

## Introduction

Phenology, the timing of seasonal life cycle events, allows organisms to synchronize important life history transitions with optimum environmental conditions (?), and is a critical component of ecosystem structure and function (??). Recent work in woody plant phenology has shown that it is not only individual phenological stages that affect these processes, but also the relationships between them (?).

One phenological relationship that has long received scientific interest (see ?), and recently, increased attention in the literature (??) is the flower-leaf phenological sequence (FLS) of deciduous woody plants. In a typical model of plant life-history, vegetative growth precedes reproduction. However, for many species in the forests of Eastern North America, it is not the green tips of new shoots that mark the commencement of the growing season, but the subtle reds and yellows of their flowers. This flowering-first FLS is common in these regions, and its prevalence suggests that this FLS has adaptive significance (?).

Understanding this phenological pattern is particularly necessary and timely now because anthropogenic climate change is altering FLSs (Fig. ??). Long-term observations show the number of days between flowering and leafout is increasing as a result of climate change, but the rate of change differs among species (Fig. ??). If FLSs are indeed an important component of woody plant fitness, this inter-specific variation will exacerbate fitness differences between species, influencing which species will persist under altered climate conditions. These long-term data showing shifts in FLSs over time also highlights how variable—within species—FLS are.

Despite recent advances in understanding the physiology and evolution of FLS (??), most research has ignored this variability—potentially styming efforts to predict how FLS patterns will respond to climate change. While some authors present general correlations between flowering and leafing phenology (??), fine-scale FLS variability has never been evaluated. We suggest that characterizing FLS variation among individuals and populations will not only improve our ability to predict how FLS patterns will change in the future, but also allow for a more biologically relevant evaluation of the current FLS hypotheses, revealing avenues for future, direct hypothesis testing.

Here we 1) Review the hypotheses of woody plant FLSs and their respective predictions, 2) Evaluate variation in FLSs, and explore how FLS variation within species, populations and individuals alters the predictions of the hypotheses, 3) Show how the incorporation of variation reveals consistencies and anomalies in support for FLS hypotheses using several case studies from temperate forests, and 4) make recommendations to improve future study of FLSs.

## FLS hypotheses and their predictions

### Hypotheses of FLS

#### Wind pollination

The most prevalent FLS hypothesis suggests that hysteranthly is an adaptation critical for effective wind pollination, with leafless flowering allowing for more efficient pollen dispersal and transfer (???). The primary evidence for this hypotheses comes from pollen diffusion studies (e.g., particle movement through closed and open canopies ???) and suggest pollen movement is indeed incomed by the maturing leaves, but exactly at what stage in the canopy delopment this effect becomes substantial has not be addressed.

Given this hypothesis, we would expect a strong association between FLS and pollination syndrome, with wind-pollination associated with hysteranthly.

## Water dynamics

Another hypothesis, emerging from the dry-deciduous tropics where flowering during the leafless season is also common (?), suggests that flowering before leaf development is an adaptation to reduce water stress associated with maintaining floral hydration while leaves are transpiring (?). We are unaware of any studies that have mechanistically evaluated the water dynamics hypothesis, though observations of flowering in the dry tropics suggest that the timing of flowering in hysteranthous taxa is associated with a plant water status recovery due to leaf drop (??). Given this hypothesis, we would expect a strong relations between FLS and metrics of drought tolerance, with increased drought tolerance associated with hysteranthly.

## Early flowering

A third possibility is that the flowering-first FLS is a physiological byproduct of selection for early flowering (?). Within this framework, there is no advantage to a species being hysteranthous vs. seranthous, as long as the absolute flowering time is the same. We are aware of no direct tests that have tried and distinguish selection for hysteranthly from selection for early flowering, but ? notes that hysteranthous, wind-pollinated species tend to also have large seed mass, and lack primary seed dormancy for germination, traits associated with early flowering in general. This raises the distinct possibility that hysteranthly may simply be one component of a larger suite of early flowering traits. Recent work from ? has demonstrated that spring flower phenology is less constrained by prior phenological events than leaf phenology, which would allow selection to drive flowering into the early season, producing the hysteranthous FLS. Given this hypothesis, we would expect a strong relationship between FLS and flowering time, with earlier flowering assoicated with hysteranthly.

## Phylogenetics

Finally, it is also possible that FLSs are highly conserved traits, and the preponderance of hysteranthly in the temperate zone is a product of phylogenetic representation of the region rather than an adaptive aspect of the trait. In a recent analysis, ? found strong phylgenetic clustering in FLS variation. Given this hypothesis, we would expect to see only weak relationships between FLS and the above mentioned traits and a strong phylogenetic signal in FLS.

Despite decades of inquiry, only limited progress has been made towards any consensus for these hypotheses. Many studies only test a single hypothesis, making comparison between them difficult. In contrast, studies testing multiple hypotheses have generally found support for more than one evolutionary driver of hysteranthly(see ??). This lack of concrete progress stems from the current way that FLS patterns are conceptualized conflict with the biological processes that research aims to test.

## The current FLS framework and its limitations

Under the current framework, FLSs are described qualitatively, and defined at the species level. This system inherently introduces observer and interpreter bias into FLS data, and cannot appropriately capture underlying biology that structures FLS. This mismatch between the framework and the biology of FLS confounds FLS hypothesis testing and is a major reason why understanding the function of FLS varation remains unresolved. Below, we highlight three main spurrious assumptions of the current FLS framework, clarify the

biological reality using long term phenological observations and suggest more appropriate alternatives for describing FLS patterns.

# **Assumption I: FLS patterns are simple to describe and biologically meaningful:**

In the current framework, there are three main FLS categories; flowers before leaves (hysteranthly, proteranthly, precocious flowering) flowers with leaves (synanthly) and flowers after leaves (seranthly) (??). Some data sources (e.g. ??) include additional categories “flowers before/with leaves” and “flowers with/after leaves” but it is unclear whether these categories are meant to describe intermediate FLS patterns or describe FLS variability in these species. While these categories are conceptually reasonable applying them to real phenological sequences is not always so straight forward.

## **The biological reality:**

Both reproductive and vegetative phenological sequences consist of multiple sub-stages, and this introduces significant ambiguity into how we interpret qualitative FLS descriptions. Consider a species with the following FLS:

**flower budburst → leaf budburst → first flowers open → leafout → peak flowering →  
end of leaf expansion**

Phenological observers could justifiably classify this species as: 1) Hysteranthous because flower budburst proceeds leaf budburst, 2) Synanthous because flowers open during the time between leaf budburst and leafout or 3) Seranthous because peak flowering occurs after leafout. This problem extends beyond this simple example to real datasets, such as the long term phenological records from Harvard Forest in Petersham, MA (?) where the same ambiguities exist (Fig. ??a). Not surprisingly then, different sources (e.g., floras) may classify the same species as synanthous or hysteranthous. For example, we compared species-level FLS descriptions in two of the most comprehensive records of FLS, Michigan Trees and its companion volume Michigan Shrubs and Vines (MTSV) (??) with The USFS Silvics manual volume II (?). Of these 49 amount of species 30% were classified differently. Such different classifications could reflect interesting temporal or geographic variability in FLSs, but—given current definitions—they could equally be an artifact of observer decision-making.

Given that the most complete FLS datasets currently come from regional guide books and floras like the ones mentioned above, these categories were most likely originally described to aid with plant identification rather than to describe functional biological processes like those underpinning the hypotheses may suggest different boundaries. Such categorization can often introduce biases in analyses (). For example, the wind-pollination hypothesis hinges on the fact that leaves create a substantial physical disruption to pollen transfer, a premise that we would not necessarily expect to be true for the early stages of leaf expansion when tiny leaf primordia would have little impact on environmental structure. Rather we expect that trees that flower during the early stages of leaf expansion would gain similar mechanical advantage to those who complete their flowering before any leaf activity. Alternatively, because transpiration intensifies as soon as leaves begin to expand in the spring (??), the water dynamics hypothesis would assert there is a significant cost to maintaining floral structures during any stage of leaf activity and only species whose flowering occurs before any leaf expansion would gain a drought advantage. It follows a statistical relationship between FLS and pollination syndrome is more likely if hysteranthly and synanthly are collapsed into a single category, and a relationship between FLS and water dynamics would be more likely when synanthly and seranthly are combined.

## Assumption II: There are no meaningful differences within FLS categories

Consider the two species depicted in figure (Fig. ??a). According to the current framework, both species would be classified as hysteranthous. Given the the FLS hypotheses, we would expect that these species should have the similar correlations between FLS and hypothesis-relevant traits. We would interpret a failure to find these associations as lack of support for a given hypothesis.

### The biological reality:

These species are classified the same simply based on sequence, but this ignores any other measure of time. Biology, however, has continually shown that the length of time also matters(). It is unreasonable to assume that the FLS of a species that flowers on average 30 days before leaf expansion is under the same selection pressures and functionally equivalent to one that flowers just 3 days before leaf expansion. This assumption obscures substantial inter-specific variation that may could in reality be the fingerprint of selection of FLS. For the X species in Harvard Forest that flower before leaf expansion, we found significant differences in the average time between flower and leaves. This inter-specific variation should be leveraged for more biologically reasonable hypothesis testing.

## Assumption III: FLS variation is a species level trait

In the current framework, FLS categories are assigned at the species level.

### The biological reality:

We investigated individual FLS variation in both the PEP725 (?) Harvard Forest data (?), and found that the time between flowering and leaf activity varied by as much as several weeks for individual species between years and sites. This variability is lost completely in the classic framework of categorization. For example, for *Q. rubra*, a species classically listed as flowering and leafing in synanth, there are some years in which flower budburst is more than a week before leaf budburst, and other years in which leaf buds burst weeks prior to floral budburst (Fig. ??). Given the variability of FLSs at the individual and population level, it is clear that considering FLS variability at only higher taxonomic levels obscures important realities about the biology of this phenological trait.

Under the current framework, any statistical relationship between FLS and other traits predicted by the hypotheses is biased by the subjectivity of the original observer, the modeler, and the possibility that the associations being tested do not reflect the biology processes that shape FLS. To illustrate this, we compared trait associations with hysteranth between the the MTSV and USFS datasets. To address how FLS definitions may bias certain hypotheses we applied two alternative FLS classification schemes; physiological hysteranth, which allowed for no overlap between floral and leaf phenophases, and functional hysteranth, which allowed for a degree of overlap. We found systematic differences between the datasets which differed in how FLS was categorized and species included in each. As predicted, we also detected categorization bias with categorization that matched the hypothesized mechanism being better. Thus a researcher using our MTSV dataset with functional FLS definitions would find strong support for the wind pollination hypothesis but one using USFS physiological data would not.

## A new framework for FLS

The shortcomings of the current FLS framework are easily rectified with a shift from categorical, species level descriptions of FLS to continuous individual level quantifications. The primary reason for this shift

is quite logical, if researchers seek to understand the *biological* processes driving FLS patterns, we will be the most successful if we use the most *biologically* relevant measures, but there are important secondary advantages of such a framework.

Quantitative measure of phenology (e.g. the BBCH scale, ?), standardize data across time and space, observer, and analyst. Adopting such measurements in the study of phenological sequences would allow for FLS patterns to be compared across larger temporal, geographic and taxonomic scales, giving researchers more power to accurately address questions about FLS variation.

One of the challenges of inter-specific comparisons that can be mitigated by an intra-specific approach is that species evolve a suite of traits for any function, and unmeasured traits might bias our results (?). For example, wind-pollinated species could compensate for pollen intercepted by a synanthous or seranthous FLS by over-producing pollen or through self-pollination. To really understand FLS across large taxonomic space, one would have to compare species across an unfeasibly large, N-dimensional trait space. Focusing on FLS variation within species holds most other traits relatively equal, avoiding this problem of tradeoffs with latent unmeasured traits.

Finally, considering the high levels of quantitative inter-and intra-specific variation is valuable for honing the currently FLS hypotheses and generating more readily testable predictions. Below, we discuss how the observed, quantitative variation below the species level may alter the existing FLS hypotheses and generate new predictions.

## How FLS variation alters predictions

### Wind pollination

Pollination syndrome is generally treated as a species-level trait, considered to be fairly immutable across ecological time and space. Because of this, we would not expect significant variation in FLS across populations or individuals because we would not expect variation in pollination syndrome. However, as discussed above, a tree with no overlap between flowering and leafing phenology does not necessarily gain a significant pollen transfer advantage over an individual with some overlap. The pollination efficiency advantage from flowering-first should diminish as the canopy fills in, but we do not know at what point during leaf expansion pollination would become significantly encumbered. It is possible that interannual and population-level variation in the timing between flowering and leaf out for hysteranthous and synanthous individuals could maintain a wind pollination advantage, as long as the overlap did not cross a certain unknown threshold. Therefore, based on the wind pollination efficiency hypothesis, we would not expect high levels of population or individual variation in FLS, but the detection of some FLS variability at these levels does not inherently challenge this hypothesis.

### Water dynamics

If FLS's are driven by water dynamics, we would expect there to be significant population-level variation in FLSs. Populations growing in drier habitats should flower earlier relative to their leaf activity than their counterparts growing in wetter areas that experience weaker selection for minimizing phenological overlap. Therefore, increased time between flowering and leafing should be negatively correlated with water availability. Water availability may also drive interannual FLS variation, with drought years increasing hysteranthous, and wetter years permitting more FLS overlap.

## Early flowering

This hypothesis predicts some variation on the population level based on local adaptation. We would expect populations in which selection for earlier phenology is stronger, perhaps those in regions with shorter growing seasons, to flower earlier relative to their leaf development. At the individual level, FLS variability could be driven by interannual variability in spring conditions. Both flowering and leaf phenology are strongly cued to temperature and photoperiod (??), but with leaf phenology constrained by xylem activity and flowering phenology relatively independent of it, we would expect a more sensitive response to environment in flowering time resulting in FLS variation. This hypothesis predicts that early flowering years or populations should be associated with an increase in the time between flowering and leafing for hysteroanthous species. It also predicts a tighter temporal correlation between flowering and leafing for seranthous species or those with mixed buds in which flower timing is constrained by leaf budburst.

## Phylogenetics

With the lack of treatment of intra-specific FLS variability in the literature, we have no strong basis for asserting whether the apparent variability in FLSs is a product of genetic or environmental controls. If there is a strong genetic component to FLS as has been shown for other phenophases (?), some population-level variation could be driven by reproductive isolation. With strong genetic control of FLS, we might also see consistent genotypic differences in FLS among individuals within a population, but would not predict high levels of interannual variation.

## Testing the new framework

To date quantitative measures of FLS and variability below the species-level have not been addressed. Yet, there are datasets widely available that allow for testing these several hysteroanth hypotheses concurrently, and at multiple taxonomic levels. To address this gap, we supplement our literature review with several analyses. First we compare biological insights generated by the old and our new framework of FLS using the long term phenological records at Harvard Forest in Petersham, MA to demonstrate and discuss how the shift to quantitative individual measures enhance our understanding of the nuances of the hysteroanth hypotheses. Then we leverage temporally and geographically explicit quantitative flowering and leaf phenology records for four species from Pan European Phenological Database (PEP725) (?) to test several of the predictions from the modified FLS hypotheses.

## Harvard Forest:

The results from the categorical version of our Harvard Forest model(2b) are in line with those from the MTSV and USFS model (Fig. 2a), showing a strong association between early flowering and hysteroanth, a weaker relationship between wind-pollination and hysteroanth, and no evidence for the water dynamics hypothesis or interactions between predictors. However, the results from our quantitative version of the Harvard Forest model paints a very different picture, pointing to key biological insights obscured by categorization. While the main effects are generally consistent, the quantitative model suggests there is some support for the water dynamics hypothesis, implying that while drought tolerance may not play a major role in determining if flowering happens before leaves, it may contribute to determining the time between these phenophases.

Most significantly, the quantitative model identifies biologically meaningful interactions between predictors. Average predictive comparisons based on this model demonstrate that water dynamics play a major role in structuring the FLS patterns of biotically-pollinated taxa but not wind-pollinated taxa (4a). This

interactions suggest that the frequency of support for multiple FLS hypotheses in the literature found may have a strong biological basis, reflecting convergent evolution in FLS. Many of the biotically pollinated trees of the temperate forests trace their biogeographic origins to the tropical regions where the water dynamics hypothesis originated (?). It may be that the occurrence of hysteroanthony among biotically-pollinated species in the temperate zone originated radically different selection environment and bears no functional relationship hysteroanthous flowering in wind-pollinated species.

## PEP725

### The Future of FLS:

Using available data, we have demonstrated the advantages of a new conceptual framework for the study of FLSs based on quantitative measures of individual variation in FLS patterns. Using these methods we have identified several biological nuances to advance our understanding of this phenological trait and provided a more comprehensive picture of where our understanding of this phenological trait is currently, and where it needs to go. Below we highlight three characteristics of FLS that we suggest, utilizing this new framework, should be incorporated into future research to improve our fundamental knowledge about this important life-history trait and better predict how alterations to FLSs will impact species in an era of global change.

#### Multiple hypotheses explain FLSs

Our results underscore other lines of evidence that show multiple hypotheses should be starting point for all future FLS research. While there is certainly value to broad taxonomic studies, and future large-scale analyses should continue, the consistent support for multiple hypotheses shows there are limits to the utility of these kinds of studies. We suggest that it is better to explore the evolutionary dynamics of hysteroanthony with a more mechanistic approach, which may mean utilizing a more taxonomically-restricted focus. The significance of interaction terms in some of our models suggest that a promising option is to look within the hypotheses to address sub-grouping of taxa in which overlap between hypotheses could be controlled. For example, we know that wind-pollination efficiency is not driving hysteroanthous flowering among biotically-pollinated taxa, so if we consider this group of species alone, we may be able to detect stronger signals from other traits that support other competing hypotheses. Incorporating a more explicit phylo-biogeographic approach would be instructive at this level; if there are phylo-geographic commonalities between the few biotically-pollinated hysteroanthous species in Eastern flora, we might better understand the function of FLS variation in these species by investigating FLS variation in their sister-taxa in their regions of origin.

#### FLS and fitness

Even with focused work on sub-groupings of species, interspecific trait-association models can only take us so far. As in most other areas of plant biology examining traits, research is hampered by the difficulty of knowing which are the right traits (). For example, we used minimum precipitation across a species' range, one of the only available quantitative drought metrics at the scale of large inter-specific models, to represent the water dynamics hypothesis but we have no way of knowing for certain that this is really a good proxy for drought tolerance.

While trait associations point to past selection, fitness is the driver of trait evolution, and at the core of each FLS hypothesis is a fitness prediction. By utilizing intra-specific comparisons and continuous measurements of FLS, we can move beyond trait associations and test the fitness consequences of FLS variation.

Variability in hysteroanthony should lead to variability into fitness outcome at the intra-specific level. For example, the wind pollination hypothesis predicts that with all else equal, years with increased time between

flowering and leafing should correlate with more pollination success. The water dynamics hypothesis suggests hysteranthous populations with a consistently larger time between flowering and leafing should better tolerate drought. These predictions could be directly assessed through well-designed experiments and field studies.

## FLS and physiology

Decades of research suggests that both floral and vegetative phenological events are cued by temperature and photoperiod (??), suggesting they are under shared genetic and physiological control. But to yield the FLS variation seen in nature, there must be systematic differences in reproductive and vegetative phenological responses to the environment. Researchers can use intra-specific variation in FLS to identify which cues dominate each phenological process and better understand the underlying genetic and physiological constraints that structure phenological sequences.

*Conclusions:* Our proposed framework provides a path to understand the drivers of FLSs in woody plants. Through examining FLS variation in more targeted taxonomic assemblages and using quantitative data with mechanistic metrics, we can refine the existing FLS hypotheses and better comprehend the causes and consequences of FLS variation at multiple taxonomic scales. This is an essential step towards a more complete understanding of the fundamental biology of temperate woody plants, and for predicting the fate of these species as global climate continues to change.

## Figures



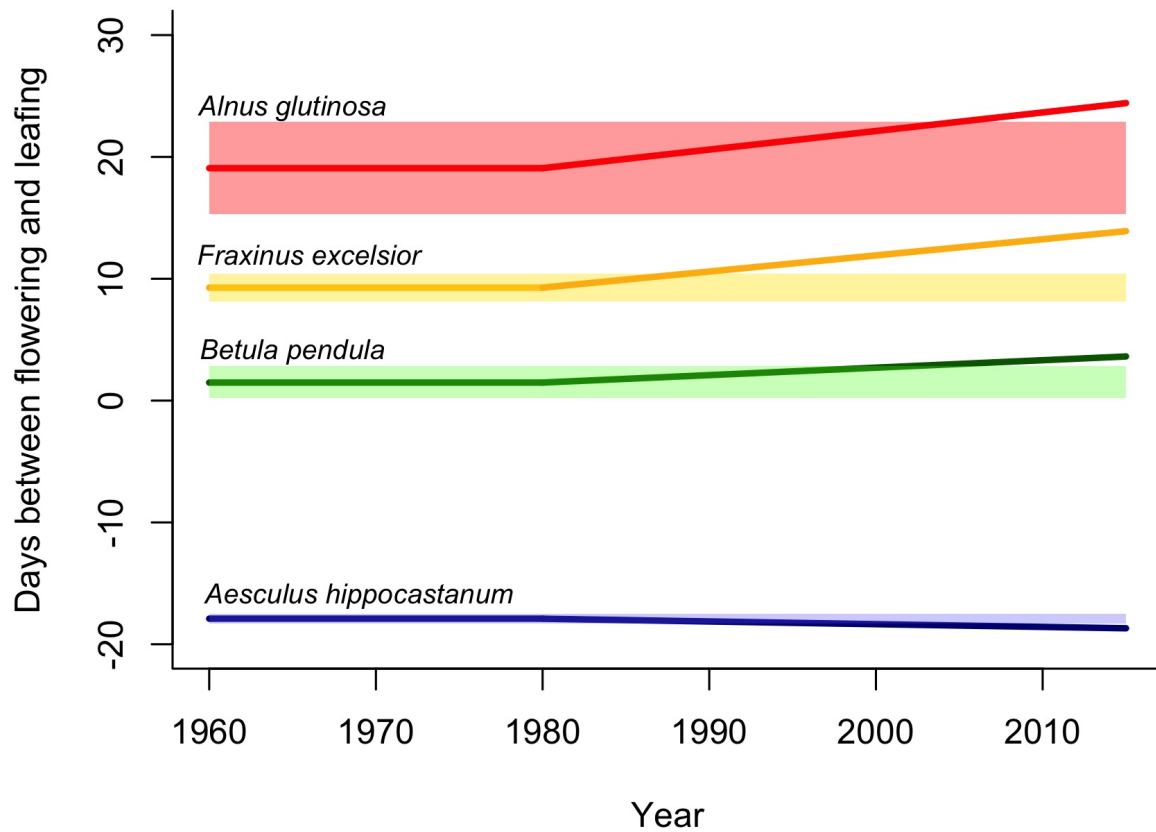
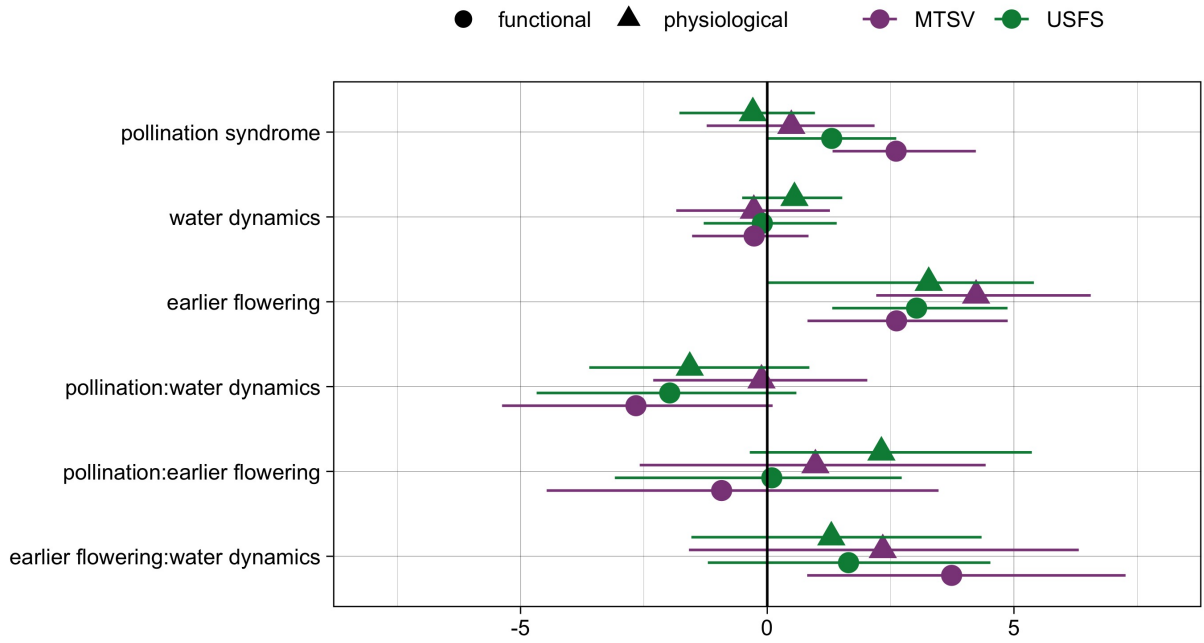


Figure 1: **FLSs across Europe for four tree species from 1960 to 2015 suggests climate change has generally increased the time between flowering and leafing**, but the direction and rate of change differs across species, which may exacerbate fitness differences within forest communities. To detect the effect of climate change on average FLS, we used models that allow for shifts in FLS after 1980. Lines represent the mean trend in FLS per species, and the highlighted regions indicate historic range of FLS variability (95% credible intervals of the pre-1980 average) from the PEP725 database (?).

a)



b)

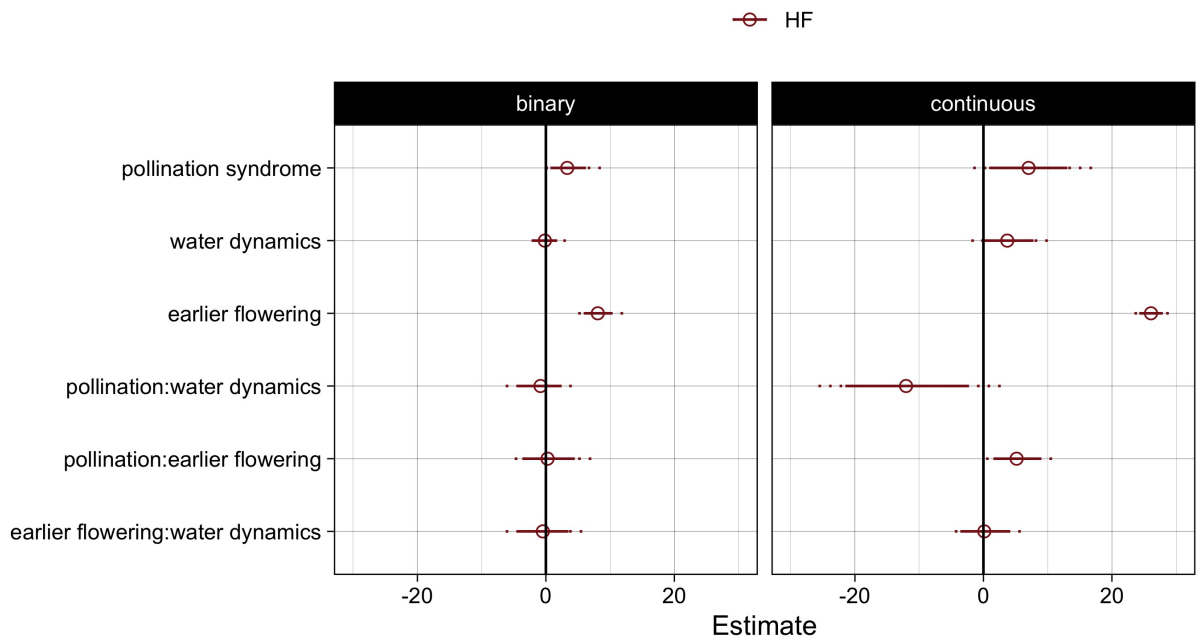


Figure 2: **Estimated effects of water dynamics (minimum precipitation across species range), pollination syndrome, and earlier flowering time on FLS patterns across three case studies.** We used phylogenetic adjustments and standardized units to make a basic comparison of datasets of different definitions of of FLS and data types (categorical and continuous).

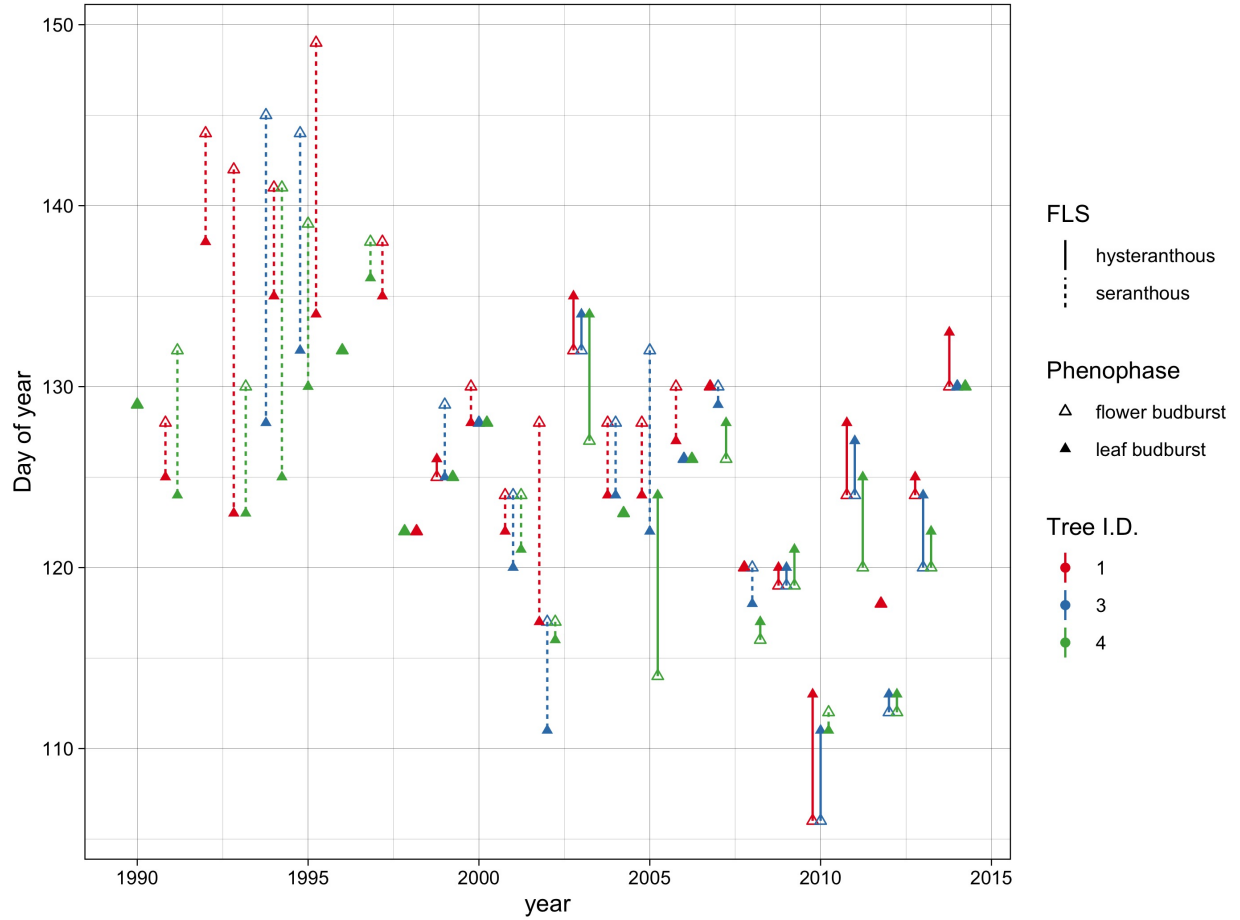


Figure 3: **Individual FLS variability over time for *Quercus rubra* at Harvard Forest.** While this species is typically classified as synanthous, we see here that the order of flower and leaf bud break, and the time between these events varies considerably for each individual over time, and between individuals in any given year. None of this variation can be accounted for in a categorical FLS classification system.

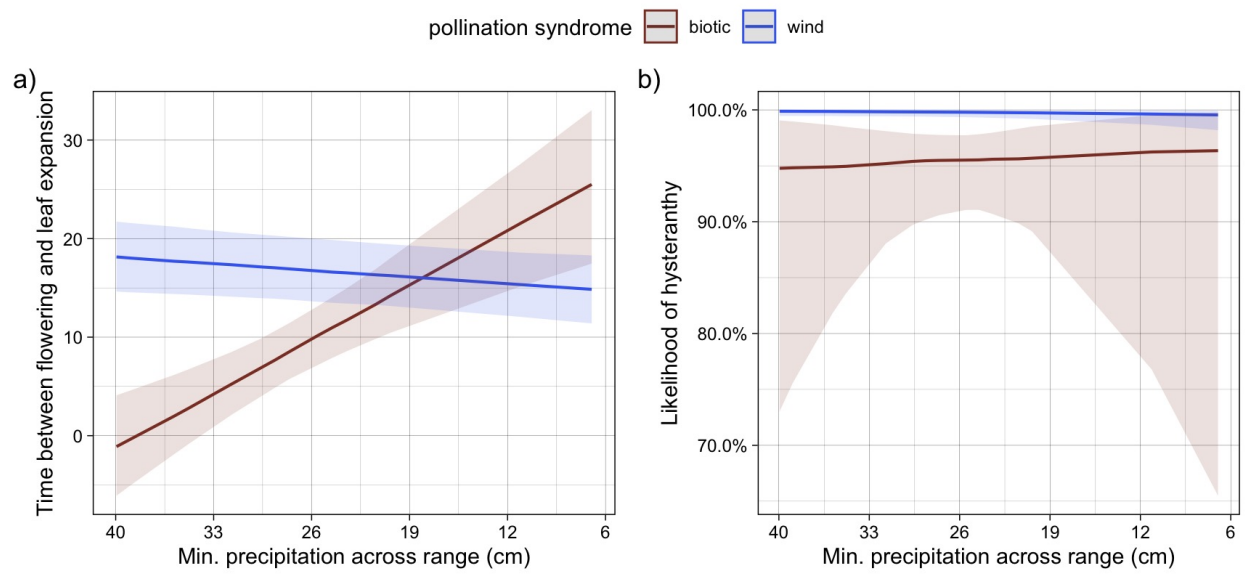


Figure 4: **Average predictive comparisons suggest that water dynamics may be a driver of hysteresis in biotically-pollinated but not in wind-pollinated taxa, suggesting future work on FLS must accommodate overlapping hypotheses.** For species flowering in mid-May (the 25 year community average) these systematic differences