LULC Analysis and Classification using Remote Sensing Data with Machine Learning

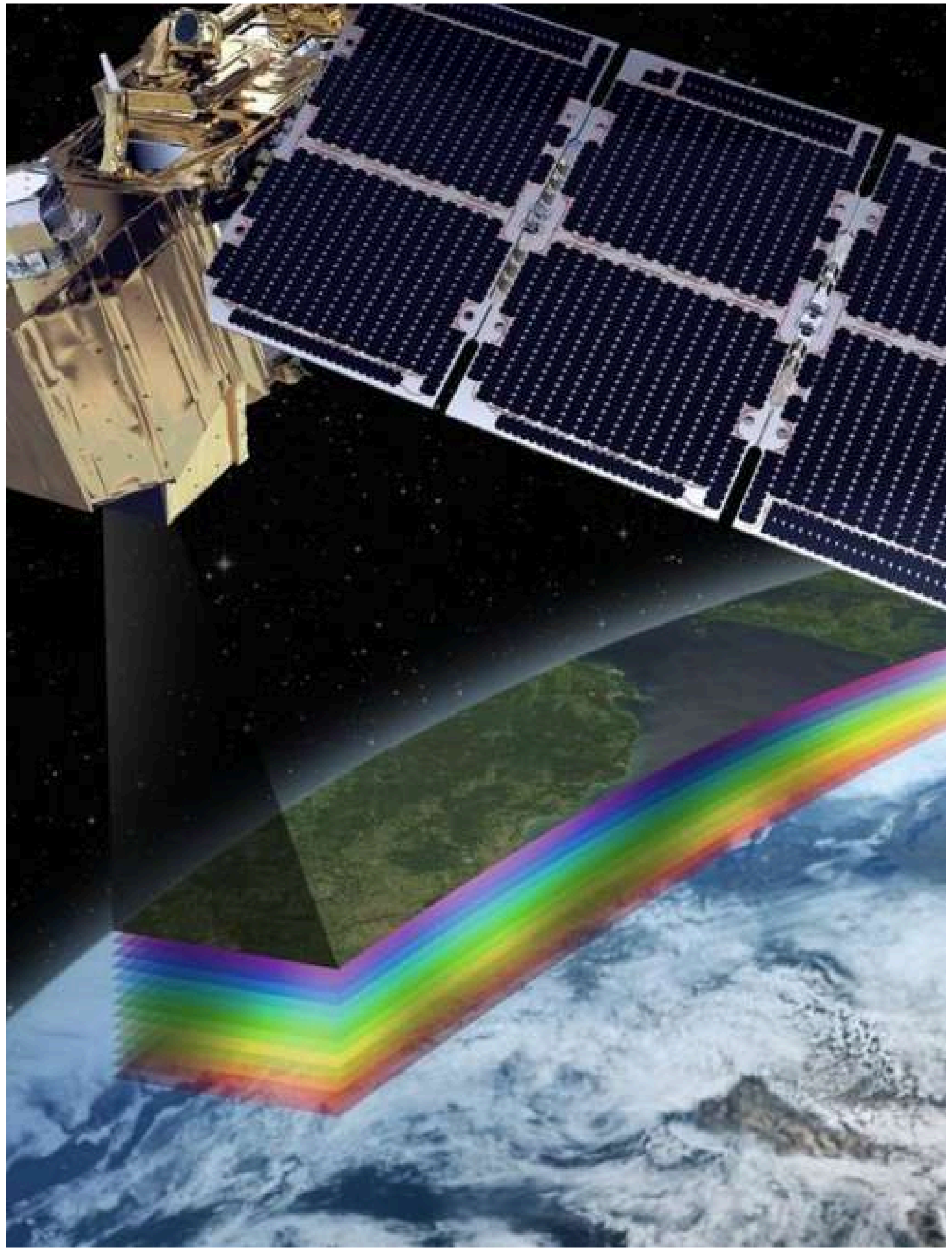
Workshop

Associate Prof. Sarawut Ninsawat

Mr. Thantham Kumyai

Ms. Krittaporn Iamsaing

GeoinfoLab, ICT, AIT

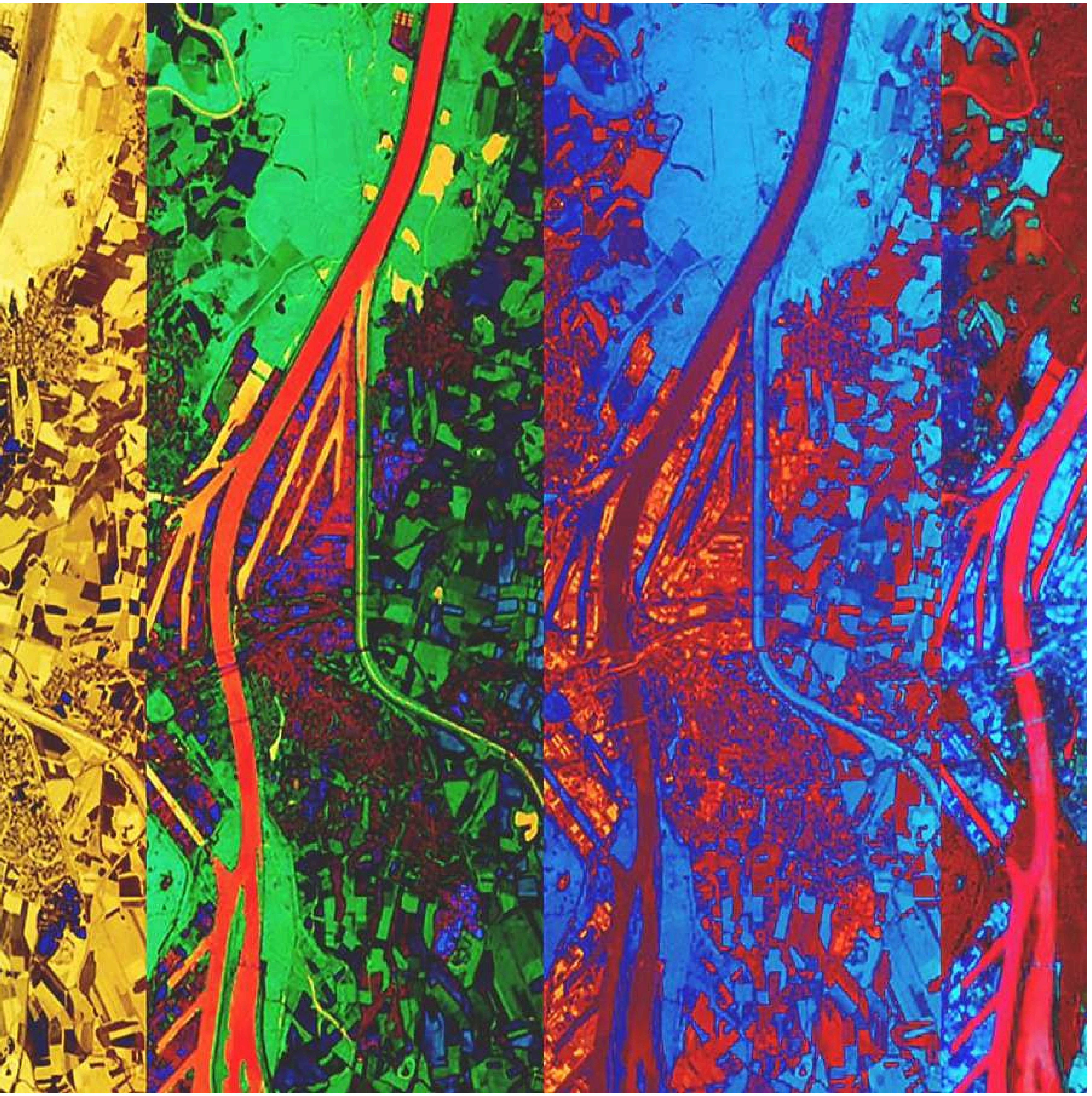


เนื้อหาการอบรม

Foundamental of Remote sensing

Remote Sensing Data Analysis

LULC Map Classification



Goals

01 Remote Sensing Programmatical Analysis

02 Machine Learning and Deep Learning

03 LULC Classification Workflow and Data Handling

เนื้อหาการปฏิบัติ

01 Basic Python and Array Math

02 Raster Math and Machine Learning

03 Deep Learning Principle

04 Model Training and Inferencing

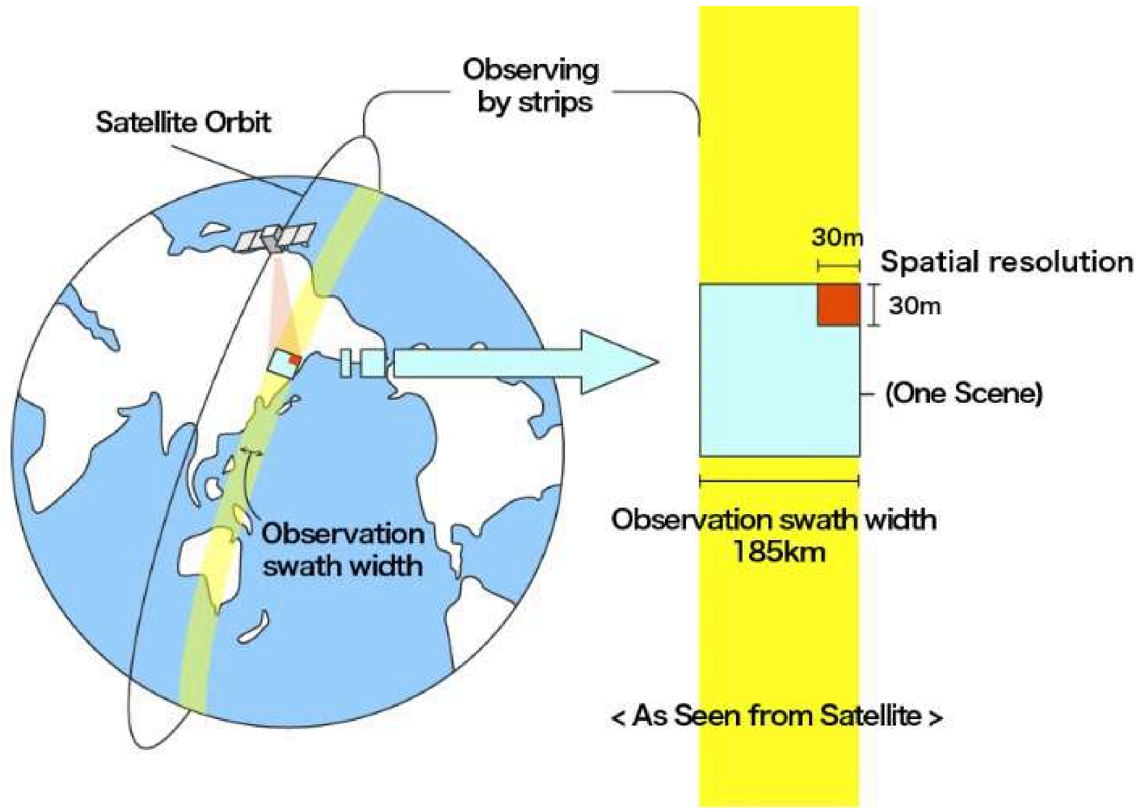
Hydro – Informatics Institute

Fundamental of Remote Sensing

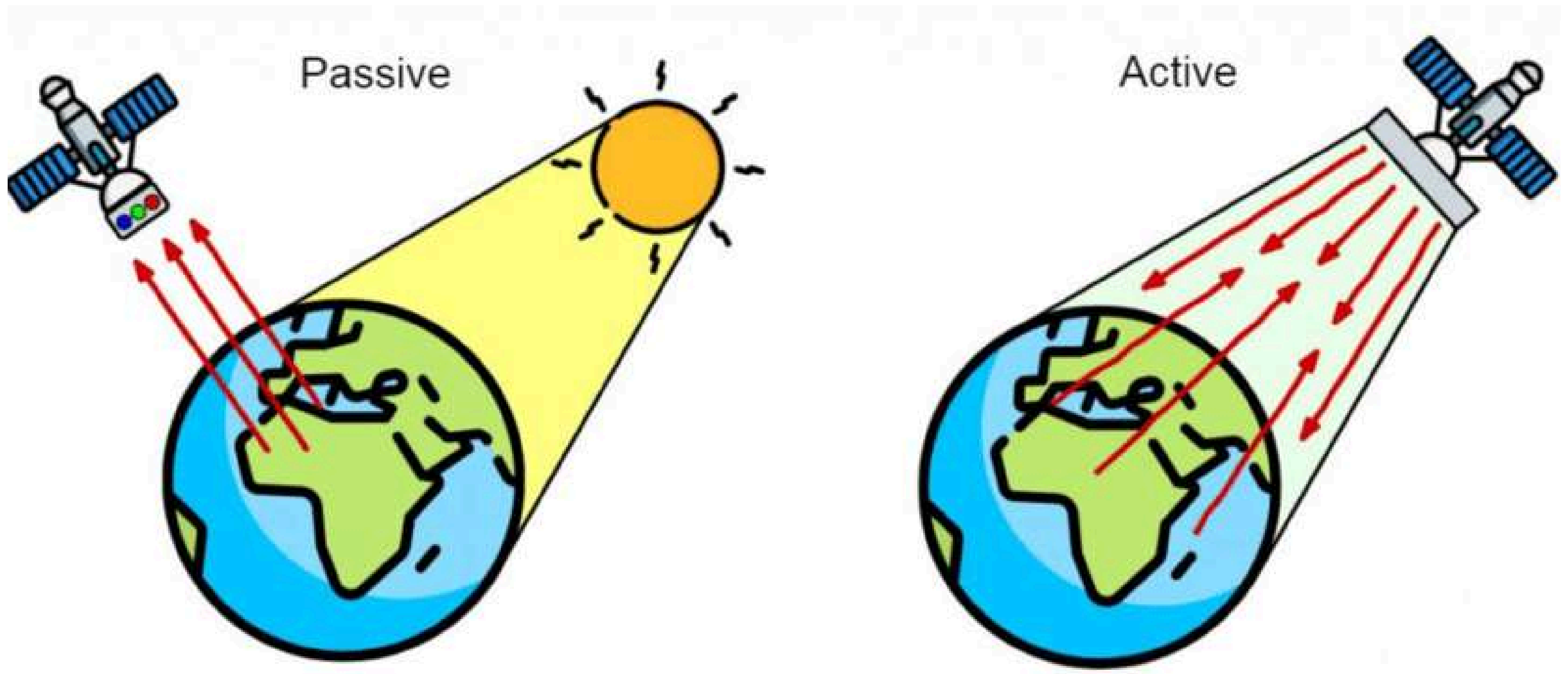
Crash Course

Remote Sensing Principle

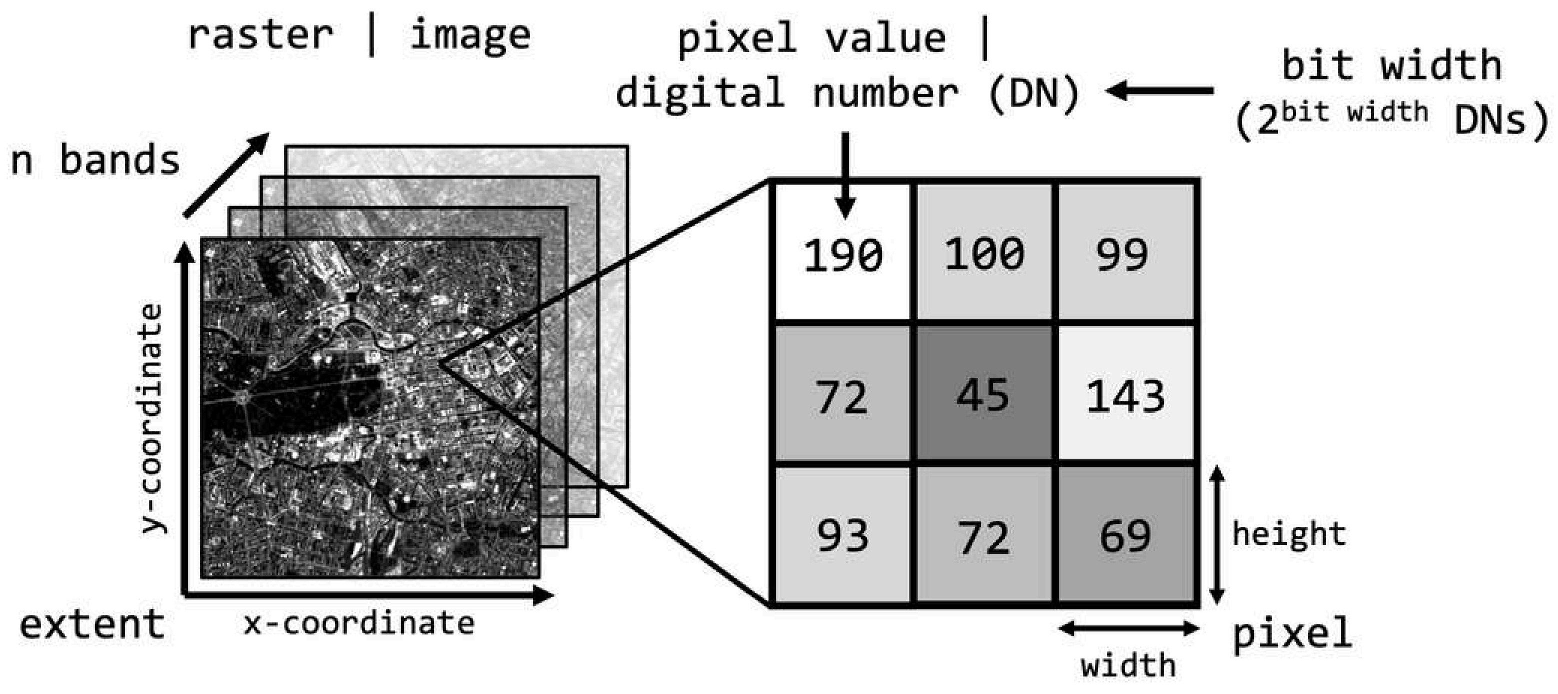
How to orbit



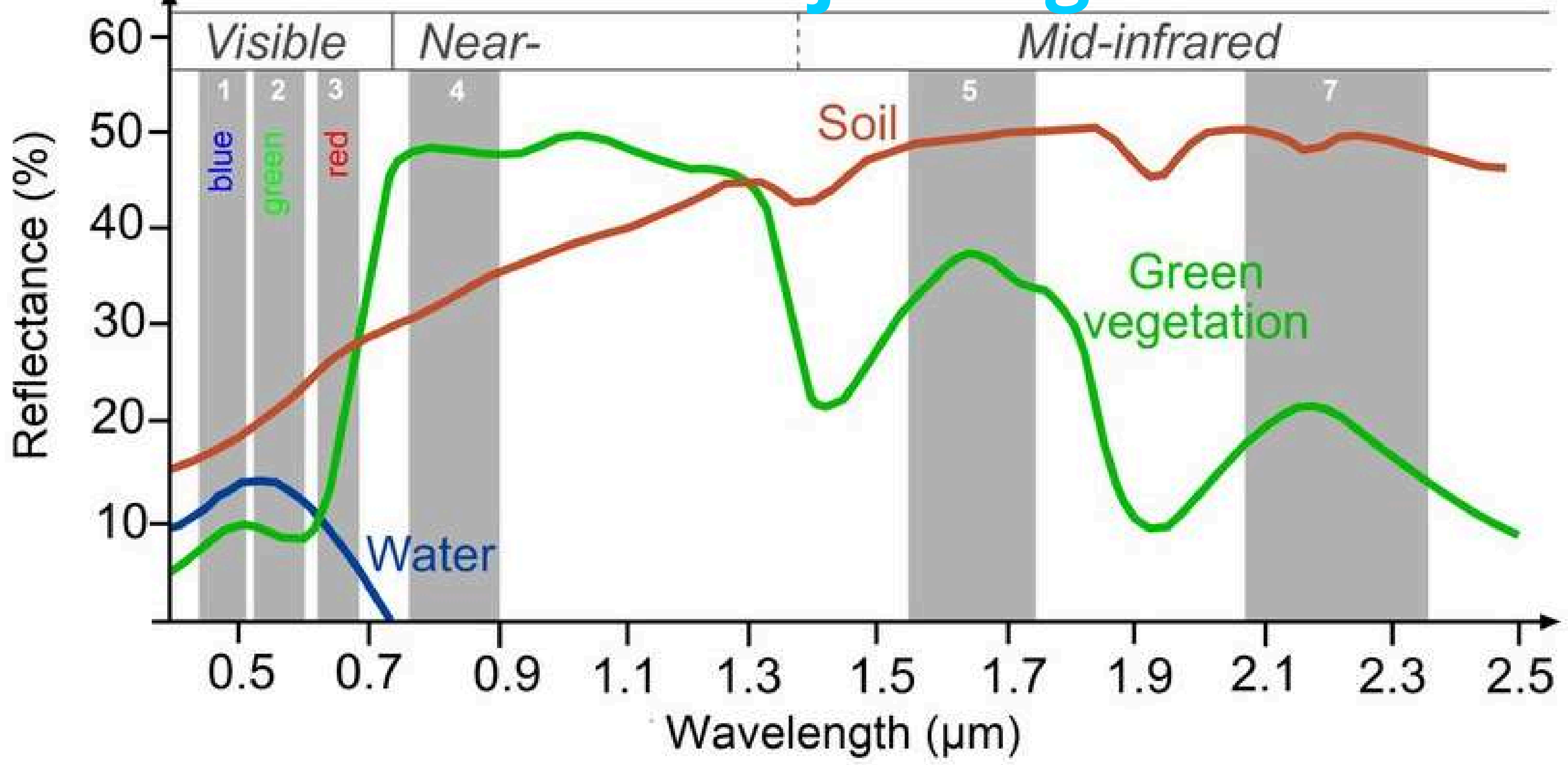
How to sense

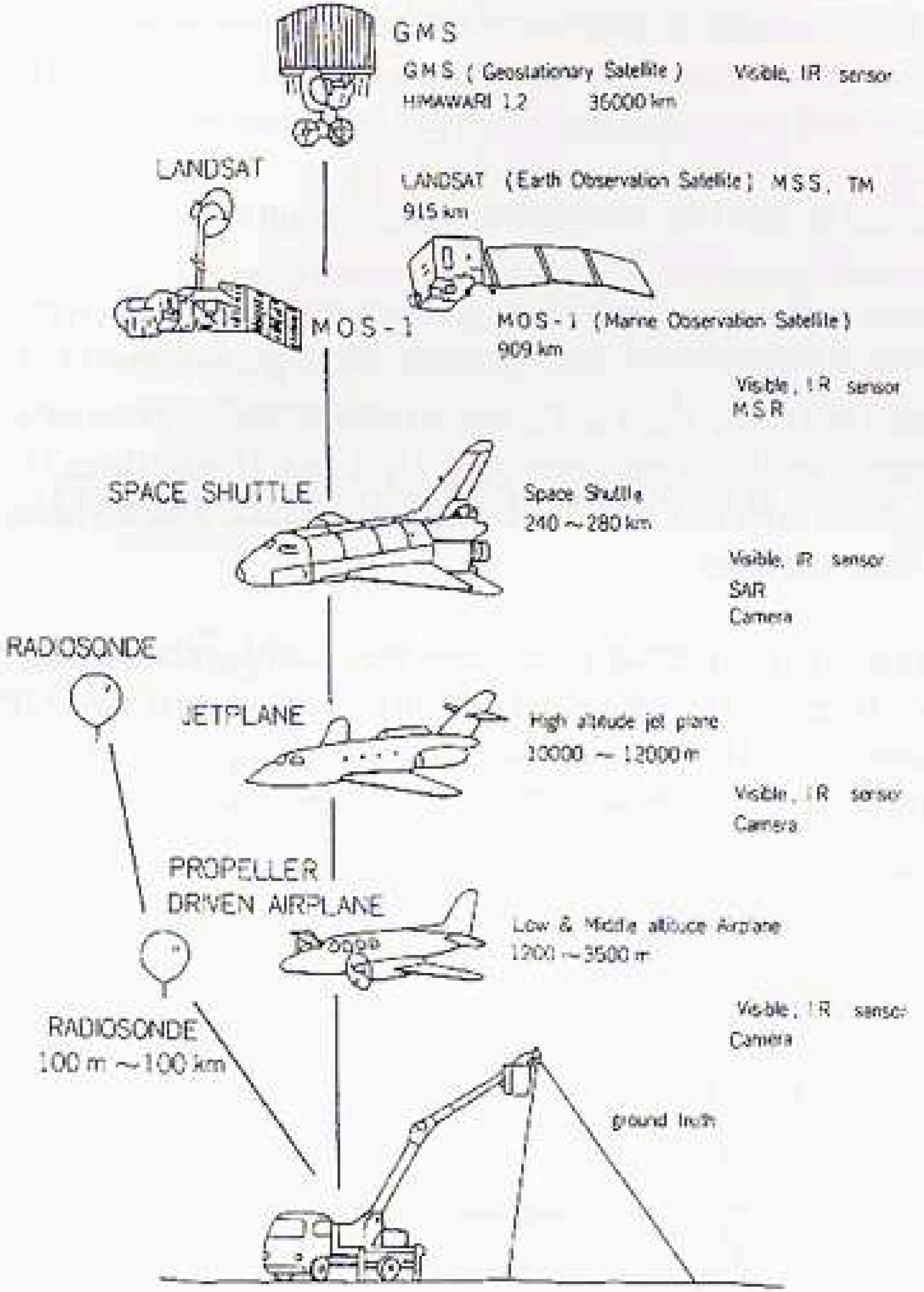


Data looks a like



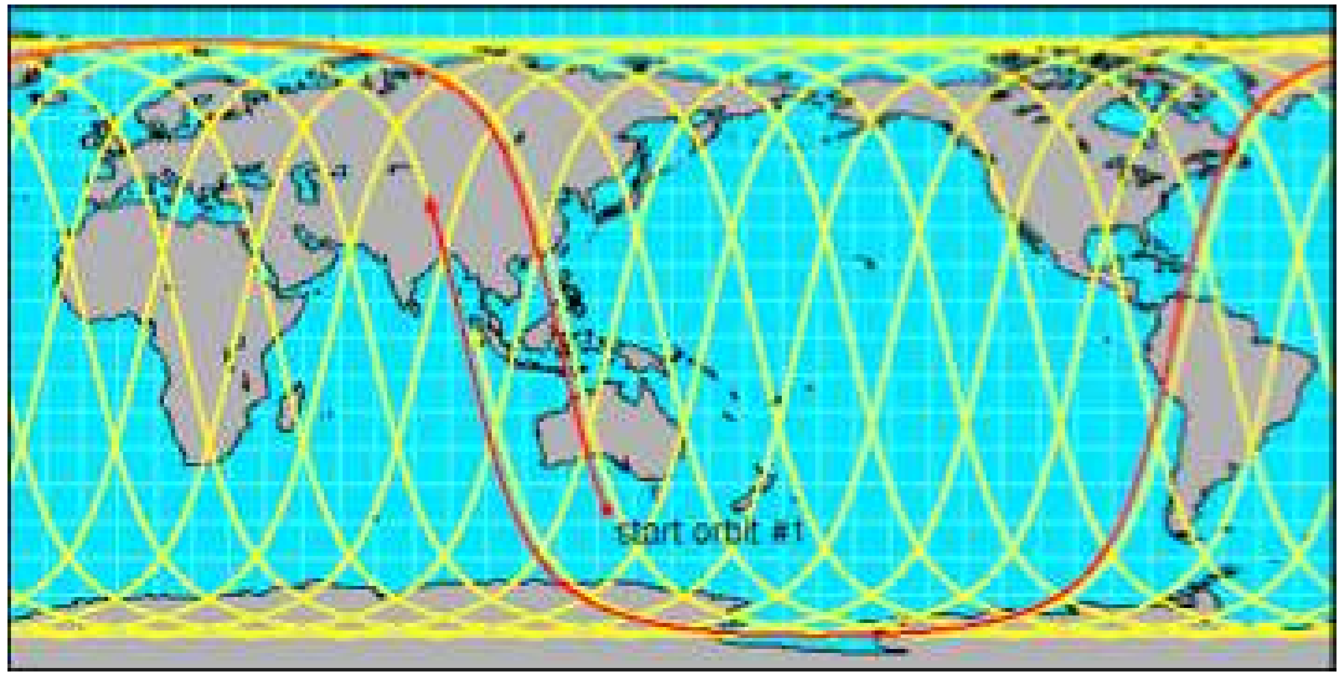
Object signatures



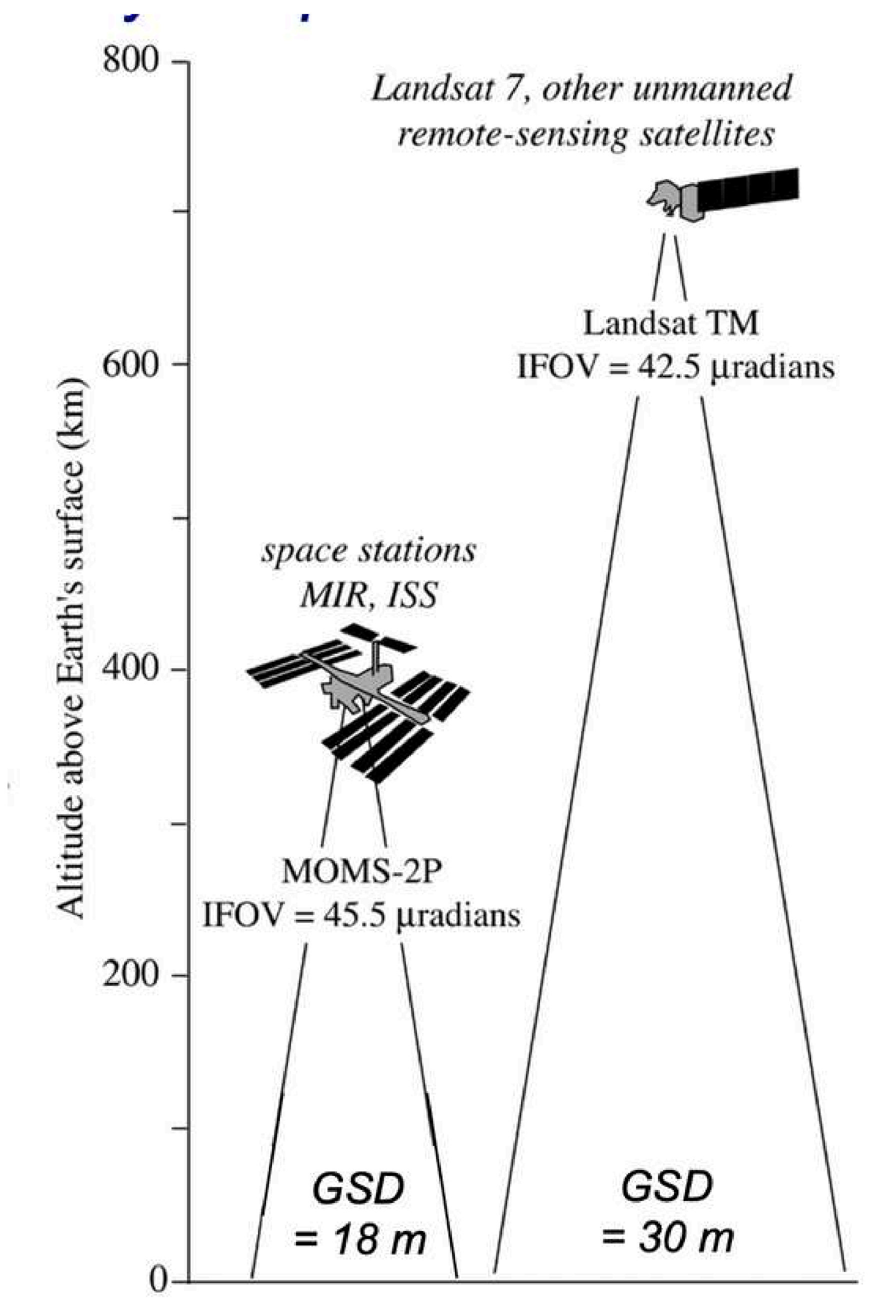


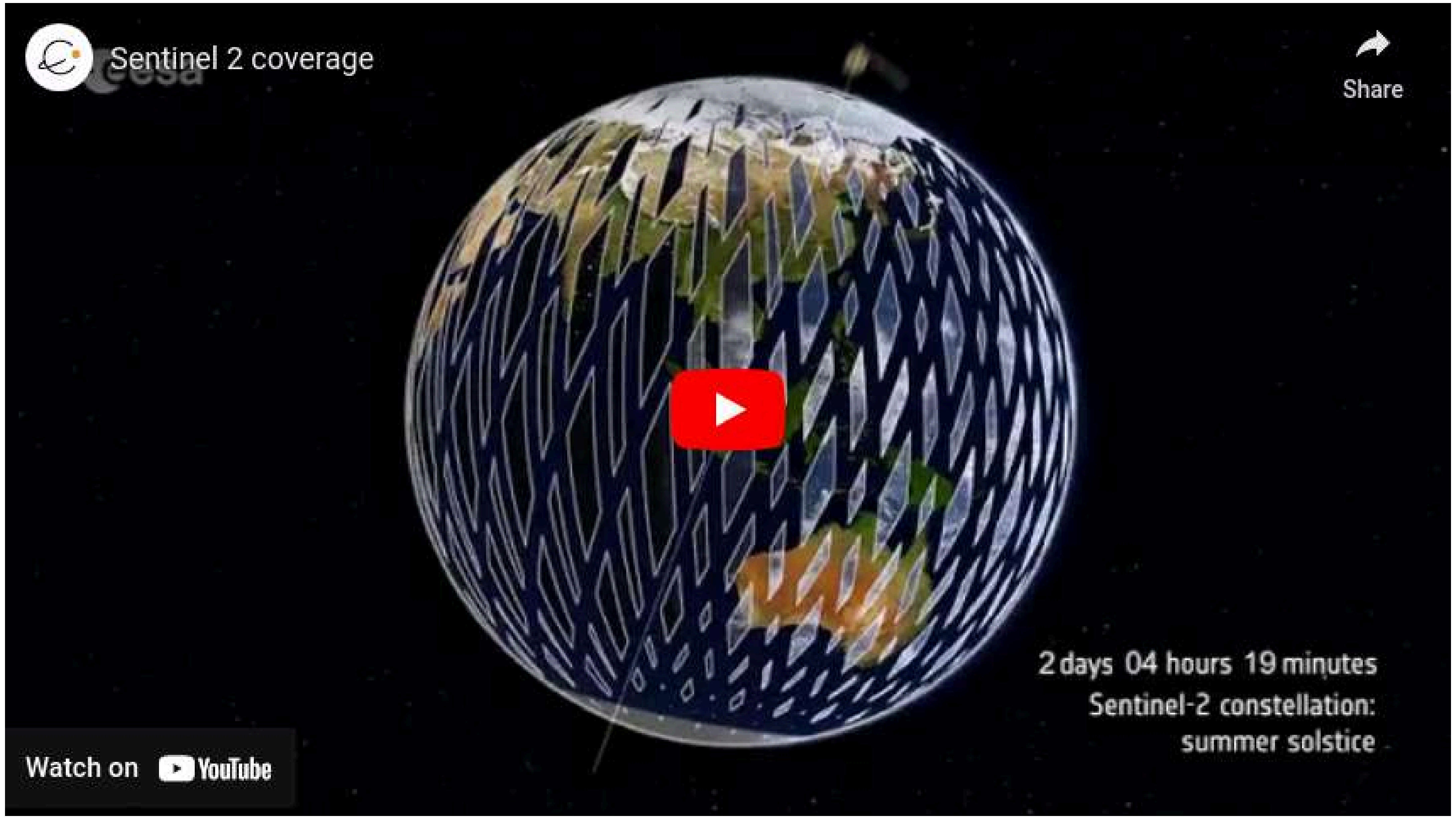
How to orbit

- Sun-synchronous orbit, Near polar-orbit, the satellite passes over the same part of the Earth at roughly the same local time.
  - why ??
- Altitude vs Spatial resolution vs Revisit



How to orbit





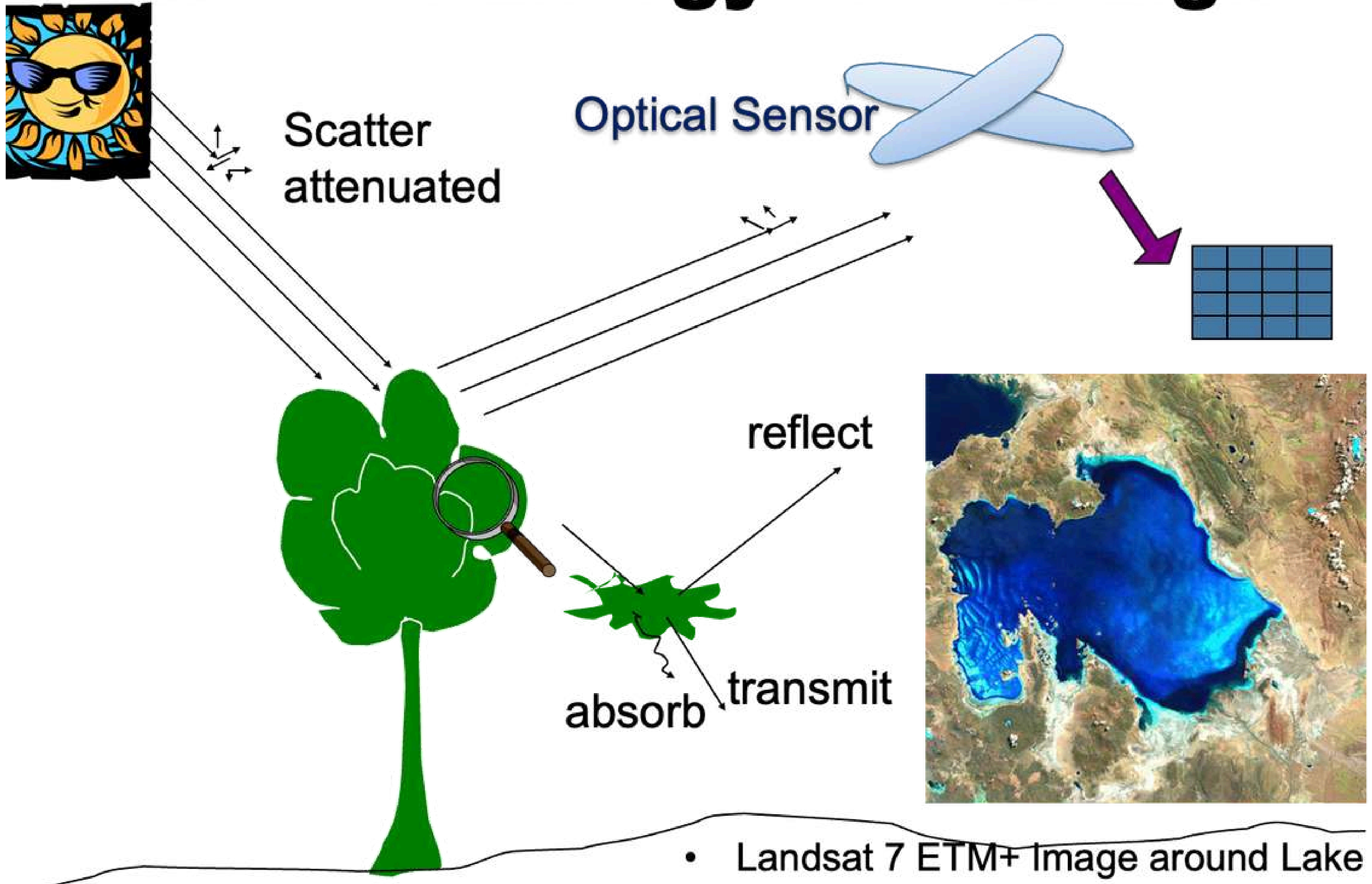
GSD : Ground Sampling Distance

Spatial resolution: ground surface area represented by one pixel

Temporal resolution : Revisit time or repeat cycle

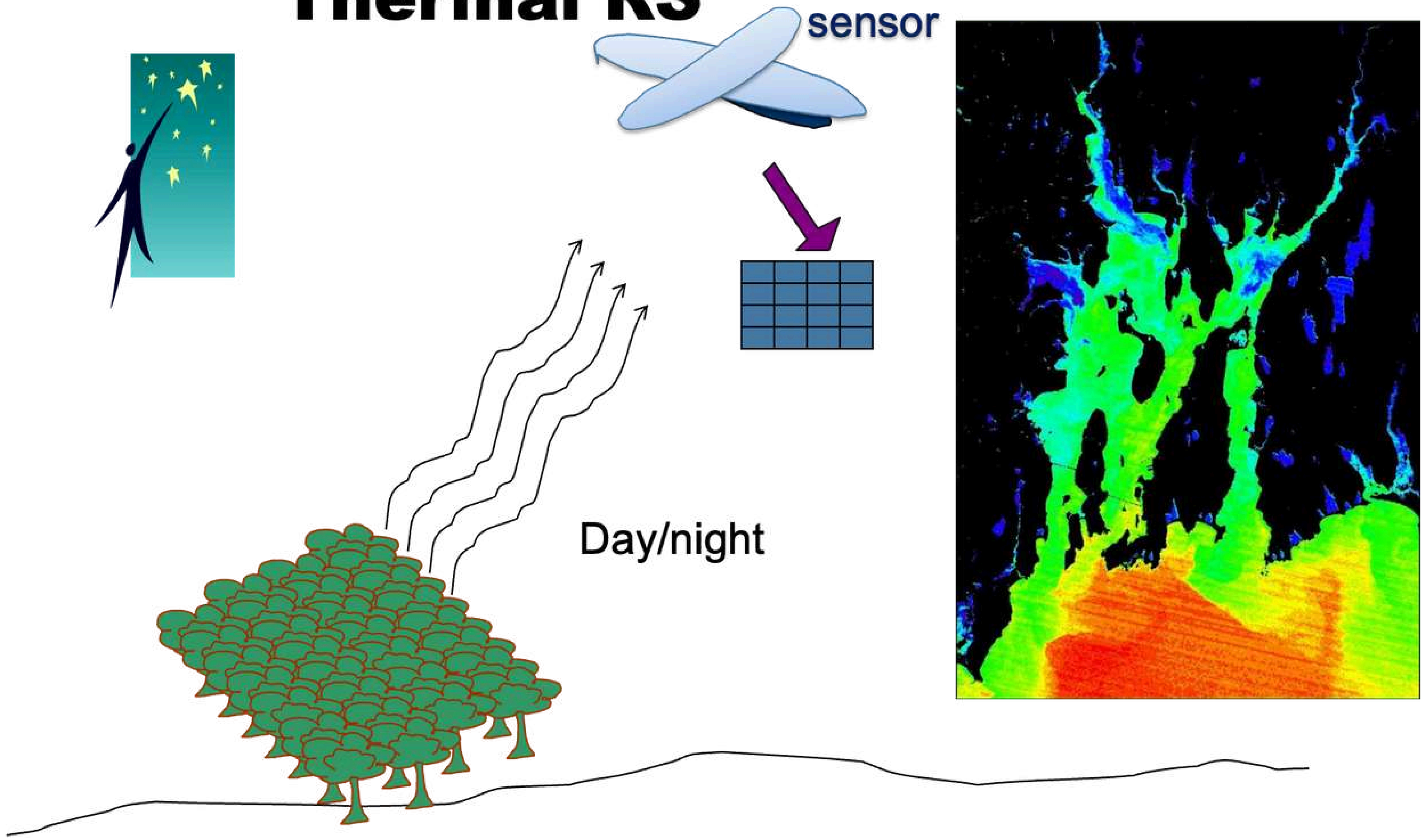
How to sense

Reflected Energy of Sun Light



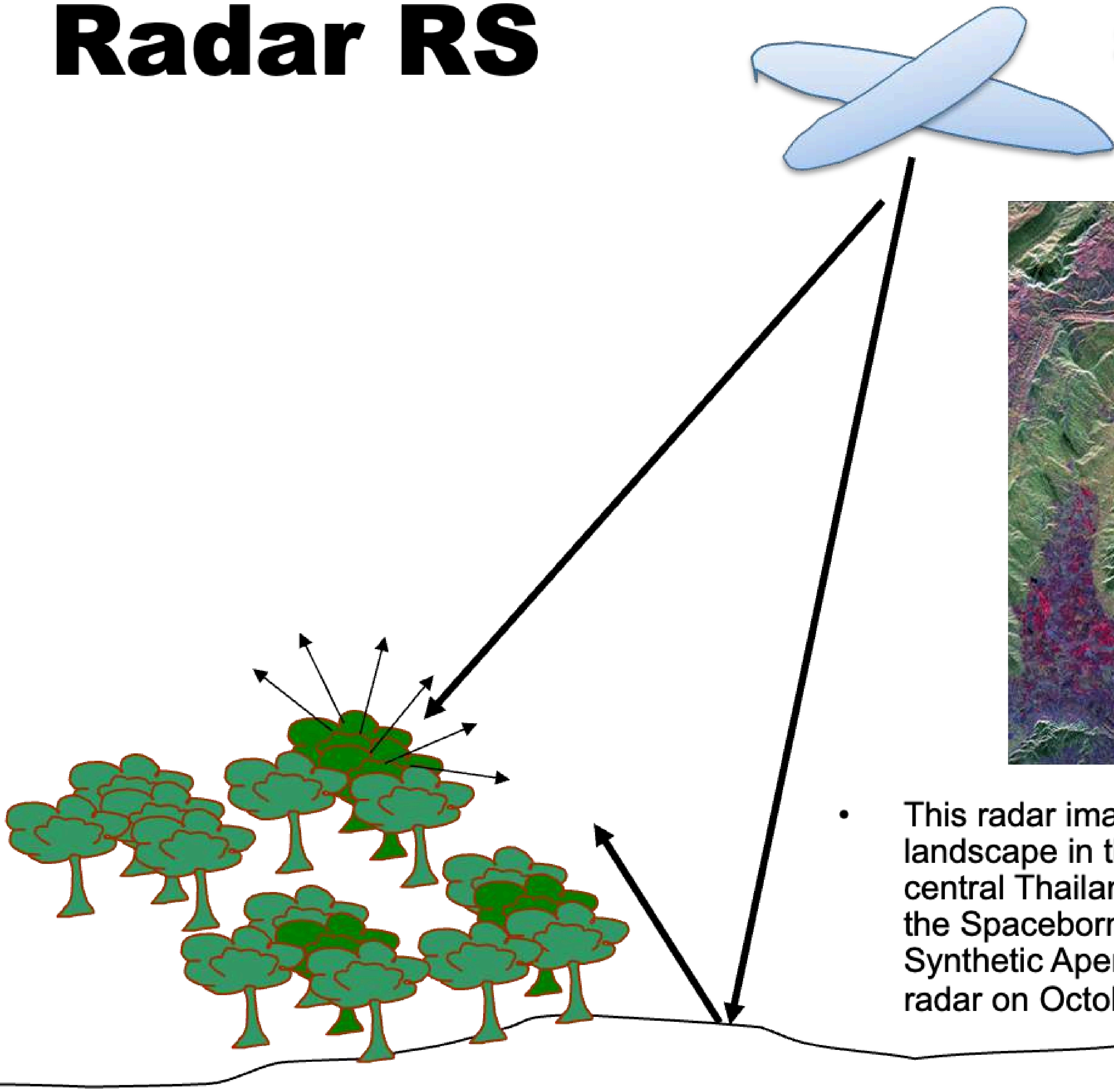
- Landsat 7 ETM+ Image around Lake Titicaca, Bolivia/Peru

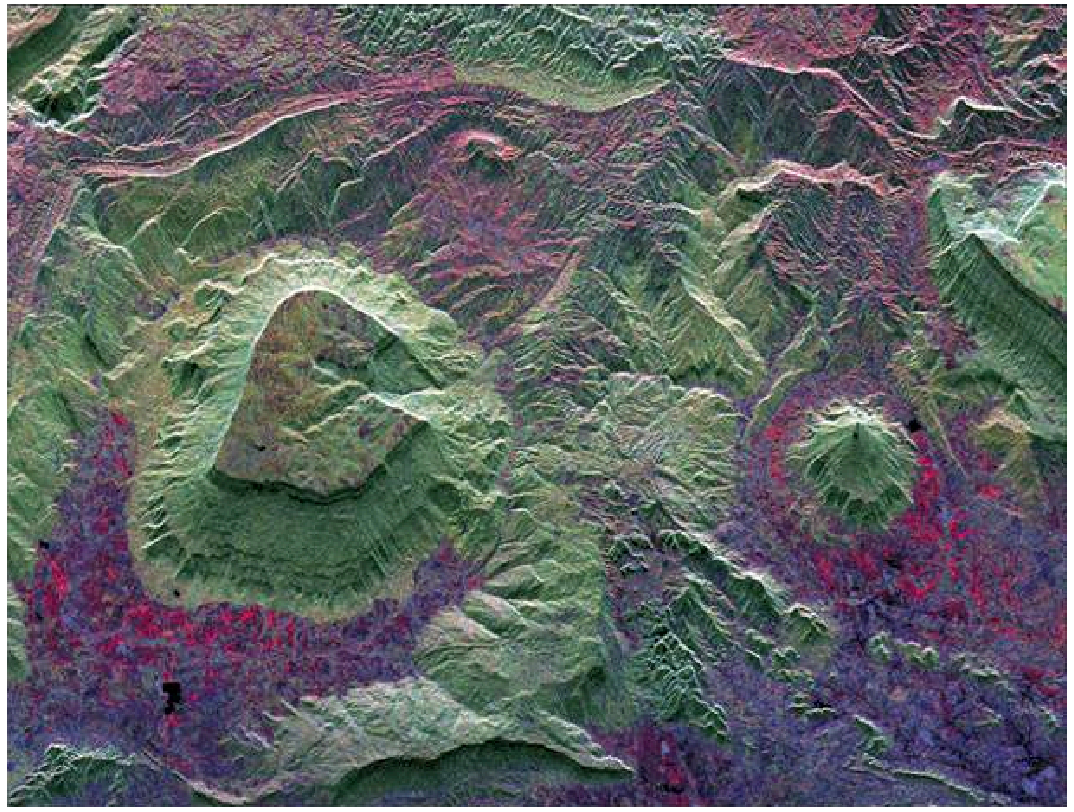
Emitted Energy
Thermal RS



Scattered Energy Radar RS

How to sense





- This radar image shows the dramatic landscape in the Phang Hoei Range of north central Thailand. The image was acquired by the Spaceborne Imaging Radar-C/X-band Synthetic Aperture (SIR-C/X-SAR) imaging radar on October 3, 1994

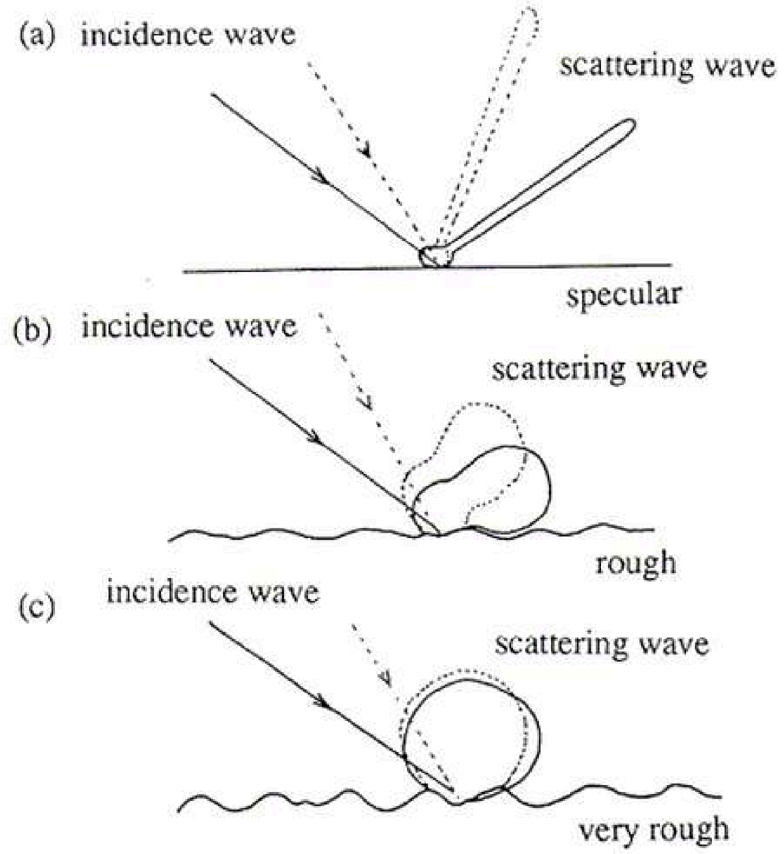
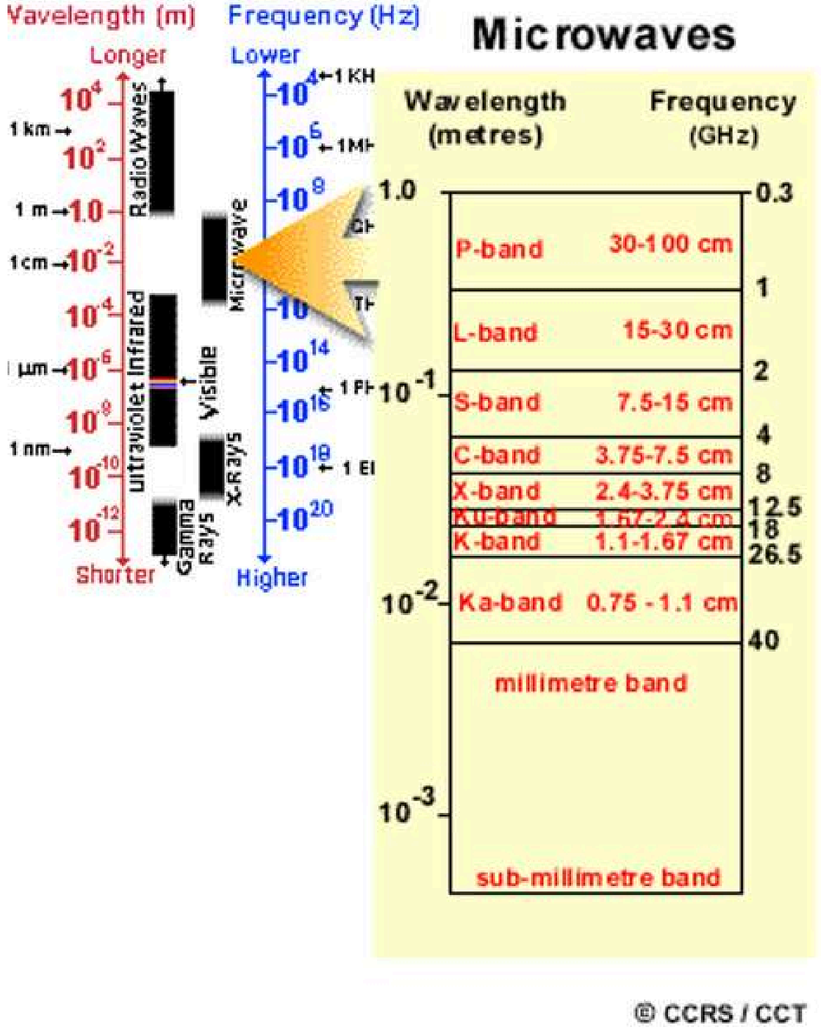
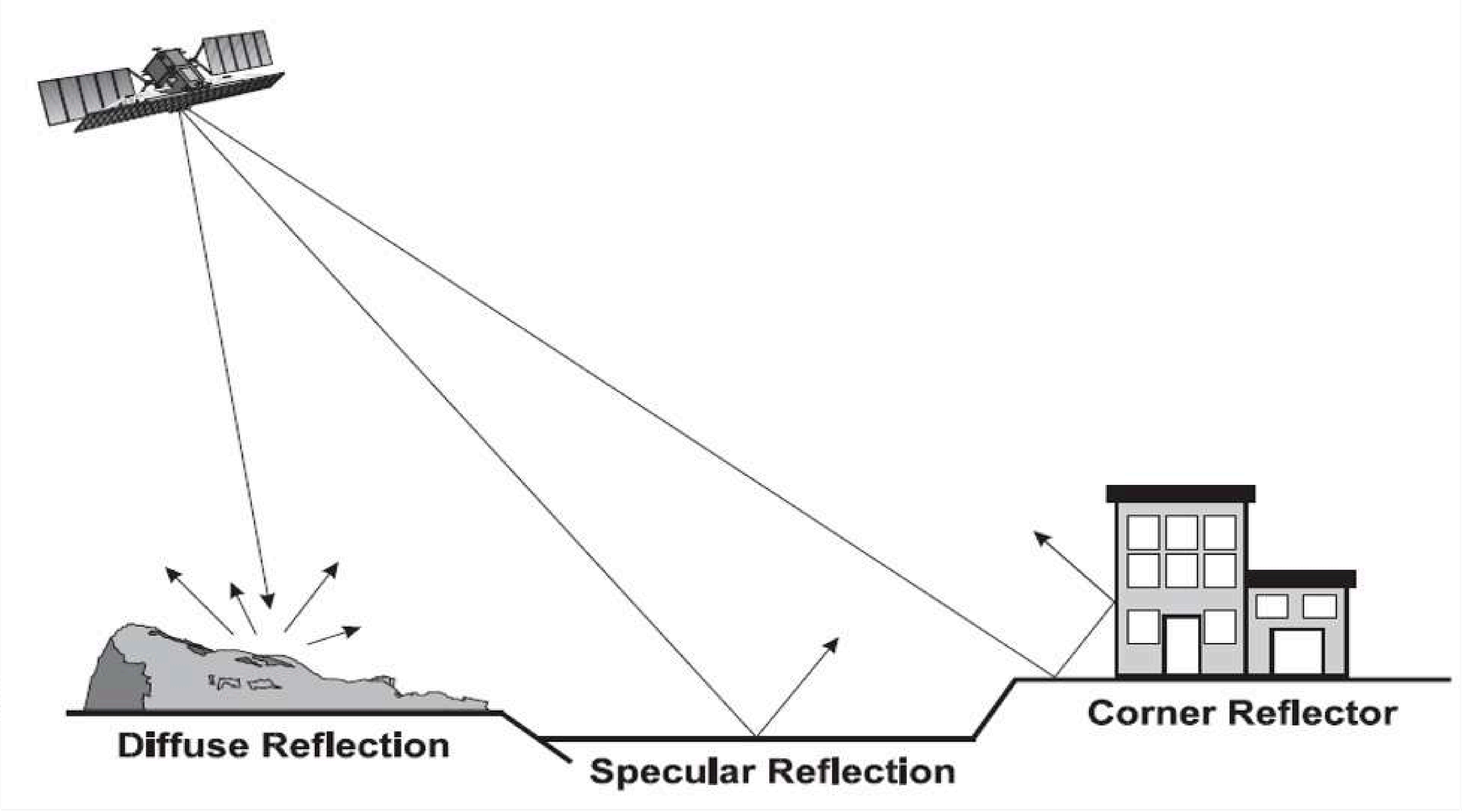


Figure 3.4.1 Surface scattering pattern with different surface roughness



How to sense

- Radar backscatter (detected intensity) is directly related to **topography**, **dielectric properties**, and **surface roughness**

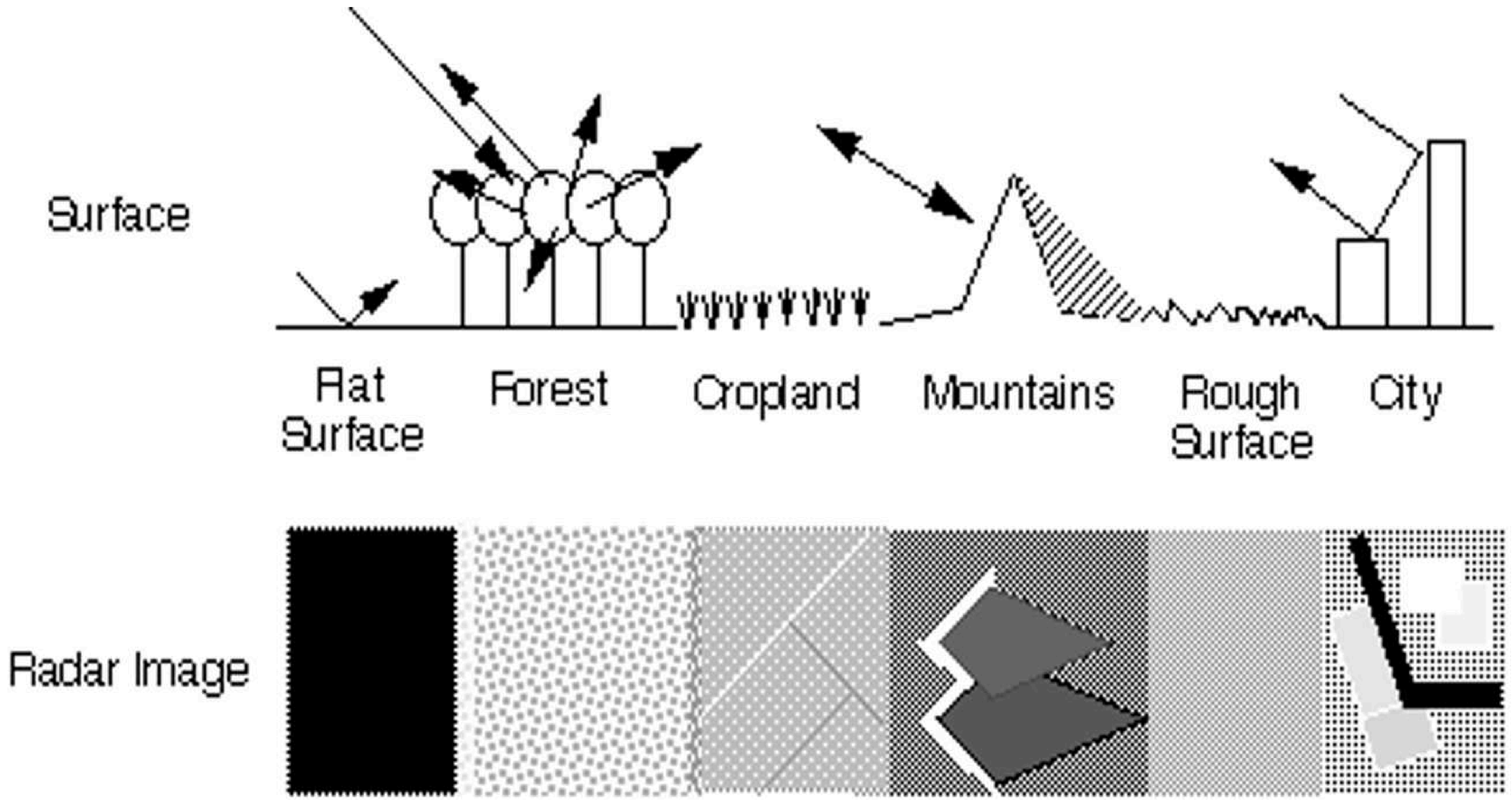


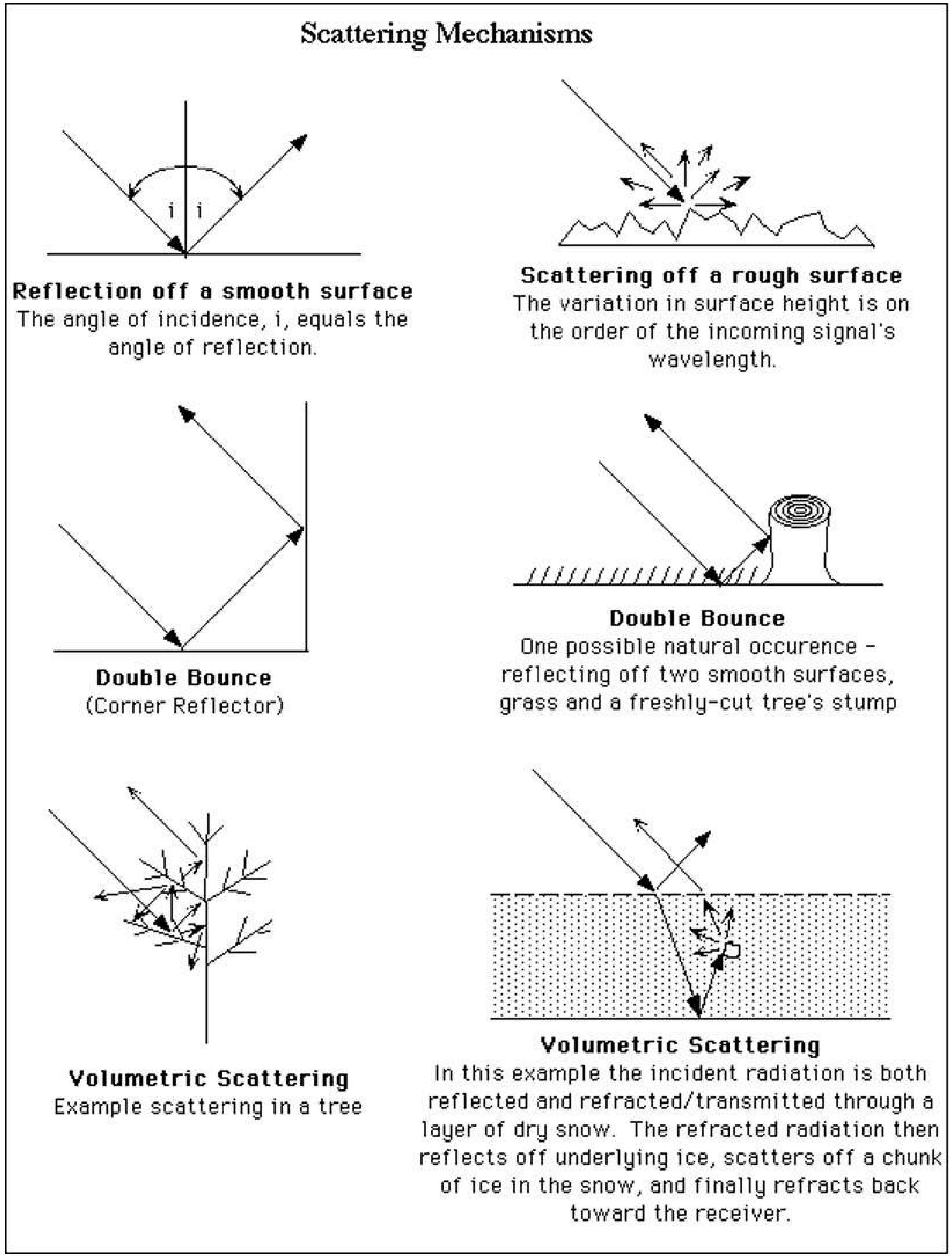
Reflection Type:

Image Appearance: **GREY – speckled** **DARK – smooth** **BRIGHT**

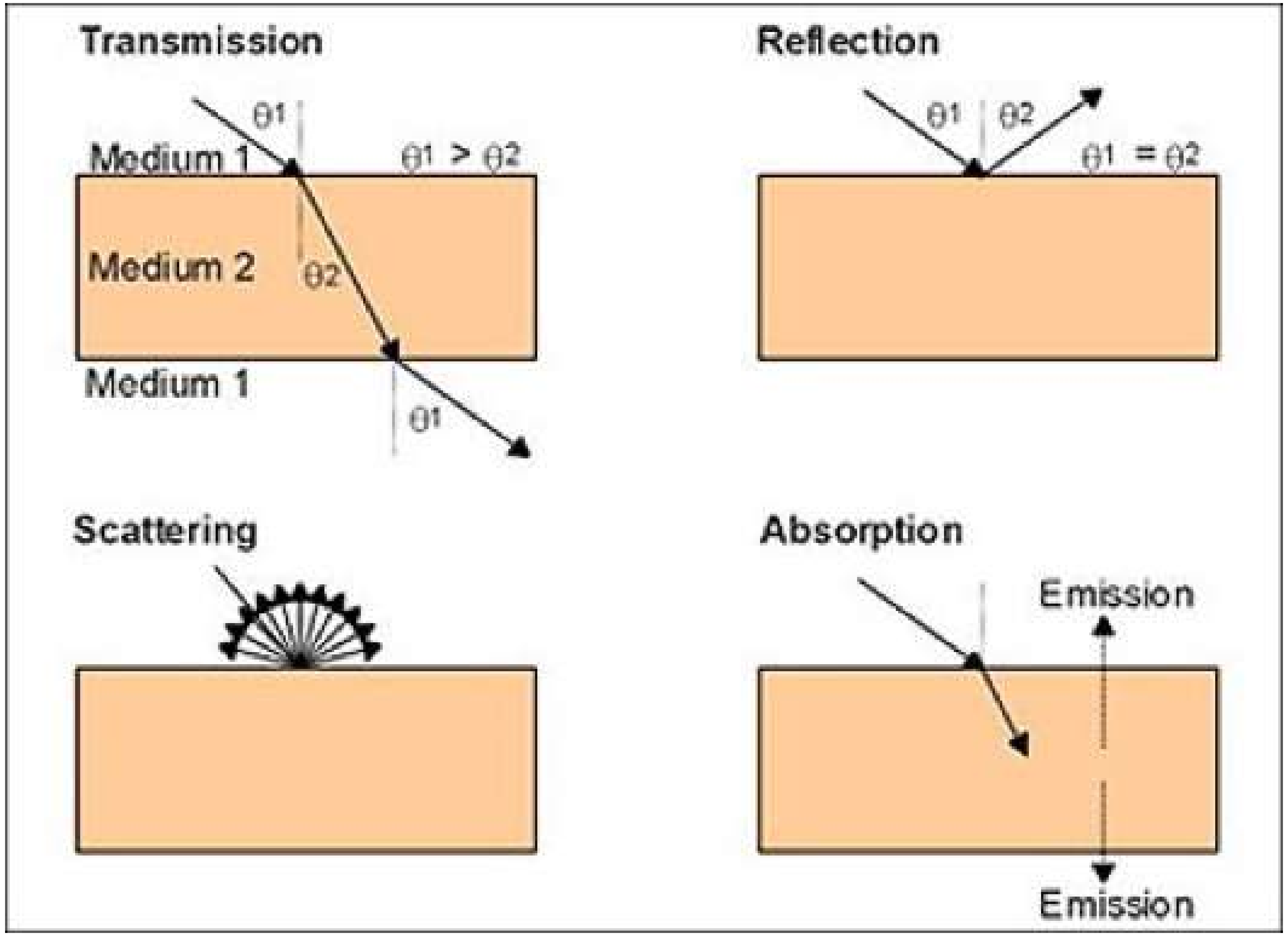
Adapted from: RADARSAT International 1996. *Radarsat Geology Handbook*. Richmond, B.C.

How to sense


Surface
Radar Image
Flat Surface
Forest
Cropland
Mountains
Rough Surface
City


Scattering Mechanisms
i
i
Reflection off a smooth surface
The angle of incidence, i, equals the angle of reflection.
Scattering off a rough surface
The variation in surface height is on the order of the incoming signal's wavelength.
Double Bounce
(Corner Reflector)
Double Bounce
One possible natural occurence - reflecting off two smooth surfaces, grass and a freshly-cut tree's stump
Volumetric Scattering
Example scattering in a tree
Volumetric Scattering
In this example the incident radiation is both reflected and refracted/transmitted through a layer of dry snow. The refracted radiation then reflects off underlying ice, scatters off a chunk of ice in the snow, and finally refracts back toward the receiver.

Interactions with Surfaces (Summary)



- There are three (3) forms of interaction that can take place when energy strikes, or is **incident (I)** upon the surface. These are: **reflection (R)**; **transmission (T)**; and **absorption (A).**

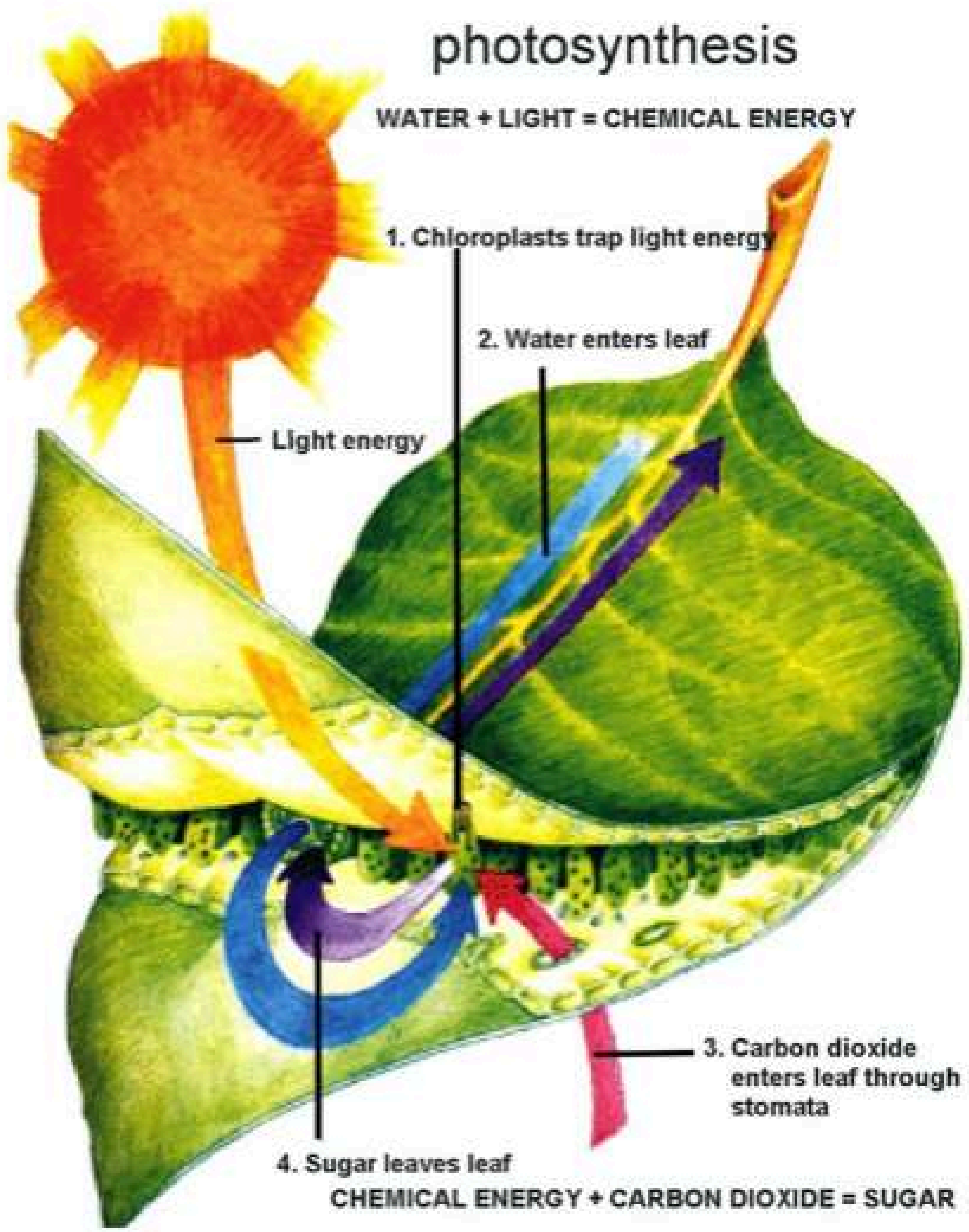
- Incident radiation passes through an object without significant attenuation, or may be selectively transmitted.

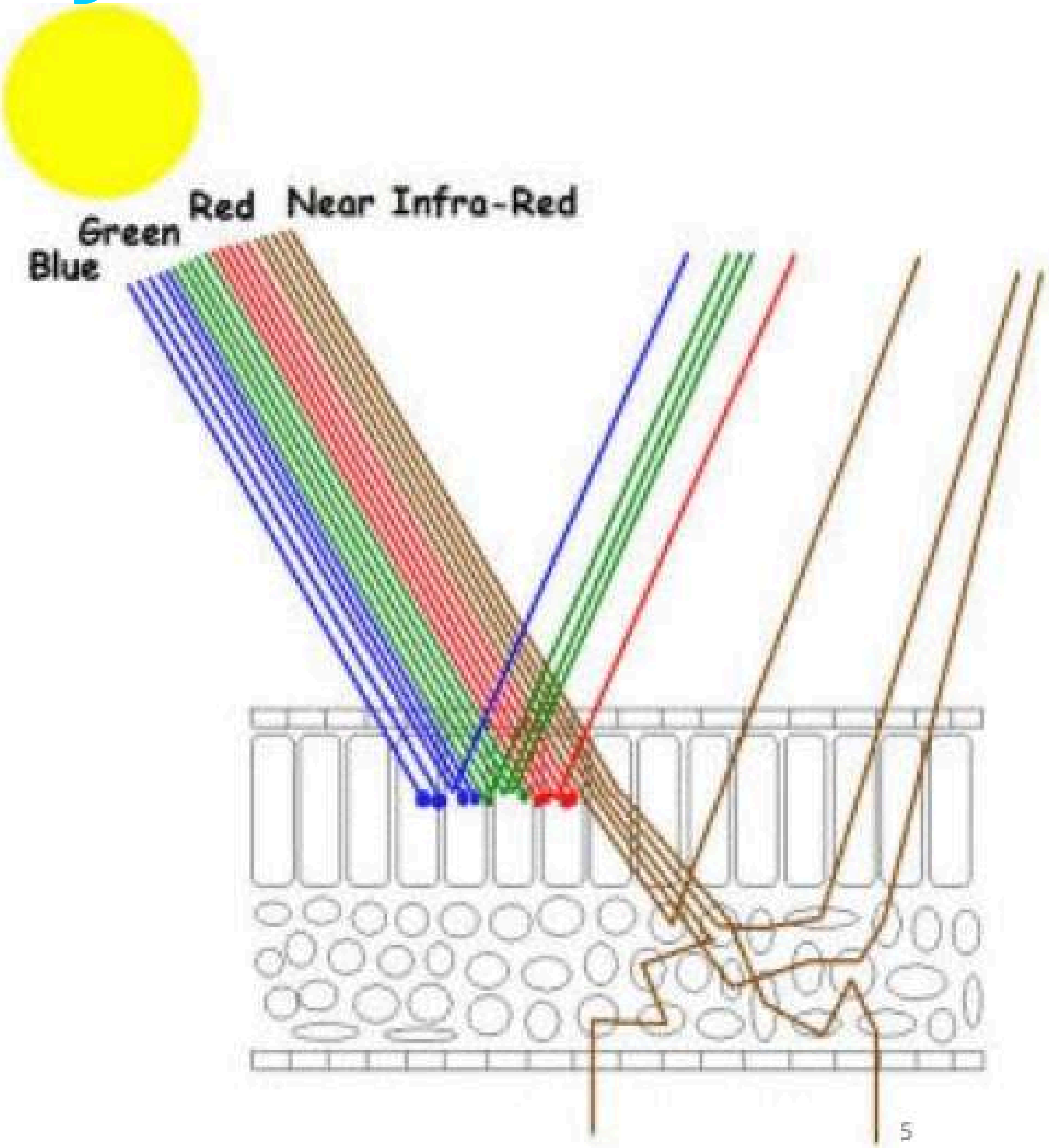
- Scattering is the redirection of electromagnetic energy by the target.

Type of Remote Sensing

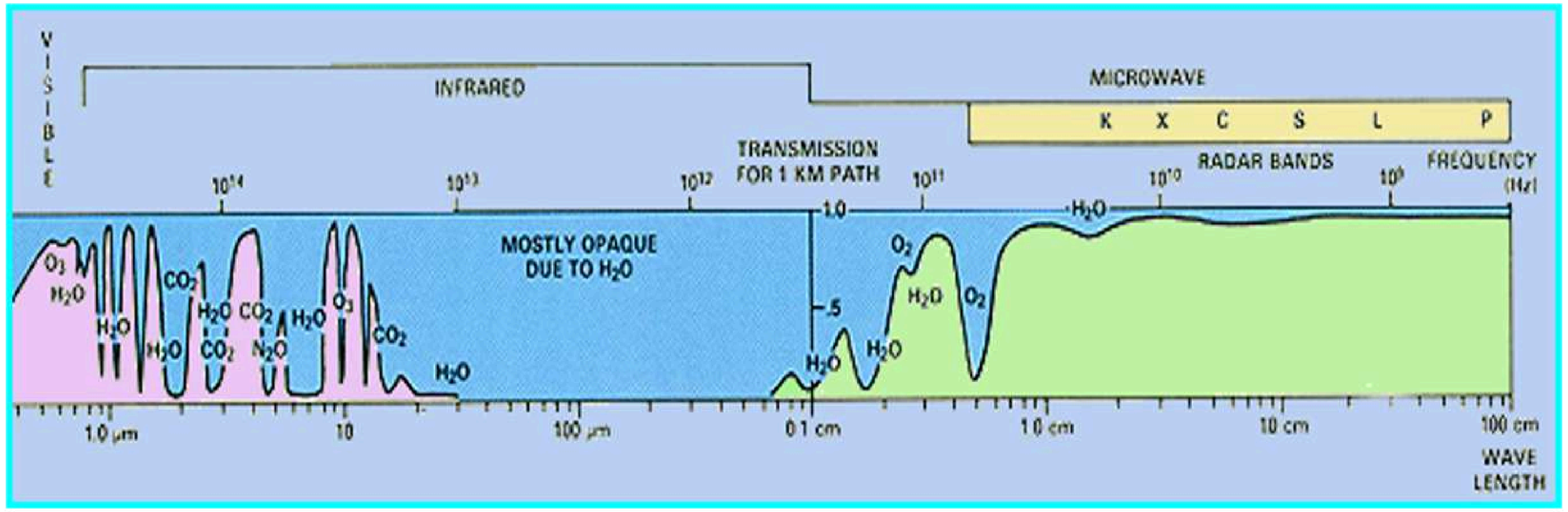
- **Optical RS** : Energy source is the sun. It operate in the visible, near infrared, middle infrared and short wave infrared portion of the electromagnetic spectrum.
- **Thermal RS** : The energy emitted from the earth features in the wavelength range of 3 μm to 5 μm and 8 μm to 14 μm.
- **Microwave RS** :
  - Passive type : Microwave radiation emitted from an object
  - Active type : Having there own sources of energy. The back scattering coefficient is detected

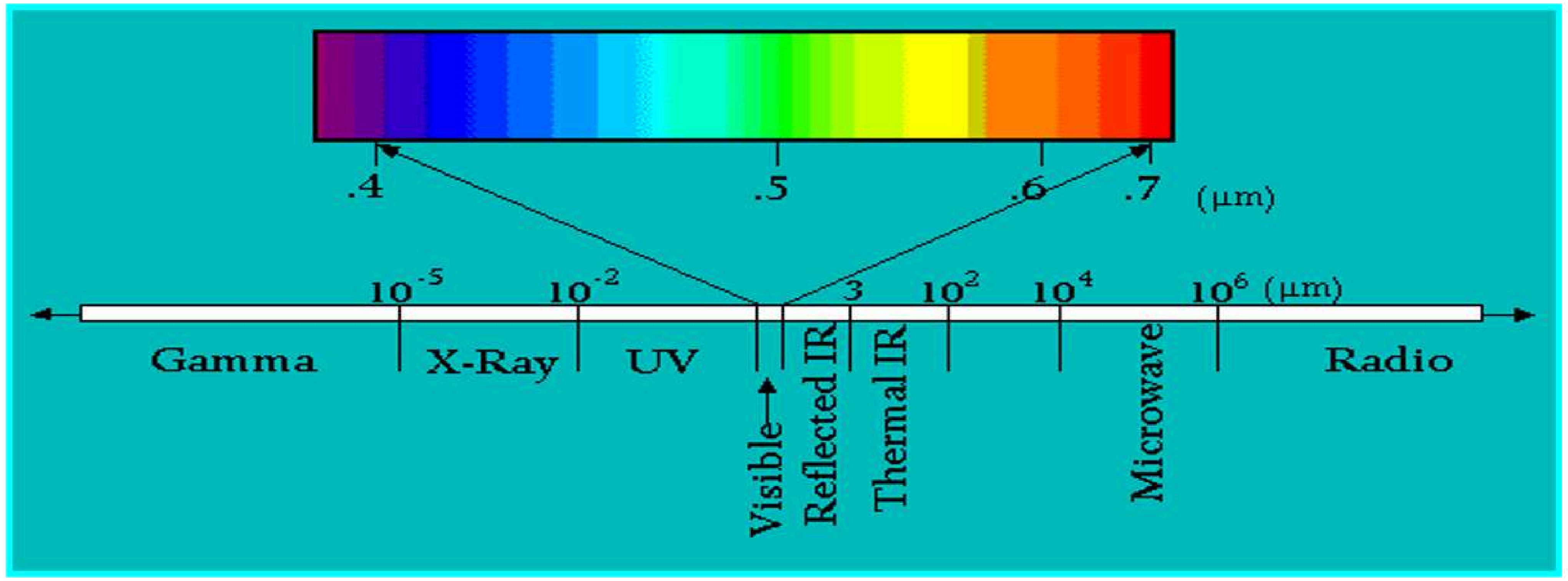
Object Signature

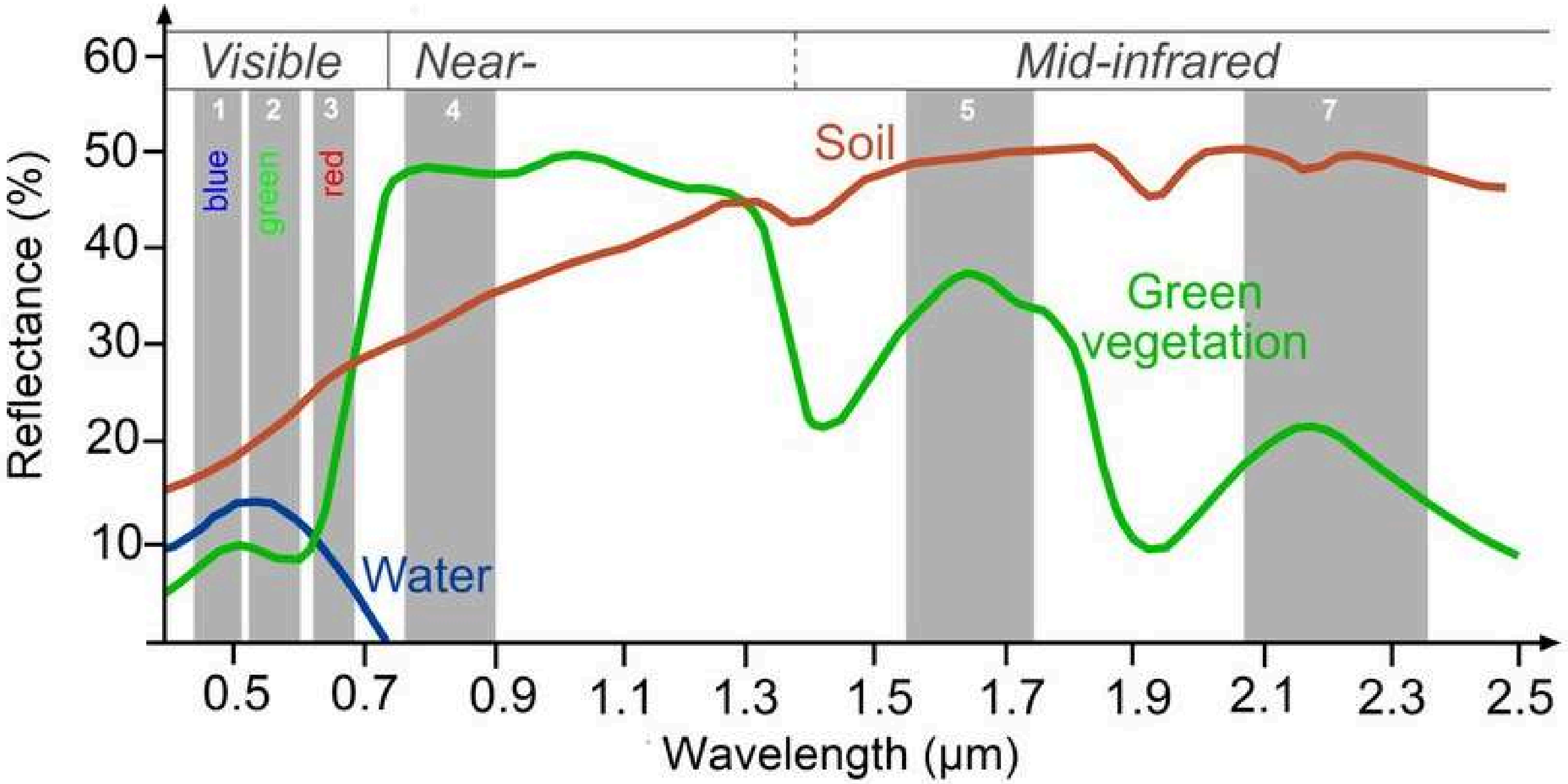


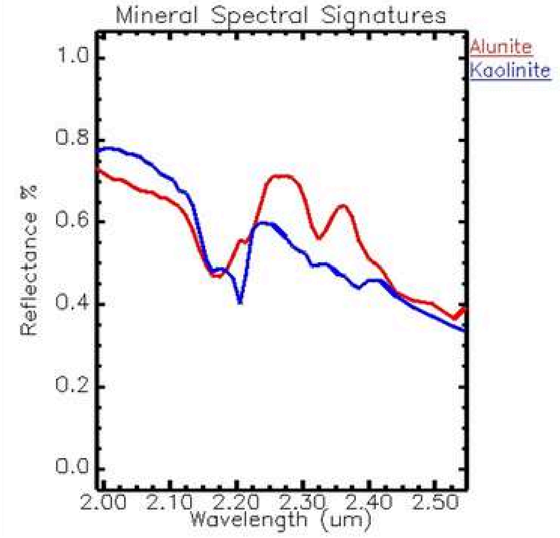


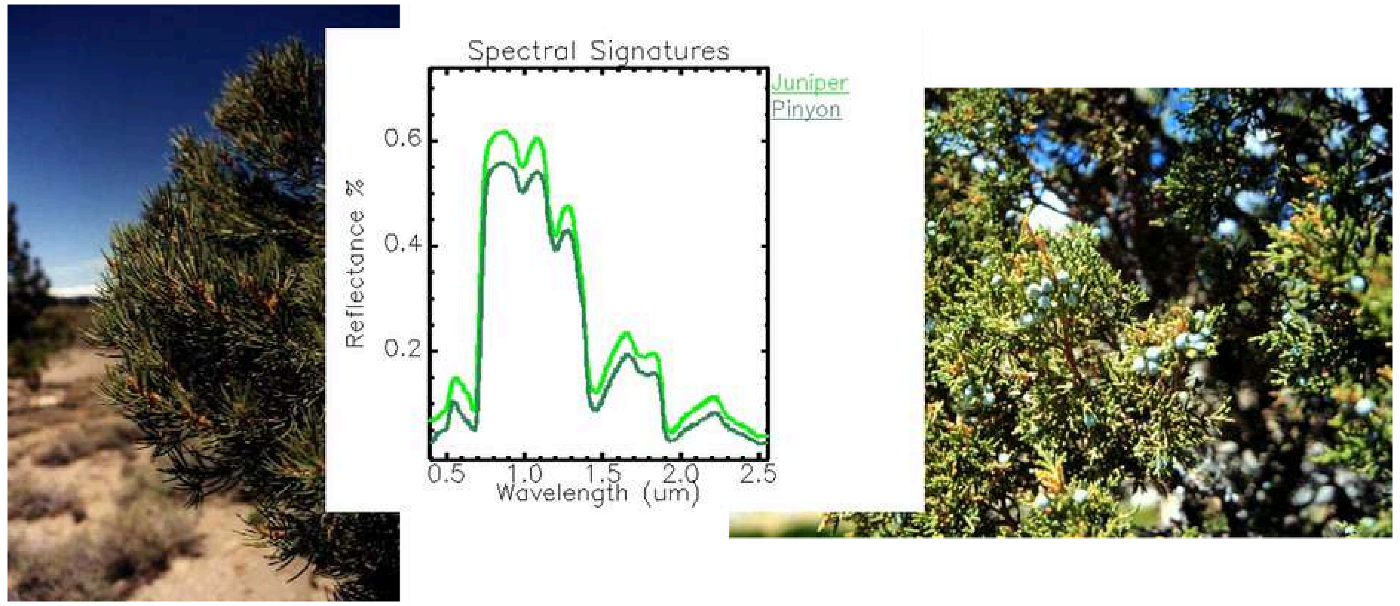
Object Signature





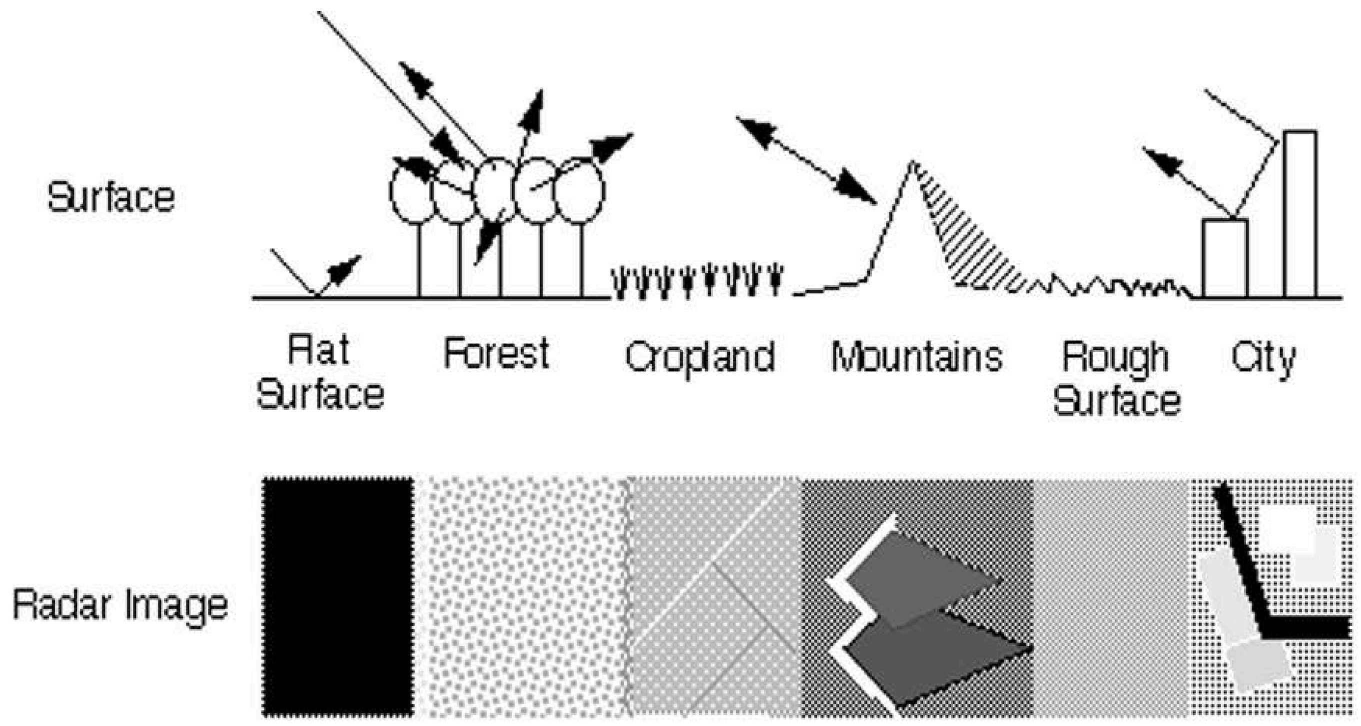




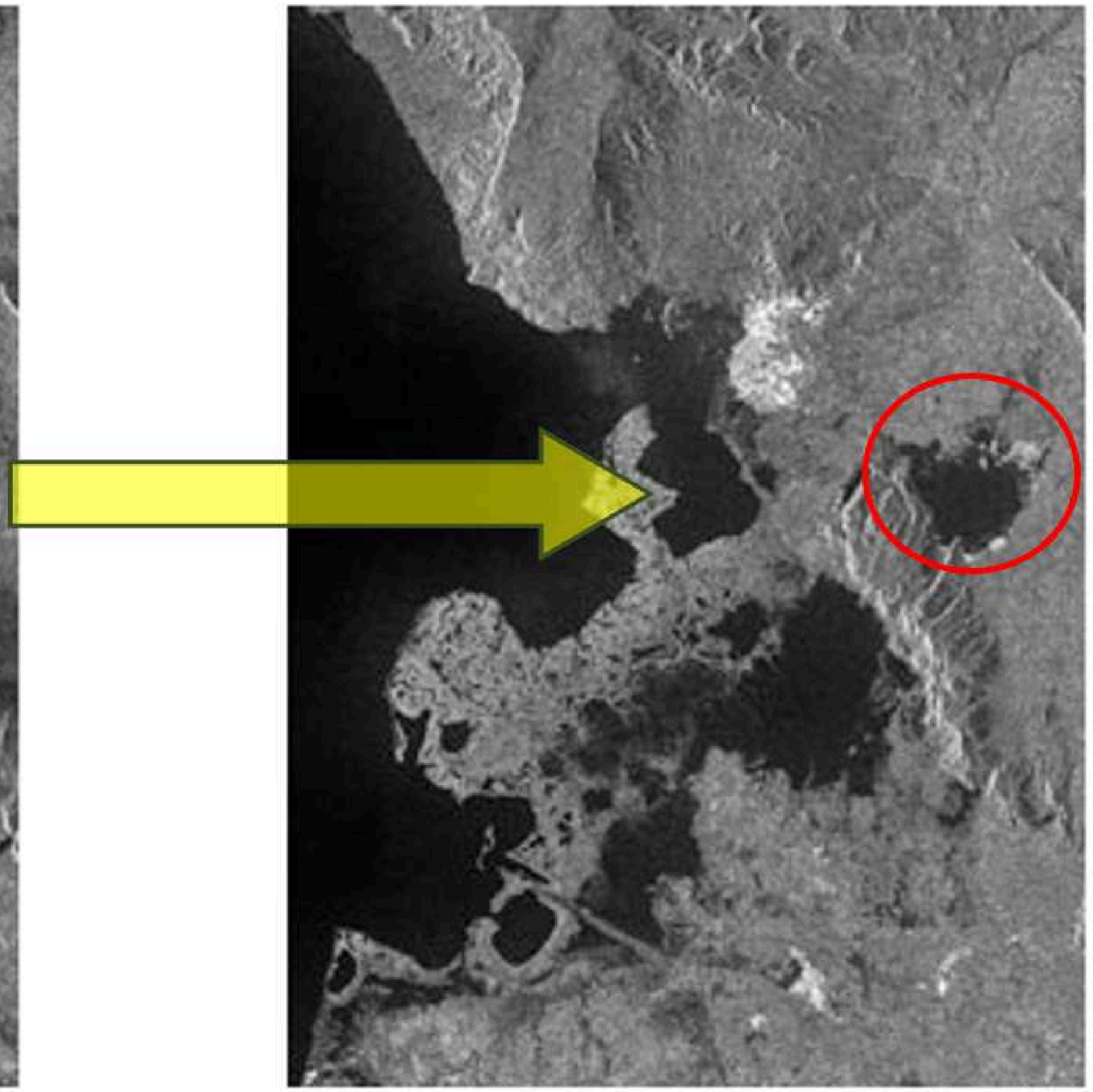


RADAR Back Scatter

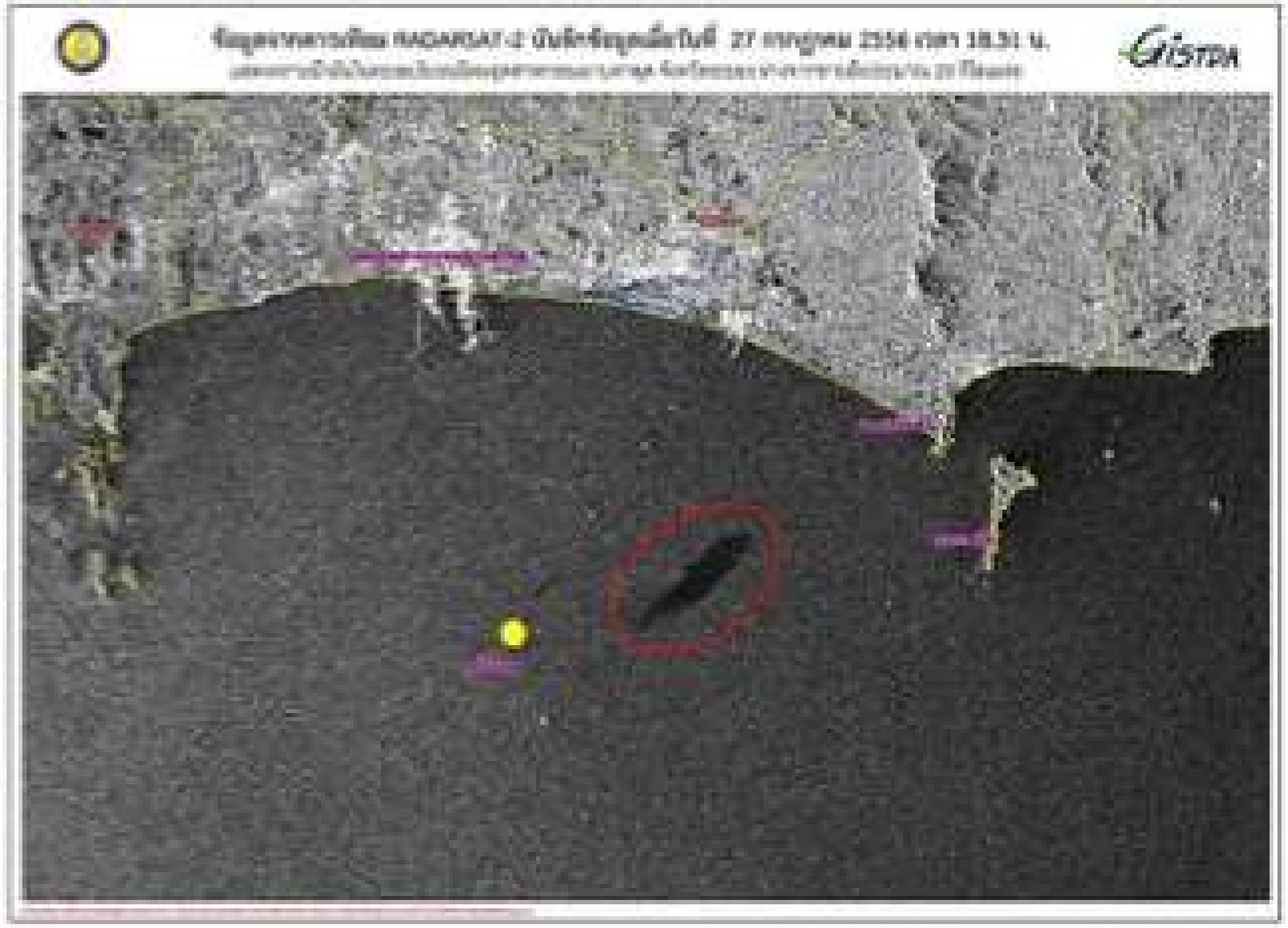
- Surface Back Scatter
  - Roughness
- Volume Scatter and Target's Structure:
  - Volume, Structure
- Dielectric Property:
  - Moisture** :
    - Variable moisture levels are represented as tonal variations in the image.
- Polarization
- Distance

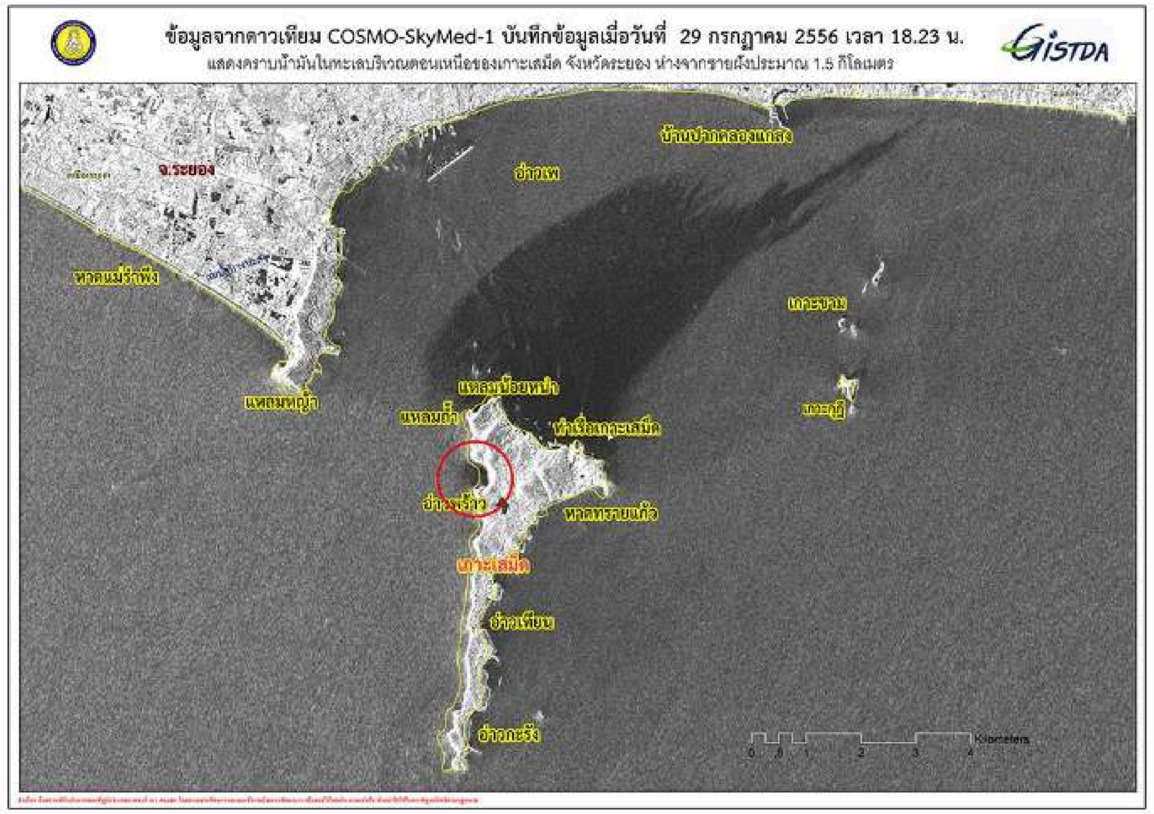


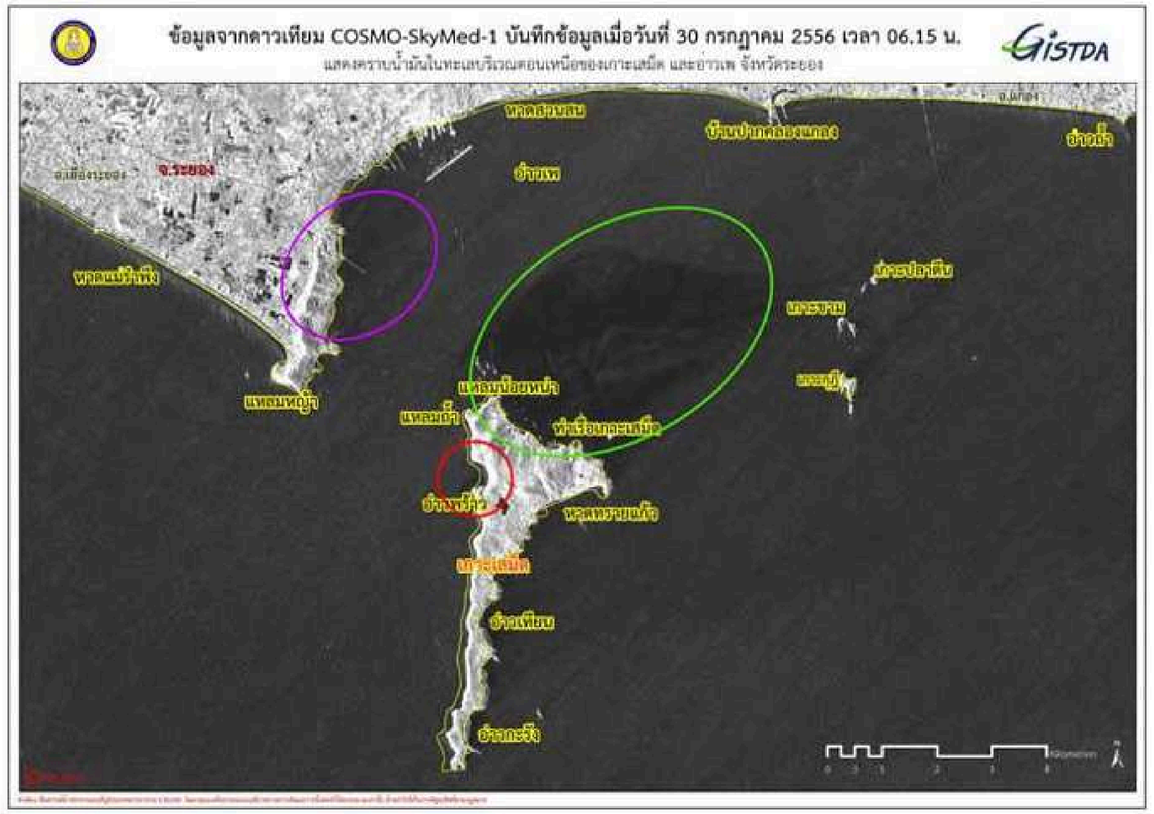
June 13, 1999



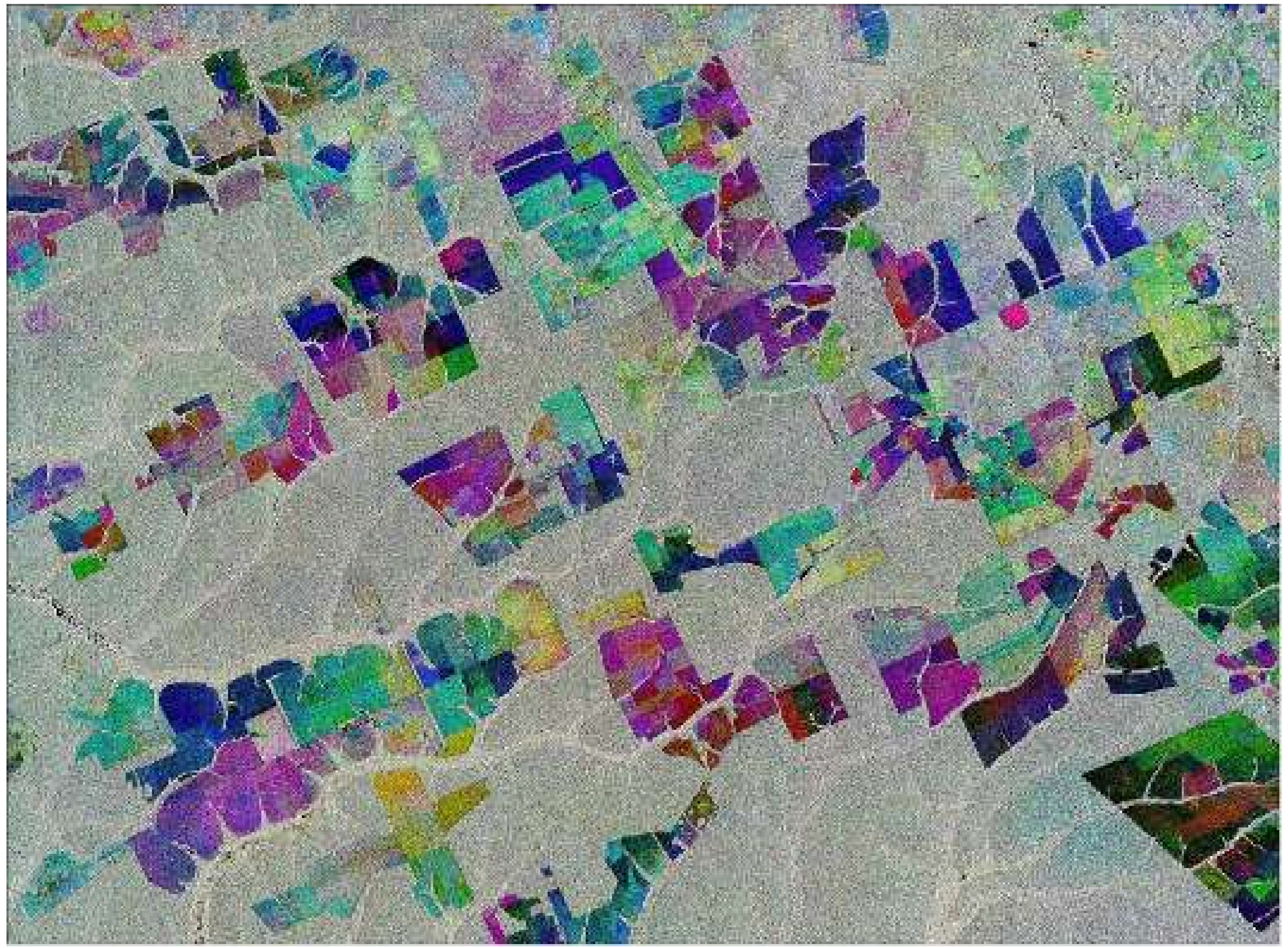
September 24, 2004



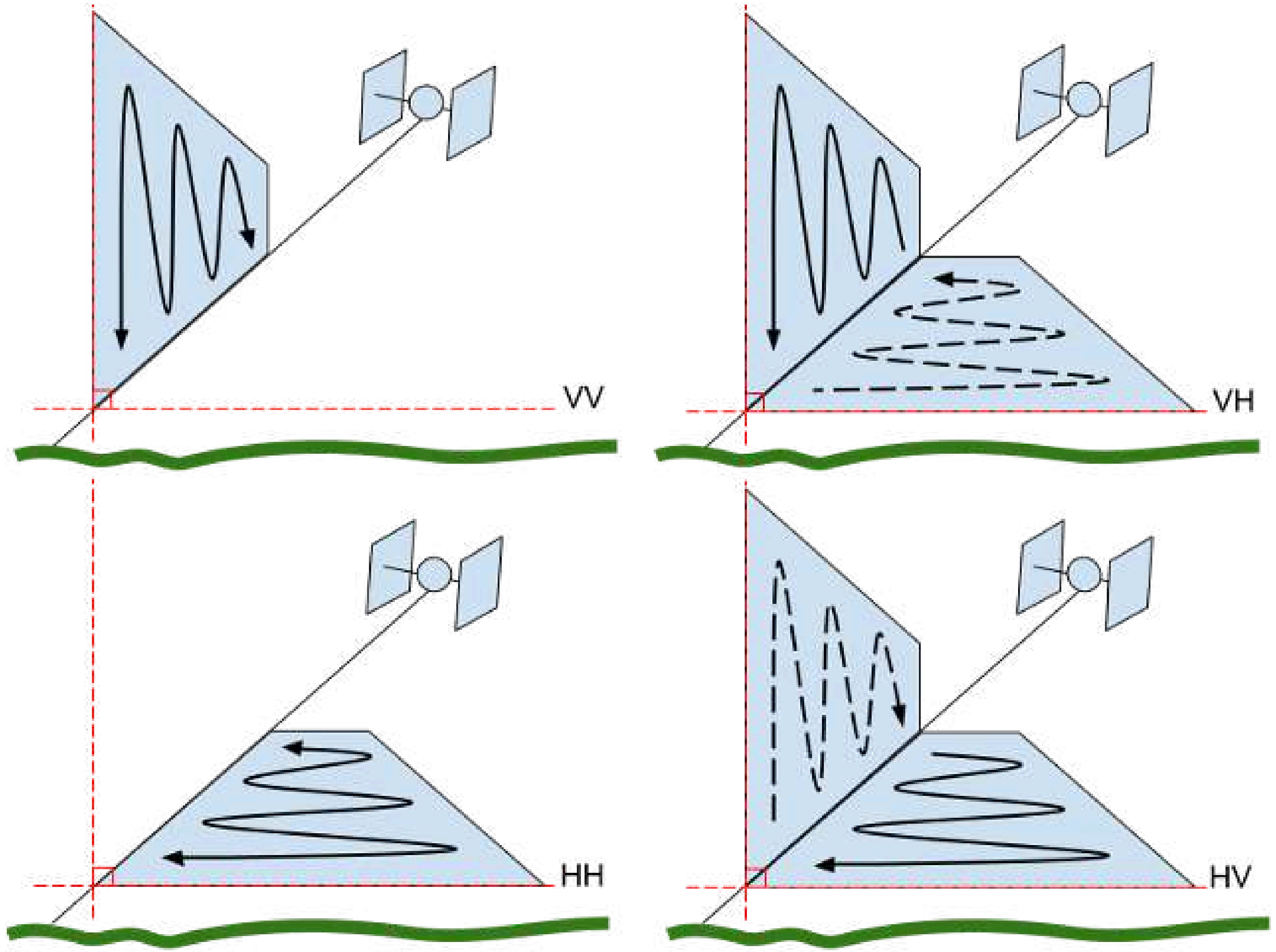




SAR Polarization



Polarimetry is an emerging field of SAR processing which is used in a number of applications such as measuring vegetation properties and changes of vegetation over time. Additional applications include oceanography, geology, and disaster response.



Data look a like

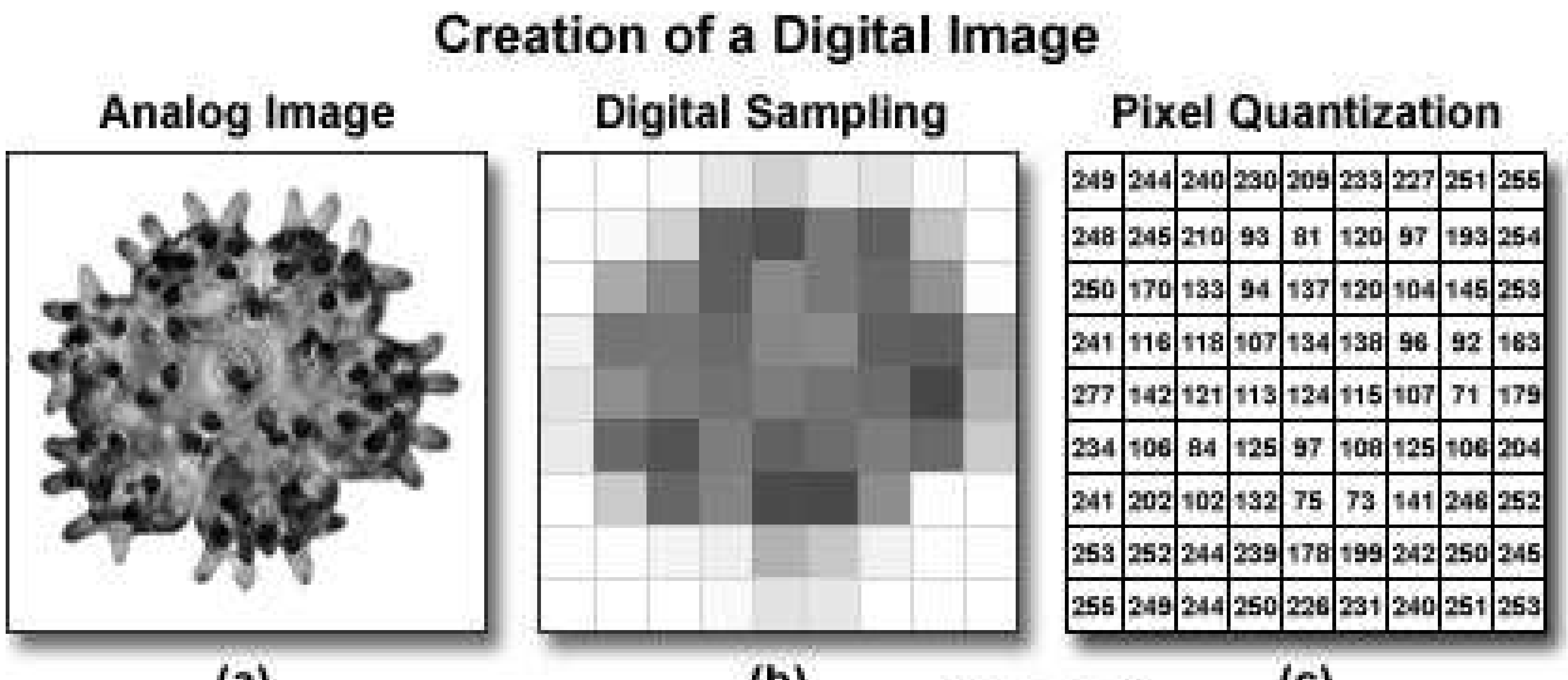


Figure 1

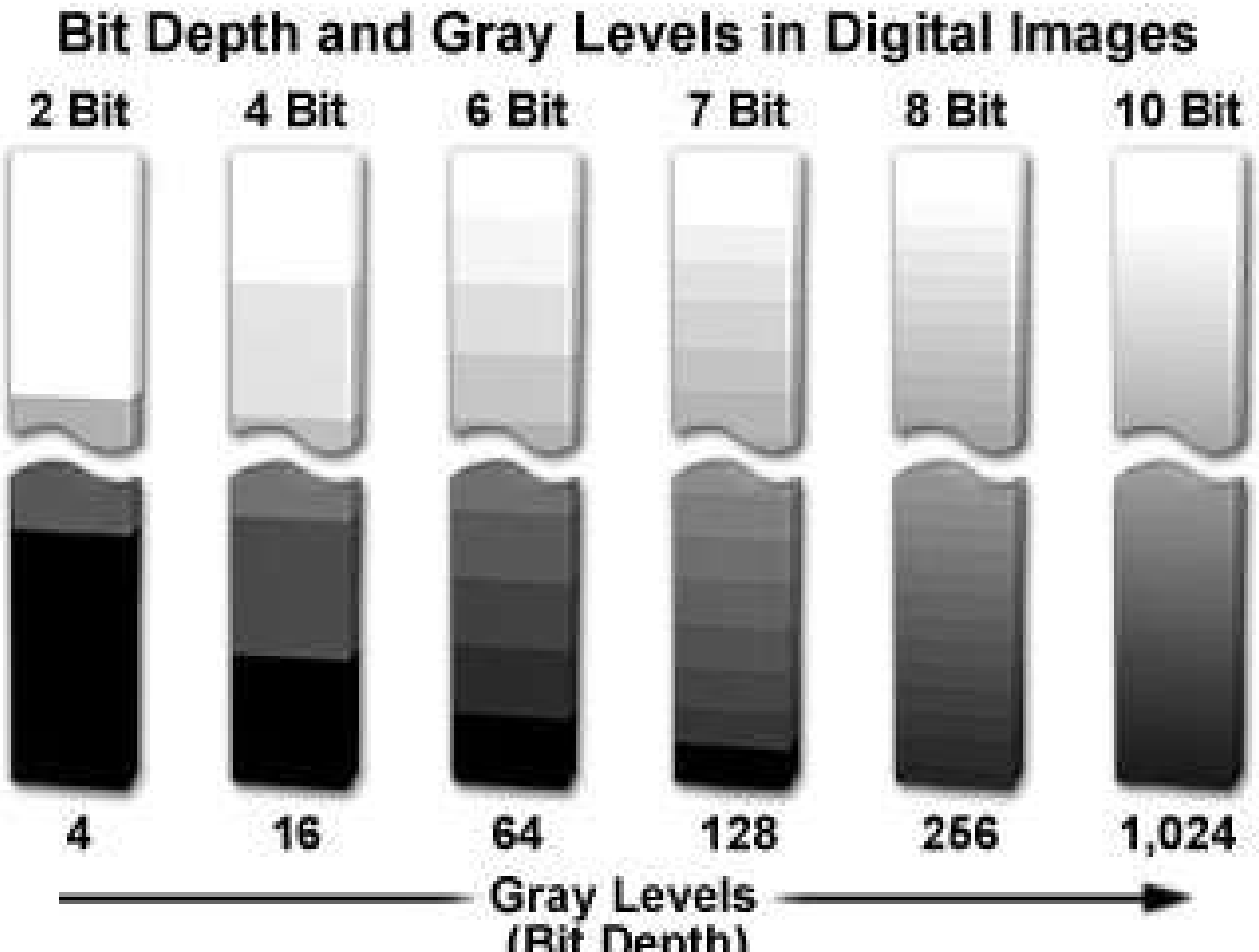
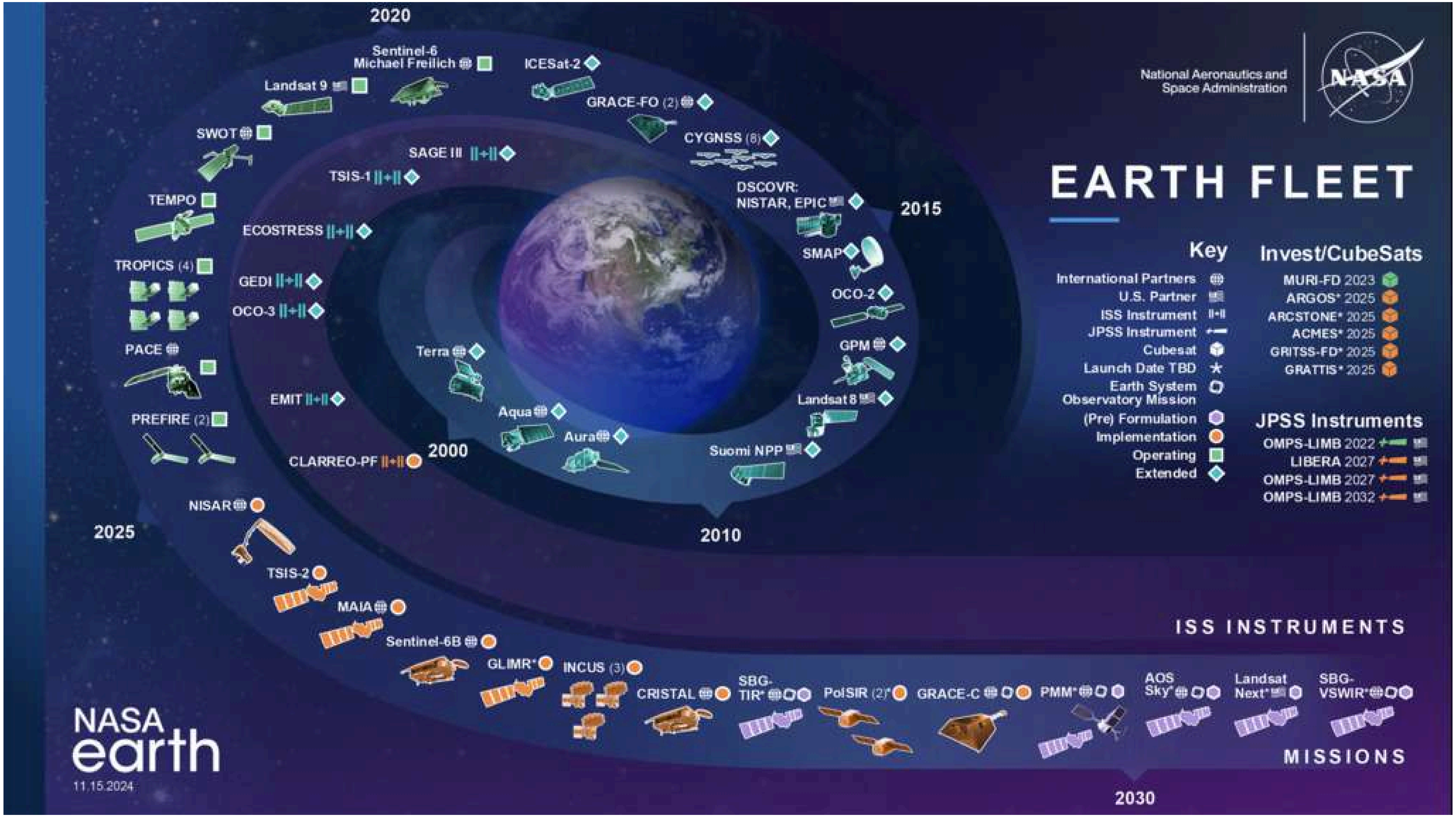


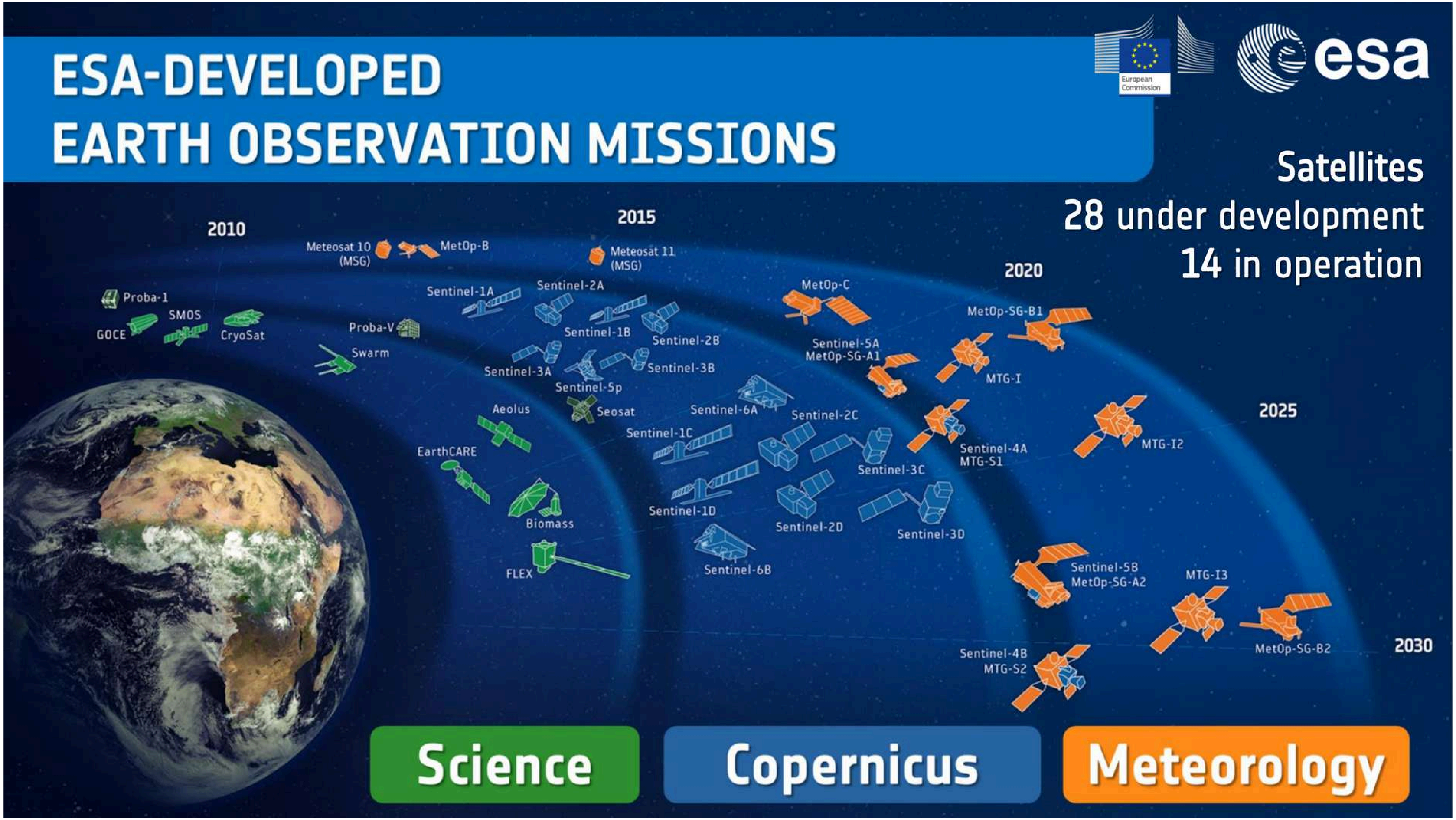
Figure 5

- A natural image captured with a camera, telescope, microscope, or other type of optical instrument displays a continuously varying array of shades and color tones.
- After an object has been imaged and sampled, each resolvable unit is represented either by a digital integer.

Remote Sensing Platform

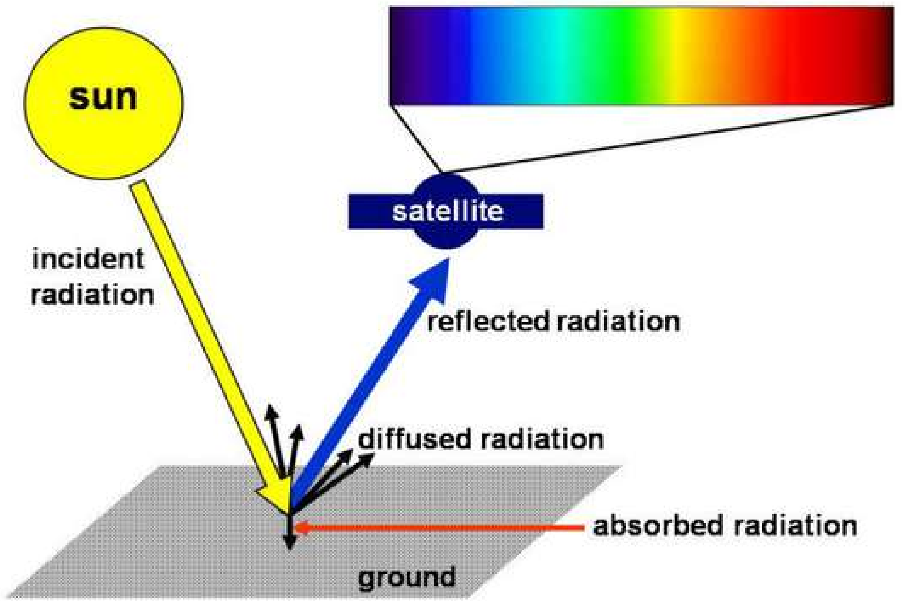


Remote Sensing Platform

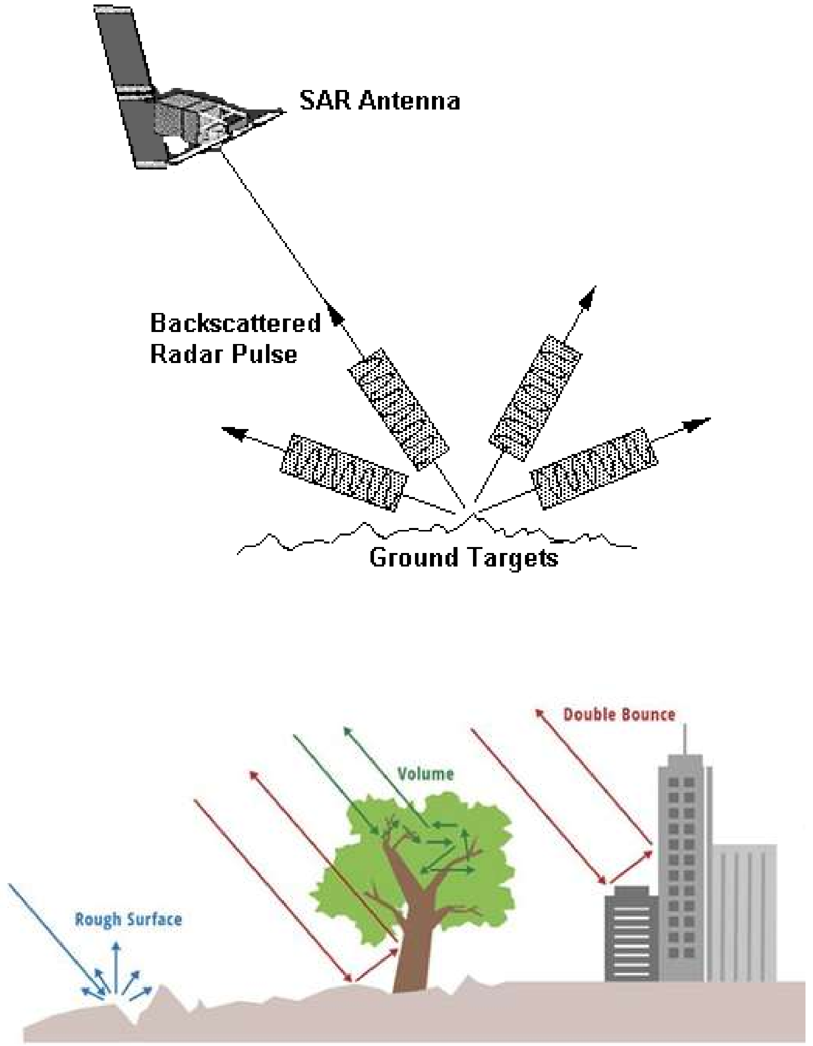


Optical (Passive) vs SAR (Active)

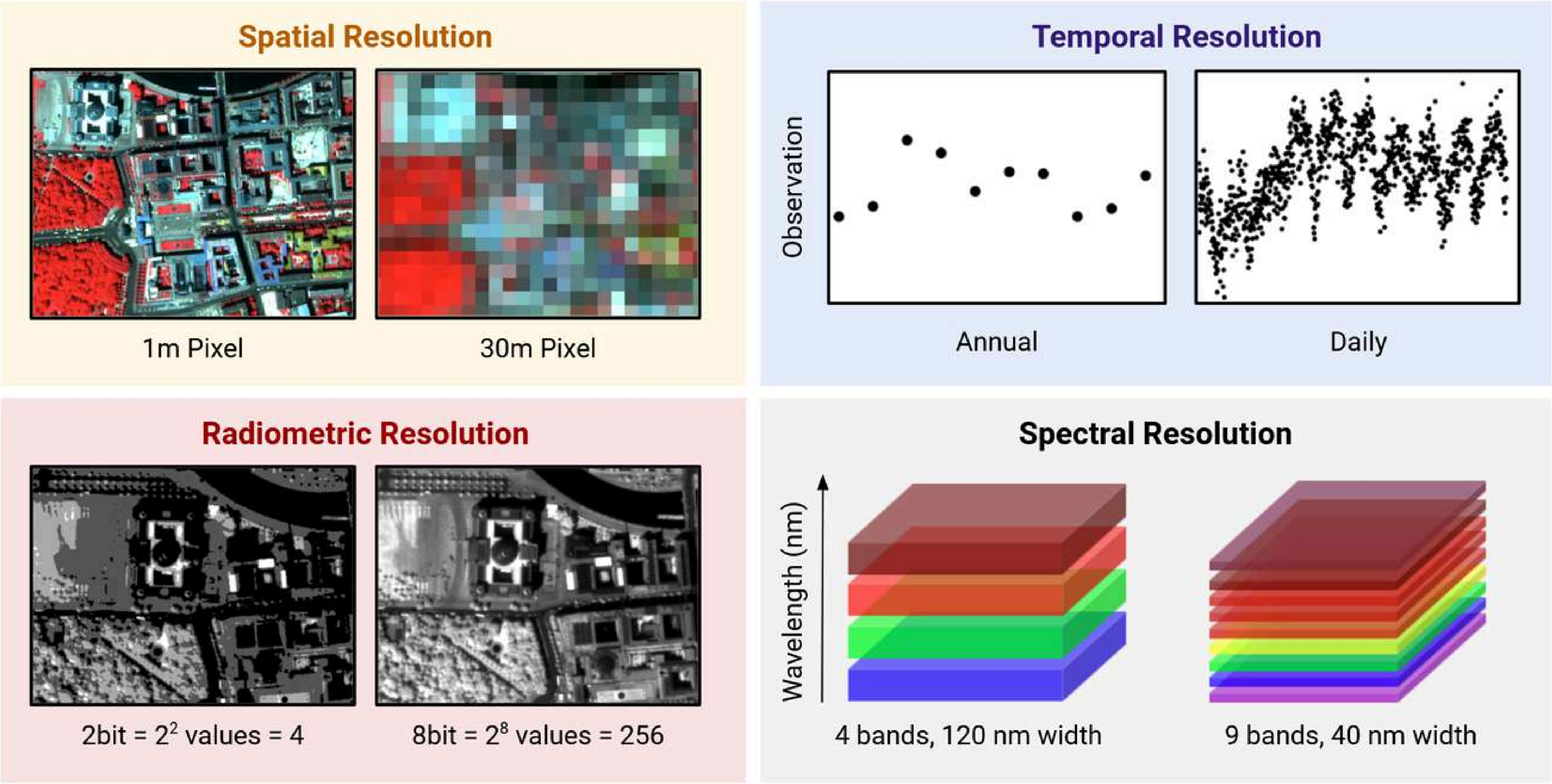
Multispectral Sensing



Physical Surface Structure Sensing



Satellite Platform Selection



Meter ?

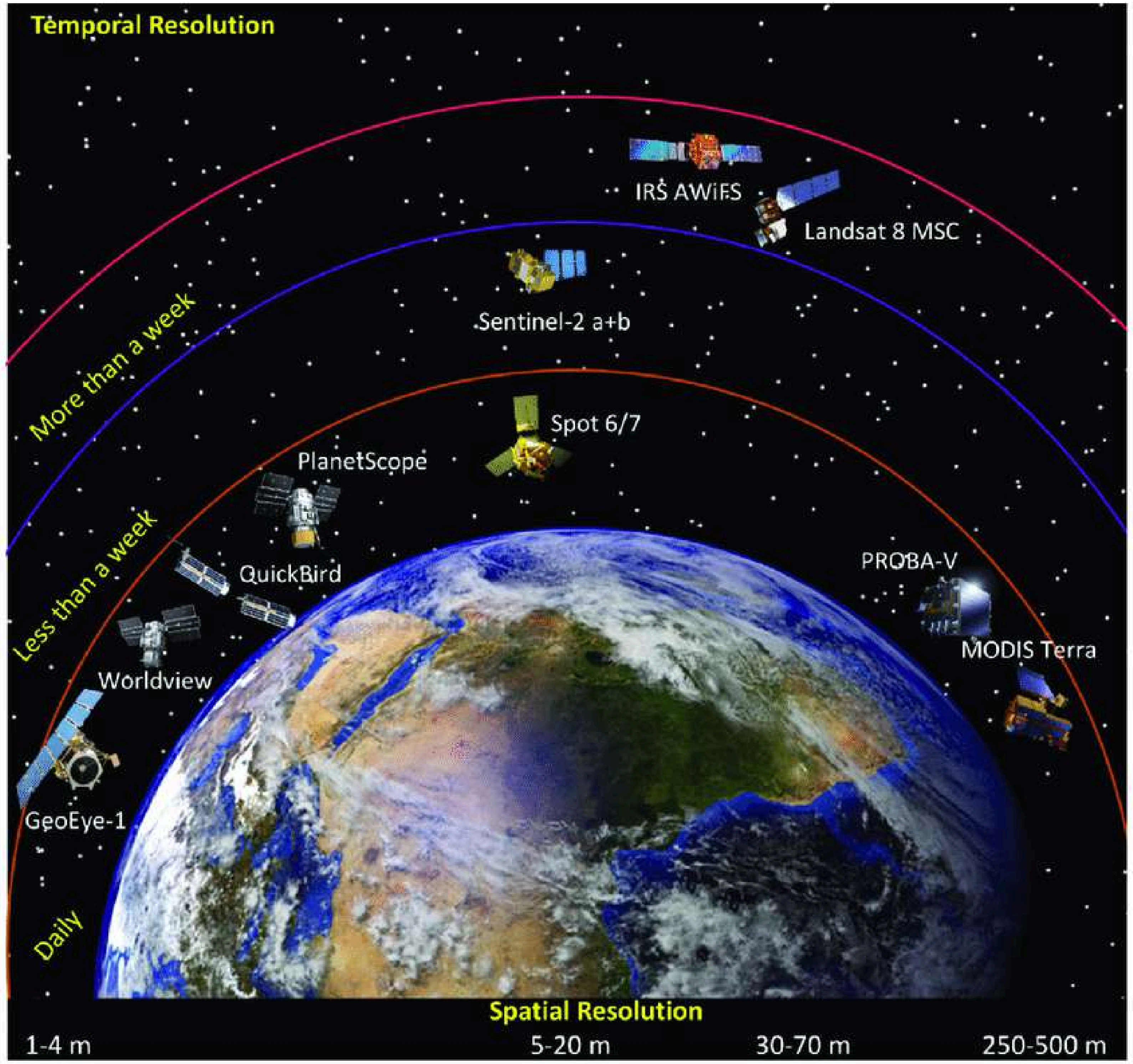
Color ?

Often ?

Sensitive ?

Observation Satellite Constellations

Satellite system	Mission start/ completion	Spectral characteristics [in µm]		Orbit height	Swath width	Resolution	Repeat Cycle	Revisit Frequency
FormoSat-2 (ROCSat-2) (NSPO, Taiwan)	2004 - *	0.45 – 0.90 0.45 – 0.52 0.52 – 0.60	0.63 – 0.69 0.76 – 0.90	888 km	24 km	PAN: 2m Others: 8m	1 day	1 day
Landsat 7 (USGS, USA)	1999 - *	0.52 – 0.90 0.45 – 0.52 0.53 – 0.61 0.63 – 0.69	0.78 – 0.90 1.55 – 1.75 2.09 – 2.35	705 km	185 km	PAN: 15m Others: 30m	16 days	16 days
Landsat 8 (USGS, USA)	2013 - *	0.50 – 0.68 0.433 – 0.453 0.45 – 0.515 0.525 – 0.60 0.63 – 0.68	0. 845 – 0.885 1.36 – 1.39 1.56 – 1.66 2.10 – 2.30	705 km	185 km	PAN: 15m Others: 30m	16 days	16 days
RapidEye (RapidEye AG, BlackBridge Germany)	2008- *	0.44 – 0.51 0.52 – 0.59 0.63 – 0.685 0.69 – 0.73 0.76 – 0.85		630 km	77 km	5m	5.5 days	1 day
Sentinel-2 (ESA, EU)	2015 - * (~ 2027)	center wavelength ; band width – band : 0.443; 0.02 – 1 0.490; 0.065 –2 0.560; 0.035 – 3 0.665; 0.03 – 4 0.705; 0.015 – 5	0.740; 0.015 – 6 0.783; 0.015 – 7 0.842; 0.115 – 8 0.865; 0.02 – 8a 0.945; 0.02 – 9 1.375; 0.03 – 10 1.610; 0.09 – 11 2.190; 0.180 – 12	786 km	290 km	B2,B3,B4,B8: 10m B5,B6,B7,B8a, B11,B12: 20m B1,B9,B10: 60m	10 days (Sentinel-2A); 5 days (Sentinel-2A & 2B)	10 days (Sentinel -2A); 5 days (Sentinel -2A & 2B)
Spot 6/7 (France)	2012 - * (~ 2025)	0.45 – 0.745 0.45 – 0.52 0.53 – 0.59 0.625 – 0.695 0.76 – 0.89			60 km	PAN: 2.2m Others: 8.8m	26 days	1-5 days
UK-DMC-1 / 2 (SSTL, UK)	2003 - *	0.52 – 0.62 0.63 – 0.69 0.76 – 0.90		686 km	650 km	32m	14 days	1 day
HJ-1A/1B (China)	2008 - *	0.43 – 0.52 0.52 – 0.60 0.63 – 0.69 0.76 – 0.90		650 km	360 km	30m	4 days	4 days
CBERS-4 (Ziyuan I-04) (China, Brazil)	2014 - *	MUXCam: 0.45-0.52 0.52-0.59 0.63-0.69 0.77-0.89	PanMUX: 0.51-0.73 (PAN) 0.52-0.59 0.63-0.69 0.77-0.89	748 km	MUXCam: 120 km PanMUX: 60 km	MUXCam: 20m PanMUX: 5m (PAN), 10m (Others)	26 days	3 – 26 days



Remote Sensing Usage

01 Image Classification and Segmentation

02 Change Detection

03 Object Detection etc.

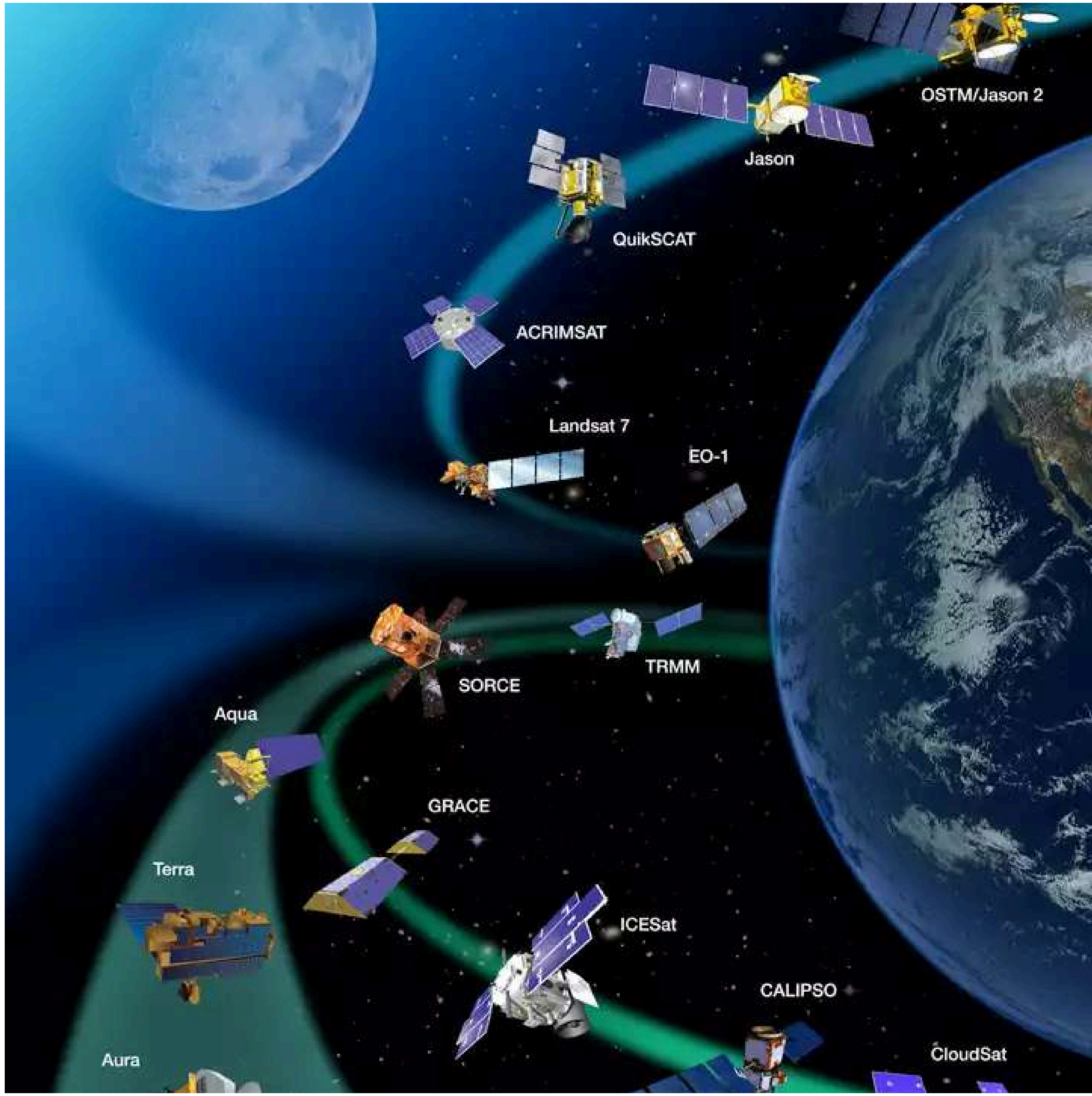
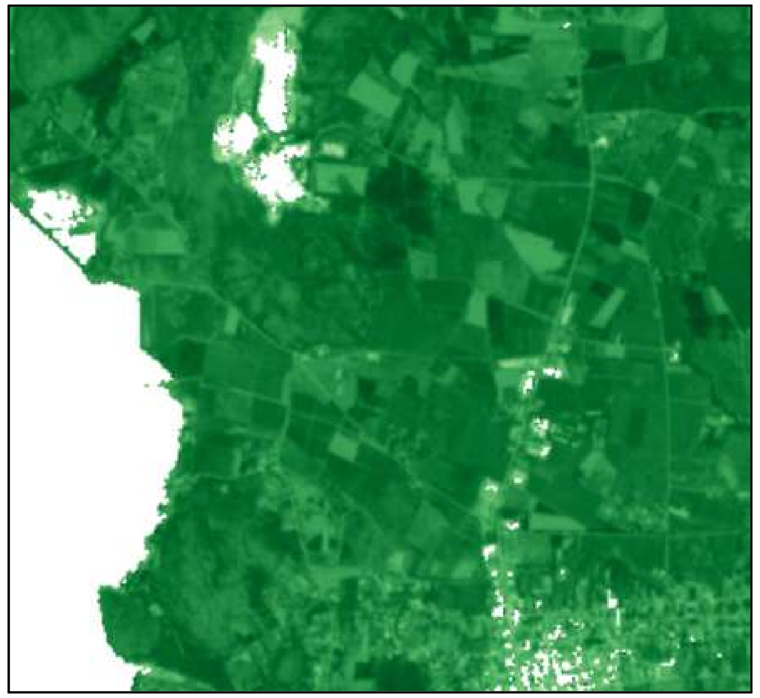


Image Indexing



NDVI

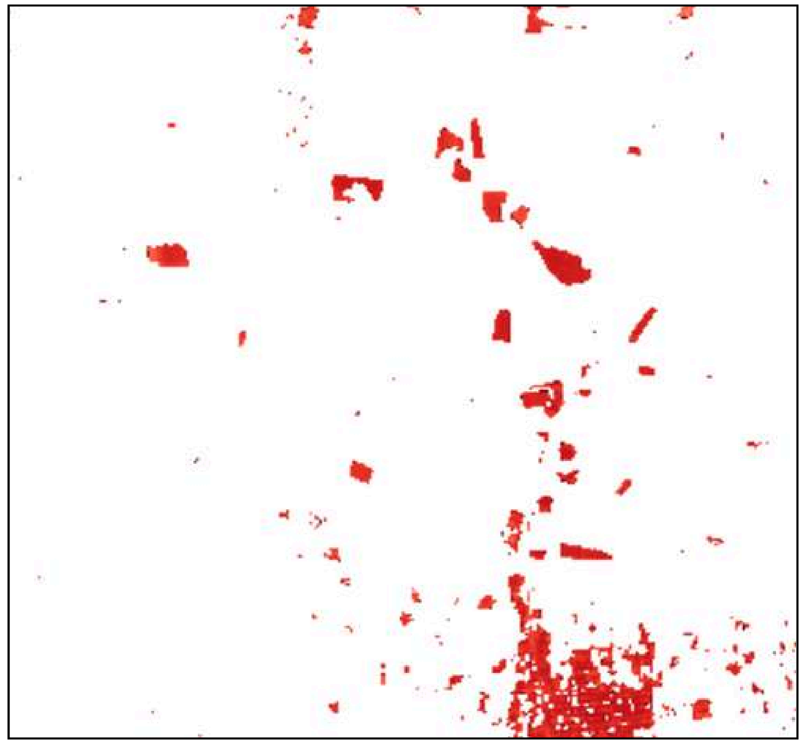
Normalized DIfference Vegetation Index

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

NDWI

Normalized DIfference Water Index

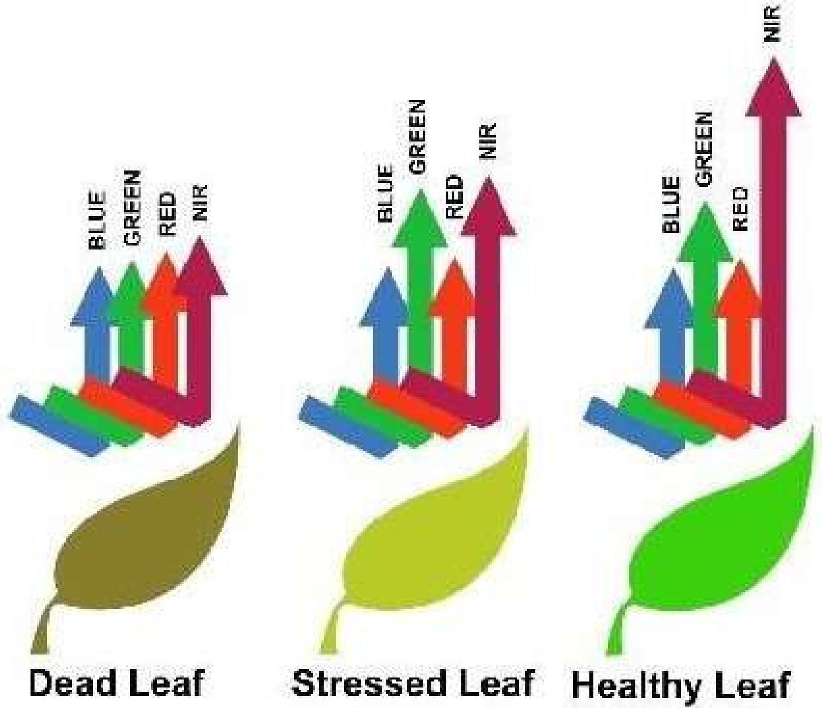
$$NDWI = \frac{Green - NIR}{Green + NIR}$$

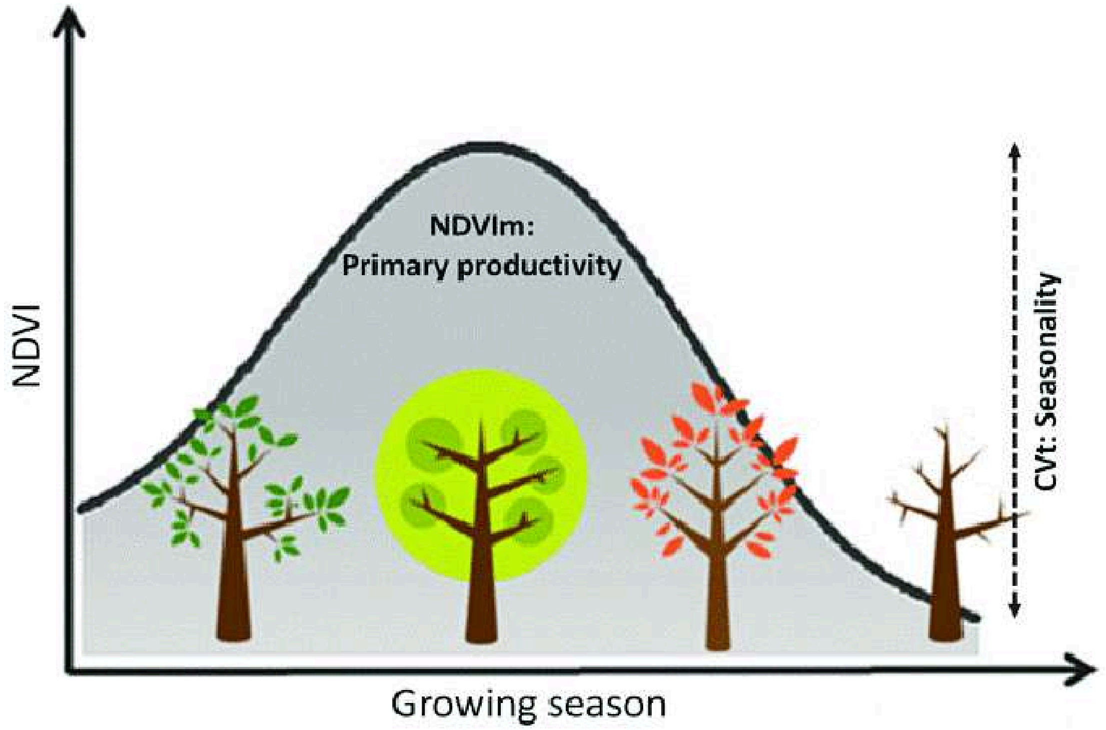


NDBI

Normalized DIfference Built-up Index

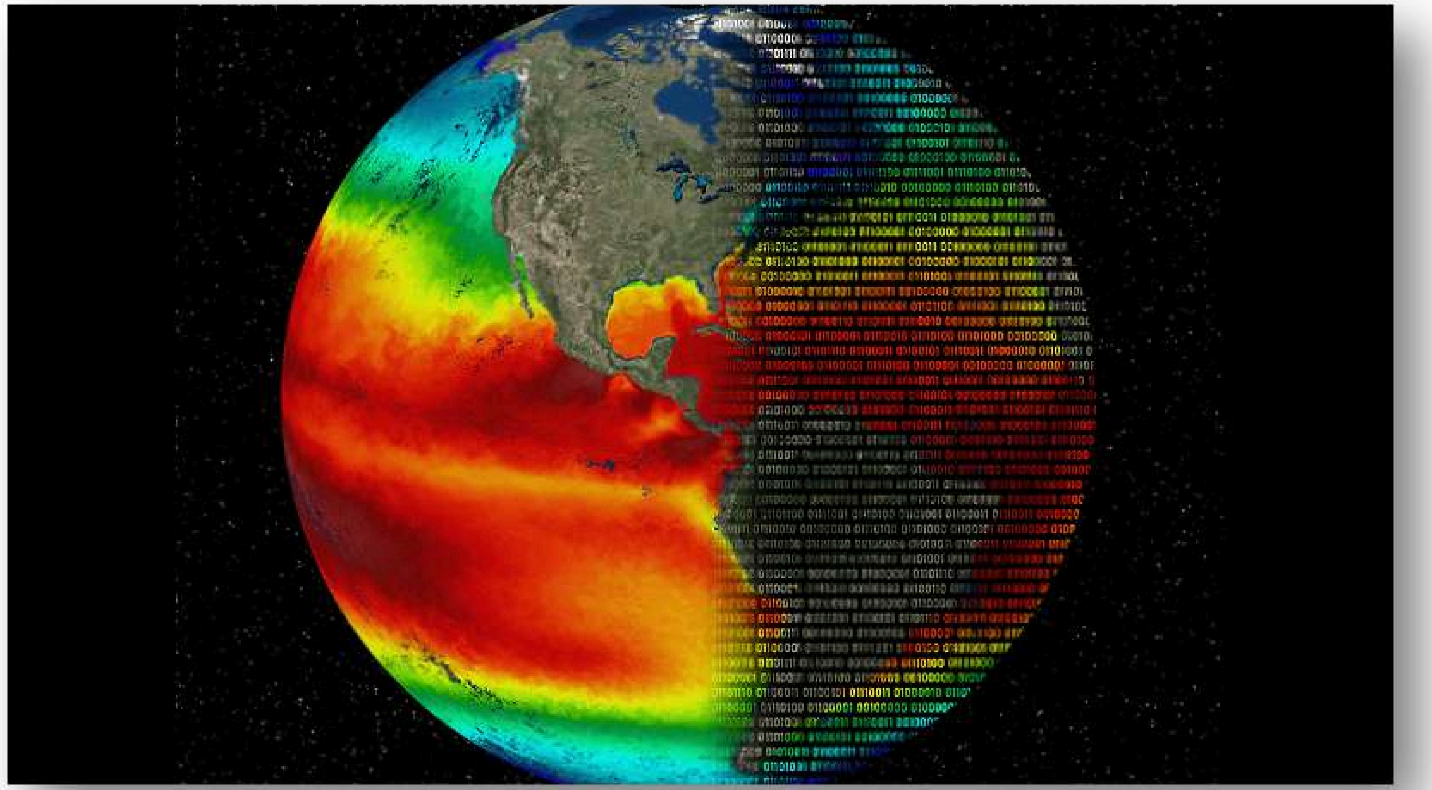
$$NDBI = \frac{SWIR - NIR}{SWIR + NIR}$$





Big Earth Data and AI Technologies

- Big Earth data are collecting at a rapid rate with challenges for understanding and using the data.
- New tools (AI algorithms) and applications are enabling analysis and improving the reliability by policy makers.

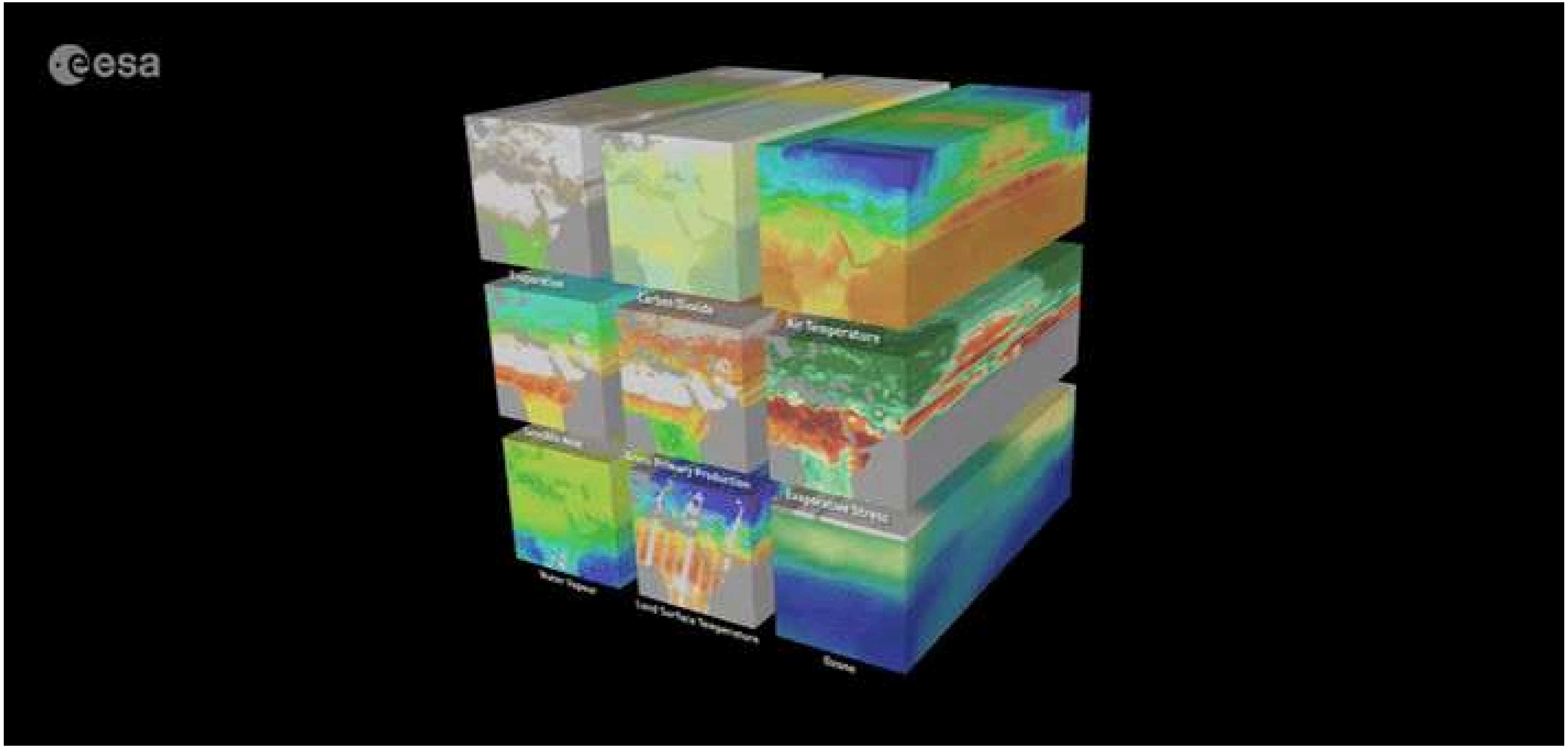


"The application of increasingly sophisticated tools for data analysis and display"

Source: Vance, T. C., T. Huang, and C. Lynnes (2022)

Big Earth Data: platform, data and processing

Euro Data Cube Facility service: the ultimate EO resource for researchers and value-adders



Sources: https://eo4society.esa.int/2019/05/21/european-data-cube-facility-service-an-eo-resource-factory/

Big Earth data characteristics

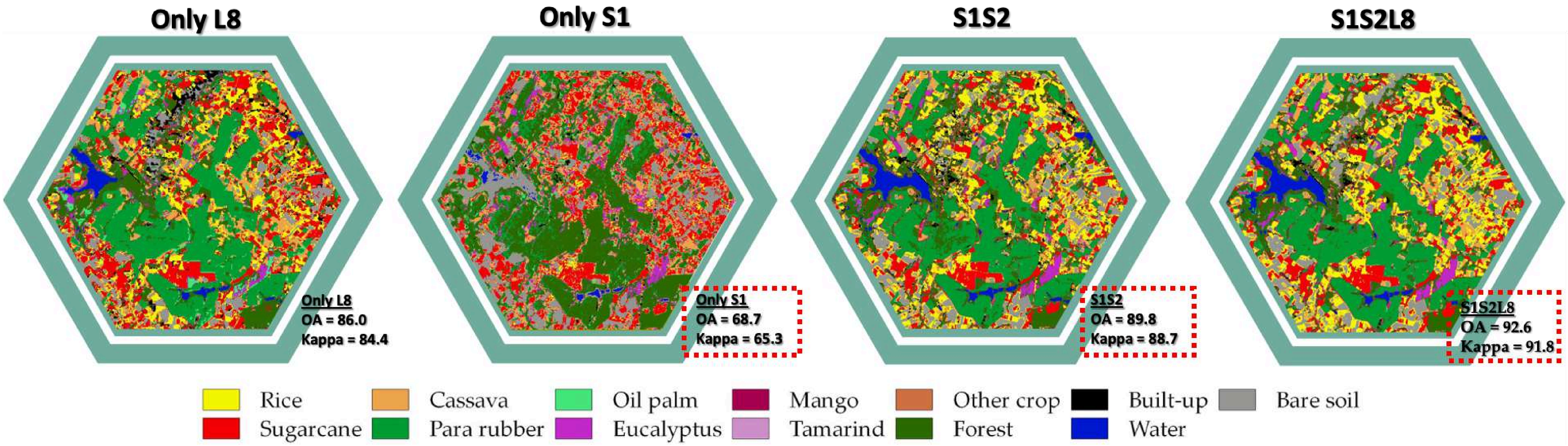
Mainly characterized by “3Vs” (volume, variety, and velocity) Six aspects:

- Large volume: On remote sensing image approximately to 1 GB. For example, only one Sentinel-2 holds about 10 GB.
- Great variety: Because of the difference of sensors, the format of EO data are Hierarchical Data Format (HDF), network Common Data Format (NetCDF), jp2, and GeoTif. For example, original file of Sentinel-2 is jp2.
- Multiple resolutions: With the improvement of remote sensing technology, data resolution is also getting better. The available sensor data are recently used such as MODIS, Landsat and Sentinel-2.
- Time series: Remote sensing satellites dynamically monitor the changes according to old and new data. For instance, Sentinel-2 A and B can cover the Earth every 3-5 days.
- Global scale: Remote sensing can detect large-scale areas from the air and even the space in a short period, and obtain valuable remote sensing data.
- Data intensive: The statistics of the EO data processing link, the rate of data preprocessing and information extraction.

EX: the combination of multi-temporal EO datasets for mapping crop types and land cover

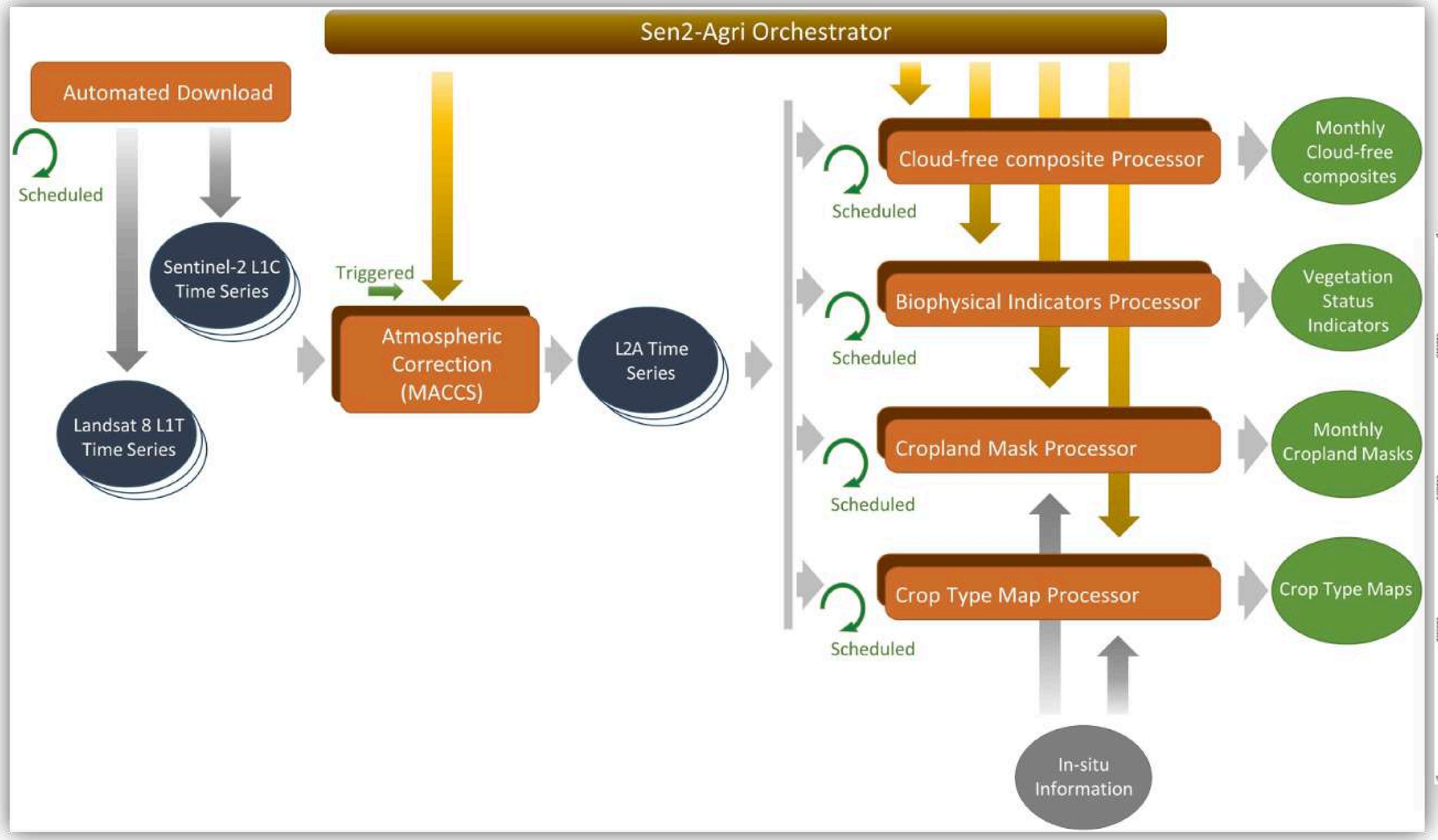
Statistic	L8	S1	S2	S1S2	S1S2L8
OA	86.0	68.7	87.5	89.8	**92.6**
Kappa	84.4	65.3	86.1	88.7	**91.8**

Crop type and Land cover map 2019 in Udon Thani, Thailand.

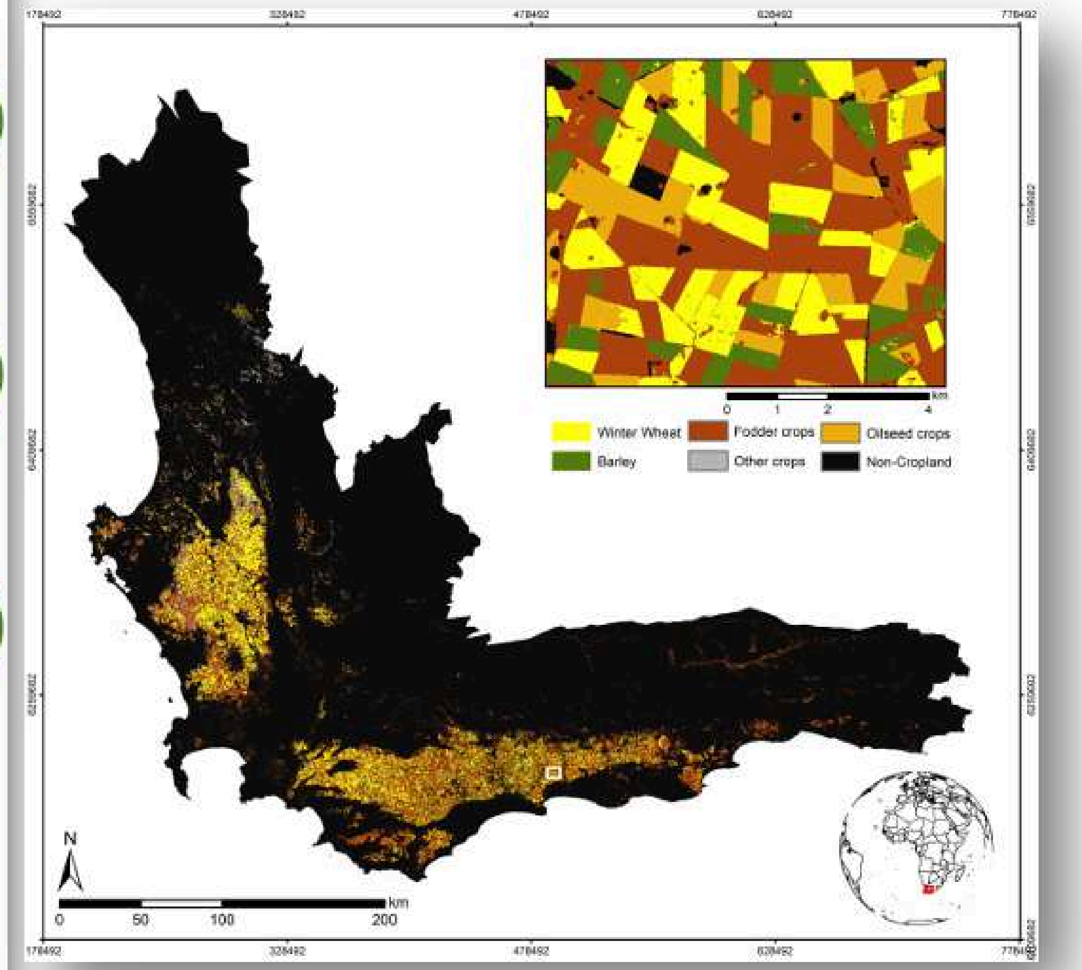


Source: Som-ard, J., et al. (2022); https://doi.org/10.1016/j.compag.2022.107083

Sen2Agri : open source system able to generate, at national scale, cloud-free composites, dynamic cropland masks, crop type maps and vegetation status indicators suitable for most cropping systems.

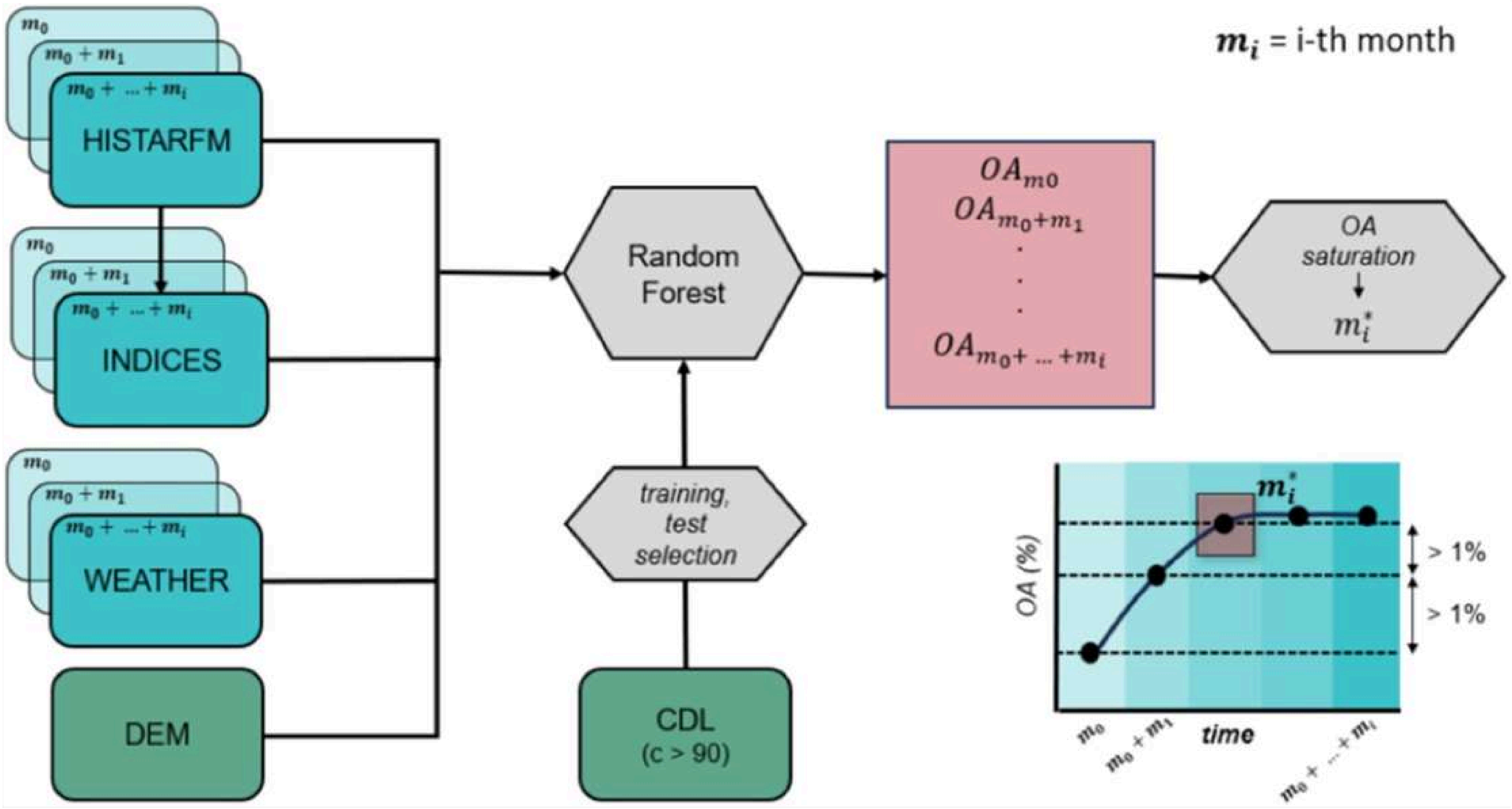


10 m crop type map for 2016 over Western Cape, the main winter grain production region in South Africa.

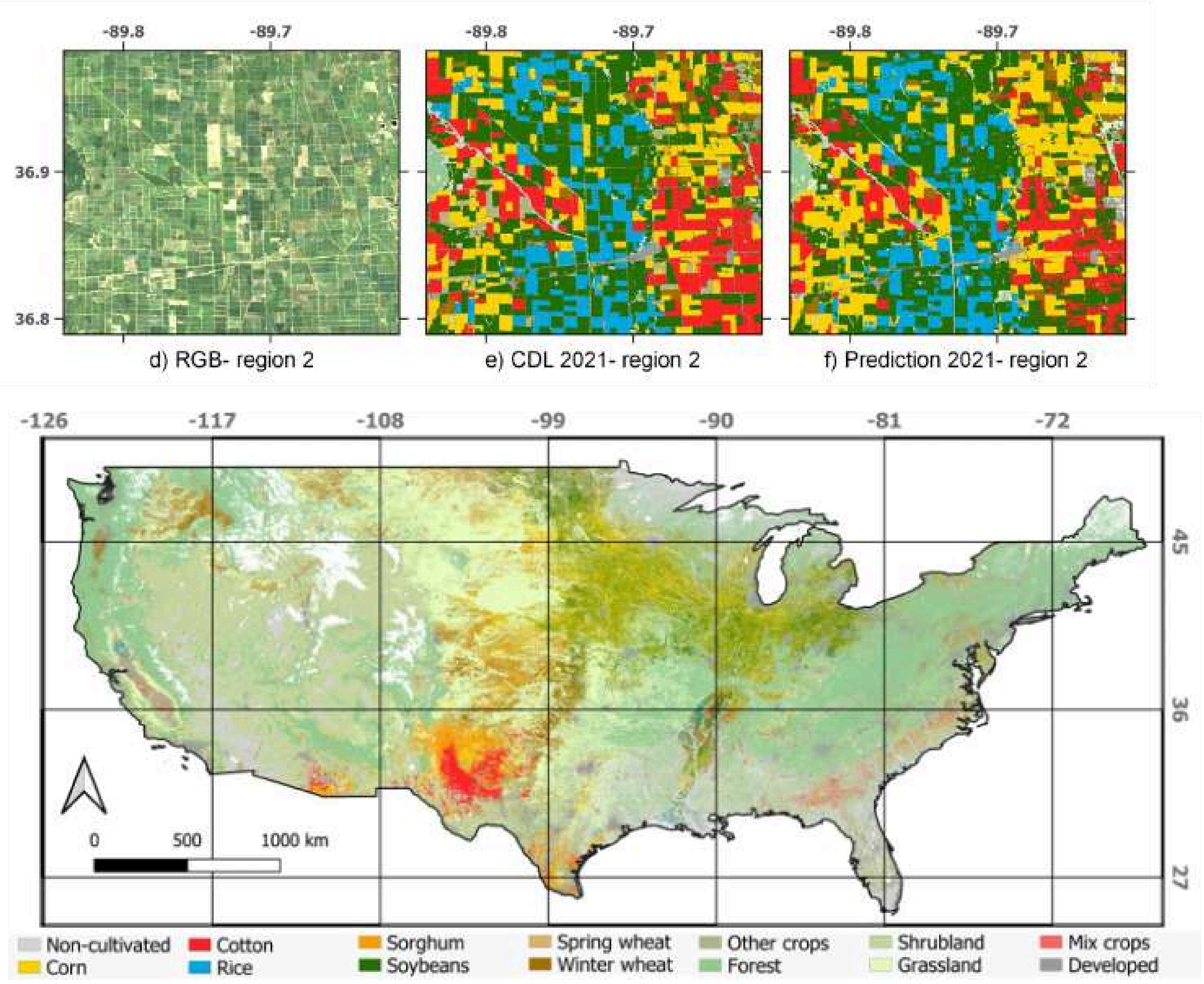


Source: Defourny, P., et al., (2019) ; https://doi.org/10.1016/j.rse.2018.11.007

Classified within-season crop monitoring at continental scale utilizing new gap-filled Landsat temporal series through the Google Earth Engine (GEE) platform.



Workflow for prompt classification of a specific year in GEE platform



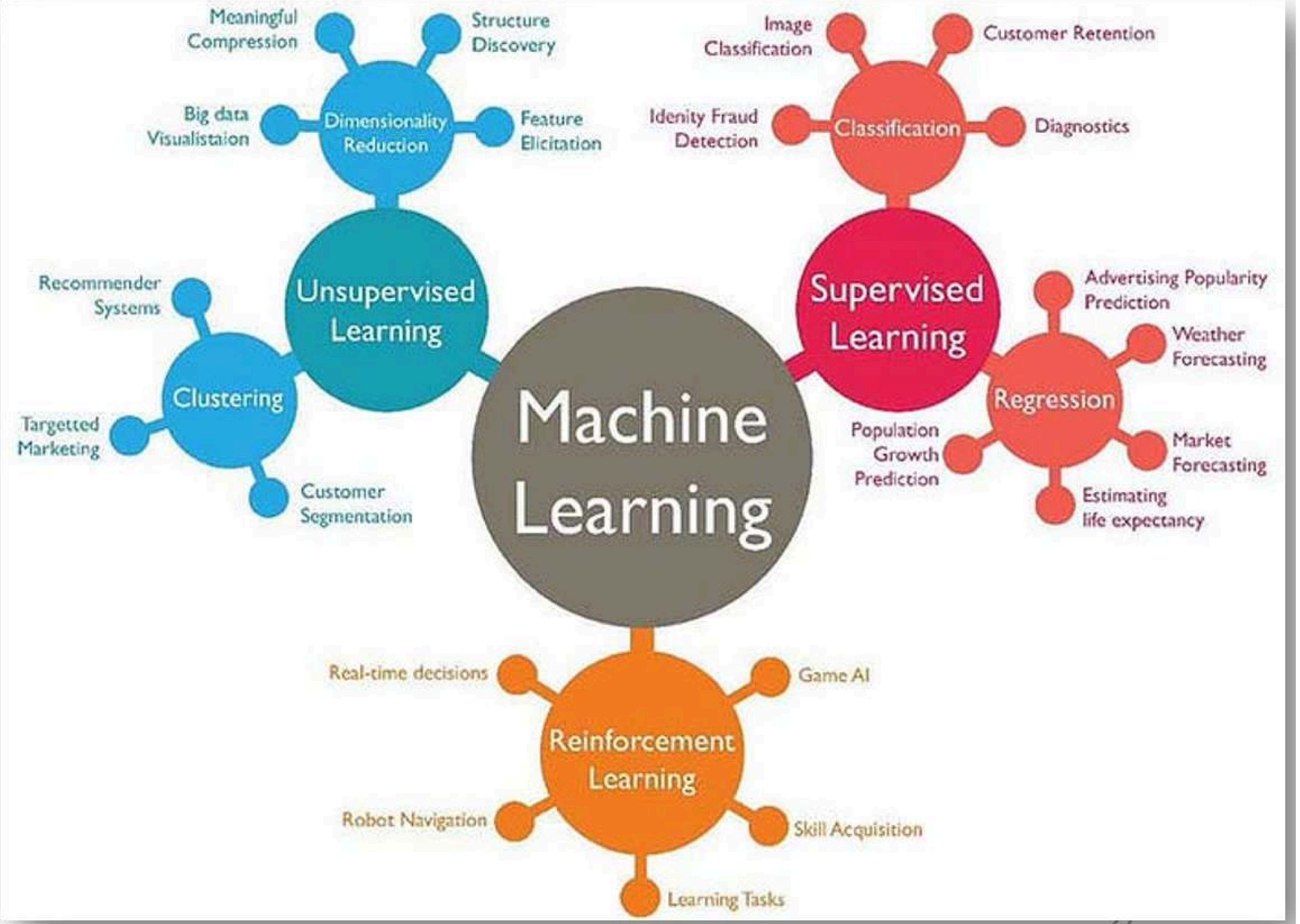
Within-season crop monitoring at continental scale

Source: Rajadel-Lambistos, C., et al.; https://doi.org/10.1080/17538947.2024.2359577

AI: Types of Machine Learning

Machine learning can be classified into 3 types of algorithms:

- Supervised Learning
- Unsupervised Learning
- Reinforcement Learning



Source: Aurélien, G. (2017); https://towardsdatascience.com/introduction-to-machine-learning-for-beginners-eed6024fdb08

Learning process

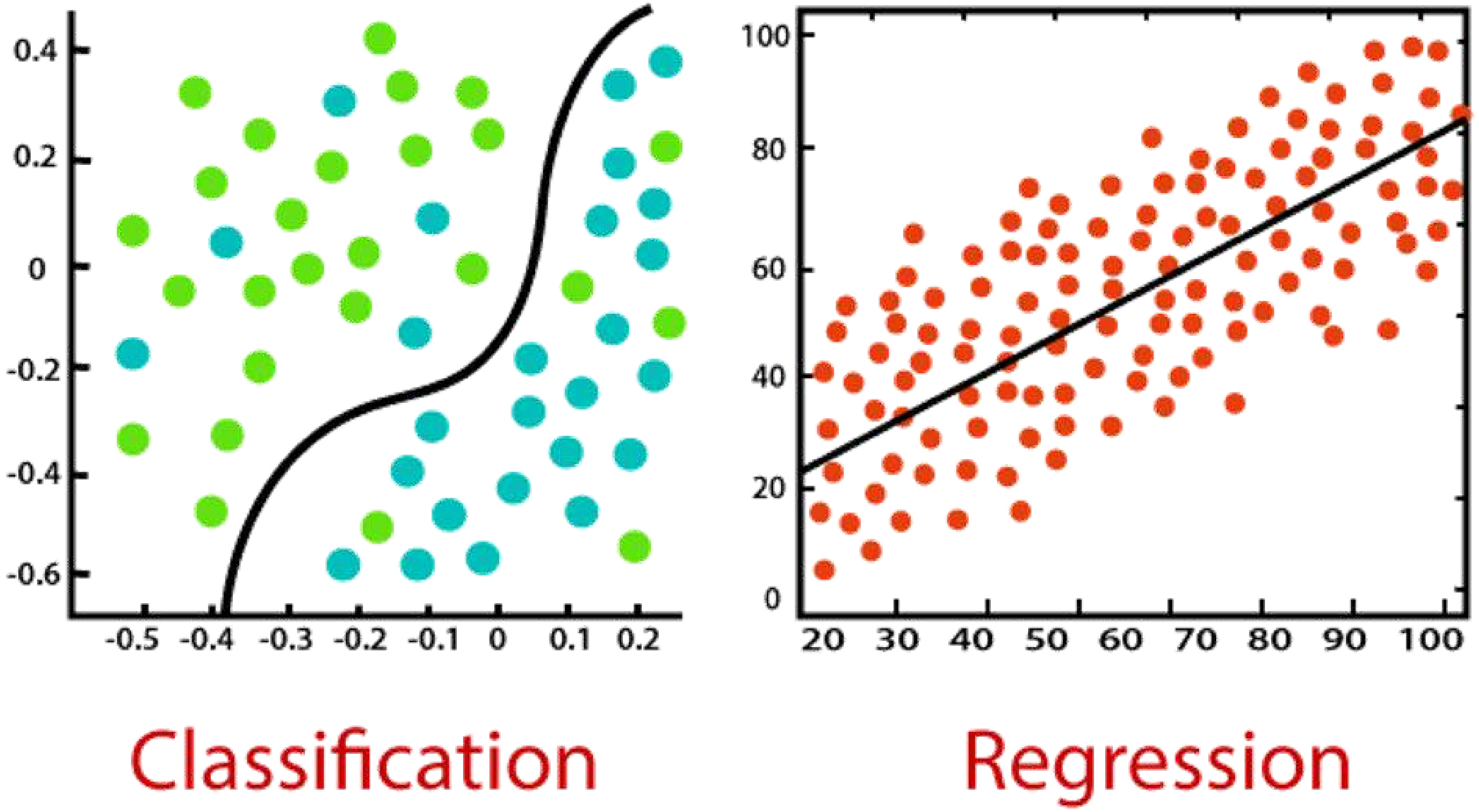
Knowledge = 0

Knowledge = 0.5

Knowledge = 1

Supervised Learning

- Supervised learning involves training a machine learning model on a labeled dataset.
- There are two types:
  - **Regression**: Used for predicting continuous values
  - **Classification**: Used for predicting categorical values

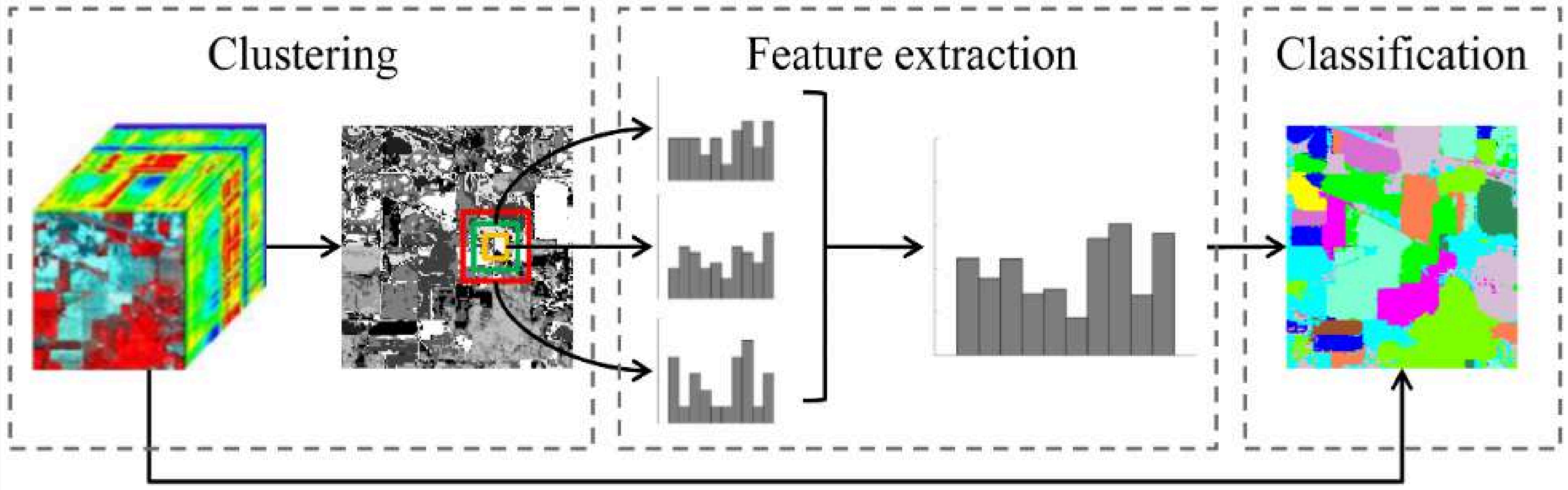


Source: Aurélien, G. (2017); https://www.javatpoint.com/regression-vs-classification-in-machine-learning

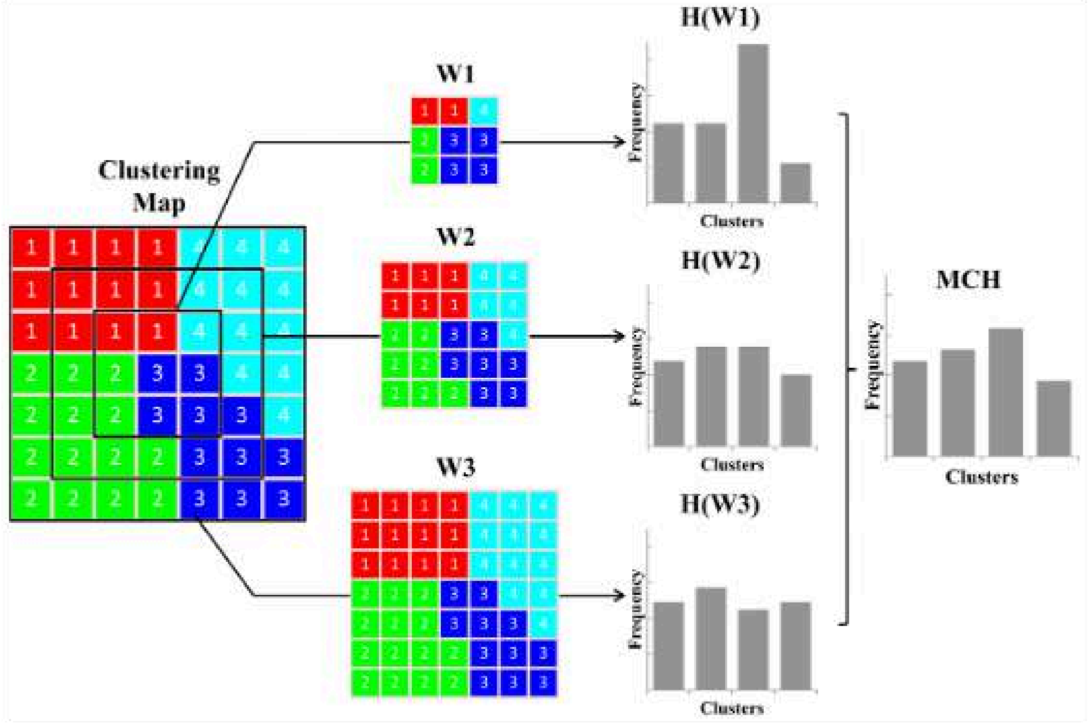
Unsupervised Learning

- Unsupervised learning involves training a model on data that does not have labeled responses.
- The model tries to learn the underlying structure of the data.
- Types of Unsupervised learning:
  - **Clustering**: Grouping similar data points together
  - **Anomaly Detection**: Identifying outliers in data

Ex: the multiscale cluster histogram (MCH) algorithm



Source: Lu, Q., Huang, X., & Zhang, L. (2014)



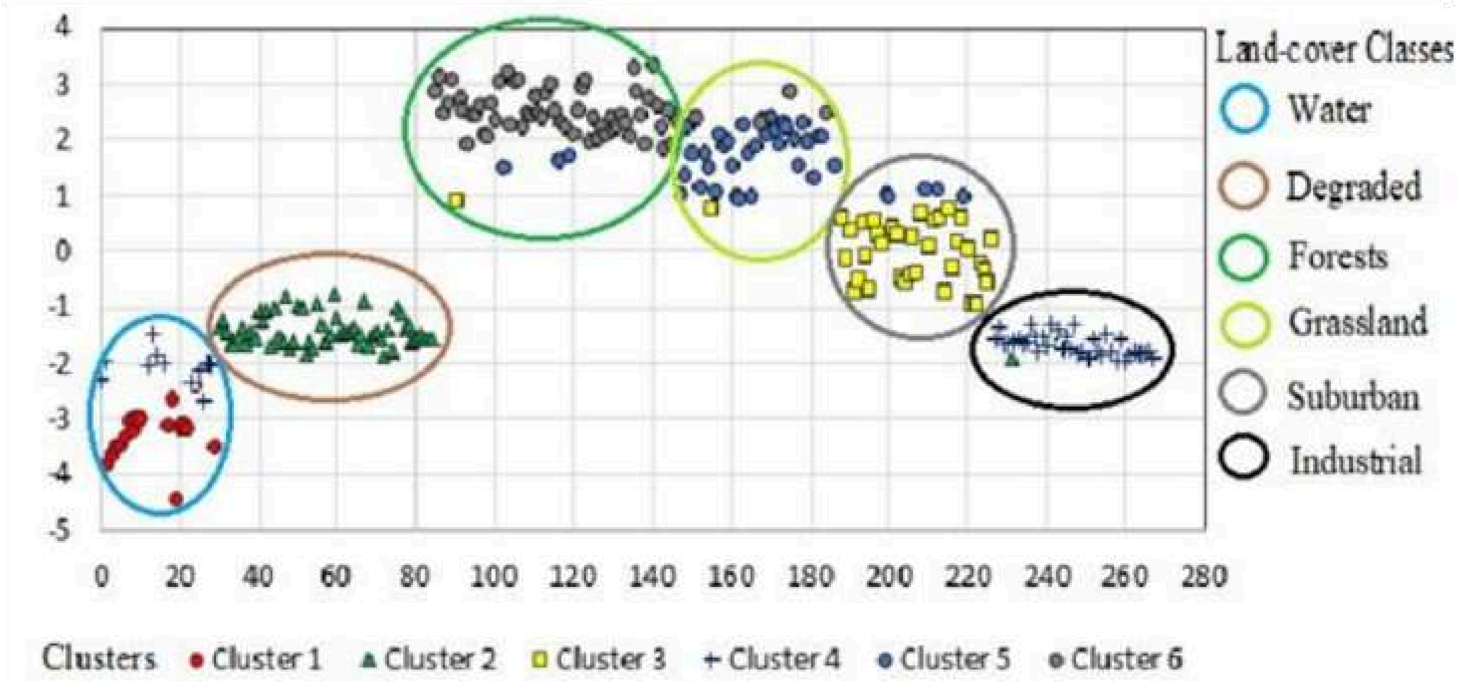
Source: Aurélien, G. (2017); https://towardsdatascience.com/introduction-to-machine-learning-for-beginners-eed6024fdb08

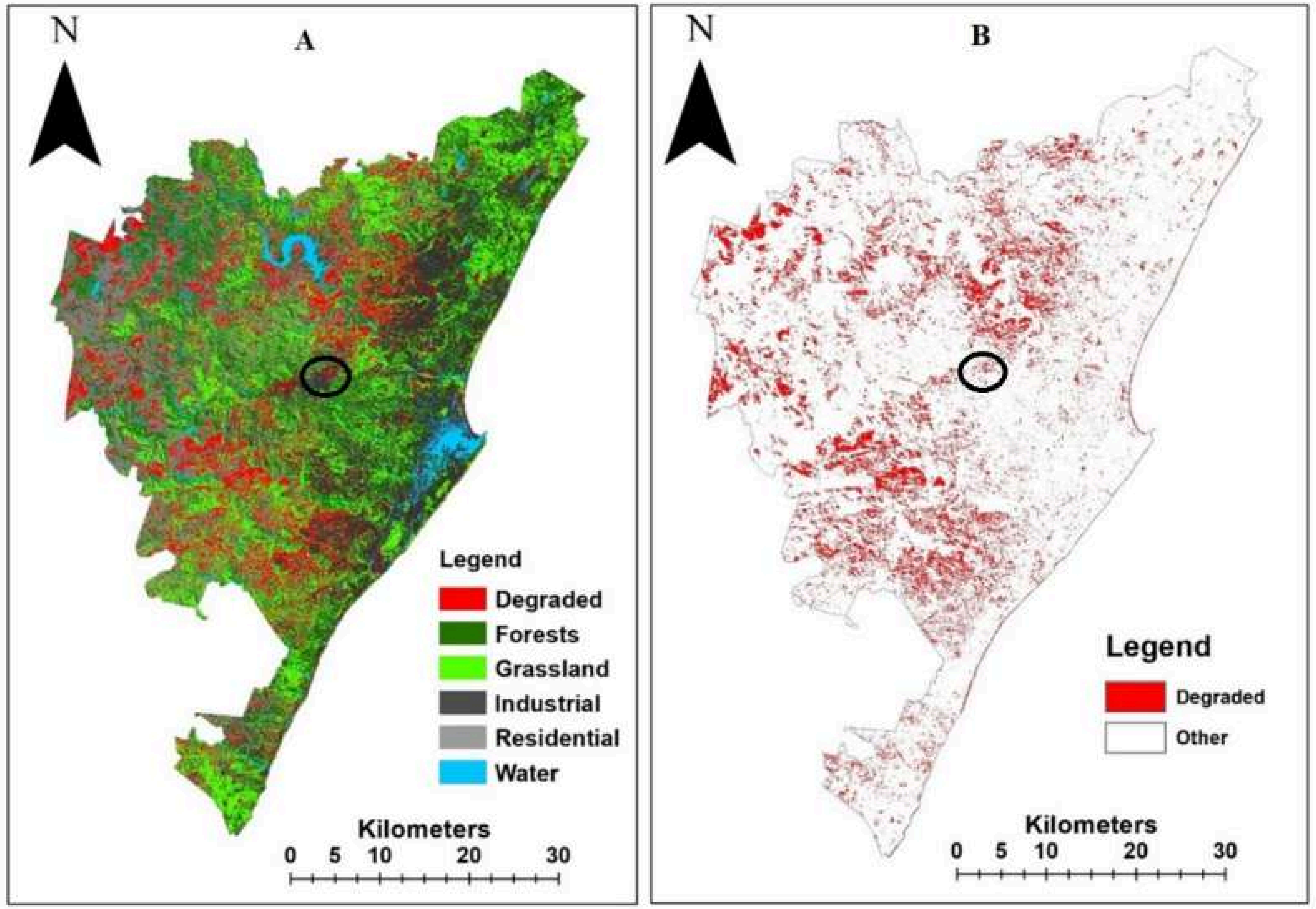
EX: Unsupervised Learning-

Unsupervised cluster algorithm and Landsat image data for monitoring land degradation in Ethekwini Metropolitan, South Africa

- This method is able to map complex land-cover classes, particularly land degradation.
- The mapped results showed the spatial distribution of degraded areas
- Its as significant tool useful for rehabilitation strategies

The groups of Land cover generated by the hierarchical clustering algorithm

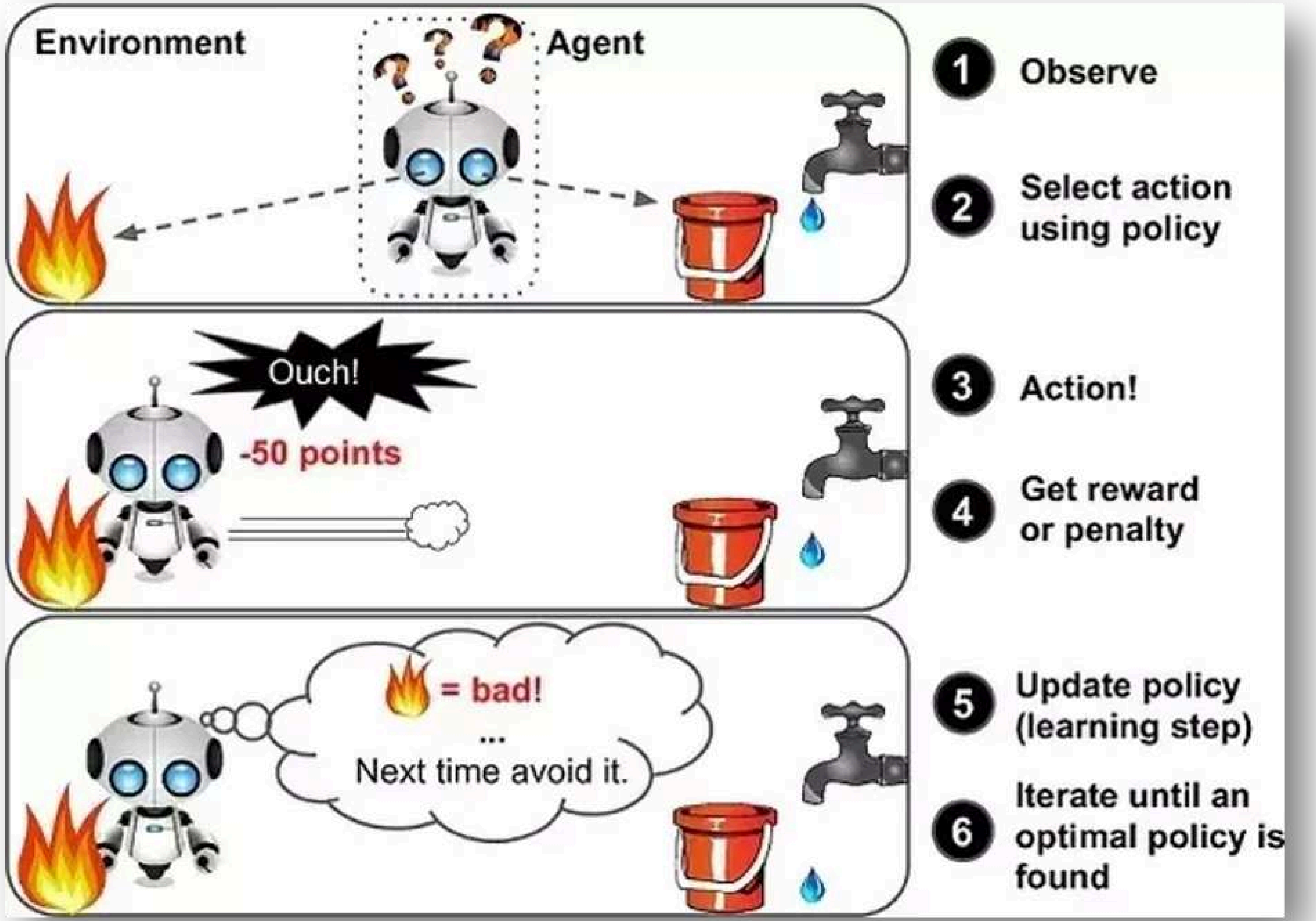




Source: Shange, P. (2020); https://researchspace.ukzn.ac.za/server/api/core/bitstreams/2ce2f59a-6211-41b2-b4f7-68042bcce53e/content

Reinforcement Learning

- Involves training an agent to make a sequence of decisions by rewarding desired behaviors and punishing undesired ones.
- The agent learns to maximize cumulative reward.

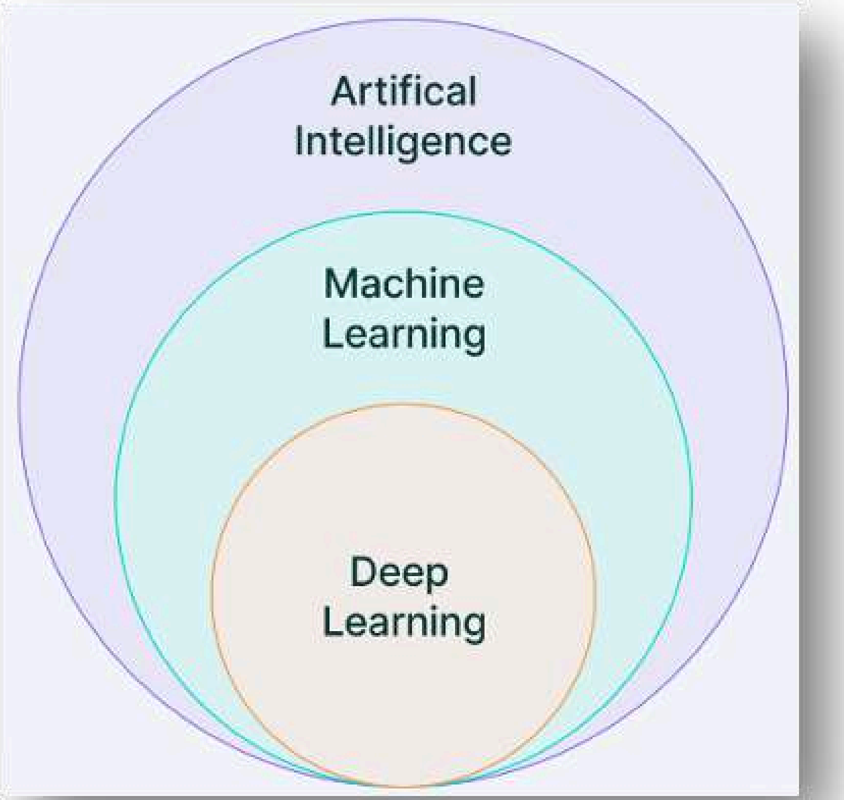


Source: Aurélien, G. (2017); https://towardsdatascience.com/introduction-to-machine-learning-for-beginners-eed6024fdb08

AI: Deep learning

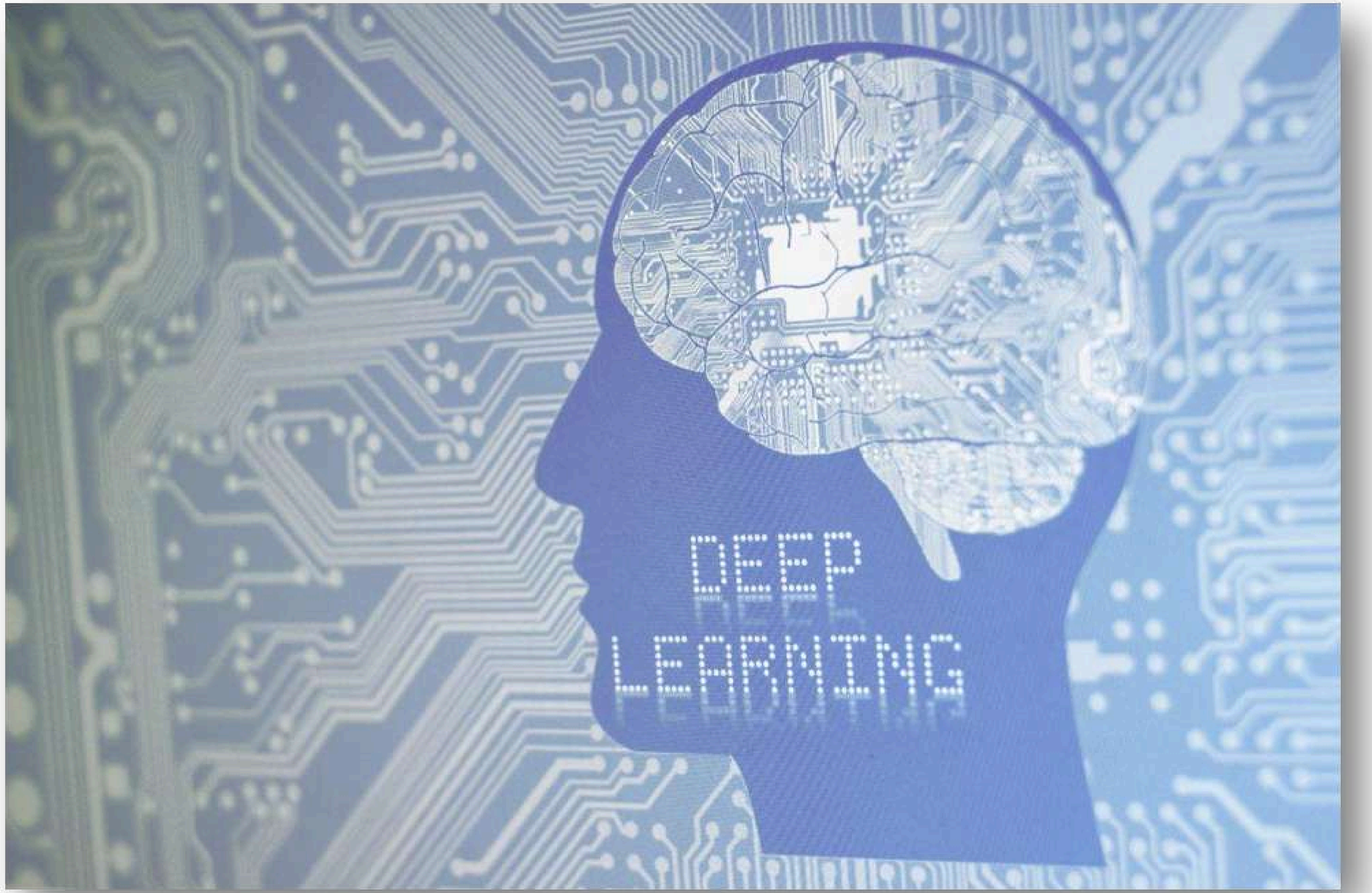
The key characteristics of Deep learning are:

- Complex models
- Larger datasets
- More computational power

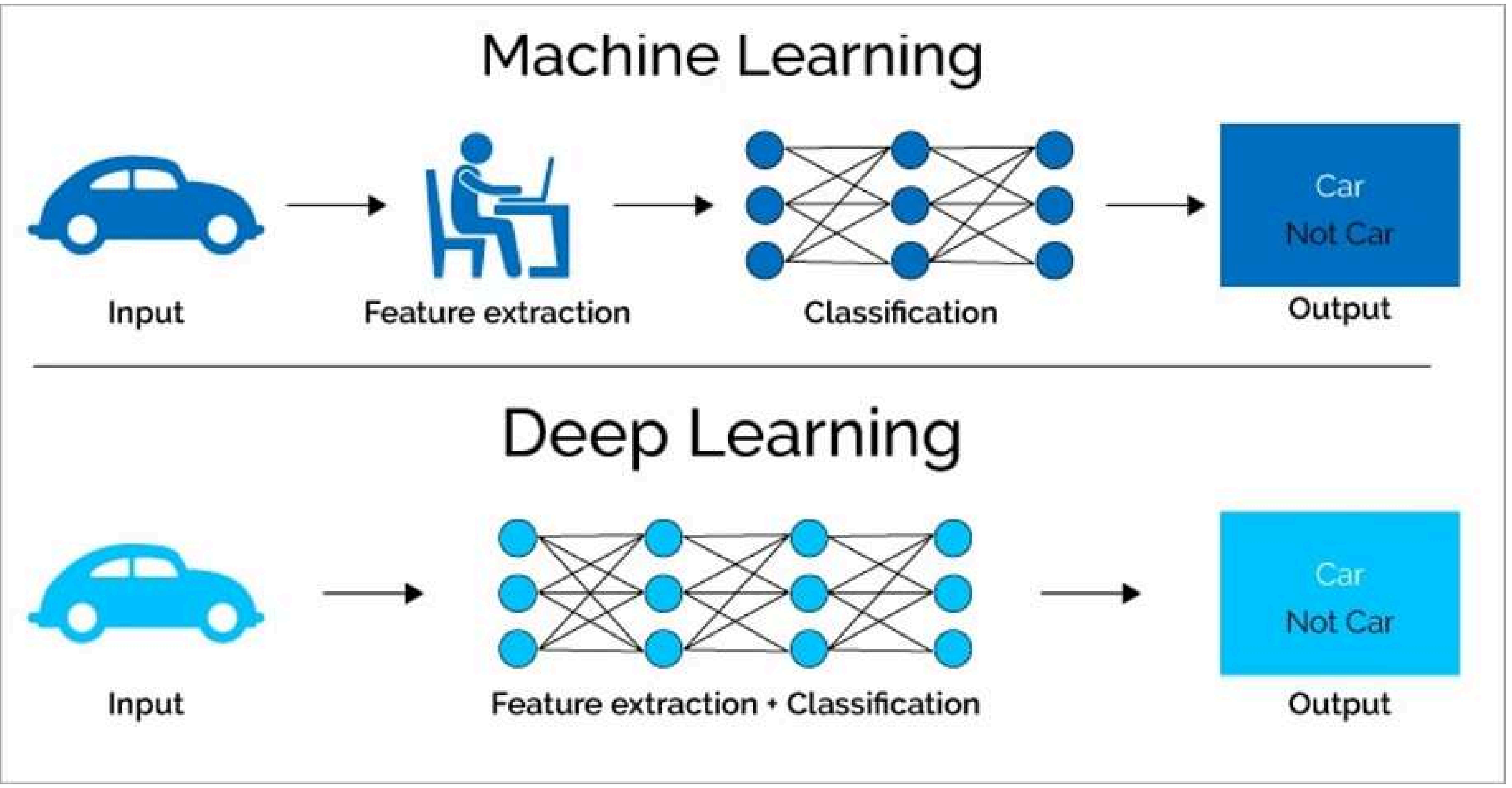


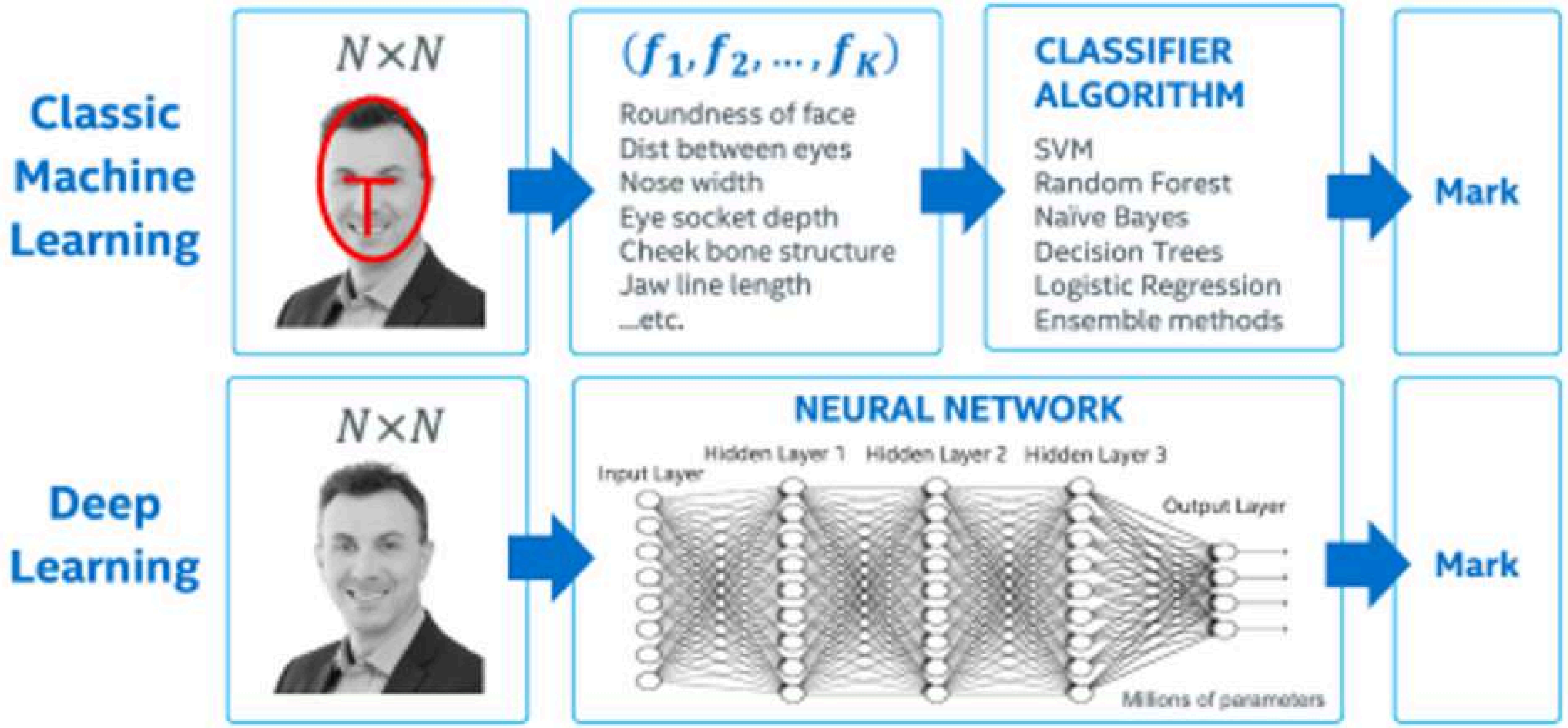
Overview: AI; ML; DL

Source: https://www.v7labs.com/blog/machine-learning-guide



Source: Aurélien, G. (2017); Son, Tim Heinrich, et al. (2023); https://doi.org/10.1016/j.scs.2023.104562

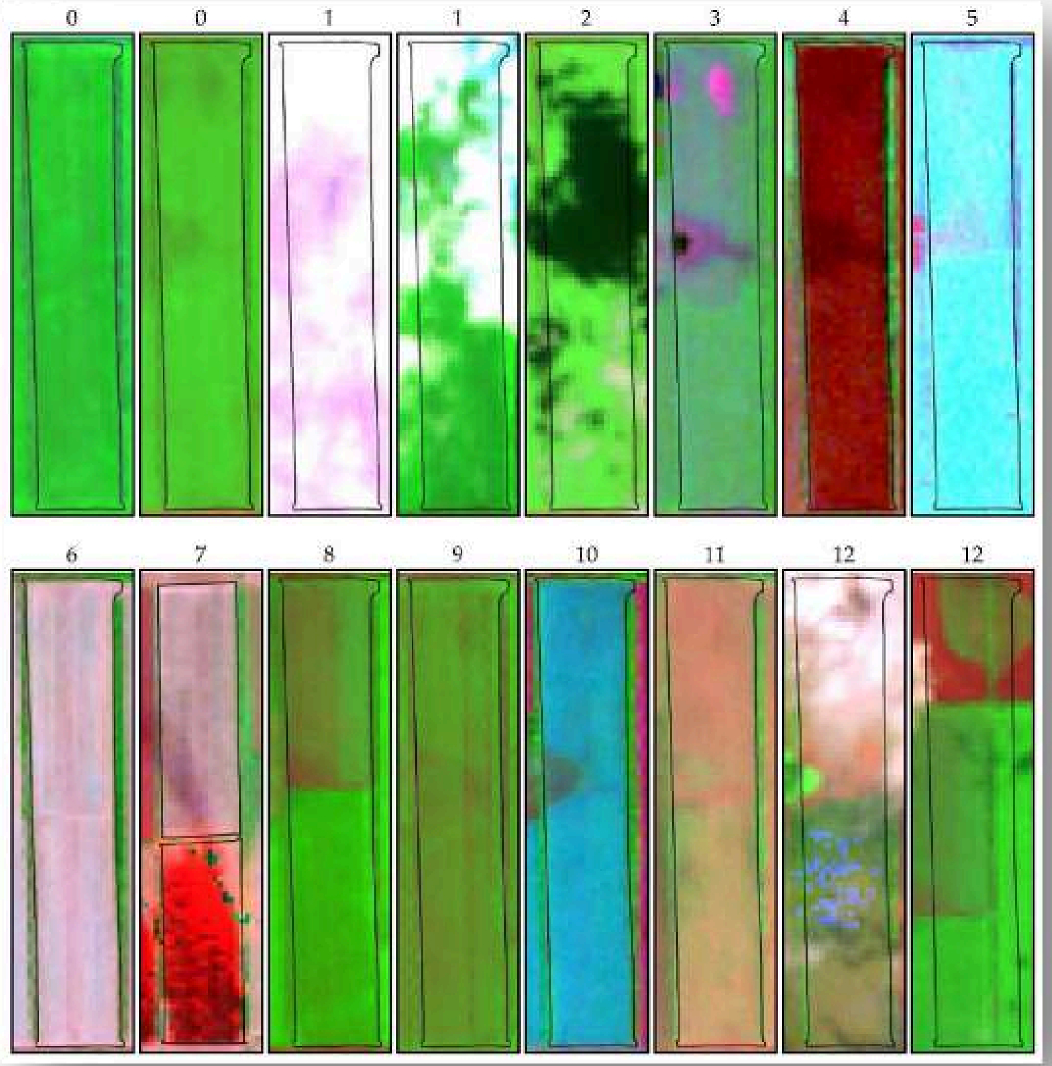

Machine Learning
Input
Feature extraction
Classification
Car
Not Car
Output
Deep Learning
Input
Feature extraction + Classification
Car
Not Car
Output

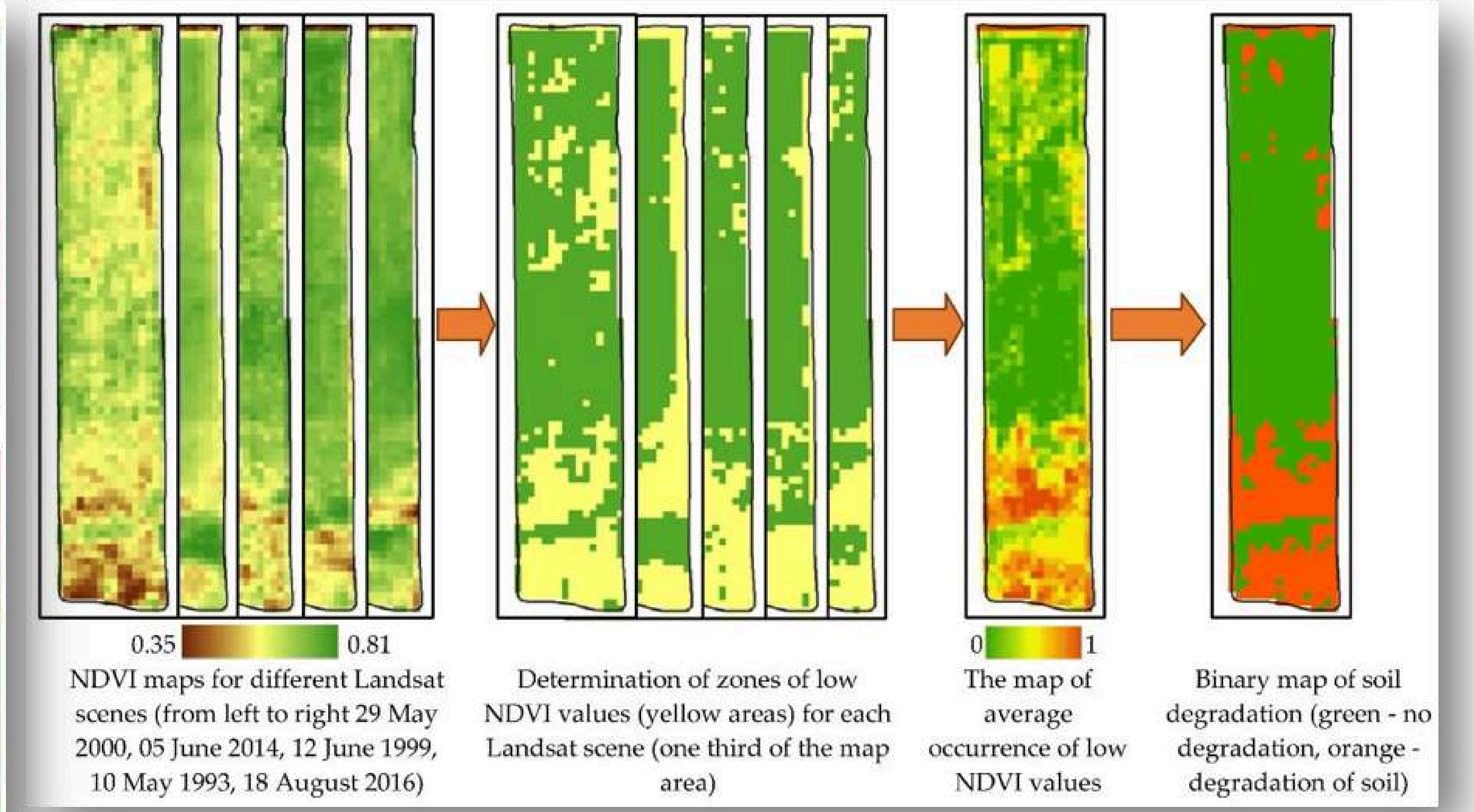

Classic Machine Learning
$N \times N$
$(f_1, f_2, \ldots, f_K)$
Roundness of face
Dist between eyes
Nose width
Eye socket depth
Cheek bone structure
Jaw line length
...etc.
CLASSIFIER ALGORITHM
SVM
Random Forest
Naïve Bayes
Decision Trees
Logistic Regression
Ensemble methods
Mark
Deep Learning
$N \times N$
NEURAL NETWORK
Input Layer
Hidden Layer 1
Hidden Layer 2
Hidden Layer 3
Output Layer
Millions of parameters
Mark

AI: Deep learning

The example Deep learning algorithms:

- **Convolutional Neural Networks (CNNs)**: CNNs are specialized neural networks designed for processing structured grid data like images.
  - **Image Recognition**: In land degradation, CNNs can be used to determine arable land degradation processes distribution.





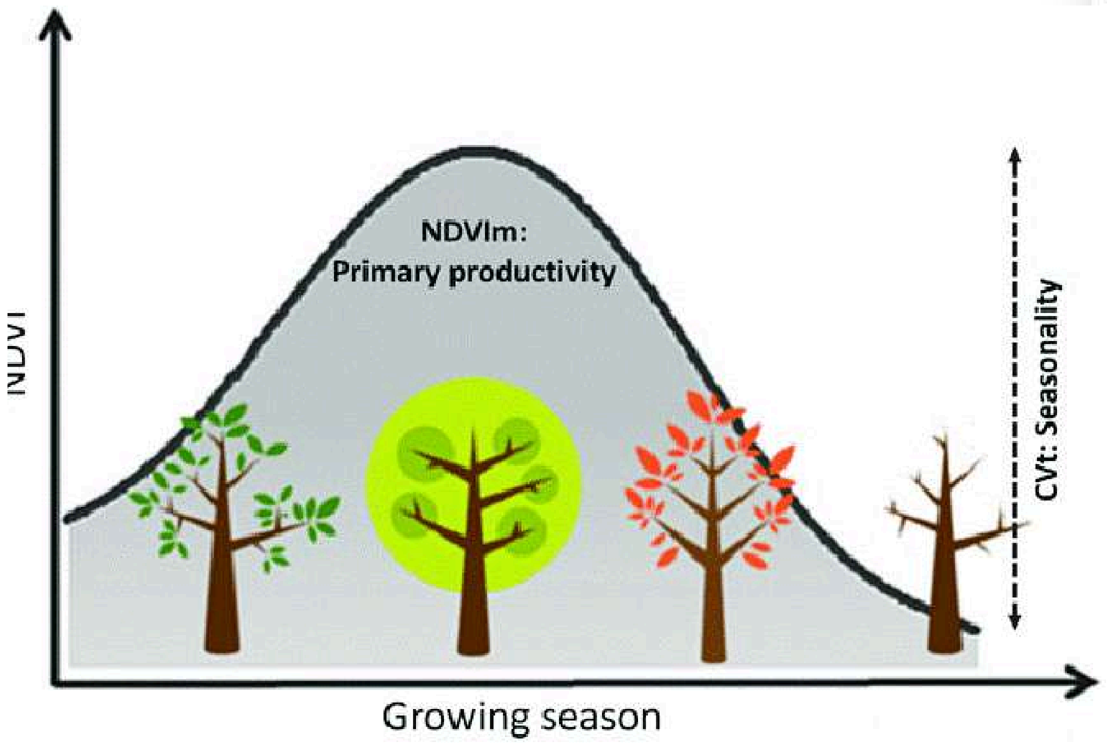
Source: Rukhovich, D. I., et al. (2021); https://doi.org/10.3390/rs13010155

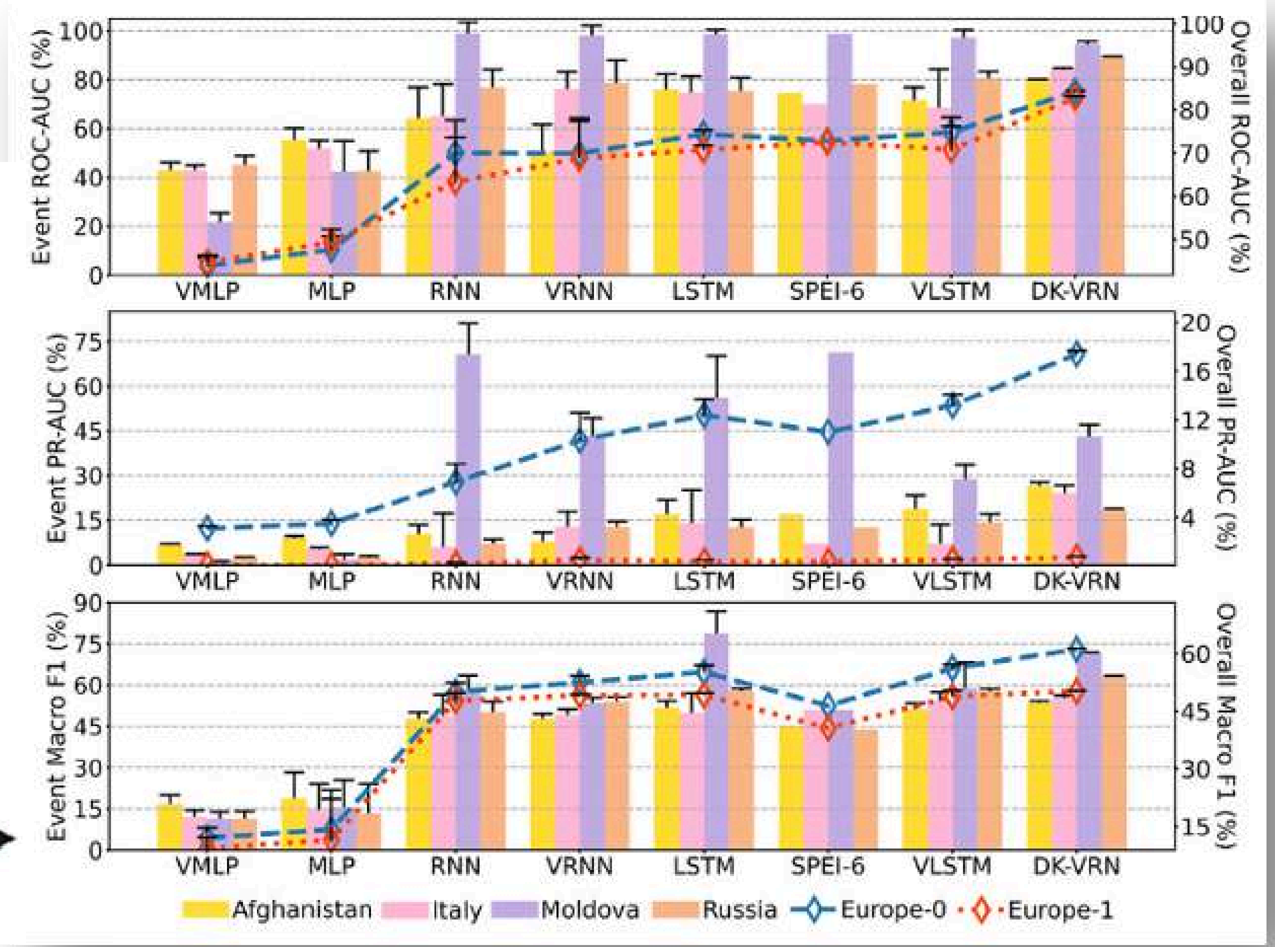
AI: Deep learning

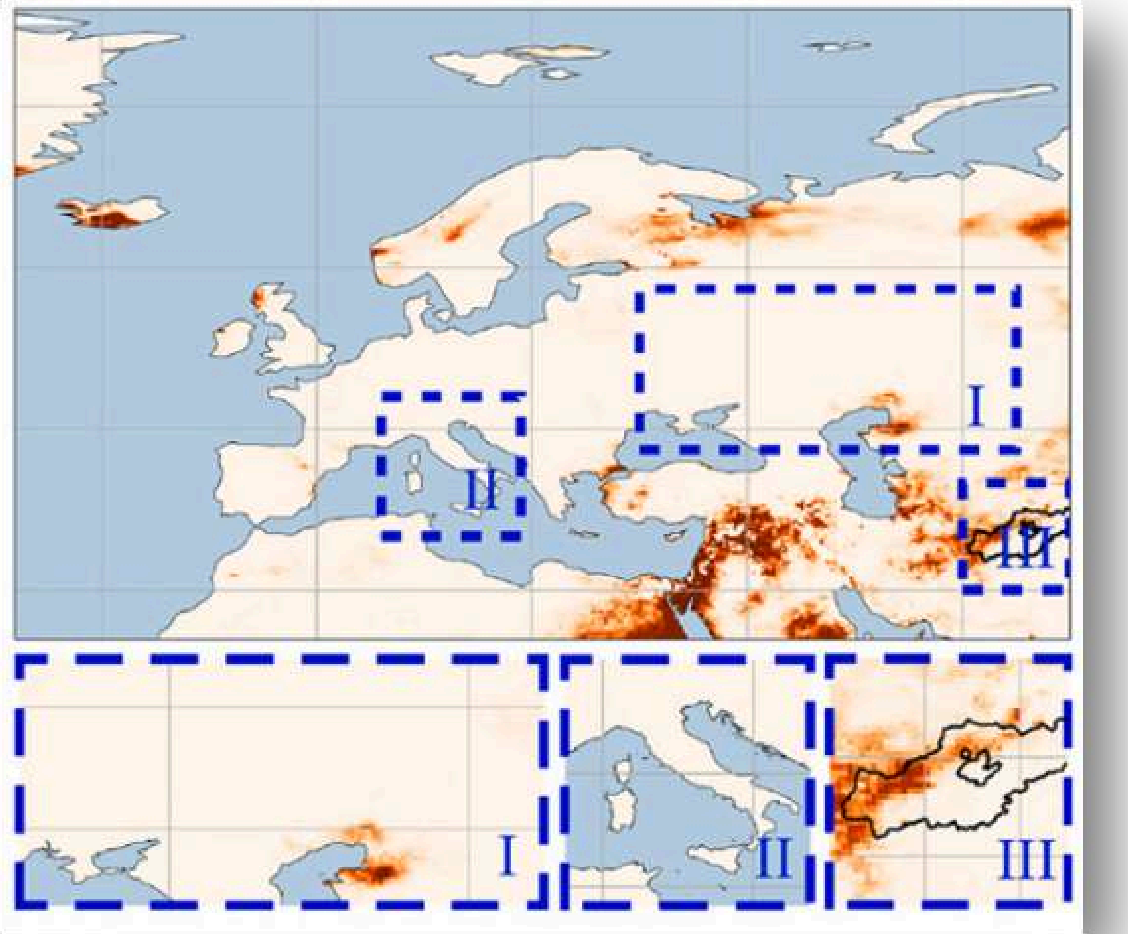
The example Deep learning algorithms:

- **Recurrent Neural Networks (RNNs)**: RNNs are neural networks designed for sequential data.

  - **Time Series Analysis**: In land degradation, RNNs can be used to analyze drought events over time, predict future drought conditions, and forecast environmental variables like rainfall and temperature.







Source: Aurélien, G. (2017); Zhang, M., Fernández-Torres, M. Á., & Camps-Valls, G. (2024); https://doi.org/10.1016/j.rse.2024.114252

Overall Objective

Macintosh HD

To develop Convolutional Neural Network Long Short-Term Memory model for monitoring mangrove forest change using multitemporal and multi-incidence angle Sentinel-1 SAR backscatter on VV and VH polarizations

Specific Objectives

1. To investigate the potential of multitemporal and multi-incidence angle Sentinel-1 SAR backscatter on VV and VH polarizations correlated with vegetation indexes in discriminating mangrove forest stages and their change

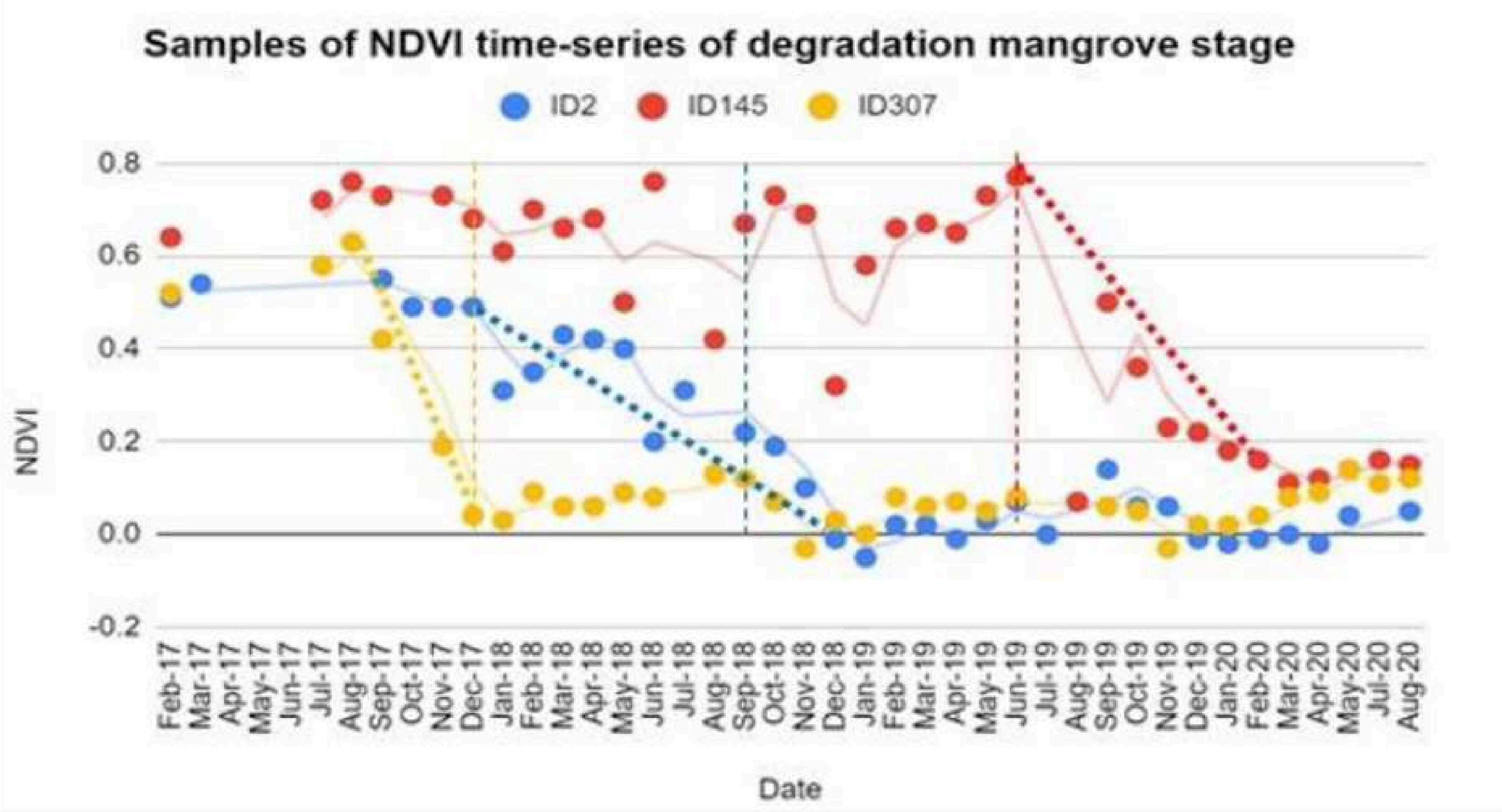
Research Questions

1. Which polarization of Sentinel-1 SAR provides better useful for the mangrove change discrimination?
2. Which incidence angle range is the most useful for the mangrove change discrimination?
3. Do single range or every range of incidence angel of Sentinel-1 SAR data improve the accuracy?
4. Do any SAR features have correlation with NDVI, NDMI, NDVIre1 or CIre1 for mangrove change discrimination?
5. Which the Sentinel-1 SAR feature has highest correlation with vegetation index and will be the input of the model?

RESULTS

- The samples plots ID2, ID145 and ID307 are representatives of degradation mangrove
- The degradation(disturbance) can be divided to natural degradation and clear cut.
- The degradation mangrove stages are defined using NDVI rule-based on time-series dimension (NDVI < 0.3)
- The blue dots illustrates gradually degraded mangrove, Red and yellow dots are the representative of clear-cut mangrove

Figure 22
Time and slope of degradation mangrove sample plots during February 2017 - August 2020 of NDVI



- Rule-based does not reveal the real staring points of mangrove degradation because NDVI values before degradation period or during the heathy stage are not the same. Some are very dense mangrove forest which NDVI values could reach 0.7
- The slope from the stable healthy point to stable degraded point, the slopes are -0.55 and -039 for plot ID145 and ID307 respectively (Clear cut mangrove). The slope for gradually degradation plot (ID2) is -0.50

RESULTS

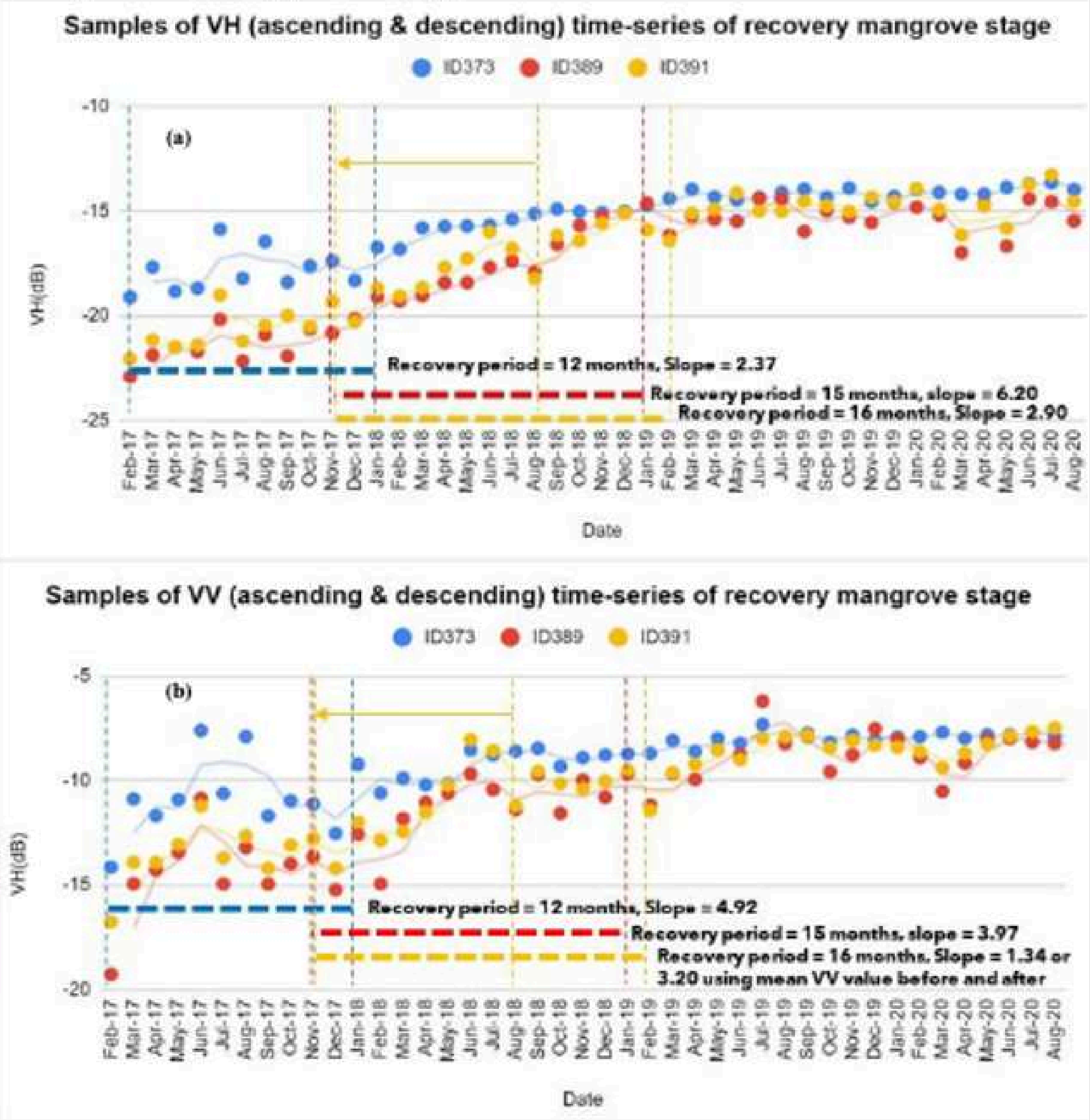
- SAR backscatter change during mangrove recovery, defined using NDVI rule-based on time-series dimension
- The observations of plot ID391 (yellow dots) are slightly higher than ID389 during recovery both in VH (B) and VV (C) backscatter
- This is opposite to NDVI time-series because plot ID391 has slightly lower NDVI

- The slope of SAR variables during recovery period are higher than NDVI
- Consequently, at the beginning stage of recovery, NDVI values shows less values due to the reflectance of bare land and water

- SAR variables have capacity to detect the stem of new mangrove generation
- In results, SAR backscatter shows higher slope in recovering period than NDVI

Figure 21

Time and slope of recovery mangrove sample plots during February 2017 to August 2020 of (a) VH and (b) VV



- SAR backscatter reach the saturated point earlier comparing to NDVI
- SAR data can detect recovery mangrove well as it could detect the volume of stem of the new mangrove generation while NDVI could detect mangrove regeneration well only when the canopy develop

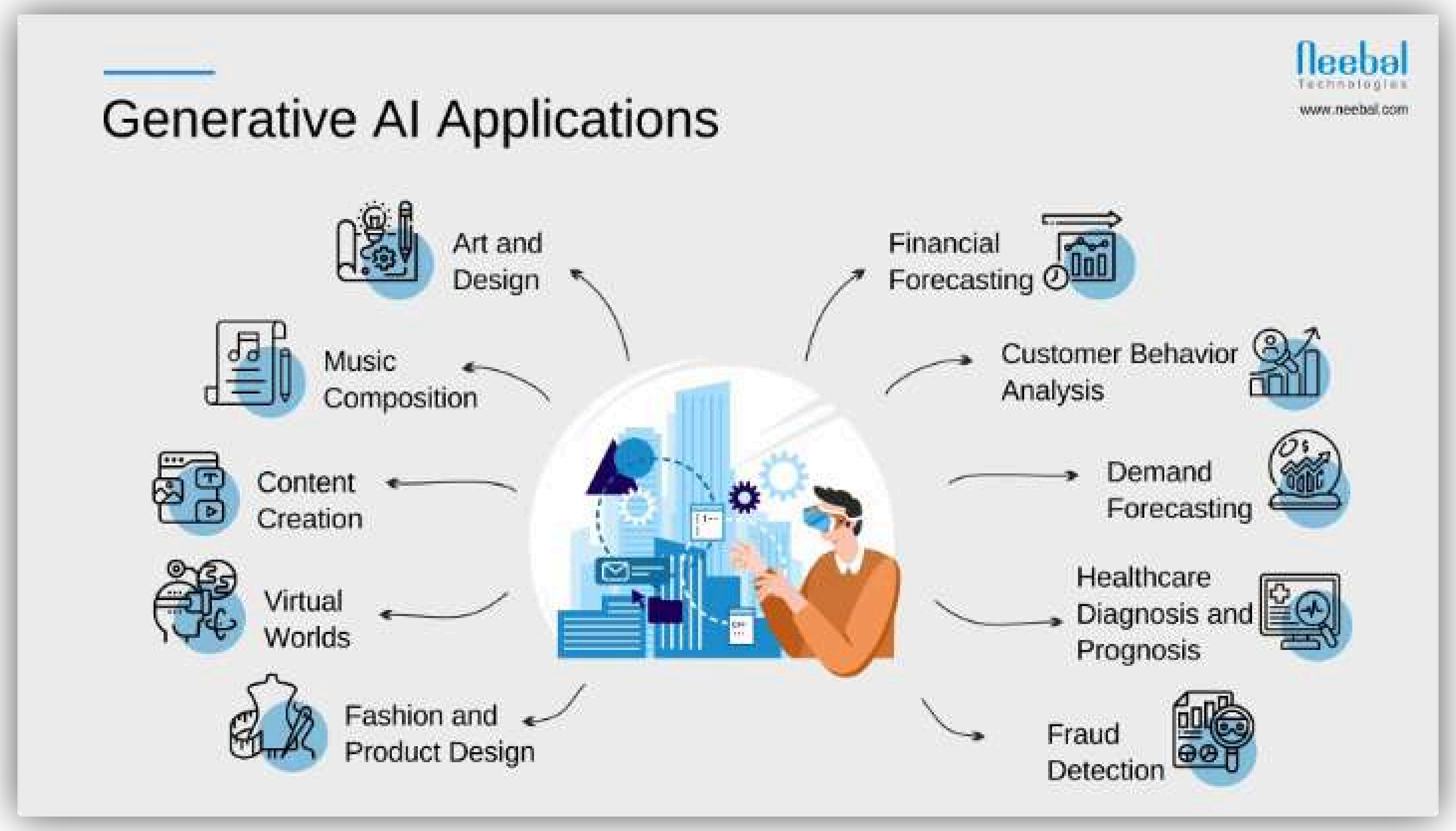
AI: Generative AI

Here's a breakdown of how generative AI works:

- Training on massive datasets
- Learning patterns and relationships
- Generating new content

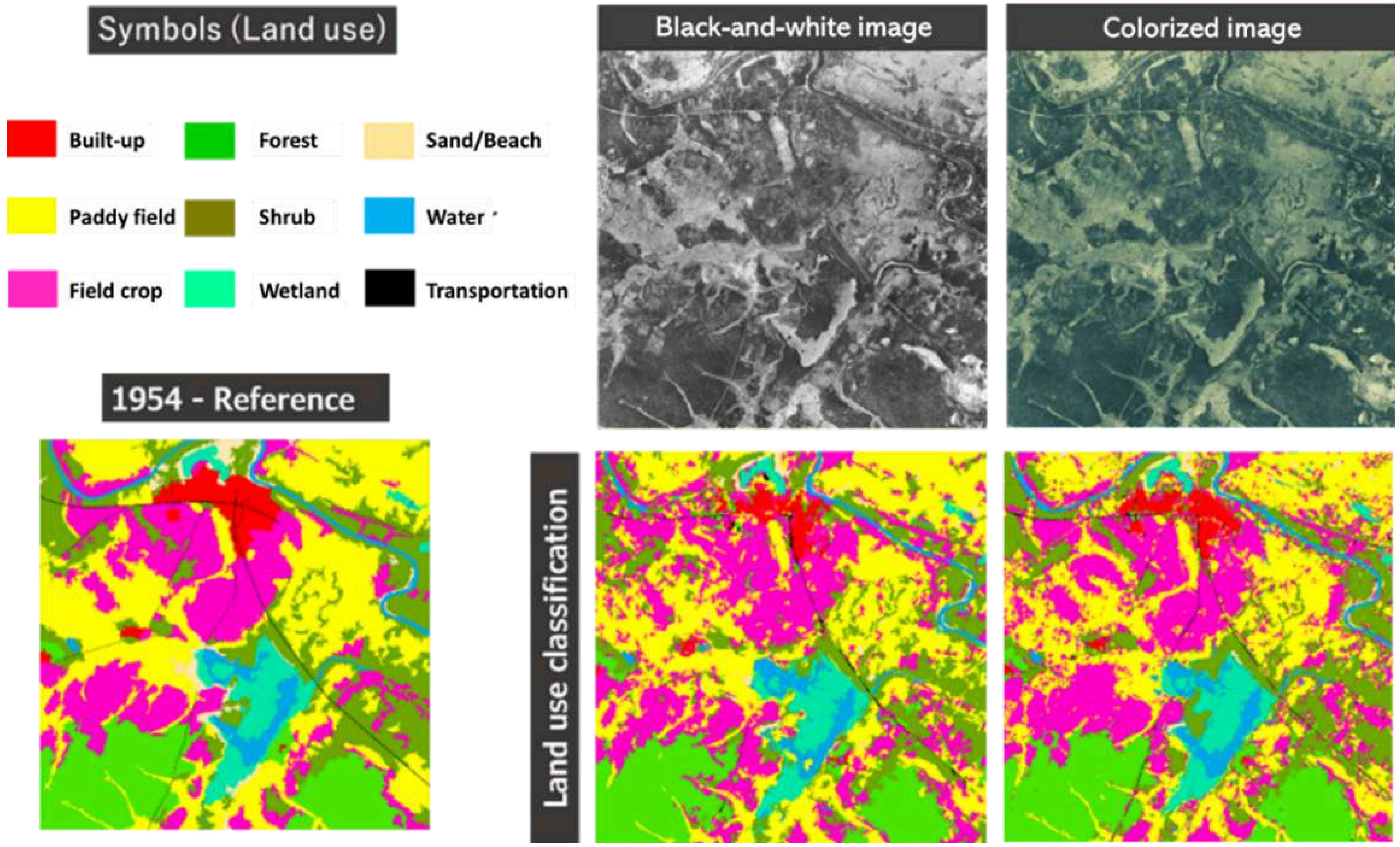
Generative Adversarial Networks (GANs) are a specific type of generative AI model.

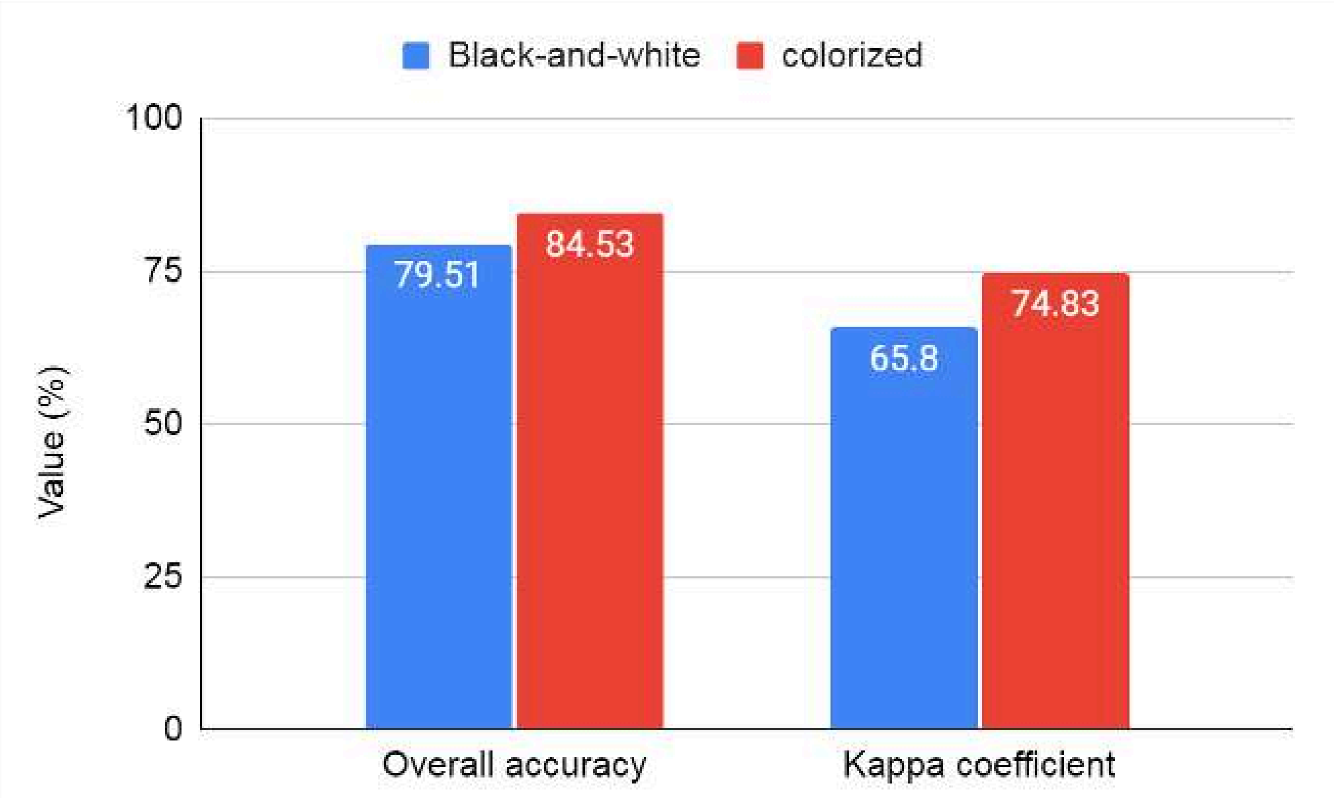
- That uses an ingenious approach to creating new data.
- GANs consist of two neural networks, a generator and a discriminator, that work together.



Source: Aurélien, G. (2017); https://www.neebal.com/blog/generative-ai-vs.-predictive-ai-unraveling-the-distinctions-and-applications

COLORIZATION OF BLACK-AND-WHITE AERIAL PHOTOGRAPHS USING DEEP LEARNING FOR OBJECT-BASED IMAGE ANALYSIS LAND USE CLASSIFICATION



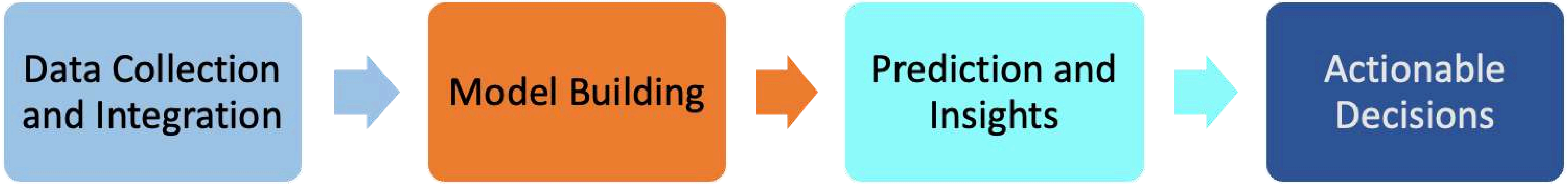


➤ ***Figure 7*** *Shows the results of land use classification both in black-and-white and colorized images and displays overall accuracy and kappa coefficients as bar chart.*

The study found that **colorized image** generated using deep learning techniques **outperformed grayscale images** in land use classification using **object-based image analysis (OBIA)** technique, with a significant improvement in overall accuracy of **+5.02%**

AI: Predictive analytics

- Predictive analytics is a powerful tool that utilizes artificial intelligence (AI), machine learning (ML).
- Big Earth Data to extract insights from historical data and make predictions about future outcomes.
- Here's how they work together:
  - Big Earth Data
  - Machine Learning
  - AI



HANDS ON

https://github.com/thanthamky/rs-lc-dl

Local Running

CONDA

Cloud Running

Workshop Materials

Main Repository

https://github.com/thanthamky/rs-lc-dl

[Chapter 1 : Python and Algebra](https://colab.research.google.com/github/thanthamky/rs-lc-dl/blob/main/1_Python_Math.ipynb)
https://colab.research.google.com/github/thanthamky/rs-lc-dl/blob/main/1_Python_Math.ipynb

[Chapter 2 : Raster EDA and Modeling](https://colab.research.google.com/github/thanthamky/rs-lc-dl/blob/main/2_EDA_Modeling.ipynb)
https://colab.research.google.com/github/thanthamky/rs-lc-dl/blob/main/2_EDA_Modeling.ipynb

[Chapter 3 : Deep Learning and Model Training](https://colab.research.google.com/github/thanthamky/rs-lc-dl/blob/main/3_Deep_Model_Training.ipynb)
https://colab.research.google.com/github/thanthamky/rs-lc-dl/blob/main/3_Deep_Model_Training.ipynb

[Chapter 4 : Inferencing and Application](https://colab.research.google.com/github/thanthamky/rs-lc-dl/blob/main/4_Fine-tuning_Inferencing.ipynb)
https://colab.research.google.com/github/thanthamky/rs-lc-dl/blob/main/4_Fine-tuning_Inferencing.ipynb

[Side Tutorial : Matplotlib](https://colab.research.google.com/github/thanthamky/rs-lc-dl/blob/main/matplotlib_tutorial.ipynb)
https://colab.research.google.com/github/thanthamky/rs-lc-dl/blob/main/matplotlib_tutorial.ipynb

[Side Tutorial : Pytorch](https://colab.research.google.com/github/thanthamky/rs-lc-dl/blob/main/pytorch_tutorial.ipynb)
https://colab.research.google.com/github/thanthamky/rs-lc-dl/blob/main/pytorch_tutorial.ipynb