Supporting Information. Bloom, Elias H., Elisabeth C. Oeller, Rachel L. Olsson, Matthew R. Brousil, Robert N. Schaeffer, Saumik Basu, Zhen Fu, and David W. Crowder. Documenting pollinators, floral hosts, and plant–pollinator interactions in U.S. Pacific Northwest agroecosystems. Ecology.

Metadata S1

Open Research Statement:

Data available on Figshare. Bloom, E; E. Oeller; R. Olsson; R. Schaeffer; M. Brousil; S. Basu; Z. Fu; and D. Crowder. 2021. Bloom et al. Ecology data paper: Documenting pollinator communities, floral hosts, and plant-pollinator interactions in Pacific Northwest United States agroecosystems. figshare. Dataset. https://doi.org/10.6084/m9.figshare.14179640.v4

Abstract:

The abundance and diversity of pollinator populations are in global decline. Managed pollinator species, like honey bees, and wild species are key ecosystem service providers in both natural and managed agroecosystems. However, relatively few studies have exhaustively characterized pollinator populations in diverse agroecosystems over multiple years, while also thoroughly documenting plant-pollinator interactions. Yet, such studies are needed to fulfill the national pollinator protection plans that have been released by the United States and other nations. Our research is among the first studies to respond to these directives by systematically documenting bee and plant biodiversity, bee-plant interactions, and bee-mediated pollen movement in farming systems of the Pacific Northwest, USA. Our data provides insight into the processes mediating pollinator and plant community assembly, persistence, and resilience across landscapes with variable crop and landscape diversity and agroecosystem management practices. These data will also contribute to the development of a United States pollinator database, supporting the United States' plan to promote pollinators. With few publicly available datasets that systematically take account of agroecosystem practices, plant populations, and pollinators, our research will provide future users the means to conduct synesthetic studies of pollinators and ecosystem function in a period of rapid and global pollinator declines. There are no copyright or proprietary restrictions for research or teaching purposes. Usage of the data set must be cited.

INTRODUCTION

Wild and managed bee pollinators are critical to angiosperm reproduction as well as crop productivity, promoting billions of dollars in pollination services within agroecosystems each year (Losey and Vaughan 2006, Klein et al. 2007, Potts et al. 2010). The conservation of these pollinators is of utmost concern, given that many wild and managed bee species are in decline where the delivery of pollination services are most critical, creating a mismatch in pollination needs and provisioning of these services (Koh et al. 2016). The impact of pollinator decline in agronomic ecosystems is particularly evident in crops where pollination services provided by managed pollinators (e.g., the western honey bee, *Apis mellifera* L.) are supplemental to those provided by their wild bee relatives (Garibaldi et al. 2013). While evidence for pollination limitations driven by pollinator declines in farming systems remains scant (but see Isaacs and Kirk 2010), the global decline of many terrestrial insects is concerning considering the importance of all insects for ecosystem function (Van Klink et al. 2020).

Coordinated efforts to document, conserve, and promote global pollinator populations have gained momentum in recent years, with at least 16 nations enacting pollinator protections plans (FAO 2020). The United Kingdom, for example, has a centralized database for pollinator data that can be used by researchers to track the status of pollinator populations over space and time (BWARS 2020). Akin to other nations, the United States released the Pollinator Research Action Plan in 2015, with research and funding directives to study and support wild and managed bee pollinators in natural and agronomic ecosystems (PHTF 2015). Our research is among the most expansive studies to respond to this directive, as we systematically documented the biodiversity of bee and plant populations, bee-plant interactions, and bee-mediated pollen movement in agroecosystems over broad spatial and temporal scales in the Pacific Northwest United States.

The digitization of pollinator collections is creating rich spatio-temporal datasets that may reveal historical patterns to aid in pollinator conservation. For example, historical datasets were used to elucidate drivers of bumble bee decline across Europe and North America due to climate change, demonstrating the importance of freely available pollinator data (Soroye et al. 2013). Lacking, however, are datasets that include a systematic accounting of agricultural management practices, plant populations, and pollinators. Such resources are useful in synthesis studies examining how pollinators respond to agricultural practices (Kennedy et al. 2013). By providing such a resource, our data can be used to gain novel insights into processes that promote pollinators and conserve ecosystem function in a period of rapid global pollinator decline (Sánchez-Bayo and Wyckhuys 2019). In particular, our dataset provides unique insight into plant-pollinator interactions, pollen transport by bees, bee pollinator and plant biodiversity, and organic management practices, over a spatial extent of more than 7,000 km^2 across 36 farms and over 200 sampling events. By making our work freely available, we also allow for the accumulation of our data in the United States wild bee pollinator database currently underway, thus adding to evidence of pollinator populations necessary for conservation at the national scale (PHTF 2015).

Decades of research suggest that the conservation of resilient, stable, and productive natural and agronomic ecosystems relies on the conservation of natural resources including bee pollinators

(Williams et al. 2019). Pollinators are critical to human health and the wellbeing of our planet more broadly (Eilers et al. 2011), and preventing pollinator declines presents itself as one of the greatest challenges facing humanity but also promises opportunities to combine ecosystem benefits. Indeed, the practices that support pollinators often enhance other beneficial insects (e.g., insect predators and biological control agents) and have links to the sequestration of carbon and the mitigation of climate change (Wratten et al. 2012). Our data provides novel insights into these concepts along with basic questions, including the mechanisms of community assembly, beta diversity, and biodiversity ecosystem function relationships (Bloom et al. 2019). By making these data publicly available, we also contribute to the larger longer-term goal of global pollinator conservation within agronomic and natural ecosystems in a changing world.

METADATA

CLASS I. DATA SET DESCRIPTORS

A. Data set identity: Western Washington, USA, Plant-Pollinator Community Dataset

B. Data set identity code: PollCommData.zip

C. Data set description: This dataset includes bee pollinator community data, plant community data, pollen network data, and pollinator visitation data from small, diversified farm sites located in the Puget Lowland of western Washington state, USA. Data collection for bees, plants, and pollen was conducted at three time periods per year, with all site measures performed on the same date per site. Farm sites were selected based on organic status (all farms were either certified organic or used certified organic production practices), time in organic management (each farm representing a different number of years in organic farming), urbanization (farms were selected along a rural to urban gradient), and distance from other farms (all farms were separated by 2 km). Thus, this dataset can be used for syntheses that examine how crop diversification and time in organic management can influence pollinator communities. For those conducting syntheses on the response of pollinators to landscape context, we include parameters describing urbanization within 1 km of each site and the approximate GPS location of each farm. By including location information, we allow the user to perform à la carte landscape analyses with the user's choice of software, landscape metrics, and questions. Bee pollinator sampling was conducted using bee bowls, traps, and netting during the spring, summer, and fall of 2014 to 2016 at each site. Bees were identified to species. Floral abundance and richness were measured on each farm in transects during the spring, summer, and fall of 2014 to 2016. Floral abundance data can be used to approximate the richness of crops within the farming system (crop management and diversification). Flowering plants were identified to species. Pollen species found on bees collected during the spring, summer, and fall of 2014 to 2016 were identified using amplicon sequencing. Bee visitation was observed in the spring, summer, and fall of 2014 to 2016. Due to logistics (the spatial extent of the study > 7,000 km^2), active sampling of bees (netting), measures of floral diversity, pollen species, and bee visitation were timed along a serpentine transect, rather than conducted in a fixed plot (e.g., quadrat). Fixed plots were impractical because these were working farms. However, because sites were generally small (mean site size = 2.88 ha), the measures from most farms represent site level samples, and each

serpentine transect traversed a similar area. To further control for this issue of sampling intensity, we include data on site size (which can be used as an offset in modeling procedures) and perimeter. Furthermore, this allows those interested in the response of pollinators to the dimension of each site to assess these relationships (e.g., edge effects and species area relationships), which would otherwise be unobtainable due to farmer privacy concerns.

Principle Investigators:

Elias H. Bloom, Dept. of Entomology, Michigan State University, East Lansing, MI, USA

David W. Crowder, Dept. of Entomology, Washington State University, Pullman, WA, USA

Abstract:

The abundance and diversity of pollinator populations are in global decline. Managed pollinator species, like honey bees, and wild species are key ecosystem service providers in both natural and managed agroecosystems. However, relatively few studies have exhaustively characterized pollinator populations in diverse agroecosystems over multiple years, while also thoroughly documenting plant-pollinator interactions. Yet, such studies are needed to fulfill the national pollinator protection plans that have been released by the United States and other nations. Our research is among the first studies to respond to these directives by systematically documenting bee and plant biodiversity, bee-plant interactions, and bee-mediated pollen movement in farming systems of the Pacific Northwest, USA. Our data provides insight into the processes mediating pollinator and plant community assembly, persistence, and resilience across landscapes with variable crop and landscape diversity and agroecosystem management practices. These data will also contribute to the development of a United States pollinator database, supporting the United States' plan to promote pollinators. With few publicly available datasets that systematically take account of agroecosystem practices, plant populations, and pollinators, our research will provide future users the means to conduct synesthetic studies of pollinators and ecosystem function in a period of rapid and global pollinator declines. There are no copyright or proprietary restrictions for research or teaching purposes. Usage of the data set must be cited.

D. Keywords:

agroecosystems, biodiversity, honey bees, insect declines, plant-pollinator interactions, pollinator policy, species interactions, wild bees

CLASS II. RESEARCH ORIGIN DESCRIPTORS

A. Overall project description

Identity: Western Washington, USA, Plant-Pollinator Community Dataset

Originators: Elias H. Bloom

Period of study: Project commenced May 2014, terminated September 2016

Objectives: Our objectives for collecting these data were to determine (i) the pollinator community, (ii) flowering plant community, (iii) pollen transported by bee pollinators, (iv) bee-plant interactions, and (v) compile these data along with agroecosystem practices to evaluate relationships between these practices and pollinators.

Sources of funding: Funding for this work was provided to EB and DC by NSF GROW (Grant Number: 121477-007), USDA Organic Transitions Program (Grant Number: 2014-51106-22096), USDA Predoctoral Fellowship (Grant Number: 2017-67011-26025), NSF GRFP Grant Number: (124006-001), Western SARE Graduate Student Grant (Grant Number: GW15-022), and an internal development grant from James Cook University, Cairns, QLD, AU.

System description: This study took place on 36 organic farms in western WA, USA, spread across a spatial extent of over 7,000 km^2. This region, commonly known as the Puget Lowland (spatial extent > 30,000 km^2), represents a matrix of land cover types [natural: 58.01%; urban: 34.5%; crop: 4.6%; other: 2.89% (USDA-NASS 2020)], and is bound by the Cascade Range, Olympic Mountains, San Juan Islands, and the terminus of the Puget Sound, in the east, west, north, and south, respectively (DNR 2021). With consistent soils (sandy loam and loam) and a maritime climate (Cuo et al., 2009), the Puget Lowland supports a diversity of small fruit and vegetable farms. Indeed, twenty five percent of organic farms within Washington state are found in this region (Granatstein & Kirby 2019). These agroecosystems are typified by small (e.g., < 13 ha in production), diversified farms (e.g., > 5 crops in simultaneous production), that use organic methods (Kirby & Granatstein 2013). Time in organic production can vary substantially across farms, ranging from transitional (0 - 3 years) to long term (> 33 years), with some practicing organic methods prior to the establishment of the Washington State Organic Program in 1987 (WSDA 2020). Furthermore, farms and urban gardens in this region also use organic methods without certification (which we term as farms that use organic practices). Over the last half century, urban expansion within the Puget Lowland has modified considerable portions of this landscape from forest to urban (city) landscapes, particularly along the corridor from Mount Vernon, WA to Olympia, WA, in the north and south, respectively (Cuo et al., 2009). However, farmlands remain prevalent in more densely rural areas to the south and east/north of Olympia and Seattle, WA, respectively (DNR 2021).

B. Survey design

Site selection and evaluation: Within the Puget Lowland, our research focused on selecting farms along an urbanization and time in organic management gradient. The response of pollinators to urbanization is increasingly well established (Wenzel et al., 2020). However, the general influence of long-term organic farming on pollinators or arthropods is unknown, making this covariate particularly foundational for syntheses studying the response of biota to organic farming over time. Practically, time in organic management is also non-trivial, because organic practices are difficult for farmers to implement and it remains unclear how the benefits (e.g., in terms of yield) change over time (Reganold and Wachter 2016). Each farm was required to be

certified organic, or use only certified organic production practices, and produce more than five different flowering crop plants simultaneously. The response of pollinators to the management of co-flowering crops and weeds, particularly on a theoretical basis, is not well understood (but see Bloom et al., 2019). Practically, however, diversification of farming systems also represents management. In other terms, weedier farms and those that are more diverse indicates a farmers pest management approach and implementation of a suite of different practices (e.g., nutrient management) based on the crops grown. Contact information of farms that met the aforementioned description were gathered through the Washington State Department of Agriculture list of certified organic farming systems (WSDA 2020), United States Department of Agriculture integrity database (USDA 2020), Seattle Department of Neighborhoods (DON 2020), and regionally specific web searches particularly for farming systems which practiced organic methods and therefore would not appear in certified organic databases. Farmers were contacted and we conducted exhaustive on-site evaluations to verify practices, create an even distribution of sites along gradients (see above), and to ensure site spatial independence (farms were required to be 2 km apart), yielding a set of 36 farms.

Urbanization was verified by calculating the percent of "developed" pixels within 1 km of each farm per year using the Cropland Data Layers (*see* USDA-NASS 2014; mean = 49.02, median = 47.52, range = 3.14 - 99.51 %). These raster layers classify single habitat types within 30 × 30 m grid cells. Additional farm characteristics were collected using geographic information systems including location (latitude and longitude), area (farm size: mean = 2.88, median = 0.97, range = 0.03 - 22.25 ha), and perimeter (length of total farm edge: mean = 688.22, median = 456.00, range = 159 - 3461 m). Organic certification and the length of time each farm had been managed organically was determined from site operators; farms with less than 1-y in organic production were given a value of 0 (organic certification: certified organic = 20, organic practices = 16 farms; time in organic management: mean = 14.67, median = 10.5, range = 0 - 43 years).

Bee diversity sampling: Bee richness and abundance were measured during three time periods (May, July, September) at each of 36 farm sites across three years (2014 - 2016). Not all farms were sampled each year, with 23 sampled in 2014, 35 in 2015, and 22 in 2016 (3 time periods × 80 sites = 240 samples). Three blue vane traps (SpringStar LLC, Woodinville, WA, USA) and 15 bee bowls (5 blue, 5 yellow, and 5 white) were placed at each site to sample the bee community at each time period. Traps were placed along a 50 m transect beginning 5 m from the field margin (Droege 2015) from 08:00 to 17:00 at temperatures above 12˚C with minimal cloud cover and wind. At each time period, bees were also netted in two 15 min bouts; one bout between 09:00 and 11:00, and another bout between 14:00 and 16:00, both beginning 5 m from the field margin and along a 1 m^2 serpentine transect (*see* Plant diversity for further details on serpentine transects). All bee specimens collected were identified to species or morphospecies.

Bee species identification: To identify bees to species, bees were first visually identified to morphotype, and up to five samples were randomly selected from each morphotype for barcoding analysis. Specimen preparation followed guidelines provided by the Canadian Centre for DNA Barcoding (CCDB, University of Guelph, 50 Stone Road East, Guelph, ON, Canada). An antenna or leg (depending on specimen size and condition) was removed from each insect and placed into a well on a microplate provided by the CCDB. Wells were pre-filled using 95%

EtOH to avoid static displacement. Subsampling tools were cleaned with 95% EtOH between each specimen. Microplates were shipped to the CCDB for single-pass COI (cytochrome c oxidase subunit I gene) barcoding analysis and specimen metadata (e.g., collection location, voucher images) were uploaded to the Barcode of Life Data System (http://boldsystems.org). The CCDB uploaded sequence data to BOLD, which then provided identifications based on sequencing results. Identifications were applied to our collection if all five voucher specimens received the same species-level identification according to sequencing results.

When bees could not be identified to species using barcoding, specimens were identified with established keys. In particular, bees from these genera were identified with taxonomic keys: *Agapostemon*, *Apis*, *Bombus*, *Coelioxys*, *Hoplitis*, *Lasioglossum*, *Megachile*, *Panurginus*, and *Triepeolus* (Cockerell 1921; Michener 1935, 1936; Hurd and Michener 1955; Girgarick 1968; Roberts 1973; McGinley 1986; Richtmyer 2008; Sheffield et al. 2011; Williams et al. 2014). However, the lack of extant keys and limited access to well-curated reference collections led to identification of some specimens by comparison with barcoded specimens outside of our collection, use of online keys, and morphospecies determination. Bees in the following genera were identified by direct morphological comparisons with barcoded specimens outside of our collection: *Andrena, Lasioglossum* (*Dialictus* + *Evylaeus*), *Colletes*, *Melissodes*, and *Osmia*. Specimens of *Hylaeus* were examined and identified by Dr. Virginia Scott at the Colorado University Museum of Natural History. Bees in *Anthidium*, *Ceratina*, *Dianthidium* and *Halictus* were identified using the Discover Life online key, and when available, confirmed by comparison with online barcoded specimens outside of our collection.

In summary, 95% of species-level identifications were made via taxonomic keys, barcoded specimens outside of our collection, online keys, and morphospecies determination. The remaining 5% were identified to species through sequencing. Regarding morphospecies determination, we were unable to identify approximately 2% of specimens to species, however, these specimens were distinct from known species within the genus, thus they were given morphospecies determinations (*see* Data Set Files for descriptions). Overall, we were unable to generate species level determinations for 9% of the bee specimens within our collection. These specimens were removed from these data during cleaning and are not displayed here. Across the collection and identification methods, bees from the genera *Apis* and *Bombus* compose approximately 70% of all specimens. The remaining 30% of specimens represent bee species from 18 families. Voucher specimens are stored at Washington State University (Pullman, WA, USA) under secure, stable, laboratory conditions.

Plant diversity sampling and identification: Plant diversity sampling was conducted at the above-mentioned frequency (*see* Bee diversity sampling) with all measures (e.g., bee and plant diversity sampling) conducted on the same date at a given site. Thus, we measured flowering plant species richness and abundance at sites three times per year in May, July, and September, across the three years of our study (2014 - 2016). Values describing the total number of sampling events are described in the bee diversity sampling section, and methods for quantifying plant diversity have been previously published in Bloom et al. (2019). In summary, plants with flowers in anthesis were recorded during 1 hr surveys of each site using a serpentine transect. Serpentine transects used for the survey began 5 m from the field margin to avoid edge effects. We then

placed a portable 1 × 1 m plot within the production row and recorded all plants with flowers in anthesis. This process was repeated every 5 m down the production row until we reached the row end. At the row end, we moved over 5 m and repeated this measure down the nearest row in the opposite direction. Thus, we moved up and down rows in a serpentine fashion. This serpentine transect was also used for bee diversity sampling (*see* above) and pollen diversity data collection (*see* below). All transects across sites were approximately 800 m long. Flowering plants were identified to species. Thus, our measure of abundance represents an estimate of the number of each plant species in anthesis on the sampling date. Approximately 6% of plants could not be identified to species. Those records have been removed from these data.

Pollen diversity data collection: To determine which type of bee was collecting which pollen type, we walked 1 hr serpentine transects starting 5 m from the field margin, stopping every 5 m for 30 sec to net bees. These bees were preserved in molecular grade EtOH and stored in the laboratory under stable conditions (-80 °C) until further processing. Pollen collection occurred at the same frequency as the aforementioned sampling measures for bees and plants. Metabarcoding of pollen was then conducted on a subset of sites (22 of 36) where sampling was conducted across the full three-year study period (2014 - 2016).

Pollen samples were aggregated by bee group (honey bee, wild bee, bumble bee), site, and pooled across years (22 sites × 3 bee groups = 66 samples). Before extracting DNA from samples stored in 1.5 ml microcentrifuge tubes, they were centrifuged at high speed (14000 rpm) for 10 min and then EtOH was decanted from each tube. Remaining EtOH was removed through careful pipetting and the tubes were then dried and lyophilized using a Freezone Benchtop Freeze Dryer (Labconco Corp., Kansas City, MO) following the manufacturer's protocol. We processed pollen pellets for DNA extraction using a DNeasy Plant Mini Kit (Qiagen, Valencia, CA). Pellets were dissolved in 400 µl AP1 solution and approximately 100 µl of 0.1 mm glass beads were added to each sample, followed by 80 µl proteinase K (1 mg/ml) and 4 µl of RNase A. After briefly mixing the buffer, samples were disrupted using a Mini bead beater (Cole-Parmer, Vernon Hills, IL) for 3 min at 30 Hz and then incubated overnight at 56 °C. The subsequent steps were performed following manufacturer's instruction. Resulting DNA concentrations of pollen samples were measured using a NanoDrop 2000 spectrophotometer (ThermoFisher Scientific Corp., Carlsbad, CA, USA).

Extracted DNA was then used as template for library preparation and amplicon sequencing, performed at the Centre for Comparative Genomics and Evolutionary Bioinformatics at Dalhousie University (Halifax, Nova Scotia, Canada). There, amplicon fragments were PCR-amplified from DNA in duplicate, using separate template dilutions (1:1 & 1:10) and high-fidelity Phusion polymerase (New England BioLabs Inc., Ipswich, MA, USA). A single round of PCR was performed using "fusion primers" (Illumina adaptors + indices + specific regions) that target the *rbcL* region with multiplexing. More specifically, primers rbcL2 (Palmieri et al. 2009) and rbcLa-R (Kress & Erickson 2007) were used, which generate amplicons of ~500 bp in length (Bell et al. 2017). PCR products were verified visually by running a high-throughput Invitrogen 96-well E-gel (Thermo Fisher Scientific Corp., Carlsbad, CA, USA). Any samples with failed PCRs (or spurious bands) were re-amplified by optimizing PCR conditions to produce correct bands in order to complete a sample plate before continuing with sequencing. The PCR reactions

from the same samples were pooled in one plate, cleaned, and then normalized using the high-throughput Invitrogen SequalPrep 96-well Plate Kit (Thermo Fisher Scientific Corp.). Samples were then pooled to make one library, then quantified fluorometrically before sequencing. Amplicon samples were then run on an Illumina MiSeq using 300+300 bp paired-end V3 chemistry. Raw sequences are available on the NCBI Short Read Archive (SRA) under BioProject PRJNA715497.

Demultiplexed sequences (total reads: 5.93×10^6; mean read depth per sample: 9.12×10^4) were initially trimmed of trailing low-quality bases using the *DADA2* pipeline (v.1.8.0; Callahan et al. 2016) in R (R Core Team 2018). Paired-end reads were then quality-filtered, error-corrected, and assembled into amplicon sequence variants (ASVs). Once assembled, chimeras were detected, removed, and taxonomic information was then assigned to each ASV using the RDP Naïve Bayesian Classifier (Wang et al. 2007), trained to a *rbcL* database (Bell et al. 2017). Moreover, potential contaminate ASVs were identified through inclusion of negative controls during sample and sequence processing, and then removed using the 'prevalence' method with the *decontam* package (Davis et al. 2017). At the end of this process, a total of 2.81×10^6 reads remained, with samples having a mean read depth of 4.33×10^4. Across samples, a total of 1216 ASVs were detected, with taxonomic information assigned at the genus and species level for 78 and 51% of ASVs, respectively.

Pollinator visitation data collection: Observations of pollinator visitation (hereafter: bee-plant interactions) was conducted at the same frequency (intensity) as the aforementioned measures (e.g., *see* Bee diversity sampling). Thus, we measured bee-plant interactions at sites three times per year in May, July, and September, across the three years of our study (2014 - 2016). The methods for quantifying bee-plant interactions, along with information regarding the number of these interactions that could not be classified, were previously published in Bloom et al. (2019). In brief, we measured bee-plant interactions by walking two 1 hr serpentine transects. Bee-plant interactions were measured from 09:00 to 11:00 and 14:00 to 16:00 (1 hr during each period). Serpentine transects began 5 m from the field margin to avoid edge effects. Every 5 m we stopped for 30 sec to record bee-plant interactions visually within 1 m^2 of the transect. Sites were working farms; therefore, we could not establish permanent plots, nor did we use portable plots to avoid disturbing bee foraging behavior. An interaction was recorded when bees made contact with the floral reproductive organs of the plant for over 0.5 sec. Bees were classified into two groups, honey bees (*Apis mellifera* L.) or wild bees (all other bees). This classification resulted in one frequency matrix for each bee group per year (3 years × 2 bee groups = 6 matrices). The sampling conditions (e.g., weather) were described above (*see* Bee diversity sampling).

Legal/organizational requirements: No permits were needed; collecting occurred exclusively on private land. All landowners must be contacted and permission granted before any sampling can occur for any subsequent study.

CLASS III. DATA SET STATUS AND ACCESSIBILITY

A. Status

Last data update: 2 September 2021

Last archival date: 2 September 2021

Latest metadata status: Metadata are complete

Data verification: Data were reviewed and corrected for any input errors

B. Accessibility

Storage location and medium: EB and DC store raw data at Washington State University under secure, stable, laboratory conditions. DC has the digital data stored on a digital drive at Washington State University and EB has digital data backed up on a commercial cloud-based server at an off-site location. EO has cleaned data backed up on a commercial cloud-based server at an off-site location.

Contact person: Elias H. Bloom, Department of Entomology, Michigan State University, East Lansing, MI, 48824, bloomel1@msu.edu

Copyright or proprietary restrictions: There are no copyright or proprietary restrictions for research or teaching purposes. Usage of the data set must be cited with the below citations.

Journal Citation: Bloom, E. H.; E. C. Oeller; R. L. Olsson; M. R. Brousil; R. N. Schaeffer; S. Basu; Z. Fu; and D. W. Crowder. 2021. Documenting pollinators, floral hosts, and plant–pollinator interactions in U.S. Pacific Northwest agroecosystems.

Data Repository Citation: Bloom, E; E. Oeller; R. Olsson; R. Schaeffer; M. Brousil; S. Basu; Z. Fu; and D. Crowder. 2021, September 2. Bloom et al. Ecology data paper: Documenting pollinator communities, floral hosts, and plant-pollinator interactions in Pacific Northwest United States agroecosystems. figshare. Dataset. https://doi.org/10.6084/m9.figshare.14179640.v4

Disclaimer: Farm names have been anonymized and location coordinates have been jittered to protect the privacy and security of farmers.

CLASS IV: DATA STRUCTURAL DESCRIPTORS

A. Data set files

Identity and size:

EB_bee_diversity_data.csv	538 KB
EB_morphospecies_data.csv	6 KB
EB_plant_diversity_data.csv	282 KB
EB_pollen_network_genus.csv	5.56 MB
EB_pollen_network_species.csv	43.29 MB
EB_visitation_honeybees_2014.csv	4 KB
EB_visitation_honeybees_2015.csv	13 KB
EB_visitation_honeybees_2016.csv	7 KB

EB_visitation_wildbees_2014.csv	4 KB
EB_visitation_wildbees_2015.csv	14 KB
EB_visitation_wildbees_2016.csv	6 KB
EB_header_info.csv	5 KB
Metadata S1.csv	5 KB

B. Variable information

Storage type: Hard copy data sheets are housed at Washington State University by DC. Digital data files are stored as comma-delimited text files (.csv).

List and definition of variable codes: Found in EB_header_info.csv

Missing value codes: Data were cleaned to remove all blank cells. Cells with "NA" indicate data were not available.

CLASS V. SUPPLEMENTAL DESCRIPTORS

A. Data acquisition

Data Forms: NA

Location of completed data: Physical notebooks and digital copies containing original raw field data are stored by EB and DC at Washington State university. EO has cleaned data backed up on a commercial cloud-based server at an off-site location.

B. QA/QC procedures: Datasets were cleaned by consolidating raw data, removing unnecessary metadata, removing blanks, checking for inconsistencies in data entry, correcting typos and consolidating unique data, anonymizing site names, and jittering site coordinates for privacy. We also removed data with missing identifications from our datasets. The procedure for data removal and percent of data removed are described above (*see* Survey design).

C. Publications using the data set: Bloom, E.H., Northfield, T.D., Crowder, D.W. 2019. A novel application of the Price equation reveals that landscape diversity promotes the response of bees to regionally rare plant species. Ecol Lett 22: 2103-2110. This paper uses bee and plant diversity data included here.

ACKNOWLEDGMENTS:

We thank the 36 sites that supported the research and the Common Acre, Washington Tilth, Seattle Department of Neighborhoods, Lettuce Link, Garden Raised Urban Bounty, Sustainable South Sound, D. Collins, and L. Patzek for helping form this network. We thank A. Pelegrin for assistance with bee species identification.

LITERATURE CITED:

Bell, K. L., V. M. Loeffler, and B. J. Brosi. 2017. An rbcL reference library to aid in the identification of plant species mixtures by DNA metabarcoding. *Applications in Plant Sciences* **5**: 1-7.

Bloom, E. H., T. D. Northfield, and D. W. Crowder. 2019. A novel application of the Price equation reveals that landscape diversity promotes the response of bees to regionally rare plant species. *Ecology Letters* **22**: 2103-2110.

BWARS (The UK Bees, Wasps and Ants Recording Society). 2020. https://www.bwars.com/. Viewed 16 June 2020.

Callahan, B. J., P. J. McMurdie, M. J. Rosen, A. W. Han, A. J. A. Johnson, and S. P. Holmes. 2016. DADA2: high-resolution sample inference from Illumina amplicon data. *Nature Methods* **13**: 581-583.

Cockerell, T. 1921. Some parasitic megachilid bees of the western United States. http://hdl.handle.net/2246/3247. Viewed 15 March 2019.

Cuo, L., D. P. Lettenmaier, M. Alberti, and J. E. Richey. 2009. Effects of a century of land cover and climate change hydrology on the Puget Sound basin. *Hydrological Processes* **23**: 907-933.

DNR (Washington State Department of Natural Resources). 2021. Puget Lowland. https://www.dnr.wa.gov/programs-and-services/geology/explore-popular-geology/geologic-provinces-washington/puget-lowland. Viewed 19 2021.

DON (Seattle Department of Neighborhoods). 2020. P-patch list. https://www.seattle.gov/neighborhoods/programs-and-services/p-patch-community-gardening/p-patch-list. Viewed 2 July 2020.

Eilers, E. J., C. Kremen, S. S. Greenleaf, *et al.* 2011. Contribution of pollinator-mediated crops to nutrients in the human food supply. *PLoS ONE* **6**: e21363.

FAO (Food and Agriculture Organization of the United Nations). 2020. FAO's Global Action on Pollination Services for Sustainable Agriculture. http://www.fao.org/pollination/major-initiatives/en. Viewed 11 May 2020.

Garibaldi, L. A., I. Steffan-Dewenter, R. Winfree, *et al.* 2013. Wild pollinators enhance fruit set of crops regardless of honey bee abundance. *Science* **339**: 1608-1611.

Girgarick, A. A., and L. A. Strange. 1968. Pollen-collecting bees of the Anthidiini of California. Cambridge University Press, London, England.

Granatstein, D., and E. Kirby. 2019. Current status of certified organic agriculture in Washington state: 2018. Center for Sustaining Agriculture and Natural Resources, Washington State University, Wenatchee, Washington.

Hurd, P., and C. Michener. 1955. The Megachiline bees of California. University of California Press, Berkeley, CA.

Isaacs, R., and A. K. Kirk. 2010. Pollination services provided to small and large highbush blueberry fields by wild and managed bees. *Journal of Applied Ecology* **47**: 841-849.

Kennedy, C. M., E. Lonsdorf, M. C. Neel, *et al.* 2013. A global quantitative synthesis of local and landscape effects on wild bee pollinators in agroecosystems. *Ecology Letters* **16**: 584-599.

Kirby, E., and D. Granatstein. 2013. Trends in Washington State organic agriculture 2004-2011. Washington State University Press, Pullman, Washington.

Klein, A. M., B. E. Vaissiére, J. H. Cane, *et al.* 2007. Importance of pollinators in changing landscapes for world crops. *Proceedings of the Royal Society B* **274**: 303-313.

Koh, I., E. V. Lonsdorf, N. M. Williams, *et al.* 2016. Modeling the status, trends, and impacts of wild bee abundance in the United States. *PNAS* **113**: 140-145.

Kress, W. J., and D. L. Erickson. 2007. A two-locus global DNA barcode for land plants: the coding rbcL gene complements the non-coding trnH-psbA spacer region. *PLoS ONE* **2**: e508.

Losey, J. E., and M. Vaughan. 2006. The economic value of ecological services provided by insects. *BioScience* **56**: 311-323.

McGinley, R. 1986. Studies of Halictinae (Apoidea: Halictidae) I: revision of new world Lasioglossum Curtis. Smithsonian Contributions to Zoology. 1-294.

Michener, C. 1935. Some pacific coast Panurginus. Canadian Entomologist **67**: 275-278.

Michener, C. 1938. American bees of the genus Heriades. Annals of the Entomological Society of America **31**: 514-531.

Palmieri, L., E. Bozza, and L. Giongo. 2009. Soft fruit traceability in food matrices using real-time PCR. *Nutrients* **1**: 316– 328.

PHTF (Pollinator Health Task Force). 2015. Pollinator research action plan. Washington, DC, The White House.

Potts, S. G., J. C. Biesmeijer, C. Kremen, *et al.* 2010. Global pollinator declines: trends, impacts and drivers. *Trends in Ecology and Evolution* **25**: 345-353.

R Core Team. 2018. R: A Language and Environment for Statistical Computing, Vienna, Austria: R Foundation for Statistical Computing.

Reganold, J., J. Wachter. 2016. Organic agriculture in the twenty-first century. *Nature Plants* **2**: 15221.

Rightmyer, M.G. 2008. A review of the cleptoparasitic bee genus Triepeolus (Hymenoptera: Apidae) Part 1. Magnolia Press, Auckland, New Zealand.

Roberts, R. 1973. Bees of northwestern America: Agapostemon (Hymenoptera: Halictidae). Oregon State University, Corvalis, OR.

Sánchez-Bayo, F., and K. A. G. Wyckhuys. 2019. Worldwide decline of the entomofauna: A review of its drivers. *Biological Conservation* **232**: 8-27.

Sheffield, C.S., R. Claudia, L. Packer, and T. Griswald. 2011. Leafcutter and mason bees of the genus Megachile Latreille (Hymenoptera: Megachilidae) in Canada and Alaska. Canadian Journal of Arthropod Identification. 18.

Soroye, P. E., T. Newbold, and J. Kerr. 2020. Climate change contributes to widespread declines among bumble bees across continents. *Science* **367**: 685-688.

USDA (United States Department of Agriculture). 2020. Organic integrity database. https://organic.ams.usda.gov/integrity/. Viewed 2 July 2020.

USDA-NASS (United States Department of Agriculture National Agricultural Statistics Service). 2014. Cropland Data Layer. https://nassgeodata.gmu.edu/CropScape/. Viewed 18 June 2021.

USDA-NASS (United States Department of Agriculture National Agricultural Statistics Service). 2020. Cropland Data Layer. https://nassgeodata.gmu.edu/CropScape/. Viewed 19 June 2021.

Van Klink, R., D. E. Bowler, K. B. Gongalsky, *et al.* 2020. Meta-analysis reveals declines in terrestrial but increases in freshwater insect abundances. *Science* **368**: 417-420.

Wang, Q., G. M. Garrity, J. M. Tiedje, and J. R. Cole. 2007. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology* **73**: 5261-5267.

Wenzel, A., I. Grass, V.V. Belavadi, and T. Tscharntke. 2020. How urbanization is driving pollinator diversity and pollination - a systematic review. *Biological Conservation* **241**: 108321.

Williams, P.H., R. W., Thorp, L. L. Richardson, and S. Colla. 2014. Bumble bees of North America: an identification guide. Princeton University Press, Princeton, New Jersey.

Williams, N. M., R. Isaacs, E. Lonsdorf, *et al.* 2019. Building resilience into agricultural pollination using wild pollinators. In: Gardner SM and Ramsden SJ (Eds). Agricultural resilience perspectives from ecology and economics. New York, NY: Cambridge University Press.

Wratten, S. D., M. Gillespie, A. Decourtye, *et al.* 2012. Pollinator habitat enhancement: benefits to other ecosystem services. *Agriculture, Ecosystems & Environment* **159**: 112-122.

WSDA (Washington State Department of Agriculture). 2020. NOP certified organic operations: producers. https://cms.agr.wa.gov/getmedia/0b56ff46-2350-461f-99f1-6f3d48d6634c/wsda_cert_org_producers. Viewed 2 July 2020.