GENERAL INFORMATION

1. Title of Dataset: Habitat mediates coevolved but not novel interactions

2. Author Information

A. Principal Investigator Contact Information

Name: Joshua P Twining

Institution: Cornell University

Address: DNR, Cornell University, Fernow Hall, Ithaca, New York, 14882, USA.

Email: [jpt93@cornell.edu](mailto:jpt93@cornell.edu) / [joshuaptwining@gmail.com](mailto:joshuaptwining@gmail.com)

3. Date of data collection (single date, range, approximate date) <suggested format YYYY-MM-DD>:

2015, 2018 and 2020

4. Geographic location of data collection <latitude, longiute, or city/region, State, Country, as appropriate>:

Across Northern Ireland (see associated site level covariate datasheet for spatial locations of sites)

5. Information about funding sources that supported the collection of the data:

British Ecological Society grant

SHARING/ACCESS INFORMATION

1. Licenses/restrictions placed on the data:

None

DATA & FILE OVERVIEW

1. File List:

Site level covariate data: This is local site level covariate information for each of the sampled sites. Each column is a different covariate.

Pinemartennoname: This is a detection-non detection matrix for the pine marten (*Martes martes)* from a multi-species surveys from 2015 – 2020

Redsquirrelnoname: This is a detection-non detection matrix for the red squirrel (*Sciurus vulgaris)* from a multi-species surveys from 2015 – 2020

Greysquirrelnoname: This is a detection-non detection matrix for the grey squirrel (*Sciurus carolinensis)* from a multi-species survey from 2015-2020.

2. Relationship between files, if important: The three detection – non detection arrays are from the same survey but for different species. All the rows of all the files align, e.g. row 34 in each file is the same site across the files. All site information for each detection – non detection can be found in the site covariate file.

3. Additional related data collected that was not included in the current data package: Detections of other non-focal species through the survey.

4. Are there multiple versions of the dataset? No

METHODOLOGICAL INFORMATION

1. Description of methods used for collection/generation of data:

We conducted a multi-species citizen science camera trap survey. The methods for this field survey can be read here: <https://besjournals.onlinelibrary.wiley.com/doi/10.1111/1365-2664.13660>

2. Methods for processing the data:

The detection – non detection data was produced by converting time and dates of detections of each species into long format binary form (with each column representing a 24hr period). The number of columns for each site represent the length of sampling. NAs show no survey was conducted that day.

3. Instrument- or software-specific information needed to interpret the data:

Bushnell Trophy Cams and Browning StrikeForce X HD were used to collect data.

4. Standards and calibration information, if appropriate:

All cameras were set to specifications which can be found here: <https://besjournals.onlinelibrary.wiley.com/doi/10.1111/1365-2664.13660>

5. Environmental/experimental conditions:

This was a landscape level observational survey so environmental conditions varied between sites and years.

6. People involved with sample collection, processing, analysis and/or submission:

Survey in 2015 was led by David Tosh, data primarily collected by citizen scientists. Survey in 2018 was led by Josh Twining in collaboration with Ulster Wildlife with data primarily collected by citizen scientists. Survey in 2020 was collected by Josh Twining due to COVID-19 data collected by citizen scientists was limited.

DATA-SPECIFIC INFORMATION FOR: Site covariates

1. Number of variables:

15

2. Number of cases/rows:

706

3. Variable List:

Site: Name of the site

1km square grid: the 1km grid location that the site fell within (using standard Irish Grid)

Bog: the proportion of the 1km grid made up of bog

Broadleaf: the proportion of the 1km grid made up of broadleaf and mixed woodlands

Built: the proportion of the 1km grid made up of urban and suburban land use

Conifer: the proportion of the 1km grid made up of coniferous woodland

Dwarf: the proportion of the 1km grid made up of heath

Rough: the proportion of the 1km grid made up of rough grassland

River: the length of freshwater ways (both canals and rivers) in the 1km grid

People: the estimated number of people inhabiting the 1km grid

Grassland: the proportion of a 1km grid made up of neutral, improved and acid grassland

Lat: the latitude of the centre of the 1km grid

Long: the longitude of the centre of the 1km grid

Year: the year sampling was conducted

Bait: the baited used

4. Missing data codes:

NA is used for missing data

DATA-SPECIFIC INFORMATION FOR: All three detection-non detection matrices, entitled “speciesname”noname.csv

1. Number of variables:

This is a detection matrix

2. Number of cases/rows:

706

3. Variable List:

Each column represents a 24hr period of sampling using a camera trap. 1 = detection of that species, 0 = no detection of that species, NA = no sampling conducted that day.

4. Missing data codes:

NA is used for missing data