Lab Report
Srinivas Konreddy 02/05/25

Question 1: For the month of July, What are the min and max rainfall values for rainfall? Where do each occur (choose a few descriptive locations for the min)?
Answer 1: The maximum value for rainfall is 356.6 mm and the minimum rainfall is 0 mm. The minimum rainfall is found in regions lying in Western Arizona and the southern part of New Mexico.

Question 2: What are TestPoints, and what is their purpose? What is the remaining data called?
Answer 2: TestPoints contains the data that is used to validate and in most cases to also calculate the accuracy of the model. It usually is a small subset of the total data that is used to build the model. The remaining data is called training data and that is used to train/build the model.

Question 3: For NearNeigh, what was the shape and size of your neighborhood? A big theme in our class is how our input affects our output, so why might you choose this particular size? What was the summary statistic you calculated using this method?

Answer 3: The shape of the neighborhood is a circle and the radius of the circle is 300km. By choosing this size we avoid using too small of a circle that might exclude significant spatial features and lead to fragmented analysis and using something bigger might cause generalization in this data. The summary statistic calculated using this method is mean.

Question 4: What is the difference between overlapping and non-overlapping moving windows? Is our overlapping, or non-overlapping? Look up Spatial Analyst → Neighborhood → BLOCK STATISTICS.
Answer 4: An overlapping moving window shifts gradually, meaning neighboring windows share some of the same data, which helps capture finer details but takes more processing power. A non-overlapping moving window moves in fixed steps without overlap, making it more efficient but possibly missing smaller patterns. We have used overlapping moving windows

Question 5: List the specific variables that you can use to describe a neighborhood in focal analysis (i.e. moving windows). List 4 different variables. (This will be a little bit of a repeat from Qs 3 and 4 above).
Answer 5: The 4 different variables would be 1) Overlapping and non-overlapping moving windows, 2) Radius, 3) Shape and 4) Number of nearest neighbors.

Question 6: These questions were answered during lecture: What would happen if my IDW parameter was set to 0—what interpolation method would I essentially be performing?
Answer 6: If the IDW parameter was set to zero, the interpolation method would essentially become nearest neighbor.

Question 7: What would my result resemble if my nearest neighbor was set to 1?
Answer 7: The result output would be a Thiessen Polygon/Voronoi.

Question 8: Did any of the interpolation methods produce values that fell outside the range of the min and max of us_precip_lamb? Which ones? Did they fall above or below the range?
Answer 8: The only interpolation method that produced values that fell outside the range was Kringing which has a negative value for rainfall which is not possible to have.

Question 9: Report on which methods best estimated the test point values.

9a. Fill out the table below.

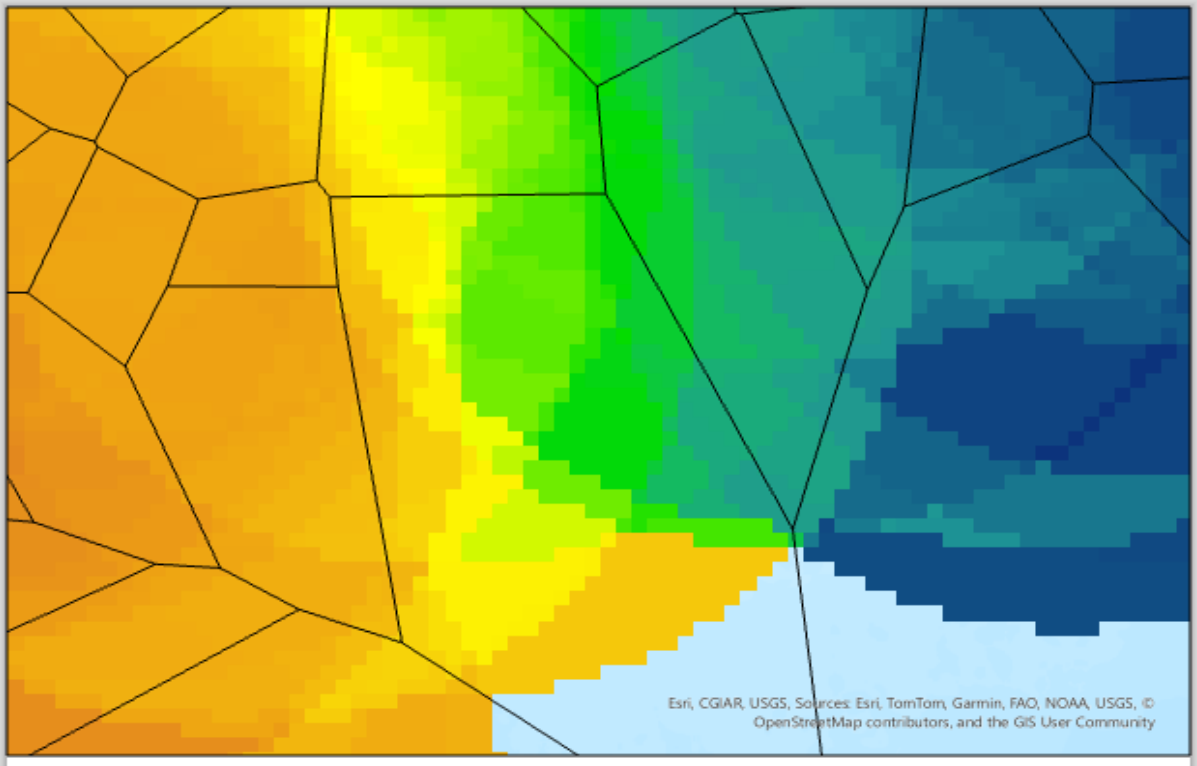
9b. Make a nice argument to describe how each method did and rank these methods from best to worst. FOR FULL POINTS, INCLUDE UNITS in your table.

Raster Method	Min Raster Value (millimeters)	Max Raster Value (millimeters)	Total RMSE Error (millimeters)
Voronoi	0	241.5	50.95
Average Nearest Neighbor	0	203.10	39.22
IDW	0	199.91	37.88
Kriging	-0.12	209.31	35.50

Out of all the interpolation methods used Kringing performed the best having the lowest RMSE values out all of the methods but it falls out of the actual range by having a negative value for rainfall which is not a possible value, IDW is the second having a slightly lower RMSE value than Kringing but did not interpolate outside of the rainfall range. Average Nearest Neighbor performed moderately well having an RMSE value greater than Kringing and IDW. Voronoi performed the worst, having the highest RMSE value, indicating it has the least accuracy amongst all the interpolation methods performed on the dataset.

In my opinion I would rank IDW as the best method, despite having a RMSE error higher than Kringing it did not interpolate outside the range of us_precip_lamb range, the second best would be Kringing as it has the lowest RMSE error but it interpolated outside the range and has a negative rainfall value. The third best would be Average Nearest Neighbor and Voronoi would be last as it has a significantly higher RMSE than the other 3 interpolation methods used in the project.

GRAPHIC 1: On your own. Turn on your vector Thiessen polygons, use only the outline. Underneath it, put your NearNeigh raster symbolized with a very distinctive gradient (use multi colors even though this upsets some people). Zoom in on a Thiessen polygon that has a big gradient change underneath it. Create a legend and a caption. For full points: In the caption, describe what you are seeing and why you think Voronoi polygons are not a great way to estimate this data. For instance, say what are Voronoi polygons better for doing? Present a convincing argument.



Legend

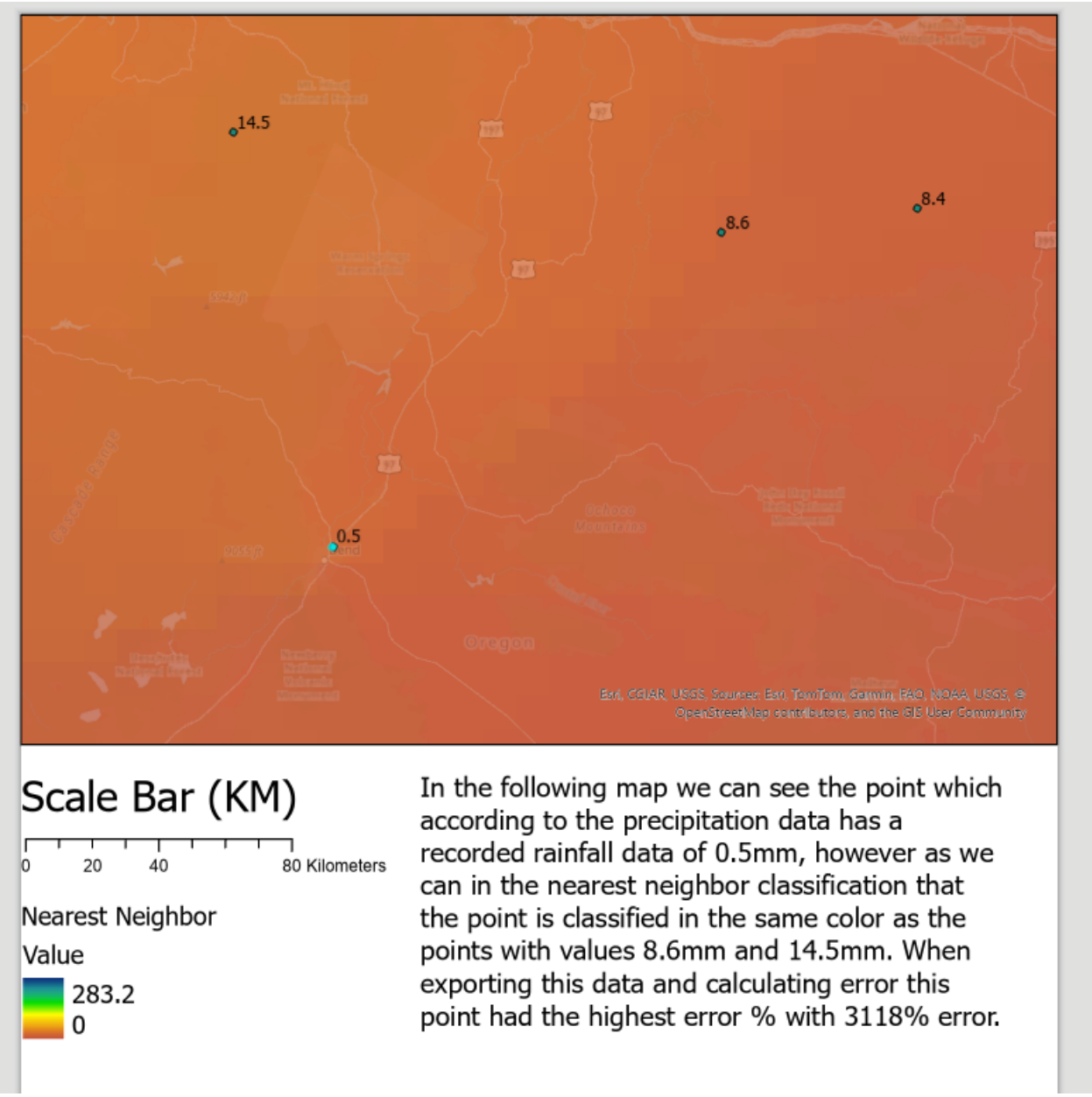
Nearest Neighbor Value

283.2

0

Thiessen polygons (Voronoi) are shown with only outlines over the raster values for the Nearest Neighbor interpolation. The gradient changes in the raster are particularly noticeable under the polygons, where some areas exhibit abrupt shifts in values, which the polygons fail to capture accurately. The Voronoi polygons work well for estimating spatial patterns in data that are more uniform, such as proximity to a set of fixed points. However, in this case, where data is highly variable (e.g., large gradients), Voronoi polygons do not effectively represent the actual distribution. For datasets with such large value fluctuations, methods like IDW or Kriging may provide better representations by incorporating more detailed proximity analysis.

GRAPHIC 2: Make a map or two (GRAPHIC 2) of the specific locations of the biggest errors in the data set. Include legend, scale bar and figure caption.



END OF LAB REPORT

Lab Challenge:

**Use your programming and looping skills to create interpolation grids for each month out of the year, and for each month, find the RSME for one or more methods.

**Use your programming and looping skills to try different sets of 'test points' (perhaps change which 10% you use, or use each number between 5%-15% etc.) and see how your results changed. This is what we call a 'sensitivity test'.

**Add "Spline" as one of your interpolation methods.

PS This is just a fun map, it does not relate to the lab report. This map has some good aspects and some not-so-good aspects…