EARTH OBSERVATION RESEARCH

Julius-Maximilians-UNIVERSITÄT WÜRZBURG

Supported by:

Federal Ministry for the Environment, Nature Conservation and Nuclear Safety

Federal Agency for Nature Conservation

EOCap 4Africa

EOCap4Africa

9 Raster Analysis

c) Land Cover Classification

Julius-Maximilians-UNIVERSITÄT WÜRZBURG

DLR

MARTIN-LUTHER-UNIVERSITÄT HALLE-WITTENBERG

UNIVERSITY OF BOTSWANA

INES Ruhengeri Institute of Applied Sciences

Learning Objectives

Understand the importance of land cover classification

Learn the differences between supervised and unsupervised classification

Explore common classification algorithms used in remote sensing

Identify the advantages and limitations of each method

Land Cover Classification

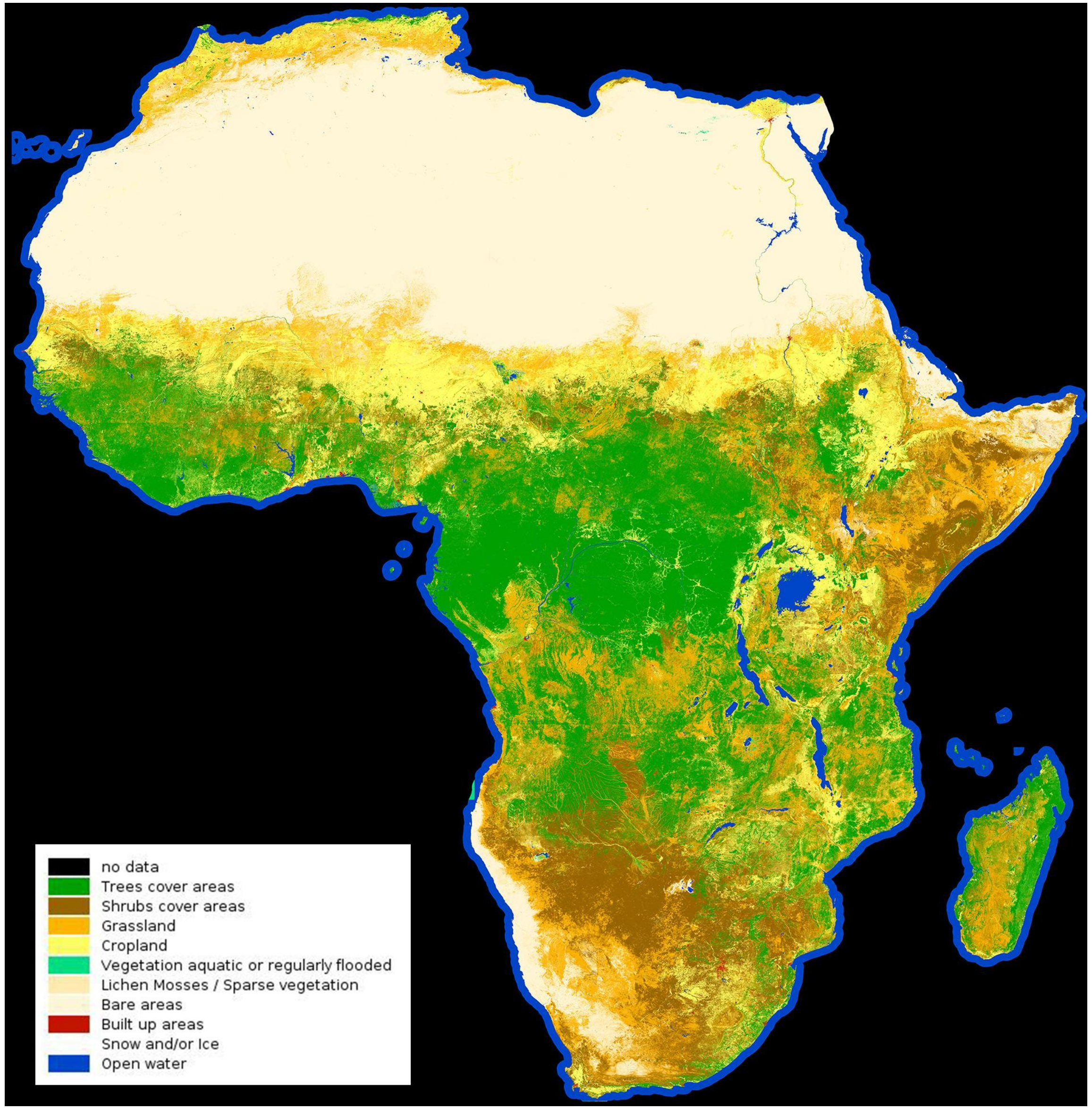
Definition

The process of categorizing and mapping the Earth's surface into distinct land cover types using remote sensing imagery and classification algorithms, which analyze the spectral, temporal, and spatial characteristics of different surfaces.

Land Cover Classification

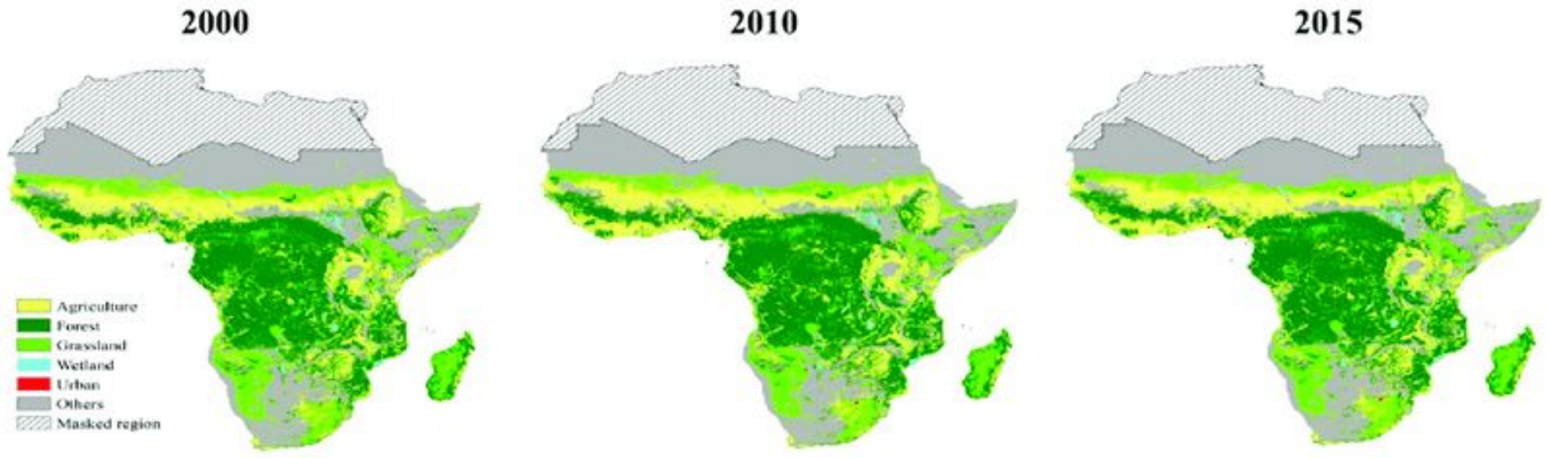
Usage

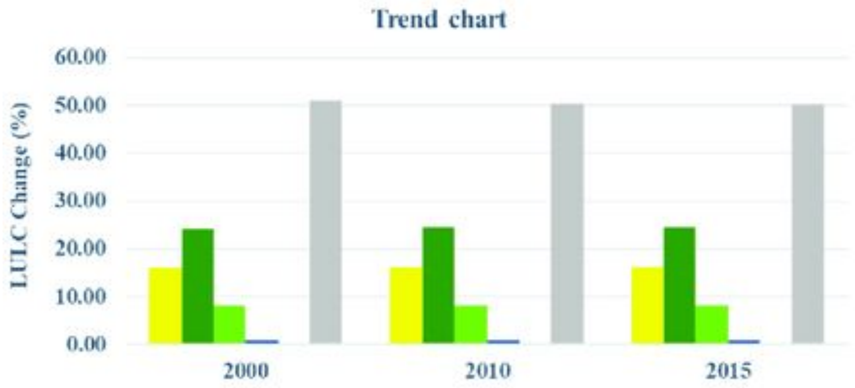
- Environmental monitoring
- Urban planning
- Agriculture and forestry
- Disaster management



(ESA 2017)

Land Cover Classification - Timeseries





	2000	2010	2015
Agriculture	15.86	16.09	16.13
Forest	24.18	24.48	24.50
Grassland	8.13	8.12	8.13
Wetland	0.83	0.84	0.84
Urban	0.08	0.13	0.16
Others	50.91	50.35	50.25

(Chiaka/Zhen 2021)

Can you think of other use cases of LCC timeseries analysis?

Land Cover Classification Methods

Supervised Classification
Requires training data
(user-defined)

Unsupervised Classification
Clusters pixels based on statistical
patterns

Supervised Classification

- User provides labeled training data (e.g., selecting known land cover types).
- Algorithm learns from these samples to classify the entire image.

Examples

- Maximum Likelihood Classification (MLC)
- Support Vector Machines (SVM)
- Random Forest (RF)
- Artificial Neural Networks (ANN)

Supervised Classification

Advantages

- High accuracy if good training data is available
- Suitable for complex land cover types
- Can incorporate expert knowledge

Disadvantages

- Requires high-quality training data
- Time-consuming data collection
- Performance depends on algorithm choice

Supervised Classification – Maximum Likelihood

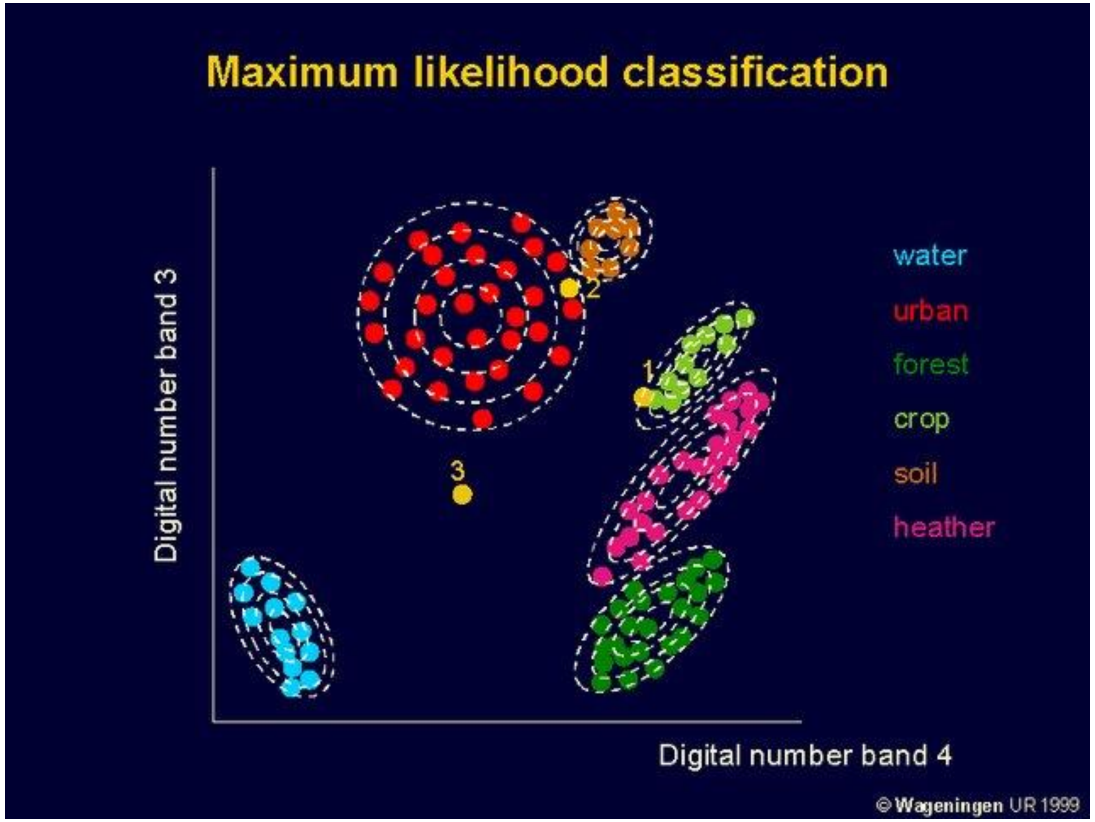
Definition

- A **probabilistic classifier** that assigns each pixel to the class with the highest probability
- Assumes that the **data follows a normal (Gaussian) distribution**
- Uses statistical parameters from training data (mean, variance, covariance)

Supervised Classification – Maximum Likelihood

How MLC Works

1. Estimates probability density functions for each class
2. Assigns pixels to the class with the highest probability
3. Uses a Bayesian decision rule to minimize classification errors



(Vahidi et al. 2023)

Supervised Classification – Maximum Likelihood

Advantages

- Well-established, widely used in remote sensing
- Provides probabilistic confidence levels
- Works well if data follows a normal distribution

Disadvantages

- Assumes normality, which may not always be true
- Performance depends on the quality of training data
- Struggles with highly heterogeneous land cover types

Supervised Classification – Random Forest

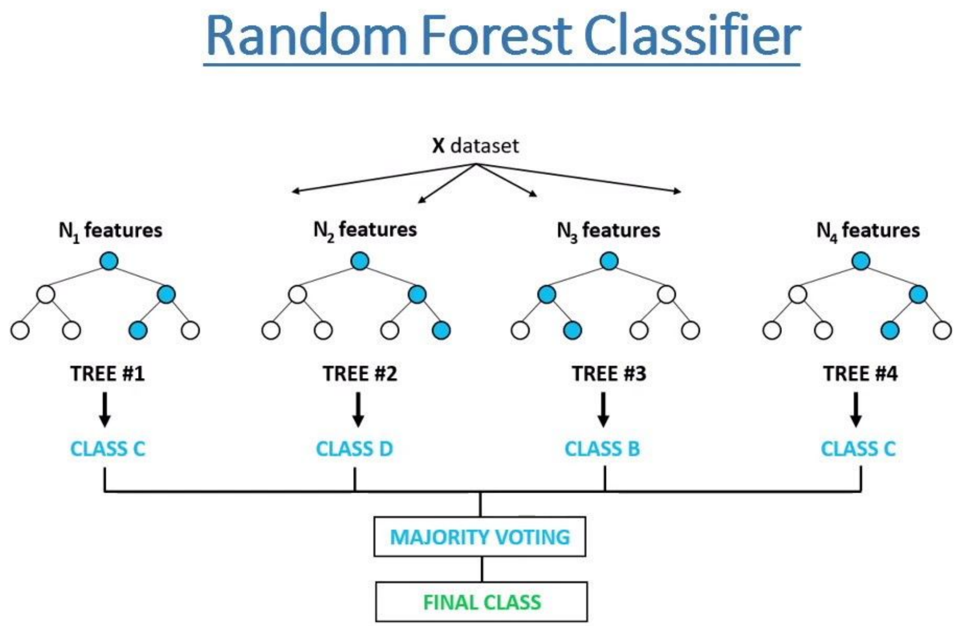
Definition

- A **machine learning algorithm** based on multiple decision trees
- Uses an **ensemble approach**, where each tree votes on the final classification
- Works well with both numerical and categorical data

Supervised Classification – Random Forest

How RF Works

1. Creates multiple decision trees from random subsets of training data
2. Each tree classifies pixels independently
3. The final classification is determined by a majority vote



(Khushvaktov 2023)

Supervised Classification – Random Forest

Advantages

- Handles large datasets with high accuracy
- Does not assume a normal distribution
- Resistant to overfitting due to ensemble averaging

Disadvantages

- Computationally intensive, especially for large images
- Requires fine-tuning (e.g., number of trees, depth of trees)
- Can be slower than simpler classifiers

Unsupervised Classification

- No prior training data required
- Algorithm automatically groups pixels into spectral clusters
- User assigns classes to clusters after classification

Examples

- K-Means Clustering
- ISODATA (Iterative Self-Organizing Data Analysis)

Unsupervised Classification

Advantages

- No need for pre-labeled data
- Faster and more automated process
- Useful for exploratory analysis

Disadvantages

- Results may not match real-world categories well
- User must interpret clusters manually
- Less control over classification rules

Unsupervised Classification – K-Means

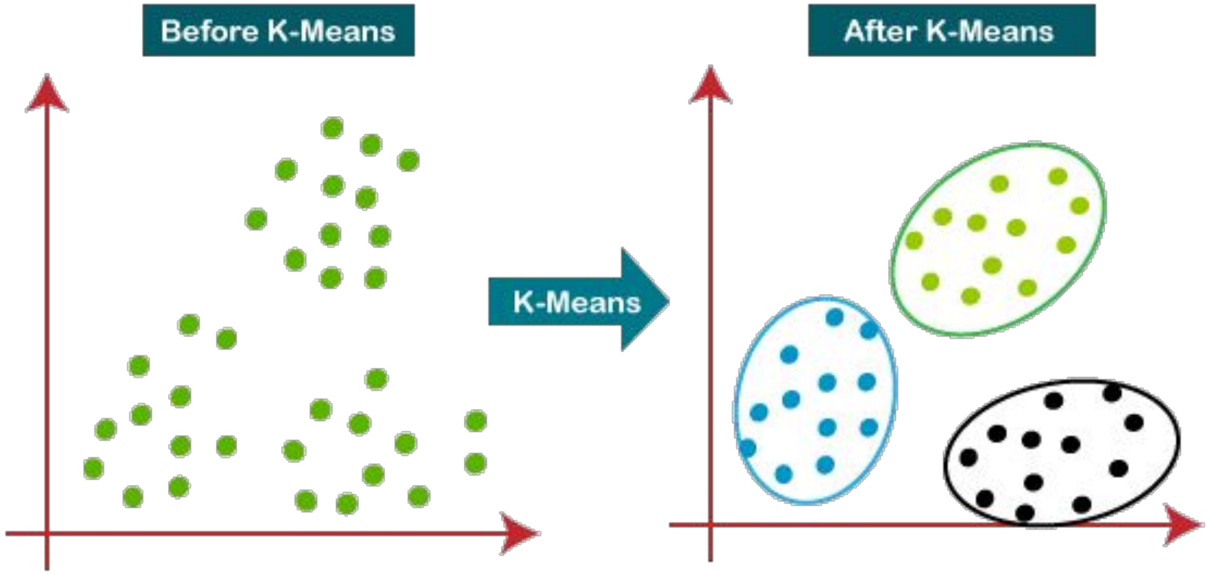
Definition

- An **unsupervised classification method** that partitions pixels into clusters based on similarity
- The number of clusters (**K**) is defined by the user

Unsupervised Classification – K-Means

How MLC Works

1. Select K cluster centers (centroids) randomly
2. Assign each pixel to the nearest centroid based on spectral distance
3. Compute new centroid positions as the average of assigned pixels
4. Repeat the process until clusters stabilize



(Patel 2021)

Unsupervised Classification – K-Means

Advantages

- Fast and efficient for large datasets
- No need for labeled training data
- Works well for initial land cover exploration

Disadvantages

- Requires predefined number of clusters (K)
- May assign similar land cover types to different clusters
- Sensitive to outliers and spectral variability

Supervised vs Unsupervised Classification

Aspect	Supervised Classification	Unsupervised Classification
Training Data	Required	Not required
User Involvement	High (training data & validation)	Lower (post-classification interpretation)
Accuracy	Typically higher	Can be lower due to spectral mixing
Best for	Thematic mapping, detailed studies	Quick analysis, unknown land cover
Example Methods	Maximum Likelihood, Random Forest, SVM	K-Means, ISODATA

Summary & Key Takeaways

Land cover classification is essential for remote sensing analysis

Supervised Classification requires training data and is more accurate

Unsupervised Classification is fully automated but requires interpretation

Choosing the right method depends o**n data availability and project goals**

Sources

European Space Agency (ESA). (2017, October 3). African land cover. ESA Multimedia. Retrieved from https://www.esa.int/ESA_Multimedia/Images/2017/10/African_land_cover

Chiaka, J. C., & Zhen, L. (2021). Land use, environmental, and food consumption patterns in Sub-Saharan Africa, 2000–2015: A review. Sustainability, 13(15), 8200. https://doi.org/10.3390/su13158200

Vahidi, S., Hatamzadeh, V., Afshinfar, A., & Nouri, P. (2023). Monitoring land cover changes in Tehran City over 5 years (2018 to 2022) using remote sensing-based spatial information. Asian Journal of Environment & Ecology, 20(3), 24-35. https://doi.org/10.9734/AJEE/2023/v20i3440

Khushvaktov, F. (2023, August 26). Introduction to random forest classification by example. Medium. Retrieved from https://medium.com/@mrmaster907/introduction-random-forest-classification-by-example-6983d95c7b91

Patel, A. (2021, June 17). K-Means clustering. Medium. Retrieved from https://imakash3011.medium.com/k-means-clustering-ef8e9258d76a

Supported by:

Federal Ministry for the Environment, Nature Conservation and Nuclear Safety

Federal Agency for Nature Conservation

EOCap 4Africa

Thank you for your attention!

Dr. Insa Otte, Hanna Schulten, and colleagues

insa.otte@uni-wuerzburg.de

Julius-Maximilians-UNIVERSITÄT WÜRZBURG

DLR

MARTIN-LUTHER-UNIVERSITÄT HALLE-WITTENBERG

UNIVERSITY OF BOTSWANA

INES Ruhengeri Institute of Applied Sciences