

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 January 2025 to 31 July 2025

Progress

In the previous project period, the project team finished initial processing of the nationwide layer (soon to be published in the USFS Research Data Archive) and 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).

Recent updates to the Ownership small area estimation (SAE) project reflect our continued efforts to refine model performance through iterative evaluation, now applied to a newly completed geospatial data layer integrating land ownership and land cover. Rather than relying on aggregate or global accuracy metrics, we have transitioned to a more localized, state-level accuracy assessment framework. This approach allows us to identify regional variation in model performance and prioritize improvements where they are most impactful.

During our iterative evaluations, we found that the lowest classification accuracies consistently occurred in low-frequency ownership classes, due to a range of contributing factors, both introduced from regional differences as well as gaps in the methodology. As a result, we are placing increased emphasis on assessing these underrepresented classes to better understand the conditions driving misclassification and to guide improvements in their predictive performance. As part of our validation strategy, we continue to overlay our model's classification outputs with published error rates from source datasets, ensuring that both model and input uncertainties are accounted for in our evaluation pipeline.

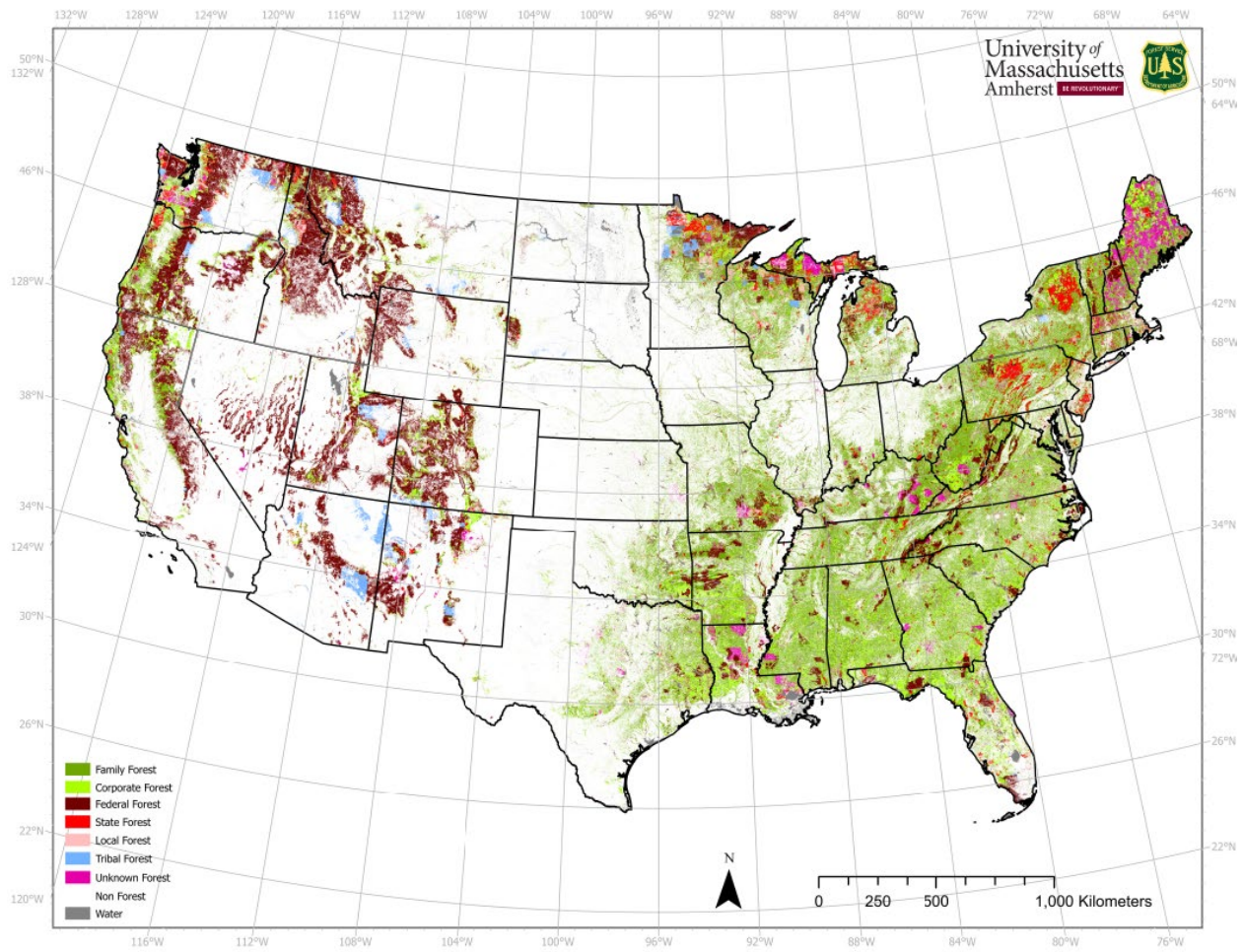


Figure 1: Completed land use and land ownership layer. CONUS.

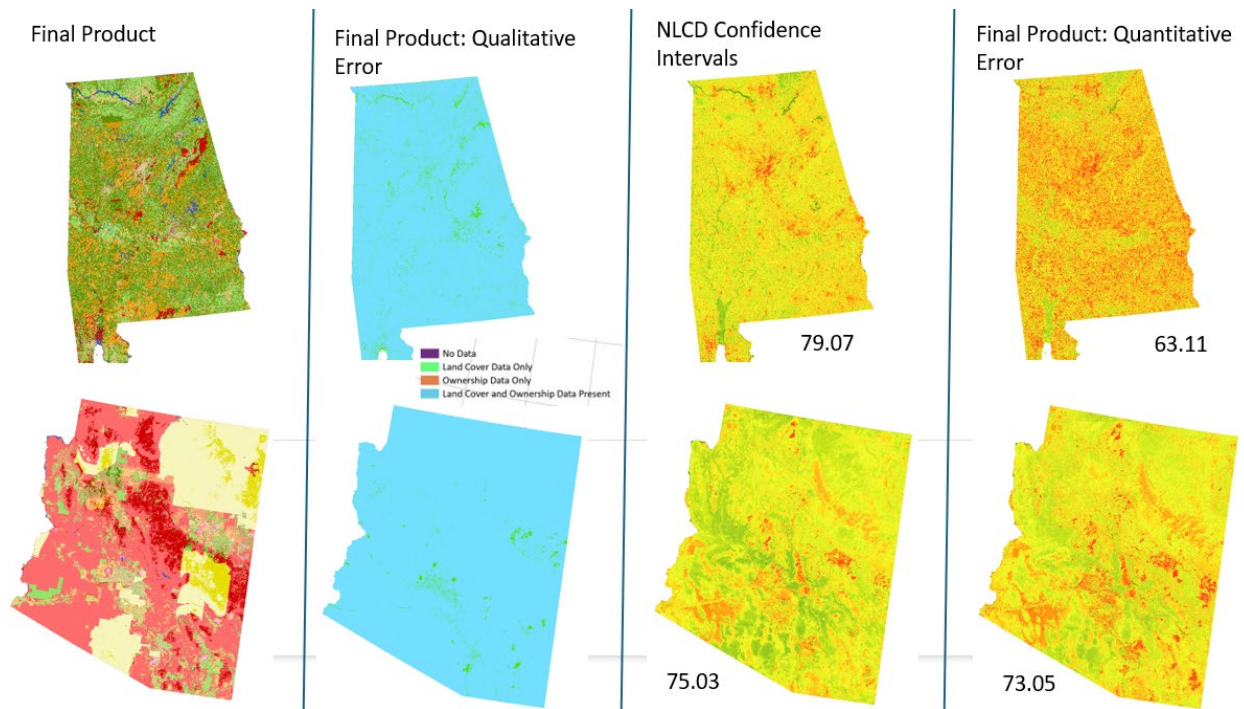


Figure 2: Land use and land ownership layer, along with qualitative and quantitative error products (BETA). Alabama (top) and Arizona (bottom).

Problems or Delays

A no-cost extension has successfully been put in place in order to make up for the delays in the previous two periods.

Next Period Plans

Over the next project period, the team will focus on refining the formulae for integrating the published error estimates for the raw NLCD data, the within-class ownership classification error rates derived from our validation efforts, and the estimated error rates associated with PAD data and third-party vendor data (currently unknown). Finally, we will execute the formulae and produce a raster error encoding our estimate of total error.