Lab Report

Title: Interpolation of NDAWN temperature data using different methods
Notice: Dr. Bryan Runck
Author: Samikshya Subedi
Date: 12/7/2022
Project Repository: https://github.com/Samikshya036/GIS5571/tree/main/Lab3
Time Spent: 13 hours

Abstract
Four different interpolation methods are used for this particular lab in order to estimate the temperature at each NDAWN station, and the results are contrasted and compared by building an ETL to pull the last 30 days of temperature data from the DNAWN site for all of the NDAWN stations and to making a map of all the station points and their average monthly temperature. The most effective interpolation technique depends on a number of factors. There is no single method that solves all problems; it depends on the type of variable and the time scale on which it is represented. It can be claimed that the IDW approach is the best choice for interpolating this set of average temperature data. With this approach, the influence of the variable being mapped lessens as it gets away from the sampling point.

Problem Statement
To build an ETL to pull the last 30 days of temperature data from the DNAWN site for all of the NDAWN stations and to make a map of all the station points and their average monthly temperature. And comparing and contrasting at least 3 of the interpolation technique and be able to run your notebook and create an interpolated temperature map for the highs and lows of the last 30 days from NDAWN in real-time.

Four different interpolation methods are used for this particular lab in order to estimate the temperature at each NDAWN station, and the results are contrasted and compared.

Table 1. Table showing data requirement

#	Requirement	Defined As	(Spatial) Data	Attribute Data	Dataset	Preparation
1	Weather Data (Temperature)	Temperature data set from NDAWN	Discontinuous	Degrees	Average Temperature	NDAWN
2	Min Temperature	Temperature data set from NDAWN	Discontinuous	Degrees	Temperature Min	NDAWN
3	Max Temperature	Temperature data set from NDAWN	Discontinuous	Degrees	Temperature Max	NDAWN

Input Data
Weather dataset for thw analysis of average temperature, min and max temperature from last 30 days in NDAWN weather stations.

Table 2. Input data set for the interpolation of Min/Max temperature and average temperature.

#	Title	Purpose in Analysis	Link to Source
1	Weather Data from NDAWN	Latest monthly Raw data for analyzing minimum/ maximum temperatures and monthly average at all NDAWN stations.	NDAWN

Methods

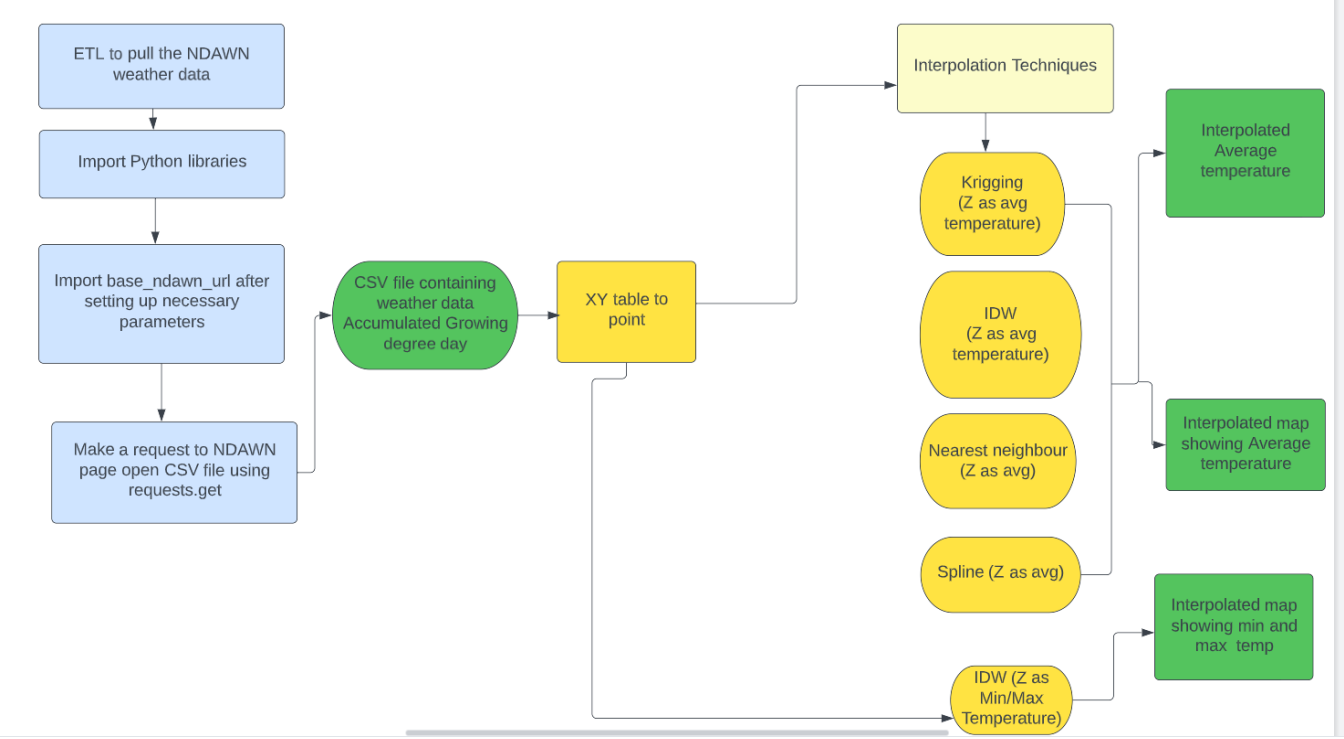

Figure 1: Figure showing workflow diagram

1. ETL pipeline in Jupyter notebook to pull weather dataset from NDAWN
2. Import python libraries
3. import requests
4. import numpy as np
5. import sys
6. import pandas as pd
7. Download CSV file
8. data_req = requests.get(ndawn_url)
9. with open('AvgTemp_AllStations.csv', 'w') as AvgTempsAll:
10. XY table to point
11. Four different interpolation techniques (IDW, Spline, Natural neighbor, Spline)
12. Interpolated average Temp map, min temperature map and max temp map.

Results

To obtain the interpolated average temperature maps, the process shown in methodoly is used. IDW performed better than Kriging, Spline, and Natural Neighbor.

Finding an appropriate interpolation technique is crucial for producing the best estimates of values at unknowable places. With the use of IDW interpolation, sample points are given weights so that the impact of one point on another decreases with increasing distance from the new point being estimated. It can be claimed that the IDW approach is the best choice for interpolating this set of average temperature data. With this approach, the influence of the variable being mapped lessens as it gets away from the sampling point.

The choice of interpolation method can also be influenced by the quality of the sample point set. If the sample points are few or poorly distributed, which may be the case below, the surface may not adequately represent the topography. Given the close proximity of the sample sites, IDW probably is the good fit for the interpolation of maximum and minimum temperature.

In [10]:
```
with open('AvgTemp_AllStations.csv', 'w') as AvgTempsAll:
    AvgTempsAll.write(data_req.content.decode('utf-8')) #decode with unicode 8 parameters
avg_temp_df = pd.read_csv('AvgTemp_AllStations.csv', skiprows = 5)
avg_temp_df = avg_temp_df.iloc[1:] #initializes df from row 2 to skip the messy labels in row index 0
avg_temp_df
```

Out[10]:

	Station Name	Latitude	Longitude	Elevation	Year	Month	Day	Max Temp	Max Temp Flag	Normal Max Temp	Departure from Normal Daily Maximum Air Temperature	Departure from Normal Daily Maximum Air Temperature Flag	Min Temp	Min Temp Flag	Normal Min Temp	Departure from Normal Daily Minimum Air Temperature
1	Ada	47.32119	-96.51406	910	2022.0	11.0	8.0	43.718	NaN	40.73	2.988	NaN	35.847	NaN	22.28	13.567
2	Ada	47.32119	-96.51406	910	2022.0	11.0	9.0	42.922	NaN	40.14	2.782	NaN	30.760	NaN	21.77	8.990
3	Ada	47.32119	-96.51406	910	2022.0	11.0	10.0	31.653	NaN	39.54	-7.887	NaN	22.874	NaN	21.37	1.504
4	Ada	47.32119	-96.51406	910	2022.0	11.0	11.0	22.933	NaN	38.94	-16.007	NaN	18.009	NaN	20.87	-2.861
5	Ada	47.32119	-96.51406	910	2022.0	11.0	12.0	19.850	NaN	38.34	-18.490	NaN	16.808	NaN	20.37	-3.562
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
4226	Zeeland	46.01351	-99.68768	2070	2022.0	12.0	3.0	20.611	NaN	31.23	-10.619	NaN	-11.020	NaN	11.52	-22.540
4227	Zeeland	46.01351	-99.68768	2070	2022.0	12.0	4.0	36.171	NaN	30.81	5.361	NaN	1.238	NaN	11.19	-9.952
4228	Zeeland	46.01351	-99.68768	2070	2022.0	12.0	5.0	22.087	NaN	30.36	-8.273	NaN	6.548	NaN	10.91	-4.362
4229	Zeeland	46.01351	-99.68768	2070	2022.0	12.0	6.0	20.197	NaN	29.94	-9.743	NaN	2.876	NaN	10.58	-7.704
4230	Zeeland	46.01351	-99.68768	2070	2022.0	12.0	7.0	3.614	NaN	29.54	-25.926	NaN	-3.856	NaN	10.28	-14.136

4230 rows × 21 columns

Figure2: Figure showing CSV file of NDAWN weather stations containing temperature datasets.

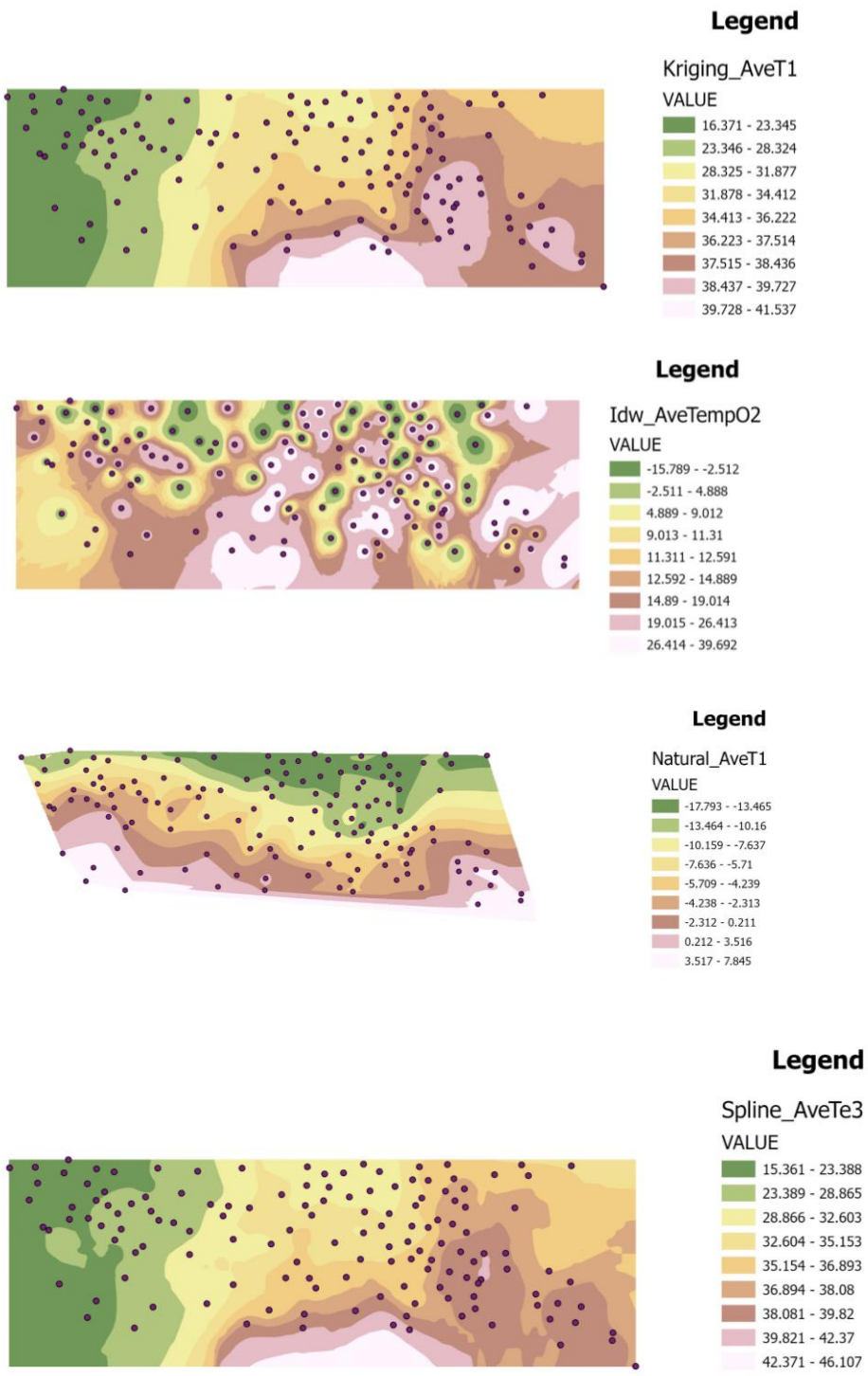

Figure3: Figure showing results on interpolated map of average temperature by using 4 different interpolation technique, Among which Given the close proximity of the sample sites, IDW probably is the good fit for the interpolation of maximum and minimum temperature.

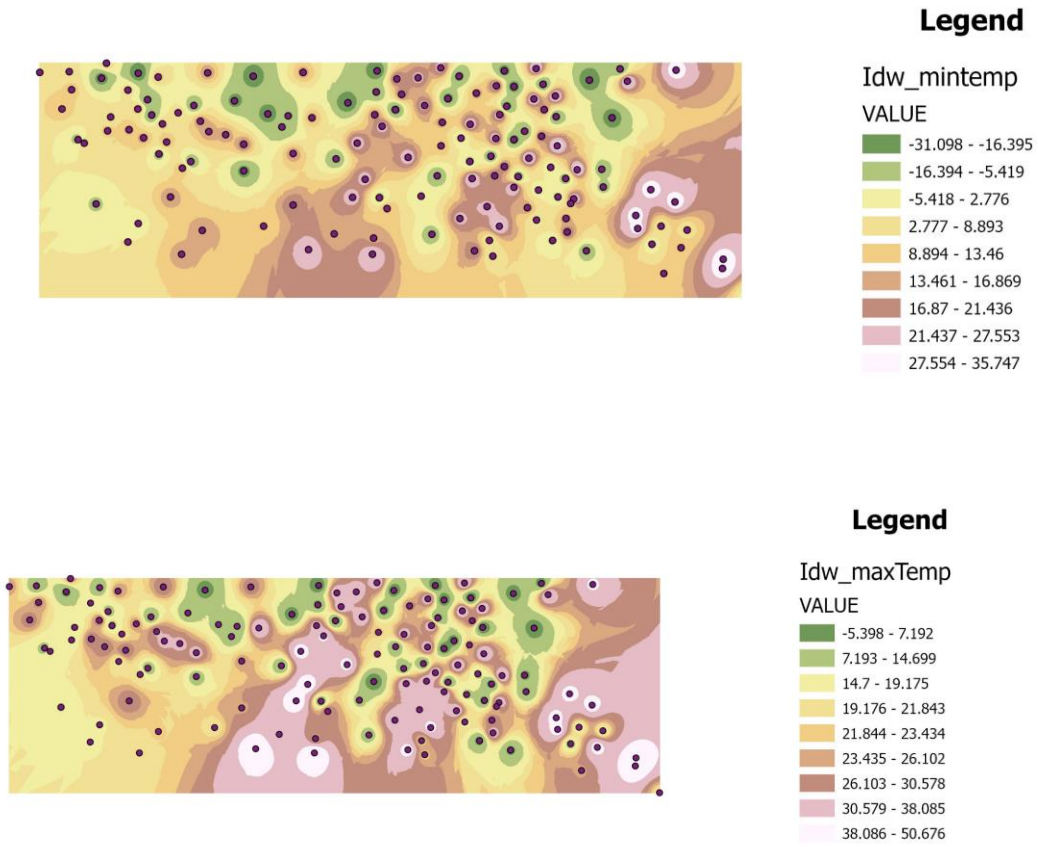

Figure 4: Figure showing results on interpolated map using IDW map with Z as min temperature in upper figure and Z as max temperature in lower figure in last 30 days in NDAWN weather stations.

Results Verification

As IDW covers the average, I visually checked the results better than other temperature maps in terms of temperature range. I ran the tools with the same parameters. For the monthly mean temperature, the IDW is the best spatial interpolation technique based on cross-validation. According to the authors in listed citations, measurement precision, density, distribution, and geographic variability all have an impact on how accurate interpolation approaches are (Chai et al., 2011; Kurtzman & Kadmon, 1999).

Discussion and Conclusion

This lab 3 part 2 is very important for me because I could apply all the methodology of this lab in my own project. I learned different interpolation techniques and learnt to choose bet fit interpolation method. The most effective interpolation technique depends on a number of factors. There is no single method that solves all problems; it depends on the type of variable and the time scale on which it is represented. Instead of thinking one interpolation approach is better than another, all interpolation strategies should be researched and their results should be compared in order to determine the best interpolation method for a specific project.

References

Chai, H., Cheng, W., Zhou, C., Chen, X., Ma, X., & Zhao, S. (2011). Analysis and comparison of spatial interpolation methods for temperature data in Xinjiang Uygur Autonomous Region, China. *Natural Science*, *3*(12), 999.

Kurtzman, D., & Kadmon, R. (1999). Mapping of temperature variables in Israel: sa comparison of different interpolation methods. *Climate Research*, *13*(1), 33–43.

How IDW works. ArcGIS for Desktop. Website. November 20, 2022. https://desktop.arcgis.com/en/arcmap/10.3/tools/3d-analyst-toolbox/how-idw-works.htm

NDAWN Station Location. https://ndawn.ndsu.nodak.edu/weather-data-daily.html

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	**24**
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	**26**
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	**17**
		100	**95**