

# Machine Learning for policy evaluation: Introduction to remote sensing

Jérémy Do Nascimento Miguel\*

\*BSE, Univ. Bordeaux, [jdnmiguel@u-bordeaux.fr](mailto:jdnmiguel@u-bordeaux.fr)

Master APP - EADD; Univ. Bordeaux - Fall 2023

# Overview

Introduction to remote sensing

Sensor Overview

Accessing data

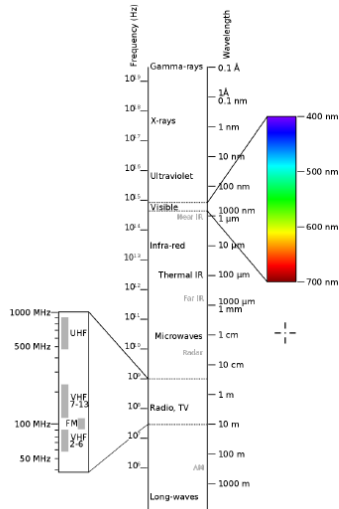
Pitfalls

# The electromagnetic spectrum

Sun's energy constantly hits the earth as electromagnetic wave

- Waves are categorized into a wide spectrum
- Only a tiny fraction is visible to the human eyes

We perceive color with specialized cone cells in the retina sensitive to different parts of the spectrum

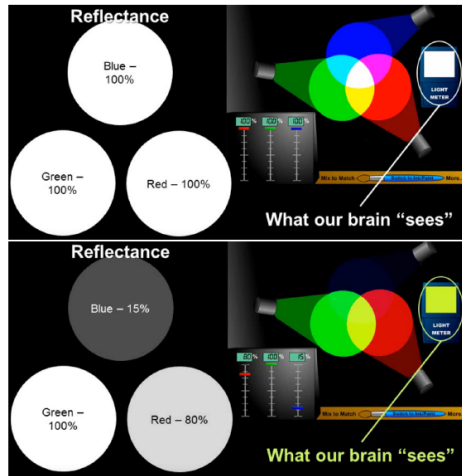


source: [https://en.wikipedia.org/wiki/Electromagnetic\\_spectrum](https://en.wikipedia.org/wiki/Electromagnetic_spectrum)

# Human vision

Combining reflectance at different intensities from different parts of the spectrum results in color vision

- Max B,R,G reflectance = white
- Intensity variation leads to color variation

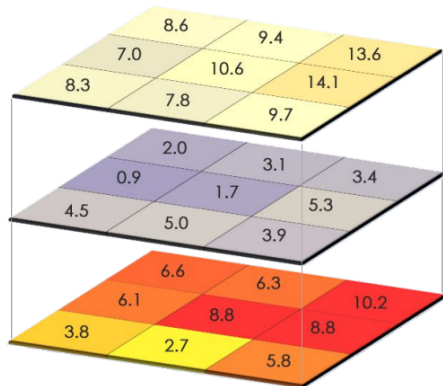


# Digital vision

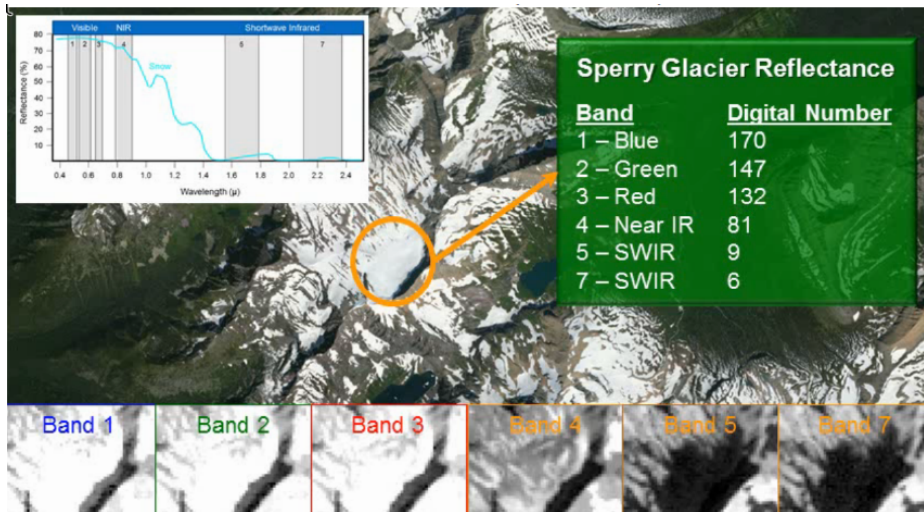
Digital cameras do exactly the same:

- Record reflectance value for each band at the pixel level (0-255)
- Different layers correspond to different bands

Satellite imagery: stacked pixels with associated coordinate reference system + transform



## An illustration: detecting snow



Snow is only "white" in bands 1-4 of the spectrum and absorbs radiation in others

# Band Combinations

## Red band:

- Absorbed by chlorophyll: little reflectance from healthy vegetation
- Absorbed by water: water bodies appear fairly dark

## Green band:

- Better reflected by water
- Strongly reflected by vegetation

## Blue band:

- Can use to differentiate soil, vegetation, and deciduous from coniferous vegetation

## Near infrared band:

- Strongly absorbed by healthy vegetation

## An example: NDVI

The Normalized Difference Vegetation Index: identify vegetation, measure its health/vitality

- Chlorophyll causes light absorption in the red part of the spectrum
- Plant leaf cells re-emit solar radiation in near-infrared part of the spectrum
- Each plant has its own spectral signature
- When vegetation is less healthy or more sparse: more red and less infrared light gets reflected

$$NDVI = \frac{NIR - RED}{NIR + RED}$$

case healthy vegetation: NDVI close to 1; if unhealthy: NIR smaller, RED larger.



## Some definition

- Coordinate reference system: rule for assigning coordinates to location on Earth. Some distorts distance and can't be use for distance calculation
- Transform: assigns pixel corners coordinates
- Spatial resolution: size of pixels in space. E.g., Landsat pixels are 30m x 30m
- Revisit rate or temporal resolution: how often a picture is taken

# Overview

Introduction to remote sensing

Sensor Overview

Accessing data

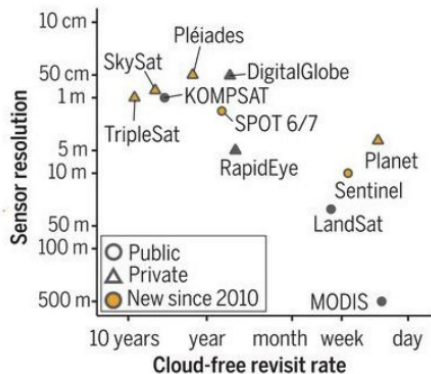
Pitfalls

## Sensor overview

Most used in economics:

- Night-time Lights
- Medium resolution multispectral: see Landsat
- High resolution visible spectrum: see Planet

**Satellite resolution and revisit rate, Africa 2019**



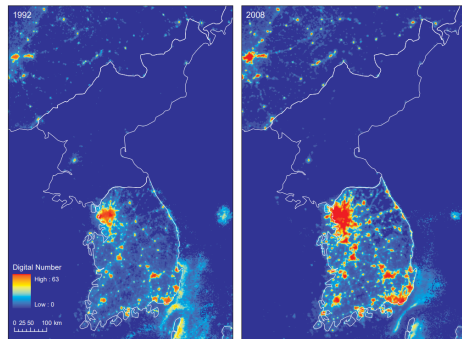
Source:

Burke et al. (2021)

# Night Lights

Widely used as a proxy for economic output

- 2 sensors: DMSP 1992-2013 and VIIRS 2012-Present
- Cons: does not work well for very developing countries and rural locations

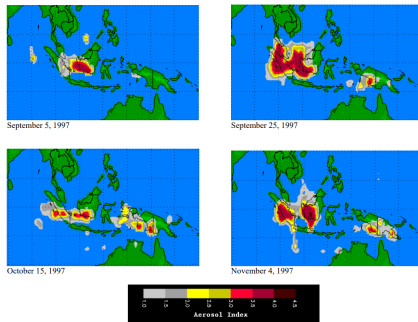


Source: Henderson et al. (2012)

# Multispectral Imagery

Imagery with multiple bands, often with some outside of the visible spectrum

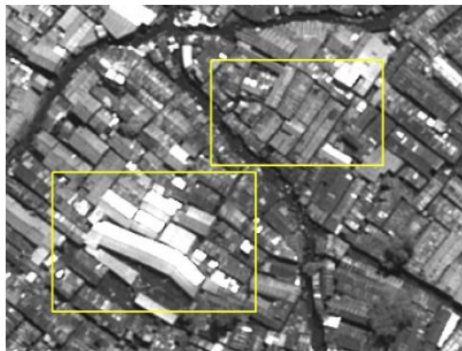
- NDVI: proxy for vegetation cover
- Aerosol Optical Depth (AOD): proxy for air pollution
- Normalized Difference Flood Index (NDFI): measures standing water
- Thermal Anomalies: Used for detecting active fires



Source: Jayachandran (2009)

# High Resolution Visible Spectrum or Panchromatic Imagery

- Visible Spectrum: Red, Blue, and Green bands used to make imagery similar to that displayed by our computers
- Panchromatic: greyscale
- Used for object detection
- For instance planet



Source: Marx et al. (2019)

# Overview

Introduction to remote sensing

Sensor Overview

Accessing data

Pitfalls

# Google Earth Engine

The image shows the Google Earth Engine web interface. At the top, there's a navigation bar with labels: 'Script manager' (pointing to the top left), 'API documentation' (pointing to the top left), 'Search for data Imports' (pointing to the search bar), 'Get a link (URL) to the script' (pointing to the 'Link' button), 'Save the script' (pointing to the 'Save' button), 'Run the script' (pointing to the 'Run' button), and 'Help button' (pointing to the 'Help' button). Below the navigation bar, the interface is divided into several panels. On the left is the 'Asset Manager' (labeled 'Asset Manager' with an arrow). In the center is the 'Code Editor' (labeled 'Code Editor' with an arrow). On the right is the 'Task manager' (labeled 'Task manager' with an arrow). Below the code editor is the 'Map' (labeled 'Map' with an arrow). To the left of the map are 'Geometry Tools' (labeled 'Geometry Tools' with an arrow) and a 'Zoom' control (labeled 'Zoom' with an arrow). On the right side of the map is the 'Layer manager' (labeled 'Layer manager' with an arrow). The 'Console output' (labeled 'Console output' with an arrow) is located below the task manager. The 'Inspect locations, pixel values, objects added to the map' (labeled 'Inspect locations, pixel values, objects added to the map' with an arrow) is located below the console output. The 'Map' panel shows a satellite view of the world with a search bar and a 'Layers' button.

Script manager

API documentation

Search for data Imports

Get a link (URL) to the script

Save the script

Run the script

Help button

Asset Manager

Code Editor

Task manager

Console output

Inspect locations, pixel values, objects added to the map

Layer manager

Geometry Tools

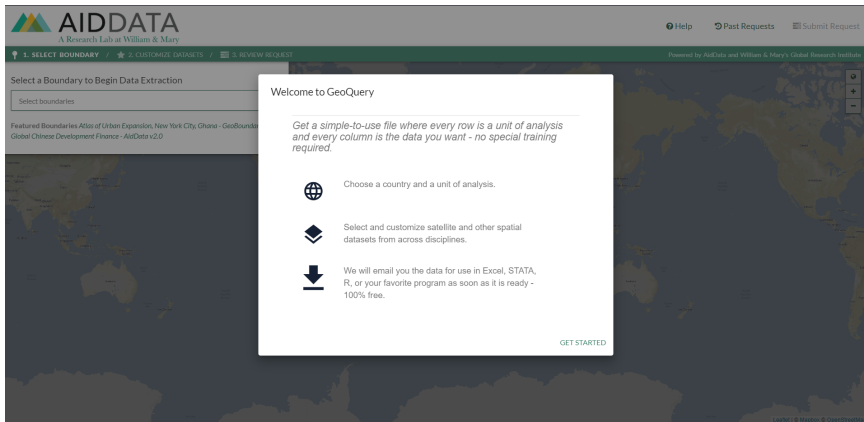
Zoom

Map

Rely on javascript but a python API exists. Lot of data freely accessible. Quite fast.



# AidData: geoquery



Good solution to start, lot of data available. But spatial limitation.

# Overview

Introduction to remote sensing

Sensor Overview

Accessing data

Pitfalls

# Main limitations

- Clouds:
  - Yield to measurement error
  - Hard to filter
  - Issue for cloudiest locations
- Time limitation:
  - Tradeoff between granularity and revisit rate
  - Higher-resolution images for recent satellites
  - Can have access to granular images but .. quite expensive 15k for Planet
- Non classical measurement error

# References I

- Marshall Burke, Anne Driscoll, David B Lobell, and Stefano Ermon. Using satellite imagery to understand and promote sustainable development. *Science*, 371(6535):eabe8628, 2021.
- J Vernon Henderson, Adam Storeygard, and David N Weil. Measuring economic growth from outer space. *American economic review*, 102(2):994–1028, 2012.
- Seema Jayachandran. Air quality and early-life mortality: Evidence from indonesia’s wildfires. *Journal of Human resources*, 44(4):916–954, 2009.
- Benjamin Marx, Thomas M Stoker, and Tavneet Suri. There is no free house: Ethnic patronage in a kenyan slum. *American Economic Journal: Applied Economics*, 11(4):36–70, 2019.