Species Distribution Modelling of European and Danish Bees

DEPARTMENT OF CHEMISTRY AND BIOSCIENCE JUNE 2021
KRISTIAN DAMHOLT MARKUSEN

AALBORG UNIVERSITY

Title:
Species Distribution Modelling of European and Danish Bees

Semester
9th and 10th semester

Project period
01.09.2020 – 04.06.2021

ECTS
60 ECTS

Supervisors
Dan Bruhn

Michael Ørsted

Participants
Kristian Damholt Markusen

Copies: 1

Pages: 45

Appendices: 0

Abstract

Bees are important in the ecosystem, in their roles as pollinators of various flora, some of which are agriculturally relevant. This makes pollinators not only valuable in terms of biodiversity and ecosystem stability, but also financially since growing more, healthier crops without the need for further land use is positive for the environment and biodiversity in general. Due to climate change, habitat fragmentation and general habitat degradation, bee populations are under pressure and declining, making them a key target for conservation efforts. To effectively plan conservation efforts, it is necessary to not only know where the intended target is present now, but also where it has the potential to be present in the future. Tools such as Species Distribution Models (SDMs) make it possible to predict the potential future distribution of species, ensuring that areas targeted for conservation efforts not only help the species currently but will benefit the species going forward. This project looks at a 50-year prediction of climate, following two different Shared Socio-economic Pathways (SSPs), specifically SSP 1-2.6 and SSP 3-7.0 (Hausfather 2019), to point to potentially beneficial areas for conservation of bees, now and in 50 years. Furthermore, this project implements species interactions into the SDMs, strengthening the accuracy of its models. The project found that Northern Europe, specifically England, the Netherlands, and Denmark, had a higher-than-average bee species richness. It also showed that the overall species richness was positively skewed. The countries with higher bee species richness should consider conservation efforts towards bees, since they can conserve the highest number of species, and these species will soon no longer be able to move their distribution northwards, due to being blocked by the ocean. Countries with a lower density of species, would still benefit from conserving bees, but efforts here might sooner become irrelevant due to changes in climate. Denmark is among the countries that have a high bee species richness, meaning that Denmark has an opportunity to conserve many species within its borders. The main species richness in Denmark is focused on Northern Jutland, with Zealand having lower than average bee species richness for areas within Denmark. When considering the distribution of the host plants of certain Danish bee species, the models show that the distribution of Danish bee species is restricted by the distribution of their host plants. The fact that bee species distribution is limited by the distribution of their host plants, means that any effort to conserve bees must take their host plants into consideration to be effective.

Indhold

Introduction

Bees

Bees belong to the order Hymenoptera, which also includes wasps, sawflies and ants; females of this order develop stingers (Michener 2007). Bees are well known for living in colonies, despite many species living solitary lifestyles. Bees primarily feed on plant pollen, which they gather from various plants. In doing so they act as pollinators, helping to facilitate plant reproduction (Ollerton, Winfree, and Tarrant 2011; Scheper et al. 2014). Some plants would not be able to set seeds without the help of pollinating animals, such as bees. In fact 78% of plant species in temperate climates are dependent on animal pollination in order to reproduce successfully (Ollerton, Winfree, and Tarrant 2011). This makes this ecological role key to the stability of an ecosystem.

Agricultural importance

When considering pollinators' roles in agriculture, bees are considered the most important for agricultural crops (Schenk, Krauss, and Holzschuh 2018). While there is not a current shortage of pollination services, agriculture is becoming more pollinator dependent (Aizen et al. 2008). However, with a declining population of European bees and a growing demand, it is only a matter of time before a shortage in pollination services will be a reality, unless efforts are made to prevent the decline in pollinator populations. Research has also shown that diversity in pollinators can help increase agricultural yield, without increases in land use or intensity (Brittain et al. 2013). By utilizing the bees' natural foraging behaviours when in the presence of another species, the likelihood of successful pollination increases. This further emphasises the need for conservation of the wild bee species, as these might be the key to keeping up with increasing demands in a sustainable way. The presence of managed bees can have a negative effect on wild bee populations by introducing pathogens to the area, changing the composition of flora, and competing for flowers and nesting resources (Mallinger, Gaines-Day, and Gratton 2017).

Interactions between bees and plants

Every species of bee has a set of potential feeding plants, they have species specific adaptations to better gather pollen from these plants. Simultaneously these plants develop to better take advantage of the pollinators, ensuring that their genetic material gets deposited and transferred via the bees that visit them. This symbiosis is important for the ecosystems, and thus it is important to know about these interactions. A Danish study has compiled a list of documented interactions between bees in Denmark and the flowers that host them (Rasmussen, Schmidt, and Madsen 2016), and a study from the Netherlands has looked into how the distribution of host-plants affect the distribution of the related bee species (Scheper et al. 2014). When looking into the conservation of bees, not paying attention to the composition of flora in the area could result in failure before the effort even begins. Not all bee species have access to the same number of host plants, some are highly specialised to gather pollen from specific plants (Miller-Struttmann et al. 2015; McAulay, Killingsworth, and Forrest 2021).

Threats and population decline

The decline of bee populations is a well-known problem (Scheper et al. 2014; Klein et al. 2017; Polce et al. 2013; Marshall et al. 2018). The use of pesticides can be a major threat to bee populations, since they often affect neural pathways important for the complex cognitive challenge of foraging pollen (Klein et al. 2017). Another factor in the decline in bee populations is climate change, and how sudden changes in climate affects bees. A common response to increasing temperatures, is migrating towards the poles, this response is lacking within bumble bees in Europe and North America (Kerr et al. 2015). This lack of adaptation to changing climate means that species' distribution area is shrinking, since the southern boundary is changed due to areas no longer supporting the species existence there. Another effect of climate change on bee species is desynchronization with interaction partners, which can lead to periods of starvation. A study found that a desynchronization of six days killed most specimens of the studied species (Schenk, Krauss, and Holzschuh 2018), their experiment was based on a study that found the average desynchronization by 2050 to be three days between bees and plants (Thackeray et al. 2016). Research also points to a mismatch in functionality, as bees specialised toward certain plants, develop more generalist traits to better cope with the pressures of a changing climate (Miller-Struttmann et al. 2015).

From the 185 species with interaction data modelled in this project, 23 are considered endangered or worse on the Danish red list (Madsen 2019).

Biotopes importance to biodiversity

Biotopes are partially defined by the flora within them, as such they also play a role in determining which pollinators are in the area. When looking at areas to focus conservation targeting biotopes could be a way to not only conserve bees but their host plants as well. Within Denmark the law of nature conservation (Miljø- og Fødevareministeriet 2009) protects the ecosystem, and under that, certain biotopes. The biotopes protected under the law of nature conservation are commonly referred to as §3 areas, as this is the part of the law that specifies these biotopes. This paragraph states that: There are not to be changes made to the condition of natural lakes of more than $100m^2$, water streams, heaths, bogs and similar, beach meadows, beach swamps, fresh meadows, or biological pastures, when these areas either alone or in conjunction reach an area of $2500m^2$.

Species Distribution Model

Species Distribution Models (SDMs) are a useful tool for conservation work, as it can help point to areas where conservation will be more likely to be effective in the future. With most species adapting their distribution as a responds to climate change, SDMs are vital to knowing where species are most likely to be in the future (Swab et al. 2015). Other than predicting the distribution of wanted species, SDMs are also used for predicting the movements of potentially invasive species (Jiménez-valverde et al. 2011). In addition to the climate focused SDM, by adding biological interactions to the models, making a Joint Species Distribution Model (JSDM), you add another layer of information to the models (Pollock et al. 2014). Such models are very useful when mapping the distribution of species with close ties to other species.

There are two different approaches to making an SDM, the correlative and the mechanistic. Both methods are interested in the niche of the species, they simply look for different stages of the niche. The mechanistic approach is concerned with the physiological limitations of a species, using these to find the niche that a species can occupy within the environment (Kearney and Porter 2009). This approach requires a tremendous knowledge about the modelled species, and thus is not viable for species that have not been thoroughly studied. The other approach, the correlative approach, seeks to decern the species niche by looking at where the species occurs. This leads to the realised niche of the species, a subset of the conditions that the species could be found in (Pulliam 2000). The correlative method only requires knowledge of where a species is present, and the environmental factors that affect that area, most often climate variables. This makes this approach more suited for modelling species that have been studied extensively, such as rare or invasive species (Morin and Thuiller 2009).

Model improvements

There are many ways to improve models after the initial run, the simplest of which is to run the model again to see whether the results match up. This is commonly done to converge on the most accurate model, the replicate runs ensure that the variability of the importance of different variables is considered (Manzoor, Griffiths, and Lukac 2018). Another way to improve the model is by simplifying the variables that are put into the model, the fewer variables used to get the same predictive power, lessens the risk of overfitting the models (Jiménez-valverde et al. 2011).

Aim

The aim of this project is to assess: 1) Where in Europe are the highest species richness for bees, 2) How does the species richness shift over time following SSP 1-2.6 and SSP 3-7.0, 3) Where in

Denmark has the highest bee species richness, 4) how do interactions between plants and bees affect the bee species distribution, 5) are endangered species present in Danish species richness hot spots.

Method

This project aims to develop models for the potential distribution of European bee species, along with that of Danish bee species and their host plants, for the current climate and that of 2061-2080 according to GCM CNRM-CM6-1's SSP 1-2.6 and SSP 3-7.0 scenarios. To do so three sets of climate data was acquired, along with occurrence data for the bees and plants relevant to this project. Models were made using the maxent software through R.

Getting the data

Observational data

Occurrence data was acquired through the Global Biodiversity Information Facility (GBIF), which is a database that collects the data of many smaller databases making it easier to gather all the data required for research such as this.

Using the families of bees; Andrenidae (GBIF 2020a), Apidae (GBIF 2020b), Colletidae (GBIF 2020c), Halictidae (GBIF 2020d), Megachilidae (GBIF 2020e), and Melittidae (GBIF 2020f), a list was compiled of all species within these families, for which the GBIF database had occurrence data, meaning verified sightings of the species. The family Stenotritidae was left out of this project, as it is restricted to Australia. This resulted in a list of 693 species of European bees, 605 species had observations that fulfilled these criteria:

- contained decimal coordinates for the observation.
    - which did not match the coordinates of a capital city, the centre of a country or an institution.
    - were not equal or zero.
    - fell outside the country the observation was assigned to.
- Had an uncertainty of 100 km or less.
- Occurred after 1900.

These criteria were set to avoid occurrence data where the coordinates were entered incorrectly, the uncertainty was too high to get a reliable distribution, or the occurrence was too old to match the climate data used in the models, thus giving the model incorrect data.

Interaction between bee and plant species

To see whether the potential Danish bee distribution was restricted by the potential distribution of their host plants, I decided to use potential distribution of known host plants as a filter for potential bee distribution in Denmark. Based on the list of interactions, defined as a bee landing on a plant, between bees and plants in Denmark found in Rasmussen, Schmidt, and Madsen (2016), a list of plant species important to the survival of Danish bees were compiled. Only interactions that were identified to species level were considered, as a matter of caution as well as to better utilise the resources available to the project. This led to a list of 163 plants that had documented interactions with one or more species of bee, out of these plant species, 159 had data that fulfilled the requirements described for the bee occurrence data. Furthermore, the interactions are only accounting for 183 of the bee species, meaning that only those species can be included in the interaction models.

Climate and soil data

The 19 climate variables used in the project were obtained from the WorldClim data website (Fick and Hijmans 2017)(Table 1), they are based on temperature and precipitation.

Table 1 descriptions of the 19 climate variables and how they are derived from temperature (T) and monthly precipitation (PPT) data, the numeric month (i) is used throughout the calculations. When a calculation aims to select a quarterly value, the dataset wraps around into the next year, to calculate all possible quarters in a full year (O'Donnell and Ignizio 2012). The variables concerning P, N and pH in topsoil, were used in the modelling of host plants, and based on LUCAS soil survey (Ballabio et al. 2019).

Bioclim annotation	Description	Unit	Calculation	Notes on interpretation
Bio1	Annual Mean Temperature	°C	$Bio\ 1 = \frac{\sum_{i=1}^{i=12} Tavg}{12}$	The monthly average is averaged over the year. This approximates the total energy input for an ecosystem.
Bio2	Annual Mean Diurnal Range	°C	$Bio\ 2 = \frac{\sum_{i=1}^{i=12}(Tmax_i - Tmin_i)}{12}$	The monthly diurnal range is averaged over the year. Help decern relevance of temperature fluctuations.
Bio3	Isothermality	%	$Bio\ 3 = \frac{Bio\ 2}{Bio\ 7} x100$	The ratio of the annual mean diurnal range to the annual temperature range. This helps quantify the day to night fluctuation compared with the summer to winter fluctuation.
Bio4	Temperature seasonality (SD)	°C	$Bio\ 4 = SD\{Tavg_1, \dots, Tavg_{12}\}$	The standard deviation of the 12 mean monthly temperatures. Helps decern the changes in temperature throughout a year.
Bio5	Max Temperature of Warmest Month	°C	$Bio\ 5 = \max(\{Tmax_1, \dots, Tmax_{12}\})$	Selecting the highest temperature of the year. Helps decern impact of warm temperature anomalies.

Bio6	Min Temperature of Coldest Month	°C	$Bio\ 6 = \min(\{Tmin_1, \ldots, Tmin_{12}\})$	Selecting the lowest temperature of the year. Helps decern impact of cold temperature anomalies.
Bio7	Annual Tempera-ture Range	°C	$Bio\ 7 = Bio\ 5 - Bio\ 6$	The range from the lowest yearly temperature to the highest. Helps decern the impact of ranges of extreme temperature ranges.
Bio8	Mean Temperature of Wettest Quarter	°C	$Q_{PPT_{max}} = \max\left(\left\lvert\begin{matrix}\sum_{i=1}^{i=3} PPT_i, \\ \sum_{i=2}^{i=4} PPT_i, \\ \ldots, \\ \sum_{i=10}^{i=12} PPT_i, \\ \sum_{i=11}^{i=1} PPT_i, \\ \sum_{i=12}^{i=2} PPT_i,\end{matrix}\right\rvert\right)$ $Bio\ 8 = \frac{\sum_{i=1}^{i=3} Tavg_i}{3}$	The three-month period with the highest total precipitation is se-lected, and the average tempera-ture of those three months are averaged. Helps decern how much environmental factors im-pact distribution.
Bio9	Mean Temperature of Driest Quarter	°C	$Q_{PPT_{min}} = \min\left(\left\lvert\begin{matrix}\sum_{i=1}^{i=3} PPT_i, \\ \sum_{i=2}^{i=4} PPT_i, \\ \ldots, \\ \sum_{i=10}^{i=12} PPT_i, \\ \sum_{i=11}^{i=1} PPT_i, \\ \sum_{i=12}^{i=2} PPT_i,\end{matrix}\right\rvert\right)$ $Bio\ 9 = \frac{\sum_{i=1}^{i=3} Tavg_i}{3}$	The three-month period with the lowest total precipitation is se-lected, and the average tempera-ture of those three months are averaged. Helps decern how much environmental factors im-pact distribution.
Bio10	Mean Temperature of Warmest Quarter	°C	$Q_{T_{max}} = \max\left(\left\lvert\begin{matrix}\sum_{i=1}^{i=3} Tavg_i, \\ \sum_{i=2}^{i=4} Tavg_i, \\ \ldots, \\ \sum_{i=10}^{i=12} Tavg_i, \\ \sum_{i=11}^{i=1} Tavg_i, \\ \sum_{i=12}^{i=2} Tavg_i,\end{matrix}\right\rvert\right)$ $Bio\ 10 = \frac{\sum_{i=1}^{i=3} Tavg_i}{3}$	The Three-month period with the highest sum of average tempera-tures is selected, the average temperatures of those three months are then averaged. Helps decern the impact of prolonged warm temperatures on distribu-tion.

Bio11	Mean Temperature of Coldest Quarter	°C	$Q_{T_{min}} = \min\left(\begin{vmatrix} \sum_{i=1}^{i=3} Tavg_i, \\ \sum_{i=2}^{i=4} Tavg_i, \\ \dots, \\ \sum_{i=10}^{i=12} Tavg_i, \\ \sum_{i=11}^{i=1} Tavg_i, \\ \sum_{i=12}^{i=2} Tavg_i, \end{vmatrix}\right)$ $Bio\ 10 = \frac{\sum_{i=1}^{i=3} Tavg_i}{3}$	The Three-month period with the lowest sum of average temperatures is selected, the average temperatures of those three months are then averaged. Helps decern the impact of prolonged cold temperatures on distribution.
Bio12	Annual precipitation	Mm	$Bio\ 12 = \sum_{i=1}^{i=12} PPT$	The sum of the monthly precipitation of a year. Helps decern the impact of water availability on distribution.
Bio13	Precipitation of Wettest Month	Mm	$Bio\ 13 = \max([PPT_1, \dots, PPT_{12}])$	Selecting the month with the highest total precipitation. Helps decern impact of extreme precipitation on distribution.
Bio14	Precipitation of Driest Month	Mm	$Bio\ 14 = \min([PPT_1, \dots, PPT_{12}])$	Selecting the month with the lowest total precipitation. Helps decern impact of extreme lack of precipitation on distribution.
Bio15	Precipitation Seasonality (CV)	%	$Bio\ 15 = \frac{SD\{PPT_1, \dots, PPT_{12}\}}{1 + \left(\frac{Bio\ 12}{12}\right)} x100$	First the standard deviation of the monthly precipitation is calculated, that is then divided with the mean monthly precipitation plus one, then that value is multiplied by 100. Helps decern the impact of variability of precipitation of the distribution.
Bio16	Precipitation of Wettest Quarter	Mm	$Bio\ 16 = \max\left(\begin{vmatrix} \sum_{i=1}^{i=3} PPT_i, \\ \sum_{i=2}^{i=4} PPT_i, \\ \dots, \\ \sum_{i=10}^{i=12} PPT_i, \\ \sum_{i=11}^{i=1} PPT_i, \\ \sum_{i=12}^{i=2} PPT_i, \end{vmatrix}\right)$	The three-month period with the highest precipitation is found. Helps decern the impact of such environmental factors on distribution.

Bio17	Precipitation of Driest Quarter	Mm	$$Bio\ 17 = \min\left(\left\| \begin{matrix} \sum_{i=1}^{i=3} PPT_i, \\ \sum_{i=2}^{i=4} PPT_i, \\ \dots, \\ \sum_{i=10}^{i=12} PPT_i, \\ \sum_{i=11}^{i=1} PPT_i, \\ \sum_{i=12}^{i=2} PPT_i, \end{matrix} \right\|\right)$$	The three-month period with the lowest precipitation is found. Helps decern the impact of such environmental factors on distribution.
Bio18	Precipitation of Warmest Quarter	Mm	$$Q_{T_{max}} = \max\left(\left\| \begin{matrix} \sum_{i=1}^{i=3} Tavg_i, \\ \sum_{i=2}^{i=4} Tavg_i, \\ \dots, \\ \sum_{i=10}^{i=12} Tavg_i, \\ \sum_{i=11}^{i=1} Tavg_i, \\ \sum_{i=12}^{i=2} Tavg_i, \end{matrix} \right\|\right)$$ $$Bio\ 18 = \sum_{i=1}^{i=3} PPT_i$$	The Three-month period with the highest sum of average temperatures is selected, the total precipitation of those three months is calculated. Helps decern the impact of such environmental factors on distribution.
Bio19	Precipitation of Coldest Quarter	Mm	$$Q_{T_{min}} = \min\left(\left\| \begin{matrix} \sum_{i=1}^{i=3} Tavg_i, \\ \sum_{i=2}^{i=4} Tavg_i, \\ \dots, \\ \sum_{i=10}^{i=12} Tavg_i, \\ \sum_{i=11}^{i=1} Tavg_i, \\ \sum_{i=12}^{i=2} Tavg_i, \end{matrix} \right\|\right)$$ $$Bio\ 19 = \sum_{i=1}^{i=3} PPT_i$$	The Three-month period with the lowest sum of average temperatures is selected, the total precipitation of those three months is calculated. Helps decern the impact of such environmental factors on distribution
P	Phosphorus in topsoil	mg/kg	From the LUCAS topsoil survey in Europe	
N	Nitrogen in topsoil	g/kg	From the LUCAS topsoil survey in Europe	
pH	pH in H_2O		From the LUCAS topsoil survey in Europe	

The resolution of the climate variable raster was 2,5 minutes, meaning the cells measured 21,44 km^2 at the equator, resulting in models of the same resolution. The climate raster were cropped to Europe and using the World Geodetic System 1984 (WGS84) to project longitude and latitude onto maps. In order to predict the potential distribution of the plants more accurately, soil variables were added to the list of variables, specifically soil pH and the content of P and N in the soil (Ballabio et al. 2019). These values were considered static in the 50-year period that the project evaluates. Due to the soil variables only being available for EU member countries, some data points from the database were no longer valid, any plant with more than 50% invalid points were not able to be modelled. This resulted in 145 plant models. For the future scenarios, the Global Circulation Model (GCM) CNRM-CM6-1 (Voldoire et al. 2019) was chosen based on its basis in Europe, furthermore the scenarios SSP1-2.6 (Hausfather 2019) was chosen to show the changes in distribution should the global temperature increase be kept to a minimum, along with SSP3-7.0 to show what would happen in the case we do not attempt to halt the increase in temperature. Models in this project concerns the period 2061-2080, and the GCMs all predict the climate up to 2100 including the climate in the period of interest (Table 2).

Table 2 The predicted warming and CO2 emissions by the year 2100 following the two shared socio-economic pathways; SSP 1-2.6 and SSP 3-7.0.

Shared Socio-economic Pathway	*Warming by 2100 (°C)*	*CO2 Emission by 2100 (Gigatonnes)*
SSP 1-2.6	1,3 – 2,9	-8,62
SSP 3-7.0	3,0 – 6,2	82,73

Characteristic biotope plants

To answer the question of which biotope contains the most important host flowers, for the distribution of Danish bees, lists of flowers from different biotopes are required. In this project the lists were based on the field charts meant to determine the quality of the protected §3 areas (Fredshavn et al. 2009). In order to determine which list contained the flowers that contributed most to the distribution of Danish bees, each list was cross referenced with the list of host plants used previously. By doing that a list of host plants for each biotope was compiled. In turn every list of biotope host plants was removed from the total list of host plants, in order to observe the changes to the distribution as a result of removing plant species from each biotope as potential interaction partners. This was done for each biotope and each climate scenario to see whether, according to the models presented in this project, the importance of each biotope remains constant or changes as a result of climate change.

Building the Models

The models in this project are made using the R package 'dismo' (Hijmans et al. 2020) and the modelling software 'MaxEnt' (Phillips, Dudík, and Schapire 2021). The models were made using the default setting of MaxEnt, with 25% percent of the occurrence data held back to train the model, by using them as unknown presences that the model must be able to predict. This meant that species with fewer than four occurrences would not be able to be modelled. In this project only one GCM was used, and no replicate runs were made of any models.

MaxEnt

MaxEnt is among the most popular tools for species distribution models, with more than 1000 published applications in the last 15 years (Merow, Smith, and Silander 2013; Hijmans et al. 2020; Phillips, Dudík, and Schapire 2021). Maxent takes an input in the form of presence-only data, along with a set of environmental data. These often come in the shape of raster, which are grid-based maps. It is important that the different environmental data raster has the same resolution. The raw output of MaxEnt when the occurrence points are changed into a raster, is a probability of presence for each cell on the grid map. The model works by taking a group of background cells, where the occurrence of a species is unknown, and training the model to differentiate between background cells and occurrence cells. This is done through a process known as machine learning, where an algorithm fine tunes the importance of different predicting factors until it can tell occurrence from background. MaxEnt will by default pick background points completely at random, which assumes that the species is equally likely to be anywhere within the area being modelled, this results in the most spatially diffuse distribution possible, leading to the largest possible range. Furthermore, MaxEnt builds a set of features from the selected predictors, by applying different feature classes such as linear, quadratic, product, threshold, hinge and categorical. By default, the types of features MaxEnt utilizes is dependent on number of presences with 80+ presences enabling all feature classes to be used. Regularization is a feature that ensures that the MaxEnt does not over-fit its models, by ensuring that empirical constraints are not too tightly fit, and by penalizing the model based on the magnitude of the coefficient. This makes regularization key in reducing the number of features that end up mattering in the final model. By default, the regularization coefficient is set for each feature class. MaxEnt assumes that every cell has the same likelihood of being sampled for occurrence, in other words that there is no sampling bias in the data, this is however rarely the case as sampling occurs more often closer to populated areas.

Output and evaluation of models

These models were then applied on the raster stack of variables for the current climate model, as well as those for SSP 1-2.6 and SSP 3-7.0. This resulted in three raster with continuous values between zero and one. To construct species richness maps these raster were converted from continuous values to binary, so that stacking the raster will give a number of species with potential presence in each cell. This is done by setting a threshold with values meeting or passing that threshold being considered potential presences, and values below that threshold being considered absences. These values are then put into a confusion matrix, that takes the known occurrence points and the predicted occurrence points and compare them:

	Predicted NO	Predicted YES
Actual NO	TNR	FPR
Actual YES	FNR	TPR

By shifting the threshold, the rates at which actual occurrence points are predicted as occurrences on the binary map changes, along with all the other rates. A cell that contains a known occurrence will always land in the 'Actual YES' row, but if the threshold is too high, it might fall in the 'Predicted NO' column, in that case lowering the threshold would put it into the 'Predicted YES' column increasing the True Positive Rate (TPR). At the same time lowering the threshold could turn cells that do not contain known occurrences from the 'Predicted NO' column to the 'Predicted YES' column, which would lower the True Negative Rate (TNR). To get the most accurate threshold you want a high true negative and true positive rate, which is why this project used True Skill Statistic (TSS) to dynamically set the most accurate threshold value for each model.

For each threshold between zero and one a TSS value was calculated using formular:

$$TSS = TPR + TNR - 1$$

Increases in either TPR or TNR will result in an increase in TSS and in a more accurate binary raster, by attempting different thresholds for the model we can find the threshold that allows for the highest TSS possible for the model, thus the most accurate binary raster.

Finding the right threshold is thus a matter of trying a series of different threshold, filling the confusion matrix to calculate the TSS, then applying the threshold that resulted in the highest TSS to the raster. The Receiver Operating Characteristic (ROC) curve is used to evaluate models independent of the threshold, by plotting the TPR as a function of the FPR. If the model is no better than random at predicting potential occurrences the Area Under the Curve (AUC) will be 0,5 or lower, but if the model is better than random the AUC will be above 0,5.

A model can only ever return a measure of potential distribution, since there is no guarantee that a species is present in a cell that supports it, the only way to truly know whether a species is precent in a cell is to physically sight it.

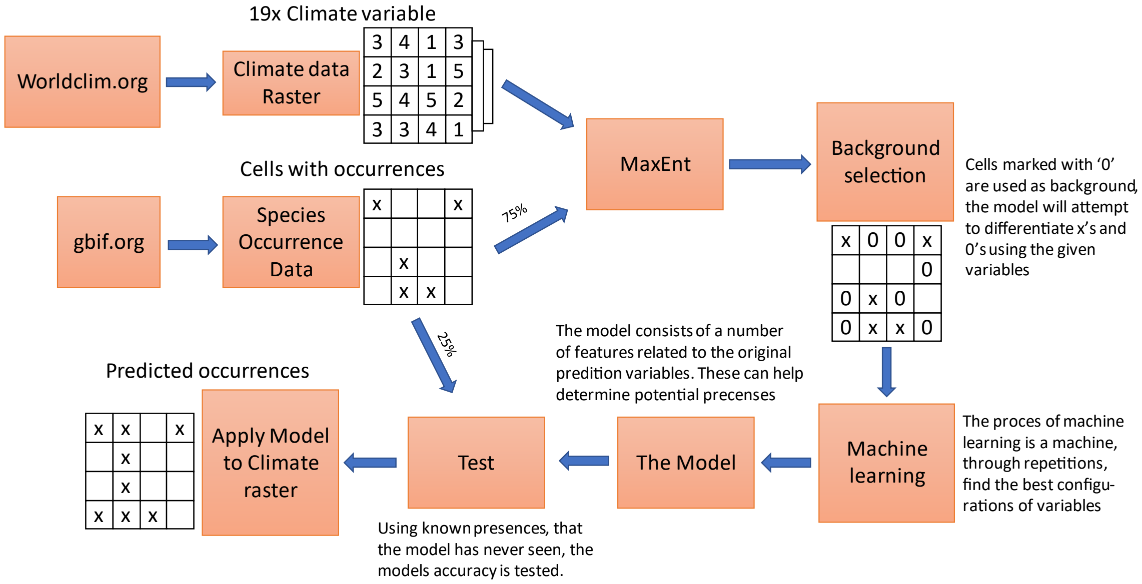

Figure 1 flowchart of the making of the predicted occurrence maps using MaxEnt. The inputs are 19 climate variable raster and the cells with known occurrences of one species.

Skewness

Skewness is a statistical measure of the evenness of a distribution, as well as which side of the mode has the more extreme values. The skewness value is either positive, negative or zero, a positive skewness means that there are more values to the right of the mode of the distribution, than what would be expected from a normal distribution. Skewness is hard to evaluate in small sample set but seeing as this project deals in very high numbers of values, it makes skewness easy to calculate. Any value +/- 0,179 is considered skewed with a sample size of 500, meaning that any map with a skew-ness beyond these values can be considered skewed (Doane and Seward 2011).

Results

689 models were constructed, 543 bee models and 146 plant models, to answer the questions posed in the aim of this project, the following results are all derived from said models.

Models of European bee species

Overview of three climate scenarios

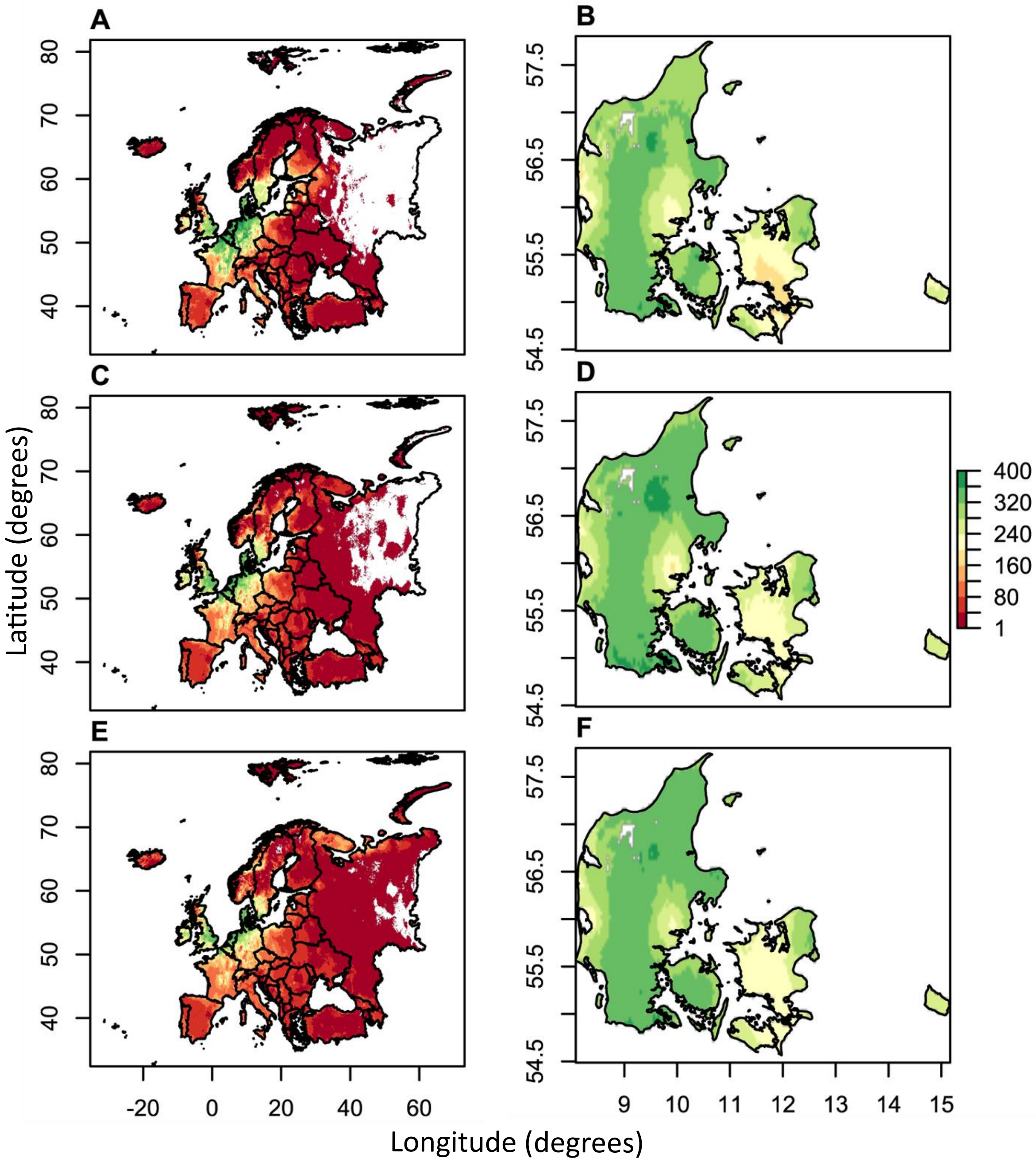

Figure 2 Potential bee species richness in Europe (A,C,E) and Denmark (B,D,F) across three different scenarios; current climate (A,B), SSP 1-2.6 (C,D) and SSP 3-7.0(E,F). Values under 1 have been coloured white, with values above 0 coloured from red to green in increments of 40.

The maps in Figure 2 are the result of stacking the binary maps of each species and summing the stack into one value. This represents the potential bee species richness of Europe and Denmark, each cell has a value between 0 and 400, with these values being designated a colour in increments of 40. The green colours are higher on the spectrum and red is lower.

The species richness map concerning the current climate period makes it clear that the highest potential species richness of bees is found in Southern England, Northern France, Belgium, The Netherlands, Germany, and Denmark (Figure 2A).

When focusing on Denmark it becomes clear that Zealand has the lowest potential bee species richness in Denmark, with parts of Northern Jutland scoring upwards of 400 potential bee species (Figure 2B).

In the SSP 1 model, the species richness is likewise higher in Southern England, Belgium, The Netherlands, Northern Germany, and Denmark. Furthermore a few species have started to move into Russia (Figure 2C). In this scenario we see little change in the distribution of potential bee species richness in Denmark (Figure 2D).

In the SSP3 scenario the species richness is highest in Southern England, the Netherlands, and Denmark. We also see that more of Russia becomes home to bees, and a patch in Sweden becomes uninhabitable to any of the modelled species of bee (Figure 2E). When looking at the Denmark map for this scenario there is no change in the distribution compared to the current climate (Figure 2F).

Table 3 Displaying the mean, median, skewness and total of each of the different climate scenarios for both European and Danish bee models.

European Bee Models	Mean Species Richness	Median Species Richness	Species Richness Skewness	Total Species Richness
Current climate	55,73	15	1,97	49406108
SSP 1-2.6 climate	55,60	27	2,03	49287168
SSP 3-7.0 climate	60,91	38	1,94	54000480

European Bee Models in Denmark	Mean Species Richness	Median Species Richness	Species Richness Skewness	Total Species Richness
Current climate	293,53	304	-0,68	1076798
SSP 1-2.6 climate	306,67	322	-0,71	1121188
SSP 3-7.0 climate	308,55	326	-0,86	1128068

Changes from current climate to the two future scenarios

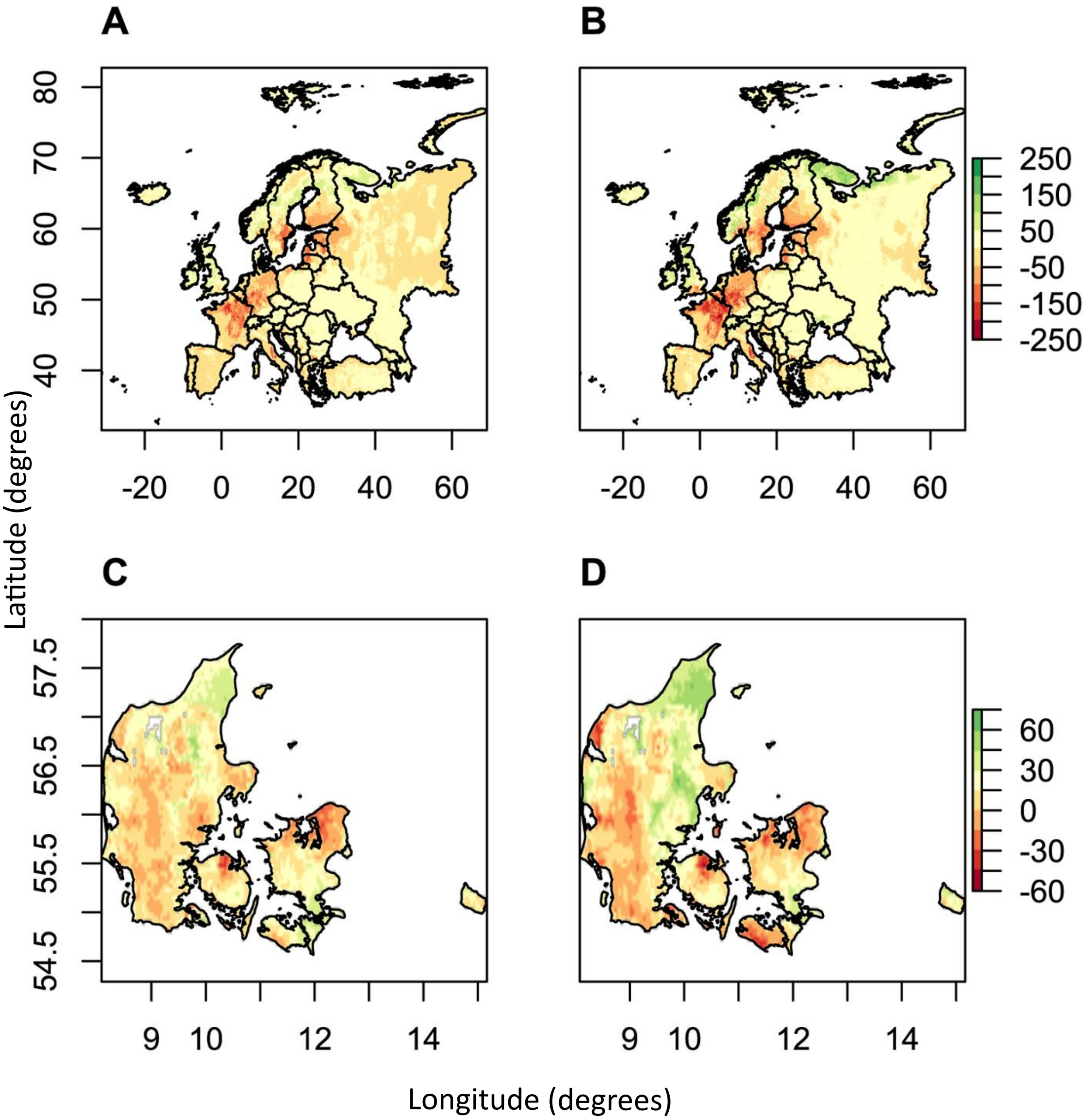

Figure 3 Changes in potential bee species richness in Europe (A,B) and Denmark (C,D), going from current climate to two different climate scenarios; SSP 1-2.6 (A,C) and SSP 3-7.0 (B,D). Negative changes are coloured red, positive changes are coloured green, the colours depict increments of 25.

To better illustrate where the changes from the current climate to the predicted climate in scenario SSP1 and SSP3, maps of the changes in value of each cell were made (Figure 3). The general image in Figure 3A is that losses are more prominent than gains, which is confirmed by the total species richness change (Table 4). Most losses occur in France and Germany.

Figure 3B shows the changes in species richness from the Current model to the SSP 3-7.0 model. Here we again see losses in France and Germany, as well as in Belgium, Sweden, Lithuania, Latvia, Estonia, and Finland. We also see gains in Russia, Norway, Great Britain, and Denmark (Figure 3B). When taking the sum of this raster there is a net gain of species richness, most likely due to the number of cells with gains in Russia.

The changes in species richness in Denmark (Figure 3C-D) are barely skewed, which can be seen in the balance between gains and losses, but regardless of climate scenario the total species richness change is positive (Table 4).

Table 4 Displaying the mean, median, skewness and total change from current climate to both SSP scenarios for European bee models in Europe and Denmark.

European Bee Models in Europe	Mean Species Richness Change	Median Species Richness Change	Species Richness Skewness Change	Total Species Richness Change
Current to SSP 1-2.6 climate	-0,13	1	-1,46	-118940
Current to SSP 3-7.0 climate	5,18	6	-0,71	4594372

European Bee Models in Denmark	Mean Species Richness Change	Median Species Richness Change	Species Richness Skewness Change	Total Species Richness Change
Current to SSP 1-2.6 climate	12,14	12	0,03	44390
Current to SSP 3-7.0 climate	14	13	0,08	51270

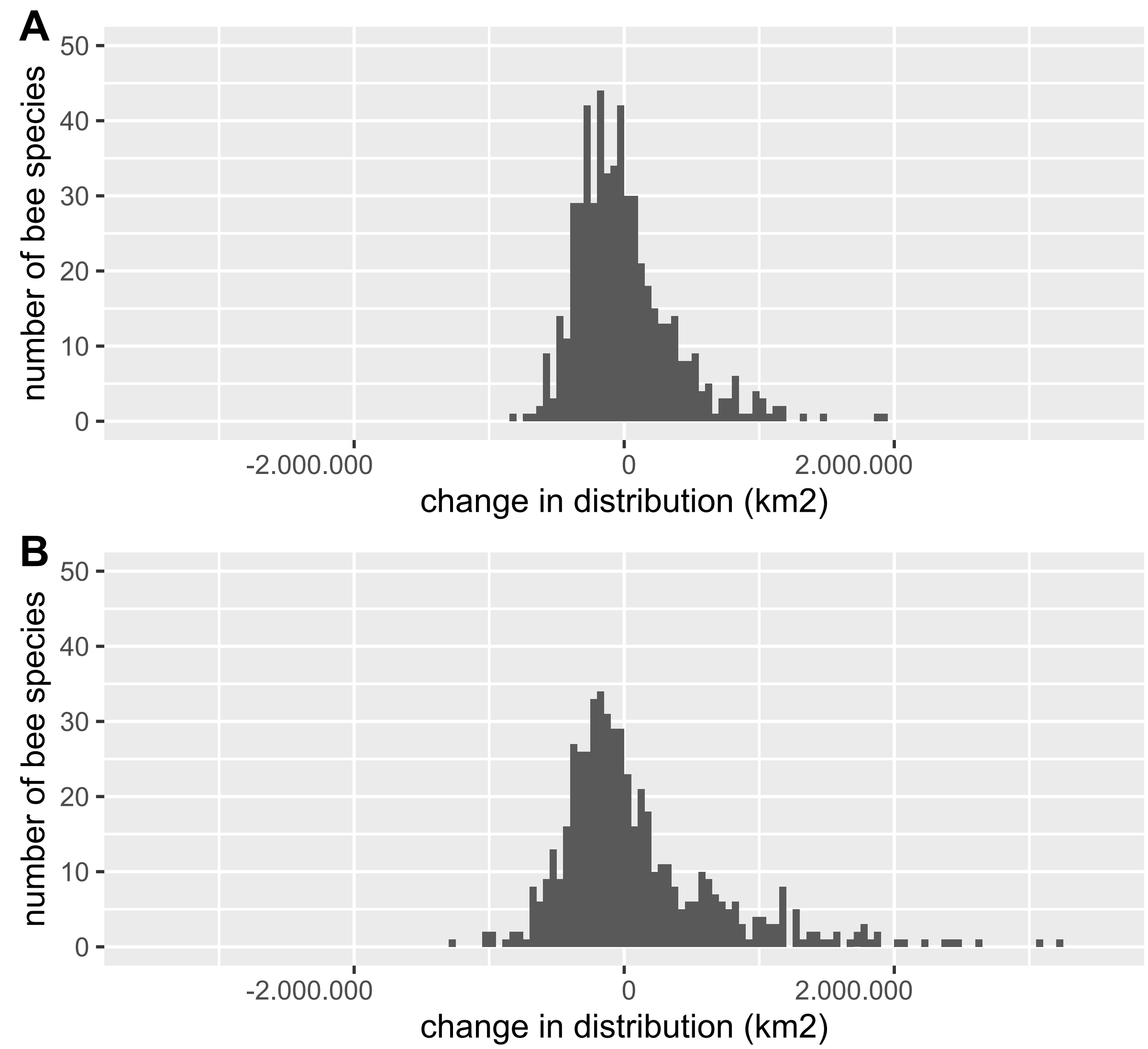

Figure 4 Histograms of the change in area of predicted potential presence for the 543 European bee species, each bar represents species that falls within a change of 50.000 km². A illustrates the change from current climate to the climate predicted for 2060 according to SSP 1-2.6, with B illustrating the change from current climate to the climate predicted for 2060 according to SSP 3-7.0.

The distribution of the change of predicted distribution of species is clearly skewed (Figure 4, Table 5), with most species losing distribution cells, but those with the greater change gaining distribution area. This effect is noticeable going to the SSP 1-2.6 climate (Figure 4A), since the mean and median changes are negative, while the skewness is positive signifying a right skewed distribution. The change going from current to SSP 3-7.0 climate (Figure 4B) shows this effect more pronounced, with the mean being positive despite half the species losing distribution area (Table 5).

Table 5 displaying the mean, median and skewness change in distribution area of the 543 European bee species modelled in this project.

	Mean Change in Distribution (km²)	Median Change in Distribution (km²)	Skewness of Change in Distribution
Current to SSP 1-2.6	-2574	-69889	1,38
Current to SSP 3-7.0	99418	-61500	1,64

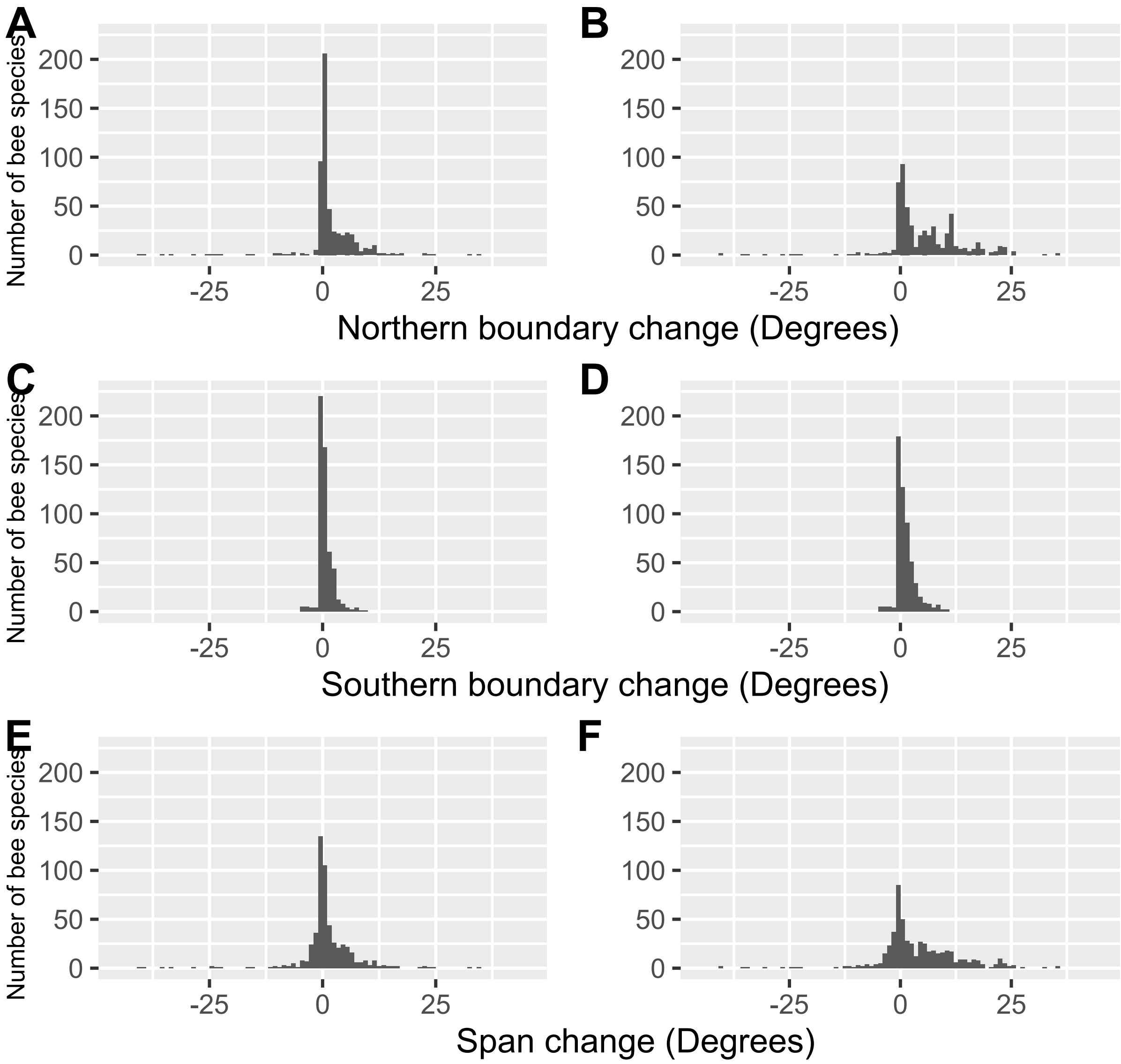

Figure 5 Histograms depicting the changes in the northern (A,B) and southern (C,D) boundaries as well as the span (E,F) of the predicted distribution off the 543 European bee species. The histograms on the left (A,C,E) illustrate the change from current climate to the climate predicted for 2060 according to SSP 1-2.6, with the histograms on the right (B,D,F) illustrating the change from current climate to the climate predicted for 2060 according to SSP 3-7.0.

There is a tendency for the predicted boundaries to move north, with all average changes being positive. Examining the skewness, some species move their northern boundaries southwards to an extreme degree. While the bee species on average travel northward, some fail to do so. On average the northern boundaries move north more than southern boundaries do, which leads to an average expanse of distribution span.

Table 6 displaying the mean, median and skewness of changes in northern and southern boundaries as well as changes to distribution span in degrees.

Northern Boundary Change	Mean	Median	Skewness
Current to SSP 1-2.6	1,74	0,75	-1,53
Current to SSP 3-7.0	5,18	2,92	-0,59

Southern Boundary Change	Mean	Median	Skewness
Current to SSP 1-2.6	0,66	0,08	1,55
Current to SSP 3-7.0	1,19	0,46	1,48

Span Change	Mean	Median	Skewness
Current to SSP 1-2.6	1,09	0,38	-1,15
Current to SSP 3-7.0	3,99	1,71	-0,33

Models of Danish bee and plant species

Overview of three climate scenarios for Danish bee species

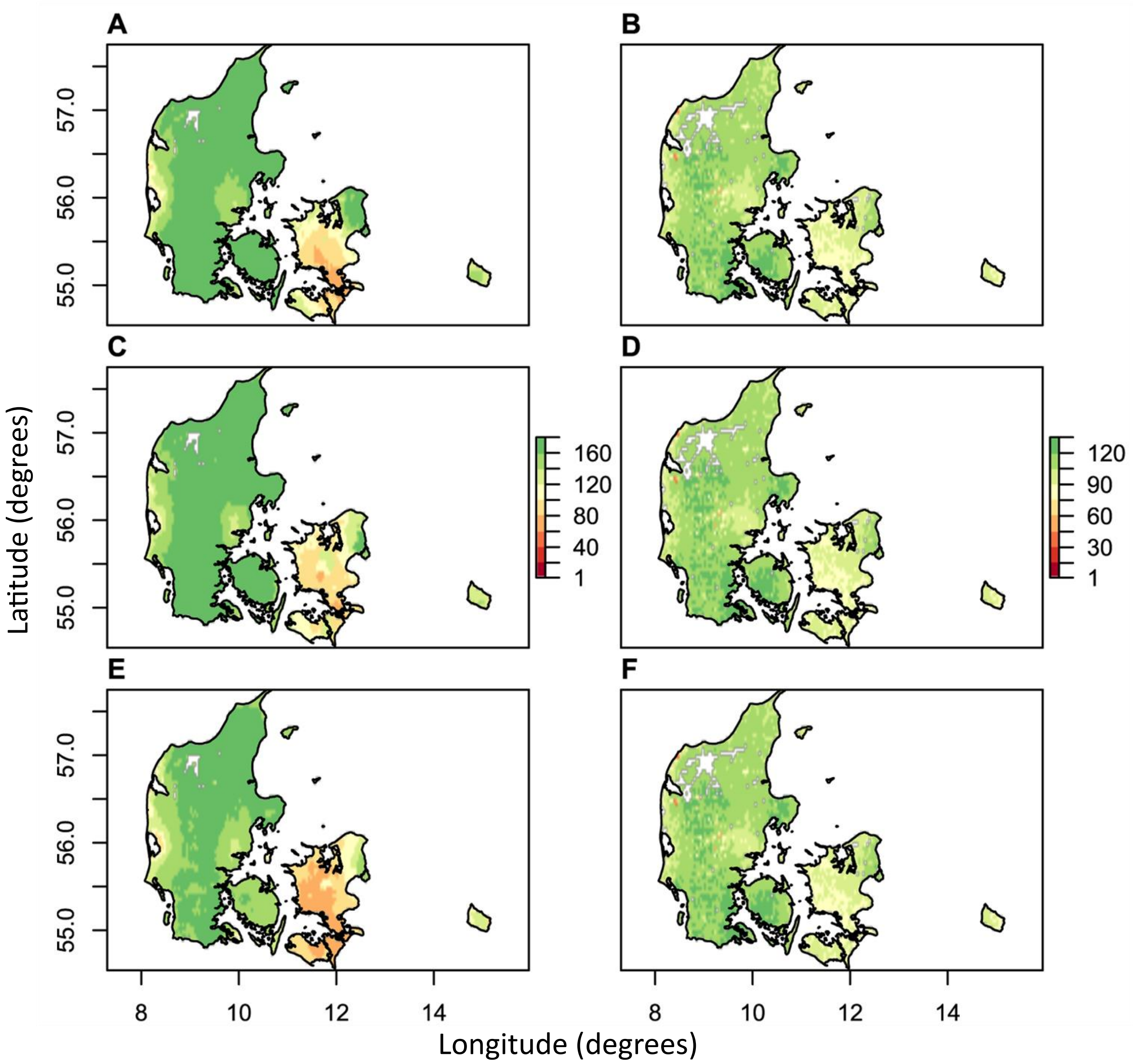

Figure 6 Potential Danish bee (A,C,E) and host-plant (B,D,F) species richness in Denmark using three different scenarios current climate (A,B), SSP 1-2.6 (C,D) and SSP 3-7.0 (E,F). Values under 1 have been coloured white, with values above 0 coloured from red to green in increments of 20.

When just looking at the species for whom we have interaction data (Rasmussen, Schmidt, and Madsen 2016), there is still higher bee species richness in Jutland and on Fyn, compared to Zealand (Figure 6 A,C,E). Furthermore, a similar trend is present in the distribution of potential plant species richness, although the spread over all is more even (Figure 6 B,D,F). When looking at the different scenarios Northern Jutland remains an area of high potential species richness regardless of which climate scenario is modelled. Losses in potential species richness occurs on Zealand along with the west coast of Jutland, especially when modelling the SSP 3-7.0 scenario (Figure 6E). In terms of potential plant species richness, no overall trend in the changes to the distribution of richness is noticeable, this is backed by the mean, median and skewness being almost unchanged (Table 7)

Table 7 displays the mean, median, skewness and total of the species richness of the Danish bees, and the plants they interact with.

Danish Bee Models	Mean Species Richness	Median Species Richness	Species Richness Skewness	Total Species Richness
Current climate	155,26	169	-1,54	567648
SSP 1-2.6 climate	153,66	168	-1,39	561783
SSP 3-7.0 climate	143,32	160	-1,27	523996

Plant Models	Mean Species Richness	Median Species Richness	Species Richness Skewness	Total Species Richness
Current climate	107,28	109	-0,79	366371
SSP 1-2.6 climate	108,28	110	-0,80	369782
SSP 3-7.0 climate	108,28	110	-0,80	369784

Including host plant distribution as restricting factor for bee distribution

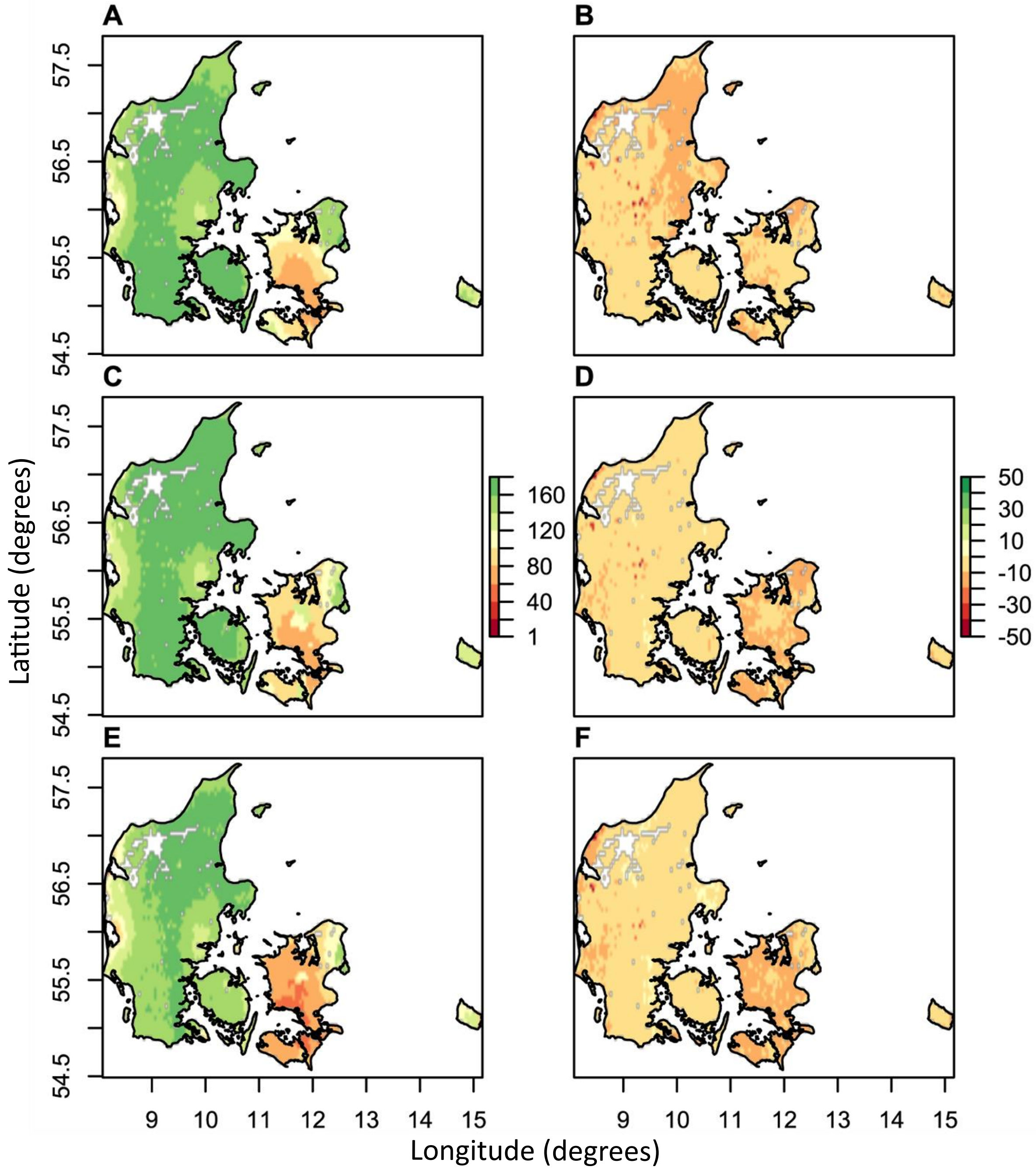

Figure 7 The Danish bee species richness when restricting the bee species distribution to cells containing at least one of its host plants (A,C,E), along with the changes going from unrestricted Danish bee species distribution to distribution restricted by distribution of host plants (B,D,F). the maps correspond to current (A,B), SSP 1-2.6 (C,D) and SSP 3-7.0 (E,F).

When looking at the maps of how interactions change the potential species richness distribution in Denmark, it is apparent that the distribution pattern remains unchanged, with the higher bee species richness in Jutland compared to Zealand (Figure 7 A,C,E). When examining the changes in each cell value from not accounting for interactions, to accounting for interactions, Jutland sees a larger loss of species due to interactions in the current climate than it would in any of the scenarios (Figure 7 B,D,F), while Zealand seems unaffected by interactions in the current climate, though the losses become more potent in the future scenarios.

Changes in distribution area correlates to number of interaction partners

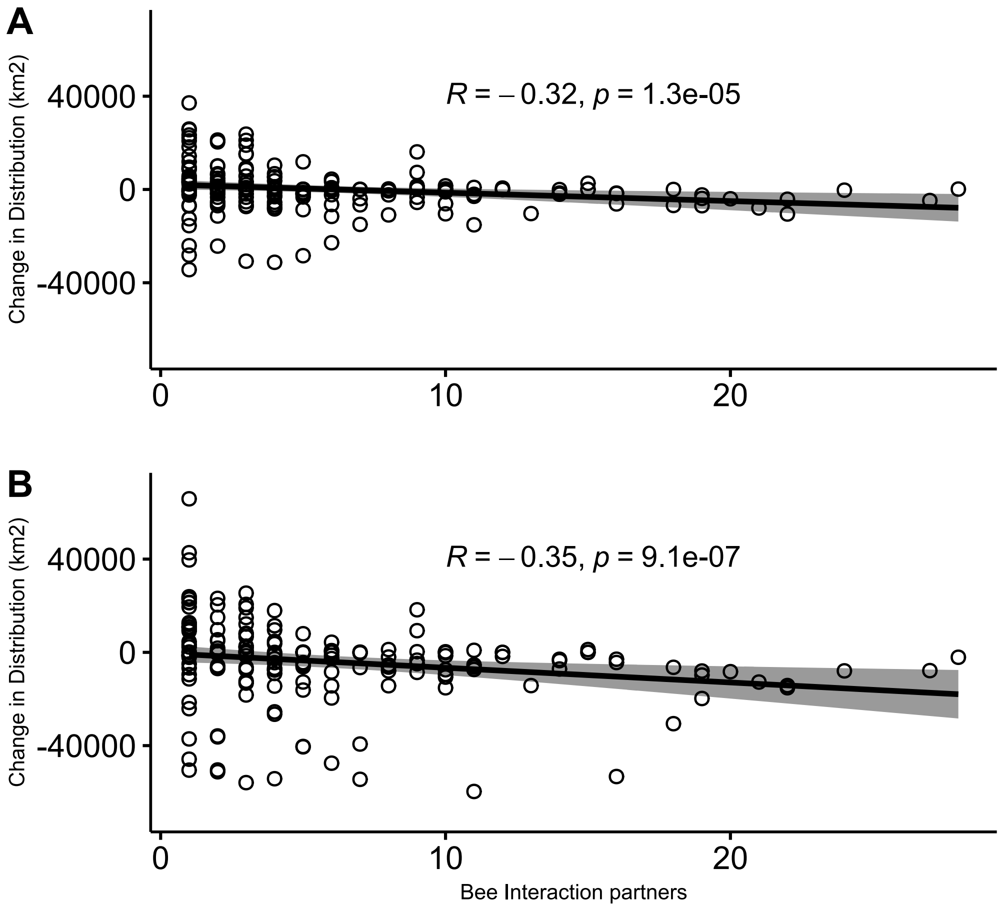

Figure 8 Scatterplot of the spearman correlation between number of interaction partners and change in distribution area in km². Each dot represents a bee species with a number of known interaction partners between 1 and 28. Each of the plots depicts the changes in distribution for one of the climate scenarios SSP1 (A) and SSP3 (B). the black line represents the linear regression between the two variables, with the grey shadow displaying the confidence interval.

To investigate the association between loss of distribution area and number of host plants, the correlation between specialisation of the bee species and the changes is their distributions area was computed. The species that have more interaction partners also saw the biggest losses in terms of distribution area due to changes in climate, both for the SSP1 (p = 1,3e-05, rho = -0,32) and SSP3 (p = 9,1e-07, rho = -0,35) scenario. We do see that there is a higher variety of responses within the species with a lower number of interaction partners.

Importance of certain biotopes on bee species richness across Denmark

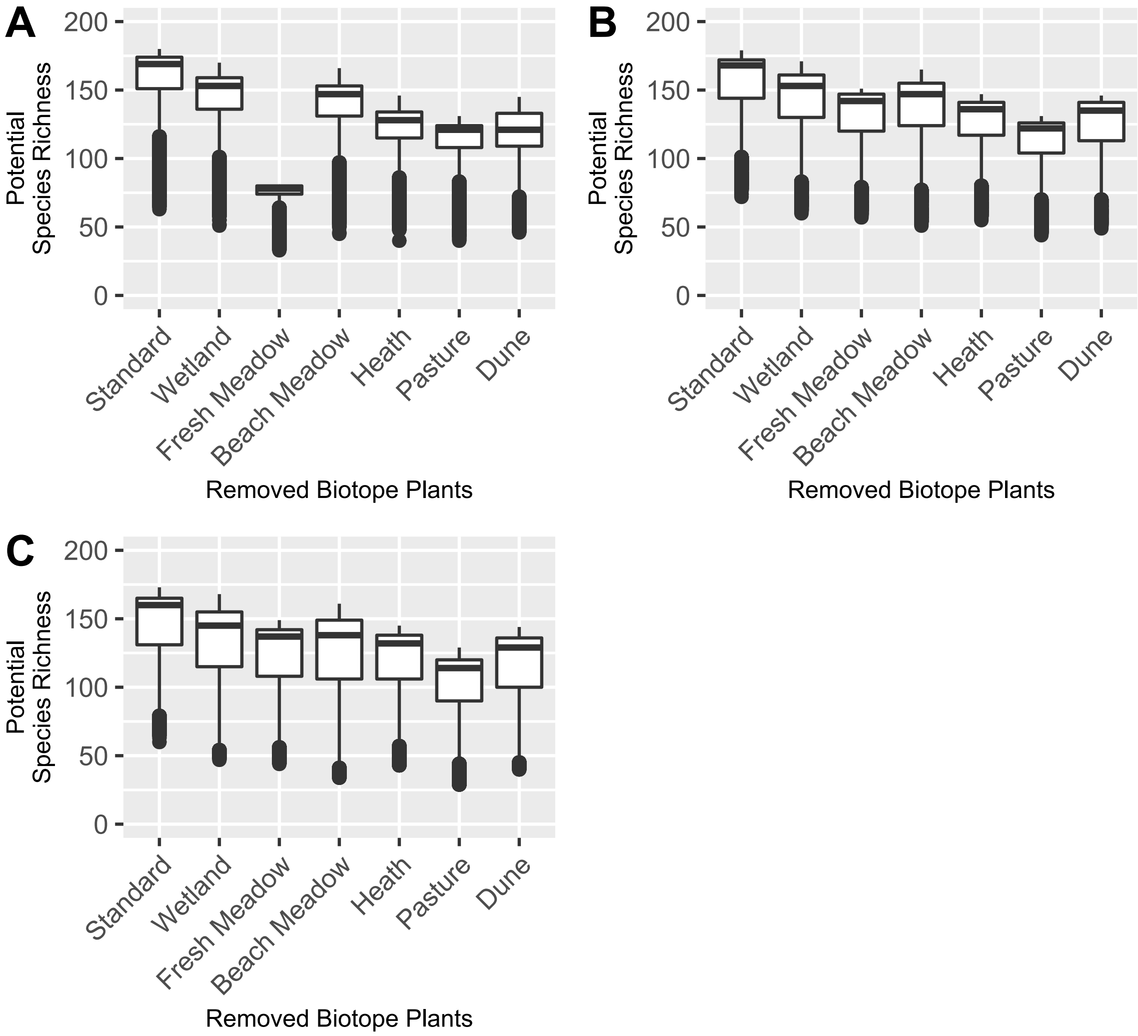

Figure 9 The values of the raster when removing the characteristic plants for the §3 protected biotopes, wetland, fresh meadow, beach meadow, heath, pasture and dune, the box labelled standard is the raster with all the plants. The three plots are depicting the three scenarios: Current (A), SSP 1-2.6 (B), and SSP 3-7.0 (C).

A boxplot consists of three elements, the box which contains the middle 50% of data from the 25th percentile to the 75th with a line inside the box marking the average value, the tails marking the 25% of data either side of the box, and the outliers, here represented with circles, these are data that fall outside of 1,5 times interquartile range. The biotope that is most important for upholding the species diversity of Danish bees seems to be fresh meadow, though that effect seems to be diminished in the future scenarios. Pasture seems to be consistently important as well, both biotopes are types of grassland. The changes in species richness are significant no matter which of the biotopes are eliminated, this is true for all the scenarios.

Changes from current climate to the two future scenarios

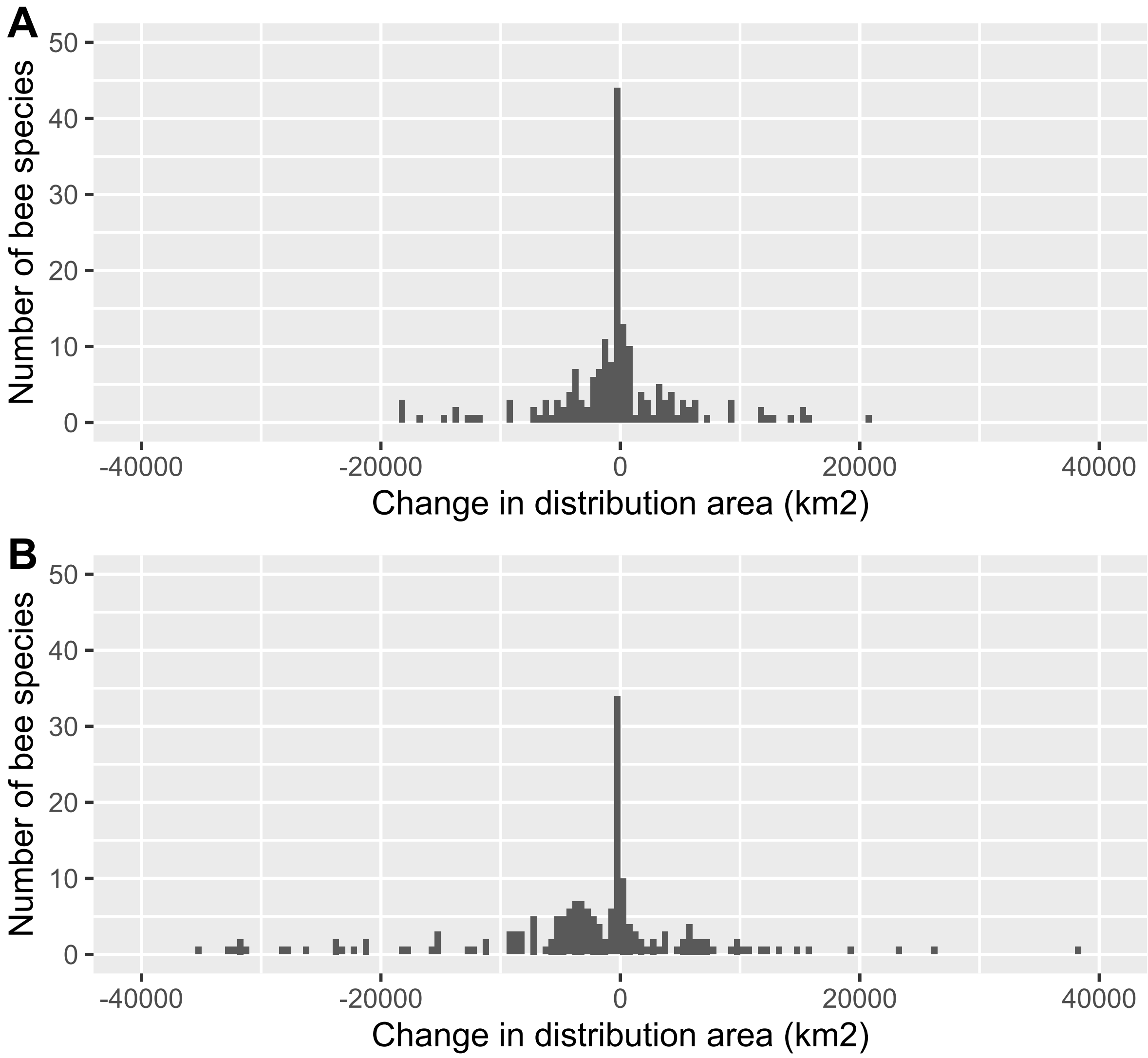

Figure 10 Histograms of the change in area with predicted presence of the 185 Danish bee species, each bar represents species that fall within a change of 500 km². A illustrates the change from current climate to the climate predicted for 2060 according to SSP 1-2.6, with B illustrating the change from current climate to the climate predicted for 2060 according to SSP 3-7.0.

When looking at the different bee species changing their distribution from current climate to the two climate scenarios, the changes are small and normally distributed around zero going to SSP 1-2.6. The changes for going to SSP 3-7.0 is leaning towards a loss in distribution, with the average, median, and skewness being negative.

Table 8 displaying the mean, median and skewness change in distribution area of the 185 Danish bee species modelled in this project.

	Mean Change in Distribution (km²)	Median Change in Distribution (km²)	Skewness of Change in Distribution

Current to SSP 1-2.6	-376,6	0	-0,02
Current to SSP 3-7.0	-2802,8	-857,8	-0,51

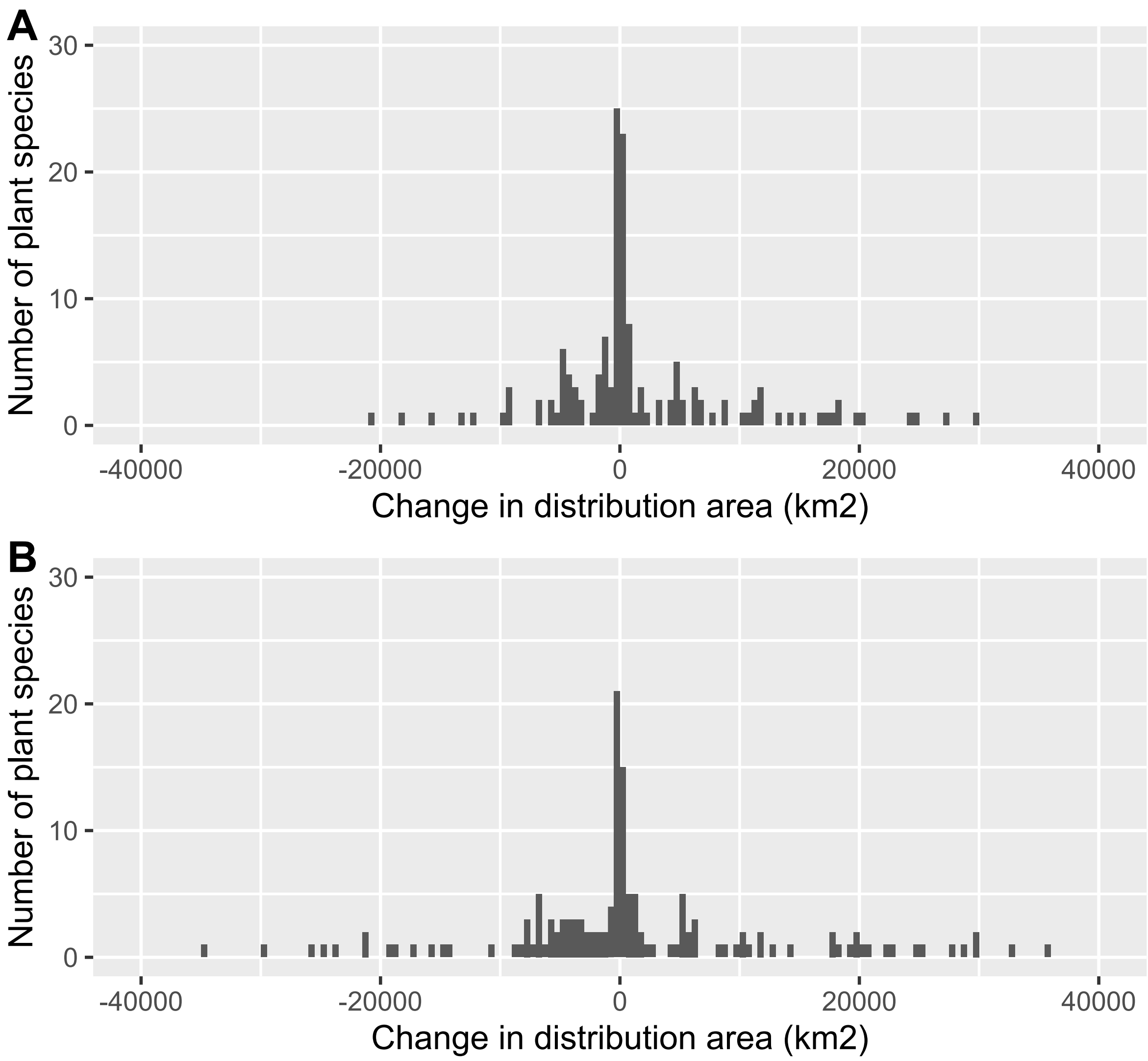

Figure 11 Histograms of the change in area with predicted presence of the 163 Danish plant species, each bar represents species that fall within a change of 500 km². A illustrates the change from current climate to the climate predicted for 2060 according to SSP 1-2.6, with B illustrating the change from current climate to the climate predicted for 2060 according to SSP 3-7.0.

When examining the changes in distribution area of plant species in Denmark, the general trend is an increase in distribution area. The general area gain is greater when modelling the SSP 1-2.6 climate (Figure 11A), with an increased average and skewness. For the SSP 3-7.0 climate a similar trend is seen (Figure 11B), although to a lesser degree.

Table 9 displaying the mean, median and skewness change in distribution area of the 163 Danish plant species modelled in this project.

	Mean Change in Distribution (km²)	Median Change in Distribution (km²)	Skewness of Change in Distribution
Current to SSP 1-2.6	1694,8	11,8	0,92
Current to SSP 3-7.0	1032	0	0,30

Distribution of endangered species with interaction data

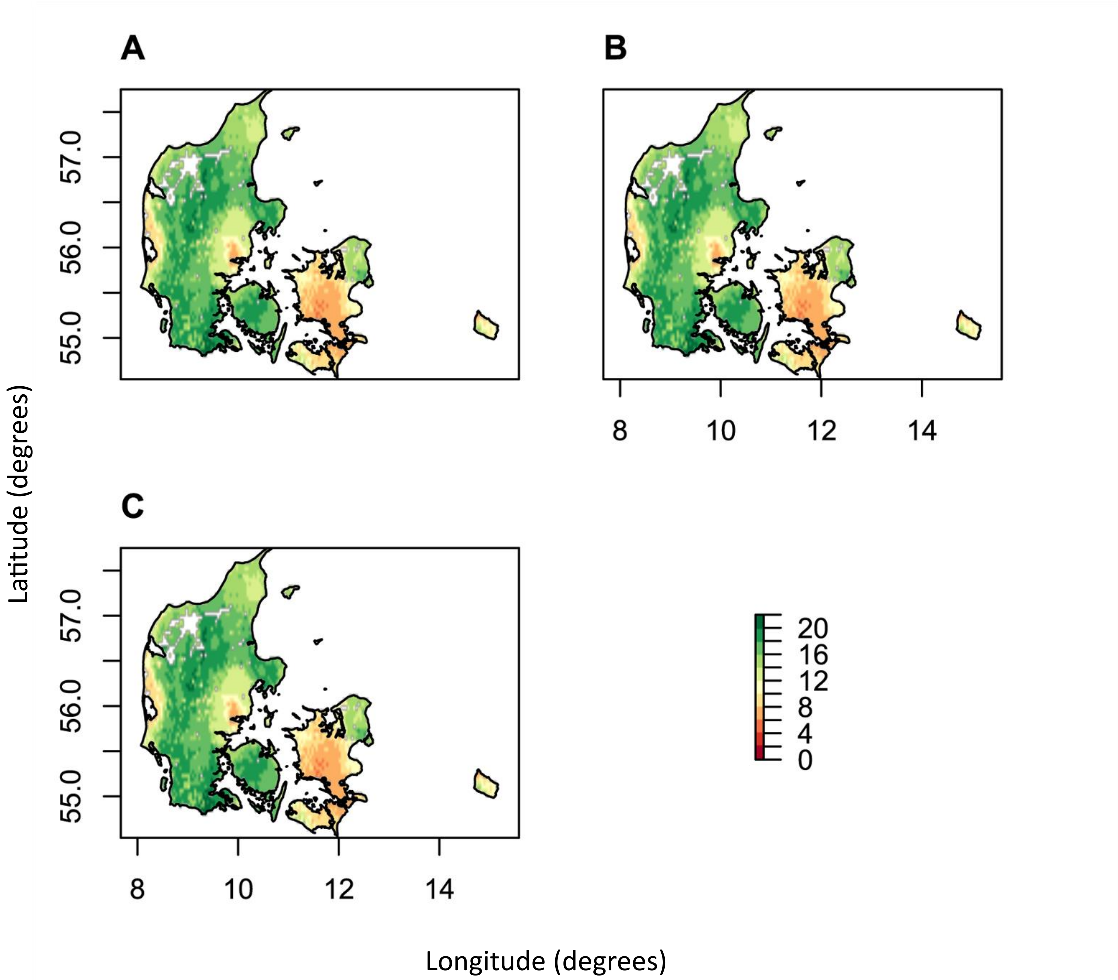

Figure 12 illustrates the distribution of the 23 endangered species modelled with interactions between host plant and bee species. The highest value on any of these distribution maps is 22, meaning no cell potentially contains all endangered species, likewise the lowest value of 2 no cell has no potential presence of an endangered species.

The distribution of the endangered Danish species, for which interaction data was available, shows that they tend to follow the distribution patterns of the general species richness. The endangered species extend their distribution hot spot further down into central Jutland, but not as far north of the Limfjord, compared to the general distribution.

Model summary

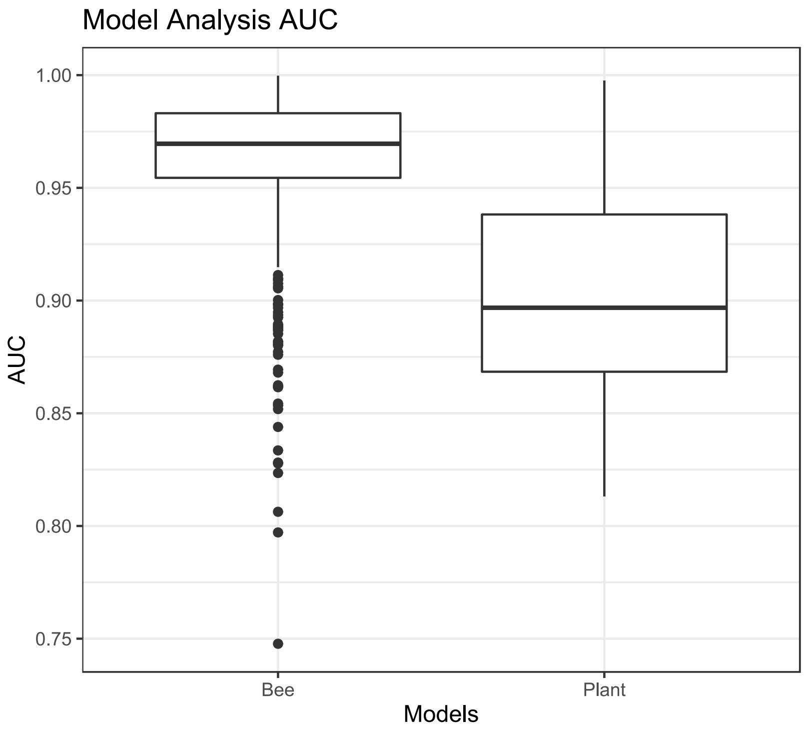

Figure 13 Boxplot of Area Under Curve (AUC) for both bee models and plant models.

The Area Under Curve (AUC) is a way to determine whether the models are better than random at predicting potential presences, with values over 0,5 signifying a better than random model. The models in this project all have an AUC value above 0,7, with most of the models reaching AUC values of above 0,85.

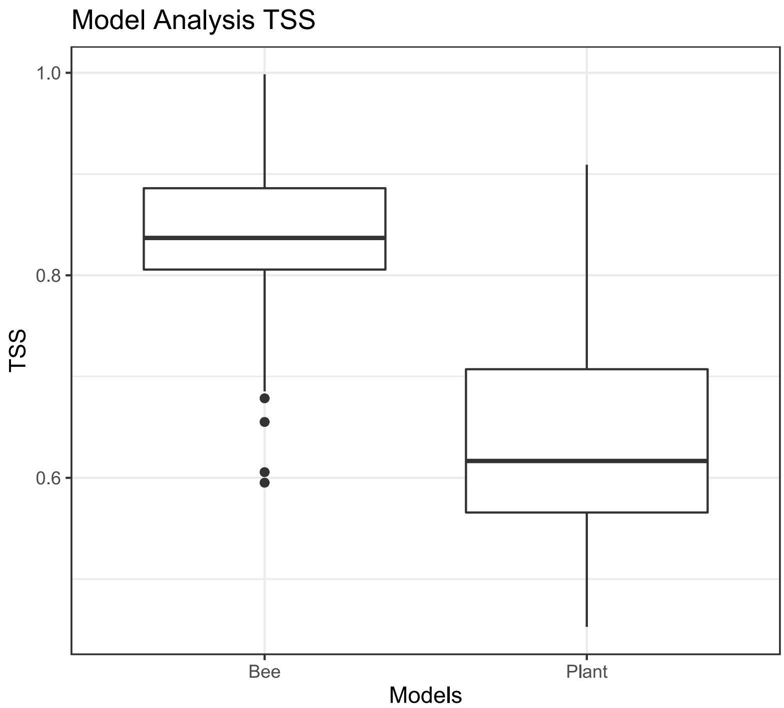

Figure 14 boxplot of True Skill Statistics (TSS) for both bee models and plant models.

The True Skill Statistic (TSS) is a result of subtracting one from the sum of the models true positive and negative rate, meaning that a TSS of 0,5 means that both true positive rates and true negative rates must be at least 0,5. This means that with most bee models having a TSS of above 0,8, can only happen when both rates are at least 0,8.

Table 10 Mean percentage contribution of variables on the bee and plant species distribution models, with the standard error in parenthesis, ordered by contribution to bee distribution models.

	Bee Mean Variable Contribution (SE)	*Danish Bee Mean Variable Contribution (SE)*	Plant Mean Variable Contribution (SE)
Precipitation Seasonality (Bio15)	13,30% (3,56)	12,59% (2,80)	5,45% (1,36)
Mean Temp Driest Quarter (Bio9)	12,17% (4,09)	7,59% (0,88)	7,10% (1,82)
Temp Seasonality (Bio4)	11,21% (2,57)	12,27% (2,24)	13,25% (2,58)
Temp Annual Range (Bio7)	10,09% (2,97)	17,49% (3,49)	11,79% (2,65)
Min Temp Coldest Month (Bio6)	8,47% (3,06)	5,71% (1,31)	0,99% (0,40)

Mean Temp Wettest Quarter (Bio8)	6,36% (1,84)	5,84% (1,27)	2,96% (0,82)
Max Temp Warmest Month (Bio5)	4,56% (1,75)	9,13% (2,24)	3,85% (1,36)
Mean Temp Warmest Quarter (Bio10)	4,22% (1,45)	7,00% (1,43)	8,00% (2,32)
Isothermality (Bio3)	4,17% (1,78)	1,58% (0,56)	2,88% (0,77)
Mean Temp Coldest Quarter (Bio11)	4,05% (2,47)	0,91% (0,24)	7,99% (2,41)
Precipitation Driest Month (Bio14)	4,04% (1,37)	5,15% (1,26)	4,02% (1,97)
Mean Diurnal Range (Bio2)	3,57% (0,92)	5,57% (1,00)	2,29% (0,85)
Precipitation Warmest Quarter (Bio18)	3,34% (0,31)	3,19% (0,79)	5,63% (1,62)
Annual Mean Temp (Bio1)	2,97% (1,86)	1,47% (0,70)	4,11% (1,26)
Precipitation Driest Quarter (Bio17)	2,47% (1,47)	1,58% (0,75)	4,03% (1,69)
Precipitation Coldest Quarter (Bio19)	1,80% (1,04)	0,73% (0,24)	1,82% (0,62)
Precipitation Wettest Month (Bio13)	1,64% (0,70)	1,56% (0,57)	0,45% (0,36)
Annual Precipitation (Bio12)	0,92% (0,93)	0,23% (0,12)	0,20% (0,11)
Precipitation Wettest Quarter (Bio16)	0,68% (0,31)	0,41% (0,17)	0,24% (0,09)
N			1,73% (0,69)
P			6,76% (1,62)
pH H2O			4,46% (1,34)

Discussion

European models

When mapping the distribution of the potential species richness of bees across Europe, it becomes clear that the species are not evenly distributed throughout Europe. This is backed up by the statistics that show us a left skewed distribution. Specifically we see that species richness is accumulating in Northern Europe, which supports research showing losses of bumblebees in Southern Europe and gains in Northern Europe (Marshall et al. 2018). The study by Marshall et al. (2018) looked at the differences when using climate only models, like the models used in present project, and using models that look land use into account. They found that climate only models tended to show a greater range gain, compared with the other models.

Research from the Netherlands, looked into correlations between host plant population trends and wild bee population trends (Scheper et al. 2014). They found that the population trend of the pollen host plants was among the most important factors for predicting the population trend of bee populations, with the relationship between the factors being positive. Research into the distribution of pollinators in England shows that the majority of their pollinators were found in the southern part of England (Polce et al. 2013), backing up the predictions of this projects models of European bee species. These areas that are predicted to have a higher-than-average species richness are important to conserve, as they can serve as safe havens for those species. Predictions showing that these areas retain high species richness for years to come, further implies that their conservation is a good investment into future pollination.

Similar to the findings of this project, other studies indicate that most species move their distribution towards the poles as a result of climate change (Feehan, Harley, and van Minnen 2009). However a study done on hoverflies in south-east Europe, found that these pollinators were resilient to climate changes effect on their distribution, with some species gaining distribution area and others losing (Miličić et al. 2018). This is also seen in the changes of distribution in this project, where most of the species' distribution changed little, although the average change was a decline in distribution area. A few species in this project saw a massive gain in distribution area, namely moving into Russia, which did not have a large bee presence in the occurrence data, this was partly why some species gained a large amount of distribution. Another study done on changes to the distribution of bumblebees, further illustrates that certain pollinators are resilient to climate changes, and that this might be to their own detriment (Kerr et al. 2015). The study by Kerr et al. (2015) shows that bumblebee species that were predicted to change their northern and southern boundaries due to changes in climate, did not follow the predicted movements for the northern boundary leading to a decreased distribution range as the southern boundary closed in on the northern boundary. In this project most bee species are predicted to move their northern and southern boundaries northwards. Most species are even predicted to have an expanding distribution range, but other studies suggest that pollinators do not adhere to these changes in boundaries, further observational studies might be necessary.

This project found that on average the most important climate variables, for European bee species, had to do with seasonality, meaning that species were more impacted by changes throughout the year than any specific extreme. The outlier in this being the highest temperature during the driest month, which indicates that dehydration could be a concern for bee species in general. A study shows that heat stress can impact the growth and foraging habits of bees, particularly more individuals are sent to forage for water (Zhao et al. 2021). With this being a response to heat stress, the combination of high temperatures and low precipitation would be a limiting factor for species distribution.

Danish models

Other studies have found that the population trends of hostplants, had a great correlation with the population trends of the bees that visit them (Scheper et al. 2014). This project found that while the lack of potential host plants was a limiting factor for bee distribution, it did not dramatically alter the distribution patterns in Denmark. Furthermore, the study by Scheper et al. (2014) points out that the level of specialisation of each bee species, does not help explain population trends. Present project found that there is a correlation between specialisation and climate impact on bees, with those bees that are less specialised losing more potential distribution area compared to those that have fewer potential partners. This stands in contrast to the findings of Scheper et al. (2014), while they do focus on population trends rather than distribution, these two factors are linked. The idea that generalists with a higher number of interaction partners, should be more prone to loss of distribution area is counter intuitive, but could be explained by those species having more potential distributions in the first place. With Denmark using most of its arable land for agriculture, it leaves little space for wild bees, with fragmentation and habitat destruction being a major threat (Jørgensen 2011). Another concern is whether managed bees, kept by beekeepers to fulfil the role of pollinators in agricultural areas, can have a negative effect on the wild bees in an area. This is difficult to say for certain, with studies reporting mixed results (Mallinger, Gaines-Day, and Gratton 2017). However it has been shown that the presence of native wild bees, can benefit agriculture by raising the potential crop yield without a need of increasing the intensity (Brittain et al. 2013). It is even possible that native wild bees can function as a reserve for pollination efforts, with wild bees being capable of providing more than 90% of the pollination service currently provided by managed bees (Winfree et al. 2007). Agriculture does however also provide a great threat to bees, in the form of pesticides. The use of pesticides on crops is a threat to wild as well as managed bees, as it can affect the cognition of the bees, and seeing as their foraging behaviour is complex and difficult to accomplish with an energy gain, the cognitive disturbance of pesticides can have grave impacts on bees (Klein et al. 2017). With how present project predicts the potential distribution of bees, both European and Danish, it makes sense to focus Danish conservation efforts on Jutland. By having areas dedicated to conserving bees and their host plants throughout the Danish mainland, we could ensure that a lot of bee species have a migration path at least until 2060, and this would not even require new areas to be protected but would mainly require an additional effort in the meadows and heaths already protected under Danish law. This is further backed by other studies that show that arable land is important to certain bee species (Marshall et al. 2018) and that grassing, mowing and cutting of grasslands such as meadows can artificially end flowering seasons early, which is a detriment to bee species with later flight periods (Scheper et al. 2014). Meadows need to be, artificially or trough grazing, kept at a stage of high biodiversity (Dahlström, Iuga, and Lennartsson 2013). This means that areas like that can provide bees with an abundance of different host plants sustaining many different species. Furthermore, the endangered species modelled for Denmark in this project, mostly follow the same distribution as the general bee species richness, meaning that efforts to protect bees generally will also protect the more vulnerable species. It is important to bring up the need for conservation of bees, and insects in general, as they are largely overlooked in this regard (Winfree 2010). The distribution of bees shown by the

SDMs in this project do not agree with the distribution pattern of Rasmussen, Schmidt, and Madsen (2016), they found that the majority of Danish species were present in North-eastern Zealand and Bornholm, compared with this project's findings that the majority of species are found in mid to north Jutland. This discrepancy could be a result of Rasmussen, Schmidt, and Madsen (2016) not modelling the species distribution, but rather looking at the occurrence points and their distribution. This gives them a good idea of the actual and historical distribution of the Danish bees, while present project shows the potential and future distribution of Danish bees.

Model summary

The models crafted in this project have a high AUC, meaning they are better than random at predicting potential presences. That said they are potentially overfitted, due to the choices made based on time constraints and limited computing power.

The choice to use the default settings in the maxent program, and thus using all potential connections between the prediction variables, can potentially lead to overfitting (Polce et al. 2013). But without the time to test whether using settings such as 'Hinge', as done in Polce et al. (2013), it was decided to use a wider setting. Furthermore, this decision was made due to the number of species modelled in this project, with each species having the potential to react to different connections between climate variables. Attempting to optimize the model to all 543 bee species is impossible and would most likely result in models optimized for the more prevalent species. At the same time, making individually optimized models is not possible for this number of species in the time allowed for this project. Another factor that could contribute to potential overfitting of this project's models, is the use of all 19 climate variables as predicting variables. In another papers it is pointed out how lowering the number of climate variables, without losing much of the variation, is preferable to avoid overfitting (Silva et al. 2014). Again, the paper only looks at the distribution of a single species, making the process of eliminating the climate variables that do not affect it easier, compared to the process of trimming climate variables for a much higher number of species. Just by looking at the differences between the important variables for the European models and the Danish models, the species do not react to all climate variables equally.

Conclusion

Pollinators are in worldwide decline, and simultaneously in increasing demand for agricultural purposes. This means that initiatives to conserve pollinators must be considered. The most important pollinator in nature is the bee and this project maps the potential species richness throughout Europe, showing that England, Belgium, the Netherlands, and Denmark are hotspots for European bee species richness. Furthermore, it was shown that these countries remain hotspots in future climates, making them ideal for conservation efforts. Diving into Denmark it was shown that Jutland is the part of Denmark with the most potential bee species richness, and as such should be the focus for conservation efforts. It was also shown that interactions do not restrict Danish bee species dramatically, but that efforts should be focused on meadows, as this biotope harbours most of the interaction partners for Danish bees. Furthermore, the distribution of endangered species was shown to follow the general distribution in Denmark, meaning that efforts to help the general species richness would also cover habitats potentially housing endangered species.

References

Aizen, Marcelo A., Lucas A. Garibaldi, Saul A. Cunningham, and Alexandra M. Klein. 2008. 'Long-Term Global Trends in Crop Yield and Production Reveal No Current Pollination Shortage but Increasing Pollinator Dependency'. *Current Biology* 18 (20): 1572–75. https://doi.org/10.1016/j.cub.2008.08.066.

Ballabio, Cristiano, Emanuele Lugato, Oihane Fernández-Ugalde, Alberto Orgiazzi, Arwyn Jones, Pasquale Borrelli, Luca Montanarella, and Panos Panagos. 2019. 'Mapping LUCAS Topsoil Chemical Properties at European Scale Using Gaussian Process Regression'. *Geoderma* 355 (December): 113912. https://doi.org/10.1016/j.geoderma.2019.113912.

Brittain, Claire, Neal Williams, Claire Kremen, and Alexandra-Maria Klein. 2013. 'Synergistic Effects of Non-Apis Bees and Honey Bees for Pollination Services'. *Proceedings of the Royal Society B: Biological Sciences* 280 (1754): 20122767. https://doi.org/10.1098/rspb.2012.2767.

Dahlström, Anna, Ana-maria Iuga, and Tommy Lennartsson. 2013. 'Managing Biodiversity Rich Hay Meadows in the EU: A Comparison of Swedish and Romanian Grasslands'. *Environmental Conservation* 40 (2): 194–205. http://dx.doi.org.zorac.aub.aau.dk/10.1017/S0376892912000458.

Doane, David P., and Lori E. Seward. 2011. 'Measuring Skewness: A Forgotten Statistic?' *Journal of Statistics Education* 19 (2): 3. https://doi.org/10.1080/10691898.2011.11889611.

Feehan, Jane, Mike Harley, and Jell van Minnen. 2009. 'Climate Change in Europe. 1. Impact on Terrestrial Ecosystems and Biodiversity. A Review'. *Agronomy for Sustainable Development* 29 (3): 409–21. https://doi.org/10.1051/agro:2008066.

Fick, Stephen E., and Robert J. Hijmans. 2017. 'WorldClim 2: New 1-Km Spatial Resolution Climate Surfaces for Global Land Areas'. *International Journal of Climatology* 37 (12): 4302–15. https://doi.org/10.1002/joc.5086.

Fredshavn, Jesper Reinholt, Rasmus Ejrnæs, Bettina Nygaard, Danmark, and Danmarks Miljøundersøgelser. 2009. *Naturtilstand på terrestriske naturarealer - besigtigelser af § 3-arealer*. Danmarks Miljøundersøgelser.

GBIF. 2020a. 'GBIF Occurrence Download of Andrenidae'. Database. GBIF. 3 November 2020. https://doi.org/10.15468/dl.xw6t7p.

———. 2020b. 'GBIF Occurrence Download of Apidae'. Database. GBIF. 3 November 2020. https://doi.org/10.15468/dl.dxpy74.

———. 2020c. 'GBIF Occurrence Download of Colletidae'. Database. GBIF. 3 November 2020. https://doi.org/10.15468/dl.e664pu.

———. 2020d. 'GBIF Occurrence Download of Halictidae'. Database. GBIF. 3 November 2020. https://doi.org/10.15468/dl.n7wk35.

———. 2020e. 'GBIF Occurrence Download of Megachilidae'. Database. GBIF. 3 November 2020. https://doi.org/10.15468/dl.me7k3k.

———. 2020f. 'GBIF Occurrence Download of Melittidae'. Database. GBIF. 3 November 2020. https://doi.org/10.15468/dl.wp9mcb.

Hausfather, Zeke. 2019. 'CMIP6: The next Generation of Climate Models Explained'. Carbon Brief. 2 December 2019. https://www.carbonbrief.org/cmip6-the-next-generation-of-climate-models-explained.

Hijmans, Robert J., Steven Phillips, John Leathwick, and Jane Elith. 2020. *Dismo: Species Distribution Modeling*. R package version 1.3-3. R. https://CRAN.R-project.org/package=dismo.

Jiménez-valverde, A., A. T. Peterson, J. Soberón, J. M. Overton, P. Aragón, and J. M. Lobo. 2011. 'Use of Niche Models in Invasive Species Risk Assessments'. *Biological Invasions* 13 (12): 2785–97. http://dx.doi.org.zorac.aub.aau.dk/10.1007/s10530-011-9963-4.

Jørgensen, Asger Søgaard. 2011. 'Beekeeping In An Agricultural Landscape—Denmark'. *Bee World* 88 (4): 74–76. https://doi.org/10.1080/0005772X.2011.11417429.

Kearney, Michael, and Warren Porter. 2009. 'Mechanistic Niche Modelling: Combining Physiological and Spatial Data to Predict Species' Ranges'. *Ecology Letters* 12 (4): 334–50. https://doi.org/10.1111/j.1461-0248.2008.01277.x.

Kerr, Jeremy T., Alana Pindar, Paul Galpern, Laurence Packer, Simon G. Potts, Stuart M. Roberts, Pierre Rasmont, et al. 2015. 'Climate Change Impacts on Bumblebees Converge across Continents'. *Science* 349 (6244): 177–80. https://doi.org/10.1126/science.aaa7031.

Klein, Simon, Amélie Cabirol, Jean-Marc Devaud, Andrew B. Barron, and Mathieu Lihoreau. 2017. 'Why Bees Are So Vulnerable to Environmental Stressors'. *Trends in Ecology & Evolution* 32 (4): 268–78. https://doi.org/10.1016/j.tree.2016.12.009.

Madsen, H.B. 2019. 'Bier'. Aarhus Universitet, DCE - Nationalt Center fr Miljø og Energi. https://bios.au.dk/forskningraadgivning/temasider/redlistframe/artsgrupperne/oevrige-leddyr/bier/.

Mallinger, Rachel E., Hannah R. Gaines-Day, and Claudio Gratton. 2017. 'Do Managed Bees Have Negative Effects on Wild Bees?: A Systematic Review of the Literature'. *PLOS ONE* 12 (12): e0189268. https://doi.org/10.1371/journal.pone.0189268.

Manzoor, Syed Amir, Geoffrey Griffiths, and Martin Lukac. 2018. 'Species Distribution Model Transferability and Model Grain Size – Finer May Not Always Be Better'. *Scientific Reports (Nature Publisher Group); London* 8 (May): 1–9. http://dx.doi.org/10.1038/s41598-018-25437-1.

Marshall, Leon, Jacobus C. Biesmeijer, Pierre Rasmont, Nicolas J. Vereecken, Libor Dvorak, Una Fitzpatrick, Frédéric Francis, et al. 2018. 'The Interplay of Climate and Land Use Change Affects the Distribution of EU Bumblebees'. *Global Change Biology* 24 (1): 101–16. https://doi.org/10.1111/gcb.13867.

McAulay, Megan K., Saff Z. Killingsworth, and Jessica R. K. Forrest. 2021. 'Understanding Pollen Specialization in Mason Bees: A Case Study of Six Species'. *Oecologia* 195 (3): 559–74. https://doi.org/10.1007/s00442-020-04786-7.

Merow, Cory, Matthew J. Smith, and John A. Silander. 2013. 'A Practical Guide to MaxEnt for Modeling Species' Distributions: What It Does, and Why Inputs and Settings Matter'. *Ecography* 36 (10): 1058–69. https://doi.org/10.1111/j.1600-0587.2013.07872.x.

Michener, Charles D. 2007. *The Bees of the World*. Baltimore, UNITED STATES: Johns Hopkins University Press. http://ebookcentral.proquest.com/lib/aalborguniv-ebooks/detail.action?docID=3318383.

Miličić, Marija, this link will open in a new window Link to external site, Ante Vujić, and Pedro Cardoso. 2018. 'Effects of Climate Change on the Distribution of Hoverfly Species (Diptera: Syrphidae) in Southeast Europe'. *Biodiversity & Conservation* 27 (5): 1173–87. http://dx.doi.org.zorac.aub.aau.dk/10.1007/s10531-017-1486-6.

Miljø- og Fødevareministeriet. 2009. *Naturbeskyttelsesloven*. *Lovtidende A*. Vol. 933. https://www.retsinformation.dk/eli/lta/2009/933.

Miller-Struttmann, Nicole E., Jennifer C. Geib, James D. Franklin, Peter G. Kevan, Ricardo M. Holdo, Diane Ebert-May, Austin M. Lynn, Jessica A. Kettenbach, Elizabeth Hedrick, and Candace Galen. 2015. 'Functional Mismatch in a Bumble Bee Pollination Mutualism under Climate Change'. *Science* 349 (6255): 1541–44. https://doi.org/10.1126/science.aab0868.

Morin, Xavier, and Wilfried Thuiller. 2009. 'Comparing Niche- and Process-Based Models to Reduce Prediction Uncertainty in Species Range Shifts under Climate Change'. *Ecology* 90 (5): 1301–13. https://doi.org/10.1890/08-0134.1.

O'Donnell, Michael S., and Drew A. Ignizio. 2012. 'Bioclimatic Predictors for Supporting Ecological Applications in the Conterminous United States'. Data Series 691. U.S. Geological Survey.

Ollerton, Jeff, Rachael Winfree, and Sam Tarrant. 2011. 'How Many Flowering Plants Are Pollinated by Animals?' *Oikos* 120 (3): 321–26.

Phillips, Steven, Miroslav Dudík, and Robert E. Schapire. 2021. *Maxent Software for Modeling Species Niches and Distributions (Version 3.4.1)* (version 3.4.1).

Polce, Chiara, Mette Termansen, Jesus Aguirre-Gutiérrez, Nigel D. Boatman, Giles E. Budge, Andrew Crowe, Michael P. Garratt, et al. 2013. 'Species Distribution Models for Crop Pollination: A Modelling Framework Applied to Great Britain'. *PLOS ONE* 8 (10): e76308. https://doi.org/10.1371/journal.pone.0076308.

Pollock, Laura J., Reid Tingley, William K. Morris, Nick Golding, Robert B. O'Hara, Kirsten M. Parris, Peter A. Vesk, and Michael A. McCarthy. 2014. 'Understanding Co-Occurrence by Modelling Species Simultaneously with a Joint Species Distribution Model (JSDM)'. *Methods in Ecology and Evolution* 5 (5): 397–406. https://doi.org/10.1111/2041-210X.12180.

Pulliam, H. R. 2000. 'On the Relationship between Niche and Distribution'. *Ecology Letters* 3 (4): 349–61. https://doi.org/10.1046/j.1461-0248.2000.00143.x.

Rasmussen, Claus, Hans Thomsen Schmidt, and Henning Bang Madsen. 2016. 'Distribution, Phenology and Host Plants of Danish Bees (Hymenoptera, Apoidea)'. *Zootaxa* 4212 (1): 1–100. https://doi.org/10.11646/zootaxa.4212.1.1.

Schenk, Mariela, Jochen Krauss, and Andrea Holzschuh. 2018. 'Desynchronizations in Bee–Plant Interactions Cause Severe Fitness Losses in Solitary Bees'. *Journal of Animal Ecology* 87 (1): 139–49. https://doi.org/10.1111/1365-2656.12694.

Scheper, Jeroen, Menno Reemer, Ruud van Kats, Wim A. Ozinga, Giel T. J. van der Linden, Joop H. J. Schaminée, Henk Siepel, and David Kleijn. 2014. 'Museum Specimens Reveal Loss of Pollen Host Plants as Key Factor Driving Wild Bee Decline in The Netherlands'. *Proceedings of the National Academy of Sciences* 111 (49): 17552–57. https://doi.org/10.1073/pnas.1412973111.

Silva, Daniel P., Victor H. Gonzalez, Gabriel A. R. Melo, Mariano Lucia, Leopoldo J. Alvarez, and Paulo De Marco. 2014. 'Seeking the Flowers for the Bees: Integrating Biotic Interactions into Niche Models to Assess the Distribution of the Exotic Bee Species Lithurgus Huberi in South America'. *Ecological Modelling* 273 (February): 200–209. https://doi.org/10.1016/j.ecolmodel.2013.11.016.

Swab, Rebecca M., Helen M. Regan, Diethart Matthies, Ute Becker, and Hans Henrik Bruun. 2015. 'The Role of Demography, Intra-Species Variation, and Species Distribution Models in Species' Projections under Climate Change'. *Ecography* 38 (3): 221–30. https://doi.org/10.1111/ecog.00585.

Thackeray, Stephen J., Peter A. Henrys, Deborah Hemming, James R. Bell, Marc S. Botham, Sarah Burthe, Pierre Helaouet, et al. 2016. 'Phenological Sensitivity to Climate across Taxa and Trophic Levels'. *Nature* 535 (7611): 241-245L. http://dx.doi.org.zorac.aub.aau.dk/10.1038/nature18608.

Voldoire, A., D. Saint-Martin, S. Sénési, B. Decharme, A. Alias, M. Chevallier, J. Colin, et al. 2019. 'Evaluation of CMIP6 DECK Experiments With CNRM-CM6-1'. *Journal of Advances in Modeling Earth Systems* 11 (7): 2177–2213. https://doi.org/10.1029/2019MS001683.

Winfree, Rachael. 2010. 'The Conservation and Restoration of Wild Bees'. *Annals of the New York Academy of Sciences* 1195 (1): 169–97. https://doi.org/10.1111/j.1749-6632.2010.05449.x.

Winfree, Rachael, Neal M. Williams, Jonathan Dushoff, and Claire Kremen. 2007. 'Native Bees Provide Insurance against Ongoing Honey Bee Losses'. *Ecology Letters* 10 (11): 1105–13. https://doi.org/10.1111/j.1461-0248.2007.01110.x.

Zhao, Hang, Guilin Li, Dezheng Guo, Han Li, Qingxin Liu, Baohua Xu, and Xingqi Guo. 2021. 'Response Mechanisms to Heat Stress in Bees'. *Apidologie* 52 (2): 388–99. https://doi.org/10.1007/s13592-020-00830-w.