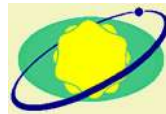


# Validácia reklasifikovaných predikčných máp


Renata Ďuračiová, Milan Muňko, Lukáš Karell

Katedra geodetických základov

Stavebná fakulta STU v Bratislave



# Motivácia

 **predikčný model → predikčná mapa**

 **reklasifikácia predikčných modelov**



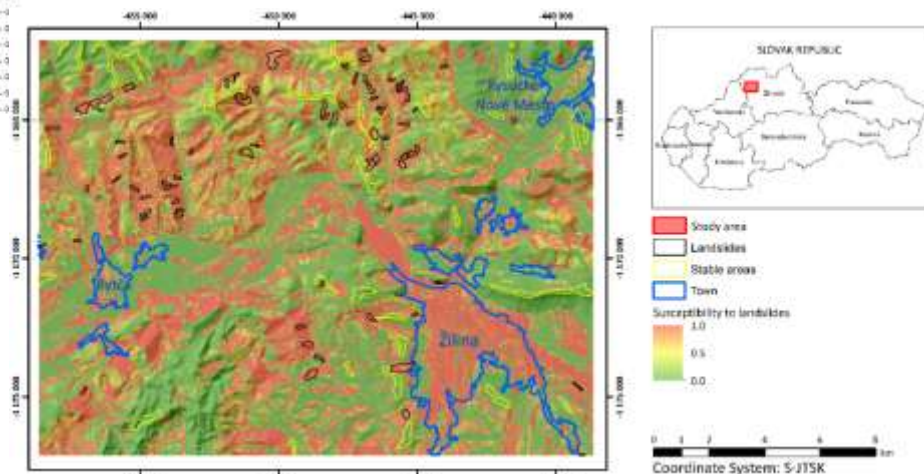
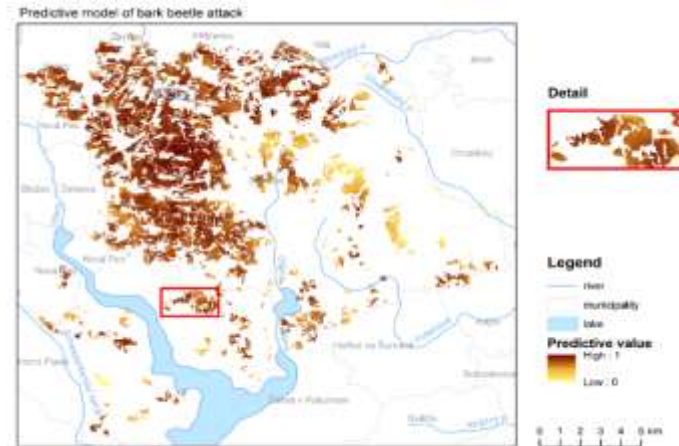
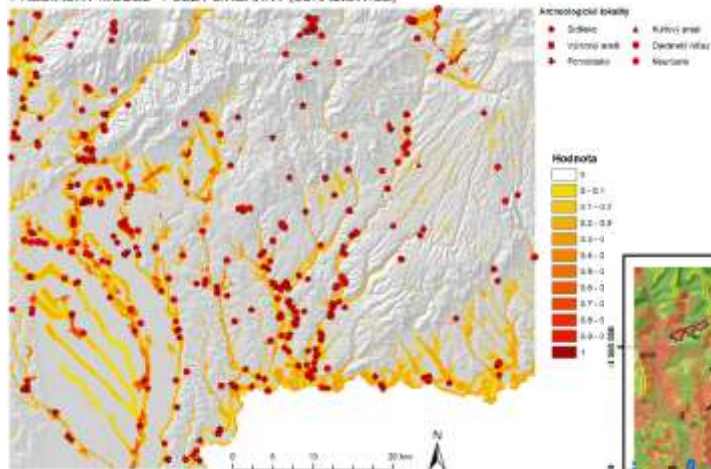
 **validácia predikčných modelov**  
(...a reklasifikovaných predikčných máp)



# Priestorové predikčné modely

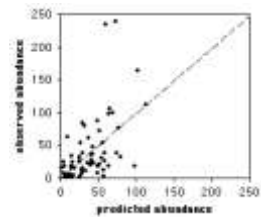
● predikcia výskytu objektov alebo javov v danej oblasti

PREDIKČNÝ MODEL - FUZZY LINEÁRNY (LUKASIEWICZ)

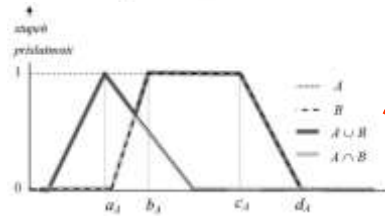


# Metódy priestorového predikčného modelovania

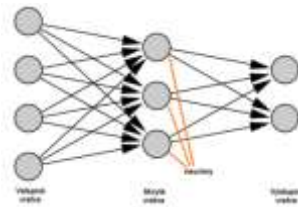
viacnásobná  
lineárna regresia



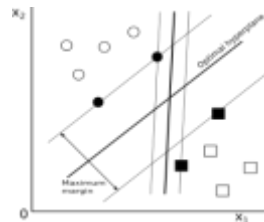
teória fuzzy  
množín



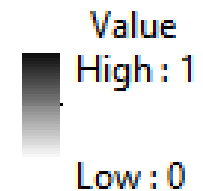
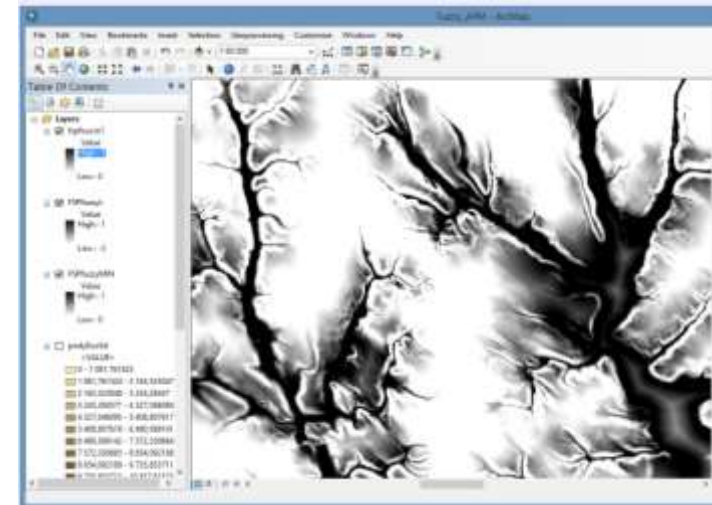
neurónové siete



SVM



...

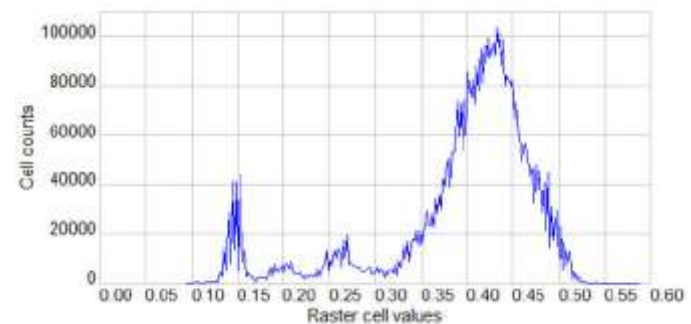


# Reklasifikácia hodnôt predikčných modelov

→ transformácia hodnôt z intervalu  $\langle 0,1 \rangle$  do konečného počtu tried (stupňov splnenia podmienky)

## Metódy

- kvantily (*Quantile*)
- rovnaké intervaly (*Equal Interval*)
- geometrické intervaly (*Geometrical Intervals*)
- metóda prirodzených hraníc (*Natural Breaks*)
- ...
- vlastné stanovenie hraníc intervalov



# Metódy validácie predikčných modelov

- kontingenčná tabuľka / klasifikačná matica

		Skutočnosť	
		P	N
Predikcia	P	True Positive	False Positive
	N	False Negative	True Negative

- presnosť  $\leftrightarrow$  správnosť

$$PR = PPV = \frac{TP}{TP + FP}$$

$$ACC = \frac{TP + TN}{TP + FP + FN + TN}$$

- senzitivita  $\leftrightarrow$  špecificita

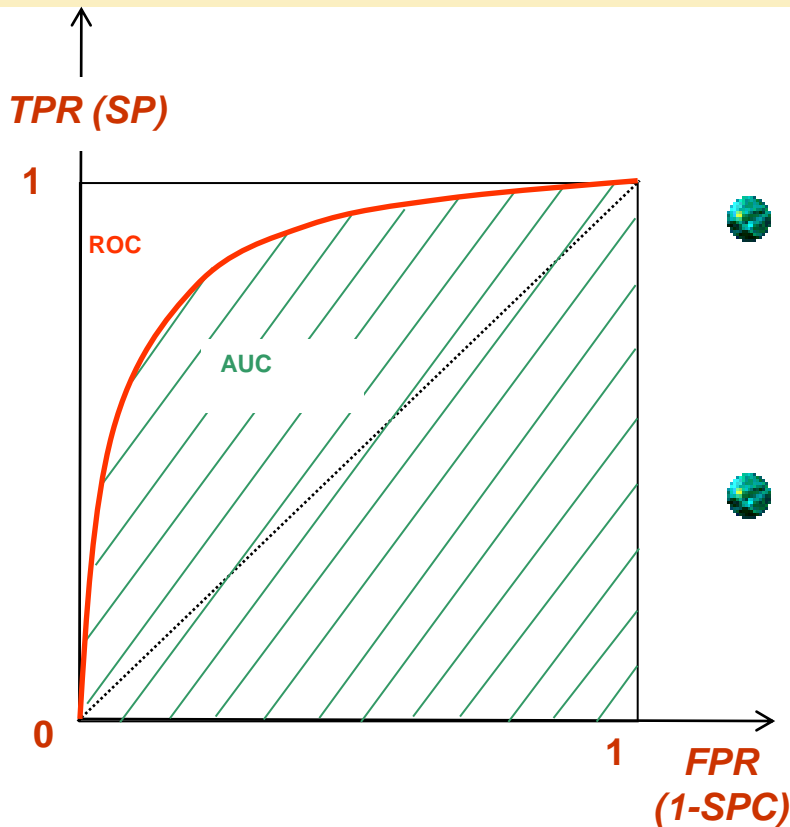
$$SE = TPR = \frac{TP}{TP + FN}$$

$$SPC = TNR = \frac{TN}{FP + TN}$$

$$FPR = \frac{FP}{FP + TN} = 1 - SPC$$

$$FNR = \frac{FN}{FN + TP} = 1 - TPR$$

# ROC krivka a AUC



- Receiver operating characteristic (ROC) curve

- Area under curve (AUC)

→ **metrika na validáciu predikčných modelov**

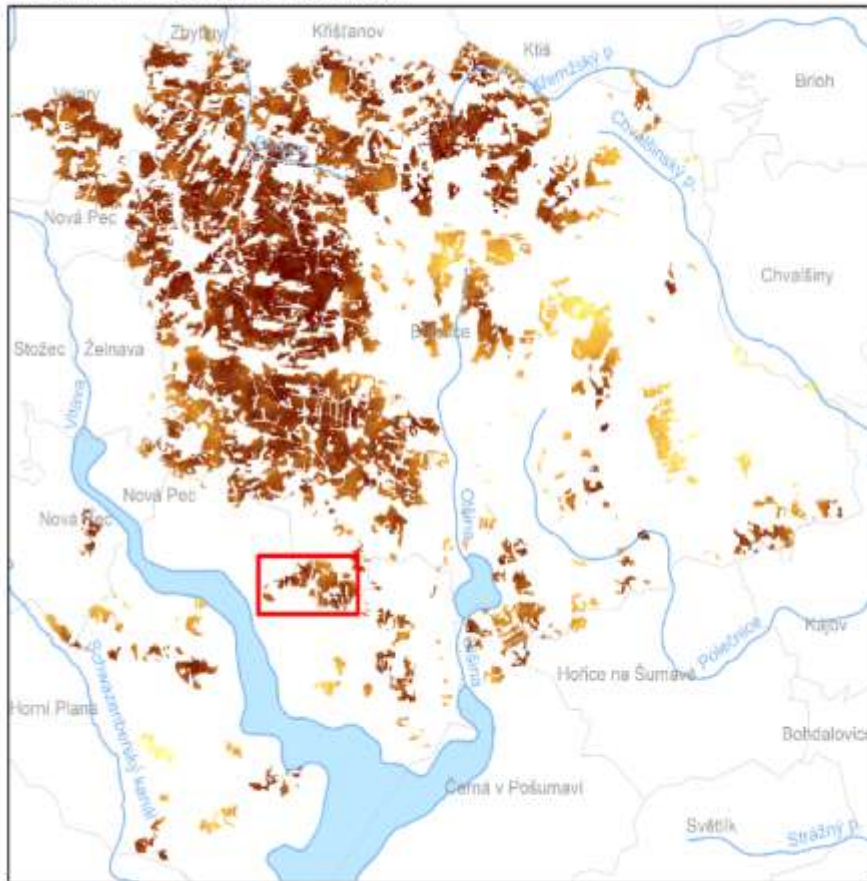
$$SE = TPR = \frac{TP}{TP + FN}$$

$$FPR = \frac{FP}{FP + TN} = 1 - SPC$$



# Predikčný model 1 (nereklasifikovaný)

Predictive model of bark beetle attack



Detail



Legend

— river  
□ municipality  
□ lake

Predictive value

High : 1  
Low : 0

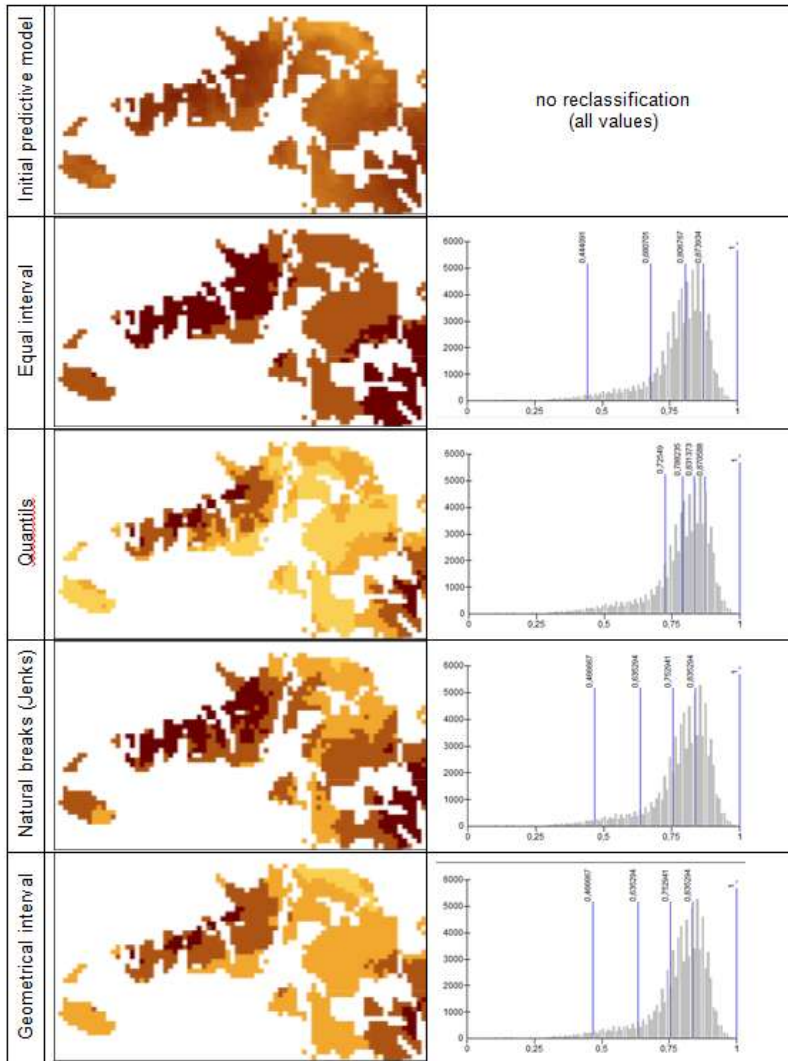
0 1 2 3 4 5 km



- náchylnosť územia na napadnutie lykožrútom smrekovým
- oblasť Horní Planá (ČR)
- metóda viacnásobnej lineárnej regresie
- model TANABBO (Jakuš et al., 2013), (Koreň, 2016)



# Reklasifikácia predikčného modelu



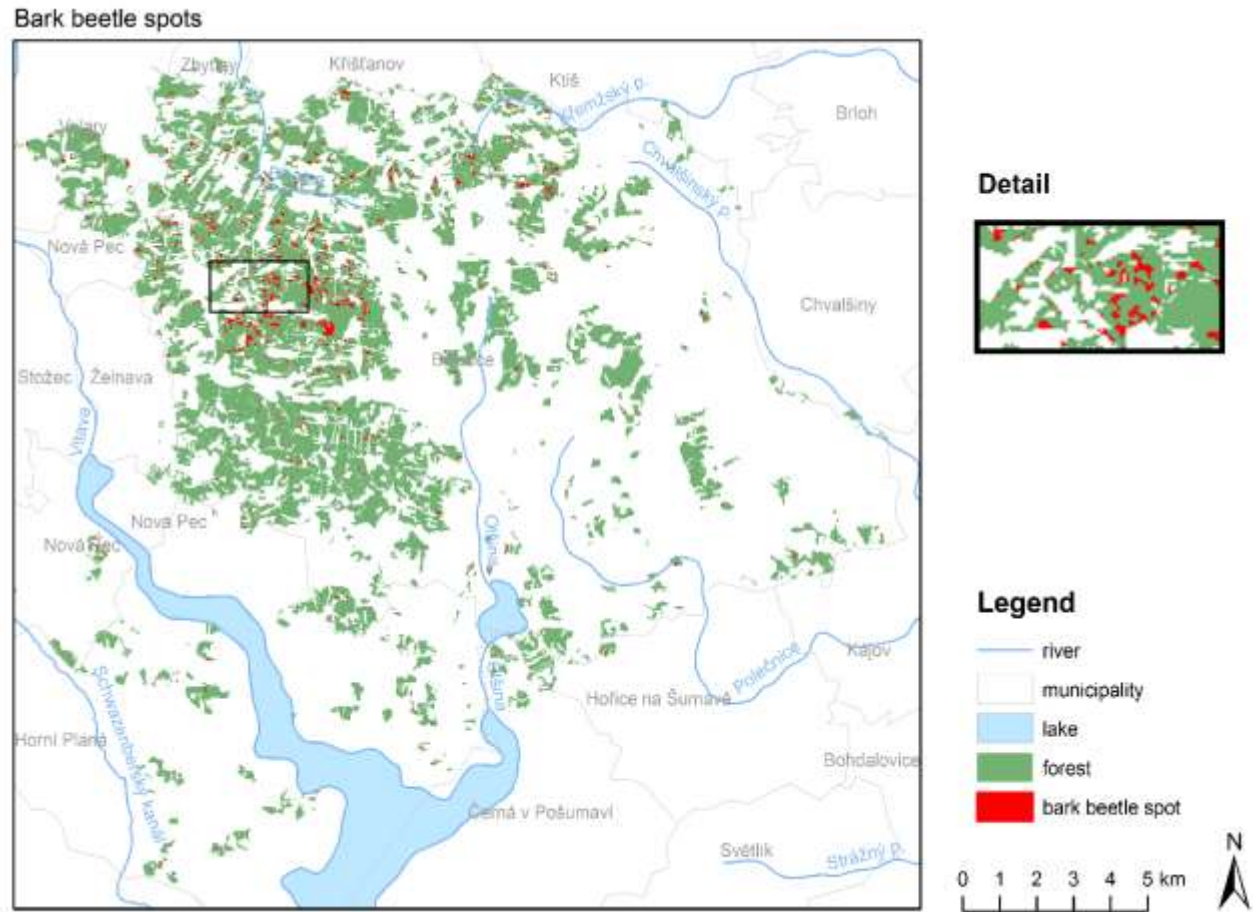
- 5 kategórií
- 4 metódy reklasifikácie

→ vplyv voľby metódy reklasifikácie na výsledok interpretácie modelu

→ **validácia**

# Validácia predikčného modelu

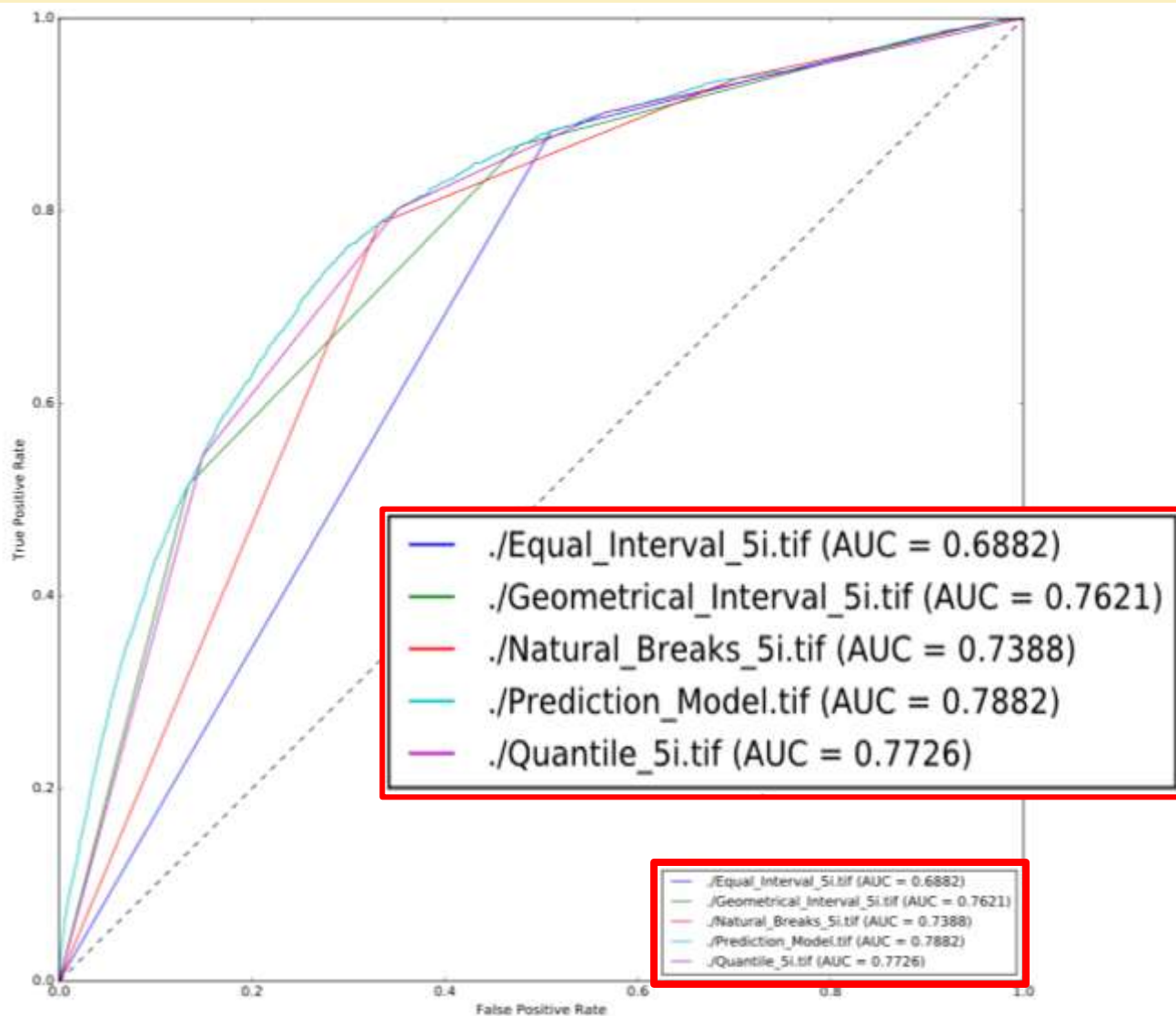
- validačné dáta
- konštrukcia ROC krivky
- parameter AUC



# ROC krivka a parameter AUC

Spracovanie:

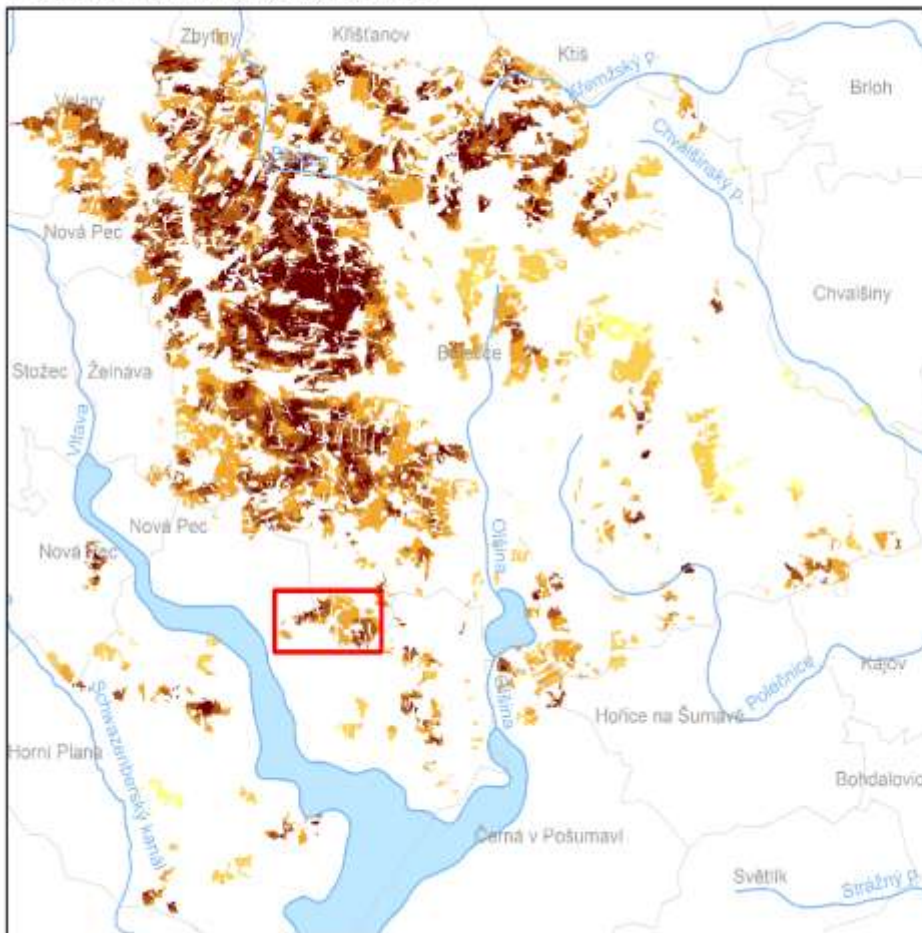
- jazyk Python
- transformácia rastrových dát na vektor hodnôt



# Reklasifikovaná predikčná mapa

→ optimálna metóda reklasifikácie

Predictive model of bark beetle attack



Detail



Legend



0 1 2 3 4 5 km



PM: náchylnosť  
územia na napadnutie  
lykožrútom  
smrekovým  
oblasť: Horní Planá  
(ČR)  
metóda: viacnásobná  
lineárna regresia  
model: TANABBO

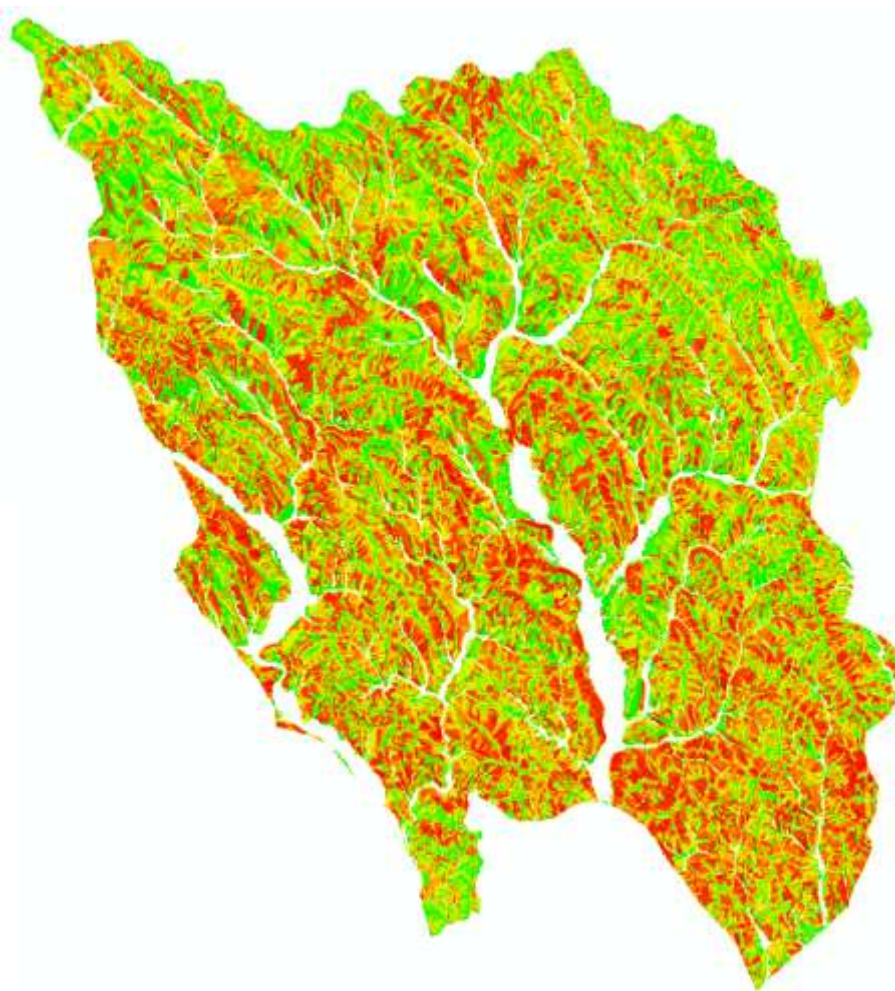
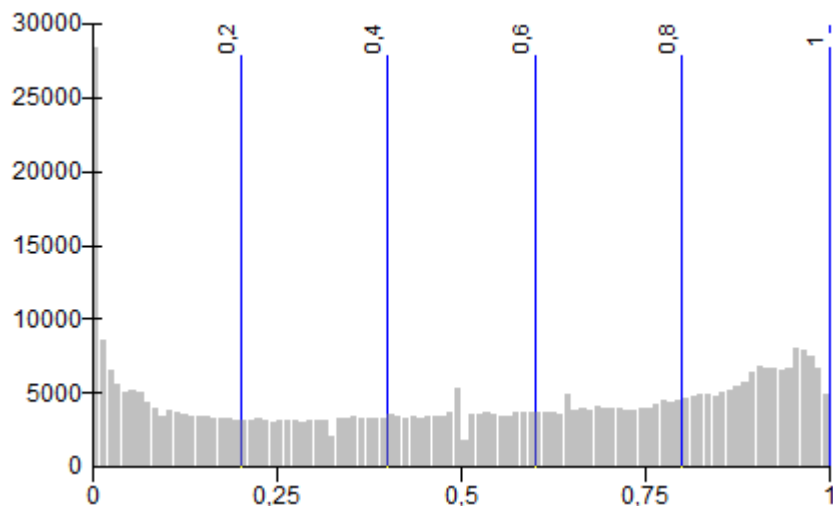
reklasifikácia: kvantily  
počet kategórií: 5  
AUC: 0,773  
AUC pôvodného  
modelu: 0,789



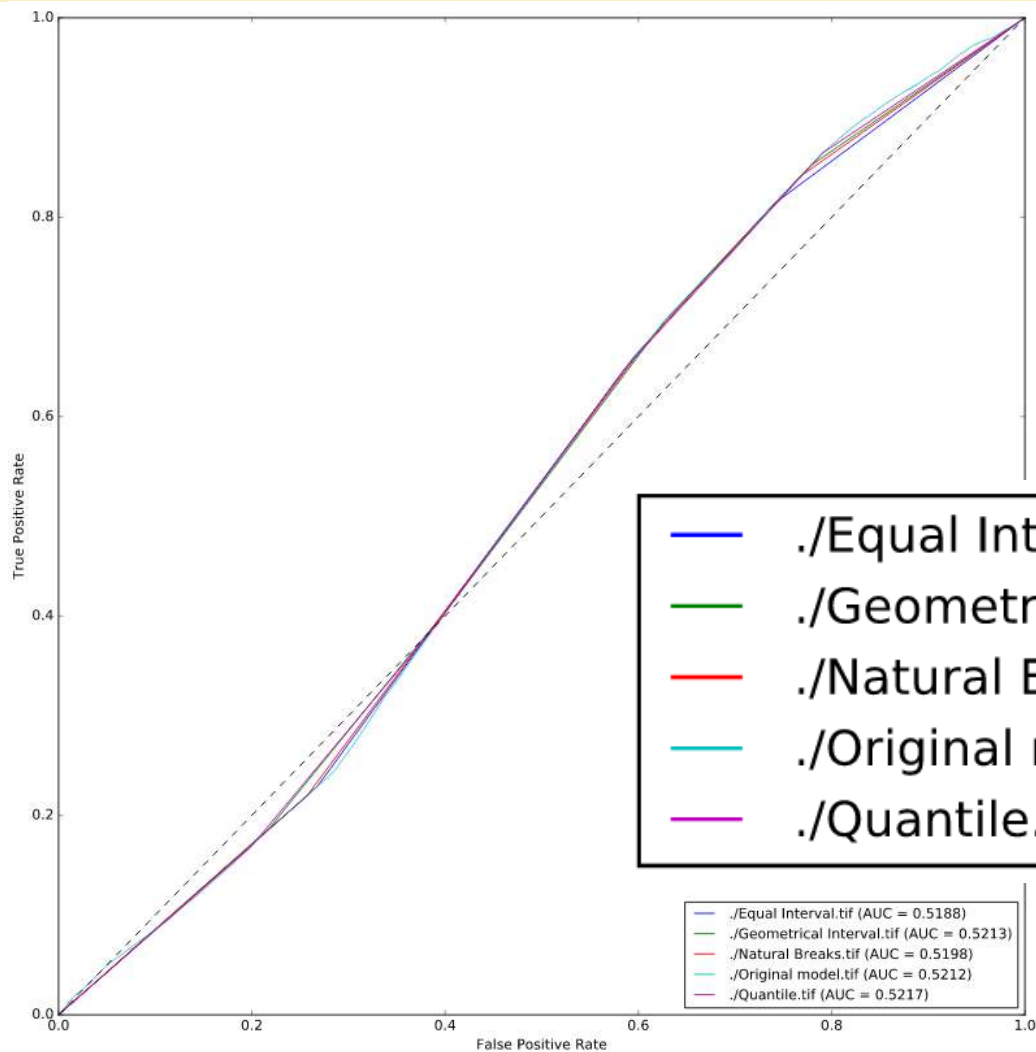
# Predikčný model 2

- predikčný model náchylnosti územia na zosúvanie
- oblasť: Ondavská vrchovina
- metóda: SVM
- 5 kategórií

→ rovnomerné rozdelenie



# Validácia PM – ROC, AUC



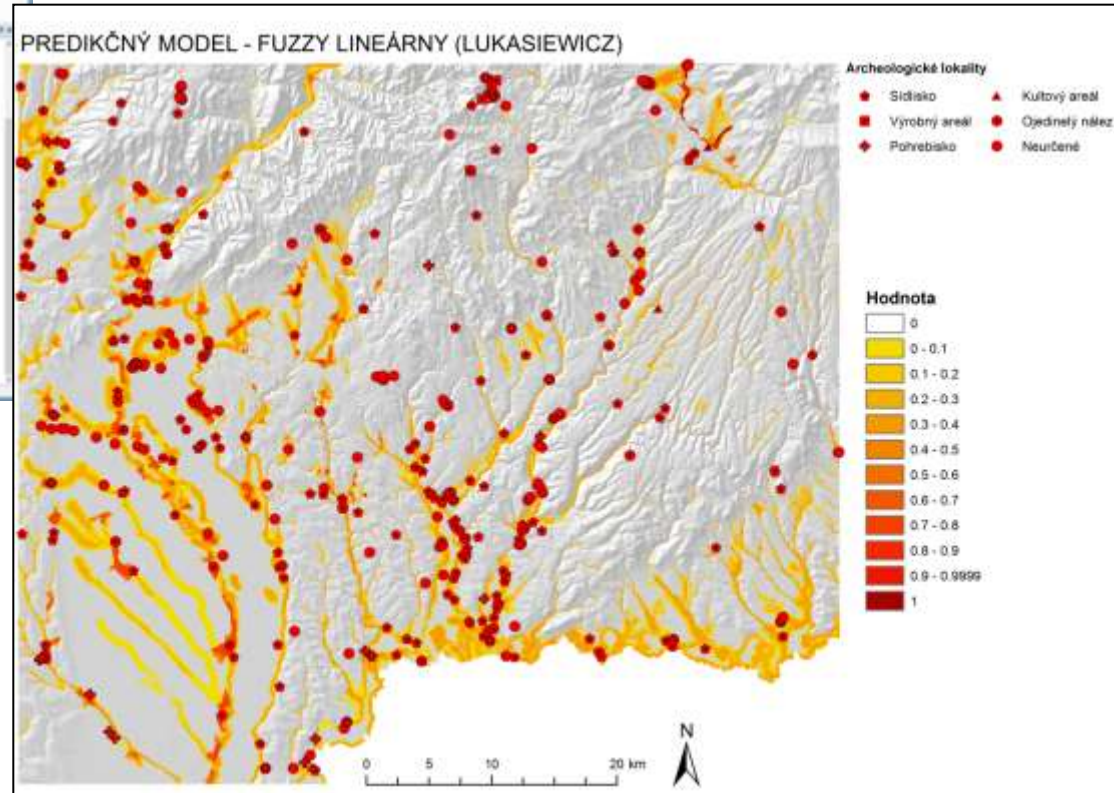
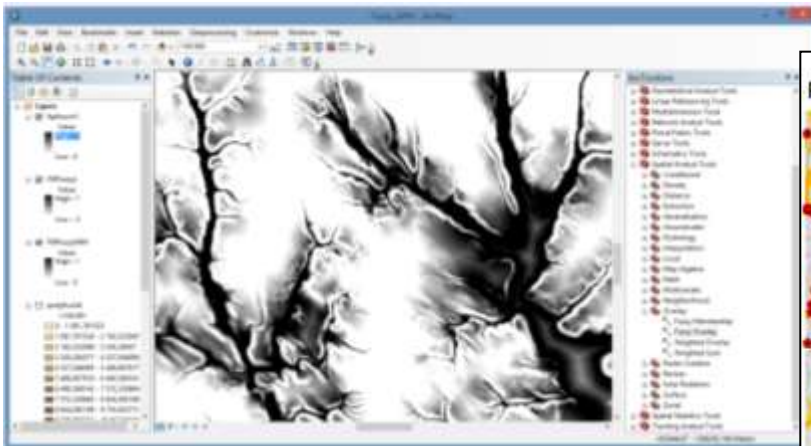
- 5 kategórií
- rovnomerné rozdelenie
- AUC ~ 0,5

./Equal Interval.tif (AUC = 0.5188)  
./Geometrical Interval.tif (AUC = 0.5213)  
./Natural Breaks.tif (AUC = 0.5198)  
./Original model.tif (AUC = 0.5212)  
./Quantile.tif (AUC = 0.5217)

./Equal Interval.tif (AUC = 0.5188)  
./Geometrical Interval.tif (AUC = 0.5213)  
./Natural Breaks.tif (AUC = 0.5198)  
./Original model.tif (AUC = 0.5212)  
./Quantile.tif (AUC = 0.5217)



# Predikčný model 3

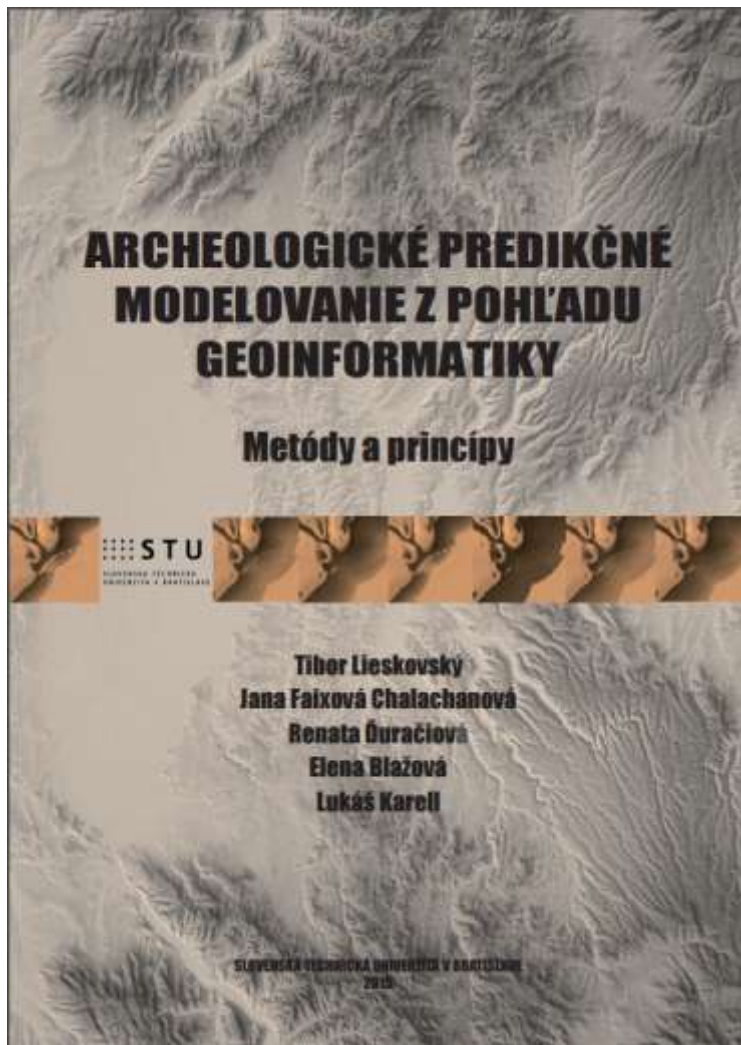


- archeologický predikčný model
- oblasť dolného Poipia a Pohronia

(Lieskovský et al., 2015), (Ďuračiová et al., 2011)



# APM z pohľadu geoinformatiky

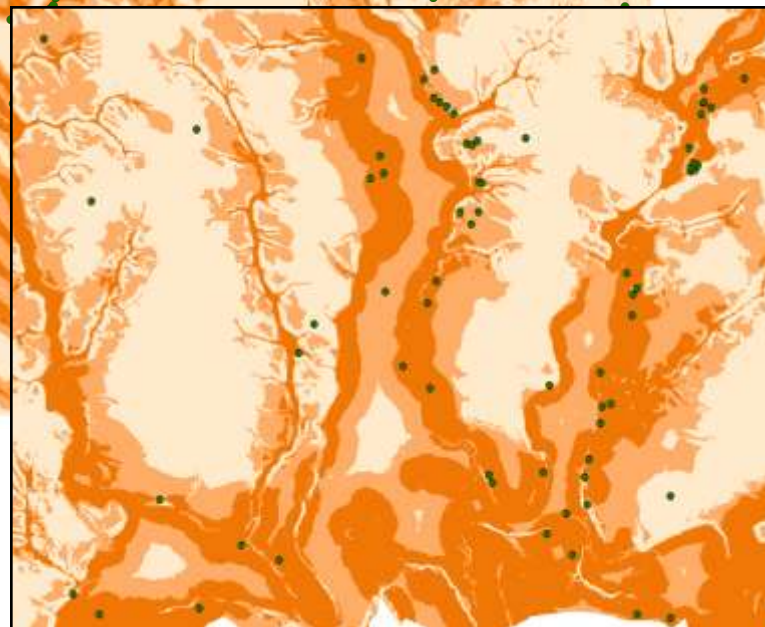
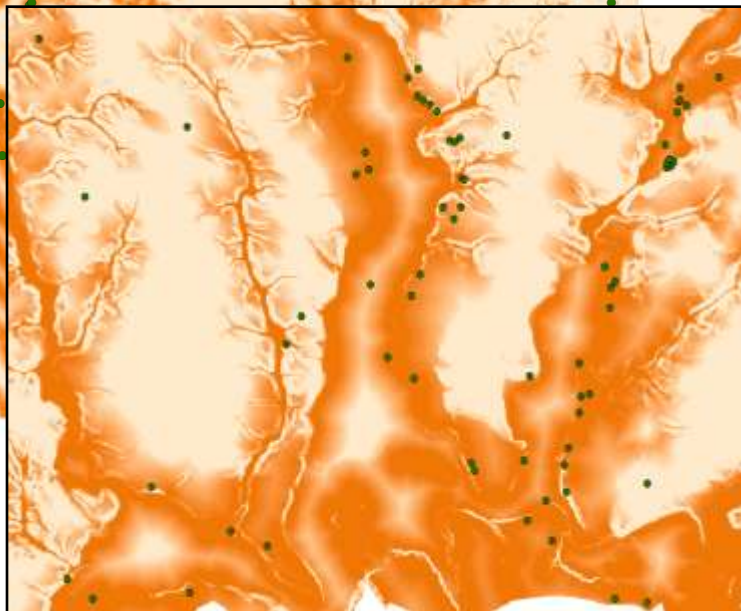
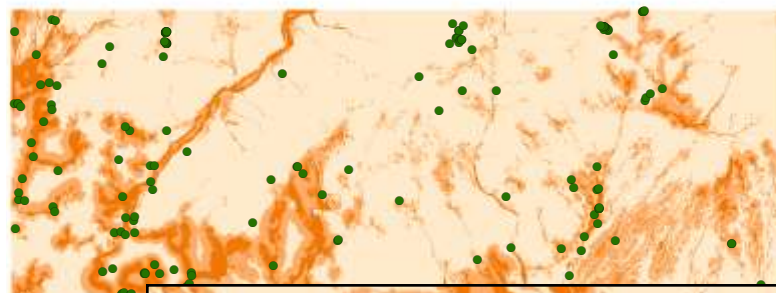
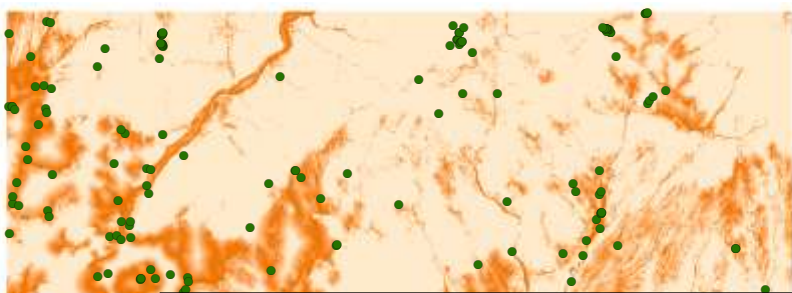


- príprava dát na predikčné modelovanie
- metódy a princípy tvorby APM
- validácia APM
- SW spracovanie
- ...

[ftp://147.175.19.174/Predikcne\\_modelovanie/Archeologicke\\_predikcne\\_modelovanie\\_z\\_pohladu\\_geoinformatiky.pdf](ftp://147.175.19.174/Predikcne_modelovanie/Archeologicke_predikcne_modelovanie_z_pohladu_geoinformatiky.pdf)

# Reklasifikácia APM

- 3 kategórie, 5 metód reklasifikácie, návrh optimálnej metódy reklasifikácie



# Modifikácia štandardných metód validácie PM

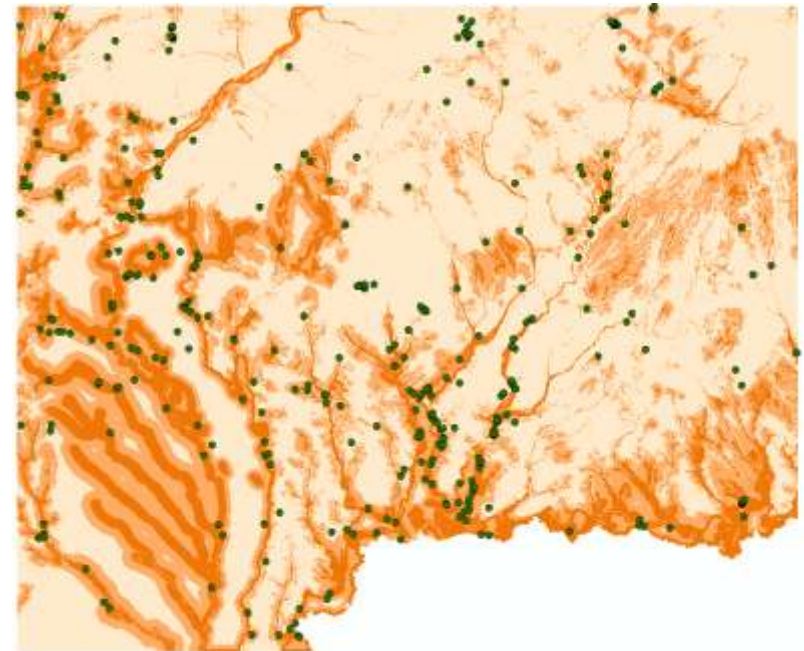
		Skutočnosť	
		P	N
Predikcia	P	True Positive	False Positive
	N	False Negative	True Negative

- fuzzy modifikácia ROC / AUC
- nové metódy validácie PM

$$e_{mf} = \frac{\sum_{i=1}^n \mu_{k_i} l_{k_i}}{\sum_{i=1}^n \mu_{k_i} k_{v_i}} - \frac{l_m}{k_m},$$

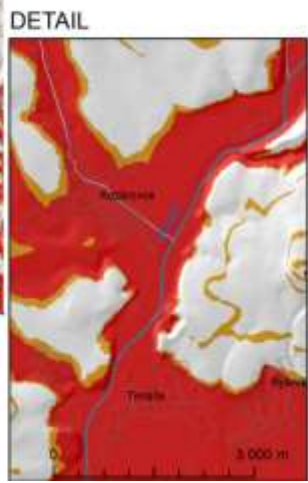
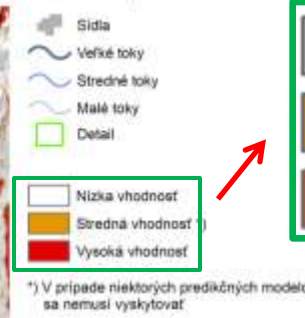
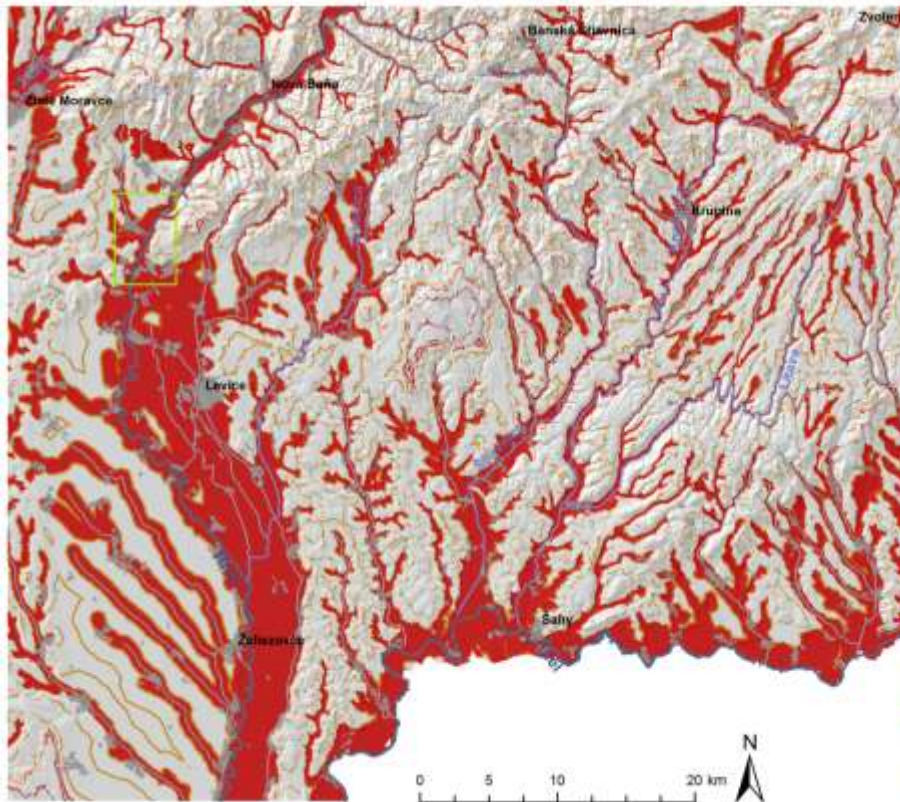
- modifikovaný Giniho index

$$G_{mf} = 1 - \frac{\sum_{i=1}^n \mu_{k_i} k_{v_i}}{\sum_{i=1}^n \mu_{k_i} l_{k_i}}.$$



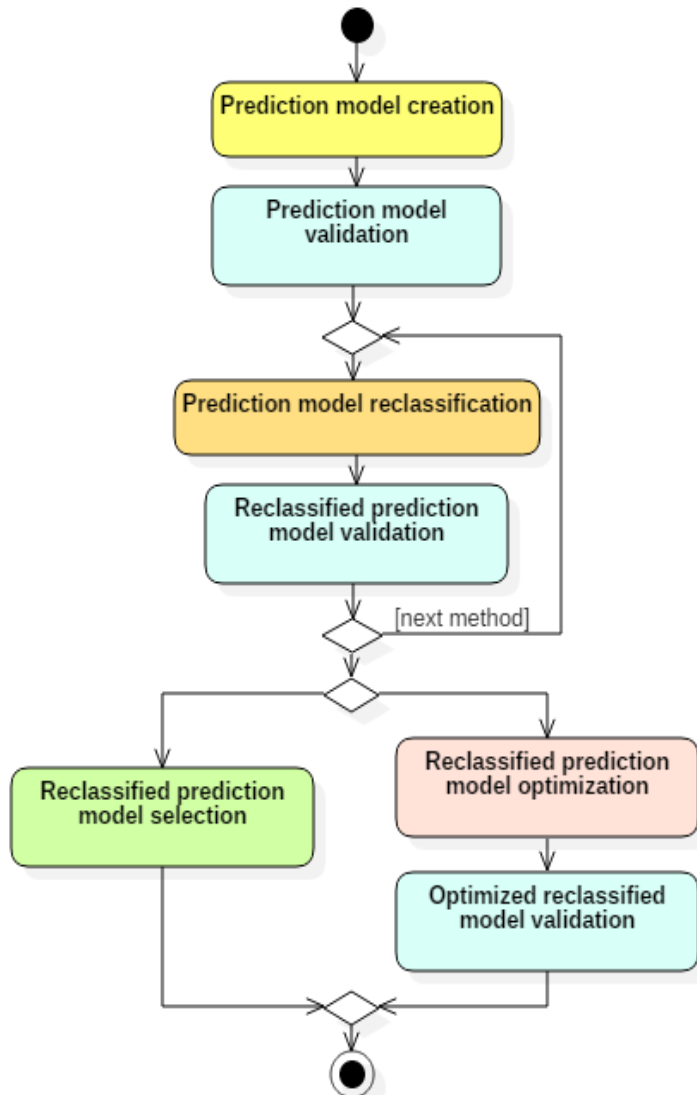


# Reklasifikovaná predikčná mapa



- metóda:  
multikriteriálne  
rozhodovanie  
prostredníctvom  
fuzzy množín
- Łukasiewiczova  
t-norma
- 3 kategórie
- manuálna  
reklasifikácia
- $G_{mf}$ : 0,64

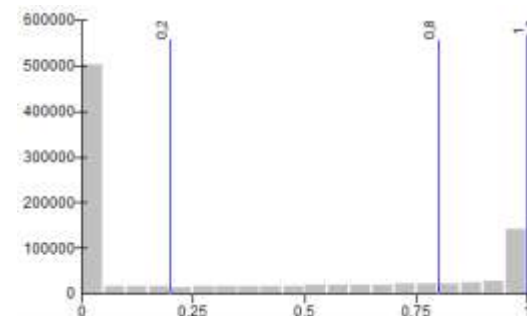
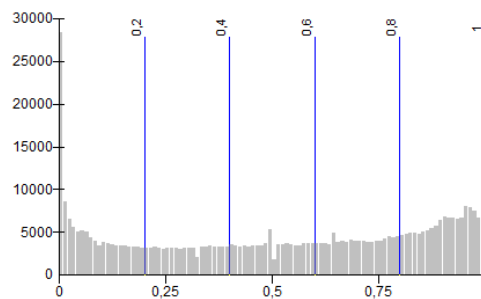
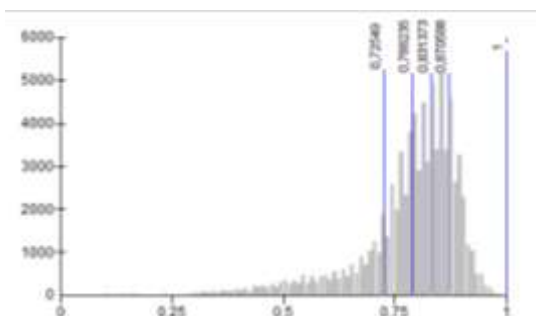
# Proces reklasifikácie a validácie PM



- tvorba PM
  - validácia PM
  - reklasifikácia PM
  - validácia reklasifikovaných PM
  - optimalizácia reklasifikácie
- ⇓
- tvorba výslednej predikčnej mapy

# Záver

- 🌐 interpretácia výsledku používateľovi
- 🌐 voľba počtu kategórií
- 🌐 výber vhodnej metódy reklasifikácie ( $\approx$  strata informácií)



**validácia (aj) reklasifikovaných predikčných máp**



**informácie o tvorbe a validácii modelu**

# ĎAKUJEME ZA POZORNOST!

renata.duraciova@stuba.sk  
milan.munko@stuba.sk  
lukas.karell@stuba.sk