GIS-5 DEM Generation

Per Halvorsen

2025-04-17

pmhalvor/geo4460
pmhalvor@uio.no
GEO4460

1 Introduction

In this report, we will look into Digital Elevation Modeling (DEM). Specifically, we'll generate DEMs using the following techniques:

- Natural Neighbor Interpolation
- Triangulated Irregular Network (TIN)
- Topo to Raster (Hutchinson's ANUDEM algorithm)

The goal of this exercise is to learn DEM generation techniques, understand how incorporating more informative data can improve DEM accuracy, and practice combining both quantitative and visual analysis methods.

2 Data

The data used to build this analysis was extracted from the N50 Dataset, by Kartverket hosted on GeoNorge.

We automated most of the workflow of this project in Python using the following software:

- Whitebox-tools: an open source Python library for geospatial analysis
- Rasterio: a Python library for reading and writing geospatial raster data
- GRASS GIS: an open source GIS software with a Python API and additional capabilities including GUI & Stream Burning
- ArcGIS Pro: for the ANUDEM algorithm and for visual quality assessments

The code for this project is found at github.com/pmhalvor/geo4460.

2.1 Preprocessing

Before we could go through the steps outlined in the project description, we needed to convert the raw N50 data to a similar geodatabase as our example data, containing feature layers `contour_arc`, `elevation_point`, `lake_polygon`, and `river_arc`.

To do so we used ArcGIS Pro to run the following steps:

1. Import the N50 data into a new geodatabase.

2. Load the land cover data, elevation points, and contour lines to the current map.
3. Create a feature layer containing a single polygon (rectangle) around the area of interest. We chose a region in Rogaland around Preikestolen.
4. Clip the N50 layers to the rectangle polygon.
5. Convert labels in the land cover data to new feature layer polygons when the object type was `innsjø`, `innsjø regulert`, or `elv`.
6. Convert river feature layers to polylines.
7. Rename relevant layers to `contour_arc`, `elevation_point`, `lake_polygon`, and `river_arc`.
8. Export the layers to a new geodatabase.

These steps were executed in ArcGIS Pro due to limitations of open-source tools for reading, parsing, and storing `.gdb` files. Also, performing these steps manually in a GUI ensured the selection of an interesting region to analyze within the boundaries of our data, providing a better foundation for understanding the results later.

Two additional preprocessing steps were performed directly in our Python workflow. These final preprocessing steps enforced the correct coordinate system across our layers and converted our contour lines to points, for more informative DEM interpolation. Generating points from contours dramatically improves the quality of simpler DEMs, as will be discussed later.

3 Methods

3.1 DEM generation

For this specific analysis, we explored 3 interpolation methods: Natural Neighbor, TIN Gridding, and ANUDEM (Topo to Raster in ArcGIS Pro). The first two methods were automated and implemented in Python using Whitebox-tools and Rasterio; the last method was manually implemented in ArcGIS Pro.

First, one DEM was generated for each of the automated methods using only the contour points. Then, another DEM using only the *elevation* points was generated for each of the automated methods. This was deemed possible after recognizing the expected inputs for the two methods required only a point layer.

For the more complex ANUDEM method, we provided all the information included in the example data: contours, elevation points, lakes, and rivers (streams). The raster output from this method extended further than the intended area of interest, so a mask was applied to extract only the region we were investigating. During this extraction, we made sure to store the output to the same directory where our other data was, in order to include this ANUDEM result in the quality assessment later.

An extra DEM was created as an attempt to approximate the ANUDEM method using the TIN method. This was achieved by merging the TIN raster with a stream-burned raster created from the river feature layer. The resulting layer yielded results comparable to the TIN gridded DEM, although still less accurate than the ANUDEM method. We only generated this DEM using the TIN grid DEM generated from contour points.

3.2 Quality assessment

The quality assessment of the DEMs was done using a combination of visual and quantitative analysis.

3.2.1 RMSE

Quantitative analysis involved calculating the Root Mean Square Error (RMSE) between each generated DEM and the original elevation points. The RMSE was calculated using the following formula:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(z_i - z_{pred})^2}$$

where (z_i) is the original elevation point value, (z_{pred}) is the predicted elevation from the DEM at that point's location, and (n) is the number of points. RMSE was calculated for all generated DEMs by comparing their predicted elevations to the known values at the original elevation point locations. A similar RMSE analysis using the contour points could have been performed but was omitted due to time constraints.

3.2.2 Hillshade

The first visual analysis involved generating hillshade rasters from each of the DEMs. Hillshade is a technique used to create a shaded relief map, which can help visualize the terrain and identify features such as ridges, valleys, and slopes. Only the DEM was needed as input, with optional parameters the azimuth and altitude for the light source. In Whitebox-tools, default values are set to 315° and 30°, which were used for our analysis. Hillshade rasters were generated for each DEM in our automated workflow, and the results were visually compared in ArcGIS Pro.

3.2.3 Slope

Similar to hillshade, slope is another means of visualizing the terrain. Again, only the DEM is needed as input, though units and zfactor could be set if needed. Simlar to before, slope rasters were generated in our workflow, and visually analyzed in ArcGIS Pro.

3.2.4 Profile analysis

Further visual analysis involved plotting elevation profiles along an arbitrary transect line. Whitebox-tools was used to generate the profile analysis, storing the output for each DEM in an HTML file. To combine all profiles into a single Matplotlib plot, Beautiful Soup was used to parse these HTML files and extract the necessary data into arrays.

A transect line was created using the `create_line` function in Whitebox-tools, which creates a line between two points. The start and stop points for the transect were automatically selected from the extent of the input data. An inset was also defined, to ensure the transect line is not too close to the edges of the DEM, which could result in some artifacts in the profile.

4 Results

4.1 RMSE results

DEM Type	RMSE (m)	N Points
Natural Neighbor (Contour)	17.45	1212
Natural Neighbor (Points)	2867.71	1212
TIN Gridding (Contour)	17.47	1212

DEM Type	RMSE (m)	N Points
TIN Gridding (Points)	2867.71	1212
Stream Burn (Contour TIN based)	17.48	1212
ANUDEM (ArcGIS Pro)	15.00	1212

Table 1: RMSE results calculated against the original elevation points for each generated DEM. Values rounded to two decimal places.
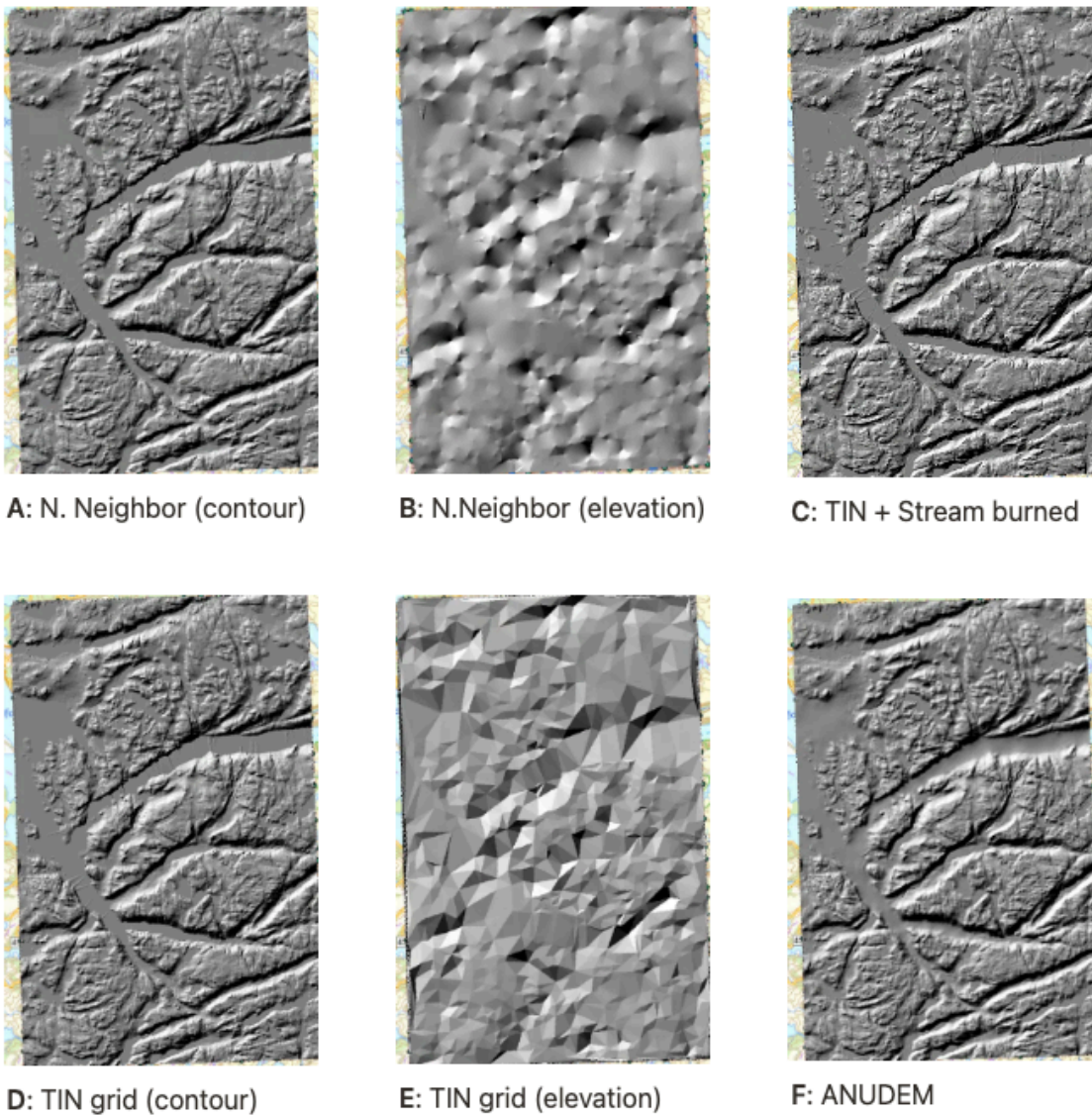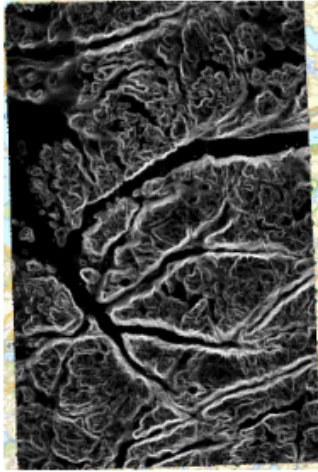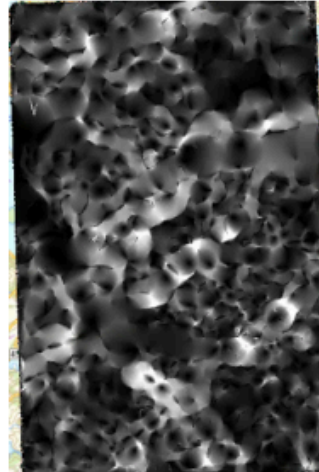
4.2 Hillshade results

A: N. Neighbor (contour)   **B: N.Neighbor (elevation)**   **C: TIN + Stream burned**

D: TIN grid (contour)   **E: TIN grid (elevation)**   **F: ANUDEM**

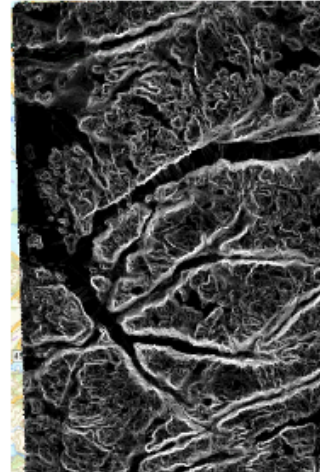**Figure 1**: Hillshade results for each generated DEM.

4.3   Slope results
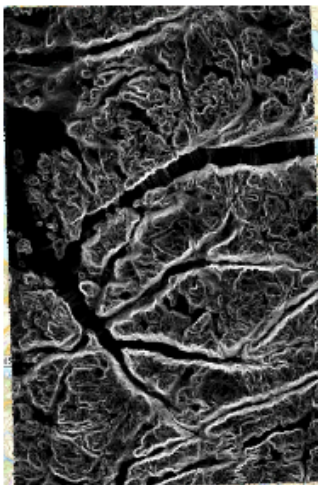

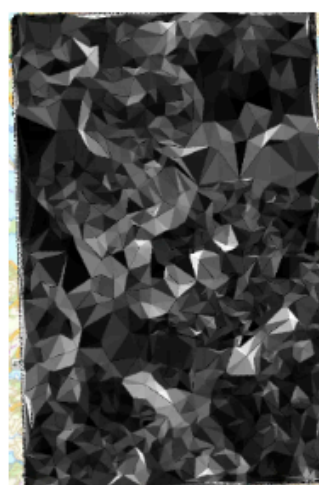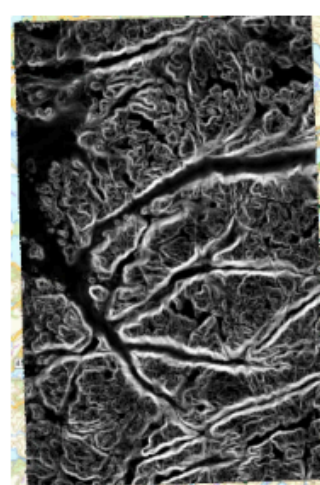A: N.Neighbor (contour)   B: N.Neighbor (elevation)   C: TIN + Stream burn

D: TIN grid (contour)   E: TIN grid (elevation)   F: ANUDEM

Figure 2: Slope results for each generated DEM.
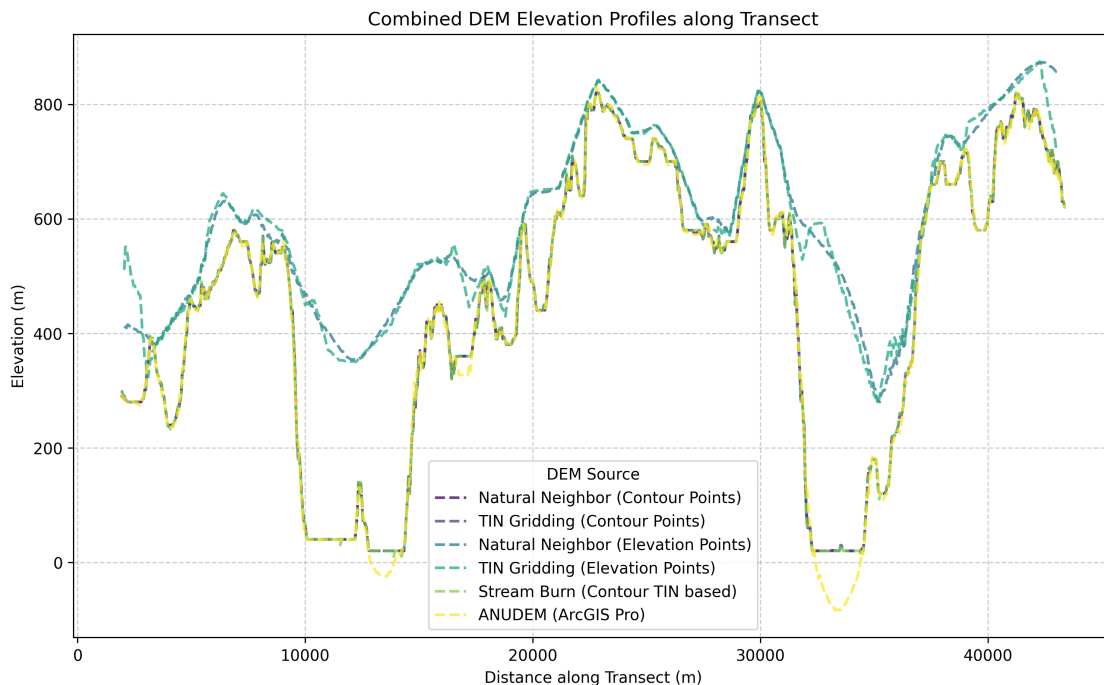
4.4 Profile results

Figure 3: Elevation profiles along the transect line for each generated DEM.

5 Discussion

The results from the RMSE analysis displayed in Table 1 show that the ANUDEM method performed best, with an RMSE of 15.00 m when evaluated against the original elevation points. However, when considering the profile analysis (Figure 3), it became apparent why this method outperformed the automated methods. Note the elevations *below zero* in Figure 3. Only the ANUDEM method estimated terrain elevations below sea level, resulting in a much more realistic DEM for this fjord. The other methods, Natural Neighbor and TIN Gridding, both estimated the terrain to be entirely above sea level, flattening out at zero elevation, since no contour data exists below sea level in our input dataset. Excluding the influence of these below-sea-level areas, our automated methods built from contour points seemed to match the terrain almost identically to the ANUDEM method, with RMSE values around 17.45 m.

As mentioned, we also wanted to compare DEMs generated solely from the sparse elevation points instead of the denser contour-derived points. The resulting RMSE values for those DEMs were significantly higher, around 2867.71 m. This large discrepancy is likely because the elevation points are not evenly distributed across the terrain and thus did not provide a good representation of the true surface when used alone for interpolation.

5.1 Hillshade and slope

Hillshade and slope results were presented in Figures 1 and 2, respectively. These visualizations further confirm that the recommended approach for generating DEMs in this context is to use contour points as the primary input.

The middle columns of both figures show the results from DEMs generated using only elevation points, which appear blurred, very coarse, and not very informative. This was expected, since minimal preprocessing was done on the elevation points, and their spatial distribution was uneven.

The stream-burning approach (applied to the TIN contour DEM) seems to yield the most fine-grained appearance in both figures. However, the standard automated methods (Natural Neighbor and TIN from contours) and ANUDEM also appear detailed and informative.

Upon closer inspection, it becomes apparent that the ANUDEM method produced the fewest visual artifacts, particularly in the fjord areas. Both of the simpler automated methods exhibited some obvious artifacts across the fjord, which were not present in the ANUDEM result. It should also be mentioned that judging by the slope results alone (Figure 2), the Natural Neighbor method (from contours) had relatively few artifacts and appeared comparable to the ANUDEM method.

6 Conclusion

In this report, we have explored different methods for generating DEMs using contour lines and elevation points. We have also performed a quality assessment of the generated DEMs using RMSE, hillshade, slope, and profile analysis.

The process was automated as much as possible using Python libraries to facilitate reproducibility and application to new datasets. However, limitations to the automation were the need to use ArcGIS Pro for the ANUDEM method and the initial manual preparation of the data as a geo-database. Everything else was possible to implement in Python, generating results comparable to the ANUDEM method, especially when using contour-derived points.

For the best overall results in this specific case, we recommend using the ANUDEM method, as it produced the most realistic DEM, effectively modeling elevations below sea level which were absent in the contour data. However, due to its proprietary nature, the ANUDEM method is not as accessible as the other methods. Simpler open-source methods like TIN Gridding and Natural Neighbor interpolation remain viable options, particularly when working with dense input data like contour-derived points.

When possible, we recommend using points derived from contour lines rather than sparse, stand-alone elevation points to generate DEMs. Contour layers typically provide more comprehensive spatial coverage, assuming they originate from a reliable source. In our case, using the N50 data from Kartverket, with a FAIR status of 0.95, we can be confident that the contour lines are a good representation of the terrain in this region.

7 References

- John Lindsay et. al, *Whitebox tools*, https://github.com/jblindsay/whitebox-tools
- Sean Gillies et. al, *Rasterio*, https://github.com/rasterio/rasterio
- GRASS Development Team, *GRASS GIS*, https://grass.osgeo.org
- ESRI, *ArcGIS Pro*, https://www.esri.com/en-us/arcgis/products/arcgis-pro/overview
- Kartverket, *N50 Kartdata*, https://kartkatalog.geonorge.no/metadata/n50-kartdata/ea192681-d039-42ec-b1bc-f3ce04c189ac?search=n50