Canopy Height Models, Digital Surface Models & Digital Elevation Models - Work With LiDAR Data in R

Learning Objectives

After completing this tutorial, you will be able to:

- Define Canopy Height Model (CHM), Digital Elevation Model (DEM) and Digital Surface Model (DSM)
- Describe the key differences between the **CHM**, **DEM**, **DSM**
- Derive a **CHM** in R using raster math.

What you need

You need `R` and `RStudio` to complete this tutorial. Also you should have an `earth-analytics` directory setup on your computer with a `/data` directory with it.

- How to Setup R / RStudio
- Setup your working directory
- Intro to the R & RStudio Interface

R Libraries to Install:

- **raster:** `install.packages("raster")`
- **rgdal:** `install.packages("rgdal")`

As we discussed in the previous lesson, LiDAR or **Li**ght **D**etection **a**nd **R**anging is an active remote sensing system that can be used to measure vegetation height across wide areas. If the data are discrete return, Lidar point clouds are most commonly derived data product from a lidar system. However, often people work with lidar data in raster format given it's smaller in size and thus easier to work with. In this lesson, we will import and work with 3 of the most common lidar derived data products in `R`:

1. **Digital Terrain Model (or DTM):** ground elevation.
2. **Digital Surface Model (or DSM):** top of the surface (imagine draping a sheet over the canopy of a forest
3. **Canopy Height Model (CHM):** the elevation of the Earth's surface - and it sometimes also called a DEM or digital elevation model.

3 Important Lidar Data Products: CHM, DEM, DSM

Digital Surface Model (DSM), Digital Elevation Models (DEM) and the Canopy Height Model (CHM) are the most common raster format lidar derived data products. One way to derive a CHM is to take the difference between the digital surface model (DSM, tops of trees, buildings and other objects) and the Digital Terrain Model (DTM, ground level). The CHM represents the actual height of trees, buildings, etc. with the influence of ground elevation removed. Graphic: Colin Williams, NEON

Digital Elevation Model

In the previous lesson, we opened a digital elevation model. The digital elevation model (DEM), also known as a digital terrain model (DTM) represents the elevation of the earth surfacel. The DEM represents the ground - and thus DOES NOT INCLUDE trees and buildings and other objects.

```r
# load libraries
library(raster)
library(rgdal)

# set working directory to ensure R can find the file we wish to import
# setwd("working-dir-path-here")
```

```r
# open raster data
lidar_dem <- raster(x="data/week3/BLDR_LeeHill/pre-flood/lidar/pre_DTM.tif")
## NOTE: rgdal::checkCRSArgs: no proj_defs.dat in PROJ.4 shared files

# plot raster data
plot(lidar_dem,
     main="Lidar Digital Elevation Model (DEM)")
## NOTE: rgdal::checkCRSArgs: no proj_defs.dat in PROJ.4 shared files
## NOTE: rgdal::checkCRSArgs: no proj_defs.dat in PROJ.4 shared files
```

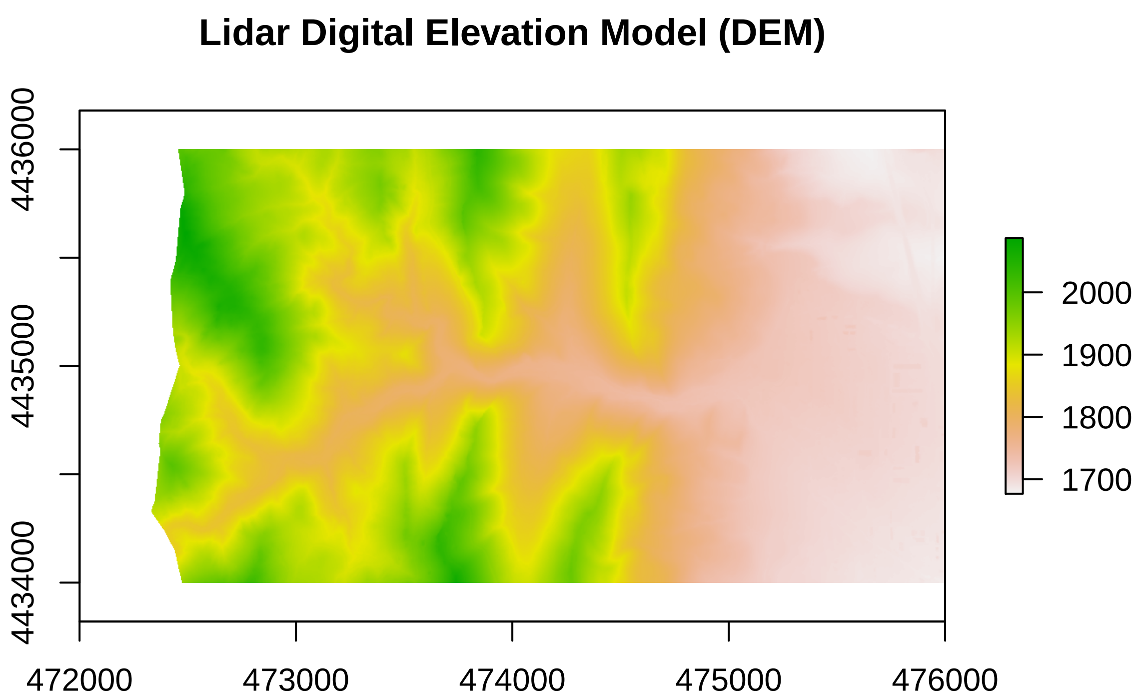

Figure 1: digital elevation model plot

Then we opened the digital SURFACE model (DSM). The DSM represents the top of the earth's surface. Thus, it INCLUDES TREES, BUILDINGS and other objects that sit on the earth.

```r
# open raster data
lidar_dsm <- raster(x="data/week3/BLDR_LeeHill/pre-flood/lidar/pre_DSM.tif")
## NOTE: rgdal::checkCRSArgs: no proj_defs.dat in PROJ.4 shared files

# plot raster data
plot(lidar_dsm,
     main="Lidar Digital Surface Model (DSM)")
## NOTE: rgdal::checkCRSArgs: no proj_defs.dat in PROJ.4 shared files
## NOTE: rgdal::checkCRSArgs: no proj_defs.dat in PROJ.4 shared files
```

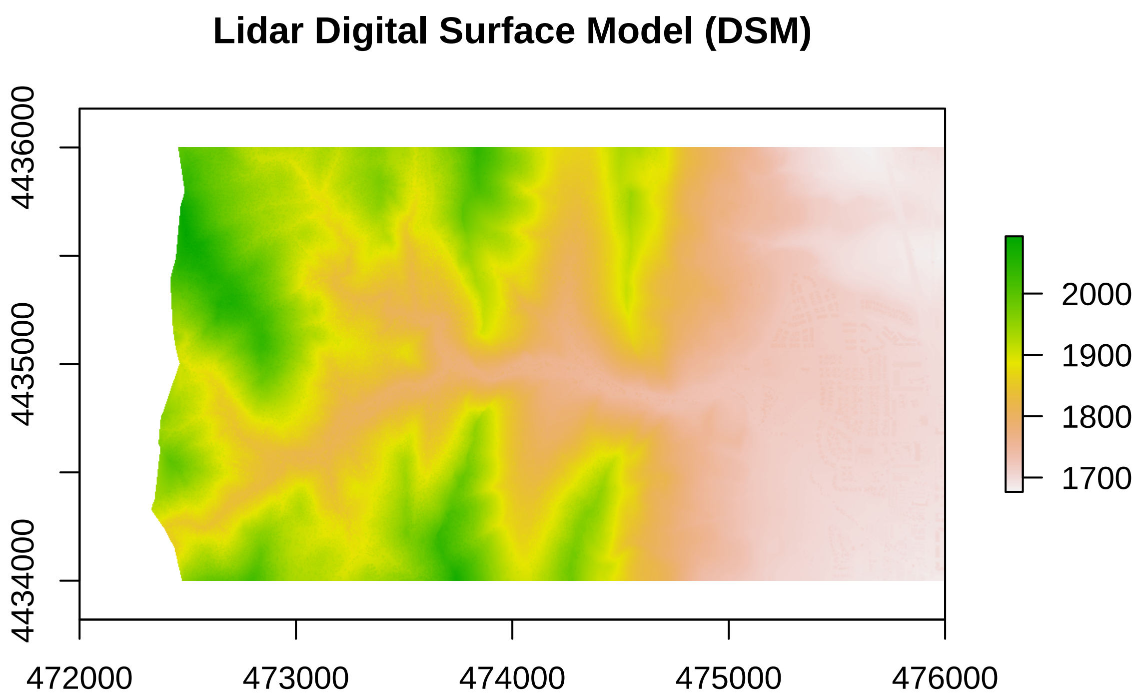

Figure 2: digital surface model plot

Canopy Height Model

The canopy height model (CHM) represents the HEIGHT of the trees. This is not an elevation value, rather it's the distance between the ground and the top of the trees. Some canopy height models also include buildings so you need to look closely are your data to make sure it was properly cleaned before assuming it represents all trees!

Calculate difference between two rasters

There are different ways to calculate a CHM. One easy way is to subtract the DEM from the DSM.

DSM - DEM = CHM

This math gives you the residual value or difference between the top of the earth surface and the ground which should be the heights of the trees (and buildings if the data haven't been "cleaned").

```
# open raster data
lidar_chm <- lidar_dsm - lidar_dem
## NOTE: rgdal::checkCRSArgs: no proj_defs.dat in PROJ.4 shared files

# plot raster data
plot(lidar_chm,
     main="Lidar Canopy Height Model (CHM)")
## NOTE: rgdal::checkCRSArgs: no proj_defs.dat in PROJ.4 shared files
## NOTE: rgdal::checkCRSArgs: no proj_defs.dat in PROJ.4 shared files
```

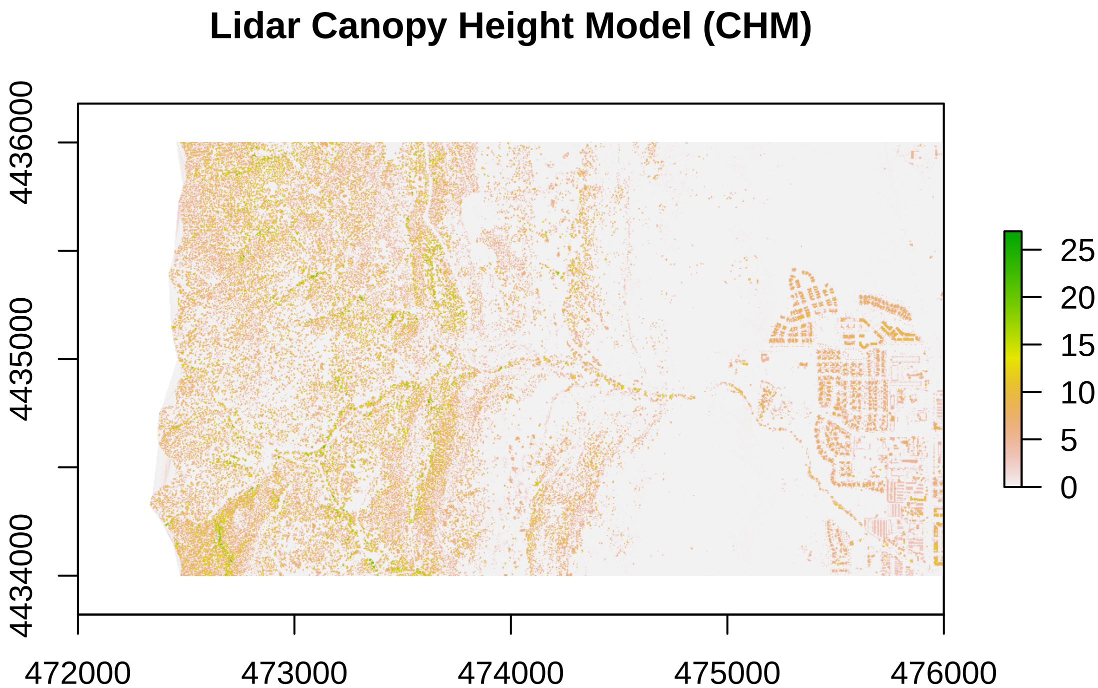

Figure 3: canopy height model plot

Plots Using Breaks

Sometimes a gradient of colors is useful to represent a continuous variable. But other times, it's useful to apply colors to particular ranges of values in a raster. These ranges may be statistically generated or simply visual.

Let's create breaks in our CHM plot.

```
# plot raster data
plot(lidar_chm,
     breaks = c(0, 2, 10, 20, 30),
     main="Lidar Canopy Height Model",
     col=c("white","brown","springgreen","darkgreen"))
## NOTE: rgdal::checkCRSArgs: no proj_defs.dat in PROJ.4 shared files
## NOTE: rgdal::checkCRSArgs: no proj_defs.dat in PROJ.4 shared files
```

Export a raster

When can export a raster file in R using the `write.raster()` function. Let's export the canopy height model that we just created to our data folder. We will create a new directory called "outputs" within the week 3 director. This structure allows us to keep things organized, separating our outputs from the data we downloaded.

```
# check to see if an output directory exists
dir.exists("data/week3/outputs")
## [1] TRUE

# if the output directory doesn't exist, create it
if (dir.exists("data/week3/outputs")) {
```

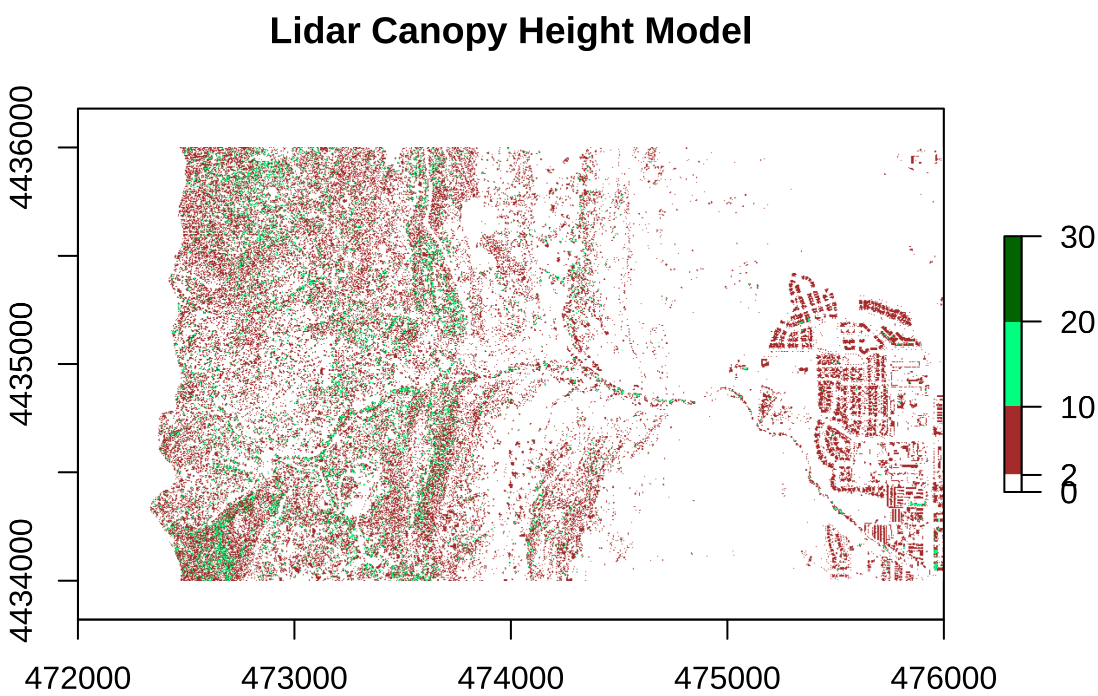

Figure 4: canopy height model breaks

```
  print("the directory exists!")
  } else {
    # if the directory doesn't exist, create it
    # recursive tells R to create the entire directory path (data/week3/outputs)
    dir.create("data/week3/outputs", recursive=TRUE)
  }
## [1] "the directory exists!"

# export CHM object to new GeotIFF
writeRaster(lidar_chm, "data/week3/outputs/lidar_chm.tiff",
            format="GTiff",  # output format = GeoTIFF
            overwrite=TRUE) # CAUTION: if this is true, it will overwrite an existing file
```

Data Tip: We can simplify the directory code above by using the exclamation `!` which tells R to return the INVERSE or opposite of the function you have requested R run.

```
# if the output directory doesn't exist, create it
if (!dir.exists("data/week3/outputs")) {
    # if the directory doesn't exist, create it
    # recursive tells R to create the entire directory path (data/week3/outputs)
    dir.create("data/week3/outputs", recursive=TRUE)
}
```

Change detection in Terrain

Now that we've learned about the 3 common data products derived from lidar data, let's use them to do a bit of exploration of our data - as it relates to the 2013 Colorado floods.

Challenge - calculate changes in terrain

- Subtract the post-flood DEM from the pre-flood DEM. Do you see any differences in elevation before the after?
- Create a CHM for both pre-flood and post-flood by subtracting the DEM from the DTM for each year. Next create a CHM DIFFERENCE raster by subtracting the post-flood CHM from the pre-flood CHM.
- Export the files as geotiff's and open them in QGIS. Explore the differences. What differences do you see between the two years?