

Why other studies are weak and available data

1. Direct test of these hypotheses in the literature are are, and support for them is mixed
 - (a) Studies tend to test single hypotheses, or present the expectation that hypotheses are mutually exclusive.
 - (b) Even the recent paper by Gougherty and Gougherty which considers multiple hypotheses models them separate (is this totally true? They also used a random forest model).
 - (c) Variability in FLS is generally ignored.
 - (d) Yet, there are datasets widely available that would allow for testing this hypotheses at multiple levels.
2. To address this gap, we supplement our literature review with re-testing some previously-used datasets to examine all hypotheses, and we leverage several widely-available datasets to test how support for these hypotheses varies across the inter-to intraspecific levels.
 - (a) Michigan Trees and its companion volume Michigan trees shrubs and vines (MTSV) contains FLS information for 195 Woody plant species. The USFS Silvics manual volume II contains FLS descriptions for 81 woody species. These data can be used to test interspecific FLS variation. Species are categorized as flowers before leaves, flowers before/with leaves and flowers with leaves, and flowering after leaves. As in previous work, we collapsed these categories to binary "hysteranthous" of "seranthous" for modeling.
 - (b) But even within this frame work, we built back in variability associated with the hypotheses. We defined 'functional hysternathy' to accommodate a degree of overlap between flowering and the early stages of leaf out as predicted by the wind pollination hypothesis (including 'flowers before leaves', 'flowers before/with leaves' and 'flowers with leaves') versus 'physiological hysternathy' (only flowers before leaves) which relates to drought tolerance hypothesis.
 - (c) Harvard Forest contains flowering and leaf phenology measurements for individuals 24 woody species over 15 year, allowing for both inter- and intra-specific comparisons. In this dataset we calculated "FLS offset" (flower DOY-leaf DOY), a continuous measure fo FLS, which allows us to better understand the impact of FLS categorization and make comparison even within categories. We approximated "functional" and "physiological" hysternathy in this data set by calculating 2 offset values (Fopn-L75) and (fbb-lbb).
 - (d) From the PEP725 database we obtained spatially and temporally explicit flowering and leaf phenology for 3 European hysternanthous species. This allows for

test only at the intra-specific level, but unlike the other datasets allow for population level variability to be assessed.

What we know, what we don't know, and what we can do to know more

1. In considering each dataset separately and in tandem, two clear trend emerge:
 - (a) Across inter-specific models, multiple hypotheses were supported. There was generally strong evidence for the early flowering and wind pollination hypothesis, poor support for the water dynamics hypothesis, and the phylogenetic signal was variable.
 - (b) This is not a surprise. We wouldn't expect the wind pollination hypothesis to explain hysteresis in biotically pollinated taxa. And geologically speaking, the relatively recent post glaciation assembly of North Eastern Forests, we have taxa with radically different bio-geographic histories.
 - (c) The relative importance of each the predictors changed significantly depending on how hysteresis was defined. This effect was minimized when continuous measure of FLS were used over categorical.
2. But perhaps more important that the results of these specific model themselves, is that through considering them together, we are provided a more comprehensive picture of where our understanding of this phenological trait is. and where it needs to go.
3. First, our analysis reveals the clear advantages of continuous data.
 - (a) As mentioned above, it minimizes the observer bias that comes with categorization.
 - (b) It reveals important interspecific differences that are masked by categorization. I.e. two hysteresis species may have dramatically different FLS offsets.
 - (c) It also reveals that there are large intra-specific difference which, as will be discussed more below, will be instructive to hypothesis testing.
 - (d) All and all, our work shows categorizing hysteresis is biased and biologically problematic, and future studies about phenological sequences should abandon these categories and treat FLS as continuous traits.
4. Because the trait modeling in large community level datasets seem to support multiple hypotheses, and are confounded by species' identities and observer bias, the utility of these data can only take us so far. A main outgrowth of our model is the realization it is instructive to test questions of hysteresis at other scales.

- (a) One option is to look within the hypotheses to address address subgrouping of taxa in which overlap between hypotheses could be controled?
 - (b) For example What drives hysternanthy among biotically pollinated taxa? It certain isn't wind pollination efficiency. Or, What traits accounts for variability in hysternanthy among wind polliated taxa?
 - (c) Further incorporating a phylo-biogeographic approach would probably be instructive at this level, for example are their phylogeopraphic commonalities between biotocally pollinated hysternanthous species in Eastern flora.
- (a) But even with sub-groupings, interspecific trait association models can only can take us so far.
 - (b) One reality of these kind of studies is that we never know we are picking the right traits. For example we used minimum P across range, one of the only available quantitative drought metrics at the scale of large interspecific models, to represent the water dynamics hypothesis. Is this really a good proxy for drought tolerance?
 - (c) Further, species evolve a suit of traits of any function, and unmeasured traits might bias our results. IE wind pollinated species could compensate for a lack of hysternanthy by over producing pollen or selfing. To really understand this trade across large taxonomic space, you would have to compare species over N-dimensional trait axis.
5. Considering hysternanthy variation at the intraspecific level overcomes many of these limitations and is the the next frontier in understanding FLS.
- (a) As predicted by evolutationalary theory the agreement between our intra and interspecific models, suggesting flowering time major driver in FLS variation, suggests that we are in the right direction.
 - (b) Further, though our datasets were taxonomically and geographically limited, they demonstrate that FLS variability is significant over time and space, .
 - (c) Looking within species holds most other traits relatively equal, avoids the problem of latent tradeoffs with unmeasured traits.
6. This allows us to move into the next state to better understand this trait, not look at the traits associated with FLS variation, but with the consequences of FLS variation through experimental manipulations and observations. you couldn't do this at the inter specific level, but you can at intra. For example.
- (a) Do years or population with more hysternanthy experience greater pollination success?

- (b) Do water stressed populations alter their FLS?
- 7. Looking at consequences will not only help clarify basic scientific hypotheses, but is important for global change because this whole thing is motivated by understand how changing hysteresis will impact species fitness.
- 1. Things I didn't sneak in but could be considered
 - (a) With the strength of flowering time across models should be abandon think of hysteresis outside of selection for early flowering?
 - (b) Community matters ie wind pollination
 - (c) other: I realize that in some way I am not suggesting we test multiple hypotheses at once. I am more suggesting we allow them to coexist by subgrouping our taxa