

Enhancing Spatial Accuracy of OpenStreetMap Data: A Geometric Approach

Maythm Al-Bakri 

Department of Surveying, College of Engineering, University of Baghdad, Baghdad, Iraq

Corresponding Author Email: m.albakri@coeng.uobaghdad.edu.iq



Copyright: ©2023 IIETA. This article is published by IIETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/mmep.100630>

ABSTRACT

Received: 22 April 2023

Revised: 27 August 2023

Accepted: 19 September 2023

Available online: 21 December 2023

Keywords:

OSM, volunteered geographic information, Baghdad, Iraq, polynomial transformation, affine transformation, conformal transformation, spatial accuracy

OpenStreetMap (OSM), recognised for its current and readily accessible spatial database, frequently serves regions lacking precise data at the necessary granularity. Global collaboration among OSM contributors presents challenges to data quality and uniformity, exacerbated by the sheer volume of input and indistinct data annotation protocols. This study presents a methodological improvement in the spatial accuracy of OSM datasets centred over Baghdad, Iraq, utilising data derived from OSM services and satellite imagery. An analytical focus was placed on two geometric correction methods: a two-dimensional polynomial affine transformation and a two-dimensional polynomial conformal transformation. The former involves twelve coefficients for adjustment, while the latter encompasses six. Analysis within the selected region exposed variances in positional accuracy, with distinctions evident between Easting (E) and Northing (N) coordinates. Empirical results indicated that the conformal transformation method reduced the Root Mean Square Error (RMSE) by 4.434 meters in the amended OSM data. Contrastingly, the affine transformation method exhibited a further reduction in total RMSE by 4.053 meters. The deployment of these proposed techniques substantiates a marked enhancement in the geometric fidelity of OSM data. The refined datasets have significant applications, extending to the representation of roadmaps, the analysis of traffic flow, and the facilitation of urban planning initiatives.

1. INTRODUCTION

The collection, uploading, and dissemination of geographic data generated by individuals, a process termed volunteered geographic information (VGI), has been facilitated by advancements in digital technology [1]. Recent years have witnessed the emergence of various VGI projects. Social media platforms such as Facebook, Twitter, and Flickr have become rich sources of geospatial data. Furthermore, collaborative mapping services like Wikimapia fall under the VGI category; however, OSM has achieved notable prominence [2].

Initiated in 2004 by Steve Coast in the UK, OSM has rapidly expanded on a global scale, underpinned by Web 2.0 technologies. These technologies denote the evolution of the internet into an interactive space characterized by user-generated content and enhanced usability, fostering more collaborative interactions among web clients, users, and content providers [3]. OSM encompasses a diverse array of data, including roads, building footprints, and land use maps. Initially, geographic data generation for OSM primarily involved the use of global positioning system (GPS) receivers to capture coordinates for upload to the database. With the advent of Bing map aerial imagery made available to OSM in November 2010, users were empowered to generate data through the digitization of aerial photographs as well [4].

The value of OSM data is underscored in contexts where governmental geographic information is prohibitive in cost or otherwise inaccessible to the public. Moreover, the financial burden of data updating often precludes national mapping agencies from performing this task with regularity [5]. Consequently, the OSM database has proven to be a pivotal resource for political and humanitarian endeavors, as well as for managing natural disasters [6].

In disaster-stricken locales, where immediate access to current and reliable geospatial information is critical, OSM data has been instrumental. The literature documents various cases where OSM mapping played a role in disaster response. Neis et al. [7] explored the application of OSM data in developing Emergency Routing Services to aid in the management of the Haiti earthquake aftermath. Westrope et al. [8] analyzed the veracity of crowd-sourced damage assessments in the Philippines following Typhoon Haiyan. Poiani et al. [9] conducted an exploratory case study on the organization of collaborative mapping efforts after the 2015 Nepal earthquake, detailing the mobilization of volunteers and the outcomes of their contributions to emergency planning. Lastly, Scholz et al. [10] provided an overview of the Red Cross and Red Crescent movement's history, current needs, and challenges, highlighting the role of OSM and digital volunteers in humanitarian missions, with a focus on disaster prevention as part of the Missing Maps project.

Despite the myriad advantages presented by OSM data, concerns regarding their quality persist, primarily due to the lack of geographic or geomatics expertise among Web 2.0 contributors [11]. Over the past decade, the maturation of OSM has been accompanied by scholarly efforts to evaluate its data quality. Jacobs [12] assessed the OSM roads' quality in the Ottawa-Gatineau region, employing a buffer method to gauge positional accuracy. It was revealed that a mere 76% of OSM roads aligned within 10 meters of the reference roads, with an even lower proportion, 70%, within a 5-meter range. Notably, the discrepancy between major and minor routes in terms of locational accuracy was found to be negligible. Al-Bakri and Sfoog [13] investigated the congruence of OSM road data with reference datasets in central Baghdad and Karbalaa city, Iraq, observing variable accuracy across different regions. Similarly, Zheng and Zheng [14] critiqued the completeness and positional accuracy of OSM data against Baidu datasets, quantifying the distribution of OSM data based on feature density. The study indicated that 71% of OSM data lacked the thoroughness of Baidu datasets, although, on average, 66% of OSM data was deemed accurate. Jasem and Al-Hamadani [15] scrutinized the positional accuracy of OSM road networks in Baghdad-Iraq, comparing them with authoritative data from the Mayorality of Baghdad (MB). Their findings pointed to substantial horizontal positional discrepancies between OSM datasets and the MB dataset, casting doubts on the viability of utilizing OSM data for updating Baghdad's road network database.

Although extensive research has been conducted on the applications and quality assessment of OSM data, there is a dearth of interest in the enhancement of its spatial accuracy. Addressing this gap, the present study concentrates on the application of geometric methods to improve the positional accuracy of OSM road network data. The methodology underpinning this research is grounded in two-dimensional transformation techniques. The outcomes of this study are intended to augment the knowledge base regarding the quality and reliability of refined OSM data across various applications.

Subsequent sections of this study are organized as follows: the methodology underpinning this study is delineated in the forthcoming section, including the two-dimensional polynomial affine and conformal transformations, along with an overview of the study area and data sources. Section 3 elucidates the development of a MATLAB program designed to compute unknown polynomial coefficients. Section 4 presents the results of the improvements made to OSM data, while the final section offers concluding remarks.

2. METHODOLOGY

2.1 Coordinate transformation in two-dimensional space

In the realm of mapping and surveying, coordinate transformations serve a pivotal role in converting spatial data between disparate coordinate systems [16]. Such transformations can be manifested through map projections, which transmute spatial coordinates from a spherical or spheroidal shape to a planar surface with rectangular (Cartesian) coordinates. Additionally, two-dimensional transformations are utilized to alter point coordinates within one rectangular system (x, y) to another (X, Y) [17].

The effects of transformations on a point set that defines a two-dimensional polygon, or a three-dimensional object,

range from simple positional and orientational adjustments, without alterations to shape or size, to uniform scaling that maintains the shape, and ultimately to transformations that modify both size and shape to varying nonlinear extents. For the purposes of this study, two transformation methods were employed to refine the OSM coordinates: the two-dimensional polynomial affine and conformal transformation. These methods are typically invoked when the coordinate systems in question exhibit non-uniformity in scale and orientation, thus necessitating the reduction of distortion within the OSM map data.

2.2 Two-dimensional polynomial affine transformation

The two-dimensional polynomial affine transformation model is applicable for the transposition of coordinates from one system (u, v) to another (x, y) . This model can be articulated as follows [18]:

$$\begin{aligned} x &= P(u, v) = \sum_{m=0}^p \sum_{n=0}^q c_{m,n} u^m v^n \\ &= c_{00} u^0 v^0 + c_{01} u^0 v^1 + c_{02} u^0 v^2 + c_{03} u^0 v^3 + \dots \\ &\quad + c_{10} u^1 v^0 + c_{11} u^1 v^1 + c_{12} u^1 v^2 + c_{13} u^1 v^3 + \dots \\ &\quad + c_{20} u^2 v^0 + c_{21} u^2 v^1 + c_{22} u^2 v^2 + c_{23} u^2 v^3 + \dots \end{aligned} \quad (1)$$

$$\begin{aligned} y &= P(u, v) = \sum_{m=0}^p \sum_{n=0}^q d_{m,n} u^m v^n \\ &= d_{00} u^0 v^0 + d_{01} u^0 v^1 + d_{02} u^0 v^2 + d_{03} u^0 v^3 + \dots \\ &\quad + d_{10} u^1 v^0 + d_{11} u^1 v^1 + d_{12} u^1 v^2 + d_{13} u^1 v^3 + \dots \\ &\quad + d_{20} u^2 v^0 + d_{21} u^2 v^1 + d_{22} u^2 v^2 + d_{23} u^2 v^3 + \dots \end{aligned} \quad (2)$$

Given that u^0 and v^0 are set to 1, this simplification reduces the number of parameters and sorts them in ascending order, i.e., first-order terms include u or v , second-order terms u^2 , v^2 , or uv , and so forth. Thus, a polynomial transformation can be formulated as:

$$x = c_0 + c_1 u + c_2 v + c_3 uv + c_4 u^2 + c_5 v^2 + c_6 u^2 v + \dots \quad (3)$$

$$y = d_0 + d_1 u + d_2 v + d_3 uv + d_4 u^2 + d_5 v^2 + d_6 u^2 v + \dots \quad (4)$$

where, $(c_0:c_6)$ and $(d_0:d_6)$ represent the unknown coefficients of the polynomial. These equations can also be delineated through matrix representations. By resolving the system matrix that embodies the polynomial, the unknown parameters c_i and d_i are discerned [19].

$$\begin{bmatrix} x_1 \\ y_1 \\ x_2 \\ y_2 \\ \vdots \\ x_n \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & u_1 & v_1 & u_1 v_1 & u_1^2 & v_1^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & u_1 & v_1 & u_1 v_1 & u_1^2 & v_1^2 \\ 1 & u_2 & v_2 & u_2 v_2 & u_2^2 & v_2^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & u_2 & v_2 & u_2 v_2 & u_2^2 & v_2^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & u_n & v_n & u_n v_n & u_n^2 & v_n^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & u_n & v_n & u_n v_n & u_n^2 & v_n^2 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \\ d_0 \\ d_1 \\ d_2 \\ d_3 \\ d_4 \\ d_5 \end{bmatrix} \quad (5)$$

To ascertain the coefficients, the equation and associated matrices can be rendered as:

$$AX=L \quad (6)$$

$$X=(At A)^{-1}(A^t L) \quad (7)$$

The dimensions and condition of the matrix system are contingent upon the number of ground control point (GCP) pairings selected and their spatial distribution within the study area. A minimum of six non-collinear GCP pairs must be chosen to establish the polynomial affine transformation framework. The selection of six GCP pairings yields a 12×12 square matrix that encodes the transformation, enabling the resolution of the polynomial equation system and the determination of the transformation parameters.

Upon the computation of the transformation parameters, the derived polynomial equations can be utilized to convert the coordinates of points from the initial system to the subsequent system.

2.3 Two-dimensional polynomial conformal transformation

In adherence to the Gauss theorem of conformal mapping and relevant complex polynomials, conformal transformations that exceed the initial power have been produced, as detailed in the analysis that follows [18]:

$$x = A_0 + A_1u - B_1v + A_2(u^2 - v^2) - B_2(2uv) + A_3(u^3 - 3uv^2) - B_3(3u^2v - v^3) \quad (8)$$

$$y = B_0 + B_1u + A_1v + B_2(u^2 - v^2) + A_2(2uv) + B_3(u^3 - 3uv^2) + A_3(3u^2v - v^3) \quad (9)$$

The partial derivatives of the above equations are:

$$\frac{\partial x}{\partial u} = A_1 + A_2(2u) - B_2(2v) + A_3(3u^2 - 3v^2) - B_3(6uv)$$

$$\frac{\partial x}{\partial v} = -B_1 - A_2(2v) - B_2(2u) - A_3(6uv) - B_3(3u^2 - 3v^3)$$

$$\frac{\partial y}{\partial u} = B_1 + B_2(2u) + A_2(2v) + B_3(3u^2 - 3v^2) + A_3(6uv)$$

$$\frac{\partial y}{\partial v} = A_1 - B_2(2v) + A_2(2u) - B_3(6uv) + A_3(3u^2 - 3v^2)$$

From these foundations, the second-order two-dimensional polynomial conformal transformation is derived, presented as:

$$x = A_0 + A_1u - B_1v + A_2(u^2 - v^2) - B_2(2uv) \quad (10)$$

$$y = B_0 + B_1u + A_1v + B_2(u^2 - v^2) + A_2(2uv) \quad (11)$$

where, $(A_0:A_2)$ and $(B_0:B_2)$ represent the polynomial's unknown coefficients. These transformations are capable of being expressed in matrix form as:

$$\begin{bmatrix} x_1 \\ y_1 \\ x_2 \\ y_2 \\ \vdots \\ x_n \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & u_1 & -v_1 & (u_1^2 - v_1^2) & -(2u_1v_1) & 0 \\ 0 & v_1 & u_1 & (2u_1v_1) & (u_1^2 - v_1^2) & 1 \\ 1 & u_2 & -v_2 & (u_2^2 - v_2^2) & -(2u_2v_2) & 0 \\ 0 & v_2 & u_2 & (2u_2v_2) & (u_2^2 - v_2^2) & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & u_n & -v_n & (u_n^2 - v_n^2) & -(2u_nv_n) & 0 \\ 0 & v_n & u_n & (2u_nv_n) & (u_n^2 - v_n^2) & 1 \end{bmatrix} \begin{bmatrix} A_0 \\ A_1 \\ B_1 \\ A_2 \\ B_2 \\ B_0 \end{bmatrix} \quad (12)$$

To determine the unknown coefficient matrix, Eq. (7) is employed, thereafter utilizing Eqs. (10) and (11) to effectuate the transformation of coordinates through the two-dimensional polynomial conformal method.

2.4 Study area and data sources

The area selected for this study is the Baghdad region of Iraq, depicted in Figure 1. As Iraq's central hub for economic, educational, cultural, and business activities, Baghdad is situated on both the Tigris River's banks and spans an area exceeding 4500 km^2 [20]. It lies longitudinally between $32^\circ 48' 00''$ E and $33^\circ 46' 00''$ E and latitudinally from $43^\circ 51' 00''$ N to $44^\circ 56' 00''$ N.

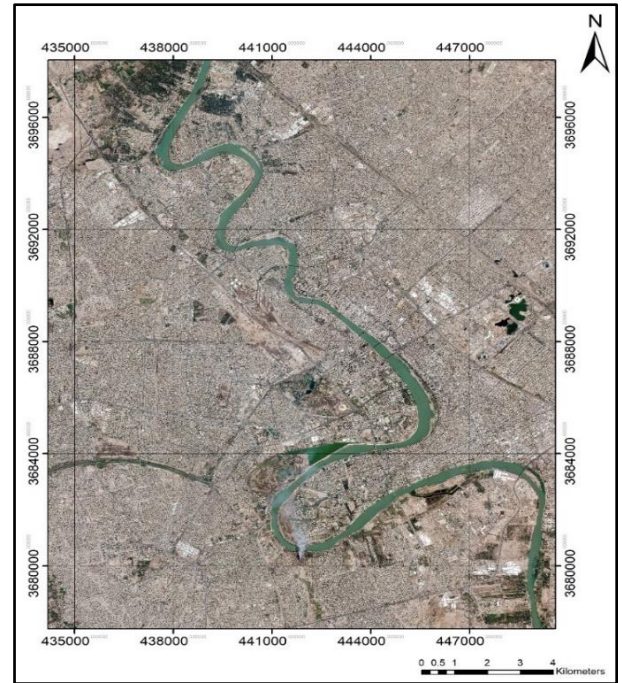


Figure 1. Study area

For the comparative study, both OSM and official reference datasets were collated for Baghdad. OSM data, sourced in May 2022, was obtained from the service provider Geofabrik, which offers data in various formats, including .osc.gz, .osm.pbf, .osm, .shp, .osh.pbf, and .poly. For this study, the ESRI shape-file (.shp) format was utilized.

The OSM line vector data encompassed all linear features such as boundaries, roads, railways, and certain aerial transport infrastructure. Given the primacy and focus of road features in OSM projects, only these were assessed in the current study. Non-road features were filtered out, retaining only those classified under the 'highway' key. Typically represented by a centerline, the OSM road data are contributed by various users and data sources.

Reference data were derived from the digitization of WorldView-3 satellite imagery, boasting a resolution of 0.3 meters. The digitizing process yielded a series of files encapsulating the digitized lines or points, complete with georeferenced positions and other necessary details. For the enhancement process of OSM data, road intersections served as point features. A total of 115 points, including control and check points, were employed for the study area. Control points are those used in model processing for referencing, while checkpoints serve to verify the accuracy of processed data by

comparison with known positions. These points were strategically positioned, both around and within the perimeter of the study area, with sufficient spacing to preclude ambiguity. Checkpoints were distributed uniformly and randomly, each clearly demarcated for easy identification on the map. Figure 2 illustrates the distribution and location of these tested points.

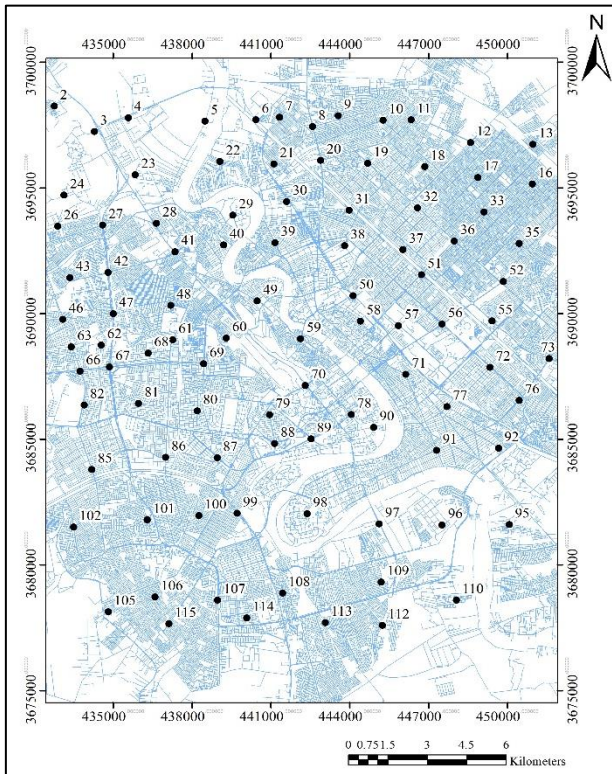


Figure 2. Distribution of the tested points

3. CODE DEVELOPMENT FOR OSM DATA IMPROVEMENT

The computation of unknown polynomial coefficients requires extensive numerical operations, often necessitating the use of a computational programming language. In this study, MATLAB, a specialized software application known for its problem-solving and graphical representation capabilities, was utilized [21]. MATLAB is commonly operated interactively, allowing the user to input expressions and receive immediate results. Additionally, scripts and functions, which are sequences of commands executed in order, can be composed in MATLAB [22].

A MATLAB script was developed to resolve Eqs. (5) and (12), which are central to the calculation of the unknown polynomial coefficients and the coordinates of check points. Prior to employing the MATLAB-based methodology, two preprocessing steps were conducted. Initially, the E and N coordinates of both the OSM data and the corresponding reference data points were extracted using ArcGIS software. These data were then projected onto the UTM zone 38 projection with the WGS84 coordinate system and exported in .xlsx format.

The programme's functionality included the loading and processing of the coordinates file for both reference and OSM datasets. Subsequently, the design (A) and observation (L) matrices were constructed. This was accomplished by applying the specialized equations to calculate the matrix

elements, with the code incorporating a for-end loop for equation application. Following this, the matrix representing the unknown polynomial coefficients (X) was computed. The final procedural step entailed the calculation of transformed coordinates for each check point, alongside the execution of statistical analyses. The results derived from the programme were saved in either .xlsx or .txt formats. Figure 3 illustrates the programme's flowchart.

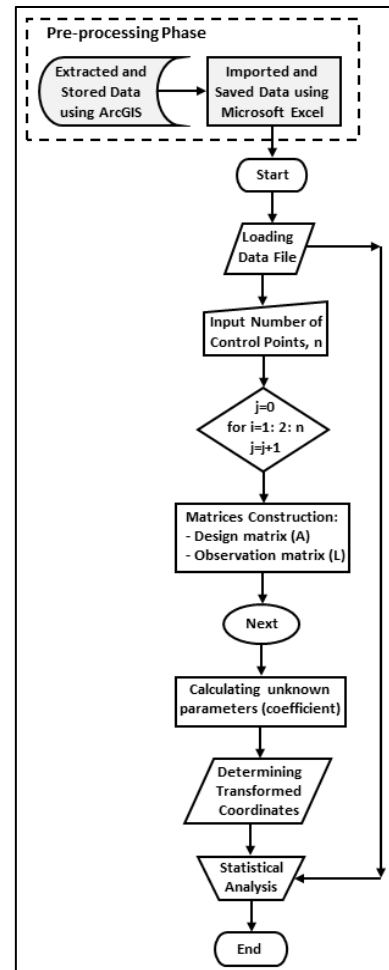


Figure 3. Flowchart of the designed programme

4. RESULTS AND DISCUSSION

Deformation errors inherent in OSM data were quantified using the derived transformation coefficients, as detailed in Tables 1 and 2. These coefficients illustrate the conversion process for enhancing OSM data.

Table 1 presents the coefficients for the two-dimensional polynomial conformal transformation, whereas Table 2 lists those for the two-dimensional polynomial affine transformation.

Table 1. Coefficients of the two-dimensional polynomial conformal transformation

Coefficients	Values
A_0	353530.4125
A_1	1.019727409
B_1	-0.193936948
A_2	-2.62E-08
B_2	4.70E-10
B_0	-79312.59952

Table 2. Coefficients of the two-dimensional polynomial affine transformation

Coefficients	Values
c ₀	327626.8478
c ₁	1.033811211
c ₂	-0.181373227
c ₃	-8.23E-09
c ₄	-3.80E-09
c ₅	2.50E-08
d ₀	-839016.636
d ₁	0.025655291
d ₂	1.452045596
d ₃	-1.69E-08
d ₄	4.13E-08
d ₅	-6.03E-08

The efficacy of the proposed transformations was assessed by calculating the RMSE, standard deviation (SD), and mean error for the study area, both pre- and post-application of the methodology, as shown in Table 3. The SD and RMSE for the *E* components were determined as follows: prior to methodology application (SD: 2.041m; RMSE: 3.252m), post-application of the two-dimensional polynomial conformal transformation (SD: 1.894m; RMSE: 2.853m), and post-application of the two-dimensional polynomial affine transformation (SD: 1.082m; RMSE: 1.732m). Similarly, for the *N* components, the corresponding values were (SD:

1.959m; RMSE: 5.107m), (SD: 2.218m; RMSE: 3.395m), and (SD: 1.836m; RMSE: 3.229m), respectively.

Discrepancies between the reference points and OSM coincident points for *E* ranged from 0.080 to 8.572 m before the application of the proposed methodology, from 0.054 to 7.589 m after the application of the two-dimensional polynomial conformal transformation, and from 0.042 to 3.922 m following the two-dimensional polynomial affine transformation. Variances in *N* were from 0.013 to 9.110m; 0.037 to 9.033m; and 0.006 to 8.874m for the data before methodology application, after conformal transformation, and after affine transformation, respectively.

Radial shifts in OSM data points, encompassing errors in both *E* and *N* coordinates, were evaluated and are presented in Table 3. Prior to the application of the proposed methods, the range of errors spanned 1.786 to 12.238m. Post-application, the ranges were reduced to 0.278 to 9.120m and 0.752 to 8.882m for the six-parameter and twelve-parameter transformation methods, respectively. The SD and RMSE of the radial errors were as follows: before the application of the methods (SD: 1.800m; RMSE: 6.054m), after the application of the two-dimensional polynomial conformal transformation (SD: 2.330m; RMSE: 4.434m), and following the two-dimensional polynomial affine transformation (SD: 1.633m; RMSE: 4.053m).

Table 3 details the coordinate discrepancies between the reference points and the OSM data points.

Table 3. Coordinate discrepancies between the reference and OSM points

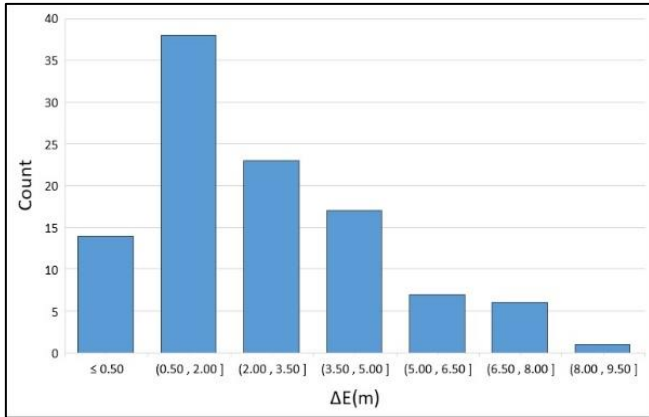
	ΔE (m)			ΔN (m)		Shift (m)			
	Before applying the proposed methodology	After applying two-dimensional polynomial conformal transformation	After applying two-dimensional polynomial affine transformation	Before applying the proposed methodology	After applying two-dimensional polynomial conformal transformation	After applying two-dimensional polynomial affine transformation	Before applying the proposed methodology	After applying two-dimensional polynomial conformal transformation	After applying two-dimensional polynomial affine transformation
Min	0.080	0.054	0.042	0.013	0.037	0.006	1.786	0.278	0.752
Max	8.572	7.589	3.922	9.110	9.033	8.874	12.238	9.120	8.882
Mean	2.531	2.133	2.197	4.716	2.569	2.657	5.780	3.772	3.710
SD	2.041	1.894	1.082	1.959	2.218	1.836	1.800	2.330	1.633
RMSE	3.252	2.853	1.732	5.107	3.395	3.229	6.054	4.434	4.053

Bar charts representing the *E* coordinate discrepancies (ΔE) for each approach were compiled and are depicted in Figure 4. The charts display the frequency of discrepancies along the horizontal axis against the magnitude of these discrepancies on the vertical axis. The *E* discrepancies of the tested data, post-application of the transformation methods, indicated improvements, particularly for discrepancies below 0.5m and those in the range of 3.5 to 8.0m. These results suggest that the applied methods have successfully reduced the variations in ΔE , thereby diminishing the magnitude of these differences.

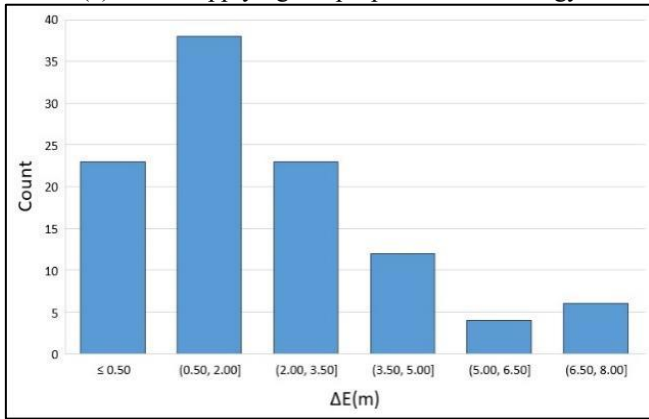
Variations in *N* discrepancies (ΔN) across OSM datasets are graphically depicted in Figure 5. Initially, discrepancies within the 0.5 to 3.5m range were relatively fewer compared to the more substantial figures observed within the 3.5 to 9.5m range. Following the application of the proposed methodology, a notable concentration of discrepancies was found within the 0.5 to 3.5m range, with a marked reduction in values from 3.5 to 9.5m. This pattern suggests that the methodology implemented tend to yield ΔN that are lower and more consistent, contrasting with the higher discrepancies noted

prior to methodology application.

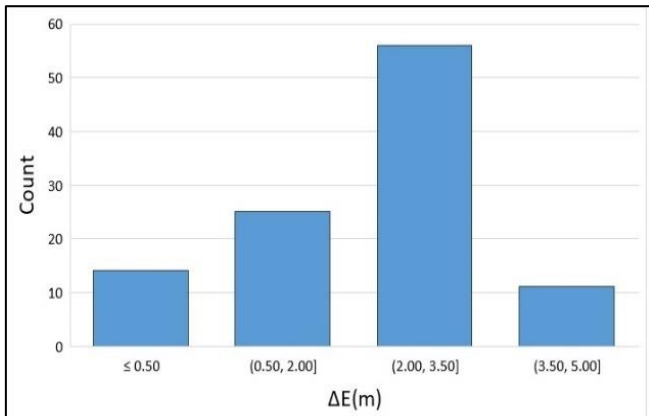
Figure 6 illustrates the radial shifts in OSM data, highlighting the impact of the transformation methods on positional accuracy. The application of a six-parameter transformation method showed an enhancement in accuracy, with a further increase observed following the twelve-parameter transformation. The largest range of errors was recorded in the datasets prior to methodology application, with shifts exceeding 12m. A discernible improvement was apparent with the six-parameter method, indicated by an increase in discrepancies less than 2m, as well as those between 2 to 4m, and a significant reduction in errors beyond 4m. The twelve-parameter method further amplified this trend, with a substantial number of shift errors below 4m, and a notable decrease in shift errors ranging from 4 to 10m. These findings underscore the effectiveness of two-dimensional transformation methods in reducing the RMSE of OSM data. The preference for these methods is attributed to their capacity to retain the true shape post-transformation.



(a) Before applying the proposed methodology

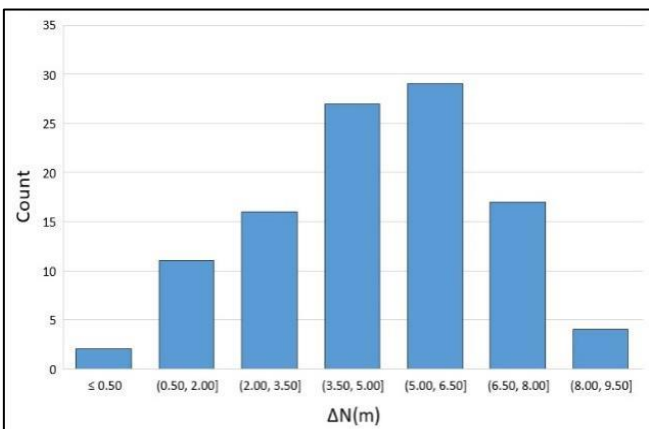


(b) After applying the two-dimensional polynomial conformal transformation

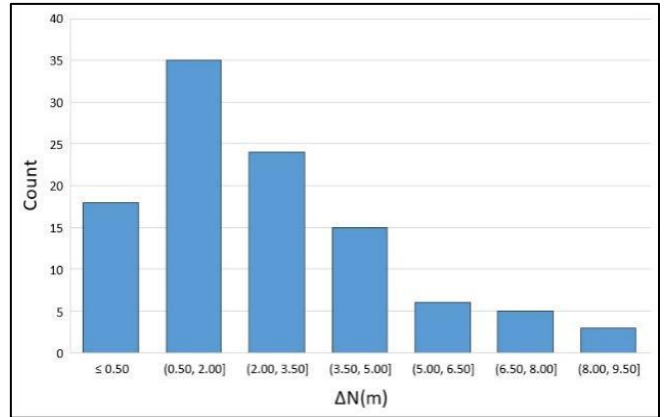


(c) After applying the two-dimensional polynomial affine transformation

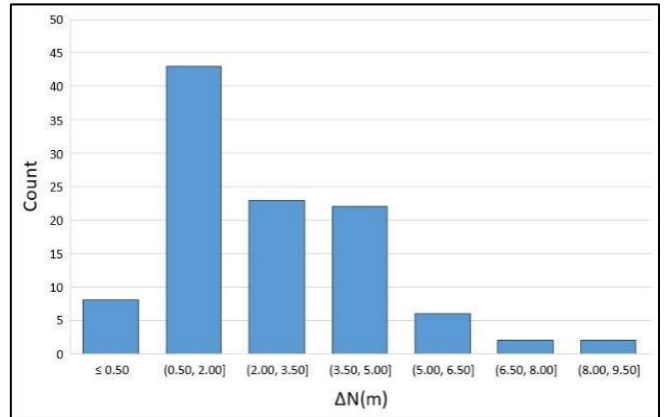
Figure 4. Bar charts of ΔE



(a) Before applying the proposed methodology

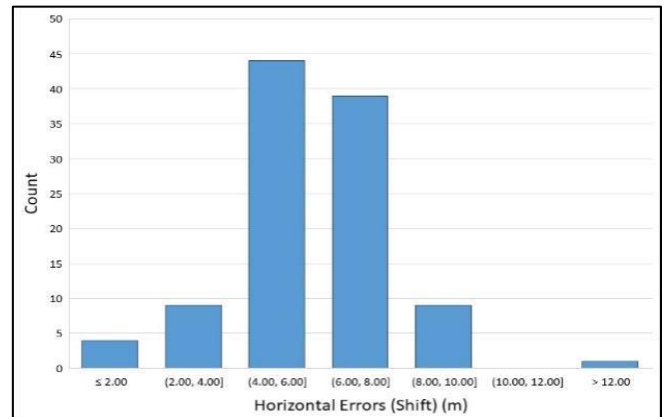


(b) After applying the two-dimensional polynomial conformal transformation

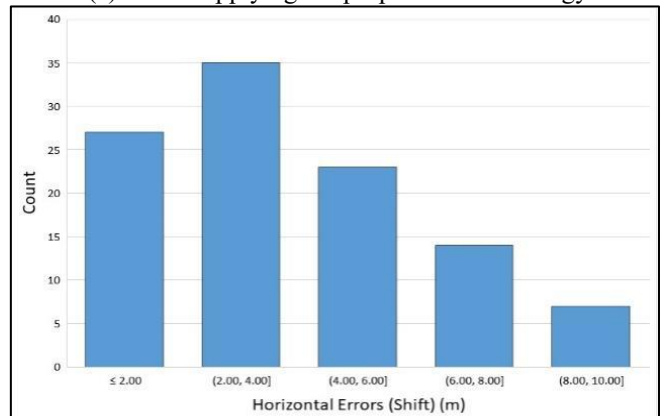


(c) After applying the two-dimensional polynomial affine transformation

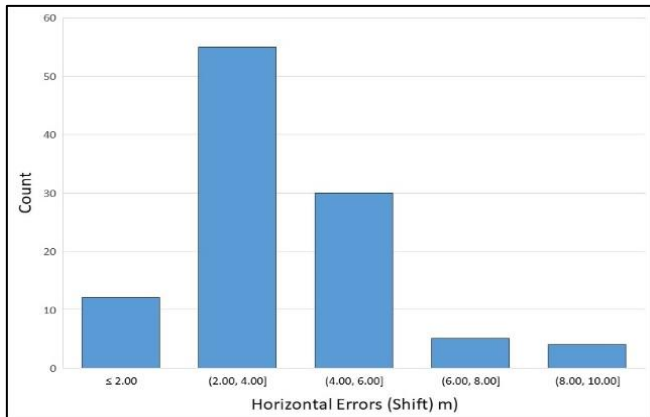
Figure 5. Bar charts of ΔN



(a) Before applying the proposed methodology



(b) After applying the two-dimensional polynomial conformal transformation



(c) After applying the two-dimensional polynomial affine transformation

Figure 6. Bar charts of the radial errors/shifts

5. CONCLUSIONS

This study has been directed at assessing and ameliorating the positional accuracy of OSM road network data, with a focus on the central area of Baghdad, Iraq. Two geometric enhancement methods were examined: the two-dimensional polynomial affine transformation, incorporating twelve unknown parameters, and the two-dimensional polynomial conformal transformation, involving six unknown parameters, using reference data for spatial accuracy improvement.

Statistical analysis of residuals, employing diverse sets of checkpoints and control points, facilitated the evaluation. This analysis included a comparison of RMSE values derived from the differences between the enhanced and original OSM datasets. The results revealed that the six-parameter transformation method attained a moderate increase in spatial accuracy, while the twelve-parameter method demonstrated greater efficacy. This was evidenced by a reduction in the RMSE of linear errors from 6.054m in the pre-enhancement phase to 4.434 and 4.053m post-application of the six-parameter and twelve-parameter methods, respectively, signifying improvements of 28% and 33%. Improvements were also observed in the *E* and *N* components, with enhancements of 12% and 47% for *E*, and 34% and 37% for *N*, respectively, when applying the two models. These enhancements are relative to the baseline accuracy of the original OSM data. While the twelve-parameter model offered substantial accuracy enhancements, it also necessitated a greater number of control points, implicating cost-effectiveness considerations.

Future research may extend the proposed methodology to enhance the spatial accuracy of other open-source datasets, such as Wikimapia, Google Earth, and Yahoo Maps. Further studies could investigate the correlation between feature type and the degree of spatial accuracy improvement in OSM data, potentially exploring elements beyond road networks, such as buildings, green spaces, railways, and water bodies.

REFERENCES

- [1] Al-Bakri, M. (2015). Ten years of OSM project: Have we addressed data quality appropriately? – Review paper. *Journal of Engineering*, 21(4): 158-175. <https://doi.org/10.31026/j.eng.2015.04.09>
- [2] Mooney, P., Corcoran, P., Ciepluch, B. (2013). The potential for using volunteered geographic information in pervasive health computing applications. *Journal of Ambient Intelligence and Humanized Computing*, 4: 731-745. <https://doi.org/10.1007/s12652-012-0149-4>
- [3] Sehra, S.S.; Singh, J., Rai, H.S. (2017). Using latent semantic analysis to identify research trends in OSM. *ISPRS International Journal of Geo-Information*, 6(7):195. <https://doi.org/10.3390/ijgi6070195>
- [4] Moradi, M., Roche, S., Mostafavi, M.A. (2022). Exploring five indicators for the quality of OSM road networks: A case study of Québec, Canada. *Geomatica*, 75(4): 178-208. <http://doi.org/10.1139/geomat-2021-0012>
- [5] Zhou, Q. (2018). Exploring the relationship between density and completeness of urban building data in OSM for quality estimation. *International Journal of Geographical Information Science*, 32(2): 257-281. <https://doi.org/10.1080/13658816.2017.1395883>
- [6] Herfort, B., Lautenbach, S., Porto de Albuquerque, J., Anderson, J., Zipf, A. (2021). The evolution of humanitarian mapping within the OSM community. *Scientific Reports*, 11(3037): 1-15. <https://doi.org/10.1038/s41598-021-82404-z>
- [7] Neis, P., Singler, P., Zipf, A. (2010). Collaborative mapping and emergency routing for disaster logistics - case studies from the Haiti earthquake and the UN portal for Afrika. In *Proceedings of the Geospatial Crossroads@ GI Forum*, Salzburg, Austria. <https://www.geog.uni-heidelberg.de/md/chemgeo/geog/gis/un-osm-emergency-routing.gi-forum2010.full.pdf>
- [8] Westrope, C., Banick, R., Levine, M. (2014). Groundtruthing OSM building damage assessment. *Procedia Engineering*, 78: 29-39. <https://doi.org/10.1016/j.proeng.2014.07.035>
- [9] Poiani, T., Rocha, R., Degrossi, L. Albuquerque, J. (2016). Potential of collaborative mapping for disaster relief: A case study of OSM in the Nepal earthquake 2015. In *Proceedings of the Annual Hawaii International Conference on System Sciences*, Koloa, HI, USA, pp. 188-197. <https://doi.org/10.1109/HICSS.2016.31>
- [10] Scholz, S., Knight, P., Eckle, M., Marx, S., Zipf, A. (2018). Volunteered geographic information for disaster risk reduction—The missing maps approach and its potential within the red cross and red crescent movement. *Remote Sensing*, 10(8): 1239. <http://doi.org/10.3390/rs10081239>
- [11] Jackson, S.P., Mullen, W., Agouris, P., Crooks, A., Croitoru, A., Stefanidis, A. (2013). Assessing completeness and spatial error of features in volunteered geographic information. *ISPRS International Journal of Geo-Information*, 2(2): 507-530. <https://doi.org/10.3390/ijgi2020507>
- [12] Jacobs, K. (2018). Quality assessment of volunteered geographic information: An investigation into the Ottawa-Gatineau OSM database. MSc. Dissertation, Carleton University, Ottawa, Ontario, Canada.
- [13] Al-Bakri, M., Sfoog, Z.Y. (2018). Matching assessment of road network objects of volunteered geographic information. *Association of Arab Universities Journal of Engineering Sciences*, 25 (1): 145-159.
- [14] Zheng, S., Zheng, J. (2014). Assessing the completeness and positional accuracy of OSM in China. In *Thematic*

- Cartography for the Society, Lecture Notes in Geoinformation and Cartography. Springer, Cham. https://doi.org/10.1007/978-3-319-08180-9_14
- [15] Jasem, M., Al-Hamadani, O. (2020). Positional accuracy assessment for updating authoritative geospatial datasets based on open source data and remotely sensed images. *Journal of Engineering*, 26(2): 70-84. <http://doi.org/10.31026/j.eng.2020.02.06>
- [16] Hussein, A.A., Mahmood, F.H. (2022). Determination local geoid heights using RTK-DGPS/leveling and transformation methods. *Iraqi Journal of Science*, 57(2C): 1604-1611. <https://ijs.uobaghdad.edu.iq/index.php/eijs/article/view/7237>.
- [17] Wolf, R., Bon A., Benjamin E. (2014). *Elements of Photogrammetry with Applications in GIS*. New York, USA: McGraw-Hill Education. <https://www.accessengineeringlibrary.com/content/book/9780071761123>.
- [18] Deakin, R.E. (2004). *Coordinate transformations in surveying and mapping*, Geospatial Science (2004). http://user.gs.rmit.edu.au/rod/files/publications/COTRA_N_1.pdf, accessed on Jan. 18, 2023.
- [19] Paniphi, N., Tripathy, S. (2002). Image registration using polynomial affine transformation. *Defence Science Journal*, 52(3): 253-259. <https://doi.org/10.14429/dsj.52.2180>
- [20] Ali, Z., Muhaimed, A. (2016). The study of temporal changes on land cover/land use prevailing in Baghdad governorate using RS & GIS. *Iraqi Journal of Agricultural Sciences*, 47(3): 846-855. <https://doi.org/10.36103/ijas.v47i3.576>
- [21] Abbas, H.K., Mohamad, H.J. (2021). Feature extraction in six blocks to detect and recognize English numbers. *Iraqi Journal of Science*, 62(10): 3790-3803. <https://doi.org/10.24996/ijs.2021.62.10.37>
- [22] Attaway, S. (2013). *MATLAB a Practical Introduction to Programming and Problem Solving*. USA: Butterworth-Heinemann.