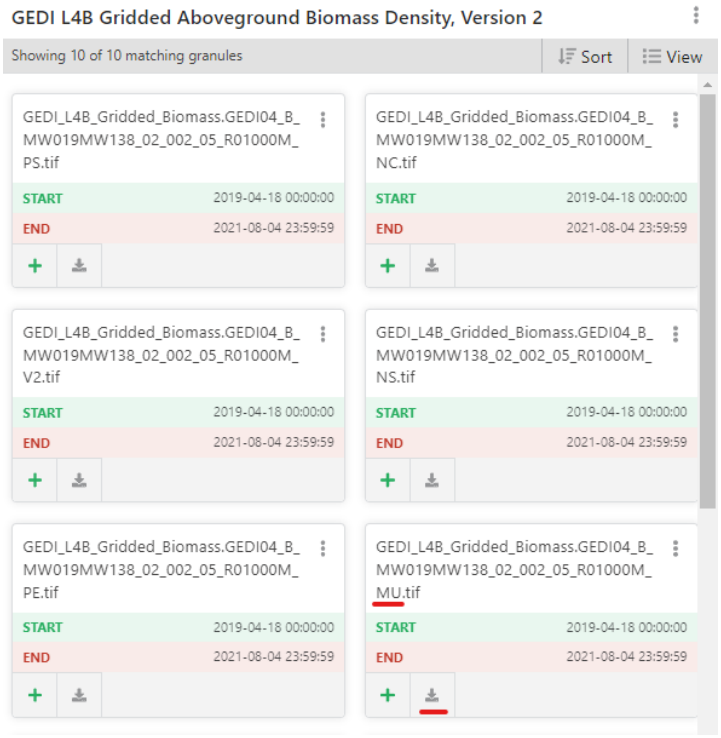
ESRM433/SEFS533

Lab 9

We are now going to download GeoTIFFs of the gridded GEDI data. The raster datasets have a resolution of 1km and are global coverage. The files are large and the resolution is course.

Download the most recent L3 Gridded Land Surface Metrics, Version 2. There are several versions on the website. The difference is the date range. Download the rh100_stddev, rh100mean, and counts raster data.





Download the GEDI L4B Gridded Aboveground Biomass Density: Mean aboveground biomass density (MU). For a full description of the L4B data product see: https://daac.ornl.gov/GEDI/guides/GEDI_L4B_Gridded_Biomass.html

A summary of all the raster outputs for the biomass density product is included as an appendix, as well as being available on the above website.

ArcGIS Pro

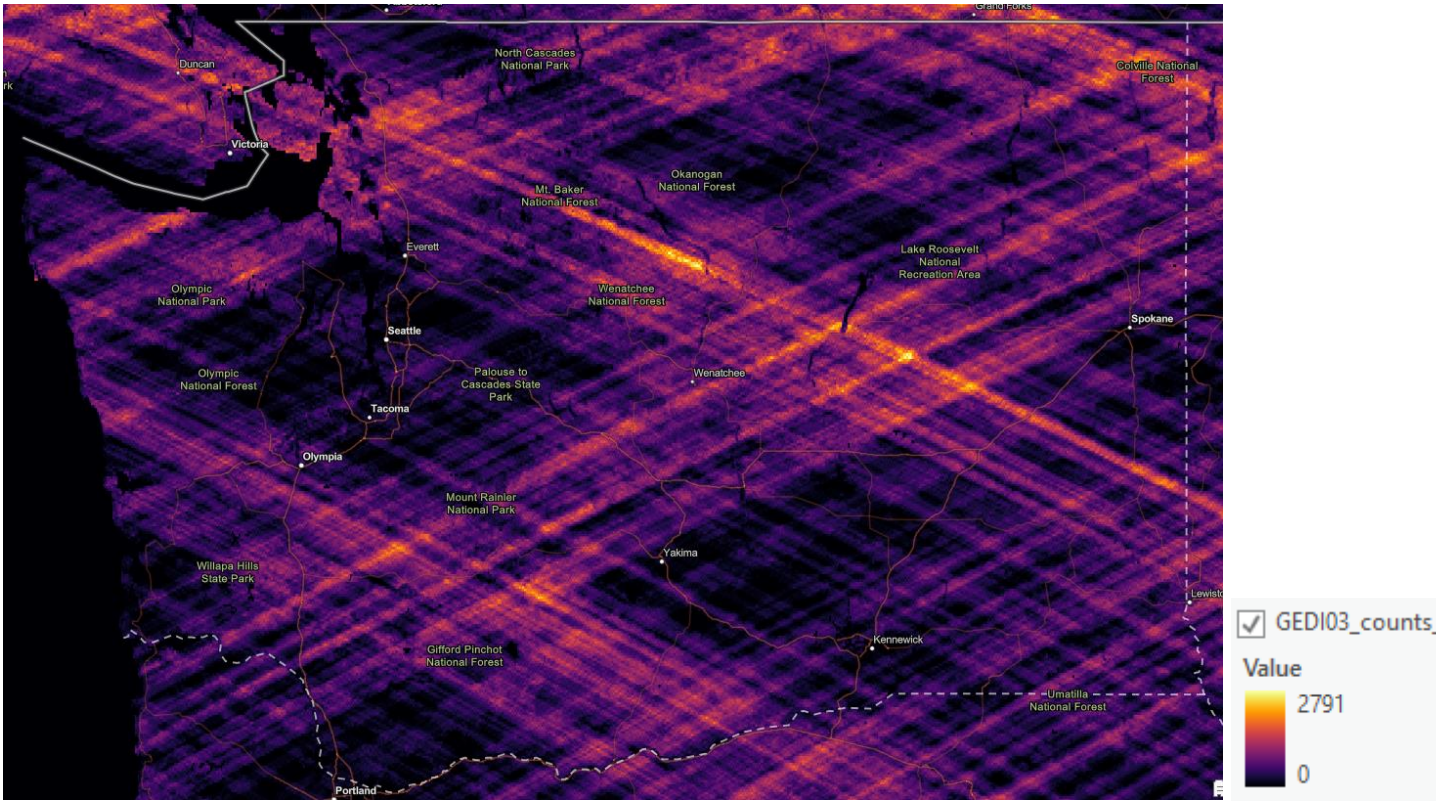
Create a new map project and bring your four downloaded rasters into ArcGIS Pro. The 1km resolution of the raster layers is very course compared to the other lidar data we have been working with so far.

QUESTION 17: What is the CRS of the downloaded rasters?

QUESTION 18: Similarly to question 12, fill out the table below in the submission template for your 6 sample locations. You are going to report the values of the raster cells located at each of your chosen GEDI footprints. Your Lat and Long will be the same from question 12. You are going to use this table again in lab 10.

Sample	Lat	Long	L3 RH100 Mean	L3 StdDev	L3 Counts	L4B Biomass
Low Veg 1						
Low Veg 2						
Med Veg 1						
Med Veg 2						
High Veg 1						
High Veg 2						

QUESTION 19: Create 4 Figures for the state of Washington of the GEDI output rasters. You must change the color ramp from black and white to something appropriate and include a legend. You can get as fancy as you want, but screenshots will suffice. They must have a descriptive caption beyond a single sentence. You can set the projection for the maps to any that you like. The CRS for the image below is WGS 84 Web Mercator (EPSG:3857), but that isn't the native CRS of the rasters.



This is an example of an figure for the # of GEDI pulse footprints within each raster cell. It is just two screen shots. Make sure to include a good figure caption that discribes the data.

GRADUATE STUDENT

GEDI data has been used for numerous ecological applications. Find a paper that is of interest to you that uses GEDI data and write a short paper review. Make sure to include the full citation.

File Names	Description	Units	No Data Value	Data Type
Data Files				
GEDI04_B_MW019MW138_02_002_05_R01000M_MU.tif	Mean aboveground biomass density (MU): Estimated mean AGBD for the 1 km grid cell, including forest and non-forest	Mg ha^{-1}	-9999	Float32
GEDI04_B_MW019MW138_02_002_05_R01000M_V1.tif	Variance component 1 (V1): Uncertainty in the estimate of mean biomass due to the field-to-GEDI model used in L4A		-9999	Float32
GEDI04_B_MW019MW138_02_002_05_R01000M_V2.tif	Variance component 2 (V2): If Mode of Inference = 1, this is the uncertainty due to GEDI's sampling of the 1 km cell. If Mode of Inference = 2, this is uncertainty owing to the model predicting biomass using wall-to-wall data, calibrated with the L4A footprint product		-9999	Float32
GEDI04_B_MW019MW138_02_002_05_R01000M_SE.tif	Mean aboveground biomass density standard error (SE): Standard Error of the mean estimate, combining sampling and modeling uncertainty	Mg ha^{-1}	-9999	Float32
GEDI04_B_MW019MW138_02_002_05_R01000M_PE.tif	Standard error as a fraction of the estimated mean AGBD (PE). If >100%, the cell values are truncated to 100.	percent	255	UInt8
GEDI04_B_MW019MW138_02_002_05_R01000M_NC.tif	Number of clusters (NC): Number of unique GEDI ground tracks with at least one high-quality waveform intersecting the grid cell			UInt16
GEDI04_B_MW019MW138_02_002_05_R01000M_NS.tif	Number of samples (NS): Total number of high-quality waveforms across all ground tracks within the grid cell			UInt16
GEDI04_B_MW019MW138_02_002_05_R01000M_QF.tif	Quality flag (QF): 0=Outside the GEDI domain 1=Land surface 2=Land surface and meets GEDI mission L1 requirement (Percent standard error <20% or Standard Error < 20 Mg ha-^{1})			UInt8
GEDI04_B_MW019MW138_02_002_05_R01000M_PS.tif	Prediction stratum (PS) determined by plant functional type and continent. PS is associated with an L4A model parameter covariance matrix that contributes to the Model Error Variance (Table 2).		0,241	UInt8
GEDI04_B_MW019MW138_02_002_05_R01000M_MI.tif	Mode of interference (MI): Method used for a particular cell. Until mission completion, only those cells where hybrid inference is possible will be populated with a mean biomass value 0=None applied 1=Hybrid Model-Based			UInt8

ESRM433/SEFS533
Lab 9