Light Detection and Ranging (LiDAR) is a mapping technology that leverages lasers to create a 3D model of the ground, vegetation, and other objects. A laser pulse is emitted and the ‘echo’ is measured after it bounces off something in the environment and returns to the sensor. Based on the difference in timing or wavelength phase between the emitted and returning pulse, and the speed or wavelength of light, it is possible to estimate the distance from the sensor to the object being measured.

In aerial LiDAR surveys, a LiDAR sensor is mounted on an aircraft – typically a plane, helicopter, or unmanned aerial vehicle. As the aircraft flies it traces laser pulses densely over a landscape, generating millions of point returns which each show the vertical position of the surface relative to sea level.

It is possible to separate ground vs. aerial LiDAR returns based on various processing techniques. Most simply, all the point returns can be sampled with a regular, square grid, and in each grid the lowest point can be assumed to be the ground. Then, the relative heights of other points in that grid cell are taken as the heights of vegetation, buildings, bridges, or other features over the local ground surface.

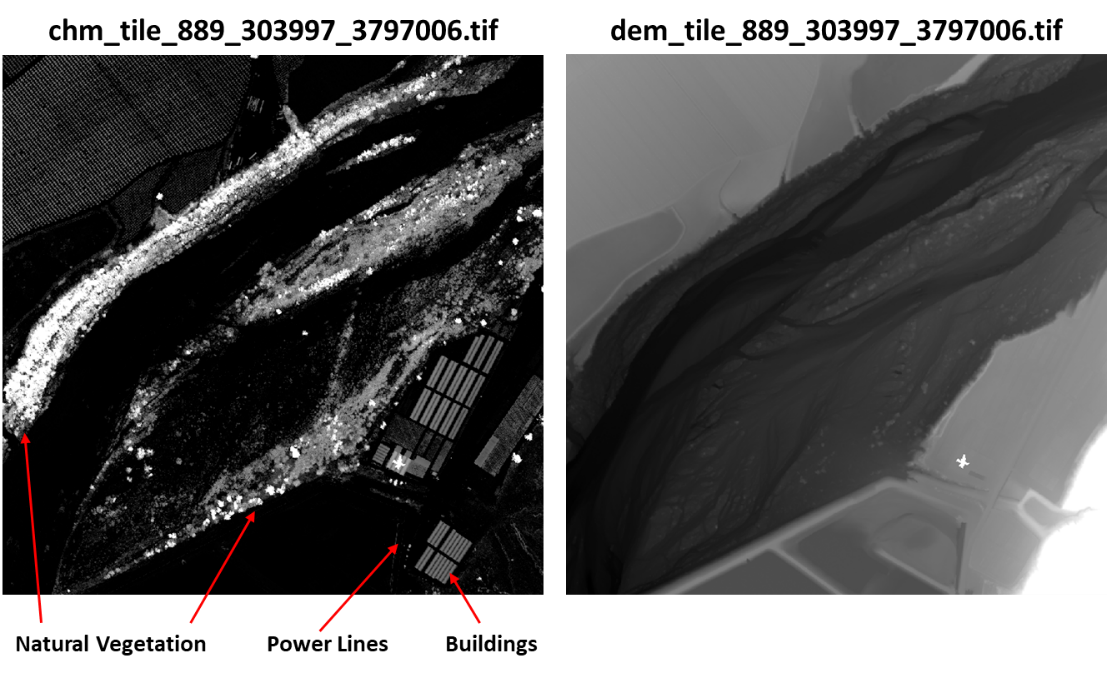
We contracted an aerial LiDAR survey of the Santa Clara River in Southern California during the summer of 2024, working with the National Center for Airborne Laser Mapping (NCALM). Our survey included more than 100 linear km of river length and approximately 134 km2 of total land area. Our initial data products include:

1. Digital Elevation Models (DEM) – raster ground elevation (in m) with 1 m pixels
2. Canopy Height Models (CHM) – raster with highest height return (in m) in each pixel RELATIVE to the local ground elevation (e.g. a pixel overlying a tree might have height = 5 to 20 m)

Both of these raster datasets are provided in tiled .tif format. Each image tile is 1000 x 1000 pixels, and each pixel is a 1 x 1 m square containing the height value in that pixel (either ground or ‘canopy height’, depending on the file type). Each tile image has a filename with the format:

<type>\_tile\_<tile\_index>\_<x-min>\_<y-max>.tif

For example, the files below are from the 889th tile, with upper left coordinate at easting=303997 and northing = 3797006. There is no overlap between adjacent tiles, which are spaced every 1000 m.



Rivers support some of the most imperiled ecosystems in the United States, and many endangered plants and animals live there. The Santa Clara in particular is a fairly ‘intact’ river, which is still in relatively good environmental condition, and it harbors protected species like the Least Bell’s Vireo and Yellow-breasted Chat that depend on healthy woodlands for survival. We are interested in studying how the woodlands and other ecosystems on the river respond to drought, climate change, and other stressors. In particular, we hope to map out vegetation types across the river, and would like to be able to distinguish between vegetation and non-vegetative material in our LiDAR imagery.

As seen in the example images above, the current canopy height models include some non-vegetative features, such as buildings, power lines, and bridges. We would like to develop an algorithm to identify and segment those features out of the imagery so that they are not included in our analysis of vegetation.

Thank you for your help! We are excited to see what you produce.