Analysis and modeling of land-use and land-cover change in Sintra-Cascais area

Pedro Cabral[1,2,3], Hélène Geroyannis[2], Jean-Paul Gilg[2] and Marco Painho[1]

[1]Instituto Superior de Estatística e Gestão de Informação – Universidade Nova de Lisboa, Portugal

[2]École des Hautes Études en Sciences Sociales - CAMS (UMR 8557 CNRS), France

[3]Contact author: pcabral@isegi.unl.pt

SUMMARY

This paper shows a Geographical Information Systems & Science (GISc) approach for modeling land use and land cover change (LUCC) in a rapid urban growing region of Portugal. The Sintra-Cascais area is integrated in the Lisbon Metropolitan Area which faced unparalleled rates of urbanization in the last thirty years. The study area is particularly sensitive as it contains the Natural Park of Sintra-Cascais which is subject to intensive anthropogenic pressure. Three Landsat images (TM 1989, TM 1994 and ETM+ 2000) were classified in three major land use/cover classes: woodland, grassland and impervious. A stochastic model was used to estimate the evolution of these classes for year 2006 and evaluate the importance of the natural park in the land use and land cover dynamics of this region. Robustness of the model is assessed by comparing estimated year 2000 classification against real year 2000 classification. Description and analysis of LUCC is made using selected spatial metrics. Results demonstrate that the natural park plays an important role in preventing the increase in the impermeabilization of soil in this region.

KEYWORDS: *GISc, remote sensing, LUCC, stochastic modeling, spatial metrics*

INTRODUCTION

In 2015, more than half of the world's population will be urban (NIC, 2000; UNECE, 2003). Urban sprawl is a phenomenon that has to be monitored and understood. Tools and scientific knowledge are available for modeling spatial dynamics. This skill can be used for planning sustainable growth of fast-growing areas. Spatial models are needed for understanding reality and comprise a temporal dimension. Models are probably the most concise and useful way to understand spatial dynamics. There are different approaches for modeling spatial dynamics. These approaches can be classified in cellular automata modeling, agent-based modeling, stochastic modeling, neural network modeling, fractal modeling and others, according to the methods available for modeling complexity and non-linearity (Weng, 2002; Cheng, 2003; Benenson & Torrens, 2004). Models cannot work without data and satellite imagery is an excelent source of data. Fortunately, there are archives of satellite imagery since 1972 with the launch of Landsat Multispectral Scanner (MSS). In the meantime, a plethora of new satelites, with different sensors and resolutions, have produced temporal series of images. Techniques are required to deal with time series of data that are produced from different sensors, thus having different spectral, temporal, radiometric, and spatial resolutions. Digital image processing and classification techniques have been improved and research has been dedicated to deal with time series of satelite data in spatial dynamics studies (Masek et al., 2000; Kaufmann & Seto, 2001; Gluch, 2002; Read & Lam, 2002; Seto et al., 2002; Herold et al., 2003; Cabral et al., 2004).

In this paper, a GISc approach is used for analysis and modeling LUCC in a fast urban growing region between 1989 and 2006 using remotely sensed data. Three Landsat images (TM 1989, TM 1994 and ETM+ 2000) are classified using a maximum likelihood algorithm with training sites obtained by image segmentation. Selected spatial metrics are used to describe LUCC dynamics of this region between 1989 and 2000. Similarity of LUCC transition mechanism is evaluated for years 1989 to 1994 and years 1994 to 2000. Having prooved this similarity, estimated values of LUCC are presented for year 2006 using a stochastic model. Finally, the influence of the Natural Park of Sintra Cascais is evaluated in the dynamics of the LUCC of this region.

STUDY AREA

The study area analyzed in this study comprises the Sintra and Cascais municipalities with an area of approximately 416 Km2. These municipalities are integrated in the Lisbon Metropolitan Area (Figure 1).

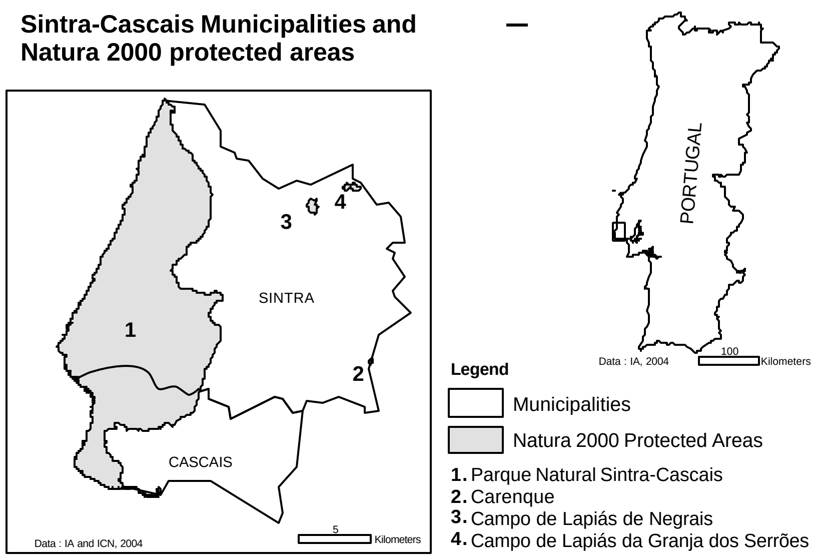

Figure 1: Study area

They are both considered touristic places and rapid urban growing areas. There was a significant increase in the number of inhabitants living in Sintra and Cascais municipalities between 1991 and 2001 (Table 1).

Municipality	1991	2001	Δ	Δ %
Sintra	260951	363749	102798	39.4
Cascais	153294	170683	17389	11.3

Table 1: Population variation between 1991 and 2001 (INE, 2003)

There are four Natura 2000 protected areas in the study area. The largest one, the Natural park of Sintra-Cascais (PNSC) has an area of 145 Km2. This value represents approximately 35% of Sintra and Cascais municipalities' total area. The vegetation of the PNSC is composed by mediterranic and western-mediterranic species from which about 10% are endemic. In this last group, are included species like the Armeria pseudarmeria, Dianthus cintranus cintyranus and the Omphalodes kusynskianae. All of them are considered threatened species at Communitary level. Endangered phauna species include Rhinolophus hipposideros, Rhinolophus euryale, Putorius putorius, Bubo

bubo, Hieraaetus fasciatus and Falco peregrinus among many others (ICN, 2004). The strong contruction pressure in recent years is threatening both touristic and natural characteristics of this area.

DATA AND METHODS

Data

Two Landsat Thematic Mapper 5 (TM) and one Landsat 7 Enhanced Thematic Mapper Plus (ETM+) images were used in this research (Table 2). The 1989 and 2000 images were downloaded from Global Land Cover Facility of the University of Maryland (USA). The 1994 TM image was specifically acquired for the purpose of this research.

Characteristics / Sensor	Landsat TM	Landsat TM	Landsat ETM +
Date	14-03-1989	08-02-1994	24-06-2000
Path – Row	204-033	204-033	204-033
Spatial resolution	30 m	30 m	30 m

Table 2: Satellite images used in this study

These are eight bit (0-255) images that measure the reflected and emitted energy from earth surface features in seven (TM) and eight (ETM+) regions of the electromagnetic spectrum. Spatial resolution is 30 m for all the bands, except for the thermal infra-red (channel 6) which has, respectively 60 m for the ETM+ and 120 m for the TM. The ETM+ image also has a panchromatic band (channel 8) with a spatial resolution of 15 m. The images are cloud-free (except image of year 1994 that has some clouds but of insignificant importance over the study area) and centered at 38º54'18''N, 8º57'46''W.

Methods

A 1:25,000 scale vectorial layer with administrative boundaries was used to create a subset of the images corresponding to the extent of Sintra and Cascais municipalities. The 1989 and 2000 images were previously geometrically and radiometrically corrected by USGS Earth Resource Observation Systems Data Center (EROS) to a quality level of 1G. The same quality level was available for the 1994 image by the European Space Agency. Both 1989 and 2000 images were already orthorectified to a UTM (Universal Transverse Mercator) projection using WGS (World Geodetic System) 84 datum. The 1994 image was co-registered to the 2000 image with a root mean square less than half a pixel (0.49). Both 1989 and 2000 images had a 28.5m pixel resolution. The image of 1994 was resampled to match this resolution using nearest-neighbour algorithm. This research is based on the detection of changes of surface reflectances of objects. This reason justifies the use of a relative radiometric correction with image regression (Jensen, 1996) over 1989 and 1994 images. Brightness values of pixels of all the bands of 1989 and 1994 images were calibrated with image of year 2000 to create a linear regression equation. This procedure minimized effects caused by using time-series of satelite data collected in different dates and with different sun angles (Jensen, 1996).

As stated by Anderson et al. (1976): *There is no one ideal classification of land use and land cover, and it is unlikely that one could ever be developed*. Classifications schemes are abstractions or generalizations of reality. We are interested in analysing the global trend of LUCC for the Sintra and Cascais municipalities. For this reason, the adopted land use/cover classification scheme included three generalized classes (Table 3).

Class	Description
Woodland	Coniferous, deciduous, and mixed forests.
Grassland	Grasses, scrubland, pastures, cropland, agriculture land, golf courses and other herbacueous vegetation
Impervious	Areas with absence of vegetation cover. Houses, roads, dispersed wharehouses, commercial and industrial buildings, airports, beaches and bare soil.

Table 3: Classification scheme used in this research

Image segmentation (Baatz et al. 2004) using Definiens - eCognition 3.0 was employed to select training samples. About 90 training samples were selected for each year. These training samples were as pure as possible and their location was maintained, when possible, over the three images. All bands, equally weighted, were used in image segmentation except the thermal and the panchromatic bands (ETM+) due to their different resolutions. Images were classified using the maximum-likelihood algorithm implemented in Idrisi Kilimanjaro software. Accuracy of classified maps was evaluated using 150 sample points sistematically distributed. These points were converted into cells with the same resolution of the satellite images (28.5m) and classified as woodland, grassland and impervious. The selected pixels had to be pure instead of mixed pixels to ensure that the correct class was identified for each pixel (Gong & Howarth, 1990). These pixels were chosen using large-scale aerial photos and 1:25,000 scale land use/cover maps. Whenever it was not a pure pixel, the closest pure pixel was selected. Confusion matrices were used to compare ground information determined by the inspection of large-scale images and 1:25,000 scale land use/cover maps with the classification results.

A crosstabulation technique was used to quantify changes in the land use/cover classes between 1989 and 2000. The statistical dependence was tested as in any contigency table (Murteira, 1990) displaying the estimated values against the measured ones. In our study, this test wasperformed to infere from the association or independence between the land use/cover classes in different years. The random variable c^2, with the chi-square distribution was defined by (1):

$$c^2 = \sum_i \sum_j ((N_{ij} - M_{ij})^2 / M_{ij}) \quad \textbf{(1)}$$

Where N will be the contigency matrix of measuredland use change, and M a contigency matrix with the estimated values of change (Murteira, 1990). c^2 measures the difference between the observed values of land use change and the estimated ones. This variable assintoptically follows the chi-square distribution for 4 degrees of freedom (m-1)*(n-1), therefore, our critical value is 0.710721 for a confidence level of 0.95.

Spatial metrics are algorithms used for quantifying spatial characteristics of patches, classes of patches, or entire landscape mosaics (McGarigal et al., 2004). They were developed in the late 1980s and include measures from information and fractal theory (Herold et al., 2003). Selected spatial metrics of classified scenes used in this study have already been used in previous researches (Parker et al., 2001; Herold et al., 2003; Cabral et al., 2004) and were calculated using FRAGSTATS public domain software (McGarigal et al., 2004). The term patch defines scale-independent homogeneous regions in a landscape (e.g., grassland, forest, urban, etc.). A short description of spatial metrics used in this study is presented (Table 4).

CLASS METRIC	DESCRIPTION
CA – Total class area	Measures the total area of a patch type in the landscape
NP - Number of patches	Equals the number of patches. They provide a simple measure of the extent of subdivision or fragmentation of the patch type
LPI – Largest Patch Index	Equals the percentage of the landscape comprised by the largest patch.
AREA – MN	MN (Mean) equals the sum, across all patches of the corresponding patch type area, divided by the number of patches of the same type.
FRAC – MN	Equals the sum, across all patches of the corresponding patch type fractal dimension, divided by the number of patches of the same type.
ENN - MN	Equals the distance (m) to the nearest neighboring patch of the same type, based on shortest edge-to-edge distance

Table 4: Spatial metrics used (McGarigal et al, 2002)

Stochastic processes generate sequences of random variables {Xn, n $\in$ T} by probabilistic laws (Kijima, 1997). In this article, index n represents time. The process is considered discrete in time and T = {0, 5, 10 …} years approximately, which is a reasonable time unit for studying land use change phenomenon. Stochastic models have been used in previous LUCC research (Flamm & Turner, 1994; Berry et al., 1996). A stochastic land use/cover map based on the transitional conditional probabilities between years 1989 and 1994 was employed to predict year 2006 land cover/use map. The model used to create this map evaluates these conditional probabilities and assumes that each class can exist at each pixel against a random distribution of probabilities. The sotchastic model used in this research is implemented in Clarke Labs - IDRISI Kilimanjaro through STCHOICE. Robustness of the model was evaluated comparing estimated land cover of year 2000 against classified year 2000 image.

RESULTS

Three land use/cover maps were produced, respectively, for years 1989, 1994 and 2000 using the maximum-likelihood algorithm (Figure 2-4).

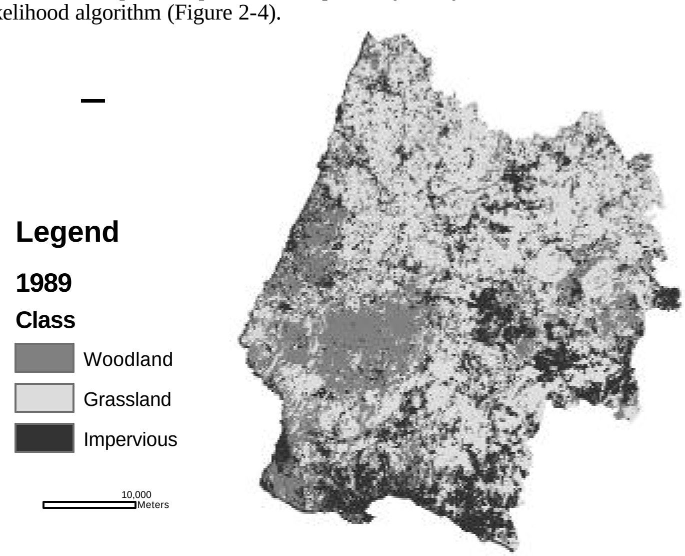

Figure 2: Land use/cover map for year 1989

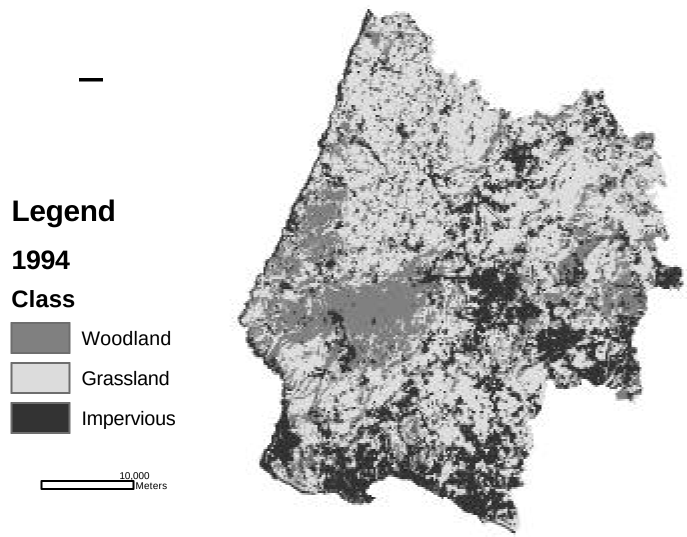

Figure 3: Land use/cover map for year 1994

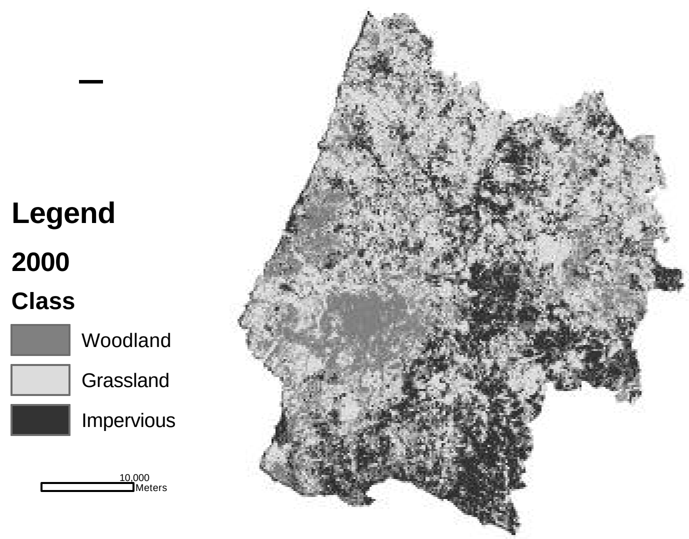

Figure 4: Land use/cover map for year 2000

The 2000 ETM+ image accuracy was assessed using large-scale orthophotos and land cover maps. There was no ground truth data available for this study to assess 1994 image accuracy. However, the classification methodology used for image of year 1989 was replicated the 1994 image. Both images were collected in the rainy season at the same time of the day using the same satellite sensor. For these reasons, we assumed that the overall accuracy of 1994 classification should be identical to the 1989 classification. Overall accuracies obtained for 1989 and 2000 images were, respectively, 88.8%

and 90.7%. The Kappa indices for years 1989 and 2000 were, respectively, 85.3% and 87.1%. These values are considered above the minimum value (85%) stipulated for interpretation accuracy in the identification of land use and land cover categories from remotely sensed data (Anderson et al., 1976). The land use/cover of Sintra-Cascais municiplities has changed significantly (34.2%) between 1989 and 2000 (Table 5). The impervious class was the most dynamic one (22.8%) followed by the woodland (-19.9%) and grassland (-4.4%) classes.

		1989					
		Woodland	Grassland	Impervious	Total	Δ Km2	Δ %
	Woodland	45.8	13.1	4.6	63.5	-15.8	-19.9
2000	Grassland	27.4	153.1	33.6	214.1	-9.9	-4.4
	Impervious	6.0	57.8	74.7	138.6	25.7	22.8
	Total	79.3	224.0	112.9	416.1	142.4	34.2

Table 5: Crosstabulation results between classified images of 1989 (columns) and 2000 (rows)

The largest patch index (LPI) of the impervious class increased 181.04 % (Table 6). This index quantifies the percentage of total landscape area comprised by the largest patch. This means that the largest impervious pacth had a greater dominance as it represents a greater percentage of total area.

Class	Class metrics	1989	2000	Δ % 89-00
Woodland	CA	7926.59	6349.52	-19.90
	NP	3801.00	4804.00	26.39
	LPI	4.05	3.28	-19.07
	AREA_MN	2.09	1.32	-36.62
	FRAC_MN	1.05	1.04	-0.52
	ENN_MN	94.75	93.80	-1.00
Grassland	CA	22401.37	21408.47	-4.43
	NP	3039.00	4344.00	42.94
	LPI	26.32	24.69	-6.18
	AREA_MN	7.37	4.93	-33.14
	FRAC_MN	1.04	1.04	0.12
	ENN_MN	75.28	67.89	-9.81
Impervious	CA	11287.11	13857.07	22.77
	NP	5485.00	5201.00	-5.18
	LPI	4.16	11.68	181.04
	AREA_MN	2.06	2.66	29.47
	FRAC_MN	1.05	1.05	0.06
	ENN_MN	81.26	77.42	-4.73

Table 6: Spatial metrics for years 1989 and 2000

Both LPI values of woodland (-19.07%) and grassland (-6.18%) classes have decreased during this time period. The area weighted mean patch fractal dimension (FRAC_AM) of all classes remained constant between the two dates under analysis. A fractal dimension greater than 1 indicates an increase in shape complexity (e.g. approaches 1 for shapes with simple perimeters as squares and approaches 2 for shapes with highly convoluted perimeters) (McGarigal et al., 2004). The number of patches (NP) for the impervious class decreased-5.18%. This means that the impervious class growth is being made by contagion of neighbouring patches. This deduction is also supported by the increase of the LPI and by the fractal dimension that was not affected by the increase of this class, remaining unchanged from 1989 to 2000. The mean euclidean nearest neighbour (ENN_MN) represents the distance (m) to the nearest neighboring patch of the same type, based on shortest edge-to-edge distance. The decrease of 9.81% and 4.73% for respectively, grassland and impervious class means that the patches of these classes are getting closer to each other between 1989 and 2000. This value remained almost identical for the woodland class.

The stability of conditional transition probabilities between years 1989 to 1994 and years 1994 to 2000 was evaluated by estimating the land use/cover of year 2000 using the conditional transition probabilities of year 1989 to year 1994. The contingency table obtained was compared against the estimated contingency table assuming the total independence between year 2000 and estimated year 2000 (Murteira, 1990). A chi-square test was used to evaluate the statistical significance of the association these two tables. A value of 0.438710699 was obtained for the chi-square statistic. This value is below the critical value of 0.710721 for a chi-square distribution with 0.95 confidence level. Knowing this, we can state with reasonable assurance that the transition process of years 1989 to 1994 is similar to the transition process of years 1994 to 2000. This result allows us to estimate 2006 with a 0.95 confidence level.

A stochastic model was used to estimate land use/cover change dynamics for year 2006 (Figure 5). This model assumes that transitional conditional probabilities between 1994 and 2000 are equivalent to the transitional conditional probabilities between 2000 and 2006 and it spatialisation is completely random.

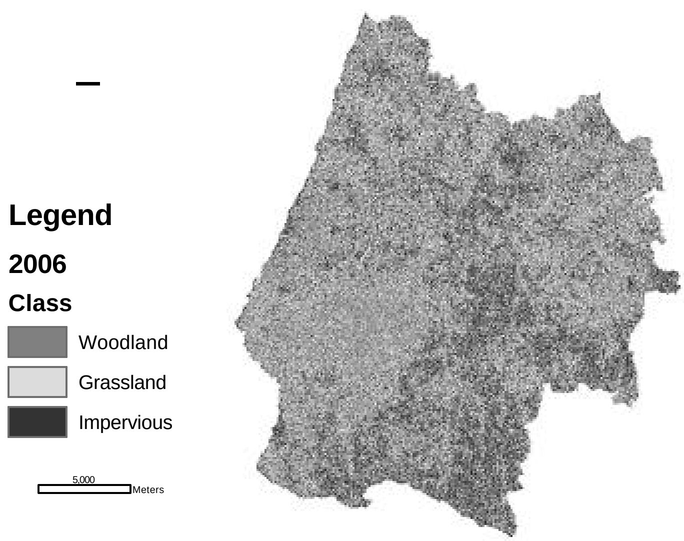

Figure 5: Predicted land use/cover map for year 2006

The result shows very sparse pixels. This is due to the absolutely random way that STCHOICE distributes the occurrences for each category. Altough this model is accurate in estimating quantities;

it does not take in consideration spatial contiguity. Classes attributed to pixels are spatially independent. According to this model in year 2006, the woodland class will decrease 345 ha (-5.44%), the grassland class will decrease 3680 ha (-1.40) and the impervious class will increase 644 ha (4.65%).

The evaluation of the importance of the park was made using a chi-square test between what happened to LUCC between years 1989 and 2000 inside and outside the park. The chi-square value obtained for these two contingency tables was 0.831 (above our critical value of 0.711). This means that what happens inside and outside the park is different for a 0.95 confidence level. Therefore the park is a significative factor in LUCC dynamics. Table 7 shows the differences in the dynamics of classes inside and outside the PNSC.

Class	Inside PNSC	Outside PNSC
Woodland	-15.1	-28.5
Grassland	16.8	-12.9
Impervious	-10.3	34.5

Table 7: Landuse/cover changes (%) between 1989 and 2000 inside and outside PNSC

The Kappa statistic was used to assess similarity between estimated map of year 2000 for the Sintra-Cascais municipalities, inside and outside the park. Values obtained were, respectively, 0.585, 0.7104 and 0.6712. Although estimated map for year 2000 has also very sparse classified pixels, the Kappa values obtained shows us that exists a reasonable similarity between the estimated map and real classified map of year 2000. This value would be about 33.3% if the classification of the model was entirely due to chance. This statistic has a higher value inside the PNSC than outside the PNSC. This fact allows us to conclude that the process of LUCC dynamics is more random outside the PNSC than inside the PNSC. The PNSC plays an important role in LUCC stability in the Sintra-Cascais area.

CONCLUSIONS

This paper aims to show a GISc approach in analysing and modeling LUCC dynamics of Sintra-Cascais municipalities between 1989 and 2006. This approach involved remote sensing, geographical informations systems, stochastic modeling and spatial metrics techniques. Results show a continued tendency for the increase of the impervious class (22.8%) between 1989 and 2000 in this area. During this period, both woodland (-19.9%) grassland (-4.4%) classes decreased. The stochastic model used to estimate LUCC for year 2006 predicts that the impervious class will continue to increase (4.65%). Both woodland and grassland classes will decrease for year 2006, respectively, -5.44% and -1.40%. These facts allow us to conclude that the natural equilibrium of the Sintra-Cascais municipalities is threatened by the strong construction pressure of the last thirty years. This analysis was extended to the PNSC to evaluate its importance in the LUCC dynamics of this region. Results show that the PNSC plays an important role in mantaining LUCC more stable in the areas inside the park. The use of spatial metrics provided quantitative information about the type of spatial dynamics that Sintra-Cascais municipalities are facing. Future reasearch may include the use of cellular automata models to better estimate spatial location of classes. During this research strong evidences were found that the process of LUCC in the Sintra-Cascais could be a Markov chain. Future work will aim to proove this hypothesis.

REFERENCES

Anderson, J. R., E. E. Hardy, et al. (1976). A Land Use And Land Cover Classification System For Use With Remote Sensor Data, U.S. Geological Survey.

Benenson, I. and P. Torrens (2004). Geosimulation: Automata-based modeling of urban phenomena, Wiley.

Berry, M., B. Hazen, et al. (1996). The Land-Use Change Analysis System (LUCAS) for evaluating landscape management decisions. IEEE - Computational Science and Engineering **3**(1): 24-35.

Cabral, P., M. Painho, et al. (2004). Spatio-temporal analysis of urban growth pattern in Lisbon Metropolitan Area. 24th Urban Data Management Symposium, Chioggia, Italy.

Cheng, J. (2003). Modeling Spatial & Temporal Urban Growth. Faculty of Geographical Sciences. Utrecht, Utrecht University**:** 203.

Flamm, R. and M. Turner (1994). Alternative model formulations for stochastic simulation of landscape change. Landscape Ecology **9**(1): 37-44.

Gluch, R. (2002). Urban growth detection using texture analysis on merged Landsat TM and SPOT-P data. Photogrammetric Engineering & Remote Sensing **68**(12): 1283-1288.

Gong, P. and J. Howarth (1990). The use of structural information for improving land-cover classification accuracies at the rural-urban fringe. Photogrammetric Engineering & Remote Sensing **56**(1): 67-73.

Herold, M., N. Goldstein, et al. (2003). The spatio-temporal form of urban growth: measurement, analysis and modeling. Remote Sensing of Environment **85**: 95-105.

ICN (2004). Parque Natural Sintra-Cascais, Instituto da Conservação da Natureza. **2004**.

INE (2003). Instituto Nacional de Estatística.

Jensen, J. R. (1996). Introductory Digital Image Processing: A remote sensing perspective, Prentice Hall.

Kaufmann, R. and K. Seto (2001). Change detection, accuracy, and bias in a sequential analysis of Landsat imagery in the Pearl River Delta, China: econometric techniques. Agriculture, Ecosystems and Environment **86**: 286-302.

Kijima, M. (1997). Markov processes for stochastic modeling. London, Chapman & Hall.

Masek, J., F. Lindsay, et al. (2000). Dynamics of urban growth in the Washington DC metropolitan area, 1973-1996, from Landsat observations. International Journal of Remote Sensing **21**(18): 3473-3486.

McGarigal, K., B. Marks, et al. (2004). FRAGSTATS.

Murteira, B. (1990). Probabilidades e estatística, McGraw-Hill de Portugal.

NIC (2000). Global Trends 2015: A dialogue about the future with nongovernment experts, National Inteligence Council. **2004**.

Parker, D., T. Evans, et al. (2001). Measuring emergent properties of agent-based land use/land cover models using spatial metrics. Seventh annual conference of the International Society for Computational Economics.

Read, J. and N. Lam (2002). Spatial methods for characterising land cover and detecting land-cover changes for the tropics. International Journal of Remote Sensing **23**(12): 2457-2474.

Seto, K., C. Woodcock, et al. (2002). Monitoring land-use change in the Pearl River Delta using Landsat TM. International Journal of Remote Sensing **23**(10): 1985-2004.

UNECE (2003). Trends in Europe and North America - The Statistical Yearbook of the Economic Commission for Europe 2003, United Nations Economic Committee for Europe. **2004**.

Weng, Q. (2002). Land use change analysis in the Zhujiang Delta of China using satellite remote sensing, GIS and stochastic modelling. Journal of Environmental Management **64**: 273-284.