Part 1 Question & Answer Session

Please type your questions in the Question Box. We will try our best to answer all your questions. If we don't, feel free to email Erika Podest (erika.podest@jpl.nasa.gov) or Sean Healey (sean.healey@usda.gov). This document will be shared to the training webpage within one week.

1. **Question: Are we measuring biomass change, or are we producing a sophisticated model-based approximation?**
   Answer: We are using model-based estimation based upon observations provided by GEDI. We propagate uncertainty of the relevant models to provide a statistical estimate of uncertainty.

2. **Question: Can space-based biomass estimates be accurate and reliable enough to support high-stakes carbon decision-making? If not, what monitoring approaches would be more appropriate? In cases where comprehensive ground-based monitoring is not feasible, can space-based biomass estimation serve as a practical alternative or complementary approach? To what extent is this approach practically applicable?**
   Answer: Longitudinal ground-based monitoring that goes back in time is the gold standard, but expense is a frequent barrier. In my opinion, OBIWAN provides a best-available estimate of biomass change in many parts of the world. In the US, where calibration with Forest Service inventory data is available, I do believe OBIWAN is a credible means of decision support. We provide probabilistic estimates of uncertainty.

3. **Question: Should temporary forest carbon storage be treated as equivalent to permanent fossil-fuel emission reductions?**
   Answer: Permanence is an important part of forest carbon accounting frameworks. "Permanent" is obviously a relative term, as disturbance is inevitable; the answer to this question depends upon the time horizon that is targeted by a particular mitigation framework.

4. **Question: Is RH50 1/2 the ELEVATION of RH100 or 1/2 the BIOMASS of the vertical column?**

Answer: Neither. RH50 is the height below which half of the canopy material is found. Midpoint of the density. Half of the energy has returned back. Depending on the canopy profile, this may be more or less than the height of the canopy top.

5. **Question: Is it possible to access the Lidar data used in the Duncanson et al. research? Is it open source?**
   Answer: Much of that data was provided under sharing restrictions, and I do not believe it is public.

6. **Question: How do you see the potential of incorporating information from Sentinel or higher-resolution satellites in addition to the Landsat model?**
   Answer: This process would likely benefit from using inputs from the Harmonized Landsat Sentinel (HLS) dataset, but that dataset does not go back far enough to support the applications described here.

7. **Question: How do I fuse Landsat imagery with GEDI data? Which software can I use to do this, and can I use different covariates from Landsat such as NDVI, EVI, etc.? Additionally, is it possible to use Sentinel-1 and Sentinel-2 instead of Landsat for biomass estimation?**
   Answer: The Google Earth Engine catalog describes the GEDI asset. That might be the easiest pathway to integration with Landsat or other sensors.

8. **Question: Would the bootstrapping time series you mention be similar to performing multiple simulations, like a Monte Carlo simulation?**
   Answer: Bootstrapping is similar to MC simulation in that there is a stochastic element generating variance around your output estimate. Bootstrapping in our context relates to “jittering” particular models whereas MC simulation can involve a broader range of simulated errors.

9. **Question: Is estimating biomass using GEDI applicable and efficient in drylands and grasslands?**
   Answer: GEDI does collect high-quality measurements in dryland forests. I think the OBIWAN fusion of GEDI and Landsat requires further validation in dryland forests.

10. **Question: I think you mentioned that the assumptions that go into the biomass estimations breakdown for areas <500 ha. Did I hear that correctly? If so, could you please speak a little more about that?**

    Answer: Patterson et al. describe our model-based methods as a large-area statistical method. Specifically, we do not account for spatial autocorrelation of model errors, limiting applicability in small areas. At finer scales, we don't trust those estimates unless at least 500 ha.

11. **Question: It seems there are quite big spatial gaps between GEDI footprints. How do you manage these gaps when producing pixel-level biomass estimates?**

    Answer: We make pixel-level biomass density predictions at the local scale, which improves model precision. For OBIWAN, that means that we make a biomass model for 10x10 km squares; even with the gaps you see, there are usually several thousand GEDI shots at that scale that we can use for model-building.

12. **Question: Why is GEDI L4A Raster Aboveground Biomass Density, Version 2.1 data not currently available after 2025-07-01? Will there be continuity in the GEDI series of data in future? Because for carbon projects GEDI AGB data is very crucial.**

    Answer: The GEDI team is working to reduce latency of public release of the data, which we recognize is a real problem for important applications. I think you should be seeing the delay decrease substantially through this summer.

13. **Question: What considerations about product resolutions do we include in the sample process to elaborate biomass models with optical images (or other covariables)?**

    Answer: The GEDI L4B product provides an estimate of mean aboveground biomass at the 1km scale. This incorporates model and sampling variability involved with the actual 25m footprint measurements. When we use GEDI footprints to fit biomass models to 30m Landsat pixels, we treat the resolutions as equivalent, and we do not account for imperfect alignment of the measurements.

14. **Question: I wonder how you validate the biomass estimation result. Are you comparing the biomass estimate with available datasets? And what is the method to do the validation?**

    Answer: This question will be comprehensively answered in the next session, but briefly, we combine the repeated biomass measurements from hundreds of field plots into a pseudo test area, and we compare the measured biomass change to the biomass change predicted by OBIWAN.

15. **Question: May I ask if a polygon can be uploaded to OBIWAN to obtain the statistics you showed before?**

    Answer: The OBIWAN API, which you will see in the next session, allows uploading of customized polygons.

16. **Question: Is it possible to use gedi 4A/B along with other fine resolution imagery such as Sentinel-2, Sentinel-1, or a DEM for machine learning input to estimate AGBD, for example, like in 1ha polygon level?**

    Answer: Many people have used GEDI biomass predictions to calibrate "fusion" with other types of imagery. Google Earth Engine, which hosts a GEDI asset that can be found in the catalog, provides an easy way to use L4A data with machine learning tools.

17. **Question: Is OBIWAN data available for every place in the world?**

    Answer: GEDI data is limited to the temperate and tropical ecosystems that the Space Station flies over. Eventually, the OBIWAN API is likely to contain bootstraps for that entire area, but right now bootstrap production has been focused on the US and Nepal.

18. **Question: I also have the question regarding the waveform signal from GEDI L1B. Does the reflected waveform always have a similar pattern? For example, is the bare soil always represented by one peak for bare soil (no vegetation area), and multiple peaks for the forestry area? And how do we know if the second/last peak is the ground?**

    Answer: With more complex waveforms (i.e., more peaks), finding the ground becomes trickier. Different ground-finding algorithms produce different sets of relative height values from the same waveform. While the GEDI mission recommends a particular algorithm for every region, it distributes sets of RH values from up to 6 ground-finding algorithms in

case researchers find that one approach works better for a particular application.

19. **Question: For the OBIWAN API, is there any documentation or source code available for calibration?**
    Answer: There is an API endpoint that we will see in the next session, which describes syntax and arguments for relevant API calls.

20. **Question: Do the L4A and L4B biomass models consider the year of the GEDI observation, or are all observations used?**
    Answer: Currently, all observations are used together. In the future, data may be split into epochs.

21. **Question: How is FIA data used if coordinates from FIA are fuzzy?**
    Answer: We use the real coordinates on secure FIA machines. The FIA calibration functions that get used in the public asset are simple linear equations (applicable to all model predictions), which do not carry any geospatial information.

22. **Question: What would you suggest to calculate biomass/height for agroforestry with 5 to 10m tree spacing?**
    Answer: GEDI waveforms are measured over approximately 25m footprints. The RH98 variable from the Level 2A product would be the best guess at overall forest height. However, there is approximately 10m location uncertainty on GEDI footprints. Also, I'm not sure biomass or height from a 25m footprint would be representative for measuring finer 5m features.

23. **Question: Which technical or methodological aspects should be considered to ensure the proper matching of field data with GEDI footprints during the calibration process?**
    Answer: The 10m uncertainty of GEDI footprints represents an important challenge when matching waveforms to plots. Bullock et al. (https://iopscience.iop.org/article/10.1088/1748-9326/acdf03/meta) provide specific recommendations for choosing the right waveform for an existing plot, but I am not aware of methods for ensuring exact co-location. This is why the L4A product was calibrated with GEDI

waveforms simulated from airborne data known to be a good spatial match with inventory data.

24. **Question: What are the limitations of GEDI when mapping dense rainforests or urban areas?**

    Answer: In dense rainforests, the ground peak of GEDI's waveform may not be distinct. In some cases, people have only used measurements from the "strong" (undithered) beams; other use only nighttime acquisitions which are less contaminated by stray light. If the ground is missed, heights may systematically be too short. Buildings can in some cases create waveforms that cannot be differentiated from trees. The L4B product uses an urban mask to remove as many spurious waveforms as as possible, but that mask is not perfect.

25. **Question: The resolution is already very impressive, but compared to Landsat's high resolution, is it possible for this model to achieve a similar Landsat-like spatial resolution?**

    Answer: The spatial dimension of a GEDI footprint is very similar to a Landsat pixel. The L4B estimates of mean biomass are at 1km because the fine-scale footprint-level measurements are used like sample plots to predict a population total with clear sample-based uncertainties.

26. **Question: If one were to apply the same statistical model in Patterson et al. 2019 to produce a mean 1km biomass per year using each year's L4A waveforms, would that be an accurate representation of change?**

    Answer: If a 1km cell had an adequate GEDI sample in every year (very few cells have this many GEDI "looks"), one could compare the semi-independent Patterson et al. estimates to get an estimate of change. That would require addressing covariance due to the dependency that the same biomass model is being applied in each time period. The method that OBIWAN uses, built around Saarela et al's estimation methods, is probably a more robust method. However, the next session will demonstrate that some method of independent calibration is required when Landsat's optical imagery is part of the process (as it is with Saarela et al.).

27. **Question: How can GEDI data improve global climate models and environmental policy decisions?**

Answer: Some models, such as Maryland's Ecosystem Demography model, learn carbon storage dynamics from an accurate representation of how tree size and biomass storage is distributed across the landscape. This model has directly ingested GEDI data and has been used to assess future mitigation benefits under different scenarios.

28. **Question: What type of variables are used in the Landsat time series model as predictors? Are these annual mosaics or also include multiannual trends, seasonality, etc?**

    Answer: The Landsat time series features are parameters from functions fitted to all clear imagery (all seasons, all years) through the CCDC algorithm (Zhu and Woodcock, 2014). CCDC outputs are stored on Google Earth Engine (Gorelick et al., 2023). We also include synthetic Landsat reflectance inputs from CCDC for two points in the year.

29. **Question: What are the minimum height detection limits - how reliably can GEDI estimate biomass for grasses?**

    Answer: Anecdotally, we see systematically different heights for grasslands at different times of the year, which makes sense. GEDI is definitely sensitive to grass cover, but I do not know what the height detection limits are.

30. **Question: I was wondering, since GEDI doesn't capture data everywhere and only follows specific tracks, do researchers use predictive models to estimate biomass for areas where there is no direct GEDI data available?**

    Answer: The fusion with Landsat time series described in the presentation represents an effort to use predictive models to understand what is happening between GEDI tracks.

31. **Question: How will RH100 and RH50 change after: 1) A Controlled Burn, and 2) A Forest/Wild Fire?**

    Answer: Assuming we had exactly aligned GEDI waveforms (a big assumption), one might generally assume: 1) no systematic change of canopy top height unless the very tops of trees were consumed; 2) some increase of lower RH values (like RH50) in cases where only understory was consumed; 3) a highly uncertain effect following a wildfire, as impacts on canopy profile can be pretty variable. Case #2 would occur as the controlled burn shifted the balance of canopy material upward. BTW,

many people avoid using RH100 (instead using RH98) because the 100th percentile can be noisy (birds, random chance of hitting the very top of pointy trees, etc).

32. **Question: What are your thoughts on transferability of biomass model configured calibration/validation beyond proposed study area?**

    Answer: This preprint (https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5236489) by Ståhl et al. suggests that extending a model calibrated in one location to another very different location (a different continent, for example) can lead to large bias that are not addressed by the estimators used by GEDI and OBIWAN. This is why the mission takes every possible measure to use models that are calibrated as locally as possible.

33. **Question: How do you see the use of ICESat-2 tracks to improve terrain correction for GEDI-derived canopy height and structure estimates? Or are terrain corrections already sufficiently incorporated into GEDI footprints, making additional correction unnecessary, which corrections do you recommend taking into account?**

    Answer: We do not perform terrain-correction for GEDI shots. Variable terrain within the footprint can indeed cause errors in derived tree height products. This is why the GEDI mission moved to higher-resolution footprints than those used in the ICESat-1/GLAS mission. Small footprints minimize the effects of terrain.

34. **Question: Could GEDI data provide information on crop production? I suppose that if we correlate field data with GEDI data and monitoring biomasa dinámica, we could have a tool to estimate crop production. Is my assumption correct?**

    Answer: As mentioned in the response to Q29, GEDI does seem to be sensitive to the height of ground cover. I suspect that latency issues and the sparseness of GEDI's sample would make it not as good as other remote platforms for this kind of application, however.

35. **Question: What are the basic assumptions that historical Biomass can be estimated using Historical Landsat and GEDI data, while Landsat usually does not contain any tree height and biomass information?**

Answer: Validation presented in the next session will confirm your suspicion that height and biomass signal in Landsat-models is minimal. However, we have seen that “correction” of those models with inventory data and/or GEDI shots independent of the modeling process can produce large-area estimates of height/biomass change that are both accurate and precise. This will be discussed next week.