



### Vybrané kapitoly z DPZ II

Alternativní metodické přístupy k analýze z obrazových dat v prostředí města

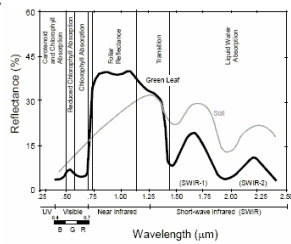
## Urban Remote Sensing



### Možné přístupy k řešení problému

1. Spektrální zvýraznění I. - indexy
2. Spektrální zvýraznění II. (PCA, MNF)
3. Obrazová spektrometrie
4. Analýza smíšených pixelů (spectral mixture analysis)
5. Objektově orientovaná klasifikace, kontextuální
6. Texture analysis, GLCM
7. Neparаметrické metody (strojové učení, neuronové sítě,...)
8. ...

## 1. Spektrální indexy



- Jaký je obecný princip vytváření spektrálních indexů?
- Jak by bylo možné zkonstruovat spektrální index zvýrazňující zastavěné plochy?
- Proč se využívá tzv. normalizovaných indexů?

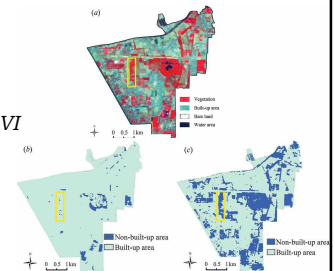
## Spektrální indexy

### Normalized Difference Built-up Index (NDBI)

Slouží ke studiu prostorového rozložení zastavěných ploch a časových změn v jejich rozsahu

$$NDBI = \frac{TM5 - TM4}{TM5 + TM4}$$

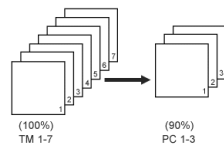
$$Built\_up\_area = NDBI - NDVI$$



Ve výsledném snímku mají zastavěné a holé plochy bez vegetace kladné hodnoty a ostatní druhy povrchů hodnoty záporné

## 2. Metody spektrálního zvýraznění

### Obrazová analýza hlavních komponent

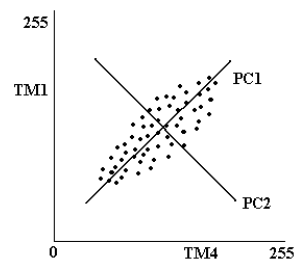


(PCA - Principal Component Analysis) se používá jako prostředek zvýraznění obrazu k vizuální interpretaci i jako metoda zvýraznění obrazu před jeho automatickou klasifikací.

Analýzou hlavních komponent lze další zpracování omezit na méně pásem bez podstatné ztráty informace - redukuje tedy rozměrnost (dimensionalitu) zpracování.

PCA je statistická metoda, která rotuje osami vícerozměrného prostoru tvořícího multispektrální snímek, a to ve směru maximálního rozptylu dat.

## Princip PCA



## Výstup PCA

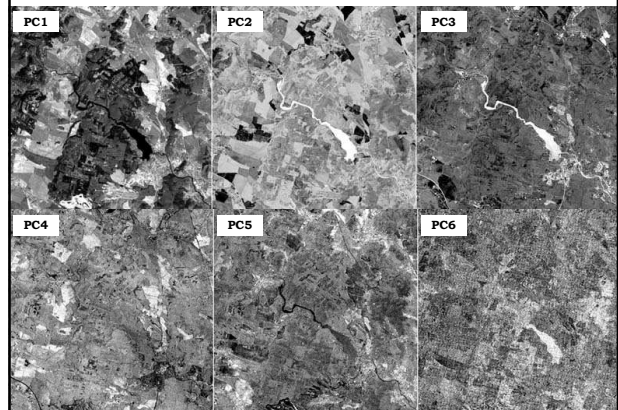
$$PC_1 = a_1TM_1 + a_2TM_2 + a_3TM_3 + a_4TM_4 + a_5TM_5 + a_6TM_7$$

$$PC_2 = b_1TM_1 + b_2TM_2 + b_3TM_3 + b_4TM_4 + b_5TM_5 + b_6TM_7$$

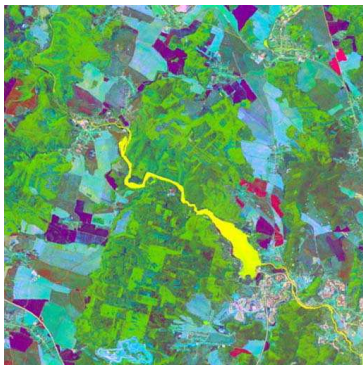
...

Číslo PC	Vlastní čísla	Procenta rozptylu	Kumulov. procenta	Zátěže						
				TM 1	TM 2	TM 3	TM 4	TM 5	TM 7	
1	2262,96	75,62	75,62	0,243	0,181	0,346	0,230	0,728	0,454	
2	682,34	22,80	98,42	0,115	0,050	0,229	-0,936	-0,012	0,237	
3	33,80	1,13	99,55	0,553	0,323	0,513	0,201	-0,531	-0,064	
4	7,79	0,26	99,81	-0,264	-0,141	-0,037	0,168	-0,432	0,833	
5	4,54	0,15	99,96	0,712	-0,102	-0,668	-0,034	0,000	0,186	
6	1,21	0,04	100,00	-0,212	0,911	-0,343	-0,044	-0,022	0,069	

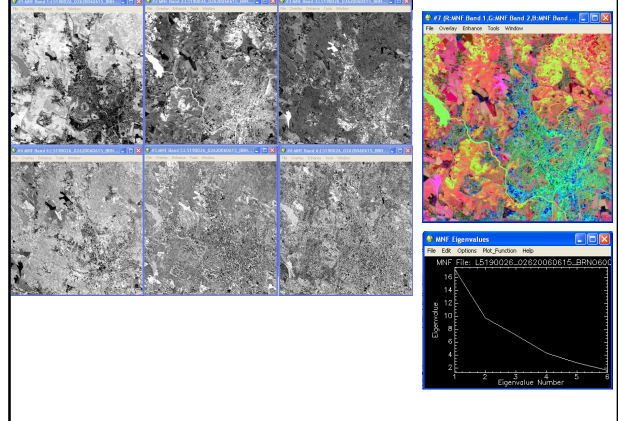
## Výstup PCA – jednotlivá transformovaná pásma



## Výstup PCA – barevná syntéza z PC1, PC2 a PC3



## Minimum Noise Fraction (MNF) transformace



## Transformace "TASSELED CAP"

Lineární transformace původních pásem obrátu. Koefficienty transformačních rovnic jsou určeny empiricky tak, aby každé z nově vypočtených pásem zvýrazňovalo informaci korelující s určitými vlastnostmi půdy a vegetace:

$$TC1 = .2043TM1 + .4158TM2 + .5524TM3 + .5741TM4 + .3124TM5 + .2303TM7$$

$$TC2 = -.1603TM1 -.2819TM2 -.4934TM3 + .7940TM4 -.0002TM5 -.1446TM7$$

$$TC3 = .0315TM1 + .2021TM2 + .3102TM3 + .1594TM4 -.6806TM5 -.6109TM7$$

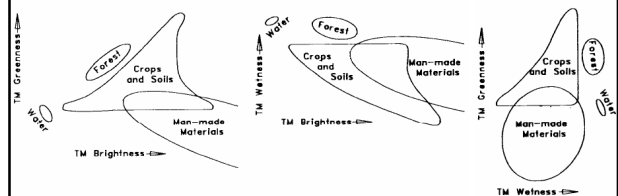
TC1 - orientováno ve směru maximálního rozptylu hodnot odrazivosti půdy a označuje se jako index „**brightness**“,

TC2 - kolmé k TC1 a je orientováno ve směru největšího kontrastu mezi viditelnou a blízkou infračervenou částí spektra. Je tedy ukazatelem množství zelené hmoty - index „**greenness**“.

TC3 - koreluje s vlhkostí půdy a vegetace - index „**wetness**“.

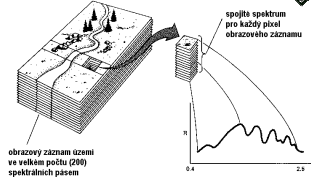
- Jak velký je problém empirického určení koeficientů?
- Jak lze algoritmus využít k identifikaci zastavěných ploch?

## Transformace "TASSELED CAP"



Priznakový prostor vybraných spektrálních indexů transformace Tasseled Cap vypočtená ze senzoru TM (<http://ciesin.org/docs/005-419/fig3.gif>)

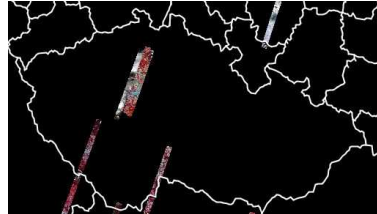
### 3. Metody obrazové spektrometrie (hyperspektrální snímky)



- Kontinuální záznam odrazových vlastností povrchů v optické části spektra
- Možnost podchytit specifika především u povrchů spektrálně „čistých“
- Identifikace jevů, které v krajíně souvisejí se změnami chemického složení
- Identifikace a lokalizace tzv. absorpčních pásů
- Identifikace antropogenních tvarů a procesů
- .....

### Příklady konkrétních systémů - družice

Družice EO-1 (NASA), skener HYPERION – hyperspektrální skener s 242 pásmy v rozsahu 0,4 - 2,5  $\mu\text{m}$  s rozlišením 30 metrů a velikostí scény 7,7 x 42 resp. 185 km, časové rozlišení 16 dní



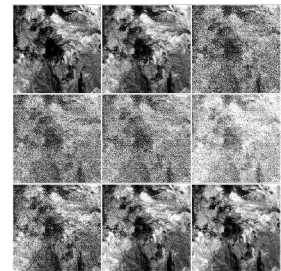
AISA Eagle – [www.usbe.cas.cz](http://www.usbe.cas.cz)  
64 spektrálních pásem, prostorové a spektrální rozlišení přibližně 0,4 m a 10 nm

### Základní etapy analýzy dat obrazové spektrometrie

- předzpracování dat
- kalibrace dat
- tvorba spektrálních knihoven
- vizualizace hyperspektrálních dat a knihoven spekter
- automatické porovnání spekter
- definování elementárních povrchů (tzv. endmembers)
- analýza a automatická klasifikace heterogenních pixelů

### Předzpracování dat

- úprava geometrie snímků, odstranění páskování (stripping), potlačení šumu



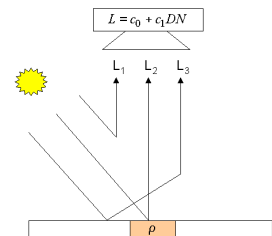
### Kalibrace dat

#### (DN hodnoty – radiance – reflectance)

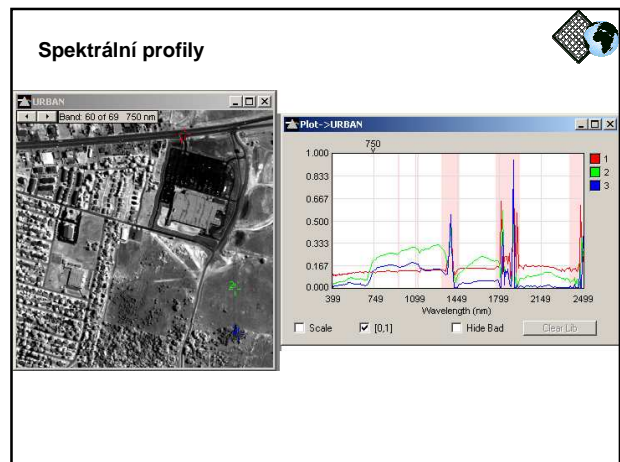
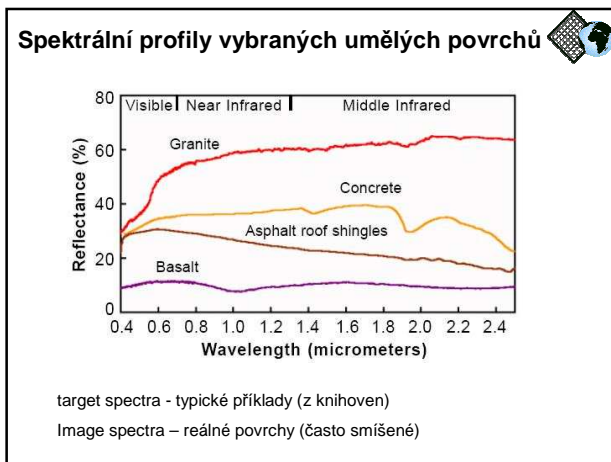
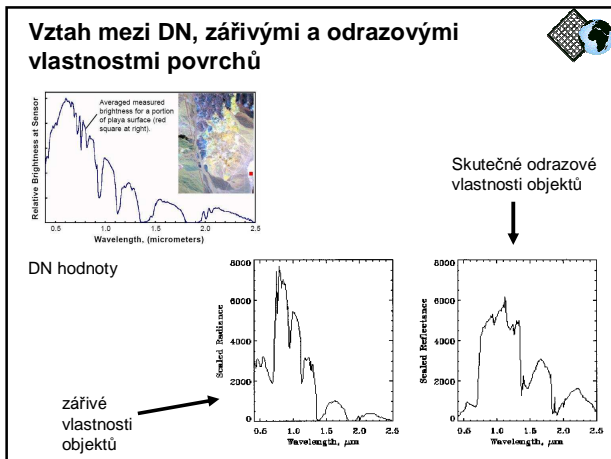
- atmosférické korekce a korekce na vlivy topografie jsou nezbytnou prvotní částí zpracování
- cílem je převést naměřená data, která obsahují charakteristiky celkového vyzařování objektů (angl. **radiance**), na data charakterizující odrazové vlastnosti objektů (angl. **reflectance**)
- **Zář (radiance)** – množství dopadající na plochu čidla (senzor)
- **Odráživost (reflectance)** – podíl mezi množstvím odražené elektromagnetické energie a množstvím energie dopadající na studovaný povrch

### Vztah mezi DN, zářivými a odrazovými vlastnostmi povrchů

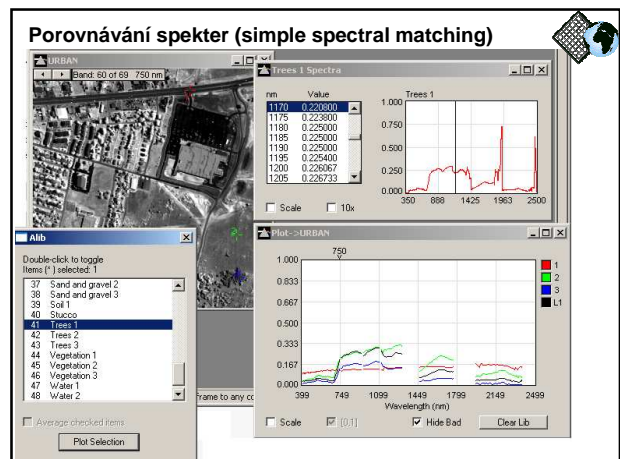
- DN
- L - zář [ $\text{W}\cdot\text{m}^{-2}\cdot\text{sr}^{-1}$ ] (radiance)
- $\rho$  - odrazivost (reflectance)



- L1 : záření rozptýlené atmosférou „path radiance“
- L2 : záření odražené snimaným povrchem
- L3 : záření odražené a rozptýlené okolními objekty
- $c_0, c_1$  : (offset, gain) – kalibrační konstanty



- ### Možné přístupy ke klasifikaci
- Porovnávání spekter (Spectral matching)
  - Klasifikace tzv. „spektrálním úhlem“
  - Analýza smíšených pixelů (mixing)
  - Metody shlukové analýzy, ...
  - Neparametrické metody (NN)
  - Rozhodovací stromy





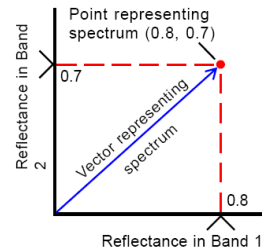
### Porovnávání spekter (simple spectral matching)

- Vyžaduje konverzi obrazových spekter na hodnoty odrazivosti
- Dobře funguje, pokud jsou v obraze větší plochy spektrálně „čistých“ povrchů
- Z vyhovujících spekter je nutné vybírat to nejhodnější – prahování resp. „fuzzy“ výsledek
- Řada povrchů je však smíšených
- Někdy vyžaduje i „mixed“ spektra ve spektrální knihovně
- Výsledek – „material map“ – mapuje dominantní druh povrchu na ploše každého pixelu.

### Spektra jako vektory ve spektrálním prostoru

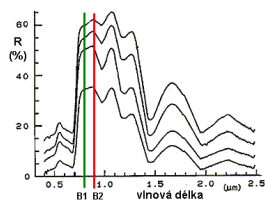
Spektrální profily jsou vhodné pro vizualizaci a porovnávání několika málo spekter

Pro analýzu a klasifikaci většího počtu spekter – N rozměrný spektrální prostor

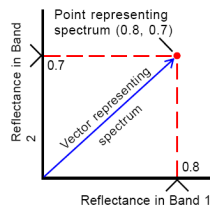


Matematicky lze tento koncept využít i pro hyper – rozměrný prostor  
Není důležitá velikost vektoru, ale jeho orientace

### Klasifikace tzv. „spektrálním úhlem“ (Spectral Angle Mapper)



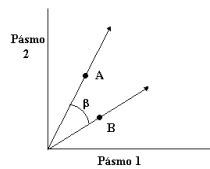
Vektor reprezentující spektrum ve 2D spektrálním prostoru



### Klasifikace tzv. „spektrálním úhlem“

Algoritmus je založen na výpočtu míry podobnosti mezi testovaným spektrálním profilem ze zpracovávaného obrazu a spektrem z knihovny

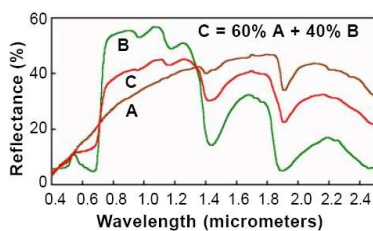
Jako míry podobnosti je využito tzv. **spektrálního úhlu**,



A - vektor známého spektra (např. z knihovny spekter), B - vektor spektra testovaného, β - spektrální úhel; data v použitých pásmech jsou korigována na vlivy atmosféry a zastiňení

### 4. Analýza smíšených pixelů – „linear unmixing“ a hledání spektrálně čistých pixelů

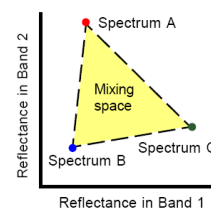
„Tradiční“ přístup - zjednodušující předpoklad, že každý jeden obrazový prvek svoji hodnotou reprezentuje pouze jeden objekt či povrch.



Křivka spektrálního chování heterogenního pixelu (mixel - mixture element) je složena z jednotlivých „spektrálně čistých“ křivek **elementárních pixelů (povrchů)** tzv. **endmembers**.

### Koncept „linear mixing“

Předpoklad – v obraze se nachází konečný, relativně malý počet elementárních pixelů a velké množství smíšených pixelů, které jsou jejich lineární kombinací



Spektra všech povrchů, které jsou kombinací tří elementárních povrchů A, B, C musí ležet uvnitř prostoru, který tyto povrchy vymezují v spektrálním prostoru

Pokud jsme schopni nalézt spektra elementárních povrchů, jsme schopni najít i jejich lineární kombinace

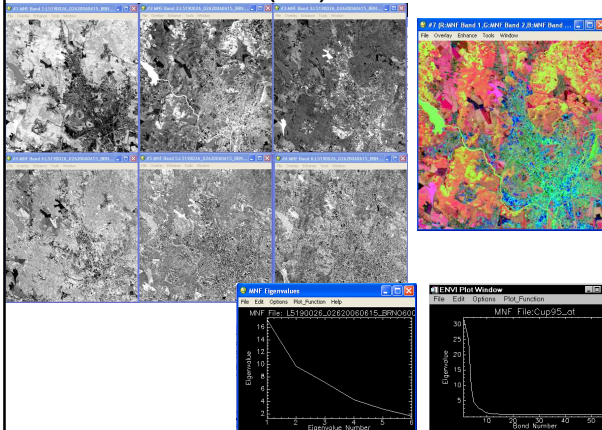
### Minimum Noise Fraction (MNF) transformace

- redukce dimensionality
- odstranění šumu v datech
- první krok k hledání „endmembers“

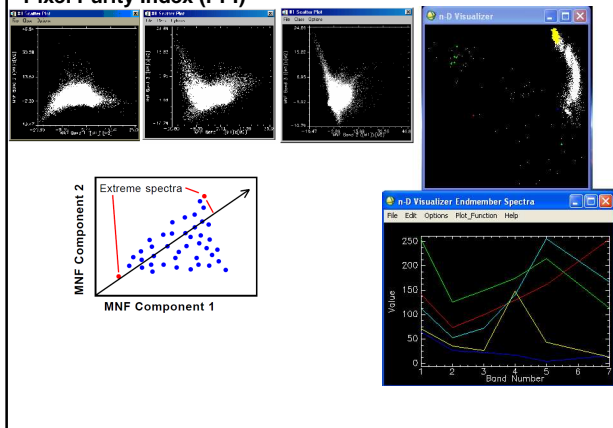
- cílem je odstranit nadbytečnou (redundantní) informaci a potlačit šum v datech
- výsledkem je menší množství tzv. MNF snímků
- MNF aplikuje metodu hlavních komponent ve dvou krocích:

- 1) Odstraní korelaci šumové složky mezi jednotlivými pásmy a transformuje snímky tak, že šum má jednotkový rozptyl
- 2) V nově transformovaných snímcích (MNF) odděluje užitečnou informaci od šumové složky

### Minimum Noise Fraction (MNF) transformace

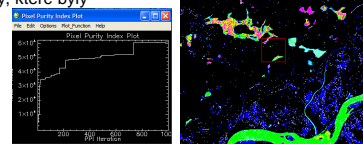
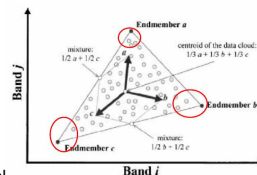


### Pixel Purity Index (PPI)

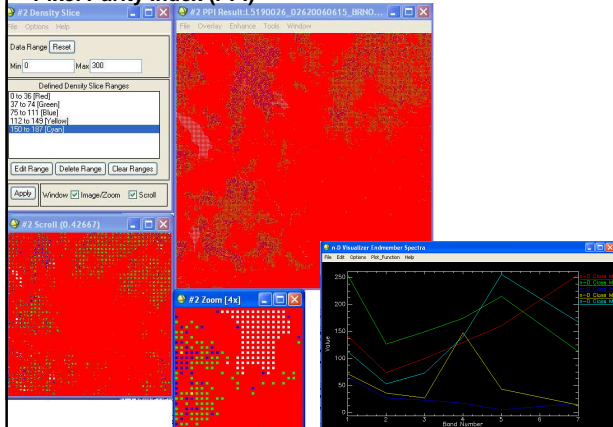


### Pixel Purity Index (PPI)

- Metoda, která pracuje s tzv. smíšenými pixely.
- Hledá ve snímku pixely, které jsou spektrálně co nejčistší
- Nejčistší pixely většinou odpovídají spektrálně odlišným povrchům.
- Metoda hledá, kolikrát se pixel objeví jako extrémní v tzv. simplexu
- PPI generuje snímek, ve kterém je zaznamenáno, kolikrát byla daný pixel extrémní
- Prahováním se oddělí pixely, které byly extrémní nejčastěji

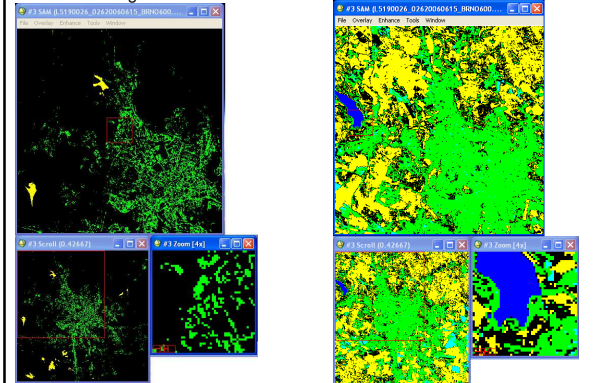


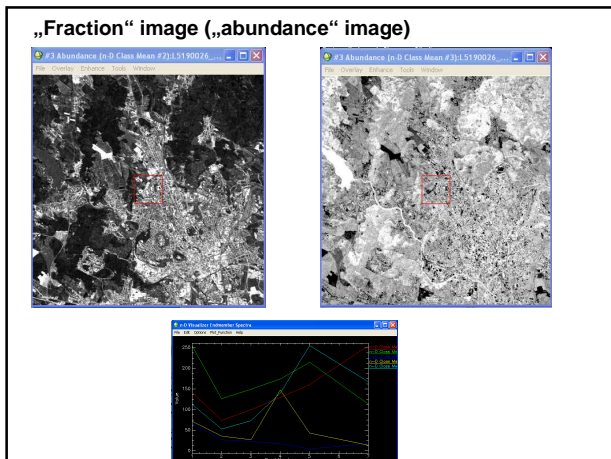
### Pixel Purity Index (PPI)



### Výsledek klasifikace spektrálním úhlem

SAM max. angle = 0.10 SAM max. angle = 0.35





### Možné přístupy k analýze smíšených pixelů

- linear mixing

Fyzikální model

Vychází z předpokladu, že spektrální informace smíšeného pixelu vzniká lineární kombinací spektrálního chování všech obsažených elementárních povrchů

The diagram shows a pixel divided into three regions: A (dark gray), B (medium gray), and C (light gray). To the right, three spectral curves are shown, labeled A, B, and C. Below them is a combined curve labeled R. The equation is  $R = 0,5A + 0,2B + 0,3C$ .

- ### Lineární kombinace spekter (linear mixing)
- Model **lineární kombinace (smíchání) spekter** - tzv. **mixing** - předpokládá, že známe spektra jednotlivých elementárních povrchů i jejich procentuální zastoupení v ploše pixelu.
  - Z těchto informací lze „smíchat“ výsledné spektrum smíšeného pixelu
  - Analýza potom spočívá v obráceném procesu označovaném jako tzv. **un - mixing**,
  - Známe pouze spektrum výsledné, obrazovou analýzou hledáme spektra jednotlivých elementárních povrchů (endmembers)
  - Z modelu „lineárního smíchání“ hledáme procentuální zastoupení jednotlivých elementárních povrchů v analyzovaném obraze.

### Matematický model lineárního smíchání výsledného spektra heterogenního pixelu

$$\begin{matrix} Z \\ m \\ \phantom{m} \\ \phantom{m} \\ n \end{matrix}
 *
 \begin{matrix} X \\ n \\ \phantom{n} \\ 1 \end{matrix}
 =
 \begin{matrix} Y \\ m \\ \phantom{m} \\ 1 \end{matrix}$$

n - počet elementárních povrchů  
m - počet zpracovávaných pásem  
Y - výsledné spektrum  
X - koeficienty určující zastoupení jednotlivých elementárních povrchů  
Z - spektrální chování n elementárních povrchů v m intervalech spektra (pásmech)

### UNMIXING – určení procentuálního zastoupení elementárních povrchů

	Z1	Z2	Y
G	46	62	52,4
R	31	42	35,4
IR	12	160	71,2

m = 3  
n = 2

The graph shows the unmixing process with equations and solutions for  $x_1$  and  $x_2$ . The equations are  $52,4 = 46x_1 + 62x_2$ ,  $35,4 = 31x_1 + 42x_2$ , and  $71,2 = 12x_1 + 160x_2$ . The solutions are  $x_1 = 0,6$  and  $x_2 = 0,4$ .

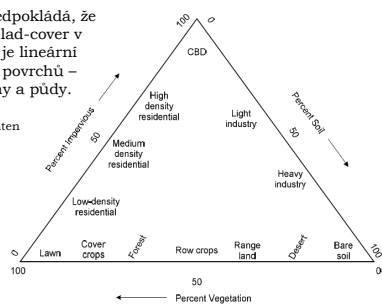
### Linear mixing

The figure shows a color image of a landscape with a red box highlighting a specific area. Below it are three grayscale images labeled "Vegetation fraction", "Water / shade fraction", and "Soil fraction".

### V-I-S (vegetation-impervious surface-soil) model for urban ecosystem analysis

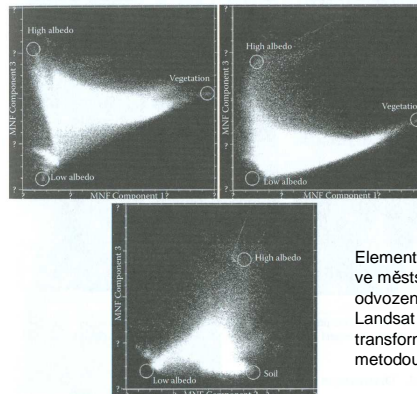
Konceptuální model – předpokládá, že každý pixel představující land-cover v oblasti městské zástavby je lineární kombinací tří základních povrchů – vegetace, zastavěné plochy a půdy.

<http://www.springerlink.com/content/t/u25162w3t3411703/>



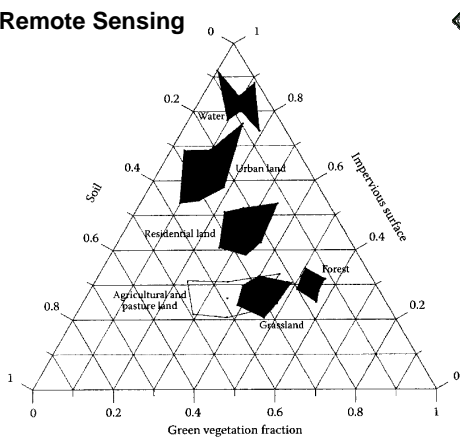
Ridd, M.K., 1995. Exploring a V-I-S (Vegetation-Impervious Surface-Soil) model for urban ecosystem analysis through remote sensing - Comparative anatomy for cities. *Int. J. Remote Sens.*, 16, pp. 2165- 2185.

### Urban Remote Sensing

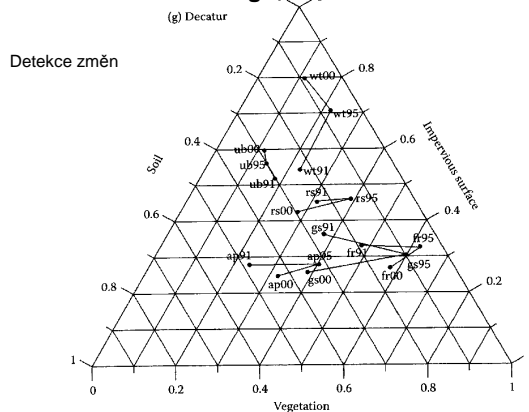


Elementární povrchy ve městské zástavbě odvozené ze snímků Landsat ETM+ transformovaných metodou MNF

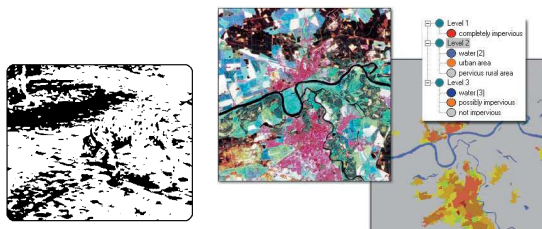
### Urban Remote Sensing



### Urban Remote Sensing



### 5. Objektová („per-object“) klasifikace



Naše rozpoznávání není založeno na postupném skládání celku z jednotlivosti.

Je založeno m.j. na:

1. zkušenosti
2. schopnosti hodnotit vztahy

### Základní východiska

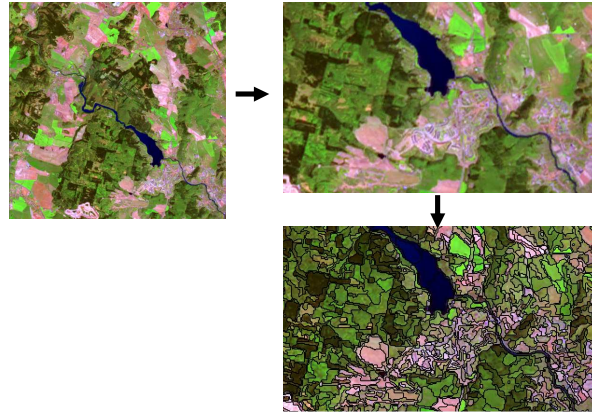
- Klasifikace založená na identifikaci jednotlivých obrazových prvků má mnohá omezení.
- Vychází z předpokladů, které již a priori vylučují úspěšnou aplikaci těchto přístupů na některé úlohy (**zastavěné plochy**).
- Informace uložená v obraze má často **fraktální povahu, hierarchii** – záleží na měřítku. (Části stromu – strom – les – krajinná mozaika).
- Analýza obrazu prozatím málo využívá **jiných charakteristik** (interpretacních znaků) než **spektrálních** (např. na radarová data nelze v důsledku značného podílu šumu použít klasický per-pixel přístup).



## Základní východiska

- Objektový přístup – základní jednotkou pro klasifikaci není obrazový prvek (pixel), ale skupina prostorově souvisejících pixelů (field, image object primitive, ...).
- Tato skupina pixelů je vytvořena procesem **segmentace obrazu**. Jejím cílem je pospojovat pixely podobných vlastností do skupin.
- Nejsou uvažovány jen vlastnosti spektrální, ale například textura, kontext, vlastnosti související s tvarem a velikostí pixelů apod.
- Vytvoření skupin pixelů podobných vlastností umožňuje následně definovat **vztahy sousedství** mezi jednotlivými skupinami

## Příklad- vytváření objektů v obraze



## Atributy objektů

Feature	Val
Layer Values	
Mean	145.88
StdDev	7.27
Form	
Area	2179.00
Length	80.26
Width	27.10
Border Length	386.00
Length/Width	2.96
Shape Index	2.07
Density	2.91
Mean Direction	3.98
Perimetry	0.4951
X-Center	221.96
Y-Center	113.81
Hierarchy	
Level	1.00
Num Higher Levels	0.00
Num Sub Levels	0.00
Num Neighbors	7.00
Num Sub Objects	0.00

Každému z objektů přísluší množina atributů, které popisují spektrální vlastnosti, tvar, topologické vazby, texturní znaky, ...

## Obecný postup objektové klasifikace obrazu

1. Spojování podobných pixelů do homogenních ploch – **segmentů**
2. **Testování** homogenity segmentů
3. Výpočet **atributů** pro každý segment
4. **Klasifikace** segmentů (objektů)

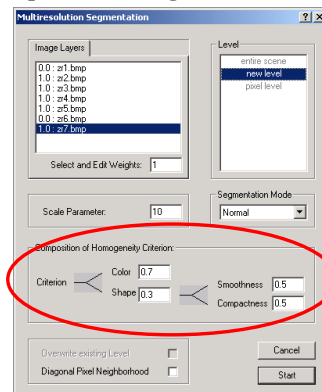
## Segmentace obrazu

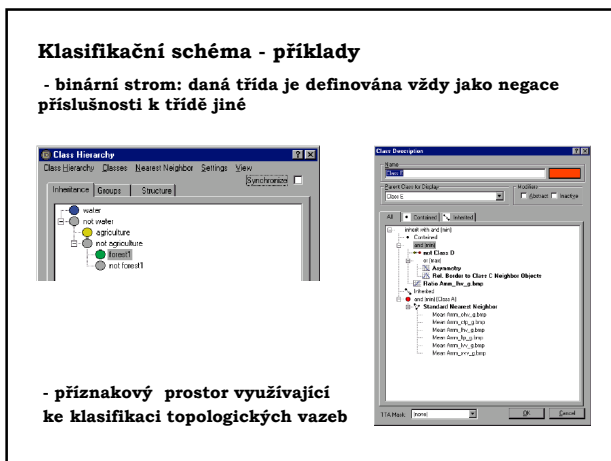
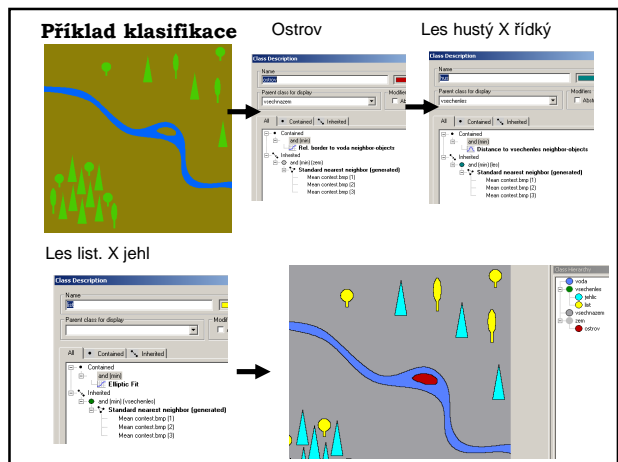
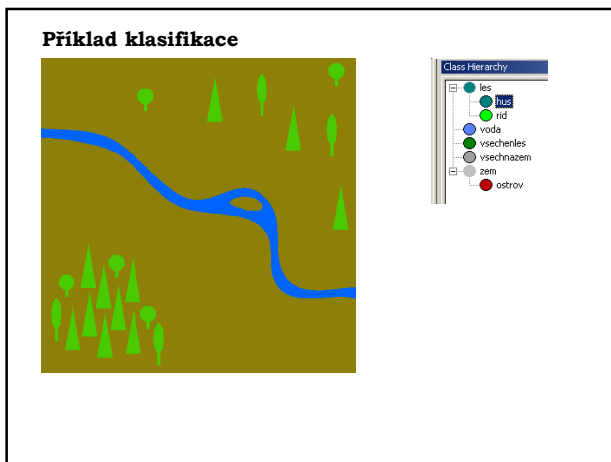
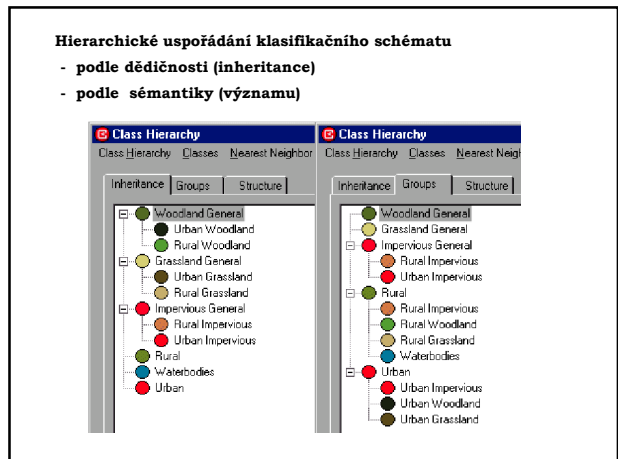
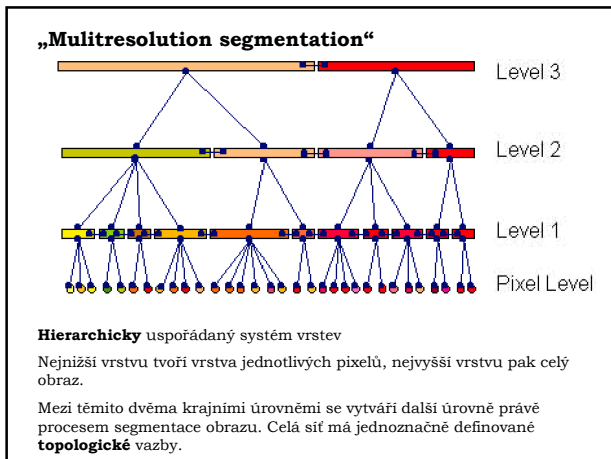
Definování základních obrazových objektů je založeno na spojování podobných pixelů.

Podobnost či homogenita je posuzována z hlediska těchto tříd informací:

- Spektrální informace
- Texturní informace
- Informace o tvaru objektů
- Informace o topologických vztazích (kontextuální)

## Nastavení parametrů segmentace obrazu



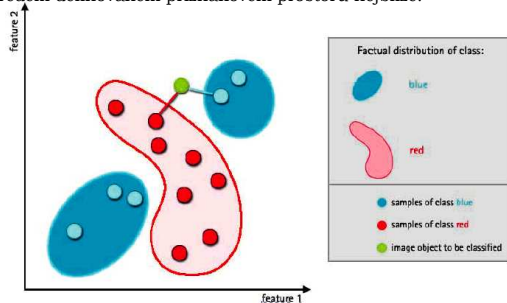


### Klasifikace objektů I.

Vlastní klasifikace může být založena na **klasifikátoru nejbližšího souseda (Nearest Neighbor)**

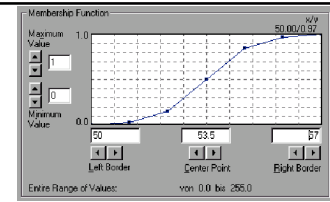
Trénovací data tvoří vybrané objekty

Klasifikátor zařadí všechny ostatní objekty do třídy, ke které má v předem definovaném příznakovém prostoru nejbližší.



### Klasifikace objektů II.

Klasifikace založená na principu **neostrých množin (fuzzy logic)**.



Příslušnost jednotlivých objektů ke každé třídě je hodnocena prostřednictvím funkce příslušnosti (membership function) pro každý z uvažovaných atributů.

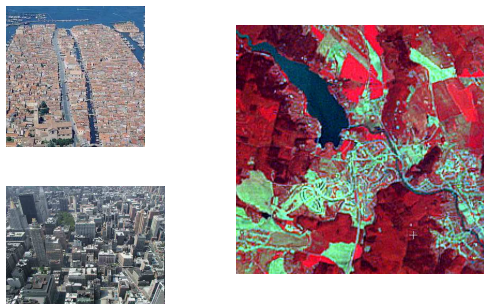
Membership function normalizuje hodnoty jakéhokoliv použitého atributu (např. DN hodnot pásma obrazu 0 až 255) do hodnot 0 až 1.

Funkce má různý průběh (např. sigmoida).

Průběh membership function definuje neostře (fuzzy) hranice a nahrazuje binární logiku (patří - nepatří).

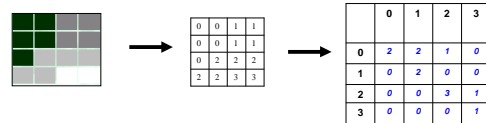
### 6. Hodnocení a klasifikace textury na snímcích

Texturní klasifikátory se snaží popsat typickou **proměnlivost** spektrálního chování



### Hodnocení a klasifikace textury na snímcích

GLCM (Grey Level Co-occurrence Matrix)



Každý prvek GLCM matice nese informaci, kolikrát se daná kombinace hodnot v okně vyskytuje.

Míry textury - **vážený průměr** buněk GLCM.

$$\text{Kontrast} = \sum_{i,j=0}^{N-1} P_{i,j} \cdot (i - j)^2$$

Interpretace: Je-li  $i$  a  $j$  stejné (na diagonále) váha je 0. Liší-li se  $i$  a  $j$  o 1 váha je 1, liší-li se o 2 váha je 4 atd. Váhy exponenciálně rostou.

### 7. Klasifikace obrazu metodami strojového učení

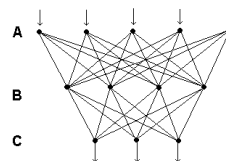
- **Neuronové sítě**
- **Rozhodovací stromy**

#### Neuronové sítě

- počítačová architektura, která se snaží napodobit procesy probíhající v nervové soustavě
- je nezávislá na statistickém rozložení dat
- je odolná proti chybám, má schopnost učit se (asociativní učení), dovede abstrahovat i generalizovat
- dokáže odhadnout nelineární vztah mezi vstupními a požadovanými výstupními daty
- umožňuje v procesu klasifikace kombinovat různé typy vstupních dat.

#### Základní pojmy

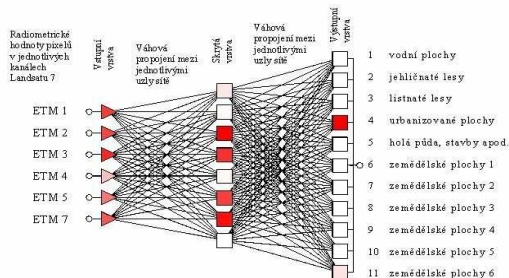
- neuron – výkonný prvek NS
- synapse – spojení neuronů
- váhové koeficienty neuronů
- adaptivní fáze – učící
- aktivní fáze – vybavovací



Schematizované uspořádání neuronové sítě A - vstupní vrstva, B - skrytá vrstva, C - výstupní vrstva

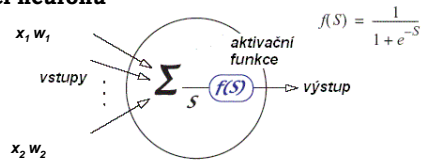
**(multi-layer perceptron)**

### NN pro klasifikaci multispektrálního snímku



Příklad třívrstvé perceptronové sítě se 6 vstupními, 8 skrytými a 11 výstupními uzly (MLP 6-8-11) s příkladem vstupu a výstupu při klasifikaci družicových snímků s vyznačenými aktivacemi zastavěné plochy

### Model neuronu



kde  $x_i$  - hodnota  $i$ -tého vstupu  
 $w_i$  - váha  $i$ -tého vstupu

$$S(x) = \sum_i w_i x_i$$

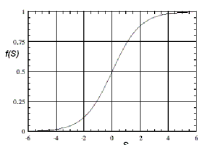
Neuron provádí tři akce:

- sumuje vstupy z jiných neuronů
- provádí prahování
- posílá výstup do jiných neuronů

### Model neuronu

- Neuron v síti jsou propojeny tzv. **váhovými koeficienty**, které zesilují nebo zeslabují signál přicházející z předchozích neuronů.
- Suma těchto vážených signálů určuje aktivaci neuronu, která ovlivňuje další výstup z neuronu.
- Výstup z neuronu je funkcí této aktivace, kdy výstup je vypočten na základě logistické aktivací funkce - **sigmoida**.
- Výstup z uzlu je realizován pouze překračuje-li určitou prahovou hodnotu.
- Váhové koeficienty jsou na počátku náhodnými čísly

Příklad aktivací funkce neuronu (sigmoida)



### Učící - adaptivní fáze

- **Učení** – váhy na spojích mezi jednotlivými výkonnými prvky sítě se mění podle určitého tzv. učícího algoritmu.
- Použije se trénovací soubor, ve kterém známe správné zařazení pixelů do jednotlivých klasifikačních tříd.
- **Učící algoritmus** – předpis, podle kterého se předkládají síti vzory k učení a podle kterého se mění váhy jednotlivých spojení mezi neurony.
- **Učení se s učitelem** – analogie řízené klasifikace Algoritmus zpětného šíření (**Back propagation**)
- **Učení se bez učitele** – analogie shlukové analýzy (neřízené klasifikace) – **samo-organizující** se sítě (Self-organizing) – Kohonenova síť
- **Ustálení** – dosažení stabilního stavu sítě

### Algoritmus zpětného šíření (**Back propagation**)

- signály nejprve vyšlou směrem dopředu
- u výstupních neuronů se porovnávají výstupy s požadovanými
- zjištěné chyby se použijí ke změně nastavení vah v síti

### Výstup ze sítě má formu vektoru:

Očekávaný výstup (z učící množiny):  $o = (0,0,0,1,0,0)$

Aktuální výstup:  $a = (0,1,0,0,0,0)$

**Hodnocení úspěšnosti učení** – chyba  $e$ :

$$e = \frac{1}{2} \sum_{j=1}^k (o_j - a_j)^2$$

- Učení probíhá iteračním způsobem – cílem je dosáhnout nulové či minimální akceptovatelné chyby.
- Adaptace sítě – úprava vah synapsí - probíhá po krocích.
- Délka kroku se nazývá **learning rate**.
- Velká délka kroku značí rychlejší, ale méně přesné učení



### Přednosti neuronových sítí

- nezávislost na statistickém rozdělení
- schopnost generalizace
- síť je tolerantní k šumu v učicích datech

### Nedostatky neuronových sítí

- problém návrhu architektury sítě (počet skrytých vrstev a počet neuronů v nich)
- dlouhá doba učení
- problém lokálního minima (oscilace)
- nastavení úvodních (náhodných) vah synapsí

### Klasifikace Rozhodovacími stromy (Decision Trees)

#### Základní pojmy:

- Název atributu - uzel
- Aritmetický či logický výraz - větve stromu
- Název třídy - list stromu

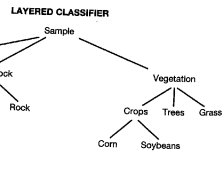
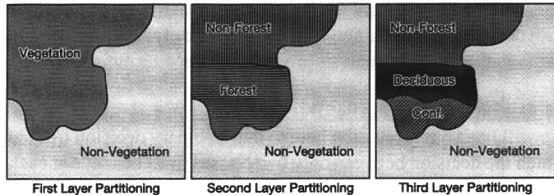
**TM3 <= 35 :**  
 | **TM4 > 99 : polev (12.0)**  
 | **TM4 <= 99 :**  
 | | **TM5 > 58 : lesi (30.0/1.0)**  
 | | | **TM5 <= 58 :**  
 | | | | **TM6 <= 12 : lesi (2.0)**  
 | | | | **TM6 > 12 : lesj (8.0)**  
**TM3 > 35 :**  
 | **TM6 <= 23 : voda (17.0)**  
 | **TM6 > 23 : poleb (26.0)**

- Mohou ale i nemusí být založeny na binárním třídění, jsou neparametrické.
- Umožňují testování sousedů - grafy sousednosti a topologických vazeb (meet, contain, overlap, ...)
- Umožňují testování atributů různé povahy
- Možnost klasifikace po vrstvách - hierarchické třídění („layered classification“)
- Problém objektivního hodnocení („ground truth“)

### Klasifikace rozhodovacími stromy

- Layered classification
- CART – Classification and Regression Tree

Klasifikace po vrstvách či hierarchická klasifikace



### CART – Classification and Regression Tree

TABLE 1. LAND-USE/LAND-COVER CLASSIFICATION SCHEME USED FOR THE STUDY

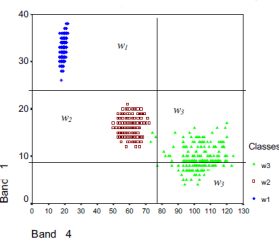
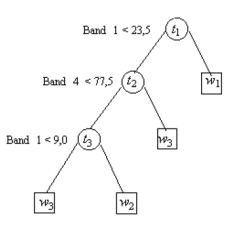
Level 1	Level 2	Level 3
Urban	Urban	Urban
Agriculture	Agriculture	Perennial agriculture
		Annual agriculture
Natural vegetation	Conifer	Mixed conifer
		Douglas-fir
	Conifer/herbaceous mix	Conifer/herbaceous mix
	Burned	Burned
	Hardwood	Aspen
		Willow
		Cottonwood
Hardwood/herbaceous mix	Hardwood/herbaceous mix	Hardwood/herbaceous mix
		Sage-grassland
		Grassland

TABLE 2. DICHOTOMOUS CLASSIFICATION TREE FOR LEVEL 1 CLASSIFICATION BASED ON CART ANALYSIS

- Tasseled Cap brightness difference < 22.5
  - Elevation < 1682 m
    - June TM band 6 < 139.5
      - June TM band 4 < 111, THEN Natural vegetation
      - June TM band 4 > 111, THEN Natural vegetation
    - June TM band 6 > 139.5
      - June TM band 2 < 32.5, THEN Agriculture
      - June TM band 2 > 32.5, THEN Agriculture
  - Elevation > 1682 m
    - June TM band 6 < 116.5, THEN Natural vegetation
    - June TM band 6 > 116.5, THEN Natural vegetation
- Tasseled Cap brightness difference > 22.5
  - Elevation < 1794 m
    - Tasseled Cap greenness difference < 1
      - June TM band 4 < 116.5
        - August TM band 5 < 138.5
          - June TM band 5 < 93
            - June TM band 1 < 66.5, THEN Agriculture
            - June TM band 1 > 66.5, THEN Natural vegetation
          - June TM band 5 > 93, THEN Agriculture
        - August TM band 1 > 77.5, THEN Natural vegetation
        - August TM band 1 < 77.5, THEN Non-vegetation
      - August TM band 5 > 118.5, THEN Agriculture
      - August TM band 5 < 118.5, THEN Agriculture
    - June TM band 4 > 116.5, THEN Agriculture
    - June TM band 4 < 116.5, THEN Agriculture
  - August TM band 1 < 77.5, THEN Agriculture
  - August TM band 1 > 77.5, THEN Non-vegetation
- Elevation > 1794 m
  - June TM band 5 < 101.5, THEN Natural vegetation
  - June TM band 5 > 101.5, THEN Natural vegetation

(Lawrence, Wright 2001)

### CART – Classification and Regression Tree



(Bittencourt, Clarke 2003)