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

Moens, M.

Chapter 6: General discussion

Wild bees and hoverflies in the Netherlands comprise a large number of species, each with their own niche, or set of abiotic conditions and biotic factors that allow them to thrive within the landscape. Even though niche theory and its applications have developed a long way since the first mention of “niche” a hundred years ago (Grinnell 1924), new challenges present themselves at every step. The aim of this thesis was to address key challenges for modelling the distribution of pollinators in complex landscapes. These challenges included the integration of biotic factors within the models and the introduction of input variables that represent the ‘complexity’ of the landscape and the ecological requirements of the bees and hoverflies living in it. Although the landscape has become more homogeneous, as seen in the intensification of agriculture, it still retains characteristics and features across both small and large scales, which together contribute to its complexity. Each chapter of this thesis addressed a different challenge and their key scientific contributions are described, as well as their contributions in terms of niche theory (section 6.1). The results presented across the different chapters can be extrapolated to other species and can infer recommendations for future modelling efforts across taxonomic groups (section 6.2). Even though SDMs can provide relevant information for pollinator conservation (see section 1.6) a gap remains between potential value for nature conservation and implementation by stakeholders (section 6.5).

6.1 Scientific contributions

Key findings of the chapters
This thesis addresses methodological challenges for mapping the distribution of pollinators, including the integration of biotic factors into models and the incorporation of ecologically relevant variables in agricultural regions. It also tackles the gap in our understanding of nature conservation and its support through interventions such as landscape elements and differences in distribution between threatened and non-threatened species. The overall aim of the thesis is to improve our understanding of how pollinators live and thrive in the Dutch landscape. As described in **Chapter 1** the theory behind the SDMs originates from the first definitions of different types of niches. Still, many of the challenges and characteristics of the niche theory are present today in the models. This thesis has made a small, but important, step towards a more robust mapping and modelling of species in space and time. The different chapters address the current challenges of modelling and mapping wild pollinators. The key findings of the chapters come from results of modelling pollinators but can infer knowledge for other taxonomic groups as well.
One of the main purposes of SDMs is as a tool for nature conservation and to understand how I can protect threatened species. In **Chapter 2**, I compared the species’ spatial distribution in relation to land use, climate and soil of both threatened and non-threatened wild bee species in the Netherlands. I found that threatened bees have a smaller area of extent, occupy areas with more extreme climatic conditions and benefit less from urban green than non-threatened species. Additionally, the threatened species were distributed among a smaller variety of natural land use types, having a stronger preference for nature types. The non-threatened species were distributed among multiple natural land use types and showed a more general preference for nature. The main negatively contributing land use types were the cover of pasture and agriculture, resulting in ‘deserts’ of predicted

biodiversity, or areas that generally are not predicted to harbour bee species. This chapter examines the differences between threatened and non-threatened bee species using a quantitative approach, employing models, and a holistic perspective by comparing a wide range of variables across most bee species in the Netherlands. As a result, the findings offer knowledge within the landscape context for different nature management policies.

Most SDMs model species in relation to climate variables but the integration of biotic factors within the models remains a challenge. The integration of biotic factors into SDMs has yielded mixed results, highlighting the need for further research on the factors that influence their significance within these models. In **Chapter 3** I integrate biotic interactions, more specifically preferred pollen plant species and hosts for parasites, within the modelling framework. Between models I compared performance, consisting of quantitative measures, evaluating the ability of the model to predict observations within the data. I found a statistically significant improvement for models including biotic interactions for plant specialists and generalists and parasitic bees. This model improvement was particularly clear for the preferred pollen host plant, as adding a random other pollinated plant would result in lower model performance. The same specificity for the importance of the known biotic interaction compared to randomized biotic interactions was found for the parasitic bees. I found that resolution of the data, degree of specialization of the bee, taxonomic resolution (species/genus) and distribution of the biotic interaction all played a role. For example, I found that specialist bees that depend on plants with a more limited distribution, showed a greater dependence on or correlation with their biotic factor. The correlations found with scale, degree of specialization and the distribution of the biotic factor may explain why other studies showed mixed success with the integration of biotic interactions within their models. These factors that play a role within the models provide important information for the integration of biotic interactions for other modellers.

The findings of predicted absences within the agricultural regions in the second chapter, begs the question whether these areas are indeed devoid of pollinators, or that there may be other underlying causes, such as biases or missing information on characteristics and features or the “complexity’ of the landscape. Improving SDM modelling methods is important in agricultural regions, as many crops are dependent on the pollination ecosystem service. In **Chapter 4** I found that agricultural regions may underpredict bee and hoverfly species if landscape elements (LE) are not considered. There was a bias in observation data towards other land uses than agriculture. I found that by integrating different LEs in the models, particularly herbaceous strips of land, the prediction maps within agricultural regions became more nuanced, showing ‘hotspots’ of biodiversity in the LEs around the agricultural fields and lower predicted biodiversity within the field. Observation data on bees and hoverflies were collected in agricultural regions to function as an independent evaluation dataset. The evaluation measures of the model performance improved when landscape element information was included.

In **Chapter 5** I compared regional differences in bee and hoverfly species composition within LEs in three agricultural regions in the Netherlands: the Ooijpolder, Alblasserwaard and the tulip fields near Lisse. In order to explain differences between LEs, I compared the characteristics of those LEs, such as dimensions and flower richness. In general, I found a higher species richness and abundance of bees and hoverflies within the LEs than within the fields, but this varied depending on the season and collection method (pan trap/ transect).

The number of species in the LEs correlated with region, season, flower richness in the herb layer, vegetation type and whether it was located near roads. The different regions had a group of common pollinators that were present in all three regions and a group of less common pollinators that were unique for the different regions.

In the context of niche theory

The results found in this thesis can be contextualized within niche theory and the theoretical framework of SDMs. The integration of biotic interactions, e.g. plant-pollinator and parasite-host relations, improves model evaluation measures, but raises issues concerning spatial scale and the influence of the modelled species on the biotic interaction. Our results showed that different factors like spatial scale and spatial distribution of the biotic interaction can influence the relevance of the biotic factor to the models. The BAM diagram was proposed by Soberon & Peterson, 2005 as the theoretical framework for SDMs and categorizes factors influencing the presence of a species in abiotic conditions, biotic factors and movement/dispersal (see figure 2 in Chapter 1.1). The actual distribution of a species depends on small-scale spatial and temporal factors, such as whether the climate is suitable or if the host plant is flowering. As a result, the general BAM diagram for a species can manifest differently across various contexts, reflecting spatial or temporal variations.
Abiotic conditions can influence the distribution of a species and their role in SDMs has changed with the availability of high-resolution data. The abiotic conditions required for a species to survive, were traditionally represented as aggregated data summarized over months, such as bioclimatic variables in many SDMs (Waltari et al. 2014). With the advent of high-resolution data, microclimates can now be represented by this high-resolution data (Lembrechts et al., 2019). As the resolution of the data increases, the data are able to represent landscape features in more detail. For example, a more detailed representation of annual rainfall would be local soil water content, which would represent a more suitable predictor for plants (Gardner et al. 2019). In the third chapter, I found that the high-resolution data within the agricultural landscape and the information on the LEs could improve model performance and provide a more nuanced view of the species richness within the landscape. Another study did not find consistent improvement in model performance for plants with high-resolution data compared with lower resolution data, except for plants dependent on specific micro-habitats, such as species occurring in screes (Pradervand et al. 2014). In general, there is a discrepancy between models with high-resolution data and with the same data aggregated at low-resolution (Meynard et al. 2023). For example, if a species is modelled with low-resolution data compared to high-resolution data, the predicted probability of occurrence increases with the coarser resolution and the resulting optimal environmental conditions differ. Several studies have paradoxically found a decrease in predictive performance at finer resolutions (reviewed in Meynard et al., 2023). The SDMs assume that the species is in equilibrium with its environment (Guisan and Thuiller 2005) and that the environmental data used are sufficient to represent the species' niche. These assumptions can be questionable at coarse resolution data (e.g. the increase in distribution of invasive species; Gallien et al., 2012) and even more problematic with high-resolution models. At a finer scale the random processes may become more apparent, such as the movement of mobile species, but also the climatic conditions of the predictors may vary more at finer resolutions. Additionally, with a more detailed dataset at a fine scale, it may become more difficult to capture all the variance within the landscape with a limited set of variables.
Besides the increase in information high-resolution data offers, the data should make sense with the movement and dispersal of the modelled species. Additionally, the rapid changes in

the environment such as current climate and land use change, challenges the assumption of a population being in equilibrium with the environment.
In terms of biotic factors, several studies (e.g. Atauchi et al., 2018; Herrera et al., 2018; Mathieu-Bégné et al., 2021; Mpakairi et al., 2017), including Chapter 3 of this thesis, have integrated biotic factors within the models. However, the addition of a biotic factor does not always yield significant model improvement consistent for all species (Giannini et al. 2013). Biotic interactions are considered to occur at a fine scale, even though many examples exist of their effect on a large-scale (reviewed in Wisz et al., 2013). In chapter 3, I found the relevance of scale for the integration of the biotic factor within the models. Additionally, at a courser scale biotic factors would, arguable, classify more as scenopoetic variables, or variables that influence the fitness of the modelled species but are not affected by its presence , instead of bionomic variables, that are influenced by the presence of the modelled species (see section 1.2). The population numbers of a prey species, for example, would be affected less at a large scale than at a fine scale. This consequence would require a revision of whether the high-resolution data represent scenopoetic or bionomic variables. The influence of the modelled species on the local population of the biotic factor would lead to a violation of the equilibrium assumption and the assumption that the niche conditions are independent of the species occurrence.

The movement/dispersal category of the BAM diagram still remains a challenge, but has been integrated by several studies (Engler et al. 2012, Shipley et al. 2022). As SDMs are developed at a finer scale, a new theoretical framework is needed for the justification of their use at these scales. Movement at fine resolution and population dynamics of interacting species can be erratic and it could be insightful to explore whether other methods that include these random processes could perform better at finer resolutions. Movement patterns can differ substantially within the group of pollinator species. Bees are central-place foragers limited to a small range around their nest where they care for their offspring (Cresswell et al. 2000). Hoverflies, on the other hand, move freely across the landscape unattached to a specific location. In fact, some species, like *Episyrphus balteatus*, show annual large distance migration (Van Der Goot 1979, Hondelmann and Poehling 2007).

6.2 Modelling perspective

With improvements in data availability and the development of new technologies, modelling methods are likely to change significantly in the coming years. In the past decades, there has been a large increase in the number of SDM studies (Lobo et al. 2010, Melo-Merino et al. 2020) and also in the development of new methods (Muscarella et al. 2014, Zurell et al. 2020). A standardized scheme has been developed for reporting SDMs (Zurell et al. 2020) and this can help researchers to improve reproducibility and transparency of their modelling methods. Even though reproducibility and transparency are not the main evaluation criteria of this thesis, they are important aspects, that allow modellers to apply methods of others studies and improve clarity, especially with future developments of novel modelling methods. The methods are improving, but the data quality is not always keeping up. There has been an increase in the use of citizen science data for ecological studies, that present many different problems related to bias (Johnston et al. 2020). Deep learning is also being more frequently applied to the problem of predicting species distributions (Rademaker et al. 2019),

and these complex machine learning methods can improve predictions, but their models are often more challenging to interpret (Ryo et al., 2021).
For trees, it is now possible with high-resolution data to identify individual trees species from images (Deur et al. 2020). Additionally, larger organisms can be identified in a similar way from satellite images (Khan et al. 2023). However, the identification of insects and herbaceous plants is likely not feasible from satellite images in the near future. For those species, species distribution modelling represents an important tool for allowing generalisations from sparse occurrences. Moreover, many countries do not have access to fine scale data or sufficient coverage of occurrence data.
The findings in my thesis show important implications for SDMs in general and in this section I provide an overview with those considerations that might be of use in future modelling work (see table 1 below).

Table 1: suggestions for other modellers from the findings of this thesis.

Category	Suggestions
Bias of the input data	• Observation data may be skewed towards nature or urban land use areas, compared to agriculture. Stratified sampling based on the contribution of different land use classes within the study area could improve models. This has been proposed by other studies as well (Husby et al. 2021, Scherber et al. 2021). • In the case of agricultural areas, I suggest exploring whether data on LEs is available for the study area. In the case of high-resolution data, the area of LEs could be represented as percentage cover or distance to. In the case of absence of high-resolution data, an aggregated measure of LEs or a measure of connectivity could be useful. The integration of connectivity and dispersal limitations in SDMs has been proposed in another SDM study (Vasudev et al. 2015). If no data is available on LEs, a measure of field edges could be explored as an approximation of the presence of LEs, as they often occur on the edges of parcels.
Biotic factor	• In the case of specialist species I suggest to include the population (or occurrence) of their biotic factor as an input variable and for the generalist species an aggregate thereof (e.g. richness of flowering plant species that are part of their diet for a generalist bee). If the biotic factor has a limited distribution it is especially important to include it in the models. • If insufficient data is available on the biotic factor, it might be an option to use genus-level data as a substitute for species level data. However, this depends on the ecology of the modelled species. For example, bee species often specialize on multiple species within a genus. Another study has used the modelled distribution of the biotic factor instead of the raw observation data (Giannini et al. 2013) to model biotic interactions of bees and this could represent another option in case of insufficient data.
Evaluation	• I suggest to use an independently collected dataset for the evaluation of the predictions made by the models, as has been done by other studies as well (Johnson and Gillingham 2005, Marshall et al. 2015, Cole et al. 2017, Lee-Yaw et al. 2022). Ideally, the dataset is collected differently and is not subject to the main biases of the original dataset. • I suggest the use of null models or randomized models, both for the performance of the models and the relevance of added variables. The null models for the evaluation of model performance have been applied to SDMs in another study (Raes and ter Steege 2007), as well as the

comparison of biotic factors against randomized interactions (Giannini et al. 2013).

6.3 Societal implications

The SDMs presented in this thesis offer significant potential for aiding pollinator conservation, mainly for nature organisations, farmer collectives and provinces/ municipalities. However, their implementation requires overcoming several obstacles and considerations. Clearly, the outcomes of SDMs cannot be used directly as inputs in conservation actions. However, the maps may present useful tools for communication between stakeholders, due to their flexibility in interpretation. The maps provide relevant information, particularly when they showcase areas of interest to stakeholders. This allows each stakeholder to focus on a specific location of interest within the broader context of the surrounding environment. The maps' flexibility in interpretation, while remaining recognizable for diverse stakeholders— is an essential criterion for facilitating communication among stakeholders (Shaw et al., 2022). An important aspect of improving implementation of modelling studies is a participatory approach (Wassen et al. 2011). The study by Wassen and coworkers found that the utilization of models was correlated with their acceptance and the prerequisite of the acceptance of the models was the applicability in decision making (Wassen et al. 2011). In other words the implementation of the models could be improved by involving stakeholders from the beginning and refining the maps to the stakeholders' needs. With hindsight, a participatory approach to promoting the implementation of the model results in this thesis could have helped raise stakeholder awareness of the value of the maps. Furthermore, the maps could have functioned as tools for initiating dialogue among various stakeholders at the landscape scale across the region. However, in this thesis, stakeholders were not included from the beginning. Yet, the resulting maps from this thesis in the agricultural regions can still be evaluated with different stakeholders, such as nature organisations, farmer collectives and municipalities. For example, the maps could be presented in a workshop around current challenges for biodiversity in an agricultural region and with the feedback they could be changed to cater to the specific needs of the stakeholders. Another obstacle for the implementation of the SDM maps is the communication of uncertainty, because through the different modelling steps, from the predictors through to the predicted habitat suitability maps, a lot of variability is introduced. It is important to communicate the exact measure that is uncertain, the underlying causes of this uncertainty and a means to express the level of uncertainty (Van Der Bles et al. 2019).
In general, the pollinator suitability maps could guide strategic landscape planning to promote biodiversity. The models identify relationships between species and certain habitats. These relationships can indirectly evaluate the impact that policies can make.
In the case of nature organisations, the pollinator maps can identify key areas for connecting nature areas and source populations of threatened pollinators. Nature conservation organizations have target species of key interest, which could be modelled to demonstrate their relevance and usefulness to these stakeholders.

The maps could also be of value for municipalities and/or the province for planning and implementing urban green spaces, which currently mainly benefit non-threatened bee species (as was found in Chapter 2). The implementation strategies could encourage the presence of threatened bees in urban green areas by creating awareness of bee friendly urban green. For example, large areas like communal gardens have been found to harbor threatened species (Lanner et al. 2020). I found that threatened bees are more often specialists, exhibiting narrower preferences for both habitat type and pollen sources. In order to attract threatened bees to urban green spaces, these spaces should present similarities to specific habitat types with local plant species, important for specialist bees.

Farmer collectives could play a critical role in the effort for pollinator conservation as 54% of the Netherlands consists of agriculture (Central Bureau of Statistics (CBS) 2015). Protecting pollinators demands landscape-level planning and change, making the role of such collectives pivotal. The low implementation of agri-environmental measures by farmers can depend on several factors. The study by Runhaar and coworkers presented a framework in which farmers should be motivated, enabled or able, legitimized, and demanded to include nature conservation in their practices (Runhaar et al. 2017). The ability and motivation are characteristics of the farmers that can be influenced by factors such as governance arrangements, but also intrinsic traits. The legitimization and demand relate to other actors such as the government and are influenced by factors such as legislation and environmental regulations (Runhaar et al. 2017). SDMs have the potential to motivate, as they can show future predictions of biodiversity in scenarios where pollinator-friendly measures, such as LEs, are taken. Additionally, they can evaluate the results of current measures and the effect of the measures on the distribution of a species of interest. Besides motivation, SDMs can also enable farmers within the collectives to implement biodiversity measures by providing information on effective measures and suitable locations of these measures. For example, the spatial characteristic of the maps allows for the identification of key areas within agricultural regions that are located next to nature areas.

From a national and European level, as outlined in Section 1.5, the European Union has implemented measures for biodiversity in agriculture, such as the requirement for at least 3% of agricultural land to feature non-productive elements. The findings of this thesis support the positive effect these measures may have on crop pollinators. The majority of crop pollinators come from relatively small group of common pollinators (Kleijn et al. 2015). The challenge lies in including the conservation of threatened pollinators and other biodiversity at the same time as securing pollination services for crops. In the report for the Ministry of Agriculture, Fisheries, Food Security and Nature I provided guidelines for the management and placements of LEs and these guidelines may be insightful for different stakeholders. In this report, I suggested measures within different ecological focus areas, which are non-productive agricultural areas, for individual pollinator groups. The measures are based on the ecology of the pollinator group and are generally focussed on providing a wide variety of reproductive habitats and food resources, while taking the surrounding landscape into account. The summary and link to the report can be found in the 'summary of report for Ministry' section in the Supplementary materials.

To conclude, SDMs are important tools for estimating how pollinators live and thrive in the Dutch landscape. However, to increase model performance, ecological relevance and the applicability of the models, several challenges have to be overcome. This thesis addresses the challenge of integrating key biotic interactions of pollinators within the models and the introduction of input variables that represent the 'complexity' of the landscape and the ecological requirements of the bees and hoverflies living in it. The knowledge obtained in this thesis can both further modelling advancements and has the potential to contribute to the protection of pollinators in the Netherlands.