

ORIGINAL ARTICLE

Editors' Choice

Insect Conservation
and Diversity

Systematic differences in phenology estimates from unstructured and structured biodiversity datasets

Mingrui Li¹ | Jonathan Bennie² | David B. Roy^{1,3} | Richard Fox^{1,4} | Robin J. Boyd^{1,3} | Chloë Smith⁴ | Richard H. Ffrench-Constant^{1†}

¹Centre for Ecology and Conservation,
University of Exeter, Cornwall, UK

²Centre for Geography and Environmental
Science, University of Exeter, Cornwall, UK

³UK Centre for Ecology and Hydrology,
Oxfordshire, UK

⁴Butterfly Conservation, Dorset, UK

Correspondence

Mingrui Li, University of Exeter, Centre for
Ecology and Conservation, Penryn Campus,
Penryn, Cornwall, TR10 9FE, UK.
Email: ml902@exeter.ac.uk

Funding information

No Funding Received

Editor: Christie A Bahlai and Associate Editor:
Jason Bried

Abstract

1. Phenology, the timing of biological events such as first flowering or emergence, plays a key role in understanding ecological responses to climate change. With the rapidly growing number of opportunistic citizen science records, it is important to understand whether and how phenological estimates differ between unstructured and structured recording approaches for animals and plants.
2. We compared phenological metrics derived from unstructured records in iRecord Butterflies with those from structured, transect-based monitoring in the UK Butterfly Monitoring Scheme (UKBMS).
3. Focusing on single-generation species, both data sources produce similar estimates of temporal trends in phenology, but unstructured data shows systematically earlier weighted mean flight dates (7.57 ± 1.57 days). It also reports earlier onsets (9.38 ± 1.24 days) and later ends (-8.46 ± 4.55 days) than UKBMS, resulting in longer inferred flight periods.
4. Results may reflect intrinsic differences in sampling between the two datasets, as well as recorder behaviour and motivations. Moreover, these differences are independent of sample size, suggesting that iRecord Butterflies observers may systematically search for early and late flying adults. Recognising and accounting for phenology sampling differences will be important when using citizen science data for phenological inference in butterflies and other taxa.

KEYWORDS

biological records, citizen science, phenology, UK butterfly monitoring

INTRODUCTION

Phenology is the study of the timing of biological events such as first flowering or first appearance, and it plays a key role in understanding ecological responses to climate change (Debinski et al., 2025; Goded et al., 2024; Löckinger et al., 2024). Detecting phenological shifts can

require large quantities of long-term monitoring data (Bishop et al., 2013; Dennis et al., 2024). In recent years, application-based citizen science platforms such as iRecord Butterflies, eBird and iNaturalist have generated rapidly growing volumes of unstructured opportunistic observations (August et al., 2020; Bowler et al., 2022; Mason et al., 2025). Moreover, these datasets are increasingly used to estimate phenology at large spatial scales (Bishop et al., 2013; Capinha et al., 2024; Mason et al., 2025; Primack et al., 2023; Taylor

[†] Deceased.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2026 The Author(s). *Insect Conservation and Diversity* published by John Wiley & Sons Ltd on behalf of Royal Entomological Society.

et al., 2019). However, apparent differences in phenology derived from citizen science records may partly reflect variation in sampling and recorder behaviour rather than true biological change (Bishop et al., 2013; Habel et al., 2025; Rosário et al., 2025). This is because detections depend not only on the presence of individuals but also on when and where observers choose to record (Dennis et al., 2017; Hadj-Hammou et al., 2017; Lundmark, 2003; Rosário et al., 2024), and phenological metrics derived from these detections can be sensitive to sampling behaviour (Park et al., 2021). Because unstructured observations now dominate the available biodiversity data (Mason et al., 2018; Prudic et al., 2018; Soroye et al., 2018; van Tongeren et al., 2023), understanding how these observational patterns influence phenological inference is increasingly important.

Structured sampling and unstructured sampling represent two different ways of data collection by citizen scientists. The UK Butterfly Monitoring Scheme (UKBMS) is a flagship structured monitoring programme for butterflies and is characterised by fixed transect locations, predetermined survey frequency and a standardised recording protocol (Brereton et al., 2011). These features ensure that temporal changes can be consistently tracked at the same sites using the same methods throughout each monitoring period (Brereton et al., 2011; Pollard & Yates, 1997). In contrast, unstructured sampling databases such as iRecord Butterflies and iNaturalist offer greater flexibility because participants are free to choose when and where to record, with no fixed transects or prescribed survey frequency (Chandler et al., 2017; Li et al., 2025; Soroye et al., 2018). Moreover, unstructured sampling generates substantially larger data volumes than the structured transects (Mason et al., 2018; Prudic et al., 2018; Soroye et al., 2018; van Tongeren et al., 2023). However, this larger volume does not automatically guarantee a reduction in systematic difference (Isaac & Pocock, 2015; Kays et al., 2021). Although the large sample sizes of unstructured sampling datasets may reduce the absolute error of phenological estimates (Bishop et al., 2013), it remains unclear whether the apparent phenology generated by the two methods is similar.

Recorder behaviour is the primary factor shaping the frequency and occurrence patterns of citizen science data. Factors such as recording completeness (Sánchez-Fernández et al., 2021), species preferences (Goldstein et al., 2024; Soroye et al., 2018), location preferences (Li et al., 2025) and observer expertise (Kühn et al., 2024) have all been shown to generate differences under opportunistic sampling. In a phenological context, recorders may also intentionally search for extreme individuals, such as the 'first butterfly of the year' or the 'last butterfly of the flight period' (Bishop et al., 2013). While these are genuine observations, phenological metrics based on such records (e.g., first and last flight dates) can be sensitive to sampling behaviour, potentially leading to earlier or later estimated onset and end dates, as well as shifts in inferred flight period length and mean flight date. However, compared with structured monitoring schemes, the magnitude and consequences of these phenological differences remain poorly understood.

Here, we analyse opportunistic citizen science butterfly records in the United Kingdom (UK) from the iRecord Butterflies app and compare them with structured monitoring data from the UK Butterfly

Monitoring Survey (UKBMS). Both datasets are managed by the UK Centre for Ecology & Hydrology in partnership with Butterfly Conservation. As true phenological timing is unknown, we focus on differences between datasets and use UKBMS as the baseline because of its structured sampling, reliability and wide use in butterfly phenology research (Roy et al., 2015; Roy & Sparks, 2000). We test whether phenological metrics, including onset date, end date and weighted mean flight date, differ systematically between structured and unstructured datasets. We also assess whether these metrics derived from unstructured data follow similar temporal trends to those from structured monitoring, while tending to be consistently earlier or later.

METHODS

Data used

In this study, we considered two data sources: UKBMS data and data from the iRecord Butterflies app. As UKBMS is well established and widely known (Pollard & Yates, 1993), only the iRecord Butterflies app is described here. The iRecord Butterflies app was developed by Butterfly Conservation, the UK Centre for Ecology & Hydrology, and Natural Aptitude to facilitate the opportunistic recording of butterfly sightings in the United Kingdom (August et al., 2020). Records can be submitted from any location and any date, with no standardised sampling protocol. When submitting records, users enter the species observed and a location name, and can add information about the life cycle stage and the number of individuals seen. The date and coordinates are automatically recorded by the device, but can be edited before submission.

We selected 18 UKBMS transects that were sampled regularly and had complete long-term data between 1 April and 30 September (inclusive) for each year from 2011 to 2024, thereby avoiding the need for gap filling in years with missing observations. In total, this dataset comprised 7740 species detection records, each representing a count of individuals recorded per species per visit. To ensure clear and robust estimation of phenological metrics, we focused on single-generation species that do not overwinter as adults, as these species exhibit a single, well-defined peak in activity, reducing ambiguity in phenological estimates derived from different data sources.

We then identified the Watsonian Vice Counties (Dandy, 1971) in which these transects are located and filtered all butterfly records (324,954 records) submitted via the iRecord Butterflies app from the same Watsonian Vice Counties. We subsequently restricted the dataset to records collected within the same survey period as the UKBMS data (1 April to 30 September, inclusive, for each year from 2011 to 2024). The study area was divided into three regions: North (N), South East (SE) and South West (SW) (Mowbray et al., 2024). The specific UKBMS transects are provided in Table 1. All species names with scientific names analysed in this article are listed in Table 2.

The phenological metrics used in this study are summarised in Table 3. For each record, the observation date was converted to 'days since 1 April' by setting 1 April of the corresponding year as day

TABLE 1 Eighteen selected transects from the BMS data with Watsonian Vice County information.

UKBMS site no.	Site name	Watsonian Vice County	Watsonian Vice County number	OS grid reference	Region
6	Yarner Wood	South Devon	3	SX770780	SW
10	Oxwich	Glamorganshire	41	SS500870	SW
19	Studland Heath	Dorset	9	SZ020790	SW
28	Walberswick	East Suffolk	25	TM470740	SE
29	Aston Rowant (N)	Oxfordshire	23	SU730960	SE
34	Potton Wood	Bedfordshire	30	TL250500	SE
45	Gibraltar Point	North Lincolnshire	54	TF560580	SE
47	Skomer	Pembrokeshire	45	SM729093	SW
54	Nagshead	West Gloucestershire	34	SO600090	SW
60	Lindisfarne	North Northumberland	68	NU130430	N
61	Insh Marshes	East Inverness-shire	96	NH810010	N
62	Northward Hill	West Kent	16	TQ780760	SE
72	Gait Barrows	West Lancashire	60	SD470770	N
84	Ampfield Wood	South Hampshire	11	SU410238	SW
85	Derbyshire Dales	Derbyshire	57	SK180650	SE
95	Wyre Forest	Shropshire	40	SO755766	SW
97	Lullington Heath	East Sussex	14	TQ540020	SE
116	Taynish	Kintyre	101	NR720820	N

TABLE 2 List of butterfly species included in this study and their scientific names.

Common name	Scientific name
Chalk Hill Blue	<i>Polyommatus coridon</i>
Dark Green Fritillary	<i>Speyeria aglaja</i>
Essex Skipper	<i>Thymelicus lineola</i>
Gatekeeper	<i>Pyronia tithonus</i>
Grayling	<i>Hipparchia semele</i>
Green Hairstreak	<i>Callophrys rubi</i>
Large Skipper	<i>Ochlodes sylvanus</i>
Marbled White	<i>Melanargia galathea</i>
Meadow Brown	<i>Maniola jurtina</i>
Purple Hairstreak	<i>Favonius quercus</i>
Ringlet	<i>Aphantopus hyperantus</i>
Silver-spotted Skipper	<i>Hesperia comma</i>
Silver-studded Blue	<i>Plebejus argus</i>
Silver-washed Fritillary	<i>Argynnis paphia</i>
Small Skipper	<i>Thymelicus sylvestris</i>
White Admiral	<i>Limenitis camilla</i>
White-letter Hairstreak	<i>Satyrrium w-album</i>

1 and calculating the difference between the observation date and this reference date (Bishop et al., 2013). We then summarised the number of records for each species by year and region, calculating

TABLE 3 Explanation of phenological terms.

Term	Definition
First day	The earliest day of the year (as days since 1 April) on which a species was recorded in the dataset indicates the onset of detectable flight activity.
Last day	The latest day of the year (days since 1 April) on which a species was recorded represents the end of the observed flight period.
Onset date	The date by which 5% of total seasonal activity has occurred; a robust and sensitive estimate of the start of the main flight period.
End date	The date by which 95% of seasonal activity has occurred defines the end of the main flight season.
Peak day	The day with the highest number of records (or highest aggregated count) corresponds to the peak of flight activity.
Flight period length	The duration between the first and last recorded days (End date – Onset date) represents the overall length of the observed flight period.
Weighted mean flight date	The mean flight date calculated by weighting each day by its corresponding record count provides a central tendency of flight activity across the season.
Count	The number of observations (or aggregated counts across years) recorded for a species on each day. UKBMS data uses individual numbers as a count
Region	The geographic grouping used in the analysis (N, SE, SW) represents broad biogeographic areas across which phenology patterns were compared.

the percentage of records relative to the maximum annual number of records for that species to represent relative recording intensity. Because of inconsistent collection of abundance information in opportunistic datasets such as iRecord Butterflies, only occurrence data were used for iRecord Butterflies, whereas UKBMS data were based on individual counts due to the scheme's standardised monitoring protocol. To account for differences in recording intensity between iRecord Butterflies and UKBMS within the selected transects and regions, we weighted the data using the method described by Brakefield (1987). Specifically, we calculated the mean flight date as a weighted average of all observation days, where the weight for each day corresponded to the number of records: daily occurrence counts for iRecord Butterflies and daily abundance counts for UKBMS (Roy et al., 2015; Roy & Sparks, 2000).

Comparison of onset and end date records

Because the first day and last day are sensitive to sample size (Chapman et al., 2015), we extracted the date when cumulative records reach 5% and 95% of the annual total records (onset date and end date) to characterise the early and late parts of the flight period. For each species in each region, we calculated onset and end dates for UKBMS and iRecord Butterflies, respectively. Then we conducted paired *t*-tests to determine whether onset and end dates differed significantly between the two datasets, using species with five or more years of data in both datasets. *p* values were adjusted for multiple testing using the false discovery rate (FDR) method. Finally, we generated figures for the onset and end dates of annual change using species which have five or more consecutive years of data in both datasets. The corresponding figures for the first day and the last day are provided in Figure S1.

Comparison of annual weighted mean flight dates

We further assessed the time series of weighted mean flight date for UKBMS and iRecord Butterflies to examine interannual variation and potential trends over time in a stepwise manner. First, we constructed annual phenology time series based on the weighted mean flight date (Bishop et al., 2013), using the daily record count as the weight for that species on each day within a given region. We used species which have five or more consecutive years of data in both datasets. This allowed us to examine interannual changes in the weighted mean flight date. We then fitted linear models to estimate the annual rate of change (days per year) in weighted mean flight date for each species in each region and directly compared the slopes between UKBMS and iRecord Butterflies to assess whether the two data sources reflect consistent directions and magnitudes of phenological shifts. We calculated Pearson's correlation coefficient and its significance level to assess whether the weighted mean flight dates derived from the two data sources exhibited similar interannual patterns. We then performed paired *t*-tests to examine whether systematic differences

existed between the UKBMS and iRecord Butterflies estimates. *p* values were adjusted for multiple testing using the FDR method.

For each species and region, we calculated the weighted mean flight dates from the two data sources for each year, retained only the shared years and calculated the absolute difference in weighted mean flight dates (annual absolute error) between them. We then estimated the mean annual absolute error with its 95% confidence interval (Bishop et al., 2013). In addition to absolute differences, we also calculated the mean difference, defined as the average of annual differences in weighted mean flight date between UKBMS and iRecord Butterflies, to indicate the direction and magnitude of difference. The mean difference was calculated as follows:

$$\text{Mean difference} = \frac{1}{n} \sum_{t=1}^n (\text{WMFD}_{\text{UKBMS},t} - \text{WMFD}_{\text{iRecord},t})$$

where $\text{WMFD}_{\text{UKBMS},t}$ and $\text{WMFD}_{\text{iRecord},t}$ denote the weighted mean flight date for a given species in year *t* derived from the UKBMS and iRecord Butterflies datasets, respectively, and *n* is the number of overlapping years between the two datasets. Then, we performed paired *t*-tests to assess whether the weighted mean flight dates differed significantly between the two datasets. Finally, for each species, we summarised the occurrence data in iRecord Butterflies as the mean annual total count by region to explore whether recording intensity might influence the magnitude of the error (Bishop et al., 2013). All statistics were compiled for each species in each region to provide a systematic assessment of the extent to which unstructured observations deviate from structured monitoring in estimating annual weighted mean flight dates.

RESULTS

We present two analyses to compare butterfly phenology metrics derived from UKBMS and iRecord Butterflies data. First, we compare the onset date and end date of the flight period derived from the two datasets. Second, we analyse annual weighted mean flight dates to evaluate overall shifts across years.

Comparison of the onset date and end date

Among all butterfly species, 14 exhibited significantly positive onset mean errors, and none exhibited significantly negative values, indicating that iRecord Butterflies consistently reports an earlier onset than UKBMS (mean difference = 9.38 ± 1.24 days). In contrast, for end dates, 1 species exhibited significantly positive mean errors and 6 species exhibited significantly negative values, suggesting that iRecord Butterflies generally records a later end (mean difference = -8.46 ± 4.55 days) to the flight period (Table 4).

At the regional level, no significant differences were detected in the N region. In contrast, significant onset differences were observed in the SE, based on five species (mean difference = 8.20 ± 1.49 days),

TABLE 4 Results of paired *t*-tests comparing onset and end dates between UKBMS and iRecord Butterflies datasets. Mean errors (days), degrees of freedom (d.f.), FDR-adjusted *p* values and flight period difference are reported for each species–region combination. *p*-values less than 0.05 are highlighted in bold. Positive mean error values indicate later phenology in UKBMS relative to iRecord Butterflies, whereas negative values indicate earlier timing. Degrees of freedom reflect the number of overlapping years used in each comparison. Flight period difference with positive values indicates a longer flight period in UKBMS, while negative values indicate a longer flight period in iRecord Butterflies. Full results are available in Table S1.

Common name	Region	Onset mean error (days)	End mean error (days)	d.f.	Onset FDR-adjusted <i>p</i>	End FDR-adjusted <i>p</i>	Flight period difference
Chalk Hill Blue	SE	8.17	−1.33	11	0.07	0.79	−9.5
Dark Green Fritillary	N	6.64	−9.64	10	0.39	0.24	−16.27
Dark Green Fritillary	SE	1.55	−7.55	10	0.48	0.04	−9.09
Dark Green Fritillary	SW	17.4	0.8	9	0.01	0.83	−16.6
Essex Skipper	SE	13	3.25	11	0.04	0.73	−9.75
Essex Skipper	SW	13.91	2.27	10	0	0.63	−11.64
Gatekeeper	N	10.1	−9.5	9	0.08	0.17	−19.6
Gatekeeper	SE	4.73	−1.09	10	0	0.75	−5.82
Gatekeeper	SW	4.18	1.55	10	0.01	0.59	−2.64
Grayling	SW	9.09	−2.09	10	0	0.75	−11.18
Green Hairstreak	SE	−0.55	2.36	10	0.9	0.63	2.91
Green Hairstreak	SW	18.1	−30.8	9	0	0.01	−48.9
Large Skipper	N	13.78	−1.67	8	0.08	0.8	−15.44
Large Skipper	SE	2.91	−0.09	10	0.25	0.97	−3
Large Skipper	SW	7.45	2.64	10	0.04	0.29	−4.82
Marbled White	SE	2.42	−3.67	11	0.39	0.12	−6.08
Marbled White	SW	−1.64	−11.73	10	0.65	0.02	−10.09
Meadow Brown	N	3.36	−10.64	10	0.21	0.04	−14
Meadow Brown	SE	9.75	9.25	11	0	0.04	−0.5
Meadow Brown	SW	8.73	−4	10	0	0.17	−12.73
Purple Hairstreak	SE	10.6	−5	4	0.3	0.59	−15.6
Purple Hairstreak	SW	6	−23.22	8	0.52	0.06	−29.22
Ringlet	N	3.45	−3.91	10	0.25	0.09	−7.36
Ringlet	SE	5.55	−4	10	0	0.04	−9.55
Ringlet	SW	3.45	−3.73	10	0	0.04	−7.18
Silver-spotted Skipper	SE	−0.9	1.8	9	0.9	0.78	2.7
Silver-studded Blue	SW	2.73	−7.91	10	0.21	0.17	−10.64
Silver-washed Fritillary	SE	4.27	1.82	10	0.11	0.69	−2.45
Silver-washed Fritillary	SW	1.42	−1.92	11	0.84	0.69	−3.33
Small Skipper	N	−1.63	−14.5	7	0.81	0.1	−12.88
Small Skipper	SE	8	1.36	10	0	0.78	−6.64
Small Skipper	SW	0.09	−4	10	0.97	0.1	−4.09
White Admiral	SE	1.33	−7.56	8	0.73	0.39	−8.89
White Admiral	SW	7.91	−1.27	10	0.03	0.78	−9.18
White-letter Hairstreak	SE	10.67	−7	8	0.08	0.1	−17.67

Abbreviations: FDR, false discovery rate; UKBMS, UK Butterfly Monitoring Scheme.

and in the SW, based on nine species (mean difference = 10.03 ± 1.77 days). Additional end date differences were observed in the SE, based on three species (mean difference = −0.77 ± 5.11 days), and in the SW, also based on three species (mean difference = −15.42 ± 8.03 days). A significantly later end date

(−10.64 days) was detected for Meadow Brown in the N region (FDR-corrected *p* = 0.038), although based on a single species. Referring to flight period length, only two species in two regions show positive values, implying a broader apparent flight window compared with UKBMS.



FIGURE 1 Dates corresponding to the onset and end date of Meadow Brown, Ringlet and Dark Green Fritillary across the three regions (N, SE and SW) for UKBMS and iRecord Butterflies. Each panel shows the dates (days since 1 April) when the cumulative number of records reached 5% of the annual total (early flight period) and 95% of the annual total (late flight period). For the other species figures, see supplementary material (Figure S3). UKBMS, UK Butterfly Monitoring Scheme; N, North; SE, South East; SW, South West.

Taking Meadow Brown, Ringlet and Dark Green Fritillary as representative examples of species with high, moderate and low total record counts (respectively), there are clear differences between UKBMS and iRecord Butterflies across the SE, SW and N regions (Figures 1 and 2). In the N, SE and SW regions, the onset date derived from iRecord Butterflies is consistently earlier than that from UKBMS, and in the SE and SW regions, the end date recorded by iRecord Butterflies is generally later, indicating a longer flight period when using unstructured data. Phenological curves for these species clearly illustrate differences across all regions, with iRecord Butterflies consistently showing an earlier first day and a later last day than UKBMS (Figure 2). This indicates that, for these species, iRecord Butterflies records reflect a longer apparent flight season in all regions. Equivalent figures for all other species are provided in the supplementary materials (Figure S3).

Comparison of annual weighted mean flight dates

Butterfly species display a range of agreement patterns across regions. Many slope pairs fall away from the $y = x$ line, indicating

absolute variation in the rate of change estimated from each dataset (Figure 3). However, based on slope difference analysis (UKBMS – iRecord Butterflies), 95% confidence intervals overlapped zero for all species region combinations (Table 5). These results show that the estimated temporal trends in weighted mean flight date are not statistically distinguishable between iRecord Butterflies and UKBMS for any species in any region.

Considering Meadow Brown, Ringlet and Dark Green Fritillary as illustrative examples, annual trends in weighted mean flight dates are broadly similar between the two datasets, especially in the SE and SW regions, where the temporal patterns align most closely (Figure 4, Table 5). However, annual comparisons (Figure 4) indicate clear differences that in the SE and SW regions, weighted mean flight dates derived from iRecord Butterflies are earlier than those from UKBMS in nearly all years, whereas in the N region, iRecord Butterflies dates are later in most years except 2014 and 2019.

Across species and regions, 22 combinations (involving 12 species) showed Pearson correlation coefficients greater than 0.5 between UKBMS and iRecord Butterflies weighted mean flight dates (Table 6).

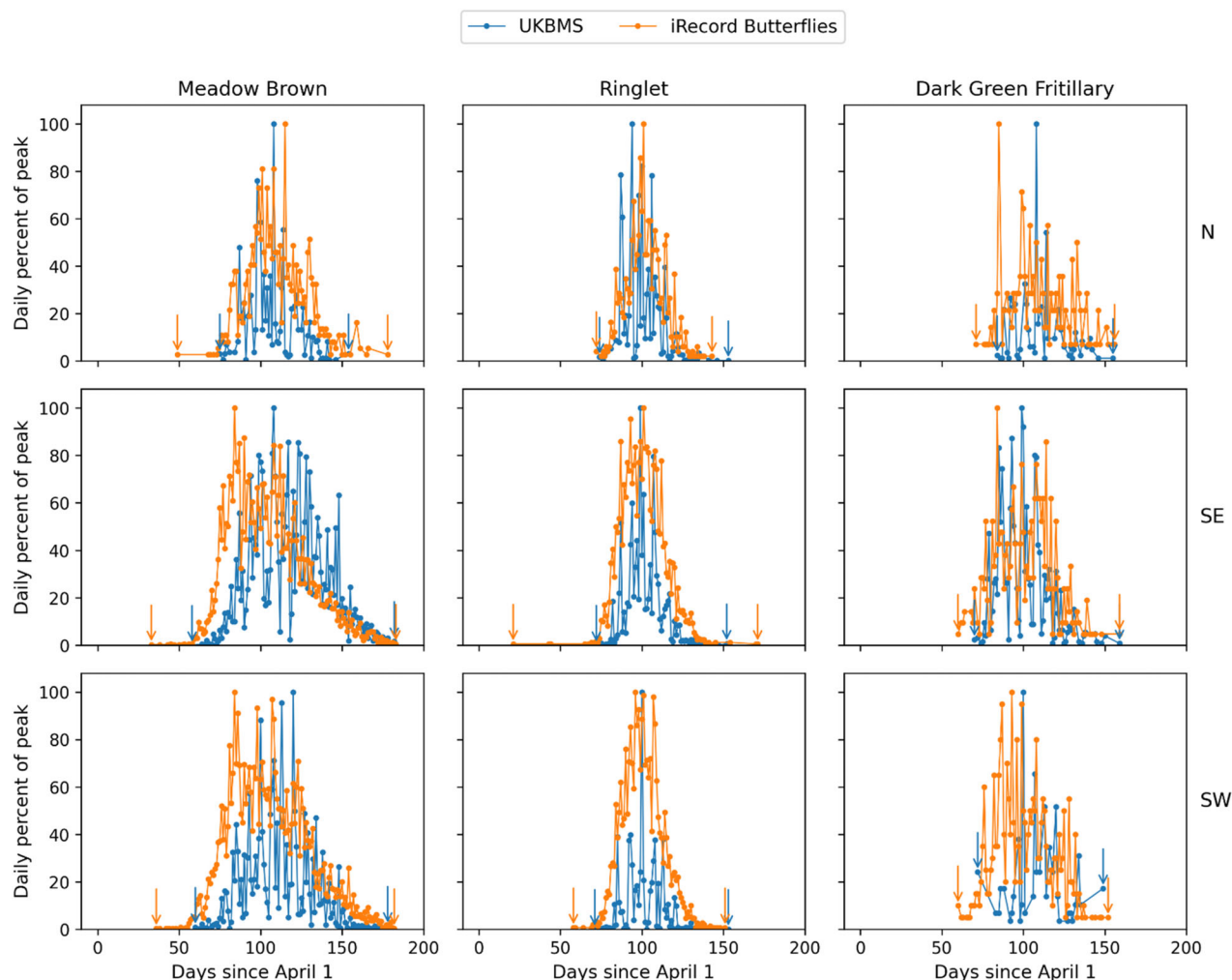


FIGURE 2 Phenology curves for Meadow Brown, Ringlet and Dark Green Fritillary, based on all available years of data, comparing structured monitoring (UKBMS) with opportunistic citizen science observations (iRecord Butterflies) across three regions (N, SE and SW). Each panel shows the daily percentage of the peak number of records within a region. Arrows indicate the first and last day of the flight period among all available years. Other species figures are available in supplementary material (Figure S1). UKBMS, UK Butterfly Monitoring Scheme; N, North; SE, South East; SW, South West.

Of these, 15 combinations (across 10 species) exhibited statistically significant correlations based on FDR-adjusted p values ($p < 0.05$). Seven combinations (six species) showed significant differences in weighted mean flight dates based on paired t -tests (FDR-adjusted $p < 0.05$), with iRecord Butterflies displaying an average advancement of 7.57 ± 1.57 days relative to UKBMS (mean error defined as UKBMS minus iRecord Butterflies). Mean differences were positive in 22 combinations (14 species), indicating systematically earlier flight dates in iRecord Butterflies, and all 7 of these combinations (across 6 species) also showed statistically significant paired t -test differences (FDR-adjusted $p < 0.05$). Notably, in the N region, the maximum annual total count was only 80.36 (for Meadow Brown), indicating substantially smaller sample sizes compared to those in the SW and SE. Despite this, a large proportion of butterfly species still showed statistically significant results after FDR adjustment when not considering regional differences.

DISCUSSION

Our study shows that estimated temporal trends in weighted mean flight date are not statistically distinguishable between iRecord Butterflies and UKBMS for any species in any region. However, iRecord Butterflies tend to show a broader apparent flight period than UKBMS (Table 4), and the weighted mean flight dates derived from iRecord Butterflies are generally significantly earlier across most species and regions. In other words, unstructured observations may reliably capture overall phenological trends (Larsen et al., 2022), but phenological metrics derived from these data may systematically differ from those estimated by the UKBMS.

One possible explanation for the earlier phenology metrics in iRecord Butterflies relates to the difference in the two datasets. Firstly, estimating the start or end of seasonal events is particularly difficult when using presence only records (Belitz et al., 2020) such as those

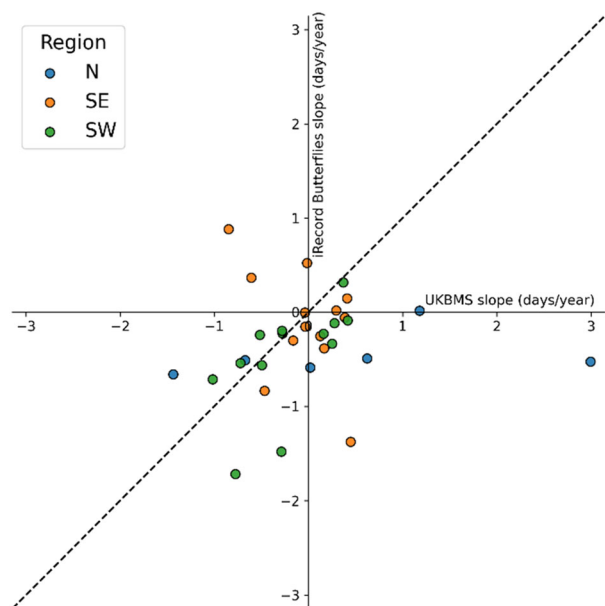


FIGURE 3 Relationship between phenological trends in weighted mean flight date estimated from UKBMS and iRecord Butterflies across three regions (N, SE and SW). Each point represents the fitted slope (days per year) of weighted mean flight date against year for one butterfly species in one region (32 points in total, 16 species; not all species occur in all regions). Colours indicate region examined, and the broken line ($y = x$) indicates equal phenological trends between UKBMS and iRecord Butterflies. UKBMS, UK Butterfly Monitoring Scheme; N, North; SE, South East; SW, South West.

provided by iRecord Butterflies, because observations around the beginning and end of the flight season are typically sparse and unevenly distributed, making records more likely to be sporadic and increasing uncertainty in identifying first and last events (Belitz et al., 2020). In contrast, UKBMS collects precise abundance counts, which more reliably capture the shape of the seasonal flight curve. Secondly, because iRecord Butterflies has more recorders, covering more of the landscape and, therefore, more populations of each species than UKBMS, it may better estimate the true overall phenology metrics of the species across the United Kingdom, whereas transects are only estimating the true phenology metrics of the monitored populations—a small subset of all populations and, therefore, presumably, a smaller subset of all the individual phenological variance present in the species. These differences in data structure constitute a further potential mechanism contributing to the observed discrepancies in phenology distributions.

Recorder motivation may also play an important role. Although the initial motivation for taking part in citizen science has been reported as self-improvement and enjoyment of social interactions (Peter et al., 2021; Vanden Berghen et al., 2024), we hypothesise that some recorders may be motivated by the desire to be the first to report a species on social media or citizen science platforms (Lowe et al., 2025), which could further encourage observations and submissions early in the flight season (Belitz et al., 2020). Declining recorder engagement over the season may also play a role, since participants

often start very enthusiastic, but motivation may wane during the mid- and late-flight period (Belitz et al., 2020).

In contrast, existing studies indicate that later end dates are generally associated with species capable of adult overwintering, high dispersal or additional late-season generations (Löckinger et al., 2024). This apparent prolongation of flight period may reflect recorder behaviour, such as competition among observers to be the first or last to submit a sighting, or a tendency to visit high abundance populations (Isaac & Pocock, 2015; Li et al., 2025) and recorders might have improved knowledge of local area so that they know where to look for early butterflies (Kühn et al., 2024).

From a regional perspective, the northern United Kingdom shows weaker phenological divergence between datasets (Table 4), likely due to shorter flight periods, lower sample size (Table 6) (Menéndez et al., 2006) and lower recorder numbers, resulting in limited sample coverage and some extreme value points (Figure 3). Correspondingly, seasonal extensions in iRecord Butterflies are evident in the SE and SW regions, yet are largely absent in the north (Table 4), where both datasets produce more comparable phenology curves across multiple species. This suggested that data scarcity may constrain detectable differences. However, sample size alone does not account for the observed patterns. Even in Meadow Brown, where data volume is substantial (Table 6), the decrease in mean absolute error indicates reduced sampling error (Bishop et al., 2013), yet the difference in mean flight date remains statistically significant. This demonstrates that increasing sample size cannot eliminate systematic differences and that divergence more likely arises from the data collection mechanism itself. Therefore, reducing phenological differences in citizen science datasets may rely less on accumulating additional observations and more on improving the structure and balance of sampling effort.

Implications for citizen science and phenology

Our first implication is that systematic phenology difference reflects the complexity of recorder behaviour. Although common species have very large numbers of records (Table 6), phenological metrics derived from iRecord Butterflies, including onset, end date and mean flight date, still differ systematically from those estimated by the UKBMS, indicating that increasing sample size alone does not eliminate this difference. Thus, unstructured datasets may require moderate guidance rather than total reliance on opportunistic sampling (Soroye et al., 2018). For example, platforms could provide reference information on typical observation periods at the point of submission to help observers judge whether early records are credible, although such guidance must be designed carefully to avoid encouraging competitive early reporting (Belitz et al., 2020). More reliable improvement may come from combining partially structured sampling (Farr et al., 2024; Goldstein et al., 2024; Kelling et al., 2019), increasing coverage during mid and late season (Belitz et al., 2020), or removing extreme records during data processing, so that data quality is strengthened without reducing participation.

TABLE 5 Difference in weighted mean flight date slopes between UK Butterfly Monitoring Scheme (UKBMS) and iRecord Butterflies across regions, with 95% confidence intervals. Noting that the slope difference is calculated from the slope of UKBMS minus the slope of iRecord Butterflies in each region for each species.

Common name	Region	Slope difference	CI95 lower	CI95 upper
Chalk Hill Blue	SE	−0.54	−2.11	1.03
Dark Green Fritillary	N	1.16	−2.31	4.63
Dark Green Fritillary	SE	0.12	−1.87	2.10
Dark Green Fritillary	SW	0.05	−3.33	3.44
Essex Skipper	SE	−1.73	−4.17	0.70
Essex Skipper	SW	0.07	−1.85	1.99
Gatekeeper	N	0.61	−2.66	3.88
Gatekeeper	SE	−0.04	−1.49	1.42
Gatekeeper	SW	0.39	−0.90	1.67
Grayling	SW	0.59	−0.86	2.03
Green Hairstreak	SE	0.14	−1.45	1.72
Green Hairstreak	SW	0.94	−6.60	8.49
Large Skipper	N	3.52	−3.14	10.18
Large Skipper	SE	0.26	−1.60	2.13
Large Skipper	SW	0.50	−0.88	1.88
Marbled White	SE	0.37	−1.38	2.12
Marbled White	SW	−0.28	−2.38	1.83
Meadow Brown	N	−0.16	−2.54	2.22
Meadow Brown	SE	0.38	−1.20	1.95
Meadow Brown	SW	0.39	−1.04	1.82
Ringlet	N	1.11	−0.30	2.53
Ringlet	SE	0.44	−1.03	1.91
Ringlet	SW	−0.05	−1.34	1.24
Silver-spotted Skipper	SE	−0.97	−3.15	1.21
Silver-studded Blue	SW	−0.08	−2.04	1.87
Silver-washed Fritillary	SE	0.27	−1.91	2.45
Silver-washed Fritillary	SW	1.19	−1.50	3.89
Small Skipper	N	−0.78	−4.99	3.44
Small Skipper	SE	0.55	−1.52	2.62
Small Skipper	SW	−0.18	−1.76	1.40
White Admiral	SW	−0.30	−2.78	2.17
White-letter Hairstreak	SE	1.83	−1.57	5.22

Our second implication is that unstructured datasets such as iRecord Butterflies can describe overall phenological trends but are less reliable for calculating phenology indices that are comparable to those derived from structured data. For many species without long-term structured monitoring (Li-Jun et al., 2013; Segre et al., 2023; van Tongeren et al., 2023), these records may capture directional shifts in flight period; however, the resulting phenological indices are more variable and cannot be directly integrated with those derived from structured datasets without appropriate adjustment.

Our third implication is that the systematic phenological metrics difference observed in unstructured datasets and structured datasets

shows strong potential for cross-platform and cross-taxon generalisation. The phenological metrics difference exhibited by iRecord Butterflies is unlikely to be unique to butterfly observations but represents opportunistic sampling databases shaped by citizen recording behaviour (Goldstein et al., 2024; Isaac & Pocock, 2015; Li et al., 2025). For example, observer preferences for recording at particular times of day or periods of the season may lead to temporal clustering of records (Hadj-Hammou et al., 2017; Rosário et al., 2024). In addition, citizen science events concentrated over a few days can result in bursts of submissions (Dennis et al., 2017; Lundmark, 2003). Both factors may introduce differences that affect phenological estimates (Park et al., 2021).

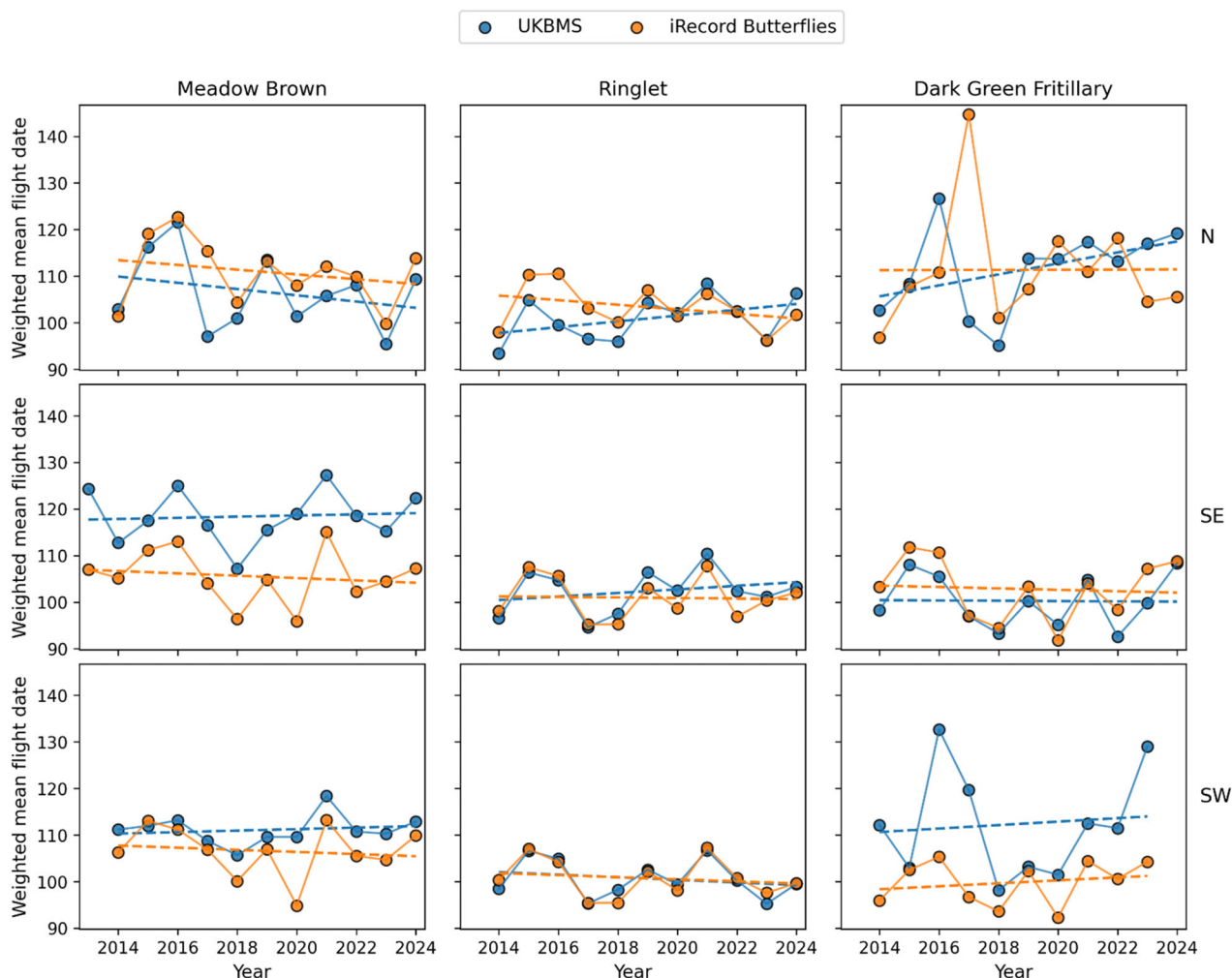


FIGURE 4 Temporal trends in the weighted mean flight date for Meadow Brown, Ringlet and Dark Green Fritillary, comparing UKBMS with iRecord Butterflies across three regions (N, SE and SW). Each panel shows the annual weighted mean flight date for the species within a region, and dashed lines indicate linear trends for each dataset. Only years for which both datasets had at least five years of overlapping observations were included in the plots. Other species figures and detailed slope information are available in supplementary material (Figure S4). UKBMS, UK Butterfly Monitoring Scheme; N, North; SE, South East; SW, South West.

Limitations and future directions

First, we did not include factors such as variation in weather conditions (Colom et al., 2022; Goded et al., 2024), nor did we explicitly account for variation in observer activity within a day (Idec et al., 2025). These factors could be incorporated by linking individual records to fine-scale temperature data or by modelling detection probability as a function of time of day. Such factors may change the results because butterfly activity and detectability are temperature-dependent and vary across the day (Idec et al., 2025).

Second, our comparison was based on UKBMS continuous monitoring sites and matched with iRecord Butterflies data at the county level rather than at a finer spatial grid. Although the repeated and standardised visits of UKBMS are likely to provide a reasonable estimate of species phenology, UKBMS does not sample the entire year or all populations (Dennis et al., 2024). Therefore, inter-population

variation in phenology may be underestimated, and spatial heterogeneity may not have been fully captured.

Third, we acknowledge that the methods used to derive phenological metrics are relatively simple. In particular, the abundance weighted mean may be less appropriate for iRecord Butterflies data, as count data from opportunistic records can be highly skewed and biased, and alternative methods may offer more robust and comparable phenological estimates.

Fourth, we recommended constructing a modelling framework that explicitly represents observer behaviour, and future research should incorporate decision rules, apply hierarchical modelling or use causal diagram-based approaches to identify potential confounding pathways to better understand how unstructured recording shapes phenology estimation and how differences might be corrected (Boyd et al., 2025). Finally, although our analyses centred on butterflies, we recommend extending this assessment to other taxa with a structured scheme to evaluate whether the same phenology difference is present.

TABLE 6 Comparison of weighted mean flight dates between UKBMS and iRecord Butterflies across butterfly species and regions. Species are ordered by mean annual total record count to show how differences between the two schemes vary with sample size. The table reports Pearson correlation coefficients (r) and FDR-adjusted p values for the Pearson correlation analysis, FDR-adjusted p values for paired t -tests, mean absolute error, mean difference in weighted mean flight date (UKBMS minus iRecord Butterflies), and annual total record counts. Pearson's $r > 0.5$ indicates moderate to strong synchrony in interannual variation. Statistical significance for both correlation and paired t -test analyses is assessed using FDR-adjusted p values. p -values less than 0.05 are highlighted in bold. The full table is available in Table S2.

Common name	Region	df	Pearson r	Pearson's FDR-adjusted p	FDR-adjusted p	Mean absolute error (days)	Mean difference (days)	Mean annual total count
Meadow Brown	SW	10	0.7	0.04	0.02	4.7	4.51	1195.64
Meadow Brown	SE	11	0.69	0.04	0	12.87	12.87	1114.75
Gatekeeper	SW	10	0.87	0	0.03	2.67	2.24	797.09
Gatekeeper	SE	10	0.78	0.02	0.5	2.25	0.9	758.91
Ringlet	SE	10	0.88	0	0.17	2.16	1.41	412.27
Ringlet	SW	10	0.94	0	0.95	1.02	-0.06	336.55
Marbled White	SW	10	0.55	0.16	0.06	7.25	-6.37	305.64
Marbled White	SE	11	0.91	0	0.62	2.39	0.7	264.5
Small Skipper	SE	10	0.78	0.02	0.02	5.24	5.24	233.91
Large Skipper	SE	10	0.73	0.03	0.33	3.68	1.92	223.18
Large Skipper	SW	10	0.46	0.26	0.02	6.2	5.83	220.55
Small Skipper	SW	10	0.54	0.16	0.3	3.81	2.42	218.18
Silver-washed Fritillary	SW	11	0.77	0.02	0.86	5.03	-0.8	139.5
Silver-washed Fritillary	SE	10	0.76	0.02	0.14	4.38	3.25	96.82
Purple Hairstreak	SE	4	0.32	0.75	0.52	9.87	5.12	86.8
Green Hairstreak	SW	9	0.34	0.47	0.17	8.36	-5.57	86.3
Grayling	SW	10	0.32	0.47	0.06	6.44	4.9	81.09
Meadow Brown	N	10	0.77	0.02	0.09	4.66	-4.32	80.36
Ringlet	N	10	0.57	0.15	0.2	3.81	-2.46	73.36
Essex Skipper	SE	11	0.03	1	0.33	9.22	5.4	60.83
Green Hairstreak	SE	10	0.76	0.02	0.38	2.75	-1.33	60
Silver-studded Blue	SW	10	0.74	0.03	0.36	4.53	-1.98	58.64
Chalk Hill Blue	SE	11	0.29	0.47	0.99	5.31	-0.04	58.33
White Admiral	SE	8	0.6	0.16	0.11	6.73	-5.96	57.33
Dark Green Fritillary	SW	9	0.57	0.16	0.02	12.53	12.53	56.1
Dark Green Fritillary	SE	10	0.87	0	0.1	3.27	-2.53	44.09
White Admiral	SW	10	0.64	0.08	0.15	6.95	4.81	39.73
Purple Hairstreak	SW	8	0.41	0.42	0.37	14.37	-7.48	39
Essex Skipper	SW	10	0.44	0.29	0.01	10.24	9.75	26.82
Small Skipper	N	7	0.4	0.47	0.16	8.59	-6.78	25.38
Dark Green Fritillary	N	10	-0.12	0.84	0.99	10.95	0.19	20.82
White-letter Hairstreak	SE	8	0.65	0.14	0.8	4.95	0.91	20.78
Silver-spotted Skipper	SE	9	0.46	0.29	0.72	4.29	1.07	19.8
Gatekeeper	N	9	-0.14	0.84	0.63	10.47	2.86	14.6
Large Skipper	N	8	-0.13	0.84	0.2	10.82	8.82	12.44

Abbreviations: FDR, false discovery rate; N, North; SE, South East; SW, South West; UKBMS, UK Butterfly Monitoring Scheme.

CONCLUSION

This study shows that although iRecord Butterflies and UKBMS data display broadly similar phenology trends, most species recorded in iRecord Butterflies exhibit earlier onset and later end dates and systematically

earlier weighted mean flight dates, resulting in a longer flight period length. This shift arises from recorder behaviour and differences between the traits of the two databases. While unstructured data can reliably capture trends in mean flight date, they are less reliable for determining the exact magnitude of advancement and therefore require cautious

interpretation. On the other hand, structured sampling may not reveal the full phenology of species because only a subset of populations is monitored. Importantly, the difference in recorder behaviours is likely to have wider relevance across platforms and taxa, and similar patterns may emerge in community science datasets. Recognising and accounting for systematic differences in phenology metrics will be important when using unstructured citizen science data for phenological inference in butterflies and other taxa.

AUTHOR CONTRIBUTIONS

Mingrui Li: Methodology; writing – original draft; writing – review and editing; formal analysis; data curation; visualization; software. **Jonathan Bennie:** Conceptualization; methodology; software; supervision; writing – review and editing; visualization. **David B. Roy:** Writing – review and editing; methodology; supervision; resources. **Richard Fox:** Writing – review and editing; resources; methodology. **Robin J. Boyd:** Writing – review and editing; methodology. **Chloë Smith:** Writing – review and editing; resources. **Richard H. French-Constant:** Writing – review and editing; conceptualization; methodology; supervision.

ACKNOWLEDGEMENTS

We thank all recorders who contributed data to the two recording schemes used in this study. Contains UK Butterfly Monitoring Scheme (UKBMS) data © copyright and database right Butterfly Conservation, the UK Centre for Ecology and Hydrology, British Trust for Ornithology and the Joint Nature Conservation Committee, 2025. The UKBMS is organised and funded by Butterfly Conservation, the UK Centre for Ecology and Hydrology, the British Trust for Ornithology and the Joint Nature Conservation Committee, and is indebted to all volunteers who contribute data to the scheme. RF is supported by the Heather Corrie Fund from Butterfly Conservation. We also thank the editor and two anonymous reviewers for their constructive comments, which helped improve the manuscript.

FUNDING INFORMATION

This research did not receive any specific grant from any funding agency in the public, commercial or not-for-profit sectors.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in datashare.zip at <https://doi.org/10.6084/m9.figshare.28783964> (Li, 2025).

ORCID

Mingrui Li  <https://orcid.org/0009-0008-8667-2217>

Richard Fox  <https://orcid.org/0000-0001-6992-3522>

REFERENCES

August, T., Fox, R., Roy, D.B. & Pocock, M.J.O. (2020) Data-derived metrics describing the behaviour of field-based citizen scientists provide insights for project design and modelling bias. *Scientific Reports*, 10, 11009.

- Belitz, M.W., Larsen, E.A., Ries, L. & Guralnick, R.P. (2020) The accuracy of phenology estimators for use with sparsely sampled presence-only observations. *Methods in Ecology and Evolution*, 11, 1273–1285.
- Bishop, T.R., Botham, M.S., Fox, R., Leather, S.R., Chapman, D.S. & Oliver, T.H. (2013) The utility of distribution data in predicting phenology. *Methods in Ecology and Evolution*, 4, 1024–1032.
- Bowler, D.E., Bhandari, N., Repke, L., Beuthner, C., Callaghan, C.T., Eichenberg, D. et al. (2022) Decision-making of citizen scientists when recording species observations. *Scientific Reports*, 12, 11069.
- Boyd, R.J., Botham, M., Dennis, E., Fox, R., Harrower, C., Middlebrook, I. et al. (2025) Using causal diagrams and superpopulation models to correct geographic biases in biodiversity monitoring data. *Methods in Ecology and Evolution*, 16, 332–344.
- Brakefield, P.M. (1987) Geographical variability in, and temperature effects on, the phenology of *Maniola jurtina* and *Pyronia tithonus* (Lepidoptera, Satyrinae) in England and Wales. *Ecological Entomology*, 12, 139–148.
- Brereton, T., Roy, D., Middlebrook, I., Botham, M. & Warren, M. (2011) The development of butterfly indicators in the United Kingdom and assessments in 2010. *Journal of Insect Conservation*, 15, 139–151.
- Capinha, C., Ceia-Hasse, A., De-Miguel, S., Vila-Viçosa, C., Porto, M., Jarić, I. et al. (2024) Using citizen science data for predicting the timing of ecological phenomena across regions. *Bioscience*, 74, 383–392.
- Chandler, M., See, L., Copas, K., Bonde, A.M., López, B.C., Danielsen, F. et al. (2017) Contribution of citizen science towards international biodiversity monitoring. *Biological Conservation*, 213, 280–294.
- Chapman, D.S., Bell, S., Helfer, S. & Roy, D.B. (2015) Unbiased inference of plant flowering phenology from biological recording data. *Biological Journal of the Linnean Society*, 115, 543–554.
- Colom, P., Ninyerola, M., Pons, X., Traveset, A. & Stefanescu, C. (2022) Phenological sensitivity and seasonal variability explain climate-driven trends in Mediterranean butterflies. *Proceedings of the Royal Society B: Biological Sciences*, 289, 20220251.
- Dandy, J.E. (1971) Watsonian vice-counties of Great Britain. *Taxon*, 20, 627.
- Debinski, D.M., Warchola, N., Altizer, S. & Crone, E.E. (2025) Implications of summer breeding phenology on demography of monarch butterflies. *Journal of Animal Ecology*, 94, 682–692.
- Dennis, E.B., Morgan, B.J., Brereton, T.M., Roy, D.B. & Fox, R. (2017) Using citizen science butterfly counts to predict species population trends. *Conservation Biology*, 31, 1350–1361.
- Dennis, E.B., Morgan, B.J., Harrower, C.A., Bourn, N.A. & Fox, R. (2024) Incorporating phenology to estimate species' population trends from snapshot citizen-science data. *Journal of Agricultural, Biological, and Environmental Statistics*, 31, 1–17.
- Farr, M.T., Zylstra, E.R., Ries, L. & Zipkin, E.F. (2024) Overcoming data gaps using integrated models to estimate migratory species' dynamics during cryptic periods of the annual cycle. *Methods in Ecology and Evolution*, 15, 413–426.
- Goded, M., Ursul, G., Baz, A. & Wilson, R.J. (2024) Changes to butterfly phenology versus elevation range after four decades of warming in the mountains of central Spain. *Journal of Insect Conservation*, 28, 935–949.
- Goldstein, B.R., Stoudt, S., Lewthwaite, J.M., Shirey, V., Mendoza, E. & Guzman, L.M. (2024) Logistical and preference bias in participatory science butterfly data. *Frontiers in Ecology and the Environment*, 22, e2783.
- Habel, J.C., Huemer, P., Schmitt, T., Rüdiger, J. & Ulrich, W. (2025) Strengths and shortcomings of citizen science data: lessons from Austrian butterflies. *Journal of Insect Conservation*, 29, 1–13.
- Hadj-Hammou, J., Loisele, S.A., Ophof, D. & Thornhill, I. (2017) Getting the full picture: assessing the complementarity of citizen science and agency monitoring data. *PLoS One*, 12, e0188507.
- Idec, J., Campbell, C.J., Belitz, M., Akshay, V. & Guralnick, R. (2025) Using citizen science data to estimate trait and climate drivers of daily activity patterns in temperate butterflies. *PLoS One*, 20, e0335856.

- Isaac, N.J. & Pocock, M.J. (2015) Bias and information in biological records. *Biological Journal of the Linnean Society*, 115, 522–531.
- Kays, R., Lasky, M., Parsons, A.W., Pease, B. & Pacifici, K. (2021) Evaluation of the spatial biases and sample size of a statewide citizen science project. *Citizen Science: Theory and Practice*, 6, 34.
- Kelling, S., Johnston, A., Bonn, A., Fink, D., Ruiz-Gutierrez, V., Bonney, R. et al. (2019) Using semistructured surveys to improve citizen science data for monitoring biodiversity. *Bioscience*, 69, 170–179.
- Kühn, E., Harpke, A., Schmitt, T., Settele, J. & Kühn, I. (2024) Counting butterflies—are old-fashioned ways of recording data obsolete? *Journal of Insect Conservation*, 28, 577–588.
- Larsen, E.A., Belitz, M.W., Guralnick, R.P. & Ries, L. (2022) Consistent trait-temperature interactions drive butterfly phenology in both incidental and survey data. *Scientific Reports*, 12, 13370.
- Li, M. (2025) Phenology paper data and code. [dataset] FIGSHARE; v10; <https://doi.org/10.6084/m9.figshare.28783964.v10>.
- Li, M., Boyd, R.J., Smith, C., Fox, R., Roy, D. & Bennie, J. (2025) Citizen scientists as butterfly predators: using foraging theory to understand individual recorder behaviour. *Ecological Modelling*, 510, 111344.
- Li-Jun, F., Hai-Gen, X. & Jian-Ling, G. (2013) History and present status of butterfly monitoring in Europe and related development strategies for China. *Yingyong Shengtai Xuebao*, 24, 2691–2698.
- Löckinger, M., Trutchnig, W., Ulrich, W., Gros, P., Schmitt, T. & Habel, J.C. (2024) Ecological performance determines phenological responses of butterflies in northern Austria. *Global Ecology and Conservation*, 54, e03114.
- Lowe, S.K., Mason, B.M., Guralnick, R., Morales, N.A. & Callaghan, C.T. (2025) Participation intensity influences motivations for contributing to iNaturalist. *Citizen Science: Theory and Practice*, 10, 25.
- Lundmark, C. (2003) BioBlitz: getting into backyard biodiversity. *Bioscience*, 53, 329.
- Mason, B.M., Mesaglio, T., Barratt Heitmann, J., Chandler, M., Chowdhury, S., Gorta, S.B.Z. et al. (2025) iNaturalist accelerates biodiversity research. *Bioscience*, 75, 953–965.
- Mason, S.C., Hill, J.K., Thomas, C.D., Powney, G.D., Fox, R., Brereton, T. et al. (2018) Population variability in species can be deduced from opportunistic citizen science records: a case study using British butterflies. *Insect Conservation and Diversity*, 11, 131–142.
- Menéndez, R., Megías, A.G., Hill, J.K., Bräschler, B., Willis, S.G., Collingham, Y. et al. (2006) Species richness changes lag behind climate change. *Proceedings of the Royal Society B: Biological Sciences*, 273, 1465–1470.
- Mowbray, S., Bennie, J., Rhodes, M.W., Smith, D.A. & French-Constant, R.H. (2024) Eyespot variation and field temperature in the meadow Brown butterfly. *Ecology and Evolution*, 14, e10842.
- Park, D.S., Newman, E.A. & Breckheimer, I.K. (2021) Scale gaps in landscape phenology: challenges and opportunities. *Trends in Ecology & Evolution*, 36, 709–721.
- Peter, M., Diekötter, T., Höffler, T. & Kremer, K. (2021) Biodiversity citizen science: outcomes for the participating citizens. *People and Nature*, 3, 294–311.
- Pollard, E. & Yates, T. (1997) *Monitoring Butterflies for Ecology and Conservation: Insect Conservation*. 1, 63–64, Springer.
- Pollard, E. & Yates, T.J. (1993) *Monitoring Butterflies for Ecology and Conservation: the British Butterfly Monitoring Scheme*. London, UK: Springer Science & Business Media.
- Primack, R.B., Gallinat, A.S., Ellwood, E.R., Crimmins, T.M., Schwartz, M.D., Staudinger, M.D. et al. (2023) Ten best practices for effective phenological research. *International Journal of Biometeorology*, 67, 1509–1522.
- Prudic, K.L., Oliver, J.C., Brown, B.V. & Long, E.C. (2018) Comparisons of citizen science data-gathering approaches to evaluate urban butterfly diversity. *Insects*, 9, 186.
- Rosário, I.T., Tiago, P., Chozas, S. & Capinha, C. (2024) Drivers of temporal bias in biodiversity recording by citizen scientists. *bioRxiv*.
- Rosário, I.T., Tiago, P., Chozas, S. & Capinha, C. (2025) When do citizen scientists record biodiversity? Non-random temporal patterns of recording effort and associated factors. *People and Nature*, 7, 860–870.
- Roy, D.B., Oliver, T.H., Botham, M.S., Beckmann, B., Brereton, T., Dennis, R.L. et al. (2015) Similarities in butterfly emergence dates among populations suggest local adaptation to climate. *Global Change Biology*, 21, 3313–3322.
- Roy, D.B. & Sparks, T. (2000) Phenology of British butterflies and climate change. *Global Change Biology*, 6, 407–416.
- Sánchez-Fernández, D., Fox, R., Dennis, R.L. & Lobo, J.M. (2021) How complete are insect inventories? An assessment of the British butterfly database highlighting the influence of dynamic distribution shifts on sampling completeness. *Biodiversity and Conservation*, 30, 889–902.
- Segre, H., Kleijn, D., Bartomeus, I., WallisdeVries, M.F., DE Jong, M., Van Der Schee, M.F. et al. (2023) Butterflies are not a robust bioindicator for assessing pollinator communities, but floral resources offer a promising way forward. *Ecological Indicators*, 154, 110842.
- Soroye, P., Ahmed, N. & Kerr, J.T. (2018) Opportunistic citizen science data transform understanding of species distributions, phenology, and diversity gradients for global change research. *Global Change Biology*, 24, 5281–5291.
- Taylor, S.D., Meiners, J.M., Riemer, K., Orr, M.C. & White, E.P. (2019) Comparison of large-scale citizen science data and long-term study data for phenology modeling. *Ecology*, 100, e02568.
- van Tongeren, E., Sistri, G., Bonifacio, M., Menchetti, M., Pasquali, L., Salvati, V. et al. (2023) Unstructured citizen science reduces the perception of butterfly local extinctions: the interplay between species traits and user effort. *Biodiversity and Conservation*, 32, 4701–4718.
- Vanden Berghen, B., Vanermen, I. & Vranken, L. (2024) Citizen scientists: unveiling motivations and characteristics influencing initial and sustained participation in an agricultural project. *PLoS One*, 19, e0303103.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Figure S1. Supplementary phenology curves for butterfly species not shown in Figure 2, based on all available years of data. Curves compare structured monitoring (UKBMS) with opportunistic citizen science observations (iRecord Butterflies) across three regions (N, SE and SW). Each panel shows the daily percentage of peak number of records within a region. Arrows indicate the first and last day of the flight period among all available years.

Figure S2. Annual variation in the first and last days of the flight period for each butterfly species across three regions (N, SE and SW). Each panel represents one species and shows the yearly changes in the first and last days of the flight period based on UKBMS and iRecord Butterflies records.

Figure S3. Dates corresponding to the onset and end dates for butterfly species not shown in Figure 1 across the three regions (N, SE and SW) for UKBMS and iRecord Butterflies. Each panel shows the dates (days since 1 April) when the cumulative number of records reached 5% of the annual total (early flight period) and 95% of the annual total (late flight period).

Figure S4. Temporal trends in the weighted mean flight date for butterfly species not shown in Figure 4, comparing UKBMS with iRecord Butterflies across three regions (N, SE and SW). Each panel shows the annual weighted mean flight date for the species within a region, and dashed lines indicate linear trends for each dataset. Only years for

which both datasets had at least 5 years of overlapping observations were included in the plots.

Table S1. Full results corresponding to Table 4, with the additional columns 'p onset raw' and 'p end raw', which indicate the unadjusted p values for onset dates and end dates before FDR correction.

Table S2. Full table corresponding to Table 6, with the additional columns 'pearson p ' and 't p value', which indicate the unadjusted p values before FDR correction.

How to cite this article: Li, M., Bennie, J., Roy, D.B., Fox, R., Boyd, R.J., Smith, C. et al. (2026) Systematic differences in phenology estimates from unstructured and structured biodiversity datasets. *Insect Conservation and Diversity*, 19(5), 1246–1259. Available from: <https://doi.org/10.1111/icad.70098>