US Forest Service Southeast Region District 8

Canopy Modeling & Parcel Databases

Contributors:
Courtney Hull, Vaz Ganushchak, Reese Beeler, Pat Hall, Johnny Burton

GEOSPATIAL ANALYSIS & MAPPING LABORATORY

Department of Geography & Environmental Sciences
COLLEGE OF LIBERAL ARTS AND SCIENCES
UNIVERSITY OF COLORADO DENVER

Contents

Project Summary

The US Forest Service has contracted the GAMLab to work on two projects. The first Project (referred to as Project 1 hereafter) is to create two canopy model products for National Forests in Alabama and Mississippi: Canopy Cover Models (CCM) and Canopy Height Models (CHM). A CCM is a raster (image) model where each cell represents the percentage of the ground covered by the tree canopy. A CHM is a raster model where each cell represents the height of the tree canopy at that point.

The second project (referred to as Project 2 hereafter) is to develop an up-to-date database of Parcel Ownership and contact information for the Gulf Atlantic Region. At the time of delivery, the scope includes parcel data for South Carolina, Florida, and Puerto Rico. The data is to be hosted in an ArcGIS Online (AGOL) Experience developed by the GAMLab.

The purpose of this report is to provide a detailed description of the products delivered to United States Forest Service (USFS) region 8, including the methodologies used and lessons learned for reference by future project personnel.

Project 1: Canopy Modeling

Purpose

The purpose of Project 1 is to produce Canopy Cover Models (CCMs) and Canopy Height Models (CHMs) that can be used by the USFS to model forest biomass density and forest health. Project 1 goals included producing canopy models for all National Forest Areas in Alabama and Mississippi, and creating and documenting the process required to do so.

Data Acquisition

LiDAR files were acquired from the United States Geological Survey (USGS) National Map. They are all products of the USGS Three-dimensional Elevation Program (3DEP). LiDAR Products were acquired for the following Forest Areas:

National Forest	Tile Count	Spatial Resolution	Temporal Resolution
Bankhead NF, AL	1023	1-meter	2017, 2020
Conecuh NF, AL	378	1-meter	2017, 2021
Oakmulgee NF, AL	399	1-meter	2016
Tabernacle NF, AL	1	1-meter	2021
Talladega NF, AL	1057	1-meter	2010, 2017, 2020
Tuskegee NF, AL	42	1-meter	2017
Bienville NF, MS	782	1-meter	2014
Delta NF, MS	583	1-meter	2018
De Soto NF, MS	129	1-meter	2016
Holly Springs NF, MS	2249	1-meter	2018
Homochitto NF, MS	890	1-meter	2018
Tombigbee NF, MS	36	1-meter	2018

All files were downloaded in the compressed .laz point-cloud format. The GAMLab's Acquisition of 3DEP Lidar files began in late August 2023, and was finished by early October 2023. Files to download were collected using USFS shapefiles to select relevant LiDAR Files on the national map. The LiDAR file names were then collected in a text file, and the torrenting application uGet (https://apps.nationalmap.gov/uget-instructions/) was used to bulk download the relevant files, which is the USGS recommended method for doing so.

Methods

To create the canopy model products, we heavily utilized the Command Line toolkit for the software package *FUSION*, developed by Rob McGaughey and the USFS. All products were created using *FUSION v 4.5.* For this process, we used Windows BatchScript files to script the workflow. Windows BatchScript is executed via the command line.

In the following sections, we will show step by step the FUSION functions that were utilized with the specific parameters broken out. Following the step-by-step processes are examples of the Windows BatchScript File used for Canopy Model Processing by the GAMLab. Appendix A shows Flowchart breakdowns for the whole process.

Canopy Modeling Process

01: Data Preparation + Canopy Modeling

FUSION only accepts certain data formats. For this process, we are starting with the compressed LAZ point-cloud format. We use the point-cloud data to derive A Digital Elevation Model (DEM) that will be used to create the CCMs and CHMs.

1. **Uncompress .LAZ file to .LAS file.**
    a. `LDA2LAS /verbose /cleanlas InputFile OutputFile`
        i. Switches Used:
            1. */verbose*
                a. Describes the process in detail as it's taking place.
            2. */cleanlas*
                a. Only outputs points that adhere to the LAS format specification (valid GPS time, return # from 1 to 5, within header extent, points not flagged as withheld. Valid for LAS format input.
        ii. Parameters used:
            1. *InputFIle*
                a. National Forest LAZ file(.laz)
            2. *OutputFile*
                a. National Forest LAS file (.las)
        iii. .LAS files can be 70-80% larger than the compressed LAZ file. To be efficient with available drive space, use a script to unpack the LAZ to a temporary LAS file. Use the LAS file to derive the DEM, CCM, and CHM for that specific tile, then delete it. Move on to the next .LAZ file and do the same thing.

2. **Filter The point-cloud file for ground elevation points.**
    a. `GroundFilter OutputFile cellSize datafile1 datfile2 …`
        i. Parameters:

1. Outputfile
    a. Name temp_groundfile.las or something similar.
2. CellSize
    a. 5
3. Datafile
    a. The temporary LAS file from step one.

ii. Similar to step one, the ground points LAS file should be deleted once used for the sake of storage efficiency.

3. Derive DEM Ground Surface from Ground point LAS.

a. *GridSurfaceCreate surfacefile cellsize xyunits zunits coordsystem zone horizdatum vertdatum datafile1 datafile2 …*

i. Parameters:

1. SurfaceFile
    a. Name for output file (i.e., “natforest_surface.dtm”)
2. Cellsize
    a. 1
3. xyunits
    a. m (meters)
4. zunits
    a. m (meters)
5. coordsystem
    a. 1 (UTM)
6. zone
    a. Depends on the area.
        i. AL and most of MS are 16, a few forests in MS fall into UTM zone 15.
7. horizdatum
    a. 1
8. vertdatum
    a. 1
9. datafile1
    a. Temp_groundfile.las produced in step 2.

4. Create CCM

a. *Cover surfacemodel outputcoverfile heightbreak cellsize xyunits zunits coordsystem Zone horizdatum vertdatum datafile1 datafile2 …*

i. Parameters:

1. Surfacemodel
    a. DEM produced in step 3
2. Outputcoverfile

a. Name of the Canopy Cover Model output(e.g., NationaForest_CCM.dtm)
3. Heightbreak
   a. 1.83
      i. (meters. Approx 6 feet)
4. Cellsize
   a. 10
5. Xyunits
   a. m
6. Zunits
   a. m
7. Coordsystem
   a. 1
8. Zone
   a. Depends on the area.
      i. AL and most of MS are 16, a few forests in MS fall into UTM zone 15.
9. Horizdatum
   a. 1
10. Vertdatum
   a. 1
11. Datafile1
   a. Temp LAS file made in step 1.

5. Create CHM

a. *CanopyModel /nofill /ground:Input_Surface_Model.dtm OutputCanopyModel cellsize xyunits zunits zone horizdatum vertdatum datafile1 dartafile2 …*
   i. Switches:
      1. /nofill
         a. Don't fill holes in the surface model.The default behavior is to fill holes in the interior of the surface model. Important for patching holes later in the merge process.
      2. /ground:input_surface_model.dtm
         a. Use the specified bare-earth surface model(s) to normalize the LIDAR data. This switch is required to create a Canopy Height Model.
   ii. Parameters:
      1. OutputCanopyModel
         a. Name of the output model. Use .dtm extension
      2. Cellsize
         a. 5
      3. Xyunits
         a. m

4. Zunits
    a. m
5. Zone
    a. Depends on the area.
        i. AL and most of MS are 16, a few forests in MS fall into UTM zone 15.
6. Horizdatum
    a. 2
7. Vertdatum
    a. 2
8. Datafile1
    a. Temp LAS file created in step 1

6. Delete the temporary LAS files created in steps one and two.

Fig. 1

```
Filename: Process.bat

@echo off

REM add Fusion commandline executables to the System Path.
REM Instruct FUSION to use the 64bit version of commandline tools (speedier
REM and more efficient for 64-bit workstations)
Path %PATH%;C:\FUSION
set FUSION64=TRUE
set LTKLOG=canopy_model.log

REM Prompt user for Forest Folder name to process
set /p NatForest=Enter Folder name for Forest:

REM Prompt user for UTM Zone input
set /p UTMZone=Enter UTM Zone:

REM Change Directory to the national forest specified
cd %NatForest%

REM Create filelist.txt with .laz files to process
DIR /b *.laz > filelist.txt
echo Compiled a list of .laz files to process in %NatForest%

REM Create output subdirectories for surface models, temp LAS files, and
output canopy Models
```

```
if not exist surface_model mkdir "surface_model"
if not exist temp_las mkdir "temp_las"
if not exist temp_ground_las mkdir "temp_ground_las"
if not exist output_canopy_cover mkdir "output_canopy_cover"
if not exist output_canopy_height mkdir "output_canopy_height"

echo "temp_ground_las" created in %NatForest%
echo "surface_model" created in %NatForest%
echo "temp_las" created in %NatForest%
echo "output_canopy_cover" created in %NatForest%.
echo "output_canopy_height" created in %NatForest%.

REM For loop that runs through all files compiled in filelist.
      for /F "eol=; tokens=1* delims=,. " %%i in (filelist.txt) do (

            REM unpack LAZ to LAS and Derive DEMs
            LDA2LAS /verbose /cleanlas %%i.laz temp_las\%%i.las

            GroundFilter temp_ground_las\temp_GroundPoints.las 5
            temp_las\%%i.las

            GridSurfaceCreate surface_model\%%i_surface.dtm 1 m m 1
            %UTMZone% 2 2 temp_ground_las\temp_GroundPoints.las

            REM Create Canopy Cover Model & Canopy Height Model

            Cover surface_model\%%i_surface.dtm
            output_canopy_cover\%%i_canopy_cover.dtm 1.83 10 m m 1
            %UTMZone% 2 2 temp_las\%%i.las

            CanopyModel /nofill /ground:surface_model\%%i_surface.dtm
            output_canopy_height\%%i_chm.dtm 5 m m 1 %UTMZone% 2 2
            temp_las\%%i.las

            REM Delete temporary LAS files
            DEL temp_ground_las\temp_GroundPoints.las
            DEL temp_las\%%i.las
      )
```

```
echo %NatForest% CCMs and CHMs processed.
set LTKLOG=
```

Fig 1 shows the Windows BatchScript used to process canopy models with the USFS FUSION-LTK *command line toolkit..*

02: Canopy Model File Format Conversion

Once all LiDAR tiles are processed into CHMs and CCMs, they need to go through post-processing to convert them to the raster ASCII format (.asc), merged and cleaned up so that they can be viewed as a contiguous area in a GIS. The ASCII format is a raster format that can be read into many GIS such as ESRI ArcGIS Pro or QGIS. All files are exported to ASCII with A 32-bit signed pixel resolution.

1. **Convert CCM & CHM files from DTM PLANS format (.dtm) to ASCII raster format (.asc).**
   a. *DTM2ASCII /raster inputfile [outputfile]*
      i. Switches:
         1. /raster
            a. Interpret the DTM points as the attribute for a cell and adjust the origin of the ASCII grid file so that the lower left data point is the center of the lower left grid cell.
      ii. Parameters:
         1. Inputfile
            a. Canopy model in DTM format
         2. [outputfile]
            a. Name of output ASCII file. (optional. Will take the name of the input file if no name is given.)
      iii. Note on Batch Conversion:
         1. Unlike some other tools in *FUSION*, *DTM2ASCII* does not have a a built-in batching feature. To run the *DTM2ASCII* function on a large selection of files, you need to write a BatchScript file with a for loop to run through your selection. See Fig.2 for an example.
      iv. Note on Conversion Speed:
         1. Because of how *FUSION DTM2ASCII* handles file conversions (loads whole file into RAM, then writes the converted file to a specified location) its faster and more reliable to run *DTM2ASCII* before merging into a larger file.
         2. Additionally, You can get a conversion speed boost if you write the new ASCII file to a different drive than the starting DTM is loaded on.

Fig 2.

```
Filename: Bulk_dtm2ascii
@echo off

REM add Fusion commandline executables to the System Path.
REM Instruct FUSION to use the 64bit version of commandline tools (speedier
REM and more efficient for 64-bit workstations)
Path %PATH%;C:\FUSION
set FUSION64=TRUE
set LTKLOG=dtm2ascii.log

Set /p selection=Converting CHMs (1) or CCMs (2)?

If %selection% == “1” (

      REM set National Forest name. Needs to be same as Forest folder
      containing REM ASCII files being converted.
      Set /p NatForest=Enter Folder name for Forest:

      REM set state folder to navigate to.
      Set /p State=Enter Folder name for state:

      REM set the output directory for ascii files. Located on different
      drive to REM optimize processing speed.
      set output=D:\CanopyHeightModels

      REM navigates to the output directory on another drive.
      cd /d %output%

      REM Creates a folder in the output directory to write output ASCII
      Canopy REM Models to.
      if not exist %NatForest% mkdir %NatForest%

      REM Change directory back to Directory storing Canopy Models in DTM
      format
      cd /d
      F:\FUSION\CanopySurface\%state%\%NatForest%\output_canopy_height

      REM Create filelist.txt with .laz files to process
      DIR /b *.dtm > dtm2ascii_filelist.txt
      echo Compiled a list of .dtm files to process in %NatForest%
```

```
        REM convert all DTMs in filelist to ASCII files
              for /F "eol=; tokens=1* delims=,. " %%i in
              (dtm2ascii_filelist.txt) do (
                          DTM2ASCII /raster %%i.DTM
                          %output%\%NatForest%\%%i.asc
                          )

        echo %NatForest% processed.
        set LTKLOG=
        )

If %selection% == “2” (
        REM set National Forest name. Needs to be same as Forest folder
        containing REM ASCII files being converted.
        Set /p NatForest=Enter Folder name for Forest:

        REM set state folder to navigate to.
        Set /p State=Enter Folder name for state:

        REM set the output directory for ascii files. Located on different
        drive to REM optimize processing speed.
        set output=D:\CanopyCoverModels

        REM navigates to the output directory on another drive.
        cd /d %output%

        REM Creates a folder in the output directory to write output ASCII
        Canopy REM Models to.
        if not exist %NatForest% mkdir %NatForest%

        REM Change directory back to Directory storing Canopy Models in DTM
        format
        cd /d F:\FUSION\CanopySurface\%state%\%NatForest%\output_canopy_cover

        REM Create filelist.txt with .laz files to process
        DIR /b *.dtm > dtm2ascii_filelist.txt
        echo Compiled a list of .dtm files to process in %NatForest%

        REM convert all DTMs in filelist to ASCII files
              for /F "eol=; tokens=1* delims=,. " %%i in
              (dtm2ascii_filelist.txt) do (
                          DTM2ASCII /raster %%i.DTM
```

```
                        %output%\%NatForest%\%%i.asc
                        )
echo %NatForest% processed.
set LTKLOG=
)
```

Fig 2. Shows the bulk_dtm2ascii.bat *script used by the GAMLab to convert Canopy Models from the PLANS DTM format (.dtm) to the ASCII raster format (.asc).*

03: Quality Assurance

Once the Canopy Models have been converted to the ASCII raster format, they need to be checked for quality before moving on to the merging and cleaning steps. Things to look out for when checking the files quality:

- Correct Coordinate Systems
- Correct coordinates
- Model corruption
- Mismatched temporal resolutions (LiDAR tiles) - use older DTMs to patch holes in new rasters

Below are steps the GAMLab used to perform Quality Assurance Checks,

1. **Create a New Mosaic Raster dataset in ArcGIS Pro**
    a. Use the *Create Mosaic Dataset* Geoprocessing tool .

2. **Load the directory containing ASCII canopy models into the Mosaic Raster Dataset.**
    a. Use the *Add Rasters to Mosaic Dataset* geoprocessing tool.

3. **Check how the rasters display in ArcGIS pro.**
    a. Are all the files located at the correct coordinates?
    b. Do the files look complete?

4. **Take corrective action.**
    a. Delete Erroneous files
    b. Fix georeferencing in places where it should be fixed.

5. **Identify date ranges for data that was collected and put them into separate .txt files as lists.**
    a. Some areas that LiDAR is downloaded for from the USGS National Map may include LiDAR data collected in different years.

i. USGS 3DEP puts the year when the LiDAR data was gathered in the file title.

b. Example:

   If you have LiDAR data for Talladega National Forest for 2015 and 2020, create two files; filelist_2015.txt and filelist_2015.txt. Have the file names listed in each.

c. When you merge rasters in the next step, this will allow the newer files to patch the older files. Fig.3 shows a snippet of batch script that can make this a bit easier.

Fig.3

```
Filename: create_model_list.bat

@echo off
setlocal enabledelayedexpansion

rem Set the directory path
set /p NationalForest=Enter name of folder holding Canopy Model Files:

rem Set the output file name
set /p year=Enter the year of files you want to collate:
set output=filelist_%year%.txt

rem Navigate to the directory holding National Forest canopy rasters
cd %NationalForest%

rem Initialize the output file
echo. > %output%

rem Loop through files in the directory
for %%F in (%NationalForest%\*%year%*.asc) do (

rem Append filenames to the output file
    echo %%~nxF >> %output%
)

echo File list generated: %output%

endlocal
```

Fig. 3 shows an example script for collating a list of files to merge in stage 04.

04: Merge Canopy models

Now that LiDAR tiles have been processed into canopy models and we have cleaned up the selection, the Canopy model files need to be merged into a single file for each national Forest Area. We achieve this using FUSION's MergeRaster tool to merge rasters in a specific order, so that the newest data is shown in the final canopy model product.

Similar to the DTM to ASCII conversion process described in stage 02, the canopy merging process runs faster if you write the merged product to a separate drive from where the original ASCII files are stored.

If, during the Quality Assurance stage, all files for a specific forest area were identified as being created in the same time period:

1. **Merge all files for a forest area using *MergeRaster.***
    a. `MergeRaster /verbose OutputFile InputFile`
        i. Switches:
            1. /verbose
                a. Echos MergeRaster process to the command line. Makes it easier to follow along with the Merge process to monitor progress.
        ii. Parameters:
            1. OutputFile
                a. Name of output merged canopy model in ASCII (.asc) format.
            2. InputFile
                a. A text file (.txt) listing all the ASCII files in the directory that you want to merge.
                b. You can also use the "*.asc" wild card variable if you want to merge all files in the directory

If, during the Quality Assurance stage, files for a specific forest area were created in separate time periods:

1. **Create multiple text files (.txt) in the directory that contains the canopy models to be merged. There should be one text file for each year represented in the canopy models.**
    a. Look to Fig. 3 for an example script to aid this process.

2. **Use the FUSION *MergeRaster* tool to create a patched raster, where you start by merging the oldest data, and then overwrite areas where newer data is available.**
    a. `MergeRaster /verbose /overlap:new OutputFile InputFile1 InputFile2 …`

i. Switches:
   1. /verbose
      a. Echos MergeRaster process to the command line. Makes it easier to follow along with the Merge process to monitor progress.
   2. /overlap: new
      a. The /overlap switch specifies how *MergeRaster* should treat areas where two or more input ASCII raster files contain values for the same cell.
      b. The *new* operator specifies that data from the input file currently being processed should overwrite any data that is currently there.

ii. Parameters:
   1. OutputFile
      a. Name of output merged canopy model in ASCII (.asc) format.
   2. InputFile
      a. To ensure that you create a "patched" Canopy model, list the input ASCII file list textfiles in order from oldest to newest.This will ensure the most current data is reflected in the final ASCII raster file.
         i. Example: filelist_2016.txt filelist_2020.txt

Fig. 4 shows an example script to create a patched canopy model ASCII raster.

Fig. 4

```
Filename: merge_ascii_patch.bat
@echo off

REM add Fusion commandline executables to the System Path.
REM Instruct FUSION to use the 64bit version of commandline tools (speedier
REM and more efficient for 64-bit workstations)
Path %PATH%;C:\FUSION
set FUSION64=TRUE
set LTKLOG=merge_ascii.log


REM set output folder for NatForest. needs to match the name of the
existing REM output folder in the output directory.
set /p NatForest=Enter Folder name for Forest:
set /p outputName=Enter output file name:

REM allows user to enter the file list names to be used in the raster
```

```
merge.
set /p FileList=Enter name of .txt file that contains files to be merged:
set /p FileList_patch=Enter name of .txt file that contains newest files to
patch older areas.

set output=G:\FinalCHM

cd %NatForest%

MergeRaster /verbose /overlap:new %output%\%outputName%.asc %FileList%
%FileList_patch%

echo %NatForest% processed.
set LTKLOG=
```

Fig. 4 shows an example script that can be used to merge Canopy models where we have older files and newer files that need to be represented. With the script presented here, pixels in the final image will represent the most up-to-date data for the space it is located in.

05: Clean Canopy Model Rasters

After the Canopy Models have been merged, they will likely represent an area that is slightly larger than the forest area we intend to look at. The final step in this process is to pull the newly created ASCII Canopy model files into a GIS software (We used ESRI ArcGIS Pro at the GAMLab). We can then use existing shape files for the desired forest areas to extract just the raster data that is relevant to our study area. We can also use this time to do a final quality check on the canopy model product.

1. **Open the merged canopy model product in ArcGIS Pro (or GIS of choice).**
    a. The first time you open the ASCII file in ArcGIS Pro, it will take longer than it will the second time. This is because ArcGIS will build pyramids and calculate statistics for the ASCII file, and write the data to an auxiliary exchange file associated with the raster. This will allow ArcGIS pro to be more efficient in displaying the raster, and allow it to open more quickly in the future.

2. **Inspect the Raster image.**
    a. Are there any holes that were not present when doing your quality Assurance check in stage 03?
    b. Does any of the data look incorrect or obviously corrupted?

3. **Use The ArcGIS Pro geoprocess *Extract by Mask* to remove extra data from beyond the National Forest Area.**
    a. *ArcGIS Pro - Extract by Mask*
    b. Parameters:

        i. Input Raster
            1. Merged forest area raster
        ii. Input Raster or Feature Mask Data
            1. Polygon of the USFS forest area
        iii. Output Raster
            1. Named to differentiate output from initial input
            2. Output raster to a geodatabase.
        iv. Extraction Area
            1. Inside
        v. Analysis Extent
            1. Use the merged Forest area raster to set extent.

4. **Export Raster as ASCII file.**
    a. *ArcGIS Pro - Raster to ASCII*
        i. Parameters
            1. Input Raster
                a. Raster created in geodatabase in step 3.
            2. Output ASCII Raster File
                a. Output to final location.

Future Considerations

The challenges of Creating Canopy models using FUSION come down largely to data management. Storing large quantities of LiDAR data and keeping it organized while using Windows batchscript is not particularly efficient, as Windows batchscript has many idiosyncrasies around file path navigation and is not commonly used enough for data management tutorials to be readily available online.

We have identified two alternative approaches that could be explored in the future; Quick Terrain Modeler (QTM), a LiDAR exploitation software from Applied Imagery, or an open-source approach using the Python packages GDAL (Geographic Data Abstraction Library) and PDAL (Point Data Abstraction Library.)

Quick Terrain Modeler

QTM is a software package explicitly for LiDAR data exploitation. It includes robust scripting abilities for batching processes on large LiDAR tile acquisitions, and there is existing research on it's effectiveness in the development of Canopy Model products. (Zaki et al., 2020)

Main advantages of QTM:

- Ease of Use
- Designed explicitly to maximize a workstations parallel processing capabilities

- Self-contained; no need to have multiple batch script files for each process run
- Created and maintained by a team of developers; plenty of up-to-date documentation on tool set

Main disadvantages of QTM:
- Paid software license

(*Features - Applied Imagery*, 2015)

The GAMLab has access to a QTM license through the University of Colorado.

Alternative Open-Source Approach

In a blog post titled *Producing a Canopy Height Model (CHM) from LiDAR* the author details a methodology for producing canopy height models using the Python packages GDAL and PDAL. (Planzer, 2020) GDAL is a library intended for working with geospatial data and PDAL is a library intended for working with Lidar and three-dimensional point data. The aforementioned paper provides a detailed guide for building pipelines to create Canopy Height Models, including the pre-processing steps of deriving DSMs (Digital Surface Model) from the point-cloud file, and the post-processing steps of merging together model outputs. (Planzer, 2020)

One area where this method, and our current method executed through USFS FUSION could be improved is data validation of LiDAR data and model outputs. A second paper, *An Improved Generalized Hierarchical Estimation Framework with Geostatistics for Mapping Forest Parameters and Its Uncertainty: A Case Study of Forest Canopy Height* discusses a process of using a training dataset to test the accuracy of a simulated LiDAR sampling technique for the creation of Canopy Height Models. The process workflow can be seen outlined in Fig 5. (Zhao et al., 2022)

Fig. 5

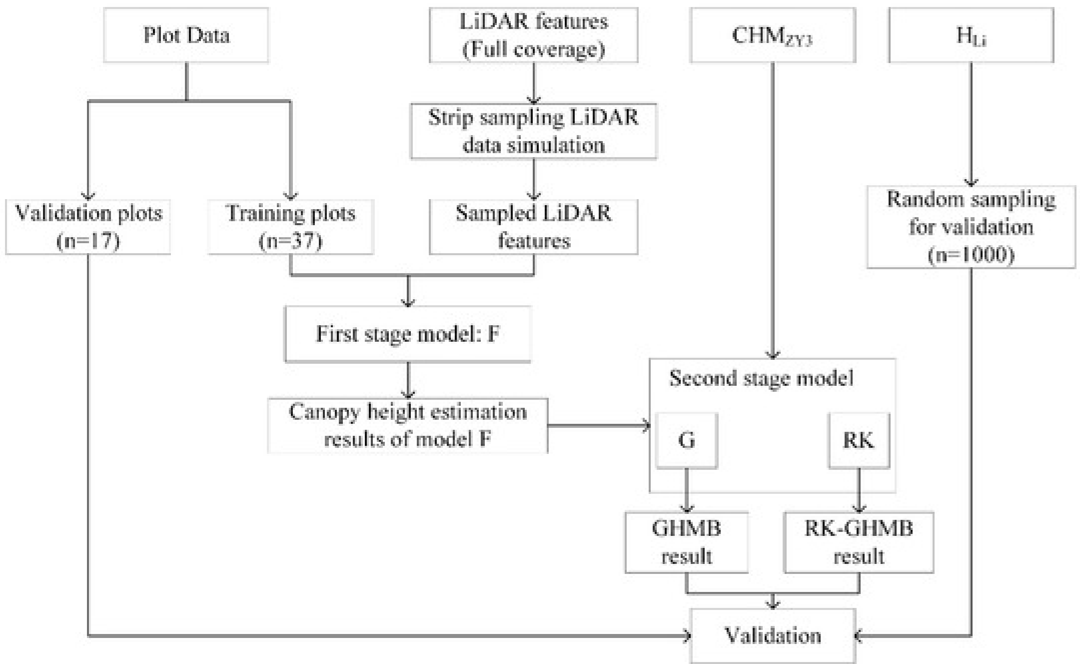


Fig. 5 shows the Zhao et al. team's study workflow, including data, models, results, and validation. (Zhao et al., 2022)

While this paper does not exactly match the kind of Canopy Modeling the GAMLab has performed for the USFS, it does present an approach that could be adapted for our use. An adapted version of this data validation process could be included as an intermediary stage in the post-processing segment of our workflow. We could use a lightweight, low-spatial resolution (10-20 meters) DSM to validate the individual DSM tiles created during processing.

In conclusion, there are many papers on small-scale and ultra-large-scale Canopy Model creation, but not so many that describe a process for developing intermediate-sized canopy models. If the USFS would like to continue using Free and Open-source software for canopy modeling, this approach utilizing existing Python libraries could make processing multiple areas more efficient and organized.

Project 2: Parcel Ownership Database

Purpose

The purpose of Project 2 is to create a searchable and comprehensive database of ownership information (owners and addresses) for parcels within national forests in Florida, South Carolina, and Puerto Rico. The final product application allows for viewing, searching, and exporting parcel polygons and attribute data. The app was designed so the average USFS employee can interact with it without needing extensive experience with GIS software.

Data Acquisition

Parcel shapefiles and up-to-date ownership data were obtained in a number of ways depending on the state/territory in question. Parcel shapefiles for Florida and South Carolina were compiled from ArcGIS Online (AGOL) and County Assessor sources, while Puerto Rico's were acquired manually (details provided in the Puerto Rico methods, *Appendix D*).

Florida's parcel ownership data can all be downloaded from the Florida Department of Revenue. South Carolina's parcels had to be acquired on a county-by-county basis, and eventually we found it more expedient and cost-effective to purchase the data from the third-party parcel data aggregator ReportAllUSA. Puerto Rico parcel ownership data were downloaded manually from the state's online tax collection map. South Carolina and Puerto Rico methods for collecting parcel ownership data are described in *Appendix D*.

Parcel Data Summary

State	Description	USFS Forest(s)	#of Counties	# of Parcels	Last Updated
South Carolina	8th State in the Union. Capital: Columbia	2: Francis Marion and Sumter National Forest, Savannah River Site	16	44,451	Dec. 2023
Florida	27th State in the Union. Capital: Tallahassee	3: Apalachicola, Osceola, Ocala	16	39,558	Dec. 2023
Puerto Rico	Sovereign State of the US. Capital: San Juan	1: El Yunque	13	679	Dec. 2023

Data Sources

State, County, and National Forest Boundaries:

- AGOL

Puerto Rico:

- PR Parcel Ownership Data: https://catastro.crimpr.net/cdprpc/
- PR Parcel Shapefiles: AGOL - PonceGIS

South Carolina:

- County Assessor Websites
- ReportAllUSA - third-party real estate aggregator

Florida:

- Florida parcel shapefiles (FL Department of Revenue): (https://floridarevenue.com/property/dataportal/Pages/default.aspx?path=/property/dataportal/Documents/PTO%20Data%20Portal/Map%20Data)
- Florida parcel ownership (FL Department of Revenue): (https://floridarevenue.com/property/dataportal/Pages/default.aspx?path=/property/dataportal/Documents/PTO%20Data%20Portal/Tax%20Roll%20Data%20Files/NAL)

Data Attributes

Attribute Name	Description
parcel_id	The number assigned to the parcel by the tax assessor of a particular jurisdiction for purposes of identification and record-keeping.
owner	The individual or organization which legally owns the land parcel.
parcel_address	The street and number of the physical address for the parcel.
parcel_municipality	The city of the physical address for the parcel.
parcel_state	The state of the physical address for the parcel.
parcel_zip	The zip code of the physical address for the parcel
mail_address1	The street and number of the owner's address
mail_address 2	Additional address information for the owner's address

mail_address3	The city, state, and zip code of the owner's address
parcel_county	The county that the parcel is within.
parcel_forest	The USFS National Forest that the parcel is within.

Methods

Creating maps and databases for parcel ownership within USFS National Forests involved acquiring parcel shapefiles, identifying which parcels are within USFS lands, collecting up-to-date parcel ownership data, appending ownership data to the parcel GIS layers, and uploading to AGOL. Parcel shapefiles were obtained from ArcGIS online, government websites, and proprietary sources. Parcel ownership data were downloaded from government websites and proprietary sources. All mapping products were produced in ArcGIS Pro.

The following section illustrates the general workflow for compiling the data required, mapping and data processing, and uploading the products to the AGOL platform. *Appendix D* provides details and screenshots to explain the specific challenges of this process for South Carolina and Puerto Rico.

General Workflow (Flowcharts in *Appendix B*)

01: GIS Layer Acquisition

1. Collect boundary layers:

a. USFS National Forests
b. States
c. Counties

2. Add layers to map in ArcGIS Pro and apply an appropriate projection for the given region.

02: Identifying Parcels in National Forests

3. Select out which counties overlap with USFS land

a. ***Select by Location*** geoprocessing tool
i. Input Features: County layer
ii. Relationship: Intersect
iii. Selecting Features: National Forest layer
iv. Right click county layer > Selection > Make Layer From Selected Features

4. Obtain parcel layers for the selected counties. Data can be found on:

a. ArcGIS Online
b. County Property Appraiser Websites
c. State Department of Revenue Websites
d. (see *Appendix D* for the Method used for Data acquisition in Puerto Rico and South Carolina.)

5. Create layers of parcels within USFS land for each county ("USFS parcel layers")

a. Add parcel layers to ArcGIS and apply map's projection
b. ***Intersect*** geoprocessing tool
   i. Input Features: (1) a county's parcel layer, (2) National Forest layer
   ii. Assign name to output feature class
   iii. Click Run
   iv. Repeat for all counties

03: Attribute Table Preparation

6. Remove all attribute table columns except:

a. Object ID, Shape, Parcel ID, Shape Length, and Shape Area
   ii. Make sure column names are the same across all USFS parcel layers
   iii. Add State, County, and Forest Name fields. Populate with ***Calculate Field***

7. Save all layers created to a geodatabase

04: Obtaining up-to-date ownership data

1. Locate and download ownership data for the given year. Data can be found on:

a. Property Appraiser Websites (free and for purchase)
b. State Department of Revenue Websites (free, if available)
c. Proprietary Websites (e.g., ReportAll.com, catastro.crimpr.net; for purchase)

05: Cleaning Data

2. For each ownership data CSV, select out columns for:

a. Parcel ID, Owner Name, Owner Address, and Parcel Address
b. Ideally do this with R. Excel automatically changes long numbers (Parcel IDs) to scientific notation and it is difficult to stop it from doing this and revert them back.
   i. Custom R script, *parcel_csv_reading_cleaning*, does this for an entire folder of CSVs

Tip: At this point change the column names of these attributes according to the data attributes listed in Project 2 // Data Acquisition section of this document

c. Ensure that the column names are the same across all ownership data CSVs.
d. Make sure that the Parcel ID column name is the same as that for the USFS parcel layer attribute tables, else joining them will not work.

3. Clean Data: remove spaces/special characters from Parcel IDs and convert them to text data.

a. The custom R script, *remove_dashes_from_data,* can remove all spaces or special characters from data in the Parcel ID columns of the ownership data CSVs.The custom R script, *converting_to_character_class,* can convert data in the Parcel ID columns to text (i.e., character) data.

4. Upload selected and cleaned CSVs to the ArcGIS project.

5. Compare CSV data to USFS parcel layer attribute tables.
Look for:

a. Parcel ID column names, make sure these match
b. Parcel ID column data types, both should be text data. The join will not work if they are different data types. If layer data is not text:
   i. Add a new blank field for parcel ID, set the Data Type to "Text"
   ii. *Calculate Field* for the blank parcel ID column, set equal to the old parcel ID column
   iii. Press Ok

Tip: Even if you don't see any spaces/special characters, remove them anyway. There are sometimes data entry errors, like spaces after the parcel IDs. These will lead to mismatches when joining the data.

06: Joining Parcel Ownership Data to USFS Parcel Layers

6. Join ownership data CSVs to the attribute tables of the USFS parcel layers

a. Use the ***Join Field*** geoprocessing tool to permanently join the tables
   i. Input Table: USFS parcel layer
   ii. Input Join Field: Parcel ID
   iii. Join Table: Ownership data CSV
   iv. Join Table Field: Parcel ID
   v. Validate Join and diagnose mismatches
   vi. Run the join

7. Merge all county USFS parcel layers with joined ownership data into one layer.

a. Use the ***Merge*** geoprocessing tool
   i. Input Datasets: enter each county's joined USFS parcel layer
   ii. Name the output layer
   iii. Run

8. Save all layers created to the geodatabase

07: Merging USFS parcel layers for all States/Territories

1. Create a new map in ArcGIS Pro

a. Add USFS parcel layers with joined ownership data for each State/Territory from their respective geodatabases

2. Merge all State/Territory USFS parcel layers into a single layer.

a. Use the ***Merge*** geoprocessing tool
   i. Input Datasets: enter each State/Territory USFS parcel layer
   ii. Name the output layer
   iii. Run

3. Save the merged layer in its own geodatabase

08: Uploading to ArcGIS Online

4. Navigate to the location of the geodatabase in the File Explorer

5. Right click it and select *Send to > Compressed (zipped) folder*

6. Log in to ArcGIS Online from a browser

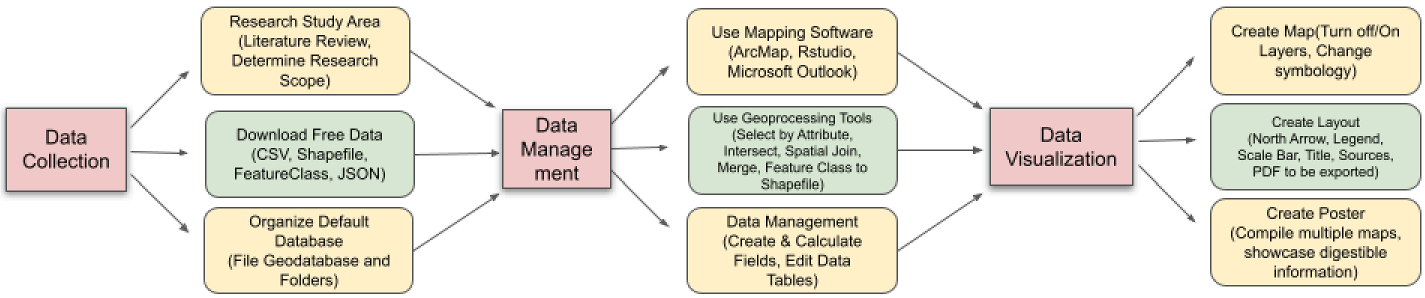
7. Go to Content

8. Press “New item” and drag/drop the zipped geodatabase or select it from “Your device”

9. When prompted for File Type, select “File geodatabase”

10. Give it a title, select the appropriate folder to save it to, add tags, and a summary

Fig. 6



General Work Flow of Data Organization. This shows the 3 phases of collecting and interpreting data.

Deliverable

Fig.7

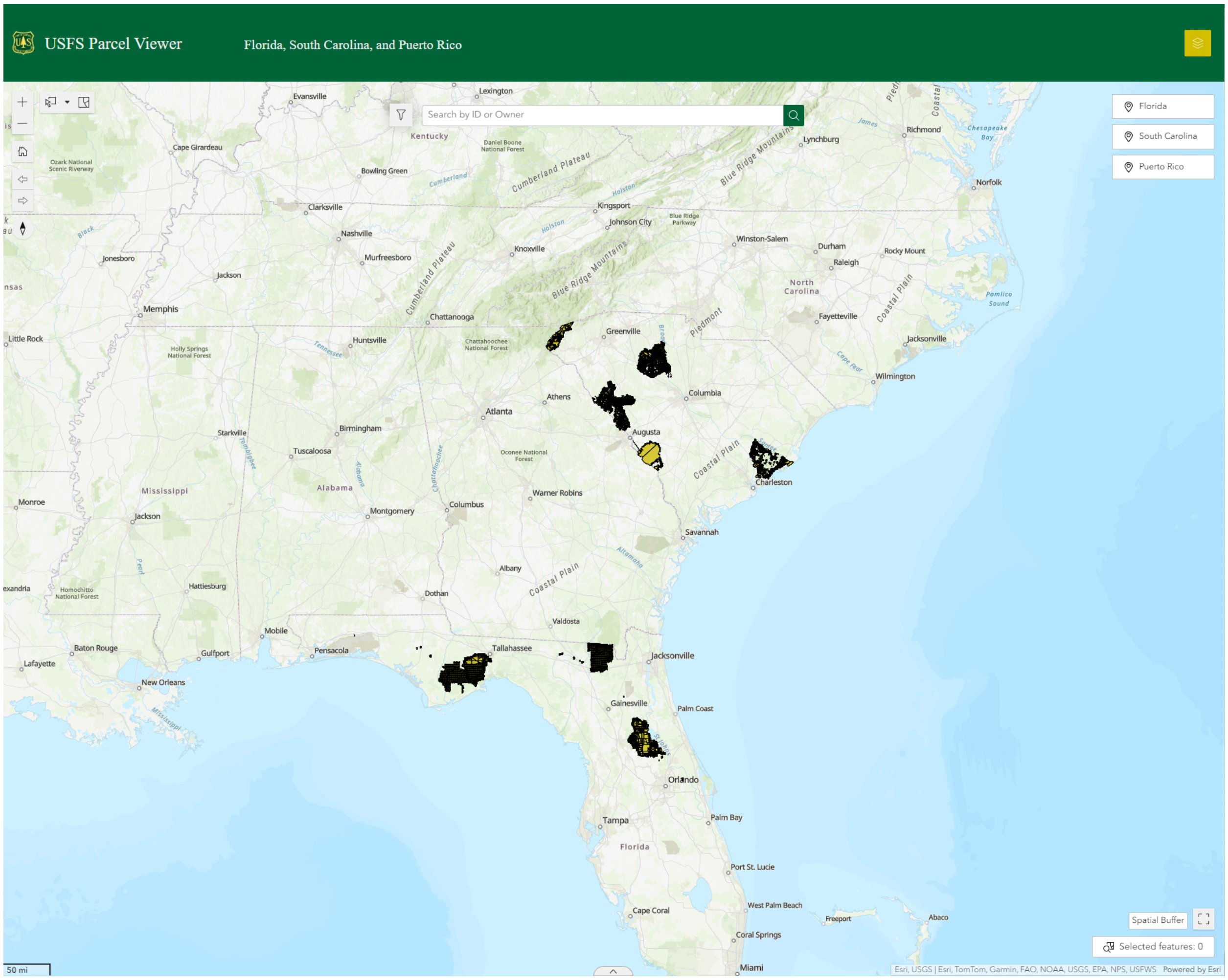


Fig. 7 shows a screencapture of the USFS Parcel Viewer Product produced by GAMLab.

USFS Parcel Viewer

The result of this project is an interactive map that allows USFS users to easily select parcels near or in National Forest Areas, and to provide users ownership contact information for those parcels.

Some Details about the USFS Parcel Viewer:

- USFS Parcel Viewer was developed in ArcGIS Experience Builder.
    - See Appendix C for resources with ESRI documentation of Experience Builder.
- Standardized data schema that can be updated with ease

- Parcel Viewer uses data from a Feature Service Geodatabase that can be accessed via ArcGIS Pro or ArcMap Desktop environments.
- For full details of data included in the USFS Parcel Viewer, refer to **Project 2 // Data Acquisition** portion of this report.

For a step-by-step guide to using the USFS Parcel Viewer tool, refer to the *USFS Parcel Viewer User Guide*.

Appendix A

Canopy Model Work Flows

Fig. 8

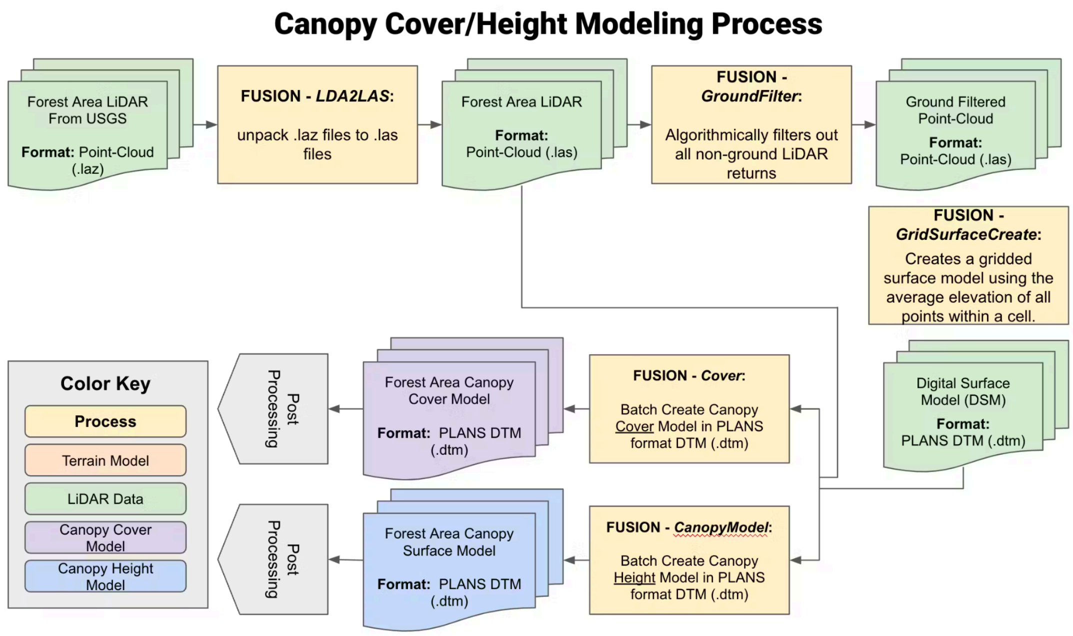


Fig. 8 shows a flowchart describing the process of deriving Canopy Model Products (CCMs &1 CHMs) from 3DEP LiDAR Files.

Fig. 9

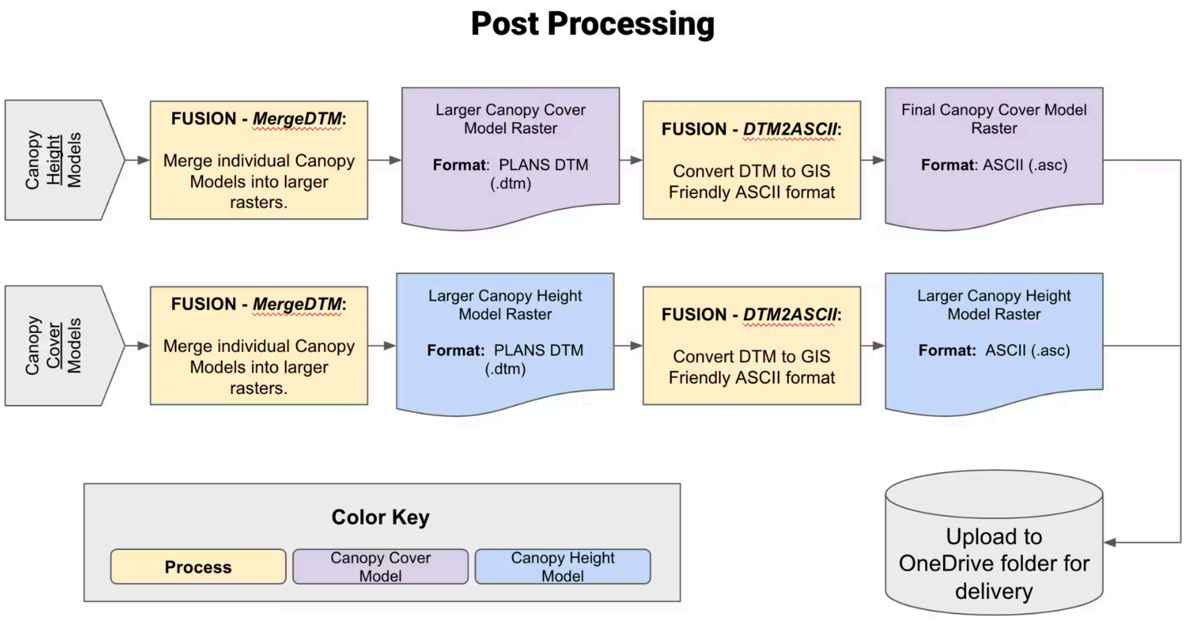


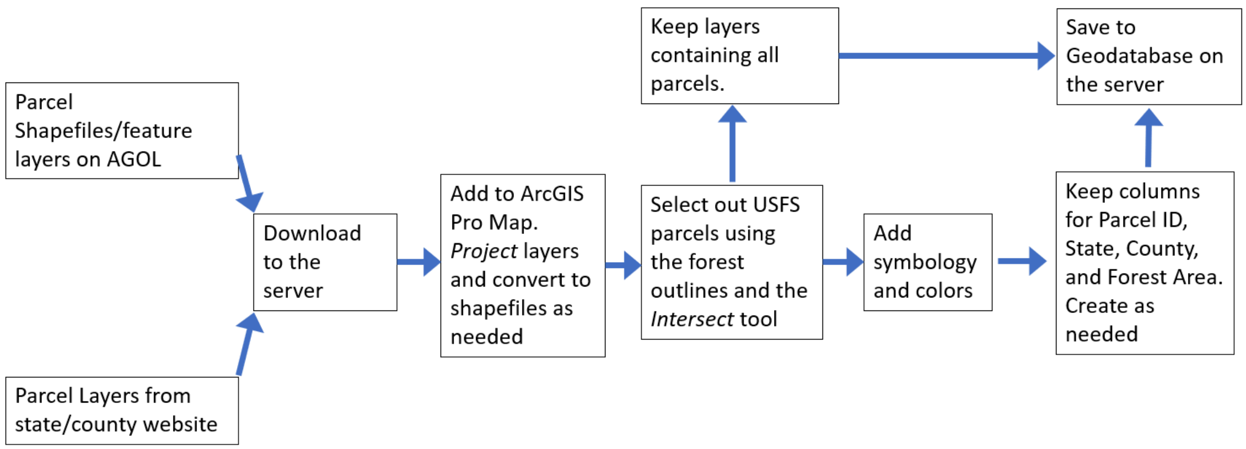
Fig. 9 shows a flowchart describing the Canopy Model Post-Processing workflow.

Appendix B

Parcel Acquisition Work Flows

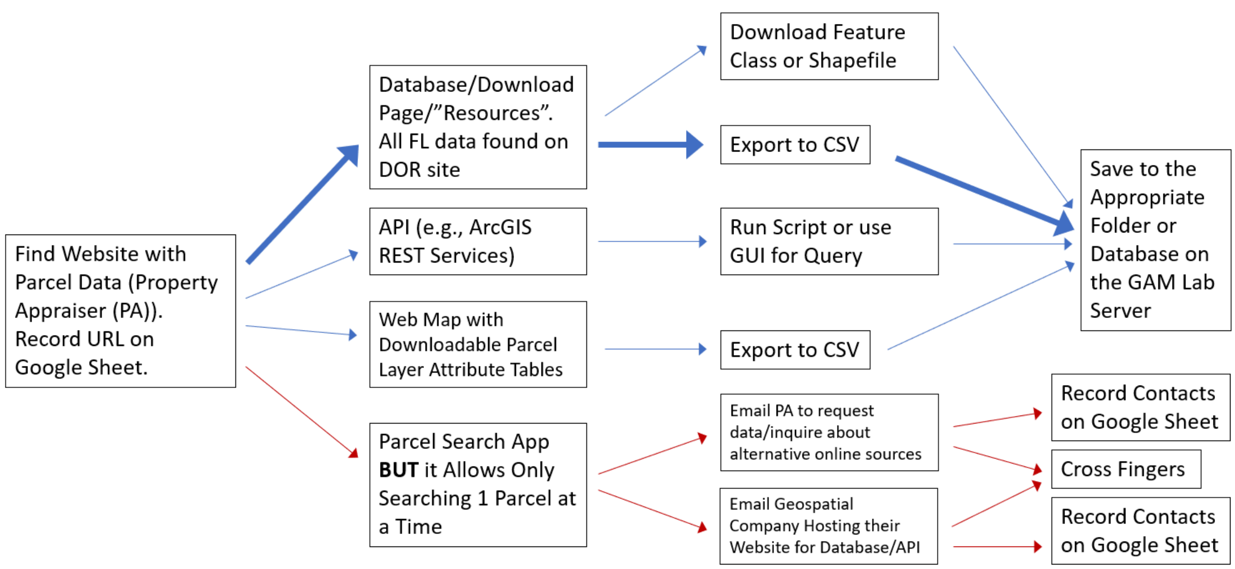
Obtaining Parcel Layers and Mapping

Fig. 10



Up-to-Date Ownership Data Acquisition

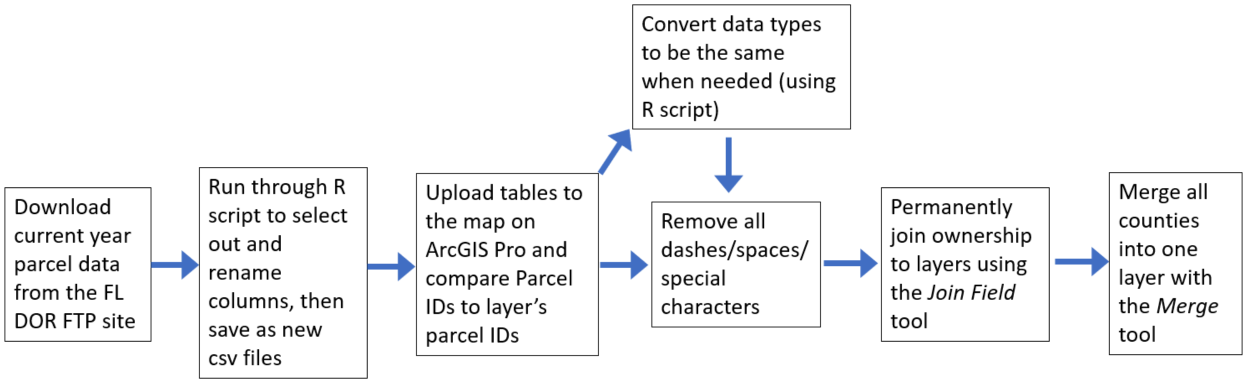
Fig. 11



The flowchart in Fig. 11 visualizes all paths to acquiring ownership data from the counties. Blue paths are more ideal. The bold path is what ended up working best for Florida.

Downloading and Joining Ownership Data to Parcels

Fig. 12



Appendix C

Software Documentation + Resources

USFS FUSION User Manual
https://www.fs.usda.gov/research/pnw/products/dataandtools/tools/fusion/ldv-lidar-processing-and-visualization-software-version-4.40

USGS instructions for using uGet for bulk data acquisition
https://apps.nationalmap.gov/uget-instructions/

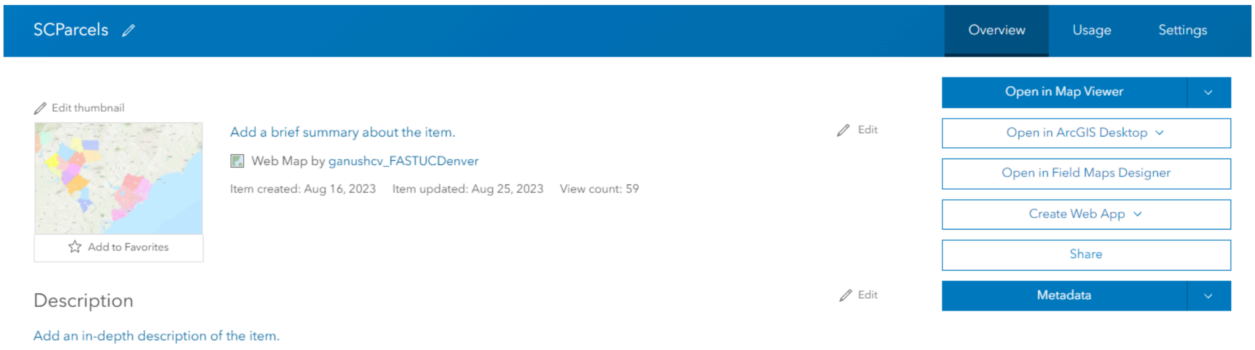
ArcGIS Experience Builder Tutorial Resources
https://www.esri.com/en-us/arcgis/products/arcgis-experience-builder/resources

ArcGIS Experience Builder Youtube Videos
https://www.youtube.com/playlist?list=PLGZUzt4E4O2K7GmDAAQagTlqBWuufDF9F

Appendix D

Alternative Parcel Data Acquisition

South Carolina



The US Forest Service encompasses 16 counties within South Carolina. In order to build the map for south carolina we downloaded the USFS boundary, SC counties and USA state boundary shapefiles. We then created a geodatabase in ArcGIS Pro and uploaded the respective shapefiles to an ocean base map available through the ArcGIS online atlas.

We intersected the forest boundaries with county boundaries and removed areas that were not relevant. As we downloaded ownership information for each county, we would jon the data to the relevant parcels we had put together in ArcGIS Pro.

Puerto Rico

Ownership data for PR was manually downloaded as parcel information from Puerto Rico's online GIS map PRIM (https://catastro.crimpr.net/cdprpc/).

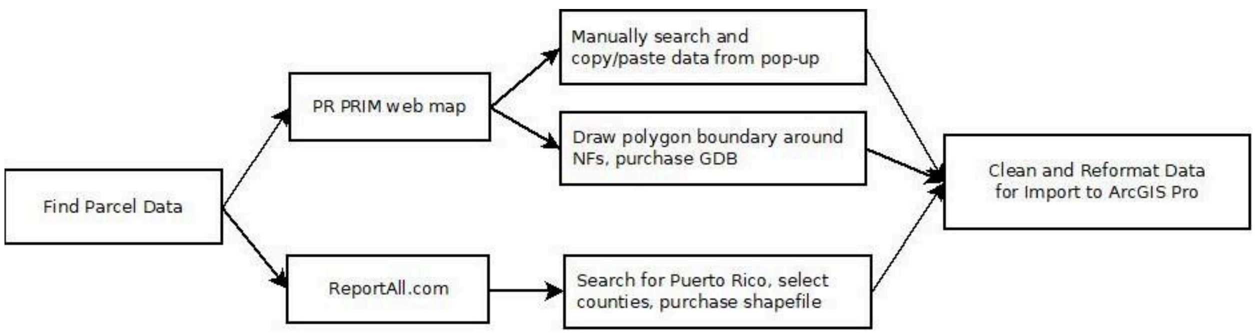
(1) A zipped file of the forest boundaries was loaded into PRIM, the extent was zoomed in until parcel boundaries appeared, and then parcels touching forest boundaries were manually selected and data from the pop-up was copy/pasted. Parcel information included parcel IDs, owner names, and addresses. Copied parcel data was then cleaned in Excel and saved as a CSV for importing into ArcPro.

(2) A parcel shapefile (Parcelas) was found via AGOL and loaded into the desktop map. This layer was then cut down to only counties (Municipios) where a National Forest existed and exported to create a new, smaller layer. New columns were added in the Attribute Table for Owner, Address, County, and GIS Area (a column for P_ID may need to be added and calculated as a Text type with info from the ParcelID field for Joining). The cleaned CSV with ownership data was then loaded, Exported, and Joined to the new Parcels layer (P_ID as Text type field may need to be added and calculated here as well). The Field Calculator tool was

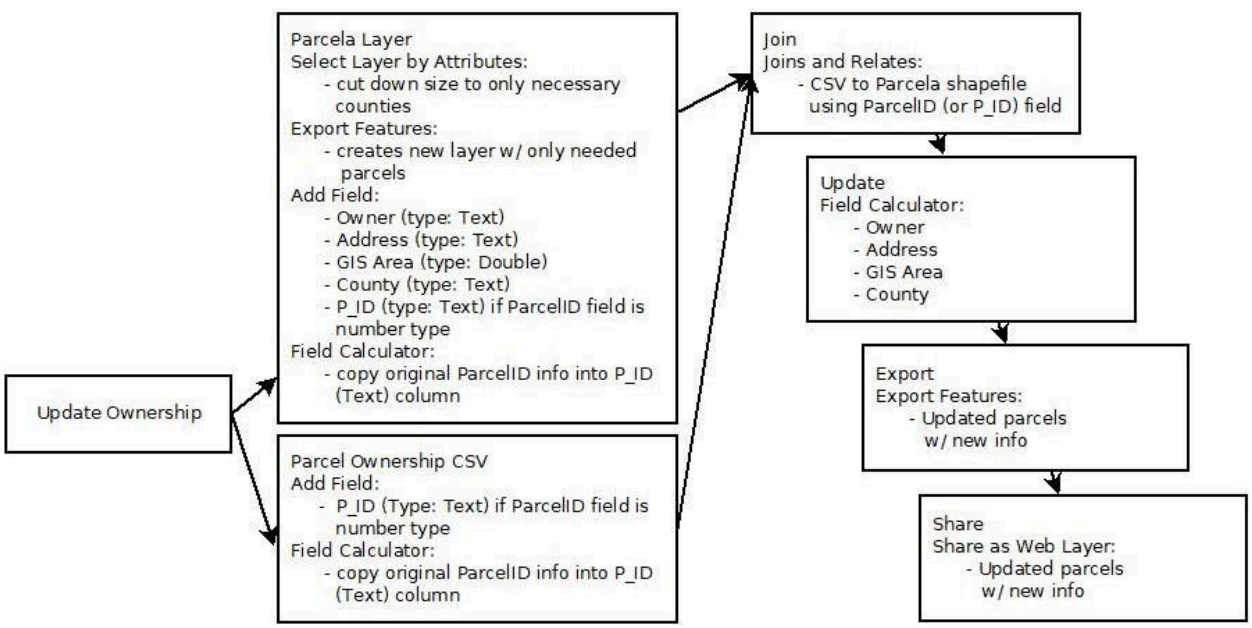
used to copy info from the Ownership Table to the Parcels shapefile and then the Join was removed.

For this project, the final updated Parcel layer was then exported and sent to GAMLab for addition to the AGOL map. In the future, a desktop map will likely exist where Ownership info can be updated and then the layer can be shared via the Share as Web Layer tool rather than exporting. Flowcharts below include steps for two other ways to update ownership data, however, these other methods cost money while the currently outlined method is free.

(1) Find Parcel Data Flowchart:

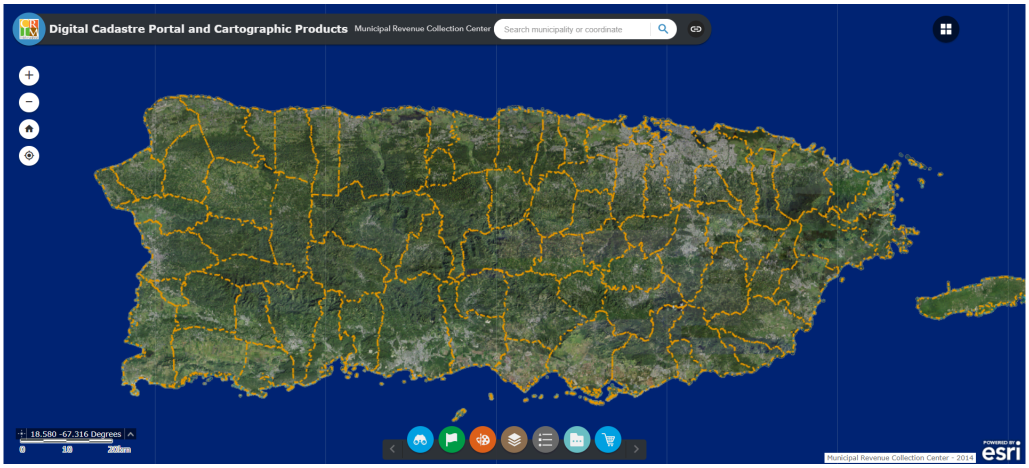


(2) Update Ownership Data Flowchart:



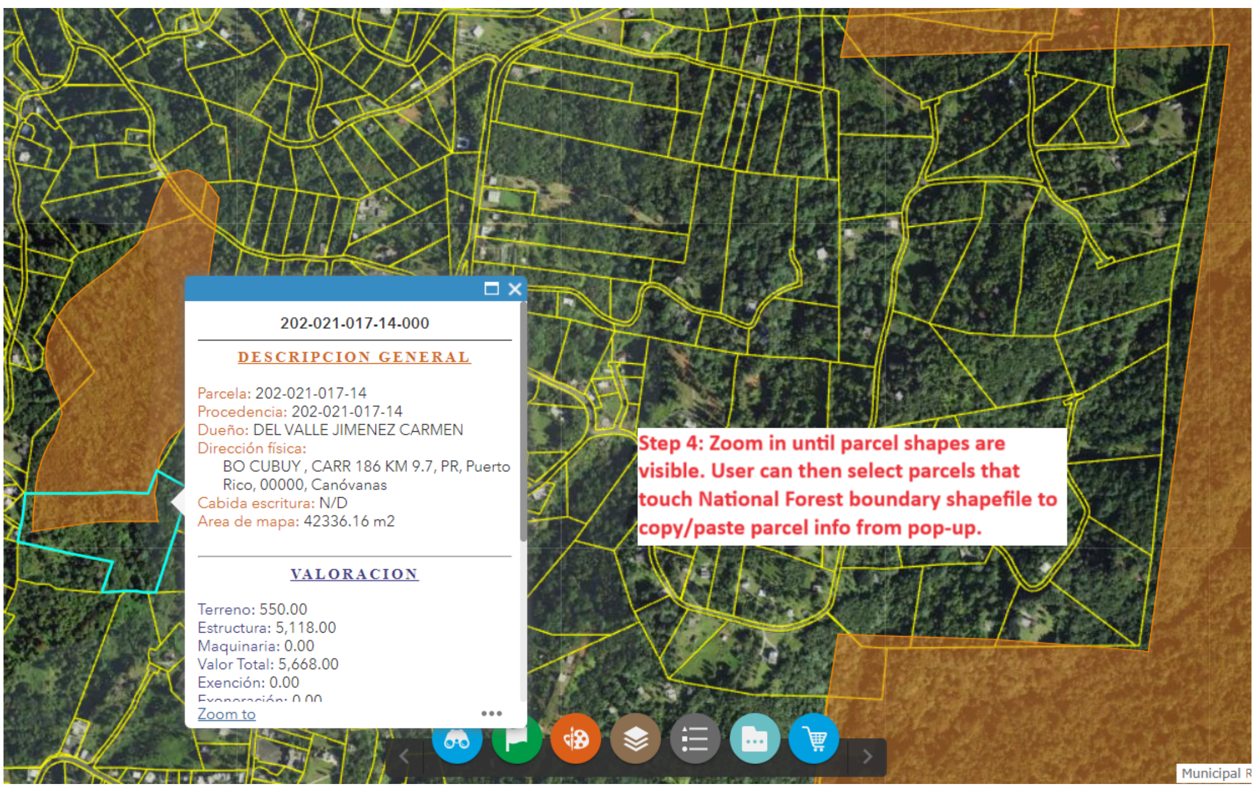
Step-By-Step Method for Puerto Rico:

Open web map from https://catastro.crimpr.net/cdprpc/:



Add National Forest zipped shapefile:

Zoom to parcel visibility and select parcels to copy/paste data:



After all parcel data has been collected from the PR web map it will need to be processed in the chosen spreadsheet application (Excel, Libre OfficeCalc, etc).

Methods

01: GIS Layer Acquisition

1. **Collect boundary layers:**
    a. a. USFS National Forests
    b. b. Counties
    c. c. Parcelas layer
        i. Open Catalog pane
        ii. Click on ArcGIS Online tab (cloud)
        iii. Search: “Puerto Rico parcela”
        iv. Select: PARCELAS layer by Ponce GIS
2. **Add layers to map in ArcGIS Pro and apply an appropriate projection for the given region.**

02: Identifying Parcels in National Forests

3. **Select out parcels that overlap with USFS land**
    a. ***Pairwise Buffer*** geoprocessing tool
        i. Input Features: National Forest layer
        ii. Outpuet Feature Class: national_forest_buffer_100ft
        iii. Distance: 100 US Survey Feet
    b. ***Select by Location*** geoprocessing tool
        i. Input Features: Parcelas layer
        ii. Relationship: Intersect
        iii. Selecting Features: national_forest_buffer_100ft layer
        iv. Right click county layer > Selection > Make Layer From Selected Features

03: Attribute Table Preparation

4. **Update attribute table columns:**
    a. Remove all columns except: Object ID, Shape, Parcel ID, Shape Length, Shape Area
    b. Check: Parcel ID column must be Text field
        i. If not: Add new field
        ii. Name: P_ID
        iii. Alias: P_ID
        iv. Type: Text
        v. Populate with ***Calculate Field*** using original Parcel ID column
    c. Add columns for: County and Forest Name

i. Populate with ***Calculate Field*** tool

5. **Save layer to geodatabase**

04: Obtaining up-to-date ownership data

1. **Locate and download ownership data for the given year. Data can be found on:**
   catastro.crimpr.net

2. **See Appendix D for Method used to obtain parcel ownership info for Puerto Rico**

05: Cleaning Data

3. **In ownership data CSV, select out columns for:**
   a. Parcel ID, Owner Name, Owner Address, and Parcel Address
   b. Ideally do this with R. Excel automatically changes long numbers (Parcel IDs) to scientific notation and it is difficult to stop it from doing this and revert them back.
      i. Custom R script, *parcel_csv_reading_cleaning*, does this for an entire folder of CSVs
4. **Clean Data:**
   a. Column Headers
      i. Translate titles within cells from Spanish to English
      ii. Highlight all data > right click > copy
      iii. Find empty cell below data > right click > paste-Transpose
      iv. Delete horizontal data
      v. Add column header row and title appropriately
      vi. Find and Replace tool: use to remove all titles from within data cells
   b. Parcel ID as Text field
      i. Click column header for Parcel ID to highlight entire column
      ii. Right click > Format Cells > Text
   c. Update Symbology
      i. Highlight all data
      ii. Choose 1 font type for all, turn off Bold or Italics, choose 1 font size
   d. Save
   e. Export as CSV
5. **Upload cleaned CSV to ArcGIS project**
6. **Update attribute tables for imported ownership CSV and USFS parcels layer:**
   a. Check for Parcel ID (or P_ID) columns names that are Text in both tables
   b. In USFS parcels layer, add columns for:
      i. Owner
      ii. Owner Address
      iii. Parcel Address

06: Joining Parcel Ownership Data to USFS Parcel Layers

7. **Join ownership data CSV to the attribute table of the USFS parcel layer**
    a. Use the ***Join Field*** geoprocessing tool to join the tables
        i. Input Table: USFS parcel layer
        ii. Input Join Field: Parcel ID (or P_ID)
        iii. Join Table: Ownership data CSV
        iv. Join Table Field: Parcel ID (or P_ID)
        v. Validate Join and diagnose mismatches
        vi. Run the join
8. **Update parcel information in USFS parcel layer:**
    a. Use ***Calculate Field*** tool to update new Owner, Owner Address, and Parcel Address fields from same fields in joined CSV
9. **Remove Join**
    a. Right click USFS parcel layer > Joins and Relates > Remove Join > Run
10. **Save layers created to the geodatabase**
11. **Merge USFS parcel layer to others**
    a. See 07: Merging USFS parcel layers for all States/Territories in Project 2: Parcel Ownership Database: Methods

Another option for updating the data is to purchase it from an online portal. The PRIM app allows for polygon drawing around the park borders to select areas for geodatabase purchase and download. This method would cost approximately $560. The other way to buy the parcel information is via ReportAll.com. The advantages to this option are that it is the quickest and most comprehensive way to gather the data. Disadvantages include that parcels must be purchased as a full county instead of selecting a specific boundary, which would require downloading much larger amounts of data, additional steps for data cleaning, and the volume of information costs a significant amount more money at approximately $2100 for all counties.

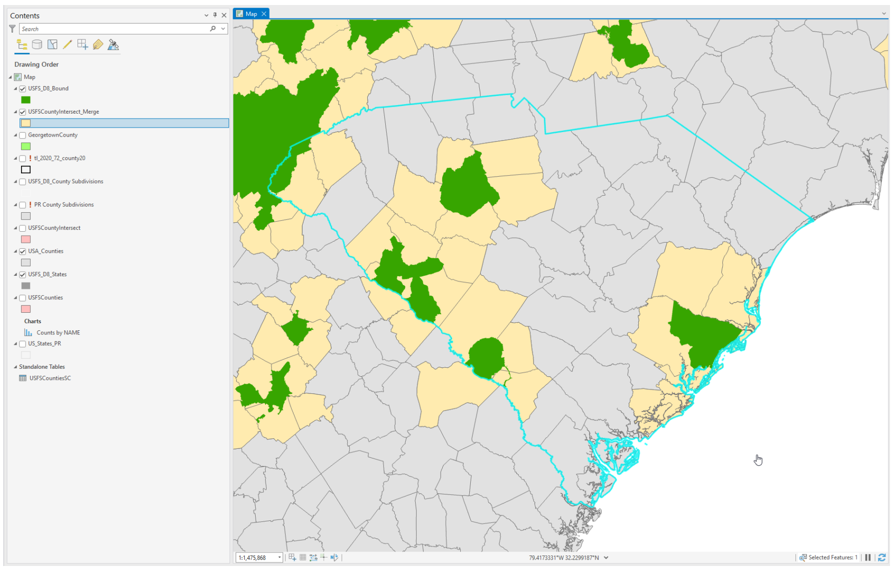
Possible ways to streamline the process of acquiring data for updating the Puerto Rico map include:

- Creating an API to pull data automatically from PRIM
- Digitizing a park boundary polygon that crosses only the touching parcels for uploading to PRIM, limiting the amount of time it requires to select and download a geodatabase from the web map
- Building a tool in ArcGIS that would quickly cut the purchased county data from ReportAll down to only necessary parcels for joining and updating.

Data Collection

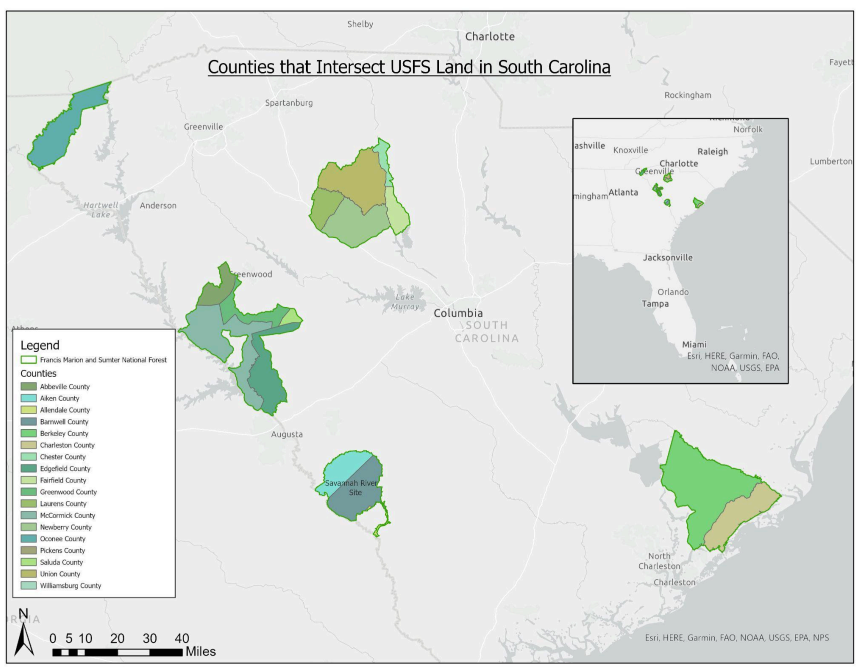
Create Relevant Databases

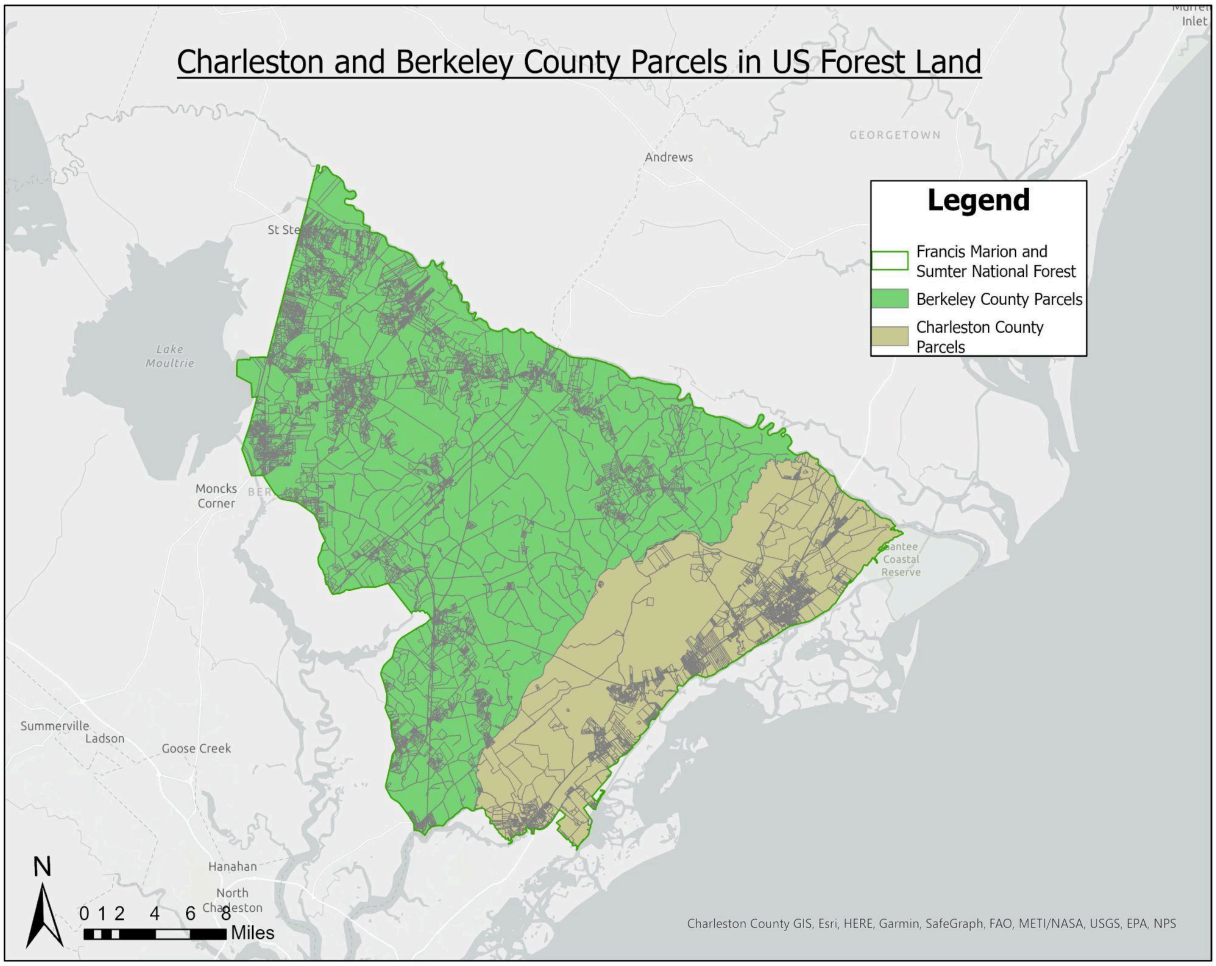
- USFS District 8 Reference.gdb
- SCParcels.gdb
- PRParcels.gdb
- FLParcels_USFS.gdb



Geoprocessing Tools

- Make Feature Layer
- Clip
- Select Layer by Location
- Select Layer by Attribute
- Intersect
- Merge
- Export Features

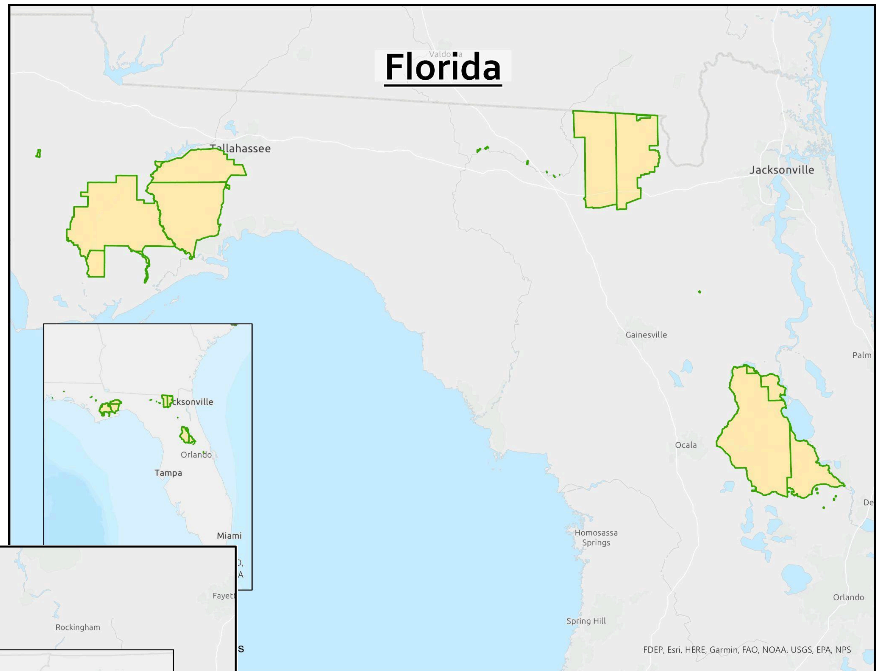


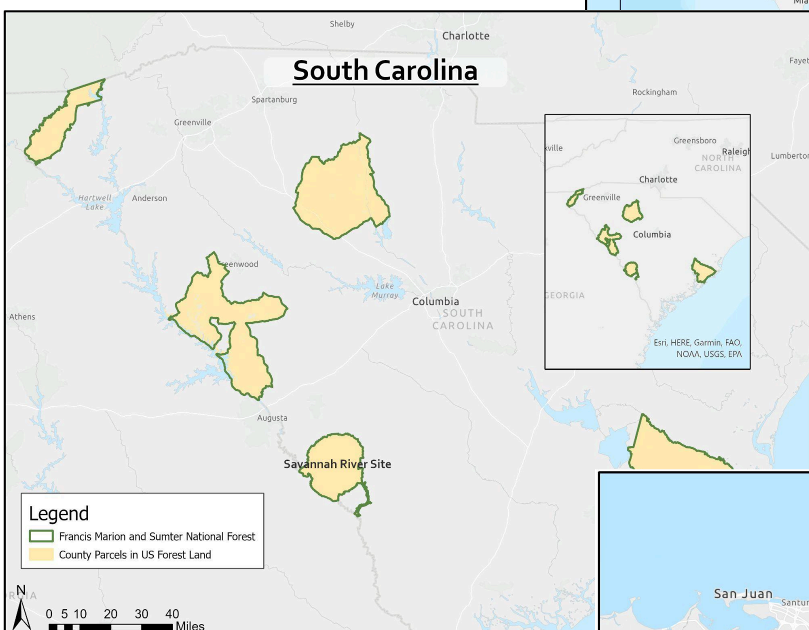


- Spatial Join
- Add Join
- Calculate Field
- Export Table
- Feature Class to Geodatabase
- Feature Class to Geodatabase
- Feature Class to Shapefile
- Feature Class to Feature Class

National Forest Areas

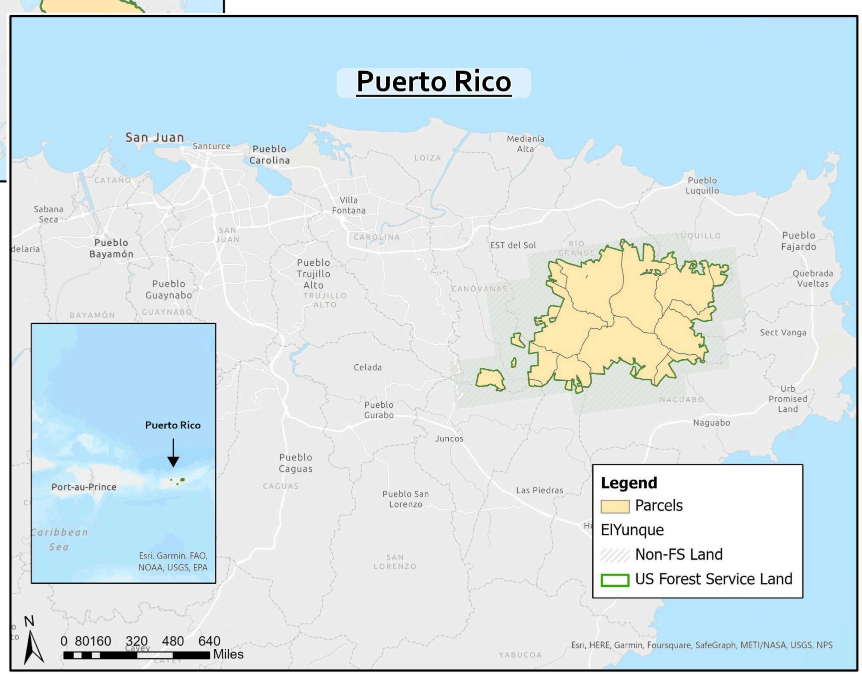
- Apalachicola
- Osceola
- Ocala





- Francis Marion and Sumter National Forest
- Savannah River Site

- El Yunque



Bibliography

Features—Applied Imagery. (2015, May 5). https://appliedimagery.com/features/

McGaughey, R. J. (2023). *FUSION/LDV: Software for LIDAR Data Analysis and Visualization*. United States Department of Agriculture. http://forsys.sefs.uw.edu/software/fusion/FUSION_manual.pdf

Planzer, S. (2020, May 7). *Producing a Canopy Height Model (CHM) from LiDAR* [Portfolio]. Simonplanzer.Com. https://www.simonplanzer.com/articles/lidar-chm/

USGS 3D Elevation Program (3DEP) Datasets from The National Map. (n.d.). [Government archive]. Sciencebase.Gov. Retrieved December 14, 2023, from https://www.sciencebase.gov/catalog/item/4f70a58ce4b058caae3f8ddb

Zaki, N. A. M., Rajuli, M. F., Latif, Z. A., Suratman, M. N., Omar, H., Norashikin, S., Zainal, M. Z., & Talib, N. (2020). Analysis of Canopy Height Model (CHM) Extraction using Quick Terrain Modeller (QTM) for Tropical Forest Area. *IOP Conference Series: Earth and Environmental Science*, *540*(1), 012045. https://doi.org/10.1088/1755-1315/540/1/012045

Zhao, J., Zhao, L., Chen, E., Li, Z., Xu, K., & Ding, X. (2022). An Improved Generalized Hierarchical Estimation Framework with Geostatistics for Mapping Forest Parameters and Its Uncertainty: A Case Study of Forest Canopy Height. *Remote Sensing*, *14*(3), 568. https://doi.org/10.3390/rs14030568