



SUSTAINABLE
FOREST
TRANSITIONS



Cornell University

Navigating data challenges in socioeconomic impact assessments of conservation regimes

Reem Hajar, Johan A Oldekop, Roberto Toto,
Lucas Alencar, Samuel D Bell, Katie Devenish,
Duong T Khuu, Mariana Hernández-Montilla,
Suhyun Jung, Sandy Nofyanza, Lok Mani Sapkota

Research Brief #3
April 2025



Availability of socioeconomic data to assess the impacts of conservation and restoration interventions remains a challenge. This “cheat sheet” provides a quick guide to available and emerging data products and how they can be used to improve evaluation efforts.

Types of data

There currently exist four broad categories of socioeconomic data:

1. National census data.
2. Representative household surveys (e.g., the World Bank’s Living Standards Measurement Study, LSMS).
3. Surveys collected by individual research programmes (e.g., the Poverty and Environment Network global dataset).
4. Gridded high-resolution human development and wellbeing data.

Four factors should be considered when using these data.

Indicator availability and consistency

Poverty measures often rely on socioeconomic proxies like a wealth index or a multidimensional poverty index (MPI). These indexes are occasionally calculated as part of more recent data collection efforts but must often be calculated manually for earlier periods to generate consistent measures over time. However, changes in data collection protocols can sometimes affect the comparability of measures over time.

Temporal resolution

Socioeconomic data should align with the timing of interventions being assessed. Panel data measuring the same unit of observation over time are preferable, but repeated cross-sectional data measuring different observations over time can provide useful alternatives when panel data are not available.

Census data provide the required periodicity but are not always easily accessible and need to be licensed. Representative household surveys often cover several time periods but rarely measure the same units of observation over time. Collecting new socioeconomic data is the most effective way to obtain optimal data for impact evaluations, but is logistically difficult, time consuming, and costly.

However, it is possible to combine prior data collection efforts with new ones to create panels if the same units of analysis can be sampled again.

Spatial resolution

Socioeconomic data should match the spatial scale at which impacts are expected to occur (e.g., households or villages). Access to fine-resolution data from censuses is often limited, and representative household surveys (e.g., LSMS) are typically only representative at broad administrative levels.

When there is a spatial mismatch between intervention and outcome data, downscaling methods, such as spatial interpolation (Fig 1A and 1B), proximity buffers (Fig 1C) and small area estimation (SAE), can be used, though each method carries risks of bias, precision loss, or modelling limitations.

A new generation of gridded poverty and wellbeing datasets has potential to revolutionise evaluations of socioeconomic outcomes. These

datasets typically combine representative household surveys, satellite imagery and other data with machine learning to create wall-to-wall high-resolution data. However, these relatively new data cannot currently be used to measure changes over longer time periods.

Format

Socioeconomic data come in various formats, creating challenges for harmonizing and linking them with spatial data. Census data often require merging with spatial boundary data

using unique identifiers like municipality codes, but sometimes only point-based spatial references are available. Thiessen polygons (Fig 1D) can be created to approximate census sampling polygons in these cases. These artificial boundaries can offer consistent, non-overlapping coverage over time. However, if representative, sample-based household surveys are used, Thiessen polygons cannot be used, so alternative methods like buffers or distance-based measures are needed to harmonize these datasets with spatial data.

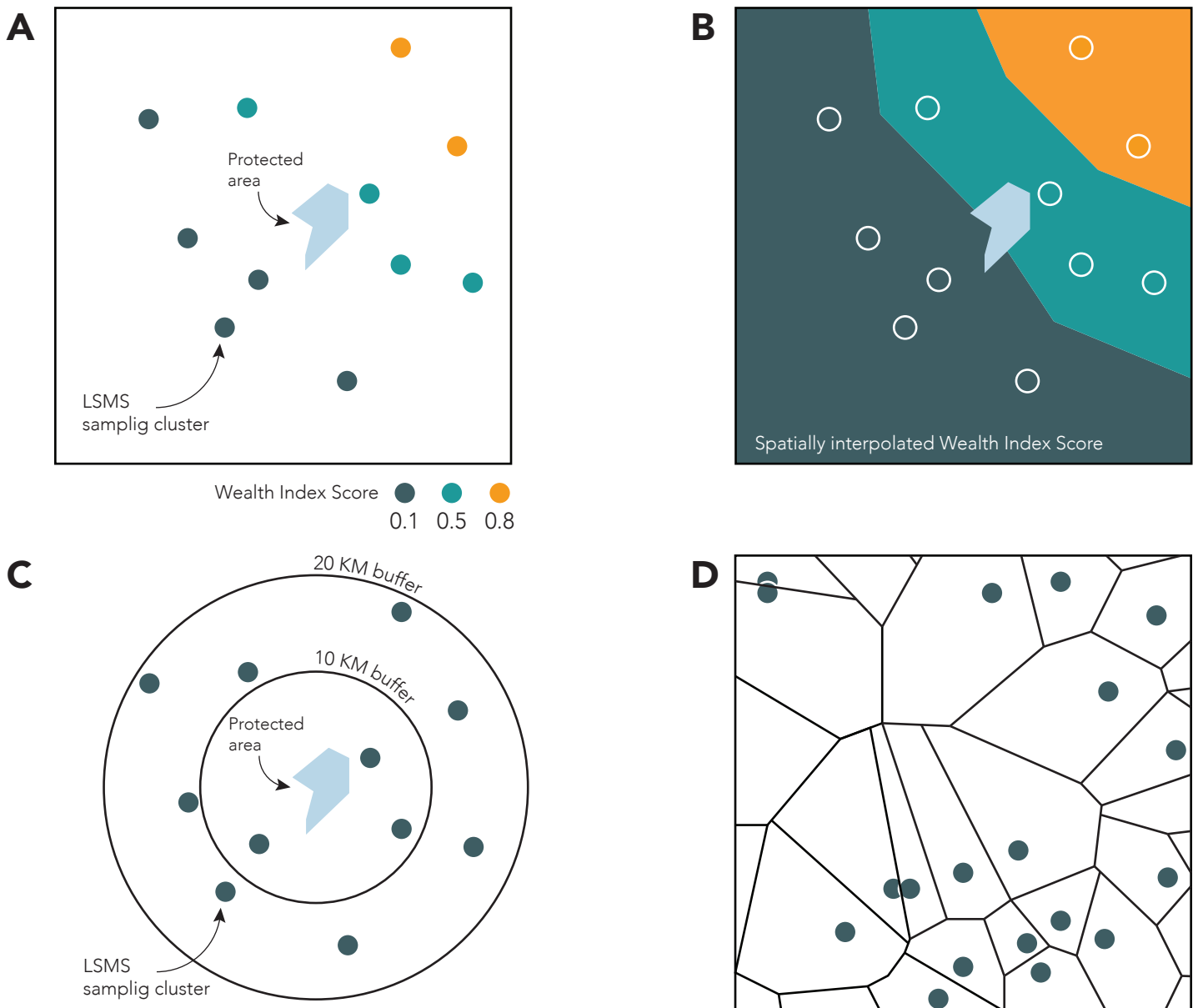


Figure 1. Spatial interpolation of Wealth Index Scores from representative household surveys (A and B); Distance buffers for representative household surveys (C); Artificial census sampling boundaries generated using Thiessen polygons when only point-based spatial information is available (D).

About the authors



Reem Hajjar is an Associate Professor in the College of Forestry at Oregon State University and the Department of Natural Resources and the Environment at Cornell University.



Johan Oldekop is an Associate Professor at the University of Manchester. He leads the Sustainable Forest Transitions project.



Roberto Toto is a PhD student in the College of Agricultural Sciences at Oregon State University.



Lucas Alencar is a research associate with the Sustainable Forest Transitions at the University of Manchester.



Sam Bell is an Assistant Professor in the College of Agricultural Sciences at Oregon State University.



Duong Khuu is a Leverhulme Trust research fellow with the Sustainable Forest Transitions project at the University of Manchester.



Katie Devenish is a research associate with the Sustainable Forest Transitions at the University of Manchester.



Mariana Hernández-Montilla is a PhD student with the Sustainable Forest Transitions project at the University of Manchester.



Suhyun Jung is an Assistant Professor in the College of Forestry at Oregon State University.



Sandy Nofyanza is a PhD student with the Sustainable Forest Transitions project at the University of Manchester.



Lok Mani Sapkota is a PhD student in the College of Forestry at Oregon State University.

About Sustainable Forest Transitions

Sustainable Forest Transitions examines how drivers of reforestation can benefit the environment and local communities, while improving the design and evaluation of forest-sector interventions. The project was selected by the European Research Council (ERC) and funded by UK Research and Innovation (UKRI) under grant number EP/X023222/1.

For more information and outputs

www.sites.manchester.ac.uk/sftresearch