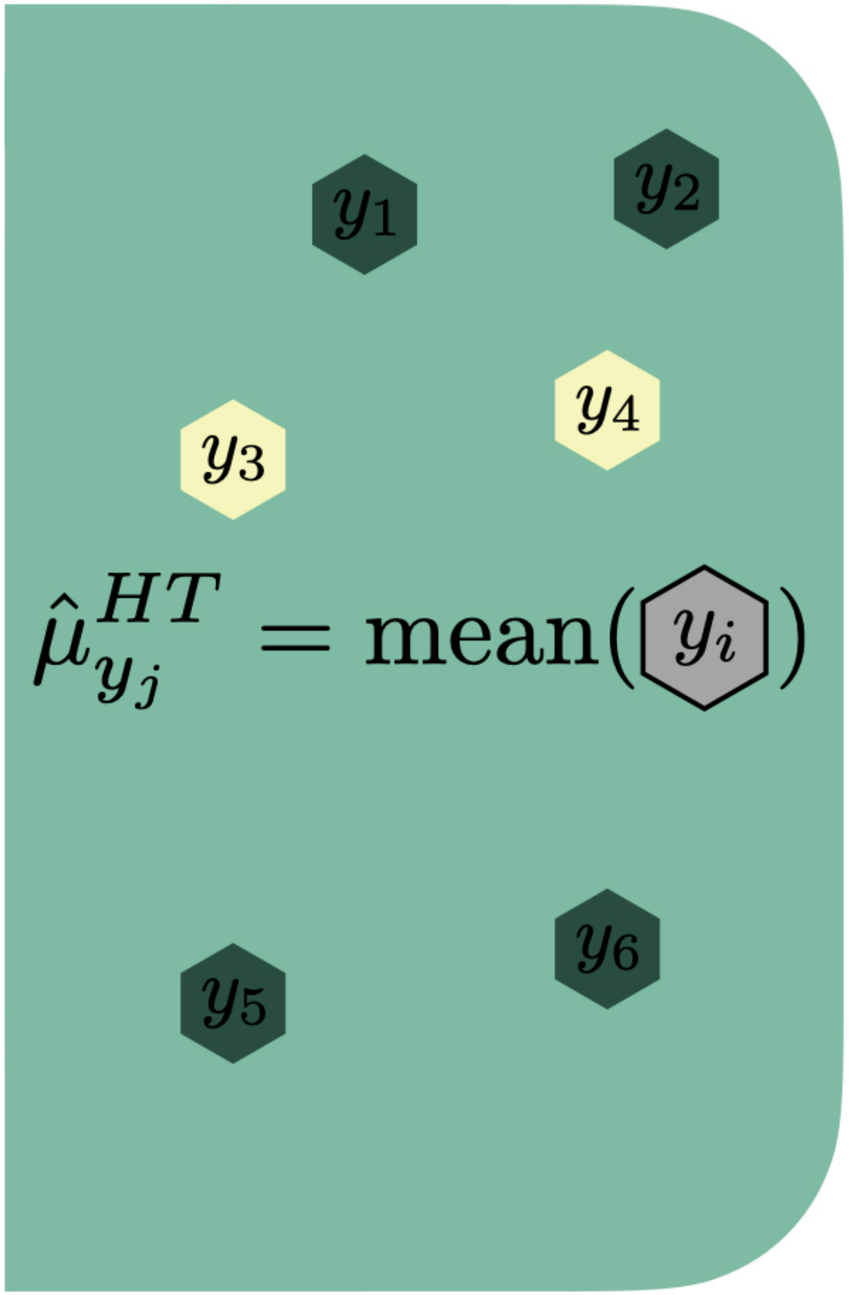
The Sample Mean

Or, "The Horvitz-Thompson Estimator"

Estimate: $$\hat{\mu}_{y_j}^{HT} = \frac{1}{n_j} \sum_{i \in s_j} y_i$$

Variance: $$\hat{V}\left(\hat{\mu}_{y_j}^{HT}\right) = \frac{1}{n_j - 1} \sum_{i \in s_j} \left(y_{ji} - \hat{\mu}_{y_j}^{HT}\right)^2$$



The Survey Regression Estimator

A "Direct" Estimator

Estimate: $$\hat{\mu}_{y_j}^{SR} = \frac{1}{N_j} \sum_{i \in U_j} \hat{y}_{ji} + \frac{1}{n_j} \sum_{i \in s_j} (y_{ji} - \hat{y}_{ji})$$

The average of the fitted model in the area j of interest

The average of the residuals over the sampled plots in the area j of interest

The Regression-Synthetic Estimators

Implicit Model-Based Indirect Estimators

Area-level Estimator

$$\hat{\mu}_{y_j}^{RSA} = \hat{\beta}_0 + \hat{\beta}_1 \bar{X}_j$$

Fit from the eco-province

Unit-level Estimator

$$\hat{\mu}_{y_j}^{RSU} = \frac{1}{N_j} \sum_{i \in U_j} \hat{y}_{ji}$$

where $\hat{y}_{ji} = \hat{\beta}_0 + \hat{\beta}_1 X_{ji}$

Fit from the eco-province

A Hierarchical Bayesian Approach to Small Area Estimation of Forest Attributes

Thesis Orals Presentation

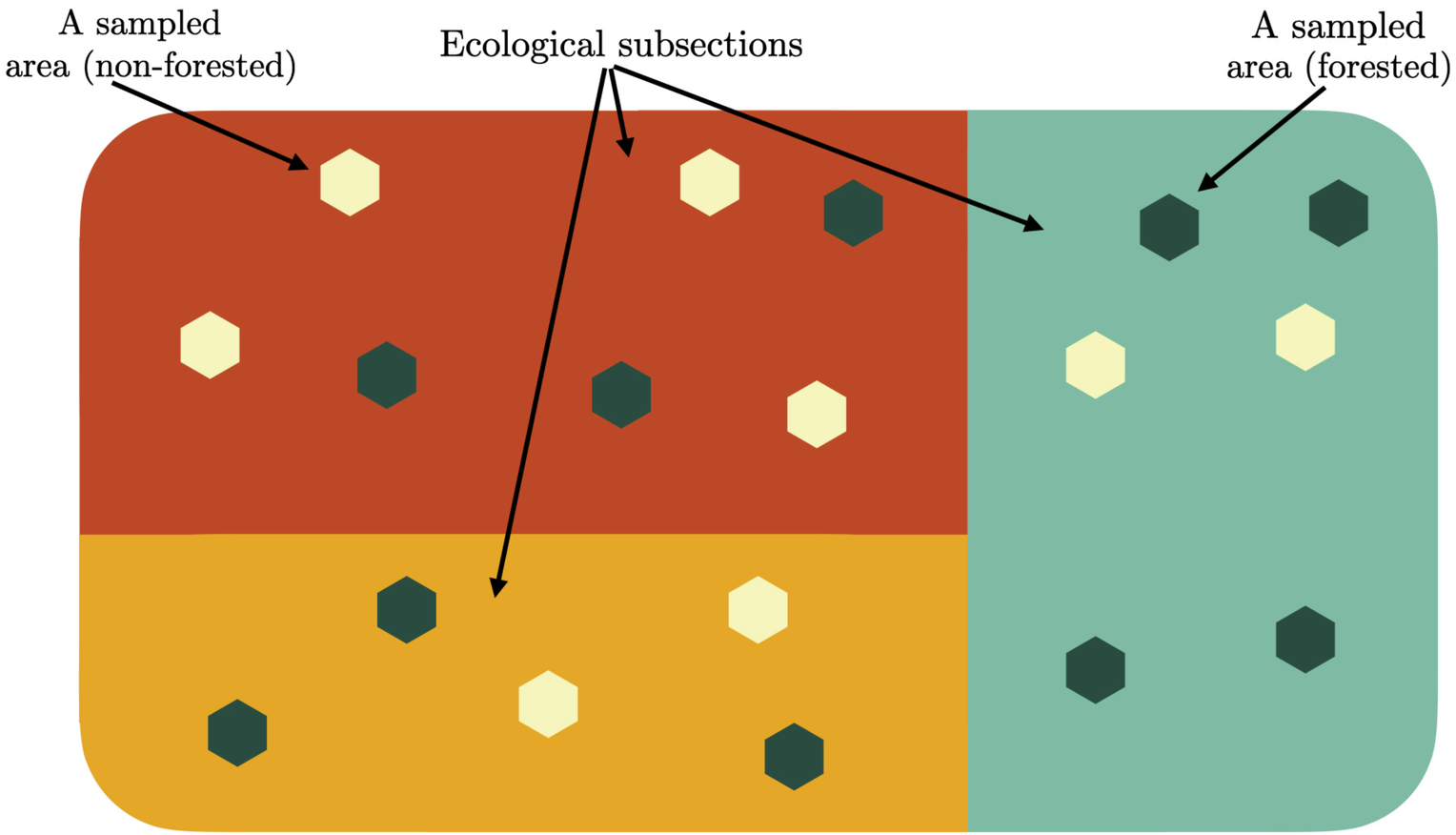
Grayson White

May 6, 2020

The Data

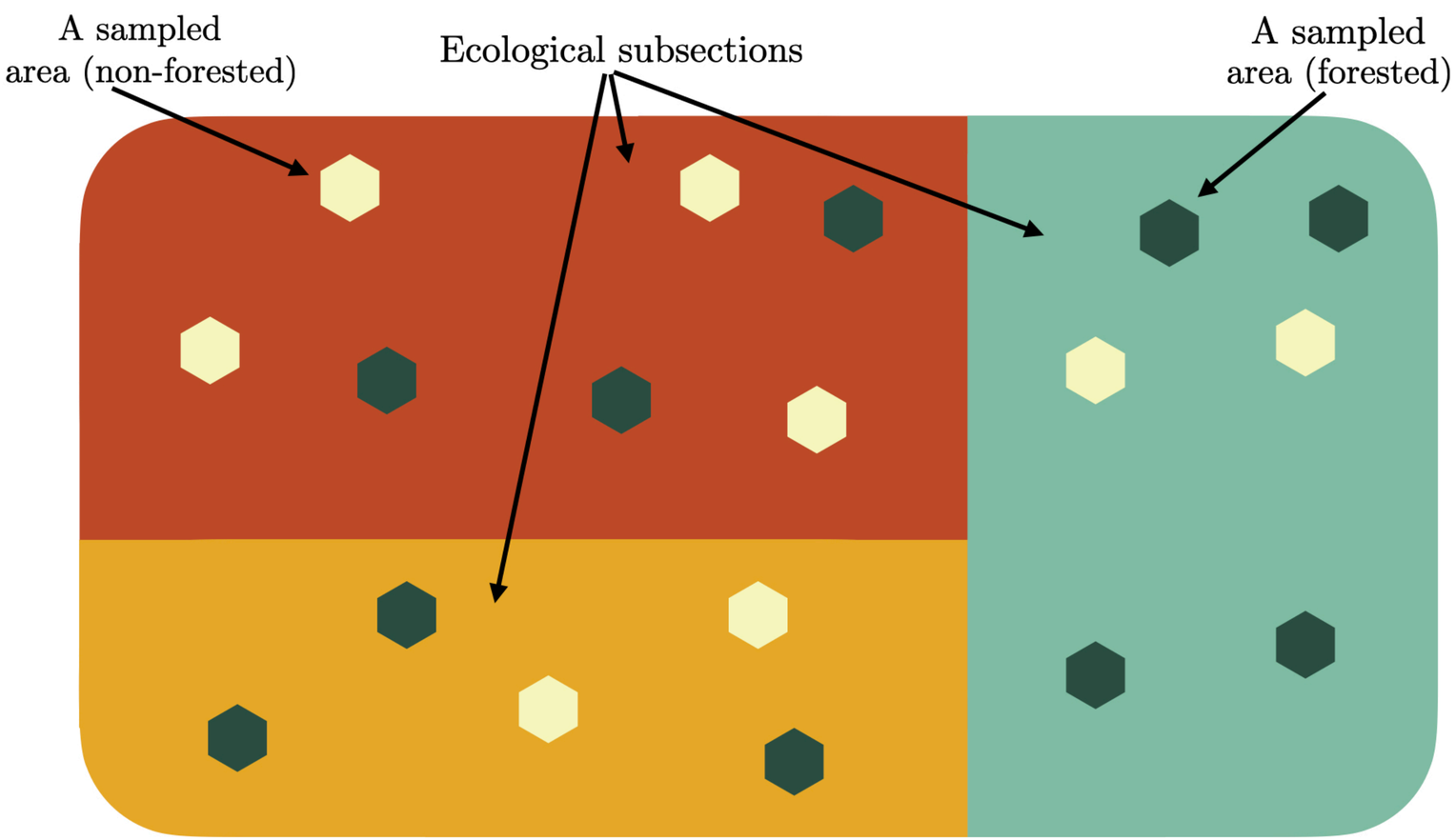
Orcas Island, Washington

The Data: A Brief Overview



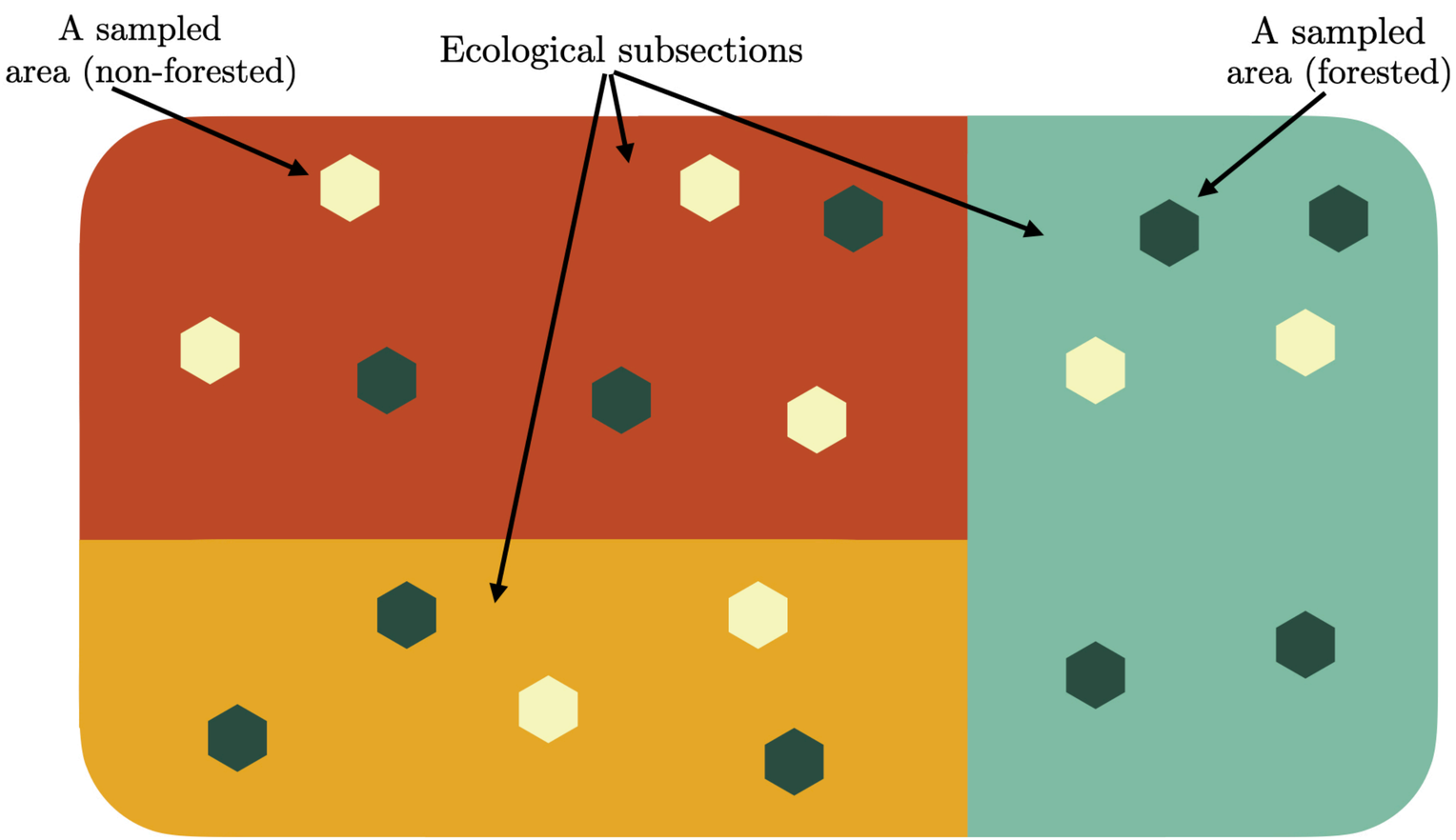
The Data: A Brief Overview

- Ground-collected, plot-level data for four response variables: biomass, tree count per acre, basal area, and volume.



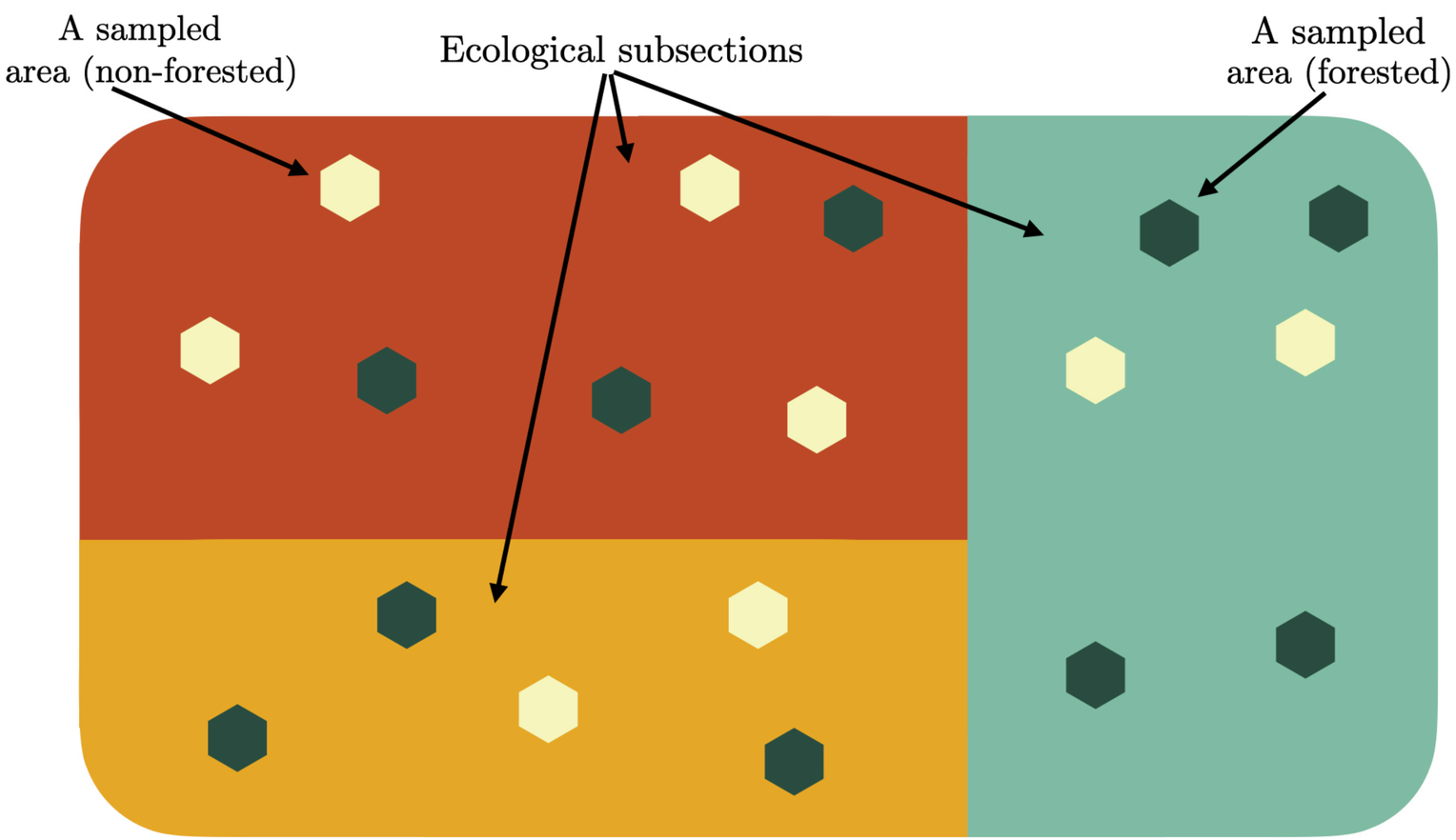
The Data: A Brief Overview

- Ground-collected, plot-level data for four response variables: biomass, tree count per acre, basal area, and volume.
- Remotely-sensed explanatory variable: total canopy cover.



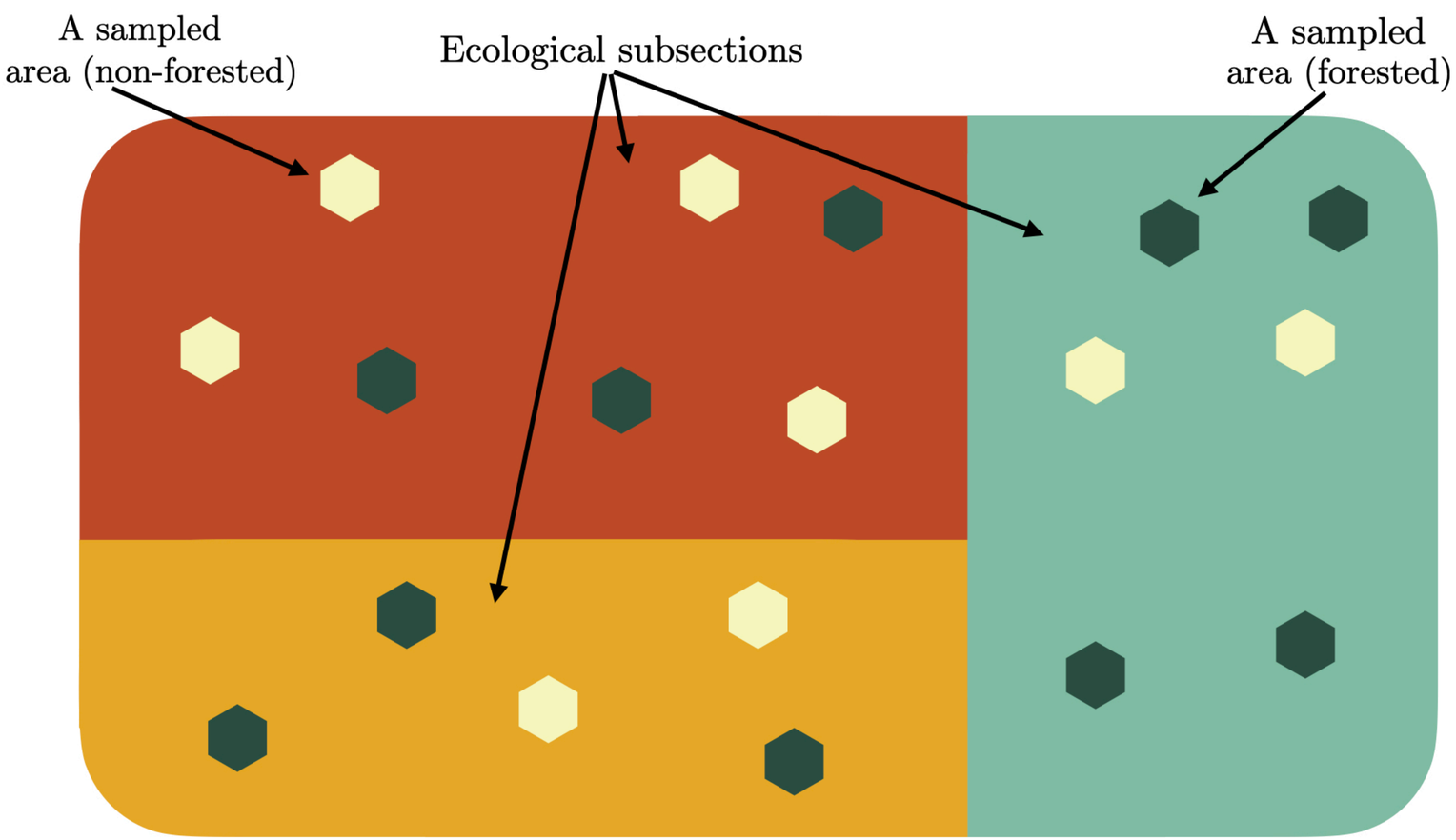
The Data: A Brief Overview

- Ground-collected, plot-level data for four response variables: biomass, tree count per acre, basal area, and volume.
- Remotely-sensed explanatory variable: total canopy cover.
- Forest/non-forest strata indicator variable for each sampled plot.



The Data: A Brief Overview

- Ground-collected, plot-level data for four response variables: biomass, tree count per acre, basal area, and volume.
- Remotely-sensed explanatory variable: total canopy cover.
- Forest/non-forest strata indicator variable for each sampled plot.
- Proportion of area that is forested for each eco-subsection.



Methods and Intuition

Washington Highway 410, edge of Mt. Rainier National Park

The Post-Stratified Estimator

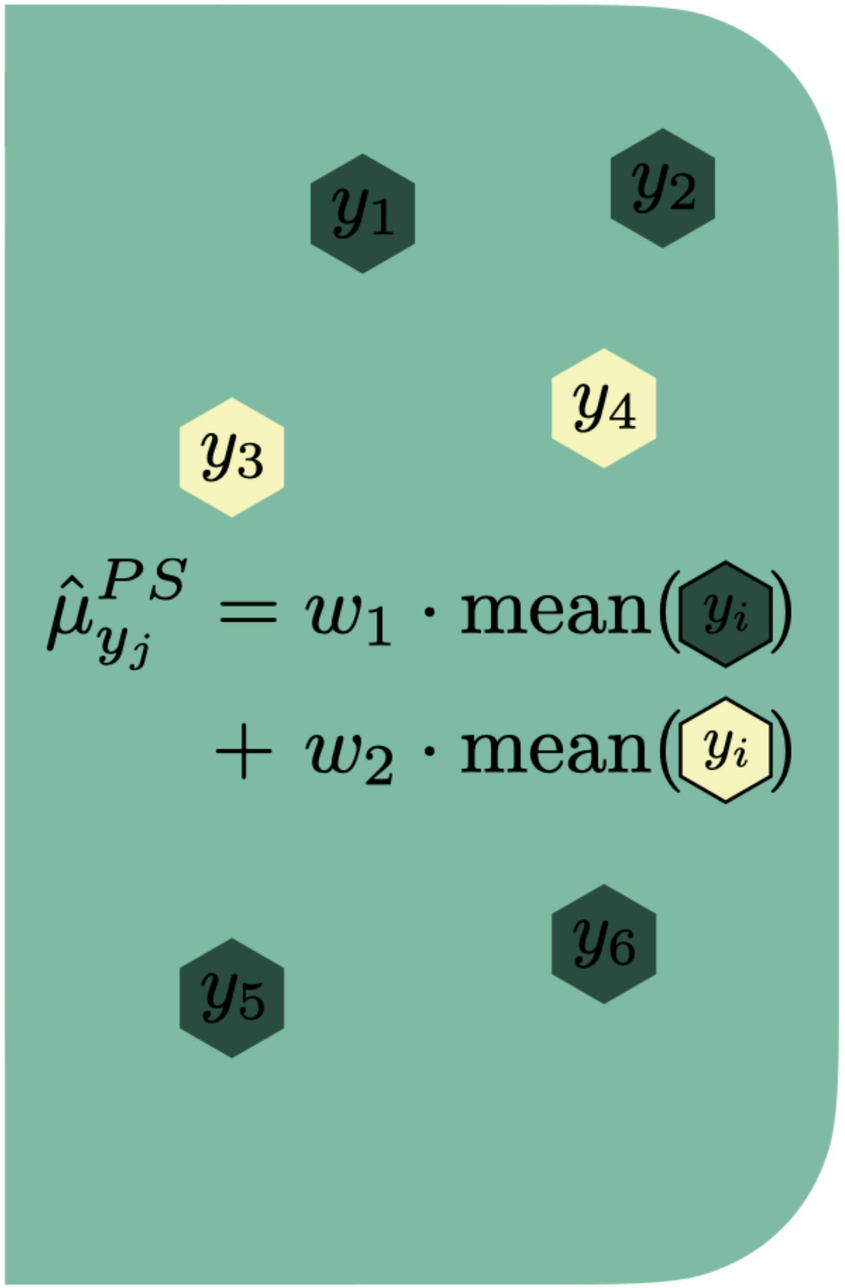
A Direct Estimator

Estimate:

$$\hat{\mu}_{y_j}^{PS} = \sum_{k=1}^{2} w_{jk} \cdot \hat{\mu}_{y_{jk}}^{HT}$$

Variance:

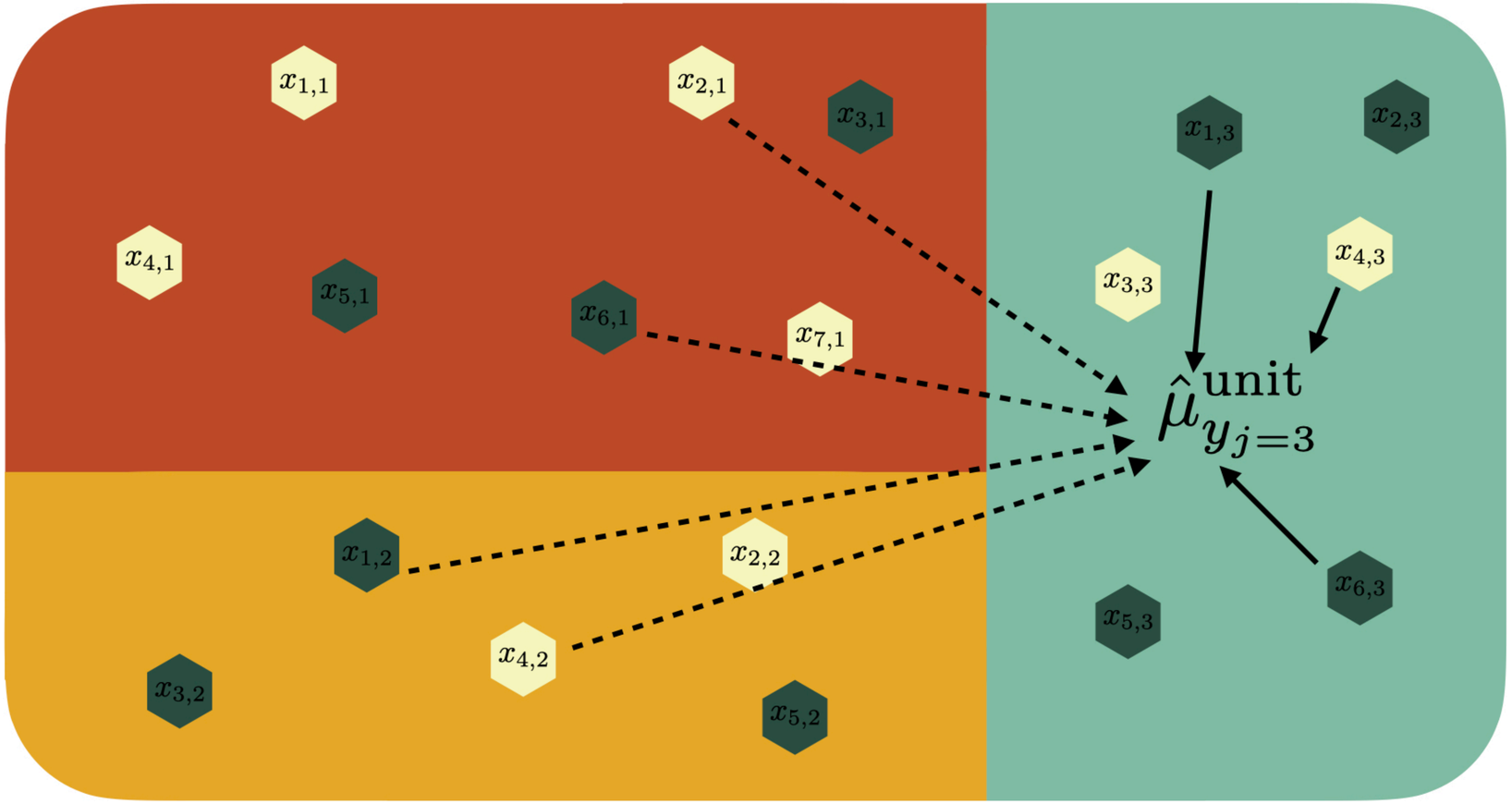
$$\hat{V}\left(\hat{\mu}_{y_j}^{PS}\right) = \frac{1}{n_j}\left(\sum_{k=1}^{2} w_{jk} n_{jk} \hat{V}\left(\hat{\mu}_{y_{jk}}^{HT}\right) + \sum_{k=1}^{2}(1 - w_{jk})\frac{n_{jk}}{n_j}\hat{V}\left(\hat{\mu}_{y_{jk}}^{HT}\right)\right)$$



The Unit-Level Hierarchical Model

Model Form and Visualization

$$y_{ji} = x_{ji}\beta + \nu_j + e_{ji}$$



The Unit-Level Mixed Model

Specification and Fit: Frequentist EBLUP

Specification:

$$y_{ji} = x_{ji}\beta + \nu_j + e_{ji}$$

Fit:

$$\hat{\mu}_{y_j}^{FRU} = \hat{\gamma}_j \hat{\mu}_{y_j}^{SR} + (1 - \hat{\gamma}_j)\hat{\mu}_{y_j}^{RSU}$$

$$\text{where} \quad \hat{\gamma}_j = \frac{\hat{\sigma}_\nu^2}{\hat{\sigma}_\nu^2 + \hat{\sigma}_e^2}$$

The Unit-Level Hierarchical Model

Specification and Fit: Bayesian

Specification:

The hierarchical model:

$$y_{ji} \mid \beta,\ \nu_j, \sigma_e^2 \sim \mathrm{N}(x_{ji}\beta + \nu_j,\ \sigma_e^2),$$

$$\nu_j \mid \sigma_\nu^2 \sim \mathrm{N}(0,\ \sigma_\nu^2),$$

$$f(\beta,\ \sigma_\nu^2,\ \sigma_e^2) = f(\beta) \cdot f(\sigma_\nu^2) \cdot f(\sigma_e^2),$$

with priors:

$$f(\beta) \propto 1,$$

$$\frac{f(\sigma_e^2)}{f(\sigma_\nu^2)} = \mathrm{IG}(\mathrm{df} = 0,\ \mathrm{scale} = 0).$$

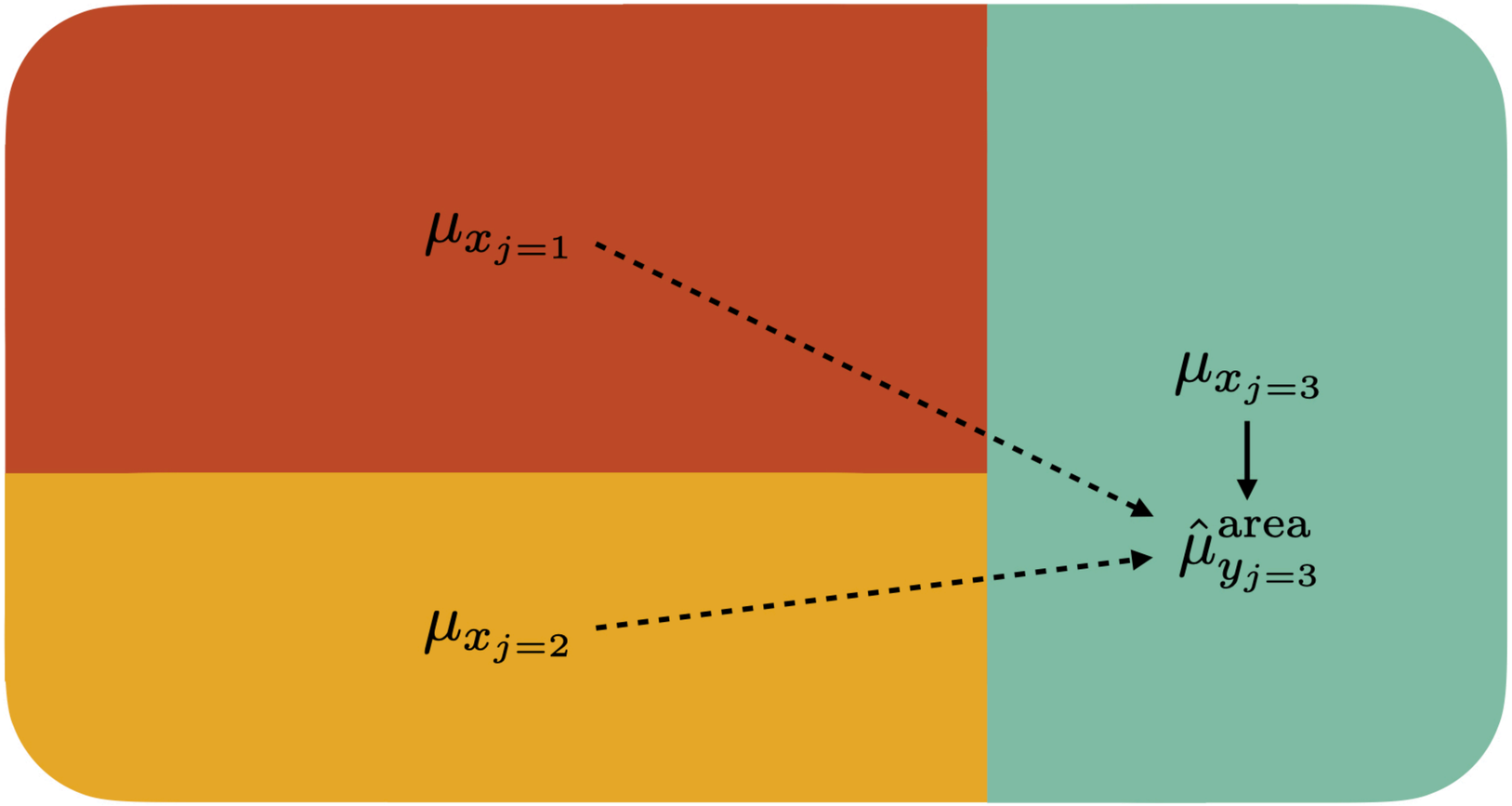
Fit:

$$\hat{\mu}_{y_j}^{HBU} = E[\mu_{y_j} \mid \mathbf{y_j}]$$

The Area-Level Hierarchical Model

Model Form and Visualization

$$\hat{\mu}_j^{PS} = x_j\beta + \nu_j + e_j$$



The Area-Level Mixed Model

Specification and Fit: Frequentist EBLUP

Specification:

$$\hat{\mu}_j^{PS} = x_j\beta + \nu_j + e_j$$

Fit:

$$\hat{\mu}_j^{FRA} = \hat{\gamma}_j\hat{\mu}_j^{PS} + (1-\hat{\gamma}_j)\hat{\mu}_j^{RSA}$$

$$\text{where} \quad \hat{\gamma}_j = \frac{\hat{\sigma}_\nu^2}{\sigma_j^2 + \hat{\sigma}_\nu^2}$$

The Area-Level Hierarchical Model

Specification and Fit: Bayesian

Specification:

The hierarchical model:

$$\hat{\mu}_{y_j}^{PS} \mid \mu_{y_j},\ \beta,\ \sigma_\nu^2 \sim \mathrm{N}(\mu_{y_j},\ \sigma_j^2),$$

$$\mu_{y_j} \mid \beta,\ \sigma_\nu^2 \sim \mathrm{N}(x_j\beta,\ \sigma_\nu^2),$$

$$f(\beta,\ \sigma_\nu^2) = f(\beta) \cdot f(\sigma_\nu^2),$$

With priors:

$$f(\beta) \propto 1,$$

$$f(\sigma_\nu^2) = \mathrm{IG}(\mathrm{df} = 10000 \cdot m,\ \mathrm{scale} = 1)$$

Fit:

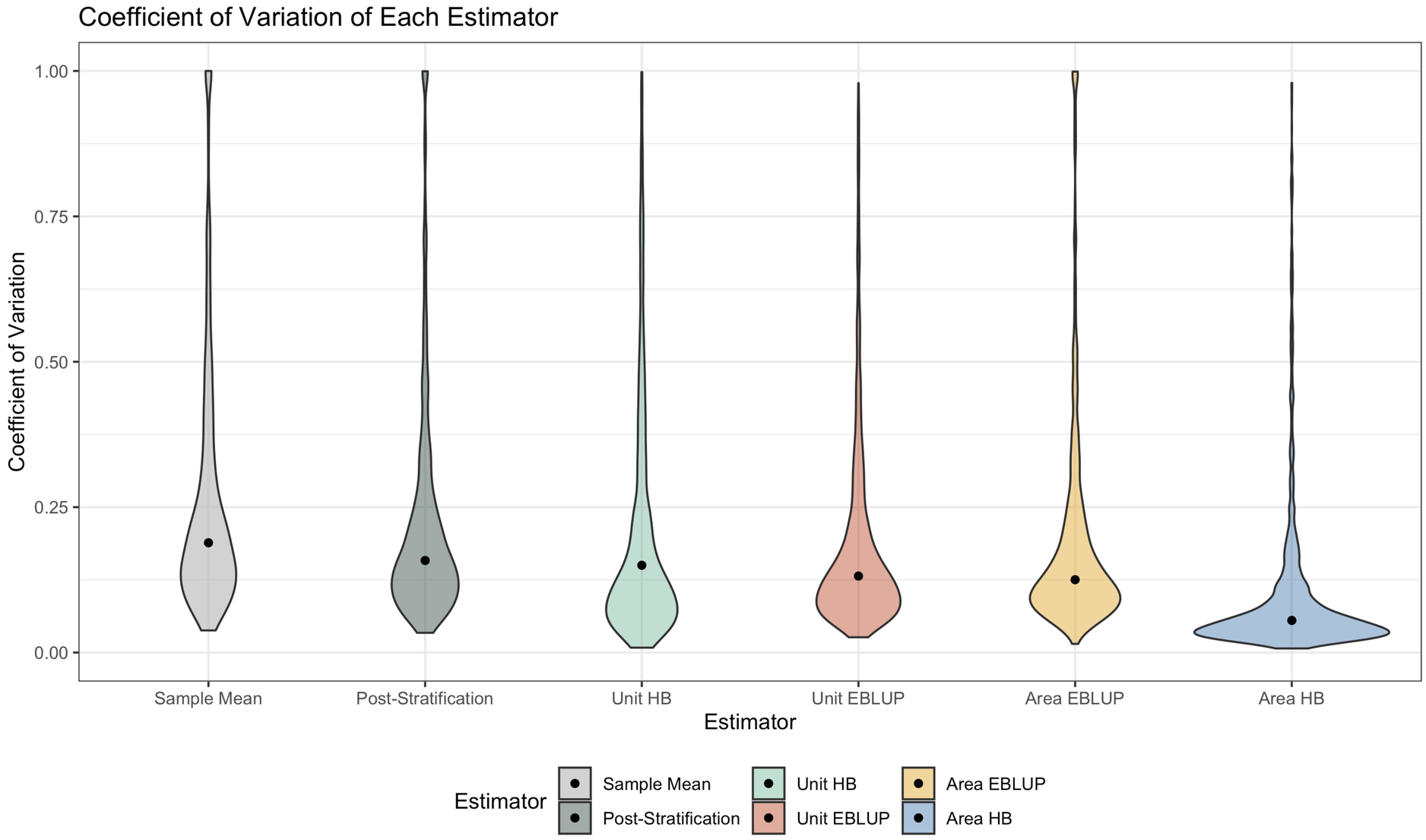
$$\hat{\mu}_{y_j}^{HBA} = E[\mu_{y_j} \mid \hat{\mu}_{y_j}^{PS}]$$

Results and Extensions

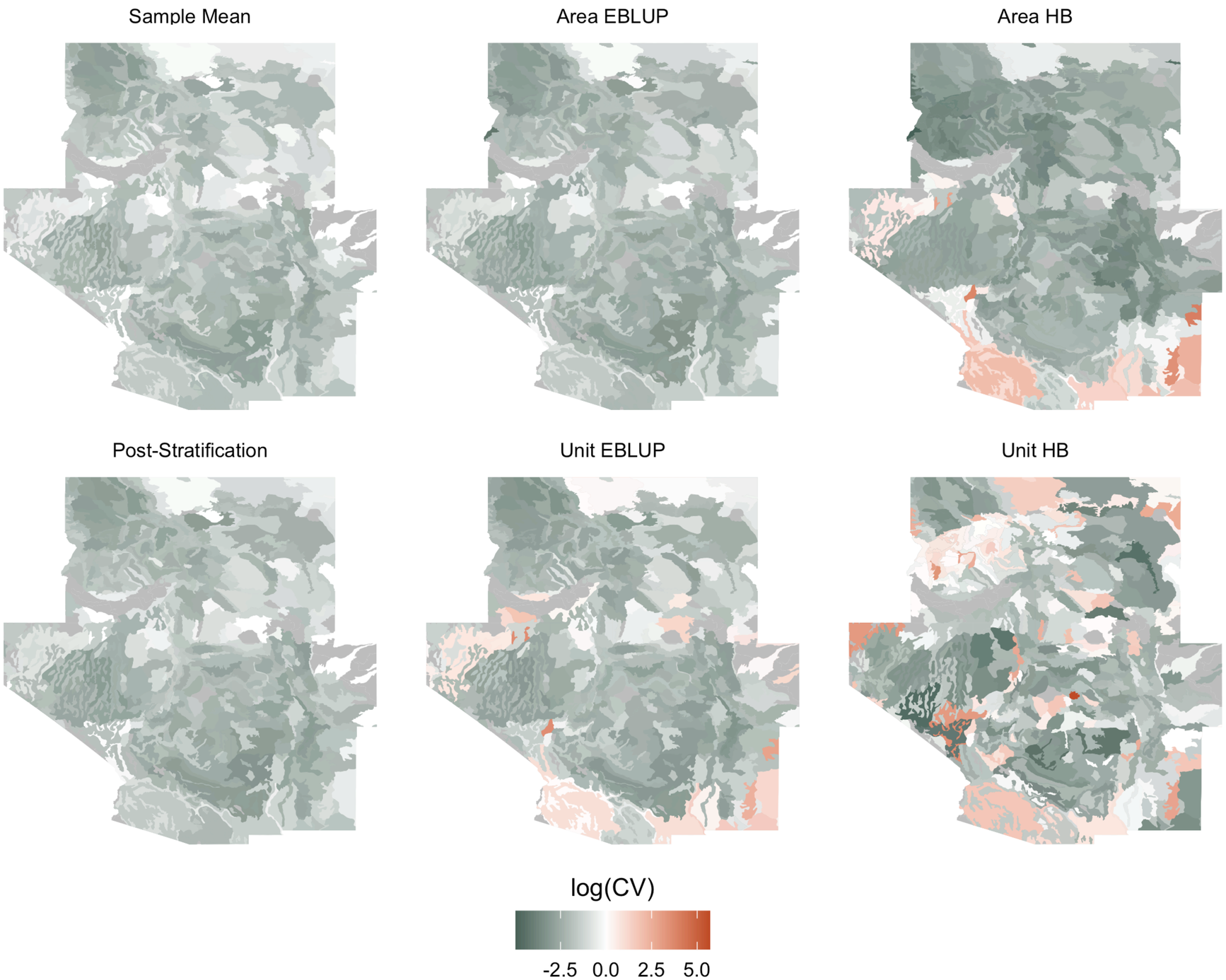
Valley of the Giants, Oregon

Cape Flattery, the most northwest part of the continental United States

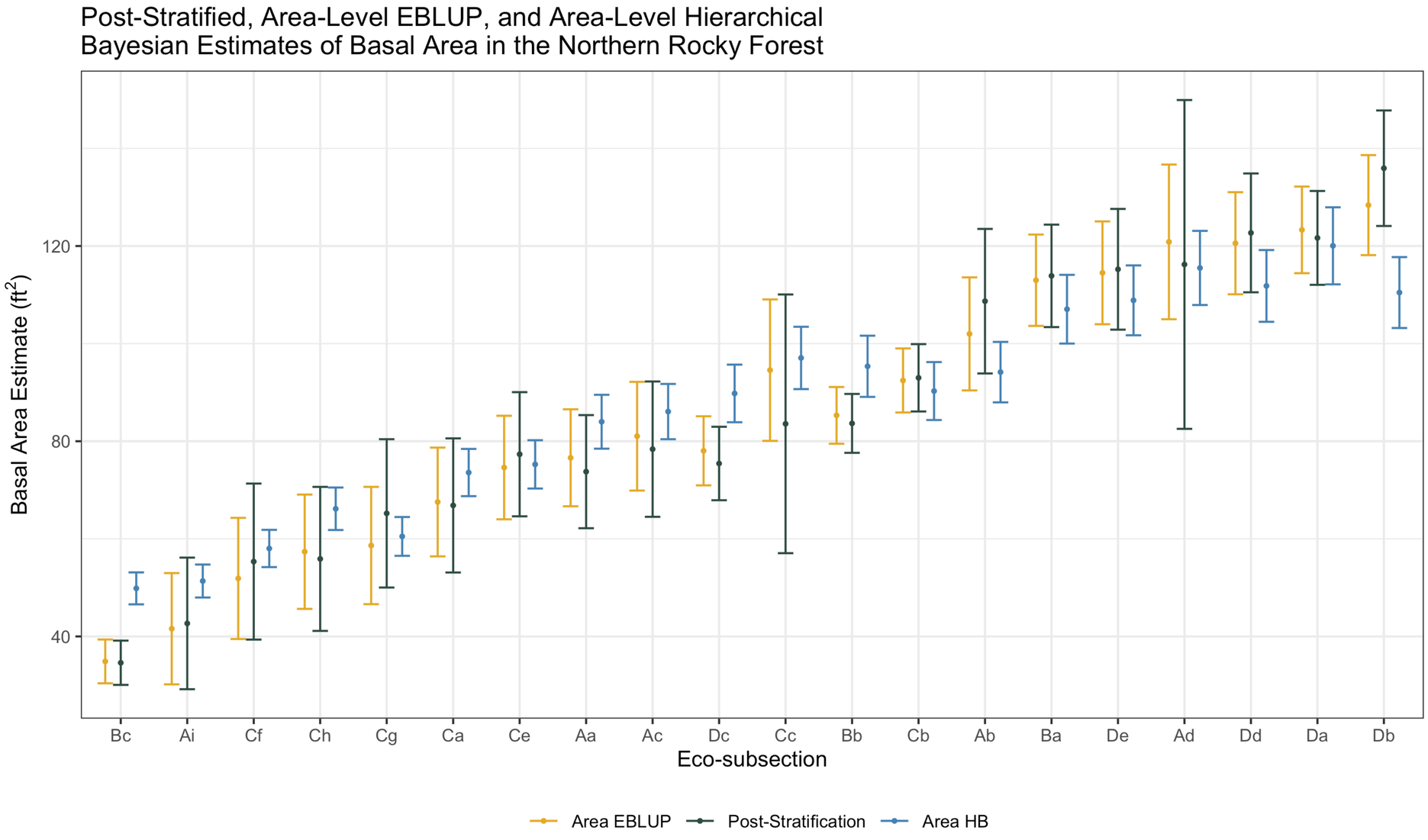
Estimator Variance: Overall



Estimator Variance: Across the Interior West

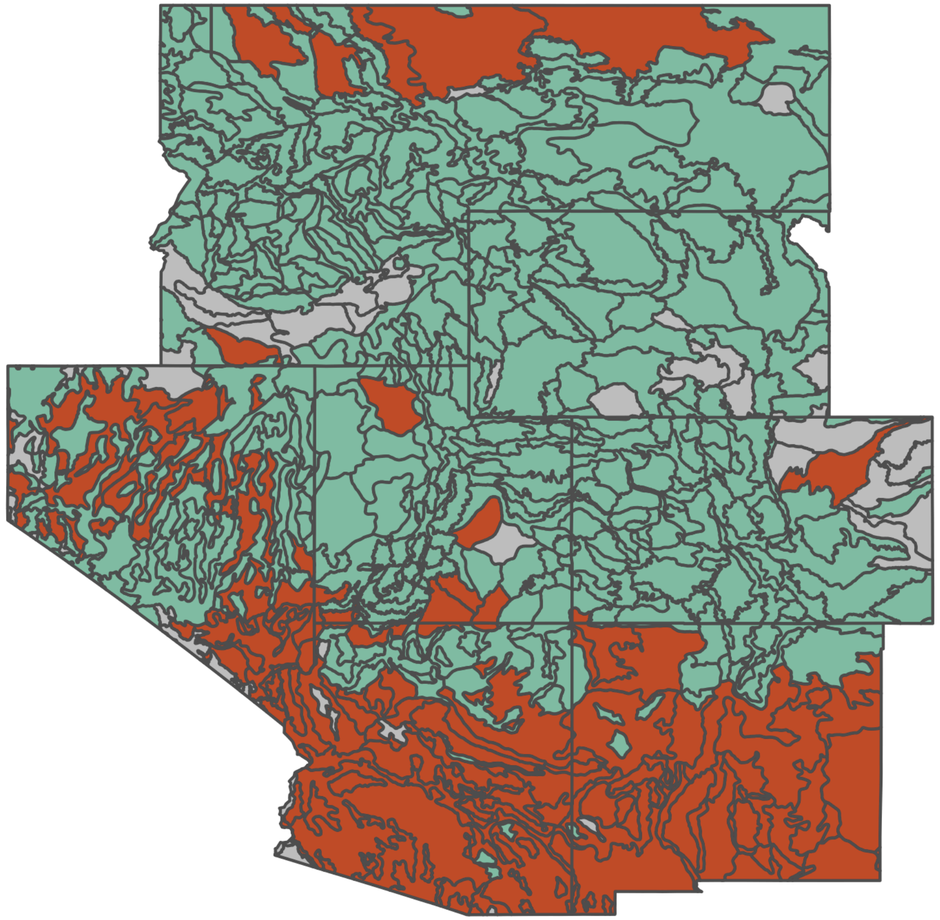


Estimates of Basal Area in the Northern Rocky Forest

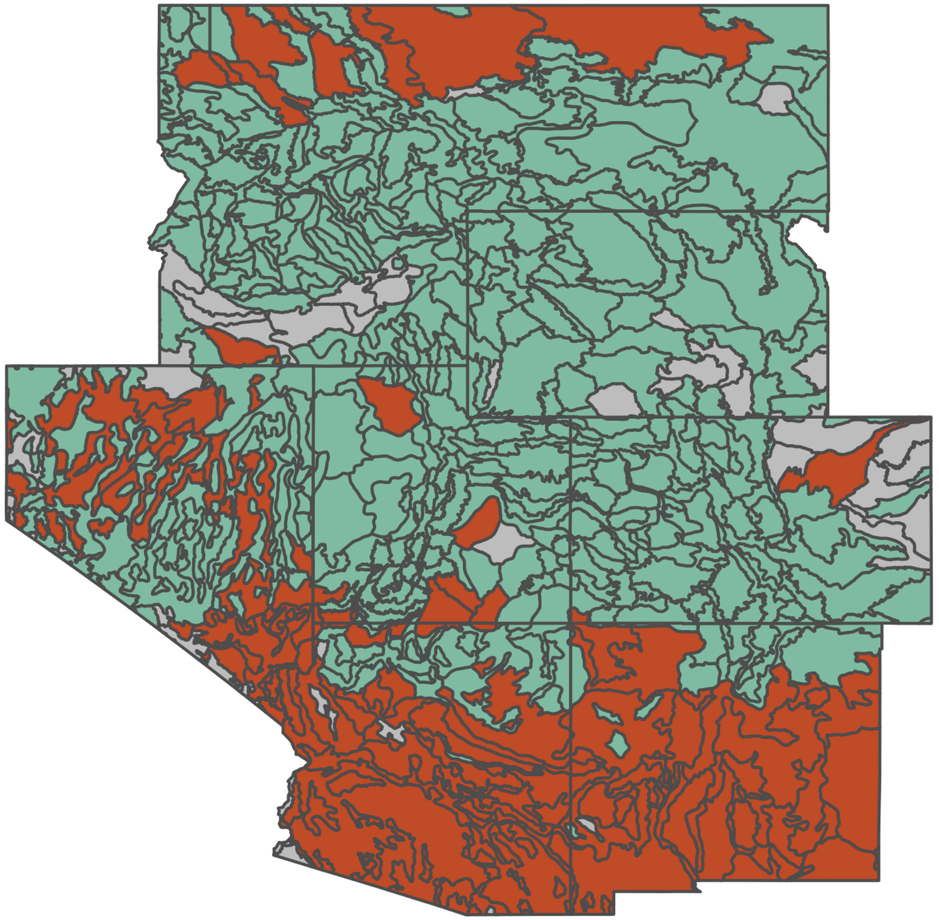


Estimator Variance Comparison

Area HB CV < Post-Strat CV



Area HB CV < Area EBLUP CV



FALSE TRUE NA

Possible Extensions and Further Work

Mt. Hood, from Bull of the Woods lookout tower

Possible Extensions and Further Work

- An R package using MCMC methods to fit these estimators through rstanarm wrapper functions.

Mt. Hood, from Bull of the Woods lookout tower

Possible Extensions and Further Work

- An R package using MCMC methods to fit these estimators through rstanarm wrapper functions.
- Applying these methods across the rest of the United States.

Mt. Hood, from Bull of the Woods lookout tower

Possible Extensions and Further Work

- An R package using MCMC methods to fit these estimators through rstanarm wrapper functions.
- Applying these methods across the rest of the United States.
- Test different response variable distributions at the unit-level.

Mt. Hood, from Bull of the Woods lookout tower

Possible Extensions and Further Work

- An R package using MCMC methods to fit these estimators through rstanarm wrapper functions.
- Applying these methods across the rest of the United States.
- Test different response variable distributions at the unit-level.
- A simulation study to understand bias.

Mt. Hood, from Bull of the Woods lookout tower