A test of six simple indices to display the phenology of butterflies using a large multi-source database

Valentina Cima, Benoît Fontaine, Isabelle Witté, Pascal Dupont, Martin Jeanmougin, Julien Touroult

A test of six simple indices to display the phenology of butterflies using a large multi-source database

Valentina Cima[1], Benoît Fontaine[1,2], Isabelle Witté[1], Pascal Dupont[1], Martin Jeanmougin[2], Julien Touroult[1]

[1] UMS Patrimoine Naturel (PATRINAT), AFB, MNHN, CNRS, CP41, 36 rue Geoffroy Saint-Hilaire 75005 Paris, France

[2] Centre for Ecology and Conservation Sciences (CESCO UMR7204), MNHN-CNRS-Sorbonne University, 55 rue Buffon, 75005 Paris, France

Corresponding author: valentina.cima@mnhn.fr; Tel.: +33 (0) 1 71 21 32 58

Abstract

Biological recording at broad temporal and spatial scales produces large volumes of species occurrence data. Multi-source datasets, which include opportunistic records, are unstructured and contain bias, mainly due to uneven and unknown observation effort, but they also provide meaningful information about species phenology. Butterflies are well known and well represented in citizen-science programs and national inventories, which makes them an interesting case for phenological studies. This work aims to find a simple, flexible, fast-rendering phenology index, which has to prove reliable when compared to standard knowledge. Six indices (two non-corrected and four corrected for observation effort) were built and implemented on butterfly records. They were analysed against blind expert opinion and a set of monitoring data. Surprisingly, all indices produced mostly realistic phenological patterns and non-corrected indices were as good as corrected ones. The number of species records divided by the number of records of all species of the group collected during the same period is the only index that should be avoided, because of an over-correction of recording intensity. Additional work is needed, in particular to refine the analysis by testing the sensitivity of the index to the amount of data, as well as by employing statistical models that are also useful for exploring trends and seasonal shifts.

Keywords: opportunistic data; citizen science; Lepidoptera; flight period; seasonality; bias correction; sampling effort

1 Introduction

The rise of biological recording schemes including broad-scale citizen-science programs has brought new possibilities to conservation and ecological research over the last decades, producing large amounts of species occurrence data (Dickinson et al. 2012, Hochachka et al. 2012, Tulloch et al. 2013, August et al. 2015, Pocock et al. 2015). When gathered into datasets that cover large temporal and spatial extents, these multi-source data (a combination of opportunistic and systematic records) may help unveiling important aspects of biodiversity state and changes (Dickinson et al. 2010, Dickinson et al. 2012, Hochachka et al. 2012, Isaac and Pocock 2015, Powney and Isaac 2015), including species phenology. However, multi-source data present different levels of standardisation (depending on the source of collection) and are by nature noisy and unstructured. They suffer from several biases (Dickinson et al. 2010, Robertson et al. 2010, Isaac et al. 2014, Isaac and Pocock 2015) which primarily relate to variation in recording intensity (Isaac et al. 2014). In fact, sampling effort may vary throughout the year, between the years and among regions and this variability is usually unknown in opportunistic datasets (Giraud et al. 2016).

Phenology is the study of periodic biological events (such as plant flowering, insect emergence and bird migration) that are regulated by environmental factors. The simplest way to represent animal phenology is by counting the number of species occurrences per period, collating all years' data. This kind of representation is sometimes employed in broad distribution Atlases (some examples: Lumaret, 1990 for dung beetles; the iNaturalist platform of crowdsourcing of data, for instance https://www.inaturalist.org/taxa/207977-Aglais-io; the Atlas of butterflies and zygens of Midi-Pyrénées, http://atlaspapillonsmidipyrenees.myspecies.info/) or in more specific studies (such as Bertone et al. 2005, Pozo et al. 2008, Archaux et al. 2011). Other works have reported phenology as weighted or mean counts (van Swaay et al. 2002, Archaux et al. 2015, Manil et al. 2015) as well as modelled counts (Dennis et al. 2013, Schmucki et al. 2016), sometimes accounting for imperfect detection or uneven recording intensity (Strebel et al. 2014).

Butterflies are relatively well known and well represented in citizen-science programs and national inventories. For this reason, they have frequently been the subject of studies based on monitoring and citizen-science data (see for example: Maes et al. 2012, Dennis et al. 2013, Schmucki et al. 2016). Here, we focus on the representation of butterfly phenology patterns with multi-source opportunistic records while accounting for sampling effort. More precisely, this study aims to explore the potential of a national biodiversity reference system, the National Inventory of Natural Heritage of France (INPN - https://inpn.mnhn.fr), for displaying the phenology of butterflies at a national scale. The goal is to develop a phenology index with the following requirements: 1) the phenology is displayed using multi-source data, including opportunistic observations; 2) the index yields overall patterns that are consistent with standard knowledge on species phenology; 3) the index must be suitable for many species and different type of seasonality and voltinism (refer to Wolda 1988 for a precise classification of seasonality patterns); 4) the phenology charts will be presented to a general audience on a web portal, so the index must be simple, easily interpretable and fast-rendering. Ideally, the index should inform on species activity over the year or display a lack of knowledge, inducing the community to collect more accurate occurrence data.

Some studies have highlighted the potential of opportunistic data to perform as well as standardised data (see for example, van Strien et al. 2010, 2013). This is generally true when opportunistic data are corrected for bias (van Strien et al. 2013, Isaac et al. 2014). Measures that simply show the number of observations are sensitive to sampling effort and likely to reflect observer activity (Dickinson et al. 2010), while those that account for temporal and spatial variation in effort should better delineate the true phenology of a species. Under this assumption, six relatively simple indices, based on both non-corrected and corrected measures, were designed with butterfly records from the INPN. The indices were analysed against blind expert opinion and a set of monitoring data. We expected corrected indices to yield more realistic patterns than non-corrected ones. The results should lead to recommendations about selecting an accurate but simple index for displaying phenology with multi-source opportunistic data.

2 Materials and methods

2.1 General analysis process

The analysis started by selecting, with the help of a lepidopterologist (PD), a panel of butterfly species that are well known and well represented in the dataset. Their records come from different data sources (Appendix 1), were collected in several years and are assumed homogeneous over France. In order to assess the versatility of the indices, the selected panel had to be diverse in terms of phenology type, ecology and latitudinal range. Based on literature review, a series of non-corrected (“raw”) and corrected indices were implemented on multi-source butterfly data and evaluated through two different approaches: a qualitative analysis, based on expert evaluation of the phenological patterns drawn by the indices, and a quantitative analysis to seek for a match between the indices and the patterns based on monitoring data. The different steps of the analysis are developed in the next sections and summed up in Fig. 1. All analyses were carried out in R v.3.5.2 (R Core Development Team 2018).

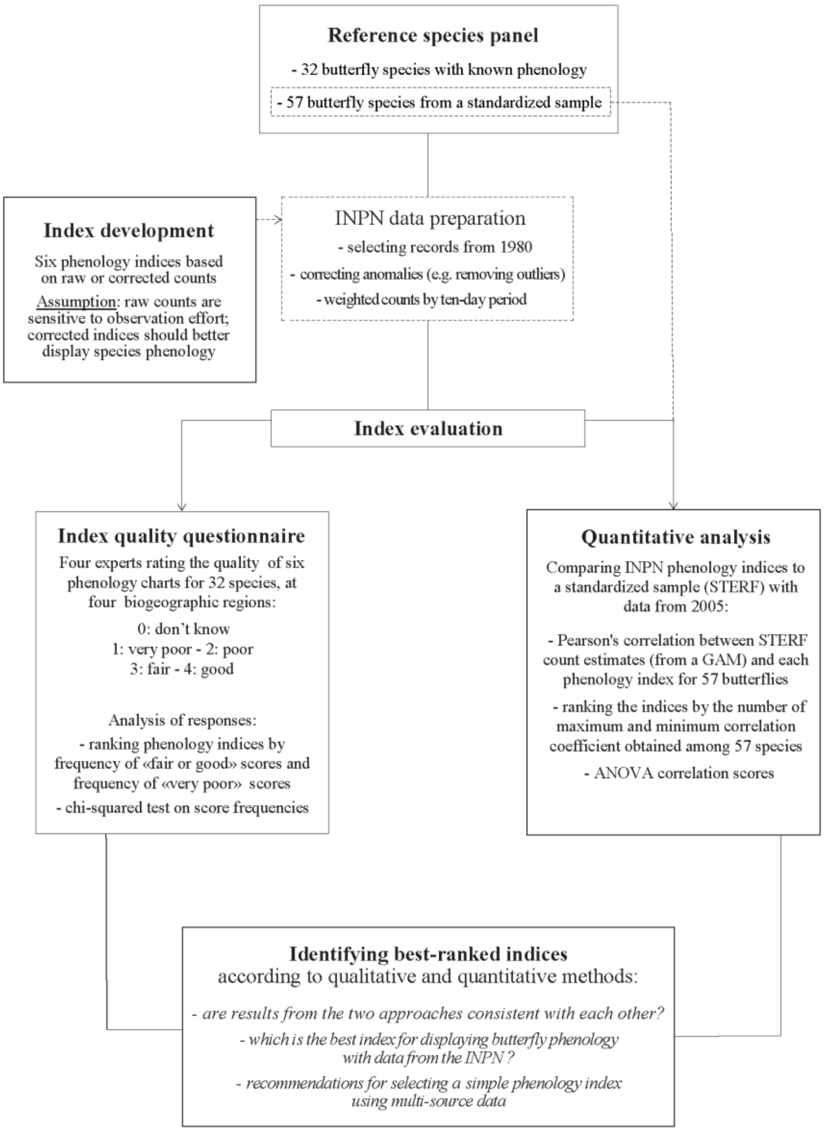


Figure 1. Scheme of the general analysis process for this study.

2.2 Data description

Data were extracted on July 2017 from the National Inventory of Natural Heritage of France (INPN - https://inpn.mnhn.fr). The INPN is a system created by the National Museum of Natural History (MNHN) and managed by MNHN and the National Agency for Biodiversity (AFB) that aims at sharing information and data about biodiversity in France.

The INPN gathers multi-source data from scientific surveys, museum collections, citizen science programs, as well as opportunistic observations. These data are collected, standardised and synthesized in order to develop a national reference bank of biodiversity data. One record in the INPN corresponds to one species occurrence, collected by one observer whatever the number of individuals. A record contains a start and an end date of collection. The interval between the two dates ("temporal resolution") may vary from 1 day ("precise date of observation"), to several days (the duration of an inventory or sampling campaign) until several years, which is the case of observations derived from literature or museum collections with uncertain temporal information. On the other hand, data from several sources may have different spatial resolution. Nonetheless, when stored in the INPN, records are assigned to 10x10 km cells, according to a national reference grid and a standardised method, which allows for reliable spatial information.

Effort correction required data from a background or target group (Ponder et al. 2001, Phillips et al. 2009, Kéry et al. 2010, Ruete 2015). Although we focused on butterfly phenology, we extracted all INPN records of diurnal lepidopterans (i.e. butterflies and diurnal moths, hereinafter referred as the "group", see Table 1) by assuming that survey methods and collector specialties within this group are similar, hence data share similar bias (Ponder et al. 2001, Ruete 2015). The group includes seven families: Pieridae, Papilionidae, Nymphalidae, Lycaenidae, Hesperiidae, Riodinidae, Zygaenidae. The French taxonomic repository TaxRef version 11.0 (Gargominy et al. 2017) was employed for taxonomic references. Assuming that naturalists and collectors have similar knowledge and bias and use analogous survey methods for the group of butterflies and diurnal moths, we defined a "field visit" as a 10x10 cell surveyed by one observer on a date, regardless of the number of species he or she had observed. In this manuscript, we refer to a species "quadrat" as a 10x10 km cell where the species was recorded at least once in a given period. In the same way, one "group quadrat" is a 10x10 km cell where at least one of the species of the group was recorded. The vocabulary used for this study is summarised in Table 1.

2.3 Data preparation and index design

In the last decades, biological recording has intensified (Isaac and Pocock 2015). In France, the mission of the MNHN to centralize information and managing a national reference bank of biodiversity started in 1979. Since those years, the collection of records at a national scale have become more frequent and rigorous, in particular with the development of national inventories and atlases (Touroult et al. 2015). On this basis, we presumed data from

1980 onwards to be more uniform and representative of current overall phenological patterns (e.g. number, position, sharpness of peaks). We restricted, therefore, the analysis to those data. Moreover, 1980-today corresponds to the time span of knowledge of the experts who took part in the survey.

In order to minimize errors, outlier data, such as records on January 1st and December 31th (probable by-default dates when the day or the duration of observation, for some reason, is not known) were discarded *a priori*. Ultimately, data for the whole group of butterflies and diurnal moths consisted in 772,307 records (Appendix 1).

Since records had different temporal resolution, we fixed a temporal unit of ten days for displaying phenological patterns. Assuming that non-precise records are still relevant for outlining overall phenological patterns, we kept all data with a temporal resolution (time interval between start date and end date, Tab. 1) up to 15 days, in order to keep as much data as possible. The uncertainty of dates was compensated with pro rata calculation. First, a record was duplicated or triplicated when its start and end date of collection overlapped two or three successive ten-day periods. Then a pro rata was calculated for each period, according to the number of days covered by the record collection dates. For instance, a record with start date 2015-07-21 and end date 2015-08-03 (temporal resolution: 14-days), is converted to one observation in the last ten-day period of July with a pro rata of 11/14 ≈ 0.78 and one record in the first ten-day period of August with a pro rata of 3/14 ≈ 0.22. Whenever start and end date are within a ten-day period pro rata is 1.

Equation 1 $$\text{WEIGHTED NUMBER OF RECORDS}_k = \sum_{i=1}^{n} prorata_i$$

Where i = 1 ... n records of a species and a ten-day period k, all years combined.

The weighting adjustment was also applied for counting quadrats (weighted number of 10x10 km cells where a species or the group was detected at least once in a ten-day period k) and field visits. For simplicity, all weighted counts are henceforth mentioned as "number of …" (records, quadrats, field visits, etc.).

Table 1. Vocabulary used in this study.

Temporal resolution	Time interval between start date and end date of an INPN record. It may vary from one day (precise date) to several years
Group	A background or target group. Survey methods and collector specialties within a group are assumed to be similar and share similar bias (Ponder et al. 2001, Ruete 2015). Examples of target groups are orchids, ground beetles, dragonflies, bats. This study focuses on "butterflies and diurnal moths", a well-known group for which data are abundant and mostly reliable

Record	Observation or collection data of a species provided by one observer on a precise date or time interval, in one locality (a 10x10 cell), whatever the number of specimens
Quadrat	Spatial unit. A 10x10 km cell where the species was recorded at least once in a period. A group quadrat is a 10x10 km cell where at least one of the species of the group was recorded in a period
Field visit	A unique survey event. One quadrat surveyed by one observer on a date for a group
Species known distribution	Set of 10x10 km cells where the species was recorded at least once since 1980

The number of records of all species of the group ("group records"), the number of group quadrats and the number of field visits (section 2.2; Fig. 2) can be employed as proxy for sampling effort (Lobo 2008, Phillips et al. 2009, Ruete 2015) and used for normalising raw counts (see below).

Six indices (Tab. 2) were built using both species number of records (M1, M2, M3), and number of quadrats (M4, M5, M6). M1 and M4 are raw relative frequencies by ten-day period. The other indices are less intuitive and data preparation requires more computing time to build them. Despite that, they should correct for bias due to uneven recording intensity over time or space (Tab. 2). They were built by normalising the number of species records or quadrats by: the number of field visits (M3); the number of group records (M2); the number of group quadrats (M5); the number of group quadrats within the species known distribution (M6), where "known distribution" is the set of 10x10 km cells where the species was recorded at least once since 1980.

Table 2. Phenology indices designed for this study.

Index	Definition	Formula (for a ten-day period *k*)	Description	Reference work that inspired the design of the index	A priori properties of the index
M1	Proportion of records per period	$\sum_{i=1}^{n} prorata_i / \sum_{k=1}^{36}\sum_{i=1}^{n} prorata_i$	Number of records (1...*n*, see equation 1) for a species in a ten-day period *k* divided by the total number of records of the same species in the dataset	Lumaret, 1990 Archaux et al. 2011 (“number of observations”) van Swaay (1990) (“number of records”)	Advantages: very simple and intuitive. Disadvantages: probably biased by uneven recording intensity (spatial and temporal bias).
M2	Ratio of records to the group	$\sum_{i=1}^{n} prorata_i / \sum_{i=1}^{m} prorata_i$	Number of records (1...*n*) for a species in a ten-day period *k* divided by the number of records of all species of the related group (1...*m*) in the same ten-day period	van Swaay (1990) (”percentage”)	Advantages: simple and intuitive; temporal bias correction. Disadvantages: may be sensitive to reporting bias, emphasizing artefact peaks.
M3	Number of records per field visit	$\sum_{i=1}^{n} prorata_i / \sum_{i=1}^{p} prorata_i$	Number of records (1...*n*) for a species in a ten-day period *k* divided by the number of field visits (1...*p*) in the same ten-day period	Archaux et al. 2015 (“mean abundance by field visit”) Strebel et al. 2014 (“naïve detectability index”)	Advantages: temporal bias correction. Disadvantages: does not correct for spatial bias; less simple and less intuitive.
M4	Proportion of quadrats per period	$\sum_{i=1}^{q} prorata_i / \sum_{k=1}^{36}\sum_{i=1}^{q} prorata_i$	Number of quadrats (1...*q*) where the species was seen at least once in a ten-day period *k* divided by the total number of quadrats where the species was observed	Archaux et al. 2015 (“number of occurrences”)	Advantages: very simple and intuitive; may correct for some bias (e.g. duplicates, quadrat oversampling). Disadvantages: still reflects uneven recording intensity.
M5	Ratio of quadrats to the group	$\sum_{i=1}^{q} prorata_i / \sum_{i=1}^{g} prorata_i$	Number of quadrats (1...*q*) where the species was seen at least once in a ten-day period *k* divided by the number of visited quadrats (1...*g*) in same ten-day period (i.e. quadrats where at least one species of the group was seen)	van Swaay (1990) (”percentage of squares”) Turin and den Boer (1988) (“corrected number of squares”)	Advantages: temporal and reporting bias correction; should limit reporting bias. Disadvantages: does not correct for spatial bias; not too simple; less intuitive.
M6	Ratio of quadrats to the group within the species known distribution	$\sum_{i=1}^{q} \boldsymbol{prorata_i} / [\sum_{i=1}^{g} \boldsymbol{prorata_i}]_{\text{skd}}$	Same as M5, but takes into account only quadrats where the species was found at least once since 1980 (species known distribution: skd)	Kéry et al. 2010 (“detection history”)	Advantages: may correct for temporal and spatial bias. Disadvantages: not too simple and not intuitive.

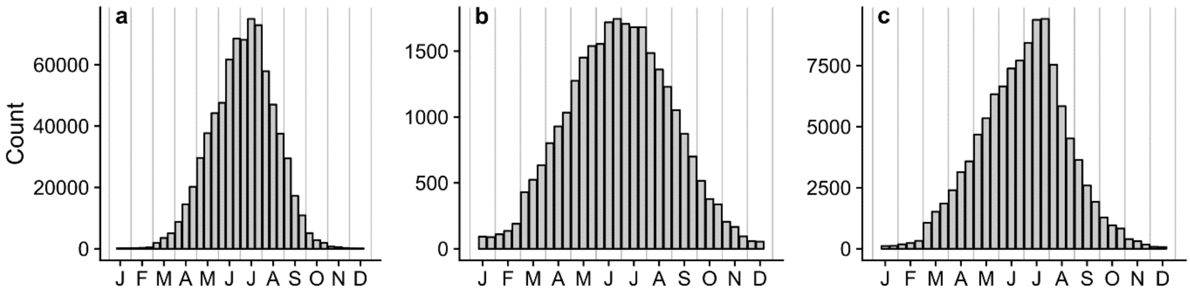


Figure 2. Seasonal distribution by ten-day period of three proxies for observation effort in diurnal lepidopterans (after data preparation). The x-labels indicate months. (a) group records: number of records of all species of the group; (b) group quadrats: number of quadrats where whatever species of the group was seen; (c) number of field visits: number of quadrats surveyed by one observer on a date for the group. These values were used for normalising counts and building, respectively, index M2, M5 and M3.

2.4 **Species selection and phenology charts for index quality questionnaire: expert analysis**

The sample selected consisted of a set of 32 univoltine, bivoltine and multivoltine butterflies whose phenology and ecology are relatively well known *a priori* (Lafranchis and Geniez, 2000; Lafranchis et al. 2015). Records of all sub-species were included in the species-rank phenology analysis (see appendix 1 for the list of species).

For species whose phenology is likely to vary with latitude or altitude, a biogeographic approach was employed. Data were analysed at four separate biogeographic regions, an aggregated version of the environmental zones defined by Metzger et al. (2005): Atlantic, Continental and Pannonian (ATCONP), Lusitanian (LUS), Alpine and Mediterranean Mountains (ALMM), and Mediterranean (MD). For such species, we considered only biogeographic regions where at least 36 total records were available since 1980. The threshold was defined assuming that 36, an average of three records per month, was the minimum sample size for displaying patterns based on pooled counts. This threshold was also a trade-off between avoiding unreliable, insufficient data and maximising information (i.e. keeping regions where some species are rarer). The indices were computed on pre-cleaned data (see previous section), yielding six different phenological patterns that were submitted in the form of bar plots to four experts. The bar plots displayed frequencies by ten-day periods and, in some cases, by biogeographic regions (Fig. 3 and Appendix 2). For every species, the four experts were asked to rate six charts, one per index, on a scale of 1 to 4 (“very poor”, “poor”, “fair”, “good”), according to their quality in representing known seasonal activity (i.e. flight phenology, in the case of diurnal lepidopterans). In order to avoid conditioning and keep their judgment unbiased, experts were not aware of index design rules. In addition, the disposition of the bar plots was randomized, so that they could not individually identify the indices (Appendix 2). Every index was ranked by counting the number of times it had been rated as "fair" or "good" (i.e. "the phenology chart is representative enough") on the one hand, and as "very poor" (i.e. "the phenology chart is not representative at all") on the other hand. In addition, a Chi-squared (χ^2) test was performed on score frequencies to seek any significant difference between the phenology indices according to expert opinion.

2.5 **Quantitative analysis: comparison with the STERF**

A comparison with an independent sample was performed in order to quantitatively analyse the pertinence of the indices and complete the analysis based on expert opinion. This sample was provided by the French Butterfly Monitoring scheme (STERF), established from 2005 onwards by Vigie-Nature (http://vigienature.mnhn.fr/page/suivi-temporel-des-rhopaloceres-de-france). The STERF provides systematic counts of adult butterflies, which should mirror true species phenology. Each observer performs at least four field visits per year: one visit per month from May to August (other visits during the year are possible). Each survey site is associated with one observer, and is either chosen by the observer or randomly selected. Butterflies are identified and individuals are counted along 5 to 15 transects selected by the observers inside the site, making sure that the habitat within each transect is uniform. Since data collection before May and after August had not been systematic, only STERF records between May and August were kept. The analysis focused on the ATCONP biogeographic region, where phenology is supposed to be uniform and STERF data are more regular and abundant. As in the previous analysis (section 2.4), we fixed a threshold of 36 records in the ATCONP region.

For each species, annual phenologies were estimated using the *rbms* R package and the regional GAM (generalized additive model) method presented in Schmucki et al. (2016). For each species, weekly basis estimate counts were hence obtained for each week of each year. These weekly and yearly count estimates were used to model the average phenology of each species across the period covered by the STERF. Thus, for each species, weekly count *c* recorded years *i* at week *t* were modeled using a GAM with a Negative-binomial distribution and log link function:

$$E[c_{it}] = \mu_{it} = exp[y_i + s(t, f)]$$

where weekly count c_{it} is a function of a year effect *y* and a penalized cubic regression splines smoothing effect over time (week) *t* with *f* degree of freedom. GAMs were computed thanks to the *mgcv* package (Wood 2017) in R 3.5.2. The respective GAM for each species was finally used to predict ten-day period count estimates that could be compared to the phenology indices calculated with INPN data. The resulting flight curve were standardized to one ($\Sigma\mu_t = 1$). This curve could be calculated for 57 species. The others were excluded for lack of data and because the first GAM failed to yield annual phenologies.

STERF count estimates at ten-day periods were compared to the phenology indices in order to verify: 1) whether the patterns were similar, 2) which index produces a pattern that best matches STERF data and can be used for representing species phenology with INPN data.

Since STERF data are already integrated in the INPN database and start in 2005, the indices were calculated excluding all STERF-derived records and other records before 2005.

For each of the 57 species, STERF count estimates at ten-day periods were compared to each of the six phenology indices computed on INPN data, using the Pearson's correlation coefficient (rho). In order to highlight which index was most correlated to STERF pattern, all rho coefficients for the 57 species were analysed together and the indices were ranked by counting the number of times they obtained the maximum and minimum rho coefficient. Table 3 illustrates with fictitious results how the indices were compared between each other according to the number of maximum and minimum coefficient obtained among several species: for example, M4 is the best correlated index to the STERF for species A and B, while it is the least correlated for species C. In this case, M4 obtains twice the maximum rho coefficient and once the minimum rho coefficient (Tab 3). In addition, the difference between rho distributions per index was tested by carrying out a one-way ANOVA (homoscedasticity assumption was met, according to a studentized Breusch-Pagan test: BP = 5.95, df = 5, p-value = 0.31).

Table 3. Fictitious Pearson's correlation coefficients (rho) table for three hypothetical species.

	M1	**M2**	**M3**	**M4**	**M5**	**M6**	**Best correlated index**	**Least correlated index**
species A	0.94	0.8	0.95	0.98	0.92	0.97	M4	M2
species B	0.77	0.5	0.82	0.93	0.75	0.77	M4	M2
species C	0.89	0.85	0.78	0.68	0.82	0.91	M6	M4
mean±SE	0.86±0.05	0.72±0.11	0.85±0.05	0.83±0.09	0.83±0.05	0.88±0.06		
Number of max rho	0	0	0	2	0	1		

Number of min rho	0	2	0	1	0	0		

3 Results

For each of the 32 species, charts were presented to the experts as in Fig. 3a or Fig. 3b, depending on whether a biogeographic effect was expected or not on species phenology. Unlike Fig. 3, the order and the design rules of the indices were concealed from the experts (see an example in Appendix 3). M6 obtained the highest number of positive scores and the smallest number of negative scores: out of the 4 experts assessments, it was rated as “fair” or “good” 62 times and as “very poor” 11 times, followed by M4 with 61 positive scores and 11 negative scores (Fig. 4). M2 obtained most negative responses (Fig. 4), which affected Chi-squared test results (χ^2= 43.73, df = 5, p-value < 0.001). In fact, when the M2-score distribution was excluded from the test, no other significant difference was found among the score distributions of the remaining five indices (χ^2 = 3.1, df = 4, p-value = 0.54).

Fig. 5 shows an example of pairwise comparison between STERF count estimates and the six phenology indices for one of the species, while overall results of this comparison for 57 butterfly species are illustrated in Fig. 6. All indices showed a fair correlation with the STERF, rho coefficient being 0.77 on average (Fig 6a). Index M1 and M5 were the best ranked, with the highest frequencies of maximum rho coefficient and few minimums (Fig 6b). Most of the times M2 resulted as the least correlated to STERF count estimates (Fig. 6b). M2 rho distribution varied also considerably, with values below the first quartile that range from -0.34 and 0.56 (Fig. 6a). Nevertheless, the statistical analysis did not highlight any significant difference (ANOVA F= 2.19, df=5, p-value=0.05).

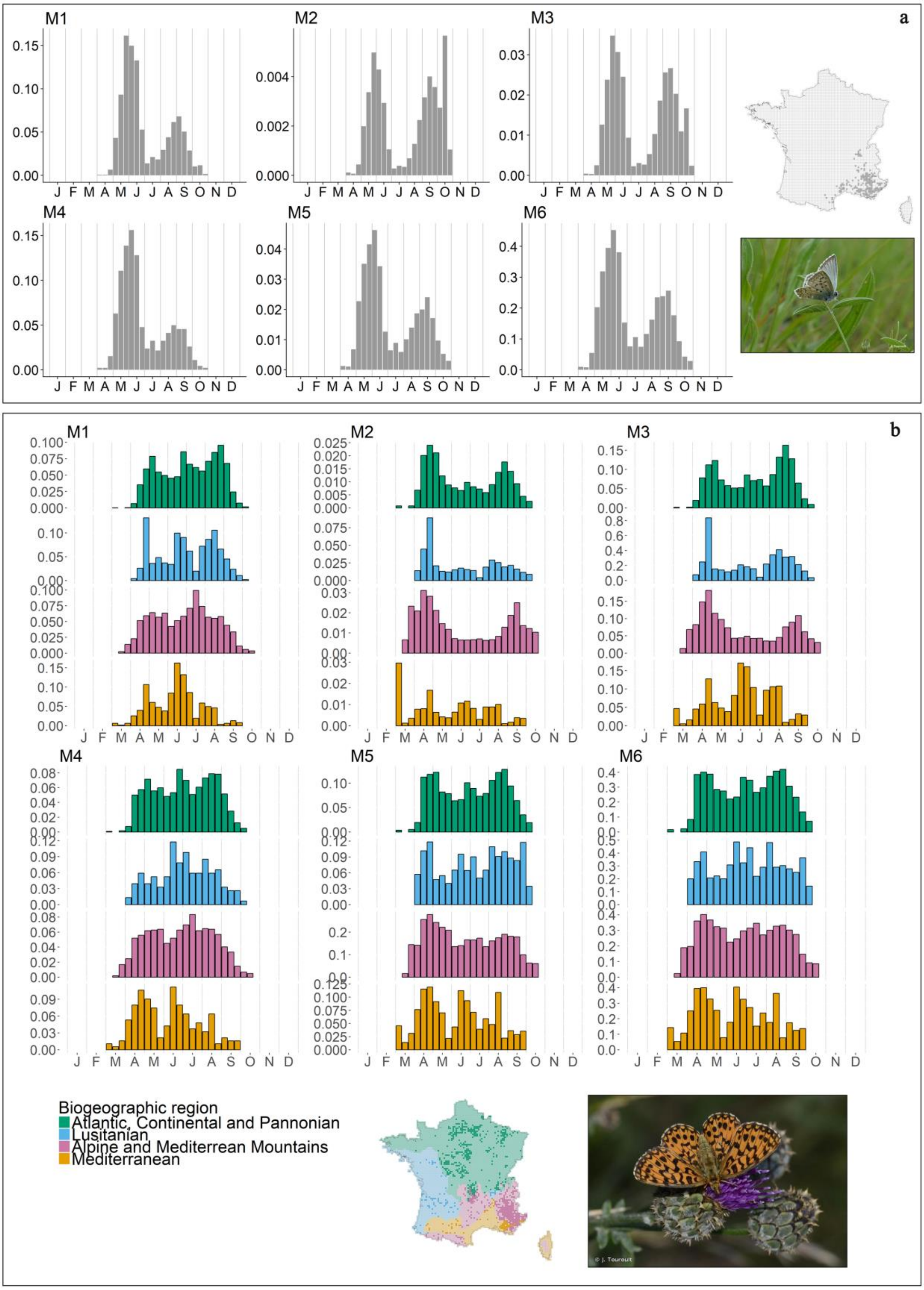


Figure 3. Phenology and distribution maps (based on INPN records after data preparation) of two butterfly species: (a) Provence chalk-hill blue (*Lysandra hispana*) and (b) Violet Fritillary (*Boloria dia*). Time unit is a ten-day period and the x-labels indicate months. Phenology is represented with six indices (see also Tab. 2): proportion of records per period (M1); ratio of records to the group (M2); number of records per field visit (M3); proportion of quadrats per period (M4); ratio of quadrats to the group (M5); ratio of quadrats to the group within the

species known distribution (M6). When a biogeographic effect was expected, such as for (b), phenology and distribution map were illustrated at four biogeographic regions.

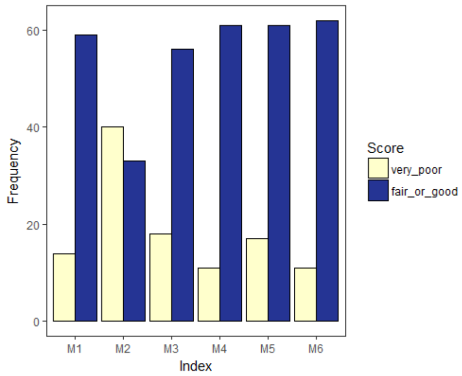


Figure 4 frequency of positive (“fair” or “good”) and negative (“very poor”) scores given by experts to six phenology indices. Index ranking according to the number of “fair or good” scores: M6, M4/M5, M1, M3, M2; index ranking according to the number of “very poor” scores: M2, M3, M5, M1, M4/M6. M2 is the only index that obtained more negative than positive scores. With M2: $\chi 2 = 43.73$, df = 5, p-value < 0.001; excluding M2: $\chi 2 = 3.1$, df = 4, p-value = 0.5.

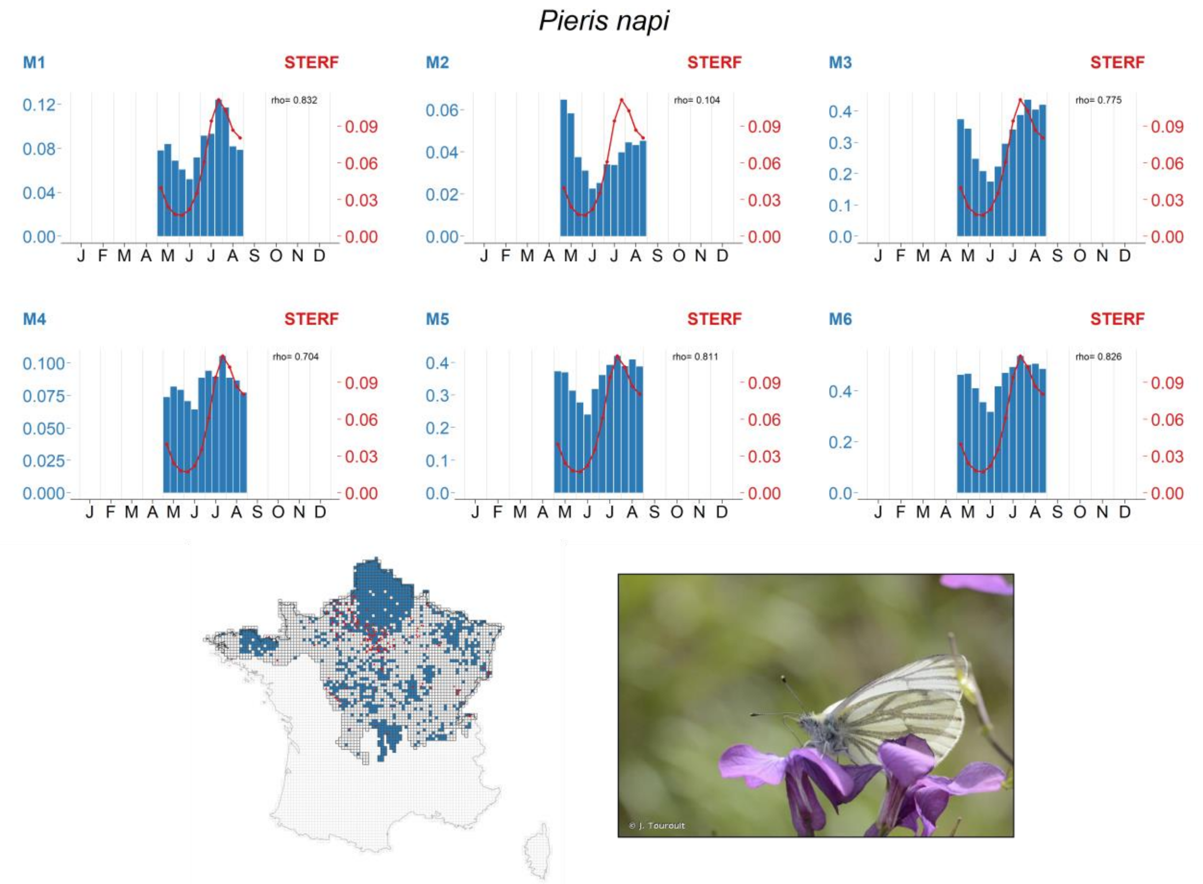


Figure 5. Comparison between the STERF and six phenology indices computed on INPN data by ten-day period for one of the species, the green-veined white (*Pieris napi*). A Pearson’s correlation coefficient (rho) was calculated between every index (blue bar plots) and STERF count estimates from a GAM (red line). Indices and STERF count estimates were calculated with data collected from 2005, and displayed

from May to August in the ATCONP biogeographic region. The map shows the geographic distribution of these data (blue quadrats for the INPN, red points for the STERF).

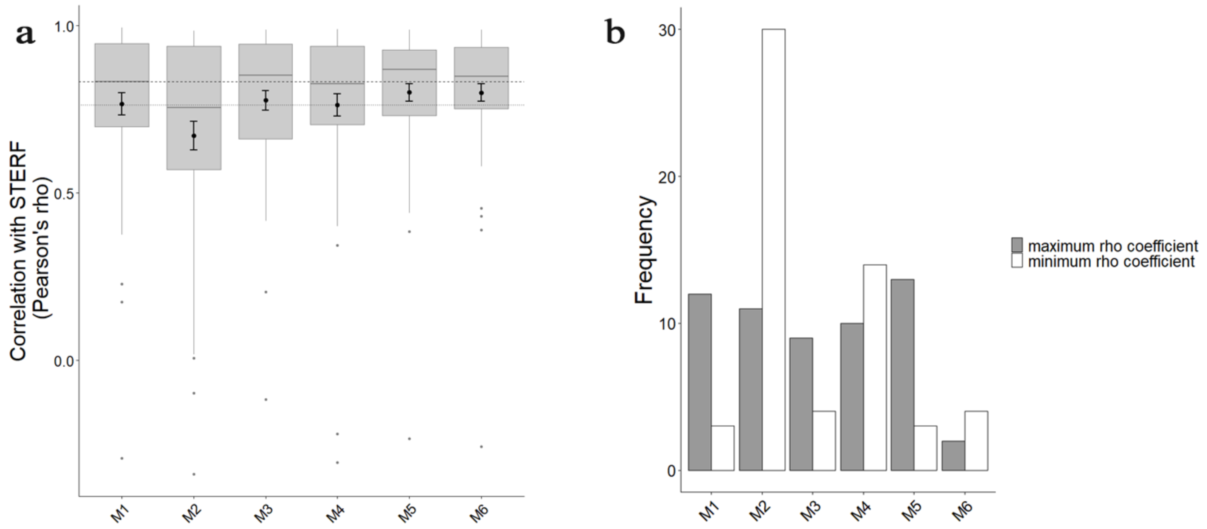


Figure 6 Comparison between the STERF and six phenology indices computed on INPN data by ten-day periods for 57 butterfly species (a) distributions and mean±se of rho coefficients by index. Dashed and dotted lines are respectively the overall median=0.83 and the overall mean=0.77 of all correlations. All phenology indices are equally correlated to STERF count estimates (ANOVA F=2.19, df=5, p-value=0.05); (b) number of times among 57 species that each index occurred to be the best or the least correlated to the STERF.

4 Discussion

4.1 *A priori* properties of the indices

Non-corrected indices are easy to understand but they may be subject to several biases when using data that originate from multiple sources and sampling techniques. Corrected indices should approximate the phenology of activity (e.g. flight period for butterflies) with greater precision than raw frequencies, which are more likely to reflect patterns of recording intensity. No robust but simple metrics based on opportunistic, unstructured data that fulfilled the required conditions was found in recent scientific literature. Dennis et al. (2013) and Schmucki et al. (2016) have produced smooth and readable seasonal patterns with GAMs. With a similar approach, we applied GAMs on butterfly monitoring data for outlining phenology patterns to compare with patterns from INPN data. However, GAMs are not suitable for multi-source opportunistic data, because they do not account for uneven recording intensity (Rothery and Roy 2001). Other approaches, such as correction for sampling effort or imperfect detection in occupancy models, could inspire research for more appropriate measures (references in Table 2). M1 (proportion of records per period; Tab.2) is the most simple and intuitive, but may reflect effort variability over time and over space. M4 (proportion of quadrats per period; Tab.2), which is also easily interpretable, may correct for some errors (duplicates, oversampled quadrats), but still reflects temporal and spatial bias. Conversely, M2 (ratio of records to the group; Tab.2) should correct for uneven recording intensity. However, this index seems to be particularly sensitive to “reporting bias” (van Strien et al. 2013). The number of group records increases sharply from near zero in winter months to near 80000 in summer months (Fig 2a), causing over-correction and making peaks shift from the centre to the edges of the distribution. This is particularly true for certain species, whose adults can emerge outside the habitual flight period, for instance on sunny winter days (such as *Aglais io* and *Aglais urticae*, Lafranchis and Geniez 2000; Lafranchis et al. 2015; phenological patterns of *Aglais io* are shown in Appendix 4). Owing to the unusual event, the species is almost the only one reported at such dates, generating artefact peaks in the phenological pattern (i.e. the ratio “species records/group records” is close to 1). The same may occur to species that are not easy to observe in general during the flight season (i.e. the ratio decreases instead of displaying a peak) and to those that are paradoxically under-reported because considered too common (Dickinson et al. 2010, van Strien et al. 2013). The proxy for observation effort used for building M3 (number of records per field visit; Tab. 2; Fig. 2c) should also correct for uneven recording intensity. This index, as the following ones, is less intuitive and may not be precise, since observer names are not always well standardized in the INPN. Such as M2 and M3, M5 (ratio of quadrats to the group; Tab. 2) is based on a proxy for observation effort, the number of group quadrats per ten-day period (Fig. 2b). As shown in Figure 2b, compared to the other proxy distributions, the number of group quadrats per ten-day period grows and shrinks slower between winter and summer periods and the shape of the distribution is wider. This should limit reporting bias and allow for a better correction of temporal bias, but it does not account for spatial effects (for example, observer activity in coastal and mountain areas is higher during summer than during spring, due to holiday habits of naturalists). Moreover, the proxies for observation effort (Fig 2) include those records or quadrats where the target species was never seen and might be actually absent. Managing data by biogeographic zone may help reducing spatial bias and over-correction. Additionally, quadrats should be restricted to those included in the known distribution of the species. This approach was applied to M6 (ratio of quadrats to the group within the species known distribution; Tab. 2), which was expected to perform better, since it deals with both temporal and spatial bias.

4.2 Which is the highest-ranked index according to the two approaches?

According to expert responses, none of the indices, except for M2 (the worst rated) could illustrate the phenology of 32 butterfly species better than the simplest index, M1 (Fig. 4). In order to refine the results and discriminate among the six methods, a quantitative analysis was carried out by comparing the phenology indices based on INPN opportunistic data to the pattern calculated from a monitoring scheme, the STERF. All indices were well correlated to the STERF in this analysis (Fig. 6). Surprisingly, in the two tests, the indices that correct for observation effort (M2, M3, M5 and M6) are not statistically better than the non-corrected ones (M1 and M4). However, even if no statistical difference was highlighted, the distribution of rho coefficients for M2 showed large variability (Fig. 6a). This result suggests that this index may perform well for some species, but it is not suitable for a large panel of species with different phenologies. In substance, the expert approach and the comparison with the STERF both lead to reject M2 as a suitable phenology index.

Other studies have compared opportunistic data with monitoring data (Dennis et al. 2017a; van Strien et al. 2013). It is possible that our comparisons lack power. The STERF provides systematic counts of adult butterflies, which should provide phenological patterns that are relatively close to reality. Nevertheless, STERF protocol and measures themselves may not be free from bias, notably because counts are recorded on a monthly basis. However, given the two approaches adopted, the number of species that have been taken into account, as well as the amount of data and the variability of expert opinions, it is not unfounded to believe that, there are no major differences between the tested indices. A possible explanation lies in the large amount of data and sources available for butterflies, which may help attenuate the bias and make raw frequencies of opportunistic observations converge towards overall realistic phenological patterns. Our method may not be suitable for other lepidopterans or other clades for which knowledge and data are much scarcer than for the butterflies analysed here. If data are opportunistic and multi-source, we suggest selecting species with many records, covered by several data sources. We fixed a threshold of 36 total records in the studied area and during the entire period of study (in our case, all data from France or from a biogeographic region, recorded since 1980), and 5 data sources, knowing that for most of the species total records and sources were more numerous (Appendix 1). Ideally, the minimum number of records and sources should be calculated and standardised. Further work is required to re-define these thresholds on a statistical basis and give more recommendations about the use of our indices for other groups than butterflies.

4.3 Selecting a rigorous but simple index for large opportunistic data and a general audience: which compromise for butterflies?

Even though they are likely to reflect bias, in our case raw indices (M1 or M4) were not less convincing than corrected ones for butterflies. Following the results and considerations discussed above, and against expectations, it could finally be reasonable to consider the use of raw frequencies. Besides, the choice should head to the most parsimonious method. The best option would probably be to display phenology through raw indices (after data preparation and weighting adjustment), alongside a visual representation of recording intensity (Fig.2b or 2c). Spatial bias may be attenuated by splitting data by biogeographic region. In this way, both experts and general public should be able to understand the graphics, while keeping a critical eye on them.

Some authors have appealed to occupancy models for studying phenology and population trends with opportunistic data (see, for example, Kéry et al. 2010, van Strien et al. 2010, Strebel et al. 2014). In fact, by estimating detectability, occupancy models help to correct for observation effort, detection and reporting bias (Kéry et al. 2010; Van Strien et al. 2013). Although we drew inspiration from occupancy models for bias correction approaches, the aim of our study was not to assess occupancy of butterfly species. Nonetheless, we do not exclude the possibility of using a modelling approach in the future, should sufficient data be available, in order to get more unbiased estimates of phenology.

4.4 Other recommendations facing database limitations

Dates in the INPN are compulsory, so when the day of observation is not known a default date (such as January 1st or December 31st) is assigned. We discarded outlier data (see section 2.3) in order to prevent the appearance of artefacts, although some small winter peaks persisted. In fact, as described in section, 4.1, some butterfly species may be recorded during winter days, even if the actual period of activity is later in the year. Unfortunately, we could not discriminate observations by stage of development or behaviour (activity versus hibernation), since this information is currently lacking in most INPN data sources. However, records of active adult butterflies in the INPN usually far exceed observations of young stage or wintering individuals, so the latter are unlikely to affect the displaying of major activity peaks. There are, however, some exceptions, such as those described in section 4.1 or the case of *Thecla betulae* (Linnaeus, 1758). This species is hardly detectable during its flight period (Lafranchis and Geniez 2000; Lafranchis et al. 2015). Conversely, the eggs are easy to recognize and they are regularly used for detecting the geographic presence of the butterfly. These occurrences are often reported as all others in datasets when achieving the INPN. This caused the appearance of improbable peaks in winter (Fig. 7). Correcting for observation effort did not provide any added value, it rather accentuated the aberrations. This underlines the need to support data producers towards a better standardisation of information.

Several studies have documented the relationship between climate and phenological shifts (Roy and Sparks 2000, Walther et al. 2002, Parmesan and Yohe 2003, Stefanescu et al. 2003, Menzel et al. 2006, Parmesan 2007, Altermatt 2009, Prodon et al. 2017, Bell et al. 2019). The impact of climate change on phenology may have major consequences on ecological systems and their conservation (Schwartz 2013), hence the importance of long-term collection of observational and monitoring data. Phenological shifts of the order of some days may have occurred in the last decades (Roy and Sparks 2000). Admittedly, we could not point out phenological shifts due to coarse temporal resolution. A finer temporal resolution would entail the loss of useful data. Furthermore, we chose to collate all years' data in order to compensate for uneven recording intensity and species detectability across the years. This also precluded the study of phenological shifts. However, we believe that our methods are adequate to provide a simple measure for a general audience that bears overall phenological patterns for many species. Nevertheless, we will consider investigating with more accuracy the phenology of those butterflies for which precise dates of observation are abundant and consistent through the years by adapting methods that use statistical models, which account for phenological changes over time, observation effort, detection and reporting bias, such as those proposed by Dennis et al. (2017b) and Strebel et al. (2014).

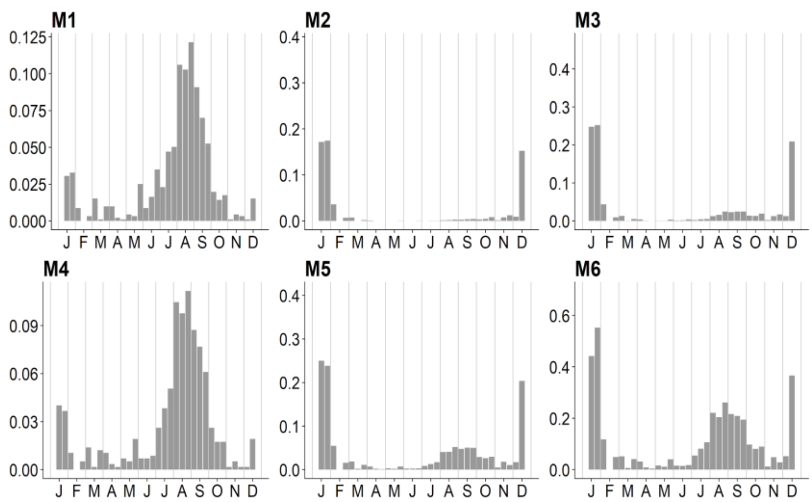


Figure 7. Phenology of the Brown Hairstreak (*Thecla betulae*). Egg and imago sightings are confused in the INPN, causing artificial peaks in winter, when the adult is not flying but eggs are easily detectable.

5 Conclusion

Multi-source national databases such as the INPN may contain bias or redundant information but they compile large volumes of data, centralize and spread knowledge on biodiversity distribution and activity. The study showed, against all odds, that raw frequencies can perform as well as corrected measures, probably due to the characteristics and the large amount of butterfly multi-source data. This is so far the first attempt at correcting large amounts of opportunistic records in order to illustrate species seasonality in France through a simple phenology index.

For groups with large amount of data and replicated visits, such as butterflies, non-corrected multi-source records (i.e. that combine standardised and non-standardised observations) probably provide sufficient information about overall phenological patterns. The next question is whether raw frequencies would fit as well on species with less available data, or on those that are aseasonal (e.g. occur constantly around the year or variations are irregular and not season-related, see Wolda, 1998). Further work will help investigate whether the properties of non-corrected and corrected indices are affected by the amount of data, as well as the variability of data sources and the type of phenology. Further analysis may also include the use of statistical models for estimating phenological shifts in connection with climate change.

Acknowledgments

We thank the respondents of the index quality questionnaire: Frédéric Archaux, Olivier Delzons, Eric Drouet and Antoine Leveque. We also thank all the respondents of a previous version of the questionnaire, which included more species and groups that are not treated in this manuscript. Thanks to Marie-Lyne Einspenner, Chiara Elettra Ferrario and Paula R. M. B. for constructive help on previous drafts of the manuscript.

References

Altermatt, F. 2009. Climatic warming increases voltinism in European butterflies and moths. Proceedings of the Royal Society B: Biological Sciences 277:1281–1287. doi: 10.1098/rspb.2009.1910

Archaux, F., P. Chatard, F. Faucheux, and A. Lévêque. 2015. Papillons du Loiret: atlas des rhopalocères et zygènes du Loiret (2000-2013). So.MOS & Alexanor.

Archaux, F., G. Jacquemin, R. Leconte, G. C. Luquet, F. Mora, F. Noël, H. Pinston, J.-C. Robert, and A. Ruffoni. 2011. Synthèse des observations récentes et anciennes de Libelloides coccajus (Denis & Schiffermüller) et L. longicornis (Linné) dans la moitié nord de la France (Neuroptera, Ascalaphidae). Bulletin de la Société entomologique de France 116:365–388.

August, T. O. M., M. Harvey, P. Lightfoot, D. Kilbey, T. Papadopoulos, and P. Jepson. 2015. Emerging technologies for biological recording: 731–749.

Bell, J.R., M.S. Botham, P.A. Henrys, D.I. Leech, J.W. Pearce-Higgins, C.R. Shortall, T.M. Brereton, J. Pickup, and S.J. Thackeray, 2019. Spatial and habitat variation in aphid, butterfly, moth and bird phenologies over the last half century. Global change biology, 25(6), pp.1982-1994.

Bertone, M., J. Green, S. Washburn, M. Poore, C. Soprenson, and D. W. Watson. 2005. Seasonal Activity and Species Composition of dung beetles inhabiting cattle pastures in North Carolina. Entomological Society of America. 98:309–321.

Dennis, E. B., S. N. Freeman, T. Brereton, and D. B. Roy. 2013. Indexing butterfly abundance whilst accounting for missing counts and variability in seasonal pattern. Methods in Ecology and Evolution 4:637–645. doi: 10.1111/2041-210X.12053

Dennis, E. B., B. J. T. Morgan, T. M. Brereton, D. B. Roy, and R. Fox. 2017a. Using citizen science butterfly counts to predict species population trends. Conservation Biology 31:1350–1361.

Dennis, E.B., Morgan, B.J., Freeman, S.N., Ridout, M.S., Brereton, T.M., Fox, R., Powney, G.D. and Roy, D.B., 2017b. Efficient occupancy model-fitting for extensive citizen-science data. PloS one, 12(3), p.e0174433

Dickinson, J. L., J. Shirk, D. Bonter, R. Bonney, R. L. Crain, J. Martin, T. Phillips, and K. Purcell. 2012. The current state of citizen science as a tool for ecological research and public engagement. Frontiers in Ecology and the Environment 10:291–297. doi: 10.1890/110236

Dickinson, J. L., B. Zuckerberg, and D. N. Bonter. 2010. Citizen Science as an Ecological Research Tool: Challenges and Benefits. Annual Review of Ecology, Evolution, and Systematics 41:149–172. doi: 10.1146/annurev-ecolsys-102209-144636

Gargominy O., S. Tercerie, C. Régnier, T. Ramage, P. Dupont, P. Daszkiewicz, L. Poncet. 2017 TAXREF v11.0, référentiel taxonomique pour la France : méthodologie, mise en œuvre et diffusion. Muséum National d'Histoire Naturelle, Paris. Rapport Patrinat 2017-116.

Giraud, C., C. Calenge, C. Coron, and R. Julliard. 2016. Capitalizing on opportunistic data for monitoring relative abundances of species. Biometrics 72:649–658. doi:10.1111/biom.12431

Hochachka, W. M., D. Fink, R. A. Hutchinson, D. Sheldon, W. K. Wong, and S. Kelling. 2012. Data-intensive science applied to broad-scale citizen science. Trends in Ecology and Evolution 27:130–137.

Isaac, N. J. B., and M. J. O. Pocock. 2015. Bias and information in biological records. Biological Journal of the Linnean Society 115:522–531. doi: 10.1111/bij.12532

Isaac, N. J. B., A. J. van Strien, T. A. August, M. P. de Zeeuw, and D. B. Roy. 2014. Statistics for citizen science: Extracting signals of change from noisy ecological data. Methods in Ecology and Evolution. doi: 10.1111/2041-210X.12254

Kéry, M., J. A. Royle, H. Schmid, M. Schaub, B. Volet, G. Häfliger, and N. Zbinden. 2010. Site-Occupancy Distribution Modeling to Correct Population-Trend Estimates Derived from Opportunistic Observations. Conservation Biology 24:1388–1397. doi: 10.1111/j.1523-1739.2010.01479.x

Lobo, J. M. 2008. Database records as a surrogate for sampling effort provide higher species richness estimations. Biodiversity and Conservation 17:873–881. doi: 10.1007/s10531-008-9333-4

Lumaret J.P. 1990 Atlas des Coléoptères Scarabéides Laparosticti de France. Muséum National d'Histoire Naturelle. Secrétariat de la Faune et la Flore Paris Paris.

Maes, D., W. Vanreusel, I. Jacobs, K. Berwaerts, and H. Van Dyck. 2012. Applying IUCN Red List criteria at a small regional level: A test case with butterflies in Flanders (north Belgium). Biological Conservation 145:258–266. doi: 10.1016/j.biocon.2011.11.021

Manil, L., A. Lerch, B. Fontaine, and R. Julliard. 2015. Suivi Temporel des Rhopalocères de France (STERF) Bilan 2005-2014. Gestion 2010:1–82.

Menzel, A., T. H. Sparks, N. Estrella, E. Koch, A. Aaasa, R. Ahas, K. Alm-Kübler, P. Bissolli, O. Braslavská, A. Briede, F. M. Chmielewski, Z. Crepinsek, Y. Curnel, Å. Dahl, C. Defila, A. Donnelly, Y. Filella, K. Jatczak, F. Måge, A. Mestre, Ø. Nordli, J. Peñuelas, P. Pirinen, V. Remišová, H. Scheifinger, M. Striz, A. Susnik, A. J. H. Van Vliet, F. E. Wielgolaski, S. Zach, and A. Zust. 2006. European phenological response to climate change matches the warming pattern. Global Change Biology 12:1969–1976. doi: 10.1111/j.1365-2486.2006.01193.x

Metzger, M. J., R. G. H. Bunce, R. H. G. Jongman, C. A. Mücher, and J. W. Watkins. 2005. A climatic stratification of the environment of Europe. Global Ecology and Biogeography 14:549–563. doi: 10.1111/j.1466-822X.2005.00190.x

Parmesan, C. 2007. Influences of species, latitudes and methodologies on estimates of phenological response to global warming. Global Change Biology 13:1860–1872. doi: 10.1111/j.1365-2486.2007.01404.x

Parmesan, C., and G. Yohe. 2003. A globally coherent fingerprint of climate change impacts across natural systems. Nature 421:37–42. doi: 10.1038/nature01286

Phillips, S. J., M. Dudik, J. Elith, C. H. Graham, A. Lehmann, J. Leathwick, and S. Ferrier. 2009. Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. ECOLOGICAL APPLICATIONS 19:181–197. doi: 10.1890/07-2153.1

Pocock, M. J. O., H. E. Roy, C. D. Preston, and D. B. Roy. 2015. The Biological Records Centre: A pioneer of citizen science. Biological Journal of the Linnean Society 115:475–493.

Ponder, W. F., G. A. Carter, P. Flemons, and R. R. Chapman. 2001. Evaluation of museum collection data for use in biodiversity assessment. Conservation Biology 15:648–657. doi: 10.1046/j.1523-1739.2001.015003648.x

Powney, G. D., and N. J. B. Isaac. 2015. Beyond maps: A review of the applications of biological records. Biological Journal of the Linnean Society 115:532–542. doi: 10.1111/bij.12517

Pozo, C., A. Luis-Martínez, J. Llorente-Bousquets, N. Salas-Suárez, A. Maya-Martínez, I. Vargas-Fernández, and A. D. Warren. 2008. Seasonality and Phenology of the Butterflies (Lepidoptera: Papilionoidea and Hesperioidea) of Mexico’s Calakmul Region. Florida Entomologist 91:407–422. doi: 10.1653/0015-4040(2008)91[407:SAPOTB]2.0.CO;2

Prodon, R., P. Geniez, M. Cheylan, F. Devers, I. Chuine, and A. Besnard. 2017. A reversal of the shift towards earlier spring phenology in several Mediterranean reptiles and amphibians during the 1998–2013 warming slowdown. Global Change Biology 23:5481–5491. doi: 10.1111/gcb.13812

R Core Developement Team. 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing Vienna, Austria.

Robertson, M. P., G. S. Cumming, and B. F. N. Erasmus. 2010. Getting the most out of atlas data. Diversity and Distributions 16:363–375. doi: 10.1111/j.1472-4642.2010.00639.x

Rothery, P. and D. B. Roy. 2001. Application of generalized additive models to butterfly transect count data. Journal of Applied Statistics, 28:897-909.

Roy, D. B., and T. H. Sparks. 2000. Phenology of British butterflies and climate change. Global Change Biology 6:407–416. doi: 10.1046/j.1365-2486.2000.00322.x

Ruete, A. 2015. Displaying bias in sampling effort of data accessed from biodiversity databases using ignorance maps. Biodiversity Data Journal 3:e5361. doi: 10.3897/BDJ.3.e5361

Schmucki, R., G. Pe’er, D. B. Roy, C. Stefanescu, C. A. M. Van Swaay, T. H. Oliver, M. Kuussaari, A. J. Van Strien, L. Ries, J. Settele, M. Musche, J. Carnicer, O. Schweiger, T. M. Brereton, A. Harpke, J. Heliölä, E. Kühn, and R. Julliard. 2016. A regionally informed abundance index for supporting integrative analyses across butterfly monitoring schemes. Journal of Applied Ecology 53:501–510. doi: 10.1111/1365-2664.12561

Schwartz, M. D. 2013. Phenology: An integrative environmental science. Phenology: An Integrative Environmental Science:53–65. doi: 10.1007/978-94-007-6925-0

Stefanescu, C., J. Peñuelas, and I. Filella. 2003. Effects of climatic change on the phenology of butterflies in the northwest Mediterranean Basin. Global Change Biology 9:1494–1506. doi: 10.1046/j.1365-2486.2003.00682.x

Strebel, N., M. Kéry, M. Schaub, and H. Schmid. 2014. Studying phenology by flexible modelling of seasonal detectability peaks. Methods in Ecology and Evolution 5:483–490. doi: 10.1111/2041-210X.12175

van Strien, A. J., C. A. M. van Swaay, and T. Termaat. 2013. Opportunistic citizen science data of animal species produce reliable estimates of distribution trends if analysed with occupancy models. Journal of Applied Ecology 50:1450–1458. doi: 10.1111/1365-2664.12158

van Strien, A. J., T. Termaat, D. Groenendijk, V. Mensing, and M. Kéry. 2010. Site-occupancy models may offer new opportunities for dragonfly monitoring based on daily species lists. Basic and Applied Ecology 11:495–503. doi: 10.1016/j.baae.2010.05.003

van Swaay, C. A. M., C. L. Plate, and A. J. van Strien. 2002. Monitoring butterflies in the Netherlands: how to get unbiased indices. Proceedings of the Section Experimental and Applied Entomology of the Netherlands Entomological Society 13:21–27.

van Swaay, C. A. M. 1990. An assessment of the changes in butterfly abundance in The Netherlands during the 20th Century. Biological Conservation 52:287–302.

Touroult, J., L. Poncet, P. Keith, V. Boullet, G. Arnal, H. Brustel and J. P. Siblet. 2015. Inventaires et atlas nationaux de distribution: pour une approche plus itérative et un rééquilibrage taxinomique. Revue d'écologie.

Tulloch, A. I. T., H. P. Possingham, L. N. Joseph, J. Szabo, and T. G. Martin. 2013. Realising the full potential of citizen science monitoring programs. Biological Conservation 165:128–138. doi: 10.1016/j.biocon.2013.05.025

Turin, H., and P. J. Den Boer. 1988. Changes in the distribution of carabid beetles in The Netherlands since 1880, II. Isolation of habitats and long-term time trends in the occurrence of carabid species with different powers of dispersal (Coleoptera, Carabidae). Biological Conservation 44:179–200. doi: 10.1016/0006-3207(88)90101-2

Walther, G.-R., E. Post, P. Convey, A. Menzel, C. Parmesan, T. J. C. Beebee, J.-M. Fromentin, O. Hoegh-Guldberg, and F. Bairlein. 2002. Ecological responses to recent climate change. Nature 416:389–395. doi: 10.1038/416389a

Wolda, H. 1988. Insect Seasonality: Why. Annual Review of Ecology and Systematics 19:1–18. doi: 10.1146/annurev.ecolsys.19.1.1

Wood, S.N., 2017. Generalized additive models: an introduction with R. Chapman and Hall/CRC

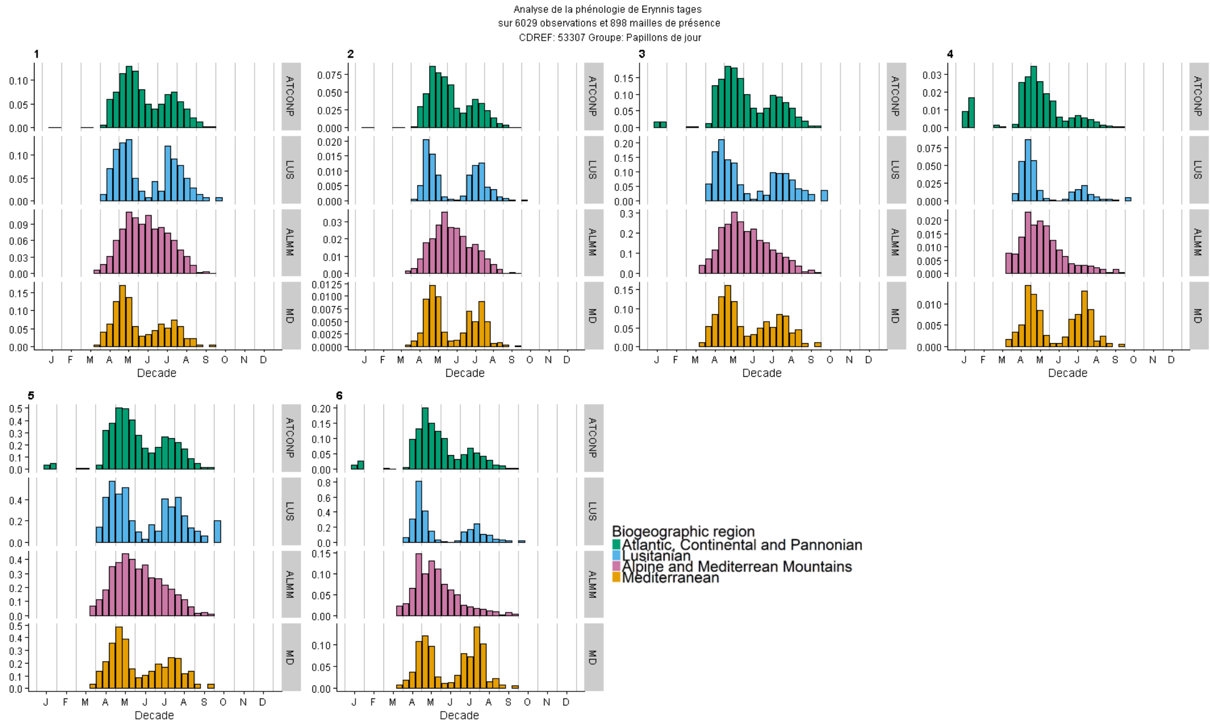
Appendix

Appendix 1. Selected butterflies, analysis approach (expert analysis, comparison with the STERF) and numbers after data preparation. No-STERF and STERF records include only data from 2005, between Mai and August.

Family	Scientific name	Expert analysis			Comparison with the STERF		Number of sources in the INPN
		Number of records	Number of quadrats	Probable biogeographic effect	Number of no-STERF records	Number of STERF records	
Hesperiidae	*Carcharodus alceae* (Esper, 1780)	2865	673	N	737	135	80
	Erynnis tages (Linnaeus, 1758)	6023	898	Y	2662	423	97
	Ochlodes sylvanus (Esper, 1777)	-	-	-	5664	1132	125
	Pyrgus malvae (Linnaeus, 1758)	-	-	-	915	154	55
	Thymelicus lineola (Ochsenheimer, 1808)	-	-	-	2476	271	93
	Thymelicus sylvestris (Poda, 1761)	-	-	-	1628	431	103
Lycaenidae	*Aricia agestis* (Denis & Schiffermüller, 1775)	9188	1321	N	3324	890	111
	Callophrys rubi (Linnaeus, 1758)	-	-	-	1472	167	104
	Celastrina argiolus (Linnaeus, 1758)	-	-	-	3353	751	118
	Cupido alcetas (Hoffmannsegg, 1804)	1216	301	N	-	-	55
	Cupido argiades (Pallas, 1771)	3092	443	Y	1287	159	60
	Cupido minimus (Fuessly, 1775)	3521	560	Y	1225	245	76
	Cyaniris semiargus (Rottemburg, 1775)	3538	681	Y	1351	104	89
	Lycaena dispar (Haworth, 1802)	2422	352	Y	-	-	50
	Lycaena phlaeas (Linnaeus, 1760)	-	-	-	2229	419	109
	Lycaena tityrus (Poda, 1761)	3169	767	Y	1383	265	92
	Lysandra bellargus (Rottemburg, 1775)	10028	856	Y	3644	885	108
	Lysandra coridon (Poda, 1761)	-	-	-	3093	835	94
	Lysandra hispana (Herrich-Schäffer, 1852)	1364	192	N	-	-	31
	Plebejus argus (Linnaeus, 1758)	2785	435	Y	-	-	67
	Plebejus argyrognomon (Bergsträsser, 1779)	1167	225	N	616	204	46
	Polyommatus icarus (Rottemburg, 1775)	27821	1991	Y	10738	2440	137
	Satyrium ilicis (Esper, 1779)	-	-	-	188	116	69
	Thecla betulae (Linnaeus, 1758)	913	325	N	-	-	54
Nymphalidae	*Aglais io* (Linnaeus, 1758)	17761	1716	Y	9001	1828	128
	Aglais urticae (Linnaeus, 1758)	16901	1423	Y	9405	1399	120
	Aphantopus hyperantus (Linnaeus, 1758)	-	-	-	7500	1161	103
	Araschnia levana (Linnaeus, 1758)	6291	965	N	4474	549	77
	Argynnis paphia (Linnaeus, 1758)	-	-	-	4526	965	121

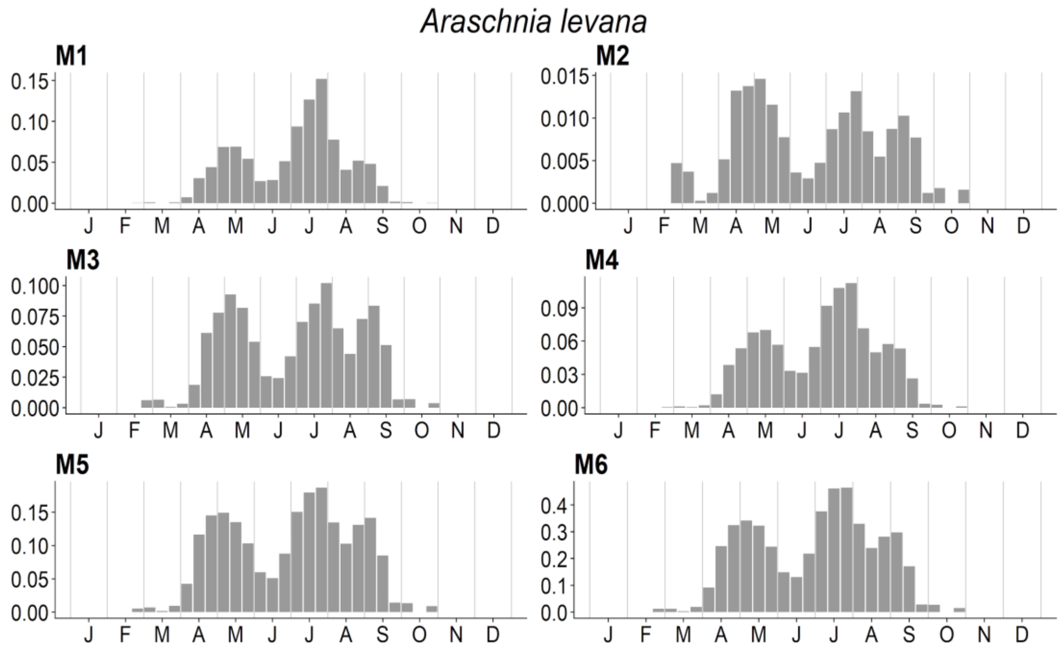
	Boloria dia (Linnaeus, 1767)	7909	811	Y	2831	567	96
	Brintesia circe (Fabricius, 1775)	-	-	-	1151	92	97
	Coenonympha arcania (Linnaeus, 1760)	-	-	-	4132	895	104
	Coenonympha pamphilus (Linnaeus, 1758)	32819	2171	N	12812	4239	142
	Fabriciana adippe (Denis & Schiffermüller, 1775)	-	-	-	392	116	74
	Issoria lathonia (Linnaeus, 1758)	4479	920	Y	1148	254	103
	Lasiommata maera (Linnaeus, 1758)	2805	464	Y	431	93	83
	Lasiommata megera (Linnaeus, 1767)	-	-	-	2976	916	130
	Limenitis camilla (Linnaeus, 1764)	-	-	-	3054	589	86
	Limenitis reducta Staudinger, 1901	3331	619	Y	-	-	81
	Maniola jurtina (Linnaeus, 1758)	-	-	-	20661	6992	143
	Melanargia galathea (Linnaeus, 1758)	-	-	-	11783	2633	136
	Melitaea athalia (Rottemburg, 1775)	-	-	-	1150	214	87
	Melitaea cinxia (Linnaeus, 1758)	-	-	-	999	272	92
	Melitaea parthenoides Keferstein, 1851	1419	418	N	190	42	70
	Melitaea phoebe (Denis & Schiffermüller, 1775)	3510	663	N	520	51	82
	Minois dryas (Scopoli, 1763)	-	-	-	1125	56	54
	Nymphalis antiopa (Linnaeus, 1758)	1213	301	Y	-	-	55
	Pararge aegeria (Linnaeus, 1758)	-	-	-	10791	3134	141
	Polygonia c-album (Linnaeus, 1758)	11012	1558	Y	5953	954	134
	Pyronia tithonus (Linnaeus, 1771)	-	-	-	11071	3774	113
	Speyeria aglaja (Linnaeus, 1758)	-	-	-	781	164	83
	Vanessa atalanta (Linnaeus, 1758)	-	-	-	12050	1905	155
	Vanessa cardui (Linnaeus, 1758)	-	-	-	7430	1819	137
Papilionidae	*Iphiclides podalirius* (Linnaeus, 1758)	7339	988	Y	1686	294	128
	Papilio machaon Linnaeus, 1758	8949	1488	Y	4002	207	135
Pieridae	*Anthocharis cardamines* (Linnaeus, 1758)	-	-	-	4280	825	136
	Aporia crataegi (Linnaeus, 1758)	-	-	-	3587	317	97
	Colias alfacariensis Ribbe, 1905	-	-	-	1292	836	84
	Colias crocea (Geoffroy in Fourcroy, 1785)	-	-	-	3691	1020	132
	Gonepteryx rhamni (Linnaeus, 1758)	19303	1838	Y	10062	1586	146
	Leptidea duponcheli (Staudinger, 1871)	540	97	N	-	-	15
	Leptidea sinapis (Linnaeus, 1758)	-	-	-	3068	972	116
	Pieris brassicae (Linnaeus, 1758)	-	-	-	6186	2522	139
	Pieris napi (Linnaeus, 1758)	-	-	-	10227	2860	136
	Pieris rapae (Linnaeus, 1758)	-	-	-	10105	6088	144
Riodinidae	*Hamearis lucina* (Linnaeus, 1758)	1603	315	Y	-	-	72
	All butterflies and diurnal moths	772307	3315	-	254508	63226	234

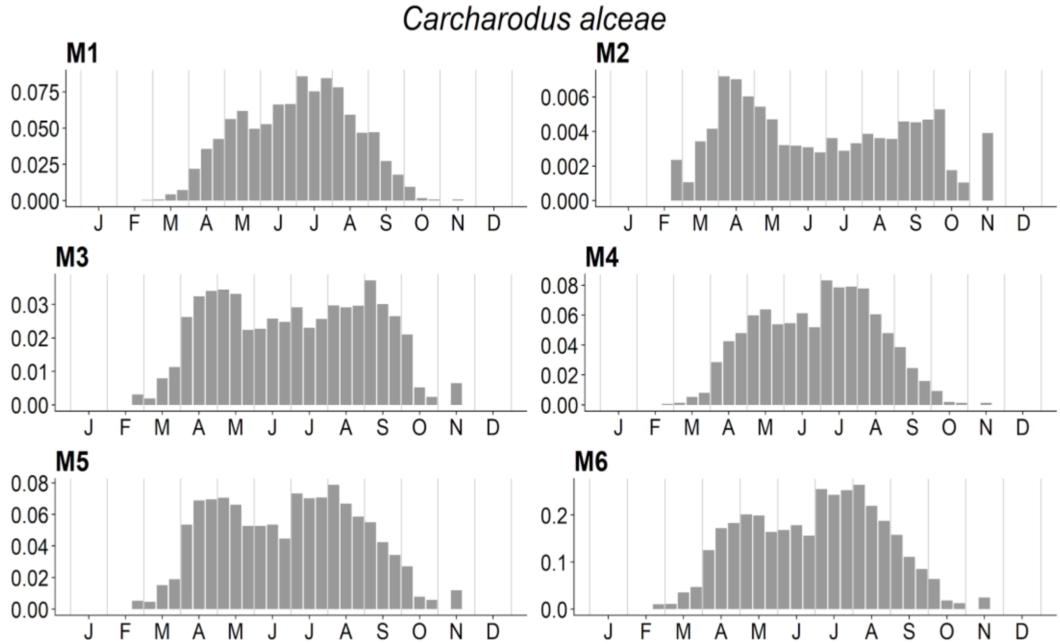
Appendix 2. Example of charts presented to the experts for rating the quality of the indices. The order of bar plots was randomized and the rules behind index building were kept from the experts so that they could not prejudge the index in itself or identify the same index among several species.



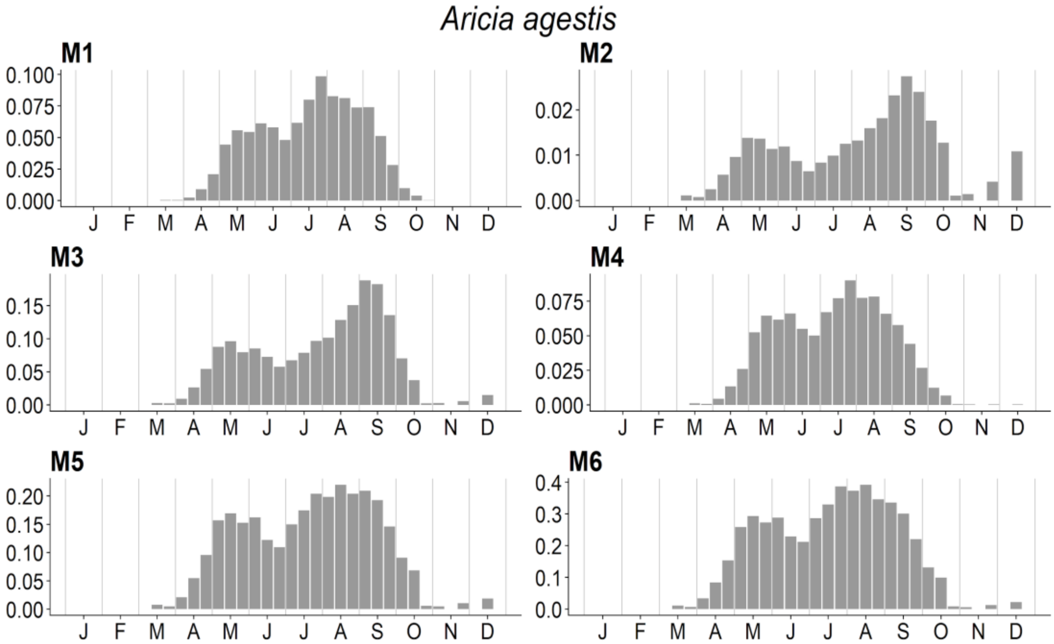
Appendix 3. Phenology of 11 butterfly species (2 Hesperiidae, 5 Lycaenidae, 3 Nymphalidae, 1 Pieridae) for which biogeographic effect was not expected. Phenology is represented with six indices (see also Tab. 2): proportion of records per period (M1); ratio of records to the group (M2); number of records per field visit (M3);proportion of quadrats per period (M4); ratio of quadrats to the group (M5); ratio of quadrats to the group within the species known distribution (M6).

Hesperiidae:

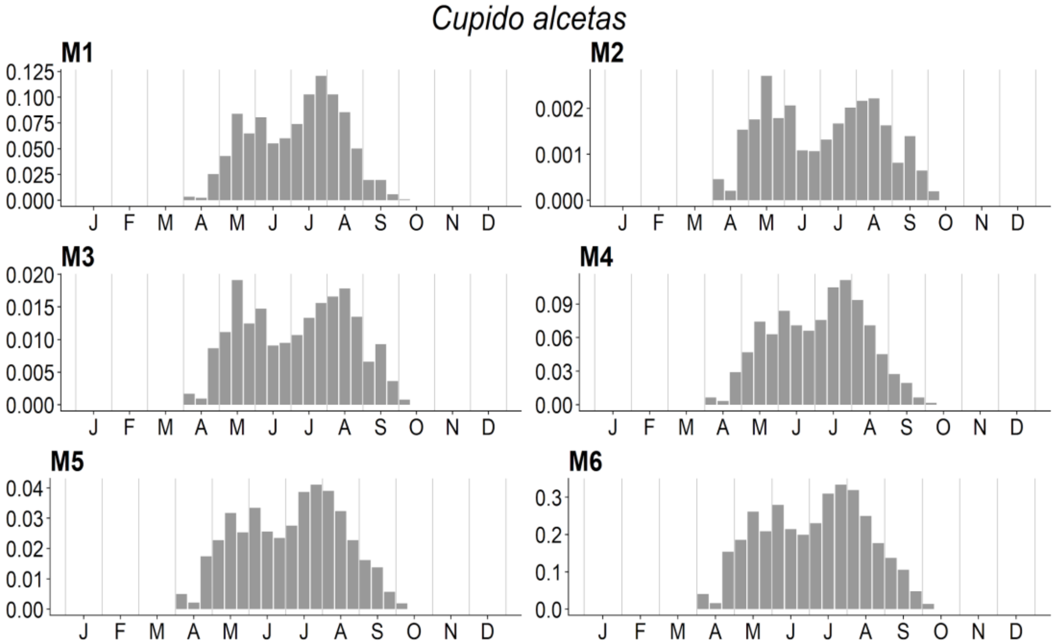




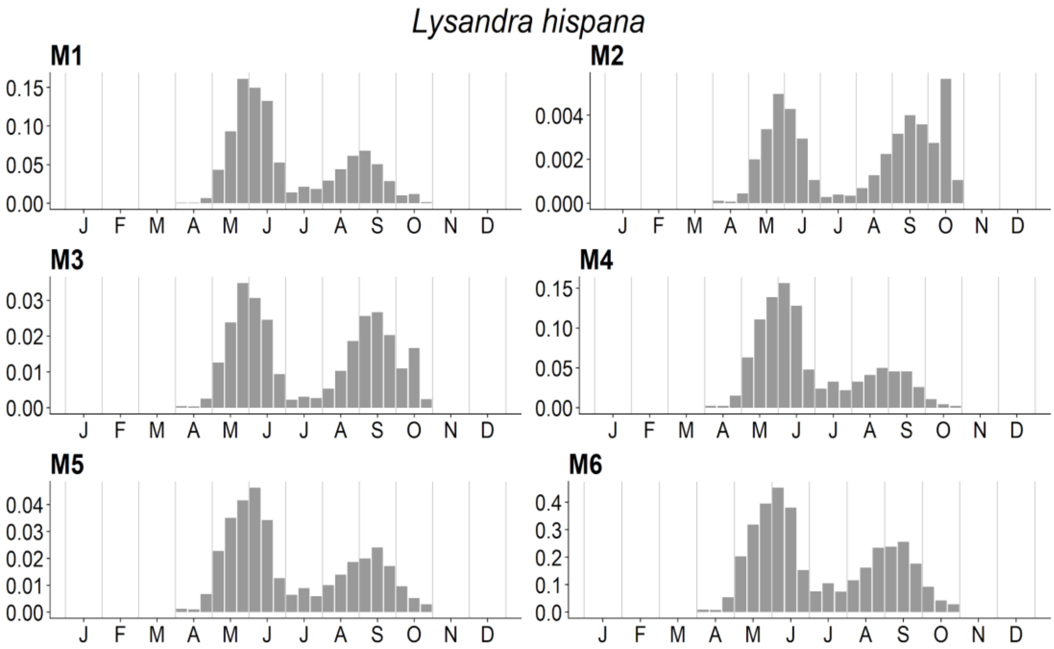
Lycaenidae:

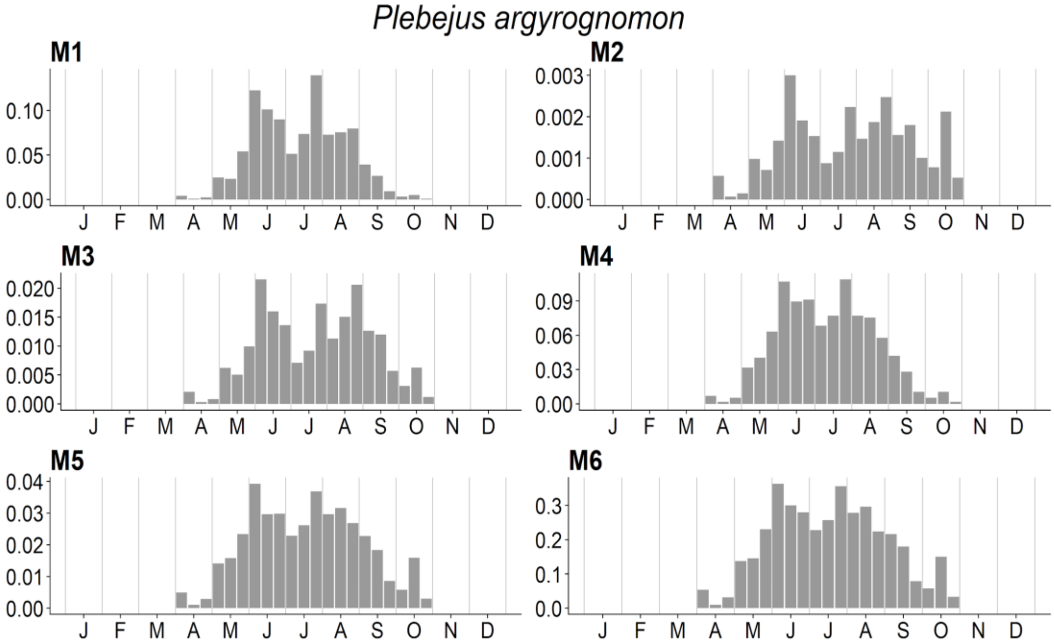


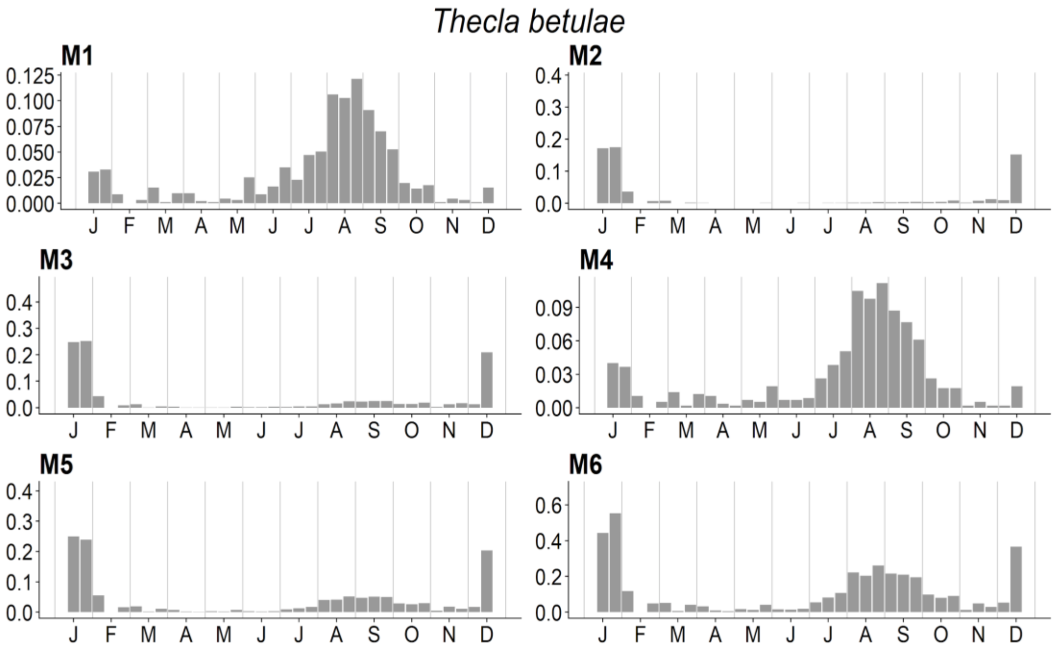
Aricia agestis
M1
M2
M3
M4
M5
M6
J F M A M J J A S O N D



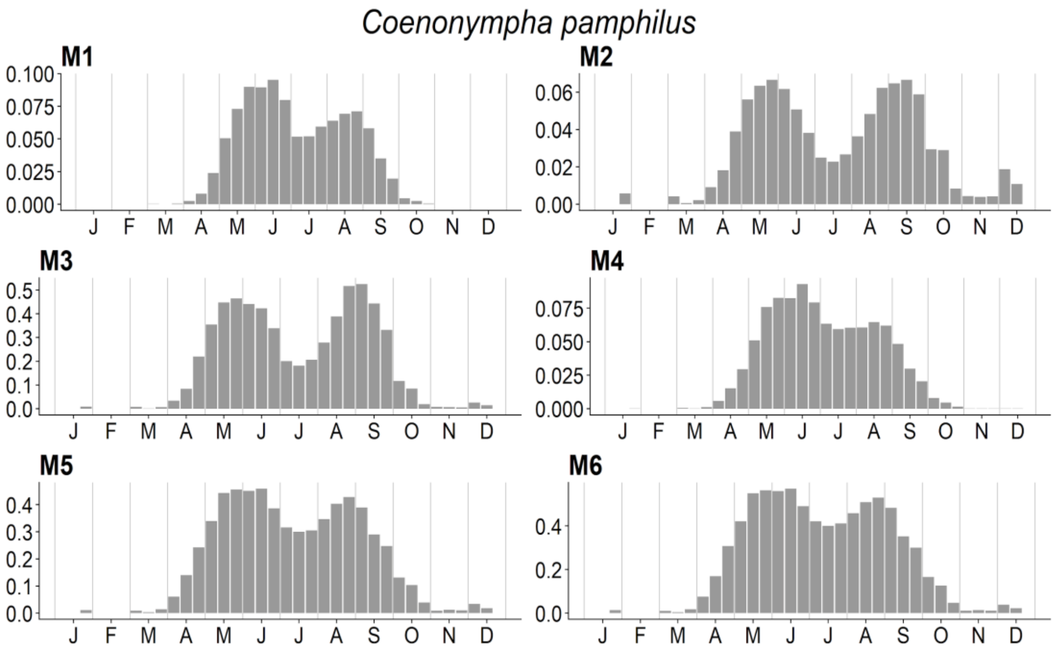
Cupido alcetas
M1
M2
M3
M4
M5
M6
J F M A M J J A S O N D

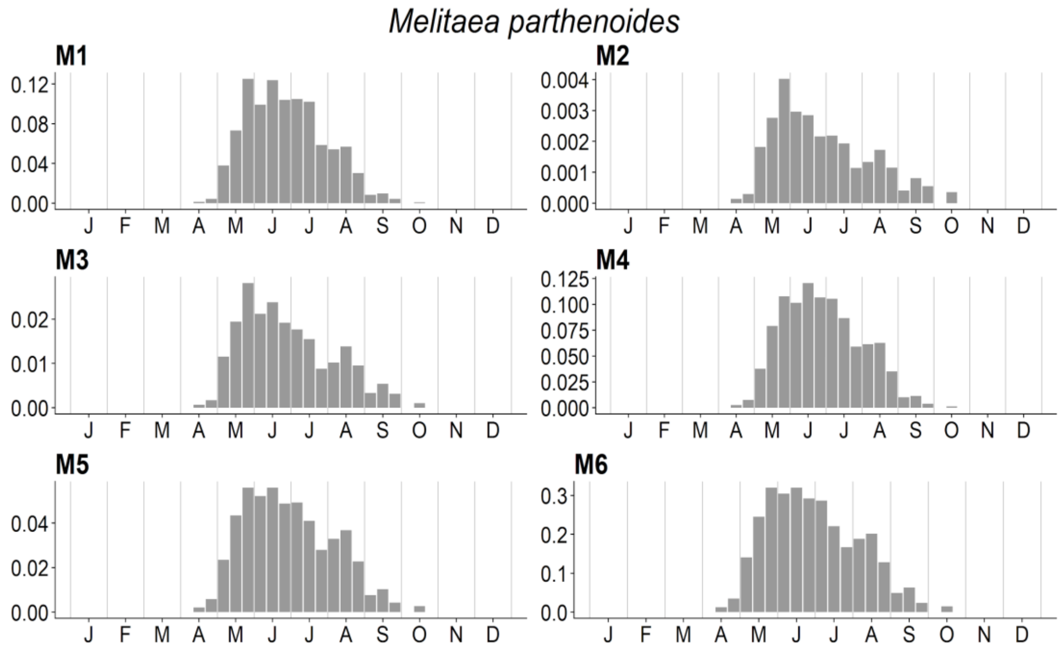

Lysandra hispana
M1
M2
M3
M4
M5
M6
J F M A M J J A S O N D

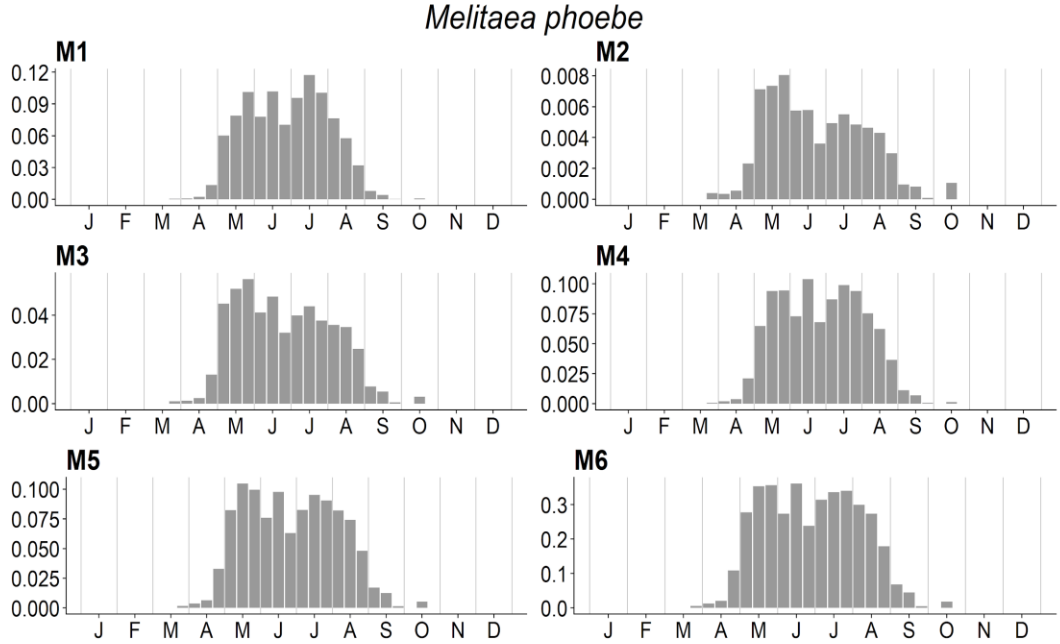

Plebejus argyrognomon
M1
M2
M3
M4
M5
M6
J F M A M J J A S O N D



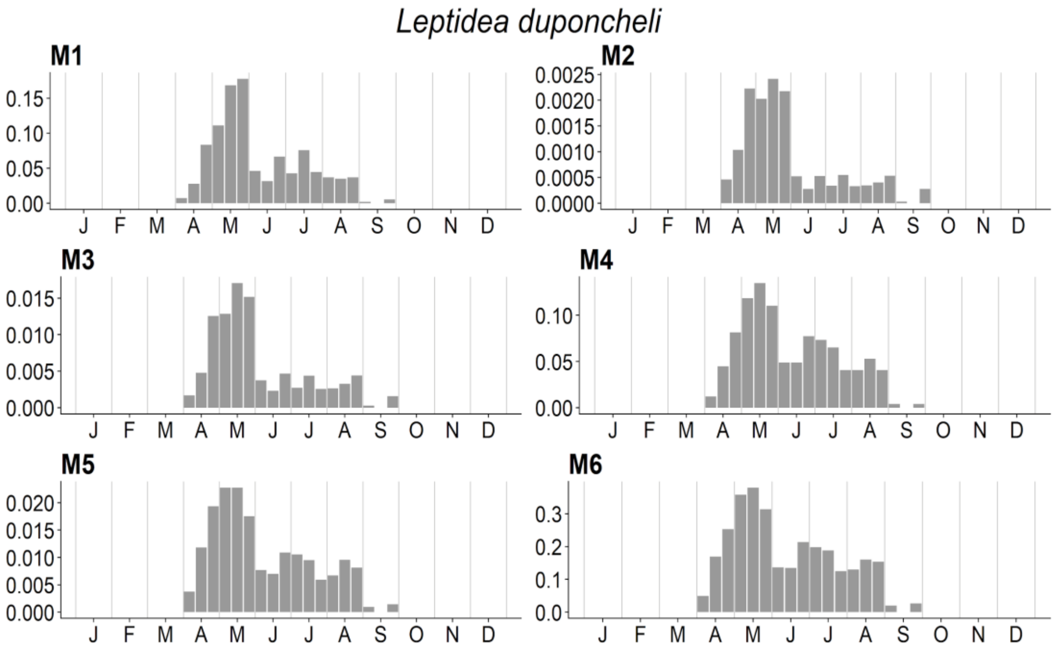
Nymphalidae:




Melitaea parthenoides
M1
M2
M3
M4
M5
M6
J F M A M J J A S O N D

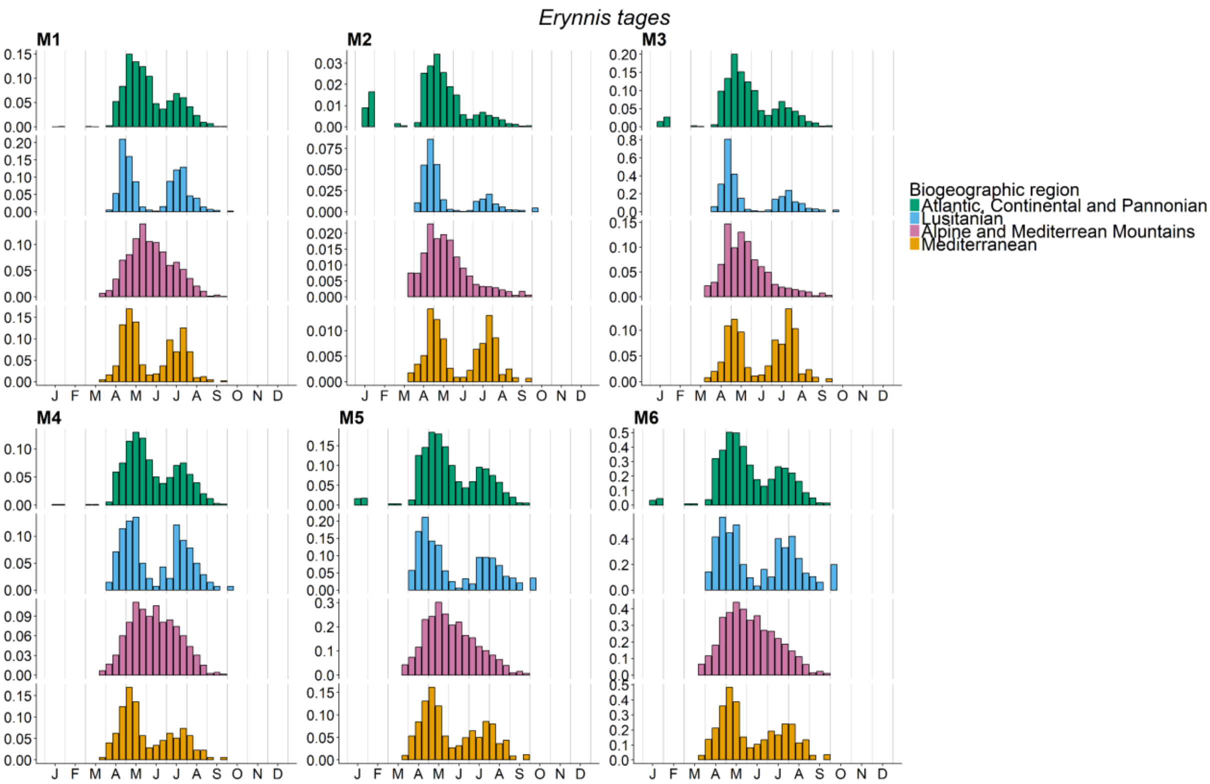

Melitaea phoebe
M1
M2
M3
M4
M5
M6
J F M A M J J A S O N D

Pieridae:

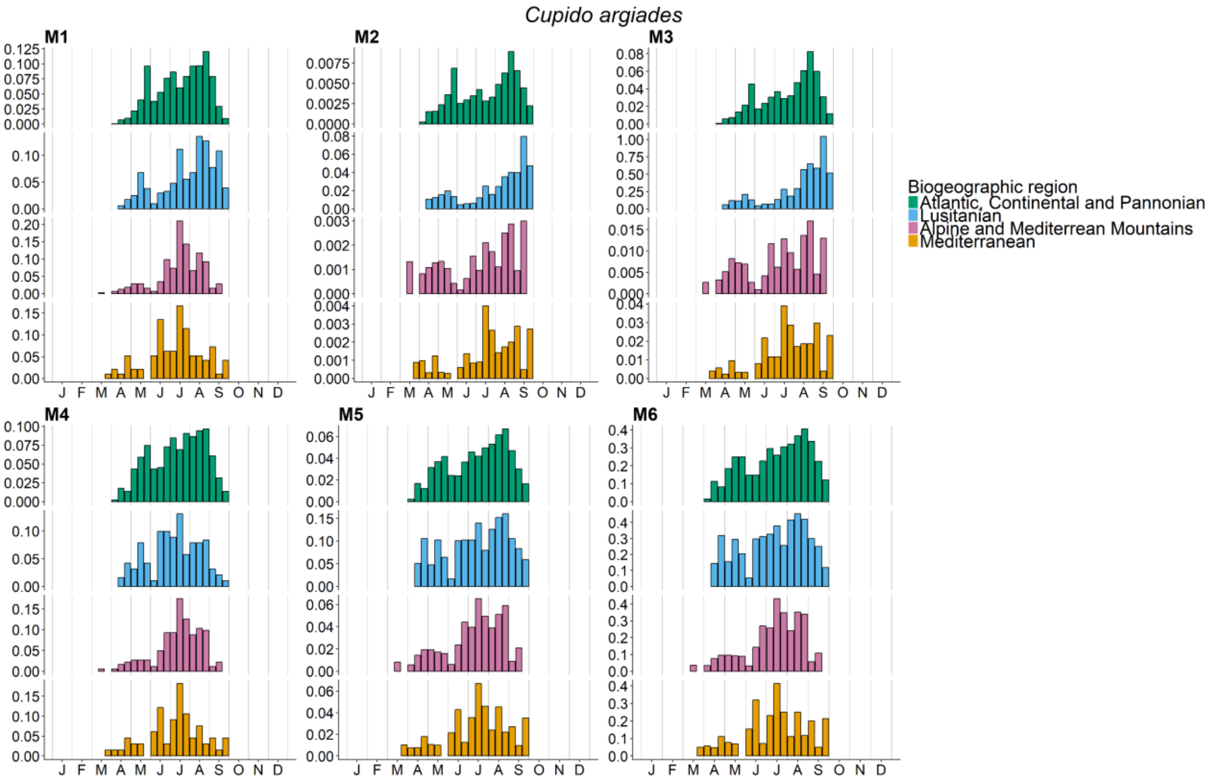


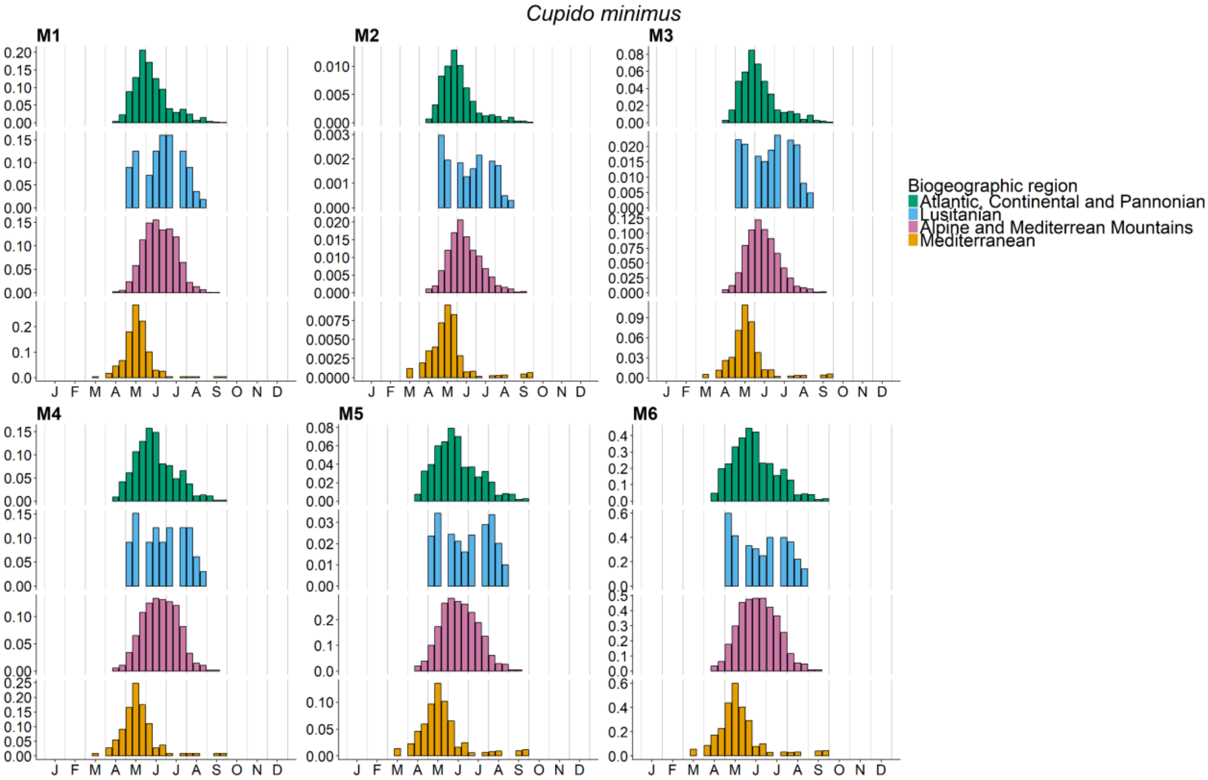
Appendix 4. Phenology of 20 butterfly species (1 Hesperiidae, 8 Lycaenidae, 7 Nymphalidae, 2 Papilionidae, 1 Pieridae, 1 Riodinidae) for which biogeographic effect was expected. Phenology is represented at four biogeographic regions with six indices (see also Tab. 2): proportion of records per period (M1); ratio of records to the group (M2); number of records per field visit (M3); proportion of quadrats per period (M4); ratio of quadrats to the group (M5); ratio of quadrats to the group within the species known distribution (M6).

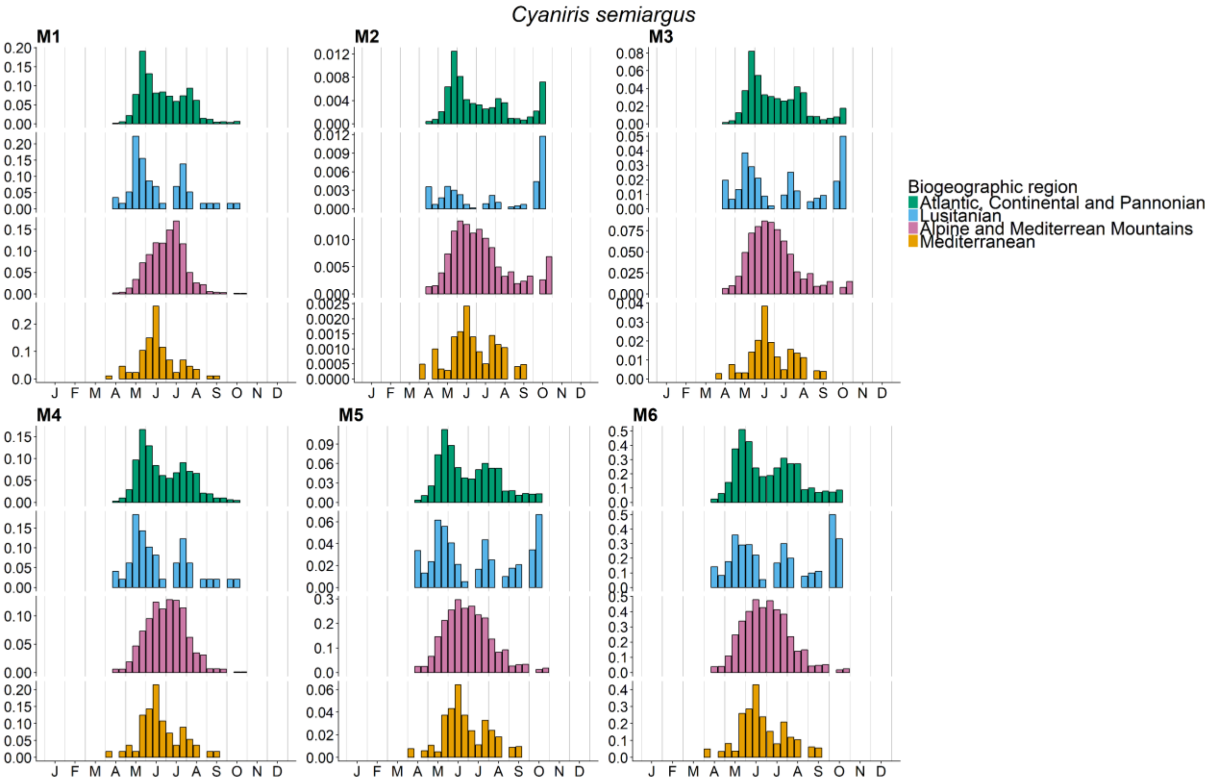
Hesperiidae:

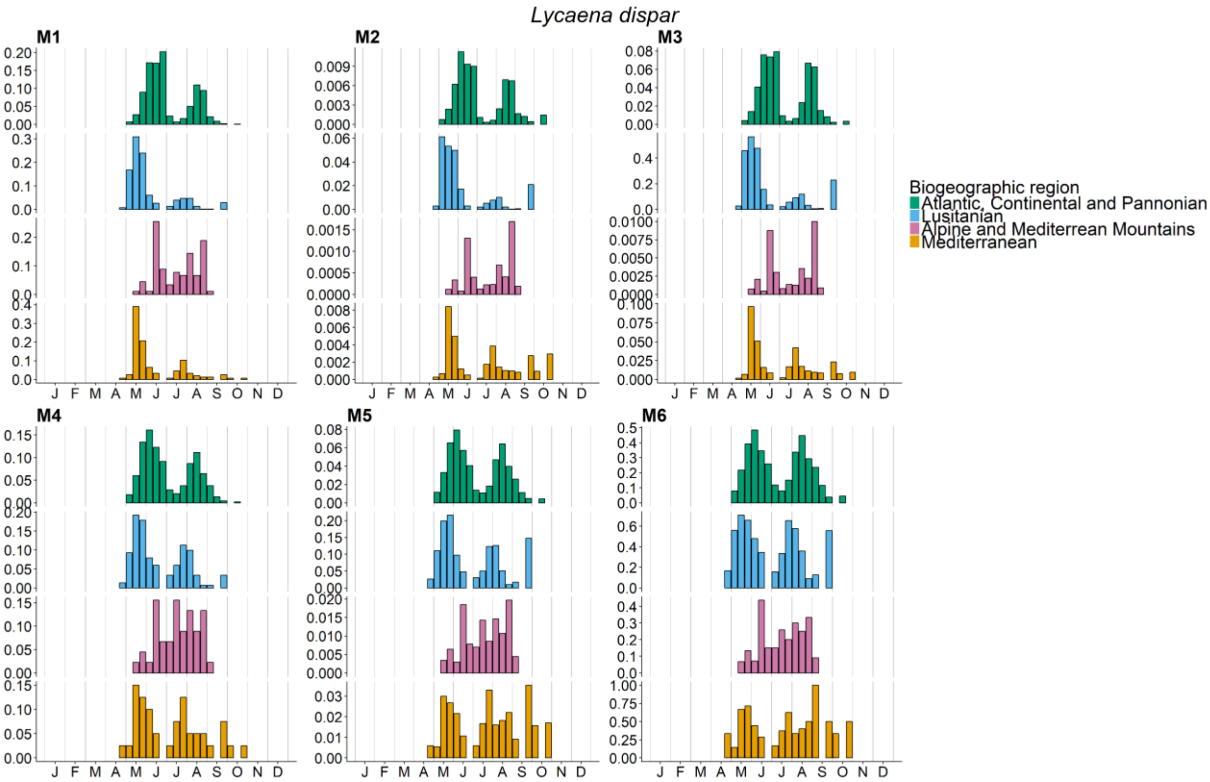


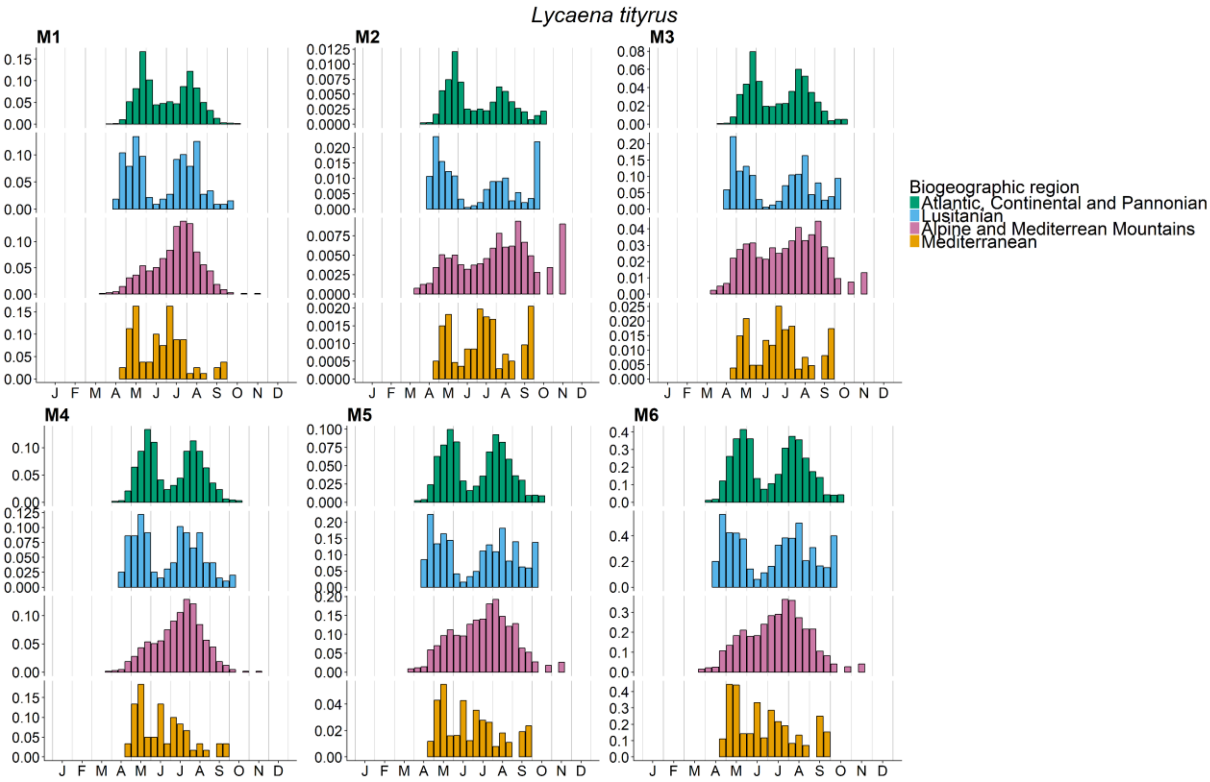
Lycaenidae:

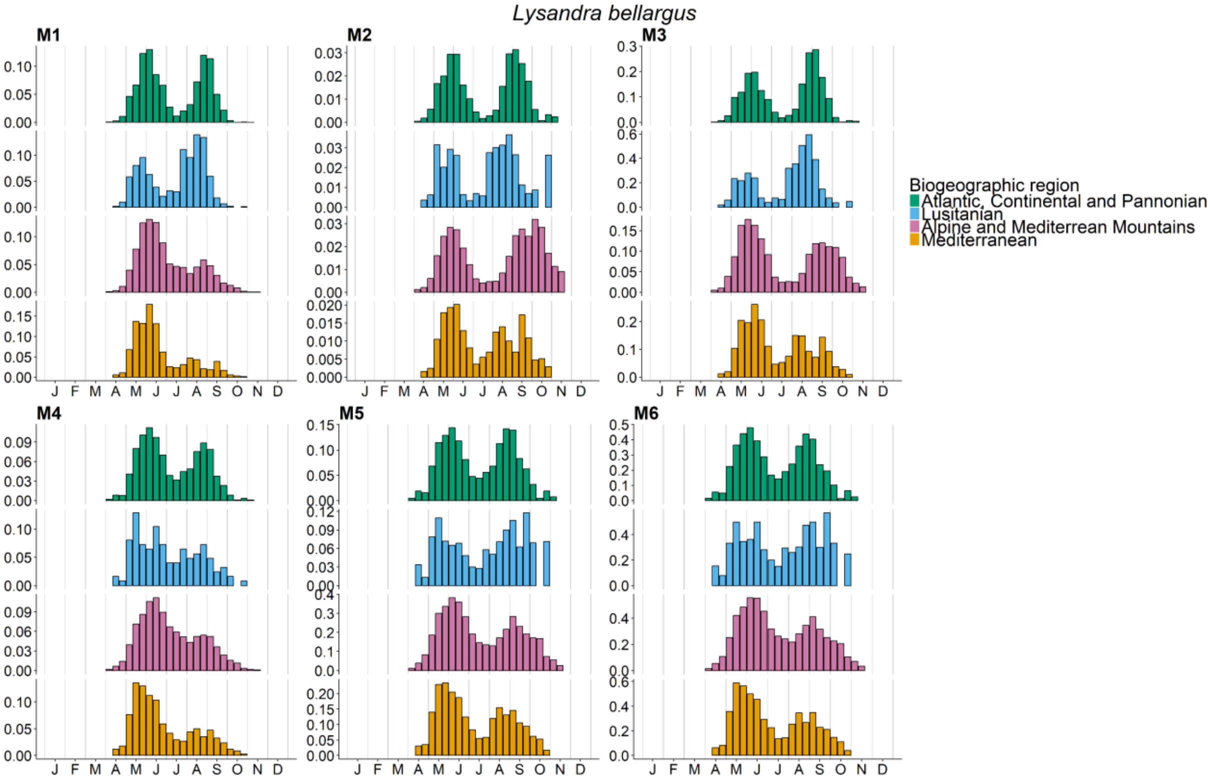

Cupido argiades
M1
M2
M3
M4
M5
M6
J F M A M J J A S O N D
Biogeographic region
Atlantic, Continental and Pannonian
Lusitanian
Alpine and Mediterrean Mountains
Mediterranean

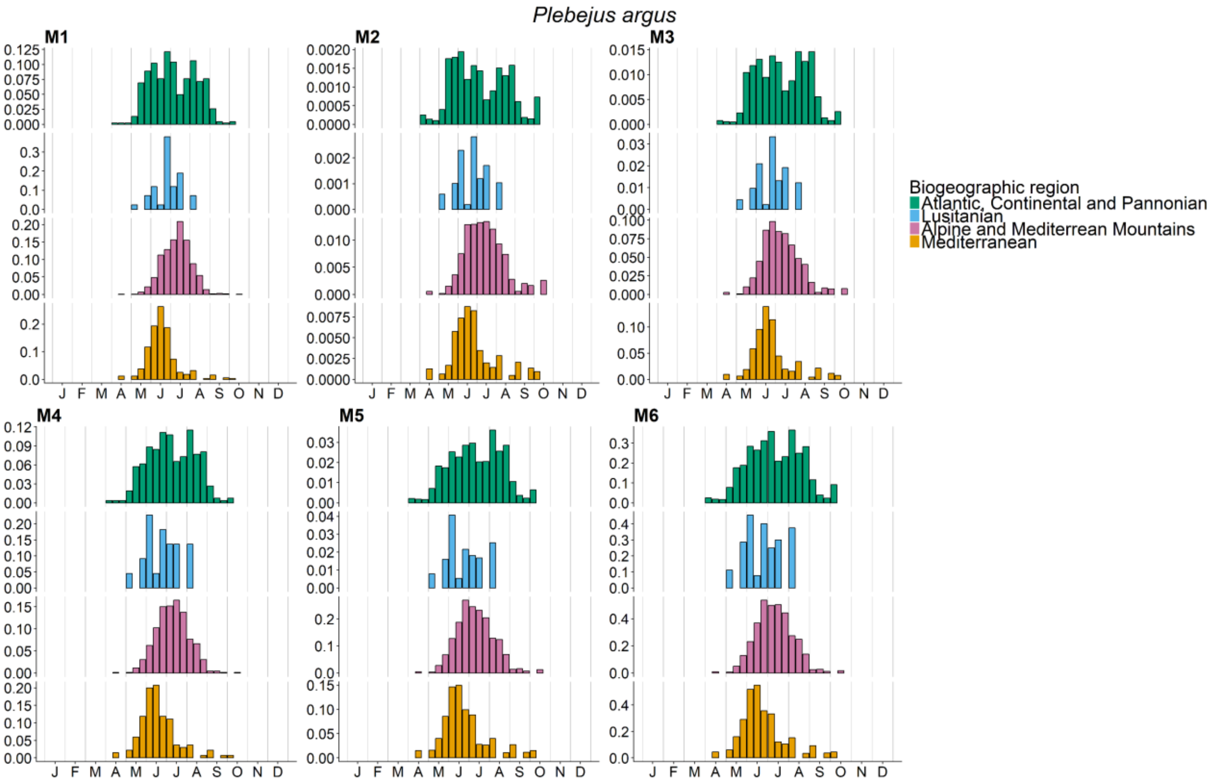

Cupido minimus
M1
M2
M3
M4
M5
M6
J F M A M J J A S O N D
Biogeographic region
Atlantic, Continental and Pannonian
Lusitanian
Alpine and Mediterrean Mountains
Mediterranean

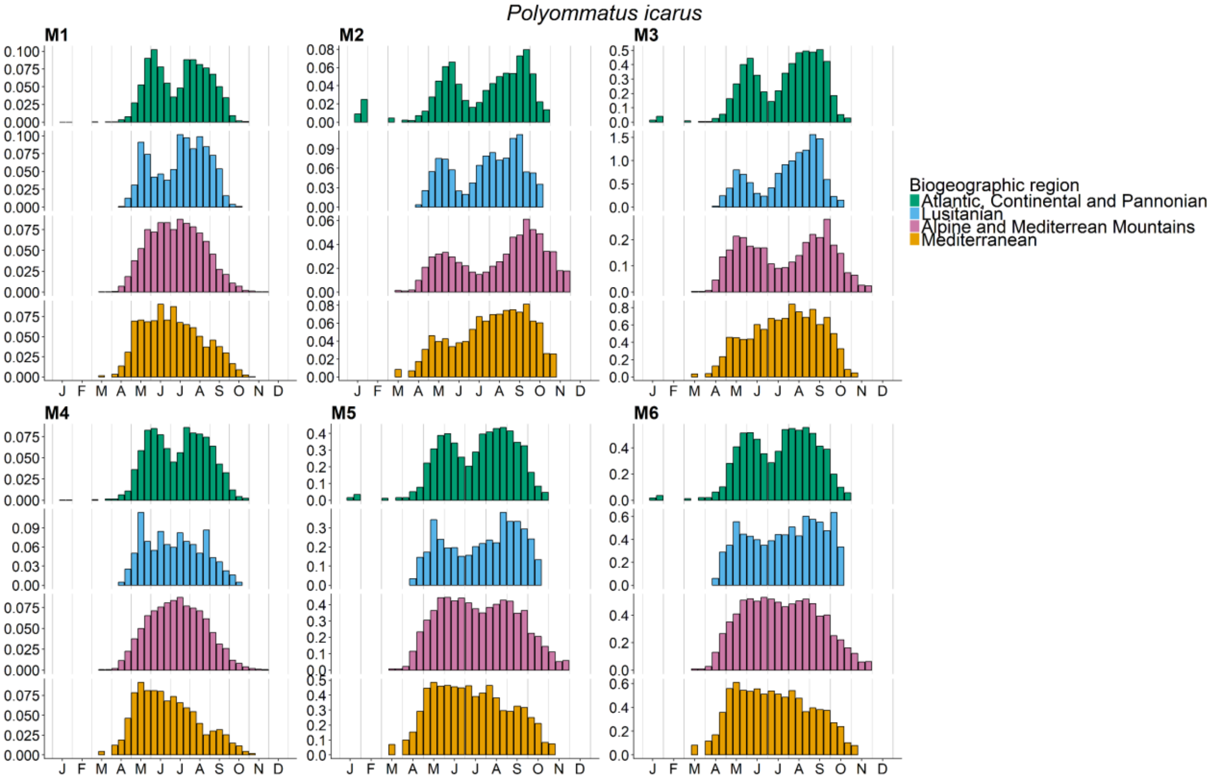

Cyaniris semiargus
M1
M2
M3
M4
M5
M6
J F M A M J J A S O N D
Biogeographic region
Atlantic, Continental and Pannonian
Lusitanian
Alpine and Mediterrean Mountains
Mediterranean


Lycaena dispar
M1
M2
M3
M4
M5
M6
J F M A M J J A S O N D
Biogeographic region
Atlantic, Continental and Pannonian
Lusitanian
Alpine and Mediterrean Mountains
Mediterranean

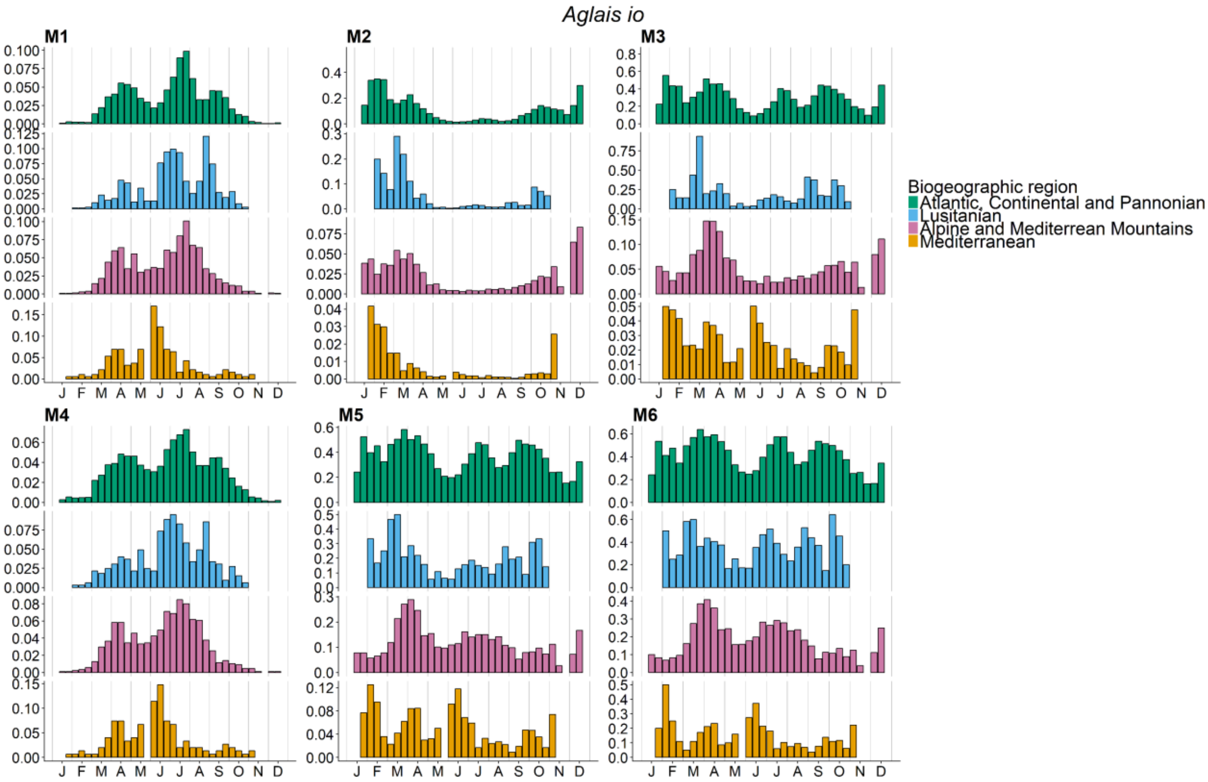

Lycaena tityrus
M1
M2
M3
M4
M5
M6
Biogeographic region
Atlantic, Continental and Pannonian
Lusitanian
Alpine and Mediterrean Mountains
Mediterranean
J F M A M J J A S O N D

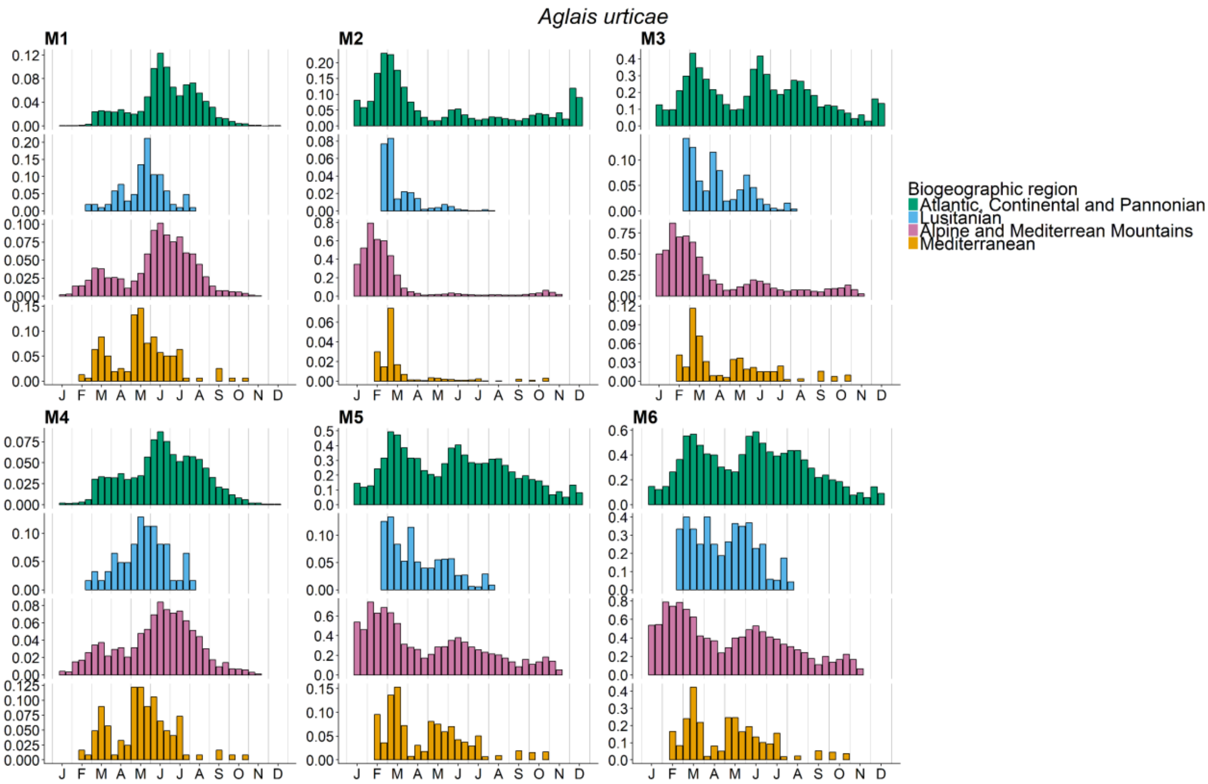

Lysandra bellargus
M1
M2
M3
M4
M5
M6
Biogeographic region
Atlantic, Continental and Pannonian
Lusitanian
Alpine and Mediterrean Mountains
Mediterranean
J F M A M J J A S O N D

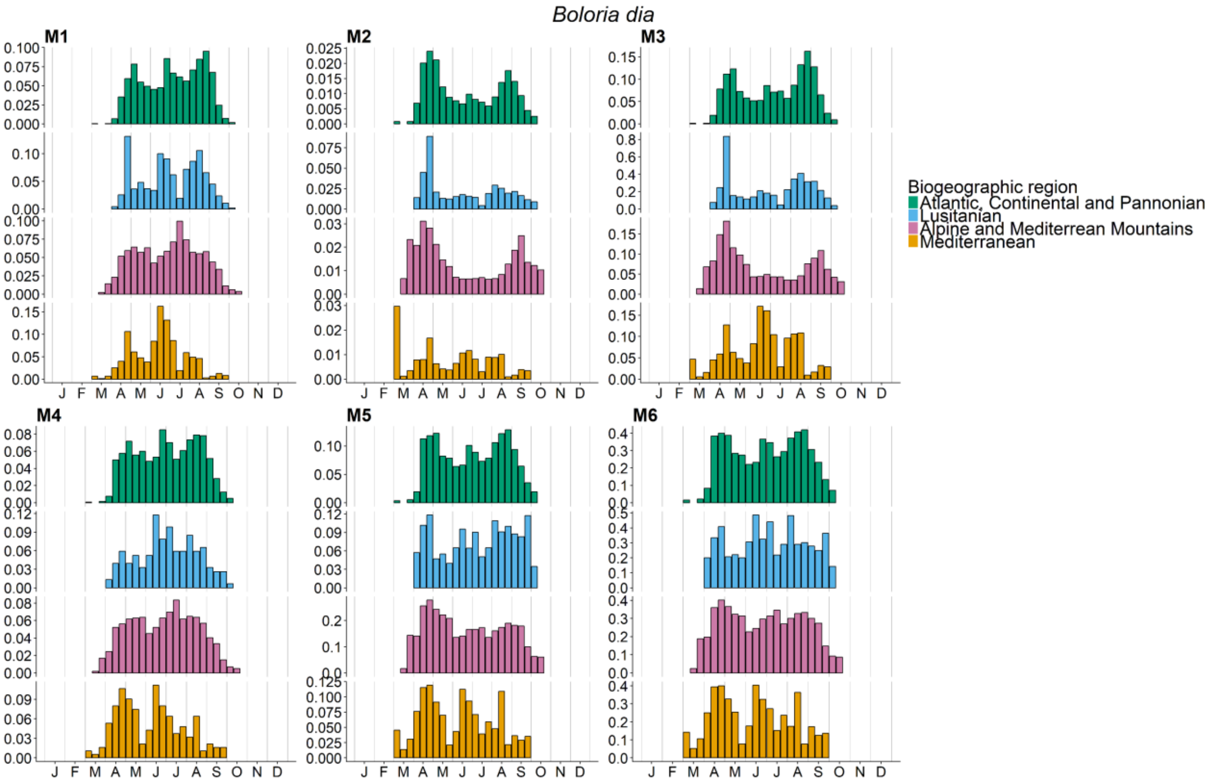


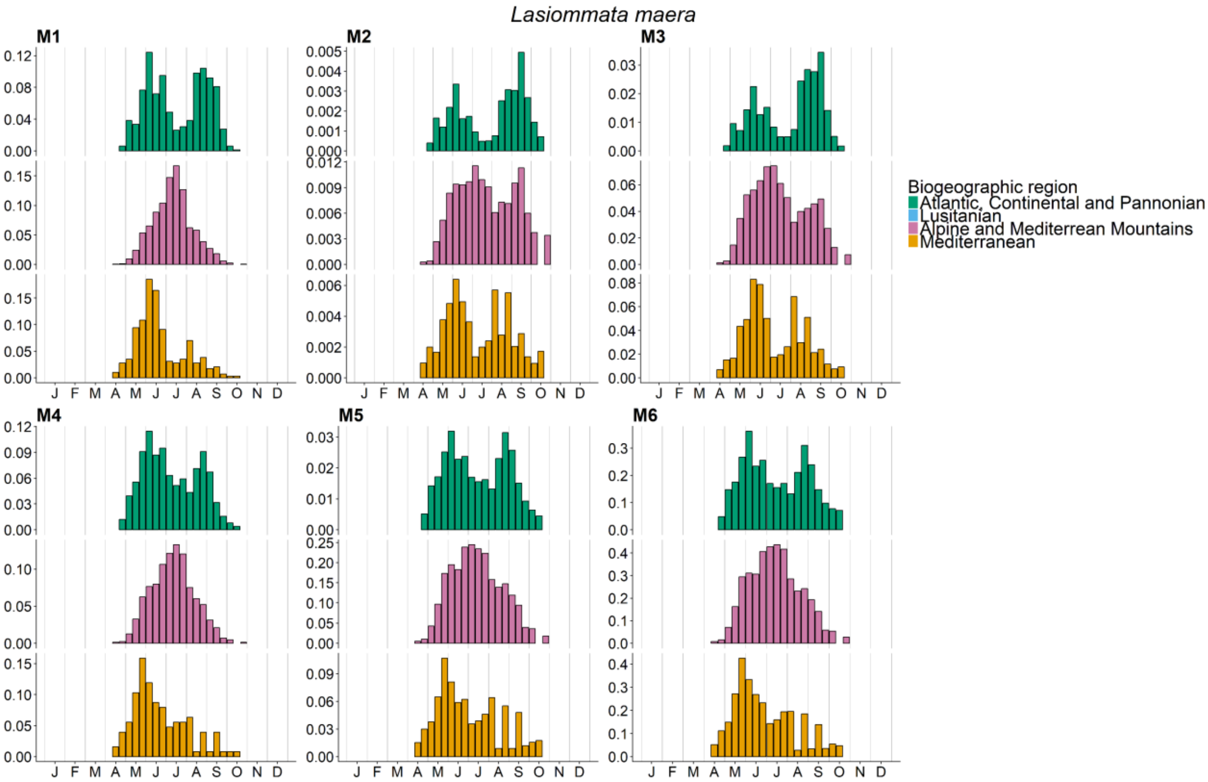


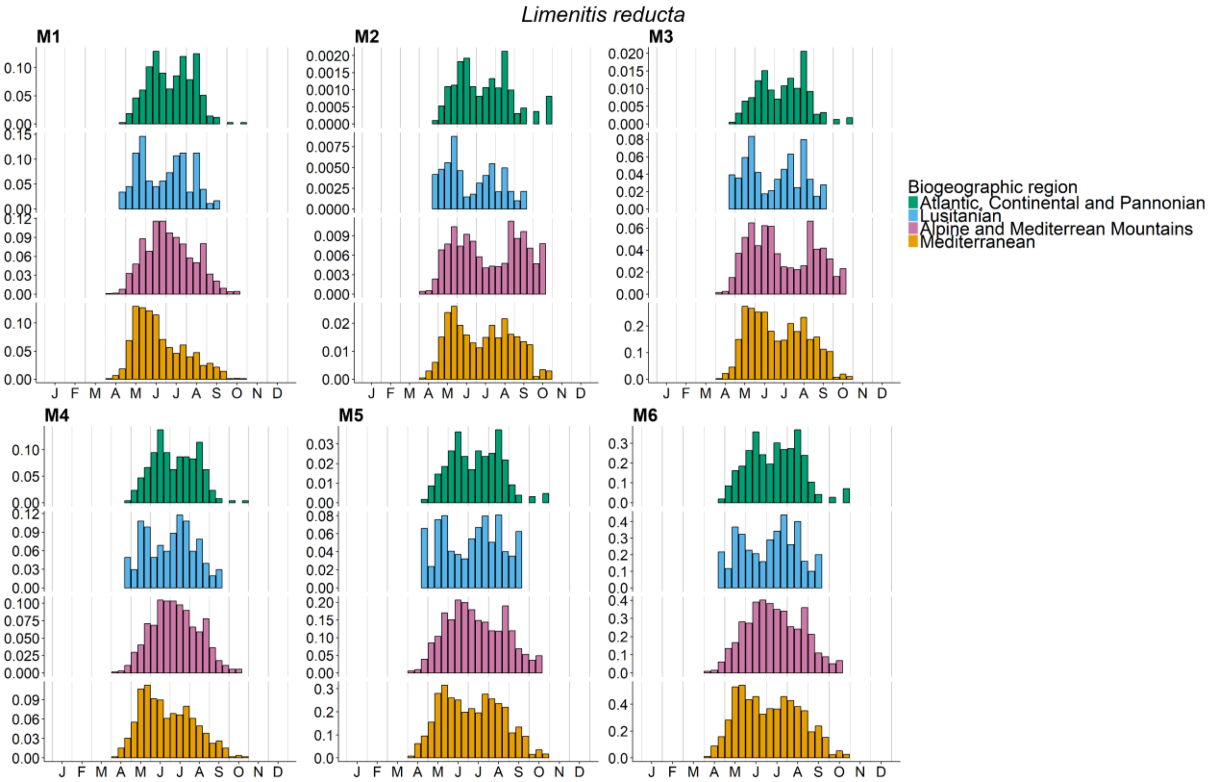
Nymphalidae:


Aglais io
M1
M2
M3
M4
M5
M6
J F M A M J J A S O N D
Biogeographic region
Atlantic, Continental and Pannonian
Lusitanian
Alpine and Mediterrean Mountains
Mediterranean

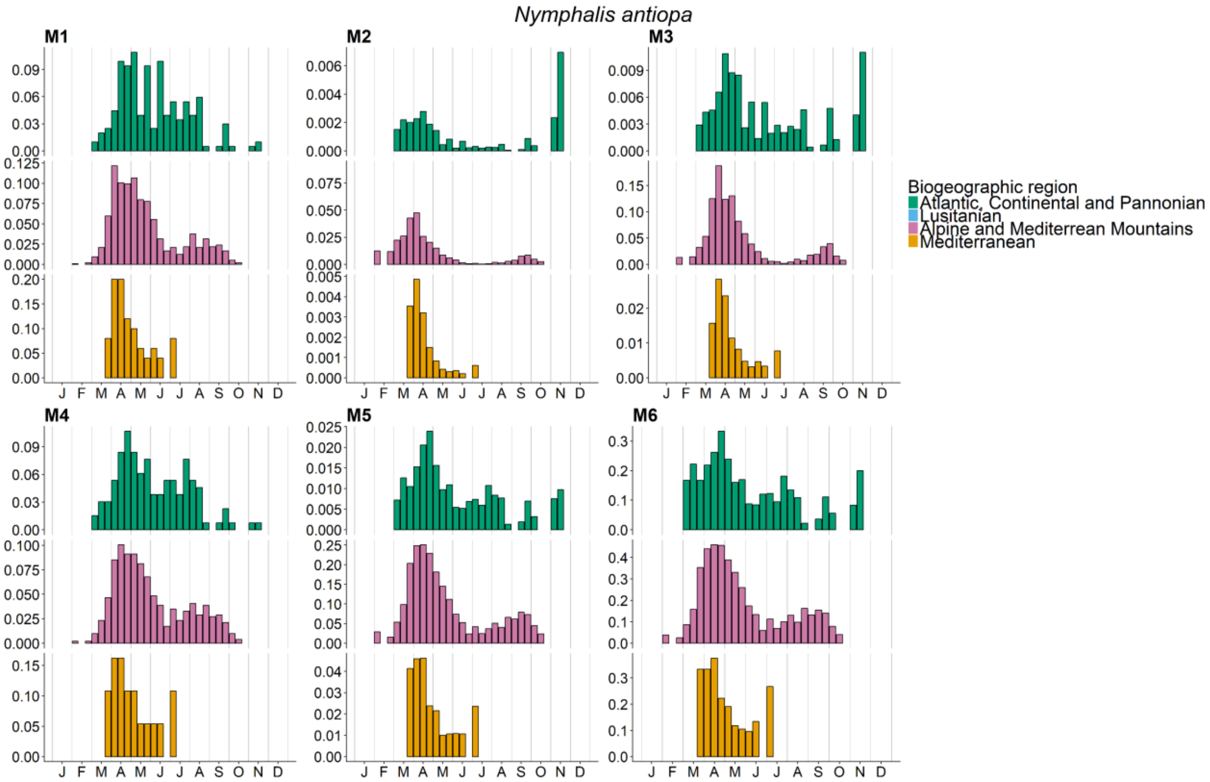

Aglais urticae
M1
M2
M3
M4
M5
M6
J F M A M J J A S O N D
Biogeographic region
Atlantic, Continental and Pannonian
Lusitanian
Alpine and Mediterrean Mountains
Mediterranean


Boloria dia
M1
M2
M3
M4
M5
M6
J F M A M J J A S O N D
Biogeographic region
Atlantic, Continental and Pannonian
Lusitanian
Alpine and Mediterrean Mountains
Mediterranean

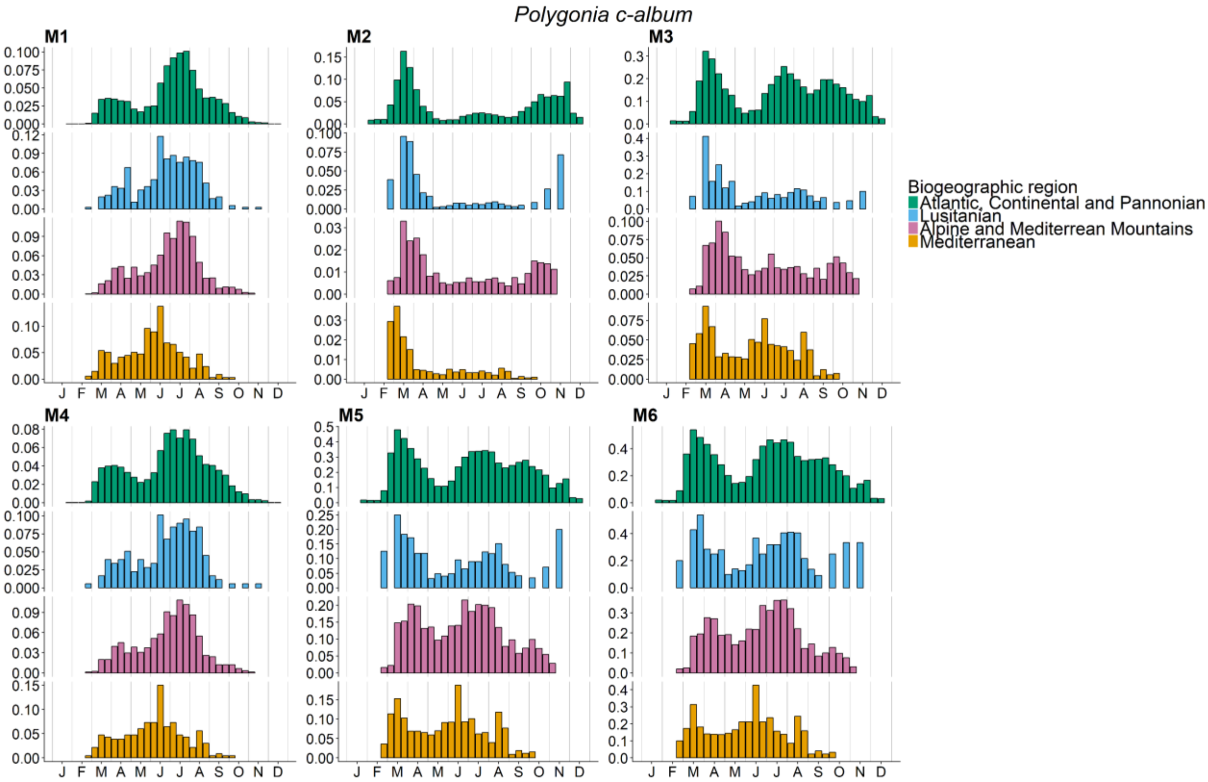

Lasiommata maera
M1
M2
M3
M4
M5
M6
J F M A M J J A S O N D
Biogeographic region
Atlantic, Continental and Pannonian
Lusitanian
Alpine and Mediterrean Mountains
Mediterranean



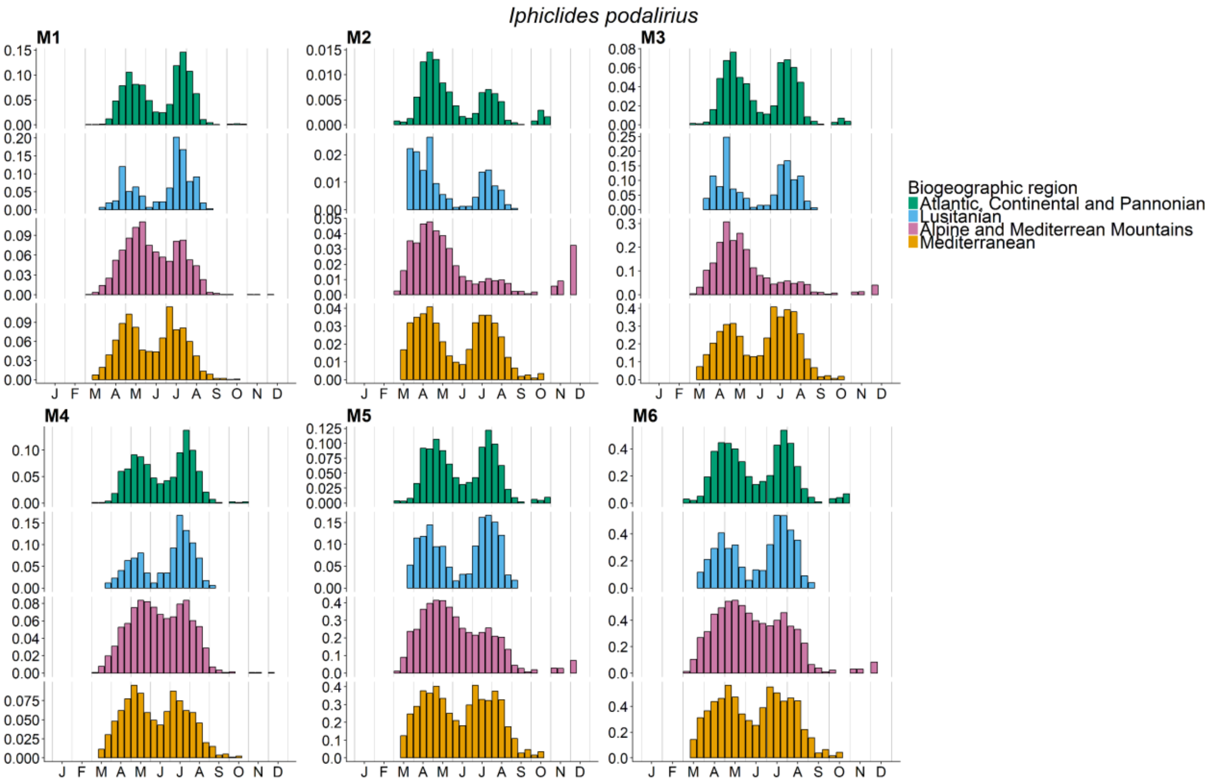
Limenitis reducta
M1
M2
M3
M4
M5
M6
J F M A M J J A S O N D
Biogeographic region
Atlantic, Continental and Pannonian
Lusitanian
Alpine and Mediterrean Mountains
Mediterranean

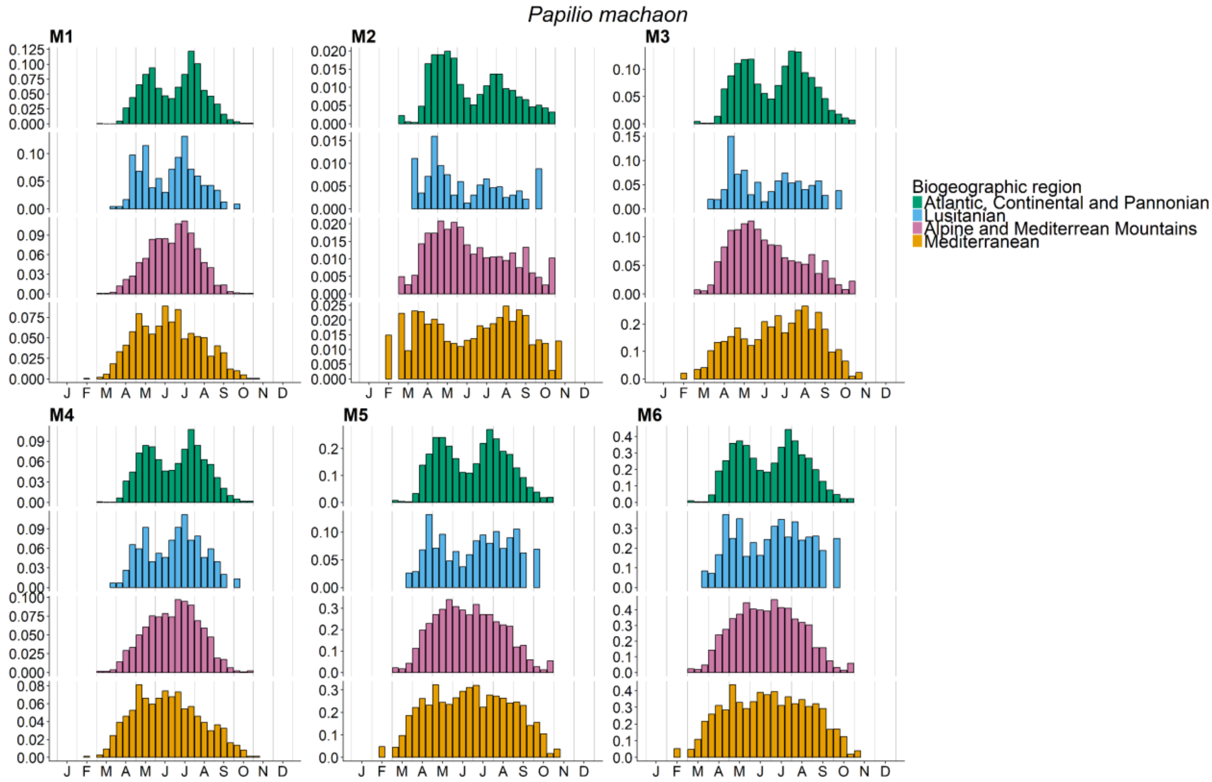


Nymphalis antiopa
M1
M2
M3
M4
M5
M6
J F M A M J J A S O N D
Biogeographic region
Atlantic, Continental and Pannonian
Lusitanian
Alpine and Mediterrean Mountains
Mediterranean

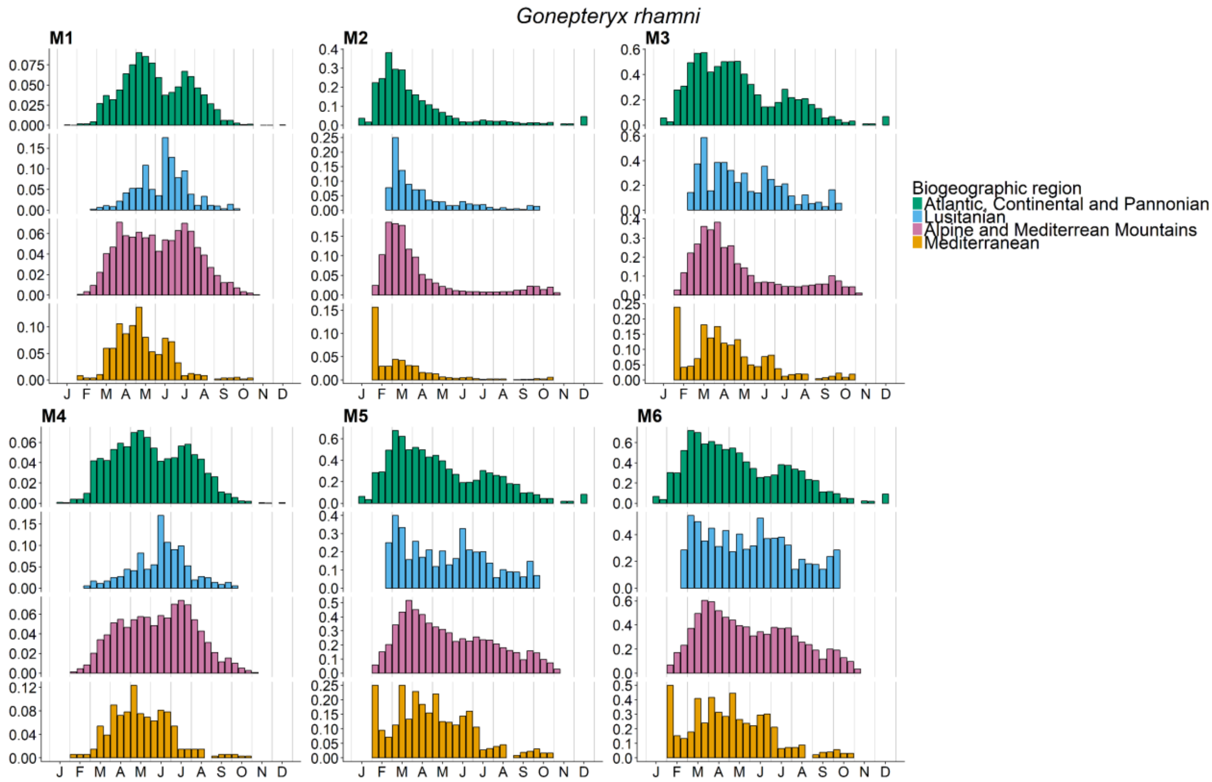


Papilionidae

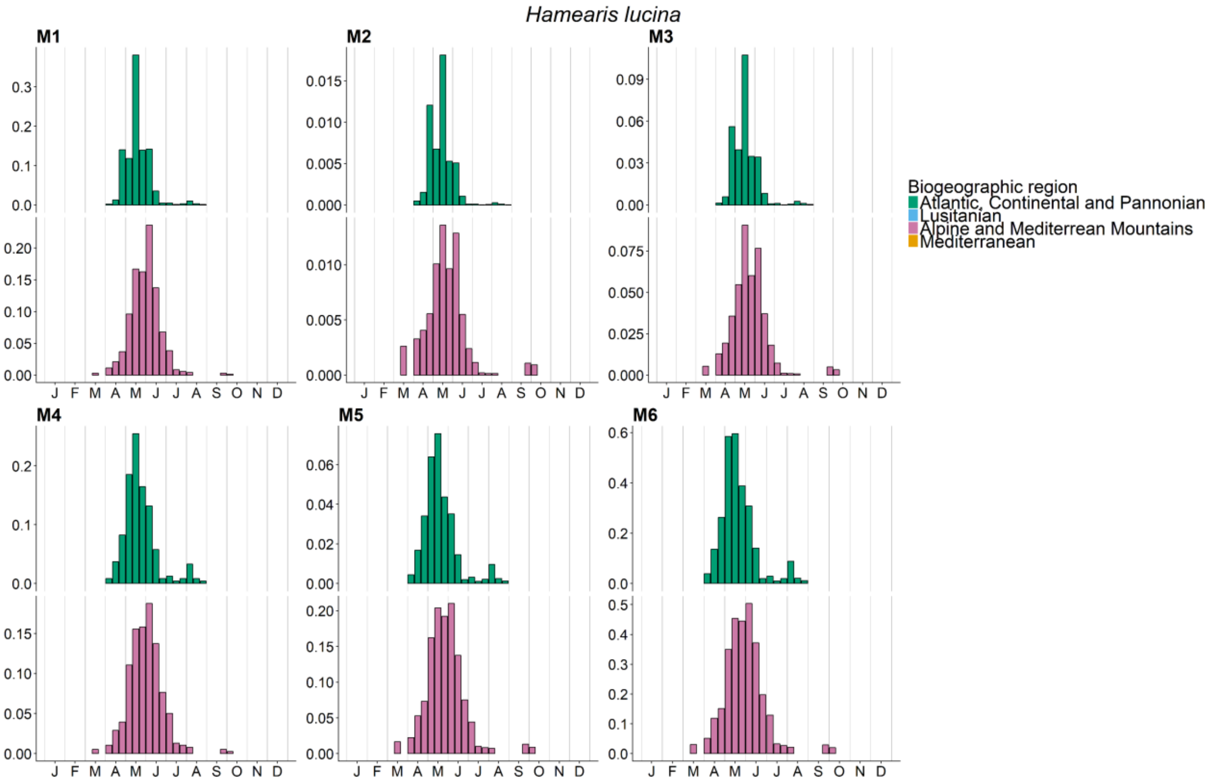




Pieridae:



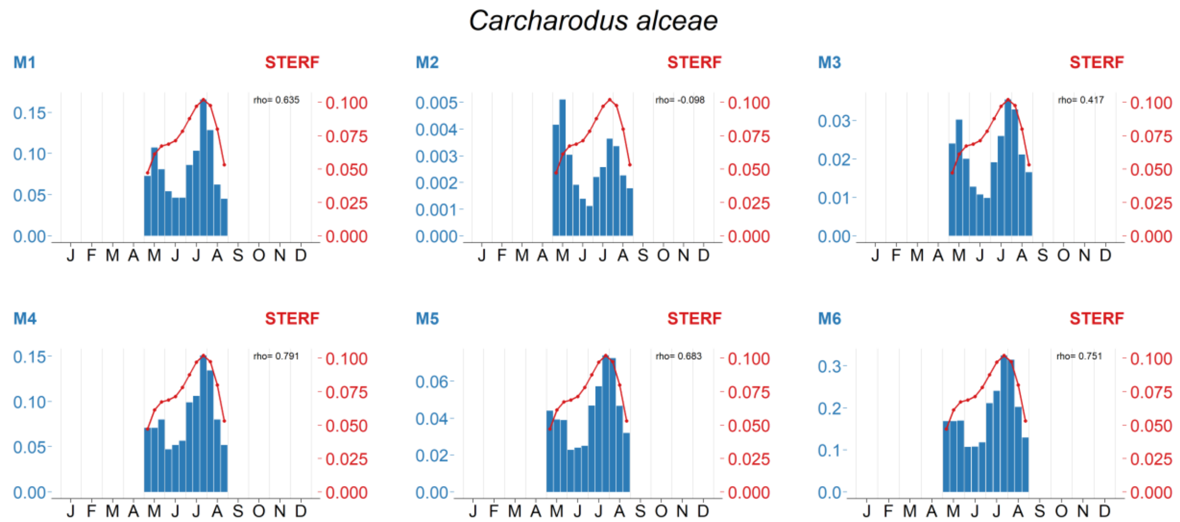
Riodinidae:

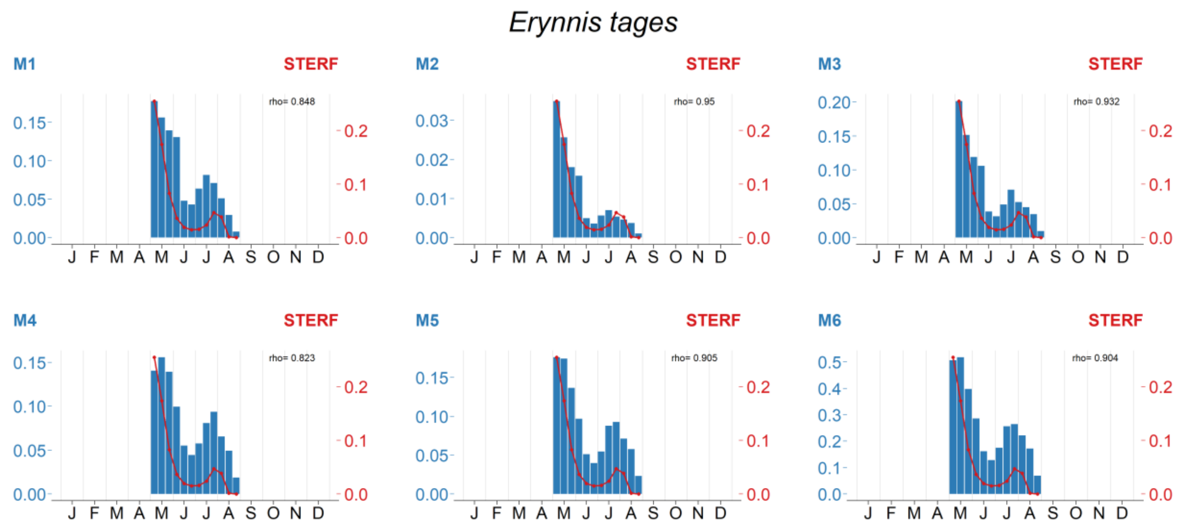


Appendix 5. Comparison between the STERF and six phenology indices computed on INPN data by ten-day periods for 57 species. A Pearson's correlation coefficient (rho) was calculated between every index (blue bar plots) and STERF count estimates (red line). Indices and

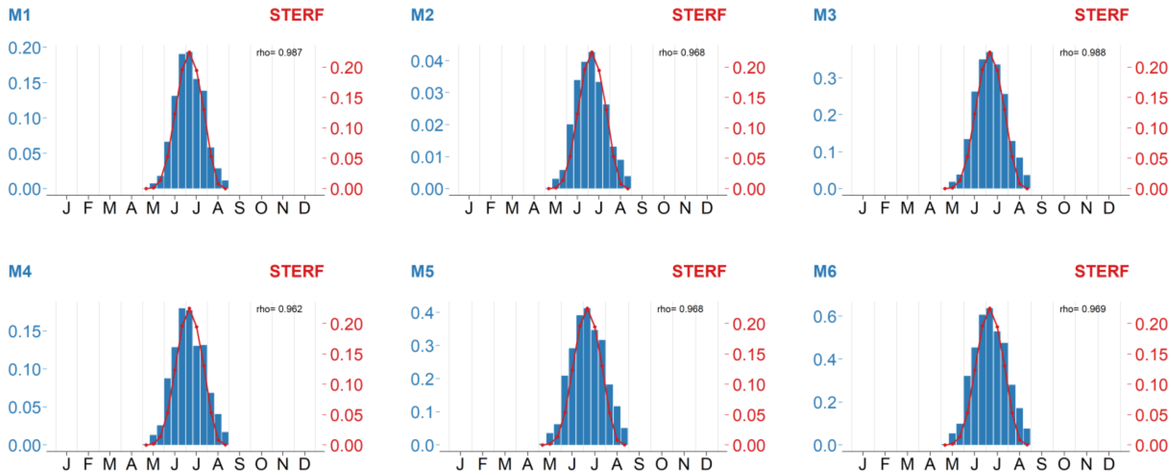
STERF count estimates were calculated from data collected from 2005, between May and August, in the ATCONP biogeographic region only.

Hesperiidae:

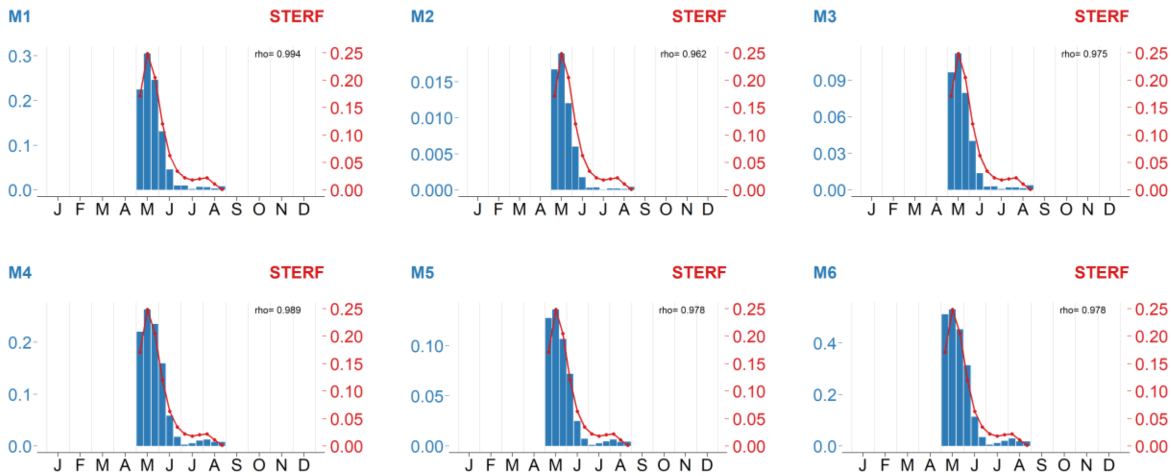




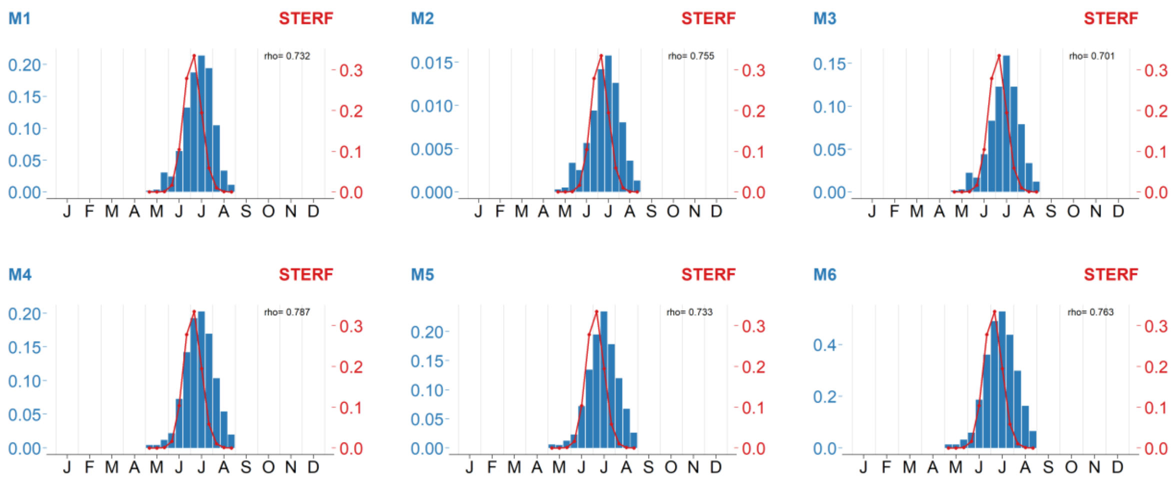
Ochlodes sylvanus

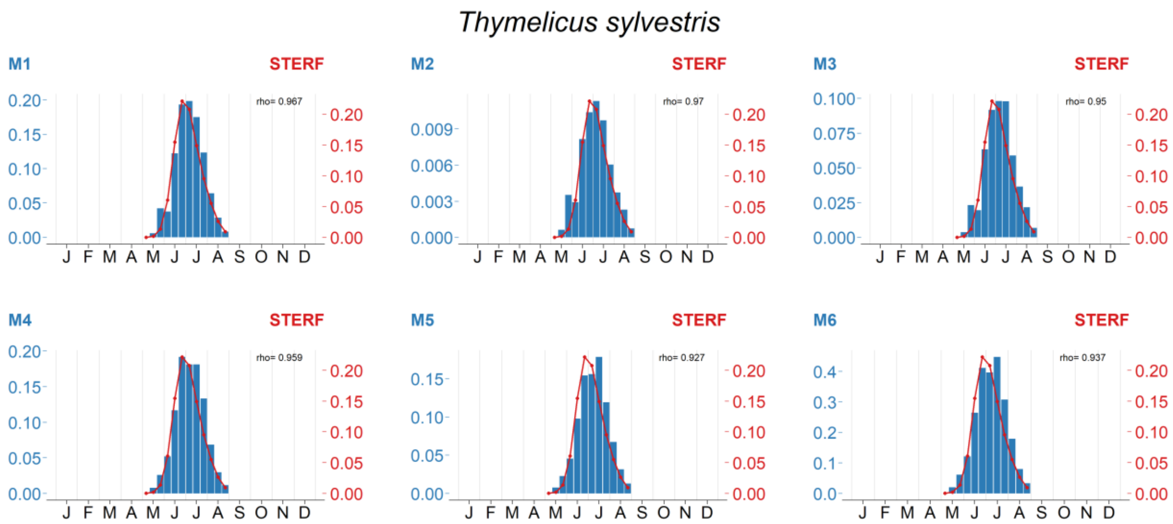


Pyrgus malvae

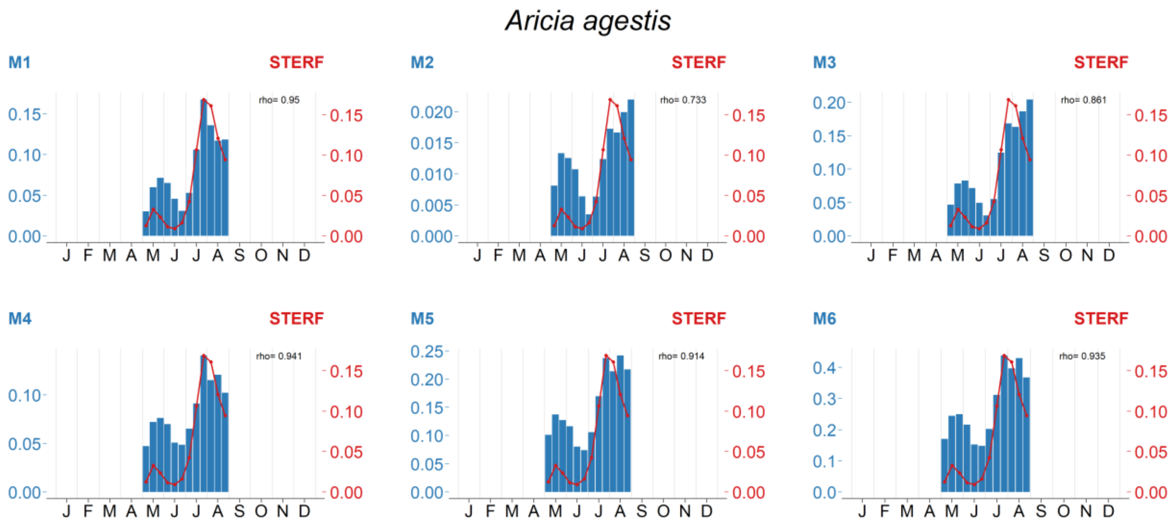


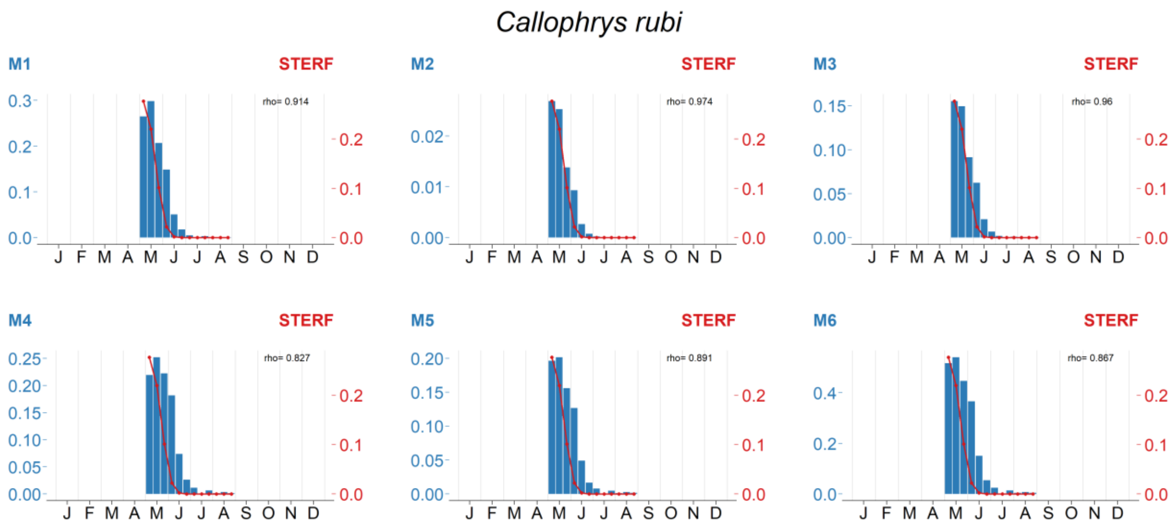
Thymelicus lineola



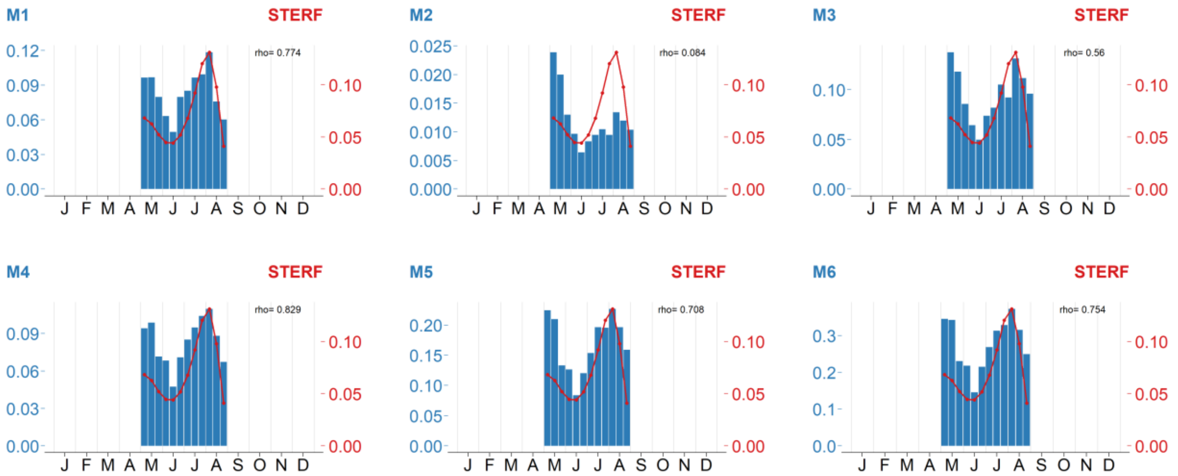

Thymelicus sylvestris
M1
STERF
rho= 0.967
M2
STERF
rho= 0.97
M3
STERF
rho= 0.95
M4
STERF
rho= 0.959
M5
STERF
rho= 0.927
M6
STERF
rho= 0.937
J F M A M J J A S O N D

Lycaenidae:

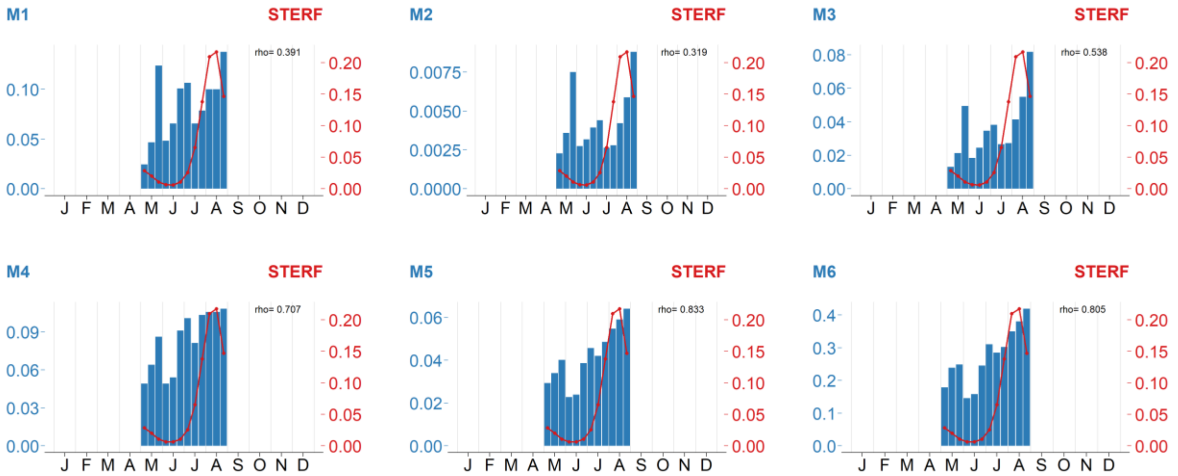

Aricia agestis
M1
STERF
rho= 0.95
M2
STERF
rho= 0.733
M3
STERF
rho= 0.861
M4
STERF
rho= 0.941
M5
STERF
rho= 0.914
M6
STERF
rho= 0.935
J F M A M J J A S O N D


Callophrys rubi
M1
STERF
rho= 0.914
M2
STERF
rho= 0.974
M3
STERF
rho= 0.96
M4
STERF
rho= 0.827
M5
STERF
rho= 0.891
M6
STERF
rho= 0.867
J F M A M J J A S O N D

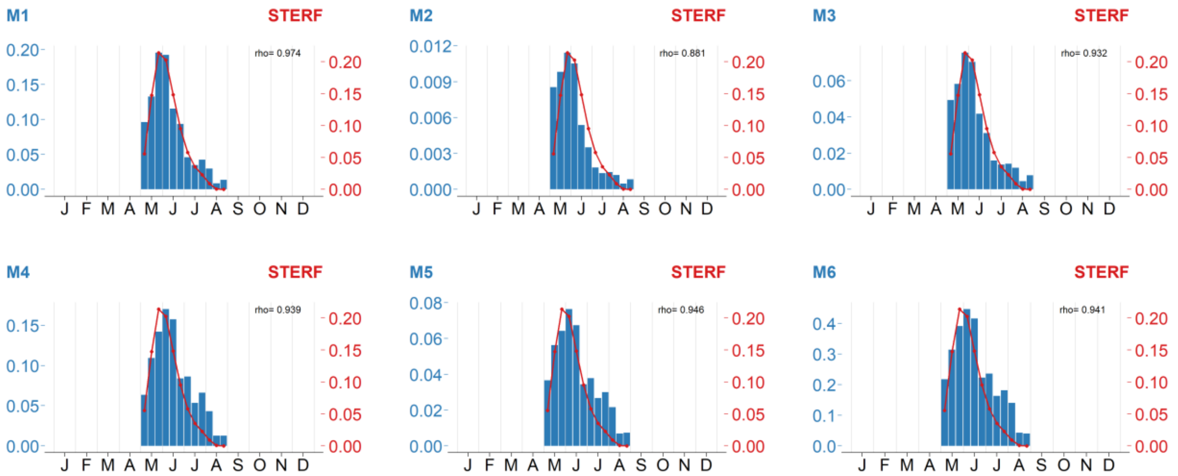
Celastrina argiolus

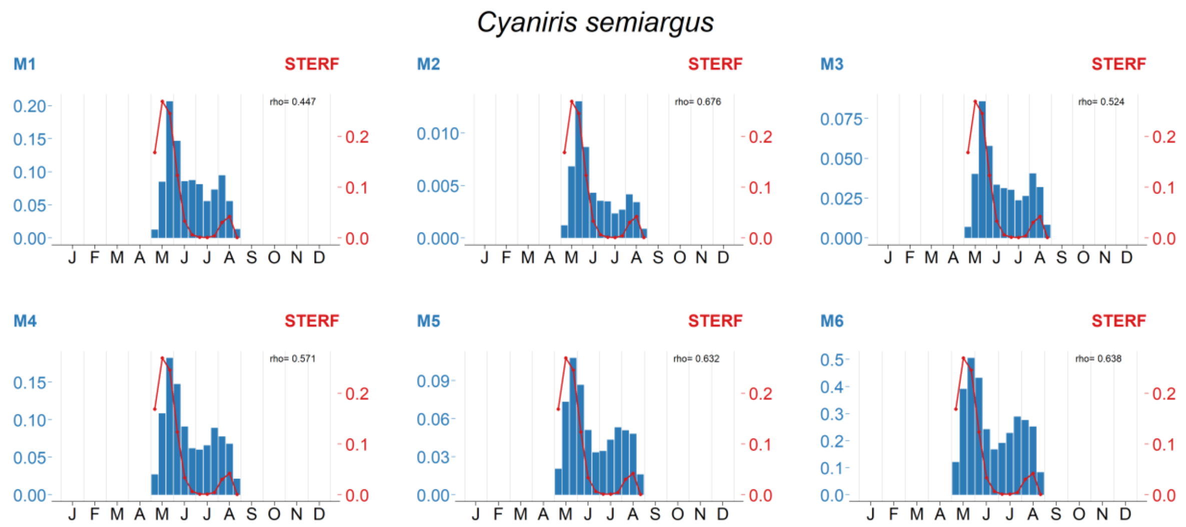


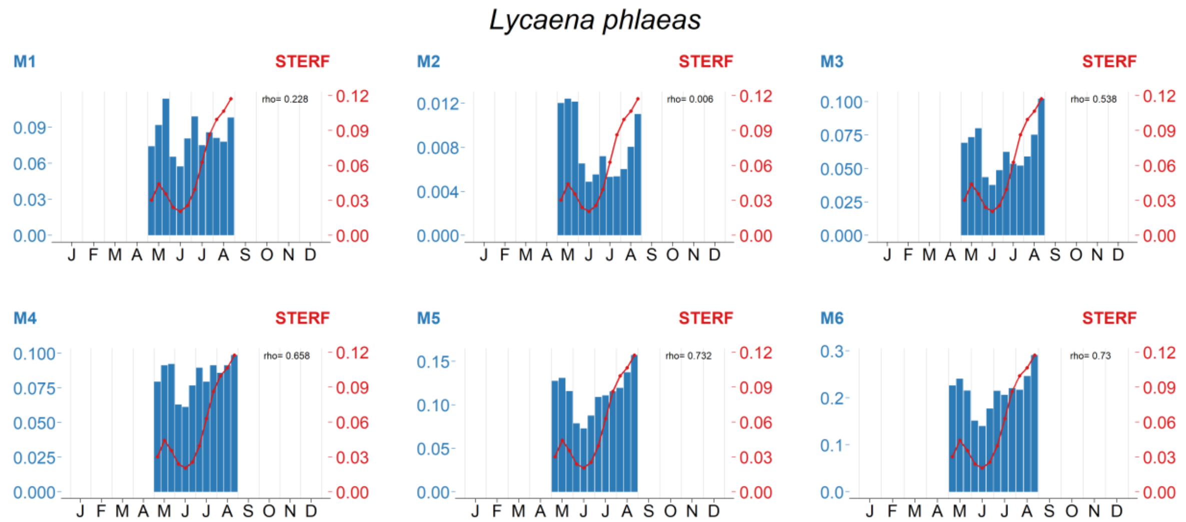
Cupido argiades

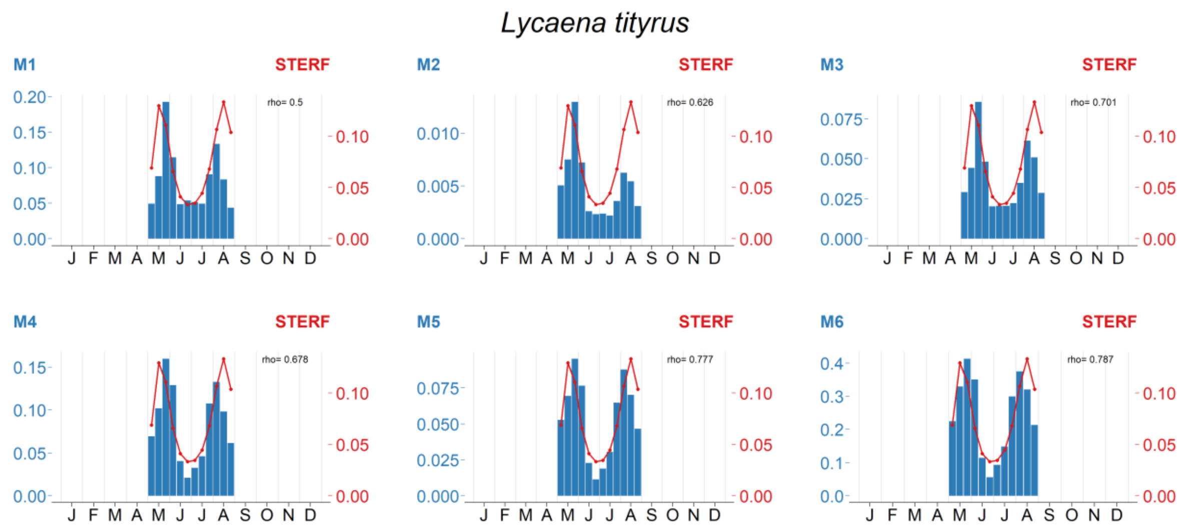


Cupido minimus

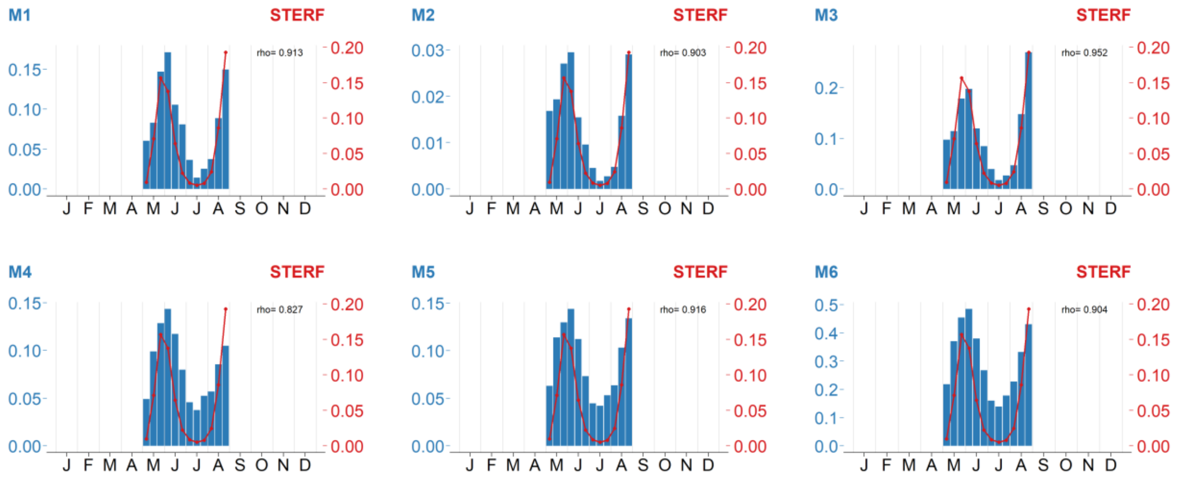



Cyaniris semiargus
M1
STERF
rho= 0.447
M2
STERF
rho= 0.676
M3
STERF
rho= 0.524
M4
STERF
rho= 0.571
M5
STERF
rho= 0.632
M6
STERF
rho= 0.638

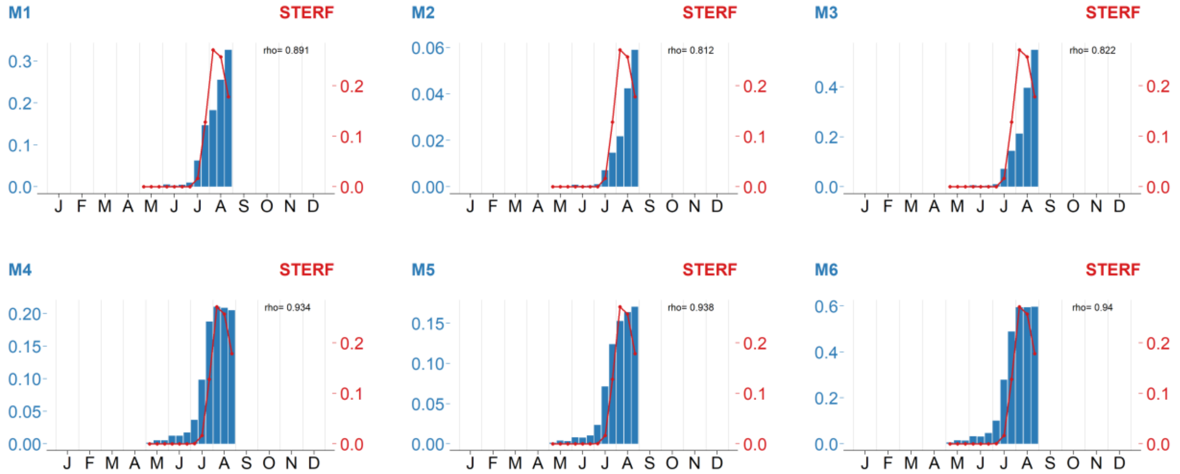

Lycaena phlaeas
M1
STERF
rho= 0.228
M2
STERF
rho= 0.006
M3
STERF
rho= 0.538
M4
STERF
rho= 0.658
M5
STERF
rho= 0.732
M6
STERF
rho= 0.73


Lycaena tityrus
M1
STERF
rho= 0.5
M2
STERF
rho= 0.626
M3
STERF
rho= 0.701
M4
STERF
rho= 0.678
M5
STERF
rho= 0.777
M6
STERF
rho= 0.787

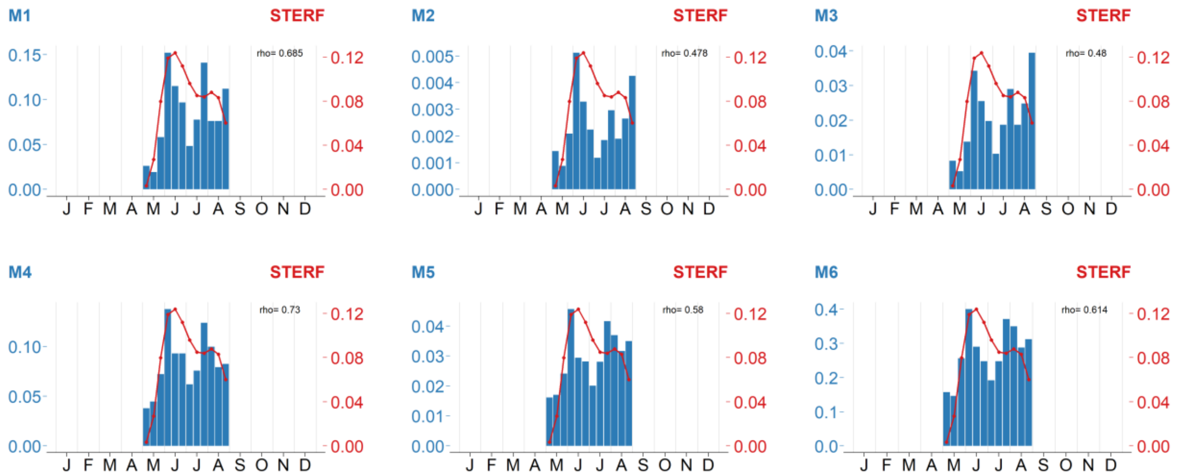
Lysandra bellargus

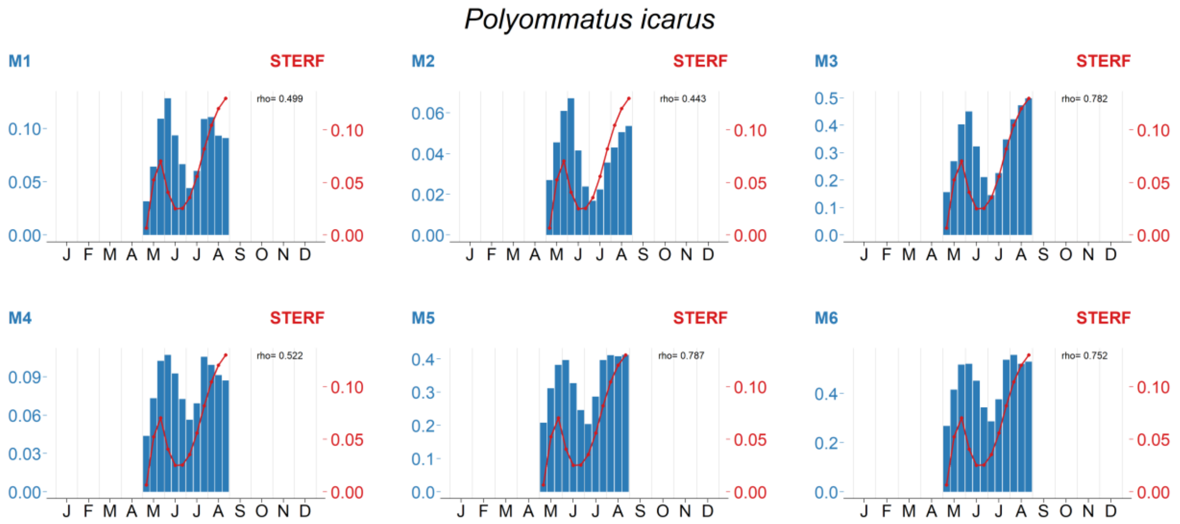


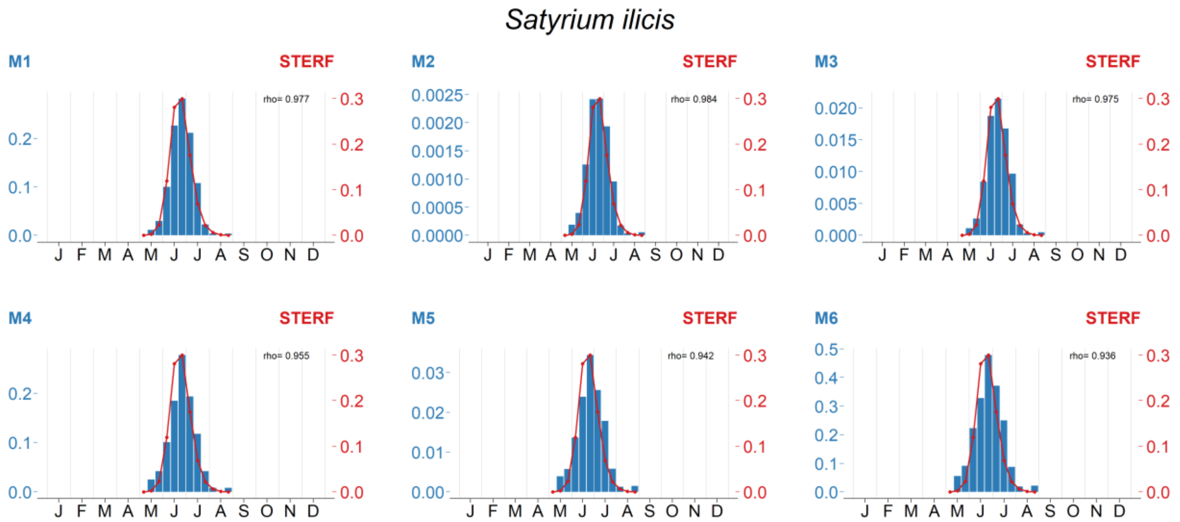
Lysandra coridon



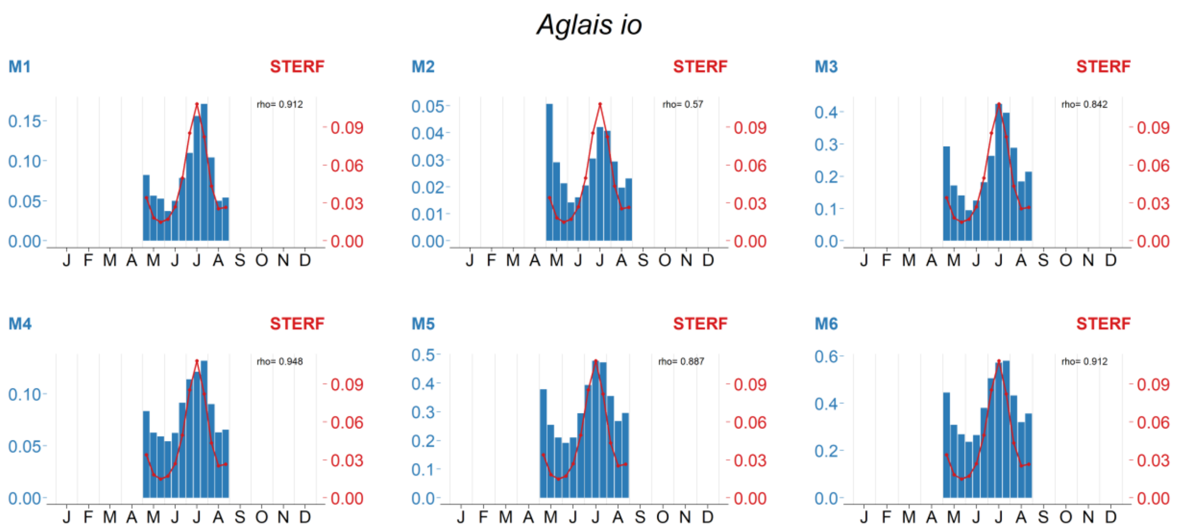
Plebejus argyrognomon





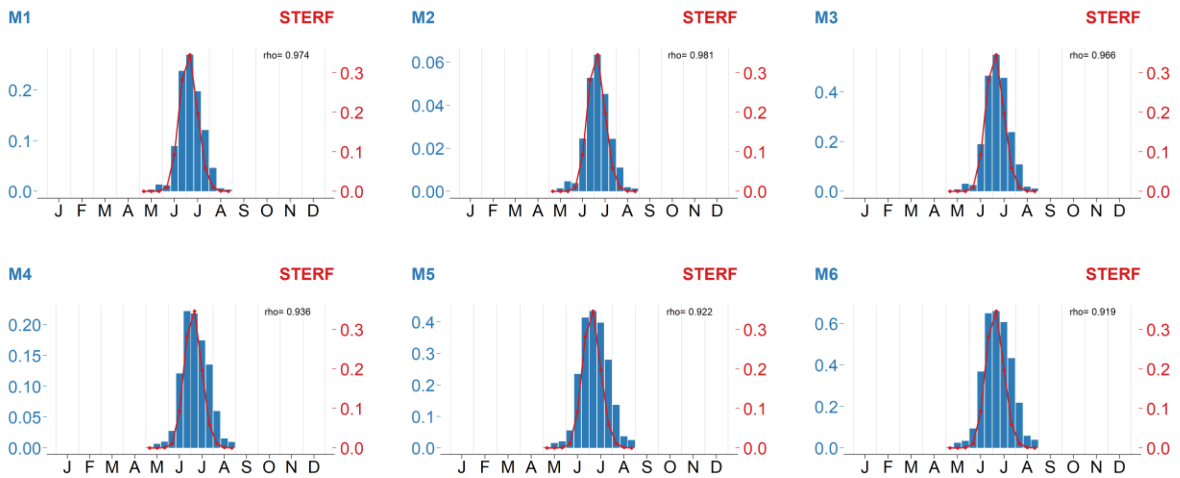


Nymphalidae:

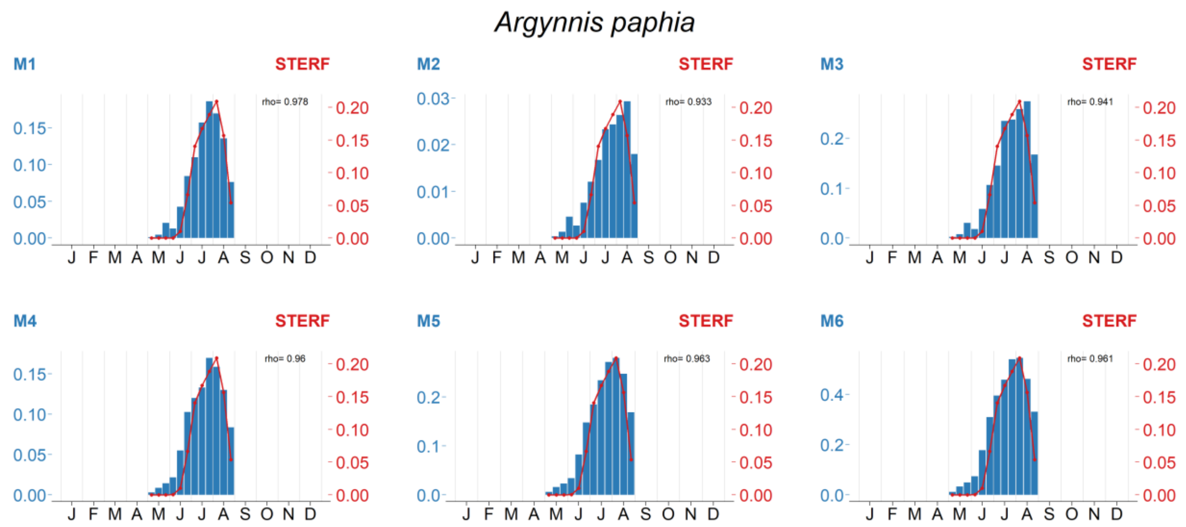


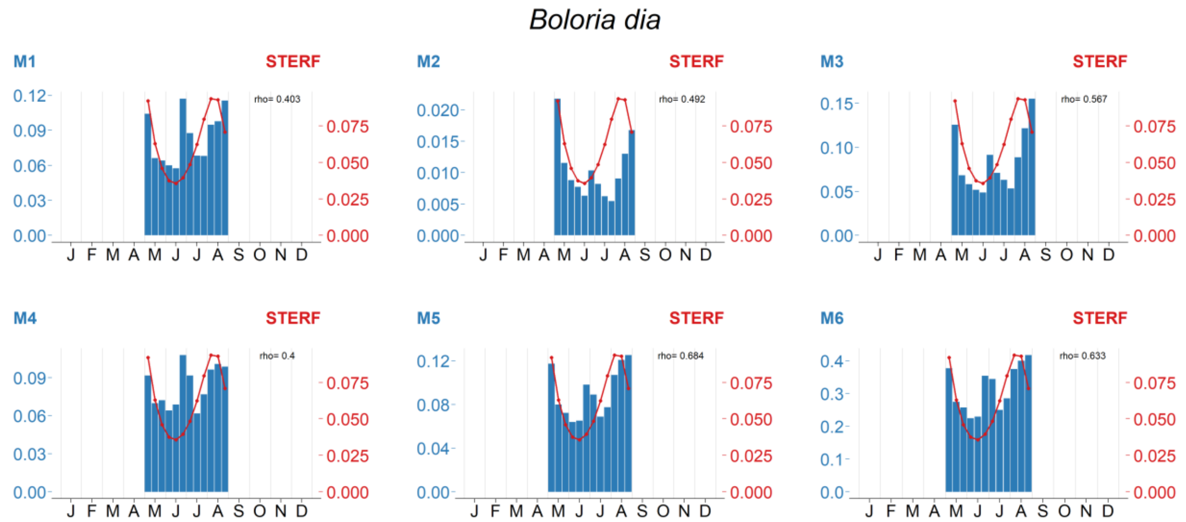
Aglais urticae

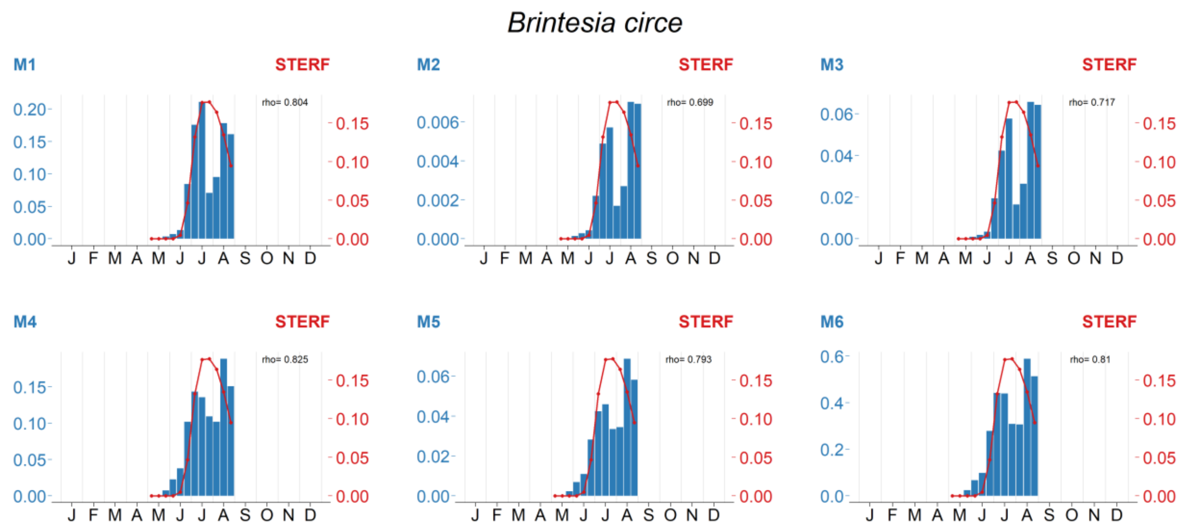
Aphantopus hyperantus

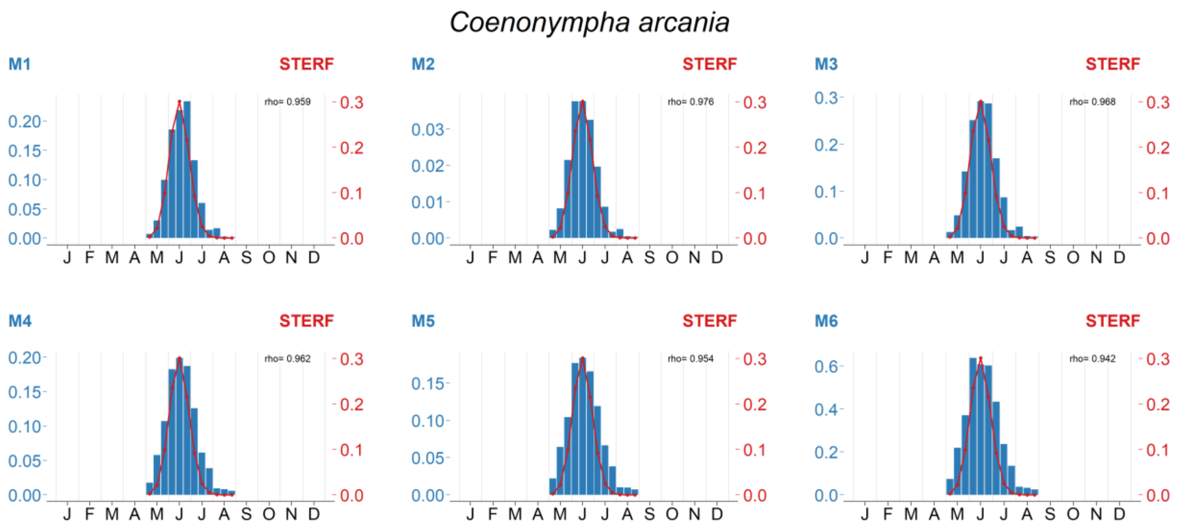


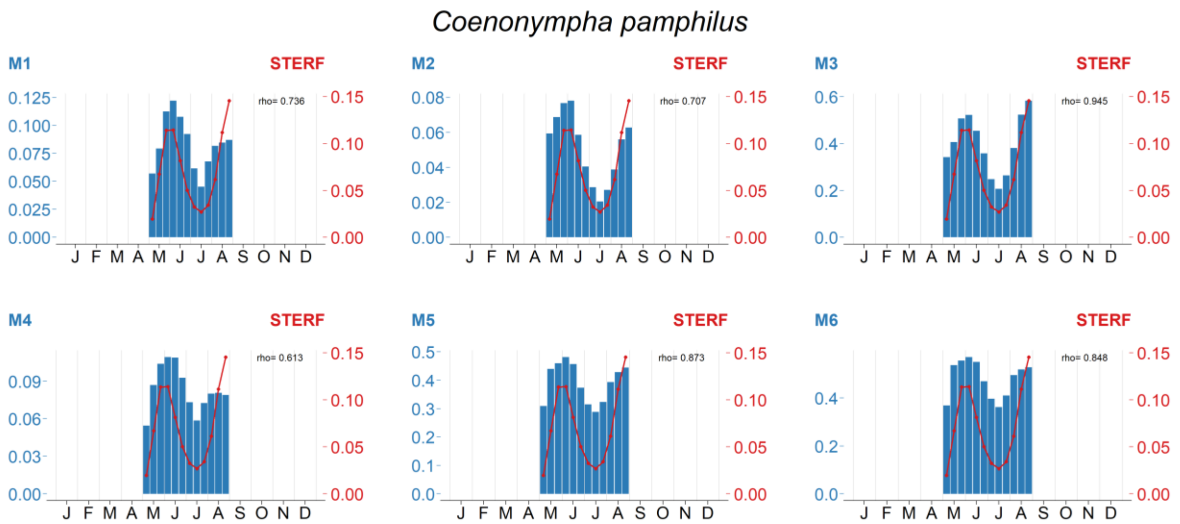
Araschnia levana

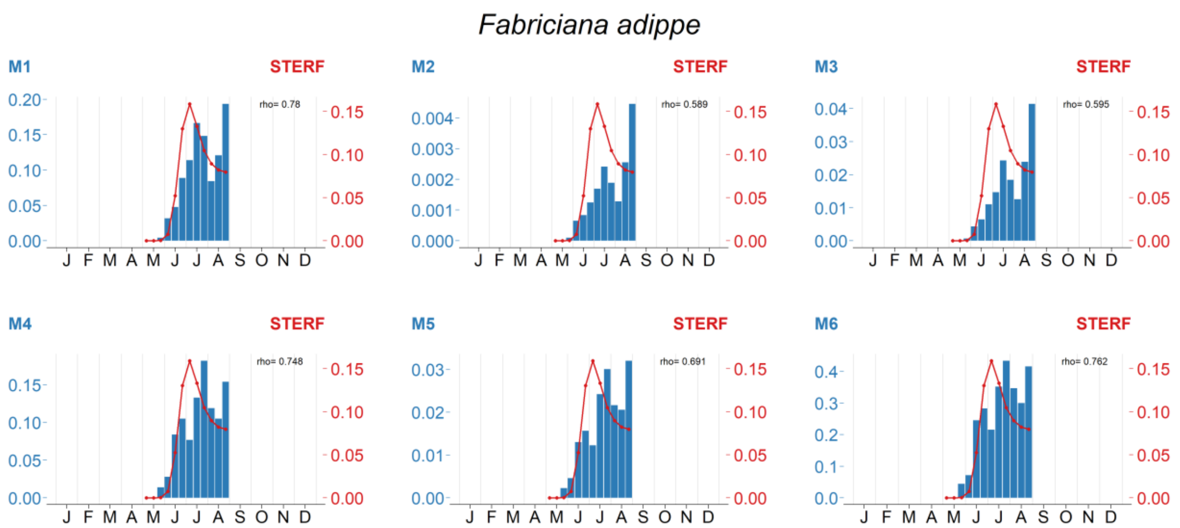

Argynnis paphia
M1 STERF rho= 0.978
M2 STERF rho= 0.933
M3 STERF rho= 0.941
M4 STERF rho= 0.96
M5 STERF rho= 0.963
M6 STERF rho= 0.961
J F M A M J J A S O N D

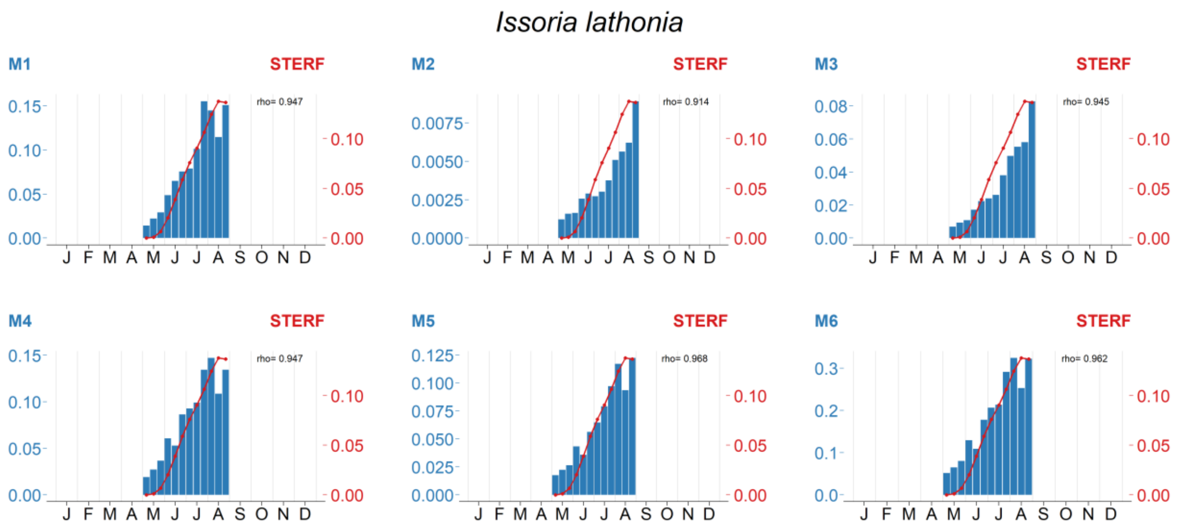

Boloria dia
M1 STERF rho= 0.403
M2 STERF rho= 0.492
M3 STERF rho= 0.567
M4 STERF rho= 0.4
M5 STERF rho= 0.684
M6 STERF rho= 0.633
J F M A M J J A S O N D

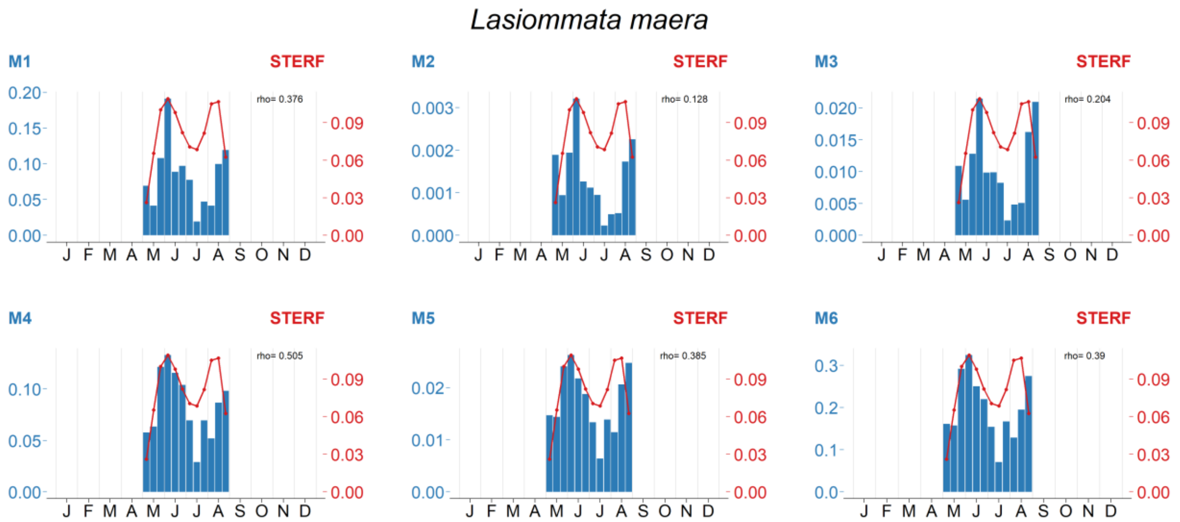

Brintesia circe
M1 STERF rho= 0.804
M2 STERF rho= 0.699
M3 STERF rho= 0.717
M4 STERF rho= 0.825
M5 STERF rho= 0.793
M6 STERF rho= 0.81
J F M A M J J A S O N D

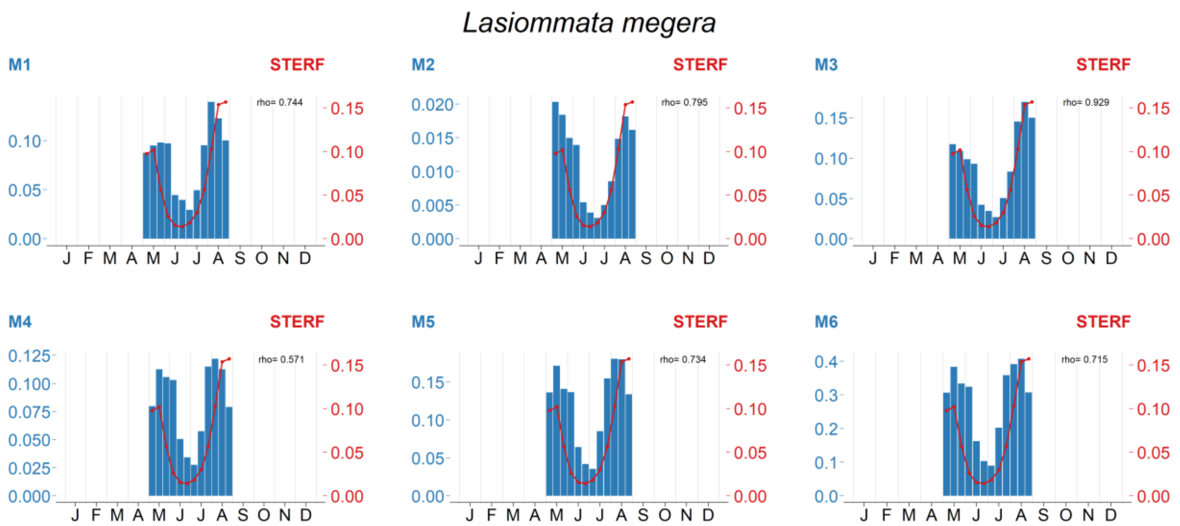

Coenonympha arcania
M1
STERF
rho= 0.959
M2
STERF
rho= 0.976
M3
STERF
rho= 0.968
M4
STERF
rho= 0.962
M5
STERF
rho= 0.954
M6
STERF
rho= 0.942
J F M A M J J A S O N D


Coenonympha pamphilus
M1
STERF
rho= 0.736
M2
STERF
rho= 0.707
M3
STERF
rho= 0.945
M4
STERF
rho= 0.613
M5
STERF
rho= 0.873
M6
STERF
rho= 0.848
J F M A M J J A S O N D


Fabriciana adippe
M1
STERF
rho= 0.78
M2
STERF
rho= 0.589
M3
STERF
rho= 0.595
M4
STERF
rho= 0.748
M5
STERF
rho= 0.691
M6
STERF
rho= 0.762
J F M A M J J A S O N D


Issoria lathonia
M1 STERF rho= 0.947
M2 STERF rho= 0.914
M3 STERF rho= 0.945
M4 STERF rho= 0.947
M5 STERF rho= 0.968
M6 STERF rho= 0.962
J F M A M J J A S O N D

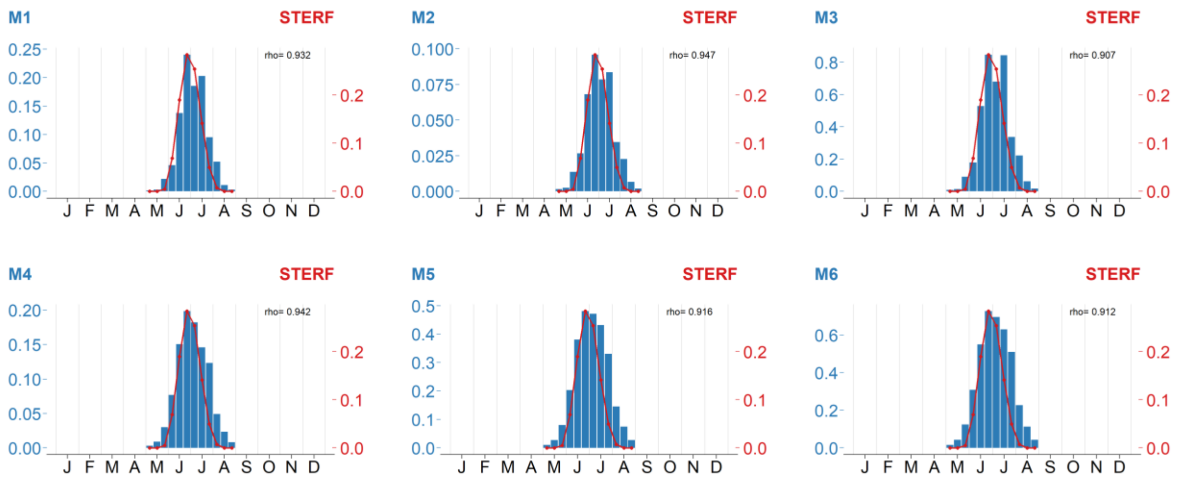

Lasiommata maera
M1 STERF rho= 0.376
M2 STERF rho= 0.128
M3 STERF rho= 0.204
M4 STERF rho= 0.505
M5 STERF rho= 0.385
M6 STERF rho= 0.39
J F M A M J J A S O N D


Lasiommata megera
M1 STERF rho= 0.744
M2 STERF rho= 0.795
M3 STERF rho= 0.929
M4 STERF rho= 0.571
M5 STERF rho= 0.734
M6 STERF rho= 0.715
J F M A M J J A S O N D

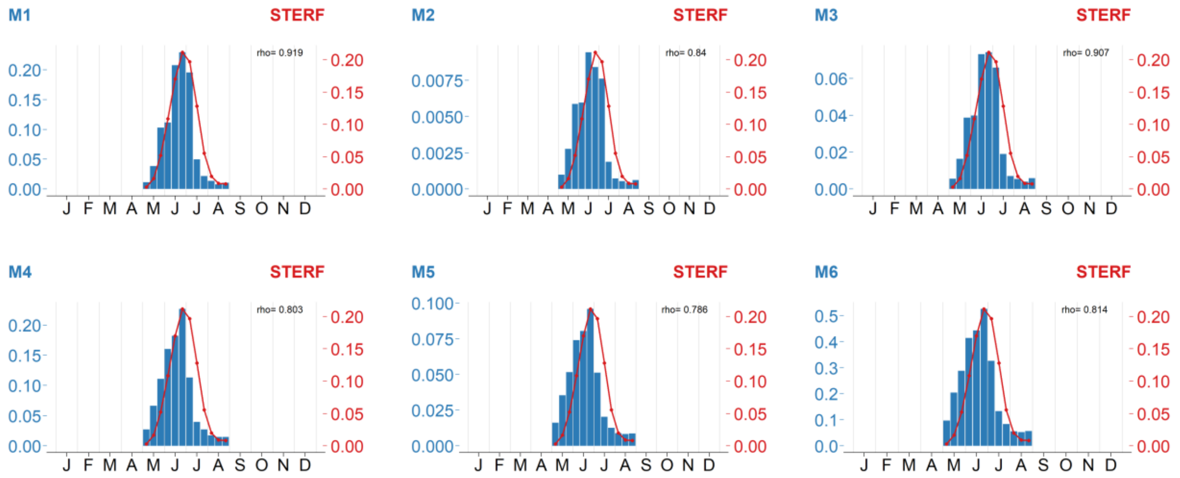
Limenitis camilla

Maniola jurtina

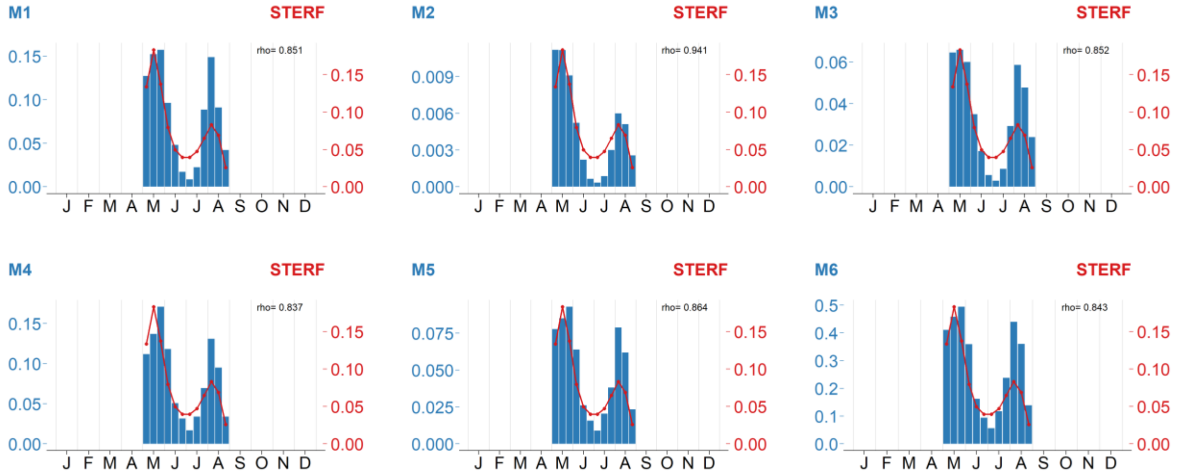
Melanargia galathea



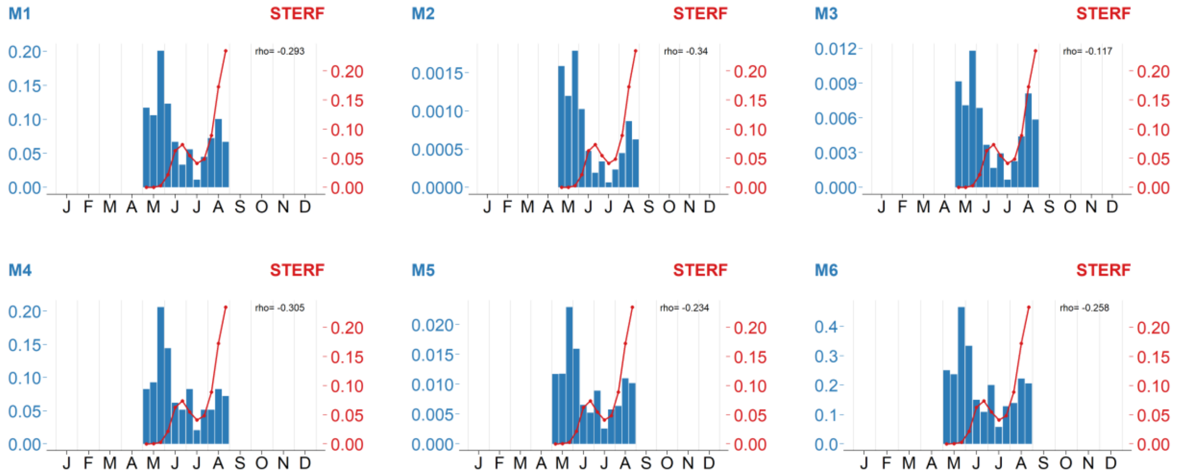
Melitaea athalia

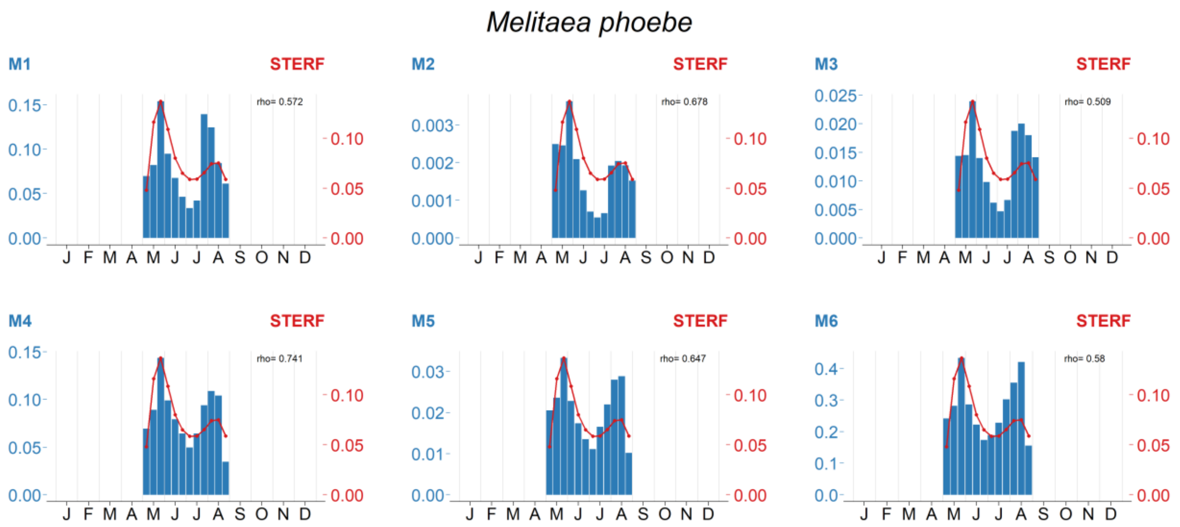


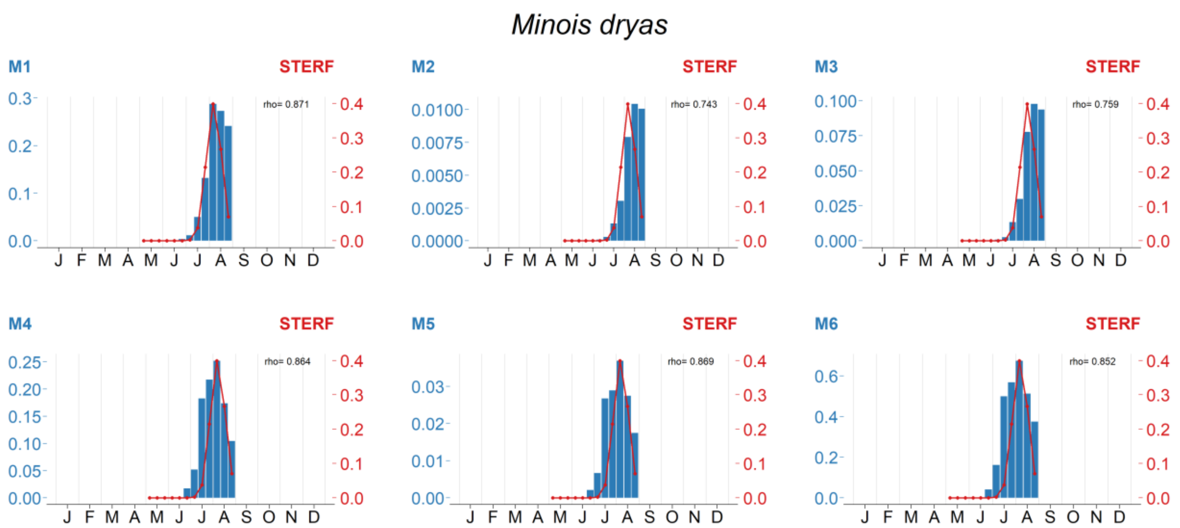
Melitaea cinxia

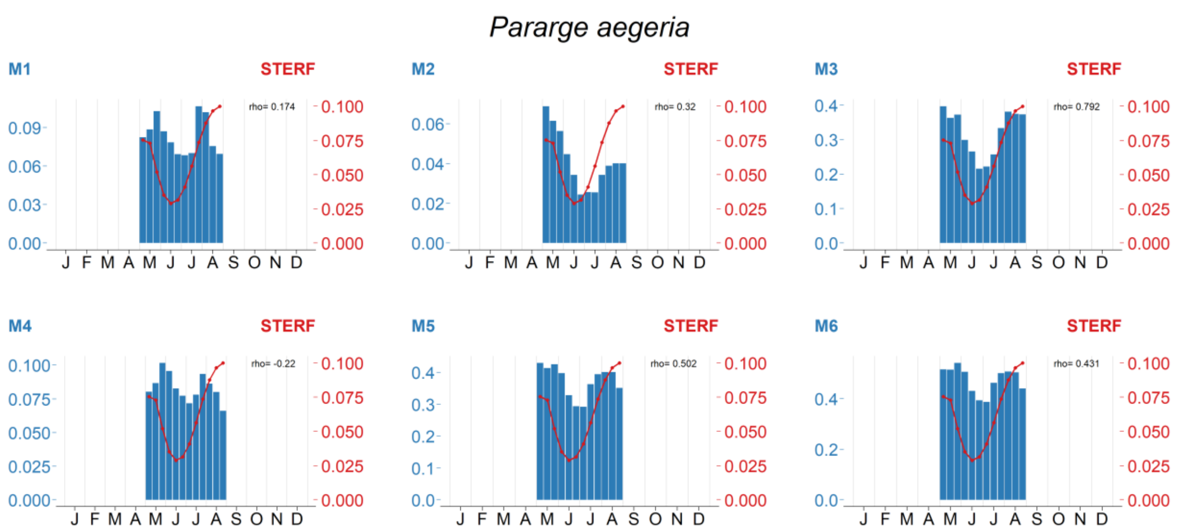


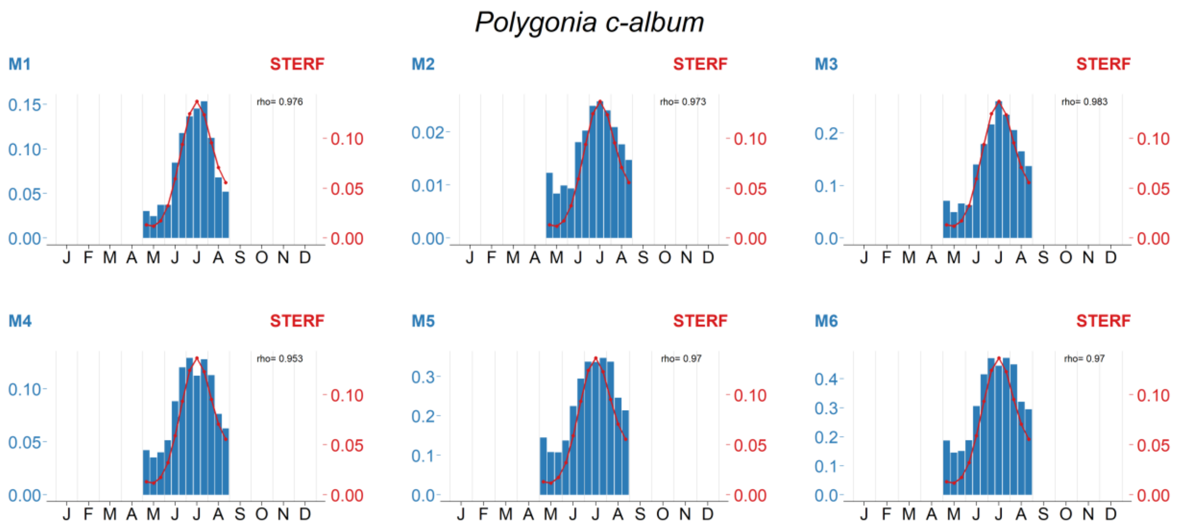
Melitaea parthenoides

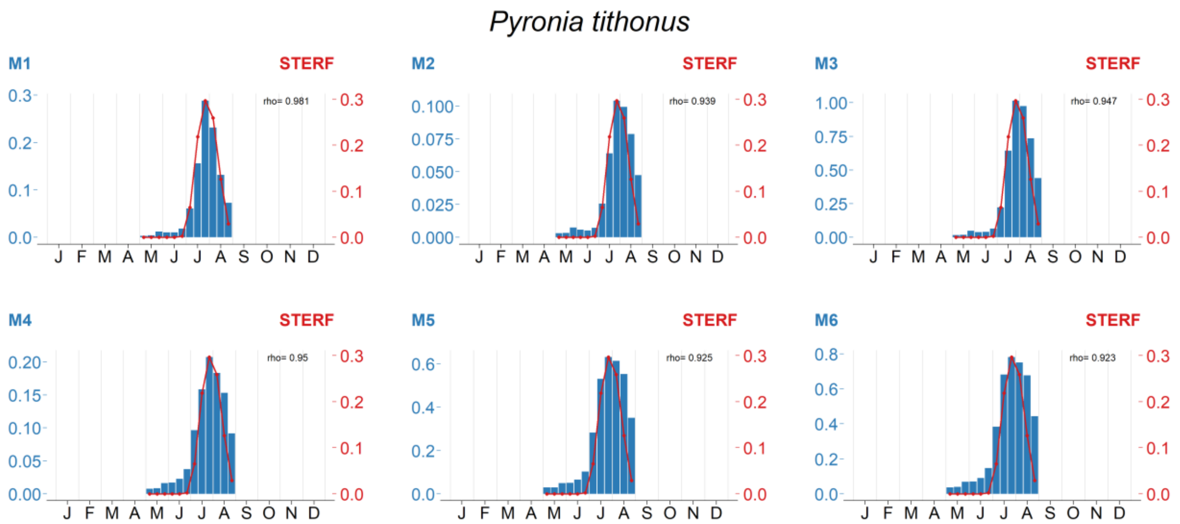


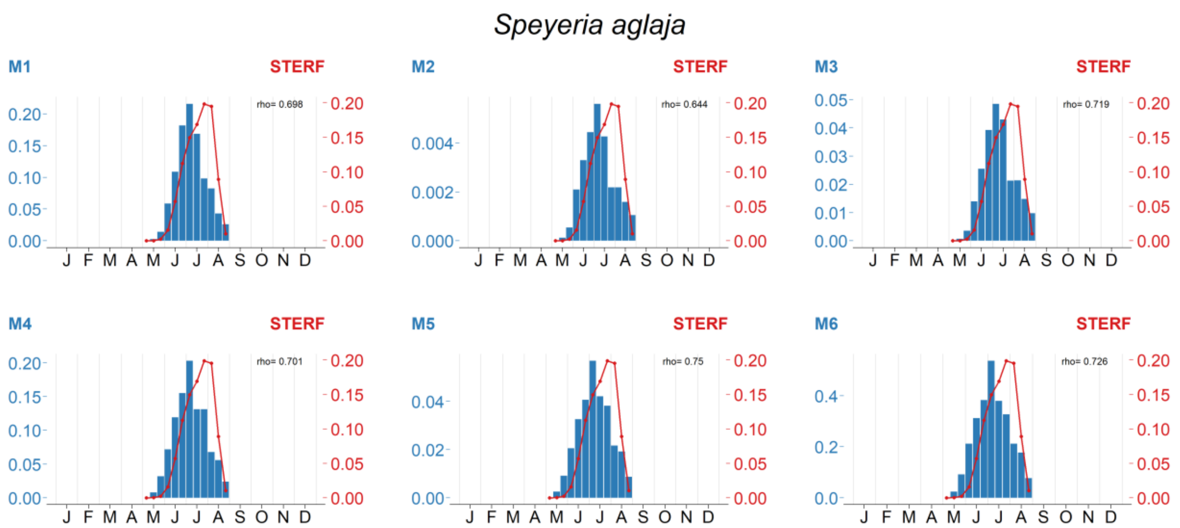

Melitaea phoebe
M1 STERF rho= 0.572
M2 STERF rho= 0.678
M3 STERF rho= 0.509
M4 STERF rho= 0.741
M5 STERF rho= 0.647
M6 STERF rho= 0.58
J F M A M J J A S O N D

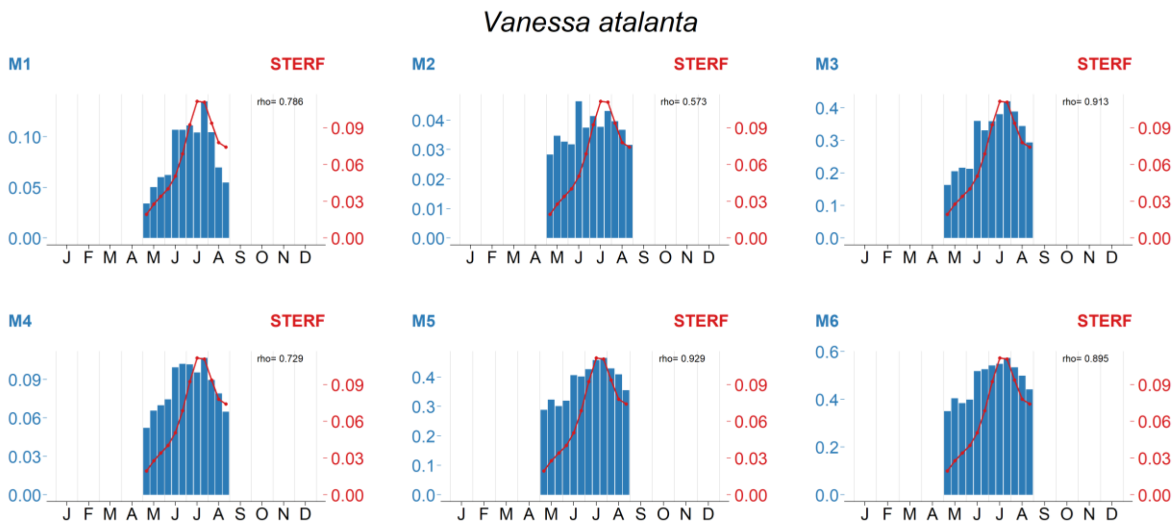

Minois dryas
M1 STERF rho= 0.871
M2 STERF rho= 0.743
M3 STERF rho= 0.759
M4 STERF rho= 0.864
M5 STERF rho= 0.869
M6 STERF rho= 0.852
J F M A M J J A S O N D

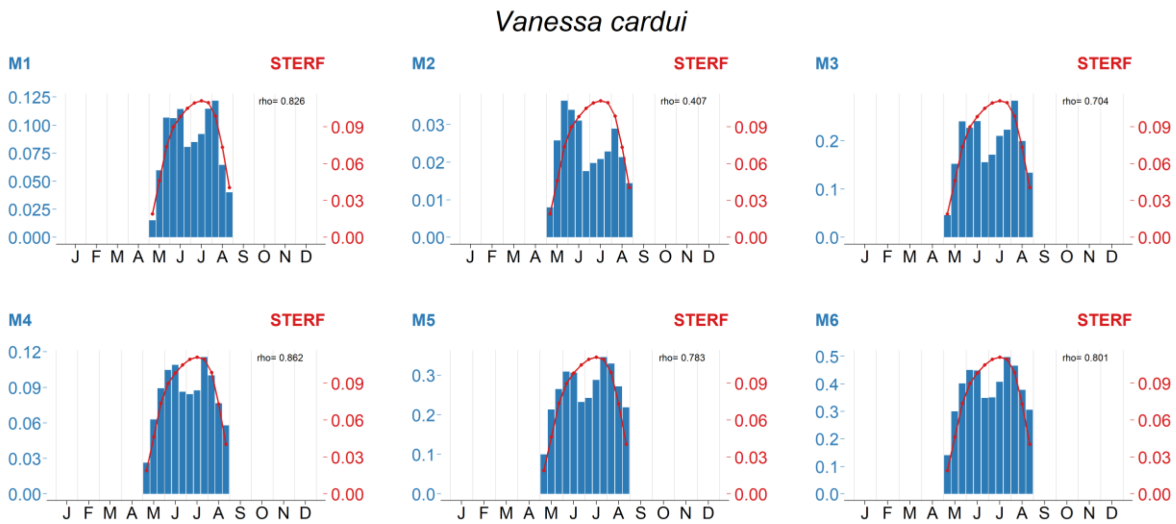

Pararge aegeria
M1 STERF rho= 0.174
M2 STERF rho= 0.32
M3 STERF rho= 0.792
M4 STERF rho= -0.22
M5 STERF rho= 0.502
M6 STERF rho= 0.431
J F M A M J J A S O N D


Polygonia c-album
M1 STERF rho= 0.976
M2 STERF rho= 0.973
M3 STERF rho= 0.983
M4 STERF rho= 0.953
M5 STERF rho= 0.97
M6 STERF rho= 0.97
J F M A M J J A S O N D

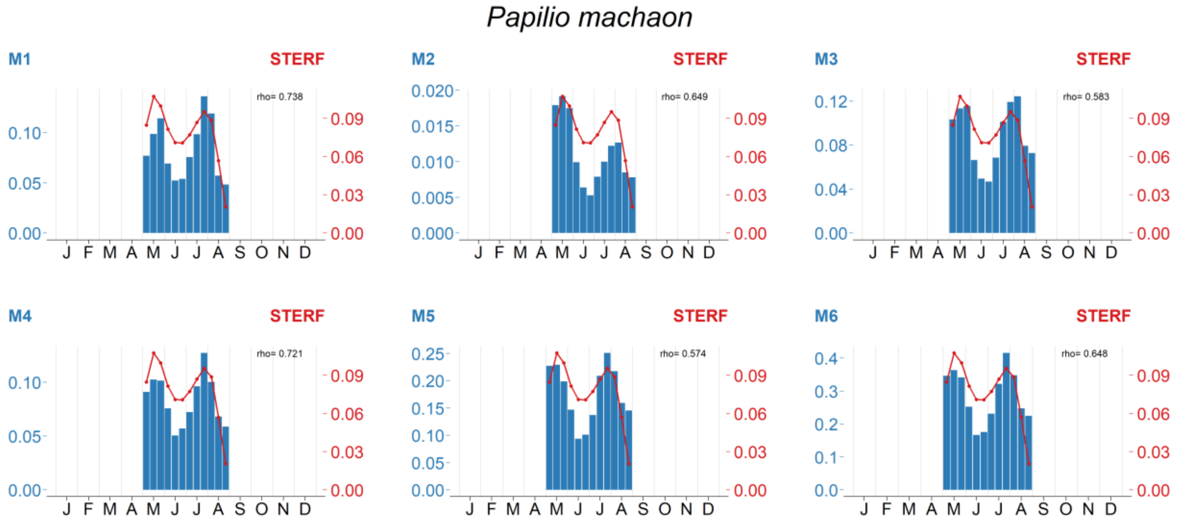

Pyronia tithonus
M1 STERF rho= 0.981
M2 STERF rho= 0.939
M3 STERF rho= 0.947
M4 STERF rho= 0.95
M5 STERF rho= 0.925
M6 STERF rho= 0.923
J F M A M J J A S O N D

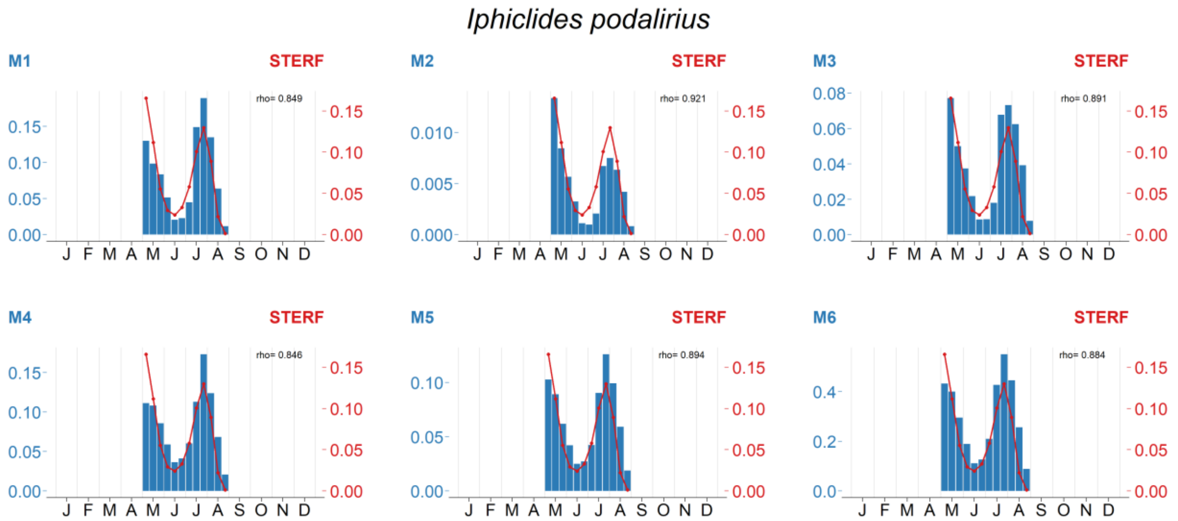

Speyeria aglaja
M1 STERF rho= 0.698
M2 STERF rho= 0.644
M3 STERF rho= 0.719
M4 STERF rho= 0.701
M5 STERF rho= 0.75
M6 STERF rho= 0.726
J F M A M J J A S O N D



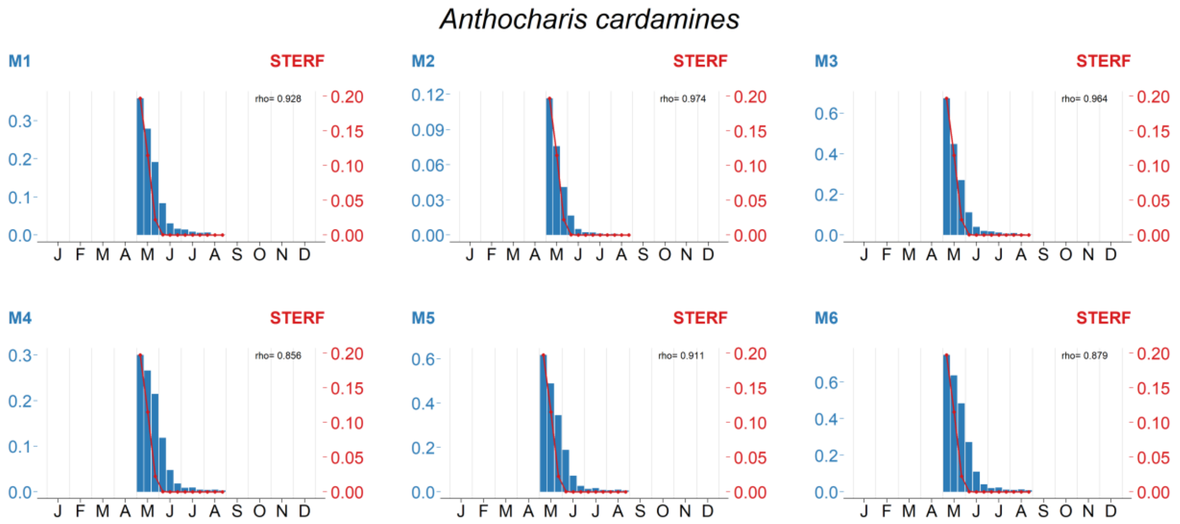


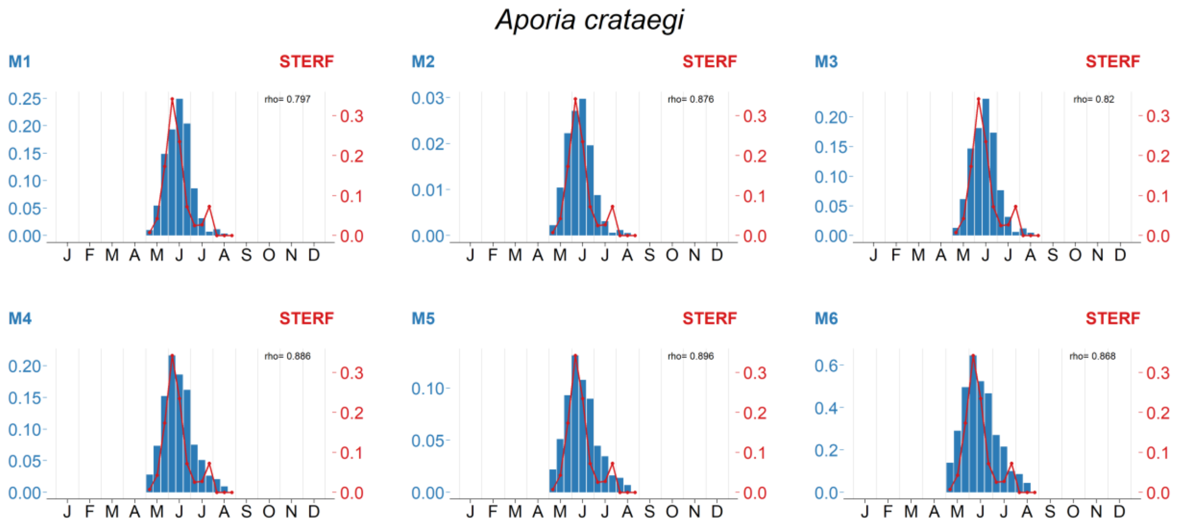
Papilionidae:

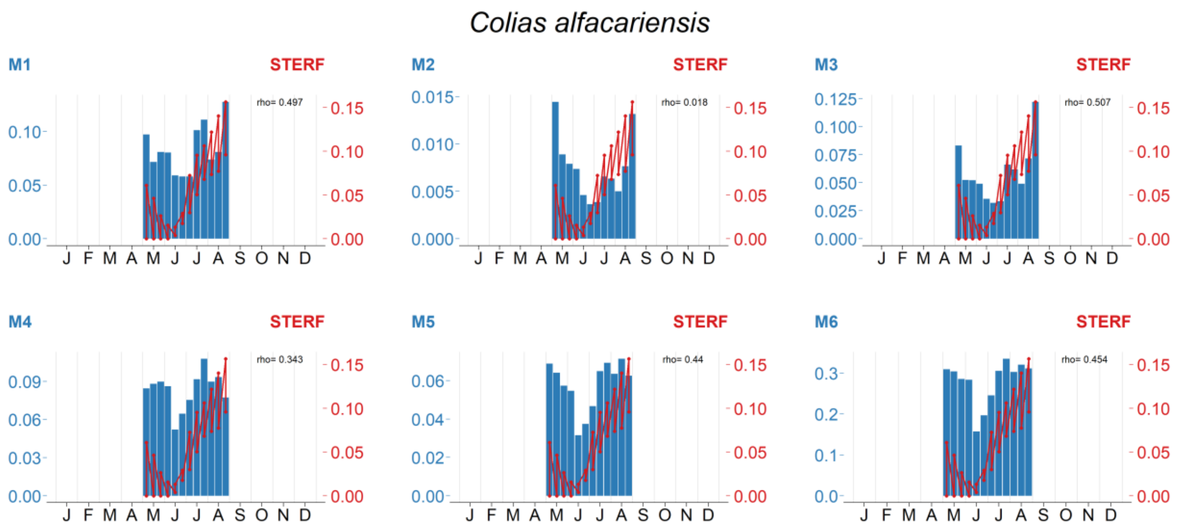

Papilio machaon
M1
STERF
rho= 0.738
M2
STERF
rho= 0.649
M3
STERF
rho= 0.583
M4
STERF
rho= 0.721
M5
STERF
rho= 0.574
M6
STERF
rho= 0.648
J F M A M J J A S O N D

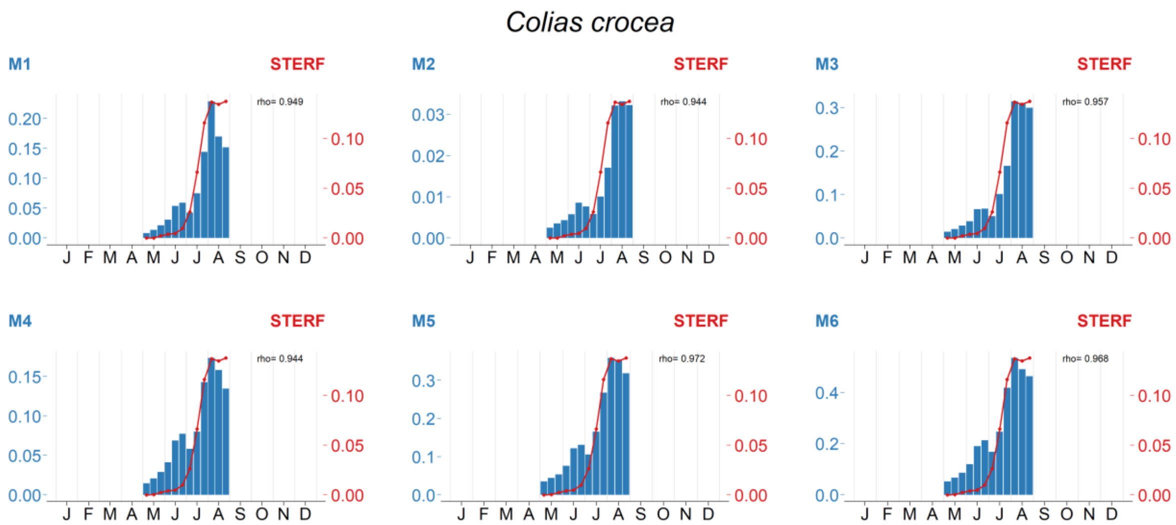

Iphiclides podalirius
M1
STERF
rho= 0.849
M2
STERF
rho= 0.921
M3
STERF
rho= 0.891
M4
STERF
rho= 0.846
M5
STERF
rho= 0.894
M6
STERF
rho= 0.884
J F M A M J J A S O N D

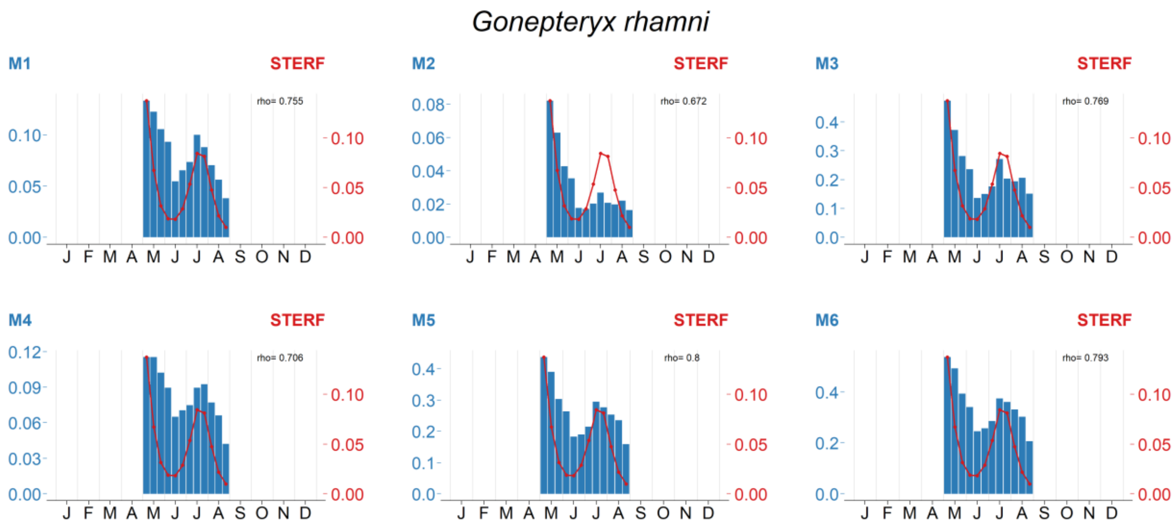
Pieridae:

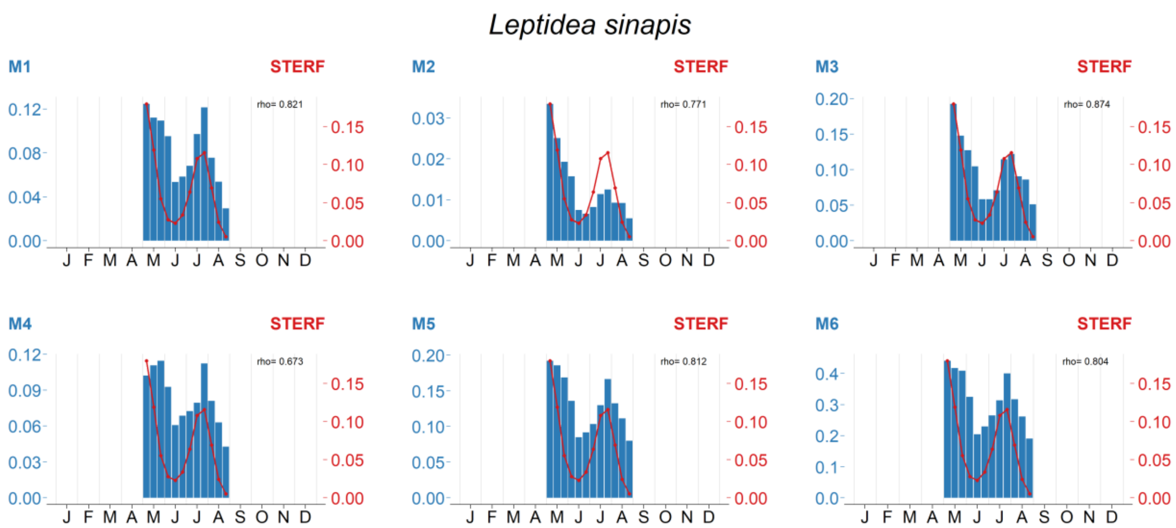

Anthocharis cardamines
M1
STERF
rho= 0.928
M2
STERF
rho= 0.974
M3
STERF
rho= 0.964
M4
STERF
rho= 0.856
M5
STERF
rho= 0.911
M6
STERF
rho= 0.879
J F M A M J J A S O N D

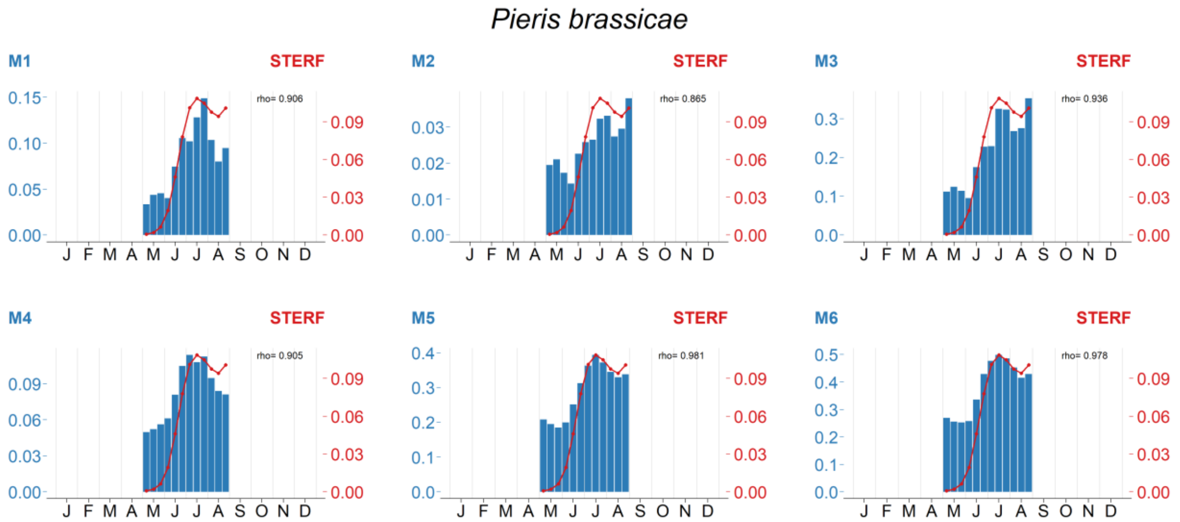

Aporia crataegi
M1
STERF
rho= 0.797
M2
STERF
rho= 0.876
M3
STERF
rho= 0.82
M4
STERF
rho= 0.886
M5
STERF
rho= 0.896
M6
STERF
rho= 0.868
J F M A M J J A S O N D

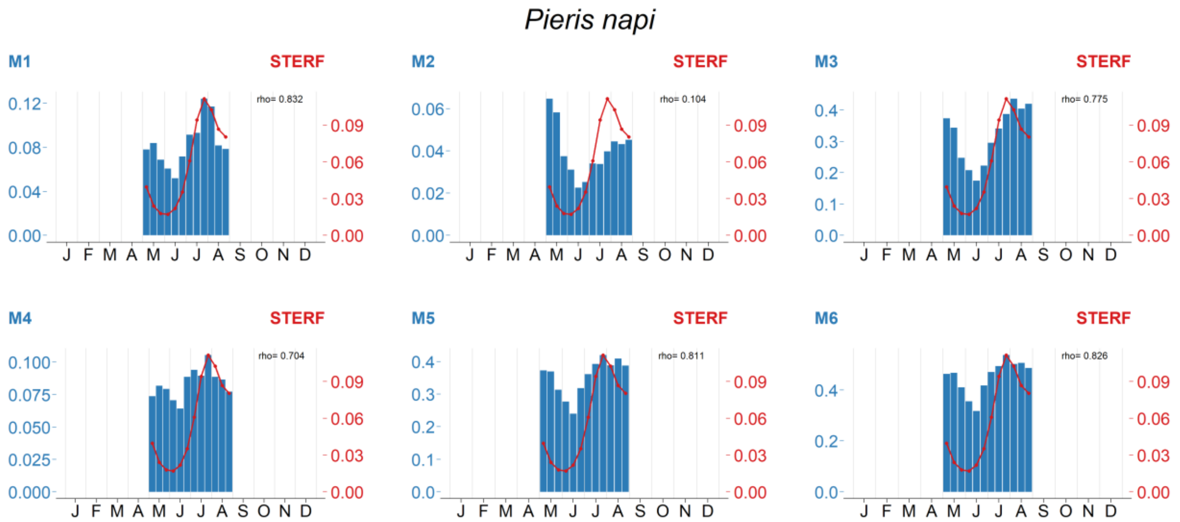

Colias alfacariensis
M1
STERF
rho= 0.497
M2
STERF
rho= 0.018
M3
STERF
rho= 0.507
M4
STERF
rho= 0.343
M5
STERF
rho= 0.44
M6
STERF
rho= 0.454
J F M A M J J A S O N D

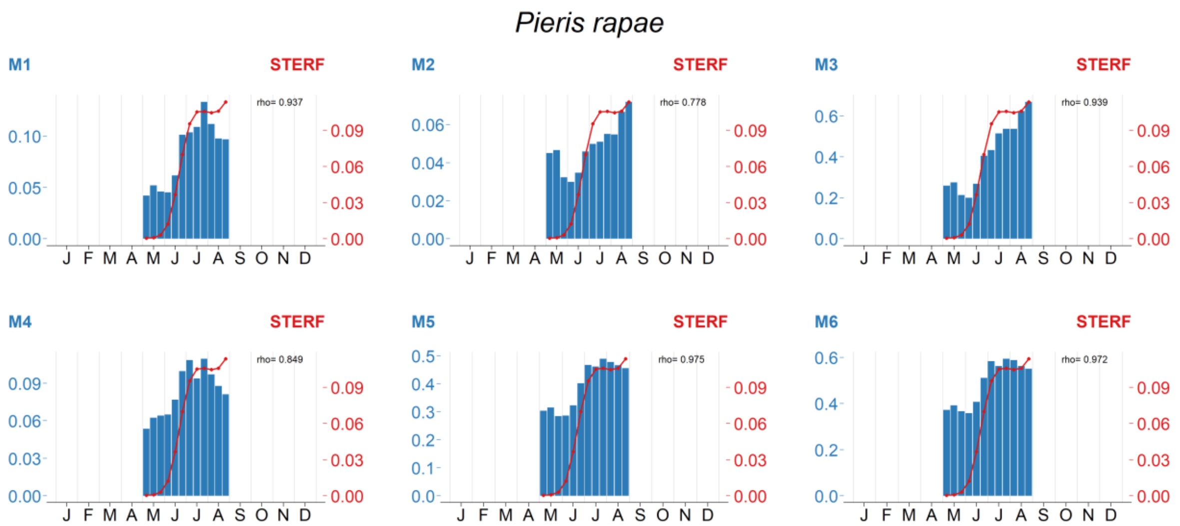

Colias crocea
M1 STERF rho= 0.949
M2 STERF rho= 0.944
M3 STERF rho= 0.957
M4 STERF rho= 0.944
M5 STERF rho= 0.972
M6 STERF rho= 0.968
J F M A M J J A S O N D


Gonepteryx rhamni
M1 STERF rho= 0.755
M2 STERF rho= 0.672
M3 STERF rho= 0.769
M4 STERF rho= 0.706
M5 STERF rho= 0.8
M6 STERF rho= 0.793
J F M A M J J A S O N D


Leptidea sinapis
M1 STERF rho= 0.821
M2 STERF rho= 0.771
M3 STERF rho= 0.874
M4 STERF rho= 0.673
M5 STERF rho= 0.812
M6 STERF rho= 0.804
J F M A M J J A S O N D


Pieris brassicae
M1
STERF
rho= 0.906
M2
STERF
rho= 0.865
M3
STERF
rho= 0.936
M4
STERF
rho= 0.905
M5
STERF
rho= 0.981
M6
STERF
rho= 0.978
J F M A M J J A S O N D


Pieris napi
M1
STERF
rho= 0.832
M2
STERF
rho= 0.104
M3
STERF
rho= 0.775
M4
STERF
rho= 0.704
M5
STERF
rho= 0.811
M6
STERF
rho= 0.826
J F M A M J J A S O N D


Pieris rapae
M1
STERF
rho= 0.937
M2
STERF
rho= 0.778
M3
STERF
rho= 0.939
M4
STERF
rho= 0.849
M5
STERF
rho= 0.975
M6
STERF
rho= 0.972
J F M A M J J A S O N D