Global Ecosystem Dynamics Investigation (GEDI) Level 1B User Guide

For GEDI L1B Data

Version 2.0

April 2021

Science Team

Principal Investigator: Ralph Dubayah[2]

Deputy Principal Investigator and Instrument Scientist: J. Bryan Blair[3]

User Guide written by Jared Beck[1] in collaboration with Bradford Wirt[1], Scott B. Luthcke[3], Michelle Hofton[2], and John Armston[2].

[1] KBR, Inc., contractor to the U.S. Geological Survey, Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota, USA. Work performed under USGS contract G15PC00012 for LP DAAC[4].

[2] Department of Geographical Sciences, University of Maryland College Park, MD 20742.

[3] Code 61A, NASA Goddard Space Flight Center, Greenbelt, MD 20771.

[4] LP DAAC work performed under NASA contract NNG14HH33I.

Document History

Document Version	Publication Date	Description
1.0	January 2020	Original For P003 L1B
2.0	April 2021	Update for GEDI V2 Release
2.1	May 2021	Addition of Known Issues

Table of Contents

List of Tables

1 Dataset Overview

The overall goal of the Global Ecosystem Dynamics Investigation (GEDI) mission is to advance our ability to characterize the effects of changing climate and land use on ecosystem structure and dynamics. The GEDI instrument produces high-resolution laser ranging observations of the 3-dimensional structure of the Earth. GEDI is attached to the International Space Station (ISS) and collects data globally between 51.6º N and 51.6º S latitudes. GEDI measurements of forest canopy height, canopy vertical structure, and surface elevation are used to characterize important carbon and water cycling processes, biodiversity, and habitat. The GEDI data also may be used in far ranging applications such as weather forecasting, forest management, snow and glacier monitoring, and digital elevation models. GEDI is led by the University of Maryland in collaboration with NASA Goddard Space Flight Center. GEDI science data algorithms and products are created by the GEDI Science Team.

This user's guide provides information about the Version 2 L1B geolocated waveform product. This product is provided in HDF5 file format. The filename convention details are provided in section 2.0 of this document.

1.1 Background

The GEDI instrument consists of three lasers that are split and beam dithered resulting in a total of eight beam ground transects that are spaced 600 meters (m) apart on the Earth's surface in the cross-track direction. Each beam consists of ~25 m footprint samples approximately spaced every 60 m along track. Captured in real time by the GEDI instrument are the shapes of the transmitted and reflected laser waveforms, enabling a precise range to the reflecting surface to be calculated for every shot (after accounting for the speed of light). The laser receive waveform represents the vertical distribution of intercepted surfaces as a function of range. Over non-complex, flat terrain the laser receive waveform shape will look similar to the shape of the outgoing (transmitted) laser pulse. Over complex or sloped terrain, photons may be reflected from multiple surfaces within the footprint resulting in a receive waveform with multiple modes. Interpretation of the laser pulse in post-processing enables precise information on surface elevation, structure, and relative heights to be derived. The L1B geolocated waveform product contains precise geolocation information for the first and last sample bins of each GEDI laser return waveform. The L2 geolocated elevation and height products contain precise geolocation for each reflecting surface within the receive waveform.

For a detailed description of the satellite and reference data, processing schemes, approaches, methods, and other information, refer to the algorithm theoretical basis documents (ATBDs) for geolocation and waveform processing and the GEDI mission website.

GEDI ATBD for GEDI Waveform Geolocation for L1 and L2 Products, Luthcke et al.

GEDI ATBD for Waveform Processing for L1 and L2 Products, Hofton et al.

1.2 Important Product Notes for Version 2 L1B

The L1B product contains all geolocated return waveforms. The L1B product provides corrected geolocated waveform returns, including transmit and receive housekeeping and relevant instrument parameters, as well as geolocation parameters and geophysical corrections. At the processing level of the L1B product, the waveform returns have not been filtered nor has noise characterization been conducted to determine if a valid land surface return exists.

GEDI Version 2 data product improvements include improved geolocation accuracy, updates to the metadata to provide spatial coordinates that allow querying in NASA Earthdata Search, and a reduction in granule size from one full ISS orbit to four segments per orbit. A full list of changes and filename convention details are provided in section 7 of this document. The Version 2 filenames have also been updated to include segment number and version number. The filename convention details are provided in section 2.1.4 of this document.

2 Dataset Characteristics

Global Ecosystem Dynamics Investigation Geolocated Waveforms (GEDI_01B) collection, granule, and dataset characteristics are described below.

2.1 GEDI Geolocated Waveforms Product

2.1.1 Collection Level

Table 1. Collection Level Data

Characteristic	Description
Collection	GEDI
Short name	GEDI01_B
DOI	10.5067/GEDI/GEDI01_B.002
Temporal Resolution	Varies

Temporal Extent	2019-04-18 – Present
Spatial Extent	Global (51.6 °S to 51.6 °N)
Coordinate System	Geographic (lat/lon)
Datum	WGS84
Geographic Dimensions	4.2 km cross-track by one fourth of an ISS orbit along-track
File size	~2 GB
File Format	HDF5

2.1.2 Granule Level

Table 2. Granule Level Data

Characteristic	Description
Number of Science Dataset (SDS) Layers	688 (86 x 8 beams)
Columns/Row	Variable
Pixel Size	25 m (footprint diameter)

2.1.3 Data Layer Characteristics

Please refer to the Level 1B data dictionary for a complete description of the product attributes, groups, and datasets, including datatype, dimensions, units, source, and short definition/description.

Each shot has a unique shot identifier (shot number) that is available within each data group of the product. The shot number is important to retain in any data sub-setting as it will allow the user to link any shot record back to the original orbit data, and to link any shot and its data between the L1 and L2 products. The shot number has been updated for Version 2 and is now in the following format:

OOOOOBBRRGNNNNNNNN

where:

OOOOO: Orbit number

BB: Beam number

R: Reserved for future use

G: Sub-orbit granule number

NNNNNNNN: Shot index

If a data packet is dropped (never received on the ground), NNNNNNNN will not save space for it; however, if we receive a packet with a bad Cyclic Redundancy Check (CRC) error detection, it will be “skipped” in NNNNNNNN in case it can be corrected in later processing.

2.1.4 Filename Convention

GEDI filenames (i.e., the local granule ID) follow a naming convention which provides useful information regarding the specific product.

In this example of a level 01B product, the filename
GEDI01_B_2019108002012_O01959_01_T03909_02_005_02_V002.h5

indicates:

- **GEDI01_B** = Product Short Name
- **2019108** = Julian Date of Acquisition in YYYYDDD
- **002012** = Hours, Minutes and Seconds of Acquisition (HHMMSS)
- **O01959** = O = Orbit, 01959 = Orbit Number
- **01** = Sub-Orbit Granule Number
- **T03909** = T = Track, **03909**= Reference Ground Track Number
- **02** = Positioning and Pointing Determination System (PPDS) type (00 is predict, 01 rapid, 02 and higher is final.)
- **005** = PGEVersion = SDPS Release Number
- **02** = Granule Production Generation Number
- **V002 =** Version Number

2.1.5 File Format

The GEDI products are stored in Hierarchical Data Format 5 (HDF5), a self-descriptive data file format designed by the National Center for Supercomputing Applications to assist users in the storage and manipulation of scientific data across diverse operating systems and machines.

3 Algorithm Description

3.1 Geolocation Algorithm Overview

The GEDI geolocation algorithm uses the instrument pointing, position, and observed range to compute the planet referenced location of the first and last samples of each laser receive waveform.

Instrument positioning and pointing information are derived from GPS and star tracker sensors located on the instrument. During post-processing, this information is combined with the laser range and various measurement model parameters to geolocate the laser footprint. Atmospheric path delays are also accounted for at this stage. The location of the laser footprint is derived in Cartesian coordinates in an Earth Centered Inertial (ECI) frame and then transformed to Earth Centered Fixed (ECF) geodetic coordinates of latitude, longitude, and height (elevation) above the reference ellipsoid, which are available in the data product.

The user is highly encouraged to refer to the GEDI ATBD for GEDI Waveform Geolocation for L1 and L2 Products for a complete description of the geolocation algorithm and definitions, and details of the geolocation and geophysical corrections parameters. In particular, Sections 3.1, 3.6, and 3.7 provide a more focused understanding of the L1B geolocation and geophysical correction parameter groups.

3.2 Waveform Algorithm Overview

A digitally recorded return laser pulse, or waveform, represents the time history of the laser pulse as it interacts with the reflecting surfaces. The waveform can have a simple (single-mode) shape similar to that of the outgoing pulse or be complex and multimodal with each mode representing a reflection from an apparently distinct surface within the laser footprint. Simple waveforms are typical in ocean or bare-ground regions and complex waveforms are typical in rough terrain or vegetated regions. The first and last modes (i.e. detected signal above noise) within the waveform are associated with the highest and lowest perceived reflecting surfaces within the footprint, respectively. The geolocation to the various surface ranging points are available in the L2 products.

The Rx waveform in the L1B data is the downlinked waveform smoothed to correct for noise bias. The waveform processing algorithms used by GEDI are described in section 1.4 of the GEDI ATBD for Waveform Processing for L1 and L2 Products and are adapted from methods developed for the analysis of waveforms acquired from NASA's Land, Vegetation and Ice Sensor (LVIS) (Blair et al., 1999).

4 Metadata

GEDI products have two sources of metadata: the embedded HDF5 metadata and the external ECS metadata. The HDF metadata contains valuable information including global attributes and dataset specific attributes pertaining to the granule. The ECS (generated by the EOSDIS Core System) .met file is the external metadata file in XML format, which is delivered to the user along with the GEDI product. It provides a subset of the HDF metadata.

The Data Set attributes contain specific SDS information such as the data range and applicable scaling factors for the data. The LP DAAC data products page provides these details within a concise document for each of the products. An HDF5 file also contains core metadata essential for search services. It is difficult for a standard HDF call to interpret HDF5 geolocation or temporal information without further knowledge of the file structure.

5 Frequently Asked Questions

How do I cite the data?

The following citation information is applicable for the GEDI01_B.002 product:

Dubayah, R., Luthcke, S., J. B. Blair, Hofton, M., Armston, J., Tang, H. (2021). *GEDI L1B Geolocated Waveform Data Global Footprint Level V002* [Data set]. NASA EOSDIS Land Processes DAAC. Accessed YYYY-MM-DD from https://doi.org/10.5067/GEDI/GEDI01_B.002.

What science datasets does the Global Ecosystem Dynamics Investigation Geolocated Waveforms product contain?

The L1B product provides corrected geolocated waveform returns, including transmit and receive housekeeping and relevant instrument parameters, as well as geolocation parameters and geophysical corrections.

The detailed product contents are defined in the GEDI L1B Product Data Dictionary.

How do I visualize the GEDI01_B waveforms?

The LP DAAC has created a tutorial to explain how to visualize the GEDI Level 01B waveforms. The tutorial can be found at the LP DAAC website under E-learning Materials.

Why is the uneven spacing between beams across track?

While the design of the instrument and its operational implementation are to provide even cross-track spacing of the beams, several factors cause the beams to have small differences in the cross-track spacing. These factors include slight variations in beam alignment in the instrument frame, changes in altitude of the ISS, and changes in ISS attitude orientation, specifically yaw.

6 Quality and Important Notes

The Level 1B product is intended for users who are interested in analyzing the entire return waveform themselves. Most users should start with the Level 2 data files. The Level 1B geolocated waveforms provide context and source data for the GEDI Level 2 products that will allow the expert user to replicate or derive new geolocated Level 2 data products using different algorithms and methodologies for ground and feature finding.

There are numerous time periods where the geolocation performance suffers due to non-optimal operating conditions. These situations include blinding or glinting of one or more instrument star trackers. The "degrade" dataset in the "geolocation" group provides an indication of when there is degraded geolocation. The "degrade" flag is a best estimate of these time periods and data could be performing better or worse in the surrounding time periods near the start and stop of the "degrade" flagged intervals. The "degrade" flag should be understood as a general indicator of a potential issue.

The first (elevation_bin0) and last (elevation_lastbin) return in the waveform are geolocated. To compute the geolocation of any point in the waveform (e.g. ground return), the user would interpolate the first and last return geolocation to the waveform location of interest. This is done in the L2 products where the geolocation of the ground return and other waveform metrics are provided.

The L1B product geolocation group contains several "error" datasets (i.e. latitude_lastbin_error). These formal propagated errors have not yet been fully calibrated. Their relative magnitude is a good indicator of relative geolocation quality, but their absolute magnitudes should not be interpreted as the actual geolocation error at this time. Future releases will contain calibrated geolocation "error" datasets that then can be used for an estimate of absolute error.

The user is encouraged to read section 3.7 of the ATBD for GEDI Waveform Geolocation for L1 and L2 Products to fully understand the application of the geophysical corrections. Of special note: **The ellipsoid height of bounce points within the /geolocation group (*elevation_bin0* and *elevation_lastbin*) has been corrected for solid earth tides, ocean loading, solid earth pole tide, and ocean pole tide.** The bounce points are NOT corrected for ocean tides and dynamic atmospheric correction. These corrections are applied by subtracting the corrections from the bounce point elevations. To remove the corrections already applied (restore the geophysical signal of interest), corrections need to be added to the bounce point elevations.

For complete and updated information regarding product quality, see the GEDI Mission Website.

7 Changes from Previous Version

Version 2 includes the following changes:

- Spatial metadata allows visualizing orbit tracks and spatial querying in NASA Earthdata Search.

- Browse images now available in NASA Earthdata Search.

- Improved geolocation: V002 mean 1-sigma horizontal geolocation error is 10.2 m with 95% of the mission weeks having 1-sigma geolocation error less than 13.2 m. In comparison V001 mean 1-sigma horizontal geolocation error is 23.8 m with 95% of the mission weeks having 1-sigma geolocation error less than 27.2 m (from the analysis of Mission Weeks 19 through 50).

- Added elevation from the SRTM digital elevation model for comparison.

- Modified shot number to include the sub-orbit granule number.

- Removed shots when the laser is not firing.

- Modified file name.

  Each orbit is divided into four sub-orbit granules, computed based on time.

  - From the orbit start time (the time on the filename), each 23.17-minute section is grouped.

  - If shots occur after 92.68 (23.17*4) minutes, they are assigned to sub-orbit granule 4.

  - Major frames are kept together, so if a major frame spans the boundary between granules all shots in that major frame are assigned to the earlier granule.

  - Individual sub-orbit granule files may be missing for an orbit if there is insufficient data in the sub-orbit granule.

  Group: /BEAMXXXX

  - Changed /BEAMXXXX/shot_number to include sub-granule number and remove minor frame number.

Group: /BEAMXXXX/geolocation

- Added dataset: /BEAMXXXX/geolocation/digital_elevation_model_srtm
- Changed /BEAMXXXX/geolocation/shot_number to include sub-granule number and remove minor frame number.
- Changed /BEAMXXXX/geolocation/degrade from a binary degraded (1) or not degraded (0) value to include specific flags for each degraded condition: Non-zero values indicate the shot occurred during a degraded period. A non-zero tens digit indicates degraded attitude, a non-zero ones digit indicates a degraded trajectory. Details are in the table below.

Table 4. Degrade Conditions

Flag	Degrade Condition
3X	ADF CHU solution unavailable (ST-2)
4X	Platform attitude
5X	Poor solution (filter covariance large)
6X	Data outage (platform attitude gap also)
7X	ST 1+2 unavailable (similar boresight FOV)
8X	ST 1+2+3 unavailable
9X	ST 1+2+3 and ISS unavailable
X1	Maneuver
X2	GPS data gap
X3	ST blinding
X4	Other
X5	GPS receiver clock drift
X6	Maneuver & GPS receiver clock drift
X7	GPS data gap & GPS receiver clock drift
X8	ST blinding & GPS receiver clock drift
X9	Other & GPS receiver clock drift

8 Known Issues

The V002 horizontal geolocation error has been significantly improved through a comprehensive calibration of the laser relative beam alignment, the time varying laser frame alignment, as well as improved estimates of range biases. The current V002 mean 1-sigma horizontal geolocation error is 10.2 m with 95% of the mission weeks having 1-sigma geolocation error less than 13.2 m. While the V002 horizontal geolocation error has been improved by over a factor of 2 from V001, analysis shows further geolocation improvement should be expected in future releases.

The range/elevation performance has been assessed through direct altimetry to known surfaces, comparisons to independent data, and crossover analysis. The fundamental instrument ranging precision is on the order of 3 cm. Improved laser channel dependent range bias estimates and laser ranging tracking point modeling have corrected the ~ -56 cm range/elevation bias observed in release V001. Release V002 elevation errors are now dominated by the contribution of the horizontal error on sloping surfaces (laser incidence angle relative to the surface). For example, the 10.2 m 1-sigma horizontal geolocation error induces a 17.8 cm elevation error on a 1 degree slope (laser surface incidence angle). In addition, the radial positioning of the GEDI instrument contributes ~ 7 cm RMS to the elevation error with excursions past 20 cm over portions of some orbits. It is expected that the radial positioning of the instrument will be further improved in future releases.

The GEDI L1 and L2 data may contain orbit segments that have elevation errors not identified by the "degrade" flag. It is the intention of the GEDI project to further identify these issues and provide a list of these segments and orbits. However, the user is strongly encouraged to edit the data using both the "degrade" flag, and a comparison of the footprint elevations to the TDX and SRTM elevations provided in the data record for each GEDI footprint. As an example, the following orbit segments with large elevation errors have been identified:

6163 - granule segment over Australia

6245 - granule segment over N America

6147 - granule segment over southern Spain

9 Dataset Access

The GEDI_01B product is available through the LP DAAC Data Pool and NASA Earthdata Search.

10 Contact Information

LP DAAC User Services
U.S. Geological Survey (USGS) Earth Resources Observation and Science (EROS) Center
47914 252nd Street Sioux Falls, SD 57198-0001
Phone Number: 605-594-6116
Toll Free: 866-573-3222 (866-LPE-DAAC)
Email: lpdaac@usgs.gov
Web: https://lpdaac.usgs.gov

GEDI Science Team Contact: Ralph Dubayah (PI), Professor of Geographical Sciences at the University of Maryland: dubayah@umd.edu

11 Citations

Dubayah, R., Luthcke, S., J. B. Blair, Hofton, M., Armston, J., Tang, H. (2020). *GEDI L1B Geolocated Waveform Data Global Footprint Level V002* [Data set]. NASA EOSDIS Land Processes DAAC. Accessed YYYY-MM-DD from https://doi.org/10.5067/GEDI/GEDI01_B.002.

12 Publications, References, and Resources

Blair J.B, Rabine D.L., Hofton M.A, The Laser Vegetation Imaging Sensor: a medium-altitude, digitisation-only, airborne laser altimeter for mapping vegetation and topography. *Isprs J. Photogramm. Remote Sens.* 54, 115–122 (1999).

Dubayah, R. O., Blair, J. B., Goetz, S. J., Fatoyinbo, L., Hansen, M. C., Healey, S. P., Hofton, M., Hurtt, G., Kellner, J. R., Luthcke, S., Armston, J., Tang, H., Duncanson, L., Hancock, S., Jantz, P., Marselis, S., Patterson, P., Qi, W., Silva, C. (2020). The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth's forests and topography. Science of Remote Sensing. https://www.sciencedirect.com/science/article/pii/S2666017220300018

Luthcke et al. (2020). ATBD for GEDI Waveform Geolocation for L1 and L2 Products

Hofton et al. (2020). ATBD for Waveform Processing for L1 and L2 Products

12.1 Project Website

The GEDI Mission maintains a website at https://gedi.umd.edu/.

12.2 Additional Resources

The GEDI L1B data dictionary can be found at: https://lpdaac.usgs.gov/documents/981/gedi_l1b_dictionary_P003_v2.html