

Analyzing Tree Deficits and Reforestation Potential in Washington State

Dave McLaughlin

Term Project, GEOG 586, Penn State University, Fall I, 2019

Abstract

Reforestation has been identified as an important component in limiting climate change, as trees are natural carbon sinks. The capabilities of a geographic information system (GIS) can be harnessed to perform analysis on multiple sets of data in order to assess current and potential forest levels and yield information for decision-makers driving reforestation efforts. The objectives of this project are to create a GIS-based analysis method to determine “tree deficits” and to use that method to analyze a region and its potential for reforestation. To accomplish these objectives, a step-by-step method for manipulating the available data and analyzing it with several geostatistical tools was developed. Desktop GIS software (QGIS and ArcGIS Pro) and data freely available via the Internet were used to manipulate data and perform the analysis on forests in the State of Washington, chosen because of its varied geography. I will discuss statewide results of the analysis as well as highlight a number of specific locations with interesting results, discussing the potential for reforestation in those locations. I will also include analysis of a few locations which expose areas for potential improvement in the method I developed.

Background

Reforestation has been proposed to have great potential as part of the solution in fighting climate change, as trees act as carbon sinks while helping moderate local climates (Bastin et al., 2019). In order to focus reforestation efforts, decision-makers need to know current and potential forest levels for regions in general, and individual locations in question. The goal of this project is to create a reproducible method of analysis to determine these levels, to use that analysis on a specific region of limited scope (the State of Washington), and to discuss forest levels and reforestation potential within that region.

The State of Washington in the Pacific Northwest region of the United States boasts a varied geography, with environments ranging from coastal temperate rainforests to alpine mountaintops and arid plateaus. When measured by basal (trunk) area per square acre, these areas support some of the highest and lowest densities of trees in the nation. Washington also has large urban areas, agricultural regions, managed forests, and areas of unlogged old-growth forest. Given this variety, I focused my analysis on this state.

Geostatistical analysis has often been used in the past in combination with satellite imagery and on-the-ground data to describe forests, including by Wilson, Lister, and Riemann (2012) to create the main input dataset for this project. Using that current tree density data, as well as ecoregion definitions, satellite imagery, and elevation, land cover, and land use data, I developed a GIS-based method for

analyzing locations which currently have fewer or smaller trees than could theoretically grow there, which I will call a location's "tree deficit." I then analyzed forests in the State of Washington and discuss reforestation potential within specific regions of the state.

Methodology

Datasets used in the project are detailed below and were acquired from various agency websites and through Esri's Living Atlas of the World service. I used ArcGIS Pro and QGIS (for some batch operations) to complete the manipulation and analysis of data, and ArcGIS Pro to generate the charts and maps included in this report.

Data

Tree Trunk Area

Live tree species basal area of the contiguous United States (2000–2009), from the U.S. Forest Service, is a set of raster imagery at 250m resolution of the estimated square feet of trunk area per acre for each of 324 species found in the U.S. (Wilson et al., 2009).

Ecoregions

Ecoregions define areas of ecosystem similarity, with Level IV being the most fine-grained definition (967 ecoregions in the conterminous U.S.). The data is available from the U.S. Environmental Protection Agency by state or region as vector shapefiles (EPA, 2012). See Omernick and Griffith (2014) for the methods used to determine the ecoregions.

Elevation

Ground Surface Elevation data was downloaded through Esri's Living Atlas of the World at 30m raster resolution (Esri, 2019a). The data is derived from the United States Geological Survey's National Elevation Dataset.

Land Cover

Land Cover data for North America at 250m resolution based on Moderate Resolution Imaging Spectroradiometer (MODIS) satellite imagery was downloaded through the Commission for Environmental Cooperation (2010).

Land Use

Land Use data for Washington was downloaded through the Washington State Department of Ecology (2010).

World Imagery

World Imagery is provided by Esri through the Living Atlas of the World and is made up of satellite and aerial imagery from many sources (Esri, 2019b). This imagery was used for visual confirmation while inspecting data and in generating some of the figures used in the Results section.

Note About Data Conversion

All data were projected to the NAD 1983 HARN StatePlane Washington North FIPS 4601 projection, and all maps were produced using this projection. Raster data available at finer resolutions than 250m were resampled to 250m using the nearest neighbor method and snapped to the tree area data before

analysis. Vector polygon data converted to raster for analysis was converted at 250m resolution (snapped to the tree area data) using the “cell center” cell assignment type.

Methods

Calculating Total Tree Density

To map the tree trunk (basal) area at locations across Washington, I needed to calculate the total based on 324 separate species datasets. To limit the computation needed, I first clipped each species’ raster to the extent of the state’s border, using the ecoregion layer’s polygons as a boundary. QGIS’ batch processing functionality sped this process immensely.

I then performed simple map algebra with the raster calculator to sum the 324 clipped rasters to one raster for the tree density across the state, then converted that raster to hold rounded integer values. To create the map algebra expression, I used the command line to generate a list of the raster image filenames, then manipulated that list in a text/code editor (Visual Studio Code) using its multi-line editing capability to insert quotation marks, spaces, and plus symbols, remove line breaks, etc., and finish with a string like: `"s11.img" + "s12.img" + "s15.img"...`

Analyzing the Distribution of Tree Density

Using ArcGIS Pro’s chart feature, I created a histogram (Figured 2) showing the distribution of tree density (basal area in square feet per acre) across the entire state. I also calculated the mean, median, and standard deviation. I would repeat this same process and create a histogram and statistics for each ecoregion later in the analysis.

Analyzing the Relationship Between Elevation and Tree Density

I predicted during the early stages of project planning that elevation would be a good predictor of tree density, at least within local regions. I created a method for assessing this relationship and started by performing regression analysis on the data statewide. I would later use the same method within ecoregions to assess the relationship locally.

I first created a set of 50,000 random points throughout the state, with a stipulation that no point be within 500m of the next, to avoid having two points in the same raster cell. I then used the Extract Multi Values to Points tool to get the values of tree density and elevation from their respective rasters and add them as attributes to the random points features. Lastly, I plotted the relationship between elevation and tree density at these points on a scatterplot (Figure 3), also showing the linear trend and calculating the R2 value. *Elevation turned out to not be a good predictor of tree density at the state level (see Results for discussion).*

Performing Analysis of Tree Density Per Ecoregion

I also predicted during planning that ecoregions would likely be very good predictors of tree density potential for locations within the same ecoregion. Unfortunately, I was not able to quickly identify a way to perform the next steps in a batch process or automated fashion. While I know it must be possible (perhaps using ArcPy), I decided to perform the following steps manually so that I could inspect the data at each step of the way while developing the process. I also chose to move on with the project and not get stuck spending more time attempting to automate the steps than it would take to just perform them manually at this scope.

I selected features from the ecoregions data set by specific ecoregion at Level IV granularity – many Level IV ecoregions are made up of multiple polygons. I created new groups and layers based on the selected features for each of the 57 ecoregions, then used those layers as the mask with the Extract by Mask tool to create a raster of tree density for each ecoregion.

I then performed the same steps identified earlier to analyze the distribution of tree density within each ecoregion (see Appendix for histograms of each ecoregion), and also created statewide maps of the mean, median, range, and standard deviation of tree density at the ecoregion level (Figures 4–7).

To be able to analyze the relationship between elevation and tree density within an ecoregion, I selected points from the random points layer (which by this time also had associated tree density and elevation attributes) and made a new layer within the ecoregion group from the selection. I then performed the same steps identified earlier to plot the relationship and calculate the trend and R2 value (Figures 8–13). *Elevation turned out to not be a good predictor of tree density at the Level IV ecoregion level (see Results for discussion).*

Calculating Tree Deficits

Within each ecoregion, I set a “goal” measure of basal area at each location to be the mean of all locations within the ecoregion plus one standard deviation. I could calculate the tree deficit for each location using its current value compared to the goal value, setting the deficit to 0 if the current value already met or exceeded the goal.

Additionally, I wanted to set the deficit to 0 for any areas covered by water, snow, or ice, since they would likely not support any increase in trees regardless of the current value recorded for that cell (usually 0). To facilitate this, I used the raster calculator on the land cover data with a conditional expression to create a raster with the value set to “0” if the type for that location was water, snow, or ice, and “1” if anything else.

For each ecoregion, I used the raster calculator to evaluate the map algebra expression:

$$d = \text{Con}((g - b) < 0, 0, (g - b) * w)$$

where d is the calculated deficit value, g is the goal, b is the current basal area, and w is the “water, snow, or ice” multiplier.

In other words, the deficit for each location is 0 if the goal minus the current basal area is less than 0 or if that location is covered by water, snow, or ice. Otherwise, the deficit for each location is the goal minus the current basal area.

I performed this raster calculation for each ecoregion, then mosaiced those 57 rasters together to create one raster of tree deficits for the entire state.

Results

Tree Density Across the State of Washington

Examining the first result of analysis, the map of tree density across the State of Washington (Figure 1), I came across few surprises. The heavily forested area in the Olympic Mountains and coastal temperate rainforest, with its massive Douglas firs, hosts the highest trunk area figures, up to 564 square feet per acre. Traveling east, numbers are depressed in low-lying, built-up urban areas, increase again through the Cascades (except for glaciated peaks) and northern Rockies, and fall to zero in much of the arid Columbia Plateau of southeast Washington.

Tree Density in the State of Washington by Trunk Area

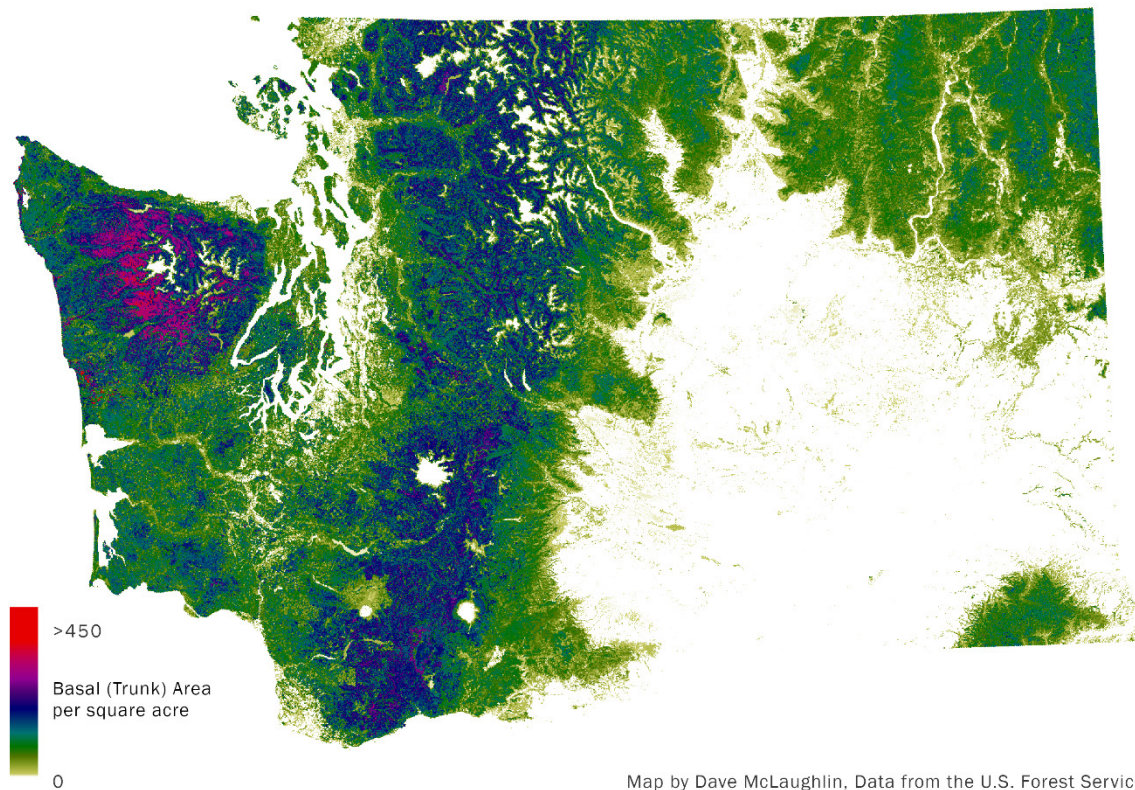


Figure 1: Mapping tree presence in the State of Washington, measured in basal (trunk) area (square feet per acre), after combining data from 324 species.

The distribution throughout the state (Figure 2) is skewed by all the “0” cells from the plateau, but a more thorough analysis of individual ecoregions will yield more readable results.

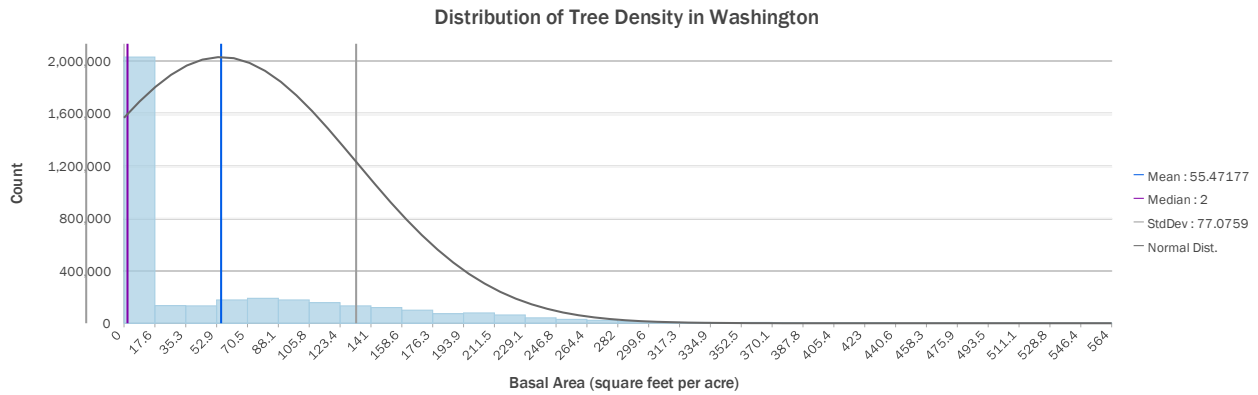


Figure 2: Plotting the distribution of tree density across the State of Washington on a histogram shows the large skewing of data by the massive number of “0” cells.

When analyzing the relationship between elevation and tree density at the state scale, I found it to be entirely insignificant, with one not being a reliable predictor of the other. Upon further examination, this is not entirely surprising – areas within the most heavily forested regions of the state share the same elevation as those in deforested urban areas and the barren plateau.

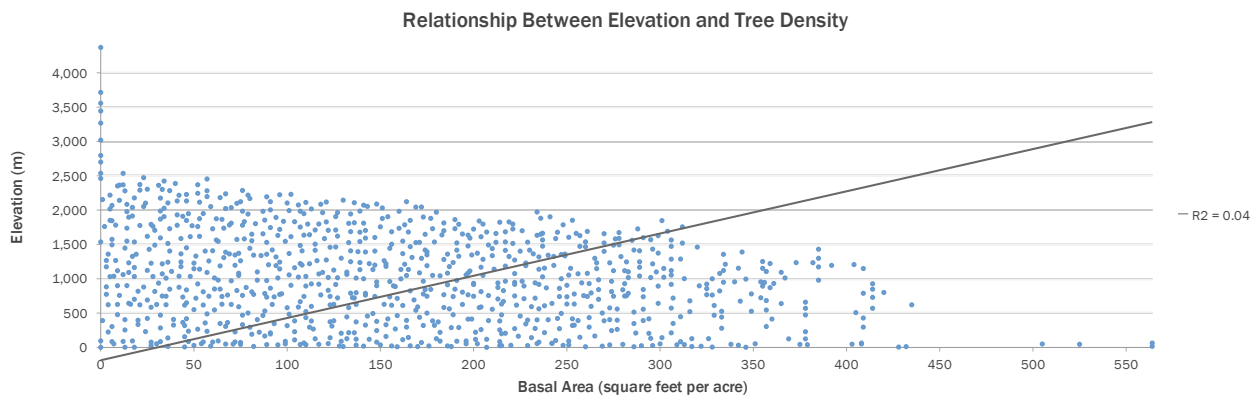


Figure 3: At the state scale, elevation has almost no relationship to tree density.

Summarizing Tree Density by Ecoregion

Summarizing the tree density data by ecoregion is useful for a broader visualization of the data, starting with the raw mean values (Figure 3). Analyzing the median (Figure 4) in the regions of the barren plateau wipes out the small stands of trees sprinkled throughout the area (which can skew the mean figures) while exhibiting little change in the more forested regions of the state.

Tree Density in the State of Washington by Trunk Area
Mean by Level IV Ecoregion

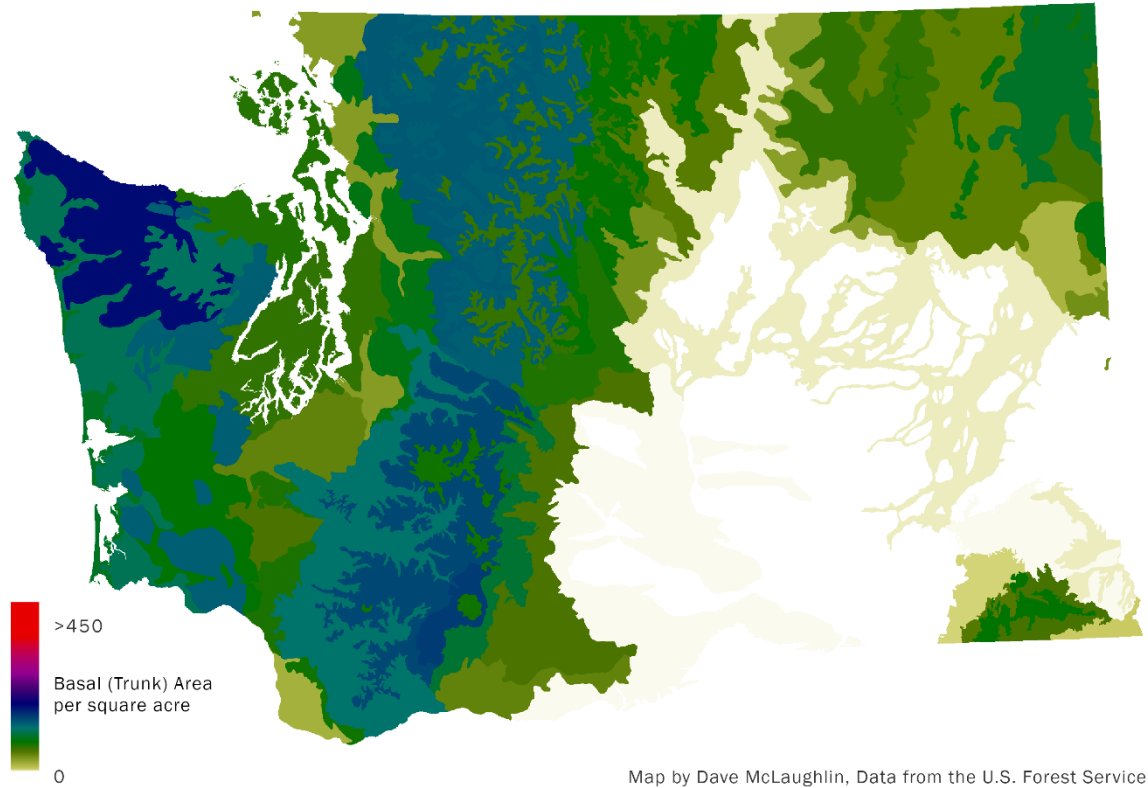


Figure 4: Summarizing tree density numbers by ecoregion yields a good broad view of the status of forests in Washington.

Tree Density in the State of Washington by Trunk Area
Median by Level IV Ecoregion

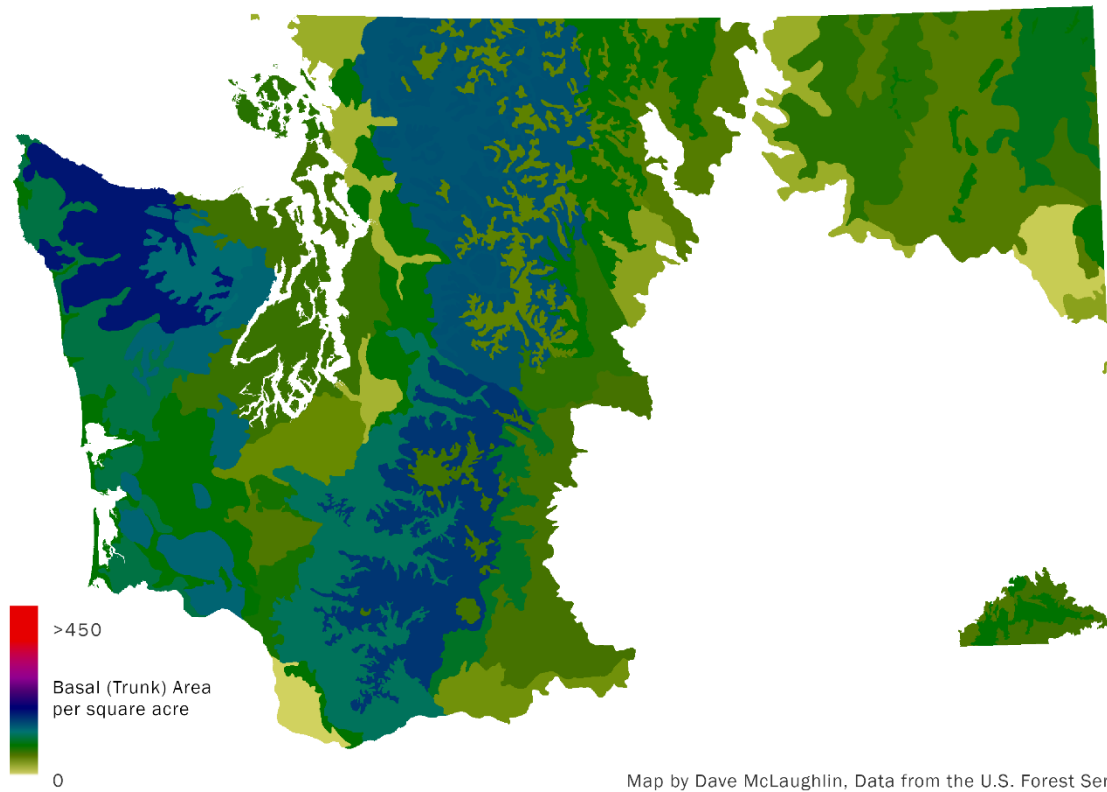


Figure 5: Mapping the median smooths out areas in the Columbia Plateau, which is mostly barren of trees.

Examining the maps of range (Figure 5) and standard deviation (Figure 6) within the ecoregions previews where we might expect to see some big figures in the deficit data, as well as possible trouble points with the model.

Tree Density in the State of Washington by Trunk Area
Range by Level IV Ecoregion

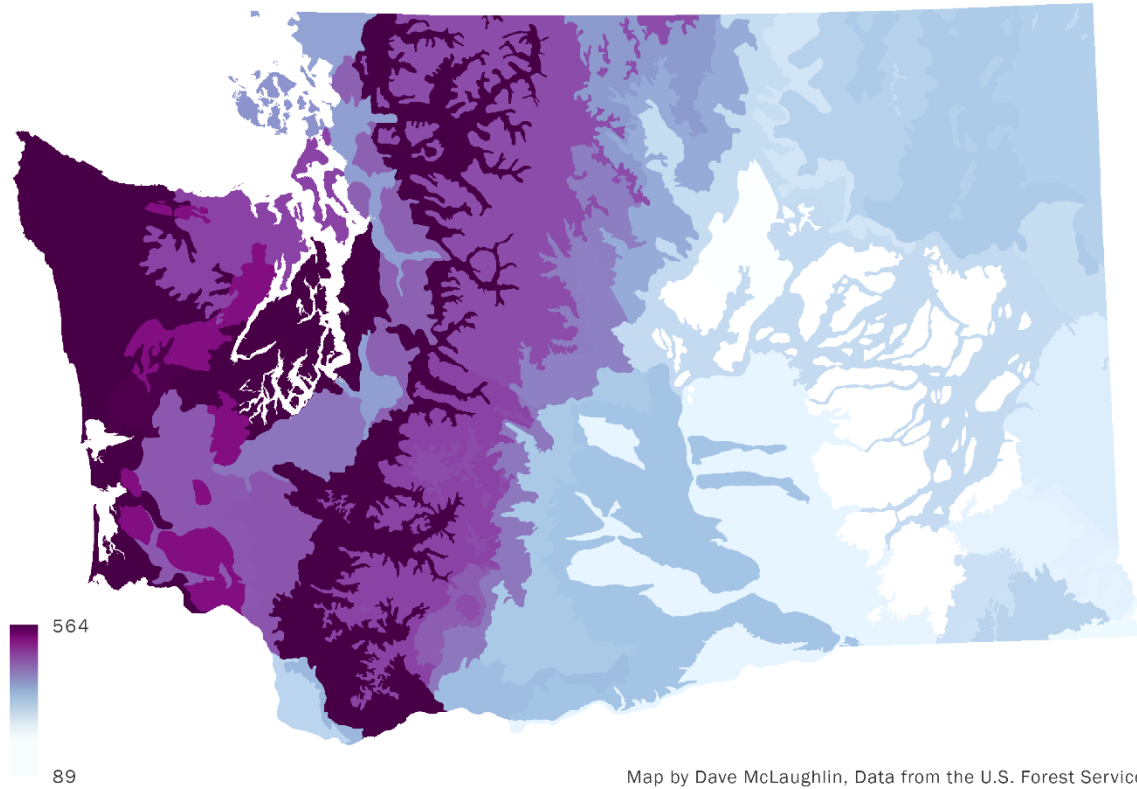


Figure 6: The Pacific Coast, Puget Lowland, and western slopes of the Cascades exhibit the greatest range in values.

Tree Density in the State of Washington by Trunk Area
Standard Deviation by Level IV Ecoregion

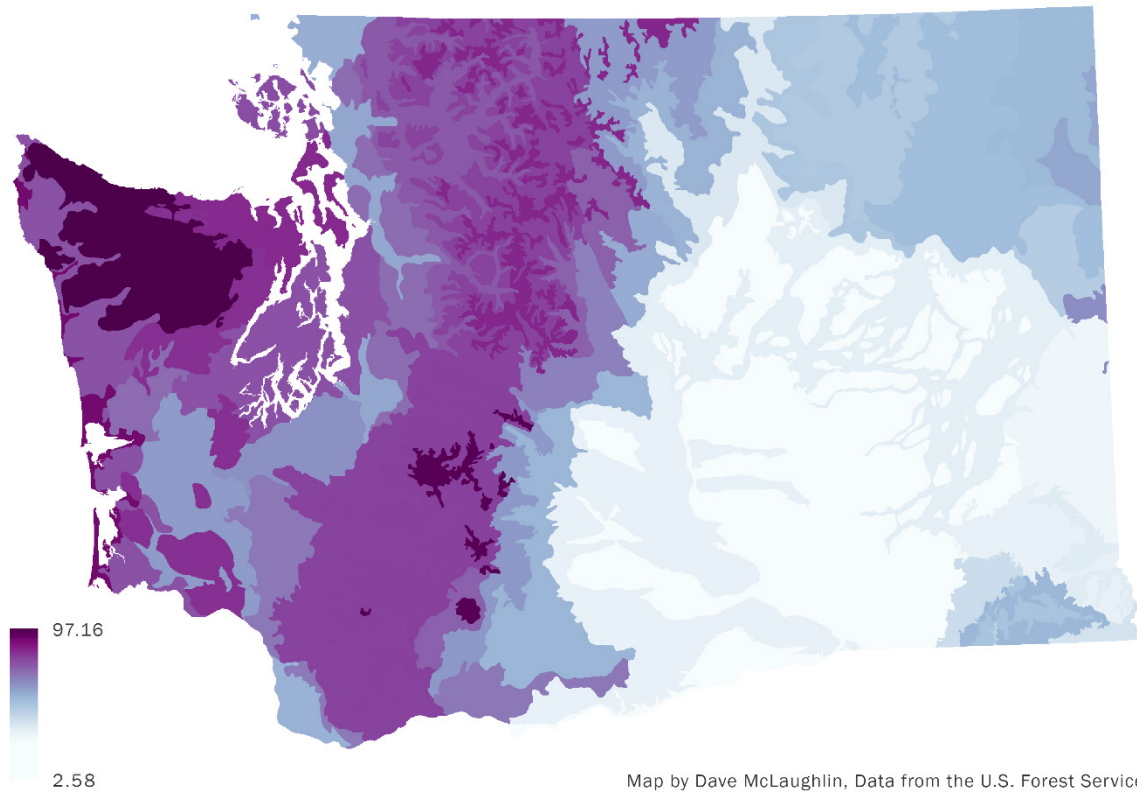


Figure 7: The greatest standard deviation numbers are found in the Olympic Mountains and some of the taller mountains of the Cascades, perhaps due to heavily forested mountainsides which give way to glaciated (barren) peaks.

I was surprised to find that even within individual ecoregions, elevation did not exhibit a strong correlation to tree density. After performing the regression analysis on six various ecoregions, only two had R2 scores that were not almost zero, after which I abandoned that route. The results of those six analyses follow.

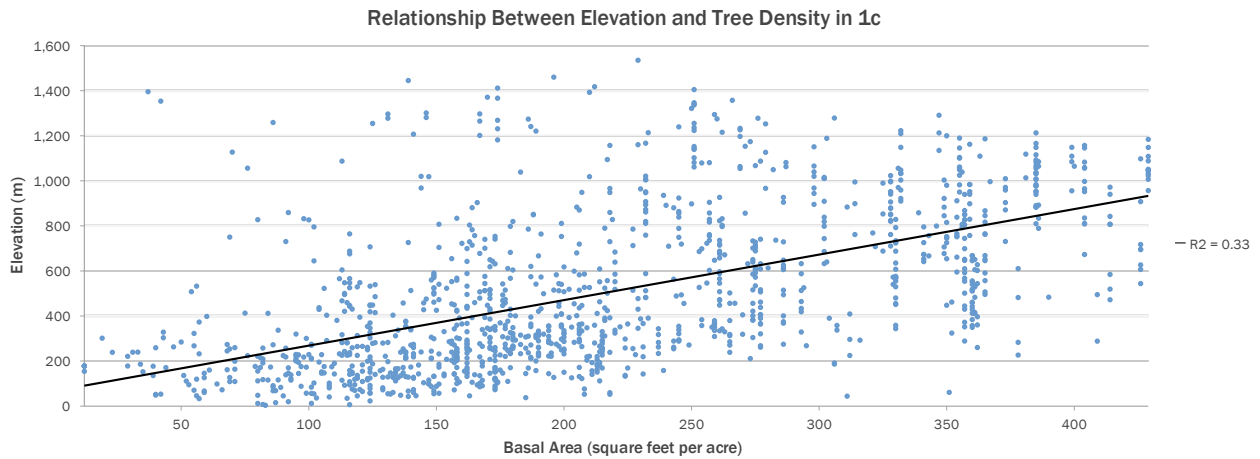


Figure 8: In the Low Olympics ecoregion, elevation has a small positive correlation with tree density. I mostly attribute this phenomenon to the fact that the lower elevations nearer the coast are home to most of the population, agriculture, and logging activity in the area, which has lowered the amount of intact forest at those lower elevations. This finding further cemented my decision to not pursue elevation as a predictor of forest potential, as it appears to be more the result of human interaction with the environment than with the environment's capacity for tree growth.

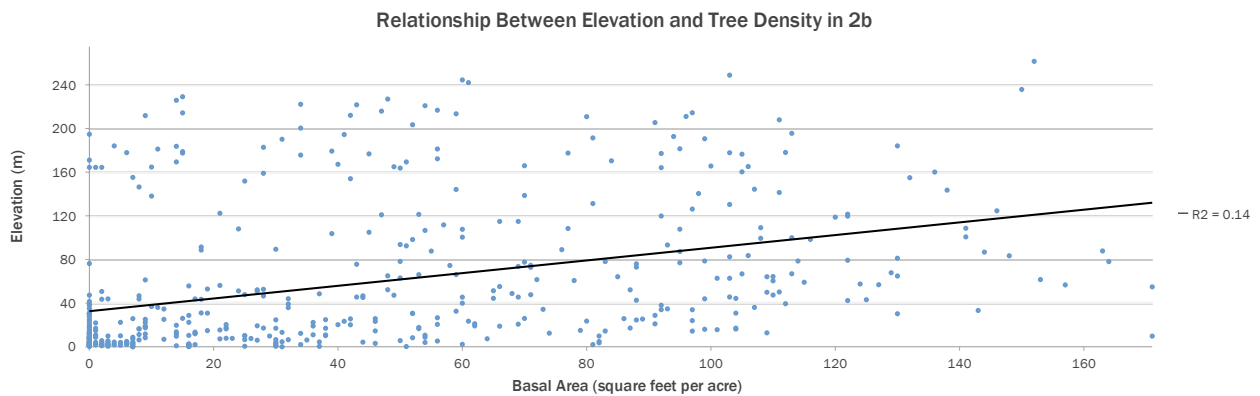


Figure 9: In the Eastern Puget Riverine Lowlands, elevation is not a strong predictor of tree density.

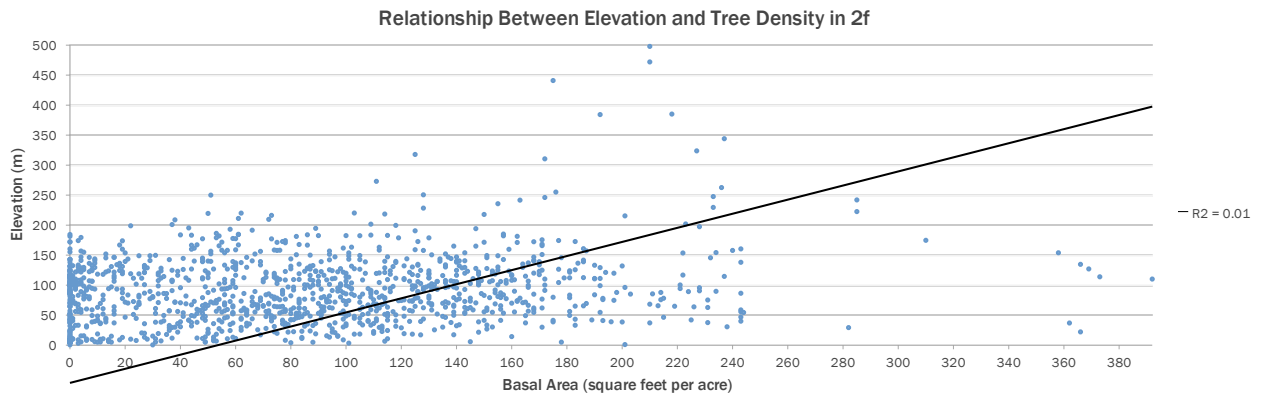


Figure 10: In the Central Puget Lowland, elevation is not at all a predictor of tree density.

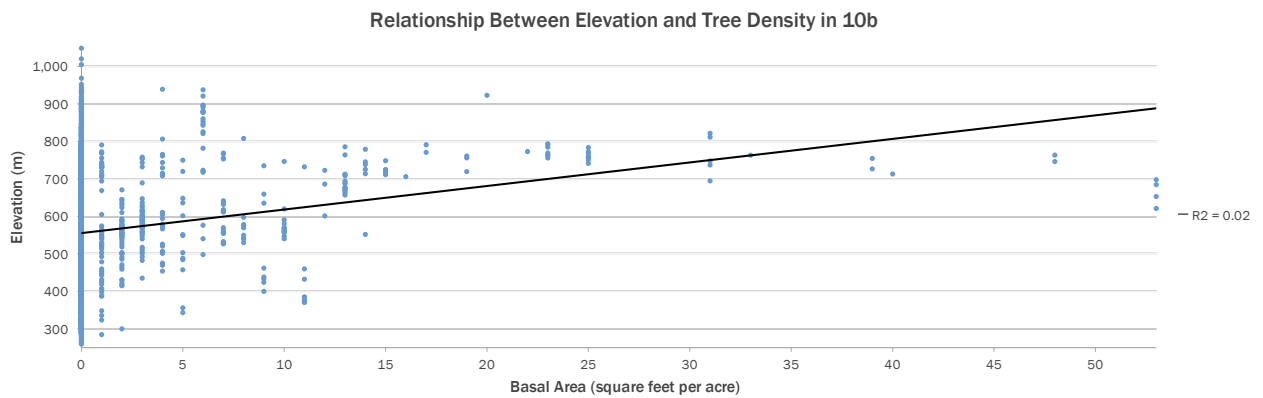


Figure 11: In the Loess Islands of the Columbia Plateau, elevation is not at all a predictor of tree density.

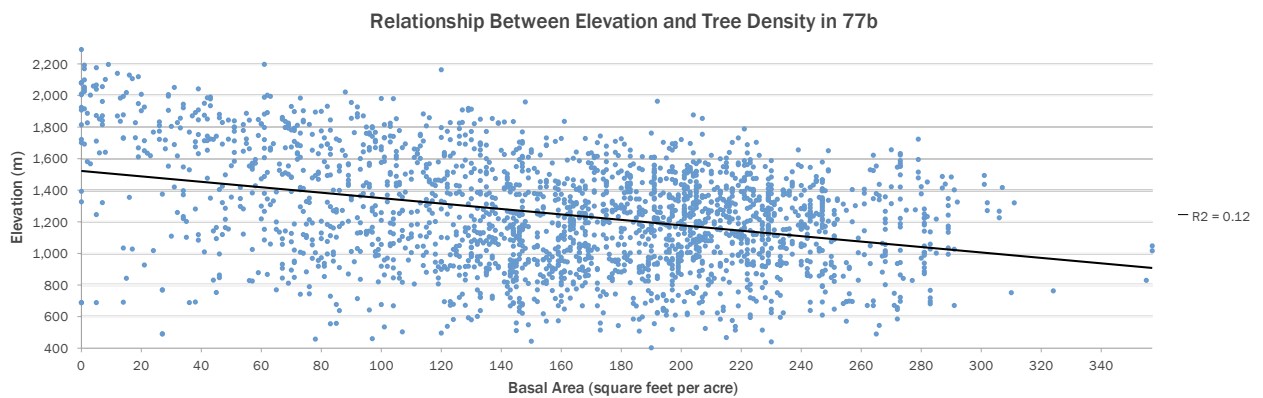


Figure 12: In the North Cascades Highland Forests, elevation is not a strong predictor of tree density.

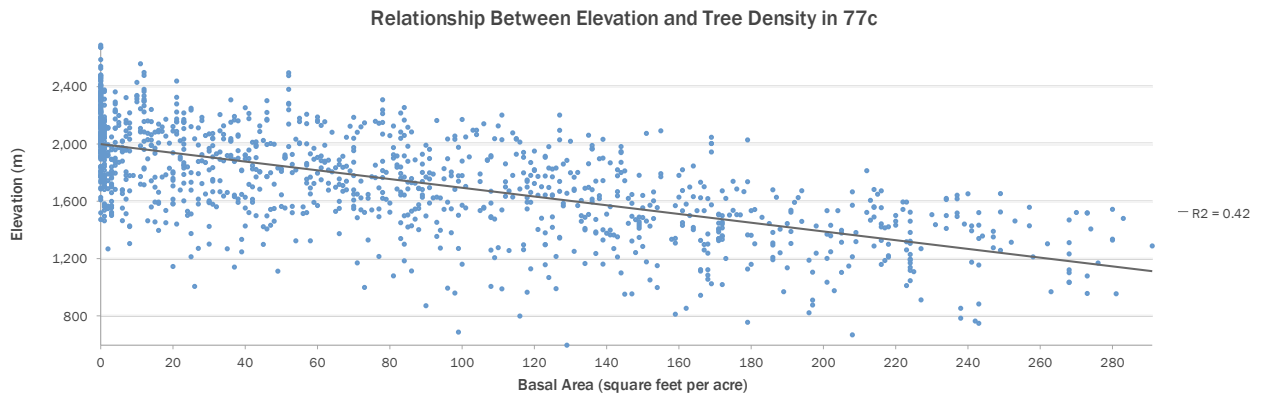


Figure 13: The North Cascades Subalpine/Alpine ecoregion exhibited the greatest correlation (negative, in this case) between elevation and tree density in my analysis, due to the inclusion of glaciated peaks above the treeline. In the end, this phenomenon was more easily and accurately accounted for using land cover data.

Tree Deficits Across the State of Washington

The main result of this project's analysis, a map of tree deficits (Figure 14), highlights a number of areas exhibiting large differences between their current and potential forestation levels. I next will examine those areas in more detail, discovering that many appear to be accurate descriptions of the state of the forest, while others highlight flaws in the model.

Tree Deficit in the State of Washington by Trunk Area

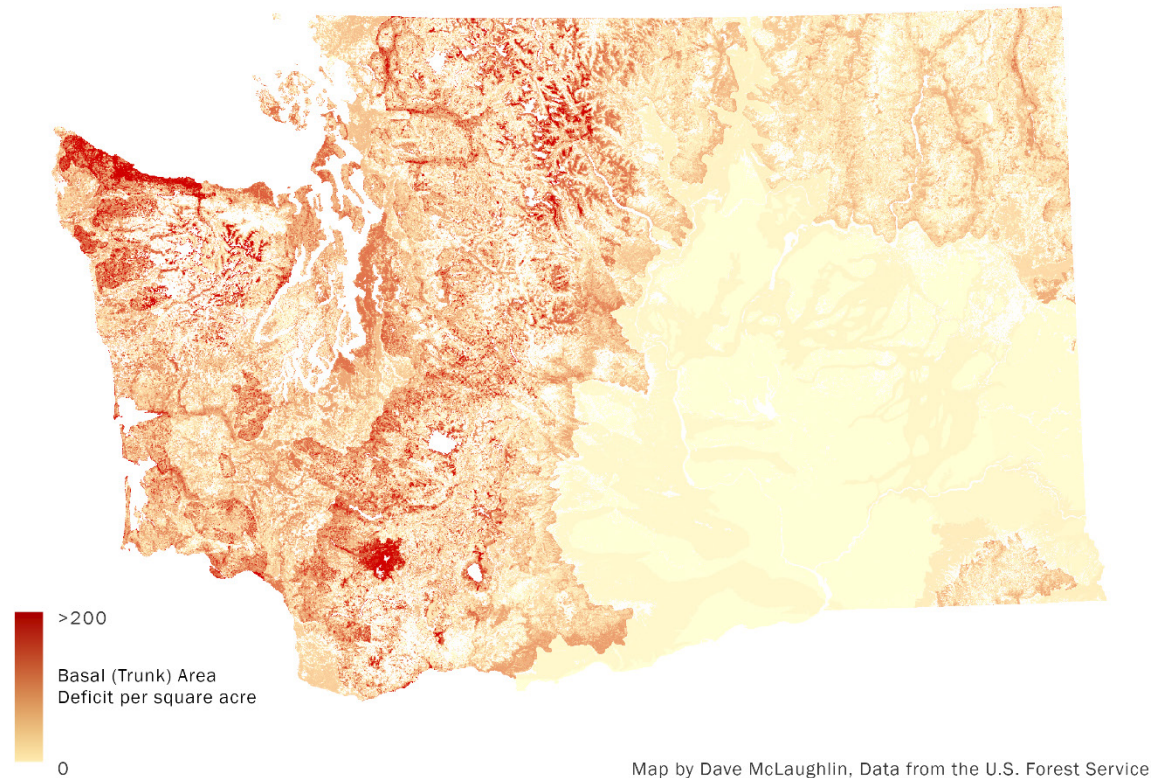


Figure 14: Tree deficits in Washington, measured in the difference between the current square feet of trunk area per acre and the potential square feet of trunk area per acre in that ecoregion.

Olympic Peninsula

The sharp bands of large deficits in a few mountainous areas are likely caused by ecoregion edge effects and a defect in the model, which I will discuss further later. However, the rest of the tree deficits identified by the analysis on the Olympic Peninsula (Figure 15) appear to be caused by human use of the land (Figure 16). The area surrounding the mountains (and Olympic National Park) is mostly managed forest where logging has occurred, causing tree deficits (see the cleared areas visible in satellite imagery in Figure 17). Even more pronounced are the areas of human settlement (mostly urban and agricultural) where some trees are present, but not in intact forests.

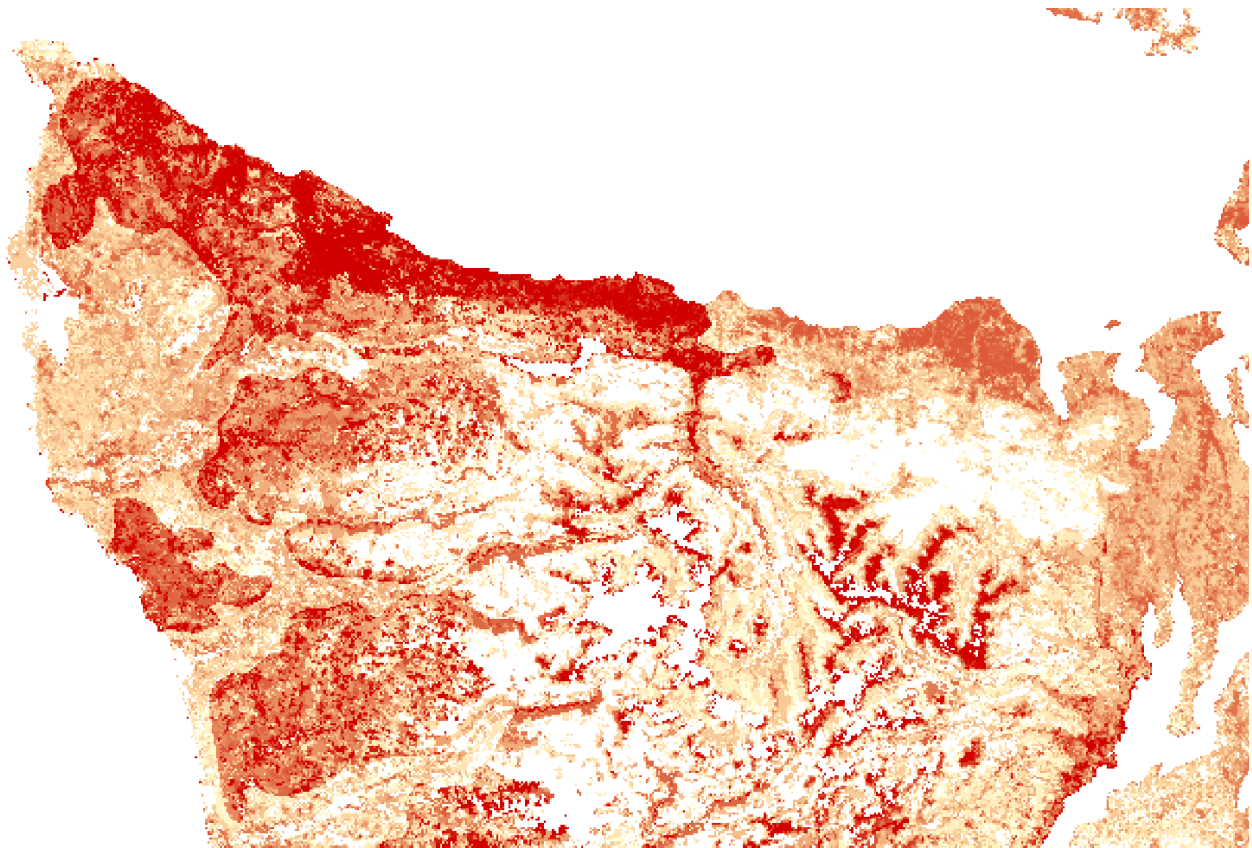


Figure 15: Close-up of tree deficits on the Olympic Peninsula.



Figure 16: Satellite imagery of the Olympic Peninsula. Note the built-up area around Port Angeles and the clear-cut areas of forest in the west peninsula area.

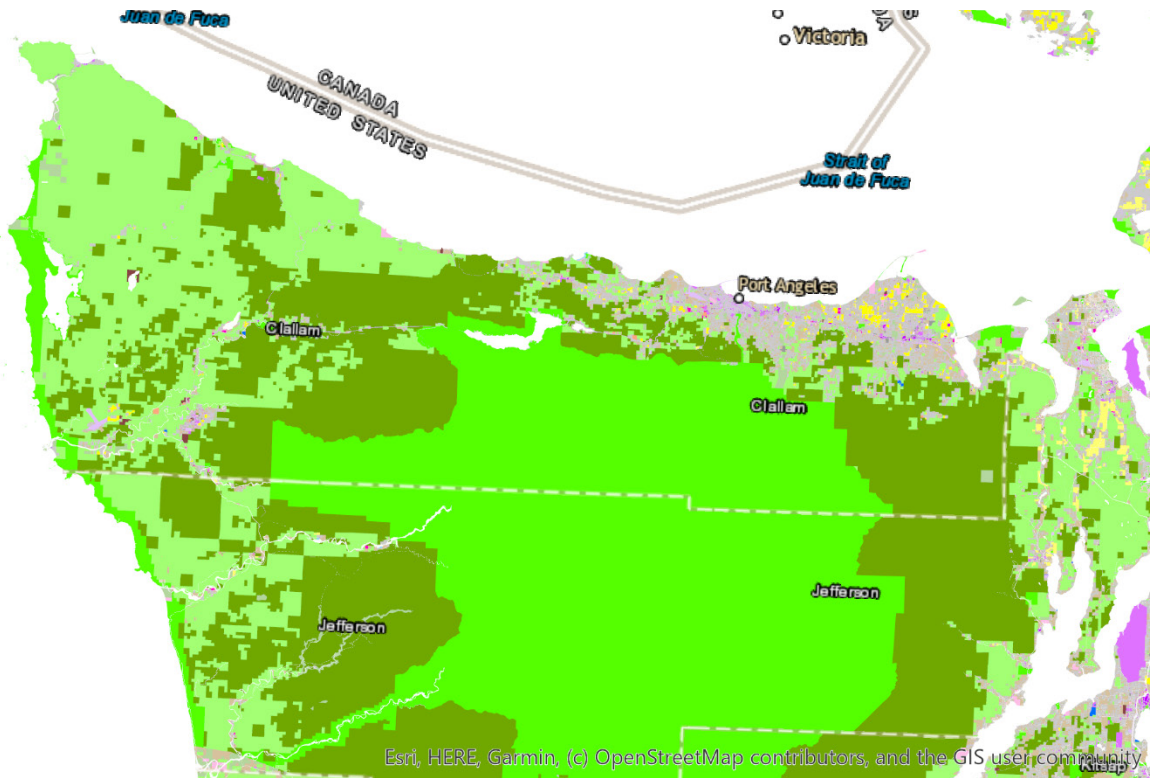


Figure 17: Land use in the Olympic Peninsula. Notable uses: parks (bright green), managed forests (public are dull green, commercial are light green), built-up urban area (gray), industry (purple), and agriculture (yellow). There would be differing considerations in reforesting each of these land use types, but the human-changed landscape has caused the areas with the theoretical greatest potential increase in tree trunk area.

Mount St. Helens

The area surrounding Mount St. Helens' peak is a bright red spot on the map of tree deficits. Considering the eruption almost forty years ago and its decimation of the surrounding forests, it is no surprise that the analysis identifies the area as having a large tree deficit. Out of the scope of this analysis, it would take a biologist to determine the true potential of this area to regrow a vigorous forest.

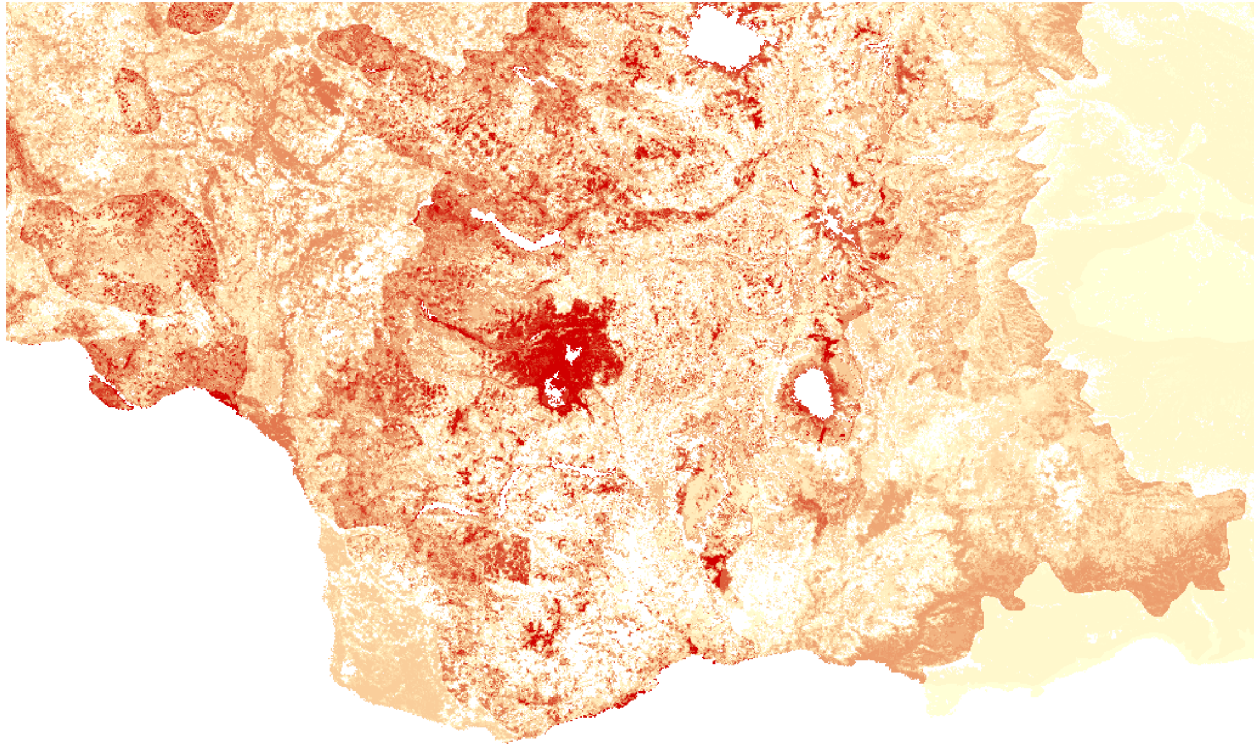


Figure 18: Tree deficits in southwest Washington, centered on Mount St. Helens.

Seattle-Bellevue-Tacoma Metropolitan Region

Not surprisingly, the analysis identifies the built-up Seattle-Bellevue-Tacoma region as having a large tree deficit (Figure 19). However, I am positive that the analysis underestimates the true deficit of trees in this region. With the metropolitan region as large as it is, the surrounding ecoregion has so few forested areas remaining from which to calculate a potential tree trunk area number.

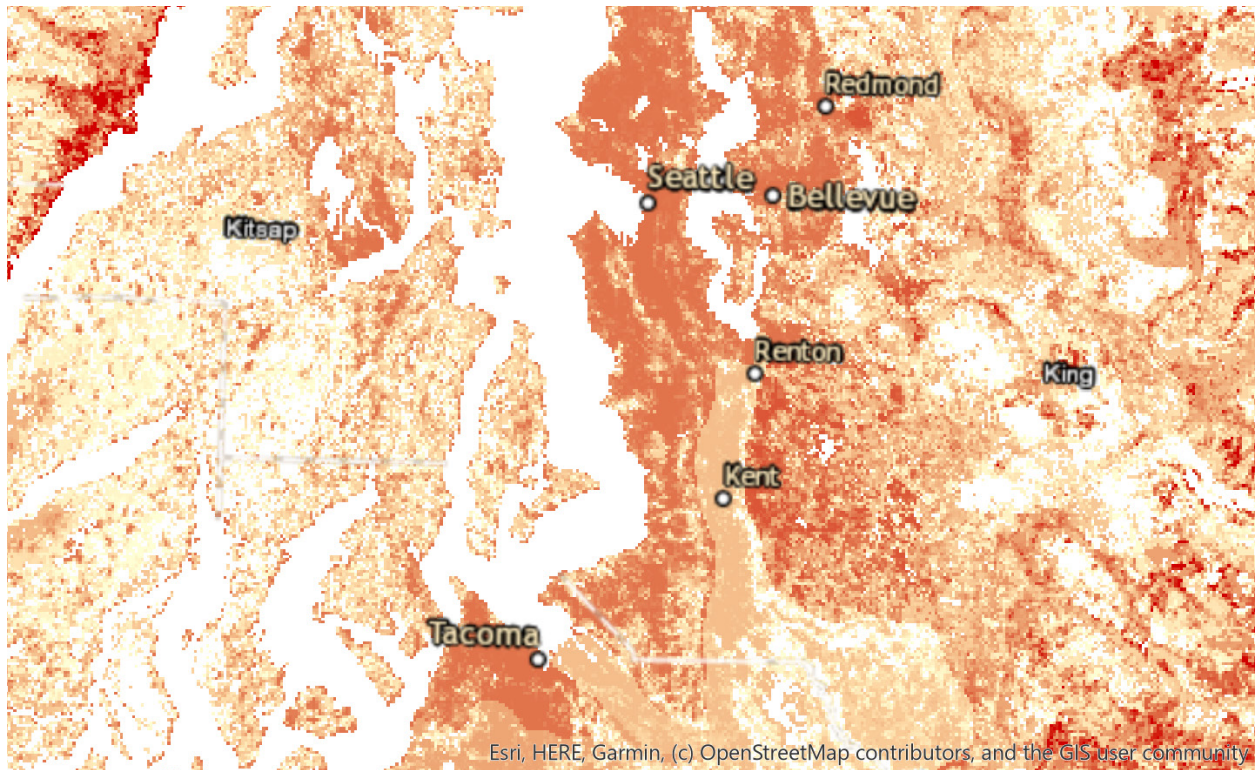


Figure 19: Tree deficits in the Seattle-Bellevue-Tacoma metropolitan region. Most of the region is devoid of forest, making the calculation of forest potential difficult.

Skagit Valley

The Skagit Valley and other similar valleys in the area show up in the analysis with significant tree deficits (Figure 20). Satellite imagery (Figure 21) also exposes some of the limits of the analysis method however, the edges of the ecoregions become clear with false hard edges in the tree deficit numbers. Surely the area around Mt. Vernon to the west could support tree growth similar to that in the Skagit Valley, yet its deficit appears to be much smaller because its ecoregion is mostly developed and the potential tree area (goal number) was set lower than what is likely its true potential.

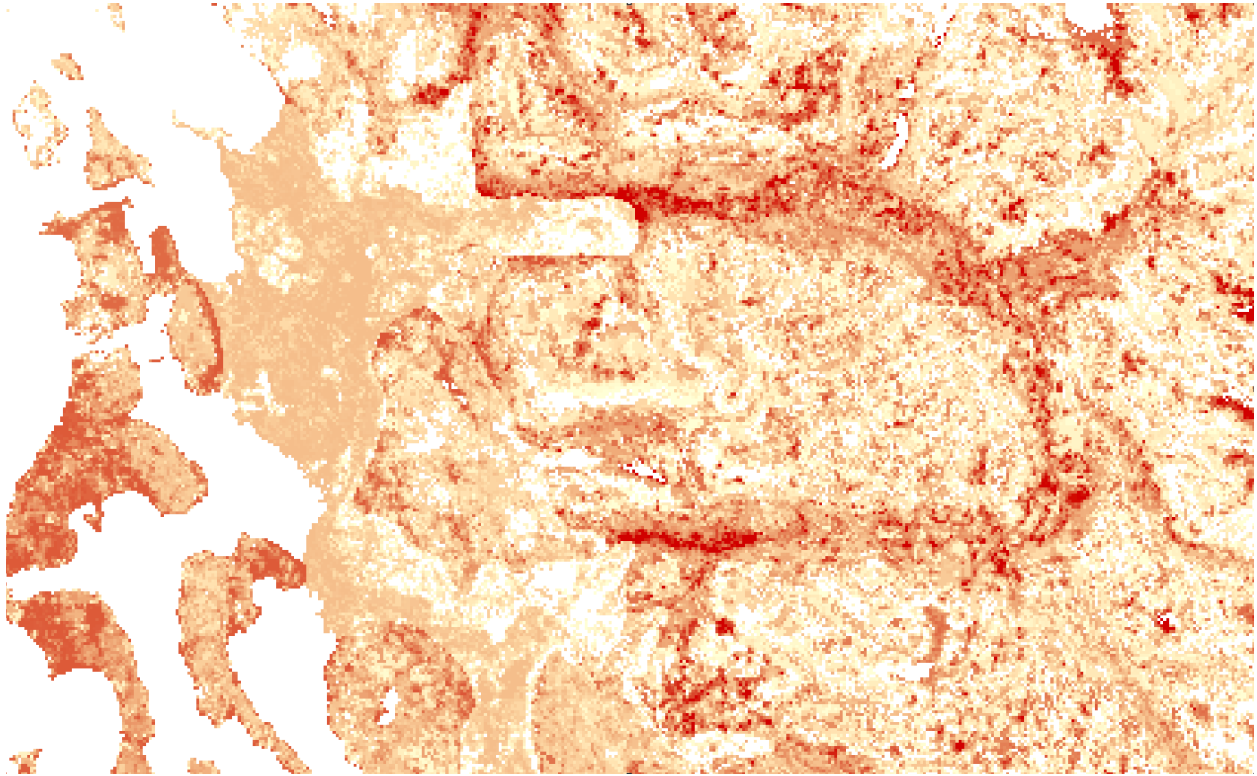


Figure 20: The analysis highlights significant tree deficits in the Skagit Valley.

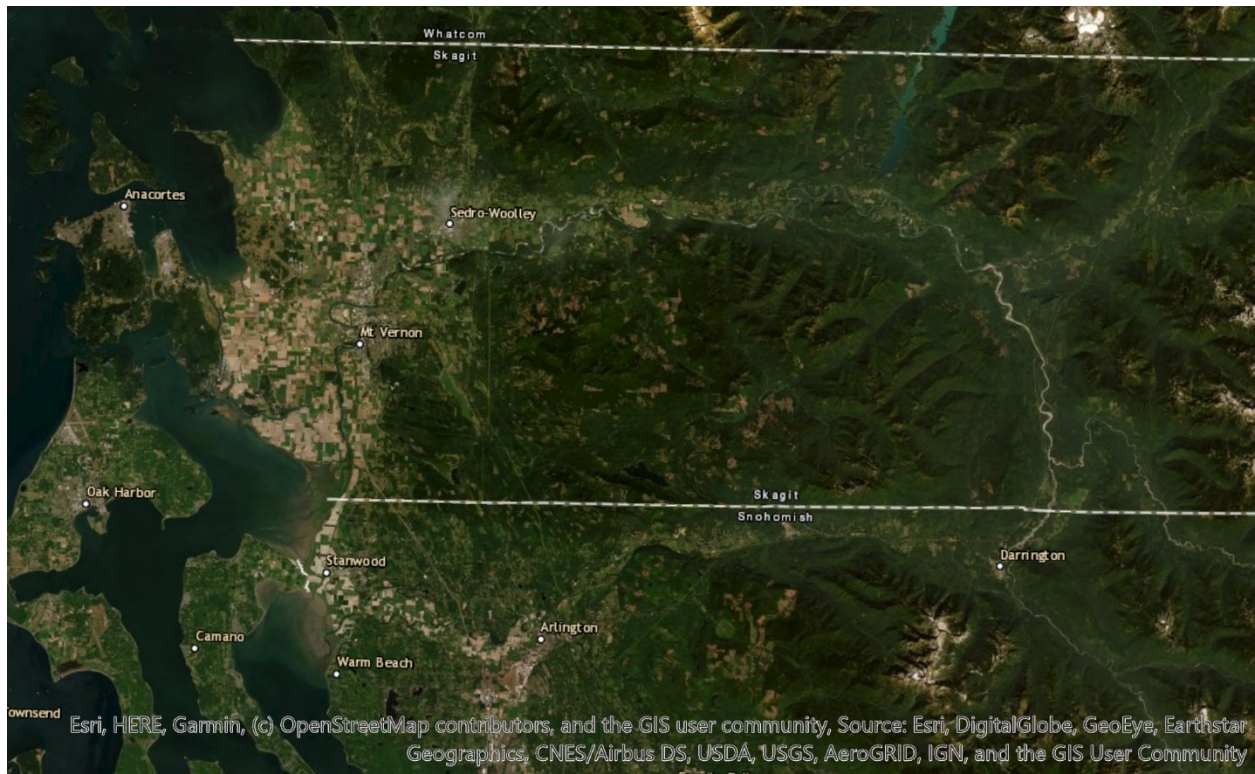


Figure 21: Satellite imagery confirms the proliferation of agriculture and other human influence in the valley.

Checkerboard Logging Patterns

Checkerboard patterns of clear-cut logging appear in the analysis in a number of places (Figure 22), and are confirmed with satellite imagery (Figure 23). Land use data denotes this land all as managed forest, but the way each parcel has been used over the last decades has more bearing on its tree deficit than its use designation, elevation, or ecoregion.

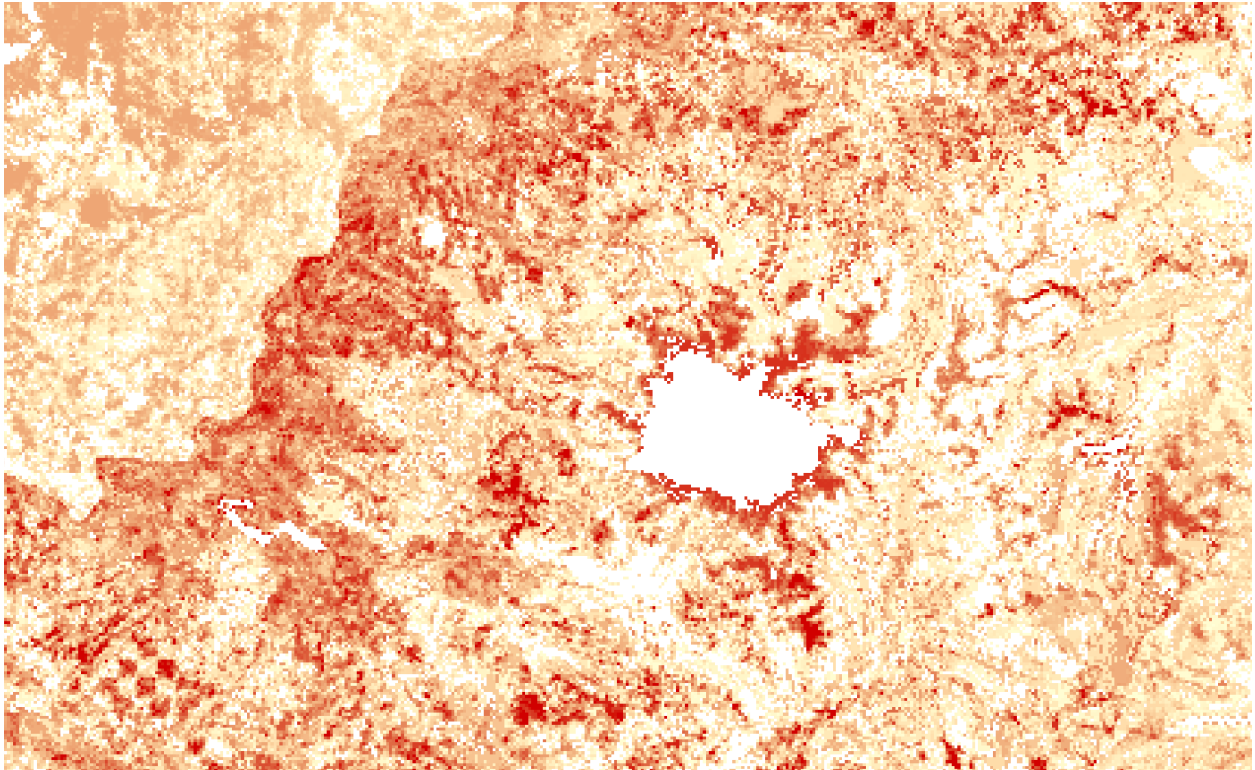


Figure 22: Tree deficits around Mt. Ranier (“blank” spot near the center). Note the checkerboard patterns of clear-cut logging in the lower-left and upper-right of the image.



Figure 23: Satellite imagery confirms the checkerboard pattern of logging which appears in the analysis in Figure 22.

Reflection

I am happy with the creation of the analysis method and its main output of a tree deficit map of the State of Washington. However, the analysis falls apart at local levels. To improve it, I would consider implementing some sort of process to smooth over the edges of ecoregions (Figure 24).

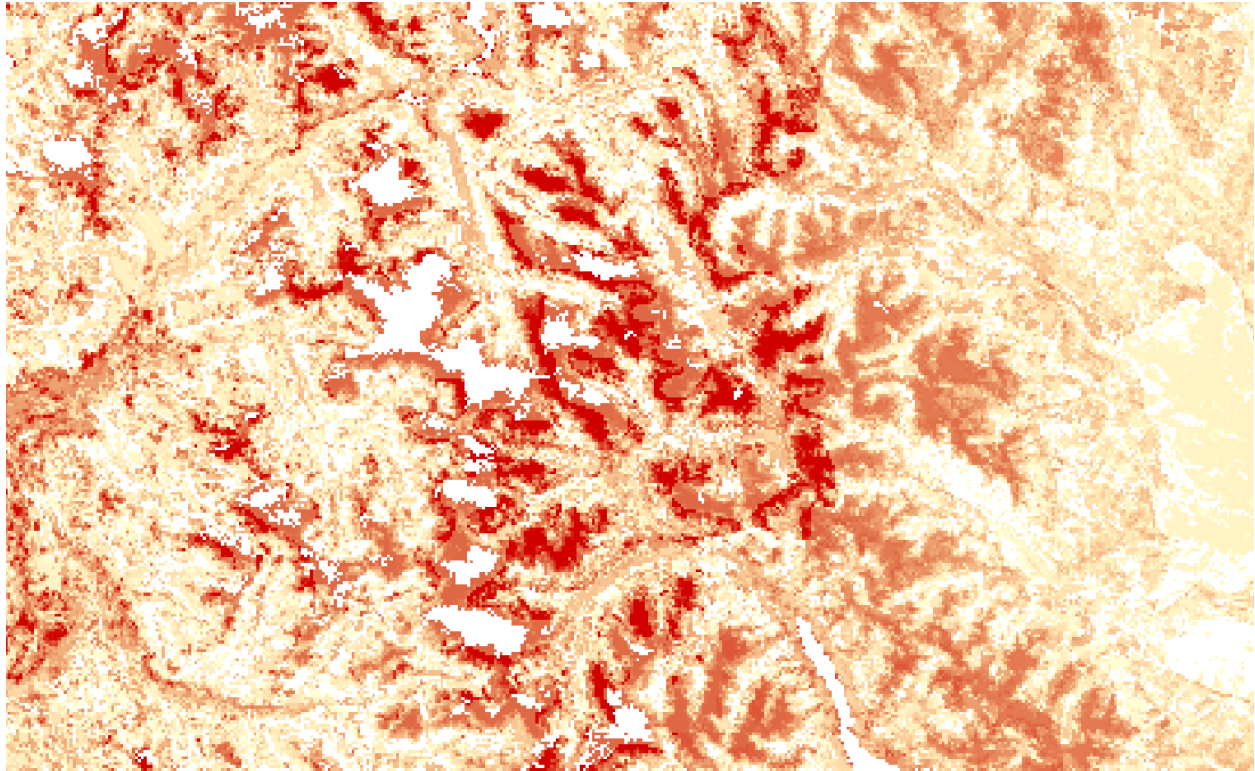


Figure 24: While perhaps difficult to interpret, this area in the Northern Cascades demonstrates a failing of the current analysis method. The areas identified as having the greatest tree deficits are the highest elevation areas of a particular ecoregion, where the trees are naturally starting to thin out but where it has not entirely transitioned to the area covered with snow or ice at the top of the mountain (and therefore excluded from showing a deficit). The analysis compares the current levels of trees in these locations to areas slightly further down the mountain with much thicker forests of larger trees. Calculating the deficit at every location within the ecoregion based on the same “goal” number leads to these false deficits.

With more geostatistical analysis it might be possible to either:

1. interpolate potential tree density across the state or ecoregion based on values at many points (either observed current data or calculated), or
2. describe potential tree density at any location with an equation.

Either of the above approaches would yield a continuous field of values instead of calculating a single value for each ecoregion, and theoretically could more accurately and smoothly describe potential tree growth increases.

I spent more time than I anticipated creating the method for analyzing the data to come up with the tree deficit values, and therefore did not have as much time as I had hoped for analyzing specific

locations and their actual reforestation potential based in land use, etc. The outputs I created in the Results discussion are also therefore not up to the quality to which I aspire.

Finally, I would like to find a way to script the manipulation of the data and its analysis, to alleviate the manual process and make it more easily reproducible.

References

Level III and IV Ecoregions of Washington

For a full map of the 57 Level IV ecoregions used in the analysis, see the Level III and IV Ecoregions of Washington map, available through the U.S. Environmental Protection Agency:

ftp://newftp.epa.gov/EPADDataCommons/ORD/Ecoregions/wa/wa_eco.pdf

Other Articles and Data Referenced

Bastin, J., Finegold, Y., Garcia, C., Mollicone, D., Rezende, M., Routh, D., Zohner, C., Crowther, T. 2019. The global tree restoration potential. *Science*, 365(6448), 76–79.

Commission for Environmental Cooperation (CEC). 2010. 2010 North American Land Cover at 250 m spatial resolution. Produced by Natural Resources Canada/The Canada Centre for Mapping and Earth Observation (NRCan/CCMEO), United States Geological Survey (USGS); *Instituto Nacional de Estadística y Geografía* (INEGI), *Comisión Nacional para el Conocimiento y Uso de la Biodiversidad* (CONABIO), and *Comisión Nacional Forestal* (CONAFOR). <http://cec.org/tools-and-resources/map-files/land-cover-2010-modis-250m>

Esri. 2019. Ground Surface Elevation – 30m, extracted from the USGS’s National Elevation Dataset. Living Atlas of the World.

<https://www.arcgis.com/home/item.html?id=0383ba18906149e3bd2a0975a0afdb8e>

Esri. 2019. World Imagery, derived from many satellite and aerial sources. Living Atlas of the World.

<https://www.arcgis.com/home/item.html?id=10df2279f9684e4a9f6a7f08febac2a9>

Omernik, J.M., and G.E. Griffith. 2014. Ecoregions of the conterminous United States: evolution of a hierarchical spatial framework. *Environmental Management* 54(6):1249-1266,

<http://dx.doi.org/10.1007/s00267-014-0364-1>

U.S. Environmental Protection Agency (EPA). 2012. Level IV Ecoregions of Washington. Corvallis, OR: Office of Research and Development, National Health and Environmental Effects Research Laboratory (NHEERL). <https://www.epa.gov/eco-research/ecoregion-download-files-state-region-10#pane-45>

Washington State Department of Ecology. 2010. 2010 Statewide Land Use. GIS Technical Services.

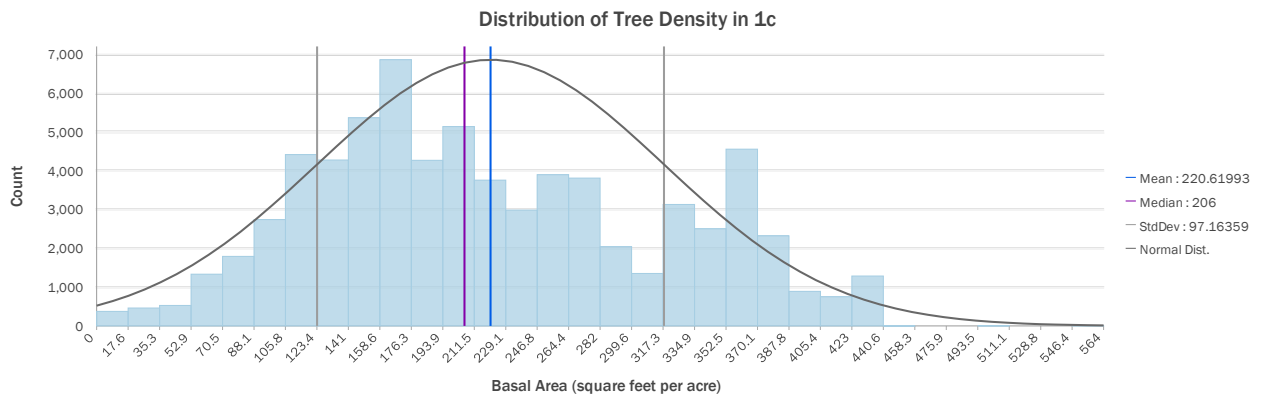
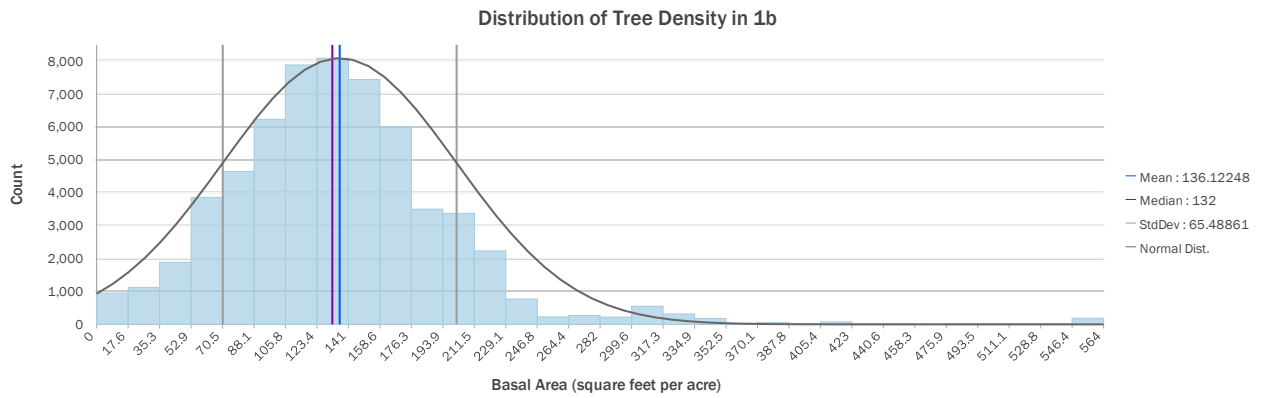
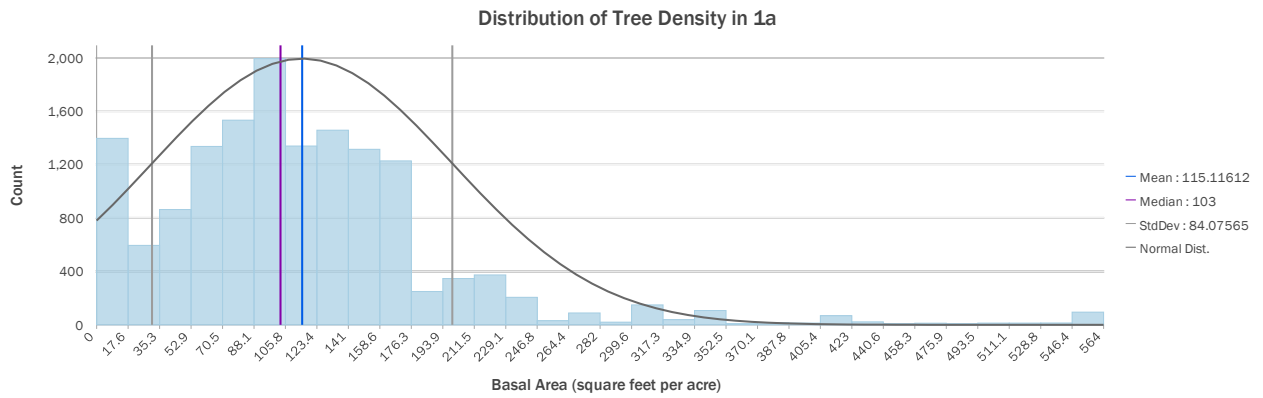
https://fortress.wa.gov/ecy/gispublic/DataDownload/ECY_CAD_Landuse2010.htm

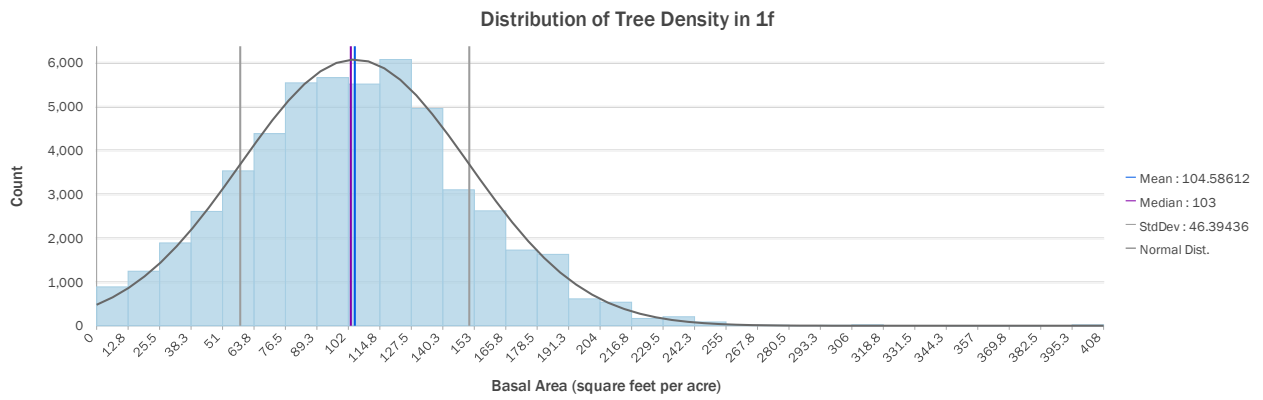
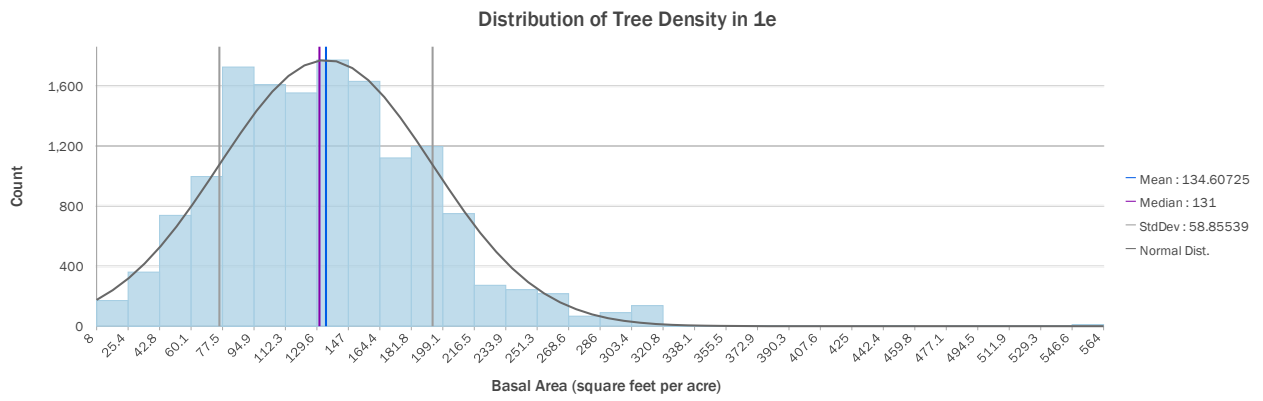
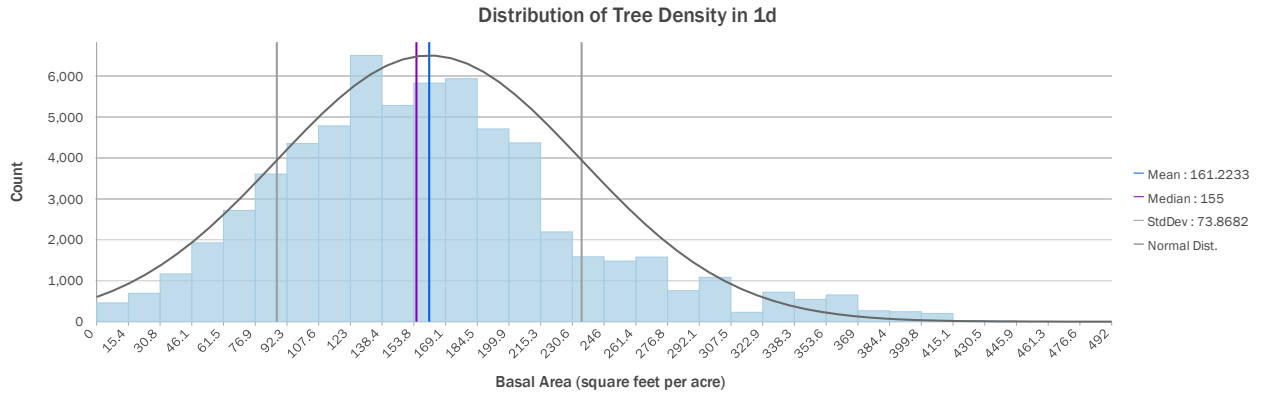
Wilson, B.T., Lister, A.J., Riemann, R.I. 2012. A nearest-neighbor imputation approach to mapping tree species over large areas using forest inventory plots and moderate resolution raster data. *Forest Ecology and Management*, 271, 182–198. <https://doi.org/10.1016/j.foreco.2012.02.002>

Wilson, Barry Tyler; Lister, Andrew J.; Riemann, Rachel I.; Griffith, Douglas M. 2013. Live tree species basal area of the contiguous United States (2000-2009). Newtown Square, PA: USDA Forest Service, Rocky Mountain Research Station. <https://doi.org/10.2737/RDS-2013-0013>

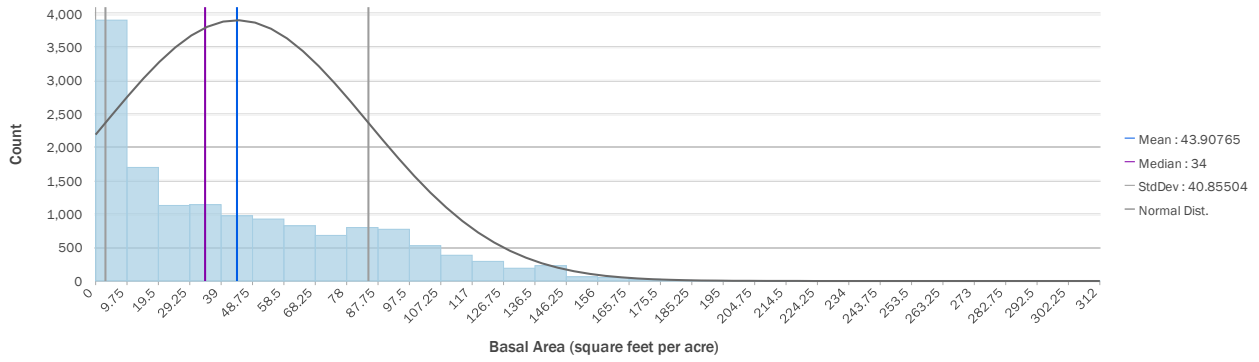
Appendix

Distribution of Tree Density by Ecoregion

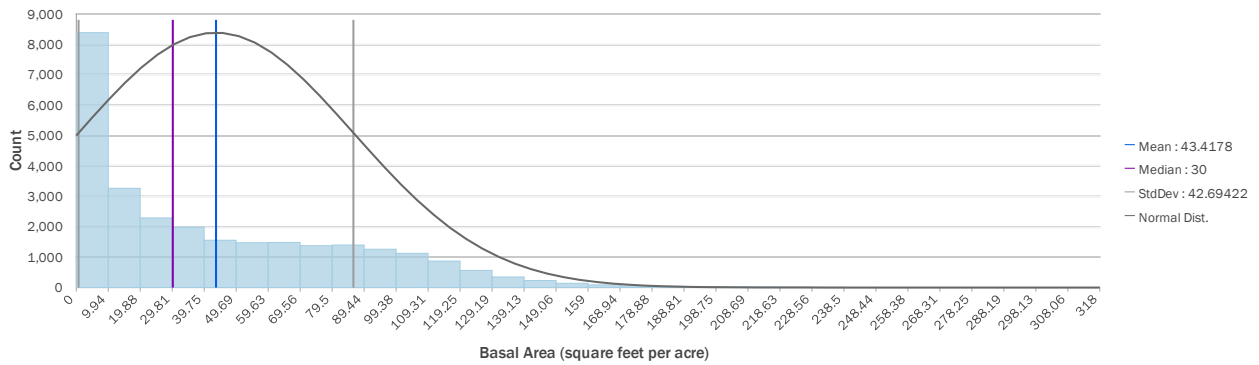




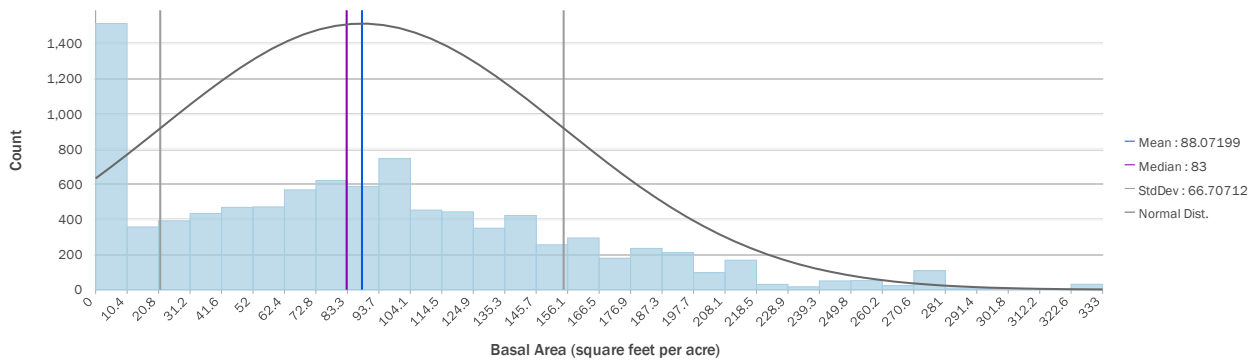
Distribution of Tree Density in 2a

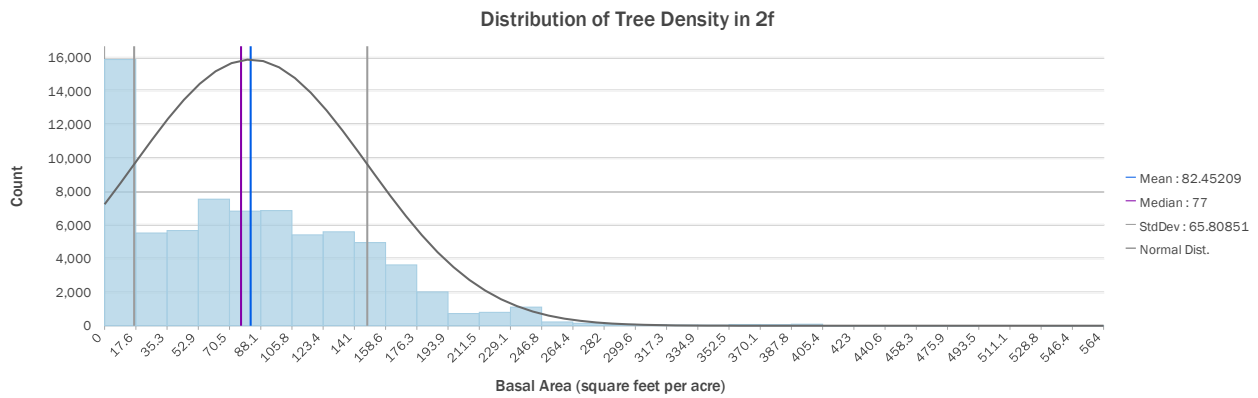
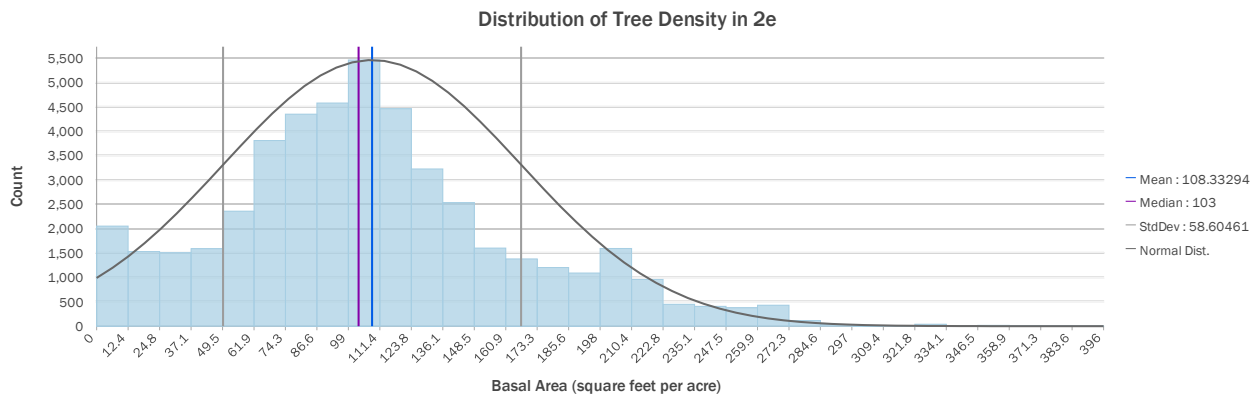
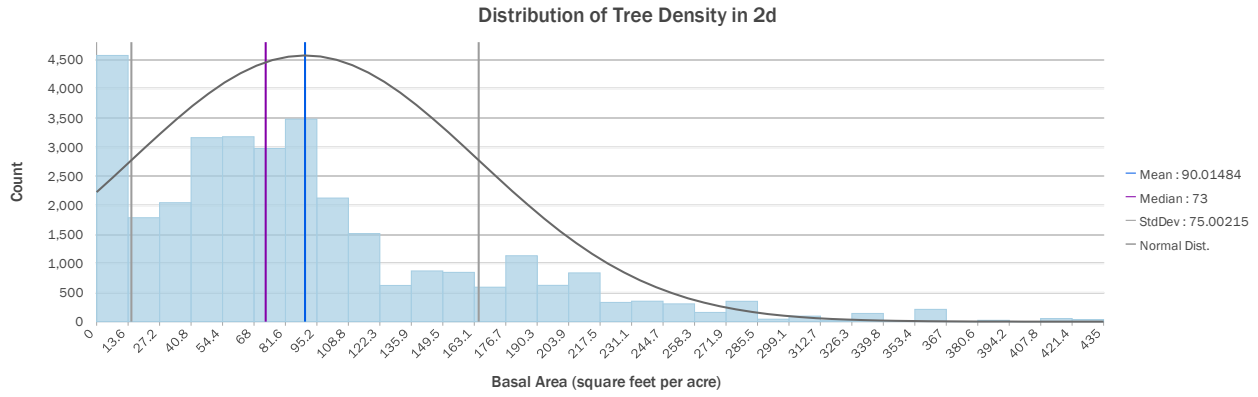


Distribution of Tree Density in 2b

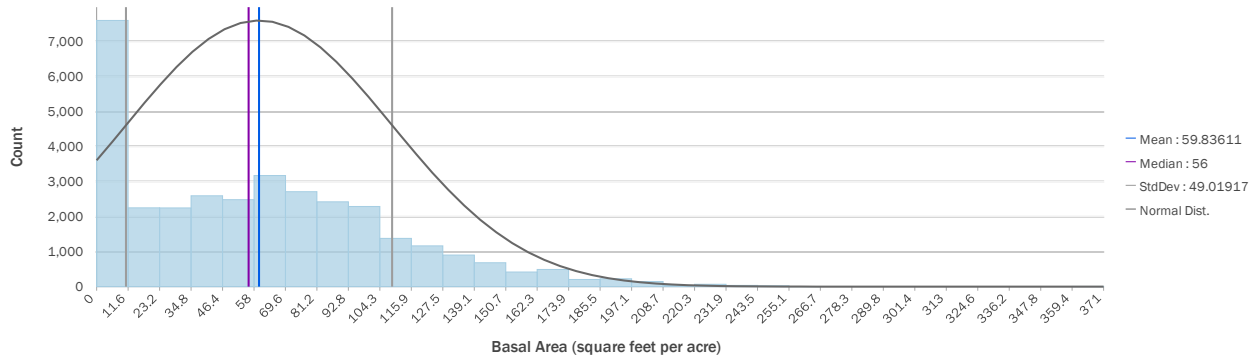


Distribution of Tree Density in 2c

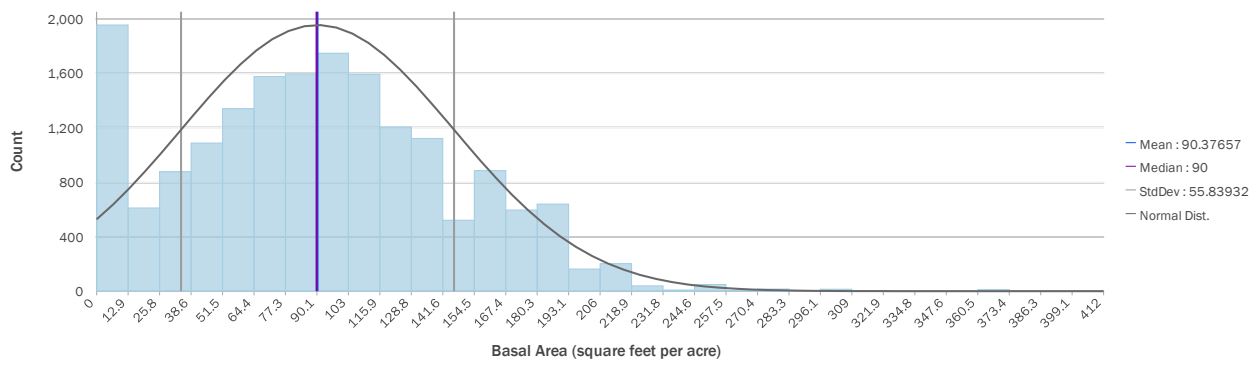




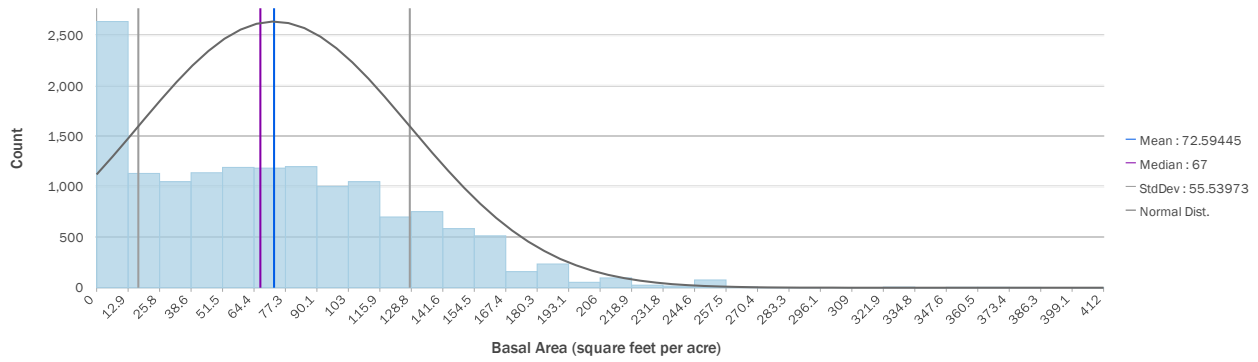
Distribution of Tree Density in 2g



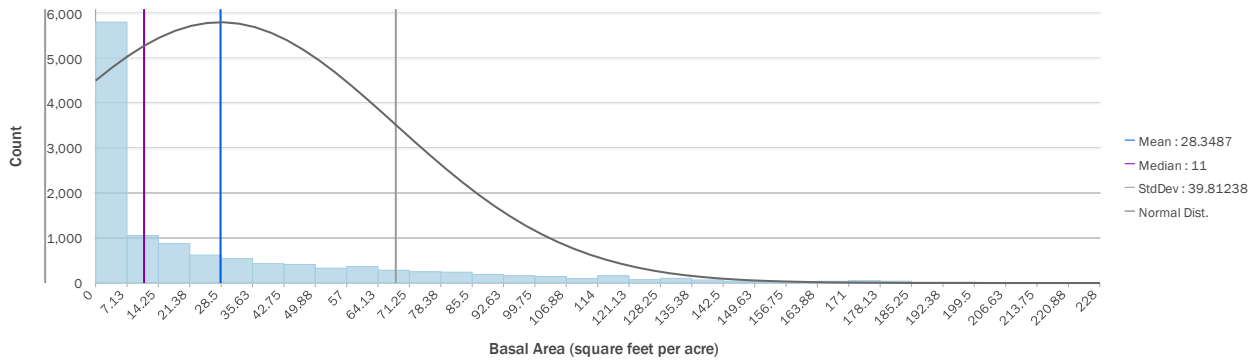
Distribution of Tree Density in 2h



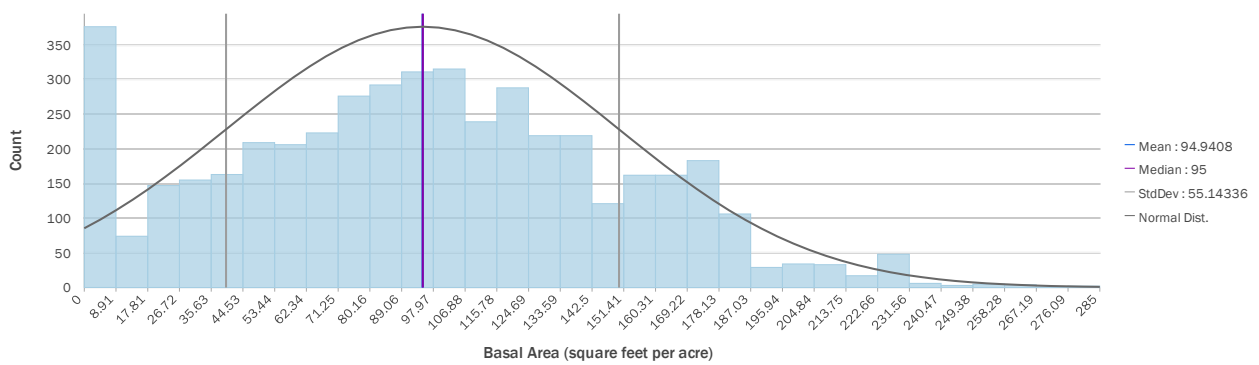
Distribution of Tree Density in 2i



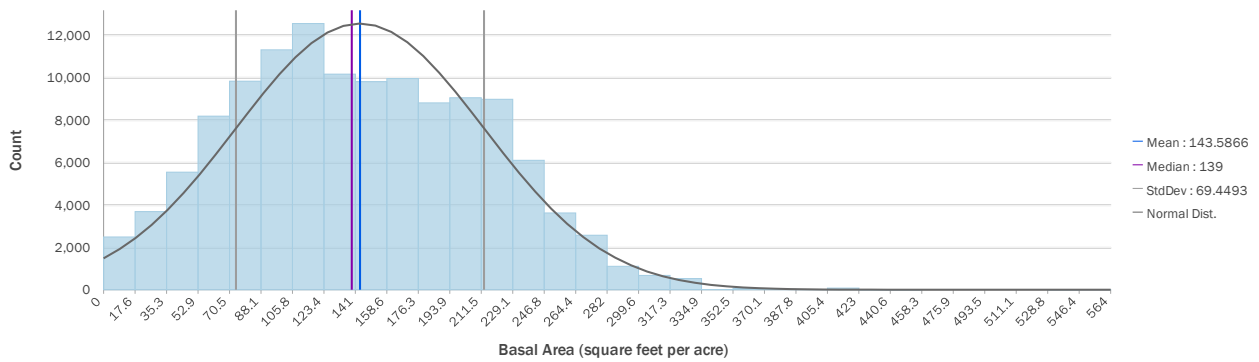
Distribution of Tree Density in 3a

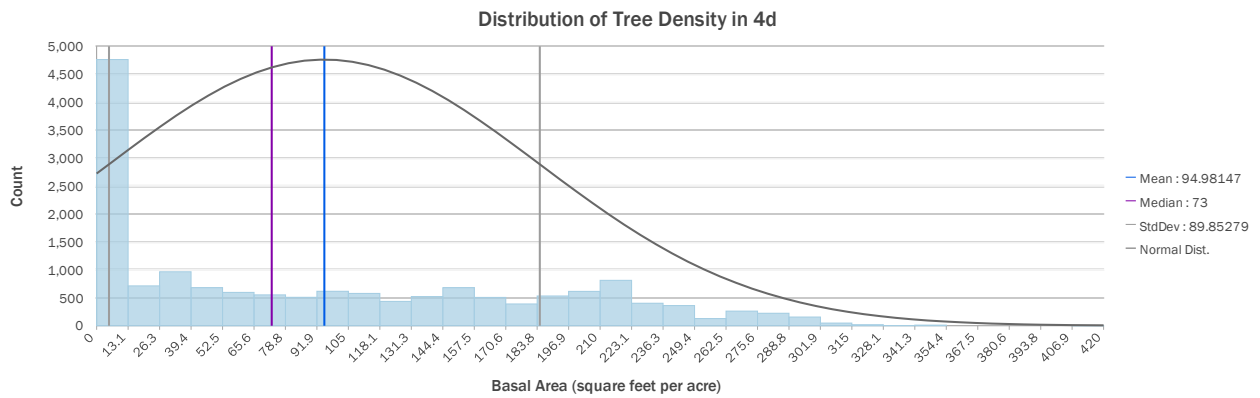
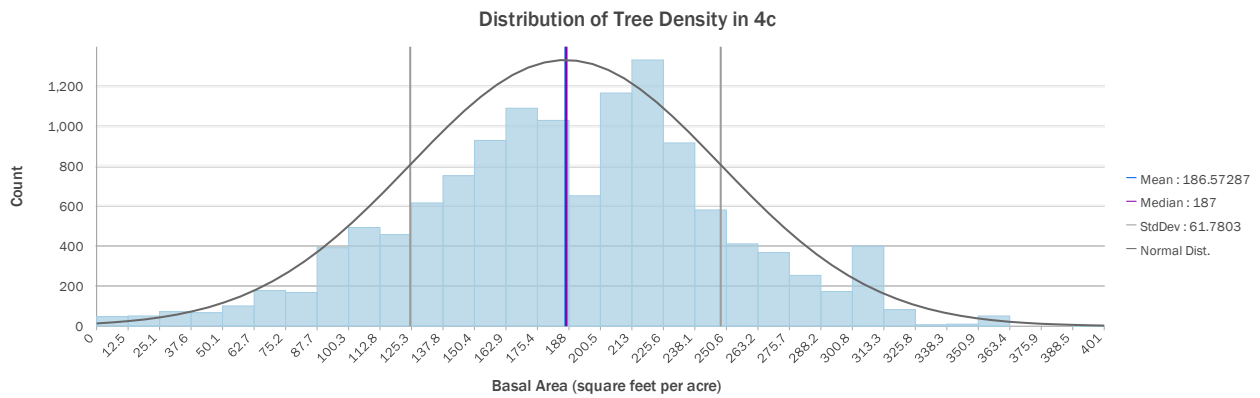
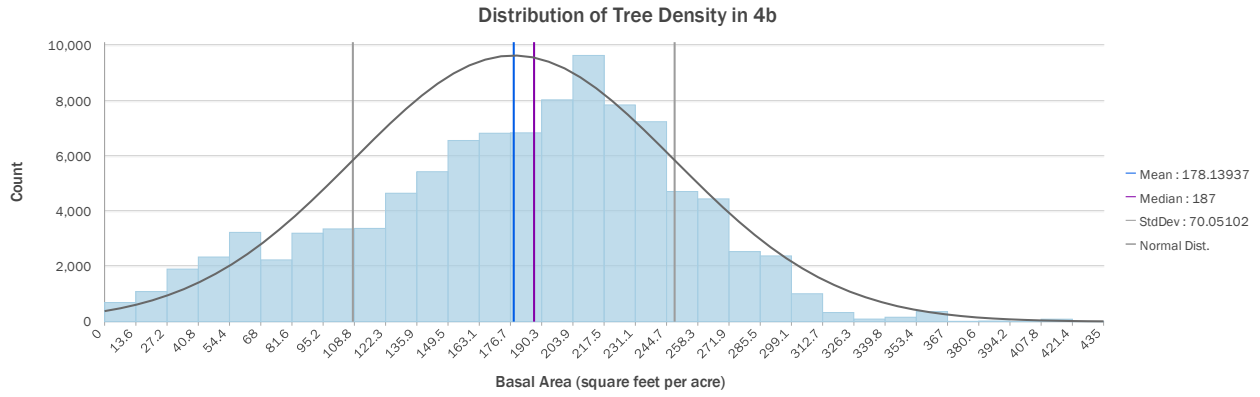


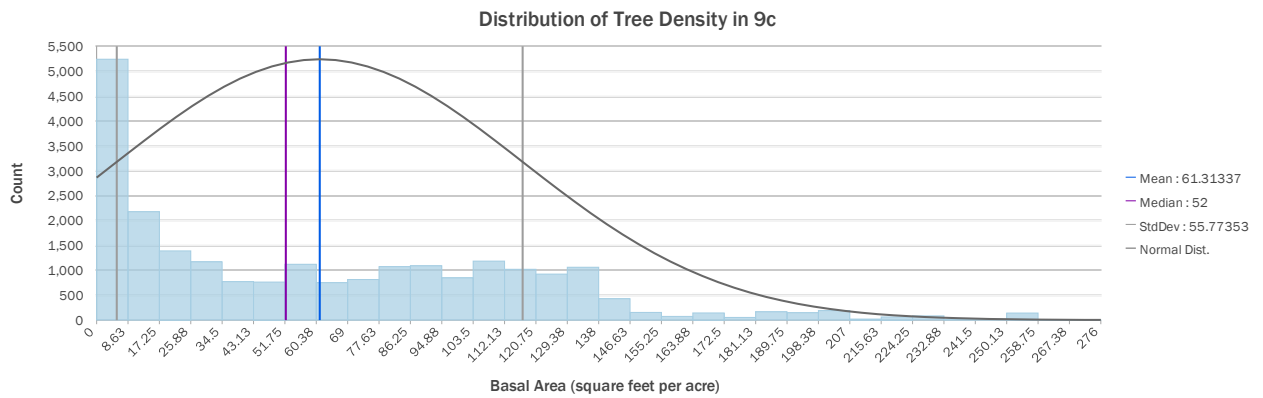
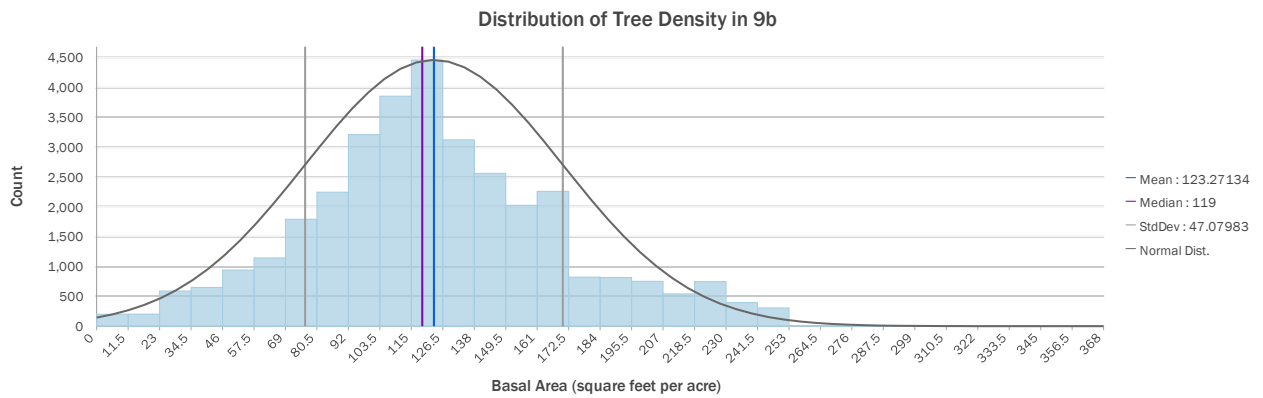
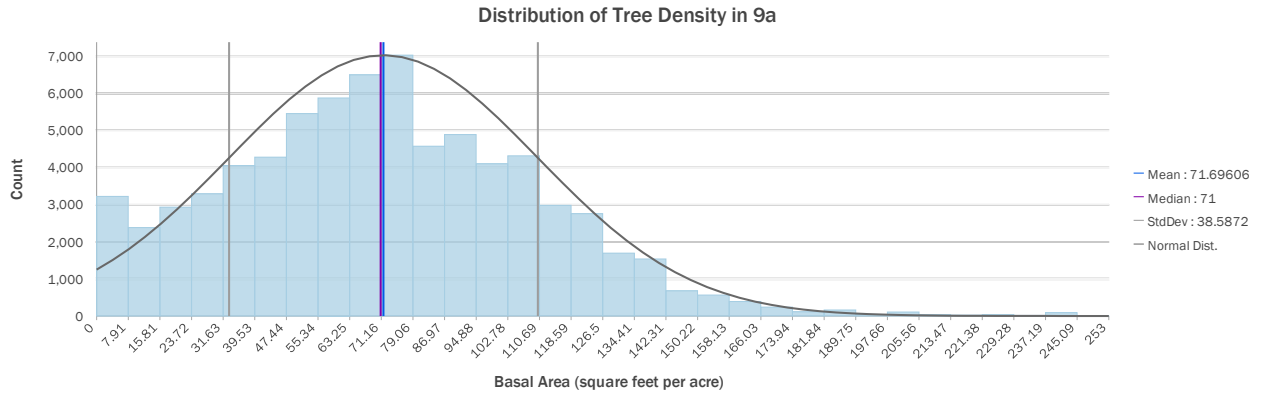
Distribution of Tree Density in 3d



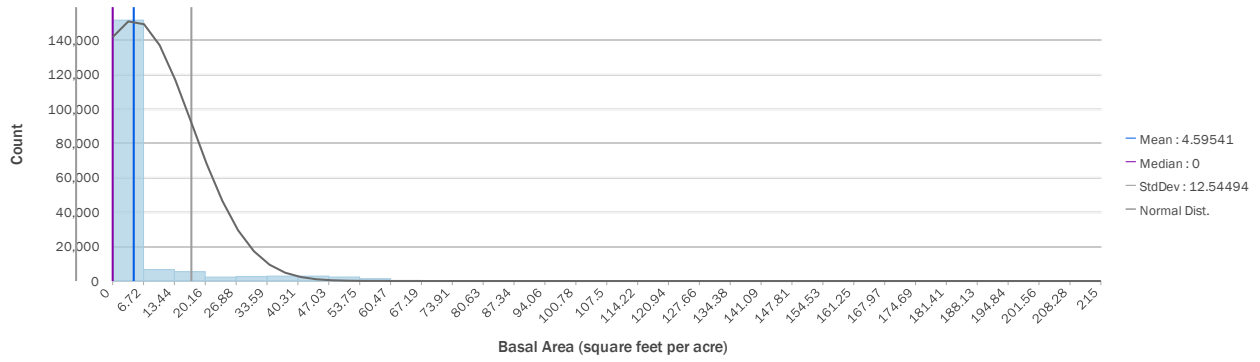
Distribution of Tree Density in 4a



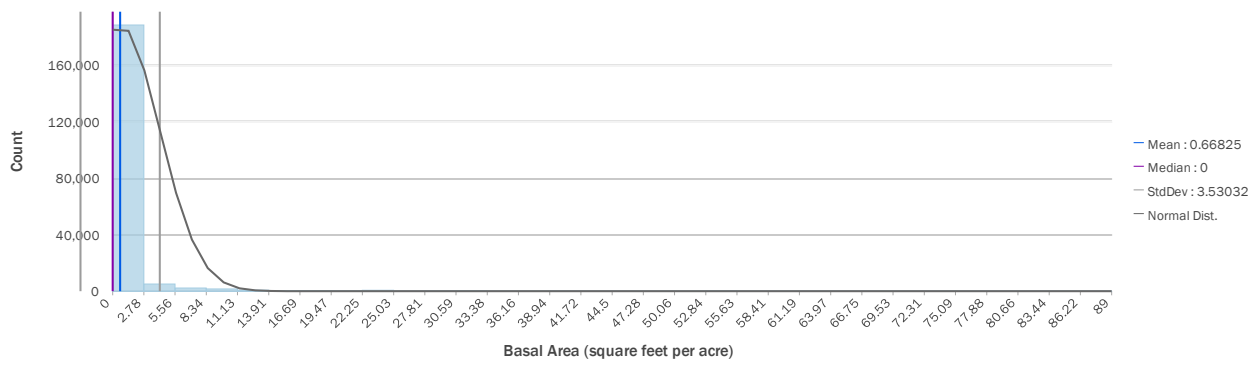




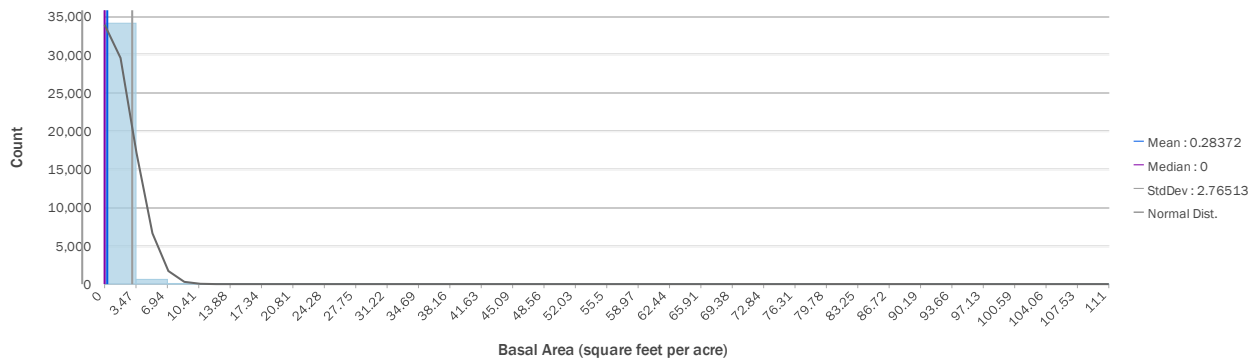
Distribution of Tree Density in 10a



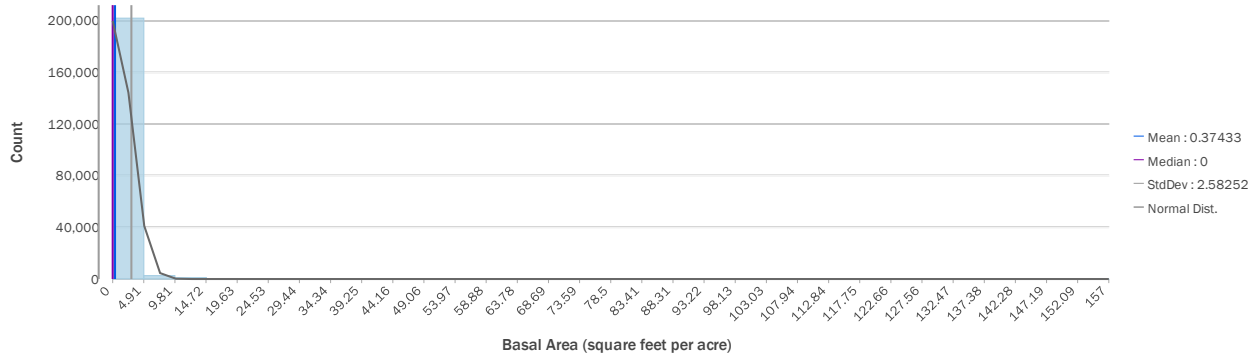
Distribution of Tree Density in 10b



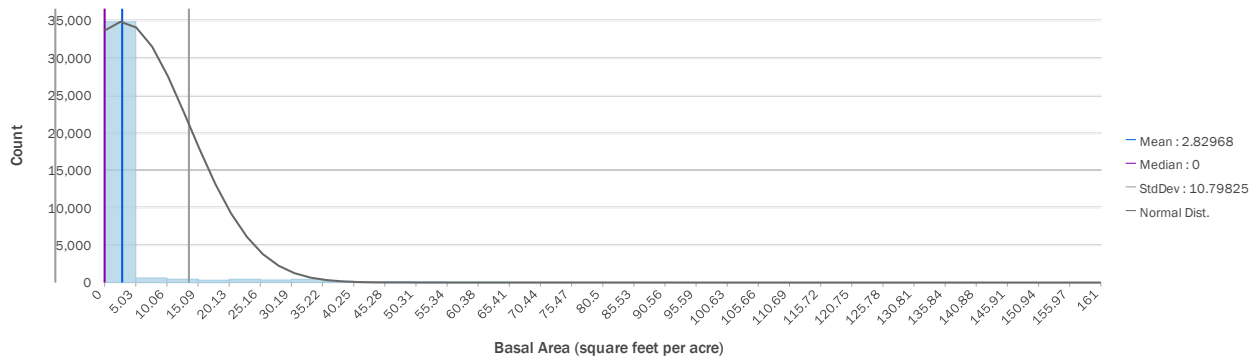
Distribution of Tree Density in 10d



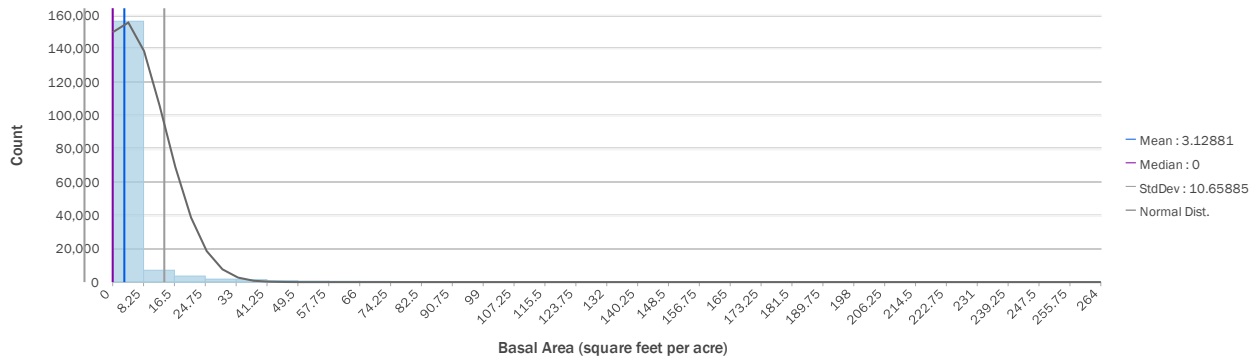
Distribution of Tree Density in 10e



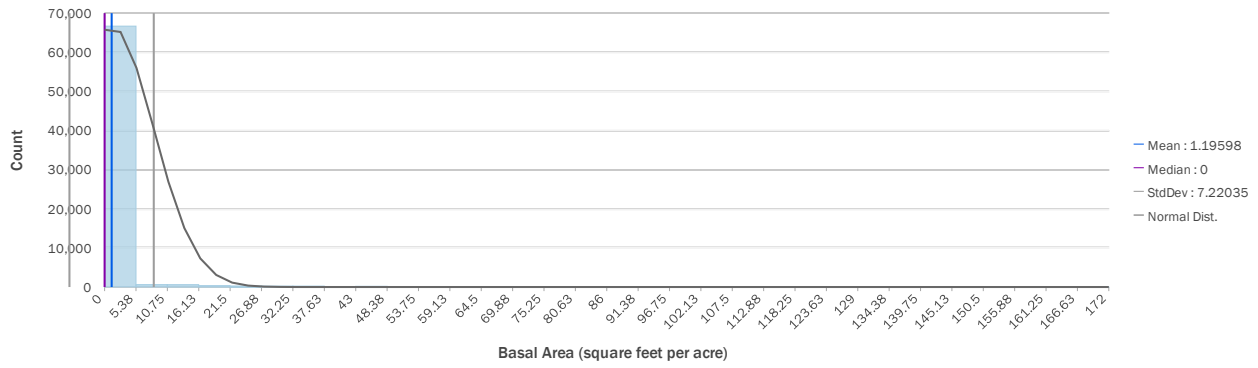
Distribution of Tree Density in 10f



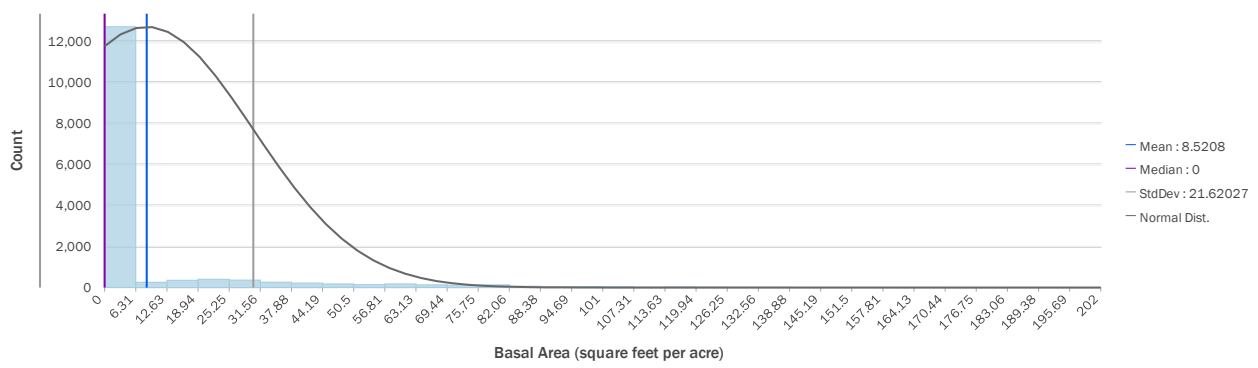
Distribution of Tree Density in 10g



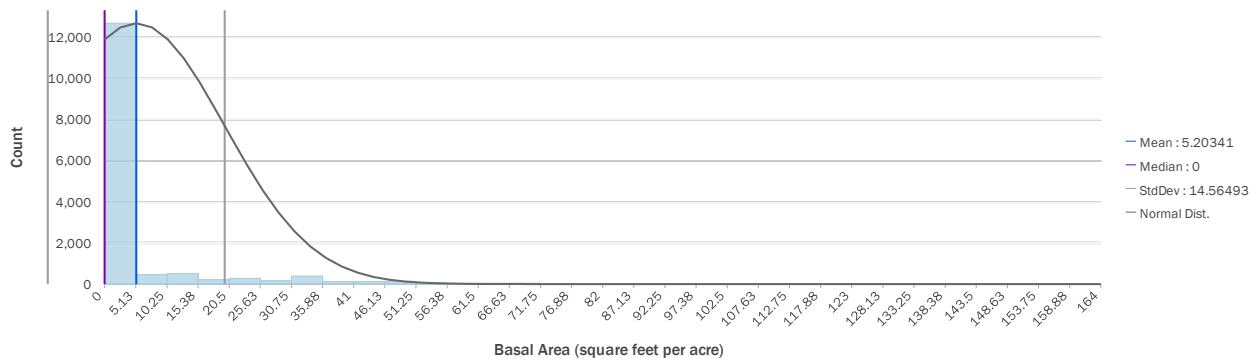
Distribution of Tree Density in 10h



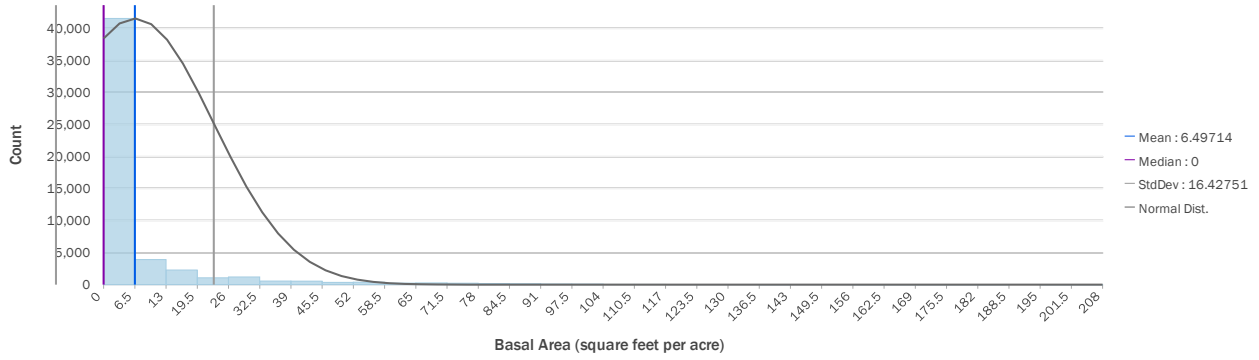
Distribution of Tree Density in 10i



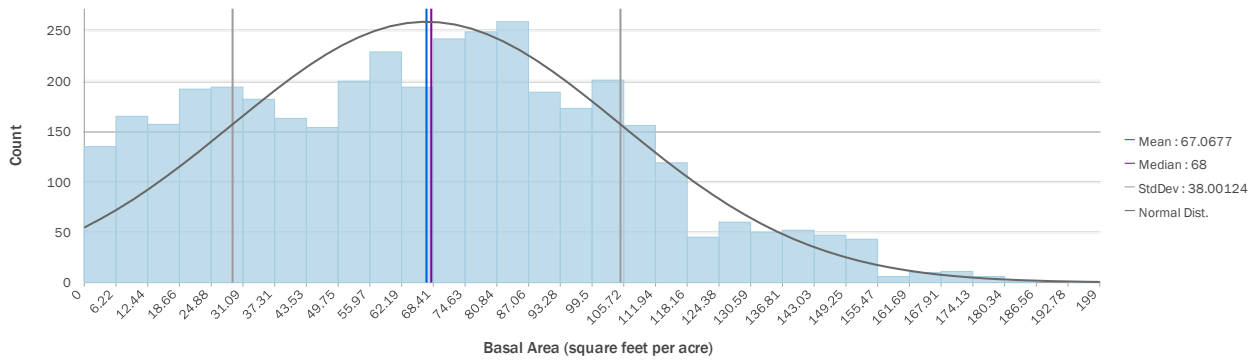
Distribution of Tree Density in 10l



Distribution of Tree Density in 10m



Distribution of Tree Density in 11c



Distribution of Tree Density in 11f

