Selection of principal stand factors as predictors for digital mapping of potentially toxic element contents in forest soils

Luboš Borůvka[1], Radim Vašát[1], Václav Tejnecký[1], Vít Šrámek[2], Milan Sáňka[3], Jarmila Čechmánková[4], Karel Němeček[1], Vít Penížek[1]

[1]Czech University of Life Sciences Prague, Czech Republic
[2]Forestry and Game Management Research Institute Strnady, CZ
[3]RECETOX, Masaryk University Brno, Czech Republic
[4]Research Institute for Soil and Water Conservation Prague, CZ

GW DSM 2016, Aarhus, Denmark

Background - Motivation

- **Process based models**
  - deterministic, physical

vs.

- **Data driven models**
  - stochastic
  - large numbers of potential predictors
  - in DSM used more frequently

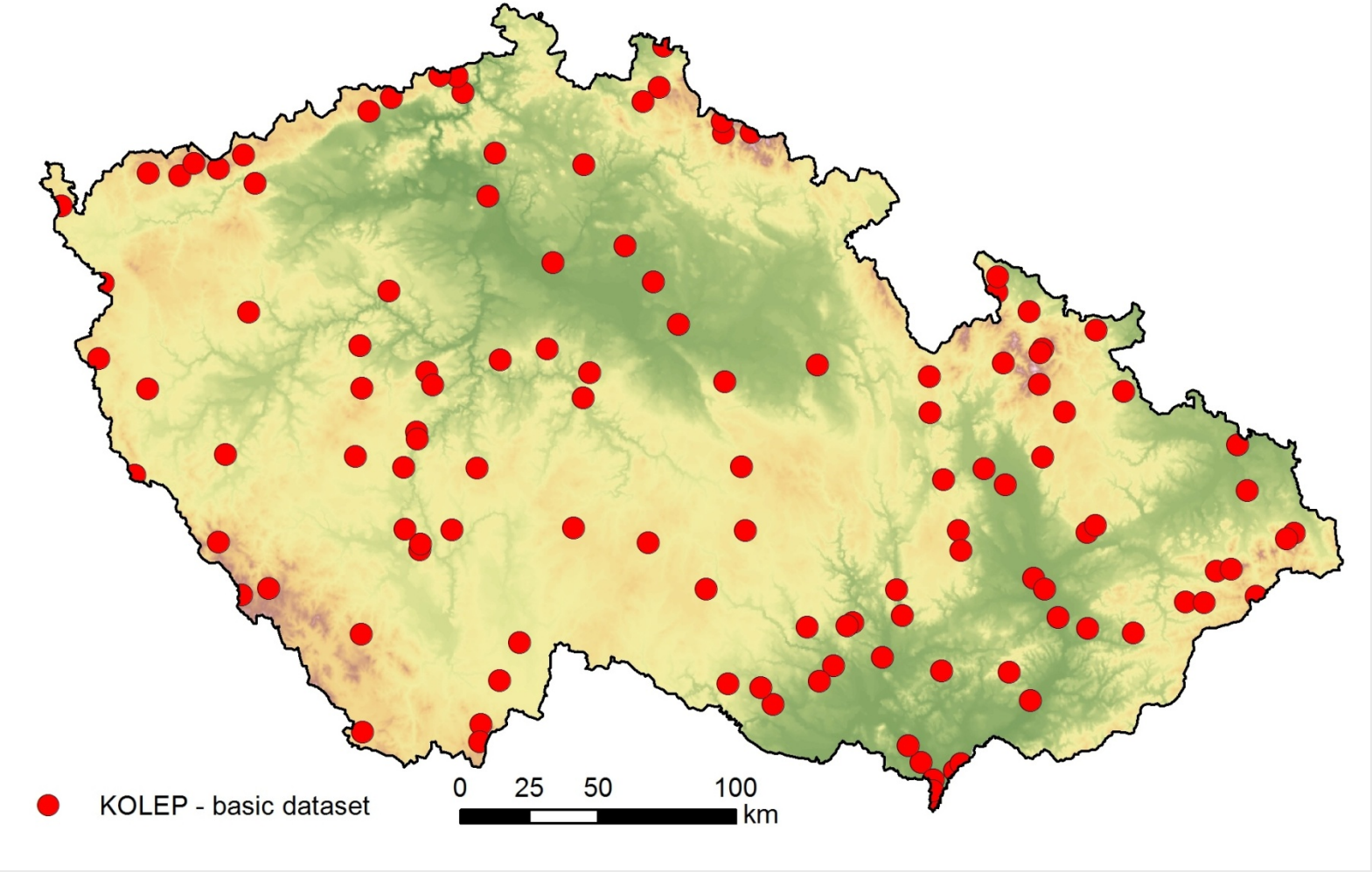
Objectives

- To predict the spatial distribution of **As, Cd, Pb, and Zn** in forest soils all over the Czech Republic
  - Polluting elements – partly anthropogenic origin
- **To analyze the importance of predictors:**
  - Differences between elements
  - Differences between depths
  - Effect of model types

Input data – element contents

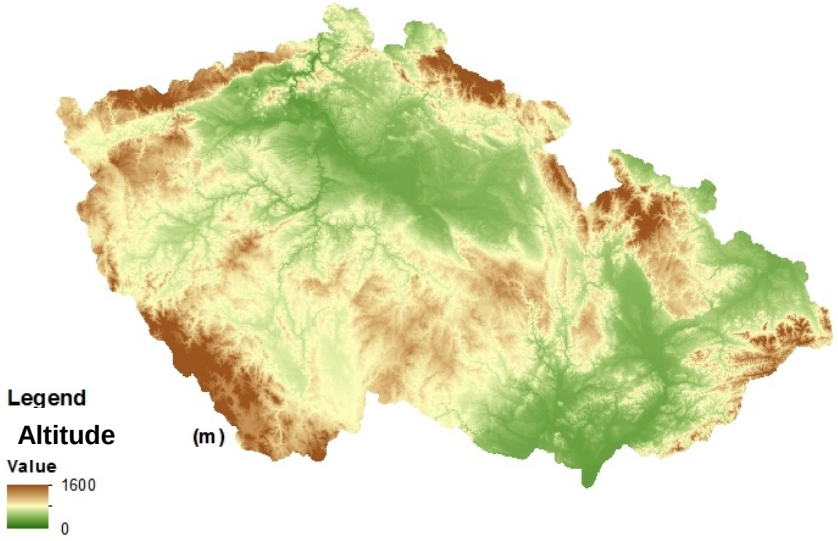
- 120 evenly distributed sampling points representing all principal **forest** types and categories
- Samples from 4 depths:
  - **1 - surface organic horizons (F+H)**
  - **2 - mineral horizons 0 to 2 cm**
  - **3 - mineral horizons 2 to 10 cm**
  - **4 - mineral horizons 10 to 20 cm**
- **As, Cd, Pb, Zn content**
  - Aqua regia (AR) digestion – pseudototal content
  - ICP-OES

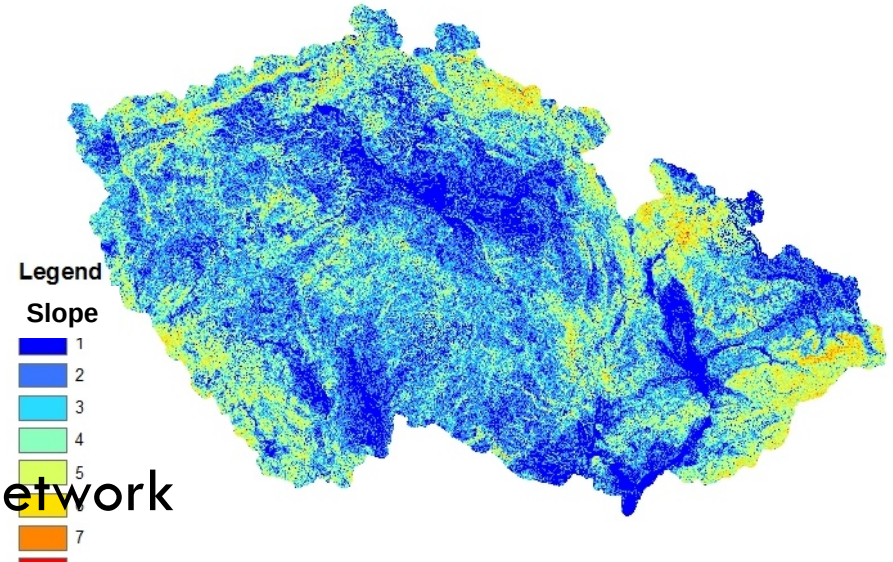
Sampling locations



Environmental covariates

- 13 terrain attributes (DTM)
  - Altitude
  - Slope
  - sin(Aspect)
  - cos(Aspect)
  - Cross-Sectional Curvature
  - Longitudinal Curvature
  - Convergence Index
  - Catchment Area
  - Topographic Wetness Index
  - LS Factor
  - Channel Network Base Level
  - Relative Slope Position
  - Vertical Distance to Channel Network

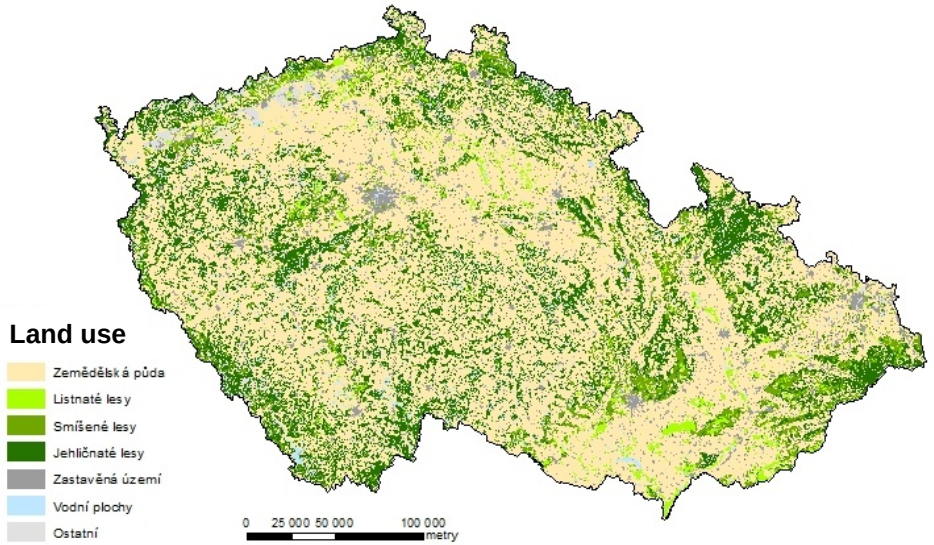


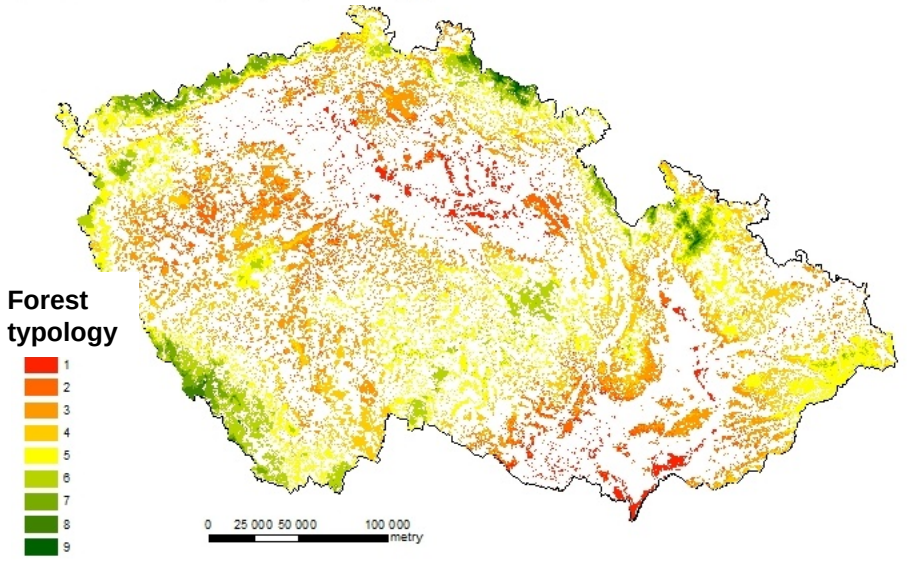


Environmental covariates

- 13 terrain attributes (DTM)
- Land use/land cover
  - Forest type
- Forest typology
  - Natural vegetation zones
- Soil class
- Parent material
  - Rock type, acidity, texture
- Position (coordinates)

- Grid: 1 x 1 km

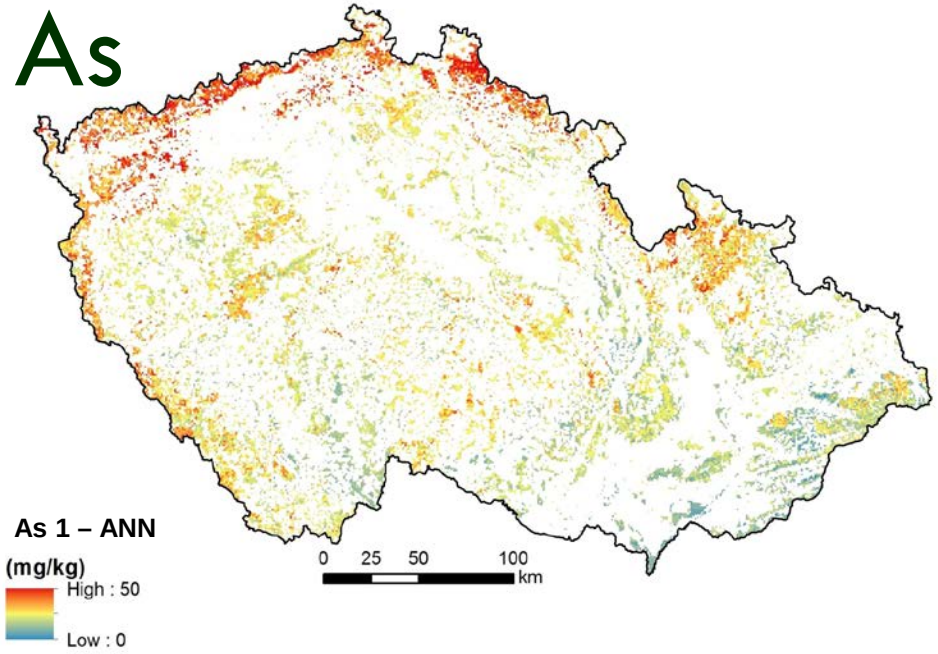


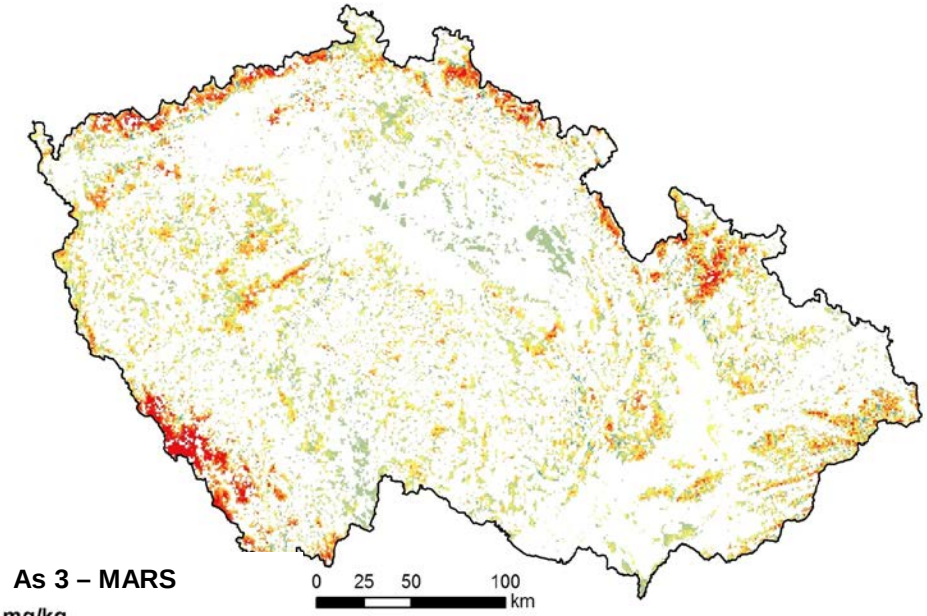


Prediction methods

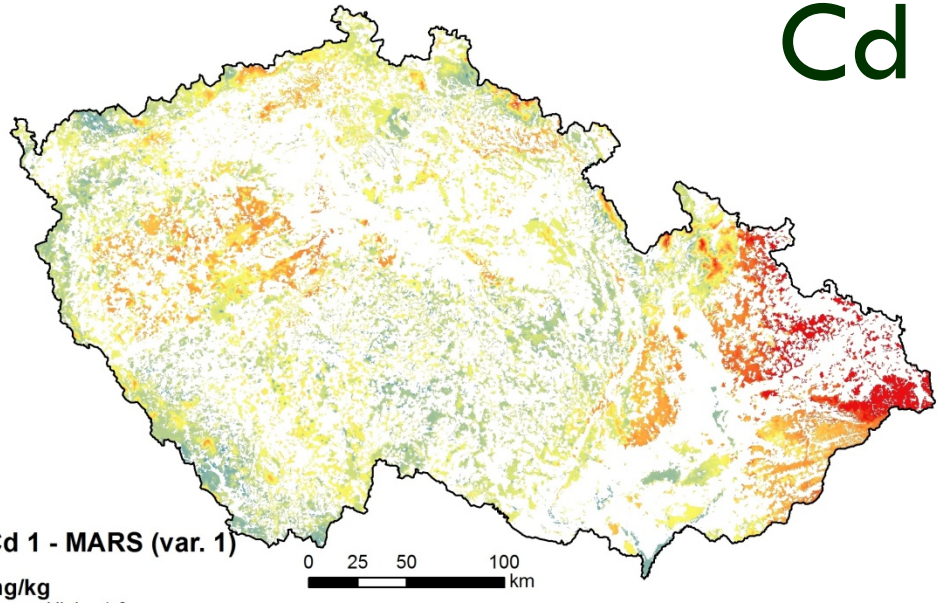
- Boosted trees (BT)
- Random forests (RF)
- Multivariate adaptive regression splines (MARS)
- Artificial neural networks (ANN)

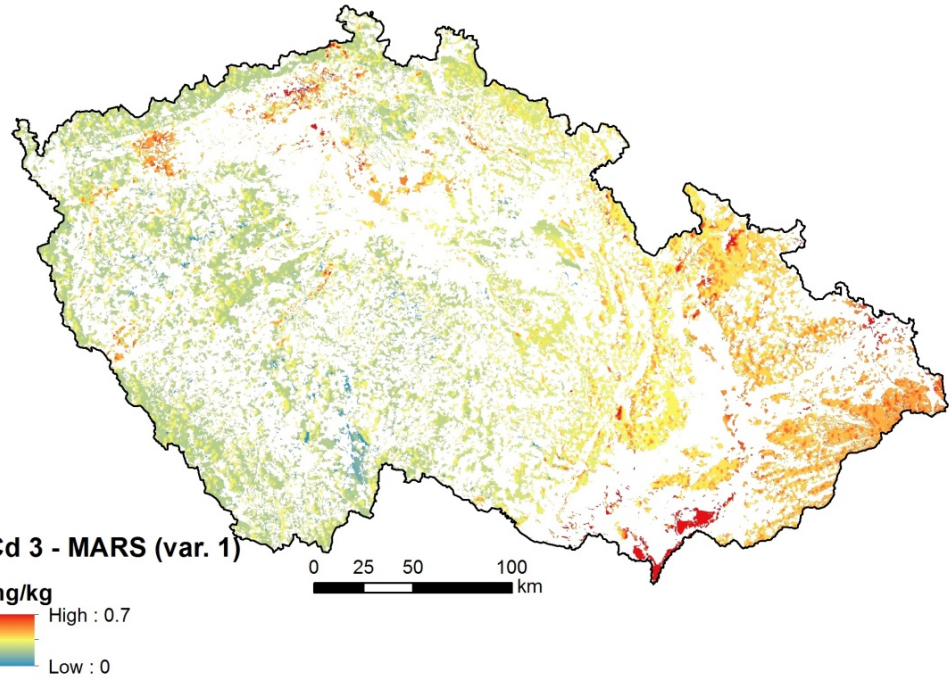
As


As 1 – ANN
(mg/kg)
High : 50
Low : 0
0 25 50 100
km

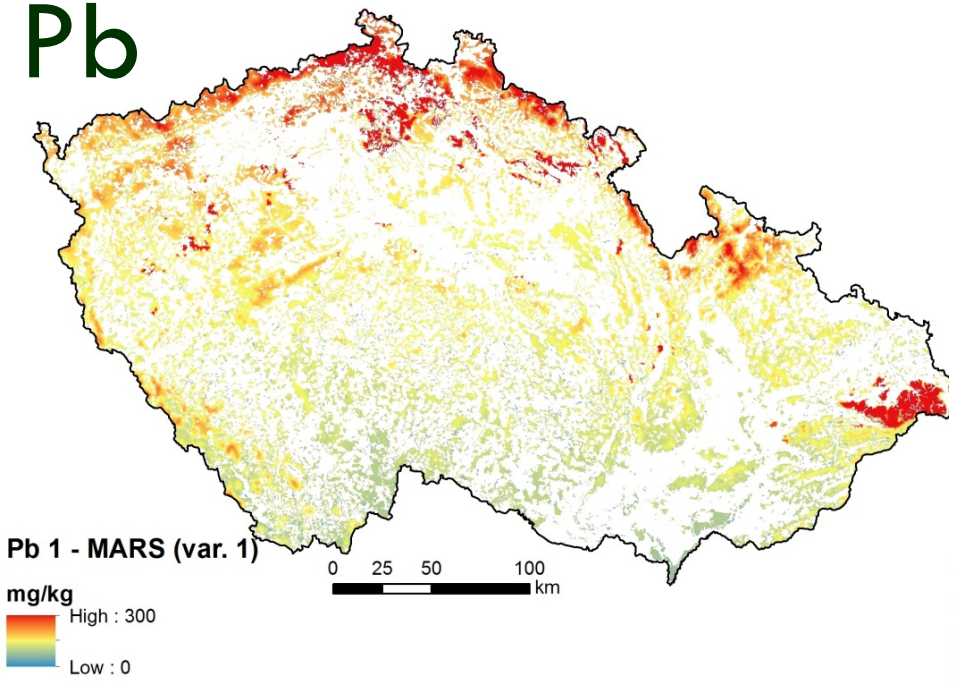

As 3 – MARS
mg/kg
High : 60
Low : 0
0 25 50 100
km

Cd

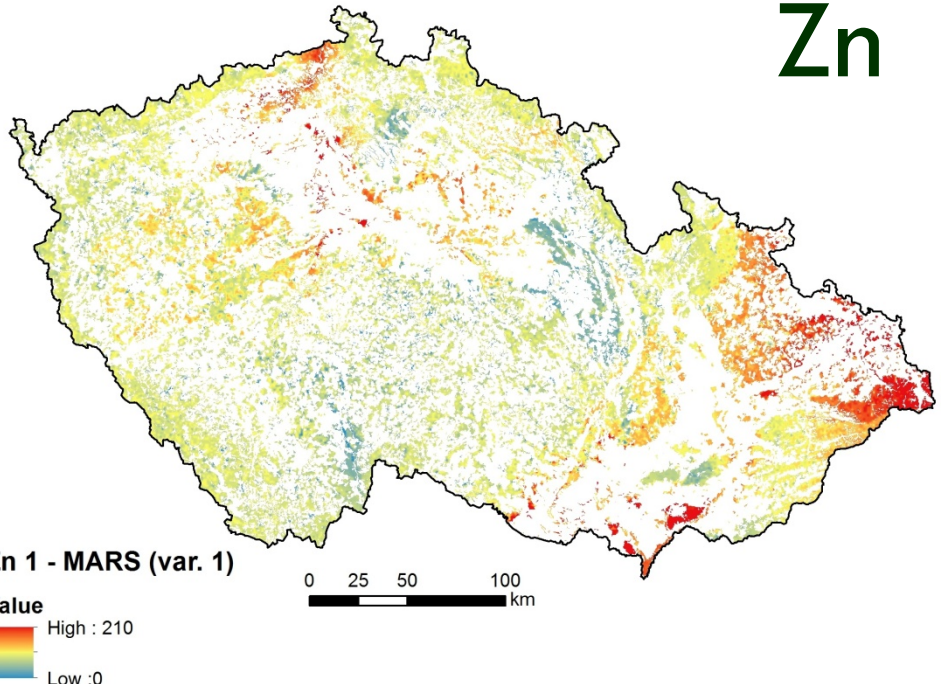

Cd 1 - MARS (var. 1)
mg/kg
High : 1.6
Low :0
0 25 50 100
km

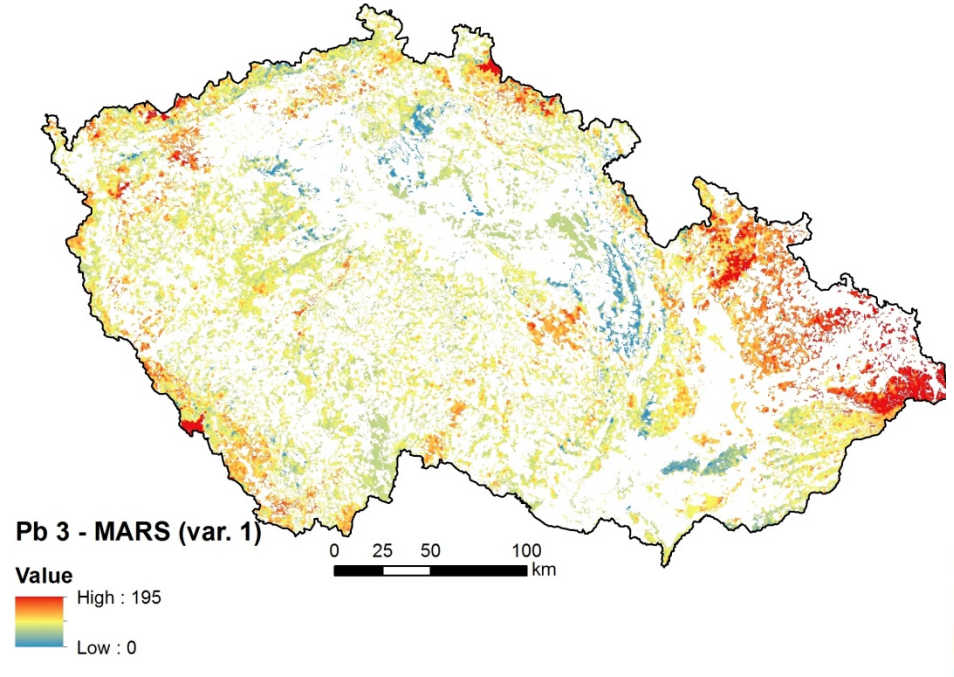

Cd 3 - MARS (var. 1)
mg/kg
High : 0.7
Low : 0
0 25 50 100
km

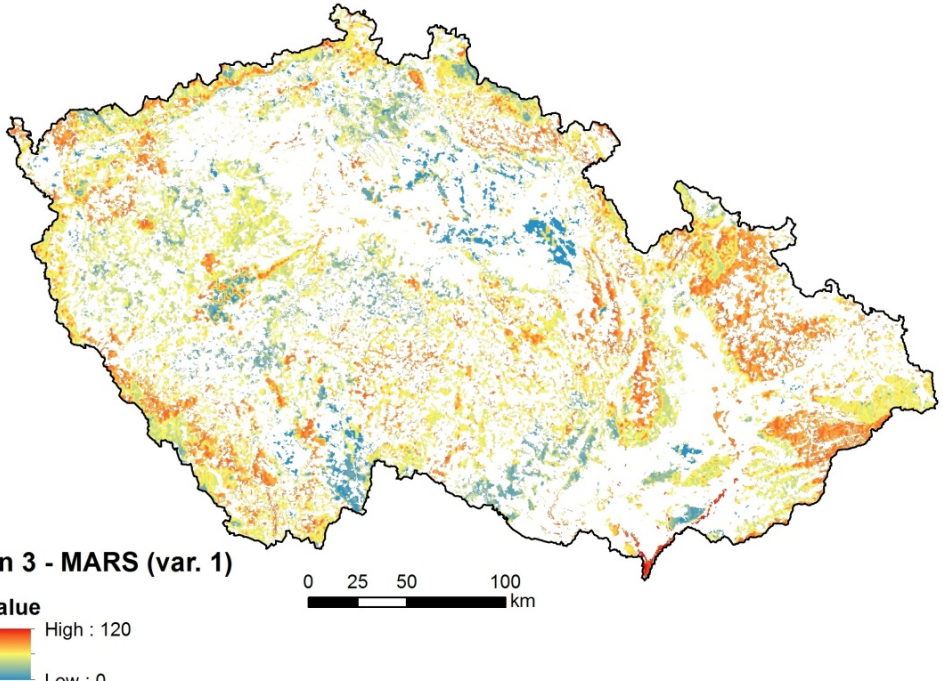
Pb


Pb 1 - MARS (var. 1)
0 25 50 100
km
mg/kg
High : 300
Low : 0

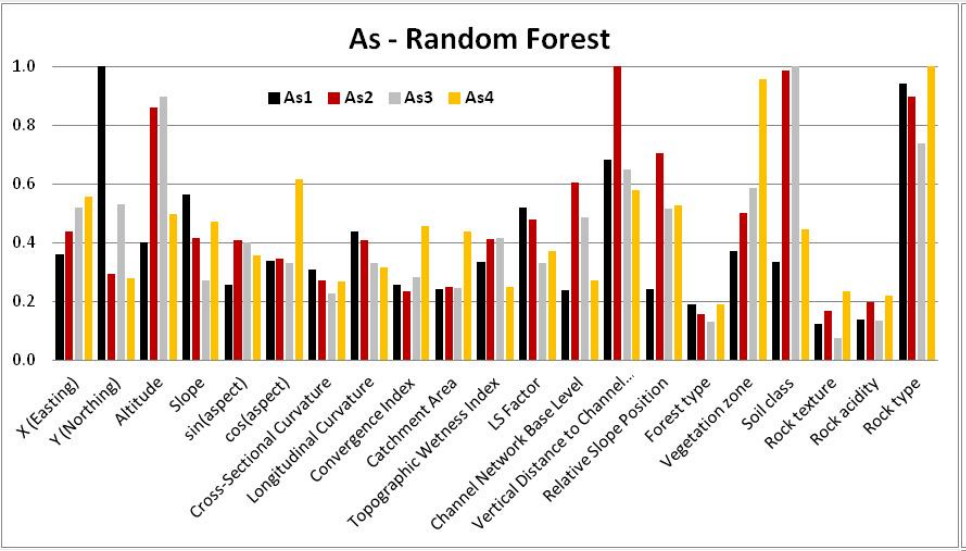
Zn

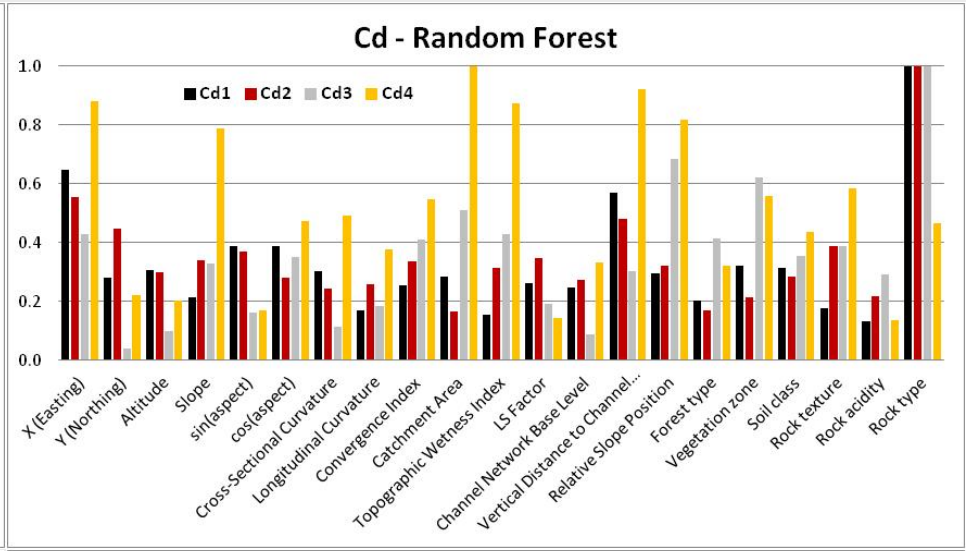

Zn 1 - MARS (var. 1)
0 25 50 100
km
Value
High : 210
Low :0

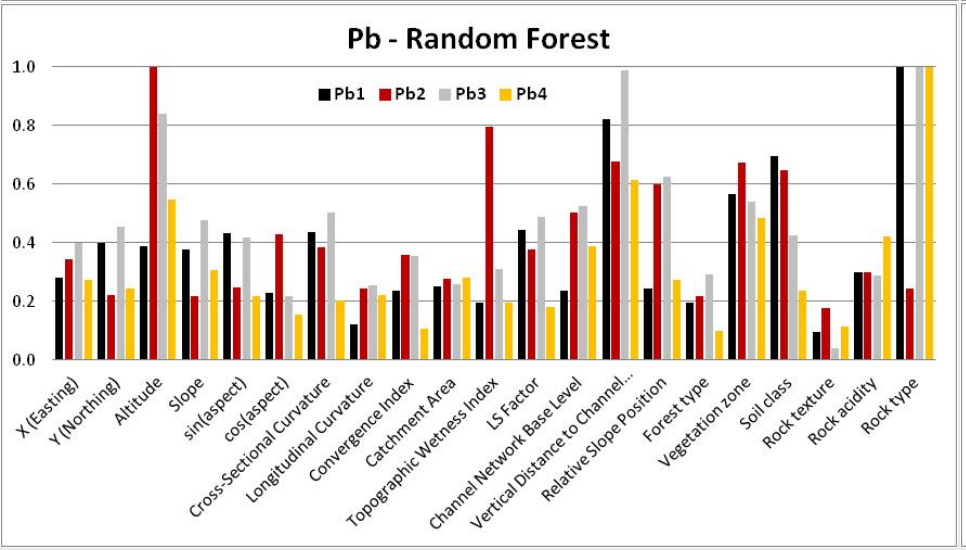

Pb 3 - MARS (var. 1)
0 25 50 100
km
Value
High : 195
Low : 0

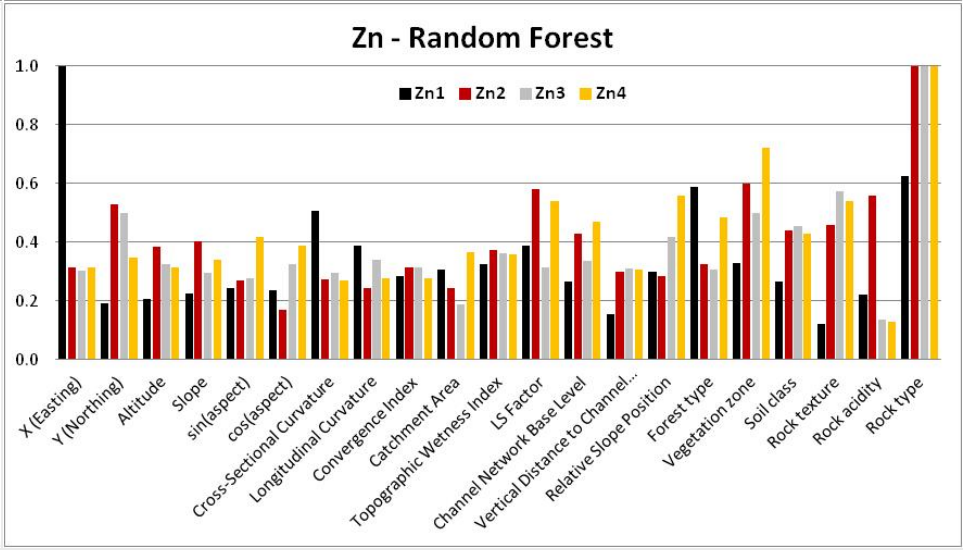

Zn 3 - MARS (var. 1)
0 25 50 100
km
Value
High : 120
Low : 0

Predictors importance – comparison between elements and depths (RF)

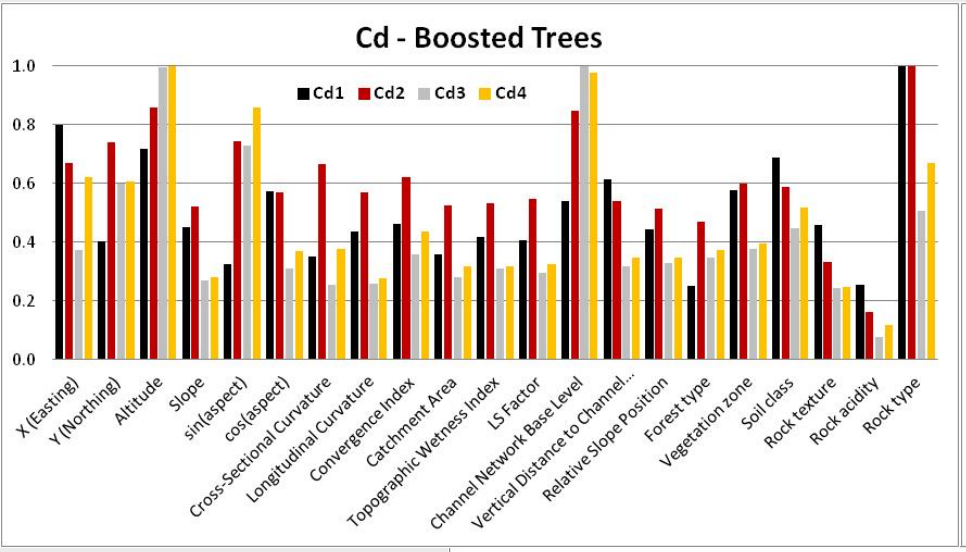


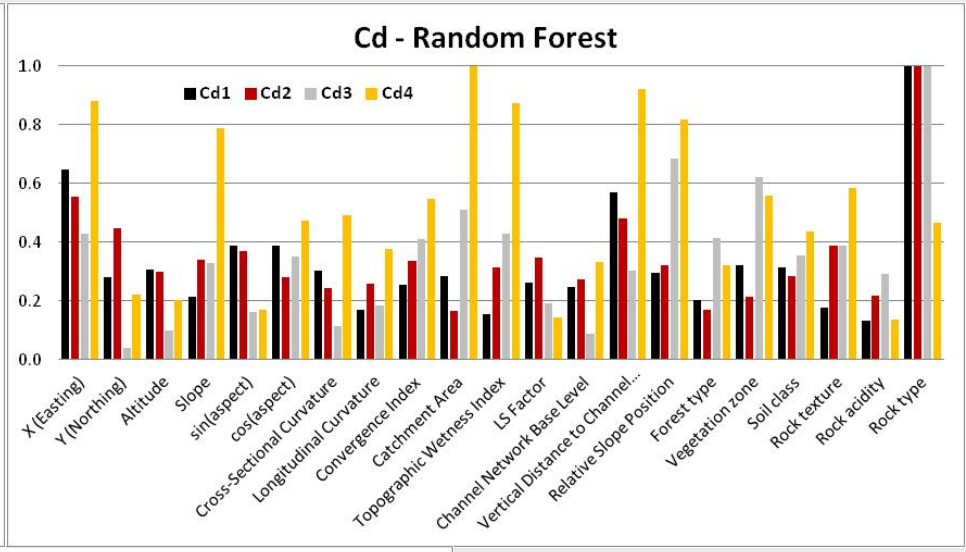




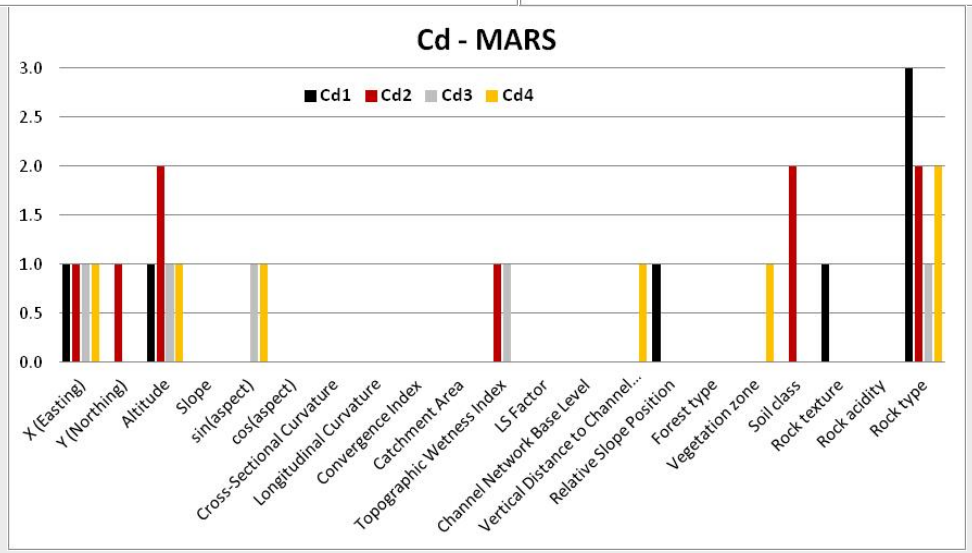


Predictors importance – comparison between models (Cd as an example)





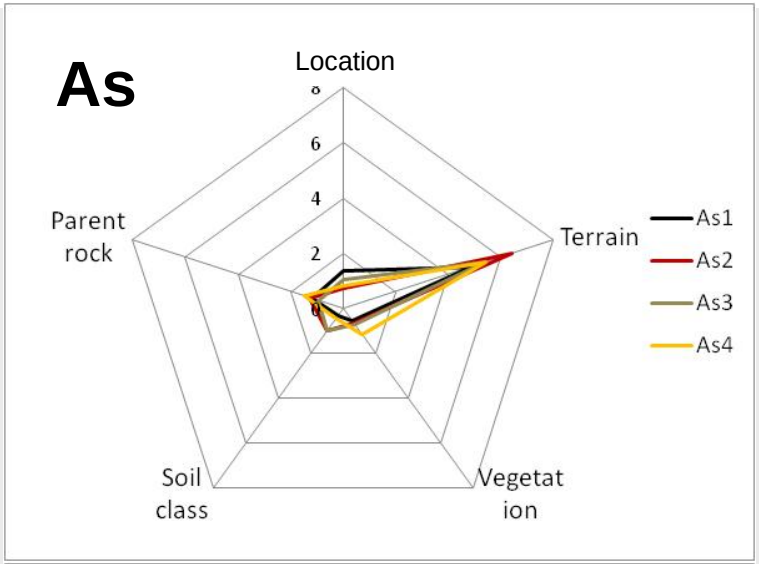
$R^2_{BT1} = 61.9$

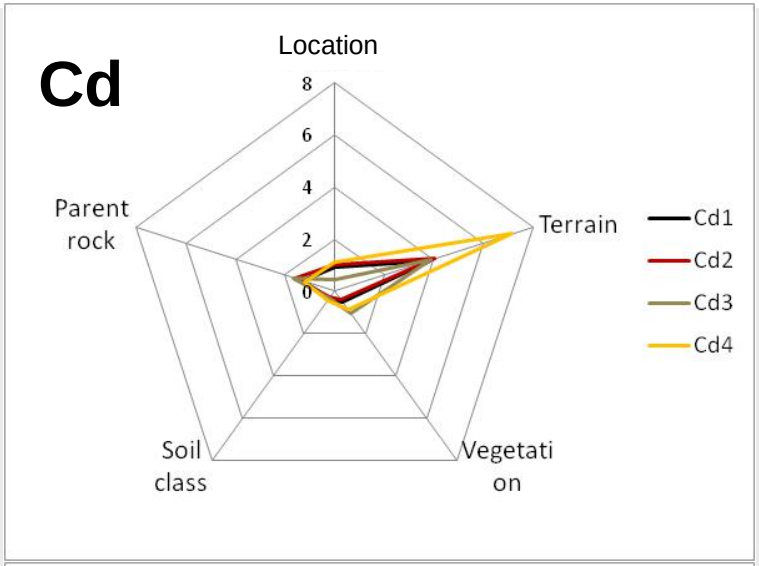


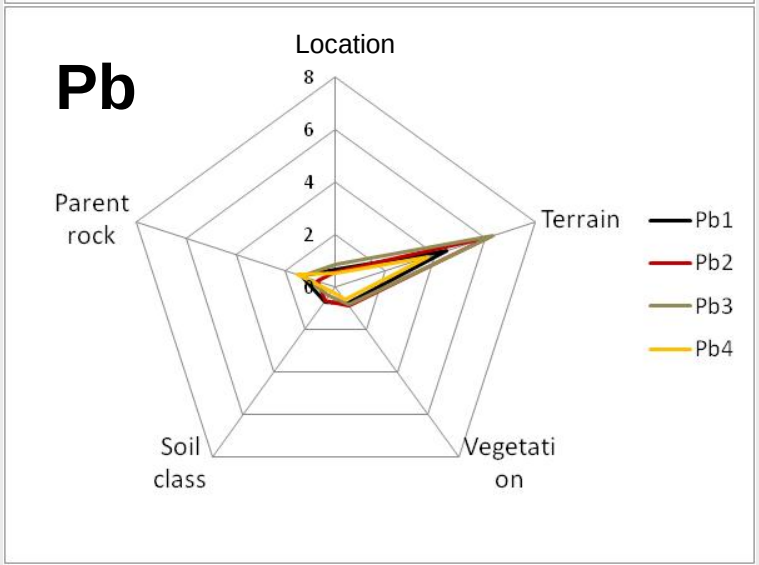
$R^2_{RF1} = 44.6$

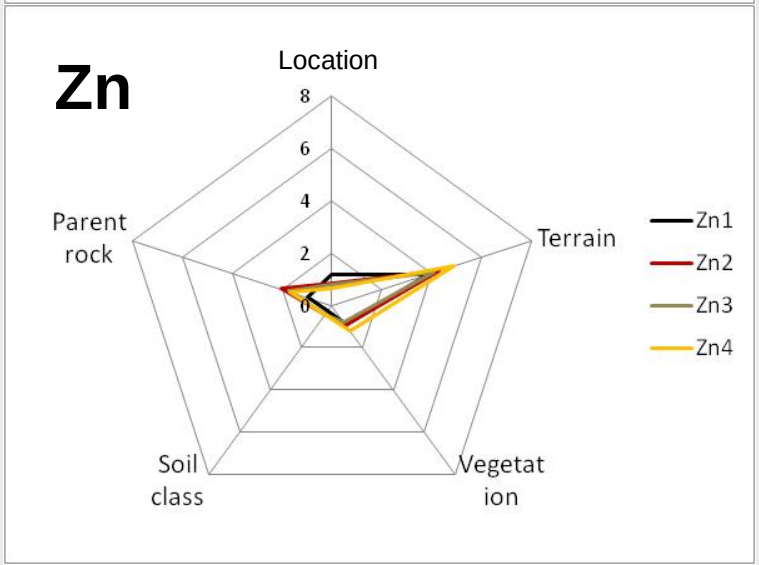
$R^2_{MARS1} = 59.6$

Predictors importance – relative contribution of different groups

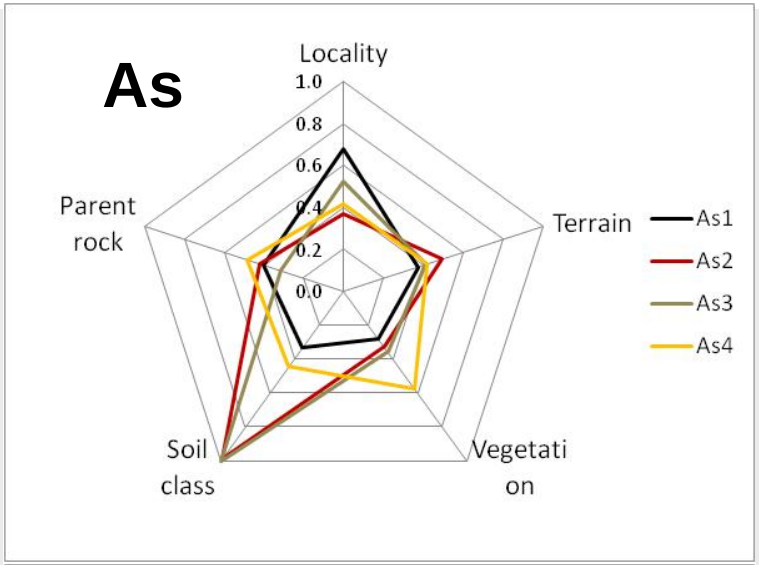


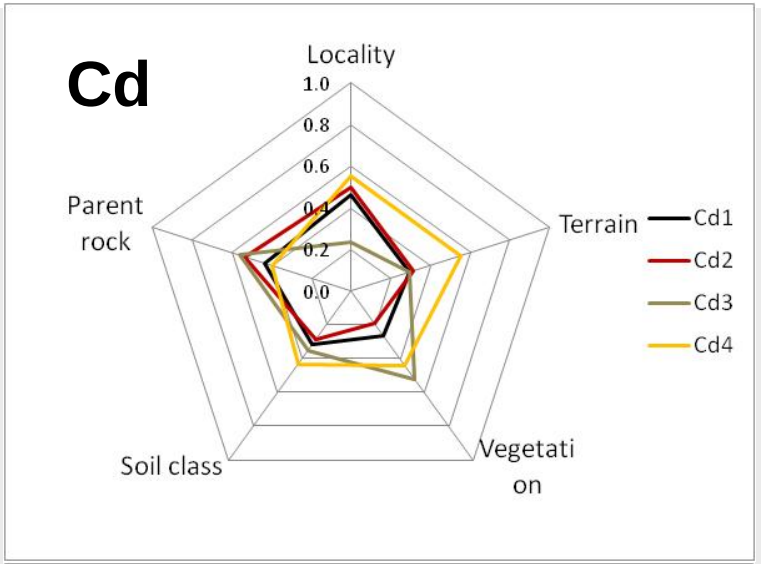


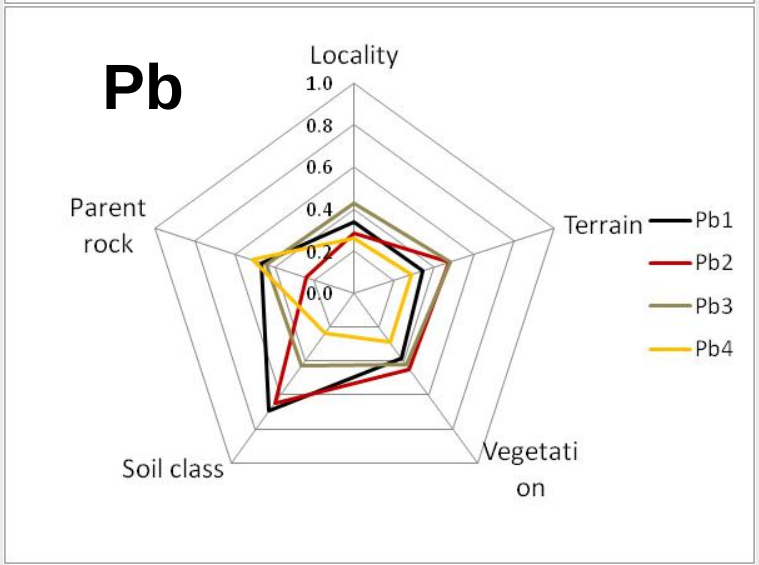


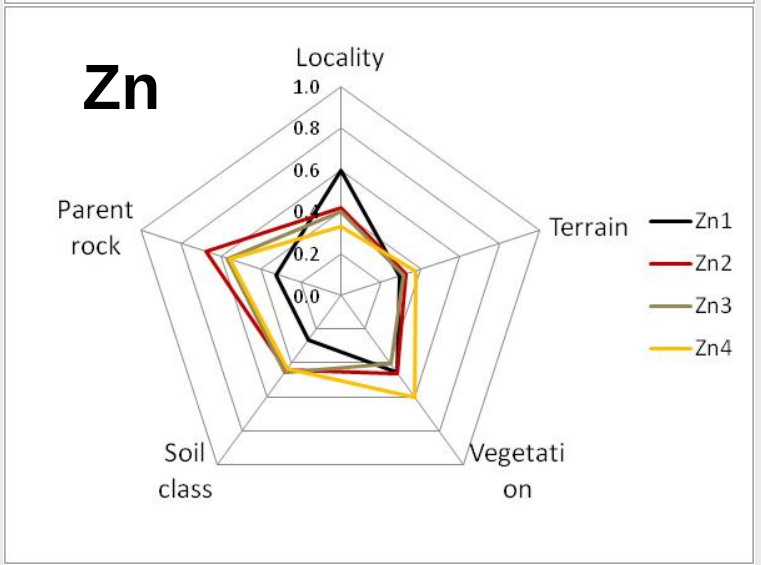


Predictors importance – relative contribution of different groups (adjusted)









Conclusions

- Even the data whose spatial distribution is strongly influenced by human activity can be spatially predicted using DTM, LU/LC etc. as predictors
- Analysis of predictors can provide another insight into the factors of spatial distribution
  - Anthropogenic vs. natural origin
  - Effect of terrain, vegetation, parent material etc.
- Appropriate selection of model types and best predictors is a crucial issue


Thank you for your attention