Characterizing habitat preferences for common wild bee species in the Iberian Peninsula

Unni Barge

Independent project • 15 credits
Swedish University of Agricultural Sciences, SLU
Faculty of Natural Resources and Agricultural Sciences
Department of Ecology
Spanish National Research Council (CSIC)
Biology and Environmental Science - Bachelor's Programme
Uppsala 2024

Characterizing habitat preferences for common wild bee species in the Iberian Peninsula

Unni Barge
https://orcid.org/0009-0008-1094-1687

Supervisors: Erik Öckinger, SLU, Department of Ecology and Ignasi Bartomeus, EBD-CSIC, Spain
Examiner: Alistair Auffret, SLU, Department of Ecology

Credits: 15 credits
Level: G2E
Course title: Independent Project in Biology
Course code: EX0894
Programme/education: Biology and Environmental Science - Bachelor's Programme
Course coordinating dept: Department of Aquatic Sciences and Assessment

Year of publication:	2024
Place of publication:	Uppsala
Cover picture:	*Andrena flavipes*, Curro Molina, EBD-CSIC
Copyright:	All featured images are used with permission from the copyright owner.
Keywords:	pollinators, habitat preference, bees, Mediterranean

Swedish University of Agricultural Sciences, SLU
Faculty of Natural Resources and Agricultural Sciences
Department of Ecology

Abstract

In the last decades, studies have shown a global decrease in pollinators. This is of great concern given the essential role that bees play in pollinating numerous crops, which is crucial for human sustenance. Habitat loss emerges as the primary threat to bees across Europe. Within this context, the Iberian Peninsula stands out in importance for pollinator research. It is a bee hotspot, comprising more than half of Europe's total bee species, but at the same time lacking basic ecological knowledge for many species. The problem is not a lack of data but rather the lack of digitization and compilation of existing datasets.

As for this, three years ago, researchers at CSIC, Spain, launched an initiative with the aim of combining available observations on bees across the peninsula and creating a large open-source database. Initial data retrieval and data cleaning were carried out in 2022 (Bartomeus et al. 2022). The subject of this bachelor's work was to tap in on this work. The objectives were to: 1) incorporate more datasets into the database using a reproducible workflow, 2) clean and process the data, 3) combine it with EU data on CORINE land cover, 4) calculate habitat preferences for a selection of wild common bee species, and 5) examine if habitat preferences could be explained by a species phenological or latitude breadth, or by its sociality.

With this bachelor's work, the database grew in size from 87,684 records to 93,803, with the incorporation of new datasets, containing bee observations from as far back as 1830 up until year 2022. Habitat preferences were calculated for the 61 most common wild bee species using null models based on the observation data.

The most preferred habitat was found to be agricultural land; arable land and pastures and dehesas, particularly, suggesting that agricultural land in the Iberian Peninsula can support biodiversity alongside crop production. *Andrena* spp. was shown to have high preferences on pastures and dehesas, highlighting this habitat importance. Natural habitats were also highly preferred, and urban environments were found to be the least preferred. Despite this, urban environments were shown to host some bee species. This wide variety of habitat preferences demonstrates the importance of maintaining high-quality landscapes with a variety of habitats to effectively support bee populations.

Correlation analyses indicated no significant relationships between habitat preferences and phenological or latitude breadth, or sociality, suggesting other traits such as floral preferences, nesting behavior, and abiotic conditions may influence habitat choice.

The large database on Iberian bees makes it possible to conduct all kinds of analyses. Future research can delve deeper into understanding the underlying traits driving habitat preferences of Iberian bee species or explore other aspects such as population trends over time.

Keywords: Anthophila, bees, database, habitat preference, Iberian Peninsula, Mediterranean pollinators

Table of contents

1. Introduction

1.1 Reasons for study

In the Anthropocene, humans alter their surroundings by depletion of natural lands and the unsustainable use of natural resources (Rockström et al. 2009; Steffen et al. 2011a). The consequent interconnected environmental stresses pose a threat to biodiversity and its direct and indirect benefits to humans; ecosystem services (Steffen et al. 2011b). One prominent example of this is pollination (Potts et al. 2010). For thousands of years, humans have been dependent on pollinators for food supply with crops including most fruits, vegetables and nut crops (Klein et al. 2007; FAO 2018), and more than 85% of the world's flowering plants, are counting on pollinators to perform their work (Ollerton et al. 2011).

In the last decades, studies have shown a global decrease in pollinators (Zattara & Aizen 2021). This is of great concern, as it jeopardizes ecosystem services provided by pollinators and their underpinning life-support mechanisms, while also diminishing their resilience (Burkle et al. 2013). However, there are many knowledge gaps regarding the status of pollinators. The European Red List of Bees released in 2014, stated that half of European bee species lacked data in order to assess their conservation status, and called for more pollinator research (IUCN 2014). The Deputy Director of the IUCN Global Species Programme declared,

> "However, our knowledge about them is incomplete as we are faced with an alarming lack of expertise and resources. […] We must urgently invest in further research in order to provide the best possible recommendations on how to reverse their decline." (Vie 2014)

Bees (Hymenoptera: Apoidea: Anthophila) are considered one of the most important pollinators (Winfree et al. 2011), and with over 20,000 known species it is a very large and diverse taxonomic group (Ascher & Pickering 2020), that requires many different conservation strategies.

The main cause of the decline of bees is thought to be the depletion and alteration of habitat (Brown & Paxton 2009; IUCN 2014), but they are also threatened by climate change, chemical exposure and invasive species, and not least the unpredictable interplay among all these threats (Potts et al. 2010). Habitat loss can reduce available nesting and feeding grounds (Potts et al. 2010), and may create fragments in the landscape, which may lead to genetic isolation and subsequent inbreeding (Lozier & Zayed 2017).

Despite habitat loss being considered the main threat, there remains a lack of knowledge about the dependence of bees on different habitats as well as basic ecological knowledge of many species (Potts et al. 2010; Winfree 2010); knowledge that is crucial in understanding the current decline in bee populations and in mitigation practices (IUCN 2014). A major obstacle has been the lack of large datasets on species-habitat associations at large spatial scales (Winfree 2010; Bartomeus et al. 2022), which means that current knowledge is often based on modeling attempts. Combining datasets from different sources into a large, rigorous database can help bridge this gap.

The need for assessments and status stands out in importance for bee populations in the Iberian Peninsula in a European context (Bartomeus et al. 2022). The Iberian Peninsula is a bee hotspot with over 1,000 known species (Bartomeus et al. 2022) which make up more than half of Europe’s total known bee species, across a peninsula that is fairly heterogenous. The few studies analyzed so far show bee declines for some species, but with sampling occurring at a few specific locations, and collected by a few specific teams (Bartomeus et al. 2018). Combining data from various sources, and thus from various locations, are therefore vital to receive a more fundamental view of the status of bees in the Iberian Peninsula (Bartomeus et al. 2018).

Over the past three years, a group of researchers in Spain have been working on combining data from museum collections, researcher sampling work, naturalists and farmers, which has generated an exceptional large dataset with bee observations on the peninsula from the 1800s to present day (Bartomeus et al. 2022). The data needs thorough cleaning but could serve as a great tool in filling the current knowledge gap in the Iberian Peninsula. It can help answer basic ecological questions about phenology and distribution, as well as understanding the dependence on different habitats (i.e. habitat preferences).

1.2 Aim and objectives

This project aims to evaluate the dependence of wild bees on different habitats (natural/agricultural/urban) in the Iberian Peninsula. In addition, the phenological breadth and latitude breadth, of a selection of species, will be investigated to determine whether these traits can explain the habitat preferences of bees.

The main objectives are to clean and filter datasets of pollinator observations in the programming language R, and to use the obtained data in calculating habitat preference and correlation analyses, using the tools of R, QGIS, the EU Corine Land Cover data, and linear models. The research questions to be answered are as follows:

RQ1 What are the habitat preferences of a selection of common wild bee species in the Iberian Peninsula?

RQ2 What are the phenological breadth and latitude breadth of common wild bee species in the Iberian Peninsula?

RQ3 Can the habitat preferences of a bee be explained by its phenological breadth, its latitude breadth, or by its sociality?

The relationship between habitat preferences and phenological breadth is expected to be positive, meaning that if a species shows a high preference for a certain habitat, it also shows a high preference for which days they fly on (i.e. narrow phenological breadth). On the contrary, if a species shows a high habitat preference, it is expected to have a low latitude breadth, as it is limited by particular habitats. Lastly, one could expect that social bees can occupy larger habitats than solitary bees and would therefore show lower habitat preferences.

1.3 Scope

This bachelor thesis is a continuation of work initiated and developed by Bartomeus et al. (2022). Some of the methods described in this study will be similar to those of Bartomeus et al. (2022).

The time frame for this project was 10 weeks. Compromises had to be made in data cleaning to allow sufficient time to perform the analyses. The dataset is very extensive, and it takes time and effort to go through all pollinator datasets. Also, due to time constraints, the different land covers are not explained in detail in this project, nor are the habitat preferences for all the different land covers.

2. Background

2.1 Bees (Anthophila)

There are more than 20,000 known bee species worldwide (Ascher & Pickering 2020), and new species continue to be found each year (see e.g. Wood et al. 2022). As such a large group, they differ a lot from one another. Some are social, meaning that they live in colonies; and some are solitary. Bees rely on flowering plants for protein, lipids, and sugars (Haydak 1970), which they receive in exchange for their service of pollination (Nicolson et al. 2007). The interplay between bees and flowers is a result of co-evolution, with correlations including emerging/blooming, flying period/flowering period, and tongue length/flower depth (Bartomeus et al. 2016).

Bees (Anthophila) is a taxonomic clade from the insect order Hymenoptera, comprising bees, wasps, sawflies, and ants. They are subdivided into six families (Andrenidae, Apidae, Collectidae, Halicitidae, Megachilidae, Melittidae). (Ascher & Pickering 2020).

Various bee species exhibit preferences for different habitats, as indicated by several studies (Strange et al. 2011; Hall et al. 2017; Collado et al. 2019; Sõber et al. 2020; Lanuza et al. 2023). These habitat preferences often vary depending on the bee's activities, such as foraging, nesting, or over-wintering (Banaszak & Twerd 2018). While some bee species favor open landscapes, others prefer more densely vegetated and sheltered environments. Additionally, each species selects nesting sites with specific characteristics; for instance, some bees exclusively nest in tree trunks or stems, while others choose ground locations (Michener 2007; Antoine & Forrest 2021). Bees may create nesting sites by digging or gnawing holes themselves or by utilizing pre-existing cavities, including shells (Hostinská et al. 2021).

The majority of bees in temperate regions have annual cycles where they go into diapause, either as a larvae or adults, before spring emergence. The opposite is perennial bees, who are active all through winter. (Michener 2007)

The most famous bee, the honey bee (*Apis mellifera*), has and continues to be of great importance to humankind with contributions to medicine, food, and, not least, pollination (Matias et al. 2017). It is also by far the most researched bee (IUCN

2014). However, the honeybee is a fairly poor indicator for other bees due to its substantially different ecological characteristics. Firstly, the honeybee is a domesticated species, and it is also one of the few perennial bee species in temperate regions. Moreover, the honey can pose a threat to wild bee populations through competition over available resources and by spreading invasive parasites or plants (Brown & Paxton 2009; Williams & Osborne 2009). Recognizing this, researchers and policymakers, such as the IUCN, emphasize the necessity of prioritizing studies on the conservation status of wild bees over honeybees, considering the latter's relatively well-established body of research (IUCN 2014).

Bees play a major role in the ecosystem, first and foremost as key pollinators for many plants (Klein et al. 2007). Wild bees also play an essential role in a social context through bee product harvesting that can be used in medicine and food (Potts et al. 2016; Klein et al. 2018), and they have an intrinsic value as part of the world biodiversity.

See Figure 1-Figure 3 for pictures of some of the bees that are included in this study (read more in chapter 3.2).

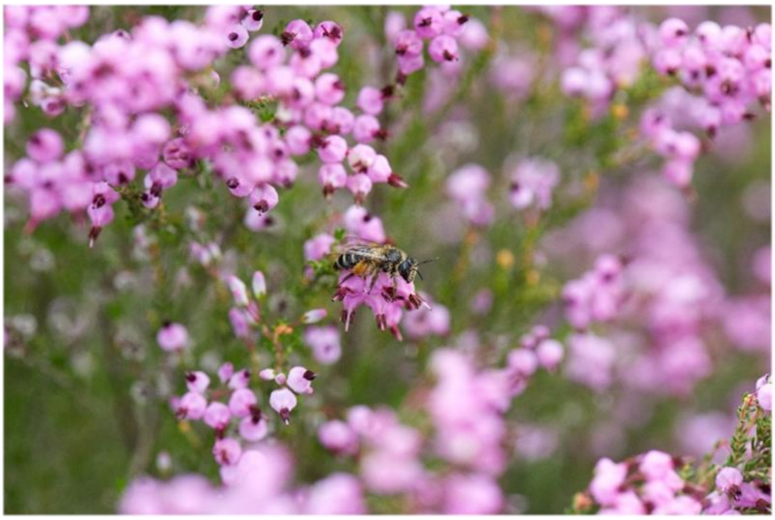

Figure 1. Andrena flavipes*, a solitary bee, on the heather* Erica umbellate. *(Curro Molina)*

Figure 2. Anthidium manicatum*, a solitary bee. (Gail Hampshire)*

Figure 3. Bombus terrestris*, a social bee, on* Halimium halimifolium*. (Curro Molina)*

2.2 Status of bees

In 2014, the IUCN released a red list of European wild bees (IUCN 2014). The assessment brought together many studies and concluded that, as of 2014, 9% of bees in Europe are threatened with extinction and that the main driver is habitat loss. The red list also listed more than half of the European bee species as ‘Data

Deficient', meaning that a status assessment of the bee couldn't be done due to a lack of data. This implies that there are species that may be threatened by extinction but are unrecorded. In fact, not many European bee extinctions have been reported (Ollerton et al. 2014). Bumblebees have the most data and also the most reported declines (Bartomeus et al. 2013).

As previously stated, habitat loss is considered the main driver of bee decline (Brown & Paxton 2009). The loss of habitat takes many forms. Urban areas are growing and may fail to maintain green areas and ecosystem services. Forests are degraded or converted to agricultural land, which can affect forest taxa (Winfree et al. 2007; Ulyshen & Horn 2023). Species-rich pastures and grasslands are among the habitats that have declined the most and have been converted to cropland or planted forest, for example (Plieninger et al. 2015). Such land conversions can lead to a loss of floral diversity, which has a direct impact on the pollinators that depend on them (Winfree et al. 2011). In this context, researchers often speak of land abandonment and agricultural degradation. The former is the process by which farmers abandon their land, for example, due to depopulation of rural areas, low productivity, and unfavorable natural conditions (Perpiña Castillo et al. 2020), and the latter occurs when the land loses quality, for example, through the use of intensive farming methods and pesticides.

Climate change, pollution, pesticides, and invasive species are likely to exacerbate the effects on habitat loss and degradation, and their impact on pollinators (Potts et al. 2010, 2016). Rising temperatures may lead to droughts, impoverished soils, a decline in the resilience of trees, pollution and invasive species that further degrade natural soils.

The importance of habitat was investigated in a study on bees in the northeastern United States (Collado et al. 2019). The study used >15,000 bee samples with known sampling coordinates and calculated the **habitat preference** for each species. A high habitat preference is defined as the occurrence of a species in a particular habitat that is more common than expected by chance. The researchers emphasize that it is not sufficient to equate the observation of a bee in a particular habitat with a preference, as it may be the only habitat available in that location, for example, therefore, null models are needed. The result showed that natural habitats had higher bee diversity than agricultural and urban areas. However, a few bee species showed preference for modified habitats, which has also been demonstrated in other studies (e.g. Hall et al. 2017; Wenzel et al. 2020).

2.3 The Iberian Peninsula and Iberian bees

The Iberian Peninsula is one of the most biodiverse regions in Europe (Araújo et al. 2007; IUCN 2014), with mountains, forests, moors and scrubland, and with large stretches of coastline exposed to the influence of the Atlantic and the Mediterranean. Almost half of the Iberian Peninsula is being occupied by agriculture (Fernández-Nogueira & Corbelle-Rico 2018). In the northern parts there are pastures, maize plantations, moors and heathland and forests, while in the southern parts there are more orchards, olive groves, vineyards, and *dehesas* (a traditional agroforestry system with pastures and scattered deciduous trees) (European Commission 2024).

The southern areas are continuously being converted into more intensive agricultural systems, for example, horticulture and olive groves (Downward & Taylor 2007), while in the northern parts, land is being abandoned (Suárez-Seoane et al. 2002).

The species richness of bees in Europe increases from north to south, with the highest species richness being seen in the Mediterranean climatic zone (IUCN 2014). There are more than 1,000 known bee species in the Iberian Peninsula, with the most commonly observed species being *Bombus terrestris* and *Bombus pascuorum* (Ascher & Pickering 2020; Bartomeus et al. 2022). Nevertheless, very little is known about the status of bees in the Iberian Peninsula (Bartomeus et al. 2022), and there is no red list of bee species in Spain in general. In Portugal, a red list for invertebrates was published in 2023; the assessment included 40 species of bee species, nine of which were found threatened (Boieiro et al. 2023). Yet, there is a rich tradition of pollinator researchers in the Iberian Peninsula, especially in Spain, and a lot of available data (Bartomeus et al. 2018).

As already mentioned, the knowledge gap in the Iberian Peninsula is not due to a lack of data, but rather to the lack of digitization of existing datasets, and the merging of different data, that can provide a bigger picture of the status of Iberian bees. There is a lot of data that can be retrieved from pinned specimens in natural collections, data from citizen sampling, and from published or unpublished work by researchers (Bartomeus et al. 2018); provided that all these data are homogenized and merged in a way that allows analysis.

2.4 Filling in the knowledge gap

In this regard, three years ago, an initiative was launched by researchers from the Spanish Research Council (CSIC), hereafter referred to as the **Iberian bees project**.

The idea was to combine data on bee observations dating back to the 19th century up to present-day records, and to merge records from all different sources, including individual researchers, historical collections, and public data, into a large database on Iberian bees, hereafter referred to as the **Iberian bees database**.

A bee observation is a record of a bee occurrence during a sampling event. Preferably, it contains not only information about the bee species observed, but also information about the date of the observation, the location of the observation, the sex, and other notable information about the observation (e.g. whether it was visiting a particular flower, altitude, or weather at the time of the observation).

The first part of the Iberian bees project began with requesting data through a **template that was sent to pollinator researchers** in Spain and Portugal. The researchers who had data should enter their data according to the template. The information requested in the template consisted of: what species of bee was observed, when and where it was observed, and any other additional information that might be relevant, see Table 1. The idea was that the researcher could fill in the template with the information available to them and send it back to the CSIC research group. The template is also available on the Iberian bees GitHub page[1] and can be viewed by anyone. See Table A1 in the Appendix for the template.

Table 1. Requested information in the template that was sent out from the Iberian bee project group to researchers, with the purpose of creating the database of Iberian bee observations. The information is per bee observation.

Mandatory information	Optional information	
Genus	Subgenus	Determined by
Species	Subspecies	Female
Country	Latitude	Male
Province	Longitude	Worker
Locality	Coordinate precision	Not specified
Year	Month	Reference doi
Authors to give credit	Day	Flowers visited
	Start date	Local ID
	End date	Any other additional data
	Collector	Notes and queries

The Iberian bees project also retrieved bee data from **natural history collections**. These collections inherit observation records in the form of pinned specimens along with a label with information about the species, and sometimes additional information such as year and locality (Bartomeus et al. 2018). Additional valuable information can be retrieved from the pinned specimens, as labels often include the

---

[1] https://github.com/ibartomeus/IberianBees.

collection location and timing. Moreover, other data can be obtained, e.g. if they still contain pollen or other traces of the collection event that can link the species to a particular flower.

Natural history collections can be an important tool to assess the historical persistence of bees (and other species) (Bartomeus et al. 2018), and to make models designed to extrapolate the distribution of bees over time more robust. A study by Bartomeus et al. (2018) showed that Spanish pollinator data were spatially biased at a few sampling sites, but using historical records and collections, some of these blind spots could be covered. The study also argues that some countries have a fairly good number of recent bee records but lack historical collections, e.g. Switzerland. Therefore, the assessment of bees in these countries should prioritize the digitization of old material before conducting data analyses (Bartomeus et al. 2018).

However, there is an obvious bottleneck in the use of records from museums: the lack of digitization (Bartomeus et al. 2018). Digitization is labor-intensive and time-consuming and often requires taxonomic identification by experts to confirm the label species (Meier & Dikow 2004). Within the Iberian bees project, 10% of the specimens deposited in the National Museum of Sciences of Madrid have been digitized and implemented (so far), together with digitization from natural collections in Leiden (Netherlands), Linz (Austria) and Washington D.C. (USA) (Bartomeus et al. 2022).

The Iberian bees project also **included data from two public platforms**: iNaturalist and the Global Biodiversity Information Facility (GBIF). Both are databases with global biodiversity datasets. The data in iNaturalist is open source, which means that it can be added not only by scientists but also by naturalists and citizens (Loarie 2022), and the data in GBIF includes data from museums and private collections, government research institutes, etc. (GBIF n.d.). The Iberian bees project downloaded all validated bee observations in the Iberian Peninsula, and added these data to the database (Bartomeus et al. 2022).

The Iberian bees database was generated in the R programming language. Figure 4 shows an overview of the Iberian bees database and its components.

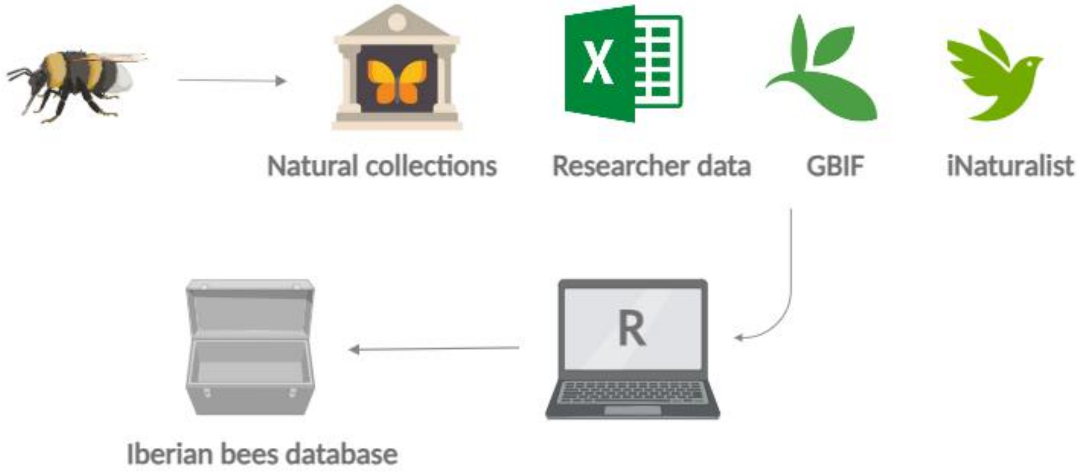

Figure 4. How the Iberian database is created and what datasets are added (read more about this in chapter 3.1). Figure created in BioRender.com.

Initial data cleaning and analysis were carried out in 2022, resulting in a journal data paper; see Bartomeus et al. (2022). At that time, the database consisted of 60 individual datasets with a total of 87 000 records collected between 1830 and 2022, and 923 unique species, with *Bombus terrestris* being the most observed one.

There are still many completed templates as well as digitized historical records that need to be incorporated into the database, which is the subject of this bachelor thesis. However, it is important to point out that the Iberian bees database is a dynamic database that is thought to be updated with each new observation or historical digitization, which means that the database in its nature is never complete.

3. Method

The Method chapter will commence by outlining the process of incorporation and cleaning of the datasets into the Iberian bees database. Following this, a section will detail the methodology used to derive habitat preferences using CORINE Land Cover data. Subsequently, another section will explain the process of calculating the phenological and latitude niches. Finally, these derived characteristics will be examined to determine potential correlations with phenology, latitude range, or sociality, aimed at gaining deeper insights into the bees' preferences.

3.1 Data cleaning

The first part of this bachelor's project was to add datasets to the Iberian database that had not been part of the cleaning that took place in year 2022. The datasets were incorporated and cleaned using R. R is a programming language most commonly used for data analysis and statistical computing, and as opposed to Excel or similar spreadsheet programs, it easily handles large datasets (Mizumoto & Plonsky 2016). With R, one can either choose to write the code in an online compiler, or in a downloaded software program. Within this bachelor's project, a software program called RStudio was used.

The added datasets were,

- **22** datasets from researchers that were retrieved from the template,
- **3** datasets from digitalized museum collections[2],

A dataset was incorporated into the database through an R script[3], meaning that each dataset had an R script linked to it. All scripts that were created within this bachelor project, can be found on GitHub[4] but an overview of the datasets, together with the number of bee records, is presented in Table A2 in the Appendix.

Even if the template with requested data sent out to researchers had a clear structure to follow, the filled-out templates had all forms of shape and diverted more or less

---

[2] The National Museum of Natural Sciences in Madrid, Oviedo Museum, Torres collection.

[3] An R script is a text file that contains the code.

[4] https://github.com/ibartomeus/IberianBees/tree/master/Scripts/1_2_Processing_raw_data, https://github.com/ibartomeus/BeeTrendsSpain/tree/main/Scripts.

from the template structure, and the vast majority of the datasets required some kind of cleaning in R. Table 2 gives examples of some of the obstacles that had to be dealt with when incorporating the datasets into the database. Even within a dataset, data could look differently, and had to require additional cleaning. Working in R, however, allowed full reproducibility, which was preferable for such large data.

Table 2. Examples of common issues and solutions/compromises with the data in the datasets. Coord. = coordinates (latitude and longitude).

Deviation from the template	Solution
Data lacks coord.	See if coord. can be retrieved from locality.
Data lacks coord. precision	See if precision can be retrieved from locality, or by number of decimals in lat/long. If number of decimals are equal to or more than 3, a precision of <100 m is assumed. If coord. precision is put as "GPS", a precision of <100 m is assumed.
Coord. faulty placed in water	Examine if coord. are written in another format than decimal degrees, and in a different coordinate system. If not, remove.
Data lacks province	See if province can be retrieved from coord.
Data lacks dates	See if sampling dates are specified in the published paper.
Dates are in weird formats	Clean and retrieve correct dates, if possible.
Data lacks species name	Remove data.
Identified species not known to live in the Iberian Peninsula	A list of these species were sent to a taxonomic expert for his judgement.
Data includes typos	Clean. Focus on typos within the mandatory information such as species name and province, and fix typos within the optional information secondly.
No information on the number of species observed, or whether it's female/male/worker	Assume one (1) specimen observed and place it in the variable Not.specified.
Data from country other than Spain and Portugal.	Remove data.

Table 3 shows an example of what a dataset could look like after cleaning. Note that it only shows some variables and rows to facilitate reading.

Table 3. Example of one dataset after cleaning. The example only includes selected variables out of a total of 28, and 5 out of 36 rows.

Species	Province	Latitude	Longitude	Date	Female/ Male/Worker
Amegilla andresi	Sevilla	37,2622	-6,2216	21-05-21	1 Male
Ammobates muticus	Sevilla	37,2622	-6,2216	21-05-21	1 Female, 1 Male
Andrena bimaculata	Sevilla	37,2622	-6,2216	21-05-21	1 Male
Andrena flavipes	Sevilla	37,2622	-6,2216	21-05-21	2 Females
Andrena fulvicornis	Sevilla	37,2622	-6,2216	21-05-21	1 Male

After the initial cleaning per dataset, all datasets were merged together to form one large dataset[5] which was followed by three different data filtering processes. In contrast to the cleaning that was done to each dataset, the purpose of the filtering processes was to prepare the data for forthcoming analyses. The second and third filtering processes will be explained later in this Method chapter. The first filtering process was based on the following criteria:

- Only observations of bee species known to the Iberian Peninsula were retained.
- *Apis mellifera* was filtered out.
- Data without coordinates was filtered out.
- Data without a specified year was filtered out.

All datasets thus far, i.e. those cleaned in 2022 and within this bachelor's project, consisted of 93,803 initial data points, and 66,578 made it through the cleaning and filtering processes, meaning that approximately 30% were discarded.

3.2 Habitat preference analysis

In order to perform the habitat preference analysis, each bee observation had to be linked to the kind of habitat that observation took place in. Habitat information was retrieved through the **CORINE Land Cover (CLC)**[6], which is a database on land coverage all across Europe, provided by the European Environmental Agency. The CLC is divided into 44 different land cover classes, including arable land, pastures, olive groves, coniferous forests, airports, and water bodies, to name a few (Kosztra et al. 2019). The first version of the inventory was launched in 1990, and updates were produced in 2000, 2006, 2012, and 2018. (European Commission 2024).

---

[5] This was done by an existing R script, developed within the initial cleaning of 2022 (Bartomeus et al. 2022). The script will not be explained in this thesis.

[6] CORINE - *Coordination of Information on the Environment*.

An additional **data filtering** (second filtering process) was carried out before the habitat preference analysis. This filtering was based on the criteria:

- Only observations from year 1985 and onwards were retained, as the first CORINE Land Cover was released in 1990 (a five-year lag was considered acceptable).
- Only observations with a known year of observation were retained.
- Species observed fewer than 100 times in total were discarded.
- Observations with a coordinate precision greater than 100 m were discarded, aligning with the resolution of the Corine Land Cover data, which is 100 m.
- Only one observation of a given species from a collection event was retained, defining a collection event as a sampling conducted on a unique day and locality (latitude and longitude). Put simply, each collection event at a specific time and place had only one observation per species in the cleaned database. This measure was implemented to ensure the independence of samples.

This filter process additionally removed almost 87% of the data. Compared to the original data of 93,803, this meant that 8,985 records made it through both the cleaning and habitat filtering processes, and a total of 61 different species.

Each observation was paired with a land cover using R code created within this project. To accomplish this, all five available CORINE Land Cover maps were downloaded and integrated into RStudio using shapefiles. The Land Cover map closest in time to the actual observation was used to extract the land cover data. For example, an observation from year 1992 retained land cover data from the CORINE Land Cover 1990, and an observation from 2017 retained land cover data from the CORINE 2018. QGIS was used to visualize the CORINE Land Cover maps.

To enhance the interpretation of the CORINE Land Cover data within the habitat analyses, land cover classes were divided into ten main categories, with three main classes, see Table 4. The division follows that of the CLC, with the following exceptions: 1) Grasslands are given their own class and are not subdivided under Scrubs; due to that grasslands in the Iberian Peninsula are distinctively different from those of scrubs; 2) Agro-forestry areas are placed under Pastures instead of Heterogeneous agricultural areas because most of the agro-forestry areas are *dehesas* or *montados*[7] which are more similar to pastures; see Figure 5 for an example of what a dehesa looks like. For a further explanation of the different classes, see Kosztra et al. (2019).

---

[7] *Dehesa* (Spanish) and *montado* (Portuguese) are Mediterranean agro-forestry systems.

Figure 5. A dehesa in Huelva, Spain. Dehesas are a cultural landscape primarily used for grazing and are characterized by scattered oak trees. (Juan Eloy Diaz)

Some classes were altogether removed, meaning that the few observations that inherited these land classes were removed from the dataset. These were: mineral extraction sites[8], coastal lagoons, estuaries, water bodies and water courses[9].

---

[8] Most likely, the sampling event has not taken place at a mineral extraction site but nearby, and the CORINE Land cover faulty assigns it a mineral extraction site due to resolution limitations; or the coordinates are faulty.

[9] Coastal lagoons and various water bodies were removed for the same reason as mineral extraction sites; most likely, the sampling event did not take place *in* actual water but in the surroundings.

Table 4. The sectioning of the CORINE Land Cover classes used for the habitat analysis.

Class 1	Class 2	Main class
Broad-leaved forest Coniferous forest Mixed forest	Forests	Natural
Moors and heathland Sclerophyllous vegetation Transitional woodland-shrub	Scrubs	Natural
Natural grasslands	Grasslands	Natural
Bare rocks Beaches, dunes, sands Burnt areas Inland marshes Intertidal flats Peat bogs Salines Salt marshes	Open spaces with little or no vegetation	Natural
Sparsely vegetated areas Pastures Agro-forestry areas	Pastures and dehesas	Natural
Non-irrigated arable land Permanently irrigated land Rice fields	Arable land	Agricultural
Annual crops associated with permanent crops Complex cultivation patterns Land principally occupied by agriculture, with significant areas of natural vegetation	Heterogeneous agricultural areas	Agricultural
Fruit trees and berry plantations Olive groves Vineyards	Orchards	Agricultural
Airports Construction sites Continuous urban fabric Dump sites Industrial or commercial units Port areas Road and rail networks and associated land	Continuous urban	Urban
Discontinuous urban fabric Green urban areas Sport and leisure facilities	Discontinuous urban	Urban

Determining a bee species' preference for a particular habitat solely based on its observations in that habitat is insufficient. Therefore, habitat preferences were calculated for each bee species using a null model approach. Initially, the true observed occurrences (i.e. the number of species in each habitat) were stored in an occurrence matrix. Subsequently, randomized matrices were generated to represent occurrences in each habitat expected by chance. This was done in R by the ‘r2dtable’ method, which maintains species richness (i.e. the total number of each species in all ten habitats) but reshuffles its values. This procedure was repeated 100 times, resulting in 100 random matrices being generated and compared to the true occurrence matrix. If a species’ true occurrence was more frequent than expected by chance, it indicated a preference for that habitat; conversely, if the species’ true occurrence was less frequent than expected by chance, it indicated avoidance of that habitat.

Based on this, a probability was calculated that a species would be found in a given habitat. The probability was calculated as the percentile of null distribution in which the observed value falls A habitat was considered preferred by a species if it had a probability greater than 0.8, avoided if less than 0.2, and no preference in-between 0.2 to 0.8, as previously used by Collado et al. (2019).

The full code for the habitat analysis can be found on the Iberian bees projects’ analysis page on GitHub[10].

3.3 Calculating phenological and latitude breadth

Similar to the preparation that was done before the habitat analysis, data filtering was also carried out for the correlation analysis (a third filtering process). This was built on the first filtering process, and based on the criteria:

- Only observations that included data for both days and months were retained.
- Only one observation of a given species from a collection event was retained, defining a collection event as a sampling conducted on a unique day and locality (latitude and longitude).

This filter process additionally removed 58% of the data retained from the first filtering process. As opposed to the filtering done for the habitat analysis (87%), the filtering for the correlation analysis didn’t exclude data before 1985, nor did it filter based on the coordinate precision or observation times. Compared to the

---

[10] https://github.com/ibartomeus/BeeTrendsSpain/tree/main/Scripts

original data of 93,803, this meant that 27,958 records made it through both the cleaning and the correlation filtering processes.

The latitude breadth (spatial distribution) of a species was obtained by simply calculating the difference between the northernmost and the southernmost observation, i.e. the difference in latitude. The phenological breadth (flight period) was calculated by the difference between the latest day of the year observation and the earliest, after excluding the 5th percentile[11]. These calculations were carried out in R.

3.4 Correlation analysis in R

To calculate whether there was a correlation between habitat preferences and phenological breadth, latitude breadth, and sociality, respectively, three linear models were used. A linear model predicts the value of a variable based on another variable, assuming normally distributed model residuals.

In order to use habitat preferences in correlation analyses, the 10 habitat preferences of a species had to be translated into a single preference. This was done by the so-called **Pielou's evenness index**, which measures how the species are evenly distributed in a community (Jost 2010). The index is defined as a number between 0 and 1: 1 represents perfect evenness, i.e. not a particular habitat is preferred, and 0 represents the largest divergence from evenness, i.e. having a strong preference for a particular habitat/s.

---

[11] Excluding the 5th percentile was done to remove potential errors in the first and last days of sampling.

4. Results

4.1 Results of the dataset incorporation and data cleaning

With this bachelor project, the Iberian bees database grew in size from 87,684 records to 93,803, due to the incorporation of new datasets. It consists of 934 different bee species (compared to 923), with the most observed genus being *Bombus* spp. The five most observed species, together with their number of records, are shown in Table 5. The first bee observation was from year 1830 up until year 2022.

The spatial distribution of the bee observations is presented in Figure 6. The observations cover large areas of Spain and Portugal, even if the observations are biased toward the coastlines and larger urban areas.

Table 5. The five most observed species in the Iberian Peninsula during years 1830-2022 and their number of records, according to the Iberian bees database.

Species	No of records
Bombus terrestris	8,147
Bombus pascuorum	3,980
Xylocopa violacea	2,316
Lasioglossum malachurum	1,604
Andrena flavipes	1,459

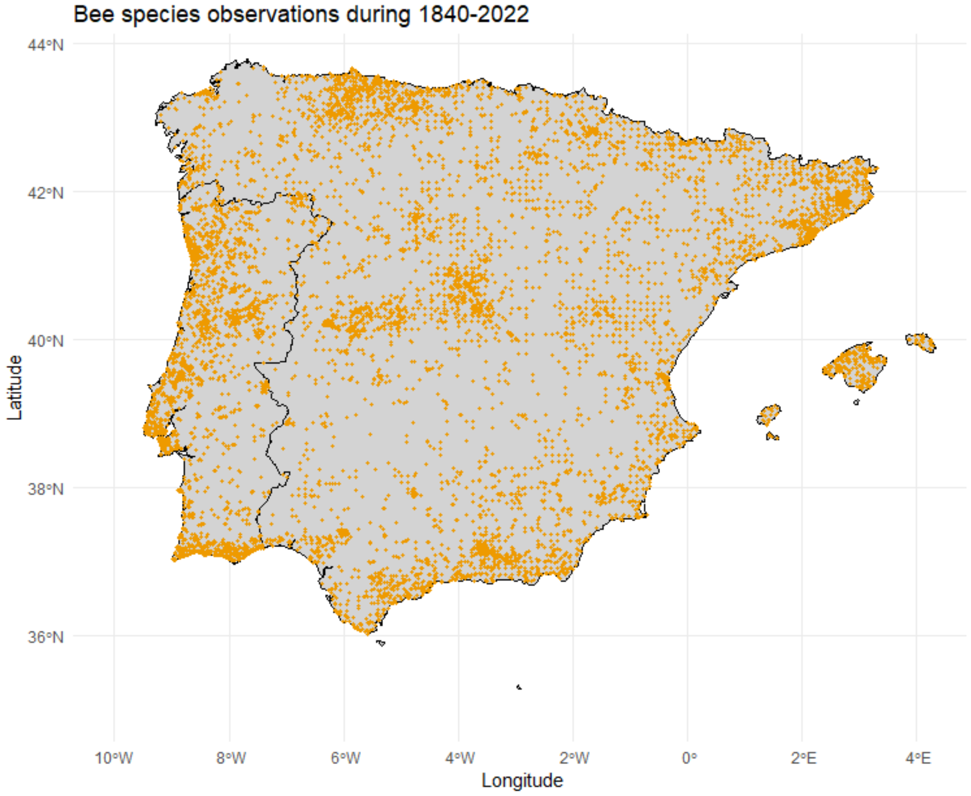

Figure 6. Bee species observations according to the Iberian bees database, after the implementation of new datasets. The observations range from 1840, and not 1830, due to that the observation of 1830 lacks coordinates.

4.2 Habitat preference analysis

The incorporation of CLC showed that bee observations mainly took place in agricultural areas (e.g. annual crop fields) and least in urban areas, see Table A3 and Table A4 in the Appendix.

The habitat preference analysis contained 61 different species after filtering. Firstly, habitat preferences were calculated for each species and for each of the main habitat classes of natural, agricultural, and urban; see results in Figure 7. 57 out of 61 species showed preferences for one or more main habitats, whereas four species lacked preferences ("habitat generalists"), i.e. had a frequency below 0.8 in all three main habitats: *Andrena agilissima*, *Andrena nigroaenea*, *Eucera elongatula*, and *Rhodanthidium septemdentatum*. Three species preferred more than one main habitat: *Bombus terrestris*, *Megachile sicula* (natural and urban), and *Panurgus calcaratus* (natural and agricultural). Agricultural land was the main habitat that was preferred by most species (33 species), and 18 species avoiding this main habitat (thus the rest 10 species out of 61 showed no preference). For natural habitat, 19 preferred this habitat, and 30 avoided it. Lastly, in urban habitats, we found 8 preferring and 45 avoiding this main habitat.

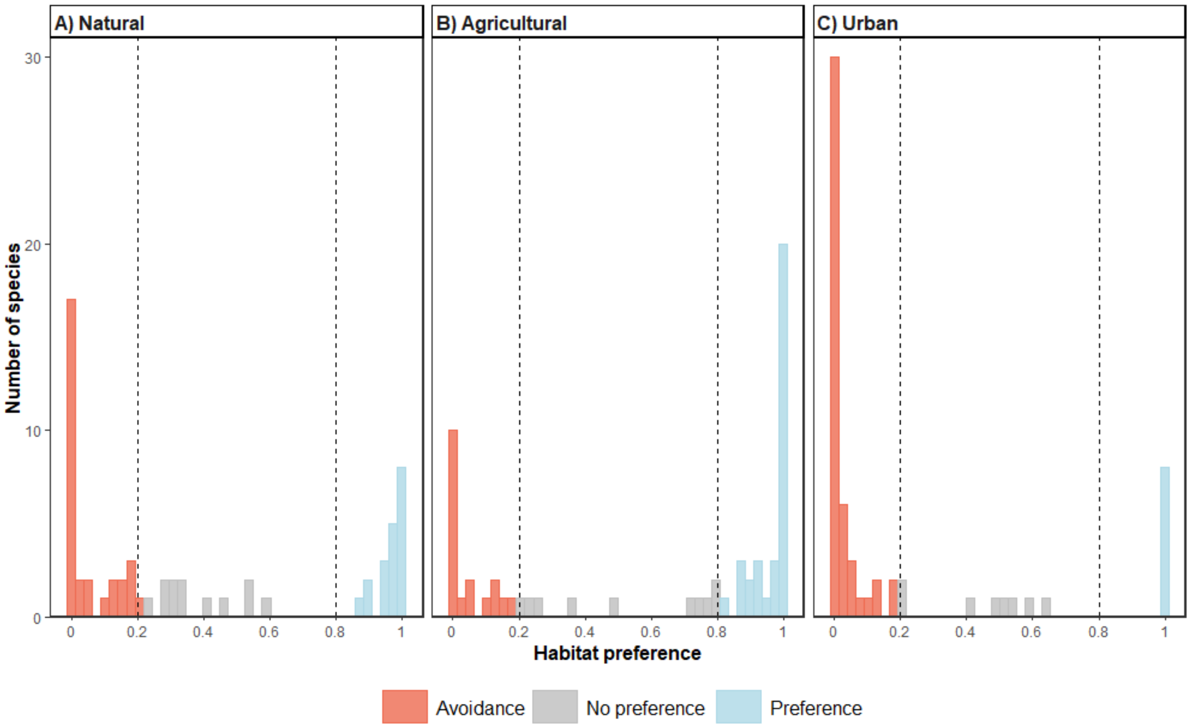

Figure 7. Habitat preference for each species in each main habitat (natural, agricultural, urban). The total number of species in the habitat analysis was 61. Red bars are the number of species avoiding that habitat class, and blue bars are the number of species preferring that habitat.

Looking more closely into the different kinds of habitat classes (based on class 2 from Table 4), one can see that 8 out of 61 species showed a preference for a single habitat, out of which 7 were for Pastures and dehesas and 1 for Arable land; see Appendix Table A5. The overall most preferred habitat was arable land (26 species preferring this), and it is also the only habitat with more species with preferences (e.g. *Lasioglossum* spp.) than avoidance (e.g. *Bombus* spp.). However, *Bombus* spp. tend to avoid agricultural habitats all together (e.g. 7 out of 8 *Bombus* spp. avoided arable lands) and instead prefer natural landscapes of primarily forests and scrubs (Table A5 in the Appendix).

The second most preferred habitat was pastures and dehesas (23 species), but it had more avoiders than Arable land (Table 6). *Andrena* spp. was shown to have a preference almost exclusively for this habitat type (Table A5). Grasslands, however, did not have as strong preferences (Table 6) and many bees that preferred this habitat were also co-dependent on e.g. Pastures and dehesas and Orchards.

The least preferred (and the most avoided) habitats were both urban, particularly by *Andrena* spp., *Bombus* spp. and *Lasioglossum* spp. Despite this, many bee species can survive in these habitats (e.g. *Andrena agilissima*) and even thrive (e.g. *Anthidium* spp. and *Xylocopa violacea*) (Table A5). Only one species prefers an urban habitat without showing a preference for another habitat: *Anthidium* *manicatum.* The rest of the bees that prefer urban habitats prefer all kinds of different habitats, with no particular specialization in any.

When examining specific habitats, it's difficult to identify consistent patterns in the similarities or differences in habitat preferences. For instance, the preference for Arable land doesn't necessarily correlate with preferences for Heterogeneous agricultural areas or Orchards. Similarly, preferences for Forests don't consistently differ from preferences for Open spaces.

Table 6. Number of species preferring and avoiding habitat classes. Presented in decreasing order of the number of preferring species. Note that a species can prefer more than one particular habitat, thus not adding up to 61 species.

Habitat	Main habitat	No of species preferring the habitat / No of species avoiding the habitat
Arable land	Agricultural	26 / 25
Pastures and dehesas	Natural	23 / 35
Orchards	Agricultural	18 / 31
Het agr areas	Agricultural	16 / 23
Open spaces	Natural	15 / 23
Scrubs	Natural	14 / 28
Forests	Natural	13 / 33
Grasslands	Natural	12 / 32
Continous urban	Urban	9 / 43
Discontinous urban	Urban	7 / 40

4.3 Phenological and latitude breadth

The individual results of the calculation of the phenological and latitude breadth will only be briefly presented here; more information will be covered in the subsequent chapter of the Correlation analyses.

The phenological breadth of the 61 species combined was 245 days. The first and last flight days, together with the flight times, are presented for each species in Table A6 in the Appendix.

Looking briefly at the latitude breadth, it can be observed that most species are found in the full study area, see Table A7 in the Appendix. A handful of species, however, are only observed in a limited range, primarily *Bombus soroeensis*, *Bombus sylvarum*, and *Lasioglossum albipes* (all in northern Spain).

4.4 Correlation analyses

The next part was to examine what the different habitat preferences could be explained by. No significant relationships were found by either phenological breadth, latitude breadth or sociality.

Firstly, the correlation with habitat preferences and phenological breadth was investigated. Recall that in this context of correlation analyses, habitat preferences were translated into an evenness index. The higher the evenness index, the less preference a species has for a particular habitat. The correlation was shown to be non-significantly negative (p-value 0.39), with very scattered data and the standard error being greater than the estimate, see Figure 8. A negative correlation suggests that a species with a higher habitat preference has a wider phenological breadth, in contrast to the hypothesis, but the very scattered data and the non-significant correlation imply that phenology is not a good indicator of habitat preferences.

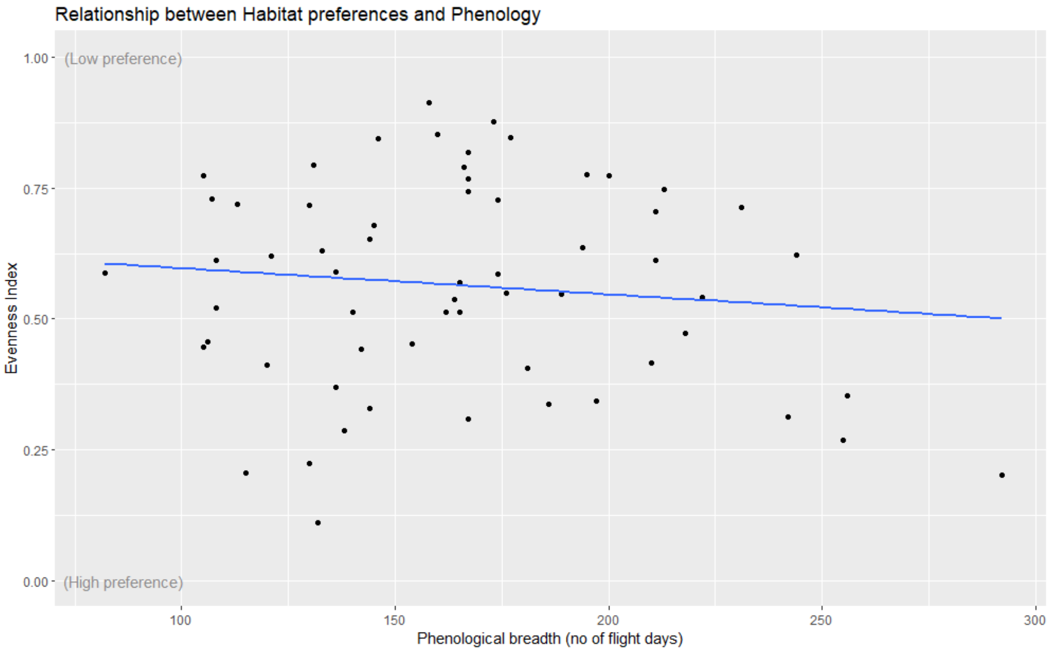

Figure 8. Correlation analysis between Evenness index (representing the Habitat preferences) and Phenology. Each black point represents one species of a total of 61 species included in the analysis. The blue line is the non-significant fitted linear regression line with the estimate -0.00049, standard error 0.00057 and p-value 0.39.

Next, habitat preferences were examined by latitude breadth. This resulted in a non-significant positive correlation (p-value 0.60), with once again the standard error being greater than the estimate, see Figure 9. A positive correlation indicates that species with higher habitat preferences have a more restricted spatial distribution (consistent with the hypothesis), but since the correlation is non-significant, it suggests that latitude breadth is not an indicator of habitat preferences. What Figure

9 tells us, though, once again, is that most of the species occur throughout the study area, as most points are concentrated in the range of 6-8 degrees latitude difference.

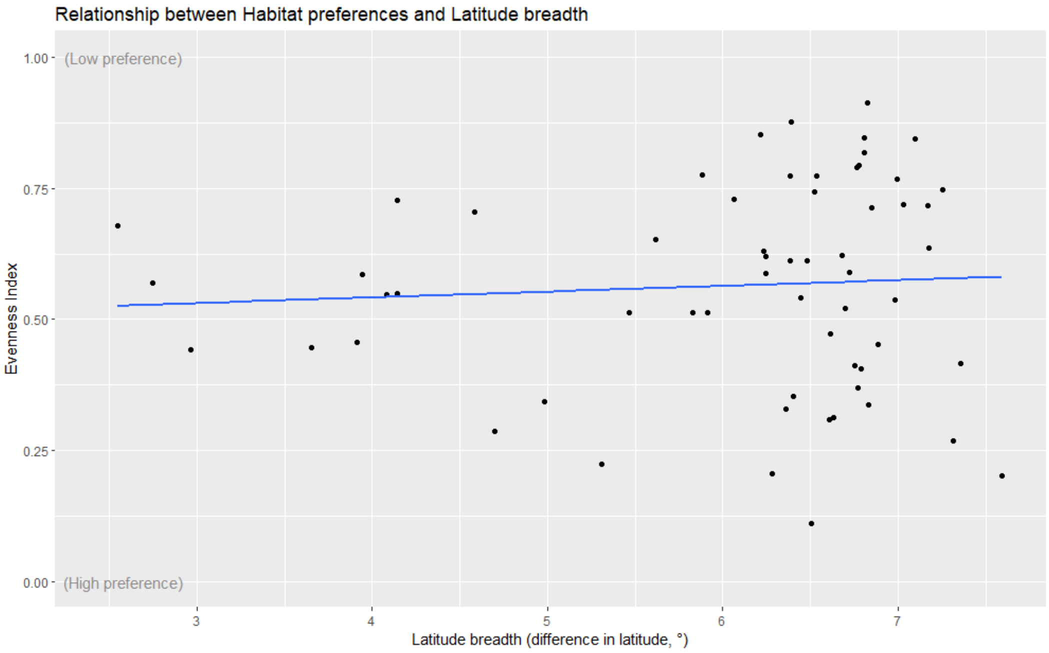

Figure 9. Correlation analysis between Evenness index (representing the Habitat preferences) and latitude breadth. Each black point represents one species of a total of 61 species included in the analysis. The blue line is the non-significant fitted linear regression line with the estimate 0.01094, standard error 0.02097 and p-value 0.60.

Lastly, habitat preferences showed a non-significant relationship with sociality (p-value 0.07), see Figure A1 in the Appendix. However, social bees show a slightly higher habitat preference than solitary bees, in contrast to the hypothesis, see boxplots in Figure 10 below.

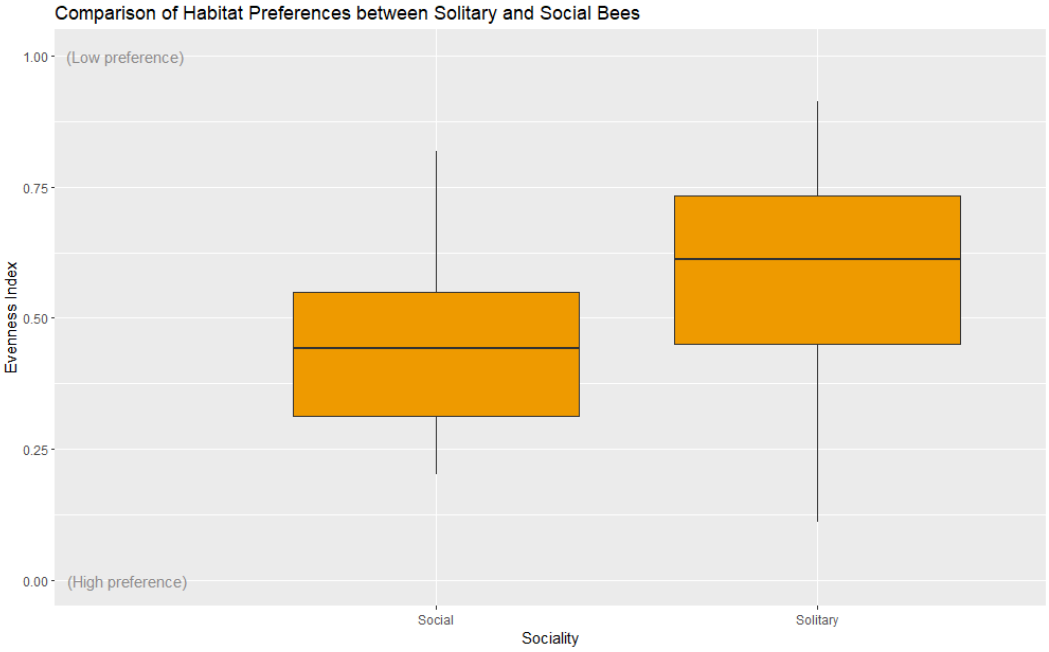

Figure 10. Comparison of Evenness index (representing the Habitat preferences) and Sociality. Due to Sociality being defined binary as either a Social or Solitary bee, its relationship (non-significant) is visualized by boxplots.

If addition, a potential correlation between phenological breadth and latitude breadth was explored (not part of the scope), see text and Figure A2 in the Appendix.

5. Discussion

5.1 Implications on habitat preferences

The most preferred main habitat was shown to be agricultural. This was somewhat surprising and also differs from the results of Collado et al. (2019) for the United States, where natural habitat was preferred by most species instead. This implies that the agricultural land in Spain and Portugal can support some biodiversity (at least bees) alongside production. Agricultural land in Spain and Portugal consists of many perennial flowering plants, such as olive trees, vineyards, and orchards, which can provide stable habitats throughout the year, in contrast to agricultural land in the northeastern United States, where predominantly annual crops (e.g. corn and soybeans) are grown that are only suitable in summers, and the majority are wind pollinated. This result highlights the importance of mitigating climate change and safeguarding traditional agricultural land in the Iberian Peninsula, which has been the target of severe droughts over the past few years (Agencia Estatal de Meteorología 2023, 2024; El Diario 2024), leading to decrease in olive oil production (Gratsea et al. 2022; EL PAÍS 2023), early almond flowering (Freitas et al. 2023), and protesting and road-blockade of farmers (El País 2024b; a).

It is not surprising that pastures and dehesas are highly preferred habitats. The high biodiversity of pastures and dehesas in the context of the Iberian Peninsula has been demonstrated in various studies (Moreno et al. 2016; Banaszak & Twerd 2018; Simonson et al. 2018; Rodríguez-Rojo et al. 2022). In recent decades, dehesas have been exposed to many threats, such as abandonment and intensification (Carmona et al. 2015; Godinho et al. 2016; Peco et al. 2017). This study shows the importance of conserving and protecting this habitat type, especially for *Andrena* spp., which almost exclusively prefers this type of habitat.

Bombus spp. prefers forests and scrubs, which is consistent with other studies (Banaszak & Twerd 2018; Sõber et al. 2020; Mola et al. 2021). Compared to open habitats, many forests and scrubs can provide food throughout the year, and they are often preferred as nesting and overwintering habitats as they offer a more protected habitat (Mola et al. 2021), which is particularly important for overwintering *Bombus* queens.

Similar to Collado et al. (2019) and not surprisingly, this study showed that urban environments are the least preferred habitats, and are especially rejected by *Bombus* spp. and *Lasioglossum* spp. Nevertheless, they may serve as an important habitat for some species (also seen in Hall et al., 2017; Collado et al., 2019; Wenzel et al., 2020), mainly *Anthidium* spp. and *Xylocopa violacea*. *Anthidium manicatum* only shows preferences for urban habitats, which is not surprising as it is the most widespread unmanaged bee in the world (Strange et al. 2011). It is native to the Iberian Peninsula but is considered invasive in other parts of the world, where it easily exploits new environments. As urban areas in the Iberian Peninsula continue to expand (United Nations 2024), more bee species may need to adapt to city environments. Therefore, it's crucial to promote bee-friendly practices in urban areas to support these populations (Banaszak-Cibicka & Żmihorski 2012).

This analysis gives a snapshot of habitat preferences between the years 1985-2022, but it does not allow a comparison of how habitat preferences may have changed over time. Agricultural degradation and land abandonment have been ongoing in Europe, particularly since the 1980s (Corbelle-Rico et al. 2012). It's possible that agricultural preferences would have been more pronounced in periods prior to the 1980s, but this study does not offer insights into such temporal shifts.

5.2 What can habitat preferences be explained by?

No significant relationships were identified between habitat preferences and phenological and latitude breadth and sociality, respectively. This suggests that bees' habitat preferences are more governed by other traits. In this study, the term "habitat" was defined as "land cover" due to the use of CORINE Land Cover data. However, it's important to recognize that a habitat encompasses more than just land cover, which may affect habitat preferences. Factors such as floral and nesting availability, climate, interactions with other species, available space, and suitability for different stages of a bee species' life cycle (such as foraging, nesting, and over-wintering) may also play significant roles, yet they were not considered in this study. Below are explorations into some of these potential traits and their potential effects on habitat preferences.

Floral preferences: studies have shown that some bees are dependent only on a few flowers, meaning that even if they are abundant all across the Iberian Peninsula (i.e. high latitude breadth), they will still only visit the places where these flowers are present (see e.g. González-Varo et al. (2016) on an example of a solitary bee's full dependence on one single flower).

Nesting behavior: bees use a variety of nests, which may impact their choice of habitat. A couple of examples from species that were in this study: *Panurgus* spp. create nests in sandy soils (Miliczky 1991) and *Rhodanthidium* spp. nest inside empty snail shells (Hostinská et al. 2021). The calculated habitat preferences (*Panurgu*s*:* grasslands, pastures and dehesas, orchards; *Rhodanthidium:* most of the agricultural habitats, urban, forests, and scrubs), however, don't showcase those kinds of details.

Abiotic conditions: factors such as soil characteristics, wind, temperature, humidity, and light, also influence bees' presence in different habitats. Soil characteristics, for example, may have an impact on the bee's choice of overwintering and nesting (Cane 1991), with influencing factors such as soil compaction and moisture (Mola et al. 2021), and pesticide residues (Anderson & Harmon-Threatt 2019). Increased wind speeds and strong summer temperatures, for example, may lead to reduced foraging efficiency, and many bees seek shelter from open habitats to more protected environments (e.g. forests) (Polatto et al. 2014; Mola et al. 2021).

5.3 Biases

It is important to note that despite the complexity of using opportunistic data, the analyses carried out in this project are based on contrasted methods for such types of data. Also, the data is thoroughly cleaned through R scripts, thus removing typos and ambiguous observations. The process of working with R makes everything reproducible and fully transparent, as opposed to, e.g. working in Excel. Nonetheless, there are some biases in the database, and in the consequent analyses that are highlighted below.

Firstly**, the individual datasets carry biases** that are being transferred to the database. All data are based on observations, all of which have biases in terms of personal interests. We also rely on the correct identification of the observed specimens, and the coordinate precisions reported by several researchers. This was accounted for by only including the 61 most observed species in the analyses (out of a total of 934) and excluding species with coordinate precision greater than 100 m. By excluding more rarely recorded species, however, we may miss some of the importance of some habitats. One can suspect that more common species (more observed species, in this context) are species that are more adaptable to a human-shaped environment, e.g. agricultural habitats, and rare species prefer more natural habitats. This was not tested in this study, but it is something that should be kept in mind.

The datasets also carry **spatial biases**, where difficult areas, such as coastlines and mountainous regions, may be under-sampled, whereas easier-access areas, such as urban regions, may be over-sampled (Figure 6). This may affect the results by exaggerating the importance of some habitats (more sampled habitats, e.g. arable land) and understating the importance of other habitats (less sampled habitats, e.g. open spaces or wetlands). In this study, however, the spatial biases are thought to be quite minor since most of the included species were found in the full study area (Table A7) and only the most observed species were included. Also note that the null models consider the sampling effort per habitat when calculating the expected occurrences, minimizing this problem.

Secondly, **processing large scripts in R may generate some errors**. For most of the datasets, a manual interpretation or inspection had to be done of the data. Many of the datasets that were sent back by researchers in the template version needed thorough data cleaning. In the case of a dataset originating from a research paper, this could mean looking at the actual paper to see what coordinate system they used when providing their latitude and longitude, or to add information on authors and collectors, or to interpret an apparent typo in dates or provinces. To account for these manual checks, wherever the interpretation or inspection didn’t give a clear answer, the observation was altogether removed; one dataset, for example, had an observation on the year “197”, and after a manual check, it still wasn’t clear if the year in question was 1977 or 1987, hence it was removed. Furthermore, all the scripts used in this bachelor project, have been peer code reviewed by Ignasi Bartomeus.

Thirdly, **the CORINE Land Cover** resolution of the downloaded land cover data was 100 m; to compensate for this, only observations with 100 m or less (i.e. 100 m or finer) were included in the habitat analysis. Shortcomings of the CORINE Land Cover data within Portugal and Spain have been studied by e.g. Caetano et al. 2005 and García-Álvarez & Camacho Olmedo 2017), where they showed that agricultural lands were underestimated in the early maps and young forests or shrubs are sometimes confused with mature forests. CORINE Land Cover data are generally recommended on a pan-European scale rather than a national scale, but can still provide a good tool for land analyses if you are aware of its flaws (Aune-Lundberg & Strand 2021). Another thing to bear in mind, is that the CORINE Land Cover has sharp lines between different habitats, but in reality, there may not be a clear transition as to where one habitat ends, and the other starts. This can lead to a potential incorrectly assigned land cover class, which may be most notable for class 1, but less for class 2 and the main class (agricultural, natural and urban) (Caetano et al. 2005). Lastly, within the incorporation of the CORINE Land Cover data, it is important to note that bees are mobile species, and its presence or avoidance in one habitat is not only a result of that particular habitat, but also the surrounding habitats. A dehesa that is surrounded by an urban region, might have a different bee appeal than another dehesa that is surrounded by other pastureland. The effect on the surrounding landscape is not considered in this habitat analysis. In the study by Collado et al. (2019), however, they showed that when considering landscape level, the results didn't differ much.

To conclude, the analyses come with some biases, but with a rigorous cleaning process that removed more than 90% of the data, the study can be seen as rather strong.

6. Conclusions

In this project, 25 datasets were incorporated into a large database of bee observations in the Iberian Peninsula and used to analyze the habitat preferences of 61 common wild bee species. The overall preferred habitats were agricultural areas, highlighting the potential of agricultural landscapes in Spain and Portugal to support biodiversity.

The diverse habitat preferences observed among individual bee species emphasize the need to consider bees as distinct entities and not as a homogeneous group. Some bees rely on low-intensity agricultural practices, while others thrive in urban environments, or forests and scrubland, demonstrating the importance of conserving diverse landscapes.

The results can also be translated into some conservation guidelines: 1) Pastures and dehesas serve as important habitats for common pollinators, highlighting the need for their protection, 2) Implementing bee-friendly practices in urban areas is important for supporting urban bee populations, 3) Recognizing the diverse habitat requirements of bees highlights the importance of maintaining high-quality landscapes with a variety of habitats to effectively support bee populations.

No correlations were found between habitat preferences and phenological and latitude breadth, and sociality, respectively, suggesting that other traits (such as floral preferences, nesting behavior, and abiotic conditions) may be more important for bee habitat choice.

The large database on Iberian bees makes it possible to carry out all kinds of analyses. Future research can delve deeper into understanding the underlying traits driving habitat preferences of Iberian bee species or explore other aspects such as population trends over time and phenological patterns.

References

Agencia Estatal de Meteorología (2023). Abril de 2023, el abril más cálido y seco desde que existen registros. Agencia Estatal de Meteorología - AEMET. Gobierno de España. https://www.aemet.es/es/noticias/2023/05/avance_abril_2023 [2024-04-01]

Agencia Estatal de Meteorología (2024). 2023 un año extremadamente cálido y muy seco. Agencia Estatal de Meteorología - AEMET. Gobierno de España. https://www.aemet.es/es/noticias/2024/01/resumen_anual_2023 [2024-04-01]

Anderson, N.L. & Harmon-Threatt, A.N. (2019). Chronic contact with realistic soil concentrations of imidacloprid affects the mass, immature development speed, and adult longevity of solitary bees. Scientific Reports, 9 (1), 3724. https://doi.org/10.1038/s41598-019-40031-9

Antoine, C.M. & Forrest, J.R.K. (2021). Nesting habitat of ground-nesting bees: a review. Ecological Entomology, 46 (2), 143–159. https://doi.org/10.1111/een.12986

Araújo, M.B., Lobo, J.M. & Moreno, J.C. (2007). The Effectiveness of Iberian Protected Areas in Conserving Terrestrial Biodiversity. Conservation Biology, 21 (6), 1423–1432. https://doi.org/10.1111/j.1523-1739.2007.00827.x

Ascher, J.S. & Pickering, J. (2020). Discover Life bee species guide and world checklist (Hymenoptera: Apoidea: Anthophila). http://www.discoverlife.org/mp/20q?guide=Apoidea_species [2024-02-20]

Aune-Lundberg, L. & Strand, G.-H. (2021). The content and accuracy of the CORINE Land Cover dataset for Norway. International Journal of Applied Earth Observation and Geoinformation, 96, 102266. https://doi.org/10.1016/j.jag.2020.102266

Banaszak, J. & Twerd, L. (2018). Importance of thermophilous habitats for protection of wild bees (Apiformes). Community Ecology, 19 (3), 239–247. https://doi.org/10.1556/168.2018.19.3.5

Banaszak-Cibicka, W. & Żmihorski, M. (2012). Wild bees along an urban gradient: winners and losers. Journal of Insect Conservation, 16 (3), 331–343. https://doi.org/10.1007/s10841-011-9419-2

Bartomeus, I., Ascher, J.S., Gibbs, J., Danforth, B.N., Wagner, D.L., Hedtke, S.M. & Winfree, R. (2013). Historical changes in northeastern US bee pollinators related to shared ecological traits. Proceedings of the National Academy of Sciences, 110 (12), 4656–4660. https://doi.org/10.1073/pnas.1218503110

Bartomeus, I., Gravel, D., Tylianakis, J.M., Aizen, M.A., Dickie, I.A. & Bernard-Verdier, M. (2016). A common framework for identifying linkage rules across different types of interactions. Poisot, T. (ed.) (Poisot, T., ed.) Functional Ecology, 30 (12), 1894–1903. https://doi.org/10.1111/1365-2435.12666

Bartomeus, I., Lanuza, J.B., Wood, T.J., Carvalheiro, L., Molina, F.P., Collado, M.Á., Aguado-Martín, L.O., Alomar, D., Álvarez-Fidalgo, M., Fidalgo, P.Á., Arista, M., Arroyo-Correa, B., Asís, J.D., Azpiazu, C., Baños-Picón, L., Beja, P., Boieiro, M., Borges, P.A.V., Bornay, G.G., Carvalho, R., Casimiro-Soriguer, R., Castro, S., Costa, J., Cross, I., Rúa, P.D. la, Pablos, L.M. de, Paz, V. de, Díaz-Calafat, J., Ferrero, V., Gaspar, H., Ghisbain, G., Gómez, J.M., Gómez-Martínez, C., González-Estévez, M.Á., Heleno, R., Herrera, J.M., Hormaza, J.I., Iriondo, J.M., Kuhlmann, M., Laiolo, P., Lara-Romero, C., Lázaro, A., López-Angulo, J., López-Núñez, F.A., Loureiro, J., Magrach, A., Martínez-López, V., Martínez-Núñez, C., Michez, D., Miñarro, M., Montero-Castaño, A., Moreira, B., Morente-López, J., Fonseca, N.N., Carbajal, A.N., Obeso, J.R., Ornosa, C., Ortiz-Sánchez, F.J., Bonilla, D.P., Patiny, S., Penado, A., Picanço, A., Ploquin, E.F., Rasmont, P., Rego, C., Rey, P.J., Ribas-Marquès, E., Roberts, S.P.M., Rodriguez, M., Rosas-Ramos, N., Sánchez, A.M., Santamaría, S., Tobajas, E., Tormos, J., Torres, F., Trillo, A., Valverde, J., Vilà, M. & Viñuela, E. (2022). Base de datos de abejas ibéricas. Ecosistemas, 31 (3), 2380–2380. https://doi.org/10.7818/ECOS.2380

Bartomeus, I., Stavert, J.R., Ward, D. & Aguado, O. (2018). Historical collections as a tool for assessing the global pollination crisis. Philosophical Transactions of the Royal Society B: Biological Sciences, 374 (1763), 20170389. https://doi.org/10.1098/rstb.2017.0389

Boieiro, M., Ceia, H., José Caramujo, M., Cardoso, P., Garcia-Pereira, P., Pires, D., Reis, J. & Rego, C. (2023). Livro Vermelho dos Invertebrados de Portugal Continental. http://lvinvertebrados.uac.pt/docs/Livro-Vermelho-dos-Invertebrados-de-Portugal-Continental.pdf

Brown, M.J.F. & Paxton, R.J. (2009). The conservation of bees: a global perspective. Apidologie, 40 (3), 410–416. https://doi.org/10.1051/apido/2009019

Burkle, L.A., Marlin, J.C. & Knight, T.M. (2013). Plant-Pollinator Interactions over 120 Years: Loss of Species, Co-Occurrence, and Function. Science, 339 (6127), 1611–1615. https://doi.org/10.1126/science.1232728

Caetano, M., Mata, F. & Freire, S. (2005). Accuracy assessment of the Portuguese CORINE Land Cover map - Global Developments in Environmental Earth Observation from Space. A. Marcal. Millpress. https://www.researchgate.net/publication/252320797_Accuracy_assessment_of_the_Portuguese_CORINE_Land_Cover_map_In_Marcal_A_ed_Global_Developments_in_Environmental_Earth_Observation_from_Space_Millpress_Rotterdam_The_Netherlands [2023-02-27]

Cane, J.H. (1991). Soils of ground-nesting bees (Hymenoptera: Apoidea): texture, moisture, cell depth and climate. Journal of the Kansas Entomological Society, 64 (4), 406–413

Carmona, C.P., Mason, N.W.H., Azcárate, F.M. & Peco, B. (2015). Inter-annual fluctuations in rainfall shift the functional structure of Mediterranean grasslands across gradients of productivity and disturbance. Bartha, S. (ed.) (Bartha, S., ed.) Journal of Vegetation Science, 26 (3), 538–551. https://doi.org/10.1111/jvs.12260

Collado, M.Á., Sol, D. & Bartomeus, I. (2019). Bees use anthropogenic habitats despite strong natural habitat preferences. Diversity and Distributions, 25 (6), 924–935. https://doi.org/10.1111/ddi.12899

Corbelle-Rico, E., Crecente-Maseda, R. & Santé-Riveira, I. (2012). Multi-scale assessment and spatial modelling of agricultural land abandonment in a European peripheral region: Galicia (Spain), 1956–2004. Land Use Policy, 29 (3), 493–501. https://doi.org/10.1016/j.landusepol.2011.08.008

Downward, S.R. & Taylor, R. (2007). An assessment of Spain’s Programa AGUA and its implications for sustainable water management in the province of Almería, southeast Spain. Journal of Environmental Management, 82 (2), 277–289. https://doi.org/10.1016/j.jenvman.2005.12.015

El Diario (2024). El mapa del calor récord de invierno en España: compara la temperatura de cada día con la media histórica de tu provincia. Sociedad. https://www.eldiario.es/sociedad/mapa-calor-record-invierno-espana-compara-temperatura-dia-media-historica-provincia_1_10865311.html [2024-04-01]

EL PAÍS (2023). El cambio climático explicado en una botella de aceite de oliva | Clima y Medio Ambiente. Clima y Medio Ambiente. https://elpais.com/clima-y-medio-ambiente/2023-12-03/el-cambio-climatico-explicado-en-una-botella-de-aceite-de-oliva.html [2024-04-01]

El País (2024a). Agricultores que no compran el discurso antiambiental: “Casi todos los problemas que tenemos en el campo son por el cambio climático.” Clima y Medio

Ambiente. https://elpais.com/clima-y-medio-ambiente/2024-03-03/agricultores-que-no-compran-el-discurso-antiambiental-casi-todos-los-problemas-que-tenemos-en-el-campo-son-por-el-cambio-climatico.html [2024-04-01]

El País (2024b). Quiénes son, qué cultivan y a cuánto venden los agricultores en España: el sector, explicado en gráficos. Economía. https://elpais.com/economia/2024-02-10/quienes-son-que-cultivan-y-a-cuanto-venden-los-agricultores-en-espana-el-sector-explicado-en-graficos.html [2024-04-01]

European Commission (2024). CORINE Land Cover. https://land.copernicus.eu/en/products/corine-land-cover/ [2024-02-20]

FAO (2018). Why bees matter - The importance of bees and other pollinators for food and agriculture. https://www.fao.org/3/i9527en/i9527en.pdf [2024-02-22]

Fernández-Nogueira, D. & Corbelle-Rico, E. (2018). Land Use Changes in Iberian Peninsula 1990–2012. Land, 7 (3), 99. https://doi.org/10.3390/land7030099

Freitas, T.R., Santos, J.A., Silva, A.P. & Fraga, H. (2023). Reviewing the Adverse Climate Change Impacts and Adaptation Measures on Almond Trees (Prunus dulcis). Agriculture, 13 (7), 1423. https://doi.org/10.3390/agriculture13071423

García-Álvarez, D. & Camacho Olmedo, M.T. (2017). Changes in the methodology used in the production of the Spanish CORINE: Uncertainty analysis of the new maps. International Journal of Applied Earth Observation and Geoinformation, 63, 55–67. https://doi.org/10.1016/j.jag.2017.07.001

GBIF (n.d.). What is GBIF? https://www.gbif.org/what-is-gbif [2024-02-21]

Godinho, S., Guiomar, N., Machado, R., Santos, P., Sá-Sousa, P., Fernandes, J.P., Neves, N. & Pinto-Correia, T. (2016). Assessment of environment, land management, and spatial variables on recent changes in montado land cover in southern Portugal. Agroforestry Systems, 90 (1), 177–192. https://doi.org/10.1007/s10457-014-9757-7

González-Varo, J.P., Ortiz-Sánchez, F.J. & Vilà, M. (2016). Total Bee Dependence on One Flower Species Despite Available Congeners of Similar Floral Shape. Borges, R.M. (ed.) (Borges, R. M., ed.) PLOS ONE, 11 (9), e0163122. https://doi.org/10.1371/journal.pone.0163122

Gratsea, M., Varotsos, K.V., López-Nevado, J., López-Feria, S. & Giannakopoulos, C. (2022). Assessing the long-term impact of climate change on olive crops and olive fly in Andalusia, Spain, through climate indices and return period analysis. Climate Services, 28, 100325. https://doi.org/10.1016/j.cliser.2022.100325

Hall, D.M., Camilo, G.R., Tonietto, R.K., Ollerton, J., Ahrné, K., Arduser, M., Ascher, J.S., Baldock, K.C.R., Fowler, R., Frankie, G., Goulson, D., Gunnarsson, B., Hanley, M.E., Jackson, J.I., Langellotto, G., Lowenstein, D., Minor, E.S., Philpott, S.M., Potts, S.G., Sirohi, M.H., Spevak, E.M., Stone, G.N. & Threlfall, C.G. (2017). The city as a refuge for insect pollinators. Conserv Biol., 31 (1), 24–29. https://doi.org/10.1111/cobi.12840

Haydak, M.H. (1970). Honey Bee Nutrition. 143–156

Hostinská, L., Kuneš, P., Hadrava, J., Bosch, J., Scaramozzino, P.L. & Bogusch, P. (2021). Comparative biology of four Rhodanthidium species (Hymenoptera, Megachilidae) that nest in snail shells. Journal of Hymenoptera Research, 85, 11–28. https://doi.org/10.3897/jhr.85.66544

IUCN (2014). European red list of bees. Publications Office. https://data.europa.eu/doi/10.2779/77003 [2024-02-17]

Jost, L. (2010). The Relation between Evenness and Diversity. Diversity, 2 (2), 207–232. https://doi.org/10.3390/d2020207

Klein, A.-M., Boreux, V., Fornoff, F., Mupepele, A.-C. & Pufal, G. (2018). Relevance of wild and managed bees for human well-being. Current Opinion in Insect Science, 26, 82–88. https://doi.org/10.1016/j.cois.2018.02.011

Klein, A.-M., Vaissière, B.E., Cane, J.H., Steffan-Dewenter, I., Cunningham, S.A., Kremen, C. & Tscharntke, T. (2007). Importance of pollinators in changing landscapes for world crops. Proceedings of the Royal Society B: Biological Sciences, 274 (1608), 303–313. https://doi.org/10.1098/rspb.2006.3721

Kosztra, B., Büttner, G., Hazeu, G. & Arnold, S. (2019). Updated CLC illustrated nomenclature guidelines. https://land.copernicus.eu/content/corine-land-cover-nomenclature-guidelines/docs/pdf/CLC2018_Nomenclature_illustrated_guide_20190510.pdf [2024-02-20]

Lanuza, J.B., Collado, M.Á., Sayol, F., Sol, D. & Bartomeus, I. (2023). Brain size predicts bees' tolerance to urban environments. Biology Letters, 19 (11), 20230296. https://doi.org/10.1098/rsbl.2023.0296

Loarie, S. (2022). What is it · iNaturalist. iNaturalist. https://www.inaturalist.org/pages/what+is+it [2024-02-21]

Lozier, J.D. & Zayed, A. (2017). Bee conservation in the age of genomics. Conservation Genetics, 18 (3), 713–729. https://doi.org/10.1007/s10592-016-0893-7

Matias, D.M.S., Leventon, J., Rau, A.-L., Borgemeister, C. & Von Wehrden, H. (2017). A review of ecosystem service benefits from wild bees across social contexts. Ambio, 46 (4), 456–467. https://doi.org/10.1007/s13280-016-0844-z

Meier, R. & Dikow, T. (2004). Significance of Specimen Databases from Taxonomic Revisions for Estimating and Mapping the Global Species Diversity of Invertebrates and Repatriating Reliable Specimen Data. Conservation Biology, 18 (2), 478–488. https://doi.org/10.1111/j.1523-1739.2004.00233.x

Michener, C.D. (2007). The bees of the world. 2nd ed. Johns Hopkins university press.

Miliczky, E. (1991). Observations on the Nesting Biology of Three Species of Panurgine Bees (Hymenoptera: Andrenidae). Journal of the Kansas Entomological Society, 64 (1), 80–87

Mizumoto, A. & Plonsky, L. (2016). R as a Lingua Franca: Advantages of Using R for Quantitative Research in Applied Linguistics. Applied Linguistics, 37 (2), 284–291. https://doi.org/10.1093/applin/amv025

Mola, J.M., Hemberger, J., Kochanski, J., Richardson, L.L. & Pearse, I.S. (2021). The Importance of Forests in Bumble Bee Biology and Conservation. BioScience, 71 (12), 1234–1248. https://doi.org/10.1093/biosci/biab121

Moreno, G., Gonzalez-Bornay, G., Pulido, F., Lopez-Diaz, M.L., Bertomeu, M., Juárez, E. & Diaz, M. (2016). Exploring the causes of high biodiversity of Iberian dehesas: the importance of wood pastures and marginal habitats. Agroforestry Systems, 90 (1), 87–105. https://doi.org/10.1007/s10457-015-9817-7

Nicolson, S.W., Nepi, M. & Pacini, E. (eds) (2007). Nectaries and nectar. Springer.

Ollerton, J., Erenler, H., Edwards, M. & Crockett, R. (2014). Extinctions of aculeate pollinators in Britain and the role of large-scale agricultural changes. Science, 346 (6215), 1360–1362. https://doi.org/10.1126/science.1257259

Ollerton, J., Winfree, R. & Tarrant, S. (2011). How many flowering plants are pollinated by animals? Oikos, 120 (3), 321–326. https://doi.org/10.1111/j.1600-0706.2010.18644.x

Peco, B., Navarro, E., Carmona, C.P., Medina, N.G. & Marques, M.J. (2017). Effects of grazing abandonment on soil multifunctionality: The role of plant functional traits. Agriculture, Ecosystems & Environment, 249, 215–225. https://doi.org/10.1016/j.agee.2017.08.013

Perpiña Castillo, C., Coll Aliaga, E., Lavalle, C. & Martínez Llario, J.C. (2020). An Assessment and Spatial Modelling of Agricultural Land Abandonment in Spain (2015–2030). Sustainability, 12 (2), 560. https://doi.org/10.3390/su12020560

Plieninger, T., Hartel, T., Martín-López, B., Beaufoy, G., Bergmeier, E., Kirby, K., Montero, M.J., Moreno, G., Oteros-Rozas, E. & Van Uytvanck, J. (2015). Wood-pastures of Europe: Geographic coverage, social–ecological values, conservation management, and policy implications. Biological Conservation, 190, 70–79. https://doi.org/10.1016/j.biocon.2015.05.014

Polatto, L.P., Chaud-Netto, J. & Alves-Junior, V.V. (2014). Influence of Abiotic Factors and Floral Resource Availability on Daily Foraging Activity of Bees: Influence of Abiotic and Biotic Factors on Bees. Journal of Insect Behavior, 27 (5), 593–612. https://doi.org/10.1007/s10905-014-9452-6

Potts, S.G., Biesmeijer, J.C., Kremen, C., Neumann, P., Schweiger, O. & Kunin, W.E. (2010). Global pollinator declines: trends, impacts and drivers. Trends in Ecology & Evolution, 25 (6), 345–353. https://doi.org/10.1016/j.tree.2010.01.007

Potts, S.G., Imperatriz-Fonseca, V., Ngo, H.T., Aizen, M.A., Biesmeijer, J.C., Breeze, T.D., Dicks, L.V., Garibaldi, L.A., Hill, R., Settele, J. & Vanbergen, A.J. (2016). Safeguarding pollinators and their values to human well-being. Nature, 540 (7632), 220–229. https://doi.org/10.1038/nature20588

Rockström, J., Steffen, W., Noone, K., Persson, Å., Chapin, F.S., Lambin, E.F., Lenton, T.M., Scheffer, M., Folke, C., Schellnhuber, H.J., Nykvist, B., De Wit, C.A., Hughes, T., Van Der Leeuw, S., Rodhe, H., Sörlin, S., Snyder, P.K., Costanza, R., Svedin, U., Falkenmark, M., Karlberg, L., Corell, R.W., Fabry, V.J., Hansen, J., Walker, B., Liverman, D., Richardson, K., Crutzen, P. & Foley, J.A. (2009). A safe operating space for humanity. Nature, 461 (7263), 472–475. https://doi.org/10.1038/461472a

Rodríguez-Rojo, M., Roig, S., López-Carrasco, C., Redondo García, M. & Sánchez-Mata, D. (2022). Which Factors Favour Biodiversity in Iberian Dehesas? Sustainability, 14 (4), 2345. https://doi.org/10.3390/su14042345

Simonson, W.D., Allen, H.D., Parham, E., De Basto E Santos, E. & Hotham, P. (2018). Modelling biodiversity trends in the montado (wood pasture) landscapes of the Alentejo, Portugal. Landscape Ecology, 33 (5), 811–827. https://doi.org/10.1007/s10980-018-0627-y

Sõber, V., Leps, M., Kaasik, A., Mänd, M. & Teder, T. (2020). Forest proximity supports bumblebee species richness and abundance in hemi-boreal agricultural landscape. Agriculture, Ecosystems & Environment, 298, 106961. https://doi.org/10.1016/j.agee.2020.106961

Southwick, E.E. & Southwick, L. (1992). Estimating the Economic Value of Honey Bees (Hymenoptera: Apidae) as Agricultural Pollinators in the United States. Journal of Economic Entomology, 85 (3), 621–633. https://doi.org/10.1093/jee/85.3.621

Steffen, W., Grinevald, J., Crutzen, P. & McNeill, J. (2011a). The Anthropocene: conceptual and historical perspectives. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 369 (1938), 842–867. https://doi.org/10.1098/rsta.2010.0327

Steffen, W., Persson, Å., Deutsch, L., Zalasiewicz, J., Williams, M., Richardson, K., Crumley, C., Crutzen, P., Folke, C., Gordon, L., Molina, M., Ramanathan, V., Rockström, J., Scheffer, M., Schellnhuber, H.J. & Svedin, U. (2011b). The Anthropocene: From Global Change to Planetary Stewardship. AMBIO, 40 (7), 739–761. https://doi.org/10.1007/s13280-011-0185-x

Strange, J.P., Koch, J.B., Gonzalez, V.H., Nemelka, L. & Griswold, T. (2011). Global invasion by Anthidium manicatum (Linnaeus) (Hymenoptera: Megachilidae): assessing potential distribution in North America and beyond. Biological Invasions, 13 (9), 2115–2133. https://doi.org/10.1007/s10530-011-0030-y

Suárez-Seoane, S., Osborne, P.E. & Baudry, J. (2002). Responses of birds of different biogeographic origins and habitat requirements to agricultural land abandonment in northern Spain. Biological Conservation, 105 (3), 333–344. https://doi.org/10.1016/S0006-3207(01)00213-0

Ulyshen, M. & Horn, S. (2023). Declines of bees and butterflies over 15 years in a forested landscape. Current Biology, 33 (7). https://doi.org/10.1016/j.cub.2023.02.030

United Nations (2024). Urbanization in Spain: Building inclusive & sustainable cities. UN-HABITAT. https://unhabitat.org/spain [2024-03-25]

Vie, J.-C. (2014). The IUCN Red List of Threatened Species (quote by Jean-Christophe, former Deputy Director, IUCN Global Species Programme). IUCN Red List of Threatened Species. https://www.iucnredlist.org/es [2024-02-22]

Wenzel, A., Grass, I., Belavadi, V.V. & Tscharntke, T. (2020). How urbanization is driving pollinator diversity and pollination – A systematic review. Biological Conservation, 241 (108321). https://doi.org/10.1016/j.biocon.2019.108321

Williams, P.H. & Osborne, J.L. (2009). Bumblebee vulnerability and conservation world-wide. Apidologie, 40 (3), 367–387. https://doi.org/10.1051/apido/2009025

Winfree, R. (2010). The conservation and restoration of wild bees. Annals of the New York Academy of Sciences, 1195 (1), 169–197. https://doi.org/10.1111/j.1749-6632.2010.05449.x

Winfree, R., Bartomeus, I. & Cariveau, D.P. (2011). Native Pollinators in Anthropogenic Habitats. Annual Review of Ecology, Evolution, and Systematics, 42 (1), 1–22. https://doi.org/10.1146/annurev-ecolsys-102710-145042

Winfree, R., Griswold, T. & Kremen, C. (2007). Effect of Human Disturbance on Bee Communities in a Forested Ecosystem. Conservation Biology, 21 (1), 213–223. https://doi.org/10.1111/j.1523-1739.2006.00574.x

Wood, T.J., Molina, F.P. & Bartomeus, I. (2022). A new Andrena species (Hymenoptera: Andrenidae) from the overlooked Doñana Protected Areas of southern Spain. Belgian Journal of Entomology, 126, 1–13

Zattara, E.E. & Aizen, M.A. (2021). Worldwide occurrence records suggest a global decline in bee species richness. One Earth, 4 (1), 114–123. https://doi.org/10.1016/j.oneear.2020.12.005

Acknowledgements

I would like to thank my supervisors, Ignasi Bartomeus and Erik Öckinger, for their support and guidance throughout this study. I am incredibly grateful for the opportunity to pursue the work in Sevilla and to collaborate with Nacho and his team. My journey through the changing seasons, from winter to spring to summer (albeit the Swedish version), in this city has been nothing short of wonderful.

Appendix

Table A1. Template sent out from the research group within CSIC, to expand the Iberian bees database. Any researcher can fill in the template with his/her pollinator data and send it back to the research group within CSIC.

In GREEN mandatory fields, In YELLOW optional fields of interest

Genus	Generic epithet
Subgenus	Subgenus if relevant
Species	Specific epithet. **Just specimens with accepted names would be considered**
Subspecies	Taxonomic subspecies if relevant
Country	Countries of the Iberian Península: Spain, Portugal, Andorra
Province	See standar list of Provinces, for Spain: http://www.ine.es/daco/daco42/codmun/cod_provincia_estandar.htm and for Portugal: https://en.wikipedia.org/wiki/Provinces_of_Portugal#Provinces
Locality	Locality description
Latitude	Decimal degrees
Longitude	Decimal degrees
Coordinate.precision	Indicate coordinate precision if relevant
Year	Year of capture or observation of the specimen
Month	Month of capture or observation of the specimen
Day	Day of capture or observation of the specimen
Start.date	In case that the specific date is unknown but an interval of dates can be provided, please specify here the start date of the interval
End.date	In case that the specific date is unknown but an interval of dates can be provided, please specify here the end date of the interval
Collector	Collector name
Determined.by	Person who id the species
Female	Number of females captured
Male	Number of males captured
Worker	Number of bee workers (just for eusocial bees)

Not.specified	Number of individuals captured without sex information
Reference.doi	Provide DOI if relevant (e.g. https://doi.org/xxxxxx)
Flowers.visited	Scientific name of the plant species where the specimen was capture or observed
Local_ID	Provide local id of the collection (if there is one)
Authors.to.give.credit	Author list contribution, as many as needed
Any.other.additional.data	Any other relevant information about the specimen
Notes.and.queries	Add a note or comment for the person that would check this data if necessary (e.g. about the species or coordinate precision)

Thanks a lot for your contribution!

Example with mandatory fields in green and optional fields of interest in yellow

Genus	Subgenus	Species	Subspecies	Country	Province	Locality	Latitude	Longitude	Coordinate.precision	Year	Month	Day	Start.date	End.date	Collector	Determzined.by	Female	Male	Worker	Not.specified	Reference.doi	Flowers.visited	Local_ID	Authors.to.give.credit	Any.other.additional.data	Notes.and.queries
Andrena	NA	ramosa	NA	Spain	Huelva	Parque Nacional de Doñana	37,136	-6,359	100 m	2021	4	1	NA	NA	Francisco P. Molina	Thomas J. Woods	0	1	0	NA	NA	Asphodelus ramosus	Azn_7	Francisco P. Molina, Thomas J. Woods and Ignasi Bartomeus	NA	New species for science!

Table A2. R scripts created in this bachelor project, and how many bee observations (data) each dataset consisted of before cleaning. The number in the beginning of the R script name is the chronological order of dataset added.

R script	Data (no of bee observations) before cleaning
59_Antonini_etal	7
60_Lopez-Angulo_etal	320
61_Wood	36
62_Pareja-Bonilla	5
63_Martinez-Lopez	81
64_Zafra	10
65_Alvarez	543
66_Alvarez_etal	770
67_life_polinizadores	251
68_Bornay	1,266
69_Alvarez	5
70_Rodrigo_etal	164
71_Rodrigo_etal	115
72_Dardon_etal	757
73_Ornosa	252
74_Ornosa	52
75_Ornosa	48
76_Ortiz	28
77_Ortiz	2
78_Ortiz	26
79_Oviedo_museum	5,385
80_Torres	12,320
81_MNCN	2,299
82_Aguado_asensio	1,116
83_Asensio	3,905

Table A3. Number of observations in each of the main habitat classes, that was included in the habitat preferences analysis. Presented in descending order.

Main habitat class	Number of observations
Agricultural	4.439
Natural	3.306
Urban	1.240

Table A4. Number of observations in each of the habitat classes, that was included in the habitat preferences analysis. Presented in descending order.

Habitat class	Number of observations
Heterogeneous agricultural areas	1.712
Forests	1.591
Pastures and dehesas	1.146
Arable land	1.136
Scrub and/or herbaceous vegetation associations	1.033
Discontinuous urban	731
Continuous urban	509
Orchards	445
Natural grasslands	345
Open spaces with little or no vegetation	337

Table A5. Habitat preferences for the Iberian bees, calculated by comparing the observed abundance to the calculated null model matrices. Blue indicates that the species shows a preference for that particular habitat (values over 0,8). Red indicates that the species shows an avoidance for that particular habitat (values under 0,2). Values in between are not highlighted (no preference).

Species	Arable land	Continuous urban	Discontinuous urban	Forests	Heterogeneous agricultural areas	Natural grasslands	Open spaces with little or no vegetation	Pastures and dehesas	Orchards	Scrub
Amegilla quadrifasciata	1	0,81	0,21	0	0,95	0,02	0,33	0	0,8	0,93
Andrena agilissima	0,91	0,26	0,71	0	0,64	0	0,83	0,09	0,82	0,65
Andrena dorsata	0	0	0,73	0,26	0,01	0,08	0,06	1	0,01	0,41
Andrena flavipes	0,78	0,01	0,24	0,7	0,14	0,85	0,04	1	0,32	0,02
Andrena humilis	0,02	0	0,02	0,01	0,29	0,25	0,08	1	0	0,11
Andrena minutula	0,25	0	0	0,44	0,21	0	0,66	1	0,01	0
Andrena nigroaenea	0,18	0,03	0,09	0,88	0,61	0,1	0,9	0,89	0,6	0,01
Andrena ovatula	1	0	0,03	0	0,05	1	0,74	0,75	0,12	0
Andrena pilipes	0,98	0	0,42	0,01	0,64	0,34	0	0,98	0,26	0
Anthidiellum strigatum	0,75	0	0	1	0,83	0	0	0	0	0,52
Anthidium florentinum	0,98	1	0,98	0	0,62	0,33	0,01	0	0,99	0
Anthidium manicatum	0,44	1	0,99	0,44	0,25	0	0,28	0	0,55	0,07
Anthophora atroalba	1	0,23	0	0,09	0,89	0,49	0,16	0	0,08	0
Anthophora bimaculata	0,91	0,03	0,18	0,62	0,03	0,35	1	0,07	0	0,54
Anthophora plumipes	0,12	0,89	1	0,49	0,99	0	0,28	0	0,62	0,15
Bombus hortorum	0,01	0	0,73	0,03	0,79	0,13	0,98	1	0,01	0,37
Bombus lapidarius	0,04	0,02	0	0,9	0	0,59	1	0	0,01	1
Bombus lucorum	0,05	0,8	0,01	1	0	0,96	1	0,05	0	0,7
Bombus pascuorum	0	0	0,02	1	1	0,13	0,55	0	0	0,95
Bombus pratorum	0	0,01	0,11	1	0,46	0,64	0,1	0,91	0	1
Bombus ruderatus	0,34	0,45	0,02	0,82	0,67	0,73	0,03	0	0,19	1
Bombus soroeensis	0	0	0	0,25	0,13	0,95	1	0,81	0	0,79
Bombus sylvarum	0,11	0	0	0,06	0,86	0,69	0,33	1	0	0,84
Bombus terrestris	0	0,13	1	1	0,87	1	0,02	0	0	1
Ceratina cucurbitina	1	0,12	0	0,94	0,03	0,64	0,96	0,54	0,2	0,2
Ceratina cyanea	0,23	0,12	0	0,2	0,06	0,84	0,68	0,85	0	0,99
Colletes nigricans	0,81	0,01	0	0,29	0,37	0	0,72	0	1	1
Eucera elongatula	0,13	0	0,28	0,64	0,77	0,18	0,97	0,02	1	0,04
Eucera notata	0,06	0	0,5	0,1	0,1	0	1	0	0,92	0
Halictus fulvipes	1	0,8	0	0,88	0	0	0,69	0,01	0,78	0
Halictus gemmeus	0,83	0,04	0,05	0,57	0,11	0	0,98	0,07	0,3	0,94
Halictus scabiosae	0,86	0,71	0,97	0,29	0,94	0,16	0,05	0	0	0,77
Halictus subauratus	0,66	0	0,29	0,69	0,28	0,41	0,14	1	0	0,09
Hoplitis adunca	0,24	0,14	0,13	0,09	0,5	0,64	0	0	1	0,95
Hylaeus clypearis	1	0	0,05	0,81	0,23	0	0,4	0,01	0,05	0,04
Hylaeus variegatus	1	0,01	0,21	0,01	0,13	0,33	0,3	0,14	0,03	0,06
Lasioglossum albipes	0,06	0,07	0	0,04	0,52	0,99	0,36	1	0	0,43
Lasioglossum albocinctum	1	0	0	0,32	0,99	0	0,32	0,03	0,53	0
Lasioglossum calceatum	0	0,01	0	0,12	0,23	0,07	0	1	0,02	0,46

Species	Arable land	Continuous urban	Discontinuous urban	Forests	Heterogeneous agricultural areas	Natural grasslands	Open spaces with little or no vegetation	Pastures and dehesas	Orchards	Scrub
Lasioglossum interruptum	0,87	0	0,04	0,18	0,91	0,21	0,22	0,22	0,95	0,22
Lasioglossum leucozonium	0,81	0,08	0	0,01	0,07	0,93	0,21	1	0,25	0,42
Lasioglossum malachurum	0,89	0	0	0,04	0	0,19	0,02	1	0,89	0,01
Lasioglossum pauperatum	0,05	0,59	0,28	0	0,55	0,73	0	1	0,63	0,14
Lasioglossum pauxillum	0,93	0	0	0	0,07	0	0	1	0	0
Lasioglossum punctatissimum	1	0	0	0,02	0,15	0	0	0,95	0,7	0,66
Lasioglossum villosulum	0,58	0	0	0,54	0,07	0,06	0,22	1	0	0,53
Megachile apicalis	1	0,28	0	0,06	0	0	0,24	0	1	0,59
Megachile pilidens	0,92	0,03	0,3	0,14	0,01	0	0,93	0,02	0,97	1
Megachile sicula	0	0,03	1	0,77	0,51	0	0,99	0	0	0,4
Osmia caerulescens	1	0	0,1	0	0,77	0	0,23	0	1	0,07
Osmia cornuta	0,97	1	0,15	0	0,93	0,59	0,28	0	0,12	0,04
Osmia niveata	0,52	0,15	0,77	0,03	0,84	0	0,18	0,12	1	0,04
Osmia submicans	0,02	0,02	0,21	0,09	0,83	0	0,04	0	1	0,03
Panurgus banksianus	1	0	0,05	0	0,37	1	0,81	0	0,05	0,29
Panurgus calcaratus	0	0	0	0,1	0	1	0,33	1	1	0,12
Panurgus cephalotes	0,05	0	0,04	0,15	0	0,8	0,19	1	1	0,38
Panurgus perezi	0	0,09	0	0	0,01	1	0	1	1	0,7
Rhodanthidium septemdentatum	0,97	0,66	0,14	1	0,24	0	0,4	0,04	0,09	0,01
Rhodanthidium sticticum	0	0,92	0,03	0	0,98	0,68	1	0	0,99	1
Xylocopa cantabrita	0	0,71	0,03	1	0,88	0,12	0,22	0	0,26	0,99
Xylocopa violacea	0,17	1	1	0	1	0	0	0	0,19	0

Table A6. Phenological breadth as number of flight days on average for the 61 species. Number of flight days are expressed by the difference between the first day of observation and the last day of observation. Flight days are expressed as day number (e.g. flight day 60 is 1st of March).

Species	First observed day (~first flight day)	Last observed day (~last flight day)	Phenological breadth
Amegilla quadrifasciata	60	330	270
Andrena agilissima	74	213	139
Andrena dorsata	74	222	148
Andrena flavipes	32	365	333
Andrena humilis	60	208	148
Andrena minutula	32	254	222
Andrena nigroaenea	32	195	163
Andrena ovatula	60	232	172
Andrena pilipes	66	250	184
Anthidiellum strigatum	98	298	200
Anthidium florentinum	37	326	289
Anthidium manicatum	58	365	307
Anthophora atroalba	46	339	293
Anthophora bimaculata	32	333	301
Anthophora plumipes	32	350	318
Bombus hortorum	37	340	303
Bombus lapidarius	91	280	189
Bombus lucorum	40	359	319
Bombus pascuorum	32	360	328
Bombus pratorum	32	319	287
Bombus ruderatus	32	328	296
Bombus soroeensis	87	238	151
Bombus sylvarum	56	280	224
Bombus terrestris	32	365	333
Ceratina cucurbitina	32	350	318
Ceratina cyanea	32	324	292
Colletes nigricans	32	305	273
Eucera elongatula	50	237	187
Eucera notata	36	186	150
Halictus fulvipes	60	365	305
Halictus gemmeus	60	365	305
Halictus scabiosae	51	314	263
Halictus subauratus	34	275	241
Hoplitis adunca	60	282	222
Hylaeus clypearis	101	295	194
Hylaeus variegatus	102	285	183

Species	First observed day (~first flight day)	Last observed day (~last flight day)	Phenological breadth
Lasioglossum albipes	102	295	193
Lasioglossum albocinctum	37	285	248
Lasioglossum calceatum	32	288	256
Lasioglossum interruptum	51	317	266
Lasioglossum leucozonium	64	287	223
Lasioglossum malachurum	33	365	332
Lasioglossum pauperatum	60	240	180
Lasioglossum pauxillum	56	272	216
Lasioglossum punctatissimum	33	295	262
Lasioglossum villosulum	44	349	305
Megachile apicalis	105	310	205
Megachile pilidens	60	295	235
Megachile sicula	42	299	257
Osmia caerulescens	60	319	259
Osmia cornuta	32	310	278
Osmia niveata	60	282	222
Osmia submicans	49	282	233
Panurgus banksianus	70	247	177
Panurgus calcaratus	60	256	196
Panurgus cephalotes	60	338	278
Panurgus perezi	105	290	185
Rhodanthidium septemdentatum	46	214	168
Rhodanthidium sticticum	44	310	266
Xylocopa cantabrita	36	356	320
Xylocopa violacea	32	365	333
Mean number of days			**245**

Table A7. Species distribution as minimum and maximum longitude and latitude within the study area.

Species	Min/Max Longitude (°)	Min/Max Latitude (°)
Amegilla quadrifasciata	-9.169 / 4.104	36.641 / 43.623
Andrena agilissima	-9.017 / 4.251	36.983 / 43.516
Andrena dorsata	-8.438 / 2.650	37.189 / 43.571
Andrena flavipes	-9.379 / 4.250	36.741 / 43.506
Andrena humilis	-6.649 / 3.097	37.282 / 43.529
Andrena minutula	-8.013 / 3.421	36.795 / 43.516
Andrena nigroaenea	-8.643 / 4.239	36.741 / 43.516
Andrena ovatula	-7.995 / 4.248	36.759 / 43.515
Andrena pilipes	-8.740 / -1.883	37.262 / 43.494
Anthidiellum strigatum	-9.025 / 2.700	36.628 / 43.237
Anthidium florentinum	-9.280 / 3.120	36.648 / 43.008
Anthidium manicatum	-9.194 / 3.157	36.633 / 43.443
Anthophora atroalba	-8.751 / -2.803	36.530 / 42.146
Anthophora bimaculata	-9.187 / 4.045	36.641 / 42.524
Anthophora plumipes	-9.312 / 4.251	36.530 / 43.559
Bombus hortorum	-9.065 / 1.871	39.423 / 43.563
Bombus lapidarius	-8.862 / 2.484	38.768 / 43.464
Bombus lucorum	-8.414 / 3.158	37.252 / 43.654
Bombus pascuorum	-9.463 / 2.993	36.980 / 43.611
Bombus pratorum	-9.184 / 2.633	37.116 / 43.561
Bombus ruderatus	-9.418 / 2.969	36.530 / 43.335
Bombus soroeensis	-6.551 / 0.869	40.680 / 43.642
Bombus sylvarum	-8.360 / 1.380	40.786 / 43.530
Bombus terrestris	-9.473 / 4.262	36.091 / 43.682
Ceratina cucurbitina	-9.025 / 4.101	36.256 / 43.508
Ceratina cyanea	-8.348 / 1.105	36.775 / 43.623
Colletes nigricans	-6.761 / 1.500	36.470 / 42.300
Eucera elongatula	-8.440 / 3.088	36.344 / 43.516
Eucera notata	-8.031 / -2.054	36.530 / 40.445
Halictus fulvipes	-7.912 / 3.430	36.433 / 41.417
Halictus gemmeus	-3.486 / 4.239	36.950 / 41.536
Halictus scabiosae	-9.420 / 4.254	36.450 / 43.625
Halictus subauratus	-8.491 / 3.129	36.702 / 43.529
Hoplitis adunca	-8.673 / 4.239	36.900 / 43.287
Hylaeus clypearis	-8.225 / 4.104	39.353 / 43.298
Hylaeus variegatus	-8.438 / 3.129	37.124 / 43.342
Lasioglossum albipes	-7.026 / -3.849	40.980 / 43.529
Lasioglossum albocinctum	-6.550 / 4.117	37.040 / 42.505
Lasioglossum calceatum	-9.026 / -4.240	39.506 / 43.587

Species	Min/Max Longitude (°)	Min/Max Latitude (°)
Lasioglossum interruptum	-8.544 / 2.035	36.859 / 43.251
Lasioglossum leucozonium	-8.996 / 3.129	37.040 / 43.562
Lasioglossum malachurum	-9.013 / 4.117	36.753 / 43.587
Lasioglossum pauperatum	-8.580 / -1.531	36.433 / 43.529
Lasioglossum pauxillum	-7.856 / 2.692	37.040 / 43.543
Lasioglossum punctatissimum	-8.959 / 4.045	36.741 / 43.529
Lasioglossum villosulum	-7.020 / 4.045	39.416 / 43.558
Megachile apicalis	-8.582 / 3.430	36.529 / 42.444
Megachile pilidens	-9.144 / 4.239	36.529 / 43.416
Megachile sicula	-6.420 / 4.288	37.235 / 40.886
Osmia caerulescens	-6.890 / 4.251	36.092 / 43.452
Osmia cornuta	-6.967 / 2.981	36.674 / 43.355
Osmia niveata	-6.550 / 4.251	36.500 / 43.494
Osmia submicans	-6.681 / 4.288	36.529 / 43.144
Panurgus banksianus	-6.783 / -2.141	36.814 / 43.515
Panurgus calcaratus	-8.840 / -0.940	36.756 / 43.529
Panurgus cephalotes	-6.783 / -4.239	37.195 / 43.443
Panurgus perezi	-8.520 / -4.988	37.124 / 42.431
Rhodanthidium septemdentatum	-9.194 / 4.251	36.228 / 42.291
Rhodanthidium sticticum	-9.242 / 3.098	36.091 / 42.372
Xylocopa cantabrita	-9.209 / 0.029	36.565 / 43.044
Xylocopa violacea	-9.449 / 4.248	36.050 / 43.368

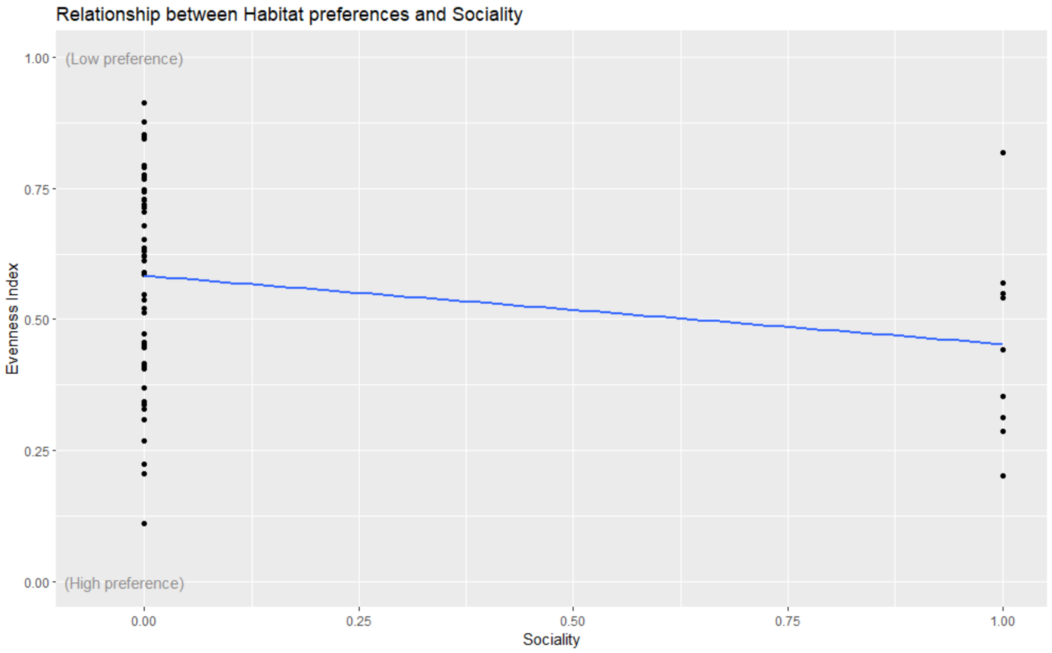

Figure A1. Correlation analysis between Habitat preferences (represented by the Evenness index) and Sociality. Sociality of 0 is a solitary be species, and sociality of 1 is a social bee species. Each black point represents one species of a total of 61 species included in the analysis. The blue line is the non-significant fitted linear regression line, with the estimate -0.13, standard error 0.07 and p-value 0.07.

Extra correlation analysis on phenological and latitude breadth

A potential correlation between phenological breadth and latitude breadth was explored to test if more northern species would have shorter seasons (hypothetically due to colder temperatures). Although this analysis was not originally planned, it was conducted to examine the potential relationship between these two traits, given that data on both were available. The correlation was found to be non-significant (p-value 0.17). Interestingly, there was a weak indication that species with a longer spatial range tended to have longer flight days, contrary to the initial hypothesis. However, it's important to note that only a limited number of species included in the analysis had a restricted spatial range, such as northern species,' which may weaken the comparison. For more details, see the correlation graph in Figure A2 in the appendix.

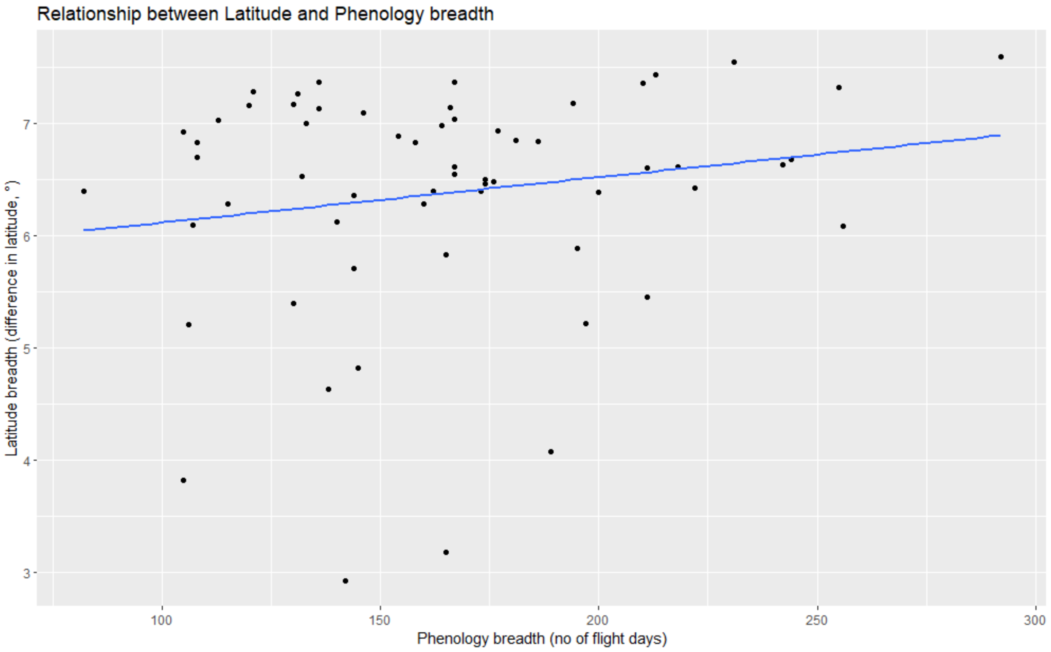

Figure A2. Correlation analysis between Phenological breadth and Latitude breadth. Each black point represents one species of a total of 61 species included in the analysis. The blue line is the non-significant fitted linear regression line, with the estimate 0.004, standard error 0.003 and p-value 0.18.

Publishing and archiving

Approved students' theses at SLU are published electronically. As a student, you have the copyright to your own work and need to approve the electronic publishing. If you check the box for **YES**, the full text (pdf file) and metadata will be visible and searchable online. If you check the box for **NO**, only the metadata and the abstract will be visible and searchable online. Nevertheless, when the document is uploaded it will still be archived as a digital file. If you are more than one author, the checked box will be applied to all authors. Read about SLU's publishing agreement here:

- https://www.slu.se/en/subweb/library/publish-and-analyse/register-and-publish/agreement-for-publishing/.

☒ YES, I/we hereby give permission to publish the present thesis in accordance with the SLU agreement regarding the transfer of the right to publish a work.

☐ NO, I/we do not give permission to publish the present work. The work will still be archived and its metadata and abstract will be visible and searchable.