Small-Area Estimation for FIA's National Woodland Owner Survey

Investigators

PI: Paul Catanzaro (paulcat@umass.edu), UMass Amherst, Family Forest Research Center

Co-PI: Vance Harris (vharris@umass.edu), UMass Amherst, Family Forest Research Center

Other collaborators include: Jesse Caputo (jesse.caputo@usda.gov) and Brett J. Butler (brett.butler2@usda.gov), USDA Forest Service, Northern Research Station Forest Inventory and Analysis Program (FIA)

Progress Period

31 July 2024 to 31 January 2025

Progress

In the previous project period, the project team finished initial processing of all 48 CONUS states and developed a procedure for stitching them together into a single nationwide raster layer (Beta version). In addition, a new classification system was developed in which land cover and land ownership were encoded in fine resolution. This classification, with 254 individual classes, includes a number of fine distinctions useful for development and validation purposes. It can be collapsed into a smaller set of 189 classes, which includes all permutations of the full set of NLCD cover classes and a slightly-reduced version of the standard FIA ownership code typology (e.g. family-owned deciduous forest, corporate-owned grassland, state-owned scrub, etc.). This collapsed layer – once finalized – will be one of the primary public products of this project. For the purposes of mapping and visualization, this layer can be further collapsed into a number of more specialized classifications. For example, Figure 1 shows the distribution of forest and non-forest land among a reduced set of private and public ownership types (13 classes). The team also began the process of error classification and quantification. As a starting point, the full classification was collapsed into a 4-class "qualitative" error layer. This is possible because the full classification encodes for each pixel whether land cover or ownership data are missing or unknown, as well as the actual classes where they are not missing. More than 72% of pixels have both ownership and land cover data associated with them. Most of the remaining pixels (> 26%) are missing ownership data, but have known land cover. Only a small number of pixels have either ownership data only (0.06%) or are missing data altogether (0.17%). Pixels with missing data (especially missing ownership data) are disproportionately distributed among the western states, and are believed with some confidence to largely represent public ownership classes.

In the current period, we began the process of taking the base output layer and the qualitative error layer, as well as outputs from the classification process and raw input data, and calculating a quantitative error layer. For example, published pixel-level error rates from NLCD and empirical classification error rates from internal validation testing were combined to calculate the base error rate for the combined land use / land ownership classification at the pixel level. Mean combined error rates for trial states include 63.11% for Alabama and 73.05 for Arizona (see Figure 2).

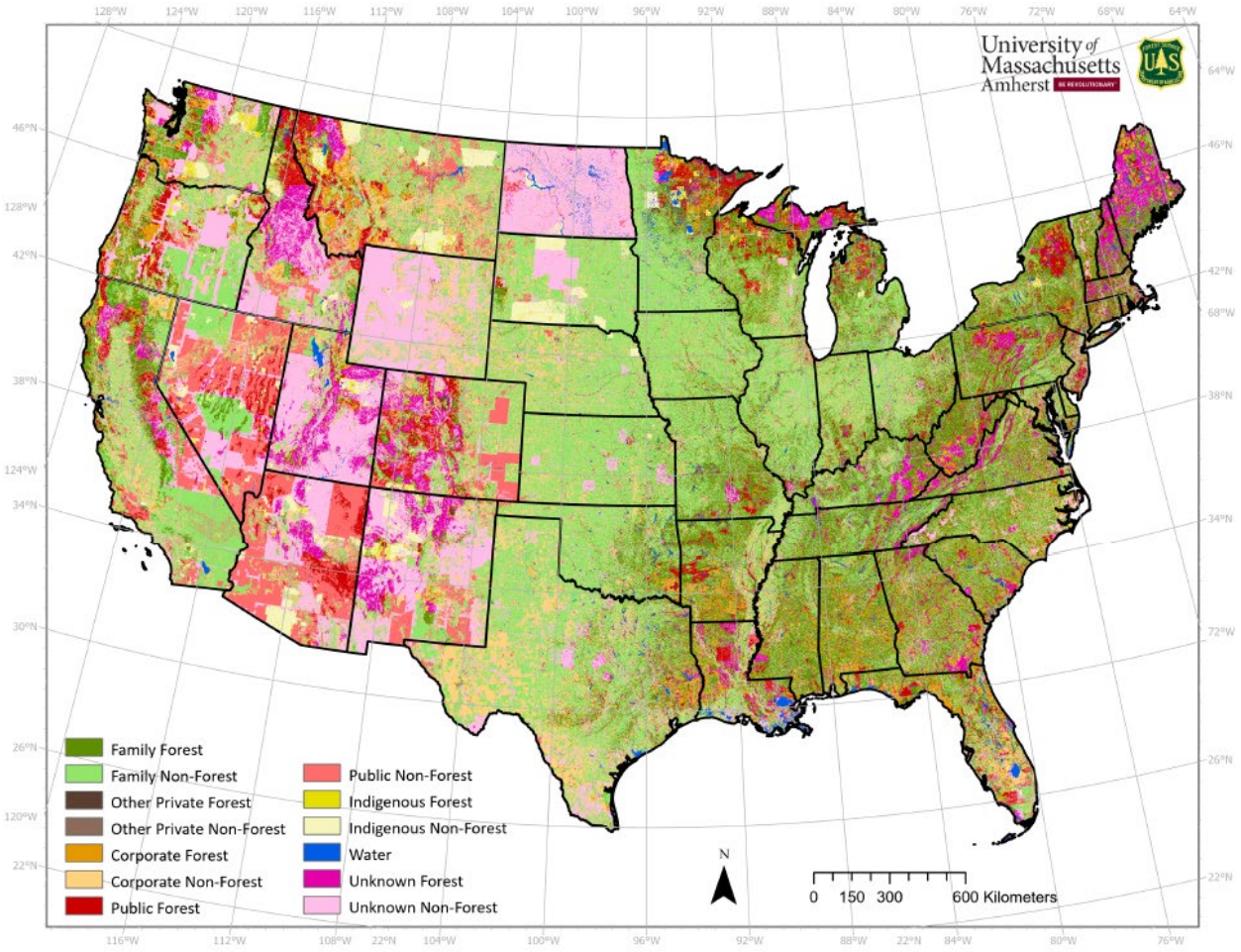

Figure 1: Completed land use and land ownership layer (BETA). CONUS.

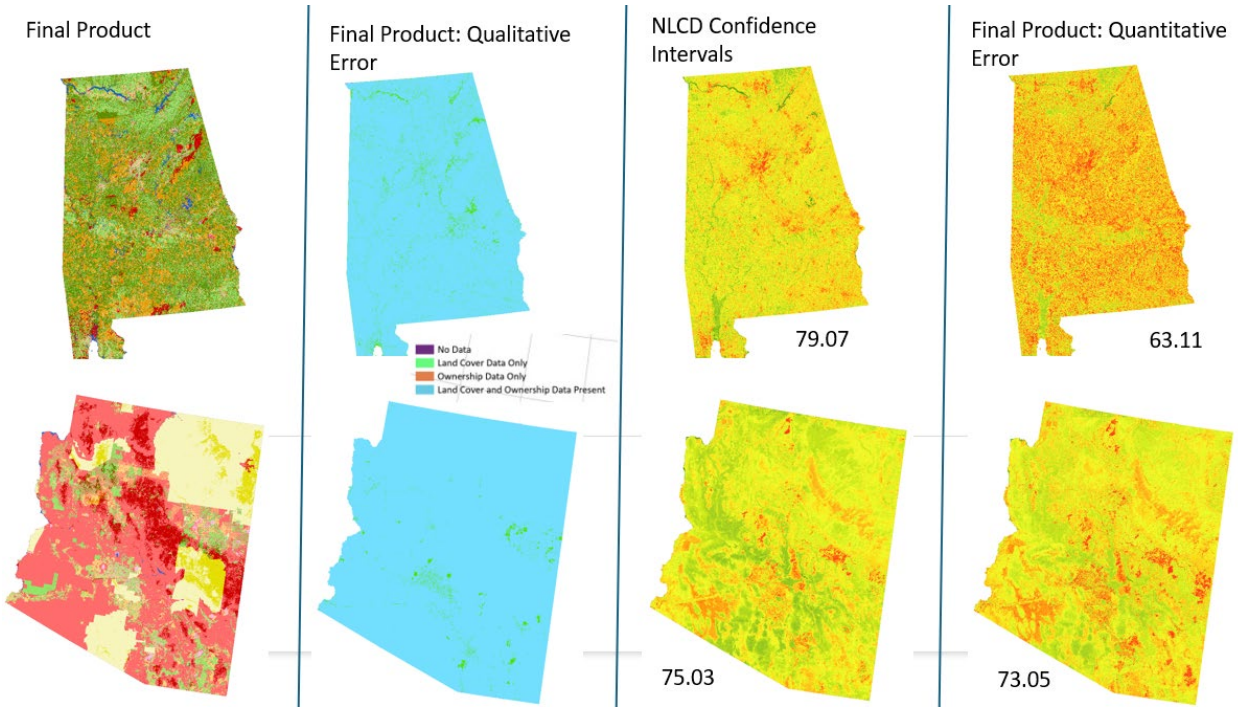

Figure 2: Land use and land ownership layer, along with qualitative and quantitative error products (BETA). Alabama (top) and Arizona (bottom).

Problems or Delays

Unfortunately, a critical error with the underlying layer generation process was discovered during this project period. The project team had to pause work on the quantitative error layer and focus time and resources on fixing and re-running the base layer. This process is now complete for most of the 48-states in CONUS and soon, efforts can once again focus on refining and completing the error estimation procedures for all states. The team has begun the process of requesting a no-cost extension in order to finish the work.

Next Period Plans

Over the next period, the team will focus on re-calculating the full data layer with the aim of locking down a final version. Once this task is complete, we will continue refining the algorithms for translating the qualitative data layer into a quantitative data layer. To do this, we will supplement the published error estimates for the raw NLCD data and within-class ownership classification error rates derived from our validation efforts with estimated error rates associated with PAD data and third-party vendor data (currently unknown). The formulae for integrating these data sources into a single quantification of total error for each pixel are still being refined.