

# Alternative Douglas-fir Diameter Growth

Greg Johnson and David Marshall

2026-01-19

## Data

We extracted and processed Forest Inventory and Analysis (FIA) data from 11 states listed in the native range of Douglas-fir in the Silvics of North America.<sup>1</sup>

After subsetting the data to censor observations with missing data, limiting the species to Douglas-fir (FIA species code 202), and remeasurement intervals  $\geq 5$  years we get the observations in Table 1.

Table 1: Douglas-fir Growth Observations by State

State Code	Observations
4	1143
6	18605
8	6774
16	13786
30	23645
32	27
35	2440
41	69362
49	2345
53	47317
56	865

<sup>1</sup>Burns, Russell M., and Barbara H. Honkala, tech. coords. 1990. Silvics of North America: 1. Conifers; 2. Hardwoods. Agriculture Handbook 654. U.S. Department of Agriculture, Forest Service, Washington, DC. vol.2, 877 p.

## Alternative Model Formulation

An alternative to the ORGANON diameter growth equation<sup>2</sup> which reduces parameter count while retaining key features of the original model is shown below. The key change is the term with a ratio of a transformation of diameter at breast height (dbh) squared to crown length. Since  $\beta_1$  is expected to be negative, this tends to slow growth as more basal area accumulates in the tree while moderating that decline by the amount of productive crown capacity as measured by crown length. Basal area in larger trees (**bal**) serves as the inter-tree competition factor, and site index (**si**) as the inherent productivity scaling factor.

Site index is flawed for a number of reasons:

1. It is not consistently obtained for each plot due to missing Douglas-fir site trees,
2. It is estimated using a number of different and not necessarily compatible **si** equations, and
3. The available **si** equations do not all use the same base age.

In the data set **si** is derived from 31 different site index equations for 25 species. Douglas-fir site index comprises 85% of the observations. There are 3 base ages used. Preliminary graphical analysis revealed that base age was most correlated with residual bias. Thus in the following, we fit two equations: one where **SIBASE** and **SISP** are treated as a random effects in a mixed model framework, and a second leaving site index out.

$$\Delta dbh = e^{(\beta_0 + \beta_1 \log(\frac{(dbh+1)^2}{(cr*ht+1)^{\beta_4}}) + \beta_2 \frac{bal^{\beta_5}}{dbh+2.7} + \beta_3 \log(si_{s,b} + 4.5))} \quad (1)$$

and

$$\Delta dbh = e^{(\beta_0 + \beta_1 \log(\frac{(dbh+1)^2}{(cr*ht+1)^{\beta_4}}) + \beta_2 \frac{bal^{\beta_5}}{dbh+2.7})} \quad (2)$$

where:

- **dbh** = diameter at breast height (inches),
- **bal** = basal area per acre in larger trees (*feet*<sup>2</sup>/*ac*),
- **cr** = crown ratio (fraction of total height),
- **ht** = total height (feet), and
- **si<sub>s,b</sub>** = site index (feet) for species **s** and base age **b**.
- $\beta_0 - \beta_5$  are parameters to be estimated.

---

<sup>2</sup>Hann, D.W., Marshall, D.D., and Hanus, M.L. 2006. Reanalysis of the SMC-ORGANON equations for diameter-growth rate, height-growth rate, and mortality rate of Douglas-fir. Forest Research Laboratory Research Contribution 49.

Nonlinear regression was used with an integrated fitting approach such that individual observations can have differing remeasurement intervals. The error to be minimized is ending dbh. Since this effectively minimizes diameter growth it can weight observations with longer remeasurement intervals more heavily. The effect of this needs to be evaluated, but putting more emphasis on longer periods may be beneficial.

The fit statistics for Equation 1 are:

```

Nonlinear mixed-effects model fit by maximum likelihood
Model: endDIA ~ est_dg(B0, B1, B2, B3, B4, B5, startDIA, startBAL, endBAL,
Data: df_subset %>% mutate(SIINT = interaction(as.factor(df_subset$SIBASE),
      AIC      BIC      logLik
 542795.5 542876.6 -271389.7

startCR, endCR, startBAL, endBAL, SIBASE, SIINT, est_dg, as.factor, interaction, mutate, %>%, %>%)

Random effects:
Formula: B3 ~ 1 | SIINT
          B3 Residual
StdDev: 0.05281403 1.038348

Fixed effects: B0 + B1 + B2 + B3 + B4 + B5 ~ 1
      Value Std.Error DF t-value p-value
B0 -6.516424 0.030995842 186299 -210.23544     0
B1 -0.594237 0.003152243 186299 -188.51235     0
B2 -0.091642 0.005197183 186299 -17.63303     0
B3  0.961590 0.024365262 186299  39.46559     0
B4  1.819121 0.007812415 186299  232.85005    0
B5  0.631865 0.010101240 186299   62.55319    0

Correlation:
      B0     B1     B2     B3     B4
B1 -0.028
B2 -0.387  0.421
B3 -0.179  0.007 -0.001
B4 -0.520  0.117  0.488 -0.035
B5 -0.360  0.409  0.992  0.001  0.438

Standardized Within-Group Residuals:
      Min        Q1        Med        Q3        Max
-87.36097231 -0.41647572 -0.09129608  0.34419613  19.12554455

Number of Observations: 186309
Number of Groups: 5

$SIINT

```

B3  
 50.FALSE 0.06107920  
 80.FALSE 0.02911915  
 100.FALSE -0.03600856  
 50.TRUE 0.03012150  
 100.TRUE -0.08431129

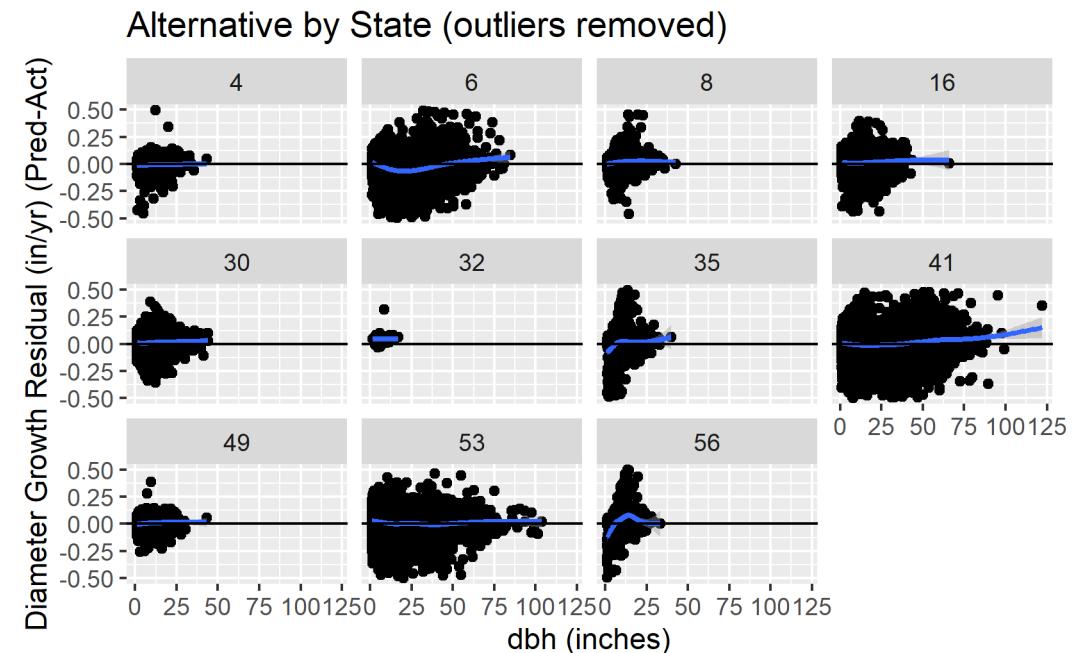
Residual Standard Error: 1.03834787514418 on 186299 degrees of freedom, AIC: 542795.5

and for Equation 2:

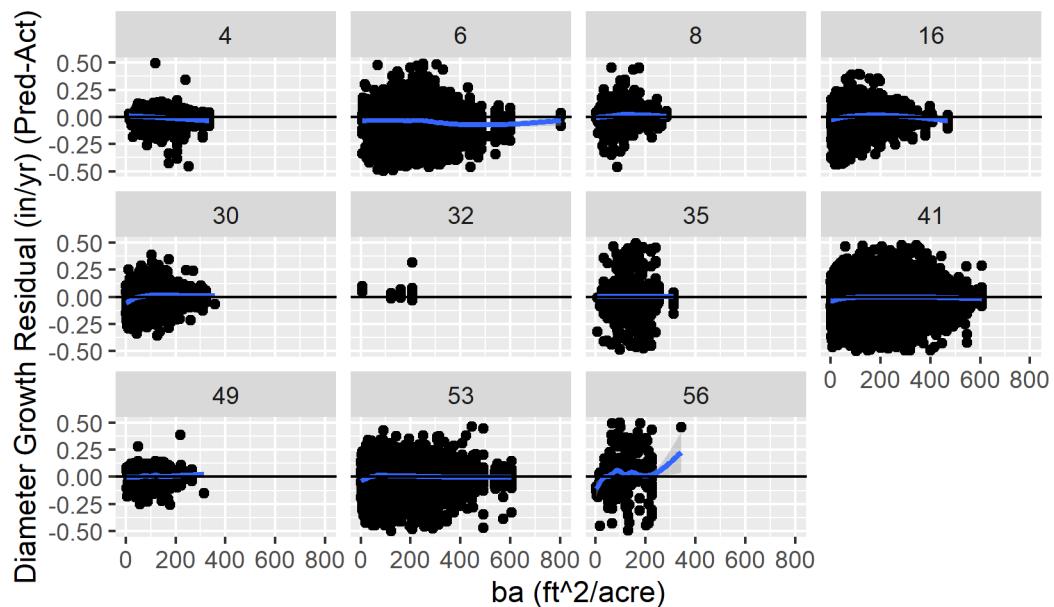
	Coef.	Std. error	t-stat.	p
B0	-3.4525848	0.0248372	-139.00844	0
B1	-0.7036591	0.0038815	-181.28419	0
B2	-0.0880806	0.0079533	-11.07471	0
B4	2.2151993	0.0083456	265.43184	0
B5	0.5636858	0.0157973	35.68249	0

Residual Standard Error: 1.14766280444999 on 186304 degrees of freedom, AIC: 580048.8

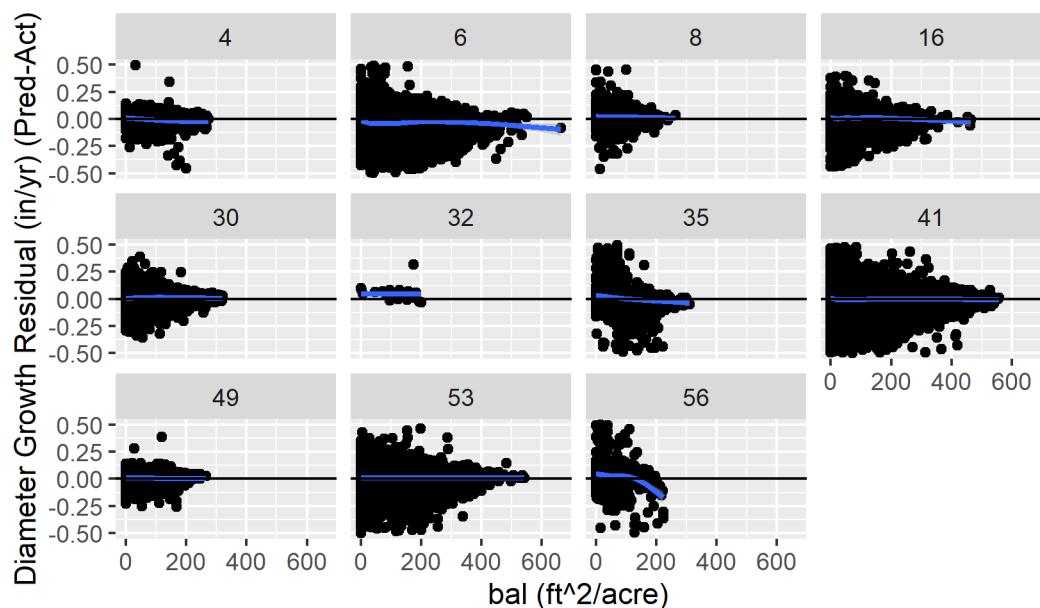
### Residual Analysis for Equation 1



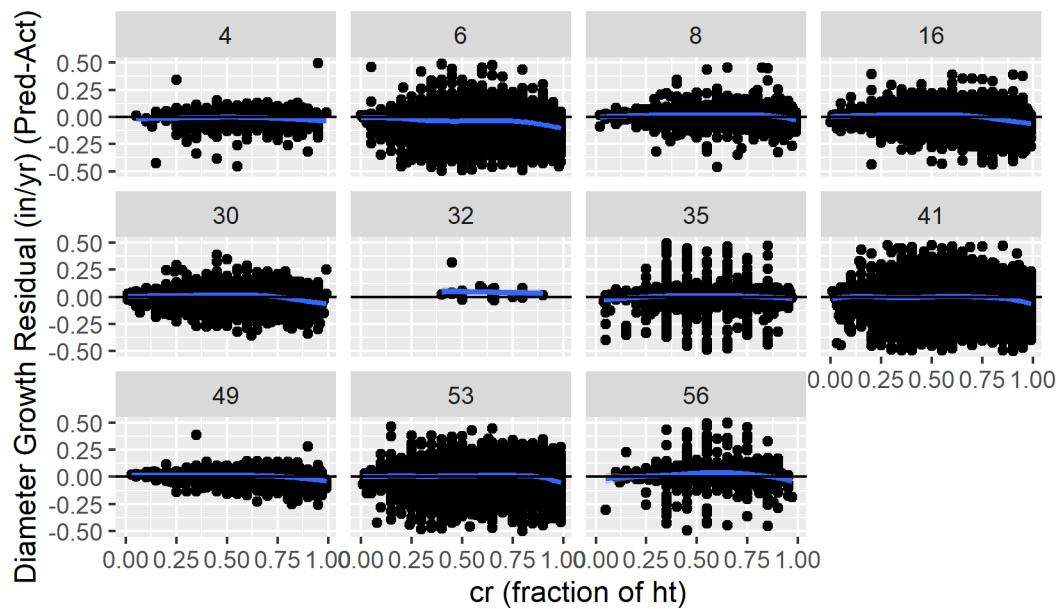
Alternative by State (outliers removed)



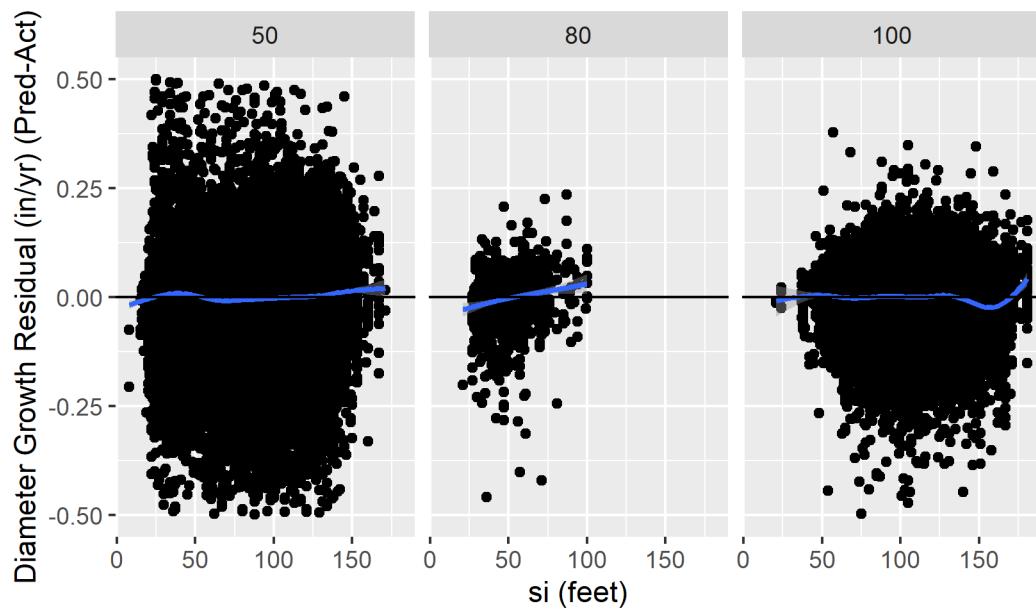
Alternative by State (outliers removed)



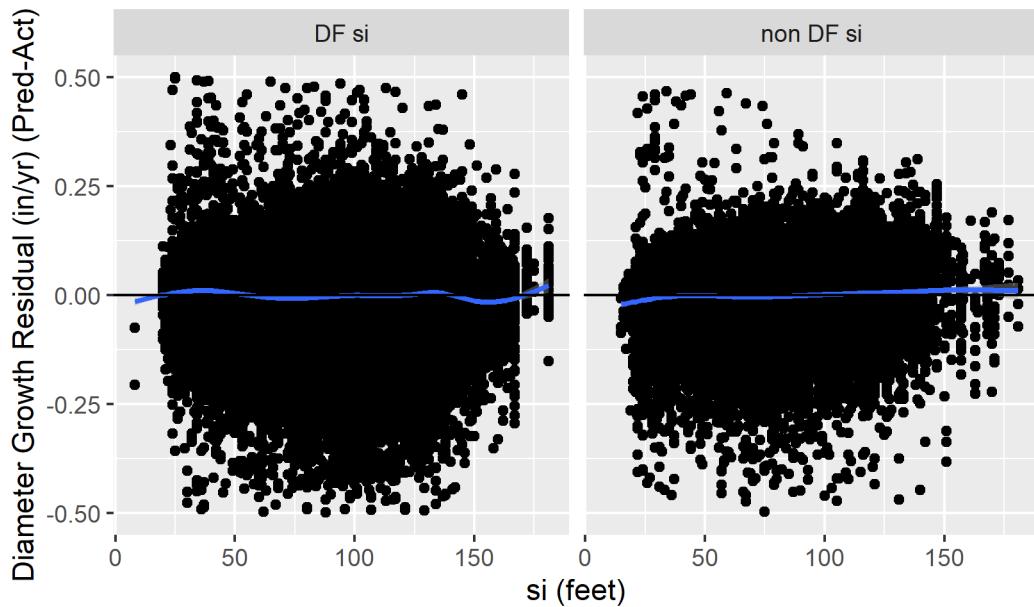
Alternative by State (outliers removed)



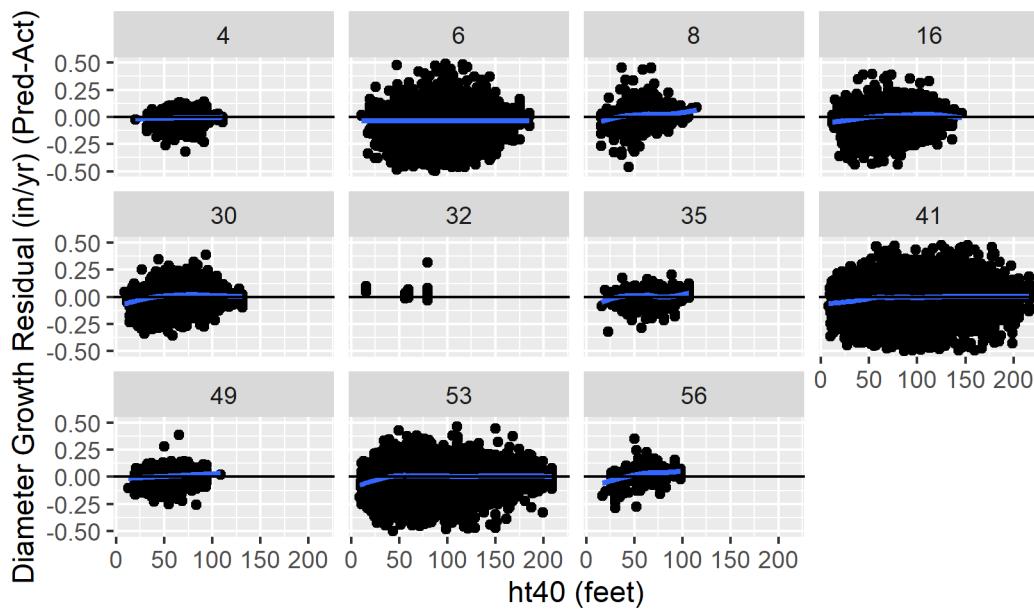
Alternative by si Base Age (outliers removed)

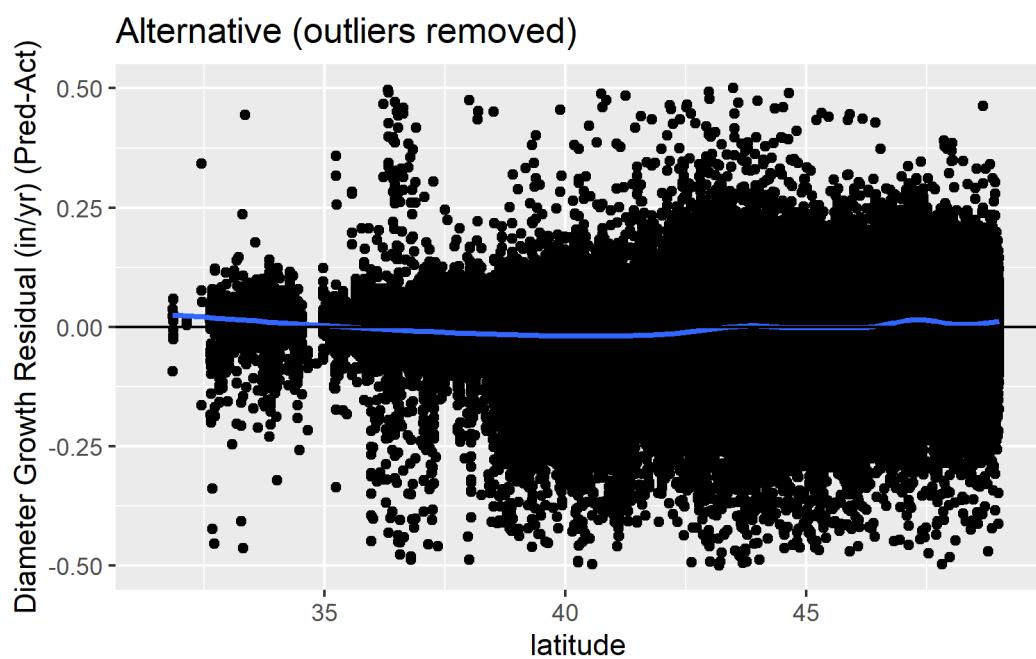
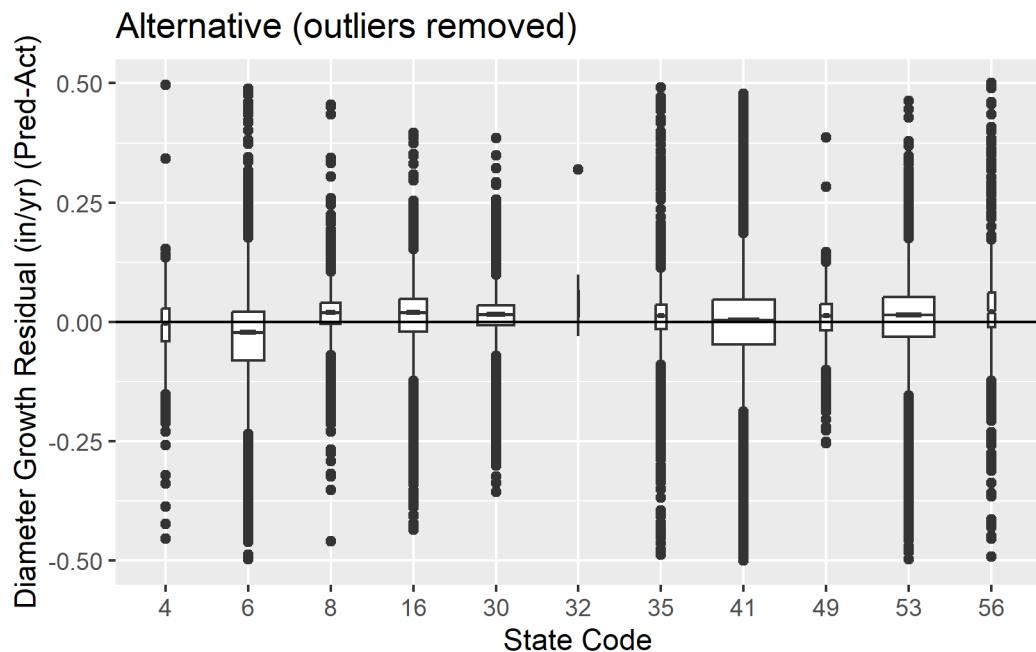


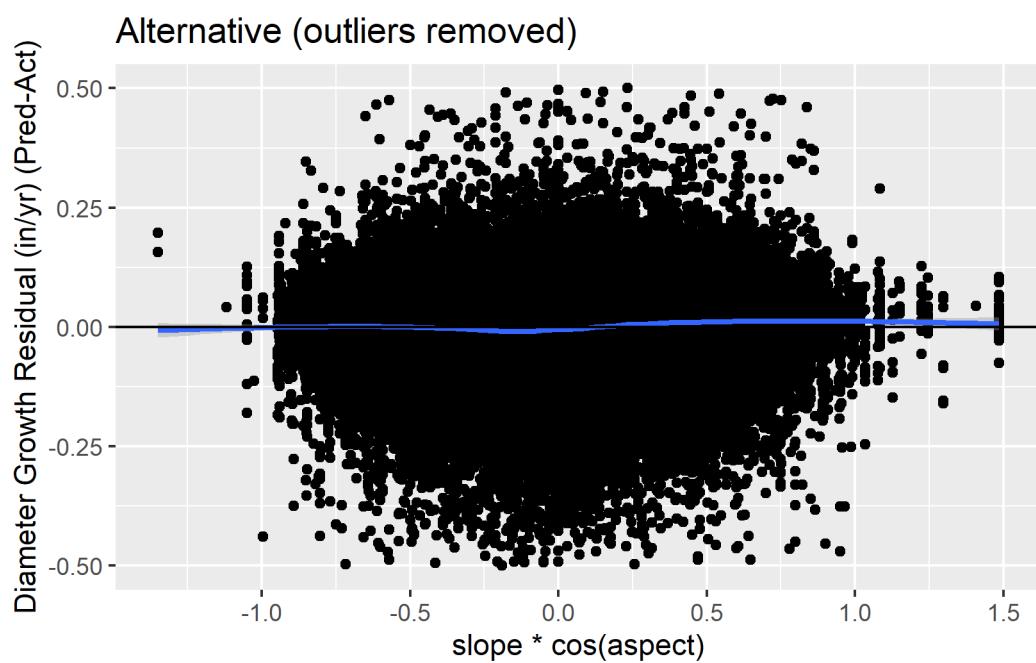
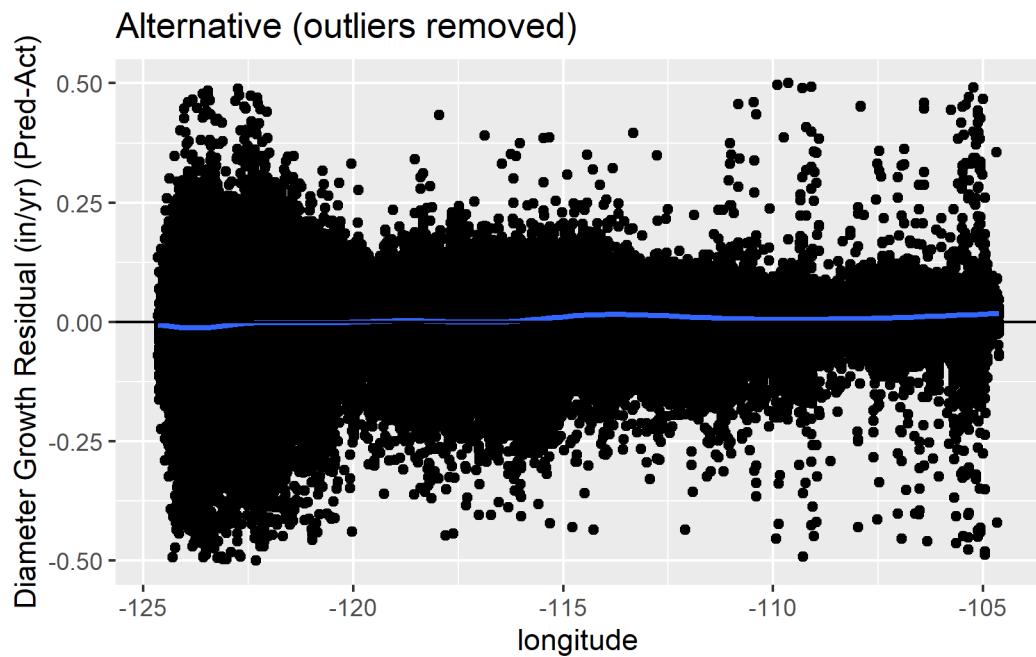
Alternative by si Species (outliers removed)

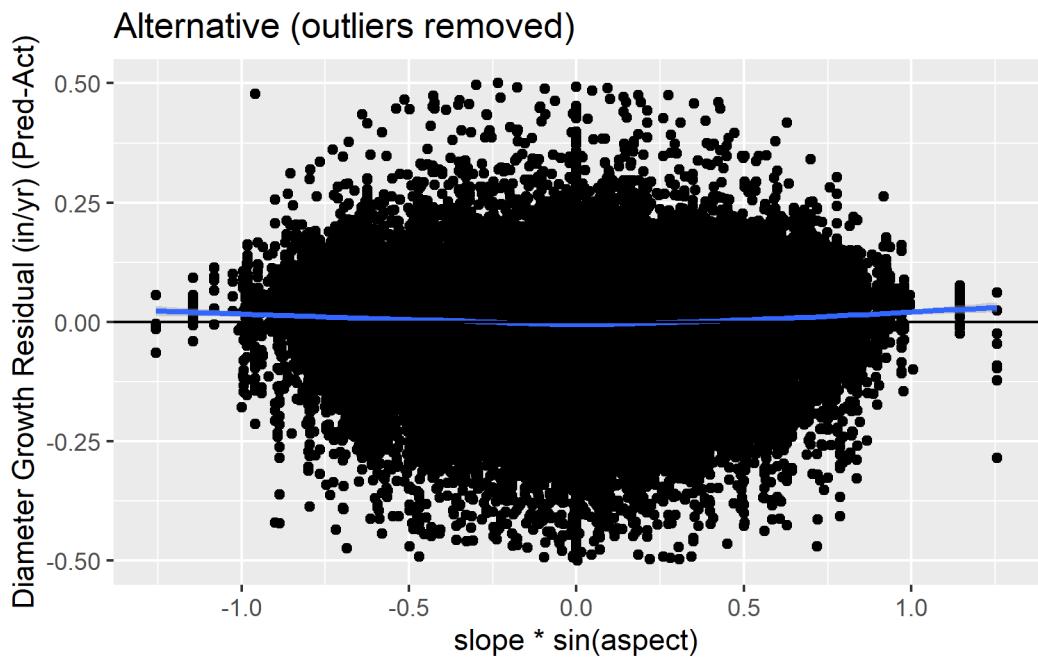


Alternative by State (outliers removed)

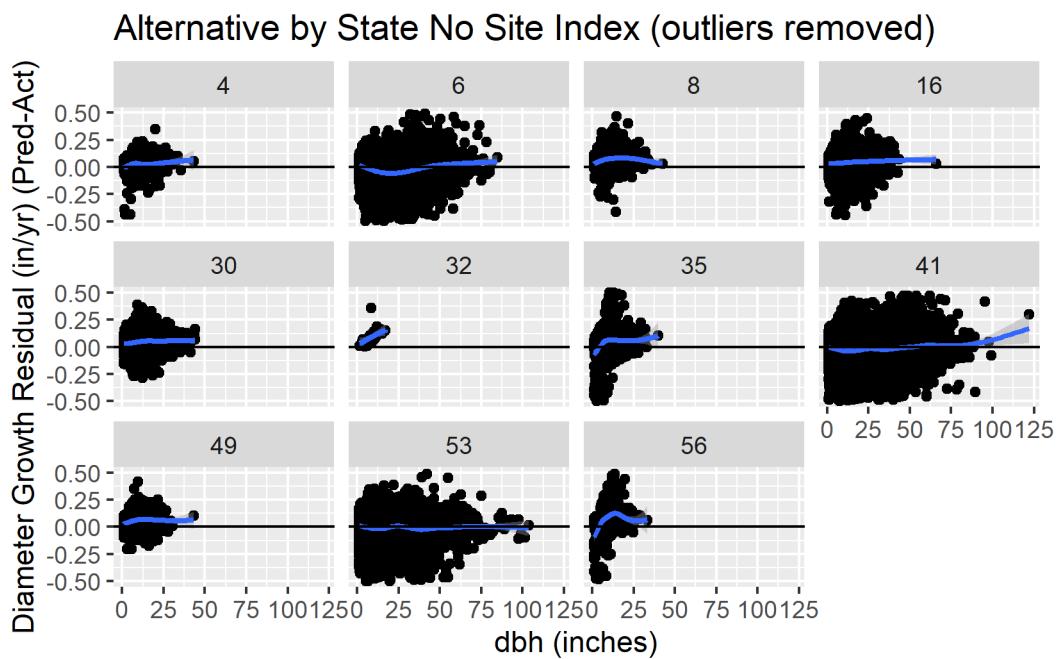




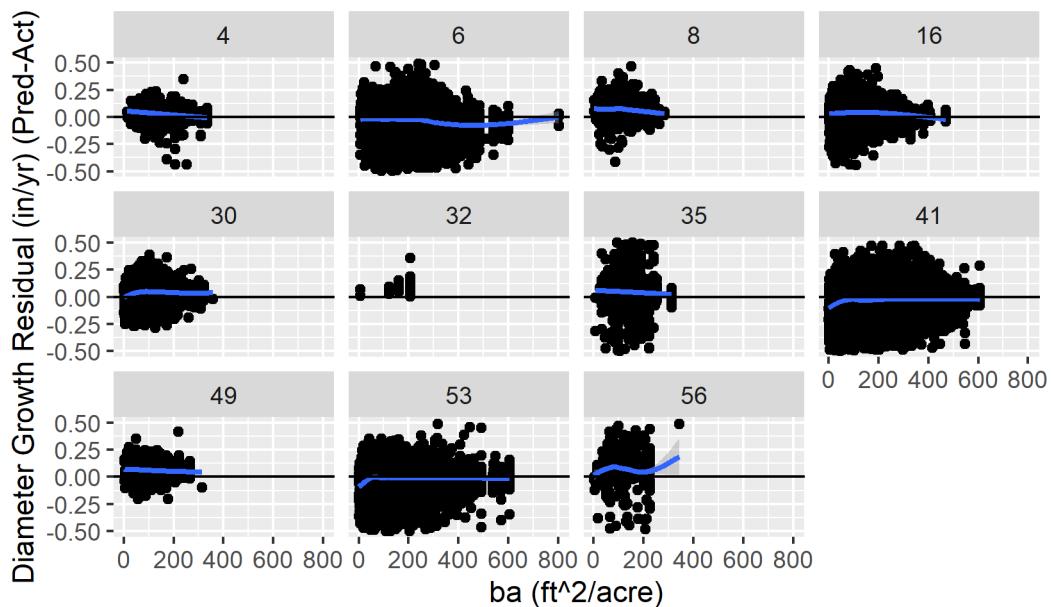




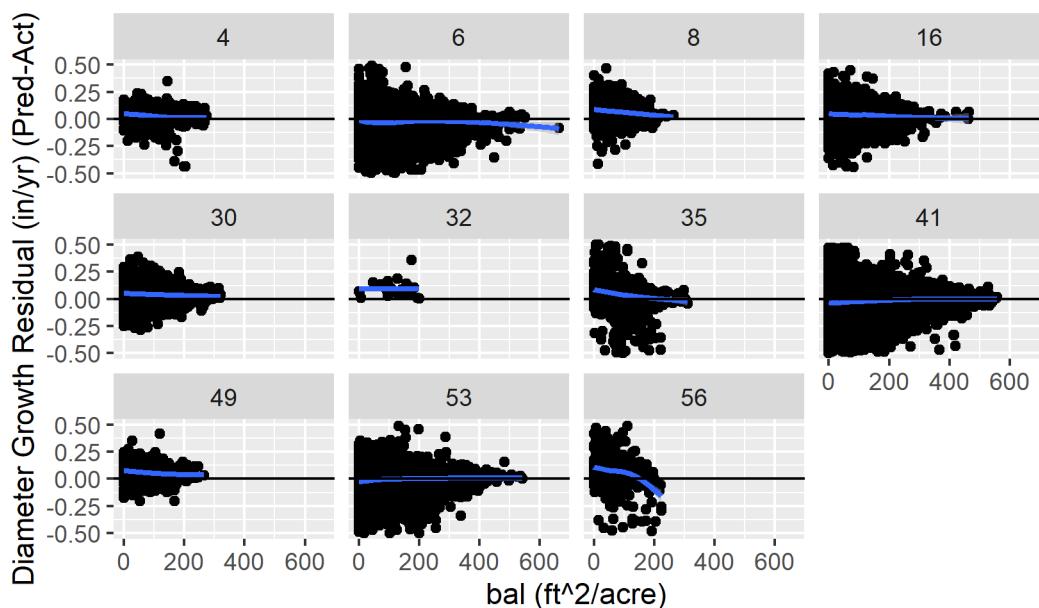
#### Residual Analysis for Equation 2



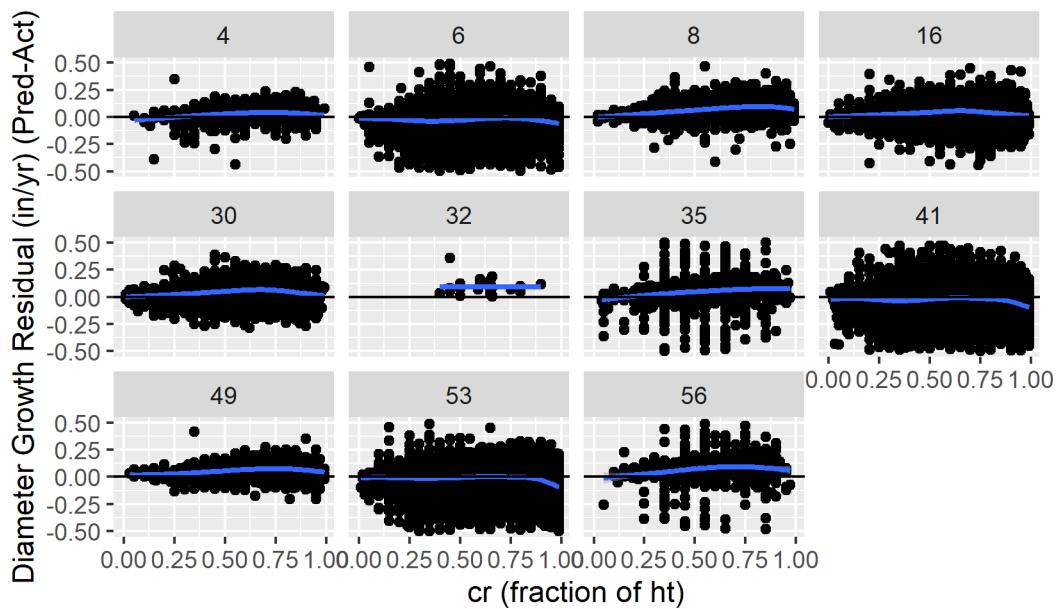
Alternative by State No Site Index (outliers removed)



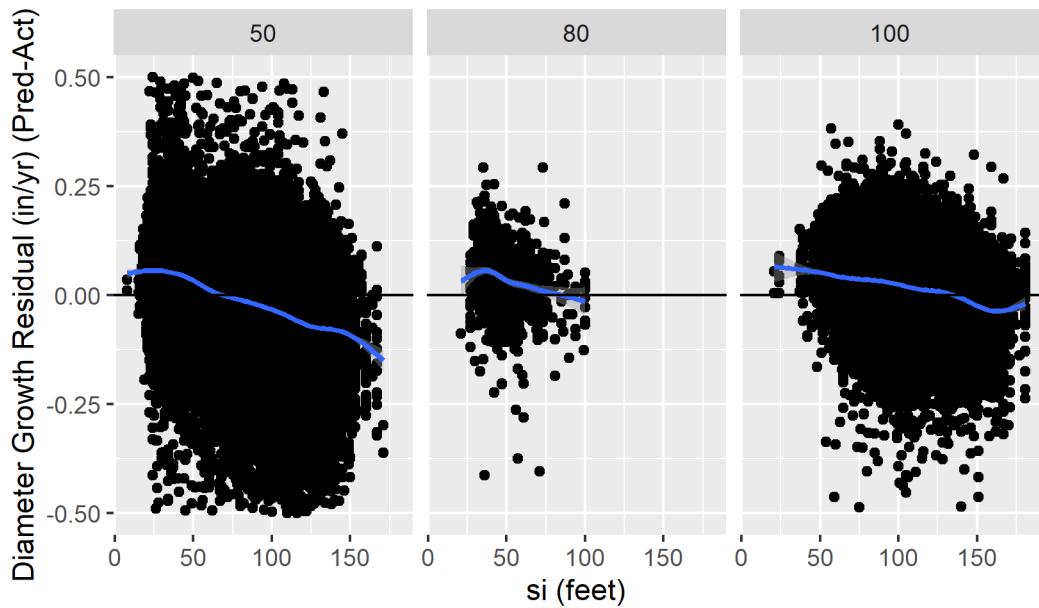
Alternative by State No Site Index (outliers removed)



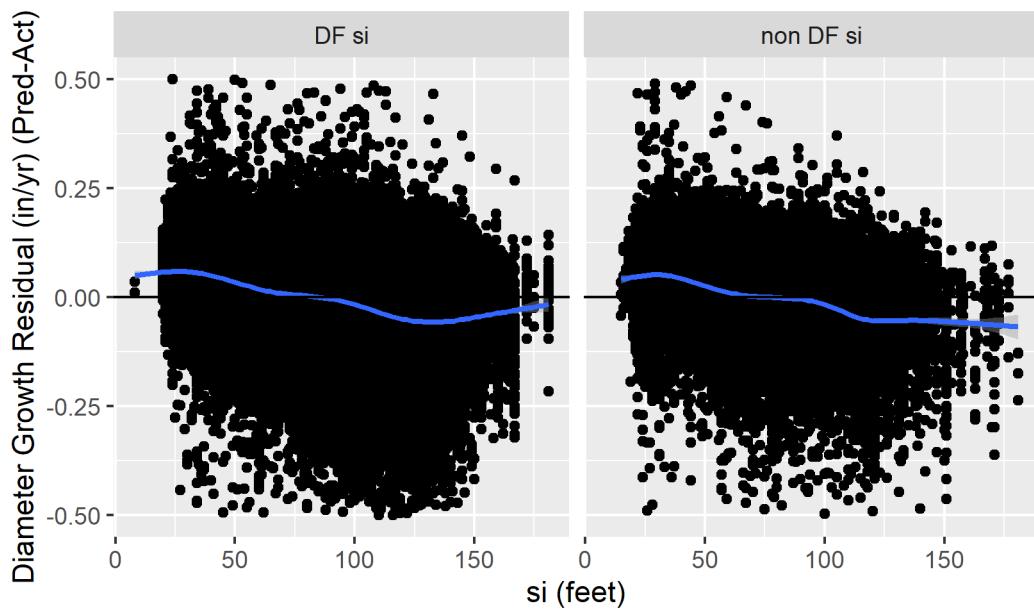
Alternative by State No Site Index (outliers removed)



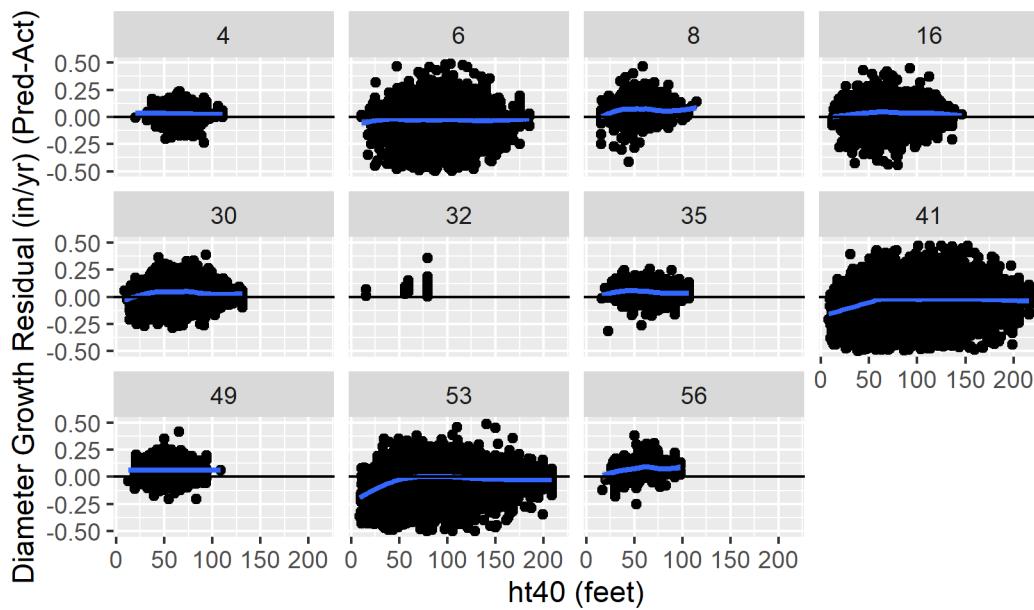
Alternative by State No Site Index by si Base Age (outliers re

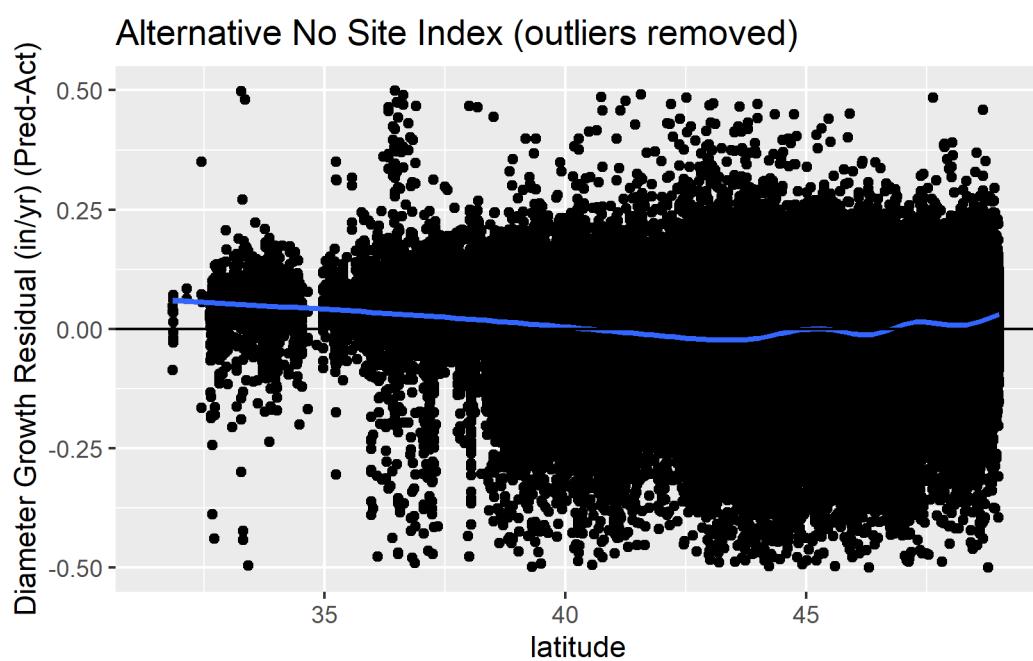
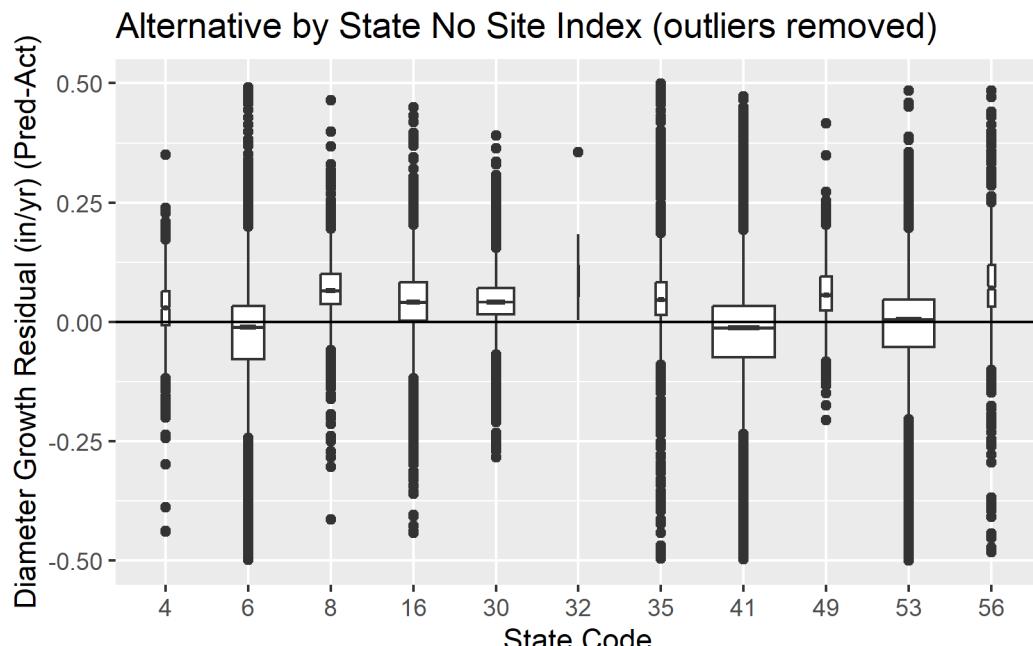


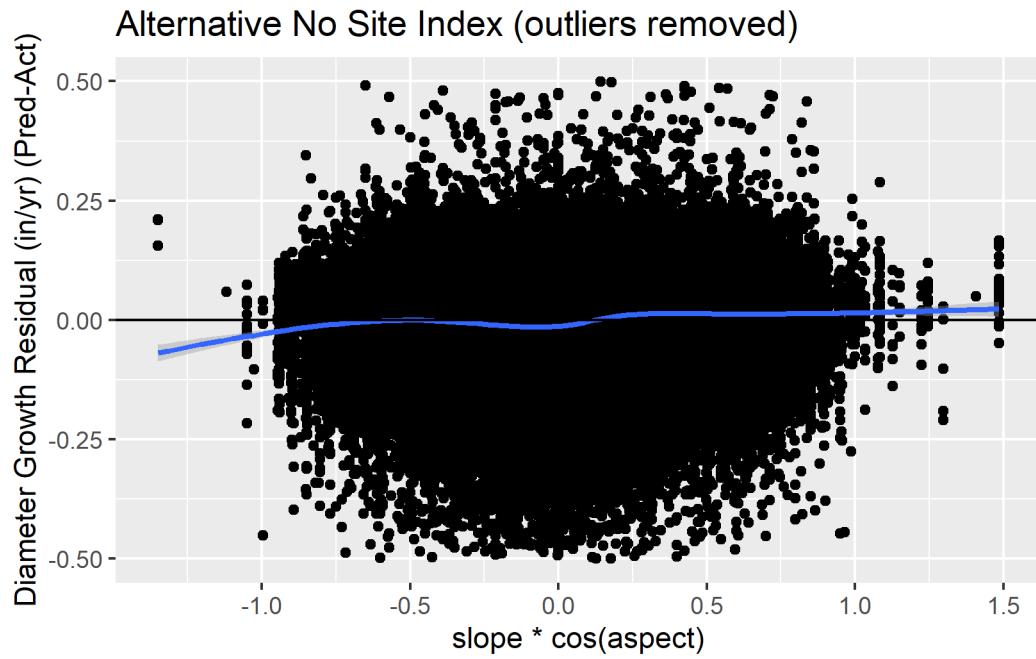
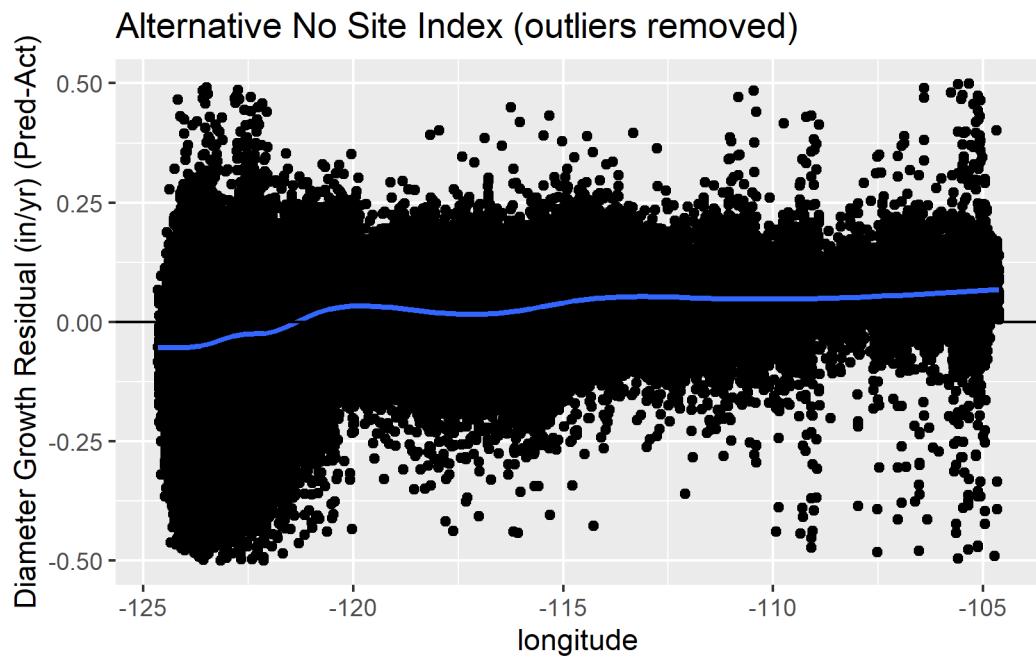
Alternative by State No Site Index by si Species (outliers removed)



Alternative by State No Site Index (outliers removed)







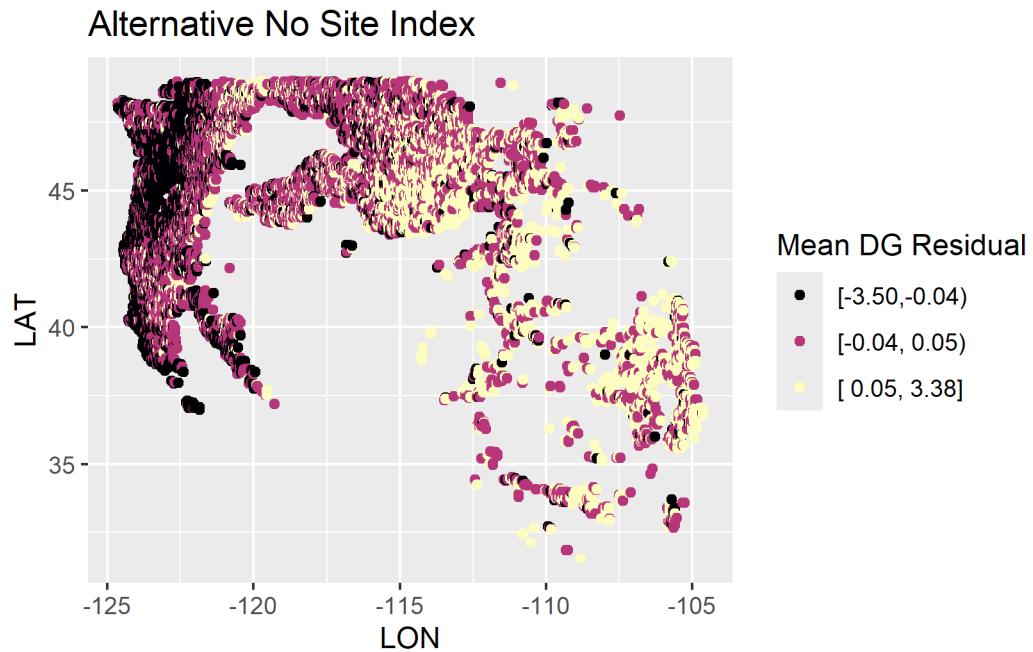
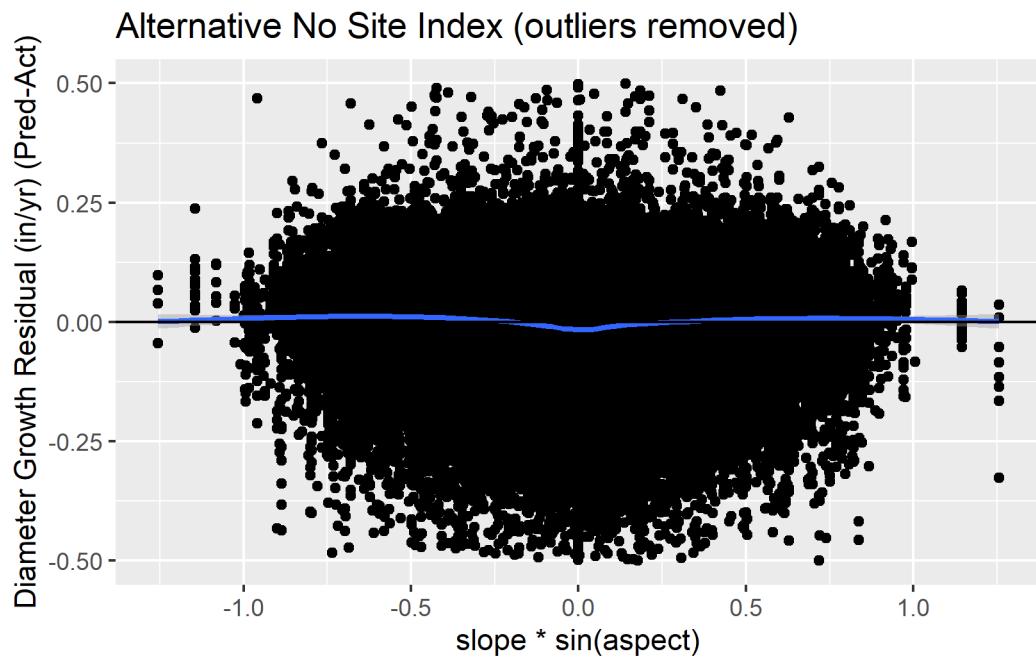


Table 3: Independent Variables for One Inch dbh Trees

Variable	N	Mean	Std. Dev.	Min	Pctl. 25	Pctl. 75	Max
ba	471	113	70	0.41	61	150	472
bal	471	112	70	0	61	150	472
ht	471	9.3	1.9	5	8	10	20
cr	471	0.55	0.25	0.05	0.35	0.8	0.99
si	471	70	30	20	44	90	159

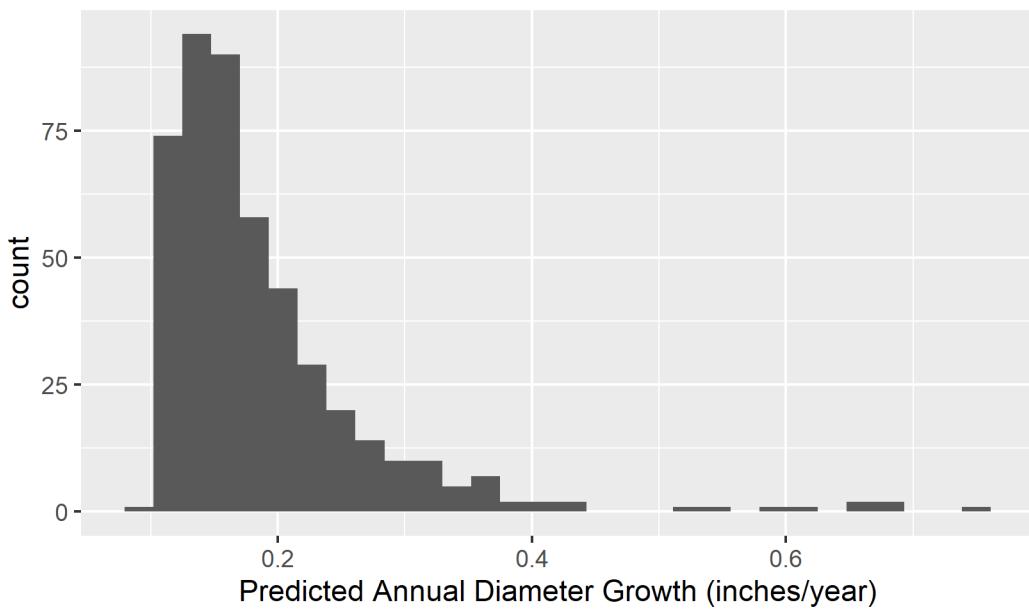
## Discussion

Removing `si` degrades the fit significantly. When `si` in its random variable form is used, the bias with the height of the largest 40 trees per acre (`ht40`) is reduced for small values. There is an average bias by state for both equations, but interestingly, the bias for states with small sample sizes are uniformly over-predicted in the equation without site index. We might expect this given that these states are on the edges of Douglas-fir's range.

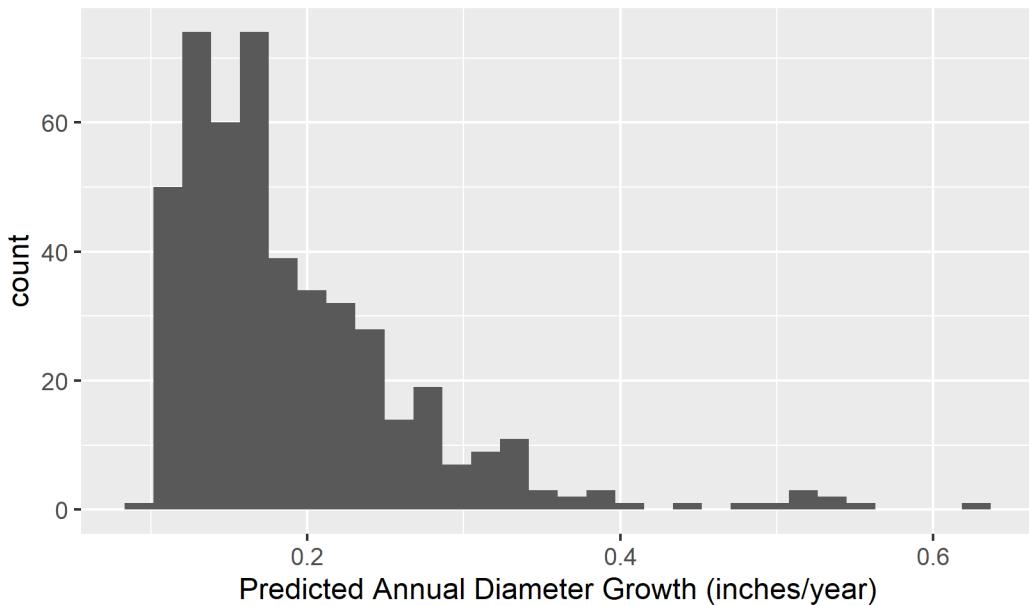
Equation 1 residual graphs show that residual trends with latitude and longitude are eliminated. This suggests that the admixture of site index values in the data set are doing a reasonable job at localization, albeit in a non-repeatable manner. It also suggests that latitude and longitude could be incorporated into Equation 2.

### Equation Behavior for Very Small Trees

Equation 1 Predictions for Trees with One Inch dbh Trees



Equation 2 Predictions for Trees with One Inch dbh Trees



## Alternative Model with Latitude and Longitude

First attempt at adding latitude and longitude to the model.

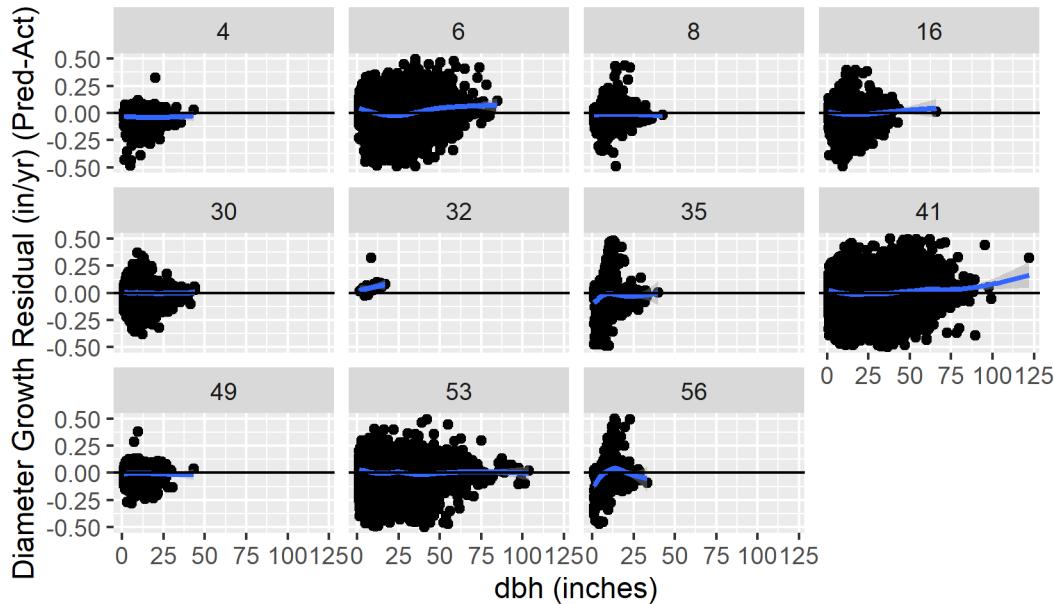
$$\Delta dbh = e^{(\beta_0 + \beta_1 \log(\frac{(dbh+1)^2}{(cr*ht+1)^{\beta_4}}) + \beta_2 \frac{bal^{\beta_5}}{dbh+2.7} + \beta_6 lat + \beta_7 long)} \quad (3)$$

The fit statistics for Equation 3 are:

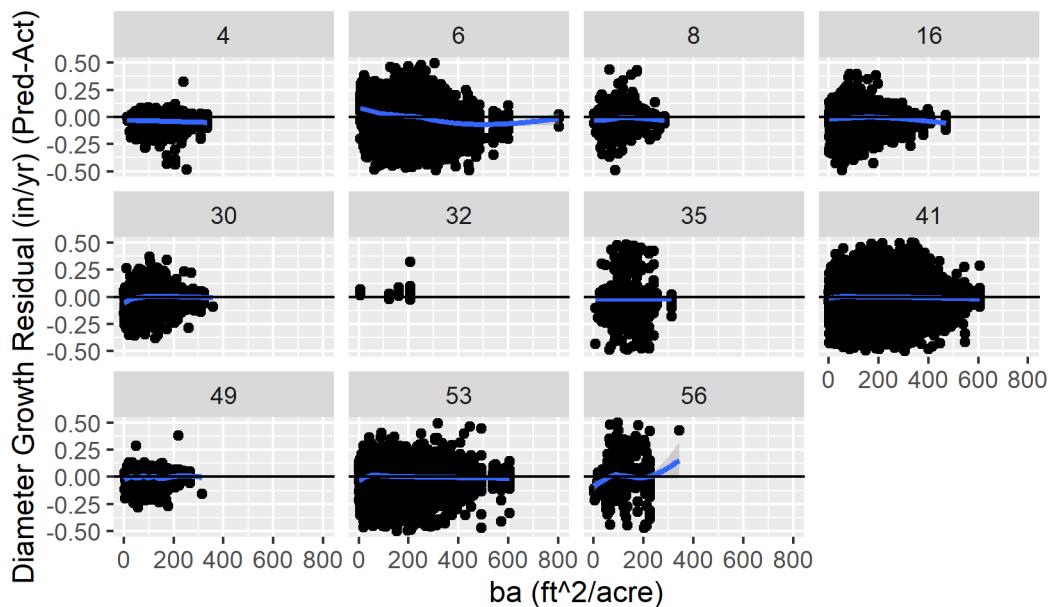
	Coef.	Std. error	t-stat.	p
B0	-12.7793177	0.0897341	-142.41312	0
B1	-0.6463581	0.0032656	-197.92700	0
B2	-0.1143400	0.0062375	-18.33119	0
B4	1.7824009	0.0075197	237.03110	0
B5	0.5959845	0.0096209	61.94655	0
B6	-0.0123034	0.0007433	-16.55241	0
B7	-0.0933456	0.0006392	-146.02864	0

Residual Standard Error: 1.05964754452461 on 186302 degrees of freedom, AIC: 550319.2

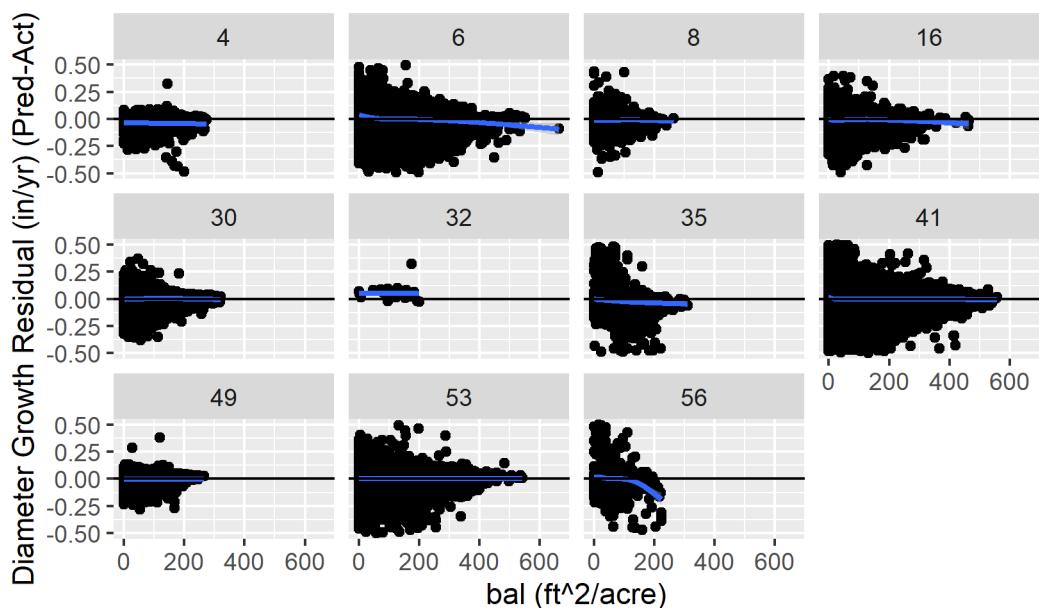
Alternative by State with Lat / Long (outliers removed)



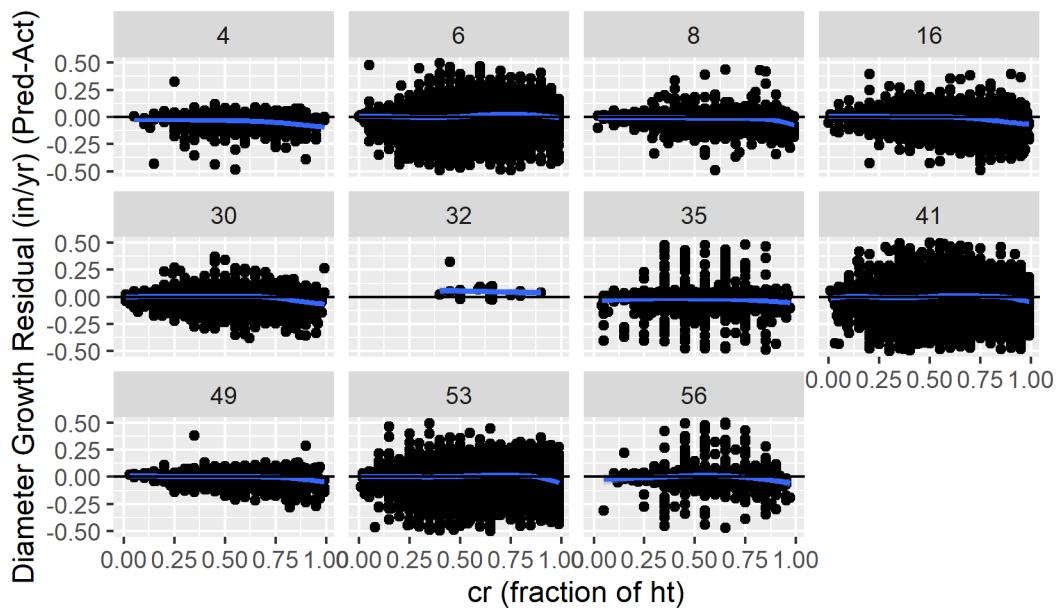
Alternative by State with Lat / Long (outliers removed)



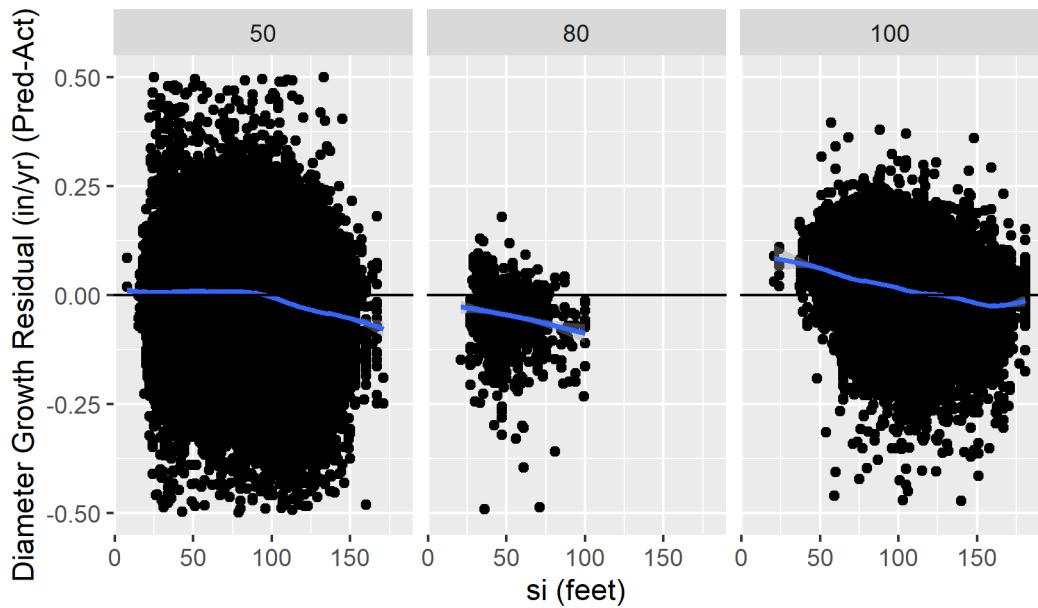
Alternative by State with Lat / Long (outliers removed)



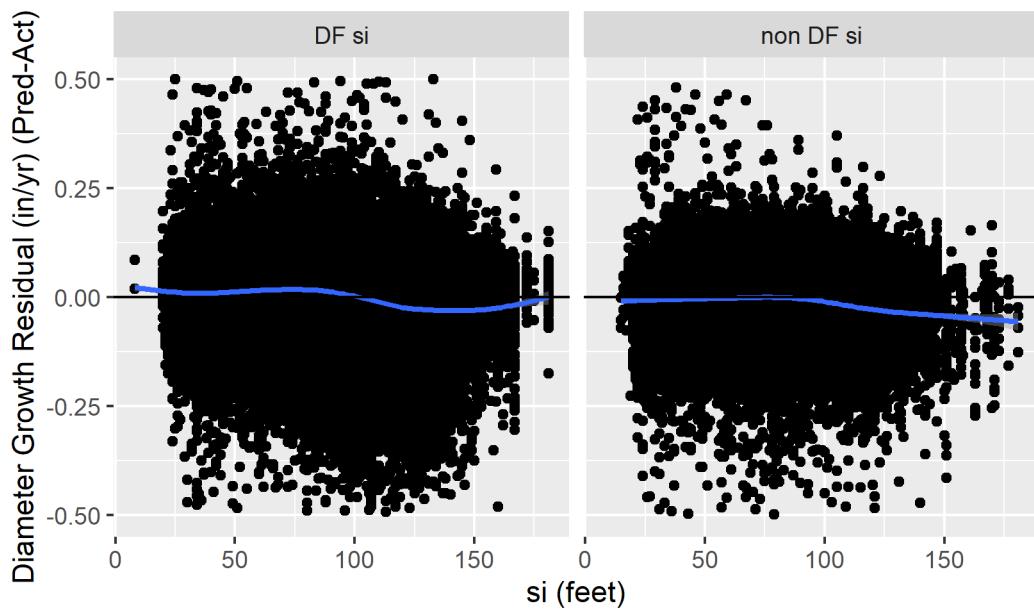
Alternative by State with Lat / Long (outliers removed)



Alternative by State with Lat / Long by si Base Age (outliers removed)



Alternative by State with Lat / Long by si Species (outliers removed)



Alternative by State with Lat / Long (outliers removed)

