Independent Evaluation Office
GLOBAL ENVIRONMENT FACILITY

Value for Money Analysis for GEF Land Degradation Projects

DECEMBER 2018

FULL REPORT

Global Environment Facility
Independent Evaluation Office

Value for Money Analysis for GEF Land Degradation Projects

December 2018

Evaluation Report No. 133

This report was presented to the Council in October 2016.

1818 H Street, NW, Washington, DC 20433
Internet: www.gefieo.org/; email: gefevaluation@thegef.org

ISBN: 978-1-64233-018-2

Contents

Tables

Foreword

The Value for Money Study was undertaken to assess if Global Environment Facility (GEF) investments and technical support have been optimally utilized to achieve the intended global environmental benefits through the GEF's land degradation focal area. Land degradation was established as a separate GEF focal area in 2002 during the third replenishment of the GEF, leading to immediate allocation of resources to directly combat the challenges associated with this global issue.

Despite the long-standing support of the GEF to these interventions to sustain and generate global environmental benefits such as increased forest cover and reduced forest degradation, or increased amount of carbon sequestrated, a huge gap exists in understanding the valuation of the actual monetary benefits accruing from these environmental benefits. The purpose of this analysis was to explicitly quantify the value for money of GEF land degradation projects as measured by biophysical indicators. This study integrated satellite and other spatial data on the geographic location of GEF land degradation projects, and related measurements on indicators aligned with United Nations Convention to Combat Desertification (UNCCD) indicators. These data, alongside related information on the geographic context and characteristics of GEF projects, were used in a matching-based quasi-observational study design to test hypotheses on the effectiveness of GEF projects.

The findings were presented to the GEF's 51st Council meeting in October 2016, as part of the Independent Evaluation Office's Semi-Annual Evaluation Report. The Council was appreciative of the evaluation and noted that similar value for money analysis in other evaluations of the GEF focal areas would make a strong case for a robust replenishment. The findings of this study also contributed to the GEF report to the UNCCD 13th Conference of the Parties, including side events in Ordos, China.

Juha I. Uitto
Director, GEF Independent Evaluation Office

Acknowledgments

Anupam Anand, Evaluation Officer with the Independent Evaluation Office (IEO) of the Global Environment Facility (GEF), led this evaluation. The value for money analysis was conducted at AidData, College of William and Mary, by Daniel Runfola, Lead Scientist. The team also included Ariel Ben, Yishay, Lead Economist; Seth Goodman, Data Engineer; and Zhonghui Lv, GIS Analyst. Jyotheshwar Nagol, Assistant Research Professor, University of Maryland, helped with the forest fragmentation analysis. The work was further supported by the AidData Geocoding Team and the High-Performance Computing Group at the College of William and Mary.

The evaluation benefited from guidance and oversight provided by Juha Uitto, Director of the IEO. Supervision and quality assurance were provided by Geeta Batra, IEO Chief Evaluation Officer. Administrative support was provided by Evelyn Chihuguyu, Program Assistant, Malac Kabir, Research Assistant, Marie-Constance Manuella Koukoui, Senior Executive Assistant. Charles Hagner edited the report, and Nita Congress designed and laid out the publication.

The team is grateful for the information, data, and insights provided by the United Nations Convention to Combat Desertification Secretariat, the GEF Secretariat, and all the GEF Agencies. The team acknowledges the support provided by NORAD to develop the value for money approach.

The GEF IEO is grateful to all of these individuals and institutions for their contributions. Final responsibility for this report remains firmly with the Office.

Abbreviations

GEF	Global Environment Facility
NDVI	Normalized Difference Vegetation Index
UNCCD	United Nations Convention to Combat Desertification

The GEF replenishment periods are as follows: pilot phase: 1991–94; GEF-1: 1995–98; GEF-2: 1999–2002; GEF-3: 2003–06; GEF-4: 2006–10; GEF-5: 2010–14; GEF-6: 2014–18; GEF-7: 2018–22.

All dollar amounts are U.S. dollars unless otherwise indicated.

Executive summary

The Value for Money (VFM) Study was undertaken to assess if Global Environment Fund (GEF) investments and technical support have been optimally utilized to achieve the intended global environmental benefits through the land degradation focal area of the GEF. Land degradation was established as a separate GEF focal area in 2002 during the third replenishment of the GEF, leading to immediate allocation of resources to directly combat the challenges associated with this global issue. Despite the long-standing support of the GEF to these interventions to sustain and generate global environmental benefits such as increased forest cover and reduced forest degradation, or increased amount of carbon sequestrated, a huge gap exists in understanding the valuation of the actual monetary benefits accruing from these environmental benefits.

The purpose of this analysis was to explicitly quantify the VFM of GEF land degradation projects as measured by biophysical indicators. This study integrated satellite and other spatial data on the geographic location of GEF land degradation projects, and related measurements on indicators aligned with the United Nations Convention to Combat Desertification (UNCCD) indicators. These data, alongside related information on the geographic context and characteristics of GEF projects, were used in a matching-based quasi-observational study design to test hypotheses on the effectiveness of GEF projects.

This study brought together economists, computer scientists, and geographers with expertise in remote sensing and impact evaluation to apply a VFM assessment of GEF land degradation projects. Leveraging methodological approaches to causal identification that has not previously been applied to land degradation, this study explicitly quantified (1) the causally identified impact attributable to GEF land degradation project locations using three indicators capturing vegetation productivity, forest fragmentation, and forest cover change; and (2) the VFM resulting from the impacts of GEF land degradation projects in terms of carbon sequestration.

The study applied a six-step procedure:

- Precise geospatial data on GEF land degradation project locations (i.e., every site at which a project operated) was generated in compliance with the International Aid Transparency Initiative standard.
- Satellite information was used to derive long-term measurements of each of the three outcomes being assessed at each geographic location (following UNCCD 2015 guidance on indicator selection and 2014 guidance from the GEF Scientific and Technical Advisory Panel on measurement.
- The data generated in the preceding steps were integrated with a wide set of geographically varying ancillary data (i.e., nighttime lights,

population, distances to roads and rivers) to enable a match of GEF land degradation project locations to control locations where no intervention occurred.

- A novel propensity score matching approach, causal trees, was employed to examine the impact of GEF land degradation project locations on each indicator of interest.
- Observed patterns between these indicators and carbon sequestration were used to estimate the contribution of each project location in terms of tons of carbon sequestered.
- A value transfer approach was applied alongside an interactive, online prototype tool to enable users to valuate individual project locations.

Key findings

Overall global positive impact. Evidence from this analysis suggests that GEF land degradation and biodiversity projects have had a global net positive impact on both forest cover and vegetation productivity—as per the Normalized Difference Vegetation Index (NDVI)—with valuations in terms of carbon sequestration and soil retention ranging from $62 to $207 per hectare affected.

Impacts vary considerably. Considerable heterogeneity exists in the absolute impact of GEF projects:

- A lag time of 4.5–5.5 years was an important inflection point at which impacts were observed to be larger in magnitude.
- Projects with access to electricity tend to have some of the largest relative positive impacts.
- The initial state of the environment is a key driver in GEF impacts, with GEF projects tending to have a larger impact in areas with a poor initial condition.
- Projects in Africa and Asia had generally positive impacts on average. Projects in Latin America and the Caribbean, North and South America, and Oceania all had positive impacts on all three indicators.

Conclusion

This study sought to estimate the VFM resulting from GEF projects implemented in the land degradation focal area. Findings suggest that the GEF has, globally, been effective in improving environmental conditions both through an increase in vegetation productivity as well as a reduction in the rate of forest cover loss. Critically, this study suggests that the local context in which programs are implemented can be assessed for suitability of interventions. By examining where projects have historically worked—or failed—better decisions as to how to site and fund projects in the future can be made.

This study represents the first step along this path, and provides general guidance to implementers regarding the contexts in which GEF projects have been most successful. The evidence presented further highlights that assessing the geospatial contexts in which projects might be placed before their implementation can result in stronger positive outcomes. By targeting funds at locations that have both the poor initial conditions and geographic characteristics for which GEF project implementations are known to provide strong outcomes, better outcomes can be achieved.

1: Introduction

1.1 Background and objective

The land degradation focal area of the Global Environment Facility (GEF) established as a stand-alone focal area during GEF-3 combines the principles of sustainable land management[1] and integrated natural-resource management[2] to maximize the global environmental benefits of combating land degradation. Since GEF-3, the GEF has supported 618 land degradation projects or multifocal area projects with a land degradation component. In dollar terms, the land degradation single focal area projects account for 5 percent of total GEF Trust Fund utilization from GEF 3 to GEF-6, which increased to 9 percent during GEF-6. Recently, there has also been an increase in land degradation multifocal area projects intended to deliver environmental benefits in multiple focal areas.

[1] According to the World Bank (2008), "Sustainable land management is a knowledge-based procedure that helps integrate land, water, biodiversity, and environmental management (including input and output externalities) to meet rising food and fiber demands while sustaining ecosystem services and livelihoods."

[2] As defined by Sayer and Campbell (2004): "Integrated Natural Resource Management is a conscious process of incorporating the multiple aspects of resource use into a system of sustainable management to meet the goals of resource users, managers and other stakeholders (e.g., production, food security, profitability, risk aversion and sustainability goals)."

With the investments and commitment to combat land degradation at scale, it becomes imperative to assess how efficient and effective these interventions are. The purpose of this value-for-money study was twofold—first, from the evaluation standpoint, this study attempts to inform us about the relevance of GEF land degradation support and whether interventions are in the areas facing desertification and land degradation. It also helps understand the impact and the contextual factors associated with the impact. Finally, it answers the efficiency and effectiveness questions by helping to quantify the value for money in terms of ecosystem services. The second purpose was to identify the causal impacts from land degradation focal area projects along three land degradation indicators that closely relate to the indicators suggested by the United Nations Convention to Combat Desertification (UNCCD's) 2015 Land Degradation Neutrality scientific framework and the proposed indicators and subindicators for the Sustainable Development Goals 15.

This report also responds to UNCCD guidance.[3] Contained in this report, and made available for future analysis, is information on the geographic

[3] United Nations Convention to Combat Desertification (UNCCD) 10-year (2008–18) strategy, para. 20 "Accounting for Land Degradation Focal Area investments in a spatially quantifiable manner will foster a more accurate picture of GEF's contribution to combating land degradation globally."

location of GEF land degradation projects, as well as related measurements following the indicators suggested in the monitoring framework of the UNCCD for measuring land degradation.

As stated in paragraph 6 of the same UNCCD strategy document, "An important aspect of linking the Strategies is therefore related to the outcomes, impacts and associated indicators, all of which serve to inform project design by all stakeholders. Annex 1 of the report is an attempt to link the expected impacts (and proposed indicators) of the UNCCD strategic objectives with the results-based management framework of the GEF Land Degradation Focal Area." This report therefore presents an operationalization of this objective, building on the project-based reporting available to date by extending such analyses to individual project locations.

However, the information—or approach—presented in this report to drive project location-level decision making should be applied with extreme caution without coupled, "bottom-up" analyses. Overall, the findings of this value-for-money analysis indicate that, on average, GEF land degradation projects have mitigated or reversed negative land degradation processes; there is also significant heterogeneity in these findings—issues that could be addressed through more in-depth inquiry in the future.

1.2 Summary

This analysis brings together economists, computer scientists, and geographers with expertise in remote sensing and impact evaluation to apply a value for money assessment to the case of GEF land degradation projects. Leveraging methodological approaches to causal identification that have not previously been applied to the study of land degradation, this report explicitly quantifies (1) the causally-identified impact attributable to GEF land degradation project locations using three indicators (capturing vegetation productivity, forest fragmentation, and forest cover change), and (2) the value for money resultant from these impacts of GEF land degradation projects in terms of carbon sequestration.

A six-step procedure is applied, in which (1) precise geospatial data on GEF land degradation project locations (i.e., every site at which a project operated) is generated in compliance with the International Aid Transparency Initiative standard; (2) satellite information is used to derive long-term measurements of each of the three outcomes being assessed at each geographic location (following UNCCD 2015 guidance on indicator selection) and GEF Scientific and Technical Advisory Panel 2014 guidance on measurement; (3) the data generated in steps 1 and 2 is integrated with a wide set of geographically-varying ancillary data (i.e., nighttime lights, population, distances to roads and rivers) to enable the match of GEF land degradation project locations to "control" locations where no intervention occurred; (4) a novel propensity score matching approach, causal trees, are employed to examine the impact of GEF land degradation project locations on each indicator of interest; (5) observed patterns between these indicators and carbon sequestration are used to estimate the contribution of each project location in terms of metric tons of carbon sequestered; and (6) a value transfer approach is applied alongside an interactive, online, prototype tool to enable users to valuate individual project locations alongside a presentation of reference values found in the literature.

The novel methodology leveraged in this approach, more regularly applied in industries, enable recommendations regarding the spatial contexts in which GEF land degradation projects result in positive outcomes. This is resultant from the combination of geographic information system methods, which enable long-term data from satellite sensors; econometric methods, which enable causal inference and identification of impacts; and computer science methods, which enable the detection of heterogeneity in impacts across different spatial contexts.

This report identifies a global positive impact of GEF land degradation projects along all three indicators examined, but also finds considerable heterogeneity in these impacts across different geographic contexts. Key findings included the following:

- A lag time of 4.5 to 5.5 years was an important inflection point at which impacts were observed to be larger in magnitude, noting some projects were still under implementation.
- The initial state of the environment is a key driver in GEF impacts, with GEF land degradation projects tending to have a larger impact in areas with a poor initial condition.
- Projects located in Africa and Asia had generally positive impacts on average, except in the case of forest fragmentation. Projects in Latin America and the Caribbean, North and South America, and Oceania all had positive impacts on all three indicators.

Across the entire globe, within 25 kilometer catchment areas, GEF land degradation projects (1) increased Normalized Difference Vegetation Index (NDVI) by approximately 0.03 (relative to an average NDVI of 0.55), (2) reduced forest loss by 1.3 percent (relative to a global mean of 2.4 percent forest loss in all areas), and (3) increased the average size of forest patches by 0.25 square kilometers (relative to a global mean of 7.3 square kilometers). The estimated carbon sequestered by the GEF was—on average—43.52 tC/ha. This equates to an estimated 108,800 tC sequestered by each GEF land degradation project location.[4]

Across the 8,093 valuations of carbon identified as a part of the value transfer approach (Costanza et al. 2014) employed to estimate project location valuations (deflated to 2014), a median dollar value of $12.90/metric ton was identified, based on academic, industrial, and government reports. Using this median dollar value, we estimate that GEF land degradation projects[5] contributed $7.5 million (2014) on average to sequestration alone—well above the average cost of most GEF land degradation projects ($4,182,887). Following these findings, this report offers two suggestions for consideration:

- In keeping with the joint goal of GEF and the UNCCD to promote the "Development of improved methods for multi-scale assessment and monitoring of land degradation trends, and for impact monitoring of GEF investment in SLM [sustainable land management]," we recommend the use of the top-down learning-based approach detailed in this document as an initial screening tool for project planning. By identifying the geographic contexts in which similar projects have historically succeeded—and failed—appropriate safeguard and mitigation efforts can be put in place a priori.

[4] This estimate is based solely on the additive impact of GEF land degradation projects on additional sequestration —i.e., the total metric tons that were sequestered due to each GEF project that otherwise would not have been sequestered. This only includes estimates of gains due to changes along the three indicators examined—forest fragmentation, NDVI, and forest land cover, and thus may not represent the full envelope of all sequestration that is attributable to GEF projects.

[5] Although the impact at each individual project location is calculated in this document, costs are only known at the project level. Thus, to calculate the average project valuation we aggregate each project's location valuation estimates.

- Echoing the joint UNCCD-GEF statement that "Accounting for Land Degradation Focal Area investments in a spatially quantifiable manner will foster a more accurate picture of GEF's contribution to combating land degradation globally," we recommend the ongoing collection of exact geographic information (latitude and longitude or geographic shape) of GEF land degradation activities. By providing exact geographic information on GEF land degradation project locations, it is possible to leverage decades of satellite and other spatial information in ways that is not otherwise possible.

1.3 Definitions and frame of analysis

The impact of GEF land degradation projects are examined along multiple indicators to capture fluctuations in natural capital, following the indicators suggested in the monitoring framework of the UNCCD for measuring land degradation (UNCCD 2015) and the GEF Scientific and Technical Advisory Panel guidance of 2014. This analysis is implemented with two-tier 1 metrics to examine impacts on land cover change (metrics of forest fragmentation and forest cover), as well as two-tier 2 metrics (vegetation productivity, carbon stocks). These are defined and discussed more extensively in annex A.

Each of these measurements is calculated with the following procedures for each geocoded GEF land degradation project:

- **Vegetation productivity.** The yearly maximum productivity for each GEF land degradation project is calculated on an annual basis from 1985 to 2015 using the Long Term Data Record NDVI product. Periods prior to GEF land degradation project implementations are used to calculate baseline trends and levels, whereas contemporary data is used to establish impacts.
- **Forest cover change.** The Hansen et al. (2013) tree cover product from University of Maryland is employed to detect forest cover change. These products are available at 30-meter resolution for 2000, and on a yearly basis for years 2001 to 2015. The absolute annual change in tree cover is calculated post-2000, whereas a baseline is calculated using the data from years prior to 2000. Additionally, data available from the Global Land Cover Facility at University of Maryland for forest change data for 1980, 1990 (Kim et al. 2014) were used.
- **Forest fragmentation.** Within the area of influence calculated for each GEF land degradation project, a regionally-varying threshold is applied to the percent tree cover. This produces a binary forest (denoted by 1) versus nonforest (denoted by 0) cover map for each time period for which forest cover change information is available. For each GEF land degradation project, the level of forest fragmentation is then calculated for each time period. For this analysis, the average patch size is used as a summary metric for fragmentation.
- **Carbon stocks and sequestration.** Using the above products, Ecofloristic Zone Carbon Fractions data set derived by the Oak Ridge National Laboratory is leveraged to estimate carbon stocks. Although these estimates have an inherent measurement error, the combination of field-based estimates and remote sensing techniques has become the primary method of examining carbon stocks and carbon sequestration (Asner et al. 2010), due to difficulties with solely field-based estimates (Saatchi al. 2011).

Following the broad scope of this assessment, as many GEF land degradation project locations as feasible are included in the analysis frame. To accomplish this goal, the GEF land degradation projects spanning from January of 2002 until January of 2014 were geocoded in compliance with the International Aid Transparency Initiative standards. These 202 projects have 1,704 project locations associated with them (figure 1.1); of these 1,704 this report focuses on 446 for which exact geographic information is available—i.e., the latitude and longitude at which the project was executed is known with a high degree of precision (figure 1.2).

FIGURE 1.1 The location of all geocoded GEF land degradation projects

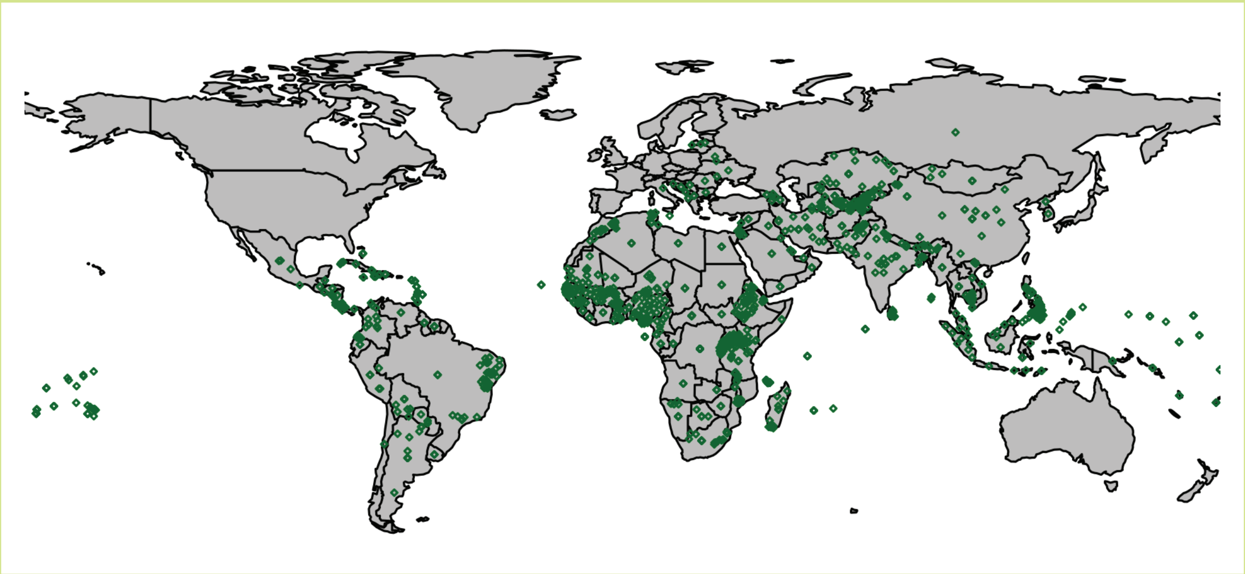
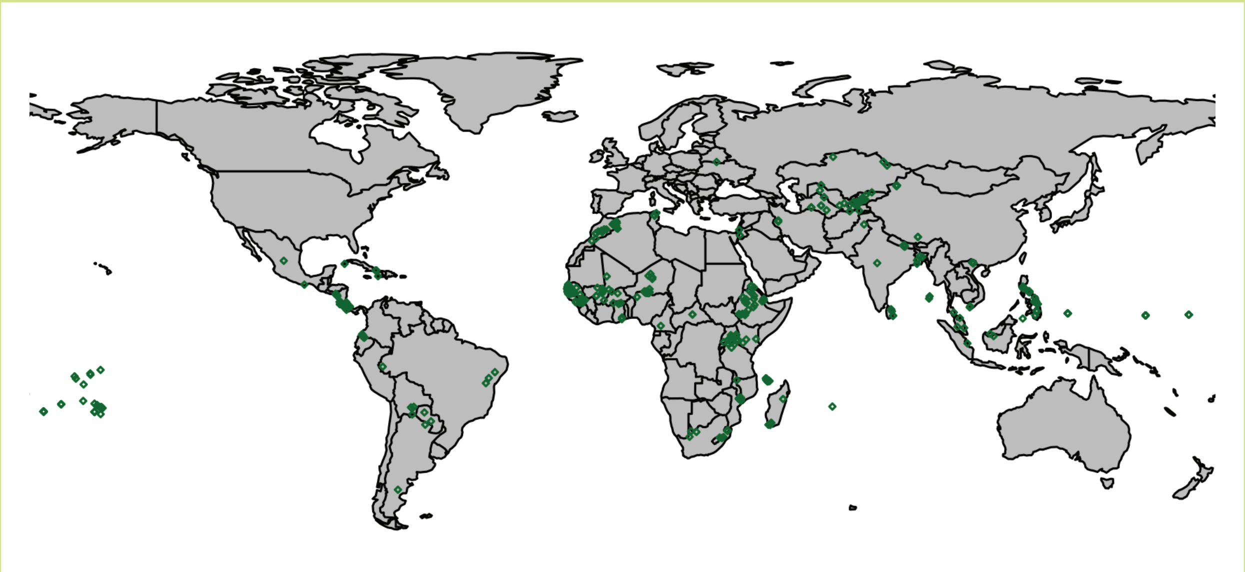


FIGURE 1.2 The location of geocoded GEF land degradation projects known with a high degree of precision



In addition to the measured locations of GEF land degradation projects, thousands of potential control cases are created in areas proximate to GEF activities, but which contained no known interventions. The geographic area from which control cases were selected are shown in figure 1.3. Eligible control locations were limited to be no further than 500 kilometers from an existing GEF land degradation project in order to provide better potential matches, but were limited to be no closer than 50 kilometers to minimize potential spillover effects.

For each GEF land degradation project location and eligible control site, the outcome metric of vegetation productivity, forest cover change, and forest fragmentation are calculated. Baseline trends and levels for each of these metrics are calculated by identifying the pre-intervention time period for each GEF land degradation project location. To further facilitate matching (for example, to ensure GEF land degradation project locations far from urban areas are compared with comparable areas), a variety of covariate information is retrieved for each location, summarized in table 1.1.

FIGURE 1.3 Locations eligible to become a control for comparison

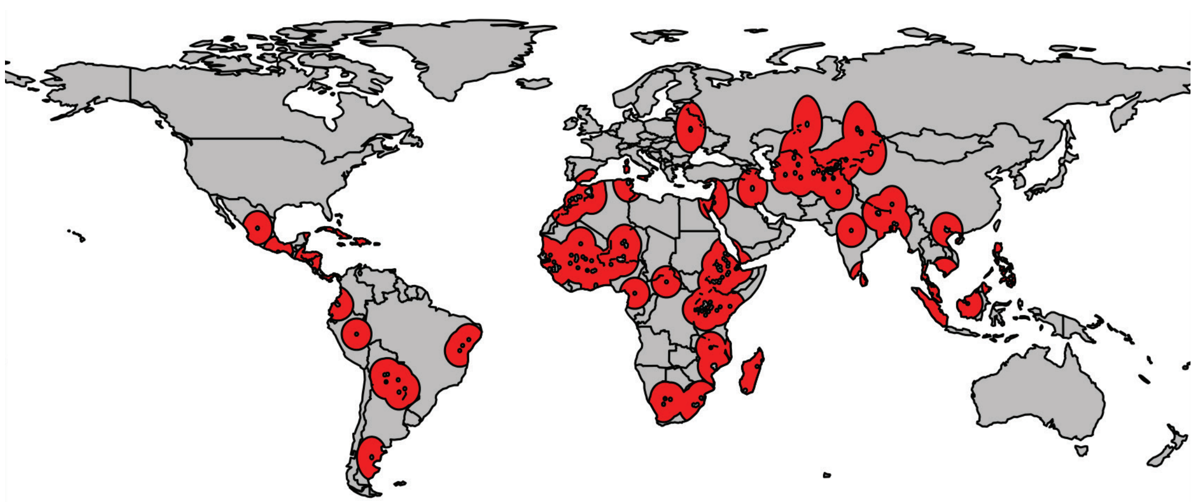


TABLE 1.1 Key covariate data sources

Domain	Source	Attribute	No. of observa-tions	Current coverage: Temporal	Current coverage: Spatial	Spatial resolution
Human development	Defense Meteorological Satellite Program – Operational Linescan System (DMSP-OLS) Visible Infrared Imaging Radiometer Suite (VIIRS)	Nighttime lights	n.a.	1992–2016	Global	Grid cell (1km; 250m)
	Global Roads Open Access Data Set (gROADS)	Road networks	n.a.	1980–2010	Global	Grid cell (~1km)
Political	World Database on Protected Areas (WDPA)	Environmental protection areas	220,453	2015	Global	Variable
Demography	Gridded Population of the World (GPW)	Population	n.a.	1990–2020 every 5 years	Global	Grid cell (5km / 1km)
Environment and natural resources	Hydrological data and maps based on Shuttle Elevation Derivatives at multiple Scales (HydroSHEDS)	River networks	n.a.	1995–2005	Global	Grid cell (~1km)
	Shuttle Radar Topography Mission (SRTM)	Elevation/slope	n.a.	2000	Global	Grid cell (500m)
	University of Delaware	Air temperature	n.a.	1900–2014	Global	Grid cell (50km)
		Precipitation	n.a.	1900–2014	Global	Grid cell (50km)

NOTE: n.a. = not applicable. For raster data sets, see spatial resolution for a more accurate depiction of measurement density.

2: Methods

Six different causal models are estimated, employing different counterfactuals and modeling approaches summarized in table 2.1. Each model estimates the impact of GEF land degradation project locations on a single indicator: (Q1) NDVI, or vegetation productivity; (Q2) forest land cover; and (Q3) the fragmentation of forests. Two different modeling approaches are used. In Case 1, each GEF land degradation project is buffered by 25 kilometers, and information is aggregated to those buffers. These 25 kilometer buffers are then compared with randomly distributed 25 kilometer buffers that did not contain a GEF land degradation project (all controls are limited to areas within 500 kilometers of GEF land degradation projects, but not less than 50 kilometers distant), and a causal tree is fit to estimate (1) the overall impacts of GEF land degradation projects, and (2) the geographic heterogeneity in these impacts. In Case 2, the same 25 kilometer treatment and control units are used in a random forest approach, in which 10,000 trees are fit with varying subsets of the data. This serves as a robustness check on the findings in Case 1, reporting both the robustness of each individual project location estimate as well as the robustness of the key indicators identified in the causal tree. More information on these approaches is provided in annexes A and B.

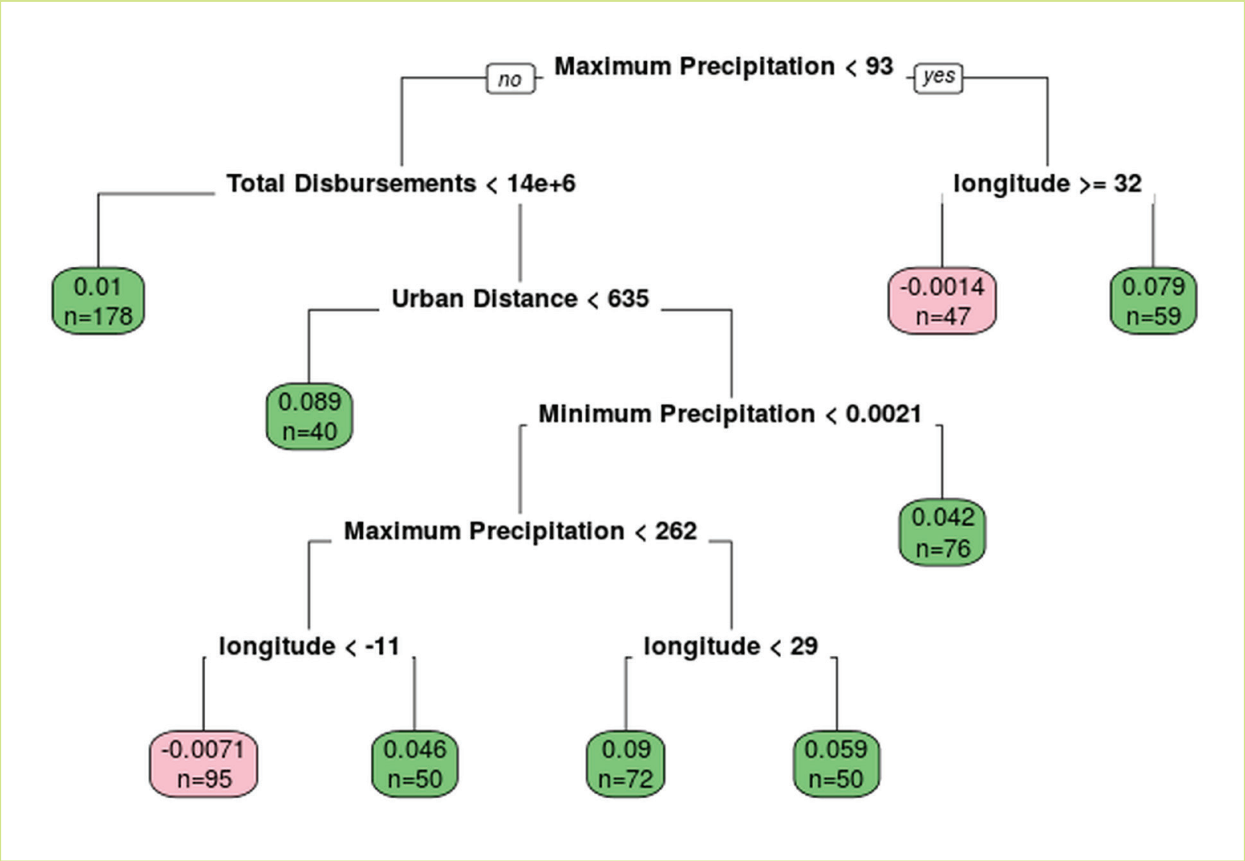
2.1 Causal model

Recent work has illustrated that—with key adjustments (e.g., figure 2.1)—tree-based approaches can be used to identify how the causal effects of an intervention (i.e., international aid; a medical treatment) vary across key parameters (such as geographic space; see Athey and Imbens 2015; Staff 2014; Shen et al. 2016). This is key for top-down, or global- scope analyses, as it is unlikely that aid projects will have the same effect across highly variable geographic contexts, and

TABLE 2.1 **Summary of conducted analyses**

Indicator	Case 1: 25 km buffer causal tree	Case 2: 25 km buffer random forest
Q1. Impact of GEF land degradation projects on vegetation productivity	Unit of observation: 25km Buffers Outcome metric source: long-term data record	Unit of observation: Watershed Outcome metric source: long-term data record
Q2. Impact of GEF land degradation projects on forest land cover	Unit of observation: 25km Buffers Outcome metric source: Hansen	Unit of observation: 25km Buffer Outcome metric source: Hansen
Q3. Impact of GEF land degradation projects on forest fragmentation	Unit of observation: 25km Buffers Outcome metric source: global land cover facility	Unit of observation: 25km Buffer Outcome metric source: global land cover facility

FIGURE 2.1 Illustrative example causal tree



the drivers of such variation may not be known. A detailed explanation of this approach is included in annex B, whereas figure 2.1 shows an example drawn from exploratory research in which a causal tree is applied to a limited subset of international aid, examining aid's impact on a maximum observed NDVI value.

This figure serves as an illustrative example of the outputs of causal tree-based approaches to identifying how impact effects may differ across a data set. Within each terminal node in figure 2.1, the difference between a weighted outcome of all treated cases (areas that received aid) is contrasted to control cases (areas that did not receive aid), and the value displayed can be directly interpreted as the causal impact of the treatment (in this example, the presence of aid) on the metric of interest (i.e., NDVI). The *n* presented in each node represents the number of units that were used to calculate this value. At each step of the tree, a statement (i.e., "Maximum Precipitation < 93 mm") is tested as true or false for each observation, and the impact of a given observation can be determined by identifying where it falls in the tree. As a simple example, the tree in figure 2.1 would provide evidence that international aid projects located in areas with a maximum yearly precipitation greater than 93 millimeters, which provide less than $1.4 million of aid, and are further than approximately a kilometer (635 meters) from an urban area tend to increase NDVI by 0.089. Further information on how these robustness checks are conducted is included in annex D.

2.2 Estimating carbon sequestration

Once an estimated impact is generated for each GEF land degradation project for all three outcomes (vegetation productivity, forest cover, and fragmentation), an additional modeling step is employed to estimate how these impacts will modify carbon stocks at each GEF land degradation project location. The two data sources used for this are the National Carbon Storage data set[1] and the United Nation's Intergovernmental Panel on Climate Change Tier-1 Global Biomass Carbon Zones (Carbon Dioxide Information Analysis Center).[2] Using these data sets, equation 2.1 is estimated across all GEF land degradation projects:

$$CS = \beta_0 = \beta_1 * NDVI + \beta_2 * ForestCover + \beta_3 * Fragmentation + D_{cz} + (D_{cz} * NDVI) + \varepsilon_i \qquad \textbf{Eq. 2.1}$$

where *CS* is the metric tons of carbon sequestered, *NDVI* is the NDVI measurement, and D_{cz} is a fixed effect for each carbon zone. This model is then used in conjunction with the treatment impact estimated for each GEF land degradation project to estimate—on a per location basis—the absolute impact in terms of sequestered carbon for each location. Some of the inherent limitations and advantages of this approach are highlighted in the Discussion.

[1] National Aeronautics and Space Administration's Jet Propulsion Laboratory; http://carbon.jpl.nasa.gov/.

[2] Carbon Dioxide Information Analysis Center; http://goo.gl/bECFSx.

2.3 Valuation

A value transfer approach is used to approximate valuations. In this approach, the value of nonmarket services is approximated through the examination of a group of studies that have been previously conducted based on similar nonmarket services. Although the authors recognize that primary data collection on valuation can provide strong, in-situ measurements of valuation, evidence suggests that the density of literature on similar services, as well as the cost-effective nature of the value transfer approach, positions value transfer as a strong "second-best" strategy (Costanza et al. 2014; see their footnote 9). In this study, we will follow the methodology put forth in Costanza et al. 2014 to select relevant studies, integrate this information with our geospatial data, and provide mapped estimates of total valuation across all four outcome measures.

3: Results

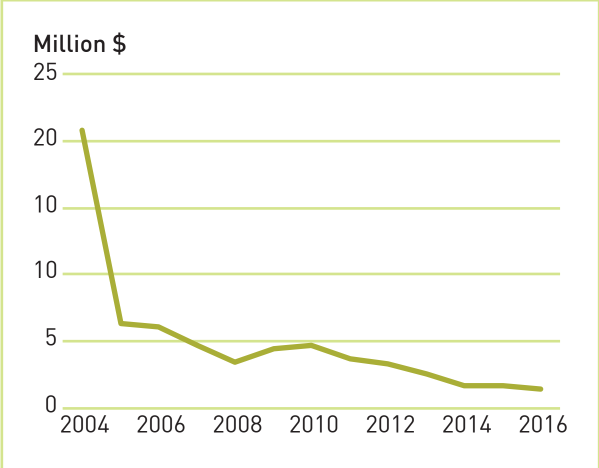
3.1 Descriptive findings

A total of 1,704 GEF land degradation project locations were included in this analysis from 445 projects, each with implementation dates between 2002 and 2014. These projects had disbursement levels ranging from $200,000 to $35.4 million. Over time, larger-scale projects tended to occur (on average) in the earlier time period, with a slight decreasing trend occurring toward 2014 (figure 3.1).

The results of a descriptive analysis examining the characteristics of GEF land degradation project locations (only considering projects for which an exact geographic location was available) can be found in table 3.1. These descriptors were based on the 25 kilometer areas around each GEF land degradation project. A few key findings are highlighted:

- GEF land degradation projects were located in areas that—on average—experienced positive increases in NDVI from 1982 to 2014. In this causal tree analysis we control for this upward trend by including information on the preproject implementation trend in NDVI as well as the level of NDVI in the year prior to project implementation.
- All GEF land degradation projects were within 25 kilometers of a designated protected area of any kind, although only a subset were within 25 kilometers of protected areas with legally empowered designations.
- GEF land degradation projects tended to be located in areas with relatively low population density and electrification.
- The physical geographic characteristics of areas the GEF operates in are highly variable, in terms of temperature, precipitation, elevation, and slope. Elevation is particularly notable in this regard, ranging from near-sea-level (~600 meters) to altitudes of 5,000 meters.
- Not all GEF land degradation projects are located in areas that have forest cover; 60 project locations were found to have no tree cover

FIGURE 3.1 Average project disbursements over time



SOURCE: GEF Project Management Information System.

TABLE 3.1 Descriptive statistics of GEF land degradation project locations

Statistic	Mean	Minimum	Median	Maximum
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	5,009.92
Slope (degrees)	3.278	0	1.974	19.173
Urban accessibility (relative)	622.302	31.078	260.602	4,644.30
Population density (2005)	184.866	0	75.209	4,179.14
NDVI (1982)	0.1756	0.0329	0.1778	0.3454
NDVI (2014)	0.1844	0.0286	0.1852	0.3982
Nighttime lights (2013)	1.651	0	0.372	32.422
Minimum air temp (2014)	17.042	-20.15	22.1	28
Maximum air temp (2014)	27.268	11.975	28	36.433
Mean air temp (2014)	22.453	-1.371	24.723	30.029
Maximum precipitation (2014)	277.145	17.7	217.275	1,470.65
Minimum precipitation (2014)	10.635	0	0.725	157.35
Mean precipitation (2014)	95.283	2.31	72.527	439.117
Protected area overlap	3.546	1	4	6
Tree cover – 2000 (percent)	17.596	0	6.772	98.076

in the initial 2000 period. However, NDVI measurements suggest that these areas did have vegetation biomass. This suggests GEF land degradation projects may be sited in degraded areas.

Although these descriptive findings do not indicate causality of GEF land degradation projects, they do provide insights into the highly varied geographies in which GEF land degradation projects operate.

3.2 Causal impacts

For each of the three models specified in table 3.2, Case 1, a causal tree is fit to identify the subsets of GEF land degradation projects for which differential treatment effects can be observed. This results in six different trees, which are summarized in this report. For the causal tree cases, we highlight the overall findings (i.e., if GEF land degradation projects in aggregate had positive, negative, or neutral impacts), as well as key findings of drivers of heterogeneity in causal impacts. For the case of random forests (Case 2), we contrast the results to facilitate a robustness check. Of key note is that, although each tree is unique, they all share the control variables identified in table 1.1 and summarized in table 3.1. If a variable is not present in a given tree, it can be interpreted as indicating that a particular variable was not key in defining subsets of the population for which the treatment varied in efficacy; however, the variable may still be important in mediating the impact in a single way across the entire population. Additionally, variables that are located in earlier splits in the tree tend to be more robust in terms of their importance in driving heterogeneity.

Not all observations were included in the causal tree analyses. The primary reason these were removed from observation was because of the date of implementation: in order to establish

TABLE 3.2 **Propensity model results**

Variable	Result
Baseline average NDVI	0.048
Baseline maximum NDVI	0.0004
Baseline minimum temperature	1.085**
Baseline maximum temperature	1.120**
Baseline average temperature	−2.146**
Baseline maximum precipitation	−0.002
Baseline minimum precipitation	0.001
Baseline average precipitation	0.018
Distance to rivers	0.00002
Distance to roads	0
Elevation	0.001
Slope	0.048
Urban accessibility	−0.003
Population density (2000)	0.002
Protected area overlap	0.218
Baseline tree cover (2000)	−0.007
Latitude	−0.009
Longitude	−0.009*

NOTE: *p<0.1; **p<0.05; ***p<0.01. Results exclude fixed effects and constant.

reasonable outcome measurements, the analysis was limited to projects that started in 2012 or earlier. Recognizing that even with this limitation, significant variation can be expected based on the number of years a project has had to make an impact, we further control for the amount of time that elapsed between the measurement of outcome and the year of implementation.

A single propensity model was fit that describes the likelihood of treatment as measured by the covariate information, and is presented in table 3.2. This model was fit using a logistic regression, in which the response variable was a binary (GEF land degradation project presence or absence). Although all variables are important in their role as controls in later stages of this analysis (see equation B.4), of note is the significant relationship between the average minimum and maximum temperature with an increased likelihood of site selection, and a relationship between average temperature and a decreased probability of selection. Furthermore, spatial patterns seem to play a role in site selection as evidenced by a significant relationship with longitude. Table 3.3 presents the pre- and post-matching difference between treatment and control groups along each ancillary variable, following a nearest neighbor matching strategy using the calculated propensity scores.

Following the indicators and subindicators suggested in the monitoring framework of the UNCCD and Sustainable Development Goals 15.3 for measuring land degradation (Minelli, Erlewein, and Castillo 2017; UNCCD 2015), three different metrics are used to ascertain the impact of GEF land degradation projects—vegetation productivity, forest cover, and forest fragmentation. Across the entire globe, GEF land degradation projects (1) increased NDVI by approximately 0.03 (relative to an average NDVI of 0.55), (2) reduced forest loss by 1.3 percent (relative to a global mean of 2.4 percent forest loss in all areas), and (3) increased the average size of forest patches by 0.25 kilometers (relative to a global mean of 7.3 square kilometers). We find that although the impact of GEF land degradation projects has been positive, there is considerable heterogeneity in impacts across different geographic contexts. Key findings for vegetation productivity included indications that projects in closer proximity to urban areas tended to be less effective; a minimum time lag of 5.5 years was an important threshold for determining impact in some contexts (with some geographic locations requiring 7.5 years), and a tendency for areas with poorer initial conditions to improve to a greater degree. When forest cover was examined, it was found that a 4.5-year lag time was influential in determining effectiveness.

TABLE 3.3 Difference in GEF land degradation projects and control locations before and after matching across covariate dimensions

Variable	Pre-matching	Post-matching	Improvement (%)
Baseline average NDVI	0.3713	−0.2136	42.4901
Baseline maximum NDVI	−139.2436	70.8344	49.1292
Baseline minimum temperature	4.8991	0.8943	81.7467
Baseline maximum temperature	1.1689	0.3929	66.3893
Baseline average temperature	2.8821	0.6062	78.9685
Baseline maximum precipitation	44.0293	10.0125	77.2594
Baseline minimum precipitation	2.9125	−0.9979	65.7382
Baseline average precipitation	15.1706	2.2311	85.2933
Distance to rivers	8250.3435	5784.3134	29.89
Distance to roads	2394.725	−3056.6765	−27.6421
Elevation	−74.4471	11.3287	84.7829
Slope	0.6414	−0.0327	94.8955
Urban accessibility	−192.9915	−6.2006	96.7871
Population (2000)	77.6916	−11.5721	85.1051
Protected area overlap	−0.1077	0.0226	79.0304
Percent tree cover (2000)	−3.9706	−0.2991	92.4672

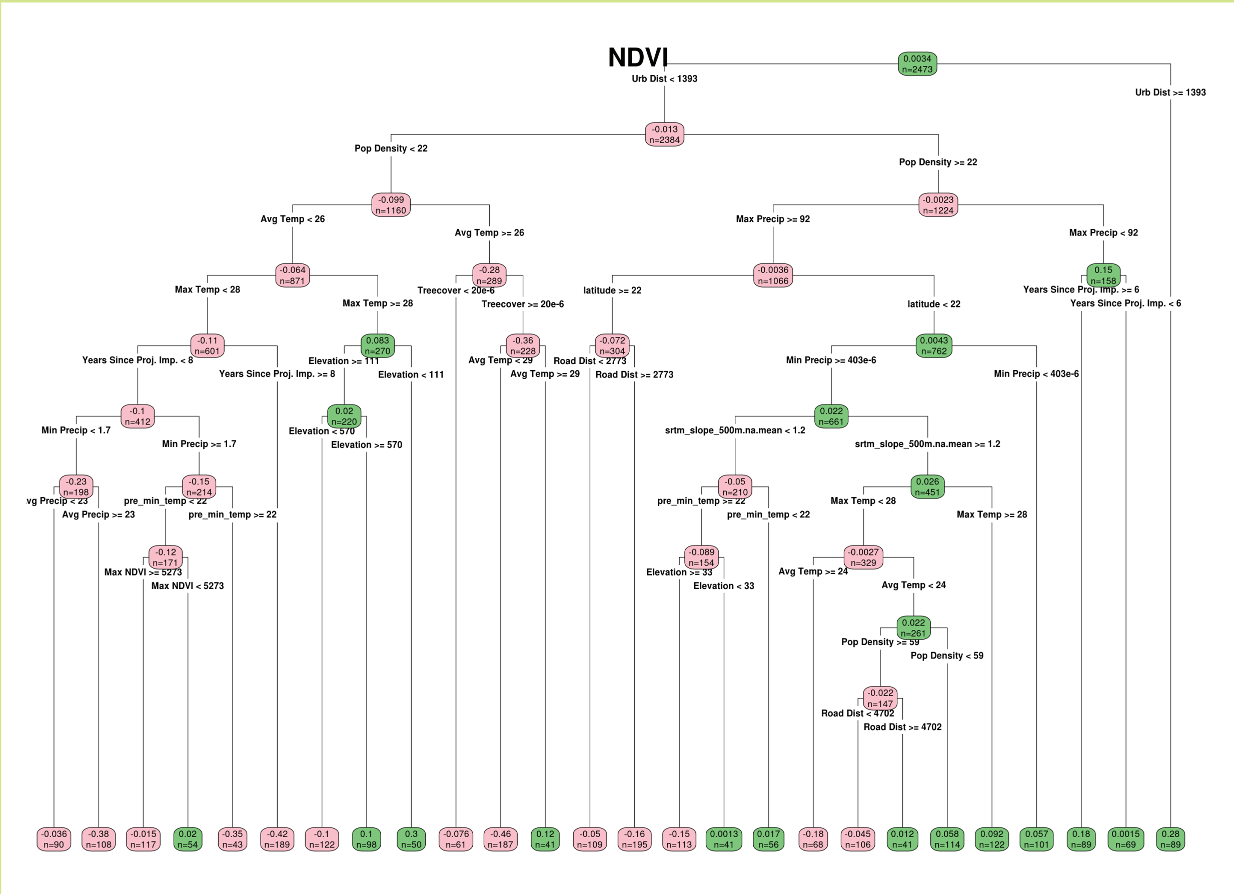
In the case of fragmentation, it was found that the initial state of fragmentation—i.e., the pretrend average—was a major factor in determining the heterogeneity in GEF land degradation project impacts.

The results of the causal tree analysis for NDVI can be seen in figure 3.2. In these results, we find that in aggregate, GEF land degradation projects had a small but positive impact on NDVI—specifically increasing NDVI by approximately 0.03 (relative to an average NDVI of 0.55). In addition to this aggregate finding, there are a number of findings in regard to the factors that mediated GEF land degradation impacts:

- In general, projects located in closer proximity to urban areas tended to be less effective than those located farther away.
- The period of time after project implementation was meaningful, with evidence suggesting that a minimum of 5.5-year time lag is an important threshold for determining the degree of impact in some contexts; the maximum time lag found to be important was 7.5 years.
- Although there is limited evidence of robustness, the analysis in this tree suggests that in limited contexts multifocal projects lead to improved outcomes.
- In some contexts, areas with poorer initial conditions (i.e., lower NDVI) saw greater improvement because of GEF land degradation projects.
- Environmental (slope, elevation, temperature, precipitation) and social characteristics (pop density, urban distance) all proved important in mediating the impact of GEF land degradation projects.

Figure 3.3 shows the causal tree describing the impact of GEF land degradation projects on forest

FIGURE 3.2 A causal tree representing impacts of GEF land degradation projects on vegetation productivity



cover. Each terminal node value represents the percent of tree cover loss that is attributable to GEF land degradation projects— i.e., a negative value indicates a GEF land degradation project slowed the rate of loss, whereas a positive value indicates it accelerated the rate of loss. As in the case of NDVI, globally there is a small but normatively positive impact attributable to GEF land degradation projects, which reduced forest loss by 1.3 percent (relative to a global mean of 2.4 percent forest loss in all areas). Key findings include the following:

- Evidence that projects with greater than 4.5 years of time since implementation had a stronger slowing effect on deforestation than more recent projects.
- Population density is a key factor driving heterogeneity in the impacts of GEF land degradation projects, but relatively few GEF land degradation projects took place in locations with extremely low population densities (less than one individual per square kilometer).
- There is some limited evidence that GEF land degradation projects closer to urban areas were slightly more successful in mitigating forest cover losses in some geographic areas.

FIGURE 3.3 A causal tree representing impacts of GEF land degradation projects on forest land cover

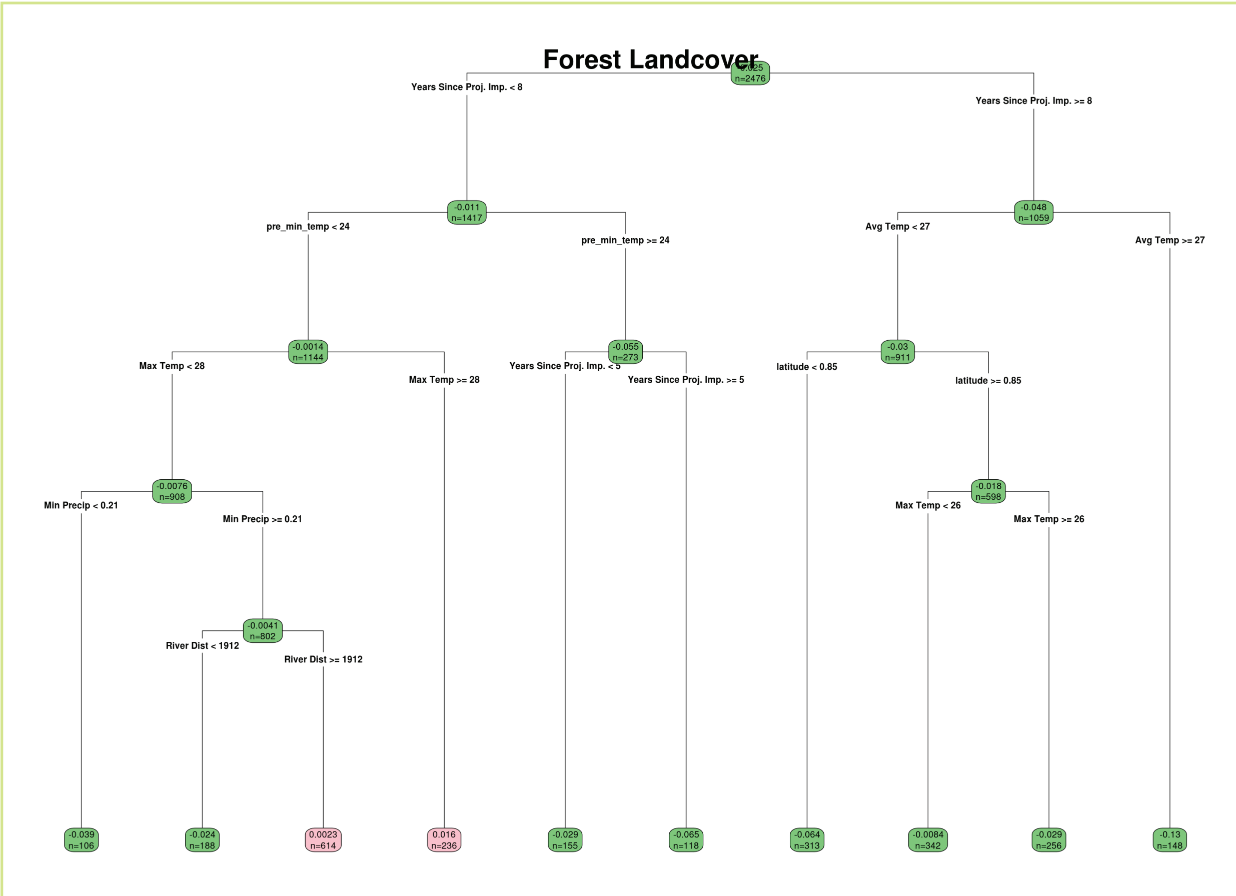
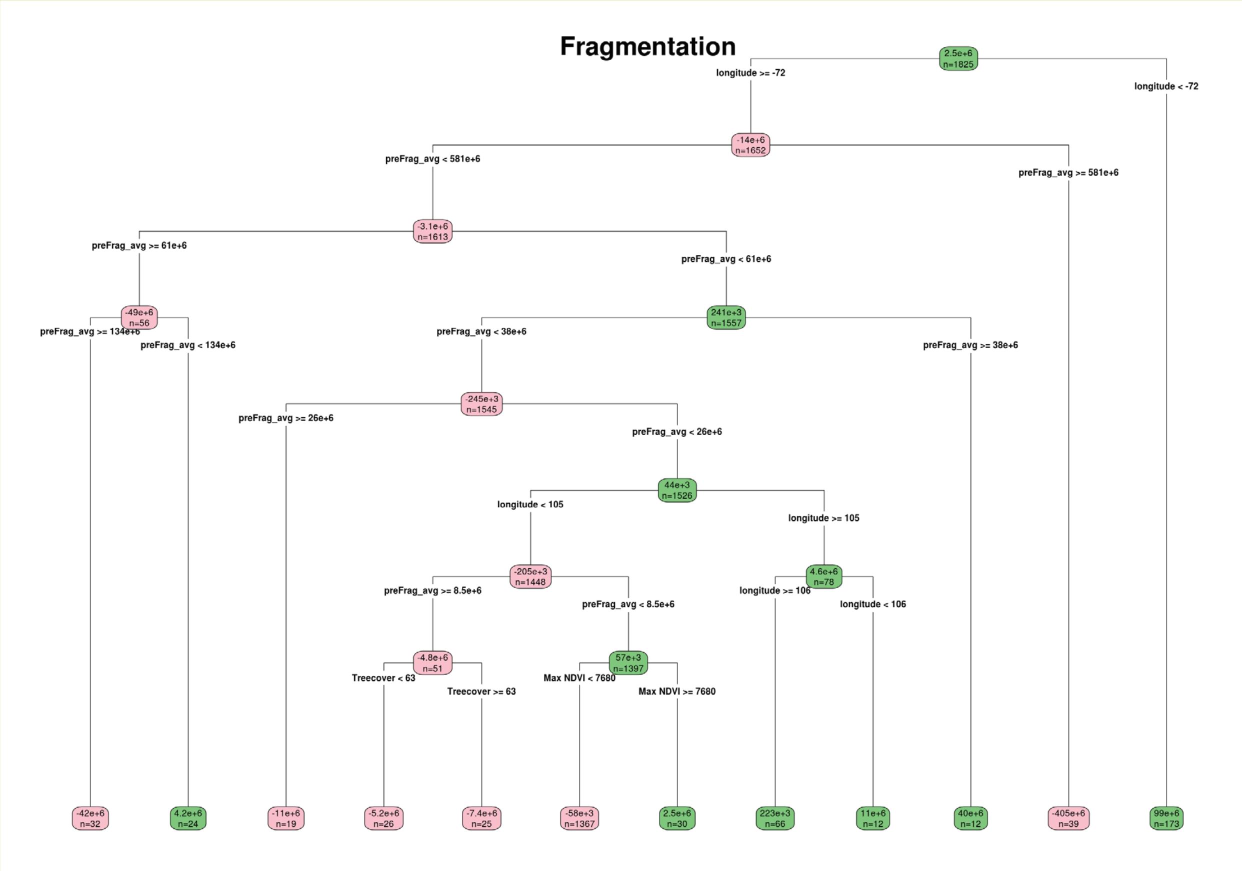


Figure 3.4 shows the causal tree describing the impact of GEF land degradation projects on forest fragmentation—specifically, the average forest patch size in 2014. In this case, positive values indicate an increase in patch size as a product of a GEF land degradation project. Globally, this analysis suggests that GEF land degradation projects positively contributed to the patch size of forests on average, but with more significant heterogeneity in impacts when compared to the other two indicators examined—i.e., many projects had negative or neutral impacts. On average, GEF land degradation projects increased the average size of forest patches by 0.25 kilometers (relative to a global mean of 7.3 square kilometers). Unmeasured geographic factors—or, strong spillover effects—tended to have a large impact in the case of forest fragmentation, with the geographic latitude and longitude of a project being a consistent driver of relative efficacy of projects. GEF land degradation projects were also heavily influenced by the initial state of forest fragmentation—i.e., the pretrend of average forest size is a major factor in determining the heterogeneity in GEF land degradation project impacts.

The model used to estimate carbon sequestration is detailed in section 3.4, and the results of this model are shown in table 3.4. Approximately 47 percent of the variation in carbon sequestration across projects can be explained by the model at the project-location scale, the most conservative

FIGURE 3.4 A causal tree representing impacts of GEF land degradation projects on forest fragmentation



unit of measurement available for this analysis. Although the model itself is purely predictive (and thus coefficient estimates and significance are not interpretable causally), the relative valuations of the different ecofloristic zones are of interest. These values indicate that the baseline values we use for estimation are highly variable by biome, an important factor for GEF land degradation projects that operate in semiarid and humid tropical areas. Furthermore, we find evidence supporting earlier academic literature that both forest cover loss and fragmentation are correlated with sequestration; NDVI plays a role in our prediction but we do not find significance in the relationship.

3.3 Valuation

A summary of identified valuations for carbon sequestration are available in annex A. Figure 3.5 shows a brief summary of these findings, which were drawn from 8,094 instances of reported carbon valuation. The documents included in the identification of these valuations ranged from official reports, to academic articles, to private-sector valuations. Carbon valuations were primarily drawn from prices of carbon taxes and carbon trading schemes. Data were drawn from a number of developed and developing countries (for example, from South African and Mexican carbon taxes to Beijing's pilot emission trading scheme),

TABLE 3.4 Carbon sequestration model

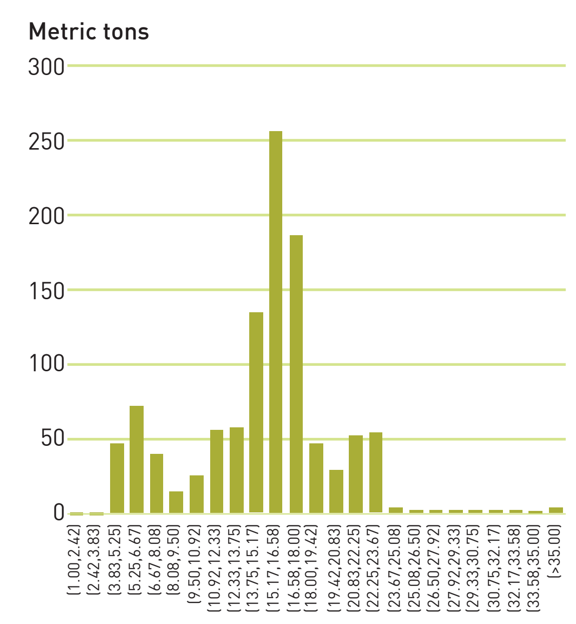
Variable	Result
Subtropical desert	–5,677.86
Subtropical mountain system	–83.87** (42.04)
Subtropical steppe	–5,754.645
Temperate desert	–100.6** (47.72)
Temperate mountain system	–103.8** (42.13)
Temperate steppe	–102.6** (46.70)
Tropical desert	–123.7*** (40.92)
Tropical dry forest	–111.7*** (39.39)
Tropical moist deciduous forest	–88.14** (39.19)
Tropical mountain system	–95.05** (39.45)
Tropical rainforest	–37.41 (39.39)
Tropical shrubland	–121.9*** (38.85)
Latitude	0.1531 (0.19)
Longitude	0.0976** (0.039)
Mean patch size (2010)	1.471×10^{-7}*** (2.94×10^{-8})
Forest cover loss (2010)	8.60×10^{-4}*** (2.81×10^{-4})
LTDR NDVI (2010)	-2.74×10^{-4} (2.01×10^{-3})
Constant	124.2*** (38.60)
adjusted R-squared	0.461

NOTE: *p<0.1; **p<0.05; ***p<0.01. Standard errors in parentheses.

but relied heavily on the European Union Emission Trading Scheme because of the availability of data.

Although the goal of this report is not to argue for any specific valuation an online tool available at the GEF IEO website enables users to select their own valuation), for the purposes of estimation we select the mean value from this data set ($12.90/metric ton), deflated to 2014. The green chart represents the density of observations, and the X-axis shows dollar values deflated to 2014 U.S. dollars.

FIGURE 3.5 Distribution of carbon valuations (per metric ton sequestered) identified in the literature

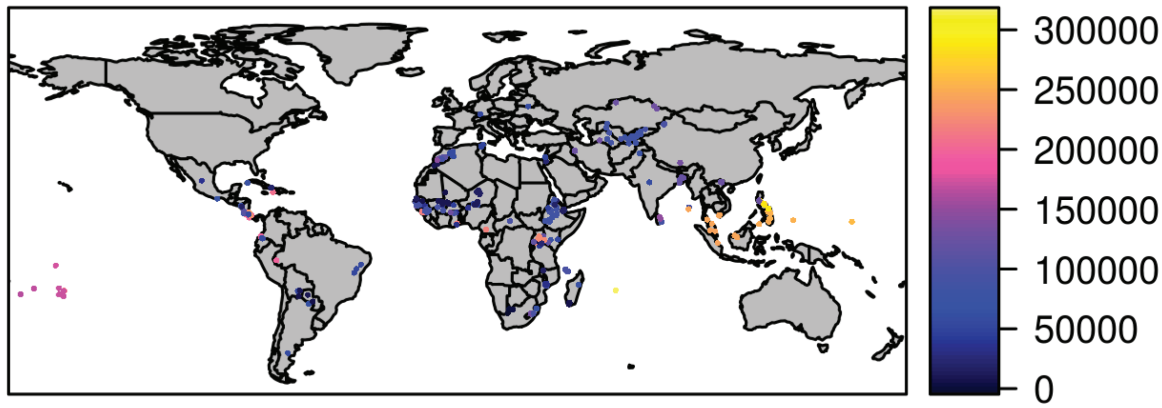


3.4 GEF land degradation project valuations

Following the methodology outlined density above, an estimate is performed for both project valuations overall as well as the valuation (and impact) for any given project location. Valuation for individual project locations is shown in figure 3.6. Based on the median estimate of $12.90 (2014 dollars) per sequestered metric ton and a 25 kilometer area of influence, GEF land degradation project locations were valued—on average—at $1,403,520, deflated to 2014 values. This ranged from a minimum of –$60,424, to a maximum of $4,108,650. At the project level, the mean valuation was $7,500,358, with a minimum of –$52,721 and a maximum of $48,653,058.

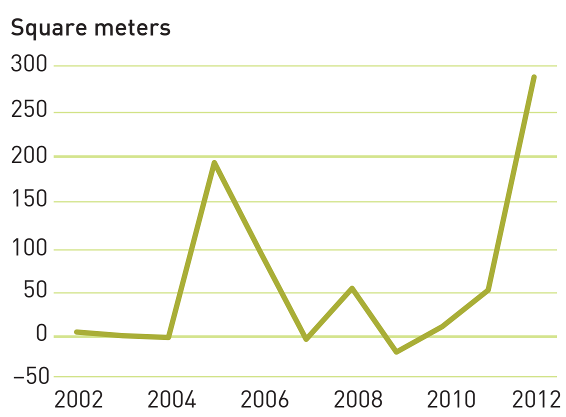
Over time, the value of a dollar of land degradation investment fluctuated across each of the three indicators assessed. In figure 3.7, the average

FIGURE 3.6 Estimated valuations of each GEF land degradation project location



NOTE: Projects can be viewed in more detail, and monetary valuation assumptions can be modified, at the GEF IEO website.

FIGURE 3.7 Shift in forest cover patch size attributable to $1 of GEF land degradation investment over time

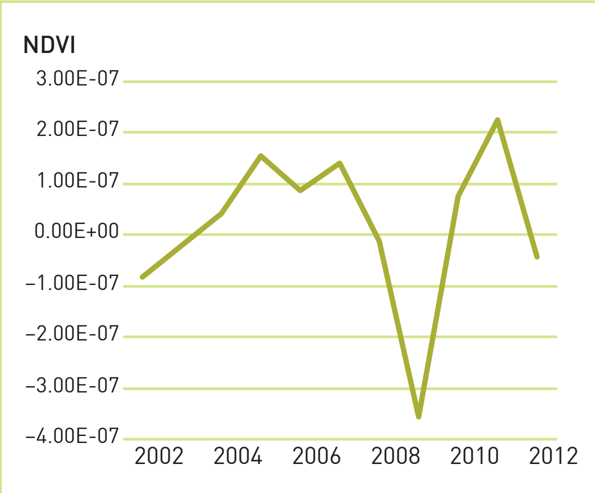


NOTE: Year is the year of project implementation; valuation is determined (in 2014 $) based on the impact of the project to 2014.

valuation of projects that started in a given year is presented (where valuation is defined based on impacts in 2014). In the case of fragmentation, projects that have been implemented most recently (2012) have apparent higher returns when contrasted to earlier projects; the causation of this pattern is beyond the scope of this analysis but may provide helpful insights for practitioners seeking to identify successful strategies.

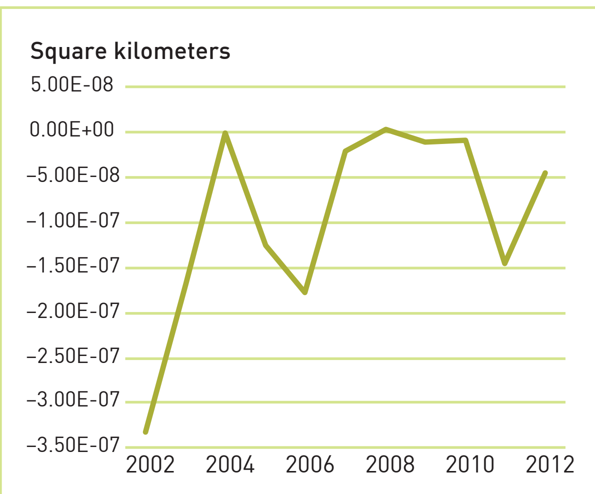
Figures 3.8 and 3.9 show the fluctuation in both NDVI and the rate of forest cover loss attributable to GEF land degradation projects, respectively. Projects that started in 2002, 2009, and 2012 all had notable negative impacts on NDVI, whereas projects in 2004 to 2007 and 2010-2011 tended to have positive impacts. In general, GEF land degradation projects have slowed deforestation, and projects that started in both 2002 and 2011 had a more positive effect overall (larger slowing of deforestation) per dollar of investment than other years. Across years, 2009 was generally the worst in terms of dollar efficacy, whereas projects that began in 2005 tended to have larger per dollar efficiencies in terms of these three indicators. Because the models presented in this report control for other potential drivers of project variation (i.e., weather), this graph suggests significant variation in the efficiency of GEF land degradation projects over time. Although it is outside the scope of this report to hypothesize what could cause these shifts, we note that an exploration of the projects that contributed to positive or negative

FIGURE 3.8 Shift in NDVI attributable to $1 of GEF land degradation investment over time



NOTE: Year is the year of project implementation; valuation is determined (in 2014 $) based on the impact of the project to 2014.

FIGURE 3.9 Shift in forest cover loss attributable to $1 of GEF land degradation investment over time

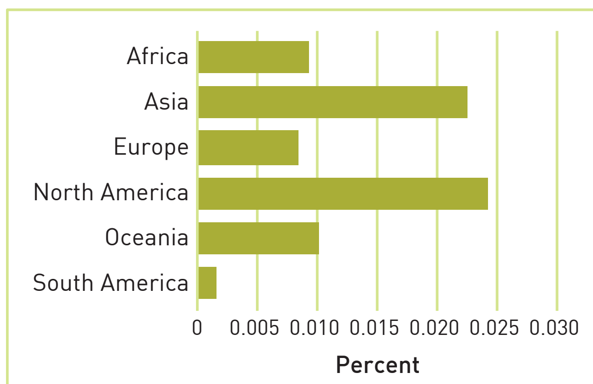


NOTE: Year is the year of project implementation; valuation is determined (in 2014 $) based on the impact of the project to 2014. Larger negative values indicate a slowing.

fluctuations could be further examined to better understand these findings.

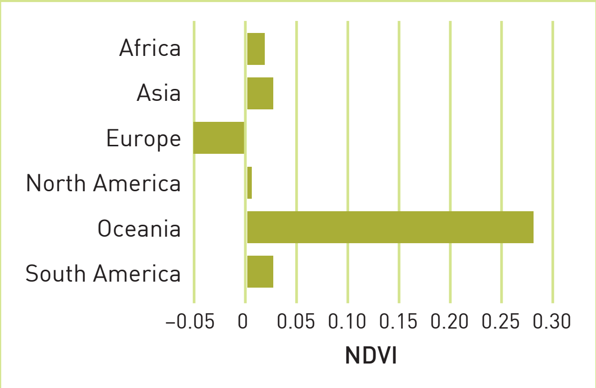
At the continental scale, there is also notable spatial variation in the impact of GEF land degradation projects. Figures 3.10, 3.11, and 3.12 describe this variation, which is generally reflective of the causal findings. Projects in Africa and Asia had generally positive impacts on average. Projects in Oceania, and North and South America all had positive impacts on all three indicators. In all regions of the world, land degradation focal area projects reduced the rate of forest loss as measured in 2014 (figure 3.10). Likewise, all regions except Europe saw improved vegetation productivity (figure 3.11). Fragmentation was the most differentiated across regions. Africa had the most fragmentation in areas of land degradation focal area projects, while North America, and South America had the largest mean patch sizes (figure 3.12).

FIGURE 3.10 Average estimated differences in rate of total forest loss at GEF intervention versus control locations



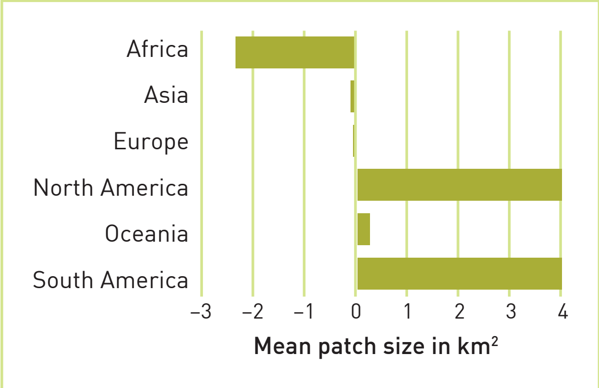
NOTE: This figure excludes a small number of projects that are not clearly delineated as affecting a single continent (i.e., where the exact degree of impact attributable to each continent is unknown).

FIGURE 3.11 Average estimated differences in increased vegetation productivity at GEF intervention versus control locations



NOTE: This figure excludes a small number of projects that are not clearly delineated as affecting a single continent (i.e., where the exact degree of impact attributable to each continent is unknown).

FIGURE 3.12 Average estimated differences in reduced forest fragmentation at GEF intervention versus control locations



NOTE: This figure excludes a small number of projects that are not clearly delineated as affecting a single continent (i.e., where the exact degree of impact attributable to each continent is unknown).

4: Discussion

Although this report provides evidence that, on average, GEF land degradation projects have mitigated or reversed negative land degradation processes, we also note the significant heterogeneity in these findings. We emphasize this heterogeneity to highlight the many opportunities for improvement that still exist by learning why and where GEF land degradation projects are leading to outcomes with relatively high benefits. These heterogeneities were found over both time—with project impacts being variable on a year-by-year basis—as well as space. As more observations are made available, we anticipate that further drivers of heterogeneity in project impact could be observed (i.e., geopolitical issues; macroeconomic trends).

The use of propensity score-matching techniques to examine the causal effects of an intervention (i.e., international aid, a new business process, a new website design) has its roots in econometric research from the early 1980s (Rosenbaum and Rubin 1983). Since their introduction, propensity-matching methods have been used for everything from better understanding customer retention and loyalty (Xerox 2005), to the testing of new medical drugs (see Baek et al. 2015), to understanding supply chain dynamics (Fałkowski 2009), and have been used extensively by researchers and practitioners seeking to understand the impact of aid (i.e., Gundersen and Sara 2016; Mensah et al. 2010). Most recently, these methods have become popular for testing websites such as eBay, Facebook, and many more to establish and test optimal website designs (Taddy 2016; Backshy 2014; Briggs 2007). Practitioners have constantly refined matching approaches to understand causality, and the most recent wave of innovation has centered around heterogeneous impact effects—i.e., how an impact might vary across different geographic areas or groups of individuals (Athey and Imbens 2015). This is coupled with a push from geographic information scientists and practitioners to apply these approaches to geographic data to more cost-effectively ascertain environmental impacts, as well as considerable increases in the quality of satellite imagery available (i.e., Hansen et al. 2013; Kim et al. 2014; Sexton et al. 2013). For example, using satellite and other geo-referenced data, propensity-score matching and difference-in-difference approaches have been used to evaluate the impact of World Bank projects on forest change in key biodiversity areas (Buchanan et al. 2016), indigenous communities' land rights on deforestation in Brazil (BenYishay et al. 2016), and land titling and land management programs in Ecuado (Buntaine, Hamilton, and Millones 2015b).

Here, we advance the state of the art by applying a joint econometric and machine learning technique (specifically, causal trees) to examine how the impacts of GEF land degradation projects vary across geography and other factors. By examining the heterogeneity in impacts—rather than exclusively estimating overall effects—we show that (1) it is feasible to conduct global- scope, top-down analyses, as traditional methods for impact evaluation require prespecification of possible factors driving heterogeneity; and (2) it is possible to distinguish between sources of positive and negative impacts.

We additionally employ state-of-the-art satellite imagery to detect changes as fine as 30 meters—a key factor when fragmentation and precise measurement of tree cover is of interest. By using the geographic information system to couple this satellite imagery with a wide variety of other, globally available data sets (table 1.1), we are able to provide geographic, contextual information that enables the identification of counterfactual cases. Furthermore, by leveraging features of geographic variance itself—i.e., the trend that locations that are closer together tend to be more similar along unmeasured variables—we argue that this approach can mitigate, although not completely remove, many challenges associated with omitted variable biases.

By coupling three approaches—econometric propensity matching techniques, computer science machine learning algorithms, and geographic information system satellite imagery analysis and data integration—we further enable more accurate valuation of the impact of GEF land degradation projects across broader scopes than has been possible to date. By providing a methodology through which the impact of individual project locations can be estimated along multiple, value-relevant indicators, valuation efforts can focus on the single (but still very difficult) challenge of valuing shifts in indicator values, rather than methods to identify the precise percentage of a shift that is attributable to any given project.

This study has a number of remaining uncertainties and limitations that could be resolved through future work. First and foremost, this analysis is top-down, using only project information that is available at a global scale. Although matching based on geography and geographic patterns can strongly mitigate omitted variable biases (i.e., by selecting treatment and control sites close together, and thus likely to experience similar conditions), nuanced, project-scale factors could still confound the results present here. We argue that, despite this limitation, the analysis presented here can be powerful in (1) identifying possible "bright spots" and "warning signs" at a relatively low cost; (2) identifying the geographic contexts in which GEF land degradation projects are most successful; and (3) providing strategic guidance as to the global and regional effectiveness of GEF land degradation projects. We strongly caution against using the information—or approach—detailed in this report to drive project location-level decision making without coupled, "bottom-up" analyses.

The scope across which GEF land degradation projects have impact is frequently unknown. Because limited geographic information has traditionally been collected on the exact geographic boundaries across which an intervention is performed, the underlying data used in this and similar analyses is point-based (i.e., a latitude and longitude coordinate). Because land degradation projects occur in a diffuse manner, the area across which project impact is anticipated lacks exact geometric representations. .Although we use a 25 kilometer buffer around each intervention, the collection of more precise geographic boundary information at the time of project implementation could result in more accurate impact estimates.

The value transfer approach (Costanza et al. 2014) leveraged to estimate the total valuations of carbon sequestration is known as a "second best" option. More advanced approaches to estimating final valuations—including more explicit, regional modeling of value, stakeholder interviews, or economic-impact analyses—could provide better insights into the final estimated valuations of each project location. Furthermore, the valuations estimated in this report only consider impacts on carbon sequestration, and do not take into consideration other benefits project locations may accrue—for example, co-benefits related to other ecosystem services and infrastructure development. Future analyses could leverage alternative remotely sensed data sets—such as Nighttime Lights—to construct indicators adequate to detect such co-benefits, or to further investigate questions of avoided emissions.

5: Conclusion

The findings of this report suggest that, in aggregate, GEF land degradation projects have had a positive impact on indicators and sub-indicators of land degradation proposed by the UNCCD for the Land Degradation Neutrality and Sustainable Development Goals 15.3—specifically vegetation productivity (measured by NDVI) and forest cover (measured directly and by mean patch size). Although these impacts vary substantially over space and time, we provide evidence that the GEF has contributed to increasing the total amount of carbon sequestered by forest cover and related biophysical processes. We estimate that, at a valuation of $12.90/metric ton, GEF land degradation projects contributed $7.5 million (2014 dollars) on average to sequestration alone— well above the average cost of most GEF land degradation projects ($4,182,887). We note that considerable heterogeneity exists in these findings.

Although examining the causal impact of international aid on environmental outcomes has been a central goal of many communities, there has been a limited engagement using spatially-explicit, geocoded aid information because of limitations in both data and methods (Athey and Imbens 2015; Corrado and Fingleton 2012). These methodological limitations primarily stem from distinctions between modeling efforts seeking to predict relationships commonly taught and accepted by the geographic community (i.e., spatial regression or classification trees), and efforts that seek to establish causal relationships similarly taught and accepted by the economics community (i.e., propensity-score matching or difference-in-difference modeling). Recent efforts have been undertaken to merge these disciplinary approaches (Buntaine, Hamilton, and Marco 2015b; Drukker, Egger, and Prucha 2013; Runfola et al. 2016), of which this report provides another example.

The methodology detailed in this report goes beyond these examples by providing an approach to capturing heterogeneity in impact effects—i.e., how GEF land degradation projects may vary in impact across different countries, regions, climate regimes, or human factors. This approach to learning based on historic GEF land degradation project implementations can additionally be flexibly applied to predict the potential impact of future projects (alongside concomitant uncertainties). As the cost of this analytical approach is lower than traditional impact evaluation, and enables the use of historic information, we believe it represents a screening step practitioners could take before project implementation.

Annex A: Definitions

A.1 Vegetation productivity

There are many different approaches to approximating vegetation on a global scale, and satellites have been taking imagery that can be used for this purpose for over three decades. Of these approaches, the most frequently used, and applied in this study, is the NDVI. The NDVI is a metric that has been used since the early 1970s, and is one of the simplest and most frequently used approaches to approximating vegetation biomass. Furthermore, it is recommended as an indicator by the GEF Scientific and Technical Advisory Panel.[1] NDVI measures the relative absorption and reflectance of red and near-infrared light from plants to quantify vegetation on a scale of –1 to 1, with vegetated areas falling between ~0.2 and 1.

The reflectance by chlorophyll is correlated with plant health, and multiple studies have illustrated that it is generally also correlated with plant biomass. In other words, healthy vegetation and high plant biomass tend to result in high NDVI values (Dunbar 2009). Using NDVI as an outcome measure has a number of other benefits, including the long and consistent time periods for which it has been calculated. Although the NDVI does have a number of challenges—including a propensity to saturate over densely vegetated regions, the potential for atmospheric noise (including clouds) to incorrectly offset values, and reflectance from bright soils providing misleading estimates—the popularity of this measurement has led to a number of improvements over time to offset many of these errors. This is especially true of measurements from longer-term satellite records, such as those produced from the MODIS (Moderate Resolution Imaging Spectroradiometer) and AVHRR (Advanced Very High Resolution Radiometer) (NASA 2015).

[1] www.stapgef.org/.

A.2 Land cover change

Understanding the relationships between "process and pattern"—i.e., the links between drivers and observations of land cover change—has long been a focus of practitioners (Lambin et al. 2001; Liverman 1998; Meyer and Turner 1996; Nagendra et al. 2004; Turner et al. 2003). Land cover change has major implications for a broad range of phenomena, including the sustainability of human development, biogeochemical cycling, and levels of greenhouse gasses (Turner et al. 1995; UNDP 2010). Investigating the many factors that influence land cover/use provides an avenue through which the human-environment interface can be better understood, but recent research has emphasized the lack of understanding of how anthropogenic processes influence land change (Nagendra et al. 2004). The impacts of land use/cover change on the vulnerability and sustainability of human-dominated landscapes is just beginning to be analyzed, and improving this understanding is a major goal of parties interested in understanding the consequences of land use change (Foley et al. 2005; GLP 2010).

Both the geographic and development economics communities have sought to understand linkages between international development and land cover change, but often using different approaches and vocabulary. Within the geographic community, limited attention has been given to causal methodologies (including matching and difference-in-difference models), but rather focused on the (1) ability to accurately measure land cover change using satellite imagery (i.e., Borak, Lambin, and Strahler 2000; Christman et al. 2015; Rogan et al. 2003; Schwert et al. 2013; Strahler, Moody, and Lambin, n.d.), (2) impacts of spatial autocorrelation on model estimates (Miller, Arun, and Timmons Roberts 2012; Waldron et al. 2013), and (3) methods for predicting the impact(s) (and related uncertainties) of international aid on land change (Laurance et al. 2002; D. M. Runfola and Pontius 2013; van Asselen and Verburg 2013). Conversely, the development economics community has focused on the application of matching (Nelson and Chomitz 2011) and difference-in-difference (Alix-Garcia, Shapiro, and Sims 2012; Nolte et al. 2013; Pfaff 1999) techniques to establish evidence of causal relationships between international aid and land cover change—methods that follow similar approaches to clinical trials with treatment and control groups.

To capture land cover change in this analysis, we leverage an analysis performed by Hansen et al. (2013), in which LandSat imagery was fused with a number of other sources to capture 30-meter-resolution, yearly estimates of tree cover loss. This land cover change analysis is widely leveraged to capture trends in deforestation, and represents one of the highest-resolution efforts for such measurements ever conducted. Furthermore, as a global analysis, this product enables a precise calculation of both (1) tree cover in the year 2000, and (2) loss from 2000 to 2013 for every GEF land degradation project location.

A.3 Forest fragmentation

Classical forest fragmentation occurs when forest patches become smaller and more isolated than those in an undisturbed landscape, a process which can be driven by both natural and anthropogenic causes (Wulder et al. 2009). Academic and policy literature has repeatedly shown that fragmentation can have significant environmental implications (Garcia et al. 2005; Mingshi et al. 2010; Riitters et al. 2012). These implications include negative impacts on the biodiversity of an area (Hanski 2005; Kolb and Diekmann 2005; Zuidema, Sayer, and Dijkman 1996), negative effects on carbon sequestration (Diaz, Hector and Wardle 2009; Matthews, O'Connor, and Plantinga 2002), as well as modified risks of natural disasters such as fire (Convention on International Trade in Endangered Species of Wild Fauna and Flora). Although there are many ways to describe fragmentation, in this analysis we examine the average patch size within the area of influence of GEF land degradation projects.

A.4 Carbon stocks and sequestration

Forests contribute significantly to carbon sequestration through holding large carbon stocks. The combination of field-based estimates and remote sensing techniques has become the primary method of examining carbon stocks and carbon sequestration (Asner et al. 2010; Maselli et al. 2006; Muukkonen and Heiskanen 2005) because of difficulties with solely field- based estimates (Gibbs et al. 2007; Houghton 2005; Saatchi et al. 2007). Carbon stocks cannot be observed directly from satellite imagery; however, they can be estimated through examining factors associated with carbon stocks, particularly vegetation biomass.

NDVI is one of the most widely used vegetation indexes to estimate carbon stocks.

To date, empirical studies employing remote sensing to estimate carbon storage have done so at a local or country level and have shown that NDVI can strongly predict carbon stocks. For example, Myeong, Nowak, and Duggin (2006) estimated carbon storage among urban trees in Syracuse, New York, and found that NDVI explains 67 percent of the variation in field-based model estimates of carbon storage. Widayati, Ekadinata, and Syam (2005) examined the relationship between carbon stocks and NDVI in Indonesia, motivated by the need to evaluate the effectiveness of community-based forest management projects in reducing deforestation. They found that NDVI explains 52.8 percent of the variation in carbon density. Wylie et al. (2003) used remote sensing to predict CO2 carbon fluxes in a sagebrush-steppe ecosystem in northeastern Idaho, finding that NDVI explains 79 percent of the variation in carbon flux, and including evapotranspiration as a predictor variable increased explanatory power to 82 percent. Gang et al. (2013) used NDVI, in combination with temperature and precipitation data, to estimate carbon stocks in the Xilingol grasslands in northern China, predicting carbon stocks with a 92.5 percent accuracy. For other studies that used NDVI to model carbon stocks, see Gilmanov et al. (2004) for estimates in Kazakhstan; Hunt et al. (2002, 2004) for estimates in Wyoming; Tan et al. (2007) and Piao et al. (2005) for estimates across China; Kanniah, Muhamad, and Kang (2014) and Hamdan et al. (2013) for estimates in Malaysia; and Verhegghen et al. (2012) for estimates of the Congo Basin.

Some researchers have moved beyond the local level to estimate global carbon stocks. Saatchi et al. (2011) estimated forest carbon stocks across 2.5 billion hectares of forests, covering Africa, Asia, and South America. They relied on 14 remotely sensed variables (including NDVI) to estimate carbon stocks and field samples from 493 field sites to develop the model. They examined the predictive power of the 14 variables across geographic regions, where they found NDVI metrics to explain most of the variation in carbon stocks in low biomass-density forests.

Other studies estimated carbon stocks around the world or at regional levels relying on remotely sensed data beyond NDVI. For example, see Baccini et al. (2012) and Ruesch and Gibbs (2008) for global estimates; Saatchi et al. (2007) for estimates of the Brazilian Amazon; Baccini et al. (2008), Brown and Gaston (1996), and Gibbs and Brown (2007a) for tropical Africa; and Brown, Iverson, and Prasad (2001) and Gibbs and Brown (2007b) for Southeast Asia. Furthermore, some researchers have found that the relationships between NDVI, forest cover, and carbon sequestration can be further permuted by forest fragmentation (Diaz, Hector and Wardle 2009; Matthews, O'Connor, and Plantinga 2002).

Annex B: Methods

B.1 Data integration

Many of the data sets used in this analysis are collected at different spatial scales, necessitating an additional step of integration so that all observations can be analyzed at the scale of GEF land degradation projects (in this case, examining a 10 kilometer x 10 kilometer region around each project). To conduct this integration, we use the piecewise approximation procedure detailed in Goodchild et al. (1993):

$$V_t = \sum_{s=1}^{S} \left[U_s \times \left(\frac{a_{st}}{a_s} \right) \right] \qquad \textbf{Eq. B.1}$$

where t is an index for the zone one is aggregating to (the GEF land degradation project area of interest), s is an index for the set of zones one is aggregating from (i.e., a satellite pixels measuring NDVI), S is the maximum index for all zones s, U_s represents the value of interest at source zone s, a_{st} is the area of overlap between the two zones, a_s is the area of the zone one is aggregating from, and V_t is the estimated value for the target zone. In our application, this procedure weights each pixel of each data set according to its overlap with each GEF land degradation project.

B.2 Causal model

Classification and regression tree approaches have been commonly employed over the last two decades to aid in the classification of remotely sensed imagery (Friedl and Brodley 1997; McIver and Friedl 2002; Gamba and Herold 2009). Here, we employ causal trees—a novel version of a Classification and Regression Tree (CART) which enables causal inferential analyses. Causal trees are implemented in a multiple step process, detailed below but simply summarized as (1) deriving a metric that indicates similarity between treatment and control groups; (2) using this metric to match pairs of treatment and control units via a tree; (3) contrasting the outcome of treated units to control units within every terminal node of the tree. Figure 2.1 shows an example drawn from exploratory research in which a causal tree is applied to a limited subset of international aid, examining aid's impact on a maximum observed NDVI value. This figure serves as an illustrative example of the outputs of causal tree-based approaches to identifying how impact effects may differ across a data set. Unlike traditional econometric approaches in which interaction terms must be prespecified to estimate differential impact effects; here, clusters of similar treatment and control units are identified dynamically. Furthermore, by including geographic factors in these trees (i.e., latitude and longitude), many unobserved geographic characteristics can be captured. As in a traditional econometric analysis in which variables can be identified as statistically significant, here variables that are significant (defined as the variables that describe the most variance in the data) are represented in the tree. All variables are controlled for through the propensity adjustment of the outcome (see equation B.4).

The primary distinction between causal trees and more traditional tree-based classifiers lies in the criterion along which splits in the tree are selected. Consider a data set with *n* independently and identically distributed units with $i = 1, \ldots, n$, and for each unit a vector of relevant covariates are measured. In a simplified case where all things other than treatment are being constant, to estimate a causal effect for each geographic location *i* we can use the Rubin causal model (Rubin 1997) and consider the treatment effect as being equal to

$$\theta_i = Y_i(W_i = 1) - Y_i(W_i = 0) \qquad \textbf{Eq. B.2}$$

where W_i is an indicator of whether a unit of observation *i* received aid (1) or did not (0). Following this simplified model, we define the expected heterogeneous causal effect for any set of units as (Athey and Imbens 2015b):

$$\theta_i = \Sigma[Y_i(W_i = 1) - Y_i(W_i = 0) \mid X_i = x] \qquad \textbf{Eq. B.3}$$

Athey and Imbens show that one can estimate the causal effect as $\theta_i = \Sigma[Y_i^* \mid X_i = x]$ where the transformed outcome Y^* is defined as:

$$Y_i^* = Y_i \frac{(W_i - e(X_i)}{e(X_i) \times (1 - e(X_i))} \qquad \textbf{Eq. B.4}$$

and the propensity score function $e(x)$ is defined as $e(x) = \Sigma[W_i \mid X_i = x]$. Several approaches to estimate the propensity score can be selected (Rosenbaum and Rubin 1983; Pan and Bai 2015)—here, we estimate $e(x)$ using logistic regression. Once the propensity score and Y_i^* have been estimated, many authors (Su et al. 2009; Athey and Imbens 2015b; Wagner and Athey 2015; Denil et al. 2014; Meinhausen 2016; Biau 2012; Wagner et al. 2014) have illustrated that classification and regression trees can be used to isolate treatment effects within sets of similar units. These trees seek to classify units of observation into clusters that are similar along covariate axes, following different splitting and optimization rules.

Using the propensity score, causal tree approaches derive a transformed outcome variable, Y^*, and use this to generate tree splits instead of (the traditionally used) *Y*. This transformed outcome is calculated following equation B.4. The causal tree replaces the traditional mean square error optimization criterion in trees by seeking to minimize the sum of $Y_i^* - \hat{\tau}(X_i)$ in each terminal node, where $\hat{\tau}(X_i)$ represents the estimated average treatment impact within a given node, i.e.:

$$\hat{\tau}^{CT}(X_i) = \sum_{i:X_i \in X_i} Y_i^{obs} \times \frac{W_i/\hat{e}(X_i)}{\Sigma i:X_i \in X_i W_i/\hat{e}(X_i)} - \sum_{i:X_i \in X_i} Y_i^{obs} \times \frac{(1 - W_i)/(1 - \hat{e}(X_i))}{\Sigma i:X_i \in X_i (1 - W_i)/(1 - \hat{e}(X_i))} \qquad \textbf{Eq. B.5}$$

This new error term is then used to split the tree in a way identical to traditional regression trees, and provides a tree that increases the similarity of control and treated units within each node, as well as node-specific estimates of impacts.

Annex C: Geocoding international aid

This project leveraged the AidData development finance and international aid geocoding methodology. In 2010, AidData developed a methodology for georeferencing development projects that International Aid Transparency Initiative later revised and adopted as its global reporting standard. Leveraging a team of trained geocoders, the geocoding methodology and online toolkit relies on a double-blind coding system, where two experts employ a defined hierarchy of geographic terms and independently assign uniform latitude and longitude coordinates, precision codes, and standardized place names to each geographic feature. If the two code rounds disagree, the project is moved into an arbitration round where a geocoding project manager reconciles the codes to assign a master set of geocodes for all of the locations described in the available project documentation. This approach also captures geographic information at several levels— coordinate, city, and administrative divisions—for each location, thereby allowing the data to be visualized and analyzed in different ways depending upon the geographic unit of interest. Once geographic features are assigned coordinates, coders specify a location class ranging from 1 to 4 for categories, including administrative regions or topographical features, along with a location type specifying the exact feature (e.g., airport, second order administrative zone, etc.). Coders then determine the location's geographic exactness value of either 1 (exact) or 2 (approximate).

AidData performs many procedures to ensure data quality, including de-duplication of projects and locations, correcting logical inconsistencies (e.g., making sure project start and end dates are in proper order), finding and correcting field and data type mismatches, correcting and aligning geocodes and project locations within country and administrative boundaries, validating place names and correcting gazetteer inconsistencies, deflating financial values to constant dollars across projects and years (where appropriate), strict version control of intermediate and draft data products, semantic versioning to delineate major and minor versions of various geocoded datasets, and final review by a multidisciplinary working group.

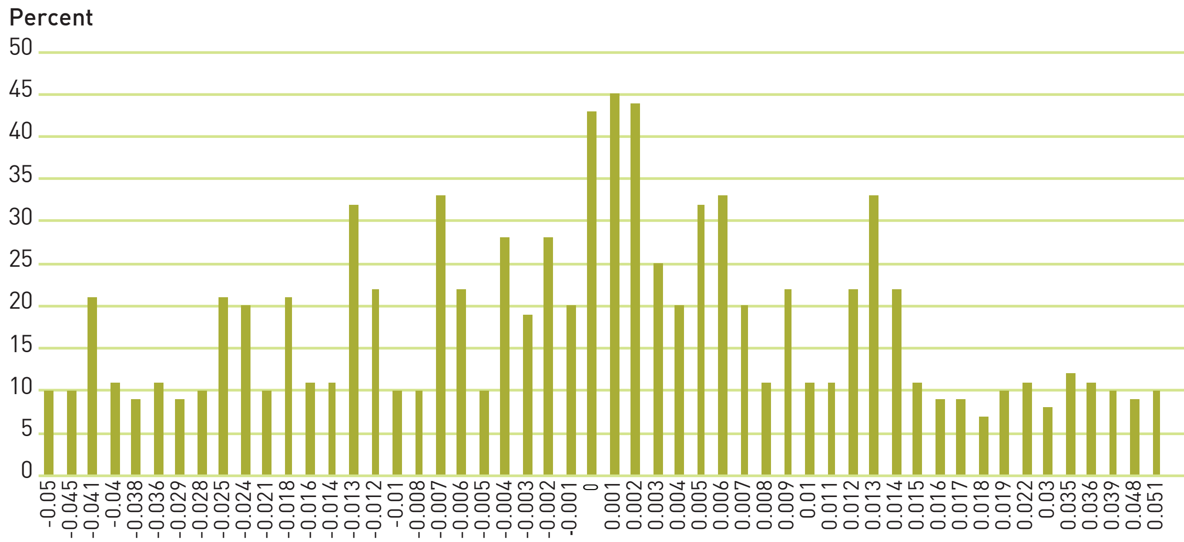
Annex D: Robustness checks

In order to test the robustness of the results presented in this document, two different approaches were followed. First, a random forest implementation of the causal tree approach was implemented. Second, the analysis was repeated using the traditional causal tree approach, but using the watershed in which each unit fell as the unit of observation (i.e., watersheds with no GEF land degradation projects contained within them were matched to watersheds that contained GEF land degradation projects). The random forest-causal tree approach takes a different approach to uncertainty than a traditional causal tree. In the random forest, a large number of trees (in this case, 10,000) are fit, each time fitting by using a different subset of the data. This approach provides two advantages. First, it allows for an estimate of the importance of different variables across trees; i.e., it can be established which variables seem to drive heterogeneity in the impacts of GEF land degradation projects. Second, it provides a range of possible values that could be estimated for each GEF land degradation project, given the potential for different matches across different subsets of the data. From these two points of evidence, it is possible to provide insight into the relative certainty of claims for any given observation, as well as the structure of the tree found in the traditional causal tree approach. The primary drawback of the random forest-causal tree approach is that it does not provide a single tree for interpretation (as in the above causal tree approach), thus limiting potential insights regarding the exact contexts in which projects succeed and fail.

Figure D.1 illustrates an example of how uncertainty because of tree construction can be captured for each individual GEF land degradation project location. We can use this distribution to calculate the percentage of observations within, for example, 1 standard deviation of the mean. Although this cannot be interpreted as a statistical significance (attributable to the lack of parametric assumptions in the underlying models and distributions, as well as differential aims of the tests), if a high percentage of observations fall in this area, we illustrate that our findings are generally robust with regard to the shape of the tree. This analysis is conducted for each of the three focal areas, as summarized in table D.1.

As table D.1 illustrates, the most robust results were found in the estimates of forest cover, with 90.5 percent of observations (across all GEF land degradation projects estimated) falling within 1 standard deviation of the mean. Both vegetation productivity (NDVI) and forest fragmentation had lower overall robustness, but both have robustness scores at 1 standard deviation greater than 80 percent. At the 2 standard deviation mark, all models had a rate of 93 percent or higher. In practice, this table suggests that although forest cover had the highest robustness, all three models can be described as robust with regard to the shape of the trees.

FIGURE D.1 Result of a random forest for one GEF observation



NOTE: Each of 1,000 iterations are plotted.

TABLE D.1 Robustness of findings by outcome measure

Outcome measure	Percentage of observations that fall within: 1 standard deviation of the mean	2 standard deviations of the estimate
Forest cover	90.5	96.3
Vegetation productivity (NDVI)	80.1	93.3
Forest fragmentation	84.3	94.8

NOTE: Higher values indicate more robust findings.

Table D.2 provides information on the relative importance of the top 10 variables across each of the random forests. For example, if a variable appears in many trees at a relatively high position, it will be rated high in this table; conversely, if it does not frequently appear or is low in the tree it is in a relatively low position. These tables can be interpreted to better understand the robustness of the shape of the trees presented in figures 3.2–3.4.

Table D.2 illustrates the robustness of the shape of the trees—specifically, if claims regarding a particular split can be determined to be robust.

Of note is that post-implementation time appears in all three models as being an important factor in distinguishing GEF land degradation project impacts. Additional variables that are important across multiple outcome measures included the year of implementation (suggesting a difference in the effectiveness of projects over time), geographic factors (latitude and longitude), and a variety of physical and environmental characteristics.

The full list of valuations considered for carbon can be found at https://tinyurl.com/ycg2r49l.

TABLE D.2 Relative importance of variables within each random forest

Variable rank	Relative occurrence in random forest (purity)		
	Land cover	Fragmentation	NDVI
1	Pretrend max NDVI	Latitude	Latitude
2	Year	Year	2000 Tree Cover
3	Pretrend avg air temp	Post-implementation time	Pretrend max NDVI
4	Latitude	Slope	Pretrend min temp
5	Post implementation time	Prelevel min air temp	Pretrend max temp
6	Slope	Prelevel avg air temp	Elevation
7	Longitude	2000 tree cover	Year
8	Urban accessibility	Pretrend avg air temp	Urban accessibility
9	Population density (2000)	Longitude	Pretrend min precipitation
10	Pretrend average NDVI	Elevation	Post implementation time
11+	All other variables	All other variables	All other variables

NOTE: The top 10 occurring variables are presented here, weighted by the location they appear in the tree (higher indicates more weight, with 1 being most influential) as well as the number of occurrences across all trees.

References

Asner, G. P., G. V. N. Powell, J. Mascaro, et al. 2010. "High-Resolution Forest Carbon Stocks and Emissions in the Amazon." *Proc Natl Acad Sci U S A* 107 (38): 16738-16742. doi: 10.1073/pnas.1004875107.

Alix-Garcia, J. M., E. N. Shapiro, and K. R. E. Sims. 2012. "Forest Conservation and Slippage: Evidence from Mexico's National Payments for Ecosystem Services Program." *Land Economics* 88 (4): 613–638.

Athey, Susan, and Guido Imbens. 2015. "Recursive Partitioning for Heterogeneous Causal Effects." *arXiv:1504.01132 [stat]*, April.

Baccini, A., N. Laporte, S. J. Goetz, M. Sun, and H. Dong. 2008. "A First Map of Tropical Africa's Above-Ground Biomass Derived From Satellite Imagery." *Environmental Research Letters* 3 (4): 045011. doi: 10.1088/1748-9326/3/4/045011.

Baccini, A., S. Goetz, W. Walker, et al. 2012. "Estimated Carbon Dioxide Emissions From Tropical Deforestation Improved by Carbon-Density Maps." *Nature Climate Change* 2 (3): 182-185. doi: 10.1038/NCLIMATE1354.

Baek, Seunghee, Seong Ho Park, Eugene Won, Yu Rang Park, and Hwa Jung Kim. 2015. "Propensity Score Matching: A Conceptual Review for Radiology Researchers." *Korean Journal of Radiology*. 16 (2): 286–296. doi: 10.3348/kjr.2015.16.2.286.

Bakshy, Eytan, Dean Eckles, Rong Yan, and Itamar Rosenn. 2012. "Social Influence in Social Advertising: Evidence from Field Experiments." In ACM Conference on Electronic Commerce (EC), June 8, 2012.

BenYishay, Ariel, Silke Heuser, Daniel Runfola, and Rachel Trichler. 2016. "Indigenous Land Rights and Deforestation: Evidence from the Brazilian Amazon." AidData Working Paper 22.

Borak, J. S., E. F. Lambin, and A. H. Strahler. 2000. "The Use of Temporal Metrics for Land Cover Change Detection at Coarse Spatial Scales." *International Journal of Remote Sensing* 21 (6-7): 1415–32.

Biau, G. "Analysis of a Random Forests Model." 2012. *Journal of Machine Learning Research* 13: 1063–1095.

Brown, S., and Gaston G. 1996. "Tropical Africa: Land Use, Biomass, and Carbon Estimates for 1980." Oak Ridge: Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy.

Brown, S., L. R. Iverson, and A. Prasad. 2001. "Geographical Distribution of Biomass Carbon in Tropical Southeast Asian Forests: A Database." ORNL/CDIAC-119, NDP-068. Oak Ridge: Carbon Dioxide Information Analysis Center, U.S. Department of Energy, Oak Ridge National Laboratory.

Buchanan, Graeme M., Bradley C. Parks, Paul F. Donald, Brian F. O'Donnell, Daniel Runfola, John P. Swaddle, Łukasz Tracewski, and Stuart H.M. Butchart. 2016. "The Impacts of World Bank Development Projects on Sites of High Biodiversity Importance." AidData Working Paper 20.

Buntaine, Mark T., Stuart E. Hamilton, and Millones Marco. 2015a. "Titling Community Land to Prevent Deforestation: An Evaluation of a Best-Case Program in Morona-Santiago, Ecuador." *Global Environmental Change: Human and Policy Dimensions* 33: 32–43.

———. 2015b. "Titling Community Land to Prevent Deforestation: An Evaluation of a Best- Case Program in Morona-Santiago, Ecuador." *Global Environmental Change: Human and Policy Dimensions* 33: 32–43.

Christman, Zachary, Christman Zachary, Rogan John, J. Ronald Eastman, and B. L. Turner II. 2015. "Quantifying Uncertainty and Confusion in Land Change Analyses: A Case Study from Central Mexico Using MODIS Data." *GIScience and Remote Sensing*, 1–28.

Corrado, L., and B. Fingleton. 2012. "Where Is the Economics in Spatial Econometrics?" *Journal of Regional Science*. Wiley Online Library. doi/10.1111/j.1467-9787.2011.00726.x/full.

Costanza, Robert, Rudolf de Groot, Paul Sutton, Sander Van der Ploeg, Sharolyn J. Anderson, Ida Kubiszewski, Stephen Farber, and R. Kerry Turner. 2014. "Changes in the Global Value of Ecosystem Services." *Global Environmental Change* 26: 152–158.

Denil, Misha, Alban Demiraj, Nal Kalchbrenner, Phil Blunsom, and Nando de Freitas. 2014. "Modelling, Visualising and Summarising Documents With a Single Convolutional Neural Network." *arXiv preprint* 1406.3830.

Diaz, S., A. Hector, A., and D. A. Wardle. 2009. "Biodiversity in Forest Carbon Sequestration Initiatives: Not Just a Side Benefit." *Current Opinion in Environmental Sustainability* 1 (1): 55–60.

Drukker, David M., Peter Egger, and Ingmar R. Prucha. 2013. "On Two-Step Estimation of a Spatial Autoregressive Model with Autoregressive Disturbances and Endogenous Regressors." *Econometric Reviews* 32 (5-6): 686–733.

Dunbar, B. 2009. "NDVI: Satellites Could Help Keep Hungry Populations Fed as Climate Changes."

Fałkowski, Jan. 2012. "Dairy Supply Chain Modernisation in Poland: What About Those Not Keeping Pace?" *European Review of Agricultural Economics* 39 (3): 397–415. doi: 10.1093/erae/jbr046.

Foley, J. A et al. 2005 "Global Consequences of Land Use." *Science* 309, 570–574. doi: 10.1126/science.1111772.

Friedl, M. A., and C. E. Brodley. 1997. "Decision Tree Classification of Land Cover From Remotely Sensed Data." *Remote Sensing of Environment* 61:399–409.

Gamba, Paolo, and Martin Herold. 2009. *Global Mapping of Human Settlement: Experiences, Datasets, and Prospects*. Boca Raton, FL: CRC Press.

García, Daniel, Mario Quevedo, J. Ramón Obeso, and Adán Abajo. 2005. "Fragmentation Patterns and Protection of Montane Forest in the Cantabrian Range (NW Spain)." *Forest Ecology and Management* 208 (1-3): 29–43.

Gibbs, H. K., and S. Brown S. 2007a. "Carbon Pools in the Forests of Tropical Africa: An Updated Database Using the GLC2000 Land Cover Product." NDP-055b. ORNL-CDIAC.

———. 2007b. "Carbon Pools in the Forests of Tropical Southeast Asia: An Updated Database Using the GLC2000 Land Cover Product." NDP-068b. ORNLCDIAC.

Gibbs, H. K., S. Brown, J. O. Niles, and J. A. Foley. 2007. "Monitoring and Estimating Tropical Forest Carbon Stocks: Making REDD a Reality." *Environmental Research Letters* 2 (4): 045023. doi: 10.1088/1748-9326/2/4/045023.

Gilmanov, T. G., D. A. Johnson, N. Z. Saliendra, K. Akshalov, and B. K. Wylie. 2004. "Gross Primary Productivity of the True Steppe in Central Asia in Relation to NDVI: Scaling Up CO2 Fluxes. *Environmental Management*. 33 (S1): S492–S508. doi: 10.1007/s00267-003-9157-7.

GLP. 2010. "Theme 2: Consequences of Land System Change Feedbacks to the Coupled Earth." (G. L. Project, Ed.) *GLP Science Plan and Implementation Strategy*. Copenhagen: GLP International Project Office.

Goodchild, M. F., L. Anselin, and U. Deichmann. 1993. "A Framework for the Areal Interpolation of Socioeconomic Data." *Environment and Planning A 25* (3): 383–397.

Gundersen, Sara. 2016. "Disappointing Returns to Education in Ghana: A Test of the Robustness of OLS Estimates Using Propensity Score Matching." *International Journal of Educational Development* 50: 74–89. doi: 10.1016/j.ijedudev.2016.05.003.

Hamdan, O., M. Khairunnisa, A. Ammar, I. M. Hasmadi, and H. K. Aziz. 2013. "Mangrove Carbon Stock Assessment by Optical Satellite Imagery." *Journal of Tropical Forest Science* 25 (4): 554–565.

Hansen, M. C., P. V. Potapov, R. Moore, M. Hancher, S. A. Turubanova, A. Tyukavina, D. Thau, S. V. Stehman, S. J. Goetz, T. R. Loveland, A. Kommareddy, A. Egorov, L. Chini, C. O. Justice, and J. R. G. Townshend. 2013. "High-Resolution Global Maps of 21st-Century Forest Cover Change." *Science* 342; 6160: 850–853. doi: 10.1126/science.1244693.

Hanski, I. 2005. "Landscape Fragmentation, Biodiversity Loss and the Societal Response." *EMBO Reports* 6 (5): 388–392. doi: 10.1038/sj.embor.7400398.

Houghton, R. A. 2005. "Aboveground Forest Biomass and the Global Carbon Balance." *Global Change Biology* 11 (6): 945–958. doi: 10.1111/j.1365-2486.2005.00955.x.

Hunt, Jr, E. R., R. D. Kelly, W. K. Smith, J. T. Fahnestock, J. M. Welker, and W. A. Reiners. 2004. "Estimation of Carbon Sequestration by Combining Remote Sensing and Net Ecosystem Exchange Data for Northern Mixed-Grass Prairie and Sagebrush–Steppe Ecosystems." *Environmental Management* 33 (S1): S432–S441. doi: 10.1007/s00267-003-9151-0.

Hunt, E. R., J. T. Fahenstock, R. D. Kelly, J. M. Welker, W. A. Reiners, and S. W.K. 2001. "Carbon Sequestration From Remotely-Sensed NDVI and Net Ecosystem Exchange." In: *From Laboratory Spectroscopy to Remotely Sensed Spectra of Terrestrial Ecosystems*. Dordrecht, Netherlands: Kluwer Academic Publishers; 161.

Kanniah, K. D., N. Muhamad, and C. S. Kang. 2014. "Remote Sensing Assessment of Carbon Storage by Urban Forest." 8th International Symposium of the Digital Earth (ISDE8). doi: 10.1088/1755-1315/18/1/012151.

Kim, D.-H., J. O. , P. Noojipady, C. Huang, A. Anand, S. Channan, M. Feng, and J. R. Townshend. 2014. "Global, Landsat-Based Forest-Cover Change From 1990 to 2000." *Remote Sensing of Environment 155*: 178–193. doi: 10.1016/j.rse.2014.08.017.

Kolb, Annette, and Martin Diekmann. 2005. "Effects of Life-History Traits on Responses of Plant Species to Forest Fragmentation." *Conservation Biology* 19 (3): 929-938.

Lambin, Eric F. , Bi L. Turner, Helmut J. Geist, Samuel B. Agbola, Arild Angelsen, John W. Bruce, Oliver T. Coomes, Rodolfo Dirzo, Günther Fischer, and Carl Folke. 2001. "The Causes of Land-Use and Land-Cover Change: Moving Beyond the Myths. *Global Environmental Change* 11 (4): 261–269.

Laurance, William F., Ana K. M. Albernaz, Schroth Gotz, Philip M. Fearnside, Bergen Scott, Eduardo M. Venticinque, and Carlos Da Costa. 2002. "Predictors of Deforestation in the Brazilian Amazon." *Journal of Biogeography* 29 (5-6): 737–748.

Liverman, D. M., E. Moran, R. R. Rindfuss, and P. C. Stern. 1998. *People and Pixels: Linking Remote Sensing and Social Ccience*. Washington, DC: National Academies Press.

Maselli, F., A. Barbati, M. Chiesi, G. Chirici, and P. Corona. 2006. "Use of Remotely Sensed and Ancillary Data for Estimating Forest Gross Primary Productivity in Italy." *Remote Sensing of Environment* 100 (4): 563–575. doi: 10.1016/j.rse.2005.11.010.

Matthews, S., R. O'Connor, and A. J. Plantinga. 2002. "Quantifying the Impacts on Biodiversity of Policies for Carbon Sequestration in Forests." *Ecological Economics* 40 (1): 71–87.

McIver, D. K., and M. A. Friedl. 2002. "Using Prior Probabilities in Decision-Tree Classification of Remotely Sensed Data." *Remote Sensing of Environment* 81 (2-3): 253–61.

Meinshausen, Nicolai. "Quantile Regression Forests." 2016. *Journal of Machine Learning Research* 7: 983–999.

Meyer, W. B., and B. L. Turner. 1996. "Land-Use/Land-Cover Change: Challenges for Geographers." *GeoJournal 39* (3): 237–240.

Mingshi,L., Chengquan Huang, Zhiliang Zhu, Weisong Wen, Da Xu, and Anxing Liu. 2009. "Use of Remote Sensing Coupled With a Vegetation Change Tracker Model to Assess Rates of Forest Change and Fragmentation in Mississippi, USA." *International Journal of Remote Sensing* 30 (24): 6559–6574.

Miller, Daniel C., Agrawal Arun, and J. Timmons Roberts. 2012. "Biodiversity, Governance, and the Allocation of International Aid for Conservation." *Conservation Letters* 6 (1): 12–20.

Minelli, S., A. Erlewein, and V. Castillo, V. 2017. "Land Degradation Neutrality and the UNCCD: From Political Vision to Measurable Targets. In: *International Yearbook of Soil Law and Policy 2016* (pp. 85–104). Cham, Switzerland: Springer. doi: 10.1007/978-3-319-42508-5_9.

Muukkonen, P., and J. Heiskanen. 2005. "Estimating Biomass for Boreal Forests Using ASTER Satellite Data Combined With Standwise Forest Inventory Data." *Remote Sensing of Environment* 99 (4): 434–447. doi: 10.1016/j.rse.2005.09.011.

Myeong, S., D. J. Nowak, and M. J. Duggin. 2006 "A Temporal Analysis of Urban Forest Carbon Storage Using Remote Sensing." *Remote Sensing of Environment* 101 (2): 277–282. doi: 10.1016/j.rse.2005.12.001.

Nagendra, H., D. K. Munroe, and J. Southworth. 2004. "From Pattern to Process: Landscape Fragmentation and the Analysis of Land Use/Land Cover Change." *Agriculture, Ecosystems & Environment* 101; (2–3): 111–15.

NASA (National Aeronautics and Space Administration). 2015. "The Land Long Term Data Record."

Nelson, Andrew, and Kenneth M. Chomitz. 2011. "Effectiveness of Strict vs. Multiple Use Protected Areas in Reducing Tropical Forest Fires: A Global Analysis Using Matching Methods." *PloS One* 6 (8): e22722.

Nolte, Christoph, Arun Agrawal, Kirsten M. Silvius, and Britaldo S. Soares-Filho. 2013. "Governance Regime and Location Influence Avoided Deforestation Success of Protected Areas in the Brazilian Amazon." *Proceedings of the National Academy of Sciences of the United States of America* 110 (13): 4956–61.

Pan, W., and H. Bai. 2015. *Propensity Score Analysis: Fundamentals and Developments*. New York: Guilford Publications.

Pfaff, Alexander S. P. 1999. "What Drives Deforestation in the Brazilian Amazon?" *Journal of Environmental Economics and Management* 37 (1): 26–43.

Piao, S., J. Fang, B. Zhu, and K. Tan. 2005."Forest Biomass Carbon Stocks in China Over the Past 2 Decades: Estimation Based on Integrated Inventory and Satellite Data." *Journal of Geophysical Research - Biogeosciences* 110 (G1): G01006. doi: 10.1029/2005JG000014.

Riitters, Kurt H., John W. Coulston, and James D. Wickham. 2012. "Fragmentation of Forest Communities in the Eastern United States." *Forest Ecology and Management* 263: 8537 (1): 26–493.

Rogan, John, Rogan John, Miller Jennifer, Stow Doug, Franklin Janet, Levien Lisa, and Fischer Chris. 2003. "Land-Cover Change Monitoring with Classification Trees Using Landsat TM and Ancillary Data." *Photogrammetric Engineering & Remote Sensing* 69 (7): 793–804.

Rosenbaum, Paul R., and Donald B. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70 (1): 41–55. doi: 10.1093/biomet/70.1.41.

Rubin, Donald B. 1997. "Estimating Causal Effects from Large Data Sets Using Propensity Scores." *Annals of Internal Medicine* 127 (8_Part_2): 757–63.

Ruesch A. S., and H. Gibbs. 2008. "New IPCC Tier-1 Global Biomass Carbon Map for the Year 2000." Oak Ridge, TN: Oak Ridge National Laboratory.

Runfola, Dan, Jeff Tanner, Matthias Leu, Graeme Buchanan, and Ariel BenYishay. 2016. "Valuing Environmental Impacts of World Bank Projects: A Case Study of Integrating Value for Money Analysis into Impact Evaluations." Washington, DC: World Bank.

Runfola, D. M., and R. G. Pontius Jr. 2013. Measuring the temporal instability of land change using the flow matrix." *International Journal of Geographical Information Science* 27.9: 1696-1716.

Saatchi S., R. Houghton, R. Alvala, J. Soares, and Y. Yu. 2007 . "Distribution of Aboveground Live Biomass in the Amazon Basin." *Global Change Biology* 13 (4): 816–837. doi: 10.1111/j.1365-2 486.2007.01323.x.

Saatchi, S. S., N. L. Harris, S. Brown, et al. 2011. "Benchmark Map of Forest Carbon Stocks in Tropical Regions Across Three Continents." *Proc Natl Acad Sci U S A* 108 (24): 9899-9904. doi: 10.1073/pnas.1019576108.

Sayer, Jeffrey, and Bruce Campbell. 2004. *The Science of Sustainable Development: Local Livelihoods and the Global Environment.* Cambridge, UK: Cambridge University Press.

Schwert, B., J. Rogan, N. M. Giner, Y. Ogneva-Himmelberger, S. D. Blanchard, and C. Woodcock. 2013. "A Comparison of Support Vector Machines and Manual Change Detection for Land-Cover Map Updating in Massachusetts, USA." *Remote Sensing Letters* 4 (9): 882–890.

Sexton, J. O., X.-P. Song, M., Feng, P. Noojipady, A. Anand, C. Huang, D.-H. Kim, K. Collins, S. Channan, C. DiMiceli, and J. R. Townshend. 2013. "Global, 30-m Resolution Continuous Fields of Tree cover: Landsat-Based Rescaling of MODIS Vegetation Continuous Fields With Lidar-Based Estimates of Eror." *International Journal of Digital Earth* 6 (5): 427–448. doi: 10.1080/17538947.2013.786146.

Shen, Changyu, Yang Hu, Xiaochun Li, Yadong Wang, Peng-Sheng Chen, and Alfred E. Buxton. 2016. "Identification of Subpopulations with Distinct Treatment Benefit Rate Using the Bayesian Tree." *Biometrical Journal. Biometrische Zeitschrift*, June. DOI: 10.1002/bimj.201500180.

Staff, The Plos One. 2014. "Correction: Inferring Tree Causal Models of Cancer Progression with Probability Raising." *PloS One* 9 (12): e115570.

Strahler, A. H., A. Moody, and E. Lambin. n.d. "Land Cover and Land-Cover Change from MODIS." In: *1995 International Geoscience and Remote Sensing Symposium, IGARSS '95. Quantitative Remote Sensing for Science and Applications*. doi: 10.1109/igarss.1995.521802.

Su, Xiaogang, Chih-Ling Tsai, and Morgan C. Wang. 2009. "Tree-Structured Model Diagnostics for Linear Regression." *Machine Learning* 74, no. 2: 111–131.

Taddy, Matt, Matt Gardner, Liyun Chen, and David Draper. 2016. "A Nonparametric Bayesian Analysis of Heterogenous Treatment Effects in Digital Experimentation." *Journal of Business & Economic Statistics* 34: 661–672. doi: 10.1080/07350015.2016.1172013.

Tan, K., S. Piao, C. Peng, and J. Fang. 2007. "Satellite-Based Estimation of Biomass Carbon Stocks for Northeast China's Forests Between 1982 and 1999." *Forest Ecology and Management* 240 (1): 114–121. doi: 10.1016/j.foreco.2006.12.018.

Turner, et al. 2003. "A Framework for Vulnerability Analysis in Sustainability Science." *Proceedings of the National Academy of Sciences of the United States of America* 100 (14): 8074–8079.

UNDP (United Nations Development Programme). 2010. "UN-REDD Programme Year in Review 2010."

Van Asselen, S., and P. H. Verburg. 2013. "Land Cover Change or Land-Use Intensification: Simulating Land System Change With a Global-Scale Land Change Model". *Global Change Biology* 19 (12): 3648–3667. doi: 10.1111/gcb.12331.

Verhegghen, A., P. Mayaux, C. de Wasseige, and P. Defourny. 2012. "Mapping Congo Basin Vegetation Types From 300 m and 1 km Multi-Sensor Time Series for Carbon Stocks and Forest Areas Estimation." *Biogeosciences* 9 (12): 5061–5079. doi: 10.5194/bg-9-5061-2012.

Wager, Stefan. 2014. "Asymptotic Theory for Random Forests." *arXiv preprint arXiv: 1405.0352*.

Wager, Stefan, and Susan Athey. 2015. "Estimation and Inference of Heterogeneous Treatment Effects Using Random Forests." *Journal of the American Statistical Association*.

Waldron, Anthony, Arne O. Mooers, Daniel C. Miller, Nate Nibbelink, David Redding, Tyler S. Kuhn, J. Timmons Roberts, and John L. Gittleman. 2013. "Targeting Global Conservation Funding to Limit Immediate Biodiversity Declines." *Proceedings of the National Academy of Sciences of the United States of America* 110 (29): 12144–12148.

Widayati, A., A. Ekadinata, and R. Syam. 2005. "Land Use Change in Nunukan. Estimating Landscape Level Carbon-Stocks Through Land Cover Types and Vegetation Density." *World Agroforestry Centre Publications*.

World Bank. 2008. *Sustainable Land Management Sourcebook: Agriculture and Rural Development*. Washington, DC: The World Bank.

Wulder, M. A., J. C. White, M. D. Gillis, et al. 2010. "Multiscale Satellite and Spatial Information and Analysis Framework in Support of a Large-Area Forest Monitoring and Inventory Update." *Environmental Monitoring and Assessment* 170 (1–4): 417–433. doi: 10.1007/s10661-009-1243-8.

Xerox. 2005. "Leveraging Customer Information to Drive Measurable Results in Revenue Growth and Loyalty: The Power of Personalization and Relevance." White paper.

Zuidema, P., J. Sayer, and W. Dijkman. 1996. "Forest Fragmentation and Biodiversity: The Case for Intermediate-Sized Conservation Areas." *Environmental Conservation 23* (4): 290–297. doi: 10.1017/S037689290003914X.

Recent GEF IEO Publications

Evaluation Reports		
132	Biodiversity Focal Area Study	2018
131	Evaluation of the Expansion of the GEF Partnership	2018
130	Evaluation of the GEF's System for Transparent Allocation of Resources	2018
129	Impact of GEF Support on National Environment Laws and Policies	2018
128	Climate Change Focal Area Study	2018
127	Evaluation of the GEF Partnership and Governance Structure	2018
126	Formative Review of the Integrated Approach Pilot Programs	2018
125	OPS6: The GEF in the Changing Environmental Finance Landscape – Executive Summary (available in English, French, and Spanish)	2018
124	Evaluation of the Multiple Benefits of GEF Support through Its Multifocal Area Portfolio	2018
123	Evaluation of Knowledge Management in the GEF	2018
122	Evaluation of GEF Support for Transformational Change	2018
121	GEF Project Performance and Progress to Impact	2018
120	Land Degradation Focal Area Study	2018
119	Evaluation of GEF Engagement with Indigenous Peoples	2018
118	Evaluation of Gender Mainstreaming in the GEF	2018
117	Program Evaluation of the Special Climate Change Fund	2018
116	Review of the GEF Policy on Agency Minimum Standards on Environmental and Social Safeguards	2018
115	Chemicals and Waste Focal Area Study	2018
114	International Waters Focal Area Study	2018
113	Evaluation of Programmatic Approaches in the GEF	2018
112	Review of Results-Based Management in the GEF	2017
111	Evaluation of GEF Engagement with the Private Sector, Volumes 1 and 2	2017
110	OPS6: The GEF in the Changing Environmental Finance Landscape – Final Report	2017
109	GEF Annual Performance Report 2015	2017
108	Evaluation of the GEF–Civil Society Organization Network, Volumes 1 and 2	2016
107	GEF Country Portfolio Evaluation: Tajikistan (1999–2015), Volumes 1 and 2	2016
106	Program Evaluation of the Least Developed Countries Fund	2016
105	FEM Evaluation du Portefeuille de Pays : Maroc (1997–2015)	2016
104	Impact Evaluation of GEF Support to Protected Areas and Protected Area Systems—Full Report	2016
103	Impact Evaluation of GEF Support to Protected Areas and Protected Area Systems—Highlights	2015
102	GEF Annual Performance Report 2014	2015
101	GEF Country Portfolio Evaluation: Eritrea (1992–2012)	2015
100	GEF Country Portfolio Evaluation: Tanzania (1992–2012)	2015
99	GEF Country Portfolio Study: Sierra Leone (1998–2013)	2015
98	GEF Country Portfolio Evaluation: Vanuatu and SPREP (1991–2012)	2015
97	Joint GEF-UNDP Evaluation of the Small Grants Programme	2015
96	Joint GEF–Sri Lanka Country Portfolio Evaluation (1991–2012)	2015
95	GEF Annual Country Portfolio Evaluation Report 2014	2015

To see all GEF Independent Evaluation Office publications, please visit our webpage.

Independent Evaluation Office
Global Environment Facility
1818 H Street, NW Washington, DC 20433, USA
www.gefieo.org
/gefieo_tweets
/gefieo