ECOGRAPHY

Research

Evaluating predictive performance of statistical models explaining wild bee abundance in a mass-flowering crop

Maria Blasi, Ignasi Bartomeus, Riccardo Bommarco, Vesna Gagic, Michael Garratt, Andrea Holzschuh, David Kleijn, Sandra A. M. Lindström, Peter Olsson, Chiara Polce, Simon G. Potts, Maj Rundlöf, Jeroen Scheper, Henrik G. Smith, Ingolf Steffan-Dewenter and Yann Clough

M. Blasi (https://orcid.org/0000-0003-3252-5402) ✉ (Maria.blasi_romero@cec.lu.se), P. Olsson, H. G. Smith and Y. Clough, Lund Univ., Centre for Environmental and Climate Research, Lund, Sweden. – I. Bartomeus, Estación Biológica de Doñana (EBD-CSIC), Sevilla, Spain. – R. Bommarco, V. Gagic and S. A. M. Lindström, Swedish Univ. of Agricultural Sciences, Dept of Ecology, Uppsala, Sweden. – M. Garratt and S. G. Potts, Reading Univ., Centre for Agri-Environmental Research, School of Agriculture, Policy and Development, UK. – A. Holzschuh and I. Steffan-Dewenter, Univ. of Würzburg, Dept of Animal Ecology and Tropical Biology, Biocenter, Am Hubland, Würzburg, Germany. – D. Kleijn and J. Scheper, Wageningen Univ., Plant Ecology and Nature Conservation Group, Wageningen, the Netherlands. – SAML, M. Rundlöf and HGS, Lund Univ., Dept of Biology, Lund, Sweden. SAML, Swedish Rural Economy and Agricultural Society, Kristianstad, Sweden. – C. Polce, Joint Research Centre, European Commission.

Ecography
44: 525–536, 2021
doi: 10.1111/ecog.05308

Subject Editor: Kyle Haynes
Editor-in-Chief: Miguel Araújo
Accepted 15 December 2020

Wild bee populations are threatened by current agricultural practices in many parts of the world, which may put pollination services and crop yields at risk. Loss of pollination services can potentially be predicted by models that link bee abundances with landscape-scale land-use, but there is little knowledge on the degree to which these statistical models are transferable across time and space. This study assesses the transferability of models for wild bee abundance in a mass-flowering crop across space (from one region to another) and across time (from one year to another). The models used existing data on bumblebee and solitary bee abundance in winter oilseed rape fields, together with high-resolution land-use crop-cover and semi-natural habitats data, from studies conducted in five different regions located in four countries (Sweden, Germany, Netherlands and the UK), in three different years (2011, 2012, 2013). We developed a hierarchical model combining all studies and evaluated the transferability using cross-validation. We found that both the landscape-scale cover of mass-flowering crops and permanent semi-natural habitats, including grasslands and forests, are important drivers of wild bee abundance in all regions. However, while the negative effect of increasing mass-flowering crops on the density of the pollinators is consistent between studies, the direction of the effect of semi-natural habitat is variable between studies. The transferability of these statistical models is limited, especially across regions, but also across time. Our study demonstrates the limits of using statistical models in conjunction with widely available land-use crop-cover classes for extrapolating pollinator density across years and regions, likely in part because input variables such as cover of semi-natural habitats poorly capture variability in pollinator resources between regions and years.

Keywords: *Brassica napus*, mass flowering crops, model predictions, permanent semi-natural habitats, transferability in ecology, wild pollinators

NORDIC SOCIETY OIKOS

www.ecography.org

Introduction

Pollination by wild animals is a key ecosystem service that is highly important for 35% of the world's crops (Klein et al. 2007), and wild insects are especially important in supporting yields (Garibaldi et al. 2013, Rader et al. 2016, Dainese et al. 2019). However, pollination could be jeopardized since some pollinators (wild and domestic) are declining in some regions of the world (Potts et al. 2010, Dupont et al. 2011, Bartomeus et al. 2019, Powney et al. 2019) due to habitat loss and current intensive agricultural practices (Winfree et al. 2011). To support wild pollinators and reduce the risk for economic vulnerability induced by low pollination levels (Gallai et al. 2009), it is recommended to change the landscape-scale land use to ensure the availability of nesting and overwintering habitats and pollen and nectar resources supplied by wild and cultivated flowering plants (Carvell et al. 2006, Smith et al. 2014, IPBES 2016). Linking landscape-scale land-use to availability of resources for pollinators across space and time (Baude et al. 2016), or more often proxies such as cover of habitat assumed to be rich in resources for pollinators, is therefore a key aspect of applied ecological research on pollinators.

Ecological models have great potential for supporting environmental policy and decision-making (Addison et al. 2013, IPBES 2016). The application of models for decision-making often requires generating predictions for locations and time periods that are distinct from those for which the available empirical data was recorded (i.e. extrapolation). Despite the critical role of model transferability, i.e. how well models generalize to new contexts, is poorly studied.

Transferring models from one region to another (requiring spatial transferability), or from past to present or future scenarios (requiring temporal transferability), is particularly important when there is data deficiency, limited research funding and urgent need to forecast species' responses to changes in land use and/or climate. In ecology, transferability of models has started receiving significant attention in the application of species distribution models (SDM) (Barbosa et al. 2009, Wenger and Olden 2012, but see Yates et al. 2018). Knowing to what degree models are transferable to different contexts is not only a prerequisite to increase ecological understanding but helps to evaluate and communicate the uncertainty associated with the predictions (Houlahan et al. 2017). For mobile ecosystem service providers such as pollinators (Woodcock et al. 2019), the study of transferability is in its infancy. This is surprising given the increasing interest in society for taking action to support pollinating insects, and the increasing availability of models and suitable datasets.

Managing landscape configuration and composition to secure complementary flowering and nesting sources is key to support pollinating insects (Smith et al. 2014). Statistical models predicting pollinator abundance usually include the cover of different types of land use in the landscape surrounding insect-pollinated crops (Holzschuh et al. 2016, Dainese et al. 2019, Martin et al. 2019). Semi-natural habitats (SNH), including permanent grasslands and forests, provide stable food and nesting resources through time (Svensson et al. 2000, Öckinger and Smith 2007, Osborne et al. 2008, Knight et al. 2009). In contrast, arable crops are frequently disturbed by ploughing and pesticide use, although mass flowering crops (MFC) can provide very significant food resources for pollinators in short bursts during parts of the year (Holzschuh et al. 2013, Rundlöf et al. 2014). The cover of semi-natural habitat, usually represented only by grasslands, and mass-flowering crops are thus often used as explanatory variables in models investigating pollinator abundance (Holzschuh et al. 2016, Shaw et al. 2020).

Studies to inform statistical models predicting pollinator abundance are often carried out at multiple sites that vary in landscape composition, but the ability of the resulting models to predict abundance under novel spatial or temporal circumstances has seldom been investigated (but see De Palma et al. 2016).

When extrapolating findings from one region to another, the quality of the predictions might be reduced for several reasons. First, there are probably substantial differences in factors not explicitly considered in the model, such as agronomic conditions, local pollinator species or climate, differences in the phenology of mass-flowering crops, or the quality of a land-cover type in terms of resources for pollinators, such that their variation has a different impact on bees across studies. Secondly, the new conditions may be outside the range of the original conditions (including their combinations). These considerations apply not only when extrapolating to another region, but also when extrapolating to another year. For instance, inter-annual variability in the explanatory variables (e.g. land cover and the degree it is indicative of pollinator resources) or in the response variable can be influenced by weather differences, introducing further unexplained variability.

Here, we aim to address this knowledge gap by assessing the transferability of pollinator models across different regions and years for a well-defined, well-studied system that is relevant to pollinator management. We use existing data on bumblebee and solitary bee abundance in winter oilseed rape fields collected over several years in four different countries to address the following questions:

1. Does cover of mass-flowering crops and semi-natural habitats in the landscape show consistent effects on wild bee abundance in oilseed rape fields across space (region) and time (year)?
2. Based on predictive performance, to what extent are hierarchical models predicting wild bee abundance across space and/or time transferable, and how well can we predict wild bee abundance in one region and year using a model that was developed for other regions or years?

Methods

Data preparation

(a) Pollinators

The data used in the present study consist of wild bee observations in winter oilseed rape *Brassica napus* which included, where reported, the varieties Excalibur, DK Expower and Compass. Our dataset includes data from six datasets from five different European regions, from four different countries (Sweden – including two different regions, Scania (Skåne in Swedish) and Skara; Germany; the Netherlands; and the United Kingdom). The original datasets come from four different studies, meaning that three datasets collected data in the same systematic way. Data were collected during three different years (2011, 2012, 2013) (Fig. 1). Hereafter, we refer to the original datasets as each region–study–year, since several, but not all regions had more than one study–year combination per region. See the Supporting information for a description of the original datasets, including the original design, sampling methods and the data selected for this study.

From each raw dataset, we selected two groups of wild pollinators known to forage and pollinate in oilseed rape crops (Supplementary Table 2, from Kleijn et al. 2015), and we divided them between *Bombus* spp. and non-*Bombus* spp. wild bees, the latter referred hereafter as 'solitary bees'. The group of solitary bees includes bees that are solitary with few exceptions since it includes solitary bees such as *Andrena* spp., *Osmia* spp., *Lasioglossum* spp., *Eucera* spp. and eusocial bees such as *Halictus* spp. and *Lasioglossum* spp., as well as those bees categorized as 'wild bees' or 'solitary bees' from original datasets when the species level was not reported.

For statistical analysis, we pooled pollinator observations by summing up counts across transects and rounds, resulting in one value per field. Since the original datasets differed in the number of rounds and transects, we standardized by selecting two transects and two rounds in all studies during the highest flower abundance. We standardized the

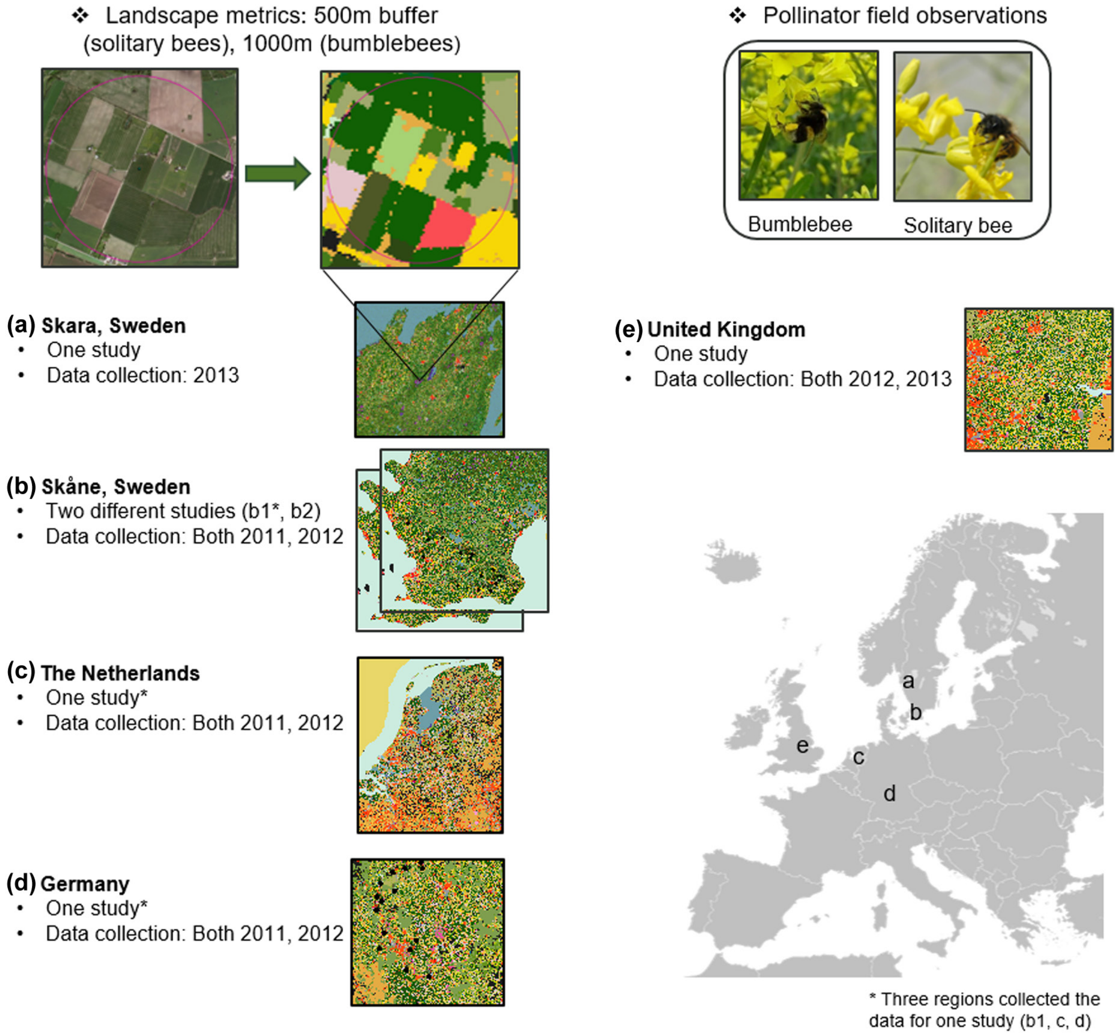

Figure 1. Overview of the dataset for the current study. Study species, bumblebees and solitary bees; and high-resolution land use maps used to calculate the proportion of land-use variables around the focal field of pollinator observations in winter oilseed rape *Brassica napus*. Our dataset includes six datasets from five different European regions, from four different countries: Sweden – including two different regions, Skara (a) and Scania (Skåne in Swedish) (b); the Netherlands (c); Germany (d); and the United Kingdom (e). Pollinator observations data from Skåne was from two separate studies, and three datasets representing three different countries collected data for the same study. Data were collected during three different years (2011, 2012, 2013).

response variable to the same units of pollinator density in the field, corresponding to the number of bees per meter per 15 min. This resulted in non-integer response variables, and we therefore used models assuming a normal distribution. To achieve normality and homoscedasticity of the residuals (Zuur et al. 2010), all observations were Log_e transformed prior to analysis.

(b) Land use maps including detailed vegetation classes

High-resolution (25 × 25 m pixel size) land use rasters from Germany, Netherlands and Sweden, were created using Integrated Administration and Control System (IACS) data. These data contain both reference and agricultural parcels, where a reference parcel could contain one or more agricultural parcels (Sagris et al. 2015). As IACS data only cover arable land, other land cover data were used to fill the gaps between the IACS data. CORINE land cover data (CLC 2012_Version 18.4) were used to fill gaps in Germany and the Netherlands. Gaps in Sweden were filled using more detailed land use information (Svenska Marktäckedata (SMD), Naturvårdsverket 2014). Land use codes were harmonized across datasets to 213 land use categories. The IACS data was retrieved from partners of the EU LIBERATION project, see further Supporting information for a description of the Geographical Information System methods used, and information on which land-use classes were used in the statistical modeling.

We focused on two land-use cover variables: percentage of winter oilseed rape and percentage of semi-natural habitats, which included forest and permanent, non-intensively managed grasslands (Supporting information). We extracted those land-use metrics in a 500 m buffer for solitary bees and a 1000 m buffer for bumblebees, both around the centroid of each field, based on their foraging distances (Osborne et al. 2008, Zurbuchen et al. 2010). Analyses were limited to the datasets for which we had land cover maps for the buffers, providing nine datasets (combinations of region–study–year) for bumblebees and ten for solitary bees. UK bumblebees were thus not included in the analysis, as the UK land use information only allowed a 500 m buffer (see Supporting information), and Sweden – Skara was removed from the analysis for solitary bees since the fields from that study contained much lower variation in oilseed rape cover within the 500 m buffers (ranging from 0.6 to 3.3%), compared to other regions (0.2–35%), to allow comparisons with the rest of the groups. To avoid study-specific approaches to land-use crop-cover classification, such as different interpretations of the term ‘grassland’, which may or may not include temporary grassland, we used a common methodology to create the spatial data and compute the landscape-scale land-use variables. Here, we refer to grassland as a permanent semi-natural habitat including the categories of pasture, mown meadow, forest pasture, mosaic pasture, natural grassland, grassland, meadow, semi-natural area), but not leys. Leys were included in a separate category (among the 213 categories), which did not contain mass flowering crops nor grasslands.

Data analysis

Statistical analyses were performed using R ver. 3.5.3 (<www.r-project.org>). We developed a model that combines all studies and we evaluated the transferability of a given model using a cross-validation approach.

Selection of model variables: habitats valuable for food and nesting resources

Oilseed rape is the crop in which the pollinator data were collected, and the most abundant early-flowering mass-flowering crop (MFC) in most of the regions. This is why we chose ‘oilseed rape’ instead of MFC in our models (see Supporting information for a correlation matrix between all habitats for each region, 500 m and 1000 m buffer). The variable cover of semi-natural habitats was created from summing the cover of grassland and forest, as a representation of permanent natural habitats in the region. Collinearity of variables in the model were checked using the variance inflation factor (VIF). All models presented VIF < 0.17, meaning no substantial collinearity between the variables (Dormann et al. 2013).

Hierarchical model: a single model for all regions, studies and years

To analyze the relationship between wild bee abundance and land-use types we built a single hierarchical model that combines all studies. We used the function *lmer* in R, from the package lme4 (Bates et al. 2015), to run a linear mixed effect model. We included oilseed rape and semi-natural habitat cover as fixed effects without interaction. We included random intercepts for a grouping variable with levels corresponding to the individual combination of country, study and year. The full model also included random slopes for both oilseed rape and semi-natural habitat, and it was used to assess the importance of spatial and temporal variation in responses across the studies. To facilitate model convergence and allow meaningful comparisons among variables with different ranges, we standardized all landscape variables to zero mean and unit variance, across all studies, before analysis.

The full model was coded as follows (once for bumblebees and once for solitary bees):

lmer ($\textit{log}_e$*(pollinator abundance) ~ oilseed rape + semi-natural habitat, (oilseed rape + semi-natural habitat | group))* where group corresponds to the combination of region, study and year.

To test the consistency in the effect of the land-use variables across the groups, we compared models with and without random slope. We tested four different models including only intercept (no random slopes), random slopes with only oilseed rape, random slopes with only semi-natural habitat, or random slopes with a combination of both variables (Supporting information). All models had oilseed rape and semi-natural habitat cover as fixed effects. We used the anova function to provide model comparison statistics for all these models using restricted maximum likelihood (method = ‘REML’) (Zuur et al. 2009). We examined the competing models within $\Delta\text{AIC} \leq 2$ (Burnham and

Anderson 2002), and we chose the simplest model after checking that the added complexity of the complex model did not add considerable explanatory power compared to the simpler model. See Supporting information for the AIC values from the anova test results. We interpreted the results from the selected model for each bee group.

Transferability of the hierarchical model

We assess the spatial and temporal transferability of the model that was selected in the hierarchical model: a single model for all regions, studies and years (Supporting information) by comparing the performance under different types of non-random cross-validation.

Cross-validation is a method that involves dividing data into training evaluation blocks (Roberts et al. 2017). Non-random cross-validation means that any validation set differs from those used for training the model in the same way as an independent dataset (Wenger and Olden 2012).

We used three non-random approaches to partition the data: 1) fully stratified, or 'balanced', where all groups (all-region and year combinations) are in both the training and evaluation set; 2) partitioning the dataset to assess temporal extrapolation (for the same region, testing from one year to another), with data from different years within the same region assigned to different partitions; and 3) partitioning the dataset to assess spatial extrapolation (given the same year, testing the prediction from one region to another) with data from different region–study combinations assigned to different partitions. The transferability analysis includes only those datasets with data from both 2011 and 2012 for both pollinator groups. This gave a total of four regions from three countries, and two years in each region, resulting in eight groups (region–study–year combinations). The partitioning was done using the R package Groupdata2 (Olsen 2017). Constrained by the partitioning for the temporal extrapolation, where we had a maximum of two years within a study, we used half of the data for training and half for evaluation.

We repeated the partitioning process 100 times for a fully stratified cross-validation, and 16 and 6 times for the temporal and regional extrapolation, respectively (the maximum number of possible combinations). The grouping variable for both fully stratified and regional extrapolation was a factor consisting of the region–study–year levels, while for temporal extrapolation the grouping variable was a factor consisting of each combination of region–study. When predicting using the information from both fixed and random effects we thus simulate prediction within the same statistical population in the case of full stratification, prediction for new regions–year combinations in the case of regional extrapolation, and prediction for new years (but within the same region) for temporal extrapolation. We consider this most representative of how ecological models are likely to be used for extrapolation in practice.

Both models for solitary bees and bumblebees included oilseed rape and semi-natural habitat as fixed effects, but differed in their random structure: only random intercepts for bumblebees, and random intercept and slope for semi-natural habitats for solitary bees. We evaluated both overall and group-specific response, meaning that group-specific response takes into account the random effects, while the overall response does not, and therefore we can compare results to contrast the contribution of fixed and random effects to the predictive performance.

To evaluate the model fit to the new data (e.g. the agreement between predictions of the training dataset with data from the evaluation dataset), we used three complementary quantitative goodness-of-fit measures: the mean absolute error (MAE), the root mean squared error (RMSE) and the coefficients of determination R^2. MAE and RMSE both describe the average model-performance error. While RMSE presents the advantage of a higher weighting of larger discrepancies, MAE is considered to be easier to interpret since it is on the same scale as the response variable and is the least ambiguous measure of average error (Willmott and Matsuura 2005). The coefficients of determination R^2 indicates the proportion of variability explained by the model. We used unadjusted R^2 since the number of model parameters remains constant. We used the R package hydroGOF (Zambrano-Bigiarini 2017) to obtain the goodness-of-fit measures. For a better interpretation of the results, data were back-transformed after the analysis. This was done both for the overall response and the group-specific response.

Results

Datasets included a total of 841 bumblebee observations from 101 fields in 9 region–study–year combinations; and a total of 534 solitary bees from 87 field observations, in 10 study–region–year combinations.

To assess whether the cover of mass-flowering crops and semi-natural habitats in the landscape show consistent effects on wild bee abundance across space and time, we analyzed the relationship between wild bees and the explanatory variables oilseed rape and semi-natural habitat by building a hierarchical model that contains all region–study–year combinations, and obtaining a parsimonious random effects structure (the results of model selection for support of random slopes (Supporting information).

For bumblebee abundance, we had two models within our model selection criteria of $\Delta\text{AIC} \leq 2$, one model including random slopes for semi-natural habitat ($\text{AIC}_{\text{SNH slope}} = 131.29$) and a model with only random intercepts ($\text{AIC}_{\text{random intercepts}} = 132.74$) (Supporting information). Both models had very similar estimates for the fixed effects, and the variance of the additional random term in the more complex model was close to zero (see Supporting information for model results), meaning that any of the two models could be used for the purpose of this paper. We selected the simplest model, the model including only random intercepts.

Bumblebee abundance decreased with an increasing cover of oilseed rape crops in the surrounding landscape within 1 km radius (Table 1, Fig. 2a, 3a), an effect that was constant across all region–study–year groups. The abundance of bumblebees increased with the cover of semi-natural habitats, but this was not consistent across all region–study–year groups (Table 1).

For solitary bee abundance, we chose the model with the lowest AIC value, which was the model including only semi-natural habitats as a random slope ($AIC_{SNH\ slope} = 114.01$). No other models had $\Delta AIC \leq 2$ (Supporting information). Solitary bee abundance decreased with the cover of oilseed rape, and increased with the surrounding semi-natural habitat cover, although these effects were not significant (Table 1, Fig. 2c–d, 3c–d).

We assessed the importance of spatial and temporal variation in responses across the studies with the full model, and we found that there was a higher interannual and spatial variation in the effect of semi-natural habitat than in the effect of oilseed rape ($\delta^2_{random\ slope\ SNH} = 0.041$; $\delta^2_{random\ slope\ OSR} = 0.002$) for bumblebees (Fig. 2a–b). A different pattern was found for solitary bees, where the interannual and spatial variation of the responses was lower for semi-natural habitats (Fig. 2d), but the variability in the effect of oilseed rape was higher than for bumblebees ($\delta^2_{slope\ SNH} = 0.018$; $\delta^2_{slope\ OSR} = 0.017$) (Fig. 2c).

Based on the predictive performance, we tested to what extent our model predicting wild bee abundance across space and/or time is transferable. We used a cross-validation method approach to study how well our model predicts. The baseline to compare our cross-validation results is predictive performance under the balanced stratification, which corresponds to prediction within region–study–year combinations without further extrapolation. The results are shown in Fig. 4.

The predicted abundances in the balanced stratification cross-validation were closest to the observed abundances (mean $MAE_{balanced\ stratification}$ bumblebees = 0.36; mean $MAE_{balanced\ stratification}$ solitary bees = 0.38). Although temporal and spatial extrapolation showed a higher mean error and a higher variance, these values were not too far from the balanced stratification (mean $MAE_{temporal\ extrapolation}$ bumblebees = 0.40; mean $MAE_{temporal\ extrapolation}$ solitary bees = 0.41; mean $MAE_{spatial\ extrapolation}$ bumblebees = 0.42 and mean $MAE_{spatial\ extrapolation}$ solitary bees = 0.41) (see Supporting information for values of RMSE, MAE and R^2).

The predicted abundances were 20% off the mean observed in the balanced stratification, with percentages ranging from 16 to 28% between the studies for bumblebees, and between 22 and 39% for solitary bees. For bumblebees, balanced stratification was followed by temporal extrapolation with an average of 24% off the mean, ranging between 19 and 31%, and finally by a not too far spatial extrapolation with an average of 25% off the mean, ranging between 17 and 38%. For solitary bees, temporal and spatial extrapolation predictions shared the same value of 33% off the mean observed, with percentages ranging from 24 to 40% (temporal) and 22 to 45% (spatial) (Supporting information).

Mean R^2 values were highest for balanced stratification, especially so for the group-specific response (0.43 for bumblebees, and 0.26 for solitary bees), compared to the overall population-level ('fixed-effect only') response (0.25 for bumblebees, and 0.11 for solitary bees). These values were followed by those of the overall temporal extrapolation (0.23 for bumblebees, and 0.07 for solitary bees), and the spatial extrapolation for the overall and group-specific response (0.19 for bumblebees, and 0.05 for solitary bees). The group-specific temporal extrapolation values were lower than the spatial extrapolation values for bumblebees (0.03), but for solitary bees, temporal extrapolation values were very similar those obtained with spatial extrapolation (0.08). The R^2 values were as expected usually higher for the group-specific than for the overall predictions, with one exception. In temporal extrapolation for bumblebees where values where higher for the overall (0.23) than for the group specific (0.03) predictions. This can be observed when predicting out-of-sample and suggests that in this case the random effects lead to overfitting.

Discussion

We assessed the contribution of land-use cover variables in predicting wild solitary bee and bumblebee abundance in a mass-flowering crop and the transferability of a statistical model predicting these effects. We found a negative relationship between the cover of oilseed rape and bee abundance for both bumblebees and solitary bees, a relationship that was more consistent across space and time for bumblebees as shown in our mixed-effects models. The effect of semi-natural habitat was variable across space and time, especially for solitary bees.

Oilseed rape is a rotational crop and its landscape-scale cover fluctuates from year to year. A potential mechanism

Table 1. Results of the linear mixed effect models testing the effect of oilseed rape cover and semi-natural habitat cover (grassland and forest) on the abundance of bumblebees and solitary bees sampled in oilseed rape fields. Only selected models are shown. Number of field observations: 841 number of bumblebees (101 fields) and 534 number of solitary bees (87 fields). Number of groups (each country–study–year-combination): 9 (for bumblebees), 10 (for solitary bees). Model estimates, standard error (SE), degrees of freedom (df) and p values are reported. p-values < 0.05 are shown in bold.

Response variable	Random structure	Land use variable	Estimate	SE	df	p-value
Bumblebee abundance	(1|group)	(Intercept)	0.88	0.13	7.30	**< 0.001**
		% Oilseed rape	−0.14	0.04	92.12	**0.003**
		% Semi-natural habitat	0.16	0.05	97.69	**0.002**
Solitary bee abundance	(1 + % Semi-natural habitat |group)	(Intercept)	0.70	0.09	8.52	**< 0.001**
		% Oilseed rape	−0.05	0.04	82.84	0.216
		% Semi-natural habitat	0.05	0.06	13.15	0.441

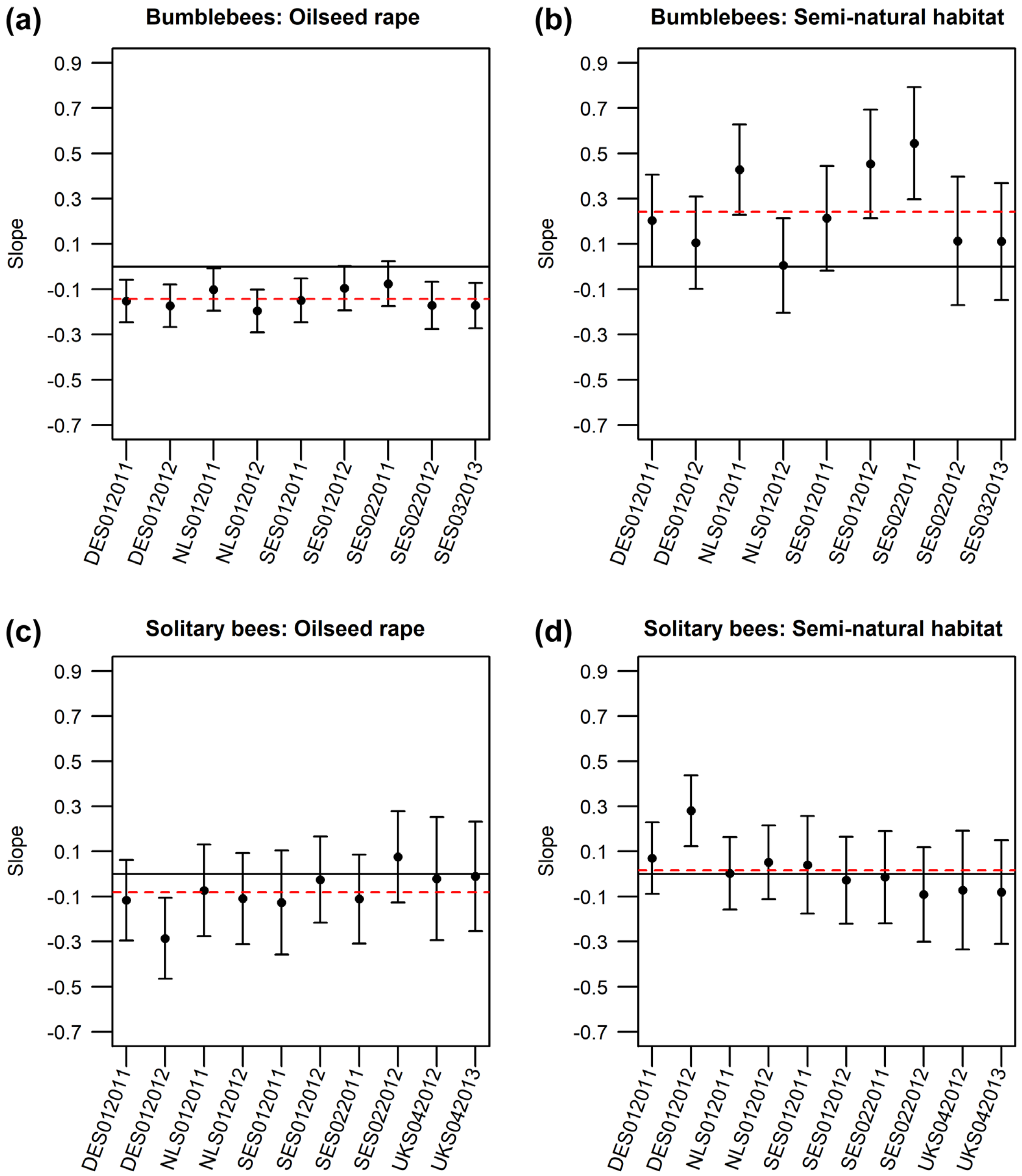

Figure 2. Slope responses from each combination of region–study for bumblebee abundance (a–b) and solitary bee abundance (c–d). We tested the effect of the surrounding oilseed rape habitat and permanent semi-natural habitats (grassland and forest), on bumblebees and solitary bees sampled in oilseed rape fields. The dashed red line is the general trend across the studies. X axis shows the different studies; and Y axis shows the standardized slope value estimated using mixed effects models.

underlying the negative effect of this mass-flowering crop on wild bee pollinators is the dilution of bees in the landscape (Riedinger et al. 2015, Holzschuh et al. 2016, Shaw et al. 2020), as a result of local pollinator numbers reacting to a large area of floral resources in the landscape. Because all oilseed rape fields within a region will flower more or less at the same time, the dilution processes due to differences in oilseed rape cover will act directly upon the local density of foraging wild bees. Since the pollinator surveys took place at the moment of maximum flowering, this likely resulted in the shared similar effect of the cover of this crop between regions. The effect was stronger in bumblebees than solitary bees, a result in line with Holzschuh et al. (2016). The longer foraging ranges of the bumblebees, which makes them better able to respond to changes in the density of food resources, may explain this, even though we tried to account for that by choosing different landscape scales for bumblebees and solitary bees. Another possible explanation could be that bees are negatively affected by pesticide use in the mass flowering crop, especially since oilseed rape is a crop where pesticide use is often high (Williams 2010, Zhang et al. 2017), which could negatively affect bee reproduction (Woodcock et al. 2017).

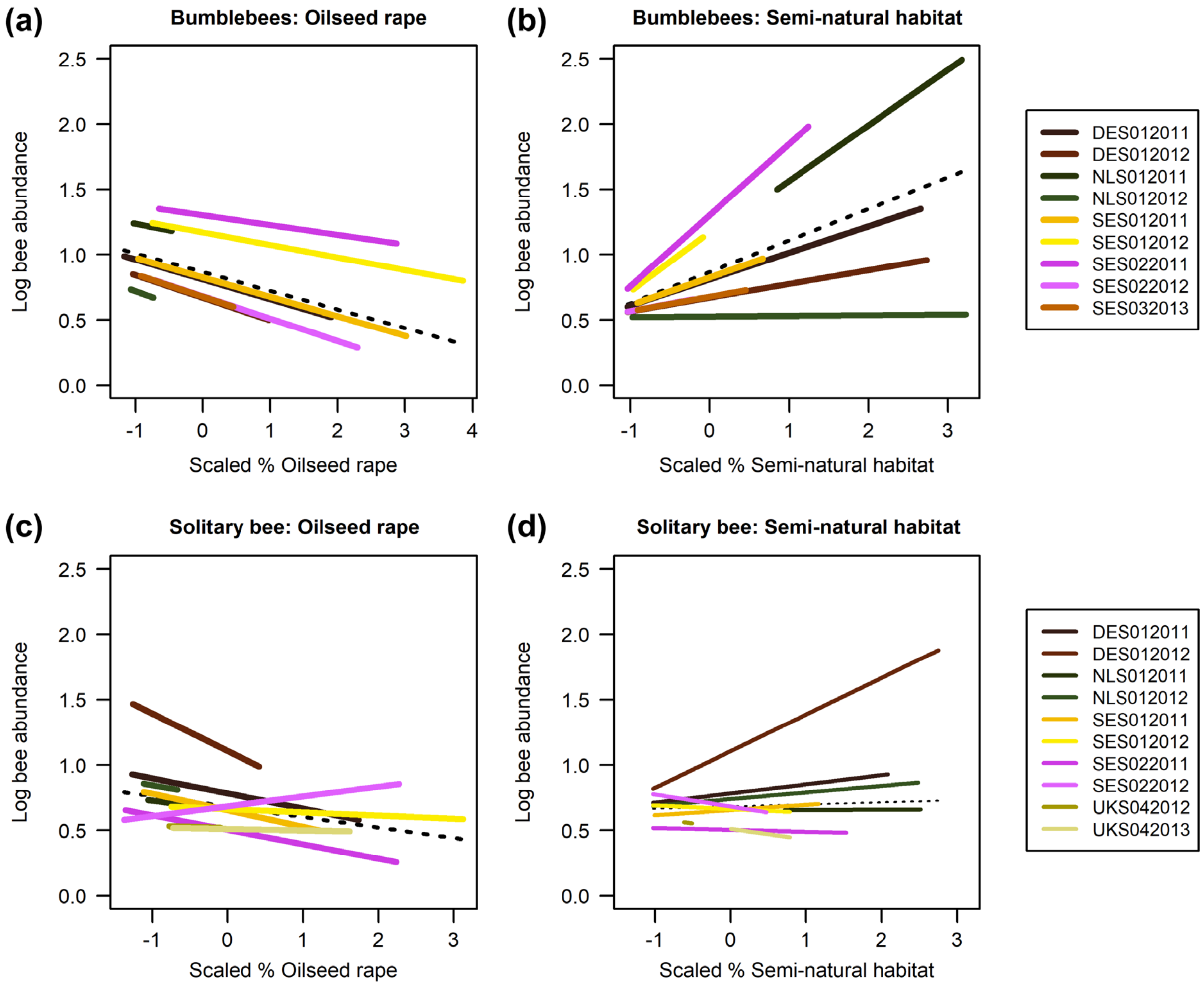

Figure 3. (a) Relationships between the standardized cover of oilseed rape and (b) semi-natural habitat (grassland and forest) (% in 1 km radius) and abundance of bumblebees. (c) Relationships between the cover of oilseed rape and (d) semi-natural habitat (% in 500 m radius) and abundance of solitary bees. The color indicates the different studies. The fitted lines are linear mixed model estimates for each region and year. The dashed line is the overall trend.

Semi-natural habitats, on the other hand, are less exposed to insecticides and other forms of disturbance (Senapathi et al. 2017). We found positive, significant relationships between wild bee abundance and semi-natural habitat for bumblebees but not for solitary bees. For solitary bees, the effect of semi-natural habitats was not significant, possibly explained by the variability among region–study–year, as shown by the random slope. The lack of consistency of semi-natural habitats across regions and years (Supporting information) could be explained by differences in the quality and quantity of semi-natural habitats. Although we included grasslands and woodlands or forests since they provide key resources for bumblebees and solitary bees (Svensson, et al. 2000, Persson et al. 2018, Donkersley 2019), the effect of heterogeneity in forest and grassland management and composition on pollinators is less well understood. The effect of the amount and quality of semi-natural habitats at the landscape-scale is highly affected by the wider landscape configuration and composition (Scheper et al. 2013, Martin et al. 2019, Sirami et al. 2019), which can lead to further inconsistency among regions and years.

Differences between bumblebees and solitary bees in response to the cover of semi-natural habitat were also shown in previous research (Shaw et al. 2020). Contrary to our results, Steffan-Dewenter et al. (2002) found that numbers of solitary bees, but not bumblebees increased with increasing semi-natural grasslands in the landscape. This was presumably due to their requirements for nesting sites that are more restricted to these semi-natural habitats (Steffan-Dewenter et al. 2002). Possibly, the higher species richness and differences in species composition between solitary bee communities in our study (Supporting information) may increase the heterogeneity of the response of this group to semi-natural habitats. The differences in species composition between the different studies, both between region and years, may be an important contributing factor to the heterogeneity of the responses to landscape-scale land-use. Better species-level coverage, and richer data, would allow running

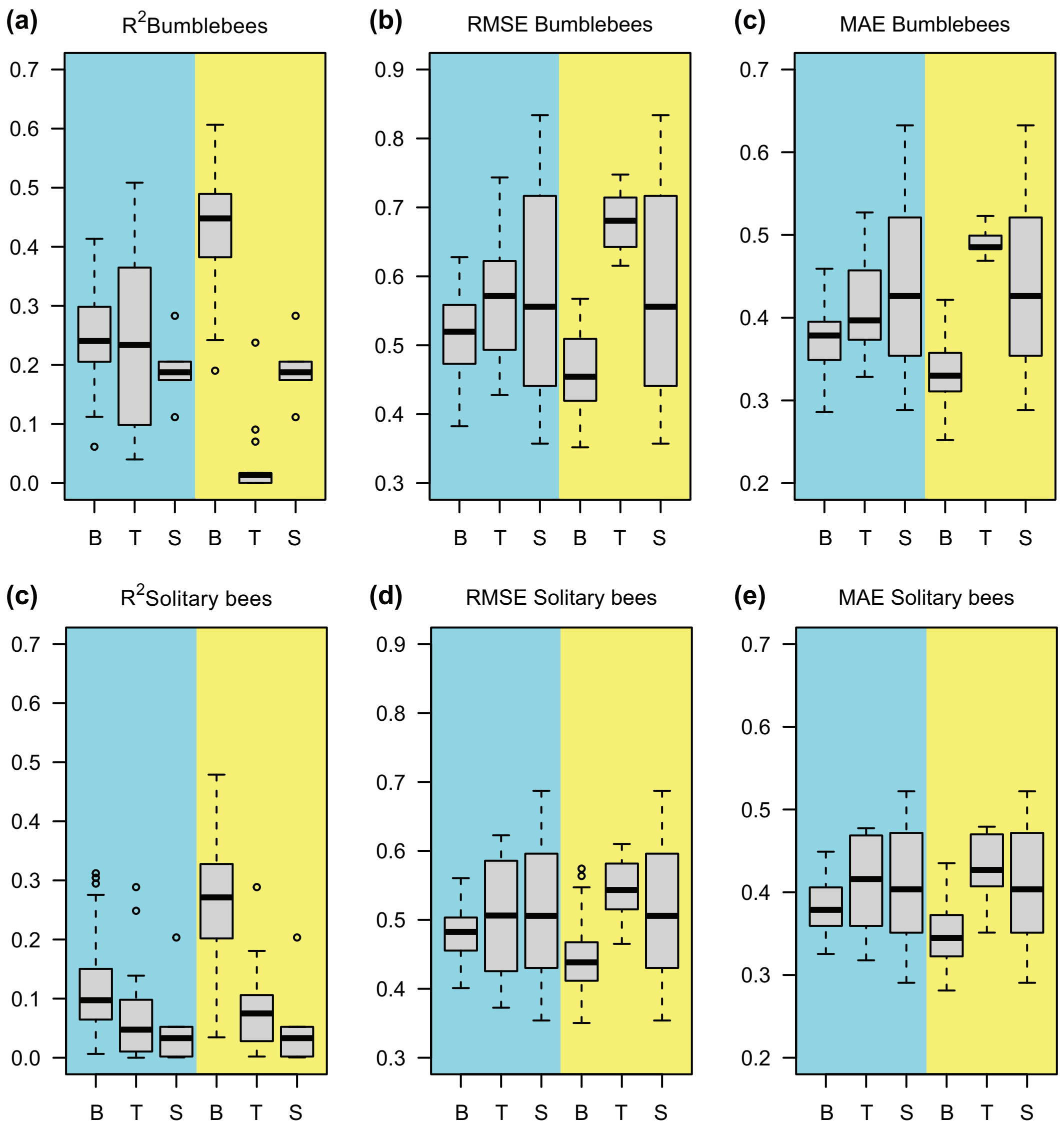

Figure 4. Cross-validation results for bumblebees (a–c) and solitary bee (d–f) abundance. Graphical representation of the goodness-of-fit measures RMSE (root mean squared error), MAE (mean absolute error) and R^2 (coefficients of determination), are shown for the overall response (blue) overall and group-specific response (yellow) of the three different cross-validation grouped data balanced stratification (B), temporal extrapolation (T) and regional extrapolation (S). The difference between overall and group-specific response is that group-specific response takes into account the random effects, while the overall response does not, and therefore we can compare results to contrast the contribution of fixed and random effects to the predictive performance. Supporting information for the mean values for the goodness-of-fit measures, and Supporting information for the interpretation of MAE.

multi-species models to capture some of this heterogeneity and increase transferability.

Using different cross-validation approaches to assess transferability across regions and years we demonstrate that transferability in space and time of these statistical models can be limited. Our study shows that one should expect considerable uncertainty when extrapolating results of models not only in space (in relatively close neighbor countries), but also in time (to predict from one single year to another). Model predictive power was low in general, likely due to the noisiness of ecological data and the simplicity of the model. When looking at the overall response among the groups, predictions for bumblebees were slightly better when predicting across years within a region, rather than across the regions within a year. This would be expected given that we can assume higher differences between regions than years in the quality of semi-natural habitats, as suggested by our analyses, but also other aspects that could not be analyzed here, such as species composition.

Whether the extrapolated predictions are useful is context-specific. For example, our 23% error prediction for temporal extrapolation might be too high for advising a particular farmer, but enough to develop a national policy. In this study, we want to point out that while the transferability of knowledge based on statistical models is very common, the transferability of statistical models is barely assessed and quantified (which is the aim of our paper). In the experience of the authors, extrapolating statistical models to new sites and years is increasingly on the table as stakeholders and policy makers request maps of populations and ecosystem services such as pollination, given land-use and/or climate inputs.

A conclusion of the transferability analysis is not only that more data is needed to generate better predictions, but more specifically that collecting more data from the regions where the models are to be applied is likely to be more useful – in particular for solitary bees.

The data used in this study, although limited, is one of the largest crop-specific pollinator datasets used for model training to date. More refined input data on the quality of the different habitats would allow building more complex models that could capture a higher share of the variance. Availability and consistency of such data are currently limiting, however, but promising approaches exist (Carrié et al. 2018).

For future studies, it is worthwhile investigating whether improved transferability can be achieved by stratifying by latitude, biogeographic region or similar amounts and types of semi-natural habitats or agricultural management, but that would require more datasets, e.g. through more systematic pollinator monitoring schemes (Carvell et al. 2017, Bartomeus and Dicks 2019, Garratt et al. 2019). Another important question is the extent to which mechanistic models that predict bumblebee or solitary bee abundance based on underlying ecological processes, can overcome the lack of transferability that we see in statistical models. Mechanistic models can capture details of the multi-scale dynamics of the system since these models predict increased pollinator abundance over time in a landscape based on additional flower resources, and the dilution/concentration effects of changes in the quantity of jointly flowering resource patches. It is worthwhile noting that the latter effect was the most consistent over space and time, but has for a long time not been included when modeling pollination potential (e.g. in the InVEST pollination model (Lonsdorf et al. 2009)). Newer models take this into account and show that this significantly improves model predictions (Nicholson et al. 2019).

Conclusions

Predictability of wild bee densities and therefore pollination service in a mass-flowering crop in response to landscape-scale land-use is limited, at least given the size of typical ecological studies as used in our analysis. We find that training models with data from the region for which predictions are required are recommended to improve model predictive performance. Our results show that including the dilution effect of simultaneously flowering crops on pollinator densities in the crops is both important for predicting pollinator visitation, and constantly so over regions and years. At present, some of the most widely used models for pollinator mapping do not take this into account (e.g. invest pollination model, Lonsdorf et al. 2009), but promising alternatives exist (Olsson et al. 2015, Haussler et al. 2017). Finally, the ability to robustly predict pollinator visitation to crops across regions and years should be more systematically used to assess both statistical and mechanistic models.

Data availability statement

Data available from the Dryad Digital Repository: <http://dx.doi.org/10.5061/dryad.qrfj6q5c1> (Blasi et al. 2020).

Acknowledgements – We thank V. Riedinger for providing data for the study. We also thank S. Redlich, M. Stjernman and T. Breeze for providing IACS data to develop the raster maps.
Funding – This work was funded by the 2013–2014 BiodivERsA/FACCE-JPI joint call for research proposals (project ECODEAL), with the national funders ANR, BMBF, FORMAS (contract 2014-1783), FWF, MINECO, NWO and PT-DL; Kungliga Fysiografiska Sällskapet i Lund, for Endowments for the Natural Sciences, Medicine and Technology – Biology; the European Community's Seventh Framework Programme under grant agreements no. 244090 for the STEP: Status and trends of European pollinators project and no. 311781, for the LIBERATION Project; the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 819374) for the DrivenByPollinators project; the Insect Pollinators Initiative funded jointly by a grant from British Biological Science Research Council, Department for the Environment Farming and Rural Affairs, Natural Environment Research Council, the Scottish Government and the Wellcome Trust, under the Living with Environmental Change Partnership (BB/I000275/1); and the 2017–2018 Belmont Forum and BiodivERsA joint call for research proposals, under the BiodivScen ERA-Net COFUND program, and with the funding organizations AEI, NWO, ECCyT and NSF.
Conflicts of interest – The authors declare no conflict of interest.
Disclaimer – Views expressed are those of the individuals and do not necessarily represent official views of the European Commission.

Author contributions

Maria Blasi Romero: Conceptualization (equal); Data curation (lead); Formal analysis (lead); Funding acquisition (supporting); Methodology (lead); Project administration (lead); Resources (supporting); Software (equal); Validation (lead); Visualization (lead); Writing – original draft (lead); Writing – review and editing (lead). **Ignasi Bartomeus**: Formal analysis (supporting); Funding acquisition (supporting); Investigation (equal); Methodology (supporting); Writing – review and editing (supporting). **Riccardo Bommarco**: Funding acquisition (supporting); Investigation (equal). **Vesna Gagic**: Formal analysis (supporting); Investigation (equal); Methodology (supporting); Writing – review and editing (supporting). **Michael Garrat**: Funding acquisition (supporting); Investigation (equal); Writing – review and editing (supporting). **Andrea Holzschuh**: Funding acquisition (supporting); Investigation (equal); Writing – review and editing (supporting). **David Kleijn**: Funding acquisition (supporting); Investigation (equal); Methodology (supporting); Writing – review and editing (supporting). **Sandra Lindström**: Investigation (equal); Writing – review and editing (supporting). **Peter Olsson**: Data curation (equal); Investigation (equal); Methodology (equal); Resources (equal); Writing – original draft (supporting); Writing – review and editing (supporting). **Chiara Polce**: Investigation (equal); Writing – review and editing (supporting). **Simon Potts**: Funding acquisition (supporting); Investigation (equal); Writing – review and editing (supporting). **Maj Rundlöf**: Funding acquisition (supporting); Investigation (equal); Writing – review and editing (supporting). **Jeroen Scheper**: Funding acquisition (supporting); Investigation (equal); Methodology (supporting); Writing – review and editing (supporting). **Henrik Smith**: Funding acquisition (supporting); Investigation (equal); Methodology (supporting); Writing – review and editing (supporting). **Ingolf Steffan-Dewenter**: Funding acquisition (supporting); Investigation (equal); Writing – review and editing (supporting). **Yann Clough**: Conceptualization (equal); Data curation (equal); Formal analysis (lead); Funding acquisition (lead); Investigation (equal); Methodology (lead); Resources (lead); Software (equal); Visualization (equal); Writing – original draft (supporting); Writing – review and editing (equal).

References

Addison, P. F. et al. 2013. Practical solutions for making models indispensable in conservation decision-making. – Divers. Distrib. 19: 490–502.

Baude, M. et al. 2016. Historical nectar assessment reveals the fall and rise of floral resources in Britain. – Nature 530: 85–88.

Barbosa, A. M. et al. 2009. Transferability of environmental favourability models in geographic space: the case of the Iberian desman *Galemys pyrenaicus* in Portugal and Spain. – Ecol. Model. 220: 747–754.

Bartomeus, I. and Dicks, L. V. 2019. The need for coordinated transdisciplinary research infrastructures for pollinator conservation and crop pollination resilience. – Environ. Res. Lett. 14: 045017.

Bartomeus, I. et al. 2019. Historical collections as a tool for assessing the global pollination crisis. – Phil. Trans. R. Soc. B 374: 20170389.

Bates, D. et al. 2015. Fitting linear mixed-effects models using lme4. – J. Stat. Softw. 67: 1–48.

Blasi, M. et al. 2021. Data from: Evaluating predictive performance of statistical models explaining wild bee abundance in a mass-flowering crop. – Dryad Digital Repository, <http://dx.doi.org/10.5061/dryad.qrfj6q5c1>.

Burnham, K. P. and Anderson, D. R. 2002. Model selection and multimodel inference. – Springer.

–Carrié, R. et al. 2018. Bee diversity in crop fields is influenced by remotely-sensed nesting resources in surrounding permanent grasslands. – Ecol. Indic. 90: 606–614.

Carvell, C. et al. 2006. Declines in forage availability for bumblebees at a national scale. – Biol. Conserv. 132: 481–489.

Carvell, C. et al. 2017. Design and testing of a national pollinator and pollination monitoring framework. Final summary report to the Dept for Environment, Food and Rural Affairs (Defra), Scottish Government and Welsh Government: Project WC1101.

Dainese, M. et al. 2019. A global synthesis reveals biodiversity-mediated benefits for crop production. – Sci. Adv. 5: eaax0121.

De Palma, A. et al. 2016. Predicting bee community responses to land-use changes: effects of geographic and taxonomic biases. – Sci. Rep. 6: 31153.

Donkersley, P. 2019. Trees for bees. – Agric. Ecosyst. Environ. 270: 79–83.

Dormann, C. F. et al. 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. – Ecography 36: 27–46.

Dupont, Y. L. et al. 2011. Quantitative historical change in bumblebee (*Bombus* spp.) assemblages of red clover fields. – PLoS One 6: e25172.

Gallai, N. et al. 2009. Economic valuation of the vulnerability of world agriculture confronted with pollinator decline. – Ecol. Econ. 68: 810–821.

Garibaldi, L. A. et al. 2013. Wild pollinators enhance fruit set of crops regardless of honey bee abundance. – Science 339: 1608–1611.

Garratt, M. et al. 2019. Capacity and willingness of farmers and citizen scientists to monitor crop pollinators and pollination services. – Global Ecol. Conserv. 20: e00781.

Haussler, J. et al. 2017. Pollinator population size and pollination ecosystem service responses to enhancing floral and nesting resources. – Ecol. Evol. 7: 1898–1908.

Holzschuh, A. et al. 2013. Mass-flowering crops enhance wild bee abundance. – Oecologia 172: 477–484.

Holzschuh, A. et al. 2016. Mass-flowering crops dilute pollinator abundance in agricultural landscapes across Europe. – Ecol. Lett. 19: 1228–1236.

Houlahan, J. E. et al. 2017. The priority of prediction in ecological understanding. – Oikos 126: 1–7.

IPBES 2016. The assessment report of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services on pollinators, pollination and food production. – In: Potts, S. G. et al. (eds), Secretariat of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services, Bonn, Germany, 552 p.

Kleijn, D. et al. 2015. Delivery of crop pollination services is an insufficient argument for wild pollinator conservation. – Nat. Comm. 6: 7414.

Klein, A.-M. et al. 2007. Importance of pollinators in changing landscapes for world crops. – Proc. R. Soc. B 274: 303–313.

Knight, M. E. et al. 2009. Bumblebee nest density and the scale of available forage in arable landscapes. – Insect Conserv. Divers. 2: 116–124.

Lonsdorf, E. et al. 2009. Modelling pollination services across agricultural landscapes. – Ann. Bot. 103: 1589–1600.

Martin, E. A. et al. 2019. The interplay of landscape composition and configuration: new pathways to manage functional biodiversity and agroecosystem services across Europe. – Ecol. Lett. 22: 1083–1094.

Naturvårdsverket. 2014. Svenska Marktäckedata. – Naturvårdsverket.

Nicholson, C. C. et al. 2019. Flowering resources distract pollinators from crops: model predictions from landscape simulations. – J. Appl. Ecol. 56: 618–628.

Öckinger, E. and Smith, H. G. 2007. Semi-natural grasslands as population sources for pollinating insects in agricultural landscapes. – J. Appl. Ecol. 44: 50–59.

Olsen, L. R. 2017. groupdata2: creating groups from data. R package ver. 1.0.0. – <https://cran.r-project.org/web/packages/groupdata2/index.html>.

Olsson, O. et al. 2015. Modeling pollinating bee visitation rates in heterogeneous landscapes from foraging theory. – Ecol. Model. 316: 133–143.

Osborne, J. L. et al. 2008. Bumblebee flight distances in relation to the forage landscape. – J. Anim. Ecol. 77: 406–415.

Persson, A. S. et al. 2018. When beggars are choosers – how nesting of a solitary bee is affected by temporal dynamics of pollen plants in the landscape. – Ecol. Evol. 8: 5777–5791.

Potts, S. G. et al. 2010. Global pollinator declines: trends, impacts and drivers. – Trends Ecol. Evol. 25: 345–353.

Powney, G. D. et al. 2019. Widespread losses of pollinating insects in Britain. – Nat. Comm. 10: 1018.

Rader, R. et al. 2016. Non-bee insects are important contributors to global crop pollination. – Proc. Natl Acad. Sci. USA 113: 146–151.

Riedinger, V. et al. 2015. Annual dynamics of wild bee densities: attractiveness and productivity effects of oilseed rape. – Ecology 96: 1351–1360.

Roberts, D. R. et al. 2017. Cross-validation strategies for data with temporal, spatial, hierarchical or phylogenetic structure. – Ecography 40: 913–929.

Rundlöf, M. et al. 2014. Late-season mass-flowering red clover increases bumble bee queen and male densities. – Biol. Conserv. 172: 138–145.

Sagris, V. et al. 2015. Registration of land for the common agricultural policy management: potentials for evaluation of environmental policy integration. – Int. J. Agric. Resour. Govern. Ecol. 11: 24–44.

Scheper, J. et al. 2013. Environmental factors driving the effectiveness of European agri-environmental measures in mitigating pollinator loss – a meta-analysis. – Ecol. Lett. 16: 912–920.

Senapathi, D. et al. 2017. Landscape impacts on pollinator communities in temperate systems: evidence and knowledge gaps. – Funct. Ecol. 31: 26–37.

Shaw, R. F. et al. 2020. Mass-flowering crops have a greater impact than semi-natural habitat on crop pollinators and pollen deposition. – Landscape Ecol. 35: 513–527.

Sirami, C. et al. 2019. Increasing crop heterogeneity enhances multitrophic diversity across agricultural regions. – Proc. Natl Acad. Sci. USA 116: 16442–16447.

Smith, H. et al. 2014. Beyond dispersal: the role of animal movement in modern agricultural landscapes. Animal movement across scales. – Oxford Univ. Press, pp. 51–70.

Steffan-Dewenter, I. et al. 2002. Scale-dependent effects of landscape context on three pollinator guilds. – Ecology 83: 1421–1432.

Svensson, B. et al. 2000. Habitat preferences of nest-seeking bumble bees (Hymenoptera: Apidae) in an agricultural landscape. – Agric. Ecosyst. Environ. 77: 247–255.

Wenger, S. J. and Olden, J. D. 2012. Assessing transferability of ecological models: an underappreciated aspect of statistical validation. – Methods Ecol. Evol. 3: 260–267.

Williams, I. H. 2010. The major insect pests of oilseed rape in Europe and their management: an overview. – In: Biocontrol-based integrated management of oilseed rape pests. Springer, pp. 1–43.

Willmott, C. J. and Matsuura, K. 2005. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. – Clim. Res. 30: 79–82.

Winfree, R. et al. 2011. Native pollinators in anthropogenic habitats. – Annu. Rev. Ecol. Evol. Syst. 42: 1–22.

Woodcock, B. et al. 2017. Country-specific effects of neonicotinoid pesticides on honey bees and wild bees. – Science 356: 1393–1395.

Woodcock, B. et al. 2019. Meta-analysis reveals that pollinator functional diversity and abundance enhance crop pollination and yield. – Nat. Comm. 10: 1481.

Yates, K. L. et al. 2018. Outstanding challenges in the transferability of ecological models. – Trends Ecol. Evol. 33: 790–802.

Zambrano-Bigiarini, M. 2017. Package hydroGOF: goodness-of-fit functions for comparison of simulated and observed hydrological time series. R package version 0.3-10. – <https://cran.r-project.org/web/packages/hydroGOF/index.html>.

Zhang, H. et al. 2017. Arthropod pest control for UK oilseed rape–comparing insecticide efficacies, side effects and alternatives. – PLoS One 12: e0169475.

Zurbuchen, A. et al. 2010. Maximum foraging ranges in solitary bees: only few individuals have the capability to cover long foraging distances. – Biol. Conserv. 143: 669–676.

Zuur, A. et al. 2009. Mixed effects models and extensions in ecology with R. – Springer.

Zuur, A. F. et al. 2010. A protocol for data exploration to avoid common statistical problems. – Methods Ecol. Evol. 1: 3–14.