gef IEO Independent Evaluation Office
GLOBAL ENVIRONMENT FACILITY

World Bank

FOCUS SESSION ON USE OF NEW TECHNOLOGIES IN M&E AND IMPLICATIONS FOR EVALUATION

# Use of Big Data in Environmental Evaluation

19th Meeting of the DAC Network on Development Evaluation

26-27 April 2016

Paris, France

**Juha Uitto**

Director, GEF IEO

# WHAT WE WILL TALK ABOUT

- What is big data?
- Why do we want big data in evaluation?
- What questions can we answer with big data?
- Challenges, limitations and lessons from using big data

# What is BIG DATA?

- data sets that are so large or complex that **traditional data processing** applications are **inadequate**
- Characterized by
  - **Volume** from various sources needing large storage
  - **Velocity** at which they are generated
  - **Variety** of unstructured formats needing additional processing
  - **Value** or meaning not immediately apparent

*Adapted from Laney 2001, [www.oracle.com](http://www.oracle.com) and [www.sas.com](http://www.sas.com)*

# Why use BIG DATA in evaluation?

- Scarcer financial resources
  - Need to target interventions where most needed
- Greater demand for transparency and country ownership
  - Need objective evidence base for decision-making
  - SDGs: 17 goals, 169 targets and 304 proposed indicators

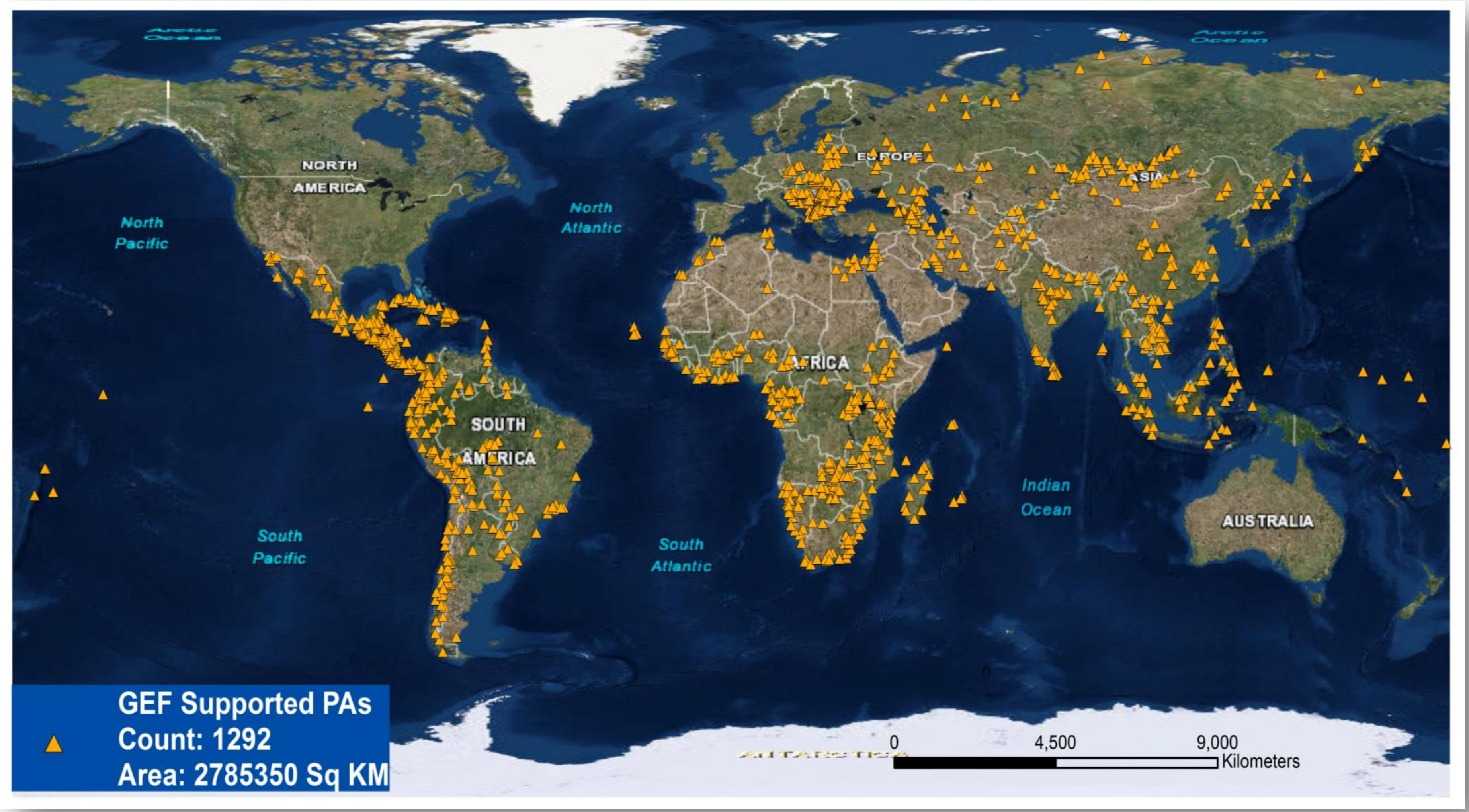
# SDGs and Earth Observation


*European Space Agency*

Big data such as from satellite imagery and sensor networks make environment and development indicators increasingly measurable

# What can BIG DATA tell us?

- Where are the funds going?
- Is funding going to the right places?
- What change occurred over time?
- Did the intervention cause the change?
- What other factors might have led to the outcome?

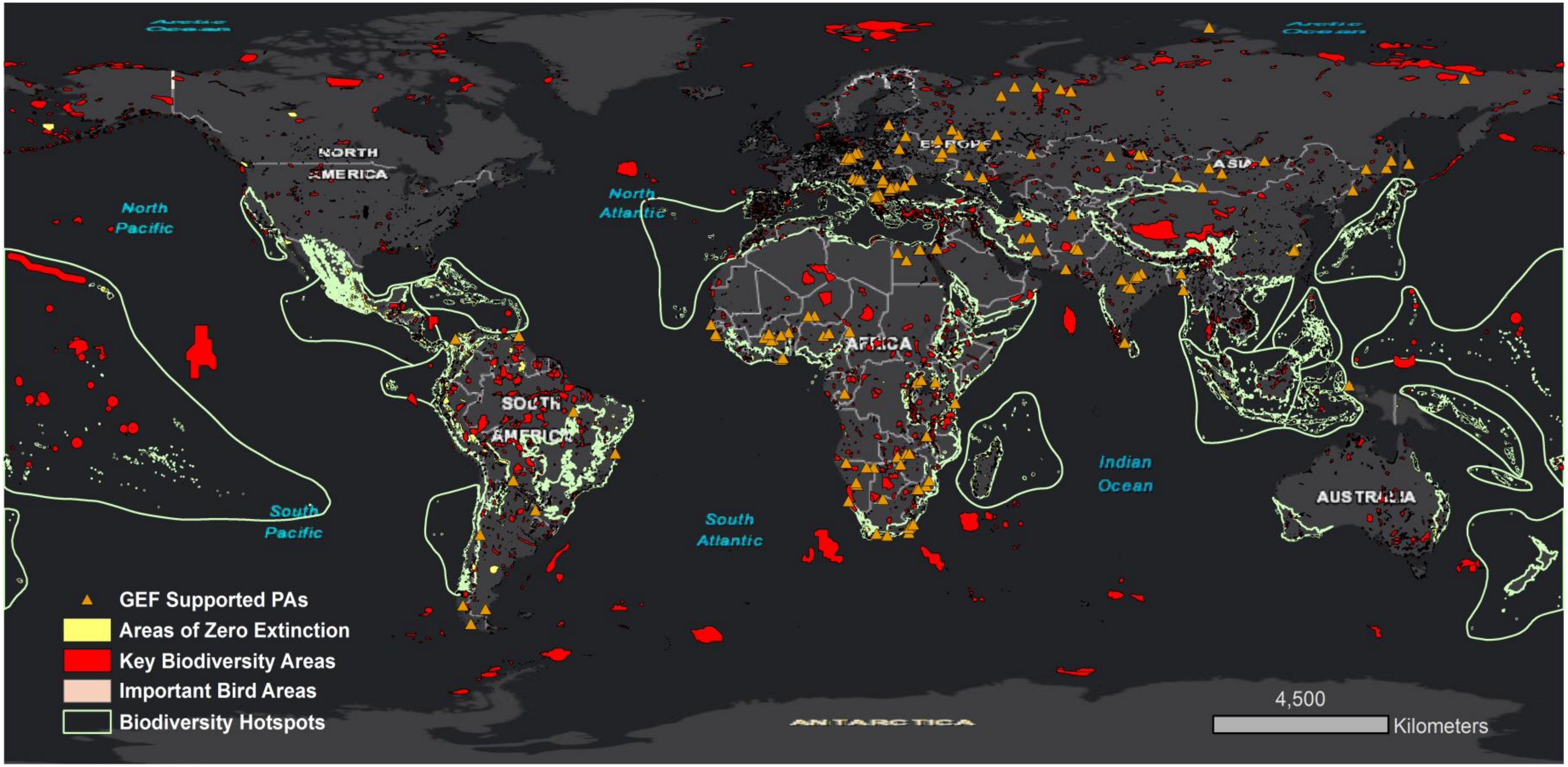



# Where are the funds going?

## Visualization of geographical context

1292 GEF-supported protected areas

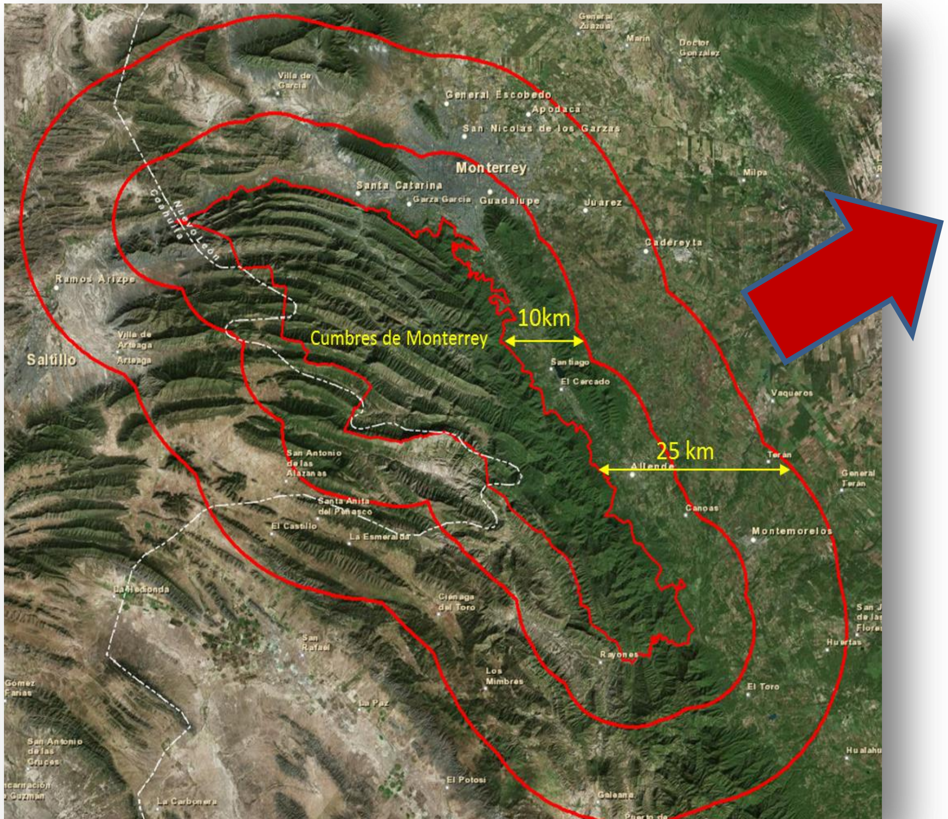
~2.8 million km$^2$ in 137 countries

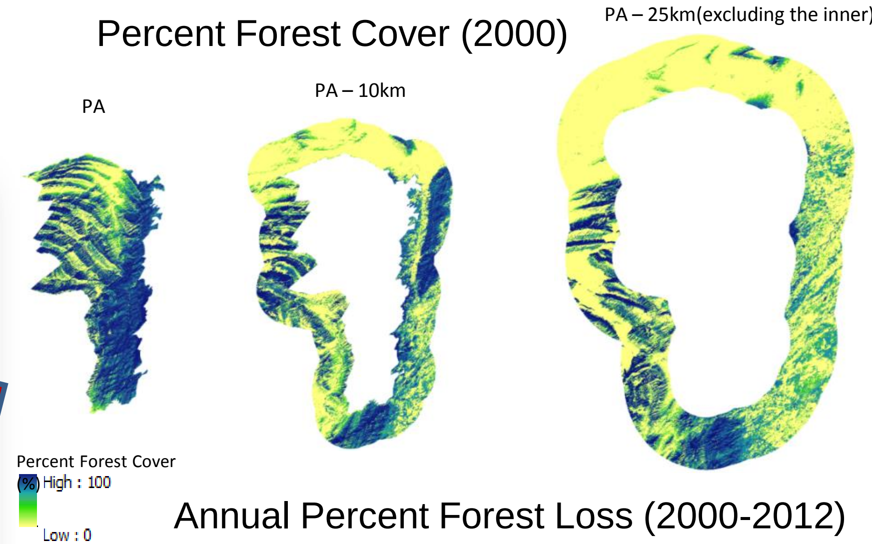


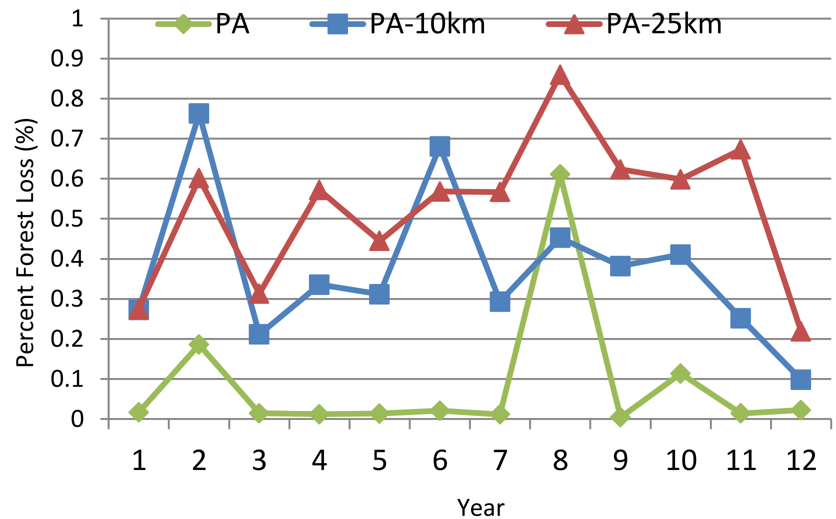

# Is funding going to the right places?

## Overlay of project sites with scientific criteria

Use of global datasets + GIS analysis to determine overlaps of GEF support with critical sites



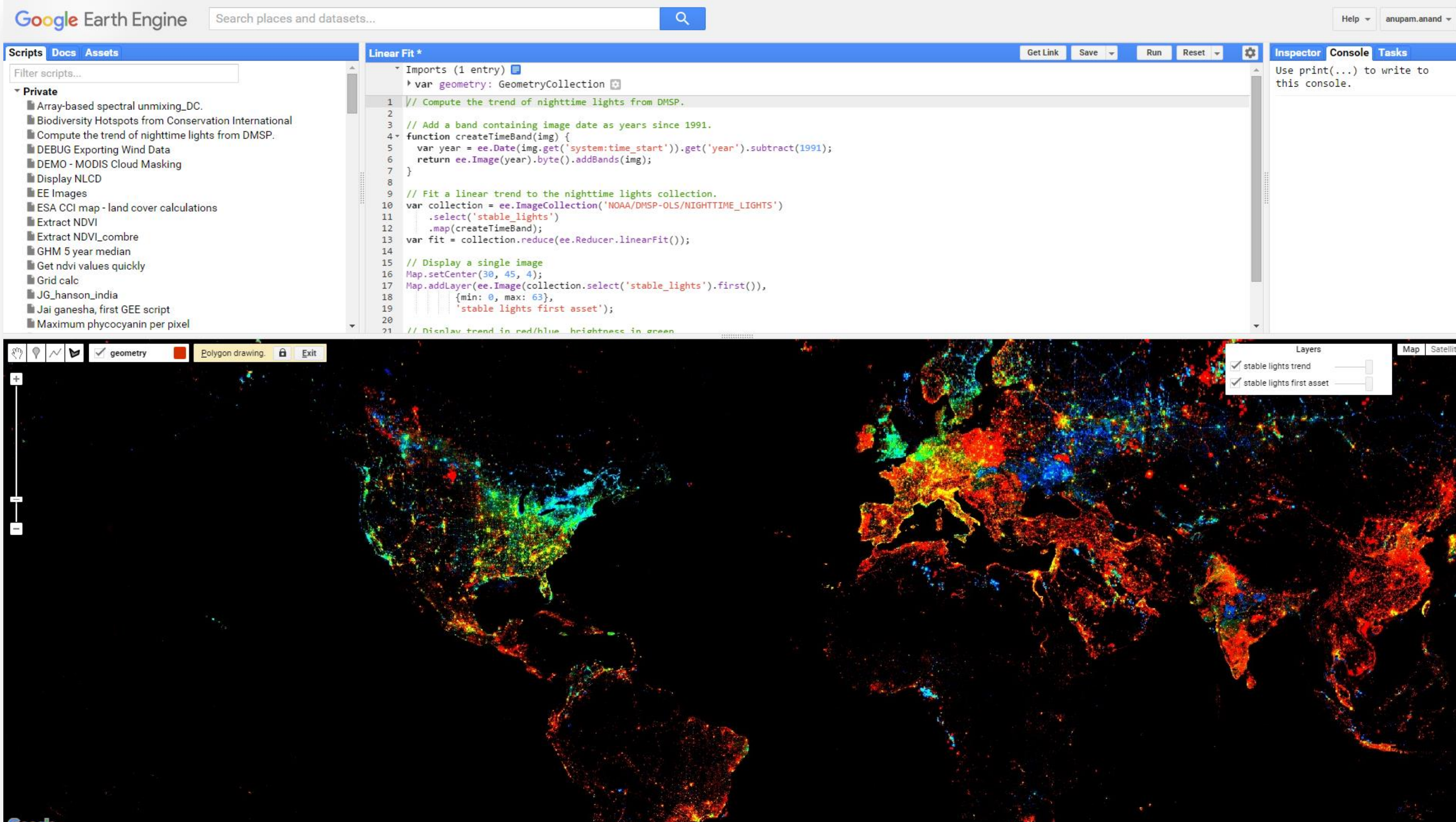




# What change occurred over time?

## Analysis of forest cover change

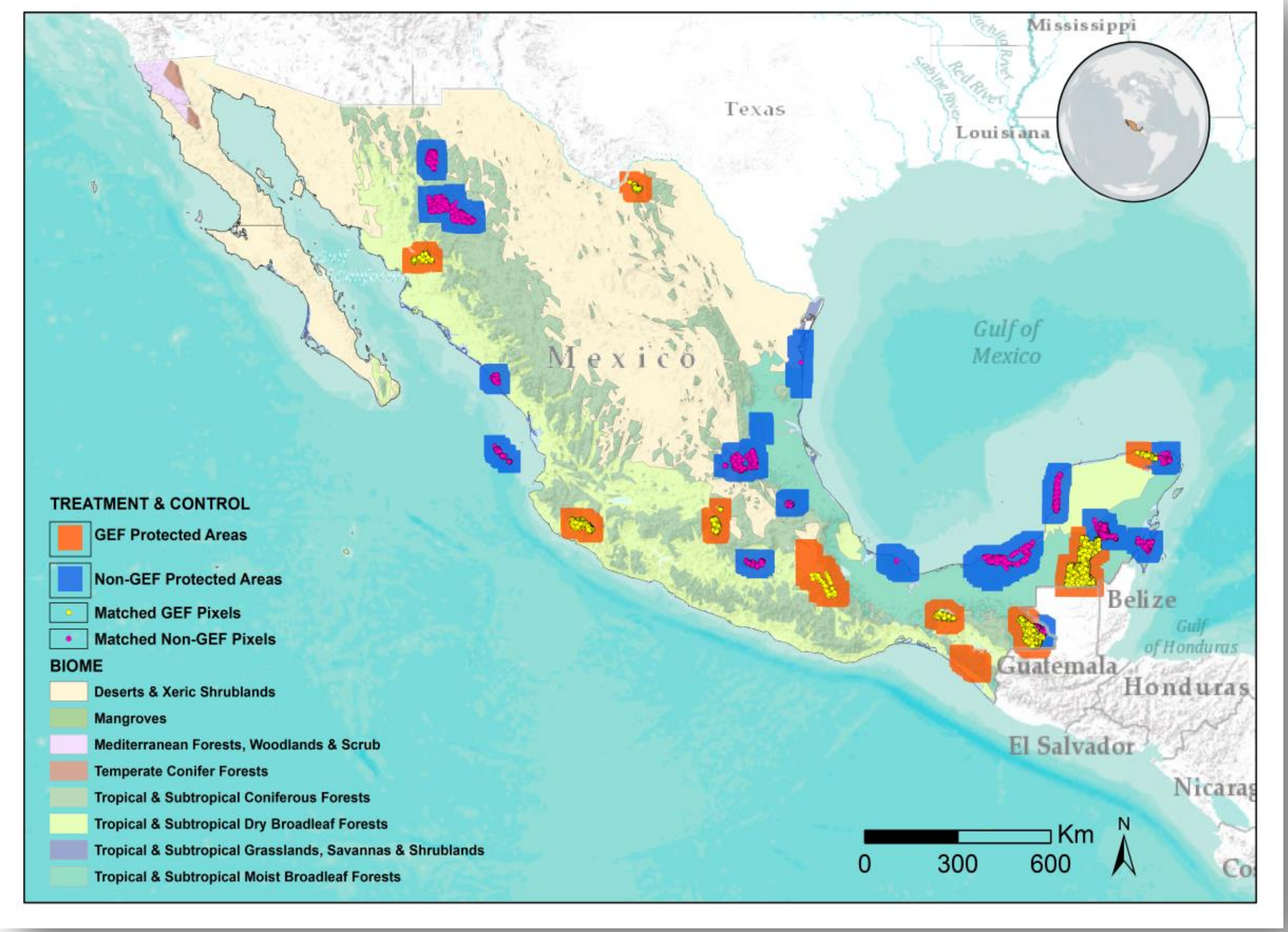
Extraction of satellite data for 30,000 GEF and non-GEF sites

30-m resolution (LANDSAT) for 12-year period




# Planetary level cloud computing with Google Earth Engine

10 years desktop computing = 7 days cloud computing




# Did the intervention cause the change?

## Quasi-experimental analysis

Propensity score matching found appropriate counterfactuals among 35,351 pixels using 9 socioeconomic and biophysical variables

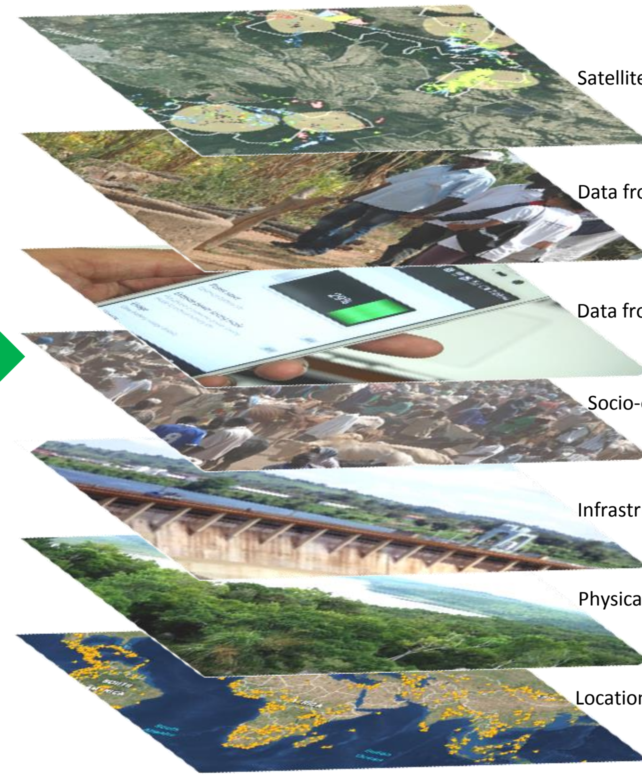
## Real World

Problem-Driven

**To assess**

- ***Impacts***
- ***Causes***
- ***Trends***

## GIS Model

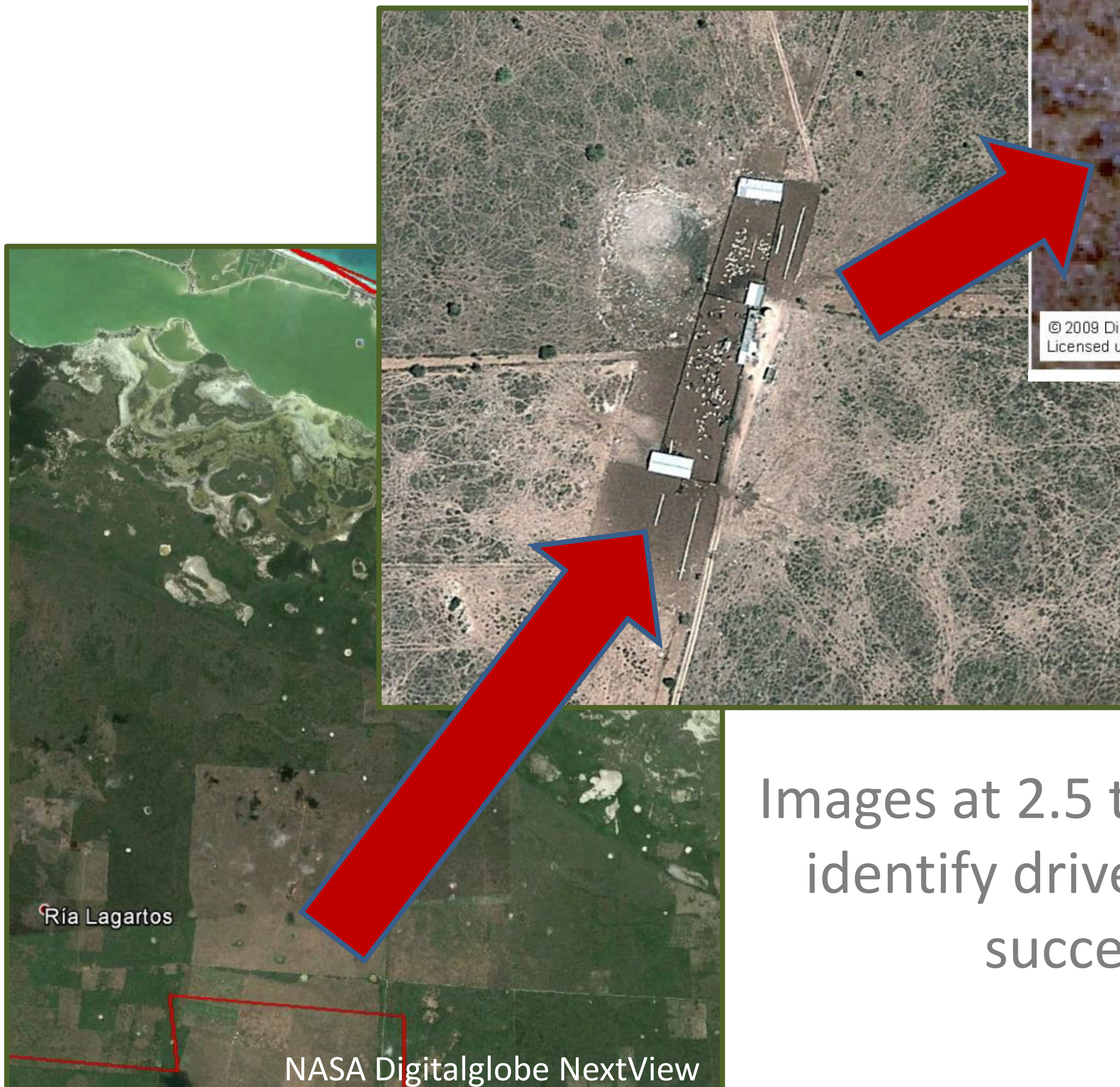


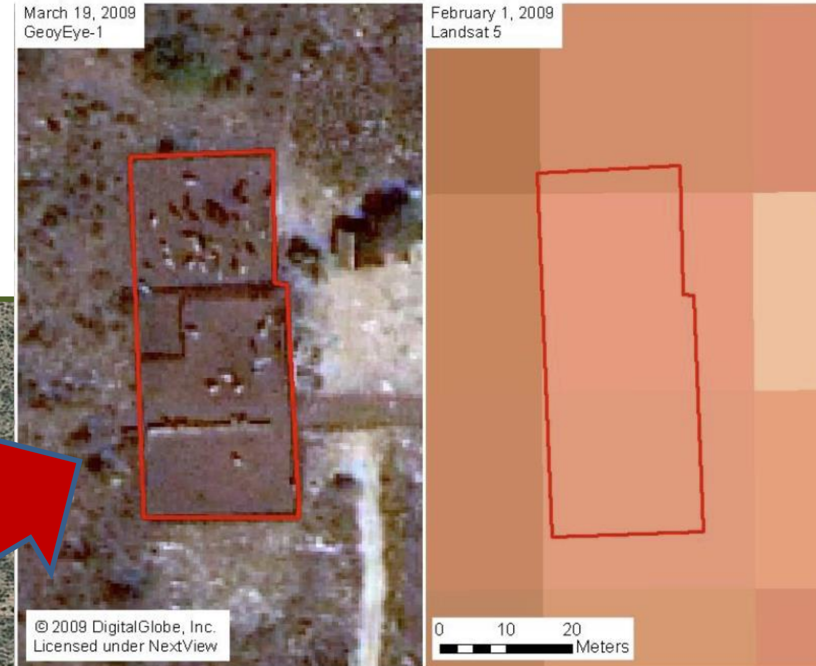

# What other factors might have led to the outcome?

## Machine learning and modelling

Data-hungry algorithms required multiple global datasets of contextual variables in different formats to assess correlations with changes

# Analysis of high-resolution images




2.5 m

30 m zoomed in to 2.5 m

Images at 2.5 to 0.5 m resolution used to identify drivers of change that hinder success of GEF support

# Challenges and Limitations

- High computing power and technical skills needed
- Uneven availability and accuracy of contextual variables
  - often vary widely across countries and sites
- Cannot answer “how” and “why” questions
- Data only as good as available resolution
  - still need to do field verification/ groundtruthing
- Still need to account for possible biases in data collection methods

# Solutions and Lessons

- Partner with global institutions with access to and infrastructure for using big data
- Used mixed approaches and methods
  - complemented global analyses with case study and portfolio analyses to triangulate findings
- Continue exploring use of new technology
  - drones, deep learning, internet of things, social media analysis, etc.
- Approach evaluation as a dynamic learning process
  - new data sets, approaches, issues will always emerge!

# GEF Independent Evaluation Office with partners

- Global Land Cover Facility, University of Maryland
- WCPA-SSC Joint Task Force on Biodiversity and Protected Areas at IUCN
- National Aeronautics and Space Administration (NASA)
- AidData