PINEMAP

Region-wide calibration of 3-PG using data assimilation

Annika Jersild & R. Quinn Thomas

Virginia Tech

Eric Ward (AIM 1)

Many others from the 3PG team (especially Evan, Carlos, and Randy)

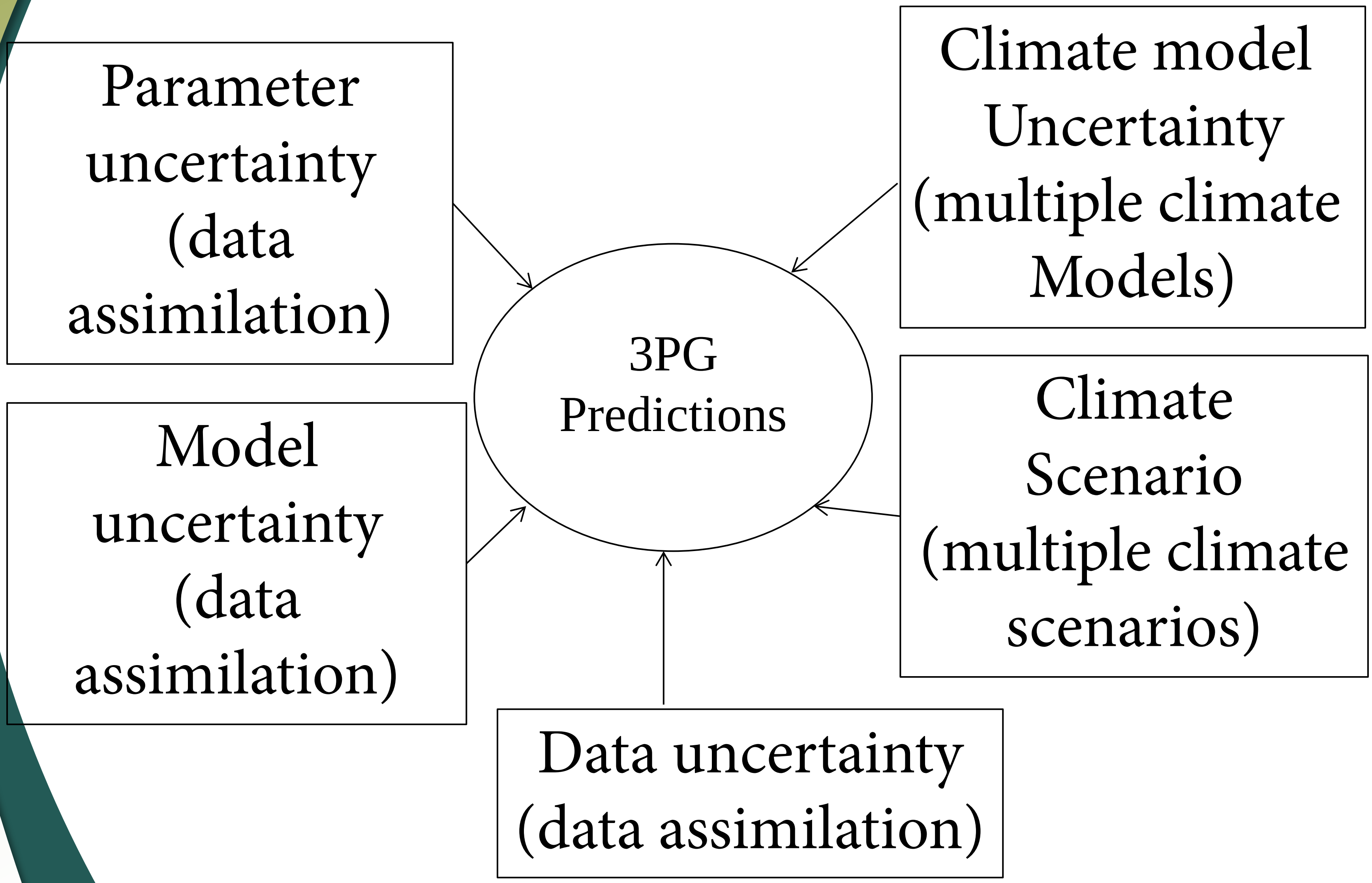
PINEMAP is a Coordinated Agricultural Project funded by the
USDA National Institute of Food and Agriculture, Award # 2011-68002-30185.

USDA United States Department of Agriculture National Institute of Food and Agriculture

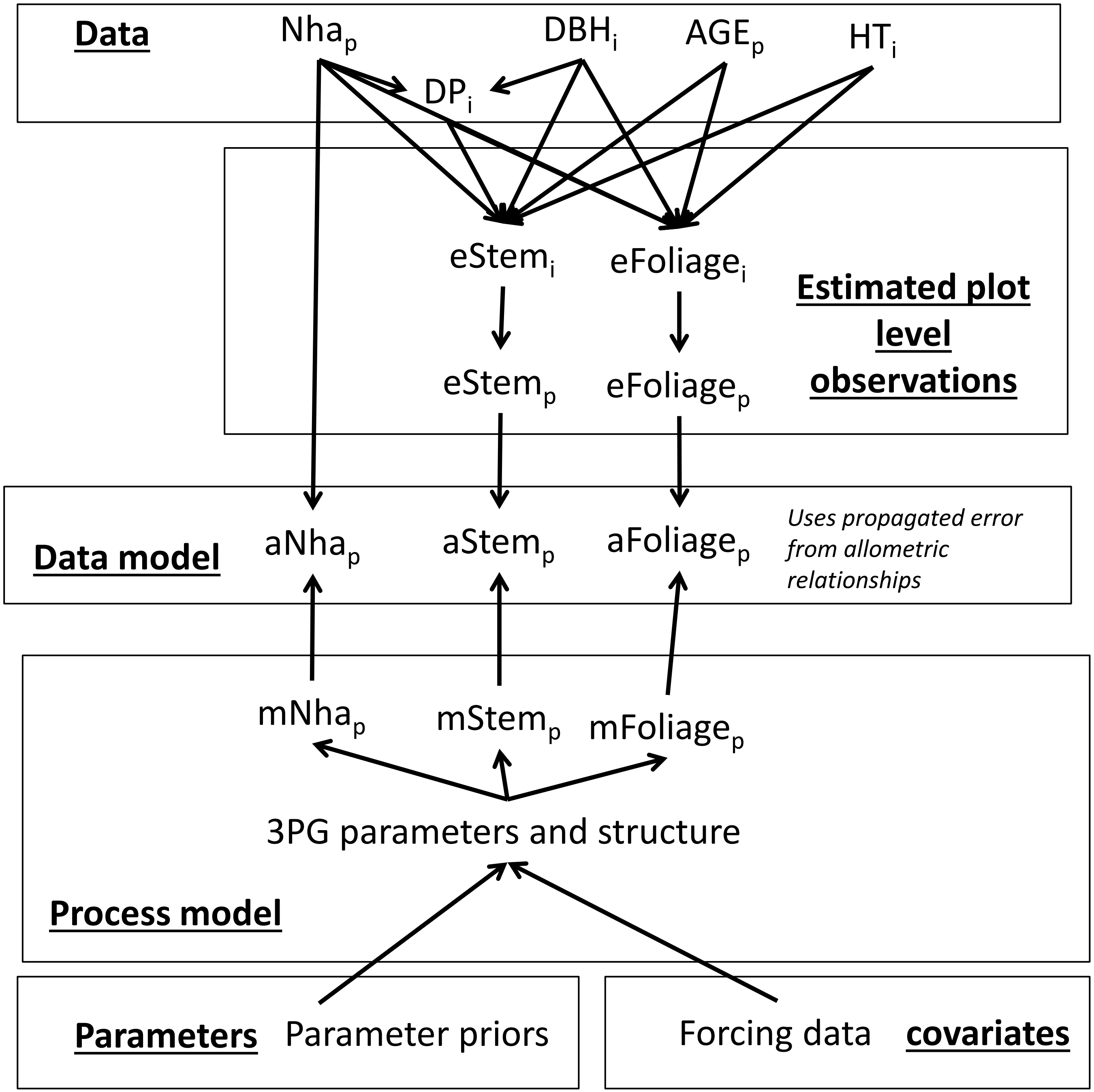
Overview

Make predictions for loblolly pine productivity with a known uncertainty that is consistent with region-wide data and prior knowledge

Known Unknowns



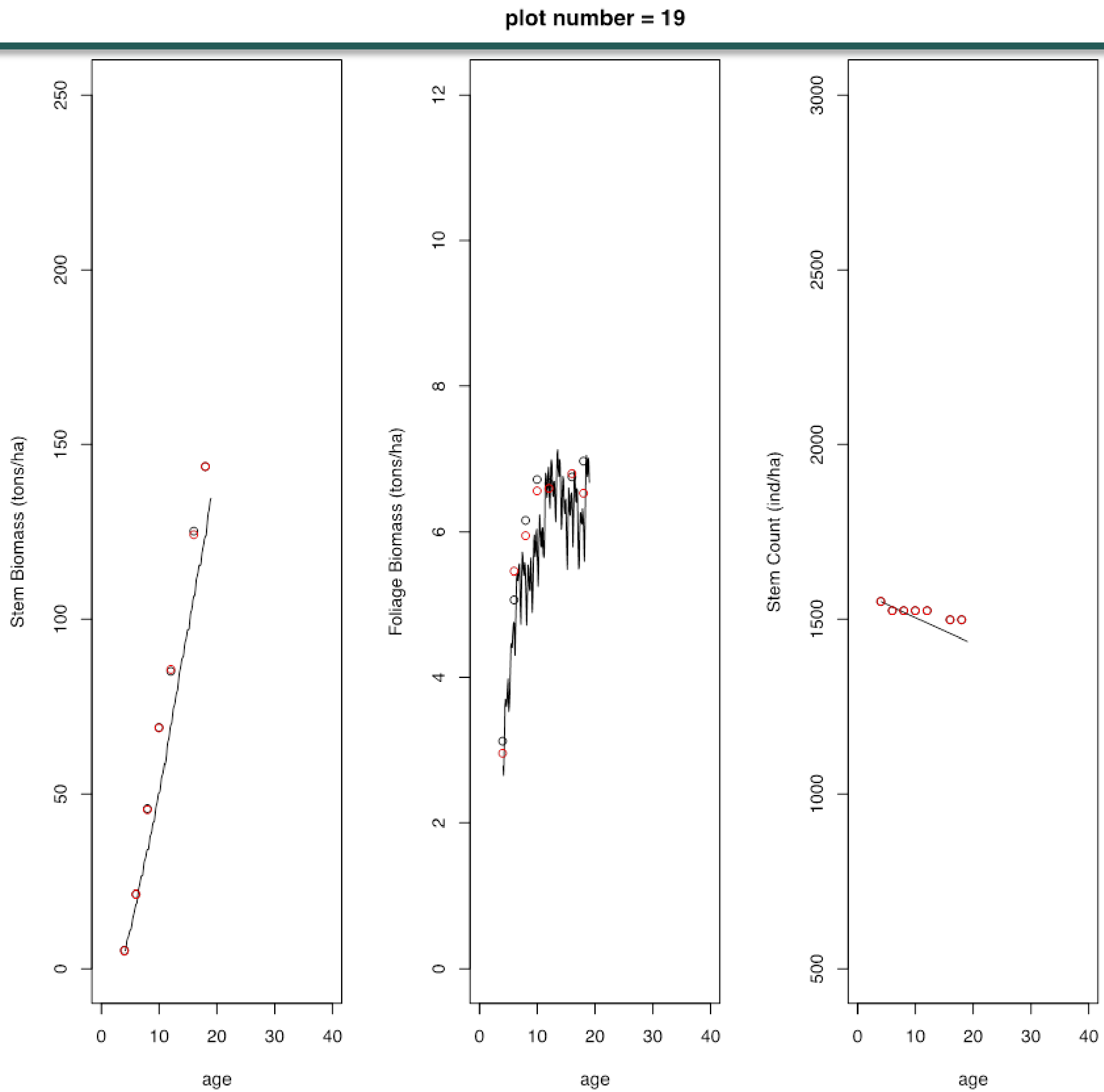
Tier 1/Tier 2 data assimilation



Completed

- Data-assimilation algorithm is running on 44 sites from Carlos set of validation data using the Flavor C.
- ~88 sites are not used in data-assimilation and are available for testing.
- Our algorithm now allows for
  - ‘data uncertainty’ associated with using allometric equations to calculate plot level biomass values
  - Estimating initial available soil water
- Capacity to make single plot predictions that integrate over parameter uncertainty
- Capacity to simulate multiple climate models and scenarios at the plot level

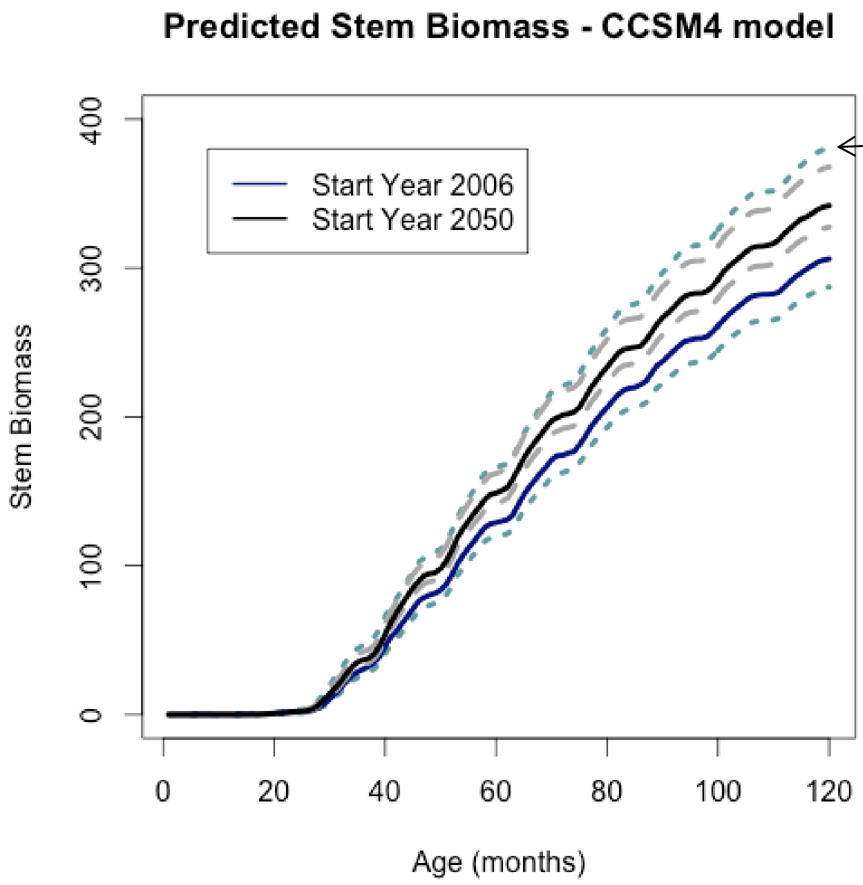
Preview



plot number = 19
Stem Biomass (tons/ha)
0
50
100
150
200
250
Foliage Biomass (tons/ha)
0
2
4
6
8
10
12
Stem Count (ind/ha)
500
1000
1500
2000
2500
3000
0
10
20
30
40
age

Preview

- Example test analysis from unfinished simulation



95% credible interval

Current status

- Working with Eric (AIM 1) to organize the Tier 3 and Duke FACE.
  - Starting with simulating the control and drought treatment at the Virginia site Tier 3
    - Focus on biomass, LAI, litterfall, and canopy transpiration
  - Simulating Duke ambient and elevated CO2 treatements
    - Focus on biomass, LAI, litterfall, and canopy transpiration
- Working with Eric and Asko (AIM 1) to organize Duke and Coastal NC flux tower and biomass data
  - Focus on winter time GPP and canopy transpiration
  - Winter allows the removal of deciduous tree influence

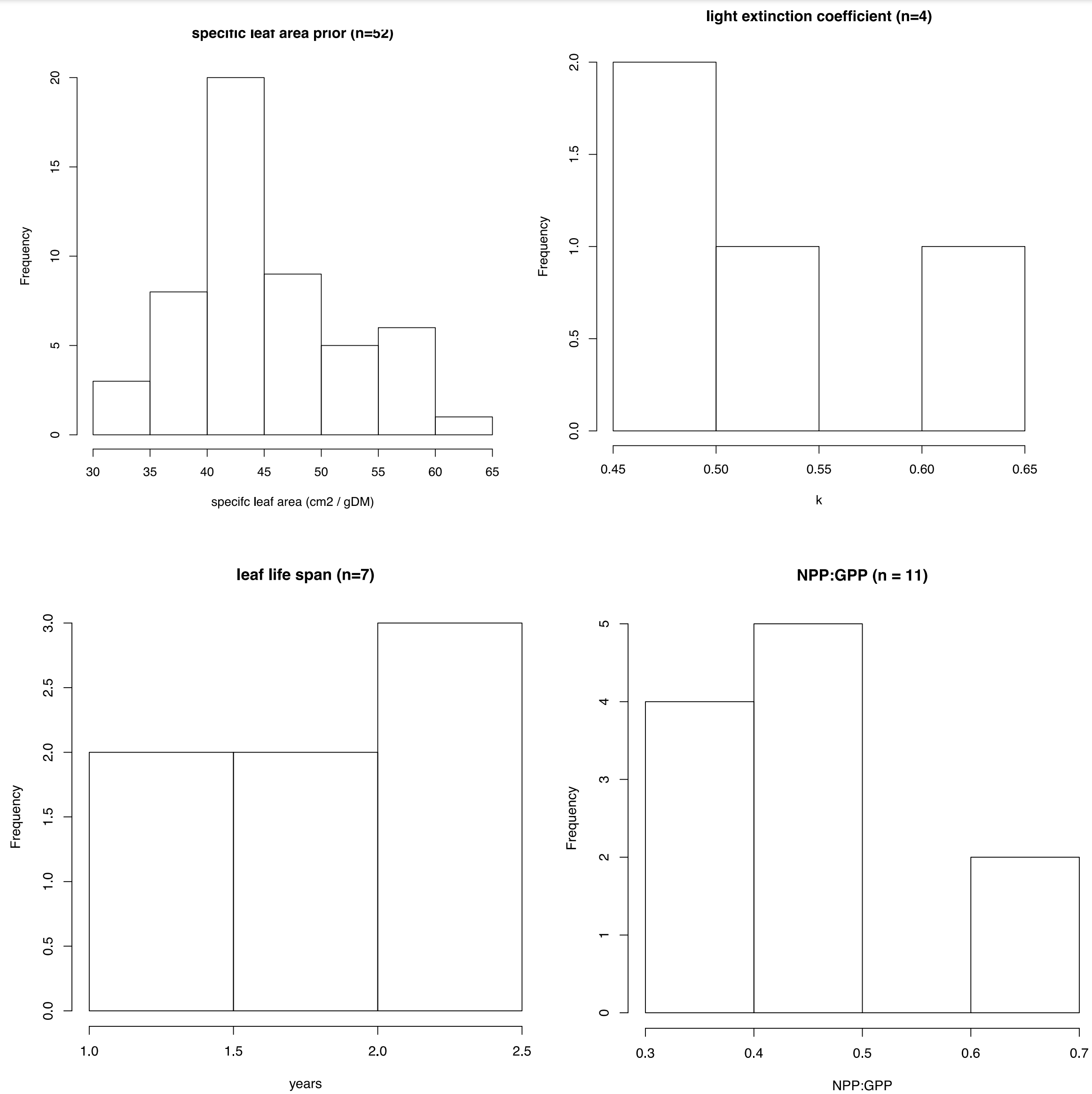
Current Status

- The integration of both observational and experimental data from across the region is unique
  - Drought response and elevated CO_2 responses will included in the parameterization
- Trade-off among parameters means that we need to fix or highly constrain parameters.
  - Example: multiple ways to fit a reduction in foliage biomass (or LAI) with age
  - Example: Minimum temperature and frost day parameters

Needs

- Priors
  - Specific leaf area
  - Leaf-life span
  - Strength of photosynthesis decline with age
  - Max and min monthly temperatures for photosynthesis
  - Root allocation or aboveground to belowground biomass ratios
  - Ratio of foliage to leaf allocation at 2 and 20 cm DBH
- What would be a good method to create the priors?
  - Sharing excel files with observations
  - Group call with interested parties

Examples of priors (but need more data)



Other needs

- Idaho ‘observed’ climate data to simulate sites where weather data is not available

Upcoming

- Three AGU oral talks in the session:
  - Constraining Ecosystem Carbon Uptake and Long-Term Storage with Integrated Modeling, Experiment, and Observation

Approaches and Constraints

Site Index Expansion

Site Index

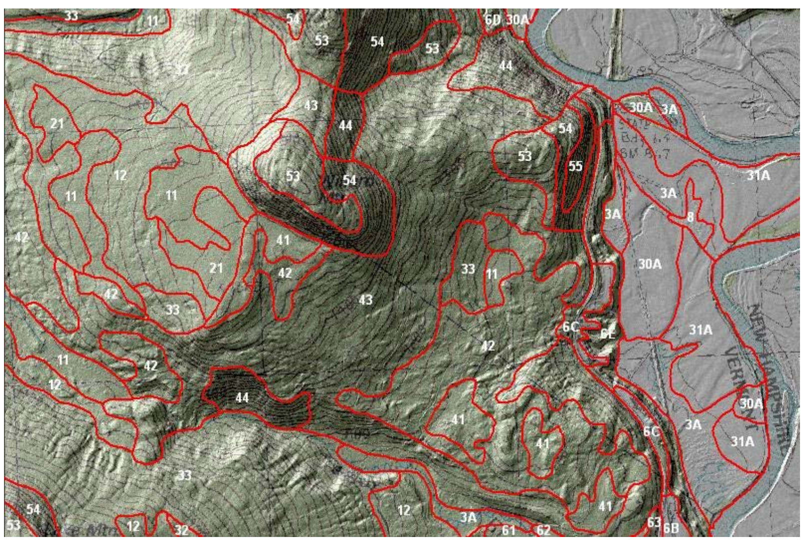
$$SI = f(Soil, Climate, Silviculture)$$

- We are already varying climate and silviculture within PINEMAP
  - Thus, we would prefer a SI map which filters out these influences on SI

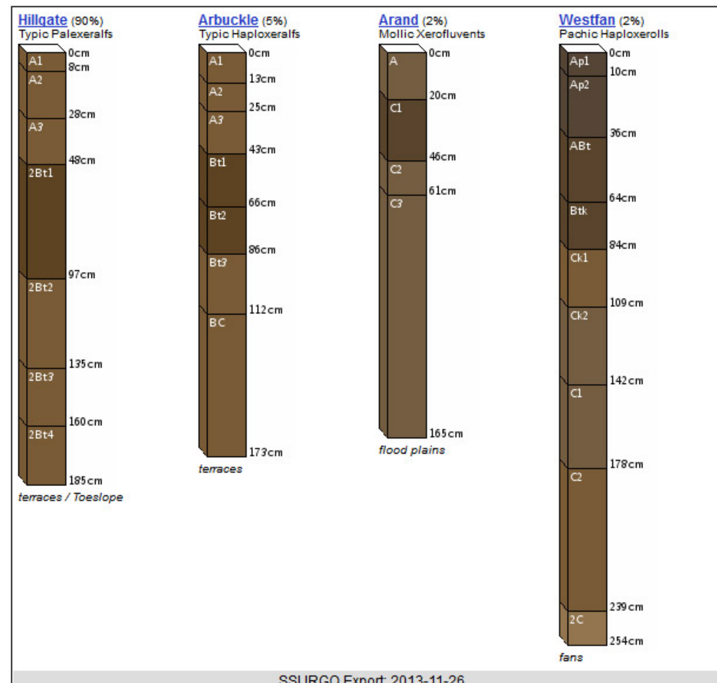
- Immediately, we have the question of whether or not estimating an extended SI for loblolly pine is within the scope of the original grant

SSURGO

- The SSURGO database contains data across the CONUS
  - Broken down into geographic “mapunits” which are in turn comprised of “components” which represent a specific sequence of “horizon” data



http://www.arcsie.com/internal_img/VectorizationFig_Final.jpg



http://casoilresource.lawr.ucdavis.edu/software/postgis-spatially-enabled-relational-database-sytem/analysis-ssurgo-data-postgis-overview/logistics-getting-connected-and-executing-queries/getting-parent-material-data-out-ssurgo/

SSURGO

- The SSURGO dataset includes a table (coforprod) which lists SI measurements by component for species observed in a particular component
  - The values for loblolly pine were used as the basis for previous SI maps for regional 3-PG runs
    - Missing values were backfilled using a state-by-state approach in which the missing value was imputed by the mean known SI for that particular component's series
- However, if we want to extend the SI map beyond places where loblolly pine data are observed, we need a different approach

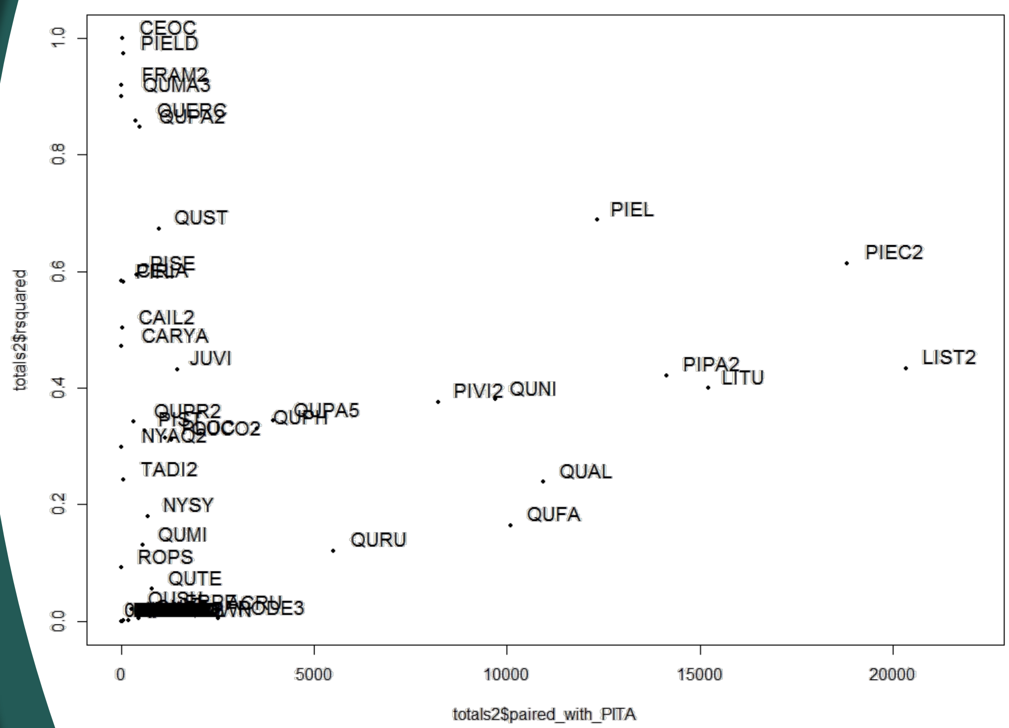
Two Methods

Method	Advantages	Disadvantages
SI matching to other species with different ranges	• Uses observed data • Straightforward processing	• Restricted by ranges and data availability for other species • Climate implicitly included • Doesn't actually impute "missing" components
Modeling SI using edaphic factors only (possibly based on "standard" climate inputs)	• Data readily available • No restrictions on range	• Generally weak fits • Confounding influences on "soil" variables

SI Matching to Other Species

- From the coforprod table for entire region of interest, convert from long to wide format
  - Each row represents a single component, with columns for the SI of each species (~150 species in all)
- For each candidate species, count the number of components which have data for that species *and* loblolly pine
  - Also generate a simple linear regression between the SIs for that species against loblolly pine

SI Matching to Other Species



By the numbers, the best candidates include…

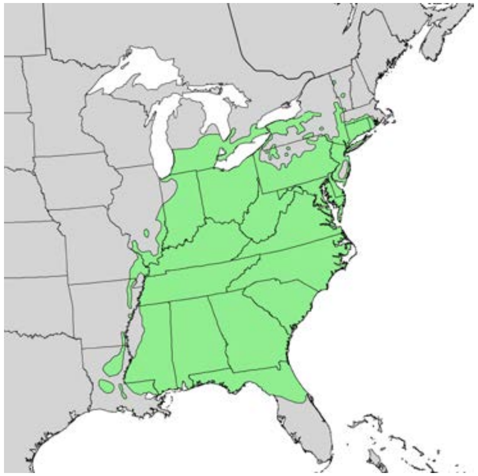
- Shortleaf Pine (PIEC2)
- Sweetgum (LIST2)
- Slash Pine (PIEL)
- Yellow Poplar (LITU)
- Longleaf Pine (PIPA2)

Other contenders might be…

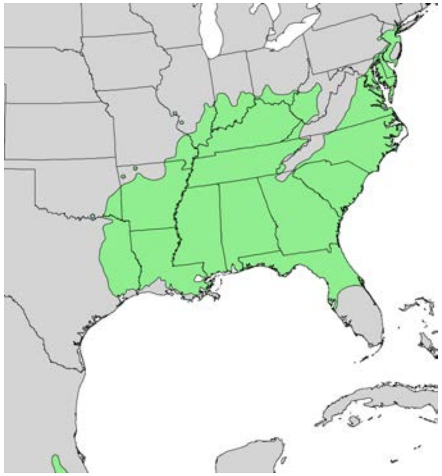
- White Oak (QUAL)
- Water Oak (QUNI)
- Virginia Pine (PIVI2)

SI Matching to Other Species

- Species with range that may be suitable for SI expansion



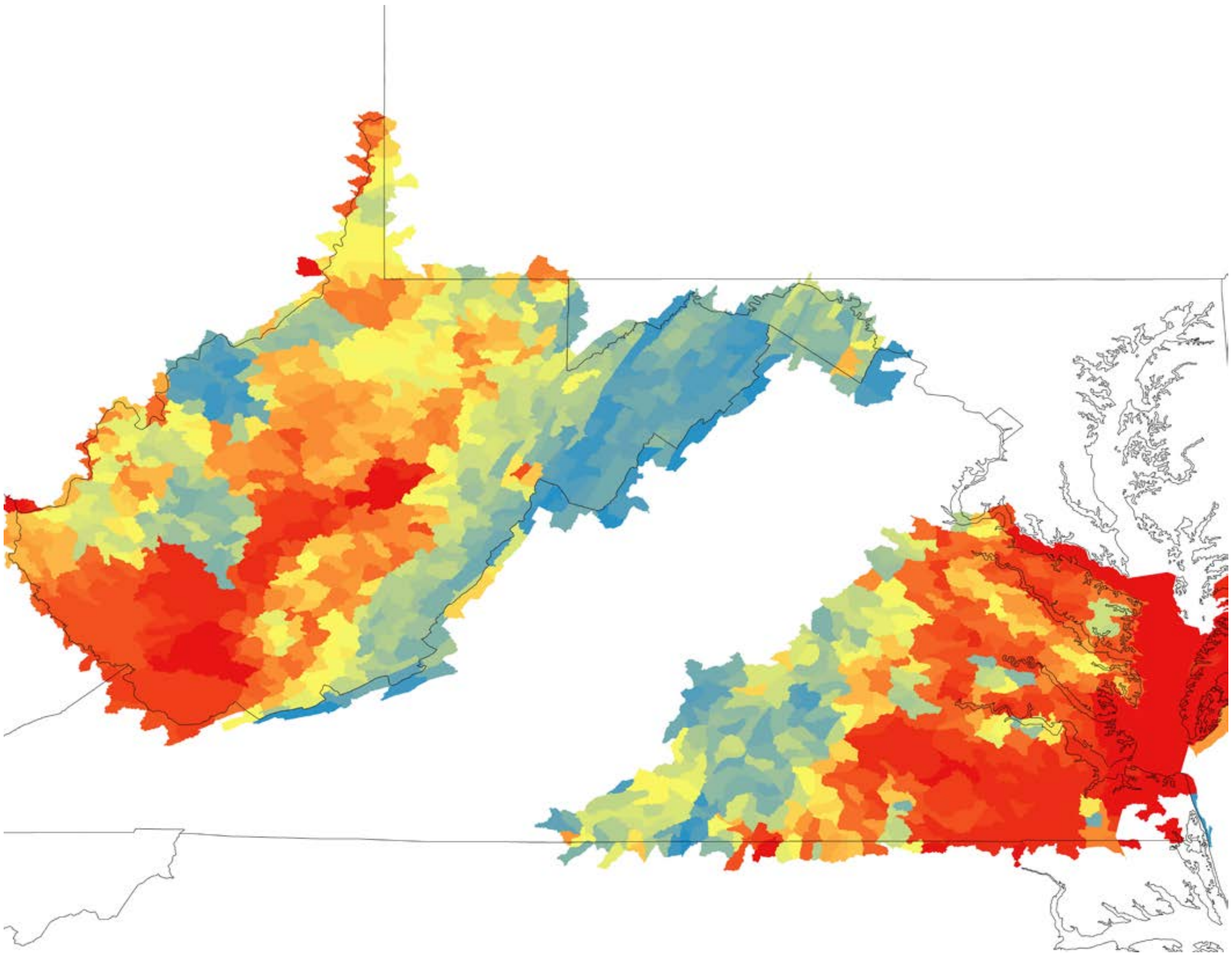
Yellow Poplar (LITU)
$R^2 \sim 40\%$



Sweetgum (LIST2)
$R^2 \sim 43\%$

White Oak (QUAL)
$R^2 \sim 24\%$

SI Matching to Other Species



Edaphic SI Modeling

- The greatest challenge in modeling SI by soil alone is that SI is intrinsically linked to climate as well
  - Need to be wary of confounding variables
  - Perhaps model SI with the climate inputs included, then use a typical or standardized set of climate inputs to get the variation due to soil

Edaphic SI Modeling

- Variables available at the component and horizon levels

Variable	Type	Table	r_{SI}
Clay	Soil	Chorizon	-0.25
Silt	Soil	Chorizon	0.11
Organic Material	Soil	Chorizon	0.11
Slope	Geographic	Component	-0.41
Elevation	Geographic	Component	-0.33
Aspect	Geographic	Component	0.09
Albedo (dry)	Soil	Component	-0.12
Air Temperature (Annual Mean)	Climate	Component	-0.11
Precipitation (Annual Mean)	Climate	Component	-0.11
Frost Free Days	Climate	Component	-0.08
Available Cations	Soil	Chorizon	-0.12

Edaphic SI Modeling

- Nonlinear model form akin to Charles Sabatia's SI function (for the G&Y model) seems desirable, but unfortunately even the initial parameter values for model fitting have been extremely difficult to estimate
  - The following results are from using a linear model

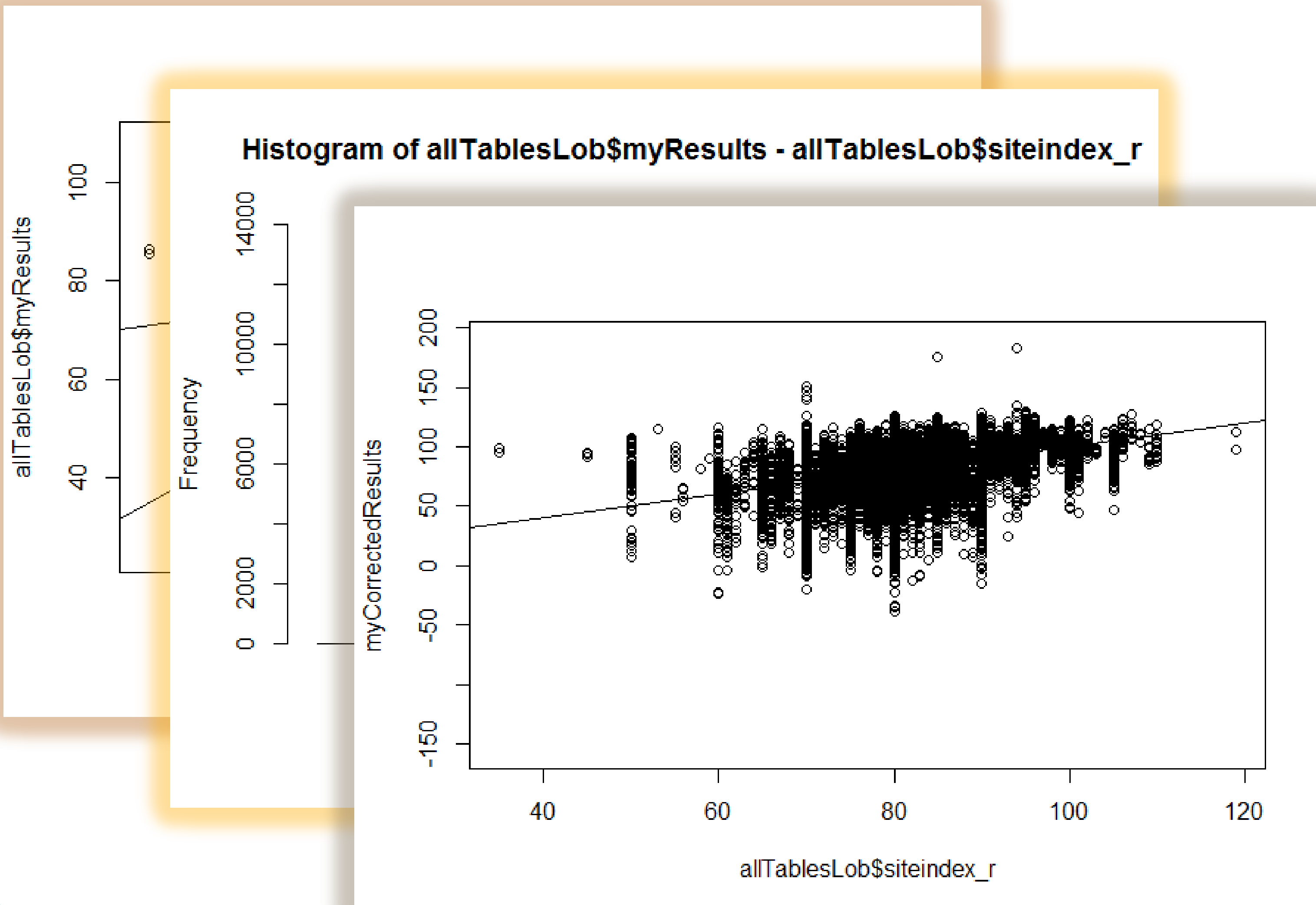
Edaphic SI Modeling

```
Residuals:
    Min      1Q  Median      3Q     Max
-56.710  -5.094   0.447   5.516  34.243

Coeffici

(Intercep
clay
silt
om
cec7
airtempa_
ffd_r
map_r
---
Signif. c

Residual
  (90271
Multiple
F-statist
```

```
Residuals:
    Min      1Q  Median      3Q     Max
-55.368  -4.885   0.772   4.902  33.269

Coefficients:

(Interc
clay
silt
om
cec7
airtemp
ffd_r
map_r
slope_r
---
Signif.

Residua
  (9027
Multipl
F-stati
```

```
Residuals:
    Min      1Q  Median      3Q     Max
-54.638  -4.682   0.563   4.527  30.706

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) 88.7193336  0.1115566 795.285  < 2e-16 ***
clay        -0.1054642  0.0026671 -39.542  < 2e-16 ***
silt         0.0838367  0.0019434  43.139  < 2e-16 ***
om           0.2832498  0.0324898   8.718  < 2e-16 ***
slope_r     -0.3789771  0.0049274 -76.912  < 2e-16 ***
elev_r      -0.0146422  0.0003069 -47.710  < 2e-16 ***
aspectrep    0.0017459  0.0004314   4.047 5.19e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 7.53 on 50989 degrees of freedom
  (65241 observations deleted due to missingness)
Multiple R-squared:  0.2566,    Adjusted R-squared:  0.2565
F-statistic:  2933 on 6 and 50989 DF,  p-value: < 2.2e-16
```

Edaphic SI Modeling



Conclusions and Questions

- Clearly, there are some major challenges and caveats to either approach to extending loblolly pine SI estimates beyond the natural range

- Questions:
  - Is it worth it?
    - Is it within the scope of PINEMAP?
  - If so, which approach should we use?
    - Yellow Poplar had the best match so far, but we still have numerous gaps to fill if we go this route