Globi meeting

Tuesday
18 May 2021

Outline

- Formatting the data
- Preliminary analysis

Formatting the data

Info on data

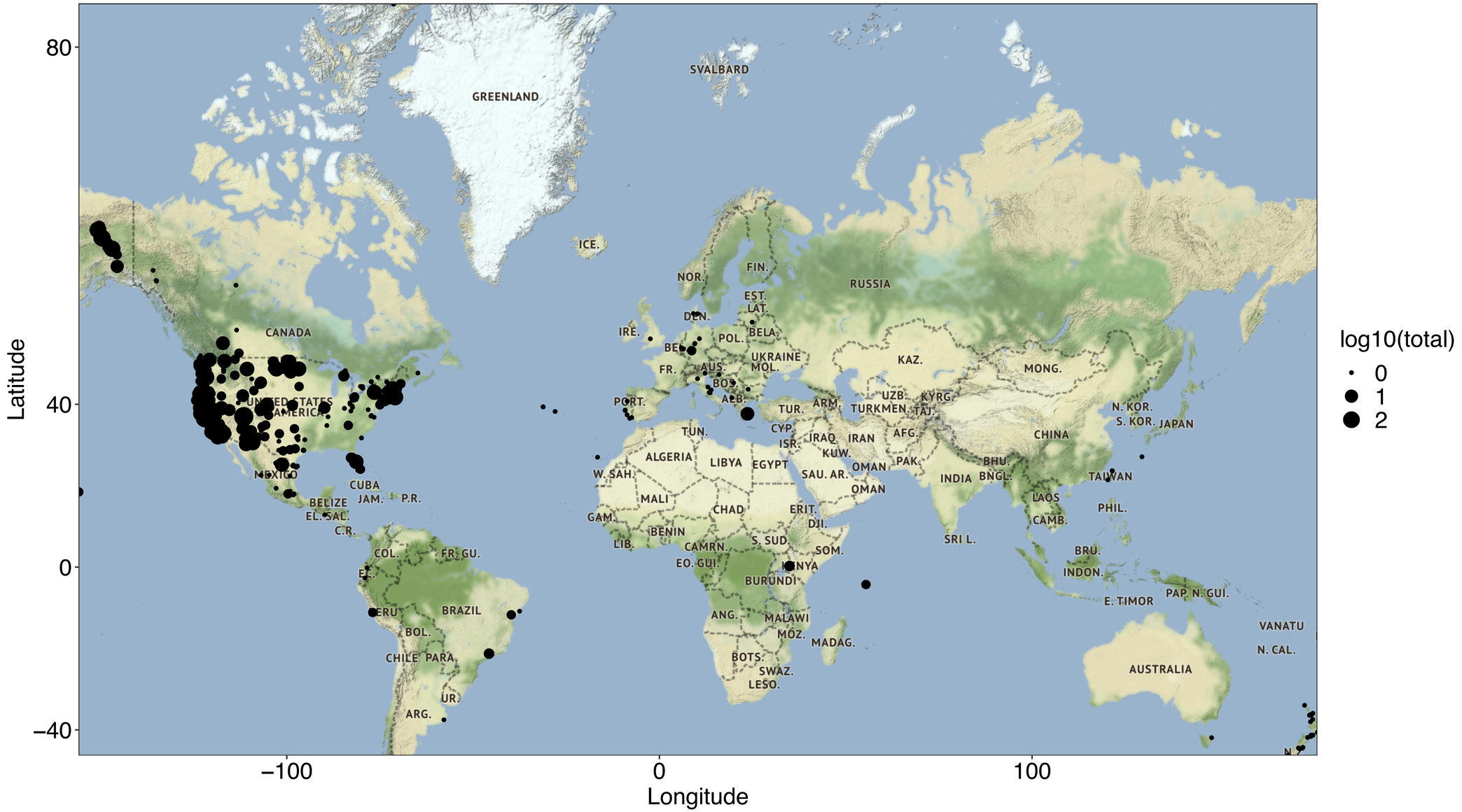
- Globi database has 304,795 entries
- Subsetting Globi to bee-plant interactions in our checklists = 8,755 entries
- Removing data with no lat/long info = 7,097 entries
    - *Where do the entries with no lat/long info come from?*

Entries with no lat/long data

- Come from 10 citations

1	Symbiota Collections of Arthropods Network (SCAN)
2	Seltmann, Katja C. 2020. Biotic species interactions about bees (Anthophila) manually extracted from literature.
3	National Database Plant Pollinators. Center for Plant Conservation at San Diego Zoo Global. Accessed via https://saveplants.org/national-collection/pollinator-search/ on 2020-06-05.
4	Global Web Database (http://globalwebdb.com): an online collection of food webs. Accessed via https://www.globalwebdb.com/Service/DownloadArchive on 2017-10-12.
5	Digital Bee Collections Network, 2014 (and updates). Version: 2015-03-18. National Science Foundation grant DBI 0956388; PBI: Phytophagous Insects as a Model Group for Documenting Planetary Biodiversity
6	Purdue Entomological Research Collection
7	A. Thessen. 2014. Species associations extracted from EOL text data objects via text mining.
8	University of California Santa Barbara Invertebrate Zoology Collection
9	California Academy of Sciences Entomology
10	https://mangal.io - the ecological interaction database.

Map of all data that matches checklists



To look at the species included in this map

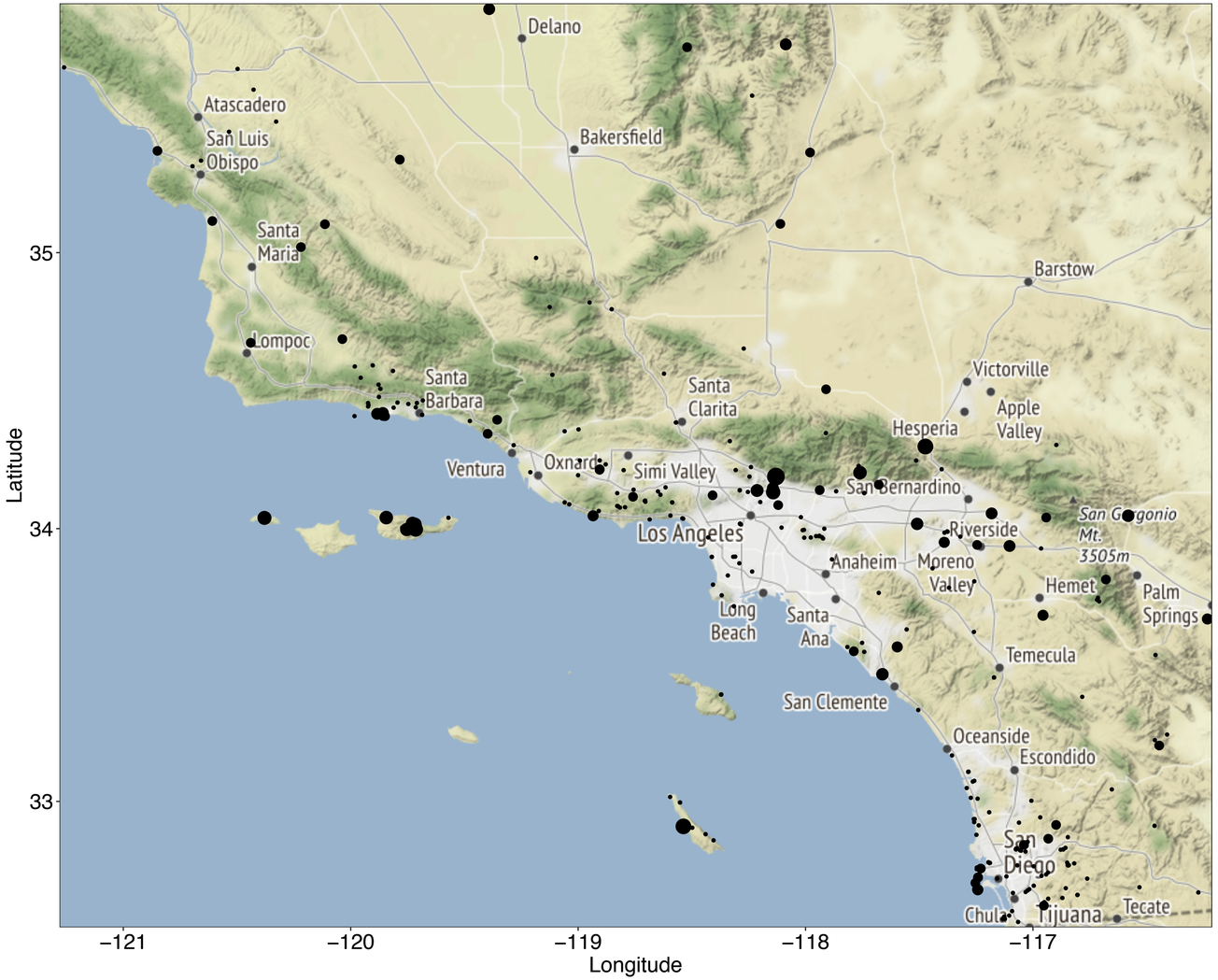
- Navigate to file: `species_names_Globi_map.csv`

Subsetting the data

- We should focus the analysis on a subset of data because the observers worldwide do not have an opportunity to document every possible interaction present in our checklists.

- Therefore, we should subset the data to a region where the observers have the possibility of documenting the bee-plant interactions in our checklists.

Subsetted CA data



What this the right region?
Check lower lab (goes to SD)

To look at the species included in this map

- Navigate to file: `species_names_Globi_CA_map.csv`

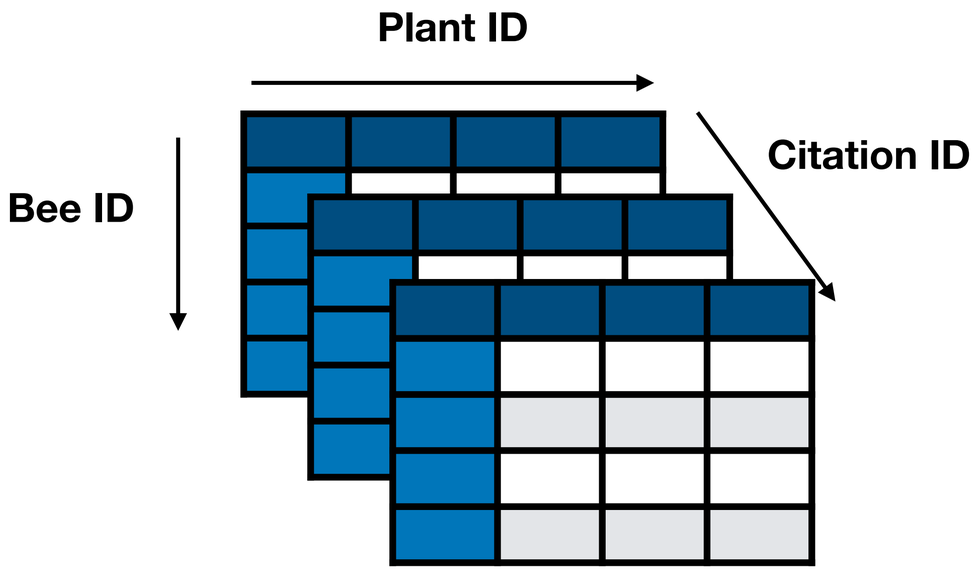
Citation list from CA subsetted data

- Come from 5 citations

1	University of California Santa Barbara Invertebrate Zoology Collection
2	Symbiota Collections of Arthropods Network (SCAN)
3	http://iNaturalist.org is a place where you can record what you see in nature, meet other nature lovers, and learn about the natural world.
4	Seltmann, Katja C. 2020. Biotic species interactions about bees (Anthophila) manually extracted from literature.
5	Digital Bee Collections Network, 2014 (and updates). Version: 2015-03-18. National Science Foundation grant DBI 0956388; PBI: Phytophagous Insects as a Model Group for Documenting Planetary Biodiversity (Insecta: Heteroptera: Miridae: Orthotylinae, Phylinae). Version: 08 Mar 2016. National Science Foundation grant DBI#0316495; Tri-Trophic Thematic Collection Network, 2014 (and updates). Version: 08 Mar 2016. http://tcn.amnh.org/. National Science Foundation grant(s) EF#1115081, EF#1115102
6	Pensoft Darwin Core Archives available via Integrated Publication Toolkit

Formatting the data for the analysis

- The analysis needs the data to be formatted in the 3-D array



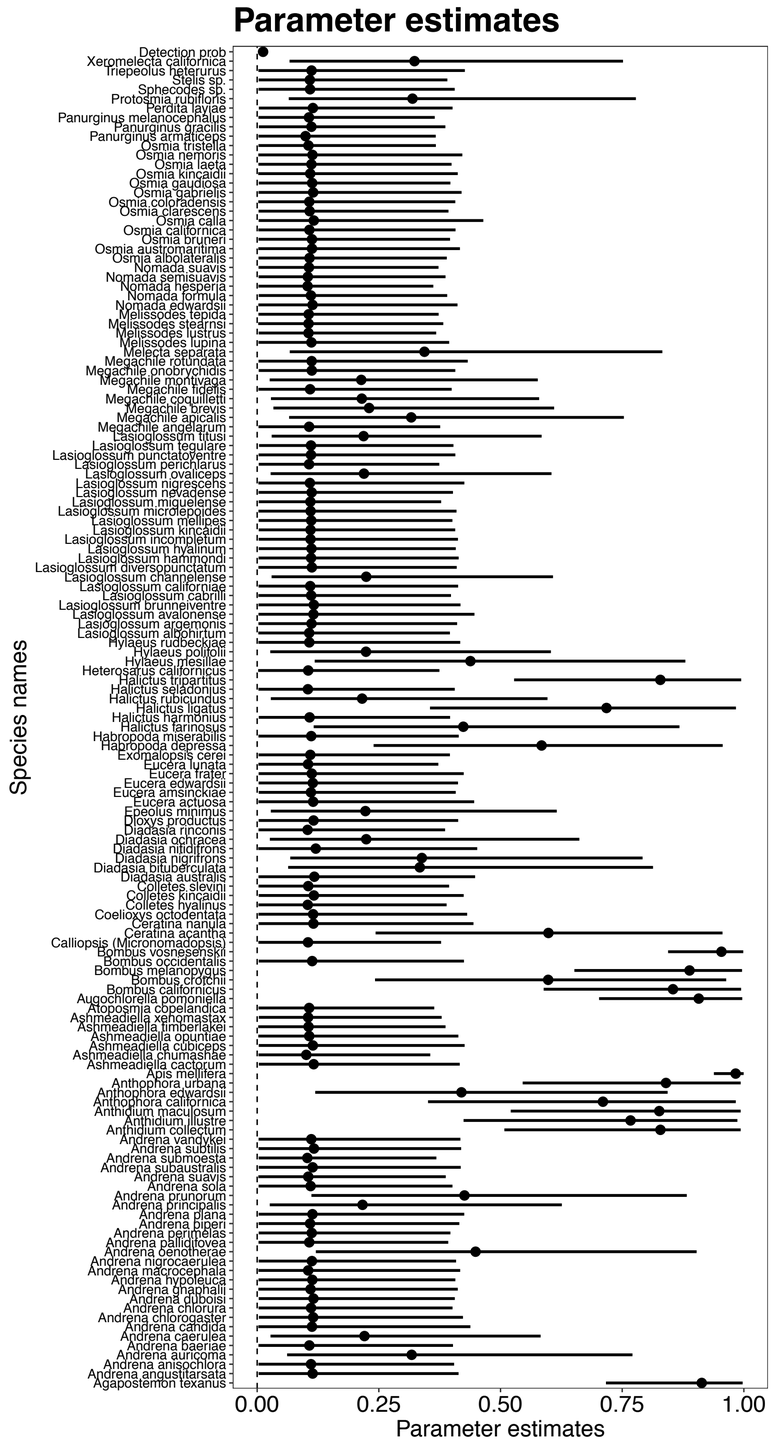
Size of the array

- 142 bee species * 1,186 plant species * 5 citations = > 1,000,000 cells!!!!
- There are a lot of plant species with NO documented bee interactions
  - 1,058 plant species with NO bee interactions
  - 128 WITH bee interactions
- 142 bee species * 128 plant species * 5 citations = more manageable

Things to consider moving forward

- Only include flowering plants in the plant checklist (i.e., remove grasses or other plants that bees do not interact with)
- What type of figure would help us answer our ultimate question?
- Are we adequately accounting for taxonomic sampling bias?
- List of covariates to include in the ecological model (what determines bee-plant interaction probability?)- covariates at bee or plant scales
- List of covariates to include in the detection model (what determines bee-plant detection probability?)- covariates at bee, plant, or citation scales

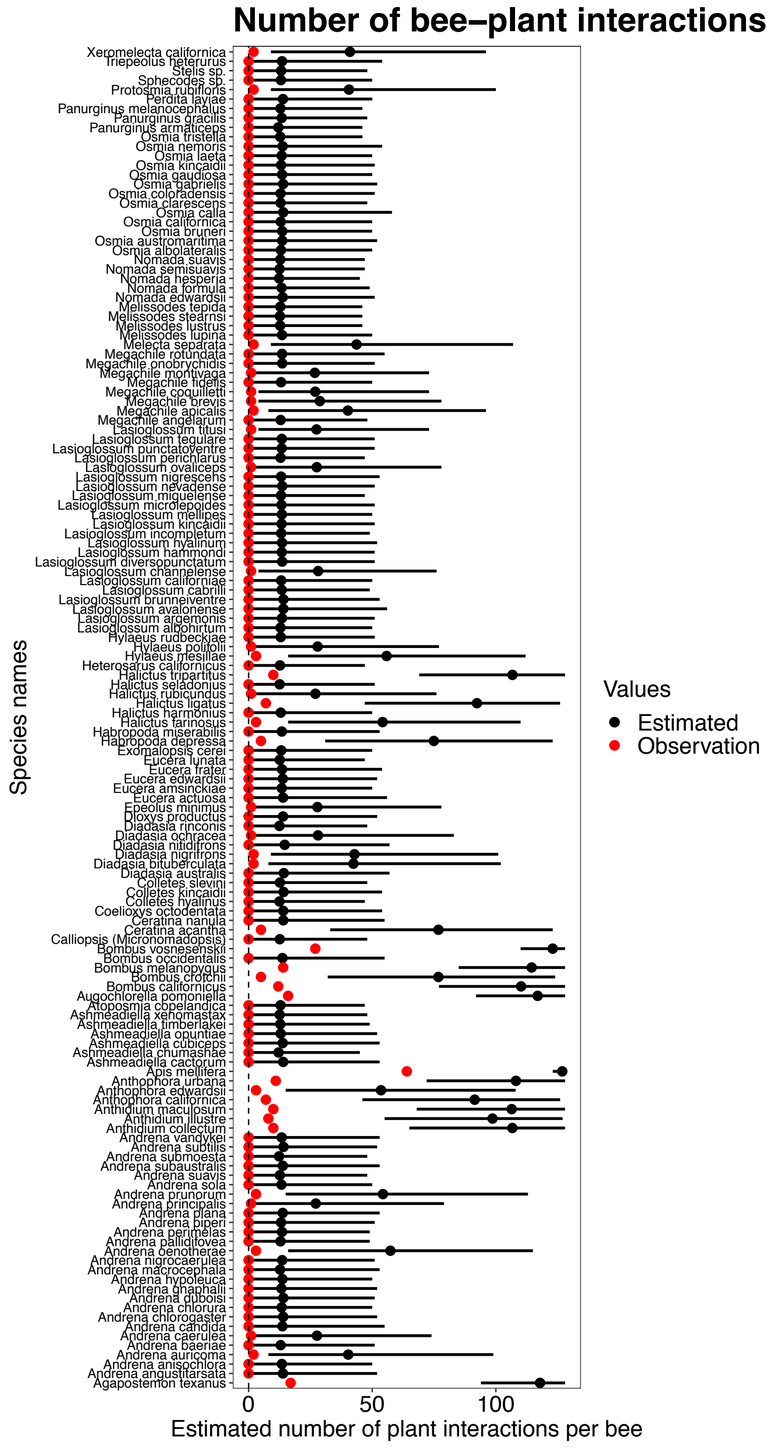
Preliminary analysis



Parameter estimates

Top = detection probability

All other = probability that the bee species interacts with different plant species



Number of bee-plant interaction

Red = observations

Black = mean and 95% Credible Interval

End