Applied Remote Sensing Training (ARSET) Program

Spaceborne Lidar for Monitoring Vegetation Structure and Biomass using GEDI

Homework Questions

Question 1

Which of the following statements best explains why full-waveform lidar is particularly well-suited for vegetation structure analysis compared to discrete return lidar?

Answers: (bold correct)

a. Full-waveform lidar has higher spatial resolution
b. Full-waveform lidar captures the continuous energy profile through the canopy, providing detailed information about vegetation density at different heights
c. Full-waveform lidar operates at different wavelengths
d. Full-waveform lidar has lower geolocation error

Feedback:
Full-waveform lidar records the complete return signal over time, creating a vertical profile that shows how laser energy interacts with vegetation at different canopy levels. This continuous profile allows researchers to analyze canopy density, layering, and structure in ways that discrete return systems (which only record peak returns) cannot achieve.

Question 2

You are planning a forest carbon assessment project in a mountainous region with steep terrain and dense canopy cover. Based on GEDI's limitations as a spaceborne lidar system, which of the following would be your primary concern?

Answers: (bold correct)

a. Spectral resolution limitations
b. Temporal resolution of 16 days
c. Geolocation error potentially affecting precise plot-level measurements
d. Limited penetration through cloud cover

Feedback

Spaceborne lidar systems like GEDI have geolocation uncertainty of ±10 meters, which becomes critical when trying to match footprints with precise ground plot locations. This is why best practices recommend using 45m diameter homogeneity buffers (25m footprint + ±10m geolocation error). In steep terrain, this uncertainty is amplified and can significantly affect the accuracy of plot-level carbon assessments.

Question 3

You are analyzing GEDI waveforms from two different forest sites. Site A shows a waveform with a sharp, high-amplitude peak near the beginning followed by several smaller peaks and a final ground return. Site B shows a waveform with multiple peaks of similar amplitude distributed throughout the signal before the ground return. Based on these waveform characteristics, what can you infer about the vegetation structure differences between these two sites?

Answers: (bold correct)

a. Site A has taller trees than Site B
b. Site A likely has a dense canopy with clear vertical stratification, while Site B has more evenly distributed vegetation throughout the canopy
c. Site B has better data quality than Site A
d. Site A is experiencing seasonal leaf-off conditions.

Feedback:

Waveform shape directly reflects vegetation structure. A sharp initial peak suggests dense upper canopy (high leaf area at the top), while multiple similar peaks indicate vegetation distributed more evenly through the vertical profile. This interpretation is fundamental to understanding how lidar energy interacts with different forest structures.

Question 4

Relative Height (RH) metrics like RH95 and RH50 are derived from the cumulative energy distribution of the waveform, where RH95 represents the height below which 95% of the canopy energy is contained.

Answers: (bold correct)

a. **True**
b. False

Feedback:
Relative Height metrics are calculated from the cumulative energy profile of the waveform. RH95 indicates the height below which 95% of the waveform energy is found, providing a robust measure of canopy height. RH50 represents the median canopy height, and these metrics are less sensitive to noise than peak-based measurements.

Part 1 Exercise

The following three questions are based on the provided GEDI L1B processing Jupyter notebook from Part 1, Exercise 1. To answer the questions, first complete the following workflow:

Data Processing:

1. Revise the Python notebook code to focus on a different area of interest – this time a tropical forest region in Ucayali, Peru east of Pucallpa. Change the aoi_bbox variable by inputting the following coordinates: aoi_bbox = [-74.2, -8.2, -74.1, -8.1]
2. Run the main processing code block to authenticate with NASA Earthdata and search for GEDI L1B granules
3. Allow the automatic retry system to find a granule with quality data in the new AOI

Question 5

Examine the "Processing Summary" output. How many files were attempted before finding quality data?

Answers: (bold correct)

a. 0
b. 1
c. 2
d. 4

Feedback:
There was 1 file (GEDI01_B_2019179175856_O03073_04_T05392_02_005_01_V002.h5) that was attempted with a 0% retention rate due to no quality data before finding quality data in the second attempt.

Question 6

What year was this GEDI overpass?

Answers: (bold correct)

a. 2019
b. 2020
c. 2021
d. 2022

Feedback:
From the "Processing Summary" output, we can see that overpass year is 2019 (with the Julian day of 220), as seen from file name: GEDI01_B_2019220015939_O03699_04_T02546_02_005_01_V002.h5.

Question 7

What is the size of the file that was successfully processed?

Answers: (bold correct)

a. 1.87 GB
b. 2.21 GB
c. 5.33 GB
d. 2.50 GB

Feedback:
From the "Processing Summary" output, we can see that the .h5 file has a size of 2.50 GB.

The following four questions are based on your results.

Question 8

How many Full-Power beams contained quality data in your final dataset and what beams were they?

Answers: (bold correct)

a. 2 beams, 'BEAM0100', 'BEAM0110'
b. 1 beam, 'BEAM0110'
c. 2 beams, 'BEAM0101', 'BEAM0110'
d. 1 beam, 'BEAM0101'

Feedback:
There were 2 Full-Power beams that contained quality data in the final dataset: 'BEAM0101', 'BEAM0110'.

Question 9

How many total high-quality shots were extracted from the HDF5 file among all beams?

Answers: (bold correct)

a. 23 high-quality shots
b. 19 high-quality shots
c. 61 high-quality shots
d. 110 high-quality shots

Feedback:
There were 61 total high-quality shots: 41 (from BEAM0101) + 20 (from BEAM0110) = 61.

Question 10

What was the elevation range (min to max) of your elevation_bin0 values?

Answers: (bold correct)

a. **255.6 to 285.1 m**
b. 225.1 to 285.1 m
c. 266.5 to 275.8 m
d. 516.2 to 655.7 m

Feedback:
The elevation range is 255.6 to 285.1 m, which can be estimated in the elevation histogram plot or seen directly from the summary output.

Question 11

Based on the 10 waveform plots generated in your analysis, which statement best describes the patterns of ground return clarity and their relationship to forest structure?

Answers: (bold correct)

a. **Shots #4 and #9 show the clearest ground returns; Shot #4 likely represents an area with minimal canopy cover (possibly a recently deforested area), while Shot #9 shows a distinct but lower-amplitude ground return with some above-ground signal.**
b. All shots show equally clear ground returns because GEDI's high-energy laser can always penetrate through any forest canopy to reach the ground.
c. Shots #2, #8, and #10 show the clearest ground returns because these locations have the tallest trees with the most complex canopy structure.
d. Ground returns cannot be reliably identified in any of the shots because the 25m footprint is too large to distinguish between canopy and ground signals.

Feedback:
Shots #4 and #9 show the clearest ground returns; Shot #4 likely represents an area with minimal canopy cover (possibly a recently deforested area), while Shot #9 shows a distinct but lower-amplitude ground return with some above-ground signal suggesting a forested area with less dense canopy that still allows complete laser penetration to the ground surface.

Ground return clarity in GEDI waveforms is directly related to canopy density and structure. Shot #4's clear, high-amplitude ground return with minimal canopy peaks suggests limited vegetation, possibly indicating a deforested area or agricultural clearing. Shot #9 shows a distinct ground return with moderate amplitude, indicating the laser successfully penetrated through the canopy despite the presence of numerous canopy layers of above-ground vegetation—this pattern is typical of forests with less dense canopies or more open understory structure.

In contrast, shots like #2, #8, and #10 show more complex multi-peak waveforms with less distinct ground returns, suggesting denser, more stratified canopy structures where the laser energy is attenuated as it interacts with multiple canopy layers. This is a common pattern in intact tropical forests like those found in Ucayali, Peru.

Key concepts demonstrated:

- Clear ground returns with minimal canopy signal = sparse vegetation or open areas
- Distinct ground returns with moderate canopy signal = penetrable forest canopy
- Ambiguous or weak ground returns with strong canopy signals = dense, multi-layered forest structure
- Multiple peaks throughout the waveform = vertical stratification and complex canopy structure

Understanding these patterns is essential for interpreting GEDI data quality and selecting appropriate footprints for your analysis. In exceptionally dense tropical forests, GEDI may experience canopy penetration failure where the signal is completely attenuated before reaching the ground, resulting in no detectable ground return. On steep slopes, the 25m footprint may span multiple elevations simultaneously, making it difficult to identify a single ground return elevation.

For extra learning, you can now try re-running the code in your own area of interest by changing the input coordinates for aoi_bbox = # [min_lon, min_lat, max_lon, max_lat]. After completing the main data processing, run the waveform visualization code block with different user_sd_from_mean values to explore shots at different elevation characteristics. High elevation areas often have different forest structures than low elevation areas, which may be reflected in waveform characteristics, though obtaining quality data in high-elevation areas can be challenging.

If you run into any challenges or technical issues, consider consulting the Q&A documents associated with this training series.

Part 2 Exercise

Introduction to the GEDI Mission and Its Derived Products

Question 12

A researcher needs GEDI data for a study comparing canopy cover and Plant Area Index (PAI) across different forest types. Which GEDI product should they use, and what is the spatial resolution of this product?

Answers: (bold correct)

a. GEDI L1B at 25m footprint resolution
b. GEDI L2A at 25m footprint resolution
c. GEDI L2B at 25m footprint resolution
d. GEDI L3 at 1km gridded resolution

Feedback:
GEDI L2B contains canopy cover fraction, vertical profile metrics, Leaf Area Index (LAI), Plant Area Index (PAI), and foliage height diversity at the 25m footprint level. L1B contains only raw/geolocated waveforms, L2A contains elevation and RH metrics but not cover or PAI, and L3 provides gridded summaries rather than individual footprint measurements needed for detailed forest type comparisons.

Part 3 Exercise

The following two questions are based on this scenario.You've downloaded GEDI L2B data for your study area and notice that many footprints have PAI (Plant Area Index) values of -9999. After reviewing the tutorial workflow, you also notice that the l2b_quality_flag has values of both 0 and 1 across your dataset.

Question 13

What combination of quality filtering steps should you apply to create a reliable dataset for vegetation analysis?

Answers: (bold correct)

a. Keep all data with l2b_quality_flag = 0 only
b. Remove all footprints where PAI = -9999, and keep only l2b_quality_flag = 1 and degrade_flag = 0
c. Remove footprints with PAI = -9999 but keep all quality flag values
d. Only filter based on beam type (Power vs Coverage beams)

Feedback:
The value -9999 is a standard "no data" indicator in GEDI products and must be removed. The l2b_quality_flag = 1 indicates the highest quality L2B retrievals, while degrade_flag = 0 ensures the laser was operating at optimal energy levels. Combining these filters ensures you're working with reliable, high-quality measurements. The tutorial demonstrates that applying these filters together typically removes a substantial portion of data but dramatically improves data reliability for analysis.

Question 14

Your analysis shows that Power beams (BEAM0101, BEAM0110, BEAM1000, BEAM1011) have a higher percentage of footprints passing quality filters compared to Coverage beams (BEAM0000, BEAM0001, BEAM0010, BEAM0011) in your study area. What is the most likely explanation for this difference?

Answers: (bold correct)

a. Power beams sample different geographic areas than Coverage beams.
b. Power beams have higher laser energy, resulting in better signal-to-noise ratios and more reliable ground detection.
c. Coverage beams are only designed for mapping topography, not vegetation.
d. Power beams operate at different wavelengths optimized for vegetation.

Feedback:
GEDI's Power beams use higher laser energy than Coverage beams, which generally provides better penetration through dense canopies and more reliable ground detection. This often results in higher quality metrics, especially in areas with dense vegetation. Both beam types operate at the same wavelength (1064 nm) and can measure both topography and vegetation, but the Power beams' higher energy makes them preferable for challenging conditions. The tutorial demonstrates this by showing separate Power and Coverage beam statistics, allowing researchers to compare data quality between beam types for their specific study area.