Using Satellite Imagery in Economics Research

Joel Ferguson

GAIN GSS
November 30, 2023

Roadmap

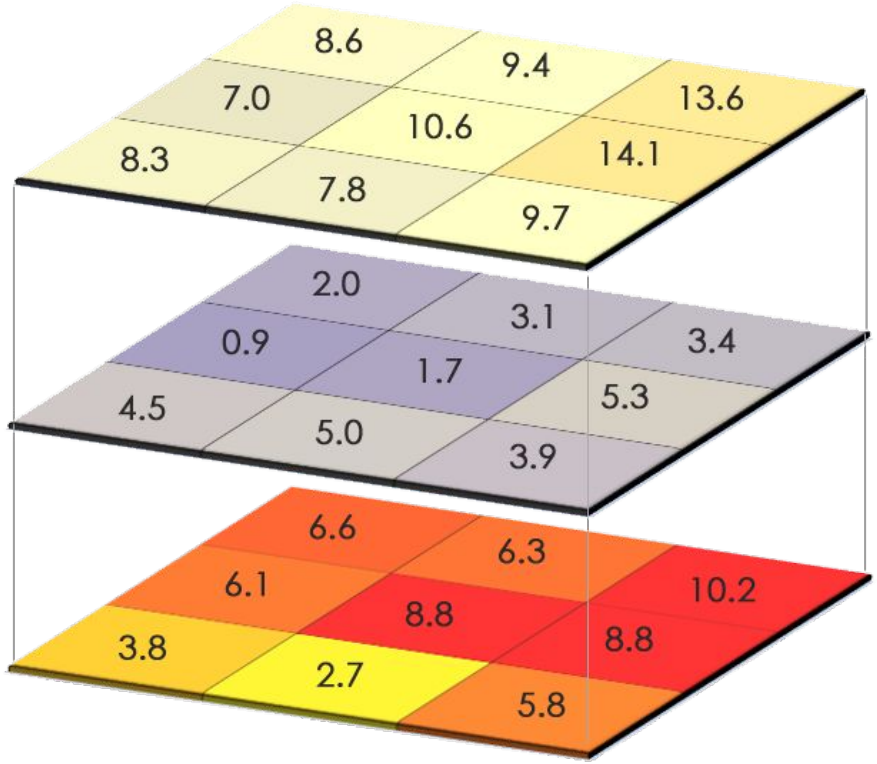
- Satellite Imagery Basics
- Overview of Available Sensors and Typical Applications
- Accessing and Working with Satellite Imagery
- Notable Pitfalls of Satellite Imagery
- Examples from my Research

Basics

Satellite Imagery Raster Data

- Easiest to think of satellite imagery as **raster data**: stacked pixels with associated coordinate reference system + transform

- Different layers correspond to different **bands**: portions of the electromagnetic spectrum



Source: GISGeography

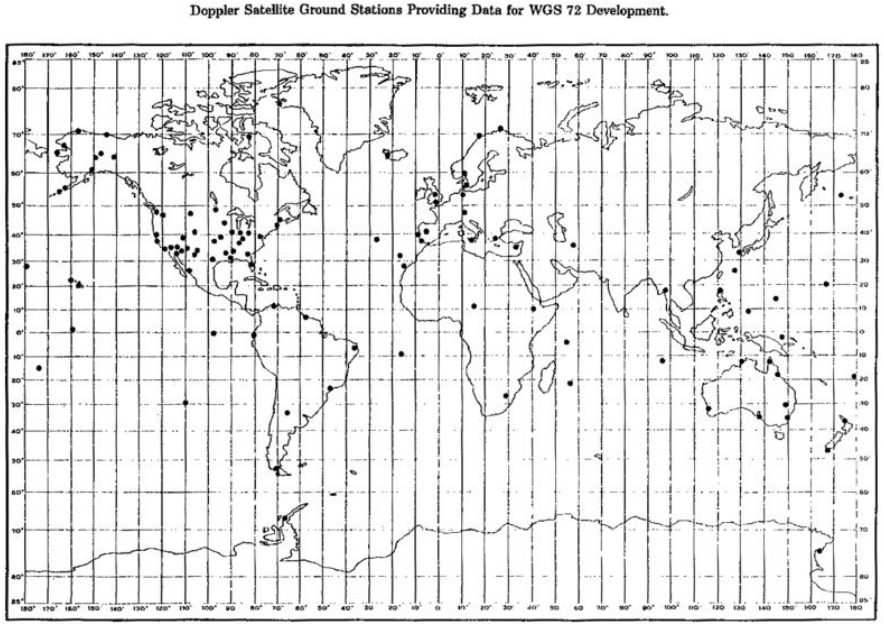
Bands of Landsat 8

Name	Units	Min	Max	Scale	Offset	Wavelength	Description
SR_B1		1	65455	2.75e-05	-0.2	0.435-0.451 µm	Band 1 (ultra blue, coastal aerosol) surface reflectance
SR_B2		1	65455	2.75e-05	-0.2	0.452-0.512 µm	Band 2 (blue) surface reflectance
SR_B3		1	65455	2.75e-05	-0.2	0.533-0.590 µm	Band 3 (green) surface reflectance
SR_B4		1	65455	2.75e-05	-0.2	0.636-0.673 µm	Band 4 (red) surface reflectance
SR_B5		1	65455	2.75e-05	-0.2	0.851-0.879 µm	Band 5 (near infrared) surface reflectance
SR_B6		1	65455	2.75e-05	-0.2	1.566-1.651 µm	Band 6 (shortwave infrared 1) surface reflectance
SR_B7		1	65455	2.75e-05	-0.2	2.107-2.294 µm	Band 7 (shortwave infrared 2) surface reflectance
SR_QA_AEROSOL							Aerosol attributes

Source: Google Earth Engine

Coordinate Reference Systems & Transforms

- **Coordinate Reference System (CRS):** Rule for assigning coordinates to location on Earth
- **Transform:** Assigns pixel corners coordinates
- Together, a CRS and a Transform allow us to map rasters

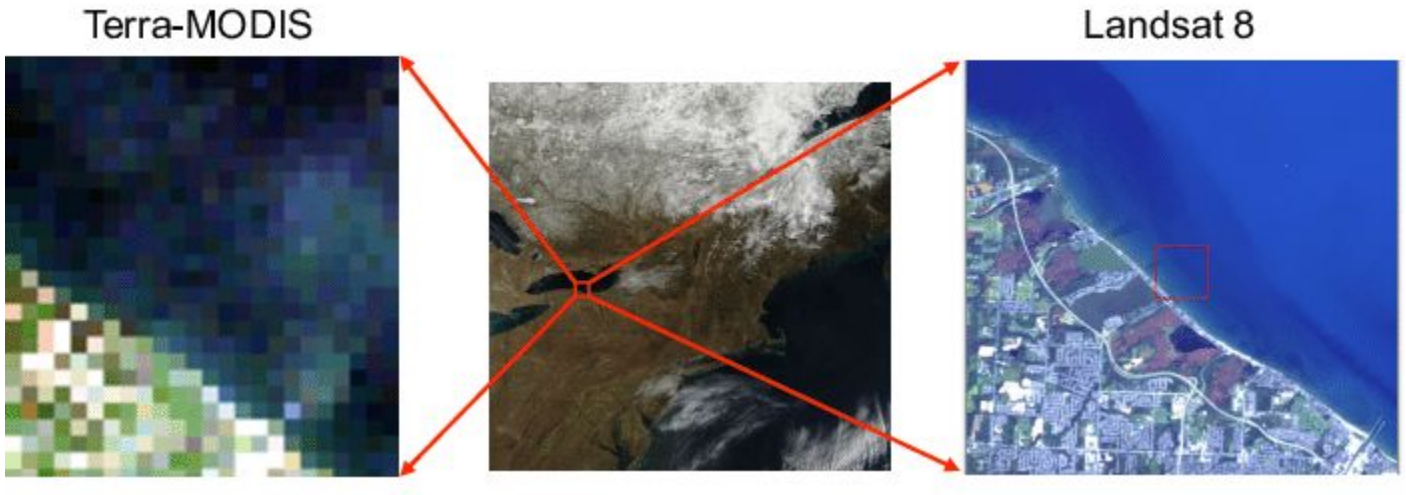


Source: NOAA

Spatial Resolution and Revisit Rate

- **Spatial Resolution:** Size of pixels in space
  E.g. Landsat pixels are 30m x 30m

- **Revisit Rate:** How often a picture is taken of a given location, also called "Temporal Resolution"



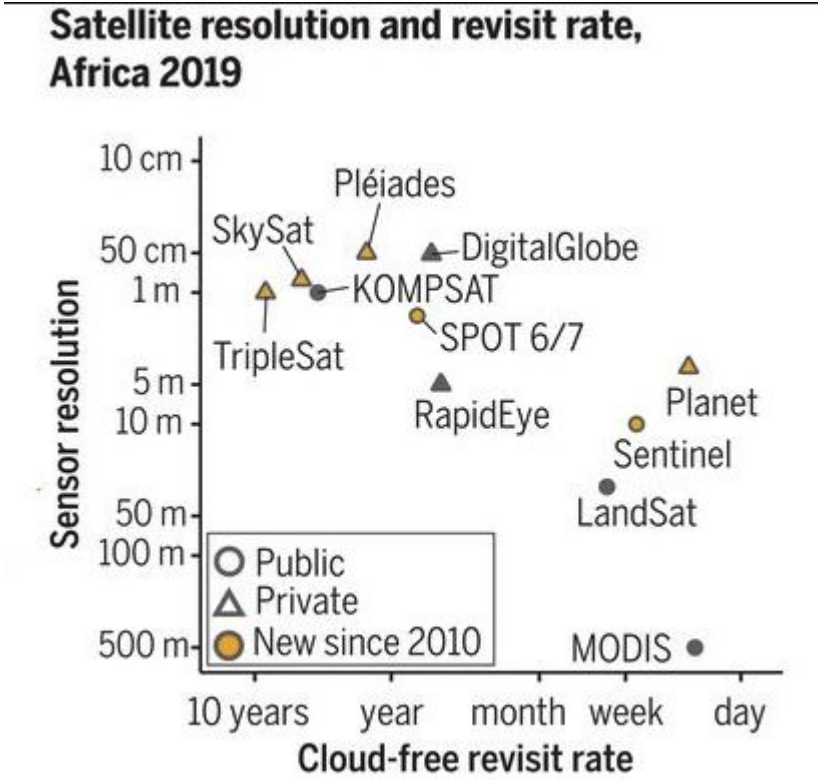
Source: Concha and Schott (2014) IGARSS

Sensor Overview

Sensor Overview

Broadly, three kinds kinds of imagery widely used in economics research:

- Night-time lights
- Medium resolution multispectral (e.g. Landsat)
- High resolution visible spectrum (e.g. Planet)



Source: Burke et al (2021) Science

Nighttime Lights

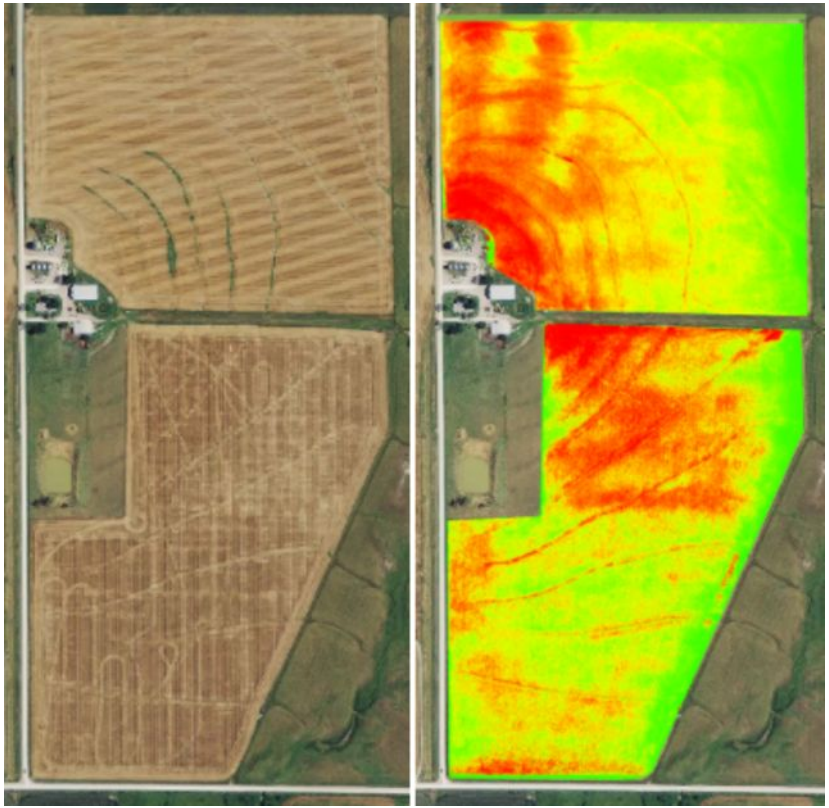
- Two sensors: DMSP 1992-2013 and VIIRS 2012-Present
- Widely used as a proxy for economic output
- Notably: does not work well for very low-income locations

Source: Yao (2019) IMF Finance & Development

Multispectral Imagery

- Imagery with multiple bands, often with some outside of the visible spectrum is referred to as **multispectral**

- Most common multispectral measure: **NDVI**, a proxy for biomass, (near infrared - red)/(near infrared + red)



Source: botlink

Other measures derived from Multispectral Imagery

- Aerosol Optical Depth (AOD): proxy for air pollution
- Normalized Difference Flood Index (NDFI): measures standing water
- Enhanced Vegetation Index (EVI): A measure more closely associated with crop biomass than NDVI
- Thermal Anomalies: Used for detecting active fires
- Optical Trapezoidal Model (OPTRAM): measure of soil moisture

High Resolution Visible Spectrum/Panchromatic Imagery

- **Visible Spectrum:** Red, Blue, and Green bands used to make imagery similar to that displayed by our computers
- **Panchromatic:** Greyscale imagery
- Used for object detection and related tasks
- Sources: Google Static Maps, Planet, Spot

Source: Suri et al (2019) AEJ:AE

Other Lesser-Known Imagery

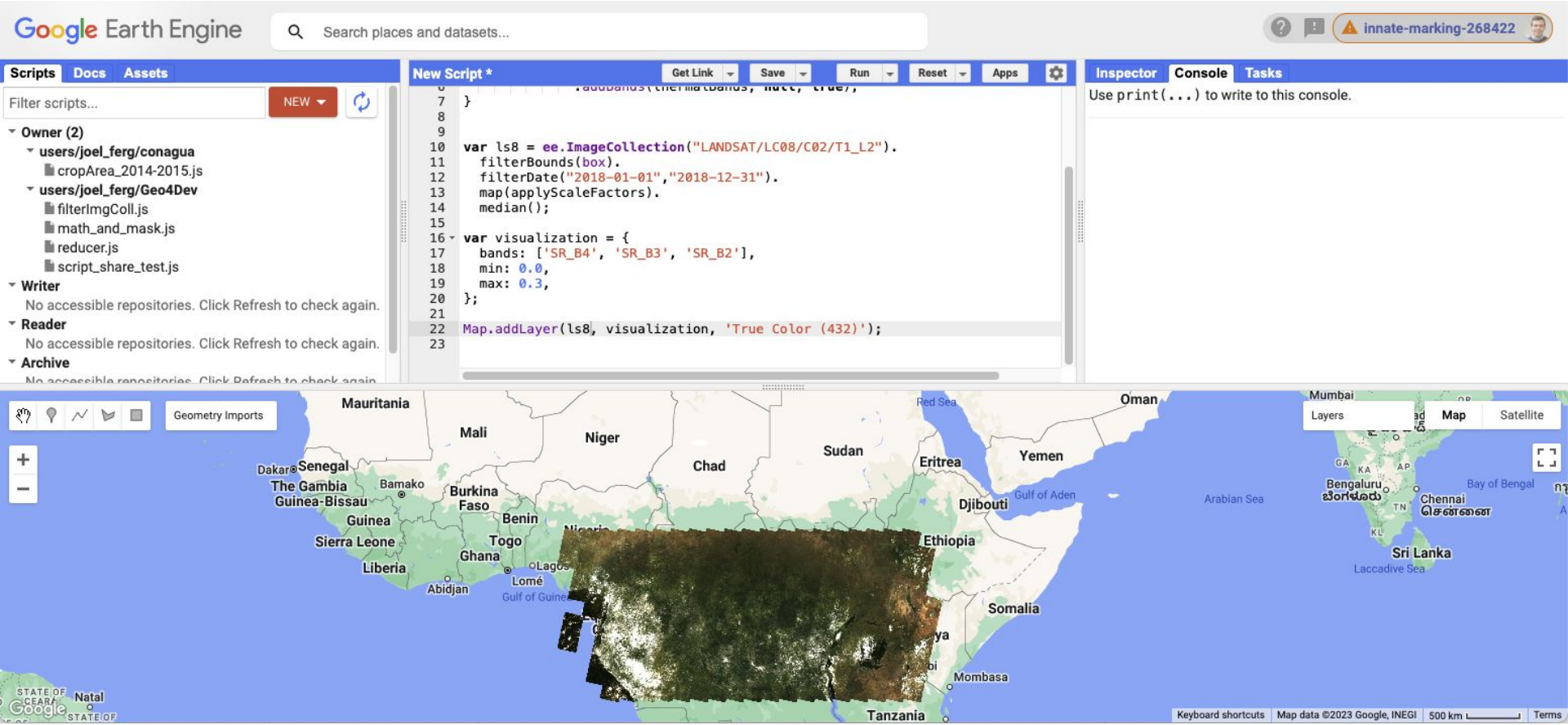
- **Hyperspectral:** Many (often 100+) bands spanning a wide range of electromagnetic spectrum

- **Radio:** Used for measuring topography (SRTM) and changes in groundwater (GRACE)

- **LIDAR:** Used to construct 3D representations of features on the ground (e.g. buildings)

Accessing Imagery

Google Earth Engine

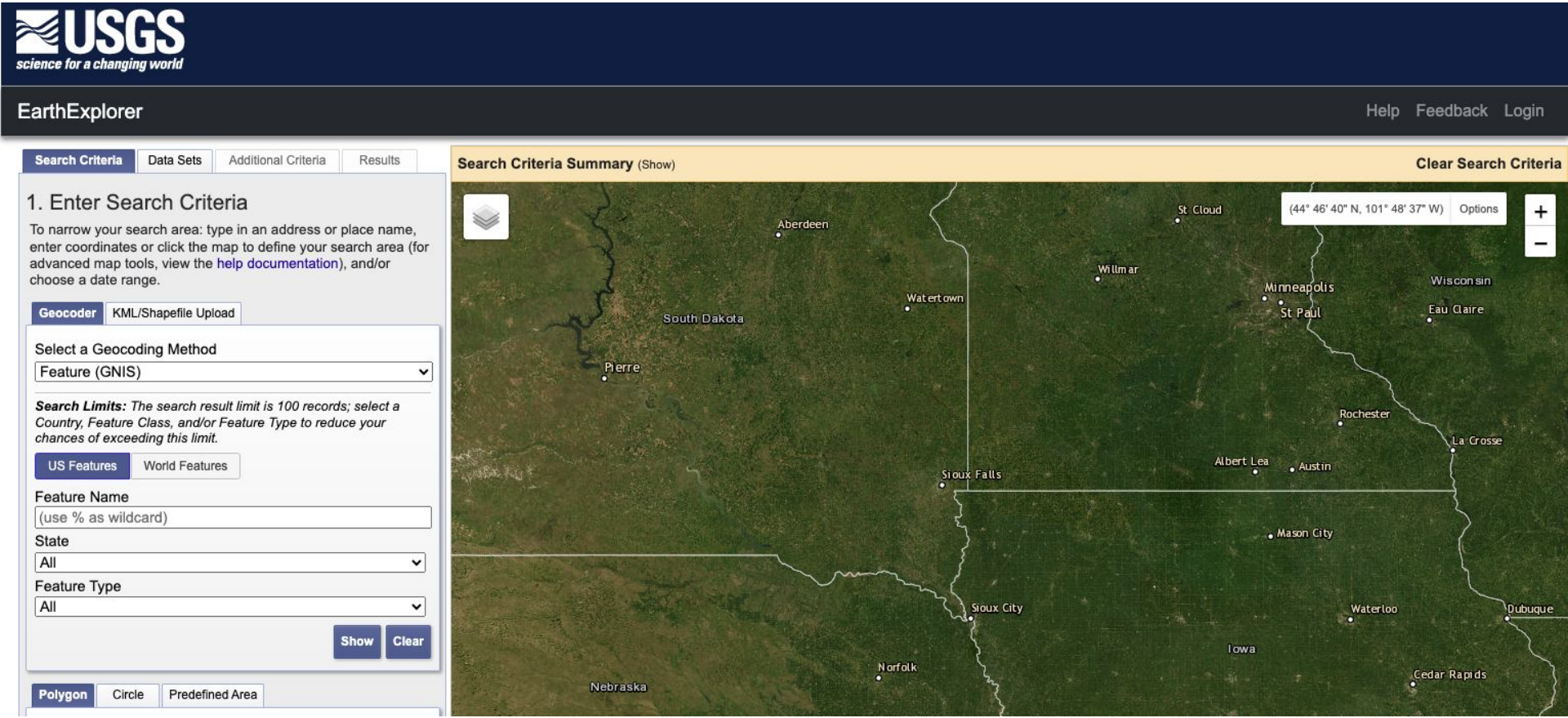


Source: Google Earth Engine

Google Earth Engine

- Browser uses Javascript, Python API also available
- Large collection of publicly available imagery and other rasters such as weather data
- Free Google Cloud compute, operations optimized to run quickly on servers

USGS EarthExplorer



Source: USGS EarthExplorer

Other Useful Tools

Python

- **Rasterio:** Working with raster data
- **Geopandas:** Working with vector data (e.g. shapefiles)
- **Pillow:** Working with non-georeferenced imagery

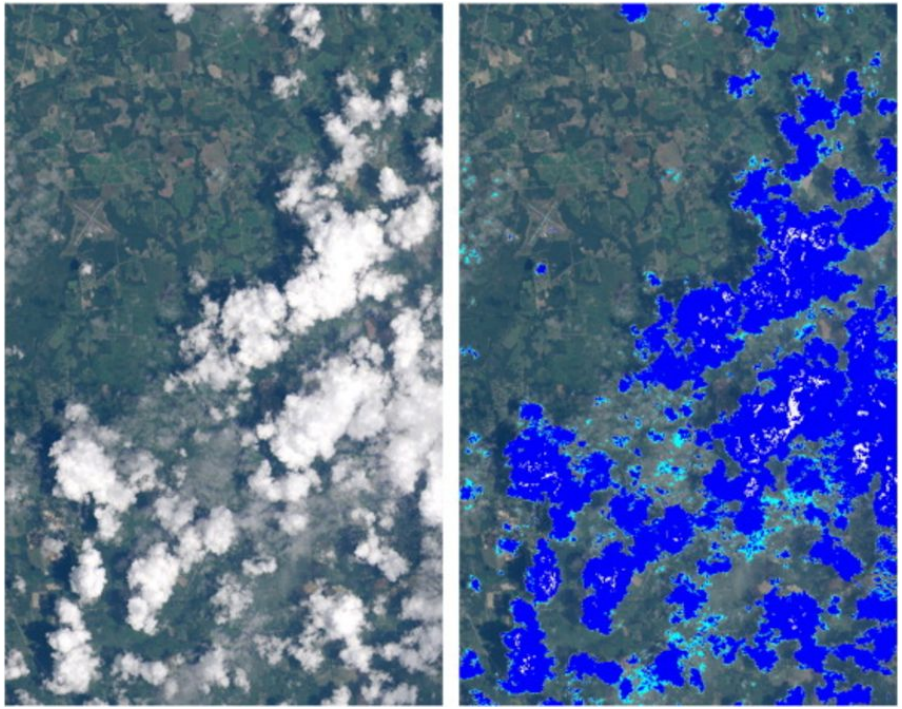
Other

- **MOSAIKS:** Satellite imagery-derived features which can be used as tabular data

Pitfalls

Clouds

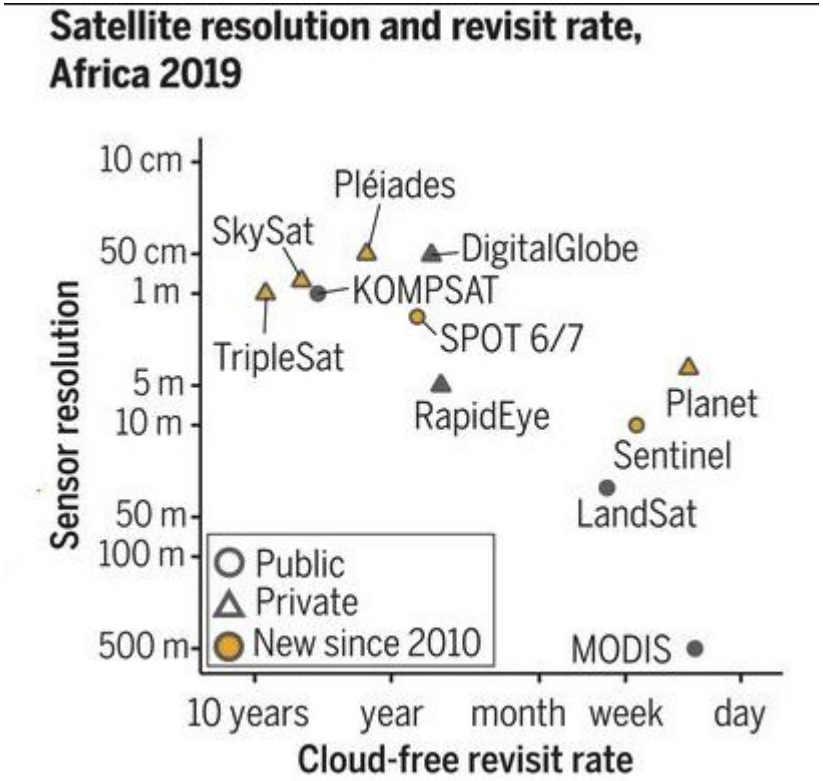
- Clouds lead to measurement error in most indices we care about
- Hard to reliably filter clouds, leads to different composition of timing of imagery over space
- Some places are especially cloudy. May need 3 years of Landsat to make a cloud-free composite in parts of East Africa!



Source: Roy et al (2018) RSE

Spatial vs Temporal Resolution and Temporal Span

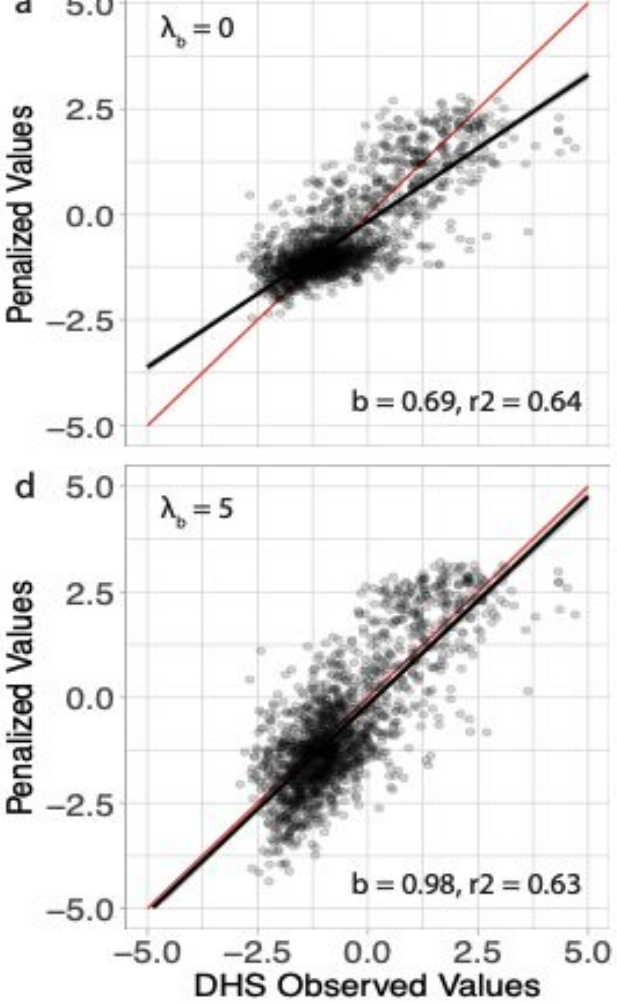
- Typically, there is a tradeoff between spatial resolution and revisit rate
- Older satellites are typically lower resolution. High resolution imagery much more available starting 2016
- Private imagery is expensive! $15k for basic Planet service



Source: Burke et al (2021) Science

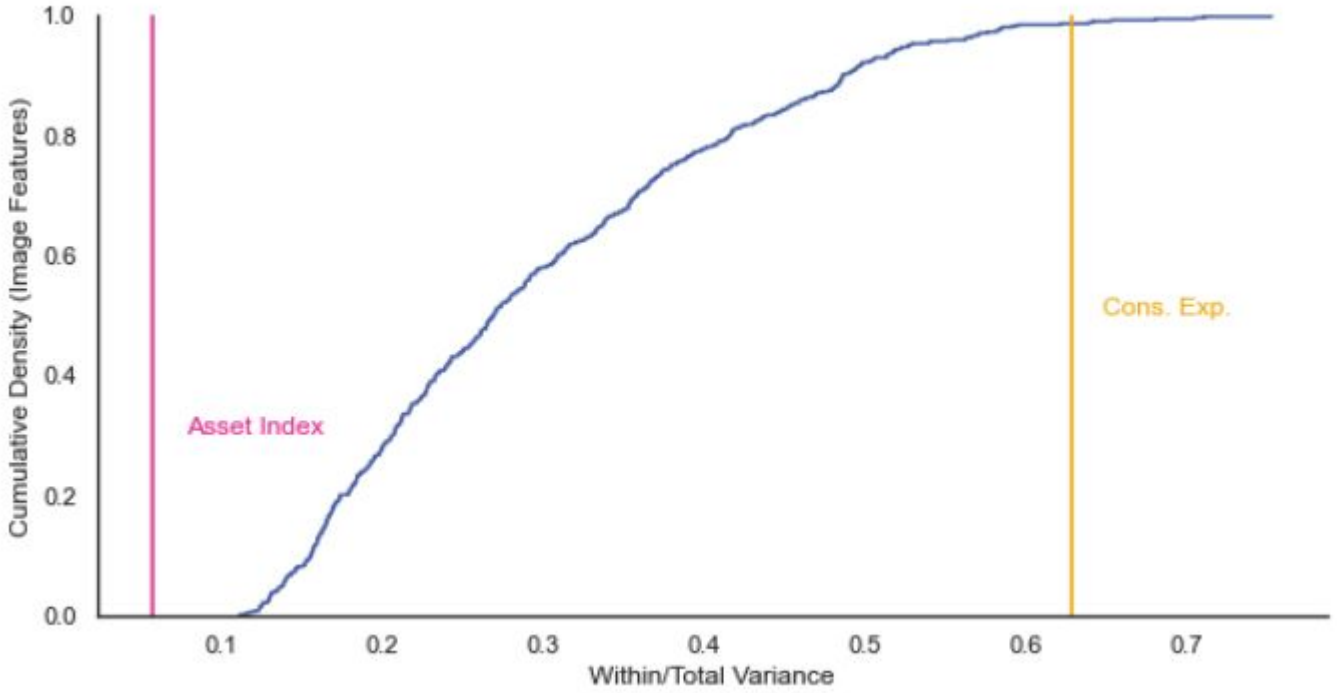
Non-Classical Measurement Error

- Often want to use satellite imagery to make a proxy for an outcome of interest
- Recently, there's been a lot of worry that the mismeasurement of outcomes leads to biases
- Typical case for continuous outcomes: Berkeson prediction leads to attenuated estimates
- Solutions: Bias-penalizing loss (Ratledge et al 2022), Multiple imputation (Carleton et al 2023)



Source: Ratledge et al (2022) Nature

Examples

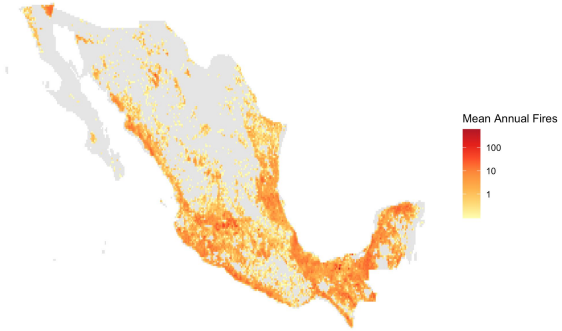


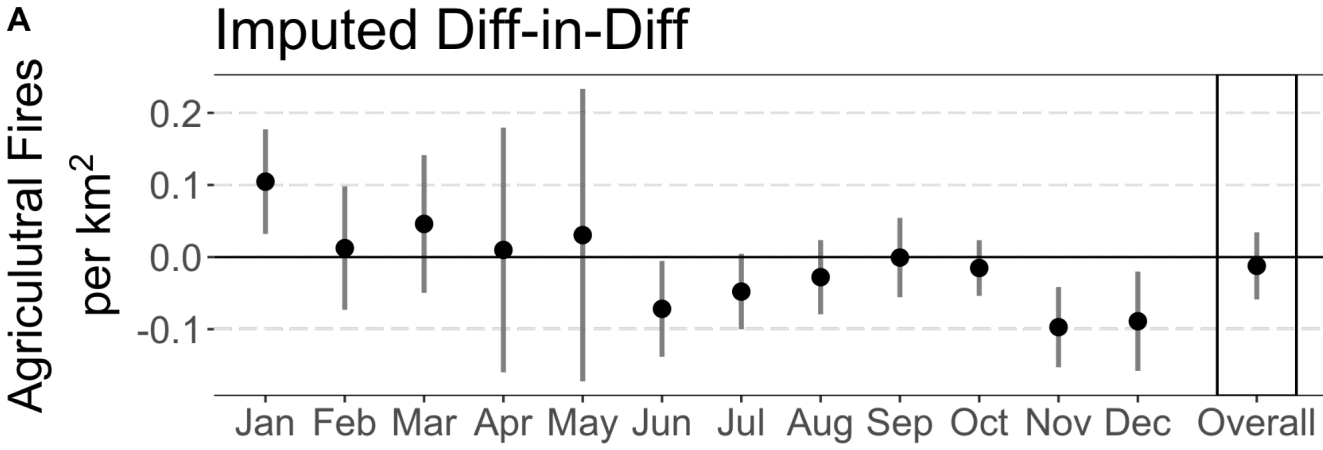
Christensen, Ferguson, and Ramírez Amaya (2021) *NeurIPS ML4D*

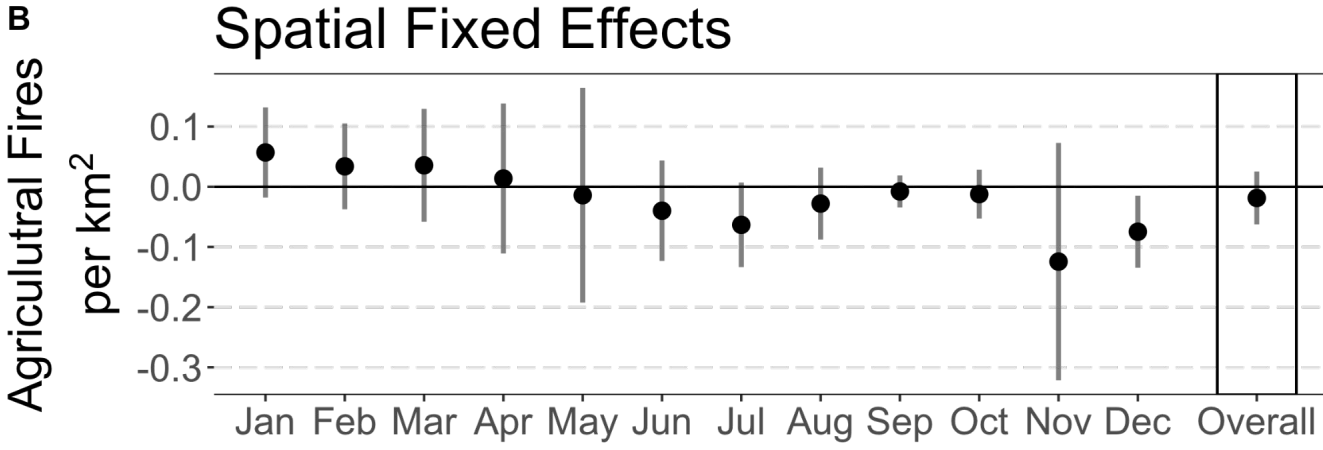
Features of Multispectral imagery that are useful for predicting asset wealth have limited within-location variation, making it hard to track changes in welfare over time

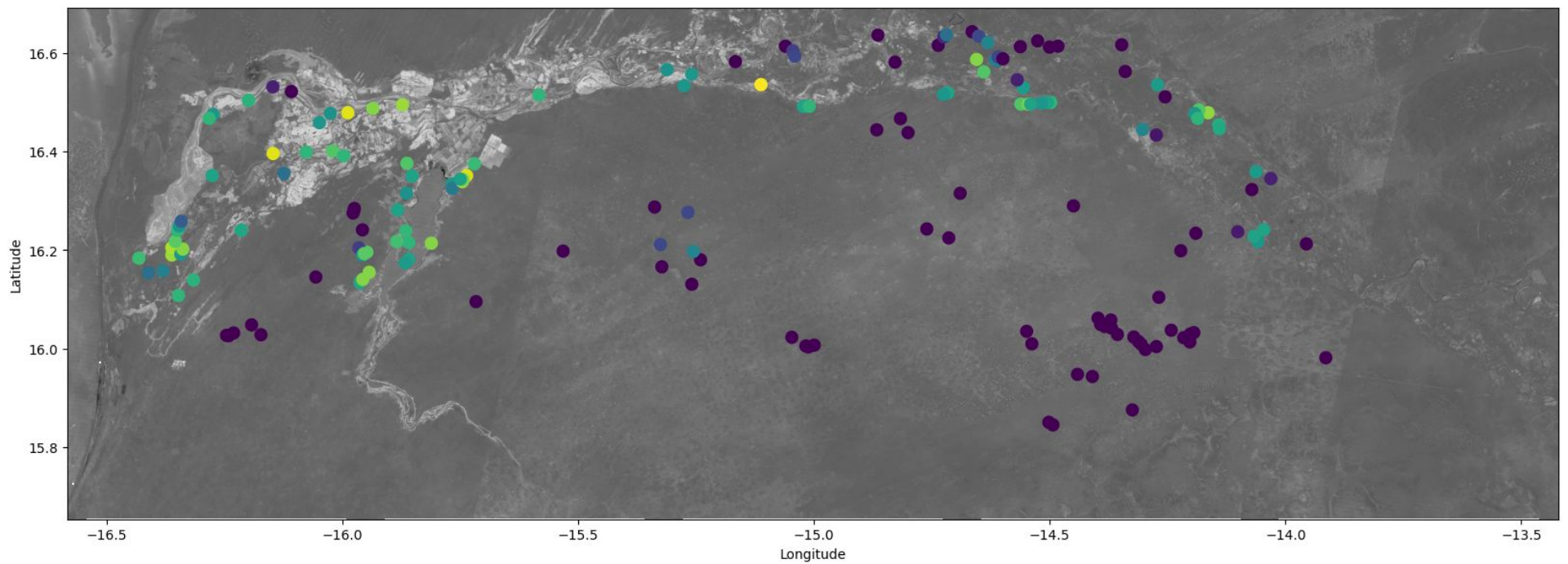
Ferguson & Govaerts (2023) Working Paper

Conservation Agriculture adoption in Mexico leads to reductions in agricultural burning





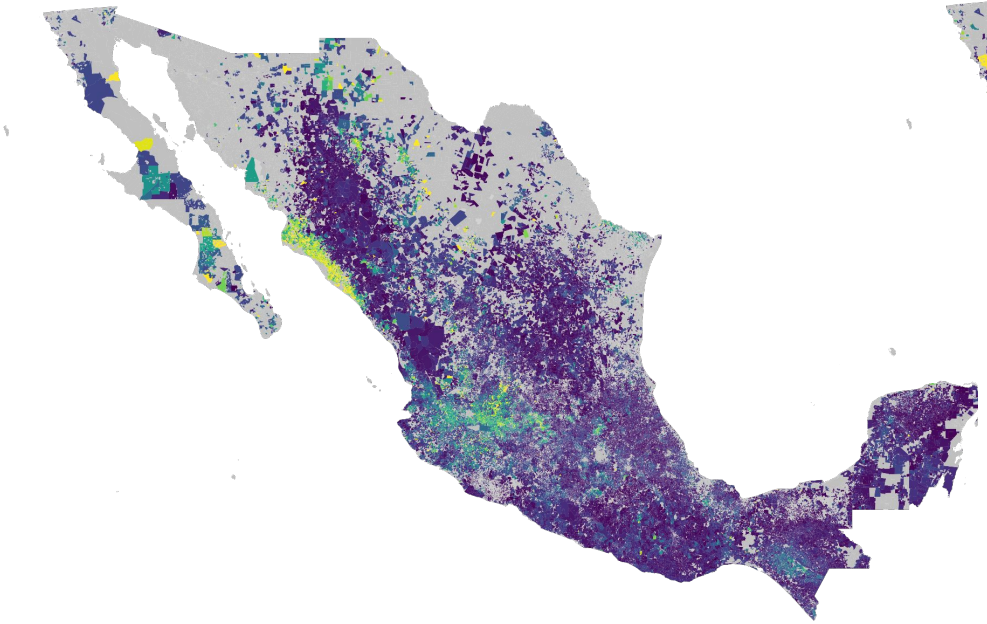




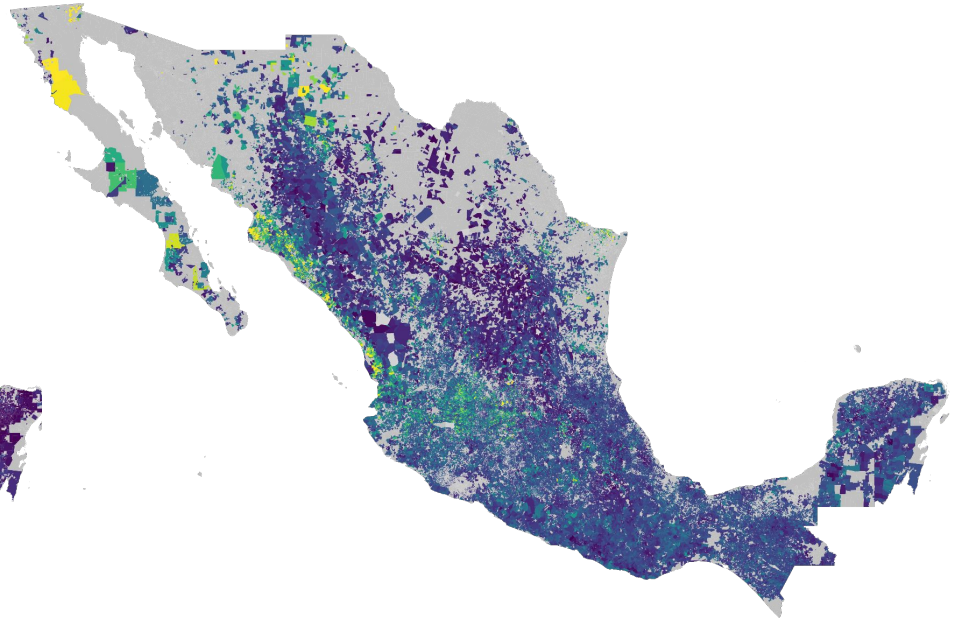
Cisse, De Janvry, Ferguson, Gonzalez-Navarro, Mbaye, Sadoulet, and Syll (2023) Working Paper

Locations exposed to greater agricultural productivity growth as a result of irrigation infrastructure development in the Senegal River Valley experienced greater increases in economic activity

Ground truth



Prediction



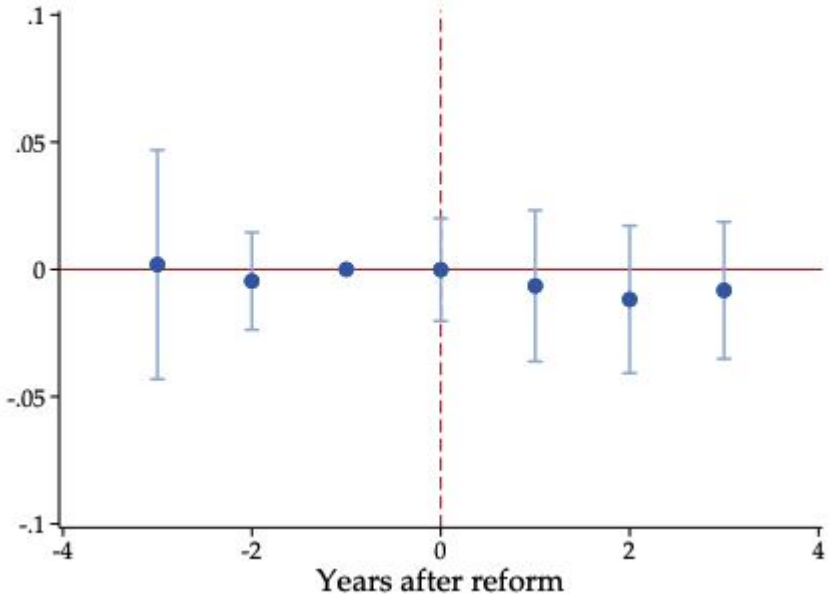
Ferguson, Sayre, and Sogomo (2023) Working Paper

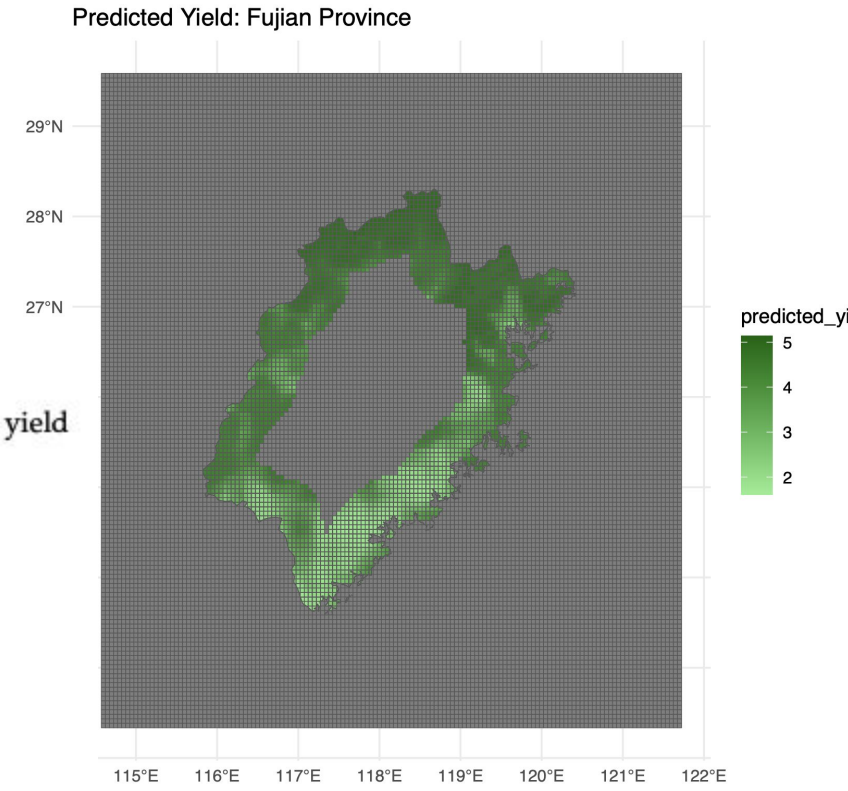
Multispectral imagery can be used to accurately downscale aggregated yield measures

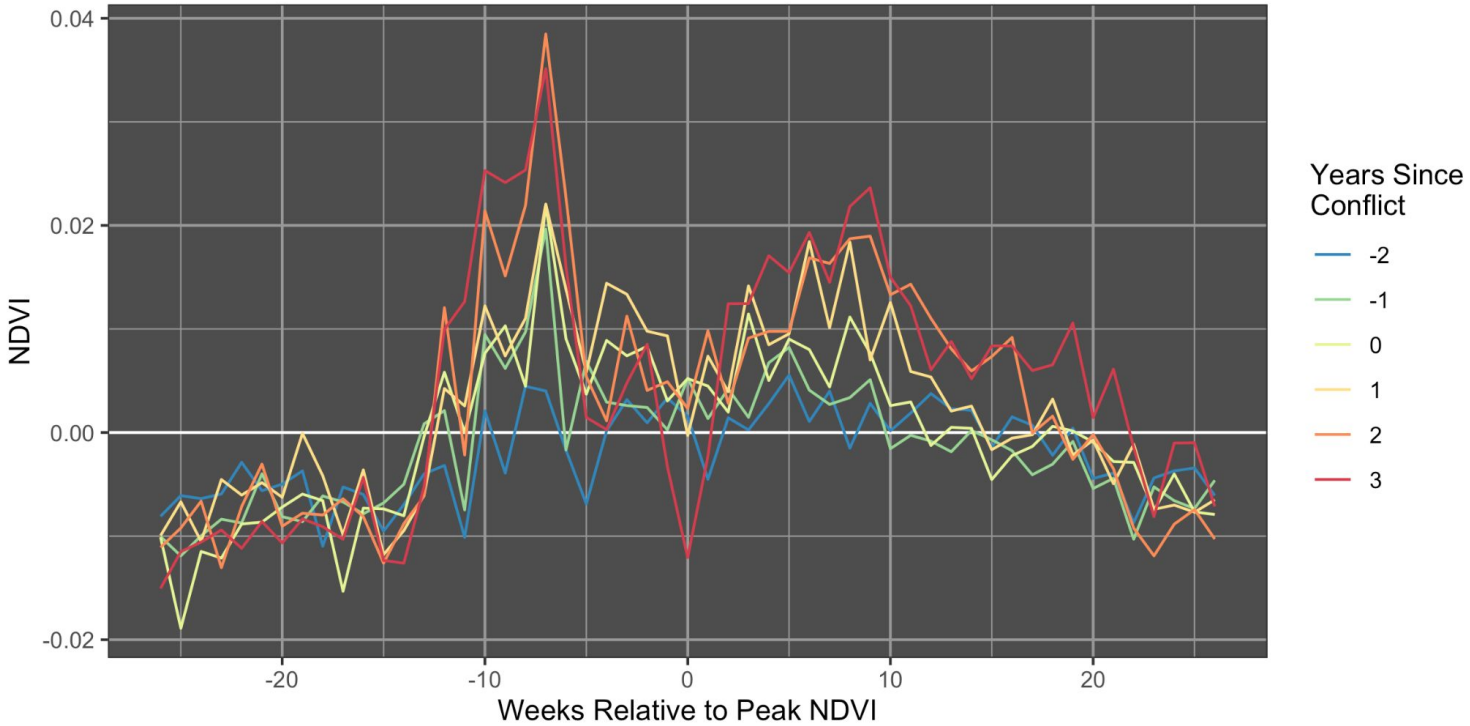
Ferguson & Kim (2023) Working Paper

Chinese Household Responsibility System had no discernible effect on agricultural productivity

Figure 7: Differences-in-discontinuities effect of HRS adoption on estimated log yield







Ferguson (2023) Working Paper

Conflict leads to land abandonment in sub-Saharan Africa

Thanks!

joel_ferg@berkeley.edu
joelferg.github.io