



Using historical data to estimate bumble bee occurrence: Variable trends across species provide little support for community-level declines

Laura Melissa Guzman^{*}, Sarah A. Johnson, Arne O. Mooers, Leithen K. M'Gonigle

Department of Biological Sciences, Simon Fraser University, 8888 University Drive, Burnaby, BC V5A 1S6, Canada

ARTICLE INFO

Keywords:

Bumble bees
species' declines
Occupancy models

ABSTRACT

Bumble bees are globally important pollinators, especially in temperate regions, and evidence suggests that many species are declining. One recent high profile study by Soroye et al. (2020) applied occupancy models to dated historical collection data to quantify declines across North America and Europe. The authors modelled 66 species across a set of sites spanning both North America and Europe, rather than confining species to sites where they might be expected to occur. In addition, they inferred non-detections for time intervals where there is no evidence that the site was visited (by forcing every site to have exactly 3 visits in each era). We use simulated data to (i) investigate the validity of methods used in that study and (ii) test whether a multi-species framework that incorporates species' ranges and site visitation histories produces better estimates. We show that the method used by Soroye et al. (2020) yields biased estimates of declines, whereas our framework does not. We use such a model to provide revised and appreciably lower estimates for bumble bee community declines, with species-specific trends more closely matching classifications from IUCN. The species level trends we provide can help inform future species-at-risk assessments. Well-parameterized occupancy models may be a powerful tool for assessing species-wide trends using curated historical collection data.

1. Introduction

Bumble bees are important globally for both natural and agricultural pollination services, as their large bodies and thick hair make them particularly efficient pollinators (Javorek et al., 2002; Williams et al., 2014). There is mounting evidence that at least some species have declined significantly (Goulson et al., 2008; Cameron et al., 2011; Graves et al., 2020); reasons are multifold and depend on species and location. Declines have been linked to habitat loss, agricultural intensification, pathogen transmission, and climate change (Cameron and Sadd, 2020), and several high profile articles have attempted to identify drivers of community change across many bumble bee species (Kerr et al., 2015; Miller-Struttman et al., 2015; Soroye et al., 2020; we note that some of these publications have also been challenged, e.g., see de Keyzer et al., 2016). Of these, the recent analysis by Soroye et al. (2020) applied occupancy models to a historical data set in an effort to test whether changes in species occurrence are linked to changes in local climatic conditions. They concluded that precipitous declines across both North America and Europe are, indeed, being driven by climatic shifts. If true, this is an alarming and important finding. This study was

ambitious both in the number of species and the spatial extent. However, the implementation of the occupancy models used may have yielded biased conclusions.

Occupancy models are a relatively new set of methods that have quickly grown in popularity and application (MacKenzie et al., 2002, 2006; Royle and Dorazio, 2008). These models, which are a derivative of mark-recapture models, consider a set of sites and estimate a species' average probability of being at any particular one of them (i.e., a number, referred to as "occupancy", that is between 0 and 1). Critically, this approach attempts to account for potential detection bias in the observations made at those sites (e.g., if detection probability is low, actual occupancy is usually estimated to be higher than it would be if inferred from raw data). Numerous methodological advances have expanded the scope and applicability of occupancy models. For example, multi-species occupancy models combine occurrence data from multiple species to enable identification of community-level trends (Iknayan et al., 2014). In contrast to post hoc comparisons of output from single-species models, in a multi-species framework, uncertainty in inferences at the species level propagates up to uncertainty in inferences at the community-level, and information from community-level trends is

^{*} Corresponding author.

E-mail addresses: laura_melissa_guzman@sfu.ca, laura_melissa_guzman@sfu.ca (L.M. Guzman).

also shared across species (Devarajan et al., 2020). In these ways, multi-species models may be both more transparent with regard to uncertainty, and more powerful at detecting real trends.

There are many community science and historical data sets with species recorded at known places and times and, in theory, occupancy models applied to these data could assess whether certain species are now more or less likely to be present across their historical range. However, the non-standardized methods typical of opportunistic collections make the data sets potentially problematic for occupancy models, as these models typically require standardized repeat visits over short time intervals (see Tingley and Beissinger, 2009 for a summary of the inherent analytical limitations associated with different types of historical data). To get around these assumptions, researchers have recently developed creative methods for analyzing historical data sets using an occupancy framework (e.g., Tingley and Beissinger, 2013; Frishkoff et al., 2014; Isaac et al., 2014; Woodcock et al., 2016; Iknayan and Beissinger, 2018; Outhwaite et al., 2018; van Strien et al., 2019; Graves et al., 2020; Soroye et al., 2020).

Approaches to deal with observation selection and changing sampling intensity through time and space have varied. For example, when using historical data, authors have directly re-surveyed sites that were surveyed in historical periods (Moritz et al., 2008; Tingley and Beissinger, 2013; Iknayan and Beissinger, 2018), or constrained analyses to locations where two or more sampling events occurred within a single calendar year (Graves et al., 2020) or locations that received visits in multiple years over the duration of the study (van Strien et al., 2013a; Isaac et al., 2014; Outhwaite et al., 2019). This replication is important for ensuring that models can estimate both occurrence and detection probability. However, depending on the availability of data, restricting to a subset of data can reduce the power of a study by reducing sample size (Isaac et al., 2014).

Given that opportunistic community science sampling and historical museum records only contain presence data, species' non-detections need to be inferred. One potential way to do this is to infer a non-detection for a particular species if a different species was observed at that same site on the same date (van Strien et al., 2013a; Kamp et al., 2016; Powney et al., 2019). The other species would likely need to be in the same taxonomic group (Outhwaite et al., 2019) but may or may not be considered to have been observed by the same observer (van Strien et al., 2010). Critically, each non-detection for each species must be assessed for each time interval at each site. An additional problem when applying occupancy models to opportunistic or historic data is that sampling effort may differ between time intervals. To account for this, authors have added covariates to detection probability that aim to capture potential sources of this variation (e.g., one might include the length of the list of all species recorded during a visit, aka "list length," as a proxy for sample effort; Isaac et al., 2014; Szabo et al., 2010; Kamp et al., 2016). A further step might be to filter the data by removing visits with only a single observation (Kamp et al., 2016) or to constrain analyses only to species that meet some minimum number of observations (Woodcock et al., 2016).

These methodological approaches have been primarily applied to atlases or opportunistic community science data (van Strien et al., 2010, 2013a, 2013b; Isaac et al., 2014; Kamp et al., 2016; Woodcock et al., 2016; Altwegg and Nichols, 2019; Outhwaite et al., 2019; Powney et al., 2019; Outhwaite et al., 2020). Such data tend to be collected by volunteers and, particularly for birds, are relatively well-structured (Gibbons et al., 2007). Historical museum records, on the other hand, present two additional problems. First, because museum records are often incomplete subsets of field collections, it is even more challenging to infer when sampling actually occurred (i.e., one cannot necessarily determine from museum specimens alone when collectors went to sampling locations, because common species are sometimes not curated (Ascher et al., 2020)). Second, it is unclear how to define "sites" over which species occupancy can then be modelled (i.e., over what spatial resolution should we lump observations together and then, given a

spatial resolution, which sites should be modelled for which species, e.g., how do we account for different species' ranges). Studies are now starting to investigate this question (Jönsson et al., 2021).

Authors have taken multiple approaches to site selection when modelling communities where many sites are outside the realistic range for many species (e.g., in a broad sense, sites on one continent are outside the range of a species not found on that continent). Frishkoff et al. (2014) used a multi-species framework that constrained bird species occupancy, such that species were not modelled in geographic regions where they had never been detected, and Graves et al. (2020) modelled the western bumble bee (*Bombus occidentalis*) using a single-species framework where each of four regions in western North America were modelled separately. Another approach is to instead include spatial covariates for modelling occupancy, such as climate or Cartesian coordinates (Dennis et al., 2017; Altwegg and Nichols, 2019).

It is in this methodological setting that Soroye et al. (2020) developed a method to estimate changes in species occupancy using historic museum specimen records. Specifically, Soroye et al. (2020) used single-species models to model each of 66 bumble bee species across a spatial grid of sites that spanned Europe and North America, and used no covariates for occupancy. Soroye et al. (2020) divided historical observations across a spatial grid into three time period subsets ("time intervals") for each grid cell ("site") in each of two time periods ("eras", historical versus modern) that differed in length (75 vs 15 years, respectively). The authors only considered sites where at least one specimen of any of their target species had been collected and then, for those sites, inferred non-detections for all other species and/or time intervals. They subsequently modelled each species across all considered sites and then assessed occupancy change, per species, at the subset of sites where that species had been detected. In doing this, the authors implicitly assumed 1) that North American species might have been present at European sites, but had simply gone undetected, and vice versa for European bees at North American sites, and 2) that every site in their final data set had been visited in every time interval. Thus, Soroye et al. (2020) took steps to deal with the challenges that arise when applying occupancy models to historical data. However, rather than validate their method first using simulated data, they applied it, untested, to field data. For the reasons mentioned above, which we explore further below, their approach was not sufficient and, in fact, yielded biased results.

We use simulated data to identify if and how temporal patterns in detection, visitation, and/or occupancy mediate biases in inferred occupancy that arise when species are modelled across sites outside their ranges or across time intervals where visits did not occur. Simulations are useful in these circumstances, because we can explore model performance in the context of known true values of the parameters that underlie trends. In the context of our simulation results, we then revisit one main question posed by Soroye et al. (2020), namely the extent to which bumble bees have declined in North America and Europe. We find that inferences about community-level bumble bee decline are strongly impacted by these particular modelling decisions. Overall, occupancy models that best mitigate the biases we uncover estimate notably weaker community-wide declines for bumble bees in both North America and Europe. Revised estimates for species-specific declines also show reduced bias and, thus, are likely closer to true values. Such revised values can potentially better inform conservation efforts, and we hope our results will help guide future conservation studies that make use of historical observations.

2. Methods

2.1. Data simulation

We simulate species occupancy for a community of N species over J sites across K eras (eras could be years or ranges of years). We further assume that, in era k , site j is visited up to I times. We assume that each

species occupies a different, potentially overlapping, subset of the J sites. Specifically, species i potentially occupies sites $R_i = \{s_{i,1}, s_{i,2}, \dots, s_{i,r_i}\}$, where $r_i \in \{1, 2, \dots, J\}$. We refer to R_i as the “range” of species i and the sites within R_i as the “relevant sites” for species i . Because our goal here is not to simulate species’ ranges, we set N , J , and R_i (for each i) to match values for the bumble bee data that we subsequently re-analyze.

We simulate occupancy for each species only across the sites in that species’ range. For species i , we set occupancy to zero for all sites $j \notin R_i$. For site $j \in R_i$, we draw occupancy in era k from a Bernoulli distribution with probability $\psi_{i,j,k}$, where

$$\psi_{i,j,k} = \text{expit}[\psi_0 + \psi_{sp}[i] + \psi_{era}[i] \times (k - 1)] \quad (1)$$

Here, ψ_0 is the baseline occupancy across species, $\psi_{sp}[i]$ is a random species-specific intercept, and $\psi_{era}[i]$ is a random species-specific effect of era. This latter term allows species to vary in their temporal occupancy trends. Note that we use $k - 1$ above to facilitate interpretation of parameters when there are only two eras. We further assume that both ψ_{sp} and ψ_{era} are normally distributed, such that:

$$\psi_{sp}[i] \sim \mathcal{N}(0, \sigma_{\psi,sp}) \quad (2)$$

$$\psi_{era}[i] \sim \mathcal{N}(\mu_{\psi,era}, \sigma_{\psi,era}) \quad (3)$$

Here, $\sigma_{\psi,sp}$ specifies the variation in occupancy across species, $\mu_{\psi,era}$ specifies the mean effect of era on species occupancy (positive/negative values imply that species are, on average, increasing/decreasing in occupancy through time), and $\sigma_{\psi,era}$ specifies the variation in temporal species-specific occupancy trends.

Among occupied sites, a species can only be detected at a site if that site is visited. We modelled site visitation in two ways. First we considered a scenario where each site is visited I times in each era (no missing visits). Second, we considered a scenario where not every site received a visit in every time interval (missing visits). Here, site visits for each of the I intervals in era k to site j were drawn from a Bernoulli distribution with probability

$$v_{j,k} = \text{expit}[v_0 + v_{era} \times (k - 1)], \quad (4)$$

where v_0 is baseline probability of visitation (on the linear scale) and v_{era} is an effect of era that allows site visitation probability to change systematically between time periods. We then simulate detections for each species only across the time intervals where visits occurred to the sites that it occupies. Specifically, for species i , if site $j \in R_i$ and species i occupies that site, we draw species’ detections across the visits that occurred in era k from a Bernoulli distribution with probability $p_{i,j,k}$, where

$$p_{i,j,k} = \text{expit}[p_0 + p_{sp}[i] + p_{site}[j, k] + p_{era} \times (k - 1)], \quad (5)$$

where p_0 is the baseline probability of detection across species (on the linear scale), $p_{sp}[i]$ is a random species-specific intercept, $p_{site}[j, k]$ is a random site-specific intercept that varies by era, and p_{era} is an overall effect of era. The random site-specific intercept allows detection probability to differ independently across sites and between eras; for historical data, this allows the model more flexibility and helps account for the fact that sites likely differ in sampling effort across space and/or time. The fixed effect of era, p_{era} further allows detection probability to change systematically between time periods. We assume that both p_{sp} and p_{site} are normally distributed, such that:

$$p_{sp}[i] \sim \mathcal{N}(0, \sigma_{p,sp}) \quad (6)$$

$$p_{site}[j, k] \sim \mathcal{N}(0, \sigma_{p,site}) \quad (7)$$

Here, $\sigma_{p,sp}$ specifies the variation in detection across species and $\sigma_{p,site}$ specifies the variation in detection across sites and eras.

2.2. Models

We compare performance of five (of many possible) models on data sets simulated using the above procedure in order to identify which models provide the most robust inference for historical data akin to those analyzed in Soroye et al. (2020).

1. $SS_{all,all}$ (the model used in Soroye et al., 2020): We first model occupancy for each species across all J sites and all I intervals using N separate single-species (SS) models. Modelling all intervals assumes that every site was visited in every time interval. In each model, we allow occupancy and detection probability to independently increase or decrease through time. Using the posterior model estimates for the latent site-specific occupancy states, we calculate the proportional increase or decrease in occupancy for each species across only the sites where that species was detected at least once (but note that posteriors are derived from models that were run across all sites). We then combine these species-specific estimates to calculate the mean proportional change in occupancy across species. We calculate this a posteriori latent state metric in order to reproduce the methods of Soroye et al. (2020). For the multi-species models described below, we calculate community change in occupancy directly from model-estimated parameter values.
2. $MS_{all,all}$: We use a multi-species (MS) framework to model occupancy of each species across all J sites and across all I intervals. This assumes all sites were visited in each of the I intervals. We include this model for completeness, but do not examine it in detail due to its relatively poor performance.
3. $MS_{range,all}$: We use a multi-species framework to model occupancy of each species only across the sites in its **range** (i.e., species i is modelled across sites R_i only) and across all I intervals. This assumes all sites were visited in each of the I intervals.
4. $MS_{range,detected}$: We use a multi-species framework to model occupancy of each species only across the sites in its **range** and only across intervals where at least one species was **detected** and so we have positive evidence that the site was, indeed, visited. Occupancy models for unstructured data typically only model visits where at least one species was detected (van Strien et al., 2010, 2013a; Kamp et al., 2016; Outhwaite et al., 2019; Powney et al., 2019).
5. $MS_{range,visits}$: We use a multi-species framework to model occupancy of each species only across the sites in its **range** and only across the intervals where **visits** actually took place. This model represents a best-case scenario where we have full information about both species’ ranges and visitation history and, thus, can include sites that have no recorded detections for any species. For historical data sets, acquiring such information may require substantial effort, if it is even possible.

For the multi-species models, we allowed for maximum biological realism by allowing both occupancy and detection probability to differ between species and also to increase or decrease through time, i.e., by fitting fixed and random effects on occupancy and detection probability that mirror the equations specified above for data-simulation (Eq. (1) for occupancy and Eq. (5) for detection). The only differentiating factor between these models is the subset of sites and time intervals that are included (see Fig. 1 for a schematic). To account for potential differences in effort across time intervals, and for consistency with Soroye et al. (2020), we also include the total number of specimens collected during a time interval as a predictor on detection probability in $SS_{all,all}$ (but note we did not include it in subsequent analyses for the reasons given below). While this measure was intended to serve as a proxy for “effort”, and indeed many authors use it to account for differences in sampling effort (Szabo et al., 2010; Isaac et al., 2014; Kamp et al., 2016; Woodcock et al., 2016; Outhwaite et al., 2019), we note that because detection probability is calculated using a logit link function, “zero effort” as modelled in Soroye et al. (2020) does not translate into a zero

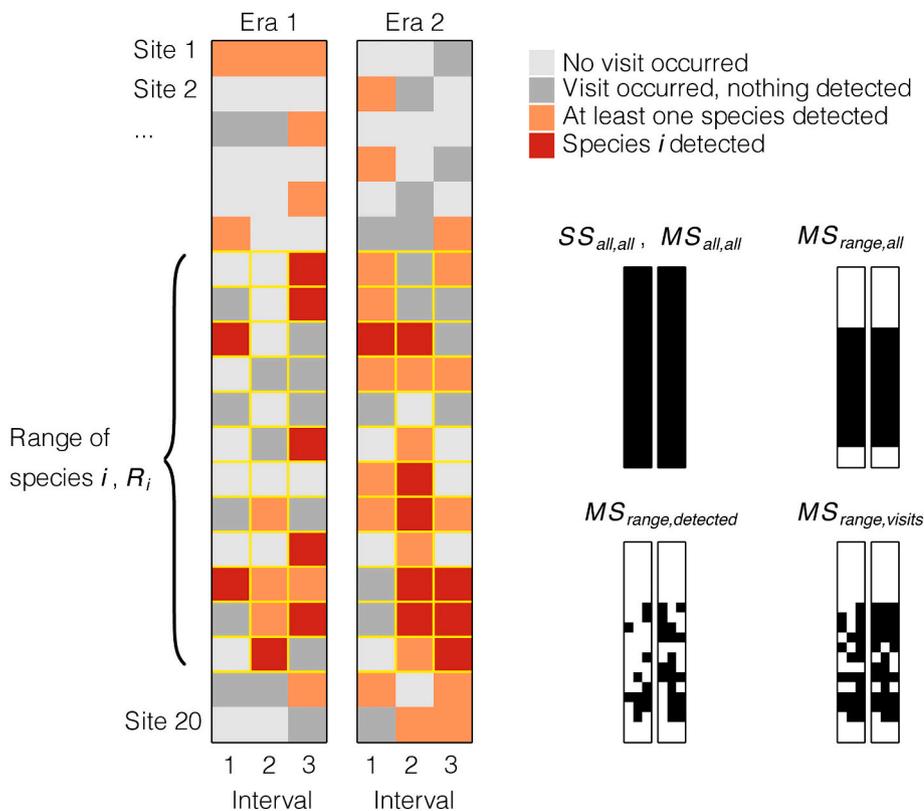


Fig. 1. Schematic illustrating the different classes of visits, with respect to species i . Each row corresponds to one of the J total sites (here, $J=20$) across $K=2$ eras and cells within each row correspond to $I=3$ time intervals where visits to those sites can potentially occur (e.g. historic Era 1 vs. modern Era 2). Potential visits to sites that are within the range of species i , R_i , are outlined in yellow. Each model we present considers a different combination of sites and time intervals, as shown in smaller panels, where black indicates the modelled sites and time intervals. Specifically, $SS_{all,all}$ and $MS_{all,all}$ model every species over every site and every time interval (all cells), $MS_{range,all}$ models each species over every time interval at the sites in that species' range (all cells outlined in yellow), $MS_{range,detected}$ models each species at every site in that species' range over every time interval where at least one species was detected, (all red and orange cells outlined in yellow), and $MS_{range,visits}$ models each species over every actual visit to a site in that species' range (all red, orange, and dark grey cells outlined in yellow). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

probability of detection, but rather a detection probability equal to $\expit(x)$ where x denotes the intercept probability of detection on the linear scale. This problem manifests when lists of length zero (e.g., time intervals where visits either did not happen or did happen, but yielded no detections) are included in the analysis. Previous analyses that included list-length as a model predictor would not encounter this problem, as they typically only include samples where a “list” actually existed, which implied at least one species detection (Kamp et al., 2016). The random effects in our multi-species models here render such a step unnecessary and, thus, we include no such term for those models.

We use uninformative priors for all parameter values (normal distributions for ψ_0 , $\mu_{\psi,era}$, p_0 , p_{era} , v_0 , v_{era} and uniform priors for $\sigma_{\psi,sp}$, $\sigma_{\psi,era}$, $\sigma_{p,sp}$, $\sigma_{p,site}$). JAGS code for data simulation and all model code is available at https://github.com/lmguzman/occupancy_bbees.

2.3. Model tests

We assess models by quantifying their performance when inferring the parameter values under which the data were simulated, as well as their performance inferring species-specific trends. We consider a range of scenarios of increasing, decreasing, or constant detection and/or site visitation and/or occupancy probability through time. Because comparisons between “historical” and “modern” time periods are of particular relevance to conservation (e.g., Tingley and Beissinger, 2013; Iknayan and Beissinger, 2018; Soroye et al., 2020), we focus here on the case of two time periods (or two “eras” in the terminology of our model).

2.4. Re-analysis of bumble bee declines

Incorporating insights from our analyses of simulated data sets, we re-analyze the subset of the data analyzed by Soroye et al. (2020) for North America and Europe. We use a multi-species framework, considering each continent separately and, within a continent, limiting each species to sites within an inferred range. We infer a species' range to be the set of all sites within the convex hull that includes all sites where that

species has been detected at least once (see Fig. S1 for a sample range for a single species). We acknowledge that there are many alternative ways one might model species' ranges and we have used the simplest one extracted from the data set. In this multi-species framework we also include only visits where at least one specimen was collected (i.e., we use $MS_{range,detected}$, and thereby assume that, if no bumble bee of any species in the dataset was collected at a site during a specific time interval, no visit occurred).

3. Results

3.1. Model performance tests

We find that, when detection probability is not constant across time periods, some models produce biased estimates of community-wide temporal changes in occupancy. Most notably, modelling species across all sites and all potential visits using the single-species framework presented in Soroye et al. (2020) ($SS_{all,all}$) incorrectly identifies changes in detection as evidence for changes in occupancy (Fig. 2). A multi-species approach that models all sites and all time intervals ($MS_{all,all}$) yields less biased estimates of changes in occupancy, but strongly biased estimates of mean occupancy (Fig. S2, Table 1). In contrast, multi-species frameworks that model each species across the sites in their respective ranges ($MS_{range,all}$, $MS_{range,detected}$, $MS_{range,visits}$) yield much improved estimates of occupancy change when detection probability changes through time (yellow points in Figs. S3, S4, S5). Because restricting models to species' ranges performs better than using all sites, we only investigate effects of visitation history (next) for the multi-species models that incorporate species' ranges.

Differences in visitation frequency across time can also bias occupancy estimates. Specifically, a decreasing frequency of visits through time leads to similar biases as decreasing detection probability for $SS_{all,all}$ (Fig. 2). Multi-species models that include all of the potential time intervals ($MS_{range,all}$) also yield biased estimates of both mean occupancy (ψ_0) and temporal trends in occupancy ($\mu_{\psi,era}$) (blue points in

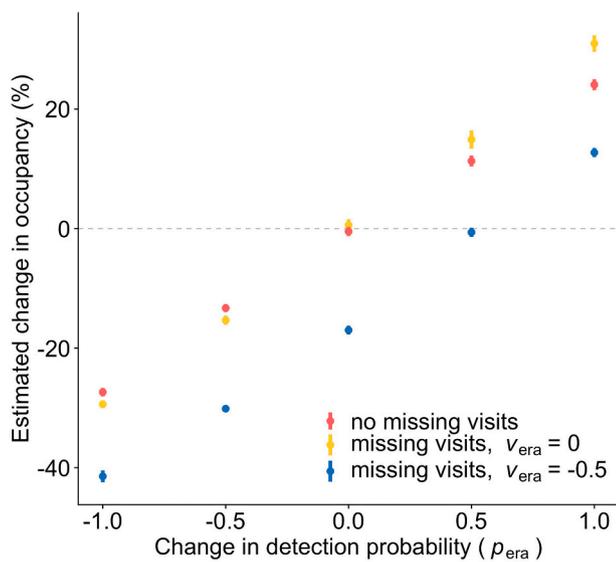


Fig. 2. The metric used to estimate mean percent change in occupancy by Soroye et al. (2020) yields biased values when detection or visitation probability are not constant through time. The dashed grey line indicates the true change in occupancy for simulated data ($\mu_{\psi,era}=0$, in all cases). Points and vertical bars show means and standard errors across 10 simulated data sets. Because these points are means across multiple individual occupancy models, we cannot easily present 95% Bayesian Credible Intervals here, as we do for other figures. Other parameter values were: $I=66$ species, $J=1905$ sites, $K=2$ eras, $I=3$ time intervals, $\sigma_{\psi,sp}=0.5$, $\sigma_{\psi,era}=0.2$, $p_0 = -0.5$, $\sigma_{p,sp}=0.5$. Values on the x-axis can be converted to mean change on the probability scale as $\text{expit}(p_0 + x) - \text{expit}(p_0)$.

Fig. S5). However, multi-species models that consider only time intervals where visits actually happened ($MS_{range,visits}$) or time intervals where at least one species was detected ($MS_{range,detected}$) yield reliable estimates of mean occupancy (ψ_0) and temporal trends in occupancy ($\mu_{\psi,era}$) (blue points in Figs. S3, S4).

While both $MS_{range,detected}$ and $MS_{range,visits}$ correctly estimate patterns in occupancy when visitation declines through time, only the latter correctly estimates detection probability (Fig. S6). $MS_{range,detected}$ consistently overestimates mean detection probability (p_0) and the change in detection probability (p_{era}) when occupancy also changes through time (Fig. S7). Thus, obtaining correct estimates of both occupancy trends and detection probabilities requires information about the actual visitation history at each site (for a summary of all the model comparisons see Table 1).

Model performance for individual species - quantified by calculating the root mean square error (RMSE) for model-estimated species-specific occupancy trends - is best (lowest error) for multi-species models when species are only modelled across their respective ranges, and over the time intervals where visits actually occurred ($MS_{range,visits}$) or intervals where at least one species was detected ($MS_{range,detected}$) (Fig. S8). Error

is high when using the single-species metric of Soroye et al. (2020) ($SS_{all,all}$) or when species are modelled over their ranges and all time intervals in a multi-species framework ($MS_{range,all}$).

3.2. Re-analysis of bumble bee declines

Our analyses on simulated data sets show that using a multi-species framework to model each species across its range produces reliable estimates of temporal patterns in occupancy when actual site visitation history is known ($MS_{range,visits}$) or when it is inferred based on species' detections ($MS_{range,detected}$). Because we do not know the actual site visitation history for bumble bees, we used the latter model to re-evaluate temporal trends in bumble bee occurrence. We find that the magnitude of estimated species' declines is appreciably lower for bees on both continents: for North America, we find a 5% decline with a 95% BCI = $[-20\%,12\%]$, compared to Soroye et al. (2020) who estimated a 46% decline, SE = 3%; for Europe we find a 6% decline, 95% BCI = $[-21\%,2\%]$, compared to Soroye et al. (2020) who estimated a 17% decline, SE = 5% (see Fig. 3 for a graphical comparison of the per-species declines and Table 2 for a summary of results, and Table S1 for a full summary of model-estimated parameters). Credible intervals for estimated declines on both continents include zero and, therefore, we cannot exclude the possibility that mean bumble bee occupancy has not changed between the periods. Until we have better estimates of actual site visitation for these data, identifying reliable trends in detection is not possible with the models we have considered here.

Soroye et al. (2020)'s estimated species-specific declines are negative and exceed 10% for all but one species (*Bombus terrestris* which does not occur in North America, see paragraph below) in North America, many of which are classified as stable by the IUCN. $SS_{all,all}$ systematically biased estimates of change in occupancy downwards when visitation declined through time (Fig. 2), and we find evidence for such a decline in visitation between eras in North America (changes in visitation are shown in Fig. 4 and discussed further below). In contrast, our multi-species approach produces estimates more in line with the IUCN classifications (Figs. 5, 6), with notable increases for species that are known to exhibit modern range expansions (e.g., *B. impatiens* with several recent western introductions, Looney et al., 2019; Palmier and Sheffield, 2019; *B. bimaculatus*, with indication of both an eastward and westward range expansion from modern community science observations, The Xerces Society, Wildlife Preservation Canada, York University, University of Ottawa, The Montreal Insectarium, The London Natural History Museum, BeeSpotter., 2017; *B. cryptarum* with evidence for a westward range expansion, Owen et al., 2012; and *B. hypnorum* colonizing the mainland of the United Kingdom, Crowther et al., 2019) or that have been assessed as increasing in previous studies (e.g., *B. bimaculatus*, *B. griseocollis*, *B. impatiens*, *B. perplexus*, Colla et al., 2012; Bartomeus et al., 2013). We estimate the most dramatic decline in North America (73%) for *B. bohemicus*, which is a specialized social parasite and, therefore, expected a priori to experience faster and steeper declines as an indicator or precursor to declines in its host(s). *B. bohemicus* is listed as Critically Endangered by IUCN for its North American range. It has

Table 1
Summary of the model results where + is overestimated, - is underestimated and = is accurately estimated.

Model	Missing visits	Bias in missing visits	Mean occupancy	Change in occupancy	Mean detection	Change in detection
$SS_{all,all}$	No			+ and -		Fig. 2
	Yes	No		+ and -		
	Yes	Yes		+ and -		
$MS_{all,all}$	No		-	+ and -		Fig. S2
	Yes			=		Fig. S5
$MS_{range,detected}$	Yes	No	=	=		Fig. S3
	Yes	Yes	=	=	+	+ and -
	Yes	No	=	=		Fig. S4
$MS_{range,visits}$	Yes	Yes	=	=		Fig. S7
	Yes	Yes	=	=		Fig. S6

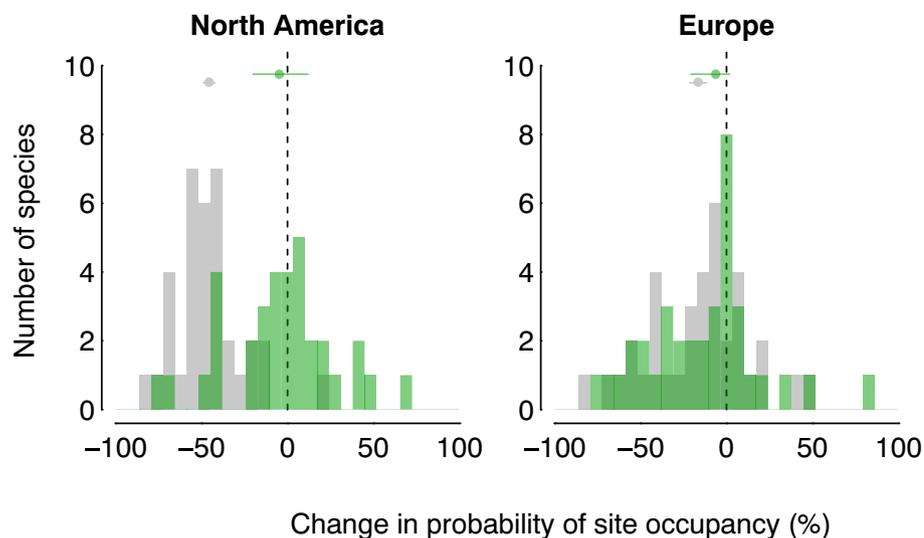


Fig. 3. Multi-species versus single species changes in bumble bee occupancy across two continents. Grey bars correspond to model results presented in Fig. 2 in Soroye et al. (2020), whereas green bars show estimates from a multi-species model where species are modelled across sites in their inferred ranges and time intervals where at least one species was detected ($MS_{range,detected}$). Points at the top of each panel indicate mean across species and horizontal bars indicate standard error (grey) or 95% Bayesian Credible Interval (green). Note that, in contrast to Soroye et al. (2020), our analysis for North America does not include *Bombus terrestris* or *Bombus lucorum* (neither is known to inhabit that continent) but does include *Bombus distinguendus* (the authors removed *B. distinguendus* from their analysis after estimating a >100% increase in occupancy). Because there are only a small number of records from these species, our conclusions are largely unchanged if we consider the same species set as Soroye et al. (2020) (see Table S2). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 2

Estimated proportional changes in occupancy for data presented in (Soroye et al., 2020) using the method presented in that paper ($SS_{all,all}$), as well as a multi-species model where species are modelled over sites in their inferred ranges and time intervals where at least one species (from any of those modelled) was detected ($MS_{range,detected}$).

		Δ Occ	SE	95% BCI
North America	$SS_{all,all}$	-45.6	3.4	-
	$MS_{range,detected}$	-4.8	-	[-20.1,11.9]
Europe	$SS_{all,all}$	-16.5	4.9	-
	$MS_{range,detected}$	-6.3	-	[-20.6,1.8]

confirmed breeding records only in colonies of *B. affinis* and *B. terricola*, and likely also parasitizes *B. occidentalis* (Williams et al., 2014), all species also estimated to be exhibiting major declines. In contrast, we estimate the most dramatic decline in Europe (78%) for *B. pomorum*, not a social parasite but a species extirpated from the UK and with evidence supporting a decline at its northern range limits (Rasmont and Iserby, 2010; Jeffers, 2017). Our model-estimated declines for this species suggests the IUCN Red List status of “vulnerable” may need to be reassessed (IUCN, 2020). Our species-specific responses for bees in Europe largely agree with those estimated by Soroye et al. (2020) (Fig. 5). This finding is potentially attributable to more consistent site visitation between time periods in Europe; for North America 64% of unique site \times time interval combinations comprised at least one species’ detection in the historical era, compared to only 27% in the modern era, whereas, for Europe 65% comprised at least one detection in the historical era, compared to 55% in the modern era (Fig. 4). Our analyses on simulated data showed that when visitation and detection are constant in time, $SS_{all,all}$ can produce reliable species-specific responses and, for Europe, at least one of these conditions appears to hold.

We also note that the inclusion of the European *B. terrestris* and *B. lucorum* in the North American analyses of Soroye et al. (2020) was likely an error, and both species have been excluded from our North American analyses (note these species are not included in Williams et al., 2014). Commercial rearing and international trade of *B. terrestris* has led to invasive populations establishing outside of their native European range (e.g., South America, Torretta et al., 2006, and Japan, Inoue et al., 2008), but *B. impatiens* replaces *B. terrestris* in North America and importation of the exotic *B. terrestris* into North America is not permitted (Goulson, 2010). *B. lucorum* is a member of a cryptic European species complex, and at one time taxonomists believed that it occurred in northwestern North America. A series of taxonomic revisions now

indicate that any *B. lucorum* recorded in North America are likely *B. cryptarum*, a separate species in the same subgenus that is known to occur in both North America and Europe (Scholl et al., 1990; Williams, 2011; Williams et al., 2012). Excluded observations of *B. lucorum*, all but one from the historical era, may thus spuriously contribute to the marked increase in occupancy we found for *B. cryptarum* in North America. Though range expansion has been documented for this species (Owen et al., 2012), the modelled increase of 71% is likely inflated. Analyses for North American bees using our revised methods on the identical set of species as considered in Soroye et al. (2020) does not change any of our conclusions here (see Table S2).

4. Discussion

Using simulated data, we found that the single-species occupancy models set up and parameterized like the one recently developed by Soroye et al. (2020) can incorrectly attribute changes in site visitation pattern and/or changes in a species’ probability of detection to changes in that species’ probability of occupancy. In other words, the model of Soroye et al. (2020) fails to accomplish the primary goal of occupancy models, which is to estimate changes in occupancy that explicitly account for changes in detection probability. In contrast, we found that the multi-species model that explicitly incorporate species’ ranges and detection histories produce reliable estimates of temporal trends in occupancy, even when visitation to these sites is inferred based on species’ detection histories. Such models do yield biased estimates of species’ detection probabilities, but appear to have comparably minimal effects on estimates of species’ temporal trends in occupancy.

Using this framework, we re-analyzed the data presented in Soroye et al. (2020), and found much of their estimated declines in North America are likely the result of a statistical artifact arising from temporal patterns in detection probability and site visitation. As expected from multi-species models, uncertainty associated with our community-level estimates was also greater than that reported in Soroye et al. (2020) for both continents.

Our multi-species analysis also seems to provide better estimates of species-specific trends when detection probability is not constant through time. While we did not explicitly investigate a single-species framework that incorporates species’ ranges and site visit histories, we note that using a multi-species approach allowed us to model species with few records. For example, in North America, *Bombus distinguendus* was only present in two sites in the current era. Robustness to the inclusion of rare species is a known advantage to multi-species occupancy

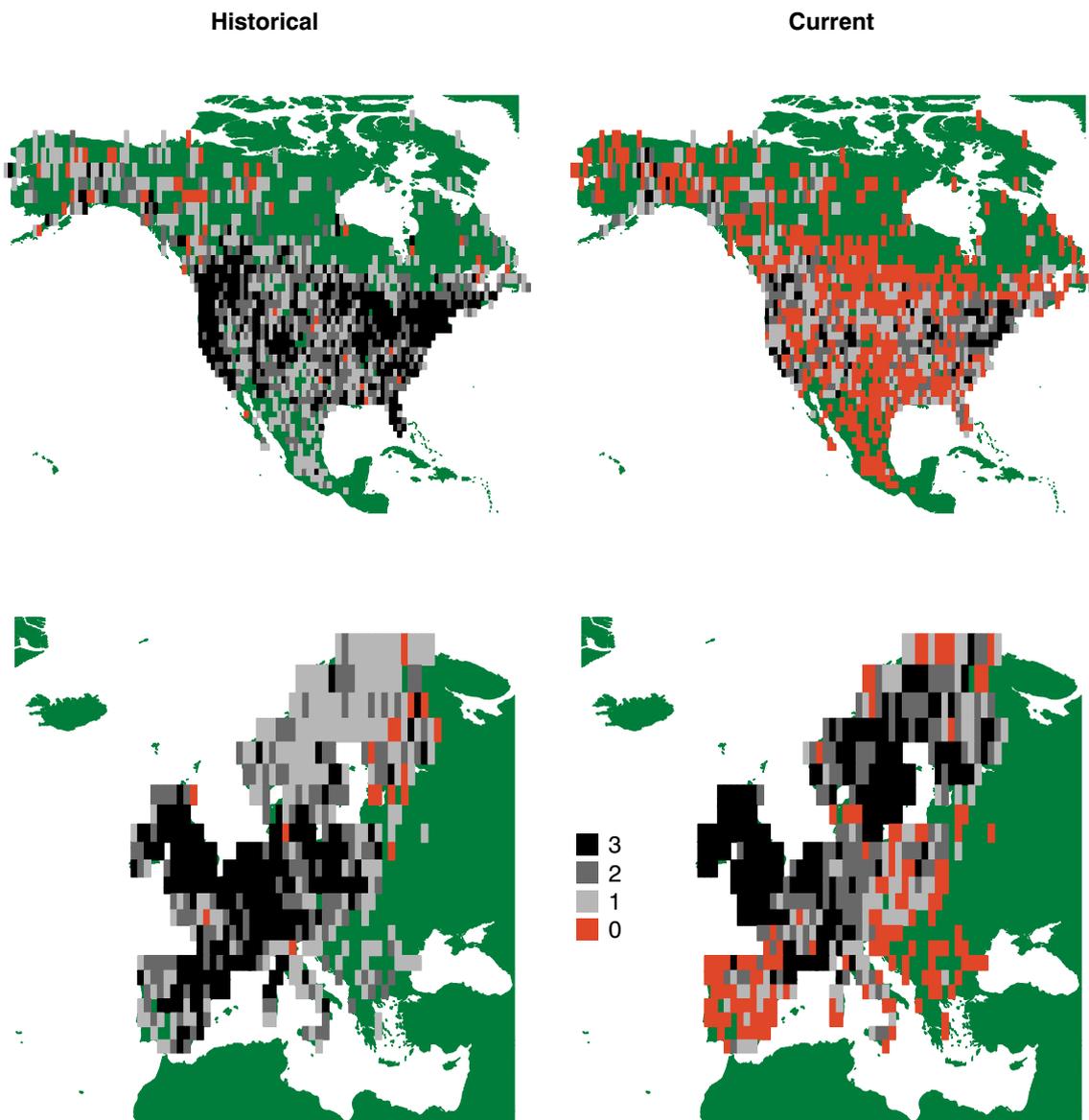


Fig. 4. Number of time intervals where a visit was known to have occurred (e.g., time intervals with recorded species' detections) at each site in each era on each continent. Red indicates that no species was detected at the corresponding site in that era and, thus, no visit may have actually happened in any of the three time intervals; light grey indicates that only one of the three time intervals was known to have contained a visit, and so on. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

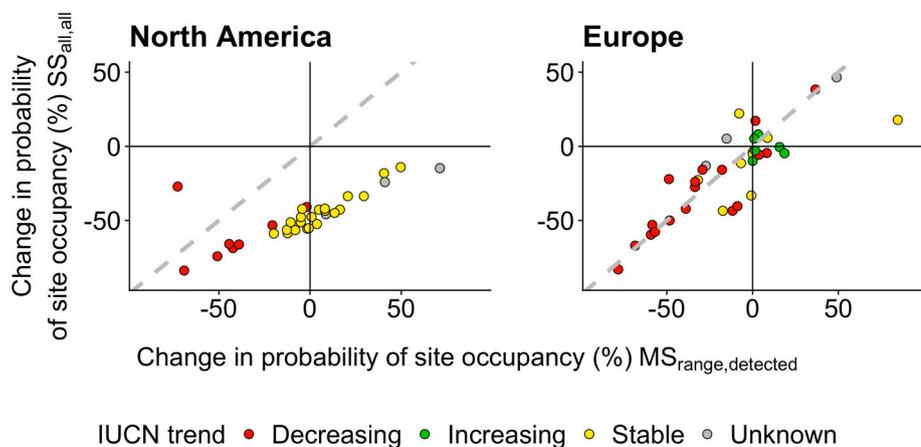


Fig. 5. Species-specific estimates of occupancy change from Fig. 3, coloured by IUCN Red List population trend. The multi-species estimates agree with Red List trends on both continents. Estimates from Soroye et al. (2020), $SS_{all,all}$, generally agree for Europe, but always overestimate declines in North America, likely due to changing probability of site visitation there. We removed *Bombus terrestris*, *Bombus lucorum*, from the North America panel, for reasons given in the Fig. 3 caption and *Bombus distinguendus* because Soroye et al.'s method yields a 700% increase, which did not fit in this plot. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

IUCN trend ● Decreasing ● Increasing ● Stable ● Unknown

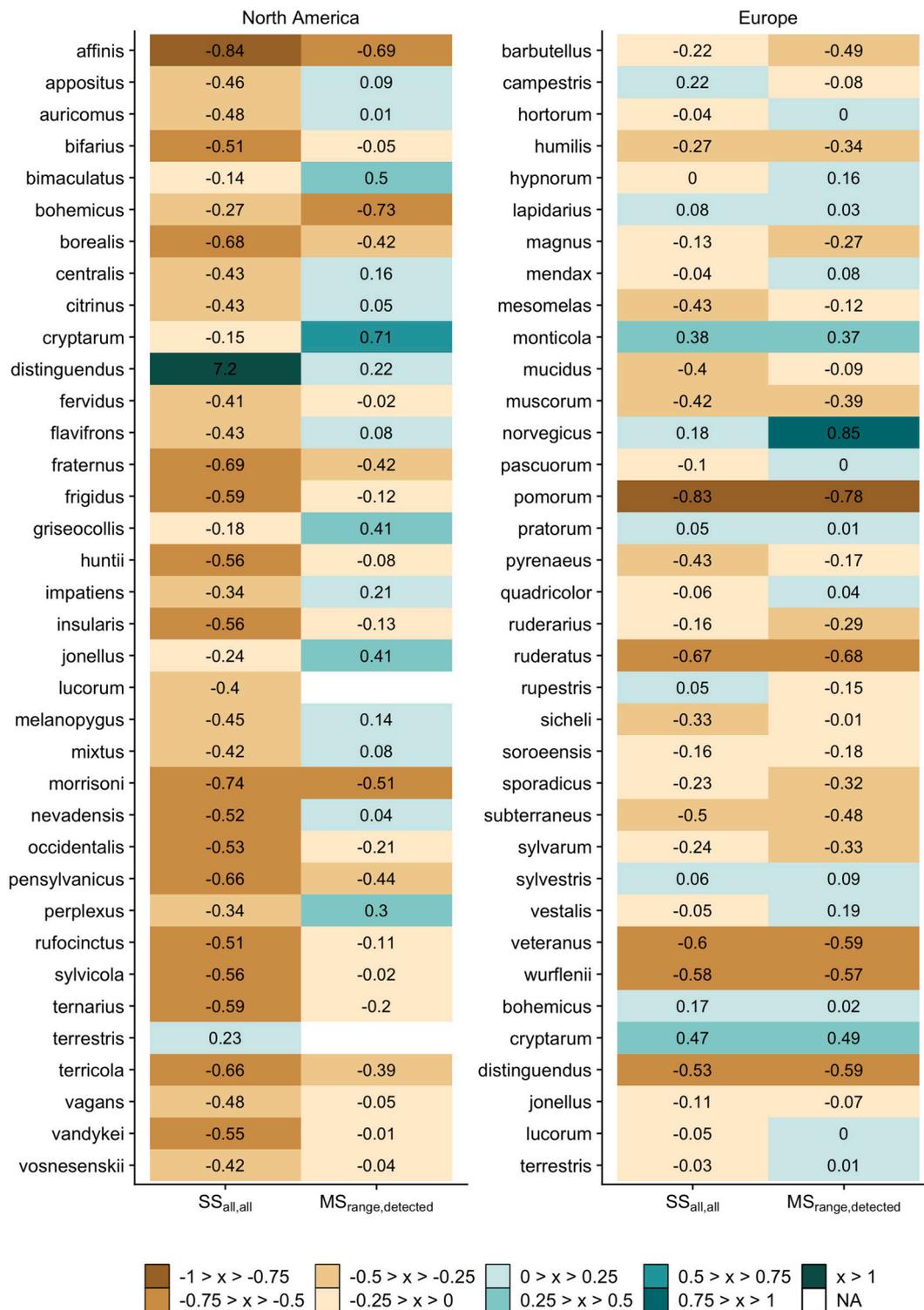


Fig. 6. Species-specific estimates of occupancy change. Estimates from Soroye et al. (2020), SS_{all,all}, generally agree for Europe, but always overestimate declines in North America, likely due to changing probability of site visitation there. We removed both *Bombus terrestris* and *Bombus lucorum* from the North America panel, for reasons given in the Fig. 3 caption.

models (Zipkin et al., 2009). In North America, Soroye et al. (2020) estimated that all but two species have experienced major declines and, of these two, they omitted one from their estimated average declines (*Bombus distinguendus*, likely omitted for the reason given above) and the other is not actually known to occur in North America (*Bombus terrestris*). These trends across species are at odds with IUCN Red List population trend reports (IUCN, 2020). Our multi-species estimates, which

are more conservative in North America, are more consistent with IUCN population trends. While we do not find major overall declines across all species, we do find that many individual species appear to be in trouble. This highlights the fact that presenting the “mean decline” across a community may not be useful for conservation applications, as increases in some species effectively cancel declines in other species. This may lead to inappropriately deprioritizing communities composed of range-

expanding common species that are replacing declining rare species (i. e., biotic homogenization, McKinney and Lockwood, 1999). Our revised estimates highlight specifically which North American species likely need conservation resources, and provide critical information for designing meaningful management actions. In contrast, in Europe, where site visitation did not change appreciably, both our and Soroye et al. (2020)'s estimates are more consistent and largely match IUCN Red List population trend reports.

Estimated declines in actual visitation and/or detection probability in North America could be, at least in part, a consequence of Soroye et al. (2020)'s decision to use time intervals that comprise different temporal durations in their "historical" baseline and "current" periods (see Ascher et al., 2020, for additional commentary related to this potential bias). Specifically, their baseline encompassed 75 years, split into three 25 year sample intervals, whereas the current era encompassed 15 years, split into three 5 year sample intervals. Sample interval durations were chosen so that their measure of overall sample effort did not differ greatly between eras. Interestingly, while the number of species' detections in the modern and historical periods did not differ dramatically, the distribution of detections across space in the modern era are much less evenly dispersed, with large portions of central North America and Mexico completely devoid of any records (Fig. 4). Further, 5 years of samples in recent years may not be equivalent to 25 years of samples in historic years. For example, a longer sample interval provides greater opportunity to detect species that fluctuate in abundance through time, which could bias estimates of occupancy in the longer time intervals upwards. While these sample periods were also used in Europe, opportunistic datasets there tend to be more complete. For example, the Bees, Wasps and Ants Recording Society is a unique dataset with a large coverage (Woodcock et al., 2016). In addition, since many opportunistic datasets are biased towards human population density and infrastructure (Geldmann et al., 2016), we would expect Europe to be more consistently sampled across time than North America (keeping in mind that in the analyses presented here and in Soroye et al. (2020), modern time intervals are 5× shorter than historical ones).

Overall, we found that modelling species across their ranges using multi-species occupancy models yields accurate estimates of temporal patterns in occurrence. However, species ranges are not always available and can change through time. Here, we approximated a species' range by including all sites within a convex hull drawn around the set of sites where that species had been detected at least once. Alternatively, explicitly incorporating ecological variables as predictors on occupancy and/or detection, such as latitude/longitude, elevation, climate, or eco-region will also help prevent spurious inferences due to the inclusion of modelled sites that are outside a species' range (Graves et al., 2020). Spatial effects have been studied extensively in joint species distribution models and results from the work of SDMs could likely also be leveraged to provide better bounds on species' ranges for occupancy models applied to historic data (Pollock et al., 2014; Tobler et al., 2019).

For the data we consider here, modelling only time intervals where at least one species in the data set was detected appears to function as a robust work-around for a lack of information about true visitation history. Ideally, one would use additional information to make inferences about the true visit history. As discussed earlier, detection records for other species (outside those under direct consideration) might be helpful in identifying which time intervals contain visits that happened but did not yield detections of any focal species. For our analysis of the bumble bee data, we assumed that no visit occurred during a time interval if there were no corresponding species' detections in that interval. By using records of specimens from other species (i.e., those from other data sets), or more information on survey procedures, one could conceivably reconstruct better visitation history for historical data sets, thereby enabling application of $MS_{\text{range,visits}}$, however, we did not do this here. Whether or not this is an appropriate assumption will depend on both the resolution of the spatial grid (and thus, the size of the "sites"), and also on the duration of the time intervals. Larger spatial grids and longer

interval durations make it less plausible that a site received no visits. We note, however, that inferring non-detections is a common strategy used for unstructured data (van Strien et al., 2010, 2013a; Kamp et al., 2016; Outhwaite et al., 2019; Powney et al., 2019), but it may be less justified to implement if we lack data on the actual observer or the date of the collection.

We also note that our conclusions about the suitability of the different models we have assessed should not be generalized beyond the parameters we have considered here, which we based on the empirical bumble bee data. For those examining other data sets, we believe that it is critical that researchers complement analyses of empirical data with analyses of simulated data using similar protocols to those that we provide, as validation of their methods. We expect our particular conclusions regarding which parameters most impact inferences may not hold for data sets with different underlying structure. However, here we present a general approach that can be used as a potential roadmap when applying occupancy models to historical museum records.

When dealing with opportunistic data and historical records, authors also have to either decide the spatial grain of sites or use the spatial grain of the observational data. For example, studies that use opportunistic data often use 1 km × 1 km grids (van Strien et al., 2010; Outhwaite et al., 2019; Powney et al., 2019; Outhwaite et al., 2020), as this is often the resolution of community science projects. Authors have also used a 5 km × 5 km resolution in order to better align with the resolution of other spatial variables (Woodcock et al., 2016) or because the data use encompassed historical museum records whose geo-reference was not accurate enough to warrant a finer resolution (van Strien et al., 2019). However, it is unclear how using a spatial grain as coarse as 100 km × 100 km might affect the trends observed. Here, we did not evaluate the effects of spatial or temporal resolution. Previous studies have found that increasing the spatial or temporal grain of the sites and time periods increases the estimates of occupancy (Hayes and Monfils, 2015; Steenweg et al., 2018). A recent empirical study suggests that varying the spatial grain from 1 km × 1 km to 10 km × 10 km and the temporal grain from 1 year to to a 10 year closure period has no effect on the estimated trends in occupancy (Jönsson et al., 2021).

Finally, we note that a critical component of the Soroye et al. (2020) study was to link changes in bumble bee occurrence at continental scales to changes in local climate. However, because Soroye et al. (2020) did not incorporate any climatic information into their occupancy models and, instead, looked for associations post-hoc, their inferences about the effects of climate depend on reliable species-specific trends. We have shown here that their parameterization may not produce reliable estimates of species-specific trends. Further, posthoc analyses of posterior means from single-species models, as implemented by Soroye et al. (2020), do not propagate uncertainty from species-level trends to community level inferences which, for these data, leads to under-estimates in uncertainty associated with community-level trends. Understanding how climate change is impacting bumble bees worldwide is of paramount importance, and we hope that our paper will provide an additional impetus for continued work on this topic.

5. Conclusions

Based on our findings here, we suggest that analyses of historical change in occupancy should incorporate species-specific ranges and information about visitation history to sites, as in the multi-species framework we applied here. If such visitation history data are not available, then we suggest modelling only time intervals where at least one species was detected. Although multi-species occupancy models are designed to explicitly incorporate changes in detection, we advocate for more work identifying best practices for combining historical observations into discrete intervals, so that the biases we find here can be avoided. While we provide revised estimates of changes in occupancy for bumble bees using multi-species models, refining these models to include environmental and species-level predictors will likely yield

further improved estimates. Including environmental and species-level predictors will tackle the open question whether climate change might be driving these changes collectively, or in a species-specific manner. Conservation practitioners have limited resources and so require accurate estimates of species' declines, so that those resources can be correctly allocated to the species that need them most.

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.biocon.2021.109141>.

Author statement

All authors conceived the study, LMG and LKM ran the analyses and wrote the initial draft of the manuscript, all authors edited the manuscript.

Declaration of competing interest

This manuscript has not been published and is not under consideration for publication elsewhere. We have no conflicts of interest to disclose.

Acknowledgements

We would like to thank Peter Soroye, Tim Newbold, and Jeremy Kerr for publishing such an extensive data set and a provocative finding, for providing us with assistance reproducing their original results, and for pointing us to IUCN Red List data as a possible validation test. We are grateful for in-depth feedback from Matt Pennell, and discussions with John Ascher, Luke Frishkoff, Hanna Jackson, Hutton Noth, Sally Otto, Lauren Ponisio, Elijah Reyes, Leif Richardson, and the SFU rEEDing Group. We acknowledge funding from Simon Fraser University (to LKM and LMG) and the Natural Sciences and Engineering Research Council of Canada (NSERC) (Discovery Grants to AOM and LKM and Vanier Scholarship to SAJ). This research was enabled in part by support from West Grid (www.westgrid.ca) and Compute Canada (www.compute.canada.ca).

References

- Altwegg, R., Nichols, J.D., 2019. Occupancy models for citizen-science data. *Methods Ecol. Evol.* 10, 8–21.
- Ascher, J.S., Marshall, L., Meiners, J., Yanega, D., Vereecken, N.J., 2020. Heterogeneity in large-scale databases and the role of climate change as a driver of bumble bee decline. (*Science E-Letter*, 2 May 2020) (URL). <https://science.sciencemag.org/content/367/6478/685/tab-e-letters>.
- Bartomeus, I., Ascher, J.S., Gibbs, J., Danforth, B.N., Wagner, D.L., Hedtke, S.M., Winfree, R., 2013. Historical changes in northeastern US bee pollinators related to shared ecological traits. *Proc. Natl. Acad. Sci. U. S. A.* 110, 4656–4660.
- Cameron, S.A., Sadd, B.M., 2020. Global trends in bumble bee health. *Annu. Rev. Entomol.* 65, 209–232.
- Cameron, S.A., Lozier, J.D., Strange, J.P., Koch, J.B., Cordes, N., Solter, L.F., Griswold, T. L., 2011. Patterns of widespread decline in North American bumble bees. *Proc. Natl. Acad. Sci. U. S. A.* 108, 662–667.
- Colla, S.R., Gadallah, F., Richardson, L., Wagner, D., Gall, L., 2012. Assessing declines of North American bumble bees (*Bombus spp.*) using museum specimens. *Biodivers. Conserv.* 21, 3585–3595.
- Crowther, L.P., Wright, D.J., Richardson, D.S., Carvell, C., Bourke, A.F., 2019. Spatial ecology of a range-expanding bumble bee pollinator. *Ecol. Evol.* 9, 986–997.
- de Keyser, C.W., Colla, S.R., Kent, C.F., Rafferty, N.E., Richardson, L.L., Thomson, J.D., 2016. Delving deeper: Questioning the decline of long-tongued bumble bees, long-tubed flowers and their mutualisms with climate change. *J. Pollinat. Ecol.* 18, 36–42.
- Dennis, E.B., Morgan, B.J., Freeman, M.S., Ridout, M.S., Brereton, T.M., Fox, R., Powney, G.D., Roy, D.B., 2017. Efficient occupancy model-fitting for extensive citizen-science data. *PLoS One* 12, e0174433.
- Devarajan, K., Morelli, T.L., Tenan, S., 2020. Multi-species occupancy models: review, roadmap, and recommendations. *Ecography* 43, 1612–1624.
- Frishkoff, L.O., Karp, D.S., M'Gonigle, L.K., Mendenhall, C.D., Zook, J., Kremen, C., Hadly, E.A., Daily, G.C., 2014. Loss of avian phylogenetic diversity in neotropical agricultural systems. *Science* 345, 1343–1346.
- Geldmann, J., Heilmann-Clausen, J., Holm, T.E., Levinsky, I., Markussen, B., Olsen, K., Rahbek, C., Tottrup, A.P., 2016. What determines spatial bias in citizen science? Exploring four recording schemes with different proficiency requirements. *Divers. Distrib.* 22, 1139–1149.
- Gibbons, D.W., Donald, P.F., Bauer, H.-G., Fornasari, L., Dawson, I.K., 2007. Mapping avian distributions: the evolution of bird atlases. *Bird Study* 54, 324–334.
- Goulson, D., 2010. Impacts of non-native bumblebees in Western Europe and North America. *Appl. Entomol. Zool.* 45, 7–12.
- Goulson, D., Lye, G.C., Darvill, B., 2008. Decline and conservation of bumble bees. *Annu. Rev. Entomol.* 53, 191–208.
- Graves, T.A., Janousek, W.M., Gaulke, S.M., Nicholas, A.C., Keinath, D.A., Bell, C.M., Cannings, S., Hatfield, R.G., Heron, J.M., Koch, J.B., Loffland, H.L., Richardson, L.L., Rohde, A.T., Rykken, J., Strange, J.P., Tronstad, L.M., Sheffield, C.S., 2020. Western bumble bee: declines in the continental United States and range-wide information gaps. *Ecosphere* 11, e03141.
- Hayes, D.B., Monfils, M.J., 2015. Occupancy modeling of bird point counts: implications of mobile animals. *J. Wildl. Manag.* 79, 1361–1368.
- Ikayan, K.J., Beissinger, S.R., 2018. Collapse of a desert bird community over the past century driven by climate change. *Proc. Natl. Acad. Sci. U. S. A.* 115, 8597–8602.
- Ikayan, K.J., Tingley, M.W., Furnas, B.J., Beissinger, S.R., 2014. Detecting diversity: emerging methods to estimate species diversity. *Trends Ecol. Evol.* 29, 97–106.
- Inoue, M.N., Yokoyama, J., Washitani, I., 2008. Displacement of Japanese native bumblebees by the recently introduced *Bombus terrestris* (L.) (Hymenoptera: Apidae). *J. Insect Conserv.* 12, 135–146.
- Isaac, N.J., van Strien, A.J., August, T.A., de Zeeuw, M.P., Roy, D.B., 2014. Statistics for citizen science: extracting signals of change from noisy ecological data. *Methods Ecol. Evol.* 5, 1052–1060.
- IUCN, 2020. The IUCN red list of threatened species. Version 2020–1. <https://www.iucnredlist.org>. Downloaded on 19 March 2020. URL. <https://www.iucnredlist.org>.
- Javorek, S.K., Mackenzie, K.E., Vander Kloet, S.P., 2002. Comparative pollination effectiveness among bees (Lymenoptera: Apoidea) on lowbush blueberry (*Ericaaceae: Vaccinium angustifolium*). *Ann. Entomol. Soc. Am.* 95, 345–351.
- Jeffers, D., 2017. The foraging behavior of a species of bumble bee (*Bombus pomorum*) that became extinct in the British Isles in the nineteenth century. *J. Apic. Res.* 56, 510–513.
- Jönsson, G.M., Broad, G.R., Sumner, S., Isaac, N.J., 2021. A century of social wasp occupancy trends from natural history collections: spatiotemporal resolutions have little effect on model performance. *Insect Conserv. Divers.* <https://doi.org/10.1111/icad.12494>.
- Kamp, J., Oppel, S., Heldbjerg, H., Nyegaard, T., Donald, P.F., 2016. Unstructured citizen science data fail to detect long-term population declines of common birds in Denmark. *Divers. Distrib.* 22, 1024–1035.
- Kerr, J.T., Pindar, A., Galpern, P., Packer, L., Potts, S.G., Roberts, S.M., Rasmont, P., Schweiger, O., Colla, S.R., Richardson, L.L., Wagner, D.L., Gall, L.F., Sikes, D.S., Pantoja, A., 2015. Climate change impacts on bumblebees converge across continents. *Science* 349, 177–180.
- Looney, C., Strange, J.P., Freeman, M., Jennings, D., 2019. The expanding Pacific Northwest range of *Bombus impatiens* Cresson and its establishment in Washington State. *Biol. Invasions* 21, 1879–1885.
- MacKenzie, D.I., Nichols, J.D., Lachman, G.B., Droege, S., Royle, J.A., Langtimm, C.A., 2002. Estimating site occupancy rates when detection probabilities are less than one. *Ecology* 83, 2248–2255.
- MacKenzie, D.I., Nichols, J.D., Royle, J.A., Pollock, K.H., Bailey, L.L., Hines, J.E., 2006. *Occupancy Estimation and Modeling: Inferring Patterns and Dynamics of Species Occurrence*. Elsevier, Burlington, Massachusetts, USA.
- McKinney, M.L., Lockwood, J.L., 1999. Biotic homogenization: a few winners replacing many losers in the next mass extinction. *Trends Ecol. Evol.* 14, 450–453.
- Miller-Struttman, N.E., Geib, J.C., Franklin, J.D., Kevan, P.G., Holdo, R.M., Ebert-May, D., Lynn, A.M., Kettenbach, J.A., Hedrick, E., Galen, C., 2015. Functional mismatch in a bumble bee pollination mutualism under climate change. *Science* 349, 1541–1544.
- Moritz, C., Patton, J.L., Conroy, C.J., Parra, J.L., White, G.C., Beissinger, S.R., 2008. Impact of a century of climate change on small-mammal communities in Yosemite National Park, USA. *Science* 322, 261–264.
- Outhwaite, C.L., Chandler, R.E., Powney, G.D., Collen, B., Gregory, R.D., Isaac, N.J., 2018. Prior specification in Bayesian occupancy modelling improves analysis of species occurrence data. *Ecol. Indic.* 93, 333–343.
- Outhwaite, C.L., Powney, G.D., August, T.A., Chandler, R.E., Rorke, S., Pescott, O.L., Harvey, M., Roy, H.E., Fox, R., Roy, D.B., et al., 2019. Annual estimates of occupancy for bryophytes, lichens and invertebrates in the UK, 1970–2015. *Sci. data* 6, 259.
- Outhwaite, C.L., Gregory, R.D., Chandler, R.E., Collen, B., Isaac, N.J., 2020. Complex long-term biodiversity change among invertebrates, bryophytes and lichens. *Nat. Ecol. Evol.* 4, 384–392.
- Owen, R.E., Otterstatter, M.C., Cartar, R.V., Farmer, A., Colla, S.R., O'Toole, N., 2012. Significant expansion of the distribution of the bumble bee *Bombus moderatus* (Hymenoptera: Apidae) in Alberta over 20 years. *Can. J. Zool.* 90, 133–138.
- Palmier, K.M., Sheffield, C.S., 2019. First records of the common eastern bumble bee, *Bombus impatiens* Cresson (Hymenoptera: Apidae, Apinae, Bombini) from the prairies Ecozone in Canada. *Biodivers. Data J.* 7, e30953.
- Pollock, L.J., Tingley, R., Morris, W.K., Golding, N., O'Hara, R.B., Parris, K.M., Veski, P.A., McCarthy, M.A., 2014. Understanding co-occurrence by modelling species simultaneously with a joint species distribution model (JSDM). *Methods Ecol. Evol.* 5, 397–406.
- Powney, G.D., Carvell, C., Edwards, M., Morris, R.K., Roy, H.E., Woodcock, B.A., Isaac, N.J., 2019. Widespread losses of pollinating insects in Britain. *Nat. Commun.* 10, 1018.
- Rasmont, P., Iserby, S., 2010–2014. *Atlas of the European Bees: Genus Bombus*, 3rd Edition. STEP Project, Atlas Hymenoptera, Mons, Gembloux <http://www.atla.shymenoptera.net>.

- Royle, J.A., Dorazio, R.M., 2008. Hierarchical modeling and inference in ecology. The analysis of data from populations, metapopulations and communities. Academic Press, London.
- Scholl, A., Obrecht, E., Owen, R.E., 1990. The genetic relationship between *Bombus moderatus* Cresson and the *Bombus lucorum* auct. species complex (Hymenoptera: Apidae). *Can. J. Zool.* 68, 2264–2268.
- Soroye, P., Newbold, T., Kerr, J., 2020. Climate change contributes to widespread declines among bumble bees across continents. *Science* 367, 685–688.
- Steenweg, R., Hebblewhite, M., Whittington, J., Lukacs, P., McKelvey, K., 2018. Sampling scales define occupancy and underlying occupancy–abundance relationships in animals. *Ecology* 99, 172–183.
- Szabo, J.K., Vesk, P.A., Baxter, P.W., Possingham, H.P., 2010. Regional avian species declines estimated from volunteer-collected long-term data using List Length Analysis. *Ecol. Appl.* 20, 2157–2169.
- The Xerces Society, Wildlife Preservation Canada, York University, University of Ottawa, The Montreal Insectarium, The London Natural History Museum, BeeSpotter., 2017. Bumble Bee Watch, a collaborative website to track and conserve North America's bumble bees. In: Available from. <http://www.bumblebeewatch.org/app/#/bees/lists>.
- Tingley, M.W., Beissinger, S.R., 2009. Detecting range shifts from historical species occurrences: new perspectives on old data. *Trends Ecol. Evol.* 24, 625–633.
- Tingley, M.W., Beissinger, S.R., 2013. Cryptic loss of montane avian richness and high community turnover over 100 years. *Ecology* 94, 598–609.
- Tobler, M.W., Kéry, M., Hui, F.K.C., Guillera-Aroita, G., Knaus, P., Sattler, T., 2019. Joint species distribution models with species correlations and imperfect detection. *Ecology* 100, e02754.
- Torretta, J.P., Medan, D., Abrahamovich, A.H., 2006. First record of the invasive bumblebee *Bombus terrestris* (L.)(Hymenoptera, Apidae) in Argentina. *Trans. Am. Entomol. Soc.* 132, 285–289.
- van Strien, A.J., Termaat, T., Groenendijk, D., Mensing, V., Kéry, M., 2010. Site-occupancy models may offer new opportunities for dragonfly monitoring based on daily species lists. *Basic Appl. Ecol.* 11, 495–503.
- van Strien, A.J., van Swaay, C.A., Termaat, T., 2013a. Opportunistic citizen science data of animal species produce reliable estimates of distribution trends if analysed with occupancy models. *J. Appl. Ecol.* 50, 1450–1458.
- van Strien, A.J., Termaat, T., Kalkman, V., Prins, M., De Knijf, G., Gourmand, A.-L., Houard, X., Nelson, B., Plate, C., Prentice, S., et al., 2013b. Occupancy modelling as a new approach to assess supranational trends using opportunistic data: a pilot study for the damselfly *Calopteryx splendens*. *Biodivers. Conserv.* 22, 673–686.
- van Strien, A.J., van Swaay, C.A., van Strien-van Liempt, W.T.F.H., Poot, M.J., WallisDeVries, M.F., 2019. Over a century of data reveal more than 80% decline in butterflies in the Netherlands. *Biol. Conserv.* 234, 116–122.
- Williams, N.M., 2011. Restoration of nontarget species: bee communities and pollination function in riparian forests. *Restor. Ecol.* 19, 450–459.
- Williams, P.H., Brown, M.J.F., Carolan, J.C., An, J., Goulson, D., Aytekin, A.M., Best, L.R., Byvaltsev, A.M., Cederberg, B., Dawson, R., Huang, J., Ito, M., Monfared, A., Raina, R.H., Schmid-Hempel, P., Sheffield, C.S., Šima, P., Xie, Z., 2012. Unveiling cryptic species of the bumblebee subgenus *Bombus s. str.* worldwide with coi barcodes (Hymenoptera: Apidae). *Syst. Biodivers.* 10, 21–56.
- Williams, P.H., Thorp, R.W., Richardson, L.L., Colla, S.R., 2014. Bumble Bees of North America: An Identification Guide. Princeton University Press, Princeton, New Jersey.
- Woodcock, B.A., Isaac, N.J., Bullock, J.M., Roy, D.B., Garthwaite, D.G., Crowe, A., Pywell, R.F., 2016. Impacts of neonicotinoid use on long-term population changes in wild bees in England. *Nat. Commun.* 7, 12459.
- Zipkin, E.F., DeWan, A., Royle, A.J., 2009. Impacts of forest fragmentation on species richness: a hierarchical approach to community modelling. *J. Appl. Ecol.* 46, 815–822.