

Lab Report

Title: Lab 04

Notice: Dr. Bryan Runck

Author: Megan Marsolek

Date: 12/1/21

Project Repository: <https://github.com/mmarsole/GIS5571>

Google Drive Link: NA

Time Spent: 7 hrs

Abstract

The objective of Lab 04 was to compare and contrast Interpolation methods, based on Air Temp data from NDawn's website (I averaged the temperature for 28 stations within MN from Nov 1st to Nov 30th). In this Lab I used Inverse Distance Weighting (IDW), Ordinary Kriging, and Global Polynomial Interpolation (GPI).

Based on the results, I would conclude that GPI isn't necessarily the best method for this set of data (it's too smooth) since our study extent is small (about 240km by 390 km area), if it were a bigger extent it may be better at generalizing air temp trends (for example across the continental US). The Kriging and IDW results were similar with kriging usually being a little bit smoother (although I might be able to change IDW smoothness by adjusting the power parameter). Overall, I would prefer to use the kriging and IDW methods since they provide more detail (variation) based on the data in our smaller extent.

Problem Statement

The Goal is to interpolate (guesstimate) the temperature for areas beyond the known temperature measurements, using 3 methods. This will result in a raster object that has assigned temperature values based on our known temperature observations (the extent is determined by the minimized bounding box that can fit all the observed data). The first step is analyzing and obtaining relevant temperature data. Once this data has been acquired, the desired data (Ave temp, Min and Max temp) is extracted.

Table 1. Data Needed

#	Requirement	Defined As	(Spatial) Data	Attribute Data	Dataset	Preparation
1	Temperature (in Fahrenheit)	Hourly recordings of temp (F) from Nov 1 st to Nov 30 th or 2021	CSV data (converted to point)	Temp (extracted from this Average temp, Min, and Max temp values	NDawn center	Had to extract the averages, Min, and Max temp values per station, reshape the CSV so each row was representative of each station location, then convert the csv to a feature class within ArcPro

Input Data

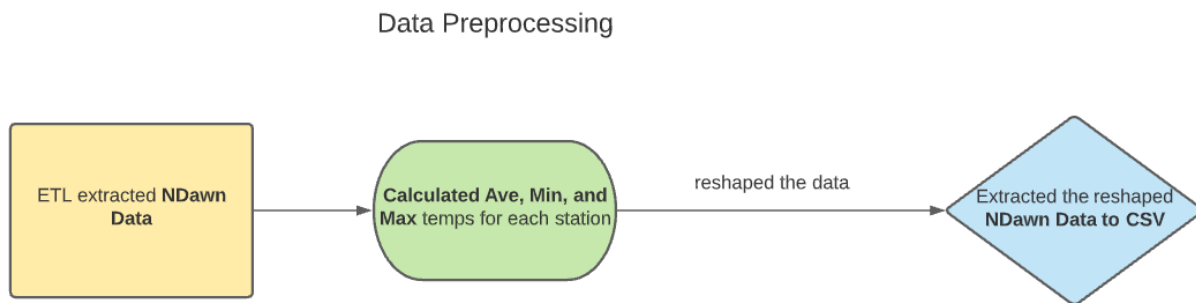
The data from NDawn was originally shaped so that each row was representative of time and location (temp was recorded each hour, so each row was the recorded temp for a given hour on a given day for each station). I reshaped the data so that each row was representative of each station (since I only needed the recorded min, max, and average temp for each location for the past month), thus resulting in a smaller dataset that was then used for interpolation (see Figure 1).

Table 2. Acquired Data

#	Title	Purpose in Analysis	Link to Source
1	Temperature Data	The NDawn data returns temperature recording on an hourly increment, from which I calculated the 30 day average (Nov 1 st to Nov30th), the Min , and the Max observed temps for each station.	NDawn center

Methods

Figure 1: Data Preprocessing

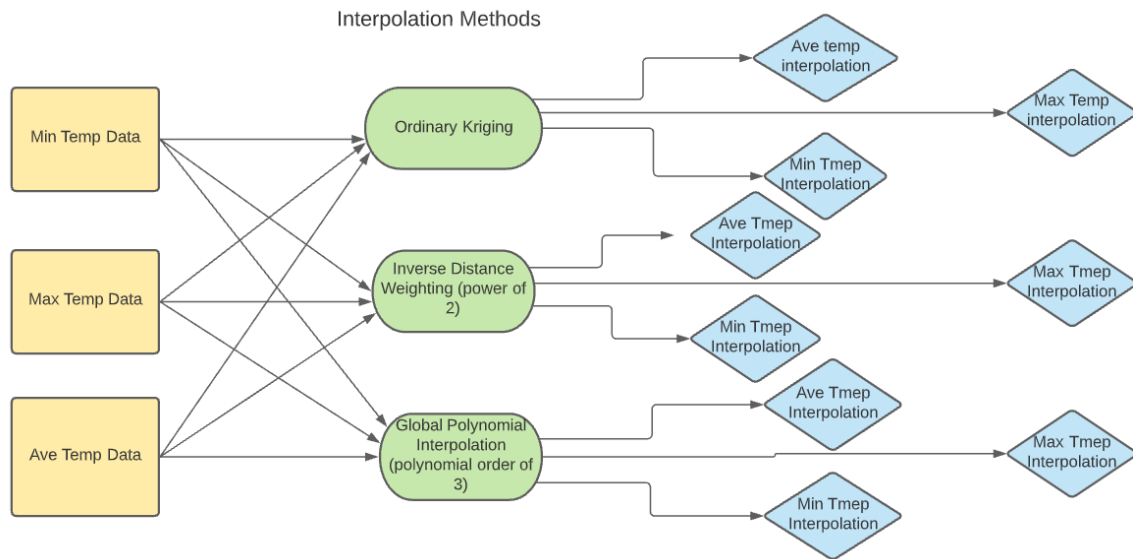


Before any Interpolation could be done, I first had to preprocess the data (see Figure 1) so that I had the Min, Max, and Ave Temp data per station. Once I had this data, I fed each observed measure into an interpolation method (see Figure 2) which resulted in an interpolation of the bounding extant for the input points (it populated temp values for the entire bounding extant of the original observations).

It should be noted that there were two important parameter that I played with in both IDW and GPI methods. The IDW can be altered by changing the power parameter, “As the power increases, the interpolated values begin to approach the value of the nearest sample point. Specifying a lower value for power will give more influence to surrounding points that are farther away, resulting in a smoother surface” (How IDW works—Help | ArcGIS for Desktop, n.d.). While increasing the order of polynomial in the GPI allows for more ‘bend’ in the fitting of a plane to the data (How global polynomial interpolation works—ArcMap | Documentation, n.d.). It is recommended that using low-order polynomial is more suitable for some physical process (such as pollution and wind direction) (How global polynomial interpolation works—ArcMap | Documentation, n.d.).

For my results, I chose an order of 2 for IDW (trying to reduce dependencies on far away points), and an order of 3 for GPI (I initially used and order of 1 because temp is a physical process, but thought the resulting interpolation was too linear, and so increased it to 3 to allow more flexibility in the resulting interpolation).

Figure 2. Interpolation Methods



Results

Looking at the interpolation results for Ave Temp (see Image 1), I would stipulate that IDW and Kriging are similar while the GPI result is starkly different from the other two, in that it is more linear (in IDW and kriging you can see the presence of more circular observations populated around select points whereas the GPI seems to have a diagonal striped interpolation throughout the extant). Based on these results I would surmise that GPI is more of an average of all points (meaning it is as the name suggests, meant for smoothing all the observed data and returning an interpolation that is best suited for a more Global, bigger, extant) when compared to kriging and IDW.

Image 1: Interpolation results for Ave temperature

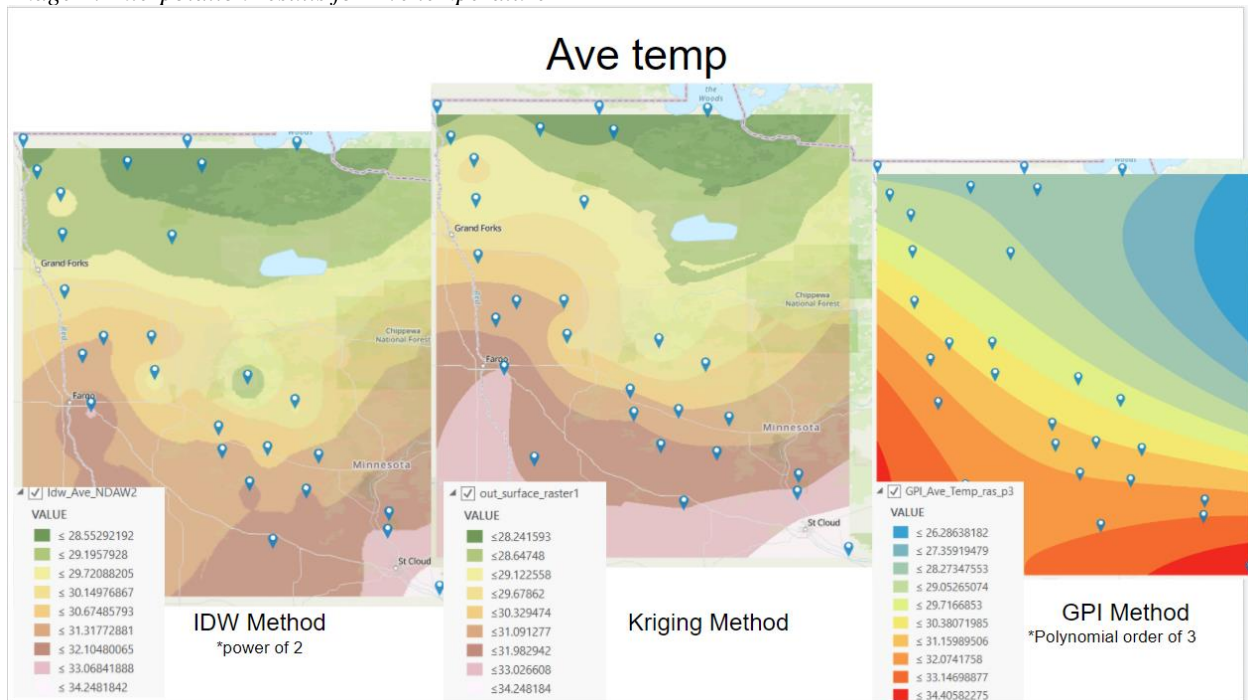
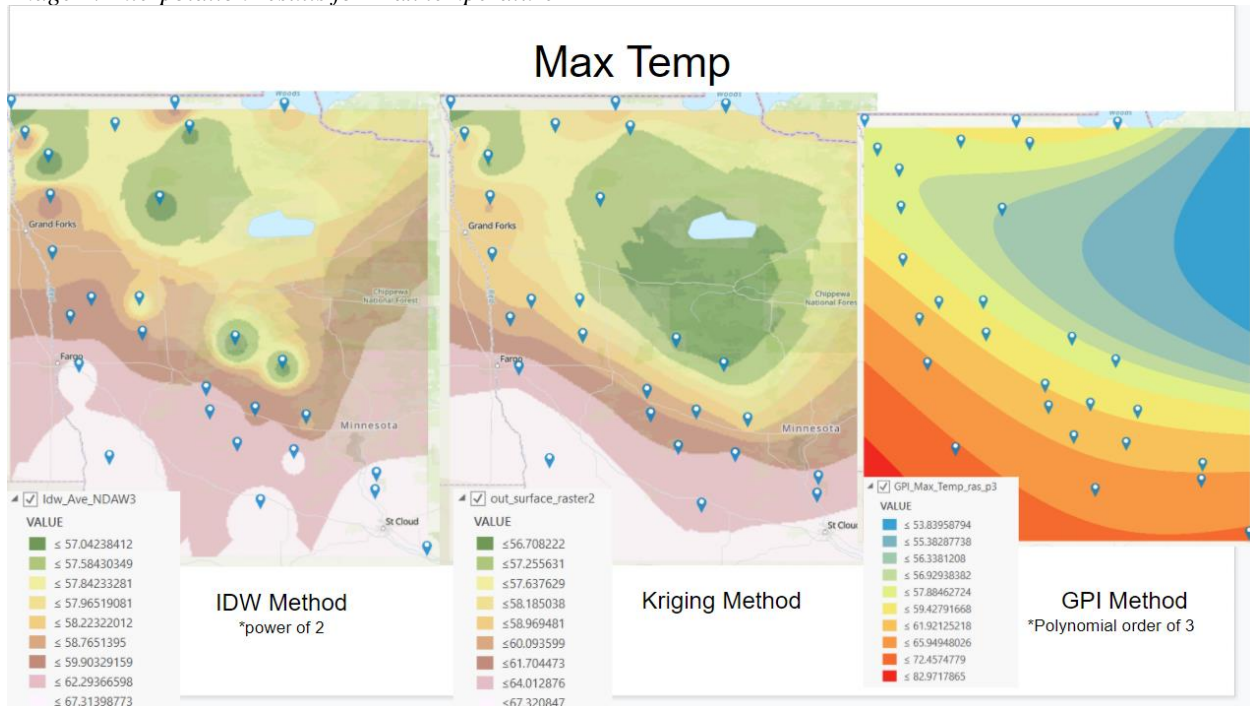
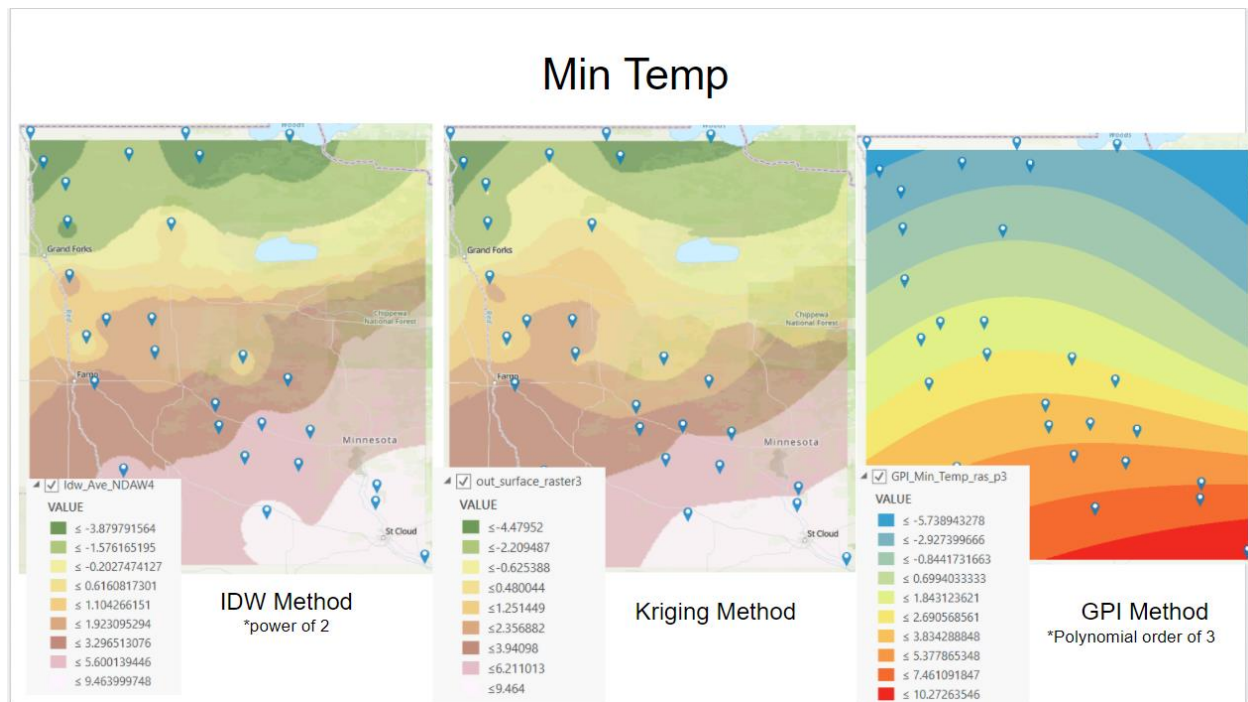


Image 2: Interpolation results for Max temperature



Based on the interpolation results for Min Temp (see Image 3), I can see that cooler min temps are more likely to be observed in the north while warmer min temps occur in the south (this supports the general understanding that that it gets colder the farther North you go). This is inversely corroborated when we look at the results for the Max Temps (see image 2). We can observe that the general highest recorded temps are situated more in the South, and as we move northward, we see the Max temps drop when compared to the more southerly temperatures (exempting the circular pattern in the IDW and Kriging results that suggest there are ‘pockets’ where lower high temps are recorded, which could suggest geological feature that may reduce this area’s general max temp like mountains or higher elevation).

Image3: Interpolation results for Min temperature



Results Verification

I surmise my results are correct though cross comparison (by comparing the general outputs of three different interpolation methods I can see if there are any obvious inconsistencies) and replication of results (I performed all interpolation within the GUI as well via python code and attained consistent results).

Discussion and Conclusion

So, I was asked to compare and contrast the 3 methods I used for interpolation. Based on Esri's article: "Classification trees of the interpolation methods offered in Geostatistical Analyst" and Eldrandaly's article: "Comparison of Six GIS-Based Spatial Interpolation Methods for Estimating Air Temperature in Western Saudi Arabia" I will summarize a few key distinctions between, IDW, Kriging, and GPI. Both IDW and GPI are deterministic interpolation methods, meaning they will always produce the same results given the same parameters, but kriging is a statistical based interpolation method (it produces probabilities and error estimates of what the interpolated value).

Additionally, GPI is a global interpolation method, meaning it doesn't weight or group neighbors (like IDW or Kriging), instead uses all available data to make its interpolation, and this may be better suited for larger extents where we want a general understanding or the 'Big Picture' instead of the local variations. As a result, GPI tends to produce the 'smoothest' (most averaged) interpolation, Kriging is somewhat smooth (it groups point pairs in groups based on distances and then calculates is semivariogram based on these grouping which in turn is then used to interpolate values), while IDW results in the least smooth (since it considers each point and weights it based on proximity to each point). See a general summary of my compare and contrast in Table 1.

Table 1: A Summary of Comparing and Contrasting three interpolation methods (based on explanations from (Eldrandaly, 2011) and (Classification trees of the interpolation methods offered in Geostatistical Analyst—ArcMap / Documentation, n.d.)

	Deterministic vs Stochastic	Spatial Autocorrelation	Resulting interpolation level of 'Smoothness'	# of Assumption made
IDW	Deterministic	Local	Not smooth	few
Kriging	Stochastic	Local	Somewhat smooth	many
GPI	Deterministic	Global	Smooth	few

References

- Classification trees of the interpolation methods offered in Geostatistical Analyst—ArcMap / Documentation.* (n.d.). Desktop.arcgis.com. Retrieved December 1, 2021, from <https://desktop.arcgis.com/en/arcmap/latest/extensions/geostatistical-analyst/classification-trees-of-the-interpolation-methods-offered-in-geostatistical-analyst.htm>
- Eldrandaly, K. A. (2011). Comparison of Six GIS-Based Spatial Interpolation Methods for Estimating Air Temperature in Western Saudi Arabia. *Journal of Environmental Informatics*, 18(1), 38–45. <https://doi.org/10.3808/jei.201100197>
- How global polynomial interpolation works—ArcMap / Documentation.* (n.d.). Desktop.arcgis.com. Retrieved December 1, 2021, from <https://desktop.arcgis.com/en/arcmap/latest/extensions/geostatistical-analyst/how-global-polynomial-interpolation-works.htm#GUID-F8D5E60F-E08A-4B20-AA7D-6E2E74FD127D>
- How IDW works—Help / ArcGIS for Desktop.* (n.d.). Desktop.arcgis.com. <https://desktop.arcgis.com/en/arcmap/10.3/tools/3d-analyst-toolbox/how-idw-works.htm>

Self-score

Category	Description	Points Possible	Score
Structural Elements	All elements of a lab report are included (2 points each): Title, Notice: Dr. Bryan Runck, Author, Project Repository, Date, Abstract, Problem Statement, Input Data w/ tables, Methods w/ Data, Flow Diagrams, Results, Results Verification, Discussion and Conclusion, References in common format, Self-score	28	28
Clarity of Content	Each element above is executed at a professional level so that someone can understand the goal, data, methods, results, and their validity and implications in a 5 minute reading at a cursory-level, and in a 30 minute meeting at a deep level (12 points). There is a clear connection from data to results to discussion and conclusion (12 points).	24	22
Reproducibility	Results are completely reproducible by someone with basic GIS training. There is no ambiguity in data flow or rationale for data operations. Every step is documented and justified.	28	25
Verification	Results are correct in that they have been verified in comparison to some standard. The standard is clearly stated (10 points), the method of comparison is clearly stated (5 points), and the result of verification is clearly stated (5 points).	20	18
		100	93