Digital Classification and Mapping of Urban Tree Cover: City of Santa Cruz

Erin Lannon

May 11, 2018

GIS: Directed Study

Middlebury Institute *of*
International Studies at Monterey

Erin Lannon | MIIS
May 11, 2018

Abstract:

The Climate Action Plan initiative intends to chart a course for climate mitigation targets for the City of Santa Cruz. The purpose of this project is to measure tree canopy cover to determine if the City can meet its goal of increasing tree canopy cover 10% by 2020. Accurate maps of tree coverage are important tools for natural resource management, urban planning and urban sustainability goals. The Center for Integrated Spatial Research (CISR) at UC Santa Cruz created an algorithm to quantify and map the City of Santa Cruz's urban tree canopy cover. To compile data for this project we used City tree planting records, urban data and satellite imagery were compiled as our reference data set. Spectral and texture layers were extracted from National Agriculture Imagery (NAIP) data sets. We used high resolution remote sensing data, aerial satellite imagery and geographic information systems to classify landcover. In this paper, we present a comprehensive literature review on the methodologies used for estimating tree canopy coverage with satellite imagery.

This paper shows that urban forestry is a key to understand the urban infrastructure. Understanding the distribution of tree canopy coverage in Santa Cruz is important in understanding the ecosystem services including: carbon sequestration, reducing urban heat-island effect, clean air, runoff filtration, water cycle regulation, wildlife diversity, increased human health and community stewardship. Furthermore, urban tree cover can improve neighborhood aesthetics and property values.

Keywords: Santa Cruz, tree canopy cover, remote sensing, urban forestry, tree benefits, ecosystem services.

Introduction:

In 2012, the City of Santa Cruz adopted the Climate Action Plan (CAP) with an overarching goal to reduce community-wide greenhouse gas emissions 30% by 2020 and 80% by 2050. The CAP outlines twelve climate mitigation actions to reduce greenhouse gases (GHG) by 30%. Quantifying the Urban Tree Canopy Project is critical to achieving Milestone 11 is to increase tree canopy cover 10% by 2020.

The Urban Tree Canopy project will compare two data sets from 2009 and 2016 to evaluate the change in tree canopy coverage. The 2009 data will serve as a baseline canopy measurement to determine the percentage change over a five-year period. Utilizing remote sensing techniques, object-based imagery analysis and GIS we were able to generate an algorithm to classify vegetation layers to determine the urban tree canopy coverage of each representative year. This project will serve as a critical piece of information to measure, maintain and improve tree canopy cover in Santa Cruz.

Data Sets

Remote sensing data

To understand the complex vegetation coverage of Santa Cruz we needed high quality remote sensing data to accurately map open and non-open tree canopy coverage. We used NAIP imagery which was used with a pixel resolution of 1 meter or 3.28 feet for the 2009 NAIP data set and 60 cm or 1.97 feet for the 2016 NAIP data set. We also purchased a Spatial Tree Canopy Coverage data set from EarthDefine at a pixel resolution of 1 meter to compare our results. In this paper our methodology combines GIS and remote sensing data to improve the accuracy of our results.

GIS Data

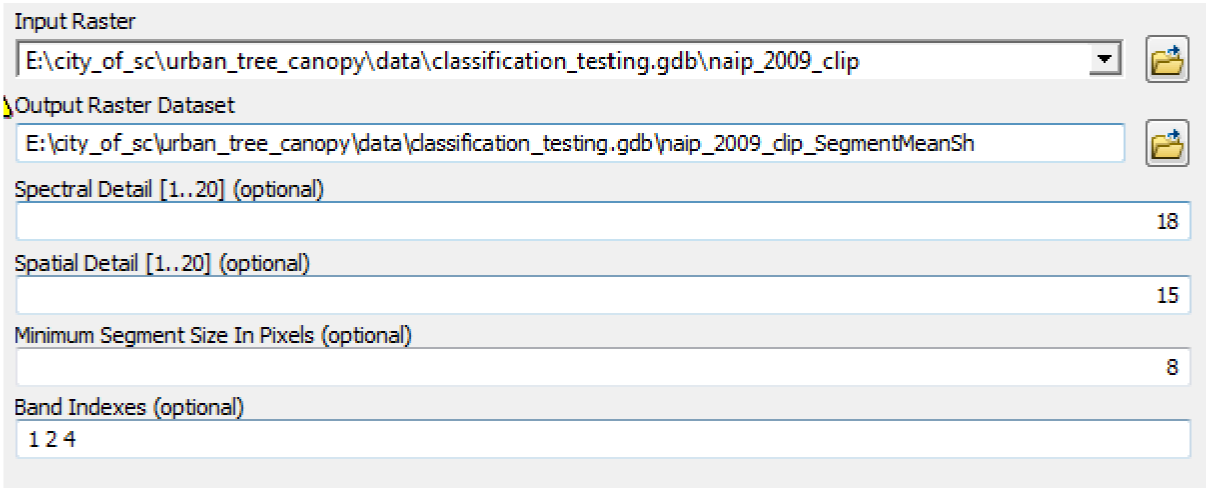
GIS layers were provided by Rich Westfall, the GIS Coordinator of City of Santa Cruz. Data layers included the boundaries of the city, parks, streets, tree plantings and land uses. All vegetation data was gathered from the National Agriculture Imagery Program (NAIP). NAIP provides aerial imagery during the agricultural growing seasons. Furthermore, we utilized data from the California Protected Areas Data Base (CPAD).

Methodology | Measuring Tree Canopy Cover

Classification Process

Following data collection, we clipped the NAIP imagery and the Santa Cruz City boundary, creating a layer of NAIP imagery for the city. Next, we used the segment mean shift tool with inputs of clipped NAIP imagery (without stretching the image) to be used to identify features in our 2009 and 2016 data sets (Figure 1). Using the NAIP Imagery data we created three separate layers to create systematic classification.

Figure 1: Clip of NAIP Four Band imagery to Santa Cruz City Boundary



Create Ancillary Raster

Following that step, we utilized the NAIP imagery layer to create a texture layer based on grouped pixels from adjacent segments to determine features or segments that had similar characteristics to create our texture layer. Then we used the NAIP data to create an output of Normalized Difference Vegetation Index (NDVI) layer. NDVI is a numerical indicator that uses the visible and near-infrared bands of the electromagnetic spectrum, which can be used is to analyze remote sensing measurements and assess whether the image is green vegetation based on reflected light. This data was used to create multiple layers based on the reflected light, see Table 1 for NAIP Parameters. Creating an additional input Raster for classifiers allows us to use a multiband raster from the following derivatives (Table 2).

Table 1. NAIP Parameters

NAIP 2009 Parameters		**NAIP 2016 Parameters**	
Spectral Detail	18	Spectral Detail	15.5
Spatial Detail	15	Spatial Detail	15
Min. Segment Size in Pixels	8	Min. Segment Size in Pixels	16
Band Indexes	1 = red, 2 = green, 4 = near-infrared(NIR)	Band Indexes	1 = red, 2 = green, 4 = near-infrared (NIR)

Table 2: NDVI Multiband Raster

NAIP bands (1,2,3,4)
NDVI data (NIR – red and NIR + red)
1st and 2nd Order Texture Metrics (ENVI) a. Variance (Band 2:naip_2009_aoiClip.tif) b. Contrast (Band 2:naip_2009_aoiClip.tif) c. Entropy (Band 2:naip_2009_aoiClip.tif) d. Correlation (Band 2:naip_2009_aoiClip.tif) e. Variance (Band 4:naip_2009_aoiClip.tif) f. Contrast (Band 4:naip_2009_aoiClip.tif) g. Entropy (Band 4:naip_2009_aoiClip.tif) h. Correlation (Band 4:naip_2009_aoiClip.tif)

The last step during this phase was utilizing the NAIP Imagery data set to create a segmented image layer. Segmentation provided an approach to extract specific features from imagery based on objects. This step allowed us to group pixels in close proximity and similar spectral characteristics in the same classes.

Following classification, we used the texture metrics and NDVI layers to generate a composite raster. The composite raster tool can be used to create a raster data set containing a subset of the original raster dataset bands. It helps to create new raster dataset with specific band combination and order to utilize the NDVI layers. The composite raster, segmented the image and training data were then used to use the Train Support Vector Machine Classifier tool.

Manually Classify Training Sites

Initial data processing involved segmenting the city of Santa Cruz into numerous quadrants and manually flagging three types of classifications (tree, shrub, urban) to determine a rough estimate of the tree canopy coverage (Table 3). This classification process had a high rate of error.

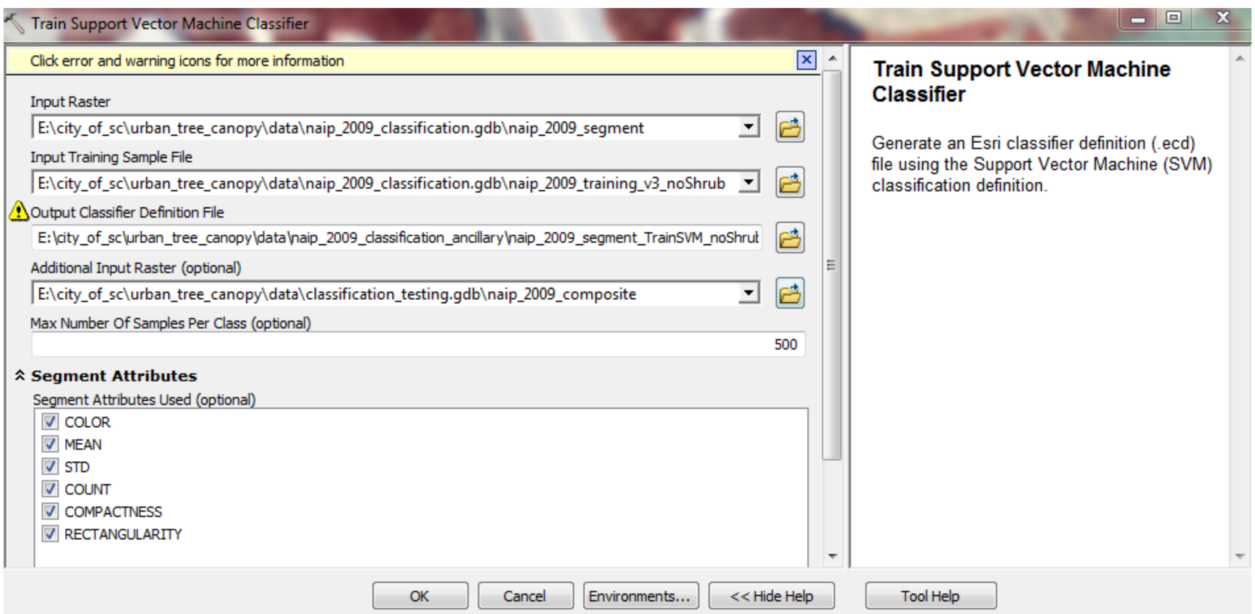
Table 3: Vegetation classes

Final vegetation use class	***Original vegetation use class***
1. Tree	1. Tree
2. Shrub	2. Shrub
3. Grass	3. Green Grass
4. Urban	4. Golden Grass
	5. Urban/Grass

Train Classifier

We utilized the image classification toolbar to create training sample set. To train the data set we used the Train Support Vector Machine (TSMV) classifier tool (see Figure 2). The TSVM tool can be used a segmented raster input or a standard image. We included the classifier in the file name to ensure that we were able to differentiate the classifications (tree, shrub, no shrub, green grass, yellow grass etc.) Always make sure to set the Max Number of Samples Per Class to zero to ensure all samples are trained.

Figure 2: Train Support Vector Machine Classifier



From here if the classification of the raster yielded unsatisfactory results so we revised the training data set to tweak our process via manual verification measures. Once we created satisfactory results we moved on to our post processing phase.

Post Processing Steps

The post processing step was an automated process. During this project Erik Lowe used an original python script he designed to smooth boundaries between classes. For example, the Python script smoothed boundaries between classes that were manually classified as shrubs that shared the majority of the boundary with trees. This additional step was needed since the automated classification misclassified trees as shrubs. Yet finding shrubs that were surrounded by trees was simpler to automate and help classify missed trees minimizing over estimation. See Table 4 below for our post processing inputs.

Table 4: Post Processing Inputs

1.	Landcover path to classify the landcover dataset
2.	Geodatabase location
3.	Landcover remapping to specify remapping for input landcover dataset
4.	Canopy hole threshold to determine the maximum size of shrub that should be reclassified as a tree (units = pixels)

Validation (QA/QC)

Classification is a statistical process that groups pixels into areas based on common characteristics. There are two types of classification human assisted (supervised) and clustering (unsupervised). Both methods were used on this project. We chose four vegetation classes based on Table 3. We ended up using multiple data sets to process the QA/QC for this project.

First, we manually identified all three types of vegetation classes based on Table 3 for the 2009 and 2016 data set. Next, we manually reviewed the training set for 2009 and 2016 to flag any misclassified tree, shrubs, grass to decrease error. Then, we manually reviewed the python algorithm in random sections to cross check the unsupervised classification. Areas that presented challenges were found near waterways, river ways, and lagoons that proved to be difficult to classify due to the verdant abundance. Finally, we compared the EarthDefine canopy estimations with our data to determine which method captured more trees. The Center for Integrated Spatial Research created computer-based tree selection through Python that quantified a more precise tree canopy estimation than the purchased EarthDefine data set.

Results (total tree canopy)

2009

EarthDefine 33.4% | 2,805 acres

CISR 36.9 % | 3,098 acres

2016

EarthDefine 32.2% | 2,703 acres | -3.64% decrease

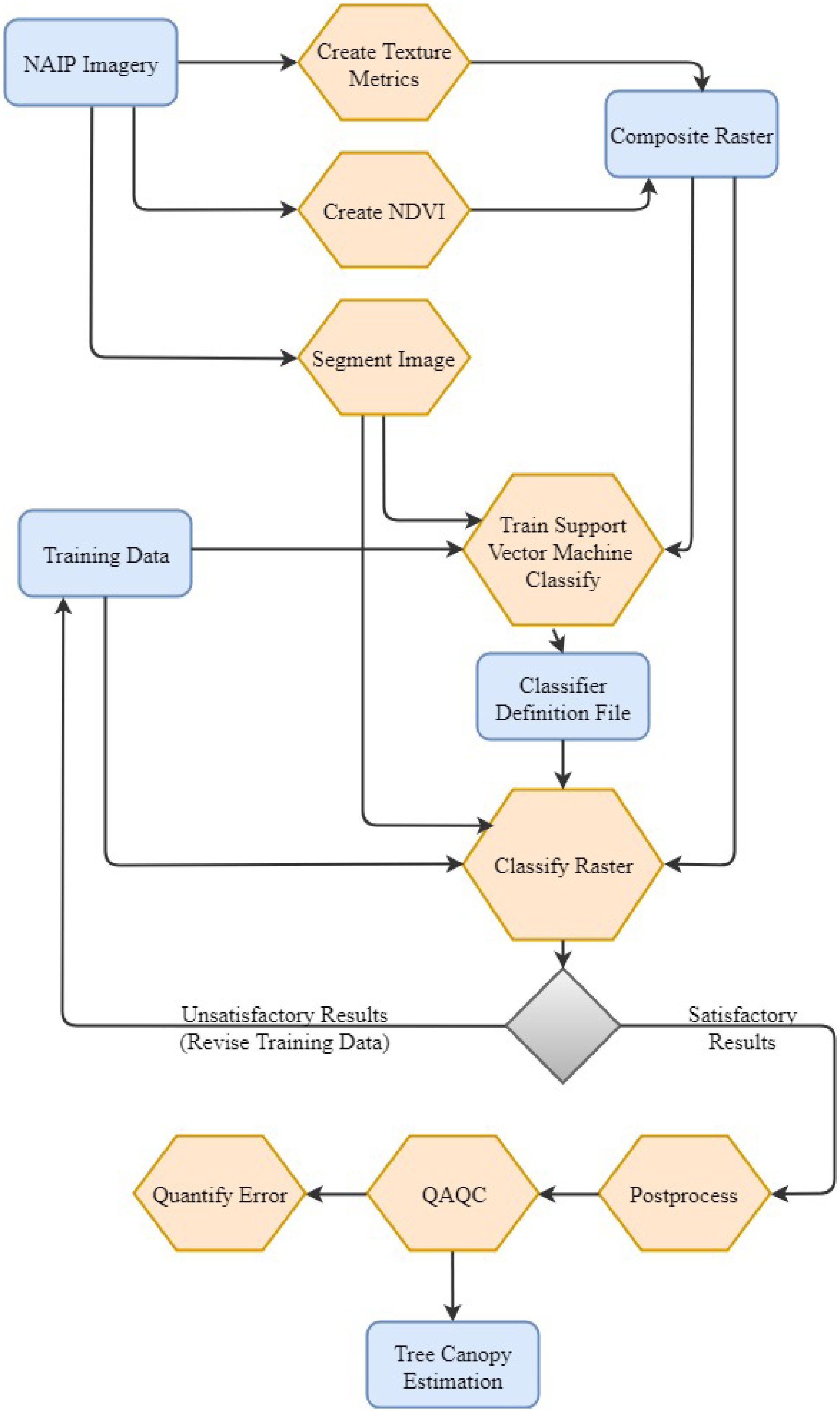
CISR 39.4 % | 3,306 acres | 6.71% increase

Conclusion

The results from this project support using NAIP imagery to monitor and asses urban tree canopy cover in the City of Santa Cruz. Although we found minor differences in tree canopy coverage estimates between EarthDefine and CISR's proprietary Python algorithm the average was within ±1.5%. Although some accuracy was lost due to precision of satellite imagery we are confident in the canopy estimation. It should be noted that some small trees that were shaded, near waterways, riverbeds and lagoons could have been missed depending on the time of year the satellite imagery was taken.

The primary purpose of the project was to determine the increase of the UTC between 2009 and 2016 to project if the City is on track to meet Milestone 11 by 2020. NAIP imagery and object-based image analysis can provide stakeholders with quantitative data that can be used to track and create realistic goals for Climate Action Planning. According to our project results the City could achieve their goal of a 10% tree canopy increase based on 500 additional trees planted in 2018 with an estimated canopy of 30' diameter.

Appendix:


NAIP Imagery
Create Texture Metrics
Composite Raster
Create NDVI
Segment Image
Training Data
Train Support Vector Machine Classify
Classifier Definition File
Classify Raster
Unsatisfactory Results (Revise Training Data)
Satisfactory Results
Quantify Error
QAQC
Postprocess
Tree Canopy Estimation