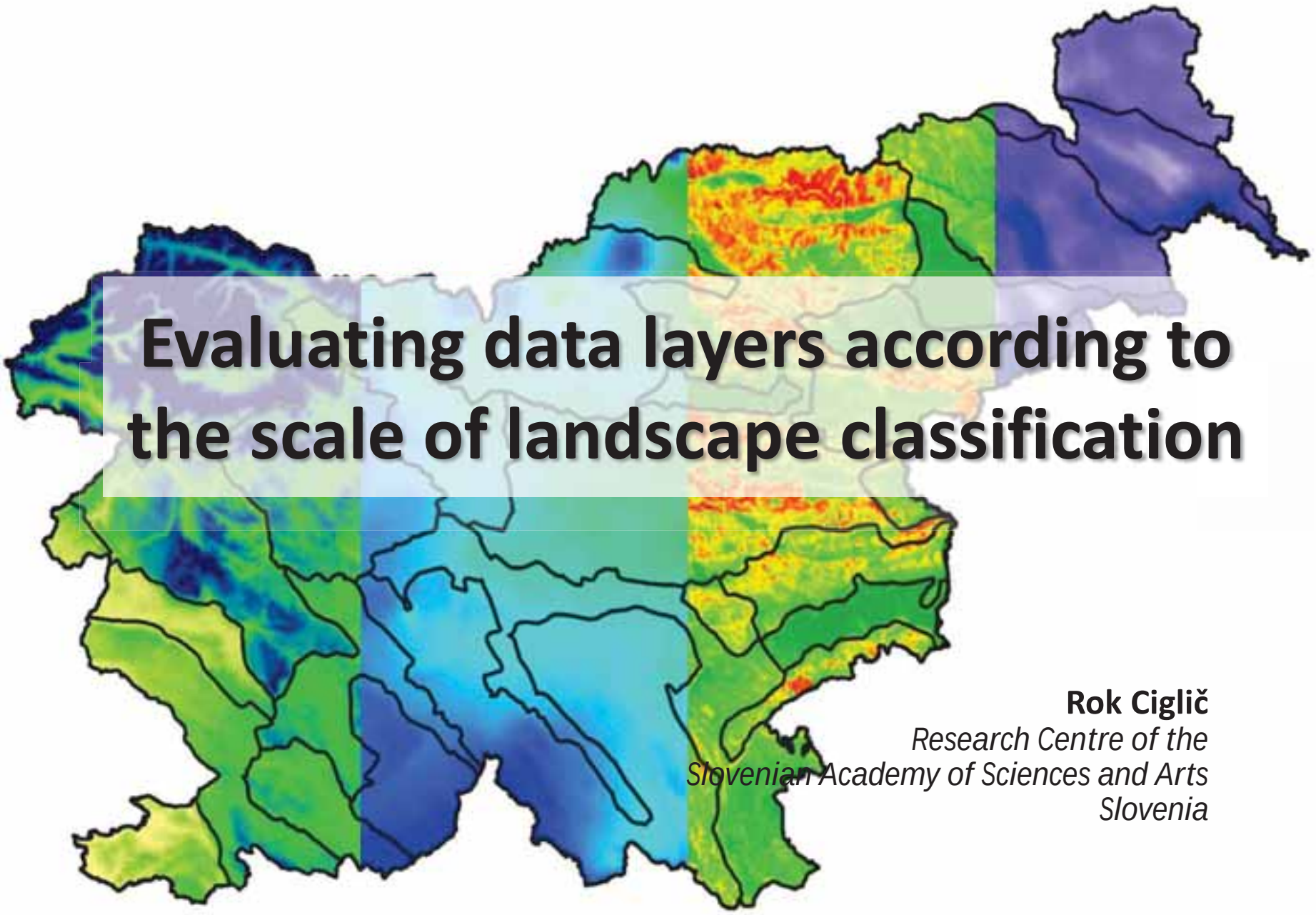


Evaluating data layers according to the scale of landscape classification

Rok Ciglič
Research Centre of the Slovenian Academy of Sciences and Arts
Slovenia

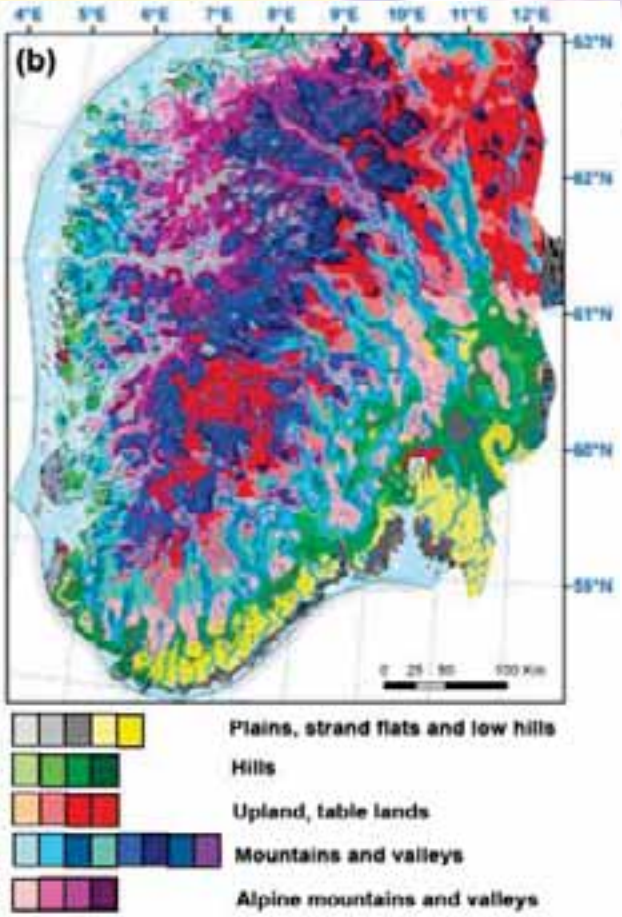
ESRI European User Conference, Oslo, 16th of October 2012

Brief overview of presentation

- Introduction to a landscape classification
- Theory and reality of landscape classification
- GIS is our friend
- Method for evaluation of data layers
- Application of the method (case study Slovenia)
- ArcGIS, Python scripting

Landscape classification examples

five major geographical regions - ***landsdeler***
(Wikipedia 2012)



25 topography types
(Etzelmüller, Romstad and Fjellanger 2007)

Marjan Garbajs

Jurij Senegačnik

Marjan Garbajs

Petra Gostinčar

50 km

(C) Google 2012

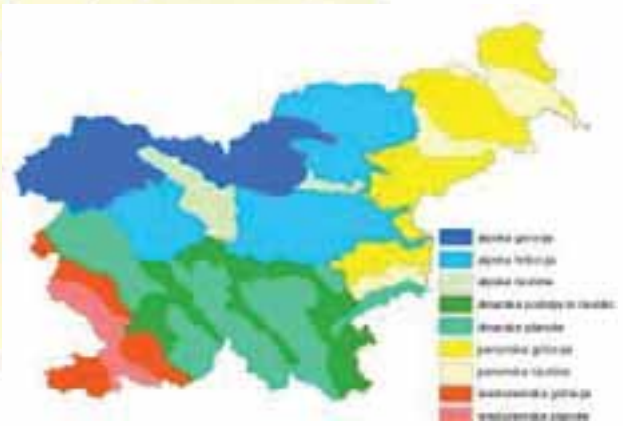
Landscape classification

... has a long tradition in Slovenian geography

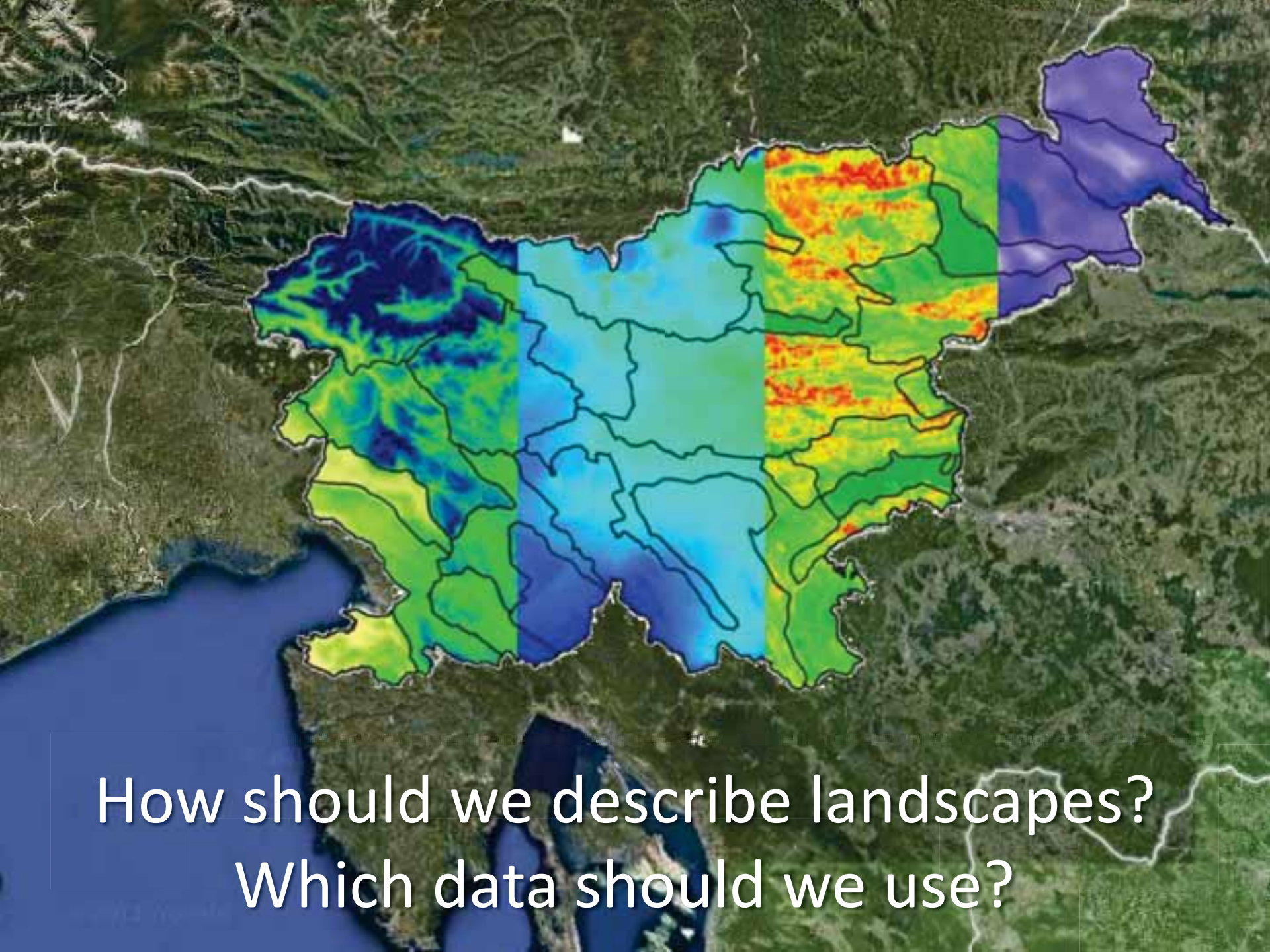
Melik, 1935

Ilešič, 1972

Plut, 1999



Perko, 1998


How should we describe landscapes?
Which data should we use?

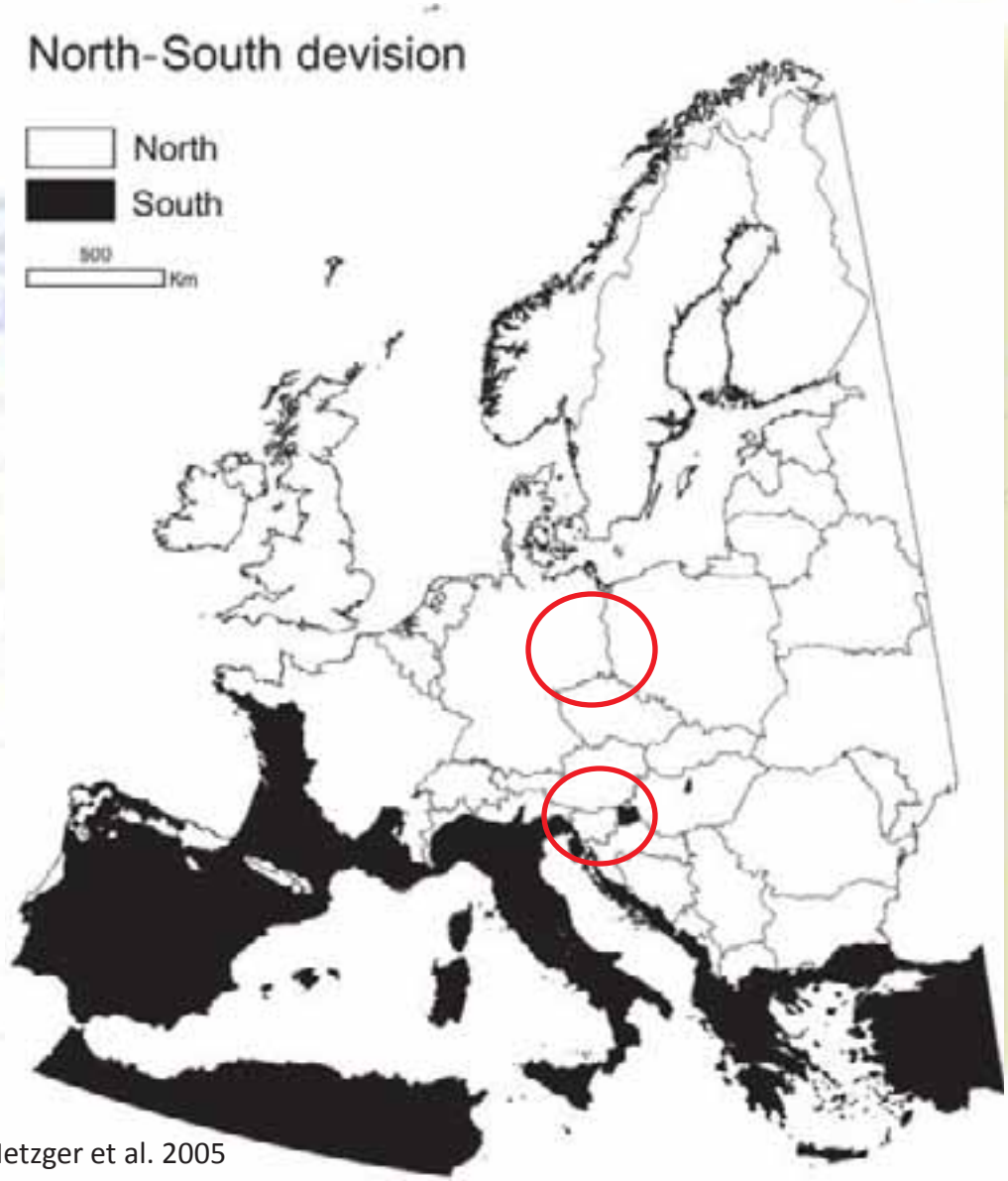
Theory ...

There are a lot of general recommendations for landscape classification.

One example (after Bailey 1996):

Spatial unit	Size	Scale	More important factors
ecoregion or zone	10^5 km^2	1 : 3.000.000	climate, continental position, altitude
landscape mosaic	10^3 km^2	1 : 250.000–1 : 1.000.000	land-surface form
site	10 km^2	1 : 10.000–1 : 80.000	topoclimate, soil moisture

… and reality



Metzger et al. 2005

Terrestrial Ecoregions of the World has 14 biomes

In Slovenia we can find 3 types of biomes.

Slovenia is on the border of two major European units.

Which data layers are more relevant?

GIS is our friend

GIS enabled landscape classification in a more quantitative way.

Input and Output

For a better result we have to chose a proper method (GIS tool) and relevant data.

In this presentation

... we are going to focus on an evaluation of digital data layers in order to get the most relevant data layers for a landscape classification of Slovenia.

Methods for evaluation of data layers

- can evaluate information values of data layers
- can evaluate correlation among data layers
- can evaluate descriptive statistics (distribution, outliers ...)

- **<u>can evaluate data layers according to the scale</u>**

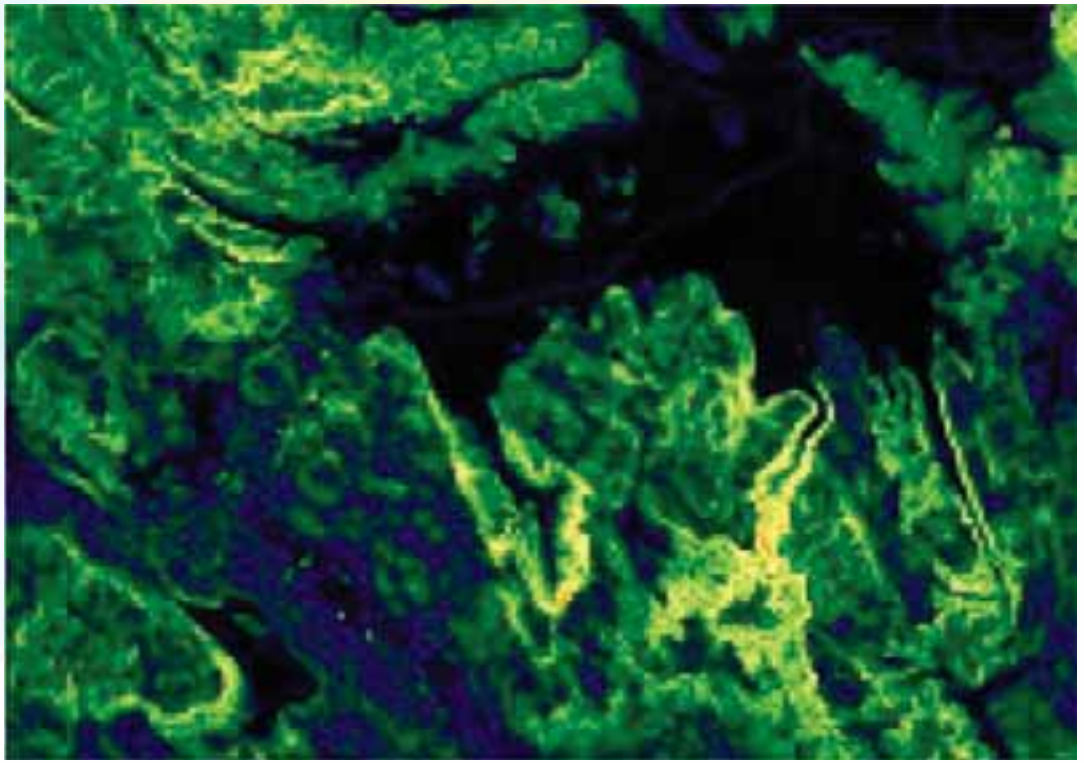
Average moderate coefficient of variation

Average moderate coefficient of variation

- Extension of coefficient of variation

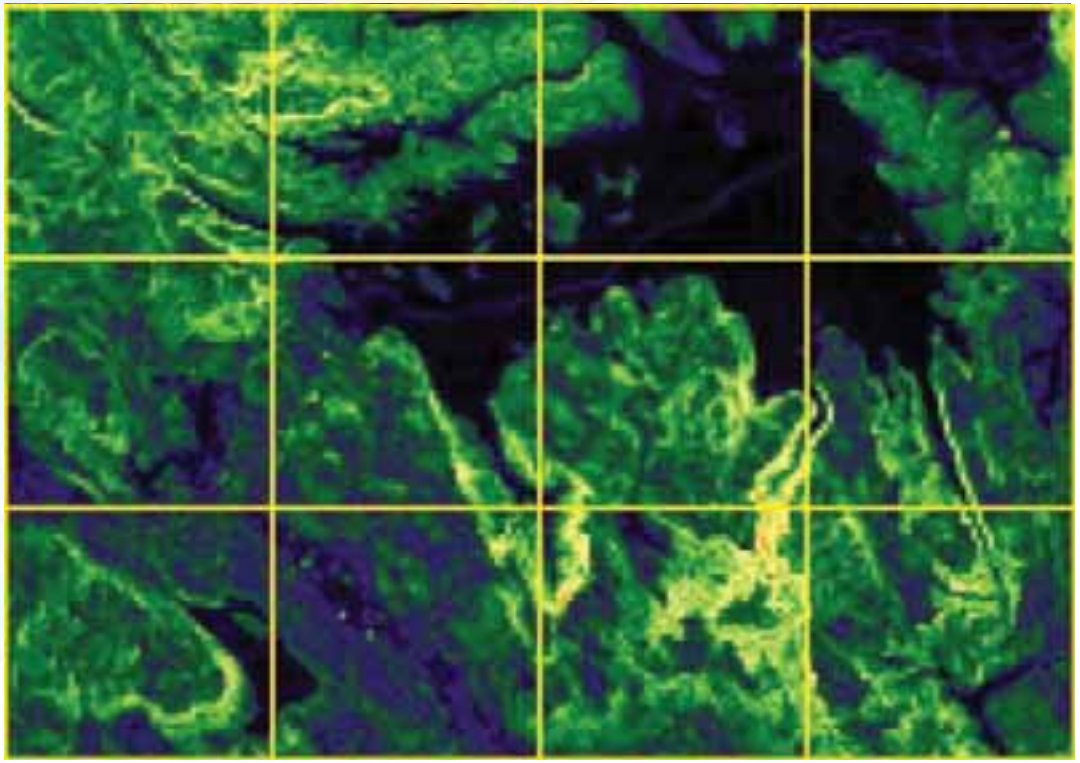
$$CV = \frac{100\sigma}{a_{all}}$$

a_{all} – *average of all cells*
σ – *standard deviation of all cells*
CV – *coefficient of variation for all cells*



Average moderate coefficient of variation

1st step: We divide our study area with (non) uniform units



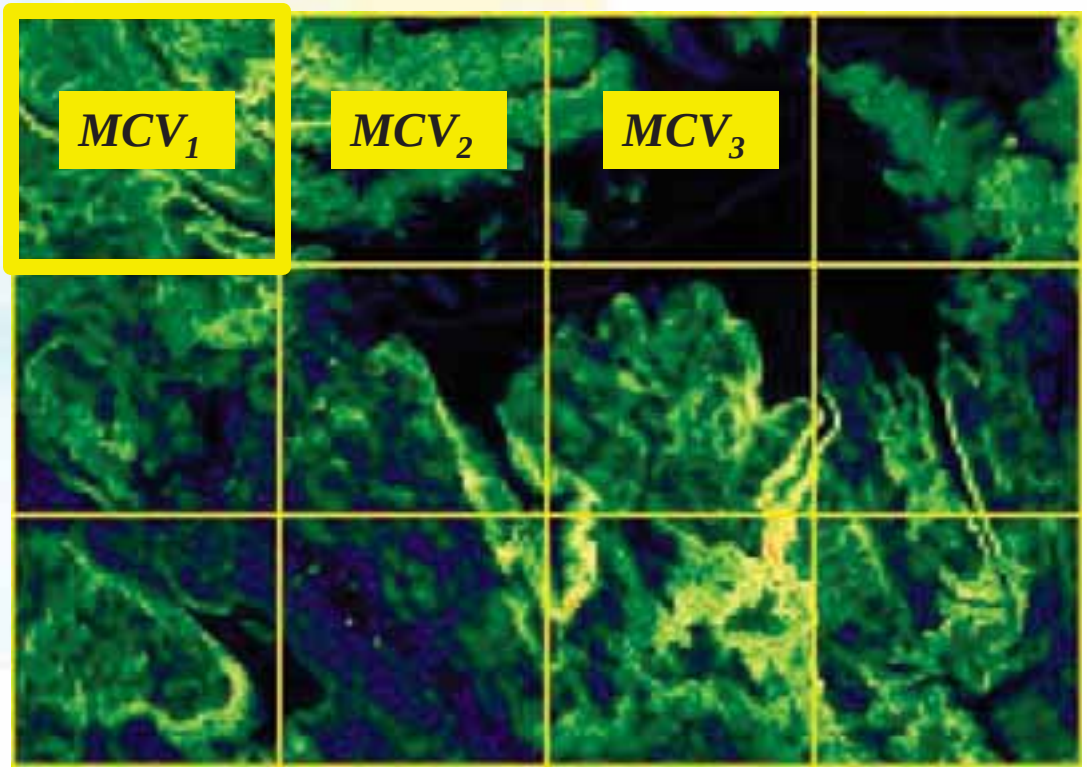
Average moderate coefficient of variation

2nd step: We calculate CV_i for each unit i

(we use average of all cells → *moderate* CV_i or MCV_i)

$$MCV_i = \frac{100\sigma_i}{a_{all}}$$

a_{all} – *average of all cells*
σ_i – *standard deviation of cells in unit i*
MCV_i – *moderate coefficient of variation for the unit i*



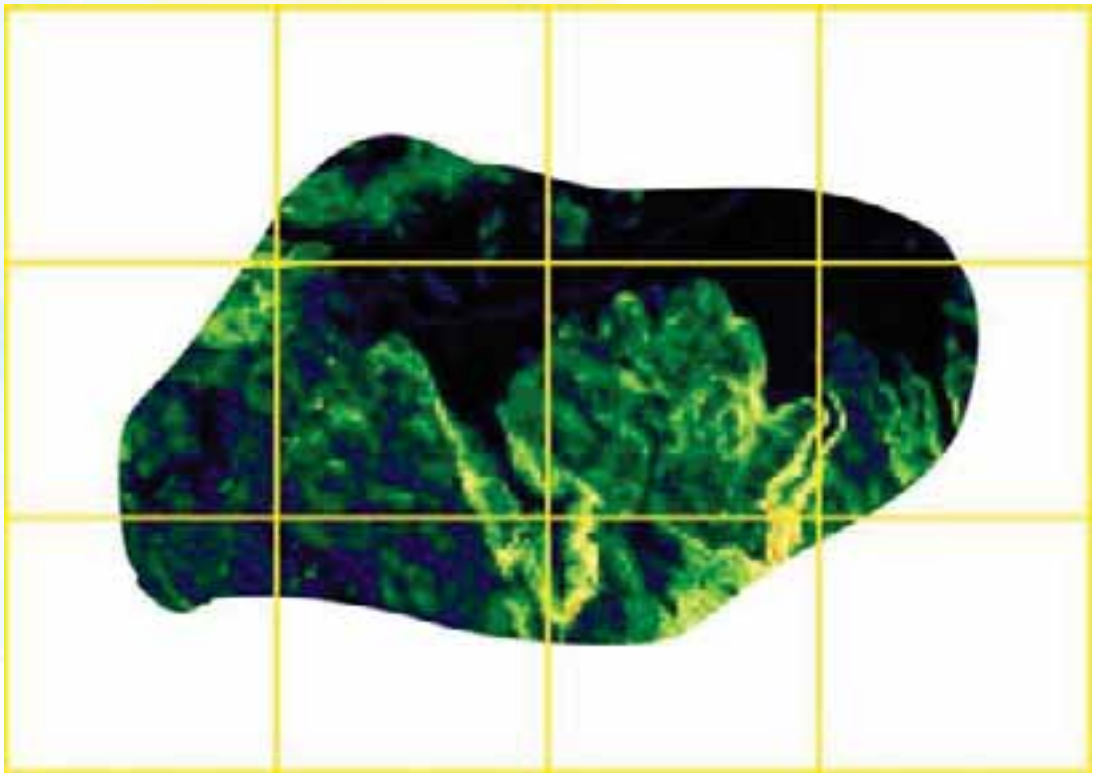
Average moderate coefficient of variation

3rd step: We calculate an average $\boldsymbol{MCV}$ for the whole layer.

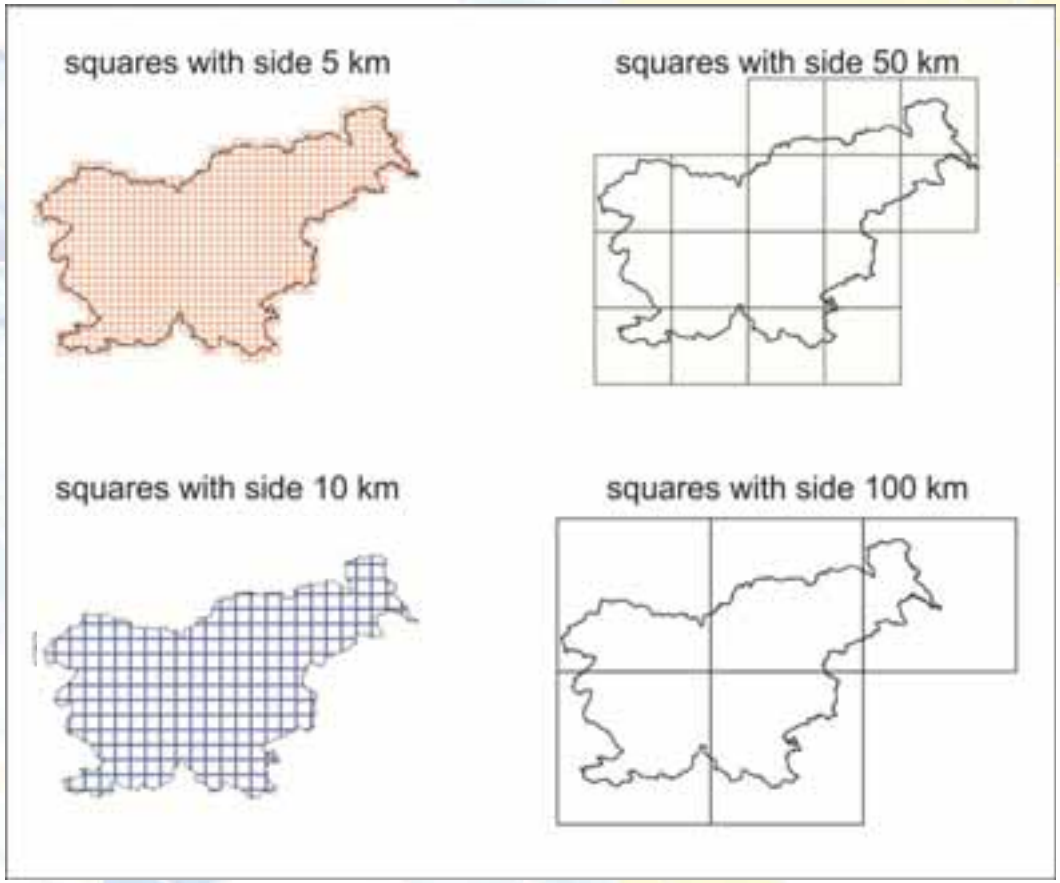
$$AMCV = \frac{\sum MCV_i}{k}$$

$$AMCV = \frac{\sum n_i \times MCV_i}{n_{all}}$$

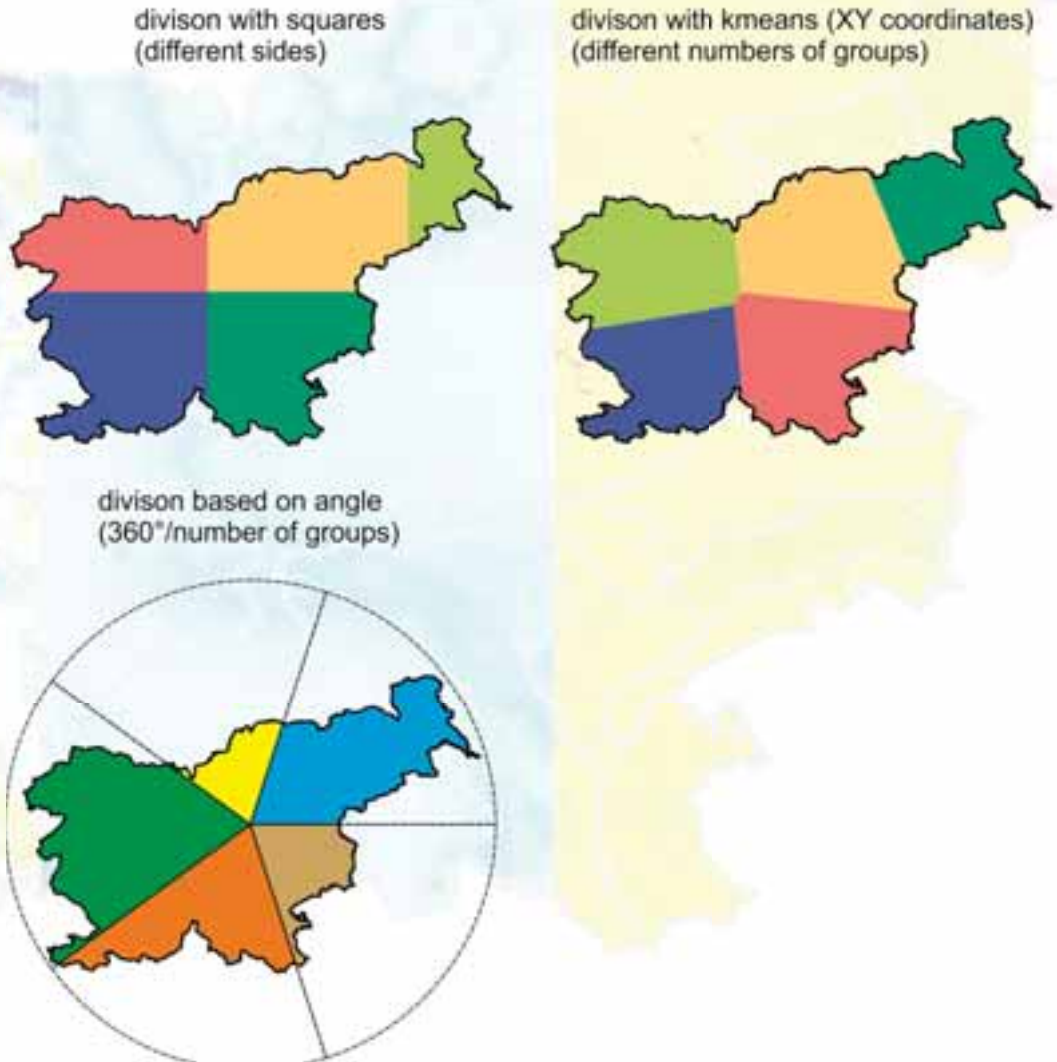
$\boldsymbol{k}$ – number of units i
$\boldsymbol{MCV_i}$ – moderate coefficient of variation for the unit i
$\boldsymbol{AMCV}$ – average moderate coefficient of variation for the whole layer
$\boldsymbol{n_{all}}$ – number of all cells
$\boldsymbol{n_i}$ – number of cells in unit i



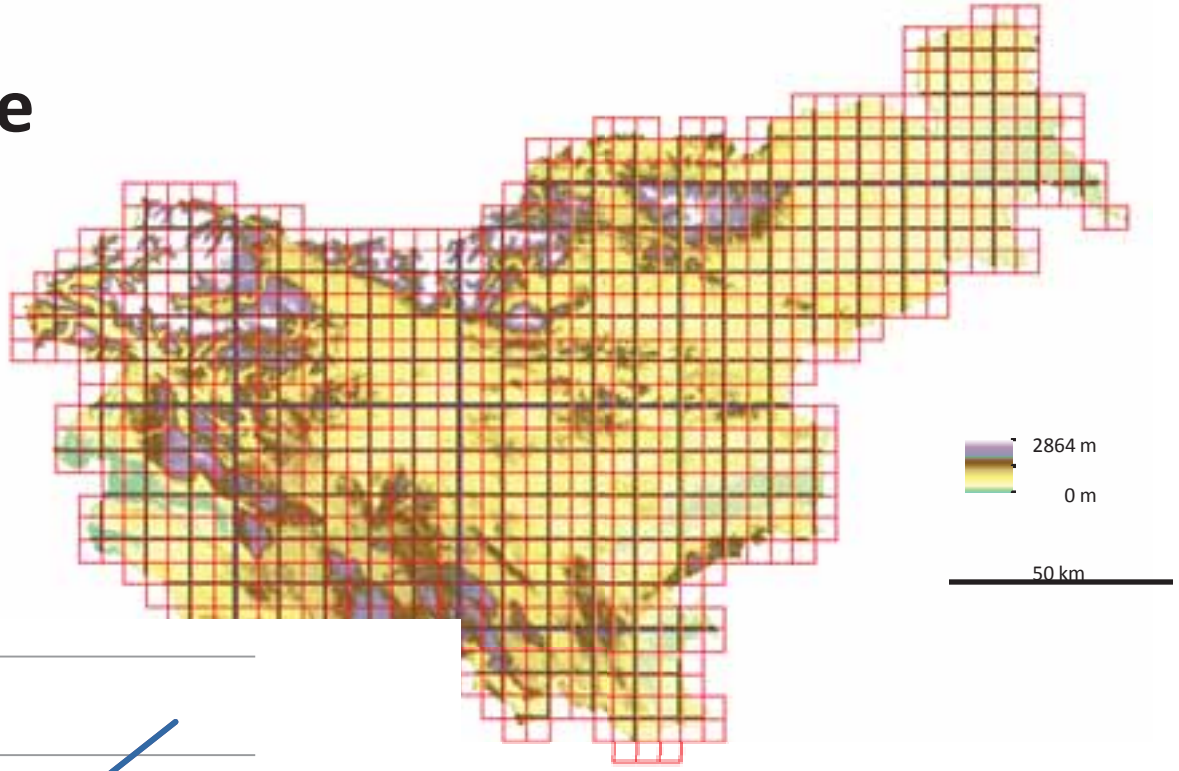
We can do this calculation for different grids (example for Slovenia)

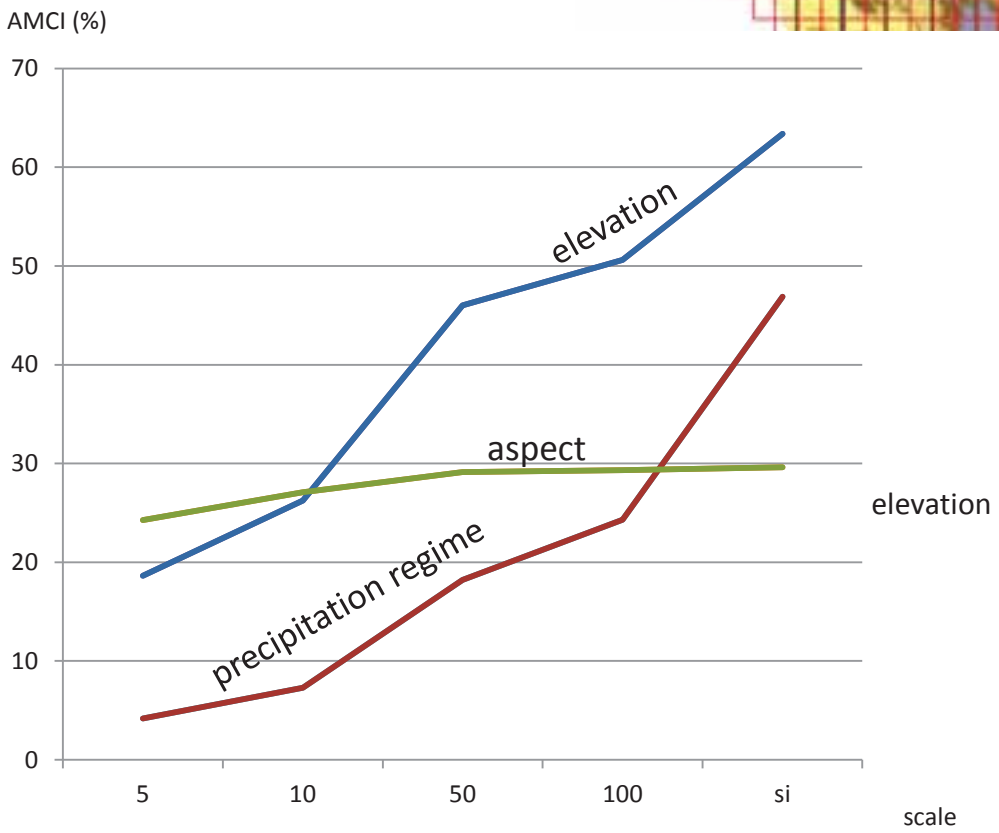


We can do this calculation for different shaped units (example for Slovenia)



Application of the method (case study for Slovenia)

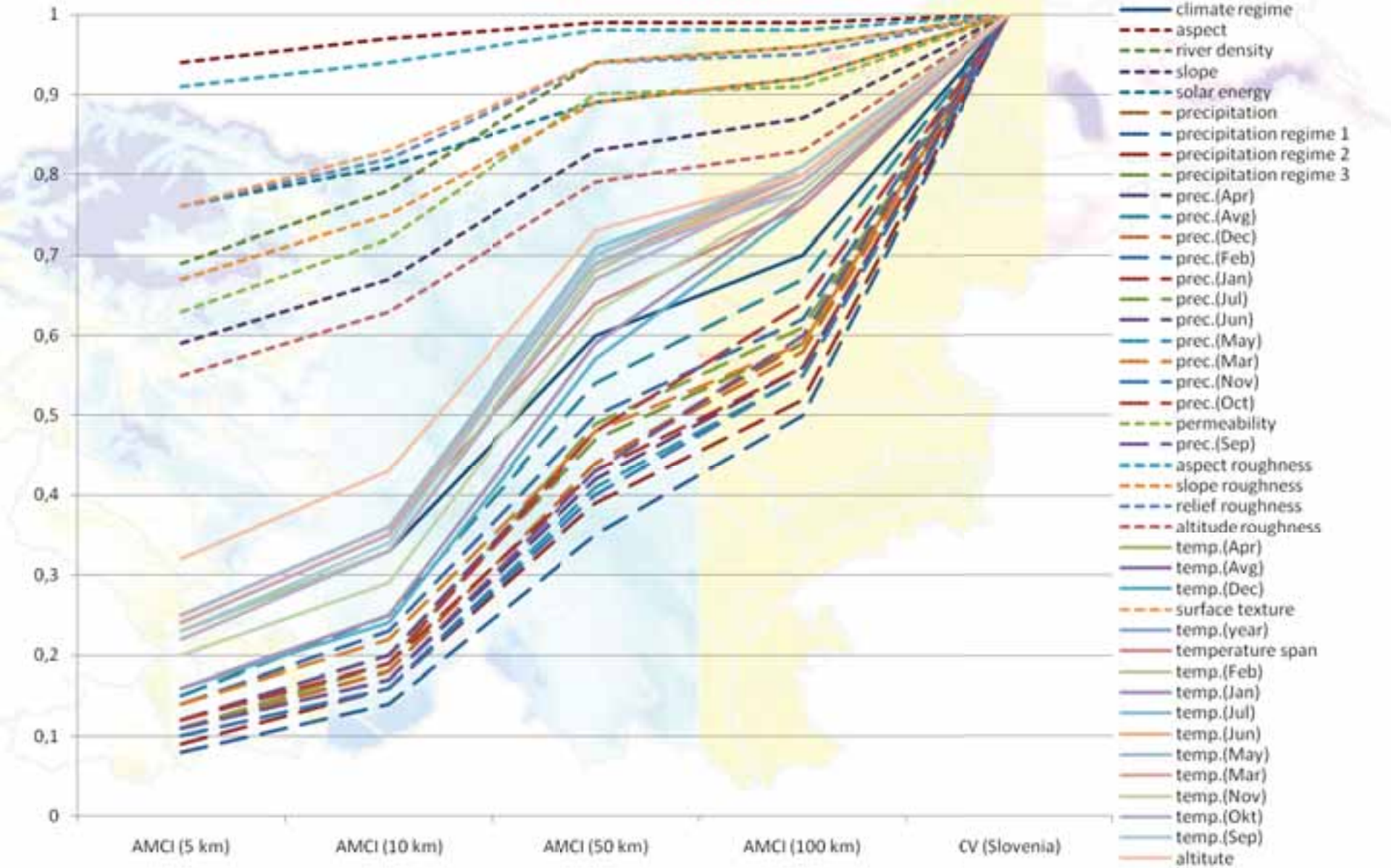




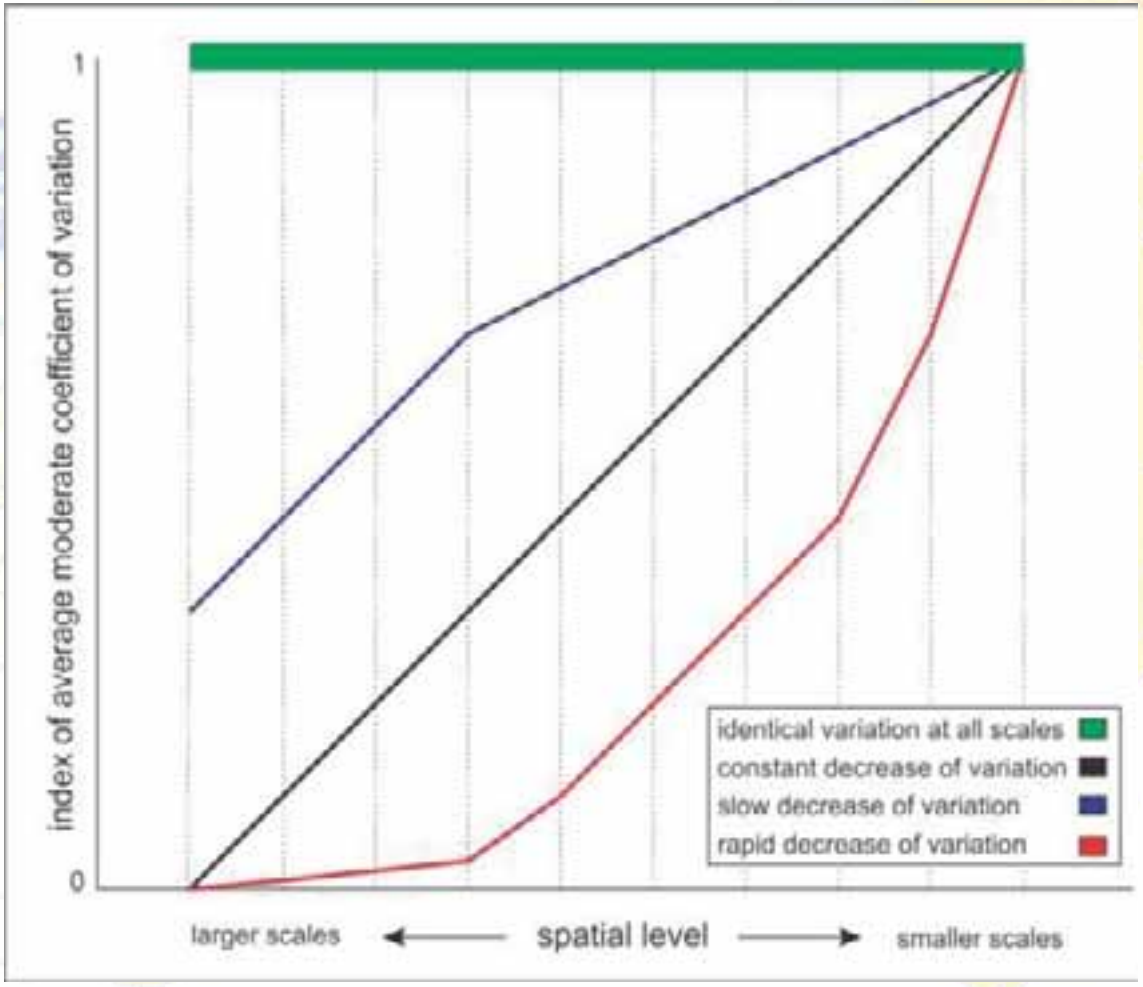
Two ways for representing results

1. Absolute values
2. Relative values (1 = *CV* for a country)

Results for Slovenia



General types of layers



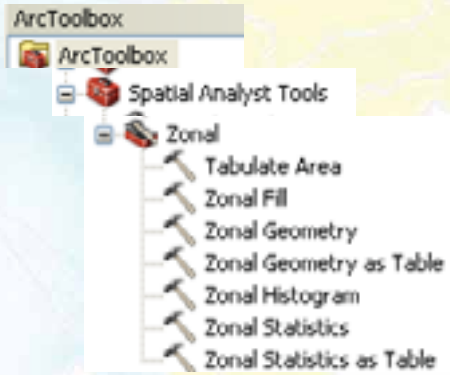
data layers that are useful at smaller scale

data layers that can be useful also at larger scales

data layers that can act as noise (very high absolute variation at all scales) or are not useful at all (very low absolute variation at all scales)

GIS tool

- If you have divisions (i.e. squares) already set you can calculate *AMCV* with ArcGIS (Zonal Statistics as Table) and then finish calculation in any tabular editor.



Python for easier work

Python script:

- **production of a grid of squares with different sides** (arcpy.CreateNormalRaster_sa, arcpy.RasterToPolygon_conversion)
- **production of kmeans units (kmeans of XY coord.)** (arcpy.RasterToPoint_conversion, numpy.array, Pycluster.kcluster, arcpy.PointToRaster_conversion)
- **calculation of a standard deviation for each unit** (arcpy.ZonalStatisticsAsTable_sa)
- **calculation of an *AMCV***

  (arcpy.Statistics_analysis, arcpy.CalculateField_management, arcpy.SearchCursor)

What's next?

- We can select data layers for our landscape classification **more objectively** with *AMCV* calculation.
- With ArcGIS tools and some other packages *AMCV* is **easy to calculate**.
- Things **to do**: to optimize some parts of the script and make it available (as a tool in toolbox) …

Thank you for your attention!

rokciglic@gmail.com