
Transformácie obrazu

Lekcia 4

Obsah lekcie

1. Analýza hlavných komponentov (PCA)
 2. Analýza kanonických komponentov (CCA)
 3. Vegetačné indexy - pomerové a vzdialenostné
 4. Tasseled Cap transformácia
-

PCA

Analýza hlavných komponentov (Principal Component Analysis, PCA)

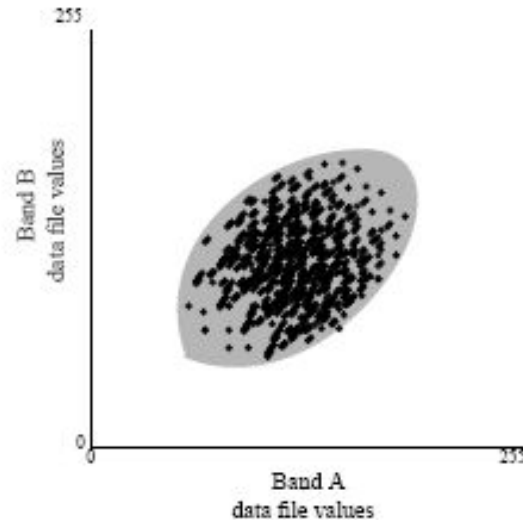
- štatistická procedúra, ktorá sa používa na redukciu veľkého počtu vzájomne súvisiacich premenných na niekoľko vzájomne nezávislých premenných - hlavných komponentov
 - pri čo najmenšej strate informácií (variability)
 - vynájdená nezávisle Karlom Pearsonom (1901) a Haroldom Hotellingom (1930)
-

PCA

- hlavné komponenty sú lineárnymi kombináciami (váženým priemerom) pôvodných premenných
 - prvý komponent vysvetľuje najväčšiu časť variability, druhý menšiu časť atď., až kým nie je vysvetlená všetka
 - komponenty sú vzájomné lineárne nezávislé (nekorelované)
 - vzájomná variabilita premenných sa počíta pomocou variančno-kovariančnej a korelačnej matice
-

PCA

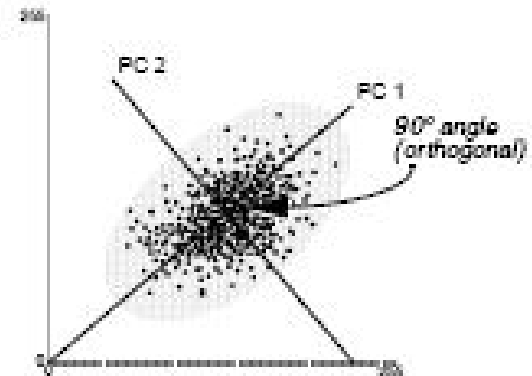
- v DPZ často používaná metóda kompresie údajov, ktoré sú redundantné (multispektrálne pásma v DPZ)
- buď na vylepšenie pred vizuálnou interpretáciou alebo ako predspracovanie pred klasifikáciou



PCA

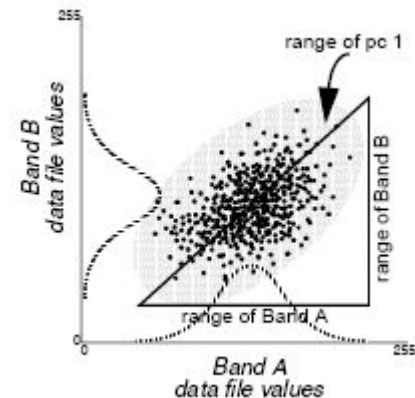
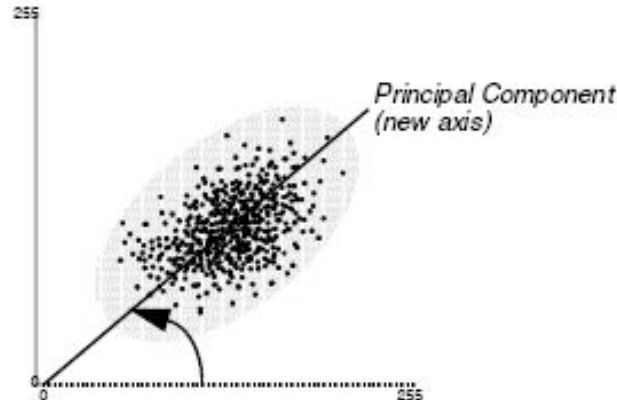
- ortogonálna transformácia n-rozmerného spektrálneho priestoru, pri ktorej sú osi rotované, pričom nové osi I a II sú rovnobežné s osami elipsy a ich počiatok je v priemere rozloženia údajov
- os I definuje smer prvého hlavného komponentu PC1, os II smer druhého hlavného komponentu PC2

Figure 6-21: Second Principal Component



PCA

- údaje pozdĺž smeru osi I majú väčší rozptyl ako údaje pozdĺž originálnych osí A a B
- os II je kolmá na os I – údaje sú nekorelované
- tak to pokračuje aj pri ďalších komponentoch (PC3, PC4 až PCn)



PCA

- PC1 obsahuje najväčší podiel celkového rozptylu scény a nasledujúce komponenty obsahujú postupne menší a menší podiel celkového rozptylu
 - to umožňuje efektívnu redukciu počtu pásiem
 - napr. Landsat TM, ETM, OLI - 6 pásiem (VI+NIR+SWIR)
 - PCA - 2 pásma
-

PCA

- matematicky - lineárna ortogonálna transformácia:

$$DN_I = a_{11}DN_A + a_{12}DN_B$$

$$DN_{II} = a_{21}DN_A + a_{22}DN_B$$

Kde

DN_I, DN_{II} – DN v novom súradnicovom systéme

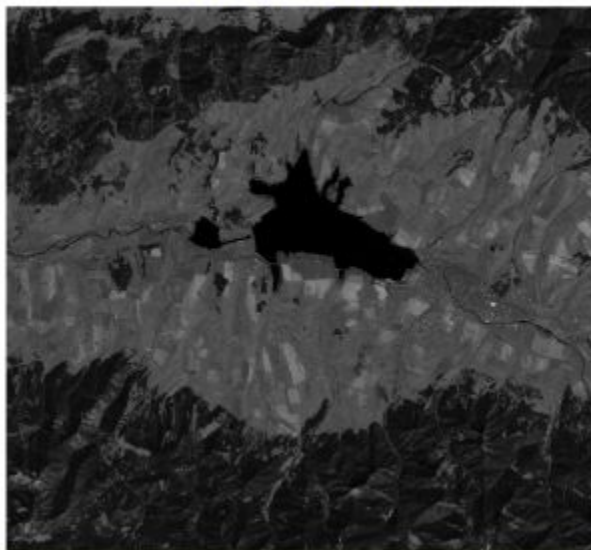
DN_A, DN_B – DN v pôvodnom súradnicovom systéme

$a_{11}, a_{12}, a_{21}, a_{22}$ – koeficienty transformácie

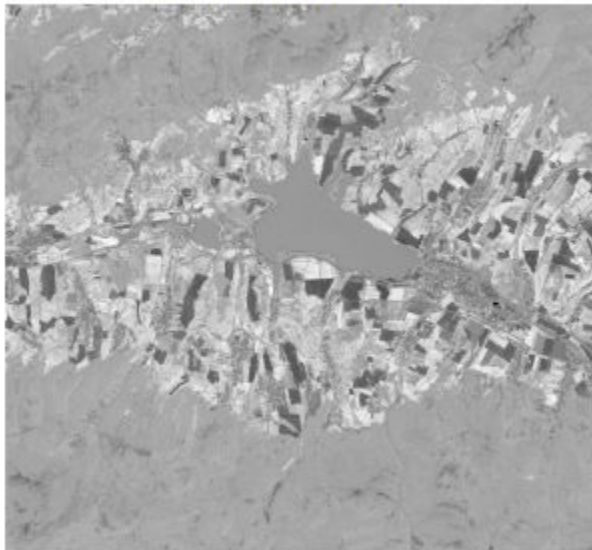
PCA - výsledky

- rastre PC1, PC2 až PCn (podľa počtu vstupných pásiem)

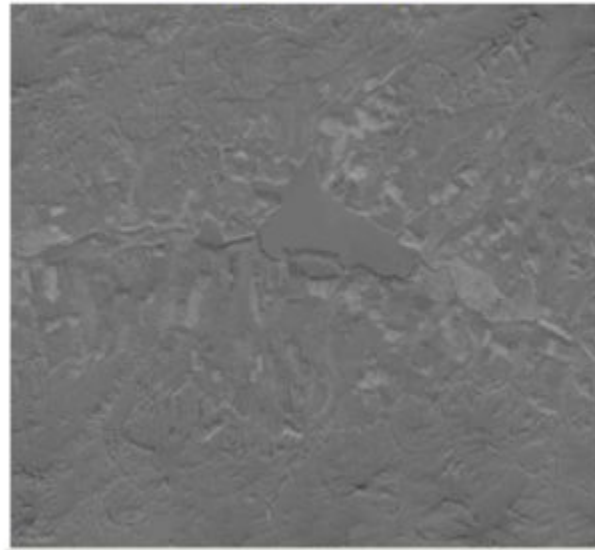
Component 1



Component 2



Component 3



PCA - výsledky

- variančno-kovariančná matica
- **kovariancia** je štatistická miera zhody rozptylu dvoch náhodných premenných, meria rozsah ich spoločnej zmeny

$$\text{cov}(x, y) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

- + x a y sa menia jedným smerom
- X a y sa menia opačným smerom
- 0 X a y sa menia nezávisle

x_i, y_i - i -ta hodnota premennej x, y

$\bar{x}, \bar{y}$ - priemerná hodnota premennej x, y

- kovariancia premennej samej so sebou je variancia (rozptyl)
-

PCA - výsledky

- korelačná matica
- **korelácia** je štatistická miera závislosti dvoch náhodných premenných
- Pearsonov korelačný koeficient - meranie lineárnej závislosti

$$r = \frac{\text{cov}(x, y)}{\sigma_x, \sigma_y}$$

Sir Francis Galton (1896)

σ_x, σ_y - štandardná odchýlka premennej x, y

- nadobúda hodnoty od -1 do 1
-

PCA - výsledky

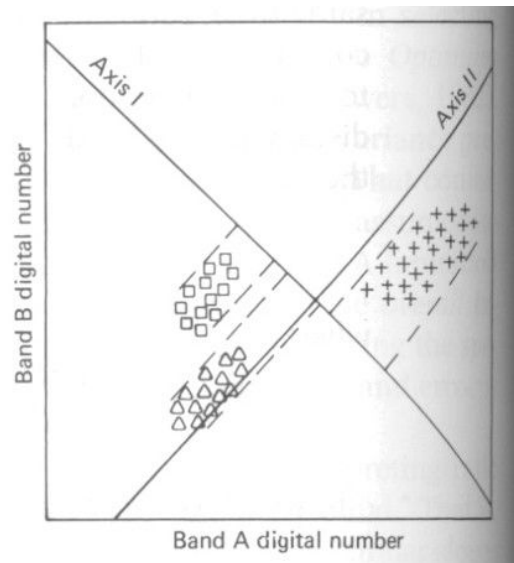
- **vlastné vektory (eigenvectors)** - podiel rozptylu z jednotlivých vstupných pásiem v komponentoch PC1 až PCn
- **vlastné čísla (eigenvalues)** - koeficienty transformácie, podiel rozptylu obsiahnutého v komponentoch PC1 až PCn

$$p = \frac{\text{eigenvalue } PC_i}{\sum_{i=1}^n \text{eigenvalue } PC_i}$$

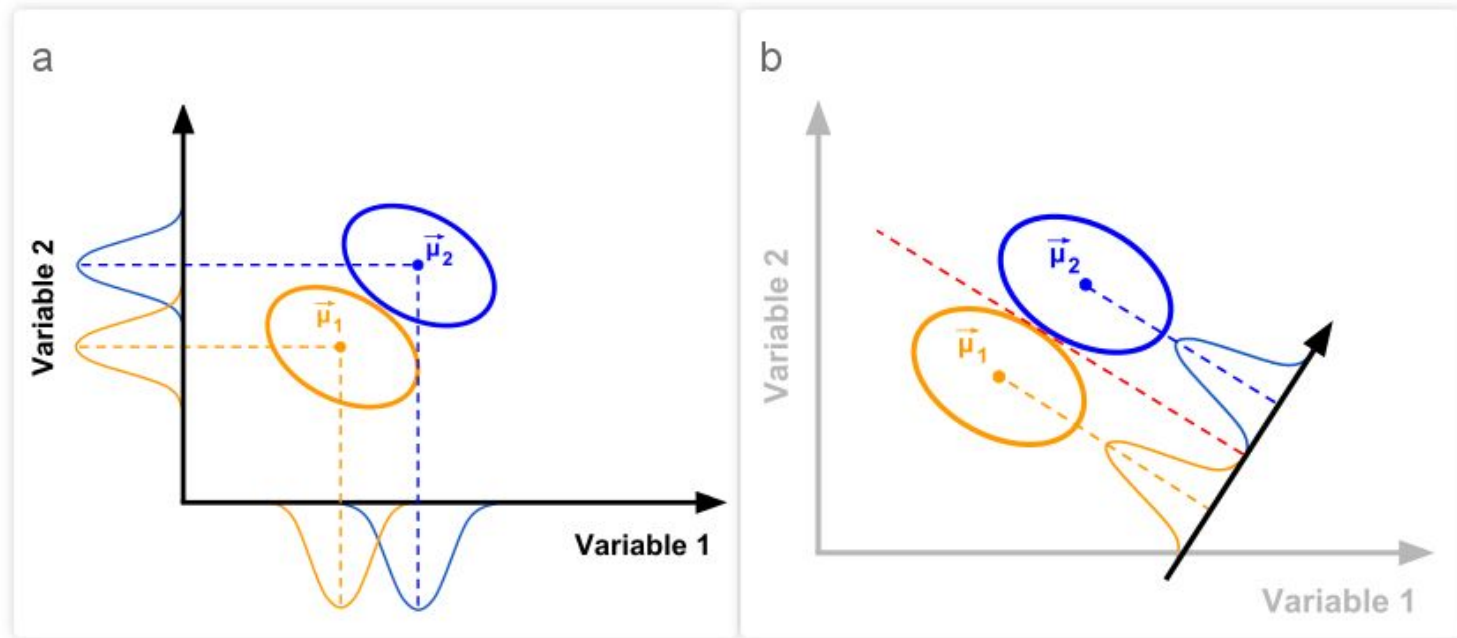
CCA

Analýza kanonických komponentov (Canonical Component Analysis, CCA)

- maximalizácia oddeliteľnosti tried
a minimalizácia rozptylu vnútri tried
- patrí do skupiny metód lineárnej
diskriminačnej analýzy (LDA)
- vychádza z kanonickej korelácie



CCA



CCA

- medzi-triedna kovariancia (between-class covariance):

$$cov_{BC} = \frac{1}{C} \sum_{i=1}^C (\mu_i - \bar{\mu})(\mu_i - \bar{\mu})^T$$

μ_i - priemerná hodnota triedy i

$\bar{\mu}$ - priemerná hodnota priemerov tried

C - počet tried

CCA

- vnútro-triedna kovariancia (within-class covariance):

$$cov_{WC} = \frac{\sum_{i=1}^C (n_i - 1) \cdot cov_i}{\sum_{i=1}^C n_i}$$

n_i - počet vektorov v triede i

cov_i - kovariancia vektorov v triede i

CCA

- pre výsledné komponenty platí:

$$(cov_{BC} - \lambda_j \cdot cov_{WC}) \cdot c_j = 0$$

c_j - j -ty kanonický komponent

λ_j - eigenvalue j -teho kanonického komponentu

- normalizačná podmienka:

$$c_j^T \cdot cov_{WC} \cdot c_j = 1$$

Vegetačné indexy

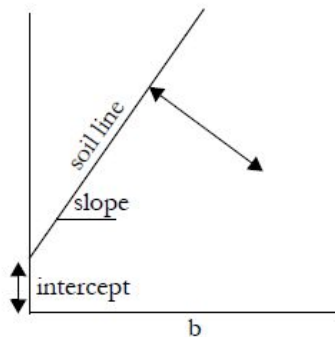
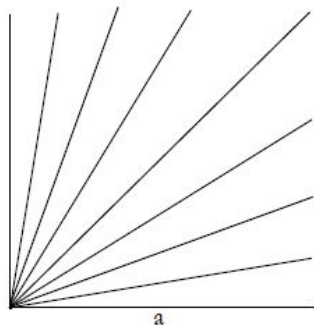
- rastlinné pigmenty (hlavne chlorofyl) absorbujú pre účely fotosyntézy slnečnú energiu vo viditeľnom pásme – oblasť tzv. **fotosynteticky aktívneho žiarenia (PAR – Photosynthetic Active Radiation)** – cca 400 až 700 nm
 - najviac absorpcie v R a B pásme, menej v G pásme
 - na jeseň sa chlorofyl degraduje, ale prídavné pigmenty (karotény, xantofyly) nie, preto sú listy žlté, oranžové, červené...
-

Vegetačné indexy

- energia v **blízkom IČ pásme** sa na fotosyntézu nepoužíva, preto ju listy odrážajú (700 až 1100 nm)
 - čím viac listov, tým väčší odraz
 - na veľkom kontraste v odraznosti medzi R a NIR pásmom sú založené **vegetačné indexy**
 - vysoké hodnoty - zdravá vegetácia
 - kvantitatívne meranie zelenej biomasy
 - biofyzikálne merania
 - ukazovatele celkovej environmentálnej zmeny
-

Vegetačné indexy

- delenie vegetačných indexov:
1. **pomerové (slope-based)** – aritmetické kombinácie pásiem
 2. **vzdialenostné (distance-based)** – rozdiel odraznosti pixla od odraznosti holej pôdy



Pomerové vegetačné indexy

Simple Ratio (SR) – Birth and McVey (1968)

$$SR = \frac{\rho_R}{\rho_{NIR}}$$

- citlivý na delenie nulou
- nelineárna mierka

Normalized Difference Vegetation Index (NDVI) – Rouse et al. (1974)

$$NDVI = \frac{\rho_{NIR} - \rho_R}{\rho_{NIR} + \rho_R}$$

- lineárna mierka
 - hodnoty od -1 do 1
-

NDVI

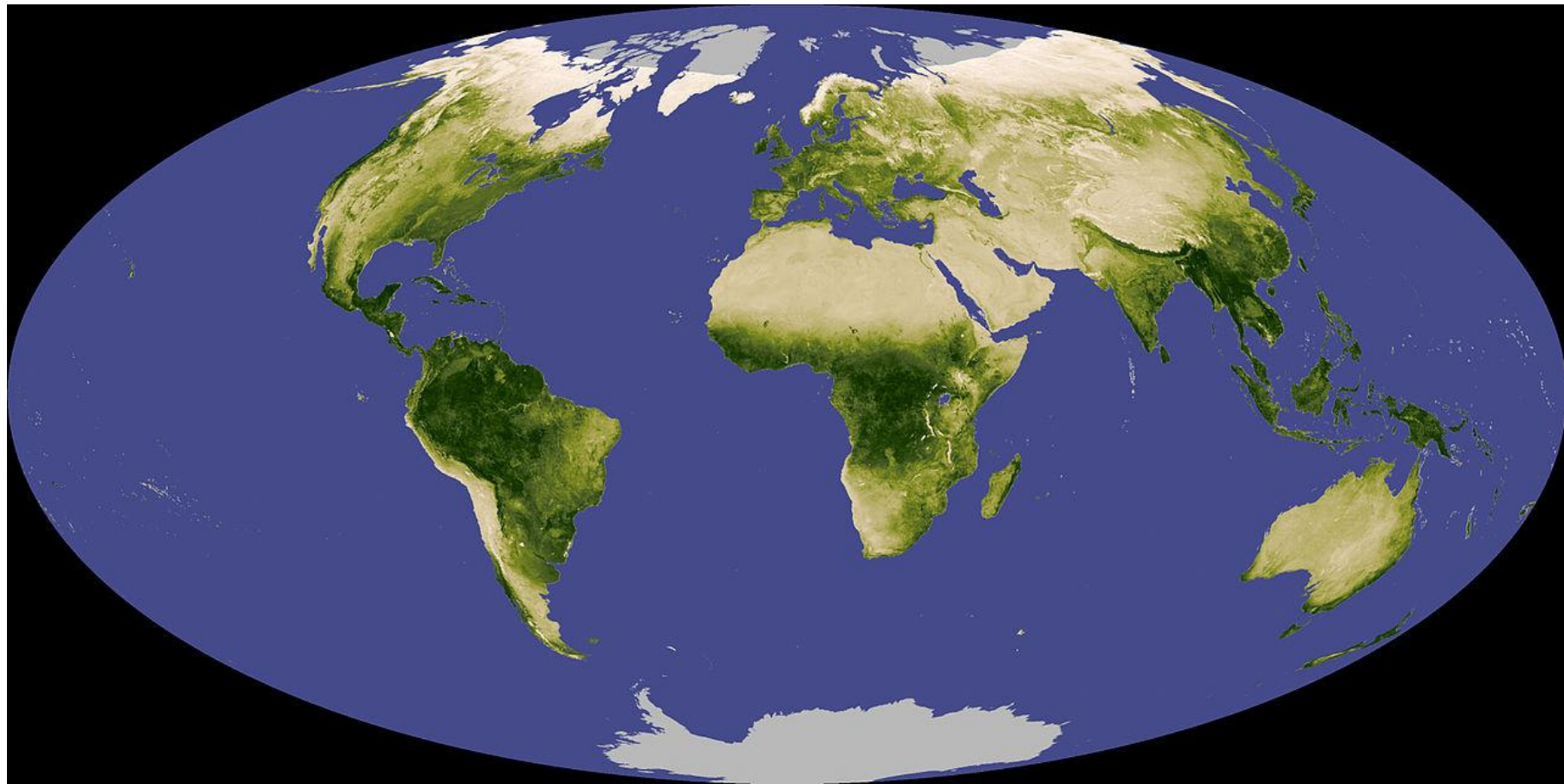
vegetácia – pozitívne hodnoty (0,3 až 0,8)

oblaky, sneh – negatívne hodnoty

voda – veľmi nízke pozitívne až negatívne hodnoty

pôda – nízke pozitívne hodnoty

- odhad Leaf Area Index (LAI), biomasy, koncentrácie chlorofylu v listoch, produktivity rastlín, frakčného pokryvu vegetácie, akumulovaných zrážok a ďalších parametrov
 - koreláciou hodnôt NDVI s pozemnými meraniami
-



Zdroj: NASA, wikipedia

EVI

Enhanced Vegetation Index (EVI) - Huete et al. (1997)

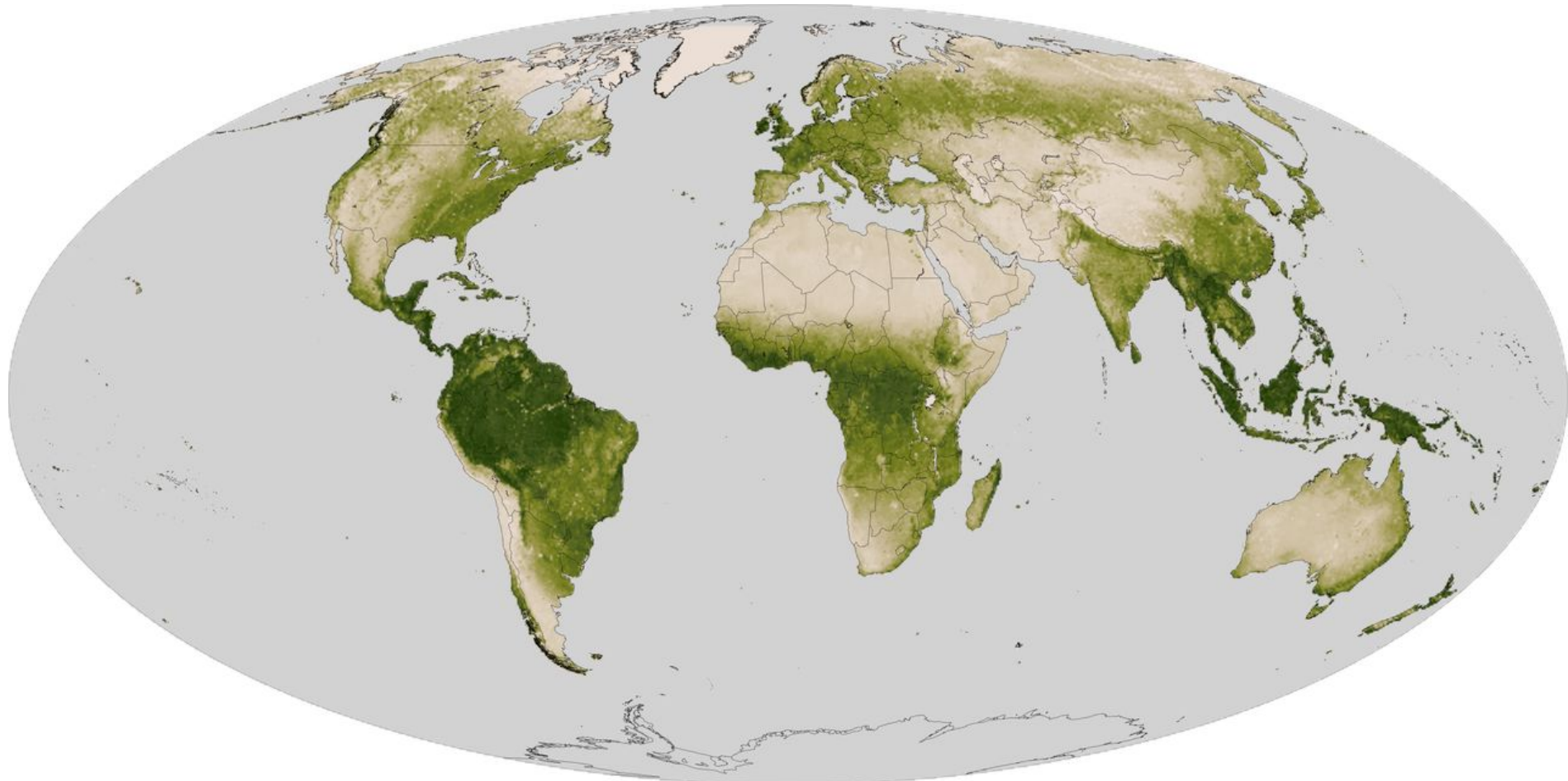
- lepšia citlivosť v oblastiach vysokej biomasy
- redukcia vplyvov atmosféry a vegetačného podkladu
- nevýhoda - potreba modrého pásma

$$EVI = G \frac{\rho_{NIR} - \rho_{red}}{\rho_{NIR} + C_1 \times \rho_{red} - C_2 \times \rho_{blue} + L}$$

G - gain factor

C_1, C_2 - korekcia vplyvov aerosolov

L - prispôsobenie vegetačnému podkladu



Enhanced Vegetation Index 2011



Zdroj: NASA USGS

NDBI

Normalized Difference Built-up Index (NDBI) - Zha et al. (2003)

- urbanizácia a rozširovanie zastavaných plôch
- spojenie Infrared Index-u (II) a NDVI

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

$$Built-up_{area} = NDBI - NDVI$$

Pomerové vegetačné indexy

Transformed Vegetation Index (TVI)

Corrected Transformed Vegetation Index (CTVI)

Thiam's Transformed Vegetation Index (TTVI)

Ratio Vegetation Index (RVI)

Normalized Ratio Vegetation Index (NRVI)

...

Vzdialenostné vegetačné indexy

- hlavným cieľom je eliminácia vplyvu jasů pôdy v oblastiach riedkej vegetácie, keď pixel obsahuje zmes zelenej vegetácie a pôdneho podkladu
- dôležité najmä v suchých a polosuchých oblastiach
- **pôdna línia (soil line)** – lineárna regresia NIR a R pásma na pixloch holej pôdy

$$Y = a + b.X$$

a – posun (v smere osi Y)

b – sklon

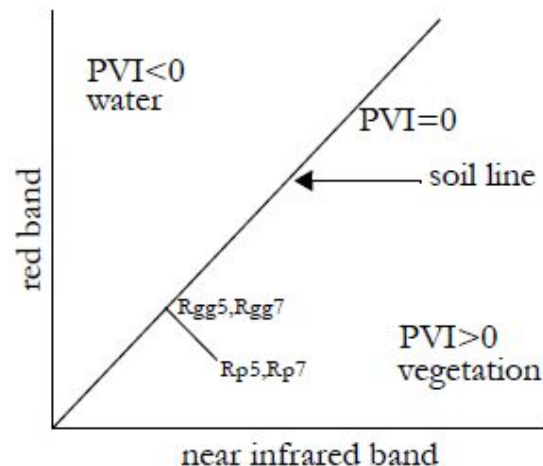
Vzdialenostné vegetačné indexy

- jedna skupina VI berie ako nezávislú premennú R pásma, druhá skupina NIR pásma

Perpendicular Vegetation Index (PVI)

Richardson a Wiegand (1977)

1. Výpočet regresnej priamky holej pôdy (pôdnej línie)
2. Výpočet rovnice priamky kolmej k pôdnej línii
3. Nájdenie priesečníka priamok
4. Výpočet vzdialenosti pixla od pôdnej línie



Vzdialenostné vegetačné indexy

vylepšenia PVI: **PVI1, PVI2, PVI3**

Difference Vegetation Index (DVI)

Ashburn Vegetation Index (AVI)

Soil-Adjusted Vegetation Index (SAVI)

Transformed Soil-Adjusted Vegetation Index (TSAVI1, TSAVI2)

Modified Soil-Adjusted Vegetation Index (MSAVI1, MSAVI2)

Weighted Difference Vegetation Index (WDVI)

Zdroje VI

- produkty vegetačných indexov:

1. AVHRR (družice NOAA) – **NDVI**

rozlíšenie 1 km USA, 8 km svet (od r.1982)

2. MODIS (družice Terra a Aqua) – **NDVI, EVI**

rozlíšenie 250 m, 500 m, 1 km (od r.2000, 8-dňové a 16-dňové kompozity)

3. VIIRS (družica NPP) - **NDVI, EVI**

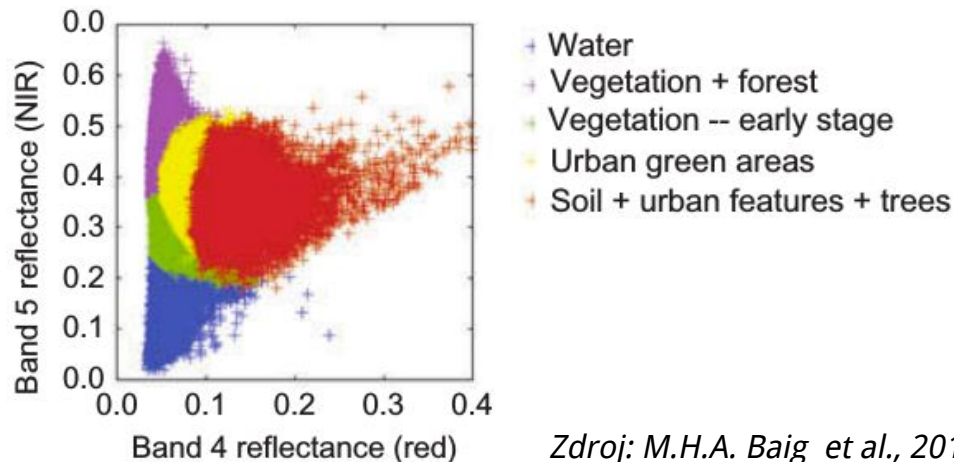
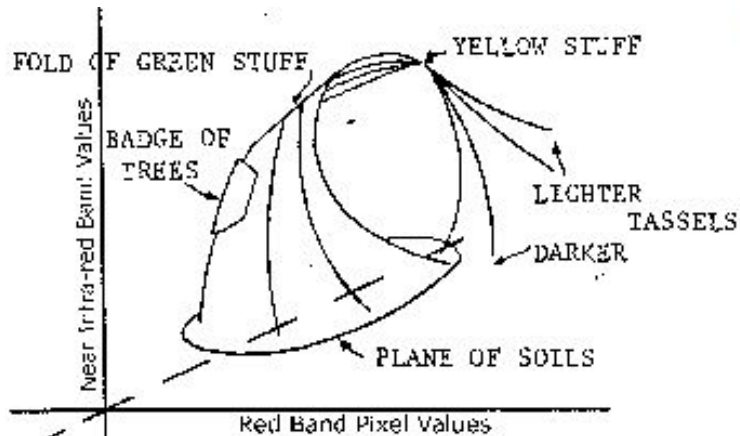
rozlíšenie 375 m (od r.2012)

4. Landsat TM, ETM – **NDVI, EVI**

rozlíšenie 30 m (od r.2016, 8-dňové a 32-dňové kompozity prístupné cez Google Earth Engine)

Tasseled Cap transformácia

- lineárna kombinácia (vážený priemer) pásiem
- potreba odvodenia koeficientov pre konkrétny snímač
- Kauth a Thomas (1976) - 4 indexy z Landsat MSS



Zdroj: M.H.A. Baig et al., 2014

Tasseled Cap transformácia

- výsledkom sú 3 indexy:

B – brightness – index svetlosti (celkový jas snímky)

G – greenness – index zelenej (fotosynteticky aktívnej) vegetácie

W – wetness – index vlhkosti (pôdy alebo povrchu)

(Y – yellowness – index suchosti (množstvo suchej vegetácie))
