Volume 3, Number 1, 2006, pp. 87-99

The Design of a Zero-Defect Sampling with Rectification Procedure in the Presence of Classification Errors: an Application to Database with Spatial Data

Linda Lee Ho
Departamento de Engenharia de Produção – EPUSP, São Paulo, SP, Brazil
E-mail: lindalee@usp.br

José Alberto Quintanilha
Departamento de Engenharia de Transportes – EPUSP, São Paulo, SP, Brazil
E-mail: jaquinta@usp.br

Roberto C. Quinino
Departamento de Estatística - ICEX, Universidade Federal de Minas Gerais (UFMG), Minas Gerais, MG, Brazil
E-mail: roberto@est.ufmg.br

Abstract

In this paper, a zero-defect acceptance sampling with rectification is used to evaluate the quality of spatial database. In such case, quadrats are the area sampling frames and the optimum sample size to be extracted in a digital file generated from a conversion process; the size of the team of inspectors is determined so as to satisfy economic criteria. For the implementation, a program was developed using the software Matlab and it is available for readers in the Appendix. The proposed procedure is illustrated by the application to digital data on the water distribution network.

Keywords: cost function, zero-defect sampling, rectification, area sampling, spatial database quality

Introduction

Consider a situation in which a digital file is generated by a conversion process (for example, printed documents/maps are converted to a digital file). This file will be used in a geographical information application and it is necessary to evaluate if the specifications stated by the user are met (for example: specification limits and restrictions for spatial features, attribute values and other relevant aspects). Evaluating the quality of digital products is not an easy task and different aspects of the quality of a database have been discussed in the literature.

Quality is commonly used to indicate the superiority of a manufactured good or as the degree of excellence of a product, service or performance. Usually in manufacturing, quality may be stated as a desirable goal to be achieved by management and by the control of the production process (by control charts, for example). Similarly, these same issues seem to be easily extended or adapted to evaluate the quality of databases, since a database can be viewed as the result of a production process, and the reliability of the process imparts value and utility to the database. The urgency for such task in database may be justified by the increase of:

- Data production in private sectors, where there are no required quality standards. In contrast, production of data by national mapping agencies (e.g., US Geological Survey; British Ordnance Survey) has long been required to conform to national accuracy standards (i.e., mandated quality control);
- The use of geographic information system (GIS) for decision support, such that the implications of using low-quality data are becoming more widespread; and
- Reliance on secondary data sources due to the development of the Internet, data translators and data transfer standards. Thus, poor-quality data are ever easier to get.

In manufacturing, the dimensions to be controlled may be easily identified and usually classified into two main groups: attributes or variables. In data quality, users face this challenge: what are the dimensions of geographical data quality since features of the real world represented by objects, points, lines, polygons or areas in the database (for example, rivers or roads are represented by lines). According to Veregin (1999), the conventional view is that geographical data is "spatial". The terms "geographical data" and "spatial data" have been used interchangeably. However, this approach is not adequate since it ignores the inherent coupling of space and time (geographical entities are actually events unfolding over space and time) and geography is really about theme, not space. Space (or space-time) is just the framework inside which theme is measured. In the absence of theme, only geometry is present. So a better definition of geographical data may include the three dimensions: space, time and theme (where-when-what). These three dimensions are the basis for all geographical observation (Berry, 1993; Sinton, 1978) and data quality must concern several components (accuracy; precision; consistency; completeness). A brief description of the components and their contexts in the dimensions of theme, space and time follows (see Veregin, 1999);

- Accuracy is the opposite of error (a discrepancy between the encoded and actual value of a particular attribute for a given entity) and its definition is based on the entity-attribute-value model (Entities = real-world phenomena; attribute = relevant property and values = quantitative/qualitative measurements). Specifically,
    1. Thematic accuracy is the accuracy of attribute values encoded in a database;
    2. Temporal accuracy is the agreement between the encoded and "actual" temporal coordinates for an entity; and

    3. Spatial accuracy is the accuracy of the spatial component of the database.
- Precision (or resolution) refers to the amount of detail that can be discerned in space, time or theme. So,
    1. Spatial resolution is well defined in the context of raster data where it refers to the linear dimension of a cell;
    2. Temporal resolution is length (temporal duration) of the sampling interval and distinct from temporal sampling rate; and
    3. Thematic resolution refers to the precision of the measurements or categories for a particular theme.
- Consistency refers to the absence of apparent contradictions in a database; it is a measure of the internal validity of a database, and is assessed using information that is contained within the database. In this sense,
    1. Spatial consistency includes topological consistency, or conformance to topological rules;
    2. Temporal consistency is related to temporal topology, e.g., the constraint that only one event can occur at a given location at a given time; and
    3. Thematic consistency refers to the lack of contradictions in redundant thematic attributes.
- Completeness refers to the lack of errors of omission in a database. It is assessed with regard to the database specification, which defines the desired degree of generalization and abstraction (selective omission). Two types of completeness are known:
    1. "Data completeness" is a measurable error of omission observed between the database and the specification. Even highly generalized databases can be "data complete" if they contain all of the objects described in the specification; and
    2. "Model completeness" refers to the agreement between the database specification and the "abstract universe" that is required for a particular database application. A database is "model complete" if its specification is appropriate for a given application.

Many papers on database quality have been published. Some of these are worth mentioning. For example, Reingruber and Gregory (1994), Chengalur-Smith, Ballou and Pazer (1999) have pointed out the influence of database quality on the decision process. Tools to control graphic objects in a quality evaluation database process are the topic of some studies, such as Brush, Hoadley and Saperstein (1990), Leung and Yang (1998), Shi and Liu (2000) and Veregin (1999 and 2000), for example. Couclelis (1992), Nugent (1995), Liu, Shi and Tong (1999), among others, have dealt with aspects related to the process of database development.

Similar to a manufacturing process, a sample of database is randomly selected using a frame for area sampling. In spatial data, the quadrat has been the most common frame employed. Each sampling unit is evaluated to verify if it satisfies the specifications previously stated. Managers first chose a rule to decide if the database meets the specifications or not. In this paper we will consider the following sampling scheme:

1. Consider an area covered by T sheets in a fixed scale. Each sheet could be divided into N independent quadrats [see: Kish (1965); Shaw and Wheeler (1985)] of a fixed size and format (in our case, a square);
2. A random sample of $m < N$ quadrats is extracted from each sheet;
3. Each quadrat is examined separately by a team of r inspectors (it is supposed that they perform similarly). Each inspector may classify the examined quadrat as conforming or non-conforming (after the inspection, r classifications are assigned to each quadrat);
4. A quadrat is finally declared conforming if $a > 0.5$ r, (that is, more than half in r inspections classify the quadrat as conforming), where a is the number of times an examined quadrat is classified as conforming; otherwise, the quadrat is declared non-conforming; and
5. If all m sampled quadrats (in each file) are declared conforming, then the related sheet is accepted; otherwise, all N quadrats of the sheet are inspected, corrected and then the file of the examined sheet is accepted.

Figure 1 illustrates the described sampling procedure. Such sampling scheme is known as zero-defect sampling scheme with rectification in the presence of classification errors and repetitive classifications. This type of scheme is usually used to evaluate high quality manufactured processes by attributes. In manufactured processes, lots are evaluated instead of sheets and items or products are examined instead of files related to the quadrats. The above proposal can be viewed as extensions of some papers: for example, in Quinino, Ho and Suyama (2005), the authors discuss a similar problem, but the sampled items were examined only by a single inspector ($r = 1$); in Anderson, Greenberg and Stokes (2001), acceptance sampling with rectification and inspection errors are presented, but the authors do not include the use of repetitive testing; in Markowski and Markowski (2002), the authors considered acceptance sampling with misclassification errors, but not a zero-defect sampling with rectification. In Greenberg and Stokes (1995), the authors introduced repetitive testing in the presence of inspection errors, but not for a zero-defect sampling with rectification; in Greenberg and Stokes (1992) and Hahn (1986), estimators of nonconformance rates after zero defect sampling with rectification are presented, but the authors do not consider the possibility of inspection errors. (regarding inspection errors, a good review is found in Johnson, Kotz and Wu (1991)).

In this paper, the problem consists in determining two parameters: sample size m (number of quadrats to be sampled) and size r of the inspection team (number of inspectors needed) so that both (r and m) minimize a cost function. The components of such function include the cost to inspect a quadrat; the costs due to the presence of non-conforming quadrats in subsets of files in accepted sheets and the costs due to classification errors.

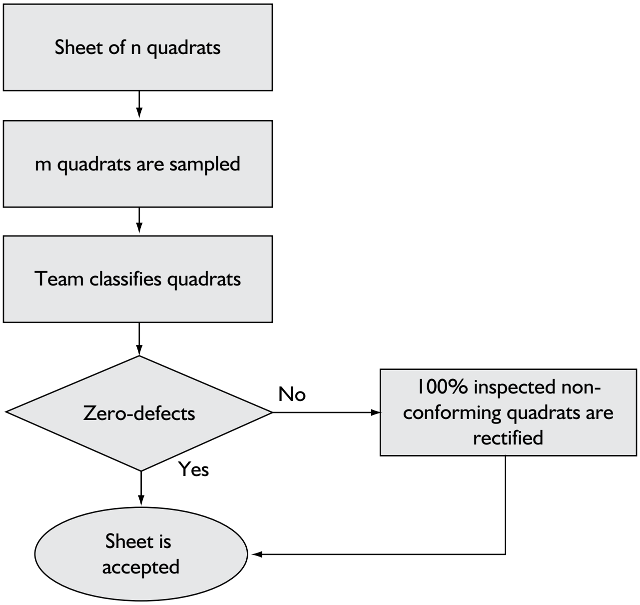

Figure 1 - Acceptance sampling: zero defect with rectification.

We introduce the notation and hypotheses considered in this paper in the next section. The expected cost function and the proposed procedure illustrated by a numerical example are the subjects of the two following sections. The last section with discussions and extensions in future works close this paper

Notation and Hypotheses

Consider an area covered by a sheet. This sheet can be divided in N independent quadrats of a fixed format and size. A random sample of m quadrats is selected. Consider the following:

D - is a binomial random variable with parameters N and p. It denotes the number of non-conforming quadrats in a file; p is the probability of a quadrat to be non-conforming. The value of p is equal to zero with probability $(1 - \pi)$. This family of distribution is flexible enough to fit well distributions related to the quality of a lot with an appropriate choice of the probability $1 - \pi$, allowing a simple interpretation and leading to a simple theory (Hald, 1981).

$D_1 \rightarrow$ denotes the number of non-conforming quadrats in m sampled quadrats;
$D_2 \rightarrow$ is the number of non-conforming quadrats in $(N - m)$ non-sampled quadrats;
$D = D_1 + D_2 \rightarrow$ is the total of non-conforming quadrats in a file;
$e_1 \rightarrow$ is the probability to classify a conforming quadrat as non-conforming;
$e_2 \rightarrow$ is the probability to classify a non-conforming quadrat as conforming;
$e_1^* \rightarrow$ is the probability to declare a conforming quadrat as non-conforming (after being examined by r inspectors);

$e_2^* \rightarrow$ is the probability to declare a non-conforming quadrat as conforming (after being examined by r inspectors);
$c_i \rightarrow$ is the cost to inspect a quadrat;
$c_{nc\_c} \rightarrow$ is the cost to classify a non-conforming quadrat as conforming;
$c_{c\_nc} \rightarrow$ is the cost to classify a conforming quadrat as non-conforming;
$Y_1 \rightarrow$ is the number of quadrats declared non-conforming in a sample of m quadrats;
$Y_2 \rightarrow$ is the number of quadrats declared non-conforming in $(N - m)$ non-sampled quadrats when they are inspected (if $Y_1 > 0$);
$Y = Y_1 + Y_2 \rightarrow$ is the total of quadrats declared non-conforming in a file;
$r \rightarrow$ is the number of inspectors (the team of inspectors); and
$D_1 | D \rightarrow$ denotes the conditional distribution of D_1 on D and assumes a hypergeometric distribution (m, D, N).

Cost Function

In this section an expected cost function (E_m) is developed employing the notations and hypotheses presented in the earlier section. The expected cost function is compounded by three parts. The first one (denoted by $E_m^{(1)}$) is related to the cost to inspect m quadrats and the possibility to inspect the $(N - m)$ non-sampled quadrats (this component is conditioned by the event $(Y_1 > 0)$, that is, when at least one of the m sampled quadrats is declared non-conforming) and it can be expressed as:

$$E_m^{(1)} = c_i \, mr \, + \, c_i \, r(N-m) \, P(Y_1 > 0) \tag{1}$$

The second part ($E_m^{(2)}$) is related to the cost of misclassifications of non-conforming quadrats classified as conforming ones. Figure 2 illustrates the possibilities of this event.

So,

$$E_m^{(2)} = c_{nc\_c} \;\; E\,[\,I_{(Y_1=0)} \;\; D\,] + c_{nc\_c} \;\; E\,[\,e_2^* \;\; I_{(Y_1>0)} \; D\,] \tag{2}$$

In which $I_{[\bullet]}$ denotes the indicator function and $E(\bullet)$, the expected value of the random variable.

And the last part ($E_m^{(3)}$) is related to the costs of misclassifications of conforming quadrats as non-conforming. In this case, all quadrats are inspected and there is the possibility of quadrats being rectified unnecessarily; this cost may be written as

$$E_m^{(3)} = c_{c\_nc} \;\; e_1^* \;\; E\,[\,I_{(Y_1>0)} \;\; (N-D)] \tag{3}$$

Summing up (3.1, 3.2 and 3.3), the expected cost function is:

$$E_m = E_m^{(1)} + E_m^{(2)} + E_m^{(3)} \tag{4}$$

where:

$$
\begin{aligned}
E_m &= c_i mr + c_i r(N-m)P(Y_1>0) + c_{nc\_c}E\,[I_{(Y_1=0)}D] \\
&\quad + c_{nc\_c}E\,[e_2^* I_{(Y_1>0)}D] + c_{c\_nc}e_1^* E\,[I_{(Y_1>0)}(N-D)] \\
&= c_i mr + c_i r(N-m)P(Y_1>0) + c_{nc\_c}E\,[(1-I_{(Y_1>0)})D] \\
&\quad + c_{nc\_c}E\,[e_2^* I_{(Y_1>0)}D] + c_{c\_nc}e_1^* E\,[I_{(Y_1>0)}(N-D)] \\
&= c_i mr + [c_i r(N-m) + c_{c\_nc}e_1^* N]P(Y_1>0) \\
&\quad + c_{nc\_c}E\,[D] - [\,c_{nc\_c}(1-e_2^*) + c_{c\_nc}e_1^*\,]E\,[I_{(Y_1>0)}(D_1+D_2)]
\end{aligned}
$$

$$P(Y_1>0) \;=\; 1-P(Y_1=0)$$

$$P(Y_1=0) = \left\{ p\,E\{[E\,[P(Y_1=0\,|\,D_1)]]\} + \;(1-p\,)\left[\sum_{x=r/2+1}^{r}\binom{r}{x}(1-e_1)^x(e_1)^{r-x}\right]^m \right\}$$

$$\left[E\left[P(Y_1=0\,|\,D_1\right]\right] = \sum_{D=0}^{N}\left[\sum_{D_1=0}^{\min(m,D)} \frac{\binom{D}{D_1}\binom{N-D}{m-D_1}}{\binom{N}{m}}\binom{N}{D}p^D(1-p)^{N-D}(1-e_1^*)^{m-D_1}(e_2^*)^{D_1}\right]$$

$$e_1^* = \sum_{x=0}^{r/2}\binom{r}{x}(1-e_1)^x(e_1)^{r-x} \;\text{ and }\; e_2^* = 1-\sum_{x=0}^{r/2}\binom{r}{x}e_2^x(1-e_2)^{r-x}$$

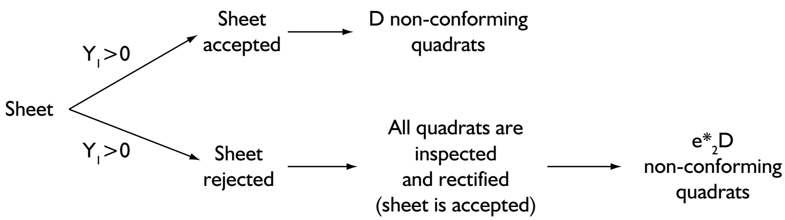

Figure 2 - Number of non-conforming quadrats when the file is accepted or rejected.

Numerical Example

The numerical example described in this section is based on the application to digital data on the water distribution network of São Paulo City, Brazil. Some layers are studied in these digital data. In GIS sense, layer is a usable subdivision of a dataset, generally containing objects of certain classes (for example, rivers, roads or geology). The GIS software use that layers to process the superimposing of two or more maps (an overlay procedure), in a layer format, through registration to a common co-ordinate system, such that the resultant layer contains the data from both layers for selected features. Figure 3 shows an example of the layers.

Volume 3, Number 1, 2006, pp. 87-99

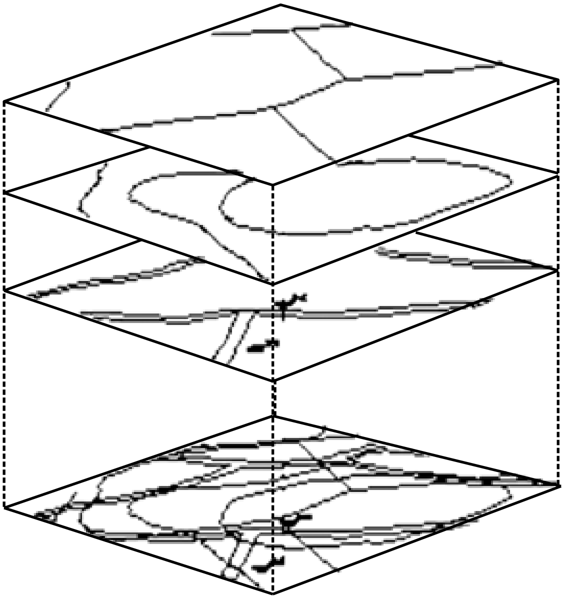

Figure 3 - Example of some layers used in GIS database (adapted from: Burrough, 1986).

One layer of interest was the block drafts, and the user wants to verify if the presence/absence of block drafts is correctly located or not. The area is covered by sheets and each one is made up by $N = 5$ thousand quadrats. They will be inspected by a zero-defect with rectification procedure and the inspection consists of checking visually the presence or absence of block drafts on the screen or in each plot. In this scenario, misclassifications are likely to occur and ideally one should design a sample plan that employs the procedure described in the previous sections. The following parameters of the process are considered:

$\pi = 0.1$; $p = 0.05$; $e_1 = e_2 = 0.0015$;

and the costs:

$c_i = \$ 1$, $c_{nc\_c} = \$ 300$ and $c_{c\_nc} = \$ 500$.

The problem is to find the optimal design: the sample size (the value of m) and the size of the team of inspectors (the value of r) that minimize the expected cost function. A program was developed in Matlab for such purpose (see the Appendix). The program provides the optimal parameters: ($m = 91$ quadrats and $r = 3$ inspectors). Such a plan results in the expected cost of \$ 1,810.80. However, if only a single inspector performed the task, a sample of 61 quadrats would be required and the expected cost increases to \$ 1,865.30 (this value is 3% higher than the previous one). Figure 4 presents the expected cost function

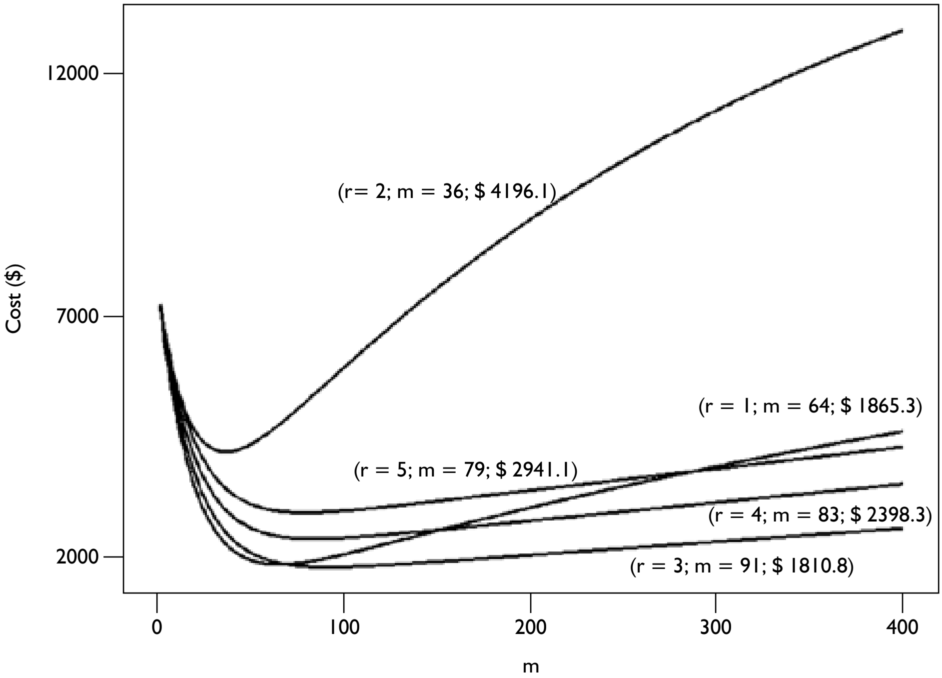

Figure 4 - Cost versus size of inspector team.

versus the sample size m given the number of inspectors (r). Note that the use of a team of two inspectors would not be recommended. It would generate a cost 2.3 times greater than the optimal policy due to the fact that the quadrat will only be declared conforming if both inspectors classify it as such.

Conclusions

This study illustrates the use of the procedure of zero-defect sampling with rectification in the presence of classification errors to evaluate the quality of a database. It is worth pointing out that misclassification errors have significant impact in determining the optimum sample size in a zero-defect with rectification procedure. As illustrated in this study, even small errors such as $e_1 = e_2 = 0.0015$, used in the numerical example, can alter significantly the value of the optimum sample size m (m^0). And in the presence of diagnosis errors, a single inspector is not the optimum design. The visual examination of the presence or absence of block drafts is not as error-free task and, in this sense, it is important to incorporate errors in the modeling in an economic perspective. Guessing the optimum parameters (the sample size and the size of the team) that may minimize a cost function is not easy, so it is advisable to run the program available in the appendix so as to draw adequately a zero-defect sampling plan.

This study can be further developed in many directions. One of them is to change the initial criteria in the sampling inspection for a limit other than zero defects, that is, $k > 0$ since a lot may be accepted if the initial sampling contains at most k > 0 non-conforming items, mainly in non high quality processes. In this study, the criteria chosen to declare a quadrat conforming were more than half in r inspections. Another possibility is to set a criterion that yields a minimum cost. In that case, the set of parameters is: the sample size, the size of the team to make the inspection and the criteria to declare an examined quadrat conforming.

Acknowledgments

The authors would like to acknowledge the referees for their worthy suggestions. This study was supported in part by *Conselho Nacional de Desenvolvimento Científico e Tecnólogico (CNPq)* and by *Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP)*.

References

Anderson, M. A.; Greenberg, B. S.; Stokes, S. L. (2001), Acceptance sampling with rectification when inspection errors are present. Journal of Quality Technology, Vol. 33, No. 4, pp. 493-505.

Berry, J. K. (1993), Beyond mapping - concepts, algorithms and issues in GIS. GIS World Inc., Fort Collins.

Brush, G. G.; Hoadley, B.; Saperstein, B. (1990), Estimating outgoing quality using the quality measurement plan. Technometrics, Vol. 32, No. 1, pp. 31-41.

Burrough, P. A. (1986), Principles of Geographical Information Systems for Land Resources Assessment. Claredon Oxford, New York, USA.

Chengalur-Smith, I. N.; Ballou, D. P.; Pazer, H. L. (1999), The impact of data quality information on decision making: an exploratory analysis. IEEE Transactions on Knowledge and Data Engineering, Vol. 11, No.6, pp. 853 - 864.

Couclelis, H. (1992), Geographic knowledge production through GIS: towards a model for quality monitoring. In: Two perspectives on data quality, Santa Barbara, National Center for Geographical Information and Analysis. (NCGIA Technical Report 92-12).

Greenberg, B. S.; Stokes, S. L. (1992), Estimating nonconformance rates after zero defect sampling with rectification. Technometrics, Vol. 34, No. 2, pp. 03-13.

Greenberg, B. S.; Stokes, S. L. (1995), Repetitive testing in the presence of inspection errors. Technometrics, Vol. 37, No. 1, pp. 102-111.

Hald, A. (1981), Statistical theory of sampling inspection by attributes. Academic Press Inc. Ltd., London.

Hahn, G. J. (1986), Estimating the percent nonconforming in the accepted product after zero defect sampling. Journal of Quality Technology, Vol. 18, No. 3, pp. 182-188.

Johnson, N. L.; Kotz, S.; Wu, X. (1991), Inspection errors for attributes in quality control. Chapman and Hall, London.

Kish, L. (1965), Survey sampling. John Wiley: New York.

Leung, Y.; Yan, J. (1998), A location error model for spatial features. International Journal of Geographical Information Science, Vol. 12, No. 6, pp. 607-620.

Liu, D. J.; Shi, W. Z.; Tong, X. H. (1999), Accuracy analysis and quality control of spatial data in GIS. Shanghai Science and Technology Documentation Press.

Markowski, E. P.; Markowski, C. A. (2002), Improved attribute acceptance sampling plans in the presence of misclassification error. European Journal of Operation Research, Vol. 139, No. 3, pp. 501-510.

Nugent, J. L. (1995), Quality control techniques for a GIS database development project. Photogrammetric Engineering and Remote Sensing, Vol. 61, No. 5, pp. 523-527.

Quinino, R. C.; Ho, L. L.; Suyama, E. (2005), Design of economically optimal zero-defect acceptance sampling with rectification when diagnosis errors are present. Pesquisa Operacional, Vol. 25, No. 1, pp. 29-44.

Reingruber, M.; Gregory, W. W. (1994), The data modeling handbook: a best-practice approach to building quality data models. John Wiley and Sons, New York.

Shaw, G.; Wheeler, D. (1985), Statistical Techniques in Geography. S. I.: Spatial indices and pattern analysis. John Wiley and Sons, New York.

Shi, W. Z.; Liu, W. B. (2000), A stochastic process-based model for the positional error of line segments in GIS. International Journal of Geographical Information Science, Vol. 14, No. 1, pp. 51-66.

Sinton, D. (1978), The Inherent Structure of Information as a Constraint to Analysis: Mapped Thematic Data as a Case Study. Harvard Papers on GIS. First International Advanced Study Symposium on topological data structures for Geographic Information Systems. Edited by G. Dutton, Vol. 7.

Veregin, H. (1999), Line simplification, geometric distortion and positional error. Cartographica, Vol. 36, No. 1, pp. 25-39.

Veregin, H. (2000), Quantifying positional error induced by line simplification. International Journal of Geographical Information Science, Vol. 14, No. 2, pp. 113-130.

Biography

Linda Lee Ho is an associate professor at the Department of Production Engineering at University of São Paulo, Brazil. She received a PhD degree from University of São Paulo, Brazil. Her current interests are statistical process control, reliability. She has published papers in the International Journal of Production Economics, European Journal of Operation

Research, Computer and Industrial Engineering, Applied Stochastic Models in Business and Industry and (in portuguese) in *Gestão & Produção and Pesquisa Operacional*.

José Alberto Quintanilha is an associate professor at the Department of Transportation Engineering of USP (University of São Paulo). He received PhD degree from University of São Paulo, Brazil. His main interests are image classification; remote sensing and spatial data quality in geographic information system applied in urban and transportation problems.

Roberto C. Quinino is an adjunct professor at the Department of Statistics of UFMG (Federal University of Minas Gerais, Brazil). He received a PhD degree from University of São Paulo, Brazil. His current interest is statistical quality control. He has published papers in the European Journal of Operation Research, Quality Engineering, Computer and Industrial Engineering, Computers & Operations Research, Applied Stochastic Models in Business and Industry and (in portuguese) in *Gestão & Produção* and *Pesquisa Operacional*.

Submitted: 20th October, 2005.
Accepted: 12th December, 2006.

Appendix

This program was developed to design a zero-defect sampling with rectification procedure in the presence of classification errors: an application to database with spatial data.

```
% Declaring global variables
clear Custo prob Um Custo0 r CustoVetor
c0 = 1; %cost of inspection
c1 = 300; % cost to classify a non-conforming quadrat as conforming
c2 = 500; % cost to classify a conforming quadrat as non-conforming
pi = 0.1;
f  = 0.05; % the probability p
N = 5000;
e1 = 0.0015;
e2 = 0.0015;
mmax = 100;
rmax = 5;
Custo0 = 1000000;
Custo = zeros(mmax,rmax);
CustoVetor = zeros((mmax*rmax),3);
Conta = 0;
for r = 1:rmax
e1Novo = binocdf((r/2),r,1-e1);
e2Novo = 1 - binocdf((r/2),r,e2);
```

```
for m = 1:mmax
clc;
mr = [m, r];
s1 = 0;
s2 = 0;
s3 = 0;
tbinom = binopdf(0:1:N,linspace(N, N, N + 1),linspace(f, f, N + 1));
for D = 0:N
minimo = min (m, D);
D1 = 0:1:minimo;
thiper = tbinom(D + 1)*hygepdf (D1,linspace(N, N, minimo + 1),linspace (D, D, minimo + 1), linspace (m, m, minimo + 1));
s1 = s1 + sum(thiper.*(1-(1-e1Novo).^(m-D1).*e2Novo.^D1).*D1);
s2 = s2 + sum(thiper.*(1-(1-e1Novo).^(m-D1).*e2Novo.^D1));
s3 = s3 + sum(thiper.*(1-e1Novo).^(m-D1).*e2Novo.^D1);
end
Um = 1-(pi*s3+(1-e1Novo)^m*(1-pi));
Custo(m, r) = c0*m*r + c0*r*(N-m)*Um + c1*pi*N*f - c1*(1-e2Novo)*pi*s1 - c1*(1-e2Novo)*pi*(N - m)*f*s2 + c2*N*e1Novo*Um - c2*e1Novo*pi*s1 - c2*e1Novo*pi*(N-m)*f*s2;
prob(m,r) = Um;
z = m;
Conta = Conta+1;
CustoVetor(Conta,1) = r;
CustoVetor(Conta,2) = m;
CustoVetor(Conta,3) = Custo(m,r);
if (Custo(m,r) < Custo0)
Custo0 = Custo(m,r);
m0 = m;
r0 = r;
end
end
end
fprintf (' r m Custo \n')
CustoVetor
fprintf (minimum Cost: %6.2f \n\n',Custo0)
fprintf ('Optimum m = %3.0f, com r = %2.0f\n',m0, r0)
```