Pollinators in complex landscapes: modelling and mapping the distribution of wild bees and hoverflies in the Netherlands

Moens, M.

Citation

Moens, M. (2025, December 10). *Pollinators in complex landscapes: modelling and mapping the distribution of wild bees and hoverflies in the Netherlands*. Retrieved from https://hdl.handle.net/1887/4284738

Chapter 1: General introduction

Wild bees and hoverflies, like all species, each have their own relationship with the landscape they live in, especially its environmental and climate conditions. This is partly based on their own traits like body size, life history and thermal tolerance, and partly based on the interactions with other species, such as their relations with hosts or host plants. Together, this results in partly predictable and partly stochastic patterns of occurrence of these wild pollinators in our landscapes. The complexity of a landscape presents itself at a small scale, such as the presence of plants and landscape elements (LEs), but also at a larger scale, such as the presence of urban green or other land uses. Capturing the relation between this landscape complexity and the occurrence of pollinators and other species, still remains a challenge. In my thesis, I will address the identified knowledge gaps and methodological challenges to further both our scientific understanding of how pollinators thrive and survive in complex landscapes and provide some avenues for better underpinning of conservation and land management strategies. To understand how pollinators are distributed across a landscape, it is essential to first examine the ecological niche each species occupies. This includes the specific role the species plays within the ecosystem, as well as the environmental conditions and resources it requires to survive and reproduce. In this introduction, I will outline the general concept of the (fundamental and realized) niche (section 1.1). Then, I explain how I can model and predict the distribution of species in a quantitative manner with species distribution models (section 1.2) to generate potential occurrence maps (section 1.3). Developing models is important, as occurrence data is sparse, unevenly distributed, and incomplete for most pollinator species, and insufficient alone to understand how pollinator communities are distributed within our landscapes. Modelling species occurrences based on incomplete data, both on the species and the factors they depend on, is a complex and challenging task (section 1.4). Ultimately, the pollinator distribution maps (and other information from models) can be important inputs in pollinator conservation strategies (section 1.5).

1.1 A brief history of niche theory

Modern ecologists have slightly different definitions of the ecological niche (Sales et al. 2021), but the niche generally refers to a species' distribution within environmental conditions that determine where it can thrive and its role within the ecosystem. The word niche was first mentioned in ecological terms in the paper by Joseph Grinnell in 1924 (Grinnell 1924) and his concept of a niche was built on concepts in earlier research (Grinnell 1917). He described the idealized distributional unit of a species or the distribution of a species with a lack of interactions with other species (Vandermeer 1972), only bound by physical and climatic restrictions. He did consider the interaction with other organisms but defined the niche as a space that was occupied before considering biotic interactions: an idealized distribution (Vandermeer 1972). Charles Elton extended the focus from only physical limitations to the place a species occupies regarding its food habits and other biotic interactions (Elton 1927). He realized how, in different places of the world, there were organisms with similar roles in the ecosystem, such as tick-eating birds that could be found in different ecosystems (Vandermeer 1972). Both Grinnell and later Gause mentioned that two species cannot occur together if they occupy a similar niche (Vandermeer 1972). The niche defined by Grinell or the Grinellian niche could be interpreted as an early notion of the fundamental niche and its cousin the Eltonian niche as the realized niche (Vandermeer 1972). Thirty years later after the first description of the Eltonian niche, a more quantitative definition of the niche as a multidimensional volume was introduced by Hutchinson (Hutchinson 1957). He introduced the term hypervolume or a space within limits of environmental variables, that allows the species a positive population growth (Hutchinson 1957). For example, a species may have a minimum and maximum temperature needed for survival and only within the limits of these temperature values can the population of this species have a population growth.

The difference between a niche and the spatial distribution of a species is that the first one relates to the requirements for the fitness of a species (Hirzel and Le Lay 2008) and the second one is the translation of this fitness within space (see fig. 1). The realized niche is the space a species occupies within a landscape and for that realization - besides the suitable abiotic conditions - several conditions need to be fulfilled. Factors that determine the presence of a species can be grouped in accessibility in terms of dispersal or movement of a species, biotic factors, abiotic conditions and evolutionary adaptability of a species to new environments (Soberon and Peterson 2005). These first three factors can be represented in the biotic, abiotic and movement diagram (BAM diagram; Soberon & Peterson, 2005; see fig. 2). An example of these abiotic conditions is climate variables often used in climate envelope models (Hijmans and Graham 2006). These abiotic conditions delineate the area with a positive population growth and these suitable conditions correspond to the fundamental niche of a species. The realized niche is the area of overlap between suitable abiotic conditions, biotic interactions and an area that is accessible to a species (see fig. 2). The biotic factors can include influences from different types of interactions, such as competition, facilitation, mutualism and predation (Wisz et al. 2013). An example of the geographic influence biotic factors may have, is the example of the puma (*Puma concolor*), whose geographic range disappeared from the Old World most likely with the expansion of the leopard (*Puma pardoides*) (Wisz et al. 2013). An area that has suitable abiotic conditions and biotic factors, but is barriered in any way for the movement of a species, is the invasible area. An example of the influence of dispersal capabilities is that of an invasive species. *Bombus terrestris* is a bumblebee native to the Mediterranean region extending upwards North (Rasmont et al. 2008). Geographic barriers can prevent it from spreading to other areas, even though suitable abiotic conditions and biotic factors may present themselves across the ocean. However, after introduction of *Bombus terrestris* by humans, this bumblebee species is now invasive in Chile and replacing local species of bumblebees (Morales et al. 2013), depicting the important role that movement and dispersal plays in the presence of a species at a given location.

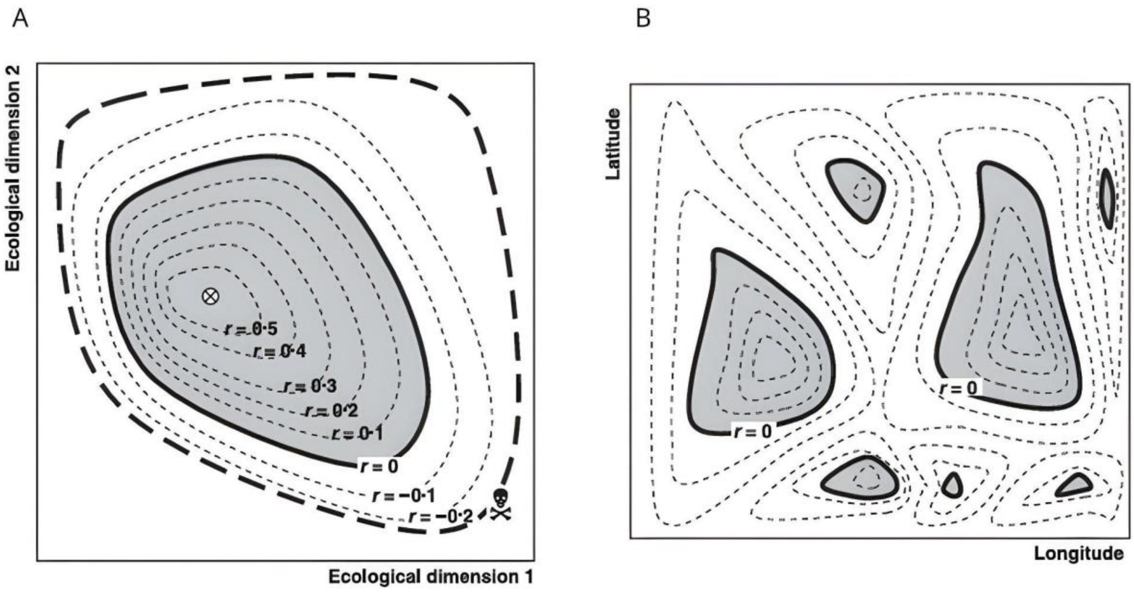

Figure 1: Graphical representation of the multidimensional space of the Hutchinsonian niche (figure 1A) and a spatial representation (figure 1B) adapted from Hirzel & Le Lay, 2008. The grey areas indicate the values in which the species shows a population growth (r>0). Figure 1B represents the locations (in longitude and latitude) that include these conditions.

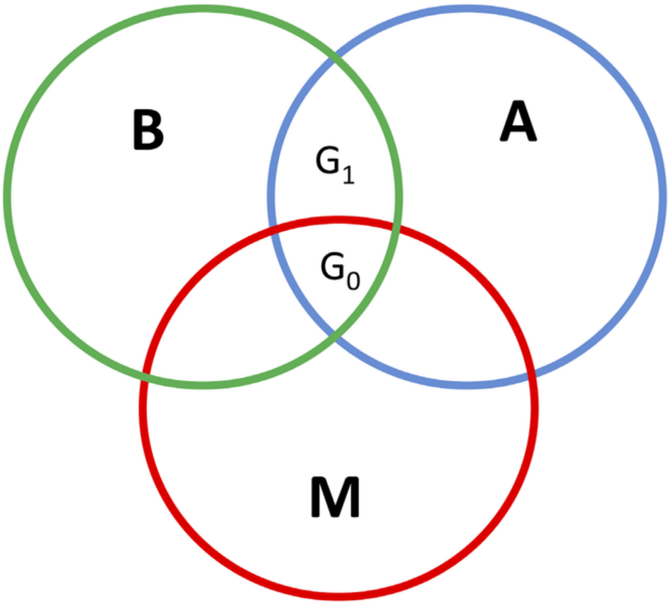

Figure 2: Schematic representation of the BAM diagram (Soberon and Peterson 2005) adapted from Ríos-Muñoz et al., 2021. The letter A represents the area with suitable abiotic conditions delineating the fundamental niche, the letter B represents the area with suitable biotic interactions and M is the area accessible to dispersal or colonization. G_0 is the realized niche and G_1 is the invasible area.

1.2 From niche to model to map

Following the more quantitative definition of a Hutchinsonian niche, tools emerged to estimate a species' niche and its distribution within a landscape. Species distribution models (SDMs) are numerical tools, combining observational data with environmental predictors (Elith and Leathwick 2009). These models either have a correlative approach in which the relationship between species and the environment is implicit, and a process-based approach in which the relations are explicitly modelled (Dormann et al. 2012). The models with a correlative approach combine environmental predictors with species occurrences to estimate correlations between the species and the environment. These predictors, which are explanatory variables of the SDMs, can be categorized into scenopoetic and bionomic variables (Soberón 2010). Scenopoetic variables, first introduced in earlier research (Hutchinson 1978), influence the fitness of the modelled species but are not affected by its presence (Soberón 2010). An example of a scenopoetic variable is mean daily temperature, that plays a key role in the survival of species and is not affected or minimally affected by the presence of a species. In contrast, bionomic variables can be influenced by the presence of the species being modelled. For instance, when modelling the distribution of red foxes researchers included the distance towards areas with rabbits as a bionomic predictor (Rosalino et al. 2010). Since rabbit populations can be impacted by the red fox population, this is a bionomic variable. The classification of biotic factors as bionomic variables is not always so clear as a biotic factor that is affected minimally by the modelled species, could arguably classify more as a scenopoetic variable (Anderson 2017).

SDMs utilize various algorithms to predict the spatial suitability of a landscape, allowing for inferences about a species' niche and predictions across both time and space. Many different types of models can be integrated spatially as SDMs, such as random forest models (Valavi et al. 2021), support vector machines (Drake et al. 2006), generalized linear models, generalized additive models (Guisan et al. 2002), extreme gradient boosting (Herdter Smith 2019) and maximum entropy modelling (Maxent; Phillips et al., 2006). Additionally, it is possible to combine different algorithms with ensemble modelling (Hao et al. 2020) or in a hierarchical framework (Perennes et al. 2021).
More complex machine learning algorithms may result in more accurate predictions, but with a lower interpretability (Ryo et al. 2021). Additionally, certain algorithms may be more effective for species with a narrow geographic distribution (Aguirre-Gutiérrez et al. 2013). These algorithms can vary significantly in their predictive power depending on ecological traits, distributional patterns and data quality of the target species (Valavi et al. 2022).

SDMs do not produce a single, uniform output; instead, their results vary depending on the modelling techniques used. Generally, the aim of the SDMs is to broaden knowledge on biogeographical and ecological theory as well as conservation ecology (Guillera-Arroita et al. 2015). In more practical terms, the output of correlative SDMs often includes environmental correlates, variable importance (Harisena et al. 2021) and habitat suitability maps or binary presence/absence maps (Meynard and Kaplan 2012). In simpler models, relationships between the environment and species may be represented as linear, facilitating interpretation of the relationship between the environment and the species. For example, the simpler models used in this thesis facilitated the interpretation of the ecology of how threatened and non-threatened bees' response to different land uses (see Chapter 2). However, in more complex machine learning models, such as random forest algorithms, the relationships are often nonlinear and more challenging to interpret (Ryo et al. 2021). The predictive map outputs are often expressed as habitat suitability. Probability of occurrence is difficult to estimate. Whether the species is observed at a location, depends on the probability of it occurring at a site, the detection probability and the probability that the specific site is sampled (Yackulic et al. 2013). Detection probability can differ with environmental covariates and the probability of a site being sampled is influenced by sampling biases (Guillera-Arroita et al. 2015). While some (immobile) organisms are relatively easy to observe (e.g. plants) and have a high detection probability, other species can easily be missed and/or are more difficult to observe (e.g. flying insects). For plants, projected maps can therefore represent more closely the probability of occurrence than for flying insects, which indicate more of a relative habitat suitability. Habitat suitability maps are often converted to binary values, resulting in presence and absence maps using a specified threshold (Guillera-Arroita et al. 2015). This threshold is chosen based on established criteria designed to balance predictive accuracy and ecological relevance. Despite these methods, converting continuous suitability scores to binary values may still result in some loss of information compared to the maps with the raw output values (Guillera-Arroita et al. 2015).

1.3 Mapping pollinators in a complex landscape

The complexity of a landscape presents itself at different spatial scales, including both abiotic conditions, such as soil characteristics, and biotic factors, such as tree species composition. Pollinators rely on diverse habitat characteristics within landscapes, which influence their behaviour, foraging patterns, and overall diversity. Pollinators comprise a wide range of pollinating animals and using estimates of global diversity around 1 in 10 of terrestrial animals are pollinators (Ollerton 2017). Many insect groups, such as bees, flies and beetles are capable of pollination (Rader et al. 2016). Within the landscape, pollinators play a vital role, they are essential for the reproduction of wild plants and as a food source for other organisms. Around 90% of all flowering plants are dependent on animal pollination (Tong et al. 2023). Besides the pollination of wild plants, pollinators fulfil an important ecosystem service: crop pollination. The estimated economic value of this ecosystem service across the world is 153 billion euros (Khalifa et al. 2021). The distribution of these organisms within the (agricultural) landscape is therefore especially important. The most important crop in terms of surface area within the Netherlands is apple (Scheper et al. 2014a). Other important crops in the Netherlands include pear and rapeseed (Scheper et al. 2014a). The pollination needs of these crops are met by various pollinators, ranging from smaller bees which typically forage within a couple of hundred meters of their nest (Zurbuchen et al. 2010), and migratory hoverflies, which can travel much further distances (Doyle et al. 2020).

To accommodate these important organisms within the landscape, their specific ecological needs must be met; landscapes provide nesting habitat, foraging grounds, overwintering sites and shelter (Cole et al. 2017). The local landscape structure, therefore, directly influences the abundance and diversity of pollinators (Cole et al. 2017). In agricultural areas the presence of semi-natural habitats provides important foraging plants for the pollinators. These feeding resources support various bees from oligolectic or specialist species that can be characterized by their dependency on pollen from a narrow selection of plants to polylectic or generalist species with broader foraging preferences (Robertson 1925).
The resources within a landscape required by pollinators are related to landscape structure and landscape elements (LEs). LEs, such as hedges, treelines and flower strips, can provide resources for pollinators throughout the season and provide large amounts of food per square meter (Timberlake et al. 2019). Hedges can be important in food provisioning for

pollinators and nesting opportunities (von Königslöw et al. 2022, Bożek et al. 2023). They commonly present shrub species providing resources throughout the year in early spring and later in the season (Bożek et al. 2023). Other LEs include field margins, banks of ditches and rivers, treelines and ponds, these can also increase pollinator diversity by providing necessary resources (Musters et al. 2009, Cole et al. 2020, von Königslöw et al. 2022).

1.4 Current modelling challenges

There are several challenges to overcome in order to improve the model evaluation criteria used in this thesis: model performance, ecological relevance and the applicability of SDMs. The model performance criterion consists of quantitative measures, evaluating the ability of the model to predict observations within the data. The ecological relevance relates to whether important ecological aspects of a species are included within the models. While the first two criteria are analysed throughout the chapters, the third criterion, the applicability of the SDMs, is addressed in the general discussion in Chapter 6 and relates to the usefulness of modelling results for stakeholders. Modelling challenges included in other studies, are appropriate integration of biotic factors (Anderson 2017, Srivastava et al. 2019), dispersal (Srivastava et al. 2019), reliability of the models (Srivastava et al. 2019) and extrapolating to future scenarios (e.g. climate and land use change) (Franklin 2013). Furthermore, the fundamental assumptions of SDMs—such as the idea that species are in equilibrium with their environments—likely does not apply in ecosystems undergoing rapid change. For example, the Dutch agricultural landscape, which is a focus of this thesis, is subject to rapid change meaning that observed pollinators may not represent stable populations. Addressing these challenges requires innovative modelling approaches and a more integrative use of ecological data. The integration of biotic factors has been attempted but with varying success (Heikkinen et al. 2007, Giannini et al. 2013, Wisz et al. 2013, Mpakairi et al. 2017). Although biotic interactions have been integrated for several species, there remains a significant gap in our understanding of how scale and other factors influence the effectiveness of incorporating these interactions. The third component of a factor leading to the presence of a species in the BAM diagram is movement/dispersal. Adding this component is challenging but has been attempted by several studies (Engler et al. 2012, Shipley et al. 2022).

Another challenge for improving model performance, ecological relevance and applicability of SDMs is how best to measure, report and integrate uncertainty. Uncertainty in SDMs can be divided into three different types of errors: measurement error, natural variability and incomplete knowledge on the ecology of the species or underlying processes (Gould et al. 2014). An example of the latter is missing data within the model or in other words an incomplete model. The measurement error relates to the way the data was collected and at which time periods and at what scale. If the ecology of a species is partly unknown not all the relevant environmental variables can be added. In general, at the landscape scale it is virtually impossible to capture all possible factors that influence a species' presence.
In this thesis, I address the challenges of integrating biotic factors within the SDMs (Chapter 3), and the absence of relevant data and model input variables in agricultural habitats (Chapter 4). Overall, across, the chapters I use relevant input data from bees and hoverflies to address the challenge of reducing model uncertainty, especially in terms of missing data that is relevant to the niche of a species.

1.5 Research questions

Modelling and mapping the distribution of pollinators across landscapes with tools such as SDMs, is essential for their conservation. These models and maps have found their use for many different species in identifying source populations, potential areas at risk for invasive species, areas for reintroduction and understanding threats and the requirements of a species. With the ultimate goal of providing tools for nature conservationists in complex landscapes and improving our knowledge on the distribution of pollinators, I formulated different research questions and sub-questions. Each sub-question is addressed in an individual chapter and can be formulated as follows:
How do pollinators live and thrive in the complex landscapes of the Netherlands?

1. Are threatened and non-threatened bees distributed differently among land use, climate and soil categories?
   To direct conservation efforts it is essential to understand the relationship between threatened bees, their environment and their distribution. In this chapter, I compare the spatial distribution between threatened and non-threatened bees. This research question addresses the challenge of including ecological relevant information in SDMs and species-specific variation in environmental responses.
2. What is the role of biotic factors, i.e. pollen source, host plant species and parasite hosts, in explaining distributions of wild bees?
   This chapter addresses the integration of biotic interactions within SDMs. Biotic interactions are a key factor in determining a species' presence and are often not included in SDMs. This case study evaluates model performance of models with and without biotic interactions for both plant-bee and parasite-host interactions.
3. What is the role of small linear landscape elements in explaining distributions of wild bees in an agricultural landscape?
   Even though agriculture comprises a large part of the Netherlands, detailed information on agricultural landscape characteristics (e.g. linear landscape elements), is rarely included in modelling species distributions. This chapter addresses the challenge of including ecological relevant, fine scale landscape information in an under sampled region by comparing model performance between models with and without small linear landscape elements.
4. How does the presence of landscape elements influence the bee and hoverfly community diversity across regions and what factors drive any difference?
   This chapter addresses the knowledge gap of what drives differences in pollinator communities within landscape elements across agricultural regions. This chapter differs from the others as it does not include SDMs. Instead, it refines the findings from Chapter 4, which focused on improving SDMs with landscape elements, by comparing the characteristics of these elements and their effects on bee and hoverfly diversity measures.

1.6 Importance of models for pollinator conservation

Many of the bee and hoverfly species are threatened and therefore modelling and mapping their distribution is especially important for their conservation. Of the 359 bee species in the Netherlands 181 of the bee species are on the Red List (Reemer 2018) and of the 317 hoverfly species, 146 species are on the Red List (Reemer et al. 2023). The most widely used Red Lists are those developed by the IUCN, serving as a tool to categorize species

based on their risk of extinction (Nieto 2014). These categories range from “Extinct” to “Least concern” with the following categories in between: Critically Endangered, Endangered, Vulnerable, Near Threatened (IUCN 2012). The Dutch Red List of bees and hoverflies is an adaptation using similar methods to estimate extinction risks for both bees (Reemer 2018) and hoverflies (Reemer et al. 2023) in the Netherlands.
The main drivers of the decline in pollinator numbers are habitat loss, pesticides, pollution, climate change, pests and pathogens (Brunet and Fragoso 2024). An additional factor for wild bee decline is the presence of managed honeybees (Weaver et al. 2022, Brunet and Fragoso 2024). In practice many bee species in the Netherlands and the rest of Europe have declined with the intensification of grasslands (Ekroos et al. 2020) and the decrease of flower rich grassland. An example is *Andrena labialis*, which is a bee species that specializes on plant species from the *Fabaceae* family, particularly clovers in grasslands. Due to the intensification and decrease in flower rich grasslands, it is now a threatened species on the Dutch Red List (Reemer 2018).

In the European Union (EU) several initiatives have been set up for the conservation of wild pollinators such as the pollinator initiative and indirect measures within the common agricultural policy (CAP). The pollinator initiative (European Commission 2023) is in line with the goals of the biodiversity strategy 2030, and it aims at improving the knowledge on the decline of pollinators and improving pollinator conservation (European Commission 2023). The CAP is a longstanding policy that started in 1962 and has developed since. In the CAP of 2023 to 2027 at least 3% of the agricultural area needs to be reserved for biodiversity and not-productive elements (European Commission 2022). Besides the different European initiatives, there have been several research projects such as MAPPY (*multisectoral analysis of climate and land use change impacts on pollinators, plant diversity and crops yields*), SPRING (*strengthening pollinator recovery through indicators and monitoring*) and ALARM (*Assessing Large-scale environmental Risks with tested Methods*). In general, the European initiatives and research projects require a better understanding of pollinator declines and their conservation. Both for understanding the decline and designing effective conservation efforts, it is paramount to understand the distribution of pollinators in their environment, both in relation with other species and the importance of habitat differences. Furthermore, more knowledge is needed on the integration of biotic interactions, such as plants, the distribution of threatened species and the importance of LEs and other semi-natural habitats. This thesis aims at broadening the knowledge on these topics and supports the ongoing efforts for the conservation of wild pollinators. The implementation of the knowledge gained throughout this thesis is discussed in section 6.5 in ‘Chapter 6: General discussion’.