Multi-Temporal LiDAR and Hyperspectral Data Fusion for Classification of Semi-Arid Woody Cover Species

Cynthia L. Norton 1,* , Kyle Hartfield 1 , Chandra D. Holifield Collins 2 , Willem J. D. van Leeuwen 3 and Loretta J. Metz 4

1 University of Arizona School of Natural Resources and the Environment, Arizona Remote Sensing Center, The University of Arizona, 1064 E. Lowell Street, Tucson, AZ 85721, USA; kah7@arizona.edu
2 Southwest Watershed Research Center, USDA-ARS, Tucson, AZ 85721, USA; chandra.holifield@usda.gov
3 University of Arizona School of Natural Resources and the Environment and School of Geography, Development & Environment, Arizona Remote Sensing Center, The University of Arizona, 1064 E. Lowell Street, Tucson, AZ 85721, USA; leewu@arizona.edu
4 USDA-Natural Resources Conservation Service, Tucson, AZ 85721, USA; loretta.metz@usda.gov
* Correspondence: cnorton1@arizona.edu

Abstract: Mapping the spatial distribution of woody vegetation is important for monitoring, managing, and studying woody encroachment in grasslands. However, in semi-arid regions, remotely sensed discrimination of tree species is difficult primarily due to the tree similarities, small and sparse canopy cover, but may also be due to overlapping woody canopies as well as seasonal leaf retention (deciduous versus evergreen) characteristics. Similar studies in different biomes have achieved low accuracies using coarse spatial resolution image data. The objective of this study was to investigate the use of multi-temporal, airborne hyperspectral imagery and light detection and ranging (LiDAR) derived data for tree species classification in a semi-arid desert region. This study produces highly accurate classifications by combining multi-temporal fine spatial resolution hyperspectral and LiDAR data (~1 m) through a reproducible scripting and machine learning approach that can be applied to larger areas and similar datasets. Combining multi-temporal vegetation indices and canopy height models led to an overall accuracy of 95.28% and kappa of 94.17%. Five woody species were discriminated resulting in producer accuracies ranging from 86.12% to 98.38%. The influence of fusing spectral and structural information in a random forest classifier for tree identification is evident. Additionally, a multi-temporal dataset slightly increases classification accuracies over a single data collection. Our results show a promising methodology for tree species classification in a semi-arid region using multi-temporal hyperspectral and LiDAR remote sensing data.

Keywords: hyperspectral; LiDAR; species classification; semi-arid

1. Introduction

Encroachment of woody vegetation into grasslands occurs worldwide with known implications to ecosystem health [1–4]. Various studies have shown an association between woody encroachment and ecosystem degradation with consequences to carbon and hydrological cycles, decreased socioeconomic potential, vegetation productivity, habitat quality of wild herbivores and increased erosion [4–11]. Therefore, mapping the spatial distribution of woody vegetation species is an important requirement for sustainable vegetation management and scientific research [12–18]. Since land management practices for planning, treating, monitoring, and evaluation are often based on individual species, mapping vegetation distribution must also be on an individual level [19,20]. Mapping species specific distribution can aid in restoration and protection efforts and improve understanding of complex species interactions and overall distributions [21–26]. Accurately creating species specific maps can give us detailed knowledge about vegetation species composition, distribution, cover and density which provide a base for management decisions and directly
influence policy [27–32]. Ground based measurements are time consuming, labor intensive, and require economic resources, while remote sensing derived measurements are an excellent alternative for providing efficient, low cost spatially explicit data [12,13,17,33–35]. Several studies have utilized various sensors such as passive multi-spectral and hyperspectral sensors, active light detection and ranging (LiDAR) and synthetic aperture radar (SAR) systems [16,17]. All these sensors provide different information that allow for estimation of biophysical parameters and tree species classification [16,17].

Studies have widely utilized multi-spectral sensors with varying spatial resolutions that offer different levels of detail but have not produced suitable accuracies when attempting to discriminate tree species [36]. Utilizing an object-based image analysis (OBIA) on 4 m multi-spectral Ikonos imagery, Kim et al. [37,38] achieved an 83% overall accuracy but was limited to three forest types and non-forested areas. A fine spatial resolution dataset can delineate some forest tree species such as presented in the study by Franklin and Ahmed [39] where OBIA and machine learning classification was used with unmanned aerial system (UAS) images to delineate five classes of tree species, resulting in an accuracy of 78%. High spectral resolution datasets are required to capture the complex variability of species spectral reflectance characteristics [40]. Utilization of hyperspectral sensors has been increasing as they offer more detailed spectral information about land surface characteristics compared to multi-spectral sensors [36,41]. A study carried out by Dian et al. [42] used the Compact Airborne Spectrographic Imager hyperspectral data in a support vector machine (SVM) classifier resulting in an overall accuracy of 86% for four woody species. Relying on spectral signals alone for species classification applications has its limitations due to penetrable canopies, multiple scattering, spectral mixing, and bright soil reflectance influences [43].

The use of LiDAR sensors to produce three-dimensional vegetation modelling offers an accurate method for measuring tree structural attributes [44]. Fusing airborne LiDAR and hyperspectral imagery is an advanced method that provides vertical and horizontal vegetation details that improve species specific distinction and classification confidence [24,44–46]. The fusion of datasets can provide complementary and synergistic capabilities for earth monitoring programs to produce vegetation land classification cover maps over large areas with species differentiation [24,43,47]. Hyperspectral datasets provide fine spectral resolution for retrieval of plant functional traits represented by vegetation indices while LiDAR provides tree structural attributes [44]. Incorporating LiDAR datasets increases classification accuracies by providing supportive structural information when optical hyperspectral remote sensing data is limited due to spectral mixing from bright background soils with sparse canopy cover [48–50]. Classification studies in forest ecosystems such as boreal, tropical, and temperate, have shown the possibilities of using a fusion of reflectance data and LiDAR datasets to improve distinguishing multiple species for classification estimates [36,51,52]. Asner et al. [51] found error rates of 18.6% or less when combining LiDAR structural data and unique spectroscopic signal for tree species detection within a terrestrial ecosystem. A study carried out by Ballanti et al. [53] explored hyperspectral and LiDAR data to classify 8 vegetation species using pixel/object-based approaches in random forest (RF) and SVM classifiers resulting in overall accuracies above 90% for all classifications. Another study carried out by Dalponte et al. [36] compared various combinations of LiDAR, hyperspectral, and multi-spectral data for tree species identification using pixel-based RF/SVM classifiers that resulted in accuracies from 85% to 93% with decreased accuracies with increased classes.

In arid or semi-arid regions, small sparse vegetation and bright (high albedo) soil backgrounds make it hard to delineate vegetation species through remotely sensed satellite images [19,20]. LiDAR and hyperspectral sensors on board UAS platforms present new technology that offers reduced cost for high spatial and spectral resolution data that are needed for ecological research [19,20]. Some efforts carried out by Sankey et al. [19,20] highlight the potential of utilizing high spatial resolution LiDAR and hyperspectral sensors on UAS drones for dryland vegetation species classification. Their study achieved
overall classification accuracies ranging from 84% to 89% for four vegetation classes when performing maximum likelihood, RF, and SVM classifications, although the approach was limited to a small spatial extent (<100 ha) due to flight time limitations and USA Federal Aviation Administration (FAA) line of sight restrictions. More research by Dashti et al. [48] created regional scale maps (23,900 ha) using an integrated LiDAR-hyperspectral approach for classifying three xeric and mesic tree species resulting in accuracies of 60% to 89% with limitations in the mesic classes and low spatial resolutions of 17 km.

The main objective of this paper is to create a framework for producing a species-specific woody vegetation map including five of the most abundant woody species in a large semi-arid region by utilizing a fusion of simultaneously acquired airborne LiDAR and high spatial resolution hyperspectral data to improve classification accuracies. Specifically, we will: (1) create spectral vegetation indices from hyperspectral reflectance bands, create canopy height models derived from raw LiDAR point clouds, and use the data for woody species classification, then (2) assess classifier performance and lastly (3) assess the classifier model, important metrics in the model, and training data combinations. This study builds on prior research studies focused on local scales utilizing similar? resolution datasets for species-specific classifications [19,20,40,48,54].

2. Materials and Methods
2.1. Study Area

This study was conducted about 80 km south of Tucson, Arizona, United States at the Santa Rita Experimental Range (SRER) (Figure 1). Between 1902 and 1988, SRER was managed by the United States Department of Agriculture Forest Service until the University of Arizona College of Agriculture and Life Sciences took over administration to support extensive ecological research as the longest active rangeland research facility in the United States. The area encompasses 21,000 hectares of semi-arid desert with an elevation gradient from 900 m in the northwest to 1400 m in the southwest. SRER has an average yearly temperature of 20 degrees Celsius and a bimodal precipitation distribution with rainfall events during summer and winter months averaging 275 to 450 mm a year.

![Figure 1. Location of Santa Rita Experimental Range (SRER), training points for the supervised species classification are symbolized in blue (A) and study area is outlined in red (B).](image)

The study area is characterized by gently sloped alluvial fans and small areas of precipitous stony hills and isolated buttes [55]. The current vegetation regime is a woody grass savanna with a mixture of scrubs, cacti, succulents, and other herbaceous species [56,57]. Higher elevations consist of savanna woodland while lower elevation physiognomy is desert scrub. Some common trees and shrubs within the study area include blue palo verde (Cercidium floridum (Benth)), velvet mesquite (Prosopis velutina (Woot.)), creosote (Larrea tridentata (Sesse & Moc.) Cov.), and graythorn (Ziziphus obtusifolia (Hook. ex Torr. [40]).
& Gray) Gray) with an abundance of cacti such as prickly pear (*Opuntia engelmanni* Salm-Dyck) and various cholla (*Cylindropuntia*) species (*Opuntia spinisior* (Engelm.) Toumey, *Opuntia fulgida* Engelm).

2.2. Raw Data

This study utilized datasets captured by both manned and unmanned aircraft systems (MAS and UAS). Hyperspectral and LiDAR data were collected by the National Ecological Observatory Network (NEON), while high spatial resolution digital color images were collected using various UASs. NEON is funded by the National Science Foundation (NSF) to collect 30 years of long-term open access ecological data for understanding, predicting, and observing U.S. ecosystem change. NEON airborne remote sensing surveys are conducted during peak greenness (90% maximum foliage) seasons at various sites on a national level, using the Airborne Observation Platform (AOP) installed on a light MAS. The NEON AOP collected three years of hyperspectral and LiDAR data in the Western United States during 2017 (collection data: 24–30 August), 2018 (collection data: 24–28 August), and 2019 (collection data: 1–13 September). Flight surveys were conducted at a low altitude (1000 m AGL) and utilized a minimum of 10 km² area flight box design (survey boundaries) to produce consistent nominal 1 m spatial resolution datasets. Data products were published categorically by levels 1, 2, and 3 (L1, L2, L3) based on the complexity of processing steps. (https://www.neonscience.org/data-collection/airborne-remote-sensing, accessed on 3 December 2021). Level 1 point cloud and level 3 hyperspectral data products were used in this study. Each year had some data gaps and bad pixels due to cloud cover and were excluded from the analysis.

2.2.1. Hyperspectral Data

The AOP imaging spectrometer (NIS) is a passive remote sensing instrument based on the Next Generation Airborne Visible/Infrared Imaging Spectrometer (NG-AVIRIS) developed by NASA’s Jet Propulsion Laboratory (JPL). Level 3 NIS products are atmospherically corrected, orthorectified, scaled by 10,000, and stored in HDF5 format. ATCOR was used to atmospherically correct data products and orthorectified using the UTM projection/ITRF00 datum coordinate system [58]. The NIS level 3 algorithm accounts for sensor position, topography, solar positions, and atmospheric scattering resulting in a NIS level 3 surface reflectance product that has 1 m resolution and is distributed as 1 km × 1 km tiles. Four hundred and twenty-six spectral bands were then subset and converted into thirty-eight tiff reflectance bands to derive spectral indices based on the strongest signal with the least amount of noise.

2.2.2. Point Cloud Data

The AOP operates three commercial LiDAR systems, two Optech Geminis and a Riegl LMS-Q780. These systems receive and transmit high pulse repetition frequencies (PRF) to provide detailed terrain sampling [59]. The Optech Gemini sensors have been active since 2013 and can return reliable point cloud data at a maximum of 100,000 PRF (100 kHz). The Riegli LMS-Q780, active since 2018, operates at a maximum of 400,000 PRF (400 kHz). The LiDAR adjacent flight lines were designed to have a 30% overlap and at least 36° full scan angle with each laser possibly generating multiple returns. PRF’s between 100,000 and 400,000 can produce 2–8 laser pulses touching the landscape every square meter. A point density of 64 points per square meter can be obtained if the ground cover produces multiple returns and overlapping areas can produce double pulses or points. The level 1 LiDAR products are published in the American Society for Photogrammetry and Remote Sensing (ASPRS) journal as an unclassified orthorectified three-dimensional point cloud in LAS meters file format.
2.2.3. Drone Images

We used seventeen high-resolution orthomosaicked images from previous UAS flights piloted by Gillan et al. [60] in 2019 using the DJI Phantom 4 RTK. The flying height was 38 m above ground level (AGL), which resulted in plot sizes from 1.6 to 7.1 ha and a spatial resolution of 1 cm. Most of the images collected were in the southeast section of the SRER and represent the vegetation dynamics in a slightly higher elevation gradient of the range. In 2021, using the DJI Mavic2 Enterprise Dual, we collected high resolution RGB images over six additional sites in the northwest section of the study area to incorporate the different abundant vegetation species within the region for training and validation data detection and selection (Figure S1). To cover a larger area, we flew at a height of 90 m, which resulted in a spatial resolution of 2 cm and plot sizes from 26.9 to 30.9 ha. Each scene collected by the drone was orthomosaicked using Agisoft PhotoScan (version 1.4.4) to create a georeferenced orthomosaicked RGB image. A total of 23 RGB images were used for this study with plot sizes ranging from 1.6 to 30.9 ha and spatial resolutions of <2 cm. The difference in UAS drone imagery collection time was not a limitation due to slow tree growth patterns and their simple usage as a visualization aid to delineate tree species locations.

2.3. Data Processing

All data were processed through the R programming language and open-source software in RStudio [61]. We utilized parallel processing on large datasets of hyperspectral and LiDAR to produce high spatial resolution classifications of vegetation species. Three years of data were clipped, mosaicked, and stacked to create overlapping raster stacks (Figures 2 and 3). We used an optimal number of cores from 8–15 at a time to speed up processing without compromising or overwhelming performance and memory.

Figure 2. Methodology Flowchart—Hyperspectral and LiDAR processing, data fusion, and classification steps.
2.3.1. Indices

To optimize hyperspectral metrics and reduce spectral data dimension, 38 spectral bands were selected for this study. The bands chosen had the clearest variation in signals and were readily available for indices calculation appropriate for the study area. We calculated 12 vegetation indices highlighted in Table 1. Vegetation indices were developed based on biophysical and biochemical signals of various absorption features that aid in detecting and mapping vegetation species. Bands selected were derived from previous research based on differences in reflectance and absorption features that represent various plant species [62,63].

Vegetation indices (VI) were calculated to highlight vegetation canopy structure, chlorophyll content, water content, leaf area index, pigments, cellulose absorption and light use efficiency to improve discrimination of different species. The normalized difference vegetation index (NDVI) highlights green plants’ chlorophyll absorption of red light and reflectance or scattering of near-infrared (NIR) light [64]. The normalized difference water index (NDWI) was used to distinguish soil from vegetation features by utilizing NIR and short-wave infrared (SWIR) bands [64]. The soil adjusted vegetation index (SAVI) minimizes soil brightness influence on vegetation indices involving red and NIR wavelengths [65–67]. The normalized difference nitrogen index (NDNI) estimates canopy nitrogen (N) and lignin contents [68,69]. We used photochemical reflectance indices (PRI) 1 and 2 to assess short term changes in photosynthetic activity with their sensitivities to xanthophyll cycle pigment de-epoxidation state and photosynthetic efficiency [70–73].

CACTI 1 and 2 were indices developed to identify cacti vegetation using near-infrared signals and water absorption in larger landscapes using hyperspectral bands [74]. The MERIS Terrestrial Chlorophyll Index (MTCI) estimates chlorophyll content with a sensitivity to the red edge position by utilizing reflectance ratio within the wavelength range of 650–700 nm [75,76]. The cellulose absorption index (CAI) was used to account for non-photosynthetic vegetation cover and soil discrimination through depths of cellulose by selecting absorption features positioned at a water absorption feature shoulders within...
2000 nm–2210 nm [75–80]. Chlorophyll index (CI) was used to measure the amount of chlorophyll in plants as an indicator of leafy vegetation health through selection of bands within the hyperspectral dataset red-edge sections [81–83]. To create a short-wave infrared index (SWIRI), near short wave infrared bands from 1648–1642 are used with 1548, 1620 or 1690 nm to detect moister sensitivity based on chlorophyll sensitivity to the fluctuations of plant moisture [80,84–87].

Table 1. Band numbers, respective wavelengths, and vegetation indices that were used in the study with their respective formulas and the references. \( \rho \) is used to reference a specific band number that corresponds to a wavelength in nm.

<table>
<thead>
<tr>
<th>HDF5 Band</th>
<th>TIFF Band</th>
<th>Wavelength (nm)</th>
<th>Indices</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>31</td>
<td>4</td>
<td>531</td>
<td>NDVI</td>
<td>( \frac{\rho_{862}-\rho_{661}}{\rho_{862}+\rho_{661}} )</td>
</tr>
<tr>
<td>35</td>
<td>5</td>
<td>551</td>
<td>NDWI</td>
<td>( \frac{\rho_{1242}-\rho_{862}}{\rho_{531}+\rho_{862}} )</td>
</tr>
<tr>
<td>38</td>
<td>6</td>
<td>564</td>
<td>PRI</td>
<td>( \frac{\rho_{551}-\rho_{531}}{\rho_{551}+\rho_{531}} )</td>
</tr>
<tr>
<td>57</td>
<td>10</td>
<td>661</td>
<td>PRI2</td>
<td>( \frac{\rho_{1070}-\rho_{970}}{\rho_{1070}+\rho_{970}} )</td>
</tr>
<tr>
<td>59</td>
<td>11</td>
<td>671</td>
<td>SWIRI</td>
<td>( \frac{\rho_{1648}+\rho_{1548}}{\rho_{1648}-\rho_{1548}} )</td>
</tr>
<tr>
<td>67</td>
<td>13</td>
<td>711</td>
<td>SAVI</td>
<td>1.5 ( \left( \frac{\rho_{862}-\rho_{661}}{\rho_{862}+\rho_{661}} \right) )</td>
</tr>
<tr>
<td>75</td>
<td>14</td>
<td>751</td>
<td>CACTI</td>
<td>( \frac{\rho_{862}-\rho_{970}}{\rho_{862}+\rho_{970}} )</td>
</tr>
<tr>
<td>97</td>
<td>16</td>
<td>862</td>
<td>CACTI2</td>
<td>( \frac{\rho_{1070}+\rho_{970}}{\rho_{971}+\rho_{1071}} )</td>
</tr>
<tr>
<td>119</td>
<td>17</td>
<td>970</td>
<td>MTCI</td>
<td>-1</td>
</tr>
<tr>
<td>139</td>
<td>18</td>
<td>1070</td>
<td>CI</td>
<td>( \frac{\rho_{711}+\rho_{751}}{\rho_{711}+\rho_{751}} - 1 )</td>
</tr>
<tr>
<td>173</td>
<td>20</td>
<td>1242</td>
<td>CAI</td>
<td>( 0.5\left(\frac{\rho_{2000}+\rho_{2210}}{\rho_{2134}} - \rho_{2134}\right)10 )</td>
</tr>
<tr>
<td>226</td>
<td>22</td>
<td>1508</td>
<td>NDNI</td>
<td>( \frac{\log\left(\frac{1}{\rho_{1508}}\right) - \log\left(\frac{1}{\rho_{1678}}\right)}{\log\left(\frac{1}{\rho_{1508}}\right) + \log\left(\frac{1}{\rho_{1678}}\right)} )</td>
</tr>
<tr>
<td>234</td>
<td>23</td>
<td>1548</td>
<td></td>
<td></td>
</tr>
<tr>
<td>254</td>
<td>25</td>
<td>1648</td>
<td></td>
<td></td>
</tr>
<tr>
<td>260</td>
<td>27</td>
<td>1678</td>
<td></td>
<td></td>
</tr>
<tr>
<td>324</td>
<td>32</td>
<td>2000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>351</td>
<td>34</td>
<td>2134</td>
<td></td>
<td></td>
</tr>
<tr>
<td>366</td>
<td>36</td>
<td>2210</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2.3.2. Canopy Height Model

LAS point cloud data were processed to create a digital terrain model (DTM) for canopy height model (CHM) calculation. Discrete LiDAR point clouds contain information on 3D coordinates of reflected laser energy locations. A digital terrain model (DTM) was created using the lidR [88] package for spatial interpolation by k-nearest neighbor (KNN) of 12 with an inverse-distance weighting (IDW) of 2 in RStudio [61]. The normalized DTM point cloud uses an algorithm that implements a point to raster (p2r) method for creating a denser collection of points resulting in a smoother CHM raster [61,88]. The CHM was created with a p2r of 0.25 and a cell resolution of 0.5 m for a model that best captured vegetation height.

2.3.3. Training Data

To validate the UAS hyperspectral data and to assess image classification outputs, plant canopies and target species of interest were geolocated and hand-delineated using field-based GPS data, an independent source of high-resolution imagery, and ground-based photographs. Accuracy assessment samples were individual pixels randomly selected within canopies across the various flight areas. A total of 3144 trees were selected within the high spatial resolution RGB drone images as training samples for the 5 most abundant woody species within the study area, bare ground, and herbaceous classes. Different dominant species were clearly identifiable through visual interpretation of the orthomosaic
imagery, vegetation indices, and the CHMs. Additionally, individual geotagged images of trees during field visits were used to aid in species delineation. Reference points were created within visually defined canopy boundaries on the CHM and vegetation indices. Landscape patterns within all the georeferenced images and high spatial resolution drone images made it easy to identify different tree species.

2.3.4. Classification

Various classification algorithms were tested for this study including Random Forest (RF), Support vector machine (SVM), and a Classification and Regression Tree model (CART). Bare ground and herbaceous cover were classified in addition to five abundant woody species selected for the study including mesquite, palo verde, creosote, lotebush, and cacti. Although clearly not a woody species, we included cacti since it is one of the dominant species in this region. We conducted a comparative approach that compares classifiers and training sample combinations. Each classifier was compared using all available data and the classifier that resulted in the highest overall classification accuracy was used to compare training sample combinations. Visual assessments of indices and CHM imagery were used to find inconsistencies in the data that could influence the final classifications. Some datasets had missing data and/or sensor issues that resulted in a final product of a mosaicked classification (white spaces within the SRER boundaries—Figure S2). Years with obvious sensor issues were removed from the classifications which were then used to fill gaps in the highest overall accuracy classification with all data available.

CART

CART is an algorithm that was developed by Breiman et al. [89] but dates to the 1960s with Automatic Interaction Detection (AID) [90,91]. Models are represented as binary trees with a two-stage procedure and a no interaction responsive formula [92,93]. Decision trees (DT) are constructed by recursively splitting at a node based on the independent variable that resulted in the largest heterogeneous reduction in the dependent variable [90–93]. The CART model with a minimum of five observations to split a node was built through the Recursive Partitioning and Regression Trees “rpart” package [92,93].

SVM

SVM is a non-parametric classifier model based on statistical learning theory that calls a hyperplane discriminative line through kernel functions to maximize separation between classes and minimize the number of misclassification errors to a proportional quantity [94–98]. The data points used to measure the margin are closest to the hyperplane and are small in numbers [94–96]. The algorithm uses a penalizing cost-classification parameter and a function specific radial basis—gamma [99,100]. Models are based on training data with an implementation of matrix and formula interface with a factor dependent variable type (y) for classification [99,100]. A linear SVM model was created using the “e1071” package with parameters best fit for desert vegetation based on Ge et al. [101] which are 0.125 cost and 128 kappa [102–106].

RF

The RF classifier uses randomly selected training data to produce multiple decision trees for an ensemble classification [104–106]. Since the RF classifier does not overfit and is computationally efficient, the user-defined number of features and trees grown (N trees) at each node can be as large as possible [97,98,107–109]. Our study used a common error stabilizing N value of 500 trees within the R package “randomForest” [97,98,106,110,111].

2.4. Accuracy Assessment

Visual inspections were conducted of each classified image to compare results and assess possible errors. Additionally, a confusion matrix based on a 5-k-fold cross validation accuracy assessment resulting in overall accuracy (OA), user accuracy (UA), producer
accuracy (PA), and kappa was used to accept or reject each classification model and find the highest accuracy classification map. K-fold cross validation subsets the data by randomly choosing equal samples in each fold then using one-fold for validation and four for training [112,113]. The training and testing steps were repeated five times with different partitions to report averaged results [114–117]. Additionally, variable importance was measured using the mean decrease Gini (MDG) coefficient to assess the contribution of each variable to the classifier’s nodes and leaves homogeneity [118].

3. Results

Classifications were performed with all years of available data using RF, SVM, and CART classifiers. Each classifier was compared and the classifier with the highest overall accuracy was used to perform eight additional classifications with various combinations of datasets. Classification performance of each classifier model is shown in Table 2. The RF classifier was successful at achieving the highest classification OA of 95.28% and a kappa of 94.17% compared to the SVM and CART classifiers. The SVM classifier achieved the lowest OA of 77.48% and kappa of 72.14% compared to CART achieving an OA of 88.55% and kappa of 85.88%. We utilized the RF classifier to assess training data combinations and produce a final classification map due to it resulting in the highest OA and kappa.

Table 2. SVM, RF and CART Classifier Overall Accuracy, kappa and processing time in minutes.

<table>
<thead>
<tr>
<th>Indices + Lidar</th>
<th>All Years</th>
<th>SVM</th>
<th>RF</th>
<th>CART</th>
</tr>
</thead>
<tbody>
<tr>
<td>OA</td>
<td>77.48</td>
<td>95.28</td>
<td>88.55</td>
<td></td>
</tr>
<tr>
<td>Kappa</td>
<td>72.14</td>
<td>94.17</td>
<td>85.88</td>
<td></td>
</tr>
<tr>
<td>Processing Time (min)</td>
<td>55.46</td>
<td>52.30</td>
<td>46.27</td>
<td></td>
</tr>
</tbody>
</table>

3.1. Training Data Combinations

Classification iterations were conducted using the RF classifier with different combinations of mono- and multi-temporal data. The highest overall classification accuracy, 95.28%, was achieved by using data from all years, while removing the LiDAR data decreased the accuracy to 93.19%. The 2018 and 2019 classifications with LiDAR resulted in the same OA of 92.62% while 2017 dropped to 89.11% OA. Removing the LiDAR training dataset caused about a 3% or more decrease in yearly classification overall accuracies to 90.61%, 88.76%, and 87.78% for 2019, 2018, and 2017 (Table 3).

Table 3. Training Data Combination OA and kappa.

<table>
<thead>
<tr>
<th>Training Data Combinations</th>
<th>OA</th>
<th>Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Years Indices + Lidar</td>
<td>95.28</td>
<td>94.17</td>
</tr>
<tr>
<td>2019 Indices + Lidar</td>
<td>92.62</td>
<td>90.88</td>
</tr>
<tr>
<td>2018 Indices + Lidar</td>
<td>92.62</td>
<td>90.88</td>
</tr>
<tr>
<td>2017 Indices + Lidar</td>
<td>89.11</td>
<td>86.90</td>
</tr>
<tr>
<td>All Years Indices</td>
<td>93.19</td>
<td>91.57</td>
</tr>
<tr>
<td>2019 Indices</td>
<td>90.61</td>
<td>88.41</td>
</tr>
<tr>
<td>2018 Indices</td>
<td>88.76</td>
<td>86.99</td>
</tr>
<tr>
<td>2017 Indices</td>
<td>87.78</td>
<td>84.87</td>
</tr>
</tbody>
</table>

3.2. Classification Assessment

Visual assessment of the classification showed some issues of missing or misclassification due to raw data having empty data pixels, sensor issues, and weather impacts. Areas with misclassification based on sensors were removed and replaced by the next best classification.

The best classification with the highest overall accuracy of 95.28% OA and a kappa of 94.17% resulted in all classes with accuracies over 90% besides mesquite and lotebush which had accuracies in the 80s with 88.03% and 86.12% (Table 4). The variables of greatest
MDG were the CACTI, CI, CHM, NDWI, and NDVI. The 2018 CHM had the highest variable importance based on MDG with 2018 and 2019 CACTI, CACTI2, NDVI, and NDWI consecutively following in importance (Table 5).

Table 4. Shows error matrix and producer/user accuracies for classification iteration with the highest overall accuracy (All Years Indices + LiDAR) and the lowest overall accuracy (2017 Indices).

<table>
<thead>
<tr>
<th>All Years Indices + LiDAR</th>
<th>Bare Ground</th>
<th>Grass</th>
<th>Mesquite</th>
<th>Cactus</th>
<th>Lotebush</th>
<th>Paloverde</th>
<th>Creosote</th>
<th>Sum</th>
<th>UA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bare ground</td>
<td>365</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>368</td>
<td>99.18</td>
</tr>
<tr>
<td>Grass</td>
<td>5</td>
<td>287</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>297</td>
<td>96.65</td>
</tr>
<tr>
<td>Mesquite</td>
<td>0</td>
<td>0</td>
<td>397</td>
<td>1</td>
<td>32</td>
<td>5</td>
<td>2</td>
<td>437</td>
<td>90.85</td>
</tr>
<tr>
<td>Cactus</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1068</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>1078</td>
<td>99.07</td>
</tr>
<tr>
<td>Lotebush</td>
<td>0</td>
<td>2</td>
<td>43</td>
<td>1</td>
<td>242</td>
<td>2</td>
<td>7</td>
<td>297</td>
<td>81.48</td>
</tr>
<tr>
<td>Paloverde</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>10</td>
<td>2</td>
<td>285</td>
<td>1</td>
<td>306</td>
<td>93.14</td>
</tr>
<tr>
<td>Creosote</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>2</td>
<td>0</td>
<td>350</td>
<td>359</td>
<td>97.49</td>
</tr>
<tr>
<td>Sum</td>
<td>371</td>
<td>294</td>
<td>451</td>
<td>1087</td>
<td>281</td>
<td>295</td>
<td>363</td>
<td>3142</td>
<td></td>
</tr>
<tr>
<td>PA</td>
<td>98.38</td>
<td>97.62</td>
<td>88.03</td>
<td>98.25</td>
<td>86.12</td>
<td>96.61</td>
<td>96.42</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5. MDG for classification iteration with the highest (All Years Indices + LiDAR) and lowest overall accuracy (2017 Indices).

<table>
<thead>
<tr>
<th>All Years Indices + LiDAR</th>
<th>MDG</th>
<th>2017 Indices</th>
<th>MDG</th>
</tr>
</thead>
<tbody>
<tr>
<td>CACTI_2019</td>
<td>131.3818</td>
<td>CACTI_2017</td>
<td>291.5153</td>
</tr>
<tr>
<td>CACTI_2018</td>
<td>128.0493</td>
<td>NDVI_2017</td>
<td>266.4296</td>
</tr>
<tr>
<td>CHM_2018</td>
<td>118.9476</td>
<td>CACTI2_2017</td>
<td>260.4622</td>
</tr>
<tr>
<td>CACTI2_2019</td>
<td>105.366</td>
<td>CI_2017</td>
<td>240.1875</td>
</tr>
<tr>
<td>CACTI2_2018</td>
<td>97.37399</td>
<td>NDVI_2017</td>
<td>168.1216</td>
</tr>
<tr>
<td>NDVI_2018</td>
<td>83.49702</td>
<td>SWIRI_2017</td>
<td>151.0334</td>
</tr>
<tr>
<td>CI_2018</td>
<td>75.12076</td>
<td>NDNI_2017</td>
<td>140.9909</td>
</tr>
<tr>
<td>NDVI_2019</td>
<td>74.52218</td>
<td>SAVI_2017</td>
<td>124.6892</td>
</tr>
<tr>
<td>NDWI_2019</td>
<td>71.98403</td>
<td>PRI2_2017</td>
<td>115.7335</td>
</tr>
<tr>
<td>NDWI_2018</td>
<td>61.93623</td>
<td>MTCL_2017</td>
<td>114.6043</td>
</tr>
<tr>
<td>NDNI_2018</td>
<td>58.84034</td>
<td>CAI_2017</td>
<td>86.10462</td>
</tr>
<tr>
<td>SWIRI_2019</td>
<td>55.42875</td>
<td>PRI_2017</td>
<td>72.31956</td>
</tr>
<tr>
<td>CI_2019</td>
<td>52.13248</td>
<td>CACTI_2017</td>
<td>291.5153</td>
</tr>
<tr>
<td>CAI_2019</td>
<td>51.43631</td>
<td>NDVI_2017</td>
<td>266.4296</td>
</tr>
</tbody>
</table>

4. Discussion

This study was performed to create a regional species-specific woody cover map in a semi-arid environment. We conducted a series of classifications with a scripted programming approach in an open-source software. In this study, SVM and CART classifiers were less successful in classifying woody species based on a k-fold validation of OA and kappa compared to the RF classifier. Results found that RF had an increased OA by about 18% compared to the SVM and 7% compared to the CART classifiers. RF MDG variable importance based on large numbers of trees and prevention of tree masking by other correlated inputs make it a powerful effective approach for classifying woody tree species in an arid environment and assessing feature selection. Additionally, the number of user-defined
required inputs are less for the RF classifier compared to the SVM classifier. Our results line up with similar remote sensing studies that compared various classifiers that resulted in RF with the highest OA compared to CART, SVM, and other classifiers [97,98,116,117,119]. Our RF classifier resulted in 8% higher OA accuracy when compared to a similar study carried out by Naidoo et al. [40] utilizing RF models to predict savanna ecosystem tree species. Our study utilized an open source R programming approach which makes it easily transferable to other larger areas with available data. Depending on data availability, Google Earth Engine (GEE) could be used apply the same methodology to larger areas and fast processing. Accuracy assessments will need to be performed to evaluate performance in other and larger areas.

To reduce processing time and noise while increasing vegetation delineation we utilize a vegetation index-based training dataset to reduce the number of bands or layers within the classifier and highlight vegetation processes [120–123]. Adding a layer of LiDAR derived CHMs for quantification of tree height provides a unique vegetation delineation to increase classification accuracy (Table 3, Figure 4). Utilizing both VI and CHM informs the classifier regarding different vegetation assets. Our study shows that tree species classification performance increases when a multiyear combination of structure and spectral information are informing the classifier [16,17,36,124–126]. Utilizing multiyear VI on their own resulted in good classifications with over 93.19% OA and 91.57% kappa but species delineation increased when CHM models were added to the predictor variables as OA of classifications increased by 3% (Table 3).

![Figure 4. Class (A) 2018 CHM differences and (B) 2019 CACTI and NDVI model with 95% confidence interval ellipses visualizing species delineation and legend indicating the color designated to each class.](image)

A similar study carried out by Sankey et al. [19] classifying four vegetation species within an arid and semi-arid region using a UAS derived LiDAR and hyperspectral data resulted in an OA of 84–89%. Compared to our study, lower OA were obtained with a smaller number of classes. Similarly, comparing our study to other studies utilizing LiDAR and hyperspectral data such as, Matsuki et al. [127] where they classified sixteen mixed forest types and Dalponte et al. [36] where they classified seven tree species in a mountainous region, resulted in 18% and 13% less OA with 82% and 77%.

MDG values show that spectral data contribute the most to classification accuracy, but structural tree height parameters further improve classification accuracies (Table 5). For our results, we conclude that the LiDAR derived CHM model is a useful compensation tool for minor loss or lack of information through spectral signals and utilizing VI training reduces the number of bands which reduces the number of features in the model for continued reduction in computation time and increase system capacity.
Additionally, we use a multi-date training classification dataset that increased classification OA between about 3% and 7% with years 2019 and 2018 variables having the highest influence on classification accuracies (Table 3). Our results are in line with other studies using multi-temporal data for classifications utilizing UAS and non-UAS imagery. Classifications with multi-temporal datasets resulted in the highest OA with all mono-temporal classification resulting in the lowest OA.

Similar results from studies carried out by Weil et al. [128] and Grybas and Congalton [129] found the optimal multi-temporal training dataset was three dates with OA leveling off after three datasets. Our results reinforce the benefits of multi-temporal classification datasets in increasing classification OA compared to a single date. However, the achieved accuracy of the multi-temporal hyperspectral and Lidar dataset stack was only 3% lower than the best single-date classification (Table 3 and Figure 4). Wolter et al. [130] and Mickelson et al. [131] made similar observations based on multi-temporal/multi-spectral datasets resulting in improved classifications. In contrast, Key et al. [132] and Modzelewski et al. [54] found that multi-temporal datasets did not outperform mono-date classifications based on their multi-spectral temporal datasets. The difference in results could indicate that collecting a single date spectral and structural remote sensing data at an optimal collection time during peak species contrast could provide similar results as multi-temporal spectral and structural datasets.

We found that VI’s CACTI, NDVI, NDWI, and CI combined with CHMs had the highest influence on the RF classifier during prediction (Table 5). CACTI, NDVI, and NDWI utilize the NIR bands while CI have a variation of NIR and RED bands similarly to the NDVI. Additionally, NDWI includes the SWIR, suggesting that red, NIR, SWIR bands, and CHM increased the capacity to delineate tree species in a semi-arid region. Confusion matrices indicate that the lowest classification OAs were measured for mesquite and lotebush which is supported by the similarities in their spectral signal and canopy height (Table 4, Figures 3 and 5).

**Figure 5.** Rainfall events at time of data acquisition and mean reflectance curves by class for hyperspectral bands used in highest variable importance (CACTI and NDVI) for the classification with the highest OA classification.
Utilizing multi-temporal datasets can introduce challenges to the classification process. Different stages of plant development or phenology during collection time, sensor noise, and weather-related data gaps are a few reasons for classification uncertainty. As shown in Figure S2, there are differences in tree positions, the number of trees/shrubs and weather leather data gaps, which can introduce error when using the same polygons for all datasets. These differences are possibly based on differences in NEON sensor characteristics. Studies have shown these differences influence results when using high spatial resolution images [133–137]. Ferreira et al. [135] had to adjust polygons at each tree sample for resampled WorldView images with varying resolutions between 0.30 and 1.2 m. Our results coincide with prior studies based on NEON 2017 having the lowest quality dataset and lowest influence on the RF forest model. Additionally, NEON 2019 and 2018 had better quality data and in turn resulted in the highest variable importance and strongest predictors within the RF model (Figures S2 and S5, Table 5). The increase in vegetation signal could be due to increased precipitation events during data acquisition in 2019 and 2018 compared to 2017 (Figure 5). Additionally, 2018 CHM had the highest influence based on the Riegl LMS-Q780 LiDAR sensor’s capacity to produce high pulse repetition frequencies (PRF) and return reliable point cloud data at 400 kHz compared to the 2017 and 2019 sensor Optech Gemini LiDAR sensors that produced point cloud data at a maximum 100 kHz. Lastly, cloud cover and sensor noise at specific wavelength bands produced misclassification and classification gaps (Figure S2).

5. Conclusions

This study compared prediction accuracy of modeling methods, various data types, and multi-temporal data for classifying dominant woody vegetation species in a semi-arid region. Our work has demonstrated the application of fine spatial resolution airborne multi-temporal LiDAR and hyperspectral data to effectively identify and classify woody vegetation species in a semi-arid desert region at the Santa Rita Experimental Range.

The results indicate that the best modelling method is the RF with all data types. The RF model results, with a combination of spectral and structural data from mono or multi-temporal datasets, resulted in satisfying accuracies. LiDAR data fused with spectral data yielded the highest accuracy and incorporating multi-temporal datasets slightly improved classification accuracy estimates. The main source of misclassifications is in the confusion of lotebush and mesquite due to their similar signals of VI and CHM. Additional indices or data available when lotebush is blooming could help it from delineating from mesquite. To apply this method in other areas, an in situ inventory of tree and species locations are needed to aid in remote identification of tree species using UAS images. Additionally, auxiliary information or multi-temporal data of equal quality may be needed to further delineate tree species more accurately.

The results of this classification study are important in providing land managers and research scientists spatially explicit tree species maps to locate species specific woody encroachment and assess factors influencing woody encroachment [137–142]. Our study considers phenology and species traits to increase classification results by using NEON multi-temporal hyperspectral derived vegetation indices and LiDAR derived canopy height data that provide useful information for delineation of tree species in a semi-arid region. In general, the methodology detailed in this paper demonstrated the ability to perform computationally fast, accurate woody species classifications of larger areas. We also demonstrated that classification accuracies were affected by dataset proxies for vegetation dynamics, data phenological signal strength, data quality, and multi-temporal data availability.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/rs14122896/s1, Figure S1. Drone images from 2019 and 2021 with data points at the entire SRER (A), a section with three images (B) and a single image (C). Figure S2. A side-by-side comparison of 2018 (A) and 2017 (B) canopy height model. Additionally, yearly NDVI, CACTI I and canopy height model for the entire SRER (C).
Author Contributions: Conceptualization, W.J.D.v.L., K.H. and C.L.N.; methodology, W.J.D.v.L., K.H. and C.L.N.; software, C.L.N.; validation, C.L.N.; formal analysis, C.L.N.; investigation, C.L.N.; resources, K.H. and C.L.N.; data curation, K.H. and C.L.N.; writing—original draft preparation, C.L.N.; writing—review and editing, W.J.D.v.L., K.H., C.D.H.C. and C.L.N.; visualization, C.L.N.; supervision, W.J.D.v.L.; project administration, W.J.D.v.L.; funding acquisition, W.J.D.v.L. and L.J.M.; All authors have read and agreed to the published version of the manuscript.

Funding: This material is based upon work supported by the U.S. Department of Agriculture, Natural Resources Conservation Service, Conservation Effects Assessment Project-Grazing Lands component, under agreement number NRC21IRA010783.

Data Availability Statement: Publicly available datasets were analyzed in this study. This data can be found here: https://github.com/cingularities/RaBEIspecies (accessed on 1 June 2022).

Acknowledgments: The authors appreciate the research support provided by the Arizona Remote Sensing Center. We are thankful to Angela Chambers who helped with data acquisition and image interpretation. We downloaded the NEON AVIRIS hyperspectral data and NEON AOP LiDAR data from the NSF NEON Data Portal.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

12. Mairs, J.W. The use of remote sensing techniques to identify potential natural areas in oregon. Biol. Conserv. 1976, 9, 259–266. [CrossRef]


Remote Sens. 2022, 14, 2896


74. Hartfield, K.; Gillan, J.K.; Norton, C.L.; Conley, C.; van Leeuwen, W.J.D. A Novel Spectral Index to Identify Cacti in the Sonoran Desert at Multiple Scales Using Multi-Sensor Hyperspectral Data Acquisitions. Land 2022, 11, 786. [CrossRef]


77. Lowe, J.J.; Walker, M. Reconstructing Quaternary Environments; Routledge: London, UK, 2014. [CrossRef]


91. Hastie, T.J.; Tibshirani, R.J. Generalized Additive Models; Routledge: London, UK, 2017. [CrossRef]


97. Vapnik, V.N. The Nature of Statistical Learning Theory; Springer: New York, NY, USA, 1995; 188p. [CrossRef]


100. Kuncheva, L.I. Combining Pattern Classifiers: Methods and Algorithms; John Wiley & Sons: Hoboken, NJ, USA, 2014. [CrossRef]


106. Torge, L. Data Mining with R: Learning with CASE Studies; Chapman and Hall/CRC: Boca Raton, FL, USA, 2011. [CrossRef]


108. Hsiao, C. Analysis of Panel Data; Cambridge University Press: Cambridge, UK, 2022. [CrossRef]


119. Illa, I.; Loupasakis, C.; Tsangaratos, P. Land subsidence phenomena investigated by spatiotemporal analysis of groundwater resources, remote sensing techniques, and random forest method: The case of Western Thessaly, Greece. Environ. Monit. Assess. 2018, 190, 623. [CrossRef]

120. Immritz, M.; Atzberger, C.; Koukal, T. Tree species classification with random forest using very high spatial resolution 8-band WorldView-2 satellite data. Remote Sens. 2012, 4, 2661–2693. [CrossRef]

121. Thanh Noi, P.; Kappas, M. Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using Sentinel-2 imagery. Sensors 2017, 18, 18. [CrossRef]


134. Yang, H.; Du, J. Classification of desert steppe species based on unmanned aerial vehicle hyperspectral remote sensing and continuum removal vegetation indices. *Optik* 2021, 247, 167877. [CrossRef]


137. Takahashi Miyoshi, G.; Imai, N.N.; Garcia Tommaselli, A.M.; Antunes de Moraes, M.V.; Honkavaara, E. Evaluation of hyperspectral multitemporal information to improve tree species identification in the highly diverse atlantic forest. *Remote Sensing* 2020, 12, 244. [CrossRef]


