

Sneak Peak

Nooriza Maharani

January 21, 2025

Table of contents

1 RS Learning Diary	4
1.1 Introduction	4
1.2 Why do I choose this module?	4
2 Getting to Know Remote Sensing	7
2.1 Summary	7
2.2 Application	8
2.3 Reflection	9
2.4 References	9
3 Xaringan and Quarto Book	10
3.1 Summary	10
3.2 Reflections	10
4 Image Correction	11
4.1 Summary	11
4.2 Application	12
4.3 Reflection	15
5 Policy	17
5.1 Summary	17
5.2 Application	20
5.3 Reflections	22
5.4 References	23
6 Introduction to GEE	24
6.1 Summary	24
6.2 Application	26
6.3 Reflection	27
6.4 References	27
7 Classification I	28
7.1 Summary	28
7.2 Application	29
7.3 Reflection	29
7.4 References	29

8 Classification II	30
8.1 Summary	30
8.2 Application	31
8.3 Reflection	31
8.4 References	31
9 Synthetic Aperture Radar	32
9.1 Summary	32
9.2 Application	33
9.3 Reflection	33
References	33

1 RS Learning Diary

This is a Quarto book to document my learning journey in **Remote Sensing Cities and Environments** course during my time at CASA UCL 24/25, offering insights learned, its applications, and my own reflections. The module is based on Dr Andrew MacLachlan github page [[here](#)].

*For those of you who also want to learn Geographic Information Science beyond ‘typical GIS’ Software, as in use R-Studio, you could also visit his other github page [[here](#)].

1.1 Introduction

Hi, I’m Nooriza, a student currently pursuing a Master’s degree in Urban Spatial Science at UCL. I graduated with a degree in Geography, specializing in Regional Development Studies. Over the past five years, I have worked as a technical consultant for various governments in Indonesia, ensuring that their decision-making is data-driven. Although the results often need to pass through political and budgeting reviews, that’s essentially what I do.

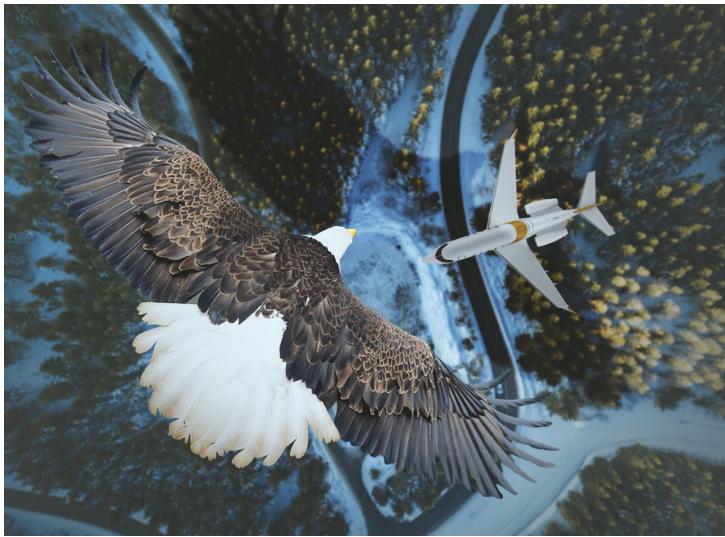
Honestly, I don’t have a specific topic of interest to add here because I have many! That’s why I chose an interdisciplinary course at CASA. One thing for sure is that I love doing analytics for social good. In the end, all the sophisticated methodologies and cutting-edge technical tools are meaningful when we use them to address challenges and solve problems, right?

1.2 Why do I choose this module?

The reason I choose remote sensing is because I want to use its vast open resources to analysis various topics. I had learned the foundations during my undergraduate degree but I haven’t delved further into it and haven’t got any experience to use GEE yet. Thus, I hope at the end of this class, I will get knowledge on to get alternative of spatial data using remote sensing plus analyse various topic across different scale using GEE.

Remote sensing is also an interesting field as it could produce wealth of information without direct contact. Don’t you think learning remote sensing makes us have the eye of the bird

even beyond? I mean we agree that remote sensing offers perspectives far beyond what our human eyes can naturally perceive : *allowing to see things from above and to see the unseen of the naked eye.*

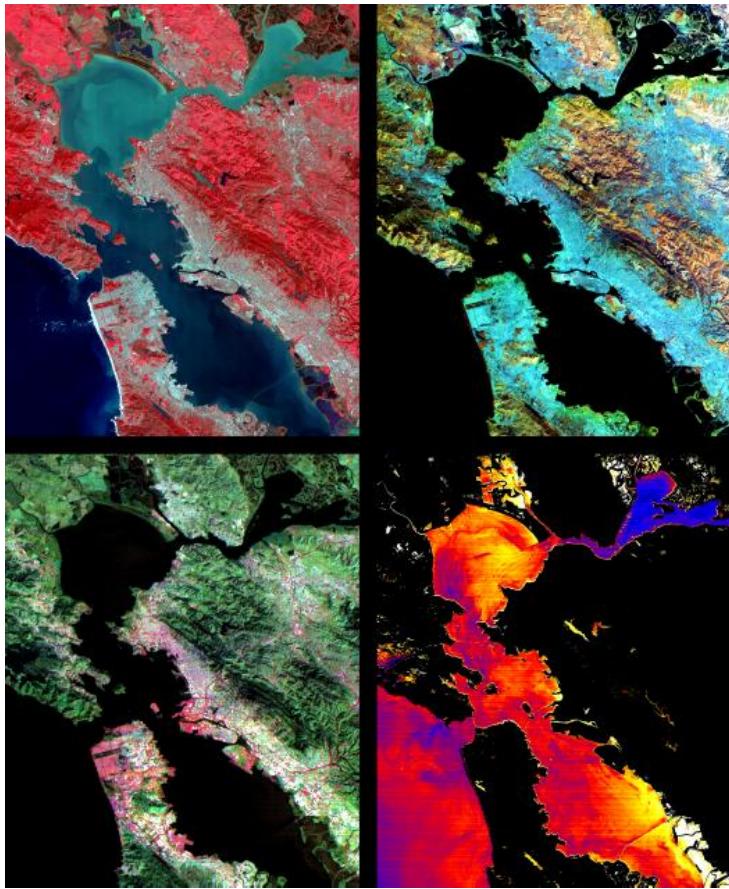


For example, see the ASTER images of San Francisco Bay below it highlights different object such as vegetation (upper left); soil & rocks in mountainous area (upper right); urban materials (lower left) ; and water temperature (lower right).

Figure 1 : ASTER images of San Francisco. source : [NASA/JPL](#)

Practically, learning this course will, hopefully, help me address the challenges I faced during my previous work in Indonesia. For example, while working on a project focused on health-care accessibility across hundreds of small islands, we struggled to obtain real-time data to identify which islands were inhabited and which were not. Additionally, we faced challenges in determining which islands had ports suitable for docking ships. I believe that applying remote sensing data is both cost- and time-efficient in helping the government maintain more precise and up-to-date data, which is particularly important in world's largest archipelago country like Indonesia.

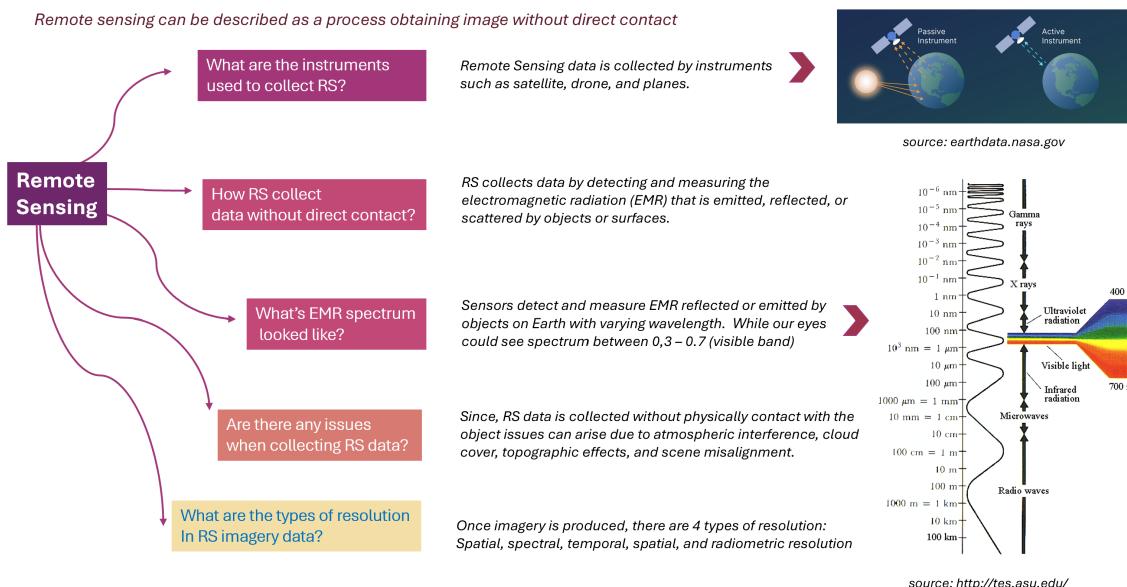
Feel free to explore my site to learn more about my learning experience. Hope it helps!



2 Getting to Know Remote Sensing

2.1 Summary

This week, the lecture covers an introduction of remote sensing, such as its vast application, instruments, collection method, and things we have to consider when we deal with remote sensing data. I tried to make the summary using visualization below to make it easier to understand.



During the practical, we explore several tools to deal with remote sensing data such as SNAP (Sentinel Application Platform) and R-studio to plot spectral signature. We are also introduced with 2 imagery : Sentinel-2A and Landsat-8. It is interesting how this two imagery has a global coverage and for FREE. Both of them has spectral bands that could be useful for vegetation monitoring, land cover classification, and agricultural applications. We could benefit from Sentinel-2 frequent observations to monitor rapid changes. Meanwhile, Landsat data allows us to do large areas and long-term vegetation monitoring as it has extensive historical archive and consistent global coverage. Below I discussed the application of both Landsat and Sentinel in a vegetation analysis.

2.2 Application

Landsat for monitoring accross vast region : Detecting of vegetation evolution across China Urban Development

- When I mentioned Landsat have a vast amount of historical data, @han2025 explores this historical archive of 30 years Landsat data (spanning of landsat 5 to 8) on 2.125 city to monitor the vegetation evolution, using reflective bands such as Blue, green, red, NIR and SWIR (1 and 2) and highlighting vegetation characteristics using NDVI, EVI, and OSAVI. The NDVI and RGB bands were further processed to derive texture variables, including variance, contrast, entropy, angular second moment, and correlation. These texture metrics capture spatial patterns and fine-scale structural details of urban vegetation that may not be visible through spectral bands alone. The findings will classify vegetation in urban area, whether it is decreasing or increasing over time. I genuinely believe this finding has the potential to serve as a framework for evaluating the effectiveness of the government's long-term plan on urban greening. For instance in my country, Indonesia we have long-term regional planning that spanning for 20 years (reviewed every 5 years), the analysis will help the policy maker to formulate more measured-target.

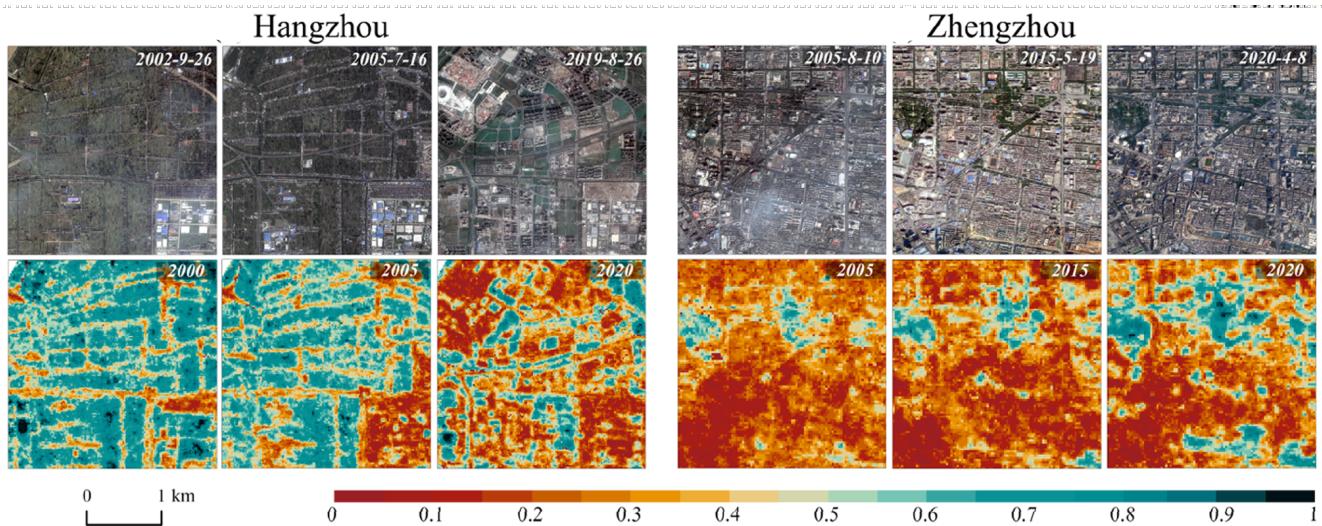


Figure 1: A sample result shows urban vegetation degradation in Hangzhou and an increase in vegetation in Zhengzhou. source : (Han et al. 2025).

However, Landsat is an optical imaging system that is often susceptible to cloud cover and has limitations in distinguishing different vegetation types based solely on spectral characteristics. Imagine studying a mountainous region where cloud cover is persistent—using optical images like Landsat for vegetation monitoring and identification would be challenging.

To address this issue, Li et al. (2023) utilized Sentinel-1, which operates with C-band Synthetic Aperture Radar (SAR), enabling vegetation mapping under all weather conditions. The SAR data from Sentinel-1, when combined with the optical imagery of Sentinel-2, allows for the production of high-resolution maps that effectively differentiate bamboo forests from other vegetation types. This integration helps overcome the limitations of optical data in vegetation monitoring, where mixed spectral characteristics often lead to uncertainty in distinguishing bamboo from other forest types.

2.3 Reflection

Remote sensing provides diversity in data source, but.... will the implementation be easy?

After exploring the application of the two selected satellites during practical this week, I have concluded that remote sensing data is particularly effective for analyzing large-scale and long-term variations. It can also help mitigate the high costs of manual data collection across vast regions. This insight made me reflect on a similar challenge in my country, Indonesia. We often have challenges to find dataset for spatial analysis as we rely much on vector data, if any it would be outdated. Using remote sensing data not only allows us to have more updated data but also allows us to explore various potential variables derived from satellite imagery : diversity in data sources. In my country, the Geospatial Agency has already utilized remote sensing to address time and cost constraint during data collection. It is legally published as a baseline regulation to address the conventional method for the collection and processing of geospatial data on shallow marine habitat. They use remote sensing to create a work map for on-field mapping habitat, enabling to be time and cost efficient.

However I must admit despite the potential of remote sensing, its adoption in the government sector—especially at the local level—remains limited. From my experience, this is largely due to a lack of human resources with the skills to process and analyze remote sensing data. For end users, the adoption of remote sensing is linked with the information they can bring on the table, and often might depend on the leadership, budgetary constraints, procedures and personal capacity[@nationalacademyofsciences.2003].

2.4 References

3 Xaringan and Quarto Book

Lecture this week reminded me of one of powerful figure in Uchiha Clan, the one who can manipulate reality once he activates this-so-called Xaringan. Well, but this Xaringan is not related to figures in Konoha's world but related to a certain library in R Studio that enable us to create neat HTML slides in R.

I think a presentation is basically a way to communicate insights to the audience, and a great presentation may even “hypnotize” the audience. Is that one of reasons it called Xaringan?

3.1 Summary

```
xaringanExtra::embed_xaringan(url = "https://nooriza16.github.io/Xaringan/Xaringan.html",
```

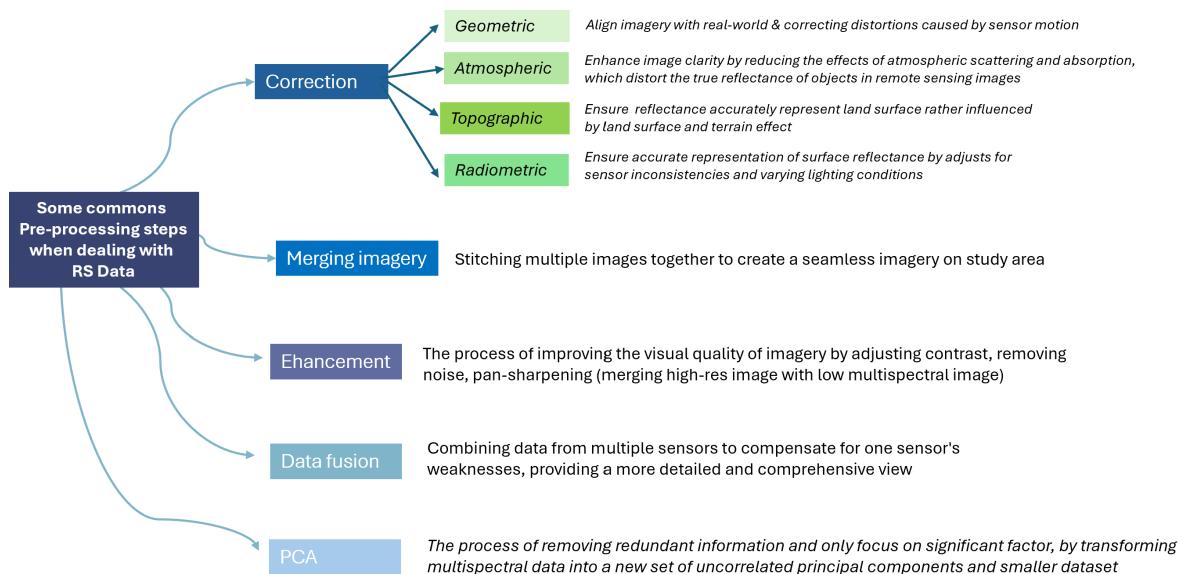
3.2 Reflections

For someone who is not familiar with html, learning Xaringan is definitely challenging compared to powerpoint, as we just usually click tabs on power point. Honestly, I still consider power point provides more themes and more visualization effects that is easily to access compared to Xaringan. However, as I delved further I realize that using Xaringan is providing us with flexibility even such as positioned our picture. So far, I feel like Xaringan is best at incorporating snippet code on presentation or interactive features that usually too heavy to load in power point. Besides, it helps me to give a sense of what html look like.

4 Image Correction

4.1 Summary

Here is my note based on this week's lecture that explores steps people usually do in image correction. Although sometimes we get a "ready-to-use data" without the need to going through all of these process, having the basic understandings of these steps would help us understand the quality our data.



This diagram is created as a note of CASA023 Lecture Week 3

I also add several new terminologies, based on my own note during the lecture, related to image processing during this week's summary

Reflectance reflectance is basically the amount of light when a surface reflect the light, while **radiance** radiance is the amount of light captured by sensor after interacting with Earth's surface

Digital Number	A raw value for a given pixel that represents the intensity of radiation received in a specific spectral band.
Digital Object Sub-straction (DOS)	Digital number is important because it serves as a basis for image classification, for example digital number close to 0 represents object that absorbs much incoming light (low reflectance) such as water bodies or shadows.
Digital Object Sub-straction (DOS)	DOS is an atmospheric correction method that subtracts pixel values based on the amount of difference between digital values of dark objects (usually water bodies) with their corresponding reflectance
Collection level, and tier	Collection will use the Landsat case to explain this terminology. In Landsat the collection would be named as Collection 1 and 2, it represents the sequence of launching time and their mission : Landsat 2 is the latest. Level 1 is a scaled digital number, while level 2 is further processed data. Meanwhile, tier 1 is the highest quality data from Landsat and suitable for time series analysis.
	source: https://www.usgs.gov/landsat-missions/landsat-collection-2-level-1-data

4.2 Application

In this week application, I would like to explore the application of remote sensing on my favorite topic : data unavailability. In developing countries like Indonesia, maintaining updated and comprehensive data is challenging due to time and cost constraints. A study in Pakistan also mentioned similar challenges, particularly data related to socioeconomic condition that often limited and sporadic Arshad et al. (2023). Honestly, this topic piques my interest because Bill Gates came across an article in The Washington Post and seems fascinated with the idea of using remote sensing to identify wealth, as he shared on X.



Figure 1 : Tweet about satellite imagery on detecting poor region. Source : <https://x.com/yohaniddawela/status/1857398431146299437>

Arshad et al. (2023) addressed this issue using imagery that automatically could derive insights about poverty. They use publicly accessible high-resolution satellite imagery (Google Maps API with 16 zoom) and Landsat 7 (low resolution data). Google Maps API provides high-resolution imagery to identify man-made features like buildings and highway which are indicators of development levels meanwhile Landsat 7 is used to train Convolutional Neural Networks (a method of feature extraction from imagery in Machine Learning) in identifying nightlight bin (low to high). Areas with higher levels of economic activity and development tend to have more lights at night. This method will produce a map that indicates a poverty and development level compared to poverty line. The results are then validated using actual socioeconomic data from surveys. Below is a map where each point represents a poverty clusters (10x10 km area), comparing predicted and actual data. Green points indicate clusters above the poverty line, while red points indicates the opposite.

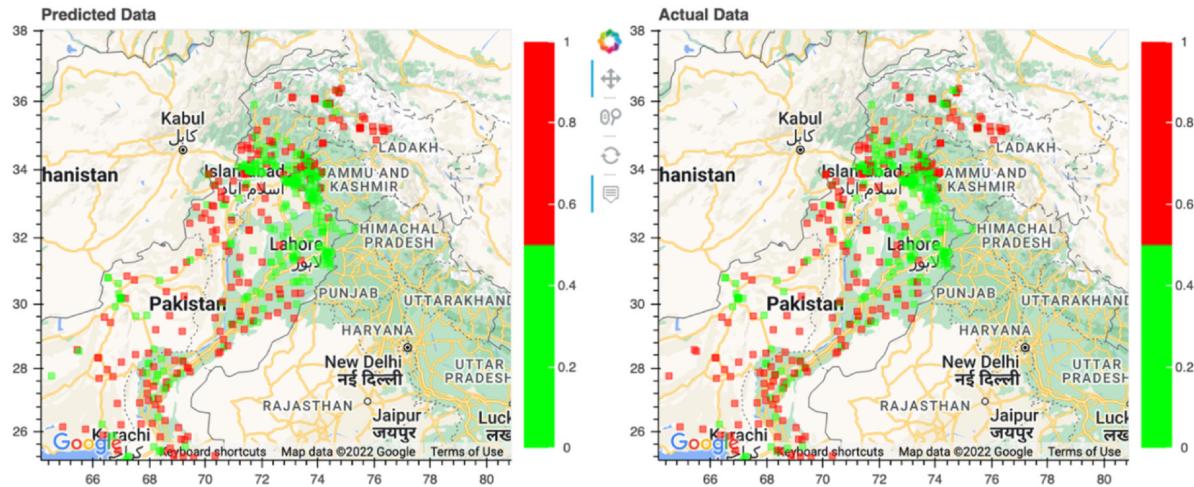


Figure 2 : Socioeconomic conditions compared to poverty line in Pakistan. Source : (Arshad et al. 2023).

When reflecting on the map above, I would prefer to visualize the classification results as polygon rather than points, as they were more intuitive. Additionally, it would be nice to map the difference across the years, Ben Abbes et al. (2024) just do this ! They use multispectral images (Landsat 5, 7,8) and Nightlight images (from Defense Meteorological Satellite Program (DMSP) and the Visible Infrared Imaging Radiometer Suite (VIIRS)) in Southeast Brazil. The classification result is represented using estimated wealth index and they could even map the socioeconomic transformation across 10 years in a single map!

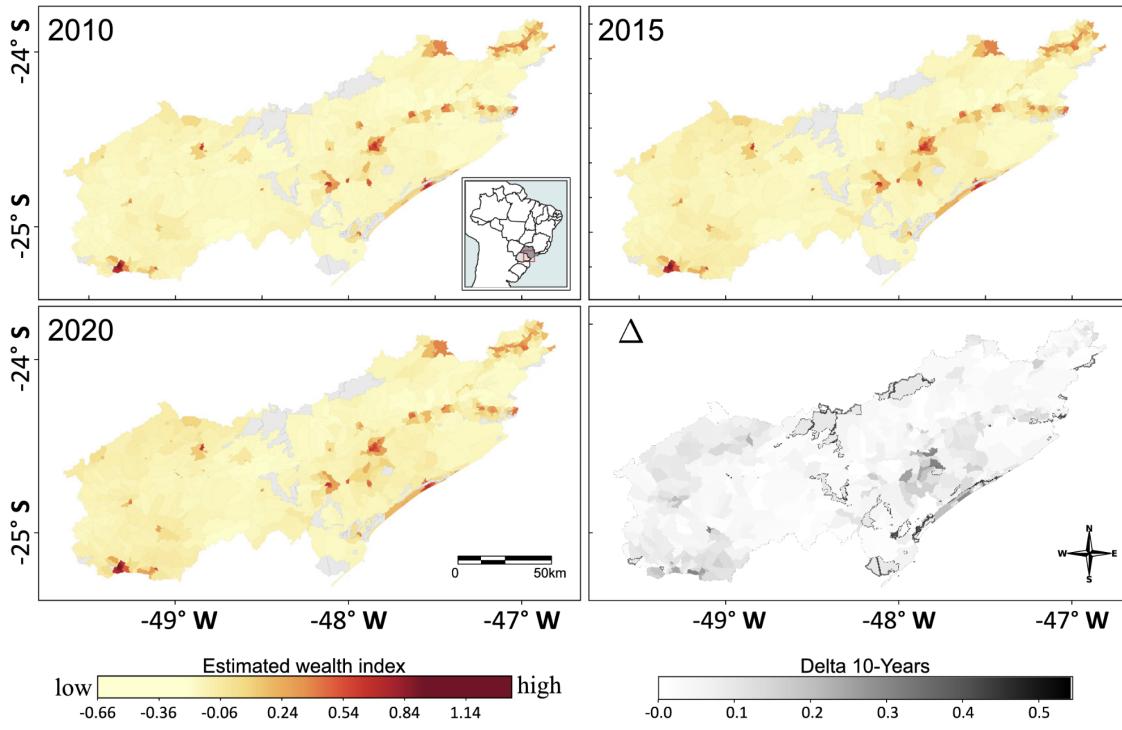


Figure 3 : Spatio-temporal mapping of wealth index estimations in Southeast Brazil. Source : Ben Abbes et al. (2024).

4.3 Reflection

I think performing Remote Sensing correction on R Studio is quite challenging, as I become more used to using ‘button’ in Remote Sensing application such as ENVI or SNAP. After this week’s lecture, I genuinely think that Remote Sensing is quite complex as it is not only an image but beyond the imagery each pixel is composed by digital number and it could be linked and better interpreted using regression too. Understanding the image classification using machine learning is also quite challenging for me, as they use new terminologies that I haven’t heard before such as convolutional neural network, epoch, and data training.

Honestly, remote sensing combined with machine learning quite scare me off. However, I try to look beyond the methodology instead focusing on how explanatory remote sensing can be when combined with classification and prediction task, tasks machine learning good at. I think looking beyond the methodology and focusing on the exciting application has also helped me me thrive on managing challenge during this Master’s. After reading the paper, I tried to delve further into the combination of remote sensing and machine learning. If I could turn back the time, I would like to deploy this combination to make my works faster. I recalled

during my works years ago, the project needed to identify thousand ports across hundred of islands in Indonesia, and we did that manually ! If I understand correctly, I could combine image classification techniques, such as convolutional neural networks, with high-resolution imagery to detect local ports used for docking ship (typically made of wood or cement) in the ocean.

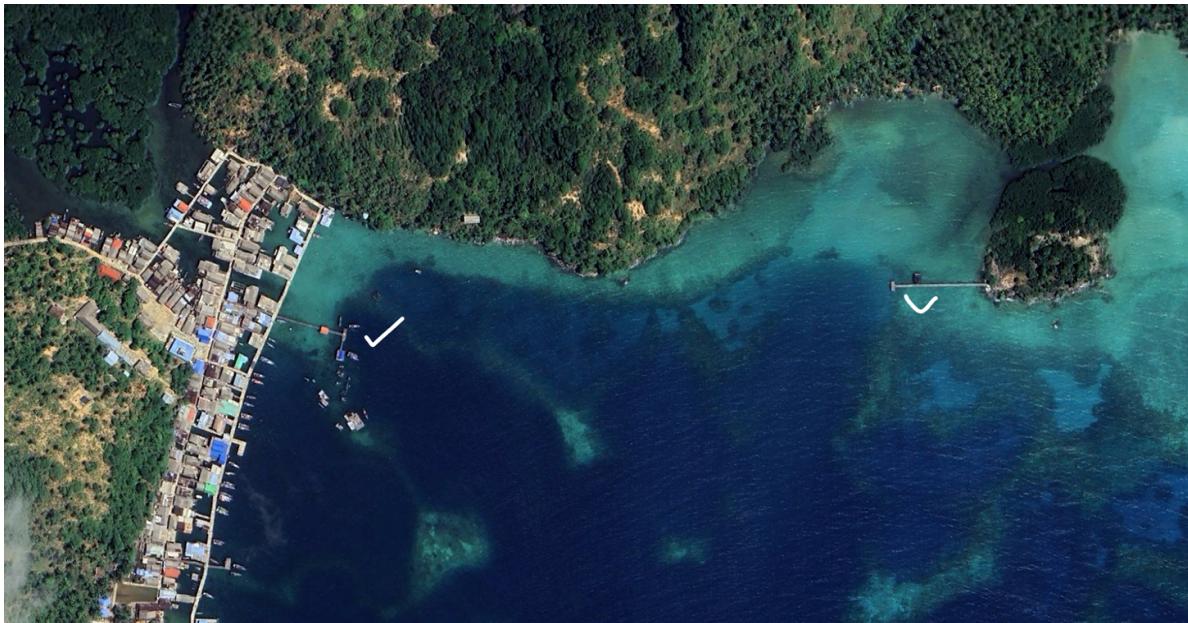


Figure 3 : Google Earth Imagery portraying local ports in Anambas Islands. Source : Google Earth, 2024

5 Policy

Project Case : A New Relocated Capital City of Indonesia ; From Jakarta to Nusantara



Source : www.nytimes.com

5.1 Summary

Recently Indonesia planned to move its capital city from Jakarta (in Java islands) into Penajam Paser Utara City (Borneo Islands), as the current capital city, Jakarta, faced an issue of sinking, land subsidence, overcrowding, low air and water quality (Bappenas 2021). The term Nusantara is used to name this new capital city, symbolizing the varied geographic settings and cultural diversities of Indonesia.

As for the time this published, Nusantara Development is on the phase 2 (2025-2029) that involved strengthening core area (housing, office, commercial zone). Thus, in the time being, Jakarta will still remain the capital of Indonesia until the Presidential Decree on the transfer of the capital to Nusantara is issued. The issuance of this decree will depend on the readiness of the new capital city, including the preparation of all supporting systems such as infrastructure, human resources, and governance systems.



Figure 1: The Relocation Settings and Vision. source: (Capital Authority 2024)

As the development is still in the initial stage, the detailed planning documents haven't been launched yet. Thus, I use available published documents regarding the detail of Nusantara's Development which all of them are publicly available, such as [Nusantara Sustainable Development Goals \(SDGs\) Voluntary Local Review Baseline \[2024\]](#) and [Nusantara Biodiversity Management Master Plan \[2024\]](#)

Policy

The new capital city, Nusantara, is designed as a **forest city**, with 75% of its designated area being green space. This design aims to create a harmonious blend of urban development and biodiversity hotspots (Borneo Island, where Nusantara is located, is famous for its tropical rainforests). However, the design of being a forest city, its proximity to the rainforest, and its drive on landscape change would present significant **challenges**. One of the major concerns is the increasing likelihood of mosquito-borne diseases (such as **malaria**) spreading in the new capital, which are prevalent in tropical regions Surendra et al. (2024). Since malaria is both a global and local challenge, certain goals should be considered to support Nusantara's sustainability, such as:

A. Global Goals : Sustainable Development Goals (SDGs) 3.3 : Fight Communicable Diseases

The SDGs propose achievable global in combating malaria with target that include reducing incident, mortality rates, eliminate malaria in 35 countries by 2030 and prevent resurgence of the disease in a malaria-free country. Meanwhile, Indonesia's estimated malaria incidence per 1000 population at risk is still on range between 1-50 incidents per 1000 population in 2023. To achieve target of Global Goals, (WHO 2021) have launched global technical strategy for malaria with framework such as:



Figure 2 : SDGs Goal and WHO Technical Strategy

B.National Level : National Action Plan for Acceleration of Malaria Elimination 2020-2026

Translating the global goals on malaria elimination, Indonesia's Ministry of Health (Ministry of Health and Control 2023) had proposed recommendations, including the new capital city

such as:

- Malaria elimination policies and implementation need basic research, operational support, and efficient technology development.
- Malaria elimination policies and implementation need basic research, operational support, and efficient technology development
- Planning and implementing malaria elimination activities are based on district endemicity stratification.
- Allocating funds from the central government to improve case finding, surveillance and vector control around Nusantara area

C. Provincial Level : Governor Regulation 53/2023 Guidelines For Malaria Elimination Implementation Number 5 point 2

- Suppress the endemic incidents (high and medium risk)
- Eliminate the number of incidents
- Maintain free malaria status (low-none risk)

5.2 Application

Remote Sensing as Baseline for detecting malaria hotspot

In malaria elimination projects, remote sensing can serve as a crucial baseline data source for mapping malaria hotspots by integrating climatic and land-use factors. @wimberly2021 proposed a framework that leverages Earth observation products to identify mosquito habitats based on climate conditions, human activities, and specific land-use patterns.

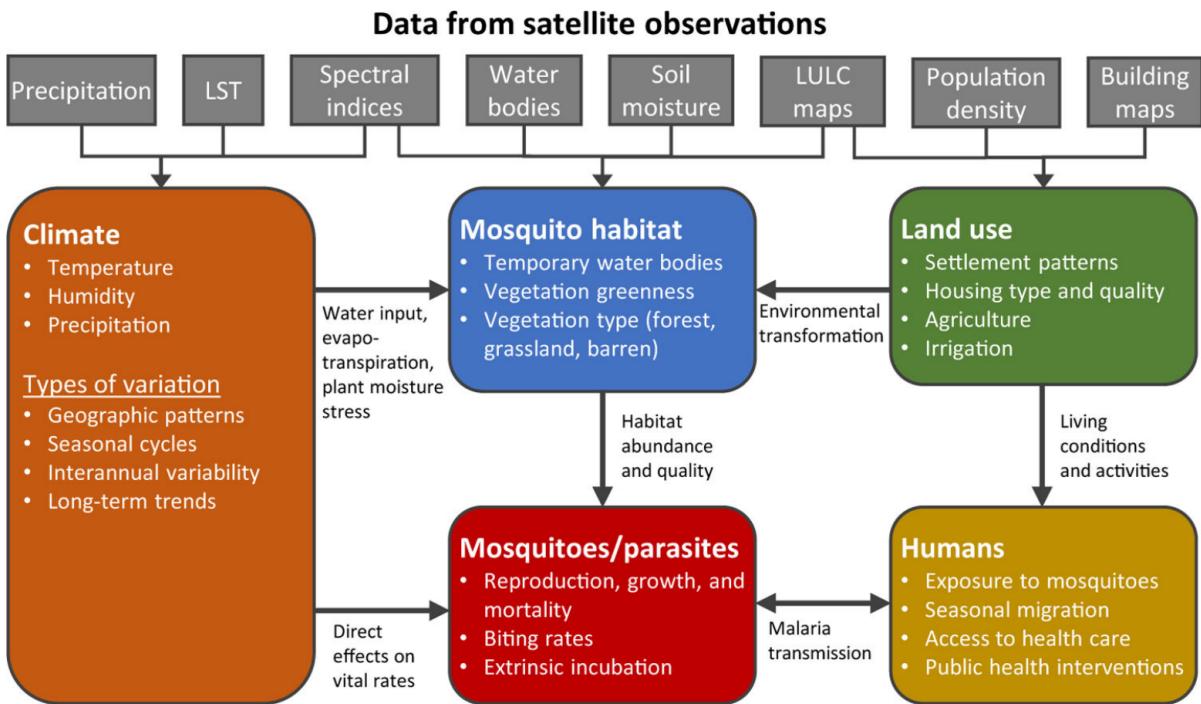


Figure 2: Framework in which Remote Sensing used in Malaria studies. source : Wimberly et al. (2021).

To address policy mentioned in section 2, I underlined some dataset that could be used to the analysis:

Data	Purpose
• Sentinel-2 (rainy season)	<ul style="list-style-type: none"> <i>Highlight water bodies and wetlands</i> – These serve as proxies for mosquito breeding sites. <i>Vegetation and land cover</i> – Provides insight into potential mosquito habitats.
• Digital Elevation Model (DEM) Data/Topography	<ul style="list-style-type: none"> <i>Surface temperature</i> – Acts as a proxy for mosquito activity. Helps to provide topography to identify potential inundation areas, which could influence mosquito breeding patterns.

Data	Purpose
Microsoft Open Buildings	<ul style="list-style-type: none"> • Useful as a proxy for human settlements and potential exposure risk.
Rainfall Data	<ul style="list-style-type: none"> • The rainfall season can be considered as a timeframe for analysis. However, if locally recorded rainfall data from the Indonesian Climatic Institution is available, it could help refine the identification of rainfall patterns, allowing for a more informed selection of the time series.

5.3 Reflections

During this week, I got a lot of reflections as I finally found lecture that explicitly bridging the gap of ‘academics’ to real-world policy. My reflections would be:

1. Combining remote sensing with GIS

Since Nusantara is still uninhabited, we could model nearby settlements to investigate the remote sensing framework. By combining the results with malaria incident data, we can validate our classification—analyzing what percentage of high-risk areas have recorded incidents and which have not. While global and local malaria elimination frameworks mention aggregating incident data and risk levels, they do not explicitly emphasize mapping. Using maps, we can overlay malaria hotspots with incident data, land use, and socio-economic factors. As (Naserrudin, Yong, et al. 2023) notes, people are exposed to malaria due to professions that require them to venture deeper into the forest.

2. Remote Sensing and GIS is good, but enriching the analysis with **affected communities** make it better

Beyond remote sensing data, incorporating local knowledge can improve the analysis. Understanding how affected communities respond to malaria provides insight into the effectiveness of mitigation efforts (Naserrudin, Lin, et al. 2023). These communities have lived near rainforests for generations and are directly affected, making their experiences valuable for practical prevention strategies.

3. Implementation challenges, the need for collaboration

One the most important key-takeaway from the lecture is that “some academics papers are too technical, without clearly addressed policy; some policy don’t include academic findings they could benefit for.” This condition lead to a gap between academics and urban governance. However, in my observation during my work with the government

the potential cause is human resources (make the adoption of academics finding hard to implement), annual budget cycles (governments prioritize immediate results and may be reluctant to invest in the long-term experimental processes typical of academia). Bridging the gap on malaria prevention requires collaboration and commitment not only between epidemiologists, healthcare, and geospatial analysts but with the governments to ensure research translates into actionable policies.

5.4 References

6 Introduction to GEE

6.1 Summary

This week's material is all about GEE (Google Earth Engine). In a simple definition, GEE is a cloud platform that allows us to access satellite imagery for the whole world and spatial-computation on Google for free. GEE hosts massive amounts of satellite imagery and we as a user request their imagery data and analyse it on the cloud-platform without have to worry about the capability of our local machine.

Basically, GEE has an architecture that collects user input (client side) and then process this input (server side). In GEE we could manage both raster and feature (vector) data. We input command on GEE mostly using Java Script programming language. As someone enthusiastic with GEE, the need to learn 'a new' programming language almost discourages me, as I worry that I might mix up all the programming languages I have learnt before.

However, when I looked at the way GEE articulates data structures, I found it quite similar to Python, with just a few additional keywords, such as 'var' to denote variable. Below are some basic pieces of information about the GEE language that we need to know from Andy's lecture. Check out this introduction from Google if you're still unconvinced [[here](#)].

```
// All the javascript you need to know (almost)

var number = 1

var string = 'Hello, World!'

var list = [1.23, 8, -3]
print(list[2])

var dictionary = {
  a: 'Hello',
  b: 10,
  c: 0.1343,
  d: list
}

print(dictionary.b)
print(number, string, list, dictionary)
```

Figure 1 : Basic javascript in GEE. source : [Andrew Maclachan Github](#)

When I was first introduced to the idea that GEE is efficient in managing raster data, I was curious about how this efficiency works. For example, if I want to use Landsat 8 (30 m resolution) to analyze the whole UK, it would require processing every 30 m pixel across the whole region, which I assumed would take longer time. However, what makes GEE faster dealing with those data actually because there is a pyramid of reduced resolution in GEE, which able to optimizes performance using lower-resolution versions of the data when full detail is not necessary and also resampling method to adapt the analysis requirement. Thus when performing analysis, GEE does not just take the original image resolution, but GEE adjusts image resolution based on our output needs. In summary, when you are zooming out and zooming in, GEE retrieves different resolution of imagery. Below is the picture of GEE interface, one thing that I notice the most is code editor : a space to write our code.

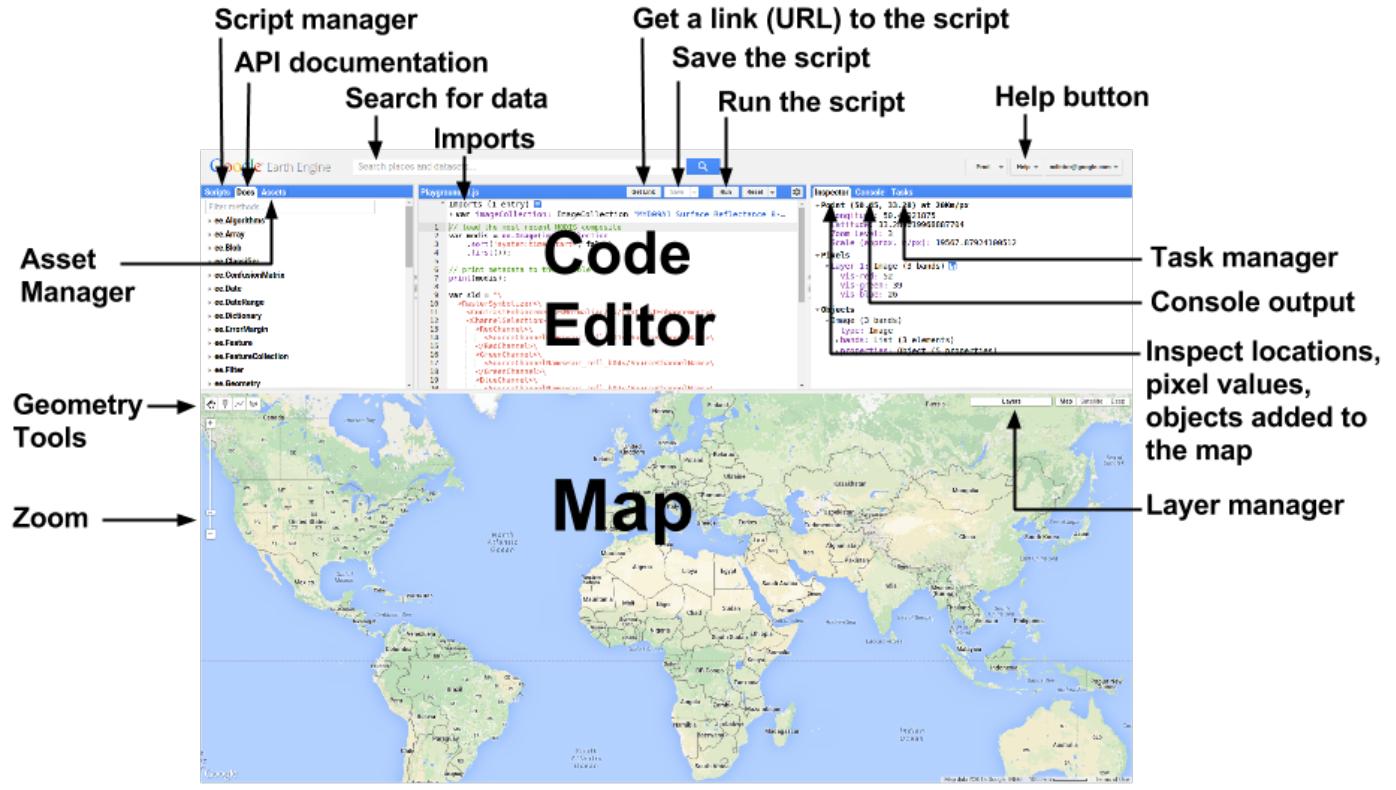


Figure 2 : GEE Interface. source : [GEE Beginner cookbook](#)

6.2 Application

In this section, I would like to explore the GEE strength on providing vast archive of imagery to investigate unplanned settlements during urban transformation projects. The vast historical of imagery enables researcher to look back at various critical date for development plans. @yilmaz2024 used collection of Landsat data (Landsat 5, 6, 7) from March 1983 to February 2007. The study selected 13 urban renewal project areas in Istanbul and examined existing project data and archival satellite images. Supervised classification was performed on 10 urban renewal project areas using the GEE platform to investigate urbanization status before development plans.

With analysis of temporal imagery, the result show a contrasting initial condition of urban renewal's targetted area and its projection. For instance, in Gaziosmanpasa district it is projected to be a forested area in the future, however after looking back up to one year before the approval of the development plans turns out it is highly urbanized area. The contrasting initial condition and urban renewal project in this area will present challenge to the success of the plan, as the development of greening area will fiercely compete with massively built up

area. This finding also underlined the issues that projected plan is not carefully consider the dominant land use of the area, which will most likely shape the future land use.

LULC status of urban renewal project areas before development plans and planning activities in project areas.

Project	1980 Metropolitan Land Use Plan planned land use type	LULC maps before development plans (Frames refer to urban renewal project areas)	Development plan information	
			Spatial plan type	Acceptance date/Planned use
Gaziosmanpasa Sarigöl I	Current Residential Area		Slum reclamation and prevention binding implementation plan	January 21, 1987 Area to be afforested
Gaziosmanpaşa Sarigöl II	Current Residential Area		Slum reclamation and prevention binding implementation plan	January 21, 1987 Area to be afforested

Figure 3 : Comparision of existing land use and its projection. Source : (Yilmaz and Alkan 2024)

[add urban expansion in china (“Urban Renewal Mapping: A Case Study in Beijing from 2000 to 2020,” n.d.)

6.3 Reflection

The most exciting part of using Google Earth Engine (GEE) for me is the ability to share analysis results through web map applications. Clients are usually more engaged when they can interact with a visual platform that helps them understand the project better. Remote sensing analysis is often more complex than regular vector processing, so I imagine the difficulty on to deliver the results with its key preprocessing step. What makes it even harder is, we often need to switch between remote sensing desktop apps and present the results.

By using a web map, we make it easier for clients, even those without a technical background, to explore and understand the analysis. The interactive map will provide a clearer view of the project, allowing clients to see the bigger picture and stay engaged throughout the process. However, I must admit GEE is only a tool to make best use of it we need a prior fundamental concept of remote sensing, such as spectral band, enhancement, fusion, PCA, and other remote sensing pre-processing steps. Because in every script we write there should be a thinking process on why we do reduction, why we do enhancement, basically need our judgement to navigate each steps.

6.4 References

7 Classification I

7.1 Summary

Today's lecture is exploring remote sensing classification. Several dataset and applications mentioned are:

- Urban expansion/sprawl with landsat data
- Surface temperature (land and sea) with sentinel 3
- Major air pollutants with sentinel 5
- Urban green spaces (high medium res, Lidar)
- Monitoring and illegal practices landsat data
- Forest fires using landsat TM image

There are several steps to extract land cover (do classification) from earth observation data. We usually combined remote sensing with machine learning techniques that include training data, build classification model, output. Today's class focus on classification and regression trees (CART) as a method of classification. It composes with 2 things :

- a. Classification trees (yes no/ discrete values)

In this method, we will have predictors and target. The decision is mapped with conditions that led to yes or no decision

- b. Regression trees (predict continuous dependant variable)

Meanwhile, in linear regression we try to fit into the regression line, however if we have data that isn't fitted into the data we subset this data into smaller chunks. When creating a decision tree, we would be mix all the leafes of several categories, quantified using gini impurity. And we aim to find the lowest gini impurity. We decide the breaks by creating imaginary vertical line among the data and identify the lowest sum of square residuals.

Things to be paid attention at = OVERFITTING, dealing with this could be

- a. Set the min number of pixels/ observations, Only split observations with **min number of 20**.
- b. Pruning

- c. Removes leave, increase alfa, and find the lowest tree score. The idea is could we keep the accuracy, when we do generalizations
- d. Dividing data : training (creating trees) and testing data (applied the trees with testing data) find the lowest sum residuals.

See the picture :

We could pick some pixels as training data and validation, the machine learning will predict the value among this train and validation pixels.

Random forest

As decision tree is not good with new data, thus we could use another method is random forest. I like to use the chatgpt analogy on this:

A Decision Tree is like asking one expert for advice.

A Random Forest is like asking 100 experts, each with a unique take, then averaging their answers.

Applying into imagery:

- a. Supervised (giving them a train data)

Support vector machine, the idea is like a logistic regression. It finds the **best boundary (decision boundary)** that separates different classes in the data. The benefit of this method is it allows some misclassification

- b. Unsupervised (I have train data, and I want 10 classification method) > see recording (43')

- a. Usually refers to clustering/ k-means.

- b. Isodata (same with kmeans but add several inputs such as (cluster, iterations)

Repeat until all pixels were classified

7.2 Application

7.3 Reflection

7.4 References

8 Classification II

8.1 Summary

Today's lecture is exploring remote sensing classification. Several dataset and applications mentioned are:

- Urban expansion/sprawl with landsat data
- Surface temperature (land and sea) with sentinel 3
- Major air pollutants with sentinel 5
- Urban green spaces (high medium res, Lidar)
- Monitoring and illegal practices landsat data
- Forest fires using landsat TM image

There are several steps to extract land cover (do classification) from earth observation data. We usually combined remote sensing with machine learning techniques that include training data, build classification model, output. Today's class focus on classification and regression trees (CART) as a method of classification. It composes with 2 things :

- a. Classification trees (yes no/ discrete values)

In this method, we will have predictors and target. The decision is mapped with conditions that led to yes or no decision

- b. Regression trees (predict continuous dependant variable)

Meanwhile, in linear regression we try to fit into the regression line, however if we have data that isn't fitted into the data we subset this data into smaller chunks. When creating a decision tree, we would be mix all the leafes of several categories, quantified using gini impurity. And we aim to find the lowest gini impurity. We decide the breaks by creating imaginary vertical line among the data and identify the lowest sum of square residuals.

Things to be paid attention at = OVERFITTING, dealing with this could be

- a. Set the min number of pixels/ observations, Only split observations with **min number of 20**.
- b. Pruning

- c. Removes leave, increase alfa, and find the lowest tree score. The idea is could we keep the accuracy, when we do generalizations
- d. Dividing data : training (creating trees) and testing data (applied the trees with testing data) find the lowest sum residuals.

See the picture :

We could pick some pixels as training data and validation, the machine learning will predict the value among this train and validation pixels.

Random forest

As decision tree is not good with new data, thus we could use another method is random forest. I like to use the chatgpt analogy on this:

A Decision Tree is like asking one expert for advice.

A Random Forest is like asking 100 experts, each with a unique take, then averaging their answers.

Applying into imagery:

- a. Supervised (giving them a train data)

Support vector machine, the idea is like a logistic regression. It finds the **best boundary (decision boundary)** that separates different classes in the data. The benefit of this method is it allows some misclassification

- b. Unsupervised (I have train data, and I want 10 classification method) > see recording (43')

- a. Usually refers to clustering/ k-means.

- b. Isodata (same with kmeans but add several inputs such as (cluster, iterations)

Repeat until all pixels were classified

8.2 Application

8.3 Reflection

8.4 References

9 Synthetic Aperture Radar

9.1 Summary

Today's lecture is exploring remote sensing classification. Several dataset and applications mentioned are:

- Urban expansion/sprawl with landsat data
- Surface temperature (land and sea) with sentinel 3
- Major air pollutants with sentinel 5
- Urban green spaces (high medium res, Lidar)
- Monitoring and illegal practices landsat data
- Forest fires using landsat TM image

There are several steps to extract land cover (do classification) from earth observation data. We usually combined remote sensing with machine learning techniques that include training data, build classification model, output. Today's class focus on classification and regression trees (CART) as a method of classification. It composes with 2 things :

- a. Classification trees (yes no/ discrete values)

In this method, we will have predictors and target. The decision is mapped with conditions that led to yes or no decision

- b. Regression trees (predict continuous dependant variable)

Meanwhile, in linear regression we try to fit into the regression line, however if we have data that isn't fitted into the data we subset this data into smaller chunks. When creating a decision tree, we would be mix all the leafes of several categories, quantified using gini impurity. And we aim to find the lowest gini impurity. We decide the breaks by creating imaginary vertical line among the data and identify the lowest sum of square residuals.

Things to be paid attention at = OVERFITTING, dealing with this could be

- a. Set the min number of pixels/ observations, Only split observations with **min number of 20**.
- b. Pruning

- c. Removes leave, increase alfa, and find the lowest tree score. The idea is could we keep the accuracy, when we do generalizations
- d. Dividing data : training (creating trees) and testing data (applied the trees with testing data) find the lowest sum residuals.

See the picture :

We could pick some pixels as training data and validation, the machine learning will predict the value among this train and validation pixels.

Random forest

As decision tree is not good with new data, thus we could use another method is random forest. I like to use the chatgpt analogy on this:

A Decision Tree is like asking one expert for advice.

A Random Forest is like asking 100 experts, each with a unique take, then averaging their answers.

Applying into imagery:

- a. Supervised (giving them a train data)

Support vector machine, the idea is like a logistic regression. It finds the **best boundary (decision boundary)** that separates different classes in the data. The benefit of this method is it allows some misclassification

- b. Unsupervised (I have train data, and I want 10 classification method) > see recording (43')

- a. Usually refers to clustering/ k-means.

- b. Isodata (same with kmeans but add several inputs such as (cluster, iterations))

Repeat until all pixels were classified

9.2 Application

9.3 Reflection

References

Arshad, Arslan, Junaid Zulfiqar, Muhammad Hassan Zaib, Ahsan Khan, and Muhammad Jahanzeb Khan. 2023. “Mapping Socioeconomic Conditions Using Satellite Imagery: A Computer Vision Approach for Developing Countries.” *Journal of Economy and Technology* 1 (November): 144–63. <https://doi.org/10.1016/j.ject.2023.11.001>.

- Bappenas, Kementrian PPN. 2021. "Buku Saku Pemindahan Ibu Kota Negara."
- Ben Abbes, Ali, Jeaneth Machicao, Pedro L. P. Corrêa, Alison Specht, Rodolphe Devillers, Jean P. Ometto, Yasuhisa Kondo, and David Mouillot. 2024. "DeepWealth: A Generalizable Open-Source Deep Learning Framework Using Satellite Images for Well-Being Estimation." *SoftwareX* 27 (September): 101785. <https://doi.org/10.1016/j.softx.2024.101785>.
- Capital Authority, Nusantara. 2024. "Nusantara Sustainable Development GOals (SDGs) Voluntary Local Review Baseline."
- Han, Yuan, Jianhua He, Xiaoping Du, Xiao Han, and Yaolin Liu. 2025. "Reconstructing Urban Vegetation Evolution in China Using Multimodal Deep Learning and 30-Years Landsat Archive." *Urban Forestry & Urban Greening* 103 (January): 128582. <https://doi.org/10.1016/j.ufug.2024.128582>.
- Ministry of Health, Directorate General of Disease Prevention, and Control. 2023. "National Action Plan for Acceleration of Malaria Elimination 2020-2026 (Revision)." https://malaria.kemkes.go.id/sites/default/files/2024-08/National%20Strategic%20Plan%20Revision_Malaria_29%20Mei%202023.pdf.
- Naserrudin, Nurul Athirah, Pauline Yong Pau Lin, April Monroe, Richard Culleton, Sara Elizabeth Baumann, Shigeharu Sato, Bipin Adhikari, et al. 2023. "Exploring Barriers to and Facilitators of Malaria Prevention Practices: A Photovoice Study with Rural Communities at Risk to Plasmodium Knowlesi Malaria in Sabah, Malaysia." *BMC Public Health* 23 (1): 1316. <https://doi.org/10.1186/s12889-023-16173-x>.
- Naserrudin, Nurul Athirah, Pauline Pau Lin Yong, April Monroe, Richard Culleton, Sara Elizabeth Baumann, Shigeharu Sato, Rozita Hod, Mohammad Saffree Jeffree, Kamrudin Ahmed, and Mohd Rohaizat Hassan. 2023. "Seeing Malaria Through the Eyes of Affected Communities: Using Photovoice to Document Local Knowledge on Zoonotic Malaria Causation and Prevention Practices Among Rural Communities Exposed to Plasmodium Knowlesi Malaria in Northern Borneo Island." *Malaria Journal* 22 (1): 166. <https://doi.org/10.1186/s12936-023-04603-5>.
- Surendra, Henry, Bimandra A. Djaafara, Helen D. Prameswari, Dedy Supriyanto, Ponco Waluyo, Setyo B. Basuki, Herdiana Herdiana, et al. 2024. "Mitigating Risks of Malaria and Other Vector-Borne Diseases in the New Capital City of Indonesia." *Nature Communications* 15 (1): 10575. <https://doi.org/10.1038/s41467-024-54891-x>.
- "Urban Renewal Mapping: A Case Study in Beijing from 2000 to 2020." n.d. <https://spj.science.org/doi/10.34133/remotesensing.0072>.
- WHO, World Health Organization. 2021. "Global Technical Strategy for Malaria 2016–2030, 2021 Update."
- Wimberly, Michael C., Kirsten M. de Beurs, Tatiana V. Loboda, and William K. Pan. 2021. "Satellite Observations and Malaria: New Opportunities for Research and Applications." *Trends in Parasitology* 37 (6): 525–37. <https://doi.org/10.1016/j.pt.2021.03.003>.
- Yilmaz, Okan, and Mehmet Alkan. 2024. "Assessing the Impact of Unplanned Settlements on Urban Renewal Projects with GEE." *Habitat International* 149 (July): 103095. <https://doi.org/10.1016/j.habitatint.2024.103095>.