UNDERSTANDING THE COMPLEXITIES OF DATA COMPILED BY RECORDING SCHEMES

STUART G. BALL[1], ROGER K. A. MORRIS[2], STEPHEN T. BUCKLAND[3] & RICHARD GLENNIE[3]

[1] *Hoverfly Recording Scheme: 255 Eastfield Road, Peterborough PE1 4BH*
[2] *Hoverfly Recording Scheme: 241 Commonside East, Mitcham, Surrey CR4 1HB*
syrphid58@gmail.com
[3] *The Observatory, Buchanan Gardens, University of St Andrews, St Andrews KY16 9LZ, UK*

ABSTRACT

This paper reviews the data within the British Hoverfly Recording Scheme and issues that research groups should consider when interpreting outputs from opportunistic data. We show, how the composition of the dataset has changed over time, and the effect that this change can have on modelling. List length emerges as an important attribute and we emphasise the need to encourage observers to record all species encountered. We also show how very large numbers of occasional records can lead to unreliable trends.

We recommend that there should be sensitivity testing of opportunistic data to assess the influence of list composition. Data for uplands and the early/late parts of the season will usually comprise very short lists and may need to be included, whereas lists for lowland Britain during peak months could be a lot longer. Whilst the use of longer lists (five or more) may need to be assessed to fully understand the implications of data composition, we think that it is more useful to exclude data from recorders who report very small numbers of species (but potentially large numbers of records). Furthermore, we suggest that a register of major data perturbations should be maintained, in order that analysts are fully aware of the possible implications that they may have.

INTRODUCTION

Demand for accurate data on biodiversity to inform and implement biodiversity policy is increasing (Wetzel *et al.*, 2015) but there is a considerable gap between need and availability (Geijzendorffer *et al.*, 2015). There are relatively few long-term monitoring programmes, targeted at a limited range of taxa, principally birds and macro-Lepidoptera. Without long-term monitoring schemes, interest has turned to opportunistic data which, in Britain, are compiled by 'recording schemes' (Pocock *et al.*, 2015a). These schemes harness the enthusiasm of natural historians, but avoid making off-putting and counter-productive demands on them (Pocock~~,~~ *et al.*, 2015b). The absence of a structured protocol means that these are 'presence only' data and any absences from the dataset can only be inferred (Hastie & Fithian, 2013).

Recording schemes were originally designed to map the distribution of their target organisms, but the data have subsequently been used to investigate wildlife trends (e.g. Biesmeir *et al.*, 2006; Powney *et al.*, 2019; Outhwaite *et al.*, 2020) and phenological changes (Hassall, Owen & Gilbert, 2017). They have also been used to investigate climate change impacts (Oliver *et al.*, 2015; Morris & Ball, 2019a). Trend analysis is largely based upon models such as Frescalo (Hill, 2012) and works on the assumption that models will smooth out data heterogeneity (e.g. van Strien, van Swaay, and Termaat, 2013; Isaac *et al.*, 2014). Each new analysis highlights subtly different aspects of a changing environment but, in so doing, some also identify contradictory results (e.g. Outhwaite *et al*., 2020; Hassall *et al*., 2017).

Some conflicting results may be explained by the ways in which data have been gathered, and in the way that biological recording in Britain has evolved. Until now, however, the complexities within the data, and their implications for analysts may not have been fully appreciated, even though Isaac *et al*. (2014) and Isaac & Pocock (2015) touch upon some of the critical issues. Using data from the British Hoverfly Recording Scheme (HRS) we investigate the issues that need to be taken into consideration when interpreting modelling outputs based on opportunistic records.

UNDERSTANDING THE DATA

Buckland & Johnston (2017) list representative sampling locations, sufficient sample size, sufficient detections of target species, a representative sample of species, and a sound temporal sampling scheme as the essential attributes of a well-designed monitoring scheme. Opportunistic data, in contrast, are not based on a pre-defined rationale.

What is a biological record and who are the recorders?

The simplest biological record comprises: the identity of the organism that was observed; the name of the recorder (if known); the location from which the record came; and the date the observation was made. Ideally, the location should include a spatial coordinate (in the UK and Eire, this usually means an Ordnance Survey of Great Britain or Ireland grid reference); but many records based on museum specimen labels, observations published in journals, or those extracted from diaries, lack this detail and the location name may be vague or abbreviated.

Biological recorders vary in expertise and in their appreciation of the value of particular data attributes. Thus, some may contribute occasional sightings that they consider noteworthy (or reliable) enough to merit a record. Conversely, some recorders log very long lists. Across this spectrum lie many other levels of expertise, but relatively few recorders develop an in-depth knowledge of more complex taxa.

Fig. 1 illustrates the evolution of the British Hoverfly Recording Scheme (HRS), the third-largest British invertebrate dataset after Lepidoptera and Odonata, with 1.3 million records. It shows how the background level of recording was extremely low prior to the scheme's launch and also how the scheme has grown in the following 45 years.

Growth in the numbers of records is mirrored by increasing numbers of contributors (Fig. 2), and a rise in the numbers of species recorded from the late 1970s onwards (when test keys were circulated). Fig. 3 demonstrates that recording has, however, shifted away from difficult taxa in favour of those that are relatively straightforward to recognise (Table 1).

Biological recording is mainly concentrated in areas of dense human population, although it often involves visits to wilder but publicly accessible places (Fig. 4a). Least well recorded areas comprise the uplands, areas of high agricultural intensification (Morris & Edwards, 2019) and where there is poor vehicular access (e.g. the more remote parts of northern and western Scotland). Recorder effort is neither geographically or temporally uniform. Figs. 4b and 4c illustrate the differences in recorder effort in the periods 1980-1989 and 2010-2019 and show how recorder activity can vary markedly; both geographically and in intensity.

Modern technology and improved taxonomic guides have led to a rapid evolution of recorder activity. Early keys were often written by a museum specialist for other specialists with similar levels of taxonomic expertise who have access to reference collections, rather than the non-specialist with limited taxonomic training and at best access to a personal voucher collection. Until the advent of keys in which individual couplets are illustrated, the aspiring specialist relied upon assistance from more experienced colleagues. When Recording Schemes were first established, the numbers of competent recorders were therefore small (Fig. 2).

Modern guides are far better illustrated and accessible; they are complemented by on-line guides and crowd-sourced identification platforms such as iSpot and Facebook (Silvertown *et al.*, 2015; August *et al.*, 2015). Engagement with recorders is also helped by social media, e-mail and on-line newsletters.

The volume of records (Fig. 1) and numbers of recorders (Fig. 2) entering the HRS in the past ten years is illustrative of the sea-change in biological recording. Unfortunately, there has been a remarkable dearth of literature on the evolution of individual schemes; a shortfall that must be rectified if the analysts are to interpret their results.

Taxonomy has also advanced: in many cases species have been found to comprise one or more species. In addition, new arrivals are reported and overlooked species are found. The species list is therefore continually changing, as is our perception of a comprehensive species list. The apparent increase in the numbers of species over the study period must therefore be taken into account. Our approach has been to work with the aggregate of species that have been recently split rather than the separate segregates and to make sure that those species that cannot be segregated without microscopy, or in one sex (usually females) are recorded as the aggregate.

Observer skills

Fig. 2 shows how the numbers of active contributors to the HRS have changed over time. The yearly proportion of records of 'difficult taxa' within the dataset has also changed (Fig. 3). These 'difficult' species are readily overlooked species within the genus *Cheilosia* and the Tribe Pipizini that often require high magnification and comparison with voucher specimens. The expertise of recorders clearly has a strong bearing upon the level of detection of some taxa: datasets will inevitably be proportionately deficient in difficult taxa if the numbers of specialists do not keep pace with the numbers of generalists.

Difficulty of identification and rarity are not linked. For example, the hoverfly *Caliprobola speciosa* Rossi is readily identified, but is confined to a small part of The New Forest in Hampshire and to Windsor Great Park in Berkshire. Unless those sites are visited at the right time of year and under suitable conditions, it will not be detected, regardless of its real frequency. Conversely, *Heringia heringi* (Zetterstedt), is a widespread species that can only be reliably separated from its close relatives by microscopic examination. Regardless of the numbers of recorders, *H. heringi* will only be recorded by those who retain specimens and examine them under the microscope.

Analysis must therefore consider the degree to which target organisms can be detected by the contributors. Defining how many species within a given Order or Family can be identified without specimens is, however, unquantified. Ball & Morris (2013) did this for hoverflies but it requires refinement because photographic recording was in its infancy and high-quality photographs can sometimes reveal features that otherwise require microscopy. There is relatively little understanding of the composition of datasets based on photographs and how these relate to recorder skill (Falk *et al.*, 2019). Some basic analysis has, however, been undertaken for hoverflies.

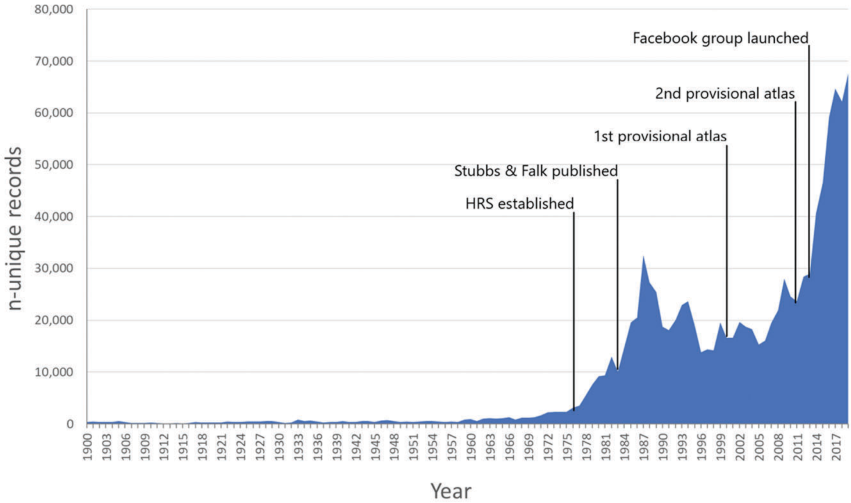


Fig. 1. The evolution of hoverfly data in Great Britain, depicting the numbers of unique records per year from 1900 to 2019. Key events are highlighted. Records that have survived and been extracted prior to the establishment of the scheme are very sparse. The publication of a user-friendly monograph in 1983 marks a first boost to recorder activity. The next big step was the founding of the UK Hoverflies Facebook group in 2013, although interest in recording by photographers using iSpot began in 2009.

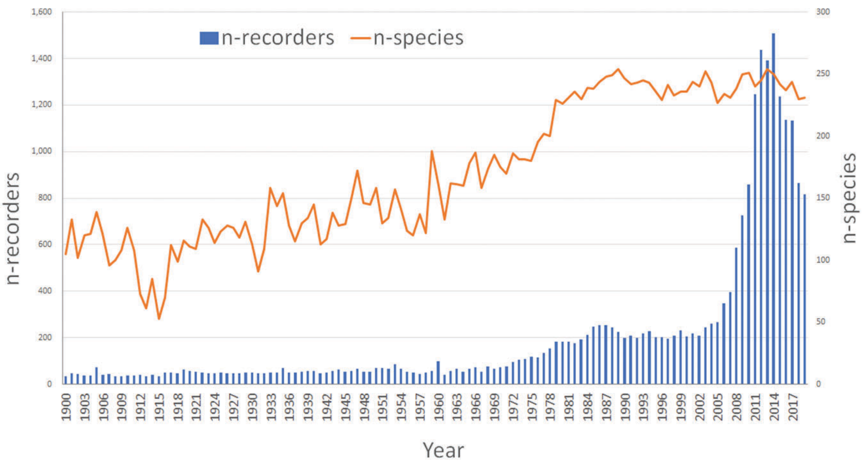


Fig. 2. The numbers of recorders contributing to the HRS compared with the numbers of species reported. The larger numbers between 2011 and 2017 reflect a period when efforts were made to scan photo-hosting websites for records; a process that stopped because it was no longer feasible with available volunteer time.

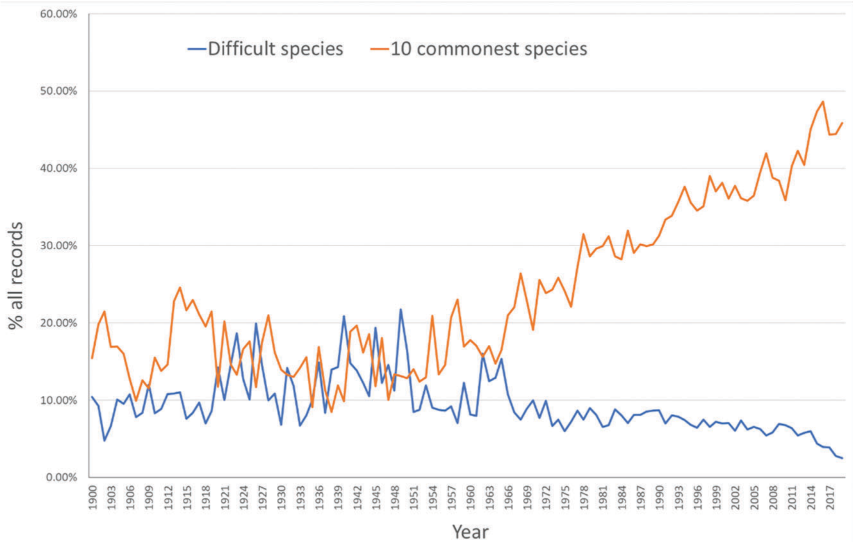


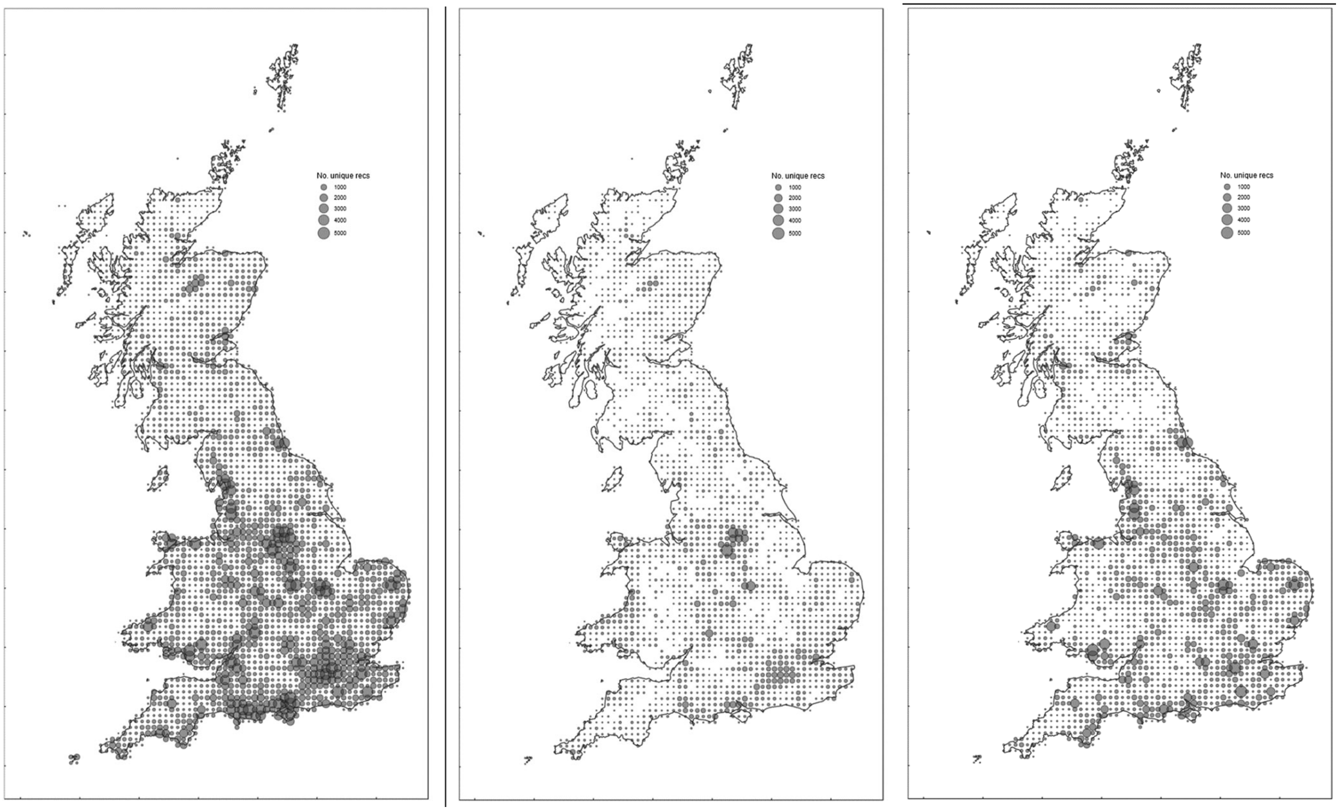
Fig. 3. The changing proportions of representation within the HRS dataset of difficult taxa (*Cheilosia* and members of the Tribe Pipizini), and the proportion of records that are of the ten most frequently reported species (Table 1).

Table 1. The ten most frequently recorded hoverfly species in the HRS dataset represented as both the numbers of records and the proportion of the total records received.

Species	n-records	% all records)
Episyrphus balteatus (De Geer)	75,303	6.54
Eristalis pertinax (Scopoli)	56,862	4.94
Eristalis tenax (L.)	49,472	4.30
Helophilus pendulus (L.)	47,251	4.11
Platycheirus albimanus (F.)	42,040	3.65
Melanostoma scalare (F.)	40,790	3.54
Syritta pipiens (L.)	38,643	3.36
Syrphus ribesii (L.)	33,971	2.95
Myathropa florea (L.)	27,157	2.36
Rhingia campestris Meigen	26,205	2.36

Photographic records on iRecord have an error rate of about 10% that are incorrectly or over-ambitiously identified (i.e. a specialist would not go further than genus) (Morris, 2019). Over the period 2015 to 2019, approximately 30% of the British hoverfly fauna was undetected photographically (Morris, 2020). This analysis also shows that imperfect detection is not confined to rare species: some extremely abundant species are poorly represented because they cannot be reliably identified from photographs. Many (but not all) of these are ‘difficult’ species illustrated in Fig. 3.

Table 2 summarises records accompanied by field photographs on iRecord that have been verified by RKAM. In this sample, 61.5% of the British hoverfly fauna



a. All records

b. 1980 to 1989

c. 2010 to 2019

Fig. 4. Density of hoverfly records across Great Britain showing the numbers of unique records received per hectad as a. all records; b. records from 1980 to 1989; and c. records from 2010 to 2019. The scale of numbers of unique records reads (smallest to largest): 1,000; 2,000; 3,000; 4,000; 5,000.

Table 2. Hoverfly records verified by RKAM on iRecord in which the record is accompanied by a field photograph. Species taken to a higher taxon are excluded (including those recorded as an aggregate species).

	Number of records										
	1–5	6–10	11–20	21–30	31–50	51–100	101–250	251–500	501–1000	1001+	Total
Species	46	34	13	3	16	21	13	16	7	5	174
Records	107	258	195	71	520	1,490	2,009	6,390	5,042	8,151	24,233

was detected, but the numbers of records are relatively small when compared against the dataset of photographs extracted directly from Facebook used by Morris (2020). Nevertheless, this small sample illustrates how the data are dominated by a few species.

A change in recorder methods can generate substantial bias in estimated trends from standard models (Ball & Morris, unpublished). The most obvious example is an increasing dependence upon photographs taken by generalists and identified by specialists, but there are others. For HRS data, the first big change was eight years after the scheme's inception in 1976 (Ball *et al.*, 2011) when a new illustrated guide (Stubbs & Falk, 1983) was published. This pivotal moment led to a significant growth in the number of active recorders (Fig. 1). Data flow, however, is dependent upon the level of feedback given to recorders: Fig. 1 illustrates how recorder activity waned when there were no goals (e.g. from 2000 to 2005) and how the advent of interactive media has provided a link to many recorders who, hitherto, would not have participated. Several classes of records and recorders can be identified, proportions of which have changed over time:

1) 'Naïve' recorders who submit occasional records of a few easily recognised or highly charismatic species such as the ubiquitous *Episyrphus balteatus* and the spectacular *Volucella zonaria* (Poda).
2) Recorders who evolve from the 'naïve' stage and take a regular interest in finding and recording a wide range of species that can either be recognised in the field, or from photographs.
3) Recorders who take a deeper interest in the subject and attempt to find as many species as possible, including those that can only be identified under high magnification and using difficult techniques such as investigation of the genitalia. These recorders can be sub-divided into:
   a) Those who submit a small number of records of what the recorder considers to be of interest, either because of its date of observation, the location seen or the rarity of the species concerned.
   b) Those who submit full lists whenever they observe hoverflies.

Whereas in the early years of biological recording schemes data were dominated by the third level, levels 1 and 2 now prevail. Fig. 3 provides part of the description of this change, whereby the proportions of 'easy' and 'difficult' species within the dataset can be seen to be diverging. Further indications of this divergence can be seen in the numbers of species recorded by individuals (Fig. 5), which illustrates the extreme paucity of recorders to have encountered a significant part of the British fauna and are therefore both taxonomically competent and capable of providing the

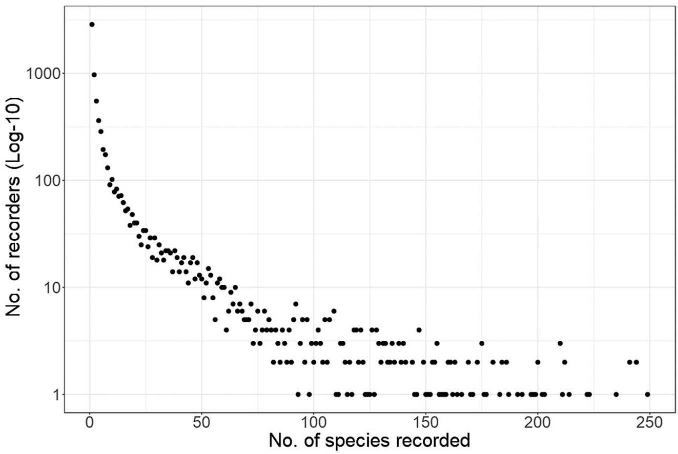


Fig. 5. Numbers of species recorded by recorders over the period 1980 to 2019. 77% of all individuals recorded 10 or fewer species, whereas just 5 individuals have recorded more than 85% of the fauna.

specialist skills required to underpin data-gathering dominated by 'naïve' and developing recorders.

Some datasets will also contain records from specialist surveys. There were three big surveys conducted by the then 'Nature Conservancy Council' covering the peatlands of Wales, the Fens of East Anglia and the shingle of Dungeness and Rye Harbour. These surveys generated large volumes of data and, at least in the case of wetland species, they mark a high point in the number of recorded observations for species such as *Anasimyia lunulata* (Meigen) that have rarely been seen since (Fig. 6).

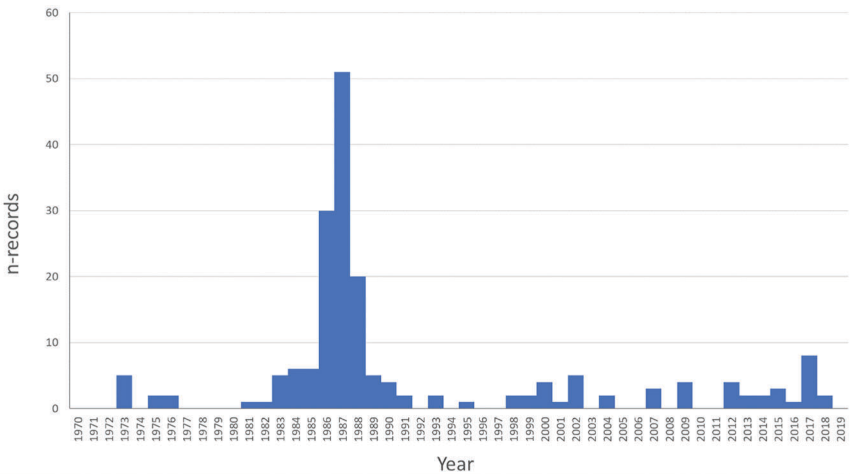


Fig. 6. Records of *Anasimyia lunulata* between 1980 and 2019, showing the impact of the Welsh Peatland Invertebrate Survey that ran from 1987 to 1989.

List lengths

The phenological tracking graphs in BirdTrack (BTO, 2020) and the use of list length as a way of investigating species declines (Szabo *et al.*, 2010) demonstrate the importance of list length. Hoverfly faunas are potentially very rich: in Britain, exceptional sites support more than 140 species but even urban gardens can be remarkably diverse. The longest published garden list (over 90 species) was from a Leicester garden (Owen, 2010; Hassell *et al.*, 2017) and monitoring by Alan Stubbs (Graham-Taylor, Stubbs & Brooke, 2008) has generated a list of more than 60 species. It therefore follows that there might be differences between analyses based on observers that generate single or occasional records, as opposed to those that attempt to record everything that they encounter at a given location on a given day.

We tested this question using Frescalo (Hill, 2012), which generates a measure (T-factor) of the frequency with which species were observed relative to the amount of recording. This can be used to generate trends (see Ball & Morris unpublished). The basic data required for this analysis comprises unique combinations of hectad (10 km square), the species and year extracted from HRS for the period 1980 to 2019. Frescalo was run using all available data and then repeated after excluding occasions when less than three, five, seven or nine 9 species were recorded from the same grid reference on the same day.

This analysis clearly illustrates the influence of single records and very short lists. When short lists are excluded, the numbers of combinations of hectad, species and year decline rapidly (Table 3); halving if the analysis is confined to data made up of lists comprise nine or more species. Moreover, the headline figures of decline (Table 4) change from 51% to 41% if lists of nine or more species are used. The choice of list length needs to be placed into context, however, because long lists may exclude some species that occur in late March or April when only a few species are flying.

Table 3. Numbers of available hectad/species/year combinations available from the HRS dataset when list length is taken into account.

List length	Hectad/species/year combinations
All records	480,593
3+ records	394,154
5+ records	337,671
7+ records	288,012
9+ records	243,550

Table 4. Trends in British hoverflies generated by separate Frescalo runs in which shorter lists are progressively excluded. The sensitivity of the model to list length is represented by the numbers of species falling into the three categories of change.

	Number of species in list				
	1+	3+	5+	7+	9+
Significant decline	121	109	107	95	88
No significant change	85	84	84	92	95
Significant increase	31	35	31	31	30
Grand Total	237	228	222	218	213

Temporal issues

There has been a dramatic recent increase in recording during winter months (Morris & Ball, 2019c): allowing more to be discerned from the data for the past decade than for the previous 3 decades! When the HRS was first established recorders largely confined their activity to favourable weather and to months when hoverfly activity was at its peak (Fig. 7). One reason for the difference is that winters have become warmer (Lorenz, Stalhandske & Fischer, 2019) and emergence times have advanced (in some cases very markedly) (Ball & Morris, 2020). Hoverfly activity also continues later in the autumn; moreover, many modern recorders are also active throughout the year.

Fig. 3 shows that relatively more records of challenging species were reported in the 1970s; perhaps reflecting that records came from reference collections. Anecdotally, however, it seems that Recording Schemes were regarded as mapping projects, as demonstrated by comparing the proportions of records of the ten most frequently recorded species in the HRS dataset (Table 1) with those that are taxonomically difficult (Fig. 3). Clearly, data submitted to the HRS has always been selective, either for rarity or for ease of identification! Rarity and difficulty were more prevalent in the 1970s and ease of identification has dominated the past decade.

It should also be noted that a very small cohort of recorders once dominated the dataset; many of whom have contributed data for almost the entire life of the scheme. Ball *et al*. (2011) report that 50% of the data submitted to the scheme had been by just 20 recorders; whereas, by 2019, 51 recorders were responsible for 50% of the data.

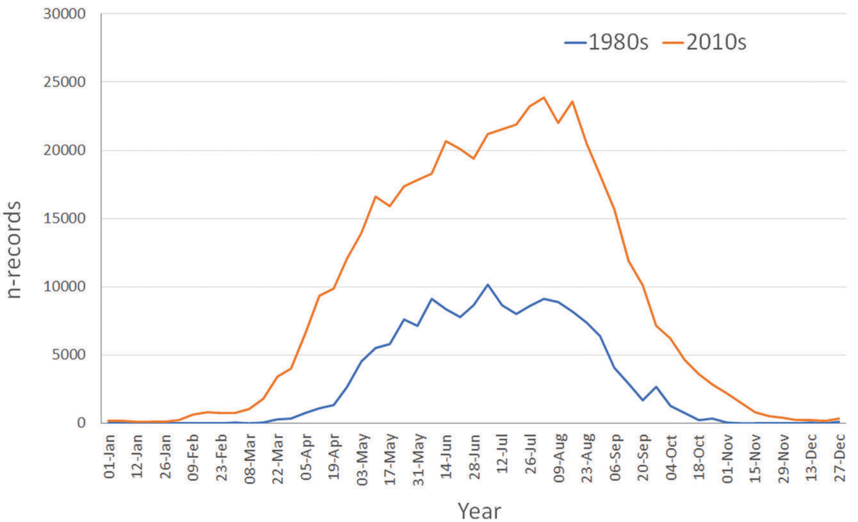


Fig. 7. Levels of recording activity in the 1980s and in the 2010s. The profile of the two graphs differs substantially, with active recording starting much earlier in the year and ending far later in the 2010s. These differences may in part be explained by warmer winters and autumns, but the shapes of the two graphs are also markedly different, with far fewer records generated in September onwards in the 1980s; these are months in which there is now a great deal of recorder activity linked to monitoring of ivy flowers.

Geographic issues

Range and variation in geographic coverage by recorders varies considerably (Johnston *et al*., 2020). Therefore, unless regional trends are quantified, reported outputs can only be the mean of the trends, even though they may differ! For example, Ball & Morris (in prep) show how the hoverfly *Leucozona glaucia* (Linnaeus) has substantially disappeared from south-east England, and yet its relative frequency is increasing in northern and western Scotland. The combined results lead to a relatively flat trend, but the regional differences are considerable. It is possible that reported trends from analyses such as Outhwaite *et al*. (2020) actually mask regional trends, perhaps explaining why their results differ from those for birds and butterflies.

Drift and underlying geology influence distribution and abundance of some plants and animals (e.g. Stubbs 1967–8). There are also assemblages and species that favour base-rich or acidic water chemistry (Ball *et al*., 2011) and others that favour saline influences (Stubbs & Falk, 1983). Some of these parameters are highly geographically constrained. For example, base-richness in Scotland and Wales is limited because the rocks are ancient and highly resilient, with essential nutrients having long-since leached away by high rainfall. Thus, even if the climate envelope becomes favourable, the potential for expanding colonisation is limited: predictive models must take account of both projected climate envelopes and water and soil chemistry.

Day-length and temperature profiles are linked to latitude and longitude. Variation in geographic and temporal spread of records will therefore affect regional phenology, which almost certainly explains the results for hoverfly phenology presented by Hassall *et al*. (2017); who reported different phenological signatures between a long-term Malaise trap programme in a Leicestershire garden (Owen, 2010), and the dataset compiled by the HRS. This result is not surprising, given that a single point of reference is compared with data covering 5 degrees of longitude and 9 degrees of latitude!

Discussion

It is important that researchers who use opportunistic data take account of the heterogeneity described here. The issue of sampling bias (Kéry, 2011; Kellner & Swihart, 2014; Morris, 2019, 2020) has received too little attention in previous analyses. Over the past 50 years, the emphasis of this variation has changed and there is a risk that some trends will be misinterpreted as real. This variation also means that comparison between highly heterogeneous data and those based on structured methods will rarely be reliable, although there may be exceptions. Increasing numbers of records do, however, mean more dots on maps for readily recognised species and also allow tracking phenological change in the better recorded species at a more refined geographic scale.

Outhwaite *et al*. (2020) compare opportunistic data with results from structured bird monitoring (Hayhow *et al*., 2017) and from butterfly monitoring (Fox *et al*., 2015). As such, it compares a dataset that scores weakly on the Buckland & Johnston criteria with datasets that more closely conform to the criteria. Heterogeneity alone may account for the problem, but there may be other reasons. For example, van Strien *et al*., (2013) compared opportunistic data with structured data for butterflies and dragonflies in The Netherlands. In both cases, the analysis involved readily recognisable and easily recorded taxa for which there are very high numbers of records. Moreover, the Netherlands is comparatively small and has fairly homogenous landforms. Equally importantly, the methods of recording may not have changed significantly within both the structured and unstructured data. Magurran *et al*. (2010) draw attention to the potential impacts of changes in sampling or analytical methods. It is clear that, within some datasets, methods have changed profoundly. At the moment, the speed of the changes means that modern datasets may differ substantially from those of previous decades, and especially those of the 1970s.

British opportunistic data cover many more species and there is considerable variation in recorder effort both geographically and temporally than for van Strien's samples. Many datasets comprise larger number of species and fewer records. In all analyses, the volume of data, and data quality are critical to the analytical outcomes, but the degree to which geography, landform, hydrology and urbanisation are homogenous will also affect the results. Thus, the arguments advanced by van Strien *et al.* must be caveated with a warning that models may not adequately cope in cases of extreme heterogeneity in these attributes.

Opportunistic data are not compiled according to a rigid set of parameters, but the degree to which they vary from the principles established by Buckland & Johnston (2017) is important. Thus, it is worth exploring each component separately:

Representative sampling locations

We have shown how the HRS dataset has a strong southerly bias and therefore question whether coverage in northern regions, especially Scotland, is adequate throughout the epochs. Johnston *et al.* (2020) draw attention to the relative paucity of data for birds in the Scottish Highlands and found that it was more difficult to 'characterise environmental relationships with occupancy'. Bearing in mind that the volume of data available for their analysis is considerably greater than those of most opportunistic datasets, it is salutary to note that recording effort was sufficiently low that precision and accuracy were affected even after the adoption of weighting.

Sufficient sample size

Bonney *et al.*, 2009 emphasise the importance of data volume for detecting trends in the natural world. Wisz *et al.* (2008) also found that as sample size decreases, model accuracy decreased. They advise that in the case of sample sizes $n < 30$, extreme caution should be employed in interpreting the results.

During the 1970s, the volume of records submitted to the HRS was comparatively small. Bearing in mind that the British Hoverfly fauna comprises at least 283 species (Ball & Morris, 2013), a sample size below 8,500 records (i.e. an average of 30 records per species) may be insufficient to generate meaningful results. This figure excludes most years prior to 1980. There is no reason to believe that a similar situation does not obtain for other recording schemes, given that the HRS is one of the bigger invertebrate datasets.

Didham, *et al.* (2020) demonstrate the importance of fluctuations in data-richness over a long time-series and how such fluctuations can affect perceived and actual trends. Trends based on data compiled in the 1970s start at a low point in recorder activity but arguably the high point in relative wildlife abundance. Conversely, assuming the combined weight of trend analyses is correct, high levels of recorder activity coincide with greatly reduced wildlife abundance. Thus, there is a danger that the 'false baseline effect' of Didham *et al.* (2020) has influenced at least some of the analyses to date.

Species composition within the sample is an additional complicating factor, and an improvement in sample size may be misleading. If the sample size grows disproportionately with numbers of single records of widespread and ubiquitous species, a cautious approach is needed. We have shown that restricting analysis to a partial dataset, based on longer lists, has a substantial impact on modelled trends.

Species for which there are insufficient data and those such as *Anasimyia lunulata* also need to be excluded from the analysis.

These problems draw attention to the weaknesses of well-meaning, but naïve, initiatives to 'increase the volume of data' for a given group of organisms. For example, the 'great British bee count' (FoE, 2020) generated large numbers of records of a tiny proportion of the known fauna (482,915 records of 50 species out approximately 270 species). If incorporated into the working dataset a tremendous imbalance would be created. Such initiatives are likely to cause changes in detection probability that will break many of the usual assumptions made (i.e. that detectability and sample size vary smoothly over time). This could, however, be handled with a covariate since the occurrence of initiatives should be known (at least to the data custodians).

Sufficient detections of target species

Opportunistic recording makes no distinction between target and non-target species. It is a mixture of whatever data are available. Using the HRS as an example, it seems that there is a data skew, in which taxonomically difficult species were better represented in the earlier data (Fig. 3). There are several approaches that might be taken to take account of the dominance of a small number of widespread and common species within the data. One option is to exclude lists composed of fewer than three species; another is to exclude contributions of recorders who concentrate on a very narrow suite of species. In addition, it may be necessary to develop new models that take account of such co-variates.

A representative sample of species

We have shown how the HRS the dataset has changed over the decades since its inception in 1976. Early data (Fig. 3) clearly included a far larger proportion of taxonomically difficult species, whereas modern data are substantially deficient in those species. Magurran *et al.* (2010) highlight the significance of detectability affecting estimates of turnover in long-term monitoring. It is clear that there have been several points in the HRS dataset when detectability has changed. Thus, over time, these hiatuses mean that the representativity of the dataset has also changed.

A sound temporal sampling scheme

Temporal variation over the life of any long-lasting recording scheme can be considerable (Fig. 4), which is certainly the case for the HRS. We have also shown how there is substantial heterogeneity in the composition of data since the scheme's inception in 1976. At the moment, the degree to which photographic recording is, or will be heterogeneous in the longer-run is unresolved. We know that the numbers of active recorders are rising and the volumes of submitted records greatly exceed those of preceding decades. We also know that the species representation is markedly smaller than the total British fauna (Morris, 2020), but it is possible that if treated as a separate block of data it will effectively be more homogenous because of the sample size and recording method.

Recommendations

We have shown that opportunistic data is highly heterogeneous, with many potential complicating influences. Of these, list length emerges as a critical influence on models. The numbers of species encountered on a given visit will vary with the time of year, local geography and latitude/longitude. The ultimate choice of the records used will therefore vary. In the case of hoverflies, if models are used that fail to incorporate major sources of heterogeneity, we consider that excluding lists of fewer than three species is optimal, but that it may be necessary to repeat analysis using longer lists to evaluate the differences. This rule is likely to be equally applicable to other pollinator assemblages, especially the bees, which have comparable representation to hoverflies within the British fauna. There is scope for other teams to investigate optimal list length for other taxa.

More robust estimation of presence-only data with existing methods such as Frescalo may be possible by excluding more data, for example sites with very few records, or very variable effort. However, the more data that are discarded, the lower the precision of resulting estimates. If feasible, our preferred option is to identify sources of heterogeneity, and then implement a model selection procedure to determine which terms should be incorporated into a model for presence-only data.

For the HRS dataset, we consider that the following sources of heterogeneity should be considered: date, time of day, temperature, recorder, location, difficulty in identifying the species, and some measure of effort, perhaps assessed by the number of species recorded or a related measure. We also consider it necessary to assess the degree to which heterogeneity has been successfully modelled, by carrying out sensitivity analyses. The ideal model would estimate very similar trends (though with varying precision) if analyses are carried out on subsets of data for which data are dropped from 1) influential recorders; 2) under-recorded sites and atypical sites with highly variable effort; 3) dates early and late in the season; 4) species that are difficult to identify.

Our analysis also highlights the need for recording schemes to make a bigger effort to encourage contributors to submit full lists (or as full lists as possible), and for modellers to make the effort to account for observer preferences. It also emphasises the risk that well-meaning initiatives to increase the volume of data do not necessarily translate into 'research-grade' data that can be used in the full spectrum of analysis.

Furthermore, as we have illustrated, there may be points in the evolution of the dataset where major perturbations such as one-off surveys or citizen science initiatives distort the smooth evolution of the dataset. These perturbations need to be identified and taken into account, either before any modelling happens or by incorporating these effects in the model.

ACKNOWLEDGEMENTS

This work would not have been possible without the contributions of more than 8,000 individual hoverfly recorders to whom we are indebted.

REFERENCES

August, T., Harvey, M., Lightfoot, P., Kilbey, D., Papadopoulos, T. & Jepson, P. 2015. Emerging technologies for biological recording. *Biological Journal of the Linnean Society* **115**: 731–749.

Ball, S. G. & Morris, R. K. A. 2020. Changes in the Phenology of Britain's Hoverflies (Syrphidae). *Dipterists Digest* (second series), **27**: 1–12.

Ball, S. & Morris, R. 2013. *Britain's Hoverflies: an introduction to the hoverflies of Britain*. Princeton University Press, Woodstock, Oxfordshire. 297pp.

Ball, S., Morris, R., Rotheray, G. & Watt, K. 2011. *Atlas of the Hoverflies of Great Britain (Diptera, Syrphidae)*. Centre for Ecology & Hydrology, Wallingford. 184pp.

Biesmeijer, J. C., Roberts, S. P. M., Reemer, M., Ohlemüller, R., Edwards, M., Peeters, T., Schaffers, A. P., Potts, S. G., Klenkers, R., Thomas, C. D., Settele, J. & Kunin, W. E. 2006. Parallel declines in pollinators and insect-pollinated plants in Britain and the Netherlands. *Science* **313(5785)**: 351354. DOI: 10.1126/science.1127863

Bonney, R., Cooper, C., Dickinson, J. L., Kelling, S., Phillips, T., Rosenberg, K. V. & Shirk, J. L. 2009. Citizen Science: A Developing Tool for Expanding Science Knowledge and Scientific Literacy. *BioScience* **59**: 977–984.

Brooks, S. J., Self, A., Powney, G. D., Pearse, W. D., Penn, M. & Paterson, G. L. J. 2017. The influence of life history traits on the phenological response of British butterflies to climate variability since the late-19th century. *Ecography* **40**: 1152–1165.

BTO (2020). https://app.bto.org/birdtrack/login/login.jsp (Accessed on 6 February 2021).

Buckland, S. T. & Johnston, A. 2017. Monitoring the biodiversity of regions: Key principles and possible pitfalls. *Biological Conservation* **214**: 23–34.

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. doi:10.1111/icad.12408

Falk, S., Foster, G., Comont, R., Conroy, J., Bostock, H., Salisbury, A., Kilbey, D., Bennett, J. & Smith, B. 2019. Evaluating the ability of citizen scientists to identify bumblebee (*Bombus*) species. PLOS ONE **14(6)**: e0218614. https://doi.org/10.1371/journal.-pone.0218614

FoE 2020. The Great British Bee Count. https://friendsoftheearth.uk/bee-count (Accessed on 6 February 2021).

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, Dorset. https://butterfly-conservation.org/sites/default/files/soukb-2015.pdf (Accessed on 6 February 2021).

Geijzendorffer, I. R., Regan, E.C., Pereira, H.M., Brotons, L., Brummitt, N., Gavish, Y., Haase, P., Martin, C.S., Mihoub, J-B., Secades, C., Schmeller, D.S., Stoll, S., Wetzel, F.T. & Walters, M. 2015. Bridging the Gap between Biodiversity Data and Policy Reporting Needs: An Essential Biodiversity Variables Perspective. *Journal of Applied Ecology* **53**: 13411350.

Graham-Taylor, L.G., Stubbs, A. & Brooke, M. de L. 2008. Changes in phenology of hoverflies in a central England garden. *Insect Conservation and Diversity* **2**: 29–35.

Hassall, C., Owen, J. & Gilbert, F. 2017. Phenological shifts in hoverflies (Diptera: Syrphidae): linking measurement and mechanism. *Ecography* **40**: 853–863.

Hastie, T. & Fithian, W. 2013. Inference from presence-only data; the ongoing controversy. *Ecography* **36**: 864–867.

Hayhow, D. B., Ausden, M. A., Bradbury, R. B., Burnell, D., Copeland, A. I., Crick, H. Q. P., Eaton, M. A., Frost, T., Grice, P. V., Hall, C., Harris, S. J., Morecroft, M. D., Noble, D. G., Pearce-Higgins, J. W., Watts, O. & Williams, J. M. 2017. *The state of the UK's birds* 2017. The RSPB, BTO, WWT, DAERA, JNCC, NE and NRW, Sandy, Bedfordshire. https://www.bto.org/sites/default/files/publications/state-of-uk-birds-_2017.pdf (Accessed on 6 February 2021)

Hill, M. O. 2012. Local frequency as a key to interpreting species occurrence data when recording effort is not known. *Methods in Ecology and Evolution* **3**: 195205.

Isaac, N. J. B., van Strien, A. J., August, T. A., Zeeuq, M. P. & Roy, D. B. 2014. Statistics for citizen science: extracting signals of change from noisy ecological data. *Methods in Ecology and Evolution* **5**: 1052–1060.

Isaac, N. J. B. & Pocock, M. J. O. 2015. Bias and information in biological records. *Biological Journal of the Linnean Society* **115**: 522–531.

Johnston, A., Moran, N., Musgrove, A., Fink, D. & Baillie, S. R. 2020. Estimating species distributions from spatially biased citizen science data. *Ecological Modelling* 422: 108927

Kellner, K. F. & Swihart, R. K. 2014. Accounting for Imperfect Detection in Ecology: A quantitative review. *PLoS One* **9**: e111436. doi: 10.1371/journal.pone.0111436

Kéry, M. 2011. Towards the modelling of true species distributions. *Journal of Biogeography* 38: 617–618.

Lorenz, R., Stalhandske, Z. & Fischer, E. M. 2019. Detection of a climate change signal in extreme heat, heat stress, and cold in Europe from observations. *Geophysical Research Letters* **46**: 8363–8374.

Magurran, A. E., Baillie, S. R., Buckland, S. T., Dick, J. McP., Elston, D. A., Scott, E. M., Smith, R. I., Somerfield, P. J. & Watt, A. D. 2010. Long-term datasets in biodiversity research and monitoring: assessing change in ecological communities through time. *Trends in Ecology and Evolution* **25**: 574–582.

Morris, R. K. A. 2020. Imperfect detection: photographic recording and its implications. *British Journal of Entomology & Natural History* **33**: 41–56.

Morris, R. K. A. 2019. Understanding common misidentifications of British hoverflies (Diptera: Syrphidae). *British Journal of Entomology & Natural History* **32**: 351–363.

Morris, R. K. A. & Ball, S. G. 2019a. Effects of the 2018 heatwave on British hoverflies (Diptera, Syrphidae). *Dipterists Digest* (second series) **27**: 139–150.

Morris, R. K. A. & Ball, S. G., 2019b. Hoverflies (Diptera, Syrphidae) visiting ivy *Hedera* species. *Dipterists Digest* (second series) **27**: 161–167.

Morris, R. K. A. & Ball, S. G., 2019c. Digital photography improves assessment of winter-active hoverflies (Diptera, Syrphidae). *Dipterists Digest* (second series) **27**: 25–33.

Morris, R. & Edwards, M. 2019. What has happened to our pollinators? *Antenna* **43**: 179–185.

Oliver, T. H., Isaac, N. J. B., August, T. A., Woodcock, B. A., Roy, D. B. & Bullock, J. M. 2015. Declining resilience of ecosystem functions under biodiversity loss. *Nature Communications* **6**: 10122. doi: 10.1038/ncomms10122

Outhwaite, C. L., Gregory, R. D., Chandler, R. E., Collen, B. & Isaac, N. J. B. 2020. Complex long-term biodiversity change among invertebrates, bryophytes and lichens. *Nature Ecology & Evolution* **4**: 384–392. https://doi.org/10.1038/s41559-020-1111-z

Owen, J. 2010. *Wildlife of a garden: a thirty-year study*. Royal Horticultural Society, Peterborough. 261pp.

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

Pocock, M. J. O., Newson, S. E., Henderson, I. G., Peyton. J., Sutherland, Noble, D. G., Ball, S. G., Beckmann, B. C., Biggs, J., Brereton, T., Bullock, D. J., Buckland, S. T., Edwards, M., Eaton, M. A., Harvey, M. C., Hill, M. O., Horlock, M., Hubble, D. S., Julian, A. M., Mackey, E. C., Mann, D. J., Marshall, M. J., Medlock, J. M., O'Mahony, E. M., Pacheco, M., Porter, K., Prentice, S., Procter, D. A., Roy, H. A., Southway, S. E., Shortall, C. R., Stewart, A. J. A., Wembridge, D. E., Wright, M. A. & Roy, D. B. 2015b. Developing and enhancing biodiversity monitoring programmes: a collaborative assessment of priorities. *Journal of Applied Ecology* **52**: 686–695.

Powney, G. D., Carvell, C., Edwards, M., Morris, R. K. A., Roy, H. E., Woodcock, B. A. & Isaac, N. J. B. 2019. Widespread losses of pollinating insects in Britain. *Nature Communications* **10**: 1018. https://doi.org/10.1038/s41467-019-08974-9

Silvertown, J., Harvey, M., Greenwood, R., Dodd, M., Rosewell, J., Rebelo, T., Ansine, J. & McConway, K. 2015. Crowdsourcing the identification of organisms: a case-study of iSpot. *ZooKeys* **480**: 125–146.

Stubbs, A. E. 1967–8. Geology as an ecological factor in the distribution of an insect. The fly *Symphoromyia immaculata* If. (Dipt. Rhagionidae). *Entomologist's Record and Journal of Variation* **79**: 292–3, 313–6; **80**: 22–25, 56–59, 80–83.

Stubbs, A. E. & Falk, S. J. 1983. *British Hoverflies: an illustrated identification guide*. British Entomological & Natural History Society, London. 253pp.

Szabo, J. K., Vesk, P. A., Baxter, P. W. J. & Possingham, H.P. 2010. Regional avian species declines estimated from volunteer-collected long-term data using List Length Analysis. *Ecological Applications* **20**: 2157–2169.

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

Wetzel, F. T., Saarenmaa, H., Regan, E., Martin, C. S., Mergen, P., Smirnova, L., Ó Tuama, É, García Camacho, F. A., Hoffmann, A., Vohland, K. &Ü%1contribdegrees%doi_link