Geospatial Impact Evaluation Methods, Tools & Applications

Day 1, Part 1: Introduction

IEO Independent Evaluation Office
GLOBAL ENVIRONMENT FACILITY

Schedule

Day 1 - Discussion, Theory and Principles of Geospatial Impact Evaluation

Goal: Provide working knowledge on what Geospatial Impact Evaluation is, its strengths and weaknesses, required inputs and types of outputs.

Day 2 - Hands on with Spatial Data

Goal: Provide hands-on experience with spatial data to expose you to the potential and limitations of a variety of spatial data sources. Teach basic GIS tools and techniques, as well as provide examples of exploratory impact evaluation.

Important Notes for Today

- **Please interrupt with any questions or comments!**

- **After today's sessions we will help you configure your personal laptops for use in the hands-on training sessions tomorrow. If you will be using a GEF-issued laptop or need to borrow one of ours, please let us know.**

- **If there is a topic you would like more depth on than provided in today's sessions, let us know during lunch or this evening. We will do our best to accommodate smaller discussion groups around specific topics.**

Why we're here today

- **Geospatial data is being used by the development community at an increasingly rapid pace, with GEF as a thought leader.**
    - **Geospatial Monitoring**
        - **The use of geographic data to describe the changing contextual circumstance around intervention sites.**
    - **Geospatial Impact Evaluation (GIE)**
        - **The use of known geographic locations of interventions for causal identification of an interventions impact.**
    - **Geospatial Targeting**
        - **The use of geographic data to identify candidate areas for interventions during project planning or implementation.**

What is Geospatial Data?

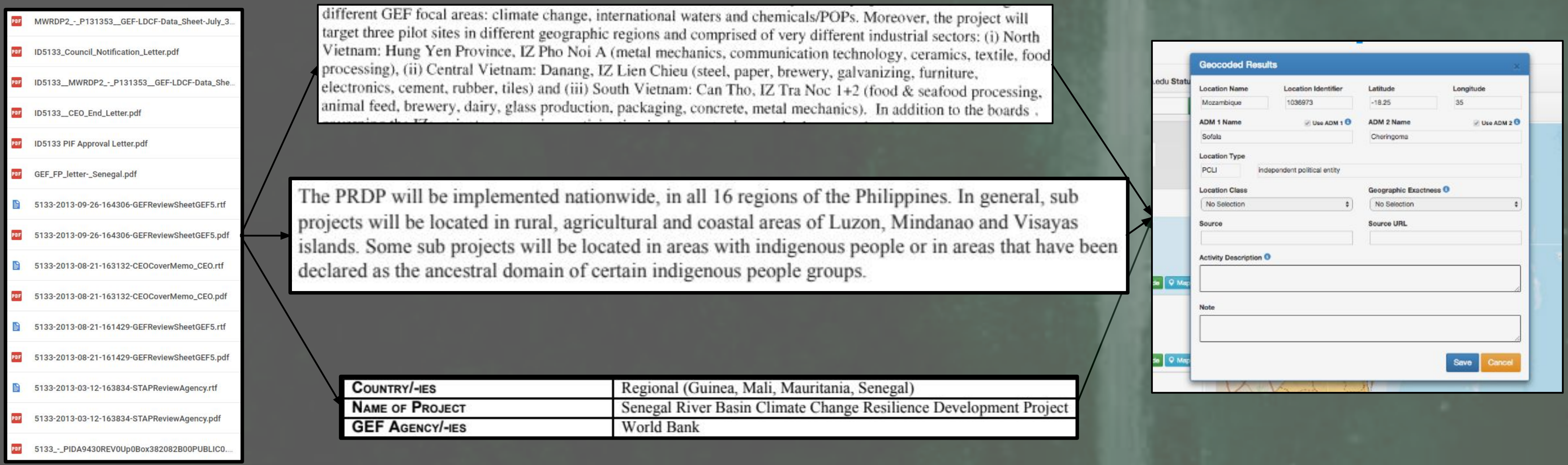
- **Any data with a latitude and longitude, line or region at which it occurred.**

- **Because all geospatial data has a spatial location, you can always merge datasets together (though this can be difficult).**

- **This merging enables a wide range of activities that would be infeasible or impossible with other data types.**

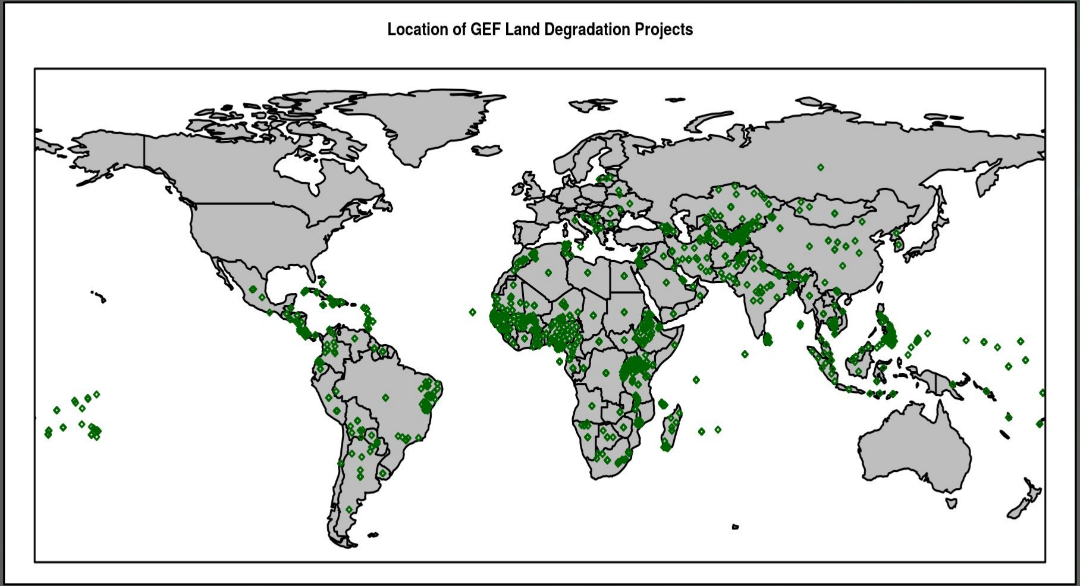
What is Geospatial Data?: Geoparsing and Geocoding

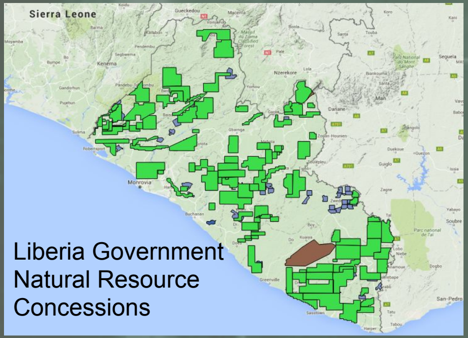
- Frequently data retrieved from PDFs or other documents describing project implementation.
- "Geoparsing" is the process of identifying place names in text documents.
- "Geocoding" is the process of making these place names (or other geographic locations) machine readable.



What is Geospatial Data?: Geoparsing and Geocoding

- Results of Geocoding can have multiple levels of geographic precision.
  - Sometimes a region, country, or administrative district is known.
  - Sometimes exact points (i.e., farming plots) or polygons are identified.
  - Other times, a line such as a road or irrigation canal is found.



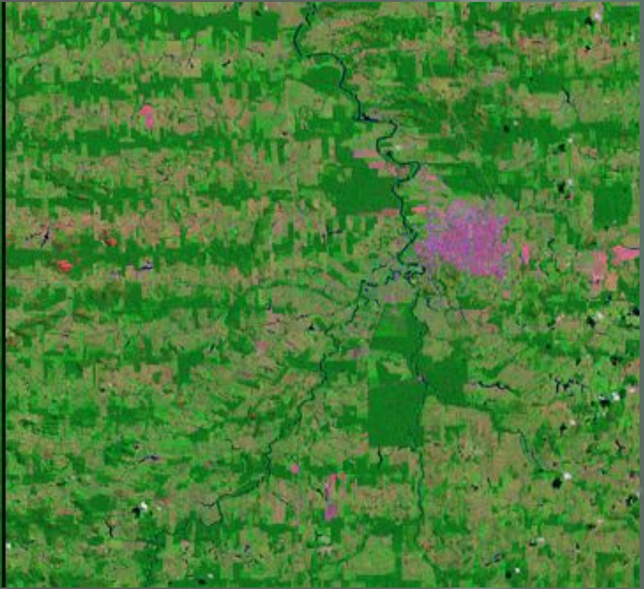


What is Geospatial Data?: Satellite Data

- Both long- and short-term records are available. Contemporary satellites (~>2000) have much better data quality.
- Can be used to establish baselines, monitor projects, as well as ascertain impact.
- Many technical challenges to the use of satellite data (more on this later).

Amazon, 1975

Amazon, 2012



What is Geospatial Data?: Survey Data

- **Over the last two decades, more surveys have included georeferenced household-level information.**

- **Examples include the Demographic and Health Surveys, AfroBarometer, Living Standards Measurement Study, and many more.**

- **Unique challenges come with using this data: frequently it is discrete ("What is this households roof made of?"), rather than continuous ("How many trees are there nearby?"). Spatial tools are generally built for continuous use cases.**

- **Anonymization is a frequent attribute of these data - geospatial errors are intentionally introduced to mask individual households, or results are only provided at coarser levels of analysis.**

Geospatial Data: Integration

- One of the biggest strengths of Geospatial Data is in integration. All of the aforementioned data types can be integrated with one another.

- However, there are many challenges.
  - Unit of observation - Buffers, administrative units, grid cells
  - Spatial resolution
  - Data error and inconsistency
  - Spatial imprecision

Geospatial Data: Defining a Unit of Observation

- In many studies, the unit of observation for an intervention is well defined:
  - Giving a pill to two individuals for a disease - the individual person is the unit of observation.
  - Applying fertilizer to a number of fields - the individual fields is the unit of observation.

- In geospatial analyses the unit of observation must frequently be theoretically defended.
  - If you build a canal, for how many meters or kilometers from that canal do you believe you could observe an impact?
  - If you conduct a training on sustainable farming techniques in a city center, over what geographic region do you anticipate impacts might be seen?

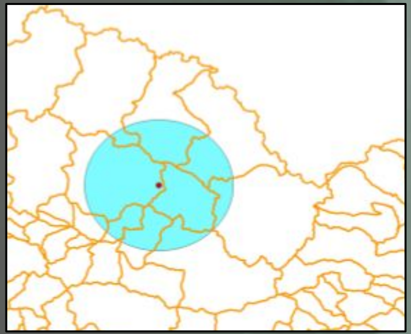
- Before any statistical analyses can be done, the question of unit of observation must first be answered.

Geospatial Data: Defining a Unit of Observation

- There are many techniques for defining units of observation.
  - Buffers
    - Used in portfolio-scale analyses for the World Bank, GEF
    - Assign a specific distance (i.e., 10km) from the known intervention point, and test for impact effects within that area.
    - Can be applied to point-based interventions (i.e., a training) or line-based interventions (i.e., a canal or road).

Geocoded Interventions (Line-based roads)

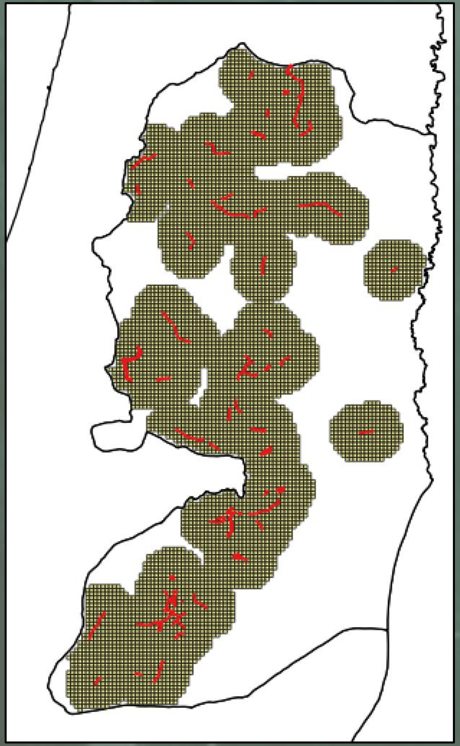
5km buffers, clipped to administrative zone.



Point-based buffer, unclipped.

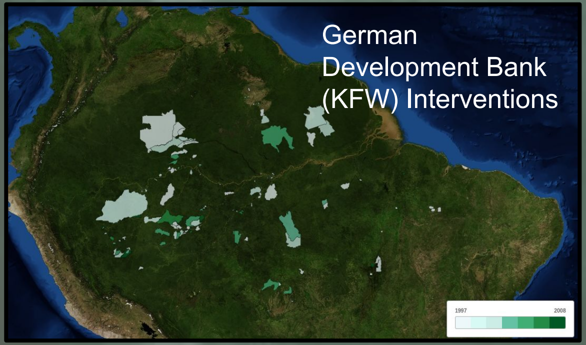
Geospatial Data: Defining a Unit of Observation

- **Grid cells over the study region.**
  - Advantages
    - Many units of observation
    - Can identify precise changes over study region.
    - Can mitigate challenge of not knowing the exact buffer radius you need to select.
  - Disadvantages
    - Both study region and grid cell size must be defended theoretically.
    - Underlying data must support chosen resolutions.
    - Statistical challenges are larger (independence of units can be scale dependent).



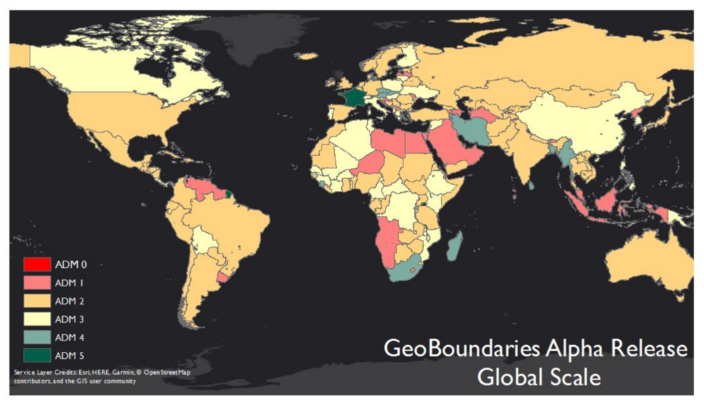
Geospatial Data: Defining a Unit of Observation

- **Explicit Polygons of Interventions**
  - Advantages
    - Works very well when the exact area an intervention would have impacted is known (i.e., in the case of the KFW indigenous lands protection program).
    - No or very little spatial imprecision / uncertainty.
  - Disadvantages
    - Many programs do not have explicitly defined spatial bounds.
    - The costs to generate explicit bounds can be high.



Geospatial Data: Defining a Unit of Observation

- **Official Boundaries**
  - Advantages
    - Helpful when an intervention is anticipated to affect an entire decision-making unit.
    - Frequently the same units used for census activities.
  - Disadvantages
    - Can change in unmeasured ways over time.
    - Can be of variable size and have variable underlying measurement qualities.



Geospatial Data: Integration

- **The analysis you undertake is limited by the interaction between your chosen unit of analysis and the quality of data you seek to integrate.**

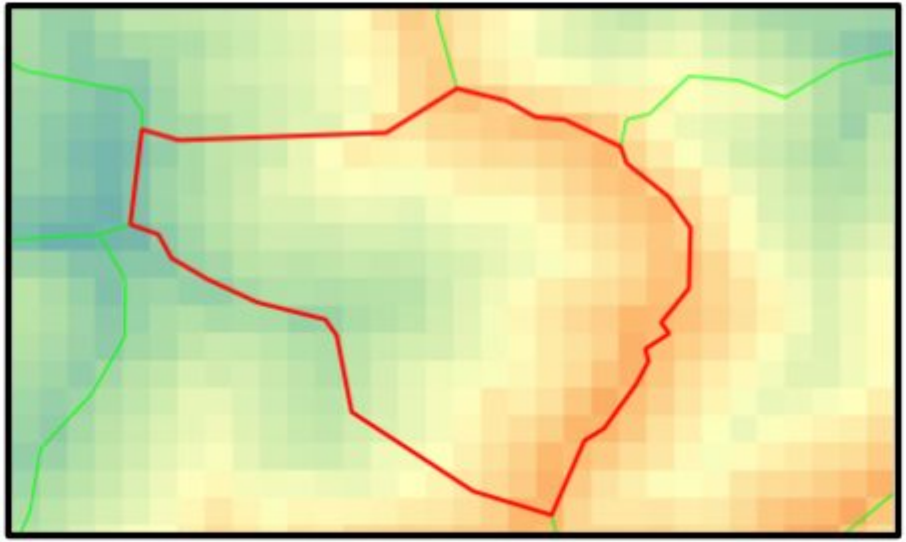
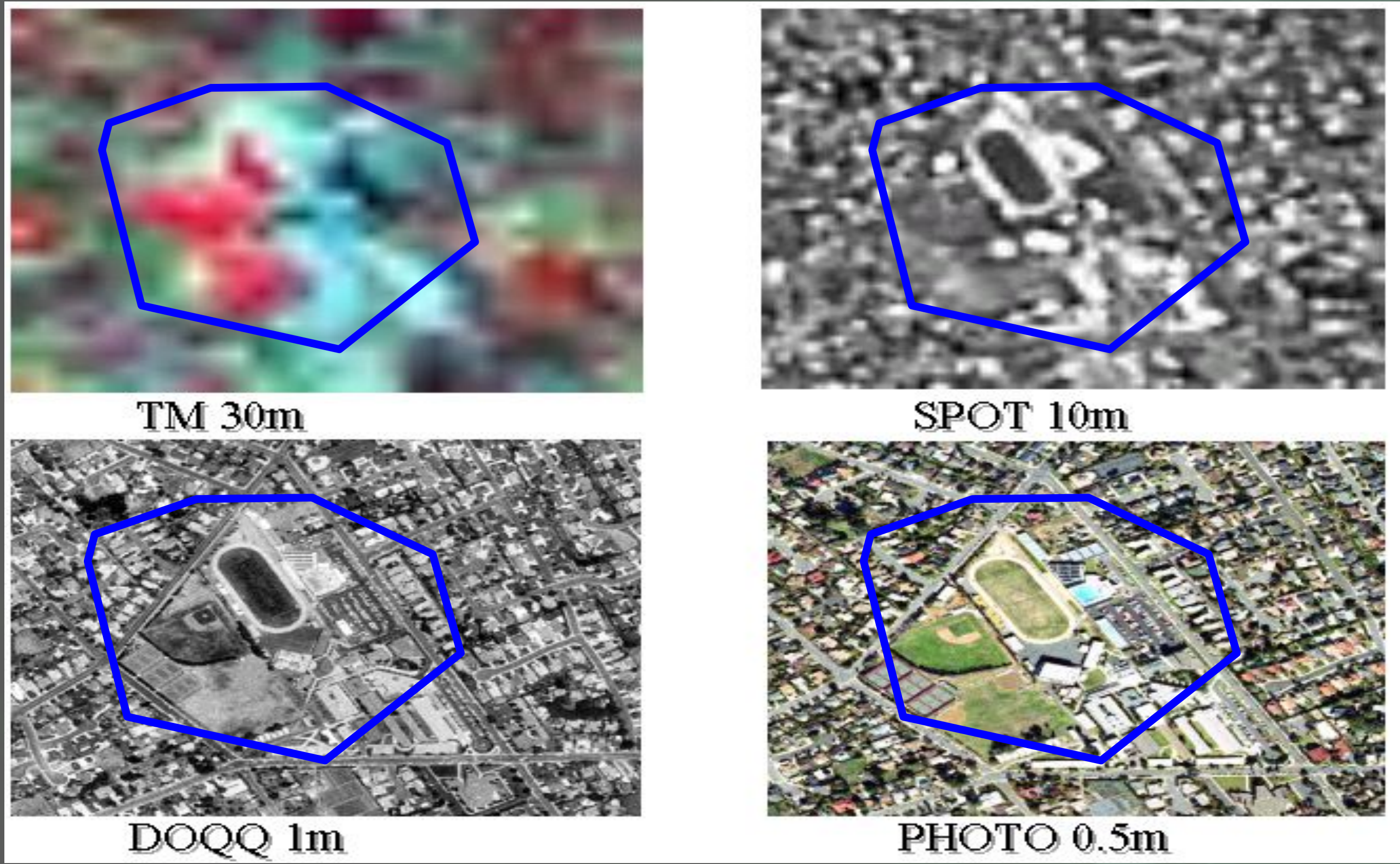


Figure 2. The village of Basa in Nepal covers approximately 185 cells of the elevation raster. Elevation values in Basa range from around 1200 meters above sea level, represented by the red-orange pixels on the right side of the village, to over 3000 meters on the left side in the blue-green cells.

Table 1. Extract Results	
Name	Basa
Count	185
Sum	393,202
Mean	2125.416

Geospatial Data: Integration



Geospatial Data: Integration

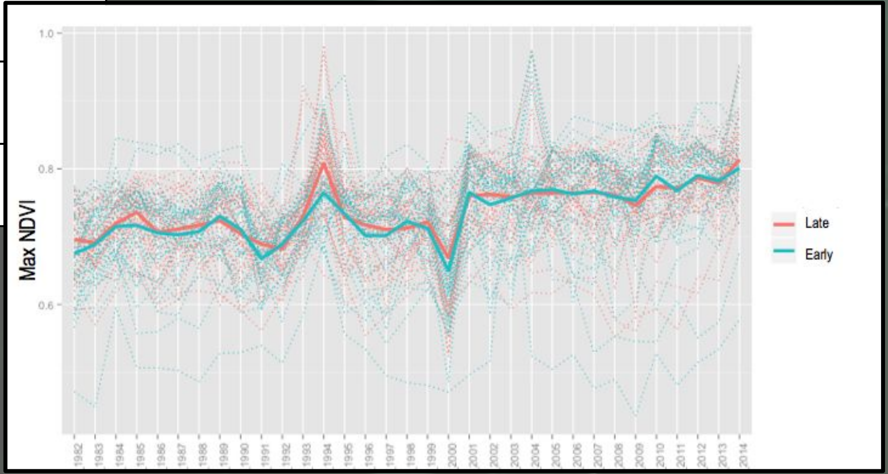
- **Ultimate goal is to get - for each unit of observation - a column of data representing a relevant spatial dataset.**

Unit of Observation	Satellite-measured vegetation in 1990	Satellite-measured vegetation in 2005
A	0.15	0.10
B	0.13	0.12
C	0.16	0.17
D	0.20	0.19

Geospatial Data: Integration

- **Ultimate goal is to get - for each unit of observation - a column of data representing a relevant spatial dataset.**

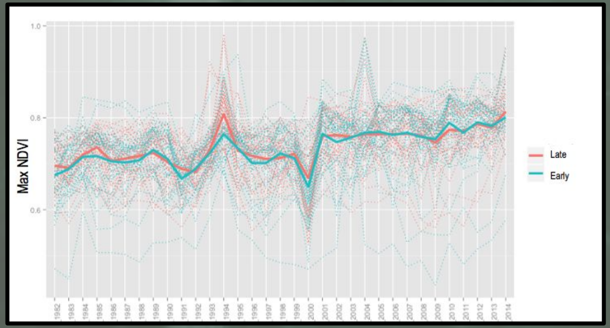
Unit of Observation	Satellite-measured vegetation in 1990	Satellite-measured vegetation in 2005
A	0.15	0.10
B	0.13	0.12
C	0.16	0.17
D	0.20	0.19



Land Rights

(Outcome: Forest Protection)

Translating your impact trajectory / theory of change into a quantifiable trend can sometimes prove a unique challenge, but is key : depending on your time frame, different spatial datasets providing measurements may or may not be available.

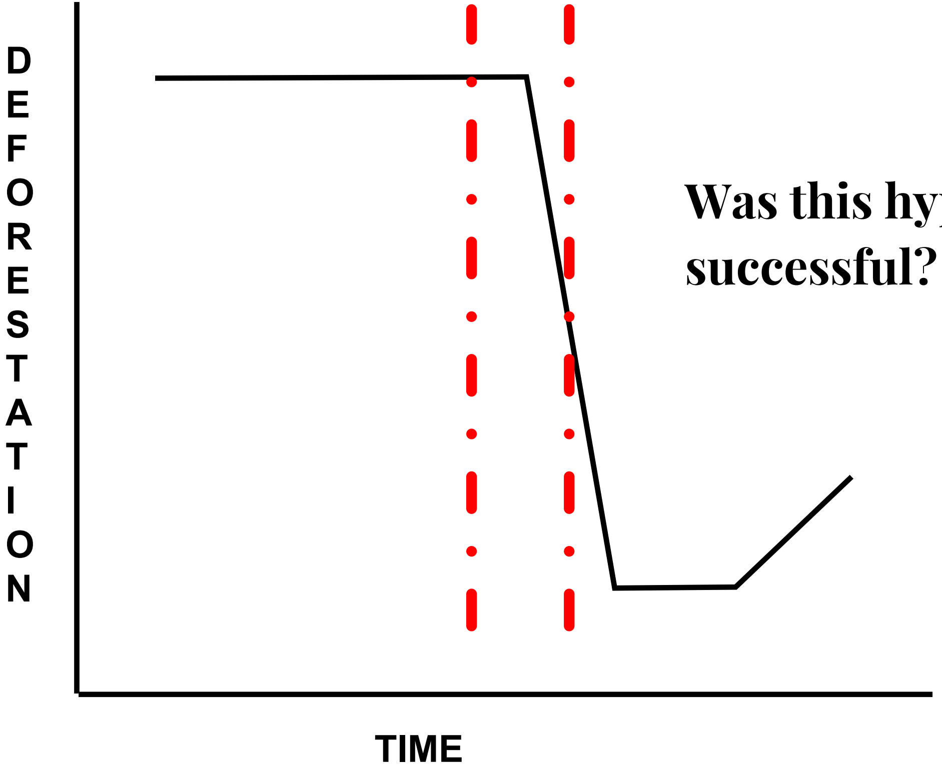


Geospatial Impact Evaluation Methods, Tools & Applications

Independent Evaluation Office
GLOBAL ENVIRONMENT FACILITY

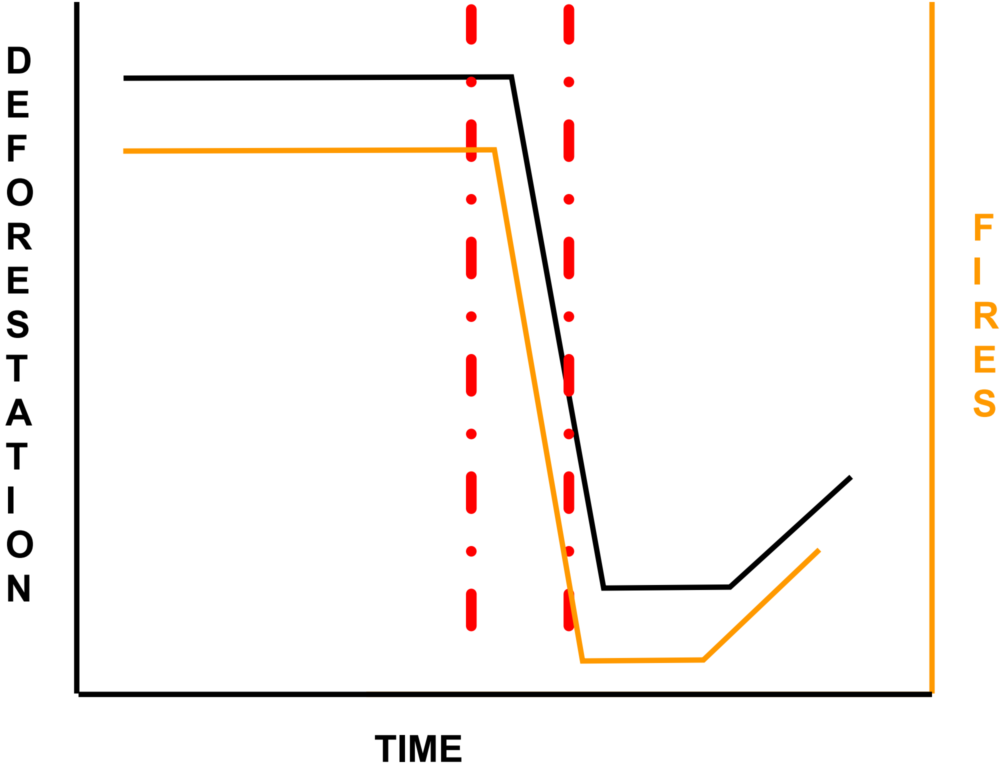
Day 1, Part 2: GIE in the GEF Context

Impact Evaluation: The Basics



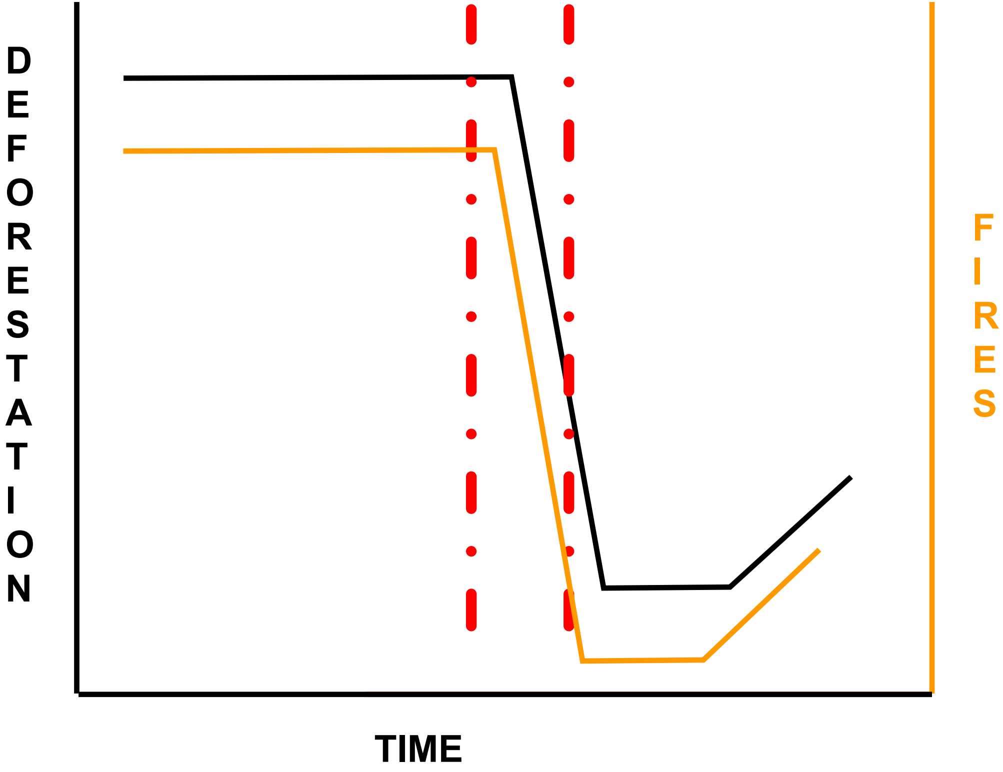
Was this hypothetical project successful?

Impact Evaluation: The Basics



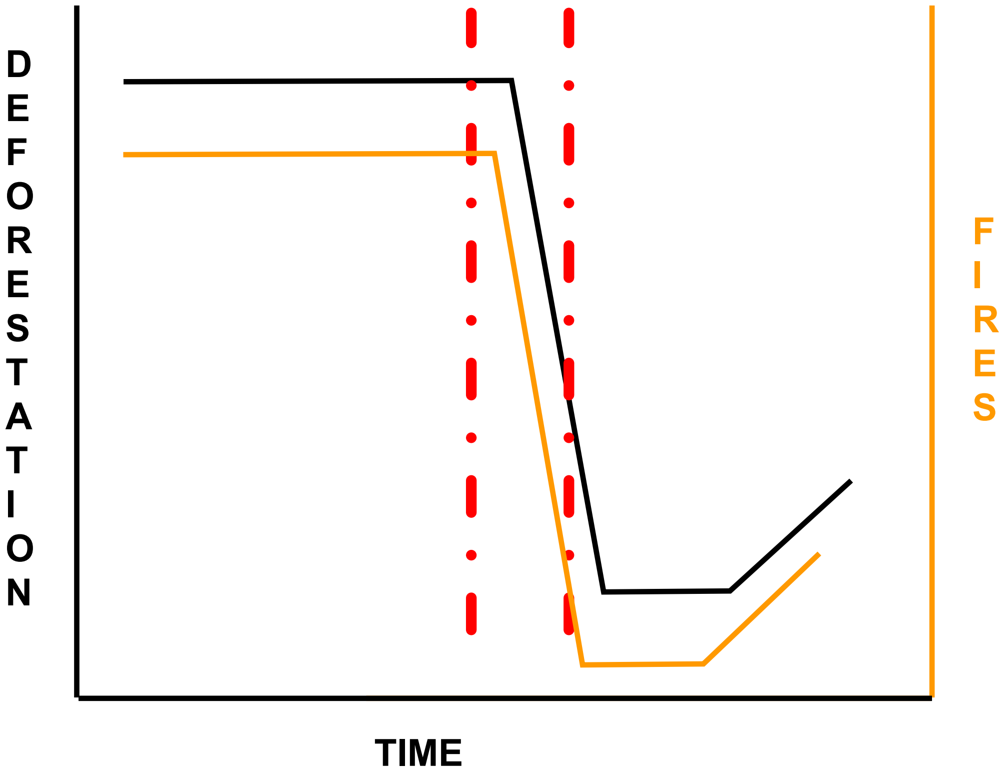
Was this hypothetical project successful?

Impact Evaluation: The Basics



What other primary and secondary drivers of deforestation can you think of?

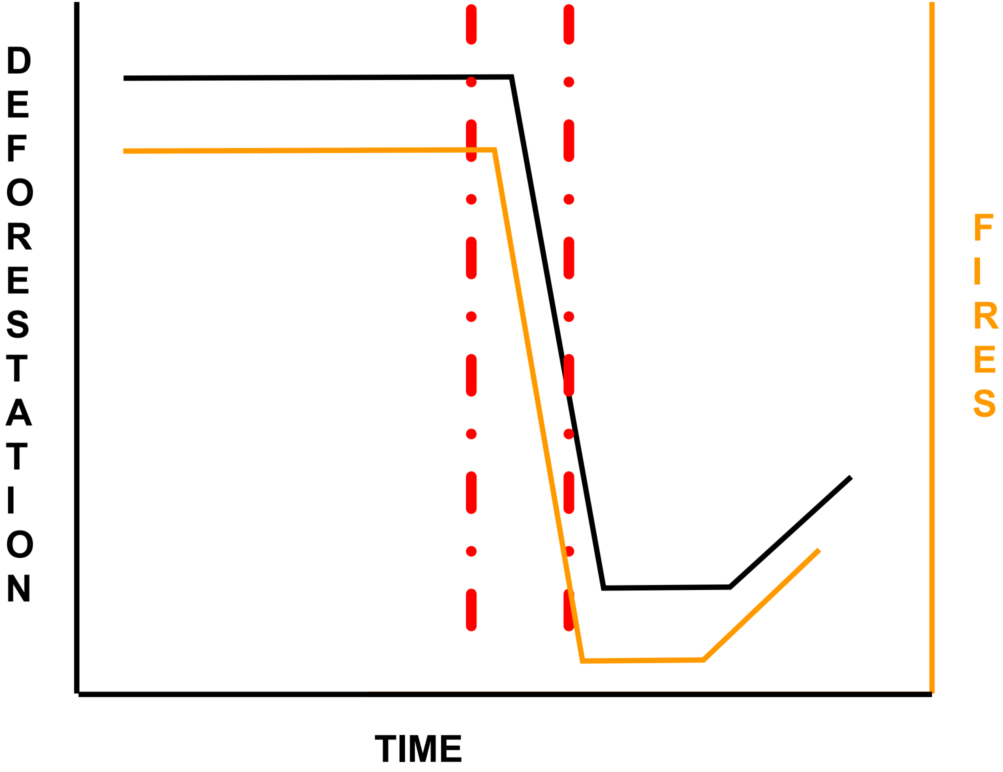
Impact Evaluation: The Basics



What other primary and secondary drivers of deforestation can you think of?

Can we measure them all?

Impact Evaluation: The Basics

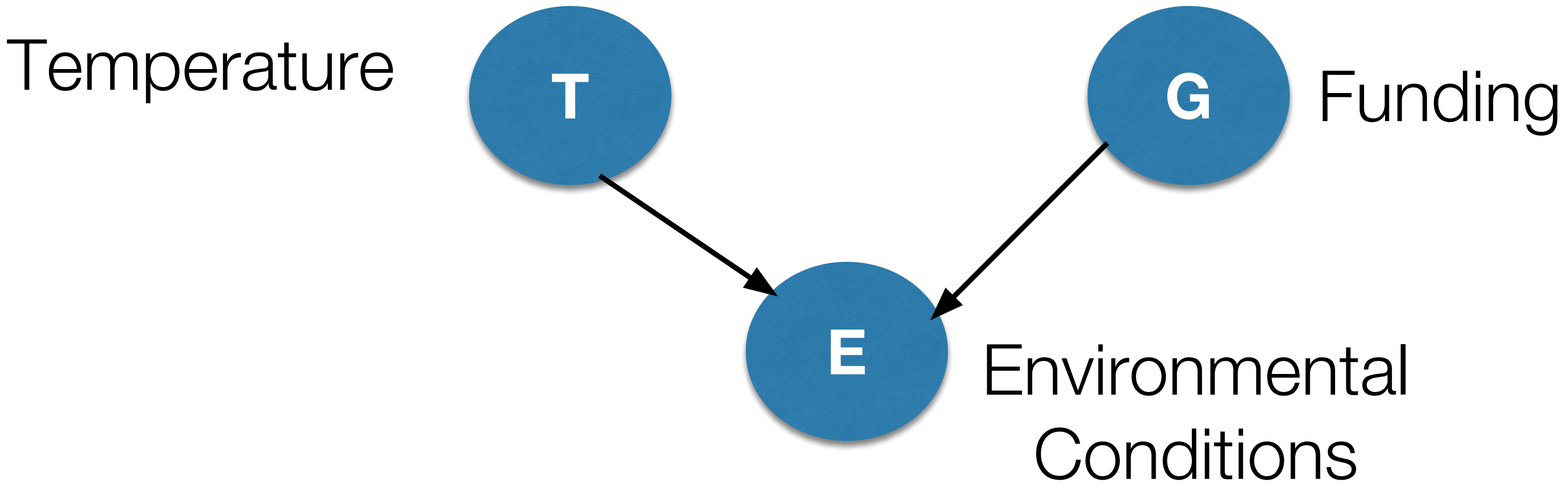


What other primary and secondary drivers of deforestation can you think of?

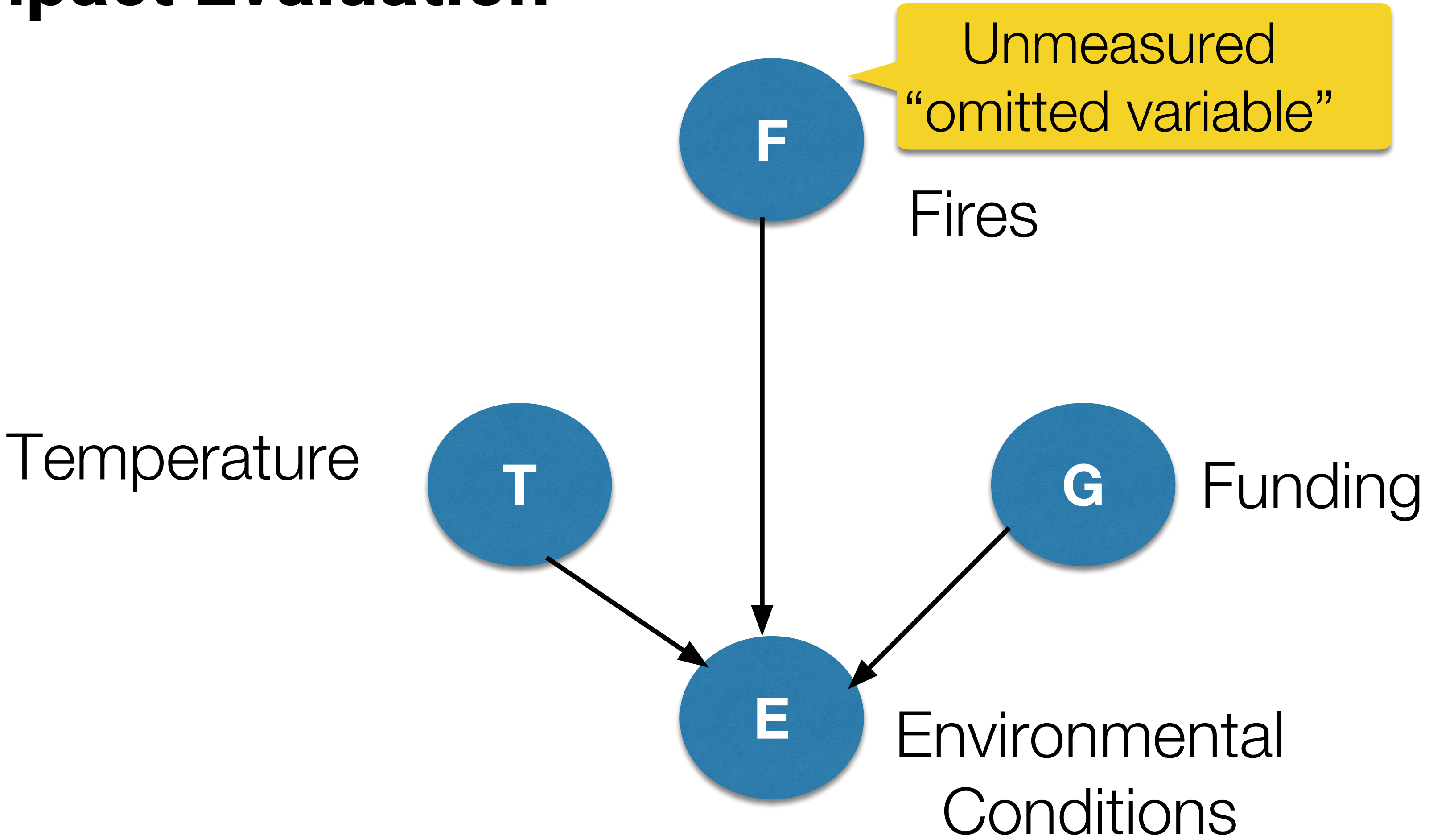
Can we measure them all?

Goal of impact evaluation is different from other types of analysis: we want to minimize the chance an unknown variable is the true cause of impact.

Impact Evaluation: The Basics



Impact Evaluation



Impact Evaluation

Two overarching designs to Impact Evaluation to overcome this challenge:

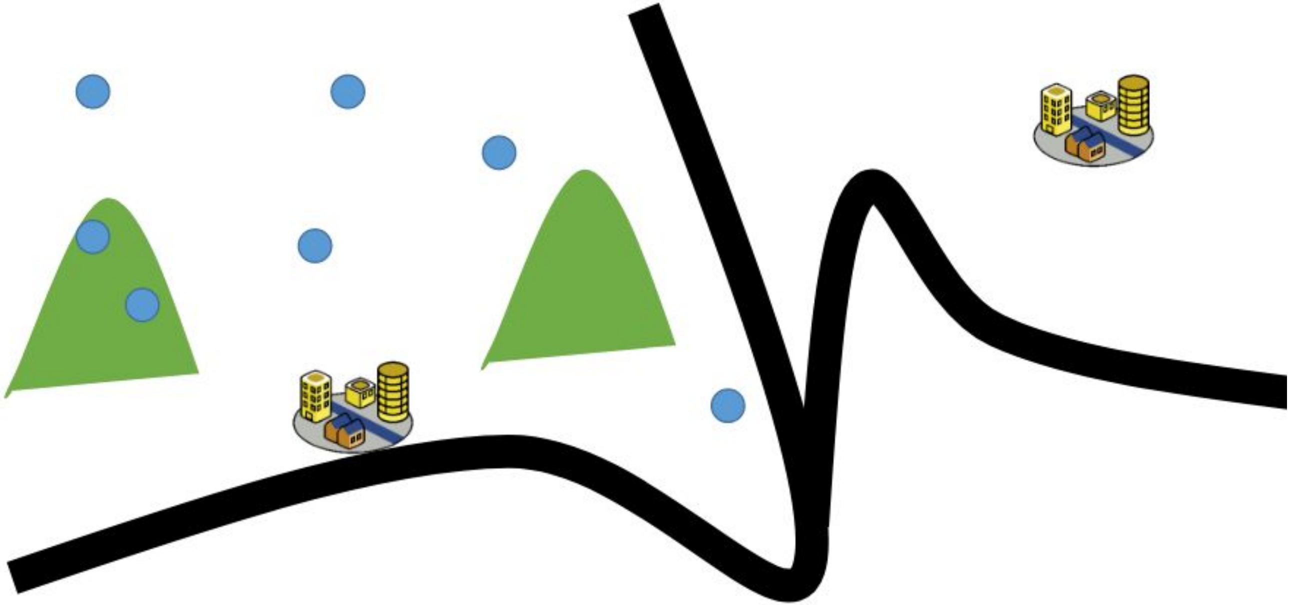
Difference - in - Difference
Compares each unit of observation to itself before an intervention occurred. A strong choice when you have time series data for all measured covariates both before and after the intervention.

Matching
Compares each unit of observation to a unit of observation that is as similar as possible ("twins"), but which had no intervention. A strong choice when you have a "wide" dataset with many attributes to match on, but limited temporal information.

Impact Evaluation : the GEF case

Matching Chosen for GEF Case Study

- Large amount of spatial data; limited temporal data for many covariates.
  - Distance to urban area
  - Distance to roads
  - GEF funding (only totals known, not exact distribution dates)



Impact Evaluation

Key decisions to be made for matching:

- From where in the study region should you draw controls?
  - How far away from your intervention do you believe you need to go before any impacts will be 0?
    - Avoid choosing controls in areas too close to interventions to be independent.
    - However, the closer the better, as places closer together are likely to be more similar.
  - Control cases can also be limited to - for example - country or district boundaries.

- What is the explicit outcome you will measure the impact on?
  - Deforestation in a 2-year period after the intervention? 5 year?
  - Difference between a pre-period baseline and post-period?

Impact Evaluation

Key decisions to be made for matching:

- Along which attributes should you match?
  - Do you match on temperature from every year, the average temperature from across multiple years, or other data manipulations?
  - Not all variables are equally important for matching - the goal is to ensure control and treatment cases are similar along dimensions most likely to be correlated with relevant omitted variables.
  - I.e., if you omit population but can match on Nighttime Lights, NTL is likely to mediate the lack of population data.
  - "Propensity Score Matching" helps you appropriately conduct this selection.
- How good of a match is "good enough"?
  - Variable "balance" across measured indicators.
  - Propensity score balance across measured indicators.
  - Do we look at match for individual countries, districts, or globally?

Impact Evaluation

Key decisions to be made for matching:

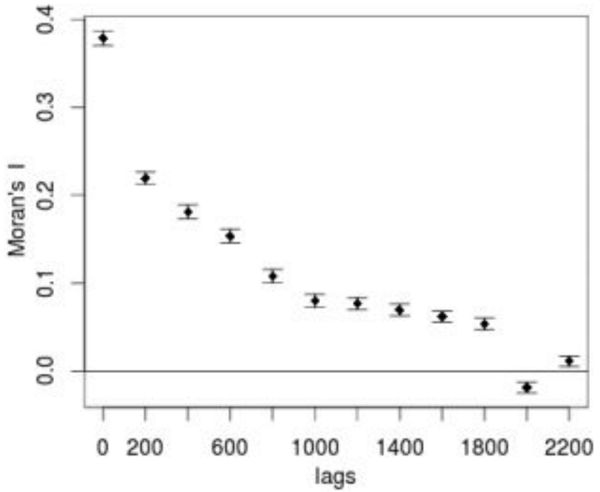
- How will you handle heterogeneity in your data?
  - If you believe there is only heterogeneity along a small number of known dimensions, interaction models can be a good choice.
  - If you are uncertain about the heterogeneity, or it can be nested (i.e., different across multiple countries), machine-learning based techniques can be best.

- How will you handle spatial imprecision in measurements?
  - Both intervention site data and satellite data can have spatial imprecision.
  - The number of dimensions of uncertainty you model increases the computational costs of your model exponentially. Choosing one or two dimensions is most manageable; expanding even to 3 can be challenging to compute, visualize, or interpret.
  - Examples might include: the size of the buffer used to represent spatial impact, the measured value of environmental quality from a satellite, the start date of a project, years used to calculate pre-trends, or any other decisions.
  - You can incorporate uncertainty as a source of error, or select one model and use uncertainty as a robustness test.

Impact Evaluation: the GEF case

Key decisions to be made for matching:

- From where in the study region should you draw controls?
    - How far away from your intervention do you believe you need to go before any impacts will be 0?
    - Global, data-driven estimate of 100km, based on spatial autocorrelation.



Impact Evaluation: the GEF case

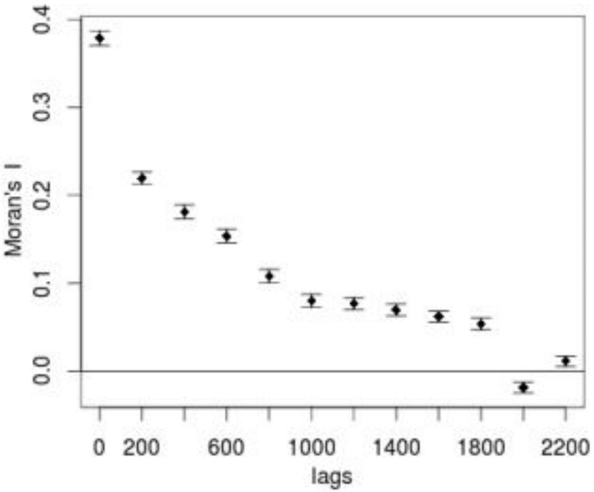
Defining Spatial Autocorrelation

- Things closer together tend to be more similar, but patterns are not the same for all regions, processes or patterns.

- “Variogram” and “Covariogram” techniques let you measure the distance at which - spatially - things are no longer correlated with one another.

- The optimal range for matching is at locations that retain correlation in *covariate* data (i.e., they have similar population), but have no spatial relationship in your *outcome* data (i.e., different levels of forest cover).

- For the GEF, this was in a concentric circle between 50 and 150 kilometers away from each GEF project, which was then limited by country boundaries.



Impact Evaluation

Key decisions to be made for matching:

- What is the explicit outcome you will measure the impact on?
    - Deforestation in a 2-year period after the intervention? 5 year?
    - Difference between a pre-period baseline and post-period?

Impact Evaluation

In the GEF case, outcome measures were determined based on available data and the types of information that would be most helpful for decision making:

1. **Outcome 1: Change in Environment Quality,** measured by NDVI (more on this metric later).
    a. “Pre-intervention period” defined as a period of no less than 5 years before the GEF intervention occurred. The maximum NDVI values within each year were averaged to give the peak-greenness estimate.
    b. “Post-intervention period” defined as the period between when the project formally started (“Actual date of implementation start” where available) until the year 2015 (the final year of satellite data available during the analysis).
    c. The outcome of interest was then defined as the post-intervention period minus the pre-intervention period. Positive values indicated a positive change in NDVI, and vice-versa.
    d. Because some projects started earlier in the time record (i.e., 2006), and others later (i.e., 2010), this meant the outcome post-period was variable. Taking advantage of this time difference enabled tests of implementation strength based on duration of implementation / post-implementation period.

Impact Evaluation

In the GEF case, outcome measures were determined based on available data and the types of information that would be most helpful for decision making:

1. **Two additional outcomes were also tested.** These included forest cover change and forest fragmentation.
   a. Enabled a triangulation of project impacts based on measurements from independent sources.
   b. Enabled a broader understanding of impact with a relatively low added cost.
      i. By leveraging open source data created by expert practitioners, both forest cover change and NDVI were available at low human resource costs (though high computational costs).
      ii. Forest fragmentation was derived in collaboration with an expert practitioner at NASA. Because fragmentation is area-dependent, this was one of the most expensive satellite data production steps, and required an exact definition of the unit of observation before it could be undertaken.

Impact Evaluation - GEF Case Recap

Interlude - key decisions made so far:

- Causal strategy: **Matching**, as data is primarily spatial with limits on temporal coverage.
- Unit of Observation: 25km **Buffer,** as at the portfolio scale exact information on spatial extent of each project is unknown.
- From where do we draw controls: In **concentric regions** no less than 50km and no more than 150km away from each GEF intervention. This is a region which is similar along our covariate dimensions, but dissimilar along the outcome dimensions.
- How do we define our outcome(s): Change in NDVI, Forest Cover, and Forest Fragmentation between a pre-intervention period and a post-intervention period, where the post intervention period can be of variable length. This enables the use of time-variation in post-intervention periods to better understand how long it takes projects to show impacts.

Impact Evaluation - GEF Case Recap

Interlude - what you have at this point:

- Once the first four decisions (causal strategy, unit of observation, control group selection, and outcome definition) are made, data integration across the required data points is conducted.

- This provides a “spreadsheet” in which every row is an intervention location, and every column is relevant satellite or other (i.e., funding amount) information about the project.

- The next goal is to match every “treated” row in the spreadsheet (locations at which GEF interventions occurred) to every “untreated” row in the spreadsheet (locations at which no known interventions occurred).

Impact Evaluation

Key decisions to be made for matching:

- Along which attributes should you match?
    - Do you match on temperature from every year, the average temperature from across multiple years, or other data manipulations?
    - Not all variables are equally important for matching - the goal is to ensure control and treatment cases are similar along dimensions most likely to be correlated with relevant omitted variables.
    - I.e., if you omit population but can match on Nighttime Lights, NTL is likely to mediate the lack of population data.
    - “Propensity Score Matching” helps you appropriately conduct this selection.

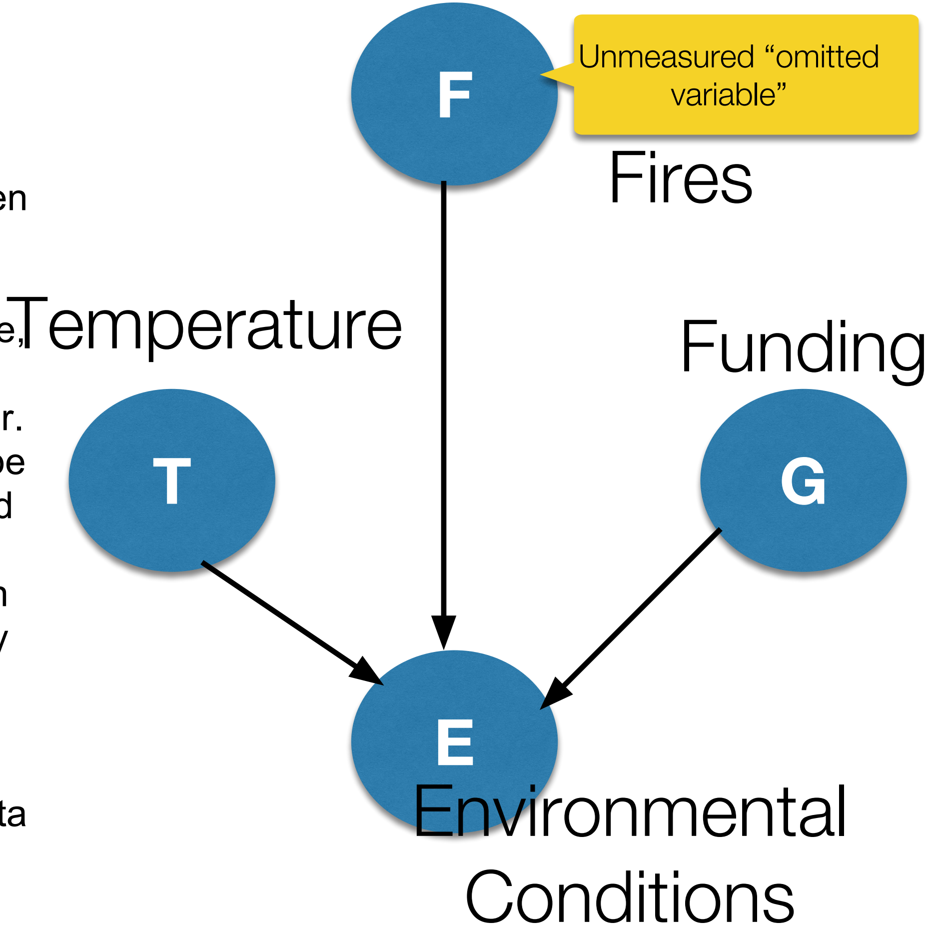
Impact Evaluation: the GEF case

- **Variables chosen for matching were chosen based on a combination of theory and data availability.**

Statistic	Mean	Min	Median	Max
Distance to Commercial River (km)	915.3	1.2	2.3	16,000
Distance to Roads (km)	36.1	0.21	2.89	994
Elevation (meters)	597.762	1.671	319.482	5,009.923
Slope (degrees)	3.278	0.000	1.974	19.173
Urban Accessibility (Relative)	622.302	31.078	260.602	4,644.295
Population Density (2005)	184.866	0.000	75.209	4,179.138
NDVI (1982)	.1756	.0329	.1778	.3454
NDVI (2014)	.1844	.0286	.1852	.3982
Nighttime Lights (2013)	1.651	0.000	0.372	32.422
Minimum Air Temp (2014)	17.042	-20.150	22.100	28.000
Maximum Air Temp (2014)	27.268	11.975	28.000	36.433
Mean Air temp (2014)	22.453	-1.371	24.723	30.029
Max Precip (2014)	277.145	17.700	217.275	1,470.650
Min Precip (2014)	10.635	0.000	0.725	157.350
Mean Precip (2014)	95.283	2.310	72.527	439.117
Protected Area Overlap	3.546	1	4	6
Treecover - 2000 (Percent)	17.596	0.000	6.772	98.076

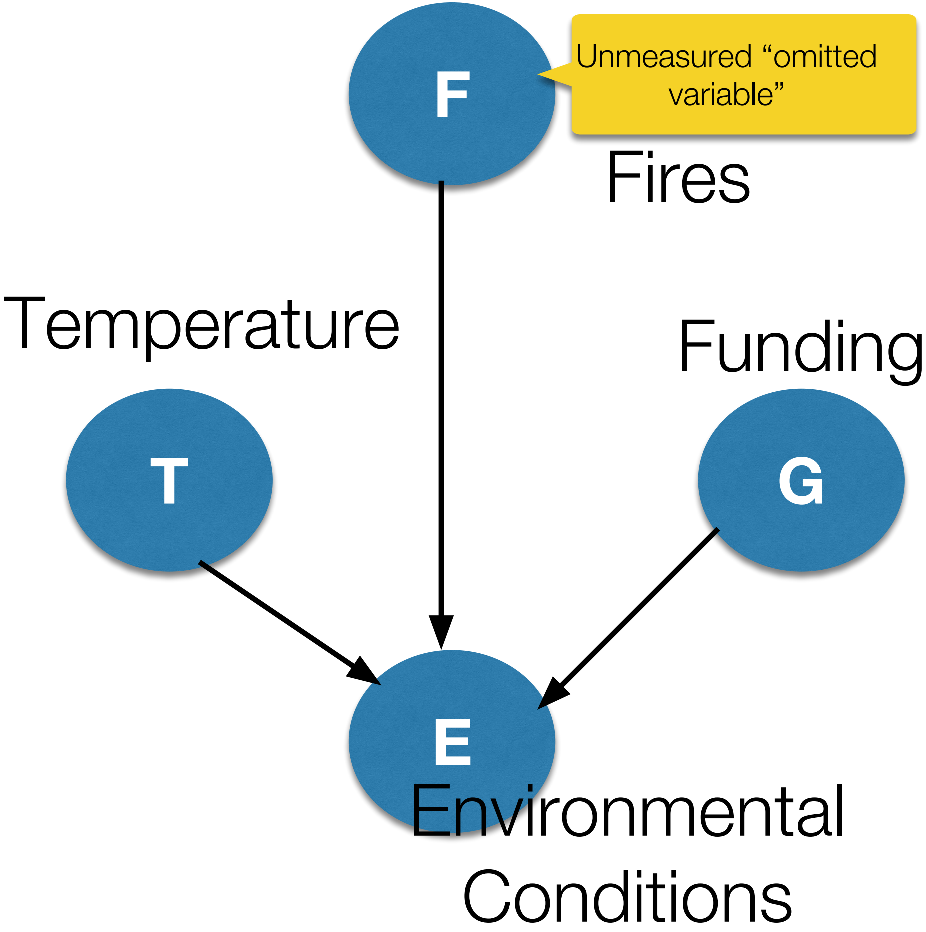
Impact Evaluation: the GEF case

- **Not all variables we might have wanted were included.**
  - One of the biggest advantages of impact evaluation methods is that even if you cannot directly measure every relevant attribute, by matching "pairs" based on everything you *can* measure, one can assume that other unmeasured elements are also similar.
  - The strength of this assumption can be challenged, but can also be measured statistically.
- For example, we did not include information on fires in this analysis, but it is very heavily correlated with nighttime lights (as they can include fires themselves).
- Thus, by matching on nighttime lights we mitigate the risk due to not including fire data (though more data is generally better!).



Impact Evaluation: the GEF case

- The risk of omitted variables is further mediated by the use of propensity score matching.
- Fundamental concept: if every GEF project was completely randomly placed spatially, you could simply compare treated and untreated sites, subtract, and have a measure of the strength of intervention.
    - Propensity score matching "weights" matches along non-random dimensions, ensuring that you are more careful to match (for example) rural to rural sites if the GEF tended to non-randomly target rural areas.
    - Propensity scores are literally the "propensity to receive treatment" for both treated and control locations (irrespective of if they did or not).



Impact Evaluation: the GEF case

Propensity Model Results (excluding year FE and constant)
(*p<0.1; **p<0.05; ***p<0.01)

Baseline Average NDVI	0.048
Baseline Maximum NDVI	0.0004
Baseline Minimum Temp.	1.085**
Baseline Maximum Temp.	1.120**
Baseline Average Temp.	-2.146**
Baseline Maximum Precip.	-0.002
Baseline Minimum Precip.	0.001
Baseline Average Precip.	0.018
Distance to Rivers	0.00002
Distance to Roads	0.00000
Elevation	0.001
Slope	0.048
Urban Accessibility	-0.003
Population Density (2000)	0.002
Protected Area Overlap	0.218
Baseline Treecover (2000)	-0.007
Latitude	-0.009
Longitude	-0.009*

- Propensity models are only descriptive, but provide some insight into project allocation strategies.
- Interpretation can be challenging if variables are interrelated (i.e., many of the temperature variables in this table are interrelated with one another).

Impact Evaluation: the GEF case

Propensity Model Results (excluding year FE and constant)
(*p<0.1; **p<0.05; ***p<0.01)

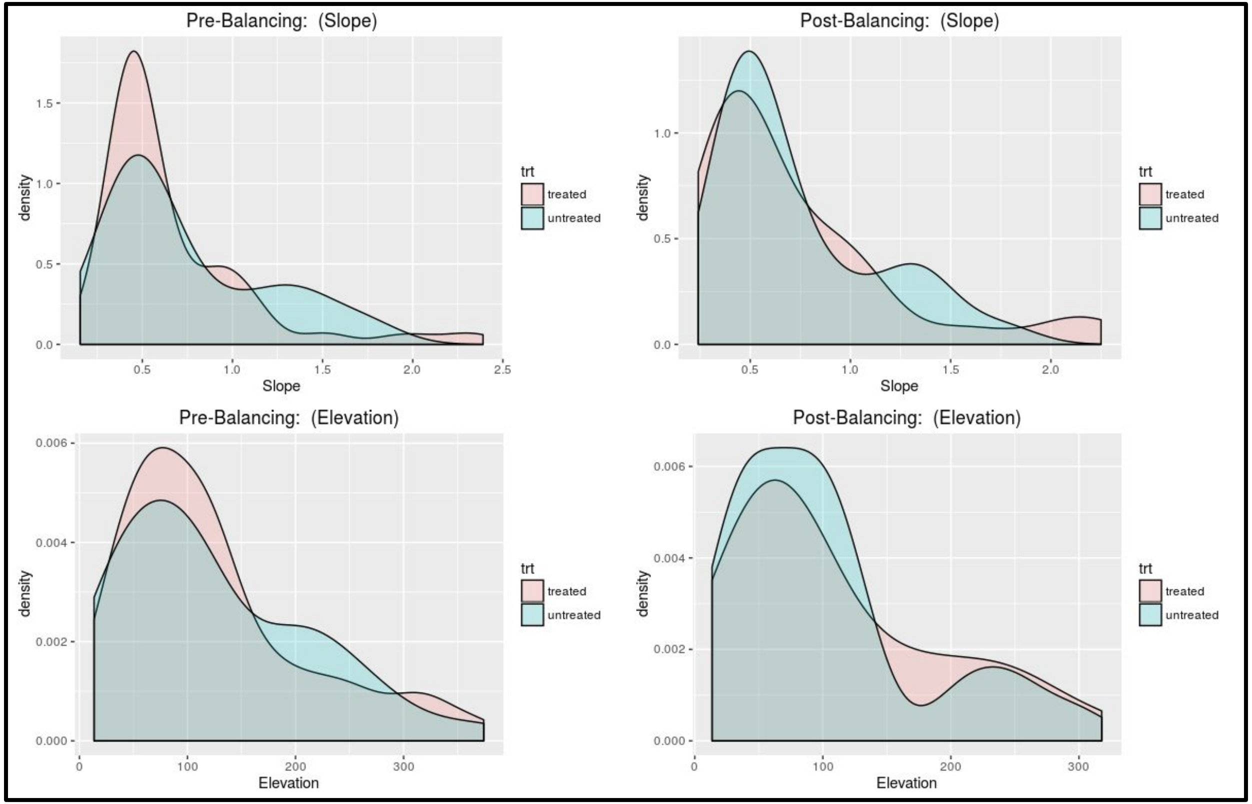
Baseline Average NDVI	0.048
Baseline Maximum NDVI	0.0004
Baseline Minimum Temp.	1.085**
Baseline Maximum Temp.	1.120**
Baseline Average Temp.	-2.146**
Baseline Maximum Precip.	-0.002
Baseline Minimum Precip.	0.001
Baseline Average Precip.	0.018
Distance to Rivers	0.00002
Distance to Roads	0.00000
Elevation	0.001
Slope	0.048
Urban Accessibility	-0.003
Population Density (2000)	0.002
Protected Area Overlap	0.218
Baseline Treecover (2000)	-0.007
Latitude	-0.009
Longitude	-0.009*

- Propensity models are only descriptive, but provide some insight into project allocation strategies.
- Interpretation can be challenging if variables are interrelated (i.e., many of the temperature variables in this table are interrelated with one another).
- However, some clear trends can be observed, i.e.:
    - GEF LD projects tend to be sited in areas with steeper slopes.
    - GEF LD projects tend to be sited in WDPA protected areas, relative to others.

Impact Evaluation - the GEF case

Matching Balance

- Matching balance is traditionally assessed through two means
  - Statistically, if you have matched well along dimensions that interventions were not randomly allocated along.
    - An example of this might be if you have evidence the GEF targeted rural areas for interventions, you want to ensure you match rural to rural and urban to urban areas.
    - This is measured by the difference in propensity scores after matching. In a perfect case, matched pairs have the same propensity.
  - Theoretically, if you believe matching along a certain dimension is important to mediate omitted variables, you want to ensure the matching captured that dimension.
    - An example of this is when you do not have measurements of fire activity, but are using nighttime lights as a proxy. You want to ensure nighttime lights balance is strong in this case.
    - This is measured by contrasting the pre-matching distribution to the post-matching distribution along relevant dimensions.



Pre-Balancing: (Slope)
Post-Balancing: (Slope)
Pre-Balancing: (Elevation)
Post-Balancing: (Elevation)
density
Slope
Elevation
trt
treated
untreated

Impact Evaluation

Modeling Heterogeneity

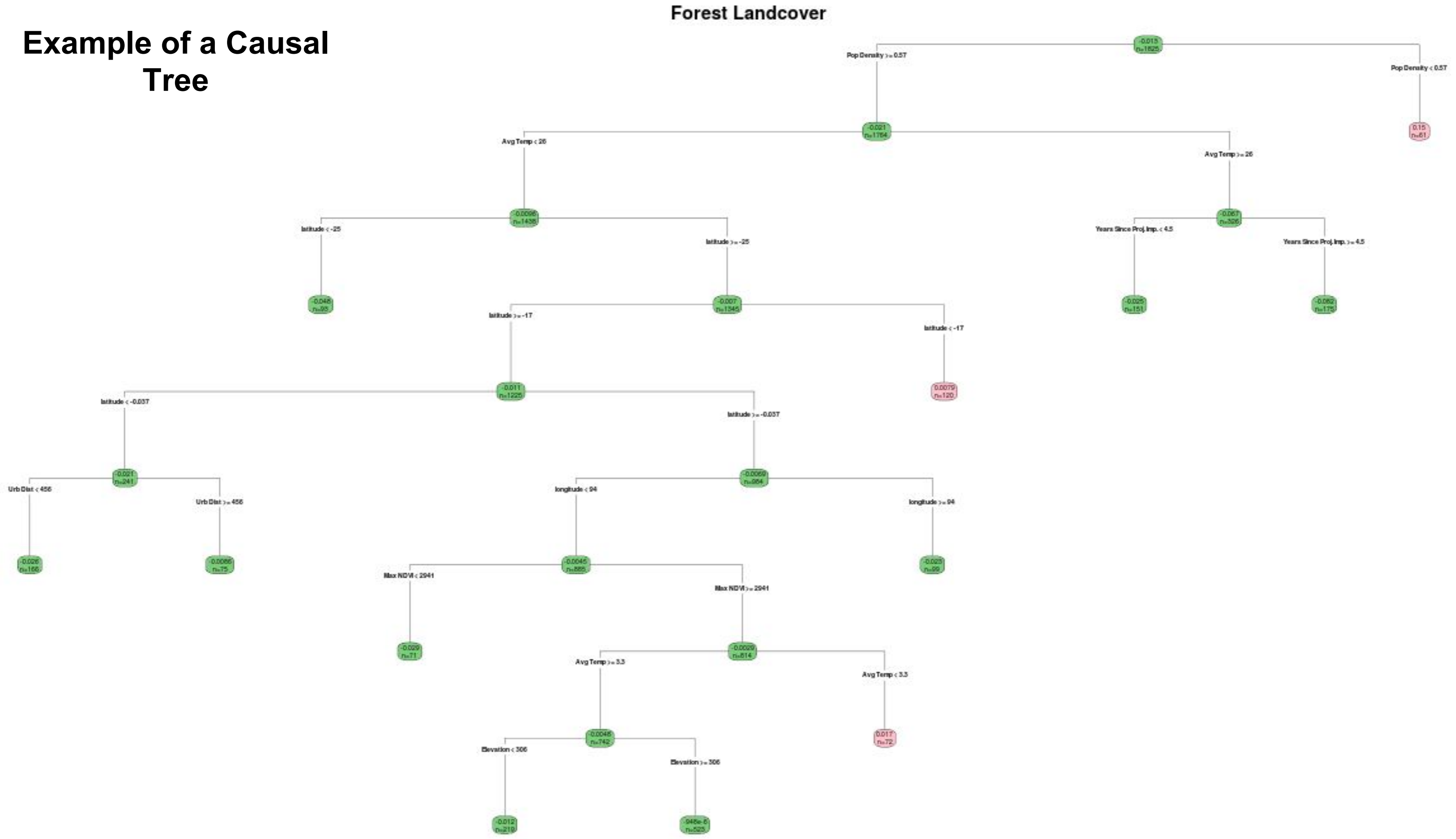
- If you have reason to believe that your intervention would not equally impact all areas of your study, it is important to model heterogeneity.
- This can impact all studies, within a country (i.e., interventions close to the current hot spots of deforestation may have different outcomes than those in the deeper forest), or across multiple countries in programmatic or regional activities (i.e., unmeasurable socio-cultural or legal frameworks may result in systematically different outcomes).
- In cases where the dimensions of heterogeneity are known (i.e., the above rainforest example), “interaction” models are a good choice.
- In cases where the dimensions of heterogeneity are unknown (i.e., a regional-scale analysis), “causal tree” models are a good choice.

Impact Evaluation - the GEF case

Modeling Heterogeneity

- In the case of the GEF, a causal tree model was chosen to model results.
  - This choice was based on the scope of Land Degradation activities, which span multiple countries and continents.
  - It is impossible to specify (or measure) all potential dimensions of heterogeneity in this case.
  - Instead, treatment impact is allowed to vary along all dimensions, and the dimensions along which the biggest differences in impact are detected are reported.
  - Further, latitude and longitude are incorporated into the model to enable the detection of heterogeneity within spatial units (i.e., countries) that may have different socio-cultural or economic contexts, without specifying or measuring such factors directly.

Example of a Causal Tree



Impact Evaluation

- How will you handle spatial imprecision in measurements?
    - While measurement imprecision is not unique to the case of spatial data, the dimensions along which imprecision can manifest are unique - for example, the malleability of the unit of observation.
- There are two classes of approaches to coping with this imprecision. While these are not mutually exclusive, they can be difficult to implement and interpret alongside one another.
    - *Spatial Imprecision as a Source of Error*
        - In this approach, spatial imprecision is explicitly introduced in a model by reducing the chance that "significant" statistical results will be found if imprecision is high.
        - Generally applied outside of machine learning models where spatial imprecision is the primary source of error being considered.
    - *Spatial Imprecision as a Test of Model Robustness*
        - In this approach, spatial imprecision is used to re-fit a model hundreds of times, each time making a different assumption about the "true" value.
        - Generally applied alongside machine learning models, or where there are multiple uncertainties that need to be tested.

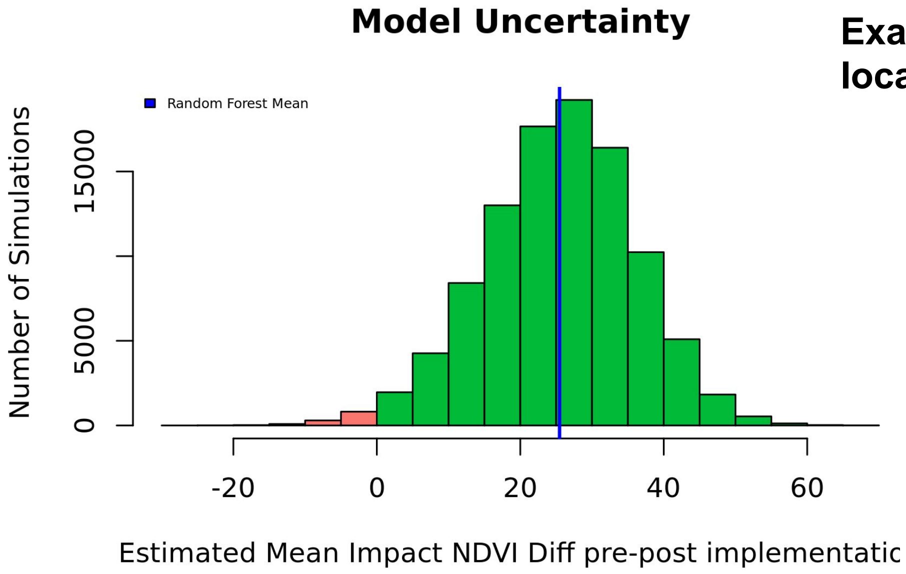
Impact Evaluation - the GEF case

- In the GEF case, we chose to incorporate spatial imprecision as a test of model robustness.
    - This decision was based on:
        - The modeling paradigm (Causal Trees)
        - The dimensions of uncertainty being tested (buffer size, ancillary variable selection, and unit of observation selection were all tested).
        - Policy needs - it was important to know if the results we found would change if these factors changed.
- Model robustness is tested following many approaches; here we use a Monte Carlo approach in which each factor is randomly permuted and a model re-fit every time.

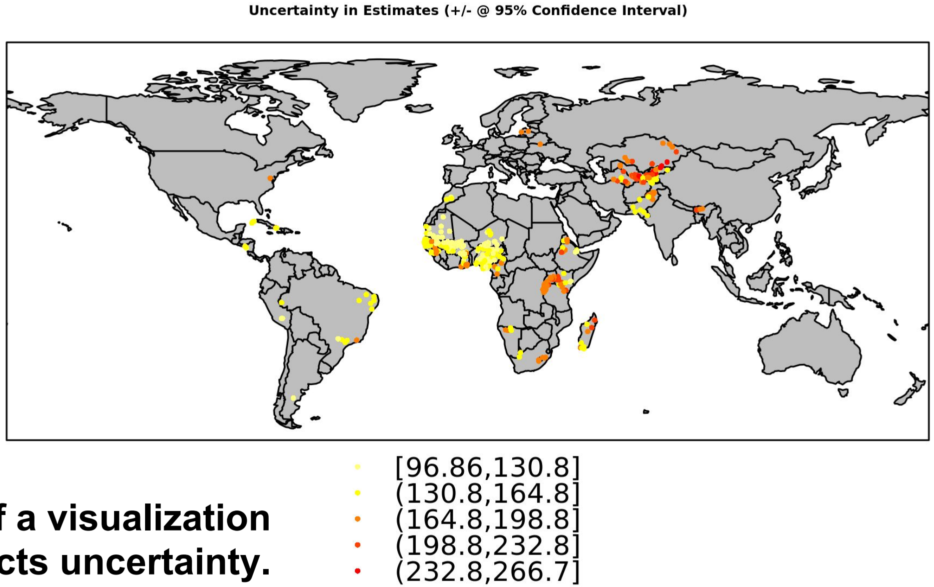
Impact Evaluation - the GEF case

- In the GEF case, we chose to incorporate spatial imprecision as a test of model robustness.
  - This decision was based on:
    - The modeling paradigm (Causal Trees)
    - The dimensions of uncertainty being tested (buffer size, ancillary variable selection, and unit of observation selection were all tested).
    - Policy needs - it was important to know if the results we found would change if these factors changed.
- Model robustness is tested following many approaches; here we use a Monte Carlo approach in which each factor is randomly permuted and a model re-fit every time.
- As a secondary robustness check, we contrast our Causal Tree global model findings to an “interaction” model to identify if model selection had an impact on our findings.

Example of a single project locations uncertainty.



Example of a visualization of all projects uncertainty.



Impact Evaluation

An econometric "interaction" model was fit to see if different modeling strategies resulted in similar findings.

In this case, the same *global* directional impact was found, and with significance.

Interaction models are called as such due to the use of "interaction terms" to capture heterogeneity. These terms must be specified by the modeler. In this case, interaction terms were specified according to the interactions identified by the Causal Tree machine learning technique.

Only the treatment variable can be causally interpreted in this (or the causal tree) approach. This is because it is the only variable that defines the groups across which we seek to match (our "twins").

Matched Model: SFA Land (Treated), Null Case Comparisons (Control)

	Dependent variable:
	NDVI Diff pre-post implementation
treatment	0.08^{***} (0.03, 0.14)
Dist. to Rivers (m)	-0.04 (-0.14, 0.07)
Dist. to Roads (m)	0.06^{*} (-0.01, 0.12)
Elevation (m)	-0.18^{***} (-0.31, -0.06)
Slope (degrees)	-0.11^{**} (-0.21, -0.02)
Urb. Dist. (rel)	-0.01 (-0.08, 0.07)
Pop. Density (2000)	0.06 (-0.04, 0.17)
Protected Area %	0.09^{***} (0.03, 0.14)
Treecover (2000, %)	0.05 (-0.04, 0.13)
Latitude	-0.09^{*} (-0.18, 0.003)
Longitude	-0.13^{***} (-0.22, -0.03)
Max Precip. (2002, mm)	-0.42^{***} (-0.58, -0.27)
Min Precip (2002, mm)	-0.08^{*} (-0.17, 0.01)
Mean Precip (2002, mm)	0.27^{***} (0.08, 0.45)
Max Temp (2002, C)	0.004 (-0.33, 0.34)
Min Temp (2002, C)	-0.28 (-0.78, 0.22)
Mean Temp (2002, C)	-0.23 (-0.98, 0.52)
Nightime Lights (2002, Relative)	-0.02 (-0.10, 0.06)
NDVI (2002, Unitless)	0.01 (-0.07, 0.10)
Urb. Dist. (rel) *Treatment	-0.004 (-0.08, 0.07)
Dist. to Rivers (m) *Treatment	-0.04 (-0.14, 0.07)
Dist. to Roads (m) *Treatment	-0.03 (-0.10, 0.04)
Pop. Density (2000) *Treatment	-0.06 (-0.17, 0.04)
Latitude *Treatment	0.03 (-0.06, 0.12)
Longitude *Treatment	0.08 (-0.02, 0.17)
NDVI (2002, Unitless) *Treatment	0.07^{*} (-0.01, 0.15)
Elevation (m) *Treatment	0.25^{***} (0.12, 0.37)
Slope (degrees) *Treatment	-0.12^{**} (-0.22, -0.02)
Treecover (2000, %) *Treatment	-0.03 (-0.11, 0.06)
Max Temp (2002, C) *Treatment	0.57^{***} (0.24, 0.90)
Mean Temp (2002, C) *Treatment	-1.05^{***} (-1.80, -0.31)
Min Temp (2002, C) *Treatment	0.80^{***} (0.30, 1.30)
Max Precip. (2002, mm) *Treatment	-0.06 (-0.21, 0.10)
Mean Precip (2002, mm) *Treatment	0.06 (-0.12, 0.25)
Min Precip (2002, mm) *Treatment	-0.12^{***} (-0.20, -0.03)
Nightime Lights (2002, Relative) *Treatment	0.01 (-0.06, 0.09)
Protected Area % *Treatment	-0.02 (-0.07, 0.04)
Constant	-0.01 (-0.06, 0.05)
Observations	966
R^2	0.30
Adjusted R^2	0.27
Note:	$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Impact Evaluation - GEF Case Recap

Recap - key decisions made:

- **Causal strategy**: Matching, as data is primarily spatial with limits on temporal coverage.
- **Unit of Observation**: 25km Buffer**,** as at the portfolio scale exact information on spatial extent of each project is unknown.
- **From where do we draw controls**: In concentric regions no less than 50km and no more than 150km away from each GEF intervention. This is a region which is similar along our covariate dimensions, but dissimilar along the outcome dimensions.
- **How do we define our outcome(s)**: Change in NDVI, Forest Cover, and Forest Fragmentation between a pre-intervention period and a post-intervention period, where the post intervention period can be of variable length. This enables the use of time-variation in post-intervention periods to better understand how long it takes projects to show impacts.
- **Along what dimensions do we match**: protected areas, temperature, precipitation, nighttime lights, and any others we theoretically believe may be correlated with omitted variables.
- How do we match and test balance: a propensity-score matching approach alongside a balance test is used.
- **How do we model**: a Causal Tree is used due to the anticipated heterogeneity in the analysis.
- **How do we model imprecision**: a Robustness Check is conducted with monte carlo analysis, to account for the range of imprecision incorporated into the analysis.

Geospatial Impact Evaluation Methods, Tools & Applications

Independent Evaluation Office
GLOBAL ENVIRONMENT FACILITY

Day 1, Part 3: Contrasting the GEF LD study to other implementations.

Case Study 1:
Indigenous Land Rights in the Amazon

Ariel BenYishay[1],Silke Heuser[2], Rachel Trichler[1], Dan Runfola[1]

[1] AidData, William and Mary
[2] KFW (German Development Bank)

	GEF LD	KFW Amazon
Causal Strategy	Matching	
Unit of Observation	25km Buffer	
Treatment Definition	Location of GEF Intervention	
Controls	Random 25km Buffers of non-intervention, between 50 and 150km from GEF intervention; within the same country.	
Outcome Variables	Change in pre- to post-period NDVI, Forest Cover, Forest Fragmentation	
Matching Dimensions	Protected Areas, Temperature Precipitation, variety of GEF-specific project characteristics	
Balance Test	Propensity Score and Distribution Diagnostics	
Model	Causal Tree, Matched interaction model to test robustness.	
Imprecision	Monte-carlo Robustness Analysis	

Objective

Does demarcating indigenous lands reduce deforestation?

- Land tenure security not widely shown to reduce deforestation
- Indigenous control / stewardship shown in several recent studies to be associated with lower deforestation rates (Nelson et al 2001, Nelson and Chomitz 2012, Nolte et al 2013, Vergara-Aseno and Potvin 2014)
- Given low rates of deforestation observed on indigenous lands, is demarcation likely to influence deforestation?

Project Details

- In 1988 constitution, Gov of Brazil committed to demarcating indigenous people’s territories
- Between 1995-2008, with funding and tech support from KfW and the World Bank, the PPTAL project identified, recognized, and studied 181 community lands.
- By 2008, 106 community lands demarcated, covering 38 million hectares (~35% of all indigenous lands in Amazon)

Project Details

- ***Demarcation***: recognition by the Min of Justice
- Followed by regularization (entry into municipal, state and federal registries)
- Varied by community between 1995 and 2008
    - Median year is 2001
- Support for Boundary Enforcement

Data

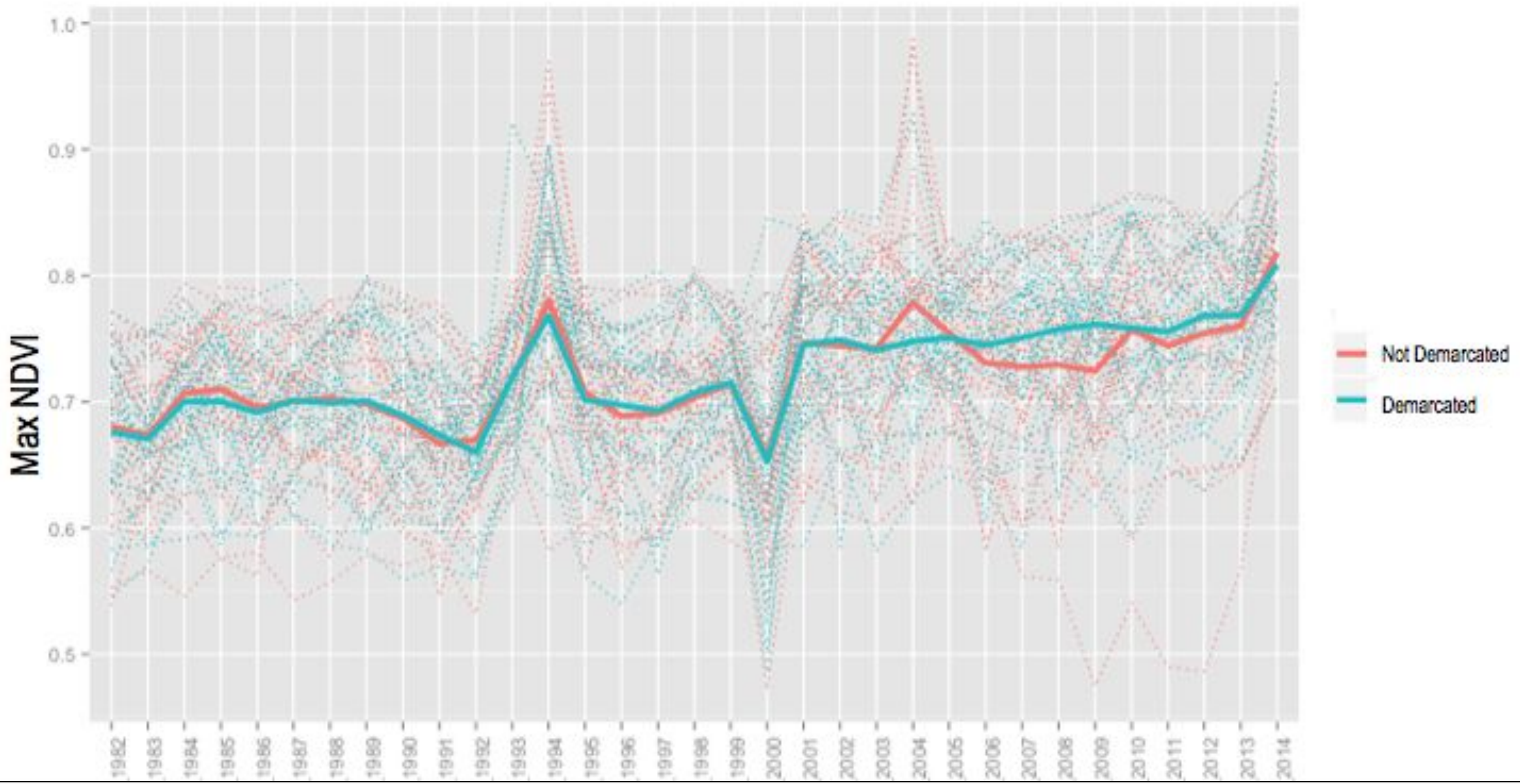
- Treatment status
  - Boundaries of community lands
  - Administrative data on demarcation dates
- Merged with satellite-based greenness measure
  - NASA Land Long Term Data Record (LTDR), 1982-2010
  - Processed to Normalized Difference Vegetation Index (NDVI)
  - Range is [0, 1] (0 = rocky, barren; 1 = dense forest)
  - Annual NDVI max and mean measures
- Covariates
  - Climate (precip., temp.); topology (elevation, slope); distance to rivers; gridded, interpolated population

Methods

- Propensity Score Matching
  - Differences over time across matched treated/comparison communities
  - Match on baseline levels, pre-trends, & covariates
  - Demarcated vs. not; “Early” (‘95-’01) vs “Late” (‘01-’08)

- Fixed effects
  - Control for time-invariant community unobservables
  - Treatment status at finer time intervals

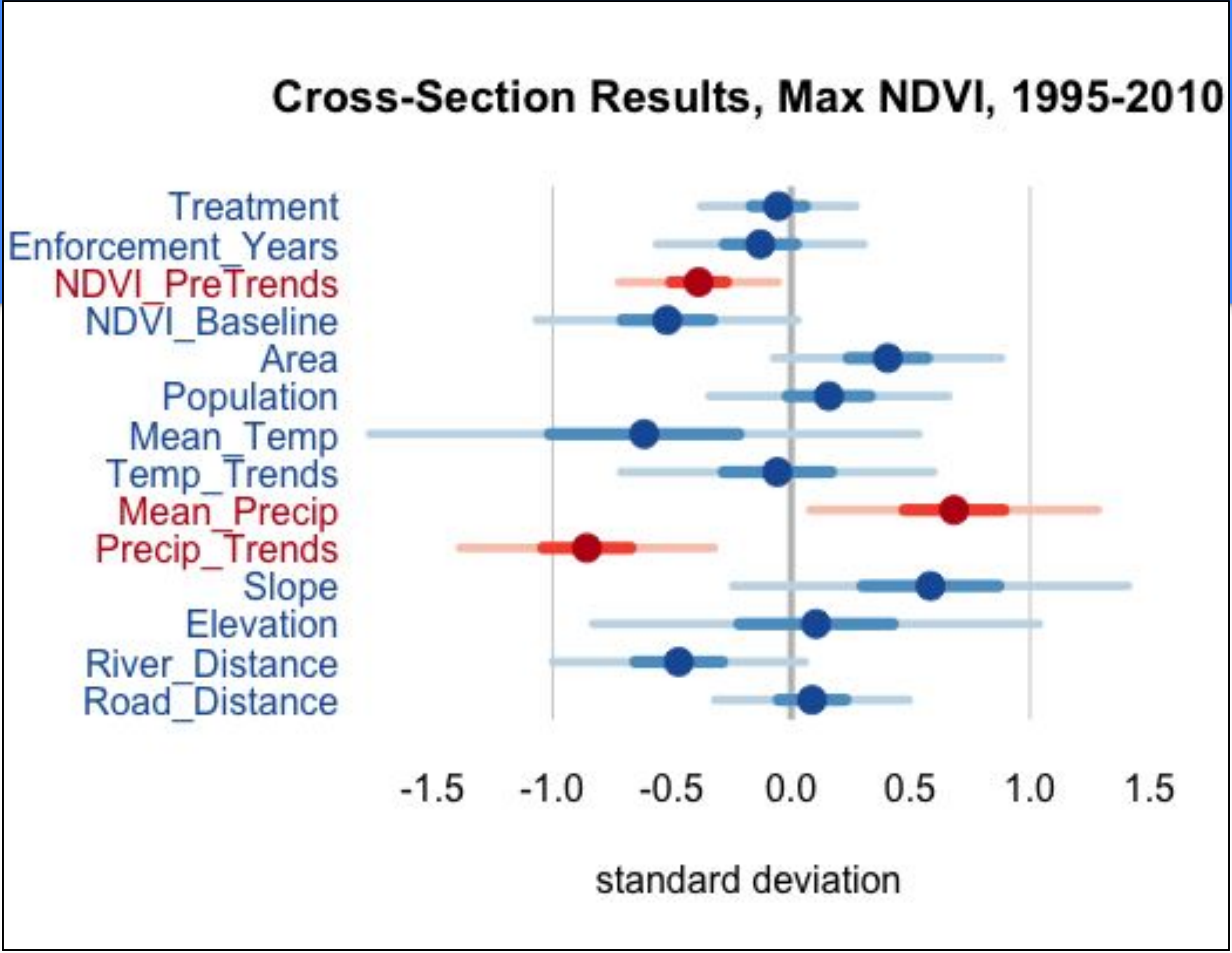

1997
2008



Max NDVI
1.0
0.9
0.8
0.7
0.6
0.5
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Not Demarcated
Demarcated

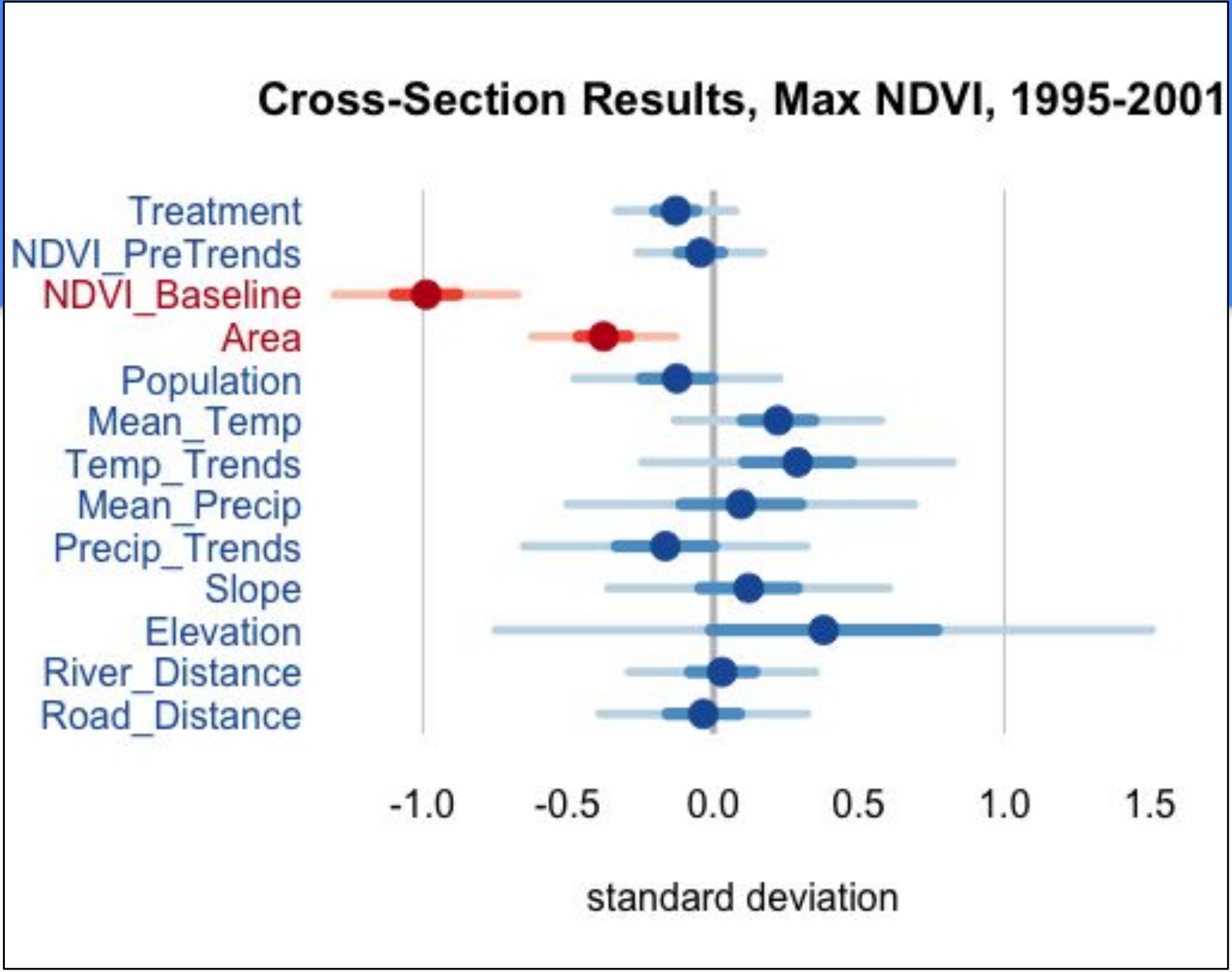
Demarcated vs. non-demarcated

- **Treatment** = Demarcated between '95 and '08
- **Outcome** = Change in mean NDVI between '95 and '10
- Control for years in demarcation status; matched pairs of demarcated and non-demarcated (n=60).



Early vs. Late demarcation

- **Treatment** = Demarcated between ’95 and ’01
- **Outcome** = Change in mean NDVI between ‘95 and ’10
- Control for years in demarcation status; matched pairs of demarcated lands (n=80).



Take-aways

- No clear, robust evidence of differences in deforestation attributable to the PPTAL project
- Much lower rates of deforestation on indigenous lands in cross-section may not be related to land tenure status of these lands (or may be mediated through multiple, complex channels)

	GEF LD	KFW Amazon
Causal Strategy	Matching	
Unit of Observation	25km Buffer	
Treatment Definition	Location of GEF Intervention	
Controls	Random 25km Buffers of non-intervention, between 50 and 150km from GEF intervention; within the same country.	
Outcome Variables	Change in pre- to post-period NDVI, Forest Cover, Forest Fragmentation	
Matching Dimensions	Protected Areas, Temperature Precipitation, variety of GEF-specific project characteristics	
Balance Test	Propensity Score and Distribution Diagnostics	
Model	Causal Tree, Matched interaction model to test robustness.	
Imprecision	Monte-carlo Robustness Analysis	

	GEF LD	KFW Amazon
Causal Strategy	Matching	Matching, Fixed Effects
Unit of Observation	25km Buffer	Known polygons of intervention
Treatment Definition	Location of GEF Intervention	(1) If a polygon was ever demarcated (2) Early vs. late demarcation
Controls	Random 25km Buffers of non-intervention, between 50 and 150km from GEF intervention; within the same country.	Known polygons of non-intervention, with no spatial restrictions.
Outcome Variables	Change in pre- to post-period NDVI, Forest Cover, Forest Fragmentation	Change in NDVI
Matching Dimensions	Protected Areas, Temperature Precipitation, variety of GEF-specific project characteristics	Temperature, Precipitation, Distance to Infrastructure and Accessibility Metrics, Size
Balance Test	Propensity Score and Distribution Diagnostics	Propensity Score and Distribution Diagnostics
Model	Causal Tree, Matched interaction model to test robustness.	Matched interaction model with fixed effects
Imprecision	Monte-carlo Robustness Analysis	Omitted

Case Study 2: World Bank Environmental Safeguards

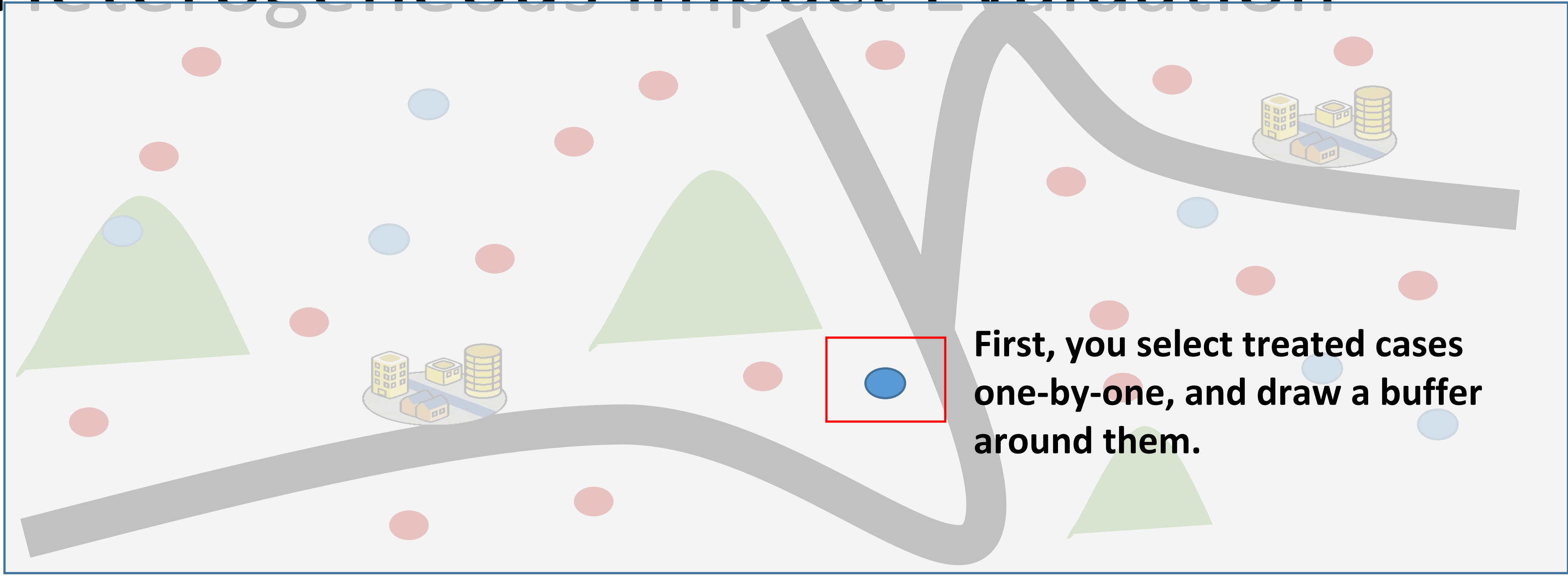
Daniel Runfola (1), Ariel BenYishay (1), Jeffery Tanner (2), Graeme Buchanan (3), Jyoteshwar Nagol (4), Matthias Leu (5), Seth Goodman (1), Rachel Trichler (1) and Robert Marty (1)

1 AidData, William and Mary,
2 Independent Evaluation Group, The World Bank
3 Center for Conservation Science, Royal Society for the Protection of Birds
4 Global Land Cover Facility, University of Maryland,
5 Department of Biology, William and Mary

Establishing Value : Heterogeneous Impacts

- In order to valuate individual project locations, the impact *of each project location* must be estimated.
- Traditional Geospatial Impact Evaluation (and IE writ large) provide a *single, overall* estimate of impact.
- Geo-VFM allows for estimating geographically heterogeneous impacts by employing Geographically Weighted Regression (GWR).

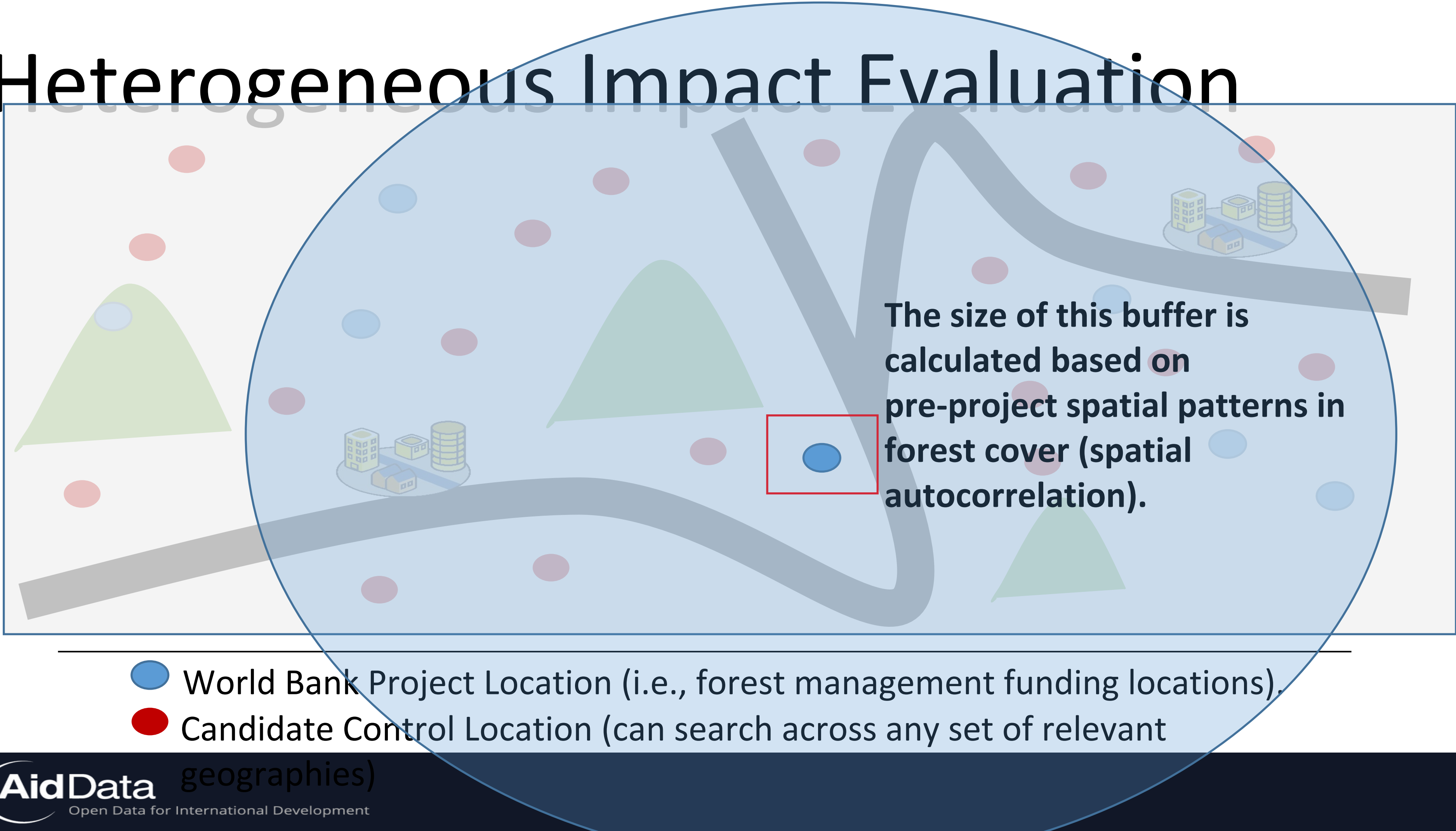
Heterogeneous Impact Evaluation



World Bank Project Location (i.e., forest management funding locations).

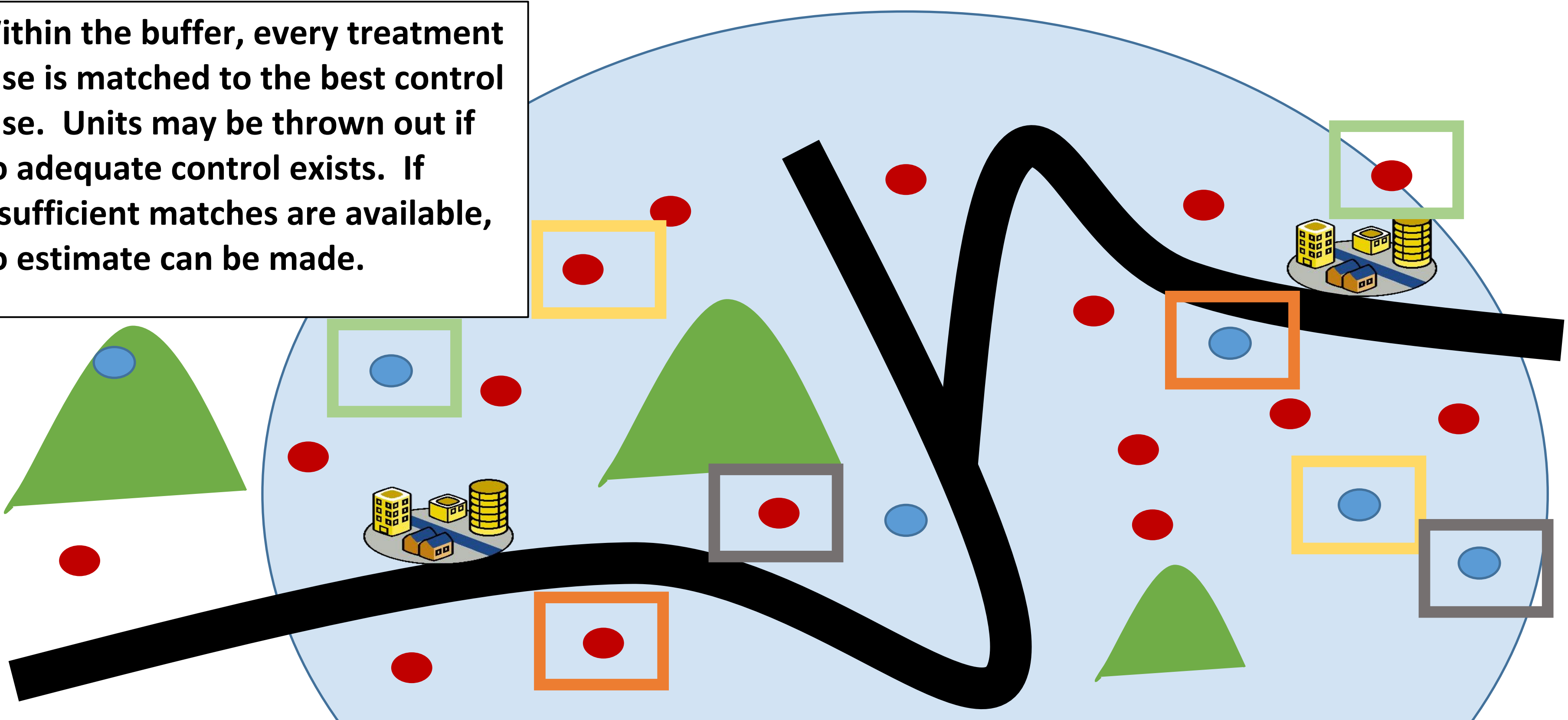
Candidate Control Location (can search across any set of relevant geographies)

Heterogeneous Impact Evaluation



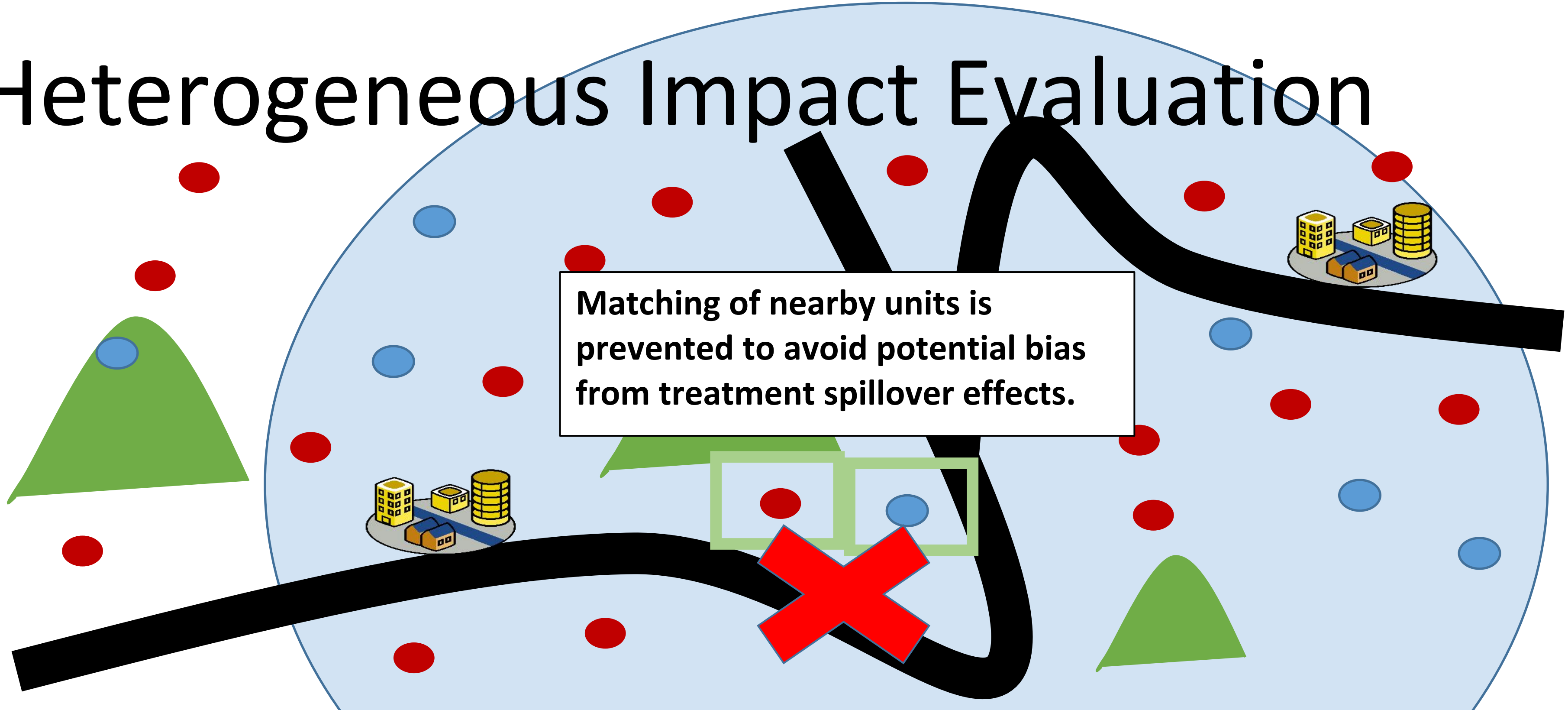
- World Bank Project Location (i.e., forest management funding locations).
- Candidate Control Location (can search across any set of relevant geographies)

Within the buffer, every treatment case is matched to the best control case. Units may be thrown out if no adequate control exists. If insufficient matches are available, no estimate can be made.



- World Bank Project Location (i.e., forest management funding locations).
- Candidate Control Location (can search across any set of relevant geographies)

Heterogeneous Impact Evaluation

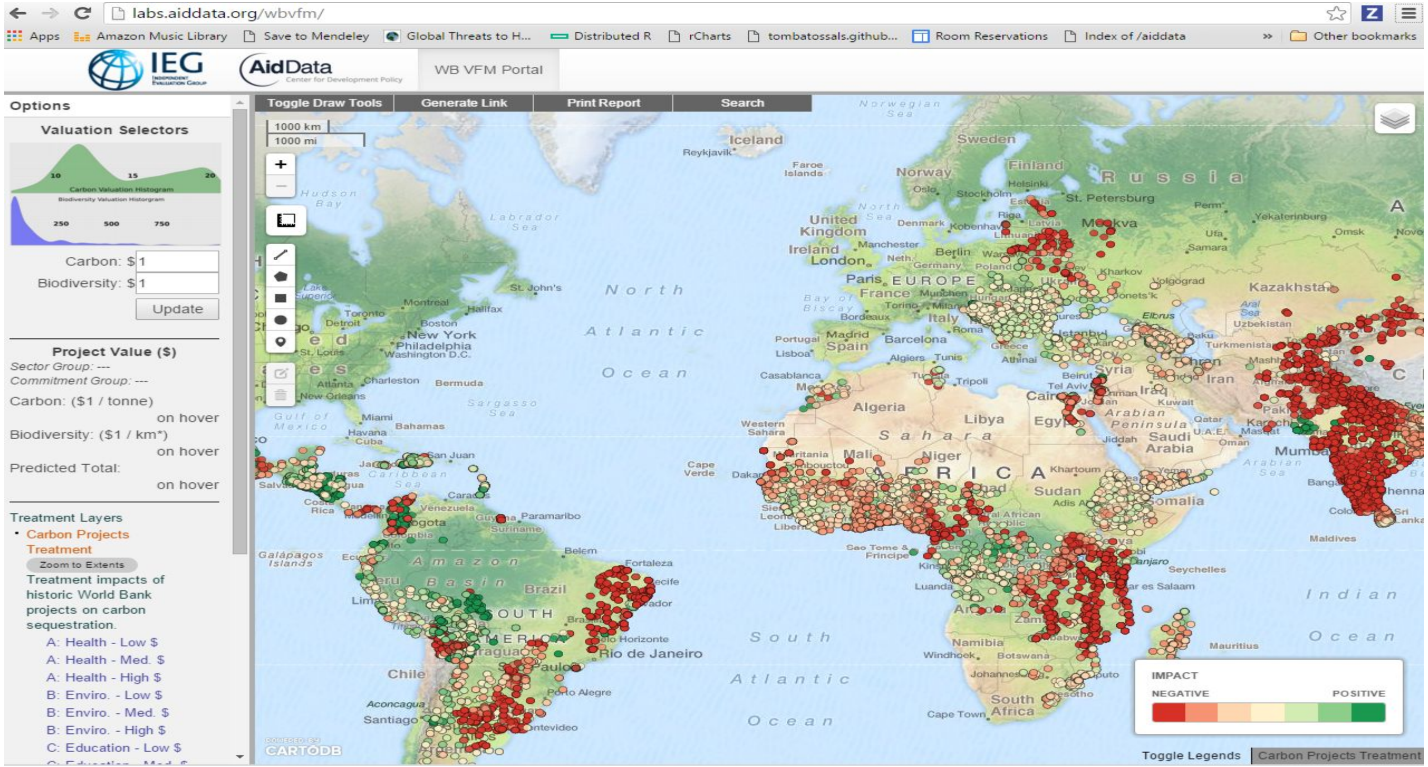


- World Bank Project Location (i.e., forest management funding locations).
- Candidate Control Location (can search across any set of relevant geographies)

Geographically Heterogeneous Design

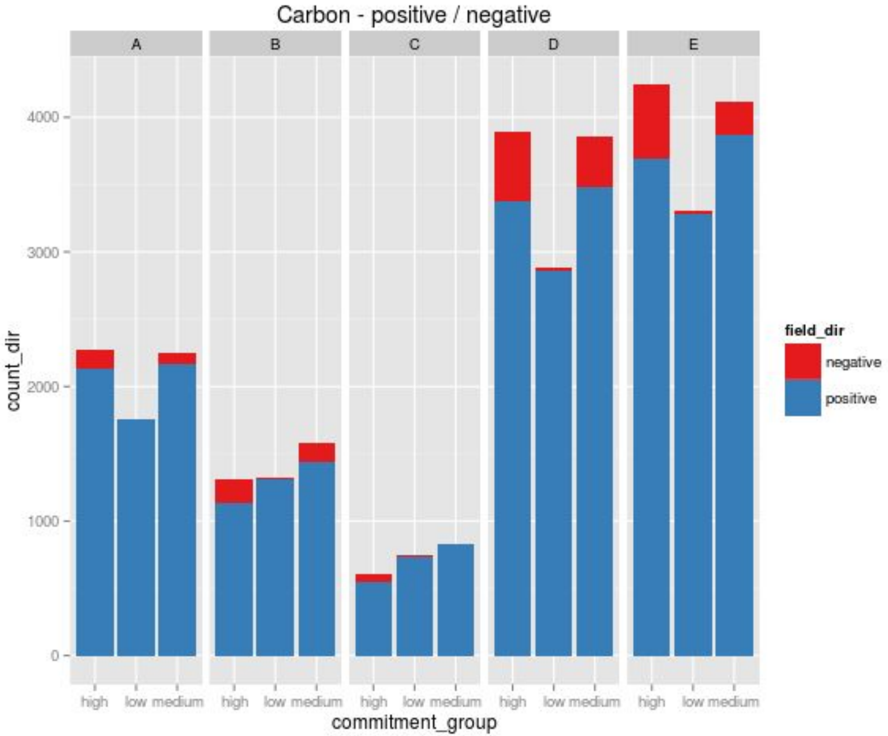
- For every "treatment" location, pairs are constructed and a second-stage analytic model is performed (i.e., an OLS model).
- The treatment effect for that treated location is then saved for every estimated treatment site, along with relevant information on standard errors.

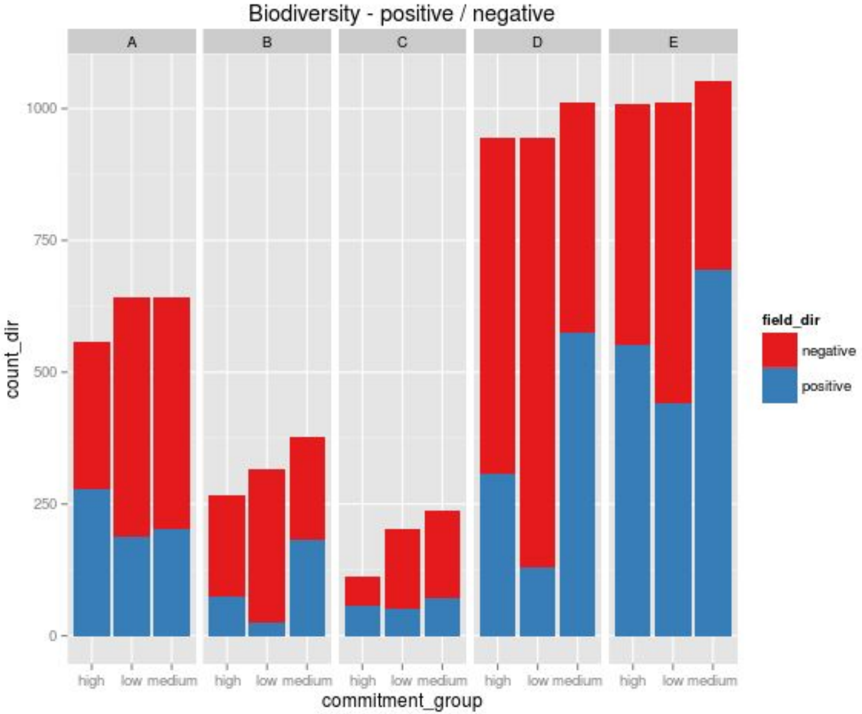
Case Study: World Bank Impacts on Forest Cover & Carbon

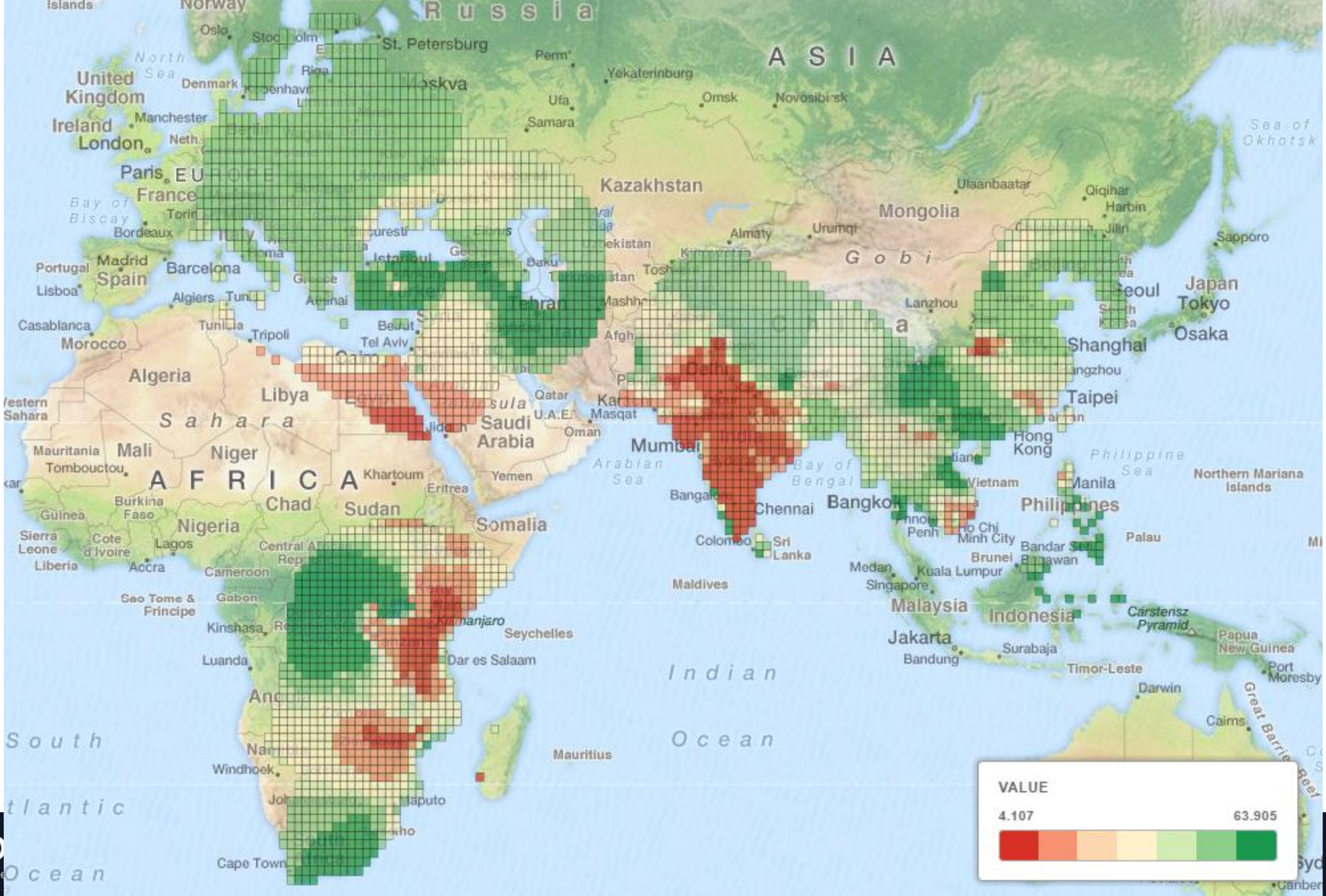


AidData

Open Data for International Development


Carbon - positive / negative
A
B
C
D
E
4000
3000
2000
1000
0
count_dir
high
low
medium
commitment_group
field_dir
negative
positive


Biodiversity - positive / negative
A
B
C
D
E
1000
750
500
250
0
count_dir
high
low
medium
commitment_group
field_dir
negative
positive


VALUE
4.107
63.905

	GEF LD	World Bank IEG
Causal Strategy	Matching	Matching
Unit of Observation	25km Buffer	10km Buffer
Treatment Definition	Location of GEF Intervention	(1) Large dollar value projects vs. Small dollar value projects
Controls	Random 25km Buffers of non-intervention, between 50 and 150km from GEF intervention; within the same country.	Known World Bank projects with low dollar value implementations.
Outcome Variables	Change in pre- to post-period NDVI, Forest Cover, Forest Fragmentation	Change in Forest Cover, In-situ measurements of Biodiversity
Matching Dimensions	Protected Areas, Temperature Precipitation, variety of GEF-specific project characteristics	Temperature, Precipitation, Distance to Infrastructure and Accessibility Metrics, variety of World Bank-specific attributes
Balance Test	Propensity Score and Distribution Diagnostics	N/A
Model	Causal Tree, Matched interaction model to test robustness.	Matched interaction model with geographic weighting
Imprecision	Monte-carlo Robustness Analysis	Omitted

Case Study 3:
Chinese Aid’s Impact in Mekong Basin

Robert Marty (1) et al.
(1) Data Specialist, DIME, World Bank

Environmental Impacts of Chinese Aid & Investment

- Funded by the John D. and Catherine T. MacArthur Foundation to evaluate the impact of Chinese development projects in the Mekong Delta

- Create a publicly accessible database of Chinese development finance and track the impact of Chinese development activities in these regions.

- Use the completed dataset to rigorously assess the impact of Chinese activity on deforestation.

Chinese Aid in Greater Mekong Basin

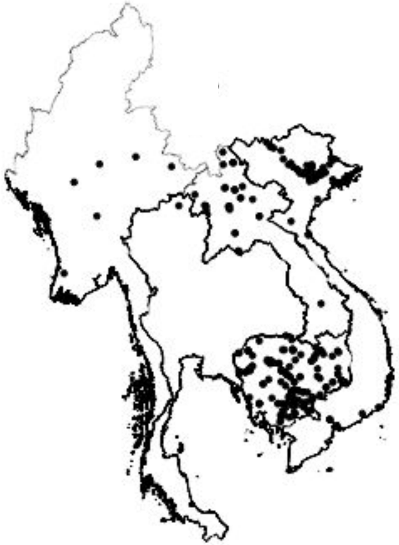
Unit of Analysis

Chinese Aid in Greater Mekong Basin

Aid was allocated . . .

Unit of Analysis

to specific coordinates

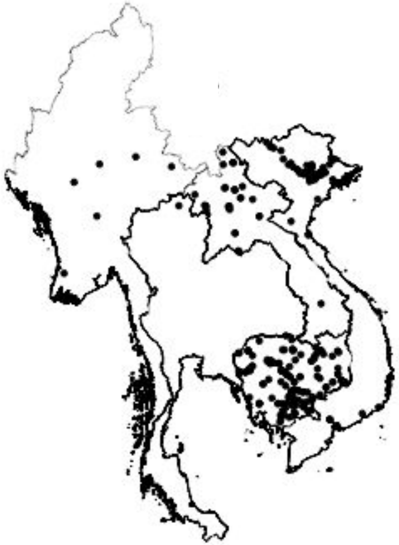


Chinese Aid in Greater Mekong Basin

Aid was allocated . . .

Unit of Analysis

to specific coordinates



somewhere in a district

somewhere in the country

Approaches to Addressing Spatial Imprecision

With regard to the cumulative disbursement figure for projects spanning multiple [traditional authorities], we have made an estimate of proportional project allocation by **weighting aid project allocation by [traditional authority] population size**."

—*The Foreign Aid Effectiveness Debate: Evidence from Malawi,* ***AidData Working Paper #6***

Approaches to Addressing Spatial Imprecision

With regard to the cumulative disbursement figure for projects spanning multiple [traditional authorities], we have made an estimate of proportional project allocation by **weighting aid project allocation by [traditional authority] population size**."

—*The Foreign Aid Effectiveness Debate: Evidence from Malawi,* ***AidData Working Paper #6***

...focusing on projects with recorded locations coded as **corresponding to an exact location or as 'near', in the 'area' of, or up to 25km away from an exact location** (precision categories 1 and 2 in Strandow et al. 2011).

—*Chinese aid and local corruption,* ***University of Gothenburg, Working Paper in Economics #667***

Approaches to Addressing Spatial Imprecision

With regard to the cumulative disbursement figure for projects spanning multiple [traditional authorities], we have made an estimate of proportional project allocation by **weighting aid project allocation by [traditional authority] population size**."

—*The Foreign Aid Effectiveness Debate: Evidence from Malawi,* ***AidData Working Paper #6***

...focusing on projects with recorded locations coded as **corresponding to an exact location or as 'near', in the 'area' of, or up to 25km away from an exact location** (precision categories 1 and 2 in Strandow et al. 2011).

—*Chinese aid and local corruption,* ***University of Gothenburg, Working Paper in Economics #667***

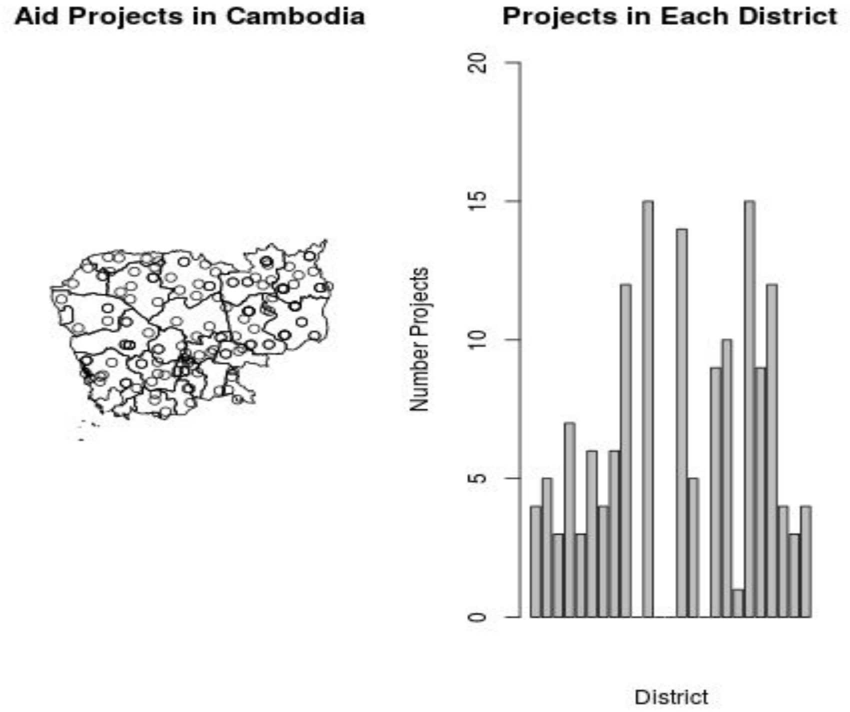
Make an assumption about aid allocation

Exclude spatially imprecise data

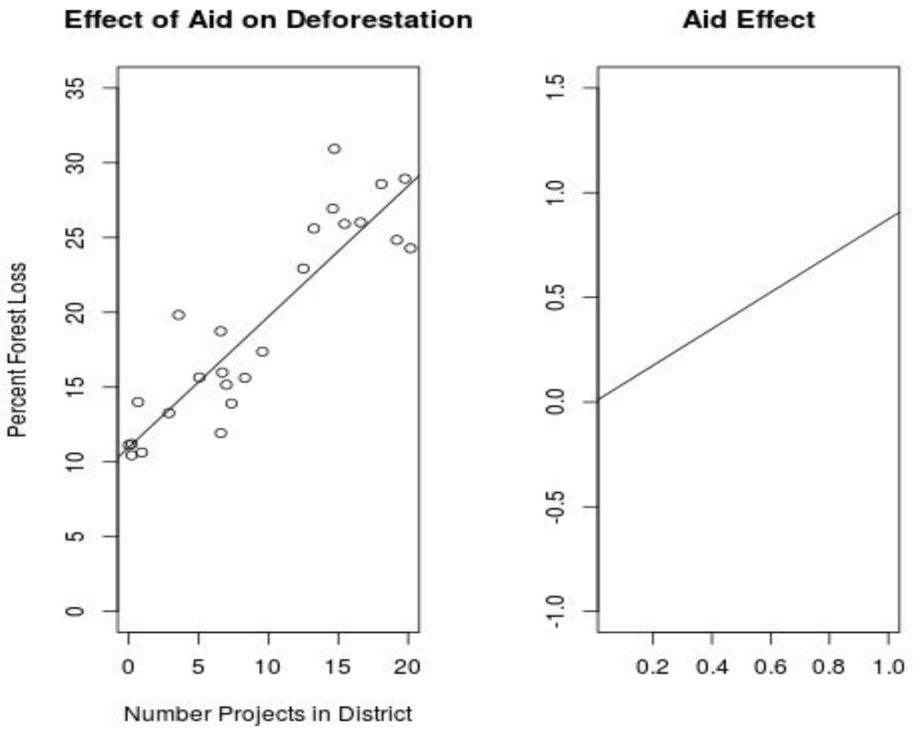
How Spatial Uncertainty Affects Research

We want to relate level of aid in a region with the level of deforestation. But...

How Spatial Uncertainty Affects Research

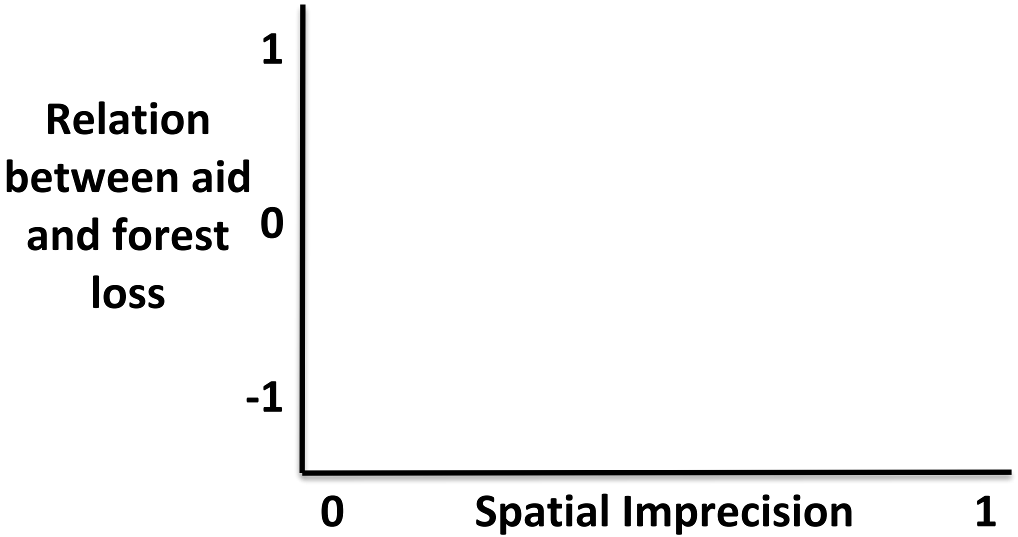


How Spatial Uncertainty Affects Research



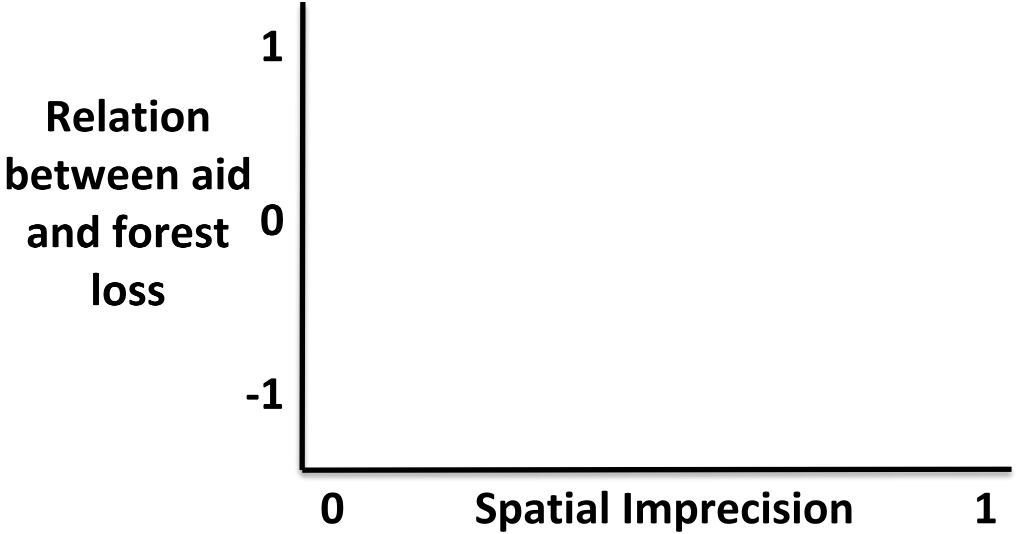
Incorporating Spatial Imprecision: geoSIMEX

Geographic Simulation and Extrapolation Method (geoSIMEX)



Incorporating Spatial Imprecision: geoSIMEX

Geographic Simulation and Extrapolation Method (geoSIMEX)

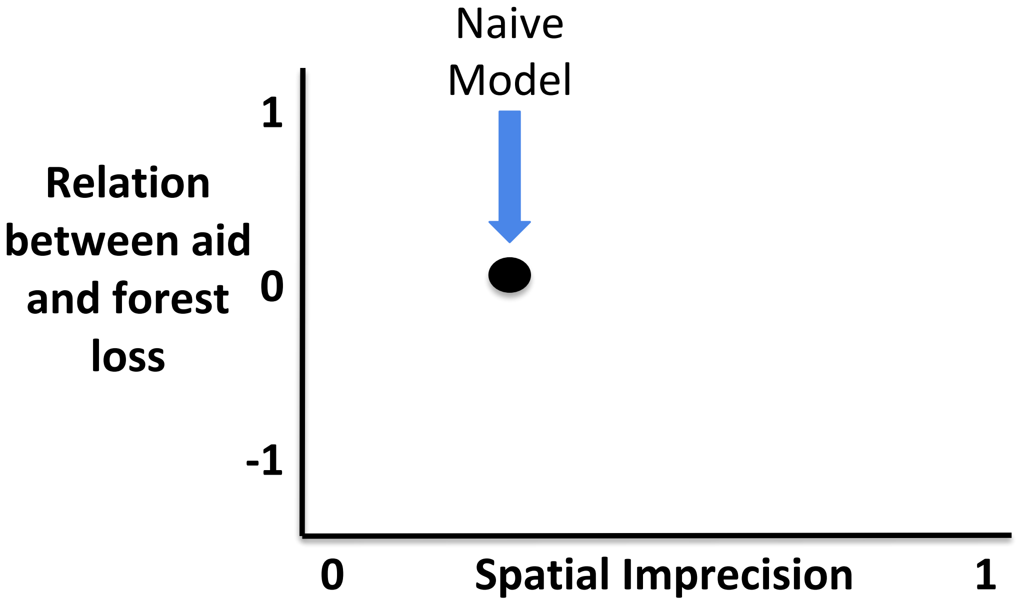


Spatial Imprecision

- *Low Spatial Imprecision:* Exact coordinates known for most projects
- *High Spatial Impression:* Many projects where only know district or region where project was allocated

Incorporating Spatial Imprecision: geoSIMEX

Geographic Simulation and Extrapolation Method (geoSIMEX)

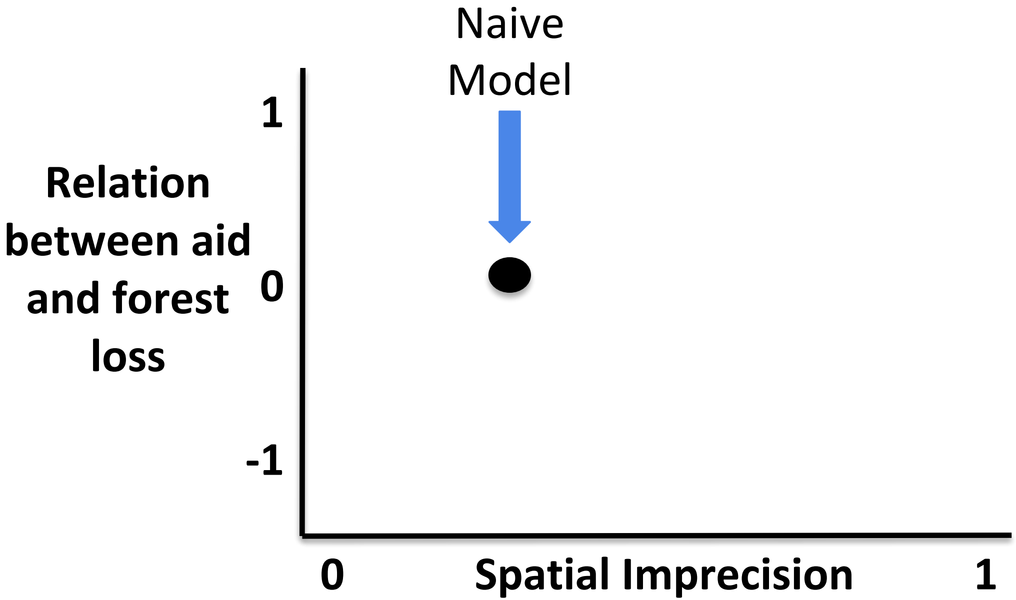


Spatial Imprecision

- *Low Spatial Imprecision:* Exact coordinates known for most projects
- *High Spatial Impression:* Many projects where only know district or region where project was allocated

Incorporating Spatial Imprecision: geoSIMEX

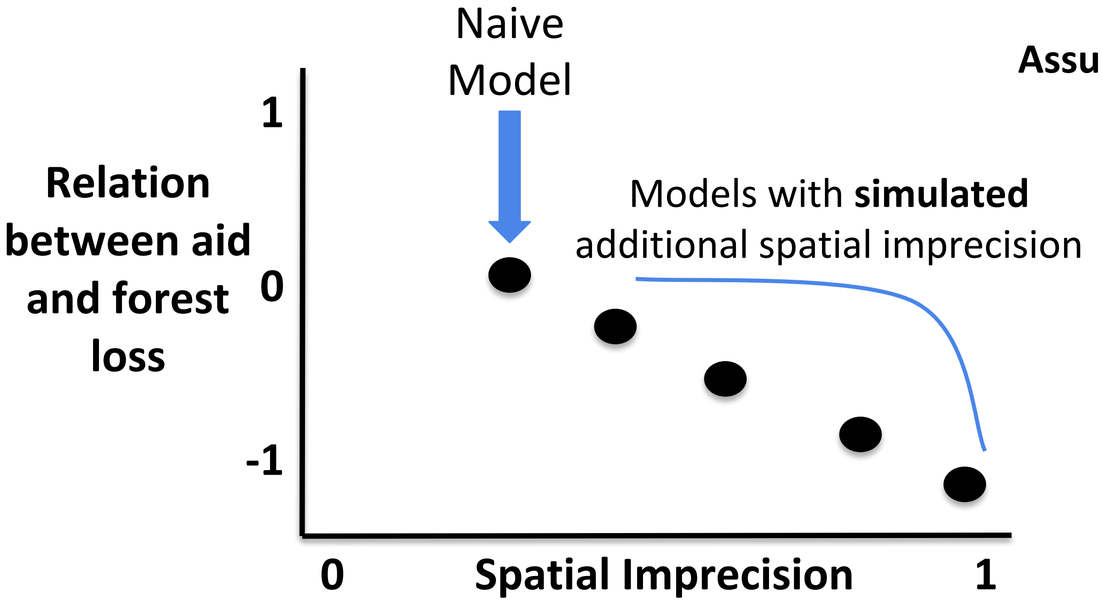
Geographic Simulation and Extrapolation Method (geoSIMEX)



Assume larger sub-counties received more aid

Incorporating Spatial Imprecision: geoSIMEX

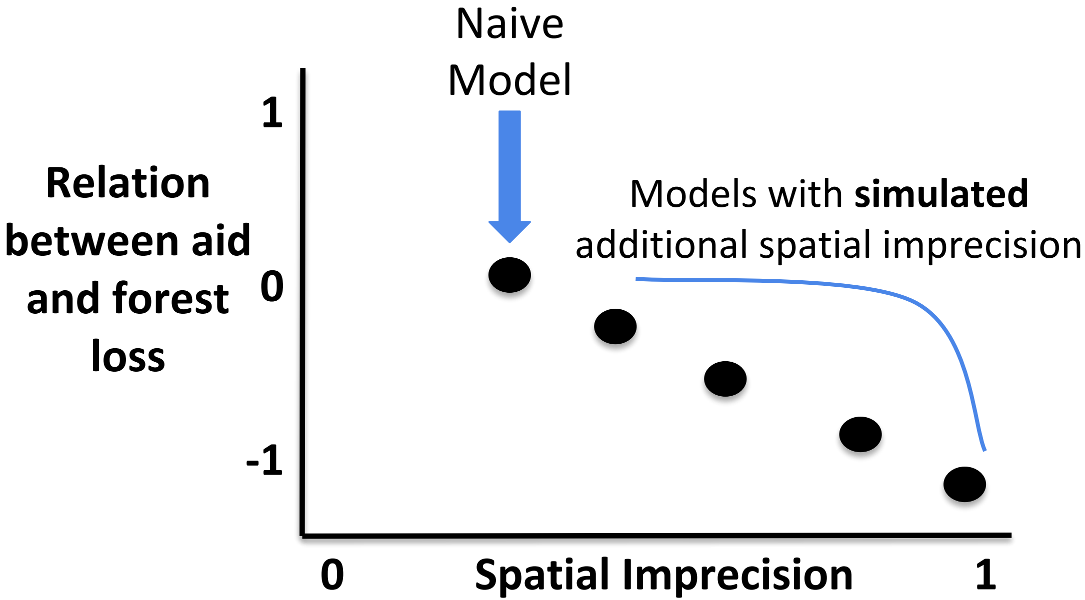
Geographic Simulation and Extrapolation Method (geoSIMEX)



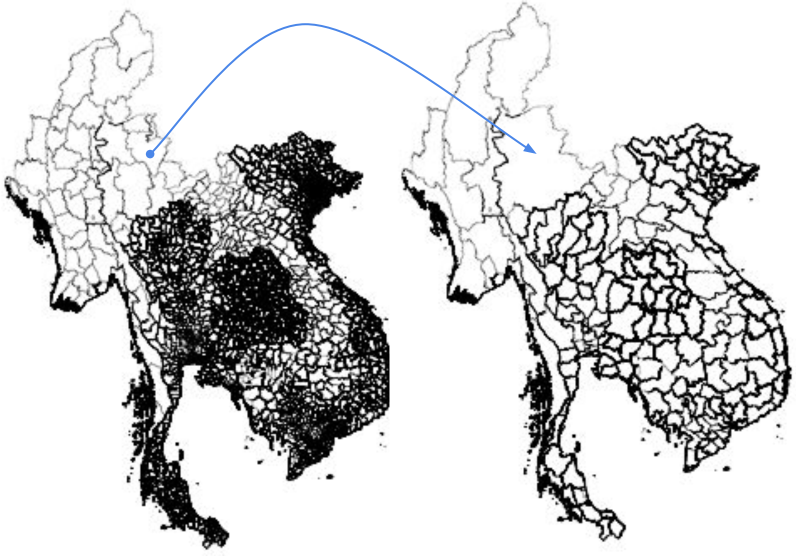
Assume larger sub-counties received more aid

Incorporating Spatial Imprecision: geoSIMEX

Geographic Simulation and Extrapolation Method (geoSIMEX)

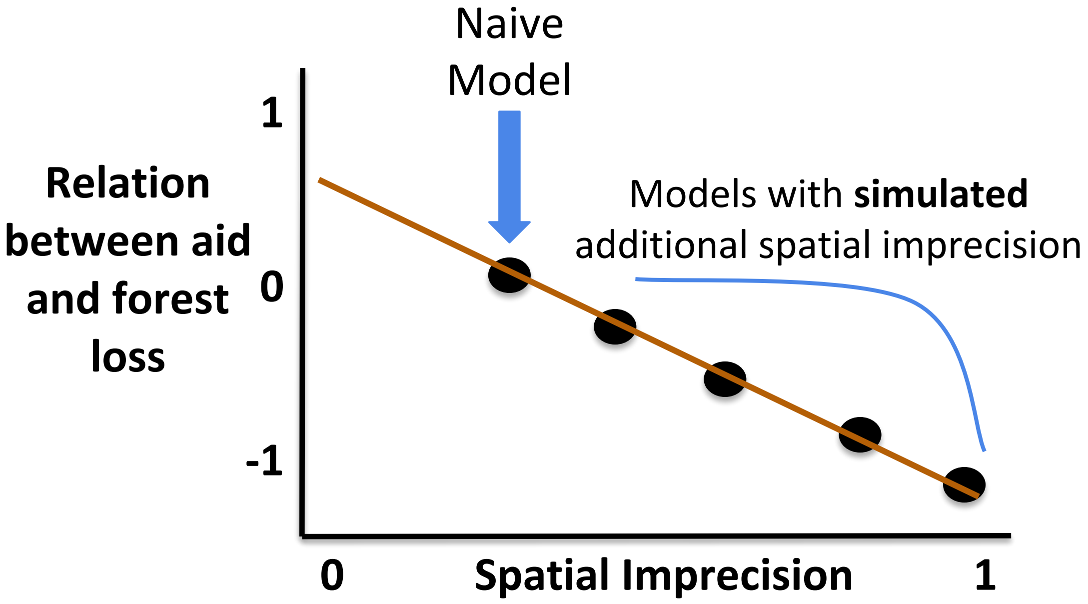


Simulating Additional Spatial Imprecision



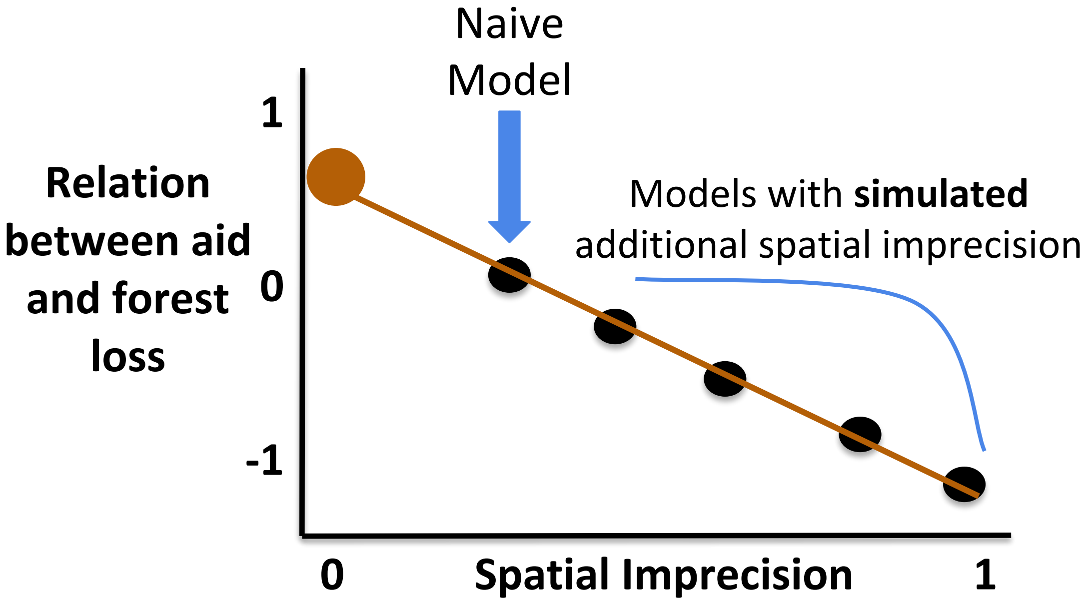
Incorporating Spatial Imprecision: geoSIMEX

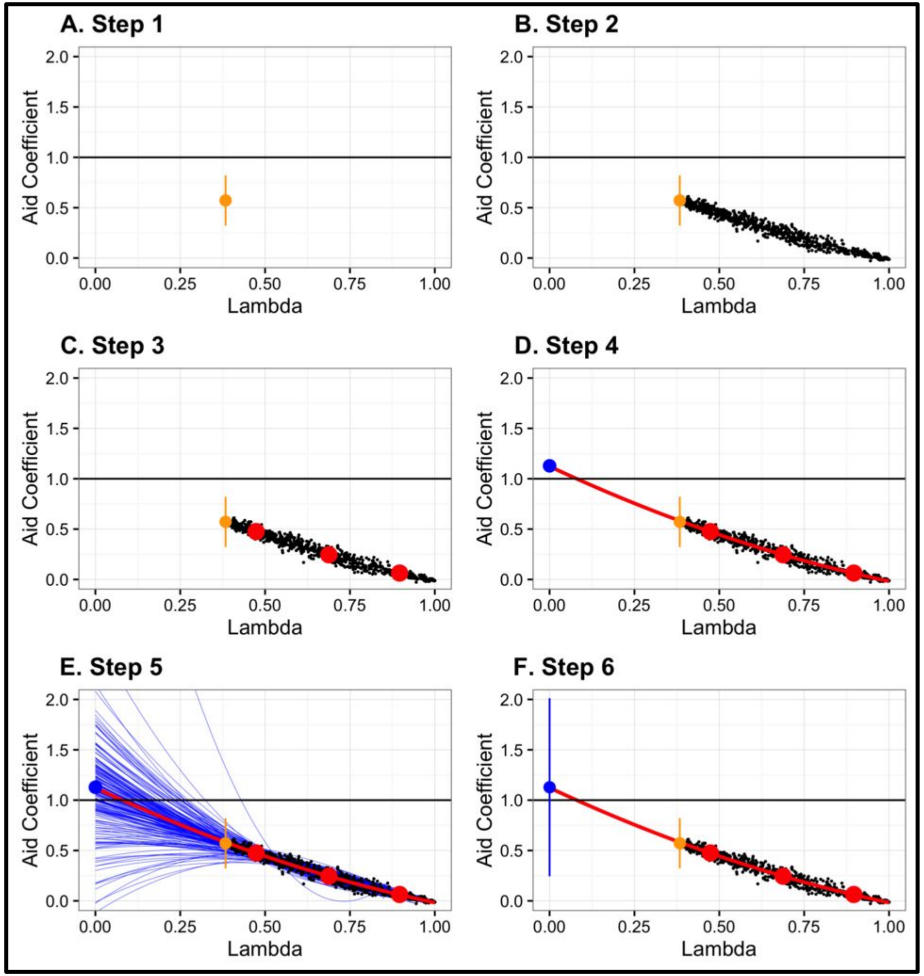
Geographic Simulation and Extrapolation Method (geoSIMEX)



Incorporating Spatial Imprecision: geoSIMEX

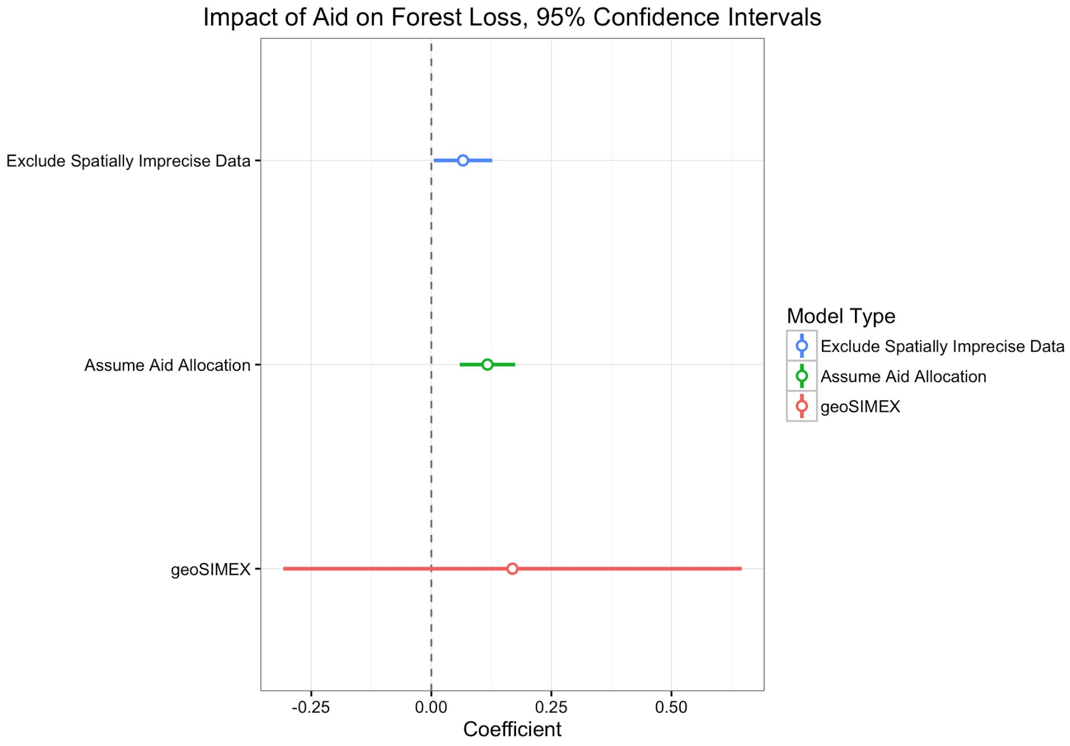
Geographic Simulation and Extrapolation Method (geoSIMEX)





A. Step 1
B. Step 2
C. Step 3
D. Step 4
E. Step 5
F. Step 6
Aid Coefficient
Lambda
0.0
0.5
1.0
1.5
2.0
0.00
0.25
0.50
0.75
1.00

Impact of Chinese Aid on Deforestation: Results



Incorporating Spatial Imprecision: geoSIMEX

Ability of Models to Capture True Relation Between Aid and Other Variable

	Assume Aid Allocation	Exclude Spatially Imprecise Data	geoSIMEX Model
% of time model captures true relation	38%	67%	95%

Capturing the true coefficient refers to the 95% confidence interval including the true coefficient

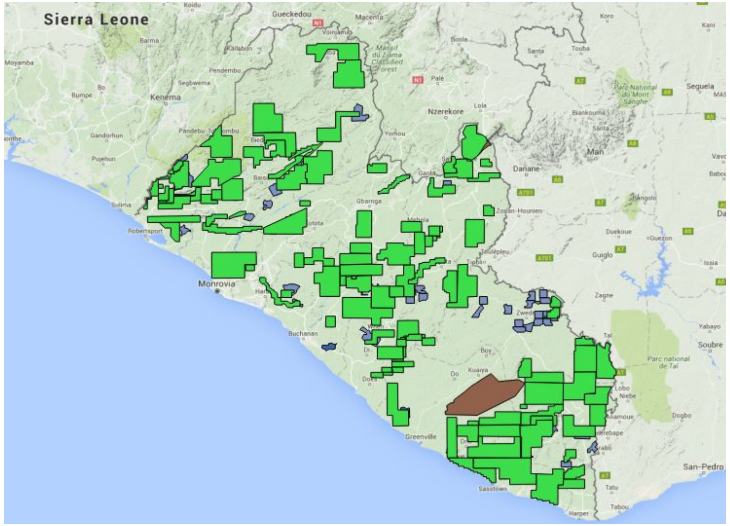
	GEF LD	MacArthur Chinese Aid & Deforestation
Causal Strategy	Matching	Panel Fixed Effects
Unit of Observation	25km Buffer	District
Treatment Definition	Location of GEF Intervention	Continuous dollar values of Chinese Aid
Controls	Random 25km Buffers of non-intervention, between 50 and 150km from GEF intervention; within the same country.	Difference-in-Difference, so each unit is contrasted to its own pre-intervention baseline.
Outcome Variables	Change in pre- to post-period NDVI, Forest Cover, Forest Fragmentation	Change in NDVI
Matching Dimensions	Protected Areas, Temperature Precipitation, variety of GEF-specific project characteristics	Temperature, Precipitation, Nighttime Lights (limited temporal data)
Balance Test	Propensity Score and Distribution Diagnostics	N/A
Model	Causal Tree, Matched interaction model to test robustness.	GeoSIMEX
Imprecision	Monte-carlo Robustness Analysis	Incorporated into Model

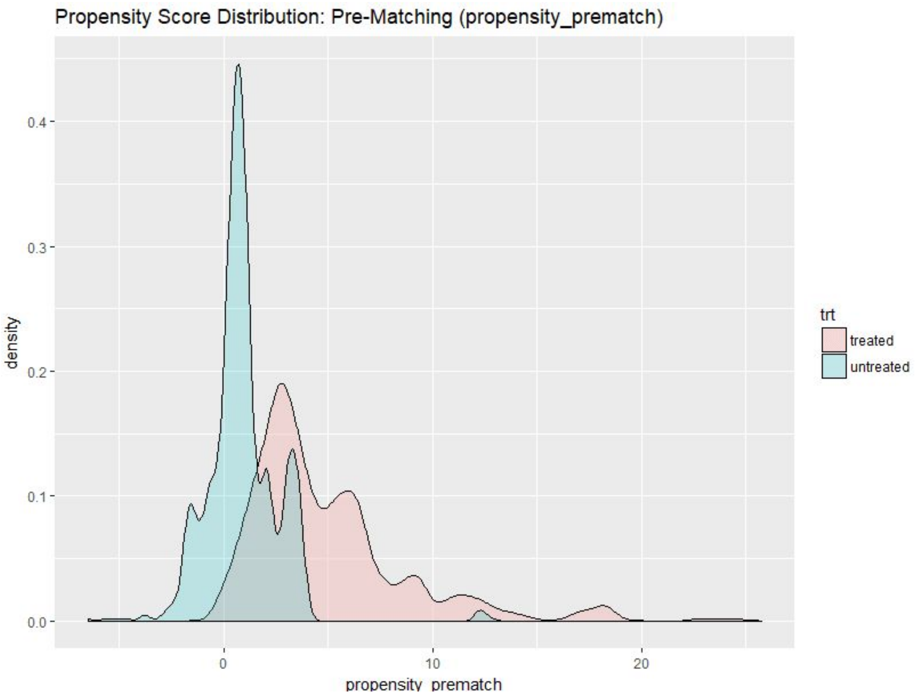
Case Study 4:
Land Concessions Impact in Liberia

Brad Parks, Jonas Bunte, Harsh Desai

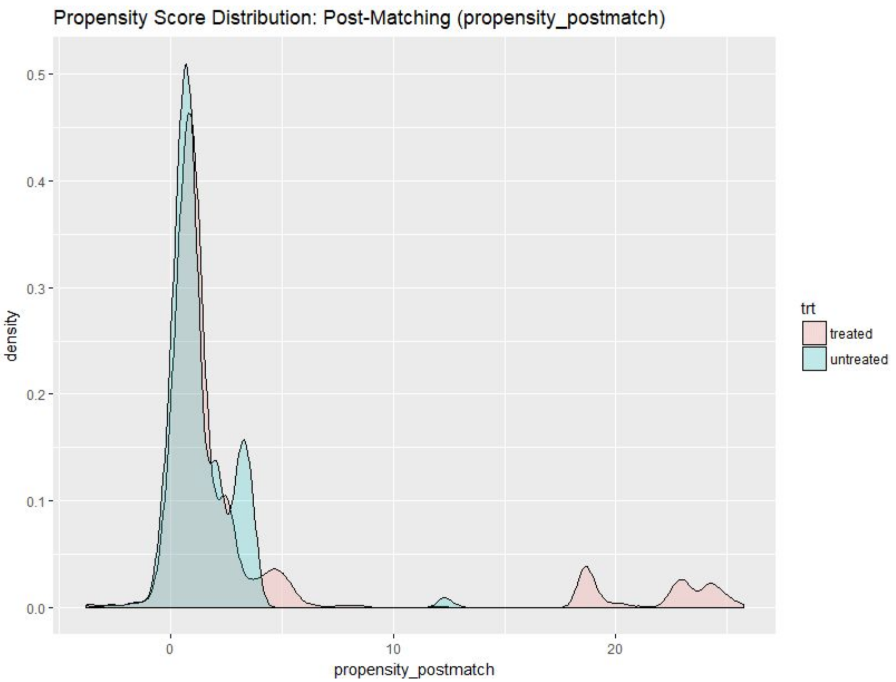
GIE in post-conflict setting: Liberia

- Conducted in collaboration with the International Growth Centre
- Impact of natural resource concessions on infrastructure outcomes.
- Digitized polygons (tracts of land granted to investors for exploration and extraction) from concession contracts made available via EITI





Propensity Score Distribution: Pre-Matching (propensity_prematch)
density
0.0
0.1
0.2
0.3
0.4
0
10
20
propensity_prematch
trt
treated
untreated



Propensity Score Distribution: Post-Matching (propensity_postmatch)
density
0.0
0.1
0.2
0.3
0.4
0.5
0
10
20
propensity_postmatch
trt
treated
untreated

	GEF LD	Liberia Concessions (IGC)
Causal Strategy	Matching	Matching
Unit of Observation	25km Buffer	Cells within 2 or 5km of DHS Enumeration Areas (accounting for spatial imprecision)
Treatment Definition	Location of GEF Intervention	In Concession Area, Distance to Concession Area
Controls	Random 25km Buffers of non-intervention, between 50 and 150km from GEF intervention; within the same country.	Areas outside of a concession, distance to concession
Outcome Variables	Change in pre- to post-period NDVI, Forest Cover, Forest Fragmentation	Change in Nighttime Lights
Matching Dimensions	Protected Areas, Temperature Precipitation, variety of GEF-specific project characteristics	Temperature, Precipitation, Distance to Infrastructure and Accessibility Metrics, DHS survey data
Balance Test	Propensity Score and Distribution Diagnostics	Propensity Score and Distribution Diagnostics
Model	Causal Tree, Matched interaction model to test robustness.	Matched interaction models
Imprecision	Monte-carlo Robustness Analysis	Monte-carlo Robustness Analysis

Geospatial Impact Evaluation Methods, Tools & Applications

Independent Evaluation Office
GLOBAL ENVIRONMENT FACILITY

Day 1: Recap

Recap

Key Concepts Covered

- **Different types of spatial information and data**
- **Key concepts of impact evaluation and spatial data**
- **Choices made - and the assumptions inherent in them - in the case study of the GEF Land Degradation Portfolio.**
- **Other development organizations experiences with GIE.**
- **Basic elements of satellite imagery analysis.**

Questions?

Wrap Up

If you are using a personal laptop tomorrow, please stop by before you leave so we can make sure it's working!

Geospatial Impact Evaluation Methods, Tools & Applications

Day 2, Part 1: Recap

Independent Evaluation Office
GLOBAL ENVIRONMENT FACILITY

Schedule

Day 1 - Discussion, Theory and Principles of Geospatial Impact Evaluation

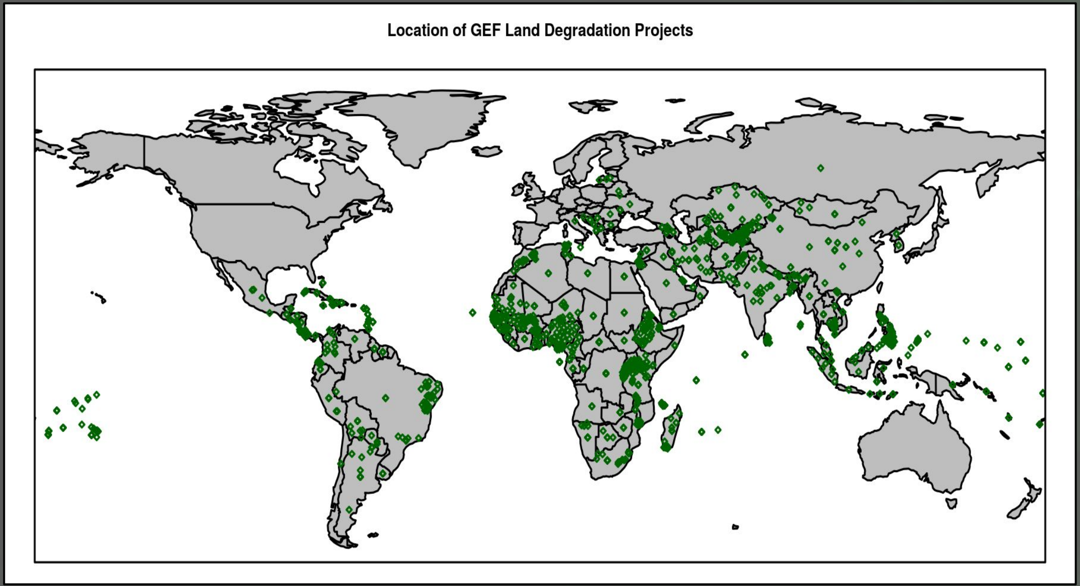
Goal: Provide working knowledge on what Geospatial Impact Evaluation is, its strengths and weaknesses, required inputs and types of outputs.

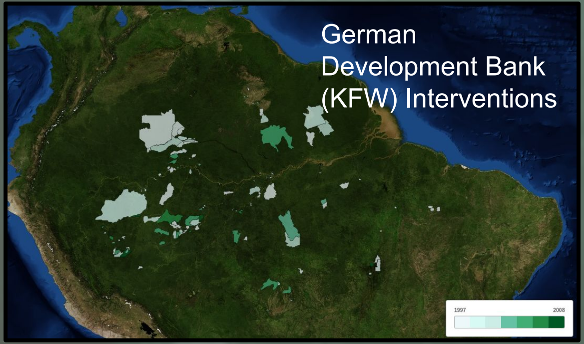
Day 2 - Hands on with Spatial Data

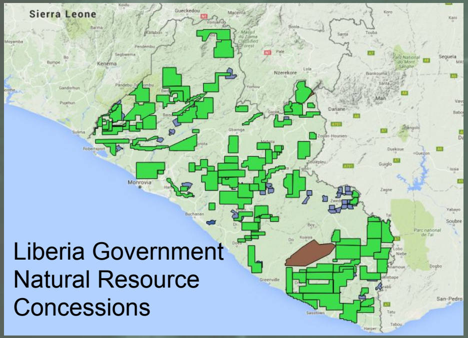
Goal: Provide hands-on experience with spatial data to expose you to the potential and limitations of a variety of spatial data sources. Teach basic GIS tools and techniques, as well as provide examples of exploratory impact evaluation.

What is Geospatial Data?: Geoparsing and Geocoding

- Results of Geocoding can have multiple levels of geographic precision.
  - Sometimes a region, country, or administrative district is known.
  - Sometimes exact points (i.e., farming plots) or polygons are identified.
  - Other times, a line such as a road or irrigation canal is found.







Geospatial Data: Integration

- **One of the biggest strengths of Geospatial Data is in integration. All of the aforementioned data types can be integrated with one another.**

- **However, there are many challenges.**
  - **Unit of observation - Buffers, administrative units, grid cells**
  - **Spatial resolution**
  - **Data error and inconsistency**
  - **Spatial imprecision**

Geospatial Data: Defining a Unit of Observation

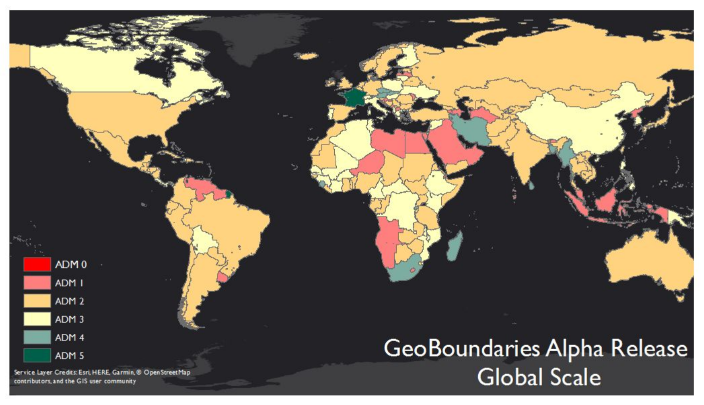
- In many studies, the unit of observation for an intervention is well defined:
  - Giving a pill to two individuals for a disease - the individual person is the unit of observation.
  - Applying fertilizer to a number of fields - the individual fields is the unit of observation.

- In geospatial analyses the unit of observation must frequently be theoretically defended.
  - If you build a canal, for how many meters or kilometers from that canal do you believe you could observe an impact?
  - If you conduct a training on sustainable farming techniques in a city center, over what geographic region do you anticipate impacts might be seen?

- Before any statistical analyses can be done, the question of unit of observation must first be answered.

Geospatial Data: Defining a Unit of Observation

- **Official Boundaries**
  - Advantages
    - Helpful when an intervention is anticipated to affect an entire decision-making unit.
    - Frequently the same units used for census activities.
  - Disadvantages
    - Can change in unmeasured ways over time.
    - Can be of variable size and have variable underlying measurement qualities.



Geospatial Data: Integration

- **The analysis you undertake is limited by the interaction between your chosen unit of analysis and the quality of data you seek to integrate.**

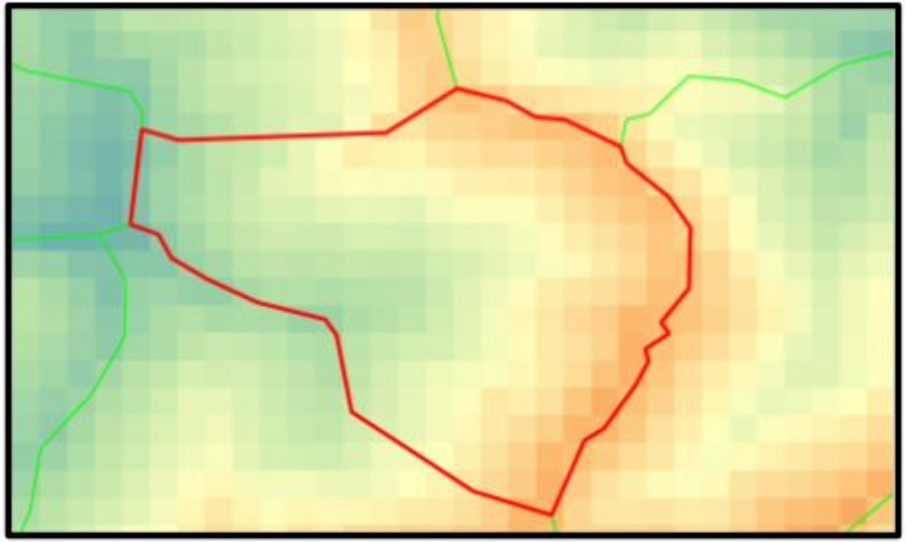
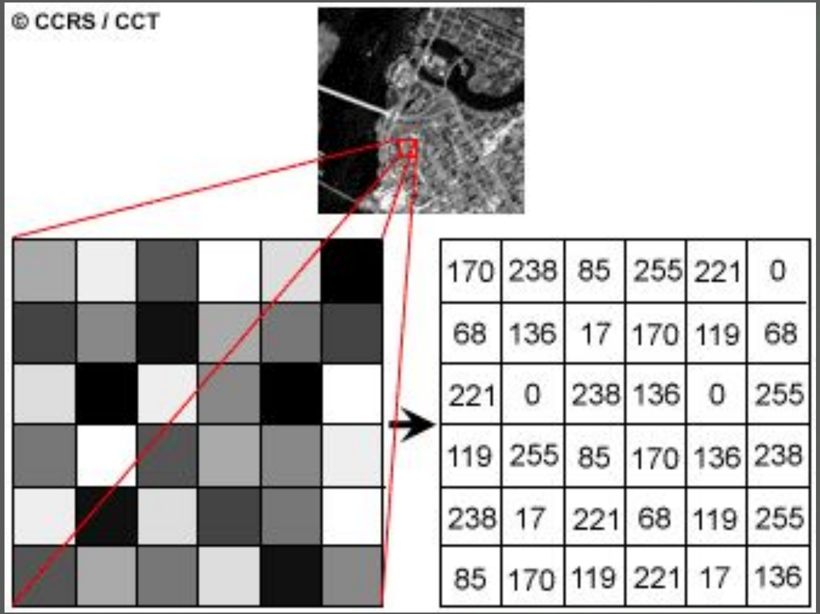


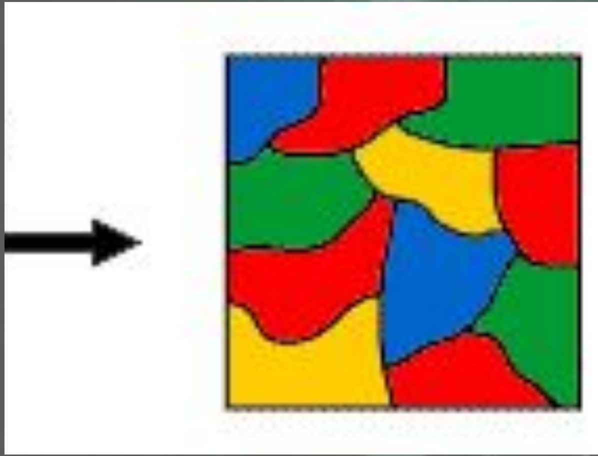
Table 1. Extract Results	
Name	Basa
Count	185
Sum	393,202
Mean	2125.416

Figure 2. The village of Basa in Nepal covers approximately 185 cells of the elevation raster. Elevation values in Basa range from around 1200 meters above sea level, represented by the red-orange pixels on the right side of the village, to over 3000 meters on the left side in the blue-green cells.

Geospatial Data: Integration

- Spatial Resolution of datasets you seek to integrate should - as a rule of thumb - be better than the unit of observation you want to work with.
- In many cases, this necessitates a trade-off between data availability and using coarser units of observation.





Geospatial Data: Integration

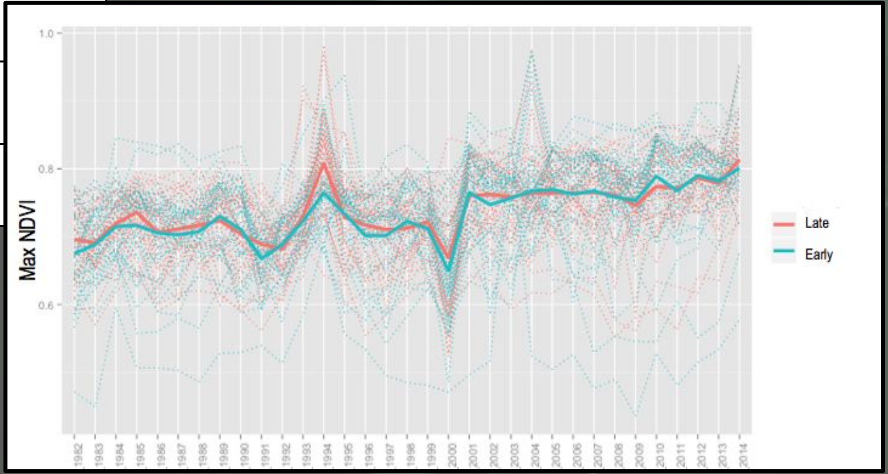
- **Ultimate goal is to get - for each unit of observation - a column of data representing a relevant spatial dataset.**

Unit of Observation	Satellite-measured vegetation in 1990	Satellite-measured vegetation in 2005
A	0.15	0.10
B	0.13	0.12
C	0.16	0.17
D	0.20	0.19

Geospatial Data: Integration

- **Ultimate goal is to get - for each unit of observation - a column of data representing a relevant spatial dataset.**

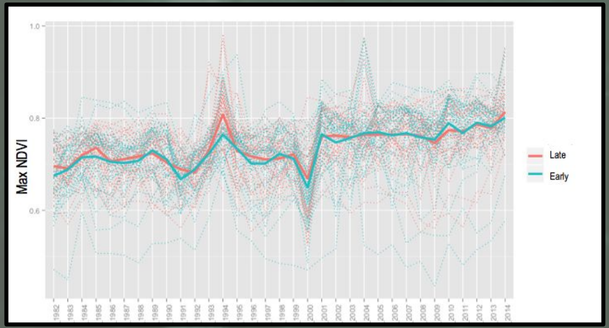
Unit of Observation	Satellite-measured vegetation in 1990	Satellite-measured vegetation in 2005
A	0.15	0.10
B	0.13	0.12
C	0.16	0.17
D	0.20	0.19



Land Rights

(Outcome: Forest Protection)

Translating your impact trajectory / theory of change into a quantifiable trend can sometimes prove a unique challenge, but is key : depending on your time frame, different spatial datasets providing measurements may or may not be available.



Impact Evaluation

Key decisions to be made for matching:

- From where in the study region should you draw controls?
  - How far away from your intervention do you believe you need to go before any impacts will be 0?
    - Avoid choosing controls in areas too close to interventions to be independent.
    - However, the closer the better, as places closer together are likely to be more similar.
  - Control cases can also be limited to - for example - country or district boundaries.

- What is the explicit outcome you will measure the impact on?
  - Deforestation in a 2-year period after the intervention? 5 year?
  - Difference between a pre-period baseline and post-period?

Impact Evaluation

Key decisions to be made for matching:

- Along which attributes should you match?
  - Do you match on temperature from every year, the average temperature from across multiple years, or other data manipulations?
  - Not all variables are equally important for matching - the goal is to ensure control and treatment cases are similar along dimensions most likely to be correlated with relevant omitted variables.
  - I.e., if you omit population but can match on Nighttime Lights, NTL is likely to mediate the lack of population data.
  - “Propensity Score Matching” helps you appropriately conduct this selection.
- How good of a match is “good enough”?
  - Variable “balance” across measured indicators.
  - Propensity score balance across measured indicators.
  - Do we look at match for individual countries, districts, or globally?

Impact Evaluation

Key decisions to be made for matching:

- How will you handle heterogeneity in your data?
  - If you believe there is only heterogeneity along a small number of known dimensions, interaction models can be a good choice.
  - If you are uncertain about the heterogeneity, or it can be nested (i.e., different across multiple countries), machine-learning based techniques can be best.

- How will you handle spatial imprecision in measurements?
  - Both intervention site data and satellite data can have spatial imprecision.
  - The number of dimensions of uncertainty you model increases the computational costs of your model exponentially. Choosing one or two dimensions is most manageable; expanding even to 3 can be challenging to compute, visualize, or interpret.
  - Examples might include: the size of the buffer used to represent spatial impact, the measured value of environmental quality from a satellite, the start date of a project, years used to calculate pre-trends, or any other decisions.
  - You can incorporate uncertainty as a source of error, or select one model and use uncertainty as a robustness test.

Discussion Point A

Cases in which spatial data is most appropriate (or not!) to use?

- **Definition of unit of observation is spatial.**

  Limits?

- **Scope of project implementation?**
- **How well defined is the theory of change? Can outcomes be quantified?**

Discussion Point B

Other modes of impact evaluation may be more effective in some contexts.

➳ **Cost benefits of GIE largely predicated on open source data that has already been analyzed or collected by experts.**

➳ **Quantitative and unbiased (mostly) due to satellite imagery; can miss qualitative outcomes.**

➳ **Satellite imagery can be expensive to process for custom, small-scale areas; practically impossible for the globe without appropriate expertise and computational resources. Even we partner with other experts!**

Discussion Point C

Implications of GIE

- **Project Evaluation**
- **Project Planning / Targeting**