GEOL0061 assessment #1

Pieter Vermeesch

Set on February 10, due on February 20, 2023

In this exercise, you will predict the erosion rate (ϵ, in mm/yr) of a number of points in a landscape with variable tree cover, soil composition and topography. You will work with two data files. Each of you has their own set of these files, stored in a zip archive with your Portico student number on Moodle:

`locations.csv`: a table with the eastings and northings of n samples as well as the local slope (S) and the number of trees (T) counted over a hectare at each location.

`kappa.csv`: a table with m rows and n columns containing m replicate measurements of the *erodibility* (κ, in t/[yr·ha]), which is a parameter that depends on the soil composition and includes factors such as the sand, silt and clay fraction, as well as the organic content of the soil.

Erosion rates can be modelled (i.e. predicted) using the following formula:

$$\epsilon = \kappa S/T \qquad (1)$$

The assessment comprises the following tasks:

1. Visualise the probability distribution of the κ values for the first four samples as kernel density estimates.

2. Summarise the κ-values of all n samples as box plots.

3. Describe the results of steps 1 and 2 in a paragraph of fewer than 100 words.

4. Average the κ-values and compute their standard error for each sample, yielding two new vectors of n values.

5. Estimate the erosion rate for each sample location using Equation 1.

6. Propagate the analytical uncertainty ($s[\epsilon]$) of the n ϵ-values obtained in step 5 assuming a 10% uncertainty of the slope measurements (i.e. $\sigma[S]/S = 0.1$), and taking into account the fact that tree density (T) has Poisson counting uncertainties.

7. Create a scatter plot of $s[\epsilon]$ vs. ϵ.

8. Visualise the predicted erosion rates on a colour-coded map using the `colourplot(X,Y,Z)` function of the `geostats` package. See `?colourplot` for further details about this function, which also does a lot of other stuff that you don't need to worry about for this assessment.

Prepare your solution using `RMarkdown` in `RStudio`. See the video on Moodle for details.

Marking criteria:

+30% for completing the exercise;

+20% if your code does not contain any programming errors;

+20% if your answer does not contain any logical errors;

+30% if you present an original solution that is tailored to your particular dataset and demonstrates a deeper understanding of distributions, data transformations, uncertainty propagation etc;

-50% for plagiarism.

This assessment can be finished in less than one hour, although most students will probably need 2–3 hours. Please do not spend more than one day on it! Read the `R` documentation.