Quick guide to FeederWatch data

Project FeederWatch is a citizen science program operated by the Cornell Lab of Ornithology in cooperation with Bird Studies Canada. More information: www.feederwatch.org

For more details about the data and how they can be accurately used in a biologically relevant manner, contact David Bonter (dnb23@cornell.edu) or Wes Hochachka (wmh6@cornell.edu).

Note that the FeederWatch protocol is a repeated measures design. Participants (ID) are reporting from the same location (LOC_ID) as often as 21 times each winter with many people reporting for many years. As such, the data are amenable to occupancy modeling.

Data quality: All FeederWatch reports (starting in 2006) are passed through geographically and temporally explicit filters to flag records that are unexpected for the location/month. Records that do not trigger a flag are given a '1' in the Valid field and enter the database without review. Records that trigger the flag ('0' in the Valid field) and are reviewed by a biologist who may clear the flag, ask the participant for supporting information, or reject the report. Reports marked Valid = 1 and Reviewed = 1 have been reviewed by a biologist and approved. Reports marked as Invalid in the database are not included in this file. [This is a similar review system as the one used by eBird.] Undoubtedly, some of the reports in the database involve incorrect identifications or reports by participants who do not completely follow the FeederWatch protocol. As with all datasets, use with caution. See Bonter & Cooper 2012. Data validation in citizen science: A case study from Project FeederWatch. *Frontiers in Ecology and the Environment* 10:305-307.

FeederWatch Protocol: In brief, participants (N ~ 20,000 in 2014) count birds at their own bird feeders (usually) as often as once/week from early November to early April annually. The count period is 2 days long. Participants report the MAXIMUM number of each species in view at any one time during the 2-day count period. This methodology avoids double-counting of the same individual birds. The complete protocol is available here: http://feederwatch.org/about/detailed-instructions/

Data fields included in the .csv file:

ID = The participants' unique Cornell Lab of Ornithology identification number. **Note that these ID numbers should not be shared or made publically available as they could, in theory, be used to access personal information of participants.** They should simply be used as a unique identifier of the individuals making observations.

Latitude/Longitude: In decimal degrees. Note that these locations are identified by the participants with varying degrees of accuracy. Prior to 2000, all sites were given the latitude and longitude of the centroid of the ZIP code (identified as "POSTCODE LAT/LONG LOOKUP" in the ENTRY_TECHNIQUE field). Since our online data entry system was developed (late 1999), we've maintained a series of different systems for identifying the count locations that are far more accurate. Contact David Bonter for more details.

StatProv = US State or Canadian Province of count location.

ENTRY_TECHNIQUE: See Latitude/Longitude above.

FW_YEAR = The FeederWatch season. Seasons run from November to April. For instance, the FW_YEAR "PFW_1992" indicates the season running from November 1991 to April 1992.

Month/Day/Year = First date of the 2-day observation period.

NHalfDays = Number of Half Days of observation during the 2-day count period. This is a measure of observer effort. Categorical (1, 2, 3, or 4 half days).

EFFORT_HRS_ATLEAST = Another measure of observer effort. Categorical (< 1, 1-4 hours, 4-8 hours, > 8 hours). Researchers at the Cornell Lab have found both measures of effort are informative in models.

BirdSpp = 6-letter species code. Contact David Bonter if the key is needed.

NSeen = Number of individuals seen (remember that this is the maximum number of the species in view at any one time over the 2-day observation period).

Valid/Reviewed. See above.

LOC_ID = Unique identifier of the location of observation. Note that participants (ID) can have multiple count locations (LOC_ID).

SUB_ID = Submission Identifier. This is a unique identifier of the entire checklist submitted by a participant from a count period. All species observations submitted on the same checklist have the same SUB_ID.

OBS_ID = Observation identifier. This is a unique identifier for the single observation. All species reported on a single checklist receive different OBS_IDs.