https://lib.uliege.be https://matheo.uliege.be

Foraging preferences of Barn swallow (Hirundo rustica) analysed by metabarcoding along an urban-rural gradient in Japan

Auteur : Deryck, Antoine
Promoteur(s) : Francis, Frédéric; 8348
Faculté : Gembloux Agro-Bio Tech (GxABT)
Diplôme : Master en bioingénieur : gestion des forêts et des espaces naturels, à finalité spécialisée
Année académique : 2021-2022
URI/URL : http://hdl.handle.net/2268.2/15450

Avertissement à l'attention des usagers :

Tous les documents placés en accès ouvert sur le site le site MatheO sont protégés par le droit d'auteur. Conformément aux principes énoncés par la "Budapest Open Access Initiative"(BOAI, 2002), l'utilisateur du site peut lire, télécharger, copier, transmettre, imprimer, chercher ou faire un lien vers le texte intégral de ces documents, les disséquer pour les indexer, s'en servir de données pour un logiciel, ou s'en servir à toute autre fin légale (ou prévue par la réglementation relative au droit d'auteur). Toute utilisation du document à des fins commerciales est strictement interdite.

Par ailleurs, l'utilisateur s'engage à respecter les droits moraux de l'auteur, principalement le droit à l'intégrité de l'oeuvre et le droit de paternité et ce dans toute utilisation que l'utilisateur entreprend. Ainsi, à titre d'exemple, lorsqu'il reproduira un document par extrait ou dans son intégralité, l'utilisateur citera de manière complète les sources telles que mentionnées ci-dessus. Toute utilisation non explicitement autorisée ci-avant (telle que par exemple, la modification du document ou son résumé) nécessite l'autorisation préalable et expresse des auteurs ou de leurs ayants droit.

FORAGING PREFERENCES OF BARN SWALLOW (*HIRUNDO RUSTICA*) ANALYSED BY METABARCODING ALONG AN URBAN-RURAL GRADIENT IN JAPAN

ANTOINE DERYCK

TRAVAIL DE FIN D'ETUDES PRESENTE EN VUE DE L'OBTENTION DU DIPLOME DE MASTER BIOINGENIEUR EN GESTION DES FORETS ET DES ESPACES NATURELS

ANNEE ACADEMIQUE 2021-2022

(CO)-PROMOTEUR(S): FREDERIC FRANCIS – AYAKO NAGASE

FORAGING PREFERENCES OF BARN SWALLOW (*HIRUNDO RUSTICA*) ANALYSED BY METABARCODING ALONG AN URBAN-RURAL GRADIENT IN JAPAN

ANTOINE DERYCK

TRAVAIL DE FIN D’ETUDES PRESENTE EN VUE DE L’OBTENTION DU DIPLOME DE MASTER BIOINGENIEUR EN GESTION DES FORETS ET DES ESPACES NATURELS

ANNEE ACADEMIQUE 2021-2022

(CO)-PROMOTEUR(S): FREDERIC FRANCIS – AYAKO NAGASE

In collaboration with Chiba University (Japan)

千葉大学
CHIBA UNIVERSITY

With the financial support of Erasmus+ scholarship

Acknowledgements

As part of my master thesis, I would like to express my gratitude to all the people who contributed in a way or another to the realization of this work.

I would like to start by thanking my supervisor at Gembloux Agro-Bio Tech Mr. Frederic Francis who allowed me to integrate the Entomology unit of Gembloux University and to work on this subject in close collaboration with Chiba University in Japan.

This study would not have been possible either without Nagase sensei, my supervisor at Chiba University, who welcomed me and provided monitoring and advice throughout the realization of my research. I am deeply indebted to you.

I could not have undertaken this journey without the experience and supervision of Grégoire Noël that I thank once again for fruitful meetings and advice but also support from the beginning to the end of this master thesis.

I would like to show appreciation to both staff, students, housemates from Chiba University and Gembloux Agro-Bio Tech for creating a supporting and pleasant working atmosphere in the System Planning Lab in Chiba as well as in the entomology unit of Gembloux. I am also thankful to my remote sensing professor Mr. Philippe Lejeune and to the remote sensing assistant professor Mr. Nicolas Latte for providing me with a high-resolution satellite image. Thank you to Kawai sensei as well for taking the time to explain the different steps of the metabarcoding analyses.

Lastly, I would be remised in not offering thanks to my parents who never stopped supporting me all along my studies and way beyond and put everything into action to enable this experience in Japan. My friends Clément, Adèle, Myo and Mathéo are to be acknowledged as well for the help and support they brought me when I needed it. Special thanks to Adrian for the proofreading of my written work.

Abstract

In the current context of urban expansion, natural environments face many changes. Loss of biodiversity, species adaptation, trophic relationships modifications and many other alterations result from this phenomenon and raise conservation concerns. Understanding the interactions between urban environments and organisms is critical in order to tackle the urbanization-related conservation problematics and to take relevant actions to mitigate the disturbances generated by the urban sprawl.

This master thesis aims at investigation the relationship between an Urban-Rural gradient and variables commonly linked to the urbanization phenomenon with the foraging preferences of *Hirundo rustica* (Linnaeus, 1758). This bird, well present in cities and appreciated for feeding on pest species display differences in morphological and breeding characteristics in urbanized areas. This work strives to measure the impact of this urban sprawl on the foraging behavior and to contribute to the diet studies realized via the DNA metabarcoding technique.

In order to achieve these goals, *Hirundo rustica* (Linnaeus, 1758) feces were collected in 75 sampling sites in the Kanto region, Japan. The data collection took place in June and July 2020. The DNA sequences from the COI-5P region present in the excrements were sequenced via the DNA metabarcoding technique and a taxonomy was assigned to these DNA fragments thanks to a Basic Local Alignment Search Tool (BLAST) on a complete reference sequence database from the National Center for Biotechnology Information and on a curated Barcode of Life Data System database limited to the Japanese insects' records. The results obtained through both databases were compared. Three times less species were identified with the curated database, and it was suggested that this specific database needed more contribution and development to accurately investigate insectivorous organisms' diet. These analyses imply that *Hirundo rustica* (Linnaeus, 1758) mostly feeds on Diptera. Coleoptera, Lepidoptera, Hemiptera and Hymenoptera represent an important part of its diet as well.

The α diversity of insects found in the fecal matter was then assessed for each sampling site. The Shannon and the observed taxa richness were calculated in this purpose. The β diversity between sampling sites was estimated as well via the Jaccard index. The relationship with the urban-rural gradient and the environmental variables has been studied with Generalized Linear Mixed Models for the α diversity and with a Permutational Multivariate Analysis of Variance for the β diversity.

The obtained results suggest that the fragmentation, the average temperature, the amount of precipitation and the radiance explain variation in the α diversity. As for the β diversity, significant statistical relationships with the average temperature, the range of temperatures and the amount of precipitation have been found.

This study attempts to provide relevant information to the scientific communities and to help understanding how trophic interactions change with disturbances generated by urbanization, providing support in decision-making process for the establishment of green infrastructures and the implementation of environmental measures in territories affected by the urban sprawl.

Résumé

Dans le contexte actuel d'expansion urbaine, les environnements naturels sont confrontés à de nombreux changements. La perte de biodiversité, les changements de comportements des espèces, les modifications des relations trophiques et de nombreuses autres altérations découlent de ce phénomène et soulèvent des préoccupations en matière de conservation. Il est essentiel de comprendre les interactions entre milieux urbains et organismes afin de s'attaquer aux problèmes de conservation liés à l'urbanisation et de prendre des mesures pertinentes afin d'atténuer les perturbations générées par ce développement urbain.

Ce mémoire vise à étudier la relation entre un gradient Rural-Urbain et des variables communément liées à l'urbanisation avec les préférences alimentaires d'*Hirundo rustica* (Linnaeus, 1758). Cet oiseau, bien présent dans les villes et apprécié en tant que régulateur d'espèces nuisibles, présente des différences de caractéristiques morphologiques et reproductives dans les zones urbanisées. Ce travail vise à mesurer l'impact de l'urbanisation sur les préférences alimentaires de ces oiseaux et à contribuer aux études de leur régime alimentaire réalisées via le metabarcoding ADN.

Afin d'atteindre ces objectifs, des excréments d'*Hirundo rustica* (Linnaeus, 1758) ont été prélevés dans 75 sites d'échantillonnage de la région de Kanto au Japon. La collecte des données a eu lieu en juin et en juillet 2020. Les séquences d'ADN de la région COI-5P présentes dans les excréments ont été séquencées par la technique de metabarcoding ADN et une taxonomie leur a été attribuée grâce à l'outil BLAST appliqué sur une base de données complète du National Center for Biotechnology Information et sur une base de données du Barcode of Life Data System restreintes aux données d'insectes collectés au Japon. Les résultats obtenus dans les deux bases de données ont été comparés. Trois fois moins d'espèces ont été identifiées avec la base de données traitée ce qui laisse penser que cette base de données spécifique a besoin de plus de contribution et de développement afin d'étudier avec précision le régime alimentaire d'organismes insectivores. Ces analyses suggèrent qu'*Hirundo rustica* (Linnaeus, 1758) se nourrit principalement de Diptères. Coléoptères, Lépidoptères, Hémiptères et Hyménoptères représentent également une partie importante de son régime alimentaire.

La diversité α des insectes trouvés dans les matières fécales a ensuite été évaluée pour chaque site d'échantillonnage. Les mesures de diversité de Shannon et de richesse observée de taxons ont été calculées à cette fin. La diversité β entre les sites d'échantillonnage a également été estimée au moyen de l'indice de Jaccard. La relation avec le gradient urbain-rural et les variables environnementales a été étudiée respectivement avec des modèles linéaires mixtes généralisés pour la diversité α et avec une analyse permutationnelle multivariée de la variance pour la diversité β.

Les résultats obtenus suggèrent que la fragmentation, la température moyenne, le taux de précipitations et la radiance soient responsable des variations de diversité α alors que des relations statistiques existent entre la diversité β et la température moyenne, les différences de température et le taux de précipitations.

Cette étude vise à apporter des informations pertinentes aux communautés scientifiques et à mesurer l'impact des perturbations générées par l'urbanisation sur les interactions trophiques afin de fournir des arguments scientifiques pour la mise en œuvre d'infrastructures vertes et de mesures environnementales dans les territoires touchés par l'étalement urbain.

Table of content

List of figures

List of tables

1. Literature review

1.1. Biodiversity decline

During its whole History, planet Earth faced five mass extinction events. The last of them is known as the "Cretaceous-tertiary extinction" and occurred 65 million years ago. Extinction is a natural process that has always existed and which, in balance with the speciation process, is critical to allow evolution (Mazzucco, n.d.). However, during mass extinction events, the rate at which the species disappear largely accelerates and causes the loss of at least 75% of all living species in a geologically brief time span often considered as lower than two million of years (Ritchie and Roser 2021). A high extinction rate ranging from 1,000 to 10,000 times the natural one and occurring during a long period might thus leads to such events and corresponds to the current estimated pace. It is thus globally considered that Earth is heading towards its sixth extinction. This probable sixth extinction event might nevertheless be slightly different than the five previous ones. While the past events have been caused by natural events such as change in tectonic activity, climate and sea-level, the sixth mass extinction should be mainly caused by threats resulting from anthropogenic activities such as resources overuse, fast-paced climate change, habitat degradation and invasion of exotic species (Martin 2019; Purvis, Jones, and Mace 2000).

This biodiversity decline impacts every animal class, but some organisms seem to be significantly more impacted than the others. This is the case of insects whose decline, although poorly studied in the past, gained the scientific community attention as well as the one from the broad public during these past few years. This rise of attention allowed scientists to reveal concerning trends about these arthropods and, in 2020, the International Union for Conservation of Nature (IUCN) ranked this taxon as the third most threatened class of the animal kingdom with 1,848 species considered vulnerable, endangered, or critically endangered behind the fishes and the amphibians with respectively 3,040 and 1,390 threatened species. Moreover, a significant amount of insect species has not been described yet and only 1% of the described ones have been taking into account in the IUCN report. The count of insect species threatened by extinction is thus believed to be massively underestimated by the IUCN, compared to the fishes and the amphibians for which correspondingly 60% and 87% of the described species have been considered (Table 1) (Ritchie and Roser 2021; Cardoso et al. 2020).

Table 1 - Amount of species classified as vulnerable, endangered or critically endangered by the IUCN for each class of the animal kingdom and associated percentage of described species taken into account. Adapted from Ritchie and Roser (2021). Data from the IUCN Red List (IUCN 2020).

Taxonomic groups	Number of species classified as vulnerable, endangered or critically endangered	Described species taken into account into the vulnerable, endangered or critically endangered species count (%)
Fishes	3,040	60
Amphibians	2,390	87
Insects	1,848	1
Birds	1,481	100
Reptiles	1,449	73
Mammals	1,317	92

The Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES) also tried to quantify the amount of insect species threatened by extinction, but its approach is conservative and tends to underestimate the reality. Its estimations suggested that about 10% of the insect species were facing the threat of extinctions and that out of one million of threatened animal and plant species, half a million were insects (Tonissen 2019; Cardoso et al. 2020). Besides a loss in species diversity, insect abundance and biomass also seem to be decreasing. Even though large-scale analyses are missing, severe depletions have been observed in locations wide spread across the world (Samways et al. 2020) and studies about some better monitored groups such as pollinators and Lepidoptera confirm these trends (Dirzo et al. 2014). This class plays key roles in the ecosystems and its decline adds up to the other factors responsible for biodiversity decline, threatening large groups of species such as aerial predators that mostly prey upon them (McClenaghan, Nol, and Kerr 2019).

1.2. Urban development

1.2.1. Context

The urbanization phenomenon is defined as the “growing concentration of the population around existing urban agglomerations” (Le Robert n.d.) and is thus characterized by the expansion of urban and peri-urban areas at the expanse of rural environments. This process has been existing for thousands of years, but its rate started to soar during the 19th century with the industrialization of the society and is still on the rise nowadays: according to the United Nations, 55% of the worldwide population currently lives in urban territories and this number is expected to reach 68% in 2050 as the Human population keeps on thriving and as rural-urban migrations are still prevailing in today’s society. The scale and the rate of this process are the reasons why the urban sprawl is an hot topic and represents a contemporary source of concern (Ritchie and Roser 2018).

1.2.2. Impact on biodiversity

Such modifications of the land-use and their associated infrastructures bring significant disturbances to the environment and its inhabitants. Some cycles such as the water, the energy and the nutrients ones find themselves modified and factors such as temperature, state of preservation of the environment, levels of light, noise, and pollution face significant changes. Typically, due to a decrease of vegetation cover in favor of impervious surfaces and high-rise infrastructures, cities represent heat traps, organic matter is present in fewer quantities, the evaporation process and thus the rain pattern are disturbed and the remaining patches of natural habitats are dispersed (Fenoglio et al. 2021; New 2015; Van Nuland and Whitlow 2014).

This change in land use and habitat type is responsible for changes in the communities assemblages (McIntyre et al. 2001; Cardoso et al. 2020) and is particularly observable on arthropod populations whose lifespans are short and reproductive rate elevated (Van Nuland and Whitlow 2014). Because of the important amount of modifications related to urbanization, generalist species tend to become predominant in these urban ecosystems. Thermophilous species also benefit from the higher temperatures and high mobility organisms tolerate better the habitat fragmentation (Fenoglio et al. 2021). These features often correspond to the ones of invasive species which take over the urban environment while the native species populations shrink (Van Nuland and Whitlow 2014). As a general result, a tendency to the homogenization and thus to a diminution of the beta-diversity of the communities is observed (New 2015). The insect populations distribution patterns also undergo modifications (Zhang and Feng 2018). These modifications are likely to lead to significant consequences since insect communities are mostly composed of a large number of uncommon species and few common species and that these communities are largely heterogenous in space due to their ability to colonize micro-habitats (Cardoso et al. 2020).

Organisms are therefore either adapting or fleeing from these extending territories. The adaptation to such environments present benefits such as the presence of infrastructures representing shelters, the low abundance of predators and the possibility to feed on anthropogenic wastes but require a high tolerance towards human disturbances and lower food quality (Lowry, Lill, and Wong 2013; Cahill et al. 2018). These variations lead to behavioral responses from individuals settling in cities and phenotypic differences between urban and rural populations emerge within species (Lowry, Lill, and Wong 2013). It has been especially observed with reproductive and foraging traits deeply altered in urban populations for birds and mammals species (Lowry, Lill, and Wong 2013).

This is the case of *Hirundo rustica* (Linnaeus, 1758), an insectivorous passerine closely associated with human infrastructures in either rural and urban environments (A. Turner 2015). This proximity with humans and its wide geographical range facilitates its studies making it a valuable and suitable case study to acutely comprehend the interactions between wildlife and urban environments (Cahill et al. 2018; Zhao et al. 2021). Their generalist and flexible diet is an asset as well in order to study the insect communities and the impact of the urban sprawl on them (McClenaghan, Nol, and Kerr 2019). Their foraging behavior has been chosen as the focus of this study for these reasons.

1.3. *Hirundo rustica* (Linnaeus, 1758)

1.3.1. Ecology and biology

Hirundo rustica (Linnaeus, 1758), also known as the barn swallow is a bird from the Hirundinidae family (BirdLife International 2019). This family gathers 20 genus and 89 species of passerines and is thus part of the passeriform order (Oiseaux.net n.d.). More than half of the living birds species belong to this large order (Gill, Clench, and Austin 2021). Individuals from the Hirundinidae family are however less numerous and share some characteristics: their body is spindle-shaped and their wings long and sharp (Britannica 2011).

The six subspecies of *Hirundo rustica* (Linnaeus, 1758) do not escape this rule and expose fusiform bodies and sickle-shaped wings. This swallow species can be distinguished from the others thanks to its indented tail and color pattern. The upper part of its body is dark while the lower part is white or reddish with a red neck and a dark band separating these two color patterns (Gailly n.d.) (Figure 1).

Figure 1 - Hirundo rustica (Linnaeus, 1758) individuals. The picture on the left captured a juvenile in a nest in a sampling site of Chiba. The picture on the right shows an adult individual feeding its offspring in a sampling site of Nishitokyo.

Another characteristic of *Hirundo rustica* (Linnaeus, 1758) is its geographical range of 251 million km² which represents the largest one within the passeriform order. It includes a large diversity of altitudes and habitats (BirdLife International 2019). These birds can be found either in rural and urban environments. The presence of open areas such as wastelands, grassy fields, pastures or even lakes is nevertheless necessary since they represent suitable foraging sites for the species. They also settle near wet areas which provide them water and mud (necessary for the nest construction) besides providing additional foraging sites (Savignac 2011). This adaptability allowed them to colonize every continent except Antarctica (Figure 2). While North America, Europe and Asia mainly correspond to their native breeding sites, South America, Africa and Oceania mostly answer their need for wintering grounds (BirdLife International 2019). *Hirundo rustica* (Linnaeus, 1758) is therefore a long-distant migrant. The period at which they can be found in the Northern Hemisphere depends on the regions but mostly corresponds to their breeding season (Savignac 2011). Once the reproduction completed, individuals reach the wintering ground in the Southern Hemisphere in order to benefit from warmer temperatures (Pancerasa et al. 2018).

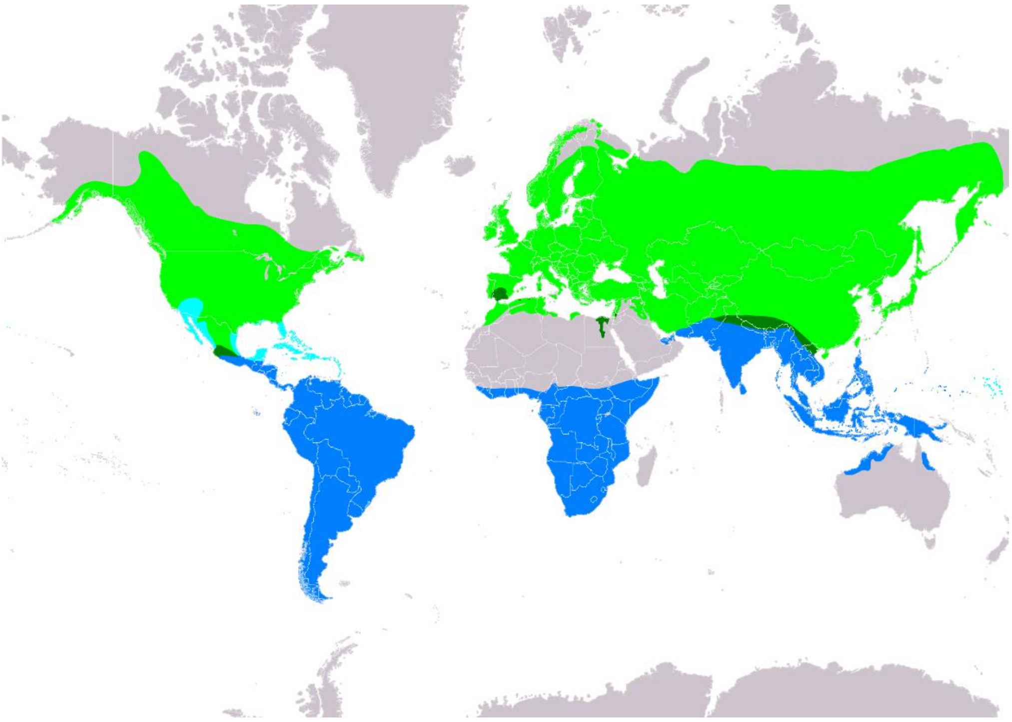

Figure 2 - Geographical range of Hirundo rustica (Linnaeus, 1758). The light green areas represent breeding sites, the dark green zones are all-year residential sites, the light blue locations match passage sites and the dark blue ones correspond to wintering sites. From Alexander Kürthy (2021).

During the breeding season, these birds generally gather in small colonies in territories closely associated with human activities. Males display a courtship behavior to attract females in order to form a pair (Møller 1990) and, once formed, the couples remain monogamous for the rest of their lives. They also show a significant nesting sites fidelity and reuse old nests as much as possible (Safran 2010). In the past, these muddy nests used to be built in natural cavities (such as caves, cliff flanges,…) but *Hirundo rustica* (Linnaeus, 1758) mostly erects them in human infrastructures (such as building eaves, frames and walls, barns,…) nowadays (Canada 2011). The couples then lay two to seven eggs every year and both parents take care of the eggs and of the offspring together. Pairs from the same colony are nevertheless territorial and cooperation within individuals from different pairs has never been observed (Moore and Breed 2015).

1.3.2. Diet and foraging behavior

Hirundo rustica (Linnaeus, 1758) is a generalist insectivore which mostly feed on the wing. Its diet is highly variable and significantly depends on the prey availability (McClenaghan, Nol, and Kerr 2019) itself hinging on different parameters such as the spatio-temporal context, the weather condition, the type of habitats and their conditions, the human disturbances,… (Van Nuland and Whitlow 2014; Zhang and Feng 2018). Some patterns are however frequent, such as the prevailing presence of species from the Diptera, Coleoptera, Hymenoptera and Hemiptera orders in their diet. Besides these orders and depending on the availability in the foraging sites, Lepidoptera can represent a significant part of their alimentation. Other insect orders such as Isoptera, Ephemeroptera and to a lesser extent Odonata, Siphonaptera, Trichoptera, Orthoptera and Blattodae have been observed in swallows' feces as well. Non-insect arthropods such as members of the Aranae and even Julida orders are not excluded from these swallows' diet either. This list of orders relying on previous studies realized in Canada, Poland, Scotland, Crimea, Malawi, South Africa and Malaysia is non-exhaustive and research in other regions from *Hirundo rustica* (Linnaeus, 1758) distribution range still have to be carried out in order to extent it (A. Turner 2010; Orłowski and Karg 2011; 2013; Law et al. 2017; Kusack 2018; McClenaghan, Nol, and Kerr 2019; Mansor et al. 2020). Even though *Hirundo rustica* (Linnaeus, 1758) diet fluctuates depending on the regions and the moment of the year, same trends have been observed in various seasons and locations. In addition to the species abundance which seems to be the most important factor influencing *Hirundo rustica* (Linnaeus, 1758) diet, preys' size and flight performances are also taken into account during the prey selection. Insects measuring between four and eight millimeters are preferred to smaller insects which provide low amounts of energy and to larger ones, too imposing for their beaks. As for the poor flight performances, they allow these swallows to catch their quarries with more ease and thus to save energy (A. Turner 2010; Law et al. 2017; McClenaghan, Nol, and Kerr 2019).

In order to catch these aerial preys, *Hirundo rustica* (Linnaeus, 1758) individuals seek for open land and water (see section 1.2.1. Ecology and biology) where insects gather. During the breeding season, they mostly forage in low numbers or individually (Savignac 2011) within a perimeter of 400 to 500 m around the nest sites (Snapp 1976; Evans, Wilson, and Bradbury 2007; Savignac 2011; Kang and Kaller 2013) and at low altitudes (generally less than 10 meters high) (Savignac 2011) since insect size and abundance are negatively correlated with the distance to the ground (A. Turner 2010). Besides these flying insects, these birds can fall back on larvae and wingless arthropods if the weather conditions are unsuitable for aerial insect gatherings or if large groups of wingless arthropods regroup in an easily accessible place. In this case, individuals will simply walk or slowly soar above their preys and capture them (Fitzsimons and Thomas 2012; A. Turner 2010).

Human-modified ecosystems facilitate *Hirundo rustica* (Linnaeus, 1758) feeding since artificial lighting lures insects that display a phototactic behavior and since anthropogenic activities such as crops burning, trimming, and tilling dislodge insects from out-of-range habitats. The presence of herds of animals (including farm animals) is also beneficial to *Hirundo rustica* (Linnaeus, 1758) since they also ensure the dislodging process by trampling the ground and produce excrements which attract their preys (especially large Diptera and Coleoptera). The diet and the foraging behavior of these swallows make their presence most of the time tolerated in cities and agricultural environments even though important quantity of feces are accumulated under the nests. As generalist insectivores able to adapt their diet depending on their habitat, they are great consumers of undesired insects such as stable flies, horse flies and mosquitoes and regulate populations of Human, crop and farm animal pests (McClenaghan, Nol, and Kerr 2019; Orłowski and Karg 2013).

1.3.3. Status and threats

Hirundo rustica (Linnaeus, 1758) has been classified as a least-concern species (LC) by the International Union for Conservation of Nature. This status means that the protection of this bird does not represent a major preoccupation at a global scale. The IUCN classification is based on the population size, the decline rate, the distribution area and the population and distribution range fragmentation. The tremendous distribution area of this species along with its important population size and its relatively moderate decline rate explains why *Hirundo rustica* (Linnaeus, 1758) does not fall into the vulnerable status. Its global population has been estimated between 290 and 487 millions of sexually mature birds and is currently shrinking at a rate lower than the one required for the species to be considered as vulnerable by the UICN (30% every 10 years) (BirdLife International 2019).

Even though the species is not considered as vulnerable at a global scale it still undergoes a significant decline. The cause of this decline is most likely a combination of various threats differing from one region to another. Habitat loss and degradation, drop in insect abundance (Savignac 2011; Imlay and Leonard 2019) along with weather change and weather events are thought to be the main threats that these birds faces. Habitat loss and degradation includes the replacement of wooden constructions (such as barns) in favor of concrete ones. Agricultural areas which represent foraging sites for *Hirundo rustica* (Linnaeus, 1758) have also been impacted by a shift of conventional farming techniques to modern ones (change in land use, utilization of chemicals and pesticides, pest-resistant cultures,..) and could play a role in the species decline (Savignac 2011). Some other threats, less studied, could also belong to the factors responsible for this decline. Indeed, collisions with cars and wind turbines, interspecific and intraspecific competition, diseases, predation and contaminants are considered liable for important amount of swallow deaths as well (Imlay and Leonard 2019). Besides this population decline, modifications of phenotypic characteristics are frequent and their impacts on individuals require to be investigated as well. In China, a negative correlation between urbanization and body weight has been observed within male individuals while this is a negative correlation between urbanization and wing length that has been determined within female individuals (Zhao et al. 2021). These phenology modifications are also believed to be linked to diet alterations due to changes in prey populations and distribution along with habitat loss and degradation (Savignac 2011; Imlay and Leonard 2019).

1.4. Metabarcoding

1.4.1. Concept and steps

The DNA metabarcoding is a technique that aims at identifying organisms from a sample on the basis of their DNA sequences. It emerged from the DNA barcoding technique and differentiated itself from it by its ability to detect not one but several species from the same sample. This can be done using high throughput sequencing, a technology which sequences thousands to millions of DNA or RNA fragments at the same time (Cristescu 2014; M. Liu et al. 2020; Haarsma, Siepel, and Gravendeel 2016). The barcoding and metabarcoding methods rely on the principle that every species is characterized by a genome and that species can be distinguished from one another thanks to differences in the nucleotides sequences of their DNA. These techniques consist therefore in sequencing a specific DNA region of the fragments present in a sample and to associate these fragments to taxa based on the nucleotides sequence of the targeted region (Cristescu 2014). Since DNA remains in the environment for a certain amount of time (variable depending on biotic and abiotic conditions) (Nielsen et al. 2007), analyzing environmental samples such as water, soil, sediment, air or biological materials by metabarcoding enables scientists to be aware of the species that were or are present in an area (as long as the stage of degradation of their DNA is not heavily advanced yet) (Bush et al. 2019).

The first pre-step of this process is to select a barcode (DNA fragment that can be found in a wide range of studied species) that is highly variable between and lowly varying within a species (Haarsma, Siepel, and Gravendeel 2016; Cristescu 2014). This barcode, also called DNA marker, should be different depending on the studied taxa and the desired taxonomic resolution. While a longer marker permits a better taxonomic resolutions, it makes it more unlikely to be sequenced correctly and thus to identify species (especially those whose DNA is degraded) (M. Liu et al. 2020; Ando et al. 2020). The occurrence of the barcode in reference sequence databases also need to be taken into account while selecting it. Once the barcode has been picked, the primers choice can start. Primers are short DNA or RNA sequences that are complementary to the extremities of the DNA marker, and which allow its amplification. They should therefore be selected in accordance with the barcode (M. Liu et al. 2020).

For insects, this barcode will most of the time be part of the COI (cytochrome c oxidase subunit I) gene (Porter and Hajibabaei 2020).

The second pre-step consists in picking or building a reference sequence database. Reference sequence databases, or taxonomic databases, are tools that associate nucleotides sequences to taxa (McClenaghan, Nol, and Kerr 2019). They are consequently critical in order to identify the species whose DNAs are present in a sample. Several databases of this kind exist and gather sequences from different genes and different taxa. Differences between databases are thus significant (Porter and Hajibabaei 2020; M. Liu et al. 2020). Some software programs allow the construction of custom databases, regrouping or filtering existing databases and generated ones, and are utilized in several studies (Li et al. 2021). The generated databases can result, for instance, from DNA sequences of organisms captured by traps located on the studied area (McClenaghan, Nol, and Kerr 2019).

The third step is defined by the start of the DNA manipulation (Figure 3). DNA extraction and its amplification by polymerase chain reaction (PCR) characterize this step. For the purposes of an optimal DNA replication, extracting the DNA from the cells is required (Keele et al. 2020). This step is commonly achieved using DNA commercial extraction kits (M. Liu et al. 2020). Following this process, the barcode of each extracted DNA fragment will be replicated by PCR in order to be abundant enough to be sequenced (Keele et al. 2020). These replicated barcodes are called amplicons (M. Liu et al. 2020).

As a fourth step, amplicons are simultaneously sequenced in order to determine their nucleotides assemblages (Figure 3). The choice of the DNA sequencer must consider the amplicons lengths and the sequencer read speed. Low-quality sequences should be removed (Ando et al. 2020).

Eventually, the fifth and last step revolves around the assignment of the read sequences to sequences from taxonomic databases and thus to identified taxa (Figure 3). A threshold of 97% matches between read sequence and reference sequence is often used to acknowledge an identification (Nearing et al. 2018; Dopheide et al. 2019; Raclariu-Manolică et al. 2021; Drake et al. 2022).

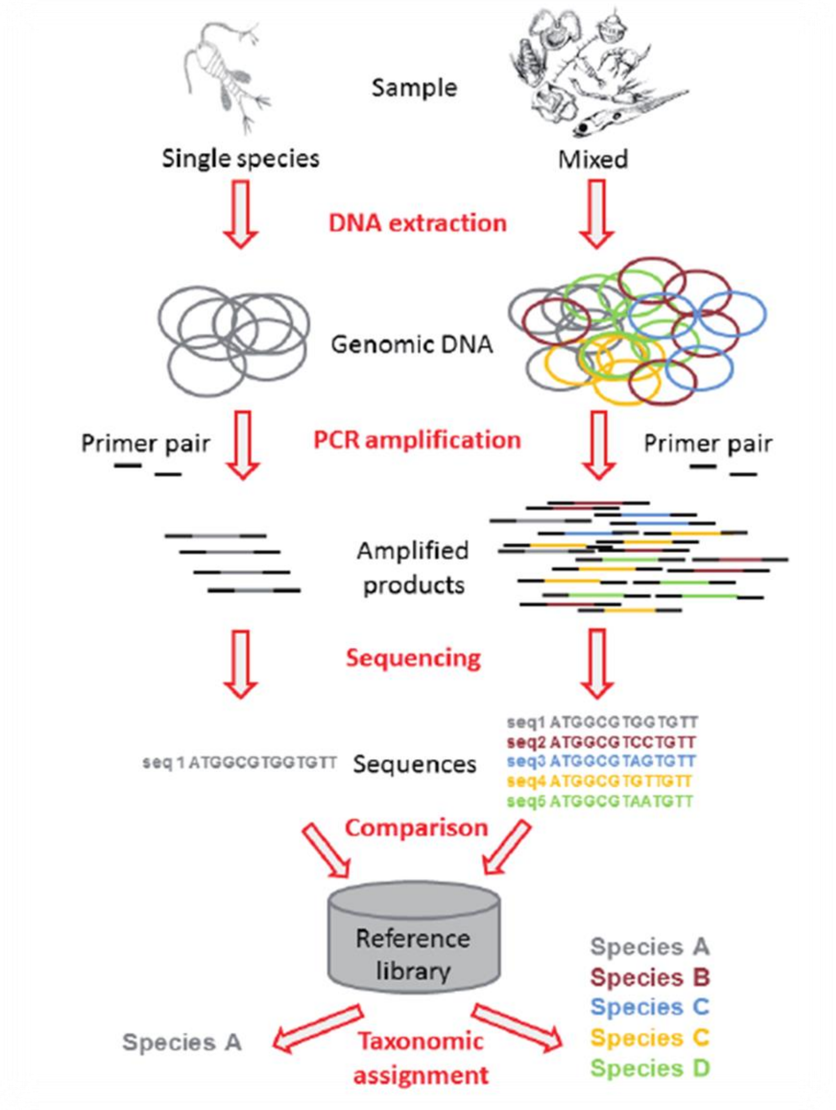

Figure 3 - Main steps of the DNA barcoding and metabarcoding manipulation process. On the left is illustrated the DNA barcoding method while the DNA metabarcoding technique is displayed on the right. From Corell and Rodriguez-Ezpeleta (2014).

1.4.2. Insect metabarcoding and metagenomic

For a long time, insect communities have been studied through visual identification. However, these last few years have seen the emergence of DNA metabarcoding techniques in entomological research. These methods show considerable advantages compared to morphological identification. Instead of identifying insects individually, DNA metabarcoding allows the simultaneous identification of most of the species present in a sample and is thereby less time-consuming. According to the same logic, DNA metabarcoding analyses require less labor than visual identification (for which entomologist experts need to be mobilized) displaying economic advantages (Trevelline et al. 2018; Nørgaard et al. 2021).

Insect communities are typically studied by DNA metabarcoding of entire specimens collected through pitfall traps (Leather 2008), funnel traps (Hufnagel and Rédei 2003; Allison et al. 2011), Malaise traps (Brandon-Mong et al. 2015; Braukmann et al. 2019; Marquina et al. 2019), nets mounted on rooftops of cars (called car nets) (Svenningsen et al. 2021) or even automatic light traps (Zenker, Specht, and Fonseca 2020; Mata et al. 2021). But generalist predators' diets analyses revealed themselves efficient as well in order to examine the biodiversity of some taxa, including the insect class (Trevelline et al.

2018; Nørgaard et al. 2021): as a generalist insectivore is likely to feed on any insect whatever the species, its diet is expected to contain most of the insect species which inhabit its foraging territory. These researches consist thus in examining the organisms remains from guts, stomachs or feces as a proxy for the diversity of an area and are specifically convenient to sample unattainable for humans but exploited by wildlife areas (such as private lands or cluttered environments) (Nørgaard et al. 2021). The usage of DNA metabarcoding techniques for this purpose is justified by the fact that, in addition to the advantages of metabarcoding already cited, digested organisms are particularly complex to identify visually (Trevelline et al. 2018; Ando et al. 2020).

Among the three kinds of diet analyses previously mentioned, the fecal metabarcoding approach gained the deepest interest among the ecologists, compared to guts or stomachs examination (Ando et al. 2020). This is explained by the non-invasive character of the method: in opposition to conventional biodiversity assessments which disrupt wildlife, the feces collection generates few disturbances and enables the monitoring of high-conservation priority species. This method has been employed in more than 155 published researches since 2009, most of which studying mammal diets (felids, small mammals and bats) and, in smaller measures, birds (Figure 4) (Ando et al. 2020). Studies of this kind applied to insectivorous species allowed to gain knowledge on a wide variety of them, their extensive range of environments but also on their preys. For instance, researches on the diets of three songbirds (*Parkesia motacilla* (Vieillot, 1809), *Hylocichla mustelina* (Gmelin, 1789), and *Empidonax virescens* (Vieillot, 1818) provided information about phenological shifts in some species abundance and the importance of aquatic insects in riparian habitats, studies on bats species such as *Plecotus auritus* (Linnaeus, 1758) permitted to characterize the shift in prey assemblages across seasons (Andriollo et al. 2019), the examination of the Pyrenean desman *Galemys pyrenaicus* (E. Geoffroy Saint-Hilaire, 1811) feces allowed scientists to determine the insect orders present in its environment, investigations on the diet of *Hirundo rustica* (Linnaeus, 1758) gave an insight of insect communities in anthropogenic areas (McClenaghan, Nol, and Kerr 2019) and the list is still long. However, relatively few bird fecal studies exist especially for the members of the Passeriformes order (Trevelline et al. 2018) which are predominantly insectivorous (Gill, Clench, and Austin 2021) and which could bring reliable information about the insect communities.

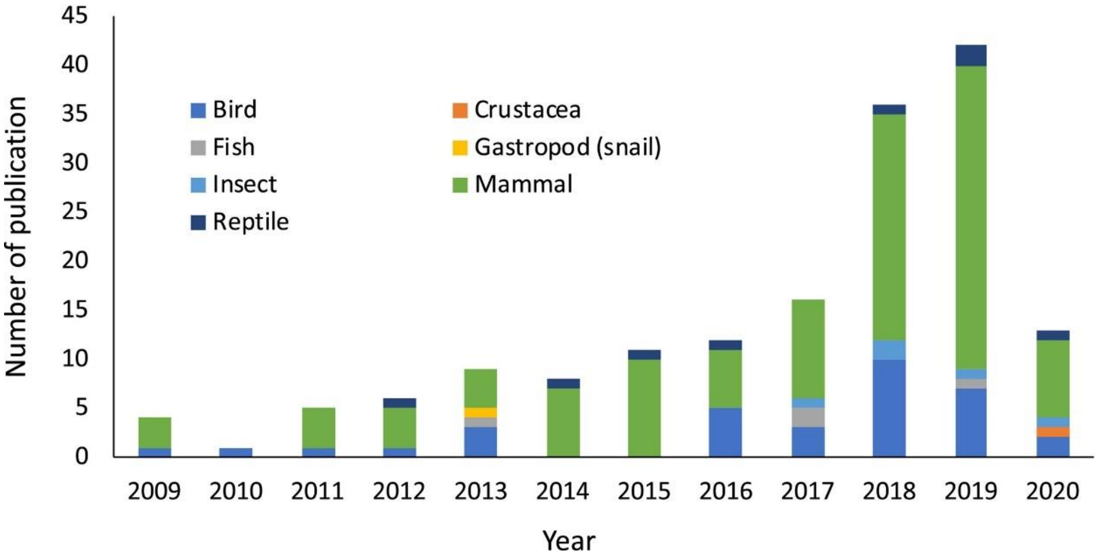

Figure 4 - Annual count of international journals publications about fecal metabarcoding analyses. The amount of articles in 2020 only takes into account the ones published before March. From Ando et al. (2020).

Most of these fecal metabarcoding analyses are carried using the ZBJ-ArtF1C/ZBJ-ArtR2c primers pair when targeting arthropods. This pair clearly outweighs the other ones and has been reported in 35 papers describing arthropods communities. Two other pairs frequently met in DNA metabarcoding researches have both been reported in two fecal metabarcoding papers and other pairs were only reported in one or less article (Table 2) (Ando et al. 2020). The success of the ZBJ-ArtF1C/ZBJ-ArtR2c partly comes from its ability to amplify the degraded DNA present in feces (Trevelline et al. 2016; M. Liu et al. 2020). As for the databases, the Barcode of Life Database (BOLD) and GenBank database (from the National Centre for Biotechnology Information) are the two largest and most utilized ones (Macher, Macher, and Leese 2017).

Table 2 - Primer pairs used in at least two fecal metabarcoding studies. Table derived from Ando et al. (2020).

Primer names	Direction	Target region	References	Number of publications
ZBJ-ArtF1c ZBJ-ArtR2c	F R	MtCO1	Zeale et al. (2011)	35
LepF1 MLepF1_Rev	F R	MtCO1	Brandon-Mong et al. (2015)	2
IN16STK-1F-mod IN16STK-1R-mod	F R	165 rDNA	Kartzinel and Pringle (2015)	2

1.4.3. Limits

Despite its broad potential, there are still several limits to the fecal DNA metabarcoding techniques and biases are therefore frequent.

Few of these biases and limits are associated with the studies operating mode and are, even though mitigable, too complex to avoid for now. As explained in the section “2.4.1 concept and steps”, several steps composed the methodology, and the biases can emerge from some of them. The feces and DNA samples collection and conservation, the choice of the primers and the sequence reference database are the main sources of inaccuracy. Feces and DNA samples are indeed subject to contamination and degradation during the collection and conservation processes. Degradation mainly happens when the samples are left in the environment for a prolonged amount of time and undergo abiotic conditions (UV exposure, temperature, soil acidity) as well as biotic conditions (actions of microorganisms, enzymes,…). Unsuitable DNA storage conditions can also lead to a significant level of degradation (Ando et al. 2020). As for the contamination, this process mainly occurs on the field where organisms are likely to get into contact with the feces (especially coprophagous arthropods) and at a lesser extent in the laboratory where samples are manipulated and chemically treated (McClenaghan, Nol, and Kerr 2019; Ando et al. 2020). Non-arthropod species that contaminated the samples are however easily discarded from the analyses either because the selected primers poorly replicate their DNA or because they are readily detected during the taxonomic assignment (Uiterwaal and DeLong 2020). The choice of the primers also has an importance since they favor the DNA replication of some species during the PCR and thus entail an overrepresentation of some taxa. For instance, the use of the ZBJ-ArtF1C/ZBJ-ArtR2C pair of primers cause a bias towards a higher detection of insects from the Diptera and Lepidoptera order (McClenaghan, Nol, and Kerr 2019). In addition to that, the choice of the sequence reference database will impact the taxonomic assignments since different databases possess different sequences and no existing databases is complete (McClenaghan, Nol, and Kerr 2019; Uiterwaal and DeLong 2020). Because they still have to be improved, most of these databases only allow the taxonomic identification at a genus or even family level for a wide range of species which represent a constrain for studies requiring species-level identification (Haarsma, Siepel, and Gravendeel 2016; Scasta et al. 2019; McClenaghan, Nol, and Kerr 2019).

Besides the DNA material from insects attracted by the swallows' feces previously mentioned, can be found DNA from organisms preyed upon by swallows' preys themselves. It means that organisms that have been ingested by predatory arthropods which have in turn been ingested by *Hirundo rustica* (Linnaeus, 1758) can remain in the feces and be identified by DNA metabarcoding. During the analyses, some species might then be considered as part of the species diet even though they're not. The same applies to parasites that could be carried by arthropods or directly consumed by the birds (McClenaghan, Nol, and Kerr 2019).

Another bias in biodiversity assessment based on predator diets can be caused by the dietary preferences of generalist predators. This bias is particularly represented in studies of omnivore diets. For example, the generalist omnivorous red foxes and European badgers feed with the purpose of meeting their nutritional needs and select their preys in accordance. They therefore do not consume every available prey species and their diet alone cannot replace invasive assessment methods (Nørgaard et al. 2021).

Eventually, the last and considerable constrain met during DNA metabarcoding studies is that it cannot be reliably used as a quantitative method. The numbers of reads of a sequence is not merely influenced by the abundance of a species but by its body size and mass, the quantity of the targeted genes in its cells, its digestibility, its abundance, and bias affecting the replications. Adjustive factors based on species body size or mock communities exist but are too imprecise and not available for enough species. Qualitative methods or semi-quantitative ones based on the incidence of the species among different samples of a defined area are consequently the sole accurate methods (Suzuki and Giovannoni 1996; Braukmann et al. 2019; McClenaghan, Nol, and Kerr 2019; Uiterwaal and DeLong 2020).

1.5. Hypotheses and objectives

Even if research on species diet composition employing fecal DNA metabarcoding practices are more and more frequent, only few of them consider the effect of an urbanization gradient on them (Noël et al. 2021; Richardson et al. 2021; Spence, Wilson Rankin, and Tingley 2022). However, knowing the extent of the urbanization phenomenon and the key roles that trophic interactions play in the ecosystems, the characterization of the impacts of this societal process on these should be of great concern for Humans (Van Nuland and Whitlow 2014). This research aims at contributing to the comprehension of these interaction by studying the foraging preferences of *Hirundo rustica* (Linnaeus, 1758) along an urban-rural gradient via DNA metabarcoding analyses of their feces. The results are expected to show that:

- The diet species richness decreases along the rural-urban gradient, a greater species diversity being observed in rural areas.
- Important differences within diet species composition of urban and rural areas exist and signal a change in the communities' structure caused by urbanization.
- Some variables influenced by the urbanization such as the degree of fragmentation or the temperature play an influence on these communities, reducing both their species richness and dissimilarities.

The analyses carried out for this thesis aims at accepting or refuting the assumptions formulated above. The primary investigation consists in determining insects preyed upon by *Hirundo rustica* (Linnaeus, 1758) in the different sampling sites. The alpha and the beta diversity of these ingested communities are then measured. The last tests estimate the influence of the urban-rural gradient mentioned earlier and environmental variables on these insect communities' composition.

The target of this study is thus to characterize the impact of urbanization on the foraging preferences of *Hirundo rustica* (Linnaeus, 1758). By better comprehending how urban expansion influences trophic relationships and biodiversity, its underlying purpose is to act as a decision-making tool for the environmental-related projects in territories affected by the urban sprawl.

2. Material and methods

2.1. Study area

This study takes place in the Kanto region in Japan where 75 breeding sites have been sampled in June and July of 2020. These sites were selected through the website of the Wild Bird Society of Japan (https://www.wbsj.org/en/). The study area covers a mosaic of natural, semi-natural and anthropic environments including megalopolis, towns, agricultural settings, forests, wetlands, and rivers. The sites are spread over seven areas (Nishitokyo, 23 cities, Ichikawa, Matsudo, Abiko, Chiba, Isumi), located in Chiba and Tokyo prefectures) and extend from Tokyo city center to rural areas (Figure 5). At the center of each of these sites can be found *Hirundo rustica* (Linnaeus, 1758) nests under which drops have been collected. As *Hirundo rustica* (Linnaeus, 1758) specimens are expected to forage in a perimeter of about 400-500 m around their nest (Snapp 1976; Evans, Wilson, and Bradbury 2007; Savignac 2011; Kang and Kaller 2013), insects present in their feces are expected to be present and preyed in this perimeter as well. This is why the studied sites are represented by circles of 500 m radius and cover a surface of 785,000 m².

Figure 5 - Sampling sites (represented by dots) for the years 2020. Map built on Google Earth Pro.

2.2. Data collection

The experimental set-up consisted in paper sheets sticked under the occupied nests from which the entirety of the fresh nesting fecal matters was collected after 24 hours. Ichikawa, Matsudo and Abiko areas were sampled by a different person than Nishitokyo, 23 cities, Chiba and Isumi. Some of the breeding sites were sampled up to five times (at different dates or targeting other nests at the same location) in 2020 resulting in 98 collected samples. They were then stored at -20°C until analyses.

2.3. DNA analyses

2.3.1. DNA extraction

The DNA present in these 98 samples was subsequently extracted. Beforehand, the uric acid and the fecal membrane were removed, and the remaining fecal matter was lyophilized with a lyophilizer freeze dryer VD-250R (TAITEC, Koshigaya, Saitama, Japan). The samples cleansed of uric acid and fecal membranes were sent to a lab to study them via metabarcoding. Therefore, the operations from the homogenization to the sequencing were conducted within the bioengineering lab of the Bioscience Technology Research company.

The samples content was then homogenized during two minutes at 1,500 rpm thanks to a ShakeMaster NEO homogenizer (bms, Shinjuku, Tokyo, Japan), diluted with a Lysis Solution F (Nippongene, chiyodaku, Tokyo, Japan), incubated at 65°C for ten minutes and then centrifuged at 12,000 x g for two minutes. The DNA extraction itself was carried out on the supernatants collected from the samples with the MPure Bacterial DNA Extraction Kit (MP Biomedicals, Irvine, CA, USA) following its guideline. The MPure-12 Automated Nucleic Acid Purification System (MP Biomedicals, Irvine, CA, USA) was subsequently utilized to purify the DNA of the samples. QuantiFluor dsDNA System (Promega, Madison, WI, USA) and Synergy LX (BioTek, Winooski, VT, USA) were used to measure the DNA concentrations.

2.3.2. DNA replication

The next step consisted in the replication of a specific DNA fragment from the Folmer region (Folmer et al. 1994) of the mitochondrial gene cytochrome c oxidase subunit 1 (COI) which is 658 base pairs long. More specifically, the ZBJ-Art primers were used in order to amplify a 211 base pairs fragment (including primers) of this region (Ribeiro, Smit, and Gilbert 2019). The COI gene is broadly used for animal species identification since it displays important nucleotide variations between species (Rennstam Rubbmark et al. 2018). It is therefore predominantly represented in sequences reference databases (Scasta et al. 2019). The same statement can be made for the Folmer region of the COI which is largely employed for insects identification (Elbrecht and Leese 2017; Ribeiro, Smit, and Gilbert 2019). This amplification process is particularly important to detect species whose DNA is present at low abundance in the environmental samples (in this case in feces) (Kelly et al. 2014).

For this study, the targeted DNA fragments were replicated by a two-step tailed polymerase chain reaction (PCR) with the purpose of creating a sequences library. The foremost PCR amplification was carried out with the ZBJ-ArtF1c and ZBJ-ArtR2c primers (AGATATTGGAACWTTATATTTTATTTTTGG and WACTAATCAATTWCCAAATCCTCC), respectively the Universal Forward and the Universal Reverse primers (Rennstam Rubbmark et al. 2018), paired with Illumina index sequences and MiSeq-specific adapters. The utmost PCR amplification was realized with index primers (AATGATACGGCGACCACCGAGATCTACAC-Index2-ACACTCTTTCCCTACACGACGC and CAAGCAGAAGACGGCATACGAGAT-Index1-GTGACTGGAGTTCAGACGTGTG).

The first PCR reactions were conducted in a 10 µL solution composed of 5.0 µL of 2x Master Mix (QIAGEN, Hilden, Germany), 0.5 µL of the ZBJ primers at a 10 µM concentration, 1.5 µL of DNA template, and 2.5 µL of double-distilled water.

The first denaturation operation which allowed the break-up of the chemical bonds between the DNA strands and therefore their dissociation was carried out on the DNA fragments at 95°C during 15 minutes. 35 PCR cycles were then realized. The fragments underwent another denaturation procedure at 94°C during 30 seconds. The subsequent step, the annealing, is the fixation of the primers on their complementary sequences and requires a temperature drop. The primers were hybridized with the single stranded DNA at a temperature of 45°C during 45 seconds. Finally, the extension of the new fragments was enabled at a temperature set at 72°C during two minutes followed by a final elongation of ten minutes at 72°C. The double purification of these PCR products was realized with the AMPure XP (Beckman Coulter, Brea, CA, USA) before the second PCR.

The second PCR reactions were conducted in a 10 µL solution composed of 1.0 µL of 10x Ex Buffer, 0.8 µL of nucleoside triphosphate dNTPs (each at 2.5 mM), 0.5 µL of the index primers at a 10 µM concentration, 2.0 µL of the first PCR product, 0.1 µL of DNA polymerase ExTaq HS at 5 U/µL (TaKaRa, Otsu, Shiga, Japan) and 5.1 µL of double-distilled water.

The steps realized through this PCR were a first denaturation at 94°C for two minutes, followed by 10 or 12 cycles of denaturation at 94°C for 30 seconds, annealing at 60°C for 30 seconds, extension at 72°C for 30 seconds, and then a last elongation at 72°C for five minutes.

The double purification of these PCR products was realized with the AMPure XP (Beckman Coulter, Brea, CA, USA). A QuantiFluor dsDNA System (Promega) and a Synergy H1 microplate reader (BioTek, Winooski, VT, USA) were then used to evaluate library concentrations while an assessment of its quality was provided by a fragment analyser (Advanced Analytical Technologies, Ankeny, IA, USA) with a dsDNA 915 Reagent Kit (Agilent, Santa Clara, CA, USA). Poor quality libraries were discarded, like the sample No.6.

2.3.3. Sequencing

Once the sequences are present in high enough density, the order of the nucleotides which make up the replicated DNA fragments needs to be determined. This process is called "sequencing" and was achieved by high-throughput sequencing (HTS) in this study. A 2×300 paired-end run on MiSeq Illumina technology (Illumina, San Diego, CA, USA) was carried on to do so.

2.3.4. Taxonomy assignment

After the sequencing operation, different procedures were used in order to curate the acquired sequences and to assign them a taxonomy. The first method aimed at assigning taxonomy from the National Center for Biotechnology Information (known as NCBI) databases by using the Basic Local Alignment Search Tool (abbreviated as BLAST). The second relied on the same tool (BLAST) but on a custom database made up of sequences from the Barcode of Life Data Systems (alias BOLD). The performances obtained through these two methods were investigated and the results compared afterwards.

Basic Local Alignment Search Tool on a BOLD modified database

This method enables the control of the treatments applied to the sequences obtained from the PCR amplification and the sequences from the BOLD database.

The sequences obtained from the HTS process were primarily treated on QIIME2 (version 2022.2), a bioinformatics platform originally developed for microbiome analyses (Bolyen et al. 2019). For each sampling sites was provided a forward and a reverse reads file. On average, the samples were composed of about 71,500 sequence reads and thus covered an important amount of organisms. The minimum read counts was higher than 15,000 and didn't require any filtering (low-frequency sequences often results from sequencing errors or contamination during the PCR and require filtering) (Drake et al. 2022). Once imported on QIIME2, several operations were carried out. In order to improve the sequences quality, the sequences were cut to the 211 first bases (length of the targeted amplicon) with the DADA2 plugin (Callahan et al. 2016). Past this length, the sequence quality drastically drops under a score of 20 which represents a 1% error rate. This was visualized thanks to the QIIME2 demux plugin (QIIME 2 development team 2022), which additionally provided information on the number of sequences per sample.

The ZBJ-primers were then trimmed from the read sequences with the DADA2 plugin again. Primers removal is important to avoid the false positive detection of chimeras, which are removed from the dataset thanks to this plugin as well (Callahan et al. 2016). Chimeras are sequences made up of several biological sequences (Ashelford et al. 2006). Consequently and since primers vary very little or are identical between sequences, amplicon with primers could be detected as such (the primers region of a sequence being wrongly considered as part of another sequence) (Jiménez 2021). In addition to the primers removal and the chimeras deletion, the reverse and forward reads were merged to form single sequences. Eventually, the DADA2 plugin identified and counted the abundance of amplicon sequence variants (ASVs), sequences observed in the dataset which show low probably of being arose from sequencing errors or other artifacts (variations from technical processes) (Callahan, McMurdie, and Holmes 2017). These ASVs correspond to the sequences which will eventually be associated to a taxonomy through a reference database. After this filtering, an average of about 63,000 sequences per sample were retained.

These treatments enabled the obtention of a table associating ASVs to the samples they were observed in, the ASVs sequences and some statistics on the carried-out operations. To assign a taxonomy to these ASVs, a custom database composed of sequences from the BOLD database was built.

The sequences from insect records from Japan were downloaded from the public data portal of the BOLD systems (https://boldsystems.org/index.php/Public_SearchTerms). Around 31,500 sequences were acquired with this operation. These sequences were then filtered on QIIME2 to solely retain the approximatively 26,400 records covering the COI-5P portion. Gaps as well as leading and trailing ambiguous bases were subsequently removed from these DNA fragments thanks to the RESCRIPt plugin (Ii et al. 2021). The RESCRIPt plugin was used for all the following operations except for the primers trimming. The next filtering process consisted in removing sequences with five or more ambiguous nucleotides and the ones containing homopolymers of nine or more nucleotides. The around 25,700 remaining sequences were then dereplicated, a process which implies the merging of identical sequences associated to different level of taxonomical description (Edgar n.d.). At this stage, slightly less than 14,500 sequences were still retained. These DNA fragments were imported on the Geneious Prime software (version 2022.1.1 Biomatters Ltd) (Kearse et al. 2012) where the primers got trimmed. Back on QIIME2, the sequences were aligned, the gaps generated during the alignment were removed and fragments of length lower than 126 bp or higher than 157 bp were removed. Around 7,000 sequences were 157 bp long which represents around 50% of the dataset. This corresponds to the expected post-trimming length. Keeping the sequences of length between 126 bp and 157 bp instead of just the sequences of 157 bp allowed to conserve 1,800 additional ones. Despite their lengths, these shorter sequences could still contain valuable information. 126 bp is therefore a trade-off between keeping long-enough sequences (nearly complete) and including a high number of sequences in the classifier (about 60% of the original sequences). The final step in the construction of the custom database was a second dereplication of the newly trimmed sequences.

The obtained classifier made up of 5,170 sequences was then utilized to assign taxonomies to the ASVs detected in the feces through BLAST. This algorithm looks for similar regions within sequences and performs an alignment based on matching nucleotides subsequences. Alignment scores based on similarity matrix are calculated. A statistical value is then given to these alignment score and the taxonomy associated to the highest value is the one assigned to the query sequence (Wheeler and Bhagwat 2007). The QIIME2 feature-classifier plugin (Bokulich et al. 2018) was utilized with a 97%

similarity identity threshold to realize this operation. The sites No.3,13,22 and 25 did not show any results from these operations.

Basic Local Alignment Search Tool on NCBI databases

For this method, the read sequences whose start entirely matched the ZBJ-primers sequences were extracted using the FASTX-Toolkit software (version 0.0.14) (Gordon and Hannon 2010). The fastq_barcode_spliltter tool and the fastx_barcode_spliltter tool were utilized for the first set and the second set (Table 3), respectively. As the query sequences were sent at two different times, slightly different operations were applied to them.

Table 3 - Sets of samples that have received different treatments.

No. of set	Samples numbers
1	1, 2, 3, 5, 9, 13, 15, 16, 17, 21, 22, 23, 25, 29, 31, 32, 37, 38, 39, 41, 42, 44, 45, 50, 51, 54, 57, 63, 65, 68, 70, 72, 73, 75, 80, 84, 86, 87, 89, 92, 93, 96, 97
2	4, 7, 8, 10, 11, 12, 14, 18, 19, 20, 24, 26, 27, 28, 30, 33, 34, 35, 36, 40, 43, 46, 47, 48, 49, 52, 53, 55, 56, 58, 59, 60, 61, 62, 64, 66, 67, 69, 71, 74, 76, 77, 78, 79, 81, 82, 83, 85, 88, 90, 91, 94, 95, 98

In the first set, the primer sequences were removed, 120 bp located at the 3'end were trimmed and the chimeric as well as the noise sequences discarded thanks to the DADA2 plugin (Callahan et al. 2016) on QIIME2 (version 2020.8). Pair-end reads were then bound with the same plug-in. Representative sequences and ASV tables were obtained from these operations.

In the second set, the same operations as for QIIME2 (version 2022.2) were applied but the primers and 120 bp at the 3' end were cut off with the fastx_trimer of FASTX-Toolkit software (version 0.0.14). Then, using the Sickle biotool (version 1.33) (Joshi and Fass 2011), the 3' end with a quality value of less than 20 were trimmed, and sequences with a length of less than 40 bases and their paired sequences were discarded. Pair reads with overlap of minimum ten bases were subsequently merged with the FLASH software (version 1.2.11) (Magoc and Salzberg 2011). Representative sequences and ASV tables were obtained after removing chimeric and noise sequences with the DADA2 plugin of Qiime2 (version 2022.2). Other parameters were performed under standard conditions.

Taxonomies from the NCBI nucleotides databases (version 2.9.0 or 2.12.0) were assigned to these ASV sequences using the Standard Nucleotide BLAST under standard conditions. The identification with less than 97% identity were discarded. The sites No.3,13,22 and 25 did not show any results from these operations.

2.4. Biodiversity analysis

For the purpose of characterizing *Hirundo rustica* (Linnaeus, 1758) diet along the urban-rural gradient and the influence of variables on it, both α and β diversity metrics were calculated based on the BLAST on the NCBI databases results. The α diversity is defined as the amount of diversity observed within a designated sample while the β diversity is described as the comparison (similarities and dissimilarities) of diversity observed within designated samples (Kessler et al. 2009). In this case, the α diversity studies the diet richness of *Hirundo rustica* (Linnaeus, 1758) while the β investigates its composition.

ASVs with the same taxonomy were grouped together and the unique taxa were retained. The Chao2 and the Shannon diversity indices were then calculated on R thanks to the microbiome package (Lahti et al. 2017) along with the observed taxa richness. The observed taxa richness was simply obtained by counting the number of unique taxa within each site. The Chao2 index takes into account the incidence of the taxa in order to estimate the species richness (1) (Vavrek 2010).

$$S_{est} = S_{obs} + \frac{F^2}{2G} \tag{1}$$

Where:

S_{est} = Estimated taxa richness

S_{obs} = Observed taxa richness

F = Singletons (taxa observed in one sampling site)

G = Doubletons (taxa observed in two sampling sites)

The Shannon-Wiener index is a metric which increases with the richness and evenness of ASVs on the sampling sites (2). A low Shannon diversity index represent a low number of taxa and an unequal repartition of these taxa biomasses (Elbrecht and Leese 2015). It is therefore influenced by rare taxa characterized by low reads counts even though a larger weight is allocated to the taxa richness (Kim et al. 2017).

$$H = -\sum_{i=1}^{S_{obs}} p_i * \ln p_i \tag{2}$$

Where:

S_{obs} = Observed species richness

p_i = Proportion taken by the taxa i in the sampling site

The beta diversity index was subsequently assessed based on distances between sites communities calculated through the Jaccard similarity index. This was realized thanks to the vegan R package (Oksanen et al. 2022). This similarity index relies on presence/absence data (3). The sites 58 and 90 were discarded prior the beta diversity assessment since no taxa were identified for them.

$$J = \frac{c}{(a+b) - c} \tag{3}$$

Where:

a = Number of taxa present in site 1

b = Number of taxa present in site 2

c = Number of taxa present in sites 1 and 2

2.5. Landscape analysis

2.5.1. Urban-rural gradient

The primary goal of this thesis being to measure the impact of an urban-rural gradient on *Hirundo rustica* (Linnaeus, 1758) diet, this gradient had to be characterized. Several ways of measuring these gradients have been utilized in scientific researches, relying on different variables (vegetation cover, impervious surfaces, population density, traffic network density,…) or a combination of several of them. Obtaining a consensus on which of these gradients are the best represented is although complex since the representation accuracy greatly depends on the spatial scale of the studies (Suarez-Rubio and Krenn 2018). However, in many cases, the percentage of impervious surfaces describes the state of urbanization with enough accuracy (Fortel et al. 2014; Geslin et al. 2016; Choate, Hickman, and Moretti 2018; de Andrade 2020; Nason et al. 2021). This measure is the one being used in this study. The percentage of impervious surface has been calculated thanks to a high-resolution (2.5 m) satellite image built from the April 2021-July 2021 period. This image results from the fusion of Sentinel-2 bands (10 meters and 20 meters resolution) and images from the PlanetScope satellite constellation (higher resolution). Once acquired, spatial analyses could be realized. Land covers can be discriminated in four categories: bare soils, vegetation cover, water surfaces and impervious surfaces (low albedo and high albedo ones) (Van de Voorde, De Roeck, and Canters 2009; Dong et al. 2021). Therefore, for this study, the bare soils, vegetation, and water surfaces will be determined, and the remaining areas will be considered as impervious surfaces.

Vegetation cover

In order to distinguish the vegetation from artificial material, bare soil and water, the normalized difference vegetation index (NDVI) was calculated (Roces-Díaz et al. 2020). This index results from a calculation based on the near infrared and red light which are respectively greatly reflected and considerably absorbed by vegetation (Pettorelli et al. 2005). The red and near infrared bands from the satellite image are therefore the ones utilized to determine the NDVI values (4).

$$NDVI = \frac{NIR - Red}{NIR + Red} \tag{4}$$

Where:

NIR = Reflectance values of the near infrared band

Red = Reflectance values of the red band

The values obtained from this calculation ranges from -1 to 1, low values corresponding to an absence of vegetation and high values to an important vegetation biomass. A threshold of 0.3 was selected in order to discriminate vegetated areas from non-vegetated surfaces. While NDVI values lower or equal to 0.3 were associated with impervious surfaces or waterways, values superior to 0.3 were associated to vegetated areas (Guo, Lu, and Kuang 2017). In order to smooth the results, a three pixels radius majority filter from the SAGA GIS module was applied in the QGIS 3.10.10 software (Senay and Elliott 2000; Passy and Thery 2018). This filter allowed the elimination of potential isolated pixels wrongly considered as vegetation. These operations enabled the estimation of the percentage of vegetation cover in each sampling site.

Water bodies

The 2.5 m resolution satellite image was also utilized to discriminate water from other surfaces. In order to detect water points and since water and non-water surfaces reflect near-infrared light differently, the reflectance values of the near infrared band of the image were explored. After examination, pixels showing reflectance values comprised between 0.100 and 0.650 were classified as water bodies (5). The utilization of the modified normalized difference water index (MNDWI) was also studied but showed reduced performance to discriminate water and buildings (Mondejar and Tongco 2019). The three pixels radius majority filter has been applied to the determined water bodies as well. The percentage of water cover in each sampling site was calculated afterwards.

$$0.100 < NIR < 0.650 \tag{5}$$

Bare soils

Bare soils have been differentiated from the other cover types using the same high-resolution image. This class of surfaces is particularly complex to discriminate from impervious materials of built-up areas. This distinction is achieved in a way which is similar to the one employed in order to spot vegetated areas. The process relies on the bare soil reflectance of the shortwave infrared and the green light which is greater than the reflectance values of the built-up areas in this part of the spectrum. To this purpose, the dry bare-soil index (DBSI), which relies on the NDVI and on these SWIR1 and green bands, was calculated (6).

$$DBSI = \frac{SWIR1 - Green}{SWIR1 + Green} - NDVI \tag{6}$$

Where:

$SWIR1$ = Reflectance values of the short-Wave Infrared band 1

$Green$ = Reflectance values of the green band

$NDVI$ = Normalized difference vegetation index

The values obtained from this calculation ranges from -2 to 2, low values corresponding to artificial materials, water or vegetation and higher values representing bare soils. The threshold of 0.26 was picked in order to differentiate bare soils from other land covers, pixel values inferior to 0.26 being associated to other land covers and values superior or equal to 0.26 being considered as bare soil values (Rasul et al. 2018). The three pixels radius majority filter was applied to these estimated bare soil surfaces. The percentages of land cover corresponding to bare soil of each sampling area was then measured.

Impervious surfaces

Finally, as land covers are classified into bare soils, vegetation cover, water surfaces and impervious surfaces (low albedo and high albedo ones) (Dong et al. 2021), the impervious surfaces have been associated with the pixels which were not classified as vegetation, water or bare soil (Figure 6). This allowed the obtention of the percentage of impervious surfaces and therefore of urban areas in each sampling site as well as the construction of the urban-rural gradient. Sites covered by an important amount of impervious surfaces are obviously considered more urbanized than sites with low percentages of impervious surfaces cover.

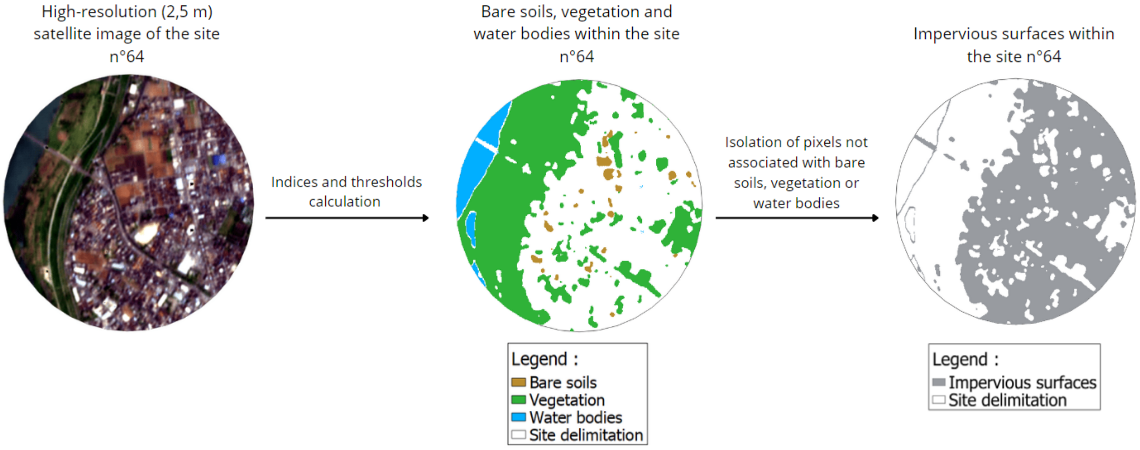

Figure 6 - Process employed to determine areas covered by impervious surfaces.

2.5.2. Climatic variables

Arthropods are ectothermic organisms which means that they are unable to produce their own heat and depends on external conditions to regulate their body temperature. Climatic conditions such as air temperature and precipitation levels have therefore an influence on insects (McIntyre et al. 2001; Zaller et al. 2014; Youngsteadt et al. 2017) and consequently on *Hirundo rustica* (Linnaeus, 1758). Weather events are also believed to be responsible for mortality episodes within migratory birds populations (García-Pérez et al. 2014). The effects of climatic variables have been studied in this work.

Precipitation and temperature data from June 2020 and July 2020 have been obtained from weather stations of the Japan Meteorological Agency (https://www.data.jma.go.jp/risk/obsdl/index.php). These stations are located in the prefectures of Chiba, Tokyo, Kanagawa, Ibaraki and Saitama. For the temperature, daily data from 31 weather stations have been compiled and interpolated. As for the precipitation, daily data from 41 weather stations have been utilized. The interpolation has been realized through the Inverse Distance Weighting (IDW) technique on QGIS 3.10.10 software (Ilayaraja and Ambica 2015). For each site, the average temperature, the minimum daily temperature, the maximum daily temperature, and the range of these temperatures have been calculated in degrees Celsius. The same computations have been applied for the average monthly and maximum daily precipitation which were measured in millimeters. The minimum daily and the range of precipitation were not utilized since the first of these two variables equals zero for every site.

Even though temperatures and precipitation are known to be influenced by urbanization, these factors depend greatly on the geographical position of the targeted area. This means that the main drivers of the climate variability among the stations and therefore among the data are also linked to spatial and geographical elements (Wood n.d.) and not only to the degree of urbanization. These elements need to be considered while studying the impact of these climatic conditions on *Hirundo rustica* (Linnaeus, 1758) diet and therefore on the insect communities.

Light pollution

Even though natural sources of light are omnipresent and participate in the fauna circadian rhythm, anthropogenic night lighting heightens night sky luminance and disrupts biological cycles and motion behaviors. Insects are particularly sensitive to this variable and this factor's influence is considered in this research (Eisenbeis and Hänel 2009; Falchi et al. 2016). The light pollution data were obtained through an artificial night sky radiance map created by the Light Pollution Science and Technology Institute and made available on the Microsoft Bing Maps Platform (https://www.lightpollutionmap.info/). This map relies on the visible Infrared Imaging Radiometer Suite (VIIRS) data obtained by the Suomi NPP and NOAA-20 satellites for the year 2021. It displays a resolution of 15 arc seconds (around 380 m in Tokyo) and expresses the radiance in 10^{-9} watt per steradian per square centimeter (Berrick 2022).

2.5.3. Landscape metrics

In addition to environmental factors, several studies highlight the effects of landscape characteristics on insect communities (Schindler et al. 2013; Su et al. 2015; Withaningsih and Rabbany 2019; González-Césped et al. 2021) and thus on *Hirundo rustica* (Linnaeus, 1758) alimentation. These characteristics are closely connected to urbanization (Kaminski et al. 2021) and studying them enables to figure out which variables are the most critical for biodiversity. Landscape metrics have therefore been calculated for each of the sampling sites. These indices allows to characterize the landscape and more specifically the spatial arrangement of vegetation patches (Jia et al. 2019). The vegetated areas determined thanks to the normalized difference vegetation index (NDVI) have been employed to compute these indexes (Rahimi, Barghjelveh, and Dong 2021).

Proportion of land covers

These indices have already been calculated in order to build the urban-rural gradient as explained in the section “2.4.1. Land covers”.

Patch size

Insects’ populations and communities seem to be influenced by the habitat patches size. More specifically, this factor has displayed a negative correlation with the insects species richness, abundance and density distribution (Fahrig and Jonsen 1998; Krawchuk and Taylor 2003; Bukovinszky et al. 2010; Burkman and Gardiner 2014). In order to study its effect, the mean patch size (MPS), expressed in square meters, has been calculated (7).

$$MPS = \frac{\sum_{i=1}^{n} a_i}{n} \tag{7}$$

Where:

a_i = area of vegetation patch i [m²]

n = amount of vegetation patches [m²]

Landscape fragmentation

Landscape fragmentation displays an influence as well on insect communities, according to some studies (Fahrig and Jonsen 1998; Krawchuk and Taylor 2003; Burkman and Gardiner 2014; Withaningsih and Rabbany 2019). For this reason, the largest patch index (LPI), which can be linked to the state of fragmentation (Jia et al. 2019; Withaningsih and Rabbany 2019) has been assessed (8). The value of this metric is comprised between 0 and 1, the highly fragmented areas showing a LPI value close to 1 (Saura and Martinez-Mlian 2001).

$$LPI = \frac{a_{max}}{A_{tot}} \tag{8}$$

Where:

a_{max} = area of the biggest vegetation patch [m²]

A_{tot} = landscape surface [m²]

Landscape heterogeneity

The landscape heterogeneity is also known to be correlated to the species richness and its effect has been studied during this work. The Shannon diversity index (SHDI) has been calculated (9) in order to represent this heterogeneity (Withaningsih and Rabbany 2019). The SHDI values will be higher for sampling sites with different kind of land covers (richness criteria) and within sites where land covers proportions are better balanced (evenness criteria).

$$SHDI = -\sum_{i=1}^{s} p_i \times \ln p_i \tag{9}$$

Where:

s = number of land cover categories in presence

p_i = proportion of land cover i

2.6. Analysis of the relationship between variables and diversity

The relationship between the environmental variables and diversity index estimated beforehand was then tested. As some variables are strongly correlated (Figure 7), a variables selection has been carried out in order to avoid further issues due to collinearity. For each combination of environmental variables showing absolute correlation values of 0.70 or more (Dormann et al. 2013), only one was retained. This way, 11 variables got selected: surface of water bodies, surface of bare soils, mean patch size (MPS), largest patch index (LPI), Shannon's diversity index (SHDI), artificial night sky radiance, average temperature, minimum daily temperature, range of temperatures, average monthly precipitation, and maximum daily precipitation. As the Chao2 and the observed taxa richness displayed a strong relationship as well (Figure 7), the relationship between the selected variables and the Chao2 index was not investigated.

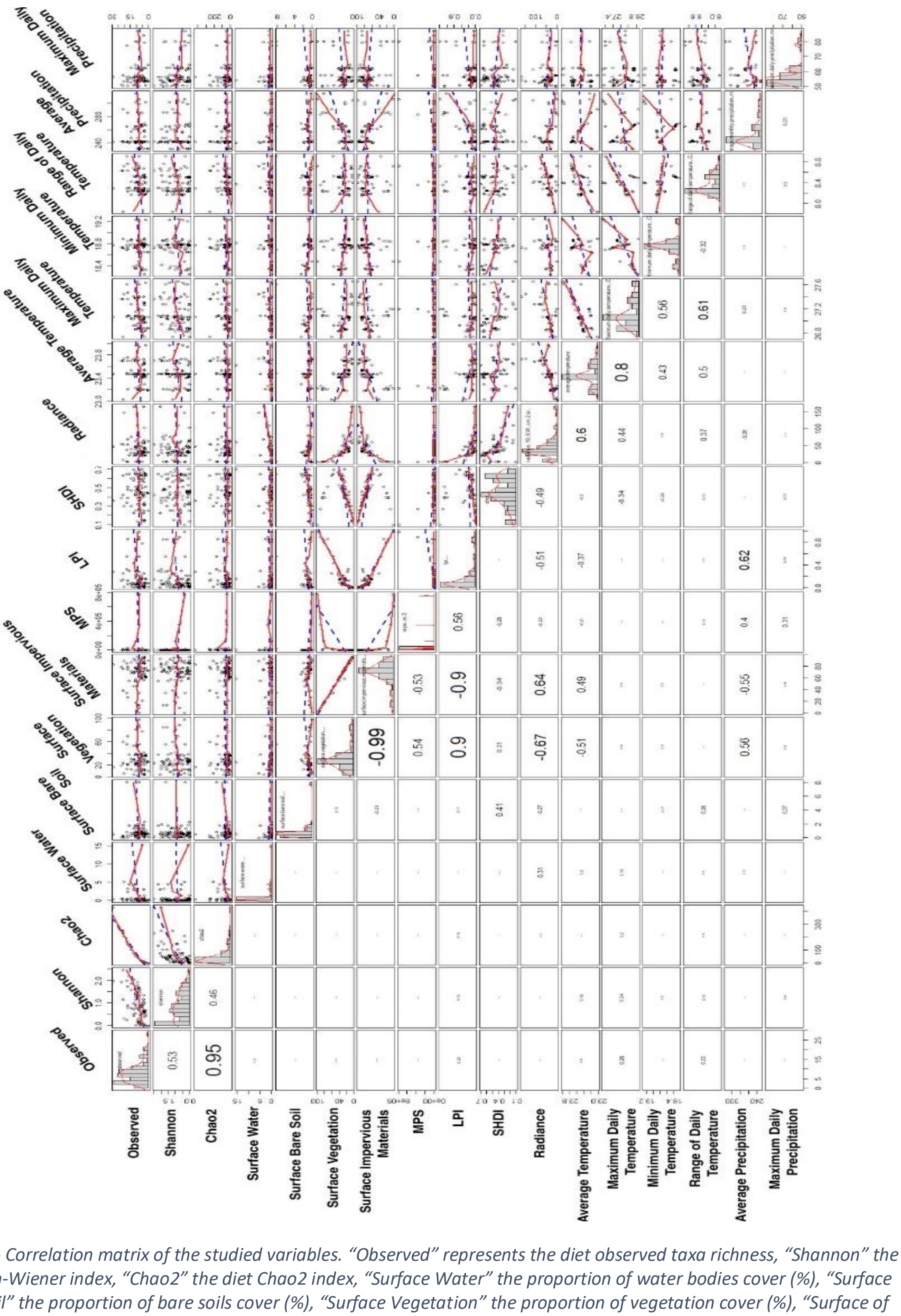

*Figure 7 - Correlation matrix of the studied variables. “Observed” represents the diet observed taxa richness, “Shannon” the Shannon-Wiener index, “Chao2” the diet Chao2 index, “Surface Water” the proportion of water bodies cover (%), “Surface Bare Soil” the proportion of bare soils cover (%), “Surface Vegetation” the proportion of vegetation cover (%), “Surface of Impervious Materials” the proportion of impervious surfaces cover (%), “MPS” the mean patch size (m²), “LPI” the largest patch index, “SHDI” the Shannon diversity index, “Radiance” the artificial night sky radiance (10^{-9} W/cm²*sr), “Average Temperature” the average monthly temperature (°C), “Maximum Daily Temperature” the maximum daily temperature (°C), “Minimum Daily Temperature” the minimum daily temperature (°C), “Range of Daily Temperature” the range of temperatures (°C), “Average Precipitation” the average monthly precipitation (mm) and “Maximum Daily precipitation” the maximum daily precipitation (mm).*

The two other alpha diversity metrics were therefore studied with the R package lme4 (Bates et al. 2022) through the construction of generalized linear mixed models (GLMM). The null hypothesis was indentical for both Shannon-Wiener diversity and the observed taxa richness: there is no statistical relationship between the predictors and the response variables. To test this hypothesis, several models were created starting by the saturated one (containing the 11 previously cited variables). New models were then generated by discarding the less significant variable of the previous one. The model selection was subsequently conducted based on the Akaike information criterion (AIC), the model with the lowest AIC being retained (Sugiura and Yamazaki 2014).

The model utilized for the observed taxa richness was a GLMM fitted with a Negative binomial distribution (logit link function) and built on three predictors (largest patch index (LPI), Shannon's diversity index (SHDI) and radiance). Random effects related to the areas containing the sites were considered by the model.

The model utilized for the Shannon-Wiener diversity was a GLMM fitted with a Gamma distribution (negative inverse link function) and built on four predictors (MPS, LPI, average monthly precipitation and average monthly temperature). A value of 1 was added to the calculated Shannon-Wiener index for the purpose of having higher than 0 values. Random effects related to the areas containing the sites were considered by the model.

The beta diversity was investigated with the vegan R package (Oksanen et al. 2022) computing a permutational multivariate analysis of variance using the distance matrice. This method relies on a permutation test with pseudo-F ratios ('Adonis: Permutational Multivariate Analysis of Variance Using... in Vegan: Community Ecology Package' 2022).

3. Results

3.1. DNA analyses

The results obtained from the two different methods showed important differences. The filtering, denoising and clustering of the sequences allowed the obtention of 1,430 different ASVs for the procedure applied prior the NCBI-based taxonomy assignment and 1,858 ASVs for the method preceding the custom BOLD database creation**.**

The BLAST applied to the custom BOLD database permitted the identification of 95 different insect species from the 1858 detected ASVs. The counterpart of this method is that a taxonomy was assigned to just about 10% of the ASVs observed in the excrements (Table 4). From these 95 species, 26 were not identified as top identity with the NCBI-based method. However, 19 of those were identified by this last method as top 2 to top 10 identity.

As previously said, the method aiming at assigning taxonomy through the BLAST on the NCBI databases enabled the detection of less taxonomic units. However, out of these 1,430 ASVs, 721 were classified as insects and 272 different insect species were assigned to these sequences (Table 4). This means that a little more than 50% of the ASVs were associated to taxa from the Insecta class. Organisms from different kingdoms such as Bacteria, Chromista, Plantae as well as non-Arthropoda animals (Mollusca, Nematoda and Rotifera) were detected as well and respectively accounted for 4.23%, 0.87%, 0.37%, 0.12%, 0.37% and 0.62% of the identified ASVs. Three other classes from the Arthropoda phylum, Arachnida, Entognatha and Malacostraca were also present and responsible for respectively 2.74%, 0.37% and 0.50% of the identified ASVs.

Table 4 - Summary of the performances of the two methods.

	NCBI databases method	Custom BOLD database method
Number of ASVs	1,430	1,858
Number of ASVs identified	803	198
Number of ASVs identified as insects	721	198
Number of insect ASVs identified to the species level	544	158
Number of insect species identified	272	95

The last mentionned method therefore enables the identification of around 50% of the sequences and outperforms the BLAST on custom BOLD database method on that matter. Both nevertheless identified orders in similar proportions. The vast majority of species whose *Hirundo rustica* (Linnaeus, 1758) feed on belong to the Diptera order which represent around 66% of the identified ASVs. The Coleoptera and Lepidoptera are the next most abundant orders detected in the droppings. According to the BLAST on NCBI method, Coleoptera represent a more important source of food than Lepidoptera (respectively representing 12.62% and 10.68% of the diet), while the opposite was observed from the BLAST on custom BOLD database results (respectively representing 10.60% and 13.64% of the diet). Eventually, Hemiptera and Hymenoptera were the last main orders commonly identified in similar proportions by both methods (approximatively 5% for the Hemiptera and 2% for the Hymenoptera order). Ephemeroptera, Trichoptera, Odonata, Orthoptera were identified in the two databases as well but separately made up less than 1% of the overall diet. The use of the NCBI databases however enabled the identification of four additional orders : Psocoptera, Raphidioptera, Thysanoptera and Neuroptera enabling the detection of 13 orders while the method relying on the BOLD curated database detected nine different orders (Figure 8).

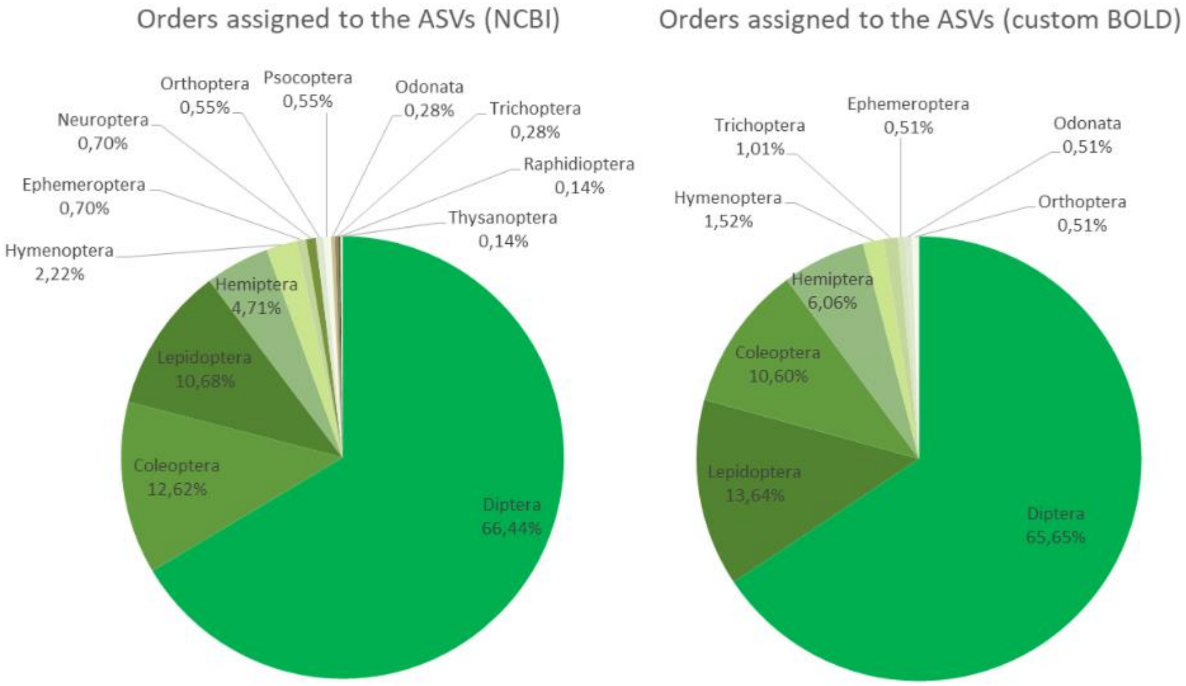

Figure 8 - Orders assigned to the ASVs by the BLAST on the NCBI databases method (on the left) and by the BLAST on the custom BOLD database method (on the right).

More important differences were observed between the families assignations. Even though both methods defined respectively Chironomidae and Drosophilidae as the main and secondary preys, the following most abundant preys belong to different families (Figure 9).

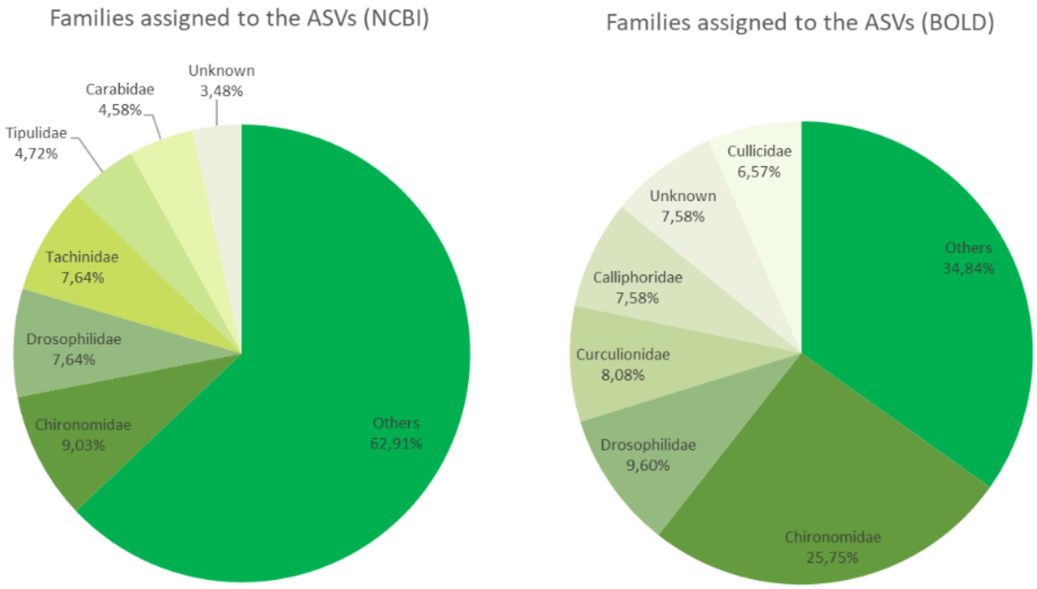

Figure 9 - Families assigned to the ASVs by the BLAST on the NCBI databases method (on the left) and by the BLAST on the custom BOLD database method (on the right). The "Others" category gathers all the families which are not among the six most abundant.

However, disregarding the relative abundances, the results were highly similar. Every 36 families identified through the curated BOLD database were present in the list of families obtained from the NCBI databases along with 70 additional identified families (Table 5).

Table 5 - List of families identified by the BLAST on NCBI method. The highlighted families are the ones commonly identified with the BLAST on the curated BOLD database.

Families and associated number of ASVs identified as such							
Chironomidae	65	Cecidomyiidae	5	Coccinellidae	2	Coreidae	1
Drosophilidae	55	Hydrophilidae	5	Coenagrionidae	2	Corylophidae	1
Tachinidae	55	Miridae	5	Cosmopterigidae	2	Cydnidae	1
Tipulidae	34	Psychodidae	5	Delphacidae	2	Diapriidae	1
Carabidae	33	Sciaridae	5	Geometridae	2	Empididae	1
Muscidae	28	Simuliidae	5	Gracillariidae	2	Ephemerellidae	1
Sarcophagidae	28	Chrysopidae	4	Heleomyzidae	2	Ephemeridae	1
Calliphoridae	23	Fanniidae	4	Hesperiidae	2	Ephydridae	1
Sepsidae	19	Psychidae	4	Hydropsychidae	2	Erirhinidae	1
Noctuidae	18	Rhiniidae	4	Lonchaeidae	2	Figitidae	1
Culicidae	17	Tephritidae	4	Notodontidae	2	Hemerobiidae	1
Curculionidae	15	Anoeciidae	3	Pentatomidae	2	Inocelliidae	1
Syrphidae	15	Braconidae	3	Phoridae	2	Lycaenidae	1
Crambidae	12	Chloropidae	3	Plutellidae	2	Monotomidae	1
Staphylinidae	11	Dermestidae	3	Pyralidae	2	Notonectidae	1
Stratiomyidae	11	Dytiscidae	3	Pyrrhocoridae	2	Nymphalidae	1
Mycetophilidae	10	Ectopsocidae	3	Scathophagidae	2	Phlaeothripinae	1
Sphaeroceridae	10	Lauxaniidae	3	Scatopsidae	2	Platystomatidae	1
Dolichopodidae	9	Limacodidae	3	Tetrigidae	2	Polymitarcyidae	1
Limoniidae	9	Lygaeidae	3	Tettigoniidae	2	Psocidae	1
Tortricidae	9	Oedemeridae	3	Tineidae	2	Rhopalidae	1
Erebidae	8	Tabanidae	3	Anthocoridae	1	Scarabaeidae	1
Ichneumonidae	8	Tingidae	3	Asilidae	1	Sphingidae	1
Anthomyiidae	7	Acanthosomatidae	2	Cerambycidae	1	Stathmopodidae	1
Nitidulidae	7	Anobiidae	2	Chalcididae	1	Tenebrionidae	1
Aphididae	6	Argidae	2	Chrysomelidae	1		
Ceratopogonidae	6	Baetidae	2	Cicadellidae	1		

3.2. Landscape analyses

The sampled sites display a large range of impervious surfaces cover proportion and thus represent a wide variety of urbanization degree. However, a tendency to impervious surfaces covers proportions higher than 50% can be observed (Figure 10).

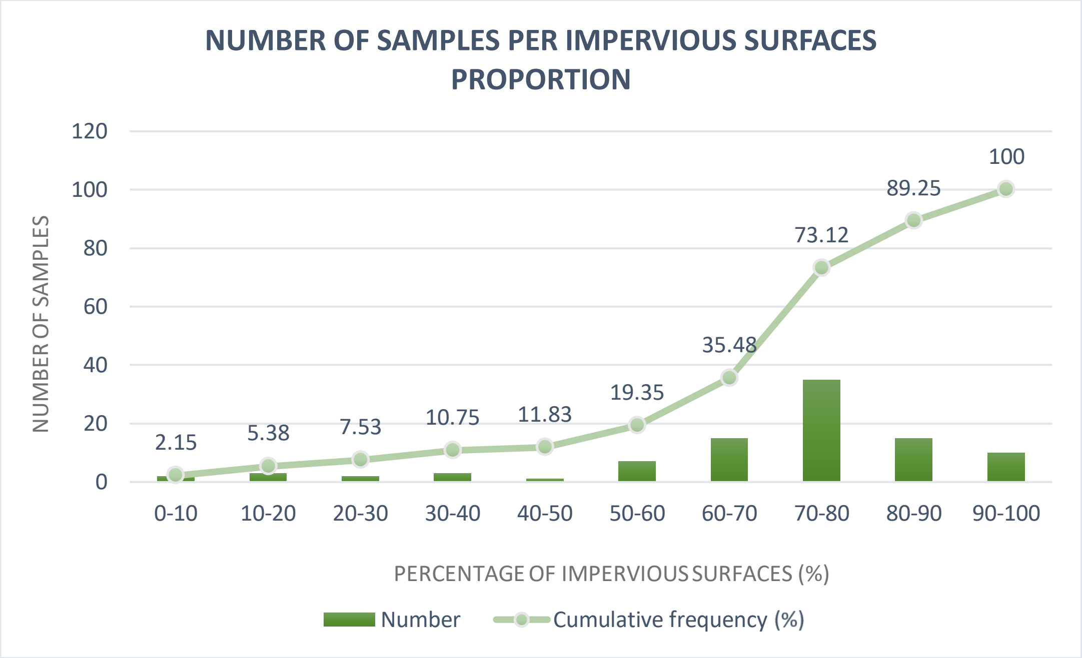

Figure 10 - Distribution and frequency of impervious surfaces cover proportion associated to samples.

Statistical analyses confirm this observation (Table 6). Indeed, despite ranging from 1.24% to 96.21% of impervious surfaces cover, the median of the data equals 72.57% and the interquartile range 19.89%. This shows that most of the data are clumped about a relatively high value of impervious surfaces proportion. The same can be observed with the mean value of 69.54% and its associated standard deviation of 19.93%. Urbanized areas are therefore better represented than rural sites.

Table 6 - Statistics describing the proportion of impervious surfaces cover of the 93 sampled sites.

Statistics	Values (%)
Mean	69.54
Median	72.57
Standard deviation	19.93
First quartile	62.72
Third quartile	82.61
Interquartile range	19.89
Minimum	1.24
Maximum	96.21

Following this logic and taking into consideration the preponderance impervious surfaces (strongly correlated to the vegetation) within the studied sites, a low discrimination of the sites and thus of the areas based on the land cover proportions can be observed (Figure 11). The climatic and other landscape variables studied for this research however allow a slightly better distinction of the areas and a more important differentiation of the sites based on their environmental conditions (Figure 12).

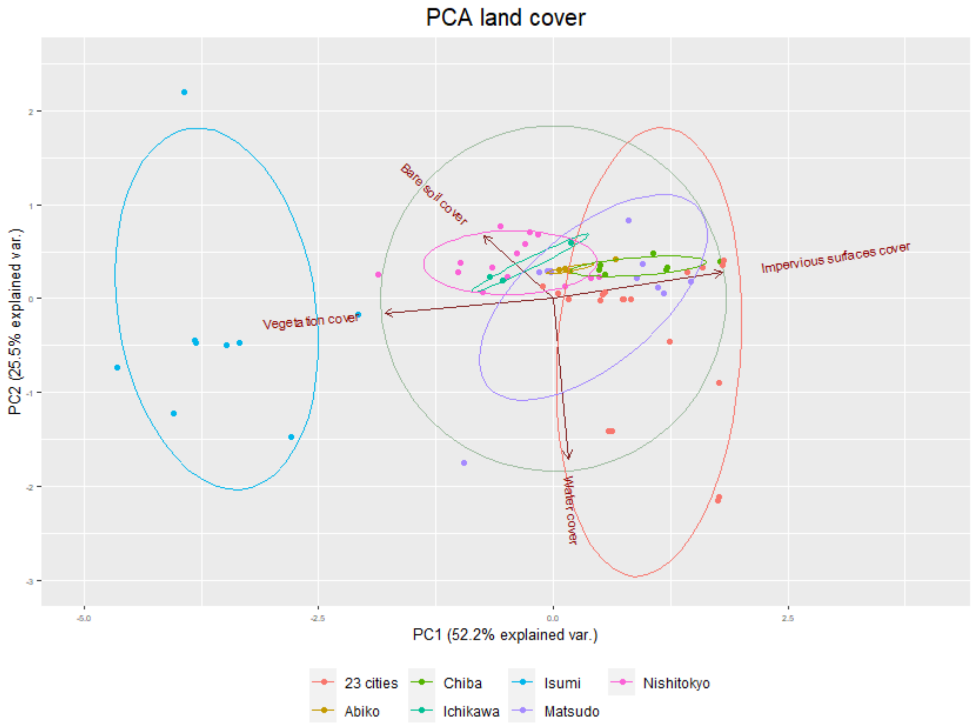

Figure 11 - Principal component analysis of the land cover proportion of the sampled sites. The different areas where the samples were collected are represented by the confidence ellipses while the correlation between variables and principal components is represented by the correlation circle.

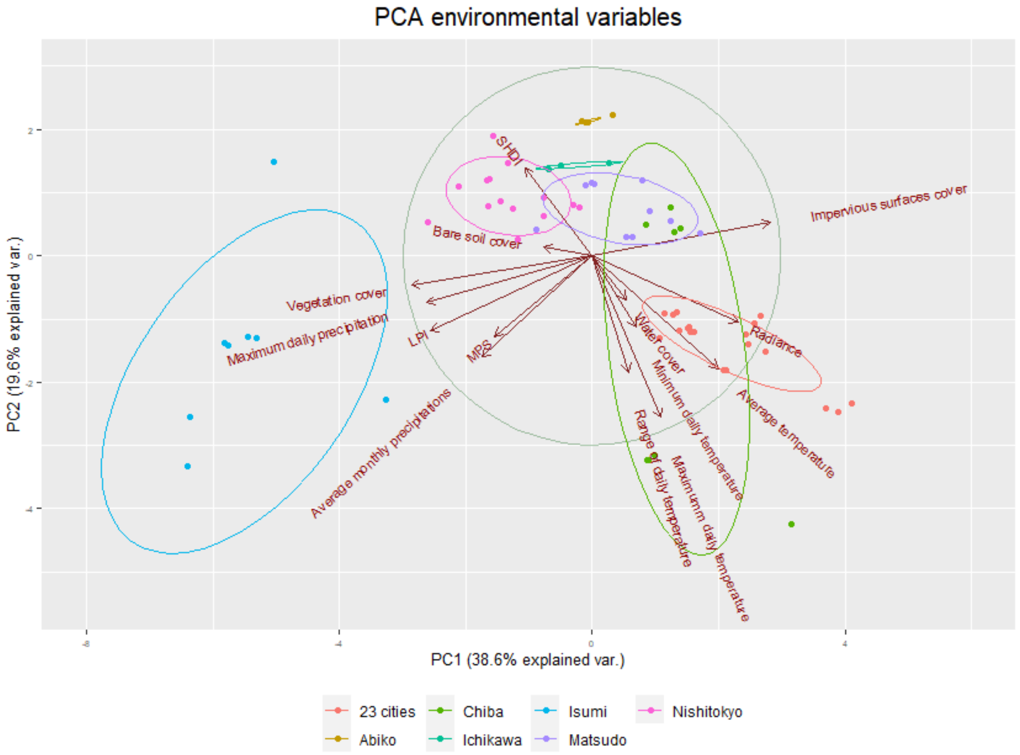

Figure 12 - Principal component analysis of the environmental and climatic variables of the sampled sites. The different areas where the samples were collected are represented by the confidence ellipses while the correlation between variables and principal components is represented by the correlation circle.

3.3. Analysis of the influence of variables on diversity

3.3.1. Alpha diversity

The generalized linear mixed models which aimed at confirming or refuting the null hypothesis that there is no statistical relationship between the selected variables and the alpha diversities showed different results for the Shannon-Wiener and the observed taxa richness diversity.

The GLMM built on the Shannon-Wiener diversity index showed statistical significance (p-values < 0.05 and |z-values| > 2) for the LPI, the average monthly precipitation and the average temperature. The three estimates were positive for these variables (Table 7).

Table 7 - Statistical parameters of the GLMM built on the Shannon-Wiener diversity index. The asterisks highlight the p-values inferior to 0.05.

Variables	Standard error	Estimate	Z-value	Pr(>|z|)
Intercept	0.0178	0.0182	30.440	< 2e-16 *
MPS	0.0205	0.0266	1.724	0.0847
LPI	0.0261	0.0269	-4.161	3.17e-05 *
Average monthly temperature	0.0320	0.0218	-3.514	0.000441 *
Average monthly precipitation	0.0281	0.0254	2.233	0.0255 *

The GLMM built on the observed taxa richness showed statistical significance (p-values < 0.05 and |z-values| > 2) for the LPI and the radiance. Both estimates were positive for these variables (Table 8).

Table 8 - Statistical parameters of the GLMM built on the observed taxa richness. The asterisks highlight the p-values inferior to 0.05.

Variables	Standard error	Estimate	Z-value	Pr(>|z|)
Intercept	2.0366	0.0982	20.740	< 2e-16 *
Radiance	0.206	0.103	2.005	0.0449 *
LPI	0.252	0.0884	2.849	0.00438 *
SHDI	0.151	0.0803	1.881	0.06

3.3.2. Beta diversity

The permutational analysis of variance studying the relationship between the selected variables and the beta diversity showed statistical significance as well (p-values < 0.05) for the average monthly temperature, the range of monthly temperatures and the average monthly temperature (Table 9).

Table 9 - Statistical parameters of the PERMANOVA conducted on the beta diversity index. The asterisks highlight the p-values inferior to 0.05.

Variables	Sum of squares	R^2	F	Pr (>F)
Surface water	0.469	0.0111	1.0167	0.416
Radiance	0.525	0.0124	1.138	0.123
Surface bare soil	0.480	0.0113	1.0408	0.344
MPS	0.492	0.0116	1.0678	0.260
LPI	0.501	0.0118	1.0875	0.255
SHDI	0.467	0.011	1.0129	0.428
Average monthly temperature	0.668	0.0157	1.449	0.005 *
Minimum daily temperature	0.566	0.0134	1.228	0.058
Range of daily temperature	0.598	0.0141	1.297	0.023 *
Average monthly precipitation	0.672	0.0158	1.457	0.002 *
Maximum daily precipitation	0.468	0.011	1.0152	0.411

4. Discussion

4.1. DNA analyses

The main goal of this object was to study the diet of *Hirundo rustica* (Linnaeus, 1758) and its range along an urbanization gradient. To achieve this, the performances of a custom BOLD database have been compared to the utilization of complete NCBI databases. The utilization of a custom database solely composed of records of insects collected in Japan had for purpose to avoid post-identification filtering, reduce computing time and decrease the likelihood of including false-positives (identification of species which were not truly present on the field) within the data (Magoga et al. 2022).

The main difference obtained through the utilization of the complete and the curated database lies in the number of species and of ASVs identified. In the case of this master thesis, the BLAST computed on the complete NCBI databases identified the largest number of taxonomic units among the Insecta class with 272 species and 50% of the ASVs identified, while the BLAST computed on a custom BOLD database detected 95 insect species and assigned taxonomy to 10% of the ASVs. Out of the 95 species identified with the curated database, 26 were not detected with the complete one. While this could be interpreted as a failure from the complete database to identify some taxa, 19 of these species were in fact identified but as top 2 to top 10 identity by the NCBI databases. This means that the complete database found a better match (correct or not) for the ASVs to which these taxonomies were assigned. Such differences probably arise from the size of the curated database (containing around 5,000 sequences and 2,500 species) which is way less important than the size of the NCBI databases since it discarded every record taken out of Japan and not identifying insect species. By doing so, records of insect species present in Japan and whose sequences are present in a global database might be discarded resulting in a not or wrongly identified ASVs. Databases are by their nature incomplete (Hestetun et al. 2020), and filtering them should be done precautiously so that it does not strongly exacerbate this deficiency. The use of incomplete databases is indeed a common source of mistakes and should be avoided, since it generates false-negatives and false-positives, counterbalancing the advantages of curated databases (Preston, Fritzsche, and Woodcock 2022).

Notwithstanding these facts, in order to know with confidence if the utilization of a curated database is justified, both methods (curated versus uncurated) accuracy should be assessed. The constitution of mock communities could reveal itself truly useful in this way. These communities are composed of identified sequences on which databases can be tested. This would enable the finding of optimal procedures to generate custom databases (O’Rourke et al. 2020). Collecting insects in the sampling sites, morphologically identifying them and comparing the species identified via metabarcoding to them is also an increasingly used methods (Cahill et al. 2018; Huo et al. 2020; Svenningsen et al. 2021). In the case of diet analyses, this technique can be used to verify if the species identified via the non-curated and curated databases were present in the sampling sites but also to informs about the foraging preferences of the bird (species present preferentially fed on) (Bryant 1973; McClenaghan, Nol, and Kerr 2019).

As *Hirundo rustica* (Linnaeus, 1758) nearly uniquely feed on flying insects (García-Pérez et al. 2014), car nets could be used to collect these insects. These collects could make the trips between sites more profitable. Other techniques such as the implementation of Malaise traps on the sampling sites could serve the same goals (Law et al. 2017). The DNA sequences of insects caught in the nets or the traps and not included in public databases such as BOLD and NCBI could subsequently be uploaded on these (Svenningsen et al. 2021) and used in local databases. The inclusion of barcodes from targeted regions and the generation of curated databases composed of them has been studied for fish in California and showed results similar to those obtained through the use of Genbank (main NCBI nucleotides database) in terms of number of ASVs identified, although showing higher accuracy (Gold et al. 2021). This alternative, even though pricier, could therefore considerably increase the performance of curated databases and should be investigated in the context of this study.

However, in most of situations including the one of this research, scientists are limited to barcodes from global databases. In this case, the results of this study suggest that the development of global reference sequence databases is still necessary for the use of custom databases limited to countries or regions. The previous recommendation, applied to several metabarcoding studies could participate to the development of these databases. But as for now, the use of a deeply curated database such as the one generated with the insect records in Japan seems too restricted and could generate even more errors. The use of a complete database on curated sequences amplified with adapted primers and whose results are filtered for insects with a 97% identity threshold already allows for an important mitigation of the number of false positives while ensuring the identification of a larger proportion of ASVs and is therefore preferred (Preston, Fritzsche, and Woodcock 2022; Drake et al. 2022). Despite showing better performances than the created curated database, the NCBI ones solely assigned taxonomy to about 50% of the sampled ASVs, expressing the need for more extensive databases.

The NCBI databases was thus utilized for these previous reasons and enabled the characterisation of *Hirundo rustica* (Linnaeus, 1758) diet. The analyses identified insects from the Diptera order as the main prey items. Coleoptera and Lepidoptera have been classified as second and third most consumed orders followed respectively by Hemiptera and Hymenoptera. Diptera have been identified as dominant order in several studies carried on *Hirundo rustica* (Linnaeus, 1758) (Kragten, Reinstra, and Gertenaar 2009; A. Turner 2010; Kusack 2018; Schmiedová et al. 2022). In some other studies though, Coleoptera (Orłowski and Karg 2011; 2013) and Hymenoptera are prevailing (Law et al. 2017). As *Hirundo rustica* (Linnaeus, 1758) is a generalist insectivore, the abundances of the different taxa in the studied areas are likely to cause these variations (Orłowski and Karg 2011). Another dissimilarity observed in these researches is the low representation of Lepidoptera in this bird diet. While Diptera, Coleoptera, Hemiptera and Hymenoptera are characterized as the four most prevailing orders in all the mentioned researches, Lepidoptera was considered part of the prevailing orders in only one paper (Kusack 2018). The large number of ASVs identified as such in this master thesis could be caused by the biases towards the detection of lepidopteran and dipteran generated by the use of the ZBJ primers pair (Brandon-Mong et al. 2015). The relatively high abundance of this taxon in Japan is also likely to play a role in this: approximatively 6,300 species of lepidopteran can be observed in Japan, making it the second most represented insect order in terms of species just behind the Coleoptera order and its 10,600 species (Motokawa and Kajihara 2016). Dipterans, hymenopterans and hemipterans are the other most represented orders with respectively 5,300, 4,500 and 3,000 species found in Japan (Motokawa and Kajihara 2016). These numbers dovetail with the assumption that *Hirundo rustica* (Linnaeus, 1758) diet greatly depends on the prey availability (McClenaghan, Nol, and Kerr 2019).

The other orders detected in this study were Ephemeroptera, Trichoptera, Odonata, Orthoptera, Psocoptera, Raphidioptera, Thysanoptera and Neuroptera. They individually represented less than 1% of the ASVs and were all identified as *Hirundo rustica* (Linnaeus, 1758) prey items in other studies (A. K. Turner 1982; A. Turner 2010; Orłowski and Karg 2011; 2013; Law et al. 2017; Kusack 2018; Gomes Lopes 2018; McClenaghan, Nol, and Kerr 2019; Mansor et al. 2020).

Looking at the results from a family point of view, the 103 families detected represent a relatively large number compared to other diet studies carried within one kind of habitat and which identified 30 (Law et al. 2017), 50 (Orłowski and Karg 2011), 100 (Kusack 2018) and 130 families (McClenaghan, Kerr, and Nol 2019). The two last researches were nonetheless the only ones relying on metabarcoding analyses and on the same primers pair which could explain the variations in the number of detected families (Zenker, Specht, and Fonseca 2020), along with their different locations and thus insect communities. The families considered as most represented in *Hirundo rustica* (Linnaeus, 1758) diet by these works were all detected within this master thesis as well. However, none of them identified Chironomidae and Drosophilidae, the two most represented family prey items of this research. According to Grigolo et al. (2020) this is because Chironomidae represent a poor-quality source of food and is therefore foraged predominately in enviromnents with low-abundance of other prey species. The presence of Drosophilidae in *Hirundo rustica* (Linnaeus, 1758) feces has, on the other side, never been documented before (the number of paper aiming to provide a detailed description of its diet being limited). But a research carried on another Hirundinidae, *Hirundo spilodera*, did not fail to identify this family as the most important component of its diet (Earl 1985), making the presence of Drosophilidae in *Hirundo rustica* (Linnaeus, 1758) diet plausible.

Non-insect taxa were identified as well in the excrements but were not all part of the bird diet. The prevailing ones were Bacterias. Their occurrence is quite common in diet studies since bacterias are present in gut or in the environment, contaminating the feces (de Sousa, Silva, and Xavier 2019). Non-arthropods animals and plants were also observed in the feaces which is quite common in diet study of this bird (Orłowski and Karg 2011). While most of them are likely to be truly part of the diet, species from the Nematoda (*Pristionchus pacificus* and *Steinernema sp.*) and the Oomycota phyllum (*Pythium spp.*, *Globisporangium rostratum* and *Hyaloperonospora sp.*) might be carried by foraged insects (Orłowski and Karg 2011; Willsey, Chatterton, and Cárcamo 2017; McClenaghan, Nol, and Kerr 2019).

Analysing feces to determine the diet composition of insectivorous birds has been characterized as a reliable method (Suarez-Rubio and Krenn 2018; Grigolo et al. 2020), specificially for hirundines (Bryant 1973; Orłowski and Karg 2011). These assumptions, along with the literature and the pertinence of the results considering the taxa abundances in Japan suggest reliability in the estimated diet composition, at least at a family level.

However, the readers of this work should stay aware of the limits imposed by the metabarcoding technique and the choices made during this work, especially for species-level identification. As written in the section “1. Literature review” and besides the primers selection (Ficetola et al. 2021), the way the data were collected, the PCR procedures, the sequences filtering and the databases utilized are likely to generate a few biases (McClenaghan, Nol, and Kerr 2019; Uiterwaal and DeLong 2020).

4.2. Environmental analyses

The environmental analyses support the fact that the selected sampling areas show similar characteristics in terms of land covers proportions and cannot be distinguished on such criterias. Isumi is indeed the only area strongly standing out of the others. 35 samples out of 98 come from sites which contain between 70 and 80% of impervious surfaces and only 35% of the samples are from sites with less than 70% of impervious surfaces cover. This lack of heterogeneity leads to an underrepresentation of the rural sites in the dataset. The discrimination of Urban-Rural gradient has been realized in many studies and some investigate the sampling sites before selecting them (Randa and Yunger 2006; Pillsbury and Miller 2008). This way, Randa and Yunger (2006) used land covers maps to generate gradients covering as much types of land covers as possible. The same method could be used to ensure a heterogeneity of the sampled sites. Once nests have been identified through the Wild Bird Society of Japan website, analyses could be carried on QGIS to ensure a diversity of environment and of impervious surfaces proportions.

4.3. Analysis of the influence of variables on diversity

The statistical analyses carried on the α and β diversity identified the average temperature, the range of daily temperatures, the average monthly precipitation, the LPI and the radiance as key factors in the diet characteristics of *Hirundo rustica* (Linnaeus, 1758).

More precisely, the results suggested that the diet richness was influenced by the LPI, negatively correlated (-0.9) to the proportion of impervious surfaces and thus to the urban-rural gradient. The more fragmented and thus the more urbanized an area, the lower the diet diversity. This supports the initial hypothesis of this work stating that diet species richness was expected to decrease along the rural-urban gradient. The taxa richness is paradoxically positively associated to artificial night sky radiance, meaning that the presence of light might be responsible for a higher number of species preyed upon by this bird. As for the Shannon diversity index, the results should be analyzed more carefully since they took into account the abundance of ASVs as well, which is correlated to a species biomass but cannot be considered as a proxy for its abundance (Elbrecht and Leese 2015). Higher Shannon diversity index, representing diets with high number of taxa and a relatively equal repartition of their biomasses, were associated to higher LPI, average temperatures and precipitation by the GLMM. According to the results of the statistical models, fragmentation impacts the diet α diversity. This is not the case for the β diversity. However, the obtained p-values implied that both temperatures (daily range and monthly average) and precipitation (monthly average) were playing a role in this diet characteristic.

To sum this up and still according to the results, fragmentation (and thus logically urbanization) negatively impacts both the number of taxa and their evenness within *Hirundo rustica* (Linnaeus, 1758) diet. Average temperature and precipitation are positively associated with the number of taxa and their evenness average temperature, and a positive relationship between radiance and number of taxa within the diet has been determined. Eventually the diet composition is influenced by the average temperature and the range of daily temperature.

Comparing these results to similar researches is complex since the impact of variables on *Hirundo rustica* (Linnaeus, 1758) diet was investigated in few studies (A. K. Turner 1983; Gruebler, Morand, and Naefdaenzer 2008; Orłowski and Karg 2011; 2013; Kusack 2018; Mansor et al. 2020; Wang, Tuanmu, and Hung 2021), and only three of them included variables investigated in this master thesis. The first one suggested an influence of the average temperature and precipitation on the prey abundance and on the energy gain rates of nestlings of this species (A. K. Turner 1983). These gain rates and alterations in prey abundances could therefore indicate an influence of these climatic variables on the diet composition and evenness of this bird, confirming the trends found in this research. Another research found complex relationships between weather conditions (temperature, precipitations and wind speed) and *Hirundo rustica* (Linnaeus, 1758) body mass and hypothesized that the body mass was directly dependent of the foraging behavior of these birds based on literature (Gruebler, Morand, and Naefdaenzer 2008). Other sources support these statements by claiming the impact of weather on the dynamics of insect populations (McIntyre et al. 2001; Liebhold and Bentz 2011; Zaller et al. 2014; Youngsteadt et al. 2017) and therefore on *Hirundo rustica* (Linnaeus, 1758) diet. The two studies cited above and targeting *Hirundo rustica* (Linnaeus, 1758) energy gain rate and body mass respectively display a positive and a negative relationship with the temperature. Comparing the direction of these relationships (and the ones resulting from the precipitations) with the results of this master thesis cannot be achieved. Indeed, while it is relatively safe to assume that energy gain rate fluctuations and body masses are influenced by the composition and evenness of the diet, their growth could either be associated to dominance of highly-energy prey or to the evenness of many, resulting in dissimilar Shannon diversity index values.

As for the landscape fragmentation, no statistical relationship with *Hirundo rustica* (Linnaeus, 1758) dietary habits had been described before. However, the responsibility of this ecological process for alterations in insect community either in terms of abundance and species richness has been reported (Didham et al. 1996). The presence of a negative relationship between fragmentation and α diversity index was consequently expected. This expectation was on top of that caused by the correlation of the fragmentation (represented by the LPI) and the urbanization (represented by the proportion of impervious surfaces), the influence of urbanization on *Hirundo rustica* (Linnaeus, 1758) being well documented (Savignac 2011; Imlay and Leonard 2019; Zhao et al. 2021).

At last, the influence of artificial night sky radiance on diversity metrics argues in favor of a more diverse diet in radiant sites. A research conducted in Taiwan ascertains this relationship by underlining an acceleration of the feeding rate and a broadening of the foraging period (past sunset) for *Hirundo rustica* (Linnaeus, 1758) individuals nesting in sites with an elevated intensity of artificial light at night. More than that, in the context of this research, the offspring demands for food seemed to occur at night, resulting in parents responding to this demand by foraging for insects drawn in by close light sources. This behavior is presumably responsible for an augmentation of the food procurement to the offspring (Wang, Tuanmu, and Hung 2021) and by this way for an enrichment of the diet in terms of number of species. The same as been observed with insectivorous bats attracted by street lights emitting ultraviolet wavelengths themselves luring positively phototactic insects at certain spatial points, lessening foraging time and resulting in fewer energy losses (Rowse et al. 2016).

Even though not correlated in this study, the close association of the climatic conditions, the artificial night sky radiance, and the urbanization should not be ignored by the readers of this work. Human settlements are indeed associated to artificial light (Horton et al. 2019), urban heat islands (Heaviside, Macintyre, and Vardoulakis 2017) and alteration in the water cycle (Ruby n.d.). Urbanization also results in habitat fragmentation (Z. Liu, He, and Wu 2016) which was illustrated by a high correlation in this work.

Eventually, an important notice about this work should be that alterations of the diet in response to variables should be interpreted as such. Assuming that a high diversity within the consumed insects always represents a high-quality diet could lead to misconceptions of *Hirundo rustica* (Linnaeus, 1758) interactions with its environment. In some cases and even if this species is generalist, individuals might preferentially prey upon an abundant taxa at the expense of the others, resulting in an inferior dietary diversity but an inflated energy gain rate (Grigolo et al. 2020). Typically, 4 to 8 mm length insects (such as Calliphoridae), possessing mediocre flight performances (such as Formicidae) and tending to aggregate (such as Calliphoridae and Formicidae) enable considerable energy gain (relatively large preys) and mitigated energy loss (uncomplicated to hunt) (Law et al. 2017).

4.4. Personal contribution

The main contribution brought personally to this master thesis was the sharing of the knowledge acquired through this work with the students from the System Planning Lab of Chiba University and with the people met in Japan, raising this way awareness on environmental-related problematics. Integrating facets of the Japanese culture and teaching them to people who did not get the chance to experience it represents, although not directly to the scientific part of this research, a precious personal contribution.

Going on the field to observe *Hirundo rustica* (Linnaeus, 1758) nests and its behavior (especially its foraging characteristics) in order to discover more about this bird and apprehend the context from a more practical point of view was a personal contribution as well.

Another personal contribution, besides the results brought by this work was the literature review conducted to determine whether a better primers pair than the ZBJ ones existed for the identification of insects. The use of a combination of the primers ZBJ-ArtF1c/ZBJ-ArtR2c and fwhF2/fwhR2 was suggested but still has to be tested.

5. Conclusion

The main goal of this work was to study the influence of an urban-rural gradient and associated factors on the diet of *Hirundo rustica* (Linnaeus, 1758). This was done successfully by estimating landscape metrics, extracting climatic variables, and characterizing the urban-rural gradient on QGIS before testing the statistical relationships of these with the taxa identified in this bird feces via DNA metabarcoding. From these operations was revealed a positive relationship between the LPI, whose higher values represent less fragmented areas, and the diet α diversity. Artificial night sky radiance, monthly temperature and precipitation averages impact the dietary behavior and display significant relationships with the α diversity diet as well. Monthly temperature and precipitation averages along with the temperatures range were them identified as key factors in the diet composition in this study context. As a secondary objective, the utilization of a curated database composed of insect records from BOLD was investigated but showed limited performances compared to the complete NCBI databases.

By determining the influence of the previously cited variables on *Hirundo rustica* (Linnaeus, 1758) alimentation, this work contributed to the knowledge acquired on this species and brought information on its diet in Japan, which has never been detailed before. Linking this bird diet to an urbanization gradient filled another gap in knowledge and could inspire scientists to test whether these modifications of dietary behavior along the gradient are responsible for the phenotypic differences observed between rural and urban populations. Using the metabarcoding technique for such studies could present numerous advantages.

The development of such a technique opens a wide range of opportunities for scientists and should be embraced. Even if several biases and limits are associated to it and were described in the framework of this work, it is essential to keep in mind that every approach brings its share of inaccuracies. Furthermore, the creation of mock communities and the analyses of barcodes from local species would permit an assessment of the result quality while contributing to the existing databases. Moreover, carrying out research on the ASVs assigned families or even orders might provide more accurate results and allow to study the diet composition along the urbanization gradient in terms of number of species preyed upon in each order or family. Eventually, calculating the sites' impervious surfaces proportion prior data collection would enable a better representation of *Hirundo rustica* (Linnaeus, 1758) diet in rural areas.

Based on these conclusions, it is safe to assume that urbanization leads to modification of trophic interactions and disrupts the balance of well-implemented ecosystems. A mitigation of its impact, for instance through the establishment of green infrastructures and the implementation of environmental measures in territories affected by the urban sprawl should be considered.

6. References

‘Adonis: Permutational Multivariate Analysis of Variance Using... in Vegan: Community Ecology Package’. 2022. 18 April 2022. https://rdrr.io/cran/vegan/man/adonis.html.

Allison, Jeremy D., C. Wood Johnson, James R. Meeker, Brian L. Strom, and Sarah M. Butler. 2011. ‘Effect of Aerosol Surface Lubricants on the Abundance and Richness of Selected Forest Insects Captured in Multiple-Funnel and Panel Traps’. *Journal of Economic Entomology* 104 (4): 1258–64. https://doi.org/10.1603/EC11044.

Ando, Haruko, Haruka Mukai, Taketo Komura, Tungga Dewi, Masaki Ando, and Yuji Isagi. 2020. ‘Methodological Trends and Perspectives of Animal Dietary Studies by Noninvasive Fecal DNA Metabarcoding’. *Environmental DNA* 2 (4): 391–406. https://doi.org/10.1002/edn3.117.

Andrade, Antonio C. de. 2020. ‘Metropolitan Lizards? Urbanization Gradient and the Density of Lagartixas (Tropidurus Hispidus) in a Tropical City’. *Ecology and Evolution* 10 (4): 1740–50. https://doi.org/10.1002/ece3.5518.

Andriollo, Tommy, François Gillet, Johan R. Michaux, and Manuel Ruedi. 2019. ‘The Menu Varies with Metabarcoding Practices: A Case Study with the Bat Plecotus Auritus’. *PLOS ONE* 14 (7): e0219135. https://doi.org/10.1371/journal.pone.0219135.

Ashelford, Kevin E., Nadia A. Chuzhanova, John C. Fry, Antonia J. Jones, and Andrew J. Weightman. 2006. ‘New Screening Software Shows That Most Recent Large 16S RRNA Gene Clone Libraries Contain Chimeras’. *Applied and Environmental Microbiology* 72 (9): 5734–41. https://doi.org/10.1128/AEM.00556-06.

Bates, Douglas, Martin Maechler, Ben Bolker [aut, cre, Steven Walker, Rune Haubo Bojesen Christensen, Henrik Singmann, et al. 2022. ‘Lme4: Linear Mixed-Effects Models Using “Eigen” and S4’. https://CRAN.R-project.org/package=lme4.

Berrick, Stephen. 2022. ‘Visible Infrared Imaging Radiometer Suite (VIIRS) | Earthdata’. March 2022. https://earthdata.nasa.gov/earth-observation-data/near-real-time/download-nrt-data/viirs-nrt/.

BirdLife International. 2019. ‘IUCN Red List of Threatened Species: Hirundo Rustica’. *IUCN Red List of Threatened Species*. https://dx.doi.org/10.2305/IUCN.UK.2019-3.RLTS.T22712252A137668645.en.

Bokulich, Nicholas A., Benjamin D. Kaehler, Jai Ram Rideout, Matthew Dillon, Evan Bolyen, Rob Knight, Gavin A. Huttley, and J. Gregory Caporaso. 2018. ‘Optimizing Taxonomic Classification of Marker-Gene Amplicon Sequences with QIIME 2’s Q2-Feature-Classifier Plugin’. *Microbiome* 6 (1): 90. https://doi.org/10.1186/s40168-018-0470-z.

Bolyen, Evan, Jai Ram Rideout, Matthew R. Dillon, Nicholas A. Bokulich, Christian C. Abnet, Gabriel A. Al-Ghalith, Harriet Alexander, et al. 2019. ‘Reproducible, Interactive, Scalable and Extensible Microbiome Data Science Using QIIME 2’. *Nature Biotechnology* 37 (8): 852–57. https://doi.org/10.1038/s41587-019-0209-9.

Brandon-Mong, G.-J., H.-M. Gan, K.-W. Sing, P.-S. Lee, P.-E. Lim, and J.-J. Wilson. 2015. ‘DNA Metabarcoding of Insects and Allies: An Evaluation of Primers and Pipelines’. *Bulletin of Entomological Research* 105 (6): 717–27. https://doi.org/10.1017/S0007485315000681.

Braukmann, Thomas W. A., Natalia V. Ivanova, Sean W. J. Prosser, Vasco Elbrecht, Dirk Steinke, Sujeevan Ratnasingham, Jeremy R. de Waard, Jayme E. Sones, Evgeny V. Zakharov, and Paul D. N.

Hebert. 2019. ‘Metabarcoding a Diverse Arthropod Mock Community’. *Molecular Ecology Resources* 19 (3): 711–27. https://doi.org/10.1111/1755-0998.13008.

Britannica, T. 2011. ‘Hirundinidae’. In .

Bryant, D. M. 1973. ‘The Factors Influencing the Selection of Food by the House Martin (Delichon Urbica (L.))’. *Journal of Animal Ecology* 42 (3): 539–64. https://doi.org/10.2307/3123.

Bukovinszky, Tibor, Rieta Gols, Andre Kamp, Filipe de Oliveira-Domingues, Peter A. Hambäck, Yde Jongema, T. Martijn Bezemer, Marcel Dicke, Nicole M. van Dam, and Jeffrey A. Harvey. 2010. ‘Combined Effects of Patch Size and Plant Nutritional Quality on Local Densities of Insect Herbivores’. *Basic and Applied Ecology* 11 (5): 396–405. https://doi.org/10.1016/j.baae.2010.04.005.

Burkman, C. E., and M. M. Gardiner. 2014. ‘Urban Greenspace Composition and Landscape Context Influence Natural Enemy Community Composition and Function.’ *Biological Control* 75: 58–67.

Bush, Alex, Zacchaeus G. Compson, Wendy A. Monk, Teresita M. Porter, Royce Steeves, Erik Emilson, Nellie Gagne, Mehrdad Hajibabaei, Mélanie Roy, and Donald J. Baird. 2019. ‘Studying Ecosystems With DNA Metabarcoding: Lessons From Biomonitoring of Aquatic Macroinvertebrates’. *Frontiers in Ecology and Evolution* 7. https://www.frontiersin.org/article/10.3389/fevo.2019.00434.

Cahill, Abigail E., John K. Pearman, Angel Borja, Laura Carugati, Susana Carvalho, Roberto Danovaro, Sarah Dashfield, et al. 2018. ‘A Comparative Analysis of Metabarcoding and Morphology-Based Identification of Benthic Communities across Different Regional Seas’. *Ecology and Evolution* 8 (17): 8908–20. https://doi.org/10.1002/ece3.4283.

Callahan, Benjamin J., Paul J. McMurdie, and Susan P. Holmes. 2017. ‘Exact Sequence Variants Should Replace Operational Taxonomic Units in Marker-Gene Data Analysis’. *The ISME Journal* 11 (12): 2639–43. https://doi.org/10.1038/ismej.2017.119.

Callahan, Benjamin J, Paul J McMurdie, Michael J Rosen, Andrew W Han, Amy Jo A Johnson, and Susan P Holmes. 2016. ‘DADA2: High-Resolution Sample Inference from Illumina Amplicon Data’. *Nature Methods* 13 (7): 581–83. https://doi.org/10.1038/nmeth.3869.

Canada, Environment and Climate Change. 2011. ‘Barn Swallow (Hirundo Rustica): COSEWIC Assessment and Status Report 2011’. Assessments;research. 3 November 2011. https://www.canada.ca/en/environment-climate-change/services/species-risk-public-registry/cosewic-assessments-status-reports/barn-swallow-2011.html.

Cardoso, Pedro, Philip S. Barton, Klaus Birkhofer, Filipe Chichorro, Charl Deacon, Thomas Fartmann, Caroline S. Fukushima, et al. 2020. ‘Scientists’ Warning to Humanity on Insect Extinctions’. *Biological Conservation* 242 (February): 108426. https://doi.org/10.1016/j.biocon.2020.108426.

Choate, Beth A., Paige L. Hickman, and Erica A. Moretti. 2018. ‘Wild Bee Species Abundance and Richness across an Urban–Rural Gradient’. *Journal of Insect Conservation* 22 (3): 391–403. https://doi.org/10.1007/s10841-018-0068-6.

Cristescu, Melania E. 2014. ‘From Barcoding Single Individuals to Metabarcoding Biological Communities: Towards an Integrative Approach to the Study of Global Biodiversity’. *Trends in Ecology & Evolution* 29 (10): 566–71. https://doi.org/10.1016/j.tree.2014.08.001.

Didham, Raphael K., Jaboury Ghazoul, Nigel E. Stork, and Andrew J. Davis. 1996. ‘Insects in Fragmented Forests: A Functional Approach’. *Trends in Ecology & Evolution* 11 (6): 255–60. https://doi.org/10.1016/0169-5347(96)20047-3.

Dirzo, Rodolfo, Hillary S. Young, Mauro Galetti, Gerardo Ceballos, Nick J. B. Isaac, and Ben Collen. 2014. ‘Defaunation in the Anthropocene’. *Science* 345 (6195): 401–6. https://doi.org/10.1126/science.1251817.

Dong, Xuegang, Zhiguo Meng, Yongzhi Wang, Yuanzhi Zhang, Haoteng Sun, and Qingshuai Wang. 2021. ‘Monitoring Spatiotemporal Changes of Impervious Surfaces in Beijing City Using Random Forest Algorithm and Textural Features’. *Remote Sensing* 13 (January): 153. https://doi.org/10.3390/rs13010153.

Dopheide, Andrew, Leah K. Tooman, Stefanie Grosser, Barbara Agabiti, Birgit Rhode, Dong Xie, Mark I. Stevens, et al. 2019. ‘Estimating the Biodiversity of Terrestrial Invertebrates on a Forested Island Using DNA Barcodes and Metabarcoding Data’. *Ecological Applications* 29 (4): e01877. https://doi.org/10.1002/eap.1877.

Dormann, Carsten F., Jane Elith, Sven Bacher, Carsten Buchmann, Gudrun Carl, Gabriel Carré, Jaime R. García Marquéz, et al. 2013. ‘Collinearity: A Review of Methods to Deal with It and a Simulation Study Evaluating Their Performance’. *Ecography* 36 (1): 27–46. https://doi.org/10.1111/j.1600-0587.2012.07348.x.

Drake, Lorna E., Jordan P. Cuff, Rebecca E. Young, Angela Marchbank, Elizabeth A. Chadwick, and William O. C. Symondson. 2022. ‘An Assessment of Minimum Sequence Copy Thresholds for Identifying and Reducing the Prevalence of Artefacts in Dietary Metabarcoding Data’. *Methods in Ecology and Evolution* 13 (3): 694–710. https://doi.org/10.1111/2041-210X.13780.

Earl, ï¿½ Roy A. 1985. ‘Foraging Behaviour and Diet of the South African Cliff Swallow Hirundo Spilodera (Aves: Hirundinidae) : Diet’. *Navorsinge van Die Nasionale Museum : Researches of the National Museum* 5 (4): 59–62. https://doi.org/10.10520/AJA00679208_529.

Edgar, Robert. n.d. ‘Dereplication’. Accessed 4 August 2022. https://drive5.com/usearch/manual8.1/dereplication.html.

Eisenbeis, Gerhard, and Andreas Hänel. 2009. ‘Chapter 15. Light Pollution and the Impact of Artificial Night Lighting on Insects’. *A. H.*, 20.

Elbrecht, Vasco, and Florian Leese. 2015. ‘Can DNA-Based Ecosystem Assessments Quantify Species Abundance? Testing Primer Bias and Biomass - Sequence Relationships with an Innovative Metabarcoding Protocol’. e1258. PeerJ Inc. https://doi.org/10.7287/peerj.preprints.1023v1.

———. 2017. ‘Validation and Development of COI Metabarcoding Primers for Freshwater Macroinvertebrate Bioassessment’. Preprint. PeerJ Preprints. https://doi.org/10.7287/peerj.preprints.2044v5.

Evans, Karl, Jeremy Wilson, and Richard Bradbury. 2007. ‘Effects of Crop Type and Aerial Invertebrates Abundance on Foraging Barn Swallows Hirundo Rustica’. *Agriculture, Ecosystems & Environment* 122 (October): 267–73. https://doi.org/10.1016/j.agee.2007.01.015.

Fahrig, Lenore, and Ian Jonsen. 1998. ‘Effect of Habitat Patch Characteristics on Abundance and Diversity of Insects in an Agricultural Landscape’. *Ecosystems* 1 (2): 197–205. https://doi.org/10.1007/s100219900015.

Falchi, Fabio, Pierantonio Cinzano, Dan Duriscoe, Christopher C. M. Kyba, Christopher D. Elvidge, Kimberly Baugh, Boris A. Portnov, Nataliya A. Rybnikova, and Riccardo Furgoni. 2016. ‘The New

World Atlas of Artificial Night Sky Brightness'. *Science Advances* 2 (6): e1600377. https://doi.org/10.1126/sciadv.1600377.

Fenoglio, María Silvina, Ana Calviño, Ezequiel González, Adriana Salvo, and Martin Videla. 2021. 'Urbanisation Drivers and Underlying Mechanisms of Terrestrial Insect Diversity Loss in Cities'. *Ecological Entomology* 46 (4): 757–71. https://doi.org/10.1111/een.13041.

Ficetola, Gentile Francesco, Frédéric Boyer, Alice Valentini, Aurélie Bonin, Albin Meyer, Tony Dejean, Coline Gaboriaud, Philippe Usseglio-Polatera, and Pierre Taberlet. 2021. 'Comparison of Markers for the Monitoring of Freshwater Benthic Biodiversity through DNA Metabarcoding'. *Molecular Ecology* 30 (13): 3189–3202. https://doi.org/10.1111/mec.15632.

Fitzsimons, James, and Janelle Thomas. 2012. 'Ground-Foraging Techniques of Welcome Swallows Hirundo Neoxena, Including an Instance of Kleptoparasitism'. *Australian Field Ornithology* 29 (September): 160–65.

Fortel, Laura, Mickaël Henry, Laurent Guilbaud, Anne Laure Guirao, Michael Kuhlmann, Hugues Mouret, Orianne Rollin, and Bernard E. Vaissière. 2014. 'Decreasing Abundance, Increasing Diversity and Changing Structure of the Wild Bee Community (Hymenoptera: Anthophila) along an Urbanization Gradient'. *PLOS ONE* 9 (8): e104679. https://doi.org/10.1371/journal.pone.0104679.

Gailly, Robin. n.d. 'Hirondelle Rustique (Hirundo Rustica) | Oiseaux | Vertébrés | La Biodiversité En Wallonie'. Accessed 24 February 2022. http://biodiversite.wallonie.be/fr/hirundo-rustica.html?IDD=50334196&IDC=305.

García-Pérez, Belén, Keith A. Hobson, Gretchen Albrecht, Michael D. Cadman, and Antonio Salvadori. 2014. 'Influence of Climate on Annual Survival of Barn Swallows (Hirundo Rustica) Breeding in North America'. *The Auk* 131 (3): 351–62. https://doi.org/10.1642/AUK-13-145.1.

Geslin, Benoît, Violette Le Féon, Morgane Folschweiller, Floriane Flacher, David Carmignac, Eric Motard, Samuel Perret, and Isabelle Dajoz. 2016. 'The Proportion of Impervious Surfaces at the Landscape Scale Structures Wild Bee Assemblages in a Densely Populated Region'. *Ecology and Evolution* 6 (18): 6599–6615. https://doi.org/10.1002/ece3.2374.

Gill, Frank, Mary Heimerdinger Clench, and Oliver L. Austin. 2021. 'Passeriform'. https://www.britannica.com/animal/passeriform.

Gold, Zachary, Emily E. Curd, Kelly D. Goodwin, Emma S. Choi, Benjamin W. Frable, Andrew R. Thompson, Harold J. Walker, et al. 2021. 'Improving Metabarcoding Taxonomic Assignment: A Case Study of Fishes in a Large Marine Ecosystem'. *Molecular Ecology Resources* 21 (7): 2546–64. https://doi.org/10.1111/1755-0998.13450.

Gomes Lopes, Cláudia Maria. 2018. 'Barn Swallows' Diet in the Montado'. Master thesis, Évora: Universidade de Évora. http://dspace.uevora.pt/rdpc/bitstream/10174/23686/1/Mestrado-Biologia_da_Conserva%C3%A7%C3%A3o-Cl%C3%A1udia_Maria_Gomes_Lopes-Barn_swallows_diet_in_the_montado%20.pdf.

González-Césped, Carlos, Alberto J. Alaniz, Pablo M. Vergara, Elizabeth Chiappa, Jorge Zamorano, and Víctor Mandujano. 2021. 'Effects of Urban Environmental Conditions and Landscape Structure on Taxonomic and Functional Groups of Insects'. *Urban Forestry & Urban Greening* 58 (March): 126902. https://doi.org/10.1016/j.ufug.2020.126902.

Gordon, A., and G. Hannon. 2010. ‘Fastx-Toolkit. FASTQ/A Short-Reads Pre-Processing Tools.’ http://hannonlab.cshl.edu/.

Grigolo, Camila, Beatrice Sicurella, Federica Musitelli, Andrea Romano, Manuela Caprioli, Diego Rubolini, Roberto Ambrosini, and Mauro Gobbi. 2020. ‘Diet Heterogeneity and Antioxidant Defence in Barn Swallow Hirundo Rustica Nestlings’. *Avocetta* 43 (January). https://doi.org/10.30456/AVO.2019203.

Gruebler, M., M. Morand, and B. Naefdaenzer. 2008. ‘A Predictive Model of the Density of Airborne Insects in Agricultural Environments’. *Agriculture, Ecosystems & Environment* 123 (1–3): 75.

Guo, Wei, Dengsheng Lu, and Wenhui Kuang. 2017. ‘Improving Fractional Impervious Surface Mapping Performance through Combination of DMSP-OLS and MODIS NDVI Data’. *Remote Sensing* 9 (4): 375. https://doi.org/10.3390/rs9040375.

Haarsma, Anne-Jifke, Henk Siepel, and Barbara Gravendeel. 2016. ‘Added Value of Metabarcoding Combined with Microscopy for Evolutionary Studies of Mammals’. *Zoologica Scripta* 45 (S1): 37–49. https://doi.org/10.1111/zsc.12214.

Heaviside, Clare, Helen Macintyre, and Sotiris Vardoulakis. 2017. ‘The Urban Heat Island: Implications for Health in a Changing Environment’. *Current Environmental Health Reports* 4 (3): 296–305. https://doi.org/10.1007/s40572-017-0150-3.

Hestetun, Jon Thomassen, Einar Bye-Ingebrigtsen, R. Henrik Nilsson, Adrian G. Glover, Per-Otto Johansen, and Thomas G. Dahlgren. 2020. ‘Significant Taxon Sampling Gaps in DNA Databases Limit the Operational Use of Marine Macrofauna Metabarcoding’. *Marine Biodiversity* 50 (5): 70. https://doi.org/10.1007/s12526-020-01093-5.

Horton, Kyle G, Cecilia Nilsson, Benjamin M Van Doren, Frank A La Sorte, Adriaan M Dokter, and Andrew Farnsworth. 2019. ‘Bright Lights in the Big Cities: Migratory Birds’ Exposure to Artificial Light’. *Frontiers in Ecology and the Environment* 17 (4): 209–14. https://doi.org/10.1002/fee.2029.

Hufnagel Levente, and Rédei D. 2003. ‘The species composition of true bug assemblages extracted with Berlese funnels (Data to the knowledge on the ground-living Heteroptera of Hungary, No 1)’. *Applied Ecology and Environmental Research* 1 (1–2): 93–113.

Huo, Shouliang, Xiaochuang Li, Beidou Xi, Hanxiao Zhang, Chunzi Ma, and Zhuoshi He. 2020. ‘Combining Morphological and Metabarcoding Approaches Reveals the Freshwater Eukaryotic Phytoplankton Community’. *Environmental Sciences Europe* 32 (1): 37. https://doi.org/10.1186/s12302-020-00321-w.

Ii, Michael S. Robeson, Devon R. O’Rourke, Benjamin D. Kaehler, Michal Ziemski, Matthew R. Dillon, Jeffrey T. Foster, and Nicholas A. Bokulich. 2021. ‘RESCRIPt: Reproducible Sequence Taxonomy Reference Database Management’. *PLOS Computational Biology* 17 (11): e1009581. https://doi.org/10.1371/journal.pcbi.1009581.

Ilayaraja, K, and A Ambica. 2015. ‘Spatial Distribution of Groundwater Quality Between Injambakkam- Thiruvanmyiur Areas, South East Coast of India’. *Nature Environment and Pollution Technology* 14 (4): 6.

Imlay, Tara, and Marty Leonard. 2019. ‘A Review of the Threats to Adult Survival for Swallows (Family: Hirundinidae)’. *Bird Study* 66 (September): 1–13. https://doi.org/10.1080/00063657.2019.1655527.

IUCN. 2020. 'IUCN Red List of Threatened Species'. 2020. https://www.iucnredlist.org/statistics.

Jia, Yuqiu, Lina Tang, Min Xu, and Xinyi Yang. 2019. 'Landscape Pattern Indices for Evaluating Urban Spatial Morphology – A Case Study of Chinese Cities'. *Ecological Indicators* 99 (April): 27–37. https://doi.org/10.1016/j.ecolind.2018.12.007.

Jiménez, Randell R. 2021. 'Practical Metagenomics: Microbiome Tutorial with QIIME2'. https://f1000research.com/documents/10-798/pdf.

Joshi, NA., and JN. Fass. 2011. 'Sickle: A Sliding-Window, Adaptive, Quality-Based Trimming Tool for FastQ Files'. https://github.com/najoshi/sickle.

Kaminski, Abigail, Dana Marie Bauer, Kathleen P. Bell, Cynthia S. Loftin, and Erik J. Nelson. 2021. 'Using Landscape Metrics to Characterize Towns along an Urban-Rural Gradient'. *Landscape Ecology* 36 (10): 2937–56. https://doi.org/10.1007/s10980-021-01287-7.

Kang, Sung-Ryong, and Michael D. Kaller. 2013. 'Foraging Habitat Use of Breeding Barn Swallow (<I>Hirundo Rustica</I>) in Farmland, Estuary, and Island'. *Open Journal of Ecology* 03 (01): 30–33. https://doi.org/10.4236/oje.2013.31004.

Kearse, Matthew, Richard Moir, Amy Wilson, Steven Stones-Havas, Matthew Cheung, Shane Sturrock, Simon Buxton, et al. 2012. 'Geneious Basic: An Integrated and Extendable Desktop Software Platform for the Organization and Analysis of Sequence Data'. *Bioinformatics* 28 (12): 1647–49. https://doi.org/10.1093/bioinformatics/bts199.

Keele, Jacque, Jamie Carmon, Sherri F. Pucherelli, and Denise Hosler. 2020. 'Identification of Unknown Organisms by DNA Barcoding: A Molecular Method for Species Classification'. 86-68220-14–14. The U.S. Department of the Interior. https://www.usbr.gov/research/projects/detail.cfm?id=45.

Kelly, Ryan P., Jesse A. Port, Kevan M. Yamahara, and Larry B. Crowder. 2014. 'Using Environmental DNA to Census Marine Fishes in a Large Mesocosm'. *PLOS ONE* 9 (1): e86175. https://doi.org/10.1371/journal.pone.0086175.

Kessler, Michael, Stefan Abrahamczyk, Merijn Bos, Damayanti Buchori, Dadang Dwi Putra, S. Robbert Gradstein, Patrick Höhn, et al. 2009. 'Alpha and Beta Diversity of Plants and Animals along a Tropical Land-Use Gradient'. *Ecological Applications* 19 (8): 2142–56. https://doi.org/10.1890/08-1074.1.

Kim, Bo-Ra, Jiwon Shin, Robin B. Guevarra, Jun Hyung Lee, Doo Wan Kim, Kuk-Hwan Seol, Ju-Hoon Lee, Hyeun Bum Kim, and Richard E. Isaacson. 2017. 'Deciphering Diversity Indices for a Better Understanding of Microbial Communities'. *Journal of Microbiology and Biotechnology* 27 (12): 2089–93. https://doi.org/10.4014/jmb.1709.09027.

Kragten, Steven, Erwin Reinstra, and Erik Gertenaar. 2009. 'Breeding Barn Swallows Hirundo Rustica on Organic and Conventional Arable Farms in the Netherlands'. *Journal of Ornithology* 150 (2): 515. https://doi.org/10.1007/s10336-009-0383-5.

Krawchuk, Meg A., and Philip D. Taylor. 2003. 'Changing Importance of Habitat Structure across Multiple Spatial Scales for Three Species of Insects'. *Oikos* 103 (1): 153–61. https://doi.org/10.1034/j.1600-0706.2003.12487.x.

Kusack, Jackson. 2018. 'Landscape-Level Effects of Agricultural Intensification on the Condition and Diet of Nestling Barn Swallows (Hirundo Rustica)'. *Electronic Thesis and Dissertation Repository*, December. https://ir.lib.uwo.ca/etd/5900.

Lahti, Leo, Sudarshan Shetty, Turaga Nitesh, Leung Eric, Gilmore Rob, Salojärvi Jarkko, Blake Tineka, et al. 2017. 'Introduction to the Microbiome R Package'. 2017. https://microbiome.github.io/tutorials/.

Law, Audrey, Miranda Threlfall, Brendon Tijman, Eric Anderson, Sean Mccann, Gary Searing, and David Bradbeer. 2017. 'Diet and Prey Selection of Barn Swallows (Hirundo Rustica) at Vancouver International Airport'. *The Canadian Field-Naturalist* 131 (July): 26. https://doi.org/10.22621/cfn.v131i1.1777.

Le Robert. n.d. 'urbanisation - Définitions, synonymes, conjugaison, exemples | Dico en ligne Le Robert'. Accessed 21 March 2022. https://dictionnaire.lerobert.com/definition/urbanisation.

Leather, Simon R. 2008. *Insect Sampling in Forest Ecosystems*. John Wiley & Sons.

Liebhold, A., and B. Bentz. 2011. 'Insect Disturbance and Climate Change.' U.S. Department of Agriculture, Forest Service, Climate Change Resource Center. www.fs.usda.gov/ccrc/topics/insect-disturbance/insect-disturbance.

Liu, Mingxin, Laurence J. Clarke, Susan C. Baker, Gregory J. Jordan, and Christopher P. Burridge. 2020. 'A Practical Guide to DNA Metabarcoding for Entomological Ecologists'. *Ecological Entomology* 45 (3): 373–85. https://doi.org/10.1111/een.12831.

Liu, Zhifeng, Chunyang He, and Jianguo Wu. 2016. 'The Relationship between Habitat Loss and Fragmentation during Urbanization: An Empirical Evaluation from 16 World Cities'. *PLOS ONE* 11 (4): e0154613. https://doi.org/10.1371/journal.pone.0154613.

Lowry, Hélène, Alan Lill, and Bob B. M. Wong. 2013. 'Behavioural Responses of Wildlife to Urban Environments'. *Biological Reviews* 88 (3): 537–49. https://doi.org/10.1111/brv.12012.

Macher, Jan-Niklas, Till-Hendrik Macher, and Florian Leese. 2017. 'Combining NCBI and BOLD Databases for OTU Assignment in Metabarcoding and Metagenomic Datasets: The BOLD_NCBI _Merger'. *Metabarcoding and Metagenomics* 1 (December): e22262. https://doi.org/10.3897/mbmg.1.22262.

Magoc, T., and S. Salzberg. 2011. 'FLASH: Fast Length Adjustment of Short Reads to Improve Genome Assemblies.' https://www.geneious.com/plugins/flash/#citing.

Magoga, Giulia, Giobbe Forni, Matteo Brunetti, Aycan Meral, Alberto Spada, Alessio De Biase, and Matteo Montagna. 2022. 'Curation of a Reference Database of COI Sequences for Insect Identification through DNA Metabarcoding: COins'. *Database* 2022 (January): baac055. https://doi.org/10.1093/database/baac055.

Mansor, Mohammad Saiful, Muhammad Rasul Abdullah Halim, Nurul Abdullah, and Rosli Ramli. 2020. 'Barn Swallows Hirundo Rustica in Peninsular Malaysia: Urban Winter Roost Counts after 50 Years, and Dietary Segregation from House-Farmed Swiftlets Aerodramus Sp.' *The Raffles Bulletin of Zoology* 68 (April): 238–48. https://doi.org/10.26107/RBZ-2020-0021.

Marquina, Daniel, Rodrigo Esparza-Salas, Tomas Roslin, and Fredrik Ronquist. 2019. 'Establishing Arthropod Community Composition Using Metabarcoding: Surprising Inconsistencies between Soil Samples and Preservative Ethanol and Homogenate from Malaise Trap Catches'. *Molecular Ecology Resources* 19 (6). https://doi.org/10.1111/1755-0998.13071.

Martin. 2019. ‘UN Report: Nature’s Dangerous Decline “Unprecedented”; Species Extinction Rates “Accelerating”’. *United Nations Sustainable Development* (blog). 6 May 2019. https://www.un.org/sustainabledevelopment/blog/2019/05/nature-decline-unprecedented-report/.

Mata, Vanessa A., Sónia Ferreira, Rebeca M. Campos, Luís P. da Silva, Joana Veríssimo, Martin F. V. Corley, and Pedro Beja. 2021. ‘Efficient Assessment of Nocturnal Flying Insect Communities by Combining Automatic Light Traps and DNA Metabarcoding’. *Environmental DNA* 3 (2): 398–408. https://doi.org/10.1002/edn3.125.

Mazzucco, Rupert. n.d. ‘Species Formation: The Driver of Biodiversity’, 1.

McClenaghan, Beverly, Kevin C.R. Kerr, and Erica Nol. 2019. ‘Does Prey Availability Affect the Reproductive Performance of Barn Swallows (Hirundo Rustica) Breeding in Ontario, Canada?’ *Canadian Journal of Zoology* 97 (11): 979–87. https://doi.org/10.1139/cjz-2019-0001.

McClenaghan, Beverly, Erica Nol, and Kevin C R Kerr. 2019. ‘DNA Metabarcoding Reveals the Broad and Flexible Diet of a Declining Aerial Insectivore’. *The Auk* 136 (1): uky003. https://doi.org/10.1093/auk/uky003.

McIntyre, Nancy, J. Rango, William Fagan, and Stanley Faeth. 2001. ‘Ground Arthropod Community Structure in a Heterogeneous Urban Environment’. *Landscape and Urban Planning* 52 (January): 257–74. https://doi.org/10.1016/S0169-2046(00)00122-5.

Møller, Anders Pape. 1990. ‘Effects of Parasitism by a Haematophagous Mite on Reproduction in the Barn Swallow’. *Ecology* 71 (6): 2345–57. https://doi.org/10.2307/1938645.

Mondejar, Jeremy P., and Alejandro F. Tongco. 2019. ‘Near Infrared Band of Landsat 8 as Water Index: A Case Study around Cordova and Lapu-Lapu City, Cebu, Philippines’. *Sustainable Environment Research* 29 (1): 16. https://doi.org/10.1186/s42834-019-0016-5.

Moore, Janice, and Michael D. Breed. 2015. ‘Comparative Social Behavior-Swallows’. In *Animal Behavior*, Second. https://www.researchgate.net/publication/316216816_Animal_Behavior_Second_Edition.

Motokawa, Masaharu, and Hiroshi Kajihara. 2016. *Species Diversity of Animals in Japan*. Springer.

Nason, Lindsay D., Perri K. Eason, Margaret M. Carreiro, Amy Cherry, and Jacob Lawson. 2021. ‘Caterpillar Survival in the City: Attack Rates on Model Lepidopteran Larvae along an Urban-Rural Gradient Show No Increase in Predation with Increasing Urban Intensity’. *Urban Ecosystems* 24 (6): 1129–40. https://doi.org/10.1007/s11252-020-01091-2.

Nearing, Jacob T., Gavin M. Douglas, André M. Comeau, and Morgan G. I. Langille. 2018. ‘Denoising the Denoisers: An Independent Evaluation of Microbiome Sequence Error-Correction Approaches’. *PeerJ* 6 (August): e5364. https://doi.org/10.7717/peerj.5364.

New, Tim R. 2015. *Insect Conservation and Urban Environments*. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-21224-1.

Nielsen, Kaare M., Pål J. Johnsen, Douda Bensasson, and Daniele Daffonchio. 2007. ‘Release and Persistence of Extracellular DNA in the Environment’. *Environmental Biosafety Research* 6 (1–2): 37–53. https://doi.org/10.1051/ebr:2007031.

Noël, Grégoire, Arnaud Mestrez, Philippe Lejeune, Frédéric Francis, Masayuki Miwa, Koichi Uehara, and Ayako Nagase. 2021. ‘Pollen Meta-Barcoding Reveals Foraging Preferences of Honeybees (Apis

Mellifera L.) along Space-Time Gradient in Japan'. bioRxiv. https://doi.org/10.1101/2021.08.05.455320.

Nørgaard, Louise, Carsten Riis Olesen, Kristian Trøjelsgaard, Cino Pertoldi, Jeppe Lund Nielsen, Pierre Taberlet, Aritz Ruiz-González, Marta De Barba, and Laura Iacolina. 2021. 'EDNA Metabarcoding for Biodiversity Assessment, Generalist Predators as Sampling Assistants'. *Scientific Reports* 11 (1): 6820. https://doi.org/10.1038/s41598-021-85488-9.

Oiseaux.net. n.d. 'Classement alphabétique de la liste des espèces de la famille des Hirundinidae - 89 espèces répertoriées. - aves'. Accessed 23 February 2022. https://www.oiseaux.net/aves/hirundinidae.html.

Oksanen, Jari, Gavin L. Simpson, F. Guillaume Blanchet, Roeland Kindt, Pierre Legendre, Peter R. Minchin, R. B. O'Hara, et al. 2022. 'Community Ecology Package'. https://cran.r-project.org/web/packages/vegan/vegan.pdf.

Orłowski, Grzegorz, and Jerzy Karg. 2011. 'Diet of Nestling Barn Swallows Hirundo Rustica in Rural Areas of Poland'. *Open Life Sciences* 6 (6): 1023–35. https://doi.org/10.2478/s11535-011-0070-4.

———. 2013. 'Partitioning the Effects of Livestock Farming on the Diet of an Aerial Insectivorous Passerine, the Barn Swallow *Hirundo Rustica*'. *Bird Study* 60 (1): 111–23. https://doi.org/10.1080/00063657.2012.748717.

O'Rourke, Devon R., Nicholas A. Bokulich, Michelle A. Jusino, Matthew D. MacManes, and Jeffrey T. Foster. 2020. 'A Total Crapshoot? Evaluating Bioinformatic Decisions in Animal Diet Metabarcoding Analyses'. *Ecology and Evolution* 10 (18): 9721–39. https://doi.org/10.1002/ece3.6594.

Pancerasa, Mattia, Roberto Ambrosini, Nicola Saino, and Renato Casagrandi. 2018. 'Barn Swallows Long-Distance Migration Occurs between Significantly Temperature-Correlated Areas'. *Scientific Reports* 8 (1): 12359. https://doi.org/10.1038/s41598-018-30849-0.

Passy, Paul, and Sylvain Thery. 2018. 'Utilisation Des Modules SAGA GIS Dans QGIS'. In *QGIS et Outils Génériques*, 1:119–63. Utilisation de QGIS En Télédétection. ISTE Editions. https://hal.archives-ouvertes.fr/hal-03102527.

Pettorelli, Nathalie, Jon Olav Vik, Atle Mysterud, Jean-Michel Gaillard, Compton J. Tucker, and Nils Chr. Stenseth. 2005. 'Using the Satellite-Derived NDVI to Assess Ecological Responses to Environmental Change'. *Trends in Ecology & Evolution* 20 (9): 503–10. https://doi.org/10.1016/j.tree.2005.05.011.

Pillsbury, Finn C., and James R. Miller. 2008. 'Habitat and Landscape Characteristics Underlying Anuran Community Structure Along an Urban–Rural Gradient'. *Ecological Applications* 18 (5): 1107–18. https://doi.org/10.1890/07-1899.1.

Porter, Teresita M., and Mehrdad Hajibabaei. 2020. 'Putting COI Metabarcoding in Context: The Utility of Exact Sequence Variants (ESVs) in Biodiversity Analysis'. *Frontiers in Ecology and Evolution* 8. https://www.frontiersin.org/article/10.3389/fevo.2020.00248.

Preston, Mark, Martin Fritzsche, and Paul Woodcock. 2022. 'Understanding and Mitigating Errors and Biases in Metabarcoding: An Introduction for Non-Specialists', April, 29.

Purvis, Andy, Kate E. Jones, and Georgina M. Mace. 2000. 'Extinction'. *BioEssays* 22 (12): 1123–33. https://doi.org/10.1002/1521-1878(200012)22:12<1123::AID-BIES10>3.0.CO;2-C.

QIIME 2 development team. 2022. ‘Demux — QIIME 2 2022.2.0 Documentation’. 2022. https://docs.qiime2.org/2022.2/plugins/available/demux/.

Raclariu-Manolică, Ancuţa Cristina, Jarl Andreas Anmarkrud, Marcin Kierczak, Nima Rafati, Birgitte Lisbeth Graae Thorbek, Audun Schrøder-Nielsen, and Hugo J. de Boer. 2021. ‘DNA Metabarcoding for Quality Control of Basil, Oregano, and Paprika’. *Frontiers in Plant Science* 12. https://www.frontiersin.org/articles/10.3389/fpls.2021.665618.

Rahimi, Ehsan, Shahindokht Barghjelveh, and Pinliang Dong. 2021. ‘Quantifying How Urban Landscape Heterogeneity Affects Land Surface Temperature at Multiple Scales’. *Journal of Ecology and Environment* 45 (1): 22. https://doi.org/10.1186/s41610-021-00203-z.

Randa, Lynda A., and John A. Yunger. 2006. ‘Carnivore Occurrence along an Urban-Rural Gradient: A Landscape-Level Analysis’. *Journal of Mammalogy* 87 (6): 1154–64. https://doi.org/10.1644/05-MAMM-A-224R2.1.

Rasul, Azad, Heiko Balzter, Gaylan R. Faqe Ibrahim, Hasan M. Hameed, James Wheeler, Bashir Adamu, Sa’ad Ibrahim, and Peshawa M. Najmaddin. 2018. ‘Applying Built-Up and Bare-Soil Indices from Landsat 8 to Cities in Dry Climates’. *Land* 7 (3): 81. https://doi.org/10.3390/land7030081.

Rennstam Rubbmark, Oskar, Daniela Sint, Nina Horngacher, and Michael Traugott. 2018. ‘A Broadly Applicable COI Primer Pair and an Efficient Single-Tube Amplicon Library Preparation Protocol for Metabarcoding’. *Ecology and Evolution* 8 (24): 12335–50. https://doi.org/10.1002/ece3.4520.

Ribeiro, Ângela M., Ben Smit, and M. Thomas P. Gilbert. 2019. ‘31° South: Dietary Niche of an Arid-Zone Endemic Passerine’. *Environmental DNA* 1 (2): 109–18. https://doi.org/10.1002/edn3.13.

Richardson, Rodney T., Tyler D. Eaton, Chia-Hua Lin, Garrett Cherry, Reed M. Johnson, and Douglas B. Sponsler. 2021. ‘Application of Plant Metabarcoding to Identify Diverse Honeybee Pollen Forage along an Urban–Agricultural Gradient’. *Molecular Ecology* 30 (1): 310–23. https://doi.org/10.1111/mec.15704.

Ritchie, Hannah, and Max Roser. 2018. ‘Urbanization’. *Our World in Data*, June. https://ourworldindata.org/urbanization.

———. 2021. ‘Extinctions’. *Our World in Data*, April. https://ourworldindata.org/extinctions.

Roces-Díaz, José V., Carlos Cabo, Covadonga Prendes, Celestino Ordoñez, and Cristina Santín. 2020. ‘Automatic Delineation of Forest Patches in Highly Fragmented Landscapes Using Coloured Point Clouds’. *Forests* 11 (2): 198. https://doi.org/10.3390/f11020198.

Rowse, E. G., D. Lewanzik, E. L. Stone, S. Harris, and G. Jones. 2016. ‘Dark Matters: The Effects of Artificial Lighting on Bats’. In *Bats in the Anthropocene: Conservation of Bats in a Changing World*, edited by Christian C. Voigt and Tigga Kingston, 187–213. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-25220-9_7.

Ruby, Emily. n.d. ‘How Urbanization Affects the Water Cycle’. California Water & Land Use Partnership. Accessed 8 November 2022. https://oehha.ca.gov/media/downloads/ecotoxicology/fact-sheet/watercyclefacts_0.pdf.

Safran, R. 2010. ‘Barn Swallows: Sexual and Social Behavior’. In *Encyclopedia of Animal Behavior*, 139–44. Elsevier. https://doi.org/10.1016/B978-0-08-045337-8.00036-X.

Samways, Michael J., Philip S. Barton, Klaus Birkhofer, Filipe Chichorro, Charl Deacon, Thomas Fartmann, Caroline S. Fukushima, et al. 2020. ‘Solutions for Humanity on How to Conserve Insects’. *Biological Conservation* 242 (February): 108427. https://doi.org/10.1016/j.biocon.2020.108427.

Saura, Santiago, and Javier Martinez-Mlian. 2001. ‘Sensitivity of Landscape Pattern Metrics to Map Spatial Extent’. *PHOTOGRAMMETRIC ENGINEERING*, 10.

Savignac, Carl. 2011. ‘Barn Swallow (Hirundo Rustica): COSEWIC Assessment and Status Report 2011’. Assessments;research. 3 November 2011. https://www.canada.ca/en/environment-climate-change/services/species-risk-public-registry/cosewic-assessments-status-reports/barn-swallow-2011.html.

Scasta, J.D., T. Jorns, J.D. Derner, S. Lake, D.J. Augustine, J.L. Windh, and T.L. Smith. 2019. ‘Validation of DNA Metabarcoding of Fecal Samples Using Cattle Fed Known Rations’. *Animal Feed Science and Technology* 255 (August): 114219. https://doi.org/10.1016/j.anifeedsci.2019.114219.

Schindler, Stefan, Henrik von Wehrden, Kostas Poirazidis, Thomas Wrbka, and Vassiliki Kati. 2013. ‘Multiscale Performance of Landscape Metrics as Indicators of Species Richness of Plants, Insects and Vertebrates’. *Ecological Indicators*, Linking landscape structure and biodiversity, 31 (August): 41–48. https://doi.org/10.1016/j.ecolind.2012.04.012.

Schmiedová, Lucie, Oldřich Tomášek, Hana Pinkasová, Tomáš Albrecht, and Jakub Kreisinger. 2022. ‘Variation in Diet Composition and Its Relation to Gut Microbiota in a Passerine Bird’. *Scientific Reports* 12 (1): 3787. https://doi.org/10.1038/s41598-022-07672-9.

Senay, G. B., and R. L. Elliott. 2000. ‘Combining AVHRR-NDVI and Landuse Data to Describe Temporal and Spatial Dynamics of Vegetation’. *Forest Ecology and Management* 128 (1): 83–91. https://doi.org/10.1016/S0378-1127(99)00275-3.

Snapp, Barbara Denniston. 1976. ‘Colonial Breeding in the Barn Swallow (Hirundo Rustica) and Its Adaptive Significance’. *The Condor* 78 (4): 471–80. https://doi.org/10.2307/1367096.

Sousa, Lara Loureiro de, Sofia Marques Silva, and Raquel Xavier. 2019. ‘DNA Metabarcoding in Diet Studies: Unveiling Ecological Aspects in Aquatic and Terrestrial Ecosystems’. *Environmental DNA* 1 (3): 199–214. https://doi.org/10.1002/edn3.27.

Spence, Austin R, Erin E Wilson Rankin, and Morgan W Tingley. 2022. ‘DNA Metabarcoding Reveals Broadly Overlapping Diets in Three Sympatric North American Hummingbirds’. *Ornithology* 139 (1): ukab074. https://doi.org/10.1093/ornithology/ukab074.

Su, Zhimin, Xiaoma Li, Weiqi Zhou, and Zhiyun Ouyang. 2015. ‘Effect of Landscape Pattern on Insect Species Density within Urban Green Spaces in Beijing, China’. *PLOS ONE* 10 (3): e0119276. https://doi.org/10.1371/journal.pone.0119276.

Suarez-Rubio, Marcela, and Renata Krenn. 2018. ‘Quantitative Analysis of Urbanization Gradients: A Comparative Case Study of Two European Cities’. *Journal of Urban Ecology* 4 (1). https://doi.org/10.1093/jue/juy027.

Sugiura, Shinji, and Kazuo Yamazaki. 2014. ‘Caterpillar Hair as a Physical Barrier against Invertebrate Predators’. *Behavioral Ecology* 25 (4): 975–83. https://doi.org/10.1093/beheco/aru080.

Suzuki, M. T., and S. J. Giovannoni. 1996. ‘Bias Caused by Template Annealing in the Amplification of Mixtures of 16S RRNA Genes by PCR’. *Applied and Environmental Microbiology* 62 (2): 625–30. https://doi.org/10.1128/aem.62.2.625-630.1996.

Svenningsen, Cecilie S., Tobias Guldberg Frøslev, Jesper Bladt, Lene Bruhn Pedersen, Jonas Colling Larsen, Rasmus Ejrnæs, Camilla Fløjgaard, et al. 2021. 'Detecting Flying Insects Using Car Nets and DNA Metabarcoding'. *Biology Letters* 17 (3): 20200833. https://doi.org/10.1098/rsbl.2020.0833.

Tonissen, Patrick. 2019. 'How Did IPBES Estimate "1 Million Species At Risk of Extinction" in #GlobalAssessment Report'. IPBES Secretariat. 22 May 2019. https://ipbes.net/news/how-did-ipbes-estimate-1-million-species-risk-extinction-globalassessment-report.

Trevelline, Brian K., Steven C. Latta, Leesia C. Marshall, Tim Nuttle, and Brady A. Porter. 2016. 'Molecular Analysis of Nestling Diet in a Long-Distance Neotropical Migrant, the Louisiana Waterthrush (Parkesia Motacilla)'. *The Auk* 133 (3): 415–28. https://doi.org/10.1642/AUK-15-222.1.

Trevelline, Brian K., Tim Nuttle, Brandon D. Hoenig, Nathan L. Brouwer, Brady A. Porter, and Steven C. Latta. 2018. 'DNA Metabarcoding of Nestling Feces Reveals Provisioning of Aquatic Prey and Resource Partitioning among Neotropical Migratory Songbirds in a Riparian Habitat'. *Oecologia* 187 (1): 85–98. https://doi.org/10.1007/s00442-018-4136-0.

Turner, Angela. 2010. *The Barn Swallow*. Bloomsbury Publishing.

———. 2015. *Swallow*. Reaktion Books.

Turner, Angela K. 1982. 'Optimal Foraging by the Swallow (Hirundo Rustica, L): Prey Size Selection'. *Animal Behaviour* 30 (3): 862–72. https://doi.org/10.1016/S0003-3472(82)80160-7.

———. 1983. 'Time and Energy Constraints on the Brood Size of Swallows, Hirundo Rustica, and Sand Martins, Riparia Riparia'. *Oecologia* 59 (2): 331–38. https://doi.org/10.1007/BF00378858.

Uiterwaal, Stella F., and John P. DeLong. 2020. 'Using Patterns in Prey DNA Digestion Rates to Quantify Predator Diets'. *Molecular Ecology Resources* 20 (6): 1723–32. https://doi.org/10.1111/1755-0998.13231.

Van de Voorde, T., T. De Roeck, and F. Canters. 2009. 'A Comparison of Two Spectral Mixture Modelling Approaches for Impervious Surface Mapping in Urban Areas'. *International Journal of Remote Sensing* 30 (18): 4785–4806. https://doi.org/10.1080/01431160802665918.

Van Nuland, Michael E., and W. Lindsay Whitlow. 2014. 'Temporal Effects on Biodiversity and Composition of Arthropod Communities along an Urban–Rural Gradient'. *Urban Ecosystems* 17 (4): 1047–60. https://doi.org/10.1007/s11252-014-0358-z.

Vavrek, Matthew J. 2010. 'Fossil: Palaeoecological and Palaeogeographical Analysis Tools', 16.

Wang, Jhih-Syuan, Mao-Ning Tuanmu, and Chih-Ming Hung. 2021. 'Effects of Artificial Light at Night on the Nest-Site Selection, Reproductive Success and Behavior of a Synanthropic Bird'. *Environmental Pollution* 288 (November): 117805. https://doi.org/10.1016/j.envpol.2021.117805.

Wheeler, David, and Medha Bhagwat. 2007. *BLAST QuickStart*. *Comparative Genomics: Volumes 1 and 2*. Vol. 1 and 2. Humana Press. https://www.ncbi.nlm.nih.gov/books/NBK1734/.

Willsey, Telsa, Syama Chatterton, and Héctor Cárcamo. 2017. 'Interactions of Root-Feeding Insects with Fungal and Oomycete Plant Pathogens'. *Frontiers in Plant Science* 0. https://doi.org/10.3389/fpls.2017.01764.

Withaningsih, Susanti, and Muhammad Rabbany. 2019. 'Correlation between Some Landscape Metrics and Insect Species Richness in Coffee Agroforests in Pangalengan Subdistrict, Bandung

District, West Java, Indonesia'. *Biodiversitas Journal of Biological Diversity* 20 (October): 3075–85. https://doi.org/10.13057/biodiv/d201042.

Wood, David. n.d. 'Climate: Classifications, Distribution & Causes'. Study.Com. Accessed 19 June 2022. https://study.com/academy/lesson/climate-classifications-distribution-causes.html.

Youngsteadt, Elsa, Andrew F. Ernst, Robert R. Dunn, and Steven D. Frank. 2017. 'Responses of Arthropod Populations to Warming Depend on Latitude: Evidence from Urban Heat Islands'. *Global Change Biology* 23 (4): 1436–47. https://doi.org/10.1111/gcb.13550.

Zaller, Johann G., Laura Simmer, Nadja Santer, James Tabi Tataw, Herbert Formayer, Erwin Murer, Johannes Hösch, and Andreas Baumgarten. 2014. 'Future Rainfall Variations Reduce Abundances of Aboveground Arthropods in Model Agroecosystems with Different Soil Types'. *Frontiers in Environmental Science* 2. https://www.frontiersin.org/article/10.3389/fenvs.2014.00044.

Zenker, Mauricio M., Alexandre Specht, and Vera G. Fonseca. 2020. 'Assessing Insect Biodiversity with Automatic Light Traps in Brazil: Pearls and Pitfalls of Metabarcoding Samples in Preservative Ethanol'. *Ecology and Evolution* 10 (5): 2352–66. https://doi.org/10.1002/ece3.6042.

Zhang, Yu, and Gang Feng. 2018. 'Distribution Pattern and Mechanism of Insect Species Diversity in Inner Mongolia'. *Biodiversity Science* 26 (7): 701. https://doi.org/10.17520/biods.2018144.

Zhao, Yanyan, Yu Liu, Elizabeth S. C. Scordato, Myung-Bok Lee, Xiaoying Xing, Xinyuan Pan, Yang Liu, Rebecca J. Safran, and Emilio Pagani-Núñez. 2021. 'The Impact of Urbanization on Body Size of Barn Swallows Hirundo Rustica Gutturalis'. *Ecology and Evolution* 11 (1): 612–25. https://doi.org/10.1002/ece3.7088.