ANALYSIS OF DIFFERENT MAP GEOREFERENCING METHODS

Alexandru TOACA, Bachelor student, Technical University of Civil Engineering Bucharest, Faculty of Geodesy, Romania
Doina VASILCA, Dr. Eng. Lecturer, Technical University of Civil Engineering Bucharest, Faculty of Geodesy, Romania, doina.vasilca@utcb.ro

Abstract*: Whenever one derives maps in other projections and at other scales compared to the base maps, the latter must be georeferenced. In the present article, we have studied multiple georeferencing methods employed through a series of professional applications such as AutoCAD Raster Design, AutoCAD Map 3D, ArcGIS Pro and QGIS, with the aim of making an educated choice with regards to their individual use cases. The residual errors obtained in the control points were analysed by applying different methods of transforming the Cartesian coordinates used in the georeferencing of a scanned map. For georeferencing, we used polynomial transformation with different order polynomials, a triangular method, adjustment transformation, spline transformation and similarity. For each method described above, the distances between the source and destination points were measured, reduced at the map scale and compared to the graphical precision of measuring a distance on a map. The present study centralizes the results obtained when using different georeferencing methods and can act as an aid for students or any party interested in georeferencing scanned images to create maps or perform different analyses based on georeferenced images.*

Keywords: *chi-squared test; control points; georeferencing methods; map; residual error*

1. Introduction

Making derived maps involves georeferencing an existing map, usually scanned, or a raster image. The points based upon which the scanned image is brought into the coordinate system of the new map are known as control points. Given their even distribution on the surface of the map, these points are easily identified (e.g. the grid nodes of intersecting meridians and parallels). Various transformations implemented using specialized software can be applied for georeferencing [1]:

- Affine transformation, producing translation, rotation, reflection, scale distortion and different scaling on *x* and *y*;
- Conformal transformation (similarity in ArcGIS, Helmert in QGIS), representing a particular type of affine transformation;
- Higher order polynomial transformations, ensuring greater accuracy;
- Spline transformation, useful when a perfect overlap is desired in the control points;
- Polynomial transformation by adjustment, providing global accuracy;
- Projective transformation, useful for georeferencing aerial photographs.

All of these methods of georeferencing were tested in a case study where we used a map of Italy made in the conic equal-area Albers map projection as a source to create two other maps: one in the equidistant conic map projection featuring the same shape of meridians and parallels as the original map, and another in an equidistant cylindrical map projection which changes the aspect of the geographical grid. As control points, we used the nodes of the grid of meridians and parallels. After performing the georeferencing, the residual errors were measured, namely the distances between the target points and the points where the source points actually arrived. Errors were compared to the graphic measurement error (i.e. 0.2-0.3 mm).

Statistical tests were then applied to the obtained values to determine which is the most suitable georeferencing method for the analysed cases.

2. Materials and Methods

The map of Italy used in the present study was scanned at a resolution of 400 dpi to increase the accuracy of the georeferencing process. The original map is split into two parts, as the original material is printed on the front and the back of the paper at a scale of 1:800 000. The map features a grid of meridians and parallels with a density of 30', though in our analysis we only used the network nodes with one degree density, due to the complexity of the georeferencing process.

The first stage in elaborating this study was to calculate the node coordinates (x, y). In the conic equidistant map projection, the representation was made on the surface of a tangent cone, the latitude of the tangent parallel $\varphi_k = 43^o$. The origin of the coordinates system was set in the apex of the cone. In the cylindrical equidistant map projection, the representation was made on a secant cylinder, with the latitude of the secant parallel $\varphi_k = 43^o$. The origin of the coordinate system was the intersection between the Equator and meridian $\lambda_o = 13^o$.

In both cases, the surface of the Earth was approximated using the WGS 84 ellipsoid. The same customizations were made for the projection coordinate systems selected in ArcGIS. Concerning QGIS, however, it bears mentioning that only the conic equidistant map projection was performed since the cylindrical equidistant map projection could not be made using the above customizations in order to obtain the same coordinates as those that were calculated.

AutoCAD with the Raster Tools plugin installed was the first application used. Any missing or incomplete part of the grid map was first reconstructed, then rasterized over the images. The maps were georeferenced using Rubber Sheet tool, which features several transformation options: triangular, first-order polynomial, second-order polynomial and third-order polynomial. This application has the option of increasing the polynomial transformations beyond the third order, but those only increase the residual errors which are the subject of this study.

The first of the above options, triangular transformation, creates a series of triangles based on the control points that are then transformed separately, thereby producing a more accurate result [2].

Polynomial transformations use a polynomial equation for transforming the control points from the local coordinate system into the project coordinate system using the least squares fitting method. The first-order polynomial is the affine transformation that only shifts, rotates, scales the image differently on the *Ox* and *Oy* axes and produces a linear distortion, while the second and third-order polynomials use more complex types of distortions to adjust the image to the control points [3].

The second application used was AutoCAD Map 3D, a different version of AutoCAD that is more specialized for GIS functions. The process took place as previously described, but the ADERSHEET command was used instead. The transformation method that AutoCAD Map 3D uses is affine and works similarly to the ALIGN command [4].

Beginning with the third application, specialized GIS Software such as ArcGIS allows one to create and assign a new map projection to the project based on an existing one. The raster data is then inserted into an ArcGIS project and georeferenced. The georeferencing methods used are akin to AutoCAD, but a few additional ones are added, and thus bring forth a greater variety of options. These methods include: zero-order polynomial, similarity, adjust, projective and spline. The zero-order polynomial shifts the image without rotating or scaling it. Similarity transformation is like the first-order polynomial transformation, but it preserves the shape of the original raster image. The adjust transformation is a combination of a polynomial transformation and a triangulated irregular network. The projective transformation warps lines so that they remain straight. Finally, the spline transformation changes the image in such way that the source points match the target points perfectly. Further research of this transformation reveals that residual errors in a spline transformation suggest that distortion is minimal only in the control points. Any area that is located far from the control points can present a higher degree of inaccuracy [5].

QGIS is the last application that was used in the present study. Its functions are similar to ArcGIS, but they are built differently. QGIS employs an integrated plugin called Georeferencer, which is more rudimentary than the Georeferencing function in ArcGIS. Furthermore, Georeferencer uses a separate window, rather than the main viewport for georeferencing. The transformations used in the georeferencing application of QGIS are almost the same as in ArcGIS, with linear transformation being the zero-order polynomial. The other three polynomial transformations are also present, as well as the spline and projective transformations, with only the Helmert transformation being the only difference. The latter is also named orthogonal transformation, as it contains four parameters (offset *x* and *y*, rotation and scale factor) [6]. By using each transformation, we could export the results and extract the residual errors, some of which are represented in pixels. Due to the known parameters of the scanned map, we were able to convert pixels into meters at a 1:1 scale such that the errors could also be represented in meters like our other extracted data. After the errors were measured or extracted from the data files, they were reduced to the desired scale. A statistical hypothesis test used to compared contingency tables with the expected results of a dataset, known as the chi-square test [7], was then applied to the data to determine which method of georeferencing is the most precise. Applying the chi-square test was done with the help of the following relations [8]:

$$\chi^2 = f\frac{S^2}{\sigma^2} \quad (1)$$

Where: $f = n - 1$ is the number of degrees of freedom;

n = total number of measurements;

σ = the root mean square error of a measurement;

S represents the selection standard deviation and is calculates as follows:

$$S = \pm\sqrt{\frac{[vv]}{(n-1)}} \quad (2)$$

v_i - deviation of the i measurement, x_i, from the mean

$$v_i = \bar{x} - x_i \quad (3)$$

$\bar{x}$ - mean value

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \quad (4)$$

The following hypotheses were considered:

the null hypothesis (H_o): $S^2 = \sigma^2$ (5)

the alternative hypothesis (H_1): $S^2 \neq \sigma^2$ (6)

If $\chi^2_{calc} < \chi^2_{f,\alpha}$, then the null hypothesis is accepted, otherwise the alternative hypothesis is accepted.

$\chi^2_{f,\alpha}$ is the critical statistic listed in the chi-square table, depending on the number of degrees of freedom, *f*, and the probability, α.

3. Results and Discussion

For georeferencing, 131 nodes of the meridian and parallel grid with a density of one degree were used as control points. The residual errors were measured in all of these points using AutoCAD Map 3D and AutoCAD Raster Design, then automatically calculated in ArcGIS and QGIS, respectively. The extreme values obtained for each georeferencing method are presented below in Table 1 for the equidistant conic map projection and the Table 2 for the equidistant cylindrical map projection.

To exemplify the distribution of residual errors in the network nodes, Figure 1 shows the diagrams obtained for the triangular, first-order polynomial and second-order polynomial transformations made in AutoCAD Raster Design.

Table 1: Extreme values of the residual errors obtained using the equidistant conic map projection

Software	Method	Minim value [m]	Maxim value [m]
AutoCAD Raster Design	Triangular	24.335	166.11
	1st order Polynomial	29.778	3506.02
	2nd order Polynomial	12.799	570.868
	3rd order Polynomial	8.855	425.391
	4th order Polynomial	32.971	391.09
AutoCAD Map 3D	Adersheet	1179.053	39807.406
ArcGIS	0 order Polynomial	39558.977	548288.034
	Adjust	0	274.013
	1st order Polynomial	103.206	3446.176
	2nd order Polynomial	11.260	545.667
	3rd order Polynomial	11.078	473.040
	Projective	48.708	2914.282
	Spline	0.000	0.000
	Similarity	119.465	9735.912
QGIS	Helmert	119.465	9735.912
	Linear	39331.893	549609.994
	1st order Polynomial	103.206	3446.176
	2nd order Polynomial	8.683	544.532
	3rd order Polynomial	11.961	472.501
	Projective	49.790	2887.159
	Spline	0.000	0.000

Table 2: Extreme values of the residual errors obtained using the equidistant cylindrical map projection

Software	Method	Minim value [m]	Maxim value [m]
AutoCAD Raster Design	Triangular	8.336	166.327
	1st order Polynomial	1208.126	23488.587
	2nd order Polynomial	26.297	570.110
	3rd order Polynomial	9.63	525.210
	4th order Polynomial	22.363	360.563
AutoCAD Map 3D	Adersheet	1685.804	47945.890
ArcGIS	0 order Polynomial	40603.555	543372.728
	Adjust	0.000	523.880
	1st order Polynomial	1354.752	23155.140
	2nd order Polynomial	29.429	1051.161
	3rd order Polynomial	4.590	523.880
	Projective	1002.226	17715.297
	Spline	0.000	0.000
	Similarity	1891.856	38194.960

Analysing the residual errors measured in the control points after georeferencing revealed that they can vary significantly depending on the transformation method being used. Most of the time, a georeferenced map does not fit perfectly into the new coordinate system and in some cases the raster data does not align at all with the new coordinate system, thereby rendering that respective transformation inappropriate for usage. For the maps analysed in the present study, these were the zero- and first-order polynomial transformations. These two transformations proved useful only when the projected coordinate system was the same as in the raster data, given that only shifting, rotation and scaling were used instead of warping the raster image. Additionally, the transformations made by the functions ADDERSHEET (AutoCAD Map 3D), SIMILARITY (ArcGIS), and HELMERT (QGIS) led to high residual values, rendering them unsuitable for georeferencing the analysed map.

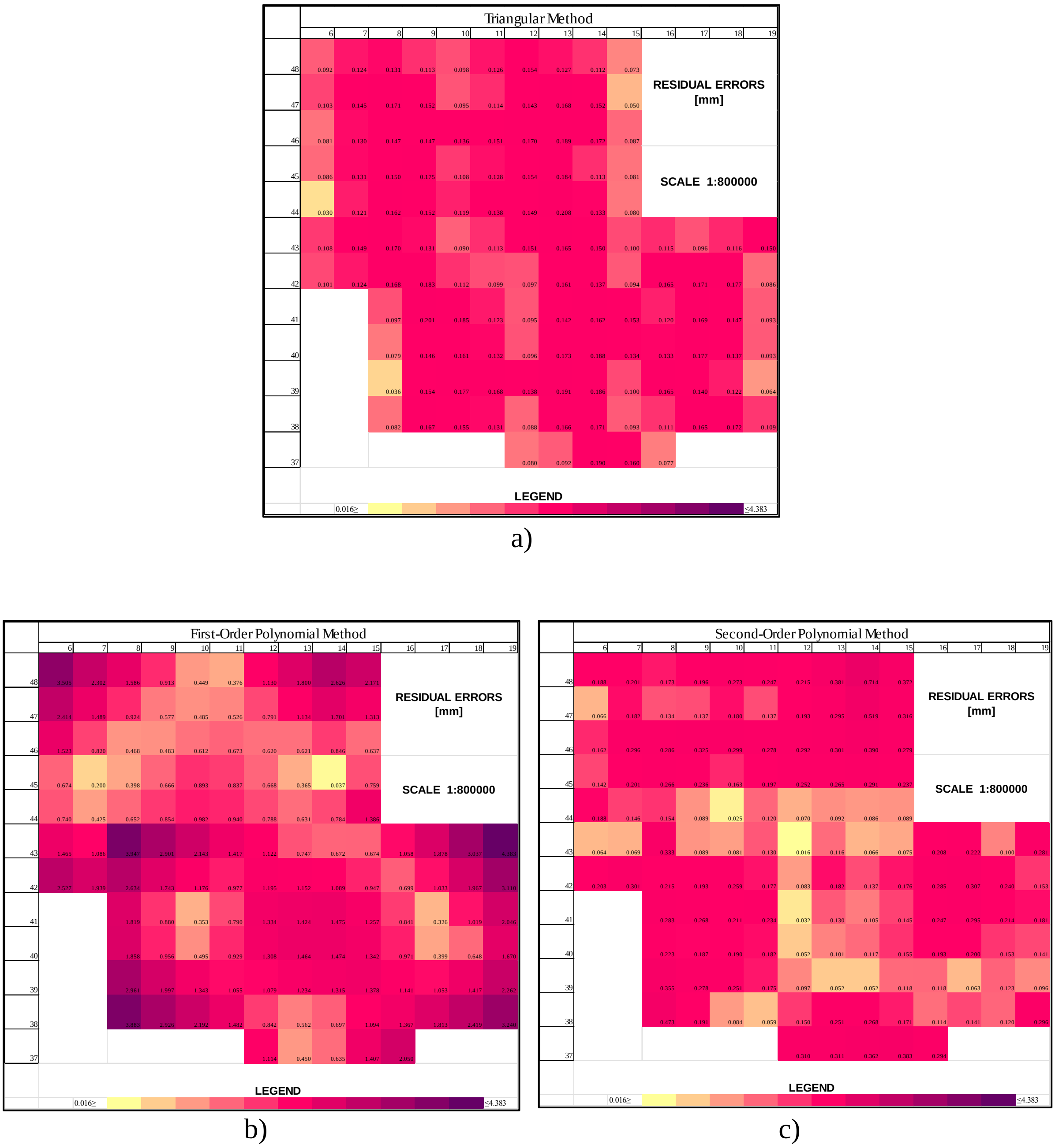

Figure 1: Residual errors in the grid nodes resulting from the triangular (a), first-order polynomial (b), and second-order polynomial (c) transformations in AutoCAD Raster Design.

If the projection of the georeferencing map and that of the new map are of the same type (e.g. azimuthal, conical, cylindrical, etc.), then the residual errors will be smaller and the raster image will suffer less distortions.

Even though Raster Design allows polynomial transformations higher than third-order, attempting to use a fifth-order polynomial will result in a severely distorted georeferencing that renders the image completely unusable and makes determining its residual errors impossible.

Due to the nature of our study and the large amount of collected data, the dataset was tested statistically. A chi-square test was conducted over its entirety to determine the accuracy of each transformation. In the present study, the contingency table was the table containing the residual errors for the employed transformations. The parameters used in the test were σ = 240 m, corresponding to the highest admitted error, namely 0.3 mm reduced at the scale map (1:800 000), and α = 0.05, representing a 95% level of confidence.

Table 3: The χ^2 calculated values for georeferencing using the equidistant conic map projection

Software	Method	χ^2_{calc}
AutoCAD Raster Design	Triangular	1.896
	1st order Polynomial	990.982
	2nd order Polynomial	16.214
	3rd order Polynomial	11.328
	4th order Polynomial	6.412
AutoCAD Map 3D	Adersheet	213631.956
ArcGIS	Polynomial 0	34226623.120
	Adjust	2.695
	1st order Polynomial	887.104
	2nd order Polynomial	15.144
	3rd order Polynomial	12.729
	Projective	624.389
	Spline	0.000
	Similarity	10382.977
QGIS	Helmert	10382.977
	Linear	34253198.360
	1st order Polynomial	887.104
	2nd order Polynomial	14.889
	3rd order Polynomial	12.700
	Projective	610.132
	Spline	0.000

Table 4: The χ^2 calculated values for georeferencing using the equidistant cylindrical map projection

Software	Method	χ^2_{calc}
AutoCAD Raster Design	Triangular	3.249
	1st order Polynomial	56590.776
	2nd order Polynomial	28.354
	3rd order Polynomial	13.720
	4th order Polynomial	6.741
AutoCAD Map 3D	Adersheet	143687.614
ArcGIS	Polynomial 0	32875657.500
	Adjust	8.444
	1st order Polynomial	58973.429
	2nd order Polynomial	104.764
	3rd order Polynomial	14.684
	Projective	29029.991
	Spline	0.000
	Similarity	125139.100

Comparing the values obtained for χ_{calc}^2 with the value extracted for the chi-square table [9], $\chi_{f,\alpha}^2 = 157.609$, it can be seen that the null hypothesis is accepted in the following cases for both map projections:

- Triangular, second, third, and forth-order polynomial transformations in AutoCAD Raster Design;
- Adjust, second, third-order polynomial and spline transformations in ArcGIS;
- Second and third-order polynomial transformations in QGIS.

In all other cases, the null hypothesis is rejected.

As the chi-square test shows, there is great variation in accuracy between the different transformations, with some of them having an extremely low precision, which is also reflected in the extremely high values that resulted from the test itself. With the exception of the spline transformation, which is a special case, the smallest values for χ^2 were obtained when using the triangular method of georeferencing for both map projections.

The spline transformation is a peculiar outlier when it comes to how precision operates. This particular transformation is very accurate in the control points, but the further a pixel is located from a control point, the larger the residential error becomes [5]. To illustrate this, a few points were chosen at various zones of the map to measure the residual georeferencing errors. The distances from the respective points to the closest meridians and parallels before and after georeferencing were measured on the map to determine the distortion (see Table 5). As expected, some of the differences were greater than the graphic precision of a measurement (i.e. 0.2-0.3 mm at the scale map), demonstrating that the respective points of the georeferenced image do not perfectly overlap the target points. Therefore, this type of transformation becomes useful mainly when overlapping the control points is important [5].

Table 5: Residual errors in points other than the control points resulting from georeferencing via the spline method

		Scale	
		1:1	1:800 000
City	Nearest meridian / parallel	Differences [m]	Differences [mm]
Livorno	43°	509.733	0.64
	10°	183.397	0.23
Rimini	44°	5.407	0.01
	12°	406.556	0.51
Venice	45°	206.001	0.26
	12°	118.158	0.15
Salzburg	47°	243.890	0.31
	13°	25.419	0.03

4. Conclusions

After analysing all of the methods implemented via Auto CAD Raster Design, AutoCAD Map 3D, ArcGIS Pro and QGIS for the georeferencing of a map of Italy extending at 12° latitude and 14° longitude, in order to create new maps in two different projections, but using the same scale as the original map, we determined that:

- Using all 131 nodes of the network of meridians and parallels as control points, with the exception for the spline method which is a special case, residual errors measuring less than 0.3 mm on the map scale were obtained in all control points when employing the following methods: triangular (AutoCAD), second- and third-order polynomial, as well as projective transformation (QGIS);

- When applying the spline transformation in ArcGIS and QGIS, residual errors equal to 0 are obtained in the control points, but in the points further away from them, these can exceed the imposed tolerance (i.e. Livorno and Rimini for the analyzed map);
- Analyzing the residual error values obtained using all georeferencing methods from a statistical point of view, it was found that the null hypothesis, $S^2 = \sigma^2$, was accepted for the following methods: spline, triangular, adjust, as well as second-, third- and fourth-order polynomial transformations for both map projections;
- The same georeferencing methods implemented in different applications led to comparable residual error values;

In conclusion, the present study can be useful to anyone wanting to create derived maps because it aids in making an educated choice with respect to the employed georeferencing method, thereby significantly reducing the time spent at this stage of the process.

5. References

1. *Imbroane, A. M., Sisteme informatice geogrefice, vol. I, Presa Universitara Clujeana, Romania, 2012, pp. 148-156;*
2. *AutoCAD Raster Design 2022 Help | Rubbersheeting an Image | Autodesk, consulted at August 12, 2024.*
3. *AutoCAD Raster Design 2024 Help | Glossary | Autodesk, consulted on August 12, 2024.*
4. *AutoCAD Map 3D 2024 Help | ADERSHEET (Rubber Sheet command) | Autodesk, consulted on August 12, 2024.*
5. *Overview of georeferencing—ArcGIS Pro | Documentation, consulted on August 25, 2024.*
6. *Helmert Transformation Definition | GIS Dictionary (esri.com), consulted on August 25, 2024.*
7. *Chi Square | Practical Applications of Statistics in the Social Sciences | University of Southampton, consulted on September 16, 2024*
8. *Danciu, V., Moldoveanu, C., Rus, T., Study concerning estimation of planimetric position and determination of the accuracy in case of extrapolation, vol. 17: Energy, Environmental and Structural Engineering Series, Turkey, 2013, pp. 232-239*
9. *https://www.medcalc.org/manual/chi-square-table.php, consulted on September 18, 2024.*