A Visual Interface Design for Evaluating the Quality of Google Map Data for some Engineering Applications

Dr. Mouayed Y. Ahmed
Assistant Professor
Department of Surveying
College of Engineering
University of Baghdad

Dr. Maythm al-Bakri
Lecturer
Department of Surveying
College of Engineering
University of Baghdad
Email: m.m.m.s.albakri@gmail.com

Luma Layth Abedulridha
Department of Surveying
College of Engineering
University of Baghdad
Email: luma_laith90@yahoo.com

ABSTRACT

Today, there are large amounts of geospatial data available on the web such as Google Map (GM), OpenStreetMap (OSM), Flickr service, Wikimapia and others. All of these services called open source geospatial data. Geospatial data from different sources often has variable accuracy due to different data collection methods; therefore data accuracy may not meet the user requirement in varying organization. This paper aims to develop a tool to assess the quality of GM data by comparing it with formal data such as spatial data from Mayoralty of Baghdad (MB). This tool developed by Visual Basic language, and validated on two different study areas in Baghdad / Iraq (Al-Karada and Al- Kadhumiyah). The positional accuracy was assessed by adopting National Standard for Spatial Data Accuracy (NSSDA). The evaluation procedure also involved one and two-sample t-test to analyze and compare the accuracy of two study areas. The findings found that the NSSDA accuracy of case study one was 15.48 m, while it was 8.71 m for case study two. This indicated that the accuracy of the GM data is different from site to site. The results also showed that the difference on mean was 6.16 m, which indicated that there is a difference in GM accuracy in different areas. It was concluded that the GM data is inappropriate for engineering applications that require high accuracy, but may be appropriate for applications that need low accuracy such as the primarily surveying of engineering design projects, tourism and reconnaissance….etc.

Key words: data quality, positional accuracy, VGI, formal data, Google Map Maker.

تصميم برنامج لتقييم جودة بيانات Google map لبعض التطبيقات الهندسية

أ.م.د. مؤيد ياسين احمد
قسم هندسة المساحة
كلية الهندسة / جامعة بغداد

م.د. ميثم مطشر شرقي
قسم هندسة المساحة
كلية الهندسة / جامعة بغداد

لمى ليث عبد الرضا
قسم هندسة المساحة
كلية الهندسة / جامعة بغداد

الخلاصة

ظهرت في الأونة الاخيرة كميات كبيرة من البيانات الجغرافية المكانية المتاحة على شبكة الإنترنيت مثل خدمة Yahoo imagery, OpenStreetMap, Google Map.... وغيرها. في اغلب الأحيان تطلق تسمية (Open Source Data) على هذه الخدمات. تكون البيانات الجغرافية المنتجة من مصادر مختلفة عادةً بمستويات دقة متباينة بسبب اختلاف طرق جمع البيانات، وهذا بدوره قد لا يلبي متطلبات المستخدمين للتطبيقات المختلفة.

يهدف هذا البحث إلى تصميم وانتاج برنامج لتقييم دقة مواقع البيانات المنتجة من ((Google Map (GM) لأغراض التطبيقات الهندسية ، هذا البرنامج تم تطويره باستخدام لغة البرمجة (Visual Basic) من أجل مقارنة مدى قرب أو بعد البيانات المنتجة من (GM) مع تلك البيانات المنتجة من المصادر الرسمية مثل أمانة بغداد. يحتوي هذا البرنامج على ثلاثة مراحل رئيسة: إدخال البيانات، التحليل والأستنتاج، اخراج النتائج على شكل تقرير.

طُبق هذا البرنامج على منطقتي دراسة في العراق/ بغداد (الكرادة و الكاظمية) من أجل التحقق فيما إذا كانت بيانات GM تمتلك دقة متماثلة في مناطق دراسة مختلفة. لقد صمم البرنامج لتقييم الدقة الموضوعية (positional accuracy) بالاعتماد على معيار (National Standard for Spatial Data Accuracy (NSSDA وبالاعتماد على threshold value لقبول أو رفض النتائج. صممت الواجهة الرئيسة للبرنامج بحيث أذا أُريد تقييم positional accuracy لأي مجموعة بيانات من (مصدر رسمي وغير رسمي) فأن البرنامج يطلب إستدعاء البيانات ويسأل عن threshold value وحدود الثقة (confidence interval) و significance of level (p-value) من أجل التحليل الأحصائي. لقد أشارت النتائج أن دقة NSSDA لمنطقة الدراسة الأولى كانت 15.48 m، بينما لمنطقة الدراسة الثانية كانت 8.71m. هذا يدل على أن دقة خرائط GM متباينة من منطقة إلى أخرى ومن خلال تطبيق الأسلوب الإحصائي two-sample t-test وجد إن الفرق في الدقة لمنطقة الدراسة الثانية عن منطقة الدراسة الأولى يساوي 6.16 m ومن خلال ذلك ، تم الأستنتاج إلى ان دقة خرائط GM غير ملائمة للتطبيقات الهندسية عالية الدقة ولكن قد يمكن أن تستخدم للتطبيقات ذات المساحات الصغيرة و المقاييس الكبيرة وكذلك تحديد المواقع الأولية لمسارات الطرق والتي عادة ما يتم تحديدها قبل عملية التصميم والتي قد تحتاج دقة موضعية تتراوح مابين5 m -50 . كما يمكن استخدامها أيضا لأغراض السياحة والاستطلاع ....وغيرها.

1. INTRODUCTION

Recent development of geospatial data collection technologies and the growth of the World Wide Web (WWW) (e.g. web 2.0) for different applications have led to a massive increase in the amount of geospatial data on the Internet, **Cartwright, 2008**. The evolution of the Web 2.0 service enables users to produce and share, download, embed and add information from different online data sources. In literature, different definitions have been suggested to describe data on web. For instance, One of the first people to define geospatial data on web was **Goodchild, 2007**, who proposed Volunteered Geographic Information (VGI) to identify spatial data which is collected and distributed on Internet. **Turner, 2006,** defined this technology as 'Neogeography' which consists of a set of techniques and tools that fall outside the realm of traditional Geographic Information Systems GIS. The term 'Neogeography' was also used by **Flanagin and Metzger, 2008,** to refer to non-traditional GIS techniques that produce geography data without geographers. **Howe, 2006,** used the term 'crowdsourcing' to define geospatial data on web. Although crowdsourcing not specifically referring to geographical data. This definition is close to those of **Antoniou, et al., 2010,** used the term 'User-Generated Content' (UGC) to refer to various types of media content which are publicly produced available on web. The general idea of different descriptions of data on Internet was on how to use the Internet to create, share, and analyze geographic information via multiple computing devices/platforms (traditional desktops, iPads, or smart phones), **Haklay, et al., 2008**.

Today there are a wide range of geospatial data sources available on the Internet such as the Google Map service, the OpenStreetMap (OSM) project, the Flickr service, the interactive Wikimapia website, Yahoo imagery and others. Web mapping have continuously evolved with the time and all of these services have been called Collaborative Maps. All of this new information is open source geospatial data; therefore it is a legitimate topic for accuracy assessment. Accuracy assessment is a problem for information in general and geographical information in particular. Hence, it becomes a major issue with increased available data on the web. In this research Google Map data has been chosen in order to evaluate its positional accuracy for engineering applications. Different techniques and procedures were followed and applied to evaluate GM quality as will be illustrated in the following sections.

2. PREVIOUS RESEARCH ON VGI DATA QUALITY ASSESSMENT

In order to assess the quality of open source data to determine the appropriate usage for such geographical Information System (GIS) processing, it is necessary to identify different elements of spatial data quality, **Delavar, and Devillers, 2010**. These elements include: positional accuracy, attribute accuracy, temporal accuracy, logical consistency and completeness. In addition, there are other three elements of non-quantitative quality: purpose, usage, and lineage, as shown in **Fig 1**. Recently, with increasing geospatial data on the Internet, several researches have shown an increased interest in assessing the quality of open source data. For example, **Haklay, 2010,** examined positional and completeness accuracy of VGI data by comparing OSM data with Ordnance Survey (OS) reference dataset in London /UK. The buffering technique was adopted to assess the positional accuracy and the results indicated that there is a slight difference between OSM and OS datasets, while the completeness analysis indicated that there are omission and commission in OSM dataset.

Ather, 2009, assessed positional accuracy of OSM data. The analysis was performed by comparing the motorways data of OSM project with those formal OS Master Map Integrated Transport Network (ITN) layer. The methodology was essentially based on buffer analysis datasets. The results of this study found that the positional accuracy of OSM data is close to OS Master Map dataset. Further quality tests were also conducted in terms of the number of users per area and road name attribute completeness. Also it showed a positive correlation between road name attribute completeness and number of users per area. Another study on VGI data quality assessment was carried out by **Kounadi, 2009,** evaluated OSM data in Athens, Greece. The quality analysis was achieved on positional accuracy, the completeness and thematic accuracy of OSM road Network. The OSM data was compared with the Hellenic Military Geographical Service (HMGS) data which is the official cartographic service in Greece. The results found that the positional accuracy of OSM data is accepted when compared to HMGS data.

Analysis of VGI data quality was also achieved by **Ciepłuch, et al., 2010**. The comparison was among data from GM, OSM, and Bing Maps (BM) in Ireland. Towns were chosen for these comparisons for five case studies. The accuracy was evaluated under three main headings: completeness, currency of the spatial information and ground-truth positional accuracy. The results found that the OSM project has shown many positive and negative characteristics in terms of providing a comprehensive mapping resource in Ireland. On the other hand, deduces that the OSM and GM projects provide the update and current road configuration, Unlike BM which estimate the data in more than one year old. In another major study, **Zielstra, and Zipf, 2010,** investigated the completeness of OSM data in Germany by comparing it with the TeleAtlas data. This work extended the studies from England by **Haklay, 2010** and **Ather, 2009**. The results showed that the geospatial data has been continues growth to freely available compared to open source spatial data in the past few years. Furthermore, the results indicated that there is still a very strong heterogeneity of OSM data in a terms of completeness .The tests showed that the larger cities is more complete than the medium –sized cities which can be indicate that the people of large cities is more interested on VGI.

Hochmair, and Zielstra, 2012, used other examples of open source data which facilitate the sharing of VGI in form of geotagged images. This measurement was conducted of 211 Flicker and Panoramio images distributed throughout six urban areas in Germany by comparing the geotagged position of photos to the position from where the photos were most likely taken. The results revealed that the Flickr provided less accurate position information than Panoramio image. In a study which set out to determine VGI data quality, **Jackson, et al., 2013,** tested VGI quality in North America. The focus was on completeness and spatial error of linear feature such as roads and walkways and point features. They used three data sources: Federal government Oak Ridge National Laboratory (ORNL) as a reference data source, OSM data, and Open Street Map Collaborative Project (OSMCP) which is product of Unite State Geological Survey (USGS). The finding showed that the automated matching methods of OSMCP data with reference data source were more accurate than OSM data with ORNL datasets. While manual matching of OSM data with ORNL datasets were more accurate than OSMCP data with ORNL datasets .The main reason for that is due to the collection methods. The OSM data do not include the quality control processes, unlike the collection methods for OSMCP undergoes to government control.

Most of these researches focused on assessing the quality of OSM project. Therefore, the current study has chosen Google Map data, as another source for open data on a web, to study and analyze its quality such as positional accuracy measurement.

3. THE PROPERTIES OF GOOGLE MAP DATA

Google Map (GM) is a project designed to create and provide spatial data on Internet for free (non-commercial usage). It includes many map-based services such as the Google maps website, Google Transit (GT), Google Ride Finder (GRF), and maps embedded of the three websites types via the Google Maps API, **Books, and Wikipedia, 2010**. For some countries around the world, online spatial data is unavailable therefore Google has decided to open up Google Maps data through Google Map Maker (GMM) service. Google Map Maker is a Google Maps service that allows user to add or edit features, such as (roads, Points of Interest (also called POIs such as restaurants, banks, hotels, etc.) and polygons). Google Map Maker is a service launched by Google in June 2008, as a way to support the improvement of existing Google map data through the expert knowledge of local citizens. It's also a proprietary project; the GMM data can be downloaded for only 216 countries, **Google Map Maker, 2013**, as presented in **Fig. 2**.

When seeing the dynamic nature of Google Maps, one might think there is something magical going on under of dynamic nature. However, there’s really nothing magical about it, it's just Hyper Text Markup Language (HTML), Cascading Style Sheets (CSS), and JavaScript working together. Google Maps presented new concept in sense of content and interactivity, which is beginning in 2005. The concept is based on the AJAX (asynchronous JavaScript and XML) which implies possibility of adding additional information on map by using a free access of programming code called Application Programming Interface (API), **Svennerberg, and Drimmie, 2010**. An API is a set of programming instructions and standards for accessing a Web-based software application or Web tool. API is available for use by programmers in form of coding in some of scripting programming languages, such as PHP, ASP.NET, or ColdFusion. The API sends information about the new coordinates and zoom levels of the map in Ajax calls that return new images, **Garrett, 2005**. Google Map helps people to navigate map information. The GM interface is simple, intuitive, and easy to use. It consists of several user controls for managing or monitoring the map, such as zoom control, scale, and gets point–to–point driving directions, **Kanduri, 2012**.

Google Map data depends on Universal Transfer Marketer Projection (UTM). When the earth is perfectly spherical, the projection would be the same as the Mercator. Google Maps uses the formula for the spherical Mercator, but cannot show the poles. Unlike Google Earth (GE) coordinates system have 3D, the GM coordinates system have only 2D, **Books, and Wikipedia, 2010**. The coordinates system used in GM is the Word Geodetic System 84 (WGS84), which is the same system of the Global Positioning System (GPS).The coordinates are expressed using latitude and longitude. **Gibson, and Erle, 2006** mentioned that the property of Mercator

Projection (MP) of GM. It treats all lines of latitude as being perpendicular to all lines of longitude, and the MP conserves angles across local areas on the map, which is making it suitable for guiding navigation. Indeed this is one major reason of the Mercator projection continues to be used after 300 year, on the despite its tendency to distort the areas around the poles.

4. SITES OF THE STUDY

In this study two different study areas were chosen located in Al-Karada-Baghdad/Iraq and Al- Kadhumiyah-Baghdad/Iraq, as shown in **Fig. 3** (a and b). The main reason for this choice was to compare the positional accuracy of GM data in two different areas. In order to assess the positional accuracy, the preparation of the datasets was included selecting a well-defined points in tested datasets such as road intersection, building corners...etc. The selected points have the same coordinate systems and same projections in both tested and reference datasets. The number of tested points and area of two case studies are shown in **Table 1.** Geospatial dataset for two study areas were obtained from Mayoralty of Baghdad (MB), department of Geographic Information System (GIS) and Google Map (GM). The dataset were shapefiles and included several layers such as parcels, main road networks, street road networks, municipality boundaries, private and public buildings. The points were extracted for the edge of parcels and centerlines of the roads which provided into two datasets as can be seen from **Table 1.**

5. STATISTICAL METHOD FOR EVALUATING POSITIONAL DISCRIPENCY

One of the major concerns of geospatial data is accuracy. Positional accuracy of data may be far more important from other elements of quality to give users the position correctly. Accuracy can be grouped into two main categories: qualitative and quantitative. Also positional accuracy can be divided into two types: horizontal and vertical accuracy of a map or geospatial datasets, **Congalton, and Green, 2009**. Many organizations have established standards in a variety of ways as national or international standards. It can be used for the positional accuracy assessment, **Taupier, et al., 1999**. Most standards were designed in order to describe GIS data quality. The Spatial Data Transfer Standard (SDTS) from the American National Standards Institute (ANSI), for example, was approved in 1998 and considered five aspects of Geographic Information (GI) quality: lineage, positional accuracy, attribute accuracy, logical consistency and completeness. National Map Accuracy Standards (NMAS) was revised in 1947 in order to set standards of planimetric accuracy for paper maps, Accuracy Standards of Large Scale Maps, the Engineering Map Accuracy Standard (EMAS), **Congalton, and Green, 2009**. These standards were helpful but not specifically designed for digital geospatial data. Therefore, a more comprehensive standard was needed, due to the fact that geospatial data can be easily manipulated formats output, and reproduced at varying scales. In 1998, a committee of the Federal Geographic Data Committee (FGDC) developed and formed the National Standard for Spatial Data Accuracy (NSSDA). This standard enabled users to test and analyze positional accuracy of digital datasets, with respect to ground geospatial data of higher accuracy.

The NSSDA presents guidelines for the distribution of tested points. It assumed that the area to be evaluated is a rectangle. The tested area is divided into four quadrants and a diagonal is to be established across the area. Tested points should be spaced at interval of at least 10 percent of the diagonal. At least 20 percent of tested points are to be located in each quadrant. In addition, the minimum number of tested points should be no less than 20 well-defined points in order to evaluate the accuracy of the datasets. The NSSDA is index of relative horizontal accuracy which is tested at the 95% confidence interval, and it can be calculated as shown in Eqs. (1) and (2):

$$RMSE_E = \sqrt{\sum_i^n (E_d - E_c)^2}\,/n\,, \quad RMSE_N = \sqrt{\sum_i^n (N_d - N_c)^2}\,/n \tag{1}$$

$$RMSE = \sqrt{\left(\sum_{i=1}^n (\delta E_i)^2 + (\delta N_i)^2\right)}\,/n \tag{2}$$

n : The number of tested points,
E_d, N_d : The coordinates of formal dataset,
E_c, N_c : The coordinates of tested dataset, and
$\delta E^2{}_i, \delta N^2{}_i$: The differences in easting and northing for $i^{th} check$ points, between formal and tested datasets.

The NSSDN accuracy can be computed for two cases as shown in Eqs. (3) and (4):

If $RMSE_E = RMSE_N, then$

$$RMSE_r = \sqrt{2(RMSE_E)^2} = \sqrt{2(RMSE_N)^2}\,, = 1.7308 * \ RMSE \tag{3}$$

If $RMSEx \neq RMSEy$, *then*

$$NSSDA\ Accuracy = 2.4477 * 0.5 * (RMSE_E + RMSE_N\ ) \tag{4}$$

In this research, tests were undertaken applying the NSSDA methodology to examine and analyze the relative positional discrepancies of tested points in both study areas. **FigS. 4 and 5** show the distribution of tested points according to NSSDA methods.

6. PROGRAM DESIGN FOR IMPLEMENTING RESEARCH METHODOLOGY

A visual tool interface or graphical user interface (GUI), by using Visual Basic Language, was implemented and designed to assess the quality of Google Map (GM) data. The intention was to create a user-friendly interface incorporating quantitative and visual analysis of GM dataset. The workflow of the designed program is illustrated in **Fig. 6**. By using designed interface, there are three main steps to determine and analyze positional accuracy, as shown in **Fig. 7**. First, data can be imported as a text file (.txt) for the coordinates of tested points. Second, the comparison and analysis of tested datasets will start by applying the methodology of this research. Third, output results (graphs and quantities values) can be exported and saved as a report.

After loading coordinates data, one can select positional accuracy assessment option from the main program interface. A window will appear to assess the positional accuracy of case study one, as illustrated in **Fig. 8**. This window has many options: Back, Run, Next, Export output, and Diagrams. These options can be used based on the needs of users. From **Fig. 8**, one can notice that the positional descriptive statistics are reported numerically, also box plot to represent t-distribution of different in easting, northing, and Euclidian distance of tested points. In addition, remarks to accept or reject the outcomes of case study one. Similarly, **Fig. 9** shows interface for evaluating the positional accuracy of case study two.

In this research, two-sample t-test was also adopted to compare the mean of the accuracy of GM data in two sites, as presented in **Fig. 10**.

7. POSITIONAL ERROR CHARACTERIZATION

The examination of positional error was conducted in a number of ways to determine the quality of Google Map (GM) data. Firstly, basic descriptive statistics were determined including mean, median, minimum, maximum, standard deviation, variance, and inter-quartile range. Secondly, statistical significance was analyzed using one-sample t-tests to investigate the relationship between GM data and Mayoralty of Baghdad (MB) data. In addition, a comparison between the two study areas was undertaken by applying two-sample t-test. Thirdly, boxes plot were created to interpret the outcomes.

Table 2 provides a comparison of the descriptive statistics of the differences in Easting and Northing (E, N) of case study one (Al-Karada-Baghdad). The mean errors of the sample were (1.016.-9.851) of differences in (E, N) respectively. The median errors were (0.888,-10.866) of differences in (E, N) respectively. Form the table below, one can also see that the median value are smaller than the mean values which indicates that the distribution of differences of E could be normal .The standard deviation values were (2.033, 3.256) of differences in (E, N) respectively. A low standard deviation indicates that the tested points tend to be very close to the mean; while a high standard deviation indicates that the tested points are spread out over a large range of values. The mean and standard deviation are very important parameters for distribution of measurement values for normal distribution. In this research, the mean and standard deviation values are almost close to each other and which indicates that they are subject to be normal distributed. Similar observations can be made for the tested data of case study two (Kadhumiyah-Baghdad), as shown in **Table 3**.

As mentioned earlier, one-sample t-test was also applied to calculate the error values between reference and tested datasets. In one-sample t-test, the null hypothesis should be stated as($Ho: \mu1 = \mu2$): where μ1 is the mean of the first dataset (case study one), and μ2 is the mean of the second dataset (case study two). On the other hand, the alternative hypothesis should be stated as($H1: \mu1 \neq \mu2$). For this project there are two values to accept or reject the null hypothesis: t-critical and p-value. the null hypothesis(H_0) assumed to be a smaller or equal to 0.6 depending on the accuracy of MB, While the alternative hypothesis(H_1) suggested to be greater than 0.6. The critical value usually obtains from the t- distribution. If the t-value falls within the non-rejection region, the null hypothesis at 95% CI cannot be rejected. It should reject the alternative hypothesis when t-value falls into the area of the rejection region. Whereas, P-value referred to significance level, it is used as a standard for accepted and rejected the null hypothesis, **Black, 2011**. For this study, 95% CI was applied according to NSSDA approach whereas P-value was 0.05. The findings in **Table 4** showed that the t-value was 24.307, t- critical was 1.960, and p-value was 0.000 which referred to that the t-value is larger than the t- critical, and the p-value is less than the $\propto$(0.05). This has sufficient evidence to reject the null hypothesis and accept the alternative hypothesis. The two-sample two-test found that the mean of the case study one (Al-Karada-Baghdad) was 10.190 m, while the mean value of the case study two (Al-Kadhumiyah-Baghdad) was 4.015m. This proves that there is a difference in the average deviation of the accuracy of the Google Map (GM) data in different study areas which was 6.175 m. Also, other statistical analysis, such as RMSE and NSSDA accuracy, showed that case study two have different values of case study one, as demonstrated in **Table 4**.

The results showed the accepted the alternative hypothesis and should be reject the Null hypothesis into both study areas based on the t-critical and p-value. This indicates that the data GM data cannot be appropriate for the purposes of engineering applications.

8. DISCUSSION AND CONCLUSION

A methodology has been developed to evaluate the positional accuracy of Google Map (GM) data. Tests were conducted using the National Standard for Spatial Data Accuracy (NSSDA) to assess the relative positional accuracy of tested dataset. Then, the RMSE values were calculated, yielding comparative measures of positional discrepancy. The linear displacement (magnitude of error) of each point was determined by measuring the Euclidean Distance (ED) between the points of reference and tested datasets. The descriptive statistics such as mean, median, standard devotion, maximum, minimum, variance, and interquartile range for the differences in Easting and Northing were computed. A tool was designed using Visual Basic program to represent and analyze the results of positional accuracy. The code of this program contents three parts: input data, analysis, and output results. Two interfaces were designed for assessing positional accuracy of two study areas in order to reduce time and efforts for comparing the mean value of two-sample t-test.

The results of this analysis showed that the informal (GM) data does not match the formal datasets (MB) in any of the case study areas. The NSSDA accuracy at 95% Confidence Interval (CI) of the first case study (Al-Karada-Baghdad) was equal to 15.48 m for 260 tested points. While in second case study (Al- Kadhumiyah-Baghdad*)* was equal to 8.71 m for 300 tested points. This refers to that 95% of tested points have an error with respect to formal data (Mayoralty of Baghdad (MB)) equal to or smaller than 15.48 m, 8.71; whereas 5% of tested points have an error larger than 15.48, 8.71 m. In general, therefore, it seems that the easting of GM data is more accurate than the northing of GM data when compared with formal dataset (MB).

The results of one-sample t-test indicated that the accuracy of GM data larger than the threshold value (0.6 m) which indicated that the GM data is not suitable for accurate Engineering applications. The findings of two-sample t-test revealed that the case study two was more accurate than case study one. The difference of mean between two study areas was 6.18 m. This provides strong indication that the accuracy of GM data does not equal over different area. It concluded that the GM data can be used for engineering applications that need low accuracy such as the preliminary surveying of projects design, tourism, and reconnaissance.

REFERENCES

- Antoniou, V., Morley, J. and Haklay, M., 2010, *Web 2.0 Geotagged Photos: Assessing the Spatial Dimension of the Phenomenon*. Geomatica, The Journal of Geospatial Information, Technology and Practice, Vol. 64**,** PP.99-110.

- Ather, A., 2009, *A Quality Analysis of Open Street Map Data,* MSc. thesis, University College London.

- Books, L. and Wikipedia, S., 2010, *Web Map Services: Google Maps, Bing Maps, Openstreetmap, Wikimapia, Geabios, Transport Direct Portal, Google Map Maker, Openseamap, Yahoo Maps,* General Books.

- Cartwright, W., 2008, *Google Maps and Mobile Devices: Can Just One Generic Design Work?*, PP.215-222.

- Ciepłuch, B., Jacob, R., Mooney, P. and Winstanley, A., 2010, *Comparison of the Accuracy of OpenStreetMap for Ireland with Google Maps and Bing Maps*, Proceedings of the Ninth International Symposium on Spatial Accuracy Assessment in Natural Resuorces and Enviromental Sciences 20-23rd July 2010, University of Leicester, UK.

- Congalton, R. G. and Green, K. ,2009, *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices,* USA, CRC Press / Taylor & Francis Group.

- Delavar, M. R. and Devillers, R., 2010, *Spatial Data Quality: From Process to Decisions,* Transactions in GIS, Vol. 14**,** PP. 379-386.

- Flanagin, A. J. and Metzger, M. J. , 2008, *The Credibility of Volunteered Geographic Information,* GeoJournal, Vol. 72**,** PP. 137-148.

- Garrett, J. J., 2005, *Ajax: A New Approach to Web Applications.*

- Gibson, R. and Erle, S., 2006, *GoogleMaps Hacks,* O'Reilly Media, Inc.

- Goodchild, M. F., 2007, *Citizens as Voluntary Sensors: Spatial Data Infrastructure in the World of Web 2.0,* International Journal of Spatial Data Infrastructures Research, Vol. 2**,** PP. 24-32.

- Google Map Maker, 2013, Available at: http://mapmaker.google.com/datadownload [Accessed 20 July 2013].

- Haklay, M., 2010, *How Good is Volunteered Geographical Information? A Comparative Study of OpenStreetMap and Ordnance Survey Datasets,* Environment and Planning B: Planning and Design, Vol. 37, PP. 682-703.

- Haklay, M., Singleton, A. and Parker, C., 2008, *Web Mapping 2.0: The Neogeography of the GeoWeb,* Geography Compass, Vol. 2**,** PP. 2011-2039.

- Hochmair, H.H., and Zielstra, D., 2012, *Positional Accuracy of Flickr and Panoramio Images in Europe*. In Jekel, T., A. Car, G. Griesebner and J. Strobl (Eds.), GI_Forum 2012: Geovisualization, Society and Learning, Berlin: Wichmann, PP. 14-23.

- Howe, J., 2006, *The Rise of Crowdsourcing* [Online]. Available at: http://www.wired.com/wired/archive/14.06/crowds.html#Replay[Accessed 10 July 2013].

- Kanduri, S., 2012, *Term Paper Is Google Maps a place for advertising? Give good examples!* HFU(Hochschule Furtwangen University).

- Kounadi, O., 2009, *Assessing the Quality of OpenStreetMap Data*, Msc thesis, Department of Civil, Environmental And Geomatic Engineering, University College of London, UK.

- Svennerberg, G. and Drimmie, R., 2010, *Beginning Google Maps API 3*, Apress.

- Taupier, R., Beliaeva, N., Gnadinger, L., Kovalenko, S. and Wheeler, M., 1999, *Parcel Mapping using GIS: A Guid to Digital Parcel Map Development for Massachusettslocal*

Goverments, University of Massachusetts, Office of Geographic Information and Analysis for MASSGIS.

- Turner, A., 2006, *Introduction to Neogeography*, O'Reilly Media, Inc.

- Zielstra, D. and Zipf, A., 2010, *A Comparative Study of Proprietary Geodata and Volunteered Geographic Information for Germany*, 13th AGILE International Conference on Geographic Information Science.

NOMENCLATURE

n = the number of tested points, dimensionless.

E_d, N_d = the coordinates of formal dataset, m.

E_c, N_c = the coordinates of tested dataset, m.

$RMSE_E$ = root Mean Square Error in Easting, m.

$RMSE_N$ = root Mean Square Error in Northing, m.

RMSE = total Root Mean Square Error, m.

$\delta E^2{}_i, \delta N^2{}_i$ = the differences in easting and northing for i^{th} check points, between formal and tested datasets, m.

Table 1. Study areas and their properties.

Data sets	Upper left corner		Lower left corner		Number of tested points	Area (km^2)
	Easting (m)	Northing (m)	Easting (m)	Northing (m)		
Study Area (1)	441998.350	3682841.827	449162.768	3683916.910	260	9.824
Study Area (2)	436255.880	3694110.799	439454.114	3688523.522	300	15.054

Table 2. Statistics computed from differences in Easting and Northing of case study one.

Statistics	Diff .in easting (m)	Diff .in northing (m)
Mean	1.016	-9.851
Median	0.888	-10.866
Standard deviation	2.033	3.256
Maximum	6.842	5.854
Minimum	-4.565	-13.839
Variance	4.131	10.605
IQR	2.379	3.107

Table 3. Statistics computed from differences in Easting and Northing of case study two.

Statistics	Diff .in easting (m)	Diff .in northing (m)
Mean	-0.116	-1.794
Median	-0.031	-0.447
Standard deviation	2.403	9.309
Maximum	6.322	6.669
Minimum	-6.676	-13.940
Variance	5.777	86.659
IQR	2.798	3.781

Table 4. Comparison of Root Mean Square Error (RMSE), National Standard for Spatial Data Accuracy (NSSDA), and t-distribution positional discrepancies for compared datasets.

Case study	RMSE (m)	NSSDA(m)	t- value	t-critical	p-value
Al-Karada_ Baghdad	10.621	15.479	51.577	1.960	0.000
Al-Kadhumiyah_ Bagdad	5.428	8.713	16.192	1.960	0.000

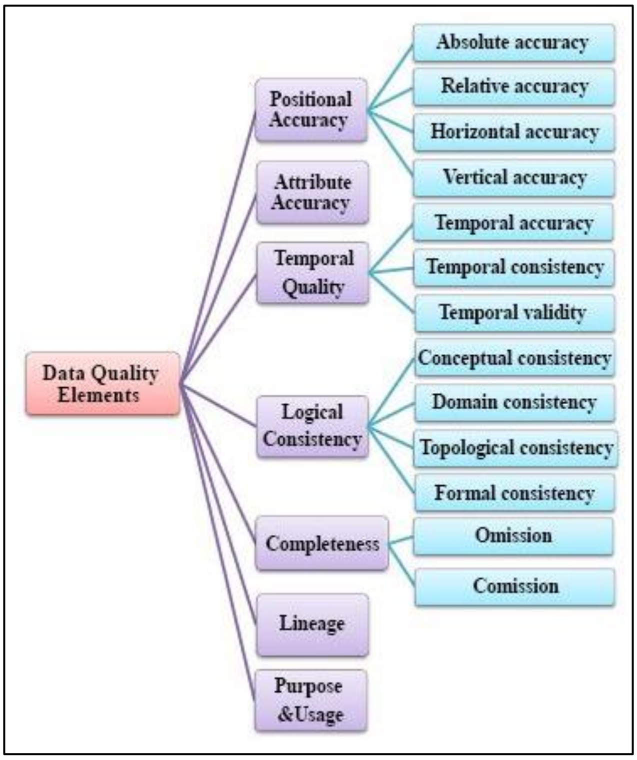
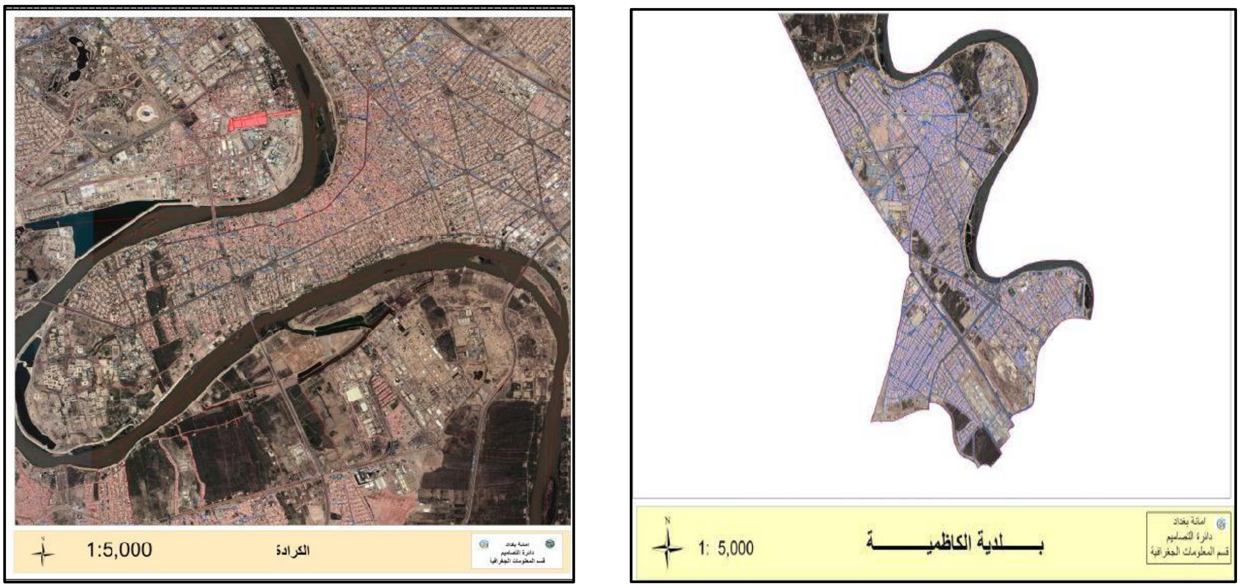


Figure 1. Data quality elements and sub-elements.

Figure 2. Google Maps Maker data availability.
(http://www.google.com/mapmaker/mapfiles/s/launched.html).



(a) Formal data Karada (case study one). (b) Formal data Kadhumiyah (case study two).

Figure 3. Formal data (source: Mayoralty of Baghdad (MB)).

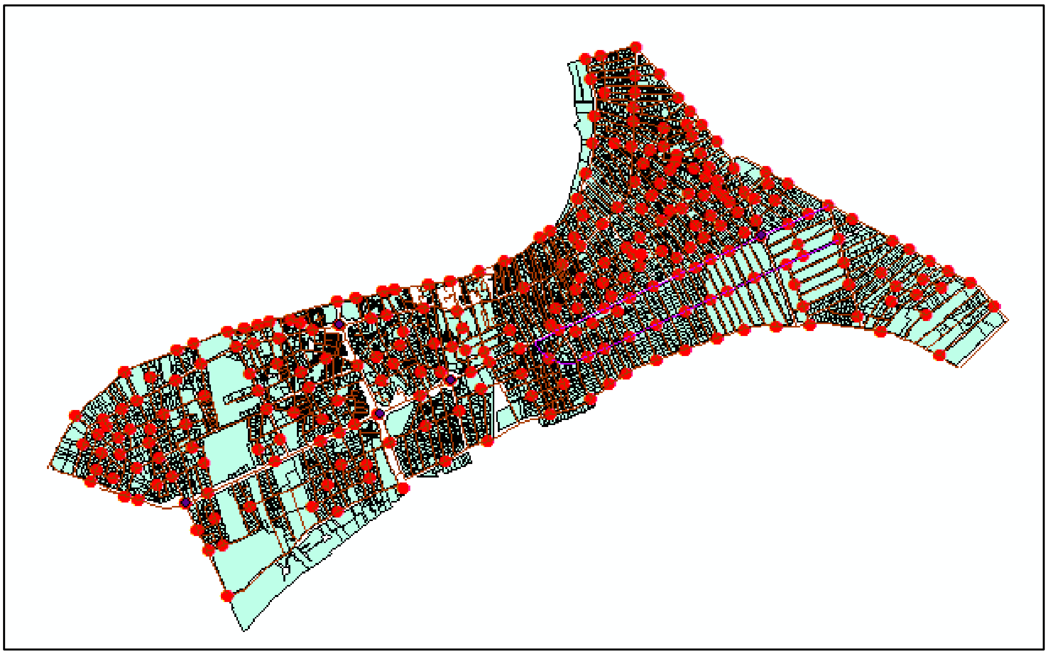


Figure 4. The distribution of tested points (case study one).

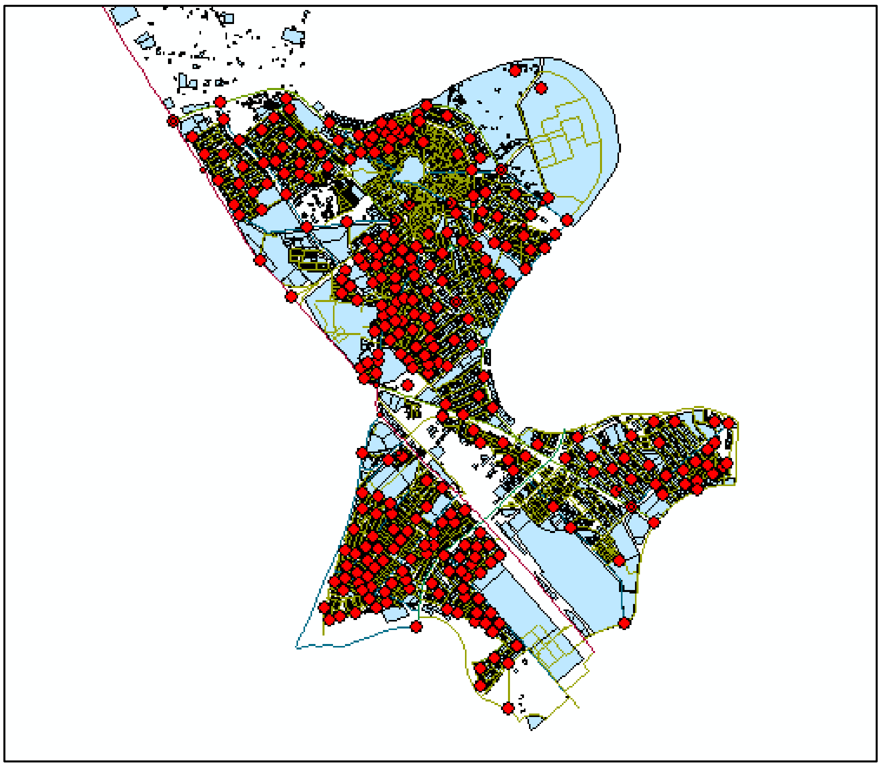


Figure 5. The distribution of tested points (case study two).

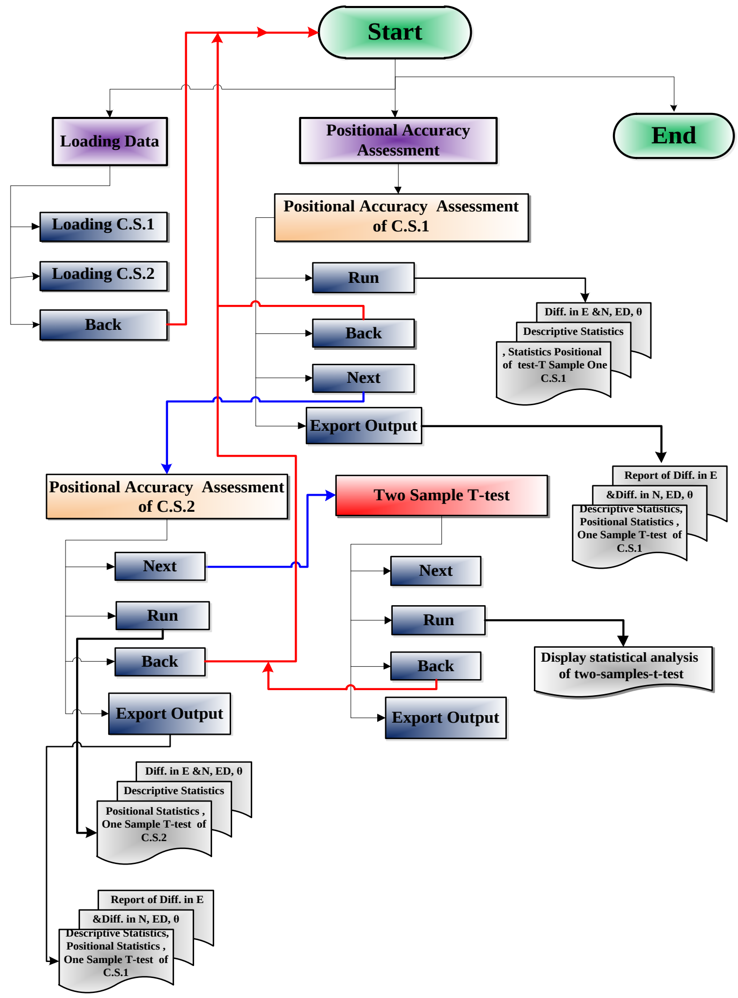


Figure 6. The workflow of the designed program.

Figure 7. The main interface of the developed program.

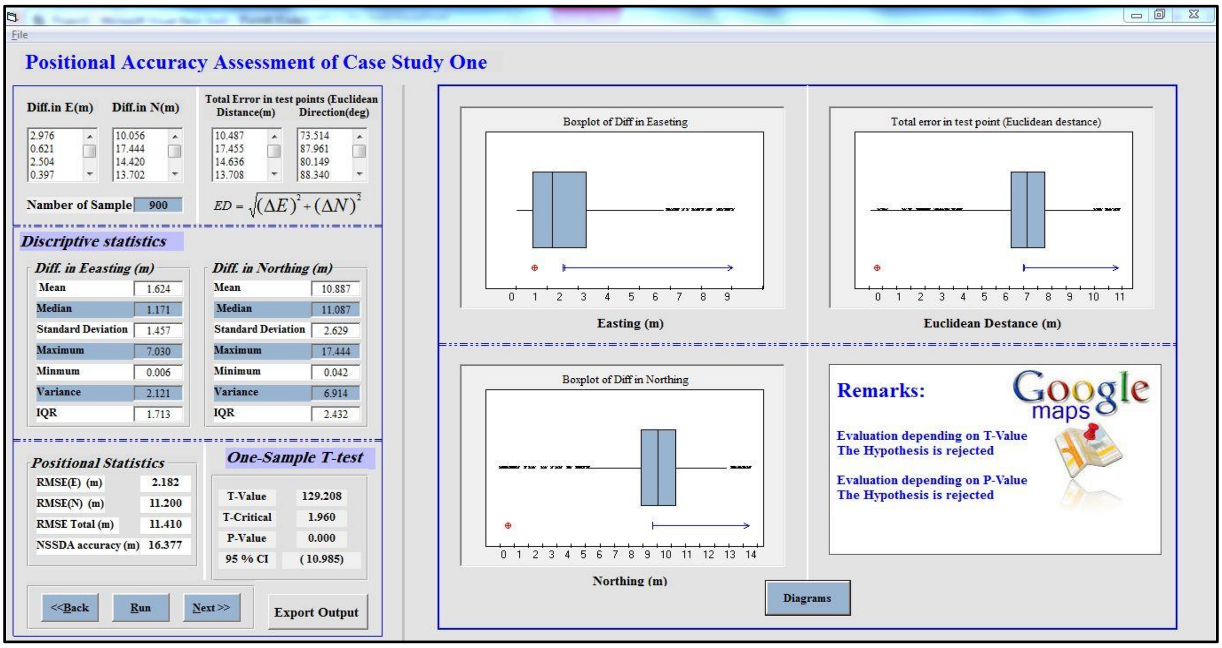


Figure 8. Interface for positional accuracy analysis (case study one).

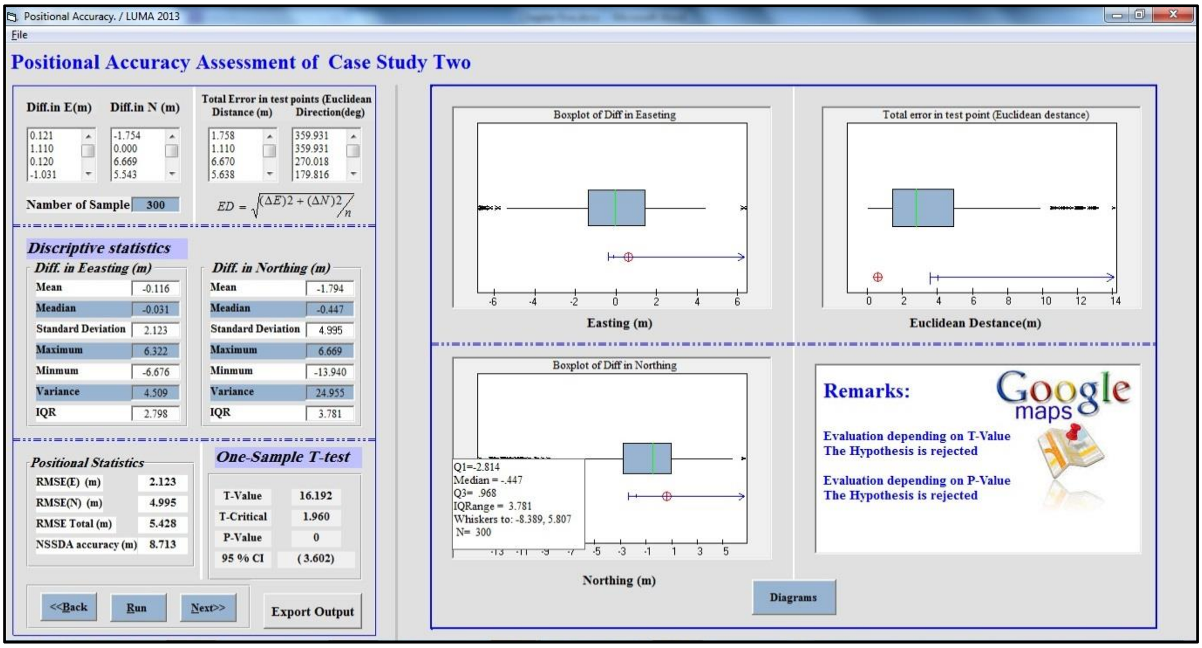


Figure9. Interface for positional accuracy analysis (case study two).

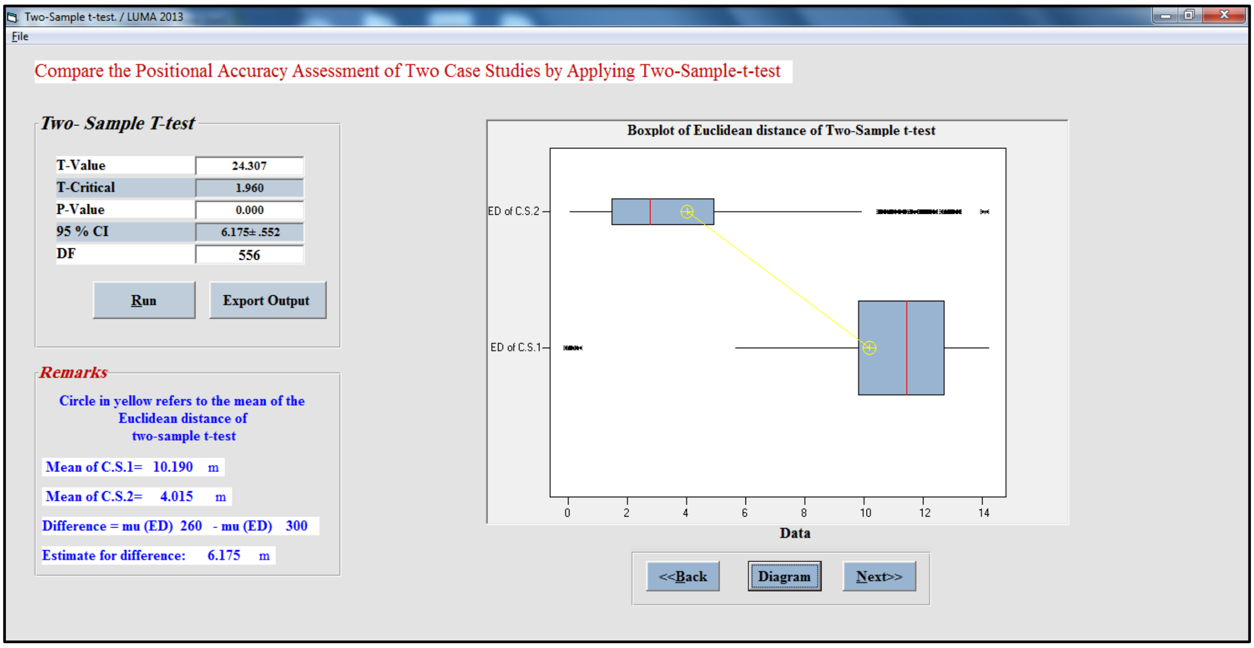


Figure 10. Interface for comparing the accuracy of GM in two different study areas.