

# Species traits influence the process of biodiversity inventorying: a case study using the British butterfly database

JORGE M. LOBO,<sup>1</sup>  CAYETANO GUTIÉRREZ-CÁNOVAS,<sup>2</sup> RICHARD FOX,<sup>3</sup>   
ROGER L. H. DENNIS<sup>4,5,6</sup> and DAVID SÁNCHEZ-FERNÁNDEZ<sup>7</sup>  <sup>1</sup>Departamento de Biogeografía y Cambio Global, Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain, <sup>2</sup>Estación Biológica de Doñana (EBD-CSIC), Sevilla, Spain, <sup>3</sup>Butterfly Conservation, East Lulworth, Dorset, UK, <sup>4</sup>Centre for Ecology, Environment and Conservation, Department of Biological and Medical Sciences, Oxford Brookes University, Oxford, UK, <sup>5</sup>The School of Life Sciences and Education at Staffordshire University, Staffordshire University, Stoke-on-Trent, UK, <sup>6</sup>UK Centre for Ecology & Hydrology, Wallingford, Oxon, UK and <sup>7</sup>Departamento de Ecología e Hidrología, Universidad de Murcia, Murcia, Spain

**Abstract.** 1. The description of how biological information is compiled over time is essential to detect temporal biases in biodiversity data that could directly influence the utility, comparability, and reliability of ecological and biogeographical studies.

2. We explore trends in species recording over time using one of the most spatially and temporally comprehensive country-level databases for any group of insects in the world – the database of butterfly occurrences from Great Britain.

3. Firstly, we used two crucial milestones (the year in which the taxonomic inventory is complete, i.e., when the last species was recorded, the year in which all species are recorded together for the first time) to delimit three main phases in the process of biodiversity recording (taxonomic, faunistic and exhaustive phases). Secondly, we aimed to quantify how far species features (attractiveness and detectability) influence the process of recording through time.

4. During the first stage of biodiversity compilation, when the main aim is to complete the taxonomic inventory (taxonomic period), entomologists tend to record attractive species more frequently. However, once the inventory is complete, particularly in the period during which more spatially and temporally comprehensive information about species distribution is amassed (the exhaustive period), the recording pattern clearly changes to more detectable species.

5. Common, highly detectable species are undersampled in the first phase of biodiversity data compilation and oversampled in the final stages. Awareness of such temporal patterns in recording is necessary in order to correctly interpret and address bias in insect biodiversity trends.

**Key words.** Attractiveness, biodiversity database, detectability, insect monitoring, temporal biases.

## Introduction

Spatial and/or temporal biases and gaps in biodiversity data can directly influence the utility, comparability, and reliability

Correspondence: David Sánchez-Fernández, Departamento de Ecología e Hidrología, Facultad de Biología, Universidad de Murcia, Campus de Espinado, Murcia, 30100, Spain. E-mail: [davidsan@um.es](mailto:davidsan@um.es)

of ecological and evolutionary studies (Hortal *et al.*, 2007). Although growing concern about biodiversity loss underscores the need to quantify and understand temporal changes (Dornelas *et al.*, 2013; Tassarolo *et al.*, 2017; Cardoso & Leather, 2019; IPBES, 2019; Montgomery *et al.*, 2020), existing research on biodiversity data quality has typically focused on spatial biases in biodiversity databases (e.g., Pardo *et al.*, 2013; Ruete, 2015; Lobo *et al.*, 2018; Shirey *et al.*, 2021). The description

of how biological information is compiled over time is essential to understand which factors influence species recording processes through time and, thus, accurately quantify biodiversity trends (Cabrero-Sañudo & Lobo, 2003; Boakes *et al.*, 2010; Costello *et al.*, 2015; Isaac & Pocock, 2015).

Discriminating the successive stages of biodiversity recording using a high quality database helps to avoid, or at least to be aware of, future bias in recorded species traits in other less complete biodiversity databases. This knowledge would enable scientists and conservation organisations to assess the workforce needed to achieve reliable biodiversity information. It is especially important in the case of insects, which despite being one of the most diverse and functionally important animal groups, are greatly under-sampled worldwide (Leather *et al.*, 2008; Cardoso *et al.*, 2011). Indeed, to date, most studies of faunistic databases have reported a dearth of complete and extensive inventories for insect taxa (e.g., Romo *et al.*, 2006; Sánchez-Fernández *et al.*, 2008; Santos *et al.*, 2010; Bruno *et al.*, 2012; Ballesteros-Mejía *et al.*, 2013; Fattorini, 2013; Lobo *et al.*, 2018). Within insects, diurnal Lepidoptera are affected by undersampling to a lesser degree than other taxa (Troutet *et al.*, 2017), in all likelihood due to their relatively large size and aesthetic appeal, and, as a result, are a particularly useful taxonomic group for investigating patterns of biodiversity recording.

In this study, we explore trends in how species are recorded over time to evaluate temporal biases in biodiversity data compilation using one of the most spatially and temporally comprehensive country-level databases for any group of insects in the world (Sánchez-Fernández *et al.*, 2021) – the Butterflies for the New Millennium (BNM) database of Great Britain (GB) (Fox *et al.*, 2015). This database contains more than 10 million species occurrence records gathered during more than 200 years for a butterfly fauna of 58 current breeding butterfly species. Dennis *et al.* (2006) found that more apparent butterflies were recorded earlier than those that are less conspicuous. Here, we go further by examining the temporal consistency and heterogeneity of the BNM database. We aim to identify some of the main foci and tendencies of entomologists and recorders driving the biodiversity inventory process by examining temporal trends in the number of records in relation to species traits. Specifically, our aims are to delimit the main periods in the process of biodiversity inventory and to quantify the relative importance through time of species traits associated with attractiveness and detectability.

## Methods

We analysed data extracted from the BNM database managed by Butterfly Conservation. The BNM database comprises butterfly occurrence records (unique combinations of species  $\times$  recorder  $\times$  location  $\times$  date) for the United Kingdom, Isle of Man and Channel Islands, many of which are the result of opportunistic, non-standardised sampling by community scientists (verified by a network of expert volunteers). The BNM database was created in 1995 but incorporates a substantial volume of historical records, particularly those gathered by a previous recording scheme that led to

the first butterfly atlas of Britain and Ireland (Heath *et al.*, 1984). The extract used for this study contained all records of the 58 breeding species in GB from 1800 to 2014 (see Fox *et al.*, 2015; Sánchez-Fernández *et al.*, 2021). Species that occur only as occasional, usually non-breeding, immigrants were excluded. Species that have become extinct in GB were also removed (including the reintroduced *Phengaris arion*) so that temporal trends could be constructed for species observed throughout the complete study period, although the data for *Nymphalis polychloros* were retained since it may have recently recolonised GB naturally after several decades of absence (see Fig. S1 in the Supporting Information). The final extract contained a total of 10 046 366 records (see Data S1 in the Supporting Information).

## Measures of butterfly attractiveness and detectability

Two approaches utilising species traits were used to assess data collection over time, here referred to as ‘attractiveness’ and ‘detectability’ (see Data S2 in the Supporting Information). Attractiveness values were obtained from Dennis *et al.* (2006). These authors used a scoring system in which the mean subjective perception of visual apparency (conspicuousness) was determined independently by 13 researchers who work on butterflies; each researcher assigned an apparency score ranging 1 (low conspicuousness) to 5 (high conspicuousness) to each species, and the results were calibrated against image analysis of the wing surfaces for a subset of the species (males,  $n = 50$ , Pearson correlation coefficient  $r = 0.89$ ; females,  $n = 42$ ,  $r = 0.87$ ;  $P < 0.0001$ ) (for details, see Dennis *et al.*, 2006).

To account for butterfly detectability, we compiled information on several trait categories that account for species functional and ecological characteristics that can be used as proxies for detectability (see Supporting Information Table S1): larval diet breadth, adult feeding, type of roost and rest sites, overwintering life cycle stage, voltinism, mobility (flight period and mobility capacity), and occurrence in different habitats (Dennis, 2010; Middleton-Welling *et al.*, 2018). These detectability traits included continuous, ordinal, multiple-choice and fuzzy-coded traits and comprise 55 specific categories (see Supporting Information Table S1). For multiple- and fuzzy-coded traits, different states were organised into various columns representing different affinity categories (e.g. Overwintering phase was divided into ‘egg’, ‘larvae’, ‘pupae’ and ‘adult’ categories). For multiple-coded traits, one or multiple options were assigned for each species in a binary way. For fuzzy coded-traits, affinity values were distributed across the categories of each trait for each species, according to the frequency of occurrence within the species. This approach is called fuzzy coding (Chevenet *et al.*, 1994) and entails compiling the intraspecific biological information available for each species. Before analysing data, multiple- and fuzzy-coded data were converted to percentages of affinity for each trait to obtain a standardised representation. To obtain a single value of detectability of each species, we summarised the trait space using a principal coordinate analysis (PCoA) selecting the first axis as a surrogate of detectability. This first axis explained 27.2% of the total variability (see the

factor scores of each trait in the Supporting Information Table S1). Gower's coefficient was used since it provides trait-based species dissimilarity calculations when binary, numerical and categorical attributes are considered simultaneously (Pavoine *et al.*, 2009; Maire *et al.*, 2015). Positive values of this factor are associated with highly detectable species, i.e., mobile, multivoltine, polyphagous species that use a broad range of nectar sources and occur in highly modified biotopes (e.g. gardens, arable crops and road banks).

### Data analysis

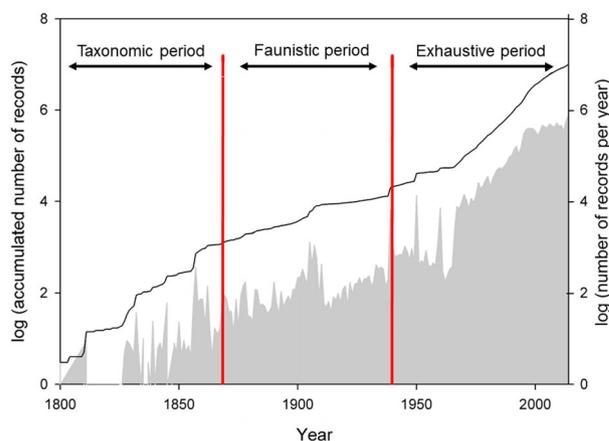
For each year, we calculated both the number of records collected and the number of species recorded. We also estimated the cumulative number of butterfly species observed in consecutive years, as well as the proportion gained annually over the total number of butterfly species. These data were used to delimit two main temporal points that could be considered crucial milestones to define phases in the process of biodiversity recording: (i) the year in which the taxonomic inventory of GB is complete, i.e., this period ends when the last species was recorded and (ii) the year in which all the species (58 butterflies) were recorded (together) for the first time. These two years delimit three key temporal phases in any biodiversity database: a first period that relates to the establishment of the national inventory (hereafter 'taxonomic period'), a second in which species occurrence data are accumulated but at a relatively slow rate, mainly devoted to establishing the distributional range of each species (hereafter 'faunistic period'), and a third period in which more spatially and temporally comprehensive information about species distribution is amassed (hereafter 'exhaustive period').

Multiple linear regressions were used to quantify the influence of attractiveness and detectability in determining the number of records accumulated for each species in each of the main periods of biodiversity inventory and over the whole time series. Before analysis, we log-transformed response variables (records collected during taxonomic, exhaustive and total periods) to reduce skewness. We displayed back-transformed fitted values to represent relationships between variables better. For each model, we obtained partitioned variance for detectability and attractiveness using the R package variancePartition (Hoffman & Schadt, 2016). This package calculates variance fractions based on the sum of squares explained by each predictor, i.e. the total amount of variance explained by each factor (unique plus attributable shared fraction). All regression models were validated by visually checking their residuals for normality and homoscedasticity.

All analyses were conducted using the R version 3.5.3 (R Development Core Team, 2019).

## Results

The BNM database extract contained butterfly species occurrence records across a period of 215 years. Most records are recent – as 50% were gathered during the last 11 years of the



**Fig. 1.** Historical variation in the number (log transformed) of database records per year (grey surface) and accumulated number of records of butterflies in Great Britain (black line). The three main periods in the biodiversity inventorying process are highlighted (red lines).

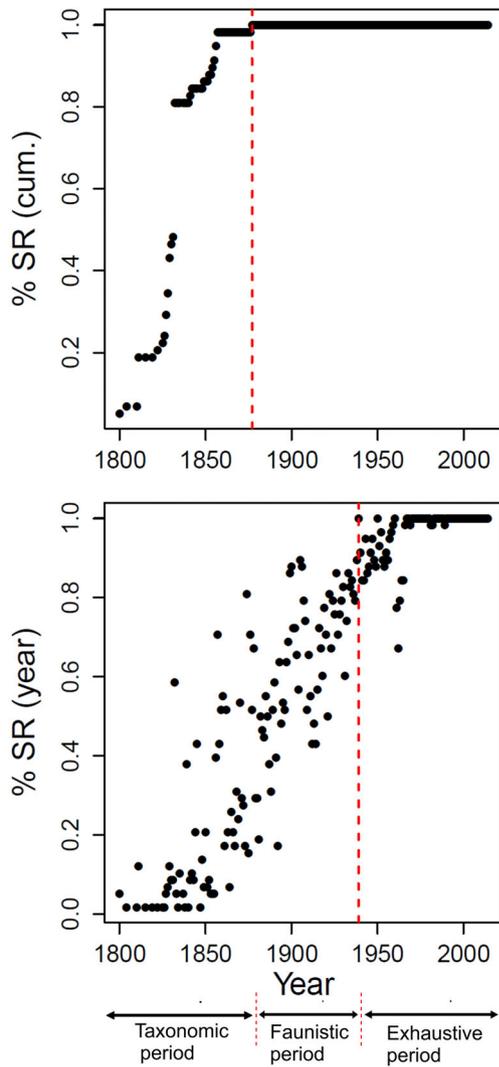
study period (2004–2014) (Fig. 1; Fig. S1 in the Supporting Information).

### Main periods in the process of biodiversity inventory

The complete GB butterfly inventory defined for this study was established in 1877 (Figs. 1 and 2). The last species to be recorded, *Thymelicus lineola*, was not recognised until 1889, but earlier specimens were subsequently identified in collections. The first year in which all the GB butterfly species were recorded at the same time was 1939 (Figs. 1 and 2). Thus, the three periods were delimited as follow: the 'taxonomic period' (up to 1877), the 'faunistic period' (from 1878 to 1939), and finally, the 'exhaustive period' period (from 1940 to 2014).

### Bias related to species features captured over time

Detectability was the most important variable in explaining the total number of records collated for each species ( $R^2 = 55.4\%$ ;  $P < 0.01$ ), with the number of records rising exponentially with increasing detectability (Fig. 3). On the other hand, butterfly attractiveness had a very low ( $R^2 = 0.2\%$ ;  $P = 0.66$ ) association with the cumulative number of records per species (see Fig. 3). However, the importance of detectability and attractiveness in explaining the variability in the number of records of each species changed over time (see Table 1 and Fig. 3). In the taxonomic period, attractiveness showed a weak, but significant, positive relationship with the total number of records of each species in the database ( $R^2 = 6.5\%$ ;  $P < 0.05$ ). In contrast, detectability showed a weak, significant negative correlation with the number of records ( $R^2 = 4.2\%$ ;  $P < 0.05$ ). After the taxonomic period, no significant relationships were found between attractiveness and number of records. There

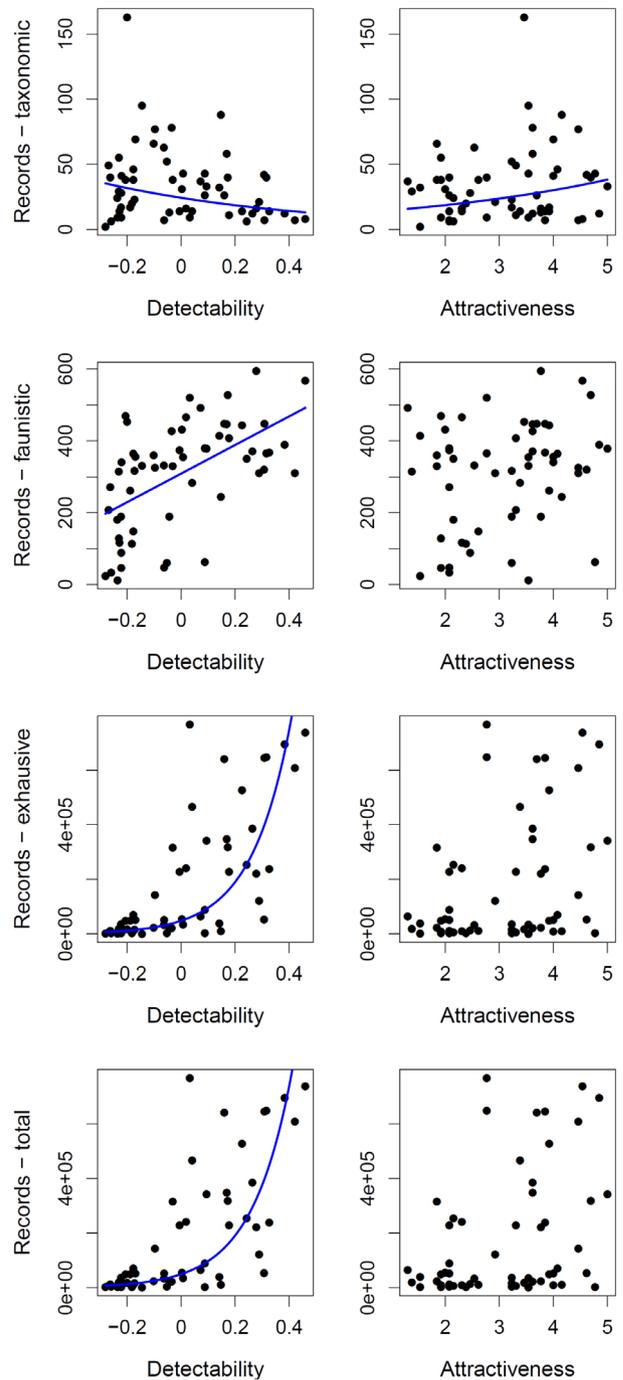


**Fig. 2.** Percentage of total species richness (SR) accumulated (above) and per year (below) included in the database. Red dotted lines indicate the first year in which each variable reached the 100%.

was a relatively strong positive linear relation between detectability and the number of records in the faunistic period ( $R^2 = 30.9\%$ ;  $P < 0.01$ ), a relationship that increased in strength during the exhaustive period ( $R^2 = 54.7\%$ ;  $P < 0.01$ ) (see Table 1; Fig. 3).

## Discussion

The butterflies of Great Britain are probably the best-studied insect group in the world with the largest amount of data per species and the longest record of observations. The temporal patterns derived from the study of such an intensively recorded, long-term dataset allow us to discriminate several temporal phases in which the attractiveness and detectability traits of



**Fig. 3.** Relationship among species traits (detectability and attractiveness) and the sampling effort (number of records) carried out in each one of the three periods of the database inventory process (taxonomic, faunistic and exhaustive) and for the total time period considered. Blue lines represent fitted values for significant relationships ( $P$ -value  $\leq 0.05$ ; see Table 1).

species variably affect recording. Our results highlight the importance of species characteristics in the historical process of biodiversity inventorying and may shed light on the future

**Table 1.** Results of the models examining the effect of species traits (detectability and attractiveness) on the sampling effort (number of records) carried out in each one of the three main periods of the database compiling.

Main period	Model terms	Coefficient	SE	<i>t</i> -value	<i>P</i> -value	Partitioned <i>R</i> <sup>2</sup>
Taxonomic	Intercept	3.19	0.11	29.405	<0.0001	
	Detectability	−0.29	0.12	−2.336	0.02	4.2%
	Attractiveness	0.24	0.12	2.004	0.05	6.5%
Faunistic	Intercept	308.64	16.56	18.640	<0.0001	
	Detectability	84.08	18.66	4.505	<0.0001	30.9%
	Attractiveness	−2.62	18.66	−0.140	0.89	0.0%
Exhaustive	Intercept	10.79	0.17	63.418	<0.0001	
	Detectability	1.44	0.19	7.532	<0.0001	54.7%
	Attractiveness	−0.09	0.19	−0.486	0.63	0.2%
Total	Intercept	10.81	0.17	64.931	<0.0001	
	Detectability	1.43	0.19	7.608	<0.0001	55.4%
	Attractiveness	−0.08	0.19	−0.444	0.66	0.2%

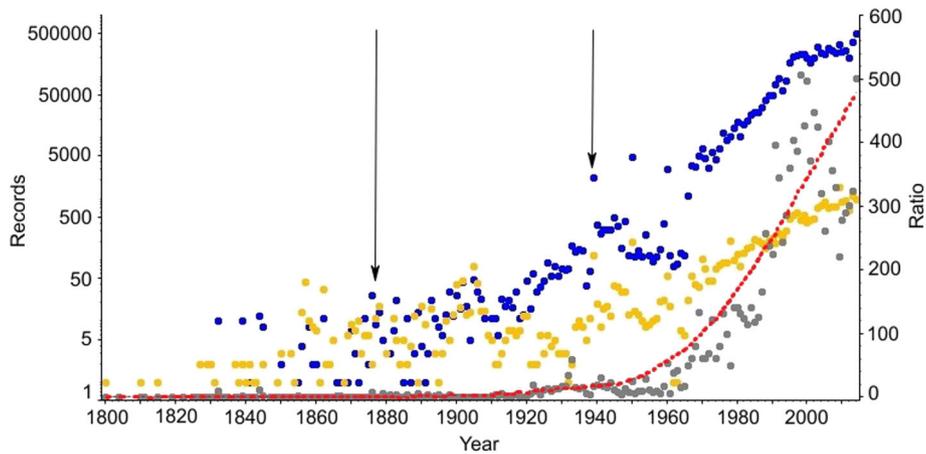
prospects of initiatives compiling information about other (including hyperdiverse) taxonomic groups around the world.

We chose two different points in time to exemplify the main phase changes in the process of biodiversity inventory. Using these time points, we identified three distinct periods in the process of accumulating butterfly biodiversity knowledge in GB. An initial taxonomic phase in which the efforts of naturalists were focussed on the discovery of new uncollected species. Although this process is typically unplanned and uncoordinated, the result of the individual actions of naturalists has the effect of making the discovery of new species more and more difficult. The attractiveness and charismatic character of the species seem to exert a moderate role in recording during this phase. This period is followed by another in which the effort of the naturalists is progressively directed towards gathering more and more information about the distribution of the species (biogeography after taxonomy). Here, the conspicuousness of different species does not seem to influence recording. Instead, detectability, assessed using species traits as proxies, becomes highly relevant. Thus, a greater number of records are gathered for polyphagous species, those that are bivoltine or multivoltine, have a high dispersal ability, and frequently occur in highly modified land-use types (such as road and river banks, gardens and arable crops) (see Supporting Information Table S1). Finally, a third phase of exhaustive study (intensive surveying, encouraged by organised recording schemes) can be differentiated, in which species detectability is even more strongly related with the number of records in each species, as occurs in the selection of species for conservation programmes (Gunnthorsdottir, 2001; Leandro *et al.*, 2017). In the taxonomic period, the most recorded species were *Euphydryas aurinia*, *Nymphalis polychloros*, *Leptidea sinapis* and *Boloria euphrosyne*, i.e. species with a relatively high attractiveness, while in the exhaustive period, the species accumulating the greatest numbers of records are *Maniola jurtina*, *Pieris rapae*, *P. brassicae*, *Pararge aegeria*, *Aglais urticae*, *A. io*, the most, abundant, widespread and easily detectable species.

Our results indicate a change in the recording process once the national inventory is taxonomically complete. It can be reasonably expected that the progressive difficulty of finding new species in a country or region and the accumulation of taxonomic

resources (keys, collections, journals, taxonomical revisions, etc.) would increasingly facilitate efforts towards the completion of regional inventories (Brunbjerg *et al.*, 2019). The variables used here as surrogates of detectability positively correlate very closely with the geographical range, the occurrence on islands, and the population size of GB butterfly species (Dennis *et al.*, 2000, 2004). Thus, after a first phase of progressive increase in taxonomic knowledge not influenced by the differential detectability of the species, a second phase begins in which the data accumulation is biased towards the most detectable species. The accumulation of observations in ‘rare’ or ‘common’ species could be an indicator of the stage of the biodiversity recording process. Identifying ‘common’ and ‘rare’ species according to their 90th or 10th percentile distribution in the number of records and occupied 5’ cells, leads to the observation that the ratio between the records of both types of species increases notably throughout the time period studied (Fig. 4).

At the end of taxonomic period, the mean number of records by species in the database was 34, and at the end of the faunistic period was 434.2. By 2014, this value had increased to 173 213. The faunistic and exhaustive phases could be difficult to discriminate and could in fact be a systematic process in which particular idiosyncratic events can be determined (Lee *et al.*, 2005). For example, there is a quirk relating to 1939 in the BNM data. When historical data were being added during the 1970s and 1980s, sources that described species occurrence as ‘before the war’ were assigned to the year 1939. Thus 1939 could be an outlier in terms of species occurrence records that determine, in our case, the first year in which all British species were recorded. However, this fact does not greatly change recording trends as the following years showed similar results. On the other hand, even in this case, the taxonomic period could not be quite over, considering the impact that new molecular techniques are having on butterfly taxonomy in Europe (Dinca *et al.*, 2011; Hernández-Roldan *et al.*, 2016). Indeed, Prof T.G. Shreeve and colleagues may well have found a unique *Polyommatus* lineage in the Outer Hebrides (Arif *et al.*, 2020). Also, we expect new butterfly species to arrive in GB with climate change (Harrison, *et al.* 2006; Sánchez-Fernández *et al.*, 2021). Be that as it may, temporal changes in the common/rare ratio of species seem to suggest that



**Fig. 4.** Temporal changes in the logarithm of the number of records of ‘common’ (blue circles) and ‘rare’ (orange circles) butterfly species in Great Britain, and temporal change in the ratio of common/rare species (grey circles). Common species are defined as those above the 90th percentile in the number of records or in the number of 5’ occurrence cells considering the whole data. Rare species are those below 10th percentile in these same variables. The dashed line represents the LOWESS fit of the ratio variable. Arrows indicate the two main temporal points delimited in the process of biodiversity recording.

faunistic and exhaustive phases could be discriminated by the existence of a turning point in this ratio which marks the beginning of the recording increase of ‘common’ (i.e. abundant and widespread) species against ‘rare’ ones. Analyses of other species occurrence data sets are required to determine the usefulness of this pattern in identifying the level of biodiversity knowledge in a country or region.

Recently, the evidence and potential consequences of declines in insect biodiversity has aroused a great scientific, media and public attention (Goulson, 2019; Habel *et al.*, 2019; Cardoso *et al.*, 2020; Warren *et al.*, 2021), with some studies calling for immediate policy responses (Forister *et al.*, 2019; Harvey *et al.*, 2020). However, noting the heterogeneity of insect responses, others studies have highlighted the need for more data (Montgomery *et al.*, 2020; Wagner *et al.*, 2021) to avoid risky overextrapolation from the limited current evidence (Didham *et al.*, 2020, Saunders *et al.*, 2020). In order to estimate insect trends accurately and thus better to understand the full extent of global biodiversity change, it is essential to know the current state of each biodiversity database as some comparative results could be simply a consequence of sampling bias. Here, we demonstrate that common, highly detectable species are undersampled in the first phase of biodiversity data compilation and oversampled in the final stages. From a general point of view, it should be noted that currently the great majority biodiversity databases in the world are still in the initial stages of data compilation (Ball-Damerow *et al.*, 2019), so these distributional databases, especially in the case of insects, are very likely in the taxonomic or at best, in the faunistic periods of data collection. Awareness of such temporal patterns in recording is necessary in order to interpret correctly and address bias in insect biodiversity trends.

### Acknowledgements

The authors are very grateful to all of the volunteers who have contributed to the Butterflies for the New Millennium project,

which is run by Butterfly Conservation with support from Natural England. D.S.-F. and C.G.-C are funded by postdoctoral contracts from the Spanish Ministry of Science and Innovation by the Ramón y Cajal (RYC2019-027446-I) and Juan de la Cierva Incorporación (IJC2018-036642-I) programmes, respectively.

### Data availability statement

Data available on request from the authors

### Supporting information

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

**Data S1:** Supporting Information

**Data S2:** Supporting Information

**Appendix S1:** Supporting Information

### REFERENCES

- Arif, A., Gerth, M., Hone-Millard, W.G., Nunes, M.D.S., Dapporto, L. & Shreeve, T.G. (2020) Multiple colonisations and *Wolbachia* infections influence the fine-scale population structure of the widespread Common Blue butterfly (*Polyommatus icarus*) in the British Isles. *bioRxiv*, 2020.09.03.267203.
- Ballesteros-Mejia, L., Kitching, I.J., Jetz, W., Nagel, P. & Beck, J. (2013) Mapping the biodiversity of tropical insects: species richness and inventory completeness of African sphingid moths. *Global Ecology and Biogeography*, **22**, 586–595.
- Ball-Damerow, J.E., Brenskelle, L., Barve, N., Soltis, P.S., Sierwald, P., Bieler, R., LaFrance, R., Ariño, A.H. & Guralnick, R.P. (2019)

- Research applications of primary biodiversity databases in the digital age. *PLoS One*, **14**, e0215794.
- Boakes, E.H., McGowan, P.J., Fuller, R.A., Chang-Qing, D., Clark, N. E., O'Connor, K. & Mace, G.M. (2010) Distorted views of biodiversity: spatial and temporal bias in species occurrence data. *PLoS Biology*, **8**, e1000385.
- Brunbjerg, A.K., Bruun, H.H., Brøndum, L., Classen, A.T., Dalby, L., Fog, K., Frøslev, T.G., Goldberg, I., Hansen, A.J., Hansen, M.D.D., Høye, T.T., Illum, A.A., Læssøe, T., Newman, G.S., Skipper, L., Søchting, U. & Ejrnæs, R. (2019) A systematic survey of regional multi-taxon biodiversity: evaluating strategies and coverage. *BMC Ecology*, **19**, 43.
- Bruno, D., Sánchez-Fernández, D., Millán, A., Ros, R.S., Sánchez-Gomez, P. & Velasco, J. (2012) Assessing the quality and usefulness of different taxonomic groups inventories in a semi-arid Mediterranean region. *Biodiversity and Conservation*, **21**, 1561–1575.
- Cabrero-Sañudo, F.J. & Lobo, J.M. (2003) Estimating the number of species not yet described and their characteristics: the case of Western Palaearctic dung beetle species (Coleoptera, Scarabaeoidea). *Biodiversity and Conservation*, **12**, 147–166.
- Cardoso, P., Erwin, T.L., Borges, P.A.V. & New, T.R. (2011) The seven impediments in invertebrate conservation and how to overcome them. *Biological Conservation*, **144**, 2647–2655.
- Cardoso, P. & Leather, S.R. (2019) Predicting a global insect apocalypse. *Insect Conservation and Diversity*, **12**, 263–267.
- Cardoso, P., Barton, P.S., Birkhofer, K., Chichorro, F., Deacon, C., Fartmann, T., Fukushima, C.S., Gaigher, R., Habel, J.C., Hallmann, C.A., Hill, M.J., Hochkirch, A., Kwak, M.L., Mammola, S., Noriega, J.A., Orfinger, A.B., Pedraza, F., Pryke, J.S., Roque, F.O., Settele, J., Simaika, J.P., Stork, N.E., Suhling, F., Vorster, C. & Samways, M.J. (2020) Scientists' warning to humanity on insect extinctions. *Biological Conservation*, **242**, 108426.
- Chevenet, F., Dolédec, S. & Chessel, D. (1994) A fuzzy coding approach for the analysis of long-term ecological data. *Freshwater Biology*, **31**, 295–309.
- Costello, M.J., Lane, M., Wilson, S. & Houlding, B. (2015) Factors influencing when species are first named and estimating global species richness. *Global Ecology and Conservation*, **4**, 243–254.
- Dennis, R.L.H., Donato, B., Sparks, T.H. & Pollard, E. (2000) Ecological correlates of island incidence and geographical range among British butterflies. *Biodiversity and Conservation*, **9**, 343–359.
- Dennis, R.L.H., Hodgson, J.G., Grenyer, R., Shreeve, T.G. & Roy, D.B. (2004) Host plants and butterfly biology. Do host plant strategies drive butterfly status? *Ecological Entomology*, **29**, 12–26.
- Dennis, R.L.H. (2010) *A Resource-Based Habitat View for Conservation. Butterflies in the British Landscape*. Wiley-Blackwell, Oxford, UK.
- Dennis, R.L.H., Shreeve, T.G., Isaac, N.J.B., Roy, D.B., Hardy, P.B., Fox, R. & Asher, J. (2006) The effects of visual apparency on bias in butterfly recording and monitoring. *Biological Conservation*, **128**, 486–492.
- Didham, R.K., Basset, Y., Collins, C.M., Leather, S.R., Littlewood, N.A., Menz, M.H.M., Müller, J., Packer, L., Saunders, M.E., Schönrogge, K., Stewart, A.J.A., Yanoviak, S.P. & Hassall, C. (2020) Interpreting insect declines: seven challenges and a way forward. *Insect Conservation and Diversity*, **13**, 103–114.
- Dincă, V., Lukhtanov, V., Talavera, G. & Vila, R. (2011) Unexpected layers of cryptic diversity in wood white Leptidea butterflies. *Nature Communications*, **2**, 324.
- Domelas, M., Magurran, A.E., Buckland, S.T., Chao, A., Chazdon, R.L., Colwell, R.K., Curtis, T., Gaston, K.J., Gotelli, N.J., Kosnik, M.A., McGill, B., McCune, J.L., Morlon, H., Mumby, P.J., Ovreas, L., Studený, A. & Vellend, M. (2013) Quantifying temporal change in biodiversity: challenges and opportunities. *Proceedings of the Royal Society B*, **280**, 20121931.
- Fattorini, S. (2013) Regional insect inventories require long time, extensive spatial sampling and good will. *PLoS ONE*, **8**, 4.
- Forister, M.L., Pelton, E.M. & Black, S.H. (2019) Declines in insect abundance and diversity: we know enough to act now. *Conservation Science and Practice*, **1**, e80.
- Fox, R., Brereton, T.M., Asher, J., August, T.A., Botham, M.S., Bourn, N.A.D., Cruickshanks, K.L., Bulman, C.R., Ellis, S., Harrower, C.A., Middlebrook, I., Noble, D.G., Powney, G.D., Randle, Z., Warren, M.S. & Roy, D.B. (2015) *The State of the UK's Butterflies 2015*. Butterfly Conservation and the Centre for Ecology & Hydrology, Wareham, UK.
- Goulson, D. (2019) The insect apocalypse, and why it matters. *Current Biology*, **29**, R967–R971.
- Gunthorsdottir, A. (2001) Physical attractiveness of an animal species as a decision factor for its preservation. *Anthrozoös*, **14**, 204–215.
- Habel, J.C., Samways, M.J. & Schmitt, T. (2019) Mitigating the precipitous decline of terrestrial European insects: requirements for a new strategy. *Biodiversity and Conservation*, **28**, 1343–1360.
- Harrison, P.A., Berry, P.M., Butt, N. & New, M. (2006) Modelling climate change impacts on species' distributions at the European scale: implications for conservation policy. *Environmental Science and Policy*, **9**, 116–128.
- Harvey, J.A., Heinen, R., Armbrrecht, I., Basset, Y., Baxter-Gilbert, J.H., Bezemer, T.M., Böhm, M., Bommarco, R., Borges, P.A.V., Cardoso, P., Clausnitzer, V., Cornelisse, T., Crone, E.E., Dicke, M., Dijkstra, K.D.B., Dyer, L., Ellers, J., Fartmann, T., Forister, M.L., Furlong, M.J., Garcia-Aguayo, A., Gerlach, J., Gols, R., Goulson, D., Habel, J.C., Haddad, N.M., Hallmann, C.A., Henriques, S., Herberstein, M.E., Hochkirch, A., Hughes, A.C., Jepsen, S., Jones, T.H., Kaydan, B.M., Kleijn, D., Klein, A.M., Latty, T., Leather, S.R., Lewis, S.M., Lister, B.C., Losey, J.E., Lowe, E.C., Macadam, C.R., Montoya-Lerma, J., Nagano, C.D., Ogan, S., Orr, M.C., Painting, C.J., Pham, T.H., Potts, S.G., Rauf, A., Roslin, T.L., Samways, M.J., Sanchez-Bayo, F., Sar, S.A., Schultz, C.B., Soares, A.O., Thancharoen, A., Tscharnke, T., Tylianakis, J.M., Umbers, K.D.L., Vet, L.E.M., Visser, M.E., Vujic, A., Wagner, D.L., WallisDeVries, M.F., Westphal, C., White, T.E., Wilkins, V.L., Williams, P.H., Wyckhuys, K.A.G., Zhu, Z.R. & de Kroon, H. (2020) International scientists formulate a roadmap for insect conservation and recovery. *Nature Ecology & Evolution*, **4**, 174–176.
- Heath, J., Pollard, E. & Thomas, J.A. (1984) *Atlas of Butterflies in Britain and Ireland*. Viking, Harmondsworth, UK.
- Hernández-Roldán, J.L., Dapporto, L., Dincă, V., Vicente, J.C., Hornett, E.A., Šíchová, J., Lukhtanov, V.A., Talavera, G. & Vila, R. (2016) Integrative analyses unveil speciation linked to host plant shift in *Spialia* butterflies. *Molecular Ecology*, **25**, 4267–4284.
- Hoffman, G.E. & Schadt, E.E. (2016) variancePartition: interpreting drivers of variation in complex gene expression studies. *BMC Bioinformatics*, **17**, 483.
- Hortal, J., Lobo, J.M. & Jiménez-Valverde, A. (2007) Limitations of biodiversity databases: case study on seed-plant diversity in Tenerife (Canary Islands). *Conservation Biology*, **21**, 853–863.
- IPBES (2019) *Summary for policymakers of the global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services*. IPBES Secretariat, Bonn, Germany.
- Isaac, N.J. & Pocock, M.J. (2015) Bias and information in biological records. *Biological Journal of the Linnean Society*, **115**, 522–531.

- Leandro, C., Jay-Robert, P. & Vergnes, A. (2017) Bias and perspectives in insect conservation: a European scale analysis. *Biological Conservation*, **215**, 213–224.
- Leather, S.R., Basset, Y. & Hawkins, B. (2008) Insect conservation: finding the way forward. *Insect Conservation and Diversity*, **1**, 67–69.
- Lee, W., McGlone, M. & Wright, E. (2005) Biodiversity inventory and monitoring: a review of national and international systems and a proposed framework for future biodiversity monitoring by the Department of Conservation. Landcare Research Contract Report LC0405/122.
- Lobo, J.M., Hortal, J., Yela, J.L., Millán, A., Sánchez-Fernández, D., García-Roselló, E., González-Dacosta, J., Heine, J., González-Vilas, L. & Guisande, C. (2018) KnowBR: an application to map the geographical variation of survey effort and identify well-surveyed areas from biodiversity databases. *Ecological Indicators*, **91**, 241–248.
- Maire, E., Grenouillet, G., Brosse, S. & Villéger, S. (2015) How many dimensions are needed to accurately assess functional diversity? A pragmatic approach for assessing the quality of functional spaces. *Global Ecology and Biogeography*, **24**, 728–740.
- Middleton-Welling, J., Wade, R.A., Dennis, R.L.H., Dapporto, L. & Shreeve, T.G. (2018) Optimising trait and source selection for explaining occurrence and abundance changes: a case study using British Butterflies. *Functional Ecology*, **32**, 1609–1619.
- Montgomery, G.A., Dunn, R.R., Fox, R., Jongejans, E., Leather, S.R., Saunders, M.E., Shortall, C.R., Tingley, M.W. & Wagner, D.L. (2020) Is the insect apocalypse upon us? How to find out. *Biological Conservation*, **241**, 108327.
- Pardo, I., Pata, M.P., Gómez, D. & García, M.B. (2013) A novel method to handle the effect of uneven sampling effort in biodiversity databases. *PLoS One*, **8**, e52786.
- Pavoine, S., Vallet, J., Dufour, A.-B., Gachet, S. & Daniel, H. (2009) On the challenge of treating various types of variables: application for improving the measurement of functional diversity. *Oikos*, **118**, 391–402.
- R Development Core Team (2019) *R Development Core Team. R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Retrieved from, Vienna, Austria. <https://www.r-project.org/>. Accessed on May 2020.
- Romo, H., Garcia-Barros, E. & Lobo, J.M. (2006) Identifying recorder-induced geographic bias in an Iberian butterfly database. *Ecography*, **29**, 873–885.
- Ruete, A. (2015) Displaying bias in sampling effort of data accessed from biodiversity databases using ignorance maps. *Biodiversity Data Journal*, **3**, e5361.
- Sánchez-Fernández, D., Fox, R., Dennis, R.L.H. & Lobo, J.M. (2021) How complete are insect inventories? An assessment of the British butterfly database highlighting the influence of distribution shifts due to climatic changes. *Biodiversity and Conservation*, **30**, 889–902.
- Sánchez-Fernández, D., Lobo, J.M., Abellán, P., Ribera, I. & Millán, A. (2008) Bias in freshwater biodiversity sampling: the case of Iberian water beetles. *Diversity and Distributions*, **14**, 754–762.
- Santos, A.M.C., Jones, O.R., Quicke, D.L.J. & Hortal, J. (2010) Assessing the reliability of biodiversity databases: identifying evenly inventoried island parasitoid faunas (Hymenoptera: Ichneumonoidea) worldwide. *Insect Conservation and Diversity*, **3**, 72–82.
- Saunders, M.E., Janes, J.K. & O'Hanlon, J.C. (2020) Moving On from the insect apocalypse narrative: engaging with evidence-based insect conservation. *BioScience*, **70**, 80–89.
- Shirey, V., Belitz, M.W., Barbe, V. & Guralnick, R. (2021) A complete inventory of North American butterfly occurrence data: narrowing data gaps, but increasing bias. *Ecography*, **44**, 1–11.
- Tessarolo, G., Ladle, R., Rangel, T. & Hortal, J. (2017) Temporal degradation of data limits biodiversity research. *Ecology and Evolution*, **7**, 6863–6870.
- Troutet, J., Grandcolas, P., Blin, A., Vignes-Lebbe, R. & Legendre, F. (2017) Taxonomic bias in biodiversity data and societal preferences. *Scientific Reports*, **7**, 2045–2322.
- Wagner, D.L., Fox, R., Salcido, D.M. & Dyer, L.A. (2021) A window to the world of global insect declines: moth biodiversity trends are complex and heterogeneous. *Proceedings of the National Academy of Sciences*, **118**, e2002549117.
- Warren, M.S., Maes, D., van Swaay, C.A.M., Goffart, P., Van Dyck, H., Bourn, N.A.D., Wynhoff, I., Hoare, D. & Ellis, S. (2021) The decline of butterflies in Europe: problems, significance, and possible solutions. *Proceedings of the National Academy of Sciences*, **118**, e2002551117.

Accepted 28 May 2021

Editor: Raphael Didham; Associate Editor: Manu Saunders