Revista Brasileira de Cartografia (2017), Nº 69/8, Edição Especial "Geovisualização, mídias sociais e participação cidadã: apoio à representação, análise e gestão da paisagem": 1610-1627.
Sociedade Brasileira de Cartografia, Geodésia, Fotogrametria e Sensoriamento Remoto
ISSN: 1808-0936

STRUGGLE FOR SPACE IN COASTAL CITIES: A CELLULAR AUTOMATA MODEL FOR SIMULATING URBAN STRATIFICATION DYNAMICS

Disputa pelo Espaço em Cidades Costeiras: Um Modelo Baseado em Autômatos Celulares para a Simulação de Dinâmicas de Estratificação Urbana

Flavia da Fonseca Feitosa[1], Raian Vargas Maretto[2], Antônio Miguel Vieira Monteiro[2]

[1] **Universidade Federal do ABC (UFABC)**
Centro de Engenharia, Modelagem e ciências Sociais Aplicadas – CECS
Alameda da Universidade s/nº - Bloco Delta, sala 331, B. Anchieta - São Bernardo do Campo, SP, cep 09606-045
flafeitosa@gmail.com

[2] **Instituto Nacional de Pesquisas Espaciais – INPE**
Divisão de Processamento de Imagens
Av. dos Astronautas, 1758, São José dos Campos, SP, cep 12227-010
raian@dpi.inpe.br, miguel@dpi.inpe.br

Received on April 4, 2017/ Accepted on August 24, 2017
Recebido em 4 de Abril, 2017/ Aceito em 24 de Agosto, 2017

ABSTRACT

The social production of Brazilian cities has historically accentuated the differences between urban locations and promoted urban stratification. This reality is particularly pronounced in coastal areas, where the locational advantages of areas near the sea create an increasingly competitive and exclusionary land market. Through methods of computational modeling and simulation, this work analyzes the production of the urban space of Caraguatatuba, a Brazilian city located on the northern coast of São Paulo State. We developed a cellular automata model called URBIS-Caragua to explore the dynamics of urban stratification bounded by urban legislation and the spatial distribution of residential location opportunities. In this process, particular attention was directed towards the empirical parameterization of the model, which relied on exploratory data analysis and statistical models for the periods 1991-2000 and 2000-2010. Through the parameterization and development of simulation experiments about past dynamics, it was possible to obtain new insights about how trends of urban expansion and stratification changed over the recent years, and the possible relation of these trends to urban legislation, locational factors and sociodemographic pressures. The experiments about future dynamics (2010-2025) indicated that the Master Plan of the city, approved in 2011, is likely to reinforce trends that have historically established an unequal distribution of opportunities for residents that belong to different social groups.

Keywords: Cellular Automata Modeling, Urban simulation, Coastal Cities, Urban Stratification.

RESUMO

A produção social das cidades brasileiras tem acentuado historicamente as diferenças entre localizações urbanas e promovido estratificação urbana. Tal realidade é particularmente observada em áreas costeiras, onde as vantagens locacionais relacionadas à proximidade ao mar acentuam a formação de um mercado de terras competitivo e excludente. Através de métodos de modelagem e simulação computacional, este trabalho analisa a produção do espaço urbano de Caraguatatuba, uma cidade brasileira localizada no litoral norte do estado de São Paulo. Um modelo baseado em autômatos celulares, denominado URBIS-Caragua, foi desenvolvido para explorar dinâmicas de estratificação urbana em relação à legislação urbana e distribuição espacial das oportunidades de localização residencial. Neste processo, particular atenção foi destinada à parametrização empírica do modelo, que se baseia em análises exploratórias de dados e modelos estatísticos construídos para os períodos de 1991-2000 e 2000-2010. Através da parametrização e desenvolvimento de experimentos de simulação sobre dinâmicas passadas, foi possível ampliar a compreensão sobre como tendências de expansão urbana e estratificação modificaram-se ao longo dos últimos anos, bem como sobre relações entre estas tendências e a legislação urbana, fatores locacionais e pressões sociodemográficas. Experimentos sobre dinâmicas futuras (2010-2025) indicaram que o Plano Diretor da cidade, aprovado em 2011, tende a reforçar tendências que historicamente estabeleceram uma distribuição desigual de oportunidades a residentes pertencentes a distintos grupos sociais.

Palavras chaves: Modelagem baseada em Autômatos Celulares, Simulação Urbana, Cidades Costeiras, Estratificação Urbana.

1. INTRODUCTION

With their increasingly important role as human habitat, cities stand out as places that are characterized by a continuous struggle for urban locations. A well-positioned urban location represents a valuable resource to its occupants because it facilitates access to the numerous opportunities and possibilities offered by the city (SANTOS, 2012). In a society that is predominantly guided by the logic of the market, the best urban locations become the targets of conflicts and disputes involving heterogeneous actors, including families that belong to different social groups, entrepreneurs, landowners and various sectors of the government (ABRAMO, 2012).

In Brazil, given the deep inequality in the economic and political power of the agents involved in these disputes, the social production of urban space has historically promoted a deeply stratified spatial structure and accentuated the differences between urban locations (CARLOS *et al.*, 2011). This situation is particularly pronounced in coastal urban areas, where the locational advantages offered by the land-sea interface accentuate the pace of Brazilian urbanization trends and create an increasingly competitive land market. In these regions, the exploitation of tourism activities and the progressive peripherization of low-income families emerge as important determinants of land-use and occupation patterns (MORAES, 2007). On the one hand, tourism and summering boost competitiveness and speculation in real estate by increasing the demand for constructing tourist facilities and second homes. On the other hand, excluded populations without access to the formal land market are forced to live in segregated areas that are often unsuitable for occupation, lack infrastructure and urban services, and concentrate several disadvantages that contribute to increasing and perpetuating poverty and vulnerability (MORAES, 2007).

Starting from the premise that the location of a family in the urban space is directly related to its capability to access the opportunities offered in the city, this paper analyzes the dynamics of urban stratification in coastal areas and emphasizes the residential distribution of social groups. A stratified urban space, as presented by Castells (1979), is characterized by an unequal distribution of opportunities among social groups. It is a spatial structure in which certain social groups occupy urban locations that favor the accumulation of advantages or disadvantages. In this work, we analyze urban stratification in the light of its relation to urban legislation, locational factors and sociodemographic pressures.

For this purpose, we selected a study area that illustrates the challenges faced by Brazilian coastal cities: the municipality of Caraguatatuba. Located on the northern coast of the State of São Paulo, Caraguatatuba has an estimated population of 115,071 inhabitants (IBGE, 2016), a number that increases up to threefold with the arrival of tourists in the summer. In addition to the pressures caused by its strong vocation for tourism, Caraguatatuba is located in a region where numerous investments to support the treatment and distribution of gas and oil have been planned and implemented, thus promoting transformations that could be related to an increased rate of population growth and urban expansion.

To explore the impact of these changes on the urban structure, more specifically on the spatial configuration of residential areas occupied by families belonging to different social groups, this work relies on methods of computational modeling and simulation. Simulation models have the advantage of overcoming the static nature of maps and measures, incorporating a dynamic dimension that offers new perspectives for urban studies (BATTY, 2013). Moreover, they are promising tools for the exchange of experience and knowledge between researchers from different disciplinary domains, decision makers and other stakeholders. Nevertheless, for the models to assume this mediation role, it is necessary to prioritize and value the *process of building and using models* instead of the traditional view of the model as a *product*. In this case, the goal of the model is not to provide precise quantitative answers but to serve as an instrument to share different views, structure discussions, test and generate hypotheses, and raise new questions (FEITOSA & MONTEIRO, 2012).

In a functional paradigm of modeling, simulation models can be considered laboratories that are able to assist in understanding a problem in two ways: (a) in the process of building this laboratory, which is continuous and demands many exploratory analyses; and (b) in the interpretation, verification and evaluation of simulation experiments. Giving a particular emphasis on item (a), this paper aims to present the process of building a model based on cellular automata to simulate dynamics of urban stratification in Caraguatatuba. This model, named as *URBIS-Caragua*, focuses on exploring possible scenarios and trends of the spatial configuration of residential areas that are occupied by families under different socioeconomic conditions.

2. BUILDING THE MODEL URBIS-CARAGUA

According to Batty (2005), cellular automata (CA) could be understood as a cell-based approach for modelling dynamic processes at the micro level. Essentially, CA are composed by a grid of cells characterized by their state. Changes in the state of cells are driven by transition rules, which take into account what exists or is happening in the neighborhood of the cell (BATTY *et al.*, 1997). During the last decades, CA models have been widely used in studies on urban dynamics, focusing mainly on the simulation of urban expansion and land use change processes, as pointed out by Santé *et al.* (2010) and Wahyudi and Liu (2016).

In addition to simulating the expansion of residential areas, the model URBIS-Caragua was developed to support analyses on how recent transformations in the northern coast of São Paulo are driving the spatial distribution of families that belong to different social groups in Caraguatatuba. An overview of the process of building and using this simulation model is shown in Figure 1. The main phases that comprise this process are discussed in the following sections.

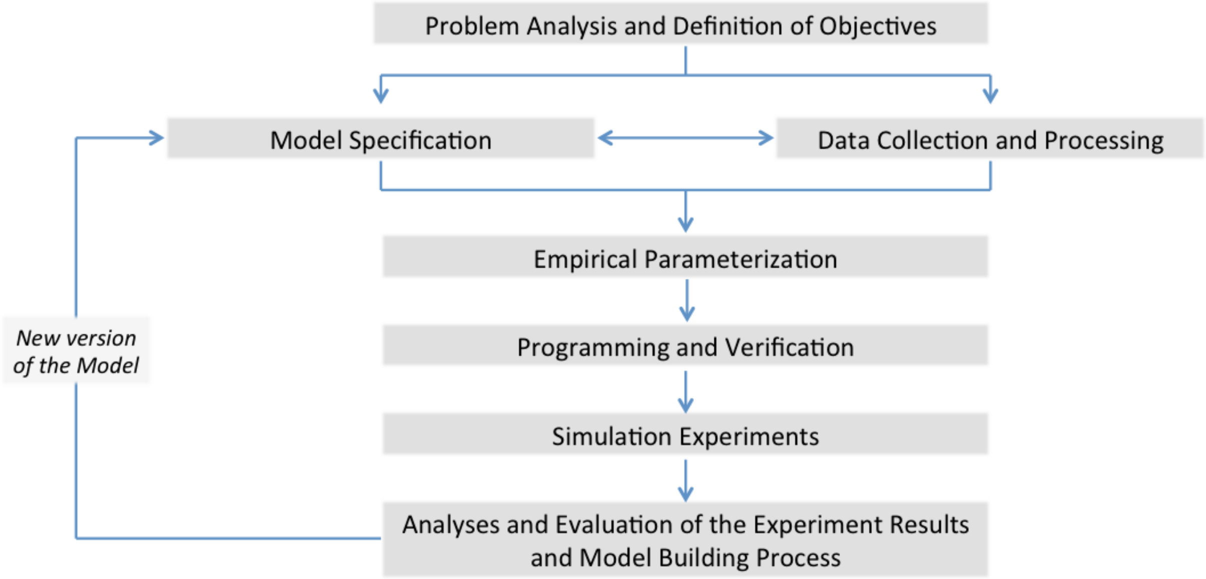

Fig. 1 - Methodological steps for building and using the model *URBIS-Caragua*.

2.1 Problem Analysis and Definition of Objectives

The typical problems arising from the production of urban space in Brazil, which has resulted in chaotic and socially unfair urban settings, tend to become particularly pronounced in coastal areas, where the main vectors and fluxes of all type converge and compose a complex mosaic of occupation patterns, land and natural resources use, and economic exploitation. The situation is aggravated when the site is characterized by fragile ecosystems, as is the case of the Northern Coast of São Paulo (Figure 2), which covers the municipalities of Ubatuba, Caraguatatuba, São Sebastião and Ilhabela. This area has significant ecological diversity, with a system of mountain ranges and escarpments covered by the Atlantic Rainforest.

The scenic resources of the area make it very appealing for projects aimed at tourism development, especially second homes, which represent approximately 42% of all dwellings in the area (IBGE, 2010). The region also hosts equipment and infrastructure that reinforce its strategic importance for the state and the country, such as the port of São Sebastião and the maritime terminal Almirante Barros (TEBAR), which specializes in the loading and unloading of petroleum and by-products and is owned and operated by the energy corporation Petrobras.

Growth in the population and tourism activities during the period of 1991-2010 was accompanied by an increase in the number of dwellings to approximately 200% (from 91,128 to 182,635, many of which meant as second homes). With this increasing pressure, the built-up areas, which occupy a narrow strip of coastal plain between the Serra do Mar and the coastline, are progressively expanding towards areas unsuitable for occupation, which are subject to floods and landslides.

The pace of these changes has remained intense with the recent construction and installation of large infrastructure projects, most of them related to supply chain activities for the oil and gas industry. Among the four municipalities of this region, our focus is Caraguatatuba (Figure 2), which represents not only the largest city in terms of population but also the city with the greatest availability of land for urban expansion and the greatest potential for population growth in the face of new investments. This role was reinforced by the process of approving the Master Plan of Caraguatatuba (PREFEITURA MUNICIPAL DA ESTÂNCIA BALNEÁRIA DE CARAGUATATUBA, 2011), which was sanctioned in November 2011 despite the criticism from many sectors of society. The plan explicitly addresses issues concerning the growing tourist activity in the region, the new industrial and logistical uses related to the oil industry, and population growth resulting from the intensification of these economic activities in the region.

In view of these transformations, this study aims to explore the following questions through the process of building a simulation model for Caraguatatuba (*URBIS-Caragua*):

a) What trends were observed in terms of urban expansion and the spatial distribution of social groups in recent decades, and what possible trends could we observe in the coming years?

b) How could the Master Plan of the city

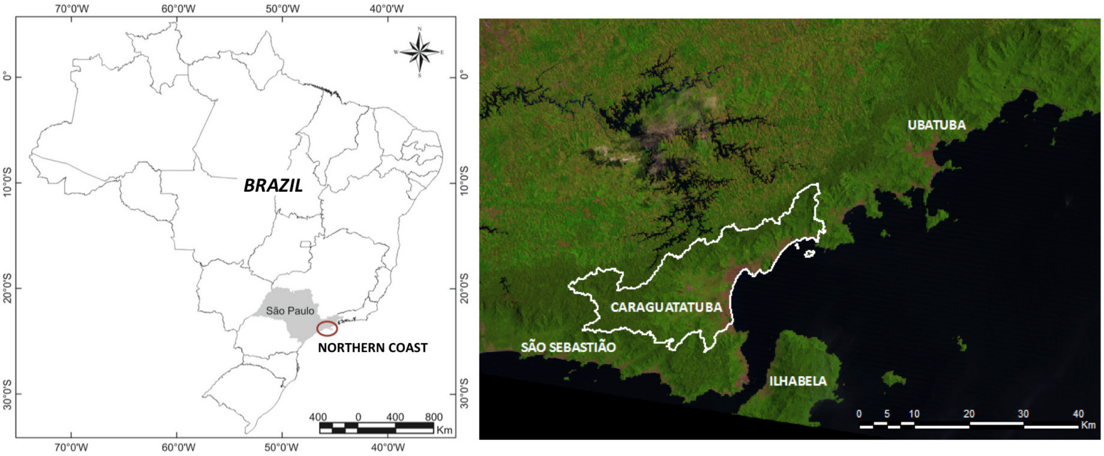

Fig. 2 - Location of the Northern Coast of São Paulo and the Municipality of Caraguatatuba.

influence these dynamics in the coming years?

c) How could these trends vary given different demographic and socioeconomic scenarios?

3 MODEL SPECIFICATION

3.1 Representation of the Study Area

This work builds a cellular automata model in which the area of the city of Caraguatatuba is represented by a grid composed of cells of 100 x 100 m, and each cell is characterized by a vector of variables. These variables, which comprise the state of the cell, can be endogenous or exogenous. The endogenous variables are those whose values are simulated by the model, whereas the exogenous variables are those determined outside the model and whose variations are significant for estimating the endogenous variables.

The main endogenous variable of the *URBIS-Caragua* model was obtained from a classification that seeks to categorize the social conditions of families living in Caraguatatuba (Figure 3). The cellular space that represents the city was divided between *built-up* and *non-built-up* areas. The built areas were split between those *without residential use* and those *with residential use*. The areas *with residential use* are of greatest interest in this study because they accommodate the families living in the city. In these areas, the residential use does not need to be exclusive or prevailing.

The areas with residential use were classified according to the social conditions of the families residing in them. To determine

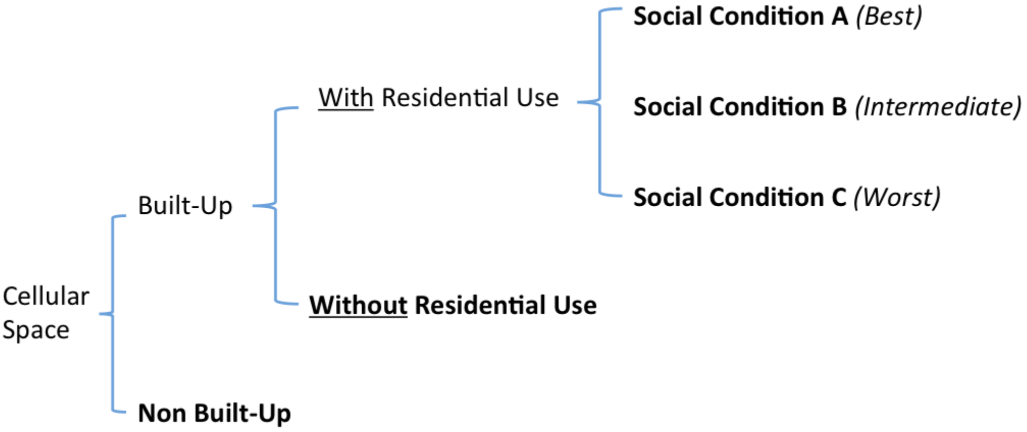

Fig. 3- Classification Tree for the Cellular Space of the *URBIS-Caragua* model.

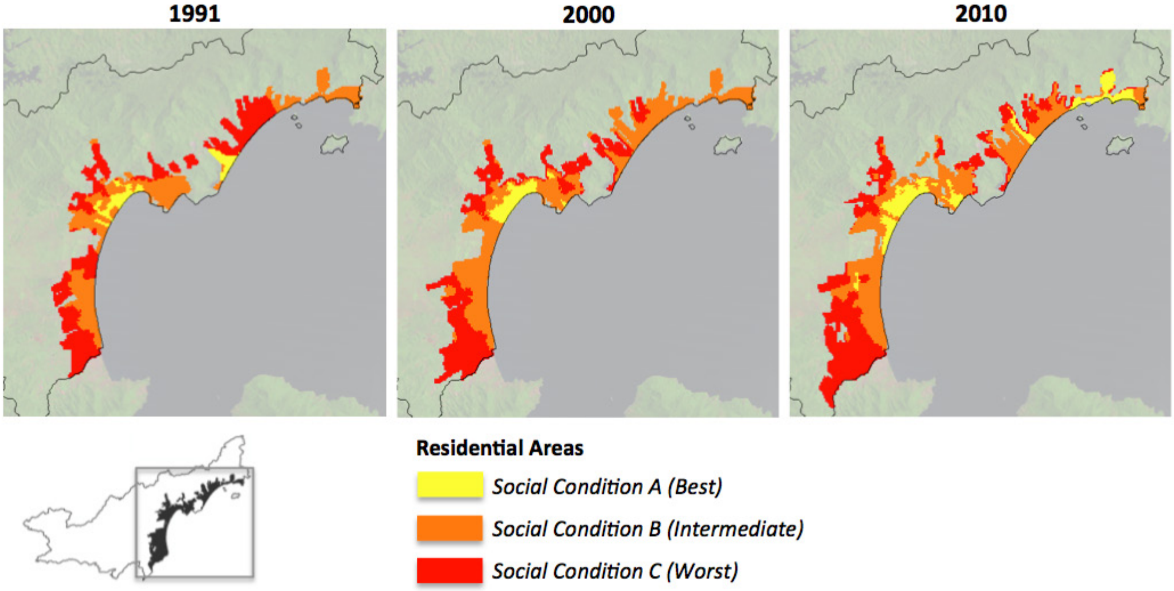

Fig. 4 - Residential areas classified according to the social condition of their residents - Years 1991, 2000 and 2010.

these classes, cluster analysis using the *k-means* method (KAUFMAN & ROUSSEEUW, 2008) was performed on a set of indicators proposed by Anazawa *et al.* (2013), which includes census data on income, education and gender of householders; youth literacy; the dependency ratio; and the rate of owned dwellings. Three residential classes were established: A (the best social condition), B (intermediate) and C (the worst). The results obtained by this process were integrated into a geographic database and redistributed on a cellular space (Figure 4).

3.2 Demand and Allocation Phase

Considering the classification presented in Figure 4, the aim of the *URBIS-Caragua* model is to simulate transitions of the spatial patterns of the residential location of families living under different social conditions. These processes are represented in this work using a top-down approach, in which the demand for residential areas presented by each group of families is allocated according to an estimated potential for each cell. The presented model is built on LuccME, an open-source framework for spatially explicit Land Use and Cover Change (LUCC) modeling, which is built on top of TerraME, a general programming environment for spatial dynamical modeling (AGUIAR *et al.*, 2012; CARNEIRO *et al.*, 2013; ANDRADE *et al.*, 2015).

Basically, the model consists of two phases: *demand* and *allocation*. The *demand phase* ("how much?") consists of determining the total demand for residential areas to be occupied by families belonging to each social group. In other words, this phase computes the total number of cells that should be allocated for each class identified in Figure 3.

The *allocation phase* ("where?") is responsible for the spatial allocation of the demands identified in the first phase. For that, it is necessary to estimate the cell suitability (or potential) for each considered class. The empirical parameterization of this step relies on binary logistic regression models to compute the local probability of having a cell belonging to a particular class. Binary logistic regression is a type of regression analysis where the dependent variable is a dummy variable (e.g., 1 = class "non-built-up area" and 0 = other class). The statistical model for logistic regression is as follows (MOORE & MCCABE, 2003):

$$\log\left(\frac{p(u)}{1-p(u)}\right) = \beta_0 + \beta_1 X_1 + \beta_n X_n \quad , \quad (1)$$

where *p(u)* is the probability that the cell belongs to class u, X_i are explanatory variable, and β_i are the logistic model parameters. The local probability of having a cell belonging to the class u, $p(u)$, represents the cell suitability for the class u and can be estimated as follows:

$$p(u) = 1 \big/ 1 + \exp(-\beta_0 - \beta_1 X_1 - \beta_n X_n) \quad (2)$$

After computing the cells' suitabilities for the considered classes, an allocation process based on competition among classes in the same cell is initiated. This process is iteratively adjusted to reach the demand estimated for each class, as described in Verburg *et al.* (2002).

4. PARAMETERIZATION AND SIMULATION EXPERIMENTS

4.1 Periods 1991-2000 and 2000-2010

While empirically based simulation models often rely directly on data about the past to parameterize experiments about the future, this work assumes that the factors that influence the dynamics of urbanization, as well as their relevance, are constantly changing. Thus, the parameterization and development of simulation experiments about past dynamics, as observed during the periods 1991-2000 and 2000-2010, subsidize a process of reflection and learning on (a) how trends of urban expansion and the spatial distribution of social groups changed over the recent years and (b) the possible relation of these trends to political, socioeconomic and demographic factors.

4.1.1 Demand Parameterization

The parameterization of the demand phase relied on the quantification of built-up areas, as well as residential areas occupied by different social groups (Table 1). For the most recent period (2000-2010), it was possible to observe a deceleration of urban expansion and, at the same time, the growing participation of poor families in newly built areas, which expanded in the direction opposite to the sea, towards the scarps (Figure 4).

Table 1- Number of cells belonging to the class "non-built-up" and residential classes "social condition A", "social condition B" and "social condition C" - Years 1991, 2000 and 2010.

Classes	**Number of cells (100 X 100 m)**		
	1991	**2000**	**2010**
Non-built-up	45,351 (91.3%)	44,586 (89.7%)	43,775 (88.1%)
Residential Classes:			
Social Condition A (*Best*)	335 (0.7%)	318 (0.6%)	832 (1.6%)
Social Condition B (*Intermediate*)	1,912 (3.9%)	2,738 (5.5%)	2,656 (5.3%)
Social Condition C (Worst)	2,087 (4.2%)	2,043 (4.1%)	2,422 (4.9%)
Total number of cells	*49,685* *(100%)*	*49,685* *(100%)*	*49,685* *(100%)*

4.1.2 Allocation Parameterization

The parameterization of the second phase of the model, *allocation*, seeks a better understanding of the relation between the spatial location of built-up areas (including the social groups that characterize them) and the factors associated with urban legislation and other local characteristics of the territory. To obtain functions that estimate the local probability of having a cell belonging to a particular class, we built binary logistic regression models based on the variables presented in Table 2, which consider the previous state of the cell and its neighborhood, the land slope, the distance from the central zone and the sea, and zoning variables. The zoning variables were obtained from the municipal

Table 2 - Binary Logistic Models: Dependent and Explanatory Variables. Periods 1991-2000 and 2000-2010.

Dependent Variables	
NR(t)[*]	Cell state in *t* = "*non built-up*"
RA(t)	Cell state in *t* = "*residential with social condition A*"
RB(t)	Cell state in *t* = "*residential with social condition B*"
RC(t)	Cell state in *t* = "*residential with social condition C*"
Explanatory Variables	
Previous Cell State (t-n):	
NR(t-n)[*]	Cell state in (*t-n*) = "*non built-up*"
RA(t-n)	Cell state in (*t-n*) = "*residential with social condition A*"
RB(t-n)	Cell state in (*t-n*) = "*residential with social condition B*"
RC(t-n)	Cell state in (*t-n*) = "*residential with social condition C*"
Neighborhood externalities in (t-n):	
NeighNR(t-n)	Proportion of neighboring cells[**] with state (*t-n*) = "*non built-up*"
NeighRA(t-n)	Proportion of neighboring cells with state (*t-n*) = "*residential with social condition A*"
NeighRB(t-n)	Proportion of neighboring cells with state (*t-n*) = "*residential with social condition B*"
NeighRC(t-n)	Proportion of neighboring cells with state (*t-n*) = "*residential with social condition C*"
Zoning:	
Special	Special management zone due to its geological fragility
Central	Central zone of the city, with higher density of fixed population, shops and services.
Subcenters	Secondary centers, where the density of fixed population, shops and services is also elevated.
Touristic	Touristic residential zone, with predominance of settlements for the floating population
FootSlope	Base of slope zone
SlopeZone	Slope zone
Agriculture	Agricultural zone
Conservation	Environmental Preservation Zone
Other factors:	
DistCenter	Distance from the central zone
DistSea	Distance from the sea
LandSlope	Land slope

[*] For the period 1991-2000: *t* = 2000 and *t-n* = 1991; For the period 2000-2010: *t* = 2010 and *t-n* = 2000.

Zoning Law which remained in effect until 2011 (PREFEITURA MUNICIPAL DA ESTÂNCIA BALNEÁRIA DE CARAGUATATUBA, 1992).

Comparing the regression models built for two different periods, 1991-2000 and 2000-2010, enables us to observe how the impact of factors that contribute to the spatial location of families belonging to different social groups changed over the years. Table 3 presents the results of the logistic models built to estimate the cell suitability for the class "non built-up" and the residential classes with "social condition A", "social condition B" and "social condition C". To allow comparison between models built for 1991-2000 and 2000-2010, the same variables were considered in the estimation of the potential for a given class. Therefore, for the model built for each class, we chose to include explanatory variables that were statistically significant at the 95% level in at least one of the two periods.

To estimate the model for the class "non built-up" (dependent variable "*NR(t)*"), we assumed that the potential of a cell to remain in this class in time *t* increases if:

(a) the neighboring cells were also not built in *t-n* (explanatory variable "*NeighNR(t-n)*"),

(b) the cell is located inside an environmentally protected area ("*Conservation*"),

(c) the cell is distant from the positive externalities associated with the city center and coastline ("*DistCenter*" and "*DistSea*"), and

(d) the cell is less suitable for occupation due to high slopes ("*LandSlope*").

The results of the model confirmed these initial expectations. Comparing the two analyzed periods, it is observed that the coefficients estimated for the variables "Conservation", "DistCenter", "DistSea" and "LandSlope" presented a smaller magnitude, or even a lack of statistical significance, in the most recent period (2000-2010). This trend may be related to a decreased availability of land, which makes the occupation of sites that are less suitable or convenient more likely, such as those that are distant from the city center or the coastline, are located in preservation areas and/or have steeper slopes.

The model built to estimate the potential of a cell belonging to the residential class "Social Condition C (RC)" shows that a cell presents a higher probability of being occupied by families in disadvantaged social conditions if it already presented this state in *t-n* (variable "*RC(t-n)*"). Similar results can be observed when analyzing the cell's neighborhood (*VizRA(t-n), VizRB(t-n), VizRC(t-n)*): the probability of belonging to the class RC increased when neighboring cells already belonged to this class in *t-n* and decreased when neighboring residential cells presented better social conditions in *t-n*.

This observation became more evident in the most recent period (2000-2010), which shows a growing importance of factors related to neighborhood externalities. These trends are usually accompanied by an increase in the levels of residential segregation and its negative impacts on poorer families. Reinforcing this evidence, the coefficients estimated for the variables concerning the central zone ("Central" and "DistCenter") and touristic residential areas ("Touristic") reveal that cells representing these zones have a decreased probability of being occupied by the poorest families.

The coefficients estimated for the zoning variable "Special", which are special management zones due to their geological fragility, reveal that in the period 1991-2000, the social groups C and B presented a higher probability of occupying these areas. In the period 2000-2010, however, this variable had a negative coefficient for social group C, likely due to the further consolidation of this area, which is now predominantly occupied by families that belong to social group B. Thus, although the geological peculiarities of the area leading to greater exposure to hazards remain, the response capacity of its residents is possibly improving.

Table 3 - Binary Logistic Models: Estimated parameters to compute the cell suitability for the classes "<u>*Non-built-up*</u>" - *NR(t)*, "<u>*Social Condition C*</u>" - *RC(t)*, "<u>*Social Condition B*</u>" - *RB(t)*, and "<u>*Social Condition A*</u>" - *RA(t)*. Periods 1991-2000 and 2000-2010.

Explanatory Variables	**Estimated Coefficients**							
	"Non-built-up" - NR(t)		***«Social Condition C» - RC(t)***		***«Social Condition B» - RB(t)***		***«Social Condition A» - RA(t)***	
	1991-2000	**2000-2010**	**1991-2000**	**2000-2010**	**1991-2000**	**2000-2010**	**1991-2000**	**2000-2010**
Constant	-6.085***	-4.48***	-1.76***	-3.01***	-2.94***	-2.61***	-2.97***	-5.99***
RA(t-n)	-	-			4.68***	2.63***	4.34***	2.83***
RB(t-n)	-	-	1.36***	3.06***	4.46***	2.5***	2.23***	3.95***
RC(t-n)	-	-	2.79***	2.42***	1.89***	1.8***		
NeighNR(t-n)	7.75***	6.91***						
NeighRA(t-n)	-	-	-3.73***	-0.29***	-6.46***	-6.04*	10.29***	8.2***
NeighRB(t-n)	-	-	-0.60**	-0.87***	-0.14	2.44***	-2.85**	-1.3***
NeighRC(t-n)	-	-	1.76***	0.22***	0.68***	0.4***	7.32***	-3.7***
Special	-	-	2.08***	-0.94***	1.13***	1.66***		
Central	-	-	-0.48*	-16.27	0.05	1.8***	0.33	0.92***
Subcenters	-	-	0.74***	0.61***	0.26	0.71***		
Touristic	-	-	-2.43***	-1.15***	1.87***	0.84***	1.96***	1.51***
FootSlope	-	-	-0.37	0.7***	1.33***	1.42***		
SlopeZone	-	-	1.20***	1.05***	0.58**	-0.24		
Agriculture	-	-	1.11***	1.29***	-16.052	-0.82**		
Conservation	1.61***	0.6***			-1.05***	1.01***		
DistCenter	$2(10^{-5})^{**}$	$-1.2(10^{-4})$	$-4.5(10^{-5})^{***}$	$1.6(10^{-4})^{***}$	$-1.1(10^{-5})$	$-1.7(10^{-4})^{***}$	$-2.5(10^{-3})^{***}$	$1.2\ (10^{-4})^{***}$
DistSea	$3.3\ (10^{-4})^{***}$	$1.5(10^{-4})^{***}$	$-4.9(10^{-4})^{***}$	$-3.3(10^{-4})^{***}$	$-4.7(10^{-4})^{***}$	$-3.3(10^{-4})^{***}$	$-2.1(10^{-3})^{***}$	$-4.3(10^{-4})^{***}$
LandSlope	$6\ (10^{-2})^{***}$	$5.2(10^{-2})^{***}$	$-4(10^{-2})^{***}$	$-1.8(10^{-2})^{***}$	$-1.1(10^{-2})$	$-3.4\ (10^{-2})^{***}$	$-1.2(10^{-1})^{*}$	$-3.1(10^{-2})^{**}$
Chi-square test χ^2	15423***	13363***	6016***	6096***	10067***	8753***	2702***	4070***
Cox & Snell R^2	0.57	0.51	0.28	0.28	0.42	0.38	0.14	0.20
Nagelkerke R^2	0.82	0.71	0.55	0.48	0.74	0.67	0.85	0.64

***, **, *and* * *indicate statistical significance at the 99%, 95%, and 90% levels.*

The models' results also reinforce previous observations about how neighborhoods that in 1991 were occupied by families living in disadvantaged social conditions (RC) went through a process of consolidation that was followed by an improvement in the social conditions of their residents, whereas the poorest families moved towards new areas far from the sea and with steeper slopes. The residential occupation of families belonging to social group B expanded towards neighborhoods that were previously characterized by disadvantaged social conditions (*NeighRC(t-n)*), a trend that became more pronounced in the period of 2000-2010. In addition, variables representing the areas located on hillsides ("SlopeZone") or intended for agricultural use ("Agriculture") had positive and significant coefficients for the models built for the residential class "Social Condition C", confirming a higher probability for having the poorest families occupying these newly built-up areas.

The models built to estimate the cells' suitability for the residential class "Social Condition A" revealed that most zoning variables were not statistically significant. The only statistically significant variables for both periods were those related to touristic residential zones ("Touristic"), the previous states of the cells and their neighborhoods (*"RA(t-n)", "RB(t-n)", "NeighRA(t-n)", "NeighRB(t-n)", "NeighRC(t-n)"*), the distance from the center and the sea (*"DistCenter" and "DistSea"*) and land slope ("*LandSlope*"). The negative estimated coefficients for the variable "DistSea" showed that the greater the distance from the sea, the lower the probability that a cell is occupied by social group A. This result is consistent with the prevailing culture of valuing the waterfront as a recreational space.

Regarding the variable "LandSlope", the greater the slope of the terrain is, the smaller the chance is that the cell is occupied by the most advantageous social group. The variable "DistCenter" showed a negative coefficient for the period 1991-2000 and a positive coefficient for the period 2000-2010. This pattern can be better understood when we observe Figure 4: whereas less vulnerable households are concentrated in the central area in 2000, they appear distributed in other parts of the coastline in 2010, which reinforces the idea that proximity to the city center is then not as important in the residential choices of this population group.

4.1.3 Simulation Experiments

Using the estimated regression coefficients presented in Table 3 as parameters for the allocation phase, simulation experiments were conducted to dynamically replicate spatial patterns observed from 1991 to 2010. To improve the fit between the simulated and real data, a calibration consisting of small changes in the allocation parameters was performed. A comparison between the simulation outputs and the real data is presented in Figure 5.

The observed trends of urban expansion and the spatial distribution of different social groups from 1991 to 2010 reveal strong intra-urban inequalities in Caraguatatuba. Regarding exposure to risks related to natural hazards, inundation occurrences (green triangles in Figure 6) have mainly affected residential areas classified as social conditions "B" and "C". The most disadvantageous families are also the ones who predominantly occupy areas that can be recognized in Figure 6 as subject to landslides or requiring special management due to its geological fragility. From the 420 residential cells under this condition (420 hectares), 205 were classified as "social condition C" and another 205 were classified as "social condition B" (total of 97.6%). In contrast, due to the valuation of areas close to the coastline, many of the areas occupied by the most advantageous social group ("A") are subject to hazards related to a rise in sea level and storm surges.

These features of the urbanization process in Caraguatatuba illustrate what has been noted as a major weakness in conception of a city: the inability to reconcile human and ecological imperatives in sustainable urbanization. In the next section, in which trends for the period 2010-2025 are simulated, it is possible to observe whether the Master Plan approved in 2011 tends to modify or reinforce the current situation.

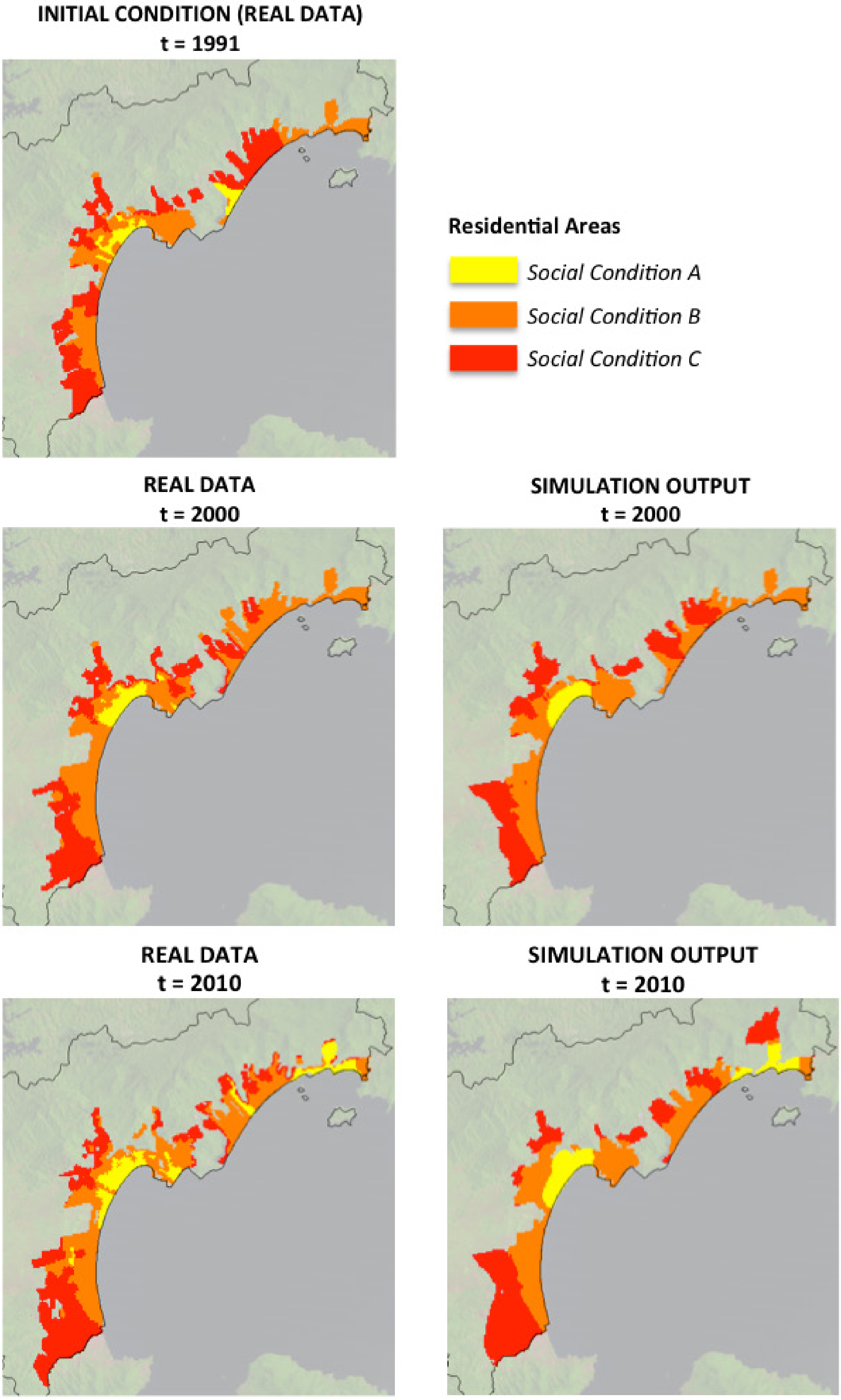

Fig. 5 - Comparison of simulation outputs with empirical data.

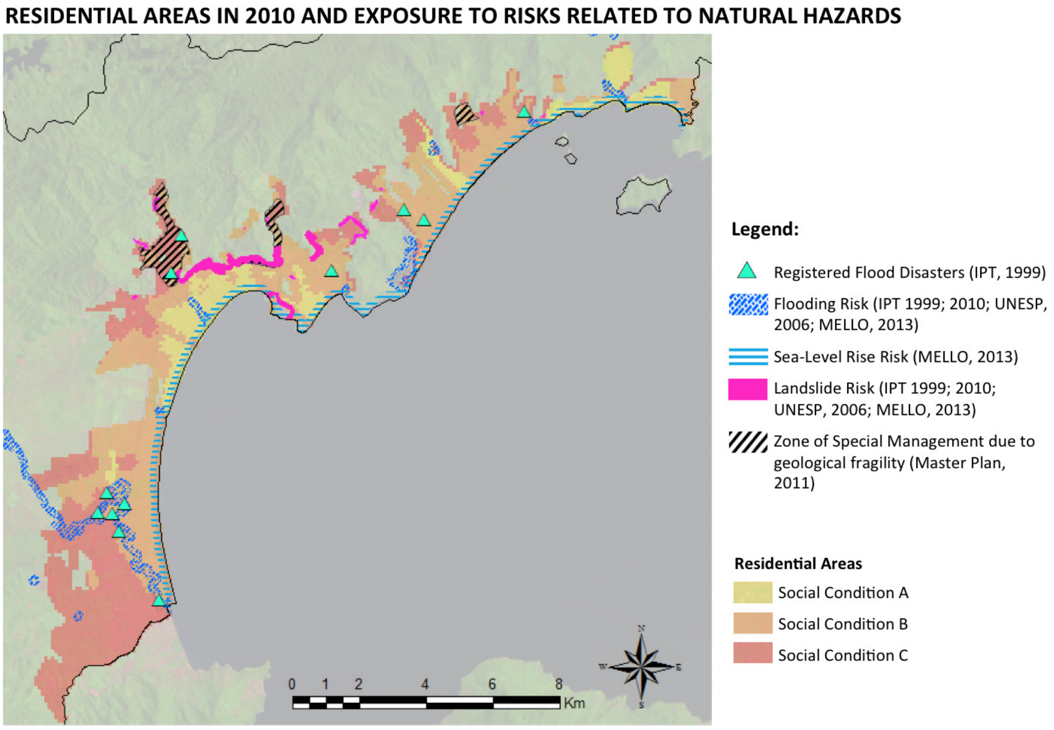

Fig. 6 - Exposure to risks related to natural hazards.
Adapted from Iwama, 2014; IPT, 2010; UNESP, 2006.

4.2 Period 2010-2025

The parameterization and design of experiments that simulate future dynamics (2010-2025) aims to shed light on the discussion of existing conflicts, possibilities in the face of different demographic and socioeconomic scenarios, and possible impacts of the Master Plan of Caraguatatuba.

To explore different demographic and socioeconomic scenarios, several demand parameters were tested (Table 4). The rate of urban expansion was changed to explore: (a) an accelerated growth of the demand for built-up areas due to growing migration levels promoted by the tourism and oil industries and (b) a decreased rate of urban expansion resulting from the densification of existing areas. To simulate scenarios of socioeconomic conditions, we also conducted experiments where the average proportions of different residential classes are modified.

To parameterize the allocation phase, a different component had to be incorporated into the analyses: the new Master Plan of Caraguatatuba (PREFEITURA MUNICIPAL DA ESTÂNCIA BALNEÁRIA DE CARAGUATATUBA, 2011). In comparison with the previous municipal law, the Master Plan of 2011 is more detailed and addresses the intensification of economic activities related to the tourism and oil industries and the growing pressure for new urbanized areas in the region. To this end, the plan increases the area designed for industrial use, reserves an area for the future construction of an airport and includes additional zoning categories, such as logistical and port support zones. In response to the expectation for a population increase in the city, areas that were previously intended for agricultural purposes are now designed as zones for urban expansion. In addition, the plan proposes the densification of many existing urban areas, which was an issue that sparked heated discussions during the Master Plan public meetings.

Based on a set of hypotheses, weights were established for each factor considered important for estimating the local probability of a cell belonging to a particular class. Although used as a reference, the weights derived from the logistic regressions could not be directly adopted in this process because the variables considered in the 2010-2025 simulations are not exactly the same as in the previous periods. This is especially due to the zoning variables, which refer to different laws depending on the analyzed period (the Zoning Law of 1992 and the 2011 Master Plan). The weights used in the allocation submodel of the simulations for 2010-2025 are summarized in Table 5.

Table 4 - URBIS-Caragua model: Demand parameters for the 2010-2025 Scenarios.

Scenario	**Urban Sprawl Rate (per year)**	**Proportion of Each Residential Class**		
		Social Condition A	**Social Condition B**	**Social Condition C**
1. Baseline (Stationary)	1.49%	14%	45%	41%
2. Increased Urban Sprawl	3%	14%	45%	41%
3. Decreased Urban Sprawl	0.7%	14%	45%	41%
4. Declining Socioeconomic Conditions	1.49%	7%	38%	55%
5. Improving Socioeconomic Conditions	1.49%	20%	60%	20%11

Table 5 - Simulation Experiments 2010-2025, Allocation Phase: Variables and weights considered when computing the cell suitability for the classes "Non Built-up" and the residential classes "Social Condition A", "Social Condition B" and "Social Condition C".

Explanatory Variables	**Non Built-up NR**	**Residential Classes**		
		Social Condition A ("Best") RA	**Social Condition B ("Intermediate") RB**	**Social Condition C ("Worst") RC**
Previous Cell State (t-n):				
NR(t-n)	3	-	-	-
RA(t-n)	-	4	1	-10
RB(t-n)	-	2	3	1
RC(t-n)	-	-1	1	3
Neighborhood externalities in (t-n):				
NeighNR(t-n)	4	-	-	-
NeighRA(t-n)	-	5	-4	-8
NeighRB(t-n)	-	-1	1	-1
NeighRC(t-n)	-	-4	-0.1	1
Zoning (Figure 8):				
ExcResidential_L	-	6	-1	-5
Social	-	-10	0	6
Residential&Mixed	-4	-	-	-
Buffer	8	-	-	-
EcoTourist	12	-	-	-
Conservation	10	-	-	-
Conservation_2	12	-	-	-
Other factors:				
DistSea	1 (10^{-4})	-3 (10^{-3})	-5 (10^{-4})	-8 (10^{-5})
LandSlope	1 (10^{-2})	-5 (10^{-2})	-2 (10^{-2})	-1 (10^{-3})

With the exception of the zoning variables, all of the factors presented in Table 5 are the same as the ones in Table 2. For all classes, we assumed that the potential of a cell to remain in the same class in time *t* increases if the cell and its neighborhood already presented this state in *t-n*. Regarding the residential classes ("RA", "RB", and "RC"), neighborhood externalities remain important, assuming that higher-income families prefer to choose neighborhoods where poorer families are not present. Therefore, the probability of a cell being occupied by the privileged residential class (social condition A - "RA") decreases when the cell is already occupied by families belonging to the social class C, and vice versa.

The explanatory variables "distance from the sea" ("*DistSea*") and "land slope" ("*LandSlope*") were considered as factors that increase the probability of a cell belonging to the class "non built-up" and decreases the probability of a cell belonging to any residential class. Nevertheless, because the residential areas distant from the sea and with higher slopes tend to be occupied by disadvantaged families (social condition C), the weights for these variables are set differently for each residential class.

Considering the zoning variables, we assumed that the potential of a cell to belong to the class "non built-up (NR)" at time *t* will decrease if it is located inside zones that, according to the Master Plan, are expected to accommodate residential uses (variable "*Residential&Mixed"*). Such areas include the new urban expansion zone and all of the residential and mixed zones.

In contrast, the potential of a cell to belong to the class "non built-up (NR)" at time *t* should increase in zones that are environmentally protected. These zones include the following:

- Permanent Conservation Zone ("*Conservation_2"*): Areas of permanent protection and conservation, with significant importance attached to environmental integrity and the conservation of the biodiversity of flora and fauna.
- Environmental Conservation Zone ("*Conservation_1"*): Integrated ecological corridors for flood control and environmental protection.
- Ecological Tourist Zone ("*EcoTourist"*): Zones for tourism and recreation in special areas for environmental integrity.
- Buffer Zone ("*Buffer"*): The area between the Ecological Tourist Zone and areas subject to urbanization.

Regarding the residential classes, two zoning types were considered particularly important in driving or consolidating the location of different social groups. The first is the "Special Social Interest Zone" ("*Social*"), which is expected to be occupied by families belonging to social condition C (class "RC"). The second is the "Exclusively Residential Zone - Low Density" ("*ExcResidential_L*"), which sets building standards that can be afforded only by the most affluent families (class "RA").

The weights presented in Table 5 were kept constant for all five simulated scenarios for the period 2010-2025: (a) baseline, (b) increasing the urban sprawl rate, (c) decreasing the urban sprawl rate, (4) improving socioeconomic conditions, and (5) declining socioeconomic conditions. The simulated outcomes obtained for the year 2025 are presented in Figures 7, 8 and 9.

The baseline scenario for 2025, which keeps the demand parameters equal to the ones observed during the period of 2000-2010, is presented in Figure 7. The scenario reveals an expansion of residential areas in the south-central region of the city, particularly in zones that, according to the new Master Plan, are intended for mixed use and urban expansion.

The newly created Zone of Urban Expansion is poorly detailed in the Master Plan and remains in conflict with an environmental legal instrument, the Ecological-Economic Zoning (SÃO PAULO, 2004), which is expected to be reviewed to accommodate new demands such as this one. For the purpose of the present simulation experiments, we included no restriction to the urbanization of these areas, and the simulation outputs indicate that they tend to be occupied by families belonging to social groups "B" and "C". However, it is important to note that most of the areas available for urban

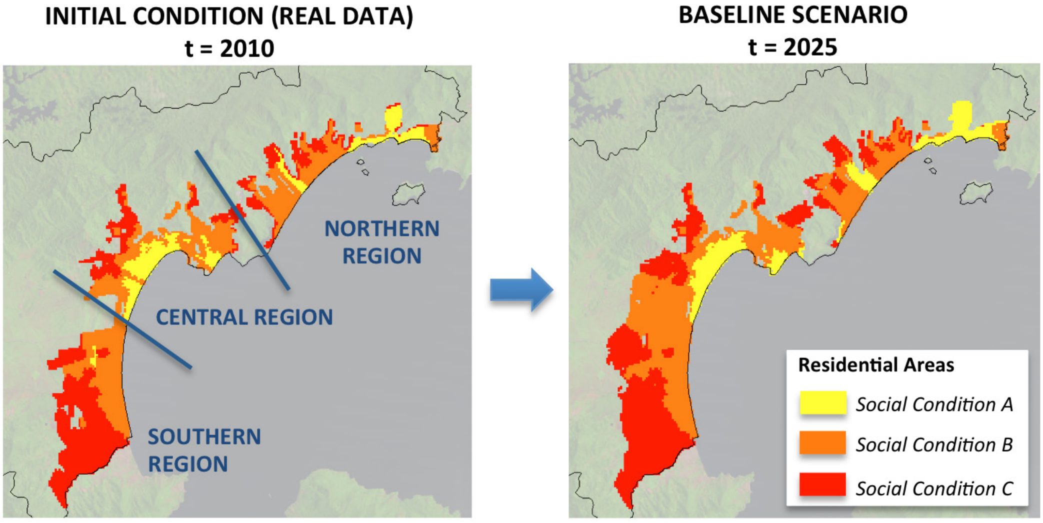

Fig. 7 - Simulation outcome (year 2025) for the baseline scenario.

expansion belong to private firms. Therefore, if innovative real estate investments for high-income families are made in these expansion areas, different urbanization trends could be observed.

Regarding the spatial distribution of social groups, the most affluent families tend to remain close to the coastline, particularly in the central and northern zones. By delimiting some areas in the northern region as low-density and exclusively residential zones, the municipal Master Plan consolidates its current vocation as a resort area for wealthy families, with many gated communities and most houses serving as second homes. Reinforcing this trend, the plan also envisages the construction of a marine facility and other recreational facilities on the northern coast of the city. This situation reveals how the segregation of the richest strengthens their power to pressure for public facilities and high-quality urban space.

In contrast, the poorest families are mainly concentrated in the southern region (Figure 7). In this region, the Master Plan defines not only residential and mixed zones but also zones intended for other uses that are not desirable in the vicinity of the richest residential areas, such as the industrial, logistical and port support zones. The largest area intended for a "Special Social Interest Zone" is also located in the south and reinforces the concentration of the poorest families in the region. It is also important to note that this particular zone is located in the plain of the River Juqueriquerê, where floods are frequent. In general, the Master Plan was not able to provide changes in the unequal distribution of risks that was previously observed in Figure 6, thus consolidating trends that have been historically driven by market forces.

Comparing the baseline scenario for 2025 with an alternative scenario that considers increasing urban sprawl rates (Figure 8), it is

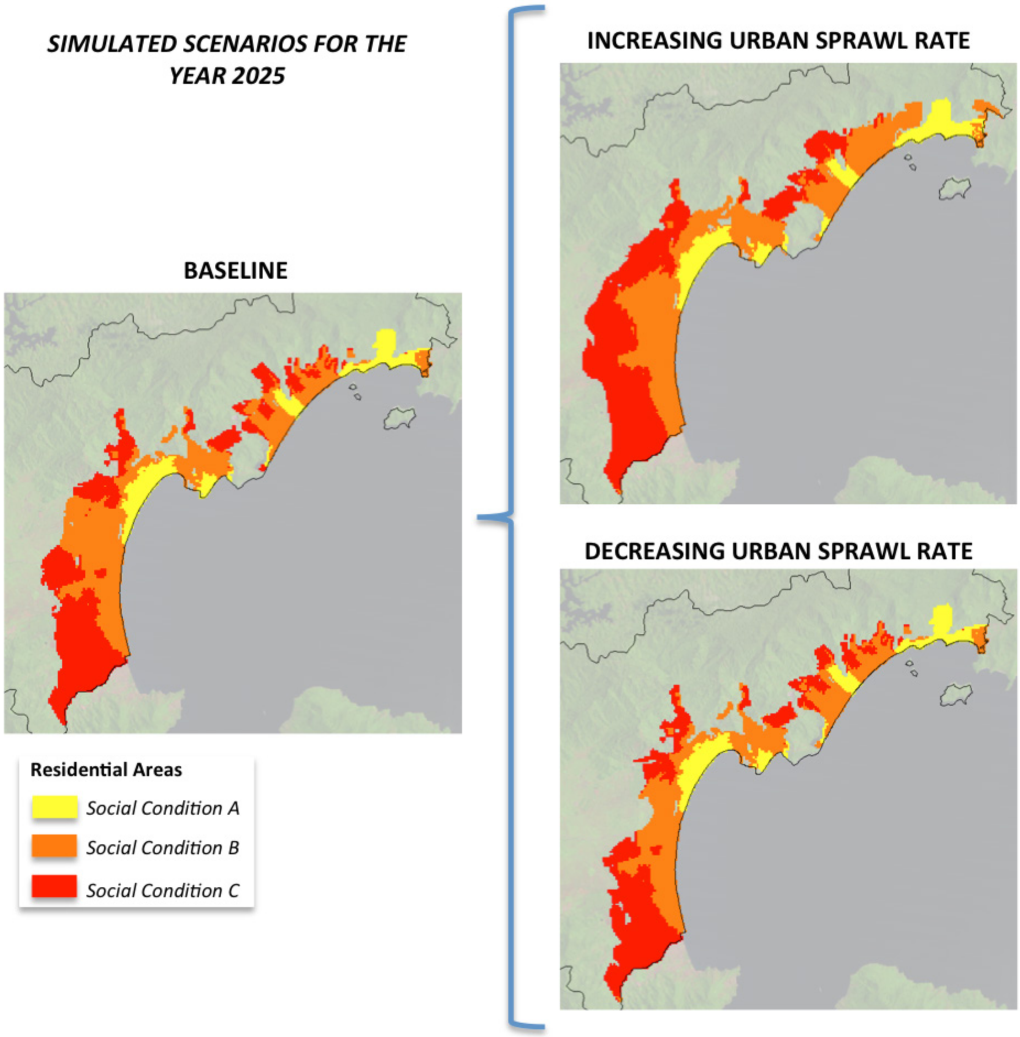

Fig. 8 - Simulation outcomes (year 2025) for the scenarios "Baseline", "Increasing Urban Sprawl Rate", and "Decreasing Urban Sprawl Rate".

possible to observe how the latter situation tends to lead to the further development of residential areas classified as "social condition C" towards the new zones of urban expansion in the south. In the northern and central regions, however, the situation becomes more dramatic. The scenario shows a higher pressure of occupation by the poorest populations toward areas of permanent preservation. Associated with this trend is an increased exposure of these populations to the risk of landslides and a greater vulnerability of the rainforest ecosystem.

The alternative scenario that considers decreasing urban sprawl rates (Figure 8) simulates a different urban growth trend that is denser and vertical and exhibits lower pressure on the new areas for urbanization. Although, on the one hand, this trend decreases urban sprawl, on the other hand, it can be criticized for meeting the demands of the real estate market, which seeks to maximize profits at the expense of the general welfare.

The results of the simulation experiments presented in Figure 9 show scenarios with improving and declining socioeconomic conditions. In the first case, it is possible to observe a stronger presence of richer families on the northern and central coastal lines. Compared with the baseline scenario, there is greater pressure for growth in the most highly valued residential areas (northern region). In the central-southern region, the expansion of social groups B and C towards zones designed as mixed or for urban expansion remains significant but is less intense. In contrast, the scenario that simulates declining socioeconomic conditions shows how these groups occupy the new expansion areas in a much more expressive manner.

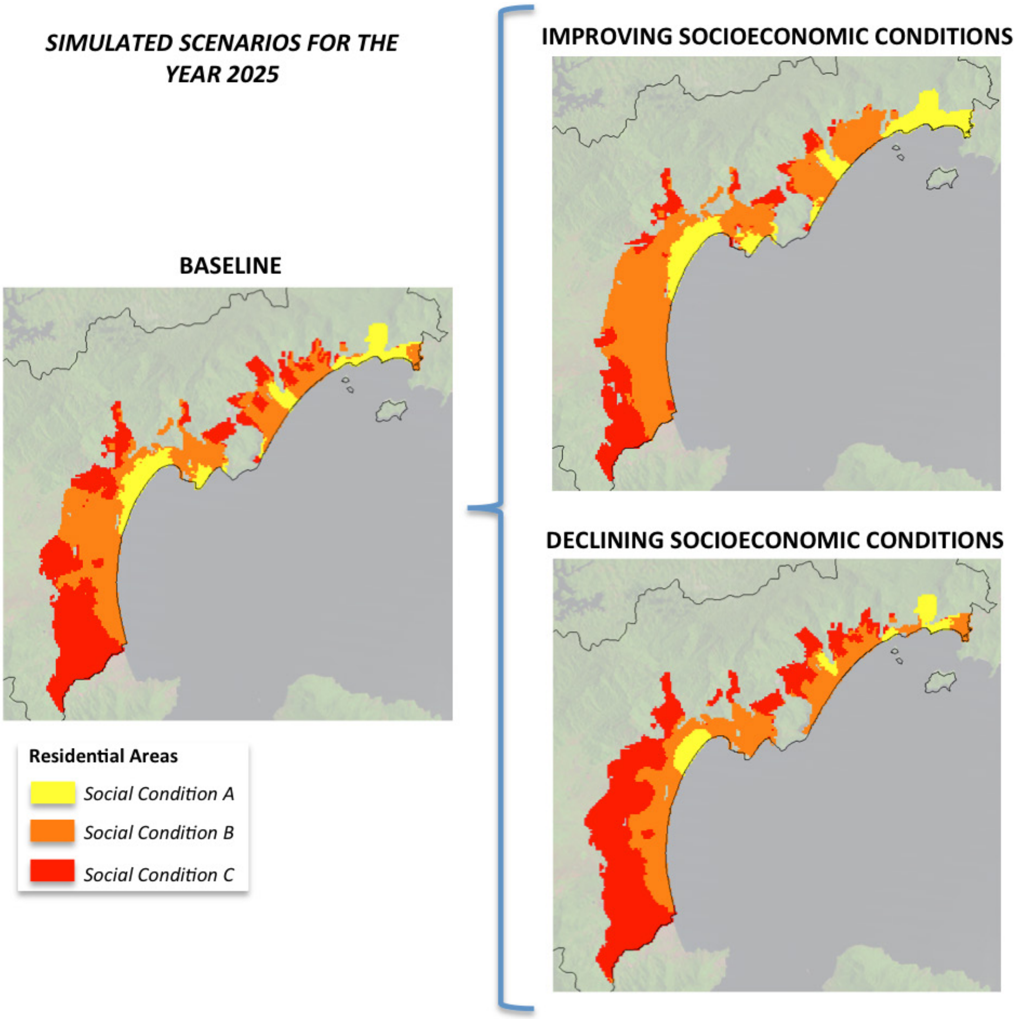

Fig. 9 - Simulation outcomes (year 2025) for the scenarios „Baseline“, „Improving Socioeconomic Conditions“, and „Declining Socioeconomic Conditions“.

5. CONCLUSIONS

This work presents the CA model *URBIS-Caragua*, which was developed to analyze and explore how recent transformations in the northern coast of São Paulo are driving the dynamics of urban stratification in Caraguatatuba. Instead of focusing on presenting the results of the model, this paper emphasizes the analytical potential of a stage that is commonly overlooked: the construction of the model.

Unlike a common modeling practice in which data about the past are used to parameterize experiments about the future, this work assumes that the relevance of conditioning factors for urbanization is constantly changing over time. Thus, the phases of parameterizing and designing simulation experiments on past and future dynamics present particular goals. On the one hand, the experiments on past periods (1991-2000 and 2000-2010) sought to support the process of discussing and learning about how trends of urban expansion and the spatial distribution of social groups have changed over the past two decades as well as the possible links between these spatial patterns and legislative, locational and sociodemographic factors. On the other hand, the experiments on future dynamics (2010-2025) sought to explore possible outcomes in the face of different demographic and socioeconomic scenarios and investigate how (or whether) the recently approved Master Plan of the city could promote changes in the predominant spatial distribution of social groups and the uneven distribution of risks and opportunities among them. As a result, the experiments indicated that the Master Plan is likely to reinforce the already prevailing trends of urban stratification.

The process of building a model provides not only a thorough analysis of numerous relevant aspects of the system of interest but also an increased awareness of the strengths and weaknesses of the approach and the most suitable way to interpret the results obtained from simulation experiments. To avoid the results from the simulations being treated as deterministic truths or, conversely, with total skepticism, it is essential to know the details of the process of building a model, including its assumptions, data, parameterization strategies and how different elements and interactions in the system are represented.

ACKNOWLEDGEMENTS

The authors would like to thank the financial support received from the National Council for Scientific and Technological Support - CNPq (Process 443052/2014).

REFERENCES

ABRAMO, P. La ciudad com-fusa: mercado y producción de la estructura urbana en las grandes metrópolis latinoamericanas. **EURE** (Santiago), v. 381, n. 114, p. 35–69, 2012.

AGUIAR, A. P. D.; CARNEIRO, T. G. S., ANDRADE, P. R.; ASSIS, T. LuccME-TerraME: an open source framework for spatially explicit land use change modelling. **GLP News**, v. 8, p. 21–23, 2012.

ANAZAWA, T. M.; FEITOSA, F. F.; MONTEIRO, A. M. V. Vulnerabilidade socioecológica no Litoral Norte de São Paulo: Medidas, superfícies e perfis de ativos. **Geografia** (Rio Claro), v. 38, n. 1, p. 189–208, 2013.

ANAZAWA, T. M.; FEITOSA, F. F.; MONTEIRO, A. M. V. Vulnerabilidade socioecológica nas áreas ocupadas nos municípios de Caraguatatuba e São Sebastião, litoral norte de São Paulo: medidas e representações. **Revista Espinhaço**, v. 3, n. 1, p. 138-151, 2014.

ANDRADE, P. R. ; CÂMARA, G. ; MARETTO, R. V. ; MONTEIRO, A. M. V. ; CARNEIRO, T. ; FEITOSA, F. F. Experiences with a Socio-Environmental Modeling Course. **Modelling in Science Education and Learning,** v. 8, n. 1, p. 71-92, 2015.

BATTY, M. **Cities and Complexity: Understanding Cities with Cellular Automata, Agent-based Models, and Fractals.** The MIT Press, 2005. 565p.

BATTY, M. **The New Science of Cities**. London, The MIT Press, 2013. 520 p.

BATTY, M.; COUCLELIS, H.; EICHEN, M. Urban Systems as Cellular Automata.

Environment and Planning B: Planning and Design, v. 24, p. 159-164, 1997.

CARLOS, A. F. A.; SOUZA, M. L.; SPOSITO, M. E. B. **A Produção do Espaço Urbano: Agentes e Processos, Escalas e Desafios**. São Paulo, Contexto, 2011. 234p.

CASTELLS, M. **The Urban Question: A Marxist Approach**. Cambridge, The MIT Press, 1979. 502 p.

CARNEIRO, T. G. S.; ANDRADE, P. R.; CÂMARA, G.; MONTEIRO, A. M. V.; PEREIRA, R. R. An extensible toolbox for modeling nature-society interactions. **Environmental Modelling and Software**, v. 46, p. 104-117, 2013.

FEITOSA, F. F.; MONTEIRO, A. M. V. Vulnerabilidade e Modelos de Simulação como Estratégias Mediadoras: Contribuição ao Debate das Mudanças Climáticas e Ambientais. **Revista Geografia** (Rio Claro), v. 37, n. 2, p. 289–305, 2012.

IBGE - Instituto Brasileiro de Geografia e Estatística. **Censo Demográfico 2010**. Rio de Janeiro: IBGE, 2010. Disponível em < http://www.ibge.gov.br/ >. Acesso: 15 abril 2016.

IBGE - Instituto Brasileiro de Geografia e Estatística. **Estimativas da população residente nos municípios e para as unidades da federação brasileiros com data de referência em 1º de julho de 2016.** Rio de Janeiro: Instituto Brasileiro de Geografia e Estatística, 2016. Disponível em < http://www.ibge.gov.br/home/estatistica/populacao/estimativa2016/ >. Acesso: 03 abril 2017.

IPT - Instituto de Pesquisas Tecnológicas do Estado de São Paulo (2010). **Mapeamento de risco 2010 - Caraguatatuba. Parecer Técnico n. 18578/301.** São Paulo, IPT, 2010. 123p.

KAUFMAN, L.; ROUSSEEUW, P. J. **Finding groups in data: an introduction to cluster analysis.** Hoboken, NJ, John Wiley & Sons, 2008. 342 p.

MOORE, D. S.; MCCABE, G. P. **Introduction to the Practice of Statistics.** New York, W. H. Freeman & Company, 2003. 828 p.

MORAES, A. C. R. **Contribuições para a Gestão da Zona Costeira no Brasil: Elementos para uma Geografia do Litoral Brasileiro**. São Paulo, Annablume, 2007. 232 p.

IWAMA, A. Y. **Riscos e vulnerabilidades às mudanças climáticas e ambientais: análise multiescalar na zona costeira de São Paulo - Brasil.** Tese de Doutorado. Núcleo de Estudos e Pesquisas Ambientais do Instituto de Filosofia e Ciências Humanas da Universidade Estadual de Campinas (IFCH/NEPAM/UNICAMP), Campinas, 2014. 353 p.

SANTÉ, I.; GARCÍA, A.; MIRANDA, D.; CRECENTE, R. Cellular automata models for the simulation of real-world urban processes: A review and analysis. **Landscape and Urban Planning,** v. 96, n. 2, p.108-122, 2010.

SANTOS, M. O **Espaço do Cidadão**. Coleção Milton Santos, 7a. ed., v.8, São Paulo, EDUSP, 2012. 176p.

UNESP - Universidade Estadual Paulista. **Elaboração de plano de redução de riscos da Estância Balneária de Caraguatatuba (SP). Relatório Técnico.** São Paulo, Fundação de Apoio à Pesquisa Ensino e Extensão – FUNEP, 2006. 108 p.

VERBURG, P. H.; SOEPBOER, W.; LIMPIADA, R.; ESPALDON, M. V. O.; SHARIFA, M. A.; VELDKAMP, A. Modelling the spatial dynamics of regional land use: The CLUE-S model. **Environmental Management**, v. 30, n. 3, p. 391–405, 2002.

PREFEITURA MUNICIPAL DA ESTÂNCIA BALNEÁRIA DE CARAGUATATUBA. **Lei Complementar n. 42, de 24 de Novembro de 2011**. 2011.

PREFEITURA MUNICIPAL DA ESTÂNCIA BALNEÁRIA DE CARAGUATATUBA. **Lei Complementar n. 200, de 22 de Junho de 1992**. 1992.

SÃO PAULO. **Decreto n. 49.215, de 07 de Dezembro de 2004**. 2004.

WAHYUDI, A.; LIU, Y. Cellular Automata for Urban Growth Modelling: A Review on Factors Defining Transition Rules. **International Review for Spatial Planning and Sustainable Development,** v. 4, n. 2, p. 60-75, 2016.