"Understanding and Modelling Urban Land Use Dynamics to Achieve Sustainability"

Cláudia Maria de Almeida[1], Antonio Miguel Vieira Monteiro[2], Gilberto Câmara[3], Britaldo Silveira Soares-Filho[4], Gustavo Coutinho Cerqueira[5], Cássio Lopes Pennachin[6]

[1]PhD Student in Remote Sensing and GIS – National Institute for Space Research (INPE) – almeida@ltid.inpe.br, phone: + 55-12-3906-5725, fax: +55-12-3906-5725; [2]Division for Image Processing – National Institute for Space Research (DPI-INPE) – miguel@dpi.inpe.br, phone: +55-12-3945-6444, fax: +55-12-3945-6468; [3]gilberto@dpi.inpe.br, phone: +55-12-3945-6499, fax: +55-12-3945-6468, Av. dos Astronautas, 1758 – 12227-010, São José dos Campos, SP - Brazil

[4]Department of Cartography – Federal University of Minas Gerais (UFMG) – britaldo@csr.ufmg.br, phone: +55-31-3499-5417, fax: +55-31-3499-5415; [5]Centre for Remote Sensing – Federal University of Minas Gerais (CSR-UFMG) – cerca@csr.ufmg.br, Av. Antônio Carlos, 6627 – 31270-900, Belo Horizonte, MG – Brazil

[6]Intelligenesis do Brasil Ltda. – cassio@intelligenesis.net, Av. Brasil, 1438-1505, 30140-003, Belo Horizonte, MG – Brazil

Abstract. This scientific paper, part of a PhD Thesis currently under execution at the Division for Image Processing of the Brazilian National Institute for Space Research (DPI – INPE), is committed with building up a methodological guideline for modelling urban land use dynamics. A medium-size town in the west of São Paulo State, *Bauru*, was adopted as case study. Its urban structure was converted into a 100 x 100 (m) resolution grid, and transition probabilities were calculated for each grid cell by means of the "weights of evidence" statistical method and upon basis mainly of the information related to the technical and social infrastructure of the town. The probabilities therefrom obtained fed a cellular automaton (CA) simulation model – DINAMICA- conceived by the Centre for Remote Sensing of the Federal University of Minas Gerais (CSR-UFMG), based on a multiscale vicinity approach and stochastic transition algorithms. Different simulation outputs for the case study town in the period 1979-1988 were generated, and statistical validation tests were then conducted for the best results, employing a multiple resolution fitting procedure.

This modelling experiment revealed the plausibility of adopting Bayesian empirical methods based on the available infrastructure knowledge to simulate urban land use change, what implies their further applicability for generating forecasts of growth trends either for Brazilian towns or cities worldwide, so as to help planners and local authorities in achieving sustainability.

Keywords: Urban Modelling,Urban Sustainability, Town Planning, Cellular Automata

1. Introduction

Recent generation models of urban dynamics have been dealing with diverse themes. According to Batty (2000), there are currently some twenty or more applications of CA to cities, such as the diffusion or migration of resident populations (Portugali et al., 1997), the competitive location of economic activities (Benati, 1997), the joint expansion of urban surface and traffic network (Batty and Xie, 1997), the generic urban growth (Clarke et al., 1997), the urban land use dynamics (Deadman et al., 1993; Batty and Xie, 1994; Phipps and Langlois, 1997; White and Engelen, 1997; White et al., 2000) , and so forth.

Specifically regarding urban land use dynamics, it is possible to identify basically three main trends of CA models in this field. A first one concerns the deterministic models, whose most evident representative is the urban growth study for the San Francisco Bay area, conducted by Clarke et al. (1997). Although this model incorporates a certain randomness in selecting the cells for urban growth and in promoting the spread of growth seeds, its transition rules, which can be spontaneous, diffusive, organic or road-influenced, are fundamentally deterministic in the sense that the cell suitability for being urbanised is not dependent upon probabilistic methods.

A second trend relates to the stochastic models with both deterministic estimations of area for land use transition and deterministic transition algorithms. A good example of this category of models is the SIMLUCIA, conceived by White et al. (2000), which is an integrated model of natural and human systems operating at several spatial scales, and was aimed at providing the officials of the Caribbean Island of Santa Lucia with a tool to explore possible environmental, social, and economic consequences of hypothesised climate changes.

In this model, a sophisticated set of equations taking into account aspects of the natural environment is formulated in order to estimate the impact of economic and demographic changes on land use. The stochasticity of this model is present in the calculation of the probabilities of land use transition for each cell, which is basically a function of the cell suitability for the new activity in question and its relative accessibility for such an activity. In the SIMLUCIA transition algorithm, cells are ranked by their highest potential, and cell transitions begin with the highest ranked cell and proceed deterministically downward, until the number of cells demanded by the above-mentioned equations is reached.

A third trend concerns the stochastic CA models with both stochastic estimations of area for land use transition and stochastic transition algorithms. The modelling experiment presented in this paper integrates this third category, in which the transition rules are randomised, the cell transition probabilities are calculated through Bayesian probabilistic methods ("weights of evidence"), and the Markov chain is in principle utilised for the definition of the transition rates for each possible type of land use change. An overview of the "weights of evidence" statistical method as well as an explanation of how it can be applied to the modelling of urban land use dynamics are presented throughout the next section.

2. Methods: A Bayesian Method-Based Cellular Automaton Model

Generalisation Procedures Applied to the Land Use Maps

The city maps provided by the Bauru local authorities presented inconsistencies due to the fact that illegal settlements are not shown on the official maps, and not all of the legally approved settlements drawn have been in fact implemented. Moreover, some urban zones refer to areas which are not yet occupied, and some other zones categories do not correspond to the prevailing use indeed encountered within their limits, reflecting just the local officials´ intention for their future use. In this way, satellite imagery arise as a feasible solution for the identification of urban settlements actually existent, as well as for the delineation of the true urban occupation boundaries of the case study town.

The following procedures were applied to the initial (1979) and final (1988) land use maps (Figure 1) used in the simulation experiment so as to render them workable by the computational model and coherent to the reality they are related to:

(i) reclassification of the zones initially assigned by the Bauru local authorities according to their dominant and effectively existent use;

(ii) reclassification of similar zones shown on official maps to only one category, e.g.: residential zones of different densities are all reclassified to residential zones only; special use zones and social infrastructure equipments zones are reclassified to institutional zones only, and so on;

(iii) adoption of eight land use zone categories: residential, commercial, industrial, services, institutional, mixed use zone, leisure/recreation, and non-urban zone;

(iv) exclusion of districts segregated from the main urban agglomeration, i.e. those which are located above 10 km from the official urban boundary;

(v) disregard of the traffic network and minor non-occupied areas in the simulations.

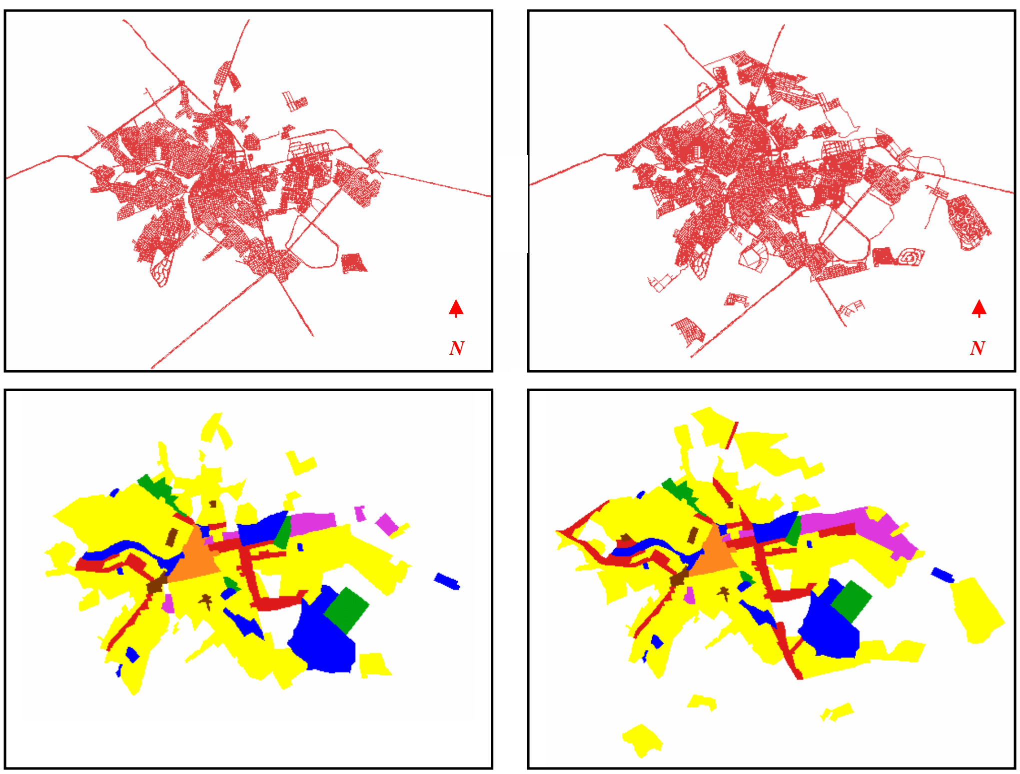


Fig. 1 – Official Bauru city maps for the years 1979 (upper left) and 1988 (upper right), and derived land use maps for 1979 (bottom left) and 1988 (bottom right).

The land use maps on Figure 1 are the results of the application of the generalisation procedures listed above on the official city maps. The yellow colour represents the residential use; the orange, the commercial use; the purple relates to the industrial use; the blue refers to the institutional use; the red corresponds to the services use zones and corridors; the brown is related to the mixed use zones; the green represents the leisure and recreation use; and the white refers to the non-urban use.

Exploratory Analysis and Selection of Variables

Some pressing constraints to the process of modelling urban land use dynamics became evident during the accomplishment of the experiment reported in this paper, namely:

- the variables available for the modelling analysis not always represent the set of necessary variables able to produce highly satisfactory results;
- the urban land use dynamics is subject to sudden and unforeseeable forces, like the landlords´ decisions to develop certain areas in disregard of others, which constitute unsuitable factors for modelling;
- areas subject to “urbanisation booms” - as it is the case of Bauru - are often regarded as chaotic or highly complex systems (Clarke et al., 1997), what render the current computational modelling technologies not best appropriate to cope with such phenomena.

Nevertheless, there is indeed a set of decisive factors for urban land use transition, suitable for modelling and commonly filed by city planning departments of local governments, and which have effectively guided the modelling experiment in question.

Some of the maps of explaining variables related to the technical and social infrastructure of Bauru and employed in the modelling analysis are presented below. Initially, these maps were scanned in the hollandaise OCE scanner (model G6035S) and digitised in AutoCad release 14. These maps were then exported as files with extension DXF to the Geographic Information System (GIS) termed SPRING, conceived by the Division for Image Processing of the Brazilian National Institute for Space Research (DPI-INPE). It is worth mentioning that these procedures were also adopted for the production of the Bauru city and land use maps presented in Figure 1. In SPRING, the maps of variables were then subjected to a preliminary processing, including vector edition, polygons identification, elaboration of distance maps and spatial statistical analysis maps like the Kernel points density estimator, etc. (Figure 2).

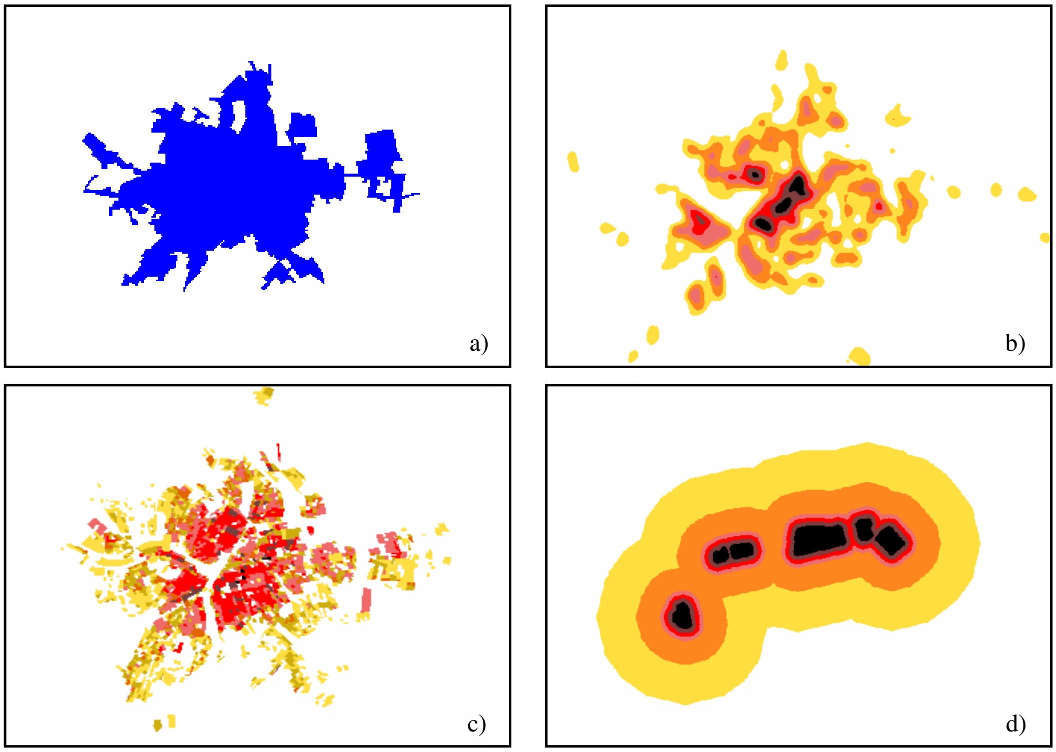


Figure 2 – Examples of maps of variables: (a) area served by water supply in Bauru, 1979; (b) Kernel estimator for the density of commercial establishments in Bauru, 1979; (c) density of occupation in Bauru, 1979; (d) map of distance to industrial zones in Bauru, 1979.

Since the "weights of evidence" statistical method (to be employed in the calculation of the cells transition probabilities) is based on the "Bayes theorem of conditional probability", the selection of variables for the modelling analysis should take into account the checking of independence amongst pairs of variables chosen to explain the same category of land use change.

For this end, two methods were used: the Cramers Coefficient (V) and the Joint Information Uncertainty (U). In both cases, it is necessary to obtain values from an area cross-tabulation between pairs of maps of variables under analysis. Let the area table between map A and map B be called matrix T, with elements T_{ij}, where there are i = 1, 2, ..., n classes of map B (rows of the table) and j = 1, 2, ..., m classes of map A (columns of the table). The marginal totals of T are defined as $T_{i.}$ for the sum of the i-th row, $T_{.j}$ for the sum of the j-th column, and $T_{..}$ for the grand total summed over rows and columns. If the two maps are independent of one another, with no correlation between them, then the expected area in each overlap category is given by the product of the marginal totals, divided by grand total. Thus the expected area $T_{ij}*$ for the i-th row and j-th column is:

$$\boxed{T_{ij}* = \frac{T_{i.} \; T_{.j}}{T_{..}}} \quad . \qquad (1)$$

Then, the **chi-square** statistic is defined as:

$$\boxed{\chi^2 = \sum_{i=1}^{n} \sum_{j=1}^{m} \frac{(T_{ij} - T_{ij}*)^2}{T_{ij}*}} \quad , \qquad (2)$$

the familiar (*observed – expected*)^2 / *expected* expression, which has a lower limit of 0 when the observed areas exactly equal the expected areas, and the two maps are completely independent. The Cramers Coefficient (V) is then defined as:

$$\boxed{V = \sqrt{\frac{\chi^2}{T_{..} \, M}}} \quad , \qquad (3)$$

where M is the minimum of (n-1, m-1).

The Joint Information Uncertainty (*U*) belongs to the class of **entropy** measures, which are also based on the area cross-tabulation matrix *T*, but can also be used for measuring associations. Suppose that the T_{ij} values are transformed to area proportions, *p*, by dividing each area element by the grand total $T_{..}$. Thus, $p_{ij} = T_{ij} / T_{..}$, and the marginal proportions are defined as $p_{i.} = T_{i.} / T_{..}$ and as $p_{.j} = T_{.j} / T_{..}$. Therefore entropy measures, also known as *information statistics* can be defined using the area proportions as estimates of probabilities. Proportions are dimensionless, so entropy measures have the advantage over chi-squared measures of being unaffected by measurement units (Bonham-Carter, 1994).

Assuming that an area proportions matrix for map A and map B has been determined from *T*, then the *entropy* of A and B are defined as:

$$\boxed{H(A) = - \sum_{j=1}^{m} p_{.j} - \ln p_{.j}} \quad \text{and} \tag{4}$$

$$\boxed{H(B) = - \sum_{i=1}^{n} p_{i.} - \ln p_{i.}} \quad , \tag{5}$$

where ln is the natural logarithm. The **joint entropy** of the combination, *H(A,B)*, is simply

$$\boxed{H(A,B) = - \sum_{i=1}^{n} \sum_{j=1}^{m} p_{ij} \ \ln p_{ij}} \quad . \tag{6}$$

Then the "***Joint Information Uncertainty***" of A and B, *U(A,B),* can be used as a measure of association and is defined as

$$\boxed{U(A,B) = 2 \left[\frac{H(A) + H(B) - H(A,B)}{H(A) + H(B)} \right]} \quad , \tag{7}$$

which varies between 0 and 1. When the two maps are completely independent, then $H(A,B) = H(A) + H(B)$ and $U(A,B)$ is 0, and when the two maps are completely dependent, $H(A) = H(B) = H(A,B) = 1$, and $U(A,B)$ is 1.

Table 1 shows the codes utilised for each map of variable employed in the modelling experiment, and Tables 2 and 3 presents the values obtained for the Cramers Coefficient and the Joint Information Uncertainty for the pairs of variables used to explain the same type of land use transition.

TABLE 1 – CODES AND MEANINGS FOR THE MAPS OF VARIABLES

Code	*Meaning*
water	*Area served by water supply.*
mh_dens	*Medium-high density of occupation (25% to 40%).*
soc_hous	*Existence of social housing.*
com_kern	*Distances to different ranges of commercial activities concentration, defined by the Kernel estimator.*
dist_ind	*Distances to industrial zones.*
dist_res	*Distances to residential zones.*
per_res	*Distances to peripheral residential settlements, isolated from the urban concentration.*
dist_inst	*Distances to social infrastructure (institutional) equipments, isolated from the urban concentration.*
exist_rds	*Distances to main existent roads.*
serv_axes	*Distances to the services and industrial axes.*
plan_rds	*Distances to planned roads.*
per_rds	*Distances to peripheral roads, which pass through non-occupied areas.*

For the Cramers Coefficient, the empirically established threshold was ***0.45***, and for the Joint Information Uncertainty, ***0.35***. As none of the association measure values surpassed the thresholds, no variables preliminarily selected for modelling have been discarded from the analysis.

In practice, the variables selection routine also include empirical procedures, based on the visualisation of distinct variables superposed on the final land use map, so as to identify those more meaningful to explain the different types of land use change (Figure 3).

TABLE 2 – *CRAMERS COEFFICIENT* FOR PAIRS OF VARIABLES EXPLAINING THE SAME TYPE OF LAND USE CHANGE – BAURU (1979-88)

	water	mh_dens	soc_hous	com_kern	dist_ind	dist_res	per_res	dist_inst	exist_rds	serv_axes	plan_rds	per_rds
water										0.3257		
mh_dens			0.046								0.2617	0.0201
soc_hous											0.1174	0.048
com_kern						0.4129	0.1142	0.1218	0.2685	0.2029		0.0434
dist_ind										0.1466		
dist_res										0.2142		
per_res								0.1487	0.0592			0.1733
dist_inst									0.0601			0.0765
exist_rds												0.0239
serv_axes												
plan_rds												0.0247
per_rds												

TABLE 3 – *JOINT INFORMATION UNCERTAINTY* FOR PAIRS OF VARIABLES EXPLAINING THE SAME LAND USE CHANGE – BAURU (1979-88)

	water	mh_dens	soc_hous	com_kern	dist_ind	dist_res	per_res	dist_inst	exist_rds	serv_axes	plan_rds	per_rds
water										0.0767		
mh_dens			0.00176								0.070062	0.000272
soc_hous											0.018788	0.004655
com_kern						0.34472	0.031004	0.05202	0.14993	0.10996		0.006365
dist_ind										0.04766		
dist_res										0.10024		
per_res								0.05592	0.00774			0.055321
dist_inst									0.01017			0.023836
exist_rds												0.001957
serv_axes												
plan_rds												0.002967
per_rds												

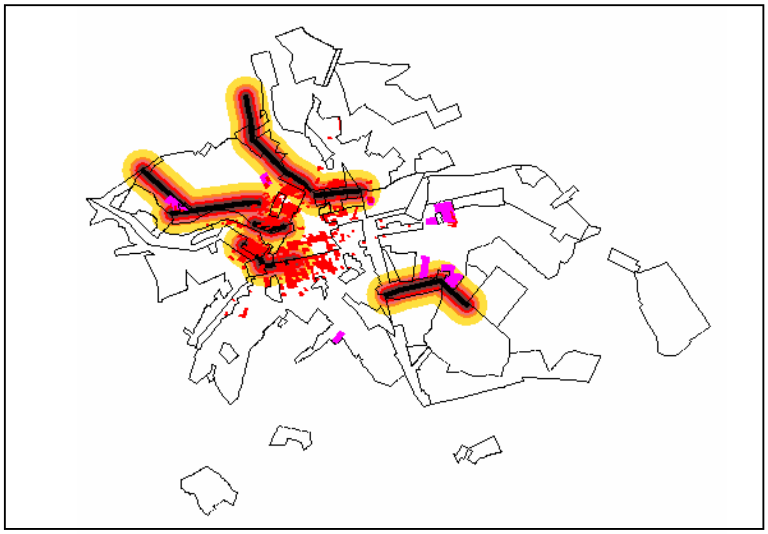


Fig. 3 – Example of the overlay (conducted in SPRING) of different explaining variables existent in the city of Bauru in 1979 on the final land use map - 1988, so as to identify those more meaningful to explain the transition “residential use - mixed use”. The distance ranges are built in relation to planned roads; the purple polygons correspond to social housing; and the red blocks refer to areas with medium-high density of occupation (25% to 40%).

Estimation of Transition Rates

As previously mentioned, eight categories of land use zones were defined for the modelling experiment: residential, commercial, industrial, institutional, services, mixed use, leisure/recreation, and non-urban use. The mixed use zone basically comprises the residential, commercial, and services uses. The leisure/recreation zone includes parks, the city zoo and other public green areas. For the specific case study town in question – Bauru – in the period 1979-1988, five types of land use change were detected (Table 4).

In order to calculate land use transition rates for the period 1979-1988, the initial and final land use maps were converted to raster files with extension TIFF and resolution 100 x 100 (m), and then exported to the IDRISI Geographic Information System. In IDRISI, a cross-tabulation operation was made between both land use maps (See Figure

1) so as to generate transition percentages for the five existent types of land use change (Table 5).

TABLE 4 - IDENTIFIED TYPES OF LAND USE CHANGE FOR THE CITY OF BAURU, IN THE PERIOD 1979-1988, AND RESPECTIVE CODES

Code	Type of Land Use Change
NU_RES	*Non-Urban to Residential*
NU_IND	*Non-Urban to Industrial*
NU_SERV	*Non-Urban to Services*
RES_SERV	*Residential to Services*
RES_MIX	*Residential to Mixed Use*

TABLE 5 – LAND USE TRANSITION RATES ESTIMATED FOR THE CITY OF BAURU, IN THE PERIOD 1979-1988

	Non-Urban	Resident.	Commercial	Industrial	Institutional	Services	Mixed Use	Leis./Recr.
Non-Urban	0.9171331	0.0697519	0	0.0095301	0	0.0035848	0	0
Resident.	0	0.9379833	0	0	0	0.0597520	0.0022647	0
Commercial	0	0	1.0000000	0	0	0	0	0
Industrial	0	0	0	1.0000000	0	0	0	0
Institutional	0	0	0	0	1.0000000	0	0	0
Services	0	0	0	0	0	1.0000000	0	0
Mixed Use	0	0	0	0	0	0	1.0000000	0
Leis./Recr.	0	0	0	0	0	0	0	1.0000000

Due to the stochastic structure of the DINAMICA transition algorithms, the envisaged transition rates established in the above table are not always reached.

For the estimation of land use percentages in the case of modelling land use change forecasts through DINAMICA, the Markov chain is to be employed. According to Hobbs (1983), the Markov chain is a mathematical model designed to describe a certain type of process that moves in a sequence of steps through a set of states, whose formula is defined as:

$$\boxed{\Pi(t+1) = P.\,\Pi(t)} \quad , \qquad (8)$$

where $\Pi(t)$ is a column vector, with n elements, that represents the system condition in a certain time t (e.g. area percentages for each n_i land use category or state); $\Pi(t+1)$ is the vector representing the occupation of n states in a given future time $t+1$; and P is the transition probabilities matrix or the table for land use transition rates.

An important constraint of the Markov model lies on the fact that, in principle, it supposes that transition probabilities do not change over time (stationary process). Moreover, given its stochastic nature, the Markov chain masks the causative variables. It is not an explanatory model, and is thus of no use in understanding the causes and driving factors of land use transition processes. On the other hand, the Markov chain analysis has the great advantage of presenting a mathematical and operational simplicity. Simple trend projection involves no more than matrix multiplication, and the only data requirement is for current land use information (JRC and ESA, 1994).

Reckoning of the Cells Land Use Transition Probabilities

As previously said, the "weights of evidence" statistical method, employed in the calculation of the cells transition probabilities, is based on the "Bayes theorem of conditional probability". Basically, this theorem concerns the favourability to detect a certain event, which can be in the current case a given category of land use change (e.g. non-urban use to residential use), provided that an evidence (e.g. water supply area), also called explaining variable, has already happened (Figure 4). The evidences or explaining variables in the experiment presented in this paper mainly refer to the technical and social infrastructure of the case study town, Bauru.

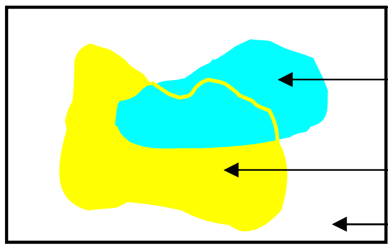


Fig. 4 – Diagram to illustrate weights of evidence calculations.

The favourability to find the event (change from non-urban to residential use) *R* given the presence of the evidence (water supply) *S* can be expressed by:

$$\boxed{P\{R/S\} = \frac{P\{R \cap S\}}{P\{S\}}} \qquad (9)$$

where $P\{R/S\}$ is the conditional probability of occurring the event *R* given the presence of the explaining variable *S*. But, $P\{R \cap S\}$ is equal to the proportion of total area occupied by *R* and *S* together. Supposing *N* is the counting of map cells (area of an event or an evidence), then the above formula can be rewritten as:

$$\boxed{P\{R/S\} = \frac{N\{R \cap S\}}{N\{S\}}} \qquad (10)$$

In order to obtain an expression relating the *posterior* probability of the event *R* in terms of the *prior* probability and a multiplication factor, we note that the conditional probability of being on the explaining variable map *S*, given the presence of the event *R* is defined as:

$$\boxed{P\{S/R\} = \frac{P\{S \cap R\}}{P\{R\}}} \qquad (11)$$

Since $P\{S \cap R\}$ is the same as $P\{R \cap S\}$, Equations *9* and *11* can be combined to solve for $P\{R/S\}$, satisfying the relationship:

$$\boxed{P\{R/S\} = \frac{P\{R\} \,.\, P\{S/R\}}{P\{S\}}} \tag{12}$$

A similar expression can be derived for the posterior probability of the event *R* occurring given the absence of the evidence *S*. Thus,

$$\boxed{P\{R/\overline{S}\} = \frac{P\{R\} \,.\, P\{\overline{S}/R\}}{P\{\overline{S}\}}} \tag{13}$$

All the equations above can be expressed in an *odds* form. ***Odds*** are defined as a ratio of the probability that an event will occur to the probability that it will not occur. The weights of evidence method uses the natural logarithm of *odds*, known as *log odds* or ***logits***. To clarify this approach, Equation *12* will be converted to *odds*. For this end, both sides will be divided by $P\{\overline{R}/S\}$, leading to:

$$\boxed{\frac{P\{R/S\}}{P\{\overline{R}/S\}} = \frac{P\{R\} \,.\, P\{S/R\}}{P\{\overline{R}/S\} \,.\, P\{S\}}} \tag{14}$$

But from the definitions of conditional probability:

$$\boxed{P\{\overline{R}/S\} = \frac{P\{\overline{R} \cap S\}}{P\{S\}} = \frac{P\{S/\overline{R}\} \,.\, P\{\overline{R}\}}{P\{S\}}} \tag{15}$$

Substituting Equation *15* in Equation *14*, yields the following:

$$\frac{P\{R/S\}}{P\{\bar{R}/S\}} = \frac{P\{R\}}{P\{\bar{R}\}} \cdot \frac{P\{S\}}{P\{S\}} \cdot \frac{P\{S/R\}}{P\{S/\bar{R}\}} \qquad (16)$$

Substituting odds into Equation *15* and cancelling leads to the desired expression:

$$O\{R/S\} = O\{R\} \cdot \frac{P\{S/R\}}{P\{S/\bar{R}\}} \qquad (17)$$

where $O\{R/S\}$ is the conditional (*posterior*) odds of R given S, $O\{R\}$ is the *prior odds* of R and $P\{S/R\}/P\{S/\bar{R}\}$ is known as the *sufficiency ratio* (*LS*). In weights of evidence, the natural logarithm of both sides of Equation *17* are taken, and $log_e\ LS$ is the *positive weight of evidence* W^+, which is calculated from the data. Then:

$$logit\{R/S\} = logit\{R\} + W^+ \qquad (18)$$

Similar algebraic manipulations lead to the derivation of an *odds* expression for the conditional probability of R given the absence of the evidence S, with the result being:

$$O\{R/\bar{S}\} = O\{R\} \cdot \frac{P\{\bar{S}/R\}}{P\{\bar{S}/\bar{R}\}} \qquad (19)$$

The term $P\{\bar{S}/R\}\ /\ P\{\bar{S}/\bar{R}\}$ is called the *necessity ratio* (*LN*). In weights of evidence, the *negative weight of evidence* W^- is the natural logarithm of LN, or $log_e\ LN$. Thus in *logit* form, Equation *19* is:

$$logit\{R/\bar{S}\} = logit\{R\} + W^- \qquad (20)$$

LS and *LN* are also called *likelihood ratios*. When events and evidences are *positively correlated*, the value of *LS* is greater than 1, whereas *LN* is in the range [0,1]. However, if an evidence is *negatively correlated* with the events, *LN* would be greater than 1 and *LS* would be in the range [0,1]. If the evidence is uncorrelated with the events, then $LS=LN=1$, and the *posterior* probability equals the *prior* probability, and the probability of an event would be unaffected by the presence or absence of a certain evidence.

Similarly, W^+ is positive, and W^- is negative, due to the positive correlation between the evidences and the events. Conversely W^+ would be negative and W^- positive for the case where a very limited part of the event occur on the evidence area than would be expected due to chance. If the events are independent of whether the evidence is present or not, then $W^+ = W^- = 0$, and the *posterior* = the *prior*, as above (Bonham-Carter, 1994).

A useful and commonly used measure of spatial association between an evidence and a given event is the contrast *C*, whose formula is given by:

$$\boxed{C = W^+ - W^-} \qquad (21)$$

The question of determining whether the magnitude of the contrast is large enough to be statistically significant can be tested by its variance, estimated from the following expression (Goodacre et al., 1993):

$$\boxed{S^2(C) = \frac{1}{N\{S \cap R\}} + \frac{1}{N\{S \cap \overline{R}\}} + \frac{1}{N\{\overline{S} \cap R\}} + \frac{1}{N\{\overline{S} \cap \overline{R}\}}} \qquad (22)$$

In the weights of evidence method, there is a specific way for calculating probability ratios (*odds*) in the case of *n* maps of evidence (M_i). The general expression for combining $i=1,2,\ldots,n$ maps is either:

$$O\,\{R\,/M_1 \cap M_2 \cap M_3 \cap \ldots \cap M_n\} = O\,\{R\}\,.\prod_{i=1}^{n} LS_i \qquad (23)$$

for the *likelihood ratios* or:

$$logit\,\{R\,/M_1 \cap M_2 \cap M_3 \cap \ldots \cap M_n\} = logit\,\{R\} + \sum_{i=1}^{n} W^{+}{}_{i} \qquad (24)$$

for the weights. In these general formulas, the *LS* becomes *LN*, and W^+ becomes W^-, if the *i-th* evidence is absent instead of present. Where data is missing for a particular map layer in some locations, the likelihood ratio is set to 1, or the weight is set to 0. These two last equations are the computing formulae for combining a set of evidence maps with the Bayes model. The principal advantages of the "weights of evidence" method are:

- the method is objective, and avoids the subjective choice of weighting factors;
- multiple evidences maps can be combined with a model that is straightforward to program with a modelling language;
- conversion of multistate evidences maps to binary maps, where each distance range is treated as a present/absent evidence;
- input maps with missing data (incomplete coverage) can be accommodated in the model;
- uncertainty a) due to variances of weights, and b) due to missing data can be modelled to show the effect on the *posterior* probability.

Some of the disadvantages of the weights of evidence modelling are:

- the tests for conditional independence between pairs of evidence maps involve contingency table calculations, generally carried out outside the GIS environment, using data files generated by the GIS;
- weights of evidence, in common with other data-driven methods, is only applicable in regions where the event (also called response variable) is fairly well known (Bonham-Carter, 1994).

For the particular case of the DINAMICA simulation model, adopted for the modelling experiment being considered, the cells transition probabilities are calculated in a different form from Equations *23* and *24* previously presented. Its transition probability expression is given by:

$$\boxed{P = \frac{e^{\sum_{i=1}^{n} W^{+}}}{1 + e^{\sum_{i=1}^{n} W^{+}}}} \qquad (25)$$

This formula shows a clear similarity with the one employed for the calculation of probability in the logistic regression method (also known as logistic function). In the above case, the sum of the positive weights of evidence (W^+) corresponds to the product of the linear regression coefficients by the independent variables adopted in the regression analysis.

The first step in the very process of calculating the cells transition probabilities using DINAMICA is to obtain a cross-tabulation map (Figure 5) between the initial and final land use maps elaborated for the city of Bauru, respectively for the years 1979 and 1988, both previously presented in Figure1.

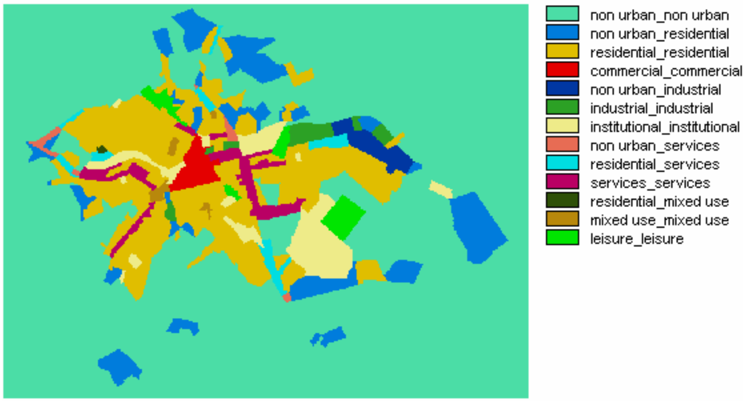


Fig. 5 – Cross-tabulation map between the initial (1979) and final (1988) land use maps elaborated for the city of Bauru. Types of land use permanence as well as transition are listed on the legend.

In IDRISI, the land use cross-tabulation map of Bauru (1979-1988) was used to generate land use transition maps for each of the five possible types of land use change presented in Table 4. This was done through reclassification tables ("edit" command), on which the following rules were observed:

- all raster values corresponding to classes of land use permanence or transition whose initial land use was different from the initial land use category in the considered type of land use change were assigned value 0 (***black colour***). This reclassification to value 0 is automatic for raster values not included in the "edit" table;
- all raster values corresponding to classes of land use transition whose initial and final land use categories were equal to the initial and final categories of the land use change at issue were assigned value 2 (***blue colour***);
- all other remaining classes of land use permanence or transition were assigned value 1 (***green colour***).

Examples of "edit" tables are shown on Figure 6, and an example of a land use transition map is presented on Figure 7.

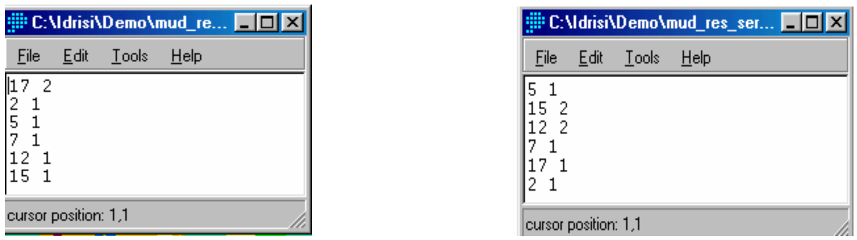
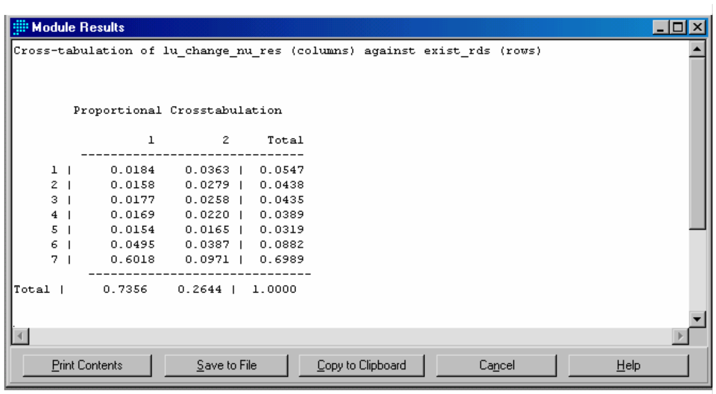


Fig. 6 – Example of an edit table for the "non-urban_residential" land use change on the left, and for the "residential_services" change on the right.

Fig. 7 – Non-urban_residential land use transition map for Bauru in the period 1979-1988.

Once all the possible types of land use transition maps were elaborated (nu_res; nu_ind; nu_serv; res_serv; res_mix), they were then subjected to partial cross-tabulations with selected explaining variables (evidences) maps, listed on Table 1, according to an apparent interdependence between a certain type of land use transition and a given explaining variable. The evidences maps, pre-processed in the SPRING Geographic Information System, were in the same manner as the initial and final land use maps converted to raster files with extension TIFF and resolution 100 x 100 (m), and then exported to IDRISI.

The partial cross-tabulations disregard the raster values 0 (black colour) in the land use transition maps and are accomplished through the "ermatt" command of IDRISI. An example of these partial cross-tabulations outcome is shown below (Figure 8).



Cross-tabulation of lu_change_nu_res (columns) against exist_rds (rows)

Proportional Crosstabulation

	1	2	Total
1	0.0184	0.0363	0.0547
2	0.0158	0.0279	0.0438
3	0.0177	0.0258	0.0435
4	0.0169	0.0220	0.0389
5	0.0154	0.0165	0.0319
6	0.0495	0.0387	0.0882
7	0.6018	0.0971	0.6989
Total	0.7356	0.2644	1.0000

Fig. 8 – Example of a partial cross-tabulation table between the land use transition map "non-urban – residential" (nu_res) and the map of distances to main existent roads (exist_rds). The numbers 1 and 2 in the first line are respectively related to classes green and blue of the land use transition map "nu_res", and the numbers 1 to 7 on the left column refer to labels for distance ranges from the main existent roads in Bauru in 1979.

The numerical values of cells proportions existing in the absence/presence of a binary evidence (e.g. water supply) or in the different ranges of distances maps and found to be overlying on either class 1 (green colour) or 2 (blue colour) of the land use transition maps are (for each cross-tabulation table) selectively transferred to EXCELL files specially created for the calculation of the weights of evidence (See Equations *17* to *20*).

Using the values for the positive weights of evidence W^+ (Table 6) concerning the several evidences maps employed in the analysis of each category of land use change, the DINAMICA simulation model will then calculate the cells transition probabilities (See Equation *25*) for the five types of land use transition.

TABLE 6 – POSITIVE WEIGHTS ESTIMATED FOR THE LAND USE TRANSITIONS IN RELATION TO THEIR RESPECTIVE EVIDENCES FOR THE CITY OF BAURU IN THE PERIOD 1979-1988.

LAND USE TRANSITION	*TYPE OF EVIDENCE (EXPLAINING VARIABLE)* / *VALUES FOR THE POSITIVE WEIGHTS OF EVIDENCE*						
RES_SERV	***AREA SERVED BY WATER SUPPLY***						
	PRESENT	***ABSENT***					
	-0.661126	*0.2883432*					
RES_MIX	***MEDIUM-HIGH DENSITY OF OCCUPATION (25% TO 40%)***						
	PRESENT	***ABSENT***					
	0.6452582	*-0.063522*					
RES_MIX	***SOCIAL HOUSING SETTLEMENTS***						
	PRESENT	***ABSENT***					
	2.4678421	*-0.32148*					
NU_RES	***DISTANCES TO RANGES OF COMMERCIAL ACTIVITIES CONCENTRATION, (KERNEL ESTIMATOR)***						
	0 -500 (m)	***500-1000 (m)***	***1000-1500 (m)***	***1500-10000 (m)***	***10000-30000 (m)***	***> 30000 (m)***	
	3.749359	*2.1061142*	*1.864181*	*0.4914826*	*-0.32329*	*0*	
NU_SERV	***0 -500 (m)***	***500-1000 (m)***	***1000-1500 (m)***	***1500-10000 (m)***	***10000-30000 (m)***	***> 30000 (m)***	
	3.4118828	*4.4690133*	*2.9118401*	*0.8777834*	*0*	*0*	
NU_IND	***DISTANCES TO INDUSTRIAL ZONES***						
	0 -500 (m)	***500-1000 (m)***	***1000-1500 (m)***	***1500-2000 (m)***	***2000-5000 (m)***	***5000-10000 (m)***	***> 10000 (m)***
	3.8624294	*4.0157265*	*3.792217*	*3.4523225*	*1.7633056*	*0*	*0*
NU_SERV	***DISTANCES TO RESIDENTIAL ZONES***						
	0 -500 (m)	***500-1000 (m)***	***1000-2000 (m)***	***2000-5000 (m)***	***5000-10000 (m)***	***> 10000 (m)***	
	2.1439745	*1.5228203*	*0.6209727*	*-0.06484*	*0*	*0*	

CONT. OF TABLE 6 – POSITIVE WEIGHTS ESTIMATED FOR THE LAND USE TRANSITIONS IN RELATION TO THEIR RESPECTIVE EVIDENCES FOR THE CITY OF BAURU IN THE PERIOD 1979-1988.

LAND USE TRANSITION	TYPE OF EVIDENCE (EXPLAINING VARIABLE) / VALUES FOR THE POSITIVE WEIGHTS OF EVIDENCE						
	DISTANCES TO PERIPHERAL RESIDENTIAL SETTLEMENTS, ISOLATED FROM THE URBAN CONCENTRATION						
NU_RES	0 -500 (m)	500-1000 (m)	1000-2000 (m)	2000-5000 (m)	5000-10000 (m)	> 10000 (m)	
	1.9675468	1.6151719	1.3924275	0.892197	-0.625689	-0.469075	
	DISTANCES TO SOCIAL INFRASTRUCTURE (INSTITUTIONAL) EQUIPMENTS, ISOLATED FROM THE URBAN CONCENTRATION						
NU_RES	0 -500 (m)	500-1000 (m)	1000-3000 (m)	3000-8000 (m)	8000-15000 (m)	> 15000 (m)	
	0.0034705	0.6003806	1.253931	0.7274862	-0.358902	-0.08934	
	DISTANCES TO MAIN EXISTENT ROADS						
NU_RES	0 -250 (m)	250-500 (m)	500-750 (m)	750-1000 (m)	1000-1250 (m)	1250-2000 (m)	>2000 (m)
	0.2305106	0.3195579	0.3527951	0.5097521	0.4432425	0.1961828	-0.084568
	DISTANCES TO THE SERVICES AND INDUSTRIAL AXES						
NU_IND	0 -250 (m)	250-500 (m)	500-750 (m)	750-1000 (m)	1000-1250 (m)	1250-2000 (m)	>2000 (m)
	2.722019	2.7998156	2.675631	2.6245036	2.5254157	1.7274357	-3.832114
NU_SERV	0 -250 (m)	250-500 (m)	500-750 (m)	750-1000 (m)	1000-1250 (m)	1250-2000 (m)	>2000 (m)
	3.5077549	3.3209421	2.9174182	1.8686185	0.450248	0	0
RES_SERV	0 -250 (m)	250-500 (m)	500-750 (m)	750-1000 (m)	1000-1250 (m)	1250-2000 (m)	>2000 (m)
	2.7801176	1.94798	1.4614056	0.8879287	-0.297012	-1.411855	-3.284054
	DISTANCES TO PLANNED ROADS						
RES_MIX	0 -250 (m)	250-500 (m)	500-750 (m)	750-1000 (m)	1000-1250 (m)	1250-2000 (m)	>2000 (m)
	3.5059276	1.8631255	0	0	0	0	0
	DISTANCES TO PERIPHERAL ROADS, WHICH PASS THROUGH NON-OCCUPIED AREAS						
NU_RES	0 -250 (m)	250-500 (m)	500-750 (m)	750-1000 (m)	1000-1500 (m)	1500-2500 (m)	>2500 (m)
	2.3770299	2.268923	2.0682195	1.9838132	1.4440168	0.8572809	-0.126596
RES_MIX	0 -250 (m)	250-500 (m)	500-750 (m)	750-1000 (m)	1000-1500 (m)	1500-2500 (m)	>2500 (m)
	1.7750028	1.6519953	1.848404	0.9032325	0	0	0

By means of the cells transition probabilities, DINAMICA will generate the respective transition probabilities maps (Figures 9 to 13) for each of the five types of land use change existing in Bauru from 1979 to 1988. These maps are seen in ERMAPPER, a Geographical Information System employed by DINAMICA for visualisation purposes.

It is worth mentioning how these probabilities maps detect considerably well the transition areas (blue colour) in the corresponding land use transition maps, for all the reddish regions in the probabilities maps relate to the very areas owning the highest transition probabilities rates.

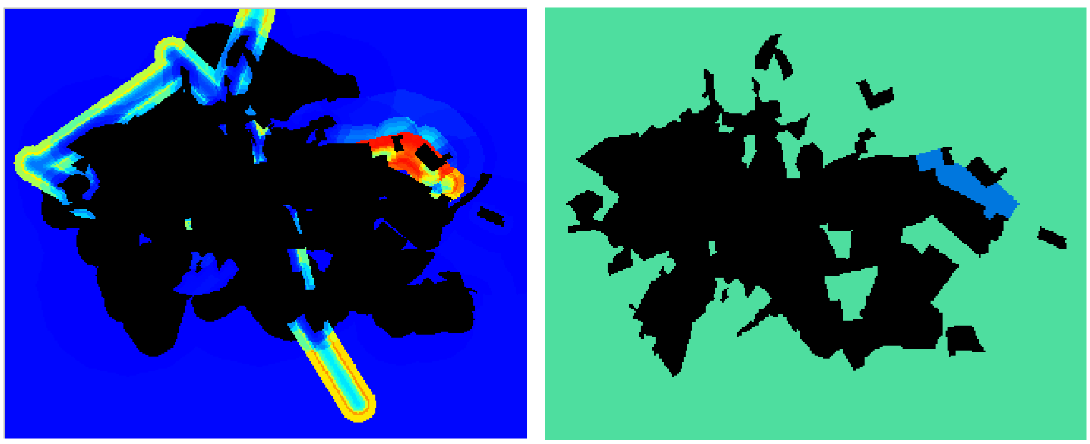


Fig. 9 - Map of cells transition probabilities, on the left, and map of land use transition "non-urban – industrial" (nu_ind), on the right.

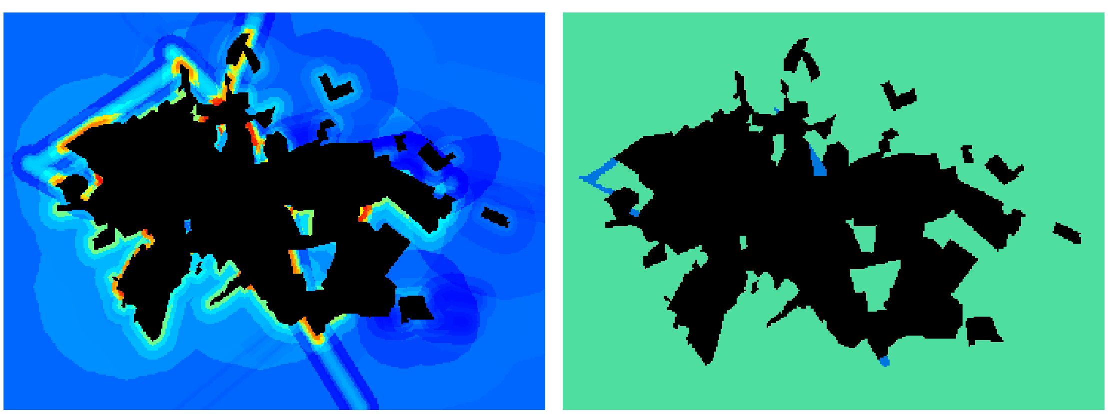


Fig. 10 - Map of cells transition probabilities, on the left, and map of land use transition "non-urban – services" (nu_serv), on the right.

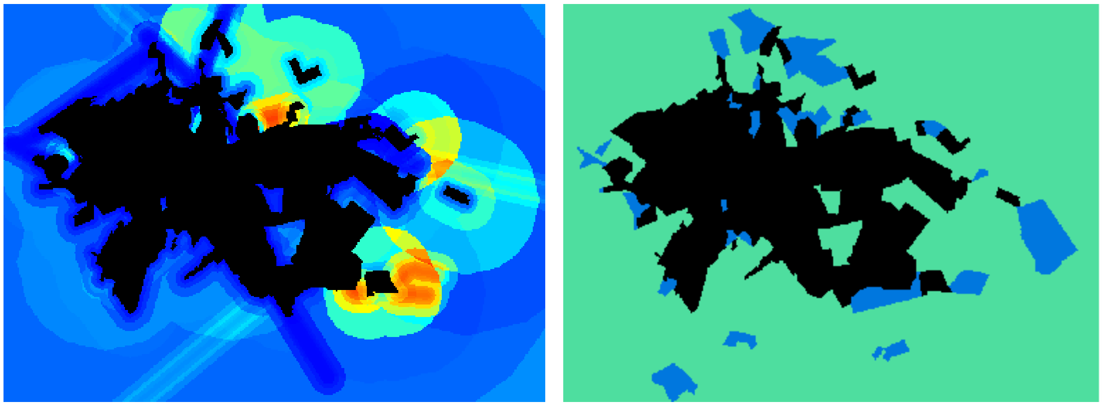


Fig. 11 - Map of cells transition probabilities, on the left, and map of land use transition "non-urban – residential" (nu_res), on the right.

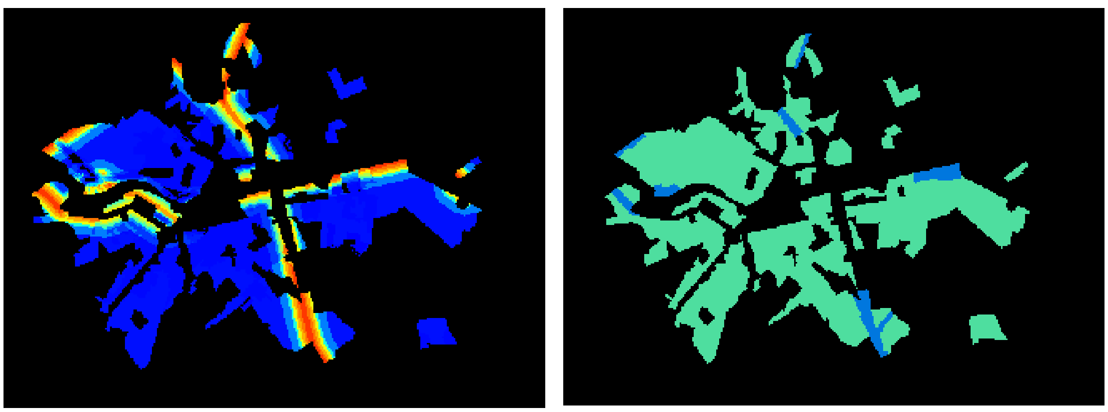


Fig. 12 - Map of cells transition probabilities, on the left, and map of land use transition "residential – services" (res_serv), on the right.

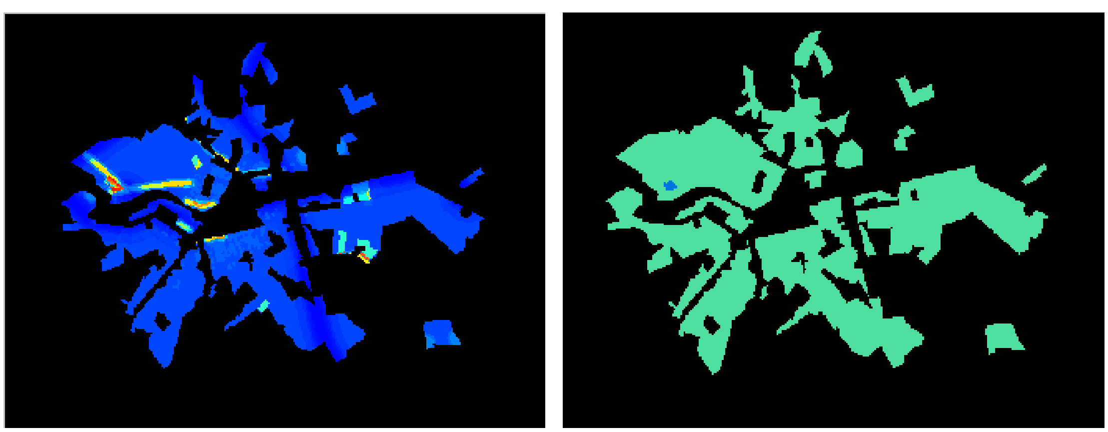


Fig. 13 - Map of cells transition probabilities, on the left, and map of land use transition "residential– mixed use" (res_mix), on the right.

Model Calibration

For the calibration of the model designed to emulate urban land use transition for the town of Bauru in the period 1979 – 1988, empirical procedures were adopted. They basically concern the visual comparative analysis, for each type of land use change, amongst the general trends of preliminary simulation results, the hints provided by both the transition probabilities map and the land use transition map, and the guideline information contained in the simultaneous overlay of different explaining variables maps upon the final land use map in vector format (Figure 14).

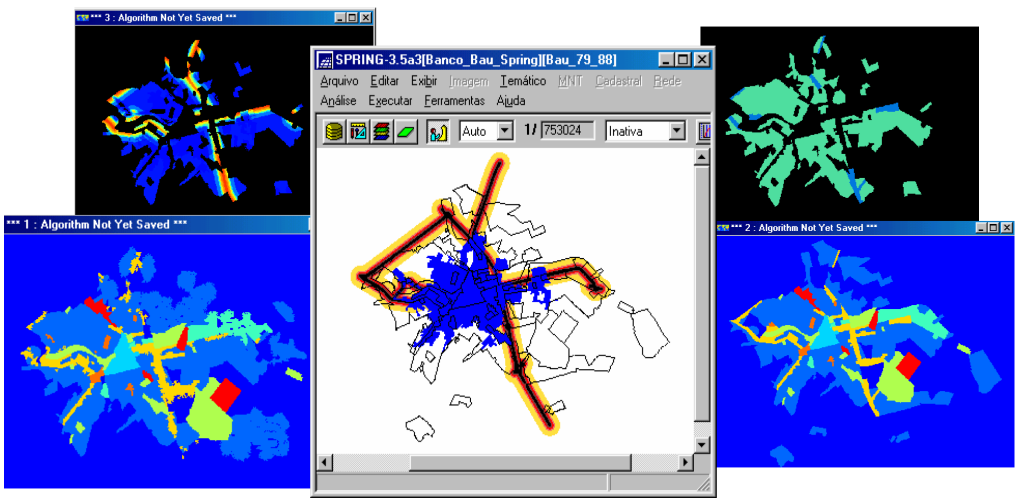


Fig. 14 – Example of visual comparative analysis for the model empirical calibration in relation to the "residential-services (res_serv)" land use change. The transition probabilities map can be seen on the upper left corner; the land use transition map, on the upper right; a preliminary simulation result, on the lower left corner; the real final land use map, on the lower right corner; and a simultaneous overlay of the water supply map and the services axes distances map upon the borders of the final land use map (1988) is seen in the centre.

The model calibration, on the other hand, is as well accomplished by the analysis of scatter plots relating subcategories of evidences (distances ranges), whenever they are available, with their respective positive weights of evidence. In a general manner, when the plots present a good fit of trendlines (which can assume different orders and types), i.e. when the lines do not demand very complex models for adjustment, the evidences to which they are associated are highly prone to be included in the model (Figure 15).

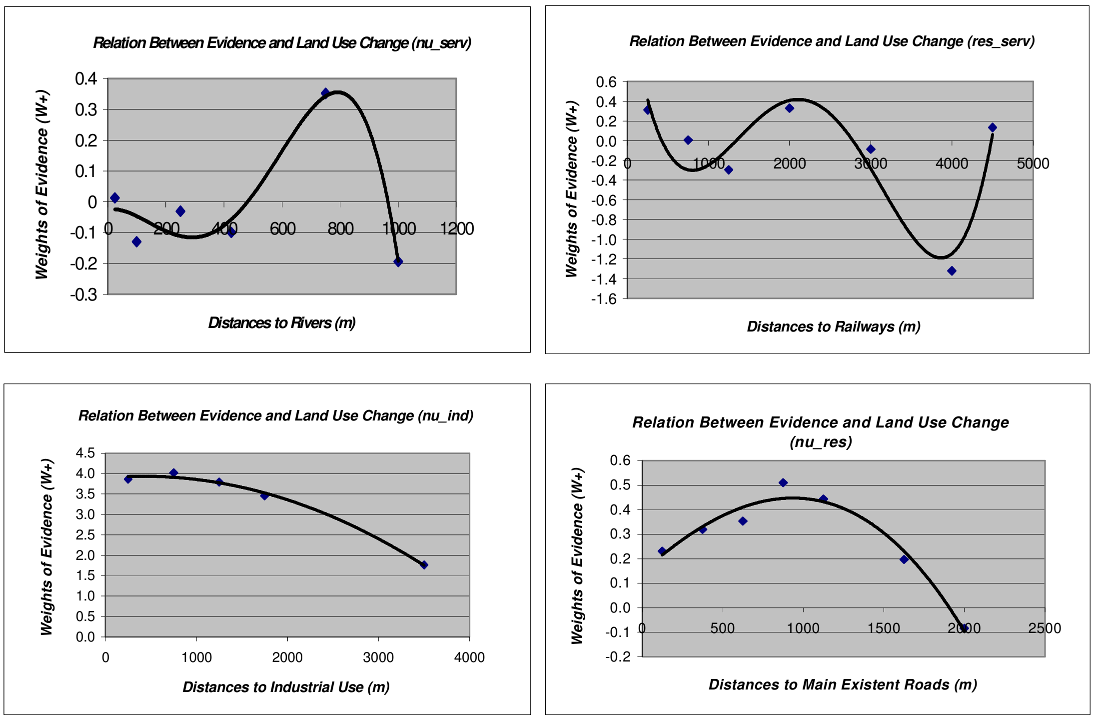


Fig. 15 – Examples of scatter plots and respective trendlines for the relations between subcategories of evidences (*X axis*) and their corresponding positive weights of evidence (*Y axis*), considering different types of land use change. The upper plots show cases of poor fit, and hence, of evidences exclusion. On the contrary, the lower plots present a good adjustment of trendlines, what implies the high probability of inclusion of such evidences in the urban land use dynamics model at issue. The final decision towards the inclusion or exclusion of a given evidence will always rely upon a broad judgement, in which the environmental importance of the evidence and its coherence concerning the phenomenon (land use transition) being modelled are analysed.

All the above-mentioned calibration procedures led to the final selection of sets of evidences (explaining variables) maps for each of the existing five types of land use change in the city of Bauru from 1979 to 1988 (Table 7).

TABLE 7 – SETS OF EVIDENCES MAPS FOR EACH TYPE OF LAND USE TRANSITION EXISTING IN THE CITY OF BAURU FROM 1979 TO 1988.

Maps	*Evidences Codes*	*Nu_Res*	*Nu_Ind*	*Nu_Serv*	*Res_Serv*	*Res_Mix*
	water				✓	
	mh_dens					✓
	soc_hous					✓
	com_kern	✓		✓		
	dist_ind		✓			
	dist_res			✓		
	per_res	✓				
	dist_inst	✓				
	exist_rds	✓				
	serv_axes		✓	✓	✓	
	plan_rds					✓
	per_rds	✓				✓

All the evidences maps presented in Table 7 (and whose codes meanings are listed on Table 1) were exported from IDRISI as raster files with extension TIFF. These maps were altogether imported in ERMAPPER as a single file, i.e. a unique data set, owning extensions ALG (regarding the visualisation algorithm) and ERS (relating to the raster data themselves), which then have been employed by DINAMICA in the generation of simulations.

In a general way, the final sets of evidences maps used in the land use transition simulation model for the city of Bauru in the period 1979-1988 can be classified according to the following analysis categories (Table 8):

TABLE 8 – CLASSIFICATION OF THE FINAL SETS OF EVIDENCES MAPS USED IN THE URBAN LAND USE DYNAMICS MODEL

ANALYSIS CATEGORIES	*SELECTED EVIDENCES*
Infrastructure	*Area served by water supply (water).*
Social Infrastructure	*Proximity to concentrations of commercial activities (com_kern) as well as to social infrastructure equipments (dist_inst).*
Accessibility	*Distances to main roads (exist_rds), to the services and industrial axes (serv_axes), to planned roads (plan_rds) and to peripheral roads (per_rds).*
Neighbourhood Influences	*Proximity to classes of residential use (soc_hous; dist_res; per_res) or of industrial use (dist_ind); occurrence of medium-high density of occupation (mh_dens).*

3. Results and Discussion

Upon basis of the carried out calibration process, it becomes evident that the probability of certain non-urban areas in the city of Bauru to shelter residential settlements ("nu_res" land use transition) largely depends on the previous existence of this type of settlements in their surroundings, on the greater proximity of these areas to commercial activities clusters as well as on the available accessibility to such areas.

As to the transition of non-urban areas to industrial use (nu_ind), there are two great driving forces: the nearness of such areas to the previously existent industrial use and the availability of road access. This can be explained by the fact that in the industrial production process, the output of certain industries represent the input of other ones, what raises the need of rationalisation and optimisation of costs by the clustering of plants interrelated in the same productive chain. Furthermore, plots in the vicinities of industrial areas are often prone to be devaluated for other uses, what makes them rather competitive for the industrial use.

Regarding the transition of non-urban areas to services use (nu_serv), three major factors are crucial: the proximity of these areas to clusters of commercial activities, their closeness to areas of residential use, and last but not least, their strategic location in relation to the N-S / E-W services axes of Bauru. The first factor accounts for the suppliers market (and in some cases also consumers market) of services; the second factor represents the consumers market itself; and the third and last factor corresponds to the accessibility for both markets related to the services use.

The transition "residential to services use" (res_serv) supposes the insertion of services into previously consolidated urban areas. In this way, since this transition type already takes place amid the suppliers and consumers markets, it will solely prioritise the strategic location in relation to the N-S / E-W services axes of Bauru, besides of course, the existence of water supply, which in the specific case of Bauru do not correspond to the whole urbanised area.

Finally, the last type of land use transition concerns the shift from residential use to mixed use (res_mix). The mixed use zones, which actually play the role of urban subcentres, constitute a sort of commercial centres consolidation, which in a later stage start to also attract services and social infrastructure equipments besides commercial activities themselves. Therefore, new mixed use zones arise in more peripheral areas, where a greater occupational gathering is at the same time assured. Thus, the decisive factors for this last type of land use change are:

- existence of medium-high density of occupation (higher density values only occur in the central commercial zone of the town or in the immediacies of already existent mixed use zones);
- presence or proximity of social housing settlements (for they shelter the greatest occupational densities in more peripheral areas, and hence, greater consumers markets);
- nearness to planned or peripheral roads, since new mixed use zones arise in farther areas of the town.

After the calibration of evidences maps sets is accomplished, a new calibration process concerning the script parameters of the DINAMICA simulation model takes place. Such parameters refer to the number of iterations (runs), proportion of cells transition by contiguity ("*expander*" operator) and by nucleation ("*patcher*" operator), average size and variance of patches to be generated either by the *expander* or *patcher* operators, etc.

The *expander* is an algorithm of the DINAMICA model which realises transitions from a state *i* to a state *j* only in the adjacent vicinities of cells with state *j*. Its procedures routine is the following:

- identification of frontier cells of class *j*;
- increase of their probability value according to the number of neighbours of class *j* found in a 3 x 3 window, i.e.:

$$\boxed{P_{final} = \frac{\textit{number of neighbours of class j}}{\textit{number of possible neighbours}} * P_{initial}} \qquad (26)$$

  where the number of possible neighbours is 8 (9 – 1);

- selection of a random number between 0 and 255. If the selected number is smaller than the cell transition probability (also ranging from 0 to 255), the cell is specially reserved to take part in a second selection process, in which the state transitions actually take place. On the other hand, if the randomly selected number is greater than the cell transition probability, the cell is discarded. In this first selection process, a total of ten times as much the number of cells required for the transition to state *j* (estimated either by a cross-tabulation table or the Markov model) are reserved for the second stage of selection;

- in the second stage, a new random number between 0 and 255 is selected. If such a number is smaller than the cell transition probability, the cell changes its state to *j*, otherwise it remains in its original state.

The *patcher* operator, on its turn, is an algorithm of the DINAMICA model that realises transitions from a state *i* to a state *j* only in the adjacent vicinities of cells with state other than *j*. Its procedures routine is the following:

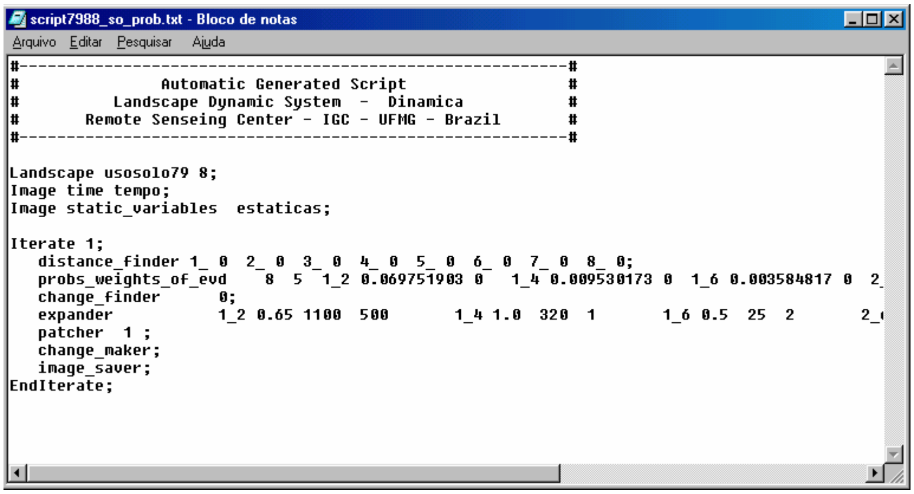
- random selection of a number between 0 and 255. If the selected number is smaller than the cell transition probability (also ranging from 0 to 255), the cell changes its state to *j*, otherwise it remains in its original state (Soares-Filho, 1998; Soares-Filho et al., 2001).

The script parameters of DINAMICA that produced the best simulation results are presented in Table 9.

TABLE 9 – FINAL PARAMETERS OF THE DINAMICA MODEL SCRIPT

LAND USE TRANSITION	Average Size of Patches	Variance of Patches Size	Proportion of "Expander"	Proportion of "Patcher"	Number of Iterations
NU_RES	1100	500	0.65	0.35	5
NU_IND	320	1	1.00	0	5
NU_SERV	25	2	0.50	0.50	5
RES_SERV	25	2	0.10	0.90	5
RES_MIX	35	2	0	1.00	5

An example of the final DINAMICA script window is shown below (Figure 16).


script7988_so_prob.txt - Bloco de notas
Arquivo Editar Pesquisar Ajuda

```
#-----------------------------------------------------------#
#                 Automatic Generated Script                #
#            Landscape Dynamic System  -  Dinamica          #
#        Remote Senseing Center - IGC - UFMG - Brazil       #
#-----------------------------------------------------------#

Landscape usosolo79 8;
Image time tempo;
Image static_variables  estaticas;

Iterate 1;
   distance_finder 1_ 0  2_ 0  3_ 0  4_ 0  5_ 0  6_ 0  7_ 0  8_ 0;
   probs_weights_of_evd    8  5  1_2 0.069751903 0   1_4 0.009530173 0  1_6 0.003584817 0  2_
   change_finder      0;
   expander           1_2 0.65 1100  500      1_4 1.0  320  1       1_6 0.5  25  2       2_
   patcher  1 ;
   change_maker;
   image_saver;
EndIterate;
```

Fig. 16 – Final DINAMICA script window containing the parameters that generated the best simulations results.

Due to the randomness of the DINAMICA transition algorithms, even though the same sets of evidences maps for each type of land use transition and the same script parameters are kept in different runs, distinct simulations results will be produced after each run of the model. In this way, the best urban land use simulation results for the city of Bauru in the period 1979–1988 are presented in Figure 17.

The *patcher* algorithm proved to be of great suitability for the modelling of residential settlements disconnected from the main urban agglomeration. Nevertheless, the shapes of these settlements in the modelling results do not strictly coincide to those observed in reality. This happens because these contours are associated with the real state properties limits. Since legal actions for the merging or split of plots may occur at any time and drastically alter their form, such boundaries can be regarded as highly unstable factors, and thus, inappropriate for modelling.

The services corridors, in light brown, were well modelled in all simulations. The industrial use zone, in light green, was considerably well detected in all of the three

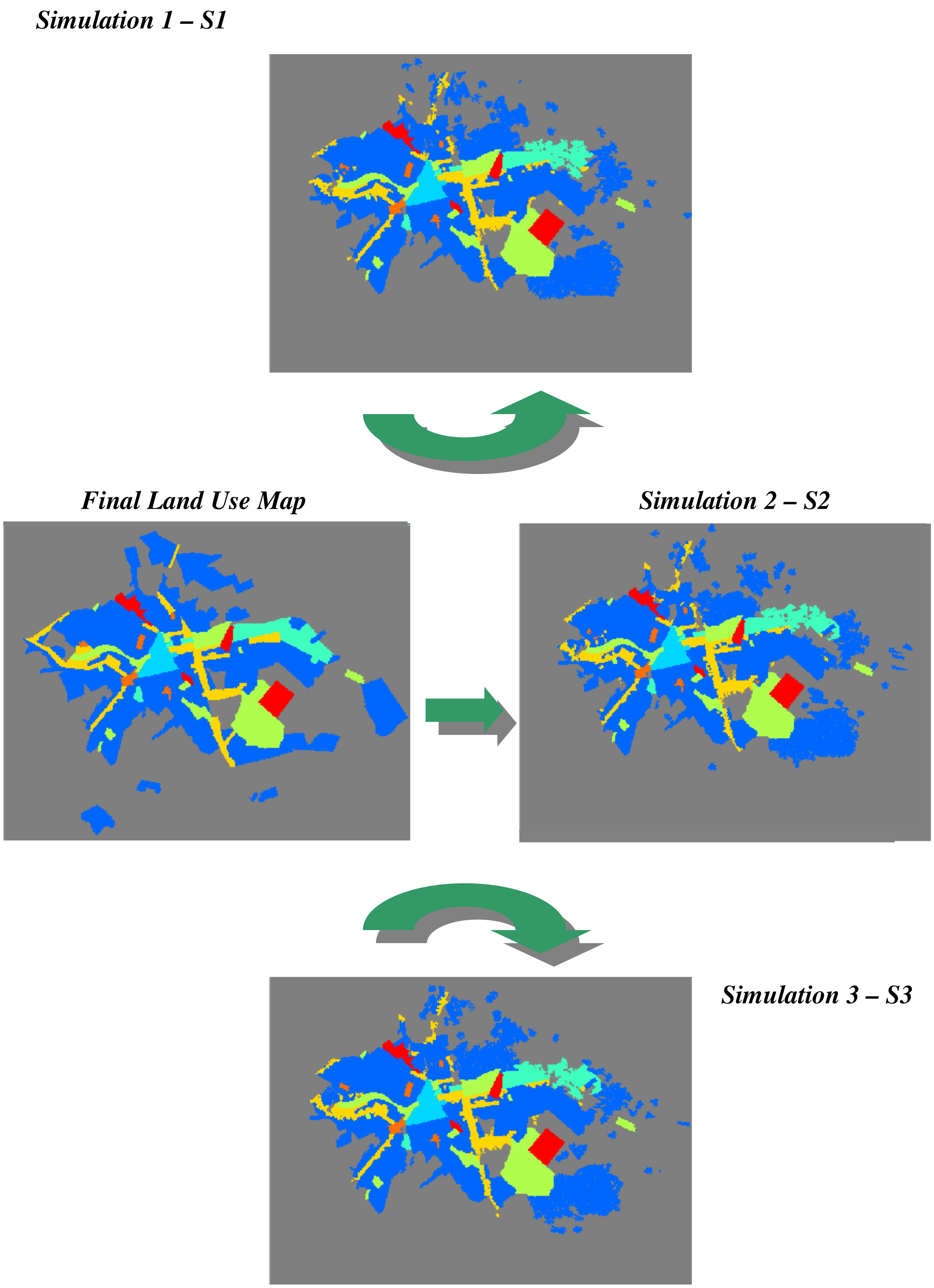


Fig. 17 – Bauru final land use map and simulations results for the period 1979-1988.

simulations results, specially in S2 and S3. The leisure and recreation zones (yellowish green), the institutional zones (red) and the central commercial zone (light blue) did not suffer any transitions. The new mixed zone that arose in the north-western part of the town during the simulation period was rather well modelled, particularly in S1 and S3.

Lastly, the shifts from non-urban areas to residential use represented the most challenging category of land use transition in the modelling experiment at issue. The reasons for the difficulties in detecting their shapes have been previously commented in this paper. It is worth remarking that 65 % of this type of transitions occur through the *expander* algorithm (Table 9). An evident shortcoming of this algorithm lies on the fact that, after the random selection of a cell for transition, neighbouring cells to it also undergo transitions regardless of their transition probability values.

The R & D team of CSR-UFMG, entrusted with the continuous upgrading programme of DINAMICA, is currently working to tackle this problem. Other enhancements such as the incorporation of fractal parameters in the transition algorithms as well as the possibility to define patches average sizes and variances for the *expander* and *patcher* algorithms separately are also envisaged.

To conclude, it is worth stressing here the wide feasibility (and the cells transition probability maps are a concrete prove) to optimise the simulations results by means of a model which embraces more refined algorithmic logics, highly suitable for the urban phenomena modelling under consideration.

4. Statistical Validation of the Model

With the purpose to conduct statistical tests for the spatial validation of models of land use dynamics, Constanza (1989) presents a procedure entitled "*Multiple Resolutions Method*", which can be applied to a wide variety of spatial resolutions through the change of size in a sampling window.

This sampling window moves over the entire images (Figure 18), and the average fit between two given scenes (the real and the simulated one) for a particular window size is calculated by the following expression:

$$F_w = \frac{\sum_{s=1}^{tw} \left[1 - \sum_{i=1}^{p} \frac{| a_{i1} - a_{i2} |}{2\, w^2} \right]}{t_w} \qquad (27)$$

where F_w is the fit for the window of size $w \times w$; a_{i1} is the number of cells belonging to class i in scene 1 (simulated image) and a_{i2} is the number of cells belonging to class i in scene 2 (real image) in the sampling window; p refers to the number of different classes found in the sampling window and t_w, to the total number of windows sampled in a scene for a window size of $w \times w$.

The total goodness of fit is then given by the equation below:

$$F_t = \frac{\sum_{w=1}^{n} F_w \; e^{-k(w-1)}}{\sum_{w=1}^{n} e^{-k(w-1)}} \qquad (28)$$

where F_t is the average of all fit measures obtained by the different window sizes employed in the analysis, F_w is the fit for sampling windows of linear dimension w, and k, a constant.

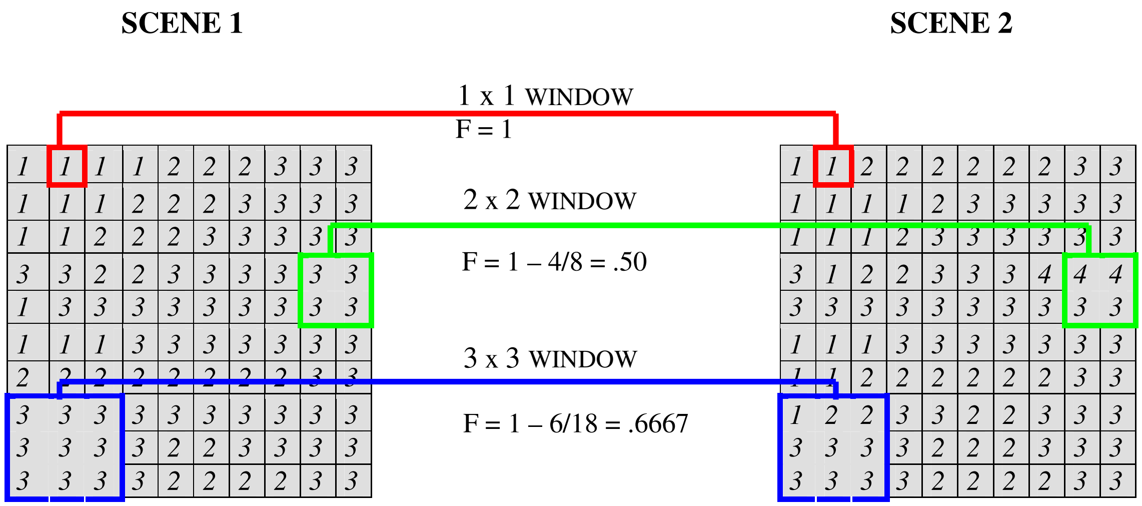


Fig. 18 – Example of the multiple resolution method for a scene of size 10 x 10 and with four classes. In this example, k=0.1 and F_t=0.84.
SOURCE: CONSTANZA (1989).

When *k* is zero, all the window sizes have the same weight, whereas when *k*=1, only the bigger windows are important. According to Constanza (1989), the values of *k* can be adjusted in function of the model objective and the data quality.

This multiple resolution method was implemented in a UNIX environment programme named FIT, developed by CSR-UFMG. FIT was applied for the best simulation results presented in Figure 17, with sampling window sizes of 3x3, 5x5 and 10x10 (Table 10).

TABLE 10 – TESTS OF THE MULTIPLE RESOLUTION GOODNESS OF FIT APPLIED TO THE BEST LAND USE SIMULATION RESULTS FOR THE CITY OF BAURU (1979-1988)

SIMULATIONS	***MULTIPLE RESOLUTION GOODNESS OF FIT (F)***
S1	*F = 0.902937*
S2	*F = 0.896092*
S3	*F = 0.901134*

5. Conclusion

The urban land use dynamics models have proved to be useful for the identification of main urban growth vectors and their general land use tendencies, what enables local planning authorities to manage and reorganise (if it comes into question) city growth according to the environmental carrying capacity of concerned sites and to their present and envisaged (future investments) infrastructure availability.

The urban expansion forecasts provided by such models also help local authorities in general, like submajors, districts administrators and municipal ministers, to establish investments goals in terms of technical and social infrastructure equipments, such as the extension of roads, the enlargement of the water supply and waste water disposal catchment areas, the creation of new bus lines, the construction of kindergartens, schools, hospitals and health centres, etc.

Decision makers from the private sphere can as well benefit from the modelling output data, since companies of transportation, conventional and mobile phones, cable TV and internet and others will have subsidies for defining priorities as to where and how intense to invest.

Also the organised civil society, either through NGOs or local associations, can profit from the modelling forecasts in order to enhance, by legal means, demanding social movements for the implementation of social and technical infrastructure, since their requests and respective arguments shall be based on realistic short- and medium-term urban growth trends.

Finally, it is worth reminding that the “weights of evidence” statistical method is not built upon rigid theories devices and does not either impose theoretical constraints to the modelling objects. Since this a wholly empirical approach, its applicability can be extended to further Brazilian and worldwide cities, provided that the minimum necessary sets of evidences maps are available.

6. Acknowledgements

The authors of this paper would like to express their thankfulness to the Bauru Planning Secretariat (SEPLAN) and the Water Supply and Waste Water Disposal Department (DAE) for providing the city maps. They are also grateful for the help and cooperation of the technical and administrative staff of the Centre for Remote Sensing of the Federal University of Minas Gerais (CSR-UFMG) as well as for the crucial financial support granted to this research by the Academic Coordination for Remote Sensing of the Brazilian National Institute for Space Research (PG-SER/INPE), the São Paulo State Foundation for Research Support (FAPESP) and the Brazilian National Foundation for the Undergraduates and Graduates Upgrade (CAPES).

7. References

Batty M., Xie Y. (1994) “From cells to cities”, *Environmental and Planning B*, 21, pp. 31-48.

Batty M., Xie Y. (1997) “Possible urban automata”, *Environmental and Planning B*, 24, pp. 175-192.

Batty, M. (2000) “GeoComputation using cellular automata”. In: Openshaw, S., Abrahart, R. J. (ed.), *Geocomputation*, New York: Taylor&Francis, chap.5, pp. 95-126.

Benati, S. (1997) “A cellular automaton for the simulation of competitive location”, *Environmental and Planning B*, 24, pp. 205-218.

Bonham-Carter, G. F. (1994) *Geographic Information Systems for Geoscientists: Modelling with GIS*, Ontario: Pergamon.

Clarke, K. C., Hoppen, S., Gaydos L. (1997) "A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay area", *Environmental and Planning B*, 24, pp. 247-261.

Constanza, R. (1989) "Model goodness of fit: a multiple resolution procedure", *Ecological Modelling*, 47, pp. 199-215.

Deadman, P., Brown, R. D., Gimblett, P. (1993) "Modelling rural residential settlement patterns with cellular automata", *Journal of Environmental Management*, 37, pp. 147-160.

Goodacre, A. K., Bonham-Carter, G. F., Agterberg, F. P., Wright, D. F. (1993) "A statistical analysis of the spatial association of seismicity with drainage patterns and magnetic anomalies in western Quebec", *Tectonophysics*, 217, pp. 285-305.

Hobbs, R. J. (1983) "Markov models in the study of post-fire succession in heathland communities", *Vegetation*, 56, pp. 17-30.

JRC (Joint Research Centre – European Commission / Institute for Remote Sensing Applications), ESA (European Space Agency / ESRIN – Earthnet Programme Office). (1994) *Modelling Deforestation Processes – A Review*, Trees Series B, Research Report n.1, Luxembourg: ECSC-EC-EAEC.

Phipps, M., Langlois, A. (1997) "Spatial dynamics, cellular automata, and parallel processing computers", *Environmental and Planning B*, 24, pp. 193-204.

Portugali, J., Benenson, I., Omer, I. (1997) "Spatial cognitive dissonance and sociospatial emergence in a self-organizing city", *Environmental and Planning B*, 24, pp. 263-285.

Soares-Filho, B. S. (1998) *Modelagem da Dinâmica de Paisagem de uma Região de Fronteira de Colonização Amazônica*, unpublished PhD dissertation, EPUSP, Universidade de São Paulo, São Paulo.

Soares-Filho, B. S., Cerqueira, G. C., Pennachin, C. L. (2001) "DINAMICA: a new model to simulate and study landscape dynamics", *Ecological Modelling*, in press.

White, R. W., Engelen, G. (1997) "Cellular automaton as the basis of integrated dynamic regional modelling", *Environmental and Planning B*, 24, pp. 235-246.

White, R., Engelen, G., Uljee I. (2000) "Modelling land use change with linked cellular automata and socio-economic models: a tool for exploring the impact of climate change on the Island of St. Lucia", [*online*].