Principal Investigators & Affiliations

Muhammed Jaslam Poolakkal, Ph.D. – University of Idaho
Mark Kimsey, Ph.D. – University of Idaho
David L. R. Affleck, Ph.D. – University of Montana
Paul A. Parker, Ph.D. – University of California–Santa Cruz
Nathaniel B. Naumann, M.S. – PotlatchDeltic Corporation

Project Title

Robust small-area estimation strategies for developing accurate stand-level diameter distributions

Reporting Period

January 1, 2025 – June 30, 2025

Progress

Over the first half of 2025, our team has laid the groundwork for a comprehensive small-area estimation framework tailored to stand-level diameter distributions. We successfully consolidated and harmonized inventory data from multiple sources: industrial partners (PotlatchDeltic, Green Diamond, Manulife), and National Forest System stand exams data across regions 1, 4, 6 (PNW & Rockies).

Simultaneously, we developed a semi-automated workflow to preprocess 3D-NAIP point clouds. This pipeline filters noise, classifies ground returns, normalizes heights, and generates a suite of gridded metrics. From these data we extracted over 30 covariates—ranging from canopy height moments and complexity indices to spectral vegetation indices (NDVI, SAVI, VARI, etc.), terrain attributes (slope, aspect, topographic wetness), soil/geology layers (gSSURGO, SGMC), and climate variables from ClimateNA.

Focusing initially on Quadratic Mean Diameter (QMD), we evaluated nine area-level estimators: REML, ML, adjusted ML variants (AMRL/AMPL), robust EBLUPs (with and without bias correction), and a PCA-driven stepwise approach. Both REML and adjusted ML delivered high explanatory power (FH-$R^2 \approx 0.95$) with low mean squared error, while PCA + Stepwise achieved the best AIC (646.4). Robust EBLUP methods demonstrated resilience to outliers, and early diagnostics flagged minor skewness in residuals for further refinement. Validation on Idaho AOIs (St. Joe National Forest, Moscow Mountain) confirms stable predictions across unseen stands.

Next period plans

- Building on this strong foundation, we will:
- Advance from QMD to Full Distributions
- Implement multivariate area-level SAE to jointly estimate distributional moments and key percentiles.
- Fit both parametric (e.g., Weibull, gamma, Johnson's SB) and finite-mixture models to capture entire diameter profiles.
- Integrate Machine Learning
- Develop multi-output ML models—such as random forests and neural networks—within the SAE framework, incorporating hierarchical random effects for stand-to-stand variability.
- Adapt weighted-loss and composite estimation strategies (e.g., random-weight Fay–Herriot) to balance survey design and ML flexibility.

Validation & Simulation

- Cross-validate models on external CFI and FIA datasets, employing bootstrap and jackknife methods for robust MSE and coverage estimates.
- Conduct simulation studies that mirror real survey designs to evaluate bias, variance, and confidence-interval performance.

Feature Refinement & Dissemination

- Apply advanced selection techniques (mRMR, embedded ML) to distill an efficient covariate suite.

Package workflows and documentation for open-source release via our GitHub repository and the PSAE cloud platform.

Problems or delays

Data Variability: Inconsistent availability and quality of 3D-NAIP across states has slowed some preprocessing. We are addressing this by sourcing supplemental LiDAR acquisitions from state surveys and industry collaborators.

Model Residuals: Minor non-normality in residuals under certain estimators suggests the need for response transformations or robust variance modeling; we will explore Y-link transformations and adjusted ML approaches.

External Data Harmonization: Aligning structure and attribute definitions between CFI stand exams and *FIA plots* remains a work in progress. We have established a staggered data delivery schedule and standardized templates to expedite this step.