

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- | | | |
|-------------------------------------|-------------------------------------|--|
| ☐ | ☒ | The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement |
| ☒ | ☐ | A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☒ | ☐ | The statistical test(s) used AND whether they are one- or two-sided
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| ☐ | ☒ | A description of all covariates tested |
| ☐ | ☒ | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| ☐ | ☒ | A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| ☒ | ☐ | For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
<i>Give P values as exact values whenever suitable.</i> |
| ☐ | ☒ | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| ☐ | ☒ | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| ☒ | ☐ | Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection	We used openly available data. Which was either downloaded manually or using R.
Data analysis	For all analysis we used R V4.3.2. For spatial manipulations we used the packages sf V1.0-15; for data manipulation and visualization we used tidyverse V2.0.0 and data.table V1.15.0; for running models, we used runjags V2.2.2-4. All code is posted on https://github.com/lmguzman/bee_occupancy and is published in Zenodo in https://zenodo.org/records/56912668535

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

North American occurrence records were acquired from Chesshire et al. 2023. Land cover data was obtained from the National Land Cover Database (NLCD) for all

years from from 2001 to 2016. <https://www.mrlc.gov/data/nlcd-land-cover-conus-all-years>. Crop cover data was obtained from the Crop Data Layer (CDL) for 2008 https://www.nass.usda.gov/Research_and_Science/Cropland/Release/index.php. Attractiveness of agricultural crops was obtained from the USDA "Attractiveness of Agricultural Crops to Pollinating Bees for the Collection of Nectar and/or Pollen, 2017" report <https://www.usda.gov/sites/default/files/documents/Attractiveness-of-Agriculture-Crops-to-Pollinating-Bees-Report-FINAL-Web-Version-Jan-3-2018.pdf>. Pesticide use data was obtained from USGS Pesticide National Synthesis Project <https://water.usgs.gov/nawqa/pnsp/usage/maps/county-level/> for all years from 1995 to 2015. LD50 estimates were acquired from the EPA ECOTOX Database <https://cfpub.epa.gov/ecotox/search.cfm>. Honey bee colony counts were obtained from The National Agricultural Statistics Service (NASS) honey colony census for years 2002 to 2017. <https://quickstats.nass.usda.gov>. Temperature and precipitation data was obtained from from the CHLSA high resolution climate data. <https://chelsa-climate.org/timeseries/>.

Human research participants

Policy information about [studies involving human research participants and Sex and Gender in Research](#).

Reporting on sex and gender

NA

Population characteristics

NA

Recruitment

NA

Ethics oversight

NA

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences

☐ Behavioural & social sciences

☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description

We used hierarchical Bayesian models to estimate species occupancy. In these models, individual species' effects are drawn from distributions whose parameters are informed by data from all species. We analyze by family because otherwise the computational burden is too high. In addition, the different families provide independent validation, and bee species responses may be taxonomically grouped, we modelled each family separately. We combined Melittidae, and Colletidae (the two smallest families), because Melittidae had too few species to run the analyses reliably. In total, we ran 15 multi-species occupancy models (five family groups and three effects we estimate).

Research sample

Our goal was to evaluate the effect of pesticides on bees, as such, we used occurrence data for any bee species found in North America.

Sampling strategy

Because we used occurrence data, this data was already collected. It comprises historical museum data, data uploaded to GBIF, and community science data.

Data collection

We used the data published by Chesshire et al. 2023, where records were downloaded from GBIF and SCAN, and were cleaned to ensure (i) taxonomic names are correct and (ii) erroneous records are removed. Further, Chesshire et al. removed records for the Western Honey Bee (*Apis mellifera*) and restricted records to only those within the contiguous United States. Chesshire et al. presented a total number of 1,923,814 occurrence records for 3,158 bee species from 1700 to 2021. Of these, many are multiple observations of the same species on the same date at the same geographic location. We reduced multiple observations to 1 per species, per year and county. This resulted in a total of 634,597 unique species x date x location combinations. Because our emphasis is on post-1990s pesticide application, we removed any observation from before 1994 or after 2016. This resulted in 207,954 observations for 2,183 bee species. Further, we removed any species that had less than 10 unique observations (unique dates and locations). This further reduced the number of species to 1,183 and 204,346 observations. Finally, we also removed species that were present in less than 3 years or 3 counties, resulting in 1,107 bee species with a total of 199,462 unique records. We also removed all *A. angelicus* records and kept only *A. texanus* records because females cannot be morphologically differentiated between the two. Finally, our models only included genera for which we had data for at least three species, which resulted in a total of 1,081 bee species across the following families: Andrenidae had 220 species, Apidae had 284 species, Colletidae 69 species, Halictidae 221 species, Megachilidae 278 species, and Melittidae 9 species.

Timing and spatial scale

Our analysis was conducted across all counties in the United States where there are bee observations, which varies by family. We modelled Andrenidae in 626 counties, Apidae in 1763 counties, Halictidae in 1046 counties, Megachilidae in 970, and Colletidae and Melittidae in 500 counties. The analysis was from 1995-2015

Data exclusions

See data collection.

Reproducibility

All code for reproducing the analysis is posted on github. All input data, code, and model outputs will be posted in figshare to aid with reproducibility.

Randomization

NA

Blinding

NA

Did the study involve field work?

☐ Yes☒ No

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involvement in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involvement in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging