

Hierarchical Bayesian Small Area Estimation Using Weakly Informative Priors in the Interior Western US

Grayson White¹, Kelly McConville²,
Gretchen Moisen³, Tracey Frescino³

¹ RedCastle Resources Inc., ² Reed College, ³ USDA Forest Service

The Forest Inventory and Analysis Program (FIA)



Orcas Island, Washington

FIA Overview

- Responsible for reporting status and trends of the nation's forests.
- Collects inventory data on and computes estimates for many forest attributes to monitor the status and trends of the nation's forests.
- Sampling design is meant for sufficient precision to provide state-level estimates through post-stratified estimation.



The Nation's Forest Census

FIA's Sampling Design

- Overlay hexagons with area of 6000 acres across the United States. Randomly select a location in each hexagon for a FIA plot.
- A FIA crew visits each plot once every 10 years to remeasure the plot.
- A 10 year timeframe where each plot has been measured is referred to as an “evaluation”.

United States
Department of
Agriculture

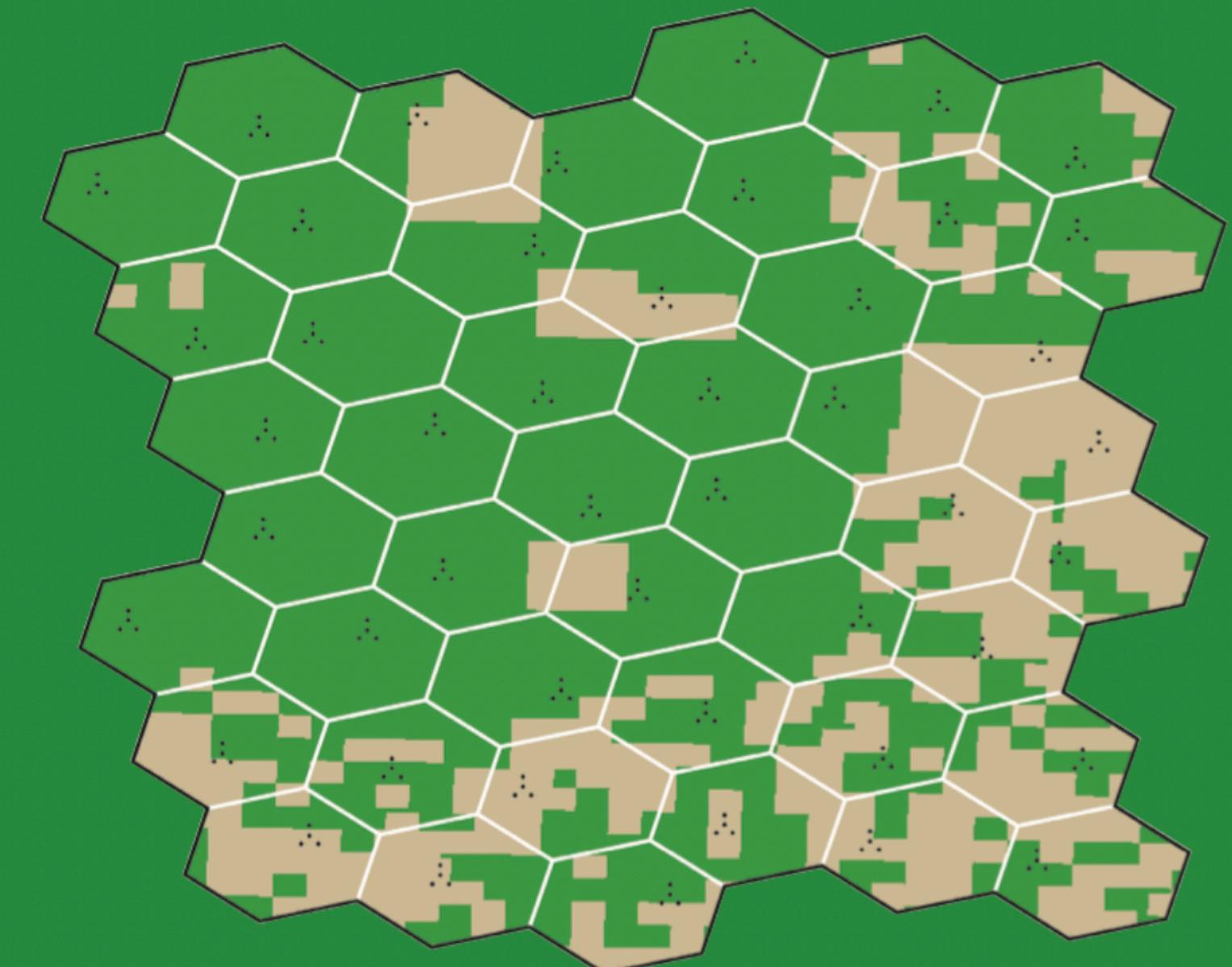
Forest Service


Southern
Research Station

General Technical
Report SRS-80

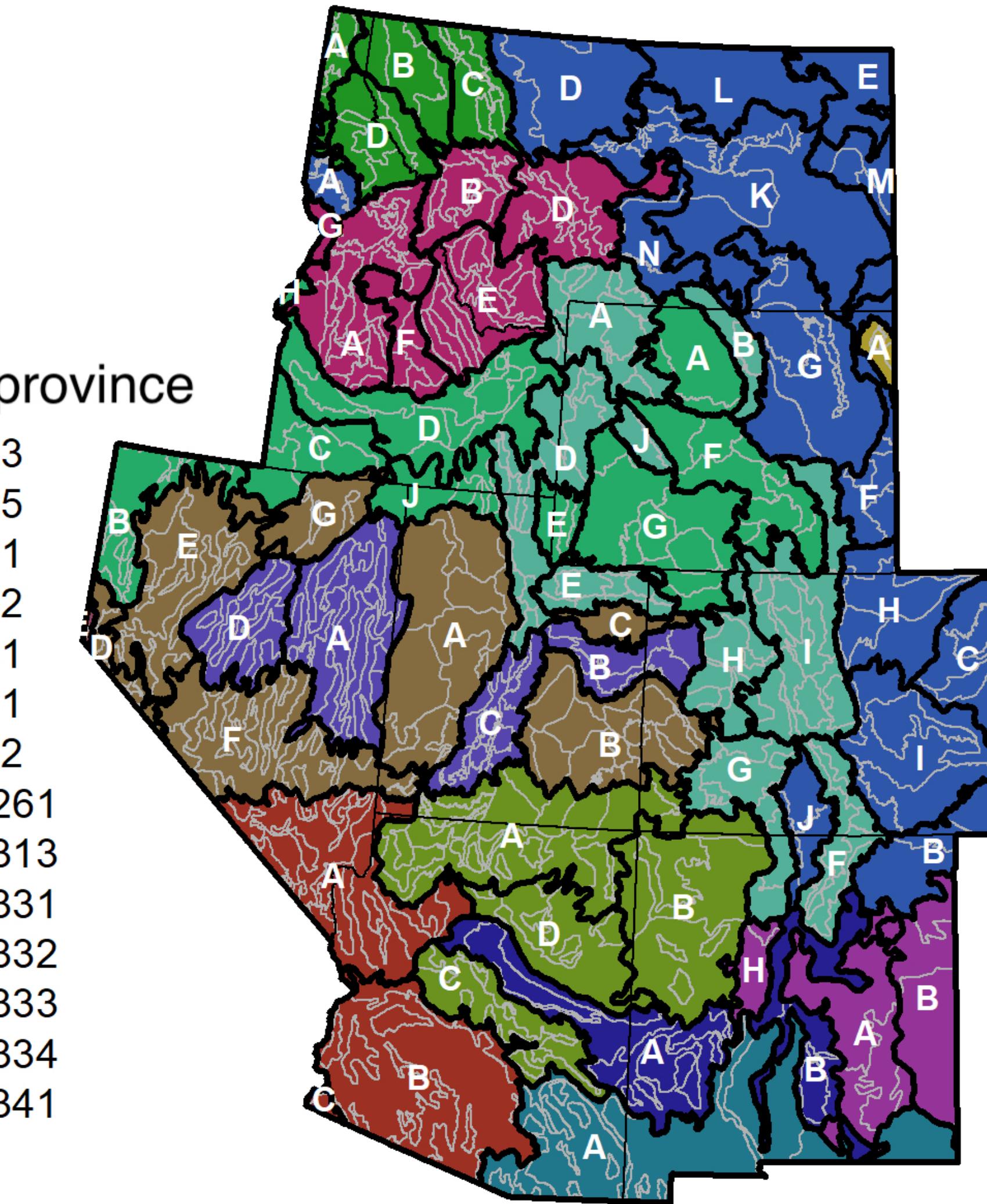
The Enhanced Forest
Inventory and Analysis
Program—National
Sampling Design and
Estimation Procedures

William A. Bechtold and
Paul L. Patterson, Editors



FIA's Small Area Estimation Needs

- FIA has a growing need for statistically defensible estimates over smaller geographic regions such as:
 - Counties
 - Watersheds
 - Ecologically defined regions
 - Recently burned areas
 - And more!



The Data



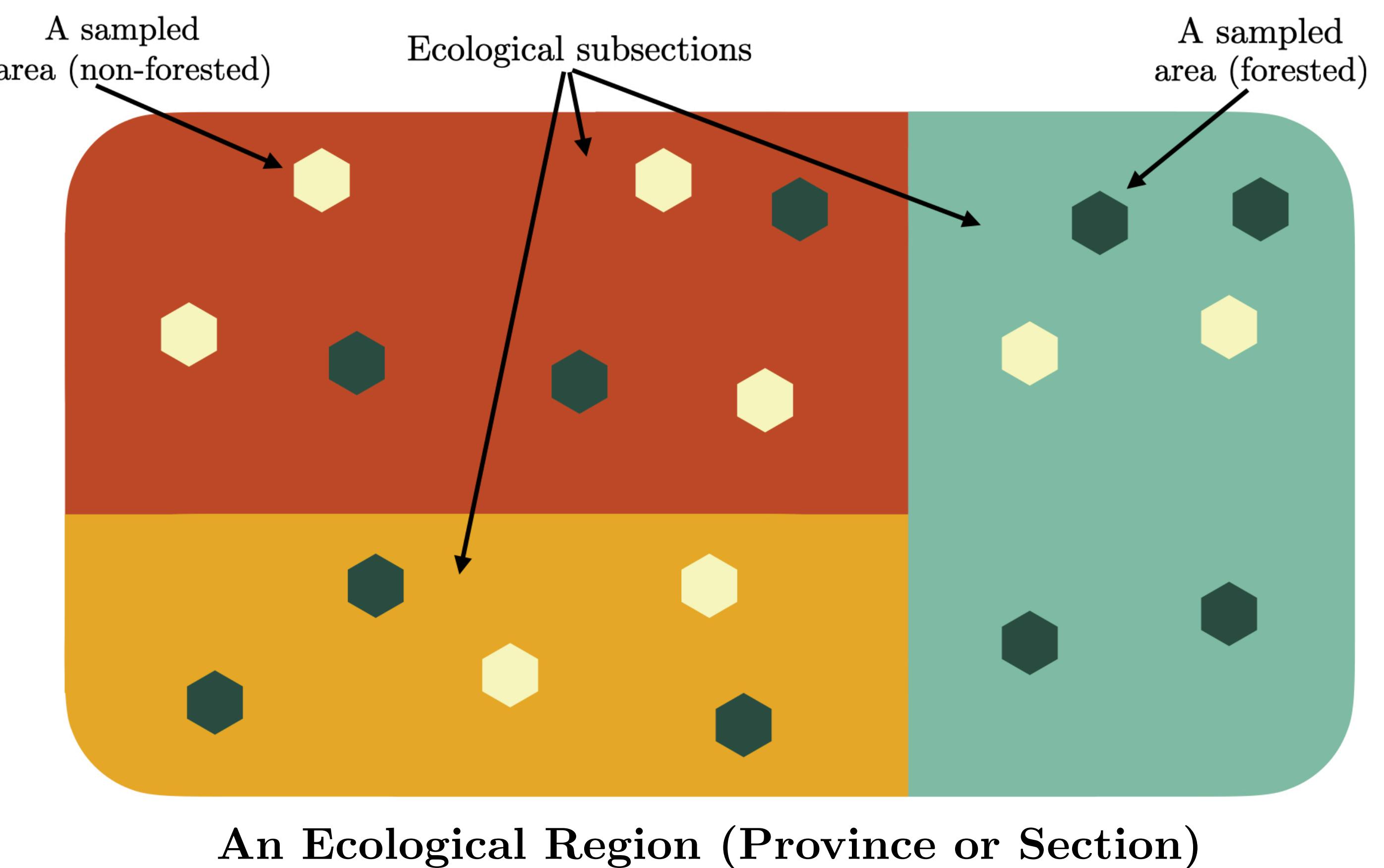
Washington Highway 410, edge of Mt. Rainier National Park

FIA Data Overview



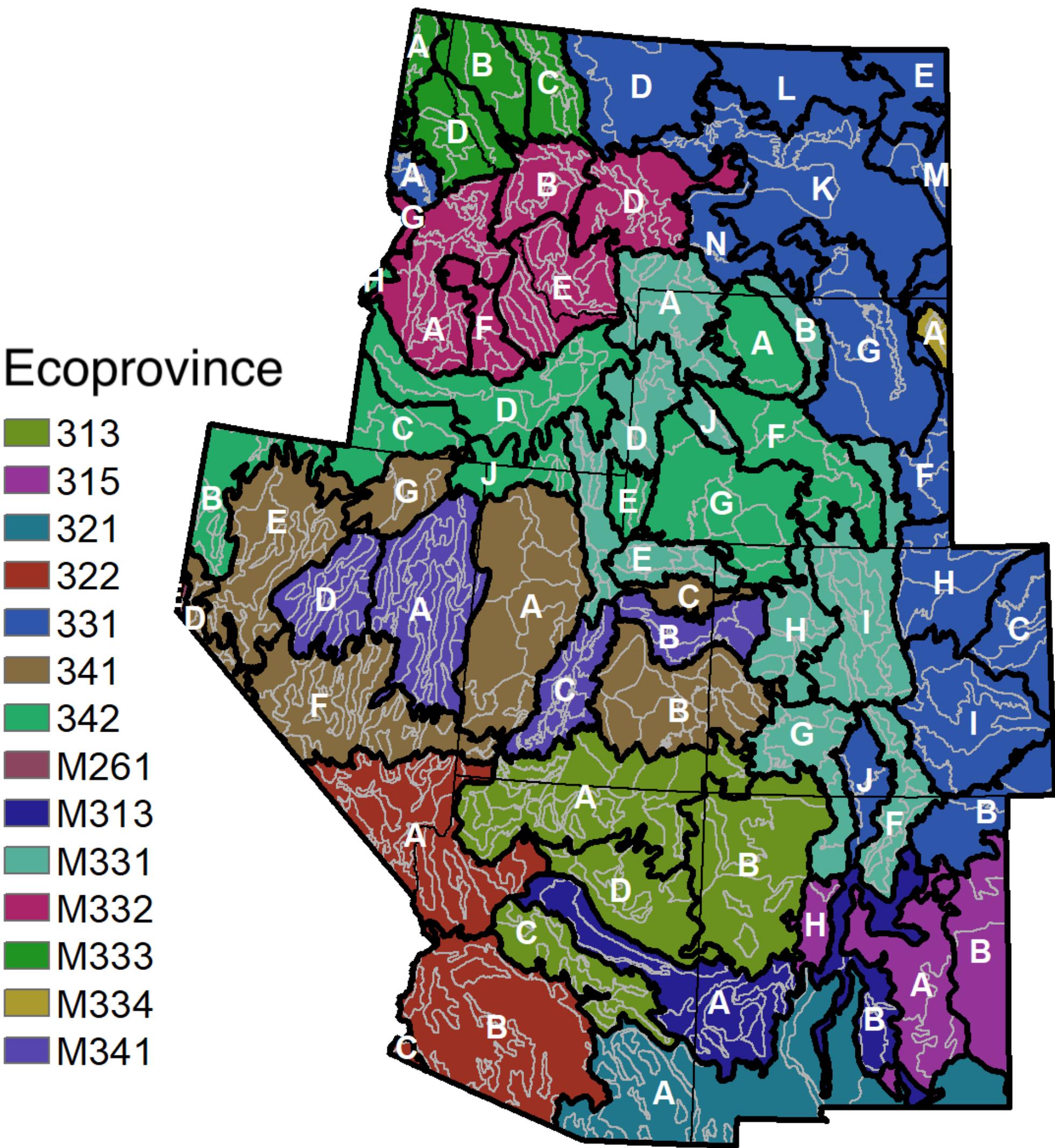
- We used FIESTA to retrieve a FIA evaluation of data from the Interior Western region of the United States from 2007 to 2017, which included 86,065 plots.

- FIESTA is FIA's custom estimation tool which allows for easy access to FIA data and many estimators, all in R.



Auxiliary Data Overview

- Remotely-sensed explanatory variable at pixel level: tree canopy cover.
- Forest/non-forest strata indicator variable for each sampled plot.
- Ecologically-defined regions with three levels of hierarchically (largest to smallest): provinces, sections, subsections.



Methods



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

The Big Idea

- Combine FIA data and auxiliary data to fit and compare multiple area-level small area models to the industry-standard post-stratified estimator for four FIA variables of interest.
- Predict means of these FIA variables (basal area, biomass, trees/acre, and net volume) of interest in ecological subsections.
- Understand the benefits of fitting hierarchical Bayesian small area models in this context: how can we leverage prior information in this specific scenario? Is it helpful to do so? When? Why?
- Compare the small area models with each other, and with the post-stratified estimator.

The Post-Stratified Estimator

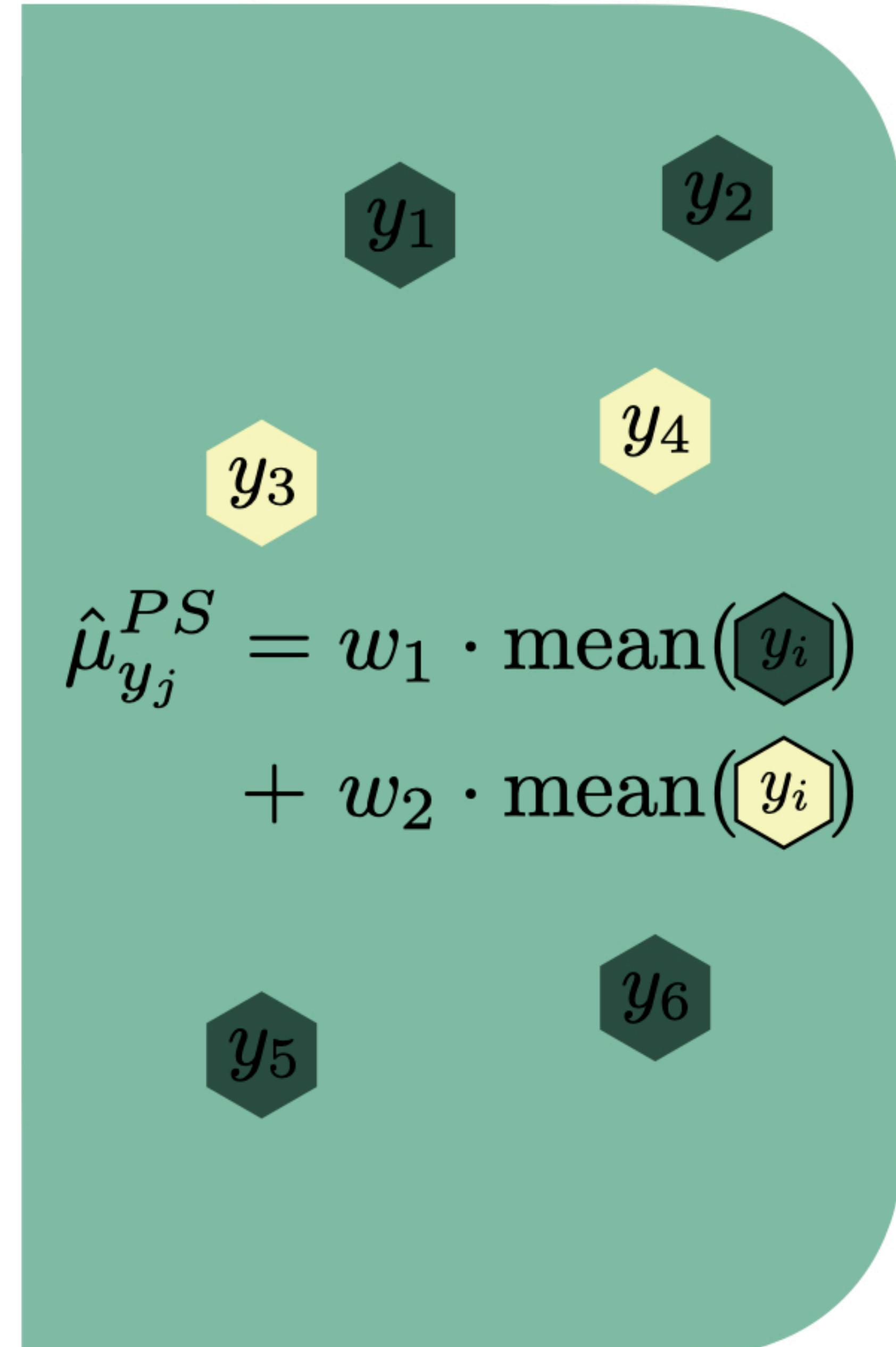
A Direct Estimator

Estimate:

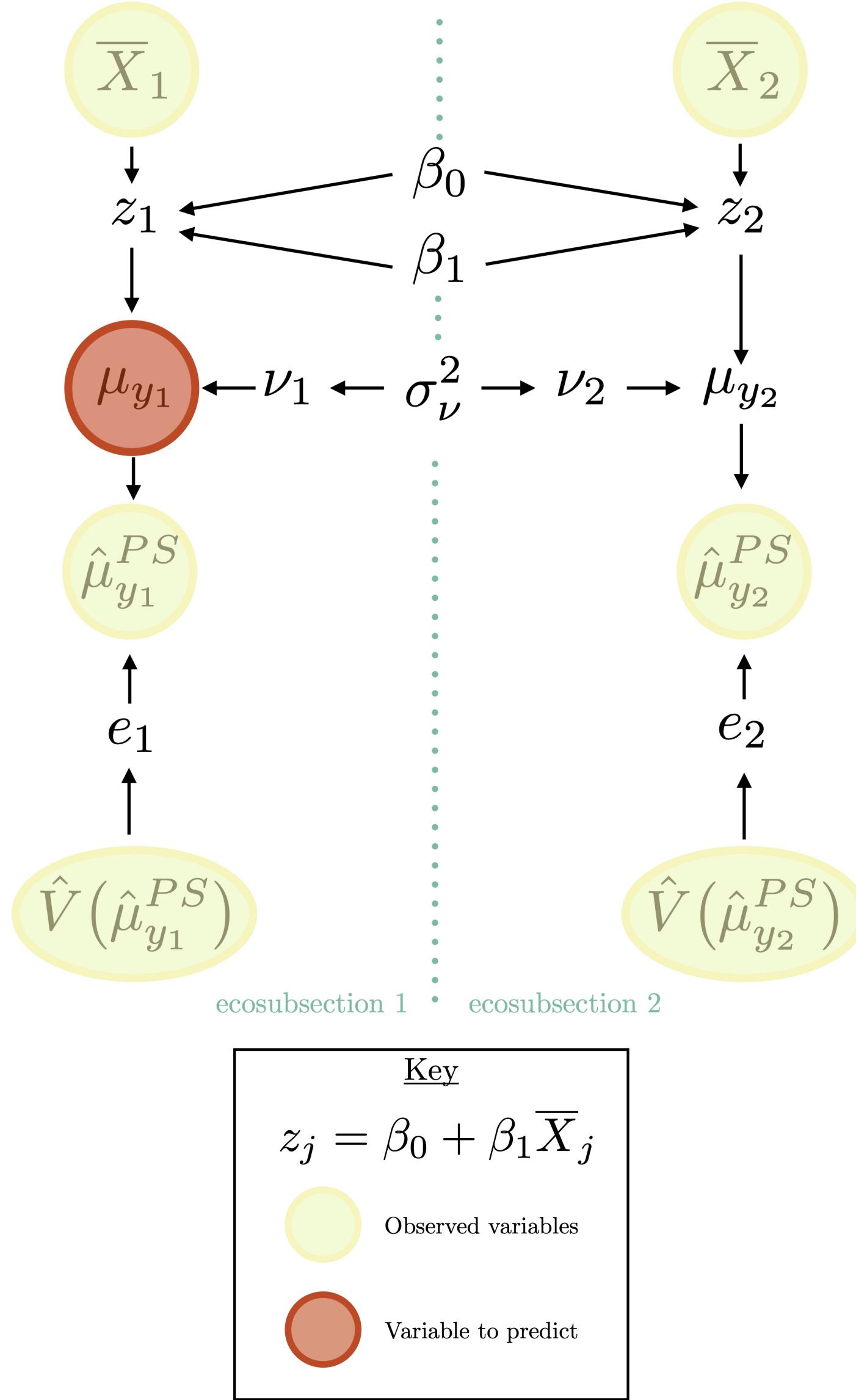
$$\hat{\mu}_{y_j}^{PS} = \sum_{k=1}^2 w_{jk} \cdot \hat{\mu}_{y_{jk}}^{HT}$$

Variance:

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



Our Fay-Herriot Models



- Six variations of the Fay-Herriot model, all which can be represented by this Probabilistic Graphical Model (PGM).

Short Name	Fitting Method	Borrow Strength to...	Prior on σ_ν^2
EBLUP Section	REML	Section	NA
EBLUP Province	REML	Province	NA
HB Section Flat Prior	MCMC	Section	$f(\sigma_\nu^2) \propto 1$
HB Province Flat Prior	MCMC	Province	$f(\sigma_\nu^2) \propto 1$
HB Section half Cauchy Prior	MCMC	Section	$\sigma_\nu \sim \text{half-Cauchy}(\text{scale} = 1)$
HB Province half Cauchy Prior	MCMC	Province	$\sigma_\nu \sim \text{half-Cauchy}(\text{scale} = 1)$

Results

Note: All results shown are for the basal area response variable

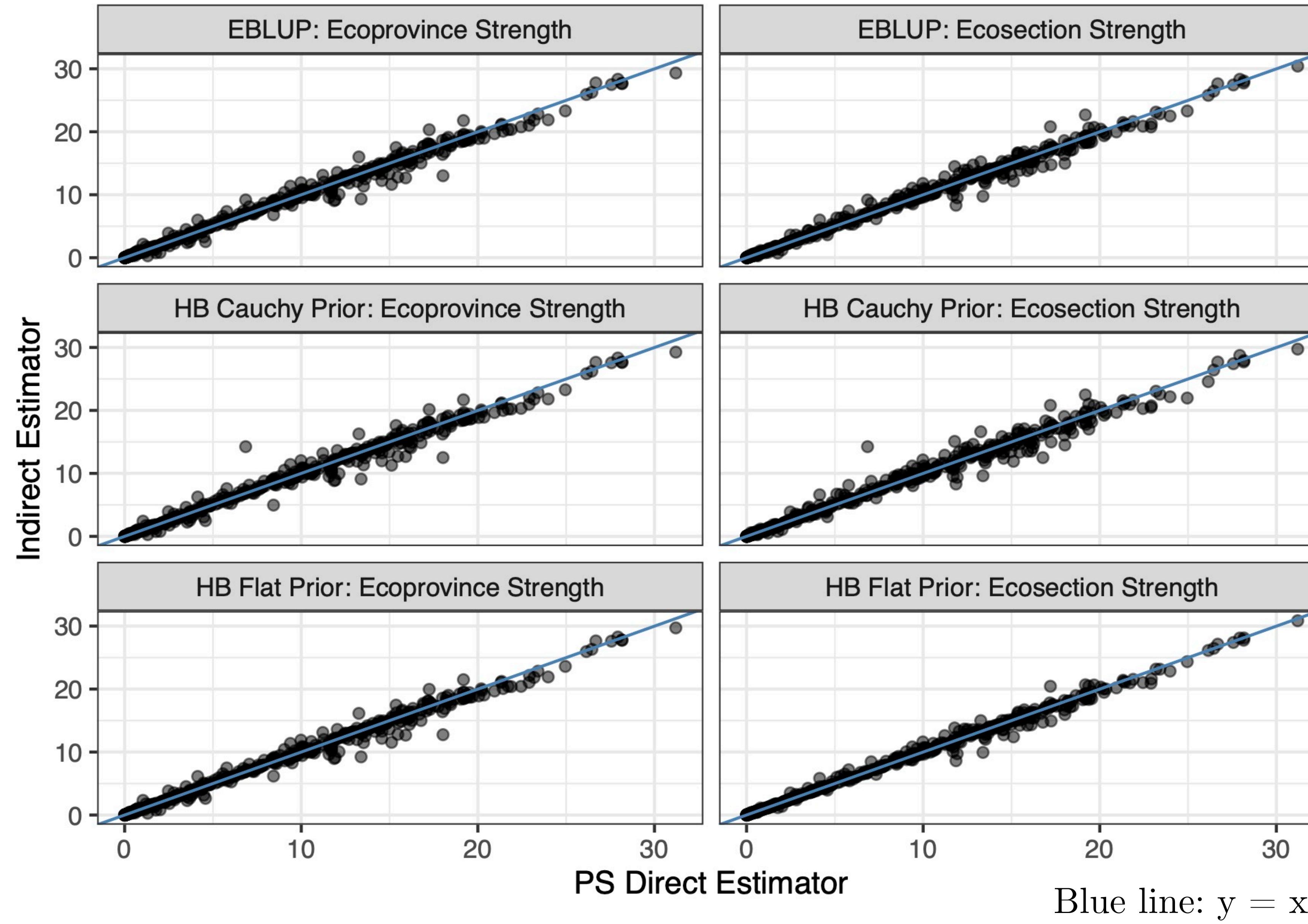


Mt. Hood, Oregon

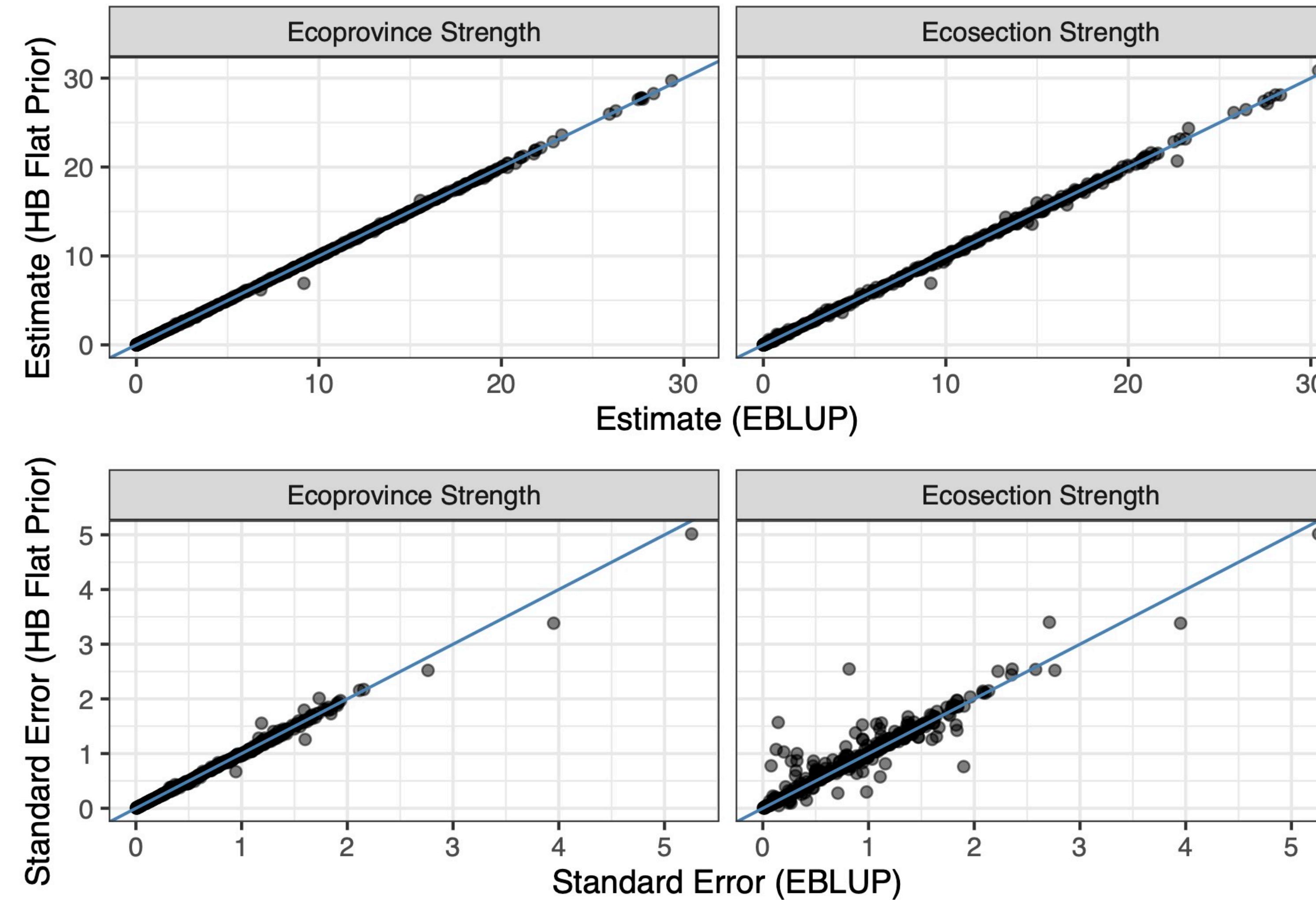


Valley of the Giants, Oregon

Assessing Bias via Agreement with Post-Stratified Estimator

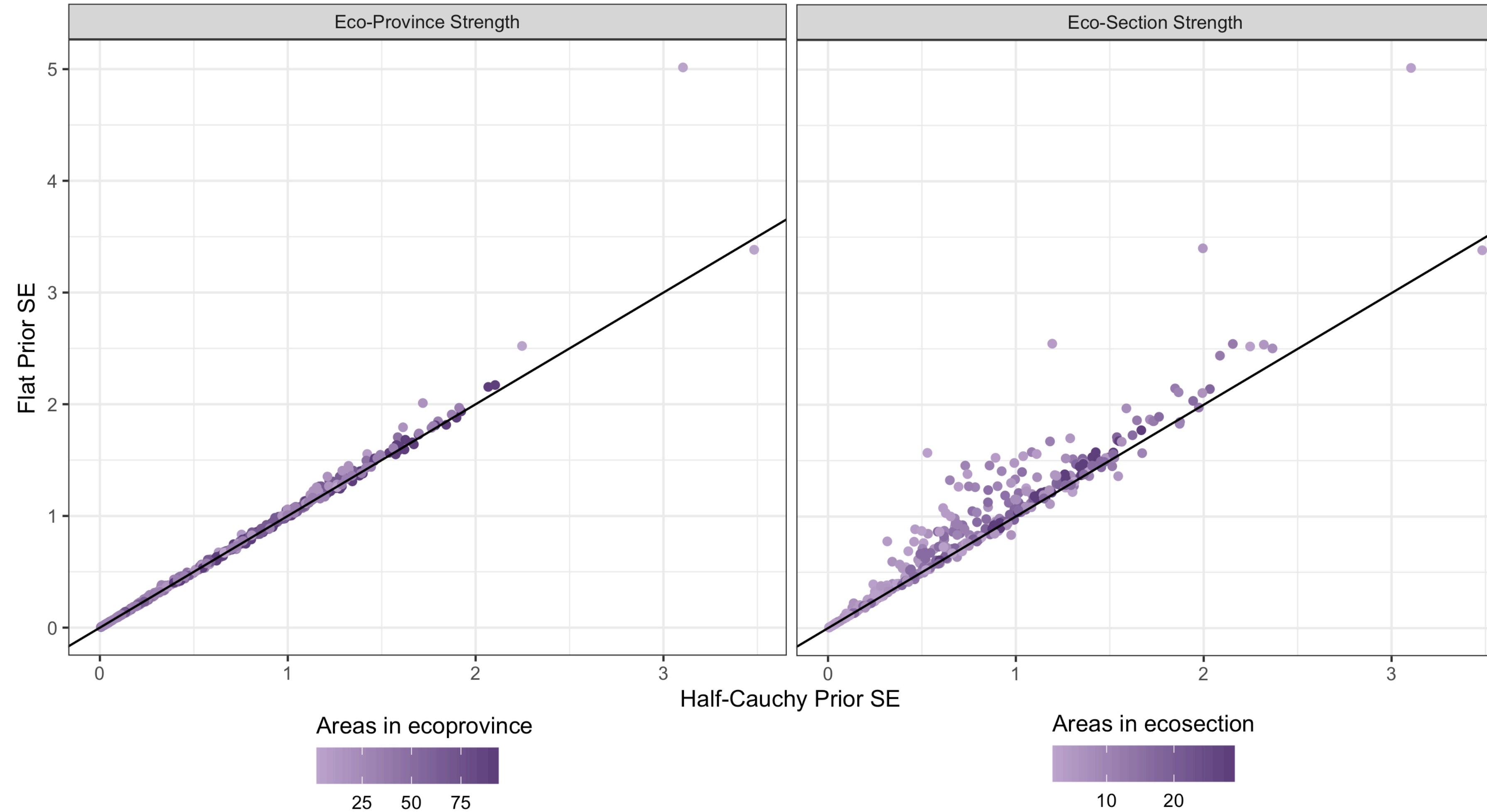


Comparing the HB Flat Prior and EBLUP Estimates (top) and Standard Errors (bottom)



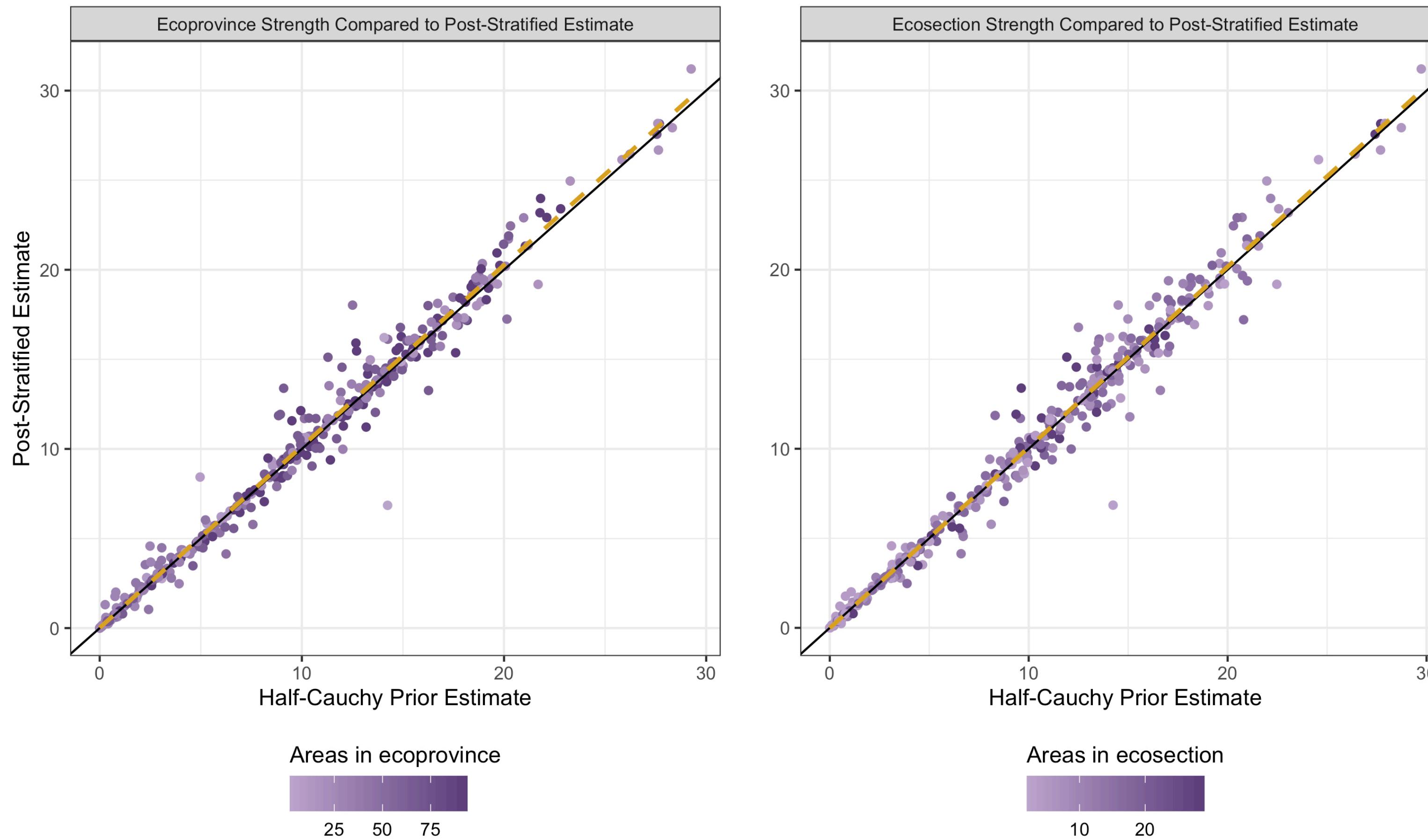
Blue line: $y = x$.

Standard Error differences in HB models



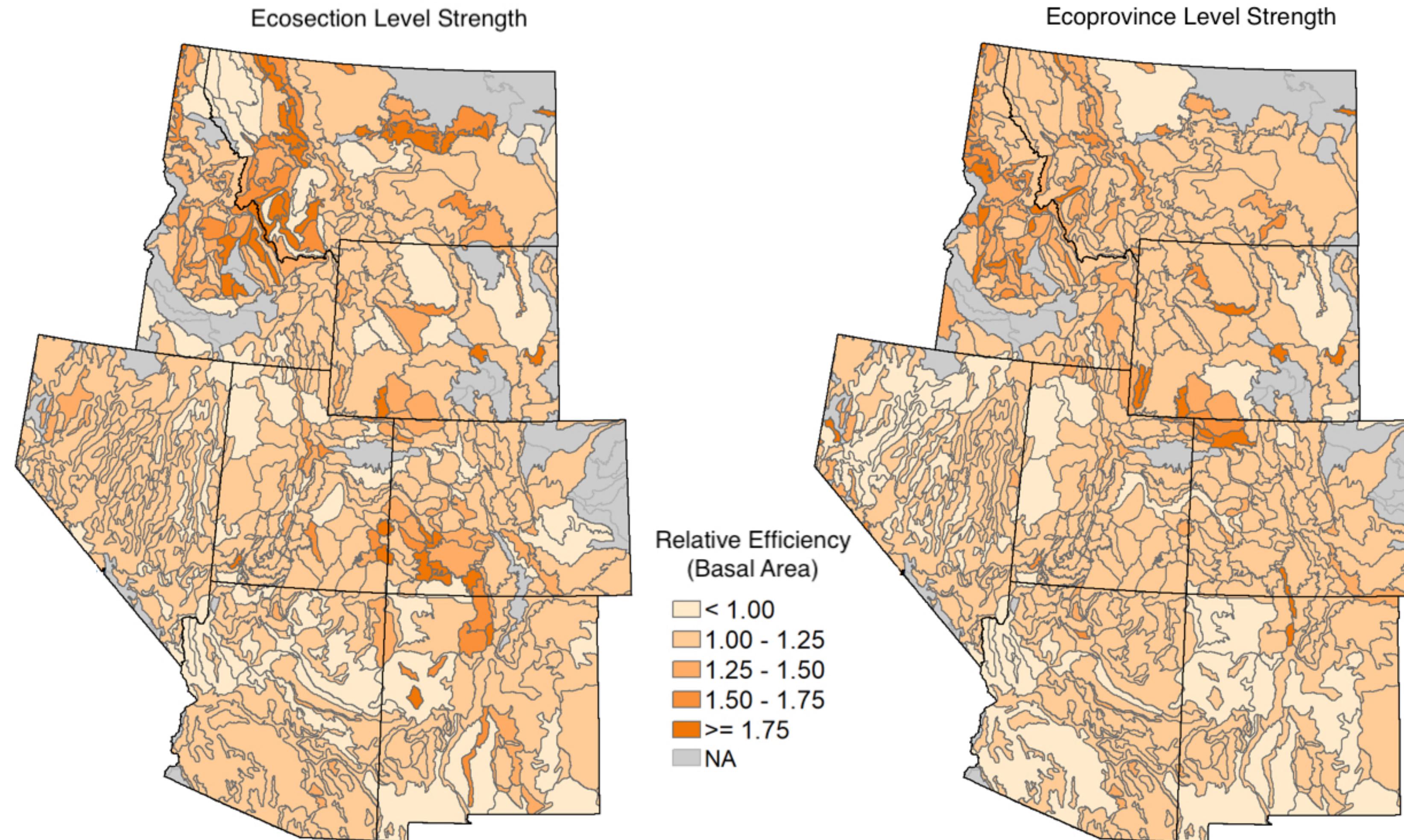
- Black line: $y = x$.
- Purple hue:
amount of areas
strength is
borrowed from.

HB half Cauchy Prior: Provincial vs. Sectional Borrowing Estimates



- Black line: $y = x$.
- Yellow dashed line: OLS best fit.
- Purple hue: amount of areas strength is borrowed from.

Relative Efficiency of HB half Cauchy Prior Estimators and Post-Stratified Estimator



Future Work



Death Valley, California

- Applying these methods across the rest of the United States.
- Unit-level small area models: zero-inflated models.
- A simulation study to understand potential bias.
- Multivariate small area models for many FIA variables of interest.

Thank you!