GAFAG

Land cover classification

Definitions and Systems

Alexander Klaus

ESA TRAINING COURSE ON RADAR AND OPTICAL REMOTE SENSING –
Vilnius/LITHUANIA, 3-7 July, 2017

Content

1. Land cover definition
2. Forest, Deforestation, Degradation definition
3. Classification definiton
4. Ratios
   - NDVI
   - NDWI
   - AVI
   - NBRI
5. Classification
   - Unsupervised
   - Supervised
   - Post Processing

Land Cover definition

Land Cover

"**Land cover** *is the observed (bio)physical cover on the earth's surface*"

"**Land use** is characterized by *the arrangements, activities and inputs people undertake in a certain land cover type to produce, change or maintain it*. **Definition of land use in this way establishes a direct link between land cover and the actions of people in their environment.**"

FAO (2000): LAND COVER CLASSIFICATION SYSTEM (LCCS): CLASSIFICATION CONCEPTS AND USER MANUAL

Definition Land Cover

When considering **land cover** in a very pure and strict sense it should be confined to describe vegetation and man-made features. Consequently, areas where the surface consists of bare rock or bare soil are describing *land* itself rather than land *cover*. Also, it is disputable whether water surfaces are real land cover. However, in practise, the scientific community usually describes those aspects under the term land cover

FAO (2000): LAND COVER CLASSIFICATION SYSTEM (LCCS): CLASSIFICATION CONCEPTS AND USER MANUAL

Definition Land Cover

Land cover is the physical material at the surface of the earth. Land covers include grass, asphalt, trees, bare ground, water, etc. There are two primary methods for capturing information on land cover: field survey and thorough analysis of remotely sensed imagery.

(Comber et al. 2005)

Definition Land Use

Land use:

- human modification of natural environment or wilderness into built environment such as fields, pastures, and settlements
- major effect of land use on land cover since 1750 has been deforestation of temperate regions
- More recent significant effects of land use include urban sprawl, soil erosion, soil degradation, salinization, and desertification
- Land-use change, together with use of fossil fuels, are the major anthropogenic sources of carbon dioxide, a dominant greenhouse gas
- It has also been defined as "the total of arrangements, activities, and inputs that people undertake in a certain land cover type"

(FAO, 1997a; FAO/UNEP, 1999).

Land Cover vs. Land Use

Land cover is distinct from land use despite the two terms often being used interchangeably.

Land use is a description of how people utilize the land and socio-economic activity - urban and agricultural land uses are two of the most commonly recognised high-level classes of use.

(Fisher et al. 2005)

Land use and land cover

Land use and land cover

- mapping based on satellite data is a core product in the monitoring and assessment of the environment
- on national and regional level.
- changes over time covered by the observation dates
- shows trends of environmental changes and indicates the areas and changes in biodiversity

Key Definitions

Definitions: Forest, Deforestation, Landuse Categories, Degradation

Forest

Quantitative parameters for Forest Definition, range of values provided through the UNFCCC Marrakesh Accords and the values for the FAO

Key aspects	UNFCCC [1]	FAO [2]
Minimum area	0.05-1.0 ha	> 0.5 ha
Minimum crown cover	10-30 %	> 10 %
Minimum tree height at maturity in situ	2-5 m	5 m

[1] FCCC/CP/2001/13/Add.1, Decision 11/CP.7; Land use, land-use change and forestry, Marrakesh Accords, p. 54.
[2] FAO 2006: Global Forest Resources Assessment 2005. Main Report, www.fao.org/forestry/fra2005

Key Definitions

What is Forest?

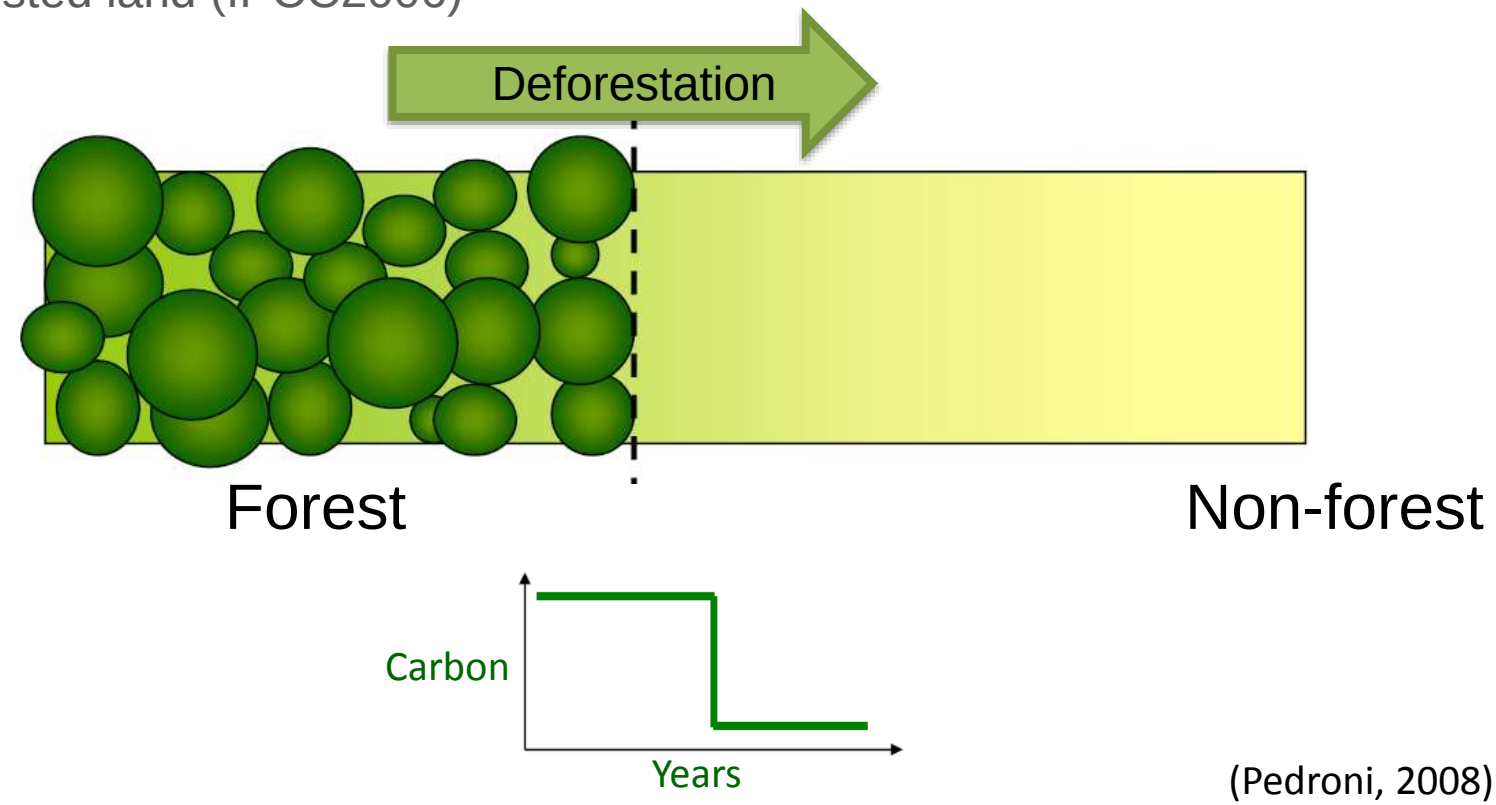
Mixture between cropland and open canopy woodland

Dense miombo type woodland under leaf-on / leaf-off condition

Managed Forest with clear borders and high crown cover density

Deforestation

Deforestation: direct human-induced conversion of forested land to non forested land (IPCC2006)



(Pedroni, 2008)

Key Definitions

Deforestation

The estimation of **deforestation** is determined by the definition of forest versus non-forest land. Under Decision 16/CMP.1 the UNFCCC defined deforestation as **"… the direct, human induced change from forest to non-forested land"**.

Deforestation is given when the forest cover of an area larger than the minimum area of the forest definition decreases below the threshold of the crown cover of the respective forest definition.

Temporarily un-stocked areas, e.g. through selected logging, with an expected re-growth are not considered as deforestation.

Decision 16/CMP.1
http://unfccc.int/resource/docs/2005/cmp1/eng/08a03.pdf#page=3

Key Definitions

Land Use and Change Categories

The applicable land-use conversions resulting from deforestation that take place between data collection intervals are as follows:

FC = Forest Land to Cropland

FG = Forest Land to Grassland

FW = Forest Land to Wetland

FS = Forest Land to Settlements

FO = Forest land to Others

The definitions of the land-use categories for greenhouse gas inventory reporting are provided by the IPCC Guidelines (2006), Chapter 3

Key Definitions

Degradation

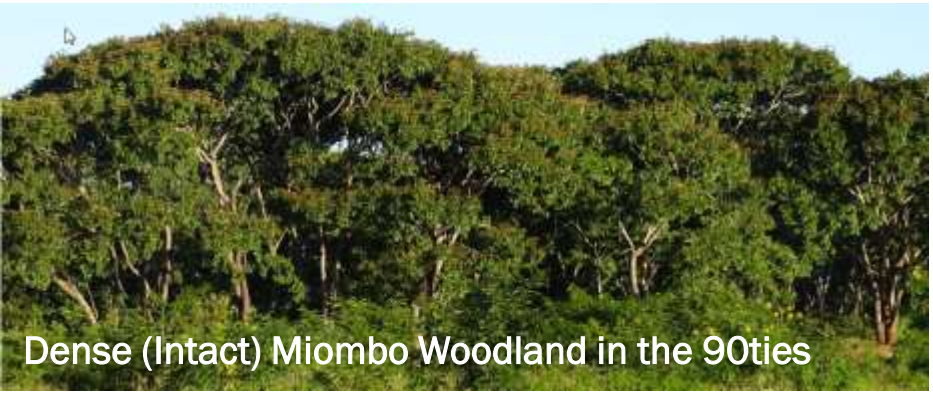
“A direct human-induced long-term loss (persisting for X years or more) of at least Y% of forest carbon stocks[and forest values] since time T and not qualifying as deforestation.”

IPCC 2003: Definitions and Methodological Options to Inventory Emissions from Direct Human-induced Degradation of Forests and Devegetation of Other Vegetation Types. The Intergovernmental Panel on Climate Change (IPCC), 2003. P. 14-15.

GOFC-GOLD, 2012: monitoring area changes of degradation is a bit more complex as opposed to permanent land use changes

Main causes of forest degradation in the SADC region are exploitation of fuel wood, subsistence agriculture, fire, selective exploitation of commercial wood

Example Forest Degradation in Malawi

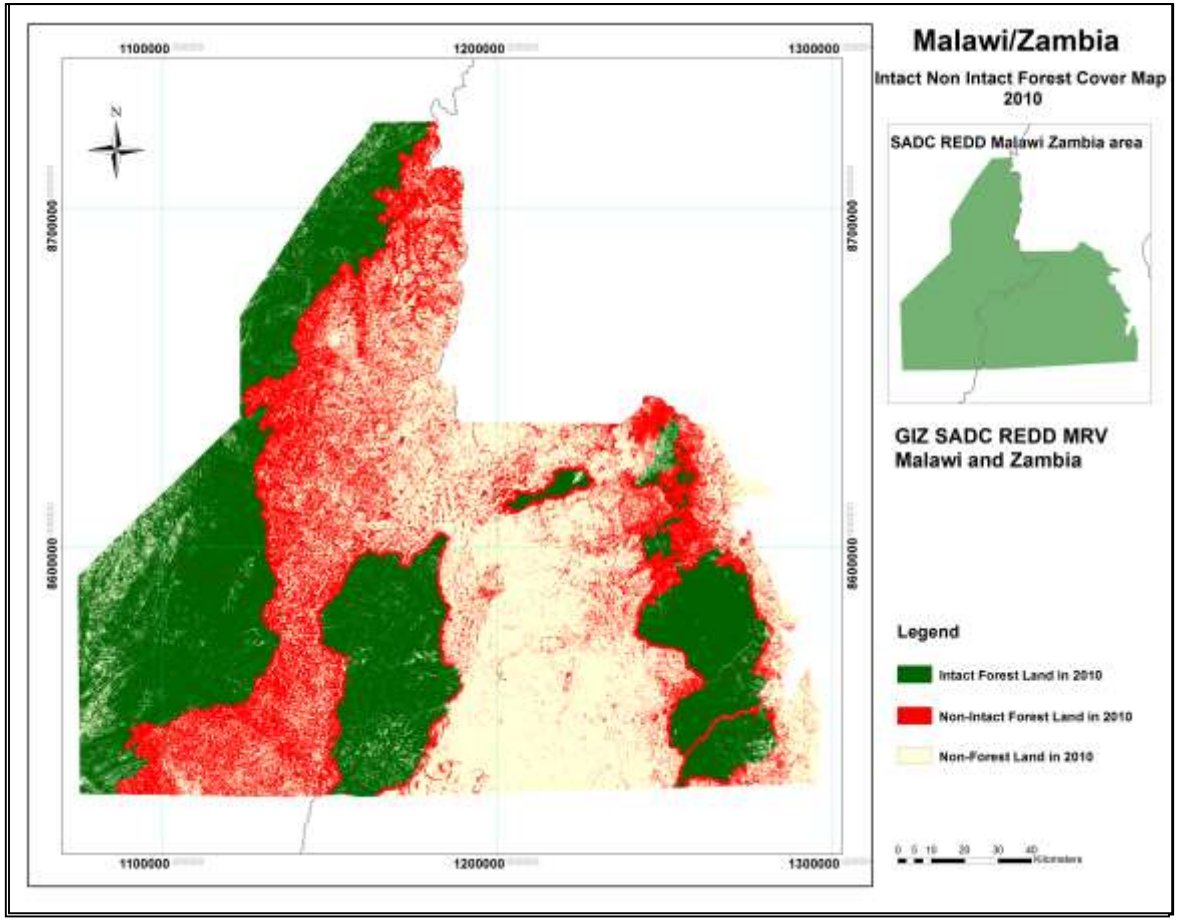

Dense (Intact) Miombo Woodland in the 90ties


Open (Non-intact) Miombo Woodland today

Extraction of trees for fuel wood consumtion and on-site production of charcoal….

Leads to degradation of Miombo woodlands in very short time, decline of biodiversity, finally to deforestation

Expansion of Degradation Area from 1990 to 2010



Key Definitions

Definitions for Intact and Non-Intact Forests:

Intact Forest: fully stocked (forest with a tree cover between 10% and 100% according to the FAO Forest Definition but must be undisturbed, i.e. there has been no timber extraction (GOFC GOLD 2012)

Non-Intact Forest: not fully stocked but tree cover must still be higher than 10% to qualify as forest according to the applied FAO Forest Definition but the forest has undergone some level of exploitation or canopy degradation (GOFC GOLD 2012).

The increase of the non-intact forest area e.g. between 2000 and 2010 will provide the area - of degradation.

Definition Classification

Hierarchical versus non-hierarchical systems

- Classification systems come in two basic formats, *hierarchical* and *non-hierarchical*.
- Hierarchical means that classes can be accumulated to a generalized class, which includes all feature of the original classes; respectively a common primary class can be divided into secondary classes
- Most systems are hierarchically structured because such a classification offers more consistency owing to its ability to accommodate different levels of information, starting with structured broad-level classes, which allow further systematic subdivision into more detailed sub-classes.
- At each level the defined classes are mutually exclusive.
  - At the higher levels of the classification system few diagnostic criteria are used,
  - whereas at the lower levels the number of diagnostic criteria increases.
  - Criteria used at one level of the classification should not be repeated at another

FAO (2000): LAND COVER CLASSIFICATION SYSTEM (LCCS): CLASSIFICATION CONCEPTS AND USER MANUAL

What is Classification?

Classification is the process of

- sorting pixels into a finite number of individual classes, or categories of data, based on their data file values.

and

- if a pixel satisfies a certain set of criteria, then the pixel is assigned to the class that corresponds to that criteria.

Why do we need classification results?

Depending on the aim / used satellite data / ancillary data

- receive the information about the Land Cover and / or Land Use of the area of interest.

Depending on the up-to-dateness of satellite data used

- receive Land Cover and / or Land Use information at a specific time or time span

Where / How do we use Land Cover / Land Use data?

- to monitor environmental changes
- as (basis) information for long-term and short-term land use plans
- for planning and management of land resources
- Land use information can be used to develop solutions for natural resource management issues
- quantifying linkages and feedbacks between LULCC, climate change, and other human and environmental components
- provide policy makers with relevant information for effective agro-environment / land use policy

Input Data

selection of input data always in relation to the envisaged:

- scope of the project
  - Base Map (status quo), Change Map (monitoring)
- mapping scale
  - large scale, small scale, entire country, selected areas
- Land Use / Land Cover classes
  - complete LU/LC, parts of the environment
- level of detail
  - Base level to very detailed
- etc

Input Data

- Orthorectified Satellite Images
  - Sentinel-2
  - Landsat (5/7/8)
  - SPOT
  - ASTER
  - DMC
  - RapidEye
- Digital Elevation Model
  - SRTM 3 arcsec
- Ratios
  - NDVI, Water-Ratio
- Ancillary Data
  - Vector Layers: AOI, settlements, roads, pipeline..

Preparations

- Develop Land Use / Land Cover categories that
  - represent areas within the country
  - are consistent to the e.g. FAO or IPCC Guidelines and to requirements to the Kyoto Protocol
  - can be derived from the data available
- Selection of most actual and cloud free areas
  - subset images with generated AOIs
  - calculate statistics (16 bit data -> “direct”)
  - enhance images for better visual interpretation
- (probably!) for large areas cut images in smaller parts (strata) that represent homogenous geographical units
  - e.g. separation of vegetated / non vegetated areas
- calculation of Ratios

Ratios

„artificial images“

- Generation of new images by dividing grey values of two channels
  - Making use of the different reflection factors (high/low reflection)
  - Brings out or highlight specific characteristics of the image that couldn’t be observed in the original data
- Channels usually used:
  - Visible (VIS) vs Near Infrared (NIR) or Short Wave Infrared (SWIR)
- Ratios as **separate or additional** image **layer** in classification processes:
  - NDVI
  - Ratio to highlight water bodies “NDWI”
- Used to **mask satellite images** in order to reduce the area to classify

NDVI

Normalized Difference Vegetation Index (NDVI)

Formula

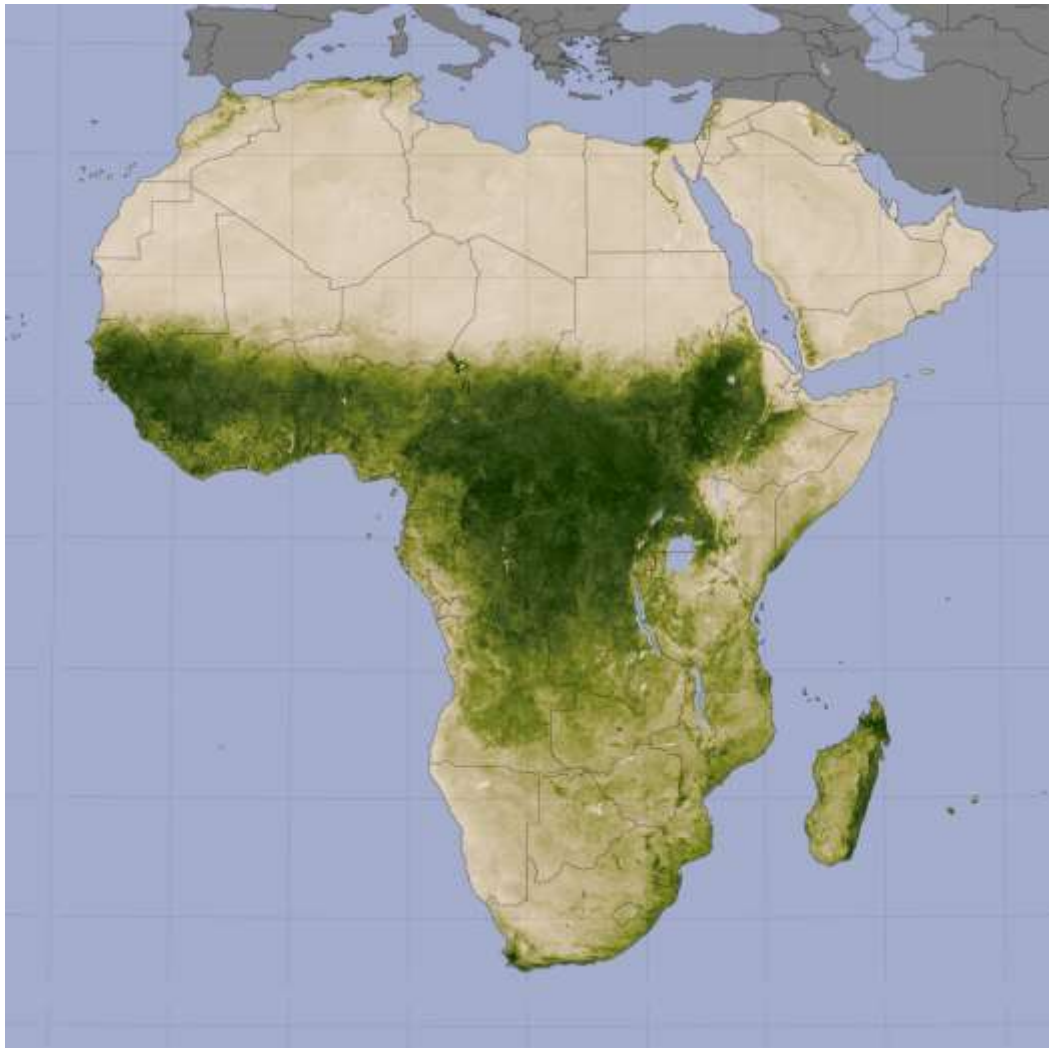
NDVI = (near IR band - red band) / (near IR band + red band)

- Index values can range from -1 to 1 but vegetation values usually range from **0.1 to 0.7**

- The **Red Edge** effect is used to detect vegetation in satellite images
  - High / Low grey values >> Vegetation / No Vegetation
- Provides information about the vitality and density of vegetation
- Enables to quantify land cover/land use
- Used as indicator/identifier for vegetation in classification processes

Ratios

NDVI for africa September 2004



NDWI

Normalized Difference Water Index (NDWI)

- The different reflection factors of water surfaces is used to detect such areas in satellite images
  - Very high / Low grey values >> Water bodies / No Water bodies

- Used as indicator/identifier for water bodies in classification processes

Ratios

The following table the near-infrared and green spectral bands from **various satellites** (Landsat 8 OLI, Landsat 5 TM, Sentinel 2 MSI) to calculate the NDWI with the following formulas (one for each satellite):

(Green band - SWIR band) / (Green band + SWIR band)

$NDWI_{L8}$ =(band 3 – band5) / (band 3 + band5) Landsat-8

$NDWI_{S2}$ =(band 3 – band8) / (band 3 + band8) Sentinel-2

AVI

Advanced Vegetation Index (AVI)

Advanced Vegetation Index (AVI) is a numerical indicator, **similar to NDVI**, that uses the red and near-infrared spectral bands. Like NDVI, AVI is used in vegetation studies **to monitor crop and forest variations over time**. Through the multi-temporal combination of the AVI and the NDVI, users can discriminate different types of vegetation and extract phenology characteristics/parameters.

Ratios

The AVI uses the red and near infrared spectral bands from various satellites (Landsat 8 OLI, Landsat 5 TM, Sentinel 2 MSI) to calculate the AVI with the following formulas (one for each satellite):

$$AVI_{L8} = (band5 * (1 - band\ 4) * (band\ 5 - band4))^{1/3}$$ Landsat-8

$$AVI_{S2} = (band8 * (1 - band\ 4) * (band\ 8 - band4))^{1/3}$$ Sentinel-2

NBRI

Normalized Burned Ratio Index (NBRI)

Forest fires are a severe manmade or natural phenomena that destroy natural recourses, live stock, unbalances the local environments, release huge amount of Green House Gases etc. The scientific community has introduced the Normalized Burned Ratio Index (NBRI) to estimate the severity of fires, mainly in forested areas. NBRI takes advantage of the near infrared and short wave infrared spectral bands, which are sensitive in vegetation changes, to detect burned areas and monitor the recovery of the ecosystem.

The NBRI must be used at least in pairs in order to extract information. One NBRI image before the fire event and one or more NBRI images after the fire event. The difference among these NBRI images will highlight the burned areas and can be used to monitor the behavior of the ecosystem as the time passes.

Ratios

In the following tables the near infrared and short wave infrared spectral bands from various satellites (Landsat 8 OLI, Landsat 5 TM, Sentinel 2 MSI) to calculate the NBRI with the following formulas (one for each satellite):

(SWIR1 band – SWIR2 band) / (SWIR1 band + SWIR2 band)

$NBRI_{L8}$ = (band 5 – band7) / (band 5 + band7) Landsat-8

$NBRI_{S2}$ = (band 8 – band12) / (band 8 + band12) Sentinel-2

Ratios

Generate the differenced (or delta) NBR:

$$dNBR = NBRprefire - NBRpostfire$$

This isolates burned from unburned areas, provides a quantitative measure of absolute change in NBR. Practical data range ≈ −500 to + 1300 when scaled by 10^3.
Apply a linear grayscale to the data range of −800 to 1100, and study this image carefully.
Define the burn perimeter using combined automated and on-screen digitizing from the dNBR.
Make an initial cut at severity thresholds in false color. A seven-tiered configuration may be useful. Ordinal severity levels and example range of dNBR (scaled by 10^3) are shown:
SEVERITY LEVEL ΔNBR RANGE

Enhanced Regrowth,	*High -500 to -251*
Enhanced Regrowth,	*Low -250 to -101*
Unburned	*-100 to +99*
Low Severity	*+100 to +269*
Moderate-low Severity	*+270 to +439*
Moderate-high Severity	*+440 to +659*
High Severity	*+660 to +1300*

(These value ranges are flexible; scene-pair dependent; shifts in thresholds +100 points are possible. dNBR less than about -550, or greater than about +1350 may occur, but usually are not considered burned. Rather, they likely are anomalies caused by miss-registration, clouds, or other factors not related to real land cover differences.)

Classification

Classification Methods

- Unsupervised Classification
- Supervised Classification

often a ***combination of the methods*** is used

Complete and / or parts of images are classified (AOI)

Supervised Classification

- Closely controled by the operator
- Classes and their spectral features are fixed **before** classification
- Generation of signatures out of “training areas“ which represent certain classes
- These training areas are created with help of ground truth information or knowlegde of terrain (**homogeneous** land use/land cover units)
- Pixels of an image are sorted into classes based on defined signatures by use of classification decision rules

Unsupervised Classification

Unsupervised Classification:

K-Means (KM) Cluster Analysis

Algorithm

The k-means clustering tool implemented in SNAP is capable of working with arbitrary large scenes.

Given the number of clusters k, the basic algorithm implemented in SNAP is:

Randomly choose k pixels whose samples define the initial cluster centers.

Assign each pixel to the nearest cluster center as defined by the Euclidean distance.

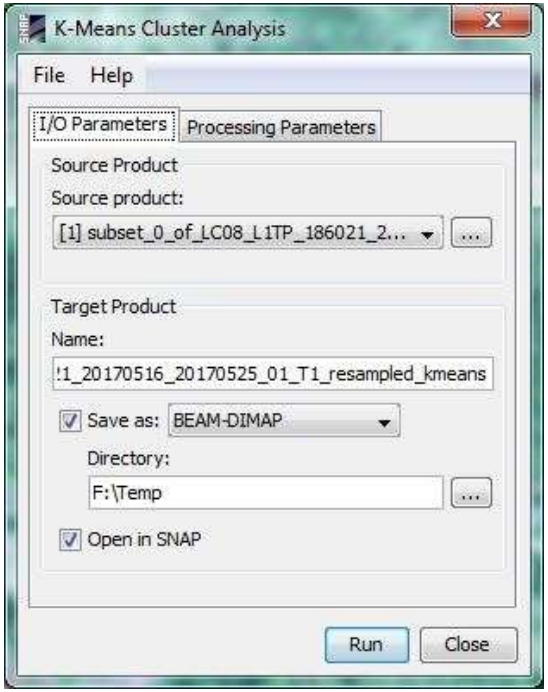
Recalculate the cluster centers as the arithmetic means of all samples from all pixels in a cluster.

Repeat steps 2 and 3 until the convergence criterion is met.

K-Means (KM) Cluster Analysis

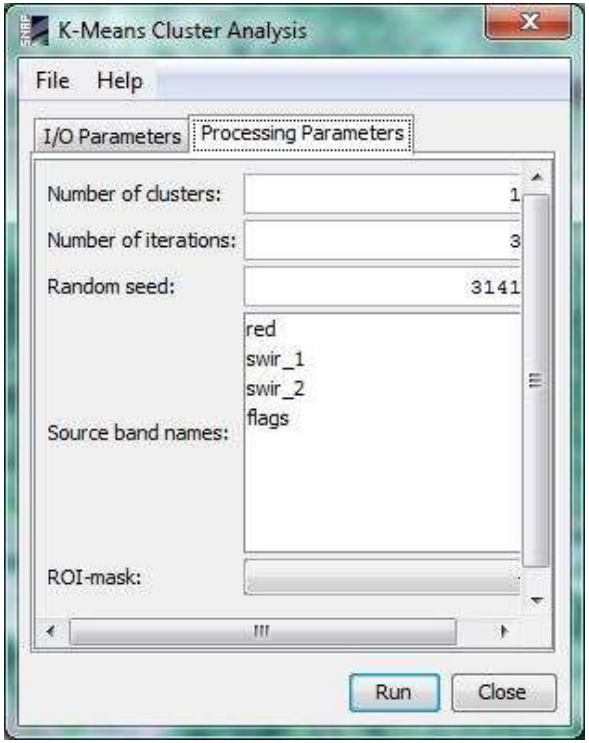
User Interface

The k-means (KM) cluster analysis tool can be invoked from the Sentinel Toolbox tool menu by selecting the K-Means Cluster Analysis command in the Image Analysis submenu. In the command line it is available by means of the Graph Processing Tool gpt which is located in the Sentinel Toolbox bin directory. Please type gpt KMeansClusterAnalysis -h for further information. Selecting the K-Means Cluster Analysis command from the Sentinel Toolbox tool menu pops up the following dialog:



K-Means (KM) Cluster Analysis

Processing Parameters Panel



Number of clusters: Use this field to specify the number of clusters. The default is 14 clusters.
Number of iterations: Use this field to specify the maximum number of iterations to be carried out. The default is 30 iterations. The cluster analysis stops when the maximum number of iterations is exceeded.
Random seed: The KM algorithm starts with a pseudo-random distribution of initial clusters. The random seed initializes the pseudo-random number generator, which is used to generate the initial clusters. By changing the random seed, you can start with different initial clusters. Any positive integral number will be a perfect random seed. The default seed is 31415.
Source band names: Use this field to specify the names of the source bands. Press the control key while selecting or deselecting individual bands.
ROI-Mask: Use this field to restrict the cluster analysis to a region-of-interest (ROI). The combo box allows to select the band which provides the ROI.

Supervised Classification

Supervised Classification Instructions

Workflow

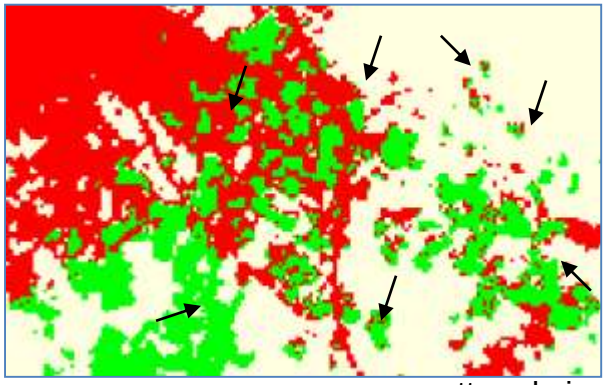
Training Stage

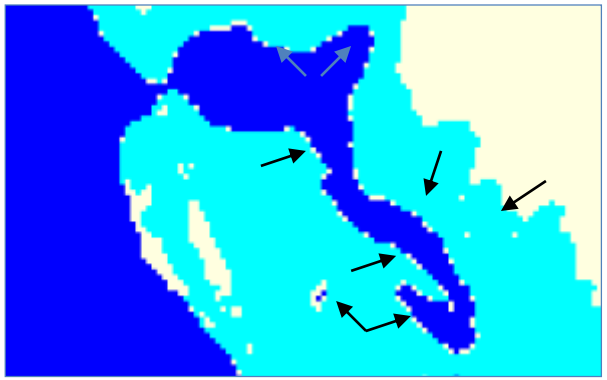
- The analyst identifies representative training areas and
- develops a numerical description of the spectral attributes of each land cover type of interest in the scene
- Workflow:
  - Open multispectral image you want to classify
  - Open ROI-Tool
  - Define proper Training Areas
  - Refine training areas with „grow“-function
  - Chose the classifier of your choice
  - Start the classifier and chose the classes defined by training areas you want to classify

Post-Classification

Post-Classification Filtering

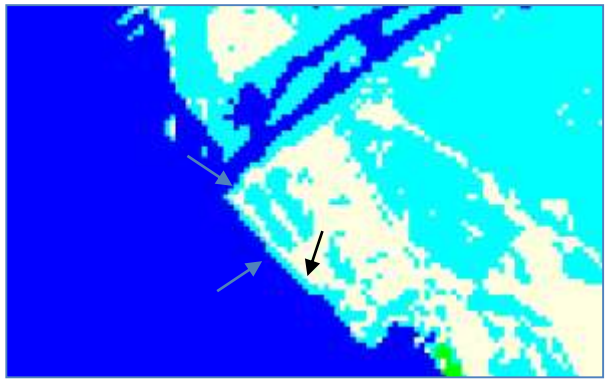
- Pixel-based classifications always contain a certain amount of „background-noise“. Often the result of incorrect classification of individual pixels or small groups of pixels or they are below the specificified minimum mapping unit (MMU)





scattered pixels inside /outside a specific class

- “Errors” occurs also on boundaries between two different land use classes, where pixels with mixed signatures or classes with great spectral variability exist



„boundary class“ between other classes