Geospatial Data to Evaluate and Target Aid

Dan Runfola, Senior Geospatial Scientist, AidData

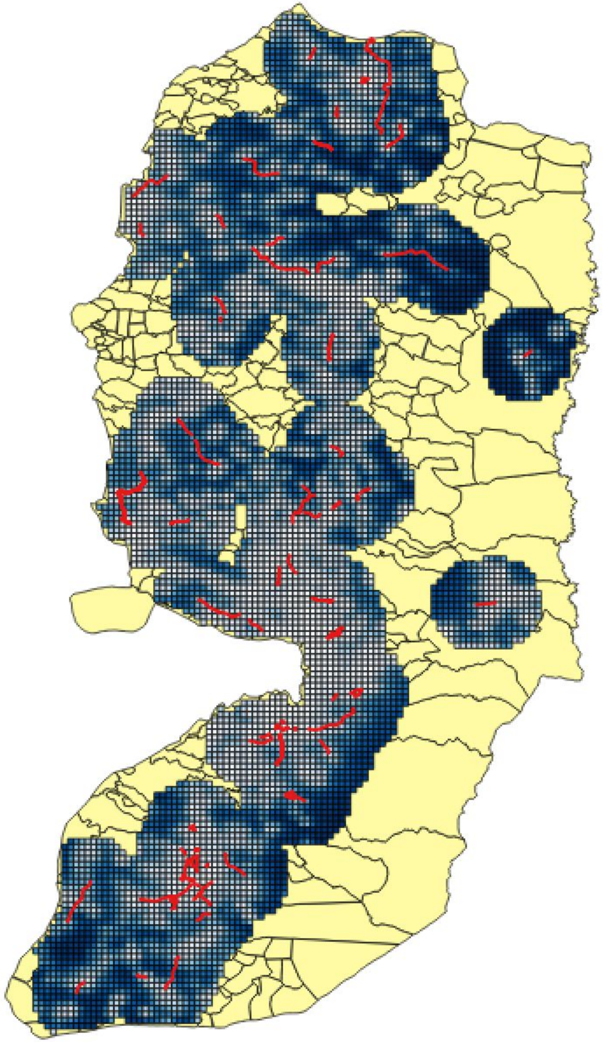
Geospatial Data is Reaching Adulthood

- Dozens of impact evaluations or targeting assessments using geospatial data over the last ~3 years.
- Geographers complaining about causal identification; Economists complaining about spatial spillovers.

No Longer Just Points

- True geometries of aid interventions are now being collected at large scales.

- Integrating this data with a wide variety of spatial data - including satellite and other products.



Examples of causally identified studies from an increasing number of sectors

- Water and Sanitation (USAID Afghanistan)
- Electrification (MCC Tanzania & Ghana)
- Poverty (IGC Liberia)
- Environment (World Bank IEG, MacArthur, GEF IEO)
- Health (Gates Foundation DRC)
- Emergent studies on governance and female empowerment (AfroBarometer)

Solutions to a growing number of methodological concerns in the use of spatial data

- **Spatial Uncertainty** - SIMEX, Bayesian Approaches
- **Spillover in treatment effects** - GeoMatch, distance-restricted control identification, new research into lag-based modeling.
- **Spatial heterogeneity in effects** - Causal matching GWR, Causal Trees
- **Data integration and access** - GeoQuery (more on this very soon!)

Case Study 1: Geospatial Impact Evaluation and Valuation of Land Degradation Projects

Juha Uitto[1],Geeta Batra[1], Anupam Anand[1], Dan Runfola[2], Ariel BenYishay[2] and Jyothy Nagol[3]

[1]Independent Evaluation office, Global Environment Facility
[2]College of William and Mary
[3] Global Land Cover Facility, University of Maryland

Objectives

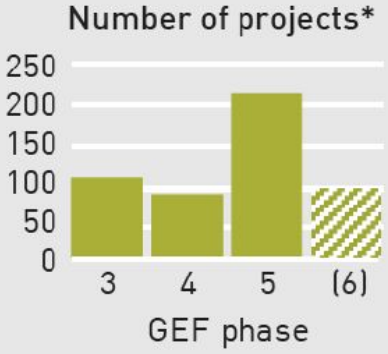
(1) Portfolio-wide impact evaluation (top-down)
(2) Identification of factors frequently associated with positive ou
(3) Valuation in terms of Carbon Sequestration

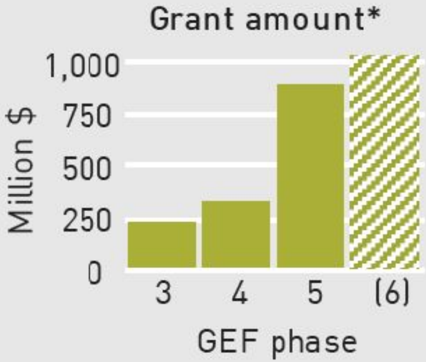
PORTFOLIO HIGHLIGHTS

237
projects

$630 million
in grant funding

$2.97 billion
in cofinancing





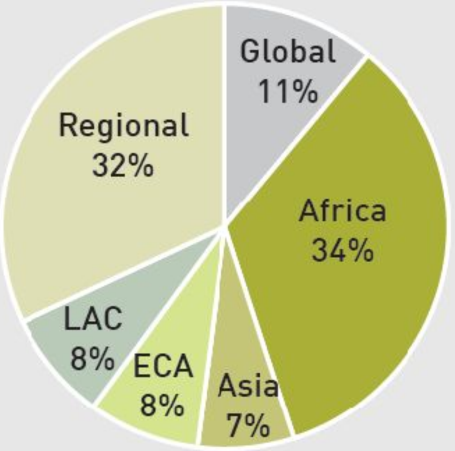
Top 3 agencies*

43% UN Development Programme
17% World Bank Group
17% UN Environment Programme

Regional distribution*

34% Africa
24% Asia
20% Latin America & Caribbean
13% Europe and Central Asia
7% Global
2% Regional

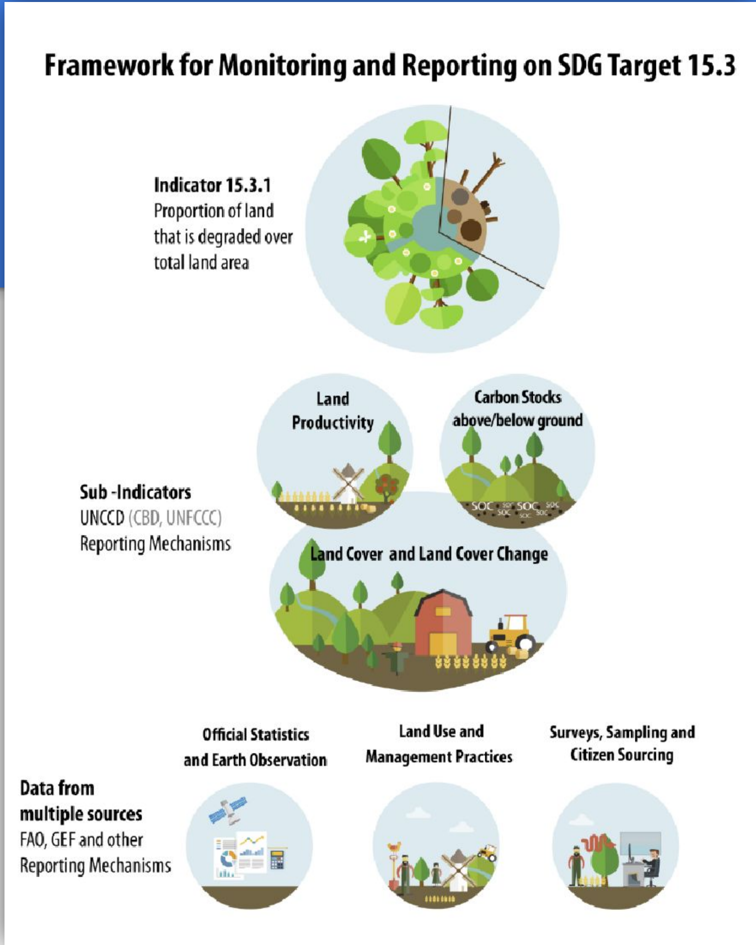
Total GEF funding per region*



*Includes LDFA and multifocal projects.

Outcomes

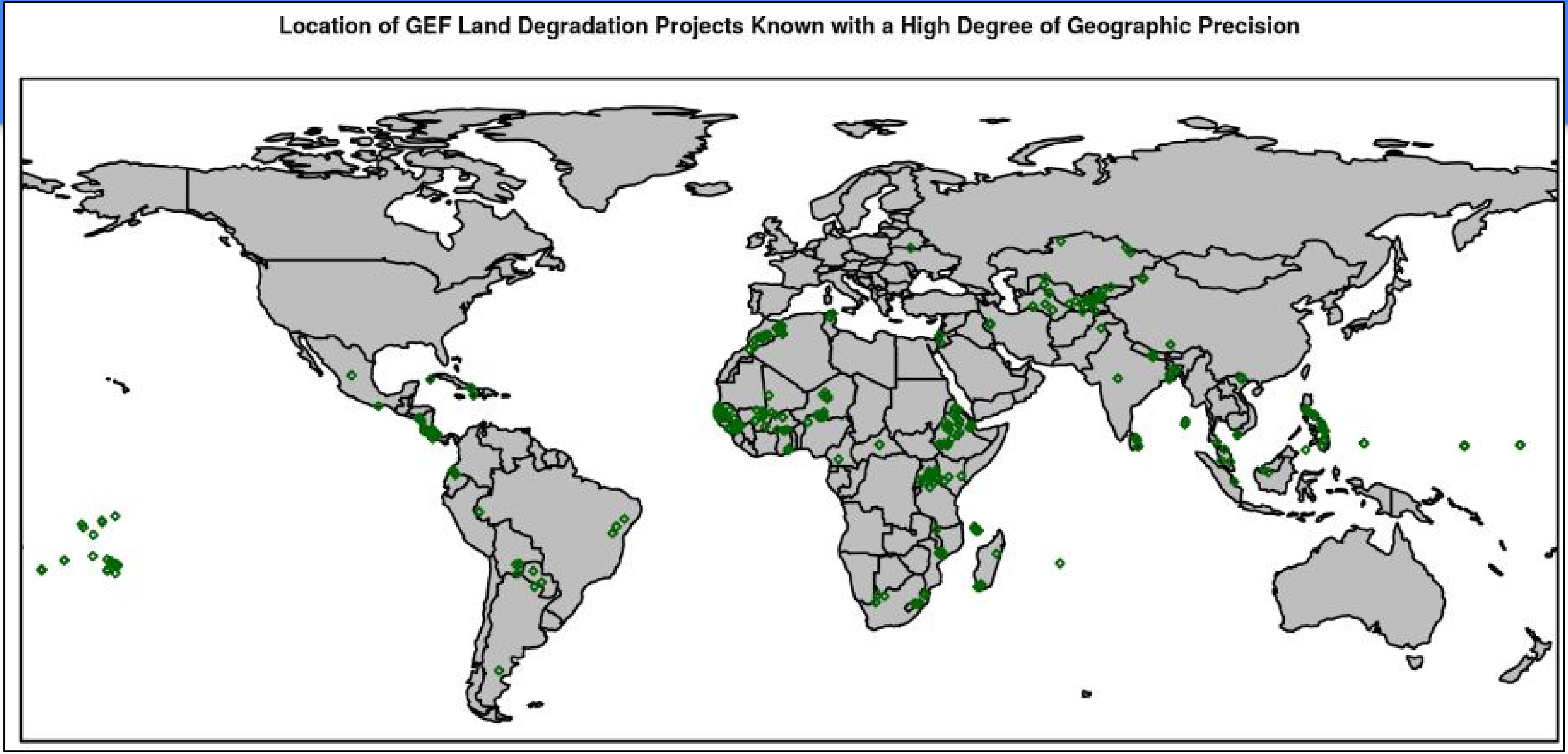
- Forest Cover
- Vegetative Density
- Forest Fragmentation



Methodological Approach

1. Geoparsing and coding GEF project locations
2. Integrating Satellite, Fragmentation, Survey, and Other Data Sources
3. Causal Inference through Cross-sectional Matching
4. Valuation of Impacts

1. Geoparsing and Coding

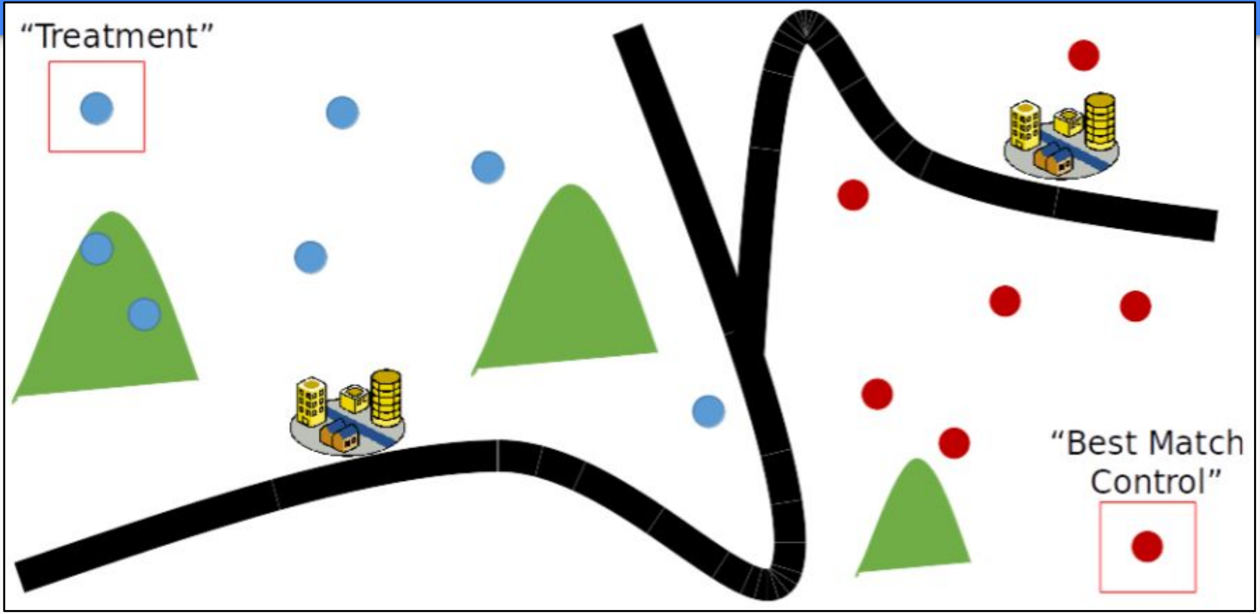


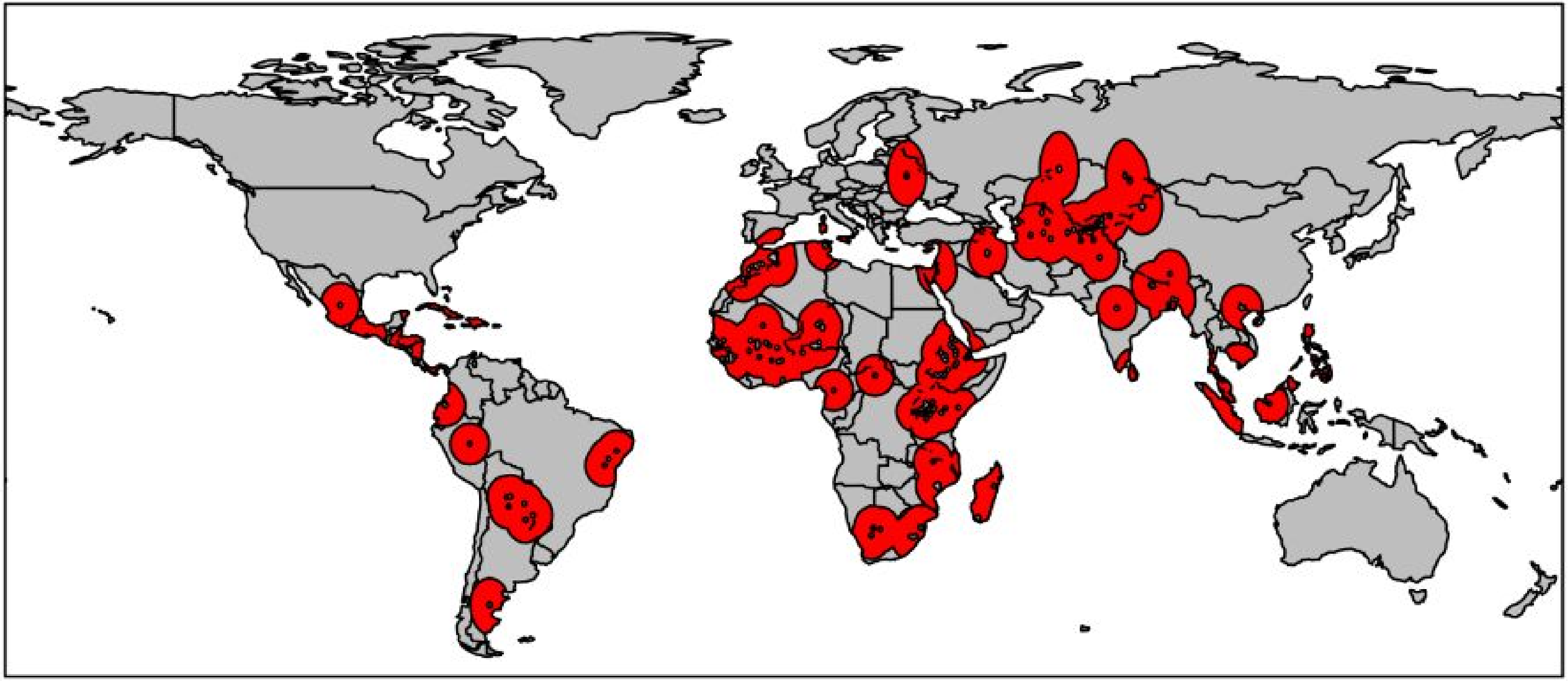
2. Data Integration

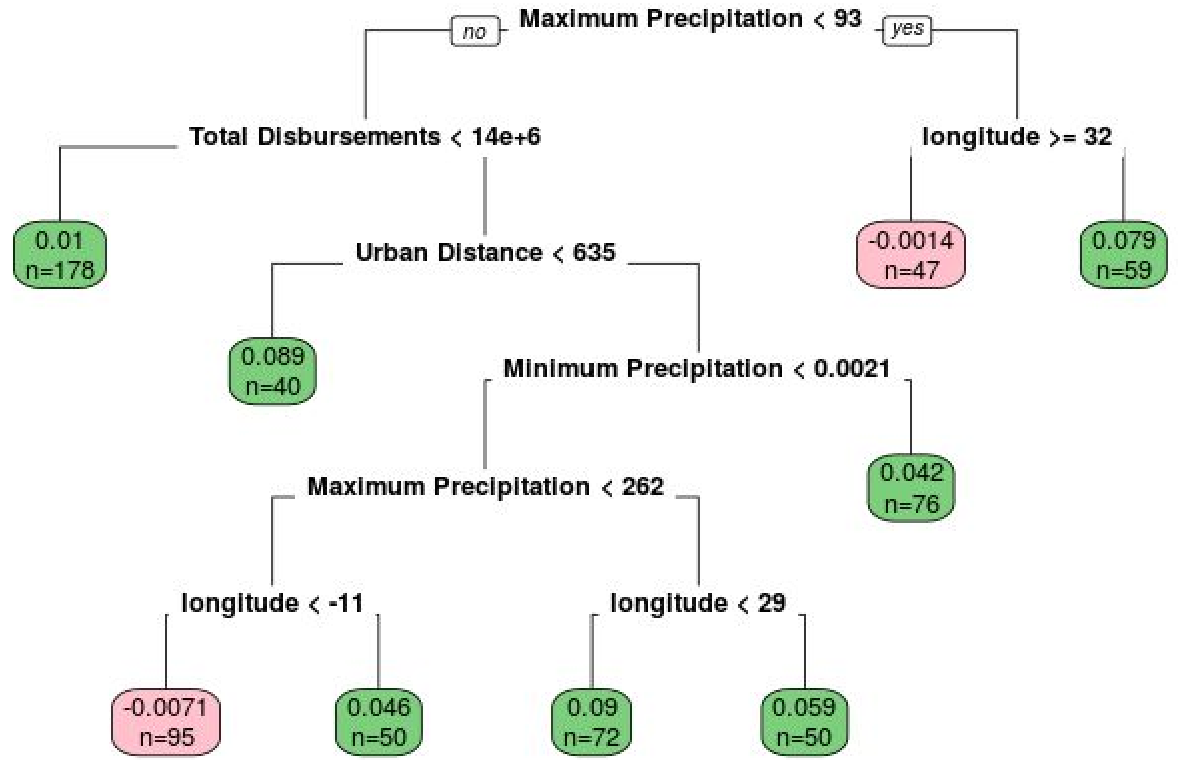
- Multi-sourced (including data from NASA, NOAA, a wide number of academic research groups, GEF project characteristics, and more.
- Multi-resolution (Monte Carlo simulation to capture uncertainties)
- Ancillary data included:
  - Distances to roads, rivers, cities, a variety of economic sites (i.e., on-shore petroleum resources, diamond mines), rainfall, precipitation, nighttime lights, GEF characteristics such as dollar value of project and year, and more.

3. Causal Model

- Cross-sectional matching (with temporal components on some dimensions).
- Propensity Score-based
- Causal Tree to capture spatial heterogeneity.

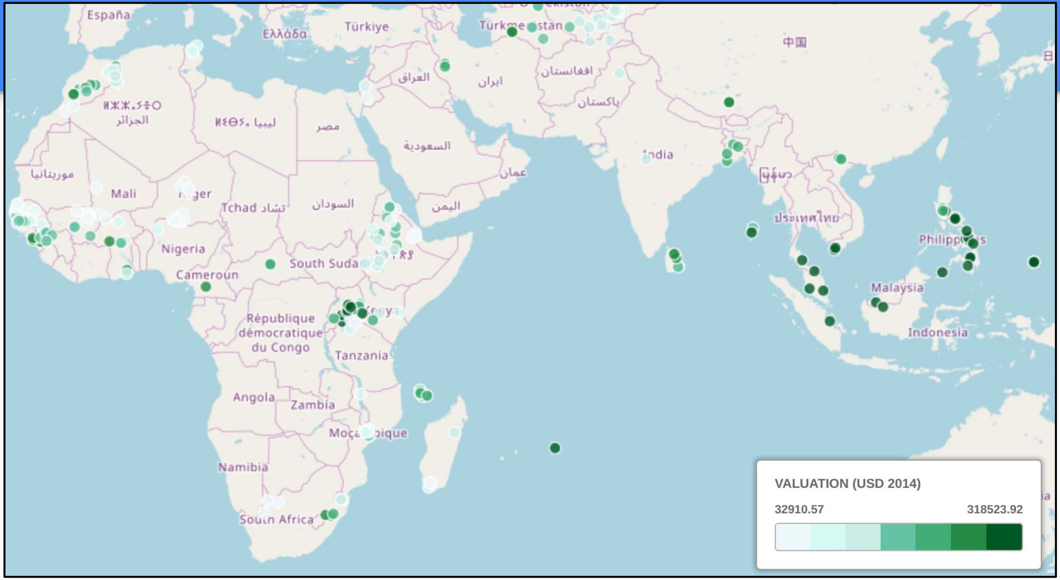





Maximum Precipitation < 93
no
yes
Total Disbursements < 14e+6
longitude >= 32
0.01
n=178
Urban Distance < 635
-0.0014
n=47
0.079
n=59
0.089
n=40
Minimum Precipitation < 0.0021
Maximum Precipitation < 262
0.042
n=76
longitude < -11
longitude < 29
-0.0071
n=95
0.046
n=50
0.09
n=72
0.059
n=50

4. Valuation

- Literature meta-review transfer approach.
- No assumptions on modeling made; tool produced to enable end-users to choose valuations.
- Mean valuation in literature used to report findings.

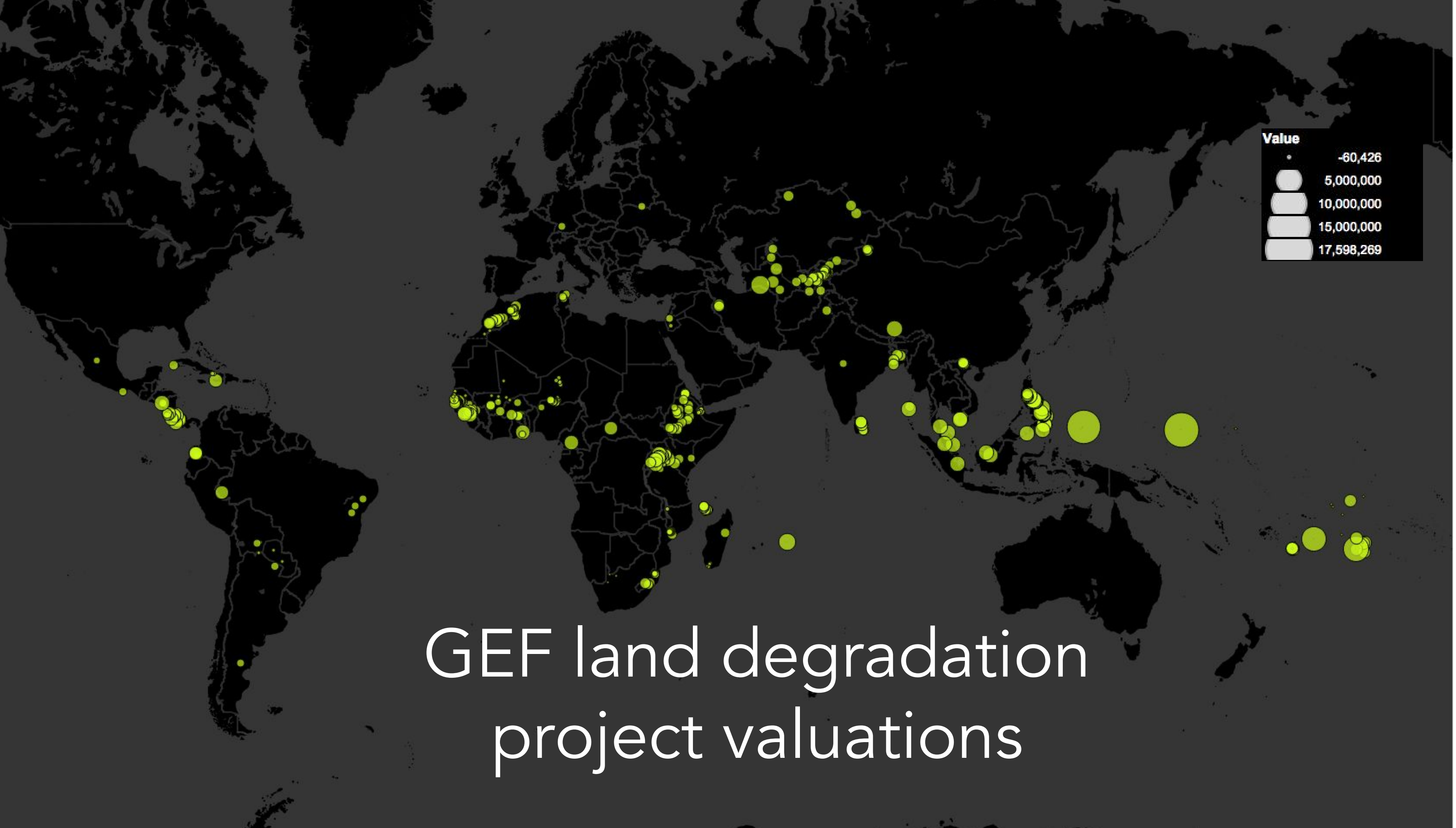


Key Findings

- Approximately 40 tonnes of carbon sequestration was attributable to GEF projects, on average per hectare.
- This resulted in approximately 100,000 tonnes of carbon per project.
- The mean valuation from the literature resulted in an estimated 7.5 million USD value of this sequestration.

Key Findings - Heterogeneity

- ~5 years was a robust breakpoint across many analyses of when impacts became most apparent.
- Projects tended to be more effective in urban or higher population density areas, though more often located in lower population areas.
- Significant heterogeneity in valuation for carbon sequestration over geographic space.
- Clear and compelling evidence for positive GEF impacts on NDVI and forest cover; limited evidence for forest fragmentation (for the Land Degradation Portfolio).



GEF land degradation project valuations

Learning from Project Success / Next Steps

- Heterogeneity in project outcomes along geographic dimensions allows for estimates of characteristics that might drive project success.
- While globally the 5 year threshold was found, many local characteristics were also positive drivers.
- For example, in some geographic contexts MFA projects outperformed SFA. While distance to roads was generally important, the distance thresholds themselves changed over different regions.

Case Study 2:
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

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

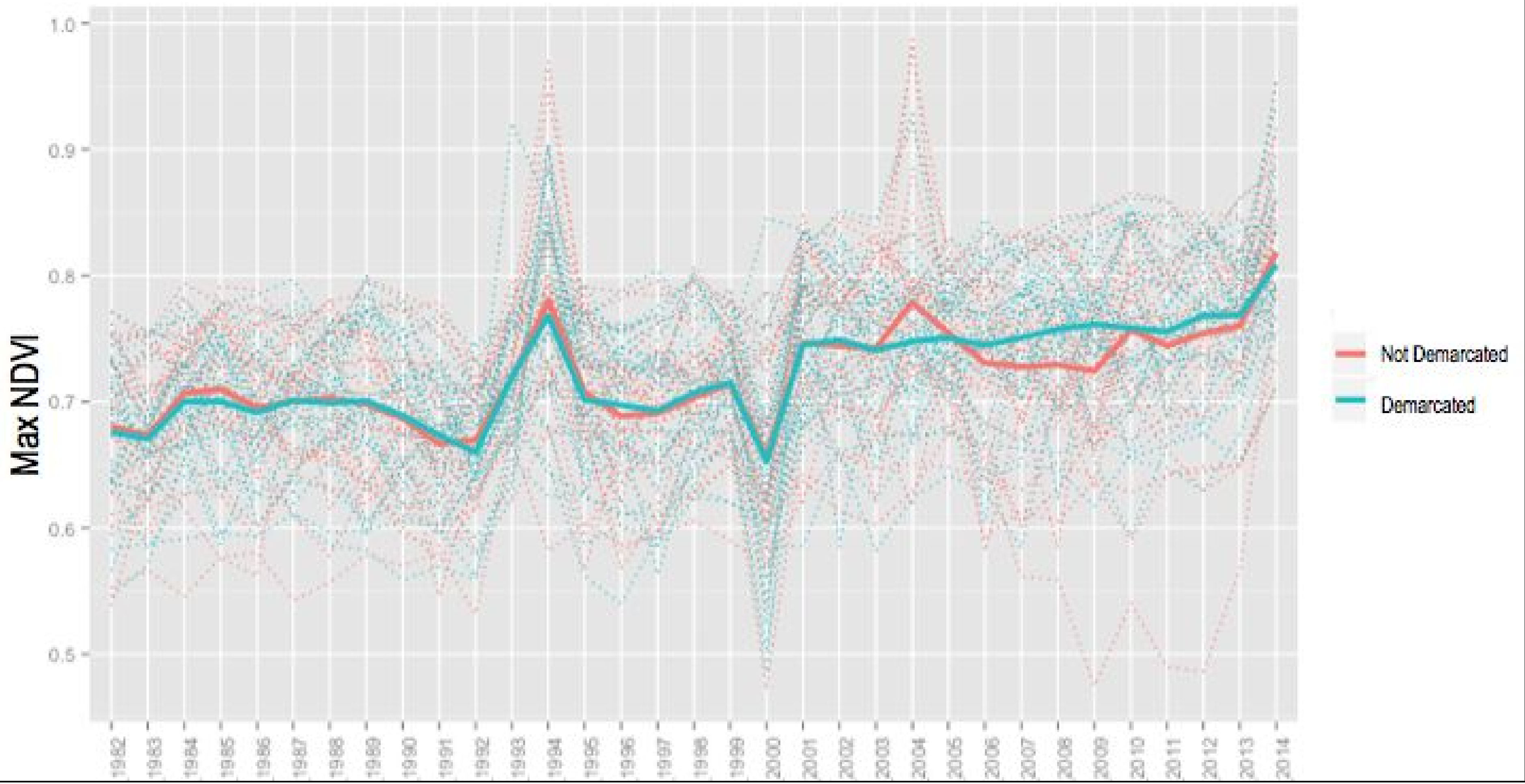
- Treatment status
  - Boundaries of community lands
  - Administrative data on demarcation dates
- Merged with satellite-based greenness measure
  - NASA Land Long Term Data Record (LTDR), 1982-2010
  - Processed to Normalized Difference Vegetation Index (NDVI)
  - Range is [0, 1] (0 = rocky, barren; 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)

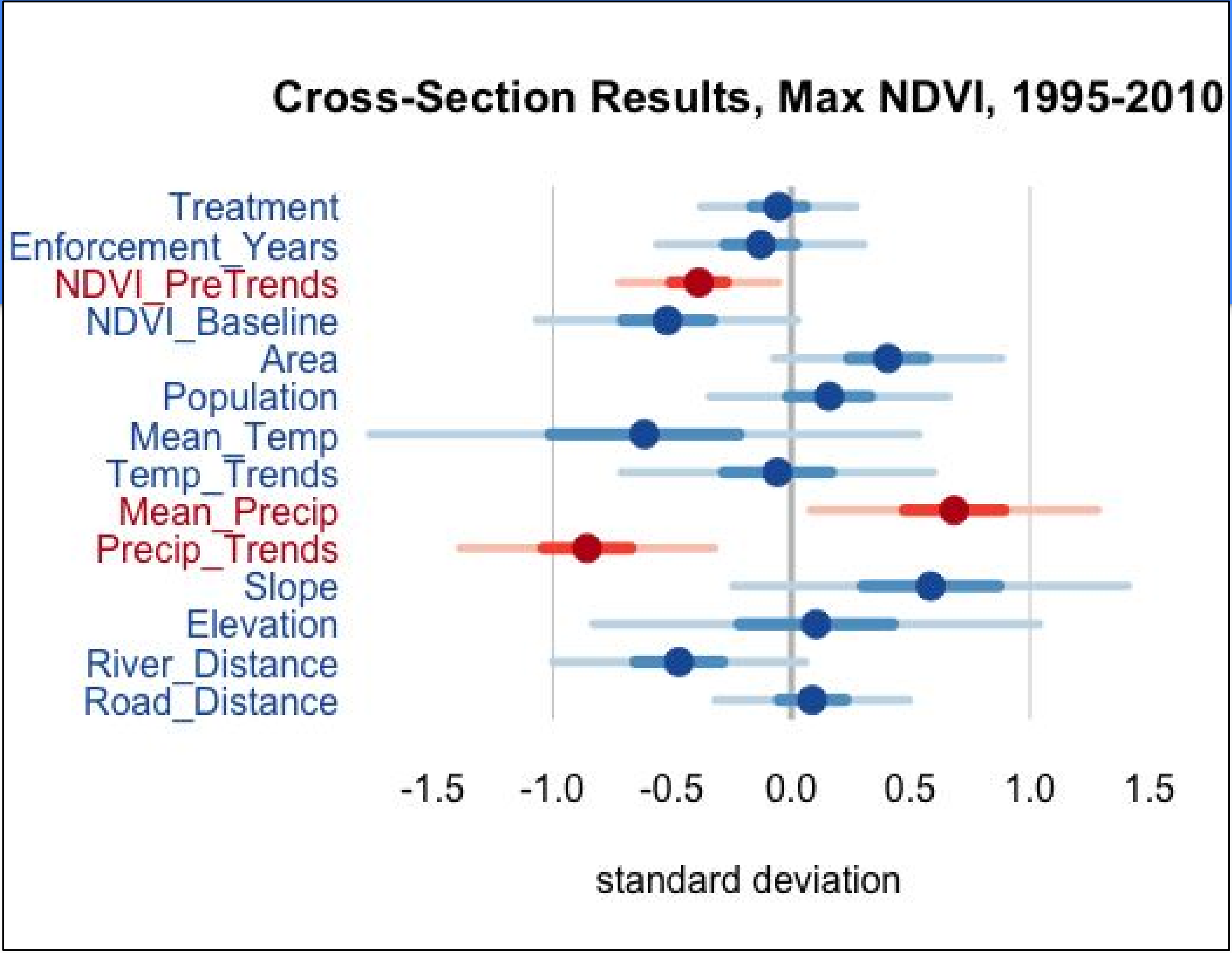
- Fixed effects
  - Control for time-invariant community unobservables
  - Treatment status at finer time intervals


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

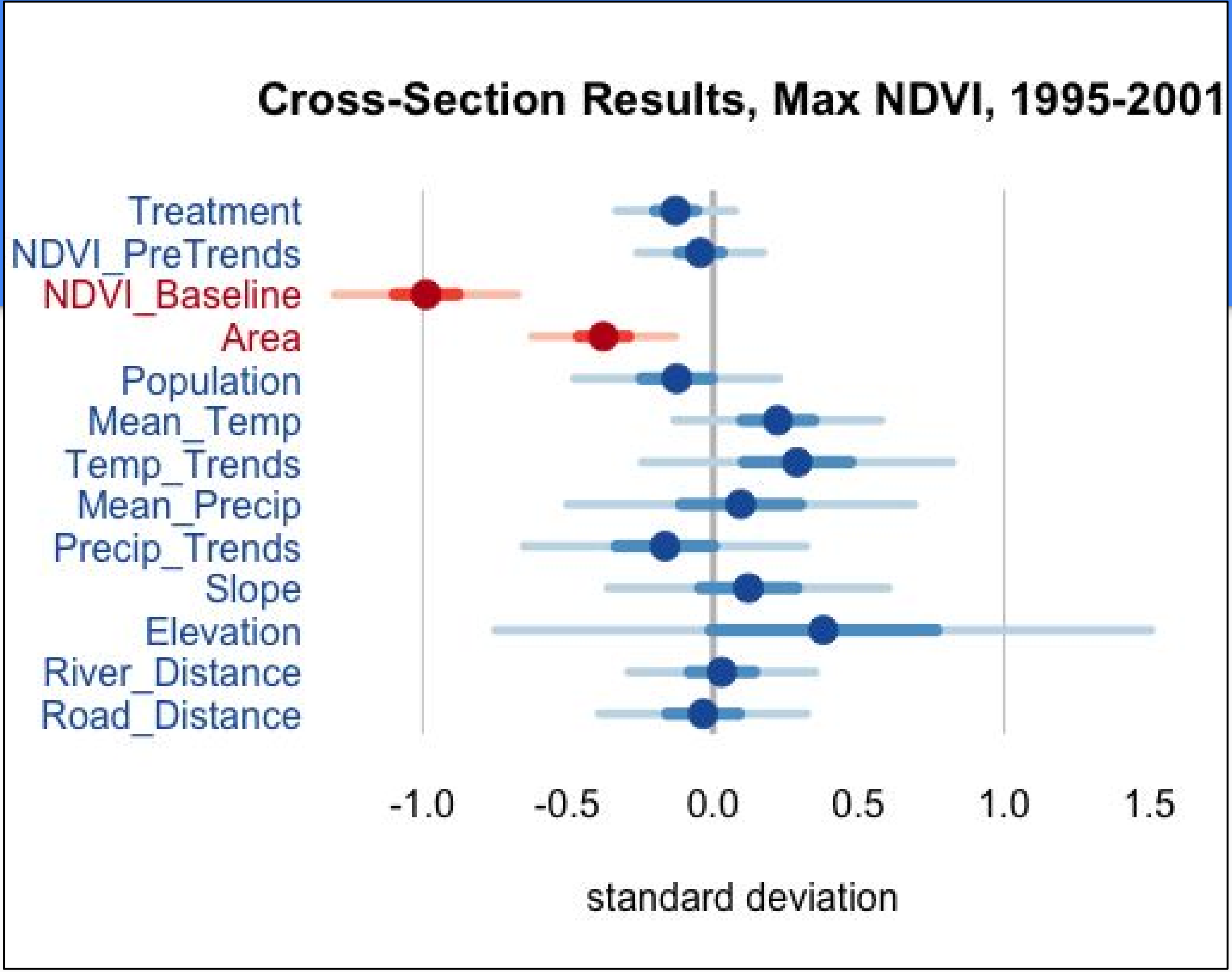
Demarcated vs. non-demarcated

- **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).



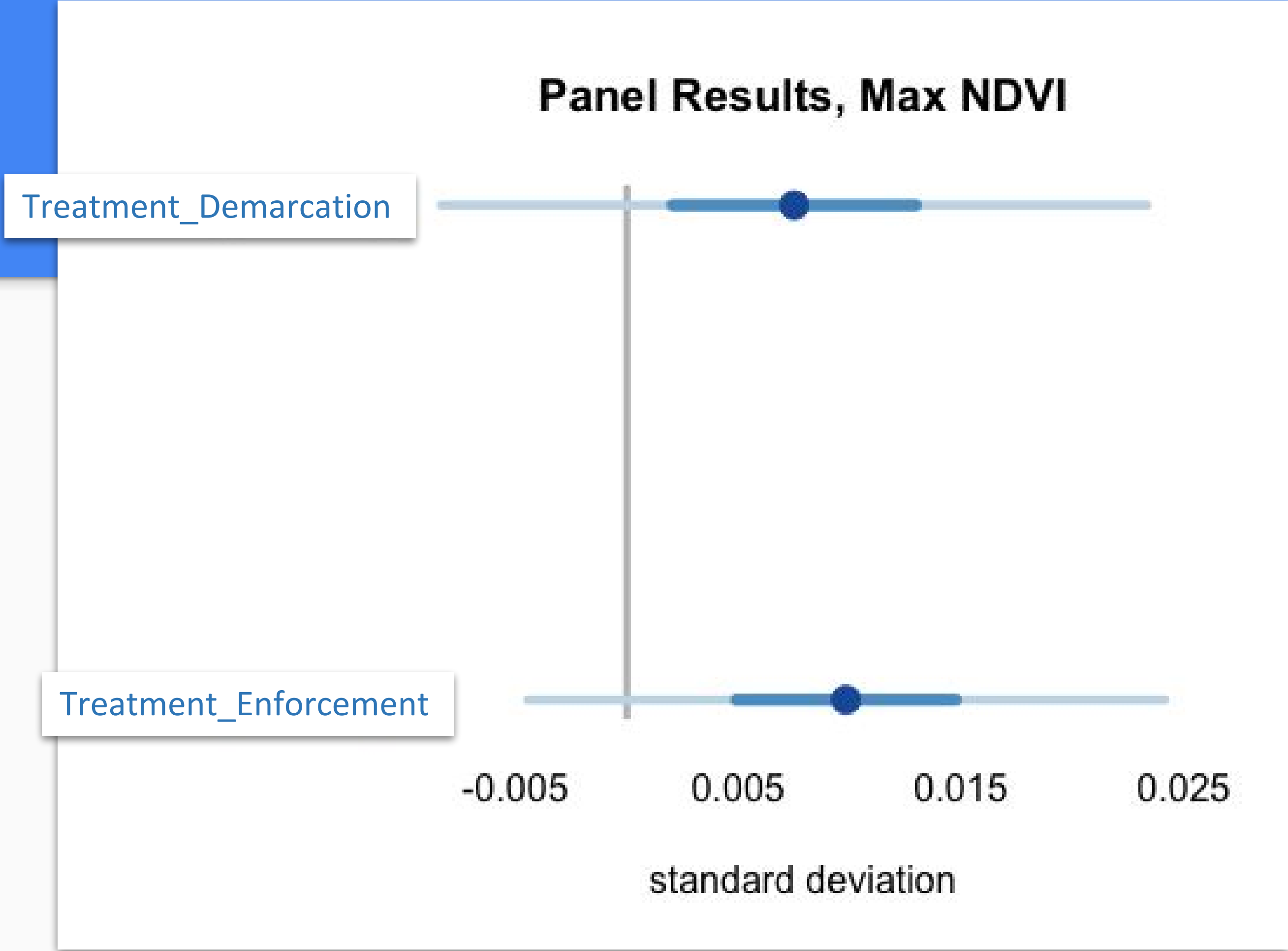
Early vs. Late demarcation

- **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).



Panel Model

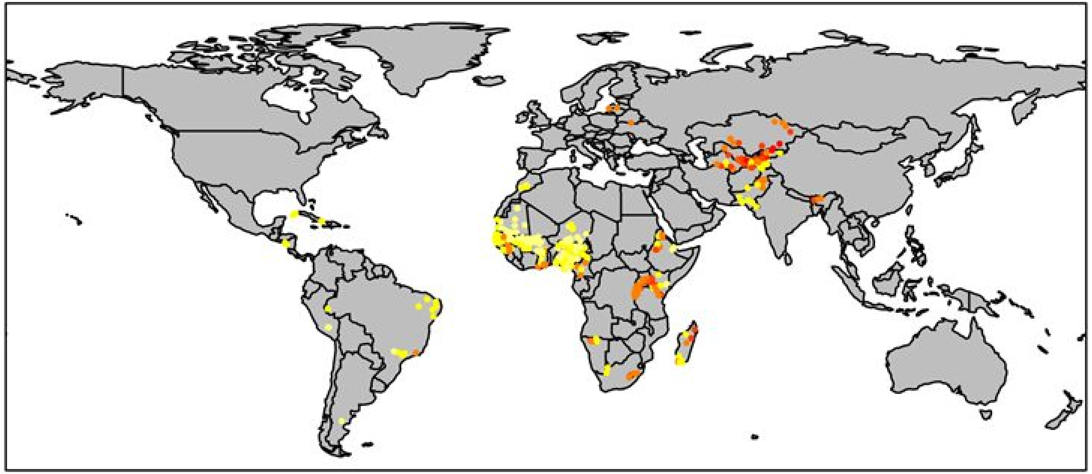
- **Treatment** = Year of demarcation.
- **Outcome** = Level of NDVI (max) each year
- Fixed effects, year trends.
- 2,128 annual observations; errors clustered by community and year.



Take-aways

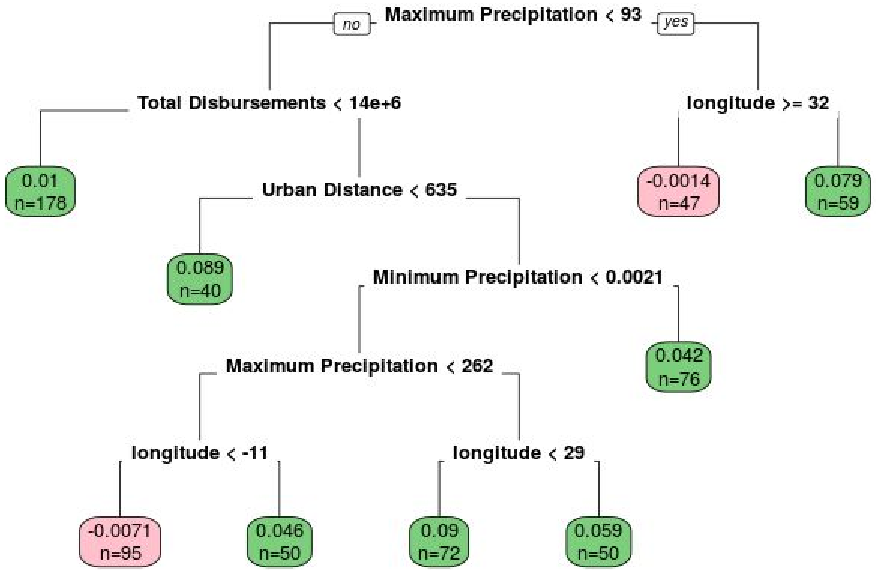
- No clear, robust evidence of differences in deforestation attributable to the PPTAL project
- Much lower rates of deforestation on indigenous lands in cross-section may not be related to land tenure status of these lands (or may be mediated through multiple, complex channels)

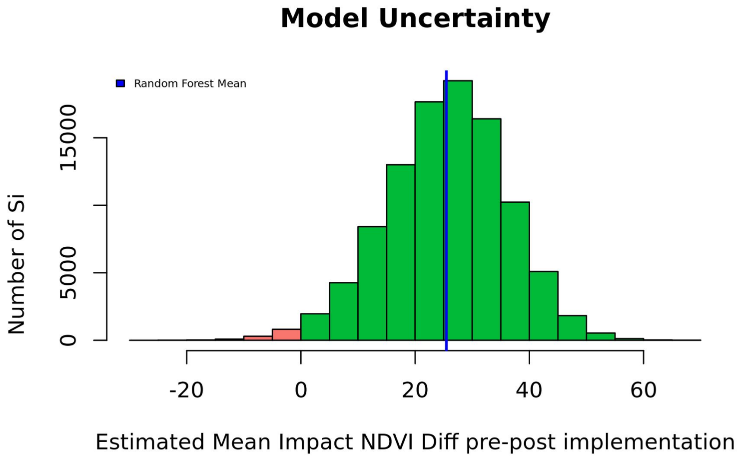
Questions



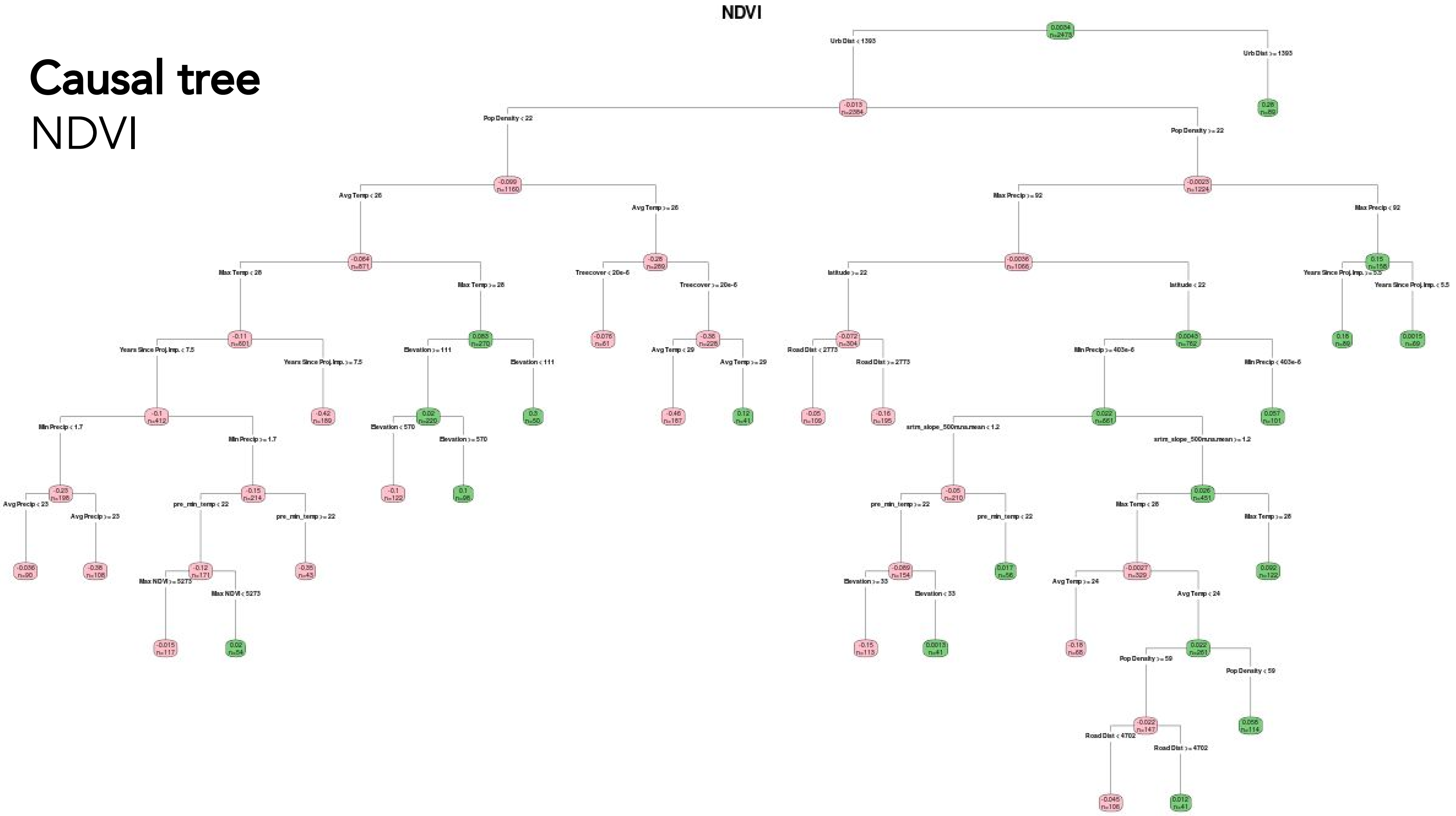
Matched Model: SFA Land (Treated), Null Case Comparisons (Control)

	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

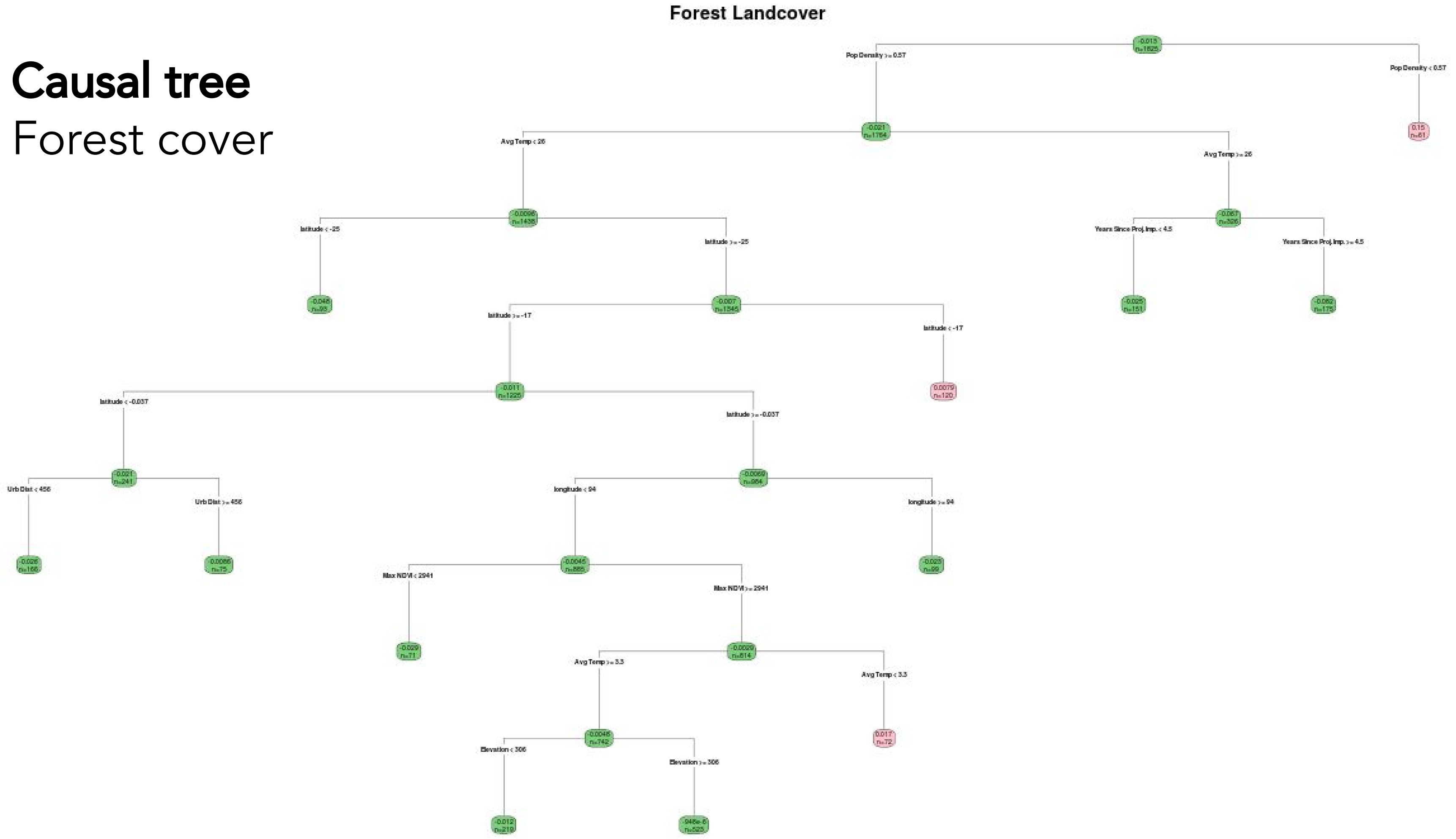

Maximum Precipitation < 93
no
yes
Total Disbursements < 14e+6
longitude >= 32
0.01
n=178
Urban Distance < 635
-0.0014
n=47
0.079
n=59
0.089
n=40
Minimum Precipitation < 0.0021
Maximum Precipitation < 262
0.042
n=76
longitude < -11
longitude < 29
-0.0071
n=95
0.046
n=50
0.09
n=72
0.059
n=50


Model Uncertainty
Random Forest Mean
Number of Si
0
5000
15000
-20
0
20
40
60
Estimated Mean Impact NDVI Diff pre-post implementation

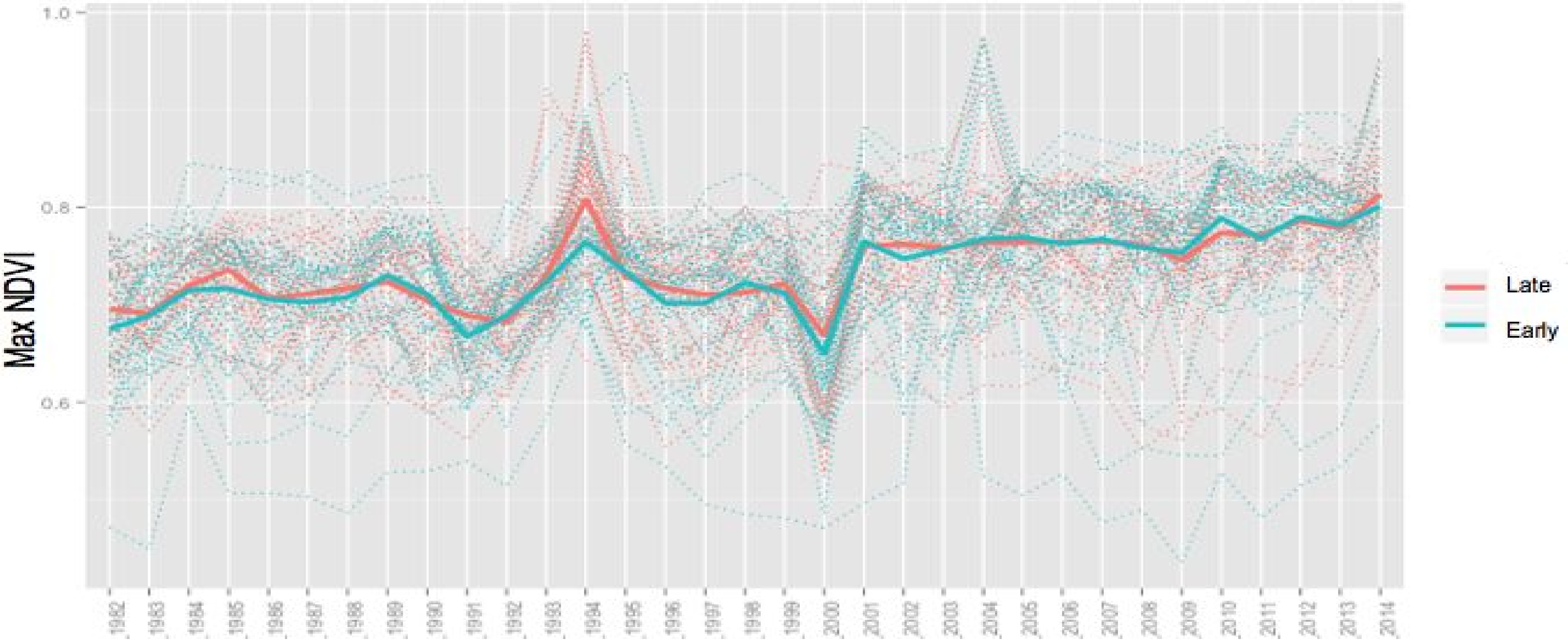
Causal tree
NDVI



Causal tree
Forest cover



NDVI trends



1.0
0.8
0.6
Max NDVI
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
Late
Early