Reliable estimation of butterfly flight period phenology from occurrence records

Elise Larsen[1] and Vaughn Shirey[1]

[1]Georgetown University

April 05, 2024

Abstract

Previous phenological analyses of GBIF records ignored data biases and produced misleading results. Here we demonstrate how data sufficiency and methodology affect inferred latitudinal patterns in lepidopteran phenology, highlighting the importance of spatiotemporal data aggregation decisions. Results indicate later and/or shorter flight period phenology at higher latitudes is common.

Reliable estimation of butterfly flight period phenology from occurrence records

Running title: Methods for phenology analysis of GBIF data

Elise A. Larsen[1], Vaughn Shirey[1]

[1] Georgetown University, Department of Biology; Washington, DC

Email: Elise A. Larsen: eal109@georgetown.edu; Vaughn Shirey: vms55@georgetown.edu

Keywords: phenology, phenometric, onset, termination, GBIF, presence-only, bias, observation process, extreme value, spatiotemporal bias, observation bias

Type of article: technical comment

Number of words in abstract: 50 (of 50)

Number of words in main text: 800 (of 800)

Number of references: 12

Number of figures: 2 main text

Number of tables: 0 main text

Corresponding author:

Elise A. Larsen, eal109@georgetown.edu, 412-519-4426 (phone), 202-687-5662 (fax)

Box 571229 Reiss Science Bldg., Room 406, 37th and O Streets, NW Washington DC 20057

Statement of authorship: EAL and VS collaboratively developed ideas for this manuscript. EAL conducted the analysis with contributions from VS. EAL wrote the manuscript, and both EAL and VS revised the manuscript.

Abstract:

Previous phenological analyses of GBIF records ignored data biases and produced misleading results. Here we demonstrate how data sufficiency and methodology affect inferred latitudinal patterns in lepidopteran phenology, highlighting the importance of spatiotemporal data aggregation decisions. Results indicate later and/or shorter flight period phenology at higher latitudes is common.

Main Body:

Larsen & Shirey 2021 (ELE13731) identified significant concerns with the phenological analysis in Fric et al. 2020 (ELE13419). In response, Fric et al. 2021 (ELE13739) presented three arguments: (1) ELE13731 switched onset and termination results in the reanalysis of ELE13419 data and when corrected, the reanalysis supports the original conclusions, (2) removing data with known errors doesn't change overall results, and (3) the benefits of including more species for life history trait analysis justifies analysing species with sparse data. These arguments fail to confront the spatiotemporal bias in the underlying data (Figure 1). ELE13739 reported that most species demonstrate later termination at higher latitudes and do not have detectable shifts in onset across latitudes. These spurious patterns resulted from coarse data aggregation and spatiotemporal variation in recorder activity (Steel et al. 2013, Isaac et al. 2014, Knape et al. 2021)

The first argument in ELE13739 is not supported by the evidence. Species-specific visualizations in ELE13731 demonstrate non-significant and positive correlations between latitude and species-specific onset, with more varied latitudinal patterns in termination (ELE13731 Figure S1). The ELE13739 reanalysis addressed data curation but ignored other ELE13731 methods; ELE13731 analytical code has been validated and results independently reproduced (Campbell and Belitz 2021).

The second argument dismissed the importance of data curation, claiming "stringent reanalysis" did not change results significantly. However, comparing glm and lmer results in ELE13739, onset and termination responses differ for 11 and 23 species/regions (of 88) respectively. Here, we quantify changes in inferred phenological trends attributable to data curation, data aggregation, phenometric estimation, and effort corrections (Supplement 1). Applying mixed effect models with random intercepts for year and species/region ("population"), we find that inferred latitudinal patterns are sensitive to methodology, particularly data aggregation, for both onset (35 populations) and termination (16 populations). Overall, we identify later onset at higher latitudes for 35 of 60 univoltine populations with annual phenometrics in at least two latitudes, and 17 of 18 populations with 10+ latitudes (Figure 2). We find no evidence of earlier onset at higher latitude for any population, refuting ELE13739 results. Latitudinal trends are rarely detected for termination metrics, which are earlier or later at higher latitude for 5 and 11 populations, respectively. Overall, different minimum data thresholds and aggregation decisions drive the significant differences between ELE13739 and ELE13731 results.

The third argument in ELE13739 presents a challenge for the research community as unstructured occurrence data become widely used. There is a trade-off between maximizing the research scope, and ensuring the quality of data input into models. Researchers must consider how variable sampling informs scope, methods, and interpretation of results, and determine when data are sufficient for analysis. These decisions are difficult; disagreement among research groups about best practices is expected. However, our analysis demonstrates that improper data aggregation and small sample size can lead to unreliable conclusions. Methodological differences are discussed further in Supplement 2.

Many factors contribute to spatiotemporal bias in occurrence data. Separating sampling factors (eg, observer density) from ecological factors (eg, longer phenological duration) is challenging; data handling and analytical decisions determine research outcomes. ELE13739 estimated phenometrics from as few as 2 occurrences, and used one pair of phenometrics for each latitudinal band in each population, removing stratification across years and elevations. Should ELE13739 results be taken at face value, the plurality of studied species experience longer flight periods at higher latitudes. This result contradicts the long-established understanding of lepidopteran thermal physiology, and does not hold up to scrutiny. ELE13731 demonstrated different phenological patterns, having aggregated within years and required 10 occurrences; this included fewer species

and latitudes, but more phenometrics per latitudinal band (Supplement 1). While previous analyses of these data did not account for observer effort, the effort metrics we include rarely affect inferred phenological responses to latitude. Misleading results in ELE13739 resulted primarily from data limitations and overly coarse data aggregation.

Recent accessibility of aggregated occurrence data requires careful consideration of how to avoid spurious patterns which may arise from data biases. Spatiotemporal variation in recorder activity in aggregated occurrence data may lead to misleading results (Isaac et al. 2014, Amano et al. 2016, Larsen and Shirey 2021). To disentangle observation patterns from ecological patterns and avoid spurious results, analyses must consider underlying observation processes (Tweedie et al. 1994). Fortunately, there is a growing body of work on how to identify and mitigate issues related to observer bias in aggregated data (Isaac et al. 2014, Shirey et al. 2021, Van Eupen et al. 2021). Similarly, several approaches have been developed for estimating robust phenometrics from occurrence data (eg, Kharouba et al. 2014, Chapman et al. 2015, Inouye et al. 2019, Belitz et al. 2020, Cima et al. 2020). With the continuing development of robust methodologies, occurrence data have great potential to expand the taxonomic and spatial domains of ecological knowledge.

References:

Belitz, M. W., Larsen, E. A., Ries, L., & Guralnick, R. P. (2020). The accuracy of phenology estimators for use with sparsely sampled presence-only observations. *Methods in Ecology and Evolution* ,*11* (10), 1273-1285. https://doi.org/10.1111/2041-210X.13448

Campbell, C. J., & Belitz, M. W. (2021) LE_Pheno2021. *Github repository* .*https://github.com/cjcampbell/ELE_Pheno2021*

Edwards, C., & Crone, E.E. (2021) Estimating abundance and phenology from transect count data with GLMs. Preprint:*https://doi.org/10.1101/2020.06.01.127910*

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*

Fric, Z. F., Rindoš, M., & Konvička, M. (2021). Radical pruning of distribution data may result in loss of knowledge (Response to Larsen & Shirey). Ecology letters.*https://doi.org/10.1111/ele.13739*

Inouye, B. D., Ehrlén, J., & Underwood, N. (2019). Phenology as a process rather than an event: from individual reaction norms to community metrics. *Ecological Monographs* , *89* (2), e01352. https://doi.org/10.1002/ecm.1352

Isaac, N. J., van Strien, A. J., August, T. A., de Zeeuw, M. P., & Roy, D. B. (2014). Statistics for citizen science: extracting signals of change from noisy ecological data. *Methods in Ecology and Evolution* , *5* (10), 1052-1060. https://doi.org/10.1111/2041-210X.12254

Knape, J., Coulson, S.J., van der Wal, R, & Arlt, D. (2021) Temporal trends in opportunistic citizen science reports across multiple taxa. Ambio https://doi.org/10.1007/s13280-021-01550-w

Larsen, E. A., & Shirey, V. (2021) Method matters: pitfalls in analysing phenology from occurrence records. Ecology Letters. https://doi.org/10.1111/ele.13731

Shirey, V., Belitz, M. W., Barve, V., & Guralnick, R. (2021). A complete inventory of North American butterfly occurrence data: narrowing data gaps, but increasing bias. *Ecography* . https://doi.org/10.1111/ecog.05396

Steel, E. A., Kennedy, M. C., Cunningham, P. G., & Stanovick, J. S. (2013). Applied statistics in ecology: common pitfalls and simple solutions. *Ecosphere* , 4(9), 1-13.

Tweedie, R. L., Mengersen, K. L., & Eccleston, J. A. (1994). Garbage in, garbage out: can statisticians quantify the effects of poor data?.*Chance* , 7 (2), 20-27.*https://doi.org/10.1080/09332480.1994.11882492*

Figure 1: Patterns between sample size and latitude, onset, and termination for population (species-region) subsets of ELE13739 data. In all panels, separate lines/colors are different populations. Panel (a) shows the variation in data density across latitudes [log(n)~lat.band:population + (1|population) has marginal r^2 of 0.35 and slope is significantly positive for 14 species.] Panels (b) and (c) show ELE13739 onset and termination DOYs respectively, according to data density [onset marginal r^2=0.03 (19 negative, 62 non-significant population-specific coefficients); termination marginal r^2=0.02 (14 positive, 67 non-significant population-specific coefficients).]
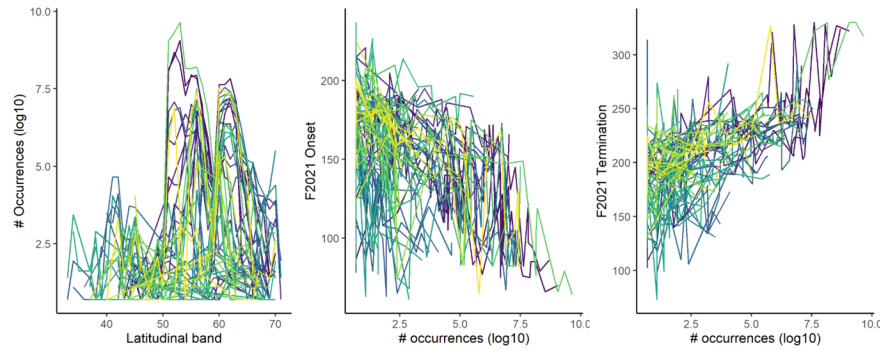

Figure 2: Population-specific coefficients for onset~latitude when data aggregation is latitude*year as related to the number of latitudes with latitude*year onset metrics for that population (requiring at least 5 occurrence records in a given latitudinal band and year). Coefficients for extreme value metrics (circles) and weibull metrics (Xs) are colored such that green represents coefficients with p[?]0.05 and gray represents p>0.05. Model results for data curations 1 & 2, extreme and weibull metrics, and models with and without effort metrics are shown for each population. From 29 populations in which at least 5 latitudes are represented in population phenometrics, onset is significantly later at higher latitudes for 22 (extreme value metrics) or 24 (weibull metrics) populations.

Latitudinal patterns by # latitudes with data