New Project

The Interplay between Sampling Design and Small Area Estimation to Improve Forestland Inventory

CAFS.24.XX

Temesgen Hailemariam, Oregon State University
Sukhyun, Joo, Oregon State University
Jacob Strunk and Bryce Frank, FIA
Dale Hogg, Green Diamond
Cristian Montes, Rayonier
Nathaniel Neumann, PotlatchDeltic
Ethan Hughes, WA DNR
Aaron Weiskittel, University of Maine
Phil Radtke, Virginia Tech
Rachel Cook, North Carolina State University

Temesgen Hailemariam, Presenter

Justification

- Small sample sizes for subpopulations of interest, defined geographically or by types (e.g., species composition, site productivity classes, or habitats), lead to unreliable estimates when using the sample data alone.

- Despite the small sample sizes, managers and analysts require reliable small-area estimates (SAE) for acquisitions, fiber supply analyses, silvicultural prescriptions, or habitat capability modeling.

- Variable selection is a critical component of SAE models. Many predictor variables are commonly acquired from climate, terrain, and remotely sensed databases. This poses an issue of selecting the optimal set of predictor variables to be included in a model.

- In acquisitions and fiber supply analyses, sampling protocols and sample size allocation affect the performance of SAE models at the stand and ownership levels.

Objectives

1) Examine variable selection methods for developing small-area estimation models that link inventory plots and remotely sensed data for timberland inventory.

2) Based on different stand characteristics and silviculture treatments, examine the performance of selected sampling designs and sample sizes for applying SAE models.

3) Examine the use of small-area estimators to either reduce sample size when precision is given or improve precision when the sample size is fixed; and

4) Explore methods to allocate sample size to subpopulation, including optimal allocation of samples in small domains.

Methods

Data: We will first use FIA data collected in Klamath and Jackson counties of western Oregon and Carbon Project Data collected by Green Diamond Resources in Klamath and then expand to other key regions in CAFS. The data include:

1. Individual inventory plot, stand and polygon-level summaries
2. Sentinel data, including mean reflectance values from selected Sentinel-2 bands and other derived variables

Variable Selection:

- Conduct a review of the literature and identify variable selection/reduction methods for estimating stand volume and site index from climate, terrain, and remotely sensed data.
- In addition to variable selection algorithms, we will select predictors based on the correlation of covariates (X) and response variable (Y) or some other metric that ranks predictors (e.g., Z-scores).
- Examine different number ad combinations of variables and reduce the number of variables before selecting a final set of variables.

Methods

Using a verification data and Monte Carlo Simulations, we will evaluate the performances of 12-16 sample selection strategies, including selecting stands proportional to volume or value, in estimating merchantable volume.

Performance of sample selection and sample allocation strategies will be compared using:

1) Bias: bias will be calculated as the mean difference between observed and predicted total merchantable volume for each stand.

2) Relative Bias: - Relative bias percentage is the ratio of bias to the total observed merchantable volume for each stand

3) Root Mean Square Error (RMSE)

4) Relative RMSE

Deliverables

Short-term (1-year)

- Literature review and background information completed
- Initial variable selection methods and comparisons completed on pilot project
- A poster on the project progress

Long-term

- Protocols that aid in linking remotely sensed data and ground data/attributes (e.g., thinning status) to reduce uncertainty and improve the quality of small-area estimates for timberland inventory
- A PhD dissertation on the interplay
- Publications in peer-reviewed literature
- Presentations at several regional professional meetings
- Summary report for member companies.

Company Benefits

- Greater understanding in borrowing strength from linking freely available remotely sensed and ground data to improve the estimation of selected stand variables and to reduce the costs of data acquisition or establishing ground plots.

- Identify protocols to operationalize small area estimation methods by linking remotely sensed and ground data and developing cost-effective variable selection algorithms solutions for improved stand-and ownership-level estimation

- The project demonstrates ways for incorporating SAE models into operational forest inventory in plantations.

- Help to translate some of the theories in small-area estimation to practice.

Summary

- The project will identify protocols to operationalize small area estimation methods by linking ground-based samples with 3D-NAIP, Sentinel, and LiDAR and developing cost-effective variable selection algorithms solutions for improved stand-and ownership-level estimation

- Additionally, the project will investigate selected sampling designs and strategies to allocate samples to small domains to reduce uncertainty in predicting selected variables.

- Help forest analysts and managers identify efficient protocols to implement a hierarchical SAE approach with FIA plots and stand-exam data

- The results will help inform decisions on allocating samples to various domains or small areas