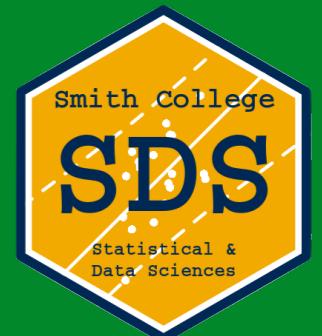


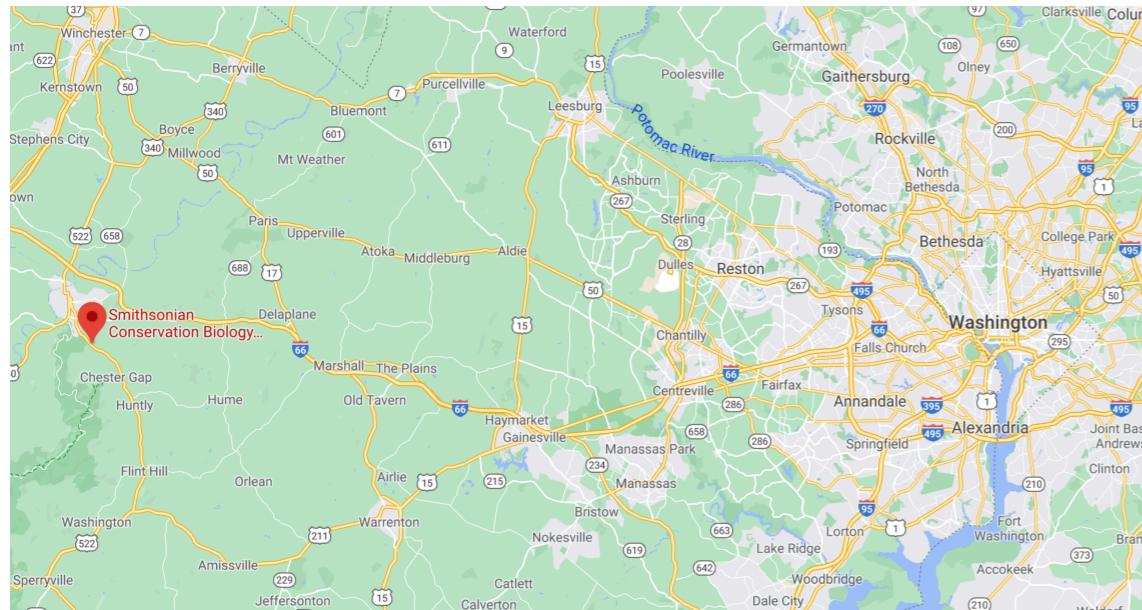
# Fusing disparate measurement data for forecasting the growth of trees via Hidden Markov Models



Prof. Albert Y. Kim [akim04@smith.edu](mailto:akim04@smith.edu)  
US Forest Service  
Wednesday, January 12 2022

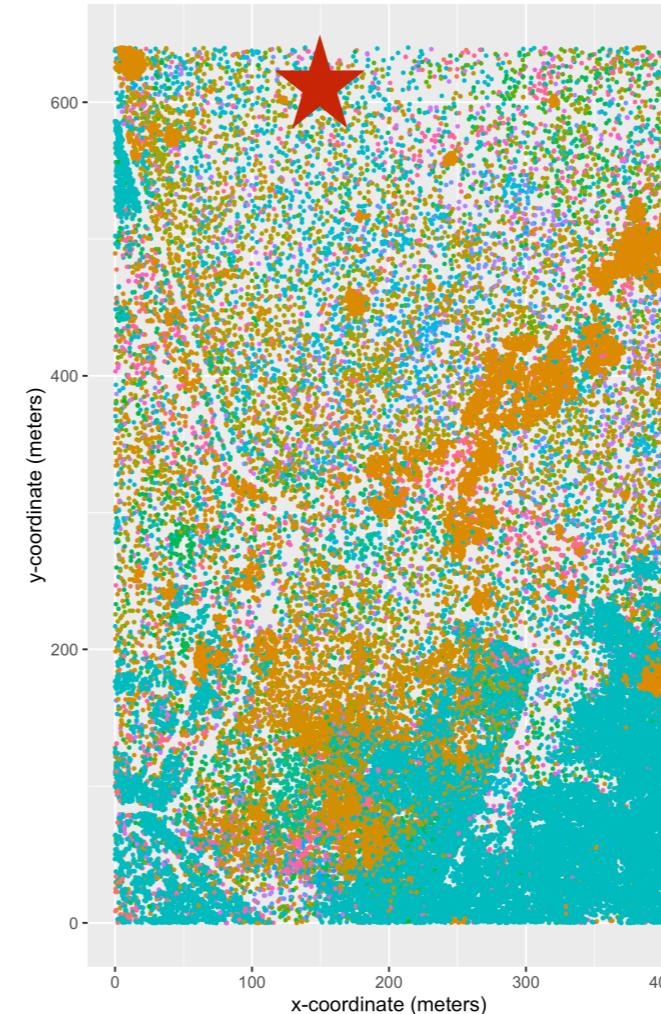


# One particular tulip poplar at SCBI



25.6 ha = 35.85 soccer fields

Census 2018: 72,555 cataloged trees



species
acne
acpl
acru
acsp
aial
amar
astr
beth
caca
caco
cade
cagl
caovl
casp
cato
ceca
ceoc
chvi
coal
coam
cofl
crpr
crsp
divi
elum
eual
fagr
fram
frni
frpe
frsp
havi
ilve
juci
juni
juvi
libe
litu
loma
nysi
pato
pipu
pist
pivi
ploc
prav
prpe
prse
prsp
qual
quco
qufa
qumi
qupr
quru
qusp
quve
rhpe
romu
rops
rual
rupe
rugh
saca
tiam
ulam
ulru
ulsp
unk
viac
virp
vire



Tag 082422

# Data on GitHub

Searched or jump to... / Pull requests Issues Marketplace Explore

SCBI-ForestGEO / SCBI-ForestGEO-Data Public

Unwatch 5 Fork 6 Star 6

Code Issues 8 Pull requests Actions Projects Wiki Security Insights

master 3 branches 4 tags Go to file Add file Code

teixeirak	Update README.md	23d07b6 on Dec 7, 2021	857 commits
R_scripts	Update visualize_tree_core_data.m	2 years ago	
leaf phenology	Revert "Merge branch 'master' of https://github.com/SCBI..."	2 years ago	
plot disturbances	Update SCBI_plot_disturbance_events.csv	11 months ago	
seedlings	Reloaded cleaned files.	16 months ago	
soils	Update README.md	3 months ago	
spatial_data	move to SCBImortality repo	7 months ago	
species traits	Update README.md	15 months ago	
species_lists	Update insects_pathogens.csv	5 months ago	
summary_data	Update README.md	3 years ago	
tree_cores	Update README.md	last month	
tree_dimensions	Revert "Merge branch 'master' of https://github.com/SCBI..."	2 years ago	
tree_main_census	Fix sp	10 months ago	
tree_mortality	Update README.md	8 months ago	
.gitignore	update	3 years ago	
DESCRIPTION	trying again with removed hyphen	8 months ago	
README.md	Update README.md	3 months ago	
SCBI-ForestGEO-Data.Rproj	create	3 years ago	
_config.yml	Set theme jekyll-theme-cayman	3 years ago	
license.txt	Create license.txt	2 years ago	

README.md

**Smithsonian Conservation Biology Institute (SCBI) ForestGEO Data**

DOI 10.5281/zenodo.4070038

This is the public data portal for the SCBI ForestGEO plot, which points to archive locations for our

About

Public data repository of the SCBI ForestGEO plot

scbi-forestgeo.github.io/scbi-forestge...

Readme CC-BY-4.0 License 6 stars 5 watching 6 forks

Releases 4

first release with hydraulic tr... Latest on Oct 6, 2020 + 3 releases

Packages

No packages published Publish your first package

Contributors 10

Environments 1

github-pages Active

Languages

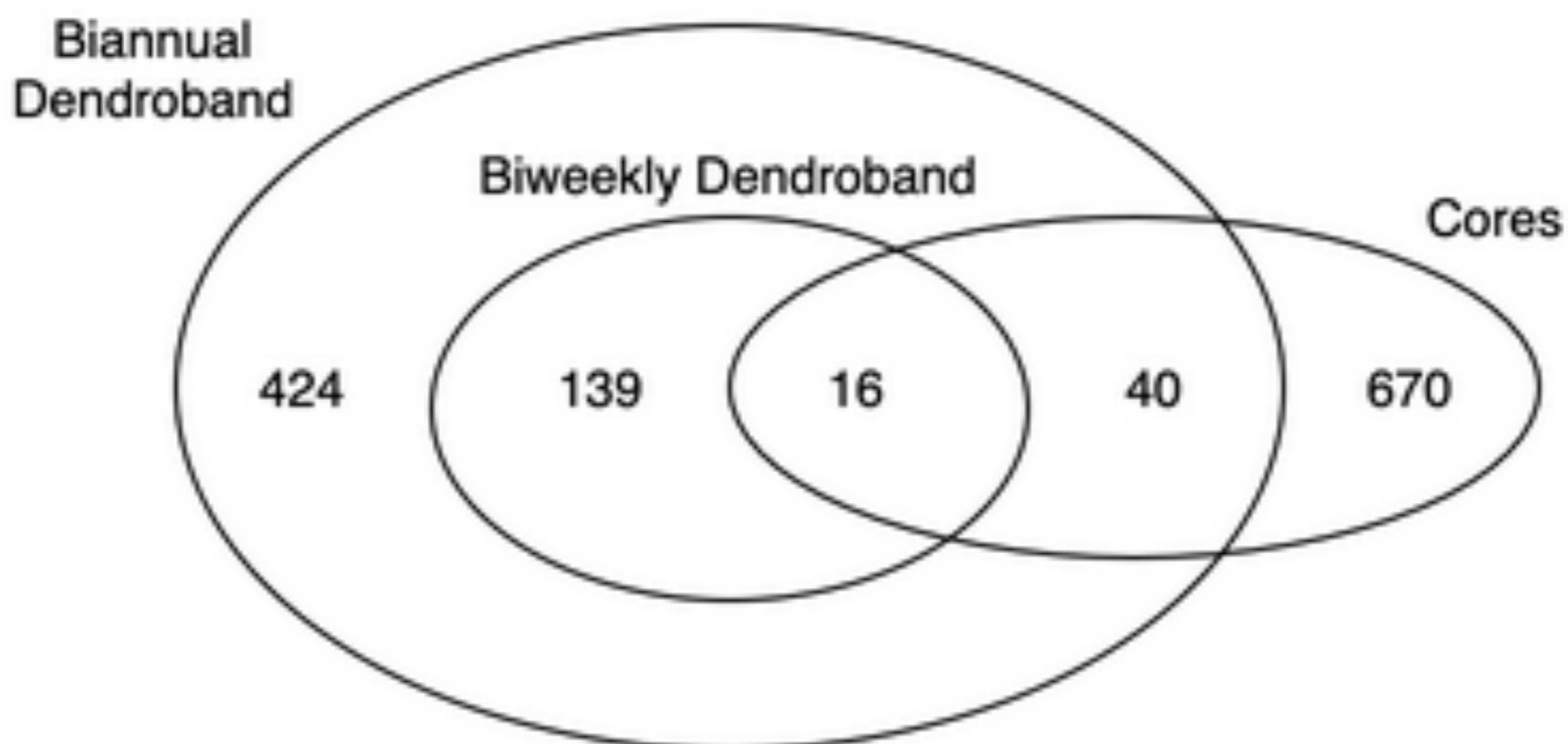
R 98.9% MATLAB 1.1%

# Comparison Chart

Data source	Measurement	Cost	Sources of Error?
	Census via tape	Diameter	Cheap Large variation in dbh 
	Tree coring	Ring width increment	Expensive Standardized, cores are dried, no bark effects
	Intraannual dendroband (every 2 weeks)	Diameter (baseline + gap size)	High setup, rapid follow-up Climate induced variation in bark & device (-'ve growth)
	Biannual dendroband (start & end of year)	Diameter (baseline + gap size)	High setup, rapid follow-up <b>Mid-year issues?</b> 

# Sample sizes

1289 stems total (all have census observations)

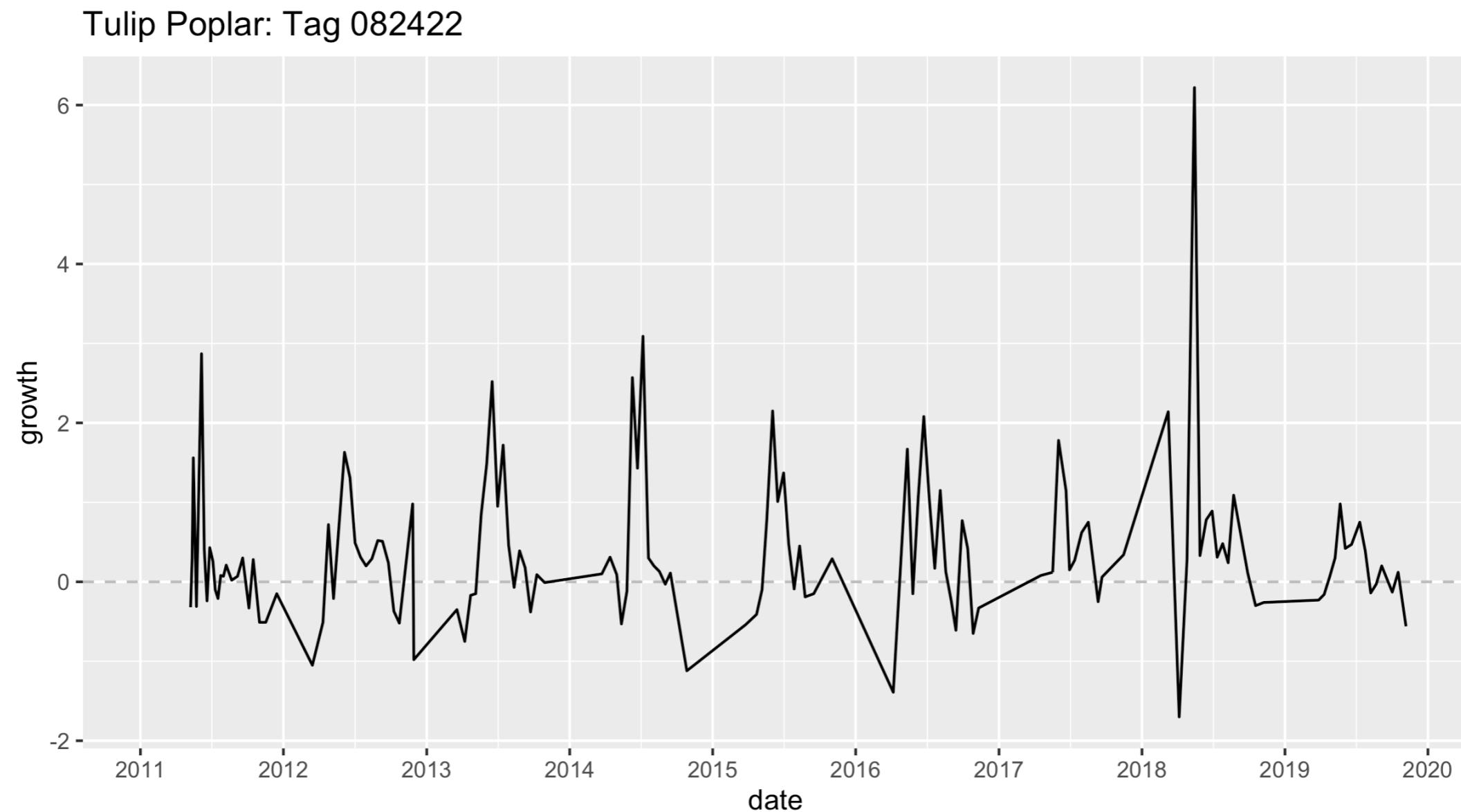


3 trees had 3 dendrobanded stems & 8 trees had 2 dendrobanded stems;  
we assume these stems are statistically independent

# Because dendroband measurements...



# ... are taken many times within-year



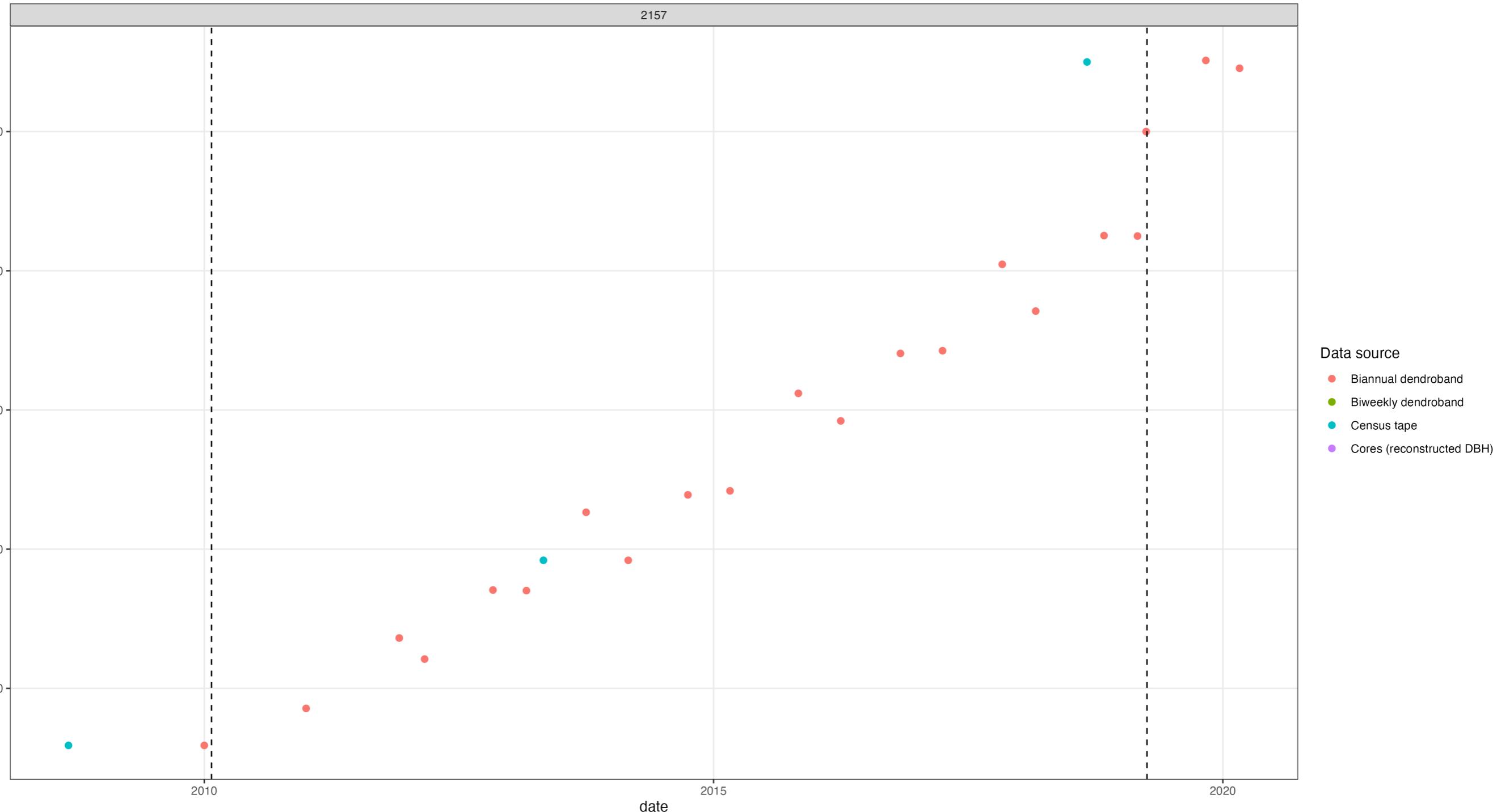
... we can link within-year variation in growth  
to within-year variation in **climate!**

# Our Data

# dbh trajectories: biannual dendro only

DBH for tag 12508: litu

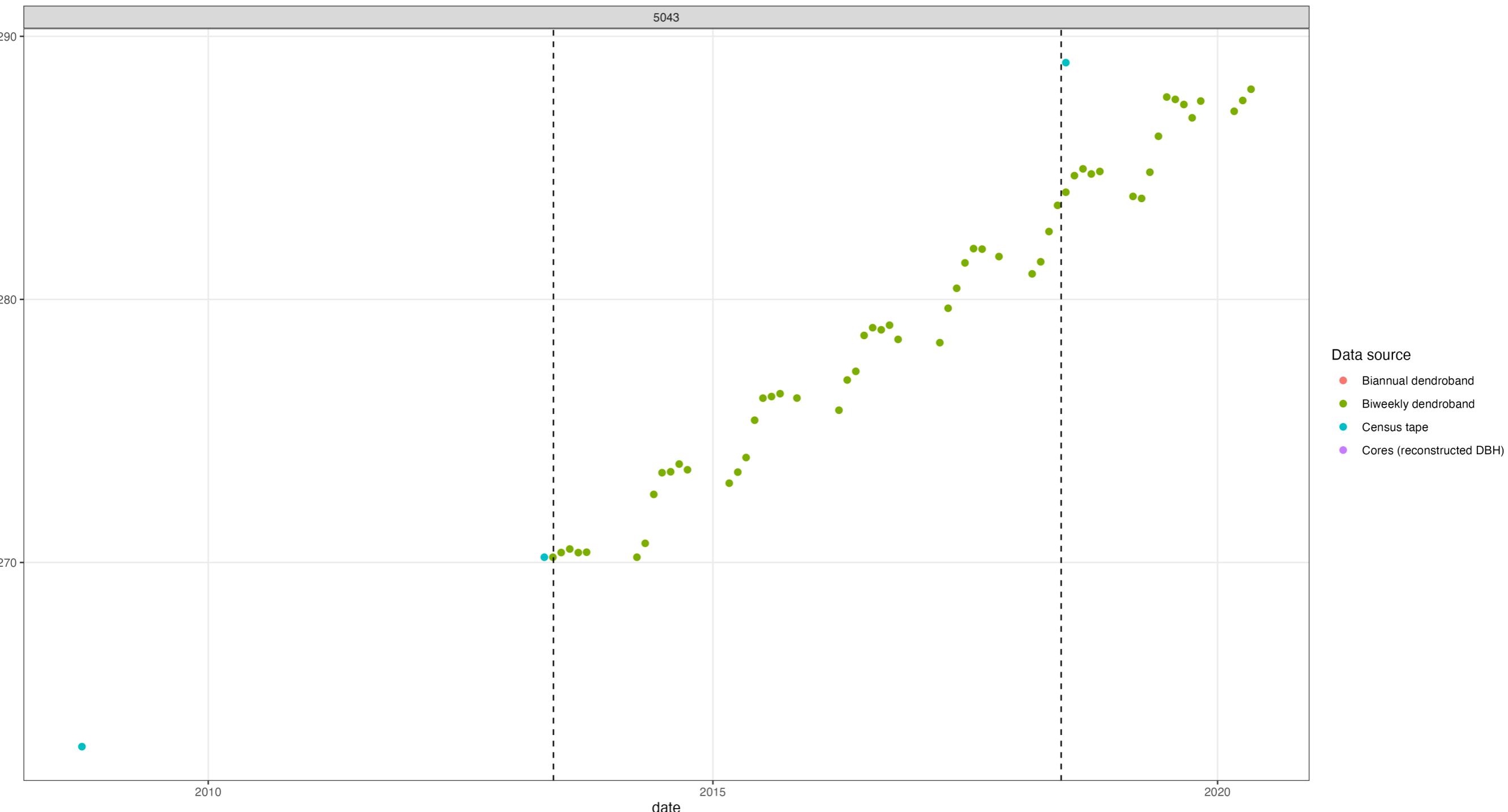
Points = observed data, dashed line = new dendroband installed



# dbh trajectories: biweekly

DBH for tag 40566: cagl

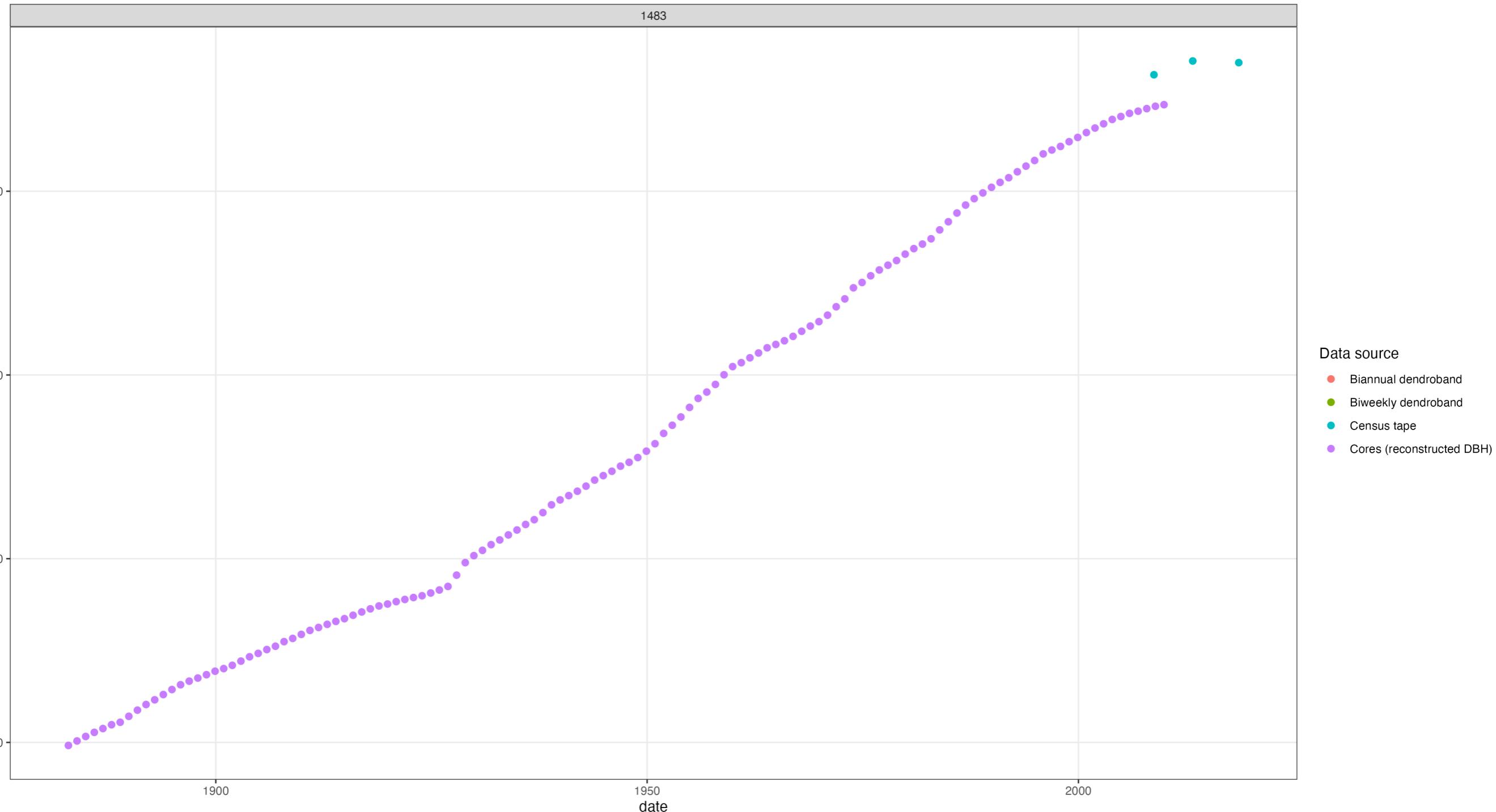
Points = observed data, dashed line = new dendroband installed



# dbh trajectories: cores (as dbh, not increments)

DBH for tag 10539: fram

Points = observed data, dashed line = new dendroband installed



# Goal



Can we fuse these disparate data sources into a single model to forecast the growth of trees?

# Hidden Markov Models

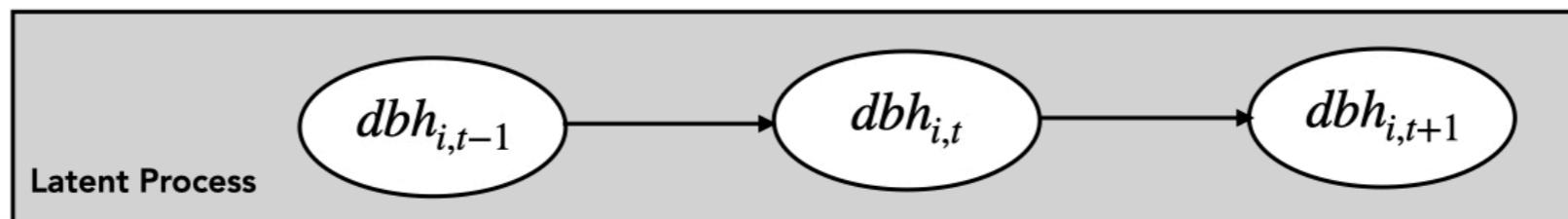
- Hidden: “Data fusion” via latent variables  
In our case: “true” dbh
- Main “process model” driving  $y_t$
- Markov property:  $y_t$  depends on  $y_{t-1}$
- Partition sources of error into those that
  - Are not of direct interest
  - Are “one and done” i.e. measurement error
  - Propagate when forecasting

# Other implementations

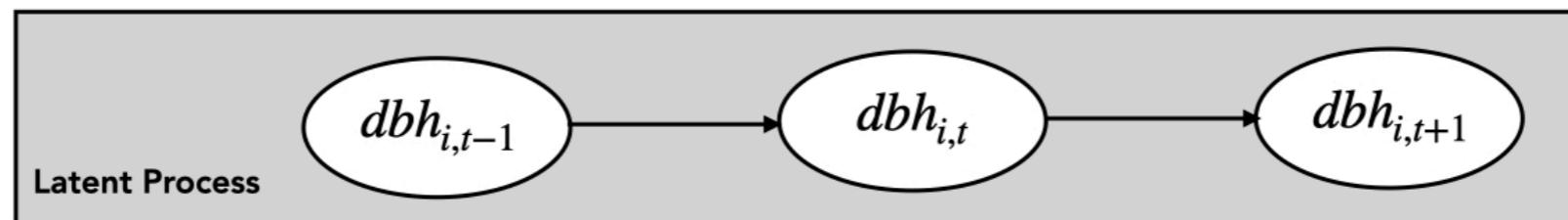
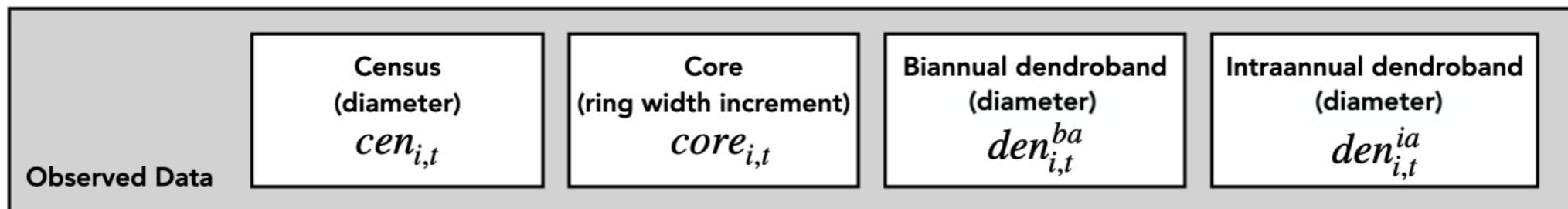
- [Clark 2007](#)
- [Schleip 2014](#)
- [Dawson 2021](#)
- [Heilman \(forthcoming\)](#)

# Spirit of Hidden Markov Model

# Latent process model

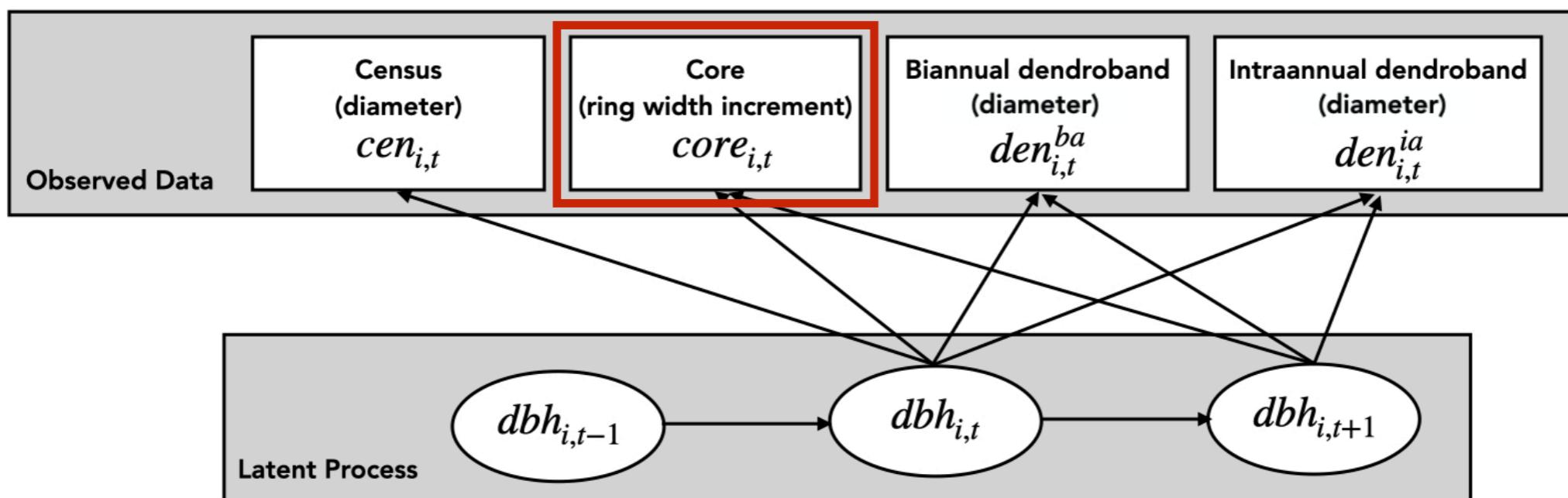


# + field collected data

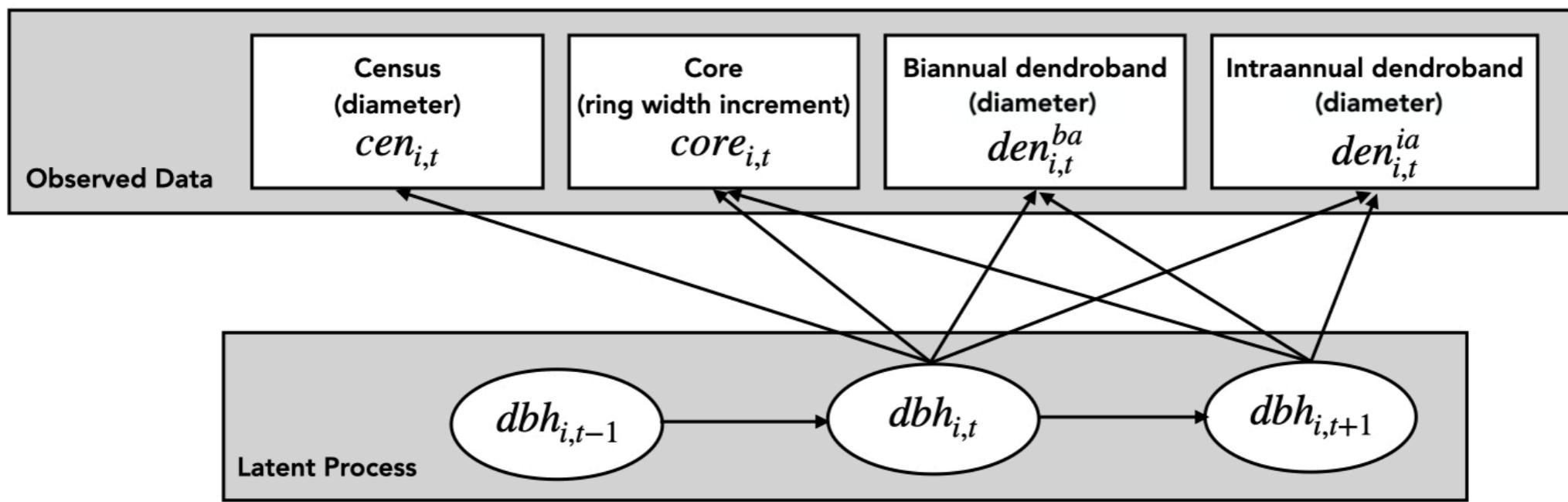


+ field collected data

$$\text{Increments} = dbh_{i,t+1} - dbh_{i,t}$$

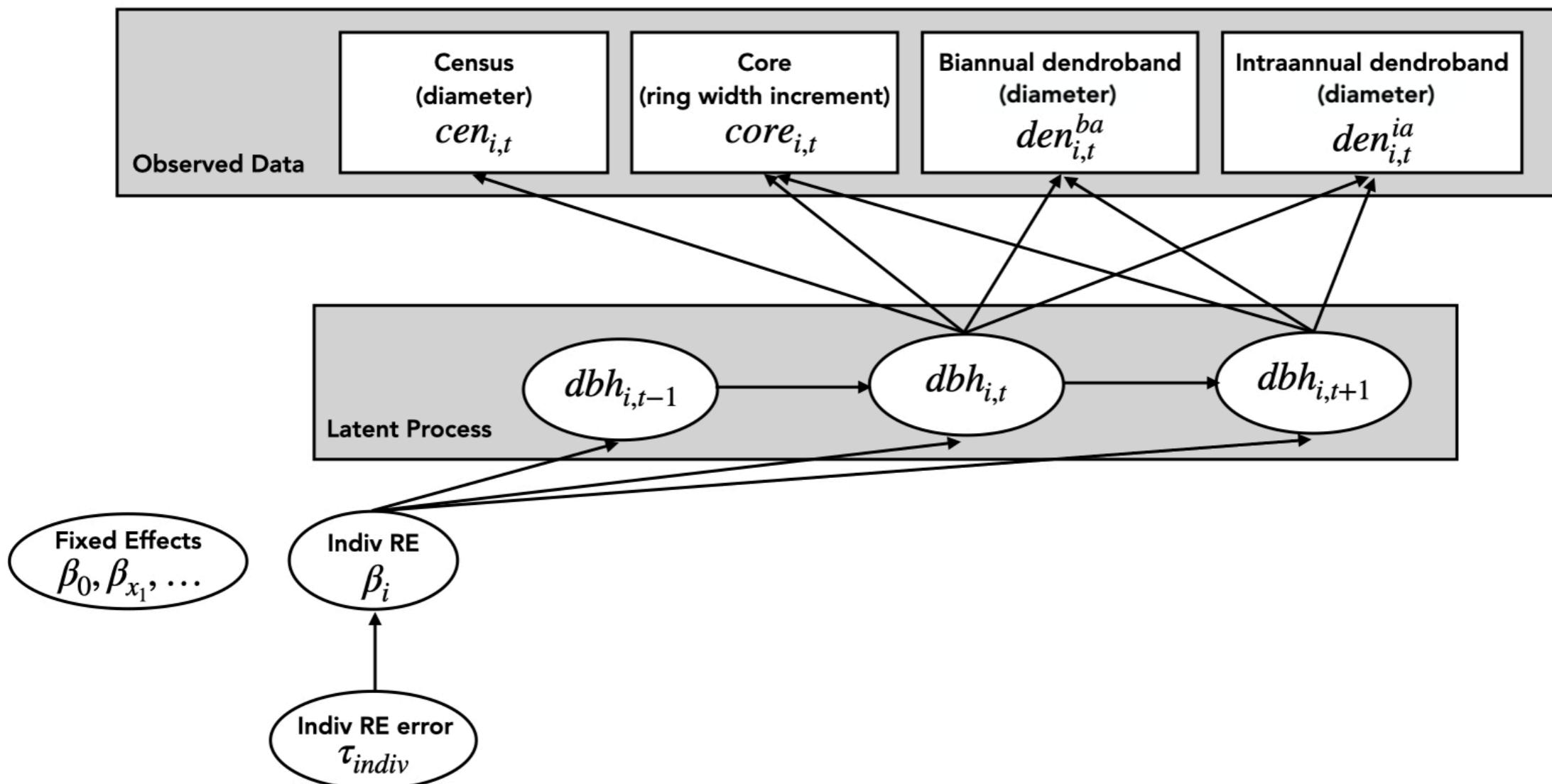


+ fixed effects (Ex: climate)

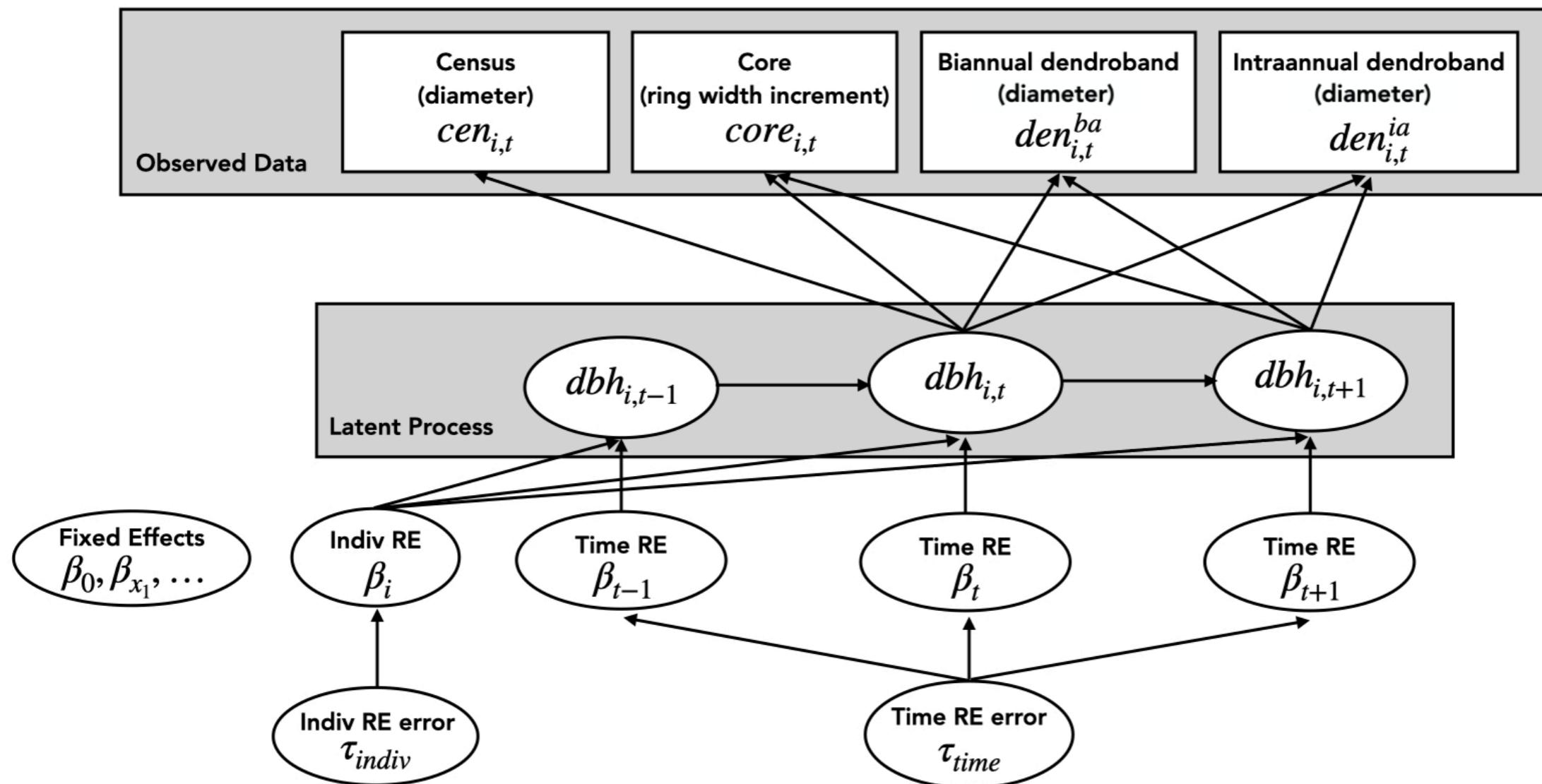


Fixed Effects  
 $\beta_0, \beta_{x_1}, \dots$

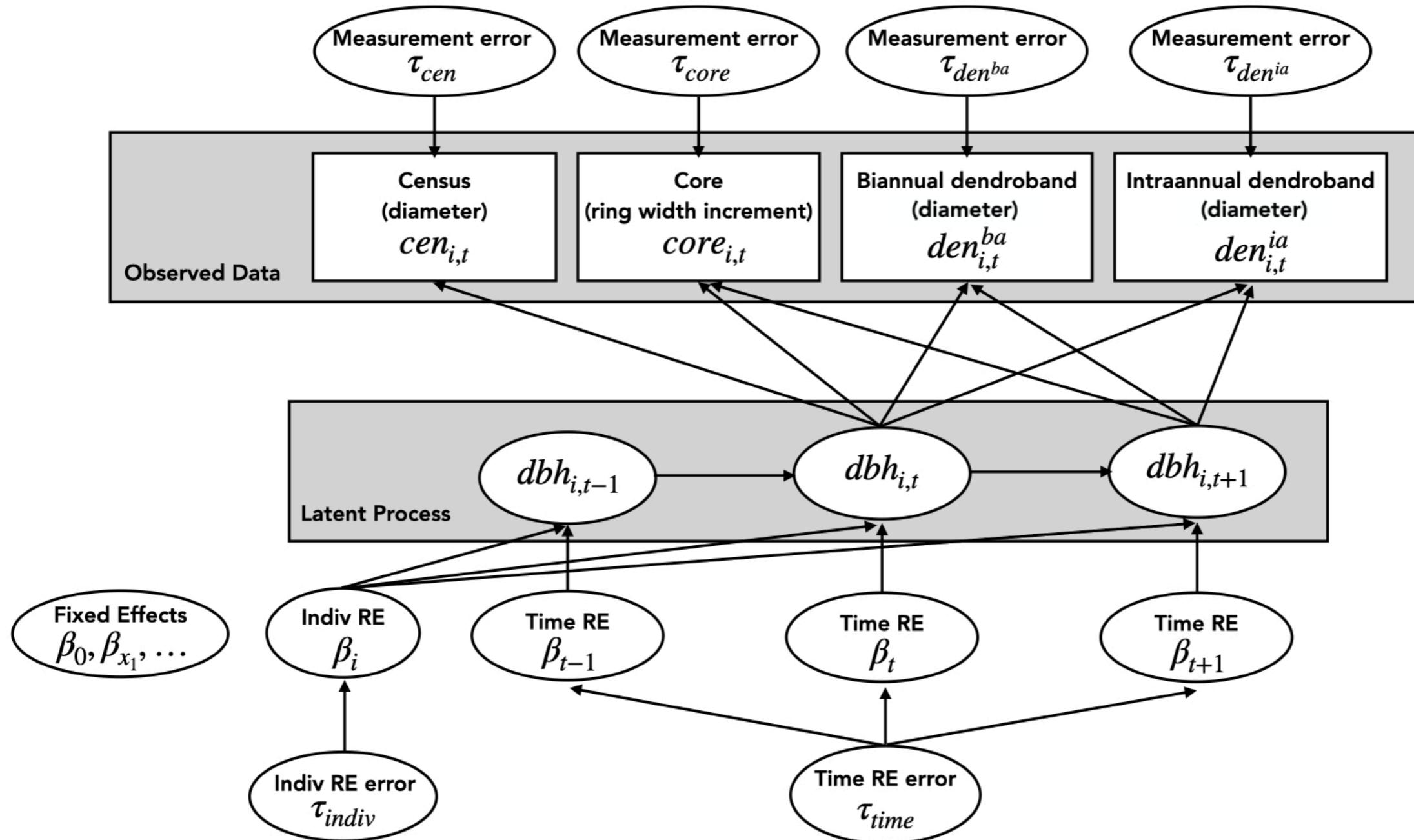
+ random effects for individual trees



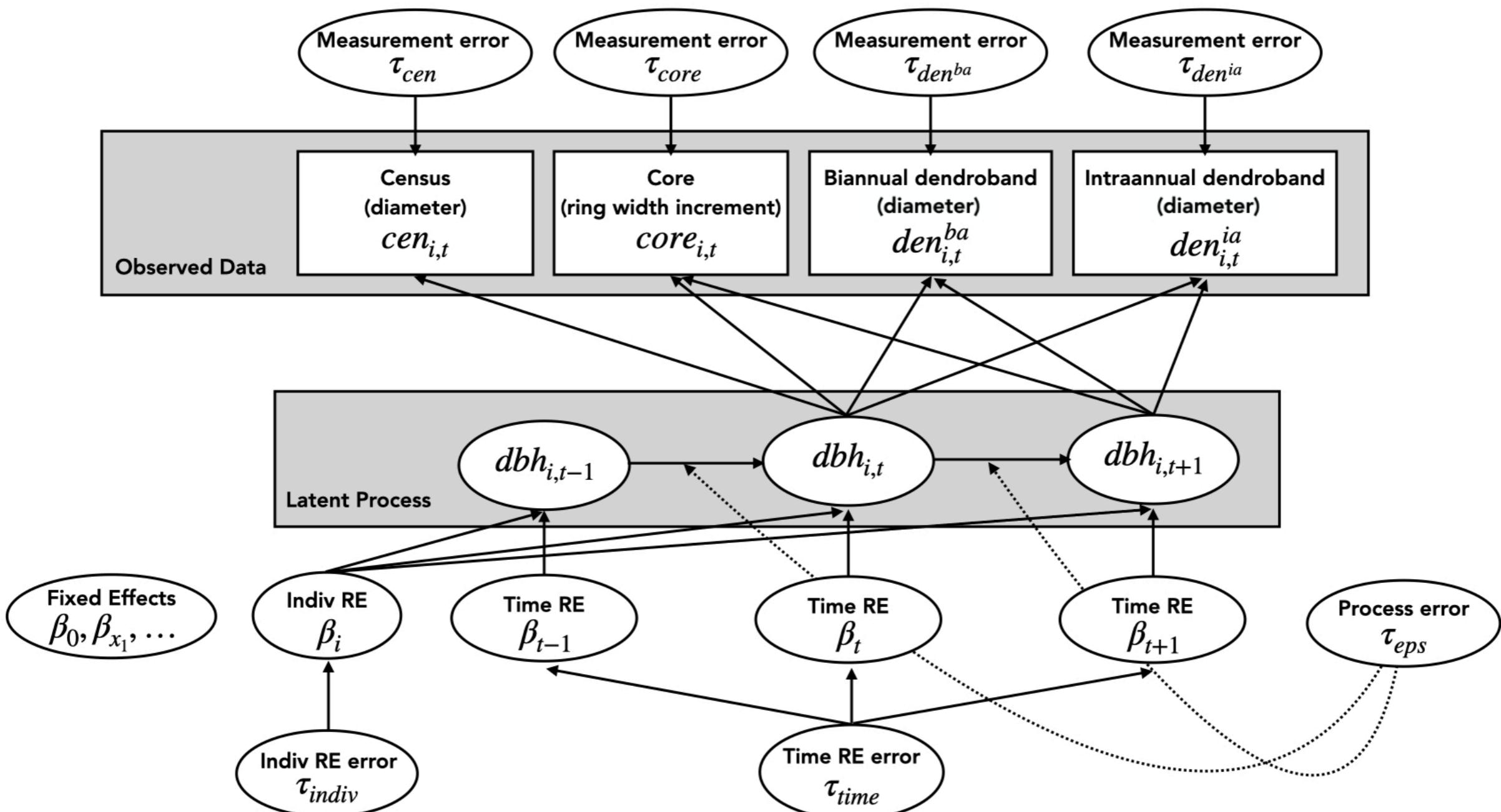
+ random effects for time points



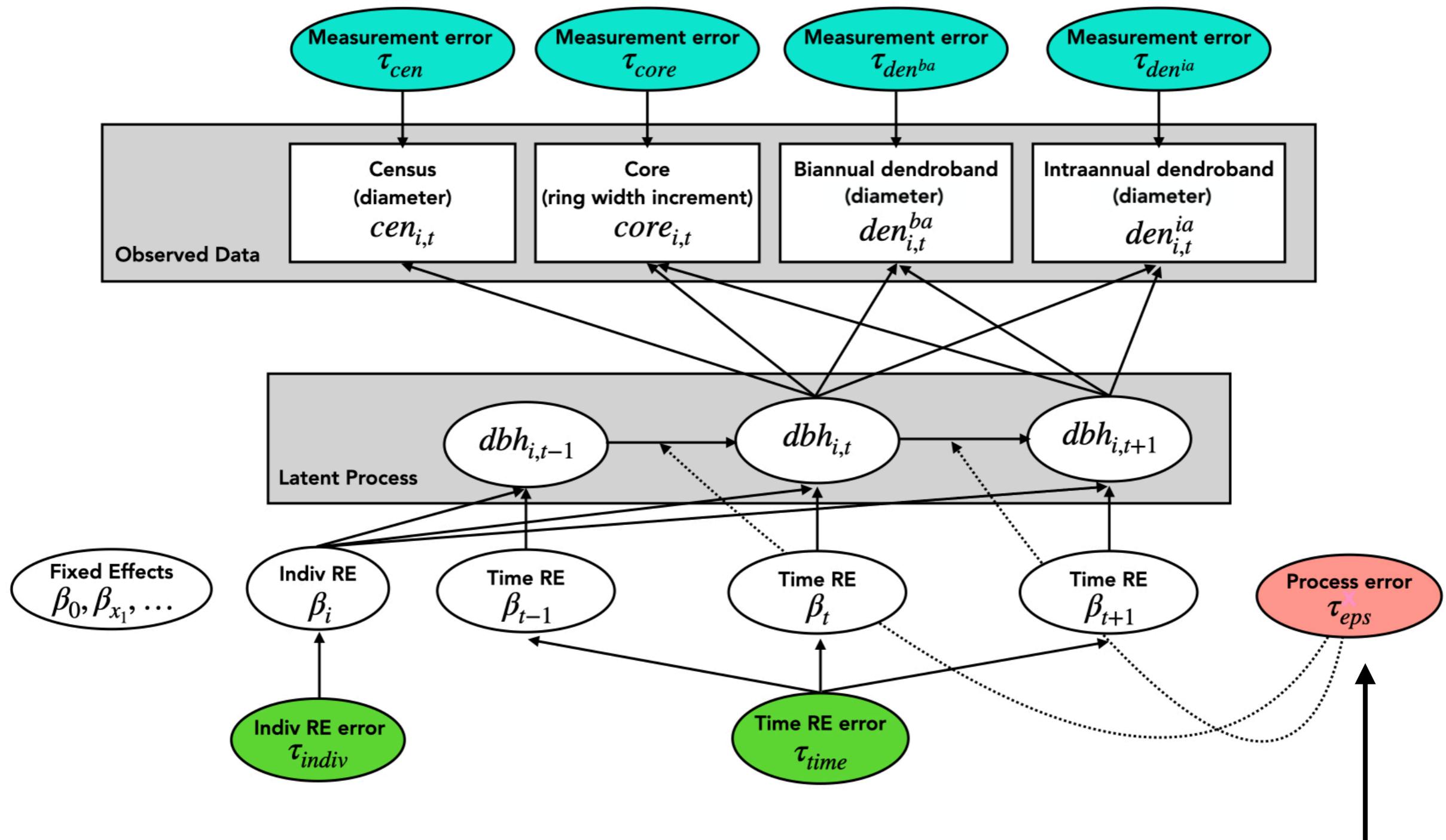
# + measurement error



+ process error



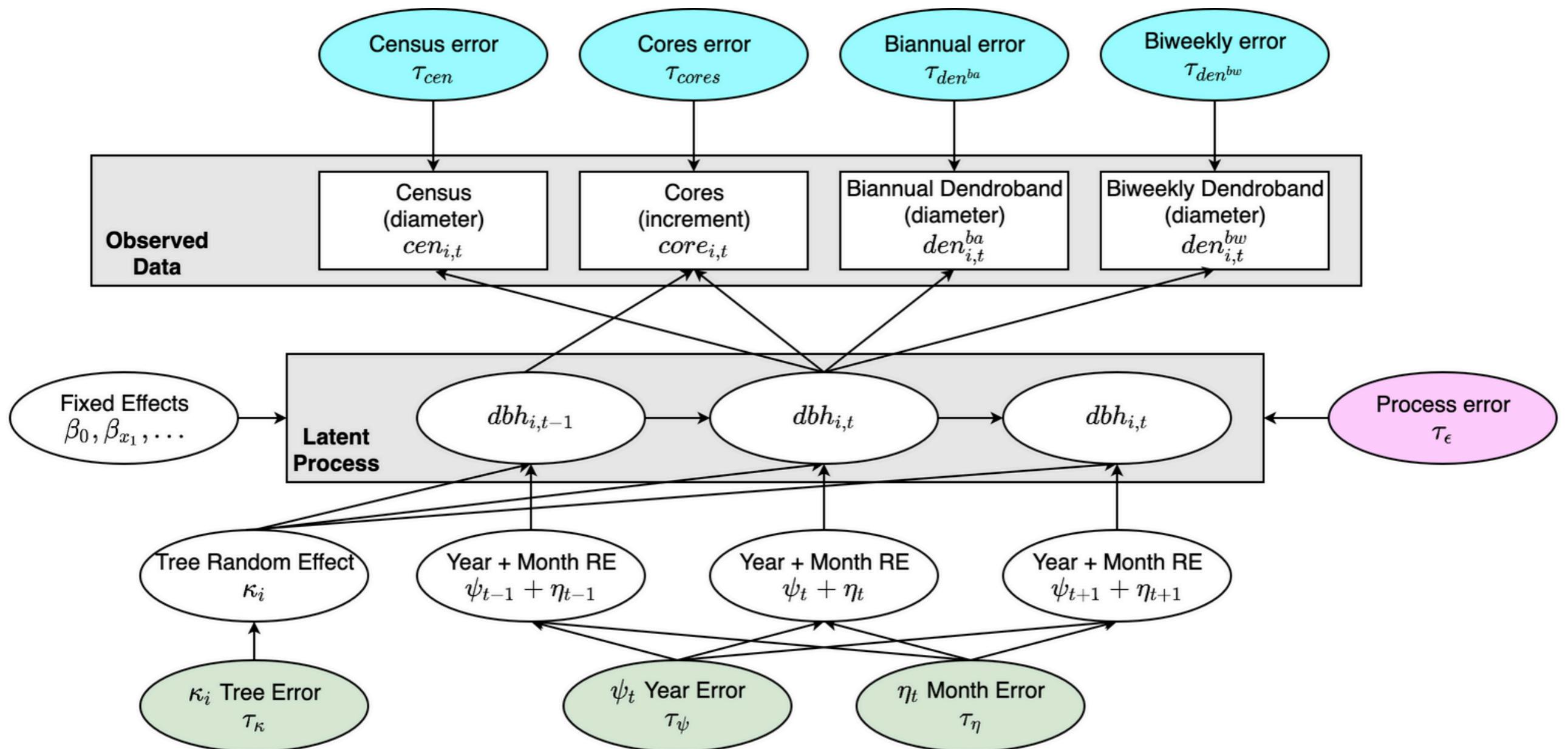
# = Hidden Markov Model



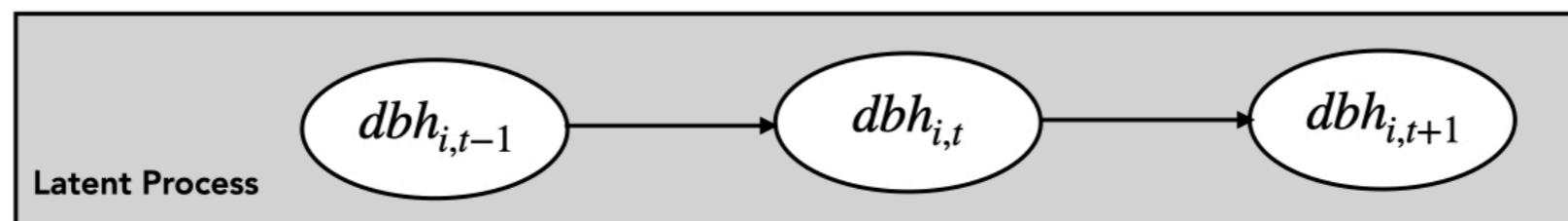
Moral: only error this propagates across time in forecasts

# Actual Model

# Model as of 2021/8/20



# Need to specify latent process model...



# Hidden Markov latent process model

$$dbh_{i,t} - dbh_{i,t-1} = \alpha + \beta \cdot dbh_{i,t-1} + \kappa_i + \psi_t + \eta_t + \epsilon_{i,t}$$

- $dbh_{i,t}$  : “True” latent dbh for individual  $i$  at time  $t$
- $\alpha$  : Baseline growth
- $\beta$  : Slope for DBH
- $\kappa_i$  : Individual tree  $i$  random effect
- $\psi_t$  and  $\eta_t$  : Year & month time  $t$  random effects
- $\epsilon_{i,t} \sim \text{Normal}(0, \sigma_\epsilon^2)$

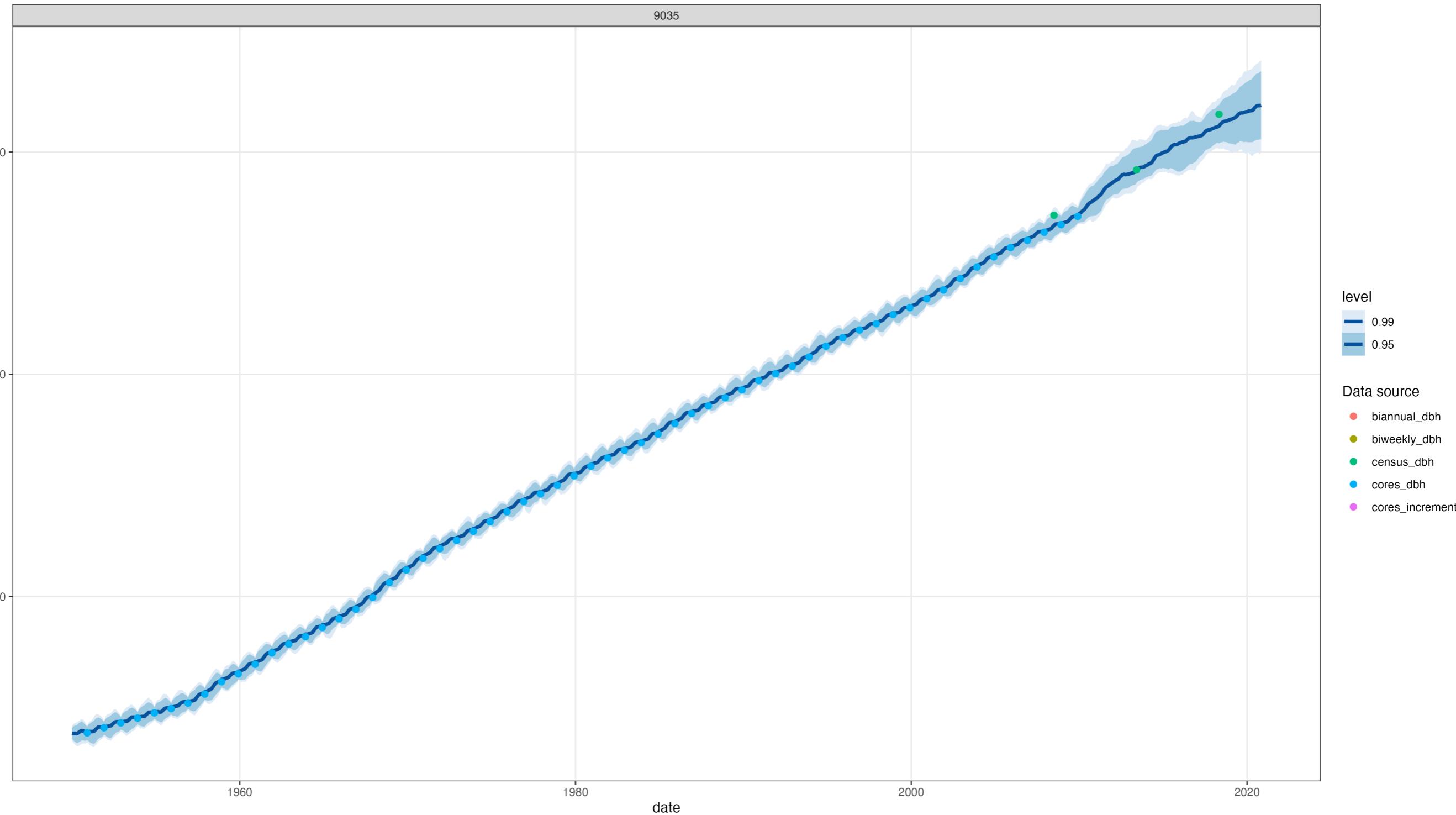
# Results

# MCMC specifications

- Only on random sample of 100 out of 1288 stems with core, dendro biweekly, or dendro biannual data only for values of cores > 1950
- 3 chains of 1k draws from posterior minus 10% burn-in
- 10min run-time
- Following results should be taken with a massive grain of salt 🤔❌🧂

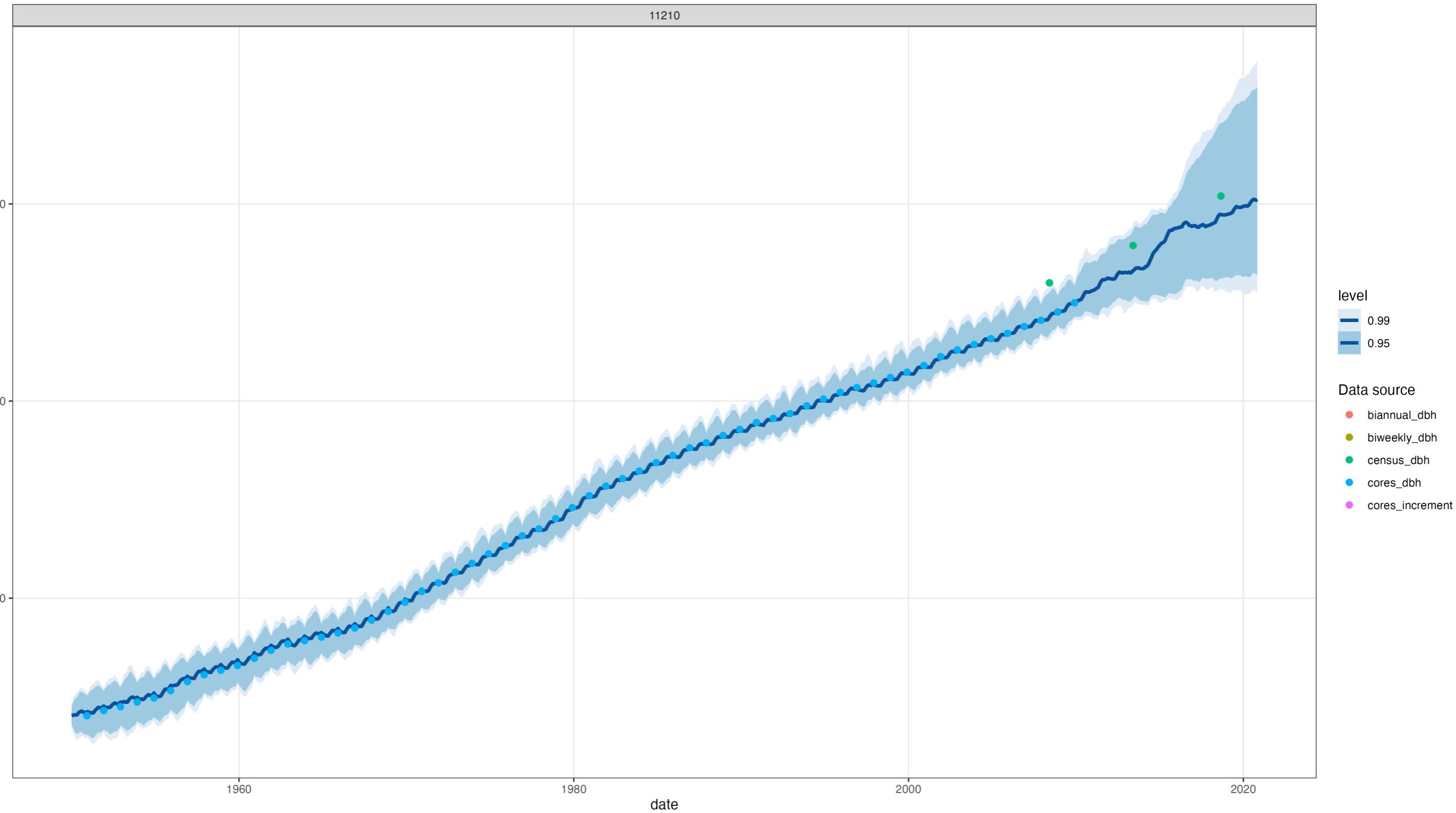
# dbh trajectories

DBH for stem 9035: quru



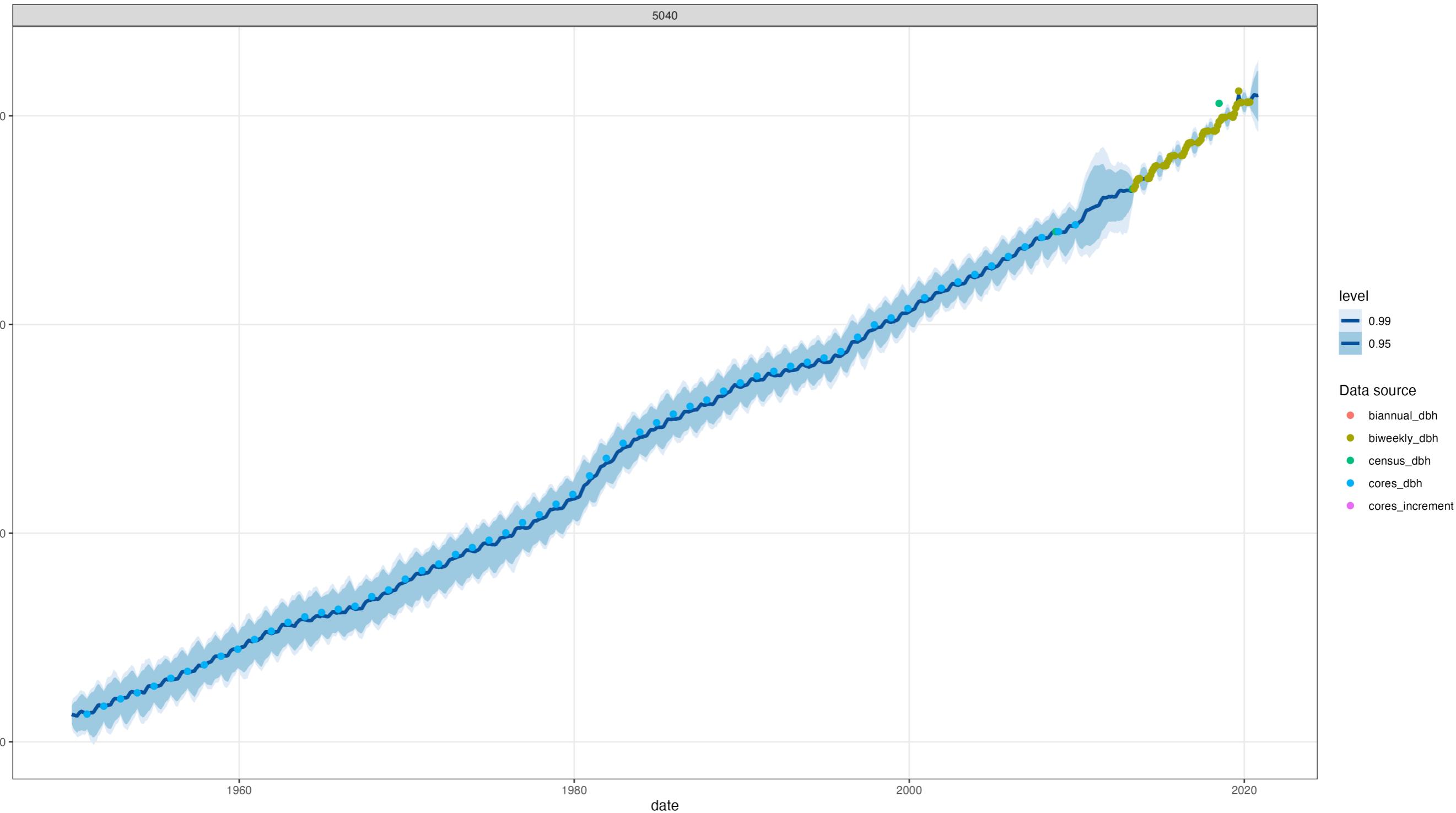
# dbh trajectories

DBH for stem 11210: qual



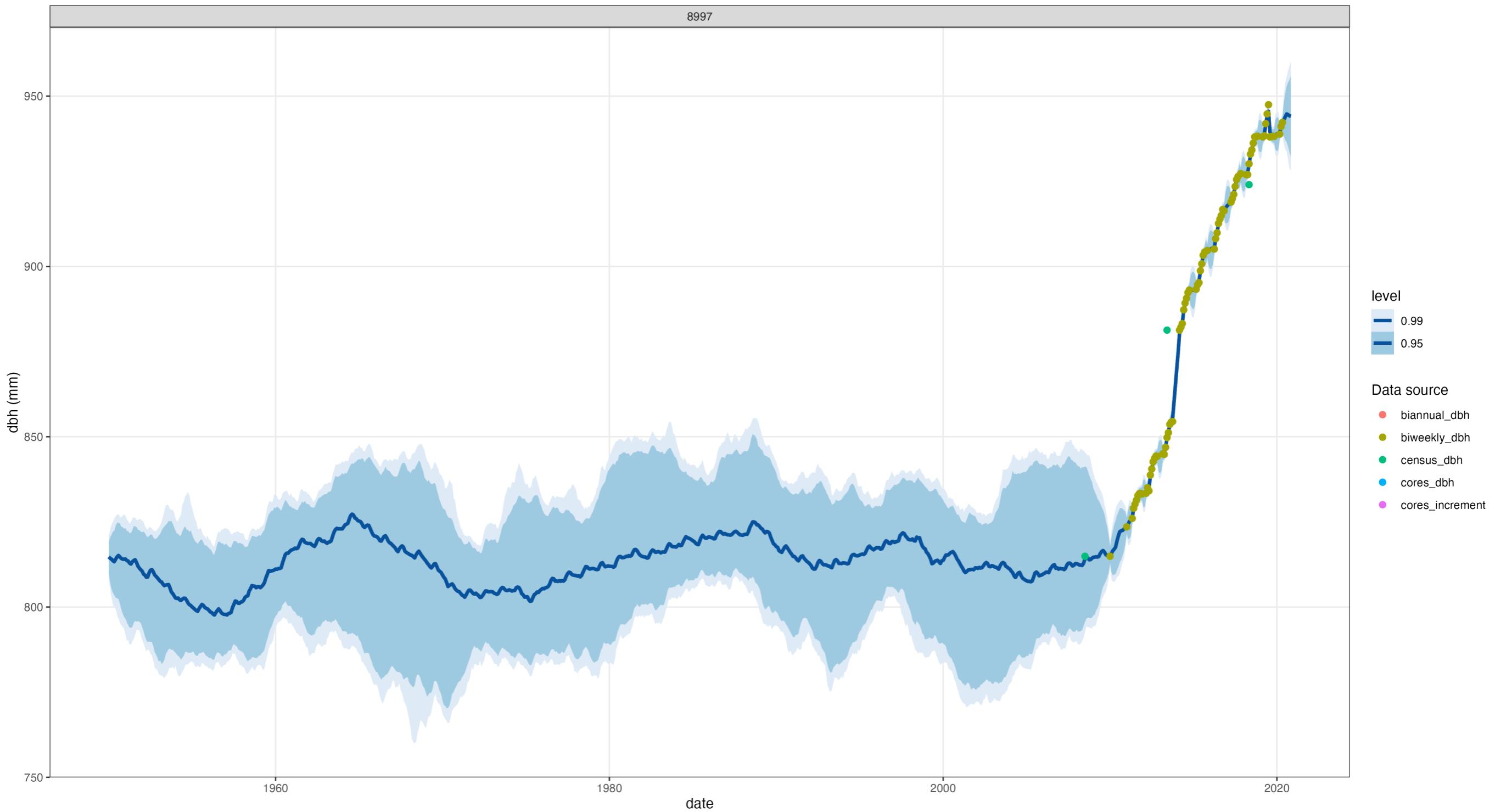
# dbh trajectories

DBH for stem 5040: fagr



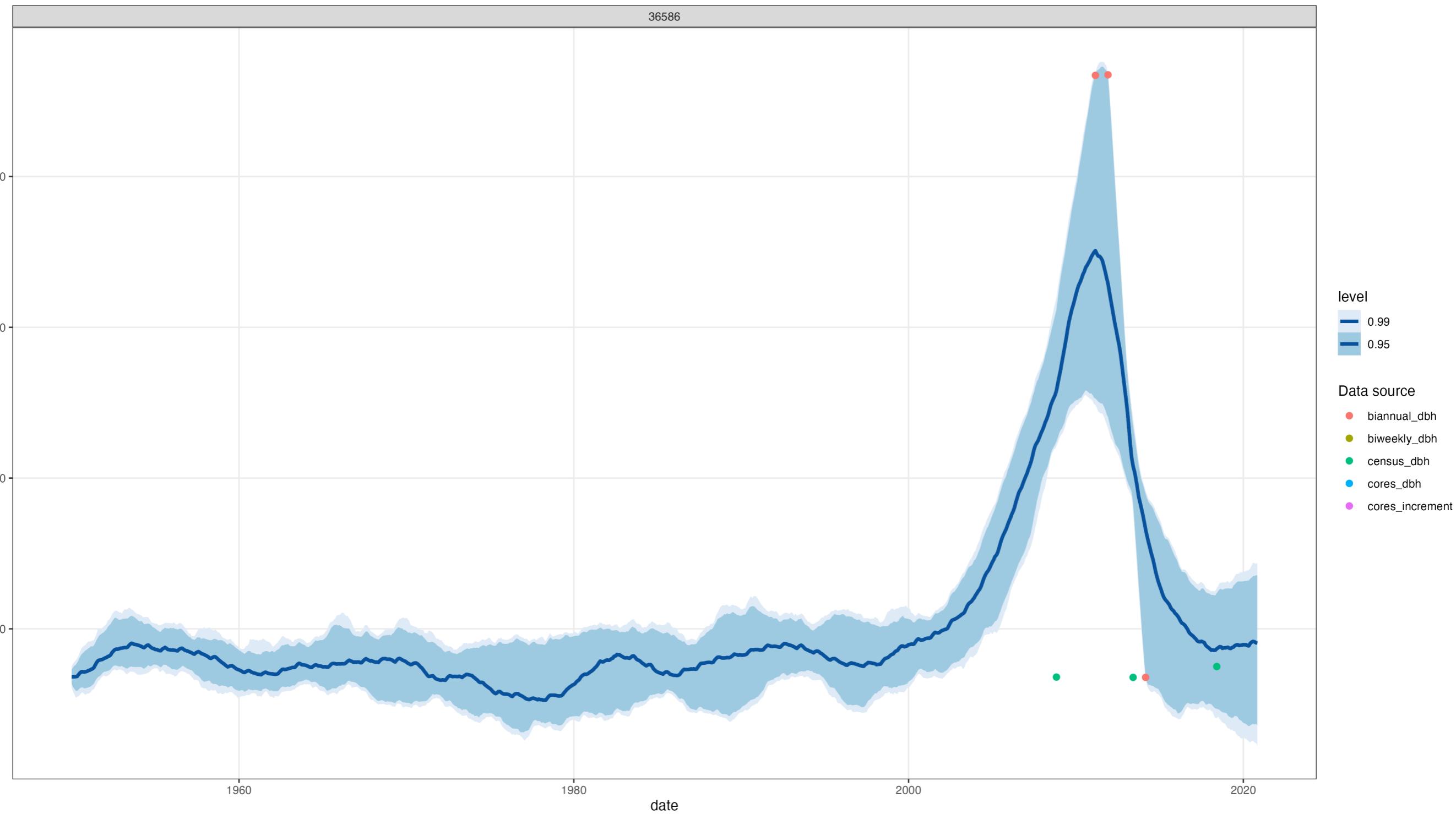
# dbh trajectories

DBH for stem 8997: quru



# dbh trajectories

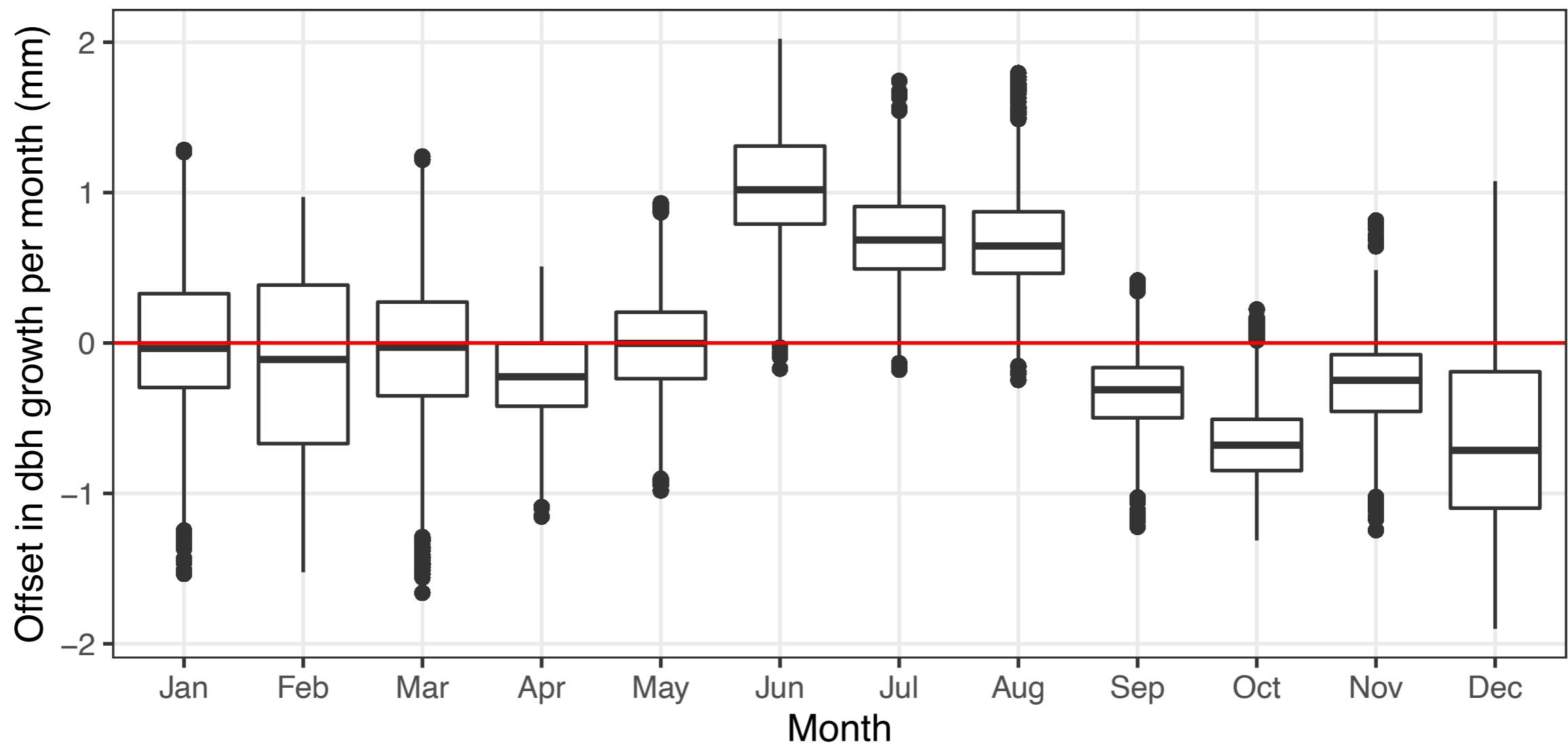
DBH for stem 36586: litu



# Posteriors: Month Random Effects

Month random effects

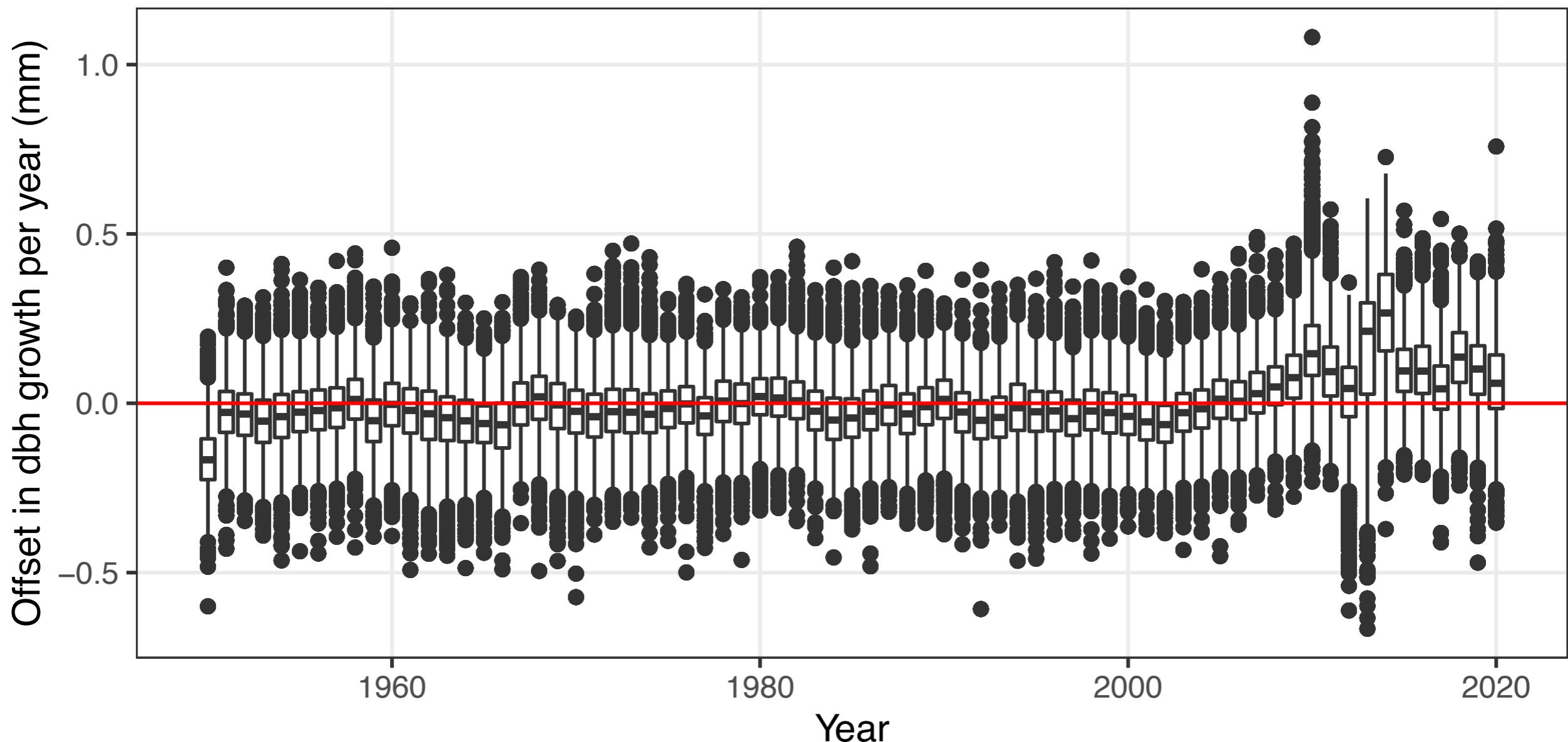
Distribution of all MCMC draws from posterior for each month



# Posteriors: Year Random Effects

Year random effects

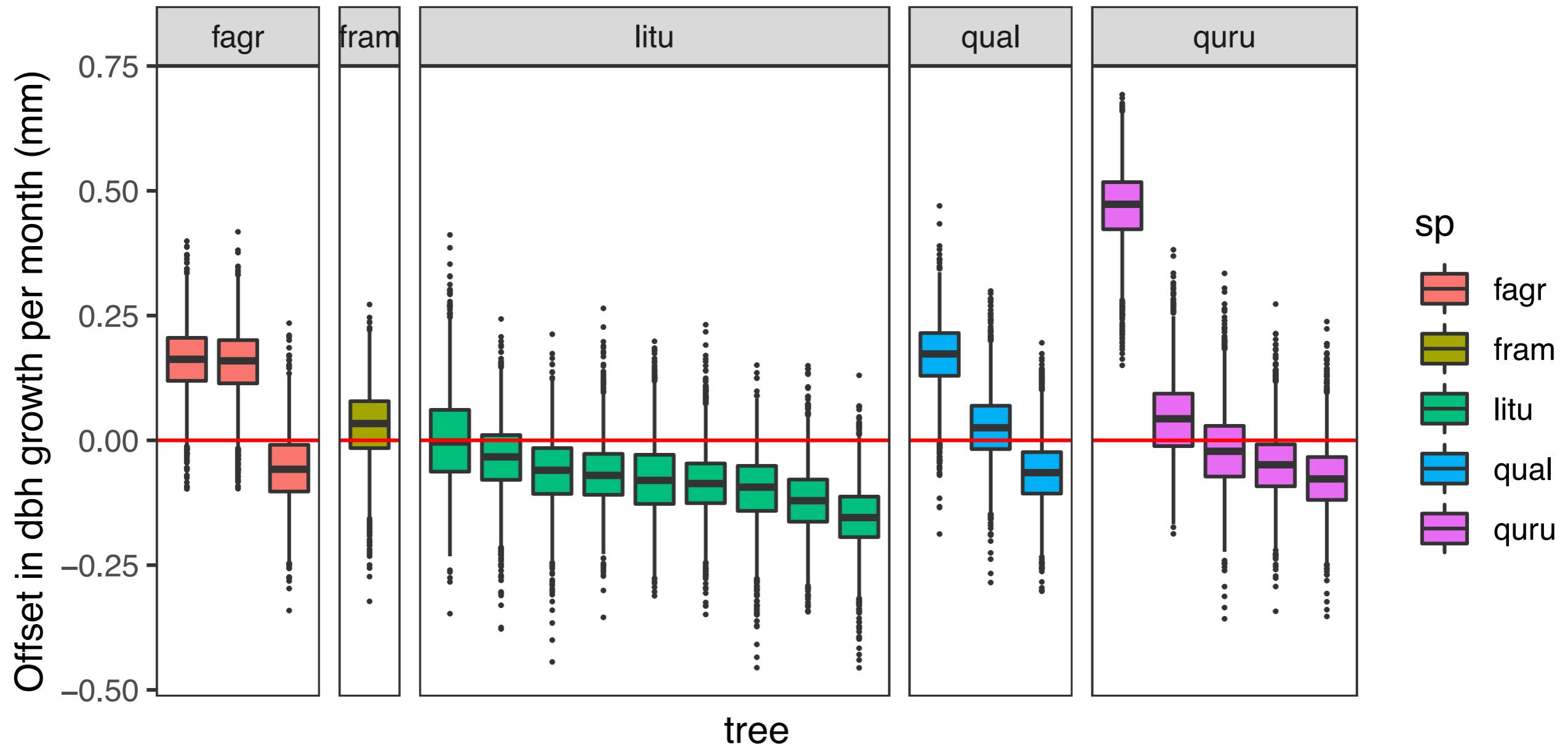
Distribution of all MCMC draws from posterior for each month



# Posteriors: Individual Random Effects

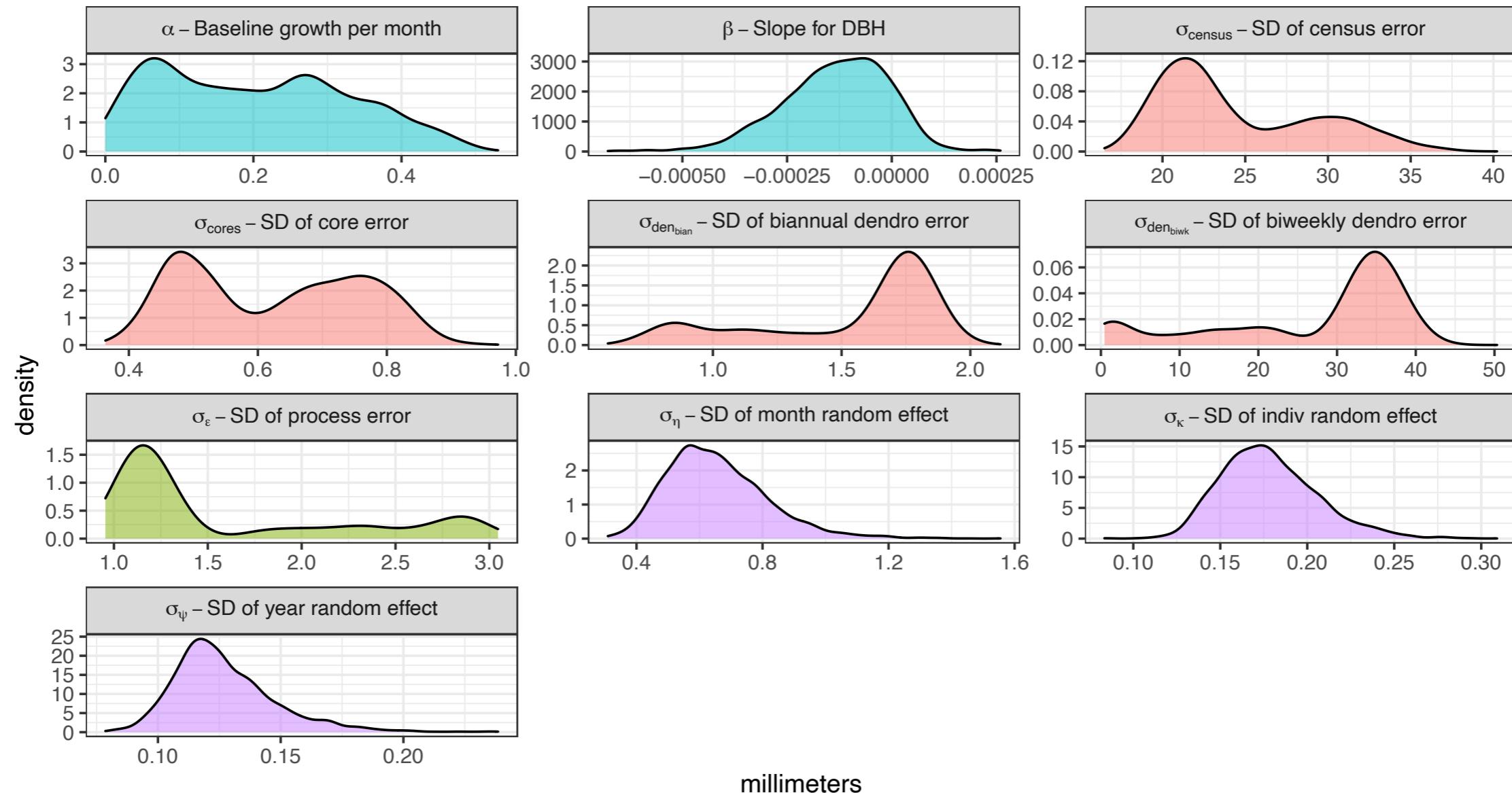
Individual tree random effects (top 7 sp)

Distribution of all MCMC draws from posterior for each tree



# Posteriors: Model parameters 🤔❌🧂

Posterior densities of all parameters



- cyan = process model parameters: baseline growth + dbh slope
- pink = measurement errors SD
- green = process error i.e. residual
- purple = random effects errors SD

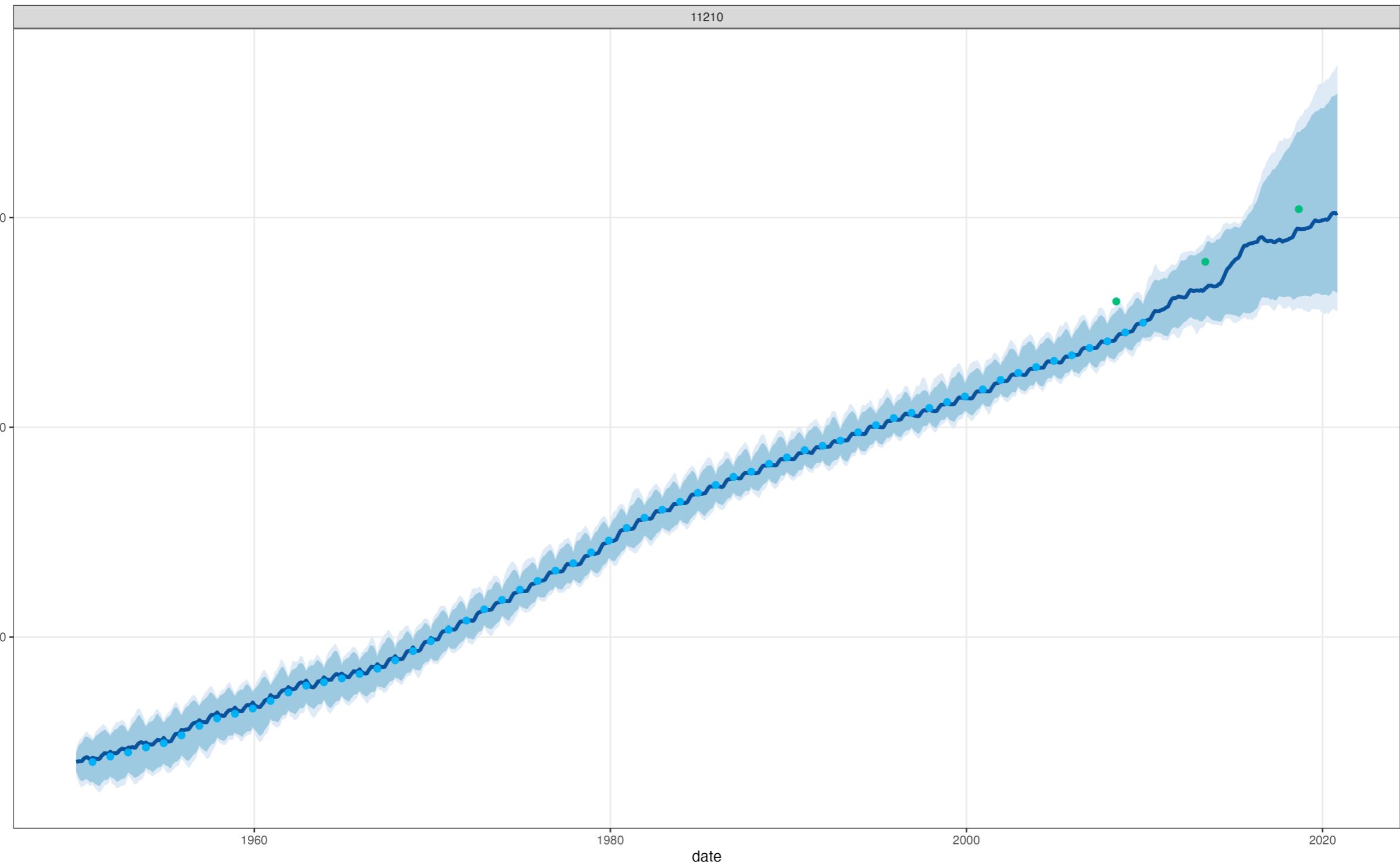
# Next steps

# TODO

1. Restrict model to have non-negative true latent growth from t to t-1
2. Add climate covariate to process model e.g.  $\beta_{temp}$  based on Helcoski 2019
3. Deal with correlation of observations/errors

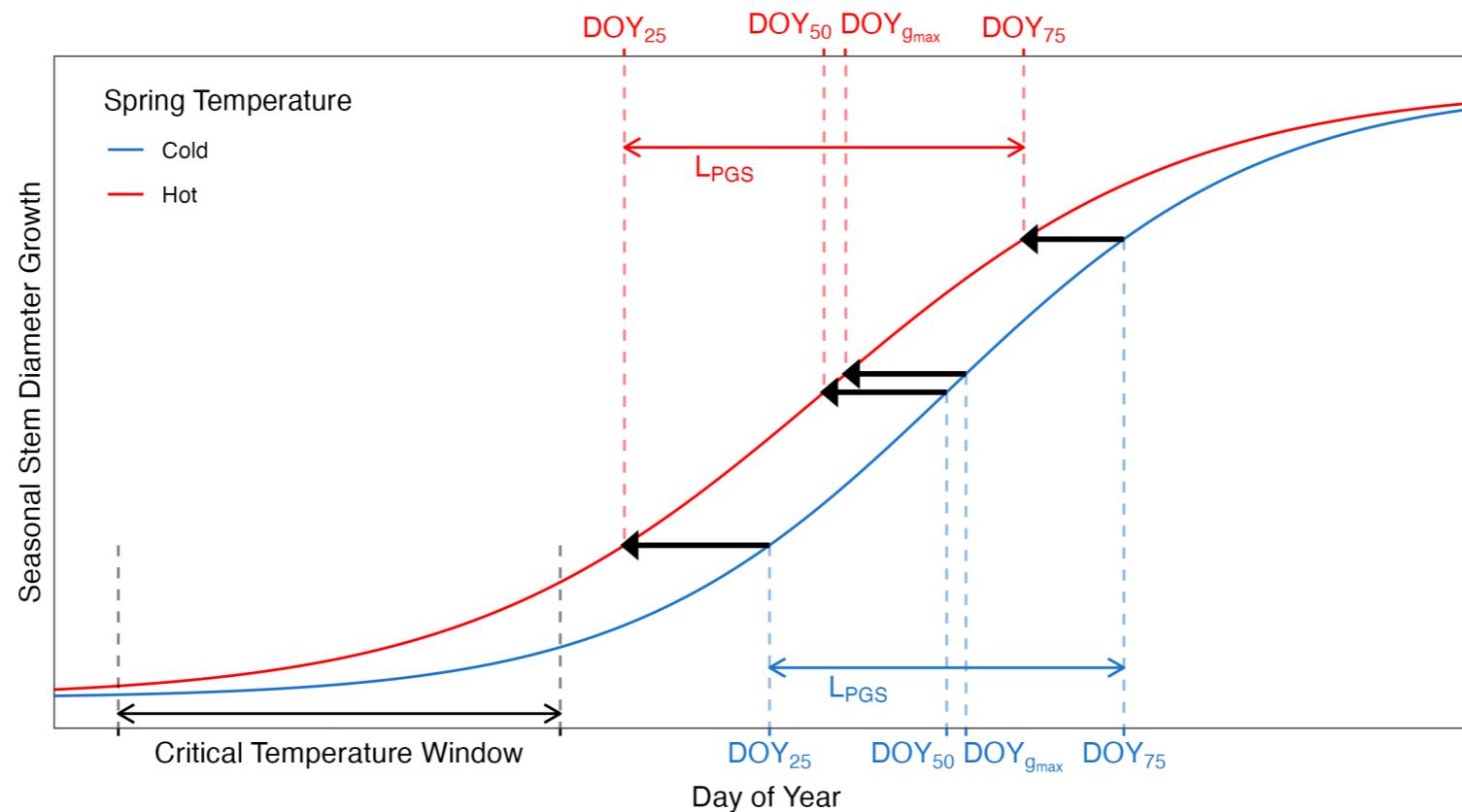
# Correlation of observations/errors

DBH for stem 11210: qual



# Related: ForestGEO Seminar Series

“Warmer spring temperatures in temperate deciduous forests cause earlier tree growth but have little effect on annual woody productivity”



Wed January 19th 9-10am ET  
[Sign up](#)

**Thank you!**

**Slides available on Twitter  
@rudeboybert**