

# Presence-only data are currently insufficient for broad-scale phenological analyses for most species of temperate butterflies

Elise Larsen<sup>1</sup> and Vaughn Shirey<sup>1</sup>

<sup>1</sup>Georgetown University

April 28, 2020

## Abstract

Large occurrence datasets provide a sizable resource for ecological analyses, but have substantial limitations. Phenological analyses in Fric et al. (2020) are flawed due to insufficient data and improper analysis. Our reanalysis of 22 univoltine species with sufficient data (from 100 original species) found substantive differences in macroscale phenological patterns.

## Main Body:

The aggregation of large datasets from museum records and community science provides new power for macroscale ecological analyses. However, “with great power must come great responsibility” (Lee 1962) and while valuable, these datasets must be treated carefully. Prior work with aggregate occurrence datasets has demonstrated spatiotemporal and taxonomic biases that must be addressed (Troudet *et al.* 2017; Ries *et al.* 2019), yet researchers still utilize these data “off the shelf” without proper curation or appropriate modelling strategies.

Fric et al. 2020 used occurrence data for 100 species aggregated in the Global Biodiversity Information Facility (GBIF) in temperate regions of North America and Europe to track phenology across latitudes. Estimating phenology metrics and trends from large occurrence datasets is possible, but requires sufficient data density and appropriate statistical methods (Taylor and Guralnick 2017). These data were too sparse to estimate phenology across the axes of interest, and were insufficiently curated. Further, an improper analytical model resulted in spurious patterns which don’t make ecological sense. For instance, after “correcting” for altitude and year, their results indicate that flight period onsets are the same at low and high latitudes for most species, contrary to expected patterns (Karlsson 2014). However, when we applied appropriate data curation and methods, our results found later onset and shorter flight periods at higher latitudes.

The data used in Fric et al. (2020) were insufficient for broad phenological analysis. Datasets, defined by species and continent, were analyzed with as few as 15 occurrence records across >20 degrees latitude and >100 years, a threshold far below any reasonable standard. Phenological “onset” and “termination” of flight periods were extracted simply as the first and last day-of-year (DOY) values within latitudinal bands, pooled across all years and altitudes. Even with sufficient data this would be problematic; here it resulted in only one observation date being used as both “onset” and “termination” of flight periods in an average of 20% of latitudinal bands per species (Figure 1), so most analyses included latitudes with flight periods of just one day.

The data in Fric et al. (2020) were also improperly curated. Spatial precision and outlier detection were incomplete. Altitudes were extracted using imprecise GIS coordinates, sometimes representing sea floor or vague place names (eg, “Mt Shasta”); other altitudes varied significantly from GBIF “verbatim elevation”. Across datasets, altitudes were skewed (95.6% of observations in species’ lowest quartiles), giving high altitude observations outsized leverage in regressions. Temporal outliers were problematic; one species’ onset at 68° N

was in January, when the next occurrence across all latitudes was in June. No sources were cited for species traits, and we found evidence documenting additional generations in portions of their range for 22 of these “obligate univoltine” species (Supplemental Table 1).

Finally, the regressions used in Fric et al. (2020) were not appropriate for their questions. The phenometrics were regressed against latitude, altitude, and year separately, without regard for spatiotemporal bias. Regressions were considered “corrections” and residuals were fed into new regressions, which failed to consider collinearity in the explanatory variables. In single regressions with latitude, 12 species’ onset appeared earlier at higher latitudes and only 26 species showed later onset at higher latitudes. Meanwhile “corrected” regressions found onset for most species unaffected by latitude (Figure 2).

In our reanalysis, we applied stricter data standards and curation to extract robust phenometrics. For 72 species (76 datasets) we confirmed as univoltine, we filtered data for altitude (0-500m) and timing (March-November). We calculated phenometrics for year-latitude combinations with at least 10 observations. Fifty-four datasets did not include at least three latitudes with data fulfilling these criteria, making them insufficient for analysis (Supplemental Table 1). For the remaining 22 datasets, all European, onset and termination were estimated from a weibull distribution using R package *phst* (Pearse 2017) and bounded by days (60,330). We modeled each species phenometric in a linear model ( $DOY \sim \text{latitude} + \text{year}$ ) using R version 3.6.2 (R Core Team 2019).

Species-specific coefficients and significance for onset & termination dates in our reanalysis varied substantially from the original results (Figure 2; Supplemental Table 2). In our model, higher latitude generally correlated with later onset (15/22 species), while correlations were mixed for termination date (Figure 2). These results indicated later and/or shorter flight periods at higher latitudes. Year was usually nonsignificant, with most significant patterns consistent with longer flight periods in recent years.

Research using large occurrence datasets with variable effort must be explicit about how data have been obtained and processed prior to analysis. Despite our critique of the study in question, we recognize that occurrence data have great potential to address many ecological questions. We enthusiastically support continued digitization and use of collection data in ecological analysis, but urge researchers to exercise caution using these data, especially when estimating effort-sensitive metrics, such as phenology.

## Acknowledgements

We appreciate the many collections, programs, and citizen scientists contributing data to GBIF which were used here. We are grateful to Leslie Ries and Mariana Abarca for providing comments improving this manuscript. We thank the original authors for openly discussing their work with us. This work was supported by the National Science Foundation Award DEB-1702708. V.S. received support from Georgetown University.

## References

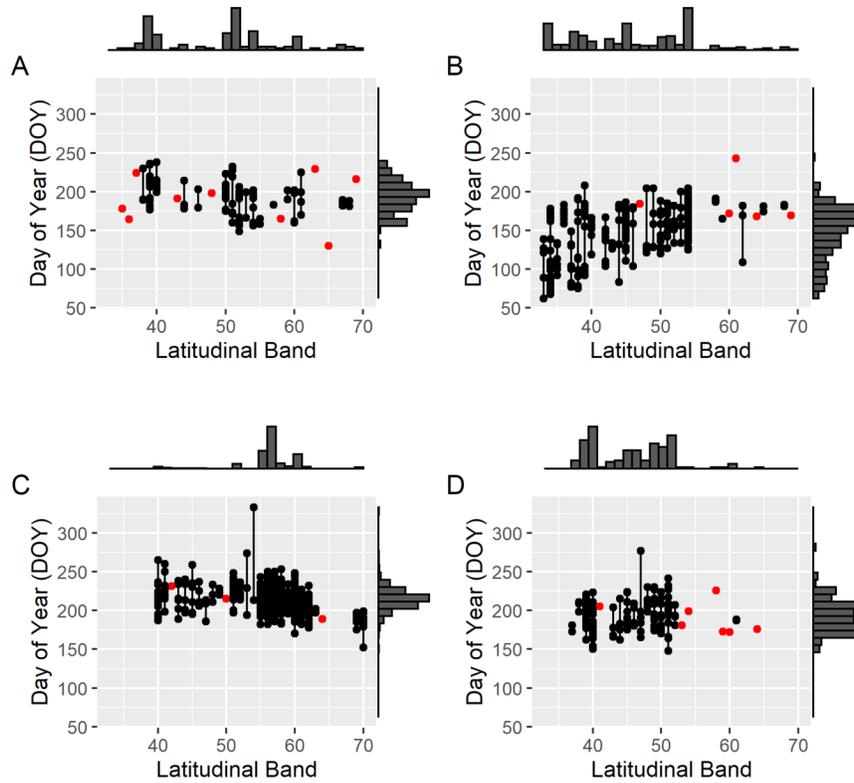
- Fric, Z. F., Rindoš, M., & Konvička, M. (2020). Phenology responses of temperate butterflies to latitude depend on ecological traits. *Ecology letters*, 23(1), 172-180. <https://doi.org/10.1111/ele.13419>
- Karlsson, B. (2014). Extended season for northern butterflies. *International journal of biometeorology*, 58(5), 691-701. <https://doi.org/10.1007/s00484-013-0649-8>
- Lee, S. (1962) *Amazing Fantasy #15*, Marvel.
- Pearse, W. D., Davis, C. C., Inouye, D. W., Primack, R. B., & Davies, T. J. (2017). A statistical estimator for determining the limits of contemporary and historic phenology. *Nature Ecology & Evolution*, 1(12), 1876-1882. <https://doi.org/10.1038/s41559-017-0350-0>
- R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>
- Ries, L., Zipkin, E. F., & Guralnick, R. P. (2019). Tracking trends in monarch abundance over the 20th century is currently impossible using museum records. *Proceedings of the National Academy of Sciences*,

Taylor, S.D., & Guralnick, R.P., 2019. Opportunistically collected photographs can be used to estimate large-scale phenological trends (preprint). *Ecology*. <https://doi.org/10.1101/794396>

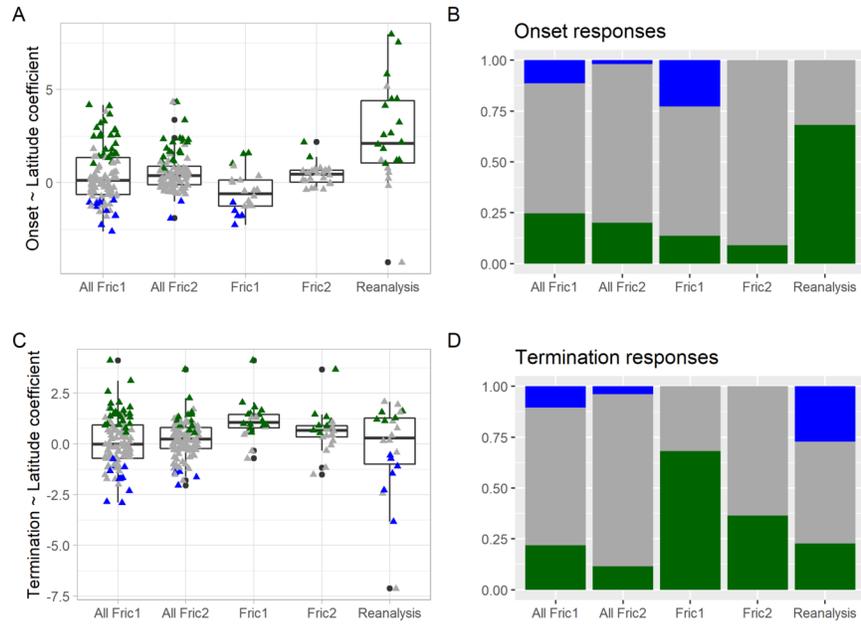
Troutet, J., Grandcolas, P., Blin, A., Vignes-Lebbe, R., & Legendre, F. (2017) Taxonomic bias in biodiversity data and societal preferences. *Nature Scientific Reports*, 7. <https://doi.org/10.1038/s41598-017-09084-6>

**Figure 1:**

page 1 of 1



**Figure 2:**



**Figure 1 Caption:**

Raw occurrence data used in original analysis and flight period lengths by latitude for 4 datasets (A: *Agriades glandon* in N. America, B: *Glaucopsyche lygdamus* in N. America, C: *Hesperia comma* in Europe, D: *Parnassius smintheus* in N. America). Results for these species were presented in Fric et al. (2020) Figure 1 to demonstrate varied phenological patterns. Red dots represent observations used as both onset and termination in the original analysis. Marginal histograms show data density across latitudes (above the plot) and time of year (to the right of the plot).

**Figure 2 Caption:**

Comparison of model parameters for onset (A, B) and termination (C, D) dates as a function of latitude. Panels A and C show values and boxplots of the model coefficients for  $DOY \sim \text{latitude}$ , colored by response: positive (green), non-significant (gray), or negative (blue) correlations. Panels B and D are stacked barplots of the response signs representing the proportion of species demonstrating positive (green), non-significant (gray), or negative (blue) correlations with latitude. Each panel shows five sets of model results. “All Fric1” represents all 105 datasets in the Fric et al. single regressions for  $DOY \sim \text{latitude}$ , while “All Fric2” represents all 105 datasets in the Fric et al. (2020) “corrected” regressions  $\text{Res}(DOY \sim \text{altitude, year}) \sim \text{latitude}$ . “Fric1” and “Fric2” correspond to the model results above, showing only results for the 22 species-region datasets used in our reanalysis. “Reanalysis” shows results from our curation and reanalysis of those 22 species-region datasets.