7th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences. Edited by M. Caetano and M. Painho.

Reducing positional error in spatio-temporal analyses

Jean-François Mas

Unidad Académica Morelia,
Instituto de Geografía, Universidad Nacional Autónoma de México,
Aquilés Serdán 382, Colonia Centro, Morelia, Michoacán, México
Tel.: + 00 52 443 317 94 23; Fax: + 00 52 443 317 94 25
jfmas@igg.unam.unam.mx

Abstract

A frequent method used to assess land use / land cover (LULC) change is the comparison (overlay) of digital maps of an area within a geographic information system (GIS). However, positional errors of the maps involved in the comparison tend to create "false" change polygons and to overestimate LULC change. In this work, a simple method to improve change estimates by detecting and correcting erroneous changes resulting from positional errors is presented. In this approach, the boundary uncertainty is managed using the epsilon band which encloses a confidence region with a specific probability of including the true location of the boundary. In order to test the method, a map of change was first generated through the overlay of two digital versions of the same analogical cartography. False changes resulting from both positional (digitizing) and thematic (labelling) errors were quantified. Finally, the method was applied to a LULC change monitoring project in a region of South-eastern Mexico, which has been undergoing important land cover changes during the past decades. In both cases, the method allowed a significant reduction of erroneous changes due to positional errors. Due to the complexity of error modelling approaches, such simple methods are likely to be the most useful in the context of practical projects aimed at assessing LULC change.

Keywords: spatial accuracy, positional error, epsilon band, land cover change monitoring

1 Introduction

Human activities dominate the Earth's global ecosystem and Land Use and Land Cover (LULC) change is one of the most pervasive and influential activities. LULC change alters nutrient cycling, hydrology, species diversity as well as ecosystem and community functioning (Jarnagin, 2004). The detection and the quantification of temporal LULC change is the subject of a growing number of studies during the last decade (Ochoa and González, 2000).

LULC change assessment can be accomplished though the comparison of LULC maps elaborated on different dates. Maps are co-registered and overlaid within a GIS in order to enhance differences between the two dates. This operation allows the mapping of the changes, the elaboration of the transition matrix and the computing of the rate of change. However, two main sources of errors can affect the change analysis: positional errors and thematic errors. Positional errors are related to the class boundary location whereas thematic errors arise from the incorrect assignment of category to map polygons. In the map of change, positional errors lead to the generation of false polygons of changes, known as "sliver polygons" due to different delimitations of the same unchanged polygon or to the translational error of the entire polygon (Figure 1). As a result, much of the change depicted by map comparison may be an artefact of these errors. For example, Mas et al. (2004) reported that the differences between the two digital versions of the same cartography, due to both thematic errors and positional errors introduced during the digitalization, affected 12% of the mapped area.

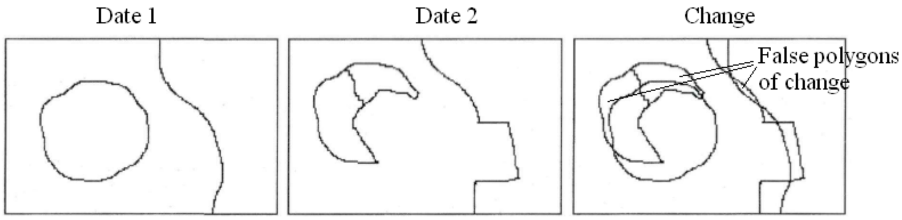

Figure 1 Effect of positional errors on LULC change assessment.

2 Methods

In the present study, the uncertainty of polygon boundary position was managed using the epsilon band approach. This band encloses a confidence region with a specific probability of including the true location of the boundary (Drummond 1990, Goodchild 1993, Aspinal and Pearson 1995). Assuming a normal distribution of errors, epsilon bands extended out by one and 1.96 standard (root mean square) deviation in either direction would be expected to include the true location of the boundary 68% and 95% of the time, respectively. As the mean deviation between two lines corresponds to the mean width of the sliver polygons created by the overlay of the two parent lines, sliver polygons were identified on the map of changes. The value of the mean width was approximated by dividing the area by half the perimeter and, assuming a normal distribution of deviation, the standard deviation was calculated as 1.48 mean deviations (Green and Hartley 2000). In this study, epsilon bands of a total width of two mean deviations (1.35 standard deviations) were generated and overlaid with the map of LULC change. The polygons of change entirely included inside the epsilon band were considered as “false change” and were deleted (Figure 2). After this procedure, polygons of change would only remain when the two parent boundaries deviated by an amount greater than two mean deviations. Assuming a normal distribution of error, only 18% of deviation exceeds twice the mean deviation and consequently most of the positional boundary errors should have been removed from the overlay analysis.

In order to evaluate this method, a map of LULC change was generated through the overlay of two digital versions of the same LULC cartography with a scale of 1:250,000. These maps present differences in the polygon boundary location (positional error) and label (thematic error) due to errors during the digitizing process. False changes resulting from both positional and thematic errors were quantified. Then the epsilon band method was applied in order to evaluate the reduction of these errors. The McNemar’s test was applied in order to test the statistical significance of the reduction of the false change polygon area. This test is based upon a two-by-two confusion matrix which describes correct and incorrect class allocations presented by the two maps being compared (Table 1 and equation 1) (Foody, 2004).

Table 1 Error matrix used in the McNemar´s test.

	Date 2	
Date 1	Correct	Incorrect
Correct	f_{11}	f_{12}
Incorrect	f_{21}	f_{22}

$$z^2 = \frac{(f_{12} - f_{21})^2}{f_{12} + f_{21}} \tag{1}$$

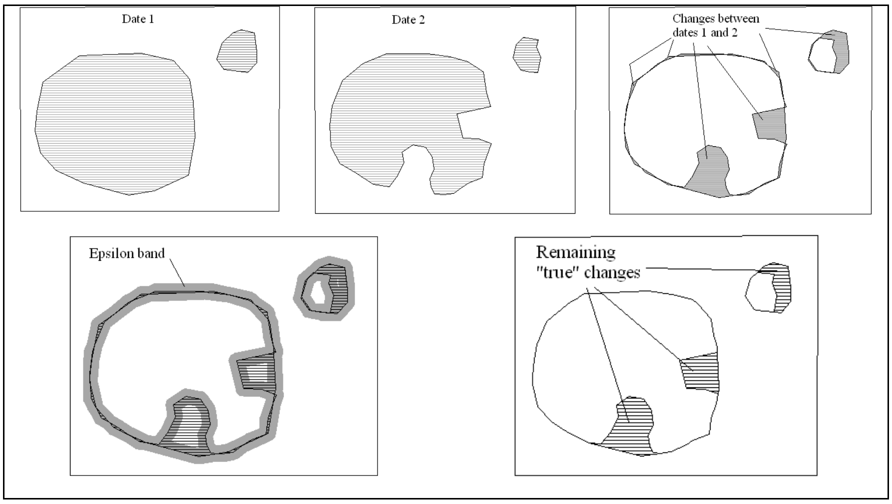

Figure 2 Detection and deletion of false polygons of change.

3 Results

In the comparison of ideal digital versions of the same map, that is perfectly identical with the original map, no change is expected. However, more than 130,000 polygons, covering 9.6% of the area, were depicted as change. This proportion of changes correspond approximately to the deforested area which occurred during a decade with an annual rate of change of 0.9%, the rate of tropical deforestation in Mexico (Mas et al. 2004).

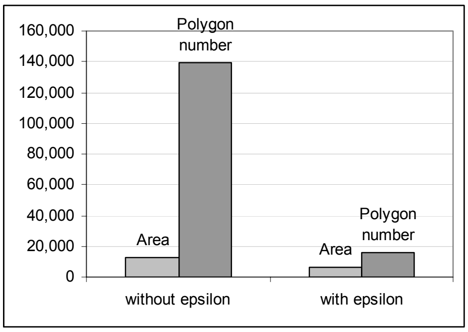

Figure 3 Statistics of the map of changes with and without the epsilon band. The method allows an important reduction of both number and area (in square kilometers) of false change.

The mean positional error was estimated in 103 m and therefore an epsilon band of 206 m width was generated. A total of 123,396 polygons, located inside the epsilon band, were eliminated, allowing a reduction of false change from 9.6 to 4.7% of the study area (Figure 3). The McNemar's test indicates that there is a significant difference in the accuracy of the maps of changes before and after the application of the epsilon band improvement. Approximately 75% of the remaining false change area corresponds to false change polygons due to labelling errors. Therefore, the method allowed the elimination of most of the erroneous change due to positional error.

The two maps of Quintana Roo were overlaid to generate the map of LULC change. The mean positional error, estimated in 83 m, was derived from the sliver polygons around water bodies whose boundaries did not show change. This estimation took into account the error due to the geometric correction of the images and the error related with digitization but not the subjectivity in delimiting the transition between different categories such as evergreen and lowland flooded forests. A total of 1,268 change polygons (1.8% of the study area) were identified as false change due to positional error (epsilon band width 83 m). The application of the epsilon did not modify the rate of change of the main categories (agriculture and pasture; evergreen forest) but the matrix of transition derived from the map of LULC change corrected by the epsilon method is more realistic because it presents fewer improbable changes, such as changes between the two types of forests. Although the method did not eliminate all of the errors, it allowed an improvement of the reliability of the map of LULC changes. A more detailed description of this method along with applications in LULC change assessment can be found in Mas (2005).

4 Conclusion

Additional improvements of this method can be obtained using different epsilon band widths depending on the type of boundary between land cover classes (larger band for ecotone or spectrally confused limits). The visual revision of the change polygons identified as false change by epsilon bands constructed with different widths and the manual editing of the map of changes also can allow better results but is time consuming. This method allows a significant reduction of the error related with LULC change assessment by map comparison. However, a large amount of error in change assessment can be due to thematic (labelling) errors and the present approach does not allow the management of this type of error.

This Method is helpful in evaluating the errors introduced by positional errors in the change estimates and in improving these estimates significantly in most cases. Although error modelling can be very complex, such simple methods are likely to be the most useful in the context of practical LULC assessments.

Acknowledgements

This study was carried out under the project "Evaluación de la dinámica espacio-temporal de la frontera agrícola de Quintana Roo (1978–2003)" reference QROO-2003-C02-13228 supported by the *Fondos Mixtos CONACYT-Gobierno del Estado de Quintana Roo*. I would like to thank José Antonio Navarrete Pacheco for elaborating the map of Quintana Roo dated 1990.

References

Aspinal, R.J. and Pearson, D.M., 1995, Describing and managing uncertainty of categorical maps in GIS. In *Innovations in GIS 2*, P. Fisher (Ed.), London: Taylor and Francis.

Drummond, J., 1990, A framework for handling errors in geographic data manipulation. In *Fundamentals of Geographic Information Systems: A Compendium*, W.J. Ripple (Ed), American Society for Photogrammetry and Remote Sensing.

Foody, G.M., 2004, Thematic Map Comparison: Evaluating the Statistical Significance of Differences in Classification Accuracy. *Photogrammetric Engineering and Remote Sensing,* 70(5), pp. 627-633.

Goodchild, M.F., 1993, Data models and data quality: Problems and prospects. In *Environmental Modeling with GIS*, Goodchild M.F., Parks B.O. and Steyaert L.T. (Eds), pp. 94-102, New York, Oxford University Press.

Green, D.R. and Hartley, S., 2000, Integrating Photointerpretation and GIS for Vegetation Mapping: Some Issues of Error, In *Vegetation Mapping: From Patch to Planet*, R. Alexander and A.C. Millington, John Wiley and Sons Ltd.

Jarnagin, S. T., 2004, Regional and global patterns of population, land use and land cover change: an overview of stressors and impacts. *Giscience and Remote Sensing* 41(3), pp. 207-227.

Mas, J.F., 2005, Change estimates by map comparison: A method to reduce erroneous changes due to positional error. *Transactions in GIS*, 9(4), pp. 619-629.

Mas, J.F., A. Velázquez, J.R. Díaz-Gallegos, R. Mayorga-Saucedo, C. Alcántara, G. Bocco, R. Castro, T. Fernández y A. Pérez-Vega, 2004, Assessing land use/cover changes : a nationwide multidate spatial database for Mexico. *International Journal of Applied Earth Observation and Geoinformation*, 5(4), pp. 249-261.

Ochoa, S. and González, M., 2000, Land use and deforestation in the highlands of Chiapas, Mexico. *Applied Geography* (20), pp. 17-42.