

Lidar-Derived Topography and Ecosystem Structure Products, California, 2013-2023

Documentation Revision Date: 2026-10-12

Dataset Version: 1

Summary

This dataset provides statewide lidar-derived topography and ecosystem structure products for California, generated from approximately 40 TB of publicly available airborne lidar point cloud data collected across 118 campaigns as part of the U.S. Geological Survey 3D Elevation Program (3DEP). The dataset includes digital terrain models (DTMs), digital surface models (DSMs), normalized digital surface models (nDSMs)/canopy height models (CHMs), slope, aspect, canopy cover, vegetation height percentiles, first-return height statistics, and supporting spatial metadata including lidar campaign year, building-fraction rasters, etc. All products are available at 1-m, 10-m, and 30-m spatial resolutions. The dataset was processed using open-source R and Python workflows on NASA Advanced Supercomputing (NAS) resources. It is designed to support a wide range of applications, including wildfire risk modeling, biodiversity and habitat monitoring, ecosystem management, urban planning, hydrologic analysis, carbon and biomass assessment, and environmental change analysis across California. The data are provided in GeoTIFF and GeoPackage format.

The dataset contains 2,486 cloud-optimized GeoTIFFs (COGs) and two GeoPackage files.

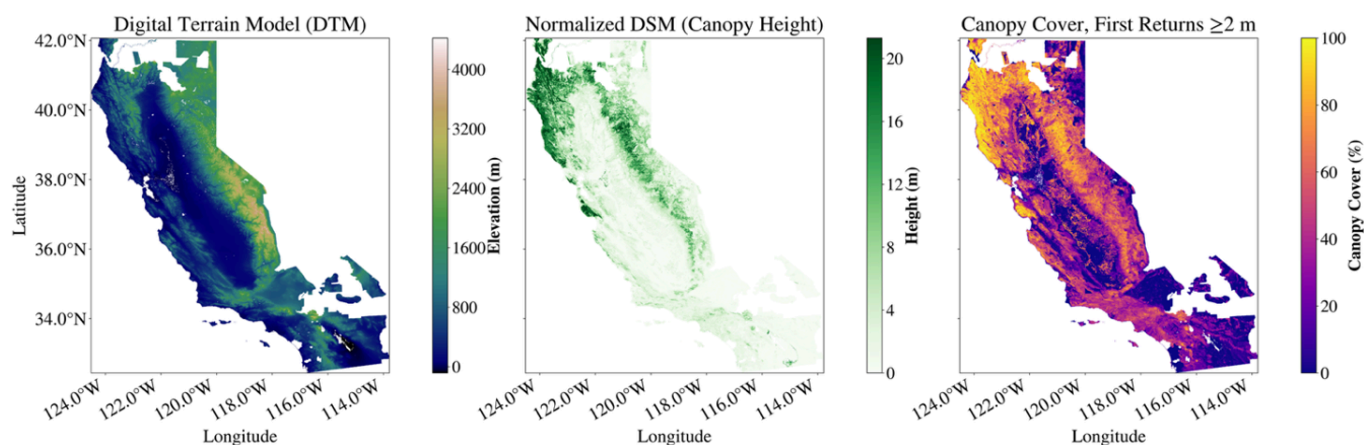


Figure 1. Statewide mosaics of the digital terrain model (left), normalized digital surface model/canopy height model (center), and first-return canopy cover greater than or equal to 2 m (right). Spatial resolution is 30 m.

Citation

Kabasares, K.K., J.T. Benik, I.G. Brosnan, A. Guzman, and T. Park. 2026. Lidar-Derived Topography and Ecosystem Structure Products, California, 2013-2023. ORNL DAAC, Oak Ridge, Tennessee, USA. <https://doi.org/10.3334/ORNLDAAC/2523>

Table of Contents

1. [Dataset Overview](#)
2. [Data Characteristics](#)
3. [Application and Derivation](#)
4. [Quality Assessment](#)
5. [Data Acquisition, Materials, and Methods](#)
6. [References](#)

1. Dataset Overview

This dataset provides a topography and ecosystem structure product suite for California. It was derived from 118 publicly available airborne lidar collections from the USGS 3DEP program (USGS, 2024). The input campaigns were acquired between 2013 and 2023 and collectively cover more than 415,000 km² when overlapping campaign areas are counted. The unique lidar coverage is approximately 331,804 km², or about 78.3% of California land area. The data products describe both the terrain surface and above-ground structure. Terrain products include bare-earth elevation, surface elevation, slope, and aspect. Ecosystem structure products include height above ground (normalized digital surface model or canopy height model), canopy cover in multiple height strata, first-return height percentiles, and first-return distribution statistics. The products are designed to provide immediately usable statewide inputs for wildfire, ecology, land management, urban, hydrologic, and risk-modeling applications.

Project: [Carbon Monitoring System](#)

The NASA Carbon Monitoring System (CMS) program is designed to make significant contributions in characterizing, quantifying, understanding, and predicting the evolution of global carbon sources and sinks through improved monitoring of carbon stocks and fluxes. The System uses NASA satellite

observations and modeling/analysis capabilities to establish the accuracy, quantitative uncertainties, and utility of products for supporting national and international policy, regulatory, and management activities. CMS data products are designed to inform near-term policy development and planning.

Acknowledgements

This research was funded by the NASA Carbon Monitoring System (CMS) (grant 80NSSC21K0964), the NASA Earth eXchange (NEX) Program (grant 80NSSC23M0230), and the NASA Wildfire, Ecosystem Resilience, and Risk Assessment Initiative (WERK) program (grant NASA SAA2-403779).

NASA Earth Exchange (NEX) is a collaborative supercomputing and data analytics platform at NASA Ames that enables researchers to analyze large-scale Earth observation datasets using high-performance computing, advanced modeling, and data science tools. NEX brings together NASA data, computational resources, and scientific expertise to accelerate research on Earth system processes, natural resources, ecosystems, wildfire, and other environmental challenges at regional to global scales.

WERK (Wildfire, Ecosystem Resilience, and Risk Assessment Initiative) is a NASA-led initiative that works with California state agencies to develop actionable, updatable, and freely accessible Earth observation-based data products for monitoring natural and working lands. Through interagency collaboration, remote sensing, and data innovation, WERK integrates diverse datasets to support wildfire resilience, ecosystem monitoring, and natural resource management across California.

2. Data Characteristics

Spatial Coverage: California, USA; limited to USGS 3DEP lidar extent (331,804 km², or about 78.3% of California land area)

Spatial Resolution: 1 m, 10 m, and 30 m

Temporal Coverage: 3DEP lidar acquired between 2013 and 2023; the most recent available acquisition is used to generate the final statewide data products.

Temporal Resolution: Epoch (2013-2023)

Study Area: Latitude and longitude are given in decimal degrees.

Site	Westernmost Longitude	Easternmost Longitude	Northernmost Latitude	Southernmost Latitude
California	-125.5829	-112.4823	42.4875	31.6057

Data File Information

The dataset contains 2,486 cloud-optimized GeoTIFFs (COGs) and two GeoPackage files. The COGs are organized by product resolution (1 m, 10 m, and 30 m) and type. The 1-m products are distributed as 30-km x 30-km tiled to increase accessibility. The 10-m and 30-m products are distributed as statewide single-band COG mosaics for each variable, with supporting acquisition year, building fraction, and spatial metadata products.

Table 1. Resolution-level product summary.

Type	1 m Products	10 m Products	30 m Products
Elevation and terrain	DTM, DSM, slope, aspect	DTM, DSM, slope, aspect	DTM, DSM, slope, aspect
Ecosystem structure	nDSM/CHM only	nDSM/CHM; first-return standard deviation, skewness, and kurtosis; first-return height percentiles P10, P25, P50, P75, P90, P95, P99; canopy cover for all returns and first returns in the >=2 m, 2-4 m, and 4-8 m height strata	nDSM/CHM; first-return standard deviation, skewness, and kurtosis; first-return height percentiles P10, P25, P50, P75, P90, P95, P99; canopy cover for all returns and first returns in the >=2 m, 2-4 m, and 4-8 m height strata
Ancillary data	Tiling system GeoPackage	Campaign boundary GeoPackage; Pixel-level lidar acquisition year raster; building-fraction raster	Campaign boundary GeoPackage; Pixel-level lidar acquisition year raster; building-fraction raster

Acronyms: digital terrain model (DTM), digital surface models (DSM), normalized digital surface model (nDSM), and canopy height model (CHM).

GeoTIFF characteristics

- Projected coordinate system: California Albers (EPSG:3310); NAD83 horizontal datum; units = meters.
- Vertical reference: Digital terrain products were corrected to GEOID18 where source campaign metadata enabled geoid correction.
- Nodata value: -9999
- Data type: Float32
- Number of bands: one

1-m GeoTIFF products

The 1-m products are distributed as 30-km x 30-km sub-tile COGs (Figure 2).

The file-naming convention is *WERK00.<#>.LTE.<Variable Code>.1m.E2013.2023.<hXXvXXsXX>.010.tif*, where

- *WERK00* is the unique product group identifier for the WERK product suite.
- *<#>* indicates the product identifier within the WERK00 product group; 1 denotes the 1 m product collection.
- *LTE* is the acronym for LiDAR-derived Topography and Ecosystem Structure Data.
- *<Variable Code>* identifies the specific variable defined in Table 2.

- "1m" indicates the spatial resolution of the product.
- E2013.2023 indicates that the product incorporates LiDAR data acquired between 2013 and 2023.
- <hXXvXXsXX> is the unique tile identifier, consisting of the horizontal (h), vertical (v), and sub-tile (s) IDs. See Figure 3 for detailed tiling information.
- 010 is the three-digit product version identifier, corresponding to Version 1.0.

Example file name: *WERK00.1.LTE.SLP.1m.E2013.2023.h08v08s12.010.tif*

Table 2. Variables in the 1-m GeoTIFF files

Band	Variable Code	Variable	Units	Description
1	DTM	Digital terrain model	m	Bare-earth elevation surface.
2	DSM	Digital surface model	m	Elevation of the upper surface, including vegetation, buildings, and other above-ground features.
3	nDSM / CHM	Normalized digital surface model / canopy height model	m	Normalized digital surface model, computed as DSM - DTM. In vegetated areas this is interpreted as canopy height above ground.
4	SLP	Slope	degrees	Terrain slope derived from the corrected DTM.
5	ASP	Aspect	degrees	Downslope compass direction derived from the corrected DTM.

10-m and 30-m Statewide Mosaic Products

For statewide mosaics (10 m and 30 m), filenames encode resolution, product family, variable, temporal range, and statewide extent. The file-naming convention is *WERK00.<#>.LTE.<Variable Code>.<RES>.E2013.2023.state.010.tif*, where

- <#> indicates the product identifier within the WERK00 product group; 2 and 3 denotes the 10 m and 30 m product collection, respectively.
- <Variable Code> identifies the specific variable in Table 3.
- <RES> indicates the spatial resolution of the product ("10m" or "30m").
- state indicates that the file represents a statewide mosaic covering California.

Example file names: *WERK00.3.LTE.DSM.30m.E2013.2023.state.010.tif*, *WERK00.2.LTE.P25.10m.E2013.2023.state.010.tif*

Table 3. Variables in the 10-m and 30-m GeoTIFF files (mosaics).

Variable Code	Variable	Units	Description
DTM	Digital terrain model	m	Bare-earth elevation surface.
DSM	Digital surface model	m	Upper-surface elevation including vegetation, buildings, and other above-ground returns.
nDSM	Normalized digital surface model / canopy height model	m	Height above ground, computed as DSM - DTM.
SLP	Slope	degrees	Terrain steepness derived from DTM.
ASP	Aspect	degrees	Downslope compass direction derived from DTM.
STD	First-return height standard deviation	m	Standard deviation of normalized first-return heights within each grid cell.
SKE	First-return height skewness	unitless	Skewness of normalized first-return height distribution.
KUR	First-return height kurtosis	unitless	Kurtosis of normalized first-return height distribution.
P10, P25, P50, P75, P90, P95, P99	First-return height percentiles	m	Selected percentiles of normalized first-return heights. P50 represents median first-return height.
CC1	Canopy cover, all returns 2-4 m	percent	Percent of all valid vegetation-related returns between 2 m and 4 m above ground.
CC2	Canopy cover, all returns 4-8 m	percent	Percent of all valid vegetation-related returns between 4 m and 8 m above ground.
CC3	Canopy cover, all returns >= 2 m	percent	Percent of all valid vegetation-related returns greater than or equal to 2 m above ground.
CC4	Canopy cover, first returns 2-4 m	percent	Percent of first returns between 2 m and 4 m above ground.
CC5	Canopy cover, first returns 4-8 m	percent	Percent of first returns between 4 m and 8 m above ground.
CC6	Canopy cover, first returns >=2 m	percent	Percent of first returns greater than or equal to 2 m above ground.

Note: "Returns" are the discrete laser measurements recorded by an airborne lidar sensor after emitted light interacts with ground, vegetation, buildings, or other objects. "First returns" are the first detected returns from each outgoing pulse and are often associated with the uppermost intercepted surface.

Supplementary Spatial Metadata (two GeoPackages and four COGs)

The GeoPackage *WERK00.LTE.CampaignInfo.gpkg* provides a subset of USGS 3DEP Work Unit Extent Spatial Metadata for the California campaigns used in these products.

The GeoPackage *WERK.L2.30km.TilingSystem.gpkg* provides the systematic tiling scheme (H: horizontal ID; V: vertical ID; S: sub-tile ID) used to produce 1-m lidar topography and ecosystem structure data across the state of California. Each tile covers a 30 km × 30 km extent (Figure 3).

Lidar acquisition year maps at 10 m (*WERK00.2.LTE.YEAR.10m.E2013.2023.state.010.tif*) and 30 m (*WERK00.3.LTE.YEAR.30m.E2013.2023.state.010.tif*) provide the source acquisition year (YYYY) for each statewide mosaic pixel.

Building-fraction rasters at 10 m (*WERK00.2.LTE.FBF.10m.E2013.2023.state.010.tif*) and 30 m (*WERK00.3.LTE.FBF.30m.E2013.2023.state.010.tif*) provide the fraction (0 to 1.0) of each pixel intersecting Microsoft Building Footprint polygons.

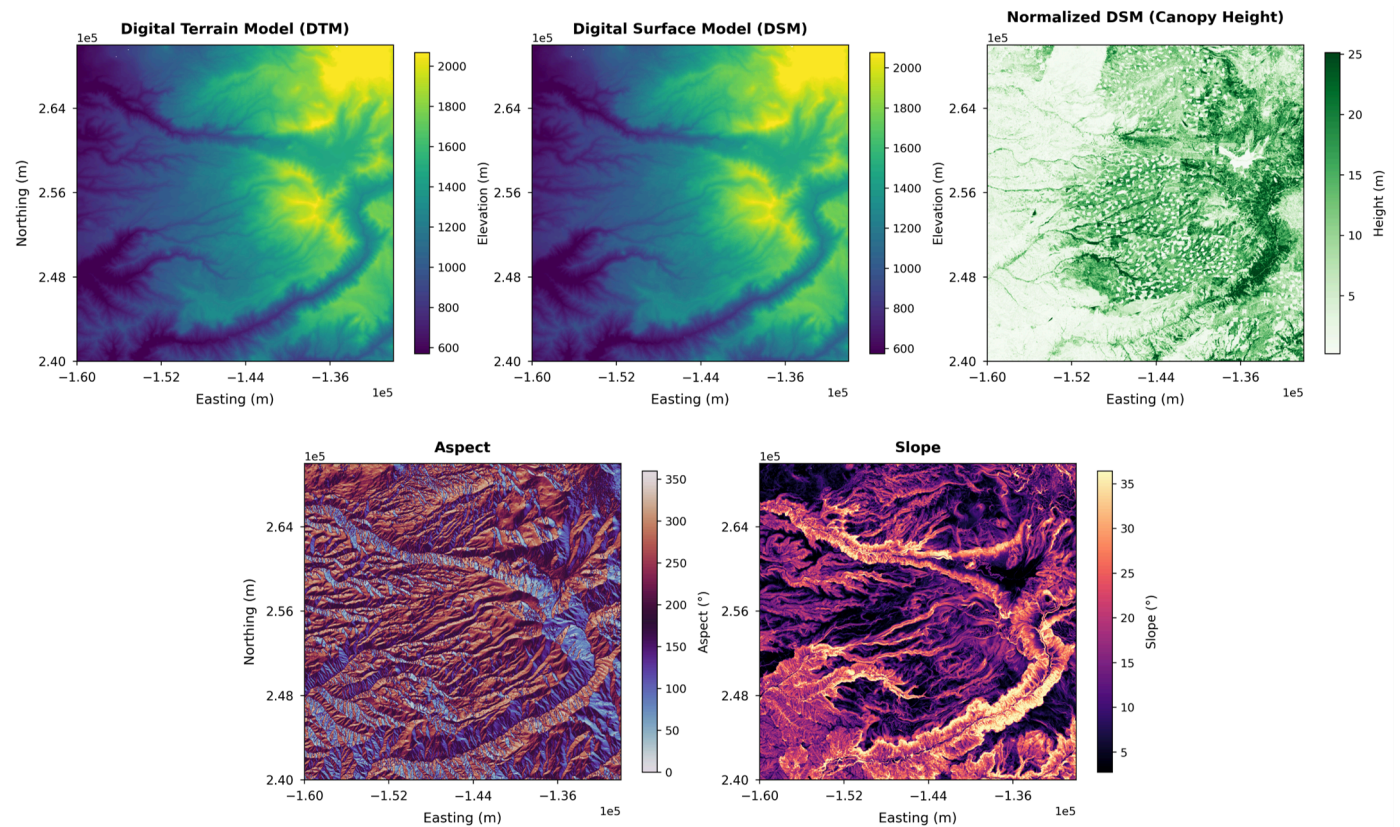


Figure 2. Example 1-meter spatial resolution maps of DTM, DSM, nDSM, Slope, and Aspect over a 30-km x 30-km tile.

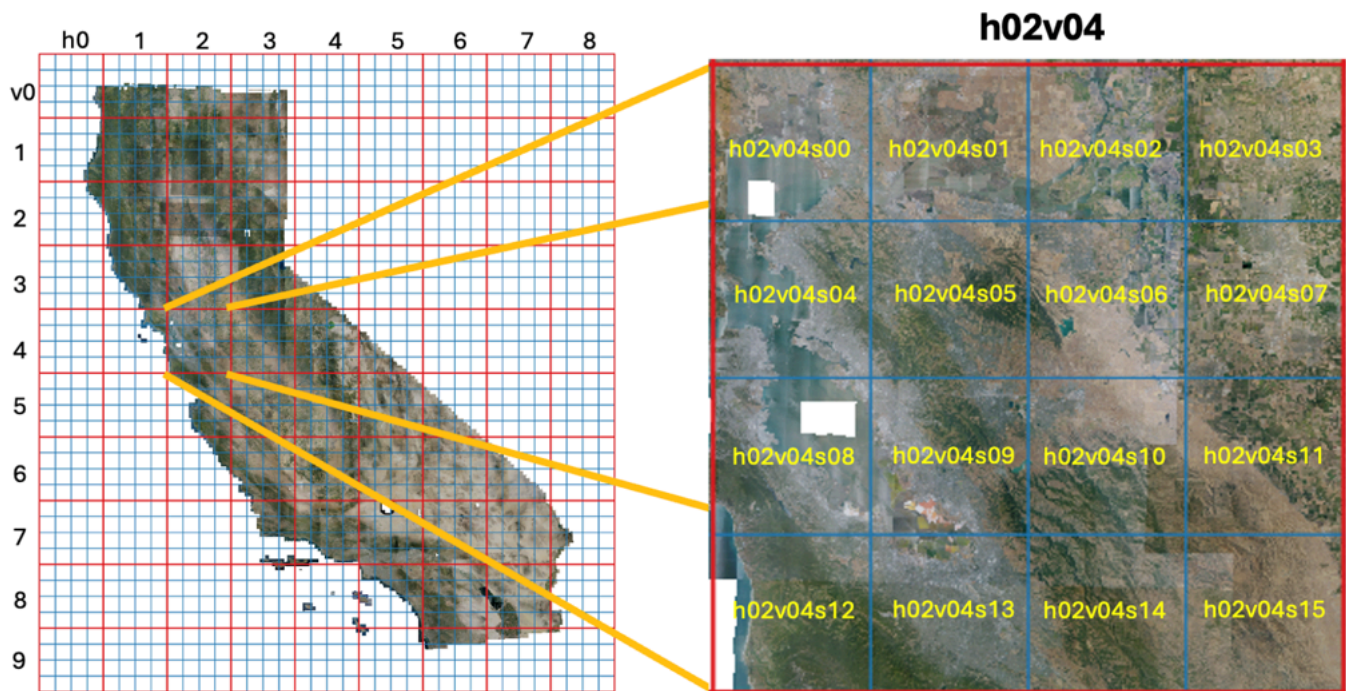


Figure 3. WERK tiling system in California (left): Level-1 WERK Tiles (red polygons) and Nested Level-2 Sub-Tiles (blue polygons). Example Level-1 and Level-2 (right) tiles. The 1-meter spatial resolution products including DTM, DSM, nDSM, Slope, and Aspect are produced over a 30 km x 30 km tile.

3. Application and Derivation

This dataset provides a statewide foundation for data-driven environmental stewardship and risk analysis. The products were designed for applications that require consistent, high-resolution terrain and ecosystem structure across California.

- **Wildfire risk modeling:** Terrain, canopy height, canopy cover, and vegetation height strata can support fuel characterization, fire behavior modeling, exposure assessment, and post-fire damage or recovery analysis.
- **Biodiversity monitoring:** Vegetation height, canopy cover, and structural complexity metrics support habitat characterization and species distribution modeling.
- **Ecosystem management and carbon assessment:** Canopy height and cover metrics support forest health monitoring, ecosystem services assessment, biomass estimation, and landscape prioritization.
- **Urban planning and infrastructure:** DTM, DSM, slope, aspect, and building-fraction layers can support infrastructure siting, flood-risk assessment, urban heat studies, and wildland-urban interface analysis.
- **Hydrologic and geohazard analysis:** High-resolution elevation, slope, and aspect products can support drainage modeling, debris-flow susceptibility analysis, and atmospheric-river impact assessment.
- **Environmental change analysis:** Multi-resolution products allow users to integrate lidar-derived structure with satellite, field, and management datasets across California landscapes.

4. Quality Assessment

Quality assessment evaluated the internal consistency of the derived raster products, verified that product values were within expected physical ranges, and, where appropriate, compared selected products with independent reference datasets. Because the source USGS 3DEP lidar point clouds undergo vendor and USGS quality assurance prior to public release, this assessment focused on the quality of the derived products rather than re-validating the source point clouds.

Post-Processing and Product Consistency Verification

- The DTM served as the spatial reference for all statewide products. All derived raster layers were aligned to the DTM grid and masked to its valid-data footprint.
- As part of quality assessment, DTM elevations were verified to fall within the known elevation range of California (approximately -86 m at Death Valley to 4,421 m at Mount Whitney). Values outside this range were assigned NoData.
- DSM elevations were verified to be greater than or equal to the corresponding DTM elevations. The nDSM and CHM products were recomputed as the difference between the DSM and DTM ($\text{DSM} - \text{DTM}$).
- Slope and aspect values were verified to fall within their expected ranges ($0-90^\circ$ and $0-360^\circ$, respectively).
- Canopy cover values were verified to fall within the range 0–100%. Hierarchical consistency was also evaluated to ensure that canopy cover ≥ 2 m was greater than or equal to canopy cover within the 2–4 m and 4–8 m height strata.
- Canopy height percentiles were evaluated for monotonicity ($P_{10} \leq P_{25} \leq P_{50} \leq P_{75} \leq P_{90} \leq P_{95} \leq P_{99}$).

External Product Comparisons

The DTM products were evaluated against independent USGS 3DEP DTMs at validation sites in Santa Clara County (2020) and Sonoma County (2013) (Table 4). Results demonstrated centimeter-level agreement and millimeter-scale mean bias following spatial alignment, reprojection to the

reference coordinate system, and clipping to the common analysis extent.

Table 4. DTM product comparison.

Metric	Santa Clara County (2020)	Sonoma County (2013)
Valid overlap pixels	99,660,837	100,175,505
Mean bias (m)	0.0018	0.0015
RMSE (m)	0.0265	0.0310
MAE (m)	0.0059	0.0157
Standard deviation (m)	0.0265	0.0310
95th percentile absolute error (m)	0.0172	0.0479
Pixels within 10 cm	99.17%	99.33%
Pixels within 25 cm	99.65%	99.78%

The nDSM/CHM products were evaluated by comparison with National Ecological Observatory Network (NEON) ecosystem structure products at three California sites: San Joaquin Experimental Range (SJER), Soaproot Saddle (SOAP), and Lower Teakettle (TEAK) (NEON, 2025) (Table 5). These sites represent a range of vegetation conditions from open oak woodland to mixed-conifer forest. The comparison does not constitute a temporally coincident evaluation because the products were derived from 2022 USGS 3DEP lidar acquisitions, whereas the NEON products were acquired in 2023. Observed differences reflect both acquisition-date differences and methodological differences, including the use of a first-return TIN-based DSM workflow for the products and pit-free canopy surface generation for the NEON products.

Table 5. nDSM/CHM product comparison to NEON products

Metric	All NEON Sites Combined
Valid co-located pixels	2,852,054
Mean bias (m)	-0.170
Median bias (m)	+0.025
Standard deviation (m)	5.737
RMSE (m)	5.739
MAE (m)	2.649
R ²	0.6735
Pixels within 1 m	61.0%
Pixels within 2 m	69.0%

Known Limitations and User Notes

- Lidar acquisition dates vary among USGS 3DEP campaigns. In areas of overlapping coverage, the most recent lidar acquisition was used to generate the statewide mosaics. When acquisition timing is relevant to an analysis, users should consult the pixel-level lidar acquisition year information or the lidar campaign GeoPackage.
- The Microsoft building footprint dataset was compiled from sources acquired between 2014 and 2018 and is not temporally consistent with all lidar acquisitions (2013–2023). Building masking reduces, but does not completely eliminate, rooftop contamination in the 10 m and 30-m vegetation structure products.
- Some original USGS 3DEP point clouds contain outlier returns classified as ASPRS Class 1 (Unclassified). Although post-processing removes most anomalous outliers, some may remain and introduce localized artifacts in the DSM and derived nDSM/CHM products.
- Small water bodies, low point-density areas, and regions with limited ground returns may contain data gaps or TIN interpolation artifacts, particularly in DTM-derived products.
- Linear artifacts may occur in areas where multiple LiDAR flightlines overlap, resulting in visible striping or subtle discontinuities in derived products. Users should consider these artifacts when interpreting fine-scale spatial patterns in the data.
- The nDSM/CHM products were generated without applying a pit-free canopy surface algorithm.
- This dataset represents the extent of USGS 3DEP lidar coverage available at the time of production and therefore does not fully cover California (78.3% of California land area).

5. Data Acquisition, Materials, and Methods

Data Acquisition and Compute Environment

Input data consisted of publicly available USGS 3DEP airborne lidar point cloud data for California obtained from the USGS Rocky Mountain Region web portal (<https://rockyweb.usgs.gov/>) and processed on NASA High-End Computing Capability (HECC, <https://www.nas.nasa.gov/>) resources. The

dataset includes 118 lidar campaigns acquired between 2013 and 2023, comprising approximately 40 TB of laz point cloud data covering 78.3% of California land area. Source campaigns differ in point density, acquisition quality, horizontal and vertical reference systems, geoid models, and nominal point spacing. Of the 118 campaigns, 71 meet USGS Quality Level 1 (QL1), 45 meet Quality Level 2 (QL2), and 2 meet Quality Level 0 (QL0).

1-m Topography and Ecosystem Structure Data

Individual LAZ point cloud tiles were processed independently and subsequently mosaicked. Tile-level raster generation was performed using R (v4.3.3) and the lidR (v4.1.2; Roussel et al., 2020) package. Noise returns (ASPRS Classes 7 and 18) were excluded prior to processing. DTMs were generated primarily from ground-classified returns (Class 2) using triangular irregular network (TIN) interpolation, whereas DSMs were generated from all non-noise returns using the dsmtin() algorithm.

Campaign-wide 1-m DTM and DSM mosaics were generated using a hierarchical GDAL workflow. To reduce memory requirements, raster tiles were processed in batches of 40. Overlapping tiles were resolved using GDAL VRT pixel functions, selecting the minimum elevation for DTMs and the maximum elevation for DSMs.

Campaign mosaics were reprojected to the California Albers coordinate system (EPSG:3310), transformed to the GEOID18 vertical datum where necessary, clipped to the 30-km × 30-km tiling system, and merged into statewide mosaics. In regions of overlapping lidar coverage, the most recent acquisition was retained. The nDSM and CHM products were computed as the difference between the DSM and DTM (DSM – DTM). Slope and aspect products were derived from the final DTM.

10-m and 30-m Topography and Ecosystem Structure Data

The 10-m and 30-m DTM and DSM products were generated by block-averaging the corresponding 1-m products to the target spatial resolution. The nDSM and CHM products were recomputed as the difference between the DSM and DTM (DSM – DTM), and slope and aspect were derived from the aggregated DTM. Campaign mosaics were reprojected to the California Albers coordinate system (EPSG:3310) and merged into statewide mosaics using the same temporal precedence rule, whereby the most recent lidar acquisition was retained in areas of overlapping coverage.

Canopy structure variables were derived directly from the original LAZ point cloud tiles. Prior to height normalization, lidar returns intersecting building footprints were excluded using the Microsoft Building Footprints dataset for California (Microsoft, 2023). Building polygons were reprojected to the coordinate reference system of each lidar tile, clipped to the tile extent, buffered by 1 m, and used to remove likely rooftop returns from vegetation structure calculations.

Following building masking, lidar point elevations were normalized using the precomputed DTM. Negative normalized heights were set to zero and returns exceeding 120 m were excluded as biophysically implausible. The normalized point cloud was filtered to retain ASPRS Classes 1 (Unclassified), 3 (Low Vegetation), 4 (Medium Vegetation), and 5 (High Vegetation). Six canopy cover metrics were computed from both all returns and first returns for the ≥2 m, 2–4 m, and 4–8 m height strata. First-return height standard deviation, skewness, kurtosis, and height percentiles (P10, P25, P50, P75, P90, P95, and P99) were computed in a single pass using R's lasR (v0.13.7; Roussel et al., 2026) package.

Building fraction rasters were generated at 10 m and 30 m by rasterizing building polygons to a 1-m binary grid and aggregating to the target resolution using the mean. Canopy cover products were weighted by vegetation fraction (1.0 – building fraction) to reduce artificial canopy cover inflation in partially urban pixels.

Canopy structure products were mosaicked by campaign using a hierarchical batching workflow, with overlapping regions merged using the mean for continuous variables.

Campaign mosaics were clipped to campaign boundary polygons prior to statewide assembly to prevent values from extending beyond valid acquisition extents. Final statewide mosaics were generated in the California Albers coordinate system (EPSG:3310), with the most recent lidar acquisition retained in areas of overlapping coverage.

6. References

- Microsoft. 2023. US Building Footprints: California. <https://github.com/microsoft/usbuildingfootprints>
- NEON (National Ecological Observatory Network). Elevation - LiDAR (DP3.30024.001), RELEASE-2025. <https://doi.org/10.48443/yra6-vw36>
- Roussel, J.-R. 2026. lasR: Fast Airborne LiDAR Data Processing. <https://github.com/r-lidar/lasR>.
- Roussel, J.-R., D. Auty, N. C. Coops, P. Tompalski, T. R. H. Goodbody, A. S. Meador, J.-F. Bourdon, F. de Boissieu, and A. Achim. 2020. lidR: An R package for analysis of Airborne Laser Scanning (ALS) data. Remote Sensing of Environment 251:112061. <https://doi.org/10.1016/j.rse.2020.112061>
- U.S. Geological Survey. 2024. 3D Elevation Program 3DEP lidar point cloud. <https://www.usgs.gov/3d-elevation-program>