



Image classification

Mário Caetano

September 4th, 2007

Lecture D2L4



Goals

1 From data to information: presentation of different mapping approaches

2 Most common problems in image classification and how to solve them

e.g. mixed pixel problem, lack of normality of the training data, Hughes phenomenon

3 Most important advances in satellite image classification

**e.g. from pixel to object, from hard to soft classifiers,
from parametric to non-parametric classifiers**

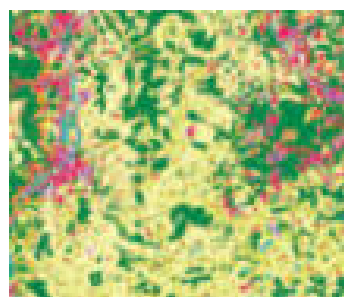


Land information extraction from satellite images

Map of
categorical
variables



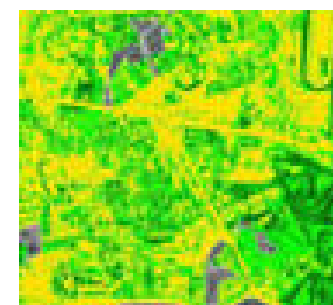
Map of thematic
classes



Land cover maps
Burned area maps
Flooded maps
Agriculture maps
Forest maps

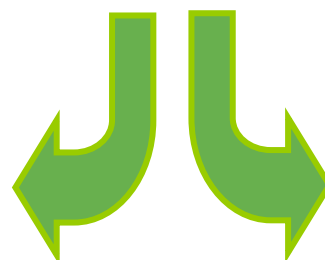
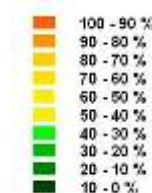
**Thematic
remote sensing**

Image classification



Leaf area index
Biomass
Tree volume

Map of
continuous
variables



**Quantitative
remote sensing**

Modelling

The traditional approach for land cover mapping

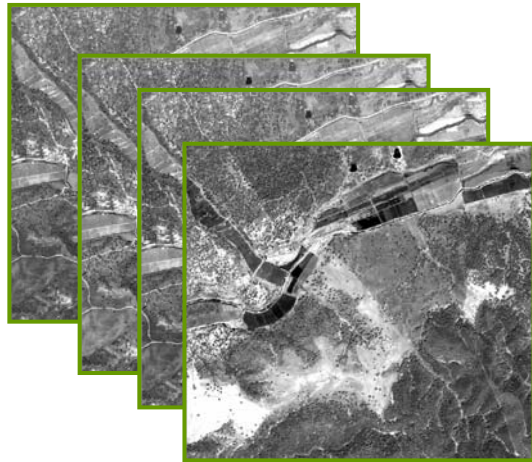
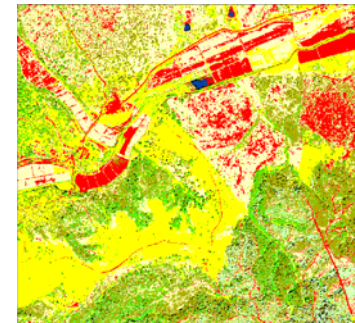


Image classification
at pixel level



**Map of categorical
classes**



Recent advances in satellite image classification

1. Development of **components of the classification algorithm**, including training, learning and approaches to class separation

e.g. artificial neural networks, decision trees

2. Development of **new systems-level approaches** that augment the underlying classifier algorithms

e.g. fuzzy or similar approaches that soften the results of a hard classifier, multiclassifier systems that integrate the outputs of several classification algorithms

3. Exploitation of **multiple** types of data or ancillary information (numerical and categorical) in the classification process

e.g. use of structural or spatial context information from the imagery, use of multitemporal data, use of multisource data, use of ancillary geographical knowledge in the overall classification system

Source: Wilkinson, 2005



For many years the research emphasis has been on the classification step itself.

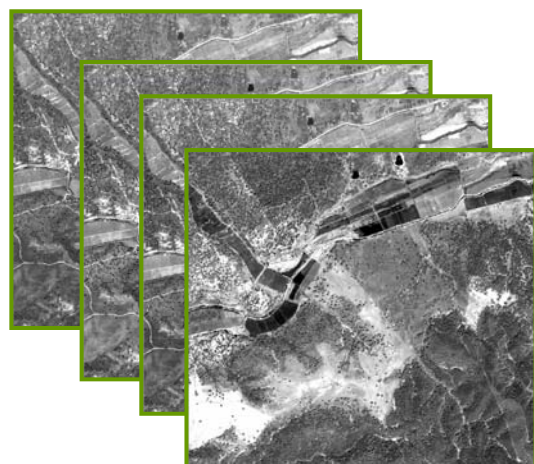


Image classification
at pixel level



**Map of categorical
classes**

Does it satisfy the user needs?

Recent
research

New classification algorithms

A new spatial unit of analysis

Spatial analysis for map generalisation



Redefine the approach
for thematic
information extraction



Thematic information extraction from satellite images

- 1 Definition of the mapping approach *
- 2 Geographical stratification
- 3 Image segmentation
- 4 Feature identification and selection *
- 5 Classification *
- 6 Ancillary data integration
- 7 Post-classification processing
- 8 Accuracy assessment *

* mandatory



Thematic information extraction from satellite images

- 1 **Definition of the mapping approach** *
- 2 **Geographical stratification**
- 3 **Image segmentation**
- 4 **Feature identification and selection** *
- 5 **Classification** *
- 6 **Ancillary data integration**
- 7 **Post-classification processing**
- 8 **Accuracy assessment** *

* mandatory



1. Definition of the mapping approach

The mapping approach has to take into account, e.g.

Characteristics of the satellite data to be used



Technical specifications of the final map (e.g. MMU)



Characteristics of the geographical area to be mapped



Availability of ancillary data



Definition of the spatial
unit of analysis

Decision on stratifying
the study area

Decision on the use of
ancillary data

NLCD = Minimum Mapping Unit



1. Definition of the mapping approach

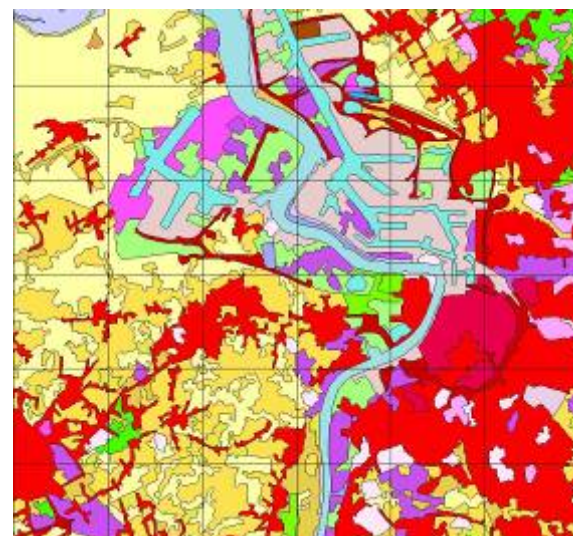
Minimum Mapping Unit (MMU)

The MMU is the smallest area that is represented in a map



In raster maps the MMU usually is the pixel

e.g. in the NLCD 2001 (USA) the MMU is 30x30 m pixel



In vector maps the MMU is the smallest object that is represented in the map

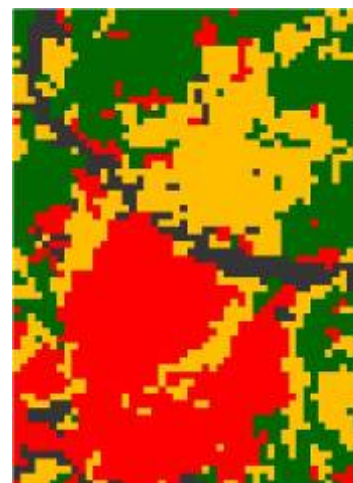
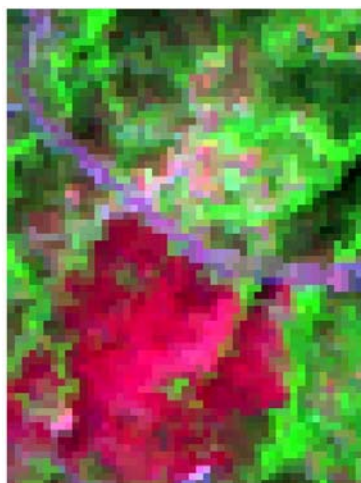
e.g. in the CORINE Land Cover (CLC) maps (from EEA) the MMU is 25 ha



1. Definition of the mapping approach

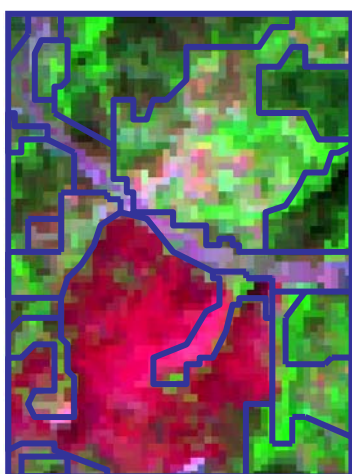
Spatial unit of analysis This is the unit to which the classification algorithms will be applied

Image pixel



Per pixel or sub-pixel classification

Object



Object oriented image classification



1. Definition of the mapping approach

The selection of the **spatial unit of analysis** depends on:

Spatial resolution of the satellite image

Type of thematic information we want to extract, e.g. land cover, land use

Format of the map we want to produce, i.e. vector or raster

Minimum Mapping Unit of the final map

Post-processing tasks that we are planning to apply



1. Definition of the mapping approach

The steps required to information extraction depend on the defined mapping approach:

Map format = raster

MMU = pixel size of input satellite data

Feature selection > Image classification > accuracy assessment

MMU > pixel size of input satellite data

Feature selection > Image classification > post-processing > accuracy assessment

↑
upsaling
↓

Map format = vector

Spatial unit of analysis = image pixel

Feature selection > Image classification > post-processing > accuracy assessment

↑
↓

Generalisation + Raster to vector conversion

Spatial unit of analysis = object

Image segmentation > Feature selection > Image classification > post-processing > accuracy assessment

↑
Generate the objects
↓

↑
Generalisation
↓



Thematic information extraction from satellite images

- 1 Definition of the mapping approach *
- 2 **Geographical stratification**
- 3 Image segmentation
- 4 Feature identification and selection *
- 5 Classification *
- 6 Ancillary data integration
- 7 Post-classification processing
- 8 Accuracy assessment *

* mandatory



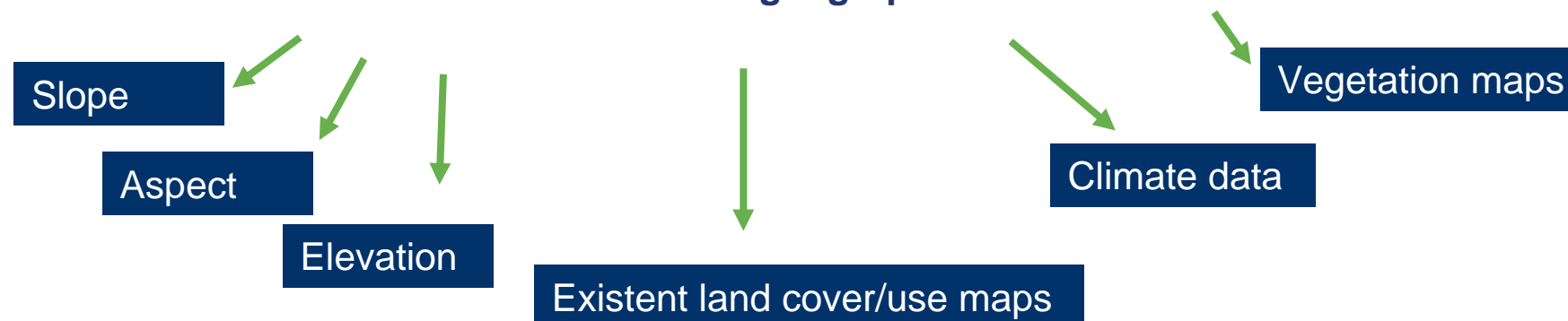
2. Geographical stratification

Geographical stratification – the study area is divided into smaller areas (strata) so that each strata can be processed independently.

Five general concepts are useful in geographical stratification:

- economics of size,
- type of physiography,
- potential land cover distribution,
- potential spectral uniformity,
- edge-matching issues.

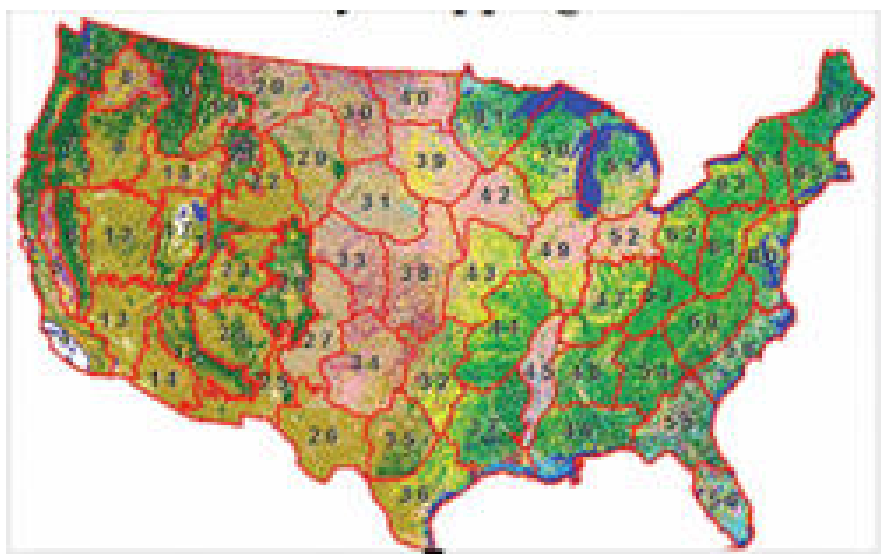
Data that can be used for geographical stratification





2. Geographical stratification

Geographical stratification used on the production of the US National Land Cover Database (NLCD) - 2001



Input data



- 83 Level III ecoregions developed by Omernik
- NLCD 1992
- AVHRR normalized greenness maps

AVHRR - Advanced Very High Resolution Radiometer

Source: Homer et al. (2004)



Thematic information extraction from satellite images

- 1 Definition of the mapping approach *
- 2 Geographical stratification
- 3 **Image segmentation**
- 4 Feature identification and selection *
- 5 Classification *
- 6 Ancillary data integration
- 7 Post-classification processing
- 8 Accuracy assessment *

* mandatory



3. Image segmentation

This step is only required if the spatial unit of analysis is the **object**.

Segmentation is the division of an image into spatially continuous, disjoint and homogeneous regions, i.e. the objects.

Segmentation of an image into a given number of regions is a problem with a large number of possible solutions.

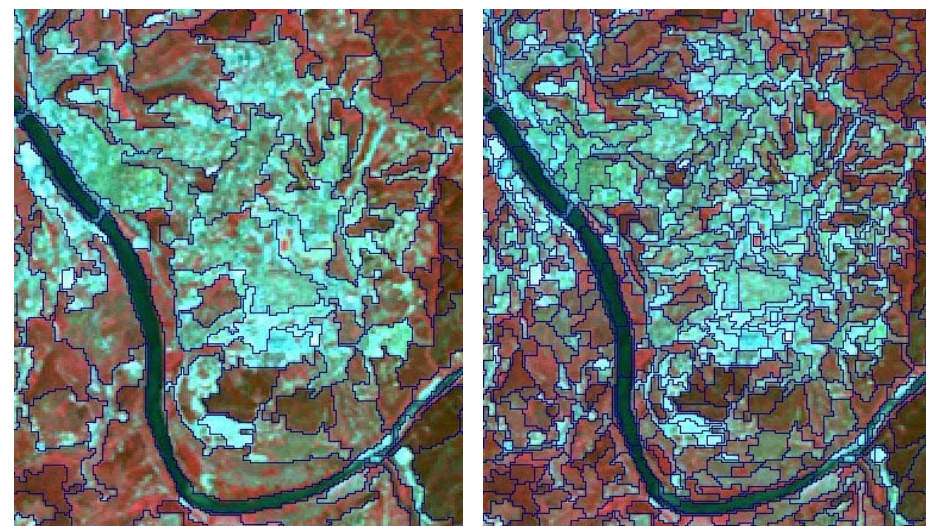
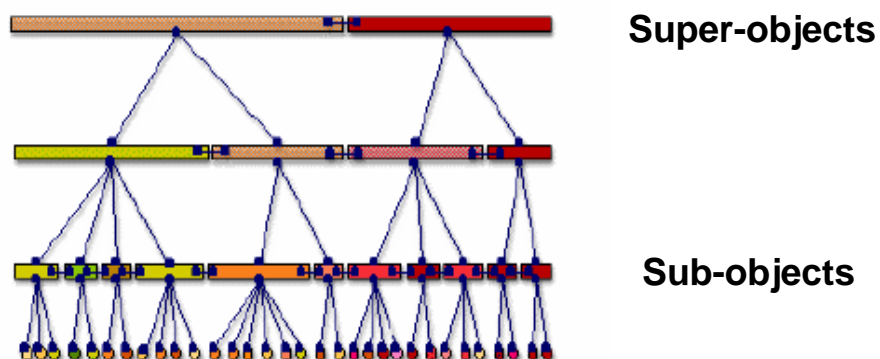


There are no “right” or “wrong” solutions to the delineation of landscape objects but instead “meaningful” and “useful” heuristic approximations of partitions of space.



3. Image segmentation

A type of segmentation that is very common is the **multi-resolution segmentation**, because of its ability to deal with the range of scales within a single image.





Thematic information extraction from satellite images

- 1 Definition of the mapping approach *
- 2 Geographical stratification
- 3 Image segmentation
- 4 **Feature identification and selection** *
- 5 Classification *
- 6 Ancillary data integration
- 7 Post-classification processing
- 8 Accuracy assessment *

* mandatory

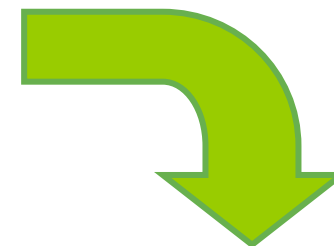


4. Feature identification and selection

What type of features can we use for information extraction?

Should we, for some reason, manipulate the feature space?

How can we select the best features for class discrimination?



Manipulation and selection of features
are used to reduce the number of
features without sacrificing accuracy



4. Feature identification and selection

What type of features can we use for information extraction?

Spectral measurements

1st order measurements

From a single date (Unitemporal approach)

From multiple dates (Multi-temporal approach)

Secondary measurements derived from the image

2nd order measurements

Measurements of the spatial unit being classified

Measurements related to the neighbourhood

Quantification of the spatial variability within the neighbourhood

Texture

Spatial features

Semantic relationships of a spatial unit with its neighbours

Ancillary information

This term is generally used for non-spectral geographical information

Data from images with different characteristics can also be considered as ancillary information. The approaches used for multisensor data may fall within **data fusion**.



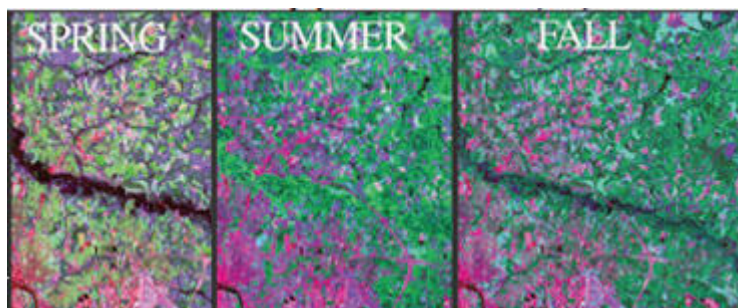
4. Feature identification and selection

What type of features can we use for information extraction?

1st order measurements

Unitemporal approach

Multi-temporal approach



The production of the US National Land Cover Database (NLCD) – 2001 is based on a multi-temporal approach



It helps to discriminate classes with different phenology

Irrigated and rain fed agriculture

Permanent and deciduous forests

Source: Homer et al. (2004)

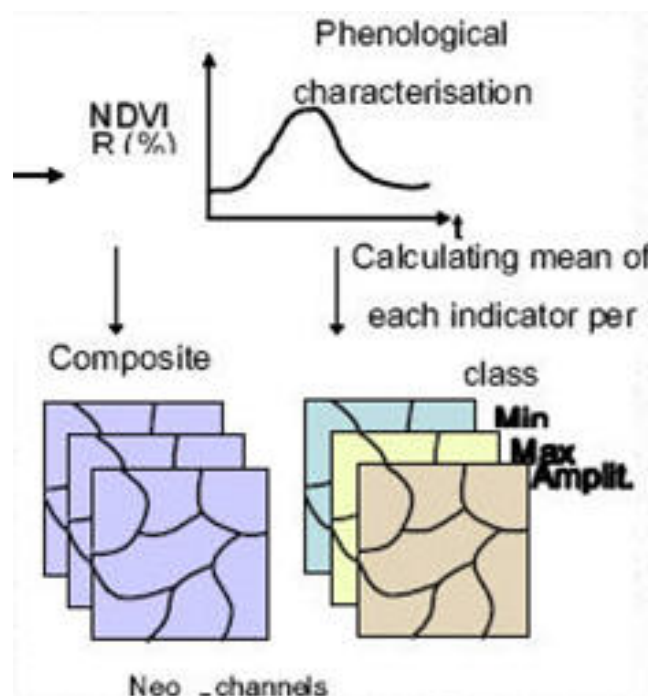


4. Feature identification and selection

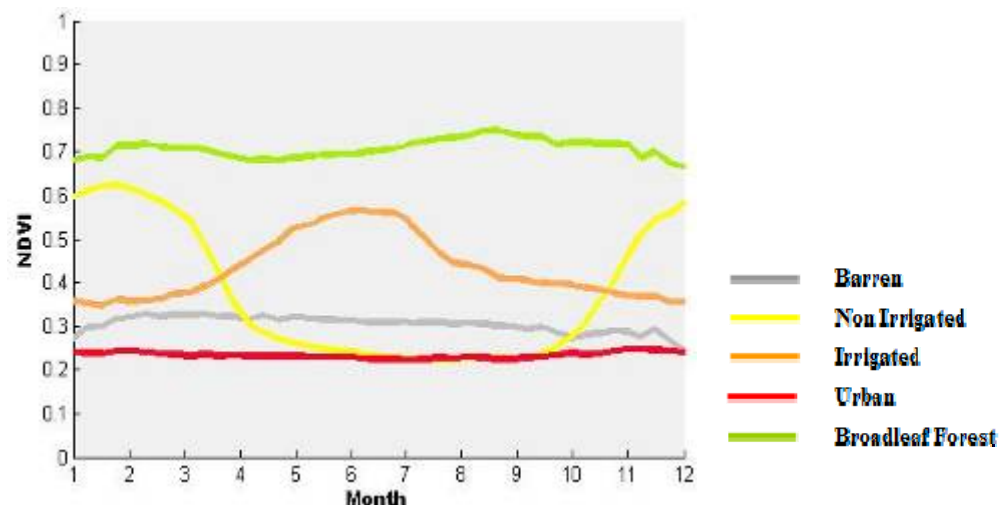
What type of features can we use for information extraction?

2nd order measurements

Measurements of the spatial unit being classified



In the GLOBCOVER project (ESA) a set of new-channels based on the annual NDVI profile are derived.



Source: Defourny et al. (2005)



4. Feature identification and selection

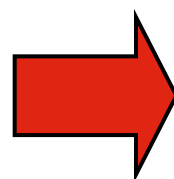
What type of features can we use for information extraction?

2nd order measurements

Measurements related to the neighbourhood (contextual information)

Most mapping approaches operate at a **pixel level**, ignoring its context

Contextual information and semantic relationships with neighbours is always used by photo-interpreters in **visual analysis**.



Several attempts have been carried out to take into automatic classification the contextual information.

Texture

First order statistics in the spatial domain

(e.g. mean, variance, standard deviation, entropy)

Second order statistics in the spatial domain

(e.g. homogeneity, dissimilarity, entropy, angular second moment, contrast, correlation)

Geostatistics

(e.g., variogram, correlogram, covariance function)

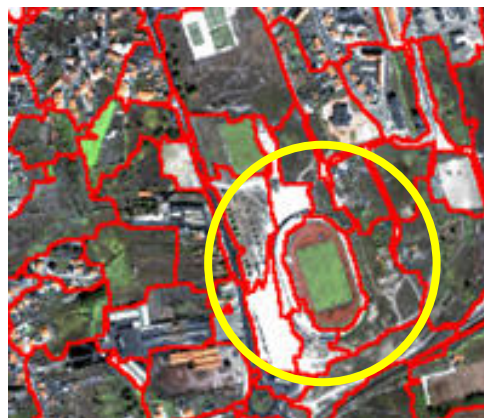
Fractals



4. Feature identification and selection What type of features can we use for information extraction?

...some considerations on object oriented image classification

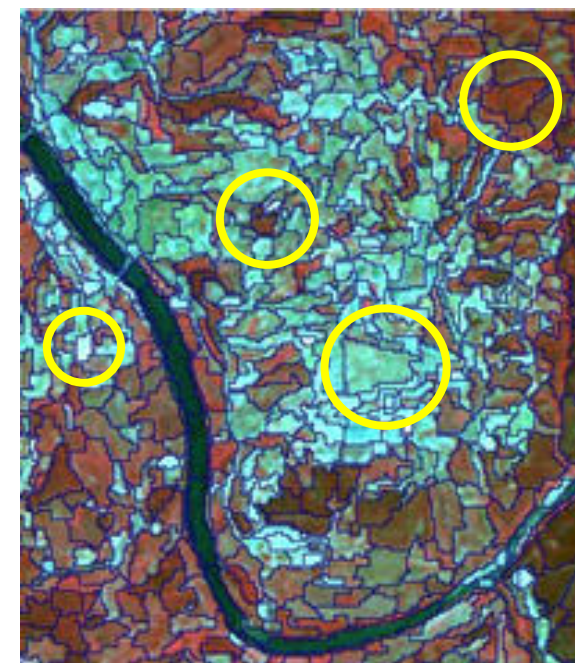
In **object oriented image classification** one can use features that are very similar to the ones used on visual image interpretation



Shape and size of the objects

Spectral homogeneity within objects

Semantic relationships of a spatial unit with its neighbours



Before object oriented image classification there was the **per-field classification**. In this approach the objects are not extracted from the satellite image through segmentation but instead from an existent geographical data base with landscape units, i.e. fields.



4. Feature identification and selection

What type of features can we use for information extraction?

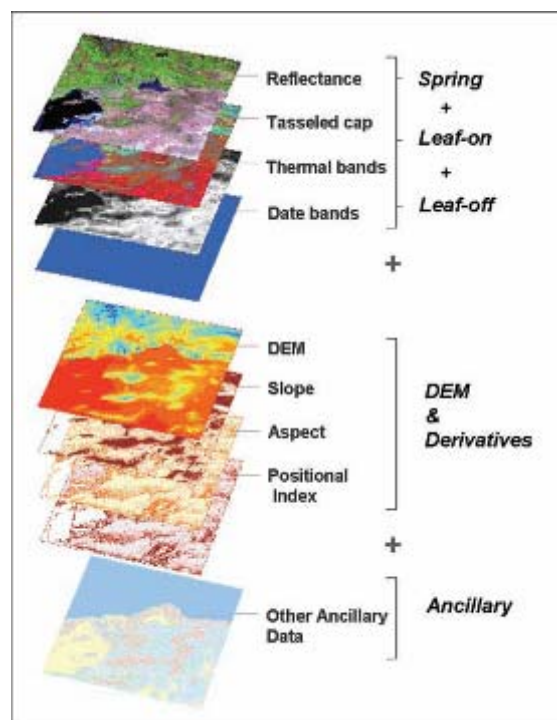
Ancillary information

continuous

e.g. elevation, slope, aspect

categorical

e.g. soil type, existent land cover maps



**US National Land Cover Database
2001**

Source: Homer et al. (2007)



Thematic information extraction from satellite images

- 1 Definition of the mapping approach *
- 2 Geographical stratification
- 3 Image segmentation
- 4 Feature identification and selection *
- 5 **Classification** *
- 6 Ancillary data integration
- 7 Post-classification processing
- 8 Accuracy assessment *

* mandatory



5. Classification

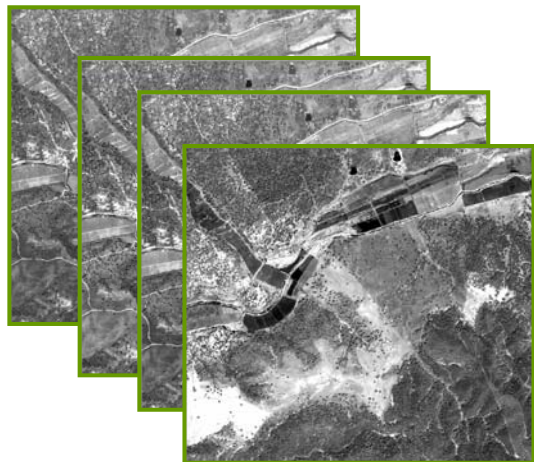
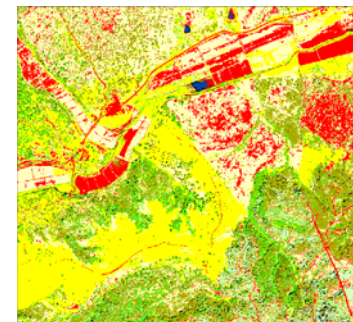


Image spatial space



Allocation of a class
to each spatial unit of
analysis (SUA)



Map of categorical
classes

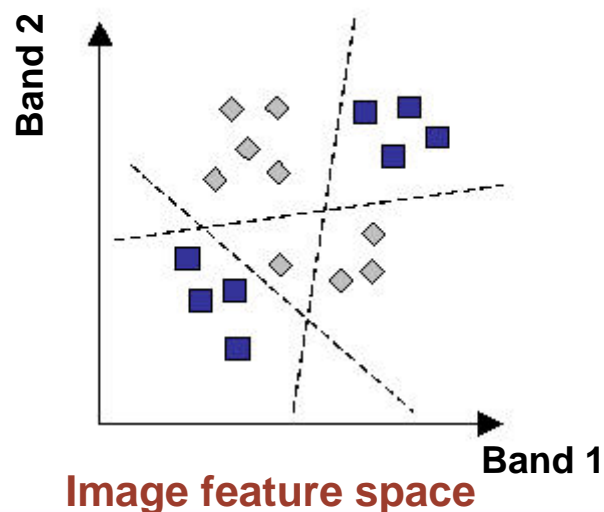


Image feature space

Each SUA is represented by a **vector**,
consisting of a set of measurements (e.g.
reflectance)

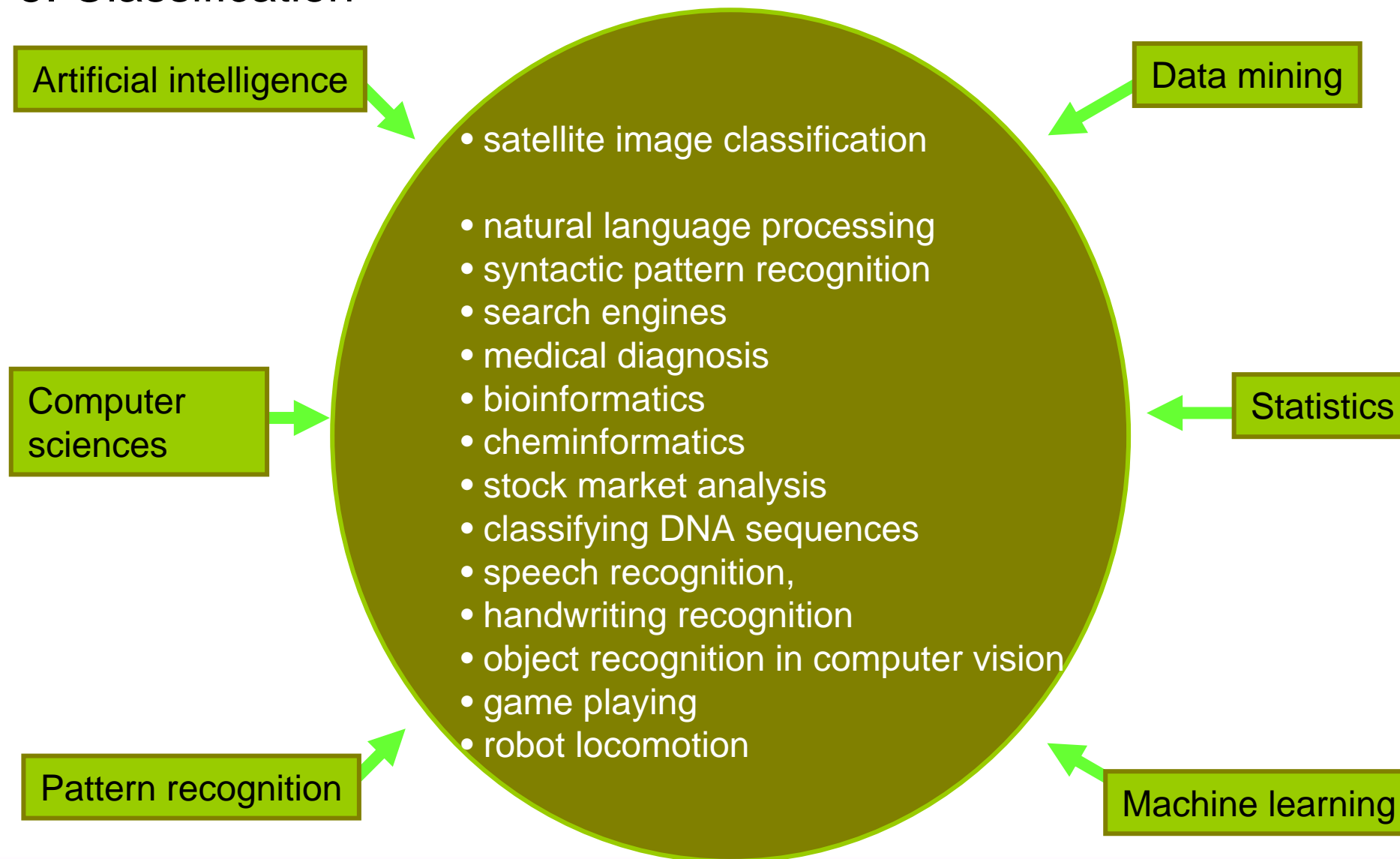
Definition of **decision boundaries** to separate
classes

Definition of the **decision rule**, i.e. the
algorithm that defines the position of a SUA
with respect to the decision boundaries and
that allocates a specific label to that SUA

The word **classifier** is widely used as a synonym of the
term decision rule

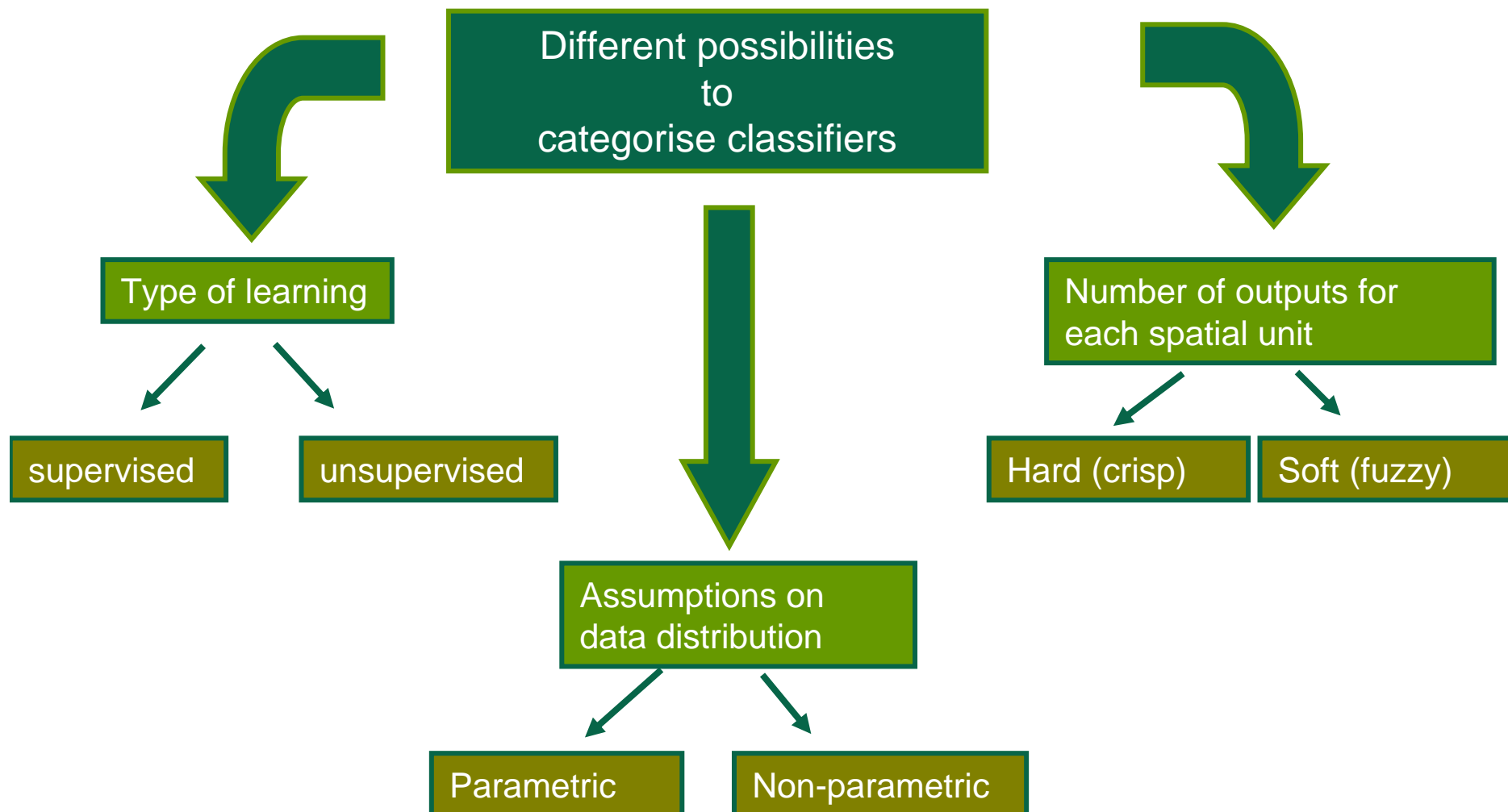


5. Classification





6. Classification



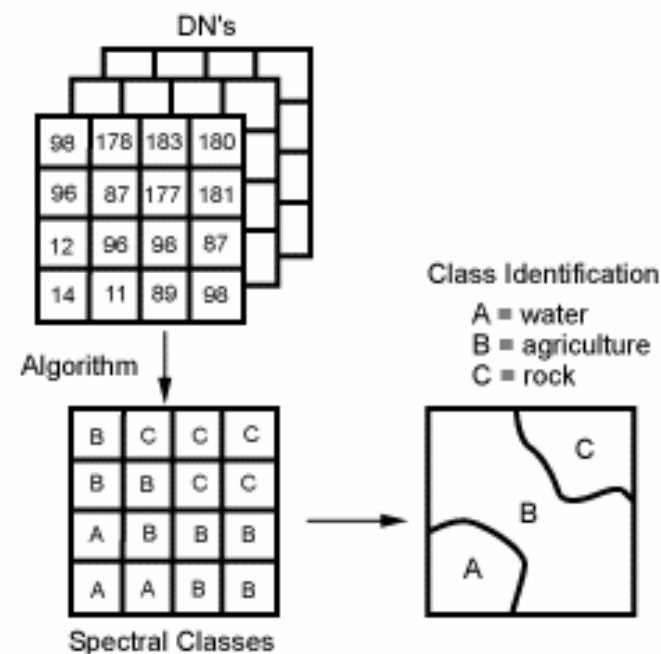
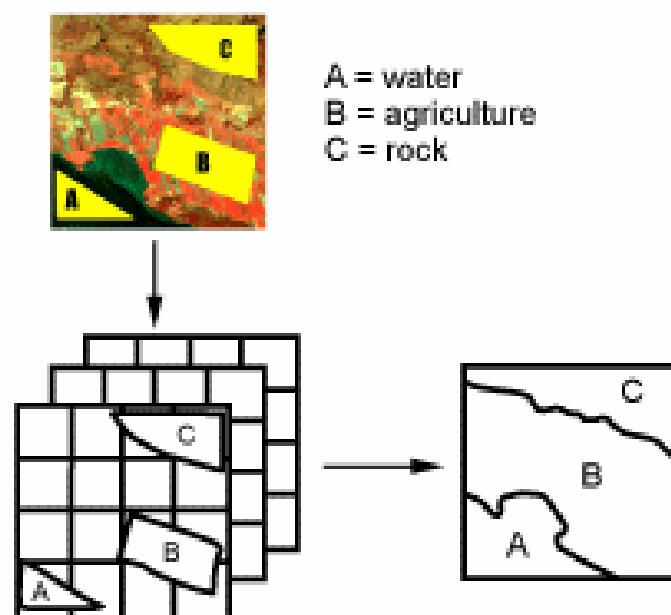


6. Classification

Type of learning

Supervised
classification

Unsupervised
classification

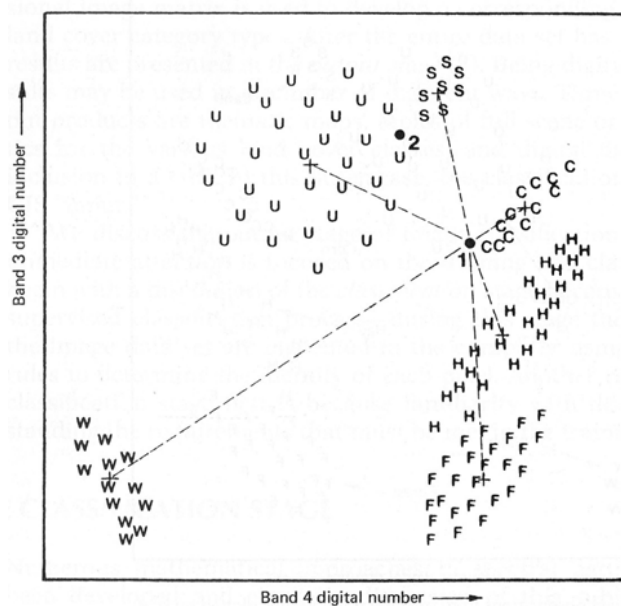


Source: CCRS

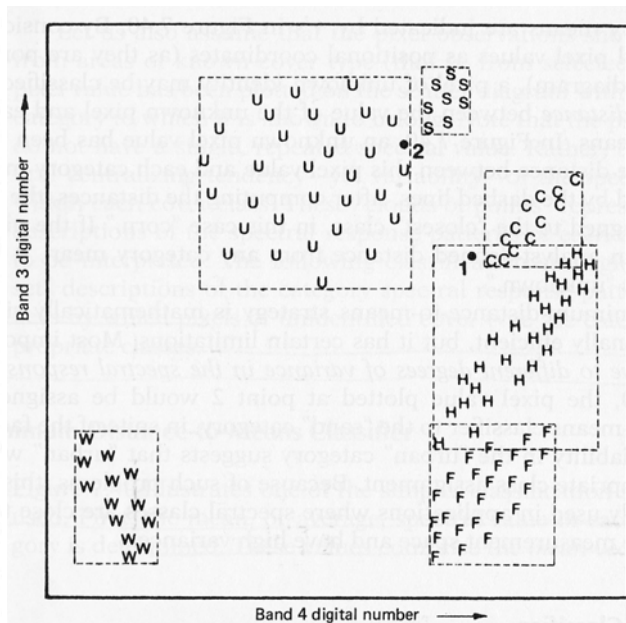


6. Classification

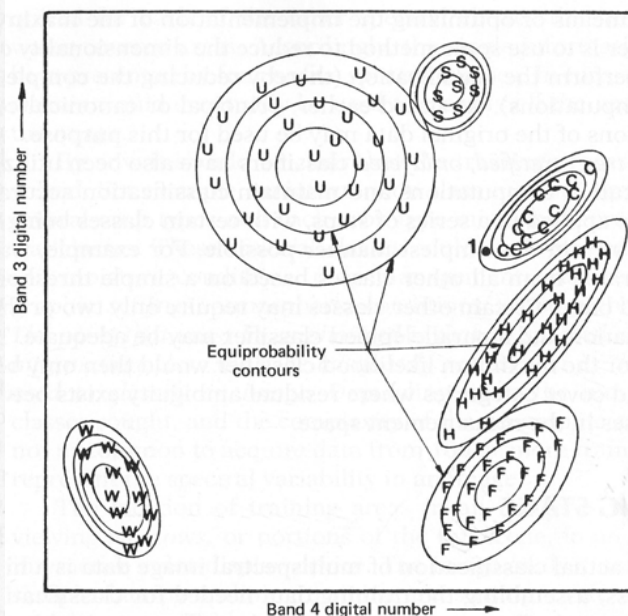
Classic supervised classifiers



Minimum distance



Parallelepiped



Maximum likelihood

Source: Jensen (1996)



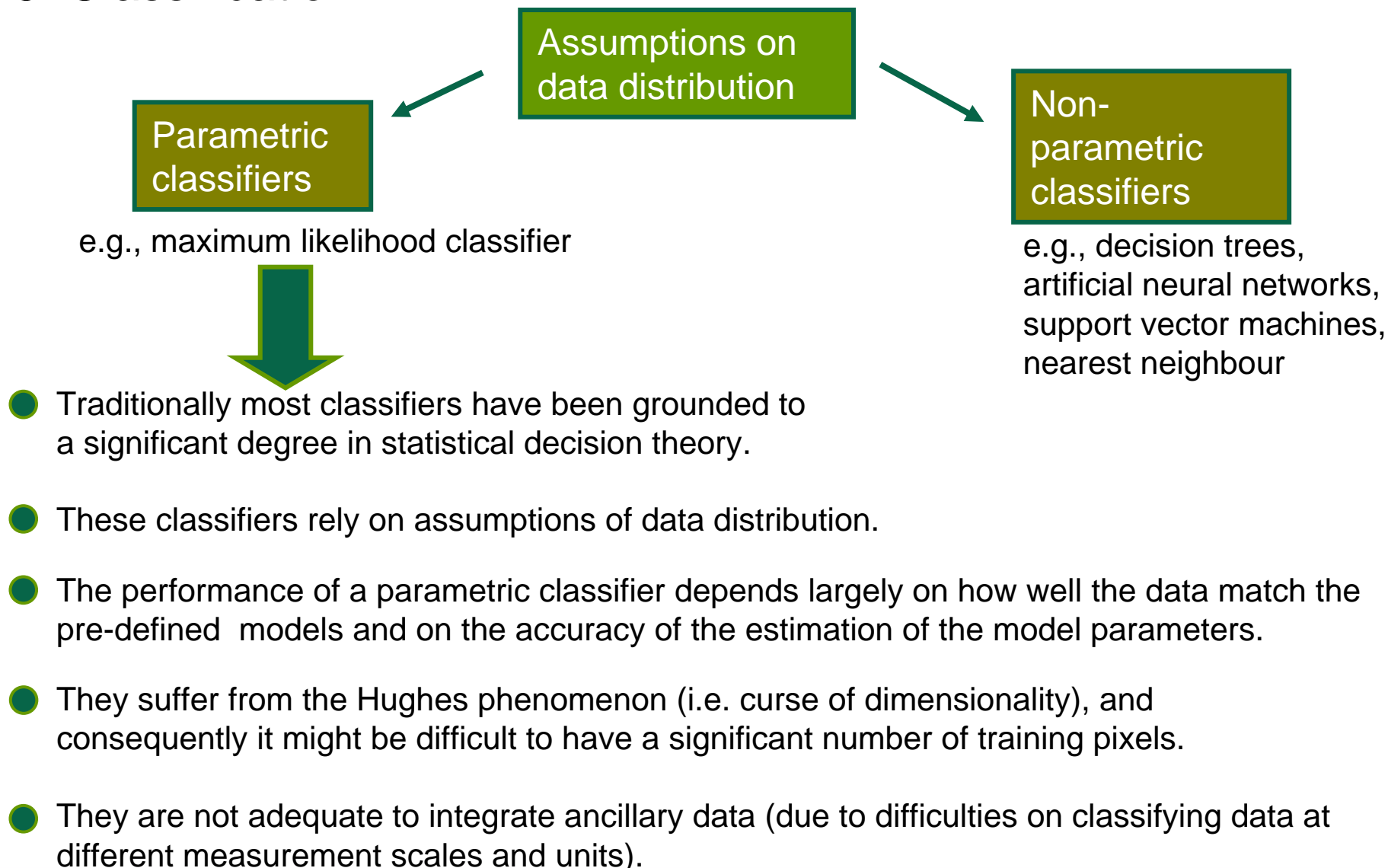
6. Classification

Some considerations on the training stage...

- The training phase is decisive on the final results of image classification. In fact, in these phase we collect the data that will be used to train the algorithm.
- The usual restrictions on sampling (cost, availability of data and accessibility) may lead to an inadequate sampling.
- In case of parametric classifiers the number of sample observations affect strongly the estimates of the statistical parameters.
- As the dimensionality of the data increases for a fixed sample size so the precision of the statistical parameters become lower (i.e., Hughes phenomenon).
- It is common that even mixed pixels dominate the image, only pure pixels are selected for training. However, this may lead to unsatisfactory classification accuracy.



6. Classification





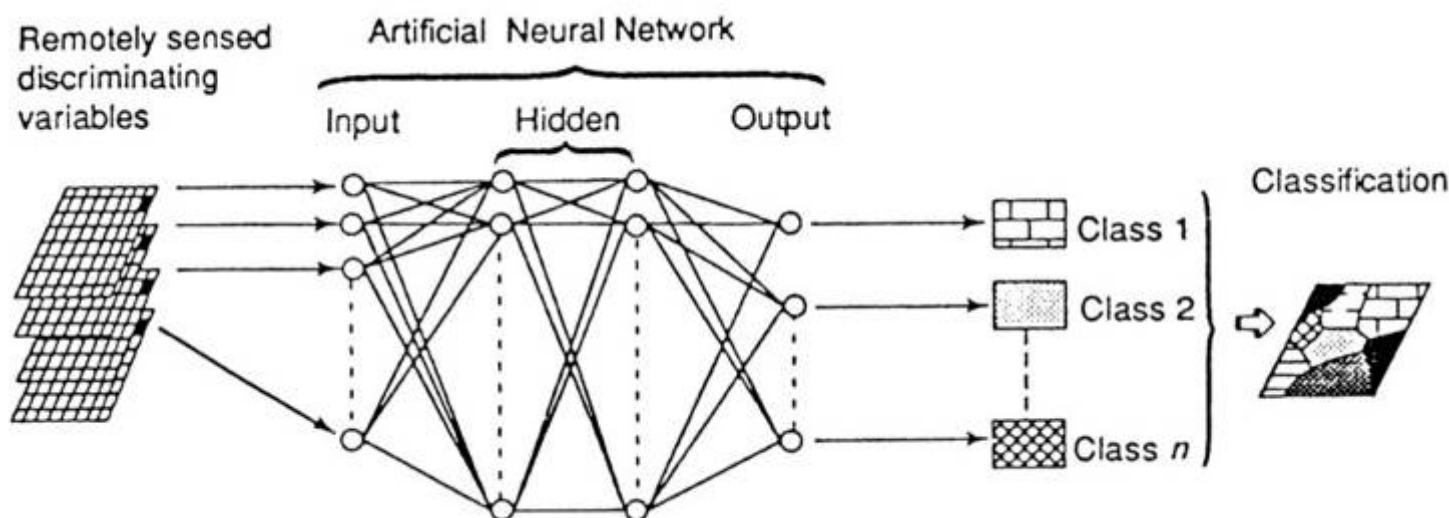
6. Classification

Non-parametric classifiers

Artificial Neural Networks

An ANN is a form of artificial intelligence that imitates some functions of the human brain.

An ANN consists of a series of layers, each containing a set of processing units (i.e. neurones)



All neurones on a given layers are linked by weighted connections to all neurones on the previous and subsequent layers.

During the training phase, the ANN learns about the regularities present in the training data, and based on these regularities, constructs rules that can be extended to the unknown data

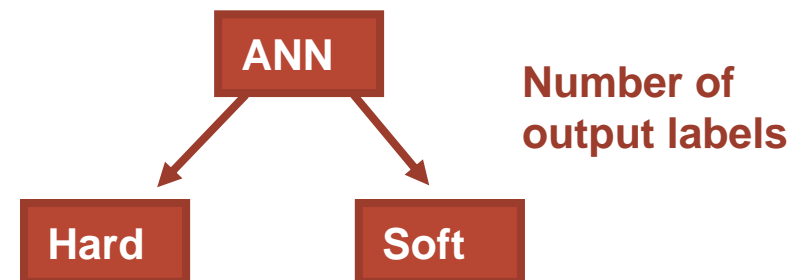
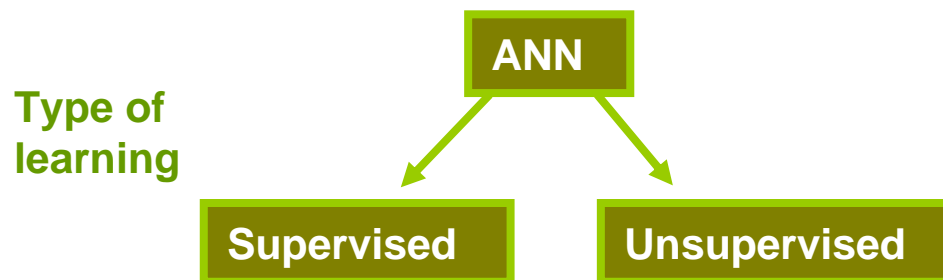
Source: Foody (1999)



6. Classification

Non-parametric classifiers

Artificial Neural Networks



Most common types of ANN

- Multi-layer perceptron with back-propagation
- Self-organised feature map (SOM)
- Hopfield networks
- ART (Adaptive Resonance Theory) Systems



6. Classification

Non-parametric classifiers

Artificial Neural Networks

Advantages of ANN

- It is a non-parametric classifier, i.e. it does not require any assumption about the statistical distribution of the data.
- High computation rate, achieved by their massive parallelism, resulting from a dense arrangement of interconnections (weights) and simple processors (neurones), which permits real-time processing of very large datasets.

Disadvantages of ANN

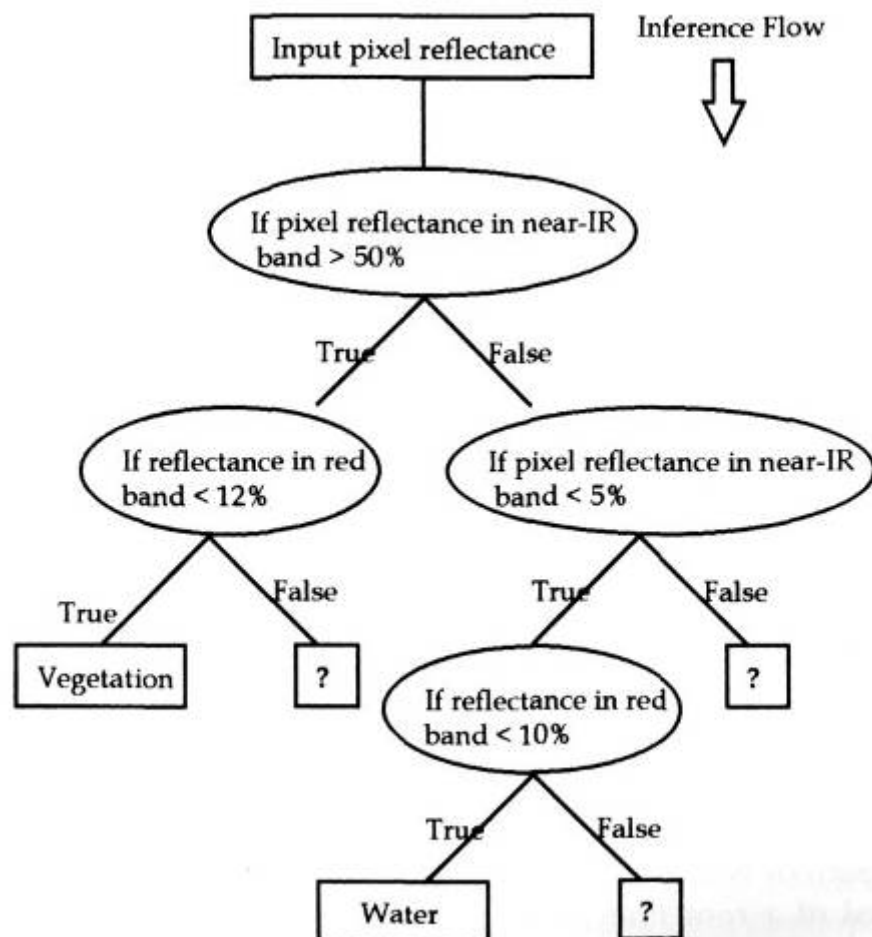
- ANN are semantically poor. It is difficult to gain any understanding about how the result was achieved.
- The training of an ANN can be computationally demanding and slow.
- ANN are perceived to be difficult to apply successfully. It is difficult to select the type of network architecture, the initial values of parameters such as learning rate and momentum, the number of iterations required to train the network and the choice of initial weights.



6. Classification

Non-parametric classifiers

Decision trees



DT are knowledge based (i.e. a method of pattern recognition that simulates the brains inference mechanism).

DT are hierarchical rule based approaches.

DT predict class membership by recursively partitioning a dataset into homogeneous subsets.

Different variables and splits are then used to split the subsets into further subsets.

There are hard and soft (fuzzy) DT.

Source: Tso and Mather (2001)

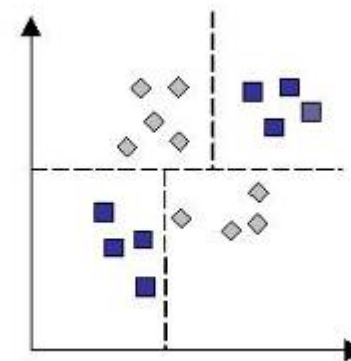
6. Classification

Non-parametric classifiers

Decision Trees

Advantages of DT

- Ability to handle non-parametric training data, i.e. DT are not based on any assumption on training data distribution.
- DT can reveal nonlinear and hierarchical relationships between input variables and use these to predict class membership.
- DT yields a set of rules which are easy to interpret and suitable for deriving a physical understanding of the classification process.
- DT, unlike ANN, do not need an extensive design and training.
- Good computational efficiency.



Disadvantages of DT

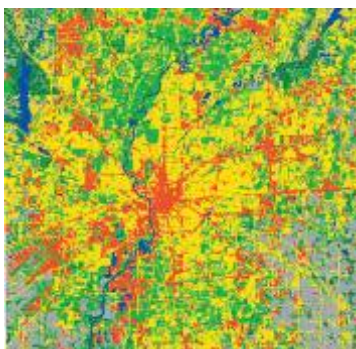
- The use of hyperplane decision boundaries parallel to the feature axes may restrict their use in which classes are clearly distinguishable.

6. Classification

Number of outputs for
each spatial unit

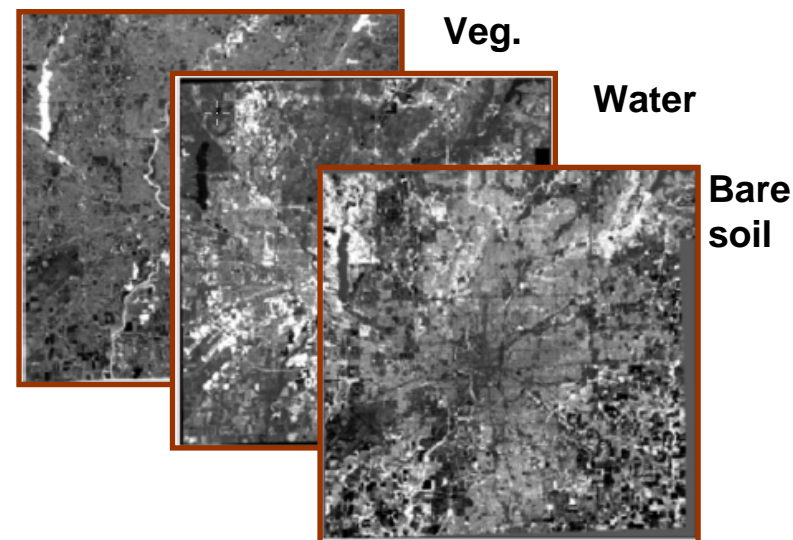
Hard (crisp)
classification

each pixel is forced or constrained to
show membership to a single class.



Soft (fuzzy)
classification

each pixel may display multiple and
partial class membership.

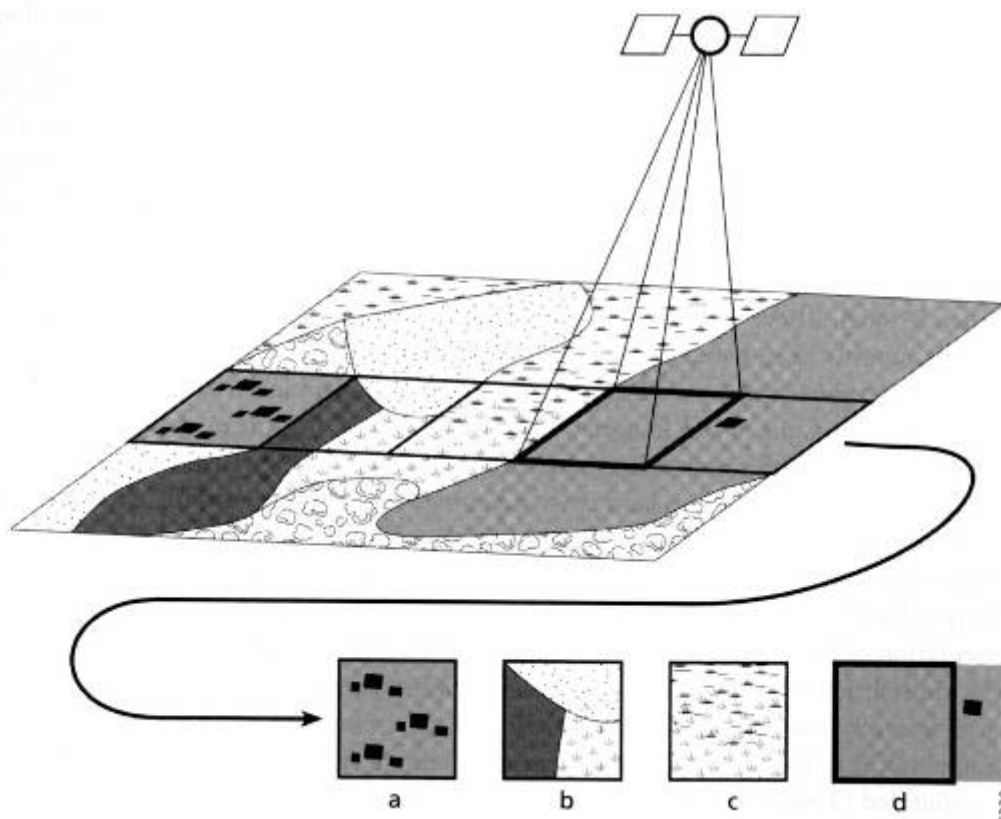


Soft classification has been proposed in the literature as an alternative to hard classification
because of its ability to deal with **mixed pixels**.



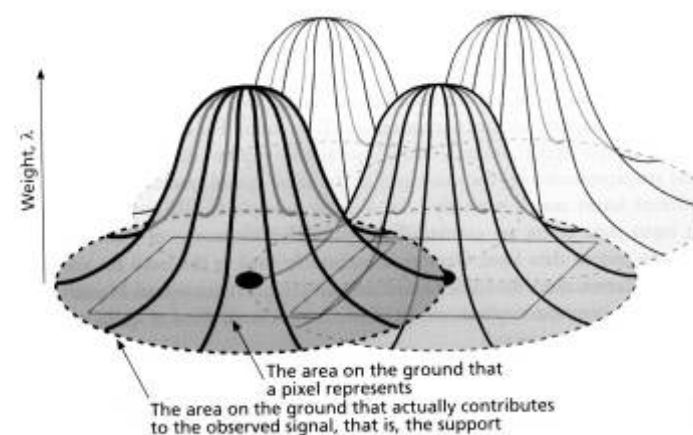
6. Classification

Topic 1:
Introduction
Topic 2:
Satellite
Topic 3:
Remote Sensing
Topic 4:
Image Processing
Topic 5:
Classification
Topic 6:
Land Use Change
Topic 7:
Remote Sensing
Topic 8:
Remote Sensing
Topic 9:
Remote Sensing
Topic 10:
Remote Sensing



The mixed pixel problem

- A – presence of small, sub-pixel targets
- B – presence of boundaries of discrete land cover classes
- C – gradual transition between land cover classes (continuum)
- D – contribution of areas outside the area represented by a pixel



Source: Foody (2004)



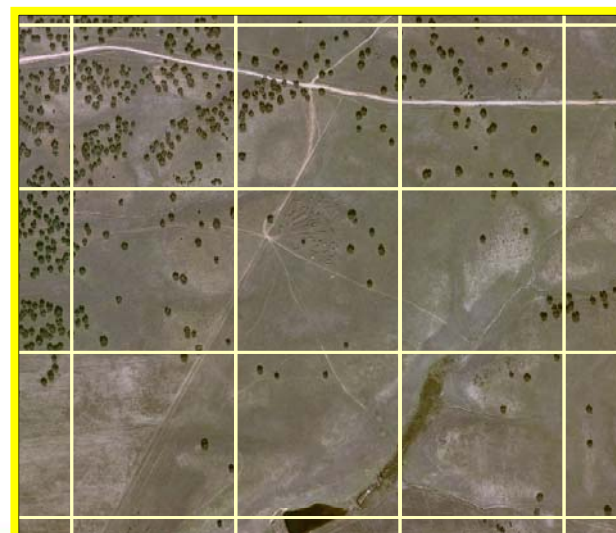
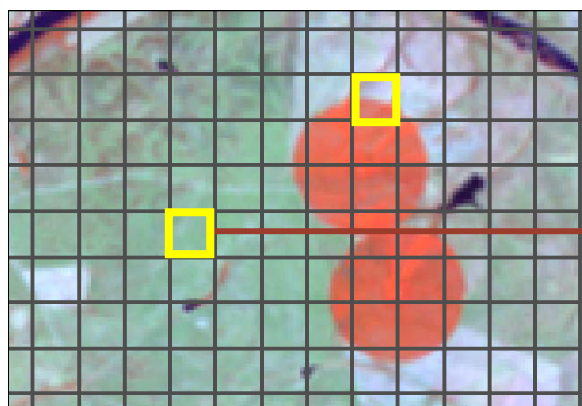
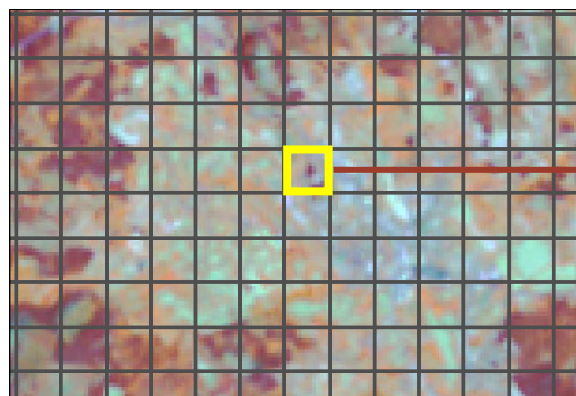
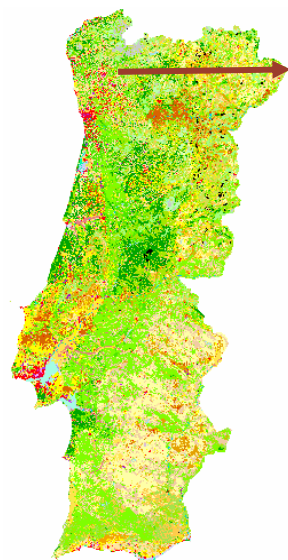
6. Classification

The mixed pixel problem

The number of mixed pixels in an image varies mainly with:

Landscape fragmentation

Sensor's spatial resolution



MERIS FR pixels

6. Classification

The mixed pixel problem

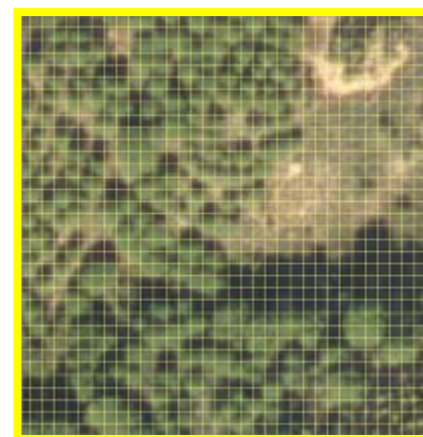
The problem of mixed pixels exist in coarse and fine resolution images:

In coarse resolution images the mixed pixels are mainly due to co-existence in the same pixel of different classes.



MERIS FR

In fine resolution images the mixed pixels are mainly due to co-existence in the same pixel of different components (e.g., houses, trees).

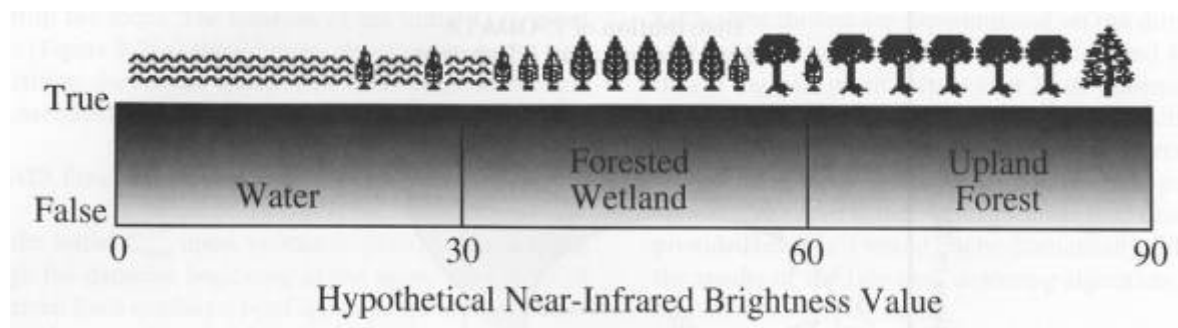


IKONOS



6. Classification

Hard classification



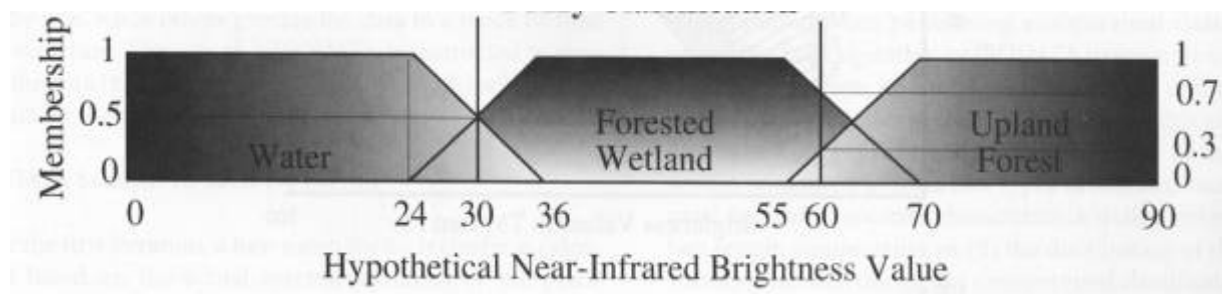
Decision rules

0 – 30 -> Water

30 - 60 -> Forest wetland

60 - 90 -> Upland forest

Fuzzy classification



Decision rules are defined as membership functions for each class.

Membership functions allocates to each pixel a real value between 0 and 1, i.e. membership grade.



But, wow can we represent the sub-pixel information?

Source: Jensen (1996)



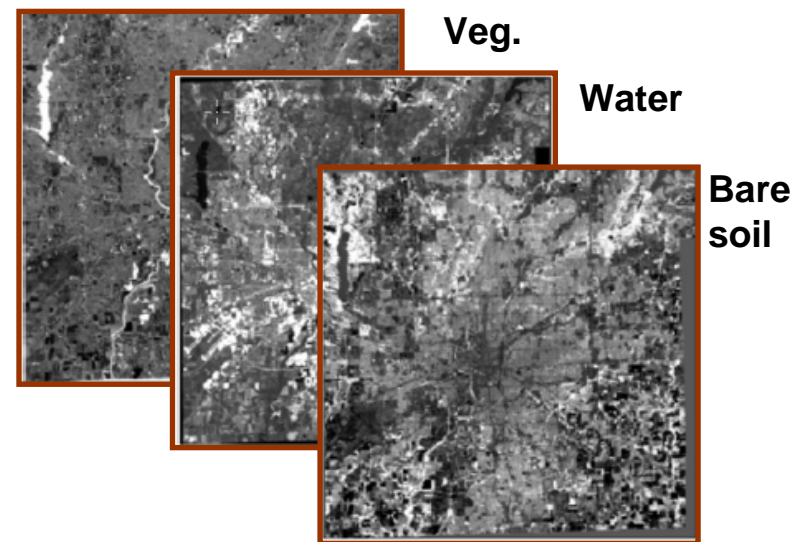
6. Classification

How can we represent the sub-pixel information?

Sub-pixel scale information is typically represented in the output of a soft classification by the **strength of membership a pixel displays to each class.**



It is used to reflect the relative proportion of the classes in the area represented by the pixel





6. Classification

How can we represent the sub-pixel information?

Map with primary and secondary classes

Entropy image

The pixel value translates a degree of mixing (entropy is minimised when the pixel is associated with a single class and maximised when membership is partitioned evenly between all of the defined classes).

Hill's diversity numbers image

The pixel values provides information on the number of classes, the number of abundant classes and the number of very abundant classes.



6. Classification **Soft classifiers**

Most common soft classifiers

- Maximum likelihood classification
- Fuzzy c-means
- Possibilistic c-means
- Fuzzy rule based classifications
- Artificial neural networks

Approaches based on fuzzy set theory



6. Classification **Soft classifiers**

The continuum of classification fuzziness

In the literature the term fuzzy classification has been used for cases where fuzziness is only applied to the allocation stage – which does not seem to be completely correct.

If we apply the concept of fuzziness to all stages of image classification we can create a continuum of fuzziness, i.e. a range of classification approaches of variable fuzziness.

**Completely-crisp
classification**



**Fully-fuzzy
classification**

Classification stages

Dominant class

Training

Individual class
proportions

Pixel is allocated to a
single class

Allocation

Membership grade to all
classes

Dominant class

Testing

Individual class
proportions

Source: Foody (2004)



6. Classification Spectral unmixing

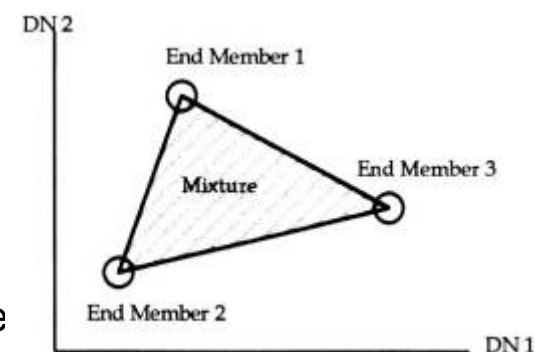
Spectral unmixing = spectral mixture modelling = spectral mixture analysis

Spectral unmixing is an alternative to soft classification for sub-pixel analysis.

Spectral unmixing is based on the assumption that spectral signature of satellite images results essentially from a mixture of a small number of pure components (endmembers) with characteristic spectra.



If so, it is then possible to use a limited number of components so that mixtures of these component spectra adequately simulate the actual observations.



Source: Tso and Mather (2000)

Linear mixture models are the most common models used in satellite image analysis

$$DN_c = \sum_{n=1}^N F_n DN_{n,c} + E_c$$

DN_c – image radiance for band c
 N – is the number of endmembers
 F_n – is the relative fraction of endmember n
 $DN_{n,c}$ – is the endmember n inner radiance
 E_c – residual fitting error

6. Classification

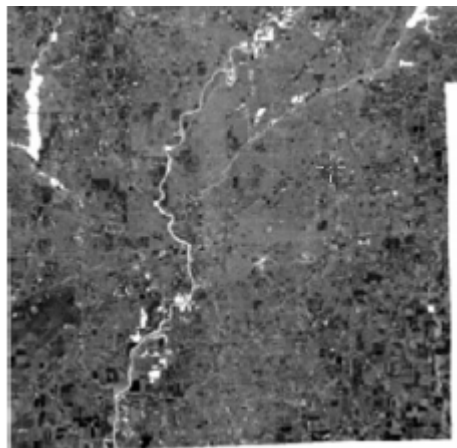
Spectral unmixing

A case study: urban mapping

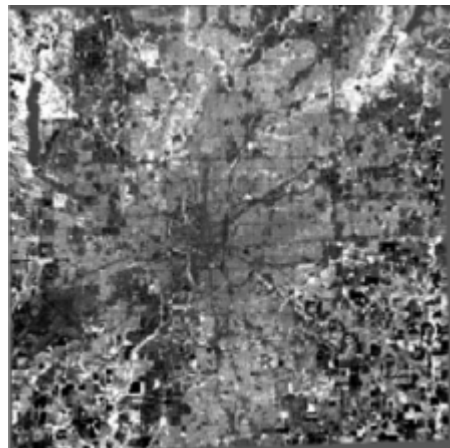
Lu and Weng (2004) used Spectral Mixture Analysis for mapping the Urban Landscape in Indianapolis with Landsat ETM+ Imagery.

SMA was used to derive fraction images to three endmembers: shade, green vegetation, and soil or impervious surface

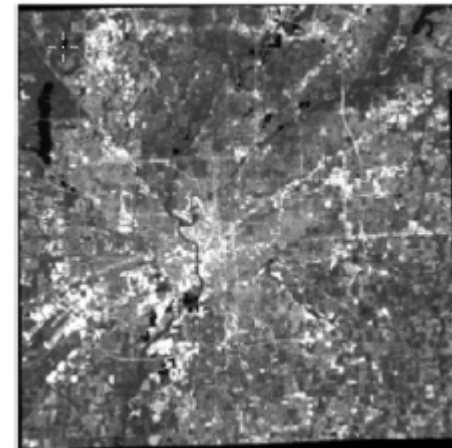
Output of spectral unmixing



Shade fraction



Vegetation fraction

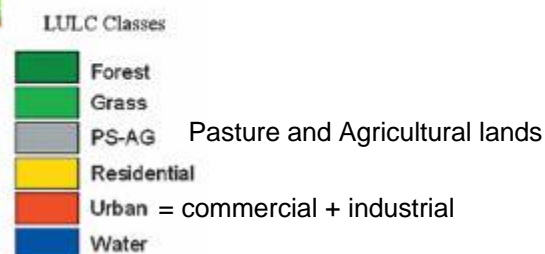
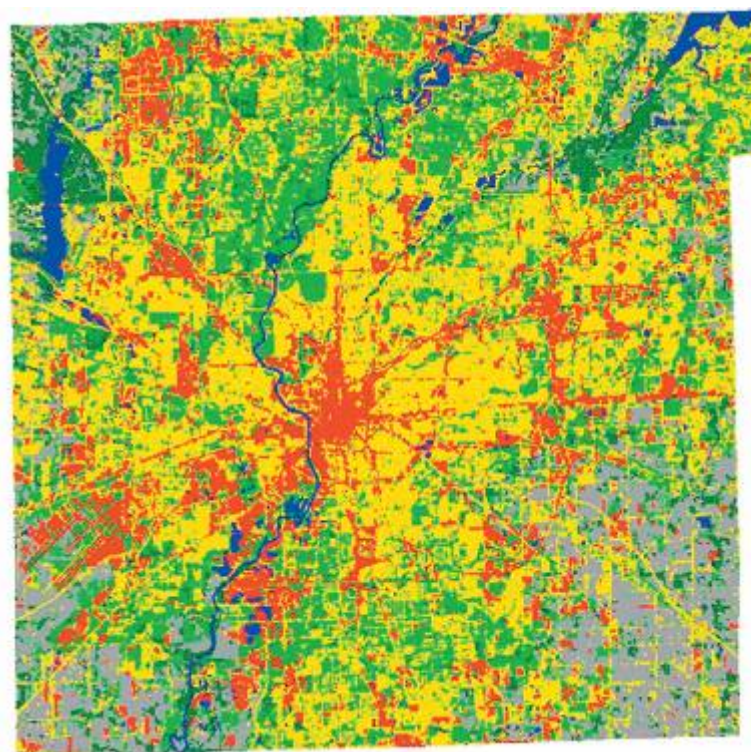


Soil or impervious surface fraction

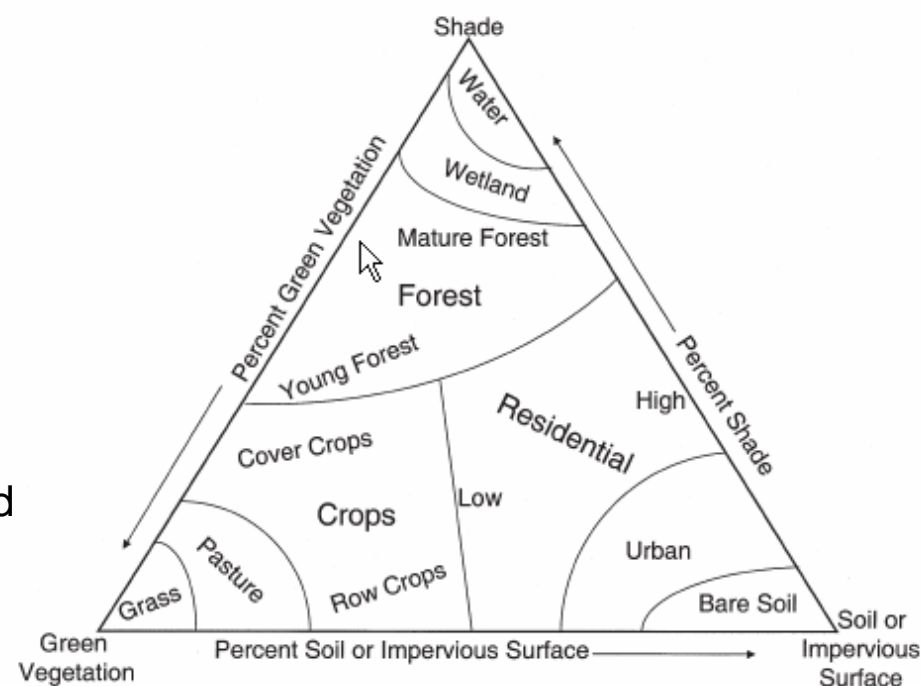
6. Classification

Spectral unmixing

A case study: urban mapping



Lu-Weng urban landscape model



The fraction images were used to classify LULC classes based on a hybrid procedure that combined maximum-likelihood and decision-tree algorithms.

Source: Lu and Weng (2004)



6. Classification **Sub-pixel classification** **Super-resolution mapping**

Although classification at sub-pixel level is informative and meaningful it fails to account for the spatial distribution of **class** proportions within the pixel.

Super-resolution mapping (or sub-pixel mapping) is a step forward.

Super-resolution mapping considers the spatial distribution within and between pixels in order to produce maps at sub-pixel scale.

Several approaches of super-resolution mapping have been developed:

- Hopfield neural networks
- Pixel-swapping solution (based on geostatistics)
- Linear optimization
- Markov random fields

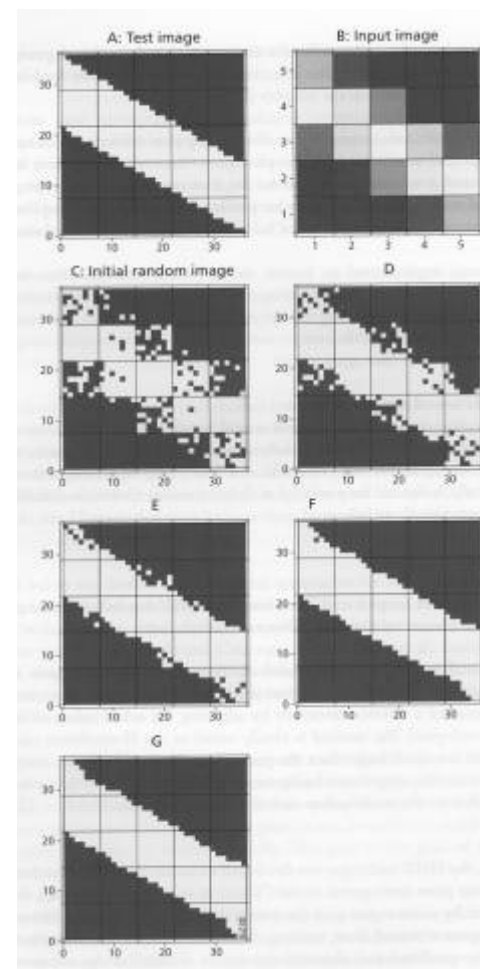
6. Classification

Sub-pixel classification

Super-resolution mapping

Pixel-swapping solution – this technique allows sub-pixel classes to be swapped within the same pixel only.

Swaps are made between the most and least attractive locations if they result in an increase in spatial correlation between sub-pixels.



Source: Atikson (2004)



6. Classification

Multiple classifiers approach

Rationale

- Different classifiers originate different classes for the same spatial unit
- There are several studies on the comparison of different classifiers
- There is not a single classifier that performs best for all classes. In fact it appears that many of the methods are complementary
- Combination of decision rules can bring advantages over the single use of a classifier

In the multiple classifiers approach the classifiers should be independent. To be independent the classifiers must use an independent feature set or be trained on separate sets of training data.

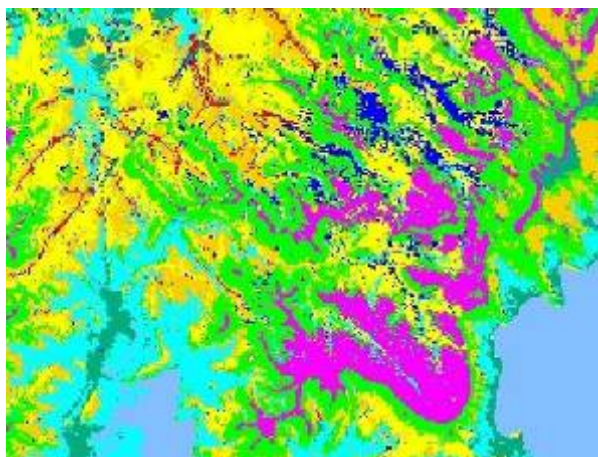


6. Classification **Multiple classifiers approach**

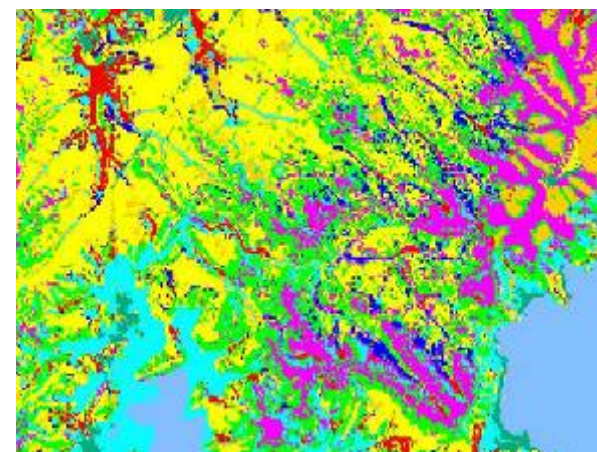
How different the results from different classifiers can be?



Maximum likelihood



Artificial Neural Networks



Decision tree

Source: Gahegan and West (1998)



6. Classification

Multiple classifiers approach

Methods for combining classifiers

Voting rules

The label outputs from different classifiers are collected and the majority label is selected (i.e. majority vote rule). There are some variants, such as the comparative majority voting (it requires that the majority label should exceed the 2nd more voted by a specific number).

Bayesian formalism

It is used with multiple classifiers that output a probability. The probabilities for a spatial unit for each class resulting from different classifiers are accumulated and the final label is the one that has the greatest accumulated probability.

Evidential reasoning

It associates a degree of belief with each source of information, and a formal system of rules is used in order to manipulate the belief function.

Multiple neural networks

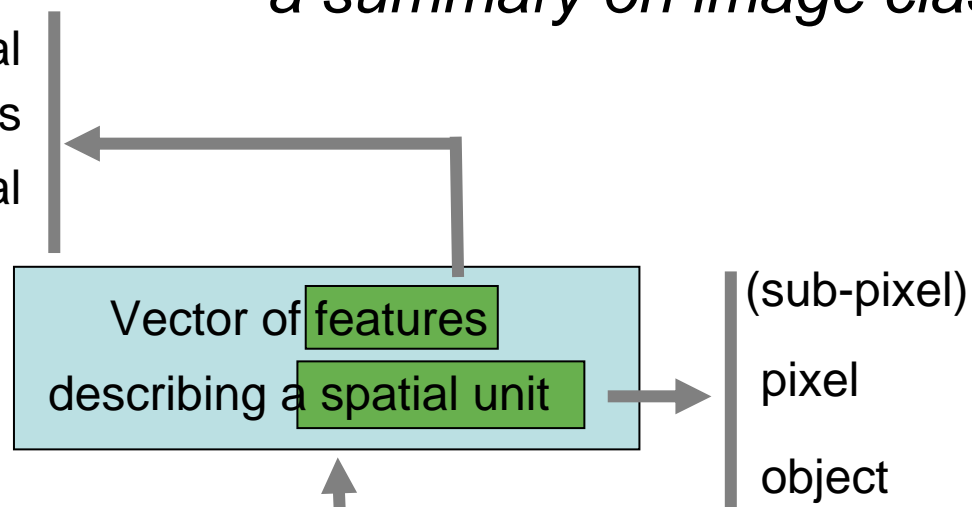
It consists on the use of a neural network to produce a single class to each spatial unit, fed with the outputs from different classifiers.



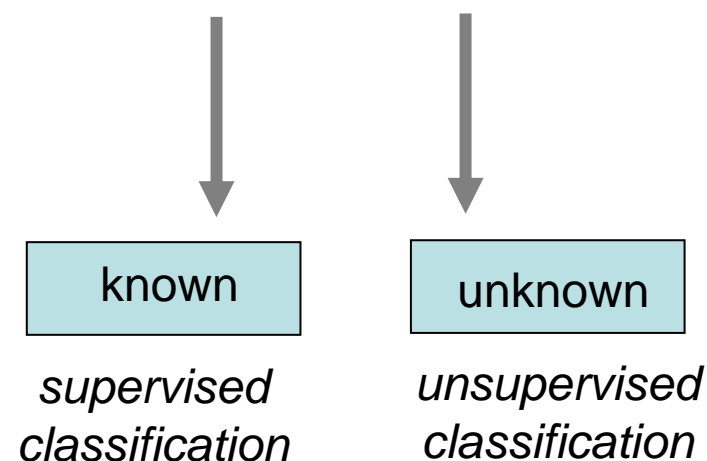
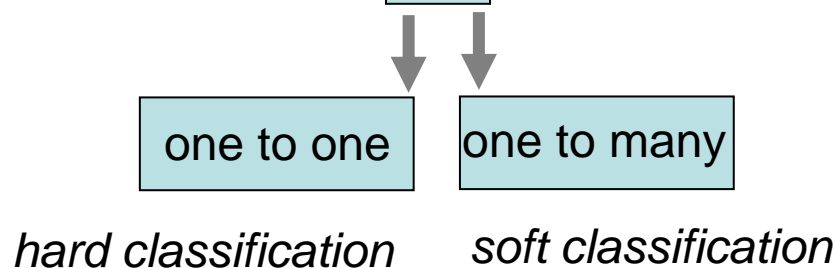
6. Classification

a summary on image classification...

spectral
secondary measurements
geographical



**The aim of pattern recognition
is to establish a link between a pattern and a class label**





Thematic information extraction from satellite images

- 1 Definition of the mapping approach *
- 2 Geographical stratification
- 3 Image segmentation
- 4 Feature identification and selection *
- 5 Classification *
- 6 **Ancillary data integration**
- 7 Post-classification processing
- 8 Accuracy assessment *

* mandatory



7. Integration of ancillary information

Ancillary data can be integrated after image classification in order to improve the results.

Post-classification sorting - application of very specific rules to classification results and to geographical ancillary data (e.g., elevation, slope, aspect)



There are several strategies based on expert systems, rule based systems and knowledge base systems



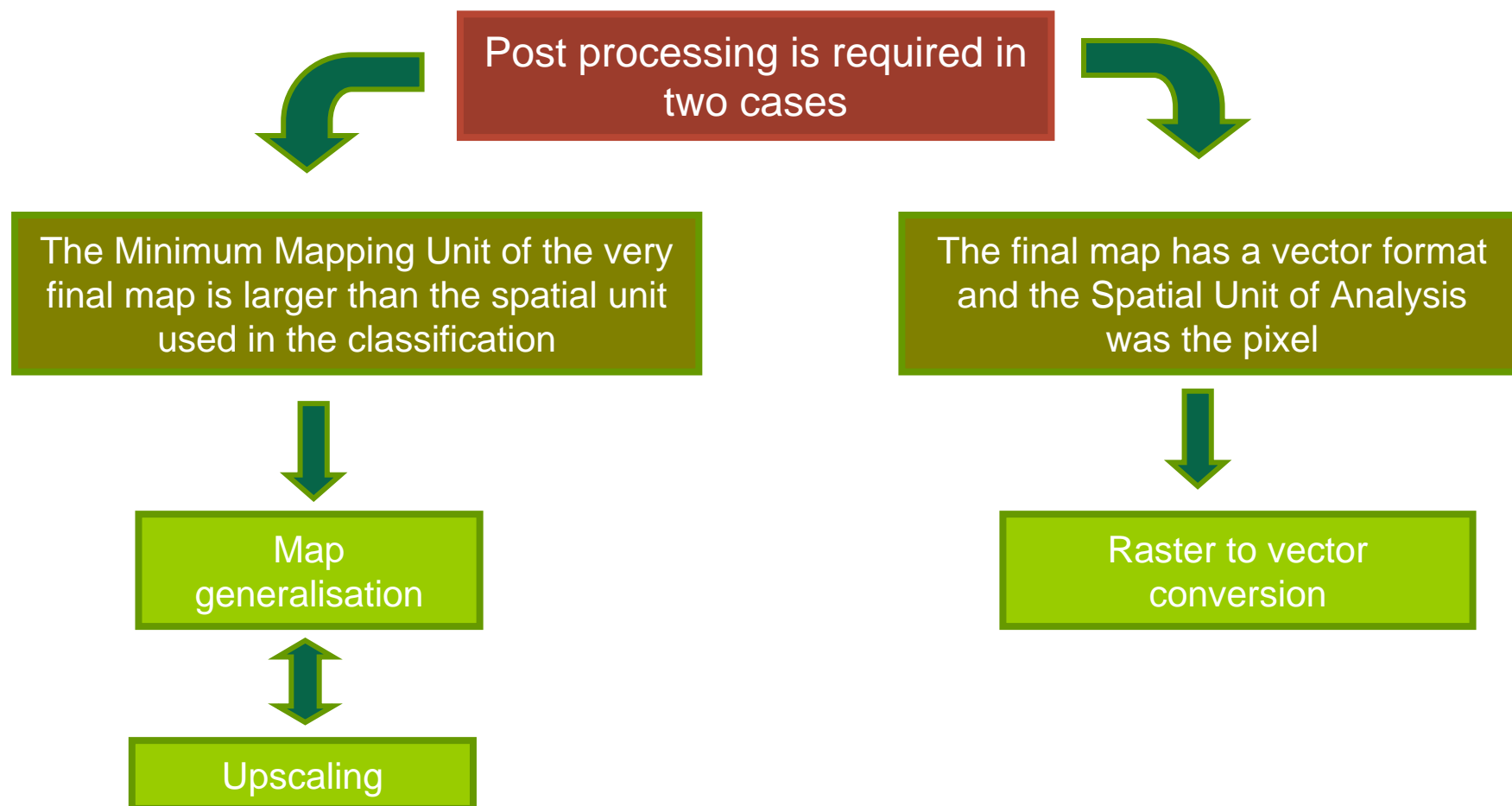
Thematic information extraction from satellite images

- 1 Definition of the mapping approach *
- 2 Geographical stratification
- 3 Image segmentation
- 4 Feature identification and selection *
- 5 Classification *
- 6 Ancillary data integration
- 7 **Post-classification processing**
- 8 Accuracy assessment *

* mandatory



8. Post-processing





1. Definition of the mapping approach

The steps required to information extraction depend on the defined mapping approach:

Map format = raster

MMU = pixel size of input satellite data

Feature selection > Image classification > accuracy assessment

MMU > pixel size of input satellite data

Feature selection > Image classification > post-processing > accuracy assessment

Map format = vector

↑
downscaling
↑

Spatial unit of analysis = image pixel

Feature selection > Image classification > post-processing > accuracy assessment

↑
downscaling
↑

Spatial unit of analysis = object

Generalisation + Raster to vector conversion

Image segmentation > Feature selection > Image classification > post-processing > accuracy assessment

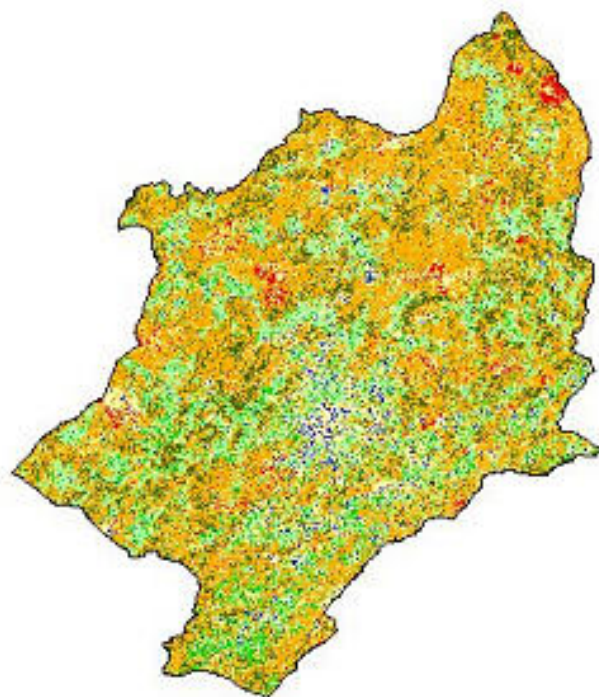
↑
Generate the objects
↑

↑
Generalisation
↑



8. Post-processing

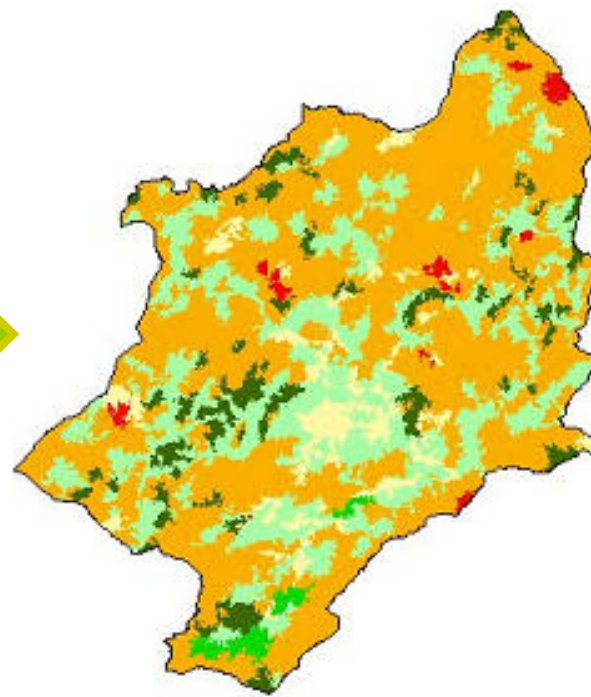
Semantic generalisation



MMU = 1 pixel (30mx30m)



Semantic
generalisation

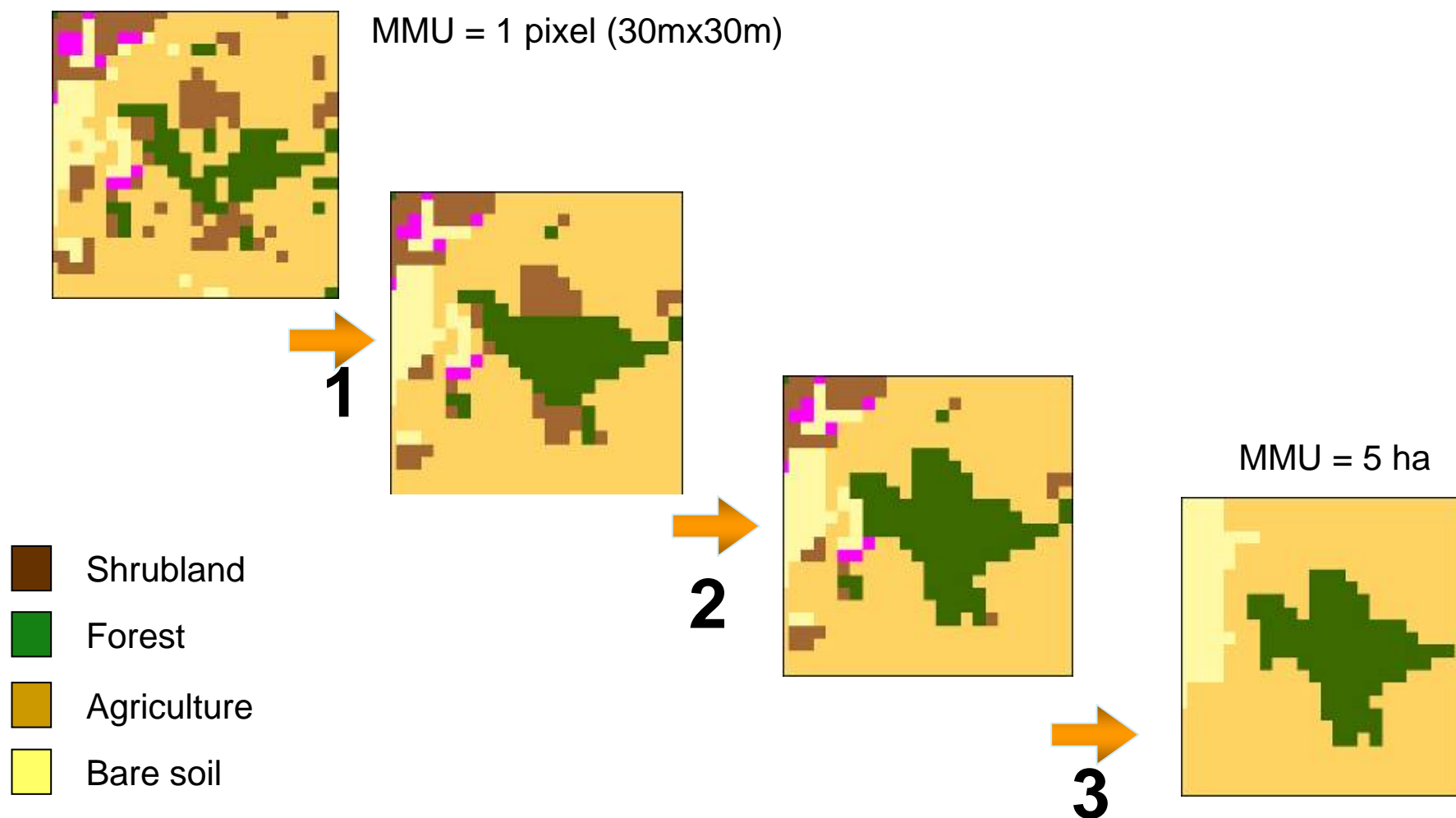


MMU = 5 ha



8. Post-processing

Semantic generalisation





Thematic information extraction from satellite images

- 1 Definition of the mapping approach *
- 2 Geographical stratification
- 3 Image segmentation
- 4 Feature identification and selection *
- 5 Classification *
- 6 Ancillary data integration
- 7 Post-classification processing
- 8 **Accuracy assessment** *

* mandatory

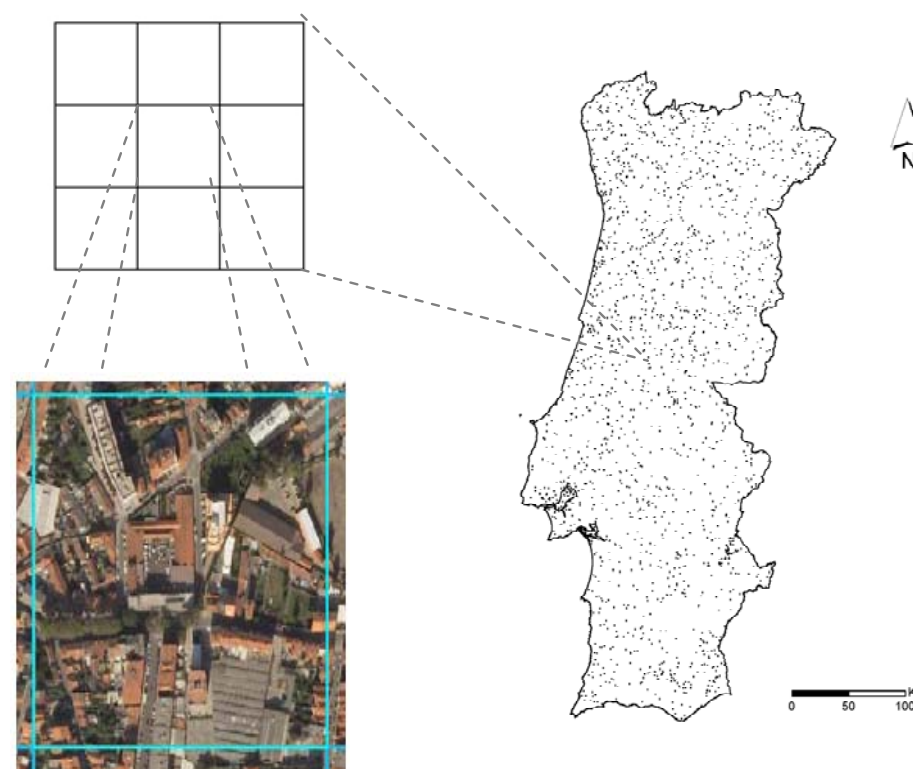


9. Accuracy assessment

Accuracy assessment allows users to evaluate the utility of a thematic map for their intended applications.

The most widely used method for accuracy assessment may be derived from a **confusion or error matrix**.

The confusion matrix is a simple cross-tabulation of the mapped class label against the observed in the ground or reference data for a **sample set**.





9. Accuracy assessment

Main steps

1 Selection of the reference sample

sampling units
sampling design

Probability sampling is necessary if one wants to extend the results obtained on the samples to the whole map.

Probability sampling requires that all inclusion probabilities be greater than zero, e.g. one cannot exclude from sampling inaccessible areas or landscape unit borders.

2 Response design

The definition of the response design depends on the process for assessing agreement (e.g., primary, fuzzy or quantitative).

3 Analysis and estimation

One has to take into account the known areas (marginal distributions) of each map category to derive unbiased estimations of the proportion of correctly mapped individuals.

Source: Stehman (1999)



Goals

1 From data to information: presentation of different mapping approaches

2 Most common problems in image classification and how to solve them

e.g. mixed pixel problem, lack of normality of the training data, Hughes phenomenon

3 Most important advances in satellite image classification

**e.g. from pixel to object, from hard to soft classifiers,
from parametric to non-parametric classifiers**



References

Atkinson, P.M., 2004, Resolution manipulation and sub-pixel mapping, in S.M. de Jong and F.D. van der Meer (eds), Remote sensing image analysis – including the spatial domain, Dordrecht: Kluwer Academic Publishers.

Defourny, P., Vancutsem, C., Bicheron, P, Brockmann, C., Nino, F., Schouten, L., Leroy, M., 2006, GLOBCPVER: a 300m global land cover product for 2005 using ENVISAT MERIS Time Series, Proceedings of ISPRS Commission VII Mid-Term Symposium: Remote Sensing: from Pixels to Processes, Enschede (NL), 8-11 May, 2006

Foody, G. M., 2004, Sub-pixel methods in remote sensing, in in S.M. de Jong and F.D. van der Meer (eds), Remote sensing image analysis – including the spatial domain, Dordrecht: Kluwer Academic Publishers.

Foody, G. M., 2002, Status of land cover classification accuracy assessment, Remote Sensing of the Environment, 80: 185-2001.

Foody, G.M., 1999, Image classification with a neural network: from completely crisp to fully-fuzzy situations, in P.M. Atkinson and N.J. Tate (eds), Advances in Remote Sensing and GIS analysis, Chichester: Wiley&Son.



References

Homer, C. C. Huang, L. Yang, B. Wylie and M. Coan. 2004. Development of a 2001 National Landcover Database for the United States. *Photogrammetric Engineering and Remote Sensing*, 70 (7): 829-840

Homer, C. C. Huang, L. Yang, B. Wylie and M. Coan. 2004. Development of a 2001 National Landcover Database for the United States. *Photogrammetric Engineering and Remote Sensing*, 70 (7): 829-840

Jensen, J.R., 1996, *Introductory digital image processing: a remote sensing perspective*, Upper Saddle River, NJ: Prentice Hall, 2nd Ed.

Lu, D. and Weng, Q., 2004, Spectral Mixture Analysis of the Urban Landscape in Indianapolis with Landsat ETM+ Imagery, *Photogrammetric Engineering and Remote Sensing*, 70 (9), pp. 1053-1062

Wilkinson, G.G., 2005, Results and implications of a study of fifteen years of satellite image classification experiments, *IEEE Transaction on Geosciences and Remote Sensing*, 43:3, 433-440

Stehman, S.V., 1999, Basic probability sampling designs for thematic map accuracy assessment, *International Journal of Remote Sensing*, 20: 2423–2441.