Does protected area connectivity moderate the efficacy of protection on tropical biodiversity? Evidence from a replication of Brodie et al. 2023

Peter Kedron, Lei Song, Wenxin Yang, Amy Frazier

2023-12-05

Abstract

This study is a *replication* of:

> Brodie, J.F., Mohd-Azlan, J., Chen, C. et al. Landscape-scale benefits of protected areas for tropical biodiversity. Nature 620, 807–812 (2023). https://doi.org/10.1038/s41586-023-06410-z

We replicate the analysis of Brodie et al. and introduce protected area (PA) connectivity as a statistical moderator of the effect PA status has on biodiversity. We find KEY FINDINGS. Using a causal framework that controls for X through matching, Brodie et al. (2023) find evidence that protected areas (PA) do preserve vertebrate biodiversity within their boundaries and in the adjacent unprotected landscape. Brodie et al. provide evidence of the efficacy of protected area status , they do not assess whether the effect they observe is altered by network connectivity.

1 Study metadata

- `Key words`: Biodiversity, Conservation, Protected Areas, Connectivity, 30x30
- `Subject`: select from the BePress Taxonomy
- `Date created`: November 8, 2023
- `Date modified`: date of most recent revision
- `Spatial Coverage`: Specify the geographic extent of your study. This may be a place name and link to a feature in a gazetteer like GeoNames or OpenStreetMap, or a well known text (WKT) representation of a bounding box.
- `Spatial Resolution`: Specify the spatial resolution as a scale factor, description of the level of detail of each unit of observation (including administrative level of administrative areas), and/or or distance of a raster GRID size
- `Spatial Reference System`: Specify the geographic or projected coordinate system for the study, e.g. EPSG:4326
- `Temporal Coverage`: Specify the temporal extent of your study—i.e. the range of time represented by the data observations.
- `Temporal Resolution`: Specify the temporal resolution of your study—i.e. the duration of time for which each observation represents or the revisit period for repeated observations
- `Funding Name`: NA
- `Funding Title`: NA
- `Award info URI`: NA
- `Award number`: NA

1.1 Original study spatio-temporal metadata

- `Spatial Coverage`: extent of original study
- `Spatial Resolution`: resolution of original study
- `Spatial Reference System`: spatial reference system of original study
- `Temporal Coverage`: temporal extent of original study
- `Temporal Resolution`: temporal resolution of original study

2 Study design

Describe how the study relates to prior literature, e.g. is it a **original study**, **meta-analysis study**, **reproduction study**, **reanalysis study**, or **replication study**?

Also describe the original study archetype, e.g. is it **observational**, **experimental**, **quasi-experimental**, or **exploratory**?

Enumerate specific **hypotheses** to be tested or **research questions** to be investigated here, and specify the type of method, statistical test or model to be used on the hypothesis or question.

3 Materials and procedure

3.1 Computational environment

3.2 Data and variables

Describe the **data sources** and **variables** to be used. Data sources may include plans for observing and recording **primary data** or descriptions of **secondary data**. For secondary data sources with numerous variables, the analysis plan authors may focus on documenting only the variables intended for use in the study.

Primary data sources for the study are to include . . . . Secondary data sources for the study are to include . . . .

Each of the next subsections describes one data source.

3.2.1 Secondary data source1 name

- `Title`: Title of data source
- `Abstract`: Brief description of the data source
- `Spatial Coverage`: Specify the geographic extent of your study. This may be a place name and link to a feature in a gazetteer like GeoNames or OpenStreetMap, or a well known text (WKT) representation of a bounding box.
- `Spatial Resolution`: Specify the spatial resolution as a scale factor, description of the level of detail of each unit of observation (including administrative level of administrative areas), and/or or distance of a raster GRID size
- `Spatial Reference System`: Specify the geographic or projected coordinate system for the study
- `Temporal Coverage`: Specify the temporal extent of your study—i.e. the range of time represented by the data observations.
- `Temporal Resolution`: Specify the temporal resolution of your study—i.e. the duration of time for which each observation represents or the revisit period for repeated observations
- `Lineage`: Describe and/or cite data sources and/or methodological steps used to create this data source

- `Distribution`: Describe how the data is distributed, including any persistent identifier (e.g. DOI) or URL for data access
- `Constraints`: Legal constraints for *access* or *use* to protect *privacy* or *intellectual property rights*
- `Data Quality`: State result of quality assessment or state "Quality unknown"
- `Variables`: For each variable, enter the following information. If you have two or more variables per data source, you may want to present this information in table form (shown below)
  - `Label`: variable name as used in the data or code
  - `Alias`: intuitive natural language name
  - `Definition`: Short description or definition of the variable. Include measurement units in description.
  - `Type`: data type, e.g. character string, integer, real
  - `Accuracy`: e.g. uncertainty of measurements
  - `Domain`: Range (Maximum and Minimum) of numerical data, or codes or categories of nominal data, or reference to a standard codebook
  - `Missing Data Value(s)`: Values used to represent missing data and frequency of missing data observations
  - `Missing Data Frequency`: Frequency of missing data observations

Label	Alias	Definition	Type	Accuracy	Domain	Missing Data Value(s)	Missing Data Frequency
variable1	...	...	...	...	...	...	...
variable2	...	...	...	...	...	...	...

3.2.2 Secondary data source2 name

... same form as above...

3.3 Prior observations

At the beginning of this analysis, we had observed the dataset provided by Brodie et al with their publication. We did not manipulate the data before beginning our replication attempt.

3.4 Bias and threats to validity

Given the research design and primary data to be collected and/or secondary data to be used, discuss common threats to validity and the approach to mitigating those threats, with an emphasis on geographic threats to validity.

These include: - uneven primary data collection due to geographic inaccessibility or other constraints - multiple hypothesis testing - edge or boundary effects - the modifiable areal unit problem - nonstationarity - spatial dependence or autocorrelation - temporal dependence or autocorrelation - spatial scale dependency - spatial anisotropies - confusion of spatial and a-spatial causation - ecological fallacy - uncertainty e.g. from spatial disaggregation, anonymization, differential privacy

3.5 Data transformations

Describe all data transformations planned to prepare data sources for analysis. This section should explain with the fullest detail possible how to transform data from the **raw** state at the time of acquisition or

observation, to the pre-processed **derived** state ready for the main analysis. Including steps to check and mitigate sources of **bias** and **threats to validity**. The method may anticipate **contingencies**, e.g. tests for normality and alternative decisions to make based on the results of the test. More specifically, all the **geographic** and **variable** transformations required to prepare input data as described in the data and variables section above to match the study's spatio-temporal characteristics as described in the study metadata and study design sections. Visual workflow diagrams may help communicate the methodology in this section.

Examples of **geographic** transformations include coordinate system transformations, aggregation, disaggregation, spatial interpolation, distance calculations, zonal statistics, etc.

Examples of **variable** transformations include standardization, normalization, constructed variables, imputation, classification, etc.

Be sure to include any steps planned to **exclude** observations with *missing* or *outlier* data, to **group** observations by *attribute* or *geographic* criteria, or to **impute** missing data or apply spatial or temporal **interpolation**.

3.6 Analysis

Describe the methods of analysis that will directly test the hypotheses or provide results to answer the research questions. This section should explicitly define any spatial / statistical *models* and their *parameters*, including *grouping* criteria, *weighting* criteria, and *significance thresholds*. Also explain any follow-up analyses or validations.

4 Results

Describe how results are to be presented.

5 Discussion

Describe how the results are to be interpreted *vis a vis* each hypothesis or research question.

6 Integrity Statement

Include an integrity statement - The authors of this preregistration state that they completed this preregistration to the best of their knowledge and that no other preregistration exists pertaining to the same hypotheses and research. If a prior registration *does* exist, explain the rationale for revising the registration here.

7 Acknowledgements

- `Funding Name`: name of funding for the project
- `Funding Title`: title of project grant
- `Award info URI`: web address for award information
- `Award number`: award number

This report is based upon the template for Reproducible and Replicable Research in Human-Environment and Geographical Sciences, DOI:%5B10.17605/OSF.IO/W29MQ](https://doi.org/10.17605/OSF.IO/W29MQ)

8 References