Mapping of the Sampling to Evaluate the Positional Accuracy in a Large Urban Area

Simone Sayuri Sato*, Claudomiro dos Santos**

* Universidade Federal de Pernambuco – UFPE
** Universidade de São Paulo - USP

Abstract. This work aims to present a part of methodology to selection of sample areas to evaluate the positional accuracy of information in maps with the scales 1:1,000 in a large extensions mapping area. The employed methods were validated in this area which is characterized by a high concentration of urban features and difficult accessibility for field survey. A major difficulty was the definition of the sample size and the selection of sample areas in accordance with the technical specification and the national standards, which made it possible to create a satisfactory solution for the producers, contractors, and end-user, called contractor. The importance of quality registers was considered and applied in each stage of the quality control processes and it was used as a support in quality assurance not only for documentation or archivation, but also on the basis of the defining criteria for evaluating the quality, especially in the choice and size in order to reduce the sampling uncertainty.

Keywords: quality criteria, sampling, quality register.

1. Introduction

The positional accuracy, created in 1984, is one of the quality elements which nowadays in Brazil supports official cartographic standards. This standard is only applied in the end of the cartographic production and the analogic format. This implies a great danger for the quality certifications when mapping large urban areas as it is causing catastrophic financial losses for everyone, especially when taking into account additional features such as high urban density, skyscrapers, assertive and intense traffic for filed surveying. Therefore it is a great challenge to control the positional quality applying this type of mapping as it is necessary to plan and to create a methodology, which evaluates those parameters, in order to be effective and efficient.

The evaluation method of the quality of the positional accuracy was created in the period of 2003-2005 while the digital mapping of São Paulo. The city is considered to be one of the world's largest metropolises with a total area of about 1.500 km^2. As described above, its map scale characteristic is 1:1,000 and 1:5,000. Moreover this mapping service was executed by a consortium of four different aerial survey companies. In the beginning the procedure was submitted in planning in order to evaluate the quality of the positional accuracy considering the roles, the quality parameters and the assessment process to accept or to reject the area of mapping delivery.

The standard assessment process of the positional accuracy, described in this work, consists in comparing the coordinates of well-defined points in the maps or the sample of the geographical data with the same measure points but obtained from an existing map which is more accurate or obtained from a specific field survey creating a set of quality control points. These points are measured with accuracy 3 to 4 times higher than the precision of the mapping.

To analyze the quality, which is the criterion to accept or reject the evaluated map, it is necessary to apply the assessment process based on the Nacional Cartographic Accuracy Standard. This process is dedicated to classify the position quality of the analogic maps by adapting the parameter class called PEC – cartographic standard accuracy – which is defined in the order-law number 89.817 of June 20, 1984. PEC is a statistical indicator (90% of probability) for horizontal accuracy, corresponding to 1.6449 times the RMSE (Root Mean Square Error) PEC = 1.6449 x RMSE). The standard error which is isolated in cartographic works will not exceed 60.8% of Cartographic Accuracy Standards. The quality is categorized in the classes A, B, C. This paper will only use the class A. The quality criteria of the first class were applied to the situation that in 90% of well-defined points selected on the map sheet and tested in the same field points, the difference of horizontal coordinates (E, N) should not exceed the error 0.5 mm of the map scale. Furthermore the standard error should not exceed 0.3 mm of the map scale. In this project the scale was 1:1000, what means that the residual values of the compared horizontal coordinates (E, N) should not exceed 50 centimeters. In the case of the PEC "Class A" concerning the Elevation (Z) to the map scale of 1:1.000 the contour lines are analyzed. So 90% of the isolated points on the contour lines, which are analyzed by comparing the measures of the map and differences of the elevation value in the same field points, may not introduce an error of half of the equidistance between the contours and 1/3 (one third) of this equidistance. It is finally the Standard Error.

In advance, it is known that the positional errors of collected elements result from the totality of the errors in previous phases of mapping, i.e., to guarantee the quality of mapping not in each step of the production but only in the end may be dangerous, as the mistakes or errors are caused by the poor specifications, errors of equipment which is used in each phase and also the inherent errors of the feature collection process may cause inaccuracies and faults. Hence, the mapping of São Paulo considered the control quality in each process. The evaluation and certification of the positional accuracy in the mapping of the total area - 1100 km² in the scale 1:1000 and 400km² in the scale 1:5000- should be a hard task. Therefore the article of Tastan (1999) was taken into consideration. The article describes two types of direct evaluation methods, the full inspection evaluation method and the sampling evaluation method. The full inspection evaluation method involves testing 100 per cent of the items in a population to determine a qualitative result. The sampling evaluation method involves testing only a sample of all the items in a population to determine a qualitative result. In the case of the mapping of São Paulo the inspection of samples was possible due to the application of quality control in each process by executing the checks and the indication of correction in each phase of mapping.

2. Methodology

2.1. Definition of the Sampling

The sampling, according to Meli (1975) in Ariza (2002) can be of different forms: discontinuous, percentile, statistical, having a random, systematic, stratified extraction, etc. The sampling can be developed for the quality control of mapping.

Due to the size and heterogeneity of the mapping area (1100 km² of urban area in the scale of 1: 1000 and 400 km² of rural area in the scale of 1: 5000), the choice to generate an oriented sampling was more effective than to generate a random one. The direction of the sample takes into consideration the following points: areas with higher density, elements and importance to the customer and the areas which had the highest error during previous phases. Stratified sampling, employed in many cases, is a well-known concept of statistics indicated by Freund & Simon (2000). The use of the Bayesian statistics concept is specific because it is established on the sample selection “based on the previous knowledge about a population" in accordance with Cavalcante & Zeppelini (2002).

It is neither common nor especially mentioned in the official national standards that the researches involve the statistic sample size to control the quality of mapping process or a guide that defines the conditions to assure quality. Since the beginning of the mapping process no technical papers or any referential information addressing this type of mapping have existed. It only had been applied to smaller map scales or in a single map (one sheet). Consequently, it is necessary to create a specific methodology, which can efficiently be applied to the mapping of 1100 km^2 and the map scale of 1:1000. This methodology was developed on this occasion.

2.2. Definition of the Basic Sampling Unit

First of all, it is necessary do select the basic sampling unit. The basic sampling unit is defined as a unit in which the quantitative and qualitative characteristics of the population are observed and measured. The sample consists of a set of sample units. Each sample unit produces a single observation of the variable of interest. The sample units can be plots of equal size, of a variable area or not superficial sample units as sampling lines or points. The definition of this basic sampling unit was considered as the real mainstay of the solution. The basic sampling unis was initially determined as a unit of an area defined in km^2. After that, certain units should be evaluated or tested. The basic sampling unit is finally composed of the set of the sampling units. The next step was to define the size of this basic sampling unit which had to be accepted by the customer and the technical team in order to implement the quality testing.

Another key point is to ensure the possibility to get a minimum number of inspection points measurable in the size of the basic sampling unit and in order to ensure statistical significance. Traditionally - based on Merchant (1982) – the number of points has been fixed on 20. This academic paper referred to the map scale of 1:20,000, which can be extrapolated to other larger scales and also considers that it is possible to identify those 20 points in a map sheet (unit of area basis).

Nonetheless, in practice it is not completely true, especially when it is applied in mapping using scales above 1:5000. Several analyzes which were carried out in the office as well as in field inspections resulted in the conclusion that it is not possible to find a number of adequate pints, proposed by Merchant, as the unit area had a size of about 0,33 km^2.

In this case, it is necessary to redefine the base units in the order to increase the area sample and also to create the possibility to select and find 20 points in this "new base units". The solution was to group 4 map sheets of the scale 1:1000 to form a square of approximately 1 km^2 of the area.

Once the base unit is defined and applied, the next task is to create a representative sample in which it is necessary to consider that the base unit must be applied in one of the areas with the same degree of probability as the positional error occurs. This was done applying a synthesis map of the probability that errors, showed in Figure 1, which are created by the grouping of classes, occur. This map was considered a solution to validate the positional accuracy of the big mapping area. The procedures to develop it was transformed a methodology of the sample selection that was created and will be mentioned in the next sections.

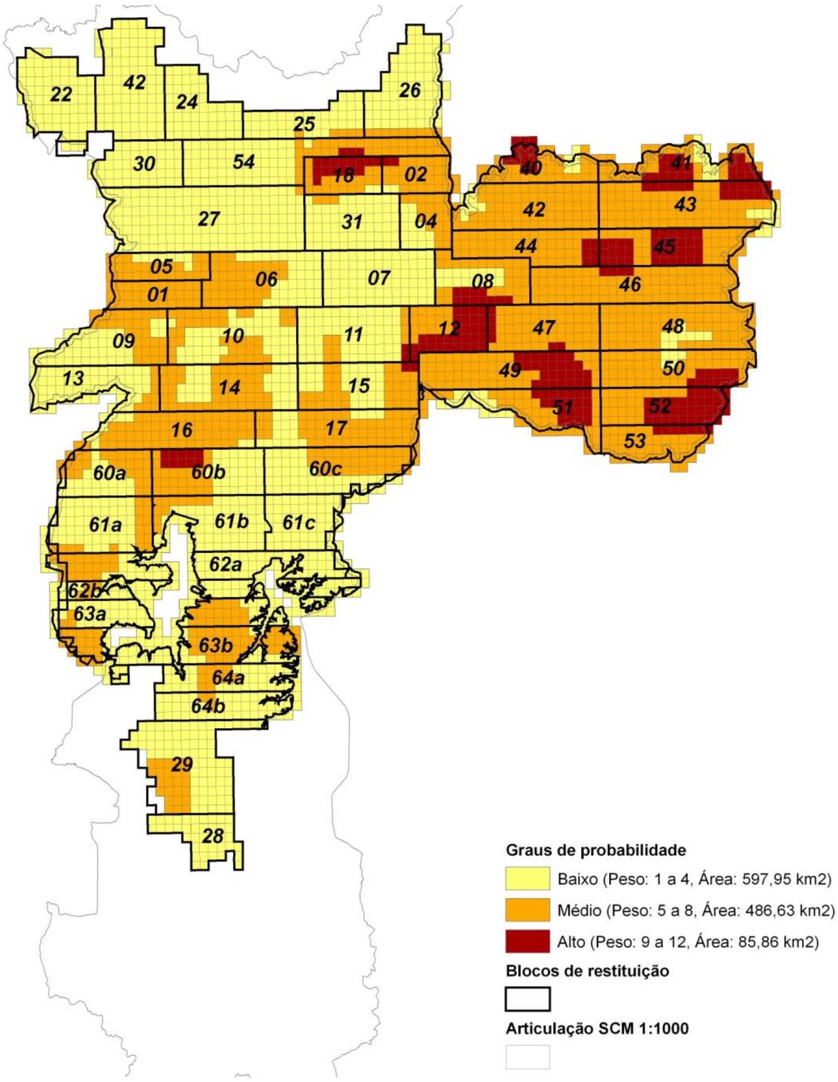

Figure 1: synthesis map of probability of errors to occur

2.3. Definition of Homogeneous Areas

To start work to definition homogeneous areas in relation to the probability that positional errors occur, some parameter values were established which are capable to identifying the errors causing similar difficulties in mapping. These parameters was divided in two parts and created the map to each group: synthesis map of classes originating aerial triangulation errors (Figure 2) and synthesis map of the Probability of positional errors to occur for the feature collection (Figure 3). The values of criteria are include in a database of feature classes of the map and listed below, are not exclusive, and are used in a combined way:

- Quality register of tie points RMS (Root Mean Square) presented in adjustment rapport.
- Quality register of the block which presents critical limits accuracy results considering GCP (ground control points) distribution and residuals;
- Quality register of the AT adjustment block considering check points residuals;
- Data quality register of aerial photo flight, scale and attitude date (reporting non-quantitative quality information and near of tolerance);
- Quality register of strategy, considering quantity and distribution of tie points and the residuals.
- Areas with greater topographic variation, resulting in greater land slope and areas of critical density of occupation or the greater economic importance, as well as areas with complex road;
- Areas of contiguity between blocks of restitution (feature collection);
- Areas division of blocks of restitution between the companies.

All of these criteria were considered to influence the quality of mapping and require redoubled attention in the areas where overlappings appear. In particular, the areas, corresponding to blocks of aerial triangulation (AT) with residual values close to the tolerance limit, should be treated with priority because already being close to the limit, to do this, it was generated the different maps and after that combined both in the end. This type of procedures were done to be submitted to be approve by suppliers and contractors of mapping, consequently it is necessary indeed each step. Those AT areas mapping may easily exceed the tolerance of positional accuracy even without considering the total amount of errors in the stage of restitution. The blocks of aerial triangulation were analyzed according to the parameters of

Table 1 in the map that resulted from the “Aerial triangulation errors originating classes" presented in Figure 2.

Quality Parameters	Analysis of the criteria		Weights
	X, Y (cm)	Z (cm)	
AT quality registers: block points RMS and orientation parameters RMS	0,07<RMSXY≤0,10	0,07<RMSZ≤0,10	0
	0,10<RMSXY≤0,15	0,10<RMSZ≤0,12	1
	0,15<RMSXY	0,12<RMSZ	2
Quality registers of the GCP	0,10<XY≤0,15	0,15<RMSZ≤0,25	0
	0,15<XY≤0,20	0,25<RMSZ≤0,35	1
	0,20<XY	0,35<RMSZ	2
AT block check points residuals	0,10<XY≤0,15	0,15<RMSZ≤0,20	1
	0,15<XY≤0,20	0,20<RMSZ≤0,25	2
	0,20<XY	0,25<RMSZ	3
Quality register of photo flight (line and attitude)	3º< drift ≤4º		1
	4º< drift ≤5º		2
Quality register of strategy of tie point distribution and quantities	Q ≥ 10		0
	10 > Q ≥ 6		1
	Q < 6		2

Table 1. Parameters and variables values determined considering AT errors class detected in the quality control to this process

The geocoding and visualization of the maps of other factors probability has resulted in a map of "Probability of Occurrence of Positional Errors for the Feature Collection" (Figure 3), which includes areas where there is high density of occupation and high land slope, classified according to the intensity of occurrences, as it is can be seen in Table 2.

Quality parameters terrain	criterious	weight
Slope	Low slope	0
	Hig slope	1
Occupation density	Low density	0
	Medium density	1
	high density	2

Table 2. Parameters and factor using to determinate the error critical areas

To produce a synthesis map of homogeneous areas, which represents the sum or combination of factors the relative importance of the parameters presented in Table 1 and 2, was considered. Thus, the combination of these two maps (Figure 2 + Figure 3) with the application of the weights shown in Table 2 and 3, resulted in a synthesis map of homogeneous areas according to the degrees of probability. The map is called "Map Summary of Probability of Positional errors to occur "Figure 1". This map is the fundamental basis in the selection of sample areas, as it indicates the locations where the probability is higher that errors occur and it is pointing to more rigorous analysis. These maps of critical areas can be changed, and the selection of sampling areas can be redirected when new critical areas are targeted, according to random criteria, ensuring, however, the minimum number of 32 samples for Area A and 3 samples for Area B.

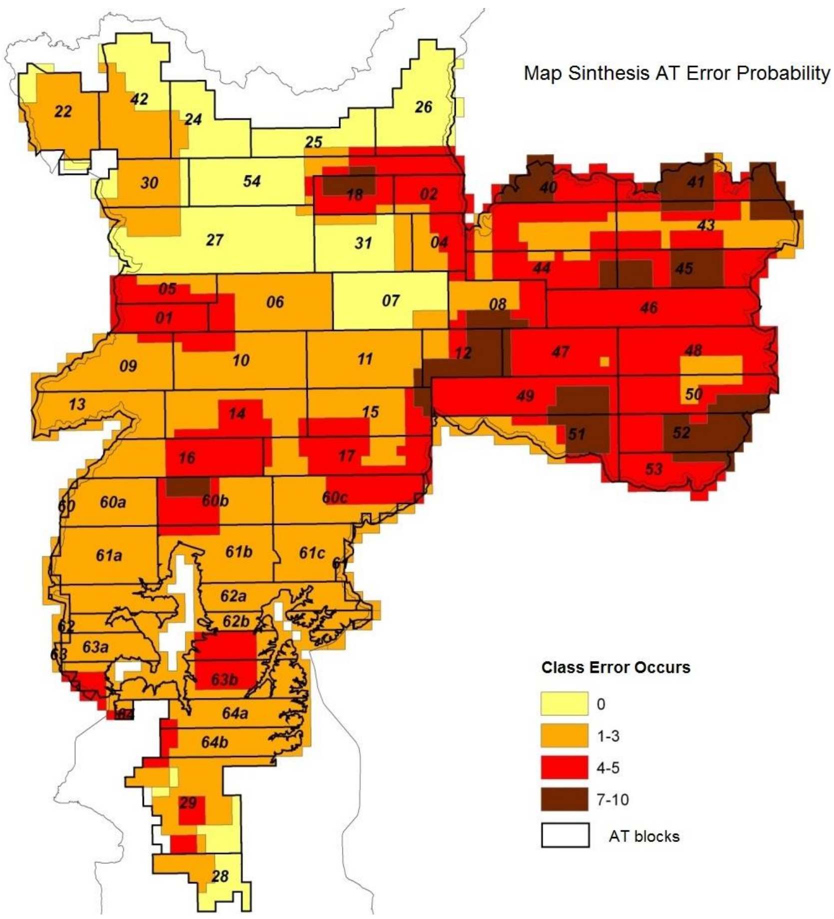

Figure 2: synthesis map of classes originating aerial triangulation errors

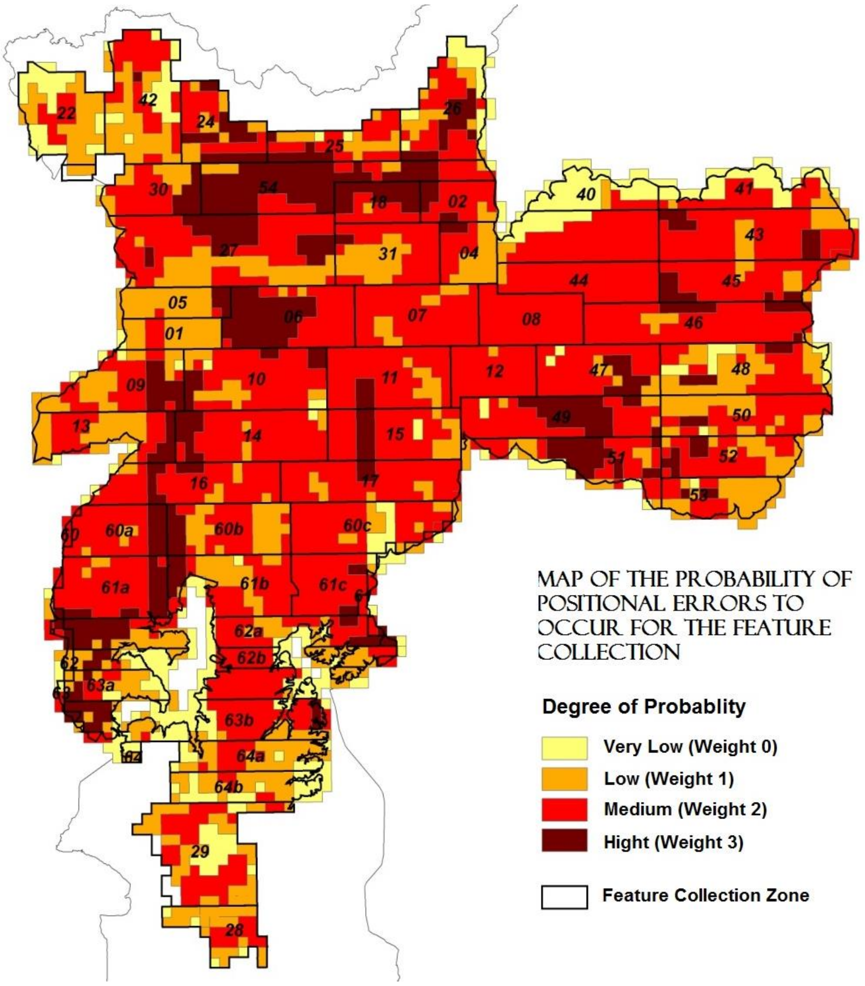

Figure 3: synthesis map of the Probability of positional errors to occur for the Feature Collection.

2.4. Definition of the sample size

The definition on the sample size is also validated, with some adaptations and innovation may be used according to the parameters and criteria defined by the national cartographical standards, like the PEC, which refers to the positional accuracy of the mappings. This subject is a second part of methodology that is not describe in this article.

3. Conclusion

The definition on the sample size is also validated, with some adaptations and innovation may be used according to the parameters and criteria defined by the national cartographical standards, like the PEC, which refers to the positional accuracy of the mappings. This subject is a second part of methodology that is not describe in this article. T

Overall, the results of this study attended the expectations of the mapping contractor and suppliers. The methodology was tested in several areas of the map, showing the criticality factor of the corresponding area, which can be considered valid for this mapping project.

References

Ariza, F. J. (2002). Calidad en la producción cartográfica, Editora Ra-Ma, Madrid.

Cavalcante, F. & ZEPPELINI, P. D. (2001). Amostragem Estatística em Auditoria – Parte I. Up-to-date N. 176.

Cavalcante, F. & ZEPPELINI, P. D. (2002). Amostragem Estatística em Auditoria – Parte II. Up-to-date N. 177.

IGAC (2001). COMITÉ TÉCNICO DE NORMALIZACIÓN DE INFORMACIÓN GEOGRÁFICA - ICONTEC 00034.http://codazzi4.igac.gov.co/. Accessed 9 July 2002.

Diário Oficial da União (1984). DOU 22-06-84. Decreto n. 89.817, de 20 de junho de 1984. Brasília, 1984.

Freund, J.E. & Simon, G.A. (2000). Estatística aplicada, Bookman, Porto Alegre.

Mplmic (Minnesota Planning Land Management Information Center). (1999). Positional Accuracy Handbook.

Merchant, D.C. (1982). Spatial Accuracy Standards for Large Scale Line Maps, in Technical Congress on Surveying and Mapping. Proceedings, v. 1, p. 222-231.

ROCHA, R. S. (2002). Exatidão Cartográfica para as Cartas Digitais Urbanas. Tese de Doutorado. Universidade Federal de Santa Catarina. Florianópolis - SC.

Teste EMAS – Engineering Map Accuracy Standard, que foi desenvolvido conjuntamente por: American Society of Civil Engineers; American Congress on Surveying and Mapping e American Society of Photogrammetry.

Veregin, H. Giordano, A. (1994). Il controllo di qualitá nei sistemi informativi territoriali. Il Cardo, Veneza - Italy.

Tastan, H. (1999). ìData Quality In Geographic Information System, Ph.D. Thesis, (Supervis or: Prof. Dr. M.Orhan ALTAN), Istanbul Technical University, Geodesy and Photogrammetry Engineering Department, Istanbul.