Lecture 6
Satellite Remote Sensing

Pierre Biscaye
Université Clermont Auvergne

Data Science for Economics

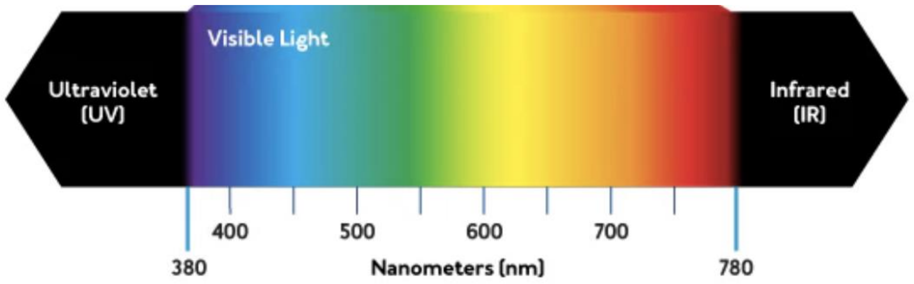
Note: Materials for this lecture are drawn from Sol Hsiang's Spatial Analysis course at UC Berkeley

Recap from Section 4

- Many types of satellite sensors:
  - Passive: observing light or other naturally occurring signals
  - Active: signals originating from sensor (e.g., radar, lidar)
- Donaldson & Storeygard (2016) discuss use in economics
- Primary advantages:
  - Access to information difficult to obtain by other means
  - Unusually high spatial resolution
  - Wide geographic coverage
  - Increasingly greater temporal frequency
- Primary disadvantages
  - Dataset size
  - Spatial dependence
  - Measurement error
  - Privacy concerns

Source

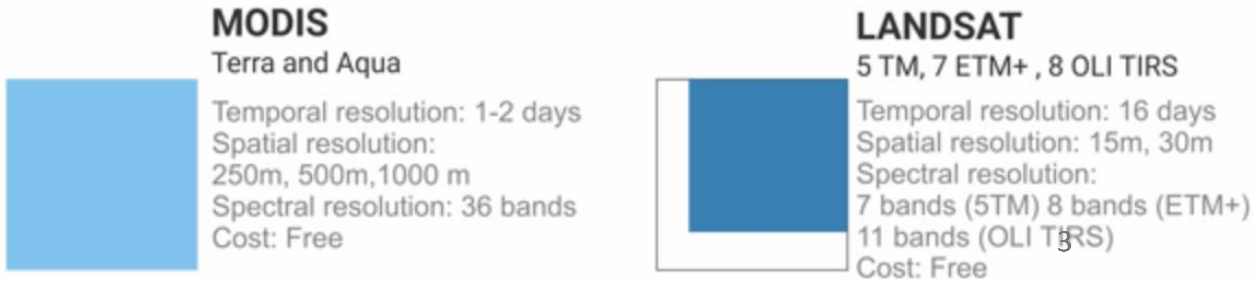
Satellite imagery data



Source

- **Light wavelength** (λ) associated with color
  - Tradeoffs between spectral and spatial resolution
- **Panchromatic** data: single wide spectral band (grayscale)
  - Can be very high resolution, useful for detecting fine details
- **RGB** data: captures 3 visible light bands
  - Combine to create natural-color images, useful for visual interpretation
- **Multispectral** data: captures images in multiple distinct spectral bands (e.g., visible and infrared)
  - Enable analysis of features like vegetation (NDVI) or water content (NDWI)
- **Hyperspectral** data: hundreds of contiguous spectral bands from visible to infrared
  - Lowest resolution but high detail useful for things like analyzing soil composition or crop types

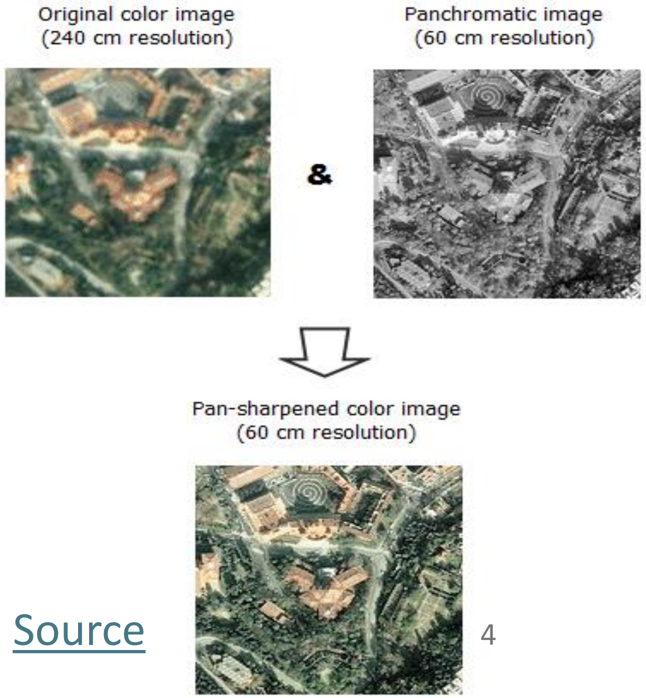
Source



Source

Common tools/tricks when using satellite imagery

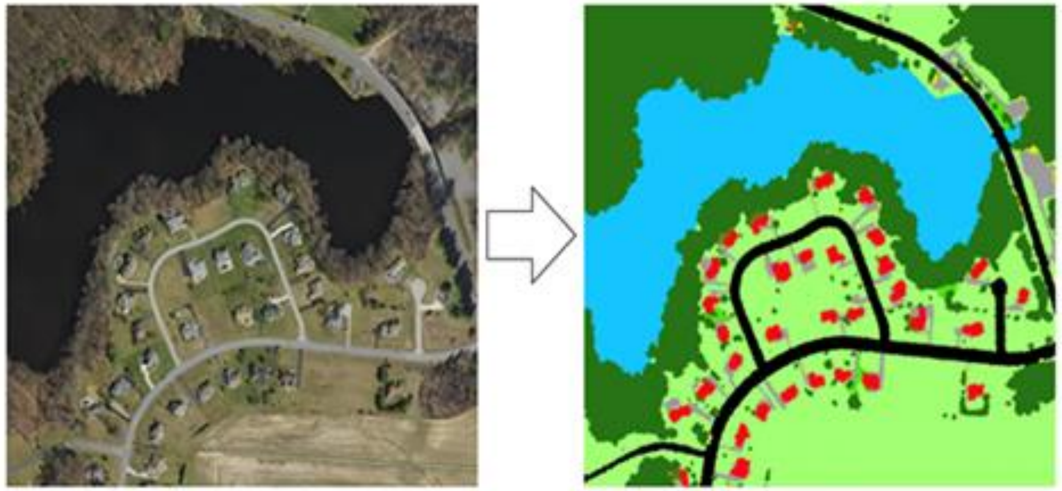
- **Change detection**: difference out features from adjacent time points
- **Thresholding**: masking out pixels above/below given threshold value
- **Cloud removal**: extract time series at given point, sort by pixel "whiteness", extract least white pixels to assemble cloud-free composite
- **Sharpening**: increase appearance of contrast between objects in an image
    - Can use a high-res panchromatic image to augment a lower-res multispectral image
    - Unsharp masking: use spatial patterns in data to improve "crispness of edges



Source

Pixel classification

- A key application with satellite imagery: identify pixel contents based on optical reflectance across different bands
- **Visual inspection**
- **Calculate indices** based on differences in color bands
- **Use machine learning** based on training images, cross-validate



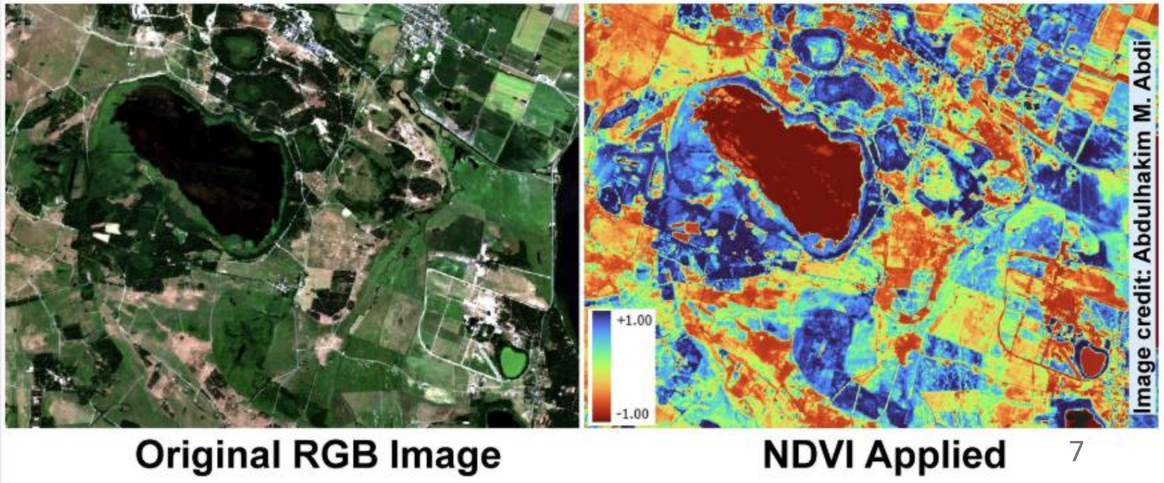
Source

Imagery-based vegetation indices

- Normalized Difference Vegetation Index (NDVI)
  - Measure of vegetation health and density → monitoring crop health, deforestation
- Enhanced Vegetation Index (EVI)
  - Measure of vegetation in areas with high biomass → vegetation monitoring in tropical forests and ag land
- Normalized Difference Moisture Index (NDMI)
  - Measures vegetation moisture content → drought assessment, wildfire risk prediction
- (Modified) Soil-Adjusted Vegetation Index ((M)SAVI)
  - Accounts for soil brightness in vegetation analysis
- Green Chlorophyll Index (GCI)
  - Measures chlorophyll concentration in vegetation → precision ag, crop monitoring
- Leaf Area Index (LAI)
  - Estimates leaf area per unit ground surface area → forest/ecosystem productivity

Normalized Difference Vegetation Index (NDVI)

- NDVI = (NIR – R)/(NIR + R)
  - Corresponds to particular bands on different satellites
  - Vegetation absorbs red light in photosynthesis: less R → more vegetation
  - Healthy vegetation strongly reflects near-infrared: more NIR → more vegetation
- Measure of plant "health"/vegetation intensity based on amount of photosynthesis happening



Source

NDVI in development economics

- Burke & Lobell (2017 PNAS), "Satellite-based assessment of yield variation and its determinants in smallholder African systems"
- Focus on maize plots
- Combined field data on yield for ground truth with satellite imagery and calculation of NDVI, GCVI, EVI
- Used random forests for land cover classification to identify maize plots
- Still active area of research
  - Ferguson et al (in progress), "Downscaling aggregate maize yields using satellite imagery and machine learning"

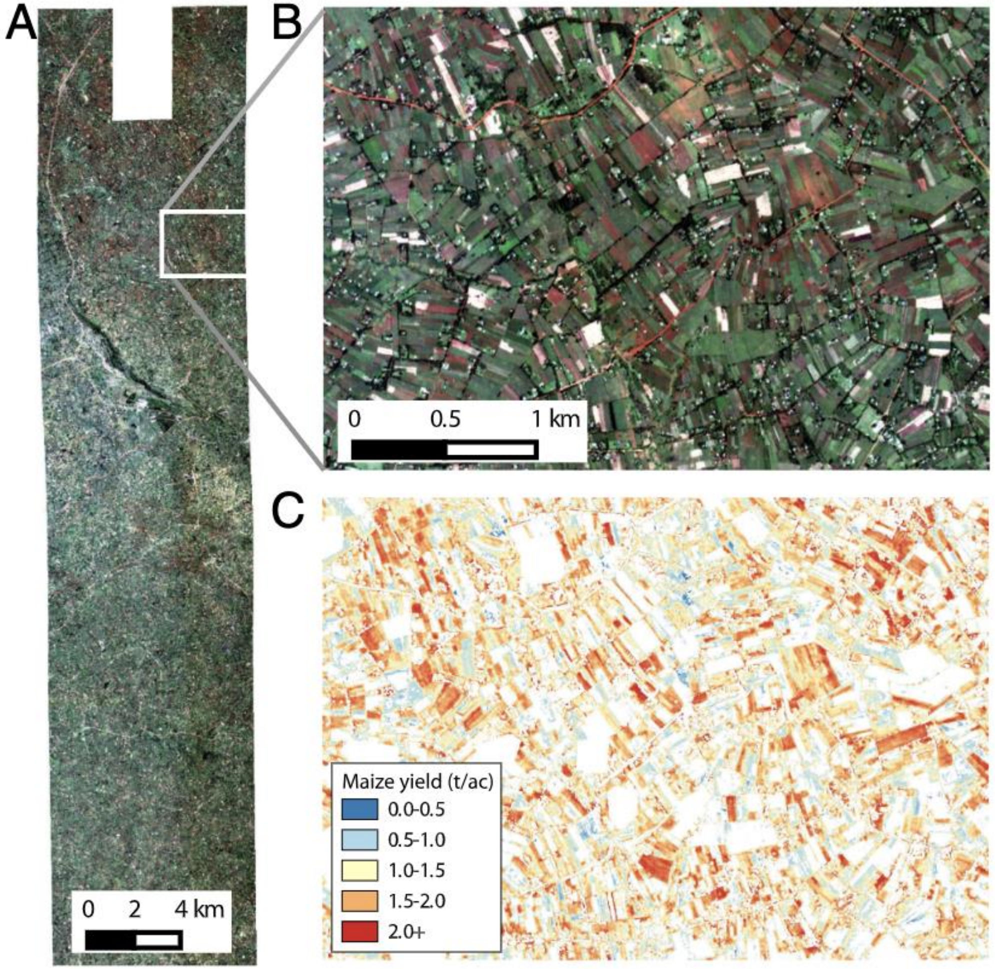


Fig. 5. Maize yield map for the study region, 2015. (*A* and *B*) One-meter image from Terra Bella of the study region (*A*) and zoom-in of that image (*B*) (see Fig. S3 for a higher-resolution version). (*C*) Yield map of the zoomed-in region for pixels classified as maize.

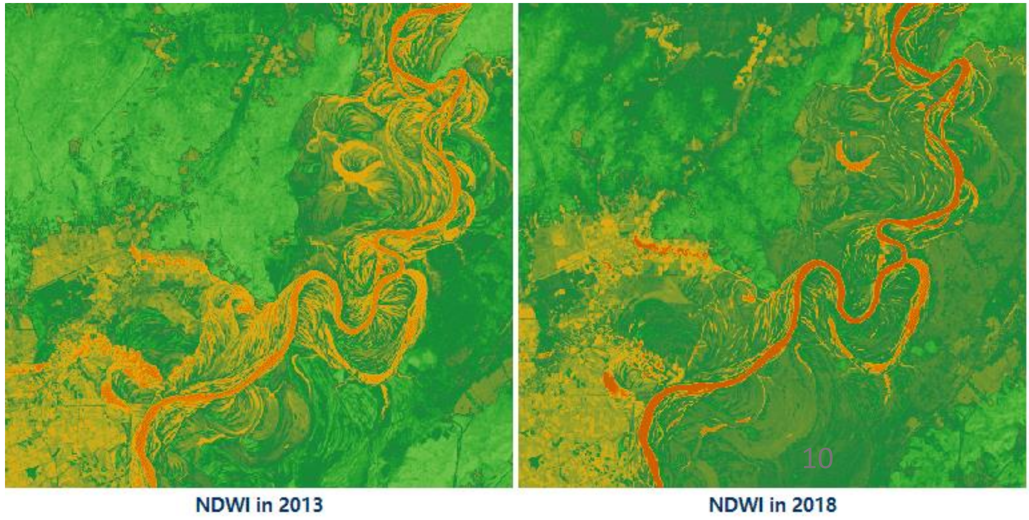
Other common imagery-based indices

- Nighttime Lights Index
  - Measures intensity of artificial light → proxy for economic activity, urbanization, electricity access
- Normalized Difference Water Index (NDWI)
  - Detects water bodies and soil moisture → flood mapping, water resource management
- Burned Area Index (BAI)
  - Detects burned areas from wildfires → fire impact and recovery
- Normalized Difference Built-Up Index (NDBI)
  - Identifies urban and built-up area → urban growth analysis, infrastructure planning
- Aerosol Optical Depth (AOD)
  - Measures concentration of aerosols in the atmosphere based on scattering and absorption of light → monitoring air quality and pollution

Normalized Difference Water Index (NDWI)

- NDWI = (G – NIR)/(G + NIR)
  - Water strongly reflects green and reflects very little NIR
  - NIR much higher for vegetation and bare soil, less green reflection: smaller and sometimes negative NDWI values
  - Positive difference in numerator suggests water, and smaller denominator amplifies this
  - Alternative for vegetation water content: (NIR – SWIR)/(NIR + SWIR)
    - Water reflects even less SWIR than NIR
- Measure of water content → indicator of water or moist surfaces

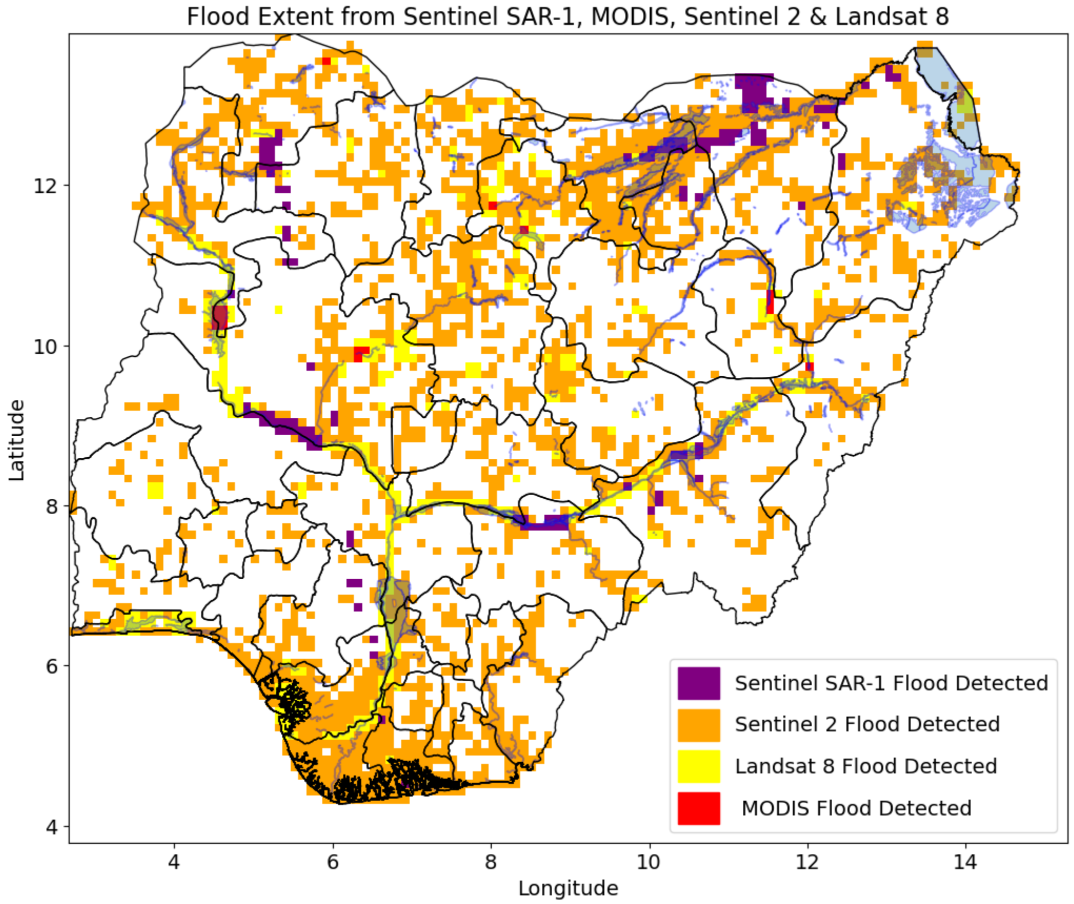
Source



NDWI in development economics

- Guiteras et al (2015 AER P&P), "Satellites, self-reports, and submersion: Exposure to floods in Bangladesh"
  - Compare survey-reported flooding to a rainfall-based proxy and a satellite measure based on MODIS imagery and NDWI
  - Find weak correlation between survey and satellite reports, weak/negative correlations between rainfall and satellite
- Ongoing work on satellite-based flood detection
  - Patel (2024 WP), "Floods" → satellite imagery + radar + surveys + ML
  - Bangalore et al (WIP), "Mapping flood exposure and human impacts" → comparisons across measures, role of definitions and measurement

Mapping flood exposure and human impacts

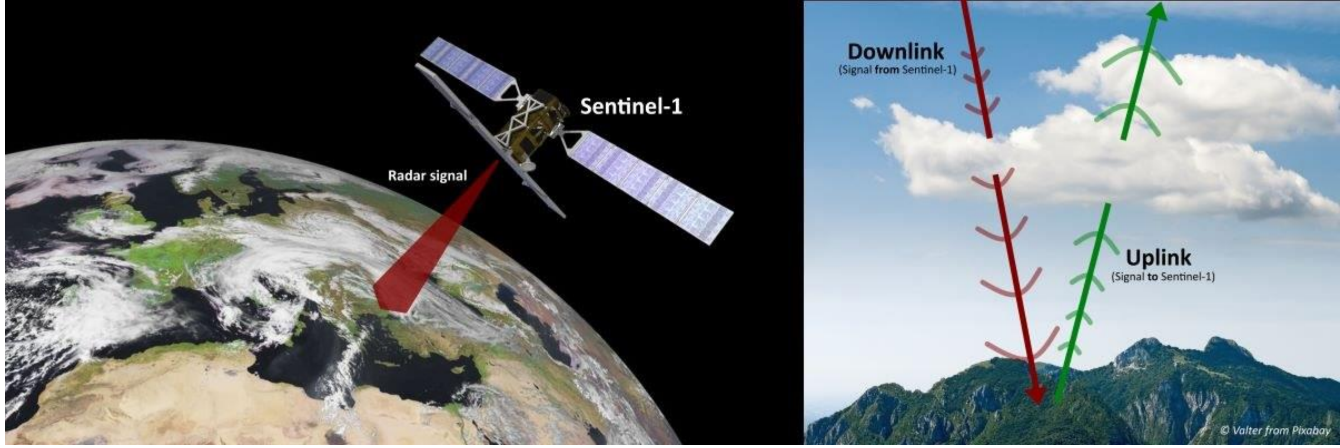


Bangalore et al (WIP)

Remote sensing and machine learning

- ML: predict Y based on some features X, when Y observed for only a small subsample
  - More on this next section
- Three types of ML applications in remote sensing
  - Classification: does an observation fall within a set of classes (e.g., crops)
  - Regression: predict a scalar value associated with an image (e.g., wealth)
  - Segmentation: where within image is something (e.g., identify roads, houses)
- Key: turning information in images into features **X** for analysis
  - Ex: generate features based on distribution of color in the image
  - Convolutional filters: generate features based on patterns in spatial arrangement of pixels; focus on "sub-images"

Satellite radar data



Source

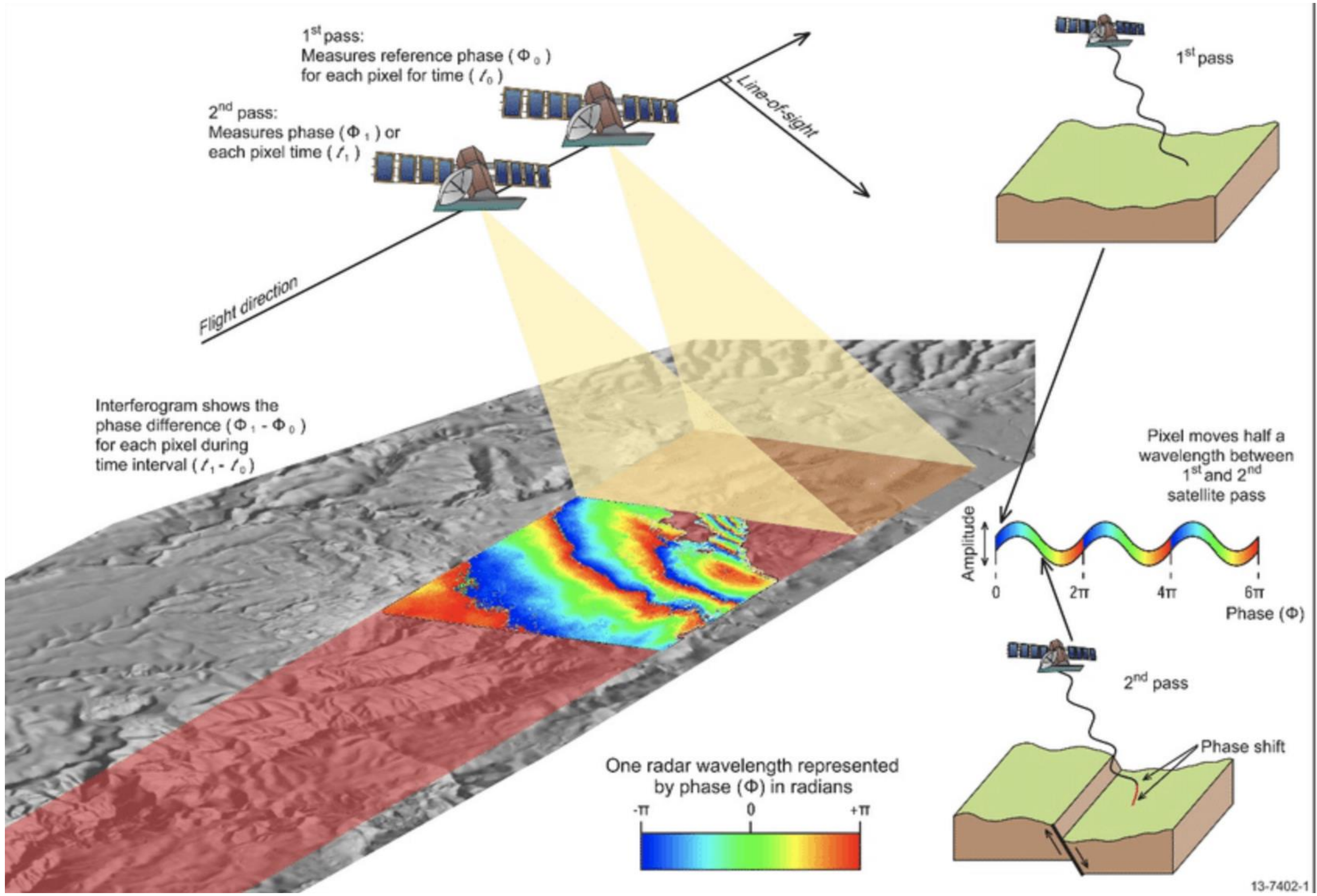
- Synthetic Aperture Radar (SAR) uses microwave signals to capture high-resolution images of Earth's surface
- Unlike optical imagery, radar can penetrate clouds, smoke, and darkness → useful in all weather conditions and at night
- Measures surface roughness, elevation, and changes in surface features
- Key outputs:
  - Backscatter intensity (surface reflectivity)
  - Interferometry (surface displacement or deformation)

Processing radar data

- Data **preprocessing**:
  - Raw radar data is processed to remove noise (speckle filtering) and georeferenced to align with maps
  - Calibration ensures backscatter values are consistent across time and locations
  - Terrain correction compensates for topographic effects on radar signals
- Data **transformation**:
  - Converts backscatter intensity into metrics such as soil moisture, water extent, or vegetation structure
  - Interferometric SAR (InSAR) techniques detect small changes in elevation, such as land subsidence or flooding

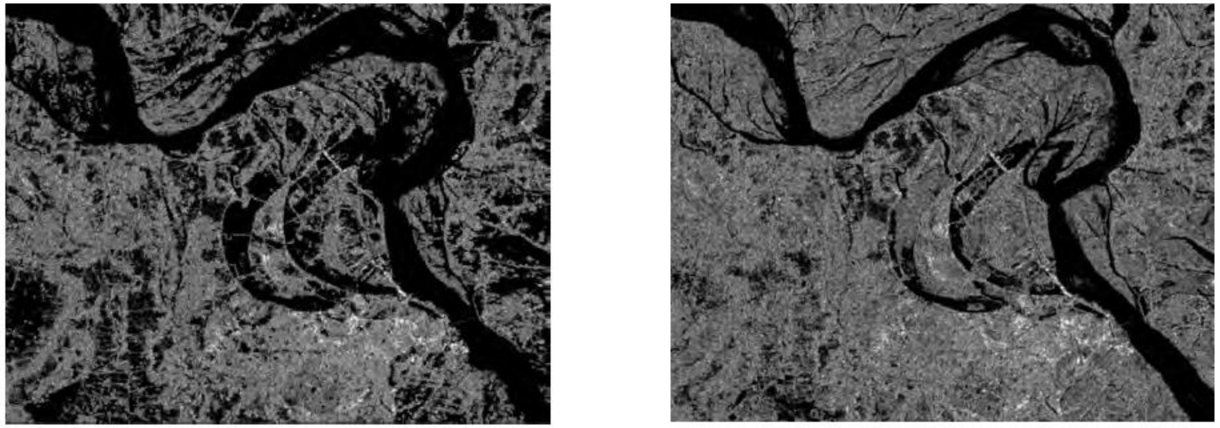
Applications of SAR in economics research

- Flood mapping: detects water extent during floods
- Agriculture and land use: tracks soil moisture and crop conditions
- Urban development and infrastructure: detects land subsidence or infrastructure changes
- Deforestation and environmental change: tracks forest loss and land degradation



Kuenzer et al 2019

SAR and flood detection (Patel 2024)


(a) Raw Data—July 27, 2020 (Flood) (b) Raw Data—July 22, 2021 (No Flood)

- SAR water detection
  - Land, buildings, and vegetation reflect radar back, surface water scatters radar
  - Big advantage: seeing through clouds and at night
  - But less frequent temporal coverage
- Patel process for identifying floods
  1. Calculate amount of surface water in each pixel for every day SAR satellite passes over
  2. Use ML to predict radar surface water using more frequent satellite imagery-based surface water measures
  3. Remove permanent water by residualizing at calendar week level, choose threshold for flood based on survey flood reports
  4. Validate measure using news reports, government flood reports, river gauges
- Then analyzes impacts of flood exposure in Bangladesh