

Data management plan

Research products

Research products for this proposal will include the development of a species home range database, as well as analysis code, data processing scripts, presentations/posters, and the resulting publications.

Reuse and redistribution

All research products, including code, data, presentations/posters, and publications will be open access and freely available for reuse without request during all stages of development under a Creative Commons Zero (CC0) or Creative Commons with Attribution (CC-BY) license, following the Panton Principles guidelines (<http://pantonprinciples.org>). The CC0 license dedicates the research products to the public domain without restriction, allowing for maximum reusability, while CC-BY allows for reuse without restriction, as long as proper attribution is given.

Code

Research products (code) will be available on a publicly available, open access GitHub repository (<http://github.com/embaldrige>), and GitHub issues and wiki will be used as an online lab notebook. The version of code used in final publication will be archived through Zenodo (<https://zenodo.org/>).

Data

The majority of the data that I will use is publicly available data, so a data management plan (for the raw data) is not necessary for those datasets. However, any scripts written to extract and process the data will be available in the code repository to make the data extraction process fully reproducible, as long as a researcher has access to the original datasets.

No new data will be collected in the field; but I will compile additional home range data from the literature to create a more comprehensive home range database than is currently available.

Data format and access

The home range database compilation will be initially compiled openly under version control on GitHub. Individual tables will be stored as .csv files, which have maximum interoperability for importing directly into a variety of analysis and data processing software/languages, including SQL databases as well as R and Python. Metadata will be written in the machine readable Ecological Metadata Language (EML), but will also

be presented in a human-readable .html form following the *Ecological Archives* guidelines to communicating caveats and best practice suggestions for data compilations that are difficult to convey with the EML format. Versions of the database used for analysis will be archived through Zenodo prior to ultimate publication in *Ecological Archives* as a data paper, which will archive the data and allow the data to be freely accessed in a more permanent format. I will write scripts to make these data, and any other publicly available datasets that I use, directly importable through the EcoData Retriever(www.ecodataretriever.org). The EcoData Retriever is a tool developed to make the reuse of ecological data simpler and more reproducible [1].

Presentations/Posters

Presentations and posters about the research will be archived on GitHub (<http://github.com/embaldrige/posters-pres>) and either figshare (posters) or Slideshare (presentations). Additionally, a pre-talk transcript will be available on the GitHub repository in advance of the presentation to accommodate neurodivergent and hearing impaired colleagues. Ideally, talks will be live-streamed with Google Hangouts on Air or similar software at the event to allow for remote participation. If this is not possible, due to venue or event restrictions, I will record the talk after the fact. In either event, the video of the talk will then be closed captioned and made publicly available on YouTube.

Publications

All publications will first be made available as pre-prints through a preprint server such as arXiv, bioRxiv, or PeerJ. Final publications will be published in journals that allow for gold open access publication and have rapid review times (e.g., Ecology, PLoS, PeerJ, etc.).

Roles and responsibilities

Elita Baldridge will be primarily responsible for data collection and management, code development and management, and dissemination of research products. Timothée Poisot (University of Montreal) will provide guidance and support in developing, managing, and sharing code. The reproducible workflow and archiving of research products for this project will provide long-term open-access to data, code, and computational tools both during the term of the grant, and after the grant ends.

1. Morris, B. D. and White, E. P. “**The EcoData Retriever: Improving Access to Existing Ecological Data**” *PloS one* 8, no. 6 (2013): e65848.