CADAstral POSITIONING ACCUracy IMPROVEMENT (PAI): A CASE STUDY OF PRE-REQUISITE DATA QUALITY ASSURANCE

N. M. Hashim 1, A. H. Omar 1, K. M. Omar 2, M. A. Abbas 3, M. A. Mustafar 3, S. A. Sulaiman 3, * 1

1 Geomatic Innovation Research Group (GnG), Faculty of Built Environment and Surveying, Universiti Teknologi Malaysia, Johor Bahru, Malaysia - shahndcdb@gmail.com, abdullahhisham@utm.my
2 Geospatial Science and Technology College, Kuala Lumpur, Malaysia – rector@gstc.edu.my
3 Centre of Studies for Surveying Sciences & Geomatics, Faculty of Architecture, Planning & Surveying, Universiti Teknologi Mara Perlis, Arau, Malaysia – (norshahrizan, mohdazwan, mohamadarsi, saifulaman) @uitm.edu.my

KEY WORDS: Positional Accuracy Improvement, Legacy Dataset; Cadastral Database Modernization

ABSTRACT:

Nowadays, there is an increasing need for comprehensive spatial data management especially digital cadastral database (DCDB). Previously, the cadastral database is in hard copy map, then converted into digital format and subsequently updated. Theoretically, these legacy datasets have relatively low positional accuracy caused by limitation of traditional measurement, adjustment technique and technology changes over time. With the growth of spatial based technology especially Geographical Information System (GIS) and Global Navigation Satellite System (GNSS) the Positional Accuracy Improvement (PAI) to the legacy cadastral database is inevitable. PAI is the refining process of the geometry feature in a geospatial dataset through integration between legacy and higher accuracy dataset to improve its actual position. However, by merely integrating both datasets will lead to a distortion of the relative geometry. Thus, an organized method is required to minimize inherent errors in fitting to the new accurate dataset. The focus of this study is to design a comprehensive data preparation for legacy cadastral datasets improvement. The elements of datum traceability, cadastral error propagation and weightage setting in adjustment will be focused to achieve the targeted objective. The proposed result can be applied as a foundation for PAI approach in cadastral database modernization.

1. INTRODUCTION

Cadastral map at the initial stage of its existence is managed and stored on a hard copy map. The emergence and development of computer technology, especially in relation to spatial data, led to the cadastral map also directly involved in the digital transformation process. Basically, managing cadastral data digitally can help in making decisions more accurately, detail and flexible in relation to the spatial strategic planning (Durdin, 1993; Effenberg et al., 1999; Ting et al., 1999).

At the beginning, this transformation process only focuses on conversion form hard copy to digital data without consider the proper coordinate system. In line with modern accuracy requirements, cadastral database is further transformed into a coordinated cadastre where the database is not only in digital form, but its coordinates are also in the actual position. In support of this aspiration, legacy cadastral datasets that are still in the plotting coordinate state or inaccurate coordinates will be transformed into a standard and high accuracy coordinate. This process is supported by the presence of GNSS technology where the process of providing control points is easier to tie to the network of cadastral databases.

In the past, the integration of new cadastral survey and databases is not a major requirement where surveyors work in surveyed lot and its adjacent lots with regardless of the entire network’s surrounding areas. The entire lot is not considered for the process of adjustment where only several control points used to control the orientation in the merging process due to the limitations of the network adjustment capability at that time. There are also issues that exist during the digitizing process of previous hard copy maps. The common error in digitizing process such as distortion of source map, digitizing operational errors and ground coordinate system constituted a combination of systematic and random errors (Tong, Shi, & Liu, 2009). Today's cadastral database is a combination of data from various sources, technologies and measurement techniques depending on the method of production and measurement used according to technological and legal changes from time to time (Sisman, 2014). As a result, all these factors will contribute errors when the legacy dataset integrating to the new high accuracy data.

The need for cadastral database updates is extremely crucial and inevitable. The development of Geographical Information System (GIS) and data sharing campaign, the need to combine spatial data from different sources and accuracy has increased dramatically. This process is very important to enable datasets that have different accuracy can be combined and analyzed together. This process requires a systematic approach so that the problems discussed above can be minimized.

![Figure 1. PAI concept (Norshahrizan et al., 2016)](image-url)
Therefore, this study was conducted to develop a systematic approach data preparation for the Positional Accuracy Improvement (PAI) of legacy dataset. In the following chapter, the concept of PAI is emphasized. Then the datum traceability requirements are discussed to control the cadastral block adjustment work in the future. Issues related to data input types are also discussed to reduce the gross and systematic error in the observation data. At the end of the literature review, the weightage system will be discussed which is an essential component of supporting high accuracy network database.

2. POSITIONAL ACCURACY IMPROVEMENT

PAI is a process of improving the position of the geometric coordinates of a feature in a geospatial dataset to represent its actual position (Rönsdorf, 2008). The PAI can be classified as the improvement of low accurate legacy dataset to more accurate dataset. Figure 1 illustrates the concept of PAI (Norsahrizan et al., 2016). PAI is a process to improve the position of geometric coordinates for geospatial to represent their true position (Rönsdorf, 2008). The success of this PAI is supported by the progress and development of the Global Navigation Satellite System (GNSS) where the process of making high precision control point distribution is possible. Additionally, the existence of high resolution remote sensing images makes it easier for the recognition of conflict area between the adjacent land parcels and this indirectly simplifies the PAI process (Hope et al., 2008).

PAI are usually used in two conditions reference data and user data (Rönsdorf, 2008). PAI reference data means improving the geometric position in a reference dataset that describes the physical or abstract features of the earth. As such, a feature is determined based on the standard coordinate reference system like Datum Geocentric Malaysia 2000 (GDM 2000) or WGS-84 in global coordinate system. Meanwhile, PAI user data illustrates the success of maintaining geometric relationships for new accurate dataset as original in legacy datasets after the integration process is done.

The Least Square Adjustment (LSA) method is often used to achieve optimum solution in PAI (Tamim 1995, Wolf and Gihlani 2006, Merritt and Masters 1999, Gielsdorf et al. 2004, Merritt 2005, Tong et al. 2005, Casado 2006, Hope et al., 2008). LSA is used widely in the disciplines of surveying where this method is more advanced technique which adjusts observations based on the laws of probability, which models the occurrence of random errors. LSA principle is the squares of the residuals must be minimized as expressed in Equation (1).

\[ \sum v_i^2 = v_1^2 + v_2^2 + \ldots + v_n^2 = \text{MINIMUM} \]  

where \( v \) = residual

LSA is used in estimating parameters for coordinate transformation (Mikhail and Ackermann (1976), Wolf and Gihlani (2006), and Koch (1999). Tamim (1995) used LSA to create a digital cadastral overlay through upgrading digitized cadastral data. Merritt and Masters (1999) and Merritt (2005) developed the spatial adjustment engine based on the LSA to improve the accuracy of cadastral data in Australia. Tong et al. (2005) presented a LSA model to resolve inconsistencies between the digitized and registered areas of cadastral parcels, and further improved the adjustment model by introducing scale parameters to reduce the influence of systematic error in the adjustment. Felus (2007) used LSA to enhance the spatial accuracy of digital cadastral maps via stochastic constraints and geometric conditions. Hope et al. (2008) proposed a method of least squares with inequalities for data integration for preserving the spatial relationships among features.

2.1 Datum Traceability

Principally, term traceability can be defined as an ability to trace something. Measurement traceability means an unbroken chain of comparisons relating an instrument’s measurements to known standards. Measurement traceability can be used to determine the bias, precision, and accuracy of the instrument. It can also be used to show the chain of custody as evidence for retrieve information.

Ses et al. (2000) says that when a measurement using international or national measurement standards, the validation of the data is known as legal traceability. Legal traceability giving users and authorities the certainty about measurements that carried out which it is very important to ensure that the representative data is indicating the actual measurement. Bissett (2008) states that legal traceability is the ability of the surveyor to measure land parcels based on the documented measurements performed by other with a common uniformity of measurement.

In Malaysia, the concept of legal traceability has been discussed in general in by Ses et al., (2000). In this study, the GNSS Cadastral Survey Guidelines 1999 are discussed which explained to surveyors how to carry out cadastral measurements using GNSS techniques. According to the author, measurements are only applicable if they have carried out various test or calibration procedures and follow the recommended practice in the field and office procedures.

Further research on traceability is discussed by Gill et al., (2016). Basically, researchers agree with the overall regulatory and cadastral survey guidelines for Malaysia. However, certain improvements can be made namely reliability of N-RTK, GNSS instrument calibration and GNSS reference station accuracy. In this study, legal traceability is very important. The CRM datum used in cadastral work is not only used as a traversing control, but it is also used as the main control point for the purpose of adjusting the cadastral database block. In addition, this CRM point will also be a common point for the integration process between different blocks so that there are connection points between adjacent blocks.

2.2 Angular Based Least Square Adjustment

Resurvey cadastral data may be the best technique for solving PAI problems. However, this method is very complicated, it takes a long time, a great workforce and a very high cost. For this purpose, it is necessary to reconstruct the cadastral survey in accordance with the new control framework and this process is an extraordinary effort and the estimated cost is very high (Arvanitis & Koukopolou, 1999). Another alternative solution is to use the original observation as contained in the surveying field book. According to Buyong and Kuhn (1990) and Durgin (1993), maintaining the old measurements and combining with the new observations may be used to support PAI efforts.
However, based on studies conducted in Israel by Perelmuter and Steinberg (1992), reprocessing of field books is impossible due to the weakness of the original datum network, data updating weaknesses and incomplete records. In addition to the economic perspective, reprocessing of field book requires many operators and takes a long time to complete (Fradkin & Doytsher, 2002).

A potential alternative for restoring the legacy datasets with reasonable cost is the PAI process where transform the existing legacy dataset into new high accuracy dataset. As discussed above, LSA is a widely used method in the PAI process. However, most of the previous studies refer to the necessity to make new measurements that need to be linked to the old measurements (bearing distance in certified plan) in the PAI process. For such cases, it is very important to ensure that the old measurements free from gross error and systematic errors so that the adjustment of the integration can be accepted. Principally, the adjustment of the measurement data is focused on random errors where the gross and systematic errors were identified and removed (Wolf and Gihlani, 2006).

In the case of cadastral database in Malaysia, the cadastral plan boundary line is represented by bearing and distance. When a block adjustment is made, the raw data used for the adjustment purpose is the control coordinates (GNSS), bearing (CP) and distance (CP). Theoretically, the data is measured directly on the field while the bearing is the data derived from the angular observations measured on theodolite and datum bearing obtained. The datum bearing according to the circular 2002 rule is either through the sun observation or the angle within 3 in good condition boundary marks. There is an error propagation concept that needs to be understood clearly where the bearing is not directly observed, but it is derived from the angular measurement as well as datum bearing as stated above. This error propagation issue needs to be viewed correctly so that results obtained from the adjustment are reliable and acceptable.

Measurements can be classified as direct and indirect. Direct measurement is made using the instrument directly to the unknown quantity such as measure the distance between two survey stations in cadastral work. Indirect measurements are performed where it is impossible or impractical to make direct measurements where desired quantities are determined from their mathematical relationship of direct measurements. In the case of this bearing observation, it is an indirect measurement where it comes from direct measurement angles (theodolite) and control bearings (solar observation and angle of 3 old boundary marks). During this procedure, errors in the direct observation will be disseminated during the calculation process into the indirect value. Hence, indirect measurement (bearing) contains errors and this concept is known as error propagation. The Equation 2 shows the basic concept of error propagation known as Law of Propagation of Variance (LOPOV).

\[
\sigma^2 = \left(\frac{\partial Z}{\partial x_1}\sigma_{x_1}\right)^2 + \left(\frac{\partial Z}{\partial x_2}\sigma_{x_2}\right)^2
\]

where \(\sigma^2\) = Variance of adjusted observation
\(\frac{\partial Z}{\partial x_1}\) = Partial derivative respected to X1
\(\sigma_{x_1}\) = Standard deviation X1
\(\frac{\partial Z}{\partial x_2}\) = Partial derivative respected X2
\(\sigma_{x_2}\) = Standard deviation X2

To solve this issue of propagation errors, this study suggests bearing observations involving the issue of error propagation being changed to the angle observation. This method is implemented as an effort to reduce issues related to error propagation where the original observation from theodolite is reused. As discussed, that retrieve data from original field book is not practical usable, then the interior angle formed of the lot in the certified plan is used. For example, if the CP is the first class survey where the maximum angular misclosure is 1'15” is used, then if there are 5 boundary marks on the lot, the standard deviation of each angle can be estimated at about 10” one station. If the lot is in the second class survey, it is still relevant because the standard deviation for each angle is about 40 “. Subsequently, it will be even more minimal if more angles are involved in the formation of a lot. The following is the misclosure limitation of the survey according to the 2002 Cadastral Survey Regulation of Malaysia.

### Table 1: Malaysia Cadastral Survey Accuracy Class

<table>
<thead>
<tr>
<th>Survey Class</th>
<th>Angular Misclosure</th>
<th>Linear Misclosure</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>1'15”</td>
<td>1:8000</td>
</tr>
<tr>
<td>2nd</td>
<td>2'30”</td>
<td>1:4000</td>
</tr>
<tr>
<td>3rd</td>
<td>5’</td>
<td>1:2000</td>
</tr>
</tbody>
</table>

The determination of the standard deviation value is reliable because it can be controlled by the angular misclosure tolerance applied in the previous work. Compared to bearing measurements, the standard deviation value set is still doubtable because of the standard deviation of the datum bearing error cannot be evidently confirmed.

Since the legacy dataset are less accurate in positioning, the integration between legacy datasets and higher updated accuracy like GNSS is one of the most possible methods to improve the legacy datasets accuracy (Hope et al., 2008). However, in her study, it is found that by simply replacing a sample of legacy dataset with more accurate version will lead to distortions of the neighboring geometry (Figure 2). In addition, often the relative geometry of the legacy datasets is better than its absolute accuracy and supposed to be the spatial relationship or relative geometry between features must be preserved.

![Figure 2: Cadastral Boundaries (Solid Lines) and Higher Accuracy (Dashed Lines). Polygon xady is Distorted If Point Abcd Are Simply Replaced By ABCD (Hope et al., 2008)](image_url)
In general, the use of angles for cadastral block adjustment is more flexible than the use of the bearing. The use of angles allows the adjustment to be more flexible where the formation of the lot is no longer controlled by a rigid direction, but only the angle between the boundary stone is used. Norshahrizan et al. (2017) proved in his study that using the angle to make adjustment of geometric relationships would be better based on the closeness of the parcel area after adjustment compared to the original parcel area in the old database. The study also found that if several combinations of block is performed adjustment, the angle method is more flexible and the potential for passing the global test is higher.

In this study, the parametric adjustment of functional model is applied. The mathematical model of horizontal angle and distance is used in the adjustment process. The mathematical model of horizontal angle and distance observation are like equations (3) and (4) respectively.

$$\Delta \theta = \tan^{-1} \frac{X_B - X_A}{Y_B - Y_A} - \tan^{-1} \frac{X_D - X_A}{Y_D - Y_A}$$

$$\Delta d = \sqrt{(X_A - X_B)^2 + (Y_A - Y_B)^2} + \sqrt{(X_D - X_B)^2 + (Y_D - Y_B)^2} + \sqrt{(X_D - X_A)^2 + (Y_D - Y_A)^2}$$

where

- $\Delta \theta =$ Horizontal angle
- $X_A, Y_A =$ Point A estimated coordinate
- $X_B, Y_B =$ Point B estimated coordinate
- $X_D, Y_D =$ Point D estimated coordinate
- $L_{AB0} =$ Estimated horizontal distance A-B
- $L_{AD0} =$ Estimated horizontal distance A-D

$$L_{AB} = \sqrt{(X_A - X_B)^2 + (Y_A - Y_B)^2}$$

where

- $L_{AB} =$ Distance
- $X_A, Y_A =$ Point A estimated coordinate
- $X_B, Y_B =$ Point B estimated coordinate
- $L_{AB0} =$ Estimated horizontal distance A-B

When a measurement is carried out, the quality of the raw data should be identified appropriately. Theoretically, the high precision observation value should be represented by a low variance value. This is important because it is high precision observations supposed to receive a slight portion of the overall correction. Likewise, measurements are low precision, having a higher variance and will gain greater portion in the correction process.

Principally, the weight of an observation represents the quality of an observation compared to others. Weights are important to control the sizes of corrections applied to measurements in an adjustment. Weight is inversely proportional to the variances, more precise an observation the higher its weight vice versa. It also follows that correction sizes should be inversely proportional to weights. This study suggests that weightage data should be set appropriately so that adjustment results are more reliable.

### 3. PROPOSE METHODOLOGY

The Cadastral PAI data preparation focuses on three main phases, datum traceability, angular based least square adjustment and suitable setting for raw data weightage. The first stage is the datum traceability which is the backbone of the cadastral block adjustment. As discussed above, datum traceability is an important component of the PAI. The changing coordinate system based on the current requirement as an example of ITRF through the process of updating several times until 2018, then the cadastral database should also be prepared for the dynamic if it still wants to be sustained in the future. To support this dynamic datum, the cadastral block adjustment should be aligned with new and more precise coordinate system and there is the ability to be transformed to any new coordinate system. Thus, the complete information on the control points is necessary and it is recommended as the study conducted by Gill et al., (2017). In general, Gill et al., (2017) agree with the existing guideline for the use of GNSS in cadastral surveying and recommend improvements as follows.

### Table 2: Additional GNSS Applications Guideline in Cadastral Survey (Gill et al., 2017)

<table>
<thead>
<tr>
<th>Issue</th>
<th>Suggestions</th>
</tr>
</thead>
</table>
| The reliability of N-RTK | i) Establishment of more stations: to reduce Inter-CORS distances to accommodate better accuracy of network corrections wherever within the network.  
ii) Improve network corrections by introducing regional atmospheric models.  
iii) N-RTK techniques are unable to provide proper quality control, as N-RTK may provide incorrect ambiguity fix and dependent on time of day and network geometry. Hence, rapid static techniques should still be preferred over N-RTK. |
| GNSS instrument calibration | i) The best verification method is still static observations connected to CORS with post-processed double differenced observables.  
| GNSS reference station accuracy | i) Update GDM2000 regularly in order to maintain the alignment of the national geocentric datum to the international standard, i.e., ITRF. |
Previous research by Norshahrizan et al., (2017) found that the standard deviation of area comparison for angular data is 0.254 which is smaller than the bearing data 0.292. The result proved that the angular based data adjustment can retain the area which proved the geometry of boundary marks are preserved.

Table 3. Coordinate and Displacement of Angular Based Adjustment (Norshahrizan et al., 2017)

<table>
<thead>
<tr>
<th>BM ID</th>
<th>NDCDB COORDINATE</th>
<th>PAI COORDINATE</th>
<th>DIFFERENCE</th>
<th>DISPLACEMENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>5219907465</td>
<td>15222.438 -60744.031 15222.440 -60743.998</td>
<td>0.002 0.003</td>
<td>0.003</td>
<td></td>
</tr>
<tr>
<td>522207765</td>
<td>15225.324 -60775.545 15225.324 -60775.555</td>
<td>0.000 0.004</td>
<td>0.004</td>
<td></td>
</tr>
<tr>
<td>528097635</td>
<td>15283.431 -60757.879 15283.442 -60757.869</td>
<td>0.011 0.010</td>
<td>0.015</td>
<td></td>
</tr>
<tr>
<td>5222507761</td>
<td>15225.294 -60773.583 15225.294 -60773.587</td>
<td>0.000 0.004</td>
<td>0.004</td>
<td></td>
</tr>
</tbody>
</table>

For the total 200 points of the boundary marks, it shows that 100% are under 0.050m where the lowest and highest values of 0.000 m and 0.033 m respectively as shown in Table 3. The standard deviation of coordinate difference is 0.009m. Based on the result, it indicates that angular based data adjustment yielded datasets that are within tolerance level compared to the NDCDB coordinate and at the same time the area and geometry preservation of original parcel is improved.

The last phase is the system weightage which appropriate with raw data for block adjustment. The weightage system is essential to ensure that the degree of permissible errors on cadastral database can be controlled. In the initial stage, it is suggested that any raw data failing the global test will be given the appropriate flexibility weightage to ensure that the initialization adjustment process can be performed. Furthermore, when there are new high accuracy measurements, it will be used to replace the legacy data which low accuracy and precision. Based on the increasingly new measurements from time to time, this will make new additional measurements of high accuracy data will be more well distribute within the cadastral block network. This is indirectly, enabling the block network to be more stable as well as enhanced numbers of high precision and high accuracy data in the database network.

4. CONCLUSION

This study proposes a new data preparation for PAI towards creating a comprehensive transformation system in converting legacy cadastral datasets to more accurate new dataset. The study outlines the three main components, datum traceability, angular based adjustment and weightage systems. This framework suggests conversion of bearing to angle for reducing the error propagation in a block adjustment where this effort can implement the PAI process without entirely relying on new measurements. Based on the proposed methodology and earlier evidence of the previous study, this method is expected to produce more flexible and economical in preparation of data PAI in the effort to transform cadastral database from low accuracy legacy datasets to high accuracy datasets and indirectly help agencies manage databases with more efficient. In order to strengthen this PAI process, a detailed study of the stochastic model for determining appropriate weightage for each observation is required to be proposed in the future.

ACKNOWLEDGEMENTS

The author would like to thank and express sincere appreciation to DSM State of Johor for providing valuable data and information regarding to the research done and Triple Axis for the funding.

REFERENCES


Revised August 2019