

CONTRIBUTED PAPER

Mapping bee diversity with landscape-level models to inform conservation

Mark A. Buckner¹  | Erin L. White² | Timothy G. Howard² |
Matthew D. Schlesinger² | Bryan N. Danforth¹

¹Department of Entomology, Cornell University, Ithaca, New York, USA

²New York Natural Heritage Program, State University of New York College of Environmental Science and Forestry, Albany, New York, USA

Correspondence

Mark A. Buckner, Department of Entomology, Cornell University, Ithaca, NY 14853, USA.

Email: mab677@cornell.edu

Present address

Mark A. Buckner, Department of Entomology, Pennsylvania State University, University Park, Pennsylvania, USA.

Funding information

Cornell Atkinson Center for Sustainability, Cornell University, Sustainable Biodiversity Fund

Abstract

Anthropogenic drivers of global change threaten bee diversity and the pollination services they provide. Despite the importance of bees, their conservation is complicated by limited and often heavily biased occurrence data. A recent state-wide survey of insect pollinators across New York, United States generated a large spatial dataset of bee species occurrence records from community scientists, historical collections, and recent surveys. To predict the distributions of most of the state's bee species, we combined the state survey records with occurrence data from across the contiguous United States before applying an ensemble modeling approach consisting of balanced random forest and small bivariate generalized linear models. We predicted the spatial distribution of bee species richness using a stacked species distribution model with climate, land cover, and soil covariates. To inform bee diversity conservation, we predicted spatial variation for each species and groups of species sharing similar life history traits. We also estimated the statewide distribution of range-size rarity, ecological uniqueness, and climate exposure. We found that the richness of modeled species is high across the state, with the greatest richness in regions with low soil clay content and intermediate forest cover. The fine spatial scale and extent of our gridded data layers match the scale of conservation action in the state, providing an opportunity to incorporate wild bee diversity into broader statewide conservation planning. Conserving bee pollinators is not straightforward, and decisions should be based on broader conservation priorities that incorporate bee biodiversity indicators into decision-making. Here, we present a roadmap for the inclusion of these vital pollinators in conservation decisions by leveraging the best available data and methods robust to small sample sizes to provide spatially explicit data products representing the distribution of bee diversity.

KEYWORDS

bee decline, bee diversity, biodiversity, climate change, pollinator conservation, protected areas, species distribution models

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2026 The Author(s). *Conservation Science and Practice* published by Wiley Periodicals LLC on behalf of Society for Conservation Biology.

1 | INTRODUCTION

The broad-scale loss of biodiversity across taxonomic groups is a global phenomenon of significant human and ecological concern (Sage, 2020). The threat of biodiversity collapse has highlighted the need to protect economically and ecologically important species including vital insect pollinators such as bees (Hall & Steiner, 2019). Public and political concern regarding bee decline came to the forefront with the onset of colony collapse disorder in the mid-2000s (Althaus et al., 2021; van Engelsdorp et al., 2009). Consequently, pollinator protection initiatives in the United States have generally focused on *Apis mellifera*, the managed western honeybee (Colla & MacIvor, 2017; Iwasaki & Hogendoorn, 2021). While such initiatives have bolstered public interest in addressing pollinator declines, misguided or even counterproductive calls for honeybee conservation disregard the loss of native bees (Geldmann & González-Varo, 2018). Yet, our awareness of humanity's role in unmanaged wild bee declines is not new (Carson, 1962; Frison, 1926; Williams, 1982), nor is our understanding of the consequences of losing wild pollinators, with the loss of wild bee-mediated pollination services noted since at least the mid-1900s (Carson, 1962; Fronk & Painter, 1960; Stephen, 1955). To date, pesticide use, pests and pathogens, habitat loss, introduced species, and climate change, among others, have all been implicated in global bee diversity declines (Dicks et al., 2021; Goulson et al., 2015; Potts et al., 2010; Vanbergen et al., 2013).

Considering the growing need to address bee declines, national or regional survey efforts may prove vital for establishing conservation baselines, formulating priorities, and enabling status assessments (Woodard et al., 2020). Although the utility of survey efforts may be limited in the near term (see Tepedino & Portman, 2021), they present a clear opportunity to address historical data limitations that have hindered native bee conservation efforts (Inouye et al., 2017). While global analyses of bee diversity exist (Bystrakova et al., 2018; Orr et al., 2021), accounting for regional context through national or sub-national assessments could better inform conservation action (Schmidt-Traub, 2021). For example, there is growing evidence of the importance of forests for pollinator diversity (Ulyshen et al., 2023), despite global bee diversity hotspots primarily occurring in midlatitude arid deserts (Bystrakova et al., 2018; Orr et al., 2021), such as the deserts of the American southwest which are home to many of the 4000 species of North American bees (Ascher & Pickering, 2020; Carril et al., 2018).

In 2021, the New York Natural Heritage Program concluded the Empire State Native Pollinator Survey (ESNPS), a four-year statewide survey focused on

evaluating the conservation status of New York State's (NYS) insect pollinators. The effort resulted in the aggregation of numerous biodiversity datasets, community science contributions, and surveys to determine the status of NYS bee, fly, beetle, and moth pollinators. To evaluate conservation statuses of the state's pollinators, the survey relied on coarse distribution trends, rarity, and expert assessment of the threats and vulnerabilities each species faces (Schlesinger et al., 2023; White et al., 2022). The effort highlighted the need for the inclusion of insect pollinators in statewide conservation planning, with 60% of native insect pollinator species ranked as vulnerable or imperiled (White et al., 2022).

Irrespective of the concerning trends, pollinators are often excluded from conservation action due in part to ineffective policy in many jurisdictions (Hall & Steiner, 2019), inadequate funding (Cardoso et al., 2011), and imperfect data (Rousseau et al., 2024). The decline of bee species jeopardizes essential pollination services, suggesting that even with the significant gaps in knowledge, conservation action is warranted to avert further declines (Drivdal & van der Sluijs, 2021). Carefully designed species distribution models (SDM) are one avenue to overcome current data limitations. Co-production of SDMs through partnerships between conservation practitioners, researchers, and other interested parties to create spatial data products for use in conservation decision making can provide essential information for directing management decisions (Beier et al., 2017; Cunningham-Minnick et al., 2024).

Here we integrate the ESNPS bee locality dataset with collection records from across the contiguous United States to model the distribution of over half of the state's bee species and generate landscape-level maps of bee diversity. We capitalize on the significant contemporary, spatially accurate records of species localities generated by recent surveys like the ESNPS which we pair with modern SDMs robust to the small sample sizes and presence-only data inherent to bee biodiversity datasets. To generate robust predictions for NYS's bee pollinators, we employ ensemble models, with balanced random forest and small bivariate generalized linear model constituents. We predict the distribution of species richness across the state, evaluate the contribution of landscape and climate variables to bee diversity, and identify regions that may be of high conservation value, including those that are important for range-restricted species and represent unique community compositions. In addition, we evaluate the threats posed by climate change to identify the bee communities experiencing the greatest pressure for climate-associated redistribution. Our approach provides a roadmap for augmenting conservation assessments following state-wide pollinator survey efforts and

presents gridded data products to facilitate the consideration of wild bee pollinators in conservation planning.

2 | MATERIALS AND METHODS

The model for NYS bee species involved five primary stages: occurrence data processing, environmental covariate acquisition and processing, background point generation and bias correction, species-wise ensemble models and stacking, and the calculation of biodiversity indicators. Figure 1 provides an overview of the complete modeling and analysis workflow including all major steps and significant methods.

New York is a geographically diverse state, encompassing a wide range of land uses and bee habitats. The southeastern portion of the state is densely populated, with urban centers including New York City and urban and suburban areas extending northward through the Hudson Valley toward Albany and along Long Island. The central and western portions of the state are a mix of agriculture and deciduous forest. Forest cover is greater in the Allegheny Plateau of the southwestern portion of the state compared to the principally agricultural Finger Lakes region. Elevation across the state is variable with peaks above 1600 m in the heavily forested Adirondacks Mountains to the northeast and 1250 m in the Catskills in the south.

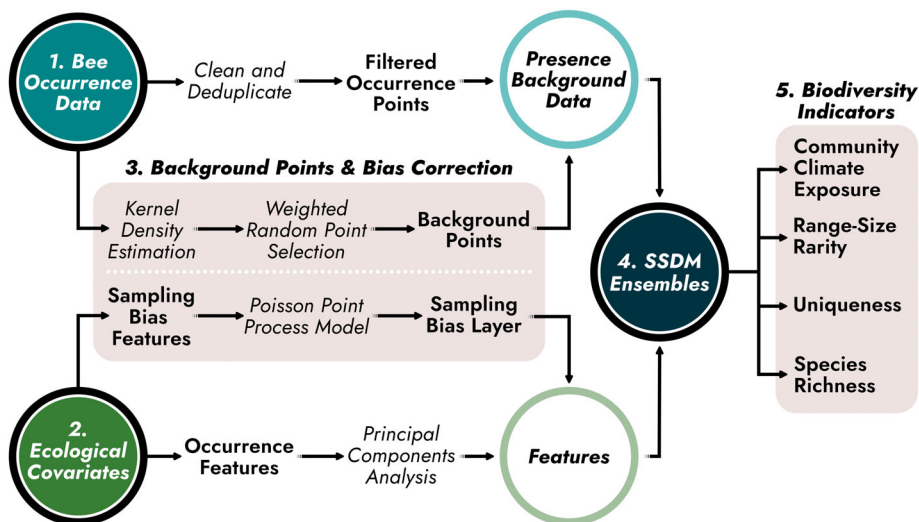
2.1 | Bee occurrence data

The New York Natural Heritage Program led a 4-year, statewide survey of native insect pollinators, which combined community science with targeted and extensive

surveys. The ESNPS dataset consists of presence-only records across the state representing bee collections stratified across all ecoregions and many unique habitats (e.g., alpine meadows, peat bogs) (Schlesinger et al., 2023). The database provides an extensive collection of records for both target species (those which were thought to be at risk) and non-target species across the state from curated datasets (e.g., Richardson, 2024), natural history museums, and community science observations (e.g., iNaturalist). White et al. (2022) and Schlesinger et al. (2023) presented a thorough overview of the data collection and integration process used in the development of the ESNPS dataset. To better represent the full distribution of environmental conditions NYS bee species occupy, we supplemented the ESNPS dataset with records from across the United States obtained from the Global Biodiversity Information Facility (GBIF; GBIF.org, 2024). We performed all analyses in R v4.1.2 (R Core Team, 2021).

We checked for georeferencing errors with CoordinateCleaner v3.0.1 (Zizka et al., 2019) by flagging coordinates that correspond to country or state centroids and biodiversity institutions. We chose to retain records from botanical gardens, museums, and university campuses as we found that the audited records corresponded to collecting events occurring on institution campuses. In addition, we removed any records within 100 m of county centroids unless they included additional locality information (e.g., Beebe Lake, Tompkins County, NY). We chose to retain records for species that have not authoritatively been identified in NYS as these may constitute early records or range expansions. Finally, we filtered the dataset to retain observations between 1971 and 2023 for species that, after removing duplicate observations, had a minimum of 10 spatially unique records.

FIGURE 1 Workflow for ensemble Stacked Species Distribution Modeling (SSDM) of New York bee species using presence-only records from the Empire State Native Pollinator Survey and GBIF. Analytical stages are numbered, specific methods are italicized, and outputs are shown in bold.



2.2 | Environmental covariates

We considered 13 environmental covariates (Table 1) that are thought to influence bee species distributions (Table S1, Supporting Information). To ensure the greatest possible temporal overlap between the climate data and observation records, we selected climate covariates derived from the 1981–2010 climate normals available in CHELSA v2.1 (Karger et al., 2017, 2021). Besides climate, we include continuous regionally relevant habitat covariates such as percent forest canopy cover (LANDFIRE, 2014) and soil texture data (Hengl, 2018a, 2018b) and derived two additional layers representing gridded terrain roughness as the standard deviation of the slope (Grohmann et al., 2011; LANDFIRE, 2020) and aridity index calculated as the ratio of annual precipitation to potential evapotranspiration (United Nations Environment Programme, 1992). We chose to exclude categorical land use data due to the small number of records for many species and to reduce the computational complexity of model fitting.

The chosen grain for SDMs should reflect a compromise between the ecological scale at which the species of interest respond to their environment and the constraints imposed by the spatial uncertainty of the occurrence data (Gábor et al., 2022). The scale at which bees interact with their environment corresponds closely with body size and sociality (Greenleaf et al., 2007; Kendall et al., 2022). Most bees forage within less than 1 km of their nest, with large, social species often foraging between 1 and 2 km (Kendall et al., 2022). Since bee records often represent foraging individuals, we accounted for the environmental

conditions thought to influence foraging and nesting habitat by aggregating each gridded dataset to approximately 990-m resolution and projected the data layers to an Albers equal area projection. To ensure that records from wetlands were not excluded erroneously for occurring within waterbodies, we masked grid cells when 90% or more of the 30-m resolution vegetation layer cells that occurred within each of the final 990 m aggregated grid cells were classified as water (LANDFIRE, 2014). Where required, we imputed the soil data surrounding water bodies using the mean of a three-by-three focal window in terra v1.7-65 (Hijmans, 2023).

Finally, we conducted principal components analysis (PCA) to mitigate the effects of collinearity and reduce the dimensionality of the environmental variables. To limit the number of constituent models, thereby reducing computational cost, we retained the minimum number of PCA components required to explain 80% of the variance in the dataset as features for model training.

2.3 | Background points and bias correction

We used two bias correction methods to address spatial sampling bias in the opportunistic collection records. First, we calculated a bias layer with kernel density estimation using a 20-km bandwidth across the United States. From this bias layer, we selected 50,000 weighted random background points with similar distribution and spatial bias to the presence records (Inman et al., 2021; Kujala et al., 2015; Valavi et al., 2021). Due to

TABLE 1 List of environmental covariates included in species distribution models prior to dimensionality reduction with a principal components analysis (PCA) and the sampling intensity point process model.

Species distribution models		Sampling intensity model	
<i>bio1</i>	Mean annual air temperature	<i>pop</i>	Population density
<i>bio7</i>	Annual range of air temperature	<i>pad</i>	Protected areas
<i>bio11</i>	Mean daily minimum air temperature of the coldest quarter	<i>inst</i>	Distance to biodiversity institutions
<i>bio12</i>	Annual precipitation amount	<i>road</i>	Distance to roadways
<i>bio15</i>	Precipitation seasonality	<i>mil</i>	Fort Drum boundary
<i>ai</i>	Aridity index	<i>x</i>	Longitude
<i>ngd10</i>	Number of growing degree days above 10°C	<i>y</i>	Latitude
<i>gsp</i>	Accumulated precipitation on growing season days		
<i>ssc</i>	Soil sand content		
<i>scc</i>	Soil clay content		
<i>rgh</i>	Terrain roughness (SD of slope)		
<i>cc</i>	Canopy cover		
<i>int</i>	Predicted sampling intensity		

the variability of sampling effort, we used an additional model-based bias correction step where we modeled sampling intensity with a spatial nonstationary Poisson point process model in spatstat v3.0-6 (Baddeley et al., 2015). Similar model-based methods have been shown to improve predictive performance and have been widely used outside of distribution modeling (Warton et al., 2013). We chose covariates that represent where researchers and community scientists often search for bees, including population density (Center for International Earth Science Information Network-CIESIN-Columbia University, 2018), protected area status (USGS GAP, 2022), distance to primary and secondary roadways (U.S. Census Bureau, 2013), distance to biodiversity institutions (Zizka et al., 2019), and whether the grid cell is part of the Fort Drum military base (U.S. Census Bureau, 2010)—a site with additional high-intensity sampling during the ESNPS. In addition, we included polynomial terms for the spatial coordinates and interactions between the distance variables and protected area status. We included the gridded intensity layer as a covariate in the SDMs alongside the principal components which encode the environmental covariates. When predicting across space, we set the intensity layer equal to its global mean such that distribution predictions assume all cells had identical sampling intensity.

2.4 | Stacked species distribution models

Due to the small number of records available for many bee species in New York, following conventional recommendations to avoid overfitting (e.g., a minimum of 10 presence-only records per covariate) would necessitate fitting unrealistically simple models (Brun et al., 2020). To overcome this limitation, we integrated additional records from across the United States and trained separate ensemble models for each species. Each ensemble consisted of a balanced random forest (BRF) model and a collection of small bivariate generalized linear models, referred to as an Ensemble of Small Models (ESM). Ensemble models have been found to perform adequately with presence-only datasets, with BRF and regression methods both performing well with small sample sizes and comparably well to other model types when large numbers of presence points are available (Valavi et al., 2022).

Both constituent model types, and ensemble models in general, have shown promise for modeling rare species (Breiner et al., 2015; Erickson & Smith, 2023; Lomba et al., 2010; Mi et al., 2017; Valavi et al., 2022). Relative to conventional random forest models, BRF performs well when applied to presence-only data with small numbers

of presence localities relative to background points (Valavi et al., 2022). The improvement in model performance when learning imbalanced data with BRF is facilitated by down-sampling the majority class (background points) to match the minority class (presence points) in each tree (Chen & Breiman, 2004). We trained the BRF models as probability forests in ranger v0.15.1 with 1000 trees and all other parameters set to their defaults (Wright & Ziegler, 2017). Additionally, we fit the ESMs as downweighed bivariate generalized linear models for each pairwise combination of covariates (Erickson & Smith, 2023; Lomba et al., 2010). We trained two sets of ESMs, one with linear terms only and one with both linear and polynomial terms. While including polynomial terms increases the computational cost of model fitting significantly, they may perform better than linear-only ESMs for discrimination tasks (Erickson & Smith, 2023).

We estimated ensemble model performance with the cross-validated area under the receiver operating characteristic curve (AUC_{ROC}) and Continuous Boyce Index (CBI; Hirzel et al., 2006) for each model. For species with greater than 50 records, we used 5-fold spatial block cross-validation, and for species with fewer than 50 records, we used nearest neighbor distance matching, a modified version of leave-one-out cross-validation (Milà et al., 2022; Valavi et al., 2019). Occasionally spatial cross-validation returned invalid folds, where one or more folds contained only a single test point. In such cases, we applied buffered leave-one-out cross-validation or, if all spatial methods failed to return valid folds, stratified k-folds implemented with caret v6.0-94 (Kuhn, 2008). All spatial cross-validation methods were performed with blockCV v3.1-3 (Valavi et al., 2019). We generated folds for each species independently and used identical folds across all constituent models.

To assemble the final ensembled predictions, we took the weighted average of each constituent model with Somers' D where $D = 2 * (AUC - 0.5)$, and excluded models where $D < 0$ (Breiner et al., 2015). We calculated the final cross validated AUC_{ROC} , and CBI for each species based on the same weighting scheme as the ensemble predictions. We then stacked the predicted distributions of each species by taking the sum of the occurrence probabilities at each grid cell (Calabrese et al., 2014). Approximating species richness in this way allowed us to estimate the variance in site-level richness by assuming richness follows a Poisson Binomial distribution; however, we were unable to include uncertainty from the occurrence probabilities themselves which are assumed fixed in this stacking scheme (Calabrese et al., 2014).

We subset the final stacked models using life-history traits and taxonomic groupings aggregated during the

ESNPS and through expert knowledge (Danforth, 2024) representing host plant specialization, sociality, native and non-native species, nesting behavior, and apple pollinators (a group of agriculturally important species). To check for differences in species distributions for each life-history group, we subdivided species richness predictions and used Pearson correlation coefficients to assess the degree to which the life-history groups differed from the NYS bee community as a whole.

2.5 | Other biodiversity indicators

We identified areas that may be of high conservation value for NYS bees based on species richness, ecological uniqueness (the contribution of cell-level community composition to overall beta diversity; Legendre & De Cáceres, 2013), and range-size rarity (RSR; the importance of each grid cell for range-restricted species; Guerin & Lowe, 2015). Where possible, we computed these metrics using probabilistic distribution predictions instead of binary threshold predictions, since applying a threshold does not improve habitat suitability estimates and may lead to over-prediction if performed before stacking (Calabrese et al., 2014; Guillera-Aroita et al., 2015). However, to estimate community composition, we applied probability ranking rules to threshold the probability predictions, resulting in presence-absence maps for each species. Probability ranking rules are a community-based method to threshold SDM predictions where species are marked present in each grid cell up to the predicted species richness based on their relative occurrence probabilities within each grid cell (D'Amen et al., 2015). To evaluate the potential influence of habitat characteristics on bee diversity, we calculated response curves for species richness across the principal components and Pearson's correlation coefficients between predicted richness and the habitat covariates.

We mapped ecological uniqueness in Julia v1.9.4 as the cell-wise local component of beta diversity, where more unique community compositions were found in cells with greater relative contribution to statewide beta diversity (Dansereau et al., 2022; Legendre & De Cáceres, 2013). To estimate the similarity in community composition between the most and least unique 2.5% of grid cells we found the proportion of the total modeled species that our model predicted to occur in both these areas. In addition, we calculated RSR to identify cell-level importance for range-restricted species (Guerin & Lowe, 2015). We calculated RSR using the continuous occurrence probability predictions, where p_{ij} represents the predicted occurrence probability for species j in grid cell i . We first estimated each species' occurrence area as the product of the area of each grid cell (A_{cell}) and the

sum of the predicted occurrence probabilities across all m cells (Stark & Fridley, 2022). Then, to estimate the cell-wise importance of each grid cell to range-restricted species, we took the sum across all species' occurrence probabilities (p) within a cell divided by their occurrence areas such that the value of RSR increases in grid cells which are likely of greater importance to range-restricted species (Equation (1)),

$$\text{RSR}_i = \sum_{j=1}^n \frac{p_{ij}}{A_{\text{cell}} * \sum_{k=1}^m p_{kj}} \quad (1)$$

To assess the potential effects of climate change, we calculated the distance-based velocity of climate change (VoCC) between the 1981–2010 and 2071–2100 climate normals across NYS as the mean of four General Circulation Models (GFDL-ESM4, MPI-ESM1-2-HR, IPSL-CM6A-LR, MRI-ESM2-0) for two shared socioeconomic pathways representing sustainable (SSP1-2.6) and fossil-fueled (SSP5-8.5) development scenarios from Coupled Model Intercomparison Project Phase 6 (CMIP6). We applied a 250-km buffer zone around NYS to ensure that analogous climates were identifiable within a reasonable dispersal range. We used multivariate VoCC based on three climate variables: mean annual air temperature, annual precipitation, and growing season precipitation following the multivariate method described by Hamann et al. (2015). To calculate species-specific climate exposures, we used the mean cell-wise climate velocity across NYS weighted by occurrence probability for each species. Next, we identified the regions of the state that have the greatest community-wide risk from climate change as Equation (2),

$$\text{CCE}_j = \frac{\sum_{i=1}^n I(x_i > \text{median}(\mathbf{x})) * t_{ij}}{n_j} \quad (2)$$

where community climate exposure (CCE) at the j th site is equal to the proportion of the total species (n), which are present ($t_{ij} = 1$) in the threshold predictions ($\mathbf{T}$) that have a ratio of climate exposure to range size ($\mathbf{x}$) above the median ratio across all species.

To evaluate the utility of permanently protected areas for bee conservation, we assessed how bee species richness, RSR, ecological uniqueness, and climate exposure within protected areas deviated from the statewide mean. The United States Geological Survey protected area inventory assigns protected areas a status code depending on the permanency of protection and allowable uses (USGS GAP, 2022). GAP status 1 and 2 are under mandates to ensure permanent protections and differ from GAP status 3 areas that are permanently protected but allow extractive uses (e.g., logging or mining) (USGS

GAP, 2022). For this analysis, we focused on protected areas with GAP statuses 1, 2, and 3 as delineated by the New York Protected Area Database (New York Natural Heritage Program, 2024).

3 | RESULTS

3.1 | Occurrence and environmental data

Our complete bee occurrence dataset contained over 213,000 unique records representing 269 bee species. The number of presence-only records varied between a minimum of 11 for *Osmia feldti* and over 35,000 records for the common eastern bumble bee, *Bombus impatiens*. Approximately 21% of the included species had 50 or fewer unique records across the United States during the 52-year data period (1971–2023) and 79% of all observations were recorded after 2010. The occurrence records showed clustering in and around urban areas and roadways suggesting significant spatial sampling bias (Figure S1). While many species had insufficient data to include in the models, we noted a similar distribution of species' life history traits in the modeled species as compared with all known species from NYS, although solitary bees were overrepresented in our dataset (Figure S2). The proportion of modeled species classified as host plant specialists was comparable to that of all known species from NYS, 23% and 21%, respectively.

Following dimensionality reduction with a PCA, the first four principal components accounted for 80.6% of the total variance in our environmental layers and were retained as features for the species distribution models (Table S2). Principal component (PC) 1 primarily measured temperature, showing strong positive associations with climate variables representing annual and winter temperatures and a negative association with annual temperature range. PC2 showed positive associations with aridity index and percent forest canopy cover. PC3 primarily represented soil texture and PC4 was dominated by growing season precipitation. Predicted sampling intensity, from the point process model, was highest near roadways and biodiversity institutions with notable peaks in intensity along roadways near institutions, in Fort Drum, and near New York City (Figure S3).

3.2 | Stacked species distribution models

We used spatial cross-validation to evaluate model performance for 97.8% of the included species. The linear-

only models performed slightly worse on average (mean AUC_{ROC} : 0.72, range: 0.51–0.96; mean CBI: 0.77, range: $-0.77 \sim 1$) than the models that included linear and polynomial terms (mean AUC_{ROC} : 0.74, range: 0.50–0.96; CBI: 0.82, range: $-0.62 \sim 1$). Models of ESNPS target species performed better on average (AUC_{ROC} : 0.75; CBI: 0.87) than non-target species (AUC_{ROC} : 0.72; CBI: 0.78). We were unable to fit satisfactory models for two species, *Andrena heraclei* and *Epeolus autumnalis*, when using linear terms only (i.e., all constituent models had an $AUC_{ROC} < 0.5$), and all models failed to reach a satisfactory AUC_{ROC} for two additional species, *Apis mellifera* and *Macropis nuda*. AUC_{ROC} and CBI showed moderate positive correlation ($r = 0.56$) with CBI identifying nine models with performance worse than random (CBI < 0 ; Hirzel et al., 2006), only three of which were target species in the ESNPS (Table S3). Models of *Protandrena andrenoides* were fit without all available GBIF records due to an unidentified taxonomic conflict between the datasets. Species that have not previously been authoritatively identified within NYS ($n = 8$) showed low predicted suitability within the state.

The predicted richness of the modeled bee species across the state was relatively homogeneous, varying between 75 and 150 species (interquartile range: 13.3 species) with a left-skewed distribution (Figure 2a). Bee richness was high throughout the state, peaking in the Adirondack foothills of northern NYS. In the remainder of the state, richness was moderately positively correlated with percent forest canopy cover ($r = 0.67$) and negatively correlated with soil clay content ($r = -0.75$). Correspondingly, we found that species richness peaked at intermediate values of PC3, dominantly representing soil texture, and intermediate values of PC2, representing forest canopy cover and aridity (Figure S4). Our models predicted low richness in the heavily agricultural Finger Lakes Region, with the lowest predicted richness occurring in the clay rich lowlands west of the Adirondacks and the “Black Dirt” region near the Hudson Valley in southern NYS. Long Island had moderately low predicted richness relative to nearby inland areas such as the Hudson Valley Ecoregion and surrounding foothills. The variance in the stacked models was even across the state with minimal variation in uncertainty from the stacking process (Figure 2b). We subset the overall richness predictions into subgroups by taxonomy or life history, including *Bombus* spp., host plant specialization, sociality, native and non-native species, nesting behavior, and apple pollinators. Except for non-native ($r = -0.09$), eusocial ($r = 0.63$), and *Bombus* species ($r = 0.69$), the spatial distribution of each subset showed a strong positive correlation with the prediction for all species ($r > 0.95$).

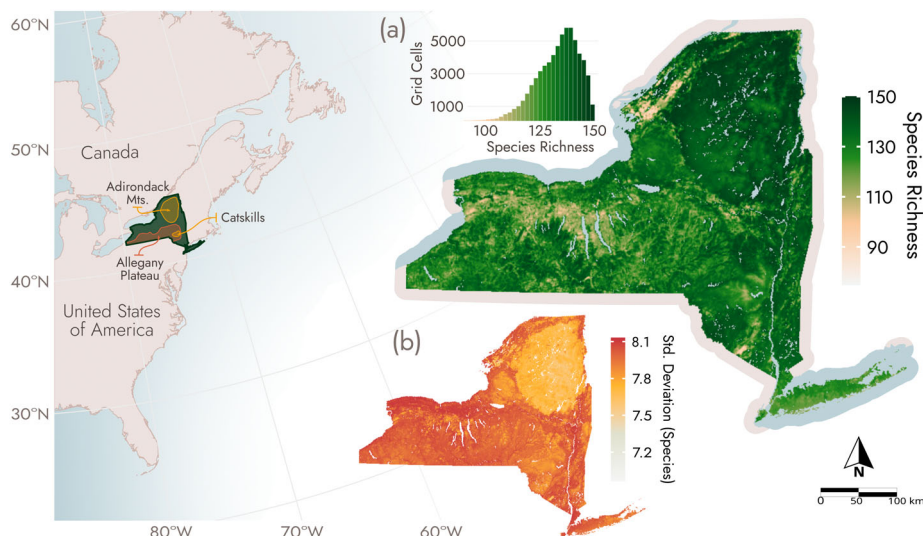


FIGURE 2 (a) Predicted bee species richness is relatively high across much of New York State with low species richness in agricultural areas with high clay soil content. The histogram shows the distribution of richness values. (b) Cell-level standard deviations of species richness across the state representing uncertainty from the model stacking process.

3.3 | Other biodiversity indicators

We found that the spatial distribution of RSR was similar to species richness (Figure 3a). In contrast, community compositions showed the inverse relationship, where uniqueness was highest in regions with low richness, including the Finger Lakes and coastal Long Island (Figure 3b). Uniqueness was also higher in some urban areas, including Albany, though interestingly it was not comparably high in the ecologically distinctive Albany Pine Bush Preserve. Overall, all bee communities showed a strong overlap with 63% of species shared between cells with the most and least unique community compositions.

Bee communities found at relatively high elevations, particularly within the Adirondacks, Catskills, and Allegheny Plateau, face high climate exposures with lower exposures in the Finger Lakes Region and lowlands south of Lake Ontario (Figure 4b). Overwhelmingly, we expect most bees across the state to track their climate niche northward (Figure 4a). The Adirondack foothills, however, present a clear exception to this pattern where climatic niches displace upward in elevation, even if this means a southward trajectory. This pattern breaks down under the fossil-fueled development scenario (SSP5-8.5) where 24% of the state, including the Adirondacks, has no climate analog within the 250-km buffer zone, compared to less than 0.1% under SSP1-2.6 (Figure 4c). A sizable portion of NYS's protected areas is within the Adirondacks, a higher elevation, mountainous ecosystem with expansive deciduous forest. The Adirondacks region contributed to higher CCEs within protected areas which also had greater richness and RSR, and lower uniqueness compared to the statewide average (Figure 5).

4 | DISCUSSION

The conservation of wild bees faces challenges from deficient occurrence data and limited knowledge about bee ecology. Through training SDMs optimized for rare species and robust to spatial sampling bias, we generated fine-scale predictions of bee richness, community uniqueness, climate exposure, and individual species distributions for 60% of NYS's bee species. The principal results of this study are our model predictions, which we provide as freely available gridded datasets at 990-m resolution (Buckner et al., 2025a). Our model indicated that bee richness was consistently high across NYS and associated negatively with clay soils and positively with percent forest canopy cover but with highest richness at values less than maximum canopy coverage (Figure S4). These patterns are consistent with habitat preference expectations for ground-nesting species which are thought to prefer exposed sandy soils and forest-associated species that show a preference for intermediate-density stands containing greater floral diversity and nest site availability (Antoine & Forrest, 2021; Ulyshen et al., 2024). Conversely, bee communities in regions with high soil clay content and those inhabiting Long Island's coastal ecosystems were among the most unique in the state. Statewide, climate change may contribute to the northward redistribution of bee ranges. A clear exception to this pattern is in the Adirondacks, where climate trajectories point toward movements upward in elevation and no analogous climate within 250 km under SSP5-8.5, indicating a potential climate trap.

The patterns of bee diversity we observed differ from the general expectations presented in global analyses of bee diversity and biogeography, which suggest that bees are most diverse in arid, seasonal habitats (Almeida

FIGURE 3 (a) Range-size rarity is lowest in the regions of New York with low bee richness. (b) Ecological uniqueness of bee community predictions is higher in areas with low predicted species richness and across Long Island. Both metrics are multiplied by 10,000 for visualization.

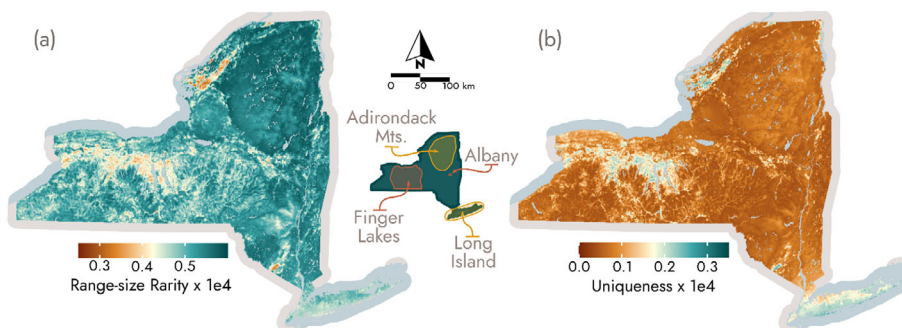
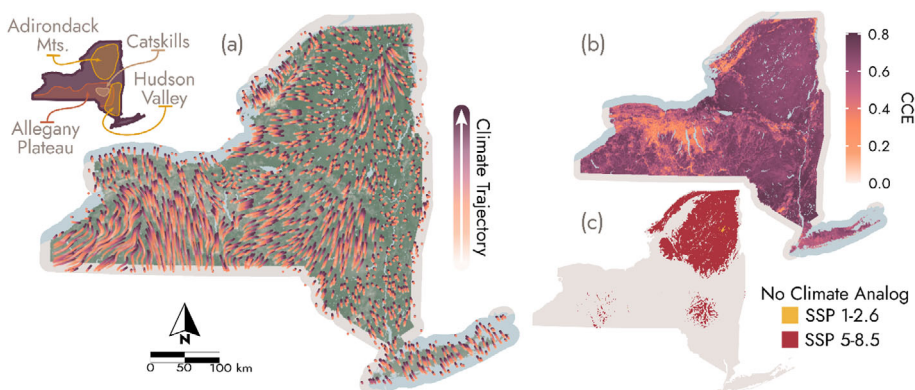


FIGURE 4 (a) Climate trajectories indicate that climate niches are expected to predominantly move northward under SSP 1-2.6. Streamline length represents the velocity of climatic niche movement. (b) CCE under SSP 5-8.5 is high across the Adirondacks, Allegheny Plateau, and Hudson Valley. (c) The proportion of the state without a climate analog within the 250-km buffer zone increases drastically with more extreme climate change.



et al., 2023; Bystriakova et al., 2018; Michener, 1979; Orr et al., 2021). Our findings highlight the importance of modeling biodiversity at scales that permit the inclusion of local ecological contexts. For bees specifically, ensuring effective conservation action requires an understanding of the habitat associations of the regional bee assemblage. In NYS, basing conservation action on the results of global assessments may overestimate the importance of grassland habitats and underestimate that of forests, which our models predict contain considerable bee diversity and which have growing evidence of their role as habitat and forage for spring flying pollinators (Harrison et al., 2018; Ulyshen et al., 2023).

Our work shows the feasibility of developing fine-scale maps for guiding conservation of data-limited taxa like bees, despite insufficient historical survey effort, dependence on unstructured collection records, and taxonomic biases (Chesshire et al., 2023; Rousseau et al., 2024). While monitoring programs are taking shape to address these data deficiencies, many large-scale conservation actions are slated to occur by the end of the decade under the 30 × 30 framework, which strives for 30% of land area to be conserved by 2030 (Dinerstein et al., 2019). The rapid timeline laid out for these targets presents a challenge for historically underprioritized wild bee species. The spatial data layers we report here are a valuable resource for including native bee diversity in imminent conservation decisions and can inform the

protection of these ecologically and economically important pollinators.

Applying the precautionary principle, we encourage prompt conservation action for the state's bee species despite the complications from the fundamental limitations in ecological knowledge. For most bee species, conservation decisions—regardless of conservation status, habitat, or host plant requirements—should focus on improving habitat quality and quantity by providing suitable nesting and foraging resources. The spatial data products we report here, richness, RSR, uniqueness, and CCE, lay a strong foundation for facilitating targeted action in areas that host the greatest bee diversity, house irreplaceable and distinctive pollinator communities, or which are likely to become increasingly important with climate change. Where funding for pollinator conservation is limited, management actions should prioritize alternatives that facilitate co-benefits with other conservation targets in the areas identified as having particularly diverse or unique bee communities or that we predict may be suitable for species of greatest concern. In these areas, existing best practices for sustainable land management can be optimized for co-benefits with pollinator conservation priorities, for example, by selective thinning and burning in forests, managing invasive plant species, and restoring diverse native flora following disturbance (Glenny et al., 2022; Ulyshen et al., 2024). Integrating pollinator conservation into ongoing

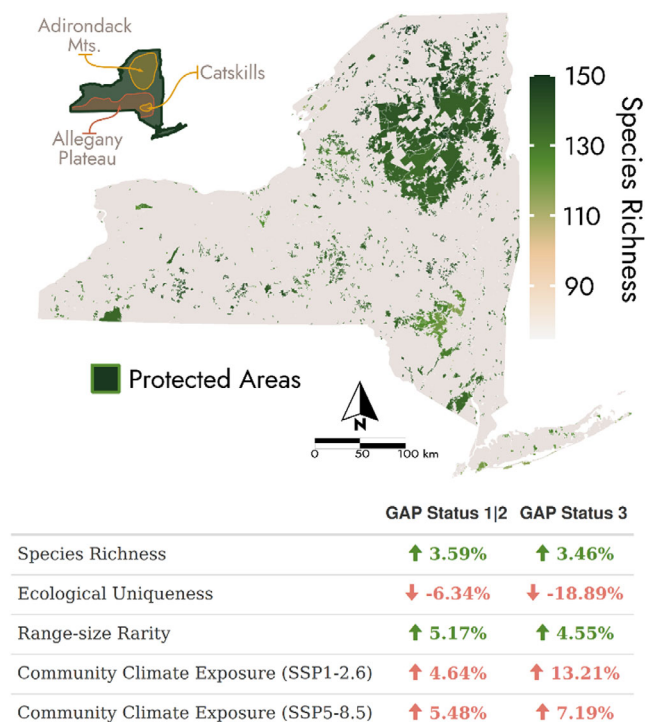


FIGURE 5 Protected areas across New York State are primarily clustered in the Adirondack Mountains and Catskills. These areas have higher species richness on average than the statewide mean, lower ecological uniqueness, greater importance for range-restricted species, and higher CCEs. The reported percentages are the deviation from the statewide average within protected areas. Areas where extractive uses are permitted (GAP status 3) showed greater deviations from the statewide average in ecological uniqueness and climate exposure than areas prohibiting extraction (GAP status 1 and 2).

management programs may provide the necessary resources for forest-dependent species without compromising the objectives of these programs and while reducing overall costs.

In December 2022, the NYS legislature passed a bill modifying the state's land acquisition policy to align with 30×30 (Department of Environmental Conservation, 2022). Our results provide an opportunity to include bee diversity as a component of broader protected area prioritizations across the state using the best available data. The current global maps of bee diversity can play a vital role in facilitating cooperative, international conservation prioritizations but are provided at too coarse of a resolution to inform local conservation action (Chaplin-Kramer et al., 2022; Wyborn & Evans, 2021). Even national biodiversity conservation programs are often hindered by a lack of maps on complementary scales (Schmidt-Traub, 2021). Previous studies that mapped bee diversity on a global scale achieved resolutions of 10 km to 100 km (Bystrakova et al., 2018; Orr et al., 2021), here we reduced the

resolution to less than 1 km. While nationwide maps of bee diversity on similar scales to what we present here are likely feasible, they will require significant data and computational overhead. By focusing on a smaller spatial area, we could achieve fine resolutions of our gridded data layers and therefore provide an opportunity to expand the utility of the recent statewide survey effort.

While the ESNPS provided a source of high-quality, recent bee observations, many of the limitations common to bee occurrence records were still present. Persistent data deficiencies prohibited us from including many of the bee species known from the state. The species that did not meet the minimum requirements are likely rare within NYS or may be specialized on host plants or habitats that are not regularly sampled. As such, we urge caution when considering our predictions of richness and RSR as we were only able to consider the most common and widely distributed species. Similarly, we note that our climate exposure results do not indicate probable species extinctions from an inability to redistribute, as all the modeled species had wide ranges outside of the Adirondacks. To better estimate the threat of climate change, further analyses and data on bee dispersal, adaptive potential, and the influence of biotic interactions are needed. Finally, the models we present were principally designed to generate spatial predictions of bee richness and diversity across NYS. Often, models are presented to predict and infer, but these two goals require different approaches (Tredennick et al., 2021). The modeled response curves reflect general trends in how habitat suitability varies across environmental gradients represented by the included features; however, they may not reflect the true response of individual species or the larger bee community to the specific covariates included in the model.

Presence-only records present a challenge for generating unbiased spatial predictions. By thoroughly accounting for sampling bias in the modeling process, we were able to limit the effects, but additional biases from variation in the sampling intensity of different environments may still be present in our models. Moreover, presence-only models cannot accurately represent demographic rates or abundances (Lee-Yaw et al., 2022), imposing limits on the conservation implications of these results. Even so, without implementing large-scale, extensive, and structured repeat surveys, such analyses are not realistic due to the lack of structured historical data. Given the substantial cost and time required to make up for the limitations in existing insect biodiversity data, we emphasize the importance of leveraging the best data currently available by employing methods robust to small sample sizes and applying thorough data cleaning and bias correction methods, as we did here.

Further refinement of our understanding of the distribution of bee diversity will require additional data collection or the application of novel modeling methodologies. Specifically, additional targeted surveys for newly identified rarities from the ESNPS could enable the inclusion of more bees of conservation interest in future models and facilitate assessments of the species that may be at the greatest risk of decline. The predicted distributions we present here provide an initial hypothesis of bee diversity and distributions across the state. These data products can aid in the protection of rare species by expediting the identification of new populations and streamlining targeted surveys of unmodeled species by prioritizing sampling in areas of high suitability for related modeled species or those sharing similar habitat and host plant preferences (Buckner & Danforth, 2022). Alternatively, with faster implementations of joint species distribution models on the horizon (Rahman et al., 2024), it may be possible to more accurately model rare species by borrowing strength from more common bees (Norberg et al., 2019). At a minimum, additional data collection through state surveys, national monitoring programs, or community science engagement will facilitate improved mapping of bee diversity, conservation, and ecosystem services, even if the resulting data are biased (Gaul et al., 2020).

The products resulting from our study represent the current best estimate of bee species richness and areas important to bee diversity across NYS to date. By leveraging species distribution modeling methods robust to small sample sizes, we were able to model over half of NYS's bee species. We present our results as gridded data layers to support statewide biodiversity conservation efforts. The data layers facilitate the inclusion of bee species in broader conservation planning and the targeted conservation and study of rare species. While invertebrate conservation can be limited by the lack of fundamental ecological knowledge and data, here we exemplify the opportunity presented by large integrated occurrence datasets for facilitating the conservation of these highly diverse and ecologically essential taxa despite these core limitations. Our work serves as an example of the extensibility of statewide pollinator surveys and a roadmap for leveraging existing occurrence datasets to ensure that bee diversity is accounted for in contemporary biodiversity conservation initiatives.

AUTHOR CONTRIBUTIONS

Mark A. Buckner: Conceptualization (lead); data curation (lead); formal analysis (lead); funding acquisition (equal); methodology (lead); software (lead); visualization (lead); writing – original draft; writing – review & editing

(lead). **Erin L. White:** Data curation (supporting); writing – review & editing (equal). **Timothy G. Howard:** Data curation (supporting); writing – review & editing (equal). **Matthew D. Schlesinger:** Data curation (supporting); writing – review & editing (equal). **Bryan N. Danforth:** Conceptualization (supporting); funding acquisition (equal); supervision (lead); writing – review & editing (equal).

ACKNOWLEDGMENTS

We thank L. Richardson and C. Wilson for providing feedback on earlier iterations of our models. Members of the Pollinator Reading Group at Cornell University and the Danforth Lab shared valuable feedback on this manuscript. The Albany Pine Bush Preserve Commission, K. Urban-Mead, and L. Richardson provided additional bee occurrence records. Preprint: bioRxiv <https://doi.org/10.1101/2024.07.15.603570>.

FUNDING INFORMATION

The Cornell Atkinson Center for Sustainability provided funding for this work through the Sustainable Biodiversity Fund awarded to MAB.

DATA AVAILABILITY STATEMENT

The raw occurrence data from the Empire State Native Pollinator Survey is available from the New York Natural Heritage Program; restrictions on the use of this data may apply. A post-processing presence/background dataset is provided with the permission of contributors for model reproducibility. The occurrence dataset and all data products produced during this study are openly available as GeoTIFFs stored on Figshare: <https://doi.org/10.6084/m9.figshare.25883932> (Buckner et al., 2025a). All associated code is archived on Figshare: <https://doi.org/10.6084/m9.figshare.25884169> (Buckner et al., 2025b), and is available on GitHub: https://github.com/bcknr/esnps_beets.

ORCID

Mark A. Buckner  <https://orcid.org/0000-0002-9692-7454>

REFERENCES

- Almeida, E. A. B., Bossert, S., Danforth, B. N., Porto, D. S., Freitas, F. V., Davis, C. C., Murray, E. A., Blaimer, B. B., Spasojevic, T., Ströher, P. R., Orr, M. C., Packer, L., Brady, S. G., Kuhlmann, M., Branstetter, M. G., & Pie, M. R. (2023). The evolutionary history of bees in time and space. *Current Biology*, 33(16), 3409–3422.e6. <https://doi.org/10.1016/j.cub.2023.07.005>
- Althaus, S. L., Berenbaum, M. R., Jordan, J., & Shalmon, D. A. (2021). No buzz for bees: Media coverage of pollinator decline. *Proceedings of the National Academy of Sciences of the*

- United States of America, 118(2), e2002552117. <https://doi.org/10.1073/pnas.2002552117>
- Antoine, C. M., & Forrest, J. R. K. (2021). Nesting habitat of ground-nesting bees: A review. *Ecological Entomology*, 46(2), 143–159. <https://doi.org/10.1111/een.12986>
- Ascher, J. S., & Pickering, J. (2020). Discover life bee species guide and world checklist (Hymenoptera: Apoidea: Anthophila).
- Baddeley, A., Rubak, E., & Turner, R. (2015). *Spatial point patterns: Methodology and applications with R*. CRC Press, Taylor & Francis Group.
- Beier, P., Hansen, L. J., Helbrecht, L., & Behar, D. (2017). A how-to guide for coproduction of actionable science. *Conservation Letters*, 10(3), 288–296. <https://doi.org/10.1111/conl.12300>
- Breiner, F. T., Guisan, A., Bergamini, A., & Nobis, M. P. (2015). Overcoming limitations of modelling rare species by using ensembles of small models. *Methods in Ecology and Evolution*, 6(10), 1210–1218. <https://doi.org/10.1111/2041-210X.12403>
- Brun, P., Thuiller, W., Chauvier, Y., Pellissier, L., Wüest, R. O., Wang, Z., & Zimmermann, N. E. (2020). Model complexity affects species distribution projections under climate change. *Journal of Biogeography*, 47(1), 130–142. <https://doi.org/10.1111/jbi.13734>
- Buckner, M. A., & Danforth, B. N. (2022). Climate-driven range shifts of a rare specialist bee, *Macropis nuda* (Melittidae), and its host plant, *Lysimachia ciliata* (Primulaceae). *Global Ecology and Conservation*, 37, e02180. <https://doi.org/10.1016/j.gecco.2022.e02180>
- Buckner, M. A., White, E. L., Howard, T. G., D. Schlesinger, M., & N. Danforth, B. (2025a). Data: Mapping bee diversity with landscape-level models to inform conservation [Data set]. figshare. <https://doi.org/10.6084/M9.FIGSHARE.25883932>
- Buckner, M. A., White, E. L., Howard, T. G., D. Schlesinger, M., & N. Danforth, B. (2025b). Code: Mapping bee diversity with landscape-level models to inform conservation. figshare. <https://doi.org/10.6084/M9.FIGSHARE.25884169>
- Bystrakova, N., Griswold, T., Ascher, J. S., & Kuhlmann, M. (2018). Key environmental determinants of global and regional richness and endemism patterns for a wild bee subfamily. *Biodiversity and Conservation*, 27(2), 287–309. <https://doi.org/10.1007/s10531-017-1432-7>
- Calabrese, J. M., Certain, G., Kraan, C., & Dormann, C. F. (2014). Stacking species distribution models and adjusting bias by linking them to macroecological models. *Global Ecology and Biogeography*, 23(1), 99–112. <https://doi.org/10.1111/geb.12102>
- Cardoso, P., Erwin, T. L., Borges, P. A. V., & New, T. R. (2011). The seven impediments in invertebrate conservation and how to overcome them. *Biological Conservation*, 144(11), 2647–2655. <https://doi.org/10.1016/j.biocon.2011.07.024>
- Carril, O. M., Griswold, T., Haefner, J., & Wilson, J. S. (2018). Wild bees of grand staircase-escalante national monument: Richness, abundance, and spatio-temporal beta-diversity. *PeerJ*, 2018(11), e5867. <https://doi.org/10.7717/peerj.5867>
- Carson, R. (1962). *Silent spring*. Houghton Mifflin.
- Center for International Earth Science Information Network - CIESIN - Columbia University. (2018). *Gridded population of the world, version 4 (GPWv4): Population density, revision 11*. NASA Socioeconomic Data and Applications Center (SEDAC). <https://doi.org/10.7927/H49C6VHW>
- Chaplin-Kramer, R., Brauman, K. A., Cavender-Bares, J., Díaz, S., Duarte, G. T., Enquist, B. J., Garibaldi, L. A., Geldmann, J., Halpern, B. S., Hertel, T. W., Khoury, C. K., Krieger, J. M., Lavorel, S., Mueller, T., Neugarten, R. A., Pinto-Ledezma, J., Polasky, S., Purvis, A., Reyes-García, V., ... Zafra-Calvo, N. (2022). Conservation needs to integrate knowledge across scales. *Nature Ecology & Evolution*, 6(2), 118–119. <https://doi.org/10.1038/s41559-021-01605-x>
- Chen, C., & Breiman, L. (2004). *Using random forest to learn imbalanced data*. University of California Berkeley. <https://statistics.berkeley.edu/sites/default/files/tech-reports/666.pdf>
- Chesshire, P. R., Fischer, E. E., Dowdy, N. J., Griswold, T. L., Hughes, A. C., Orr, M. C., Ascher, J. S., Guzman, L. M., Hung, K. J., Cobb, N. S., & McCabe, L. M. (2023). Completeness analysis for over 3000 United States bee species identifies persistent data gap. *Ecography*, 2023(5), e06584. <https://doi.org/10.1111/ecog.06584>
- Colla, S. R., & MacIvor, J. S. (2017). Questioning public perception, conservation policy, and recovery actions for honeybees in North America. *Conservation Biology*, 31(5), 1202–1204. <https://doi.org/10.1111/cobi.12839>
- Cunningham-Minnick, M. J., Milam, J., Fassler, A., & King, D. I. (2024). Best management practices for bee conservation in forest openings. *Conservation Science and Practice*, 6(11), e13231. <https://doi.org/10.1111/csp2.13231>
- D'Amen, M., Dubuis, A., Fernandes, R. F., Pottier, J., Pellissier, L., & Guisan, A. (2015). Using species richness and functional traits predictions to constrain assemblage predictions from stacked species distribution models. *Journal of Biogeography*, 42(7), 1255–1266. <https://doi.org/10.1111/jbi.12485>
- Danforth, B. (2024). New York bee traits (Unpublished data).
- Dansereau, G., Legendre, P., & Poisot, T. (2022). Evaluating ecological uniqueness over broad spatial extents using species distribution modelling. *Oikos*, 2022(5), e09063. <https://doi.org/10.1111/oik.09063>
- Department of Environmental Conservation. (2022). Thirty by thirty conservation goal, 49-0113 New York Environmental Conservation Law.
- Dicks, L. V., Breeze, T. D., Ngo, H. T., Senapathi, D., An, J., Aizen, M. A., Basu, P., Buchori, D., Galetto, L., Garibaldi, L. A., Gemmill-Herren, B., Howlett, B. G., Imperatriz-Fonseca, V. L., Johnson, S. D., Kovács-Hostyánszki, A., Kwon, Y. J., Lattorff, H. M. G., Lungharwo, T., Seymour, C. L., & Potts, S. G. (2021). A global-scale expert assessment of drivers and risks associated with pollinator decline. *Nature Ecology & Evolution*, 5(10), 1453–1461. <https://doi.org/10.1038/s41559-021-01534-9>
- Dinerstein, E., Vynne, C., Sala, E., Joshi, A. R., Fernando, S., Lovejoy, T. E., Mayorga, J., Olson, D., Asner, G. P., Baillie, J. E. M., Burgess, N. D., Burkart, K., Noss, R. F., Zhang, Y. P., Baccini, A., Birch, T., Hahn, N., Joppa, L. N., & Wikramanayake, E. (2019). A global deal for nature: Guiding principles, milestones, and targets. *Science Advances*, 5(4), eaaw2869. <https://doi.org/10.1126/sciadv.aaw2869>
- Drivdal, L., & van der Sluijs, J. P. (2021). Pollinator conservation requires a stronger and broader application of the precautionary principle. *Current Opinion in Insect Science*, 46, 95–105. <https://doi.org/10.1016/j.cois.2021.04.005>

- Erickson, K. D., & Smith, A. B. (2023). Modeling the rarest of the rare: A comparison between multi-species distribution models, ensembles of small models, and single-species models at extremely low sample sizes. *Ecography*, 2023(6), e06500. <https://doi.org/10.1111/ecog.06500>
- Frison, T. H. (1926). Contribution to the knowledge of the interrelations of the bumblebees of Illinois with their animate environment. *Annals of the Entomological Society of America*, 19(2), 203–235. <https://doi.org/10.1093/aesa/19.2.203>
- Fronk, W. D., & Painter, L. I. (1960). Some characteristics of alkali bee nesting sites. *Journal of Economic Entomology*, 53(3), 424–425. <https://doi.org/10.1093/jee/53.3.424>
- Gábor, L., Jetz, W., Lu, M., Rocchini, D., Cord, A., Malavasi, M., Zarzo-Arias, A., Barták, V., & Moudrý, V. (2022). Positional errors in species distribution modelling are not overcome by the coarser grains of analysis. *Methods in Ecology and Evolution*, 13(10), 2289–2302. <https://doi.org/10.1111/2041-210X.13956>
- Gaul, W., Sadykova, D., White, H. J., Leon-Sanchez, L., Caplat, P., Emmerson, M. C., & Yearsley, J. M. (2020). Data quantity is more important than its spatial bias for predictive species distribution modelling. *PeerJ*, 8, e10411. <https://doi.org/10.7717/peerj.10411>
- GBIF.org. (2024). GBIF occurrence download. <https://doi.org/10.15468/dl.rm72jb>
- Geldmann, J., & González-Varo, J. P. (2018). Conserving honey bees does not help wildlife: High densities of managed honey bees can harm populations of wild pollinators. *Science*, 359(6374), 392–393. <https://doi.org/10.1126/science.aar2269>
- Glenny, W., Runyon, J. B., & Burkle, L. A. (2022). A review of management actions on insect pollinators on public lands in the United States. *Biodiversity and Conservation*, 31(8), 1995–2016. <https://doi.org/10.1007/s10531-022-02399-5>
- Goulson, D., Nicholls, E., Botías, C., & Rotheray, E. L. (2015). Bee declines driven by combined stress from parasites, pesticides, and lack of flowers. *Science*, 347(6229), 1255957. <https://doi.org/10.1126/science.1255957>
- Greenleaf, S. S., Williams, N. M., Winfree, R., & Kremen, C. (2007). Bee foraging ranges and their relationship to body size. *Oecologia*, 153(3), 589–596. <https://doi.org/10.1007/s00442-007-0752-9>
- Grohmann, C. H., Smith, M. J., & Riccomini, C. (2011). Multiscale analysis of topographic surface roughness in the Midland Valley, Scotland. *IEEE Transactions on Geoscience and Remote Sensing*, 49(4), 1200–1213. <https://doi.org/10.1109/TGRS.2010.2053546>
- Guerin, G. R., & Lowe, A. J. (2015). “Sum of inverse range-sizes” (SIR), a biodiversity metric with many names and interpretations. *Biodiversity and Conservation*, 24(11), 2877–2882. <https://doi.org/10.1007/s10531-015-0977-6>
- Guillera-Arroita, G., Lahoz-Monfort, J. J., Elith, J., Gordon, A., Kujala, H., Lentini, P. E., McCarthy, M. A., Tingley, R., & Wintle, B. A. (2015). Is my species distribution model fit for purpose? Matching data and models to applications. *Global Ecology and Biogeography*, 24(3), 276–292. <https://doi.org/10.1111/geb.12268>
- Hall, D. M., & Steiner, R. (2019). Insect pollinator conservation policy innovations at subnational levels: Lessons for lawmakers. *Environmental Science & Policy*, 93, 118–128. <https://doi.org/10.1016/j.envsci.2018.12.026>
- Hamann, A., Roberts, D. R., Barber, Q. E., Carroll, C., & Nielsen, S. E. (2015). Velocity of climate change algorithms for guiding conservation and management. *Global Change Biology*, 21(2), 997–1004. <https://doi.org/10.1111/gcb.12736>
- Harrison, T., Gibbs, J., & Winfree, R. (2018). Forest bees are replaced in agricultural and urban landscapes by native species with different phenologies and life-history traits. *Global Change Biology*, 24(1), 287–296. <https://doi.org/10.1111/gcb.13921>
- Hengl, T. (2018). Clay content in % (kg/kg) at 6 standard depths (0, 10, 30, 60, 100 and 200 cm) at 250 m resolution (version v0.2). *Zenodo*. <https://doi.org/10.5281/ZENODO.2525663>
- Hijmans, R. J. (2023). terra: Spatial data analysis (version 1.7-65). <https://rspatial.org/>
- Hirzel, A. H., Le Lay, G., Helfer, V., Randin, C., & Guisan, A. (2006). Evaluating the ability of habitat suitability models to predict species presences. *Ecological Modelling*, 199(2), 142–152. <https://doi.org/10.1016/j.ecolmodel.2006.05.017>
- Inman, R., Franklin, J., Esque, T., & Nussear, K. (2021). Comparing sample bias correction methods for species distribution modeling using virtual species. *Ecosphere*, 12(3), e03422. <https://doi.org/10.1002/ecs2.3422>
- Inouye, D., Droege, S., & Mawdsley, J. (2017). Words alone will not protect pollinators. *Science*, 355(6323), 357. <https://doi.org/10.1126/science.aam6132>
- Iwasaki, J. M., & Hogendoorn, K. (2021). How protection of honey bees can help and hinder bee conservation. *Current Opinion in Insect Science*, 46, 112–118. <https://doi.org/10.1016/j.cois.2021.05.005>
- Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., Zimmermann, N. E., Linder, H. P., & Kessler, M. (2017). Climatologies at high resolution for the earth's land surface areas. *Scientific Data*, 4, 170122. <https://doi.org/10.1038/sdata.2017.122>
- Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., Zimmermann, N. E., Linder, H. P., & Kessler, M. (2021). Climatologies at high resolution for the earth's land surface areas. <https://doi.org/10.16904/envdat.228>
- Kendall, L. K., Mola, J. M., Portman, Z. M., Cariveau, D. P., Smith, H. G., & Bartomeus, I. (2022). The potential and realized foraging movements of bees are differentially determined by body size and sociality. *Ecology*, 103(11), e3809. <https://doi.org/10.1002/ecy.3809>
- Kuhn, M. (2008). Building predictive models in R using the caret package. *Journal of Statistical Software*, 28, 1–26. <https://doi.org/10.18637/jss.v028.i05>
- Kujala, H., Whitehead, A., & Wintle, B. (2015). Identifying conservation priorities and assessing impacts and trade-offs of potential future development in the Lower Hunter Valley in New South Wales. <https://doi.org/10.13140/RG.2.1.1417.4884/1>
- LANDFIRE. (2014). *Existing vegetation type layer (version LANDFIRE 1.4.0)*. U.S. Department of the Interior, Geological Survey, and U.S. Department of Agriculture. https://landfire.gov/version_download.php
- LANDFIRE. (2020). *Slope degrees (version LANDFIRE 2.2.0)*. U.S. Department of the Interior, Geological Survey, and U.S. Department of Agriculture. https://landfire.gov/version_download.php
- Lee-Yaw, J., McCune, J. L., Pironon, S., & Sheth, S. N. (2022). Species distribution models rarely predict the biology of real

- populations. *Ecography*, 2022(6), e05877. <https://doi.org/10.1111/ecog.05877>
- Legendre, P., & De Cáceres, M. (2013). Beta diversity as the variance of community data: Dissimilarity coefficients and partitioning. *Ecology Letters*, 16(8), 951–963. <https://doi.org/10.1111/ele.12141>
- Lomba, A., Pellissier, L., Randin, C., Vicente, J., Moreira, F., Honrado, J., & Guisan, A. (2010). Overcoming the rare species modelling paradox: A novel hierarchical framework applied to an Iberian endemic plant. *Biological Conservation*, 143(11), 2647–2657. <https://doi.org/10.1016/j.biocon.2010.07.007>
- Mi, C., Huettmann, F., Guo, Y., Han, X., & Wen, L. (2017). Why choose Random Forest to predict rare species distribution with few samples in large undersampled areas? Three Asian crane species models provide supporting evidence. *PeerJ*, 5(1), e2849. <https://doi.org/10.7717/peerj.2849>
- Michener, C. D. (1979). Biogeography of the bees. *Annals of the Missouri Botanical Garden*, 66(3), 277–347.
- Milá, C., Mateu, J., Pebesma, E., & Meyer, H. (2022). Nearest neighbour distance matching leave-one-out cross-validation for map validation. *Methods in Ecology and Evolution*, 13(6), 1304–1316. <https://doi.org/10.1111/2041-210X.13851>
- New York Natural Heritage Program. (2024). New York protected areas database (NYPAD). <https://nypad.org>
- Norberg, A., Abrego, N., Blanchet, F. G., Adler, F. R., Anderson, B. J., Anttila, J., Araújo, M. B., Dallas, T., Dunson, D., Elith, J., Foster, S. D., Fox, R., Franklin, J., Godsoe, W., Guisan, A., O'Hara, B., Hill, N. A., Holt, R. D., Hui, F. K. C., & Ovaskainen, O. (2019). A comprehensive evaluation of predictive performance of 33 species distribution models at species and community levels. *Ecological Monographs*, 89(3), e01370. <https://doi.org/10.1002/ecm.1370>
- Orr, M. C., Hughes, A. C., Chesters, D., Pickering, J., Zhu, C. D., & Ascher, J. S. (2021). Global patterns and drivers of bee distribution. *Current Biology*, 31(3), 451–458.e4. <https://doi.org/10.1016/j.cub.2020.10.053>
- Potts, S. G., Biesmeijer, J. C., Kremen, C., Neumann, P., Schweiger, O., & Kunin, W. E. (2010). Global pollinator declines: Trends, impacts and drivers. *Trends in Ecology and Evolution*, 25(6), 345–353. <https://doi.org/10.1016/j.tree.2010.01.007>
- R Core Team. (2021). *R: A language and environment for statistical computing [Computer software]*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Rahman, A. U., Tikhonov, G., Oksanen, J., Rossi, T., & Ovaskainen, O. (2024). Accelerating joint species distribution modeling with Hmsc-HPC: A 1000x faster GPU deployment. *bioRxiv*: <https://doi.org/10.1101/2024.02.13.580046>
- Richardson, L. (2024). Bumble bees of North America: An identification guide. <https://www.leifrichardson.org/bbna.html>
- Rousseau, J. S., Woodard, S. H., Jepsen, S., Du Clos, B., Johnston, A., Danforth, B. N., & Rodewald, A. D. (2024). Advancing bee conservation in the US: Gaps and opportunities in data collection and reporting. *Frontiers in Ecology and Evolution*, 12, 1346795. <https://doi.org/10.3389/fevo.2024.1346795>
- Sage, R. F. (2020). Global change biology: A primer. *Global Change Biology*, 26(1), 3–30. <https://doi.org/10.1111/gcb.14893>
- Schlesinger, M. D., White, E. L., Corser, J. D., Danforth, B. N., Fierke, M. K., Greenwood, C. M., Hatfield, R. G., Hietala-Henschell, K. G., Mawdsley, J. R., McFarland, K. P., Niver, R., Rozen, J. G. J., Van Dyke, M., & Howard, T. G. (2023). A multi-taxonomic survey to determine the conservation status of native pollinators. *Frontiers in Ecology and Evolution*, 11, 1274680. <https://doi.org/10.3389/fevo.2023.1274680>
- Schmidt-Traub, G. (2021). National climate and biodiversity strategies are hamstrung by a lack of maps. *Nature Ecology & Evolution*, 5(10), 1325–1327. <https://doi.org/10.1038/s41559-021-01533-w>
- Stark, J. R., & Fridley, J. D. (2022). Microclimate-based species distribution models in complex forested terrain indicate widespread cryptic refugia under climate change. *Global Ecology and Biogeography*, 31(3), 562–575. <https://doi.org/10.1111/geb.13447>
- Stephen, W. P. (1955). Alfalfa pollination in manitoba1. *Journal of Economic Entomology*, 48(5), 543–548. <https://doi.org/10.1093/jee/48.5.543>
- Tepedino, V. J., & Portman, Z. M. (2021). Intensive monitoring for bees in North America: Indispensable or improvident? *Insect Conservation and Diversity*, 14(5), 535–542. <https://doi.org/10.1111/icad.12509>
- Tredennick, A. T., Hooker, G., Ellner, S. P., & Adler, P. B. (2021). A practical guide to selecting models for exploration, inference, and prediction in ecology. *Ecology*, 102(6), e03336. <https://doi.org/10.1002/ecy.3336>
- U.S. Census Bureau. (2010). 2010 TIGER/line shapefiles.
- U.S. Census Bureau. (2013). 2013 TIGER/line shapefiles.
- Ulyshen, M., Urban-Mead, K. R., Dorey, J. B., & Rivers, J. W. (2023). Forests are critically important to global pollinator diversity and enhance pollination in adjacent crops. *Biological Reviews*, 98(4), 1118–1141. <https://doi.org/10.1111/brv.12947>
- Ulyshen, M. D., Ballare, K. M., Fettig, C. J., Rivers, J. W., & Runyon, J. B. (2024). The value of forests to pollinating insects varies with forest structure, composition, and age. *Current Forestry Reports*, 10(5), 322–336. <https://doi.org/10.1007/s40725-024-00224-6>
- United Nations Environment Programme. (1992). World atlas of desertification. <https://wedocs.unep.org/20.500.11822/42137>
- USGS GAP. (2022). *Protected areas database of the United States (PAD-US) (version 3.0)*. U.S. Geological Survey data release. USGS. <https://doi.org/10.5066/P9Q9LQ4B>
- Valavi, R., Elith, J., Lahoz-Monfort, J. J., & Guillerá-Arroita, G. (2019). blockCV: An r package for generating spatially or environmentally separated folds for k-fold cross-validation of species distribution models. *Methods in Ecology and Evolution*, 10(2), 225–232. <https://doi.org/10.1111/2041-210X.13107>
- Valavi, R., Elith, J., Lahoz-Monfort, J. J., & Guillerá-Arroita, G. (2021). Modelling species presence-only data with random forests. *Ecography*, 44(12), 1731–1742. <https://doi.org/10.1111/ecog.05615>
- Valavi, R., Guillerá-Arroita, G., Lahoz-Monfort, J. J., & Elith, J. (2022). Predictive performance of presence-only species distribution models: A benchmark study with reproducible code. *Ecological Monographs*, 92(1), e01486. <https://doi.org/10.1002/ecm.1486>
- van Engelsdorp, D., Evans, J. D., Saegerman, C., Mullin, C., Haubruge, E., Nguyen, B. K., Frazier, M., Frazier, J., Cox-Foster, D., Chen, Y., Underwood, R., Tarpy, D. R., & Pettis, J. S. (2009). Colony collapse disorder: A descriptive study. *PLoS One*, 4(8), e6481. <https://doi.org/10.1371/journal.pone.0006481>

- Vanbergen, A. J., Garratt, M. P., Vanbergen, A. J., Baude, M., Biesmeijer, J. C., Britton, N. F., Brown, M. J. F., Brown, M., Bryden, J., Budge, G. E., Bull, J. C., Carvell, C., Challinor, A. J., Connolly, C. N., Evans, D. J., Feil, E. J., Garratt, M. P., Greco, M. K., Heard, M. S., & Wright, G. A. (2013). Threats to an ecosystem service: Pressures on pollinators. *Frontiers in Ecology and the Environment*, 11(5), 251–259. <https://doi.org/10.1890/120126>
- Warton, D. I., Renner, I. W., & Ramp, D. (2013). Model-based control of observer bias for the analysis of presence-only data in ecology. *PLoS One*, 8(11), e79168. <https://doi.org/10.1371/journal.pone.0079168>
- White, E., Schlesinger, M., & Howard, T. (2022). *The Empire State native pollinator survey 2017–2021*. New York Natural Heritage Program.
- Williams, P. H. (1982). The distribution and decline of British bumble bees (*Bombus* Latr.). *Journal of Apicultural Research*, 21(4), 236–245. <https://doi.org/10.1080/00218839.1982.11100549>
- Woodard, S. H., Federman, S., James, R. R., Danforth, B. N., Griswold, T. L., Inouye, D., McFrederick, Q. S., Morandin, L., Paul, D. L., Sellers, E., Strange, J. P., Vaughan, M., Williams, N. M., Branstetter, M. G., Burns, C. T., Cane, J., Cariveau, A. B., Cariveau, D. P., Childers, A., ... Wehling, W. (2020). Towards a U.S. national program for monitoring native bees. *Biological Conservation*, 252, 108821. <https://doi.org/10.1016/j.biocon.2020.108821>
- Wright, M. N., & Ziegler, A. (2017). ranger: A fast implementation of random forests for high dimensional data in C++ and R. *Journal of Statistical Software*, 77, 1–17. <https://doi.org/10.18637/jss.v077.i01>
- Wyborn, C., & Evans, M. C. (2021). Conservation needs to break free from global priority mapping. *Nature Ecology and Evolution*, 5(10), 1322–1324. <https://doi.org/10.1038/s41559-021-01540-x>
- Zizka, A., Silvestro, D., Andermann, T., Azevedo, J., Duarte Ritter, C., Edler, D., Farooq, H., Herdean, A., Ariza, M., Scharn, R., Svantesson, S., Wengström, N., Zizka, V., & Antonelli, A. (2019). CoordinateCleaner: Standardized cleaning of occurrence records from biological collection databases. *Methods in Ecology and Evolution*, 10(5), 744–751. <https://doi.org/10.1111/2041-210X.13152>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Buckner, M. A., White, E. L., Howard, T. G., Schlesinger, M. D., & Danforth, B. N. (2026). Mapping bee diversity with landscape-level models to inform conservation. *Conservation Science and Practice*, e70239. <https://doi.org/10.1111/csp2.70239>