Biological Conservation 290 (2024) 110455

Contents lists available at ScienceDirect

Biological Conservation

journal homepage: www.elsevier.com/locate/biocon

Understanding and addressing shortfalls in European wild bee data

Leon Marshall[a,b,1,*], Nicolas Leclercq[b,1], Luísa G. Carvalheiro[c,d], Holger H. Dathe[e], Bernhard Jacobi[f], Michael Kuhlmann[g], Simon G. Potts[h], Pierre Rasmont[i], Stuart P.M. Roberts[b], Nicolas J. Vereecken[b]

[a] Naturalis Biodiversity Center, Darwinweg 2, 2333 CR Leiden, Netherlands
[b] Agroecology Lab, Université libre de Bruxelles (ULB), Boulevard du Triomphe CP 264/02, B-1050 Brussels, Belgium
[c] Departamento de Ecologia, Universidade Federal de Goiás, Goiânia, Brazil
[d] Center of Ecology, Evolution and Environmental Changes, Universidade de Lisboa, Lisboa, Portugal
[e] Senckenberg Deutsches Entomologisches Institut, Eberswalder Straße 90, 15374 Müncheberg, Germany
[f] Dieckerstraße 26, D-46047 Oberhausen, Germany
[g] Zoological Museum of Kiel University, Hegewischstr. 3, D-24105 Kiel, Germany
[h] Centre for Agri-Environmental Research, School of Agriculture, Policy and Development, University of Reading, Reading RG6 6AR, UK
[i] Laboratory of Zoology, Université de Mons, B-7000 Mons, Belgium

ARTICLE INFO

Keywords:
Knowledge gaps
Big data
Online tool
Biodiversity decline
Citizen science
Biodiversity monitoring

ABSTRACT

Understanding and reversing biodiversity decline in the Anthropocene requires robust data on species taxonomic identity, distribution, ecology, and population trends. Data deficits hinder biodiversity assessments and conservation, and despite major advances over the past few decades, our understanding of bee diversity, decline and distribution in Europe is still hampered by such data shortfalls. Using a unique digital dataset of wild bee occurrence and ecology, we identify seven critical shortfalls which are an absence of knowledge on geographic distributions, (functional) trait variation, population dynamics, evolutionary relationships, biotic interactions, species identity, and tolerance to abiotic conditions. We describe "BeeFall," an interactive online Shiny app tool, which visualizes these shortfalls and highlights missing data. We also define a new impediment, the Keartonian Impediment, which addresses an absence of high-quality in situ photos and illustrations with diagnostic characteristics and directly affects the outlined shortfalls. Shortfalls are highly correlated at both the provincial and national scales, identifying key areas in Europe where knowledge gaps can be filled. This work provides an important first step towards the long-term goal to mobilize and aggregate European wild bee data into a multiscale, easy access, shareable, and updatable database which can inform research, practice, and policy actions for the conservation of wild bees.

1. Introduction

Biodiversity is declining globally (Díaz et al., 2019). However, data availability and suitability often severely restrict a full evaluation of biodiversity trends, requiring information on taxonomic identity, distribution, ecology, population dynamics, interactions between species as well as their evolutionary relatedness (Hochkirch et al., 2021; Tylianakis et al., 2008). However, spatial and temporal sampling biases are ubiquitous features of species occurrence data (Hughes et al., 2021), partly due to lack of digitalization of massive amounts of data that are collected and stored in public or private collections, and partly due to lack of sampling (Meyer et al., 2015). Historically, the primary source of data has been opportunistic collections. Even within digitized data, there can be significant taxonomic, geographic, temporal, and methodological biases (Isaac and Pocock, 2015). For example, evidence shows that there has been a historical focus of taxonomists on larger, more charismatic species (Gaston, 1992) such as birds and mammals (Cardoso et al., 2011), and research on insects has been lagging behind. Within insect groups however, Western Europe and North America have seen greater collection and digitization of data (Rocha-Ortega et al., 2021) notably through opportunistic sampling such as citizen science and voluntary recording schemes (Pocock et al., 2015) that generate millions of

---

* Corresponding author.
E-mail addresses: leon.marshall@naturalis.nl (L. Marshall), nicolas.leclercq@ulb.be (N. Leclercq), nicolas.vereecken@ulb.be (N.J. Vereecken).

[1] Both authors contributed equally to the manuscript.

https://doi.org/10.1016/j.biocon.2024.110455
Received 30 May 2023; Received in revised form 23 December 2023; Accepted 5 January 2024

occurrence records every year (Kamp et al., 2016). Historical data collected by naturalists is also often biased towards rarer species and species with greater morphological differences between genders and developmental stages (Pocock et al., 2015). While citizen science approaches have a high intrinsic value to raise awareness about the diversity of living organisms around us, such opportunistic samplings are associated with important observation and detection biases because of the variable sampling intensity per survey event, uneven temporal distribution, the different levels of discoverability among target species (e.g., large vs small, colorful vs dull, widespread vs with a narrow range) (see e.g. Higgs and Attrill, 2015), the observational skills of the recorders (Geldmann et al., 2016), and the reliability of downstream validation opportunities by experts.

Navigating databases consisting of non-random occurrence data is perhaps one of the most pressing challenges of contemporary biodiversity research. To provide a framework on how to systematically address and fill the persistent knowledge gaps, Hortal et al. (2015) proposed a classification of "biodiversity shortfalls" into seven categories that best represent the multifaceted nature of the problem, with a focus on species identity, their distribution, population dynamics, evolution, behavioral/ecological traits, environmental tolerances, and interactions. These shortfalls are a crucial hindrance to biodiversity research as they cause difficulties for predicting declines and assessing threat status (Hortal et al., 2015; Cardoso and Leather, 2019). This, in turn, prevents a comprehensive understanding of species niches (Hortal et al., 2008) and their evolutionary histories (Diniz-Filho et al., 2013), and so remains an obstacle towards a global characterization of known and yet undiscovered insect diversity (e.g., Kass et al., 2022). Because these data gaps are not mutually exclusive in both their origins and solutions (Hortal et al., 2015; Diniz-Filho et al., 2013), it is of pivotal importance to acknowledge and quantify these shortfalls both individually and in combination, through novel and refined approaches of data collection, sharing and mapping.

Despite major advances over the past few decades in our understanding of bee diversity, their key role in the pollination of wild flowering plants and many crop species (Ollerton et al., 2011), and the drivers of their decline, important biases and knowledge gaps still plague the field of wild bee research and hamper targeted conservation efforts (Potts et al., 2016). Wild bee observation data are key for understanding ecological requirements (Winfree et al., 2009; Vereecken et al., 2021a, 2021b), modeling projected changes under global change (Marshall et al., 2018), understanding their role as pollinators (Weekers et al., 2022), resolving their evolutionary relationships (Bossert et al., 2022), informing conservation planning, practice, and policy (Potts et al., 2016), and raising public awareness to their importance. Last, the availability of tools tailored to the analyses and for outreach is key to stimulating the collection of data and to prioritize where data acquisition efforts are most needed. To date, available wild bee biological observations are primarily sourced from opportunistic collections or from detailed, local, scientific surveys, resulting in knowledge gaps, which follow geographical patterns (Orr et al., 2021a; Wetzel et al., 2018). To ensure the maintenance of wild bee diversity in a context of global changes, long-term monitoring is essential to fill the existing gaps (Potts et al., 2016; Potts et al., 2021).

Here, we use a unique occurrence dataset for the 2000+ species of wild bees in Europe, to (1) provide a first list and quantification of seven shortfalls (Fig. 1) relevant to European wild bees, and (2) develop the "BeeFall" tool, an interactive online map, available as a Shiny app, designed to visualize shortfalls independently and report missing data. We also describe an additional new shortfall related to the availability of high-quality in situ photographs and diagnostic traits illustration, a key asset to promote biological recording through citizen sciences. We discuss how the "BeeFall" tool can be used to prioritize field surveys, research efforts and conservation actions.

2. Methods

2.1. Study area

We established the study area as the European Union and all other countries included within the geographical bounds of the European Union. We opted for these geographical boundaries to align our study to the geographical scope of the EU Pollinator Monitoring Scheme (EU-PoMS) which includes all European Member States (Potts et al., 2021). Other countries within the boundaries of the EU member states are included to provide a continuous space by which to classify the data. Some non-contiguous regions (not independent countries) for which there was a species checklist were included at this scale of the analysis. This includes the Åland Islands, Corsica, Crete, East Aegean Islands, Balearic Islands, Canary Islands, Azores, Madeira, Isle of Man, Northern Ireland, Sardinia, and Sicily (hereafter all are referred to as countries). The provincial (the actual naming conventions at this scale vary by country but hereafter we refer to all as provinces) map used for visualization purposes is based on NUTS2 administrative units for the countries and regions. The checklist of wild bee species for each country was taken from Reverté et al. (2023). This publication lists the most up-to-date checklist for all countries and spatially distinct regions within our study area, see also Ghisbain et al. (2023) for greater detail. For a full list of the countries used in this study and their species richness values see Supplementary Table S1.

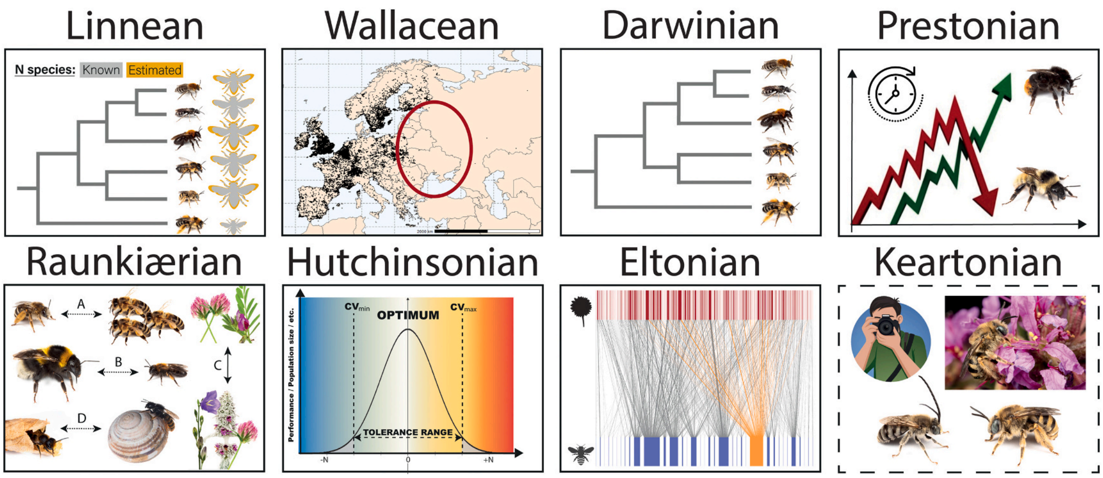

Fig. 1. Visual representation of the seven shortfalls discussed in this manuscript. See Hortal et al., (2015) for a detailed description of the first seven; we propose and introduce the Keartonian impediment.

2.2. Compiling a database of spatially explicit species occurrence records

We compiled several datasets from public resources or from different projects and experts. The public data were obtained via the Global Biodiversity Information Facility (GBIF) platform (https://doi.org/10.15468/dl.umg2ny) (66.2 % of all observations before aggregating). The remaining data were compiled from the Status and Trends of European Pollinators (STEP) project (Potts et al., 2011) for the most part (17.8 %). We supplemented it with occurrence records from; (i) the Bees, Wasps & Ants Recording Society (BWARS) program for the United Kingdom's data (www.bwars.com, 10.7 %); (ii) the occurrence dataset of the Bee Library - Big-Bee Project (Seltmann et al., 2021; 1.9 %) (iii) database of Iberian bees (Bartomeus et al., 2022); (iv) the National Biodiversity Data Centre (NBDC) program for Ireland data (www.biodiversityireland.ie, 1.1 %); (v) the *Hylaeus* spp. databases of H. Dathe (0.6 %); (vi) the *Colletes* spp. databases of M. Kuhlmann (0.12 %) and; (vii) Apoidea of the collections of Lyon, Aix-en-Provence, Marseille and Toulon Museums of Natural History (Meunier et al., 2023; 0.1 %). See Leclercq et al. (2023) for more details. Bee collection data are likely to have several inconsistencies (Dorey et al., 2023). We cleaned the distribution data by first only selecting observations within the study area, removing all observations that were duplicated in species, year, and coordinates, and removing observations that did not match the country checklist in Reverté et al. (2023). Using the clean coordinates function from the Coordinate Cleaner R package (Zizka et al., 2019) we also removed observation that; (a) matched institutional (excluding botanical gardens) locations and GBIF headquarters, (b) that matched country centroids, and (c) that were unique observations further than 250 km from any other observation of the same species. Although these datasets are geographically biased and potentially non-exhaustive, they cover the whole of Europe and are easily accessible and available and represent all widely known, publicly available datasets.

2.3. Establishing shortfalls

Following Hortal et al. (2015), we set out to explore seven common shortfalls of biodiversity data. For each shortfall, we measured different data impediments which provide a quantitative indication of its importance/magnitude, as well as its geographic patterns. We also introduce a shortfall related to the availability of high-quality illustrative material to aid in identification, a particularly useful asset for the development of citizen science initiatives. The Hutchinsonian shortfall was inappropriate to measure geographically here, given the data available, and is only discussed. Each of the shortfalls and the data impediments we use in this paper are described as follows (Fig. 1).

2.3.1. Linnean

The Linnean shortfall highlights an absence of described species and directly informs all other shortfalls (Hortal et al., 2015). While the Linnean shortfall is likely to be high for wild bee species at the global scale, Europe represents what is probably the best described wild bee fauna of any continent (Leclercq et al., 2023; Nieto et al., 2014). However, new species are regularly and still found in Europe (Rasmont et al., 2017; Wood et al., 2020), and the use of molecular tools will probably reveal new taxa among some of the current cryptic wild bee species complexes (Pauly et al., 2015) as well as potentially splitting or lumping existing species (Williams, 2022). Therefore, it is vital to support wild bee systematics and taxonomists in Europe, and the field should adapt to utilize new technological advances (Orr et al., 2021b). We assessed the Linnean shortfall exclusively at the country scale, using the year of species description to determine the rate of species description for each country, as an indicator of species discovery. We first computed the cumulative species richness per year, using the year a species was described for each year (Fig. S1), and then determined the annual growth rate as a percentage. Subsequently, we averaged the growth rate percentages for each area since 1990 as a proxy for the likelihood of discovering new species in each region. Higher values indicate ongoing species description and a greater probability of discovering undescribed species in those regions relative to areas with lower values.

2.3.2. Wallacean

The Wallacean shortfall concerns an absence of knowledge on the geographic distributions of species (Hortal et al., 2015; Lomolino, 2004). Historical, museum, and private collections are fundamental for our understanding of the niche and ecology of wild bees (Bartomeus et al., 2019). For wild bees, the Wallacean shortfall is often driven by biases in these historical collections, whereby certain regions, and countries with a long history of natural history collections, such as Germany, UK and the Netherlands, have a far more complete historical representation of species' geographic distributions (Meyer et al., 2015). Although, many of these collections, even in countries with large collections, are largely undigitized. Filling these gaps depends on the digitization of historical collection records (Meyer et al., 2015; Ascher et al., 2020), a costly and time-consuming process.

We quantified three impediments to calculate the Wallacean shortfall for wild bees. Firstly, we compared the discrepancy between the database of wild bee observations and the national checklists for each country. We recorded the shortfall as the proportion of known species with geo-referenced location data in comparison to the total known species. Secondly, we compared modeled predictions of taxonomic wild bee diversity against observed diversity based on occurrence records. Full methods for this section are available in Supplementary Notes S1 and in Leclercq et al. (2023). In short, we first created a grid consisting of 9563 cells of 25 × 25 km covering our study area. Then, we estimated species richness corrected for sampling completeness for all unbiased cells in Europe which represented the observed diversity (Supplementary Notes S1). We used a selection of the best-sampled of these cells to predict taxonomic diversity for all cells using a Generalized Linear Model (GLM). To compute the second proxy of the Wallacean shortfall, we calculated the differences between observed diversity and predicted diversity for each unbiased cell and averaged these at provincial and country scale. Cells that overlapped multiple regions were assigned to the region that covered it the most. We then normalized the log of the difference to between 0 and 1 and took an inverse percentage between 0 and 100, with 100 being the area with the greatest negative difference between observed and predicted and 0 the greatest positive difference. Finally, we calculated sampling completeness (coverage) for each 25 × 25 km covering our study area using the iNext package in R (Chao et al., 2021). We present this as sampling incompleteness by taking the inverse of the coverage (100 - coverage).

2.3.3. Darwinian

The Darwinian shortfall concerns the absence of knowledge on the evolutionary relationship between species (Hortal et al., 2015). The evolutionary relationships between wild bee species is useful to understand evolutionary rates (Danforth et al., 2013), co-evolution with flowering plants (Kooi and Ollerton, 2020), the origins of key traits in wild bees, e.g. eusociality (Danforth et al., 2003), variation in population (Lecocq et al., 2013), the impact of environmental changes on diversity (Hoiss et al., 2012) and the provision of ecosystem services (Grab et al., 2019) among others. The rapid increase in DNA sequencing capabilities has made understanding the relationship between species more accurate and easier (Yang and Rannala, 2012). In terms of wild bees, DNA sequencing has led to a much clearer understanding of the relationships between the seven wild bee families (Danforth et al., 2013). To address the Darwinian shortfall of European wild bees, we used the availability of publicly accessible gene sequences for all species as a data impediment for building phylogenies. We chose the commonly used COI gene sequence and its availability from the Barcode of Life Data System (BOLD; Ratnasingham and Hebert, 2007). While the COI gene may not necessarily be the best way to construct full phylogenies at the species level (Danforth et al., 2013), it is a conservative representation of the genetic data available for all European wild bee species. We worked under the assumption that the absence of COI genes was a reliable indicator of missing genetic data (number of species missing data) with which to build species level phylogenies at the European scale. However, the absence of DNA barcodes does not necessarily prevent the possibility of building accurate and useful phylogenies at a broad spatial scale using a diverse selection of species (e.g., Almeida et al., 2023; Henríquez-Piskulich et al., 2023) or using Linnean taxonomic hierarchies (Vereecken et al., 2021a, 2021b; Leclercq et al., 2023).

2.3.4. Prestonian

The Prestonian shortfall concerns the population dynamics of a species in time and space and requires having long-term estimates of species abundance (Cardoso et al., 2011; Hortal et al., 2015). In general, long-term abundance data for wild bees is not available at provincial or national scales (Potts et al., 2021) and represents perhaps the largest knowledge gap in understanding wild bee decline (Potts et al., 2016; Nieto et al., 2014). Here, we used two closely related data impediments related to the IUCN Red List, as a proxy for the absence of meaningful population data. We chose data deficiency in the European Red List of bees (Nieto et al., 2014). as a proxy for absence of meaningful population data. According to the IUCN (2001), a species is classified as data deficient when there is inadequate information to make a direct, or an indirect, assessment of its risk of extinction based on its distribution and/or population status". However, data deficiency on the red list is most probably an underestimation of abundance and population gaps; 57 % of species in Europe were classified as data deficient but 79 % of species had no data on population trends (Nieto et al., 2014). Using the updated database of observations, we recalculated population trend data using a simplified approximation of IUCN criteria. Therefore, the shortfall was measured as; (i) the number of species without a minimum of three observations for at least ten separate years after 1950 and (ii) number of data deficient species in each province and country.

2.3.5. Raunkiærian

Hortal et al. (2015) characterize the Raunkiærian shortfall, (also called Raunkiæran shortfall), as the lack of ecologically significant data concerning functional trait variation within and between species. Traits which affect a species interaction with environment and other species to influence performance and fitness can be classified as functional traits (McGill et al., 2006). For wild bees, this includes variation in traits which influences their ability to access food and nesting resources, and their ability to provide pollination services. To quantify the Raunkiærian shortfall for European wild bees, we focused on six ecologically relevant functional traits which have been used in several studies to understand decline, distribution patterns and community diversity (e.g., De Palma et al., 2015; Marshall et al., 2015; Aguirre-Gutierrez et al., 2016; Weekers et al., 2022). We extracted traits from the "European bee traits database" (established by ALARM, www.alarm-project.ufz.de, developed by STEP, www.STEP-project.net and maintained and updated by S. P.M Roberts). The six traits were: (i) body size (averaged intertegular distance of females, where the wings join the thorax); (ii) sociality (a. solitary, b. communal, c. primitively eusocial, eusocial, and parasitic); (iii) nesting habit (above- or below-ground, renters or excavators); (iv) larval feeding specialization, i.e., pollen collection, (a. polylectic species, focused on a variety of different plant taxa, b. oligolectic species, focused on various plant species, but mainly on one family, and c. monolectic species, focused on a single plant taxon (genus or species); (v) voltinism (number of generations per year, a. uni-, b. bi- or c. multivoltine) and; (vi) how they transport pollen (a. accidental, b. corbiculae, c. crop, d. legs and body, e. legs only, and f. abdomen). The shortfall was measured per province and country as the both the average percentage of 'trait completeness' (how many of the six traits have been recorded) and whether any one of the six traits is missing for species known from that area. The traits we have selected here are commonly used to understand trait diversity in bees, but they are biased towards easily measurable traits. More difficult-to-measure traits that may have a stronger influence on fitness, such as lifespan and offspring production, are likely to be missing for all but the most common and well-studied species.

2.3.6. Hutchinsonian

The Hutchinsonian shortfall is defined by Hortal et al. (2015) as an absence of knowledge about the tolerance of species to abiotic conditions. A bee's tolerance to abiotic conditions can be estimated in two ways. The first is through experimental work of physiological responses to different conditions, which so far is mostly constrained to a selection of European bumblebees (Martinet et al., 2021). The second is through environmental niche modeling (ENM) where occurrence records are used to obtain a statistical representation of the limits of the abiotic niche (Elith and Leathwick, 2009). Again, at the European scale, this has mostly been done for bumblebees (Marshall et al., 2018). Additionally, for ENM, both the Wallacean and Prestonian shortfalls are critical to how well we can model abiotic tolerances. For these reasons, we did not explicitly represent the Hutchinsonian shortfall in this work.

2.3.7. Eltonian

The Eltonian shortfall concerns knowledge of the interactions that the taxon of interest has with other species (Hortal et al., 2015). For wild bees, a key interaction is the visitation of flowering plants to collect pollen or nectar (Ollerton et al., 2011). These are often measured by observing interactions in situ and recording every time a bee is seen to be visiting a certain plant species (Dafni, 1992). This data can be = sed to determine feeding preferences of certain wild bee species and create plant-pollinator networks, quantifying how the entire community of wild bees and plants interacts (Memmott, 1999). Here, we used a broad data impediment for our knowledge of wild bee interactions in Europe. For each species recorded in Europe, we checked whether there was any recorded interaction with other species on GloBi (Poelen et al., 2014), Mangal (Vissault et al., 2019), web of life (www.web-of-life.es), DoPI (Balfour et al., 2022) and FlorAbeilles (Gombault et al., 2018). We then quantified and mapped the sum of species without any known interactions for each province and country.

2.3.8. Keartonian impediment

Here, we propose to add a new impediment, strongly connected with the other shortfalls, which concerns the availability of visual representations of what a species looks like. We have named this the Keartonian impediment after Richard (1862–1928) and Cherry Kearton (1871–1940), who have been credited as establishing the field of wildlife photography (Bevis, 2016). We propose and argue that visual representations of species represent fundamental knowledge related to "what species look like", "which traits to look for" and contribute to our understanding of species ecology and evolution. These data are valuable for identification, monitoring, and outreach, particularly in countries without access to established museum collections. When such representations are lacking, it can hinder our understanding of a species' appearance, behavior, and habitat, making it more challenging to study, conserve, and communicate about.

The majority of the 20,000+ bee species on Earth (Ascher and Pickering, 2020), their general habitus nor their diagnostic traits have never been photographed. This is one of the issues that prompted the European Commission to support the ORBIT project (ORBIT, 2022) with the goal to merge the available illustrations on the 2000+ species of bees in Europe through networking with photographers and field trips aimed to bring back the first photos ever of a wide range of known species. These visual representations can highlight key diagnostic characteristics, convey specific information (size, color, diagnostic traits, etc.) and can help fill other shortfalls in ecological and evolutionary knowledge. Photographs for reference purposes, alongside new photographs for identification will be key aspects of large-scale monitoring schemes

(Potts et al., 2021). They are also necessary to train machine learning tools focused on the automatic identification of species from photographs, technological advances are making this a reality for insects, including bees (Spiesman et al., 2021). Here, we use the absence of high-quality in-situ photographs as measure of this data impediment. This could also include videos and 3D models. We have utilized two sources independently; (i) 'flickr.com' as a photographic database to check the availability of photographs for all species, focusing only on those with expert validated identifications; (ii) iNaturalist observations classified as research grade (community agreement on the identification from the photo). The two sources are complementary, with Flickr.com having good photos and poor IDs in general, while iNaturalist has a much greater proportion of poor photos but many are good enough to provide IDs. Some of the iNaturalist photos may represent pinned specimens not in situ but we expect that this number is small and therefore still functions as useful proxy. The impediment was measured as the number of species without publicly available, validated photographs on either of the two platforms.

2.4. BeeFall online shiny tool (beefall.org)

The BeeFall tool is an online Shiny app and was created using R statistics (v4.1.3; R Core Team, 2022) and Shiny (v1.7.1; Chang et al., 2022) to provide an accessible, spatially explicit visualization of the European bee data shortfalls. We took the aforementioned shortfalls, which were then aggregated to country and provincial shapefiles in order to visualize the shortfalls as maps. For the country scale, we used the checklist as an overview of what species were expected to occur, and at the provincial scale, we used the occurrence records from each administrative unit as a proxy checklist. Currently, the tool allows users to select the shortfalls and scale of interest and interact with the map. Both countries and administrative units can be clicked on to show the list of species missing data for each shortfall. At the country scale, it is possible to visualize the whole checklist and see which species are missing occurrence records (Wallacean). For the other shortfalls, a table displaying all species with missing data is available. For the Raunkiærian shortfall, it is possible to see for each species which trait data is missing. Additionally, we provide the option to visualize two shortfalls simultaneously by using bivariate choropleth maps. The goal of the tool is to stimulate researchers to provide novel data to fill the gaps. The quantification of shortfalls is only useful if it can be used to stimulate additional research to fill them.

3. Results

The gaps in European wild bee data show clear and convergent geographic patterns. In general, greater data gaps exist in Southern Europe. We found the greatest shortfalls at the national scale in countries where the wild bee fauna is relatively well known but poorly studied compared to the North of Europe, such as in Spain, Portugal, Italy, and Greece (Fig. 2a). Conversely, countries in Eastern Europe exhibit fewer overall data gaps, but this was mostly because spatial data was missing for much of the fauna. The final checklist for our selection of European countries includes 2030 species of wild bees. Missing data for one or more shortfalls was the default for most species. In terms of the Wallacean shortfall, there was publicly available geo-referenced occurrence data for 81 % of the bee species on the checklist for our selection of European countries (Table 1, Fig. 3a). On average, *per* country, the discrepancy between the checklist and the availability of digitized occurrence records was 33 %. However, this discrepancy varied massively between countries (SD = 24 %). The five countries with the greatest discrepancy were, in order, Albania (72 % missing), Serbia (71 %), the island of Sicily (67 %), Montenegro (63 %) and the island of Sardinia (63 %). Conversely, the five countries with the highest occurrence records of the checklist were, Sweden (0 %), Finland (1 %) Norway (2 %), Isle of Man (2 %), and Switzerland (2 %), (Fig. 2a). When looking at modeled sampling completeness, based on the difference between observed and predicted richness, we saw the same high variation between countries. The five countries with lowest values and therefore, the biggest discrepancy between the observed and predicted diversity were mostly islands: Iceland, Azores, Madeira, Bosnia and Herzegovina, and the Canary Islands. Although this is likely a cause of the fact that the model does not specifically consider island characteristics, most importantly isolation, resulting in overpredictions. The other 4 mainland areas with the lowest values were Albania, Serbia, Bulgaria, and Montenegro. The top five countries were Belgium, Switzerland, Lichtenstein, Luxembourg, and the Netherlands, mostly well-sampled, relatively small countries. In terms of coverage the countries with the lowest sampling completeness, averaged across all 25 × 25 km squares in the country, were Greece (56 ± 27 %) Cyprus (58 ± 28 %), Sicily (62 ± 29 %), the East Aegean Islands (62 ± 24 %) and Montenegro (64 ± 19 %). On the other hand, the countries with the highest sampling completeness were Belgium (99 ± 1 %), the Netherlands (99 ± 2 %), Luxembourg (98 ± 1 %), Switzerland (97 ± 3 %), and Liechtenstein (97 ± 1 %) (Fig. 3b). The Linnean shortfall was highest in Southern Europe, specifically islands, with the Canary Islands, Madeira, East Aegean Islands,

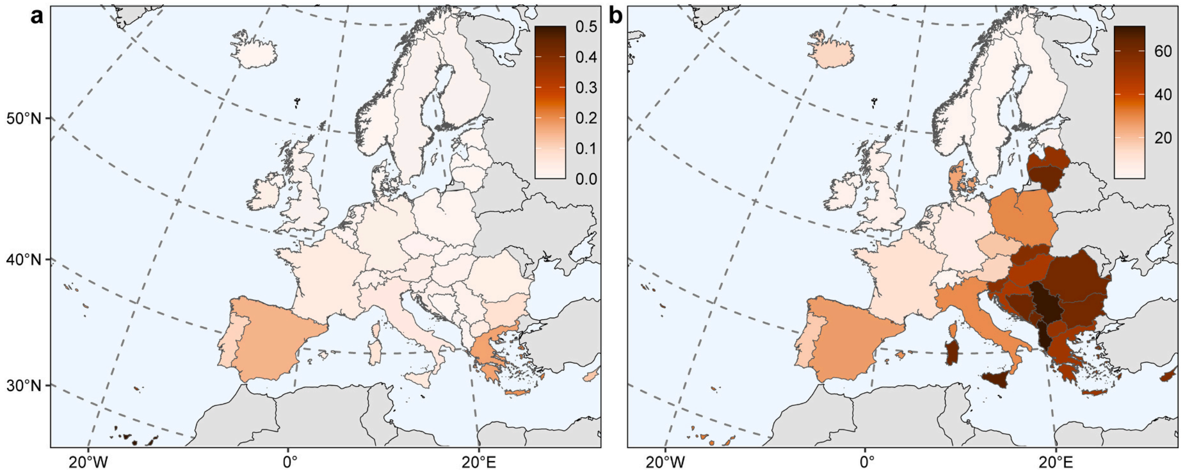

Fig. 2. Country level shortfalls: new species description rate (%) and checklist incompleteness (%). (a) New species description rate (Linnean shortfall) measures the average percentage rate of new descriptions per year in each country since 1990. (b) Checklist incompleteness (Wallacean shortfall) measures the percentage of species on the checklist for which there is no distribution data in the digitized datasets available. For further detail on spatial variation within and between countries see BeeFall Online shiny tool (beefall.org). European map projection: EPSG:3035, ETRS89-extended / LAEA Europe.

Table 1
Number of species per rank of number of observations in full distribution dataset.

Number of observations	1	2	3–5	6–10	11–25	26–50	51–100	101–500	>500
Number of species	94	68	158	1238	177	141	162	320	434

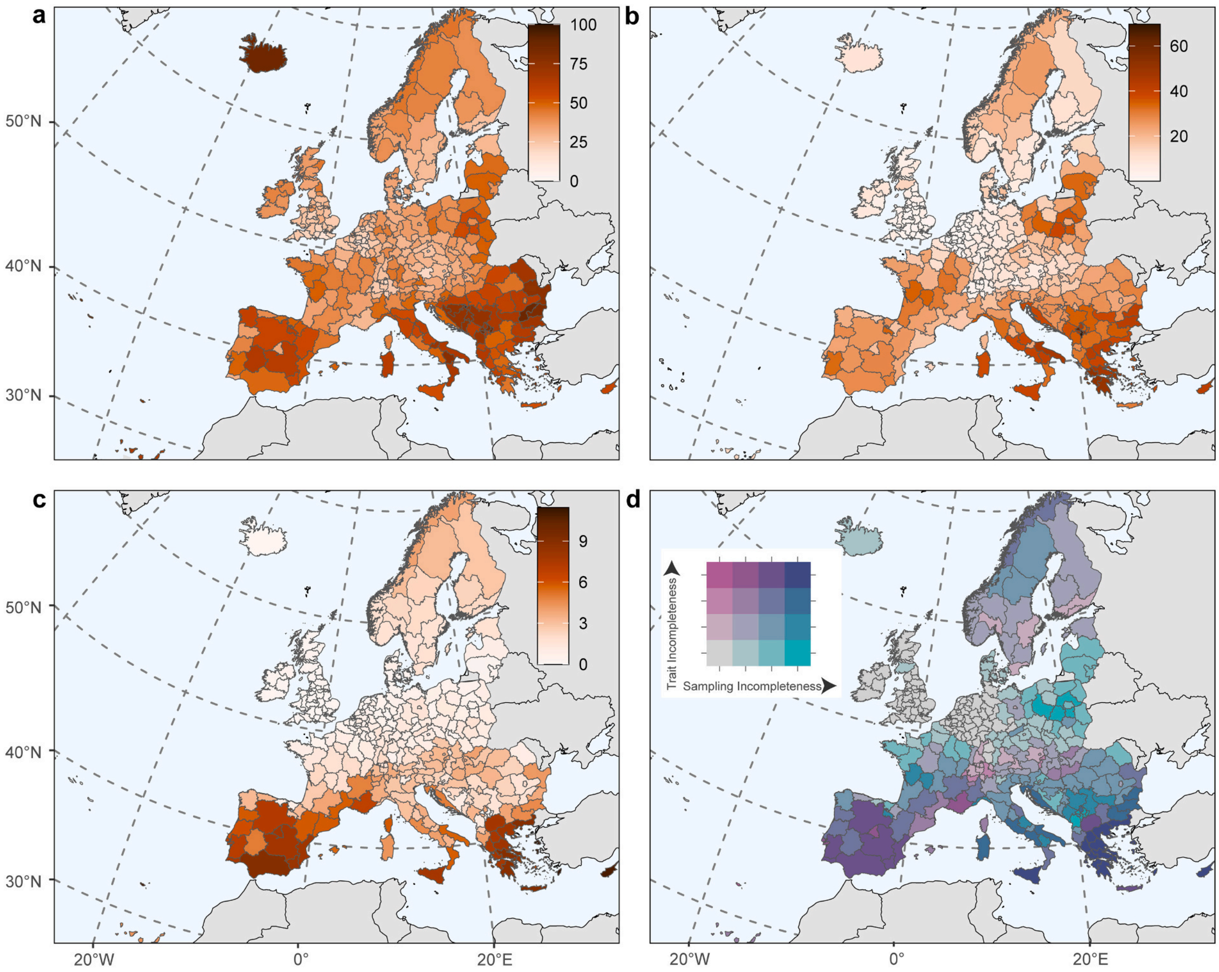

Fig. 3. Percentage provincial level shortfalls: predicted vs actual diversity (%), sampling incompleteness (%) and trait incompleteness (%). (a) Wallacean shortfall, shortfall in distribution data as measured by percentage (%) difference between predicted diversity and actual diversity and then scaled between 0 and 100. (b) Wallacean shortfall, shortfall in distribution data as measured by sampling incompleteness (%; inverse coverage). (c) Raunkiærian shortfall, trait incompleteness measured as the percentage of missing traits from these 6 traits; (1) body size (intertegular distance of females, where the wings join the thorax), (2) sociality (solitary or social), (3) nesting habit (above- or below- ground, renters or excavators), (4) feeding specialization (oligolectic, feeding on one plant species or polylectic, feeding on multiple plant species), (5) voltinism (number of generations per year, uni-, bi or multi-voltine) and (6) how they transport pollen. (d) Example of a bivariate map showing the relationship between sampling incompleteness and trait incompleteness. N·B: Iceland was removed from the predicted vs actual diversity map because it represented a significant outlier. For further detail on spatial variation within and between countries see BeeFall Online shiny tool (beefall.org). European map projection: EPSG:3035, ETRS89-extended / LAEA Europe.

Crete, Azores, Greece, Spain, Cyprus, and Portugal having a mean discovery rate since 1990 above 0.1 % (Fig. 2b).

The other shortfalls consisted of counts of species missing the appropriate data (Table 2). At the national scale, we show this in relation to the checklist and at the provincial scale in terms of collected and digitized occurrence data. The shortfall currently with the greatest number of species missing data was the Prestonian shortfall, it would not be possible to calculate population trends for at least 69 % of species, while 48 % of species were data deficient on the 2014 Red List for European wild bees. Next came the Eltonian shortfall, with 58 % of species not having online publicly available interaction data. This is followed by the Raunkiærian shortfall, 57 % of species missing data from at least one ‘key’ trait, although across all 2004 species the average trait completeness of the 6 traits was 94 % ± 12 % (Fig. 3c). Finally, the Darwinian shortfall had the fewest species missing data, 42 % of species without published COI gene sequences. (Fig. 4). For the newly proposed Keartonian impediment, 54 % of species were missing high-quality *in natura* (i.e., in their habitats) or in situ (e.g., photographed on a white background locally) photos, according to Flickr. This increases to 63 % when we consider iNaturalist as the photo source. Joining the two databases together decreases this value to 49 %. All shortfalls measured show clear relationships and display country and provincial correlations (Fig. 5 & S2). Specifically, we saw the number of species missing data increased with total known richness within a country. Overall, Greece and its larger island groups consistently had the largest knowledge gaps and small north-western areas, such as the Isle of Man, Northern Ireland, the fewest (Table 2). We detected some anomalies, for example the East Aegean Islands and Cyprus have moderately high richness, but a higher

Table 2
Summary table of shortfalls per country.

		Number of Species			Percentage of Checklist		
Shortfall	Definition	Average	Lowest	Highest	Average	Lowest	Highest
Linnean[a]	Absence of described species	–	–	–	0.06[a] (±0.10)	0[a] (Multiple[d])	0.51[a] (Canary Islands)
Wallacean (Checklist)	Absence of knowledge on geographic distributions, checklist incompleteness	154.4 (±159.4)	1 (Multiple[e])	553 (Greece)	32.5 (±24.3)	0.36 (Sweden)	71.6 (Albania)
Wallacean[b] *(Model)*	Absence of knowledge on geographic distributions, difference between modeled diversity minus observed	–	–	–	25.4[b] (±52.7)	−114.8[b] (Belgium)	93.3[b] (Azores, Iceland)
Wallacean[c] *(Coverage)*	Absence of knowledge on geographic distributions, measured as sampling incompleteness (inverse coverage)	–	–	–	19.8 (±12.3)	0[c] (Belgium)	43.8[c] (Greece)
Darwinian	Absence of knowledge on the evolutionary relationship between species	59.1 (±79.4)	0 (Multiple[f])	394 (Greece)	15.7 (±15.9)	0 (Multiple[f])	54.1 (Canary Islands)
Prestonian (Population trend)	Absence of knowledge on long-term estimates of species abundance	135.6 (±156.5)	0 (Multiple[g])	662 (Greece)	24.7 (±21.5)	0 (Multiple[g])	76.3 (Canary Islands)
Prestonian (Data Deficient)	Absence of knowledge on threat status based on IUCN Red List	131.0 (±112.4)	1 (Iceland)	474 (Greece)	27.6 (±79.0)	10.4 (Northern Ireland)	54.8 (Canary Islands)
Raunkiœrian (trait incompleteness)	Proportional absence of ecologically relevant information on inter- and intra-species trait variation	–	–	–	4.5 (±3.2)	0 (Iceland, Isle of Man)	11.6 (Greece)
Raunkiœrian (missing traits)	Any absence of ecologically relevant information on inter- and intra-species trait variation	115.6 (±122.7)	0 (Iceland, Isle of Man)	515 (Greece)	21.5 (±13.5)	0 (Iceland, Isle of Man)	47.4 (East Aegean Islands)
Eltonian	Absence of data on species interactions	117.8 (±119.7)	0 (Iceland, Northern Ireland)	495 (Greece)	22.3 (±12.5)	0 (Iceland, Northern Ireland)	44.7 (Greece)
Keartonian[h]	Absence of high-quality, correctly identified in-situ photos of species	91.6 (±101.2)	0 (Iceland, Northern Ireland)	482 (Greece)	16.6 (±11.8)	0 (Iceland, Northern Ireland)	44.4 (Cyprus)

[a] Does not show the percentage of the checklist but shows the mean percentage rate of discovery per year since 1990.
[b] Does not show the percentage of the checklist but shows the percentage difference between predicted diversity and observed diversity.
[c] Does not show the percentage of the checklist but shows the inverse of coverage or sampling incompleteness.
[d] Alend Islands, Andorra, Bosnia & Herzegovina, Denmark, Estonia, Iceland, Isle of Man, Latvia, Liechtenstein, Lithuania, Northern Ireland, and Norway.
[e] Iceland, Isle of Man, Northern Ireland, Sweden.
[f] Alend Islands, Estonia, Finland, Iceland, Ireland, Isle of Man, Liechtenstein, Northern Ireland, and Norway.
[g] Iceland, Ireland, Isle of Man, Northern Ireland.
[h] The Keartonian Impediment row joins the information obtained from both Flickr and iNaturalist.

proportional number of gaps compared to other countries (Fig. 5a). In terms of the relationship of other shortfalls to the Wallacean shortfall, we found clear relationships but again with anomalies. The better sampled a country was the fewer gaps there were, countries which were poorly sampled (observed richness lower than predicted) fall into two groups those with few gaps and those with many (Fig. 5b). A similar pattern was observed for checklist completeness (proportion of checklist with available occurrence records). Countries which had both high and low discrepancies in checklist completeness had few gaps in the other shortfalls. Most countries with large values for the other shortfalls had intermediate checklist completeness (Fig. 5c). At the provincial scale, this quadratic relationship could be clearly seen, countries with a high Wallacean shortfall did not have enough extra knowledge to be able to classify the other shortfalls (e.g., eastern countries). On the other hand, countries such as Spain, Greece, and Italy, which were comparatively well studied, had enough knowledge to be able to classify the many gaps there were for the other shortfalls (Fig. S2).

4. Discussion

Here, we highlight and quantify for the first time the major shortfalls impairing knowledge and conservation of European wild bees, and we provide a user-friendly, interactive online tool to visualize the geographic patterns of these shortfalls and raise awareness. The results indicate that despite the large amount of wild bee occurrence records available, there are still many critical gaps in our knowledge that are hampering adequate monitoring and conservation of all species. Furthermore, we observe that the availability of data concerning wild bee diversity at the European scale shows extreme spatial variability. For some countries such as Spain, Italy, and Greece, we have reached a good understanding of the species composition. This situation highlights the many gaps we have for the other shortfalls, in comparison, for other countries, including the Baltic and Balkan states, due to a poorer understanding of the species composition, there are fewer gaps in the other shortfalls, suggesting unknown gaps. These results call for increased investment into wild bee research and monitoring in these areas, with targeted national and international support to deploy fundamental approaches to generating, digitizing, and sharing data. It should be noted that wild bees are one of the better studied insect groups, especially in terms of their relationship with ecosystem services (Noriega et al., 2018), yet there are still many knowledge gaps, implying that these gaps are likely much larger for many other insect groups.

The shortfalls we highlight here limit our understanding of wild bee diversity and conservation but also present clear pathways for ensuring these gaps can be filled. The Wallacean (absence of distribution data) shortfall provides the greatest difficulties in terms of generating bee data to fill the other shortfalls as well. With current data, we are likely severely underestimating the ranges of some species, specifically towards the East. The consequences of such underestimations are that certain taxa and/or areas should be considered as conservation priorities (Riddle et al., 2011), and so prioritized for surveying and monitoring. The Wallacean shortfall has however been made a priority at the

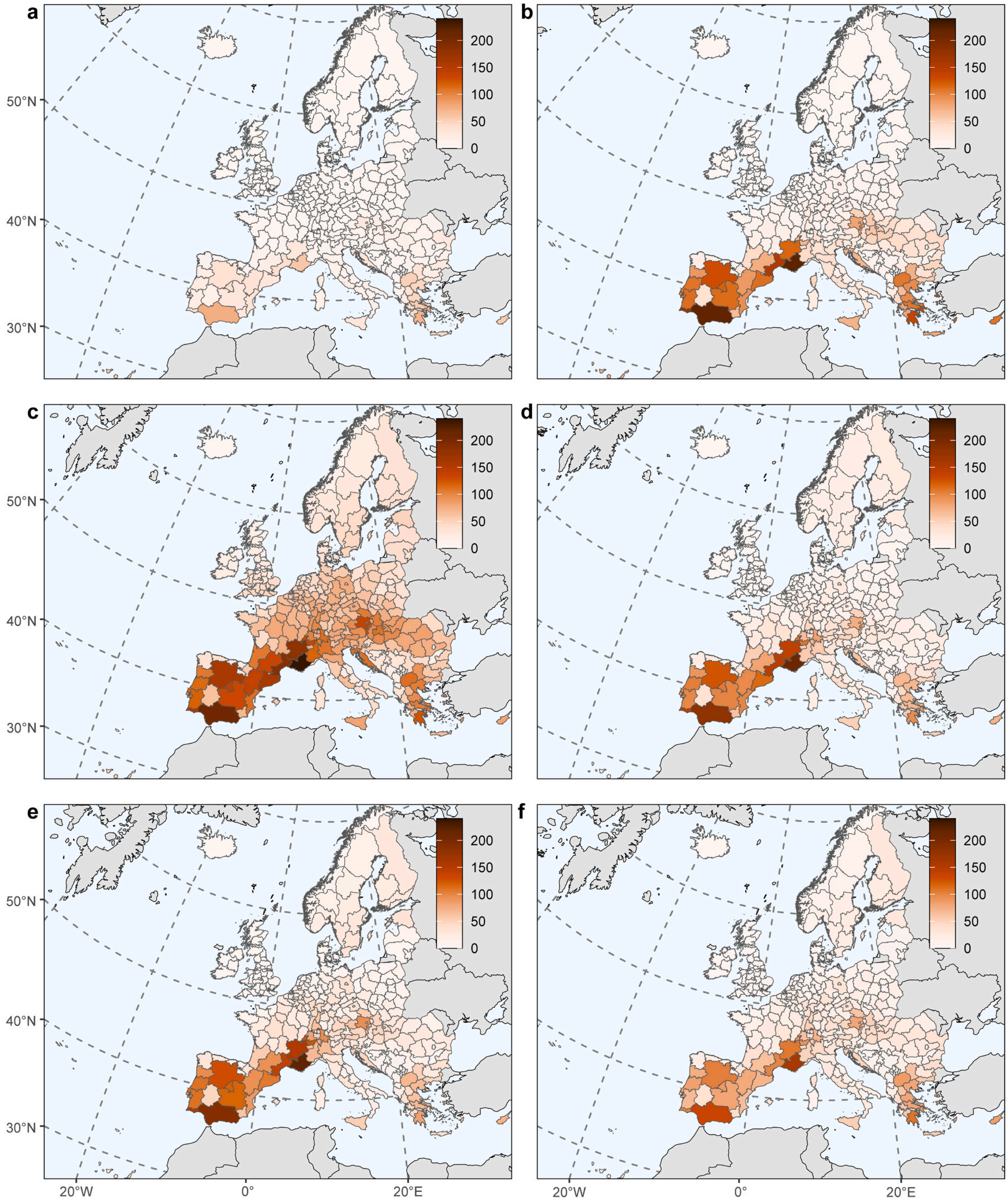

Fig. 4. Species count provincial aggregation of wild bee shortfalls in Europe. (a) Darwinian shortfall, count of species missing genetic information, measured as species with COI gene data on BOLD. (b) Prestonian shortfall, count of species for which population trends cannot be calculated, i.e., the number of species without a minimum of 3 observations for at least 10 separate years after 1950. (c) Prestonian shortfall, count of species missing long-term population measured as species 'data deficient' in the EU Red List (Nieto et al., 2014). (d) Raunkiærian shortfall, count of species missing one or more of the six key traits; (1) body size (intertegular distance of females, where the wings join the thorax), (2) sociality (solitary or social), (3) nesting habit (above- or below- ground, renters or excavators), (4) feeding specialization (oligolectic, feeding on one plant species or polylectic, feeding on multiple plant species), (5) voltinism (number of generations per year, uni-, bi or multi-voltine) and (6) how they transport pollen. (e) Eltonian shortfall, count of species without information of the biotic interactions with plants on publicly accessible databases (mangal.io and web-of-life.es). (f) Keartonian impediment count of species missing high-quality publicly available photographs (on flickr.com and inaturalist.org). For further detail on spatial variation within and between countries see BeeFall Online shiny tool (beefall.org). European map projection: EPSG:3035, ETRS89-extended / LAEA Europe. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

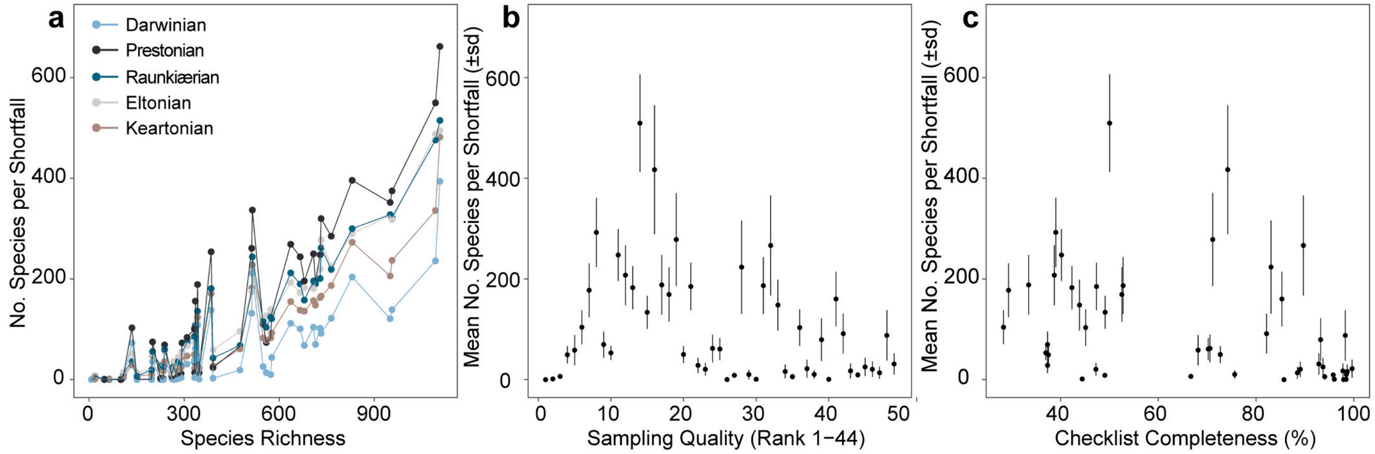

Fig. 5. Relationship between the different shortfalls. (a) The number of species missing for each shortfall per country against the overall species richness of that country. (b) The mean and standard deviation of the number of species missing for the five shortfalls shown in *(a)* per country against the overall sampling quality, measured as a rank of countries in terms of lowest to highest difference between observed and predicted richness. (c) The mean and standard deviation of the number of species missing for the five shortfalls shown in *(a)* per country against the overall checklist completeness, measured as a percentage of how many species on a country's checklist have digitized occurrence data.

continental scale as part of the European Union pollinator monitoring scheme (EU PoMS), which plans to monitor 2000–3000 sites across all European member states for insect pollinators, including wild bees (Potts et al., 2021).

The EU PoMS will also provide abundance data and will help to fill the Prestonian (absence of population data) shortfall. The Prestonian shortfall when measured as data deficiency in the European Red List had the fewest gaps overall but this is in part due to the data impediment we used. Due to the shortfalls, we have outlined here, the threshold to be assessed for the European Red List of wild bees was permissive and did not incorporate dynamics across time as a specific requirement (Nieto et al., 2014). Therefore, this estimation of the Prestonian shortfall is likely severely underestimated and may more closely resemble an example of the Wallacean shortfall or even the Ostromian Shortfall proposed by Lopes-Lima et al. (2021) regarding the application and effectiveness of conservation measures. Abundance data is likely to be severely limiting and indicates the difficulty in establishing whether wild bees are in decline in all areas (Potts et al., 2016). Therefore, our alternative measure related to consistent sampling of multiple specimens across multiple years provides a more realistic representation of this shortfall. We have not addressed another notable knowledge gap, known as the Haeckelian shortfall (Faria et al., 2021). This shortfall pertains to a lack of knowledge of a species' semaphoronts, which are distinct life cycle stages. For bees this shortfall appears to be substantial. Nests are known from only a small fraction of the fauna and our understanding of the development stages is even more limited. However, it is worth noting that this shortfall is likely less critical for biodiversity analyses and data acquisition as developmental stages are restricted to the nest.

The simplified proxies for the shortfalls we use here effectively highlight the gaps in our knowledge but may in turn underestimate the scale of the gaps. By focusing only on the availability of COI barcodes for bees we do not necessarily highlight the biggest challenges establishing and using robust phylogenies of wild bees including improving our understanding of rates of evolution in wild bees (Murray et al., 2018) and trait evolution (Litman et al., 2011). Although, alternatively we may be overestimating the Darwinian shortfall, as it is possible to build complex phylogenies, which can answer important ecological and biogeographical questions about bee diversity, even with a subset of species (e.g., Hedtke et al., 2013; Almeida et al., 2023; Henríquez-Piskulich et al., 2023). Our estimation of the data impediment for the Eltonian (absence of interaction data) shortfall is likely an underestimation of the known interactions for many species due to the dispersed nature of the biotic interactions databases, which is the case across taxa, not only for plant-pollinator interactions (Poisot et al., 2021). Yet, there are several databases (as used here) which are growing in data that can be a more reliable proxy for the Eltonian shortfall in the future (Poelen et al., 2014; Vissault et al., 2019).

The underlying causes of the shortfalls described above can be summarized and attributed to the absence of one or more of three steps. Firstly, there might be a genuine lack of specimen observations and therefore a lack of actual data. This is where establishing Europe-wide monitoring programs such as the EU PoMS is so vital, as it will generate a new wave of systematic biological records, including from under-surveyed regions. Secondly, the data exists but has not been digitized and therefore not been made available. Recent decades have seen a massive increase in the digitization of biodiversity data which has stimulated new research into large-scale patterns of wild bee diversity (Nelson and Ellis, 2019). Digitizing museum collections of wild bees in Europe will likely contribute to generating more robust analyses of biodiversity trends over time, filling in knowledge of historical distributions and filling both the Wallacean shortfall and Prestonian shortfall (Bartomeus et al., 2019; Johnson et al., 2023). For example, Naturalis Biodiversity Center (Netherlands) hosts and curates one of the largest natural history collections in Europe, yet of the bees, only the bumblebees have been digitized, leaving hundreds of thousands of non-*Bombus* specimens ranging from the 18th century to the present undigitized (J.C. Biesmeijer, personal communication, August 2022). Finally, the digitized data might be available, but they have not (yet) been shared. Wild bee data in Europe may represent an economically important resource which disincentivizes communities to share data, such as in academic research, where holding on to data can ensure publications and grants (Wetzel et al., 2018). Incentives to share data including correct citations, attributions and authorship credits may help to mobilize this data into continental databases (Wetzel et al., 2018). Newly described or revised species may not have distribution data available immediately through online databases, which creates a time lag. As a result, even countries with well-sampled data may not have accessible georeferenced distribution data available for all species on the checklist. Directly digitizing and uploading specimen location records used in species descriptions to GBIF, would minimize this problem.

The diversity, and particularly both the magnitude of individual shortfalls and impediments, as well as their interconnectedness, hamper our comprehensive understanding of bees in Europe. First, the combination of these shortfalls creates a substantial barrier to a comprehensive understanding of bees in Europe by limiting our ability to conduct meaningful biodiversity assessments, hindering the formulation of effective conservation policies and practices. For example, limited information on the functional traits of different bee species hinders our ability to understand the ecological roles of different bee species, their adaptations to specific environments, and their responses to environmental changes. Likewise, the lack of comprehensive data on where different bee species are distributed across Europe hampers efforts to assess the regional/national/continental diversity of bee species, highlight critical habitats, identify drivers of biodiversity across larger areas and formulate evidence-based, targeted conservation strategies. Perhaps even more critically, our inability to accurately identify bee species using expert literature (keys) and illustrations of diagnostic traits (photos) can jeopardize efforts to study the ecological roles of different species and attribute specific functions to different bee species to understand their unique contributions to ecosystems. Second, we argue that the interconnected nature of these shortfalls has the potential to exacerbate their individual impacts such as those described above. Indeed, the lack of data on one aspect, such as species identity or biotic interactions, can cascade into difficulties in addressing other shortfalls. For instance, without accurate species identification (Linnean), understanding evolutionary relationships (Darwinian) and ecological interactions (Eltonian) becomes more challenging, if not altogether impossible. Also, until a more comprehensive traits and geographic distribution database will be available for the bees, we will struggle to address pressing issues in the emergent field of functional biogeography explain gradients in trait diversity/distribution (see Leclercq et al. (2023) on species and phylogenetic diversities), predict ecosystem functioning and services across larger region, explore how bee populations will respond to changes in their habitats and environmental conditions, and create conservation actions with a functional basis.

Although knowledge shortfalls tend to be addressed individually, it is important to note that they are not mutually exclusive, and that they in fact interact significantly and in multiple directions (Hortal et al., 2015). The most important shortfall in terms of leverage will always be the Linnean (absence of described species) shortfall and the identification, description, and delineation of species through traditional, but increasingly, modern taxonomy (e.g., biosystematics) techniques remains a fundamental task (Orr et al., 2021b). Moreover, the Linnean shortfall can be thought of as two components; (i) the discovery of unknown species and (ii) the splitting and lumping of already described species, both of which impact the other shortfalls in different ways (Diniz-Filho et al., 2013). The former may represent a smaller knowledge gap for a well-studied region such as Europe compared to other continents. Whereas the latter may represent significant future changes to our knowledge of bee diversity in Europe, for example how different species concepts have been applied through time to bumblebees has led to periods where species were more likely to be split or lumped together (Williams, 2022). Splitting of species also poses challenges with respect to the validation of historical records. The verification of historical occurrences requires specimen examination by an expert, a process that is often not possible and, when feasible, is time-consuming and logistically difficult. It is not possible to correctly attribute all records from species that have undergone splitting in the shiny tool, instead we have highlighted the species that have recently been updated (Ghisbain et al., 2023) to indicate that there may be some discrepancies in terms of their distribution.

By introducing the Keartonian impediment we hypothesize that high-quality visual representations of species *in natura* or in situ can be a hitherto overlooked but effective tool to leverage the other shortfalls, by aiding in species identification, by encouraging citizen scientists and expert amateurs, and by acting as a tool to help highlight the existing gaps and importance of this data (Seppänen and Väliverronen, 2003). The availability of large numbers of photographs illustrating a wide taxonomic range of wild bee species is still under development; it is often a particularly challenging and sometimes tedious task as it requires to get series of photographs *in natura* or in situ showing the diagnostic morphological characters, or to collect the specimens photographed, have their identification confirmed by an expert, and curate the specimens appropriately in a reference collection. Yet, despite these obstacles, the availability of these photographs is increasingly considered as a key asset for outreach, but also to feed artificial intelligence and machine learning algorithms and programs that aim to generate automatic identification (Spiesman et al., 2021). Speeding up the ability of collectors to make accurate identifications will directly benefit the Linnean and Wallacean shortfalls and in turn all the others. Single specimen records can be utilized to fill multiple gaps in our knowledge, and those collecting specimens should be encouraged to not only provide location and taxonomic data but to also record ecological and morphological traits, observed interactions, and to make the specimens available for further research and genetic analyses. In the case of the Raunkiærian (absence of trait data) shortfall, much progress could be made on morphological traits made by taking measurements from bee specimens housed in museums and private collections (Suarez and Tsutsui, 2004; Kendall et al., 2018). Additionally, future updates of the tool should include measures of intraspecific trait variation, which can influence overall functional trait variability and obfuscate interspecific distinctions (Albert et al., 2010). The inclusion of intraspecific variability alongside interspecific comparisons also provides critical insights into species' adaptive capacity, population dynamics, and ecological responses to environmental changes (Siefert et al., 2015). Finally, there is a need for increased research into traits that directly affect how species respond to environmental changes (i.e. response traits, Lavorel & Garnier, 2002) such as those affecting demographic performance and population trends (e.g. lifespan, fecundity) and ability to adapt (e.g. flight range, nesting, and diet generalization level) as well as traits that affect how species affect ecosystem functioning (i.e. effect traits, Lavorel & Garnier, 2002), such as those affecting pollination efficiency (e.g. body size, tongue length). In terms of interactions with flowers a more varied classification of diet specialization in terms of pollen as larval food source and nectar consumption alongside other uses, such as oil collecting, could provide a more complete picture of the relationship between fitness and floral resource use. A combination of these improvements will extend our understanding of the functional relationship between traits and pollination services (Chase et al., 2023). Population trends will likely remain the hardest to fill (Bartomeus et al., 2019) and that is why the focus of the EU PoMS scheme is based around this goal (Potts et al., 2021). The BeeFall tool will be effective in highlighting areas that would benefit most from detailed long-term surveys.

We envisage BeeFall as an essential first step in raising awareness and providing an impetus to fill the gaps in European wild bee data, as well as to illustrate and help recognize that there is much work still to be done, both in categorizing these shortfalls as well as in establishing programs to generate the data to fill them. Data holders are encouraged to share their private data to the online repositories included in this study (Table S2) and this information can be added to future updates of the tool (see e.g., Callaghan et al., 2022). Furthermore, we believe that BeeFall can also serve as an evaluation tool to measure progress in data collection at the EU scale. However, it is important to highlight that these shortfalls should not be used as evidence that existing research into European wild bees is inaccurate or lacking. Data gaps are universal, particularly for vast groups of organisms like insects, using model specimens and testing ecological theories is a necessary and important step towards a better use of currently available information (Diniz-Filho et al., 2010). BeeFall and its underlying databases should be maintained and regularly updated to be useful as a measure of progress; to this end, we will also develop a "flagging" feature in BeeFall to encourage users to flag errors or inaccuracies. To make BeeFall a permanent surveillance tool at the European scale, the underlying shortfalls databases, and the tool itself will require longer-term financial support, governance and oversight, in particular good communication and cooperation between experts and data holders in Europe (Costello et al., 2014). We also see the possibility to highlight the gaps in particular locations and generate citizen programs, such as iNaturalist projects, which can be used to engage many different types of participants (Salmon et al., 2021; Vereecken et al., 2021a, 2021b).

Europe has one of the most extensively studied bee faunas globally (Leclercq et al., 2023). This makes it an ideal area for contrasting biodiversity gaps at different administrative levels. We hypothesize that the global Linnean and Wallacean shortfalls for wild bees exceed those in Europe, accentuating other data impediments and shortfalls. However, we expect variations between well-studied and less-explored regions to be evident. North American countries have a rich bee diversity and endemism (Freitas et al., 2009; Ascher and Pickering, 2020). A completeness analysis for wild bees in the US showed overall low completeness, which varied spatially and at different taxonomic levels (Chesshire et al., 2023). Yet, taxonomic challenges on bee fauna are likely less accentuated in regions with colder climates than in the more biodiverse rich subtropical regions of North America. Asia also has regions that are better studied than others and faces taxonomic challenges (Warrit et al., 2023). Similarly, in South America, knowledge is contrasted between well-researched regions, e.g., within Chile (Marshall et al., 2023; López-Aliste et al., 2021) and Brazil (Oliveira et al., 2016; Pereira et al., 2021), and areas with fewer standardized methodologies yielding notable shortfalls, often driven by a high Linnean shortfall (Freitas et al., 2009). In Africa, substantial Linnean and Wallacean shortfalls are likely (Eardley et al., 2009), although well-documented regions exist, e.g., Morocco (Lhomme et al., 2020) and southern Africa (Kuhlmann, 2009). Much African data within European institutions await digitization which could be used to calculate shortfalls (Tshibungu et al., 2023).

Overall, we see this study as an important first step towards the long-term goal to mobilize and aggregate European wild bee data into an easy access, shareable, and updatable database to be used at various spatial scales for research into the diversity, ecology and conservation of wild bees and to help inform conservation practice and supporting policies. The framework presented here is fully adaptable to other species groups and locations. We hope that this work will inspire regional gap-assessments of wild bees and other organisms worldwide.

CRediT authorship contribution statement

Leon Marshall: Conceptualization, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing. **Nicolas Leclercq:** Conceptualization, Formal analysis, Investigation, Methodology, Writing – review & editing. **Luísa G. Carvalheiro:** Data curation, Methodology, Writing – review & editing. **Holger H. Dathe:** Data curation, Writing – review & editing. **Bernhard Jacobi:** Data curation, Formal analysis, Methodology, Writing – review & editing. **Michael Kuhlmann:** Data curation, Writing – review & editing. **Simon G. Potts:** Data curation, Methodology, Writing – review & editing. **Pierre Rasmont:** Data curation, Writing – review & editing. **Stuart P.M. Roberts:** Data curation, Writing – review & editing. **Nicolas J. Vereecken:** Conceptualization, Methodology, Writing – review & editing, Supervision, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

All data and code is available at https://doi.org/10.5281/zenodo.10470867 The online shiny tool presenting the results of the study is available at: beefall.org. For code and data of updated versions of the online shiny tool please see https://github.com/lmar116/BeeFall.

Acknowledgments

LM, NL and NJV were supported by the FNRS/FWO joint program “EOS – Excellence of Science” for the project “CliPS: Climate change and its effects on Pollination Services (project 30947854)”. LM was also supported by F.R.S.-FNRS fellowships (Chargé de recherches) and the “BeeConnected” project (VI.Veni.222.141) which is financed by the Dutch Research Council (NWO). NL is additionally supported by “Fonds David et Alice Van Buuren” and “Fondation Jaumotte-Demoulin”. LGC is supported by the Brazilian National Council for Scientific and Technological Development, CNPq (307625/2021-4). SGP was supported by the Safeguard project (H2020 Grant Agreement 101003476). Wild bee collection data was obtained from: the European Commission Framework Programme (FP) 7 via the Status and Trends of European Pollinators (STEP) collaborative project (grant no. 244090, www.STEP-project.net). Organisations involved in managing and sharing this data included the Bees, Wasps and Ants Recording Society (BWARS, www.bwars.com), the Banque de Données Fauniques Gembloux-Mons (BDFGM, www.atlashymenoptera.net), the Swedish Species Information Centre (SSIC, www.artdatabanken.se), the European Invertebrate Survey—Nederland (EIS-NL, www.eis-nederland.nl), the Centre Suisse de Cartographie de la Faune (CSCF, www.cscf.ch), the National Biodiversity Data Centre (NBDC, www.biodiversityireland.ie), the Finnish Museum of Natural History (FMNH, www.luomus.fi), the Norwegian Species information Centre (NSIC, www.biodiversity.no), the Global Biodiversity Information Facility (GBIF.org, 2023, www.gbif.org). We would like to thank all those involved with the incredibly important and fundamental work of collecting, managing, curating, digitizing, correcting, and sharing bee collection data from Europe that makes this work possible. For a full list of all contributors to the full databases please see the appropriate wild bee family pages at http://www.atlashymenoptera.net/europeanbees.aspx, the GBIF link and the publications/websites of the other data sources. We would also like to thank Sara Reverté Saiz for her assistance in adapting the present study to the latest update of the European wild bee checklist. We are also grateful to all the iNaturalist users who collect and validate wild bee data, including the three most active collectors (as of Jan 2024) in Europe ‘dierkesp’, Henk Wallays (‘henkwallays2’), and Michael Knapp (‘waldgeist’) and the three most active validators (as of Jan 2024) John S. Ascher (‘johnascher’), NABU Nordvorpommern (‘nabu-nvp’), and ‘frank007’. Finally, we would like to thank everyone who uploads data to any of the publicly available sources used in this study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.biocon.2024.110455.

References

Aguirre-Gutierrez, J., et al., 2016. Functional traits help to explain half-century long shifts in pollinator distributions. Sci. Rep. 6, 13.

Albert, C.H., Thuiller, W., Yoccoz, N.G., Douzet, R., Aubert, S., Lavorel, S., 2010. A multi-trait approach reveals the structure and the relative importance of intra-vs. interspecific variability in plant traits. Funct. Ecol. 24 (6), 1192–1201.

Almeida, E.A., Bossert, S., Danforth, B.N., Porto, D.S., Freitas, F.V., Davis, C.C., Orr, M.C., 2023. The Evolutionary History of Bees in Time and Space. Current Biology.

Ascher, J.S., Pickering, J., 2020. Discover life bee species guide and world checklist (Hymenoptera: Apoidea: Anthophila). In: Draft 55. Website. Accessed Oct 2022.

Ascher, J., Marshall, L., Meiners, J., Yanega, D., Vereecken, N., 2020. Heterogeneity in large-scale databases and the role of climate change as a driver of bumble bee decline. Science (E-Letter, 2 May 2020).

Balfour, Nicholas J., Castellanos, Maria Clara, Goulson, Dave, Philippides, Andrew, Johnson, Chris, 2022. DoPI: the database of pollinator interactions. Ecology 103 (11), e3801 (Accessed November 2023).

Bartomeus, I., Stavert, J.R., Ward, D., Aguado, O., 2019. Historical collections as a tool for assessing the global pollination crisis. Philos. Trans. R. Soc. B 374, 20170389.

Bartomeus, I., et al., 2022. Base de datos de abejas ibéricas. Ecosistemas 31, 2380.

Bevis, J., 2016. The Keartons: Inventing Nature Photography. Uniformbooks.

Bossert, S., et al., 2022. Phylogeny, biogeography and diversification of the mining bee family Andrenidae. Syst. Entomol. 47, 283–302.

Callaghan, C.T., et al., 2022. The benefits of contributing to the citizen science platform iNaturalist as an identifier. PLoS Biol. 20, e3001843.

Cardoso, P., Leather, S.R., 2019. Predicting a global insect apocalypse. Insect Conserv. Divers. 12, 263–267.

Cardoso, P., Erwin, T.L., Borges, P.A., New, T.R., 2011. The seven impediments in invertebrate conservation and how to overcome them. Biol. Conserv. 144, 2647–2655.

Chang, W., et al., 2022. shiny: WebApplication Framework for R.

Chao, A., et al., 2021. Measuring temporal change in alpha diversity: A framework integrating taxonomic, phylogenetic and functional diversity and the iNEXT. 3D standardization. Methods Ecol. Evol. 12 (10), 1926–1940.

Chase, M.H., Fraterrigo, J.M., Harmon-Threatt, A., 2023. Bee functional traits and their relationship to pollination services depend on many factors: A meta-regression analysis. Insect Conserv. Divers. 16, 313–323.

Chesshire, P.R., et al., 2023. Completeness analysis for over 3000 United States bee species identifies persistent data gap. Ecography, e06584.

Costello, M.J., et al., 2014. Strategies for the sustainability of online open-access biodiversity databases. Biol. Conserv. 173, 155–165.

Dafni, A., 1992. Pollination Ecology: A Practical Approach. Oxford University Press.

Danforth, B.N., Conway, L., Ji, S., 2003. Phylogeny of eusocial Lasioglossum reveals multiple losses of eusociality within a primitively eusocial clade of bees (Hymenoptera: Halictidae). Syst. Biol. 52, 23–36.

Danforth, B.N., Cardinal, S., Praz, C., Almeida, E.A.B., Michez, D., 2013. The impact of molecular data on our understanding of bee phylogeny and evolution. Annu. Rev. Entomol. 58, 57–78.

De Palma, A., et al., 2015. Ecological traits affect the sensitivity of bees to land-use pressures in European agricultural landscapes. J. Appl. Ecol. 52, 1567–1577.

Díaz, S.M., et al., 2019. The Global Assessment Report on Biodiversity and Ecosystem Services: Summary for Policy Makers.

Diniz-Filho, J.A.F., De Marco Júnior, P., Hawkins, B.A., 2010. Defying the curse of ignorance: perspectives in insect macroecology and conservation biogeography. Insect Conserv. Divers. 3, 172–179.

Diniz-Filho, J.A.F., Loyola, R.D., Raia, P., Mooers, A.O., Bini, L.M., 2013. Darwinian shortfalls in biodiversity conservation. Trends Ecol. Evol. 28, 689–695.

Dorey, J.B., Fischer, E.E., Chesshire, P.R., et al., 2023. A globally synthesised and flagged bee occurrence dataset and cleaning workflow. Sci. Data 10 (1), 747.

Eardley, C.D., Gikungu, M., Schwarz, M.P., 2009. Bee conservation in Sub-Saharan Africa and Madagascar: diversity, status, and threats. Apidologie 40 (3), 355–366.

Elith, J., Leathwick, J.R., 2009. Species distribution models: ecological explanation and prediction across space and time. Annu. Rev. Ecol. Evol. Syst. 40, 677–697.

Faria, L.R.R., Pie, M.R., Salles, F.F., Soares, E.D.G., 2021. The Haeckelian shortfall or the tale of the missing semaphoronts. J. Zool. Syst. Evol. Res. 59 (2), 359–369.

Freitas, B.M., Imperatriz-Fonseca, V.L., Medina, L.M., Kleinert, A.D.M.P., Galetto, L., Nates-Parra, G., Quezada-Euán, J.J.G., 2009. Diversity, threats, and conservation of native bees in the Neotropics. Apidologie 40 (3), 332–346.

Gaston, K.J., 1992. Taxonomy of taxonomists. Nature 356, 281–282.

GBIF.org, 28 November 2023. GBIF Occurrence Download. https://doi.org/10.15468/dl.umg2ny.

Geldmann, J., et al., 2016. What determines spatial bias in citizen science? Exploring four recording schemes with different proficiency requirements. Divers. Distrib. 22, 1139–1149.

Ghisbain, G., Rosa, P., Bogusch, P., Flaminio, S., Le Divelec, R., Dorchin, A., Müller, A., 2023. The New Annotated Checklist of the Wild Bees of Europe (Hymenoptera: Anthophila). Zootaxa.

Gombault, C., Morison, N., Guilbaud, L., Vaissière, B.E., 2018. FlorAbeilles: Base de données en ligne sur les interactions plantes-abeilles en France métropolitaine. In: Inrae, Unité Abeilles et Environnement. France, Avignon. http://www.florabeilles.org (Accessed November 2023).

Grab, H., et al., 2019. Agriculturally dominated landscapes reduce bee phylogenetic diversity and pollination services. Science 363, 282–284.

Hedtke, S.M., Patiny, S., Danforth, B.N., 2013. The bee tree of life: a supermatrix approach to apoid phylogeny and biogeography. BMC Evol. Biol. 13 (1), 1–13.

Henríquez-Piskulich, P., Hugall, A.F., Stuart-Fox, D., 2023. A supermatrix phylogeny of the world's bees (Hymenoptera: Anthophila). bioRxiv 2023.06.16.545281. https://doi.org/10.1101/2023.06.16.545281.

Higgs, N.D., Attrill, M.J., 2015. Biases in biodiversity: wide-ranging species are discovered first in the deep sea. Front. Mar. Sci. 2, 61.

Hochkirch, A., et al., 2021. A strategy for the next decade to address data deficiency in neglected biodiversity. Conserv. Biol. 35, 502–509.

Hoiss, B., Krauss, J., Potts, S.G., Roberts, S., Steffan-Dewenter, I., 2012. Altitude acts as an environmental filter on phylogenetic composition, traits, and diversity in bee communities. Proc. R. Soc. B Biol. Sci. 279, 4447–4456.

Hortal, J., Jiménez-Valverde, A., Gómez, J.F., Lobo, J.M., Baselga, A., 2008. Historical bias in biodiversity inventories affects the observed environmental niche of the species. Oikos 117, 847–858.

Hortal, J., et al., 2015. Seven shortfalls that beset large-scale knowledge of biodiversity. Annu. Rev. Ecol. Evol. Syst. 46, 523–549.

Hughes, A.C., et al., 2021. Sampling biases shape our view of the natural world. Ecography 44, 1259–1269.

Isaac, N.J., Pocock, M.J., 2015. Bias and information in biological records. Biol. J. Linn. Soc. 115, 522–531.

IUCN, 2001. IUCN Red List Categories and Criteria: Version 3.1. (IUCN Species Survival Commission).

Johnson, K.R., Owens, I.F.P., THE GLOBAL COLLECTION GROUP, 2023. A global approach for natural history museum collections. Science 379, 1192–1194.

Kamp, J., Oppel, S., Heldbjerg, H., Nyegaard, T., Donald, P.F., 2016. Unstructured citizen science data fail to detect long-term population declines of common birds in Denmark. Divers. Distrib. 22, 1024–1035.

Kass, J.M., et al., 2022. The global distribution of known and undiscovered ant biodiversity. Sci. Adv. 8, 9908.

Kendall, L.K., Rader, R., Gagic, V., Cariveau, D.P., Albrecht, M., Baldock, K.C.R., Freitas, B.M., Hall, M., Holzschuh, A., Molina, F.P., Morten, J.M., Pereira, J.S., Portman, Z.M., Roberts, S.P.M., Rodriguez, J., Russo, L., Sutter, L., Vereecken, N.J., Bartomeus, I., 2018. Pollinator size and its consequences: robust estimates of body size in pollinating insects. Ecol. Evol. 9, 1702–1714.

Kooi, C.J. van der, Ollerton, J., 2020. The origins of flowering plants and pollinators. Science 368, 1306–1308.

Kuhlmann, M., 2009. Patterns of diversity, endemism, and distribution of bees (Insecta: Hymenoptera: Anthophila) in southern Africa. S. Afr. J. Bot. 75 (4), 726–738.

Lavorel, S., Garnier, E., 2002. Predicting changes in community composition and ecosystem functioning from plant traits: revisiting the Holy Grail. Funct. Ecol. 16 (5), 545–556.

Leclercq, N., et al., 2023. European bee diversity: taxonomic and phylogenetic patterns. J. Biogeogr. 50 (7), 1244–1256.

Lecocq, T., et al., 2013. Patterns of genetic and reproductive traits differentiation in mainland vs. Corsican populations of bumblebees. PLoS One 8, 65642.

Lhomme, P., Michez, D., Christmann, S., Scheuchl, E., El Abdouni, I., Hamroud, L., Dathe, H.H., 2020. The wild bees (Hymenoptera: Apoidea) of Morocco. Zootaxa 4892 (1), 1–159.

Litman, J.R., Danforth, B.N., Eardley, C.D., Praz, C.J., 2011. Why do leafcutter bees cut leaves? New insights into the early evolution of bees. Proc. R. Soc. B Biol. Sci. 278, 3593–3600.

Lomolino, M.V., 2004. Conservation Biogeography. Frontiers of Biogeography: New Directions in the Geography of Nature, 293, p. 293.

Lopes-Lima, M., et al., 2021. Major shortfalls impairing knowledge and conservation of freshwater molluscs. Hydrobiologia 848, 2831–2867.

López-Aliste, M., et al., 2021. Wild bees of Chile: a database on taxonomy, sociality, and ecology. Ecology 102, e03377.

Marshall, L., et al., 2015. Testing projected wild bee distributions in agricultural habitats: predictive power depends on species traits and habitat type. Ecol. Evol. 5, 4426–4436.

Marshall, L., et al., 2018. The interplay of climate and land use change affects the distribution of EU bumblebees. Glob. Chang. Biol. 24, 101–116.

Marshall, L., Ascher, J.S., Villagra, C., Beaugendre, A., Herrera, V., Henríquez-Piskulich, P., Vereecken, N.J., 2023. Chilean bee diversity: contrasting patterns of species and phylogenetic turnover along a large-scale ecological gradient. Ecosphere 14 (5), e4535.

Martinet, B., et al., 2021. Global effects of extreme temperatures on wild bumblebees. Conserv. Biol. 35, 1507–1518.

McGill, B.J., Enquist, B.J., Weiher, E., Westoby, M., 2006. Rebuilding community ecology from functional traits. Trends Ecol. Evol. 21, 178–185.

Memmott, J., 1999. The structure of a plant-pollinator food web. Ecol. Lett. 2, 276–280.

Meunier, J.-Y., et al., 2023. Apoidea of the collections of Lyon, Aix-en-Provence, Marseille and Toulon Museums of Natural History (France). Biodivers. Data J. 11, e99650.

Meyer, C., Kreft, H., Guralnick, R., Jetz, W., 2015. Global priorities for an effective information basis of biodiversity distributions. Nat. Commun. 6, 1–8.

Murray, E.A., Bossert, S., Danforth, B.N., 2018. Pollinivory and the diversification dynamics of bees. Biol. Lett. 14, 20180530.

Nelson, G., Ellis, S., 2019. The history and impact of digitization and digital data mobilization on biodiversity research. Philos. Trans. R. Soc. B 374, 20170391.

Nieto, A., et al., 2014. European Red List of Bees. Publication Office of the European Union, Luxembourg.

Noriega, J.A., et al., 2018. Research trends in ecosystem services provided by insects. Basic Appl. Ecol. 26, 8–23.

Oliveira, U., Paglia, A.P., Brescovit, A.D., de Carvalho, C.J., Silva, D.P., Rezende, D.T., Ascher, J.S., et al., 2016. The strong influence of collection bias on biodiversity knowledge shortfalls of Brazilian terrestrial biodiversity. Divers. Distrib. 22 (12), 1232–1244.

Ollerton, J., Winfree, R., Tarrant, S., 2011. How many flowering plants are pollinated by animals? Oikos 120, 321–326.

O.R.B.I.T, 2022. ORBIT project - taxonomic resources for european bees. https://orbitproject.wordpress.com/about-the-project/.

Orr, M.C., et al., 2021a. Global patterns and drivers of bee distribution. Curr. Biol. 31, 451–458.e4.

Orr, M.C., et al., 2021b. Taxonomy must engage with new technologies and evolve to face future challenges. Nat. Ecol. Evol. 5, 3–4.

Pauly, A., Devalez, J., Sonet, G., Nagy, Z.T., Boevé, J.-L., 2015. DNA barcoding and male genital morphology reveal five new cryptic species in the West Palearctic bee Seladonia smaragdula (Vachal, 1895. Halictidae). Zootaxa 4034, 257–290.

Pereira, F.W., Goncalves, R.B., Ramos, K.D.S., 2021. Bee surveys in Brazil in the last six decades: a review and scientometrics. Apidologie 52 (6), 1152–1168.

Pocock, M.J., Roy, H.E., Preston, C.D., Roy, D.B., 2015. The biological records centre: a pioneer of citizen science. Biol. J. Linn. Soc. 115, 475–493.

Poelen, J.H., Simons, J.D., Mungall, C.J., 2014. Global biotic interactions: an open infrastructure to share and analyze species-interaction datasets. Eco. Inform. 24, 148–159 (Accessed November 2023).

Poisot, T., et al., 2021. Global knowledge gaps in species interaction networks data. J. Biogeogr. 48, 1552–1563.

Potts, S.G., et al., 2011. Developing European conservation and mitigation tools for pollination services: approaches of the STEP (Status and Trends of European Pollinators) project. J. Apic. Res. 50, 152–164.

Potts, S.G., et al., 2016. The Assessment Report of the Intergovernmental Science-policy Platform on Biodiversity and Ecosystem Services on Pollinators, Pollination and Food Production.

Potts, S.G., et al., 2021. Proposal for an EU pollinator monitoring scheme, EUR 30416 EN. 23859–1. In: Publications Office of the European Union.

R Core Team, 2022. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing.

Rasmont, P., Devalez, J., Pauly, A., Michez, D., Radchenko, V.G., 2017. Addition to the checklist of IUCN European wild bees (Hymenoptera: Apoidea). In: Annales de la Société Entomologique de France (NS), vol. 53. Taylor & Francis, pp. 17–32.

Ratnasingham, S., Hebert, P.D.B., 2007. BOLD: the barcode of life data system. Mol. Ecol. Notes 7, 355–364 (Accessed November 2023).

Reverté, S., et al., 2023. National Records of 3000 European Bee and Hoverfly Species: A Contribution to Pollinator Conservation. In press, Insect Conservation and Diversity.

Riddle, B.R., Ladle, R.J., Lourie, S.A., Whittaker, R.J., 2011. Basic biogeography: estimating biodiversity and mapping nature. In: Conservation biogeography (eds. Ladle, R. J. & Whittaker, R. J.), 47–92. Blackwell Publishing.

Rocha-Ortega, M., Rodriguez, P., Córdoba-Aguilar, A., 2021. Geographical, temporal and taxonomic biases in insect GBIF data on biodiversity and extinction. Ecol. Entomol. 46, 718–728.

Salmon, R.A., Rammell, S., Emeny, M.T., Hartley, S., 2021. Citizens, scientists, and enablers: a tripartite model for citizen science projects. Diversity 13, 309.

Seltmann, K.C., Allen, J., Brown, B.V., Carper, A., Engel, M.S., Franz, N., Gilbert, E., Grinter, C., Gonzalez, V.H., Horsley, P., Lee, S., Maier, C., Miko, I., Morris, P., Oboyski, P., Pierce, N.E., Poelen, J., Scott, V.L., Smith, M., Talamas, E.J., Tsutsui, N. D., Tucker, E., 2021. Announcing Big-Bee: an initiative to promote understanding of bees through image and trait digitization. Biodivers. Inf. Sci. Stand. 5, e74037 https://doi.org/10.3897/biss.5.74037.

Seppänen, J., Väliverronen, E., 2003. Visualizing biodiversity: the role of photographs in environmental discourse. Sci. Cult. 12, 59–85.

Siefert, A., Violle, C., Chalmandrier, L., Albert, C.H., Taudiere, A., Fajardo, A., de L. Dantas, V., et al., 2015. A global meta-analysis of the relative extent of intraspecific trait variation in plant communities. Ecol. Lett. 18 (12), 1406–1419.

Spiesman, B.J., et al., 2021. Assessing the potential for deep learning and computer vision to identify bumble bee species from images. Sci. Rep. 11, 7580.

Suarez, A.V., Tsutsui, N.D., 2004. The value of museum collections for research and society. BioScience 54, 66–74.

Tshibungu, A., et al., 2023. The Megachilidae (Hymenoptera, Apoidea, Apiformes) of the Democratic Republic of Congo curated at the Royal Museum for Central Africa (RMCA, Belgium). Zootaxa 5392 (1), 1–103.

Tylianakis, J.M., Didham, R.K., Bascompte, J., Wardle, D.A., 2008. Global change and species interactions in terrestrial ecosystems. Ecol. Lett. 11, 1351–1363.

Vereecken, N.J., et al., 2021a. Five years of citizen science and standardised field surveys in an informal urban green space reveal a threatened Eden for wild bees in Brussels, Belgium. Insect Conserv. Divers. 14, 868–876.

Vereecken, N.J., et al., 2021b. Insect biomass is not a consistent proxy for biodiversity metrics in wild bees. Ecol. Indic. 121, 107132.

Vissault, S., Gravel, D., Poisot, T.E., 2019. Mangal: An Open Infrastructure for Ecological Interactions. Biodiversity Information Science and Standards.

Warrit, N., et al., 2023. Opportunities and challenges in Asian bee research and conservation. Biol. Conserv. 285, 110173.

Weekers, T., et al., 2022. Ecological, environmental, and management data indicate apple production is driven by wild bee diversity and management practices. Ecol. Indic. 139, 108880.

Wetzel, F.T., et al., 2018. Unlocking biodiversity data: prioritization and filling the gaps in biodiversity observation data in Europe. Biol. Conserv. 221, 78–85.

Williams, P.H., 2022. Novel splitting/lumping index reflects the history of species concepts applied to bumblebees (Insecta: Apidae). Zool. J. Linnean Soc. 196 (2), 704–719.

Winfree, R., Aguilar, R., Vázquez, D.P., LeBuhn, G., Aizen, M.A., 2009. A meta-analysis of bees' responses to anthropogenic disturbance. Ecology 90, 2068–2076.

Wood, T.J., Cross, I., Baldock, D.W., 2020. Updates to the bee fauna of Portugal with the description of three new Iberian Andrena species (Hymenoptera: Apoidea: Anthophila). Zootaxa 4790 (zootaxa.4790.2.1).

Yang, Z., Rannala, B., 2012. Molecular phylogenetics: principles and practice. Nat. Rev. Genet. 13, 303–314.

Zizka, A., Silvestro, D., Andermann, T., Azevedo, J., Duarte Ritter, C., Edler, D., Farooq, H., Herdean, A., Ariza, M., Scharn, R., Svantesson, S., Wengström, N., Zizka, V., Antonelli, A., 2019. CoordinateCleaner: standardized cleaning of occurrence records from biological collection databases. Methods Ecol. Evol. 10 (5), 744–751.