Analysis of User and Producer Risk when Applying the ASPRS Standards for Large Scale Maps

Francisco Javier Ariza López, Alan D.J. Atkinson, José Luis García Balboa, and José Rodríguez Avi

Abstract

Using a statistical simulation process, the behavior of the Accuracy Standards for Large Scale Maps (ASLSM) of the American Society for Photogrammetry and Remote Sensing is analyzed according to the sample size, as well as the relation between the limiting errors (thresholds), stated by the standard as a root mean squared error, and the actual root mean squared error of the product. When the root mean squared error of the product equals the threshold of the standard the simulation results show the ASLSM is very restrictive, classifying 75 percent of products as Class 2 maps instead of Class 1 maps. If the variability of the product is greater or lesser than this threshold, results can be depicted by a family of acceptance curves. These curves can be employed by users to determine the sample size needed to limit their acceptance risk, but also by producers to analyze their rejection risk.

Introduction

The positional accuracy of cartographic products has always been of great importance. It is, together with logical consistency, the quality element of geographic information most extensively used by the National Mapping Agencies (NMAs), being also the more commonly evaluated (Jakobsson and Vauglin, 2002). Positional accuracy is a matter of renewed interest because of the capabilities offered by the Global Navigation Satellite System (GNSS) and the need for a greater spatial interoperability for supporting the spatial data infrastructures. Different positional behaviors of geographic data sets mean the existence of an inter-product positional distortion and a barrier to interoperation (Church *et al.*, 1998). This barrier is not only for the positional and geometric aspects, but also for thematic ones which are greatly affected by position (Carmel *et al.*, 2006). For these reasons many NMAs are currently involved in the development of positional accuracy improvement programs (EuroSDR, 2004).

In a Geographic Data Base (GDB) the position of a real world entity is described with values in an appropriate coordinate system. Positional accuracy represents the nearness of those values to the entity's "true" position in that system. The positional accuracy requirements for a GDB are directly related to its intended use(s). Positional accuracy is determined by means of a statistical evaluation of random and systematic errors (DOD, 1990) and specified by means of the Root Mean Squared Error (RMSE) or by the mean value of errors (μ) and their standard deviation (σ). Comparison with an independent source of greater accuracy is the preferred method for assessing positional accuracy (ANSI, 1998).

Since positional accuracy is essential in cartographic production, mapping agencies have used statistical methods for its control. In the United States of America there are several recognized standard methodologies which can be used for specifying spatial data products, and resultant positional accuracy compliance criteria, or controls, such as: the National Map Accuracy Standard (NMAS) (USBB, 1947), the Accuracy Standards for Large Scale Maps (ASLSM) (ASPRS, 1989), the Engineering Map Accuracy Standard (EMAS) (ASCE, 1983), or the more recent National Standard for Spatial Data Accuracy (NSSDA) (FGDC, 1998). The NSSDA was developed by the FGDC to replace the NMAS, the ASLSM forming the basis for revision of the NMAS (FGDC, 1998). The increasing importance of Digital Terrain Models during recent years, and the new acquisition technologies (e.g., lidar), support demands for the adaptation of previous standards or the development of new ones (FEMA, 2003, NDEP, 2004; ASPRS, 2004; Maune, 2007).

Standards should be taken into account when seeking economic optimization of the quality of geographic information (Krek and Frank, 1999); with a quality standard the producer provides the product according to the known specifications and characteristics, as defined in the standard. This assures a certain level of reliability and certainty, allowing the acquirer, or user, to avoid excessive measuring of the quality and thus reducing the measuring cost and shortening the decision-making process. But standards are not problem free. As demonstrated recently by Congalton and Green (2009), they commonly use statistical terms in a confusing way, implement statistical bases incorrectly, or are prone to mis-estimation of accuracy parameters. Also, in general, standards contain problems mainly arising from the lack of (Ariza and Atkinson, 2008a): definition (formalism, previous hypothesis testing, bias, and outliers treatment), and the explanation of their own behavior. For this reason is very important for both acquirers and suppliers be well-informed about the main features of the standard they are working with.

Grupo de Investigación en Ingeniería Cartográfica, Universidad de Jaén, Campus "Las Lagunillas" s/n 23071 Jaén (Spain) (fjariza@ujaen.es).

Photogrammetric Engineering & Remote Sensing
Vol. 76, No. 5, May 2010, pp. 625–632.

0099-1112/10/7605–0625/$3.00/0

The objective of this paper is to analyze the ASLSM from the perspective of the user's and producer's risk in order to clarify its acceptance/rejection behavior. For this purpose we will use an analysis methodology based on a simulation strategy. Using the same methodology of simulation some recent studies have analyzed the EMAS (Ariza *et al.*, 2008) and the NSSDA (Ariza and Atkinson, 2008b) offering information about these characteristics of the aforementioned standard. These are new aspects that have been never addressed before for this kind of control methodologies. For this reason it is interesting to develop the same analysis for the ASLSM in order to obtain a similar degree of knowledge about its specific acceptance/rejection behavior.

The ASLSM establishes a set of positional accuracy requirements in the form of limiting root mean squared errors (*LE*) for planimetry and altimetry. These *LE* depend on the product scale. It is a very simple and functional standard which offers a clear classification system based on assessed product accuracy. The ASLSM was developed in a period of transition from analogous to digital cartographic systems, and the proposed computational method and system of tolerances and classes can be applied to digital as well as analogous data. This standard is still in use in the US and has also been applied in other countries. So, despite its age (20 years) and that the fact that it has been surpassed by the NSSDA standard, we think it is interesting to analyze its acceptance/rejection behavior. Furthermore, the NSSDA allows data producers to use other standards such as the ASLSM, if they are considered truly applicable to digital geospatial data. Our study focuses on the planimetric case and uses a unitary threshold ($LE = 1$ m), so that results can be easily expanded to other values.

This document is organized into five main sections. The first explains the methodology of the ASLSM in order to highlight its major features. The next section presents the simulation methodology. Once the simulation process is known, it is applied in order to analyze the ASLSM acceptance behavior when the positional uncertainty of the product and the limiting error threshold of the standard are the same (third section) or different (fourth section). Finally the main conclusions are given.

The Accuracy Standards for Large Scale Maps Methodology

The ASLSM was developed during the 1980s by the Specifications and Standard Committee of the American Society for Photogrammetry and Remote Sensing. It was intended as an alternative to the NMAS for large scale topographic maps (Merchant, 1987). As stated in ASP (1985), the interest in spatial accuracy standards for large-scale maps was prompted by a series of court decisions, in the sense that the lack of a generally understood and accepted standard of map accuracy, in terms of quantifiable and verifiable error definitions, was a primary factor in these litigations. During the 1980s several drafts of the standard were published in order to facilitate and encourage comments. The first (ASP, 1985) showed a statistical methodology, based on compliance testing, similar to that of the EMAS, which was developed at the same time by the ASCE Surveying and Mapping Committee on Cartographic Surveying. The second draft (Merchant, 1987), which was finally approved in 1990 (ASPRS, 1990), showed a different methodology, in which computations were simplified because the ASLSM assessment was based on an RMSE calculated for each coordinate (*X* and *Y*). This index is easy to calculate and takes account of bias and dispersion in a unique value. In fact, it is a double estimation process, for the *X* and *Y*, and a double pass/fail test, but in this new version, no statistical compliance tests were included. If the absence of bias is considered, $RMSE = \sigma$, the estimation of an RMSE is similar to the estimation of a population standard deviation. This estimator is a random variable with a specific density function which depends on the sample size (*n*) and the true value of σ (Johnson *et al*, 2005).

Using the structure proposed by Giordano and Veregin (1994), the ASLSM can be summarized as follows (the vertical component is omitted in the explanation):

- Common applications: The ASLSM specifies the planimetric accuracy of large scale maps as adopted and recommended by the American Society for Photogrammetry and Remote Sensing.
- Comparison method: The accuracy compliance is tested by comparing coordinates of well-defined points on the product to the coordinates of the same points as determined by a check survey of higher accuracy (at least one-third of the limiting error).
- Positional component - Planimetric: The *X* and *Y* coordinates are evaluated separately.
- Class of control elements - Points: Well-defined points are used for the control. This term refers to features that can be accurately identified as discrete points. Points that are not well-defined are excluded from the map accuracy test.
- Correspondence to accuracy levels: The standard establishes planimetric coordinate accuracy requirements (see Table 1) for the *X* and *Y*, in ground units, for well-defined points. These threshold values are mean limiting errors stated as root mean squared errors. These requirements are for the so called "Class 1 maps." Products compiled within limiting RMSE of twice or three times those allowed for a Class 1 map shall be designated Class 2 or Class 3 maps, respectively.
- Overview: The ASLSM establishes a methodology for the assessment of positional error and for the classification of the product. In relation to assessment, the quantification of error is based on the RMSE derived from a sample of control points. Some guidelines are given for the control elements (well-defined points), the sample size, the sample spatial distribution, and blunder management. In relation to the classification, it is a multiple pass/fail classification system defined upon a set of threshold values related to a nominal map scale.
- Procedure: The main steps of the procedure are:

1. Select a sample size (*n*) of, at least, $n = 20$ well-defined and well-distributed points both in the product and on the ground.
2. For each point *i* ($i \in \{l, \ldots n\}$) calculate the discrepancies between the coordinates of each point as determined from the product and by the check survey:

$$dx_i = X_{product;\,i} - X_{check;\,i},$$

$$dy_i = Y_{product;\,i} - Y_{check;\,i}.$$

TABLE 1. PLANIMETRIC COORDINATE ACCURACY REQUIREMENT (GROUND *X* OR Y) FOR WELL-DEFINED POINTS. CLASS 1 MAPS LIMITING ROOT MEAN SQUARED ERRORS (*LE*)

Typical map Scale	*LE* [m]
1:50	0.0125
1:100	0.025
1:200	0.050
1:500	0.125
1:1000	0.25
1:2000	0.50
1:4000	1.00
1:5000	1.25
1:10 000	2.50
1:20 000	5.00

3. Calculate the RMSE for each coordinate:

$$RMSE_X = \sqrt{\frac{\sum dx_i^2}{n}},$$

$$RMSE_Y = \sqrt{\frac{\sum dy_i^2}{n}}.$$

4. For each coordinate, compare the assessed *RMSE* with the *LE* or acceptance thresholds established by the standard (Table 1) for the nominal scale of the product. The rules are:
   - R1: If $RMSE \leq LE$ the product can be classified, or accepted, as Class 1 map,
   - R2: If $LE < RMSE \leq 2 \times LE$ the product can be classified, or accepted, as Class 2 map,
   - R3: If $2 \times LE < RMSE \leq 3 \times LE$ the product can be classified, or accepted, as Class 3 map.
5. If the product was checked and found to conform to this standard the following statement can be used: *"This map was checked and found to conform to the ASPRS Standard for Class __ Map Accuracy."*

Simulation Methodology

Simulation for this project was used as the base tool for analyzing the behavior of the ASLSM methodology in the planimetric case. For our simulation we applied the Monte Carlo Method. Simulation is an easy method to understand because it is based on the reproduction of a known process or system, so that theoretical aspects are obviated. Simulation can be defined as the construction of a mathematical model to reproduce the characteristics of a phenomenon, system or process, using a computer in order to infer information or solve problems (Ríos *et al.*, 1997). There is no single prescribed methodology under which simulation studies are conducted. Most simulation studies follow four major steps (Ríos *et al.*, 1997): formulating the problem, developing the model, running the model, and analyzing the output.

Our research focuses on the study of the acceptance or classification of the estimated positional accuracy, but only horizontal, of a GDB when applying the ASLSM to different sample sizes. So we are going to replicate the ASLSM application process under controlled circumstances by means of synthetic populations of errors. Here some basic assumptions are made:

- In order to enable the generalization of results, synthetic Normal ($\mu_p = 0$, $\sigma_P^2 = 1$) distributed populations of data are used as if they were positional errors of *X* and *Y* coordinates derived from a control survey of higher accuracy. Notice that the sub-index *p* in parameters means population. Note that (a) it is considered that ASLSM controls are applied to GDBs with no bias ($\mu_P = 0$), which means that $RMSE_p = \sigma_p$, and (b) μ and σ units of measure are meters.
- The simulation does not work with simulated control points distributed in a geographic space, but rather with their simulated positional errors. This means that positional errors and the spatial distribution of control points are uncoupled. By means of this artifact, it is not necessary to observe the spatial distribution of control points. Because of this simulation, it is possible to imagine that the spatial distribution of control points is always adequate, not affecting the validity of results.

The model, or the simulation process, basically consists of (Figure 1):

- Simulation of populations (A): Populations of well known parameters ($\mu_P = 0$, $\sigma_P^2 = 1$) are derived from a controlled random values generation process. Single population values are considered positional error values. The total of synthetic populations that will be realized is NR_{SP}.
- Simulation of samples (B): Samples of different sizes ($n = 10$, 20, 30, and so on) are randomly extracted from each population. The ASLSM is applied to each sample as if it were a single positional control test or assessment. For each sample size NR_{SAM} samples will be extracted for each population.

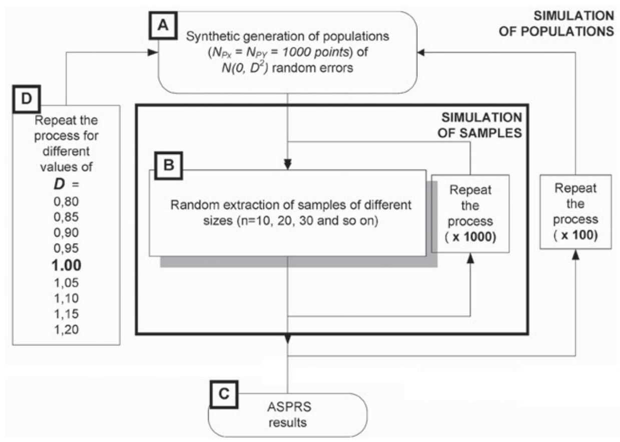

Figure 1. Simulation process flowchart.

The Monte Carlo method is conceptually related to sampling, so the number of times or runs (*NR*) a process must be iterated can be derived from sampling recommendations. Heuvelink (1998) presents a discussion on *NR*. Here $NR_{SP} = 100$ and $NR_{SAM} = 1{,}000$; so the total number of runs considered for each sample size is $NR = NR_{SP} \times NR_{SAM} = 100{,}000$; this in order to obtain statistically sounder results with the lesser variations of estimated deviations (stability), and because the numerical load is not a problem here.

The output of the process is analyzed from the statistical computations (C in Figure 1). The process is essentially variance estimation, but it is applied two times, one for the *X* coordinate and another for the *Y* coordinate. After this step, a multiple pass/fail test is applied (remember Point·No. 4 of the ASLSM summary). Because the simulation is performed using a Normal $N(\mu_P = 0, \sigma_P^2 = 1)$ distributed population, the theoretical value to be estimated by the ASLSM is $RMSE_{X\ (theoretical)} \approx RMSE_{Y\ (theoretical)} \approx 1.0$ m. For each sample size ($n = 10$, 20, 30, and so on) results values of ASLSM are aggregated by deriving the mean value of $RMSE_X$ and $RMSE_Y$, and the mean value of the deviation (variability) of them, taking into account the total number of runs (*NR*). Also the stability of the simulation process (quality of the process) is assessed by deriving the variability between populations.

For the analysis performed in the section dealing with the case where uncertainty of the product and *LE* are different, we also considered the possibility of different population deviations, meaning different variational behaviors. In this case the simulation of populations (synthetic generation of populations) is repeated for different values of variance (labeled D in Figure 1).

ASLSM Acceptance Behavior when the Positional Uncertainty of the Product and Limiting Error Threshold of the Standard are the Same

As stated in Point·No. 4 of the ASLSM summary, acceptance comes from the comparison of the assessed $RMSE_X$ and $RMSE_Y$ of the sample with the *LE* indicated by Table 1 for the nominal scale of the product. Here we are going to use an $LE = 1$ m as a general case. If we consider that the sampling is representative, the assessed *RMSE* equals the actual *RMSE* of the product. But the last is the result of a production process, that is, a unique combination of tools, materials, methods and people engaged in producing a measurable output. We can thus speak about a process capability index (C_p) which can be defined as the ability of a process to meet its specifications, and it can be expressed here as the ratio $C_p = LE/RMSE$. If $RMSE < LE$; it means that $Cp > 1$, and the process increases its probability in order to meet its specifications. If $RMSE > LE$ means that $Cp < 1$ and the process decreases its probability to meet its specifications. If $Cp = 1$ we are producing with the same quality as expected, and here the random variations coming from sampling will affect our final pass/fail decision.

The result of an estimation process can be expressed as a mean value for the statistical estimator and its deviation or variability from the mean value. Both mean and deviation values are affected by sample size, but especially deviation. High variability of an estimator means that the estimation is not fine, and for a given case it can take values which could create problems through the loss of accuracy incurred. Results for our simulation process are shown in Figures 2 and 3:

- Figure 2a shows the estimated mean values for the $RMSE_x$ (continuous light gray line) and $RMSE_y$ (dashed dark gray line) as a function of the sample size (n), taking into account the total number of runs *NR*. Both estimations are very similar in value and in behavior in relation to the sample size.
- The stability (variability between populations) in both mean *RMSE* values is depicted in Figure 2b. Both curves (for *X* and *Y*) show a decreasing variability and greatly reduced values, which means a high stability of the simulation process.
- Figure 3a shows the estimated mean values of the deviation (variability) of $RMSE_x$ and $RMSE_y$ as a function of the sample size (n), taking into account the total number of runs *NR*. This figure is very important because it shows the variability of the estimation process. For instance, it shows that we need 200 control points if we wish to limit the variability of the estimation to ±5 percent.
- The stability (variability between populations) in both mean values of deviation is depicted in Figure 3b. With values always less than 1 percent, Figure 3b also confirms a very high stability for the simulation process

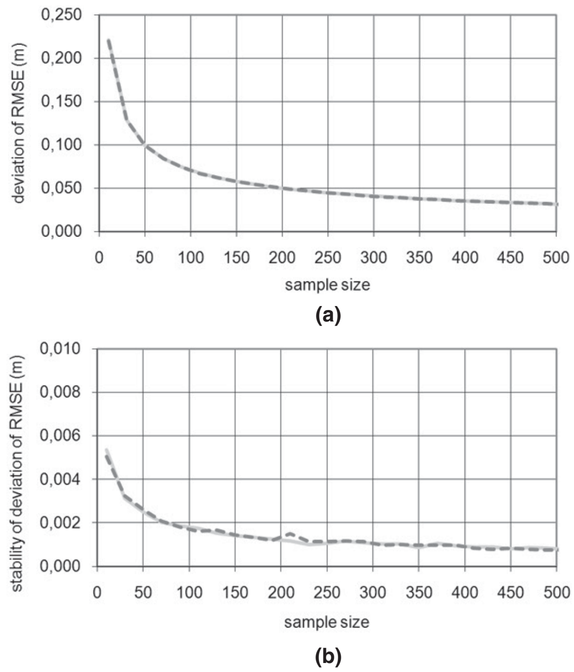

Figure 3. (a) Estimated mean values of the deviation of $RMSE_x$ and $RMSE_y$ (meters, vertical axis) for each sample size (n, horizontal axis, and (b) Stability (variability between populations) of the mean value of the deviation of $RMSE_x$ and $RMSE_y$ of Figure 3a.

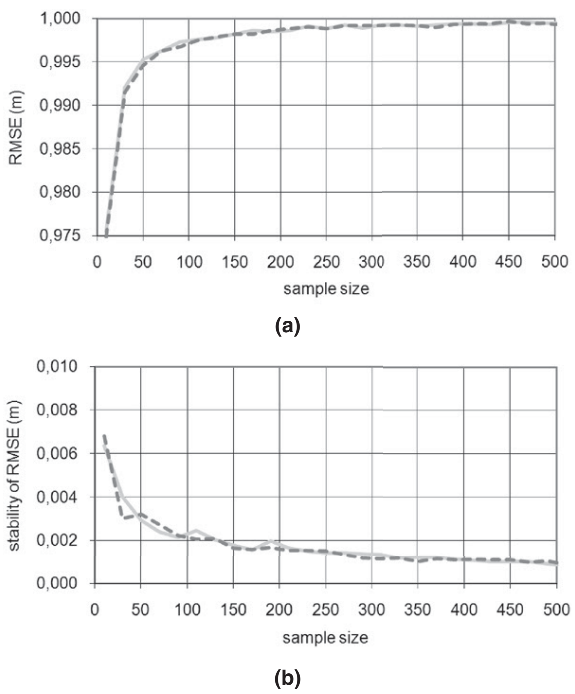

Figure 2. (a) Estimated mean values of $RMSE_x$ (continuous light gray line) and $RMSE_y$ (dashed dark gray line) (meters, vertical axis) for each sample size (n, horizontal axis), and (b) Stability (variability between populations) of the mean value of $RMSE_x$ and $RMSE_y$ of Figure 2a.

Because we are working with a simulation process based on synthetic and controlled data, we know in advance the actual *RMSE* and the class to which a GDB should be assigned (in our case a Class 1 map). So the acceptance behavior means the actual classification given by the standard in each iteration of the simulation process. This classification is not always the same as expected, and this is the important behavioral characteristic we are interested in for this paper. Figures 4 and 5 show the simulation results for the assignation to a Class 1 map.

- Figure 4.a shows the estimated mean probability of classification into Class 1 as a function of the sample size (n), taking into account the total number of runs *NR*. As can be observed, the probability for each coordinate (*X*, continuous light gray line, and *Y*, dashed dark gray line) approaches 50 percent when n increases sufficiently. The value 50 percent comes from the ASLSM definition itself. In our simulation and acceptance processes we are producing populations with a $\sigma_P = 1$, which implies an $RMSE \approx 1$ estimation and accepting or classifying results as Class 1 map if actually $RMSE \leq 1$.

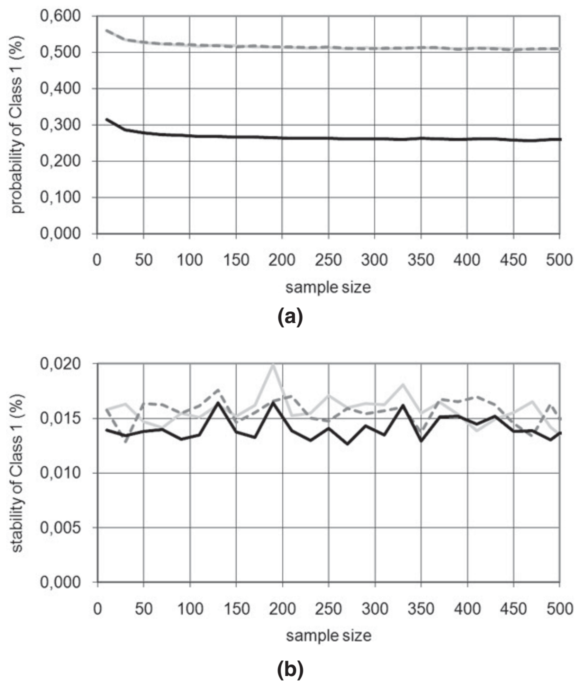

Figure 4. (a) Estimated mean values of probability of classification into Class 1 (%, vertical axis) for each sample size (*n*, horizontal axis) and for *X*, *Y* coordinates separately (continuous light gray line and dashed dark gray line) and together (continuous black line), and (b) Stability (variability between populations) of the mean values of probability of Figure 4a.

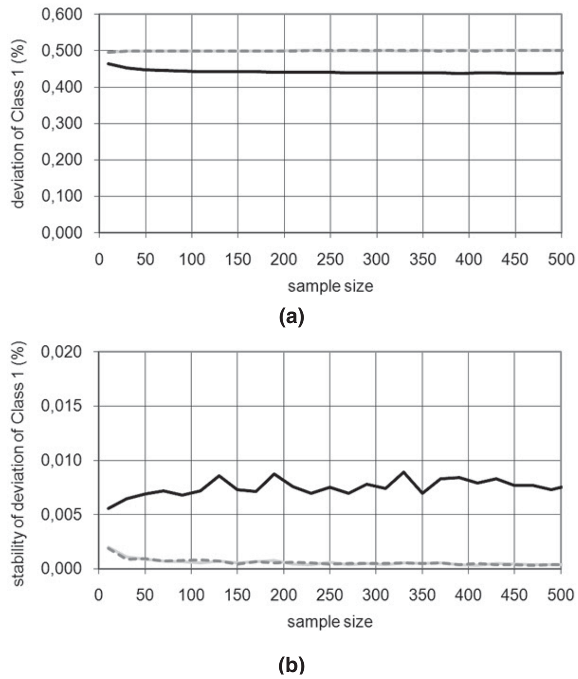

Figure 5. (a) Estimated mean values of the deviation of the probability of classification into Class 1 (%, vertical axis) for each sample size (*n*, horizontal axis) and for *X*, *Y* coordinates separately and together, and (b) Stability (variability between populations) of the mean values of deviation of the probability of Figure 5a.

Here the estimation of the *RMSE* is similar to the estimation of a mean value of a population. The standard deviation follows a *N* (0, 1) and the probability $P(\sigma_P <= 1) = 50\%$. When considering both coordinate pairs together, the global value approaches 25 percent because of a double standard deviation estimation. This value means independence between both coordinates ($P(\sigma_{xP} \leq 1) \times P(\sigma_{yP} \leq 1) = 0,5 \times 0,5 = 0,25$). This situation is depicted in Figure 4a by the continuous black line which is the final probability for classifying into Class 1 map for populations distributed as *N* (0, 1).

- Figure 4b shows the stability (variability between populations) of the mean probability value of classification into Class 1 which is always below 2 percent.
- Figure 5a depicts the estimated mean value of the deviation (variability) of the probability as a function of the sample size (*n*), taking into account the total number of runs *NR*. It is very close to 50 percent for each coordinate separately. This value is obvious for the same reasons given above.
- Figure 5b shows the stability (variability between populations) of the mean values of deviation. As can be observed these values are always less than 1 percent, which confirms a great stability in the simulation process.

It is very important to notice here that the rest of the cases are classified as belonging to the following class, which means Class 2 maps in our case. These results evidence that the ASLSM is very restrictive: we are working with synthetically created error populations belonging to Class 1, but the standard assigns them to a lower class (Class 2). This means that there is no user's risk but a high producer's risk. There is no user's risk, because we know that the errors of the product are always distributed following a Normal distribution with $\mu_P = 0$ and $\sigma_P^2 = 1$, to which corresponds a Class 1 map classification. When we accept a product as a Class 1 map it always belongs to this class. There is a high producer's risk because an average of 75 percent of good products, products whose error populations belong to Class 1 map, are rejected as Class 1 maps and assigned to Class 2. This problem arises because the process capability index equals one ($C_p = 1$).

At this point it is interesting to analyze the situation for the minimum sample size ($n = 20$) recommended by the standard. Figure 2a shows a sub-estimation of the *RMSE* but only in the order of 2 percent, but Figure 3a gives us more problematic information; the variability of estimations is in the order of 20 percent, which is a significant value. Regarding the classification process, here only 30 percent of cases are classified as Class 1 maps and the complementary 70 percent as Class 2 maps when they actually are Class 1 maps. We obtain 30 percent instead of 25 percent because the estimator of the *RMSE* is not an unbiased estimator, and the bias correction factor is not included (Weisstein, 2008). Regarding the variability of the classification process, it is estimated at 45 percent, which is also very high. This is a bad position for the producer because good products are underclassified, and for the user because producers always translate costs to buyers.

ASLSM Acceptance Behavior when the Positional Uncertainty of the Product and Limiting Error Threshold of the Standard are Different

Until now we have worked under the assumption of having a $N(\mu_P = 0, \sigma_P^2 = 1)$ distributed population of errors and analyzed what can occur when estimating from a sample of a given size *n*. Now we are going to analyze the behavior of

TABLE 2. VARIABILITY RATIOS AND CORRESPONDING POPULATION DEVIATIONS AND CAPABILITY INDEXES USED FOR TESTING THE ASLSM WHEN THE POSITIONAL UNCERTAINTY OF THE PRODUCT AND THE *LE* THRESHOLD OF THE STANDARD ARE DIFFERENT

LE [m]	Population Deviation (*D*) [m]	Variability ratios [%]	*Capability index Cp*
1	0.80	−20%	1.25
1	0.85	−15%	1.18
1	0.90	−10%	1.11
1	0.95	−5%	1.05
1	0.98	−2%	1.02
1	0.99	−1%	1.01
1	**1.00**	**0%**	**1.00**
1	1.01	+1%	0.99
1	1.02	+2%	0.98
1	1.05	+5%	0.95
1	1.10	+10%	0.91
1	1.15	+15%	0,87
1	1.20	+20%	0.83

the ASLSM when expecting $N(\mu_P = 0, \sigma^2_P = 1)$ distributed population but actually working with other normal $N(\mu_P = 0, \sigma^2_P = D^2)$ distributed population where $D \neq 1$. In this way we are going to analyze the influence of the C_p in the classification assigned by the ASLSM estimations, and this analysis will allow us to derive the acceptance curves for the same.

For this analysis we have used a simulation process similar to that mentioned above, but changing the variation behavior when creating random populations. So a set of normally distributed populations has been synthetically created following a $N(\mu_P = 0, \sigma^2_P = D^2)$, where $D = 0.8$; 0.85; 0.9; 0.95; 1.00; 1.05; 1.10; 1.15, and 1.20. For each synthetic population a thousand samples of different sizes ($n = 10$, 20, 30, and so on) were extracted. The ASLSM was applied to each sample as if it were a single positional control test.

The different values of *D* can be considered, in relation to $D = 1$ m, as variability ratios implying a detected nominal accuracy value when applying the ASLSM, and vice versa. This idea is presented in Table 2.

In order to understand the acceptance index in the form of probability (percentage) we consider the following rules:

- If $D < 1$, the accuracy of the population is better than expected ($C_p > 1$), and this means that it would be considered as satisfactory or accepted. So when performing the simulation, we will take into account the number of cases where $\{RMSE_X < 1$ m *and* $RMSE_Y < 1$ m$\}$ and so results in a Class 1 map. The number of such cases will be expressed as an acceptance percentage of total cases.
- If $D > 1$, the accuracy of the population is worse than expected ($C_p < 1$), and this means that it would be considered as not satisfactory or rejected. So when performing the simulation, we will take into account the number of cases where $\{RMSE < 1$ m *and* $RMSE_Y < 1$ m$\}$ and so results in a Class 1 map classification. The number of such cases will be expressed as an acceptance percentage of total cases.

The results of this process are presented in Figure 6, which shows a family of curves we call acceptance curves. The horizontal axis corresponds to sample size (*n*), the vertical to percentage of acceptance (%) and the labels to C_p. The wider black curve in the middle (labeled 1.00), corresponds to the case where $D = 1$, a previously studied situation where population follows a $N(\mu_P = 0, \sigma^2_P = 1)$. Here it can be observed that this curve is somewhere above 25 percent. This has been commented on previously. Curves above the wider line correspond to those cases were $D < 1$, labeled 1.01 up to 1.20, and curves below to those where $D > 1$, labeled 0.99 down to 0.80.

Figure 6 shows the general acceptance behavior of the ASLSM standard. When $D < 1$ ($C_p > 1$) quality is better than expected (lesser variability), and user acceptance increases when sample size *n* also increases. Here there is a risk for the producer because the product is good enough in variability but is not accepted in a percentage that equals 100 percent minus acceptance (%) (a good product can be rejected in this percentage). On the other hand when $D > 1$ ($C_p < 1$) quality is worse than expected (greater variability), user acceptance is at risk (a bad product can be accepted in this percentage), but the risk decreases when sample size *n* increases. Figure 6 also shows that however lesser or greater the deviations are in

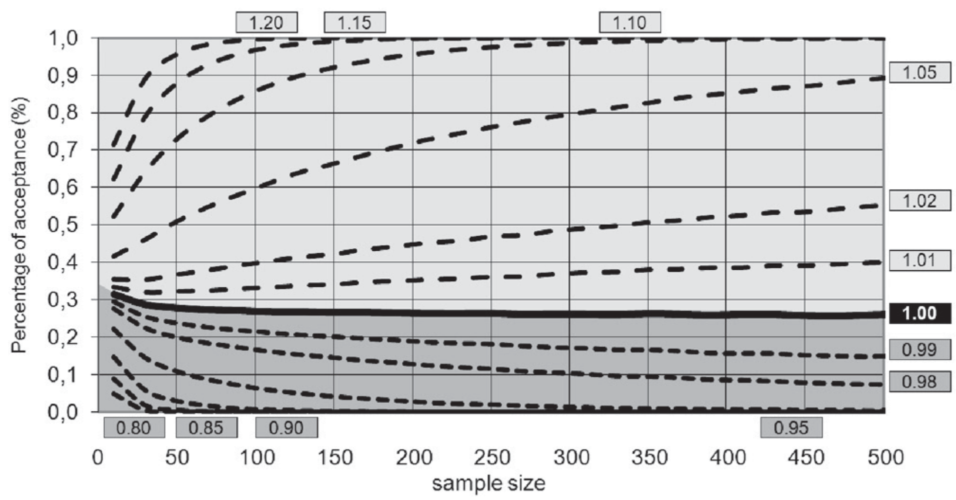

Figure 6. Acceptance curves for the ASLSM (%, vertical axis) for each sample size (*n*, horizontal axis). Each curve is labeled with its *Cp* value. Dark gray color of the background represents the user's risk (below the curve with *Cp* = 1.0). Light color of the background represents the producer's risk (above the curve with *Cp* = 1.0).

relation to the expected, the smaller sample size is needed for a given acceptance or rejection risk.

One example can help to understand the above. Let us consider a batch of products to which the ASLSM is applied. If the positional error of the product follows a $N(\mu = 0, \sigma^2 = 0.95^2)$ and we expect a behavior of a $N(\mu = 0, \sigma^2 = 1^2)$, the variability of this data is 5 percent less than expected (it means it is 5 percent better) and the acceptance follows the curve of Figure 6 labeled 1.05. For this situation, if the sample size of control points is $n = 100$, the percentage of acceptations in the batch is on average in the order of 60 percent, which also means a risk of 40 percent for the producer, because the product is actually better than expected. Consider now a product whose variability is about 10 percent less than expected (it means it is about 10 percent better) and the acceptance follows the curve of Figure 6 labeled 1.10. For this situation, if the sample size of control points is $n = 200$, the percentage of acceptations in the batch is on average in the order of 95 percent, which also means a risk of 5 percent for the producer.

In general, all the curves of Figure 6 below the wider black curve in the middle (labeled 1.00) represent user's risk (background in dark gray color). For example, consider a product whose positional errors follows a $N(\mu = 0, \sigma^2 = 1.05^2)$, and where we expect a behavior of a $N(\mu = 0, \sigma^2 = 1^2)$, the variability of this data is 5 percent more than expected (it means it is 5 percent worse), the capability index C_p is 0.95, and the acceptance follows the curve labeled 0.95. For this situation, if the sample size of control points is $n = 50$, the percentage of acceptations in the batch is on average in the order of 11 percent, which is the risk of the user.

In industrial processes the producer's and user's risk are commonly limited to 5 percent or 10 percent, respectively. So bearing our curves in mind, unless the difference between the actual and expected variability is greater than 15 percent, the sample size provides the best protection for the interests of both producer and user. This can be seen from a different point of view: the interest for the producer to create GDBs with positional accuracies at least in the order of 15 percent better than the stated threshold for the assessment, and thus to reduce his rejection risk, which is the same as increasing the capacity of his production process. This also allows a reduction of the sample size of the positional accuracy assessment. So there is a need for an agreement between the user (acquirer) and the producer in order to decide where to direct the resources: to control or production. For particular cases the answer comes from the cost functions of both processes. In general, the solution is quality assurance in production and reduction of controls.

As stated in the previous subheading, it is very important to notice that the rest of the cases are designated as belonging to other classes. Table 3 shows the assignation given by the ASLSM when applied to the same cases used for compiling Figure 6. It is a complementary view of this figure. For each sample size, the other columns show, for values of *LE* from 0.5 to 1.3, the acceptance or classification percentage for the indicated class (row below C_p) under the supposition of populations distributed as a $N(\mu = 0, \sigma^2 = 1^2)$. The last row of Table 3 indicates the alternative class assigned. Classification occurs always between two adjacent classes, for instance, between Class 1 and Class 2 or between Class 2 and Class 3. Classification between Class 1 and Class 3 is very improbable. From the column with $C_p = 0.8$ to the right of the table we can see how the classification into Class 1 maps increases as C_p increases. It is interesting to notice that the column corresponding to $C_p = 0.5$ shows the same behavior as the column with $C_p = 1.0$, but assignations occur between Class 2 and Class 3, instead of between Class 1 and Class 2.

TABLE 3. NUMBER OF MAPS (%) BELONGING TO A CLASS OF THE ASPRS STANDARD IN RELATION TO THE SAMPLE SIZE AND THE CAPABILITY INDEX C_p

	C_p								
	0.5	0.6	0.7	0.8	0.9	**1.0**	1.1	1.2	1.3
Sample size (*n*)	% ∊ Class 2	% ∊ Class 2	% ∊ Class 2	% ∊ Class2	% ∊ Class 2	**% ∊ Class 1**	% ∊ Class 1	% ∊ Class 1	% ∊ Class 1
10	30.8	72.6	93.1	94.5	86.5	**31.7**	52.5	71.7	85.3
20	28.8	83.4	98.7	98.6	91.5	**29.3**	59.5	82.7	95.0
30	27.9	90.3	99.8	99.6	94.4	**29.3**	64.7	89.3	98.3
40	27.9	94.4	99.9	99.8	96.6	**28.5**	69.8	93.9	99.4
50	27.5	96.8	99.9	100	97.4	**27.7**	73.7	96.1	99.8
60	27.1	98.1	100	100	98.5	**27.4**	77.2	97.5	99.9
70	27.1	98.9	100	100	98.8	**27.4**	80.7	98.5	99.9
80	27.0	99.2	100	100	99.1	**27.1**	83.7	99.1	100
90	27.0	99.6	100	100	99.4	**27.2**	85.8	99.5	100
100	27.5	99.7	100	100	99.6	**27.6**	87.2	99.7	100
150	26.7	100	100	100	100	**27.0**	94.5	100	100
200	26.6					**26.6**	98,1		
250	26.8					**26.7**	99,2		
300	26.7					**26.7**	99,9		
350	25.9	100	100	100	100	**26.7**	100	100	100
400	26.3					**26.7**	100		
450	26.6					**26.6**	100		
500	27.0					**27.0**	100		
The remaining samples ∊ to the Class:	***3***	***3***	***3***	***1***	***1***	**2**	**2**	**2**	**2**

Conclusions

The behavior of a methodology should be known for its appropriate application. As with other positional accuracy assessment methodologies, the ASLSM does not give any information about that. By using a simulation-based methodology the ASLSM classification estimation behavior has been analyzed. The statistical analysis is based on the use of normal distributed synthetic populations, which ensures the control of the process, the generality of results and their easy applicability to real cases. The main conclusions derived from our research are:

- The ASLSM is based on a double estimation process of standard deviations (*X* and *Y* coordinates). These estimations are similar to the general estimation process applied to mean values.
- Classes of map accuracy are established using a mean limiting error calculated as a *RMSE*, but no advice is given about the relation between a product *RMSE* and the limiting error to consider for the ASLSM application. This relation (*LE/RMSE*), named here as the process capability index is the key factor for understanding the ASLSM behavior.
- If LE equals the *RMSE* of the population, the ASLSM has a very restrictive behavior, classifying 75 percent of products as Class 2 maps instead of Class 1 maps.
- For the minimum proposed sample size ($n = 20$ points), and under the hypothesis $LE = RMSE$, only 30 percent of cases are classified as Class 1 maps and the complementary 70 percent as Class 2. This is a bad position for the producer because good products are labeled as having lower accuracy.
- In order to obtain at least a 95 percent of acceptance we need a product which satisfies (a) the ratio $LE/RMSE \approx 1.20$ and a sample size of 50 points, (b) the ratio $LE/RMSE \approx 1.15$ and a sample size of 100 points, and (c) the ratio $LE/RMSE \approx 1.10$ and a sample size of 200 points.
- If the variability of the population is greater or lesser than expected the behavior of the ASLSM results is depicted by a family of acceptance curves. These curves can be entered using the relation *LE/RMSE*, so that they can be considered of general application.
- Acceptance curves can be employed by users to determine the sample size for limiting their acceptance risk, but also by producers to analyze the tradeoffs between their product's quality and acceptance ratios in order to decide and establish the capacity of the production process.

Acknowledgments

This work has been partially funded by the Ministry of Science and Technology of Spain under Grant No. BIA2003-02234 and by the Regional Ministry of Innovation, Science and Enterprise of Andalusia (Spain) under Grant No. P08-TIC-4199.

The authors also wish to acknowledge the Regional Government of Andalusia (Spain) for their constant financial support since 1997 to their research group (Ingeniería Cartográfica, Code TEP-164).

References

American National Standards Institute (ANSI), 1998. *Spatial Data Transfer Standard (SDTS) – Part 1, Logical Specifications, ANSI NCITS 320-1998*, Washington, D.C.

American Society of Civil Engineers (ASCE), 1983. *Map Uses, Scales and Accuracies for Engineering and Associated Purposes*, ASCE Committee on Cartographic Surveying, Surveying and Mapping Division, New York.

American Society of Photogrammetry (ASP), 1985. Accuracy specification for large-scale line maps, *Photogrammetric Engineering & Remote Sensing*, 51(2):195–199.

American Society for Photogrammetry and Remote Sensing (ASPRS), 1990. Accuracy standards for large scale maps, *Photogrammetric Engineering & Remote Sensing*, 56(7):1068–1070.

American Society for Photogrammetry and Remote Sensing (ASPRS), 2004. *ASPRS Guidelines, Vertical Accuracy Reporting for Lidar Data*, American Society of Photogrammetry and Remote Sensing.

Ariza López, F.J., and A.D.J. Atkinson, 2008a. Analysis of some positional accuracy assessment methodologies, *Journal of Surveying Engineering*, 134(2):45–54.

Ariza López, F.J., and A.D.J. Atkinson, 2008b. Variability of NSSDA estimations, *Journal of Surveying Engineering,* 134(2):39–44.

Ariza López, F.J., A.D.J. Atkinson, and J. Rodríguez Avi, 2008. Acceptance curves for the positional control of geographic data bases, *Journal of Surveying Engineering*, 134(1):26–32.

Carmel, Y., C. Flather, and D. Dean, 2006. A methodology for translating positional error into measures of attribute error, and combining the two error sources, *Proceedings of Accuracy 2006: 7th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences* (M. Caetano and M. Painho (editors), 05–07 July, Lisbon, Portugal, pp. 3–17.

Church, R., K. Curtin, P. Fohl, C. Funk, M. Goodchild, P. Kyriakidis, and V. Noronha, 1998. *Positional Distortion in Geographic Data Sets as a Barrier to Interoperation*, Technical Papers ACSM, American Congress on Surveying and Mapping, Bethesda, Maryland.

Congalton, R.G., and K. Green, 2009. *Assessing the Accuracy of Remotely Sensed Data, Principles and Practices*, Second edition, CRC Press, Boca Raton, Florida.

Department of Defense (DOD), 1990. *Mapping, Charting and Geodesy Accuracy*, MIL STD 60001, Washington, D.C.

European Spatial Data Research (EuroSDR), 2004. *Positional Accuracy Improvement: Impacts of improving the Positional Accuracy of GI Databases*, Publication No. 48, URL: *http://www.eurosdr.net/km_pub/no48/html/positional/positional_index.htm* (last date accessed: 14 February 2010).

Federal Emergency Management Agency (FEMA), 2003. *Guidelines and Specifications for Flood Hazard Mapping Partners*, Federal Emergency Management Agency.

Federal Geographic Data Committee (FGDC), 1998. *Geospatial Positioning Accuracy Standards, Part 3, National Standard for Spatial Data Accuracy,* FGDC-STD-007, URL: *http://www.fgdc.gov/standards/projects/FGDC-standards projects/accuracy/part3/chapter3* (last date accessed: 14 February 2010).

Giordano, A., and H. Veregin, 1994. *Il Controllo di Qualitá nei Sistemi Informative Territoriali, Come valuare e mantenere l' accuratezza del database*, Il Cardo, Venezia, Italy.

Heuvelink, G., 1998. *Error Propagation in Environmental Modelling*, Taylor and Francis, London, United Kingdom.

Jakobsson, A., and F. Vauglin, 2002. *Report of a Questionnaire on Data Quality in National Mapping Agencies*, Report prepared for CERCO Working Group on Quality, Comite Europeen de Responsibles da la Cartographie Officielle, Marne-la-Vallée, France.

Johnson, N.L., S. Kotz, and A.W. Kemp, 2005. *Univariate Discrete Distributions*, Vol. 1. Wiley, New York.

Krek, A., and A.U. Frank, 1999. Optimization of quality of geoinformation products, *Proceedings of 11th Annual Colloquium of the Spatial Information Research Centers*, *SIRC '99* (P.A. Whigham, editor), 13–15 December, Dunedin, New Zealand, pp. 151–159.

Merchant, D.C., 1987. Spatial accuracy specification for large scale topographic maps, *Photogrammetric Engineering & Remote Sensing*, 53(7):958–961.

Maune, D. (editor), 2007. *Digital Elevation Model Technologies and Applications: The DEM Users Manual*, Second edition, American Society for Photogrammetry and Remote Sensing. Bethesda, Maryland

National Digital Elevation Program (NDEP), 2004. *Guidelines for Digital Elevation Data*, Version 1.0, National Digital Elevation Program.

Ríos, D., S. Ríos, and J. Martín, 1997. *Simulación, Métodos y Aplicaciones*, Ra-Ma, Madrid, Spain.

U.S. Bureau of the Budget (USBB), 1947. *United States National Map Accuracy Standards*, Washington, D.C.

Weisstein, E.W., 2008. Standard Deviation Distribution, *MathWorld*, URL: *http://mathworld.wolfram.com/StandardDeviationDistribution html* (last date accessed: 14 February 2010).

(Received 17 April 2009; accepted 21 September 2009; final version 01 December 2009)