NASA G-LiHT Yucatan LiDAR Reprocessing — Dataset README

Version: 1.1 (October 2026)
License: CC BY 4.0
Repository: OpenTopography

NOTE:

This report describes the fully merged data product. However, OpenTopography hosts it as two datasets divided by UTM zone, so values like file count and total area will reflect the combined whole rather than either half. Processing methods and other metadata are applicable to both datasets.

- **G-LiHT Yucatan West 2013: Reprocessed Ground Points, DEMs and Visualizations (https://doi.org/10.5069/G9Z899N1)**
- **G-LiHT Yucatan East 2013: Reprocessed Ground Points, DEMs and Visualizations (https://doi.org/10.5069/G9TH8JX9)**

Project Overview

This dataset contains reprocessed LiDAR ground points, digital elevation models (DEMs) and archaeological visualization composites for 456 NASA Goddard's LiDAR, Hyperspectral and Thermal (G-LiHT) airborne transects. The transects were acquired over the Yucatan Peninsula, Mexico, in April–May 2013.

The original NASA G-LiHT acquisitions used automated ground classification, which misclassified many ground returns beneath dense tropical forest canopy. This dataset applies the NCALM ground classification methodology (Fernandez-Diaz et al.; Estrada-Belli et al. 2025) to recover bare-earth surfaces suitable for Maya archaeological feature detection and landscape analysis.

NASA's ground-filtered DTMs for these transects were published at 1 m per pixel, although the ground returns are dense enough to support finer grids. This dataset therefore rasterizes the reprocessed ground surfaces at 33 cm, 50 cm and 100 cm.

Credits and Acknowledgments

Principal Investigator:
Benjamin Britton, MFA, MA
Department of Geography, University of Cincinnati

Reprocessing methodology:
Bare-earth point cloud classification was performed using TerraScan/TerraBatch (Terrasolid Ltd., academic license); 33 cm, 50 cm, and 100 cm Digital Terrain Models (DTMs) were then generated from the classified points by kriging in Golden Software Surfer. The classification workflow replicates the ground filtering and macro parameters described by Estrada-Belli et al. (2025) (see main text and Supplementary Document #2, Paragraph 2) as utilized in their processing of regional airborne LiDAR datasets. This implementation represents an independent replication of the published methodology applied to the source point cloud data.

Source data:
NASA Goddard's LiDAR, Hyperspectral and Thermal (G-LiHT) Airborne Imager
Original acquisitions: April–May 2013, Yucatan Peninsula, Mexico
NASA G-LiHT Program: https://gliht.gsfc.nasa.gov

Visualization methodology:
G1 composite recipe: Britton et al. 2025 (see citation below)
RVT visualization toolkit: Relief Visualization Toolbox (rvt-py), ZRC SAZU

Funding:
University of Cincinnati Bearcat AI Grant Program (award UC_BearcatAI_2025)
(This dataset was produced in part through support from the Bearcat AI Grant Program, University of Cincinnati. Please cite this program when using or redistributing this dataset.)

Citation

When using this dataset, please cite:

Dataset(s):

Britton, B. (2026). G-LiHT Yucatan West 2013: Reprocessed Ground Points, DEMs and Visualizations. NASA Goddard Space Flight Center G-LiHT (G-LiHT). [Dataset]. Distributed by OpenTopography. https://doi.org/10.5069/G9Z899N1.. Accessed YYYY-MM-DD

Or

Britton, B. (2026). G-LiHT Yucatan East 2013: Reprocessed Ground Points, DEMs and Visualizations. NASA Goddard Space Flight Center G-LiHT (G-LiHT). [Dataset]. Distributed by OpenTopography. https://doi.org/10.5069/G9TH8JX9. Accessed YYYY-MM-DD

Methodology:

> *Britton, B. et al. (2025). [Citation in progress.]*

Source data:

> *Cook, B.D., Corp, L.A., Nelson, R.F., Middleton, E.M., Morton, D.C., McCorkel, J.T., Masek, J.G., Ranson, K.J., Ly, V., Montesano, P.M. (2013). NASA Goddard's LiDAR, Hyperspectral and Thermal (G-LiHT) Airborne Imager. Remote Sensing, 5(8), 4045–4066.*

Funding acknowledgment (required):

> *Supported in part by the University of Cincinnati Bearcat AI Grant Program (award UC_BearcatAI_2025).*

Dataset Contents

The dataset has three parts, delivered as four zip files, plus a checksum manifest and a per-transect metadata table. The DEMs are split by resolution across two zips so that each zip stays under 50 GB.

Part	File	Files	Size
Ground points	`GLiHT_Yucatan_LAZ_v1.zip`	456	7.56 GB
DEMs, 0.33 m	`GLiHT_Yucatan_DEM_0p33m_v1.zip`	456	31.00 GB
DEMs, 0.5 m and 1 m	`GLiHT_Yucatan_DEM_0p5m_1m_v1.zip`	912	17.77 GB
G1 composites	`GLiHT_Yucatan_G1_v1.1.zip`	1,368	23.47 GB
Checksum manifest	`GLiHT_Yucatan_OT_manifest_sha256.csv`	1	< 1 MB
Metadata table	`GLiHT_Yucatan_OT_metadata.csv`	1	< 1 MB

All rasters (DEMs and G1 composites) are Cloud Optimized GeoTIFFs (COG): DEFLATE-compressed (lossless), tiled in 512 × 512 blocks, with internal overviews.

Version 1.1 replaces the G1 zip of version 1.0 (`GLiHT_Yucatan_G1_v1.zip`). The G1 pixel values are unchanged; version 1.1 declares 0 as the no-data value in every G1 file and rebuilds the overviews accordingly (see section 3). It also adds the manifest and the metadata table. The point clouds and DEMs are unchanged from version 1.0.

1. Point Clouds — LAZ/

456 files, 4.40 billion points, 7.56 GB. LAS 1.4, point data format 6, LASzip-compressed (lossless). The coordinate scale is 0.001 m in x, y and z. There is one file per G-LiHT transect.

These files contain a new ground classification of the NASA G-LiHT point clouds, made for this dataset. They are not a subset of NASA's original classes. Every DEM in this dataset was interpolated from the points in these files.

Classification procedure (TerraScan / TerraBatch, Terrasolid Ltd.):

1. All original NASA classifications were discarded: every return was reset to class 1 (unclassified).
2. Ground was classified from scratch with TerraScan's adaptive TIN ground routine, using NCALM-derived parameters (Fernandez-Diaz; Estrada-Belli et al. 2025):

- maximum building size 70 m
- terrain angle 88°
- iteration angle 9°
- iteration distance 3.0 m

A second pass at a 12° iteration angle was also run, but it added no points.

3. Returns lying within ±0.20 m of the classified ground surface were assigned to class 8.
4. Only classes 2 and 8 were written out. All other returns (vegetation, structures, noise and unclassified points) are omitted.

TerraScan wrote the classified points in its own FastBin format. These were then converted to standard LAZ 1.4 with LASzip.

Class	Contents
2	Ground, as classified by the procedure above
8	Near-ground: returns within ±0.20 m of the ground surface that were not classified as ground

What is not included. These files are not complete point clouds. NASA's original, fully classified G-LiHT point clouds remain available from the G-LiHT data portal (https://gliht.gsfc.nasa.gov) and are not redistributed here. Users who need canopy, vegetation-structure or first-return data should obtain the originals there.

Coordinate reference system. Each file carries a compound CRS, WGS 84 / UTM zone 15N or 16N plus EGM96 height (EPSG:32615+5773 or EPSG:32616+5773), as an OGC WKT VLR (record 2112) with the global-encoding WKT bit set, as LAS 1.4 specifies for point format 6. See the Coordinate Reference System section below.

Files by region:

Folder	Files
Yuc_South	103
Yuc_North2	91
Yuc_Centro	87
Yuc_Chiaps1	64
Yuc_Norte	49
Yuc_Out_Yuc	38
Yuc_Campeche	24

Point counts range from 72,679 (`ChiapsCampeche_NFI_l0s472`) to 27.3 million (`Centro_GLAS_l1s465`) per file.

File naming: `{Region}_{Dataset}_{TransectID}_GND_v1.laz`, for example `South_GLAS_l0s395_GND_v1.laz`.

Software: CloudCompare, QGIS, PDAL, LAStools, ArcGIS Pro, Global Mapper, or Python (laspy with the lazrs or laszip backend).

2. Digital Elevation Models — `DEM/`

1,368 files (456 transects × 3 resolutions), 48.76 GB. The files are single-band float32 Cloud Optimized GeoTIFFs, DEFLATE-compressed (lossless, floating-point predictor), tiled in 512 × 512 blocks, with internal overviews.

Resolutions:

Suffix	Nominal cell size	Cell sizes recorded in the files
`_DEM_0p33m_v1.tif`	0.33 m	0.3299–0.3302 m
`_DEM_0p5m_v1.tif`	0.5 m	0.4997–0.5004 m
`_DEM_1m_v1.tif`	1.0 m	0.9993–1.0008 m

Cell sizes are nominal. The exact spacing of each grid differs slightly from transect to transect and is recorded in each file's geotransform. Users mosaicking transects should resample to a common grid.

Interpolation. Each DEM was interpolated from all points in the corresponding LAZ file (classes 2 and 8) by kriging in Golden Software Surfer. Settings not listed here were left at Surfer's defaults.

- 20 m isotropic search radius
- 1 to 64 points per search
- a single search sector

These parameters follow Estrada-Belli et al. 2025 (Supplement).

Edge treatment. Interpolated values along transect edges showed artifacts, so a 21 m strip inside every transect boundary was set to no-data in each DEM. The boundary was detected along both rows and columns, so that edges running in any direction are

treated. In transect `Centro_GLAS_l0s441`, for example, the strip removes about 10% of the interpolated cells. The unmasked surfaces are not distributed.

No-data value: `1.70141e+38` (Surfer's blanking value), declared in each file's GDAL no-data tag. Software that reads that tag treats these cells as empty.

Coverage and values. The 1 m DEMs contain 1,128.9 km² of valid cells. Elevations range from −48.39 m to 1,661.98 m across the dataset.

Vertical reference: EGM96 geoid heights, meters (EPSG:5773). See the Coordinate Reference System section below.

Coordinate reference system. Each file carries a compound CRS, WGS 84 / UTM zone 15N or 16N plus EGM96 height (EPSG:32615+5773 or EPSG:32616+5773), written as EPSG-coded GeoTIFF keys.

File naming: `{Region}_{Dataset}_{TransectID}_DEM_{resolution}_v1.tif`, for example `South_GLAS_l0s395_DEM_0p33m_v1.tif`.

Software: QGIS, ArcGIS Pro, Global Mapper, GDAL, Python (rasterio), R (terra).

3. G1 Archaeological Visualization Composites — `G1/`

The G1 composite is a four-layer blended visualization designed for Maya archaeological feature detection, published in Britton et al. 2025 (Table 3). It combines multiple topographic derivatives to reveal subtle relief, shadow and surface texture at the same time.

Composite recipe (Britton et al. 2025, Table 3):

Layer	Blend Mode	Opacity	Brightness	Contrast	Gamma
Multi-directional Hillshade (base)	Normal	100%	+10	+20	2.0
Slope	Difference	67%	+10	+20	3.0
Positive Openness	Overlay	67%	−30	+55	1.0
Sky View Factor	Multiply	67%	+1	+30	0.5

All layers are min/max stretched before the brightness, contrast and gamma adjustments. Output is uint8 RGB (0–255 per channel), displayed as a grayscale-balanced composite.

There are three resolutions per transect, matching the DEMs above.

Format. 1,368 files (456 transects × 3 resolutions), 23.47 GB. Each file is a three-band uint8 Cloud Optimized GeoTIFF, DEFLATE-compressed (lossless, horizontal predictor), tiled in 512 × 512 blocks, with internal overviews. The three bands hold identical values.

The composite is grayscale, stored as RGB so that tiles can be used directly as training samples by object-detection frameworks that expect three channels.

Edge treatment and no-data. Each G1 composite was masked with the same 21 m edge strip as its DEM, so its valid area matches the DEM's exactly. Masked and outside cells are 0 in all three bands, and 0 is declared as the no-data value on every band. Before masking, any pure-black pixel inside the valid area was raised to 1, so a value of 0 occurs only where there is no data.

File naming: `{Region}_{Dataset}_{TransectID}_G1_{resolution}_v1.tif`, for example `South_GLAS_l0s395_G1_0p33m_v1.tif`.

4. Checksum Manifest and Metadata Table

`GLiHT_Yucatan_OT_manifest_sha256.csv` lists every data file in the four zips (3,192 rows) with its SHA-256 checksum. Columns:

Column	Contents
`relpath`	Path inside the zip, e.g. `DEM/Yuc_South/South_Glas/South_Glas/South_GLAS_l0s395_DEM_0p33m_v1.tif`
`zip`	Name of the zip that holds the file
`bytes`	File size in bytes
`sha256`	SHA-256 checksum (64 hexadecimal characters)
`check`	Agreement with the checksum recorded when the file was produced

`GLiHT_Yucatan_OT_metadata.csv` has one row per transect (456 rows): `stem`, `region`, `dataset`, `transect_id`, `epsg` (compound, e.g. `32616+5773`), `utm_zone`, the bounding box in UTM meters (`bbox_min_x`, `bbox_min_y`, `bbox_max_x`, `bbox_max_y`, from the 0.33 m DEM), `centroid_lon` and `centroid_lat` (WGS 84), the 0.33 m DEM's `dem_0p33m_width_px` and `dem_0p33m_height_px`, the point counts `input_returns` (in NASA's source file), `ground_points` (class 2), `near_ground_points` (class 8), `retained_points` (classes 2 and 8) and `retained_pct`, and the paths inside the zips of the transect's seven files (`laz`, `dem_0p33m`, `dem_0p5m`, `dem_1m`, `g1_0p33m`, `g1_0p5m`, `g1_1m`).

Across the 456 transects, NASA's source files hold 12,404,104,097 returns, of which 4,403,882,767 (35.50%) are retained: 2,595,444,817 ground and 1,808,437,950 near-ground.

Coordinate Reference System

Two UTM zones are used, depending on geographic location:

EPSG	Zone	Central Meridian	LAZ files
EPSG:32615	WGS 84 / UTM Zone 15N	−93°	222
EPSG:32616	WGS 84 / UTM Zone 16N	−87°	234

Each transect uses the zone NASA assigned it.

Vertical datum: EGM96 geoid heights, meters (EPSG:5773). NASA's G-LiHT campaign metadata gives terrain elevations as "m, EGM96 geoid". As a check, the median difference between this dataset's DEMs and NASA's own G-LiHT DTMs on three transects was between 0.00 and 0.13 m.

In the files:

Part	CRS recorded	Form
LAZ	EPSG:326xx + 5773 (compound)	OGC WKT VLR (record 2112)
DEM	EPSG:326xx + 5773 (compound)	EPSG-coded GeoTIFF keys
G1	EPSG:326xx (horizontal)	EPSG-coded GeoTIFF keys

The G1 composites are images, not heights, so they carry the horizontal CRS only. No manual CRS assignment should be needed in standard GIS software.

Transects Not Included

NASA's archive provides 458 unique transects for this campaign. Two are absent from every part of this dataset, because their source data contain too few ground returns to support a surface:

- `Chiaps1_NFI_l0s422`: 1 ground return
- `ChiapsCampeche_NFI_l2s505`: 0 ground returns

File Organization

```
GLiHT_Yucatan_LAZ_v1.zip
└── LAZ/
    └── Yuc_South/South_Glas/South_Glas/South_GLAS_l0s395_GND_v1.laz ...
GLiHT_Yucatan_DEM_0p33m_v1.zip
└── DEM/
    └── Yuc_South/South_Glas/South_Glas/South_GLAS_l0s395_DEM_0p33m_v1.tif ...
GLiHT_Yucatan_DEM_0p5m_1m_v1.zip
└── DEM/
    └── Yuc_South/South_Glas/South_Glas/South_GLAS_l0s395_DEM_{0p5m,1m}_v1.tif
...
GLiHT_Yucatan_G1_v1.1.zip
└── G1/
    └──
Yuc_South/South_Glas/South_Glas/South_GLAS_l0s395_G1_{0p33m,0p5m,1m}_v1.tif ...
GLiHT_Yucatan_OT_manifest_sha256.csv
GLiHT_Yucatan_OT_metadata.csv
```

The region folders are the same in all three parts; the two DEM zips unpack into the same `DEM/` tree. Most regions nest as `Region/Sub/Sub/`. The `Yuc_Campeche/Campeche/` files sit one level higher.

Contact

Benjamin Britton
Department of Geography, University of Cincinnati
benjaminbritton@yahoo.com