APPLICABILITY OF ARTIFICIAL NEURAL NETWORKS FOR PREDICTING CONCENTRATIONS OF CHEMICAL ELEMENTS (VARIOUS CASE STUDIES)

Robert Šajn & Jasminka Alijagić

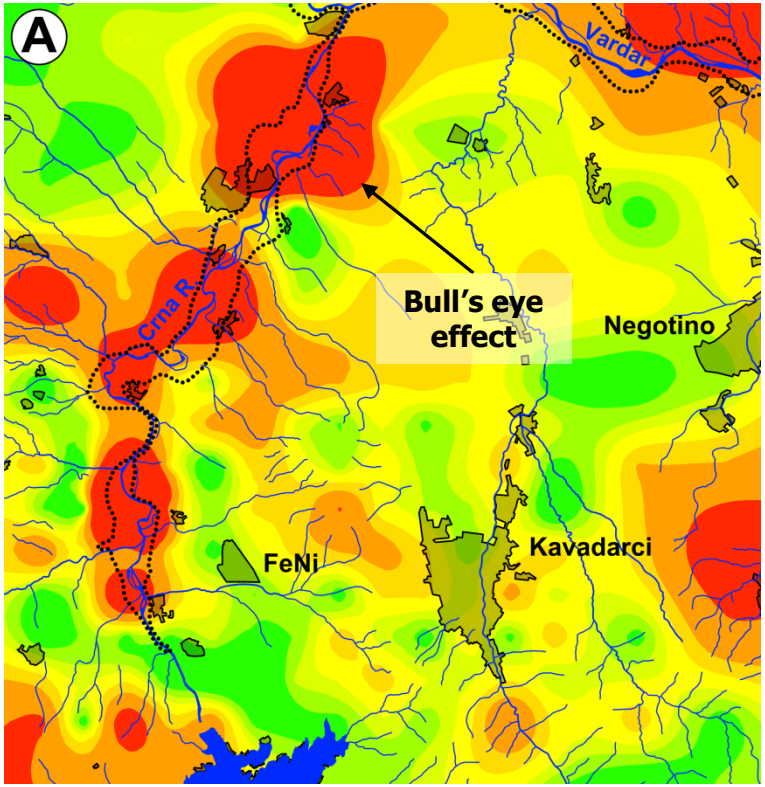
29th Task Force meeting of the ICP Vegetation, Dubna 1.3. – 3.3. 2016

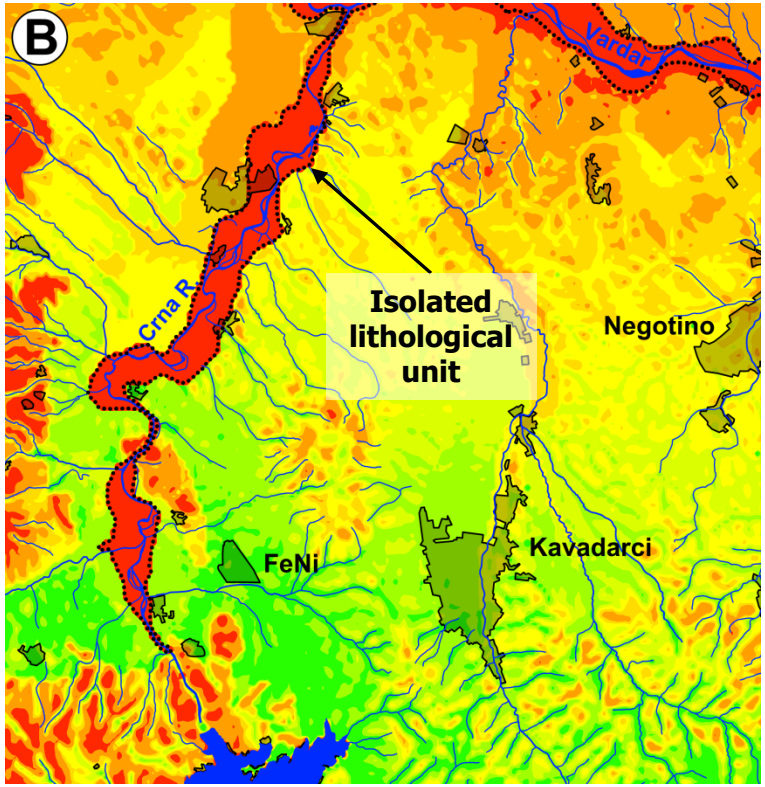
Geološki zavod Slovenije

Linear mathematical method vs. ANN

Inverse Distance
Mouving Average
Triangulation
Polinomial Regresion
Kriging

ANN



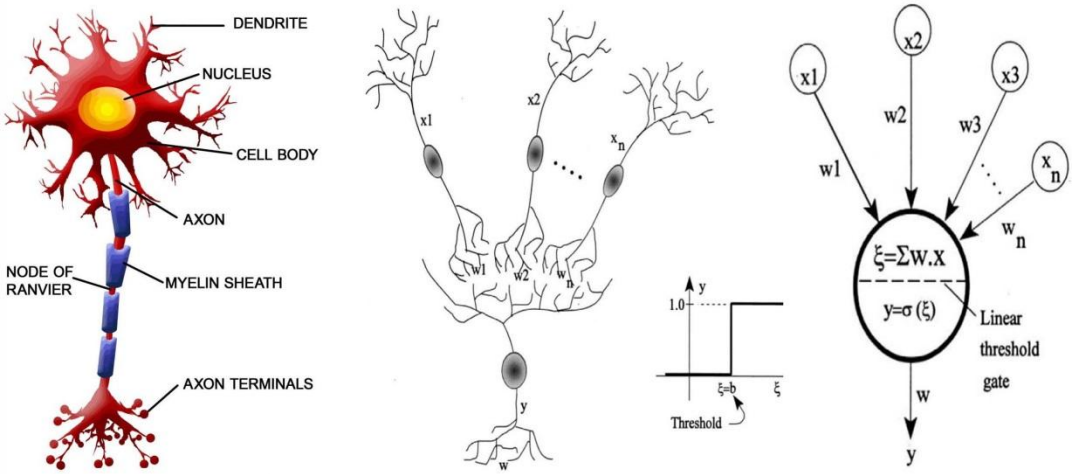


Main problem of linear methods is that their concentrations depend only on distance. Each of this method will give similar results. Different things can influence the results what can lead to the wrong interpretation, such as the Bull's eye contours. In other side ANN is giving much better results

GeoZS

Artificial neural networks

Artificial Neural Network - A computer simulation of human neurons. A system (implemented in software or hardware) that is intended to emulate the computing structure of neurons in the human brain. The main challenge is to actually produce a modelling system that can handle a large number of input and output parameters.

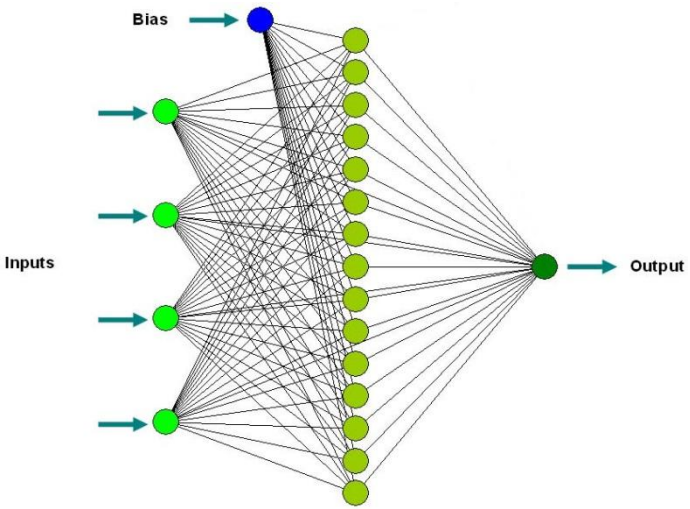


Biological neuron and mathematical model of McCulloch and Pitts neuron

A neuron is a processing unit in a neural network. It is a node that processes all fan-in from other nodes and generates an output according to a transfer function called the activation function. The activation function represents a liner or nonlinear mapping from the input to the output. A neuron is linked to other neurons by variable synapses (weights). Simple neuron model have been proposed by McCulloch and Pitts.

GeoZS

Multilayer Perceptron



Multilayer perceptron architecture

REASONS FOR APPLICATION

They can model extremely complex systems, which cannot be modeled by methods based on linear algebra.

No problems with the dimensionality - it can be arbitrary.

Due to well developed learning algorithms they are easy to use.

Input data - secondary attributes sourced from the DEM, land use, and remote sensing in combination with sparse and expensive soil measurements

Due to high cost and time consuming nature of soil sampling, research in developing methods for the creation of soil maps based in various prediction methods is becoming increasingly important. Each aforementioned applied modelling technique by itself helped us in reconstruction simultaneously different processes that influenced the entire study area.

GeoZS

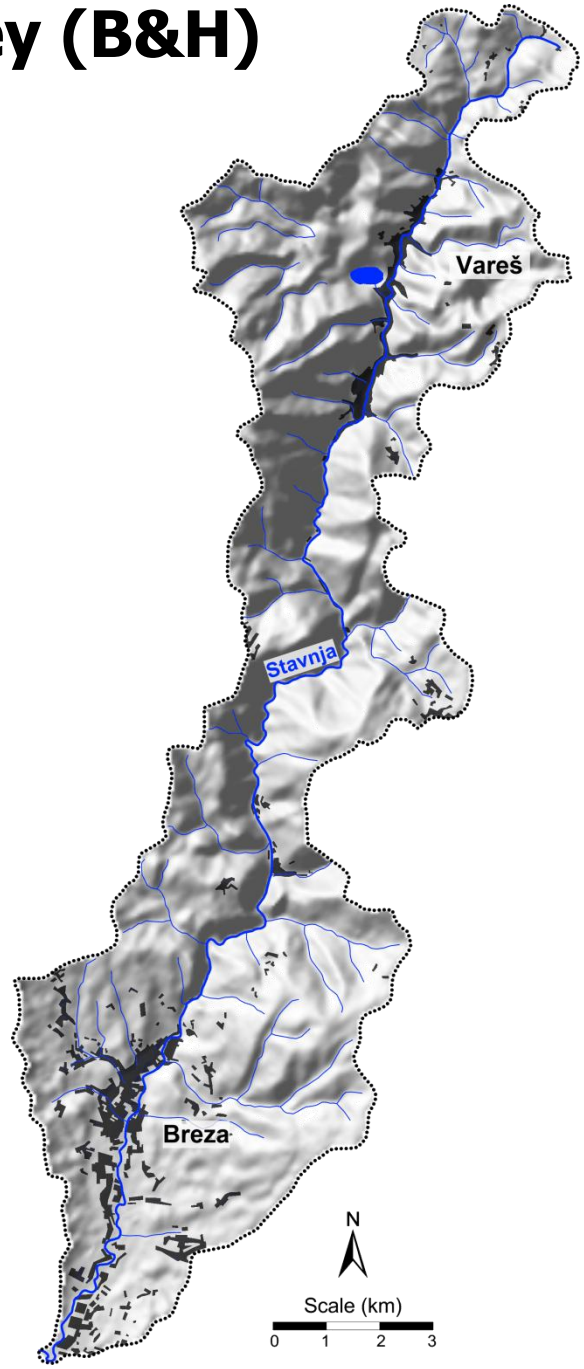
Case study Stavnja valley (B&H)

Small area (104 km^2)

Case study Stavnja valley (B&H)

Small area (104 km^2)

Shape of the Stavnja Valley (narrow valley)



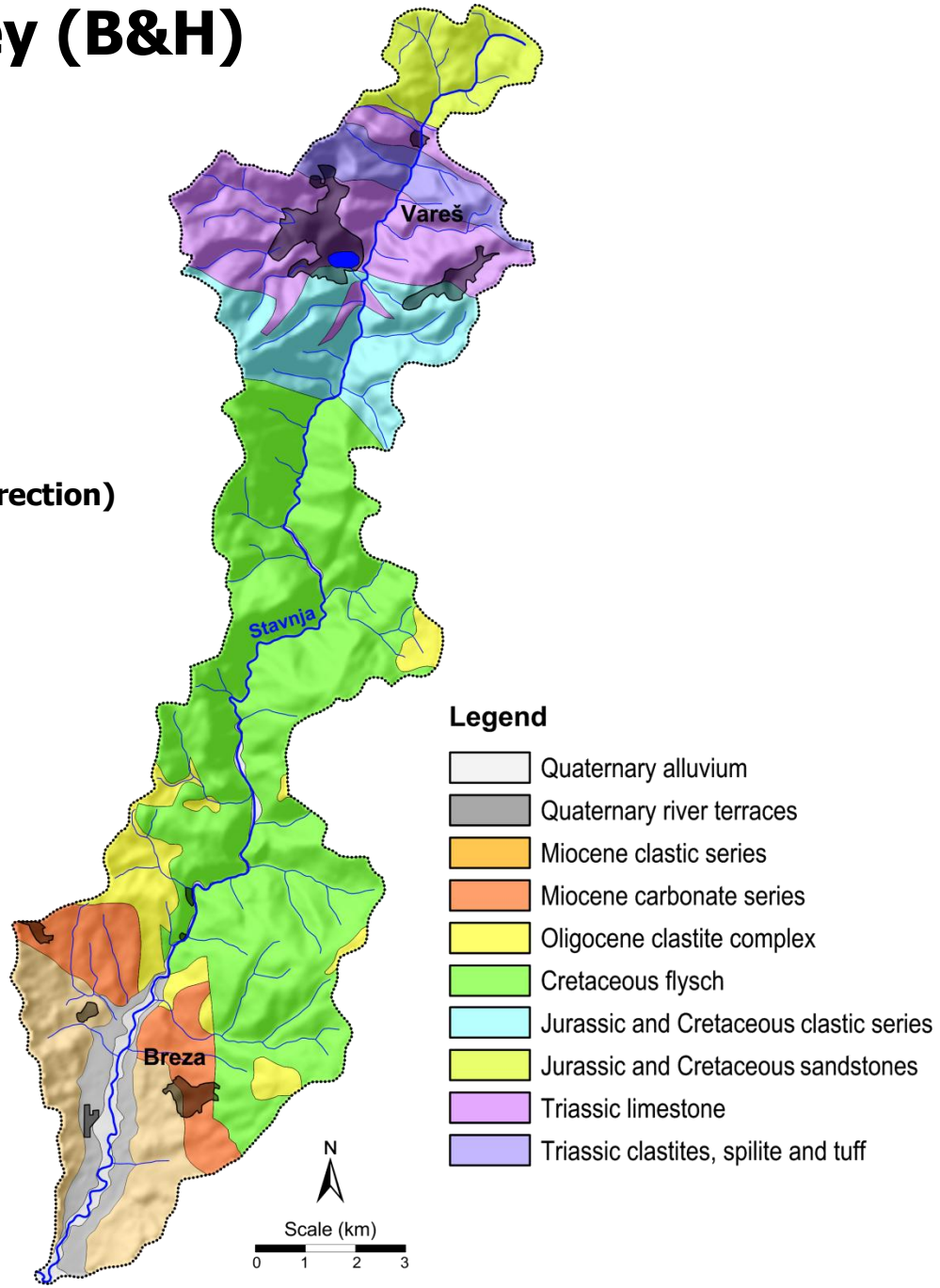
GeoZS

Case study Stavnja valley (B&H)

Small area (104 km^2)

Shape of the Stavnja Valley (narrow valley)

Geological background
(geological settings perpendicularly at the valley direction)



GeoZS

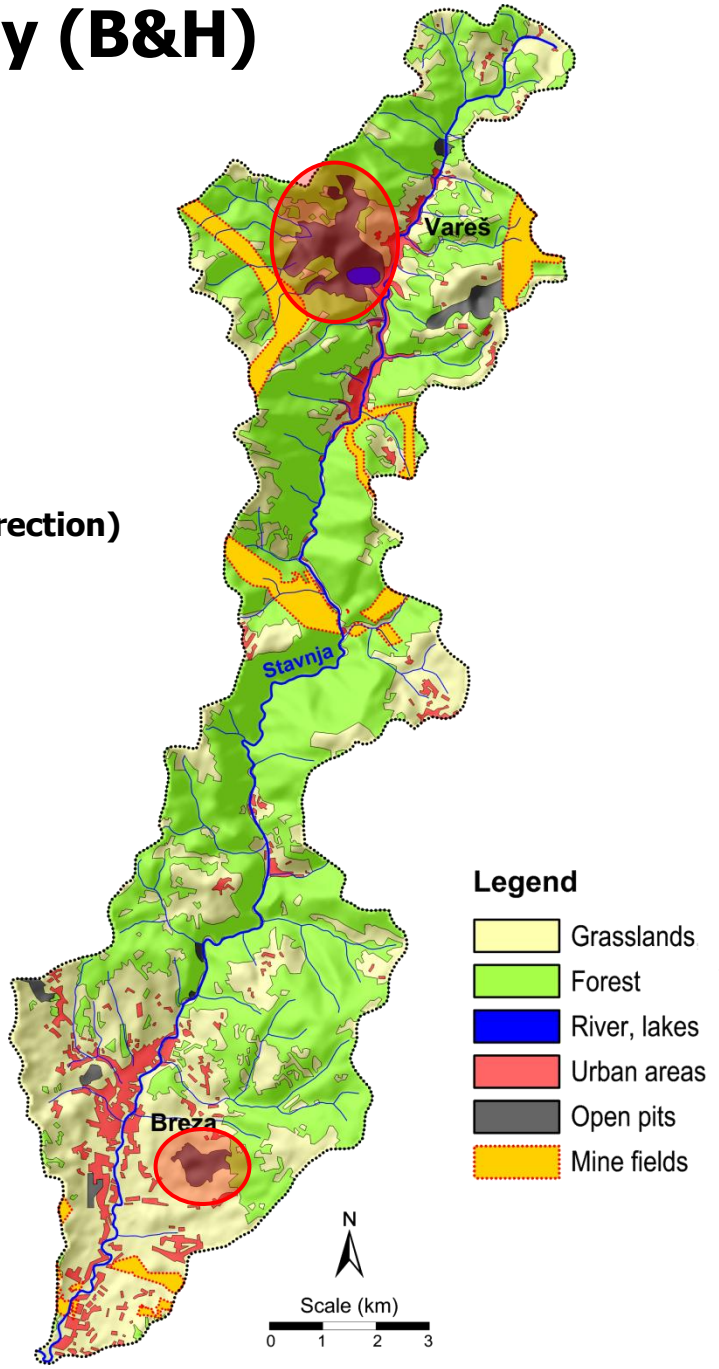
Case study Stavnja valley (B&H)

Small area (104 km^2)

Shape of the Stavnja Valley (narrow valley)

Geological background
(geological settings perpendicularly at the valley direction)

Intensive mining and metallurgical activities
Fe ore deposits (Smreka, Droškovac, Brezik)
Pb–Zn–Ba open pit (Veovača)
Coal mines Breza



GeoZS

Case study Stavnja valley (BiH)

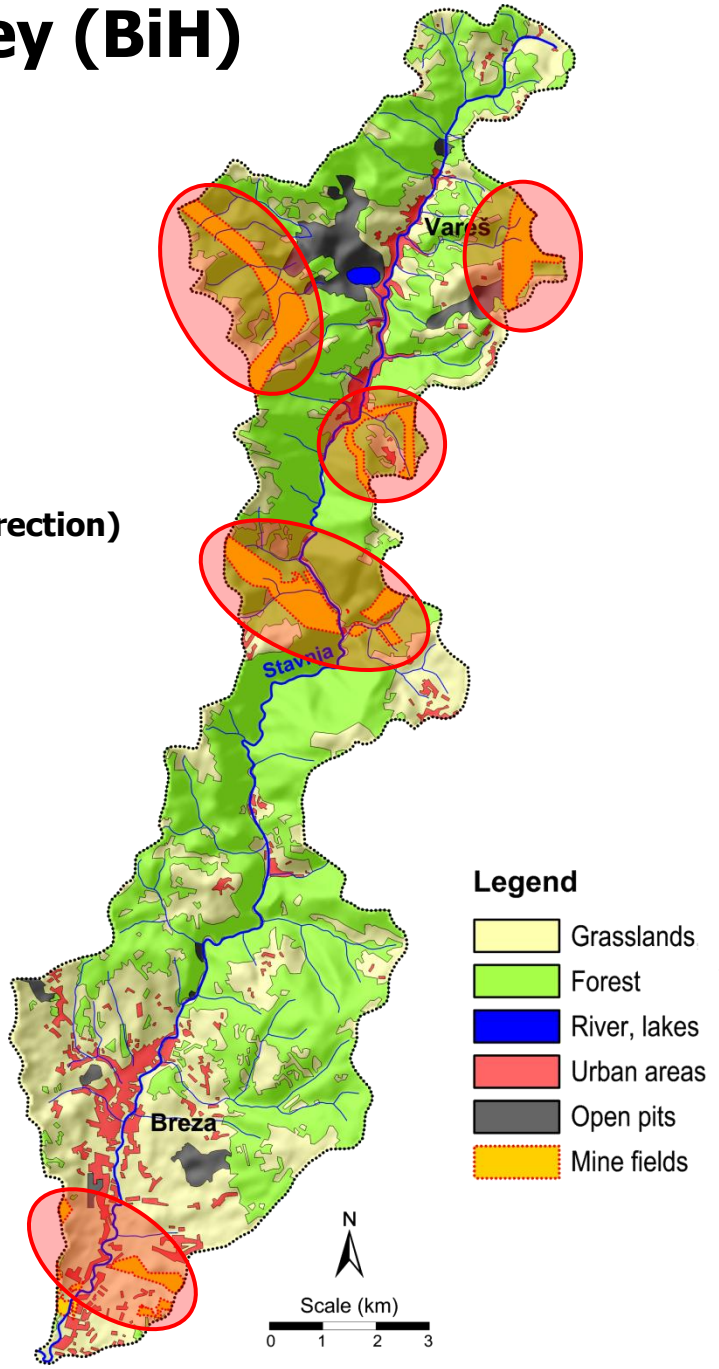
Small area (104 km^2)

Shape of the Stavnja Valley (narrow valley)

Geological background
(geological settings perpendicularly at the valley direction)

Intensive mining and metallurgical activities
Fe ore deposits (Smreka, Droškovac, Brezik)
Pb–Zn–Ba open pit (Veovača)
Coal mines Breza

Area of former milatary operation
The remain mine fields



Case study Stavnja valley (B&H)

Small area (104 km^2)

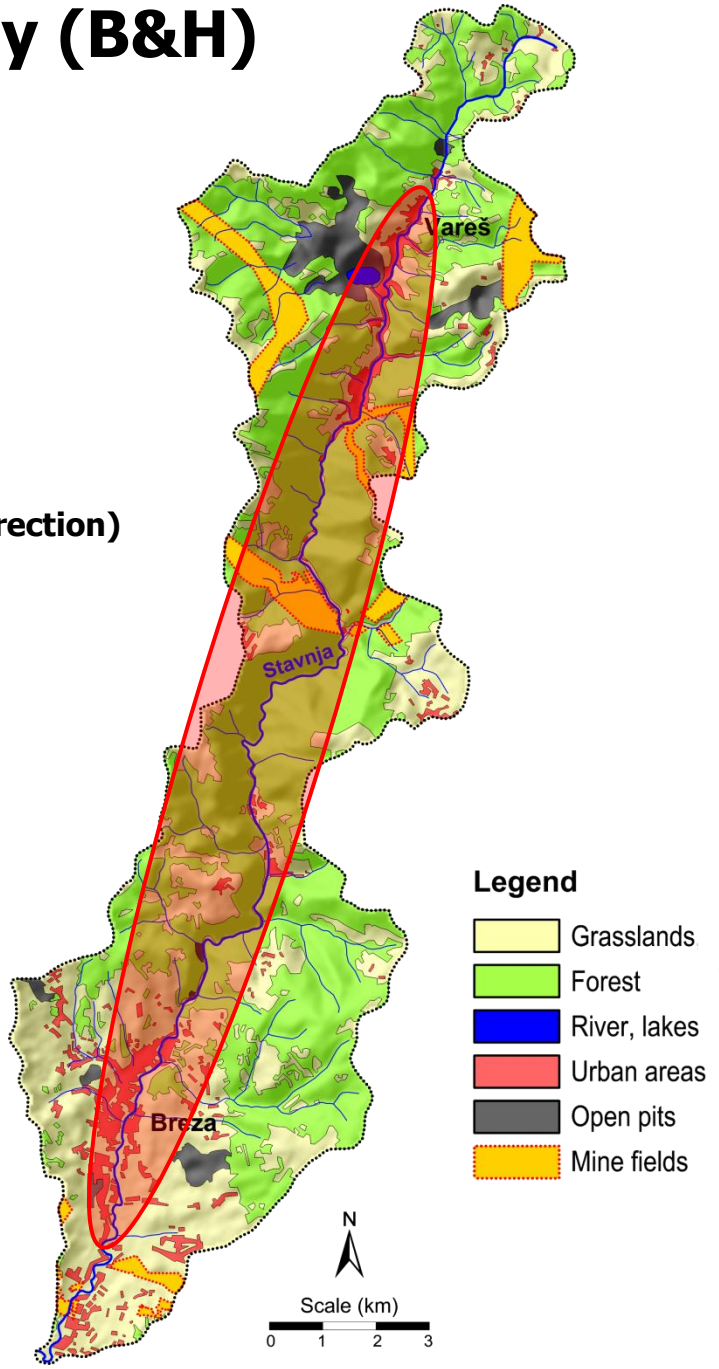
Shape of the Stavnja Valley (narrow valley)

Geological background
(geological settings perpendicularly at the valley direction)

Intensive mining and metallurgical activities
Fe ore deposits (Smreka, Droškovac, Brezik)
Pb–Zn–Ba open pit (Veovača)
Coal mines Breza

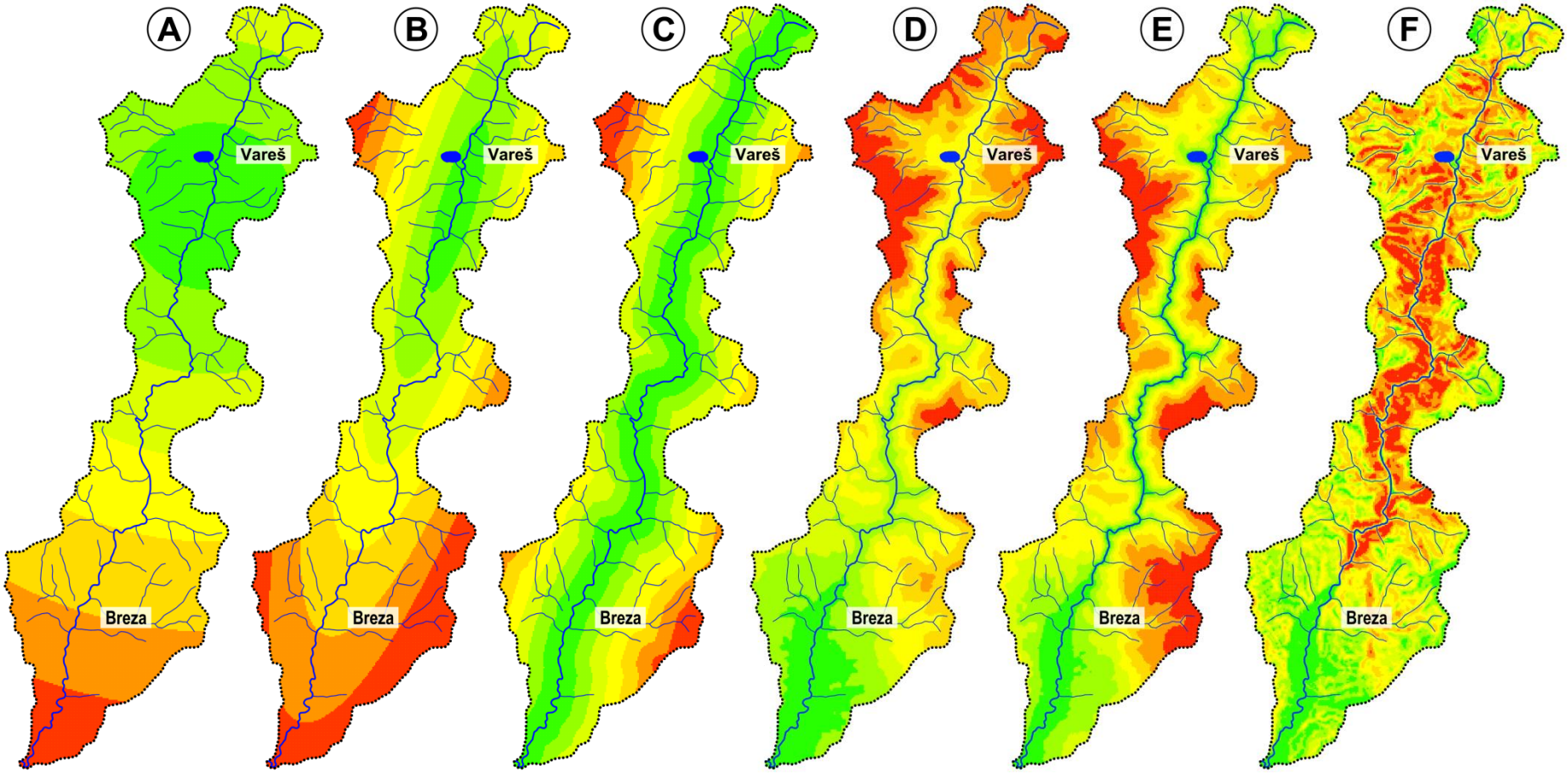
Area of former milatary operation
The remain mine fields

Expecting anthopogenic impact
along the valley direction



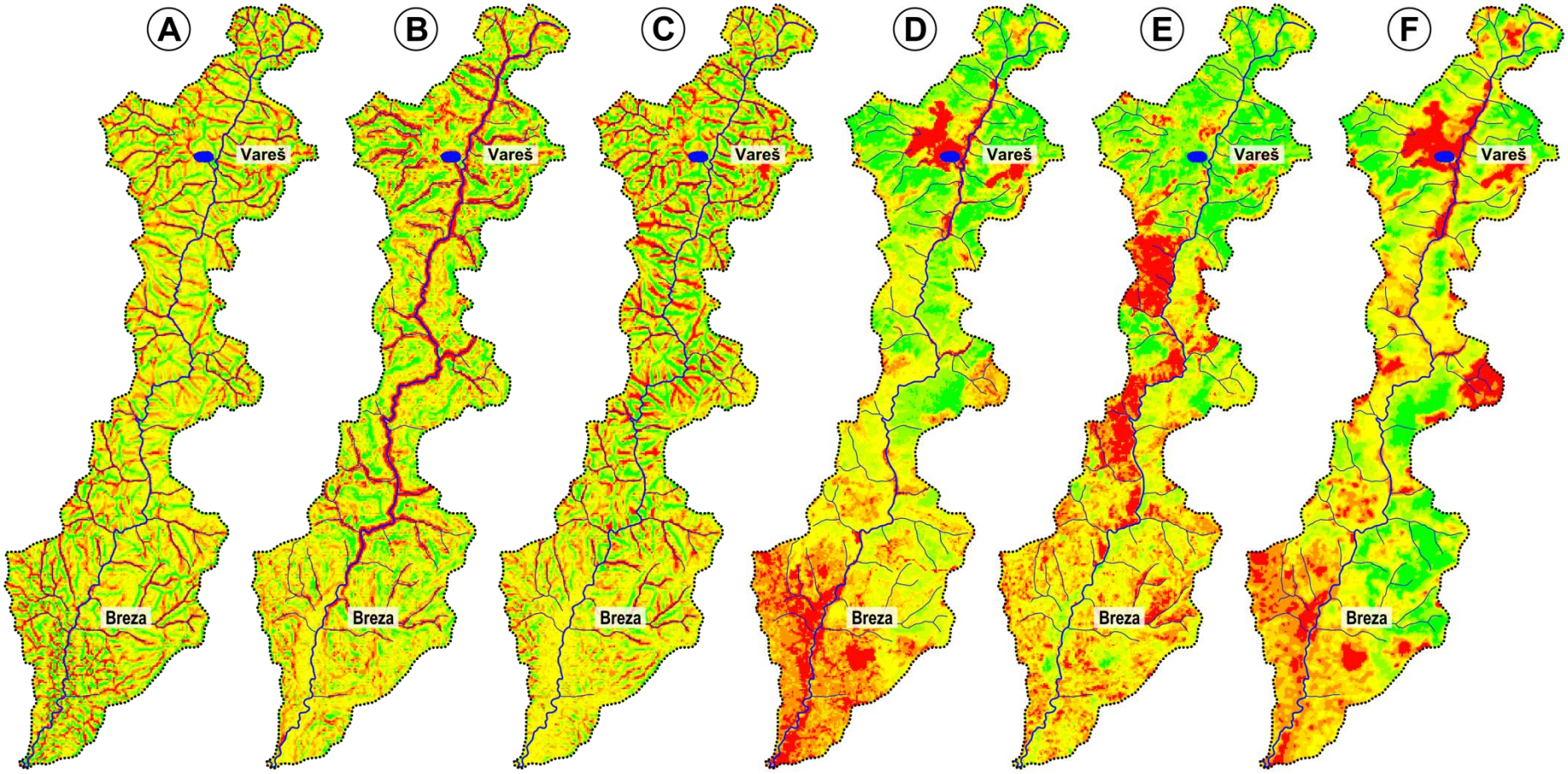
GeoZS

ANN-MP input data (I)



A – Absolute distance from the ironworks chimneys
B – Elliptical distance from the ironworks chimneys
C – Distance from the river Stavnja
D – Altitude above the see level (absolute)
E – Altitude above the bottom of Stavnja valley (relative)
F – Terrain Slope

ANN-MP input data (II)



A – Plan terrain curvature
B – Profile terrain curvature
C – Tangent terrain curvature

D – Visible spectrum, 0.45 – 0.69 μm (LandSAT)
E – Infrared spectrum, 0.76 – 0.90 μm (LandSAT)
F – Thermal radiation, 10.4 - 12.5 μm (LandSAT)

ANN calculation summary

Categorical input data

Geological map
Landuse map

Numerical input data

Distances
Digital elevation model and its derivates
LandSat satellite images (B10-B70)

Learning and recall

Learning data: 101 sampling points (automorphic soil)
10 sampling points (alluvial soil)
Recall data: 41 471 locations (Grid 50 X 50 m)

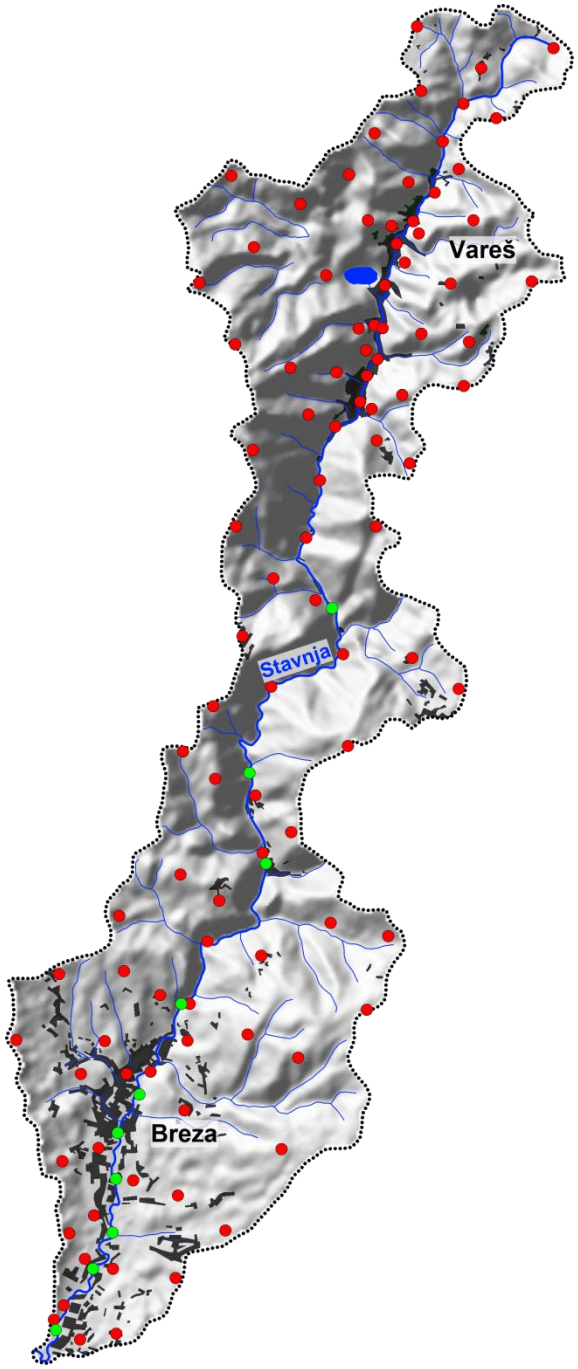
ANN Training

22 parameters

Multilayer perceptron - Hidden units - 220

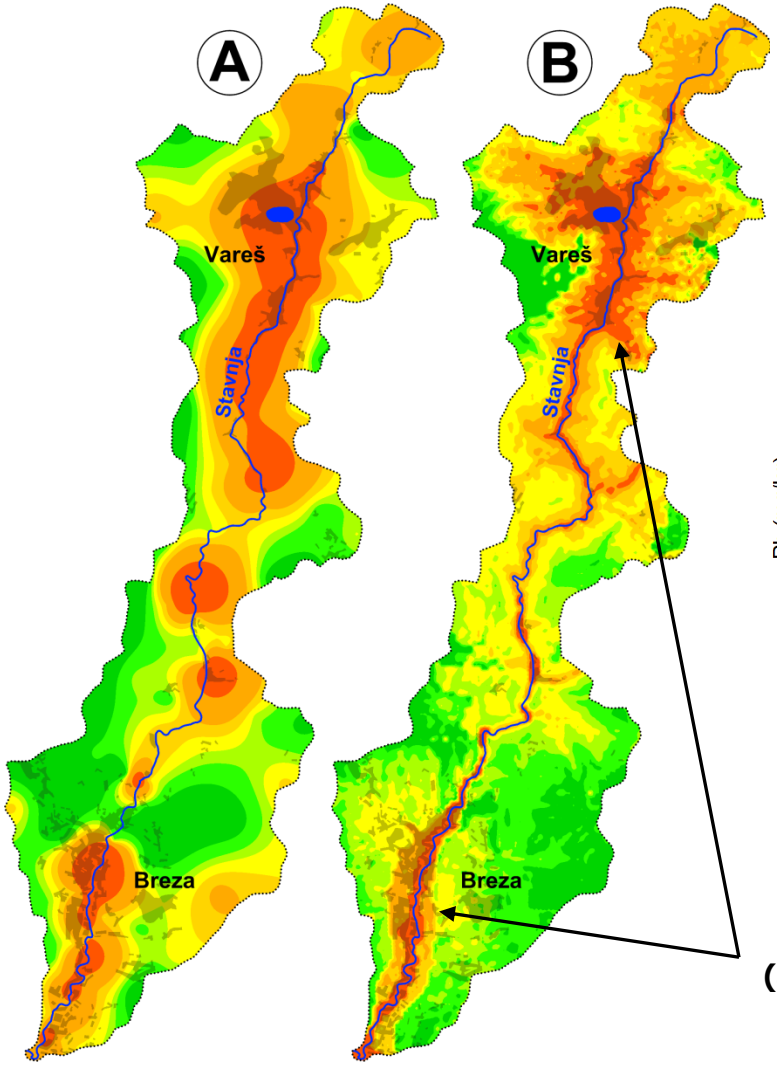
Train networks - 25 (the final model representing an average of 5 most logical solution)

Each training model contain: Training perfection, Test perfection, Validation perfection, All perfection, Training error, Test error, Validation error, Training algorithm, Hidden activation, and Output activation.

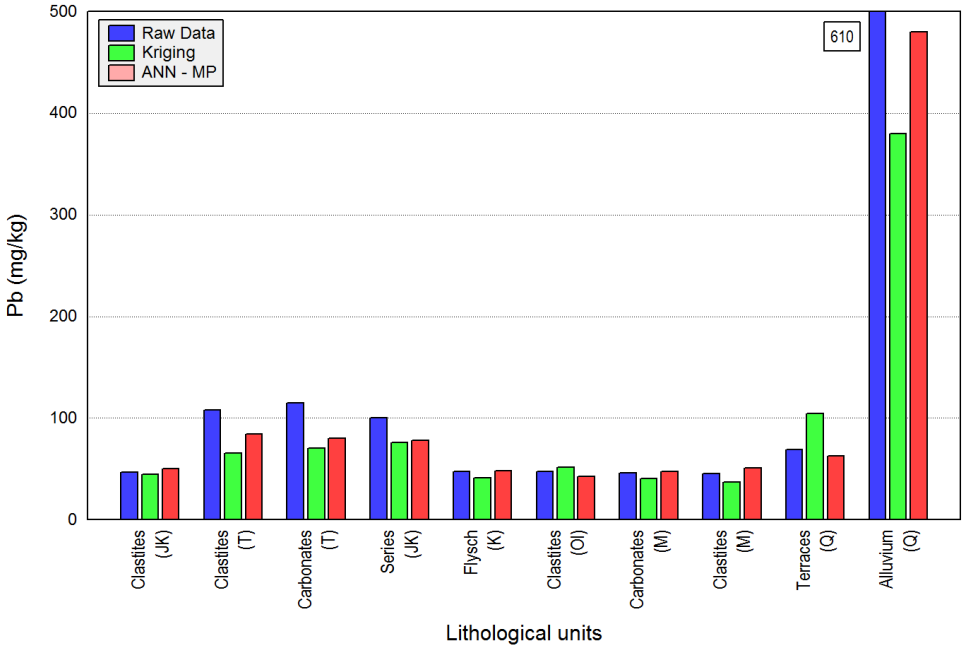


GeoZS

Spatial distribution of lead (Pb)



A – Standard kriging method
B – ANN (multilayer perceptron)



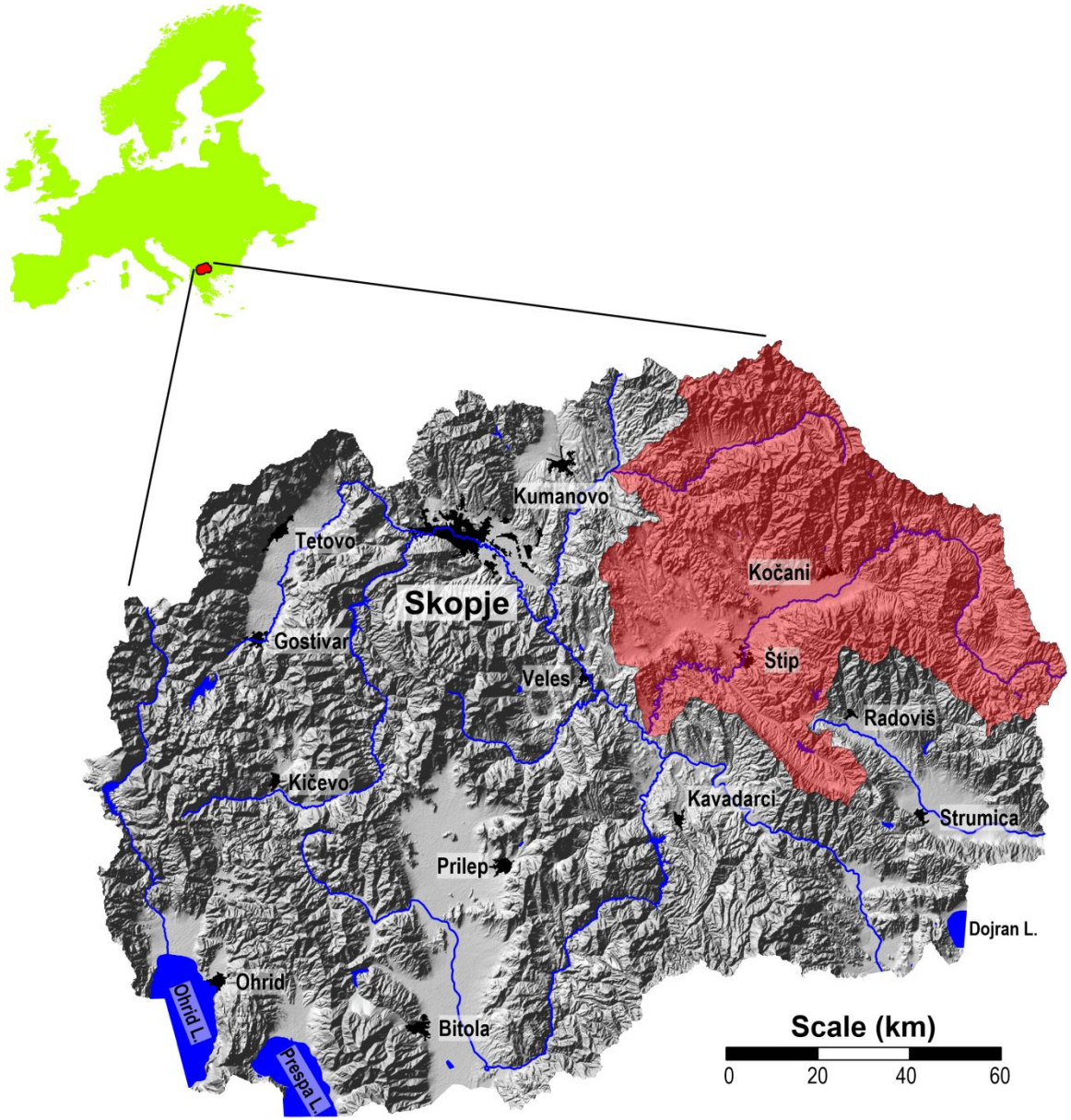
High concentrations are related to the smelter environs (atmospheric transport) and alluvial plain (river transport)

GeoZS

ALIJAGIĆ, J. 2013. Application of multivariate statistical methods and artificial neural network for separation bedrock background and influence of mining and metallurgy activities on distribution of chemical elements in the Stavnja valley (B&H), dissertation.
http://www.ung.si/~library/doktorati/okolje/33Alijagic.pdf

Case study NE Macedonia

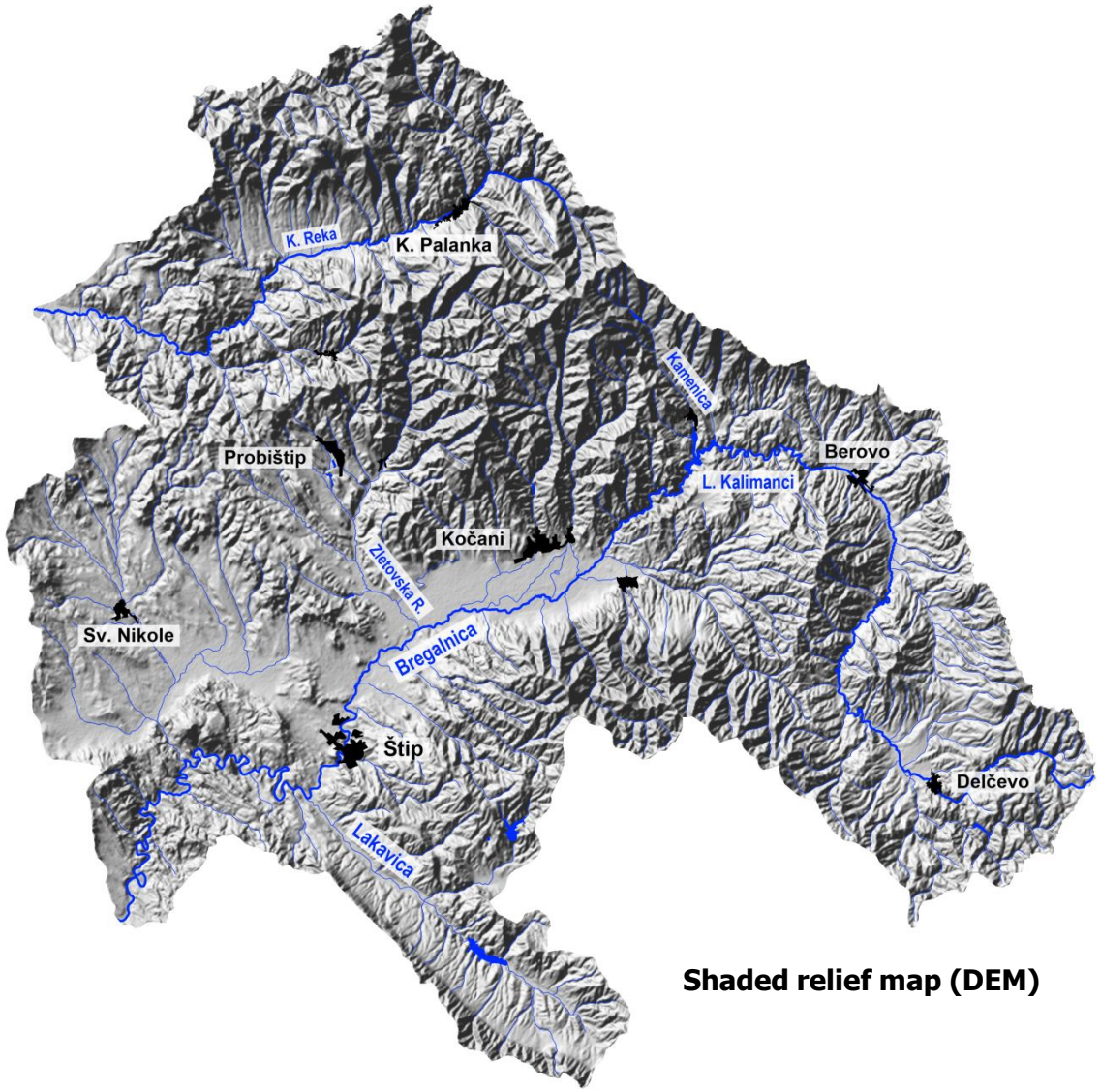
Large area (c. 5400 km^2)



Case study NE Macedonia

Large area (c. 5400 km^2)

Complex morphological area



Shaded relief map (DEM)

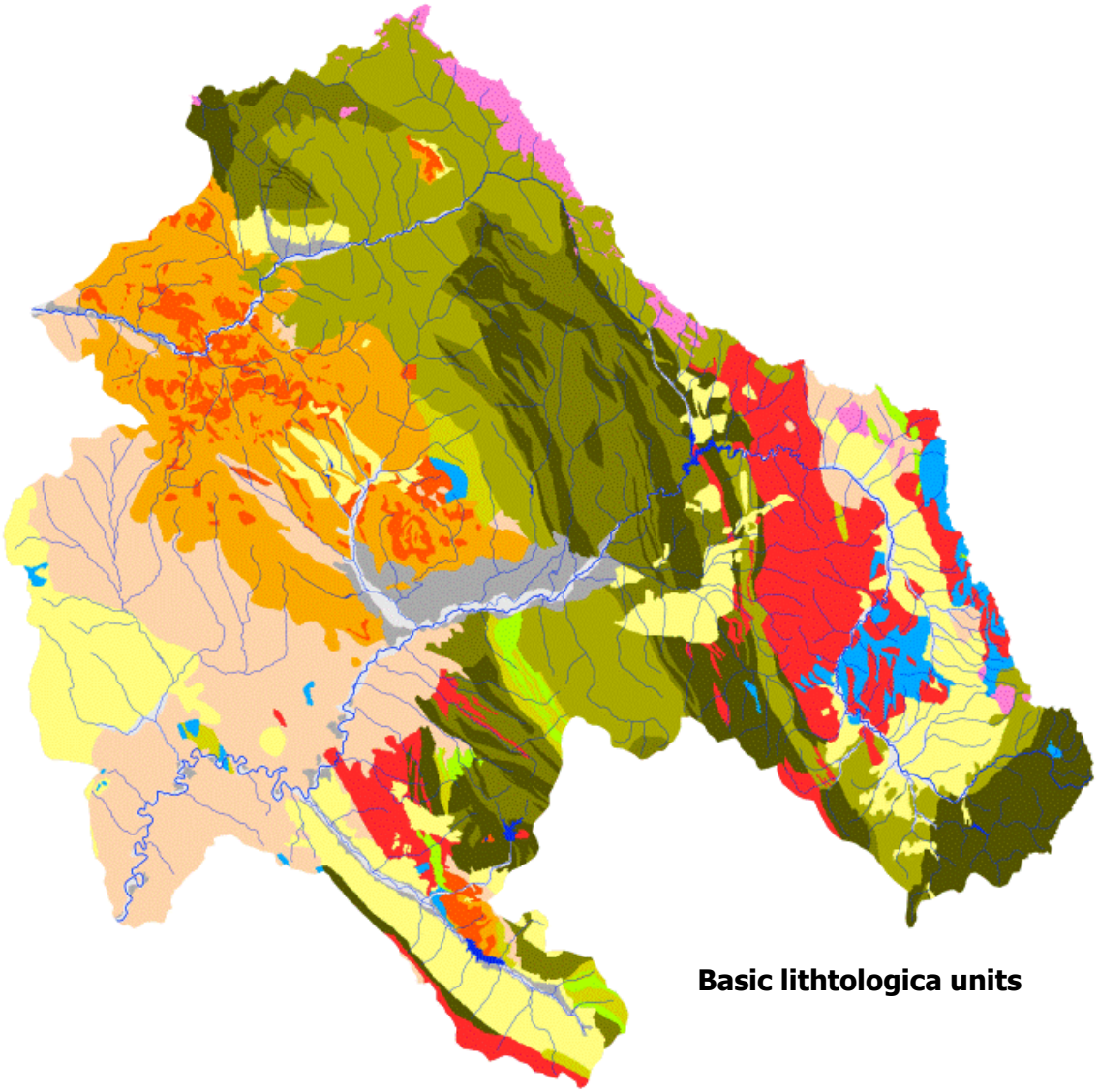
GeoZS

Case study NE Macedonia

Large area (c. 5400 km^2)

Complex morphological area

Geological background -
many different lithological units (Pt-Q)



Basic lithtologica units

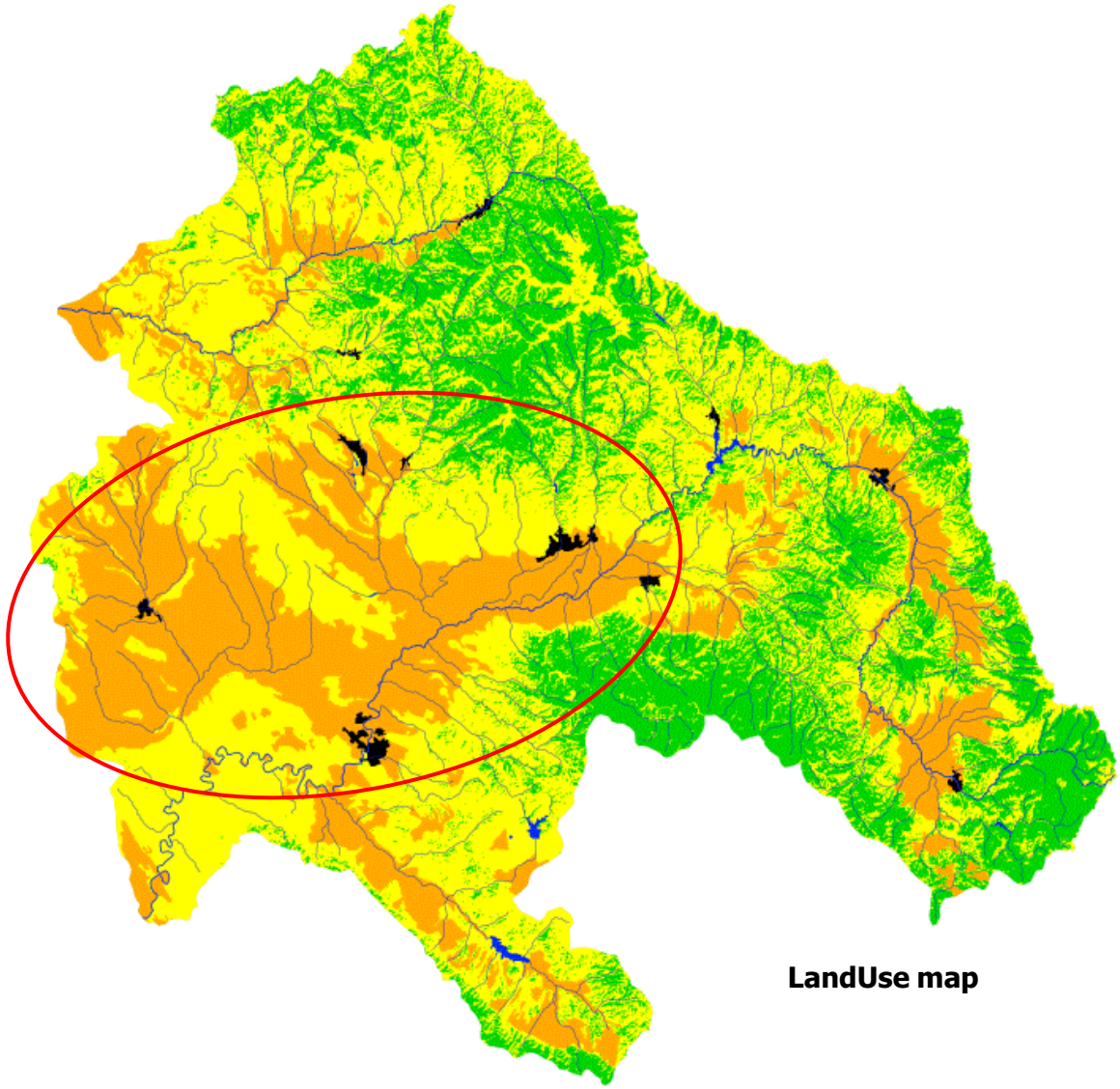
Case study NE Macedonia

Large area (c. 5400 km^2)

Complex morphological area

Geological background
many different lithological units (Pt-Q)

Area of important agricultural activities



LandUse map

GeoZS

Case study NE Macedonia

Large area (c. 5400 km^2)

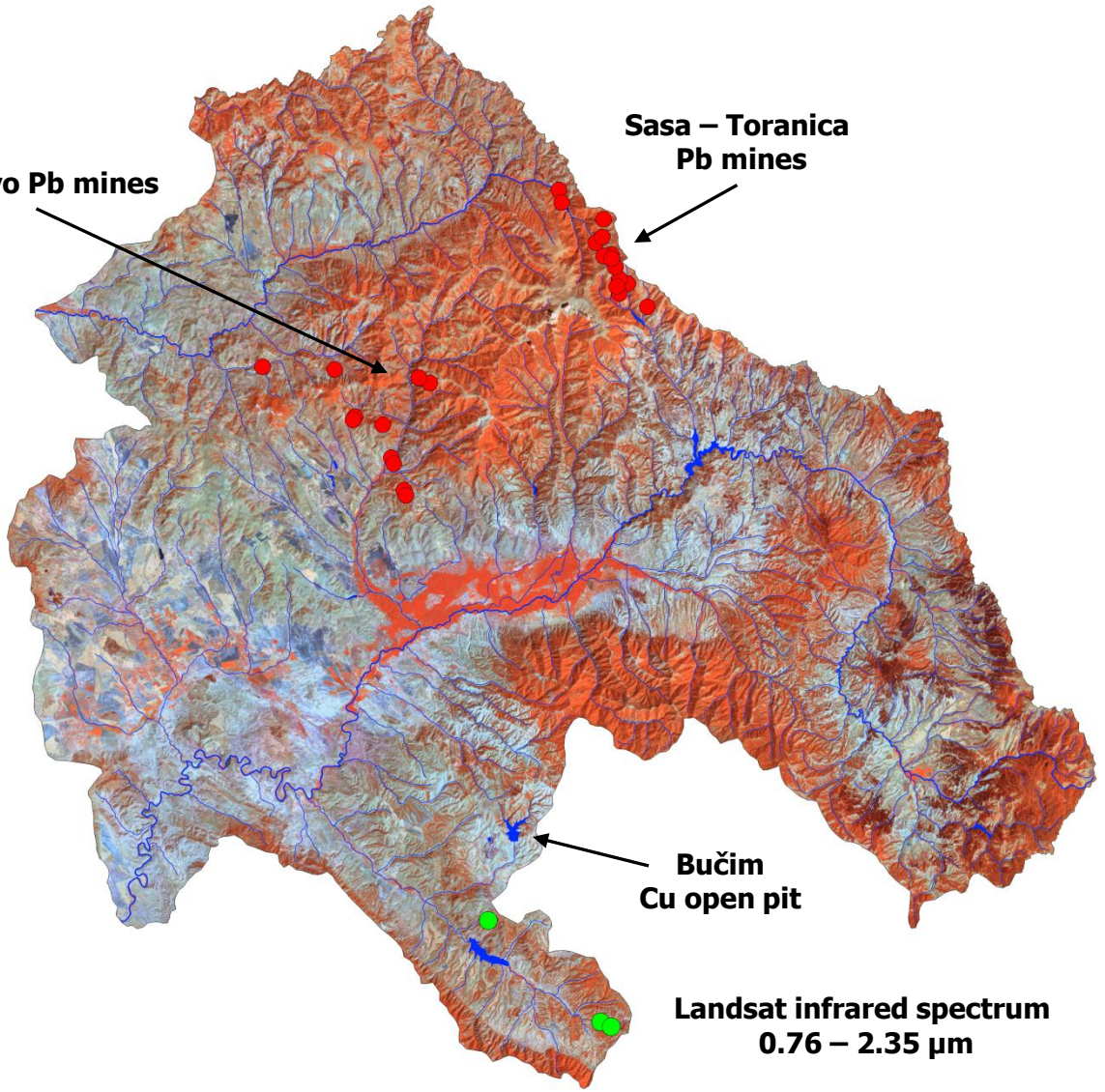
Complex morphological area

Geological background -
many different lithological units (Pt-Q)

Area of important agricultural activities

Intensive mining activities
Pb mines Sasa, Toranica, Zletovo
Cu open pit Bučim
Pb-Zn-Cu mineralisation outcrops

Presence of natural enrichment
and complex anthopogenic impact
in various directions



Landsat infrared spectrum
0.76 – 2.35 µm

GeoZS

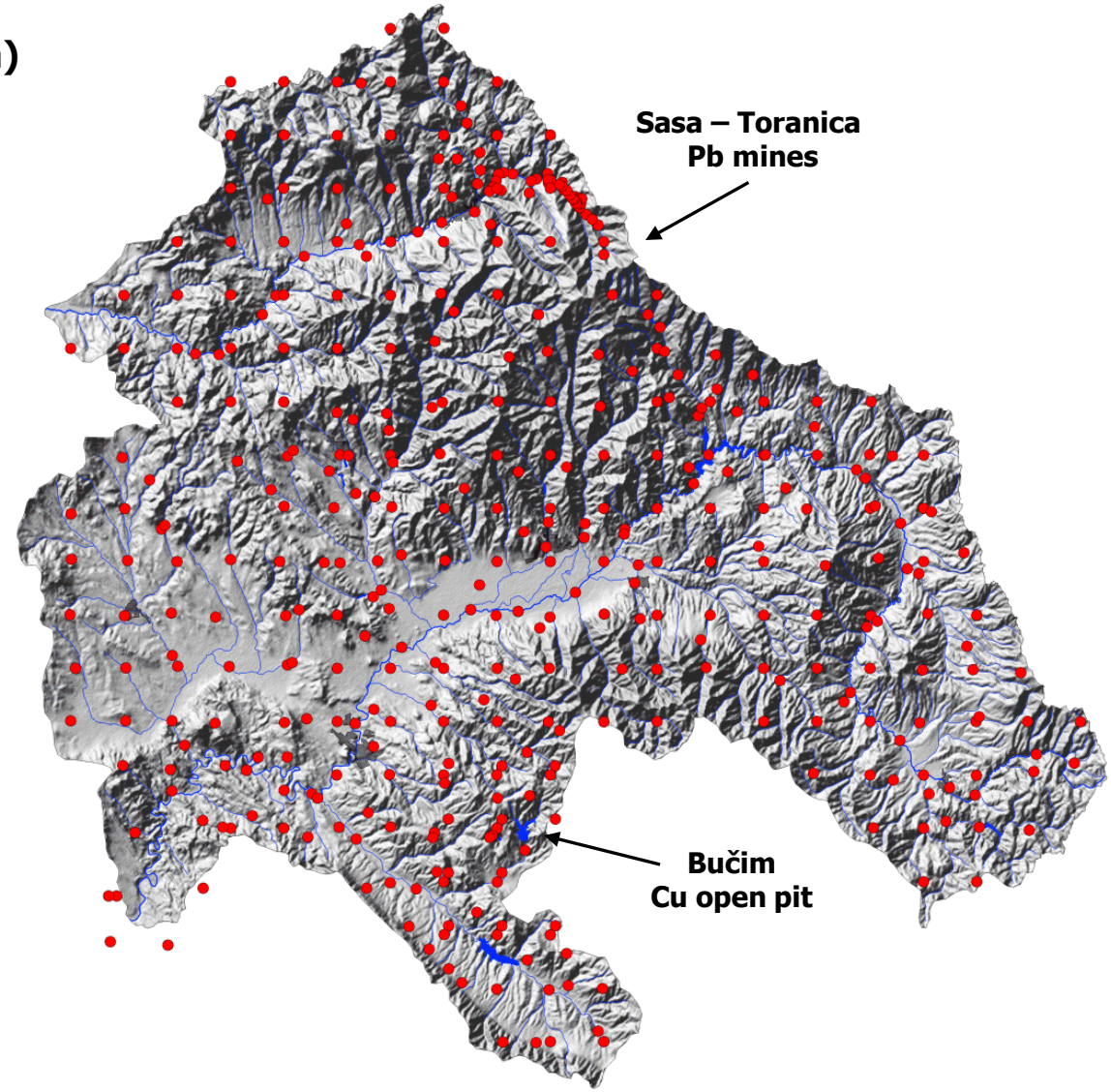
ANN-MP input data (topsoil)

409 sampling points (learning data)

218 – regular grid 5 x 5 km (geochemical map of Macedonia)

126 – other investigation (mainly around mines)

65 – alluvial soil investigation



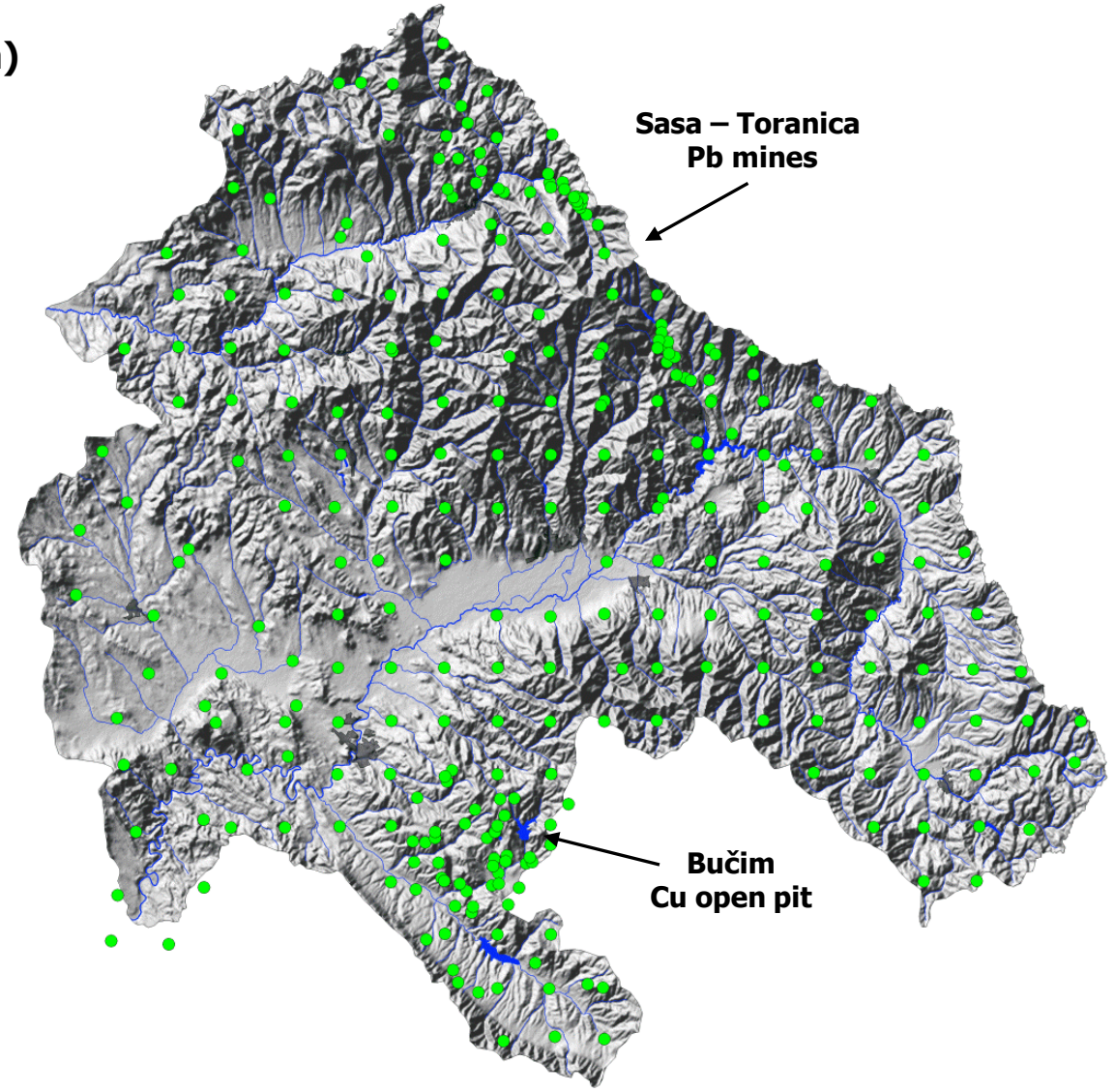
GeoZS

ANN-MP input data (Moss)

286 sampling points (learning data)

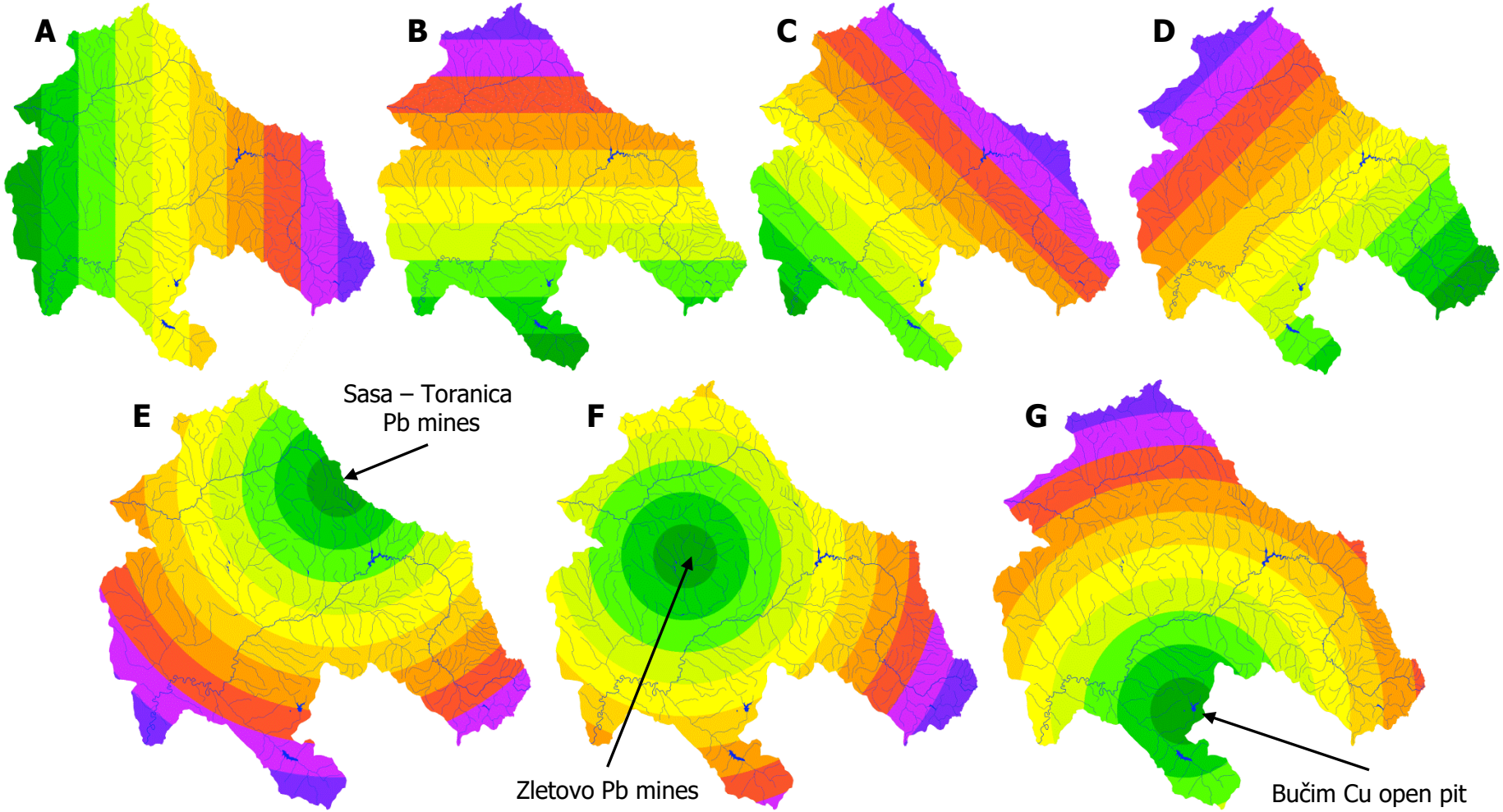
195 - regular grid 5 x 5 km
(geochemical map of Macedonia)

91 – other investigation
(mainly around mines)



GeoZS

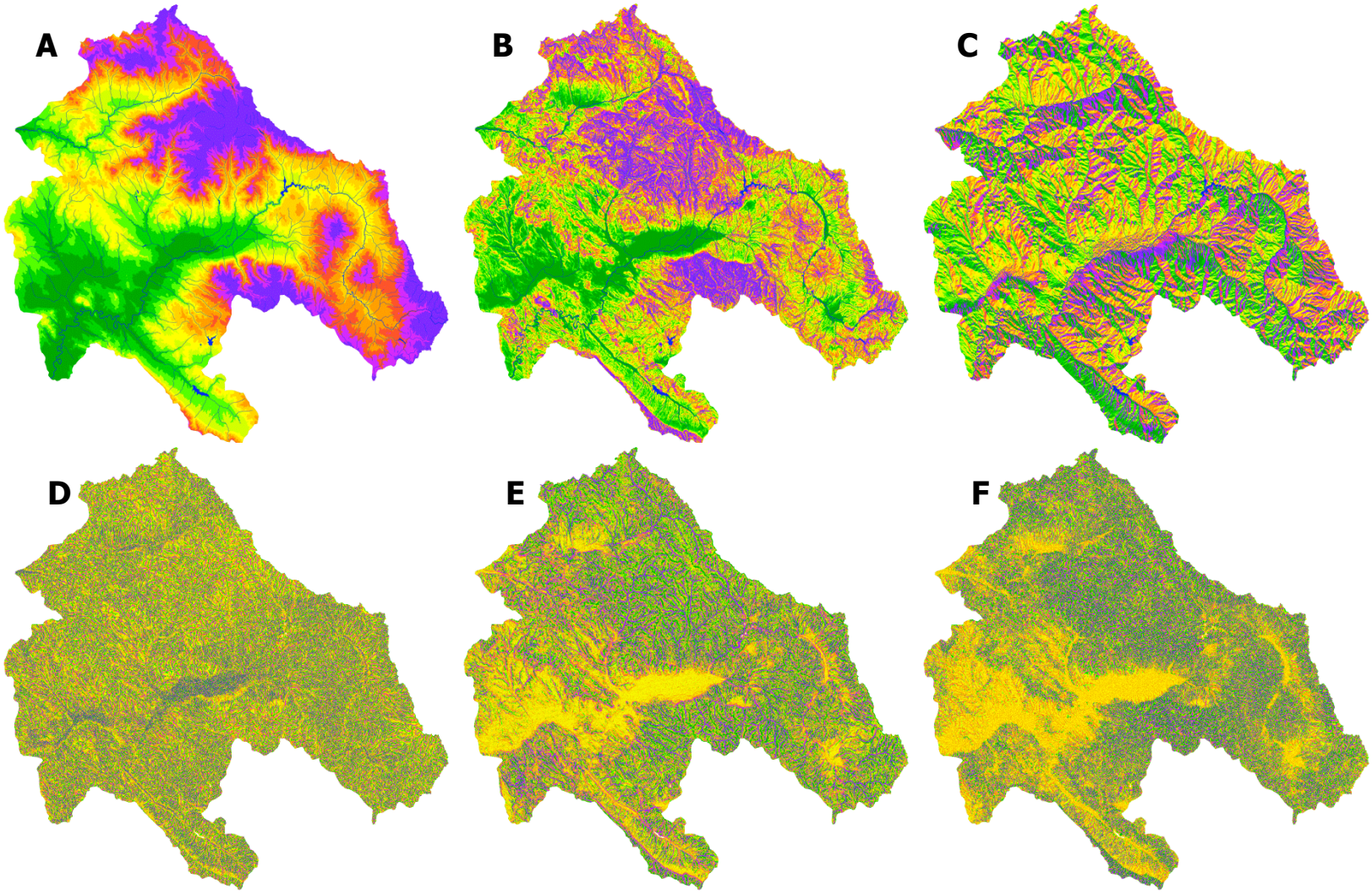
ANN-MP input data (distances)



A – Distance in W-E direction
B – Distance in S-N direction
C – Distance in SW-NE direction
D – Distance in SE-NW direction

D – Distance from the area of Sasa – Toranica Pb mines
E – Distance from the area of Kratovo Pb mines
F – Distance distance from the Bučim Cu open pit

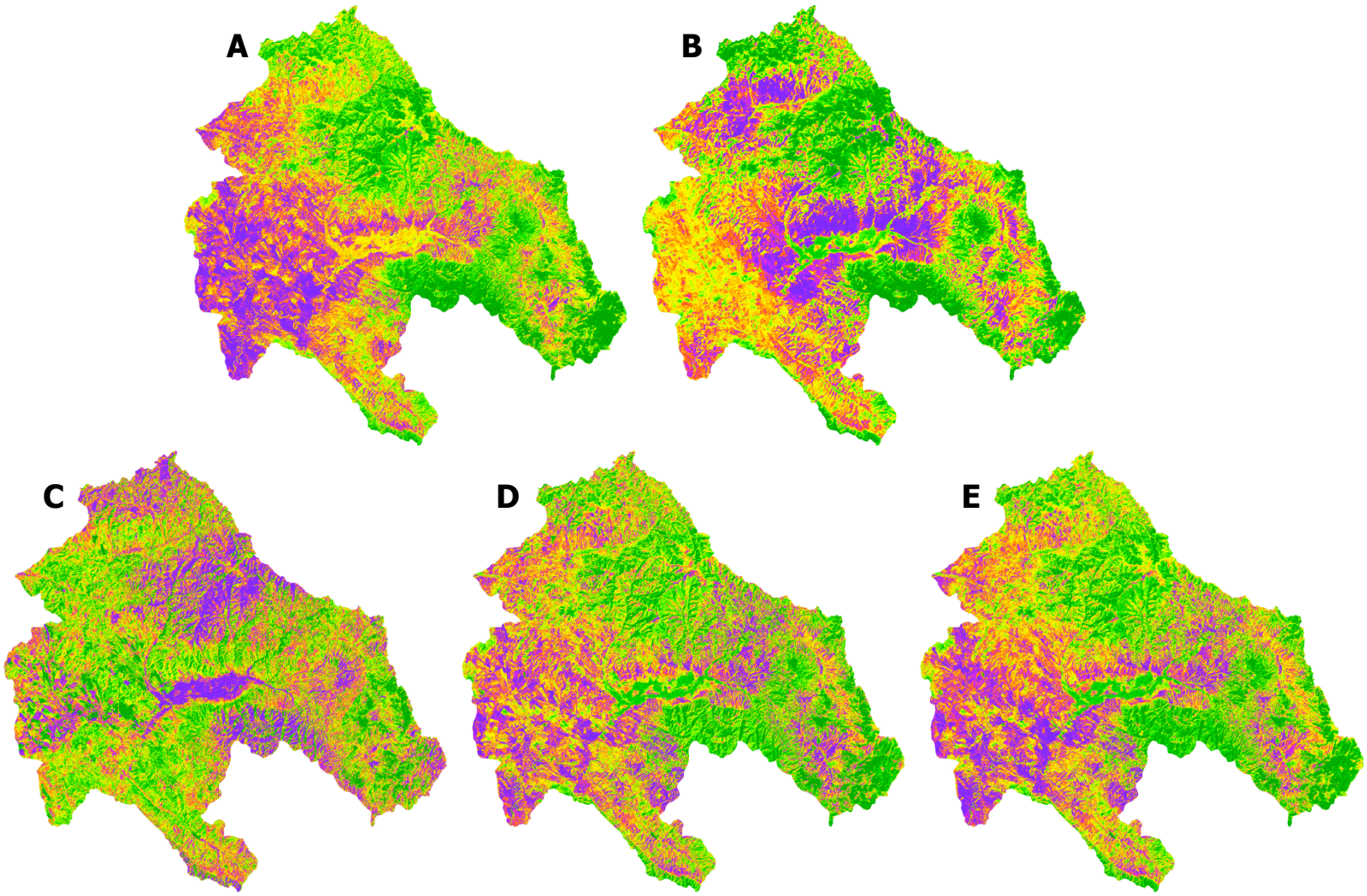
ANN-MP input data (DEM)



A – Altitude above the see level (absolute)
B – Terrain slope
C – Aspect (insolation)

D – Plan terrain curvature
E – Profile terrain curvature
F – Tangent terrain curvature

ANN-MP input data (LandSat)



A – Join visible spectrum, 0.45 – 0.69 μm (B10-B20-B30)
B – Thermal radiation spectrum, 10.4 - 12.5 μm (B60)

C – Infrared spectrum, 0.76 – 0.90 μm (B40)
D – Infrared spectrum, 1.55–1.75 μm (B50)
E – Infrared spectrum, 2.08–2.35 μm (B70)

ANN calculation summary

Categorical input data

Geological map
Landuse map

Numerical input data

Distances
Digital elevation model and its derivates
LandSat satelit images (B10-B70)

Next step (improvement of model)

CORINE Landuse map (categorical data)
Pedological map (categorical data)
Mean annual precipitation map (numeric data)
Mean annual temperature map (numeric data)
Mean annual wind magnitude (speed) map (numeric data)
Mean annual wind vector (direction) map (numeric data)

Learning and recall

Learning data: 409 sampling points (Topsoil)
Learning data: 286 sampling points (Moss)
Recall data: 540 497 locations (Grid 100 X 100 m)

ANN Training

Multilayer perceptron - Hidden units - 120

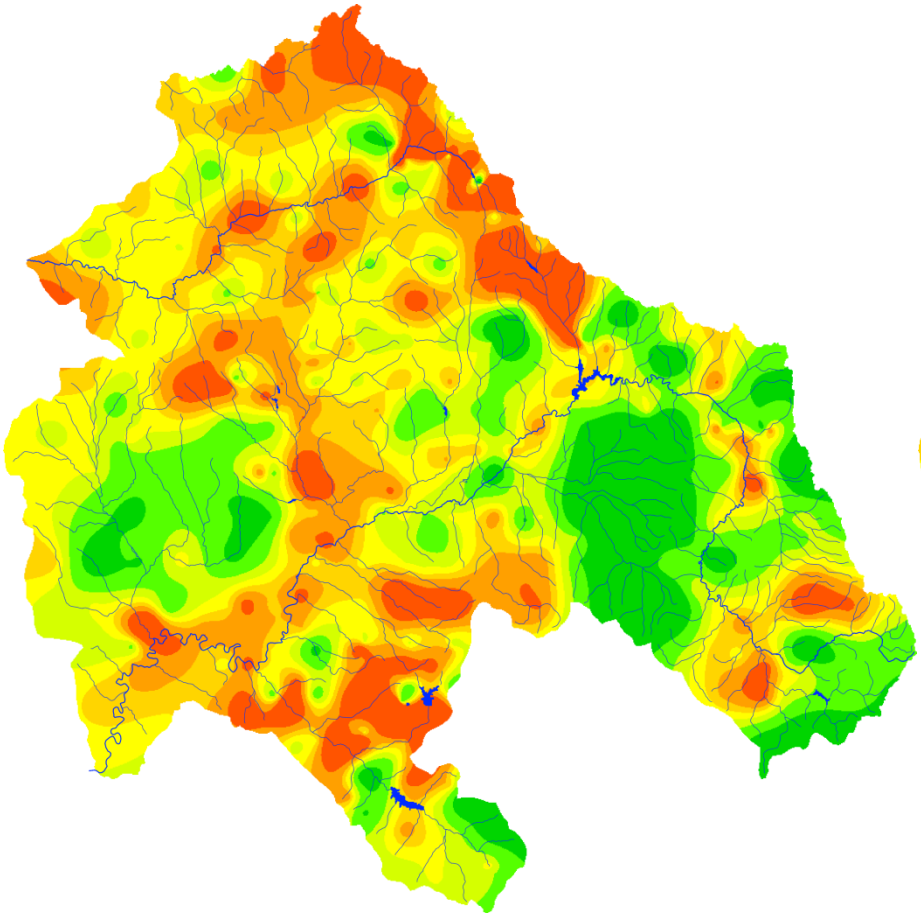
Train networks - 25 (the final model representing an average of 5 most logical solution)

Each training model contain: Training perfection, Test perfection, Validation perfection, All perfection, Training error, Test error, Validation error, Training algorithm, Hidden activation, and Output activation.

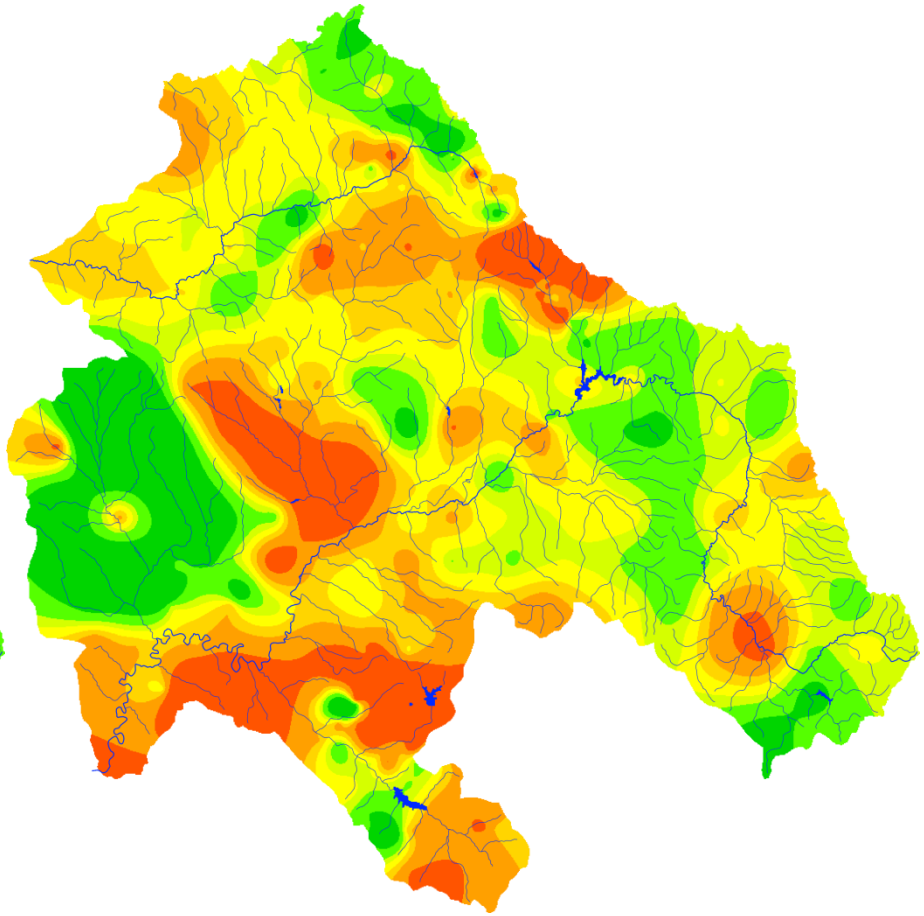
GeoZS

Distribution of copper (Universal Kriging)

Topsoil



Moss

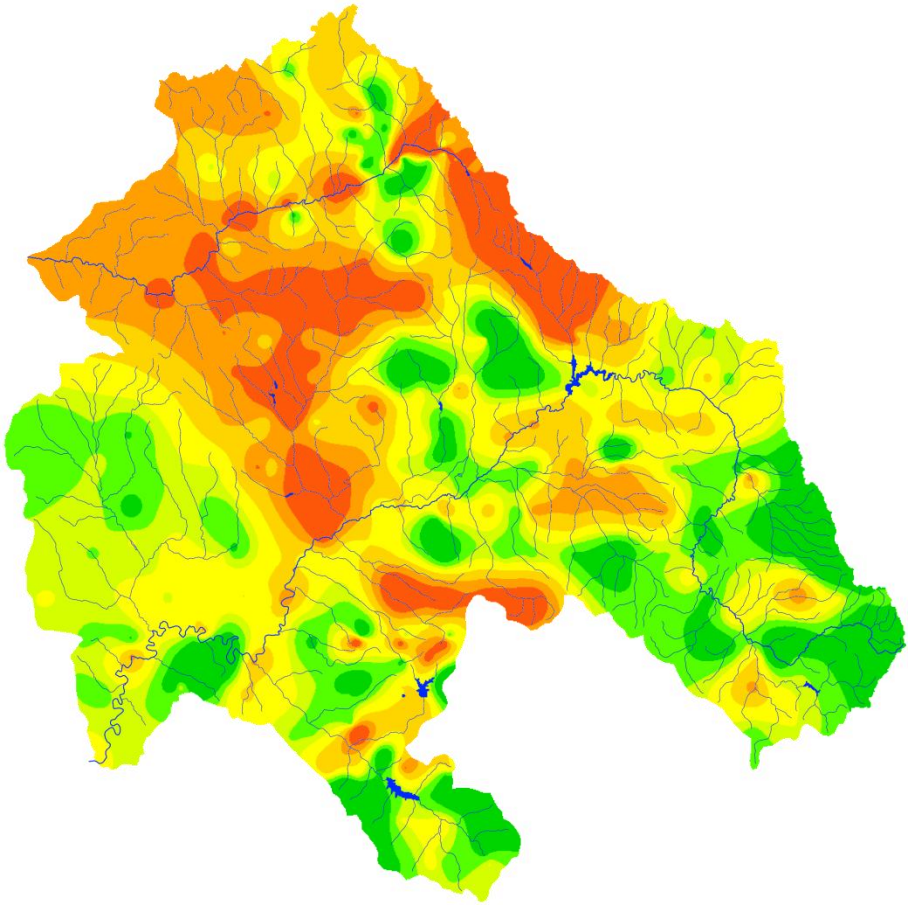


The results are difficult to interpret in both sampling materials due to typical mistakes of linear interpolations – Bull's-eye effect.

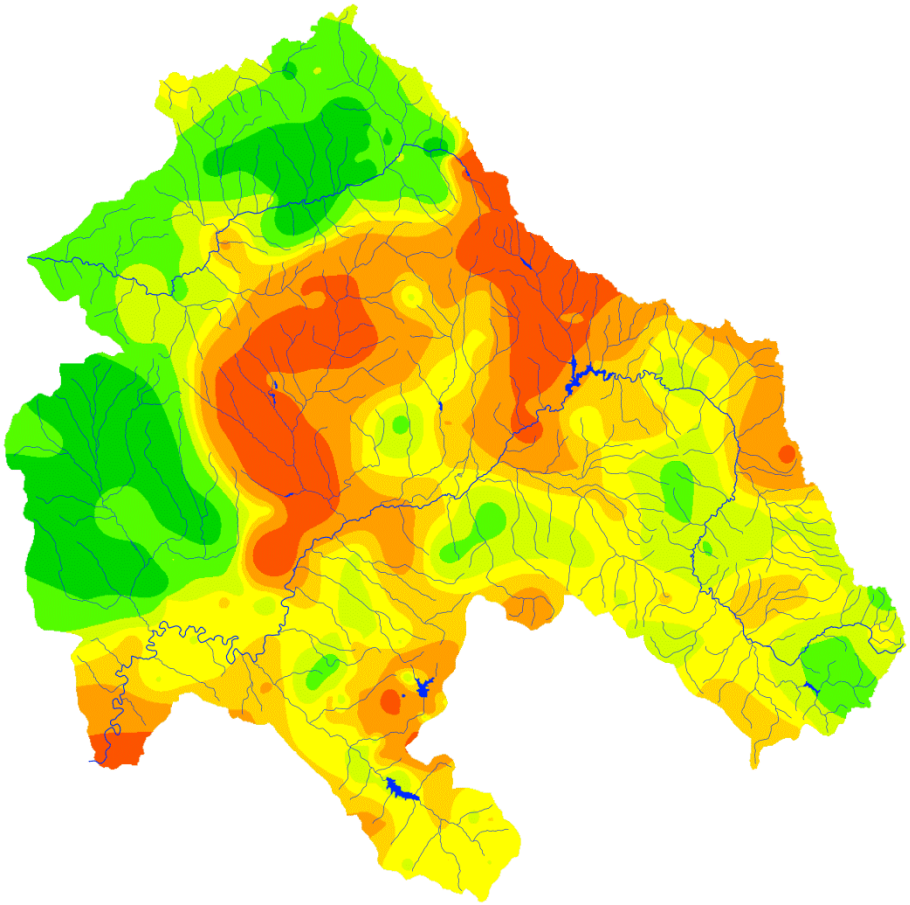
GeoZS

Distribution of lead (Universal Kriging)

Topsoil



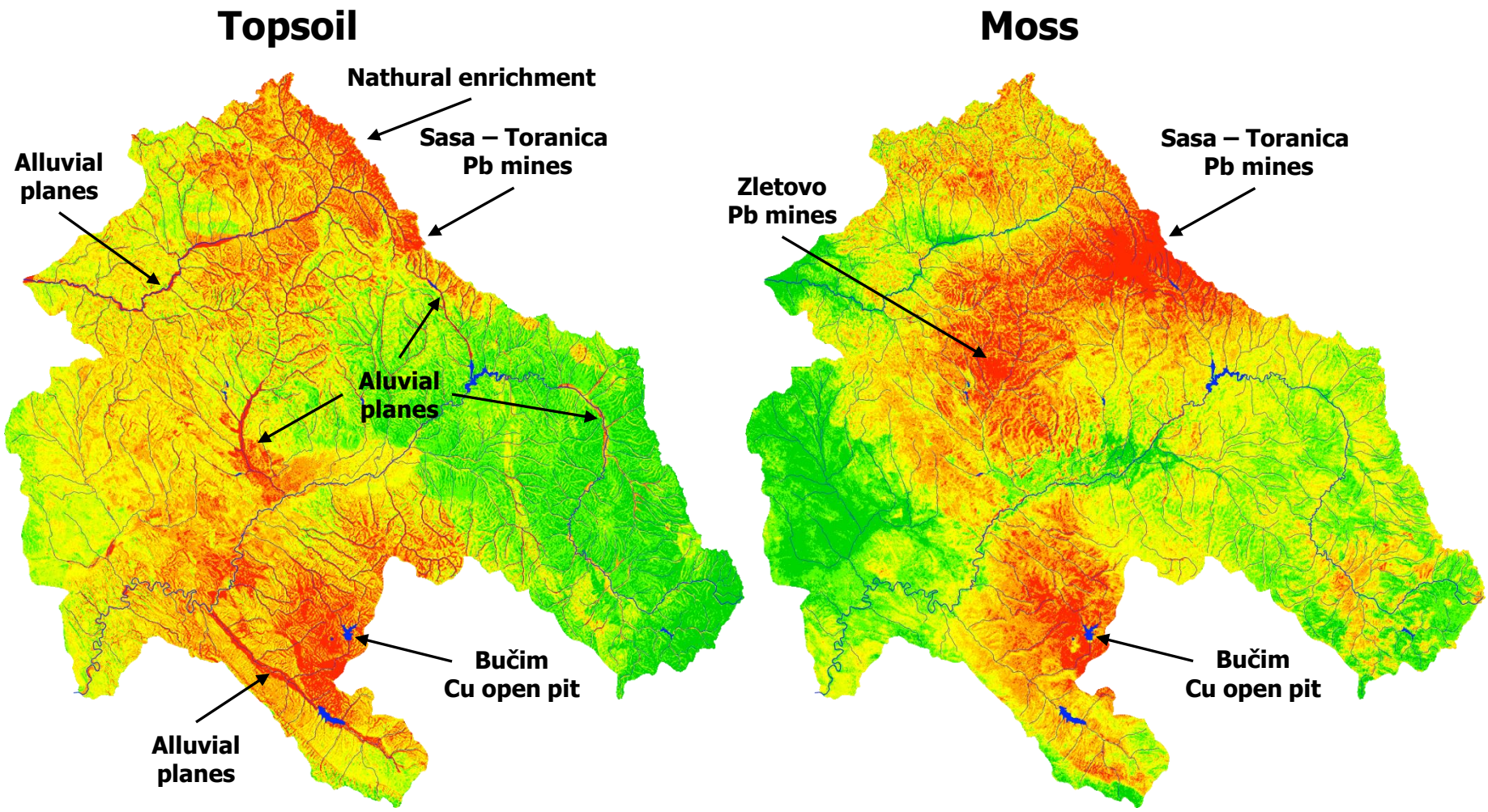
Moss



Spatial distributions of Lead are more logical than the previous one. The high concentrations are connected to the natural enrichment on particular lithological units or Pb mining areas .

GeoZS

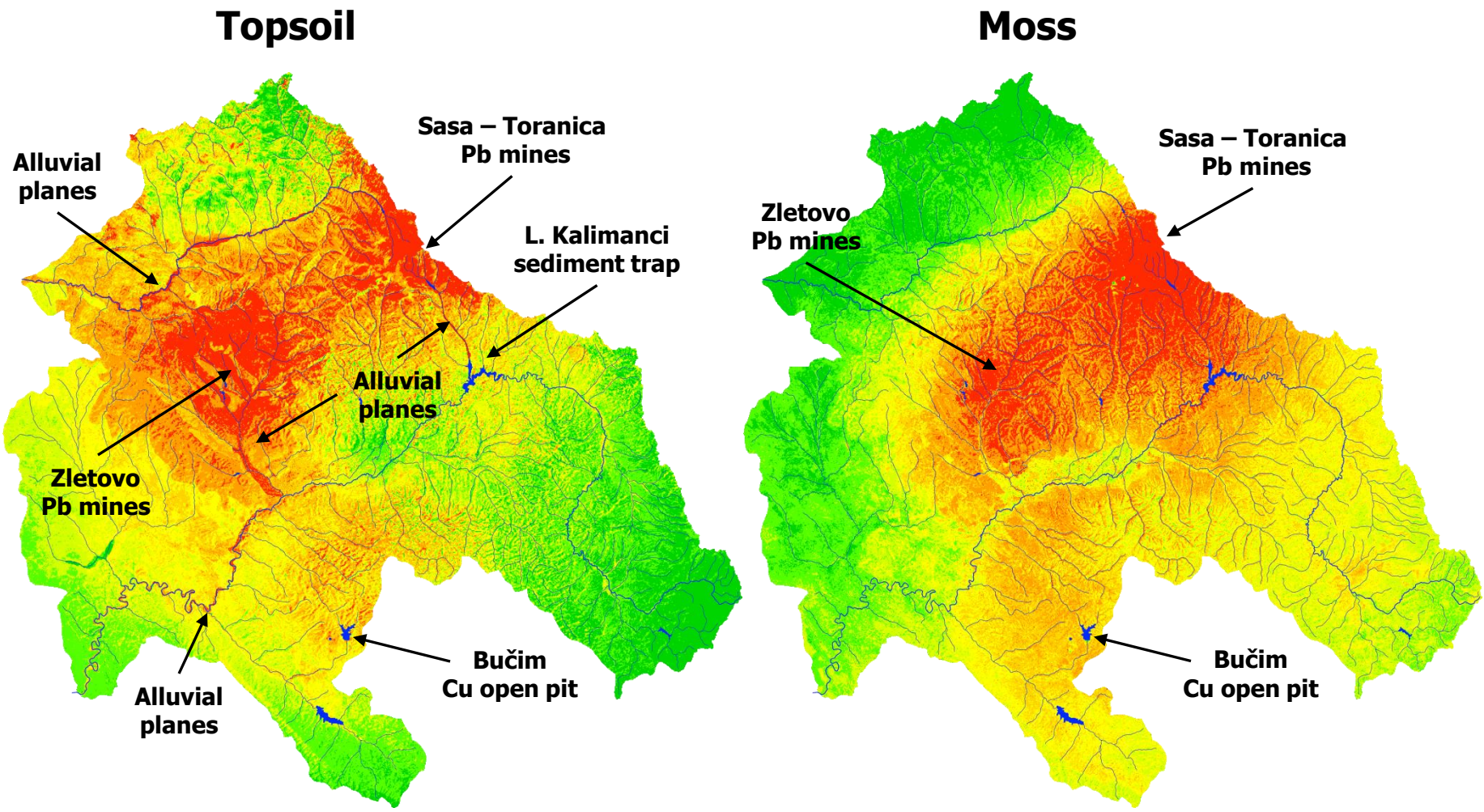
Distribution of copper (ANN-MLP)



Model obtained by ANN is significant and logical. Cu enrichment is connected to the Cu open pit and some lithological units and along the rivers (alluvial planes) – what indicate presence of river transport.

Areal distribution is more significant for the moss. Atmospheric enrichment is connected to the mining areas. The high concentrations are not connected to the lithological units.

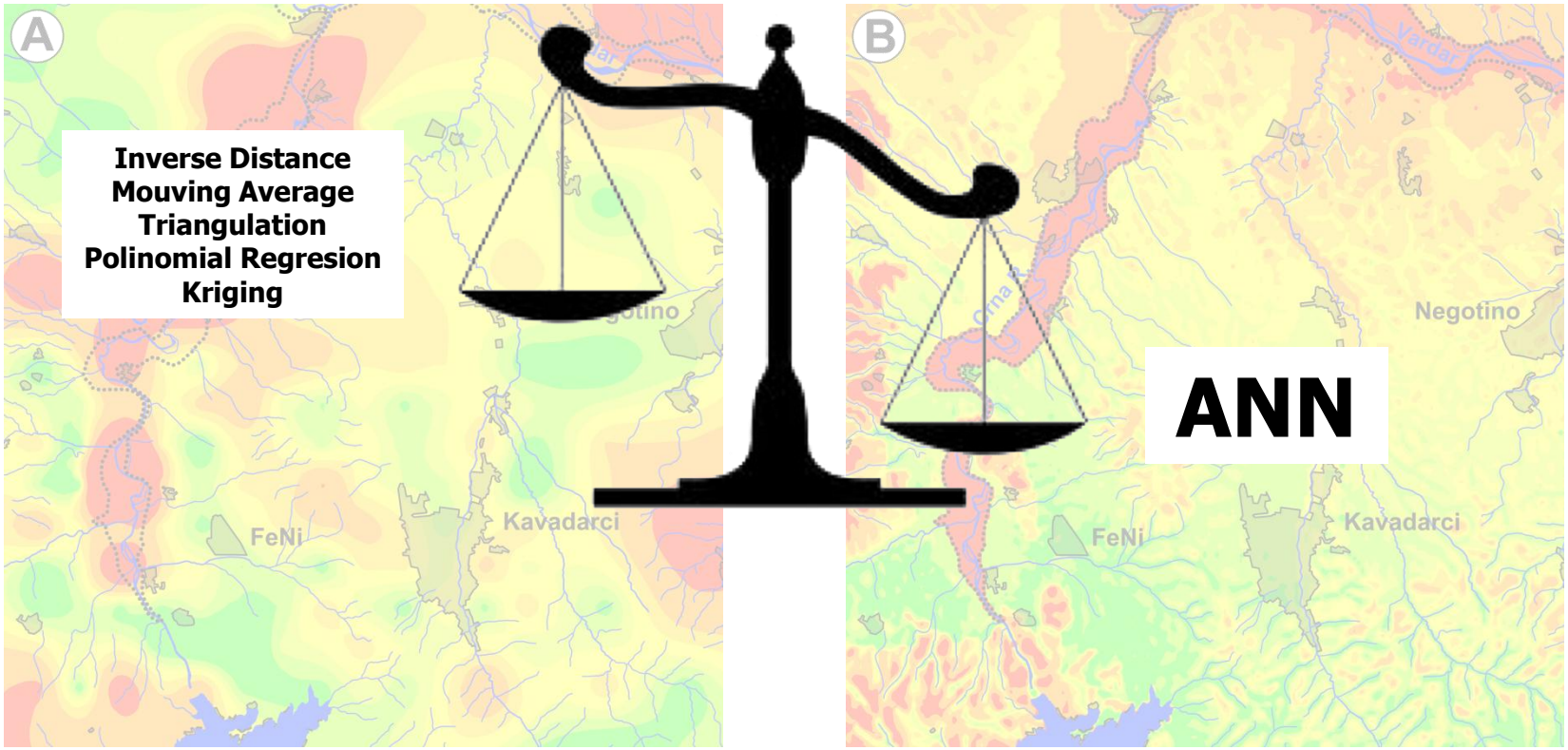
Distribution of lead (ANN-MLP)



Pb enrichment is connected only to Pb mining areas, some lithological units and along the rivers (alluvial planes). At the middle flow of the river Bregalnica the ANN didn't isolated the high concentrations. This means that the sediments are trapped in the lake Kalimanci and polluted sediments accumulate in the lake.

Atmospheric enrichment is connected to the mining areas. The high concentrations are not connected to the lithological units.

Conclusion



Everything is a matter of choice !


Thank you !