**Lab Report 3.2**

Title: Interpolating Temperature Data

Notice: Dr. Bryan Runck

Author: Timothy Johnson

Date: 26 November 2024

**Project Repository:**<https://github.com/tjjohnson1415/GIS5571>

**Time Spent:** 3.5 hours

**Abstract**

Temperature is a quantity that we often want to know the value of at places where it is not measured. Interpolation methods can be used to estimate temperature at these points. Data from NDAWN is used to run three interpolation methods (inverse-distance weighting, natural neighbor, and Kriging) and compare the results. The results of the three methods have the same general pattern, but will have different interpolated values.

**Problem Statement**

Temperature forms a continuous surface, but we can only measure it at a finite number of locations. If we want to know the temperature at a location where we did not or cannot measure it, we must interpolate based on the known temperatures at other locations. Surfaces for 30-day high temperature, 30-day low temperature, and 30-day average temperature will be produced based on temperature data at weather stations that report to NDAWN. These surfaces will be produced using three methods: inverse-distance weighting, Kriging, and natural neighbor interpolation.

Table 1: Data used in this lab.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **#** | **Requirement** | **Defined As** | **(Spatial) Data** | **Attribute Data** | **Dataset** | **Preparation** |
| 1 | 30 days of temperature data | Daily weather station data from NDAWN | Weather station point locations | Average temperature, high temperature, low temperature | [NDAWN](https://ndawn.ndsu.nodak.edu/table.csv?station%3e0&variable=ddavt&variable=ddmxt&variable=ddmnt&variableyear=2024&ttype=daily&quick_pick=30_d&begin_date=2024-11-25&end_date=2024-11-25) |  |

**Input Data**

The only dataset used in this lab is a temperature dataset from NDAWN. The records span a 30-day period from October 27, 2024, to November 25, 2024. Each record contains a station’s single day high temperature, low temperature, and average temperature.

Table 2: Data

|  |  |  |  |
| --- | --- | --- | --- |
| **#** | **Title** | **Purpose in Analysis** | **Link to Source** |
| 1 | NDAWN Temperature Data | Raw input dataset for determining the 30-day high, 30-day low, and 30-day average temperature at each reporting weather station. | [NDAWN](https://ndawn.ndsu.nodak.edu/table.csv?station%3e0&variable=ddavt&variable=ddmxt&variable=ddmnt&variableyear=2024&ttype=daily&quick_pick=30_d&begin_date=2024-11-25&end_date=2024-11-25) |

**Methods**

**A diagram of a process

Description automatically generated**

Figure 1: Data flow diagram for processing NDAWN data and interpolating temperature surfaces.

A .csv file containing the data was read into a Pandas dataframe. Because the data includes one record for each station for each day, the data needed to be aggregated by station. During the aggregation, the maximum of the high temperatures, the minimum of the low temperatures, and the mean of the average temperatures were each calculated to be included in the new aggregated dataframe. The data was converted to a spatially enabled dataframe using the latitude and longitude provided for each station. Some stations did not have data for all thirty days, so these stations were removed because we cannot know the 30-day high, low, or average temperatures. The spatially enabled dataframe was converted to a feature class and three interpolation methods (inverse-distance weighting, natural neighbor, and Kriging) were used to create interpolated surfaces for each 30-day high, 30-day low, and 30-day average. The symbology was standardized across the three methods to make the surfaces visually comparable.

**Results**

Figure 2 shows the results of the three interpolation methods (inverse distance weighting, natural neighbor, and Kriging) for the 30-day high temperature. Figure 3 shows the results for interpolating the 30-day low temperature. Figure 4 shows the results for interpolating the 30-day average temperature.

IDW

**A map of the world

Description automatically generated**

**A map of the earth

Description automatically generated**

Natural Neighbor

**A map of the earth

Description automatically generated**

Kriging

Figure 2: Results of interpolation for 30-day high temperature.

**A map of different colors

Description automatically generated**

IDW

A map of different colors

Description automatically generated

Natural Neighbor

Kriging

A map of different colors

Description automatically generated

Figure 3: Results of interpolation for 30-day low temperature.

A map of different colors

Description automatically generated

IDW

Natural Neighbor

Kriging

A map of the earth

Description automatically generated

A map of different colors

Description automatically generated

Figure 4: Results of interpolation for 30-day average temperature.

**Results Verification**

The results were verified qualitatively by visually comparing the outputs of the three interpolation methods. Because the color scheme was standardized across outputs, we can expect that each color should appear in approximately the same area from each of the three interpolation methods. Similar patterns are observed for each method, which suggests that the output is correct.

**Discussion and Conclusion**

All three interpolation methods were very simple to run using arcpy, and they produced similar results. The same general patterns are observed across the three interpolation methods for each 30-day high temperature, 30-day low temperature, and 30-day average temperature. While inverse-distance weighting and Kriging can be used to extrapolate beyond the spatial extent of the known points, natural neighbor cannot, and instead produces a surface with the shape of a polygon that fits all the known points.

**References**

*IDW (Spatial Analyst)—ArcGIS Pro | Documentation*. (n.d.). Retrieved November 26, 2024, from <https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-analyst/idw.htm>

*Kriging (Spatial Analyst)—ArcGIS Pro | Documentation*. (n.d.). Retrieved November 26, 2024, from <https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-analyst/kriging.htm>

*Natural Neighbor (Spatial Analyst)—ArcGIS Pro | Documentation*. (n.d.). Retrieved November 26, 2024, from <https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-analyst/natural-neighbor.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 | **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 | **28** |
| **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 | **97** |