08-Jan-2024  
  
Dear Ms. Lee,  
  
Thank you very much for submitting your manuscript "Leveraging local and regional sources of data and a multi-species occupancy model to explore bee-plant interactions" (EAP23-0567) to Ecological Applications. We appreciate the work you have accomplished. Based on the reviewer comments, the manuscript requires significant revisions and I will not be able to accept this manuscript for publication at this time. However, I would be willing to consider a major revision based on your response to the comments outlined in this letter.  
  
Two reviewers agree that this is a novel way of using occupancy models. However, Reviewer 1 in particular highlighted a number of issues of high concern. One of the main issues mentioned by Reviewer #1 is that the data used to test the suggested model is too limited to make robust inferences. The reviewers list many other revisions needed to have a much stronger paper. These are more than can be dealt with in a straightforward revision.  
  
Should you decide to revise the manuscript for further consideration here, your revisions should address the specific comments outlined in this letter. You must include a "Response to Reviewer Comments" (you will find a corresponding field for this in ScholarOne) that shows your responses to the reviewer comments and the changes you have made in the manuscript. If you disagree with a comment, explain why. A clear and concise response to review comments will help the Subject-matter Editor when reviewing your resubmission.  
  
Please note that the field in ScholarOne does not retain type formats such as italics, boldface, or colors, so please format the responses accordingly. We suggest you upload a separate file for your "Response to Reviewer Comments" and use our template: [https://www.esa.org/wp-content/uploads/2021/04/Author-Response-to-Reviewers-Template.docx](https://nam10.safelinks.protection.outlook.com/?url=https%3A%2F%2Fwww.esa.org%2Fwp-content%2Fuploads%2F2021%2F04%2FAuthor-Response-to-Reviewers-Template.docx&data=05%7C02%7Cgdirenzo%40umass.edu%7Cf98698c60e95487501bd08dc107310ae%7C7bd08b0b33954dc194bbd0b2e56a497f%7C0%7C0%7C638403334546520490%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C3000%7C%7C%7C&sdata=DawgVXhOfdyoR%2FGDZUA2XBr1CiiSNkRbCfv5%2B3e8Sa8%3D&reserved=0).  
  
Please include a "track changes" version of your manuscript. The "track changes" files should contain "track changes" in the file name, be given the file designation "Additional File for Review but NOT for publication", and placed at the end of the file list. The "clean" copy of your manuscript should be labeled as your "Main Document".  
  
To resubmit after revising your manuscript, log in to [https://mc.manuscriptcentral.com/ecologicalapps](https://nam10.safelinks.protection.outlook.com/?url=https%3A%2F%2Fmc.manuscriptcentral.com%2Fecologicalapps&data=05%7C02%7Cgdirenzo%40umass.edu%7Cf98698c60e95487501bd08dc107310ae%7C7bd08b0b33954dc194bbd0b2e56a497f%7C0%7C0%7C638403334546520490%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C3000%7C%7C%7C&sdata=rzwVnkX0u3JINngU10lI%2FNmcYZ%2FAHxkINp0zKDX09aQ%3D&reserved=0) and go to your Author Center. You will find the option to resubmit under "Manuscripts with Decisions".  
  
The due date for your resubmission is 04-Sep-2024, which is nine months from the day this decision letter was sent. If you are not able to resubmit by the listed due date, please email us and we’ll extend the due date in the system to ensure you can easily resubmit. This can be done before or after the due date has passed. You can find the current due date for the resubmission in your Author Center under “Manuscripts with Decisions”.  
  
The resubmission would be assigned a new manuscript number and likely be sent out for additional review.  
  
Sincerely,  
  
Dr. María Silvina Fenoglio  
Subject-matter Editor, Ecological Applications  
  
\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*  
  
Editorial staff note: In addition to revisions according to the review comments below, you must make revisions to your paper to conform to the journal style at this stage, especially the tables, figures, and supporting information. Closely following our manuscript preparation guide at this stage would expedite the production of your paper for publication, should it be accepted. Please review the full guide here: [https://www.esa.org/wp-content/uploads/2022/05/ESA-Manuscript-Preparation-Guide.pdf](https://nam10.safelinks.protection.outlook.com/?url=https%3A%2F%2Fwww.esa.org%2Fwp-content%2Fuploads%2F2022%2F05%2FESA-Manuscript-Preparation-Guide.pdf&data=05%7C02%7Cgdirenzo%40umass.edu%7Cf98698c60e95487501bd08dc107310ae%7C7bd08b0b33954dc194bbd0b2e56a497f%7C0%7C0%7C638403334546520490%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C3000%7C%7C%7C&sdata=CkfJsUo0MdHwa7386jKuDNL3jnnnp4x0hjzHw2y9YsM%3D&reserved=0).  
  
\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*  
  
-------------------------------------------------  
  
Reviewer #1 Comments to the Author:  
  
This manuscript leverages state-of-the-art occupancy models with opportunistic public datasets to investigate plant-pollinator interactions. Occupancy models (and similar variants) are widely used to model species interactions when detection is imperfect. The clever gist of the paper is to use multiple sources of opportunistic data to assess detectability. However, I am unsure that the use of multiple collection types could substitute for multiple surveys (or multiple surveillants). I am not suggesting it is wrong conceptually, but given the different nature of collection types (citizen science vs natural history collections), I wonder how useful they are to model detectability. The paper would be much stronger if this is investigated, for example using simulations. In particular, assumption 3 (line 297) seems quite strong and might warrant some testing. However, if this is well known in the literature of occupancy models, just reinforcing that such type of contrasting multiple observations which use different temporal and spatial scales have the power to correctly model detectability is fine. In the same line, I wonder if simply creating a dummy variable assigning half of the data to one set of observations and the other half to another set would do the trick. All these questions remain open to me in order to generalise the method.

* Outside of scope
* Think about simulations
* Other papers use citizen science data

Other than this technical details, my main concern is that the data used to test the suggested model is too limited to make robust inferences. After the cleaning process, less than 400 observations remain in the dataset for a network of a large size (hundreds of species, with thousands of potential links observed). I wonder if the lack of clear results could emerge from the fact that the data is too sparse. In fact, the result about sociality can be interpreted as a bias in the dataset, which contains much more information on bumblebees (iconic, big and easy to observe species) than other taxa.

* Using all globi dataset as all possible interactions
  + And then doing a local analysis of –
* Start with all dataset
  + Filtering steps
  + Except lat/long list
  + Filtering step – based off of all possible interactions and active periods of pollinators/bees (SCI) – see how many interactions we loss from all possible interactions

Going further, I miss a clear indication on how good are the models (e.g. something similar to variance explained). A robust and more convincing test would split data into training and test datasets and validate how good is your model at predicting links on the test dataset using only the train dataset.

* Mention Bayesian p-value

Finally, I understand the models can’t contain many variables due to the limited sample size. However, some of the covariates are poorly justified. A clear example is the use of Yellow vs any other color for flower colours, stripness as a trait, which is unclear its ecological meaning or comparing Asteracea (which tend to be yellow) with all other families. However, maybe even more important is to capture trait-matching traits. That is, there is a large literature suggesting the trait match between e.g. pollinator and flower size, is more explanatory than the absolute values of this traits. I think the work by Ben Weinstein on occupancy models and trait-matching on hummingbirds is not cited, and is a very good reference for this type of work.

* Carpenter bees chew the bottom of flowers
* We can cite more literature – these traits that people often look at or reference as important functional traits that modify or change their adaptability to various plants/conditions

Minor in line comments:  
  
-Both the abstract and intro start with pollinator declines, but the paper is about modelling interactions. I find this initial sentences too far from the paper topic, but this might be a matter of taste, so ignore the comment if you do not agree. specially lines 58-60 suggest we do not know why bees are declining, when I think there is a vast literature showing land use change, climate change, pesticides and other factors all contribute to the observed declines.  
  
-Lines 41-43 I find “source citation type“ is too specialised for the abstract if not defined there.  
  
-line 143: Here you mention yellow flowers are more generalised without a reference. Is this robustly tested?  
  
-line 158: there is an extra “that”.  
  
-line 175: It should be noted somewhere that bees do not see colours as we do, at least as a discussion point.  
  
-line 225: Nov 2020 seems a quite old data retrieving date. Did the number of records increase in the last two years?  
  
-line 373: Why 3%? Which is the justification?  
  
-line 376: Which kind of null models you use and why? This information is missing, and crucial to understand and interpret the method.  
  
-line399: The fact that 80% of species are without interactions in your dataset is indicative that data is not sufficient?  
  
-------------------------------------------------  
  
Reviewer #2 Comments to the Author:  
  
In the manuscript “Leveraging local and regional sources of data and a multi- species occupancy model to explore bee-plant interactions” the authors used multi-species occupancy models and data from Globi to evaluate potential plant-pollinator interactions and their predictors. The authors found that non-solitary bees were more likely to interact, as well as interactions coming from observations (e.g. iNaturalist) had a higher detection probability.  
  
Overall I really like this study. This is a novel way of using occupancy models, and the process the way the authors structured the analysis makes a lot of sense. I also commend the authors for presenting transparent assumptions.  
  
Major comments:  
  
- Can you provide more detail on how phenology is calculated? At the moment the methods only have: “Bee phenology was inferred using bee occurrence data downloaded from GBIF (GBIF 2020) from specimens collected on Santa Cruz Island.” But there is no mention of the method to estimate phenology. If the estimation is only done from the minimum and maximum dates found in historical data then you should use more robust methods to estimate phenology. Some papers you can look for methods are:  
  
Belitz, M. W., Larsen, E. A., Shirey, V., Li, D., & Guralnick, R. P. (2023). Phenological research based on natural history collections: Practical guidelines and a lepidopteran case study. Functional Ecology, 37(2), 234-247.  
Belitz, M. W., Larsen, E. A., Ries, L., & Guralnick, R. P. (2020). The accuracy of phenology estimators for use with sparsely sampled presence‐only observations. Methods in Ecology and Evolution, 11(10), 1273-1285.  
Pearse, W. D., Davis, C. C., Inouye, D. W., Primack, R. B., & Davies, T. J. (2017). A statistical estimator for determining the limits of contemporary and historic phenology. Nature Ecology & Evolution, 1  
  
- To clarify, in equation 2 the intercept has an i underscript. (ui) does this mean that the intercept was different for each bee species but not for each plant species? And if this is the case, can you justify why it was done this way? – also I have the same question for the intercept in detection.  
  
- Equation 5 has a fixed effect of month and this is justified by “We included a quadratic term for survey month, assuming that detection probability of bee- plant interactions peaks sometime during the year and then decrease.” I am wondering whether this assumption makes sense for all interactions. I would assume that each interaction has its ``phenology” and it should be separate across bee and plant species. Maybe this is worth doing as a random effect? (i.e. as the intercept was done).  
  
- For Figure 1, I am surprised that the possible number of interactions is so high for all of the bee species. Perhaps by estimating species phenology more robustly you can diminish some of the potential interactions?. Also can you explain why there is so little variation in the number of potential interacting plants for all bees?  
  
- L 419 The probability of detecting a bee-plant interaction was much higher for observational citations (e.g., iNaturalist) than for collections (e.g., museums). – How much is due to the initial distribution of the data? I could not find the number of how many observations were collected in iNaturalist vs. museums.  
  
- L426 We found a moderate effect of bee size and stripiness on bee-plant detection probability, where larger bees had a lower detection probability than smaller bees (Fig. 2; Table 1) and stripped bees had a slightly lower detection probability than not striped bees. – I find this result very counter intuitive, and in the discussion L475-478 the authors merely mention that this is in contrast to other sources, can the authors provide a hypothesis of why this is the case?  
  
Minor comments  
  
- In the line: “Lastly, we hypothesized that flower color and shape would play a role in bee-plant interaction probability because yellow flowers and bowl shapes or open flowers are more generalized than non-yellow or not bowl-shaped flowers.” I am not totally sure what the authors mean with more generalized.  
  
- L196 - how many more interactions did you find when you extended this range?  
Is it worthwhile filtering by habitat type?  
  
- I am confused about figures S2 and S3 because they say plant-bee detection probability but the columns just refer to the bees.  
  
- Figure S4 is really hard to see the species names.  
  
- L 429 contradicts the previous sentence regarding bee stripiness.  
  
Laura Melissa Guzman