PCL transects with process_pcl

Jeff Atkins

2020-01-30

This vignette is aimed at those working with portable canopy LiDAR (PCL) data. Raw PCL data is most typically stored in comma separated values (.csv) files that contain a series of ASCII strings of return distances and return intensities. These return distances are interspersed with data markers that correspond to known distance measures on the ground as described in Hardiman et al. (2013). The spacing between these markers, typically at 10 m, is used as an input into the forestr processing algorithm (Atkins et al. in review).

This vignette will walk you through how to analyze a PCL or portable canopy LiDAR file in two ways: 1) an automated way using the `process_pcl()` function, and 2) by processing the file step-by-step.

The `process_pcl()` function has the advantage that all of the output from the file is written to disk in an output directory that is created in the working directory which you can access via `getwd()` in your console. Approaching the problem step-by-step allows you to keep sections of the process in the workspace, but is more cumbersome and is not recommended as all relevant files are written to the output directory and can be worked with seperatly if so desired.

Using `process_pcl`

To analyze a single PCL transect, call the function `process_pcl` with the following information:

- `marker.spacing` is the distance between markers in the dataset (default is 10 meters). In PCL data, data markers which are default values below -9999 in the code, indicate the end of defined sections along the transect.
- `user_height` is the offset, based on the user's own height differences and directly refers to the distance of the front portion of the laser from the ground. The value is added to all the distances in the data set.
- `max.vai` is the maximum vegetation area index at the 1 x 1 meter scale that is used to adjust sections of the data under saturated conditions (i.e. when the LiDAR fails to penetrate the canopy and records only canopy hits). A default value of 8 is used and has been found to be well-representative for a broad selection of forests.
- `pavd` and `hist` refers to the plant area volume density graph with histogram option. If `pavd` is set to TRUE, a plot will be written to the output folder. If `hist` is set to TRUE, a histogram will be superimposed over the PAVD curve.

The `process_pcl()` function will print multiple CSC metrics to the screen, including canopy rugosity, rumple, porosity, etc., but will also create an output directory named output within the working directory where it will store four things:

1. The Summary Matrix - a .csv file of columnar values of mean leaf heigh, maximum height, vegetation area index (VAI), etc.
2. The Hit Matrix - a .csv file containing VAI by rows where each row correpsonds to the x, z position of VAI in the canopy.
3. The Output File - this is a .csv file that contains all CSC metrics in output form.
4. The Hit Grid - this is a PNG image file of the vertical hit grid as explained in Hardiman et al. (2013) and is a graphic representation of the vertical distribution of leaf density throughout the canopy. This can be recreated and modified using the source code with the Hit Matrix file as the input.

```
require(forestr)
#> Loading required package: forestr
# Link to stored, raw PCL data in .csv form
uva.pcl <- system.file("extdata", "UVAX_A4_01W.csv", package = "forestr")

# Run process complete PCL transect, store output to disk
process_pcl(uva.pcl, marker.spacing = 10, user_height = 1.05, max.vai = 8, pavd = FALSE, hist = FALSE, s
#> how many in base df have NA
#> [1] 253
#> Transect Length
#> [1] 40
#> Table of sky hits
#>
#> FALSE  TRUE
#> 14323   253
#> Intensity Statistics
#> Mean Intensity
#> [1] 81.19242
#> Median Intensity
#> [1] 83
#> Standard Deviation of Intensity
#> [1] 15.08965
#> Max intensity
#> [1] 126
#> Minimum intensity
#> [1] 8
#> Intensity Skewness
#> [1] -0.8123838
#> Intensity Kurtosis
#> [1] 4.636563
#> RAW LiDAR metrics -- WARNING
#> Mean Return Height (m) of raw data
#> [1] 9.114691
#> Standard Deviation of raw  Canopy Height returns-- meanStd in old code
#> [1] 5.130027
#> Median of raw  Canopy Height returns
#> [1] 7.8185
#> Max Measured Canopy Height (m)
#> [1] 26.511
#> Scan Density
#> [1] 364.4
#> OPENNESS AND COVER METRICS
#> Sky Fraction (%)
#> [1] 1.736326
#> Cover Fraction (%)
#> [1] 98.26367
#> Rumple
#> [1] 4.925
#> Mean Gap Fraction ---as error check should be same as porosity
#> [1] 0.7392857
#> now we replace the 0's with 1's so when we take the ln they = 0
#> Clumping Index
#> [1] 0.9249189
```

```
#> Transect Length (m)
#> [1] 40
#> HEIGHT METRICS
#> Mean Leaf Height (H) - plot mean of column mean leaf height
#> [1] 11.97755
#> Height2 (H[2]) - standard deviation of column mean leaf height
#> [1] 4.922554
#> H [median] - median column mean leaf height
#> [1] 11.08111
#> Mean Leaf Height variance (H[var]) - variance of column mean leaf height
#> [1] 24.23154
#> Root Mean Square Mean Leaf Height (H[rms]) - the root mean square or quadratic mean of column mean l
#> [1] 12.92624
#> Max canopy height (m)
#> [1] 26.511
#> Mean Outer Canopy Height (m) or MOCH
#> [1] 17.5717
#> Median Column Max Ht
#> [1] 19.038
#> AREA AND DENSITY METRICS
#> Mean VAI - mean VAI for entire transect
#> [1] 7.300039
#> SD VAI - SD of VAI for entire transect
#> [1] 1.364013
#> Median VAI - median VAI for entire transect
#> [1] 8
#> Maximum VAI x,y entire transect -- max el!
#> [1] 8
#> Mean Height of VAI[max] - modeEl
#> [1] 11.95
#> Mode 2- The standard deviation of VAImax or MaxEl
#> [1] 6.636921
#> Median of VAI max or Maxel
#> [1] 10.5
#> Maximum VAI x,y entire transect -- max el!
#> [1] 7.551586
#> Mean Peak VAI for entire transect
#> [1] 3.43842
#> SD of peak VAI for entire transect
#> [1] 1.414623
#> Median Peak VAI for entire transect
#> [1] 3.164244
#> CANOPY AND OPENNESS METRICS (cont.)
#> Deep Gaps
#> [1] 0
#> Deep Gap Fraction (0-1)
#> [1] 0
#> ARRANGEMENT METRICS
#> Canopy porosity
#> [1] 0.7392857
#> Square of leaf height variance (stdStd from old script)
#> [1] 489.5471
#> Mean Standard deviation of leaf heights -- meanStd
```

```
#> [1] 15.44022
#> Canopy Rugosity
#> [1] 15.84761
#> Surface Rugosity--TopRugosity
#> [1] 6.539101
#> Effective Number of Layers:
#> [1] 20.6851
#> No. of NA values in hit matrix
#> [1] 828
```

Graphing capabilities

`process_pcl()` creates two types of figures:

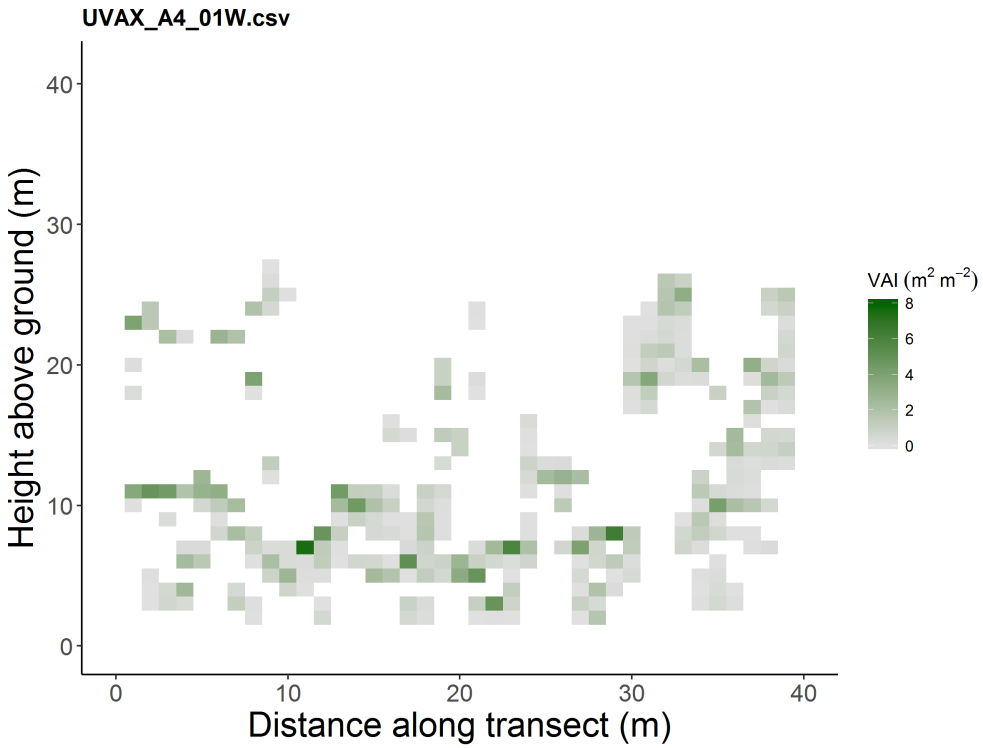

Figure 1: A canopy hit grid that shows the density of VAI aggregated to 1 meter pixels in the canopy.

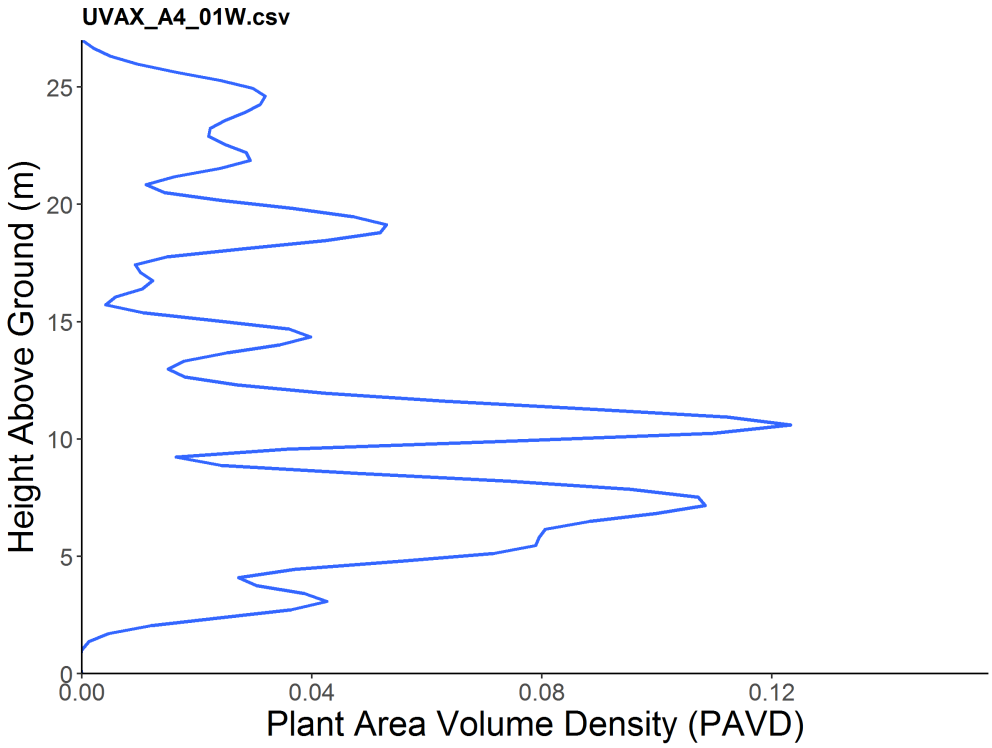

Figure 2: A plant area volume density plot

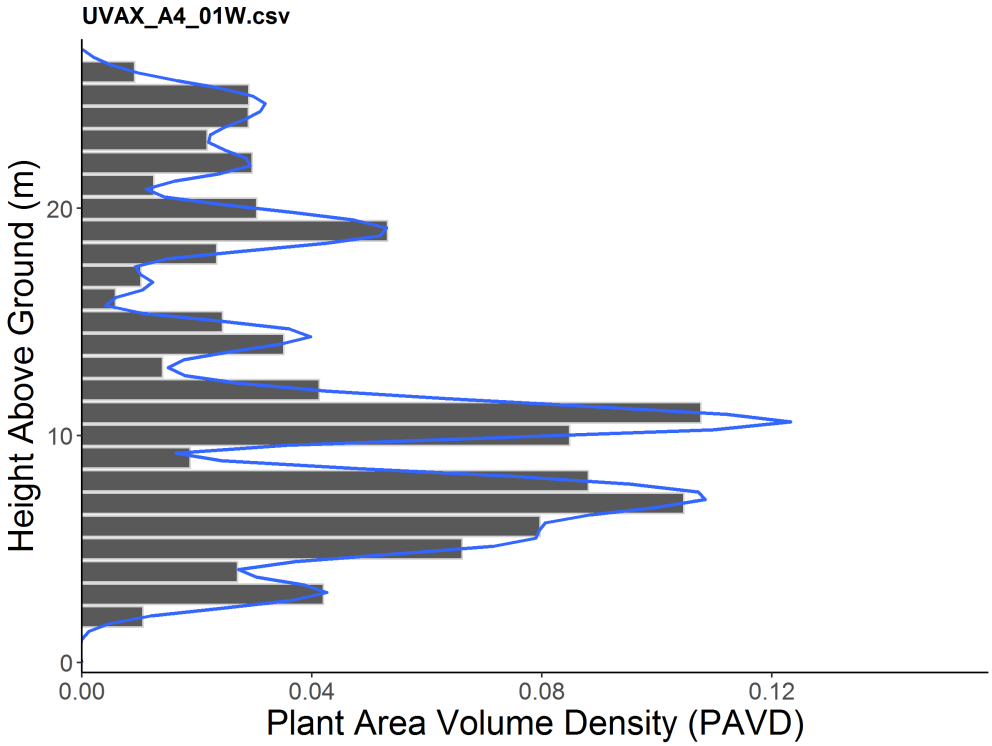

Figure 3: A plant area volume density plot with histogram