Understand uncertainty

Learning Objectives

After completing this tutorial, you will be able to:

- Be able to list atleast 3 sources of uncertainty / error associated with remote sensing data.
- Be able to interpret a scatter plot that compares remote sensing values with field measured values to determine how "well" the two metrics compare.
- Be able to describe 1-3 ways to better understand sources of error associated with a comparison between remote sensing values with field measured values.

What you need

You will need a computer with internet access to complete this lesson and the data for week 5 of the course.

Understanding uncertainty and error.

In science, it is important to think about error and uncertainty.

Uncertainty

Uncertainty: the range of values within which the value of the measurand can be said to lie within a specified level of confidence. The uncertainty quantitatively indicates the "quality" of your measurement.It answers the question: "how well does the result represent the value of the quantity being measured?"

Tree height measurement example

So for example let's pretend that we measured the height of a tree 10 times. Each time our tree height measurement may be slightly different? Why? Because maybe each time we visually determined the top fo the tree to be in a slightly different place. Or maybe there was wind that day during measurements that caused the tree to shift as we measured it yielding a slightly different height each time. or... what other reasons can you think of that might impact tree height measurements?

What is the true value?

So you may be wondering, what is the true height of our tree? The true value? In the cause of a tree in a forest, it's very difficult to determine the true height. So we accept that there will be some variation in our measurements and we measure the tree over and over again until we understand the range of heights that we are likely to get when we measure the tree

```
tree_heights <- data.frame(heights=c(10, 10.1, 9.9, 9.5, 9.7, 9.8,
                                     9.6, 10.5, 10.7, 10.3, 10.6))
# what is the average tree height
mean(tree_heights$heights)
## [1] 10.06364
# what is the standard deviation of measurements?
sd(tree_heights$heights)
```

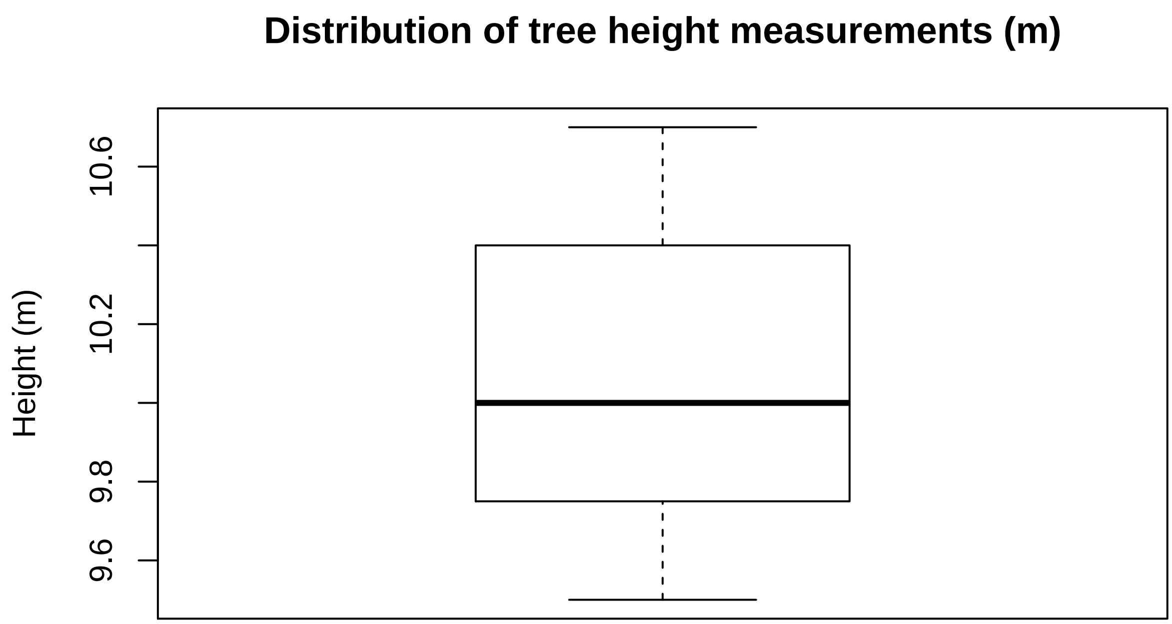

Figure 1: Distribution of tree heights.

```
## [1] 0.4129715
boxplot(tree_heights$heights,
        main="Distribution of tree height measurements (m)",
        ylab="Height (m)")
```

In the case above of our tree heights, our mean value is towards the center of our distribution of measured heights. Thus we might assume that our mean represents the average measured value well. And that our range of expected values or uncertainty can be seen in the box plot.

```
hist(tree_heights$heights, breaks=c(9,9.5,10.5,11),
     main="Distribution of tree height values",
     xlab="height (m)", col="purple")
```

- http://www.ece.rochester.edu/courses/ECE111/error_uncertainty.pdf
- https://www.nde-ed.org/GeneralResources/ErrorAnalysis/UncertaintyTerms.htm

Conventional on the ground methods to measure trees are resource intensive and limit the amount of vegetation that can be characterized. Source: National Geographic

An example LiDAR waveform. Source: NEON, Boulder, CO.

Cross section showing LiDAR point cloud data (above) and the corresponding landscape profile (below). Graphic: Leah A. Wasser

Study site location

Study area plots

QGIS imagery layer

qgis.utils.iface.addRasterLayer(“http://server.arcgisonline.com/arcgis/rest/services/ESRI_Imagery_World_2D/MapServer?f=json&pretty=true”,“raster”) <qgis._core.QgsRasterLayer object at 0x12739ee90>

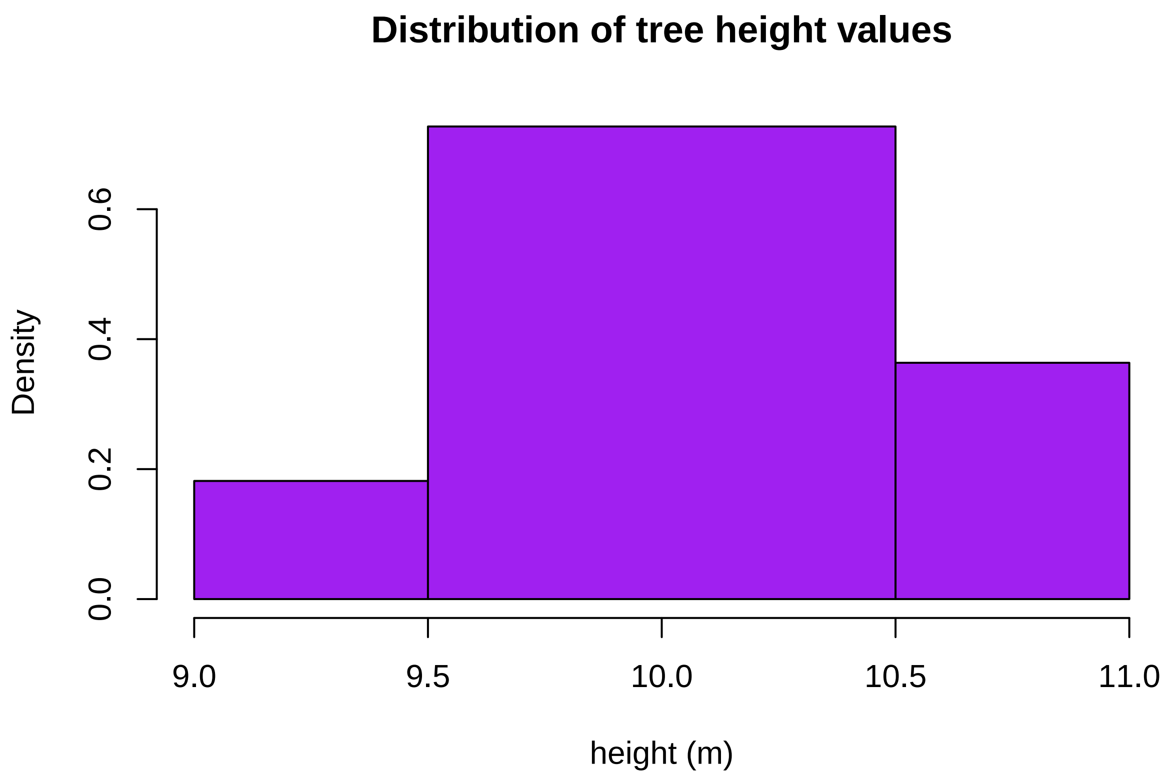

Figure 2: Tree height distribution

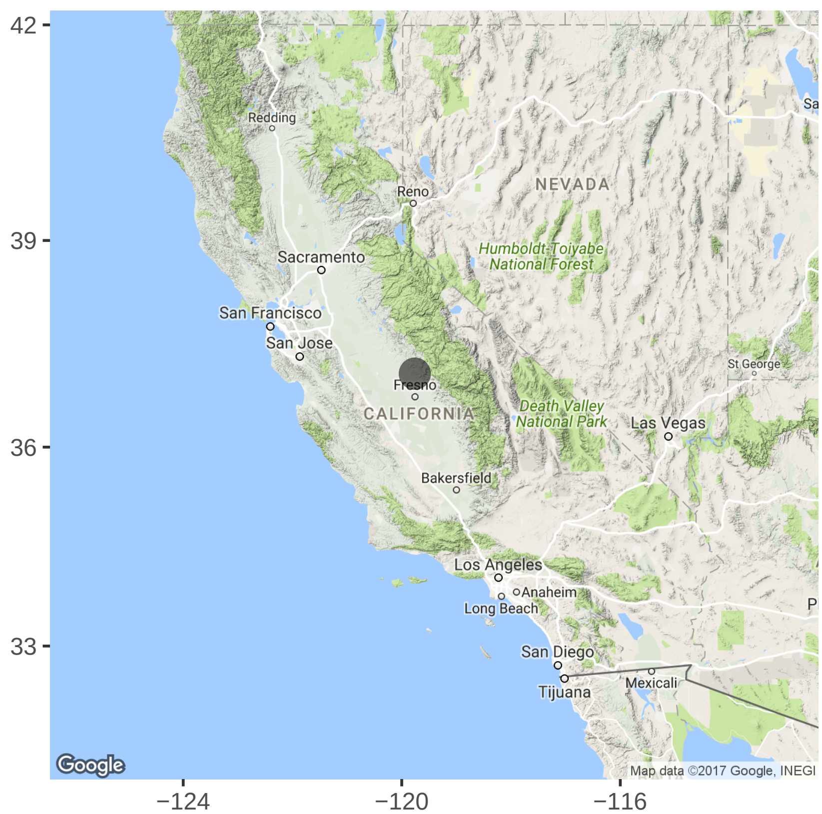

Figure 3: ggmap of study area.

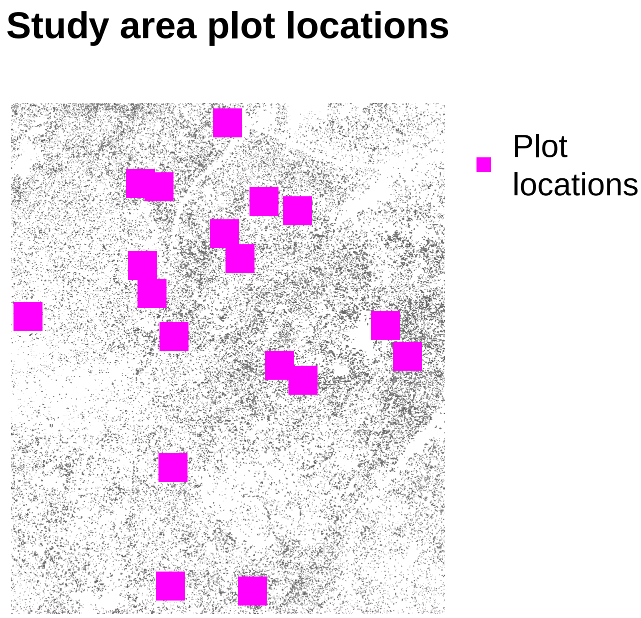

Figure 4: plots

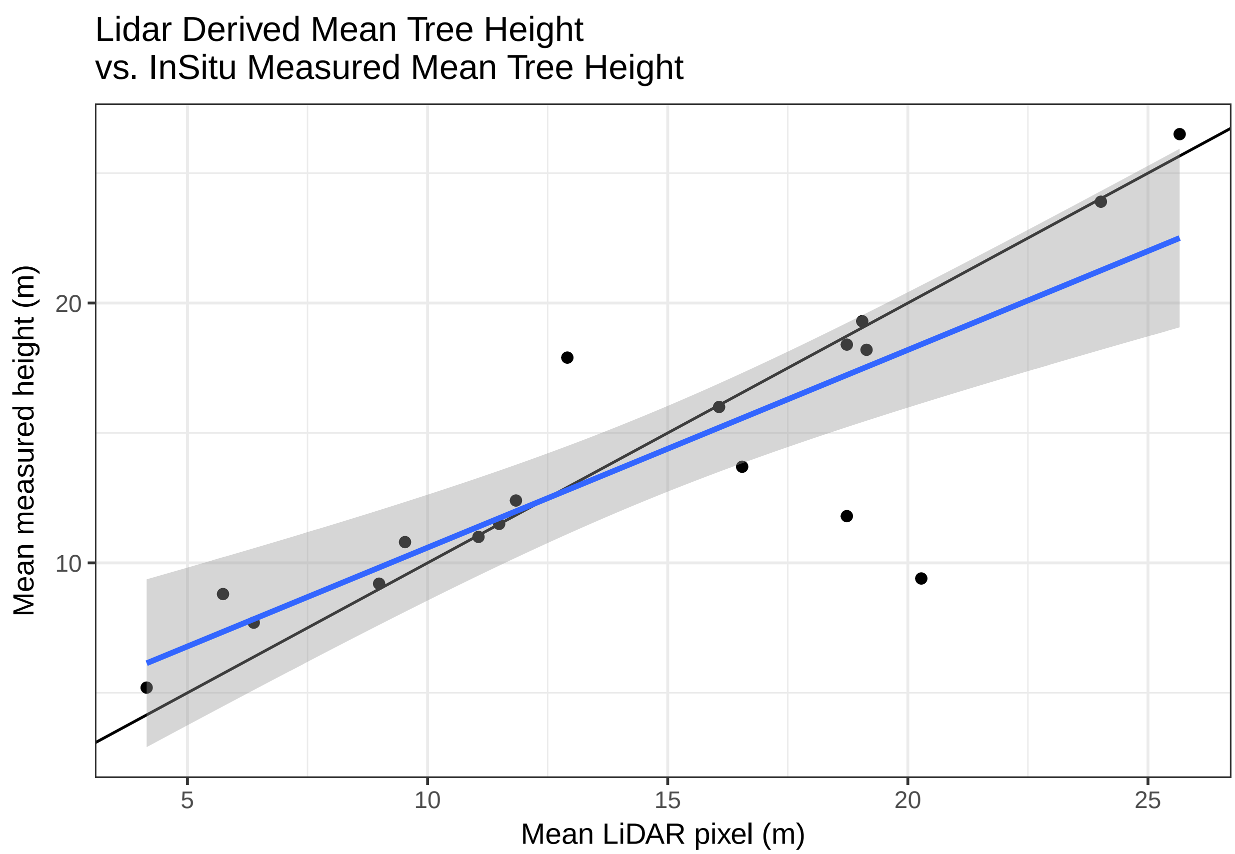

Figure 5: final plot

View interactive scatterplot

View scatterplot plotly

View interactive difference barplot

View scatterplot differences