Analysis methods for characterizing the climate sensitivity of annual growth and woody productivity using tree-rings

Kristina Anderson-Teixeira, Valentine Herrmann*

Note: This analysis lives in a GitHub repository (currently private): https://github.com/EcoClimLab/ForestGEO-climate-sensitivity. You're all welcome to see full code & results there, and to raise/ comment on issues. Ask KAT for access (you'll need a GitHub profile).

Goals / principles

- predict growth based on climate drivers *and* tree size, by species, so that we can subsequently scale to plot level to predict woody stem productivity $ANPP_{stem}$
- allow non-linear fits of growth to climate variables (note that this contrasts with traditional dendro methods)
- consider multiple relevant climate drivers simultaneously

Overview

There are 3 basic steps:

1. Identify most important climatic drivers of individual growth

2. Identify best multivariate statistical model accounting for climate and DBH

3. Scale to the ecosystem by combining statistical model with census data

These are outlined below ("Current methodology"), with additional considerations and alterative approaches outlined later ("Considerations and alternative approaches").

R script is here: https://github.com/EcoClimLab/ForestGEO-climate-sensitivity/blob/master/scripts/Analysis.R. (ask KAT for access if you can't access)

Current methodology

Data inputs

Raw data:

needed from collaborators

1. Cross-dated tree-ring records of radial growth increments (typically, .rwl files)

2. DBH of cored trees (any year). If not available, we can estimate DBH for trees on which the cores hit center.

helpful to get from collaborators

3. Bark thickness, preferably bark thickness allometries (particularly important for trees with very thick bark)

requested if collaboratos believe they have something better than what we're using

4. Biomass allometries. Note that we've compiling biomass allometries for ForestGEO sites in "allo-db", and have already sought out recommendations of site PIs.

5. Monthly climate data spanning the length of tree-ring records. We're currently using data from the CRU database, which has ~10 climate variables available, but in some cases it may be better to use alternative data sources.

Processed data that we put into model:

1. Tree growth - We start with raw data on radial growth increments (i.e., tree-ring widths). Optionally (probably preferably), we convert this to aboveground biomass increments (ΔAGB) prior to analysis. This requires DBH (see below) and a biomass allometry. (*Note: We are still working on the allometries*)

2. Tree diameters - This requires reconstruction of the historical DBH. We need DBH for any year. Ideally, we want to use a bark thickness allometry to correct for changing bark thickness.

3. Monthly climate data. We're currently using the ~10 climate variables available through CRU, but could consider others.

Step 1. Identify most important climatic drivers

1.1. "Detrend" individual tree growth records ($log[radial.growth.increment]$ or $log[\Delta AGB]$) by fitting a spline using a GAM, thereby producing residuals. This parallels detrending of traditional dendro analyses, which seems necessary to pull out a good climate signal (see "Considerations and alternative approaches"). However, rather than averaging by species as per traditional dendro methods, we will analyze individual-level residuals.

1.2. We use climate data from CRU v. 4.04. Based on how variables are related to one another, and how they are derived in CRU (Harris et al. 2020), we grouped variables as follows, using climwin to select the best variable and time window for each: - Precip group (describes precip and its frequency): PRE, WET - temperature group (describes temperature): TMP, TMIN, TMAX, PET - temperature range group (describes variation in temperature, linked to cloud cover): DTR, CLD, PET

1.3. Use the climwin R package (van de Pol et. al. 2016), to identify the climate variables and time windows that most strongly influences tree growth *for all species combined*. Climwin can check for linear and quadratic relationships. We checked quadratic only because (i) it speeds up the process, (ii) preliminarily tests revealled that quadratic was usually the one with lower AIC, and when not there tended to be little difference in AIC and the curve would look like a straight line anyway, (iii) quadratic makes more biological sense, (iv) it would be too tricky to code the gam to adapt to any circumstance. We put the residuals derived in step 1.1 in *climwin* with a mixed effects model using species and tree identity as random effect: `residual ~ [climate] + (1 | sp) + (1 | treeID)`. Outputs are saved here. Example output is shown in Fig. 1.

1.3. For each group of candidate climate variables, we move forward with the best variable over the time window identified by climwin as a candidate climate variable for the multivariate models. When PET comes out as the best variable in both temperature and temperature range groups (should always be for the same time frame), then there are only 2 candidate variables for the GAM.

(*Note: This process will tend to pick out current growing season variables, but it may be that once we have 1-2 variables for current growing season, others would be more important (e.g., previous growing season, start/ end of growing season). Thus, we may want to modify how we select climate variables. (GitHub issue # 19)*)
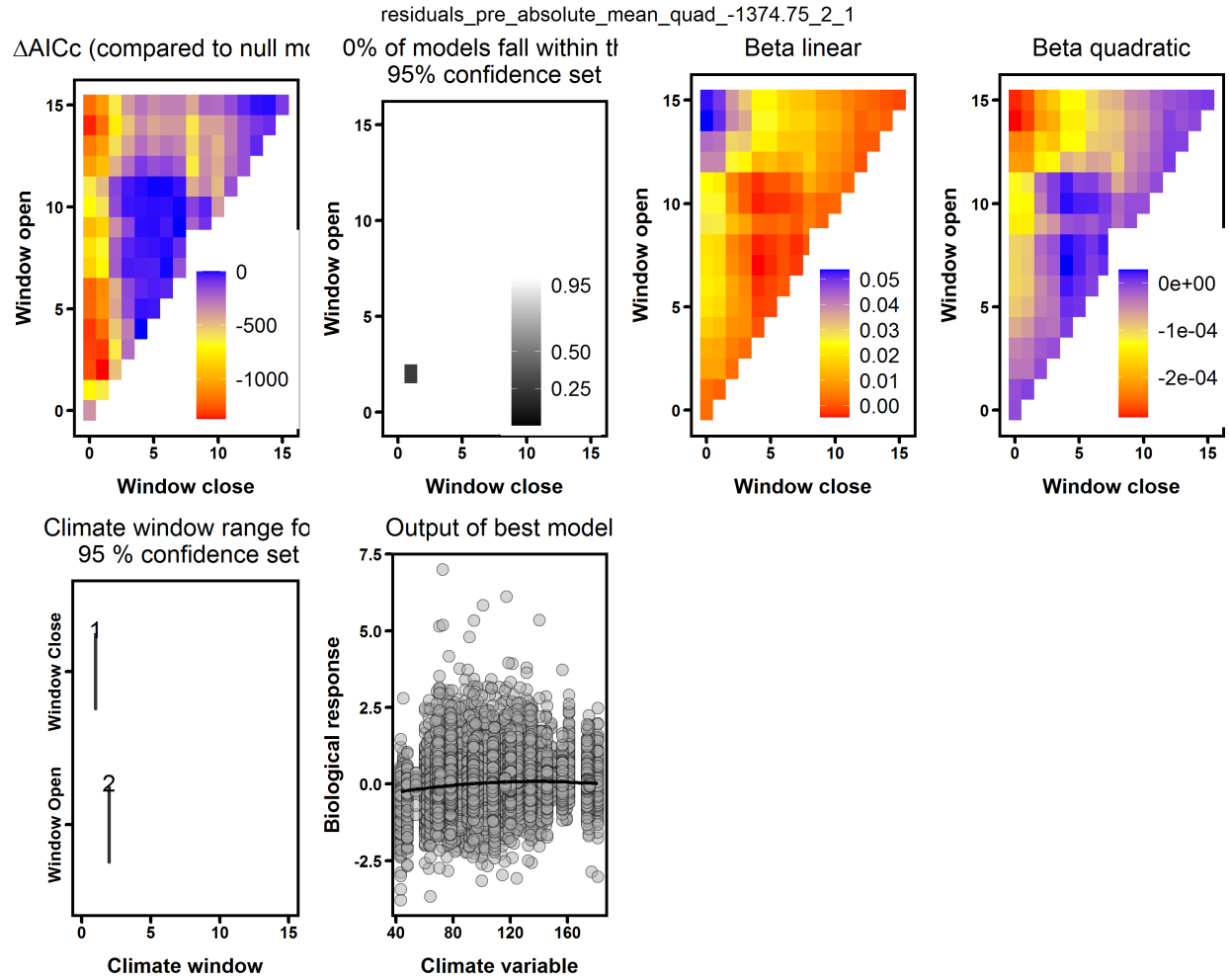
Fig. 1. Example output from Climwin. This example, from SCBI, looks at the influence of precipitation on radial growth increments. It identifies June-July of the current growing season as the most important window for precipitation's influence on annual growth. On axes, window open and window close refer to months prior to August of current gowing seasoon. Detailed explanation of output can be found in van de Pol et. al. (2016).

Step 2. Identify best multivariate statistical model accounting for climate and DBH

2.1. The 2-3 candidate variables for the GAM should then go through the same process of checking for collinearity. If there is redundancy, we removed any variable with VIF > 10 (more standard way than looking at AIC).

2.2. For each species independently, run GAM models including every combination of climate variables (as a 2-degree polynomial to allow quadratic curve), spline of DBH (3 knots), and splines of year fit to each individual as a random effect. `ln[response variable] ~ spline(DBH) + quad([climate variables]) + spline(year, with treeID as random effect)`

2.3. Across all possible models, sum variable AICc weights to determine which drivers are most important. Consider the model containing those variables as the top model (even if it's not top in terms of AIC)
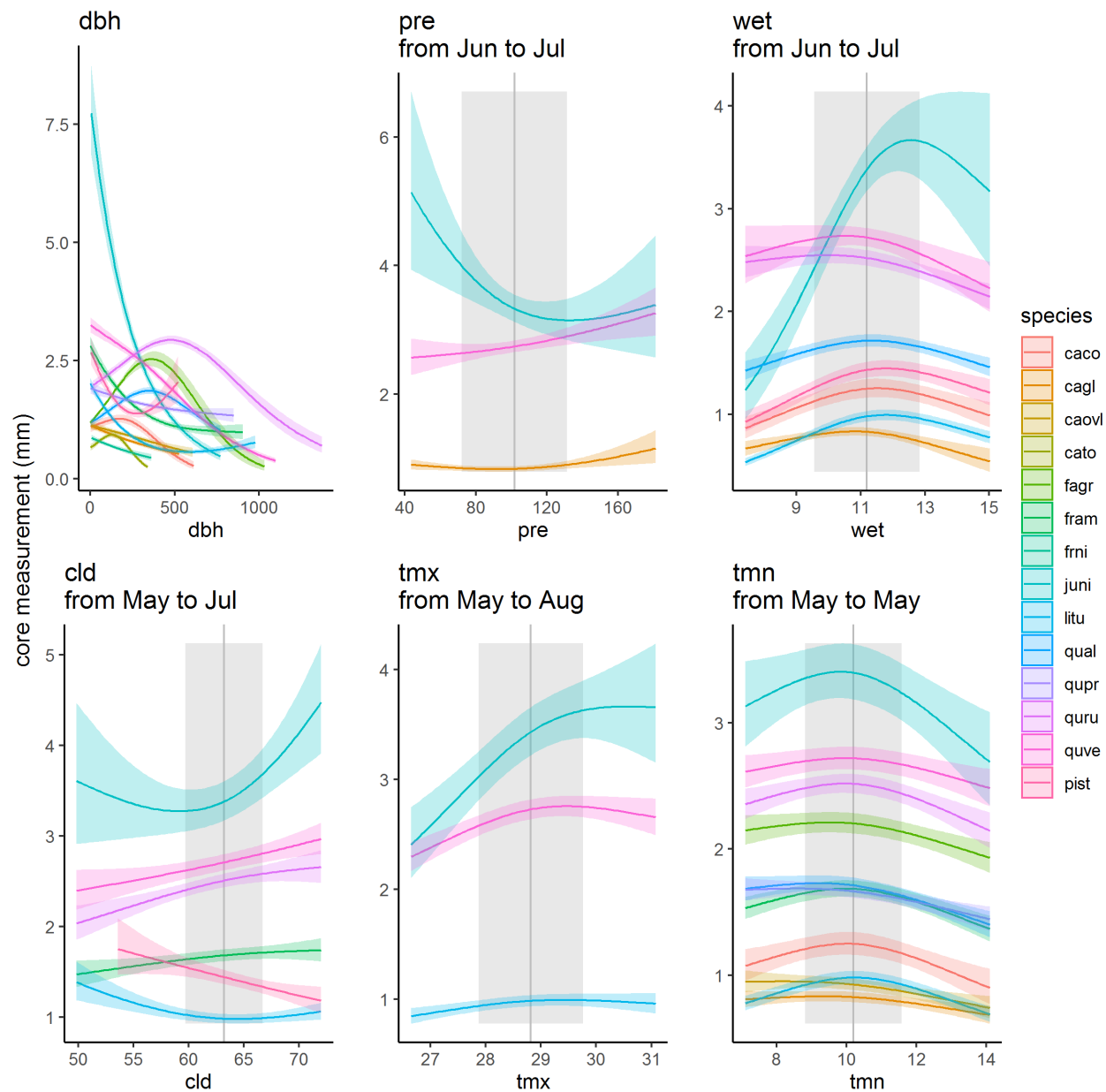
2.4. Plot best model for each species (Figs. 2-3)

Fig. 2. OUTDATED Current SCBI output for radial increment. For each species, relationships are plotted if included in top model. Grey bars indicate long-term mean climate, shading indicates 1 SD. Time windows for each climate variable (CRU abbreviations) are indicated above plot.

Fig. 3. OUTDATED Current SCBI output for ΔAGB. For each species, relationships are plotted if included in top model. Grey bars indicate long-term mean climate, shading indicates 1 SD. Time windows for each climate variable (CRU abbreviations) are indicated above plot.
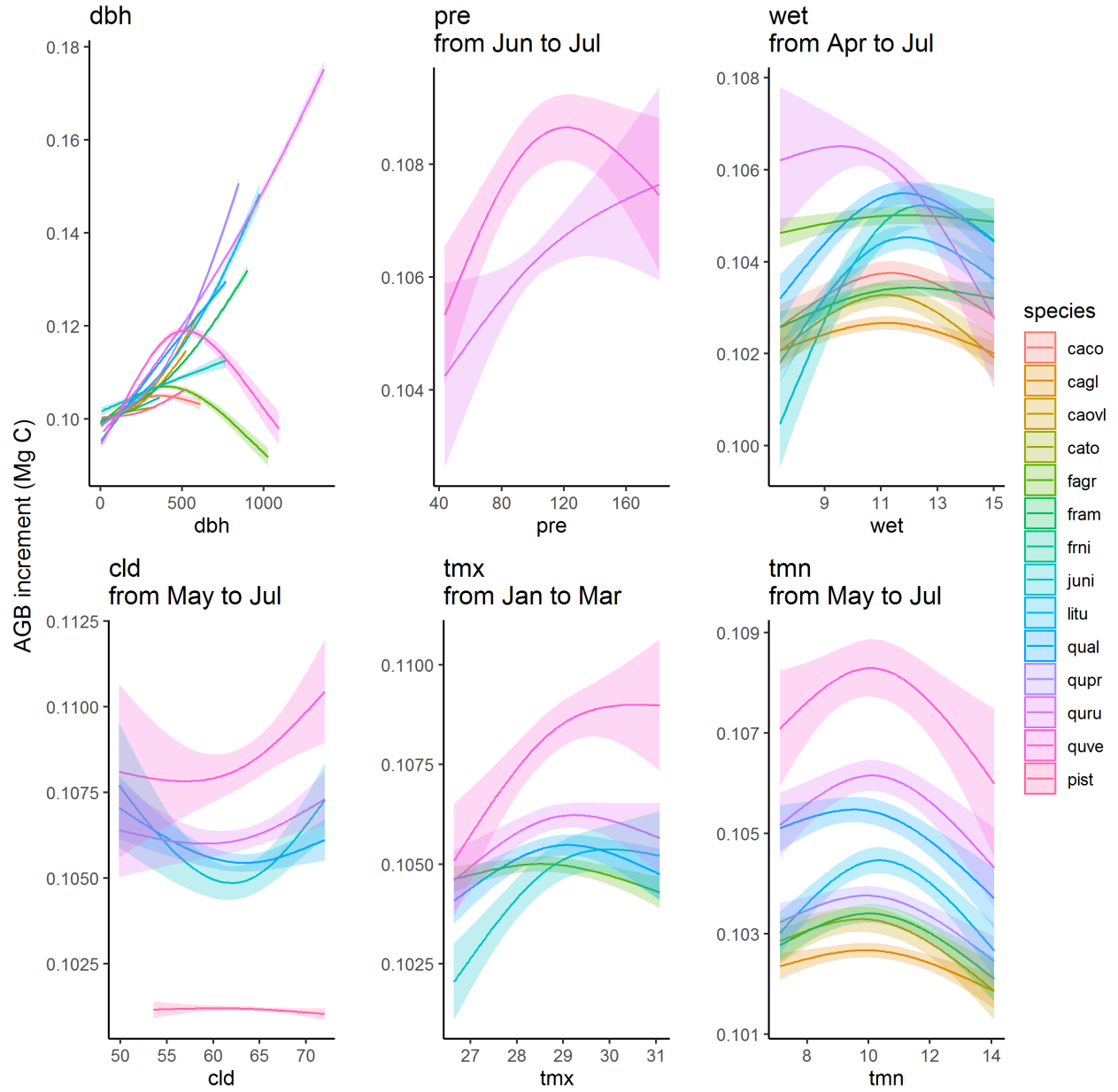
Step 3. Scale to the ecosystem by combining statistical model with census data

(We haven't started this yet. Basically, we'll use the statistical models selected in step 2 to predict growth for each individual tree under a range of conditions, sum to get $ANPP_{stem}$.)

Considerations and alternative approaches

Species- versus individual-level analysis

- Species-level chronologies seem to provide the clearest picture of climate sensitivity, using methods optimized by dencrochronologists (e.g., Helcoski et al. 2019). However, they aren't compatible with analysis of the roles of ecological factors, including tree size.

- Individual-level residual analysis seems to work fine for pulling out the climate signal, and is our current approach.

Response metric

- **ring width (radial growth)** - Annual ring width from individual-level chronologies. Helene Muller-Landau argues that, from tree scaling/geometry perspective, this metric is the closest to being size-independent (as opposed to BAI) (see Francis et al. (2017)). Empirically, this seems to be the case in closed-canopy broadleaf forests.

- **biomass increments (ΔAGB)** - Ultimately, we want to predict ΔAGB (and then sum to get $ANPP_{stem}$). The question is whether it's better to convert to ΔAGB before or after running the statistica model. Christy Rollinson recommends before, and we are leaning that way. However, we're continuing to run models both ways for comparison. It's good to see whether we get similar results both ways.

  - Foster et al. (2016) used indivdiual-level biomass increments.

  - Teets et al. 2018 summed biomass increments from cored trees at the species and stand level. They state, "Summing annual biomass increments maintains the signals from individual trees while homogenizing the effect of tree-level factors such as tree size and crowding." This logic makes sense. However, it requires that sampling be proportional to community composition.

- **percent biomass change** - potential option that would be less strongly dependent on size than biomass increments

- **basal area increment (BAI)** - (Currently don't plan to use this.) BAI is commonly considered by dendrochronologists to be independent of DBH (or at least moresy than ring width) after trees pass a juvenile growth phase, the logic being that if a tree produces a constant amount of stem wood area each year, BAI will hold steady and ring width will decline. However, while this may hold true in some cases, evidence from closed-canopy broadleaf forests around the world (ForestGEO) suggests that radial/ diameter growth *increases* with DBH (e.g., Muller-Landau et al. 2006, Anderson-Teixeira et al. 2015). There are some likely differences between trees represented by tree-rings (survivors) and the community as a whole, but in Helcoski et al. 2019 we show a positive relationship of radial growth to DBH from tree-rings. The implication–based on basic geometry–is that BAI increases even more strongly with DBH than does radial growth hence should not be considered independent of DBH. Furthermore, BAI in itself is not very meaningful, as wood is added in three dimensions, and BAI only accounts for two. Thus, KAT sees no *a priori* motivation to use BAI over ring width or biomass. BAI was used by and Alexander et al. (in review) but without assumption that BAI is independent of DBH.

Identifying the most important climatic drivers

Helcoski et al. (2019), (GitHub repo with full results) analyzed climate sensitivity using traditional dendro methods and gave a reliable (as far as we know!) picture of climate sensitivity at SCBI. From our explorations of other analysis methods on the same data, it seems that pulling out a reliable climate signal *requires* indivdidual-level detrending of cores. Our attempt to do this with raw growth data, using climwin and a mixed effects model to account for DBH and year, resulted in weak or absent climate correlations that didn't align with Helcoski results. However, if we fit splines to individual core sequences and then analyze these data in *climwin*, we get results that are generally very consistent with Helcoski (except that they don't pull out an opposite signal from the previous growing season).

We have two completementary ways of looking at climate sensitivity:

1. The climwin R package, described in this paper, is currently our best known option for identifying the most important climate drivers. This works with species-level chronologies or individual-level analysis, and can integrate with most types of regression models that can be implemented in R. We're currently using it with *lmer* (see issue #1).

2. The "quilt plots" in Helcoski et al. 2019 work well for visualizing the big-picture of multiple variables/ months / species, and are a useful complement to *climwin*. So far, we have only done this for species-level residual chronologies, not with individual-level analyses. However, there is no reason (other than the time required to code it) why we couldn't make similar "quilt plots" for individual-level analyses.

Potential alternatives

Mixed-effects model in climwin on raw data (not optimal)

This approach did not work well for pulling out climate signals.

- Used *climwin* with with mixed effects model: `diameter.growth ~ [climate] + DBH + year + treeID (random effect)`

- Some of the results aligned with Helcoski et al. 2019, but many don't (e.g., positive effect of DTR, where Helcoski found negative).

- This analysis could have potentially worked better if we'd worked harder at it (e.g., removing outliers, testing various ways of running it); however, it didn't seem like it was even getting close to picking out the signals identified in Helcoski et al. 2019.

- Conclusion: individual-level detrending seems to be the way to go for identifying the most important climate drivers.

Accounting for DBH (and other factors)

- DBH must be accounted for in the model. Peters et al. (2015) reviews some methods

- Should we account for age too? Probably. See issue #21.

Scaling to ecosystem level

What others have done:

- Helcoski et al. 2019 used tree-rings to charactize growth as a function of DBH, and assumed that the proportional response to climate was consistent across size classes. We want to move beyond that.
- Teets et al. 2018 used stratefied random sample based on community composition. They summed across individuals to get values proportional to ecosystem-level productivity. This approach does not work when sampling is not carefully stratefied relative to community composition.
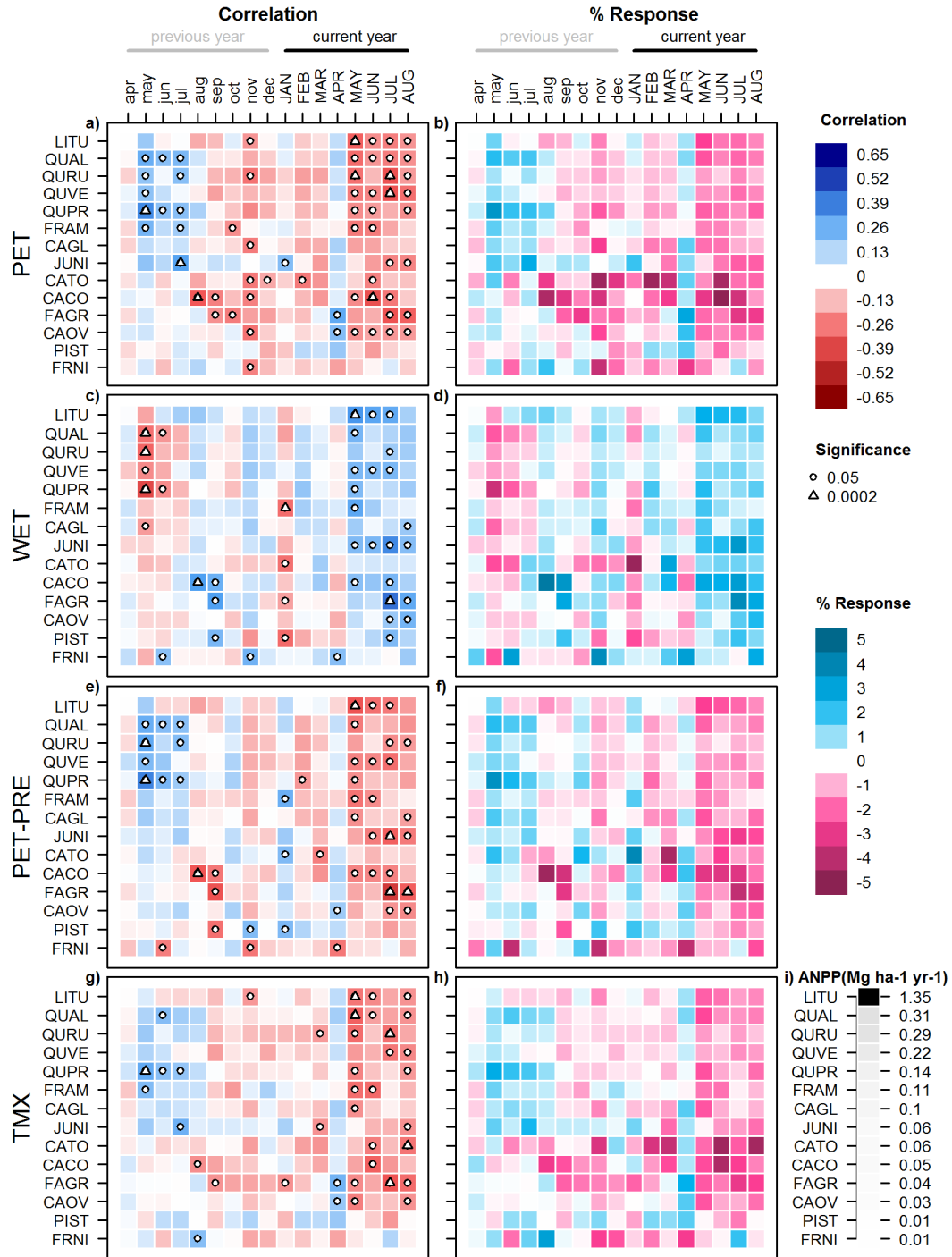
Fig. 4. Quilt plots of Helcoski et al. (2019), which used traditional dendro methods to characterize climate sensitivity of growth at SCBI. Shown are month-by-month correlations of species-level residual chronologies to climate of current and previous year for the 14 species analyzed.

Other statistical models of individual-level tree-ring data:

- M. Ross Alexander and colleagues (in review-TREAT CONFIDENTIALLY) "used generalized additive models (GAMs) using the gam function in the mgcv R package to allow for non-linear climate and size effects without prescribing a priori functional forms (Polansky & Robbins 2013; Peters et al.2015, Simpson 2018, Wood 2003). ....Year effects were estimated using a 4-knot thin plate regression spline fit by sampling plot to capture stand dynamics such as suppression or release events that may have influenced growth through time. Size effects were estimated for each species using 3-knot thin-plate regression splines fit to reconstructed DBH (Wood 2003)."

  Peters, R.L., Groenendijk, P., Vlam, M. & Zuidema, P.A. (2015). Detecting long-term growth trends using tree rings: a critical evaluation of methods. Glob Change Biol, 21, 2040–2054.

  Polansky, L. & Robbins, M.M. (2013). Generalized additive mixed models for disentangling long-term trends, local anomalies, and seasonality in fruit tree phenology. Ecology and Evolution, 3, 3141–3151.

  Simpson, G.L. (2018). Modelling palaeoecological time series using generalised additive models. Frontiers in Ecology and the Environment, 6, 1–21.

  Wood, S.N. (2003). Thin plate regression splines. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 65, 95–114.

- Foster et al. (2016) used an individual level Baysian heriarchical model to characterize to what extent whole-tree growth depends on climatic, tree, and stand structural variables.