The Geodata Decision Tree
A Guiding Framework for the Use of Geodata in Evaluation
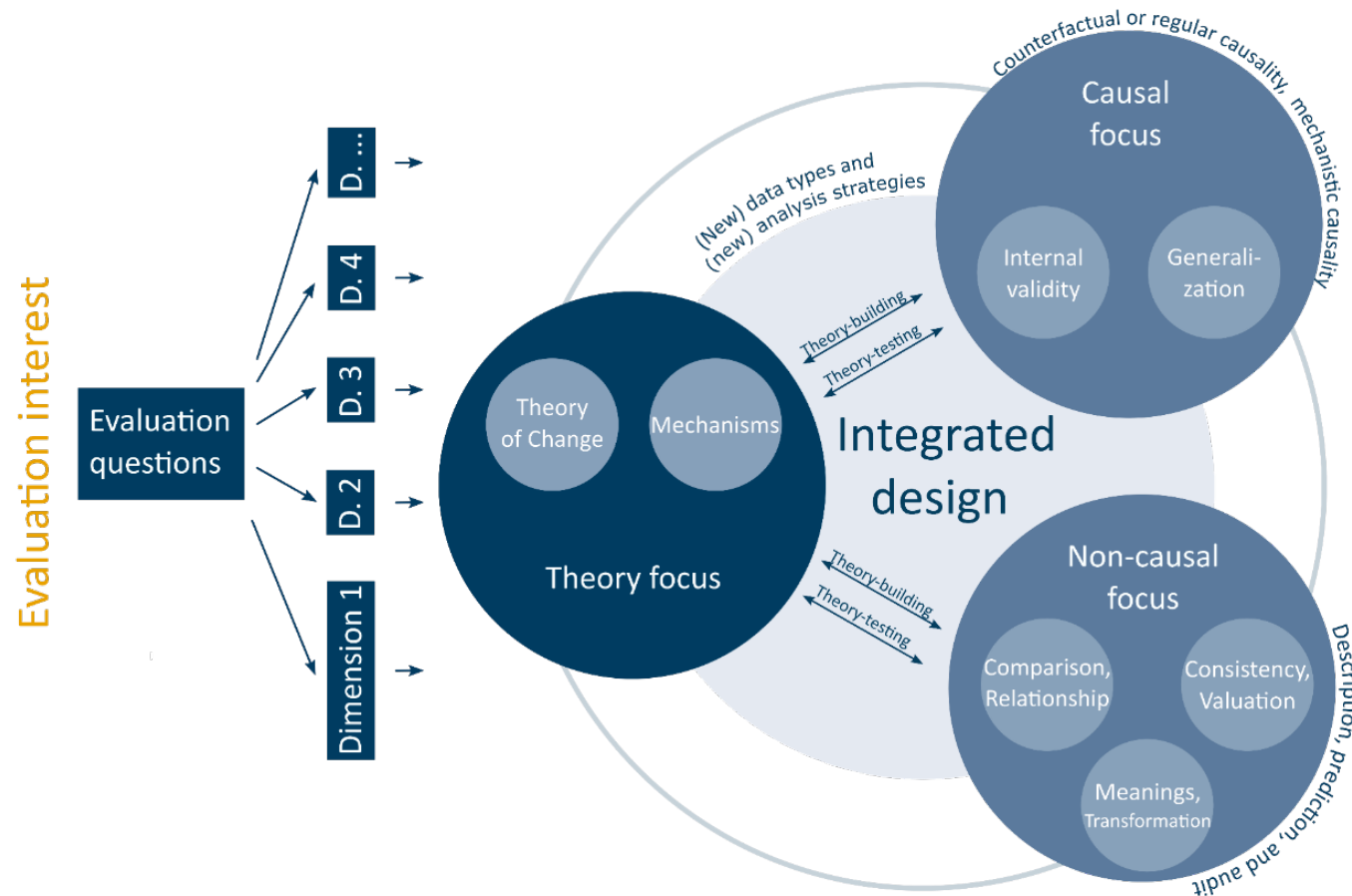
Kai Rompczyk

DEval, German Institute for Development Evaluation

IEG Symposium April 11, 2024

Unlocking the Potential of Geospatial Analysis for Evaluation

DEval | German Institute for Development Evaluation

- Founded in November 2012, Bonn

- Evaluation and federal research institute

- Works at the interface between science, politics and implementing organizations

- Aims at providing scientifically robust evidence

- Pursues improvements at the strategic level with its evaluatory work

ECD

Methods and standards

Disseminating knowledge and networking

Evaluations

Integrating geospatial analysis in complex Evaluation Designs

Attributes of evaluation object, context and stakeholder constellation

Evaluation interest

Evaluation questions

D. ...
D. 4
D. 3
D. 2
Dimension 1

Counterfactual or regular causality, mechanistic causality

(New) data types and (new) analysis strategies

Causal focus
Internal validity
Generali-zation

Theory-building
Theory-testing

Theory focus
Theory of Change
Mechanisms

Integrated design

Non-causal focus
Comparison, Relationship
Consistency, Valuation
Meanings, Transformation

Theory-building
Theory-testing

Description, prediction, and audit

Practicability and efficiency

Can we use geodata?

(a) Original image before disaster (T0)   (b) LCLU classification before disaster (T0)   (c) LCLU classification directly after disaster (T1)
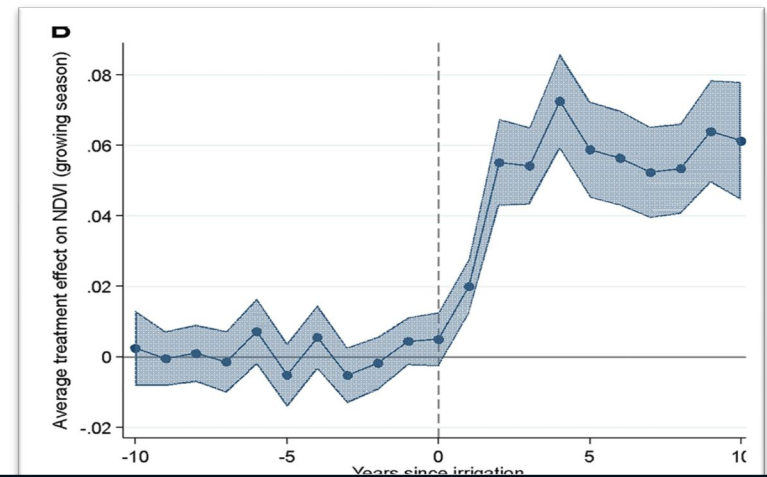
Geospatial correlation

Machine learning based approach for land-use and land-cover classification after typhoon Hayan on the Phillippines (Lech, 2020)

Descriptive Mapping

Deforestation in Cameroon – (Global Forest Watch 2023)
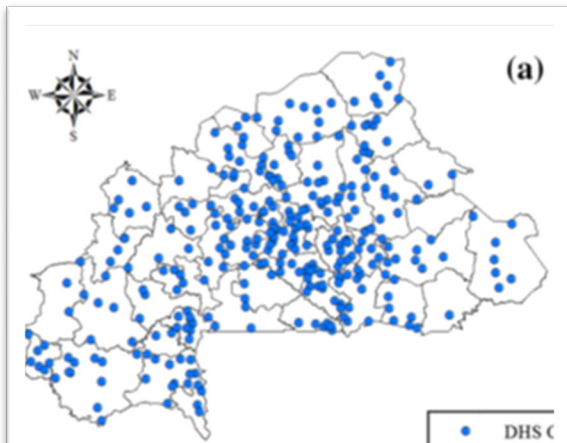
Can we use geodata?

Data integration

DHS Cluster Points Burkina Faso (Nawrotzki 2016)

Geospatial Impact evaluation

Effects of irrigation projects in Mali (NDVI) (BenYishay et al., 2024)
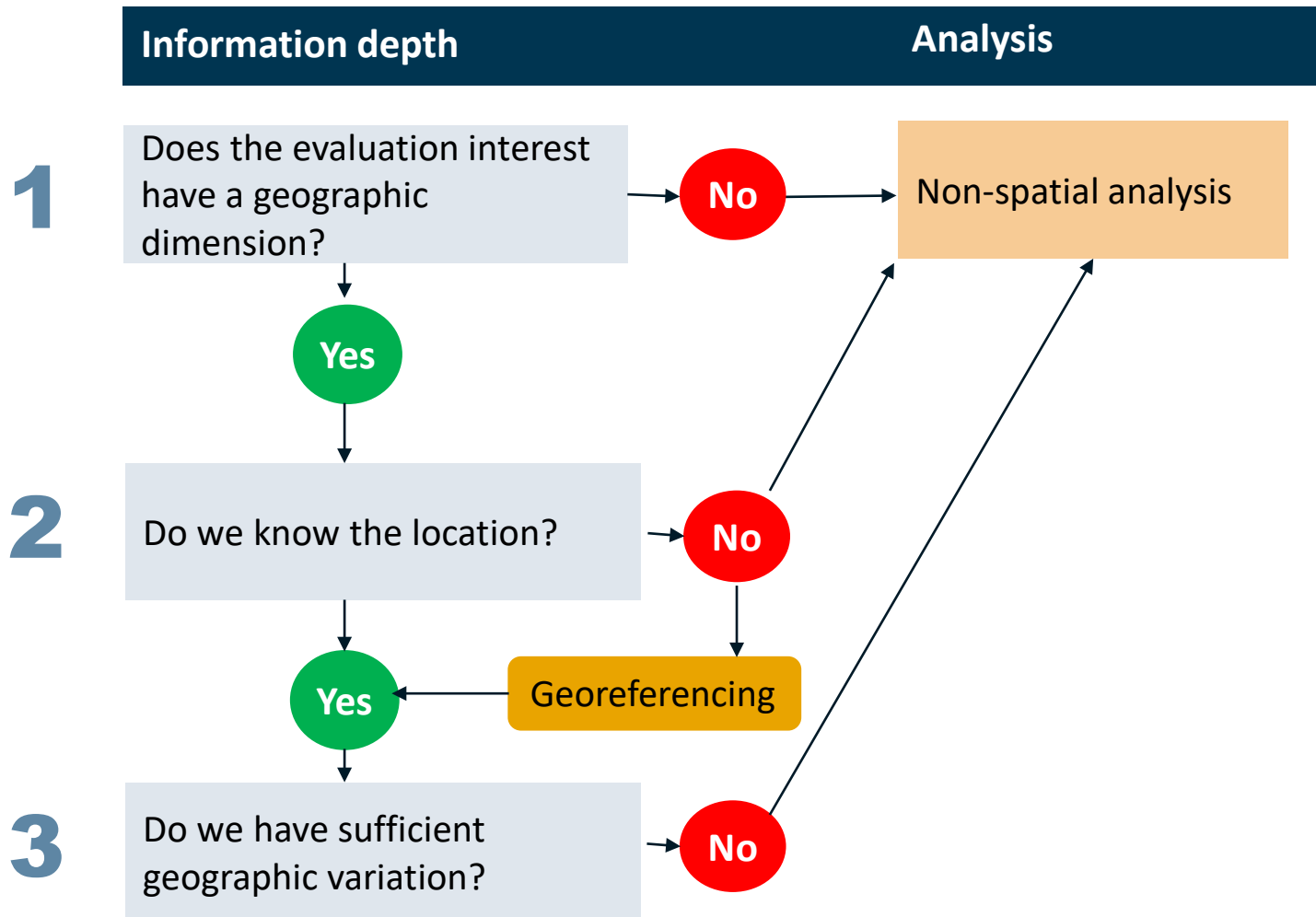
Can we use geodata?

The Geodata decision tree

- An orientation framework

- 5 guiding questions

- Depending on the depth of information in the geodata

- Recommendations for analysis types
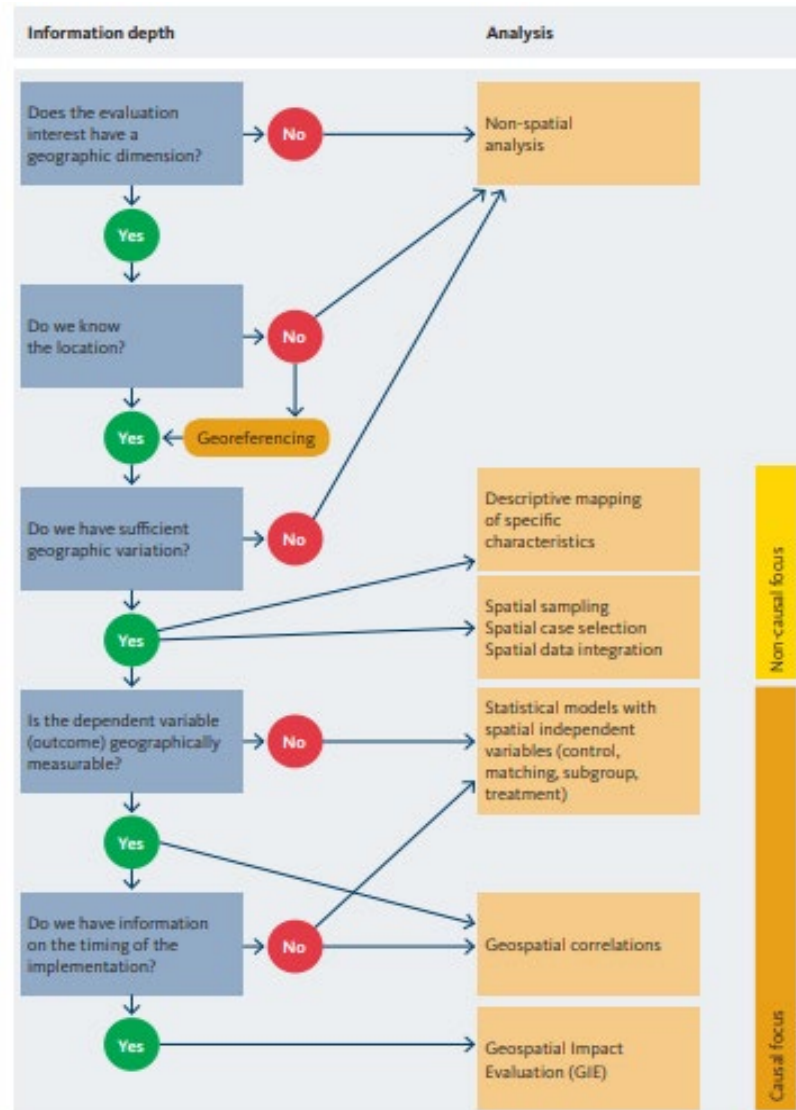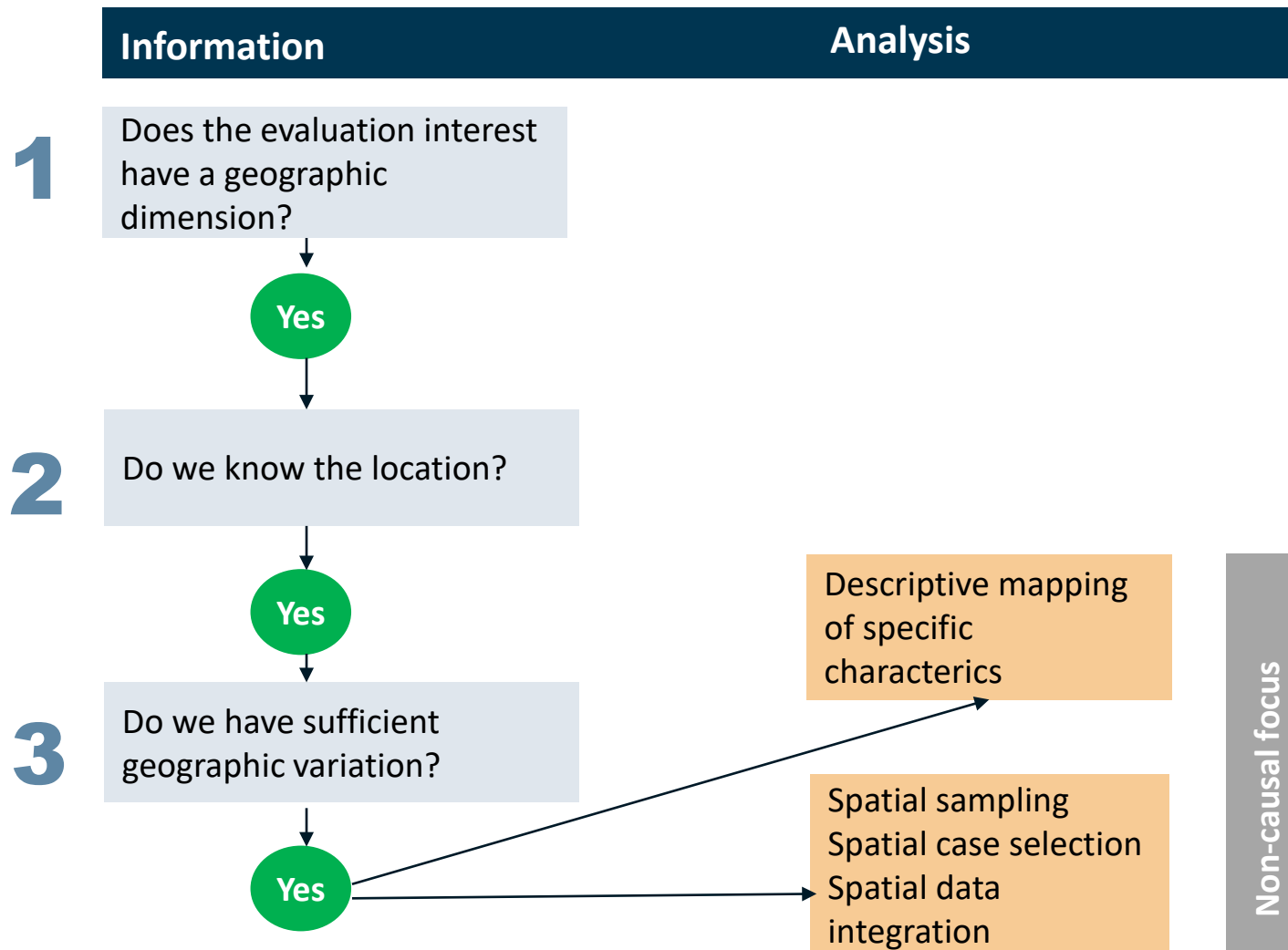

Nawrotzki 2019

When we cannot use geodata

Information depth	Analysis

1 Does the evaluation interest have a geographic dimension? → **No** → Non-spatial analysis

↓ **Yes**

2 Do we know the location? → **No**

↓ **Yes**

No → Georeferencing → **Yes**

3 Do we have sufficient geographic variation? → **No**

Georeferenced data

Georeferenced data can be sourced from a variety of origins.

- Manually georeferencing

- Incorporating geocoding in the data collection process

- Gathering data that include geospatial information

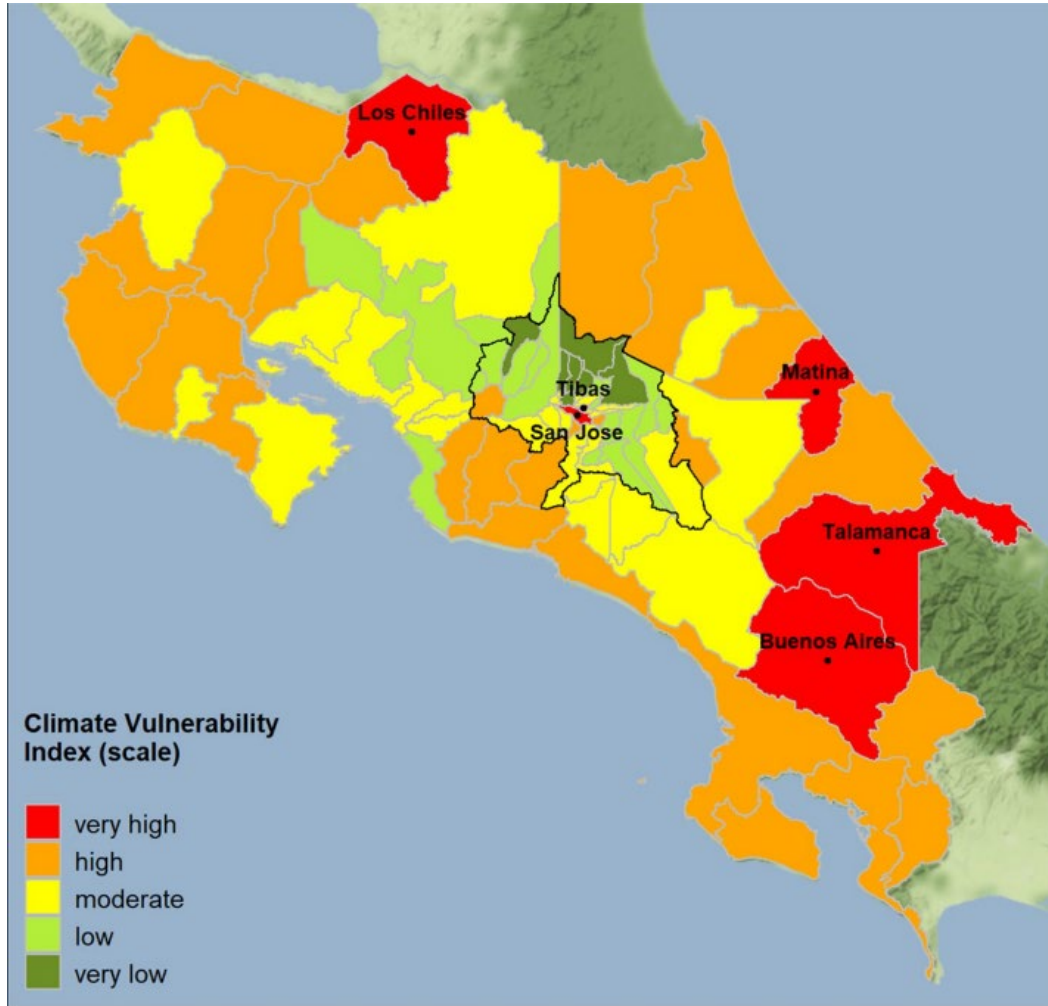- Databases with georeferenced information

Georeferenced data

- Survey data
- Social media
- Mobile phone data
- Public media
- Event data

- Observational data/ground truth
- Satellite data
- Project data
- Photos
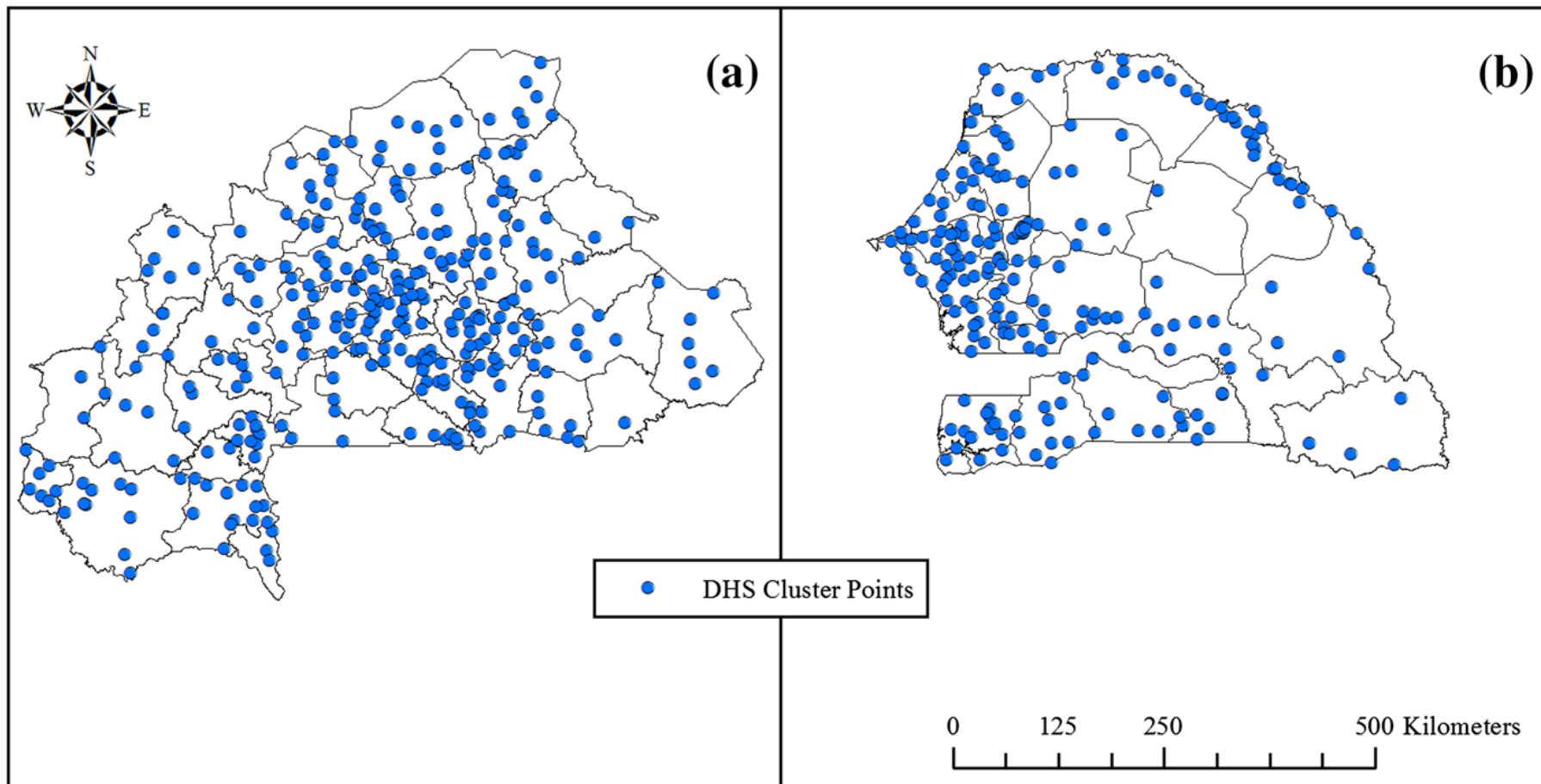
Analysis with non causal focus

Information	Analysis

1 Does the evaluation interest have a geographic dimension?

↓

Yes

↓

2 Do we know the location?

↓

Yes

↓

3 Do we have sufficient geographic variation?

↓

Yes

Descriptive mapping of specific characterics

Spatial sampling
Spatial case selection
Spatial data integration

Non-causal focus

Descriptive Mapping

Dimension	Component	Relationship
Exposure	Heat months	+
	Drought months	+
	Flood risk	+
Sensitivity	Asset index	−
	Work in climate sensitive industry	+
	Population density	+
	Tree cover	−
Adaptive capacity	Employment	+
	Literacy	+
	Remittances received	+
	Infant mortality	−
	Road density	+
	Distance from health center	−

Nawrotzki et al. (2023)

Integrating Different Data Sources

Burkina Faso   Senegal
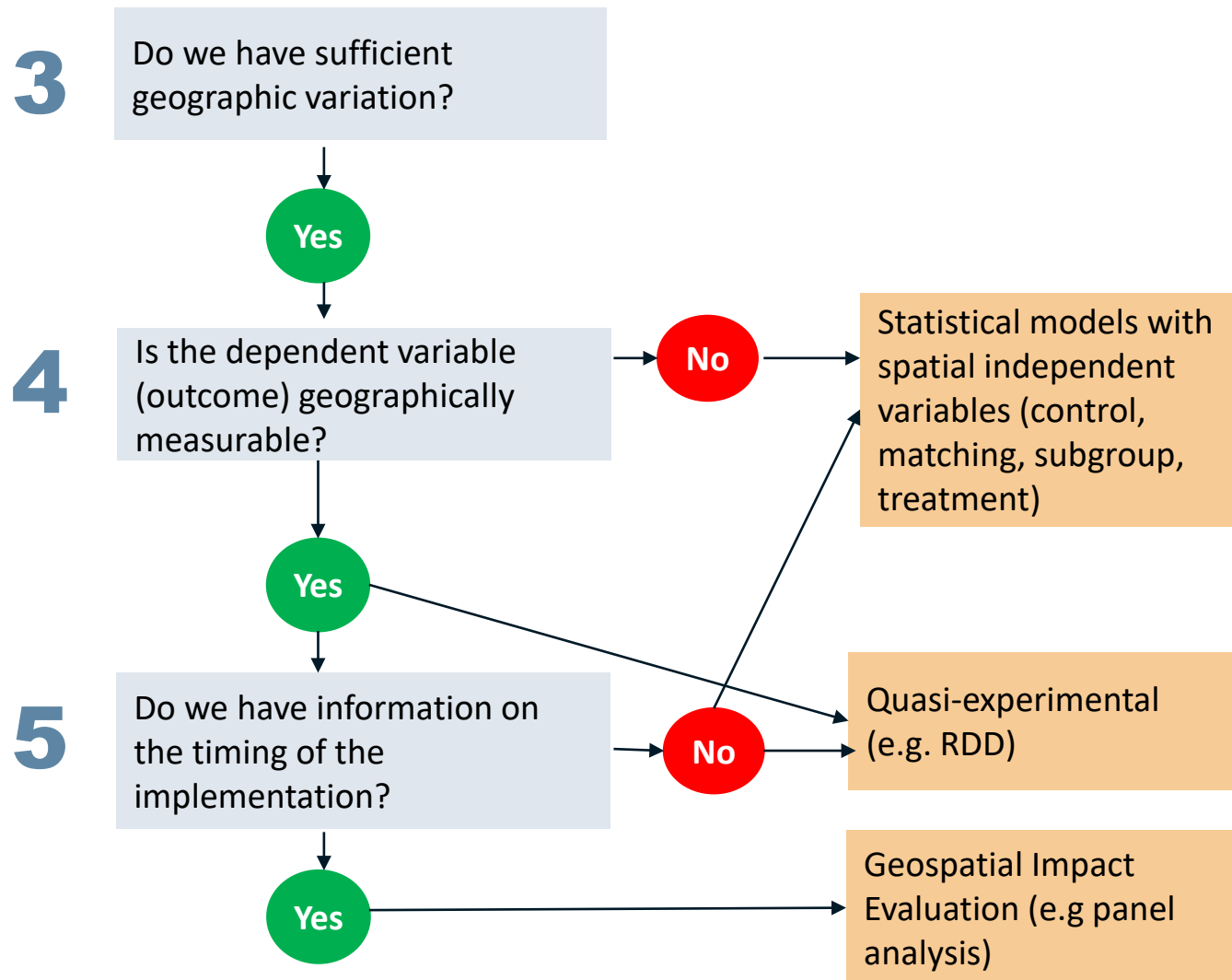
Nawrotzki et al. (2016)

Case selection
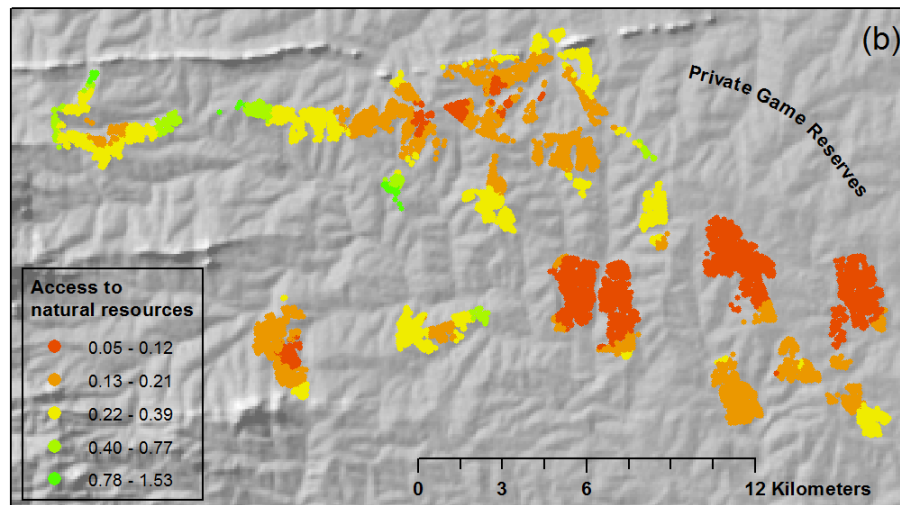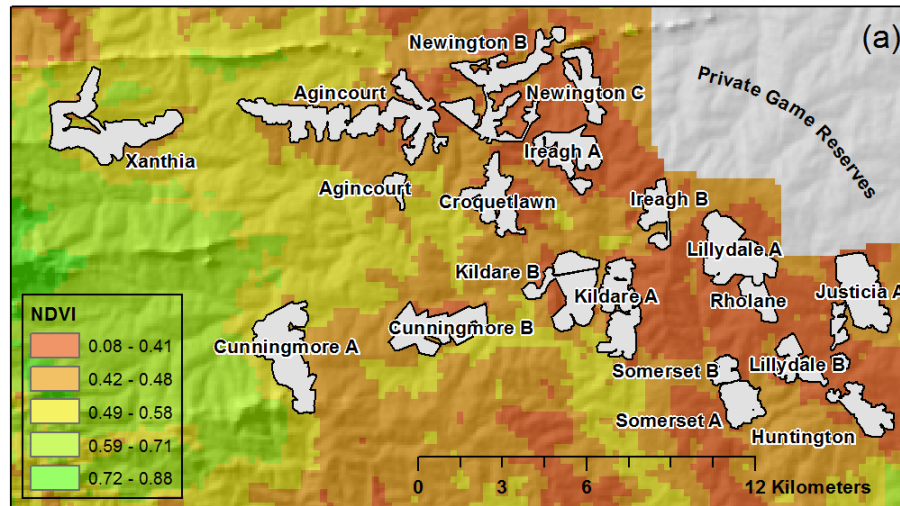
Nawrotzki et al. (2014)   **NDVI , South Africa**

Analysis with causal focus

3 Do we have sufficient geographic variation?

→ **Yes**

4 Is the dependent variable (outcome) geographically measurable?

→ **No** → Statistical models with spatial independent variables (control, matching, subgroup, treatment)

↓ **Yes**

5 Do we have information on the timing of the implementation?

→ **No** → Quasi-experimental (e.g. RDD)

↓ **Yes**

→ Geospatial Impact Evaluation (e.g panel analysis)

Analysis with causal focus

3 Do we have sufficient geographic variation?

Yes

4 Is the dependent variable (outcome) geographically measurable?

No

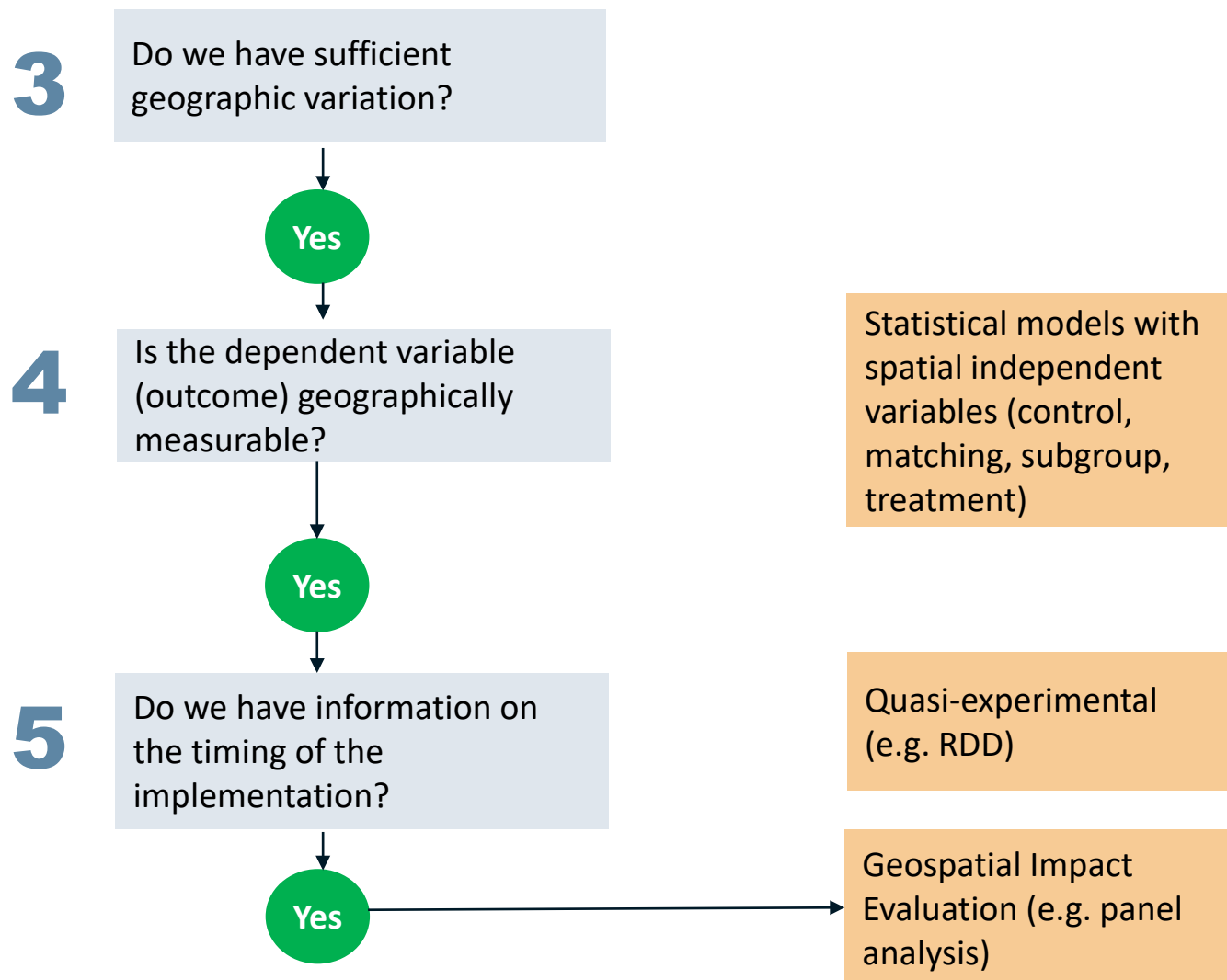Statistical models with spatial independent variables (control, matching, subgroup, treatment)

Spatial independent variables

- Rural vs. Urban
- Country
- District
- Climate zones
- Proximity to next school
- …

# Statistical models with spatial independent variables
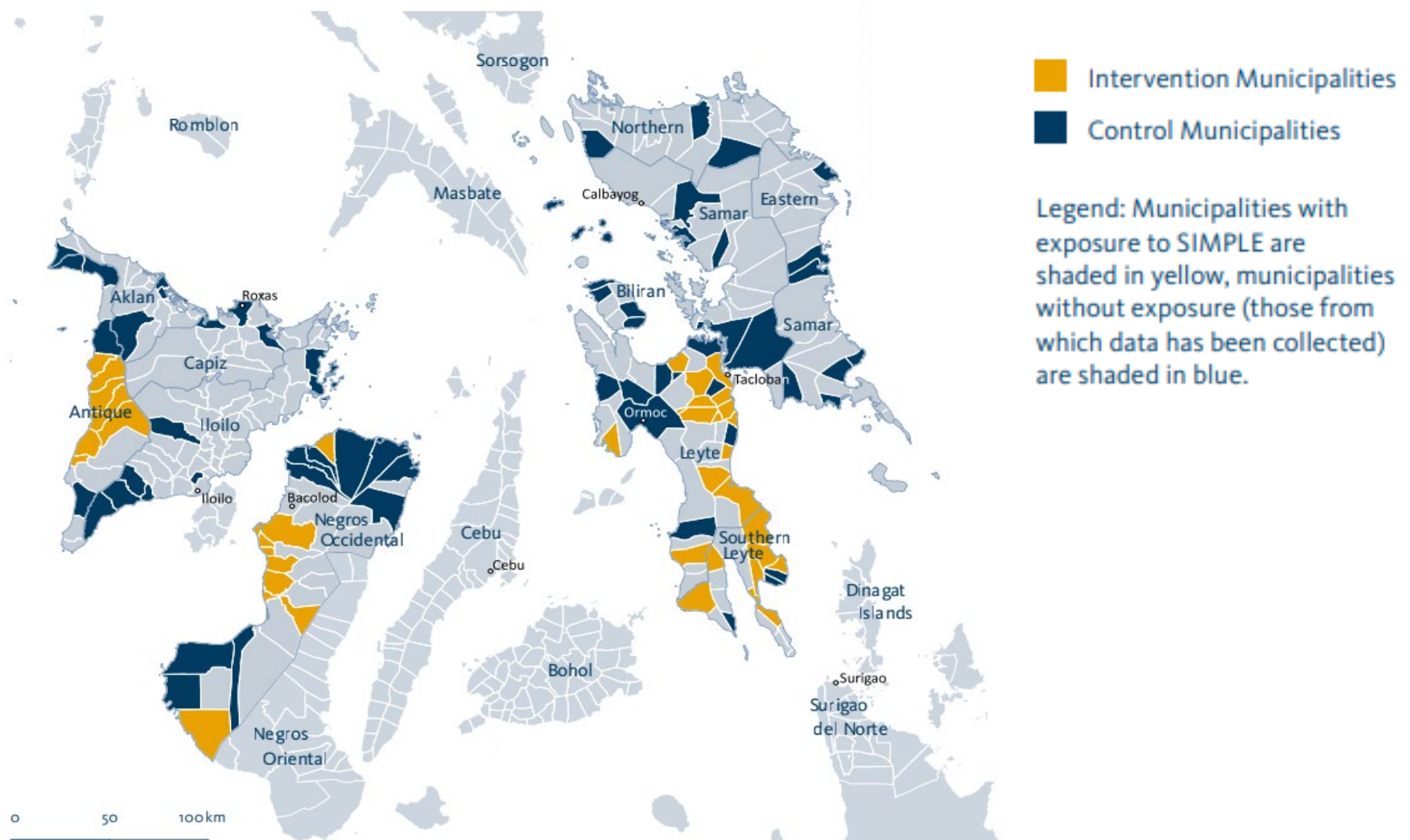
Schustereder (2016): Donor-assisted land-use planning in the Philippines

Analysis with causal focus

3 Do we have sufficient geographic variation?

Yes

4 Is the dependent variable (outcome) geographically measurable?

Statistical models with spatial independent variables (control, matching, subgroup, treatment)

Yes

5 Do we have information on the timing of the implementation?

Quasi-experimental (e.g. RDD)

Yes

Geospatial Impact Evaluation (e.g. panel analysis)

Geospatial impact evaluation

"A GIE attempts to causally connect the intervention with geographically measurable changes in the environment (BenYishay 2017)
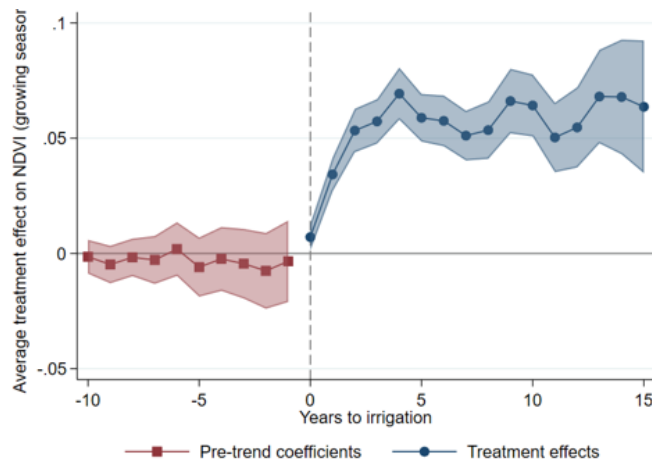

BenYishay et al. 2023
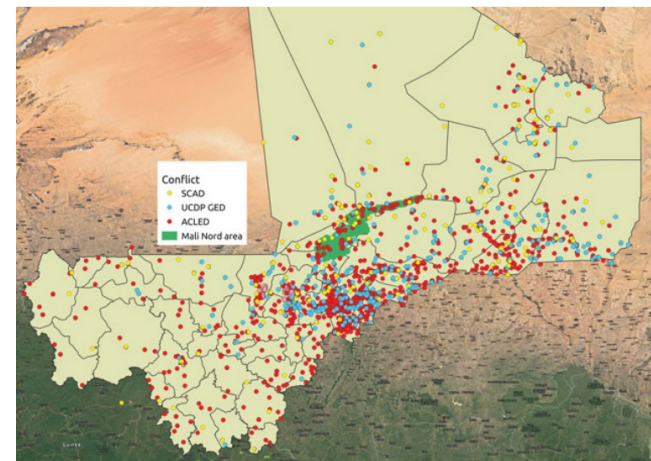
Effects of Irrigation in Mali

Conlusion

The geodata decision tree

- An **orientation framework** in the inception phase

- Determine wether the evaluation question has a **geographic dimension**

- Ascertain the **availability** of the necessary geographic data

- Where feasible, utilize and support **georeferencing**

- Geodata can enhance both causal and non-causal evaluation questions, demonstrating its strength in **method integration**
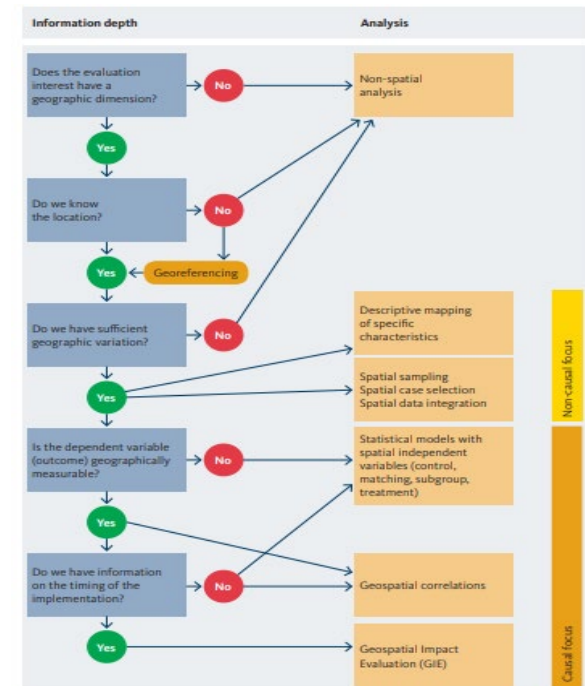
BenYishay, Ariel, Daniel Runfola, Rachel Trichler, Carrie Dolan, Seth Goodman, Bradley Parks, Jeffery Tanner, Silke Heuser, Geeta Batra, and Anupam Anand (2017), A Primer on Geospatial Impact Evaluation Methods, Tools, and Applications. AidData Working Paper #44. Williamsburg, VA: AidData at William & Mary.

BenYishay, A., Sayers, R., Singh, K., Goodman, S., Walker, M., Traore, S., Rauschenbach, M., Noltze, M. (2024), Irrigation strengthens climate resilience: Long-term evidence from Mali using satellites and surveys, PNAS Nexus, Volume 3, Issue 2.

Lech, M. et al. (2020), "A Proof-of-Concept of Integrating Machine Learning, Remote Sensing, and Survey Data in Evaluations. The Measurement of Disaster Resilience in the Philippines", *DEval Discussion Paper 1/2020*, German Institute for Development Evaluation (DEval), Bonn.

Nawrotzki, R. (2019), "The Geodata Decision Tree: Using Geodata for Evaluations", *DEval Policy Brief 3/2019*, German Institute for Development Evaluation (DEval), Bonn.

Nawrotzki, R. J. et al. (2023), "Climate change vulnerability hotspots in Costa Rica: constructing a sub-national index", *Journal of Environmental Studies and Sciences,* Vol. 13, No. 3.

Schustereder, G. (2016), "Donor-Assisted Land-use Planning in the Philippines: Insights from a Multi-Level Survey", German Institute for Development Evaluation (DEval), Bonn.
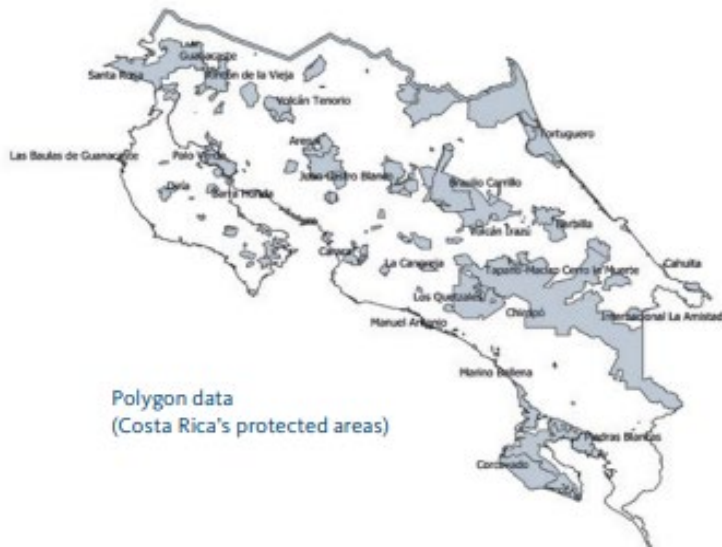
BACKUP

Types of geodata
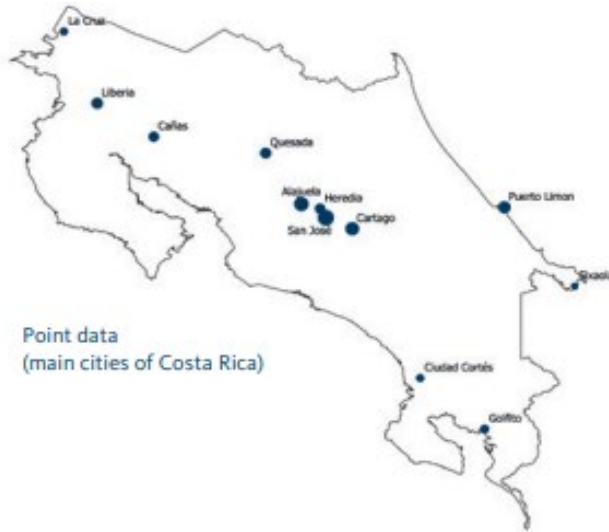
Point data
(main cities of Costa Rica)
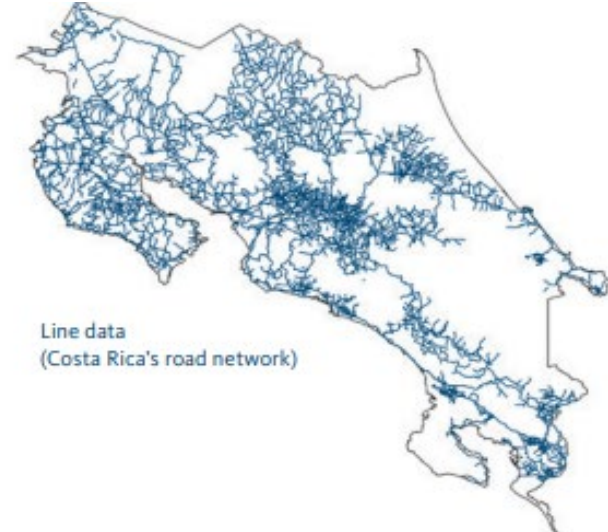
Line data
(Costa Rica's road network)

Polygon data
(Costa Rica's protected areas)

Raster data
(tree cover in Costa Rica)