

Managing Spatial Data Quality in the Portuguese National Ecological Reserve

Miguel PEIXOTO, Marco PAINHO, Ana Cristina COSTA, Thomas BARTOSCHEK

Abstract

The National Ecological Reserve (NER) is a basic and diversified bio-physic structure, defined and protected by Portuguese national law through land use constraints in areas with specific ecological characteristics. These areas were originally mapped in paper and this information is often required to manage several human activities, such as mineral extraction, real estate, industry, tourism, etc.

Managing spatial data in paper maps is quite different from managing digital spatial information. The objective of this experience was to produce digital cartography of Ecological National Reserve in Portugal (Alentejo), assuring its spatial data quality and conformity with the original cartography.

To assure and measure spatial data quality, a map comparison was made using a spatial sample stratified by ecological class assuring proportional area representation of each class. The spatial sampling methodology is described and the final results are shown.

At the end some conclusions of this experience are drawn.

1. Framework and Objectives

The natural patrimony preservation and appraisal can only be reached through an effective and adjusted legal protection in tune with the physical reality. The community's public conscience for this matter also plays an important role in environmental protection. The public conscience can only be enlightened by spreading specific information on this topic to the Society. The quality of the information is an important factor, because it allows to have confidence in the use of information and guarantees the adequacy of the results of this use. The NER constitutes an ideal set of geographic information to make this awareness effort, however, no so long ago, this information was in paper maps. Spreading and managing spatial data in paper maps is quite different from spreading and managing digital spatial information.

In September 14, 2001 a agreement between Instituto Superior de Estatística e Gestão de Informação of Universidade Nova de Lisboa (ISEGI-UNL) and Direcção Regional do Ambiente do Alentejo (DRA Alentejo), nowadays Comissão de Coordenação e Desenvolvimento Regional do Alentejo (CCDR Alentejo) was established and signed. The main objective of this agreement was to turn available an accurate digital geographic database of the NER in Portugal of the Alentejo region, to the Portuguese community.

The digital geographic database was created from the NER original paper cartography through digital scanning, geo-referencing and manual vectorization (Painho et al. 2005), all these processes are well known for adding error to spatial data information, thus accuracy assessment was a particularly important issue and played a major role during all this effort. This article explains the methodology used to assess the spatial data quality and conformity of the NER's digital geographic database with it's the original paper cartography.

2. The Portuguese National Ecological Reserve

The NER established in the Portuguese national law protects areas with a diversified bio-physical structure and specific ecological characteristics. Portugal has a law that defines areas of NER where these sensitive ecosystems are located stating also what kind of protection they need. In those areas it is forbidden the land division in parcels, sewer construction, building construction, hydraulic works, road construction, land cut and the vegetation destruction. The delimitation of those areas was a manual processes (Painho and Gomes, 1999) done more than 15 years ago.

The NER was first legally established in July 5, 1983 by the Decree-law nr. 321/83. Official journal 152/83 Series I., issued by the Minister of Quality of Life and was defined as “[...]an essential area for the ecological stability and sustainable management of natural resources [...]” (Decreto-Lei n.º 321/83, 1983). NER’s definition continues in a more up to date Decree-law as “[...] a basic and diverse biophysical structure that, by conditioning the use of areas with specific ecological characteristics, secures the protection of ecosystems and the remaining and intensification of the biological process indispensable to a well-balanced framing of human activities. [...]” (Decreto-Lei n.º 93/90, 1990).

The NER law was created and shaped to “[...] enable resources and land use exploitation, protecting certain functions and potentialities for which the ecological balance and the biophysical structure of regions depend upon, as well as the permanency of many of their economic, social and cultural values.[...]” (Decreto-Lei n.º 93/90, 1990).

NER covers 28 classes of ecological importance, examples are: islands, areas with erosion, areas with tendency of floods. NER maps were originally produced in paper, characterizing the classes with different shaded markers and were vectorized to produce digital cartography. This study started with the vectorization of 153 NER maps of the Alentejo-region in south Portugal, where each county produced its own maps with mostly slightly different legends. This fact, as well as a lot of cases of overlaying classes could add error to data and accuracy during the manual vectorization. And we know that, the most studied errors in GIS are those introduced during digitizing, particularly by inaccurate placement of the digitizer (Goodchild, 1989). The objective of this study is to measure the spatial accuracy and to assure the conformity with the original cartography.

3. Sampling Methodology

The purpose of the accuracy assessment is to identify and quantify positional and thematic errors in the NER - Portuguese Ecological Reserve digital geographic database, by comparing the NER classification of the digital database versus the reference paper cartography at various spatial locations.

The accuracy of the NER digital geographic database was assessed through a stratified random sampling scheme (Peixoto *et al.* 2006). This section details the sampling methodology, and in particular the statistical details for the computation of the global sample size, the stratum sample sizes, the positional and thematic errors proportions estimates and the sampling error.

Nusser and Klaas (2003) state that, ideally, accuracy estimates are based on probability samples and statistical estimation methods that provide a measure of the precision of the estimated accuracy rate. In fact, a probability-based sampling strategy has multiple advantages from which the most important one is that it

allows estimates of parameters from a limited number of units to be extrapolated, with known uncertainty, to the entire population. Among the probability sampling designs, stratified random sampling is a powerful and flexible method that is widely used in practice.

In a stratified random sampling approach, the target population is divided into H non-overlapping subpopulations, called strata. A simple random sample without replacement of units is independently selected from each stratum, where the stratum sample sizes are n_h ($h=1, \dots, H$), and the global sample size is $n=\sum n_h$.

There are several reasons for the popularity of stratified random sampling, including the following (Cochran, 1977; Lehtonen and Pahkinen, 1996): (i) stratification can enhance the precision of estimates if each stratum is homogeneous; (ii) stratification provides unbiased estimates of parameters not only for the whole population, but also for each stratum; (iii) stratification can guarantee representation of small subpopulations in the sample if desired.

The amount of increase in precision of sample estimates accomplished by stratification will depend on the degree of homogeneity that is achieved within strata, i.e. stratification enhances precision when within-stratum variation is low. When geographic strata are used, a positive spatial correlation is beneficial since positive spatial correlation indicates that geographically nearby elements have similar response values, thus producing the favourable within stratum homogeneity (Stehman, 2000).

For the accuracy assessment of the NER digital geographic database, a stratified sampling design is particularly recommended since reasonable estimates are needed for NER classes that are inherently complex or have a poor readability in the original paper map. Furthermore, since stratification ensures that a specified sample size is obtained in each stratum (NER class), including those small ones that would not be prevalent in a simple random or systematic sample without stratification, this strategy also guarantees unbiased estimates for small NER classes.

In this study, the target population is the region represented by the NER digital cartography map, and the area sampling frame is the NER region from the digital geographic database, with 739 strata defined as NER classes and regions not classified as NER, within each county. The parameters to be estimated are the digital cartography, positional and thematic, errors proportions for the entire area of interest and within each stratum. The global study area accounts for 25.510,48 Km² in Alentejo, Southern Portugal

The area sampling frame provides a description of the population and strata boundaries. The sampling protocol used with an area frame is based on first selecting a sample of spatial locations, followed by associating a sampling unit with each sampled location (Stehman and Czaplewski, 1998). The sampling algorithm and the response design implemented are detailed by Bartoschek *et al.* (2006).

The comparison of the NER classes from the digital geographic database and reference classifications is conducted on the scale of a sampling unit. The sampling units are defined as site locations (spatial coordinates) within the boundaries of the 43 counties from the digital geographic database. Therefore, the statistical population associated with this point sampling unit is viewed as continuous, rather than partitioned into discrete spatial units (e.g. pixels).

Accuracy assessment problems are usually treated in a finite population sampling framework, since the

sampling units employed, or recommended for various accuracy assessment projects, are usually defined as pixels, polygons, and fixed-area plots (Stehman and Czaplewski, 1998). The sampling design and inference for a continuous population, such as the NER region, requires a special statistical framework.

Typically, the size N_h of stratum h is the number of sampling units in stratum h . Since points are the units being sampled, N_h is the area (in square-meters) of stratum h , and N is the whole NER area from the digital geographic database. Thus, the weight W_h of stratum h is equal to its relative size:

$$W_h = N_h/N, \quad h = 1, 2, \dots, H \quad (1)$$

The global sample size n is the total number of units to be sampled. Allocation provides a tool for determining the number of sample units to be taken from each stratum under the constraint that the total number of units to be sampled is n . Proportional allocation means that strata are sampled in proportion to their size. Thus, the number of sample units n_h in stratum h is given by:

$$n_h = W_h \times n, \quad h = 1, 2, \dots, H \quad (2)$$

In this study, a projected global sample size is first determined taking into consideration the desired sampling error and the sample proportional allocation. Let us first introduce some additional notation and definitions. The most important parameters to be estimated are the digital geographic database, positional and thematic, errors proportions for the entire area of interest, which are denoted by p_p and p_t , respectively. The parameters $p_{p,h}$ and $p_{t,h}$ denote the digital geographic database, positional and thematic, errors proportions within each stratum h , respectively. Since the formulations for the parameters p_p and p_t ($p_{p,h}$ and $p_{t,h}$) are analogous, the notation p (p_h) will be used to designate any of these errors proportions.

The estimator of a given parameter θ is denoted by $\hat{\theta}$, the variance of $\hat{\theta}$ by $V(\hat{\theta})$, and the estimator of this variance is denoted by $\hat{V}(\hat{\theta})$. Hence, the sampling error of the estimation of a given parameter θ is defined by:

$$d = z_{1-\alpha/2} \sqrt{V(\hat{\theta})} \quad (3)$$

where $z_{1-\alpha/2}$ is the $1-\alpha/2$ quantile of the standard normal distribution, for a given confidence level $(1-\alpha)\%$.

For stratified random sampling, an unbiased estimator of p is:

$$\hat{p} = \sum_{h=1}^H W_h \hat{p}_h \quad (4)$$

where W_h is the weight of stratum h given in equation (1), and \hat{p}_h is the errors proportion observed in the sample of the h stratum.

The variance of this estimator is given by:

$$V(\hat{p}) = \sum_{h=1}^H W_h^2 (1-f_h) \frac{p_h(1-p_h)}{n_h} \quad (5)$$

and its unbiased estimator is:

$$\hat{V}(\hat{p}) = \sum_{h=1}^H W_h^2 (1-f_h) \frac{\hat{p}_h(1-\hat{p}_h)}{n_h-1} \quad (6)$$

where $(1-f_h)$ is the finite population correction factor for stratum h ($f_h = n_h/N_h$ is the stratum h sampling rate).

For stratified random sampling with proportional allocation (STP), the variance of the estimator of p

becomes (by replacing n_h from equation (2) in equation (5)):

$$V(\hat{p})_{STP} = (1-f) \frac{S_{within}^2}{n} \quad (7)$$

where $(1-f)$ is the finite population correction factor ($f=n/N$ is the sampling rate), and the within-stratum variance is given by:

$$S_{within}^2 = \sum_{h=1}^H W_h p_h (1-p_h) \quad (8)$$

In the infinite population sampling framework the $(1-f)$ correction factor is equal to 1 (as the sampling rate is equal to zero), thus equation (7) becomes:

$$V(\hat{p})_{STP,\infty} = \frac{S_{within}^2}{n} \quad (9)$$

Replacing the variance of the STP estimator of p (from equation 9) in equation (3), and then solving for sample size, produces the formula to compute the global sample size with the desired sampling error precision for the estimation of p , under an infinite population STP approach:

$$n = \left(\frac{z_{1-\alpha/2}}{d} \right)^2 S_{within}^2 \quad (10)$$

Sample size depends on the within-stratum variance which assumes the knowledge of the true proportion of errors within each stratum. This cannot be done without assuming a value for the unknown true accuracy. The most conservative strategy is to consider an accuracy of 50% for all strata (i.e. $p_h=0.5$ for all strata), because $p_h(1-p_h)$ is always less than or equal to 0.25.

In order to overcome this limitation, and to take into consideration the complexity and uncertainty inherent to each NER class, the results from a previous pilot study were used to set up the value of p_h for each stratum h . The pilot study used a simple random sample of 1000 units which was post-stratified by major NER classes within each county. The overall positional and thematic accuracy estimate was approximately equal to 98% ($\hat{p}_{pilot} = 0.02$). Not distinguishing positional from thematic errors, the errors proportion within each post-stratum were also estimated and the highest one was equal to 17%.

The computation of the projected global sample size assumed a sampling error precision for the estimation of p equal to 1% (d value), a 95% confidence level ($z_{1-\alpha/2} = 1.96$), and the value of p_h for each stratum h was set up using the following criterions:

- p_h was assumed equal to 0.02 for strata corresponding to regions that were not classified as NER within each county, because they are expect to have a high accuracy rate in the digital geographic database and the overall pilot study accuracy estimate was 98%;
- p_h was assumed equal to its corresponding estimated value in the pilot study (strata matching major NER classes within each county) whenever that estimate was different from zero;
- p_h was assumed equal to 0.07 whenever the corresponding pilot study estimated value (strata matching major NER classes within each county) was equal to zero or the reference paper cartography had poor readability;
- p_h was assumed equal to 0.17 for strata with high complexity and uncertainty (small strata that result

from crossing major NER classes within each county), because that was the highest error rate estimated in the pilot study.

The global sample size determined using this methodology is equal to 1668 units. Consequently, after the sample proportional allocation, several strata sample sizes were computed as zero or one, due to their small areas. However, at least two units per stratum are required to estimate the within-stratum variability. Therefore, the sample sizes of those strata were set to two, in order to assess the precision of the errors proportions estimates in all strata. Thus, the final global sample size is equal to 2668 units and, consequently, the sampling error is expected to be below 1%, for a 95% confidence level.

The accuracy of the NER digital geographic database was assessed through equation (4), which was used to estimate both positional and thematic errors proportions for the entire area of interest; and a 95% confidence interval was computed as:

$$\hat{p} \pm 1.96\sqrt{\hat{V}(\hat{p})} \quad (11)$$

where the $\hat{V}(\hat{p})$ formulation is given by equation (6) with the finite population correction factor for each stratum h set to one. The sampling error of the estimation of p is the half confidence interval width.

Note that an estimator of $V(\hat{p})_{STP,\infty}$ given by equation (9) can no longer be used, for this sampling design, since the final global sample was not proportionally allocated to strata.

Regarding the estimation of the remaining parameters of interest, each \hat{p}_h is computed as the digital geographic database errors proportion observed in the sample of the h stratum. In order to assess the precision of those estimates, the variance of each \hat{p}_h is estimated by:

$$\hat{V}(\hat{p}_h) = \frac{\hat{p}_h(1-\hat{p}_h)}{n_h-1} \quad (12)$$

Strata sampling errors and 95% confidence intervals for all p_h can be computed using an analogous formulation of equation (11).

4. Computational Processes

The manner in which spatial data is represented in an information system is key to the efficiency of the computational processes that will act upon it (Worboys and Duckham, 2004). The application works with spatial data given in shape or coverage format, which represents a simple geodatabase. Upon this database the calculations and randomizing computations can be done very efficiently and interoperability is granted by using the same output formats. Figure 1 gives a brief overview about the computed working processes, their input and output.

4.1 Data Preparation and Calculations

To assure a sufficient global sample size and probable testing conditions we first need to know about the classes and class combinations in the NER data in each county. The input data contains overlaying polygons with each having one single class as attribute. As described below we use an object-oriented algorithm to get all existing class combinations and to calculate their respective area. As part of the equation for the global and stratum sample size we need these areas and their relation to the stratum size.

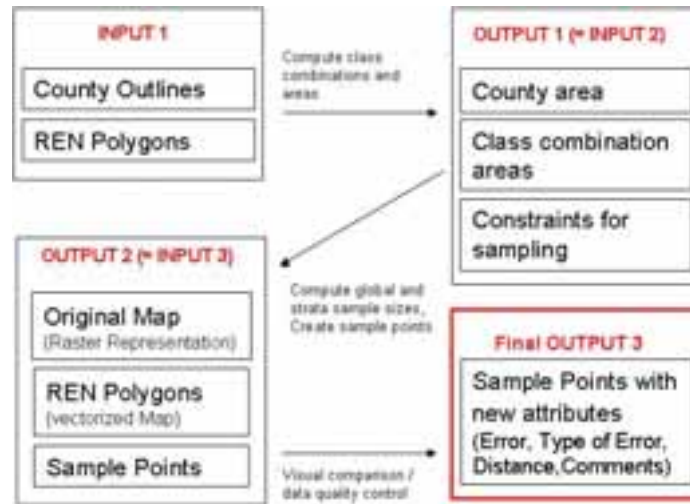


Figure 1 - Input/Output diagram for computational processes

4.1.1 Algorithm

The algorithm uses two classes representing a NER-class combination and a collection of NER-class combinations. It can be applied to all counties at the same time, if they are added as different layers to the map. First a collection of NER-class names is set. For each layer in the map (county) we set a new object of the combinations-collection-class. Then for each record we encode a binary string with 1 representing a contained class and 0 a not contained, respective to the sorting of the collection of NER-class names. If the encoded string exists in the combinations-object, its area is added to the area saved before, if not, a new combination object is created and its name and area is saved. After the computation the results can be printed in an ASCII file or shown on the screen.

The result of this algorithm is a list of all single classes and class combinations with the respective area for each stratum, in our case county.

With the information of the number of classes and class combinations, the area of each class and class combination, the area of non-NER polygons, the total county area and constraints to assure probable testing of spatial accuracy and data quality assessment the sample size per class combination is calculated, following the equations in topic 3. Sampling Methodology.

4.2 Sampling

The inputs in this stage of processing are the areas of all NER classes and class combinations, for each county and the county area. This data is used to create some constraints for the sampling like global sample size, stratum sample size (for each county), assurance of considering each class and class combination with at least one point. These constraints are discussed in Peixoto *et al.* (2006). After following the constraints the sampling is done by the use of a special randomizing algorithm explained as followed.

4.2.1 Randomizing Algorithm

The input for the algorithm is a list of NER classes and class combinations with the respective number of sample points. Out of the string, providing the NER class and class combinations, we create a SQL

expression which is passed to a randomize-function together with the number of points for this class or class combination. The SQL expression is used to query the layer and create a selection containing the proper classes. This recordset is mostly composed by many polygons. The randomize function chooses first randomly one of these polygons, gets its extent, creates randomly a point using the polygons extent coordinates, checking if the point is inside the polygon (the extent coordinates are rectangular – the polygons are not), if the point is inside it continues until the number of points per class is reached. If the point is outside the polygon, it creates new points randomly until one is inside the polygon. Finally a Shapefile with the sample points and their class as an attribute is created and added as a layer to the map.

4.3 User Interface

An application's interface is of extreme importance. It is inextricably related to the usefulness of a geographic information system. An interface should be easy to learn, appear natural, and be independent of implementation complexities such as data structures and algorithms (Egenhofer and Frank, 1988). The applications User Interface design is kept simple; it contains basic GIS-Viewer capacities like zooming, panning, measuring and an identification button in the toolbar, to provide a deeper map analysis and interaction. A legend and a map control are centred and buttons for visual comparison are placed in the bottom. An extra form containing a zoom on the raster maps legend can be opened too facilitate the data quality control.

4.3.1 Functionalities

To facilitate the comparison of the raster map versus the vectorized representation the User Interface takes some work usually done on the user side. After creating the random points some map properties are changed automatically. The NER-layer is shown in a bright transparent colour, the names of classes/class combinations are centred in each polygon. The user adds the proper raster image to the project and can open an extra form with a zoom on the raster legend. This helps comparing the raster map without scrolling to the legend or printing it. After adding the map the user sees the raster map in the background, above it are the NER polygon layer and the sample points layer, with point attributes (class and class combinations names) are shown. Beside that the user has two arrows to navigate through the actual dataset. By clicking analyze points, the application makes a zoom to the first record, centering at the point and allowing the user to compare directly the class or class combination on the raster map with the class or class combination in the point sample, which was read out from the NER layer. If the class / class combination is equal the user continues comparing the points. If the user finds an error, he has to classify it (thematic or spatial error) and can save a comment as attribute into the shape file. The newly changed class or class combination is also saved in the shapefile and in the case of spatial errors the distance to the probable polygon with the correct class is calculated.

4.3.2 Implementation

The tool was implemented with Visual Basic 6.0 due to its capacities in progress effectiveness, using the component software technology Map Objects® 2.2. from ESRI.



Figure 2 - RENalyzer during visual comparison process

5. Results and Discussion

The result of this study is a stand-alone application, providing help and facilitation in data quality assessment by implementing the statistical framework and automating a lot of work processes. The class names and sample sizes are variable and easily changeable so the tool is applicable to different cases in different frameworks. The output of the tool is a shapefile for each county containing the sample points, their original attributes from the vectorized map. In case of thematic error the corrected attributes are saved and in case of spatial error the distance to the nearest correct feature is also saved. The shape format allows interoperability with other GIS tools for deeper data analysis and accuracy assessment. The data quality control is done by visual comparison through the user, so the source of probable errors could produce new errors during comparison. But, controlling the points out of the global sample size is comparatively few work and errors are little too.

5.1 Thematic Accuracy

The thematic accuracy estimate of the NER digital geographic database, for the entire area of interest, is equal to 99,7%. For a 95% confidence level, the sampling error of this estimation is equal to 0,21% (which is below 1%, as expected). Hence, a 95% confidence interval for the thematic errors proportion, for the entire area of interest, ranges from 0,06% to 0,47%.

Only 2,8% of the 739 strata defined have thematic errors. Within those strata, the minimum estimate for the thematic errors proportions is equal to 5%. Conversely, there are 9 strata (1% of all strata) with an estimated proportion of errors equal to 100%, concerning to very small and complex NER classes.

The average thematic accuracy estimate is approximately equal to 65,6%, within the 21 strata with thematic errors. The average sampling error of this estimation is equal to 45%, for a 95% confidence level.

Strata corresponding to regions that were not classified as NER within each county do not have positional errors. Only 3 of those 43 strata have thematic errors, and their maximum errors proportion estimate is equal to 16.7%. The average thematic accuracy estimate of those 43 strata is approximately equal to 99%, as expected. Figure 3 gives a brief overview about the thematic errors found by county in Alentejo.

5.2 Positional Accuracy

The positional accuracy estimate of the NER digital geographic database, for the entire area of interest, is equal to 99,2%. For a 95% confidence level, the sampling error of this estimation is equal to 0,41% (which is below 1%, as expected). Consequently, a 95% confidence interval for the positional errors proportion, for the entire area of interest, ranges from 0,37% to 1,2%.



Figure 3 - Thematic errors found by county in Alentejo.

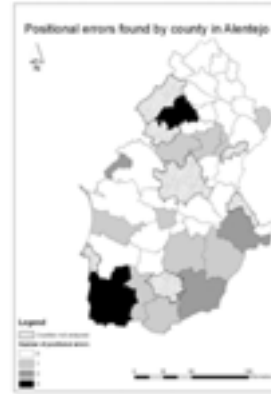


Figure 4 - Positional errors found by county in Alentejo.

Just 3,1% of the 739 strata defined have positional errors. Within those strata, the minimum and maximum estimates for the positional errors proportions are equal to 2% and 50%, respectively.

The average positional accuracy estimate is approximately equal to 33,7%, within the 23 strata with positional errors. The average sampling error of this estimation is equal to 33,7%, for a 95% confidence level. Figure 4 gives a brief overview about the positional errors found by county in Alentejo.

5.3 Overall Accuracy

The overall accuracy of the NER digital geographic database is assessed through the sum of the positional and thematic errors proportions estimates, for the entire area of interest, and it is approximately equal to 98,9%.



Figure 5 - Overall errors found by county in Alentejo.

For a 95% confidence level, the sampling error of this estimation is equal to 0,46%. Thus, a 95% confidence interval for the positional plus thematic errors proportion, for the entire area of interest, ranges from 0,59% to 1,51%.

Only 5,8% of the 739 strata defined have some kind of error. Within those strata, the minimum estimate for the errors proportion is equal to 2%. The average accuracy estimate is approximately equal to 50%, within the 43 strata with some kind of error. The average sampling error of the estimation, within strata with errors, is equal to 56%, for a 95% confidence level.

6. Conclusion

In view of these results, we conclude that the NER digital geographic database, once compared with its original map version in paper, is highly accurate. Moreover, the stratified random sampling design implemented turned out to be an appropriate and valuable framework for the spatial accuracy assessment of digital cartography. These positive results help to conclude that RENalyzer is a tool that provides helpful functions for spatial accuracy assessment in digitalized cartography, the work process is much shorter than the analysis in a full GIS, which costs a lot of manual operations and lots of time and money also.

References:

- Bartoschek, T., Painho, M., Henriques, R., Peixoto, M. and Costa, A. C., 2006, RENalyzer: A tool to facilitate the spatial accuracy assessment of digital cartography, in: Spatial Accuracy Research Group (Eds.), Proceedings of the Seventh International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences.
- Cochran, W. G., 1977, Sampling Techniques. 3rd Edition, John Wiley & Sons, New York.
- Decreto-Lei n.º 321/83, Diário da República n.º 152/83, Série I - de 05 de Julho de 1983.
- Decreto-Lei n.º 93/90, Diário da República n.º 65/90, Série I - de 19 de Março de 1990.
- Egenhofer, M. and Frank, A., 1988, Designing Object-Oriented Query Languages for GIS: Human Interface Aspects, Proceedings of the Third International Symposium on Spatial Data Handling, International Geographical Union Commission and Mapping; Falls Church, VA.
- Goodchild, M., Gopal, S., (1989), Accuracy of Spatial Databases, London: Taylor & Francis.
- Lehtonen, R. and Pahkinen, E. J., 1996, Practical Methods for Design and Analysis of Complex Surveys. Revised Edition, November 1996, Statistics in Practice, John Wiley & Sons, Chichester.
- Nusser, S. M. and Klaas, E. E., 2003, Survey methods for assessing land cover map accuracy. Environmental and Ecological Statistics 10, 309–331.
- Painho, M. and Gomes, R., 1999, "Using GIS on the delimitation of a National Ecological Reserve", Third Annual Conference of the Society for Conservation GIS, Idlewild, CA July 24-26, 1999.
- Painho, M., Peixoto, M., and Brito, P., 2005, "Relatório Final: SIG de Ambiente e Recursos Naturais para o Alentejo" para a Direcção Regional de Ambiente e Ordenamento do Território do Alentejo, ISEGI.
- Peixoto, M., Costa, A. C., Painho, M., Bartoschek, T., 2006, A stratified sampling approach to the spatial accuracy assessment of digital cartography: an application to the Portuguese Ecological Reserve, in: Spatial Accuracy Research Group (Eds.), Proceedings of the Seventh International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences.
- Stehman, S. V., 2000, Practical Implications of Design-Based Sampling Inference for Thematic Map Accuracy Assessment. Remote Sens. Environ. 72, 35–45.
- Stehman, S. V. and Czaplewski, R. L., 1998, Design and Analysis for Thematic Map Accuracy Assessment: Fundamental Principles. Remote Sens. Environ. 64, 331–344.
- Worboys, M. and Duckham, M., 2004, GIS A Computing Perspective – 2nd ed., CRC Press LLC.

Author Information:

Miguel PEIXOTO, Dr - mpeixoto@isegi.unl.pt

Marco PAINHO, PhD - painho@isegi.unl.pt

Ana Cristina COSTA, MSc - ccosta@isegi.unl.pt

Instituto Superior de Estatística e Gestão de Informação - Universidade Nova de Lisboa

Campus de Campolide, 1070-312 Lisboa, Portugal

Telephone: +351 21 387 04 13

Fax: +351 21 387 21 40

Thomas BARTOSCHEK, Dr - bartoschek@uni-muenster.de

Institute for Geoinformatics - University of Münster

Robert-Koch-Str. 26-28, 48149 Münster, Germany

Telephone: +49 (0) 251 83 33 083

Fax: +49 (0) 251 83 39 763