



HAL
open science

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

► To cite this version:

Valentina Cima, Benoît Fontaine, Isabelle Witté, Pascal Dupont, Martin Jeanmougin, et al.. A test of six simple indices to display the phenology of butterflies using a large multi-source database. *Ecological Indicators*, 2020, 110, pp.105885. 10.1016/j.ecolind.2019.105885 . mnhn-03082080

HAL Id: mnhn-03082080

<https://mnhn.hal.science/mnhn-03082080v1>

Submitted on 5 Jan 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 A test of six simple indices to display the phenology 2 of butterflies using a large multi-source database

3 Valentina Cima¹, Benoît Fontaine^{1,2}, Isabelle Witté¹, Pascal Dupont¹, Martin Jeanmougin², Julien
4 Touroult¹

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

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

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

10 Abstract

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

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

29 **1 Introduction**

30 The rise of biological recording schemes including broad-scale citizen-science programs
31 has brought new possibilities to conservation and ecological research over the last decades,
32 producing large amounts of species occurrence data (Dickinson et al. 2012, Hochachka et al.
33 2012, Tulloch et al. 2013, August et al. 2015, Pocock et al. 2015). When gathered into
34 datasets that cover large temporal and spatial extents, these multi-source data (a combination
35 of opportunistic and systematic records) may help unveiling important aspects of biodiversity
36 state and changes (Dickinson et al. 2010, Dickinson et al. 2012, Hochachka et al. 2012, Isaac
37 and Pocock 2015, Powney and Isaac 2015), including species phenology. However, multi-
38 source data present different levels of standardisation (depending on the source of collection)
39 and are by nature noisy and unstructured. They suffer from several biases (Dickinson et al.
40 2010, Robertson et al. 2010, Isaac et al. 2014, Isaac and Pocock 2015) which primarily relate
41 to variation in recording intensity (Isaac et al. 2014). In fact, sampling effort may vary
42 throughout the year, between the years and among regions and this variability is usually
43 unknown in opportunistic datasets (Giraud et al. 2016).

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

56 Butterflies are relatively well known and well represented in citizen-science programs
57 and national inventories. For this reason, they have frequently been the subject of studies
58 based on monitoring and citizen-science data (see for example: Maes et al. 2012, Dennis et al.
59 2013, Schmucki et al. 2016). Here, we focus on the representation of butterfly phenology
60 patterns with multi-source opportunistic records while accounting for sampling effort. More
61 precisely, this study aims to explore the potential of a national biodiversity reference system,

62 the National Inventory of Natural Heritage of France (INPN - <https://inpn.mnhn.fr>), for
63 displaying the phenology of butterflies at a national scale. The goal is to develop a phenology
64 index with the following requirements: 1) the phenology is displayed using multi-source data,
65 including opportunistic observations; 2) the index yields overall patterns that are consistent
66 with standard knowledge on species phenology; 3) the index must be suitable for many
67 species and different type of seasonality and voltinism (refer to Wolda 1988 for a precise
68 classification of seasonality patterns); 4) the phenology charts will be presented to a general
69 audience on a web portal, so the index must be simple, easily interpretable and fast-rendering.
70 Ideally, the index should inform on species activity over the year or display a lack of
71 knowledge, inducing the community to collect more accurate occurrence data.

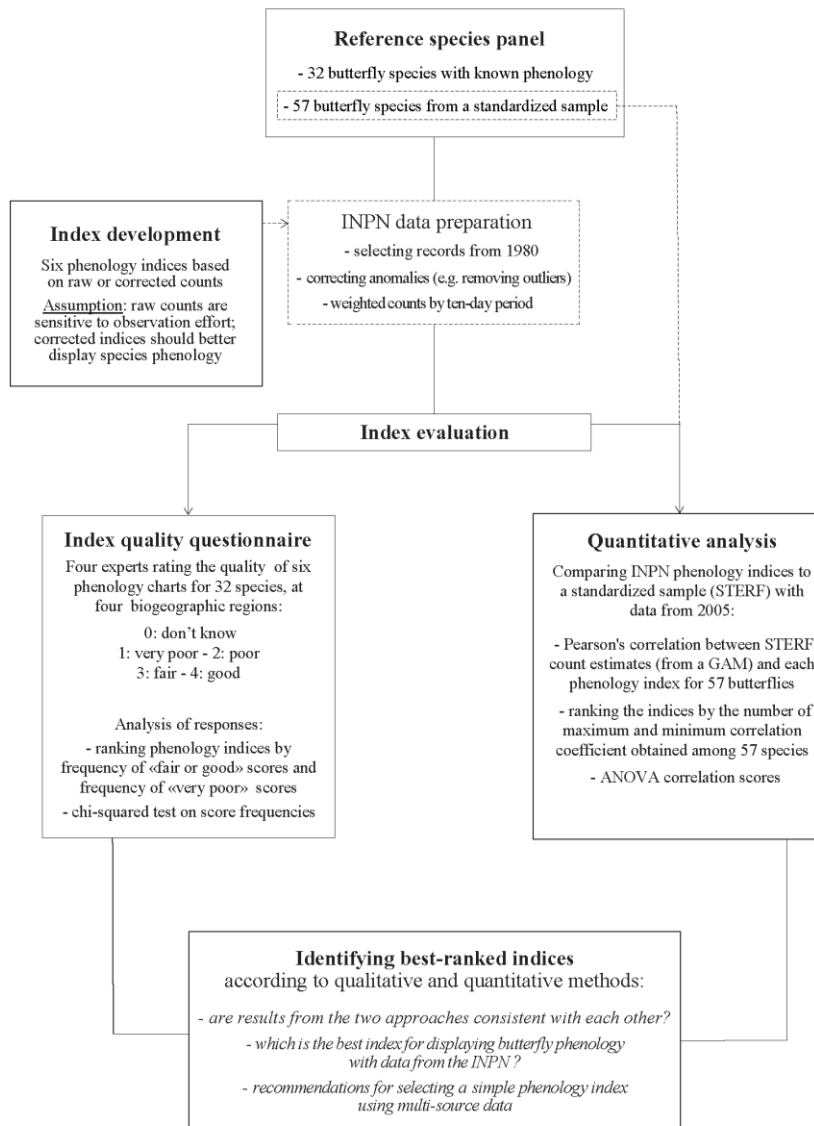
72 Some studies have highlighted the potential of opportunistic data to perform as well as
73 standardised data (see for example, van Strien et al. 2010, 2013). This is generally true when
74 opportunistic data are corrected for bias (van Strien et al. 2013, Isaac et al. 2014). Measures
75 that simply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

84 **2 Materials and methods**

85 **2.1 General analysis process**

86 The analysis started by selecting, with the help of a lepidopterologist (PD), a panel of
87 butterfly species that are well known and well represented in the dataset. Their records come
88 from different data sources (Appendix 1), were collected in several years and are assumed
89 homogeneous over France. In order to assess the versatility of the indices, the selected panel
90 had to be diverse in terms of phenology type, ecology and latitudinal range. Based on
91 literature review, a series of non-corrected (“raw”) and corrected indices were implemented
92 on multi-source butterfly data and evaluated through two different approaches: a qualitative
93 analysis, based on expert evaluation of the phenological patterns drawn by the indices, and a

94 quantitative analysis to seek for a match between the indices and the patterns based on
 95 monitoring data. The different steps of the analysis are developed in the next sections and
 96 summed up in Fig. 1. All analyses were carried out in R v.3.5.2 (R Core Development Team
 97 2018).



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

99 2.2 Data description

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

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

116 Effort correction required data from a background or target group (Ponder et al. 2001,
117 Phillips et al. 2009, Kéry et al. 2010, Ruete 2015). Although we focused on butterfly
118 phenology, we extracted all INPN records of diurnal lepidopterans (i.e. butterflies and diurnal
119 moths, hereinafter referred as the “group”, see Table 1) by assuming that survey methods and
120 collector specialties within this group are similar, hence data share similar bias (Ponder et al.
121 2001, Ruete 2015). The group includes seven families: Pieridae, Papilionidae, Nymphalidae,
122 Lycaenidae, Hesperidae, Riodinidae, Zygaenidae. The French taxonomic repository TaxRef
123 version 11.0 (Gargominy et al. 2017) was employed for taxonomic references. Assuming that
124 naturalists and collectors have similar knowledge and bias and use analogous survey methods
125 for the group of butterflies and diurnal moths, we defined a “field visit” as a 10x10 cell
126 surveyed by one observer on a date, regardless of the number of species he or she had
127 observed. In this manuscript, we refer to a species “quadrat” as a 10x10 km cell where the
128 species was recorded at least once in a given period. In the same way, one “group quadrat” is
129 a 10x10 km cell where at least one of the species of the group was recorded. The vocabulary
130 used for this study is summarised in Table 1.

131 2.3 Data preparation and index design

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

137 1980 onwards to be more uniform and representative of current overall phenological patterns
138 (e.g. number, position, sharpness of peaks). We restricted, therefore, the analysis to those
139 data. Moreover, 1980-today corresponds to the time span of knowledge of the experts who
140 took part in the survey.

141 In order to minimize errors, outlier data, such as records on January 1st and December
142 31th (probable by-default dates when the day or the duration of observation, for some reason,
143 is not known) were discarded *a priori*. Ultimately, data for the whole group of butterflies and
144 diurnal moths consisted in 772,307 records (Appendix 1).

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

157 **Equation 1**
$$\text{WEIGHTED NUMBER OF RECORDS}_k = \sum_{i=1}^n \text{prorata}_i$$

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

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

163 **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

164

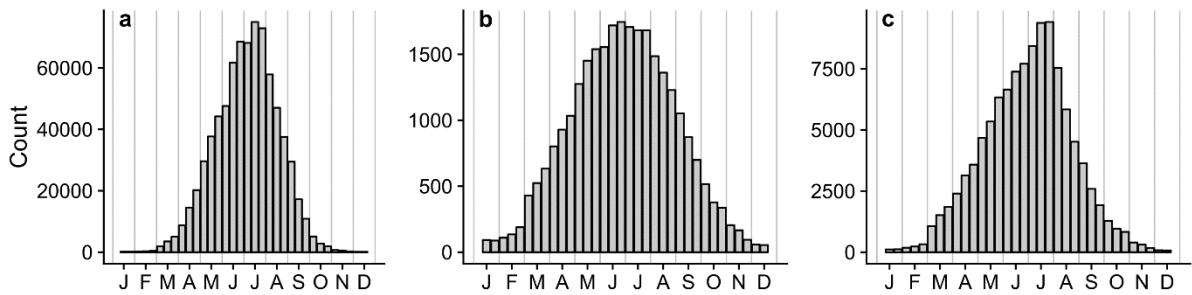
165 The number of records of all species of the group (“group records”), the number of
166 group quadrats and the number of field visits (section 2.2; Fig. 2) can be employed as proxy
167 for sampling effort (Lobo 2008, Phillips et al. 2009, Ruete 2015) and used for normalising
168 raw counts (see below).

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

178 **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	$\sum_{i=1}^q prorata_i / [\sum_{i=1}^g prorata_i]_{skd}$	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.

179



180

181 **Figure 2.** Seasonal distribution by ten-day period of three proxies for observation effort in diurnal lepidopterans (after data preparation). The
 182 x-labels indicate months. (a) group records: number of records of all species of the group; (b) group quadrats: number of quadrats where
 183 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.
 184 These values were used for normalising counts and building, respectively, index M2, M5 and M3.

185 **2.4 Species selection and phenology charts for index quality questionnaire: expert**
 186 **analysis**

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

191 For species whose phenology is likely to vary with latitude or altitude, a biogeographic
 192 approach was employed. Data were analysed at four separate biogeographic regions, an
 193 aggregated version of the environmental zones defined by Metzger et al. (2005): Atlantic,
 194 Continental and Pannonian (ATCONP), Lusitanian (LUS), Alpine and Mediterranean
 195 Mountains (ALMM), and Mediterranean (MD). For such species, we considered only
 196 biogeographic regions where at least 36 total records were available since 1980. The threshold
 197 was defined assuming that 36, an average of three records per month, was the minimum
 198 sample size for displaying patterns based on pooled counts. This threshold was also a trade-
 199 off between avoiding unreliable, insufficient data and maximising information (i.e. keeping
 200 regions where some species are rarer). The indices were computed on pre-cleaned data (see
 201 previous section), yielding six different phenological patterns that were submitted in the form
 202 of bar plots to four experts. The bar plots displayed frequencies by ten-day periods and, in
 203 some cases, by biogeographic regions (Fig. 3 and Appendix 2). For every species, the four
 204 experts were asked to rate six charts, one per index, on a scale of 1 to 4 (“very poor”, “poor”,
 205 “fair”, “good”), according to their quality in representing known seasonal activity (i.e. flight
 206 phenology, in the case of diurnal lepidopterans). In order to avoid conditioning and keep their
 207 judgment unbiased, experts were not aware of index design rules. In addition, the disposition

208 of the bar plots was randomized, so that they could not individually identify the indices
209 (Appendix 2). Every index was ranked by counting the number of times it had been rated as
210 “fair” or “good” (i.e. “the phenology chart is representative enough”) on the one hand, and as
211 “very poor” (i.e. “the phenology chart is not representative at all”) on the other hand. In
212 addition, a Chi-squared (χ^2) test was performed on score frequencies to seek any significant
213 difference between the phenology indices according to expert opinion.

214 2.5 Quantitative analysis: comparison with the STERF

215 A comparison with an independent sample was performed in order to quantitatively
216 analyse the pertinence of the indices and complete the analysis based on expert opinion. This
217 sample was provided by the French Butterfly Monitoring scheme (STERF), established from
218 2005 onwards by Vigie-Nature ([http://vigienature.mnhn.fr/page/suivi-temporel-des-](http://vigienature.mnhn.fr/page/suivi-temporel-des-rhopaloceres-de-france)
219 [rhopaloceres-de-france](http://vigienature.mnhn.fr/page/suivi-temporel-des-rhopaloceres-de-france)). The STERF provides systematic counts of adult butterflies, which
220 should mirror true species phenology. Each observer performs at least four field visits per
221 year: one visit per month from May to August (other visits during the year are possible). Each
222 survey site is associated with one observer, and is either chosen by the observer or randomly
223 selected. Butterflies are identified and individuals are counted along 5 to 15 transects selected
224 by the observers inside the site, making sure that the habitat within each transect is uniform.
225 Since data collection before May and after August had not been systematic, only STERF
226 records between May and August were kept. The analysis focused on the ATCONP
227 biogeographic region, where phenology is supposed to be uniform and STERF data are more
228 regular and abundant. As in the previous analysis (section 2.4), we fixed a threshold of 36
229 records in the ATCONP region.

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

$$237 E[c_{it}] = \mu_{it} = \exp[y_i + s(t, f)]$$

238 where weekly count c_{it} is a function of a year effect y and a penalized cubic regression
239 splines smoothing effect over time (week) t with f degree of freedom. GAMs were computed
240 thanks to the *mgcv* package (Wood 2017) in R 3.5.2. The respective GAM for each species

241 was finally used to predict ten-day period count estimates that could be compared to the
 242 phenology indices calculated with INPN data. The resulting flight curve were standardized to
 243 one ($\sum\mu_t = 1$). This curve could be calculated for 57 species. The others were excluded for
 244 lack of data and because the first GAM failed to yield annual phenologies.

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

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

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

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

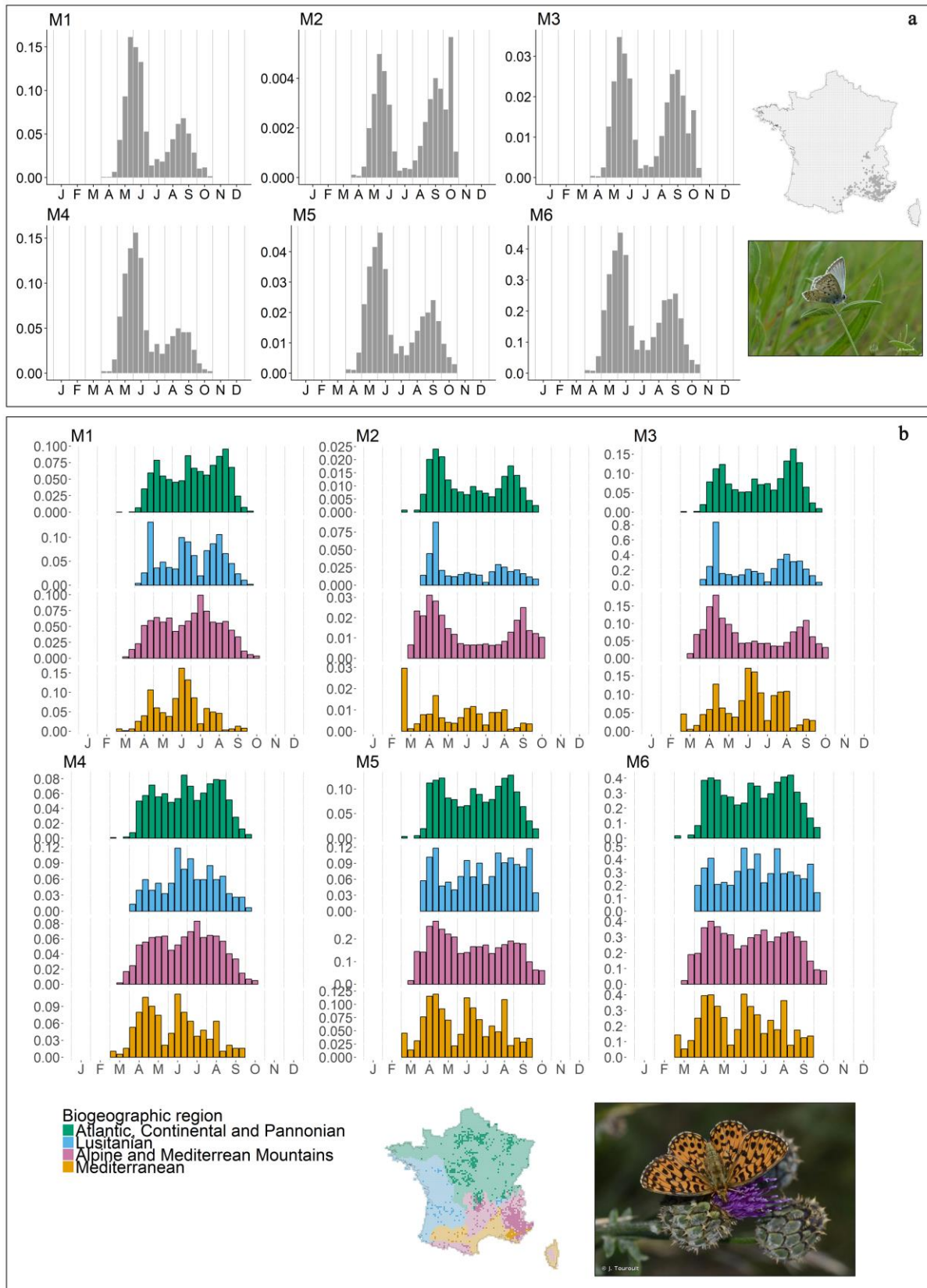
	M1	M2	M3	M4	M5	M6	Best correlated index	Least correlated index
species A	0.94	0.8	0.95	0.98	0.92	0.97	M4	M2
species B	0.77	0.5	0.82	0.93	0.75	0.77	M4	M2
species C	0.89	0.85	0.78	0.68	0.82	0.91	M6	M4
mean±SE	0.86±0.05	0.72±0.11	0.85±0.05	0.83±0.09	0.83±0.05	0.88±0.06		
Number of max rho	0	0	0	2	0	1		

Number of min rho	0	2	0	1	0	0		
----------------------	---	---	---	---	---	---	--	--

265 3 Results

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

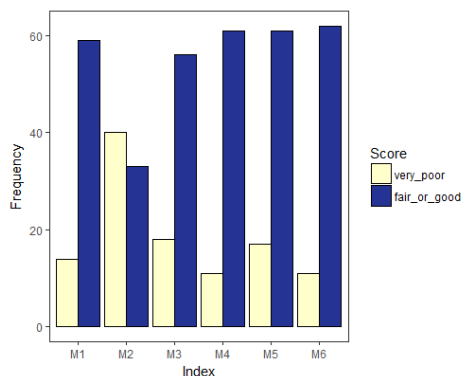
276 Fig. 5 shows an example of pairwise comparison between STERF count estimates and
 277 the six phenology indices for one of the species, while overall results of this comparison for
 278 57 butterfly species are illustrated in Fig. 6. All indices showed a fair correlation with the
 279 STERF, rho coefficient being 0.77 on average (Fig 6a). Index M1 and M5 were the best
 280 ranked, with the highest frequencies of maximum rho coefficient and few minimums (Fig 6b).
 281 Most of the times M2 resulted as the least correlated to STERF count estimates (Fig. 6b). M2
 282 rho distribution varied also considerably, with values below the first quartile that range from -
 283 0.34 and 0.56 (Fig. 6a). Nevertheless, the statistical analysis did not highlight any significant
 284 difference (ANOVA $F= 2.19$, $df=5$, $p\text{-value}=0.05$).



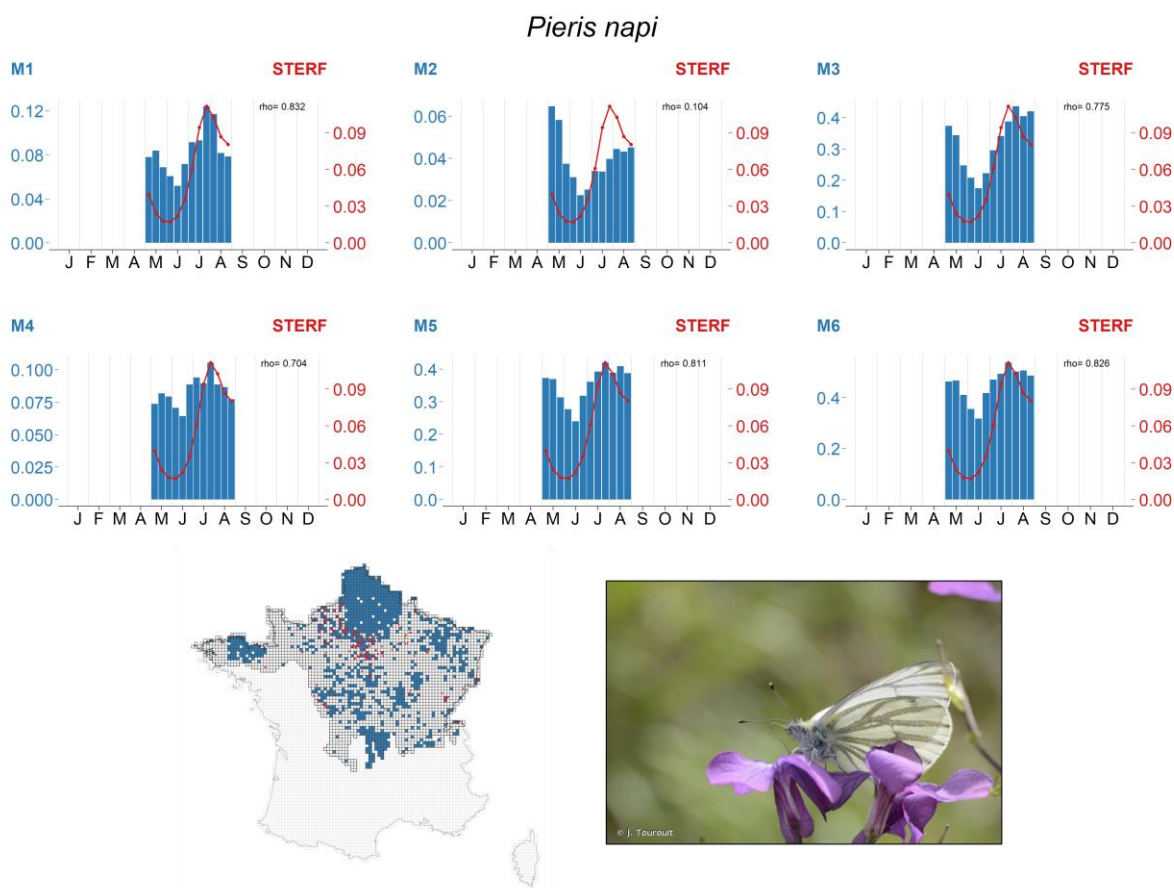
285

286 **Figure 3.** Phenology and distribution maps (based on INPN records after data preparation) of two butterfly species: (a) Provence chalk-hill
 287 blue (*Lysandra hispana*) and (b) Violet Fritillary (*Boloria dia*). Time unit is a ten-day period and the x-labels indicate months. Phenology is
 288 represented with six indices (see also Tab. 2): proportion of records per period (M1); ratio of records to the group (M2); number of records
 289 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

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

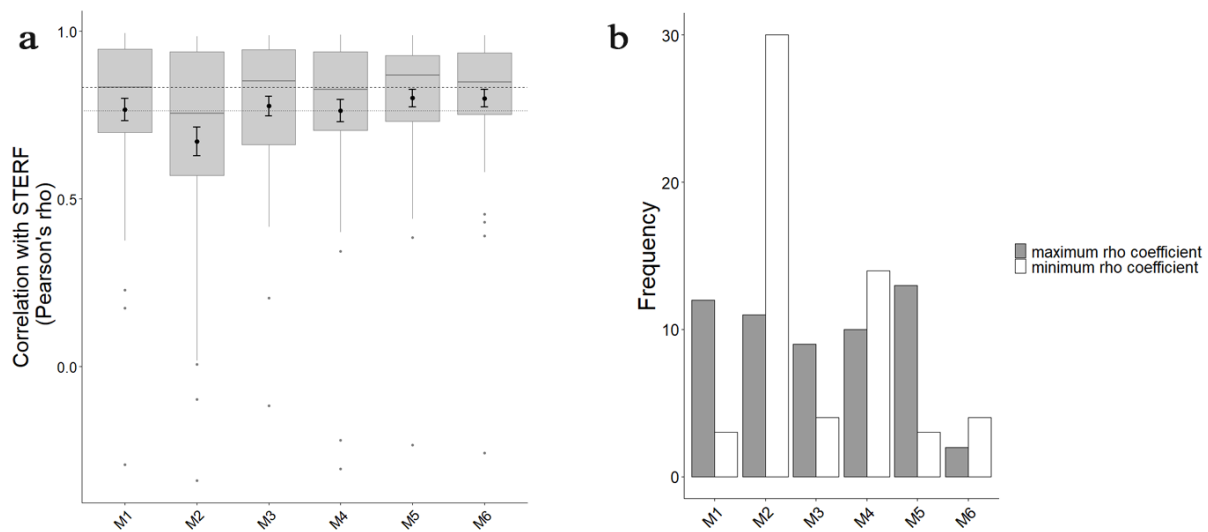


292
 293 **Figure 4** frequency of positive (“fair” or “good”) and negative (“very poor”) scores given by experts to six phenology indices. Index ranking
 294 according to the number of “fair or good” scores: M6, M4/M5, M1, M3, M2; index ranking according to the number of “very poor” scores:
 295 M2, M3, M5, M1, M4/M6. M2 is the only index that obtained more negative than positive scores. With M2: $\chi^2 = 43.73$, $df = 5$, $p\text{-value} <$
 296 0.001 ; excluding M2: $\chi^2 = 3.1$, $df = 4$, $p\text{-value} = 0.5$.



297
 298 **Figure 5.** Comparison between the STERF and six phenology indices computed on INPN data by ten-day period for one of the species, the
 299 green-veined white (*Pieris napi*). A Pearson's correlation coefficient (rho) was calculated between every index (blue bar plots) and STERF
 300 count estimates from a GAM (red line). Indices and STERF count estimates were calculated with data collected from 2005, and displayed

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



303
 304 **Figure 6** Comparison between the STERF and six phenology indices computed on INPN data by ten-day periods for 57 butterfly species (a)
 305 distributions and mean \pm se of rho coefficients by index. Dashed and dotted lines are respectively the overall median=0.83 and the overall
 306 mean=0.77 of all correlations. All phenology indices are equally correlated to STERF count estimates (ANOVA $F=2.19$, $df=5$, p -
 307 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.

308 4 Discussion

309 4.1 *A priori* properties of the indices

310 Non-corrected indices are easy to understand but they may be subject to several biases when
 311 using data that originate from multiple sources and sampling techniques. Corrected indices
 312 should approximate the phenology of activity (e.g. flight period for butterflies) with greater
 313 precision than raw frequencies, which are more likely to reflect patterns of recording
 314 intensity. No robust but simple metrics based on opportunistic, unstructured data that fulfilled
 315 the required conditions was found in recent scientific literature. Dennis et al. (2013) and
 316 Schmucki et al. (2016) have produced smooth and readable seasonal patterns with GAMs.
 317 With a similar approach, we applied GAMs on butterfly monitoring data for outlining
 318 phenology patterns to compare with patterns from INPN data. However, GAMs are not
 319 suitable for multi-source opportunistic data, because they do not account for uneven recording
 320 intensity (Rothery and Roy 2001). Other approaches, such as correction for sampling effort or
 321 imperfect detection in occupancy models, could inspire research for more appropriate
 322 measures (references in Table 2). M1 (proportion of records per period; Tab.2) is the most
 323 simple and intuitive, but may reflect effort variability over time and over space. M4
 324 (proportion of quadrats per period; Tab.2), which is also easily interpretable, may correct for

325 some errors (duplicates, oversampled quadrats), but still reflects temporal and spatial bias.
326 Conversely, M2 (ratio of records to the group; Tab.2) should correct for uneven recording
327 intensity. However, this index seems to be particularly sensitive to “reporting bias” (van
328 Strien et al. 2013). The number of group records increases sharply from near zero in winter
329 months to near 80000 in summer months (Fig 2a), causing over-correction and making peaks
330 shift from the centre to the edges of the distribution. This is particularly true for certain
331 species, whose adults can emerge outside the habitual flight period, for instance on sunny
332 winter days (such as *Aglais io* and *Aglais urticae*, Lafranchis and Geniez 2000; Lafranchis et
333 al. 2015; phenological patterns of *Aglais io* are shown in Appendix 4). Owing to the unusual
334 event, the species is almost the only one reported at such dates, generating artefact peaks in
335 the phenological pattern (i.e. the ratio “species records/group records” is close to 1). The same
336 may occur to species that are not easy to observe in general during the flight season (i.e. the
337 ratio decreases instead of displaying a peak) and to those that are paradoxically under-
338 reported because considered too common (Dickinson et al. 2010, van Strien et al. 2013). The
339 proxy for observation effort used for building M3 (number of records per field visit; Tab. 2;
340 Fig. 2c) should also correct for uneven recording intensity. This index, as the following ones,
341 is less intuitive and may not be precise, since observer names are not always well
342 standardized in the INPN. Such as M2 and M3, M5 (ratio of quadrats to the group; Tab. 2) is
343 based on a proxy for observation effort, the number of group quadrats per ten-day period (Fig.
344 2b). As shown in Figure 2b, compared to the other proxy distributions, the number of group
345 quadrats per ten-day period grows and shrinks slower between winter and summer periods
346 and the shape of the distribution is wider. This should limit reporting bias and allow for a
347 better correction of temporal bias, but it does not account for spatial effects (for example,
348 observer activity in coastal and mountain areas is higher during summer than during spring,
349 due to holiday habits of naturalists). Moreover, the proxies for observation effort (Fig 2)
350 include those records or quadrats where the target species was never seen and might be
351 actually absent. Managing data by biogeographic zone may help reducing spatial bias and
352 over-correction. Additionally, quadrats should be restricted to those included in the known
353 distribution of the species. This approach was applied to M6 (ratio of quadrats to the group
354 within the species known distribution; Tab. 2), which was expected to perform better, since it
355 deals with both temporal and spatial bias.

356 4.2 Which is the highest-ranked index according to the two approaches?

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

369 Other studies have compared opportunistic data with monitoring data (Dennis et al.
370 2017a; van Strien et al. 2013). It is possible that our comparisons lack power. The STERF
371 provides systematic counts of adult butterflies, which should provide phenological patterns
372 that are relatively close to reality. Nevertheless, STERF protocol and measures themselves
373 may not be free from bias, notably because counts are recorded on a monthly basis. However,
374 given the two approaches adopted, the number of species that have been taken into account,
375 as well as the amount of data and the variability of expert opinions, it is not unfounded to
376 believe that, there are no major differences between the tested indices. A possible explanation
377 lies in the large amount of data and sources available for butterflies, which may help attenuate
378 the bias and make raw frequencies of opportunistic observations converge towards overall
379 realistic phenological patterns. Our method may not be suitable for other lepidopterans or
380 other clades for which knowledge and data are much scarcer than for the butterflies analysed
381 here. If data are opportunistic and multi-source, we suggest selecting species with many
382 records, covered by several data sources. We fixed a threshold of 36 total records in the
383 studied area and during the entire period of study (in our case, all data from France or from a
384 biogeographic region, recorded since 1980), and 5 data sources, knowing that for most of the
385 species total records and sources were more numerous (Appendix 1). Ideally, the minimum
386 number of records and sources should be calculated and standardised. Further work is
387 required to re-define these thresholds on a statistical basis and give more recommendations
388 about the use of our indices for other groups than butterflies.

389 4.3 **Selecting a rigorous but simple index for large opportunistic data and a general** 390 **audience: which compromise for butterflies?**

391 Even though they are likely to reflect bias, in our case raw indices (M1 or M4) were not
392 less convincing than corrected ones for butterflies. Following the results and considerations
393 discussed above, and against expectations, it could finally be reasonable to consider the use of
394 raw frequencies. Besides, the choice should head to the most parsimonious method. The best
395 option would probably be to display phenology through raw indices (after data preparation
396 and weighting adjustment), alongside a visual representation of recording intensity (Fig.2b or
397 2c). Spatial bias may be attenuated by splitting data by biogeographic region. In this way,
398 both experts and general public should be able to understand the graphics, while keeping a
399 critical eye on them.

400 Some authors have appealed to occupancy models for studying phenology and
401 population trends with opportunistic data (see, for example, Kéry et al. 2010, van Strien et al.
402 2010, Strebel et al. 2014). In fact, by estimating detectability, occupancy models help to
403 correct for observation effort, detection and reporting bias (Kéry et al. 2010; Van Strien et al.
404 2013). Although we drew inspiration from occupancy models for bias correction approaches,
405 the aim of our study was not to assess occupancy of butterfly species. Nonetheless, we do not
406 exclude the possibility of using a modelling approach in the future, should sufficient data be
407 available, in order to get more unbiased estimates of phenology.

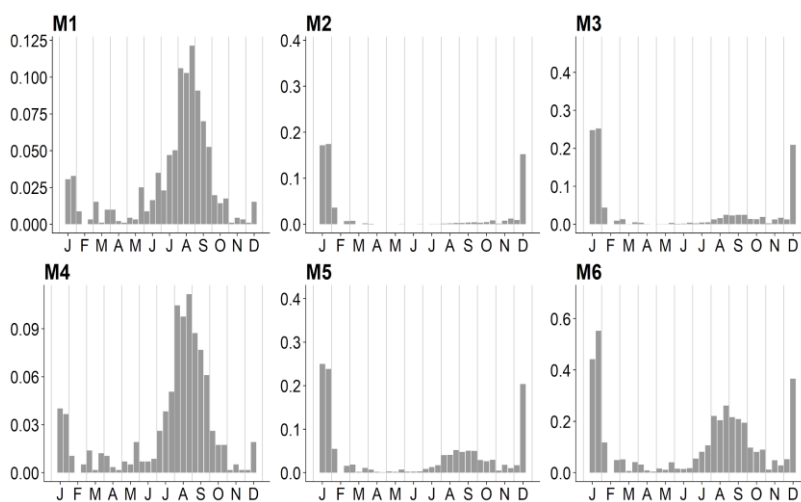
408 4.4 **Other recommendations facing database limitations**

409 Dates in the INPN are compulsory, so when the day of observation is not known a
410 default date (such as January 1st or December 31st) is assigned. We discarded outlier data (see
411 section 2.3) in order to prevent the appearance of artefacts, although some small winter peaks
412 persisted. In fact, as described in section, 4.1, some butterfly species may be recorded during
413 winter days, even if the actual period of activity is later in the year. Unfortunately, we could
414 not discriminate observations by stage of development or behaviour (activity versus
415 hibernation), since this information is currently lacking in most INPN data sources. However,
416 records of active adult butterflies in the INPN usually far exceed observations of young stage
417 or wintering individuals, so the latter are unlikely to affect the displaying of major activity
418 peaks. There are, however, some exceptions, such as those described in section 4.1 or the case
419 of *Thecla betulae* (Linnaeus, 1758). This species is hardly detectable during its flight period
420 (Lafranchis and Geniez 2000; Lafranchis et al. 2015). Conversely, the eggs are easy to

421 recognize and they are regularly used for detecting the geographic presence of the butterfly.
422 These occurrences are often reported as all others in datasets when achieving the INPN. This
423 caused the appearance of improbable peaks in winter (Fig. 7). Correcting for observation
424 effort did not provide any added value, it rather accentuated the aberrations. This underlines
425 the need to support data producers towards a better standardisation of information.

426 Several studies have documented the relationship between climate and phenological
427 shifts (Roy and Sparks 2000, Walther et al. 2002, Parmesan and Yohe 2003, Stefanescu et al.
428 2003, Menzel et al. 2006, Parmesan 2007, Altermatt 2009, Prodon et al. 2017, Bell et al.
429 2019). The impact of climate change on phenology may have major consequences on
430 ecological systems and their conservation (Schwartz 2013), hence the importance of long-
431 term collection of observational and monitoring data. Phenological shifts of the order of some
432 days may have occurred in the last decades (Roy and Sparks 2000). Admittedly, we could not
433 point out phenological shifts due to coarse temporal resolution. A finer temporal resolution
434 would entail the loss of useful data. Furthermore, we chose to collate all years' data in order
435 to compensate for uneven recording intensity and species detectability across the years. This
436 also precluded the study of phenological shifts. However, we believe that our methods are
437 adequate to provide a simple measure for a general audience that bears overall phenological
438 patterns for many species. Nevertheless, we will consider investigating with more accuracy
439 the phenology of those butterflies for which precise dates of observation are abundant and
440 consistent through the years by adapting methods that use statistical models, which account
441 for phenological changes over time, observation effort, detection and reporting bias, such as
442 those proposed by Dennis et al. (2017b) and Strebel et al. (2014).

443



444

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

447 **5 Conclusion**

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

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

465 **Acknowledgments**

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

471 **References**

472 Altermatt, F. 2009. Climatic warming increases voltinism in European butterflies and moths.
473 *Proceedings of the Royal Society B: Biological Sciences* 277:1281–1287. doi:
474 10.1098/rspb.2009.1910

- 475 Archaux, F., P. Chatard, F. Faucheux, and A. Lévêque. 2015. Papillons du Loiret: atlas des
476 rhopalocères et zygènes du Loiret (2000-2013). So.MOS & Alexanor.
- 477 Archaux, F., G. Jacquemin, R. Leconte, G. C. Luquet, F. Mora, F. Noël, H. Pinston, J.-C.
478 Robert, and A. Ruffoni. 2011. Synthèse des observations récentes et anciennes de
479 Libelloides coccajus (Denis & Schiffermüller) et L. longicornis (Linné) dans la moitié
480 nord de la France (Neuroptera, Ascalaphidae). Bulletin de la Société entomologique de
481 France 116:365–388.
- 482 August, T. O. M., M. Harvey, P. Lightfoot, D. Kilbey, T. Papadopoulos, and P. Jepson. 2015.
483 Emerging technologies for biological recording: 731–749.
- 484 Bell, J.R., M.S. Botham, P.A. Henrys, D.I. Leech, J.W. Pearce-Higgins, C.R. Shortall, T.M.
485 Brereton, J. Pickup, and S.J. Thackeray, 2019. Spatial and habitat variation in aphid,
486 butterfly, moth and bird phenologies over the last half century. *Global change biology*,
487 25(6), pp.1982-1994.
- 488 Bertone, M., J. Green, S. Washburn, M. Poore, C. Soprenson, and D. W. Watson. 2005.
489 Seasonal Activity and Species Composition of dung beetles inhabiting cattle pastures in
490 North Carolina. *Entomological Society of America*. 98:309–321.
- 491 Dennis, E. B., S. N. Freeman, T. Brereton, and D. B. Roy. 2013. Indexing butterfly abundance
492 whilst accounting for missing counts and variability in seasonal pattern. *Methods in*
493 *Ecology and Evolution* 4:637–645. doi: 10.1111/2041-210X.12053
- 494 Dennis, E. B., B. J. T. Morgan, T. M. Brereton, D. B. Roy, and R. Fox. 2017a. Using citizen
495 science butterfly counts to predict species population trends. *Conservation Biology*
496 31:1350–1361.
- 497 Dennis, E.B., Morgan, B.J., Freeman, S.N., Ridout, M.S., Brereton, T.M., Fox, R., Powney,
498 G.D. and Roy, D.B., 2017b. Efficient occupancy model-fitting for extensive citizen-
499 science data. *PloS one*, 12(3), p.e0174433
- 500 Dickinson, J. L., J. Shirk, D. Bonter, R. Bonney, R. L. Crain, J. Martin, T. Phillips, and K.
501 Purcell. 2012. The current state of citizen science as a tool for ecological research and
502 public engagement. *Frontiers in Ecology and the Environment* 10:291–297. doi:
503 10.1890/110236
- 504 Dickinson, J. L., B. Zuckerberg, and D. N. Bonter. 2010. Citizen Science as an Ecological
505 Research Tool: Challenges and Benefits. *Annual Review of Ecology, Evolution, and*
506 *Systematics* 41:149–172. doi: 10.1146/annurev-ecolsys-102209-144636
- 507 Gargominy O., S. Terceirie, C. Régnier, T. Ramage, P. Dupont, P. Daszkiewicz, L. Poncet.
508 2017 TAXREF v11.0, référentiel taxonomique pour la France : méthodologie, mise en
509 œuvre et diffusion. Muséum National d’Histoire Naturelle, Paris. Rapport Patrinat 2017-
510 116.
- 511 Giraud, C., C. Calenge, C. Coron, and R. Julliard. 2016. Capitalizing on opportunistic data for
512 monitoring relative abundances of species. *Biometrics* 72:649–658.
513 doi:10.1111/biom.12431
- 514 Hochachka, W. M., D. Fink, R. A. Hutchinson, D. Sheldon, W. K. Wong, and S. Kelling.
515 2012. Data-intensive science applied to broad-scale citizen science. *Trends in Ecology*
516 *and Evolution* 27:130–137.

- 517 Isaac, N. J. B., and M. J. O. Pocock. 2015. Bias and information in biological records.
518 *Biological Journal of the Linnean Society* 115:522–531. doi: 10.1111/bij.12532
- 519 Isaac, N. J. B., A. J. van Strien, T. A. August, M. P. de Zeeuw, and D. B. Roy. 2014. Statistics
520 for citizen science: Extracting signals of change from noisy ecological data. *Methods in*
521 *Ecology and Evolution*. doi: 10.1111/2041-210X.12254
- 522 Kéry, M., J. A. Royle, H. Schmid, M. Schaub, B. Volet, G. Häfliger, and N. Zbinden. 2010.
523 Site-Occupancy Distribution Modeling to Correct Population-Trend Estimates Derived
524 from Opportunistic Observations. *Conservation Biology* 24:1388–1397. doi:
525 10.1111/j.1523-1739.2010.01479.x
- 526 Lobo, J. M. 2008. Database records as a surrogate for sampling effort provide higher species
527 richness estimations. *Biodiversity and Conservation* 17:873–881. doi: 10.1007/s10531-
528 008-9333-4
- 529 Lumaret J.P. 1990 Atlas des Coléoptères Scarabéides Laparosticti de France. Muséum
530 National d’Histoire Naturelle. Secrétariat de la Faune et la Flore Paris Paris.
- 531 Maes, D., W. Vanreusel, I. Jacobs, K. Berwaerts, and H. Van Dyck. 2012. Applying IUCN
532 Red List criteria at a small regional level: A test case with butterflies in Flanders (north
533 Belgium). *Biological Conservation* 145:258–266. doi: 10.1016/j.biocon.2011.11.021
- 534 Manil, L., A. Lerch, B. Fontaine, and R. Julliard. 2015. Suivi Temporel des Rhopalocères de
535 France (STERF) Bilan 2005-2014. *Gestion* 2010:1–82.
- 536 Menzel, A., T. H. Sparks, N. Estrella, E. Koch, A. Aasa, R. Ahas, K. Alm-Kübler, P.
537 Bissolli, O. Braslavská, A. Briede, F. M. Chmielewski, Z. Crepinsek, Y. Curnel, Å.
538 Dahl, C. Defila, A. Donnelly, Y. Filella, K. Jatzcak, F. Måge, A. Mestre, Ø. Nordli, J.
539 Peñuelas, P. Pirinen, V. Remišová, H. Scheifinger, M. Striz, A. Susnik, A. J. H. Van
540 Vliet, F. E. Wielgolaski, S. Zach, and A. Zust. 2006. European phenological response to
541 climate change matches the warming pattern. *Global Change Biology* 12:1969–1976.
542 doi: 10.1111/j.1365-2486.2006.01193.x
- 543 Metzger, M. J., R. G. H. Bunce, R. H. G. Jongman, C. A. Múcher, and J. W. Watkins. 2005.
544 A climatic stratification of the environment of Europe. *Global Ecology and*
545 *Biogeography* 14:549–563. doi: 10.1111/j.1466-822X.2005.00190.x
- 546 Parmesan, C. 2007. Influences of species, latitudes and methodologies on estimates of
547 phenological response to global warming. *Global Change Biology* 13:1860–1872. doi:
548 10.1111/j.1365-2486.2007.01404.x
- 549 Parmesan, C., and G. Yohe. 2003. A globally coherent fingerprint of climate change impacts
550 across natural systems. *Nature* 421:37–42. doi: 10.1038/nature01286
- 551 Phillips, S. J., M. Dudik, J. Elith, C. H. Graham, A. Lehmann, J. Leathwick, and S. Ferrier.
552 2009. Sample selection bias and presence-only distribution models: implications for
553 background and pseudo-absence data. *ECOLOGICAL APPLICATIONS* 19:181–197.
554 doi: 10.1890/07-2153.1
- 555 Pocock, M. J. O., H. E. Roy, C. D. Preston, and D. B. Roy. 2015. The Biological Records
556 Centre: A pioneer of citizen science. *Biological Journal of the Linnean Society* 115:475–
557 493.

- 558 Ponder, W. F., G. A. Carter, P. Flemons, and R. R. Chapman. 2001. Evaluation of museum
559 collection data for use in biodiversity assessment. *Conservation Biology* 15:648–657.
560 doi: 10.1046/j.1523-1739.2001.015003648.x
- 561 Powney, G. D., and N. J. B. Isaac. 2015. Beyond maps: A review of the applications of
562 biological records. *Biological Journal of the Linnean Society* 115:532–542. doi:
563 10.1111/bij.12517
- 564 Pozo, C., A. Luis-Martínez, J. Llorente-Bousquets, N. Salas-Suárez, A. Maya-Martínez, I.
565 Vargas-Fernández, and A. D. Warren. 2008. Seasonality and Phenology of the
566 Butterflies (Lepidoptera: Papilionoidea and Hesperioidea) of Mexico's Calakmul
567 Region. *Florida Entomologist* 91:407–422. doi: 10.1653/0015-
568 4040(2008)91[407:SAPOTB]2.0.CO;2
- 569 Prodon, R., P. Geniez, M. Cheylan, F. Devers, I. Chuine, and A. Besnard. 2017. A reversal of
570 the shift towards earlier spring phenology in several Mediterranean reptiles and
571 amphibians during the 1998–2013 warming slowdown. *Global Change Biology*
572 23:5481–5491. doi: 10.1111/gcb.13812
- 573 R Core Development Team. 2018. R: A language and environment for statistical computing.
574 R Foundation for Statistical Computing Vienna, Austria.
- 575 Robertson, M. P., G. S. Cumming, and B. F. N. Erasmus. 2010. Getting the most out of atlas
576 data. *Diversity and Distributions* 16:363–375. doi: 10.1111/j.1472-4642.2010.00639.x
- 577 Rothery, P. and D. B. Roy. 2001. Application of generalized additive models to butterfly
578 transect count data. *Journal of Applied Statistics*, 28:897-909.
- 579 Roy, D. B., and T. H. Sparks. 2000. Phenology of British butterflies and climate change.
580 *Global Change Biology* 6:407–416. doi: 10.1046/j.1365-2486.2000.00322.x
- 581 Ruete, A. 2015. Displaying bias in sampling effort of data accessed from biodiversity
582 databases using ignorance maps. *Biodiversity Data Journal* 3:e5361. doi:
583 10.3897/BDJ.3.e5361
- 584 Schmucki, R., G. Pe'er, D. B. Roy, C. Stefanescu, C. A. M. Van Swaay, T. H. Oliver, M.
585 Kuussaari, A. J. Van Strien, L. Ries, J. Settele, M. Musche, J. Carnicer, O. Schweiger, T.
586 M. Brereton, A. Harpke, J. Heliölä, E. Kühn, and R. Julliard. 2016. A regionally
587 informed abundance index for supporting integrative analyses across butterfly
588 monitoring schemes. *Journal of Applied Ecology* 53:501–510. doi: 10.1111/1365-
589 2664.12561
- 590 Schwartz, M. D. 2013. Phenology: An integrative environmental science. *Phenology: An*
591 *Integrative Environmental Science*:53–65. doi: 10.1007/978-94-007-6925-0
- 592 Stefanescu, C., J. Peñuelas, and I. Filella. 2003. Effects of climatic change on the phenology
593 of butterflies in the northwest Mediterranean Basin. *Global Change Biology* 9:1494–
594 1506. doi: 10.1046/j.1365-2486.2003.00682.x
- 595 Strebel, N., M. Kéry, M. Schaub, and H. Schmid. 2014. Studying phenology by flexible
596 modelling of seasonal detectability peaks. *Methods in Ecology and Evolution* 5:483–490.
597 doi: 10.1111/2041-210X.12175
- 598 van Strien, A. J., C. A. M. van Swaay, and T. Termaat. 2013. Opportunistic citizen science
599 data of animal species produce reliable estimates of distribution trends if analysed with

- 600 occupancy models. *Journal of Applied Ecology* 50:1450–1458. doi: 10.1111/1365-
601 2664.12158
- 602 van Strien, A. J., T. Termaat, D. Groenendijk, V. Mensing, and M. Kéry. 2010. Site-
603 occupancy models may offer new opportunities for dragonfly monitoring based on daily
604 species lists. *Basic and Applied Ecology* 11:495–503. doi: 10.1016/j.baae.2010.05.003
- 605 van Swaay, C. A. M., C. L. Plate, and A. J. van Strien. 2002. Monitoring butterflies in the
606 Netherlands: how to get unbiased indices. *Proceedings of the Section Experimental and*
607 *Applied Entomology of the Netherlands Entomological Society* 13:21–27.
- 608 van Swaay, C. A. M. 1990. An assessment of the changes in butterfly abundance in The
609 Netherlands during the 20th Century. *Biological Conservation* 52:287–302.
- 610 Touroult, J., L. Poncet, P. Keith, V. Bouillet, G. Arnal, H. Brustel and J. P. Sibley. 2015.
611 Inventaires et atlas nationaux de distribution: pour une approche plus itérative et un
612 rééquilibrage taxinomique. *Revue d'écologie*.
- 613 Tulloch, A. I. T., H. P. Possingham, L. N. Joseph, J. Szabo, and T. G. Martin. 2013. Realising
614 the full potential of citizen science monitoring programs. *Biological Conservation*
615 165:128–138. doi: 10.1016/j.biocon.2013.05.025
- 616 Turin, H., and P. J. Den Boer. 1988. Changes in the distribution of carabid beetles in The
617 Netherlands since 1880, II. Isolation of habitats and long-term time trends in the
618 occurrence of carabid species with different powers of dispersal (Coleoptera, Carabidae).
619 *Biological Conservation* 44:179–200. doi: 10.1016/0006-3207(88)90101-2
- 620 Walther, G.-R., E. Post, P. Convey, A. Menzel, C. Parmesan, T. J. C. Beebee, J.-M.
621 Fromentin, O. Hoegh-Guldberg, and F. Bairlein. 2002. Ecological responses to recent
622 climate change. *Nature* 416:389–395. doi: 10.1038/416389a
- 623 Wolda, H. 1988. Insect Seasonality: Why. *Annual Review of Ecology and Systematics* 19:1–
624 18. doi: 10.1146/annurev.ecolsys.19.1.1
- 625 Wood, S.N., 2017. *Generalized additive models: an introduction with R*. Chapman and
626 Hall/CRC

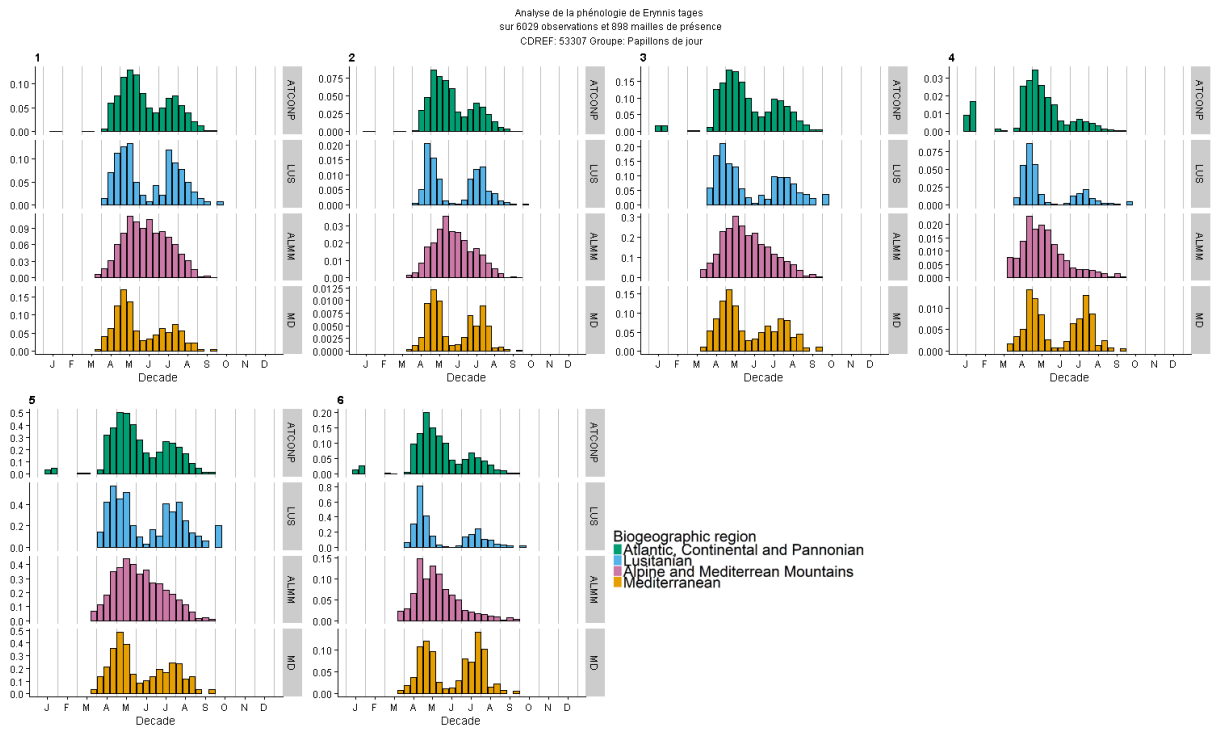
627 **Appendix**

628 **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
 629 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	<i>Carcharodus alceae</i> (Esper, 1780)	2865	673	N	737	135	80
	<i>Erynnis tages</i> (Linnaeus, 1758)	6023	898	Y	2662	423	97
	<i>Ochlodes sylvanus</i> (Esper, 1777)	-	-	-	5664	1132	125
	<i>Pyrgus malvae</i> (Linnaeus, 1758)	-	-	-	915	154	55
	<i>Thymelicus lineola</i> (Ochsenheimer, 1808)	-	-	-	2476	271	93
	<i>Thymelicus sylvestris</i> (Poda, 1761)	-	-	-	1628	431	103
Lycaenidae	<i>Aricia agestis</i> (Denis & Schiffermüller, 1775)	9188	1321	N	3324	890	111
	<i>Callophrys rubi</i> (Linnaeus, 1758)	-	-	-	1472	167	104
	<i>Celastrina argiolus</i> (Linnaeus, 1758)	-	-	-	3353	751	118
	<i>Cupido alcetas</i> (Hoffmannsegg, 1804)	1216	301	N	-	-	55
	<i>Cupido argiades</i> (Pallas, 1771)	3092	443	Y	1287	159	60
	<i>Cupido minimus</i> (Fuessly, 1775)	3521	560	Y	1225	245	76
	<i>Cyaniris semiargus</i> (Rottemburg, 1775)	3538	681	Y	1351	104	89
	<i>Lycaena dispar</i> (Haworth, 1802)	2422	352	Y	-	-	50
	<i>Lycaena phlaeas</i> (Linnaeus, 1760)	-	-	-	2229	419	109
	<i>Lycaena tityrus</i> (Poda, 1761)	3169	767	Y	1383	265	92
	<i>Lysandra bellargus</i> (Rottemburg, 1775)	10028	856	Y	3644	885	108
	<i>Lysandra coridon</i> (Poda, 1761)	-	-	-	3093	835	94
	<i>Lysandra hispana</i> (Herrich-Schäffer, 1852)	1364	192	N	-	-	31
	<i>Plebejus argus</i> (Linnaeus, 1758)	2785	435	Y	-	-	67
	<i>Plebejus argyrognomon</i> (Bergsträsser, 1779)	1167	225	N	616	204	46
	<i>Polyommatus icarus</i> (Rottemburg, 1775)	27821	1991	Y	10738	2440	137
<i>Satyrrium ilicis</i> (Esper, 1779)	-	-	-	188	116	69	
<i>Thecla betulae</i> (Linnaeus, 1758)	913	325	N	-	-	54	
Nymphalidae	<i>Aglais io</i> (Linnaeus, 1758)	17761	1716	Y	9001	1828	128
	<i>Aglais urticae</i> (Linnaeus, 1758)	16901	1423	Y	9405	1399	120
	<i>Aphantopus hyperantus</i> (Linnaeus, 1758)	-	-	-	7500	1161	103
	<i>Araschnia levana</i> (Linnaeus, 1758)	6291	965	N	4474	549	77
	<i>Argynnis paphia</i> (Linnaeus, 1758)	-	-	-	4526	965	121

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

631 **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
 632 rules behind index building were kept from the experts so that they could not prejudice the index in itself or identify the same index among
 633 several species.



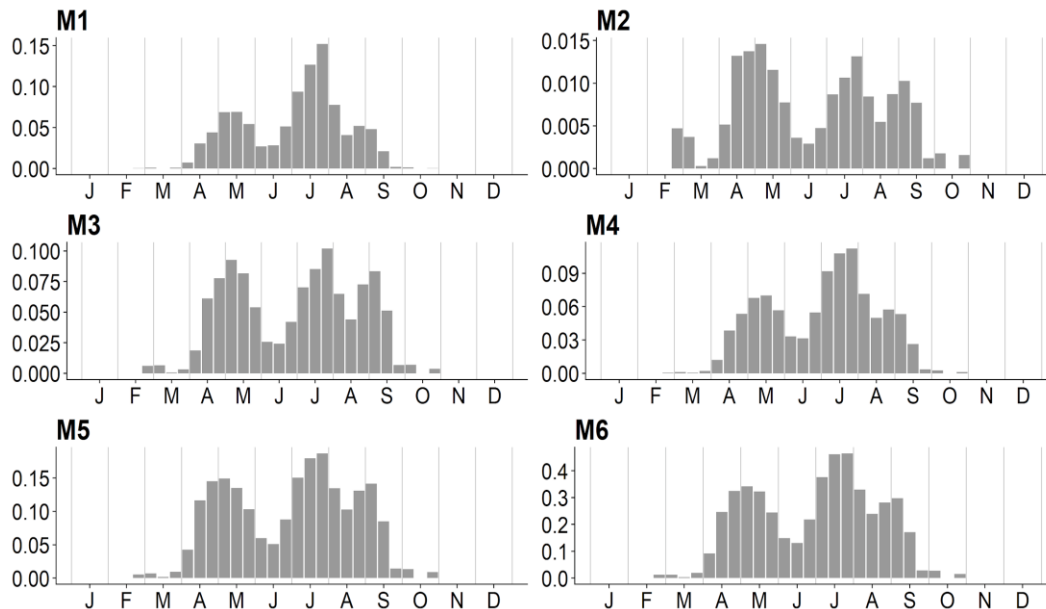
634

635

636 **Appendix 3.** Phenology of 11 butterfly species (2 HesperIIDae, 5 Lycaenidae, 3 Nymphalidae, 1 Pieridae) for which biogeographic effect was
 637 not expected. Phenology is represented with six indices (see also Tab. 2): proportion of records per period (M1); ratio of records to the group
 638 (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
 639 the group within the species known distribution (M6).

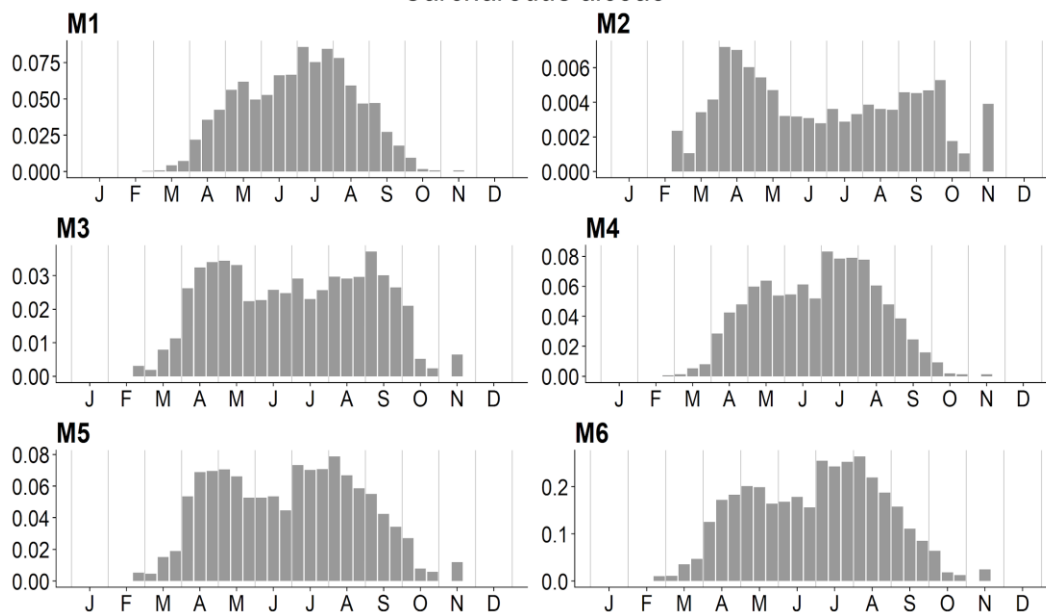
640 HesperIIDae:

Araschnia levana



641

Carcharodus alceae

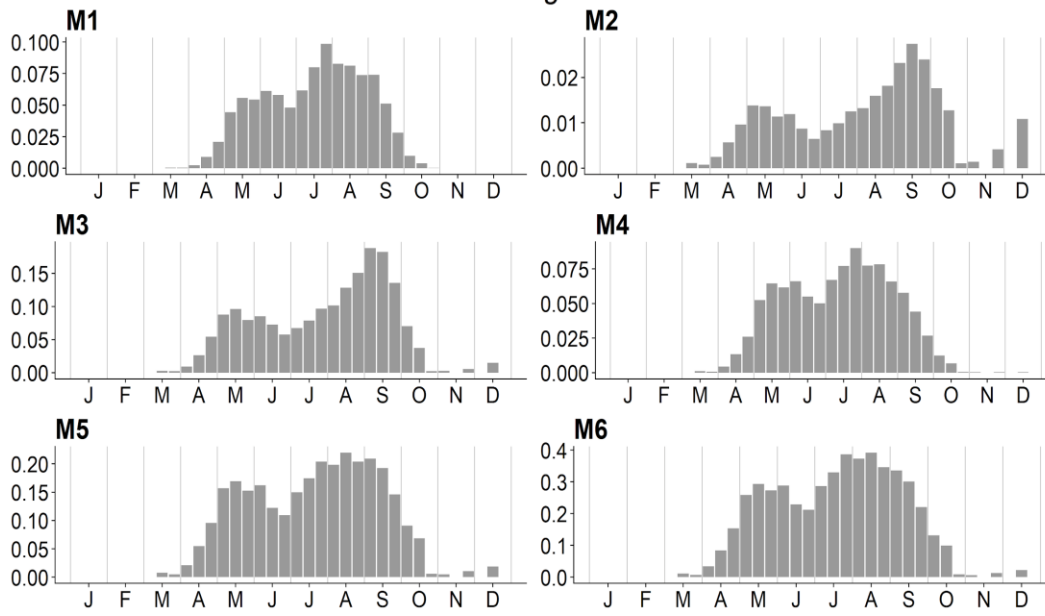


642

643 Lycaenidae:

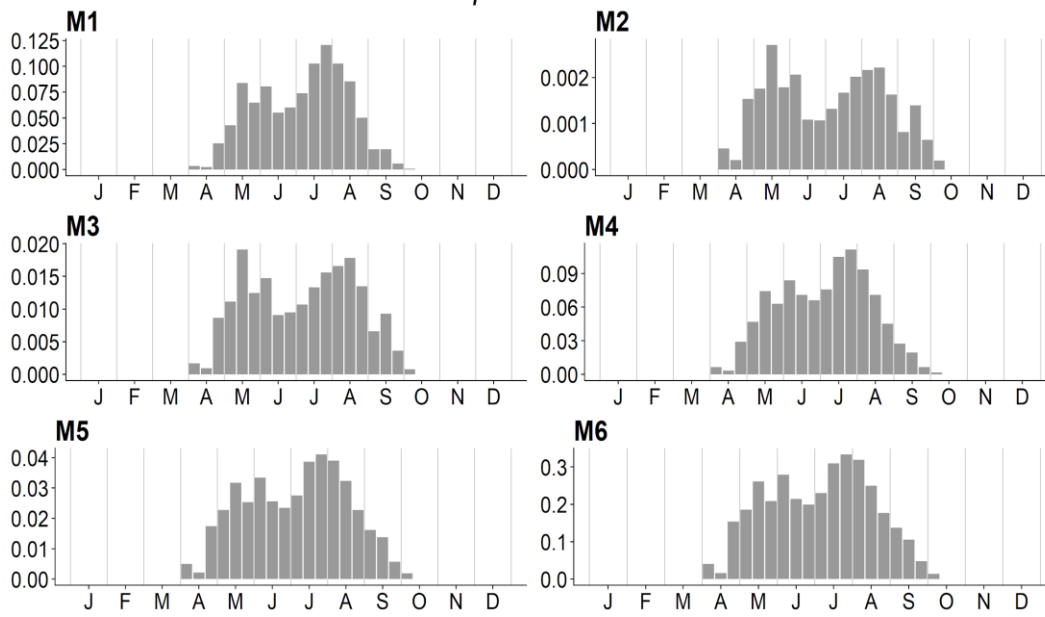
644

Aricia agestis



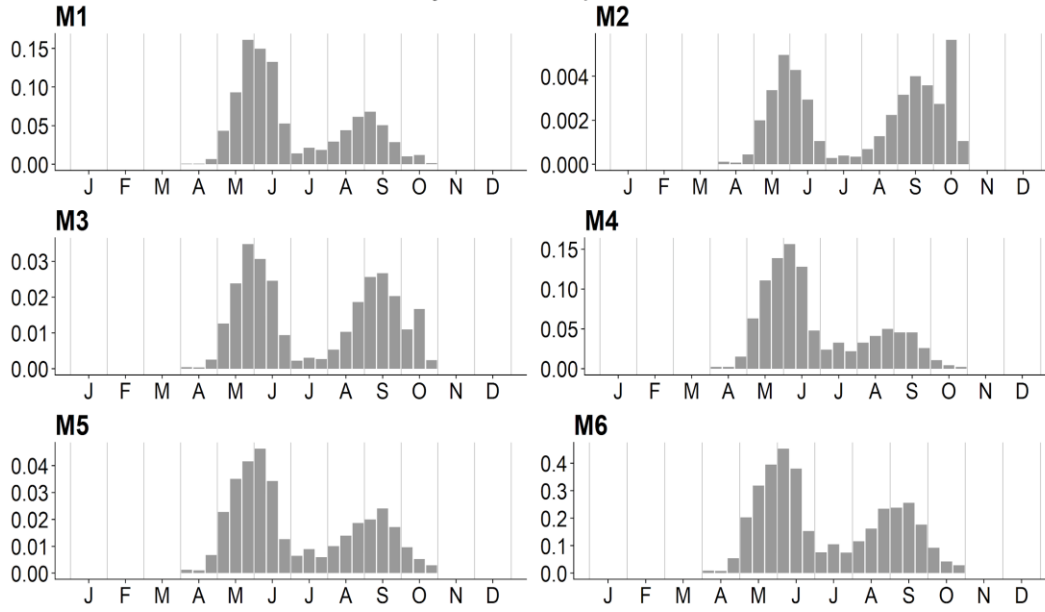
645

Cupido alceas



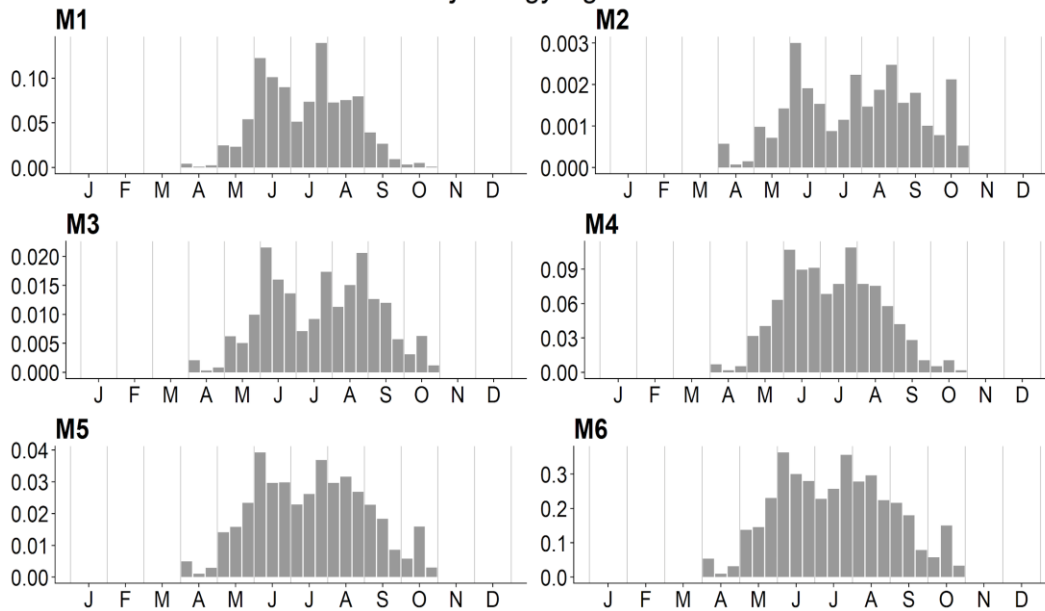
646

Lysandra hispana



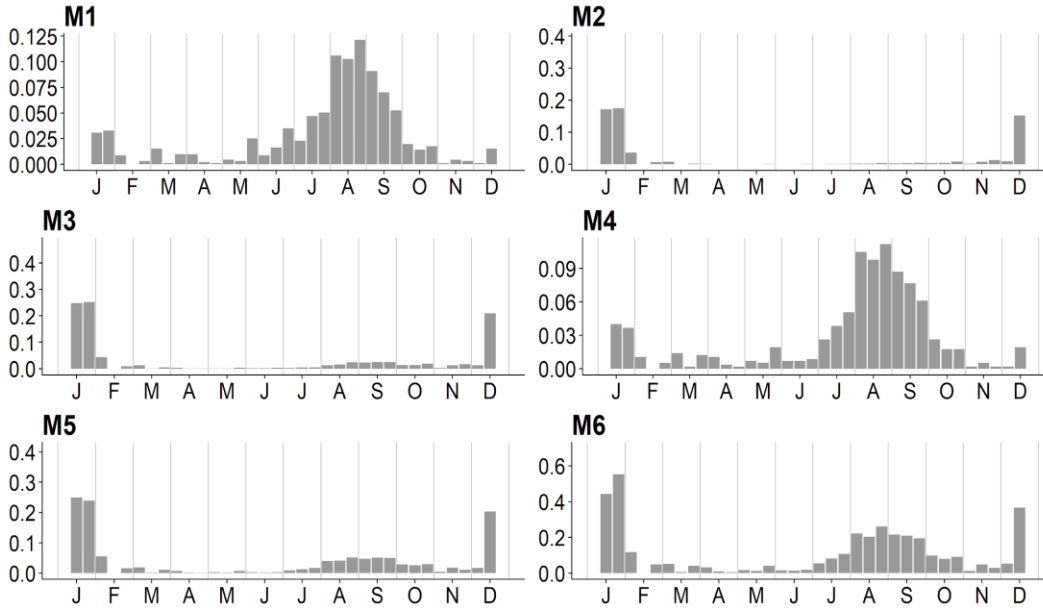
647

Plebejus argyrognomon



648

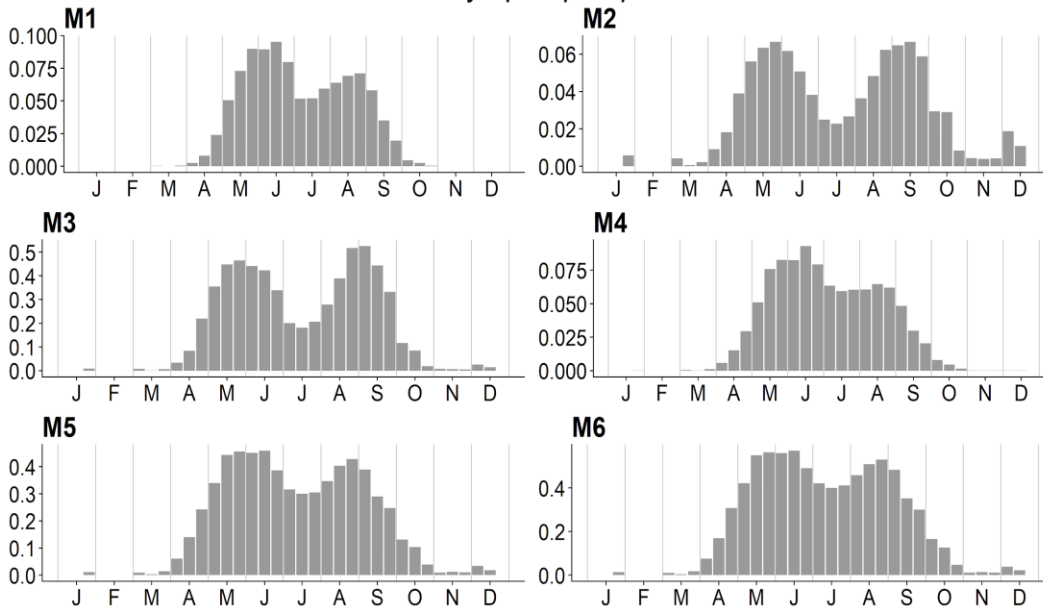
Thecla betulae



649

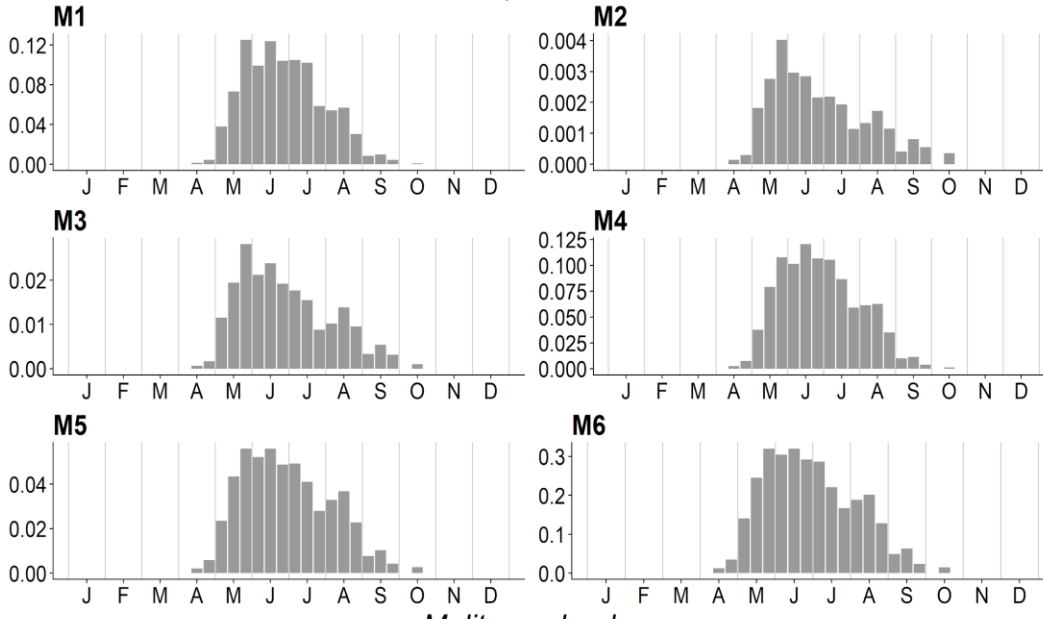
650 Nymphalidae:

Coenonympha pamphilus



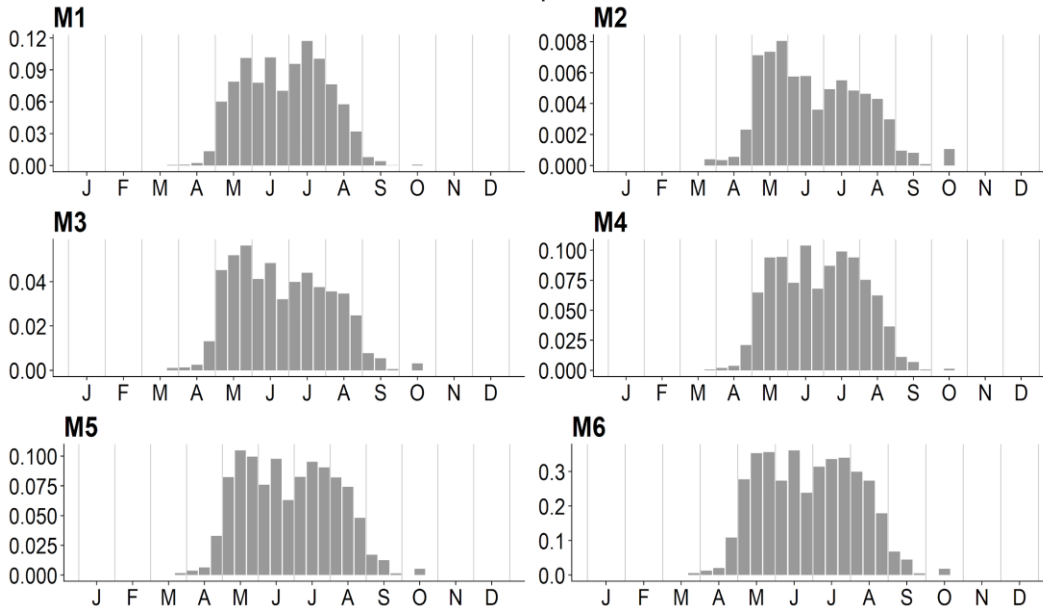
651

Melitaea parthenoides



652

Melitaea phoebe

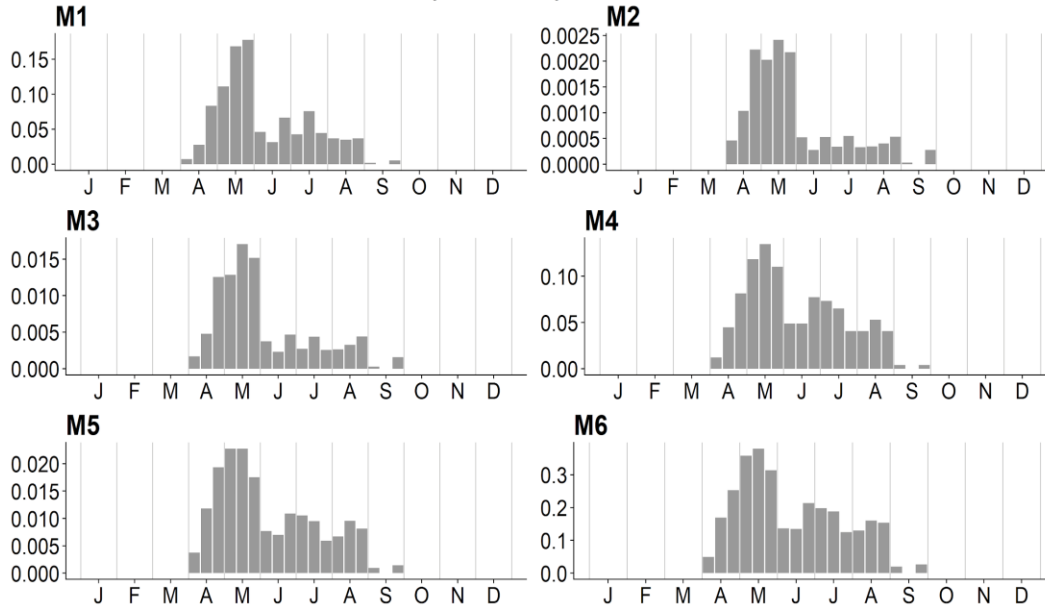


653

654 Pieridae:

655

Leptidea duponcheli

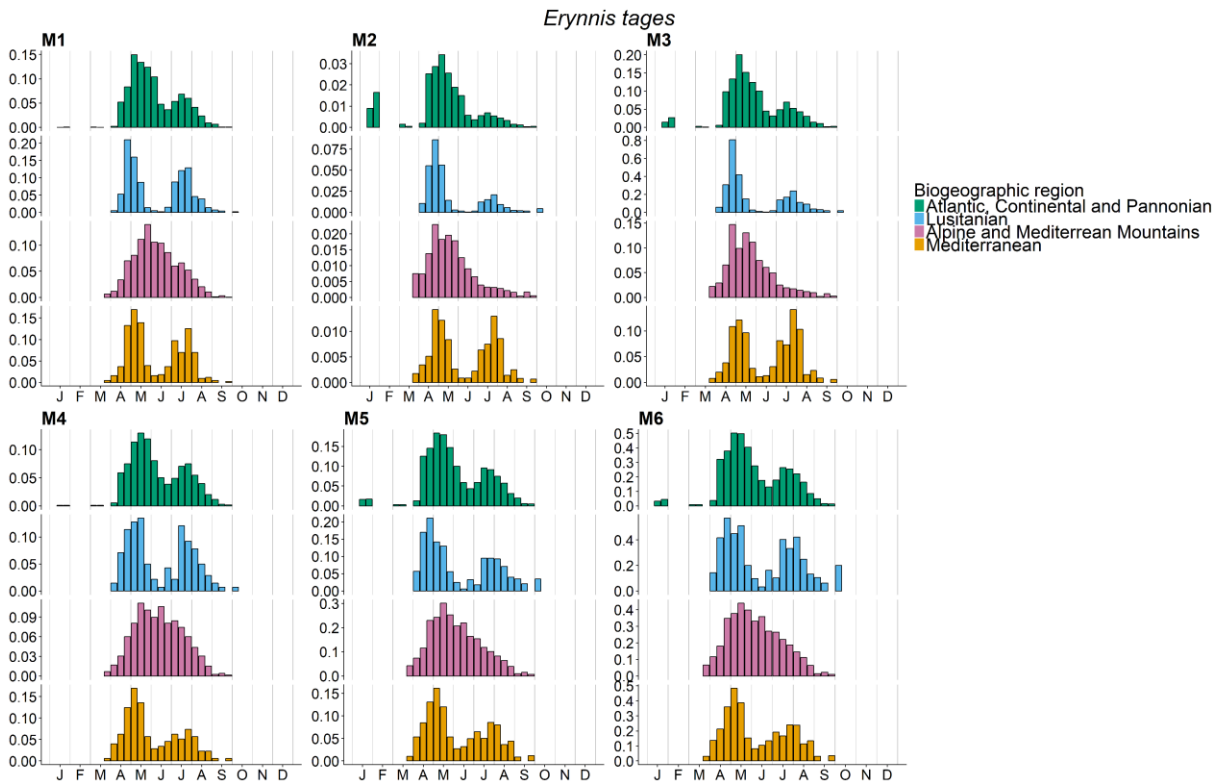


656

657

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

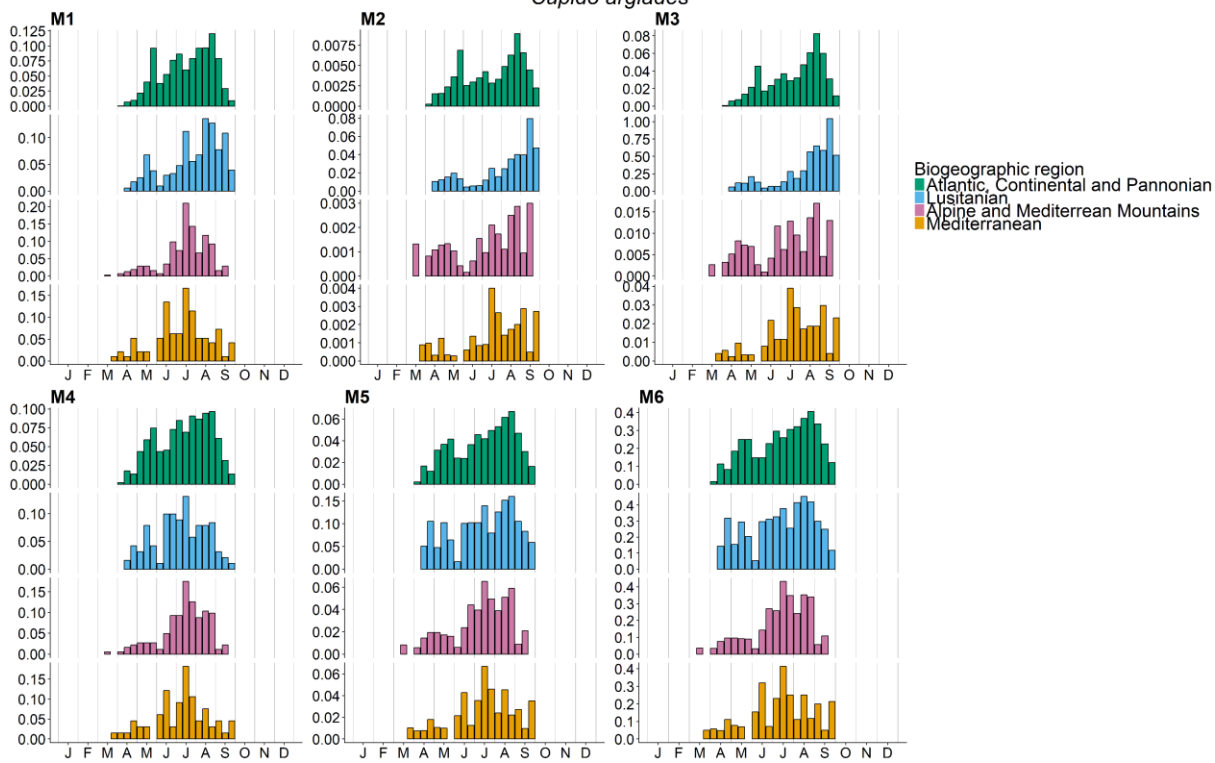
662 Hesperidae:



663

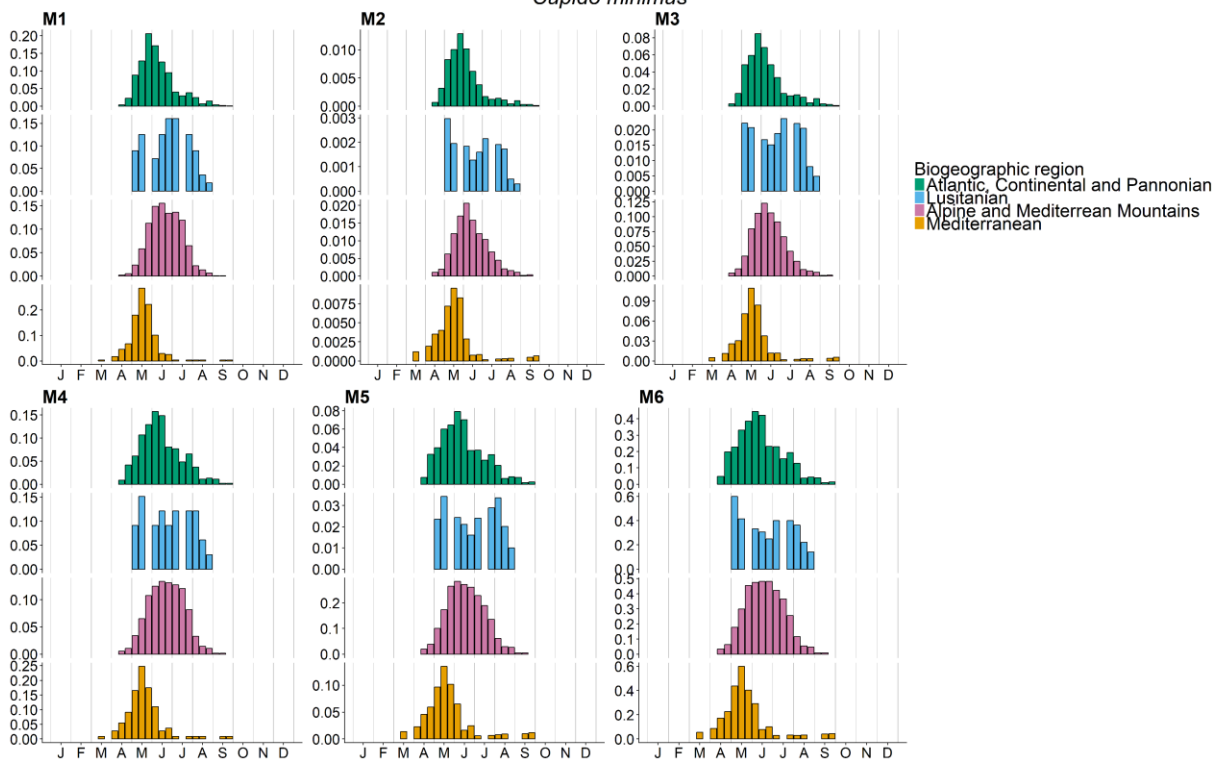
664 Lycaenidae:

Cupido argiades



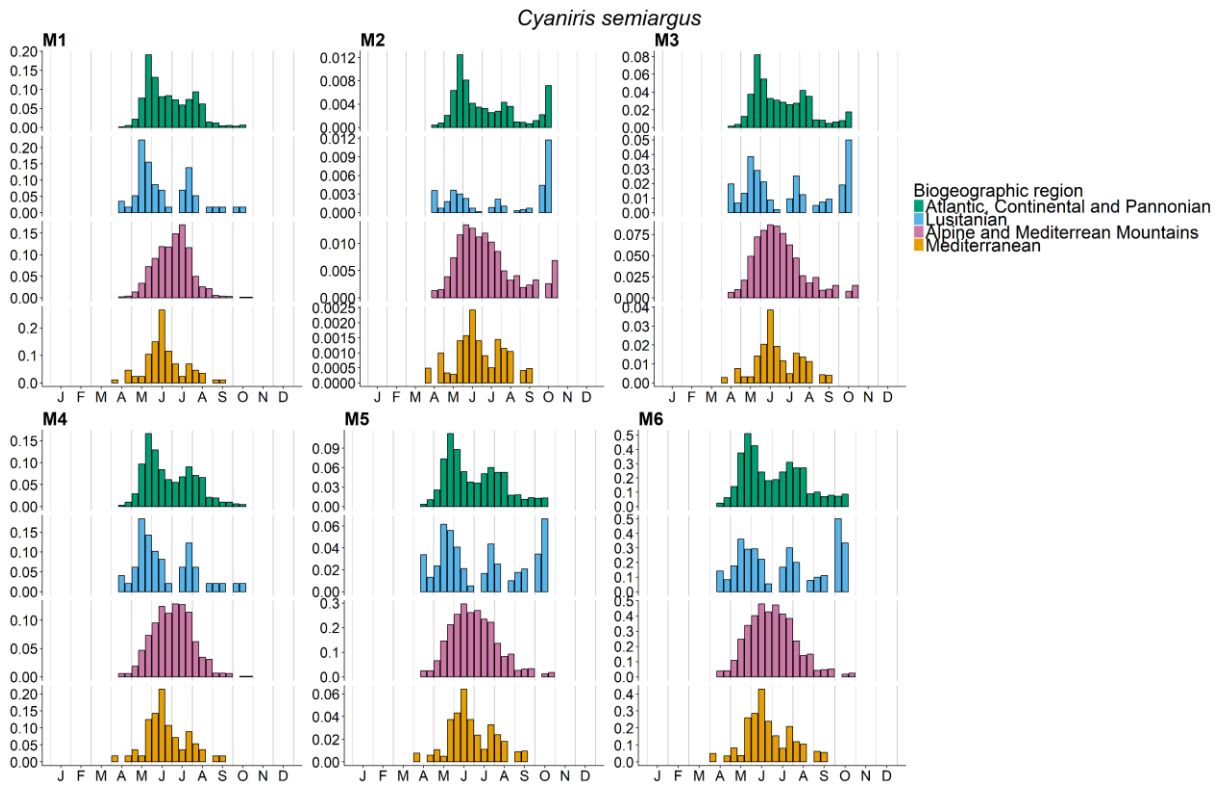
665

Cupido minimus

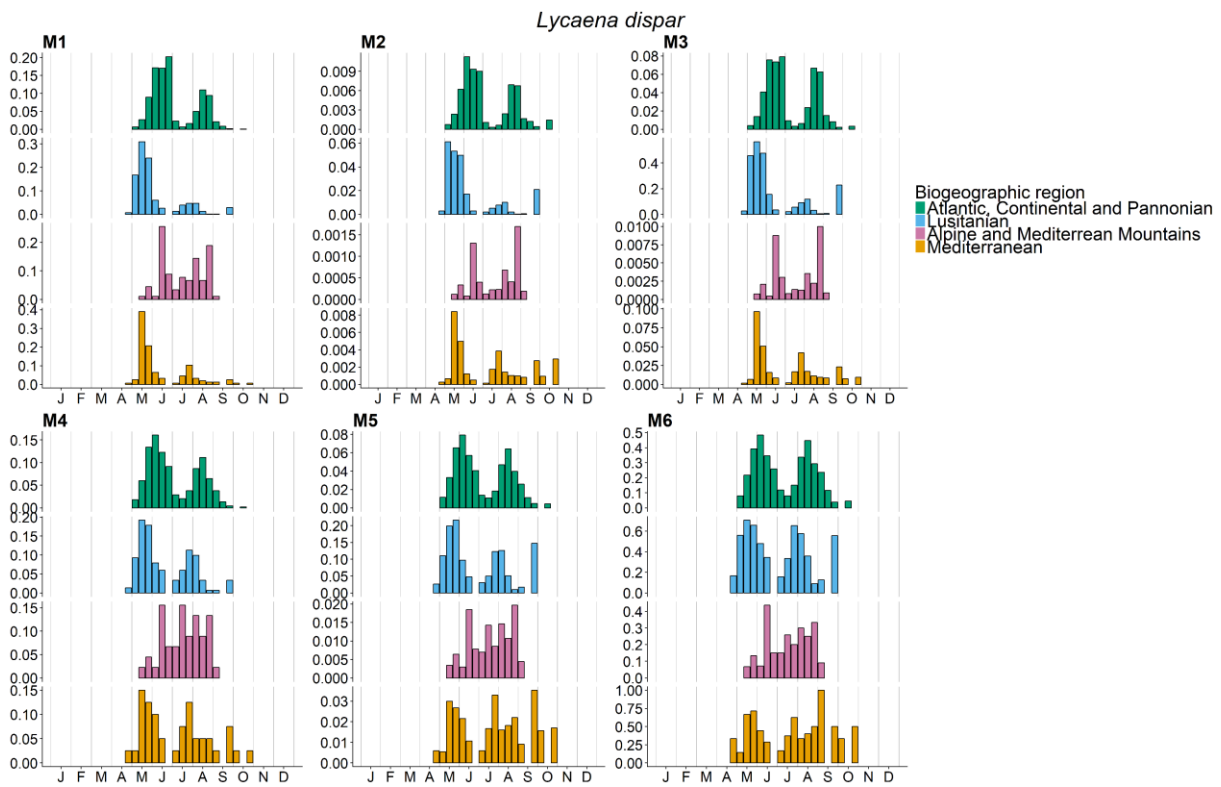


666

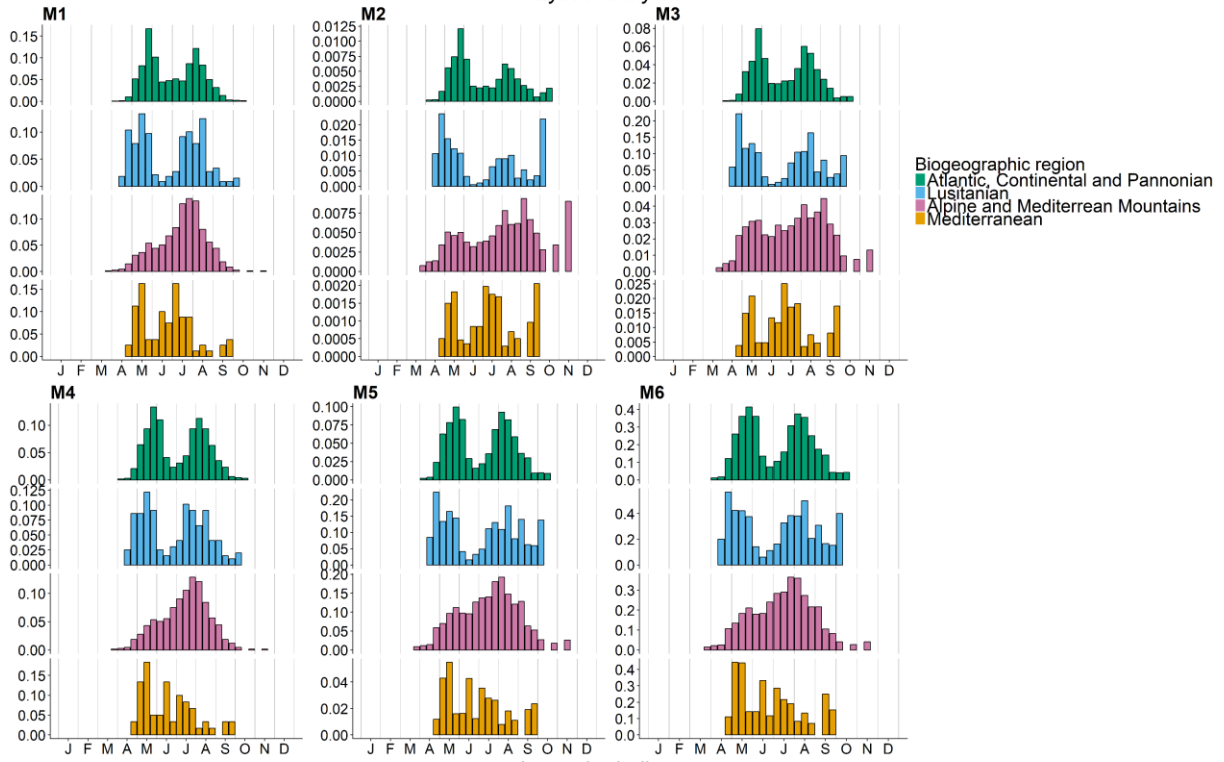
667



668

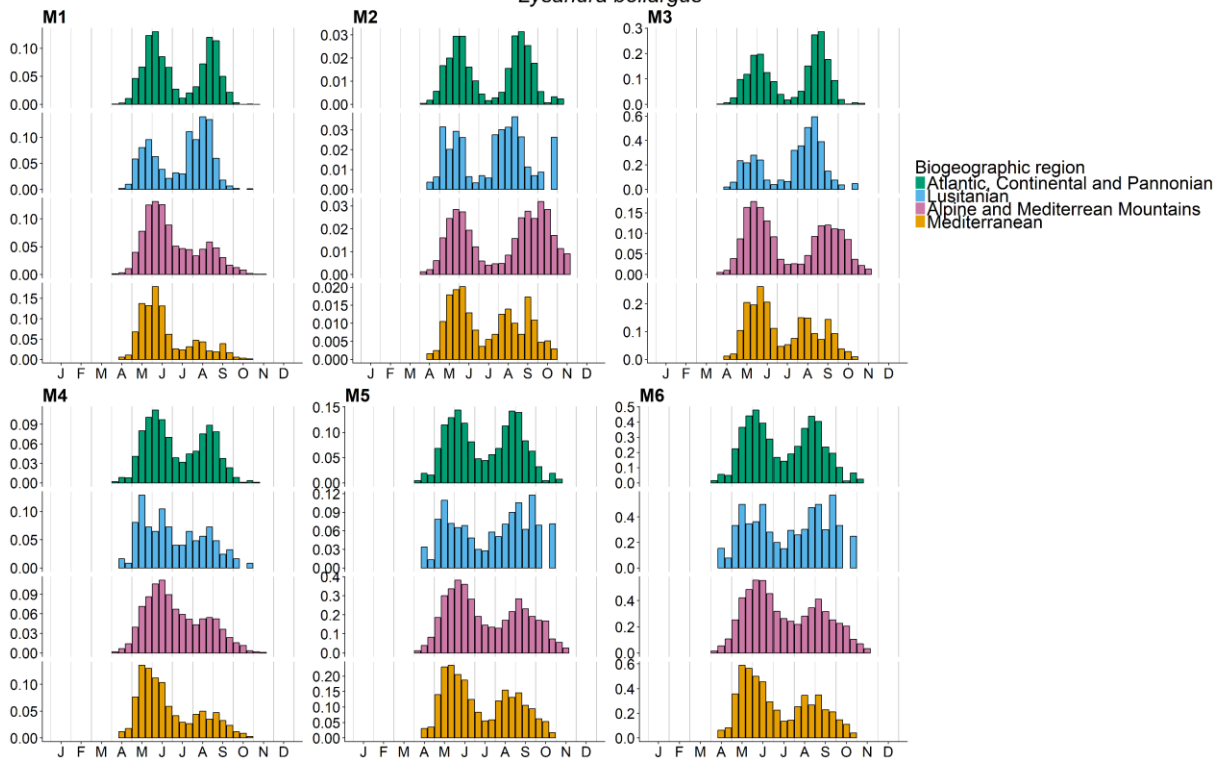


Lycaena tityrus



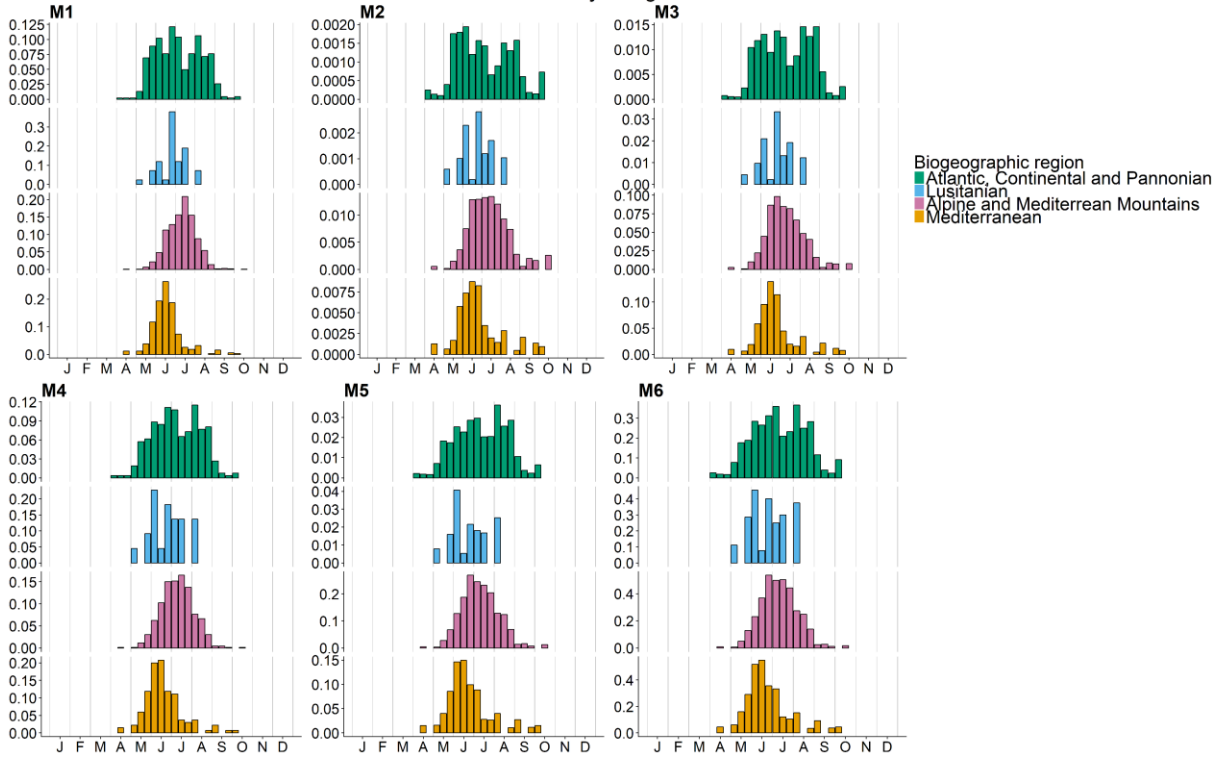
669

Lysandra bellargus



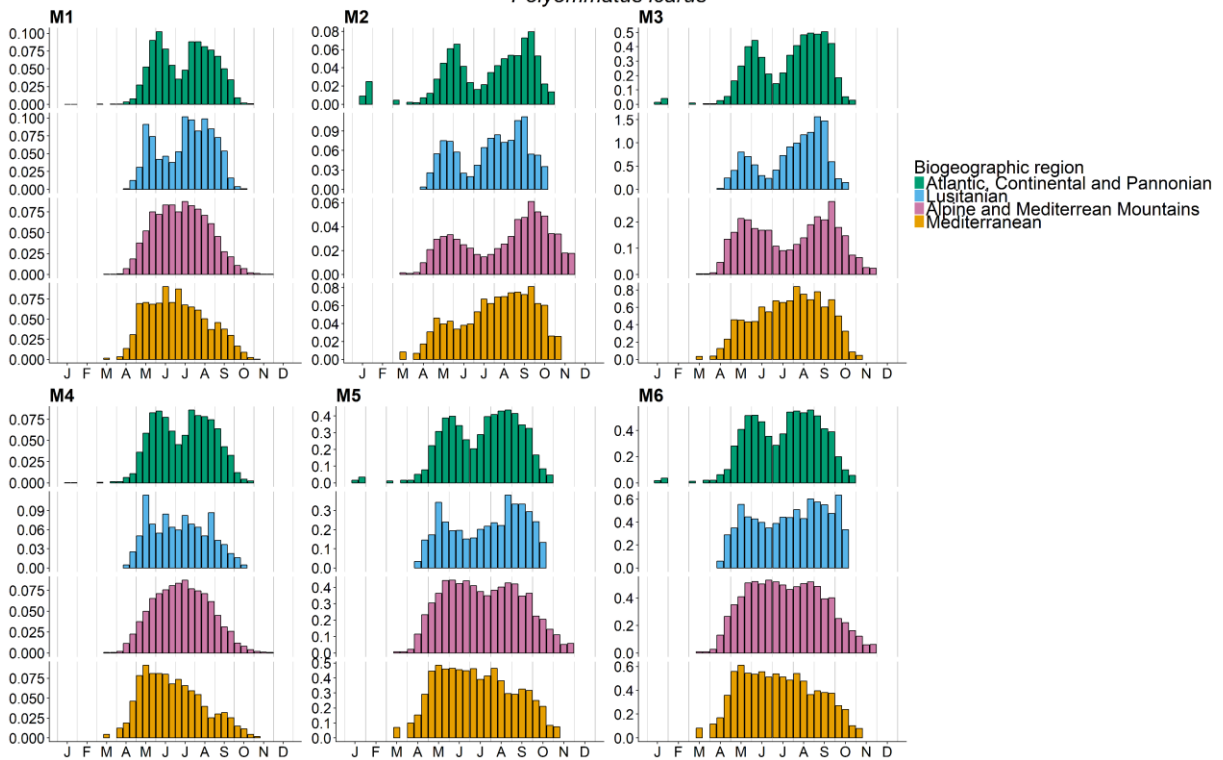
670

Plebejus argus



671

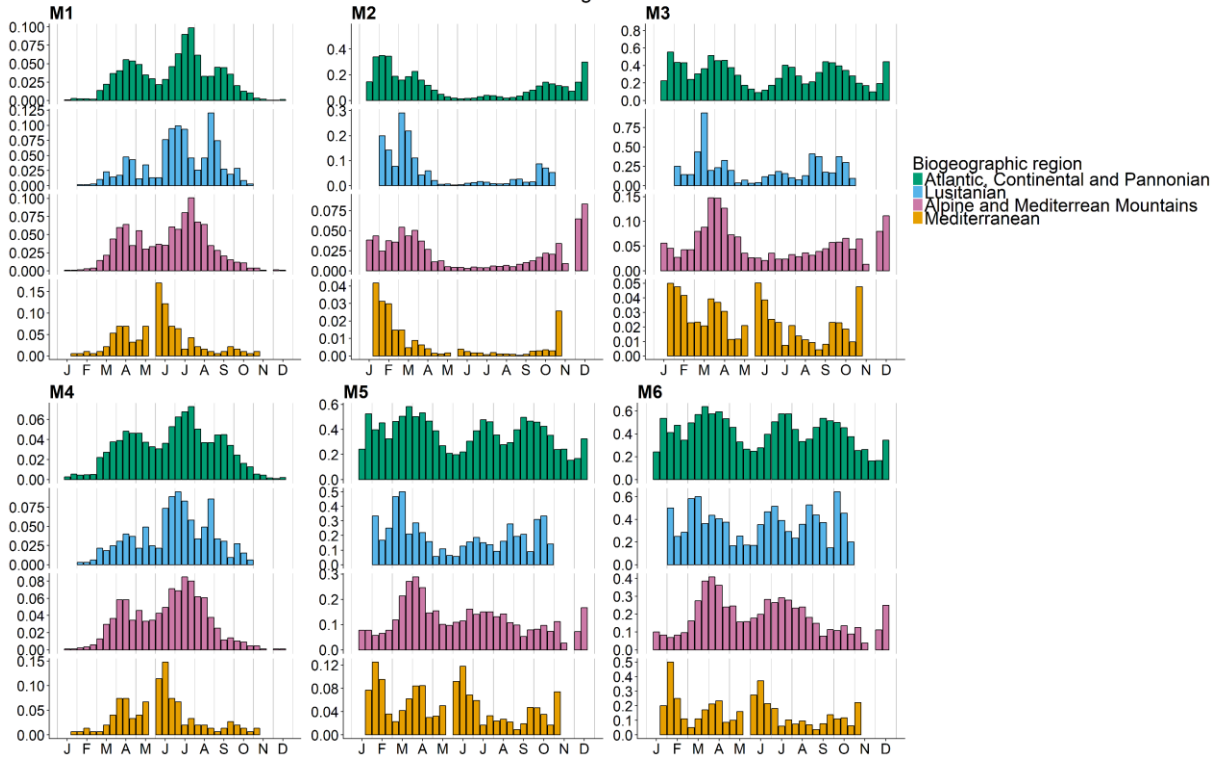
Polyommatus icarus



672

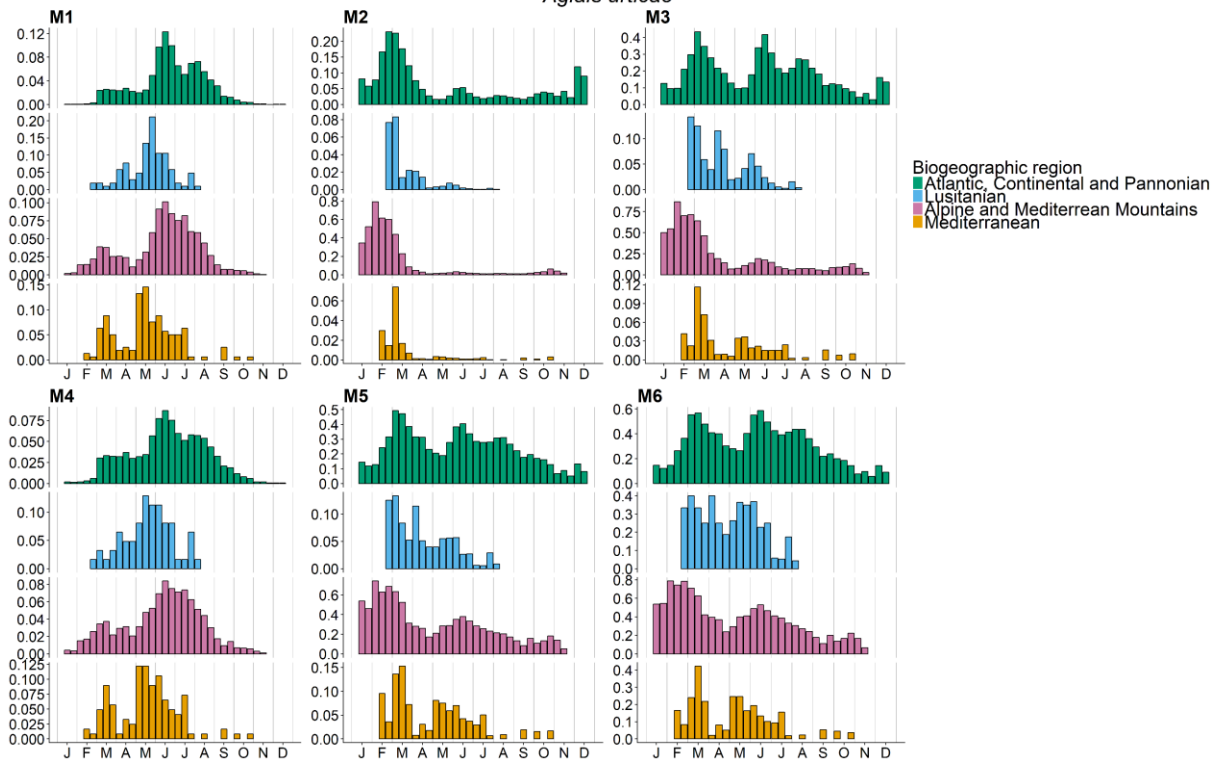
673 Nymphalidae:

Aglais io



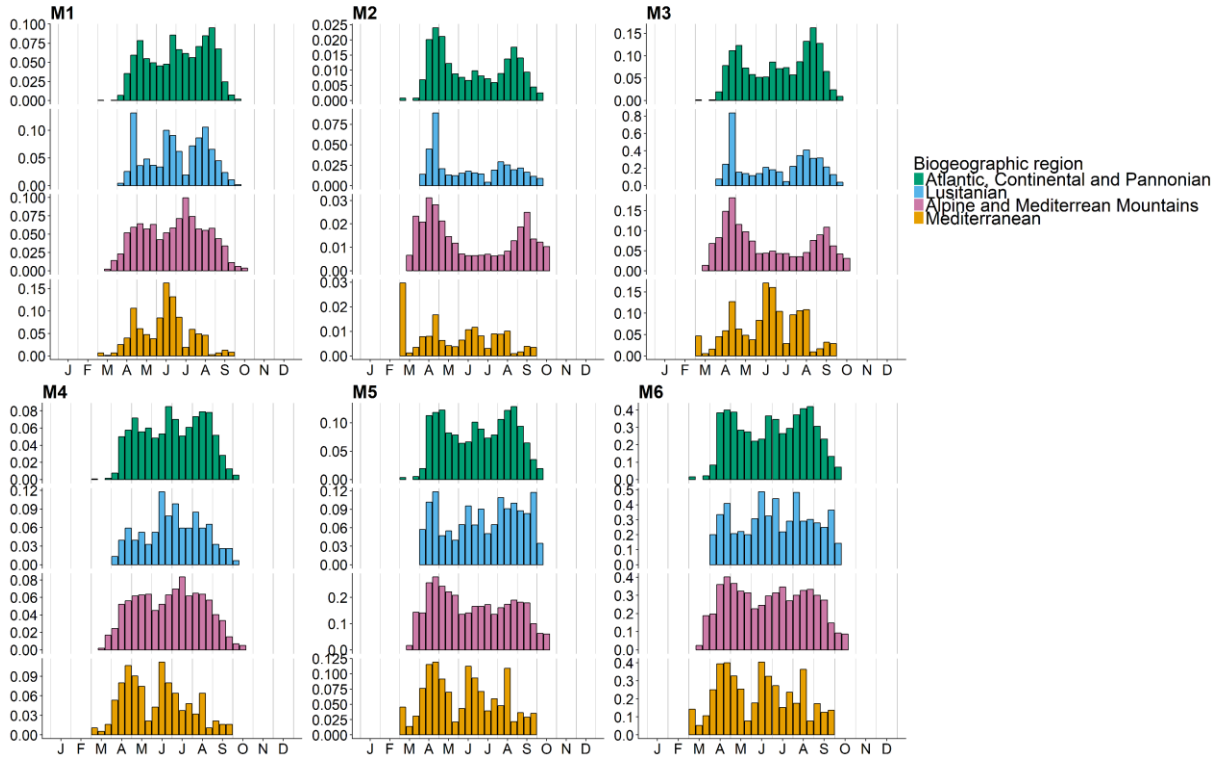
674

Aglais urticae

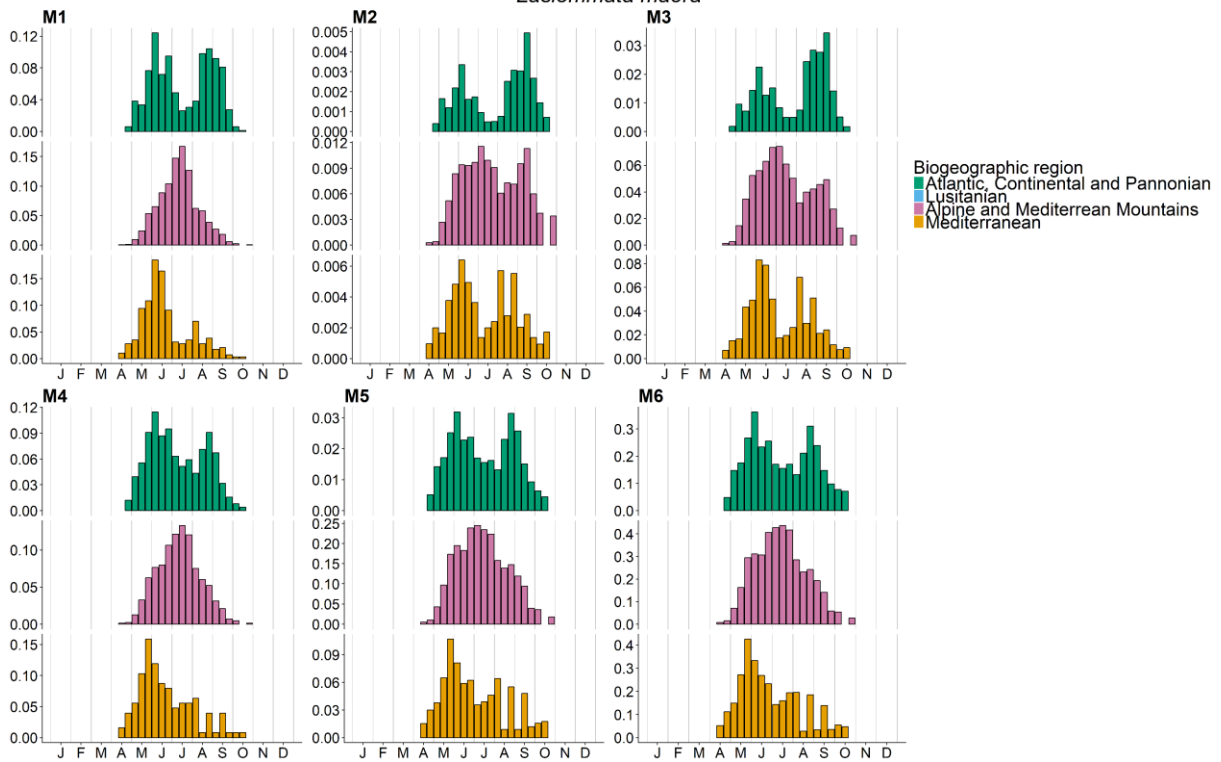


675

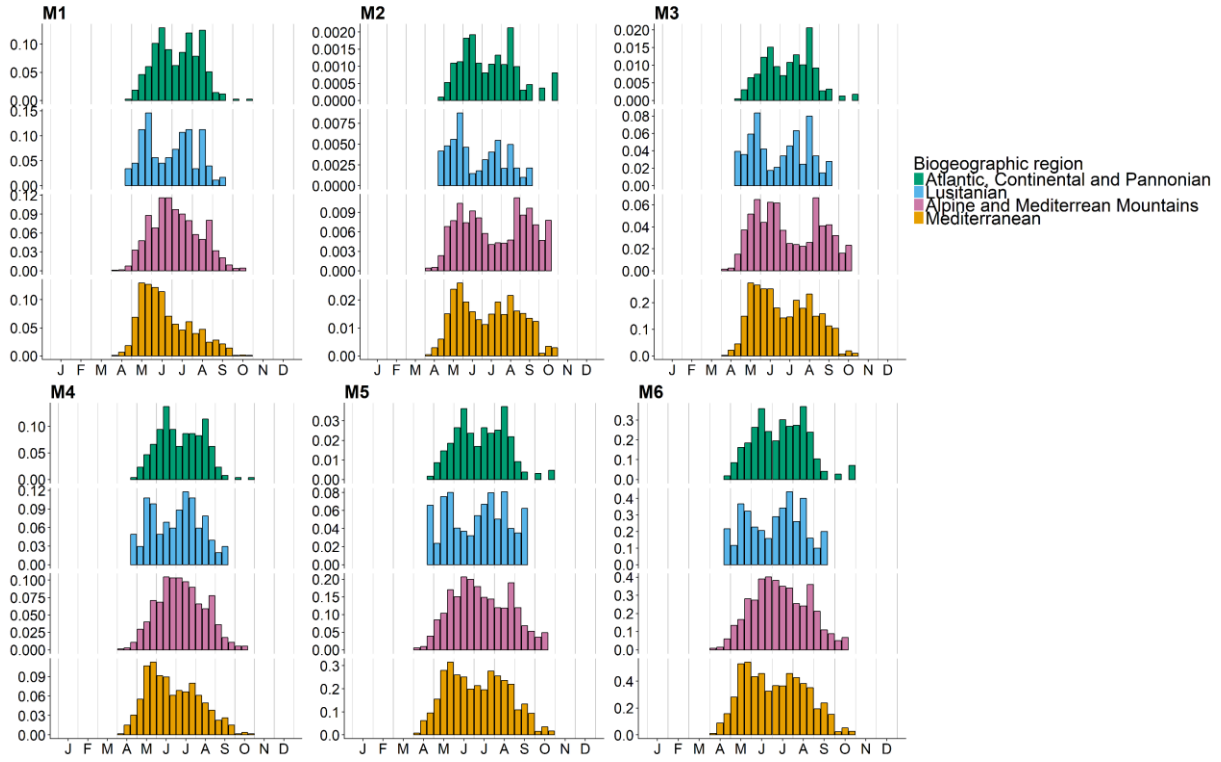
Boloria dia



Lasiommata maera

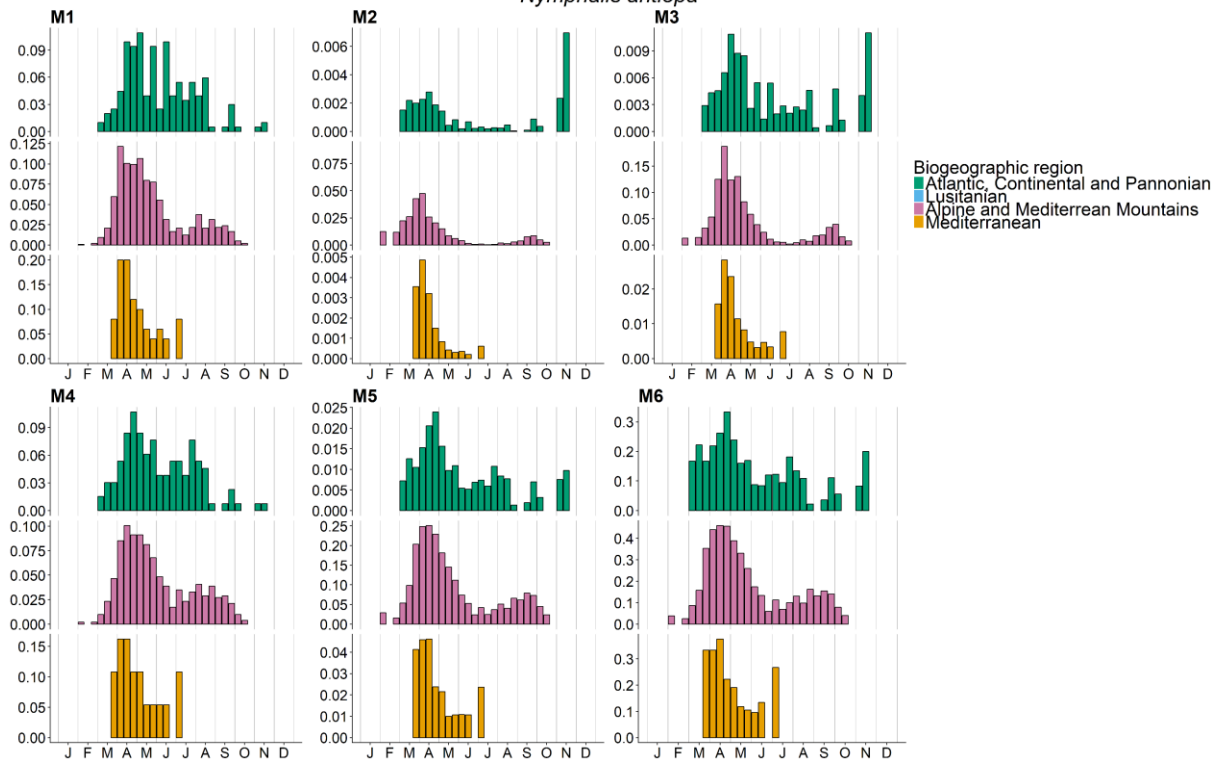


Limenitis reducta

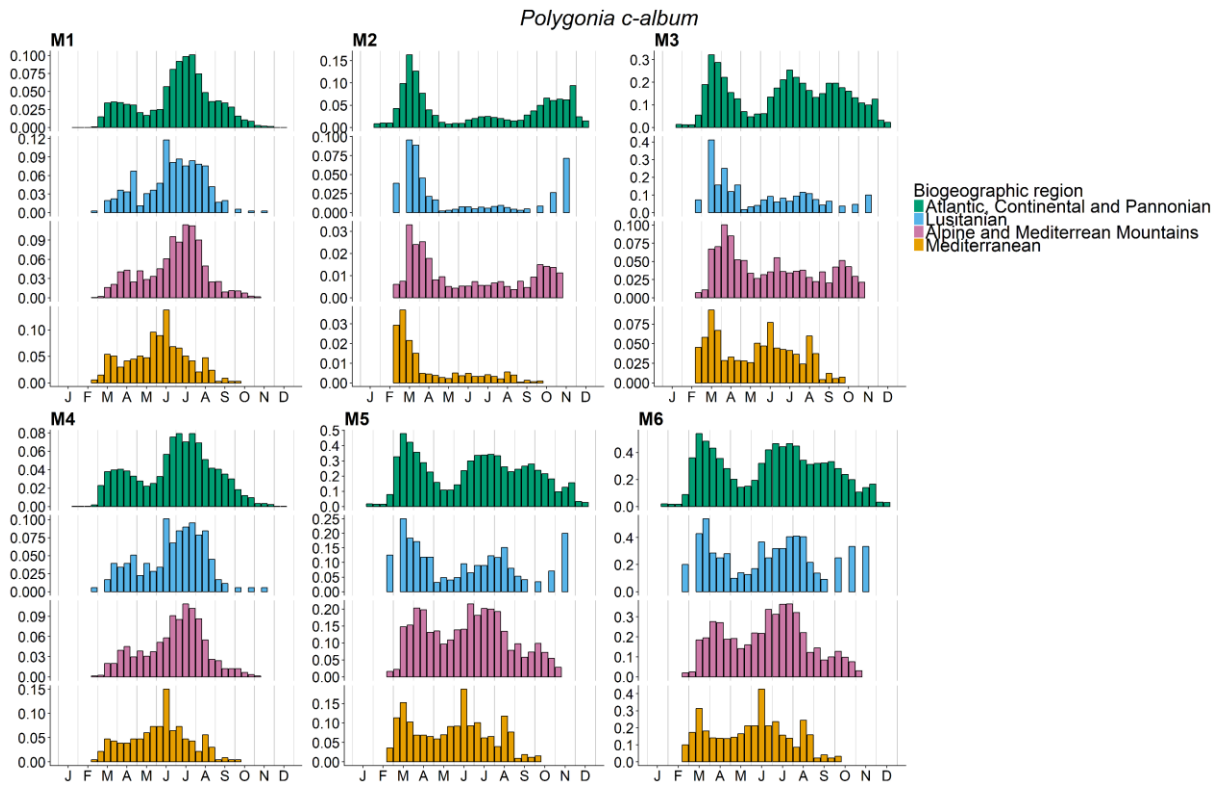


678

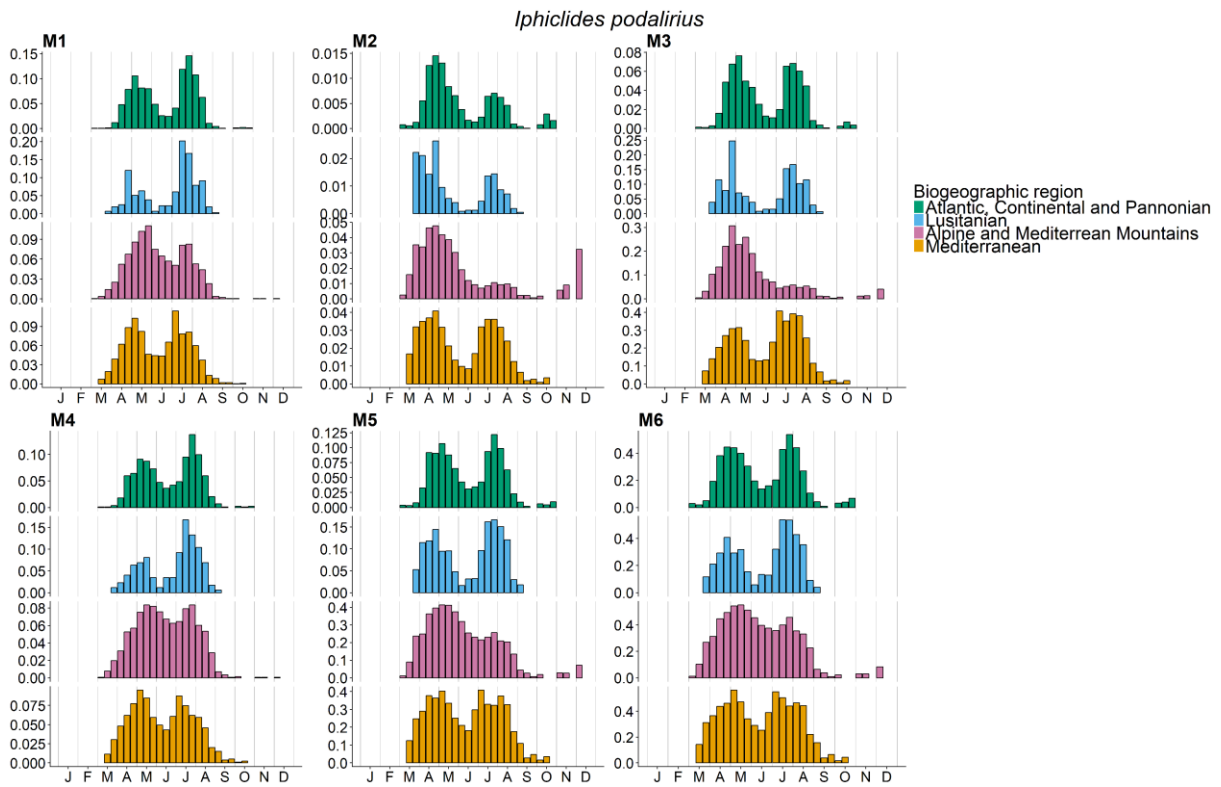
Nymphalis antiopa



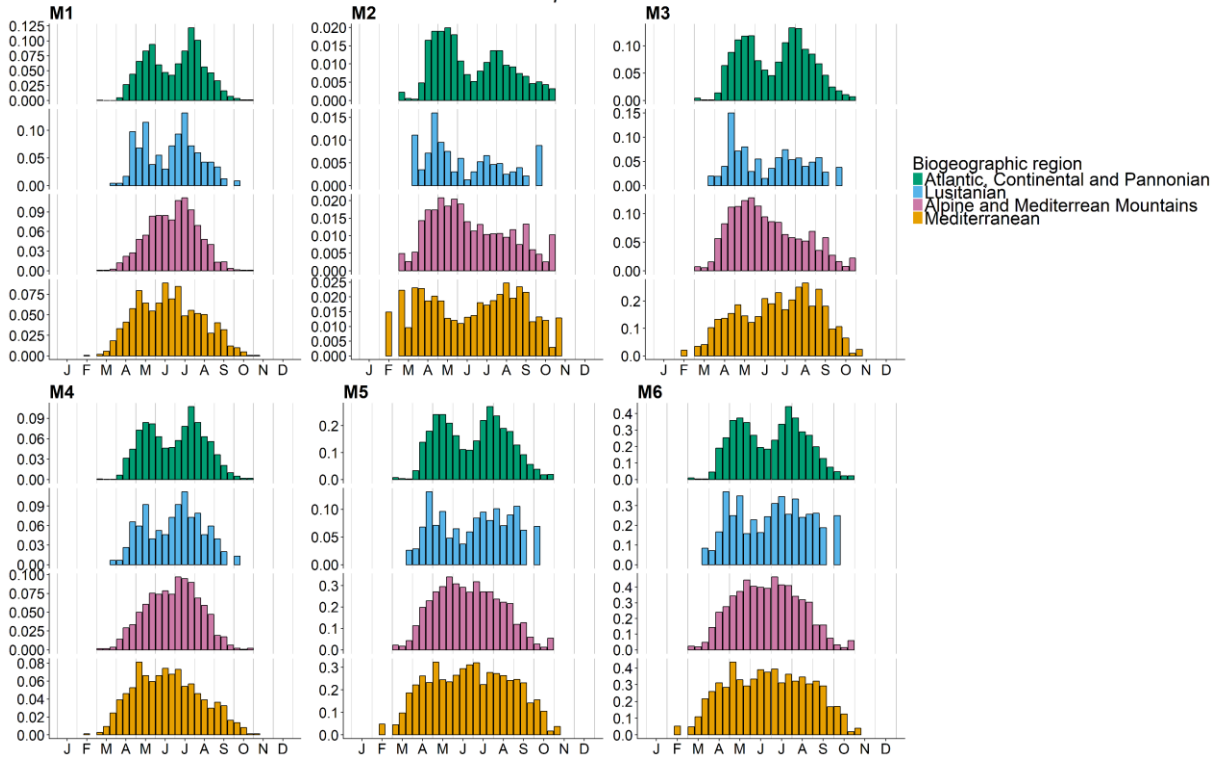
679



681 Papilionidae



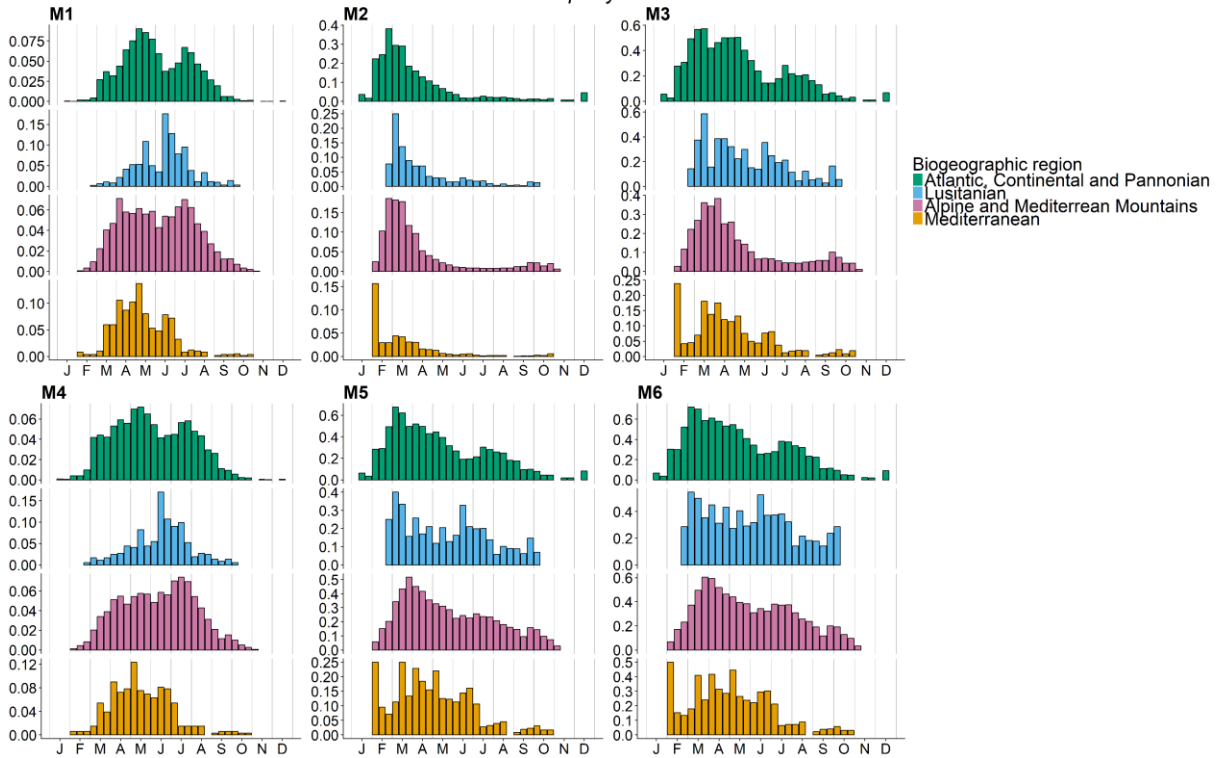
Papilio machaon



683

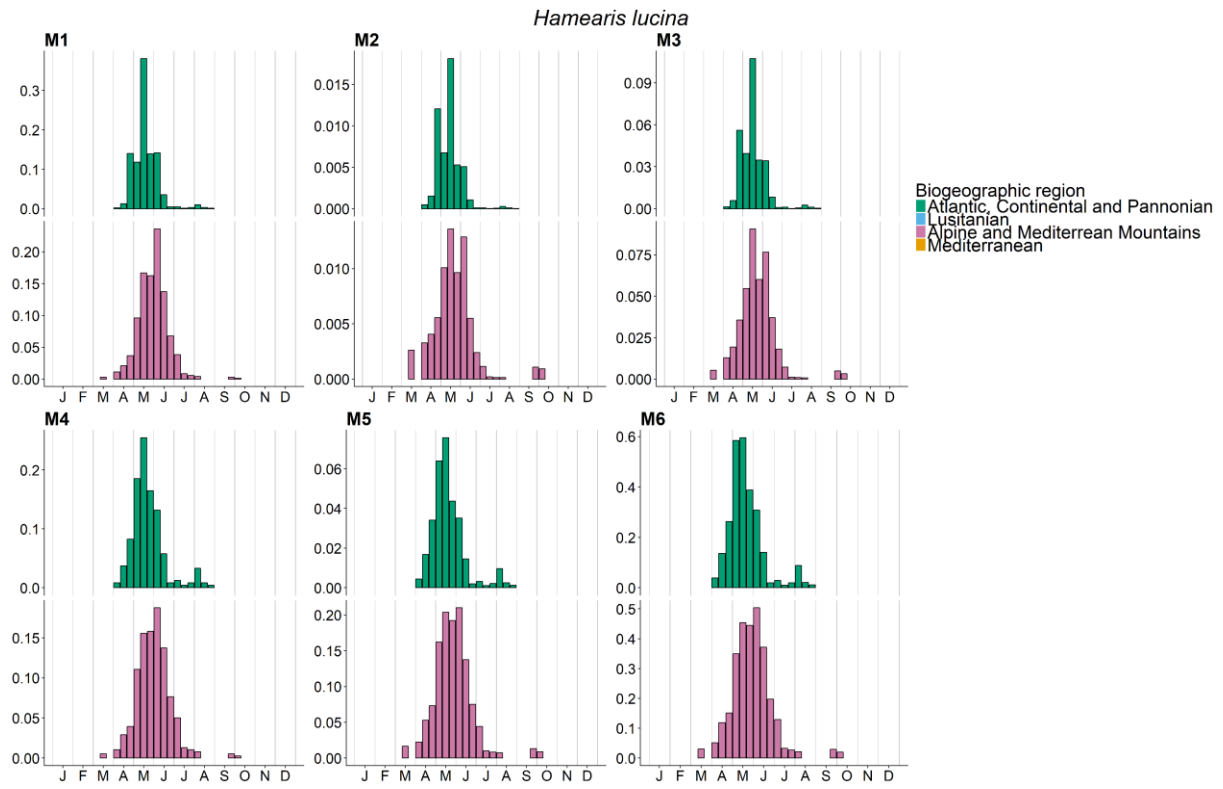
684 Pieridae:

Gonepteryx rhamni



685

686 Riodinidae:



687

688

Appendix 5. Comparison between the STERF and six phenology indices computed on INPN data by ten-day periods for 57 species. A

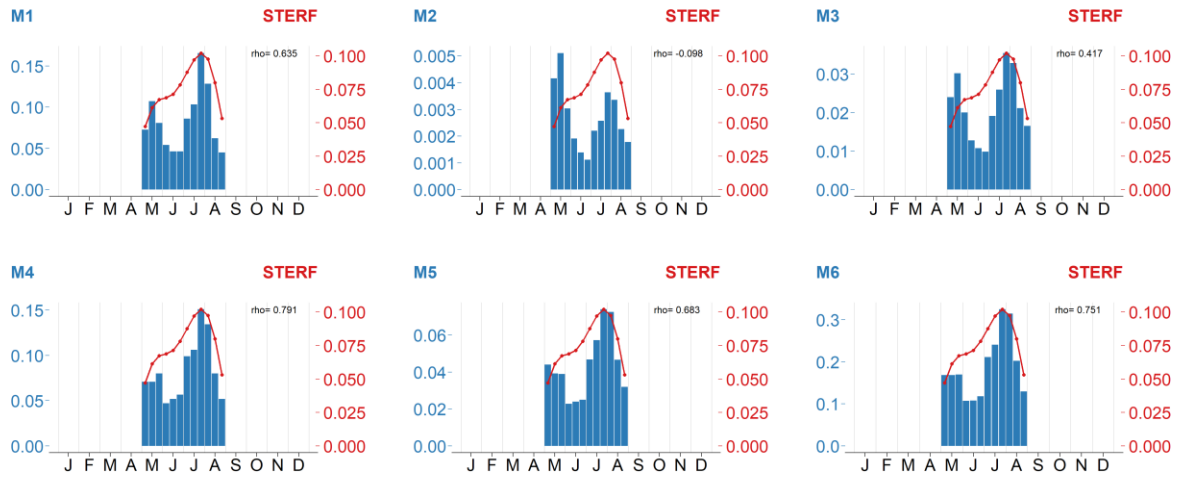
689

Pearson's correlation coefficient (ρ) was calculated between every index (blue bar plots) and STERF count estimates (red line). Indices and

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

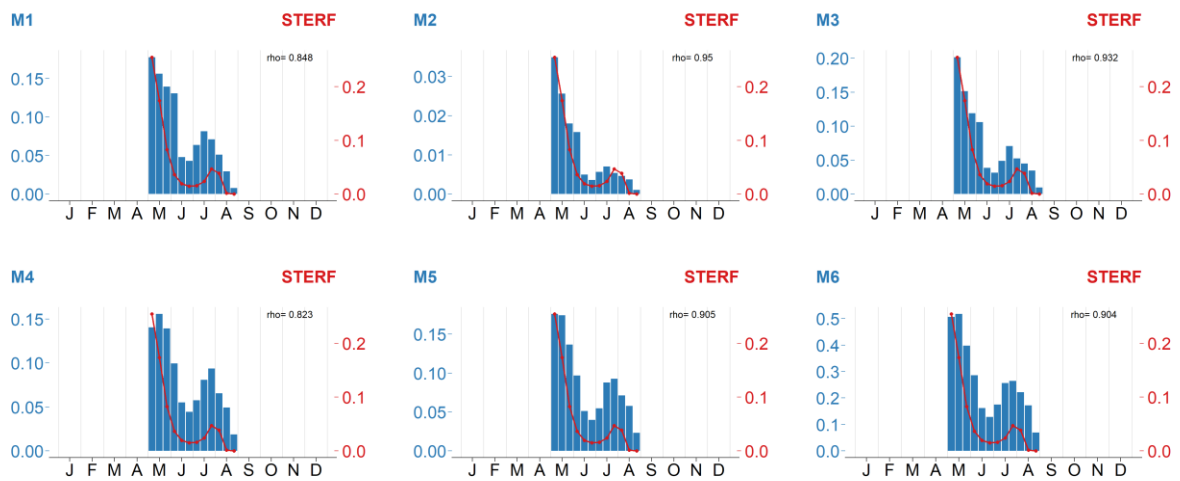
692 Hesperidae:

Carcharodus alceae



693

Erynnis tages



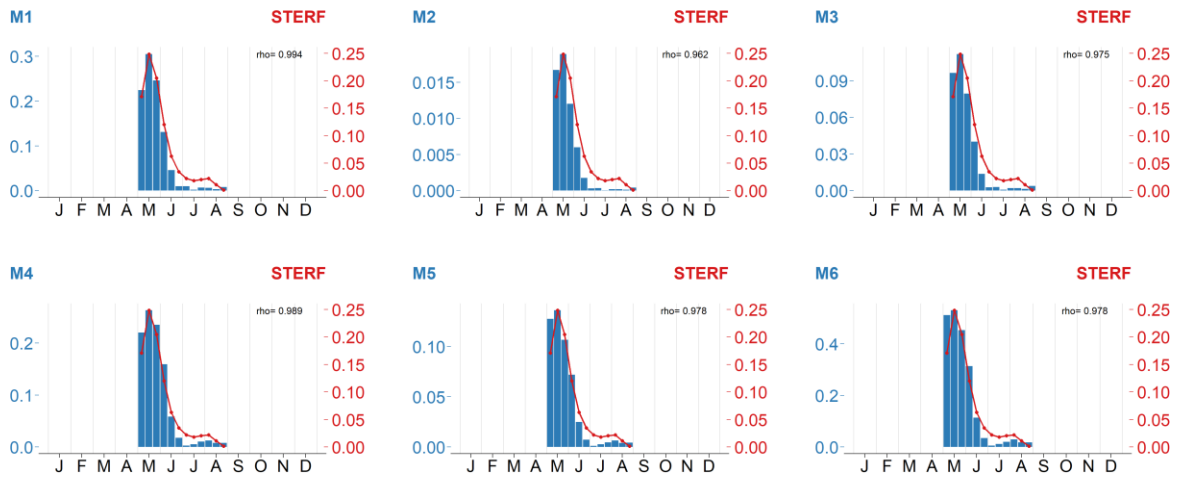
694

Ochlodes sylvanus



695

Pyrgus malvae



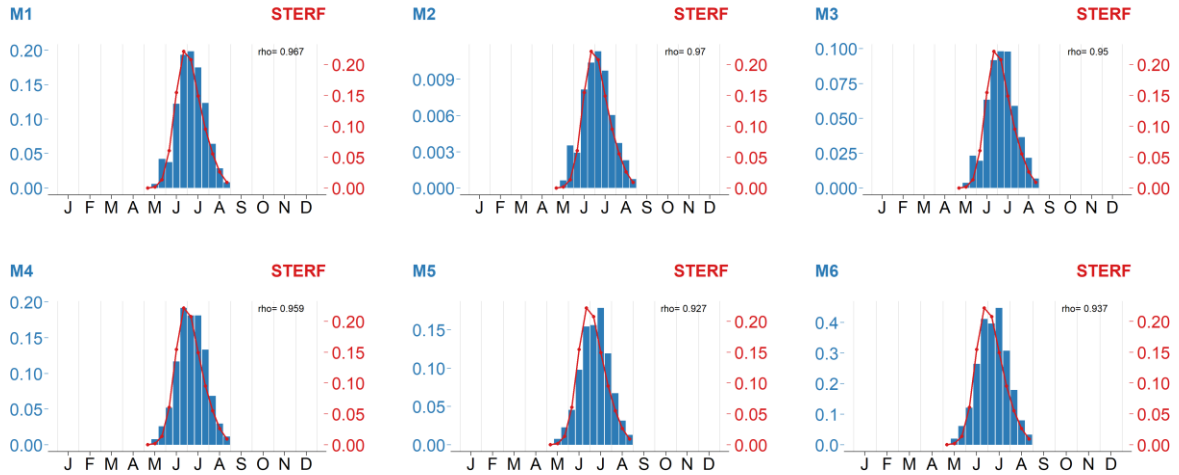
696

Thymelicus lineola



697

Thymelicus sylvestris



698

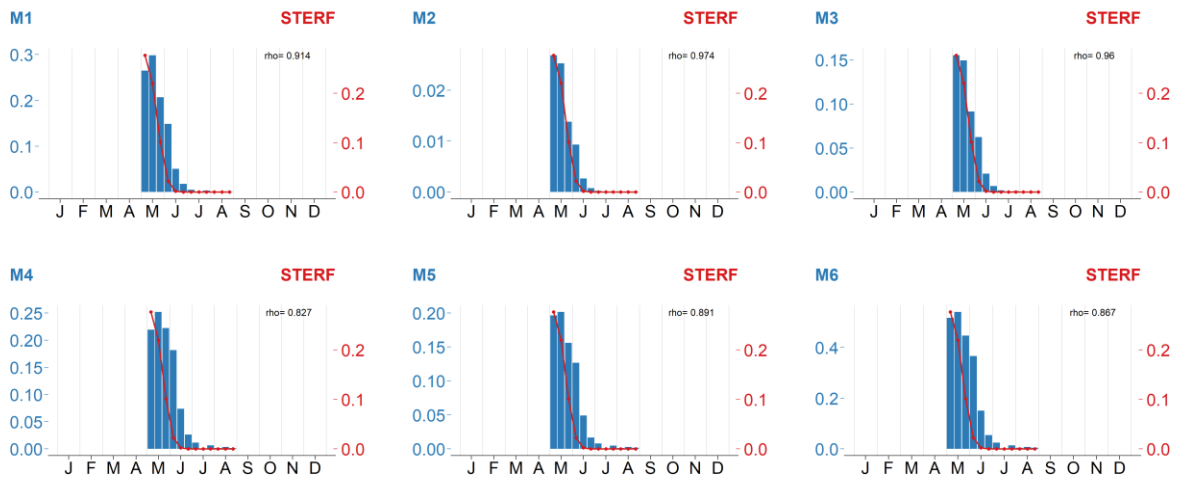
699 Lycaenidae:

Aricia agestis



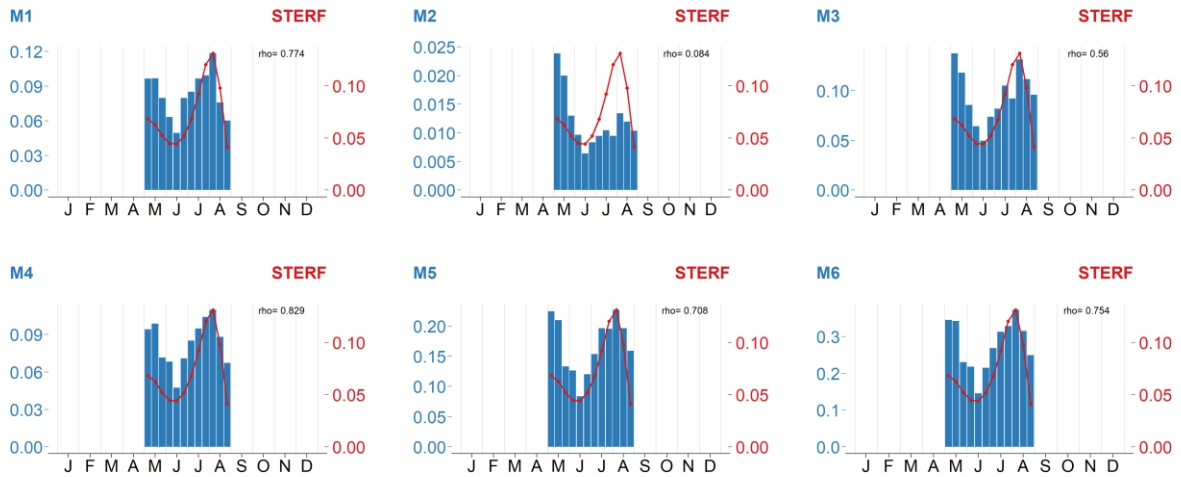
700

Callophrys rubi



701

Celastrina argiolus



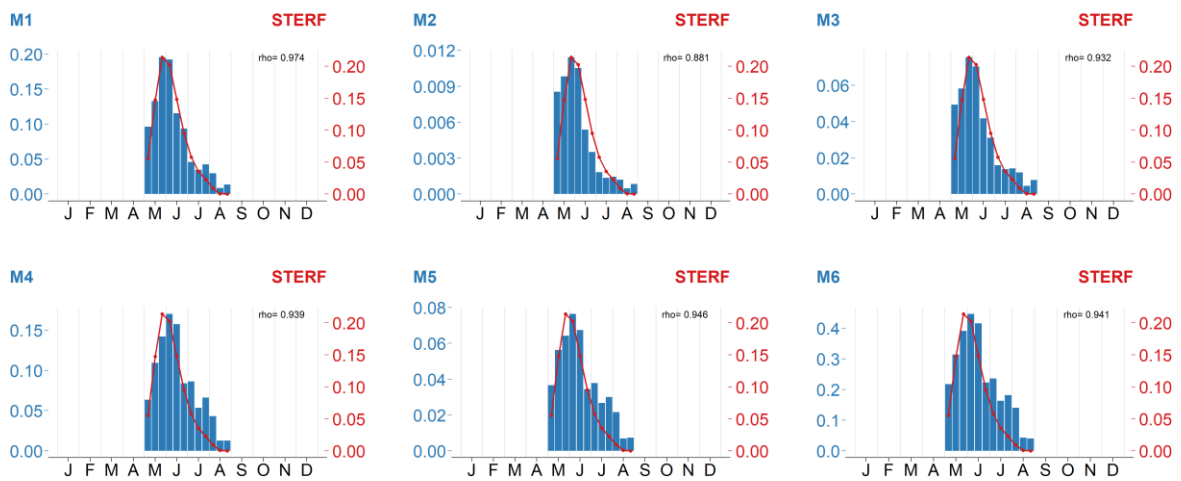
702

Cupido argiades



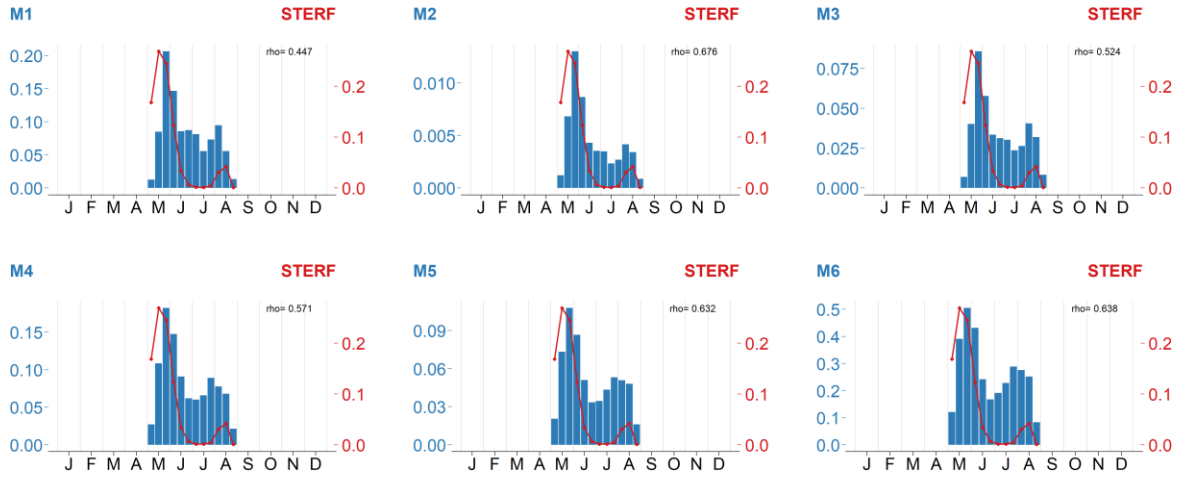
703

Cupido minimus



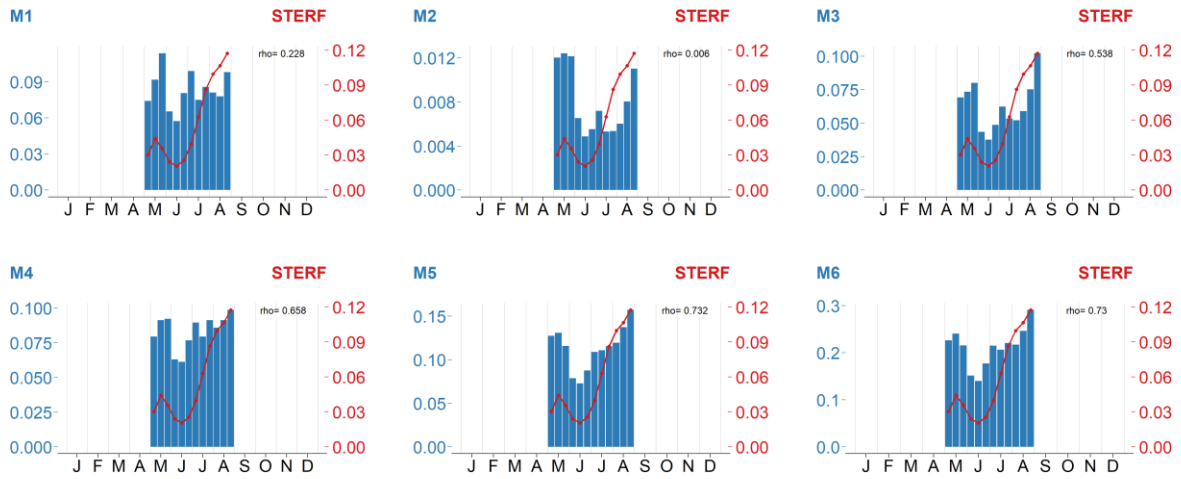
704

Cyaniris semiargus



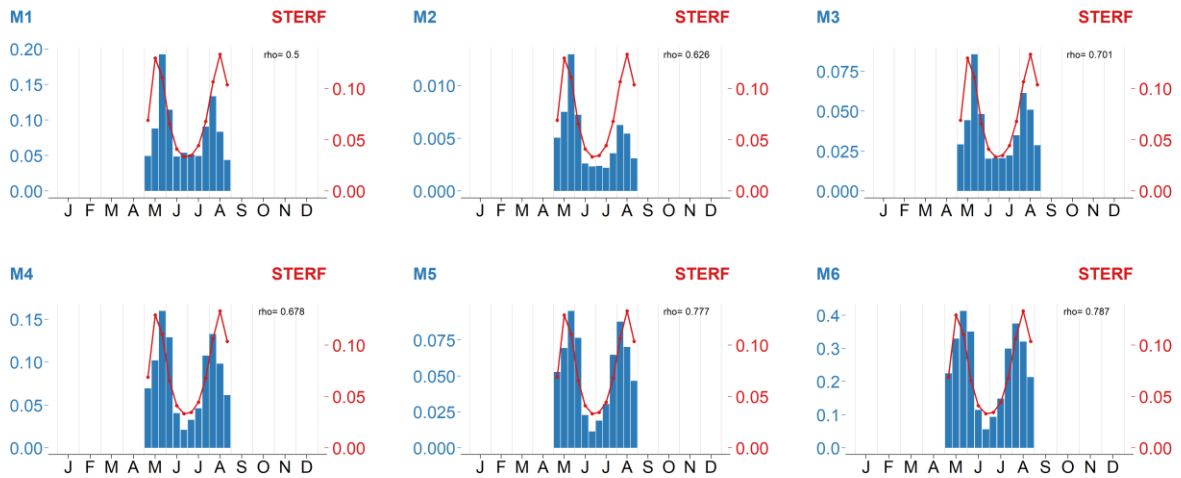
705

Lycaena phlaeas



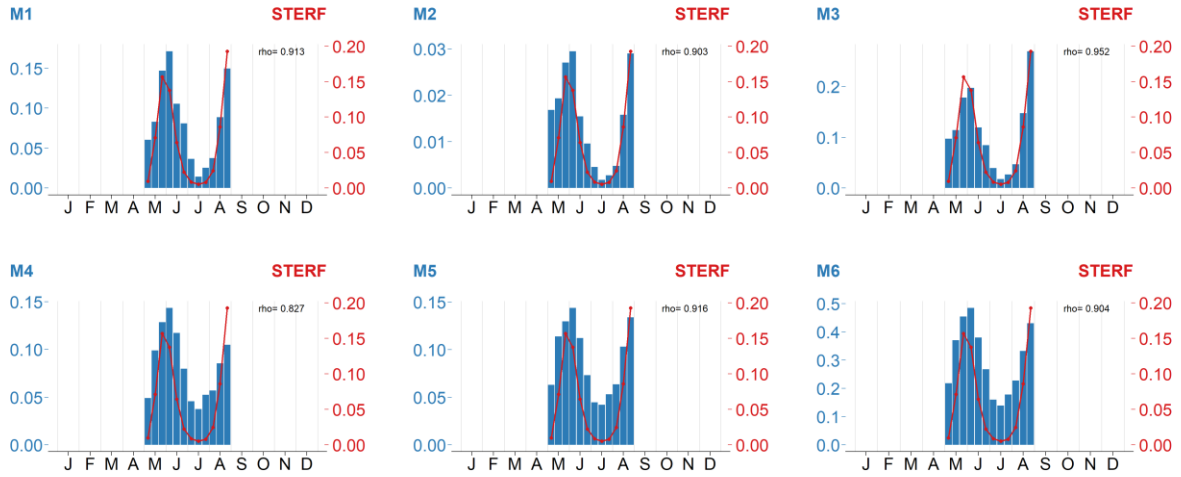
706

Lycaena tityrus



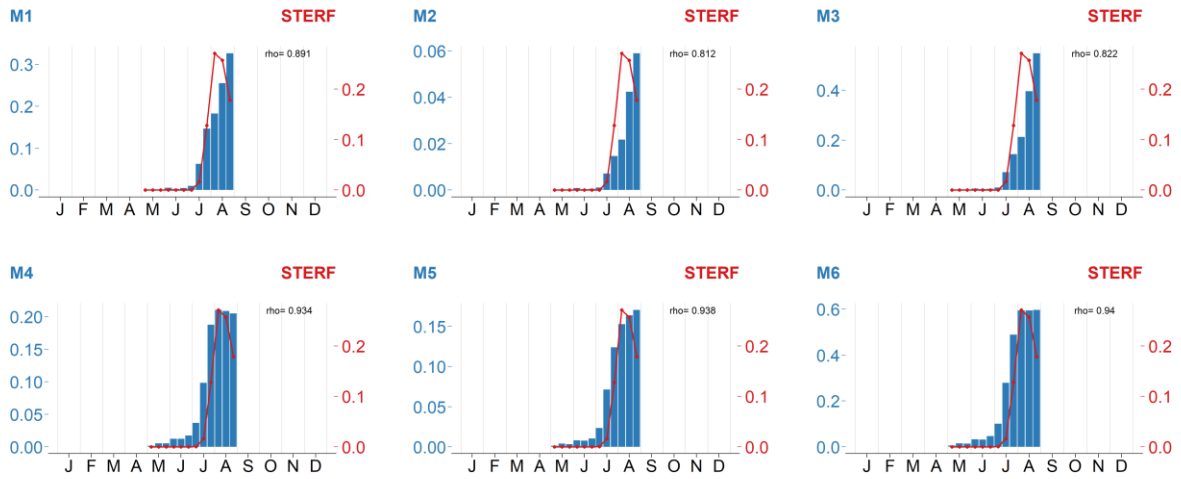
707

Lysandra bellargus



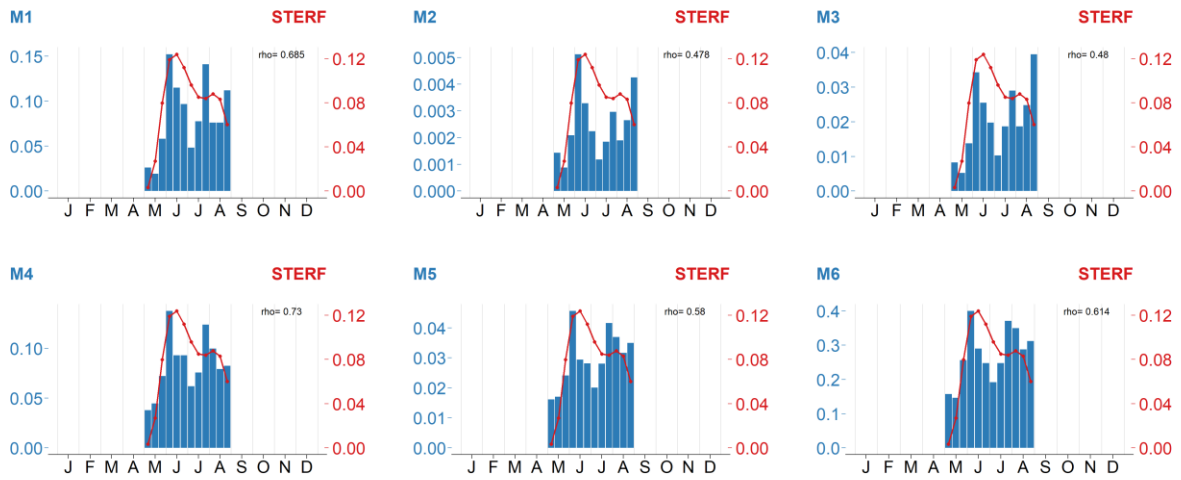
708

Lysandra coridon



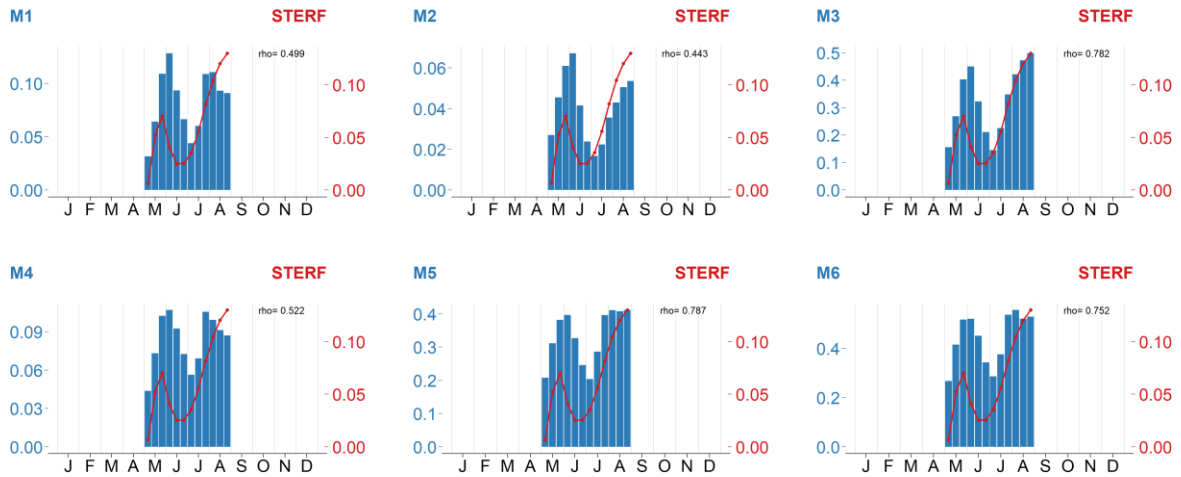
709

Plebejus argyrognomon



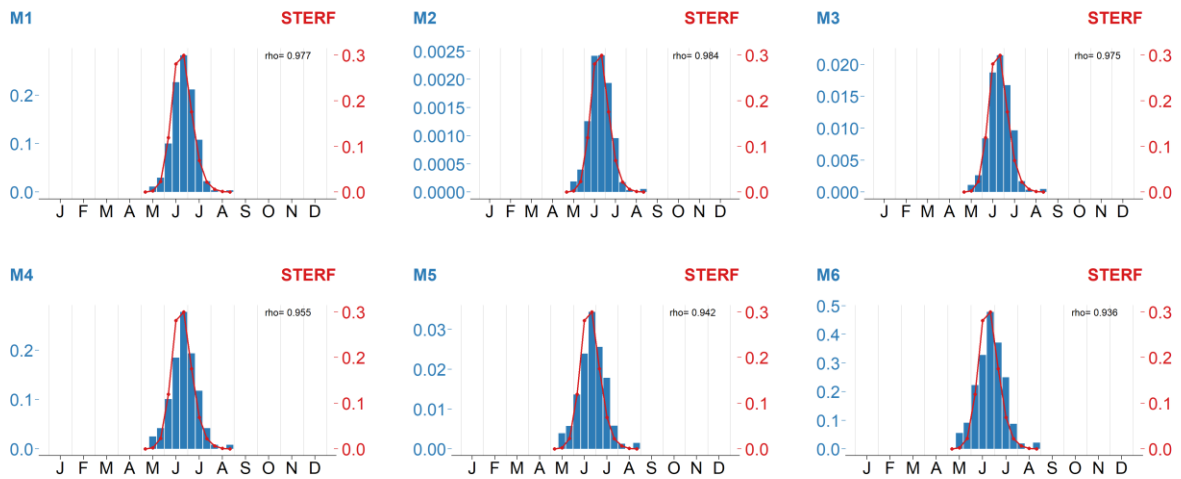
710

Polyommatus icarus



711

Satyrium ilicis



712

713 Nymphalidae:

Aglais io



714

Aglais urticae



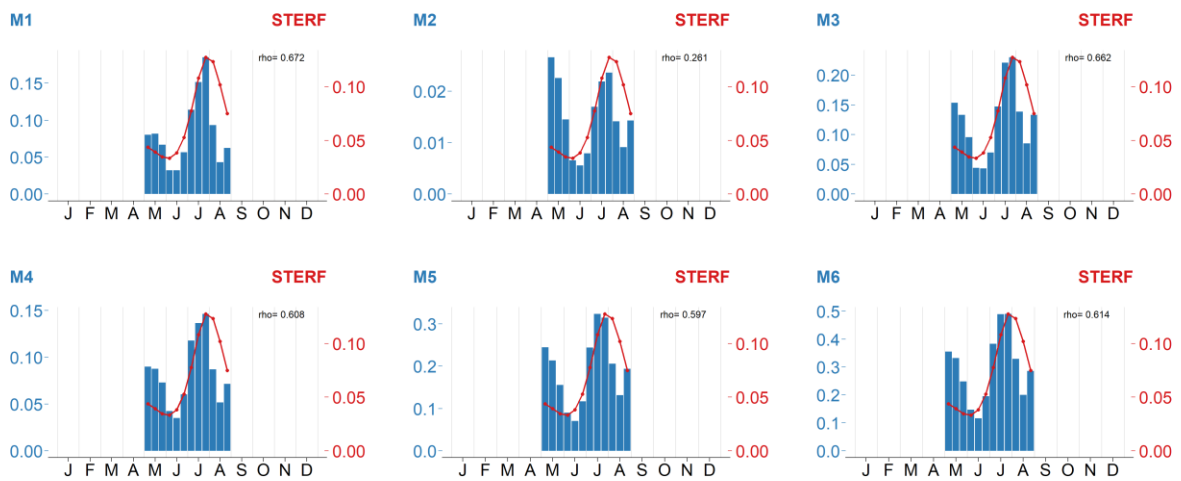
715

Aphantopus hyperantus



716

Araschnia levana



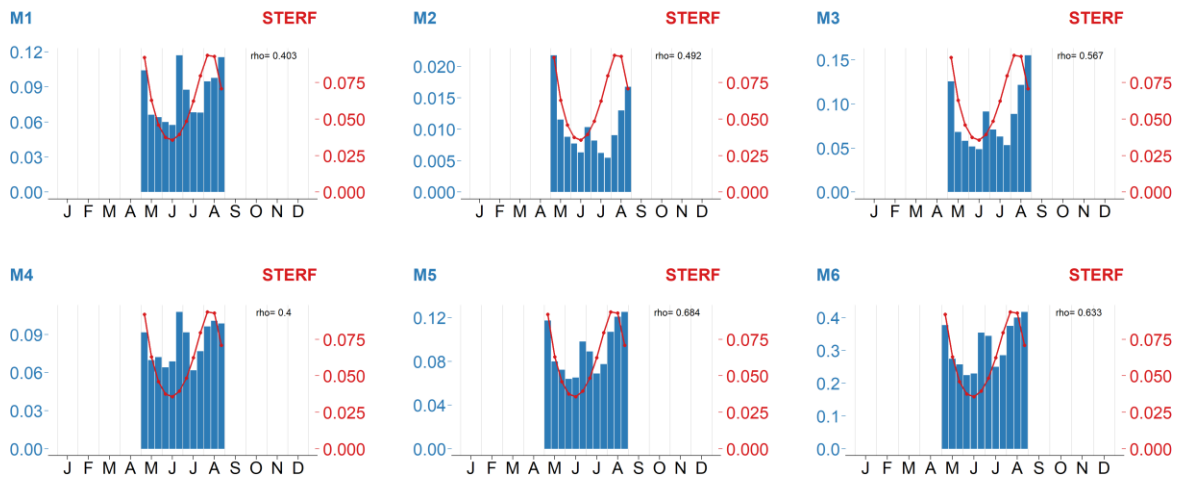
717

Argynnis paphia



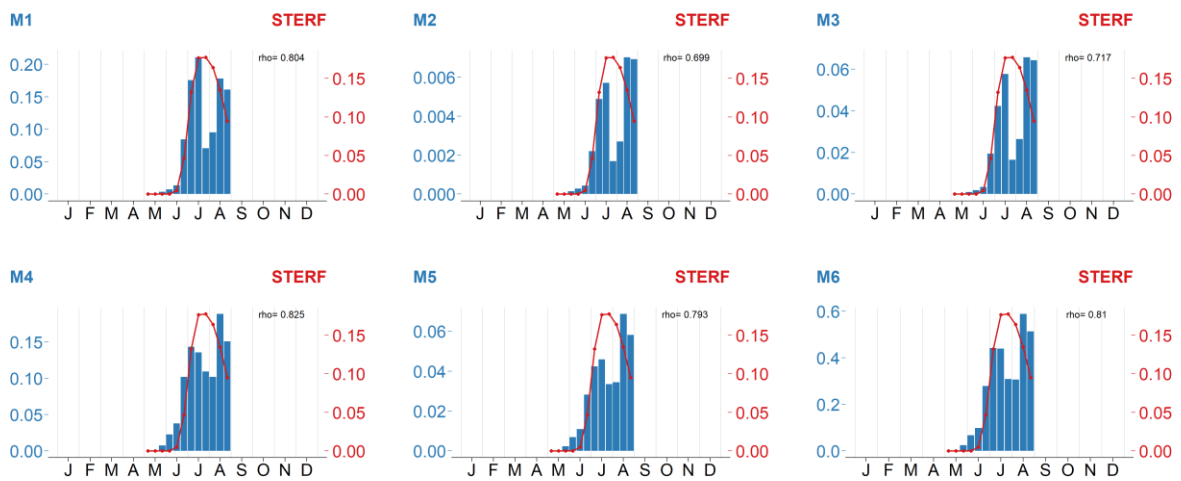
718

Boloria dia



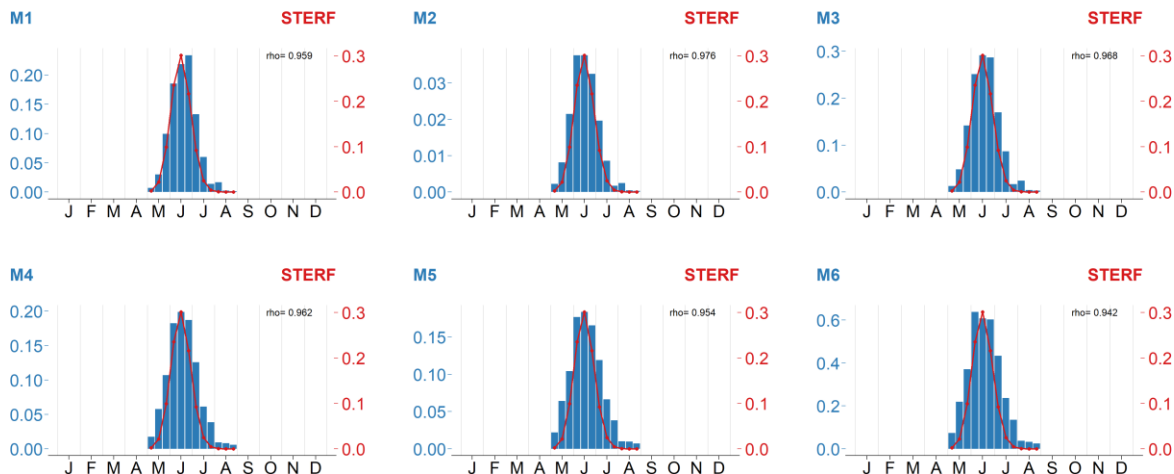
719

Brintesia circe



720

Coenonympha arcania



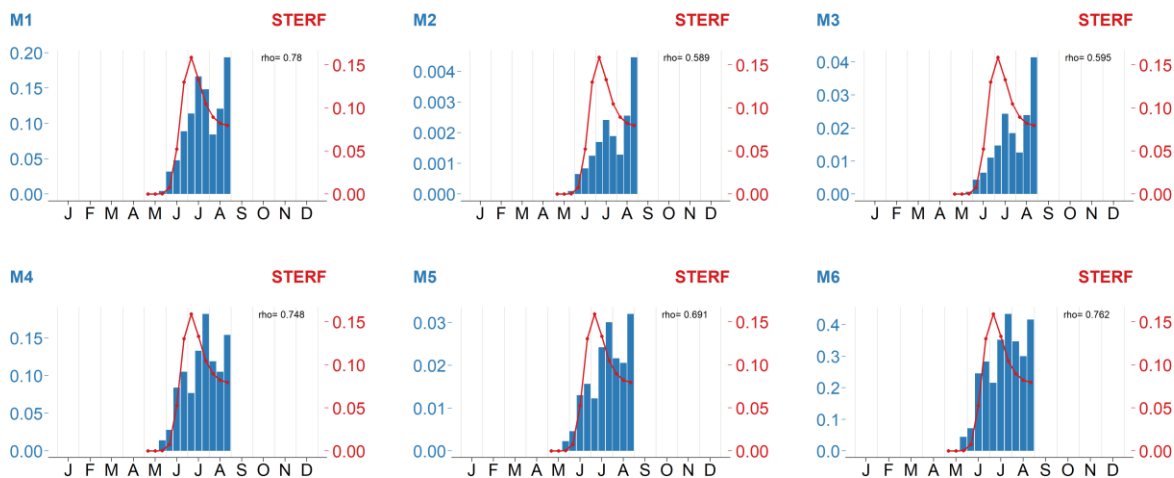
721

Coenonympha pamphilus



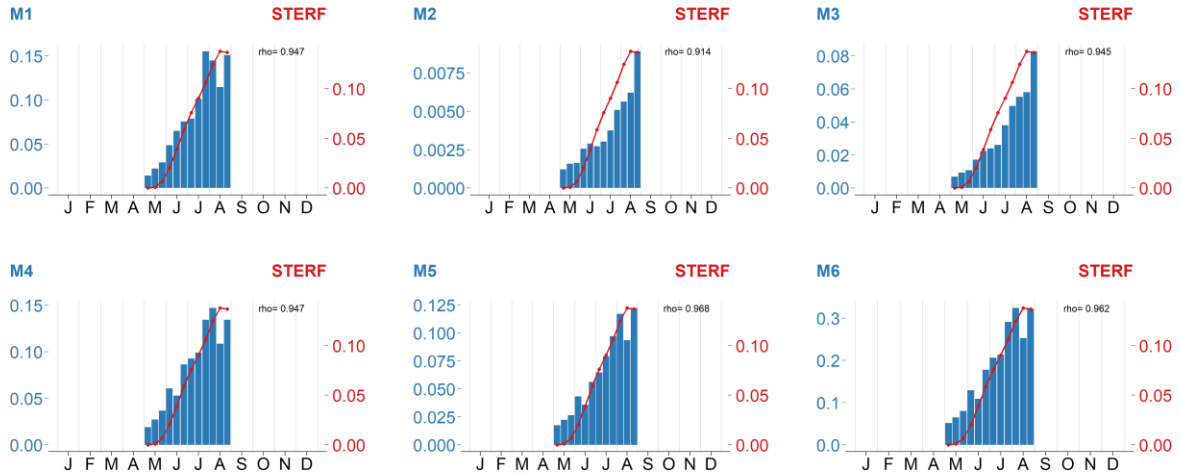
722

Fabriciana adippe



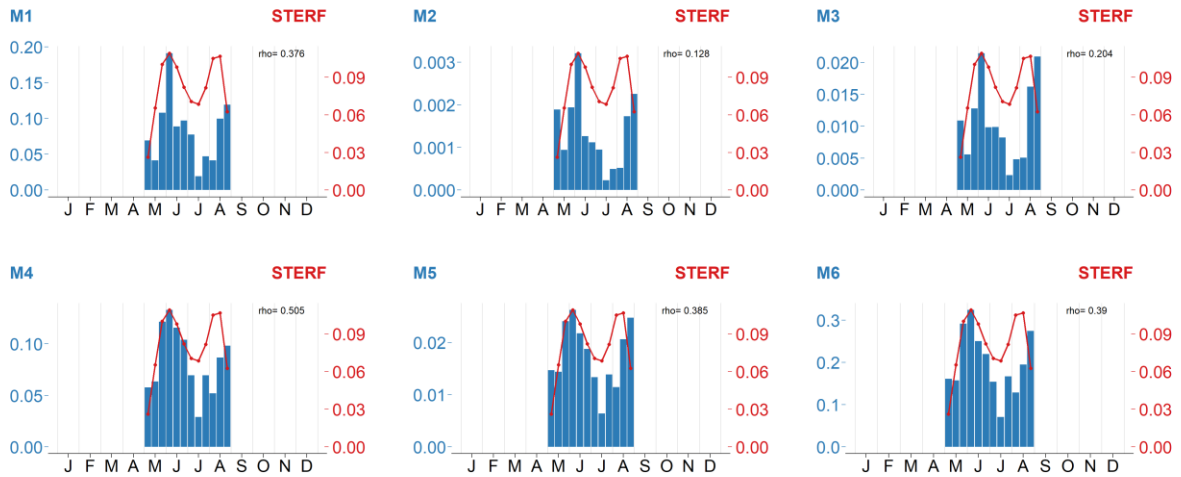
723

Issoria lathonia



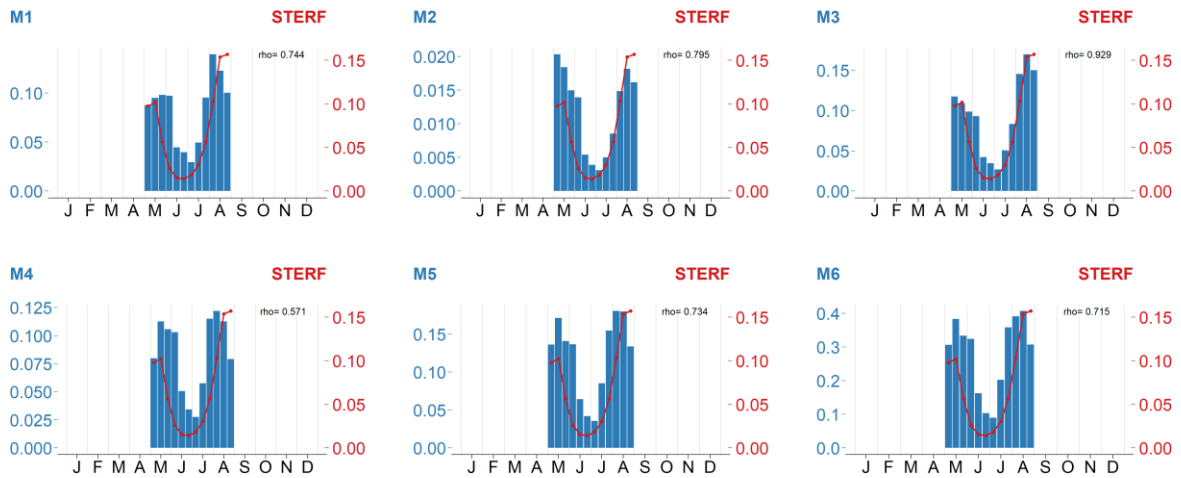
724

Lasiommata maera



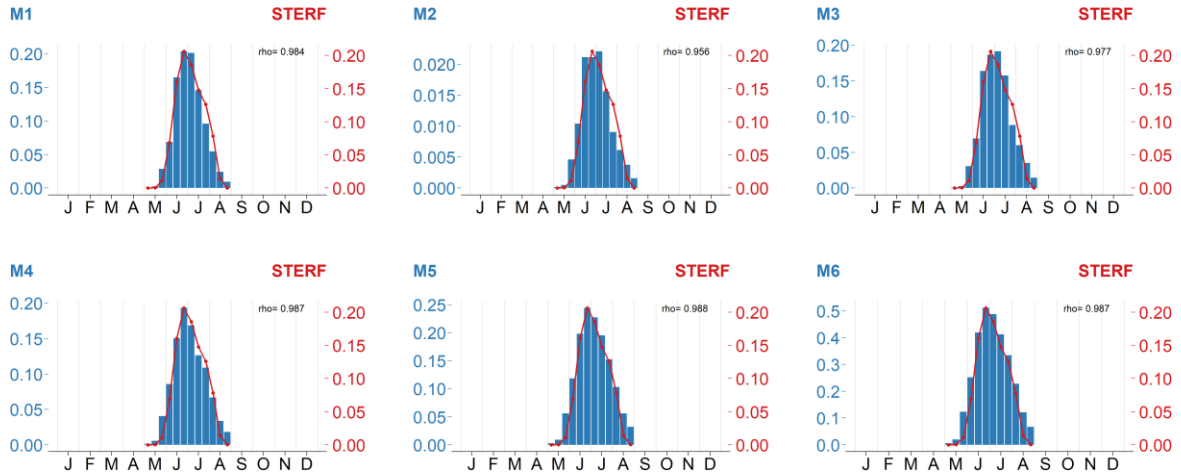
725

Lasiommata megera



726

Limenitis camilla



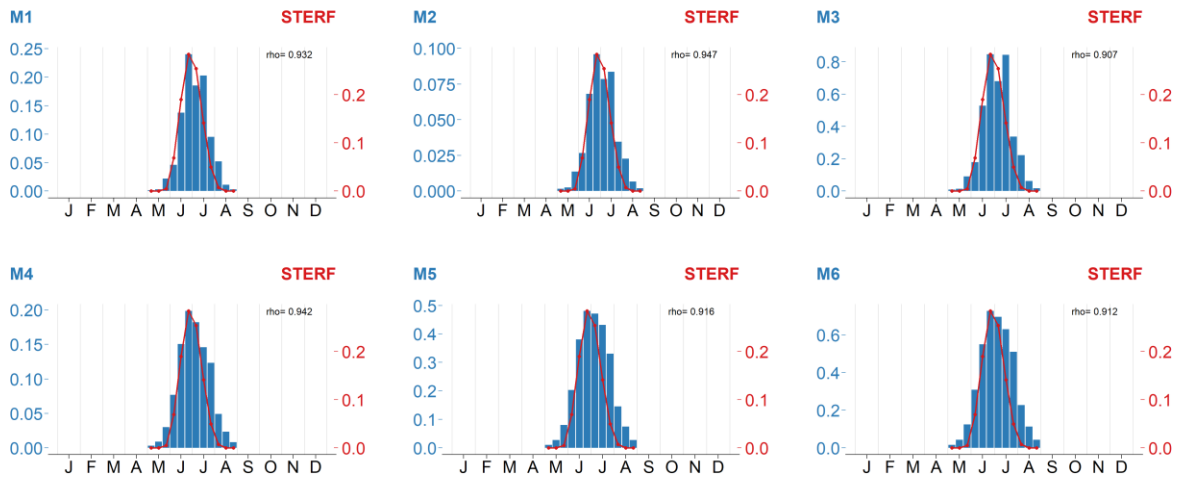
727

Maniola jurtina



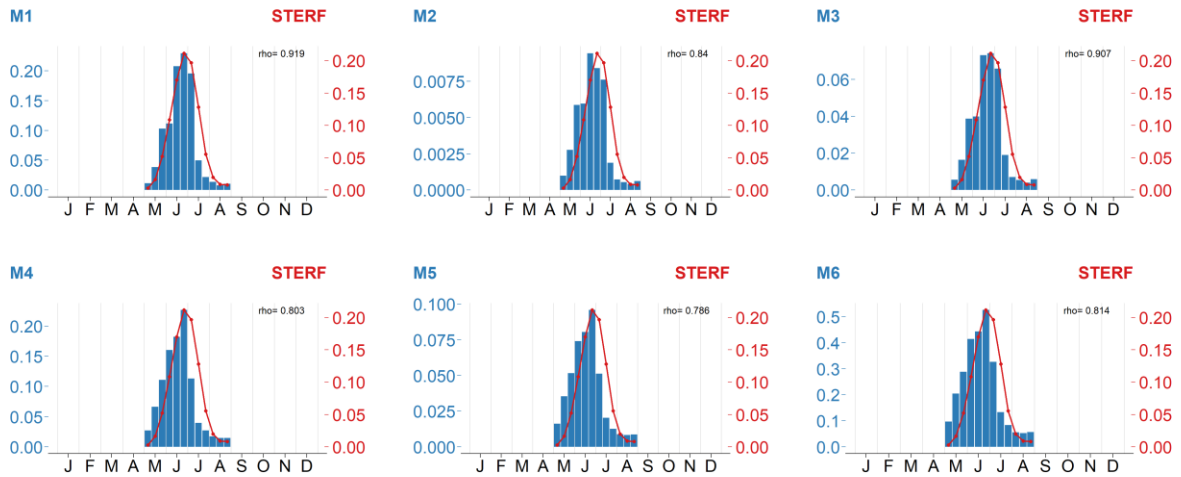
728

Melanargia galathea



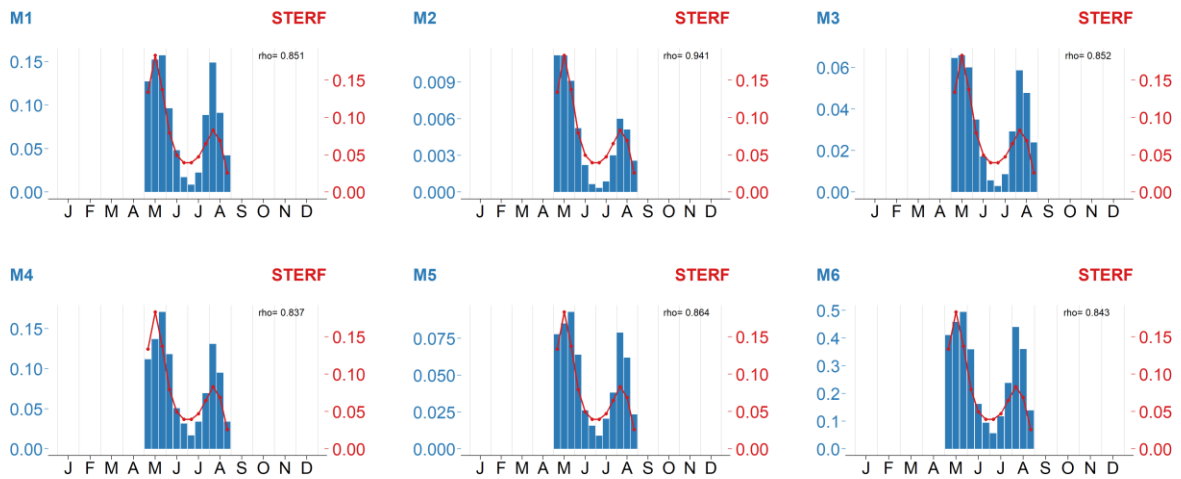
729

Melitaea athalia



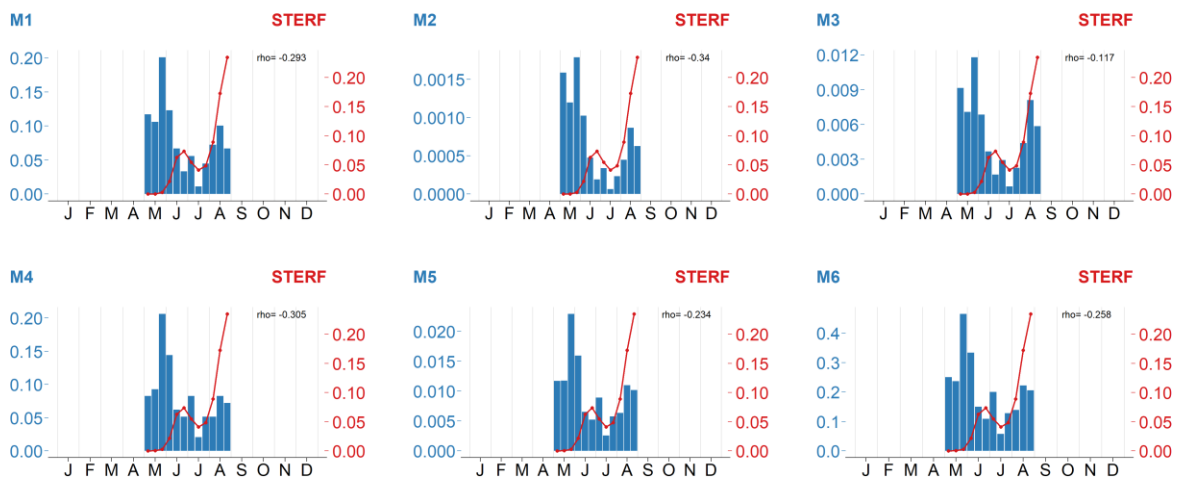
730

Melitaea cinxia



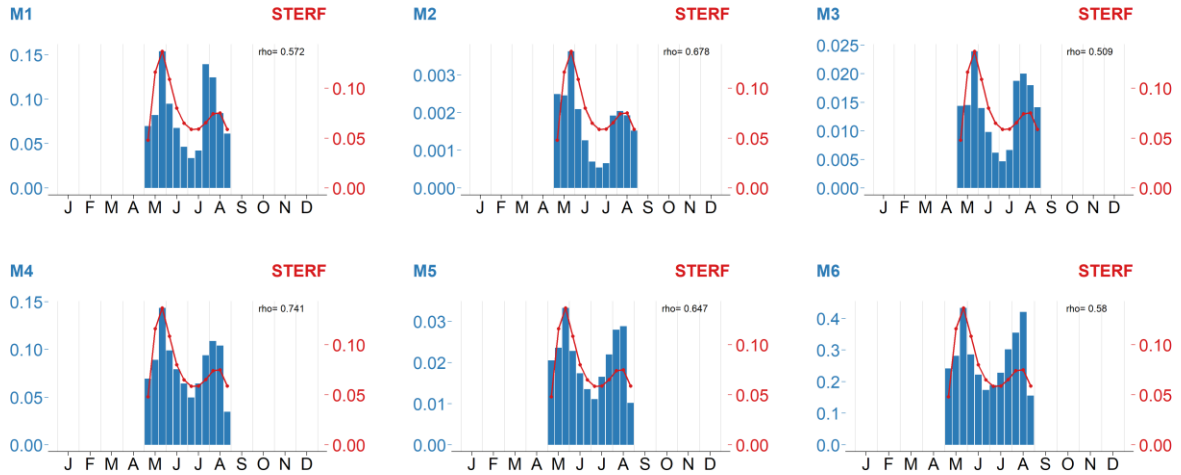
731

Melitaea parthenoides



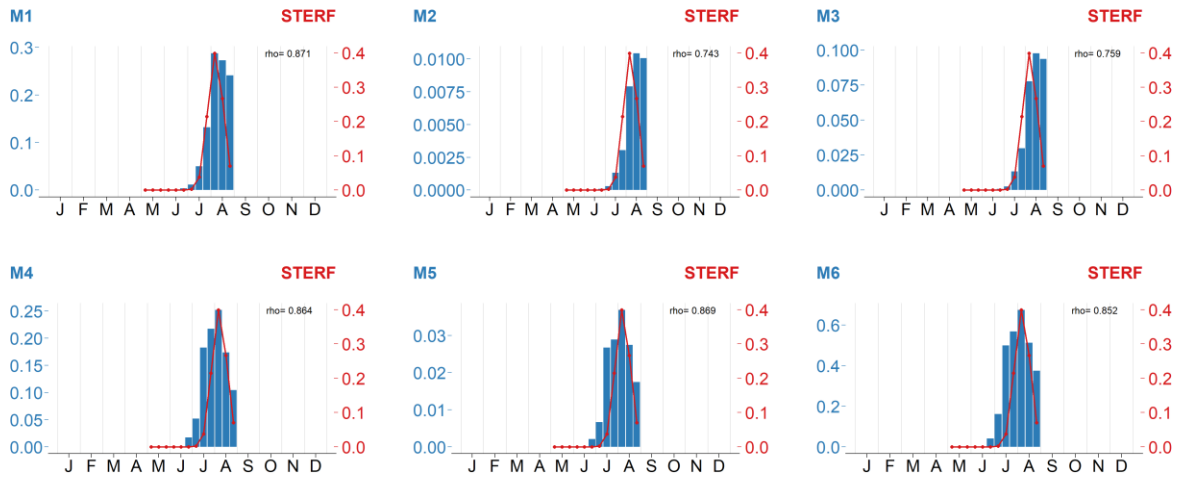
732

Melitaea phoebe



733

Minois dryas



734

Pararge aegeria



735

Polygonia c-album



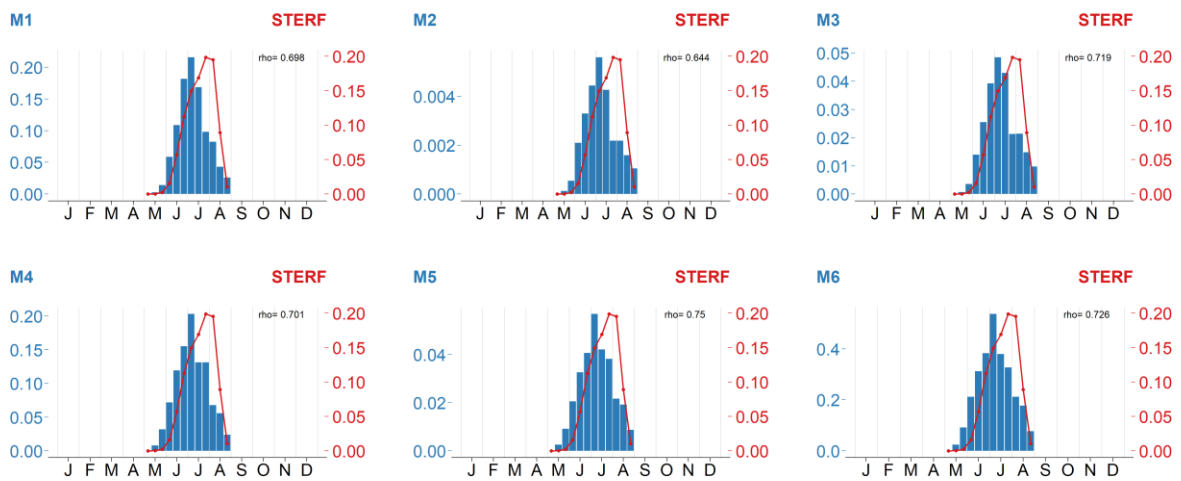
736

Pyronia tithonus



737

Speyeria aglaja



738

Vanessa atalanta



739

Vanessa cardui

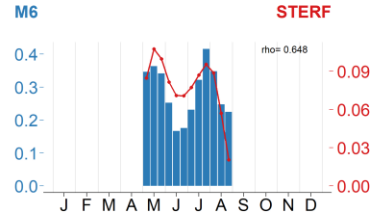
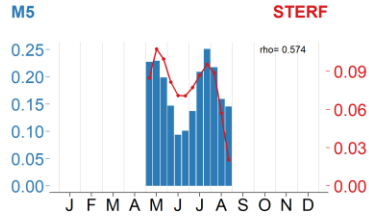
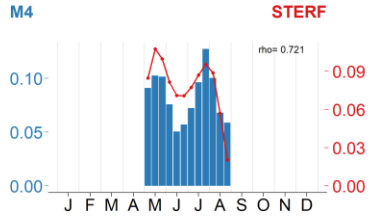
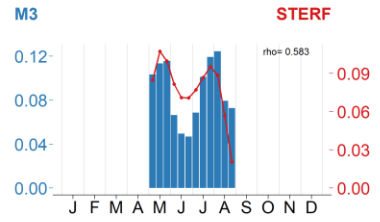
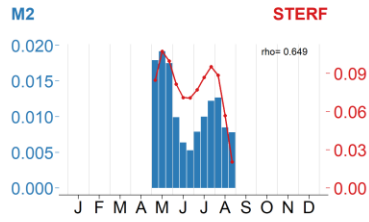
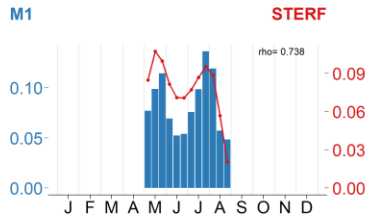


740

741 Papilionidae:

742

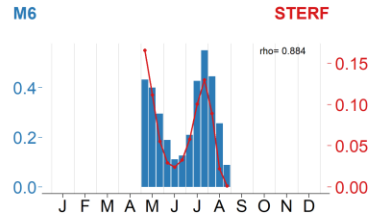
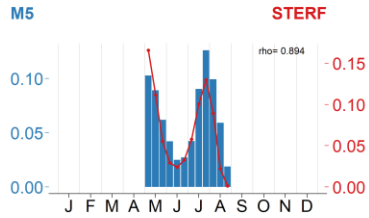
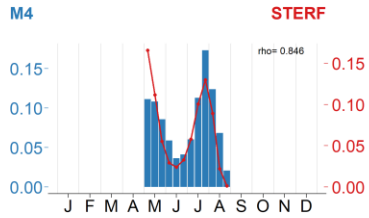
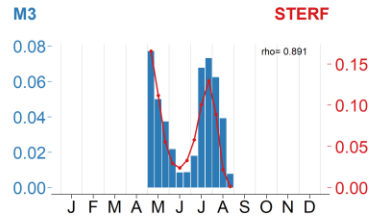
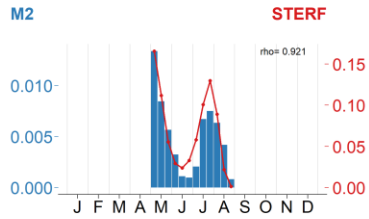
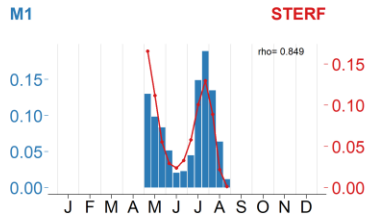
Papilio machaon



743

744

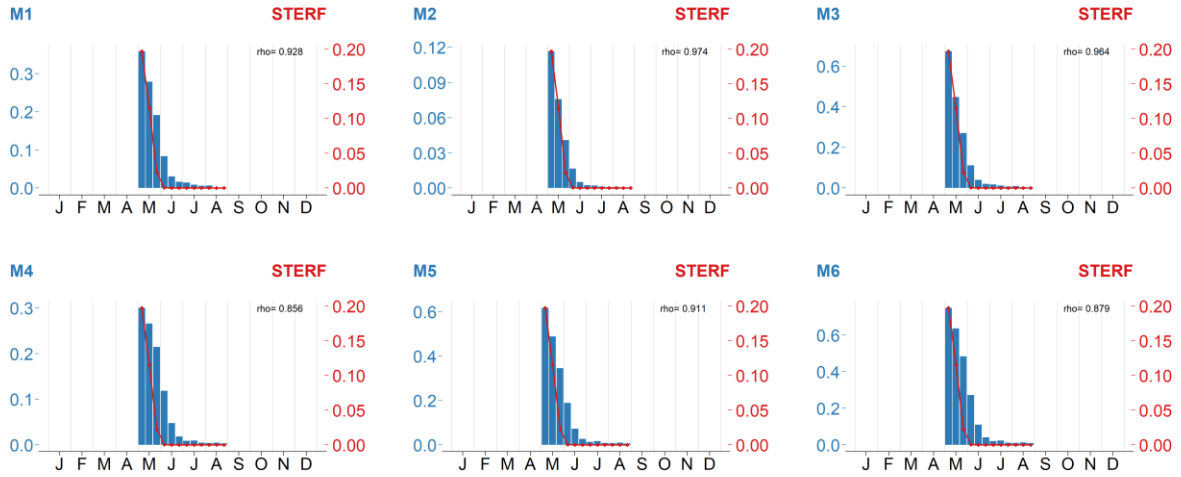
Iphiclides podalirius



745

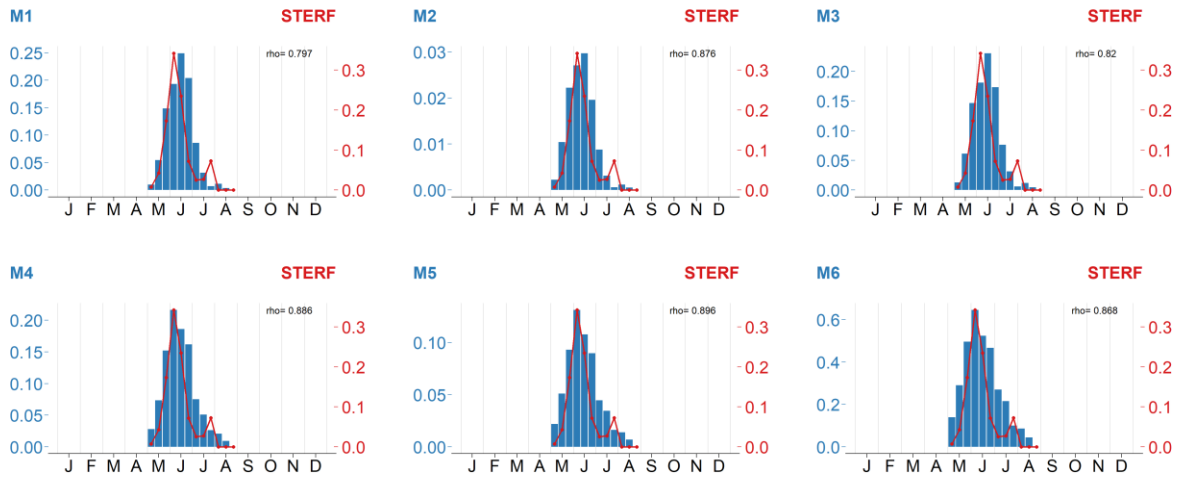
746 Pieridae:

Anthocharis cardamines



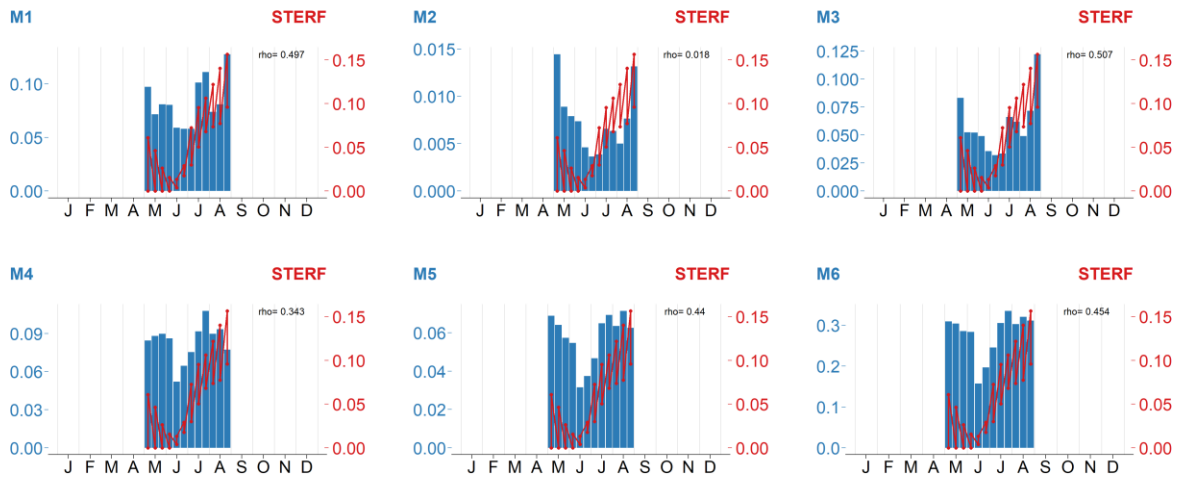
747

Aporia crataegi



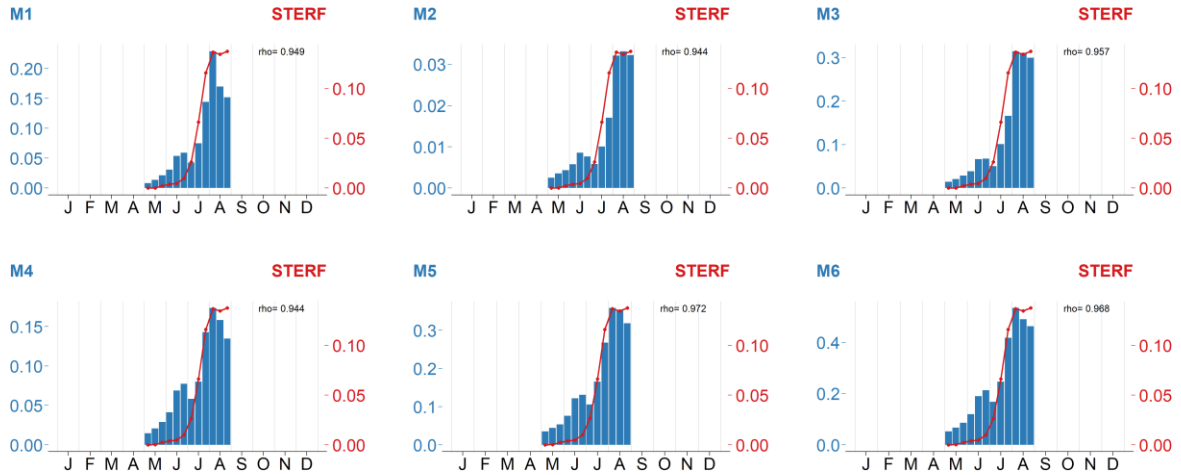
748

Colias alfacariensis



749

Colias crocea



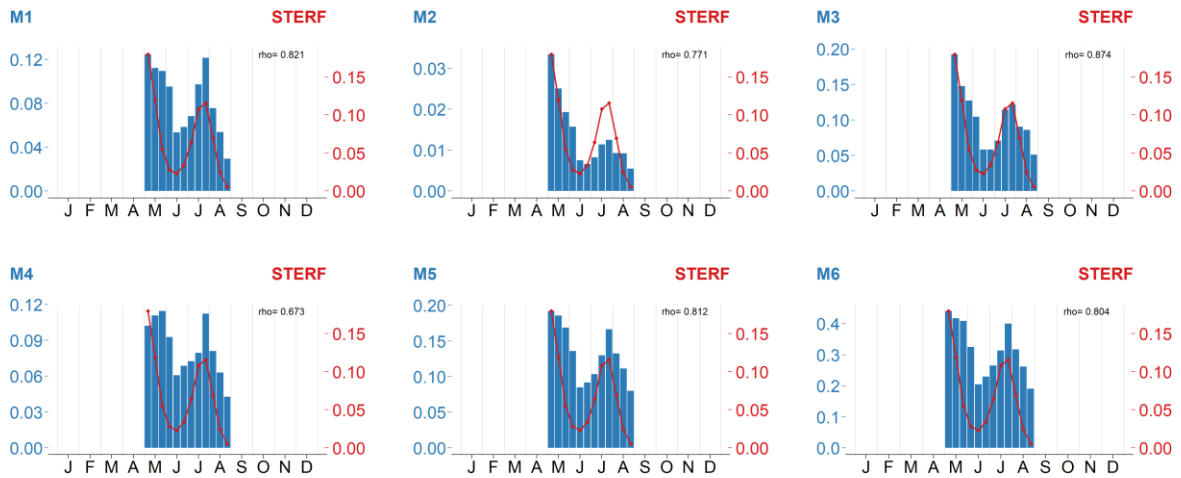
750

Gonepteryx rhamni



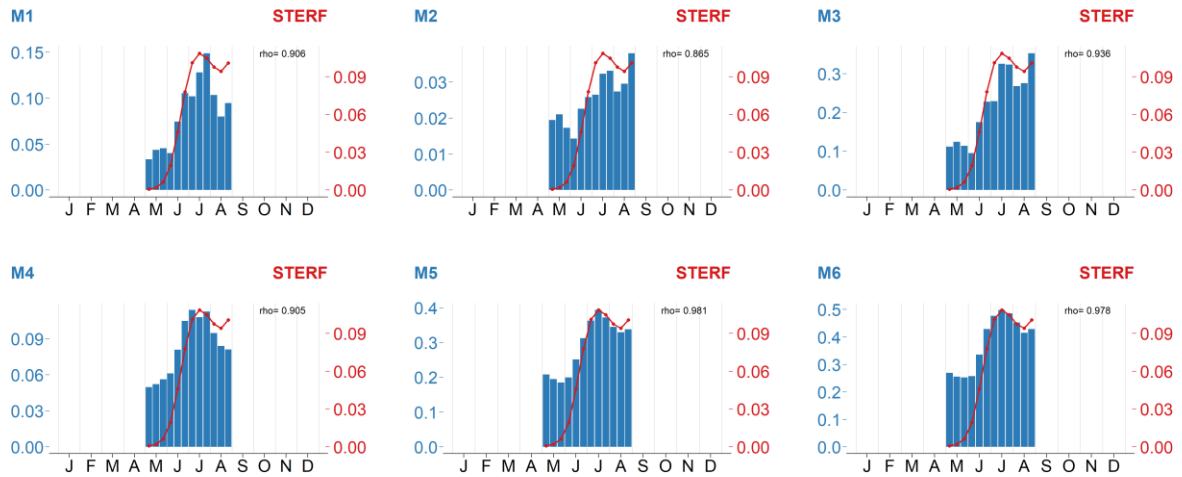
751

Leptidea sinapis



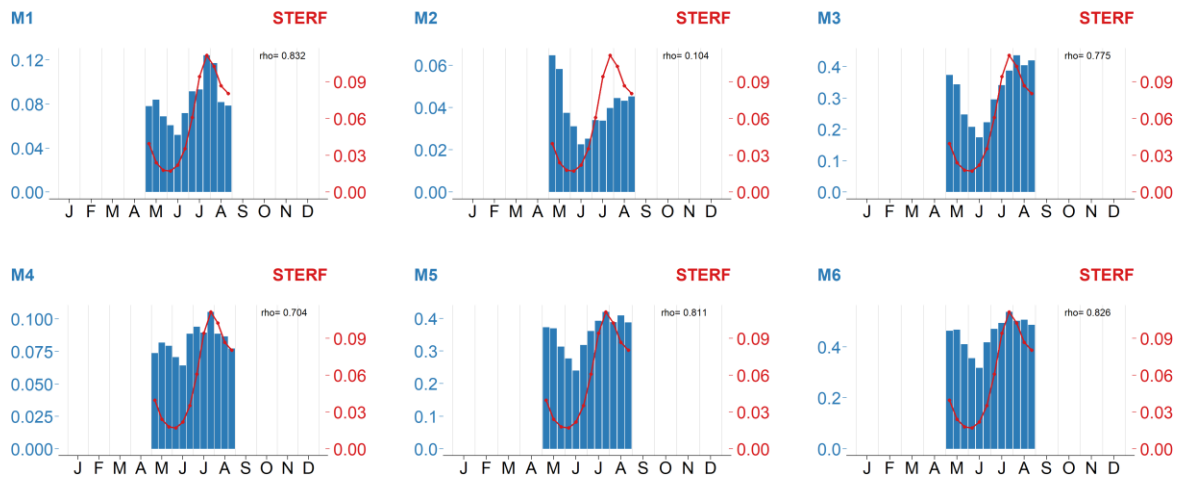
752

Pieris brassicae



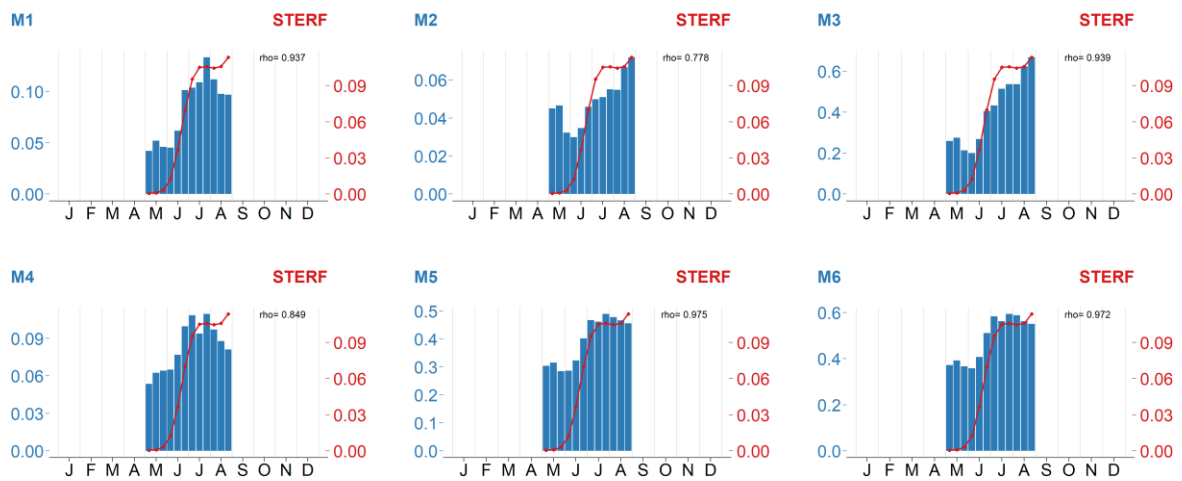
753

Pieris napi



754

Pieris rapae



755

756