Estimating Positional Accuracy of Linear Features

Andriani Skopeliti
Surveying Engineer

Lysandros Tsoulos
Assistant Professor NTUA

Faculty of Rural and Surveying Engineering - Cartography Laboratory
National Technical University of Athens
H. Polytechniou 9, 157 80 Zographou Campus, Athens, Greece
Tel: +30 +1+772-2730 Fax: +30+1+772-2734
Email: askop@central.ntua.gr, lysandro@central.ntua.gr

Positional accuracy is considered as one of the main elements of spatial data quality. For linear features, it can be analyzed to horizontal position accuracy, vertical position accuracy and shape fidelity. In this paper, the assessment of linear features' positional accuracy due to data acquisition (digitization) and generalization is studied. For the study of the results of the above-mentioned cartographic processes, the change in horizontal position can be measured by displacement measures and the change of line shape can be evaluated through measures describing line character. In each phase, experimental analysis is conducted to identify the relationship between the amount of error and the shape of the cartographic line/segment. The analysis is based on the results of the methodology for the parametric description of the shape of cartographic lines and the segmentation of cartographic lines in homogeneous parts, developed by the authors. The linear features used in the experiment initially undergo the segmentation and line complexity identification procedures. Linear features segments are grouped into clusters with regard to their complexity. These groups, representing different degrees of complexity, are utilized for the estimation of the data acquisition and generalization error.

1. INTRODUCTION

Cartographic features' accuracy, is the result of a number of cartographic transformations such as generalization, projection change etc. These transformations have an overall influence on map reliability, which can be assessed through the development of a model quantifying the relevant error components. The study elaborated here, focuses mainly on two components of positional accuracy, as they are affected by cartographic transformations: horizontal position accuracy and the preservation of shape. An experiment is carried out to qualify and quantify the influence of data acquisition on positional accuracy due to manual digitization and the consequences of cartographic generalization performed with the use of the simplification operator.

2. METHODOLOGICAL APPROACH

2.1 SEGMENTATION

In order to control the behavior of the cartographic line, it has to be segmented in homogeneous segments. Segmentation is accomplished through a methodology, which refers to natural features and is based mainly on fractal dimension variation (Skopeliti and Tsoulos, 1999). This methodology is implemented in three steps:

a. Self - similar segments identification along the linear feature

b. Formation of groups of self - similar segments using cluster analysis

c. Selection of a representative segment for each group of segments (the one with fractal dimension value closer to the average value of the cluster).

When the partitioning in homogeneous segments is completed, the linear segments are grouped in similar shape clusters, based on the values of three parameters used for their description: average magnitude angularity, error variance and the ratio of length and the base line length (Skopeliti and Tsoulos, 2000). These parameters are calculated for each line segment and hierarchical cluster analysis is conducted to identify the number of data groups with the same character. Finally, line segments are classified with non- hierarchical analysis.

2.2 ASSESSMENT OF LINEAR FEATURES POSITIONAL ACCURACY

Linear features positional accuracy assessment is based on shape preservation and horizontal position accuracy. Positional accuracy is measured through comparison with high quality data. For the estimation of positional accuracy due to generalization, generalized data are compared with the initial data, whereas for the estimation of data acquisition accuracy, manually digitized data are compared with data collected through scanning. Data, which are used in the experiment, will be referred to from now on as "test data".

Shape modification is examined (Tsoulos and Skopeliti, 2000a; Tsoulos and Skopeliti, 2000b):

- at the line segment level, by observing the change in the classification of test data segments applying non - hierarchical classification with the clusters centers of the reference data (modification compared to initial shape) and hierarchical classification (inter - cluster modification),
- at the line group level, by observing the cluster center change. The distance, between the cluster centers of the "test data" and the reference data, measures the cluster center change. This last measure is objective but can be used only for comparison.

Horizontal position accuracy is measured by the deviation from the position of the reference data. It can be assessed by distance measures such as average Euclidean distance, Hausdorff distance (Abbas et al., 1995), and areal displacement measures such as the ratio of the area between the original and the generalized line to the length of the original line (McMaster, 1987). The values of these measures can be judged on the basis of the map scale.

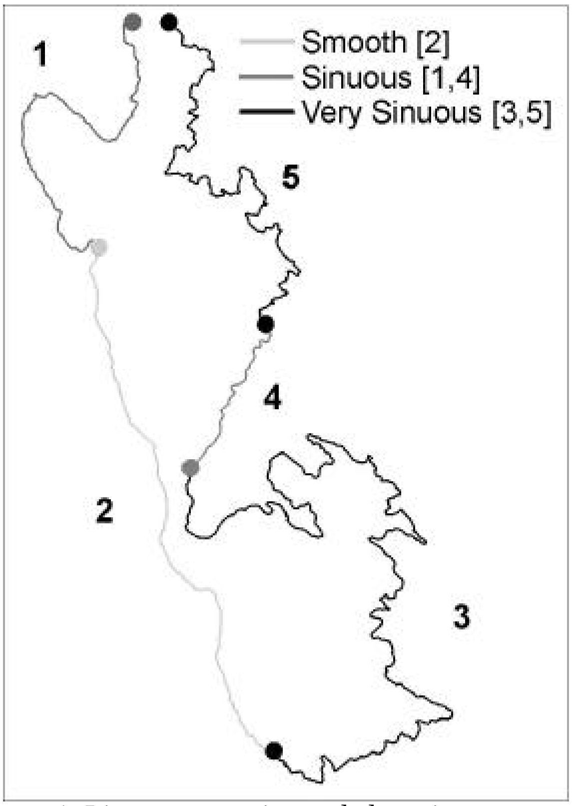

Figure 1. Line segmentation and clustering.

	Line Groups	Line Code
1	S	2
2	SIN	1
		4
3	VSIN	3
		5

Table 1. Segments Classification.

Line Code	Haussdorf distance (meters)	Group
1	24.51	SIN
2	18.53	S
3	37.73	VSIN
4	23.4	SIN
5	35.97	VSIN

Table 2. Heads–up digitization results.

3. CASE STUDY

The data set used in the experiment, is the coastline of the Greek island Ithaki (Figure 1), digitized from a 1:100 000 scale map. The specific line was selected due to its complex configuration. The coastline is segmented into five (5) segments (Figure 1), which are classified into three (3) groups according to their shape: "smooth" (S), "sinuous" (SIN) and "very sinuous" (VSIN) (Table 1).

3.3 DATA ACQUISITION

An experiment was conducted to study the results of heads-up digitization, a well-known method of data encoding. Heads-up digitization enables the operator to zoom into the line and thus acquire a detailed encoding of the line footprint. On the other hand, this facility can cause problems, if the operator uses a zoom factor greater than the data resolution implied by the map scale. This leads to the collection of an excessive amount of vertices, which does not always result to a more accurate recording.

Heads up digitization was conducted by eighteen [18] different operators in the AutoCAD environment, using a color raster image of the source map in the background. Instructions were given to the operators working in the CAD environment to avoid vertices congestion (procedural knowledge). The number of operators involved in the experiment although is not big enough, does not influence the validity of the process. The results will be used as indicators of the operators' attitude towards line complexity.

The line segments recorded through heads up digitization, are classified into three groups utilizing cluster analysis. They have the same synthesis with the clusters of the initial segments. The same synthesis is preserved even when these line segments are classified using non - hierarchical analysis and the clusters centers of the nominal segments. It becomes evident that noticeable changes of line shape are not present.

The Hausdorff distance measures the positional error of linear segments. From Table 2, it becomes apparent that lines belonging to the group of very sinuous segments exhibit the greatest horizontal deviation. On the other hand, sinuous lines exhibit moderate values and smooth line segments exhibit smaller values. It can be concluded that error in data acquisition measured with the Hausdorff distance follows the pattern of the line classification in three groups based on their complexity. The operators participated in the experiment, handled the three line categories - which were identified through parametric line description and cluster analysis - differently. Taking into account the source map scale, it is estimated that positional error is close to the legibility threshold and no gross errors exist.

3.4 GENERALIZATION

Generalization of linear segments is carried out with the utilization of the most commonly used simplification algorithms such as: Douglas Peucker (DP), Reuman – Witkam (RW) and Euclidean distance (ED), for a number of scales. The relevant parameter values are selected equal to the legibility threshold (estimated to be 0.25 mm), at each scale.

The shape of each generalized segment is compared with the segments belonging to different groups, through hierarchical classification (Table 3). Three different cases are observed in relation to the change of the parameter value: a. membership in the three initial groups is retained, b. membership in the three initial groups changes and c. classification leads to the formation of two groups. Shape modification for each segment compared to the initial segments groups, is examined through non-hierarchical classification of generalized lines using the cluster centers of the initial lines. It becomes apparent (Table 4), that generalized lines become smoother. They are classified to the "smooth" group and the "sinuous" group of the initial line segments. The "very sinuous" group does shades off, after a certain parameter value, different for each algorithm.

The effect of simplification algorithms on the horizontal position is measured by the average Euclidean distance between the initial and the generalized line. Figures 2 and 3, illustrate the relationship between horizontal deviation and shape modification as results of generalization for each algorithm and line segment group. As it is derived from figures 2 and 3:

- Shape modification of the “smooth group” is temperate for all algorithms.
- The effect on positional accuracy caused by the Euclidean algorithm is more gentle compared to that of Douglas Peucker. The Reuman – Witkam algorithm causes larger values in horizontal deviation.
- Shape modification of the “smooth group” is not significant using the Euclidean Distance algorithm.
- Shape modification for the sinuous and very sinuous group, is similar for all algorithms applied. Up to the point where the simplification parameter takes the value 75, the modification in both axes of the “sinuous” group is discrete and temperate compared to the “very sinuous” group (Figure 2). For simplification parameter values larger than 100, modification for the "sinuous" and "very sinuous" group is very similar (Figure 3).

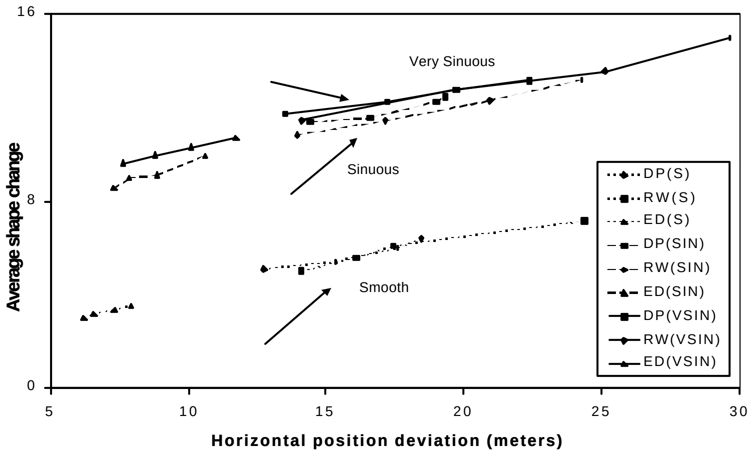

Figure 2. Horizontal position deviation and average shape change for simplification parameter values: 37.5, 50, 62.5 and 75 meters.

Algorithms Parameters Values (meters)	Line Groups	Line Code	Simplification Algorithms		
			DP	ED	RW
37.5-125	S	2	1	1	1
	SIN	1	2	2	2
		4	2	2	2
	VSIN	3	3	3	3
		5	3	3	3
187.5	S	2	1	1	1
	SIN	2	2	2	2
		4	2	2	1
	VSIN	3	3	3	3
		5	3	3	3
250	S	1	1	1	1
	SIN	2	2	2	1
		4	1	2	1
	VSIN	3	3	3	3
		5	3	3	3

Table 3. Hierarchical classification of simplified lines.

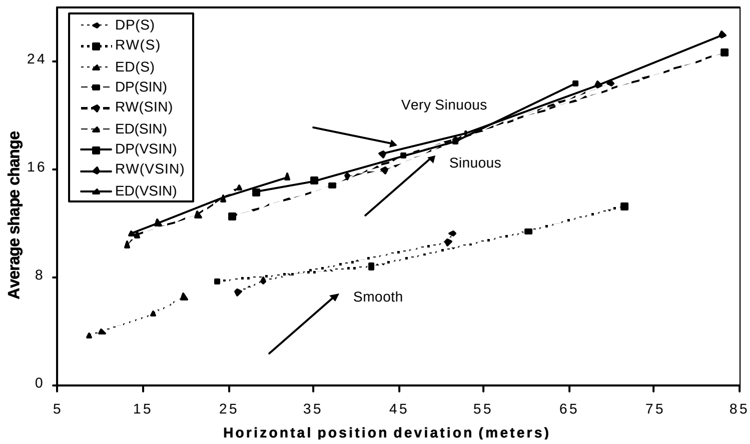

Figure 3. Horizontal position divergence and average shape change for simplification parameter values: 100, 125, 187.5 and 250 meters.

Algorithms Parameters Values (meters)	Line Groups	Line Code	Simplification Algorithms		
			DP	ED	RW
37.5	S	2	1	1	1
	SIN	1	1	1	1
		4	1	1	1
	VSIN	3	2	2	2
		5	2	3	3
50 -100	S	2	1	1	1
	SIN	1	1	1	1
		4	1	1	1
	VSIN	3	2	2	2
		5	2	3	2
125-250	S	2	1	1	1
	SIN	1	1	1	1
		4	1	1	1
	VSIN	3	2	2	2
		5	2	2	2

Table 4. Non-hierarchical classification of simplified lines.

4. CONCLUSIONS

The above-described case study shows that the results of the simplification algorithms differ; depending on the algorithm/parameter value applied and the corresponding linear feature shape. The necessity for segmentation of the linear features to homogeneous segments - based on their shape - before generalization has become apparent. It has been shown that the proposed methodology can be used for the assessment of the consequences of a generalization schema, regarding the horizontal positional accuracy and the shape of linear features.

Data acquisition with manual digitization shows that there is a pattern between positional error due to manual digitization and linear features complexity. Thus in order to develop a valid model, segmentation of lines in homogeneous parts is a prerequisite.

This methodology for the assessment of the positional error due to cartographic generalization of linear features, can be applied independently of the generalization algorithms and operators involved. Comparison of the results between generalization operators, algorithms, parameters values and complexity of linear segments groups can be carried out. Thus, the suitability of generalization alternatives can be evaluated and knowledge can be acquired for the automation of cartographic generalization. Regarding positional error due to data acquisition, future research on this subject will involve the attempt for the development of the relevant model.

REFERENCES

Abbas, I., Grussenmeyer, P., Hottier, P., 1995, Controle de la planimetrie d' une base de donnes vectorielles: une nouvelle methode basee sur la distance de Hausdorff: la methode du controle lineaire, in: Bul. S.F.T.P. N°137 (1995-1), pp.6-11.

McMaster, R. B., 1987, Automated line generalization, Cartographica 24(2), pp.74-111.

Skopeliti, A., Tsoulos, L., 1999. On the parametric description of cartographic line shape. Cartographica 36 (3), pp. 57-69.

Tsoulos, L. and A. Skopeliti, 2000a. "Developing a model for quality assessment of linear features." In Accuracy 2000 Proceedings of the 4th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences, G.B.M. Heuvelink and M.J.P.M. Lemmens, eds., Amsterdam, The Netherlands, 655-658.

Tsoulos, L. and A. Skopeliti, 2000b. "Assessment of data acquisition error for linear features." In Proceedings of the XIXth ISPRS Congress, Amsterdam, The Netherlands, 1087-1091.