Positional Accuracy Improvement for Heterogeneous Geodata Integration

Thi Bich Phuong DONG, Vietnam

Key words: Positional accuracy improvement, integration heterogeneous geodata, cadastral, rubbersheeting

SUMMARY

This paper clarifies the role of positional accuracy improvement (PAI) in heterogeneous geodata integration in Vietnam and proposes an improvement workflow for cadastral data to fit with more accurate national geographic data using rubbersheeting transformation. Test results of enhancement planar accuracy of cadastral map 1: 10 000 according to national geographic data of same scale and higher accuracy demonstrate that positional accuracy improvement workflow can be used in integrating multi-source data and enabling utilization of map legacy.

SUMMARY

Bài báo làm rõ vai trò của cải thiện độ chính xác vị trí không gian dữ liệu địa lý trong tích hợp dữ liệu đa nguồn gốc và đề xuất quy trình chuẩn hóa dữ liệu địa chính để tích hợp với dữ liệu nền địa lý có độ chính xác cao hơn sử dụng phép chuyển đổi rubbersheeting. Kết quả thử nghiệm cải thiện độ chính xác vị trí không gian dữ liệu địa chính tỷ lệ 1: 10 000 theo CSDL nền địa lý cùng tỷ lệ chứng minh quy trình cải thiện độ chính xác vị trí không gian có thể sử dụng trong tích hợp dữ liệu đa nguồn gốc và sử dụng hiệu quả di sản bản đồ, dữ liệu địa lý

Positional Accuracy Improvement for Heterogeneous Geodata Integration

Thi Bich Phuong DONG, Vietnam

1. INTRODUCTION

As countries have completed high-precision seamless topodata as the reference frame for thematic applications, the usage of map and geodata legacy requires improved positional accuracy of the objects which has been captured or shown in the database or geographic map (Rönsdorf, 2005). The Ordnance Survey (UK), the Bureau of Census (USA) and Swisstopo are the National Map Agencies (NMAs) pioneering in research and application of PAI techniques for spatial integration multi-source, multi-time data (Belussi and Migliorini, 2012). In United States, the most important PAI program had conducted by the Census Bureau as part of the MAF / TIGER Project. This program aims to improve the accuracy of the TIGER (Topologically Integrated Geographic Encoding and Referencing System) database with a reported error of up to 150 meters over the results of the Differential-GPS test. The result of the program has increased the spatial accuracy of TIGER data to reach RMSE 3.8m after PA1.

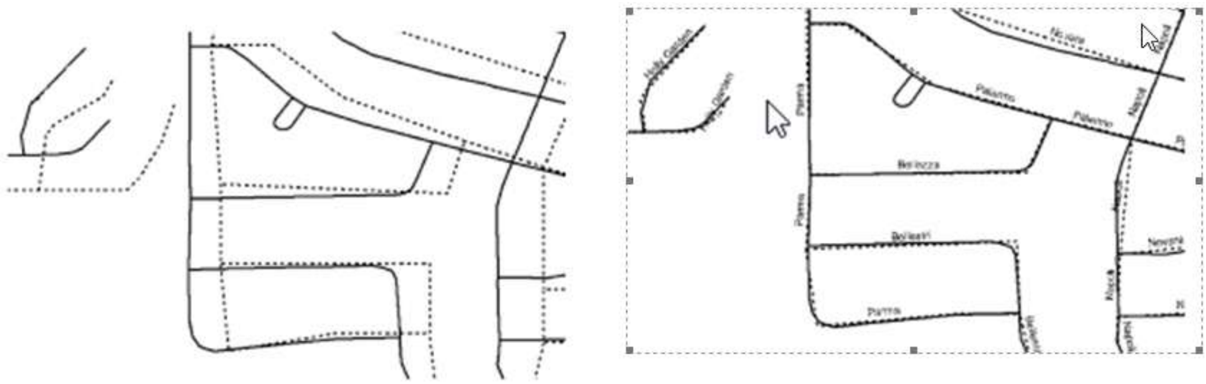


Figure 1. Improved TIGER road network data accuracy (solid line) according to Department of Transportation data (dotted line) [Shekhar, 2007].

Ordnance Survey's Positional Accuracy Improvement Programme was introduced in (Ordnance Survey, 2016). While GPS technology allows locating objects with accuracy up to decimeters, the absolute position of the objects on the maps provided by the Ordnance Survey includes LandLine and OSMaster Map at a scale of 1: 2500. RMSE is up to 2.8m. However, this system has a low relative accuracy of 1.2 m. In a program lasting from 1999 to 2006, the Ordnance Survey had improved the accuracy of the 1: 2500 maps using GPS results, with Absolute Positional RMSE of 1.1 m in rural areas and 0.4 m in urban areas, and relative error 0.4 - 1.0m.

2. RUBBERSHEETING TRANSFORMATION

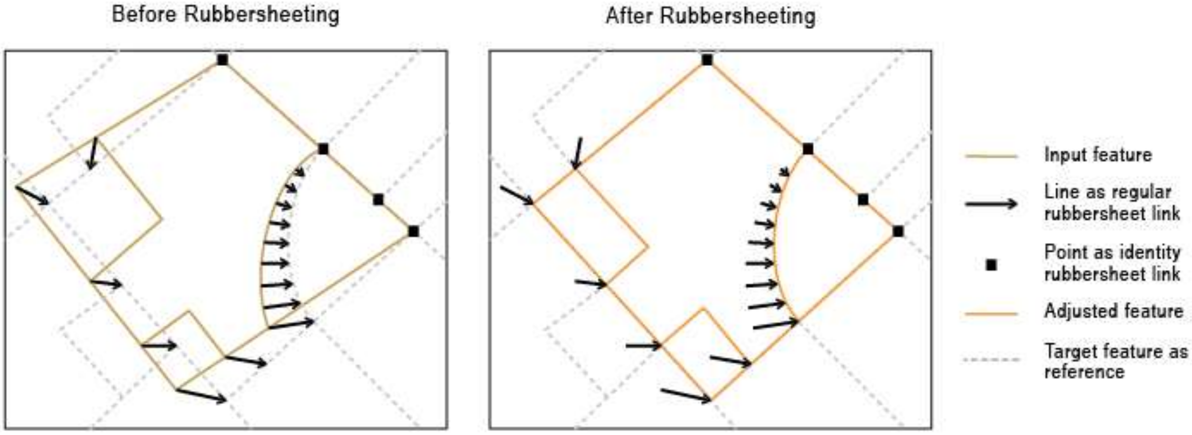


Figure 2. Conversion of Rubbersheeting for vector-2-vector integration (ArcGIS Pro Tool Reference)

Geographic data integration in GIS systems is the process of processing different data sources to achieve accurate data, minimizing redundancy and resolving spatial conflict. Requirements for geographic data integration also appear during data updating or integrating raster data with vector data. The geometric properties of features in geographic information systems like digital maps, databases, are mainly described by point coordinates in a given reference system. In fact, the coordinates of features are not directly measurable but always are the result of computation. Improved spatial position accuracy can be considered as a transition, defining mathematical rules that allow correction and recalculation of pairs of source data coordinates based on tie points coordinates.

Accuracy of the map object location is influenced by many factors, such as projection distortion, error of measurement, acquisition, editing .... and can only be controlled by the error margin in technical regulations. Rubbersheeting transformations are often used to perform coordinate transformations, improve location accuracy, integrate data and map legacy with data sources of higher precision. For example, Doytsher and Gelbman (1995) used Rubbersheeting to improve digitizing cadastral maps. Hashim et all (2016) proposed improve cadastral legacy of Malaysia according to modernized reference geodata by a chain of transformation including Rubbersheeting.

Unlike conventional transformations in image processing and map digitizing, rubbersheeting are a set of local transformations based on the Delaunay Triangle Network. Z value at each triangle node is the difference in coordinates δX, δY of the linked points between the source and destination data. Within each triangle, the difference in coordinates δX_i, δY_i is interpolated by one of the common transformation algorithms, such as the Helmert or Affine transformation.

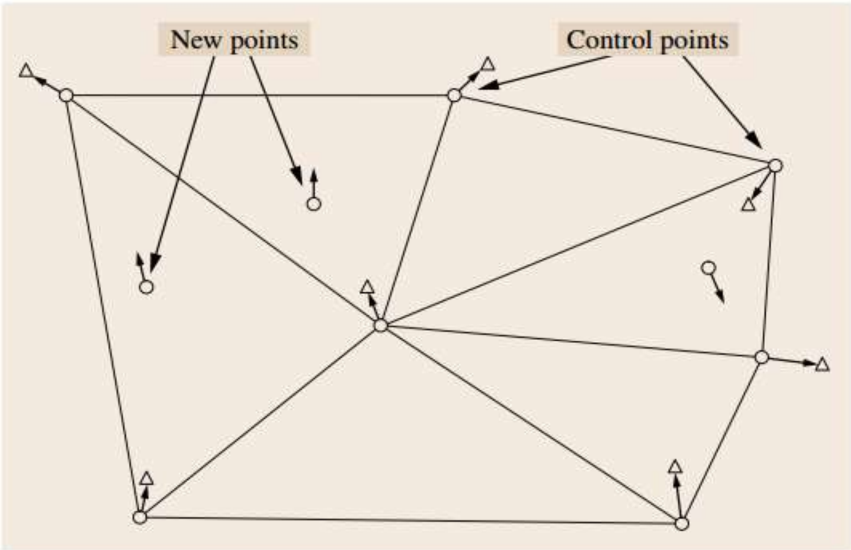


Figure 3. Rubbersheeting transformations (Kresse and Danko, 2012)

3. POSITIONAL ACCURACY IMPROVEMENT FOR CADASTRAL MAP LEGACY

In Vietnam, cadastral maps are constructed at scales from 1: 200 to 1: 10 000. Important resources in land administration are cadastral maps established between 1999 and 2009. Basic cadastral map of 1: 10 000 was established for forest land with factors related to forest land boundaries such as topography, traffic, irrigation, dividing lines... in a less detailed level than topographic maps of the same scale.

Regular cadastral maps are available at a scale from 1: 200 to 1: 5 000. In urban areas, maps and cadastral data generally have a larger scale than national topographic maps. They are used as a basis for the establishment of public transport maps, underground works and urban planning. Accuracy of absolute and relative locations of objects on these thematic maps depends entirely on the cadastral map accuracy. Due to long term acquisition, various technologies and different technical regulations, maps and cadastral data have inconsistencies in positional accuracy as well as geographic reliability.

Table 1 below shows a comparison of the positional accuracy between the national geographic database 1:10 000 and digital cadastral maps established according to 1999 and 2008 regulations.

Table 1. Comparison of regulations on accuracy of locating objects on digital cadastral map and geographic database 1:10 000

Accuracy	Cadastral map 1999	Cadastral map 2008	National geographic database
RMSE of control points	0.1mm-0.15mm x M	0.1mm-0.15mm x M	None

RMSE of parcel points	0.6mm x M	0.6mm -0.8mm x M	0.5-0.7 mm x M
RMSE of other points	0.9 mm x M	0.5mm - 0.75mm x M	0.5-0.7 mm x M
Relative error	0.4 mm x M	0.7mm -0.9mm x M	None

For effective use of cadastral and related maps legacy should include solutions to improve spatial location and integrate cadastral data with reference geographic data. There is a real need of integration cadastral data with more accurate geographical data for local land administration and building LIS. In the case of cadastral maps, obligatory condition is that parcel areas and segment lengths left unchanged.

4. INTEGRATION CADASTRAL DATA AND NATIONAL GEOGRAPHICAL DATABASE

Based on the analysis of regulations on data accuracy for cadastral maps, land database and geographic database in recent years and the actual requirements for integrating these important data sources, there is a proposal for improvement workflow of cadastral data according to national geographic data of the same scale. The general procedure is shown in Figure 4 and main steps are as followed:

Step 1. Material and data evaluation

Material and data evaluation should be based on the overlaying of cadastral data and geographical data, documents related to the process of establishment of both types of data such as technical design, metadata data, check records acceptance... Points of difference beyond the allowable limit are required to be checked by the reference document like remote sensing images and field surveys.

Step 2. Selection of linked objects.

Boundary objects, hydrography, transport and elevation points are selected for link creation. Distance between two identical points must be smaller than the limit positional error of source dataset.

Step 3. Link creation

Link creation is the task of defining the relationship between source and destination coordinate pairs. The important parameter in link creation is Search Distance. In order to control the displacement of parcel points and the location of other objects, the search radius must be smaller than the limit positional error the plane position of target dataset. In case of national geographic database, limit plane positional error is set as 3* RMSE

Step 4. Improvement.

After checking, links are used to improve the location of the original data.

Step 5. Evaluation

Overlay datasets is checked in the right scale with the naked eye, statistics on topology errors, points in excess of permissible errors. Defining the maximum and average parcel vertex displacement, maximum, average and cumulative parcel area change, maximum and average parcel segment change.

Steps from 2 to 5 are repeated if needed.

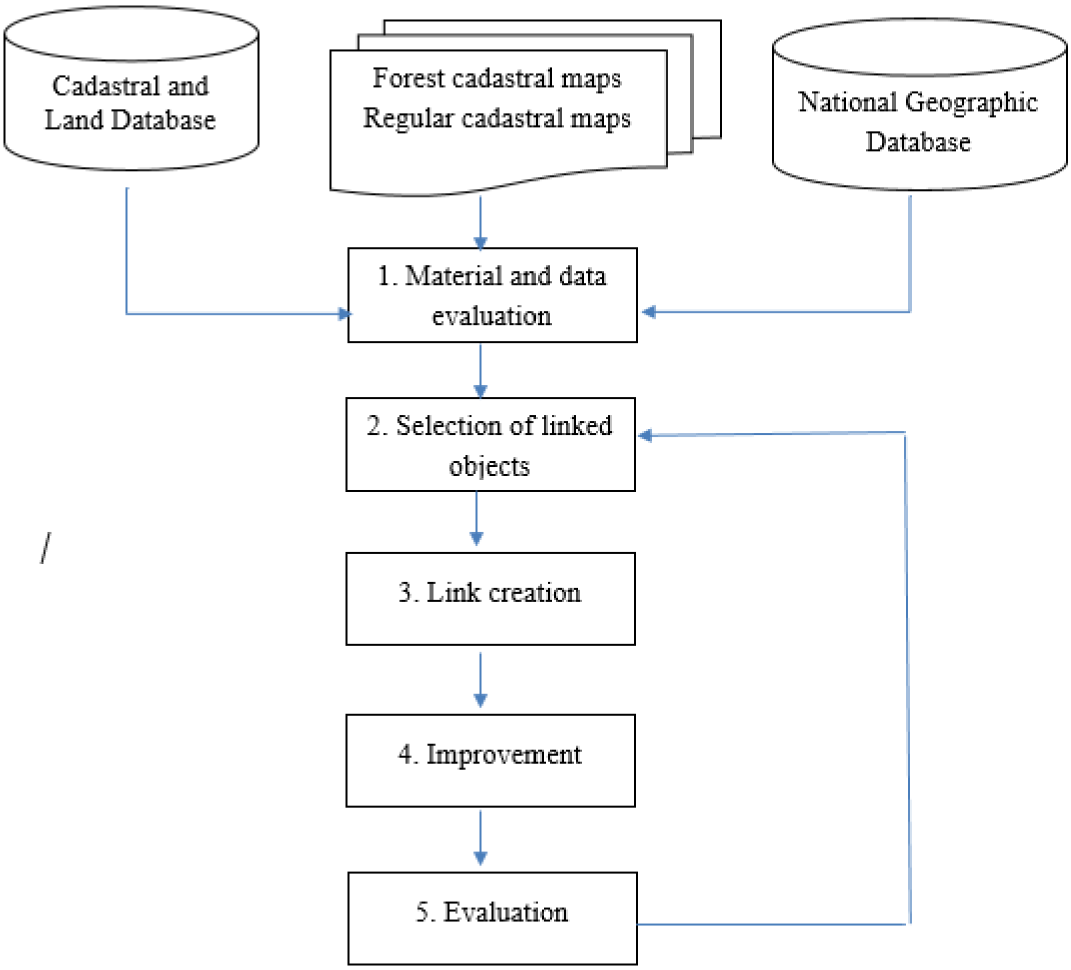


Figure 4. Positional accuracy improvement workflow for cadastral data

5. TEST

Data used for the experiment are 1: 10,000 cadastral map established in 2004 of Yen Lang commune, Dai Tu district, Thai Nguyen province and national geographic database of Thai Nguyen province at the scale of 1: 10 000 (2012). The tool used for improvement is the ArcGIS 10.2 Conflation Toolkit with the Generate RubberSheet Links and Rubbersheet Features tools.

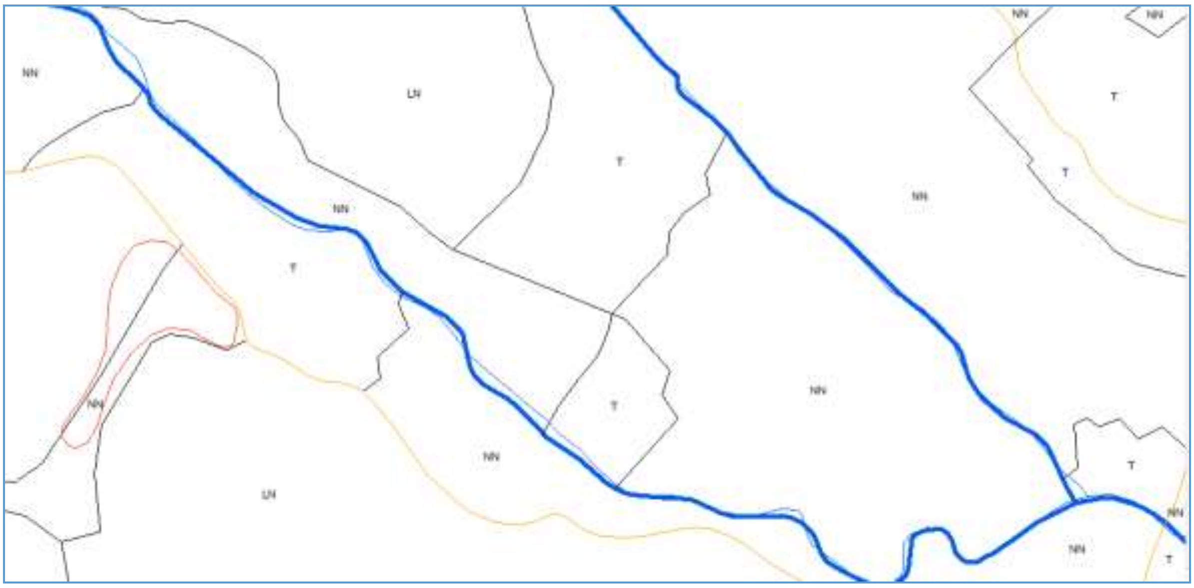


Figure 5. Overlay cadastral and national geographic dataset

Overlaying and comparing source and target show that hydrography and traffic features in two data files are not identical. The difference can be seen with the naked eye at a scale of 1: 10 000 in Figure 5, where two sets of rivers are showing by blue polylines .

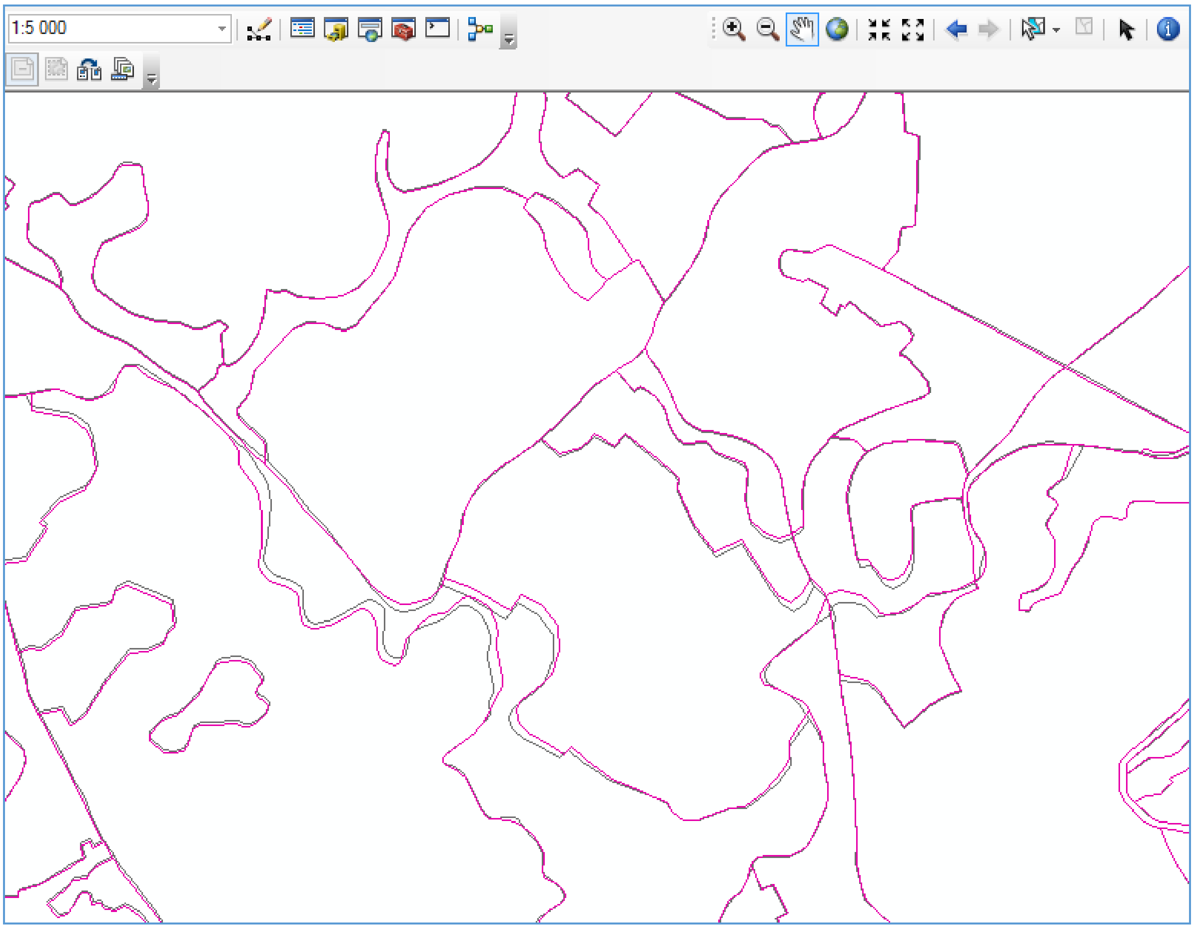
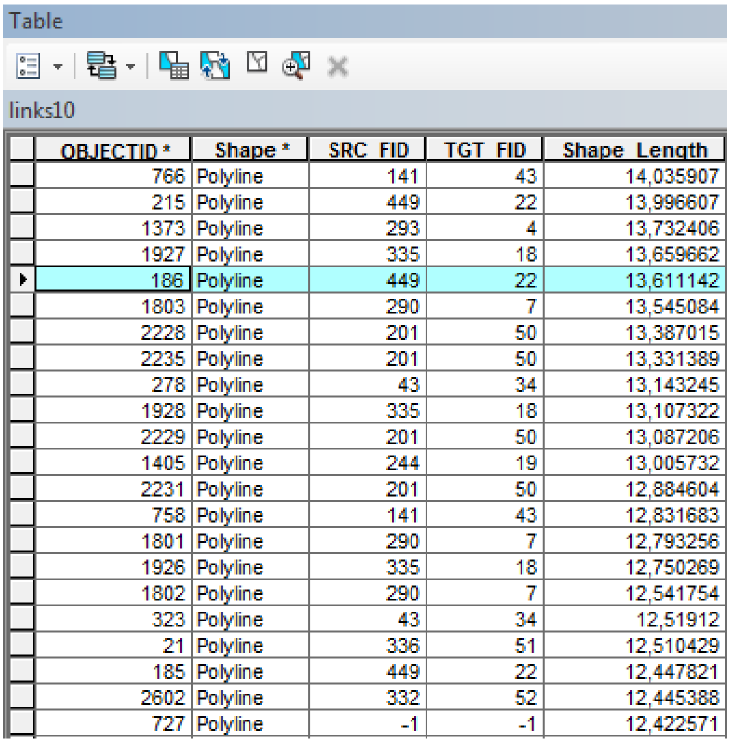


Figure 6. Overlay of land parcel boundaries before (purple) and after (black line) PAI at 200% magnification

Land features, roads, rivers and elevated elevation points are selected for links using Search radius of 10m. Links are used to improve each data layer of the original map with Rubbersheet Features tool.

Search Distance choice is experiment-based, in this case Search Distance is set as 2*RMSE. With Search Distance = 10m test result at 100% magnification shows no topology and exceeding the limit error. Rivers, roads, boundaries, and elevation are completely identical. Overlay of source and target sets at 200% magnification is showed in Figure 6.



Table

links10

OBJECTID *	Shape *	SRC_FID	TGT_FID	Shape_Length
766	Polyline	141	43	14,035907
215	Polyline	449	22	13,996607
1373	Polyline	293	4	13,732406
1927	Polyline	335	18	13,659662
186	Polyline	449	22	13,611142
1803	Polyline	290	7	13,545084
2228	Polyline	201	50	13,387015
2235	Polyline	201	50	13,331389
278	Polyline	43	34	13,143245
1928	Polyline	335	18	13,107322
2229	Polyline	201	50	13,087206
1405	Polyline	244	19	13,005732
2231	Polyline	201	50	12,884604
758	Polyline	141	43	12,831683
1801	Polyline	290	7	12,793256
1926	Polyline	335	18	12,750269
1802	Polyline	290	7	12,541754
323	Polyline	43	34	12,51912
21	Polyline	336	51	12,510429
185	Polyline	449	22	12,447821
2602	Polyline	332	52	12,445388
727	Polyline	-1	-1	12,422571

Figure 7. Parcel points displacement control by length of linked edges

Table 2. Comparison of parcel area before and after PAI

Code	**Before PAI**	**After PAI**	**Difference**
LN	6602487.226	6602487.226	0.00000131316
LN	3841426.715	3841426.715	-0.00000466779
NN	13739.975	13739.975	-0.00000004532
NN	11875.000	11875.000	0.00000001679
NN	4346.122	4346.122	0.00000000383
NN	5212.040	5212.040	-0.00000000034
T	22054.178	22054.178	-0.00000001482
T	20849.628	20849.628	-0.00000004737
NN	548.699	548.699	-0.00000000006
NN	8815.387	8815.387	-0.00000000089
NN	5380.908	5380.908	-0.00000000206

NN	4220.364	4220.364	-0.00000000012
NN	16117.061	16117.061	0.00000004433

The maximum point shift is controlled by the Search Distance when creating the link. With Search Distance = 10m, the maximum point shift is 15m (Figure 7).

Comparison before and after PAI demonstrated that parcel areas changed completely negligible (Table 2). Maximum land area change does not exceed 4.10^{-6} m^2, what in practice can be considered unchanged and completely acceptable.

Table 3. Comparison of edge length before and after PAI

Before PAI	**After PAI**	**Difference**
450.478416	450.478416	-2.2078E-10
133.5717059	133.5717059	-3.8338E-10
262.1534668	262.1534668	2.62048E-10
124.4946157	124.4946157	-1.16472E-10
114.5228538	114.5228538	-3.54561E-11
285.7205002	285.7205002	2.18279E-10
198.3007729	198.3007729	-4.41673E-11
322.6371061	322.6371061	3.84432E-10
67.29056393	67.29056393	3.38218E-11
204.4331488	204.4331488	3.73518E-10
221.1031837	221.1031837	-2.18989E-10
240.156492	240.156492	-4.35818E-10
212.0106365	212.0106365	-2.69722E-11
132.6158322	132.6158322	-3.31994E-10
17.75472943	17.75472943	-2.8912E-11
640.5861061	640.5861061	-2.57728E-10
93.90648061	93.90648061	-2.16147E-11
249.4654547	249.4654547	1.32104E-10

235.1753461	235.1753461	4.47955E-10
178.2299235	178.2299235	-3.11275E-10

Comparison of parcel segment length after PAI shows that the length of the parcel edges has changed completely negligible. At the trial site, the maximum edge distance shift does not exceed 5.10^{-9} m as showing in Table 3. In practice it can be considered unchanged and completely acceptable.

6. DISCUSSION AND CONCLUSION

Test results show that PAI workflow improves positional accuracy of cadastral data according to reference geographic data of the same scale, keeping parcel areas and segment lengths unchanged.

The proposed workflow of improving positional accuracy of geodata can be applied when the cadastral data has a lower plane position accuracy than the national geographic database. The workflow can be used in the development of a land database to minimize the cost of standardization work. The general procedure may be applied to improve thematic geospatial data accuracy according to data with higher accuracy.

This research is funded by Vietnam Ministry of Natural Resources and Environment.

REFERENCES

1. Belussi, A., & Migliorini, S. (2012). A framework for integrating multi-accuracy spatial data in geographical applications. GeoInformatica Volume 16, Issue 3
2. Doytsher Y., Gelbman E (1995). Rubber-Sheeting Algorithm for Cadastral Maps. Journal of Surveying Engineering 121(4).
3. Hashim N. M., Omar A. H., Omar, K. M., Abdullah N. M.,Yatim H. M. (2016). Cadastral Positioning Accuracy Improvement: A Case Study In Malaysia. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences.
4. Kresse, W. D., Danko, D.M. (2012). Handbook of Geographic Information. Springer.
5. MAF/TIGER Database. US Census Bureau. https://www.census.gov/geo/reference/gtc/gtc_maftiger.html
6. Ordnance Survey (2016). Positional Accuracy Improvement Programme (www.ordnancesurvey.co.uk/PAI).
7. Rönsdorf, C. (2005). Positional accuracy and integration of geographic data. Coordinates, Volume I, Issue 4, Sep.
8. Shekhar, S. (2007). Encyclopedia of GIS. Springer.
9. ArcGIS Pro Tool Reference. http://pro.arcgis.com/en/pro-app/tool-reference/editing/rubbersheet-features.htm

BIOGRAPHICAL NOTES

Dr. DONG Thi Bich Phuong works in Vietnam Institute of Geodesy and Cartography as Head of Department of Cartography and GIS. She is a member of Vietnam Association of Geodesy - Cartography - Remote Sensing. She competed her undergaduate and PhD degree in Russia. Her research interest includes mathematic cartography, automatic generalisation and spatial data integration.

CONTACTS

Dr. DONG Thi Bich Phuong
Vietnam Institute of Geodesy and Cartography
Address 479 Hoang Quoc Viet Road, Cau Giay District
City: Ha Noi
VIETNAM
Tel. +84912482545
Email:dbphuong@yahoo.com
Web site: vigac.vn