Page 1 of 1

**Project Name:**  Bee conservation paper

**Project Description:**  Assess the value of publicly available bee data for conservation assessments.

**Analysis completed by:** Josée Rousseau, with guidance from co-authors

**Project leader:** Josée Rousseau

**Start and end date:** 2021 - 2024

**Program(s) used:** Mainly R, also qGIS for final maps (cartography)

**ANALYSIS FLOW CHART**

|  |  |  |
| --- | --- | --- |
| **Data file(s) used** | **Program / Tool used** | **Comments (details, file locations, etc)** |
| **Download, format, filter, and compile data** | | |
| Downloaded records from SCAN: https://scan-bugs.org/ , from each of the 6 families of bees on August 24, 2022. | ./Programs/BeeData\_Conservation\_paper/Explore\_Format\_BeeData/ FilterExplore\_SCAN\_Bee\_datasets\_20231004.Rmd | Combines all SCAN files and select records for 2021.  Output:   * here("Data", "BeeData\_Conservation\_paper", " SCAN\_Bees\_2021\_USA\_20231004.csv") |
| SCAN\_Bees\_2021\_USA\_20231004.csv  Queried GBIF website to download records representing bees from all 6 families, located in USA, and from 1992 to 2021. Records were queried on August 24, 2022. File name = “GBIF\_Bees6Families\_USA\_1992to2021\_20220824”  Chesshire et al. bee data: supplement in her paper – saved: here("Data/BeeData\_Conservation\_paper/Chesshire\_2023/contiguousRecords\_high\_Only.csv") | ./Programs/BeeData\_Conservation\_paper/Explore\_Format\_BeeData/ CombineBeeDatasets\_ChesshireSCANGBIF\_20231004.rmd | Takes records from SCAN, GBIF, and Chesshire et al. (2023), combines them, removes duplicates, and assess the species that are not included in the publication from Chesshire et al. 2023 (Ecography).  Chesshire et al. contains data up to (and including) 2020. The new 2021 species were vetted by Bryan Danforth (Cornell University).  This program appends the 2021 data from SCAN and GBIF to the Chesshire dataset.  Filters used:   * Valid latitude and longitude, within the USA * Remove Honeybees (Apis mellifera) * Assess accuracy of location info and add value under new column ‘finalCoordinateUncertaintyInMeters’ when the latitude AND longitude had only 1 to 3 decimal places. * Removed records with location uncertainty of more than 15 km.   Output:   * here("Data", "Bioindicator\_Chesshire2021\_checklists\_easternHalfUSA", " bees\_Chesshire2021\_USA\_20231011.csv")   This program also summarizes and output the number of records per institutionCode and year. |
| **Summaries of bee species** | | |
| bees\_Chesshire2021\_USA\_20231011.csv | ./Programs/BeeData\_Conservation\_paper/Species\_Summaries/ perSpecies\_summaries\_20231010.rmd | Summarizes the number of records per species in the USA (not by state).  Species summaries:   * Number of decades for which we have records for each species, * Number of records, per species, and trap type, for past 10 years, * Number of records, per species, with effort, for past 10 years, * Number of unique locations, per species, for past 10 years.     Also assesses if we have enough records per species to complete a   * convex polygon, * species distribution model, and * occupancy models.   The number of species with sufficient records for each type of analysis is printed in the program. |
| bees\_Chesshire2021\_USA\_20231011.csv | ./Programs/BeeData\_Conservation\_paper/Species\_Summaries/mapBees\_USA\_20240224.rmd | Summarizes the number of bee records across the contiguous USA and years.  Specifically:   * Creates hexagon grid using 25 km apothem * Calculates number of bees per hexagon cell for the past ten years (2012 to 2021) and across all years * Plot histogram of number of records with protocol (trap type) and effort, with protocol only, without protocol and effort, and from iNaturalist (no records have effort only). * Extract summaries about iNaturalist data |
| bees\_Chesshire2021\_USA\_20231011.csv | ./Programs/BeeData\_Conservation\_paper/Species\_Summaries/beesPerStateSummaries\_20231010 | Summarizes species information per state, within the contiguous USA.  Specifically:   * Calculates total number of species per state, all years included * Calculates the number of species per state, per decade, for the most recent decades. * Creates a table of the bee species that were detected at least 3 times prior to 2012 but are not longer detected since 2012 |
|  |  |  |