GBIF

Integration of Biodiversity Data for the Management and Conservation of Wild Bee-Plant Interactions in Mexico

Programme: BID
Project ID: BID-CA2020-021-NAC
Project lead organization: National Commission for the Knowledge and Use of Biodiversity
Project implementation period: 1/7/2021 - 30/6/2023
Report approved: 20/6/2022

Narrative Midterm report

Executive Summary

The project progressed as it was planned and the midterm goals were reached. The achievements to date are 1) The data model for the integration of the ecological information about the wild bee-plant interactions in Mexico was finished; 2) The data model was designed with feedback received from potential model users who participated in the first workshop named "Towards to a data model for monitoring, management, and conservation of the pollination by wild bees in Mexico" (Spanish name: "Hacia un modelo de información para el monitoreo, manejo y conservación de la polinización por abejas nativas en México"); 3) Integration of information for data sets began from the sources proposed by the project (Naturalista, Literature and biological collections); 4) To data model deployment, the project agreement four specialists (two wild bees taxonomists, one plant taxonomist and one ecologist specialists); 5) One of our team members received the GBIF certification in Data Mobilization, and 6) We published the first data set obtained from the Naturalista platform. The last product contains metadata, 1716 occurrences for 265 taxa of the wild bee host plant, one table with ecomorphological traits data for 881 plant occurrences, and a table with ecological information about the wild bee-plant interactions. The monitoring and evaluation of the project's progress have been carried out with the review of the contract commitments of each specialist and researcher; these checks were carried out by CONABIO. The project was inserted into the National Strategy for the Conservation and Sustainable Use of Pollinators. Therefore, it has been guaranteed that the data model and data mobilization of wild bee-plant interactions will have an implementation in the medium and long term.

Progress against milestones

Has your project published at least one dataset through GBIF.org?: Yes

Dataset published:

Dataset	DOI
Biological records of potentially host plants of mexican wild bees identified from iNaturalist	https://doi.org/10.15468/m4r9h2

Has at least one member of your project team received certification following the BID capacity enhancement workshop?: Yes

Name of the workshop participant: Juan Martín Barrios Vargas

Certification obtained: Advanced Badge

Report on Activities

Activity progress summary

To date, we started all the activities committed and completed all the activities scheduled for the mid report, except that related to the publication of the datasets. In detail, one of the six datasets was published and includes 1716 records of plants identified from observations of native bees available in Naturalista. However, we made progress in data integration for all committed datasets (10% - 80% depending on the data set), and all of them are in the process of being reviewed for publication following GIBF standards. The publication of the pending datasets will be completed between September and December 2022. Thus, we have progressed in constructing the datasets that will mobilize information from the three sources considered by the project: Naturalist, Collections, and Literature. On the other hand, we completed the following activities: 1) the first workshop of the project "Towards a data model for monitoring, management, and conservation of the pollination by wild bees in Mexico"; 2) the design of the data model for integration biodiversity and ecological data, and 3) Human resource training began with the incorporation of four specialists in plant taxonomy, native bees and database management and ecological analysis. <u>Eliminated Activities</u>: A dataset will not be built from the collection of Kansas as the records are already available in GBIF. Instead, we will publish records on both wild bees and host plants visited with at least 2,000 (new) records available in the ECOSUR wild bee collection. These records will include information on the bees and plants visited.

Completed activities

Activity: Model design for integrating biodiversity and ecological data

Description: Definition of relevant information fields to integrate data on wild bees-plant interactions. These fields provide data on the type of interaction (e.g. floral visitor vs pollinator) and eco-morphological traits of both wild bees and plants (e.g. body size, life form, floral traits, plant type: native, exotic, cultivated). The data model design is an adaptive process between potential users and structures of old and new compiled data.
Start Date - End Date: 1/7/2021 - 30/6/2022
Verification Sources: GitHub - XicotliData/schemas: Schemas to save biological interaction data

Activity: Workshop 1: What do we need to know about wild bee - plant interactions?

Description: Organization of a workshop to identify data requirements from different stakeholders. The results from this workshop will be the guideline to the interactions data model specification and possibly the identification of new data providers that were not considered at the start of this project.
Start Date - End Date: 1/10/2021 - 30/11/2021
Verification Sources: Links: https://www.youtube.com/watch?v=BhK_8sJ4XAc&list=PLDJg7F5BO5iDuDSNtwrkelS7-I-KAA6qe ;
Documents:1. List of the assistants (PDF); 2. Summary report of the workshop 1 (PDF).

Activity: Human resource training

Description: The activities to carry out the proposed work will involve four interns in biology or data/computational science.
Start Date - End Date: 1/8/2021 - 31/5/2022
Verification Sources: Documents: the contracts for four profesionals intenrs (PDF).

Report on Deliverables

Deliverables progress summary

Below are details on deliverables completed or in process:

1) Dataset Title: Naturalista based wild bee-plant interactions dataset. To build this dataset, a project called 'Xicotli Data: Mexican bees and their flowers' was created within Naturalista. The compiled observations were processed according to the following protocol: 1) observations with photographs of bees perched on or near a flower were detected. For this task, Conabio's experts in biodiversity big data adjusted and implemented an artificial intelligence algorithm that selected the most relevant observations in a semi-automatized way, and 2) the taxonomic curation of the observations of bees and plants included in the photograph was carried out by a project team. In the case of the bee component, all observations were worked on, including those classified as 'research-grade' (i.e. when more than two-thirds of the identifiers agree on the identification). The experts on bees confirmed or corrected the taxonomic identifications previously made in Naturalista. Since there was no prior identification of the plants, the identification was done by the team's plant experts for the first time. For each observation, taxonomic and eco-morphological information was entered as Fields of Observation.

The observation fields were defined from a first proposal generated by the team, which was complemented by the feedback obtained in the workshop "Hacia un modelo de información para el monitoreo, manejo y conservación de la polinización por abejas nativas en México". Additionally, an illustrated handbook was designed for Naturalista users, which contains the list of fields, as well as the description of the type of information of each field, both for bees and host plants (see attachment documents). This handbook is expected to guarantee the continuity of the 'Xicotli Data project: Mexican bees and their flowers'.

2) Dataset Title: ECOSUR Mexican wild bee dataset. This dataset is being compiled from the Ecosur Bee Collection database (ECOAB), which has been compiled and partially curated by Dr. Remmy Vandame and various collaborators, including Jorge Mérida, and both are part of the team of this project. Records containing information about the bee and its host plant were filtered from the database. To this data, the experts in bees and plants reviewed the taxonomic determination and the ecological information. The Conabio's experts in biodiversity big data also participated, by verifying that the names of the bee and plant species agreed with the Conabio catalogs of species.

3) Dataset Title: Wild bee eco-morphological traits. The eco-morphological data of bee species are being integrated from three sources of information, Naturalista, collections, and bibliographic references. From the source of information of collections (Ecosur Bee Collection database -ECOAB -), a list of 313 species of bees from the Apidae family, 58 species from the Halictidae family, and 64 species from the Megachilidae family were extracted. So far, information for at least one of the 6 defined traits has been compiled for 80% of these species. This information is obtained from the specialized literature. Another source of information is the bibliographical references. An exploration of nearly 20 scientific articles yielded data on eco-morphological traits for 27 bee species from the families of interest. For Naturalista and collections the data integration focused on the features defined as follows. Based on an extensive literature review, internal work sessions, and feedback obtained from the project's first public Workshop "Hacia un modelo de información para el monitoreo, manejo y conservación de la polinización por abejas nativas en México", the team defined a set of eco-morphological traits of bees that meet the following criteria: 1) feasibility (traits that can be assigned relatively quickly and can be identified by anyone after brief training); 2) generality (traits that can be extracted for the vast majority of genus or species from more than one source of information), and 3) unambiguity (traits that can be precisely defined and cannot be confused with others). We generated a handbook that specifies the eco-morphological traits defined for bees and an illustrated guide for users of the Project 'Xicotli data: native bees and their plants'.

4) Mexican wild bee-plant interactions from literature. So far, 840 records of wild bee-host plant interactions have been integrated from bibliographical references, but we haven't started their publication yet. Virtually all interaction records obtained from this information source are checklists. In other words, very few occurrence records of bee-plant interaction can be obtained from the literature. Initially, a literature search was carried out on the Internet, according to a series of 'search words' defined expressly. The search was focused on scientific articles that included studies carried out in Mexico on species from three families of bees: Apidae, Halictidae, and Megachilidae. As a result, we found around 200 articles, which were reviewed and filtered to have a collection of wild bee-host plant interaction articles. The storage of the scientific articles and their management have been carried out using the specialized software 'Zotero'. After analyzing the type of information contained in these articles, together with the Conabio's experts in biodiversity big data, the relevant information for the project was defined. Additionally, the corresponding fields available in the DarwinCore standard were identified in order to integrate the information. A template was created to compile the information and tools were enabled to extract the information from the tables contained in the articles. For the extraction and integration of the data, the Conabio's experts in biodiversity big data developed a text mining tool that facilitates the identification, extraction, and curation of the information. Until now, only scientific articles have been reviewed, and information from the thesis and other types of documents will be incorporated later.

5) Wild bee host plants' eco-morphological traits. The eco-morphological data of host plant bee species are being integrated from three sources of information. The first is the Project 'Xicotli data: Mexican bees and plants', which was created in Naturalista (https://www.naturalista.mx/projects/xicotli-data-abejas-mexicanas-y-sus-flores). There, information on eco-morphological traits was integrated into 1,716 records of host plants. These data are published as an eco-morphological attributes table and are part of the first published dataset: 'Registros biológicos de plantas potencialmente nutricias de abejas nativas mexicanas identificadas a partir de observaciones de Naturalista' (https://zenodo.org/record/6522118#.YoVpUKjMLIU). The second source of information is the Ecosur Bee Collection database (ECOAB). From the ECOAB database, a list of the 50 species of host plants of bee species of the Apidae family was extracted. The ECOAB database previously contained information about a set of eco-morphological traits of plants (corolla color, flower size, etc.). These traits were complemented with those defined for the "Xicotli Data: Native Bees and their flowers" project. This information will continue to be compiled by plant experts and the list of host plants will be expanded for the other two families of bees in the project: Halictidae and Megachilidae. This information is obtained from the specialized literature. The third source of information is the

bibliographical references. An exploration of nearly 20 scientific articles yielded data on eco-morphological traits for 20 plant species. For the first two sources of information (i.e., Naturalista and collections) the data integration focused on the features defined as follows. Based on an extensive literature review, internal work sessions, and feedback obtained from the project's first public Workshop "Hacia un modelo de información para el monitoreo, manejo y conservación de la polinización por abejas nativas en México", the team defined a set of eco-morphological traits of plants that meet the same criteria defined for bees: 1) feasibility (traits that can be assigned relatively quickly and can be identified by anyone after brief training); 2) generality (traits that can be extracted for the vast majority of genus or species from more than one source of information), and 3) unambiguity (traits that can be precisely defined and cannot be confused with others). We generated a handbook that specifies the eco-morphological traits defined for plants and an illustrated guide that makes it easier for users of the Project 'Xicotli data: native bees and their plants' to correctly fill in the fields.

6) Wild bee – Plant interactions data model. The process to generate the data model consisted in, first, defining the bee's features relevant for identifying the type of interactions with host plants (mainly morphological features). These features provide data on the type of interaction (e.g. floral visitor vs pollinator), and eco-morphological traits of both bees and plants (e.g. body size, life form, floral traits, plant type: native, exotic, cultivated). We generated a first version of the list of the features which was completed according to the feedback received by the attenders of the workshop "Hacia un modelo de información para el monitoreo, manejo y conservación de la polinización por abejas nativas en México". The data model maximizes the use of the DarwinCore fields and GBIF – Extensions, and also includes tools available in Plinian Core and Audubon Core, the latter designed to represent metadata for biodiversity multimedia resources and collections. The data model fulfilled the project's main objective, which consists in providing a dynamic source of information for the study, management, and conservation of pollination, and the monitoring of the wild bee-plant interactions.

7) Workshop 1: What do we need to know about wild bee-plant interactions? Workshop. The final title of the workshop was "Hacia un modelo de información para el monitoreo, manejo y conservación de la polinización por abejas nativas en México". The workshop was held on January 13, 2022 in virtual mode (10:00 - 16:30 hrs) (v=BhK_8sJ4XAc&list=PLDJg7F5BO5iDuDSNtwrkelS7-I-KAA6qe). The workshop was attended by 70 persons, representing 5 sectors of society, education (52% of attendees); research (20%), agriculture, beekeeping-meliponiculture (15%), government (10%) and others (3%). Nearly 35 institutions and different regions of the country (15 states of the Mexican Republic) were represented in it. As part of the preparations for the workshop, a national survey to identify the information needs of different potential users of the datasets were designed and applied (https://docs.google.com/forms/d/1QhDX3A691ndcc29PsOOSfI39fIJCK66pCKyqqqmT7H8/edit). The survey was answered by 203 people. The aim of the workshop - to identify data requirements from different stakeholders-, was successfully fulfilled and that information was used to complement our first data model version. Also, the workshop allowed the identification of key actors and potential users of the data, the first step for an Inter-institutional and multi-sectoral collaboration. Additionally, the workshop is one of the activities associated with the Mexican National Strategy for the Sustainable Use of Pollinators (ENCUSP), particularly in strengthening the monitoring and conservation of the wild bee-plant interaction.

Progress towards deliverables

Dataset deliverables

Naturalista based wild bee-plant interactions dataset
Dataset type: Occurrences
Dataset scope: Regarding native bees, the taxonomic determination will be carried out to the genus level, and if possible, down to the species level. Regarding plants, the determination will be done to the Family and Genus level.
Number of records: 1,716
Data holder: GBIF
Data host institution: CONABIO
% complete: 50%
Status update: Published one dataset
DOI: https://doi.org/10.5281/zenodo.6522118
Expected date of publication:

ECOSUR Mexican wild bee dataset
Dataset type: Occurrences
Dataset scope: Specimen collected dataset with information about plant interaction. This collection is focused on Mexican native wild bees, their trait information and interactions.
Number of records: 3,000
Data holder: ECOSUR
Data host institution: ECOSUR
% complete: 80%
Status update: The number of records of wild bee-host plant interactions that were integrated and curated in this dataset was 2,404, all belonging to species of the Apidae family

DOI:

Expected date of publication: 2022-07-31

Wild bee ecomorphological traits

Dataset type: Checklist

Dataset scope: Mexican wild bee checklist. Describe wild bee ecomorphological traits useful for the identification of their type of interaction with plant species. Record findings of bee-plant interactions.

Number of records: 400

Data holder: SNIB CONABIO; entomology collections from INECOL, CIIDIR-Oaxaca, ECOSUR, Faculty of Science - UNAM; the National Collection of Insects, Institute of Biology, UNAM

Data host institution: CONABIO

% complete: 20%

Status update: From the source of information of collections (Ecosur Bee Collection database - ECOAB -), a list of 313 species of bees from the Apidae family, 58 species from the Halictidae family, and 64 species from the Megachilidae family were extracted.

DOI:

Expected date of publication: 2022-12-15

Mexican wild bees-plant interactions from literature

Dataset type: Occurrences

Dataset scope: From each bibliographic source, the scientific names of the wild bees and host plants will be extracted together with the geographic data and the type of the interaction.

Number of records: 150

Data holder: Literature (articles, technical reports, thesis, books)

Data host institution: CONABIO

% complete: 10%

Status update: 840 records of wild bee-host plant interactions have been integrated from bibliographical references, but we haven't started their publication yet.

DOI:

Expected date of publication: 2022-09-30

Wild bee host plants' ecomorphological traits

Dataset type: Checklist

Dataset scope: Describe ecomorphological traits of plants for the identification of syndromes and the potential type of interaction with wild bee species. Characteristics as the habit of a plant, the flower (shape, color, symmetry), the floration season will be included. Record findings of bee-plant interactions.

Number of records: 200

Data holder: SNIB CONABIO, herbariums from UNAM and INECOL, literature (articles, technical reports)

Data host institution: CONABIO

% complete: 70%

Status update: This data is not yet ready to be published. But the progress of its construction is included as a complementary table of the first dataset published by the project. Adiotionally, plants eco-morfological data was integrated from the ECOSUR collection (25 species).

DOI:

Expected date of publication: 2022-11-30

Other deliverables

Wild bee – Plant interactions data model

Description: Data model that defines wild bee's features relevant for identifying type of interactions with host plants (mainly morphlogical features). Data model for wild bees' host plants that helps identify syndromes and relations with each species of wild bee. Data model for wild bee - plant interactions, this model defines all the features instrinsic to an interaction.

% complete: 100%

Status update: Done

Sources of verification: GitHub - XicotliData/schemas: Schemas to save biological interaction data

Workshop 1: What do we need to know about wild bee - plant interactions? Workshop

Description: Workshop to identify data requirements from different stakeholders. The result from this workshop should be a guideline to the interactions data model specification.

% complete: 100%

Status update: Done

Sources of verification: Link: v=BhK_8sJ4XAc&list=PLDJg7F5BO5iDuDSNtwrkelS7-I-KAA6qe
Documents: 1. List of the assistants (PDF); Summary report of the workshop 1 (PDF).

Events

Open-House

Dates: 2021-09-30 - 2021-09-30
Organizing institution: Instituto de Ecología (INECOL)
Country: México
Number of participants: 300
Comments: Open House is an annual scientific divulgation event aimed at the general public. In 2021 it was held virtually. Our participation was carried out with an interactive talk entitled: "Xicotli Data abejas silvestres y sus plantas" (Xicotli Data: Wild Bees and their plants).
Website or sources of verification: https://youtu.be/lzw5WbKXV1Q?t=1780

Events

Workshop "Toward a data model for monitoring, management, and conservation of the pollination by wild bees in Mexico"

Dates: 2022-01-13 - 2022-01-13
Organizing institution: CONABIO, INECOL, ECOSUR, and CIIDIR-IPN Oaxaca
Country: Mexico
Number of participants: 70
Comments: The workshop was attended by 70 assistants, representing 5 sectors of society: education (52% of attendees), research (20%), agriculture, beekeeping-meliponiculture (15%), government (10%), and others (3%). Nearly 35 institutions and different regions of the country (15 states of the Mexican Republic) were represented in it.
Website or sources of verification: https://www.youtube.com/watch?v=BhK_8sJ4XAc&list=PLDJg7F5BO5iDuDSNtwrkelS7-I-KAA6qe

Events

Survey: Wild Bees and their flowers

Dates: 2022-01-01 - 2022-06-30
Organizing institution: INECOL, CONABIO, ECOSUR, CIIDIR-IPN Oaxaca
Country: Mexico
Number of participants: 204
Comments: The survey was answered by 203 people. The aim of the workshop - to identify data requirements from different stakeholders-, was successfully fulfilled and that information was used to The survey was answered by 203 people. The aim of the workshop - to identify data requirements from different stakeholders-, was successfully fulfilled and that information was used to complement our first data model version. Also, the survey allowed the identification of key actors and potential users of the data, the first step for an Inter-institutional and multi-sectoral collaboration.
Website or sources of verification:
(https://docs.google.com/forms/d/1QhDX3A691ndcc29PsOOSfI39fIJCK66pCKyqqqmT7H8/edit) and (https://docs.google.com/forms/d/1_rNDY_iuI0AH0LjeIYiVRjv4xJ5Aw3XD0RaOz92Hbi8/edit)

Events

Mexico-Mesoamerica Bee Biodiversity Initiativemeeting

Dates: 2021-09-13 - 2021-05-13
Organizing institution: Scientists and citizen initiative
Country: USA
Number of participants: 27
Comments: This initiative brought together a large group of specialists in studying bees native to North America. Our participation focused on highlighting the use of the DwC standard to integrate and mobilise ecological data from bee-plant interactions. Thus, we managed to provide further visibility to our project.
Website or sources of verification: https://youtu.be/DWxMe8wtUyM?t=3321

Communications and visibility

The results of the project are being communicated through different activities: First of all, the workshop "Towards to a data model for monitoring, management, and conservation of the pollination by wild bees in Mexico" (Spanish name: "Hacia un modelo de información para el monitoreo, manejo y conservación de la polinización por abejas nativas en México" brought together key potential users,

which strategically made it possible to publicize the existence of this project, socialize its objectives, publicize the Naturalistra project, Xicotli data: las abejas mexicanas y sus flores. The second workshop contemplates reinforcing the dissemination of the results of the project with this group of participants and extended to other institutions. The results of the workshop, as well as dissemination activities such as the INECOL Open House, are duly available for public consultation shared with the project stakeholders and broader GBIF community on our website (https://www.gbif.org/ project/BID-CA2020-021-NAC/integration-of-biodiversity-data-of-wild-bee-plant-interactions-in-mexico). There, we provide additional information including links to resources.

Monitoring and evaluation

Monitoring and evaluation findings

The project has progressed as planned, with most midterm goals being achieved. The data model for integrating the ecological information about wild bee-plant interactions in Mexico was finished, with key stakeholders and potential users being involved in the process. The data model fulfilled the project's main objective, providing a dynamic source of information for the study, management, conservation, and monitoring of wild bee-plant interactions. The integration of information from the three sources (Naturalista, Literature, and collections) had advanced as planned. The necessary adjustments related to the availability of the data and the technical details for its publication in the GBIF platform were conducted. Thus, we developed a plan to complete the number of wild bee-plant interaction records pledged in the project. We published the first data set from the Naturalista platform. The process of preparing and publishing the dataset has been done in close interaction with the assistance of the project mentor, Leonardo Buitrago. We want to highlight that in this period, the team acquired the necessary training, experience, and eco-informatic tools that will allow it to intensify the search and integration of information and fulfil the commitments planned for the next stage. We also want to highlight that we have identified key actors and potential data users in this period, this is the first key step for an Inter-institutional and multi-sectoral collaboration. Additionally, the workshop we conducted allowed our project activities to become associated with the Mexican National Strategy for the Sustainable Use of Pollinators. Our participation in this strategy will help to strengthen the monitoring and conservation of wild bee-plant interactions in the country.

Impact of COVID-19 pandemic on project implementation

The pandemic has not had negative effects on the project, since this situation was taken into account in its planning. The workshop was planned as a face-to-face or online activity, depending on the circumstances. The fact that it was carried out online allowed people from 15 different states of the Mexican Republic to participate, that is, probably more attendees from the province than if it had been face-to-face. The team meetings have been online which is not a problem, given the nature of the project. On the contrary, permanent online meetings have facilitated communication between project participants who are in 6 different cities.