Lab03: Regression Kriging

**Handed out:** Thursday, April 2, 2020

**Return date:** Wednesday, April 16, 2020

**Grading:** This lab counts 16 % towards your final grade

**Objectives:** First you will be using the global trend-surface interpolator and then the local Kriging interpolator on the remaining error component to obtain a combined DEM. Furthermore, you will be evaluating the spatial prediction uncertainty.

**Format of answer:** Your answers (statistical figures and verbal description) should be submitted as Word document to the Lab04 SUBMIT tool in eLearning. Add a running title with the following information: Lab04, your name and page numbers. You may use this document as template. Copy the requested statistical figures into your document. No trial and error answers are permitted. Label each answer properly with the bold task headings. You are expected to hand in professionally formatted answers: use a fixed pitch font, like **Courier New**, for any R code the use mathematical type-setting when equations are required. Copy and paste figures into your document. Make sure that each figure has a proper ***caption*** describing its content.

# Data Description

The zipped file **KansasDEM.zip** contains a DEM elevation points and the river network. UTM zone 14 is the underlying projection, so spherical distortions can be ignored in this example. The elevation (see ELEVATION) is measured in meters. Another variable (see RIVERDIST) measures the distance in miles from each grid cell to the nearest river. See the elevation map below.

|  |
| --- |
| DEMSample |

# Part I: Trend Surface Model (8 points)

Task 01: You may use the script **SampleByClick.r** to pick manually in total 100 sample points. Alternatively, use your favorite GIS program. This task requires careful planning and perhaps a ***nested sampling strategy***. The predictive quality of your model very much depends on the selected sample points. (2 points)

Clearly justify your selection strategies of sample points based on the criteria listed below: (2 points)

(i) You want to avoid any bias in the predicted surface. Therefore, the average predicted elevations should match closely the average observed elevations in the study area. *How can you try to avoid this potential bias?*

(ii) The extrapolation problem should be avoided and the prediction error, in particular at the edges of the study area, need to be minimized. *How many sample points should be assign to control for this problem?*

(iii) The rapid topographic variation along a transection of the river valleys and ridges needs to be captured properly. *How many sample points should be assigned to model this variability, where should they be placed and which variable in the data set measures it?*

(iv) In order to build a well-defined variogram all spatial scales of the inter-sample point distances need to be represented. *How many sample points should be assign to fill in missing distance ranges and where should these points be placed?*

Finally, show the map of your sampling points and the given elevations similar to the map shown above.

*Should you encounter any problems with Task01, please let me know. I can provide you with a set of sample points so you can perform the remaining tasks.*

Task 02: Estimate the 1st, 2nd and 3rd order trend-surface models. Include the distance to the nearest rivers as covariable. (1 point)

Task 03: Map the three predicted trend-surfaces. Use a meaningful color ramp. (1 point)

Task 04: Decide with the partial *F*-test, which of the three surface models is most appropriate for your given sample points. Interpreted the selected trend-surface regression model. (1 point)

Task 05: Evaluate the prediction quality of your most appropriate trend surface model. Does the histogram of observed elevations match that based on the predicted values? Does your prediction model lead to biased overall elevation estimates? If yes, what may be the cause? (1 point)

Task 06: For your most appropriate model, map the ***standard errors*** of the prediction surfaces. Use a meaningful color ramp. Interpret the general pattern in the standard errors. In particular evaluate the standard errors at the edges of the study area relative to those in the center? (1 point)

Task 07: For your most appropriate model, calculate the error component (residual surface: observed DEM minus predicted trend DEM). Map this pattern with a bipolar map theme (zero is the neutral value) and overlay the river network onto your residual map. Interpret this residual pattern. (1 point)

# Part II: Variogram Estimation (2 points)

Task 08: Estimate the variogram function based on the error component at the sampling locations from Part I. Show the necessary plots and interpret them by exploring possible anisotropy, range, sill and nugget effects.

# Part III: Kriging Interpolation of the error component (3 points)

Task 09: Predict the error component by Kriging for all locations. Justify your choice of the Kriging model. Map the surface of the predicted error component with an appropriate color ramp. (1.5 point)

Task 10: Estimate the uncertainty of the error component for all locations. Map the uncertainty surface with an appropriate color ramp. (1.5 point)

# Part IV: Combining Frist and Second Order Components (3 points)

Task 11: Combine the predicted trend-surface with the predicted error component to obtain the overall predicted DEM surface. Map this predicted surface with a proper color ramp. (1 point)

Task 12: Combined the trend-surface prediction uncertainty with the kriging uncertainty in the standard deviation scale. Map the uncertainty surface with a proper color ramp. (1 point)

Task 13: Calculate the root mean squared error of your overall predicted DEM values by comparing it against the observed DEM value of the Kansas topographic surface. (1 point)