United States Department of Agriculture

GTAC-10149-RPT1 April 6, 2018

Riparian Existing Vegetation (REV) Mapping on the Apache-Sitgreaves NF

Clark, A.; Goetz, W.; Maus, P.; Megown, K.; Triepke, J.; Matthews, B.; Muldavin, E.; 2018. Riparian Existing Vegetation (REV) Mapping on the Apache-Sitgreaves National Forest. GTAC-10149-RPT1. Salt Lake City, UT: U.S. Department of Agriculture, Forest Service, Geospatial Technology and Applications Center. 9 p.

United States Department of Agriculture

Abstract

Existing vegetation products were developed to better understand the spatial distributions of vegetation types, height classes, and canopy cover on the Apache-Sitgreaves National Forest. The vegetation maps comprise of six vegetation types, four leaf retention types, five canopy cover classes for trees and shrubs, and five vegetation height classes for trees and shrubs. The existing vegetation products discussed in this document will help users to better understand the extent and distribution of vegetation, and disclose the methods and summaries of these products.

Authors

Adam Clark is a remote sensing specialist employed by RedCastle Resources at the Remote Sensing Applications Center (RSAC) in Salt Lake City, Utah.

Wendy Goetz is the vegetation mapping group leader employed by RedCastle Resources at RSAC.

Paul Maus is a contract leader and a principal of RedCastle Resources at RSAC.

Kevin Megown is the Resource Mapping, Inventory, and Monitoring program leader at RSAC.

Jack Triepke is the Regional Ecologist for the Forest Service Southwestern Region in Albuquerque, New Mexico.

Bart Matthews is the Photogrammetry Program Specialist for Forest Service Southwestern Region.

Esteban Muldavin is the Director and Ecologist for Natural Heritage New Mexico, Biology Department, University of New Mexico in Albuquerque.

Table of Contents

Introduction

Existing vegetation maps characterizing riparian lifeform, leaf retention, canopy cover, and canopy height for the Apache-Sitgreaves National Forest were developed using geospatial data including imagery, topographic and LiDAR data, photo-interpreted reference data, and modeling algorithms. These maps provide basic information on vegetation structure and composition patterns for analysis of current conditions and trends, per the 2012 Plan Rule, and to supplement R3 monitoring needs.

This work is a continuation of a previous project completed on the Gila National Forest (Clark et al. 2016) which assessed the efficiency of digital surface models (DSM) produced from stereo image pairs for mapping canopy cover and canopy height. For this project, both LiDAR data and digital stereo imagery was available. Where LiDAR data was available, covering 75% of the project area, canopy height was directly assigned to the mapping segments. For the remaining area where digital stereo imagery was available, the inverted image-derived DSM method was used to model canopy height. If neither data source was available, a traditional imagery based modeling approach was used.

Study Area

The study area is located in the Apache-Sitgreaves National Forest in east-central Arizona. The area included all riparian corridors within the Apache-Sitgreaves National Forest as defined by boundaries created by the Regional Riparian Mapping Project (RMAP), encompassing about 59,826 acres/24,210 hectares and an elevation range from 1,047 to 3,482 meters (3,435 to 11,423 feet). RMAP was produced in 2013 using topographic

Figure 1 Apache-Sitgreaves NF boundary within Arizona

information and photo-interpretation methods to delineate all riparian corridors in the Forest Service Southwestern Region (Triepke et al. 2013).

Data Collection

This project used a wide variety of geospatial data including high resolution resource imagery, Landsat 8 imagery, topographic data, and photo-interpreted data. All data were projected to a NAD 83 UTM Zone 12 projection.

Landsat Seasonal Coefficients

Landsat scenes from 2014-2016 were compiled into a time series using Google Earth Engine. Angle, a derivative from the Tasseled Cap Transformation, was calculated for each scene and a harmonic regression equation was then built for each pixel. These equations used the cosine and sine of time as independent variables and angle values as the dependent variable. These equations then represented the seasonal variability (speed, magnitude, and longevity of green-up and senescence). The equations each had three coefficients (slope of cosine, slope of sine, and y-intercept) which were represented as individual bands in an image.

NED

The National Elevation Dataset (NED) is a seamless elevation dataset for the entire United States provided by the USGS (Siddiqui and Garrett, 2008). Multiple sources such as lidar, contour maps, and data from the Shuttle Radar Topography Mission were used to create this dataset (Gesch, 2002). Although throughout the US different spatial resolutions are available, within the Apache-Sitgreaves NF the NED data resolution is 1/3-arcsecond (about 10 meters). Slope was created for use as a predictor variable in the modeling phase.

Resource Imagery

A stereo aerial image dataset of 25 cm resolution covering the Apache-Sitgreaves NF was collected by a Zeiss RMK A 21/23 sensor in summer 2008. The imagery contained three spectral bands (blue, green, red) in 8-bit GeoTIFF format scanned from film at 14 microns. To decrease processing time and memory storage required to develop the DSM, the minimum number of image pairs was chosen for the required coverage of the study area. Where possible, one image pair was chosen as coverage for an area, when four or more could have been included.

Lidar

Lidar data intersecting 75% (44,250 acres) of the RMAP project area were provided by the Southwestern Regional Office. These data were acquired in 2013 and 2014 and had an average pulse density of over 8 pulses per square meter. LiDAR data were used to generate training samples, assign height classes, and for model validation.

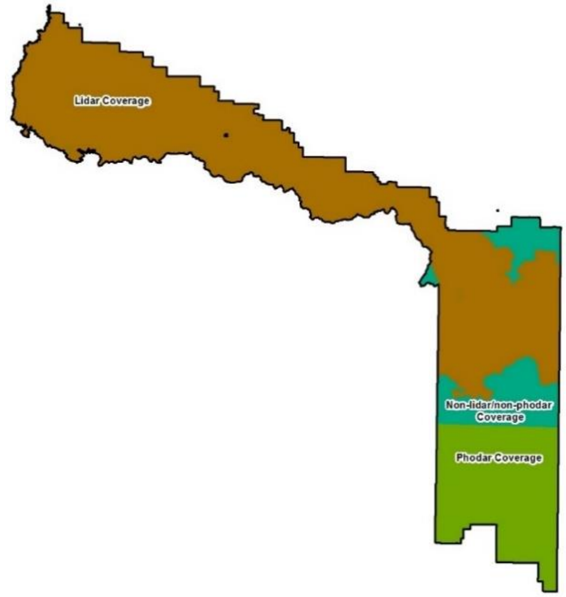

Figure 2 Map of data coverages within Apache-Sitgreaves NF

Reference Data

Reference data for this project were comprised of approximately 9,500 photo-interpreted sites. These plot data were synthesized to represent project map themes of lifeform, leaf retention, and plant height. A quality check was done to ensure all plots represented the entire mapping segment in which they were located.

Methods

The development of the final vegetation maps was accomplished in three main phases. First predictor layers to aid in modeling from a number of sources were gathered and produced. Second, modeling units were generated. Third, the lifeform, leaf retention, canopy cover, and canopy height maps were produced using Random Forest classification and photo-interpreted reference data. The final maps clipped to the RMAP boundary and filtered to a .25 hectare minimum map feature size.

Phase I: Development of Predictor Layers

Landsat mosaics and composites were processed in Google Earth Engine. NAIP, DEM, lidar were all processed and resampled using tools in ArcGIS and ERDAS Imagine.

Stereo imagery was converted into point clouds using Photoscan SfM algorithms. SfM uses overlapping imagery and sensor orientation, location, and correction data from the associated 'block files' to create data points with an x, y, and z coordinate if the same feature is 'matched' in more than one image. The outputs from SfM are in the same file format as a lidar point cloud (.las) and can contain millions of data points. The point cloud can then be fed into a software application to create a digital surface model (DSM). The imagery and block files were provided by the Southwestern Region and were used to create .las file point clouds for all riparian corridors within the extent of the 2008 imagery in the Apache-Sitgreaves NF. This extent included about 25% of the RMAP areas within the forest. Image-derived point clouds were not created in areas that already had lidar, as lidar height values were substituted for a modeled CHM. There were some areas (about 15% of the study area) with neither coverage from lidar nor the 2008 imagery. The height classes for these areas were modeled using training data, spectral, and topographic data layers. A column identifying the source of the height data is included in the dataset.

Phase 2: Image Segmentation

Image segmentation is the process of partitioning digital imagery into spatially cohesive polygonal segments (modeling units) that represent discrete areas or objects on a landscape (Ryherd and Woodcock 1996). The goal of developing segments is to simplify complex images comprised of millions of pixels into more meaningful objects. Modeling units (segments) were produced in eCognition using NAIP imagery, lidar, and the image-derived CHM data. A minimum size filter of approximately 40 square meters was used to screen out the segments that were too small to be useful. Quality was determined by assessing the modeling unit's homogeneity in lifeform, leaf retention, canopy cover, and canopy height.

Phase 3: Modeling

The modeling phase developed the statistical relationships between the reference data and the geospatial predictor data. These statistical relationships were then applied to the full extent of the census data to build a map. The first step involved producing zonal statistics (mean and standard deviation) for each modeling unit using Landsat 8 imagery and seasonal coefficients, resource imagery, Top Off CHM, and the topographic data. Random Forest algorithm was then used to assign lifeform, leaf retention, canopy cover, and canopy height classes (Breiman 2001). Where lidar was available canopy height was assigned the mean lidar value for the segment. Where lidar was not available height was modeled using image-derived height models where available as well as spectral data. Each model output was carefully evaluated for inconsistencies or misclassification using the high resolution imagery. Areas that were misclassified were reassessed, new training data added, and new models developed. This modeling procedure was repeated until the maps were considered satisfactory. The map was finalized by clipping it to the RMAP boundary and aggregating and filtering the map features to the minimum feature size.

Results/Discussion

Mapping

Map attributes characterizing lifeform, leaf retention, canopy cover, and canopy height were developed using the Random Forest algorithm. Random Forest is an advanced machine-learning algorithm based on the recursive generation of classification and regression trees. The resulting map products provide for continuous vegetation information for the RMAP area. These maps were produced using all of the geospatial data, including the image-derived canopy height information, and photo-interpreted training data. The lifeform map was produced first, followed by leaf retention, canopy cover, and canopy height classes. The final map was aggregated and filtered to the .25 hectare minimum map feature size.

Lifeform Attribute

The final map contained six lifeform classes (Figure 3). Of the total 59,826 acres, 49% percent or approximately 29,000 acres were mapped as tree or shrub. The herb class was more common in lower elevation, wide riparian corridors, adjacent near private land that was used for grazing. The sparse vegetation class was typically found along the river banks as sand bars or as dry stream beds for ephemeral streams. The shadow class was the least dominant, occurring only in areas with tall vegetation or large cliffs along the edges of riparian corridors.

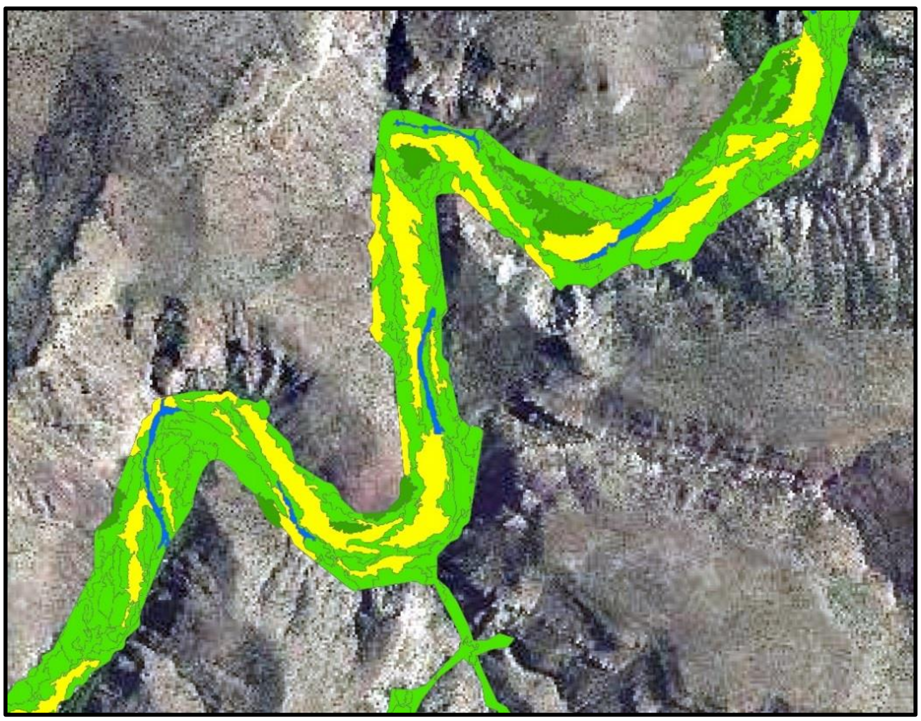

Figure 3Lifeform map

Lifeform	Acres	%
Tree	24,716	41.2%
Shrub	4,675	7.8%
Herb	16,662	27.8%
Sparse Vegetation	13,093	21.9%
Water	776	1.3%
Shadow	0	0%

Table 1 Lifeform acre summaries

Leaf Retention Attribute

Classes describing leaf retention were assigned to tree and shrub polygons identified in the lifeform classification (Figure 4). Evergreen was the most common leaf retention type and was mainly mapped in higher elevations surrounding headwater streams, while the deciduous type was more commonly found in the wide flat riparian corridors. Mixed evergreen-deciduous occurred the least. This may have been a result of limited training points, as well as segments that succesfully captured leaf retention homogeneity.

Leaf Retention	Acres	%
Non Tree-Shrub	30,532	51%
Deciduous	8,545	14.3%
Evergreen	18,544	30.9%
Mixed Evergreen-Deciduous	2,301	3.8%

Table 2 Leaf Retention acre summaries

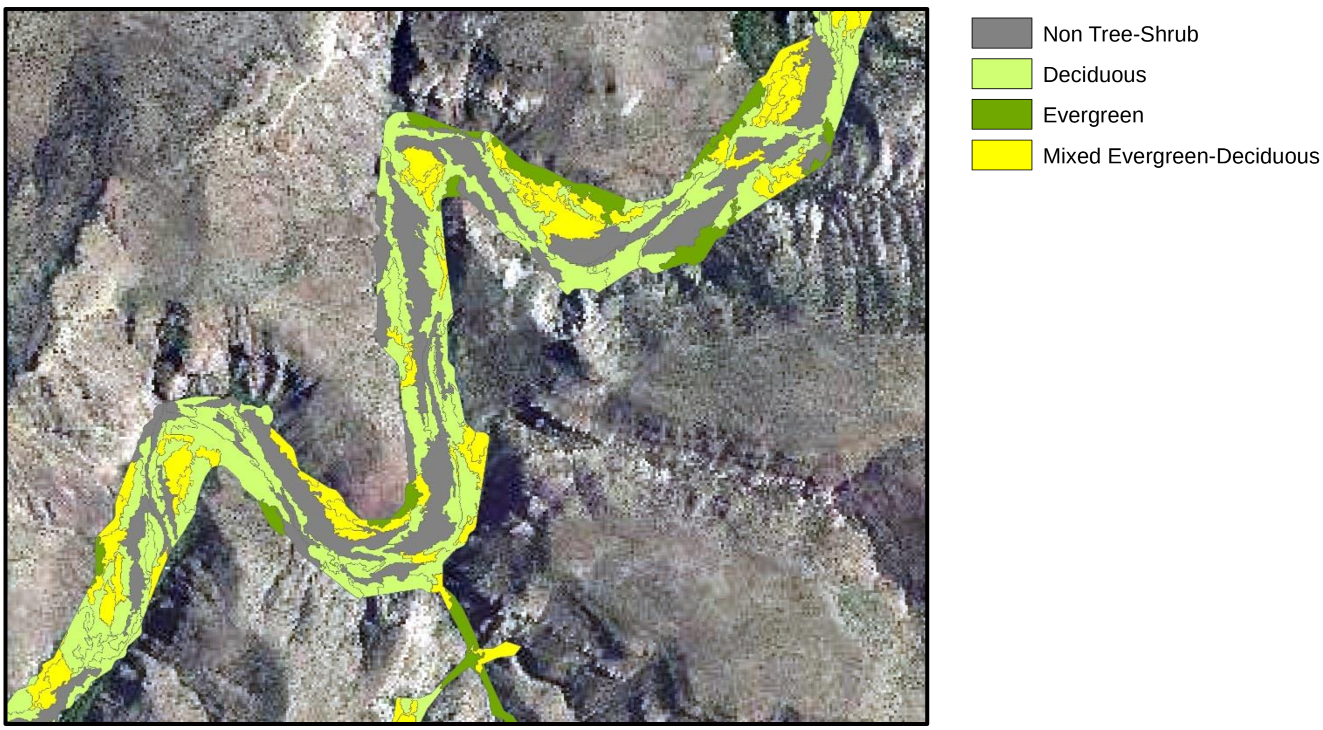

Figure 4 Leaf Retention map

Canopy Cover Attribute

Canopy cover classes were assigned to tree and shrub lifeform polygons (Figure 5). The 25-50% class was the most dominant, occupying approximately 10,800 acres or 18% of the project area.

Canopy Cover	Acres	%
0) Non Tree-Shrub	30,532	51%
1) 10 – 25%	10,044	16.8%
2) 25 – 50%	10,825	18.1%
3) 50 – 75%	5,418	9%
4) 75-100%	3,103	5.2%

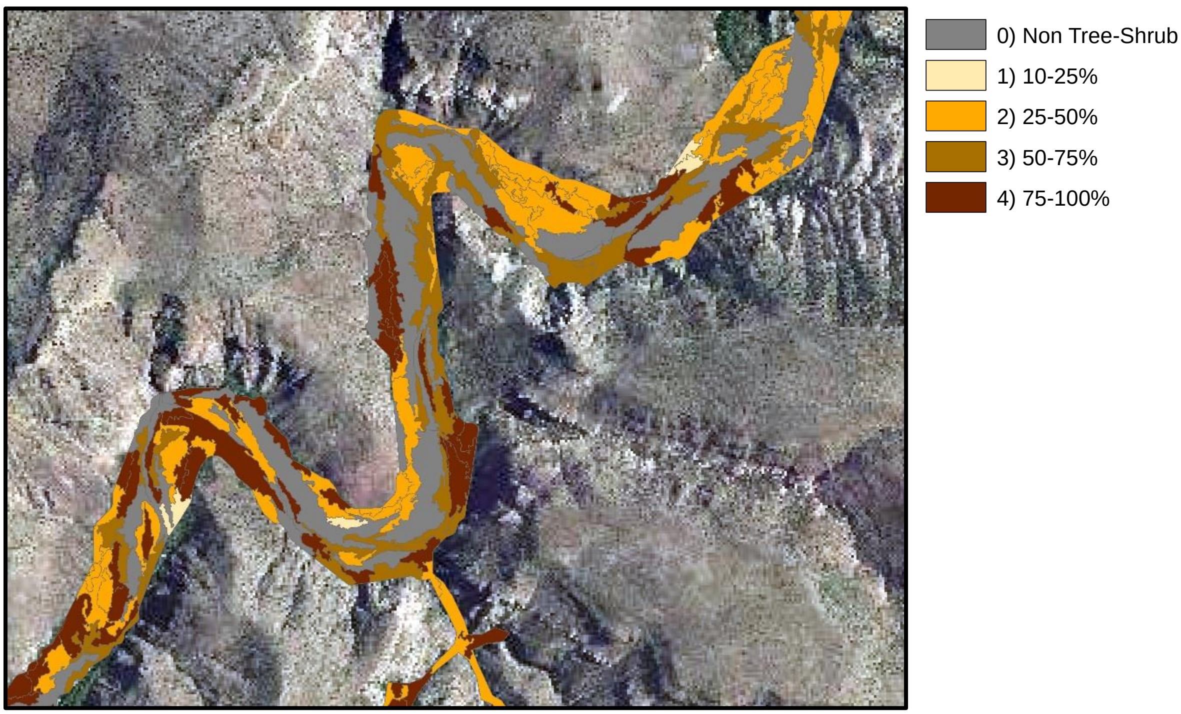

Figure 5 Canopy cover map

Canopy Height Attribute

Canopy height classes were assigned to tree and shrub lifeform polygons (Figure 6). The least dominant class was the 12+ meter class, occupying about 2,000 acres and 3% of the project area. The 5-12 meters class was the most dominant, occupying approximately 13,000 acres or 22% of the project area.

Canopy Height	Acres	%
0) Non Tree-Shrub	30,532	51%
1) 0 - 0.5 m	2,193	3.7%
2) 0.5 – 5 m	11,919	19.9%
3) 5 – 12 m	13,237	22.1%
4) 12+ m	2,041	3.4%

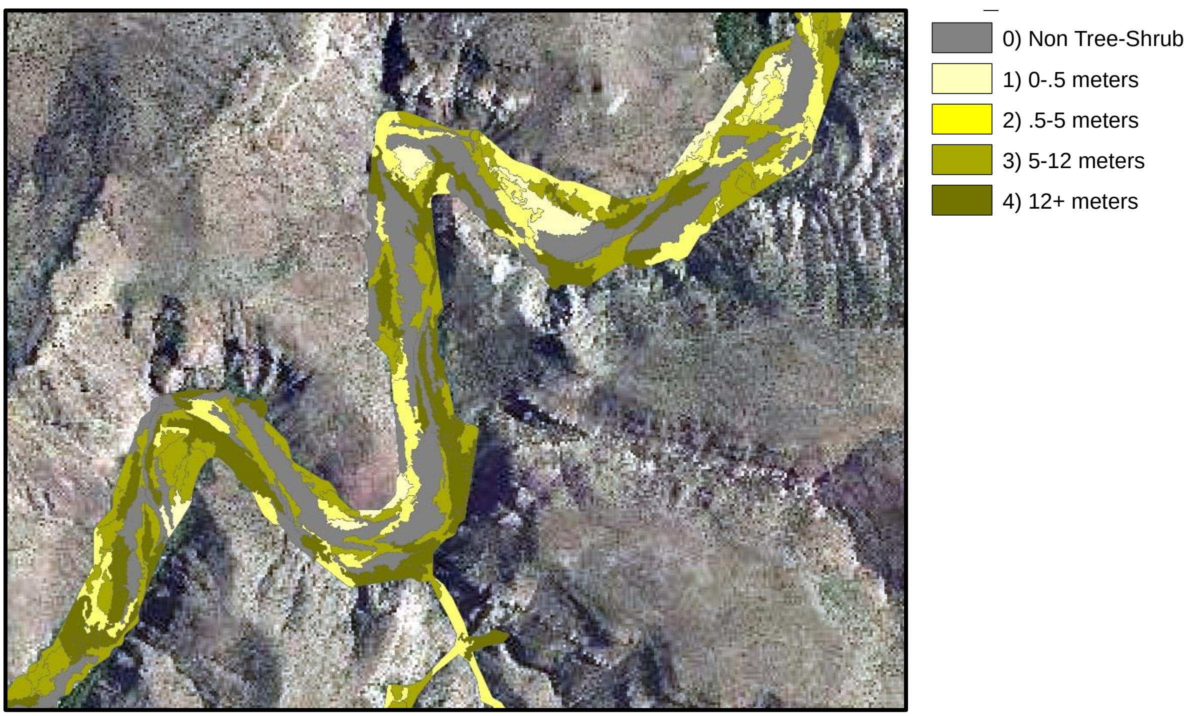

Figure 6 Canopy Height map

Conclusion

Understanding the current structure and composition of riparian areas is key to riparian resource management. Riparian corridors make up a small area but can house the largest amount of biodiversity in a forest. Using machine learning algorithms on spatial data combined with detailed information from local experts, a riparian vegetation map identifying lifeform, leaf retention, canopy cover, and height features were created.

References

Breiman, L. 2001. Random Forests. Machine Learning. 45: 5–32.

Clark, A.; Stam, C.; Goetz W.; Maus, P.; Megown, K.A.; Triepke, J.; Matthews, B.; Muldavin, E. 2016. Mapping riparian vegetation on the Gila National Forest using photogrammetric techniques. RSAC-10121-RPT1. Salt Lake City, UT: U.S. Department of Agriculture, Forest Service, Remote Sensing Application Center. 9 p.

Gesch, Dean, et al. "The national elevation dataset." *Photogrammetric engineering and remote sensing* 68.1 (2002): 5-32.

Gobakken, Terje, Ole Martin Bollandsås, and Erik Næsset. "Comparing biophysical forest characteristics estimated from photogrammetric matching of aerial images and airborne laser scanning data." *Scandinavian Journal of Forest Research* 30.1 (2015): 73-86.

Hirschmüller, Heiko. "Stereo processing by semiglobal matching and mutual information." *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 30.2 (2008): 328-341.

Pyysalo, Ulla, and H. Hyyppa. "Reconstructing tree crowns from laser scanner data for feature extraction." *International Archives Of Photogrammetry Remote Sensing And Spatial Information Sciences* 34.3/B (2002): 218-221.

Siddiqui, Yusuf, and Mick Garrett. 2008 "DATADOORS: A SYSTEM FOR CATALOGING, ACCESSING, PROCESSING, AND DELIVERING LARGE AMOUNTS OF IMAGE DATA."

Triepke, F.J., M.M. Wahlberg, D.C. Cress, and R.L. Benton. 2013. RMAP – Regional Riparian Mapping Project. USDA Forest Service project report available online < http://www.fs.usda.gov/main/r3/landmanagement/gis>. Southwestern Region, Albuquerque, NM. 53 pp.

Westoby, M. J., J. Brasington, N. F. Glasser, M. J. Hambrey, and J. M. Reynolds. 2012. 'Structure-fromMotion' photogrammetry: a low-cost, effective tool for geoscience applications. Geomorphology 179:300-314. DOI 10.1016/j.geomorph.2012.08.021. http://www.sciencedirect.com/science/article/pii/S0169555X12004217

Ziegler, Michaela, et al. "Assessment of forest attributes and single-tree segmentation by means of laser scanning." *AeroSense 2000*. International Society for Optics and Photonics, 2000.