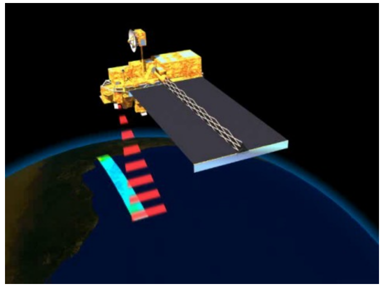
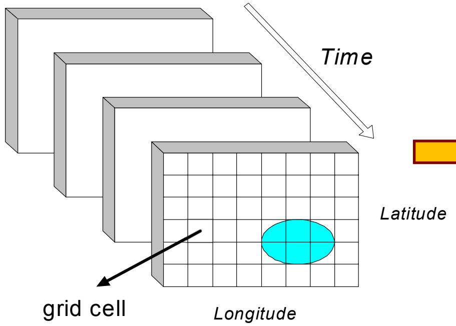
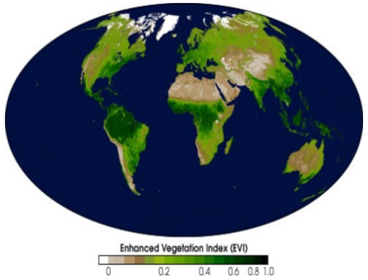
Data Type 3: temporal snapshot model

Xiaowei Jia
University of Pittsburgh
Xiaowei@pitt.edu

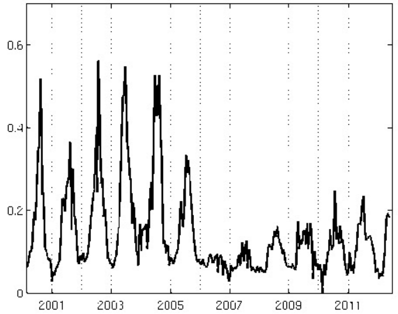
Big Data in Remote Sensing – temporal snapshot model







Different indices measures the surface biomass, temperature, soil conditions, air quality, etc.



This **time series** of indices captures temporal dynamics

MODIS covers ~ 5 billion locations globally at 250m resolution daily since Feb 2000.

Data	Type	Coverage	Spatial Resolution	Temporal Resolution	Spectral Resolution	Duration	Availability
MODIS	Multispectral	Global	250 m	Daily	7	2000 - present	Public
LANDSAT	Multispectral	Global	30 m	16 days	7	1972 - present	Public
Hyperion	Hyperspectral	Regional	30 m	16 days	220	2001 - present	Private
Sentinal - 1	Radar	Global	5 m	12 days	-	2014 - present	Public
Quickbird	Multispectral	Global	2.16 m	2 to 12 days	4	2001 - 2014	Private
WorldView - 1	Panchromatic	Global	50 cm	6 days	1	2007 - present	Private

Remote Sensing for Health

- How is climate change related to important health problems
- New opportunities to use machine learning and data mining methods to extract important characteristics of the Earth system that are indicative of disease infection
- Use remote sensing to track changes before and after the pandemic

Outline

- Monitoring the spatio-temporal pattern of disease infections using remote sensing data
  - **Potential of using remote sensing data in monitoring large-scale changes**
  - RS for studying effects of climate change on epidemics
  - RS for studying impacts of the pandemic
  - Future directions

Example: Global Forest Fires Mapping

Monitoring fires is important for climate change impact

A record number of more than 150 countries signed the landmark agreement to tackle climate change at a ceremony at UN headquarters on 22 April, 2016.

SEARCH The New York Times

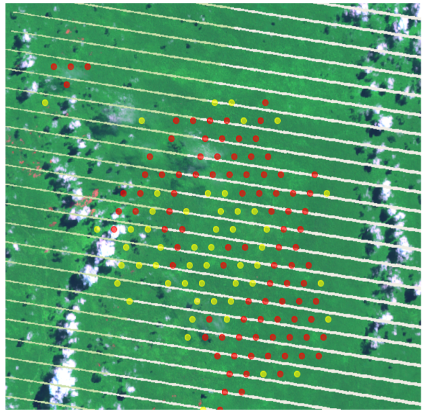
ENVIRONMENT

Delegates at Climate Talks Focus on Saving the World's Forests

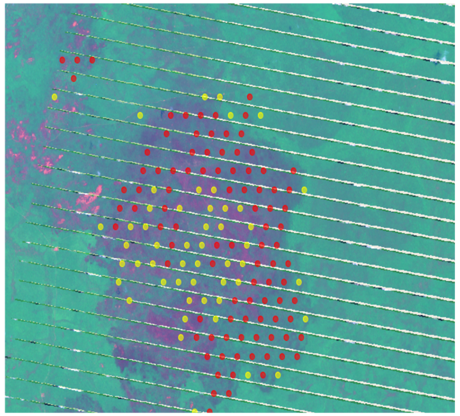
By JUSTIN GILLIS DEC. 30, 2015

The canopy of the forest in Puerto Viejo, Costa Rica, in October 2014. Climate change negotiations in Paris could lead to a sweeping effort to save the world's forests. Adriana Zehbrauskas for The New York Times

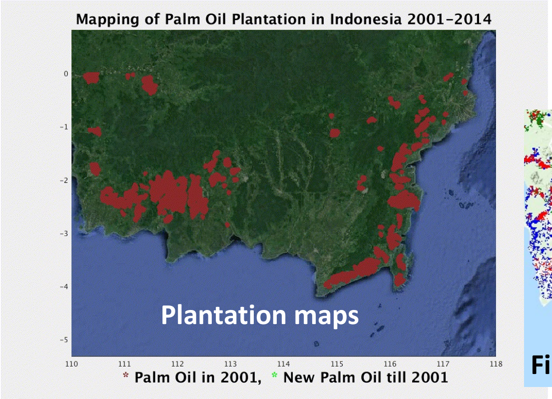
Mithal et al. 2017



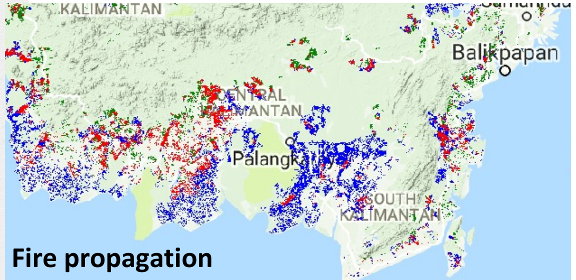
Before Fire Event



After Fire Event



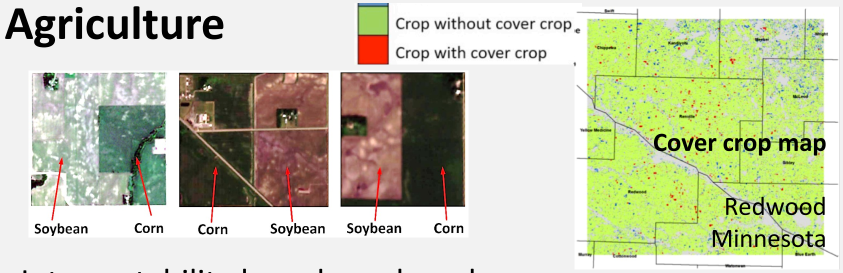
Jia et al. 2020



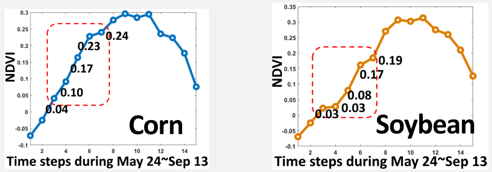
Mithal et al. Rapt: Rare class prediction in absence of true labels. TKDE, 2017.
Jia et al. Automated plantation mapping in southeast asia using modis data and imperfect visual annotations. Remote Sensing. 2020

Example: Land Use and Land Cover Study

Agriculture

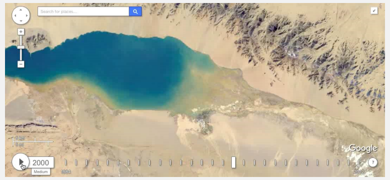


Interpretability based on phenology:



Urbanization

Water monitoring

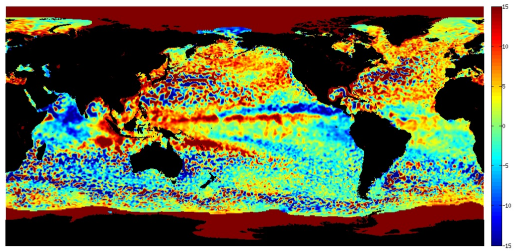


Water dynamics

Water quality

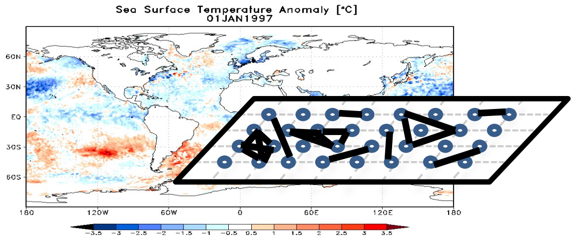
Understanding Climate Change: A Data-driven Approach

Research Highlights



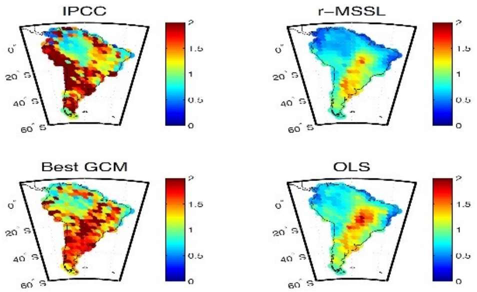
Pattern Mining: Monitoring Ocean Eddies

- Spatio-temporal pattern mining using novel multiple object tracking algorithms
- Created an open source data base of 20+ years of eddies and eddy tracks



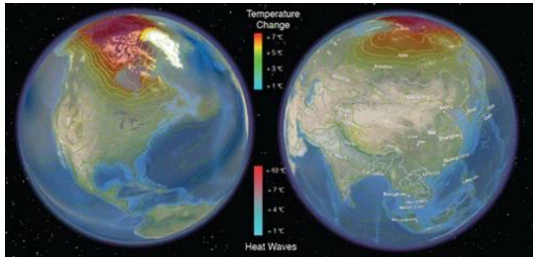
Network Analysis: Climate Teleconnections

- Scalable method for discovering related graph regions
- Discovery of novel climate teleconnections
- Also applicable in analyzing brain fMRI data



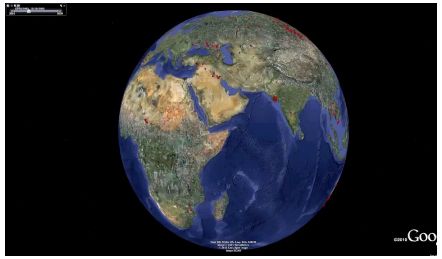
Sparse Predictive Modeling: Precipitation Downscaling

- Hierarchical sparse regression and multi-task learning with spatial smoothing
- Regional climate predictions from global observations



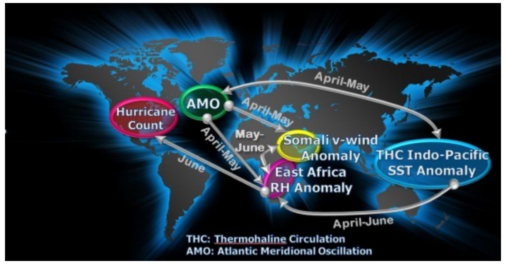
Extremes and Uncertainty: Heat waves, heavy rainfall

- Extreme value theory in space-time and dependence of extremes on covariates
- Spatiotemporal trends in extremes and physics-guided uncertainty quantification



Change Detection: Monitoring Ecosystem Distrubances

- Robust scoring techniques for identifying diverse changes in spatio-temporal data
- Created a comprehensive catalogue of global changes in surface water and vegetation, e.g. fires and deforestation.



Relationship mining: Seasonal hurricane activity

- Statistical method for automatic inference of modulating networks
- Discovery of key factors and mechanisms modulating hurricane variability

http://climatechange.cs.umn.edu/

Outline

- Monitoring the spatio-temporal pattern of disease infections using remote sensing data
  - Potential of using remote sensing data in monitoring large-scale changes
  - **RS for studying effects of climate change on epidemics**
  - RS for studying impacts of the pandemic
  - Future directions

Effects of climate change on epidemics

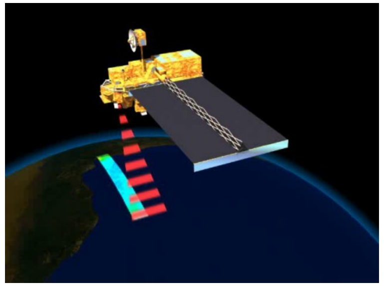
- Cryptosporidiosis (Milwaukee, Wisconsin, USA, in 1993) and Escherichia coli O157 infection (Walkerton, Ontario, Canada, in 2000): Both events were preceded by heavy rains; had highly concentrated sources of pathogens in the form of untreated sewage and animal waste, respectively. (Wilson et al.)

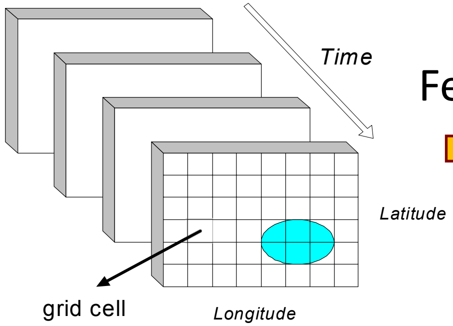
Wilson ME et al. Disease in evolution: global changes and emergence of infectious diseases. New York: New York Academy of Sciences; 1994

Effects of climate change on epidemics

- Cryptosporidiosis (Milwaukee, Wisconsin, USA, in 1993) and Escherichia coli O157 infection (Walkerton, Ontario, Canada, in 2000): Both events were preceded by heavy rains; had highly concentrated sources of pathogens in the form of untreated sewage and animal waste, respectively. (Wilson et al.)

- Earth observation data can be leveraged to estimate environmental variables that influence the transmission cycle of diseases.





Spatio-temporal data

Feature extraction

Environmental variables and representative indices
Land covers, sea surface temperature, vegetation indices, precipitation

Predict

Outbreaks/infection of target diseases

Wilson ME et al. Disease in evolution: global changes and emergence of infectious diseases. New York: New York Academy of Sciences; 1994

Feature extraction from RS data

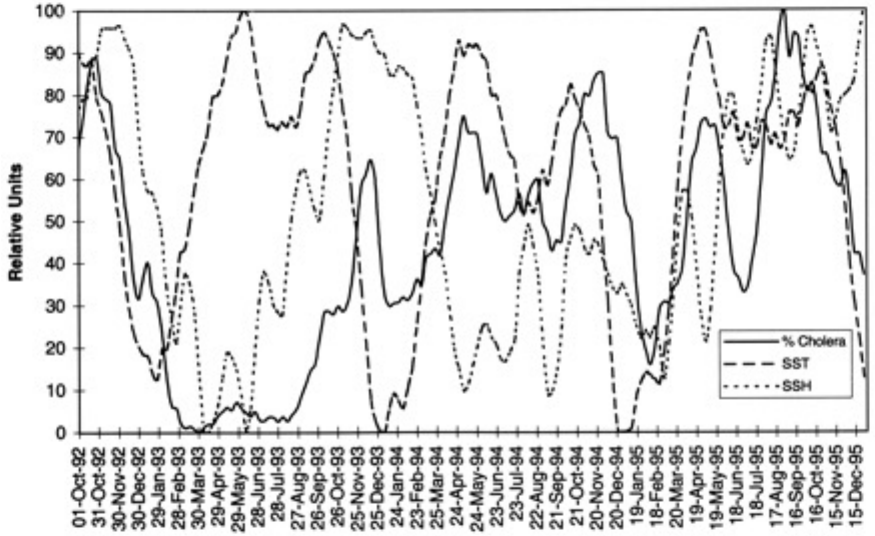
- Objective: extract representative features (vegetation, land surface temperature, atmospheric moisture and rainfall indices) from satellite imagery
- Domain knowledge-based extraction from multiple channels (e.g., NDVI, Land Surface Temperature Indices, Moisture Indices)
  - For example, live green plants appear relatively bright in the near-infrared while clouds and snow tend to be rather bright in the red (as well as other visible wavelengths) and quite dark in the near-infrared

$$\mathrm{NDVI} = \frac{NIR - Red}{NIR + Red}$$

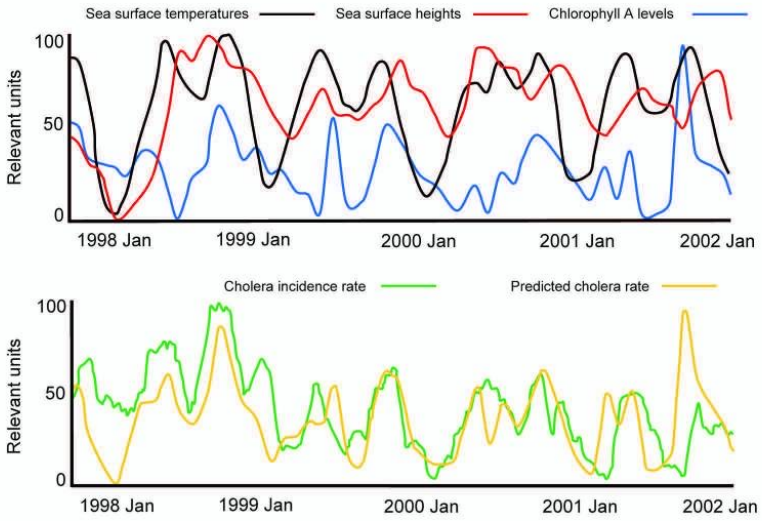
- Traditional ML approaches, spatio-temporal data mining approaches.

Use RS to predict cholera outbreaks

- Motivation: Flooding is the most frequent natural weather disaster (30%–46% of natural disasters in 2004–2005), affecting >70 million persons worldwide each year.
- Lobitz et al. extract sea surface temperature and sea surface height from satellite data, and then use them as input to statistical model to predict Cholera cases in Bangladesh.
- Ford et al. use sea surface temperature, sea surface height, and chlorophyll A levels to cholera outbreaks in South America.
- Can be used for other diseases associated with floods such as diarrhea, typhoid, hepatitis (jaundice), and leptospirosis.



Lobitz et al.



Ford et al.

Lobitz et al. Climate and infectious disease: Use of remote sensing for detection of *Vibrio cholerae* by indirect measurement. PNAS. 2000
Ford et al. Using Satellite Images of Environmental Changes to Predict Infectious Disease Outbreaks. Emerg Infect Dis. 2009

RS research for vector-borne disease

- Objective: Study the spread of mosquito-borne diseases, including Malaria, Dengue, West Nile Virus
- Input:
  - 1) environmental/climate variables (e.g., air temperature, soil temperature, SST, precip, NDVI, EVI),
  - 2)non-environmental variables: Population density (estimated from the intensity of nighttime light), running water, hygienic services, etc.
- Output labels: epidemiological data (disease incidence, prevalence or case, mortality data)
- Modeling approaches:
  - Simple regression models
  - Statistical models: ARIMA, spatial statistics
  - Probabilistic graphical models
  - ML methods (SVM, ANN, and ensemble approaches)

Parselia et al. Satellite Earth Observation Data in Epidemiological Modeling of Malaria, Dengue and West Nile Virus: A Scoping Review. Remote Sensing. 2019

Other references

Modeling other climate-related variables

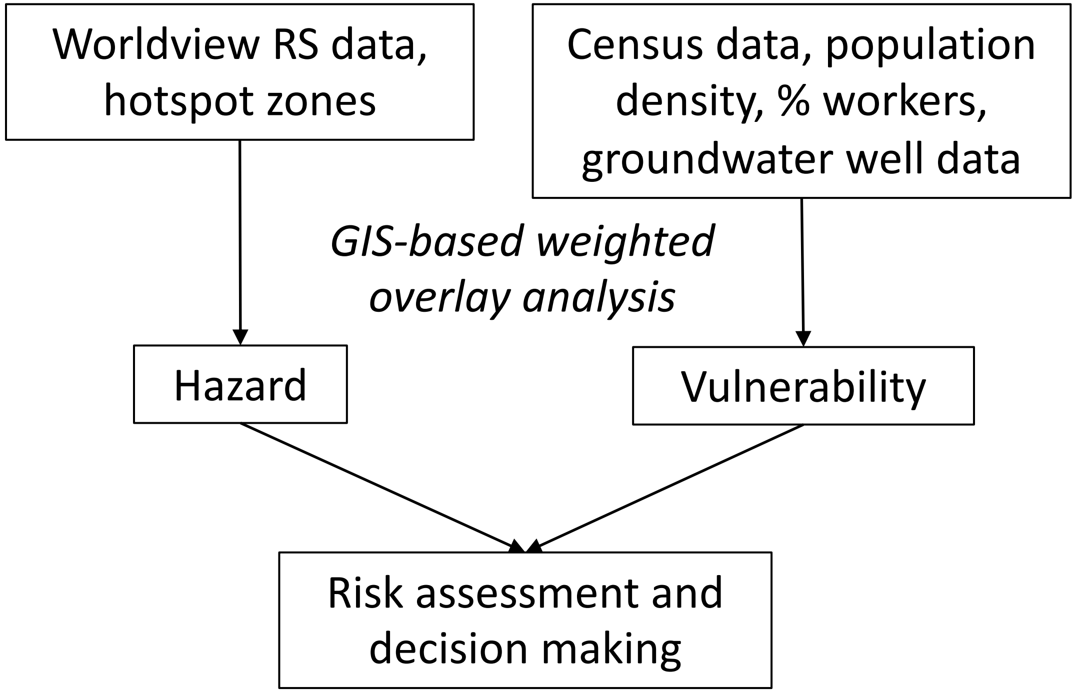
- Lin et al. "Mining public datasets for modeling intra-city PM2. 5 concentrations at a fine spatial resolution." In *Proceedings of the 25th ACM SIGSPATIAL international conference on advances in geographic information systems*, 2017.
- Kotchi et al. "Using Earth observation images to inform risk assessment and mapping of climate change related infectious diseases." *Canada Communicable Disease Report*, 2019.

Study on other diseases

- (Vector-borne disease) Ceccato et al. "Data and tools to integrate climate and environmental information into public health." Infectious diseases of poverty, 2018.
- (Brucellosis, ANN) Wang et al. A Remote Sensing Data Based Artificial Neural Network Approach for Predicting Climate-Sensitive Infectious Disease Outbreaks: A Case Study of Human Brucellosis, Remote Sensing, 2017

Remote sensing for COVID-19

- How to use remote sensing in analyzing COVID infection?



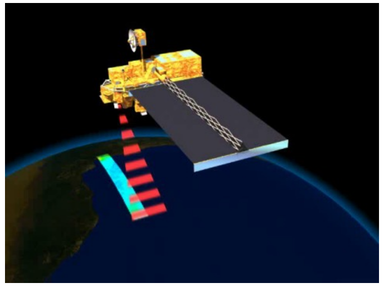
Kanga et al. Analyzing the Risk to COVID-19 Infection using Remote Sensing and GIS. Risk Analysis. 2021

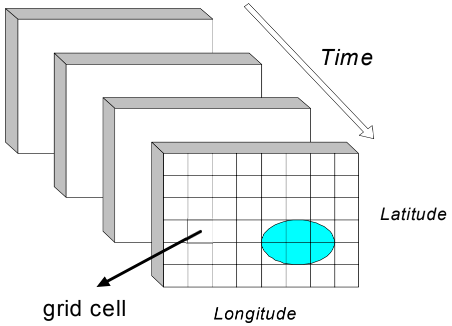
Outline

- Monitoring the spatio-temporal pattern of the disease infections using remote sensing data
  - Potential of using remote sensing data in monitoring large-scale changes
  - RS for studying effects of climate change on epidemics
  - **RS for studying impacts of the pandemic**
  - Future directions

RS for studying impacts of the pandemic

Spatio-temporal data





Day t_1

Variables related to human activities, economy, etc.

Day t_2

Variables related to human activities, economy, etc.

. . .

Day t_T

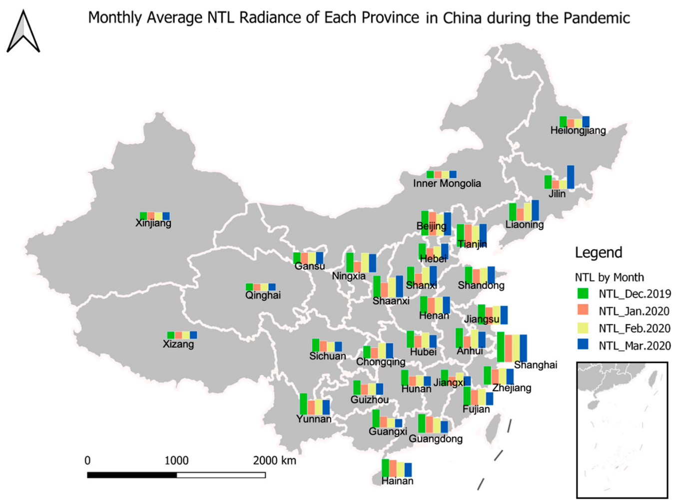
Variables related to human activities, economy, etc.

Before pandemic

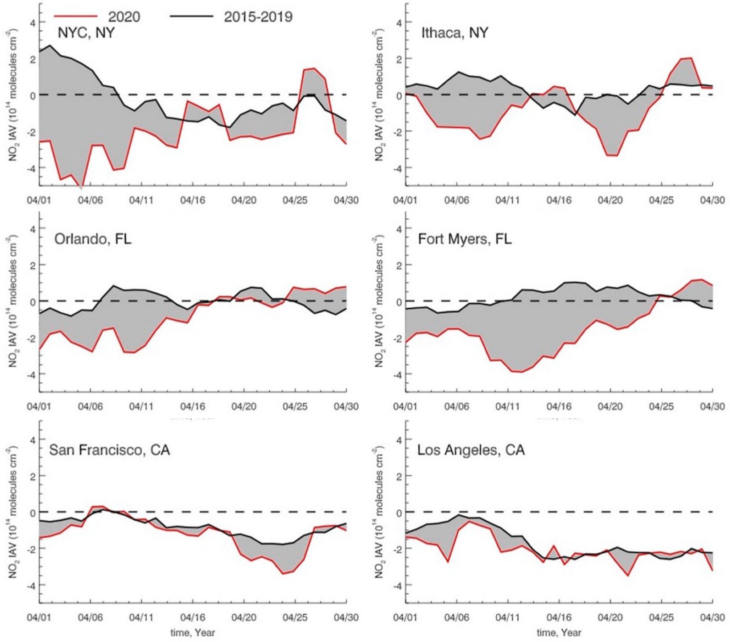
After pandemic

RS for studying impacts of the pandemic

- Human activities: Nighttime Light radiance before and during the pandemic in mainland China (Liu et al.)
- Air quality: reduction of CO and NO_2 in traffic-intensive states (NY, IL, FL, TX, and CA) during the pandemic (Elshobany et al.).



Liu et al.



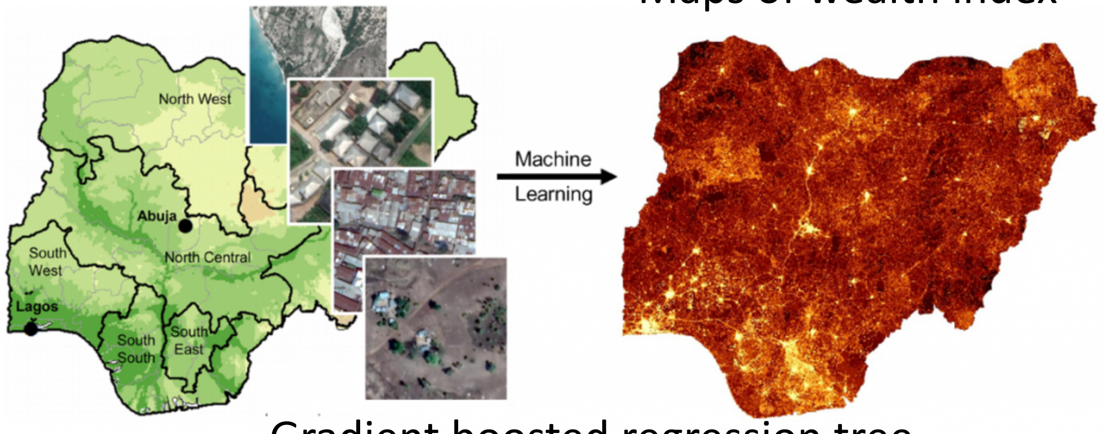
Elshobany et al.

Liu et al. Spatiotemporal Patterns of COVID-19 Impact on Human Activities and Environment in Mainland China Using Nighttime Light and Air Quality Data. Remote Sensing. 2020
Elshobany et al. The Status of Air Quality in the United States During the COVID-19 Pandemic: A Remote Sensing Perspective. Remote Sensing. 2021

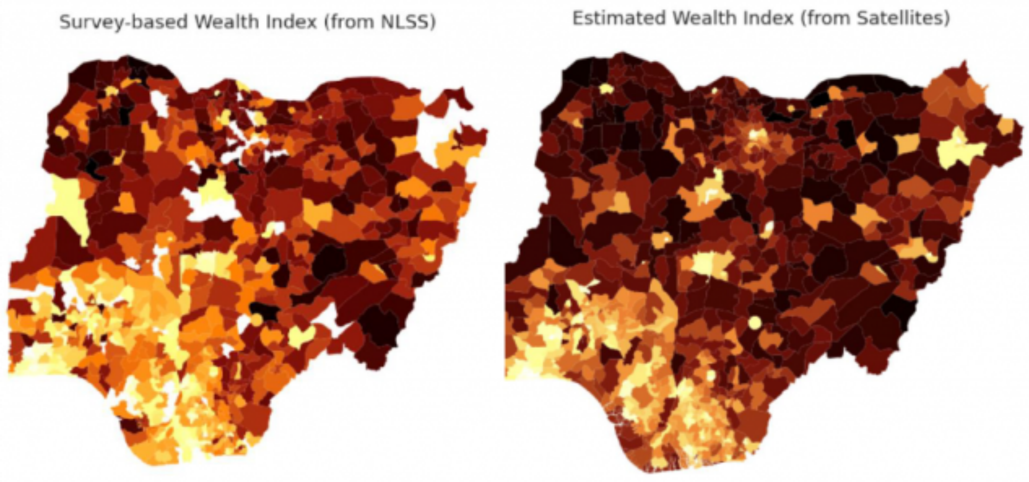
Poverty and Economy

- Objective: generate up-to-date poverty maps for Nigeria using satellite imagery.

Maps of wealth index



Gradient boosted regression tree



- Other works: Nighttime light for studying declines and recovery in economy (Elvidge et al., 2020)

Blumenstock et al. Using Big Data and machine learning to locate the poor in Nigeria.
Elvidge et al. The Dimming of Lights in China during the COVID-19 Pandemic. Remote Sensing. 2020

Other references

- Sussman et al. Can We Measure a COVID-19-Related Slowdown in Atmospheric CO_2 Growth? Sensitivity of Total Carbon Column Observations. Remote Sensing. 2020
- Liu et al. Spatiotemporal Patterns of COVID-19 Impact on Human Activities and Environment in Mainland China Using Nighttime Light and Air Quality Data. Remote Sensing. 2020
- Li et al. Estimating the Impact of COVID-19 on the $PM_{2.5}$ Levels in China with a Satellite-Driven Machine Learning Model. Remote Sensing. 2021

Outline

- Monitoring the spatio-temporal pattern of disease infections spread using remote sensing data
    - Potential of using remote sensing data in monitoring large-scale changes
    - RS for studying effects of climate change on epidemics
    - RS for studying impacts of the pandemic
    - **Future directions**

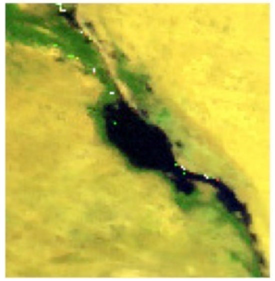
Future opportunities

- Advanced machine learning algorithms
- Remote sensing datasets
- Knowledge-guided machine learning

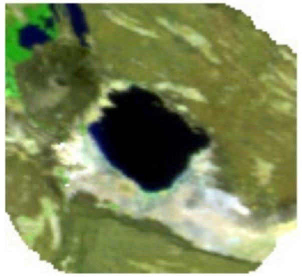
Advanced machine learning algorithms

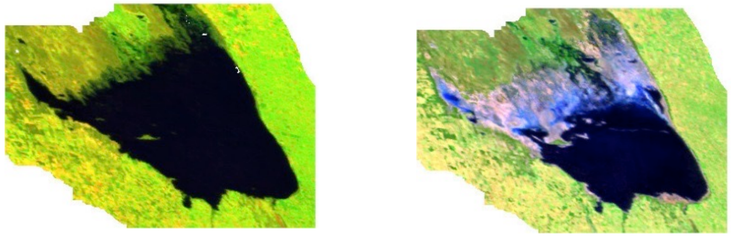
Major challenges in remote sensing

- Spatial and temporal heterogeneity
- Limited and noisy ground-truth labels
- Noisy data


Great Bitter Lake, Egypt

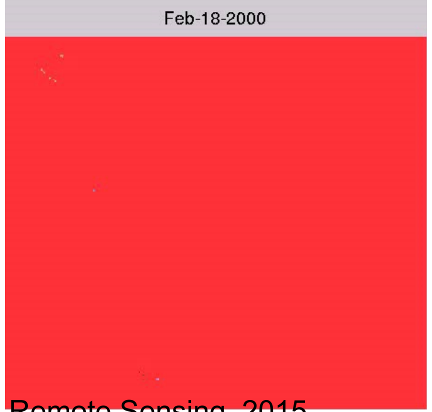

Lake Tana, Ethiopia


Lake Abbe, Africa


Mar Chiquita Lake, Argentina in 2000 (left) and 2012 (right)

Advanced ML on RS

- Transfer learning (Hu et al.)
- Zero-shot learning (Li et al.)
- Weakly-supervised learning (Schmitt et al.)
- Others (Ghosh et al.)



Hu et al. Transferring deep convolu-tional neural networks for the scene classification of high-resolution remote sensing imagery. Remote Sensing, 2015.
Li et al. Zero-shot scene classifi-cation for high spatial resolution remote sensing images. TGRS, 2017.
Schmitt et al. Weaklysupervised semantic segmentation of satellite imagesfor land cover mapping–challenges and opportunities. 2020.
Ghosh et al. Land Cover Mapping in Limited Labels Scenario: A Survey. 2021

Remote sensing datasets

Existing datasets

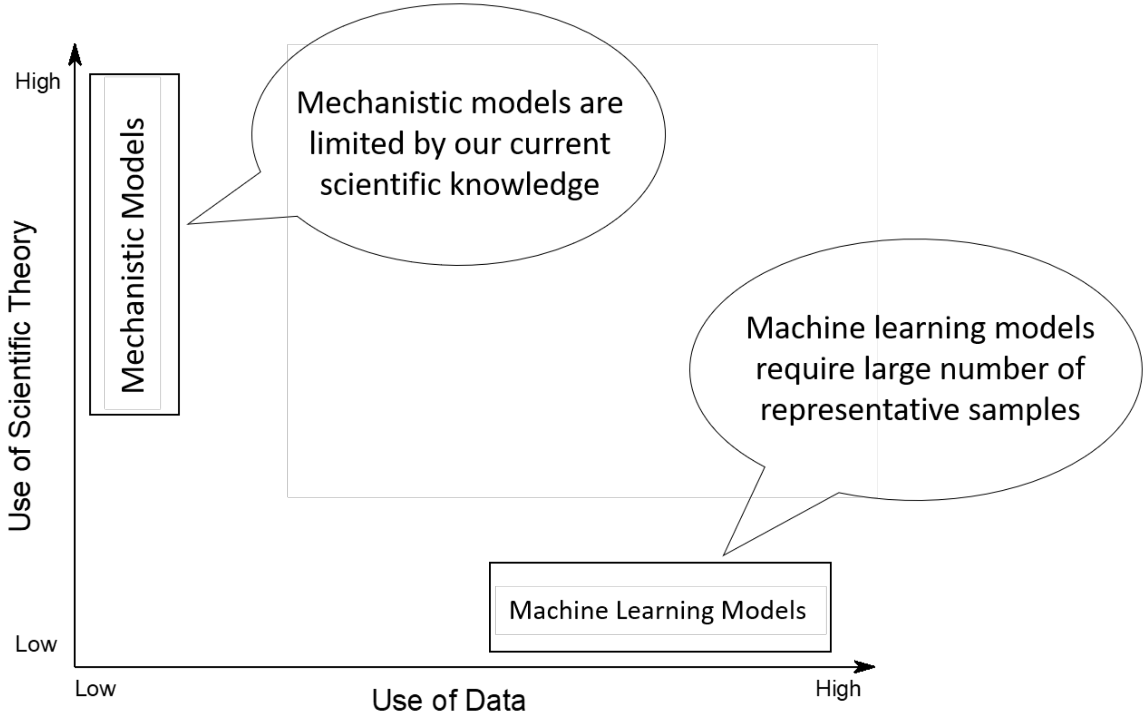
- Sen12MS
- DeepGlobe
- UC Merced Land Use Dataset
- WHU-RS Dataset

Other RS data sources

Data	Type	Coverage	Spatial Resolution	Temporal Resolution	Spectral Resolution	Duration	Availability
MODIS	Multispectral	Global	250 m	Daily	7	2000 - present	Public
LANDSAT	Multispectral	Global	30 m	16 days	7	1972 - present	Public
Hyperion	Hyperspectral	Regional	30 m	16 days	220	2001 - present	Private
Sentinal - 1	Radar	Global	5 m	12 days	-	2014 - present	Public
Sentinel-2	Multispectral	Global	10 m	5-10 days	12	2016 - present	Public
Quickbird	Multispectral	Global	2.16 m	2 to 12 days	4	2001 - 2014	Private
WorldView - 1	Panchromatic	Global	50 cm	6 days	1	2007 - present	Private

Knowledge-guided machine learning

- Mechanistic models have been widely used to study epidemiology, climate changes, traffic systems and economy.
- Machine learning models commonly require sufficient training data at desired spatial and temporal resolution.
- KGML for leveraging complementary strengths of two types of models. (Willard et al.)



Willard et al. Integrating Physics-Based Modeling with Machine Learning: A Survey. 2020

Q&A

Conclusion

- **Data type 1: event or process model** -- Human routine behavior modeling
  - Challenges of behavior modeling
  - Properties of spatiotemporal data
  - Inverse reinforcement learning
  - Behavior patterns and epidemic spreads
- **Data Type 2: temporal change model** -- Structural learning on networks
  - Susceptible-infected-recovered (SIR) – like models
  - Graph neural networks (GNN) for epidemiology
  - GNN, SIR, and PDE
- **Data Type 3: temporal snapshot model --** Remote sensing
  - Potential of using remote sensing (RS) data in monitoring large-scale changes
  - RS for studying effects of climate change on epidemics
  - RS for studying impacts of the pandemic

Thank you