
Prediktivní modelování struktury biologických společenstev na základě referenčních dat – metodický příspěvek k hodnocení ekologických rizik

Klára Kubošová

Prediktivní modelování

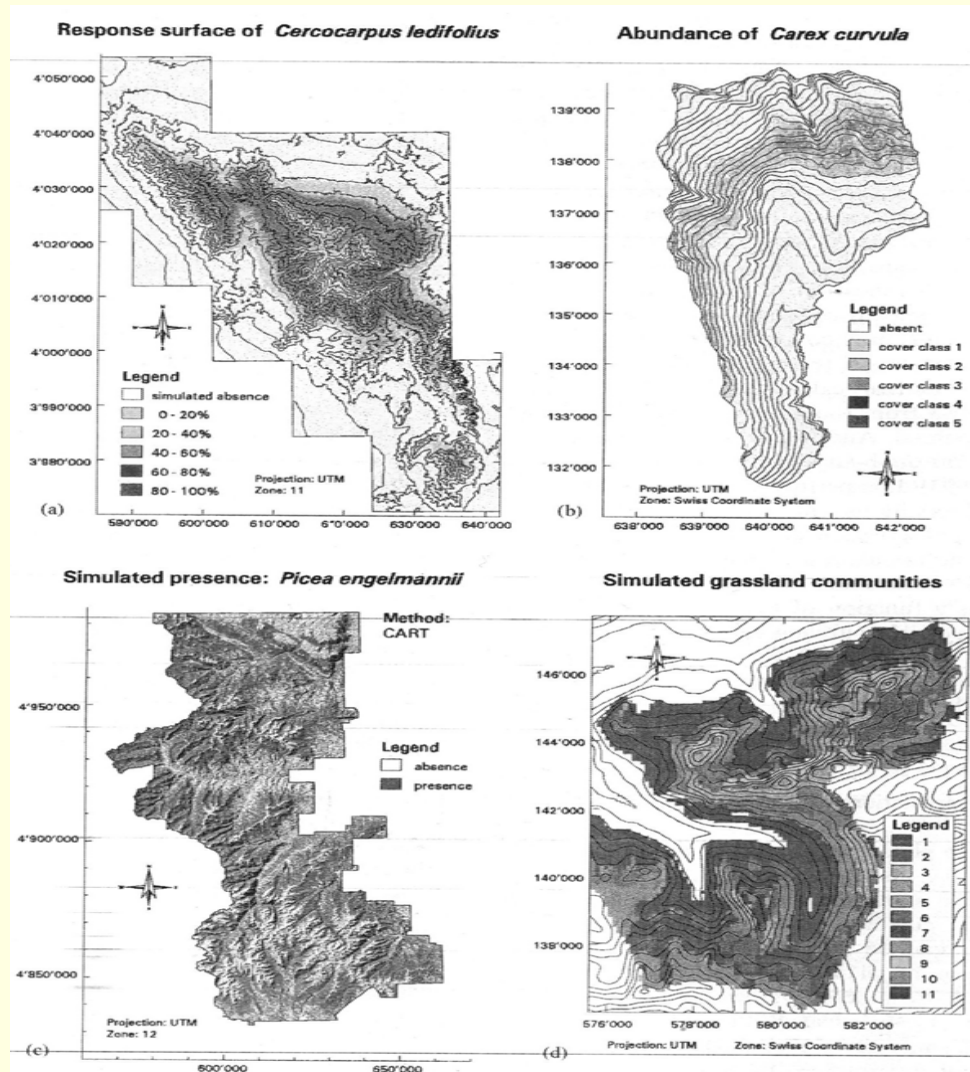
- Výhody a nevýhody metod využívaných k tvorbě predikčního modelování struktury a stavu biologických společenstev v různých prostorových měřítcích
- zhodnocení stavu referenčních databází u případových studií (rostlinná společenstva a společenstva makrozoobentosu): tvorba biologicky relevantních predikčních modelů, definice nejvhodnějších sad abiotických proměnných využitelných v modelech
- cross-validační studie srovnávající predikční schopnosti různých modelovacích metod
- vytvoření a validace nejvhodnějšího v praxi využitelného predikčního modelu
- metodické výstupy směřující do oblasti hodnocení ekologických rizik, především fáze hodnocení expozice a biologických účinků
- vývoj nových přístupů k této problematice s ohledem na charakter dostupných dat

PHDMs -potential habitat distribution maps

- Modely, které předpovídají potenciální rozšíření společenstev se označují PHDMs

- Můžou být definovány několika způsoby:
 - 1) pravděpodobnost výskytu druhů
 - 2) podle rozložení abundance druhů
 - 3) předpokládaný výskyt založený na nepravděpodobnostních metrikách (presence/absence)
 - 4) nejvíce pravděpodobná entita

PHDMs -potencial habitat distribution



Modelovací přístupy

- Ordinační techniky (PCA, RDA, CA, CCA, ENFA)
- Regrese (GLM, GAM, regresní stromy, neuronové sítě)
- Klasifikační techniky
 - diskriminační analýza,
 - klasifikační stromy
 - náhodné lesy
 - neuronové sítě

Praktické aplikace

- Prediktivní modelování potenciální distribuce rostlinných asociací
- Výběr indikačních taxonů pro říční habitaty

Predictive modelling of potential distribution of plant associations in the Czech Republic

- data were taken from the Czech National Phytosociological Database at the Department of Botany
- 36 associations – with 4700 phytosociological plots
- abiotic environmental variables - elevation, average temperature (June, January, annual), annual precipitation and soil acidification
- geographical information system ArcGIS
 - 37 685 squares of 2.21 km²

The data set

- training x testing
- outliers of predictive variables were excluded
- presence/absence - comparing the phytosociological plots assigned to the given association with plots assigned to other associations

Models

- **variable with binary response**

 - ordination method - Canonical Correspondence Analysis (CCA)

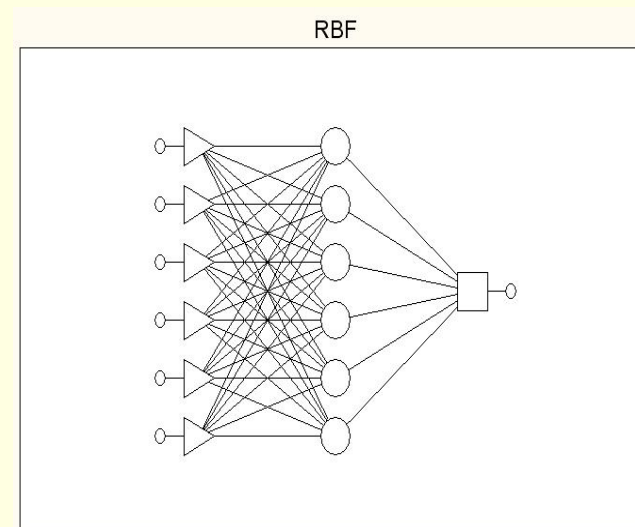
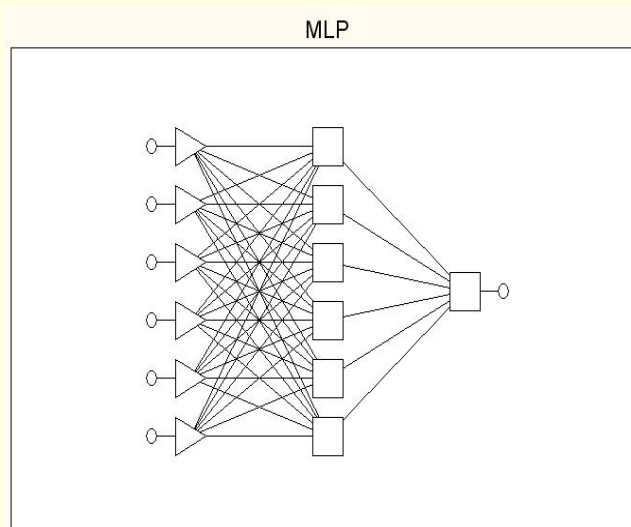
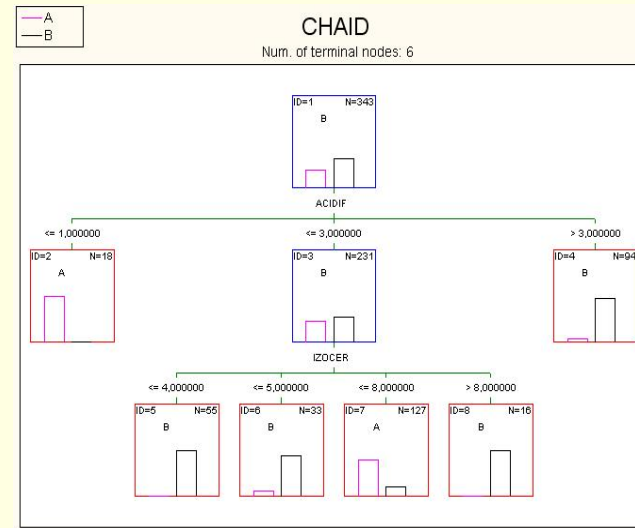
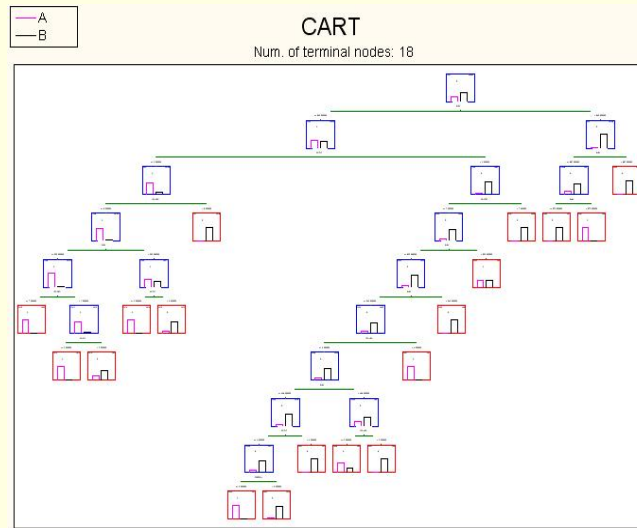
 - regression techniques - Generalized Linear Models (GLM)
Generalized Additive Models (GAM)

- **categorical variable**

 - neural network - Multilayer Perceptron (MLP),
Radial Basis Function (RBF)

 - classification trees - CART, CHAID

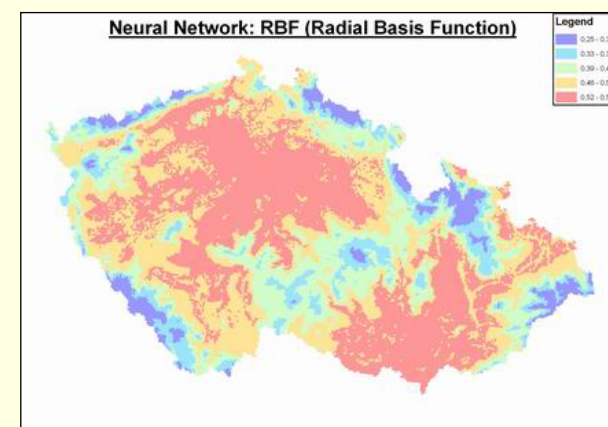
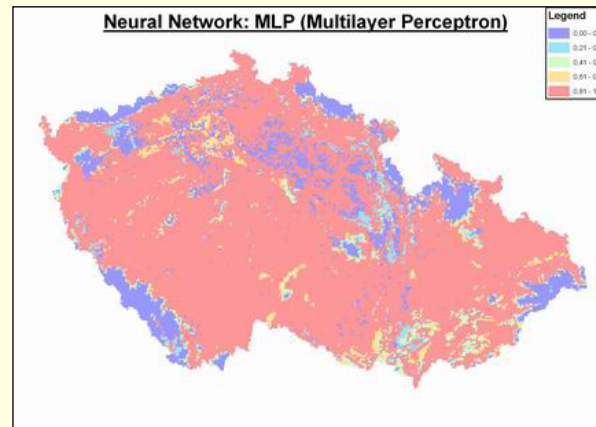
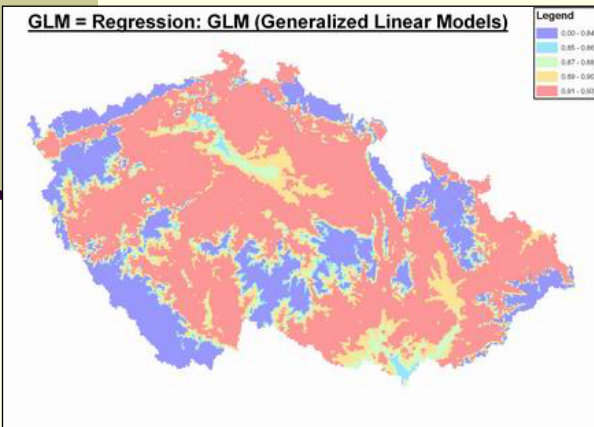
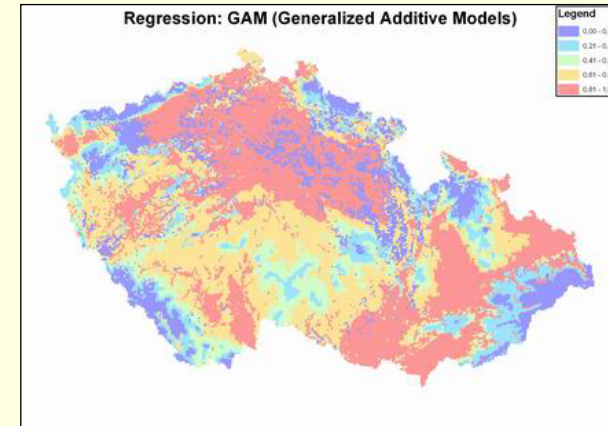
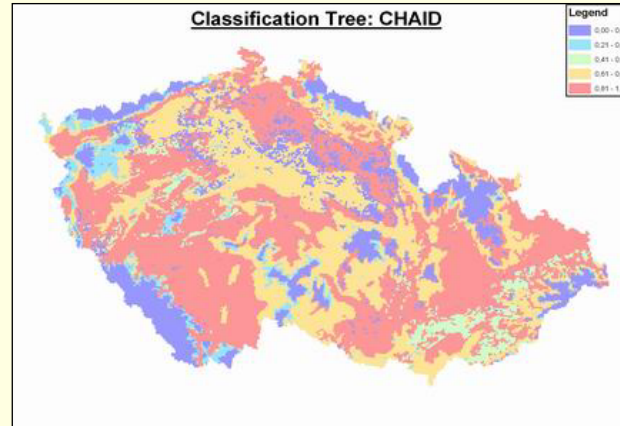
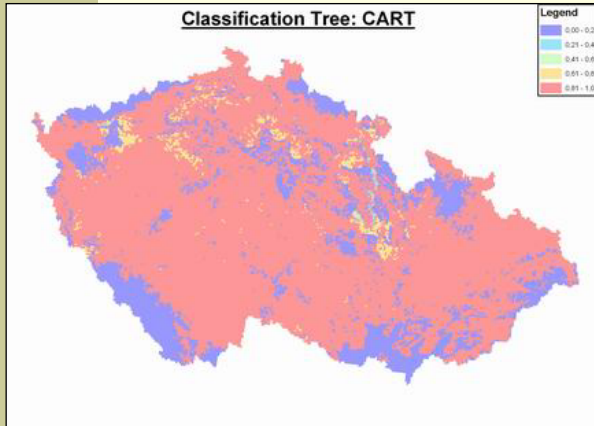
Classification methods for association of *Aphano arvensis*-*Matricarietum chamomillae*



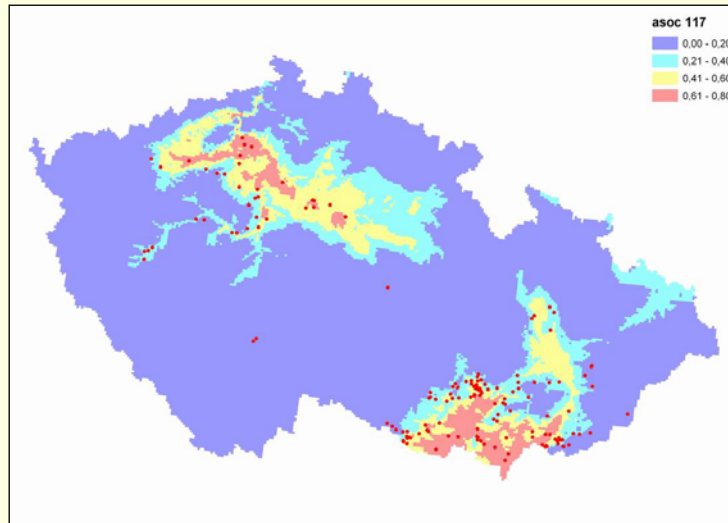
Comparison of predictions for ten associations.

Associations	GLM accuracy	CART accuracy	CHAID accuracy	MLP accuracy	RBF accuracy
<i>Setario viridis-Fumarietum</i>	87,1	76,50	74,19	65,03	64,38
<i>Euphorbio exiguae-Melandrietum noctiflori</i>	85,4	77,50	78,25	64,39	64,15
<i>Vaccinio-Callunetum vulgaris</i>	84,1	79,59	76,53	64,81	66,05
<i>Angelico sylvestris-Cirsietum palustris</i>	82,3	66,67	50,00	61,56	62,10
<i>Chaerophyllo hirsuti-Filipenduletum ulmariae</i>	82,2	80,98	50,00	64,41	62,71
<i>Echinochloo crus-gali-Setarietum pumilae</i>	79,6	54,28	62,85	59,59	60,96
<i>Filipendulo ulmariae-Geranietum palustris</i>	64,8	62,30	67,20	36,30	53,08
<i>Aphano arvensis-Matricarietum chamomillae</i>	62,4	76,60	68,80	50,00	33,73
<i>Ranunculo bulbosi-Arrhenatheretum elatioris</i>	57,6	75,40	59,80	45,80	50,38
<i>Poo-Trisetetum flavescens</i>	56,9	65,00	66,10	41,10	47,85

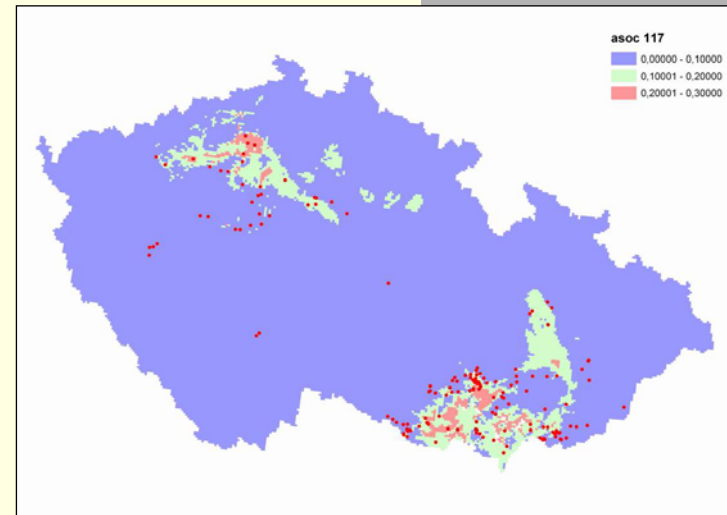
Potential distribution maps of *Aphano arvensis* - *Matricarietum chamomillae* in the Czech Republic



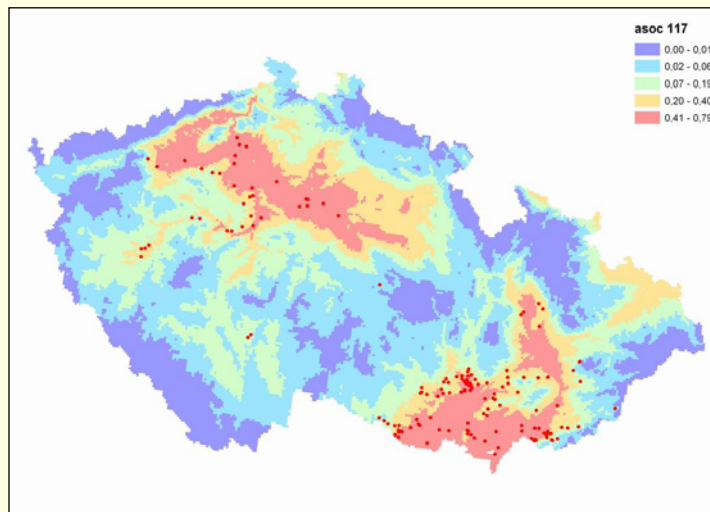
Classification tree CART



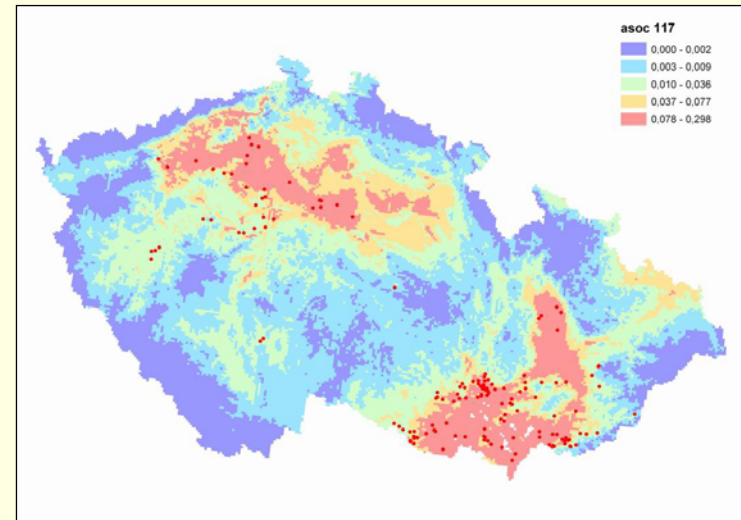
Classification tree CHAID



Regression technique GLM



Regression technique GAM



Výběr indikačních taxonů

- srovnání statistických metod pro výběr indikačních taxonů charakteristických pro abioticky definované říční habitaty
- Random Forests (Breiman, 2001)
- Metoda indikačních druhů (Indicator Species Analysis; Dufrene & Legendre, 1997)

Datový soubor

- do analýzy vstupovala kategoriální závisle proměnná, obsahující 54 vzorků rozdělených do 4 typů habitatů (-definovaných na základě polohy odběrových míst v rámci koryta toku a lokálních hydraulických podmínek)
- prediktory byly abundance 180 taxonů makrozoobentosu, zjištěných v těchto vzorcích

Habitat H1 – peřeje

Nejproudivější místa v korytě hydraulicky charakterizovaná hodnotou tzv. Froudova čísla vyšší než 0,18. Kvantita usazené organické hmoty a řasových nárostů velikosti a délce trvání zvýšených průtoků.



Habitat H2 - příbřežní tišiny

Mírné proudění až stagnace se projevuje usazováním jemných sedimentů. Potravní i úkrytové podmínky jsou ovlivňovány přítomností pobřežní vegetace (opad, dřevní hmota, kořeny).



Habitat H3 – tůně

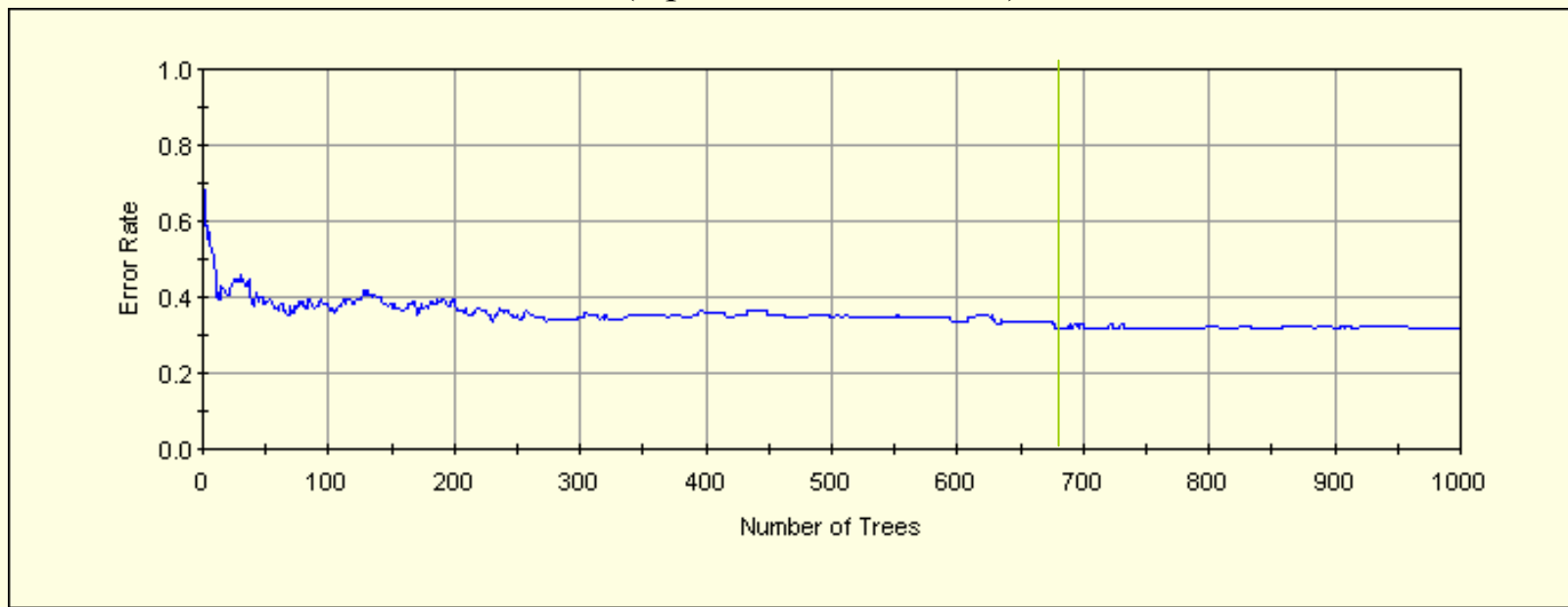
Habitaty hlavního koryta charakteristické laminárním prouděním a Froudovým číslem nižším než 0,18.

Habitat H4 - boční ramena

Morfologie habitatu je utvářena erozním účinkem vysokých průtoků. Po většinu roku jsou slepá ramena povrchově propojena s hlavním korytem pouze na spodním konci. Převažujícím substrátem jsou jemné sedimenty a organická hmota.

Results from RF

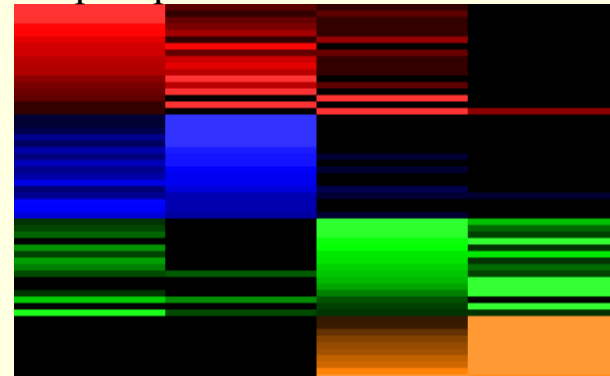
Selection of correct number of trees (4 predictors, 680 trees)



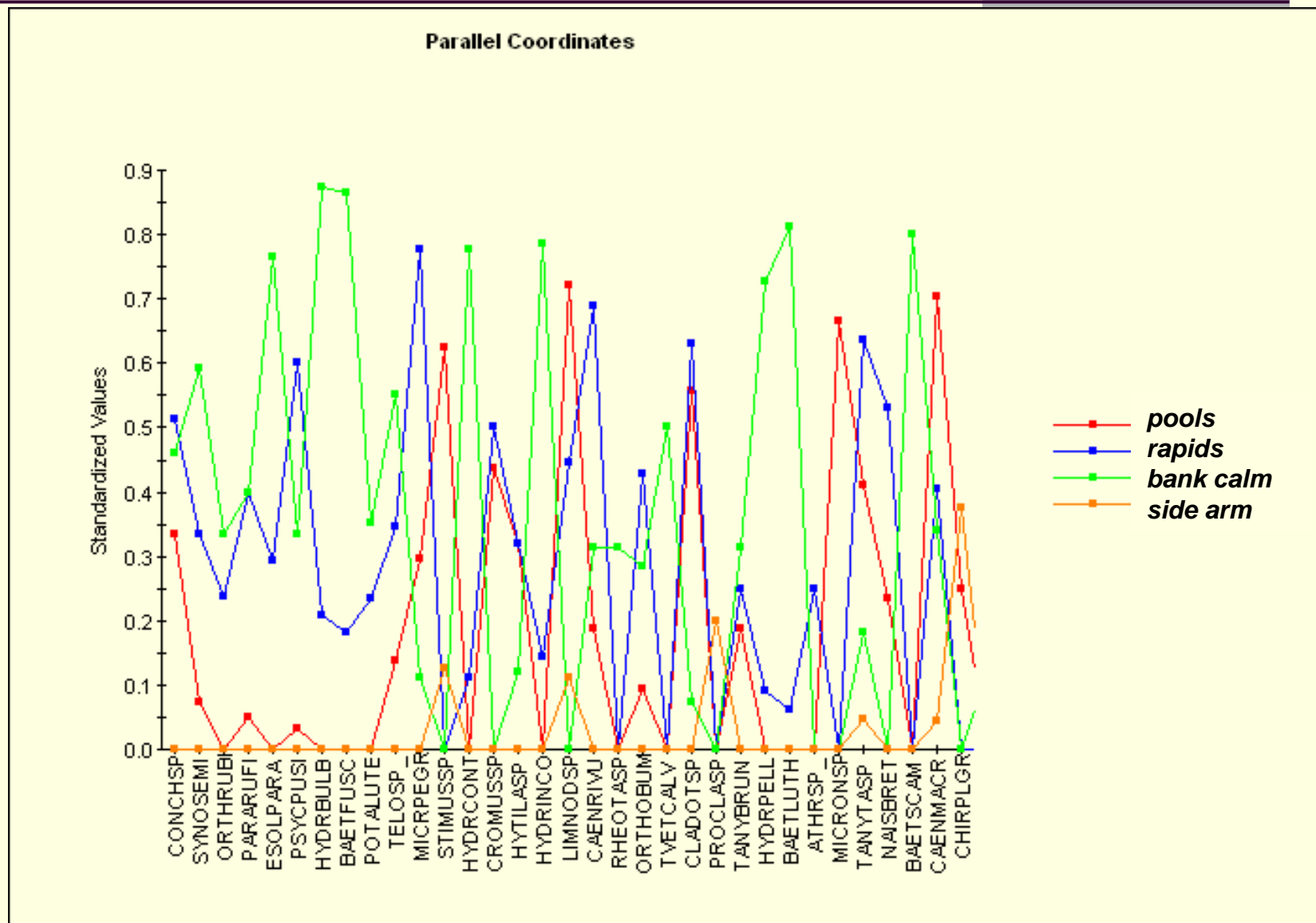
Random Forests Prediction Success

Actual Class	Total Cases	Percent Correct	pools	rapids	bank calm	side arm
pools	27	70.4	19	3	4	1
rapids	40	65	12	26	1	1
bank calm	25	40	5	3	10	7
side arm	15	93.3	0	0	1	14
	107	67.2	36	32	16	23

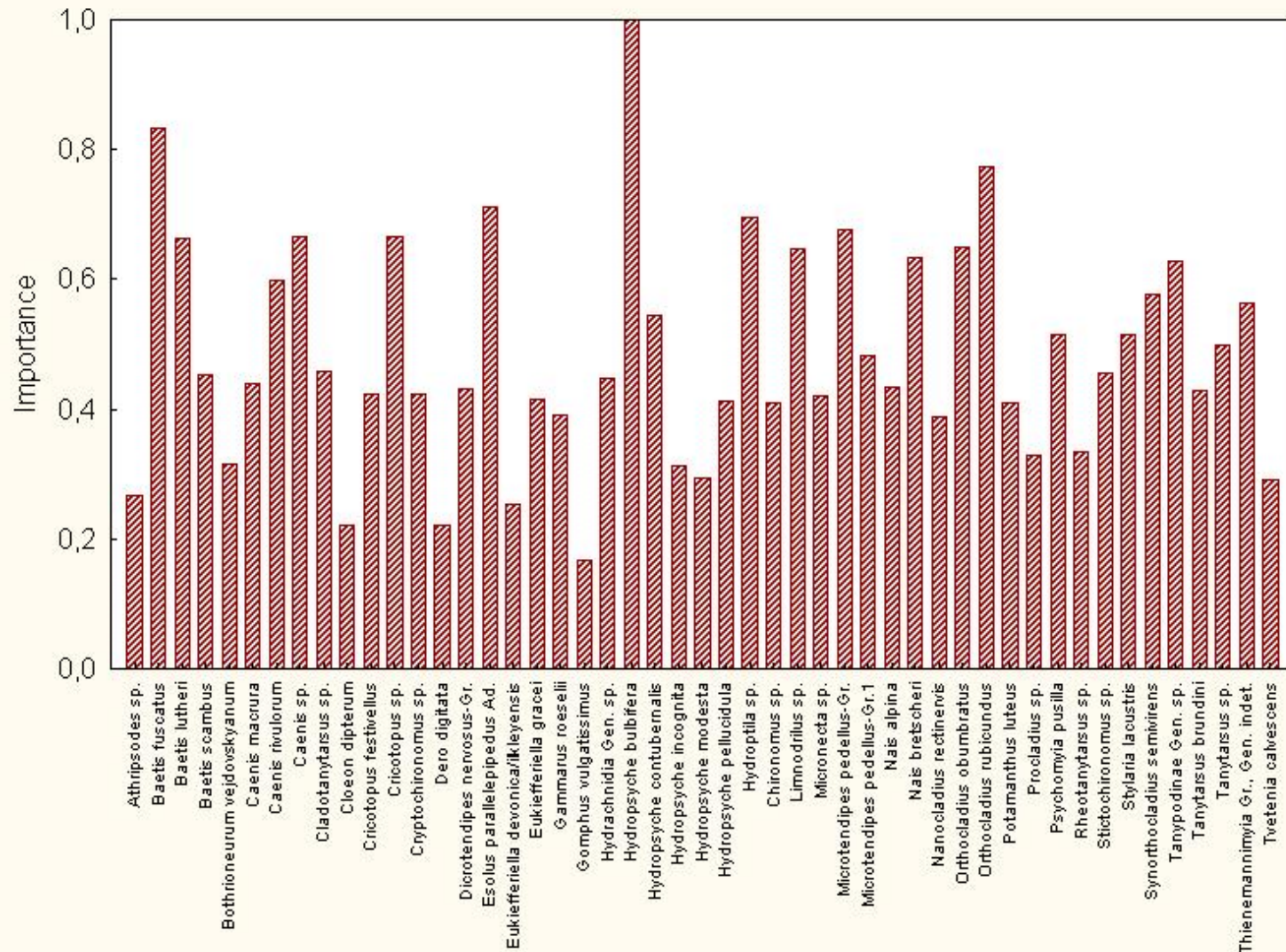
Map of proximities



Results from RF



Importance value (i.e. indicative power of taxa) for selected taxa



Výsledky

- bylo vybráno 50 taxonů z celkového počtu 180 taxonů, které byly alespoň v jedné z těchto metod určeny jako významné
- tyto dvě metody se shodovaly v 85% vybraných taxonů
- tůňové habitaty v hlavním korytě (H3) se ukázaly z hlediska indikátorových taxonů nejméně výrazným typem habitatu
- z uvedených hodnot je patrné, že použité metody, dávají velmi podobné výsledky

další cíle

- otestování dostupných metod na společenstvech makrozoobentosu – projekt ARROW (výběr indikačních taxonů, váhy pro parametry...)
- predikce lesních společenstev (na presenčních datech, srovnání upravené PCA a CA s GLM na absenčních datech z celé ČR)
- klimatické modelování (BIOCLIM, CRES-Centre for Resource and Environmental Studies)
- modelování perzistentních organických polutantů v půdách pro území ČR