XXIV CONFERENZA ITALIANA DI SCIENZE REGIONALI

LAND USE DYNAMICS: A STOCHASTIC MODEL BASED ON KNOWLEDGE FROM SOM NEURAL NETWORKS

Lidia DIAPPI[(1)] and Paola BOLCHI[(2)]

[(1)] Department of Architecture and Planning – Polytechnic of Milan, Via Bonardi 3, 20131 Milano – tel. +390223995471, fax +390223995454 – email: lidia.diappi@polimi,it
[(2)] Department of Architecture and Planning – Polytechnic of Milan, Via Bonardi 3, 20131 Milano – tel. +390223995478, fax +390223995454 – email: paola.bolchi@polimi,it

SUMMARY

It is widely accepted that the spatial pattern of settlements is a crucial factor affecting quality of life and environmental sustainability, but few recent studies have attempted to examine the phenomenon of sprawl by modelling the process rather than adopting a descriptive approach. The new approaches of artificial intelligence, and in particular, systems involving parallel processing (neural networks, Cellular Automata and Multi Agent Systems) defined by the expression "*Neurocomputing*", allow problems to be approached in the reverse, bottom-up, direction by discovering rules, relationships and scenarios from a database.
The present article describes a dynamic model of urbanisation based on transition rules learned from an autopoietic neural network (SOM, self-organizing map) which processes changes in land use occurring in the area being studied over a certain period; a stochastic model then allocates the land-use changes in the subsequent period (forecast) by applying learned rules.
The results have discovered that the spontaneous process follows a logic of expansion around urbanised nuclei and road axes and does not appear to exhibit undifferentiated diffusion, contrary to the concerns expressed in much of the literature.

1 INTRODUCTION

The recent tendency of settlement patterns to sprawl is a central issue for all policies endeavouring to adopt sustainable approaches. It is surprising that there have been few efforts attempting to predict future urbanisation patterns even in the case of a "no action" scenario. The question was partly addressed by the integrated models of land use and transportation developed in the 1970s and 1980s. Recent applications of these models (Putman, 1979), models developed by the school of Dortmund (Wegener, 1998), TRANUS (De la Barra, 1989) and the very recent URBAN SIM model (Waddell P. and Gudmundur F., forthcoming), just to mention a few examples – indicate that there is still lively interest being shown by experts, planners and politicians to find out more about the future functional and spatial patterns of cities and regions. A principal objective of these models is to produce spatial patterns which are in accordance with theoretical principles.

- This vast amount of scientific work has some common features:
- An assumption that the complexity of urban systems can be limited to some key measurable variables: population and dwellings, employment, economic activities, and space mainly assumed in terms of distance and available area;
- A formalisation in equations that is derived from a well-established theoretical basis (spatial interaction, principles of hierarchy, basic economic theory, random utility theory, etc.);
- An economic efficiency criterion guides the choices of populations and activities.

An aggregate scale of description and representation (a meso-scale) applies, which generates a defined transformation scenario for categories or classes.

In recent years new territorial problems and renewed epistemological needs have resulted in significant innovations to models of land use.

Environmental issues – brought to notice by the question of sustainable development, particularly in cities – have focused attention on the consumption of non-renewable resources, with consumption of land being a top priority.

Non-urbanised land, no longer considered a residual space or passive support, becomes a crucial asset and an important ecological and amenity resource, but is being steadily encroached upon by the pressure of dispersed urbanisation. However, if suitable settlement models could be identified, dispersed urbanisation could be compatible with the requirements of sustainability.

On an epistemological level, dissatisfaction with the results achieved so far using earlier models to capture the actual dynamics of land use has prompted growing interest in alternative approaches and the emergence of a new paradigm based on developments in artificial intelligence (AI), which is able to examine the structure of data, discover significant

relationships and to extrapolate them to construct scenarios in complex situations marked by self-organisation and uncertainty.

In traditional modelling, the model is derived from theory in a top-down process, whereas the IA approaches reverse the order and construct knowledge endogenously through a bottom-up process. They start from data and the transformation rules – or more exactly, the transition from one state to another – are only discovered *a posteriori*. These approaches adopt distributed architecture such as cellular automata (CA), neural networks (NN) and multiagent systems (MAS), which process data occurring at the level of individual interactions on the micro-scale. Among these approaches, neural networks have proved to be a very effective and reliable tool since, via a learning phase, they can quantify and model complex data and behaviours. Neural networks have already been widely used as a method for classifying fuzzy data and as statistical estimators, but have only recently been applied in the planning area for forecasting land use. The LTM model (Pijanowski, Brown, Snellito, Manik, 2002) uses NNs applied to GIS to learn how factors such as roads, motorways and the local road system, recreational facilities, and agricultural settlements can influence urbanisation patterns, assessing their effects and effectiveness at different scales.

The authors of this article have already tested a model for forecasting land use which was based on supervised neural networks applied to a database containing information organised according to a cellular automata method. This assumed that the change in land use in a certain part of the territory was a function of land use in the neighbouring cells at the previous temporal threshold (Diappi, Bolchi, Franzini, 2002). However, the experiment was limited by the type of supervised neural network used and the results were not very satisfactory. In the present work we wished to follow a different approach which used the neural network's ability to identify transition rules and, on this basis, to predict the spatial patterns of urbanisation via a stochastic simulation where the rules were translated into transition probabilities.

2 METHODOLOGY

The aim of the methodological approach chosen was to identify the transition rules from the information describing land-use patterns at two temporal thresholds. The databases used referred to the southern area of the province of Milan described by main types of use and was subdivided into cells in a square grid.

The neural network method, and in particular, the SOM (self-organizing map) method, enables the typologies of transitions from an initial state at time t to the subsequent state at time t+1 to be identified.

The type of information found in typical cellular automata hypotheses was used – what influences the transformation of one portion of the territory from one land use to another is a

function of the initial state of the cell (and therefore of its land use) and of the state of neighbouring cells, i.e. of the surroundings at the initial timepoint. The logic of spatial localisation according to CA is therefore "local"; for example, a subject will decide to build his or her dwelling in a particular place on the basis of its proximity to other dwellings, services and roads. But in the classical form of territorial modelling based on CA (Cecchini, 2003; Engelen and Ulje, 1993), the rules are formulated *a priori* and exogenously. The resulting transformation mechanisms are necessarily oversimplified and the scenarios presented do not adequately reflect complex realities.

In this article, while we accept the "local" logic underlying CA and therefore use the same input information, our transformation rules are discovered *a posteriori*, by exploring the relations between the initial states of the cell and its neighbours, and the final state of each individual cell.

If the data is processed using SOM networks, which basically classify the data, some transformation typologies or classes emerge, which group all the individual cells according to similarities in their dynamic behaviour.

In particular, for each class, the SOM NN identifies a prototype, called a codebook, which represents the dynamic behaviour of that group of cells and expresses a "transition rule" which contains the "average" initial state of cells and neighbourhood and the expected "average" final state.

The rules, which are tested according to appropriate procedures, are the fundamental instrument used, and are applied to the most recent temporal threshold, and the pattern of future urbanisation.

So, starting at an initial state that is changed at time t+1, the cells, that in the meantime have obviously also changed, are allocated to the SOM classes based on their similarity to the prototype only for the initial state.

At this point the easiest solution would consist in simply applying the use predicted by the class rule to each cell. But as mentioned, the codebook expresses an average transition behaviour, from which the cells of the group deviate to variable extents.

It therefore seemed preferable to build a simulation model, based on a Monte Carlo procedure, where the change in use is extracted from within a probability distribution calibrated from transformations that actually occurred for the cells of the group in the previous time period. The whole procedure is depicted in fig. 1.

3 THE CASE STUDY AND THE DATA

The area under study covered the territory extending in a semi-circle round the south of Milan, but excluding the city of Milan itself, containing a large amount of agricultural land and protected Natural Park areas (Parco del Ticino, Parco Sud). The area is being increasingly

eroded by urbanisation caused by an exodus from Milan and by the creation of shopping centres and industrial developments. The objective of examining not only the amount, but also the form of future probable patterns of development, clearly meets the objective of assessing future patterns in the context of sustainability objectives.
The model uses a basic grid of cells with a side of 500 metres, for a total of 2703 cells. The land uses considered are: residential, commercial, industrial and unbuilt and are expressed as a percentage of the total cell area. Only two temporal thresholds were considered: 1980 and 1994.
The following information was provided to the neural network:

- Land uses of cell *i* at time *t* (1980)
- Land uses of neighbouring cells at time *t*
- Average distance from roads at time *t*
- Land uses of cell *i* at time *t+1* (1994)

In summary, the record contains 7 pieces of information (on the cell and neighbourhood) related to the initial state and three pieces of information (only on the cell) related to the final state.

4 NEURAL NETWORKS

Neural networks are powerful instruments which use the learning approach of an artificial intelligence machine to quantify and model complex patterns and behaviours. In recent years neural network models have generated increasing interest in the field of urban and regional analysis. The idea of training a system to independently develop cognitive abilities, such as adaptive responses to a set of information, is a fundamentally new approach to processing information. Openshaw (1992 and 1993), White (1989) and Fischer (1992) were the first to appreciate the potential of neurocomputation applied to regional sciences.
A neural network is a system where information is processed in parallel. It consists of processing points (*nodes*) which can possess a logical memory and can process localised information. They are characterised by being able to adapt and are interconnected by transmission channels of unidirectional signals called *connections.*
A weight is associated with each connection, specifying the intensity of the link.
There are two distinct phases in the operation of an artificial neural network (ANN): the learning or *training phase,* when the network learns the rules (and thus defines the intensity of the connection or weights) starting from "examples" and the *querying phase*, when the network processes the input data using the learned rules. The two phases are split by a *test* phase when the network's learning ability is assessed.
The ANN dynamic varies according to the *learning rules* adopted, which in turn depend on the objective of the network and thus on the type of output the network is required to learn to

generate. The learning dynamic of networks is usually divided into two classes: *supervised networks* and *non- supervised networks*.

If the vectors of experimental data for which the network has to identify the connections are defined as "models", then (Hecht-Nielsen, 1991; Buscema 1994a):

- *supervised networks* assume predefined targets to which every model in the network has to conform. These are output targets decided by the external environment and are not deduced by the input vector or any part of it.
- *monitored networks* are networks whose ideal targets are predefined for each input model, but each target is taken by the input vector of the network or part of it.
- *autopoietic networks* are networks whose targets are formed dynamically during the learning phase and cannot therefore be defined *a priori*. They are networks that stabilise on the basis of the statistical regularities of their input models.

Most of the applications of NNs In the field of regional studies have used supervised networks (Griguolo, forthcoming; Fischer, 1992 and 1997) and monitored networks (Buscema, Diappi, Ottanà, 1998). Some work has recently appeared using autopoietic, or self-organising, neural networks (Diappi, Bolchi, Franzini, 2001) such as SOM networks, which belong to the more general set of "mapping networks", i.e. networks able to identify a mathematical interpolating function from a set of data.

To verify the ability of NNs to identify connections, records are first used to train the network, and then, deprived of the output variables, resubmitted to the network for the scenario phase. This makes it possible to calculate the error based on the number of cells which are "wrong" when compared with the output proposed to the network when being trained.

5 SOM NETWORKS (SELF-ORGANISING MAPS)

Some of the properties of SOM networks (Kohonen, 1995) are summarised here and the main results obtained in the work on Milan are outlined; a full description has already been presented in a previous publication (Diappi, Bolchi, Franzini, 2002).

Since SOMs use a learning technique which is non-supervised and self-organised, they are also called autopoietic networks, since the target is dynamic and changes during training. Using an architecture composed of two layers, one input and one output, the SOM network dynamically constructs clusters through systematic collation, but it does this randomly among records that have been acted on by 'winner' records which organise the most similar records around them. The only, but most sensitive, exogenous decision is selecting the number of classes desired, which obviously involves balancing the need to produce differentiated classes against the need to have significant class sizes. In the present work, a process of trial and error led to 16 classes being chosen. One objection, which needs to be immediately answered, is

what advantage there is between this instrument and other well-known statistical methods, such as multivariate statistical analysis, which meet the same statistical objective of identifying classes of similar phenomena. The ability of NNs to gather nonlinear relations, unlike the above-mentioned methods, means that NN are often used as an alternative or as a complement to them. Many studies have compared the effectiveness of approaches based on statistical estimates with those using NNs, for example when calibrating models. The range of different results obtained shows that they are instruments that operate with different logics and they measure different relations but they converge in indicating the "best fit" of some models compared to others if the choice is presented.

After analysing the study data, classification of records showed there is a spatial logic. Cells belonging to the same class are positioned contiguously or in consistent positions relative to urbanised areas: for example at the edge of the urban centre or along the main transport routes.

The results are summarised in fig. 1, where the profiles and codebooks of each group are shown, and in fig. 2 where the spatial map of cells according to their class is shown.

In fig. 2, diagrams of the codebooks of an individual cluster superimposed on the profiles of cells belonging to it, enable the prototypical properties of the group to be read, while the set of profiles of individual cells, through their deviations or agreement with the codebook, enables the distinguishing variables, the influences on the classification, to be assessed,

In fig. 2, the 16 groups are each shown in a diagram with the x axis representing state variables (in order from the origin: land use of cell in 1980, proximity to roads and land use of neighbourhood in 1980, land use of cell in 1994). The y axis shows the degree of activation of individual variables. The envelope of all the records in the group is shown in grey, around the codebook in yellow.

As can be seen, the first row shows cells which are unbuilt and remain so at the second temporal threshold; they are mainly agricultural areas. However, on moving to the right, proximity to roads increases and this indicates classifying for "probability of future urbanisation".

Going vertically down the first class 1-1, increasing proximity to roads and an increase in residential use over the considered time period can be seen. More precisely, transformations of agricultural areas to residential use occur in class 1-2, of already partially urbanised areas in class 1-3 or of areas already residential in 1980 in class 1-4.

In the last row, class 4-4 shows an increase in industrial development in areas which were already industrialised in 1980. In this case too, the final column gradually takes account of modifications towards the end use, which only emerges as a final state in classes 3-4 and 4-4.

From the description of the possible land-use dynamics identified by the SOM, "transition rules" can be extracted which can be used as input for a model forecasting subsequent urbanisation.

As mentioned, a spatial logic to transformations also emerges. The map in fig. 3 shows spatial distribution of cells according to their class. It can be seen how new residential settlements tend to favour expansion outwards from urban centres and proximity to roads while industrial development appears to have a tendency to cluster together.
In general it can be seen that cells of the same transition class are also spatially contiguous, thus confirming the validity of the method being tested.

6 THE PROBABILISTIC MODEL

We now face the problem of how to apply the discovered transition rules in the form of a dynamic model. The dynamic could be formalised as a stochastic process using transition probabilities between initial and final states that can be extracted by the codebooks identified by SOMs. The codebook in fact perfectly describes the average profile, both of the initial state and the final state.
However, this would result in an excessively oversimplified procedure; as mentioned above, these are average values and the real transitions occurring in the period deviate from them considerably. In fact the transformations in each class can be better described by a probability distribution of land-use activations rather than a single average value.
It therefore seemed preferable to adopt a simulation procedure using probability distributions to allocate the land uses to the various cells. The initial state is therefore 1994 and the temporal threshold for the forecast, assuming the same time lag, will be 2008, the final state of the system.
The procedure adopted was as follows:

1. From the SOM classification, probability distributions of the growth in the three types of land use were extracted for each class, based on the frequencies of observed increases in area.
2. For each cell, new records were built containing information about the cell and its neighbourhood (7 fields) for 1994, which now became the first temporal threshold;
3. The new records were assigned to the SOM classes previously discovered, on the basis of similarity to the "initial state" part of the codebooks.
4. To assign a final state to each cell, a Monte Carlo procedure was used on the cumulative probability functions and the percentage of growth relative to the different land uses was extracted.
5. The model was tested using the 1980-1994 data to check its reliability.
6. The model was implemented and
7. The results have been evaluated by different criteria and indexes.

7 RESULTS

The difficulties in interpreting results obtained from models are well known, particularly when neural networks are used, since their 'black box' nature makes 'objective' validation of results difficult to achieve. It is therefore appropriate to rely on qualitative evaluation to check whether the results achieved make sense and are consistent.

Various approaches have been followed in this work to adequately evaluate the results. Some are based on the aggregate area data for different end uses, others use SOM classes repeated at different temporal thresholds to analyse changes of class and position of cells, and other approaches use statistical indicators of compactness and dispersion to evaluate trends in spatial patterns.

1. Probability of change

Before analysing the final results, it is worth making a few comments on the probability distributions obtained from evaluating the groups of SOM networks. fig. 4(1,2,3) show probability distributions of growth for all SOM groups. The x axis shows the percentage area activated in the considered time period (1980-1994) relative to the total area of the cell and the vertical axis shows the different probabilities. For the residential dynamic, it can be seen that the main increases occurred for classes C31 and C22 (extra-urban areas of residential settlement), C41 and C42 (urban areas of residential settlement), followed by C32 (extra-urban commercial areas). These classes of cells are in fact the ones producing greatest residential increases. The other groups display probabilities mainly below 10%.

The analogous distributions for industrial use display greatest increases in classes C34 (industrial areas near to roads) and C44 (infilling of existing industrial areas far from roads) and, to a lesser extent, in C24 (non-urbanised areas); increases for other classes are negligible (below 10%).

For commercial uses the main increases are concentrated in class C32 (extra-urban commercial areas).

2. The land use dynamics

A first examination of the results focuses on a comparison between total areas dedicated to the three end uses estimated for 2008 and the historical data of 1980 and 1994.

	residential	industrial	commercial	total
1980	33.86	15.72	1.86	*51.44*
1994	47.04	24.56	3.34	*74.94*
2008	3.31	35.85	6.15	*105.31*
increase 80-94	13.18	8.84	1.48	*23.50*
increase 94-08	16.27	11.29	2.82	*30.37*
incr % 80-94	38.9%	56.3%	79.4%	*45.7%*
incr % 94-08	34.6%	46.0%	84.4%	*40.5%*

For residential and industrial uses, the increase would basically appear to continue, though at a reduced rate. Commercial use would grow at a constant rate.

The statistical distribution for the size of new urbanised areas according to function is shown in fig.5(1-2-3-4). The x axis shows the percentage of growth and the y axis the number of cases. For all three functions the average lot sizes appear to be larger in the first period than the second.

3. The change in urbanisation dynamic via SOM classes

Since application of the model also required reclassifying the SOM of cells based on the initial state in 1994, another interesting comparison is to examine the number of cells present in each class at the two temporal thresholds (forecast and calibration).

It can be seen in fig. 6 how all the classes have decreased relative to non-urbanised land (C12-C14), with the exception of class C11, which represents permanently green areas; this result is predictable since part of the new urbanisation has obviously occupied unbuilt areas. The analysis of the intermediate classes (C21-C34) is more interesting, with an alternating tendency to increase and decrease being evident. Class 23, with a clear increase, is an exception and corresponds to transitions from pure residential to commercial/ residential areas. The last four classes, which can be defined as zones of mixed residential/industrial use, from the more central, mainly residential ones (C41), to the more outlying ones of mainly industrial development (C44), all increase significantly..

If we consider the dynamics of cells and their neighbours between the three temporal thresholds 1980, 1994 and 2008 and analyse which SOM classes they belong for each period, it is possible to evaluate the class changes that have occurred in the period. In the first period of time the principal transitions are of cells at the edges of inhabited centres or nuclei of new urbanisation mainly along road axes.

Between 1994 and 2008, however, transitions are mainly of cells located further away from historic centres and situated around the nuclei which emerged in the previous period.

4. The degree of compactness

A significant help in interpreting the results obtained so far may be provided by applying an index of compactness for the historical thresholds 1980, 1994 and 2008, applying it first to the entire urbanised area and then just to the newly urbanised areas of the two periods.

The indicator used (Xuan Thinh, Arlt, Heber et al., 2002) measures the spatial configuration of urban functions applied to a grid of cells and expressing the area dedicated to different uses as a percentage.

For each pair of cells *i* and *j* (*i*=1(1) *N*-1, *j*=*i*+1(1) *N*) with respective areas Z*i* e Z*j,* attraction is calculated using an inverse square relationship analogous to the law of gravitation:

$$A_{ij} = \frac{1}{C}\frac{Z_i Z_j}{d_{ij}^2}$$

where d_{ij}^2 is the Euclidean distance between the centres of the cells i and j and $C = 100$ m^2 is a proportionality factor to make the term A_{ij} dimensionless. The degree of compactness is defined as the mean value of the gravitational matrix:

$$T = \sum \frac{A_{ij}}{N(N-1)/2}$$

where N is the total number of urbanised cells. T is a measure of centrality for the spatial interaction between aggregate areas. If urban structures are more extensively dispersed and dot-like, then the interaction between urban aggregates is weaker. Conversely, as the compactness of urban structures increases then T will be larger.

The index, applied to the entire area and to the whole existing built area at the three temporal thresholds, shows (fig.7 (1-2)) increasing compactness in all uses, particularly commercial, and highlights how new urbanisation processes fill areas and aggregate to old centres. When the same index is applied just to new construction in the two periods, there is a decrease, indicating that the settlement models are more dispersed. This second index is not very significant however since the position of new urbanisation clearly depends on pre-existing structures and not on the functions created in the same period.

5. The degree of dispersion

Assuming as centroids the main historical centres of the area (Abbiategrasso, Binasco Melegnano, Rozzano, Paullo, San Giuliano, San Donato, Trezzano) and tracing circles of variable radius, but sufficient to include both the urbanised area and the surrounding agricultural area (fig. 8), the following index was calculated for each use:

$$\frac{\sum_i S_i d_{ic}^2}{\sum_i S_i}$$

where S_i is the built area and d_{ic} is the distance from the centroid. The indexes applied to the total built area at the three thresholds show varying behaviour. For residential use, dispersion generally increases slightly in Abbiategrasso, Binasco and San Donato, decreases in Paullo and Trezzano and is stable in Rozzano. When the index is applied to increases occurring in the period it shows that dispersion decreases everywhere, except for San Donato, San Giuliano and Binasco (fig. 9.1 a and b).

Industrial settlements show settlement models similar to the residential ones, with some variability. Total dispersion is again stable or increases slightly, but Melegnano and Trezzano go against the trend, as is more evident in the diagram of just the increases occurring in the two periods (fig.9.2 a and b).

The location of commercial functions is difficult to interpret since just a few large centres are involved and they are unlikely to display a common trend. Nonetheless, in this case too, when the index is applied to existing structures it seems to indicate greater dispersion, i.e. new

extra-urban locations, but the increase diagrams do not support this hypothesis (fig. 9.3 a and b).
In short, this does not appear to show that most new settlements result in greater dispersion of settlements and hence greater use of land.

6. The ratio perimeter/area

A further measure of compactness/dispersion is given by the simple ratio perimeter/area of urbanised cells, an index which indicates the degree of separation of urban agglomerates. The diagram of the existing situation at the three historic temporal thresholds shows a decrease of intermediate data and hence evidence of growth (fig. 10 a and b). This would seem to confirm the hypothesis that in the period 1980-1994 new settlements occurred via a process of infilling of existing urban areas and in the subsequent period there was a process of peri-urban urbanisation away from historic centres. This hypothesis is confirmed by the trends shown in the increase diagram.

In summary, some distinct trends can be seen to emerge in the dynamic of the area under consideration:

- While the process certainly involves encroachment on agricultural areas surrounding urban centres, it is not a process of uncontrolled diffusion of settlements over the territory. New urbanisation appears to form and consolidate in fairly compact fashion around urban nuclei, as is indicated by the indexes of compactness. The two phases – the calibrated and the simulated – show an initial period where urbanisation occurs in nuclei external to the inhabited area with settlement extending in large lot sizes and a second phase where infilling occurs and the nuclei expand with smaller size lots.
- The index of dispersion around historic centres is not inconsistent with this interpretation because the new nuclei which are established outside the centres along roads, while they are compact, nonetheless increase the value of the index of dispersion, which increases with the increase in distance of settlements from the centroid;
- There is a high functional mix in new urbanised areas, which bring together residential, industrial and commercial uses. The choice of residential location seems to be more focused on a search for privacy and proximity to services and workplace than the social contacts and environmental quality generally offered by exclusively residential zones.

8 CONCLUDING REMARKS

The method presented here aimed to combine the significant investigative power of neural networks in organising knowledge with a stochastic simulation model able to produce urbanisation scenarios based on rules learned by NNs.

The NNs showed they were able to capture the fundamental characteristics of urbanisation, identifying fuzzy typologies in the spatial dynamic. These were then used as ‘rules’ of change and entered into the stochastic simulation model.
It was then decided to adopt an ‘eclectic’ approach, combining methods of geocomputation, with knowledge being built from the bottom up, and traditional ‘top-down’ methods which apply explicitly-defined rules.
The benefits of this are that rules are transparent and it is therefore easier to interpret and evaluate the results.
In the context of understanding the phenomenon of “urbanisation”, it is important to have discovered that the spontaneous process follows a logic of expansion around urbanised nuclei and road axes and does not appear to exhibit undifferentiated diffusion, contrary to the concerns expressed in much of the literature.

9 REFERENCES

Buscema M, Diappi L., Ottanà M.(1998) “A neural Network investigation on the crucial Assets of urban Sustainability”, *Substance Use and Misuse*, vol. 33, pp. 793-817 ISSN: 1082-6084 Special Issue on Artificial Neural Networks and Social Systems.

De la Barra T. (1989) *Integrated Land Use and Transportation Modelling*, Cambridge University Press

Diappi L, Bolchi P., Franzini L.(2002) “The urban Sprawl Dynamics: does a Neural Network understand the spatial Logic better than a Cellular Automata?”, 42th European Regional Science Association Congress, Dortmund, August 27-31, CD-ROM

Diappi L, Bolchi P., Franzini L.(2001) “Urban dynamics: behavioural rules of Land Use change in the South Metropolitan Area of Milan through Neural Networks SOM”, 41th European Regional Science Association Congress, Zagreb 29th August-1st September 2001, CD-ROM

Fischer M., Gopal S. (1992) “Neural network models and interregional telephone traffic comparative performance between multi-layer feed-forward network and the conventional spatial interaction model”, in *WSG Discussion Papers* **27**, Institut für sozial Geographie - Wirtschafts Universität - Vienna.

Fischer M. (1997) "Computational Neural Networks: a new paradigm for spatial analysis", *Environment and Planning A*, 29, 1873-1891

Fischer M.M., Leung Y. (eds) (2001) *Geocomputational Modelling*, Springer Verlag, Berlin-Heidelberg

Griguolo S. (2004) “Neural classifiers and land cover recognition: merging radiomenric and ancillary information”, in L.Diappi (ed) *Evolving cities: geocomputation in planning process*, Ashgate Pub, Aldershot

Xuan Thinh N., Arlt G., Heber B., Hennersdorf J., Lehmann I. (2002) “Evaluation of urban land-use structures with view to sustainable development”, *Environmental Impact Assessment Review*, 22, 475-492

Openshaw S., Abrahart R.J. (2000) *Geocomputation*, Taylor & Francis, London and New York

Hech-Nielsen R. (1990) *Neurocomputing*, Addison-Wesley, Reading, Ma

Kosko B.(1992) “Neural Networks and fuzzy Systems:a Dynamical System Approach to Machine Intelligence”, Prentice Hall, Englewood Cliffs, NJ, USA

Openshaw S. (1993) “Modelling spatial interaction using neural net”, in Fischer M., Nijkamp P. (ed.), *GIS, spatial modelling and policy evaluation,* Springer Verlag, Heidelberg, 147-164.

Openshaw S. (1998) “Neural Networks, Genetic and Fuzzy Logic Models of spatial Interaction”, *Environment and Planning a*, vol.30, 1857-1872

Putman S.H. (1979) *Urban residential location models*, Martinus Nijhoff Publishing

Wegener M. (1998) “Applied Models of urban land use, transport and environment, state of the art and future developments”. In Lundqvist L., Mattson L.-G., Kim T. J (Eds.): *Network Infrastructure and the urban Environment: Recent Advances in Land Use/Transportation Modelling*. Berlin, Heidelberg, New York, Springer Verlag, 245-267

White R., Engelen G. (2000) “High-resolution integrated modelling of the spatial dynamics of urban and regional systems” *Computers, Environment and Urban Systems*, 24, pp 383-400.

White R., Engelen G., Uljee I. (1993) “Cellular automata modelling of fractal urban land use patterns: forecasting change for planning applications”, in Eight European Colloquium on Theoretical and Quantitative Geography, Budapest

Waddell P. and Gudmundur F. Ulfarsson (forthcoming), “Introduction to Urban Simulation: Design and Development of Operational Models. Handbook in Transport”, Volume 5: *Transport Geography and Spatial Systems*, Stopher, Button, Kingsley, Hensher eds., Pergamon Press (www.urbansim.org)

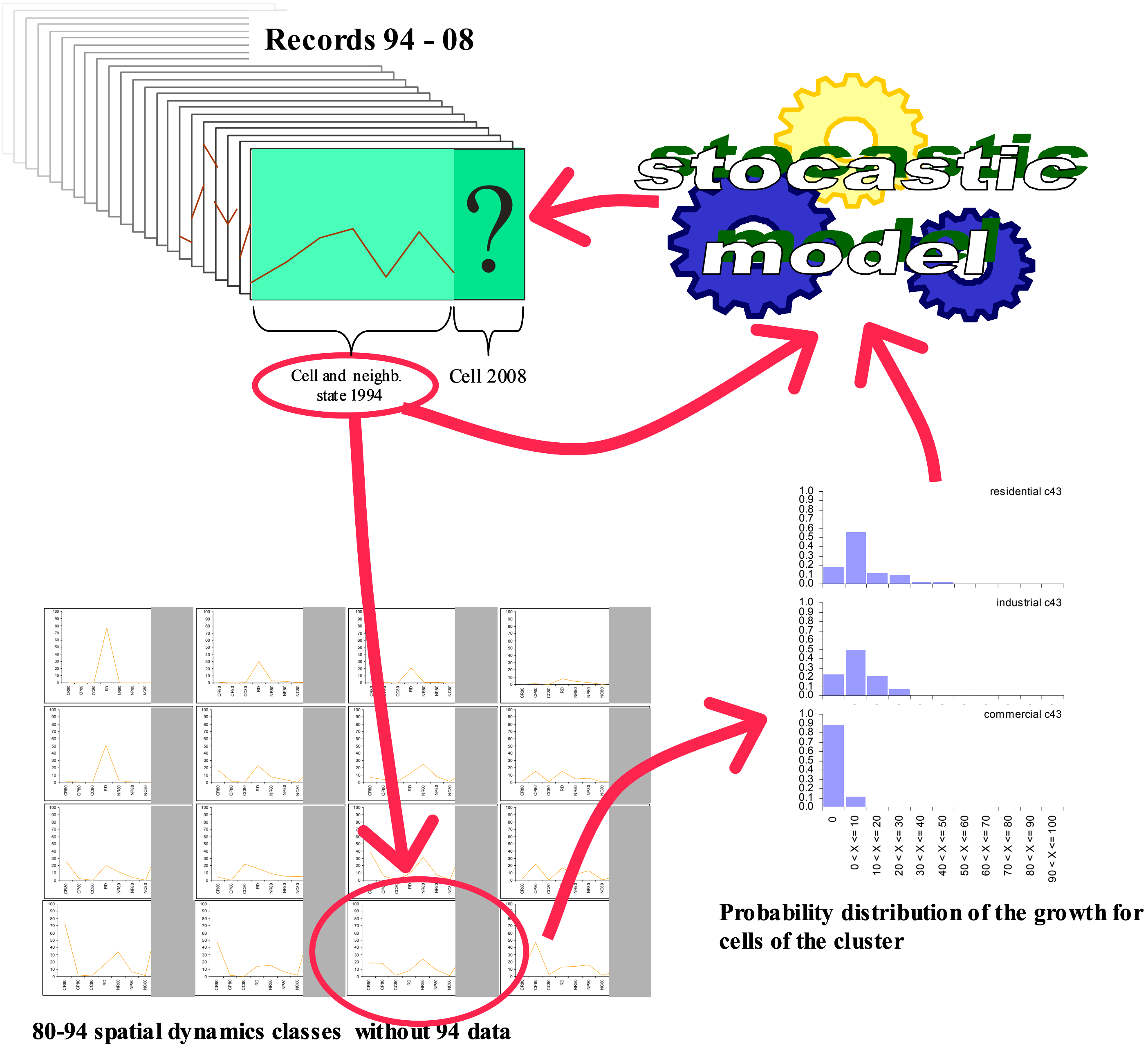
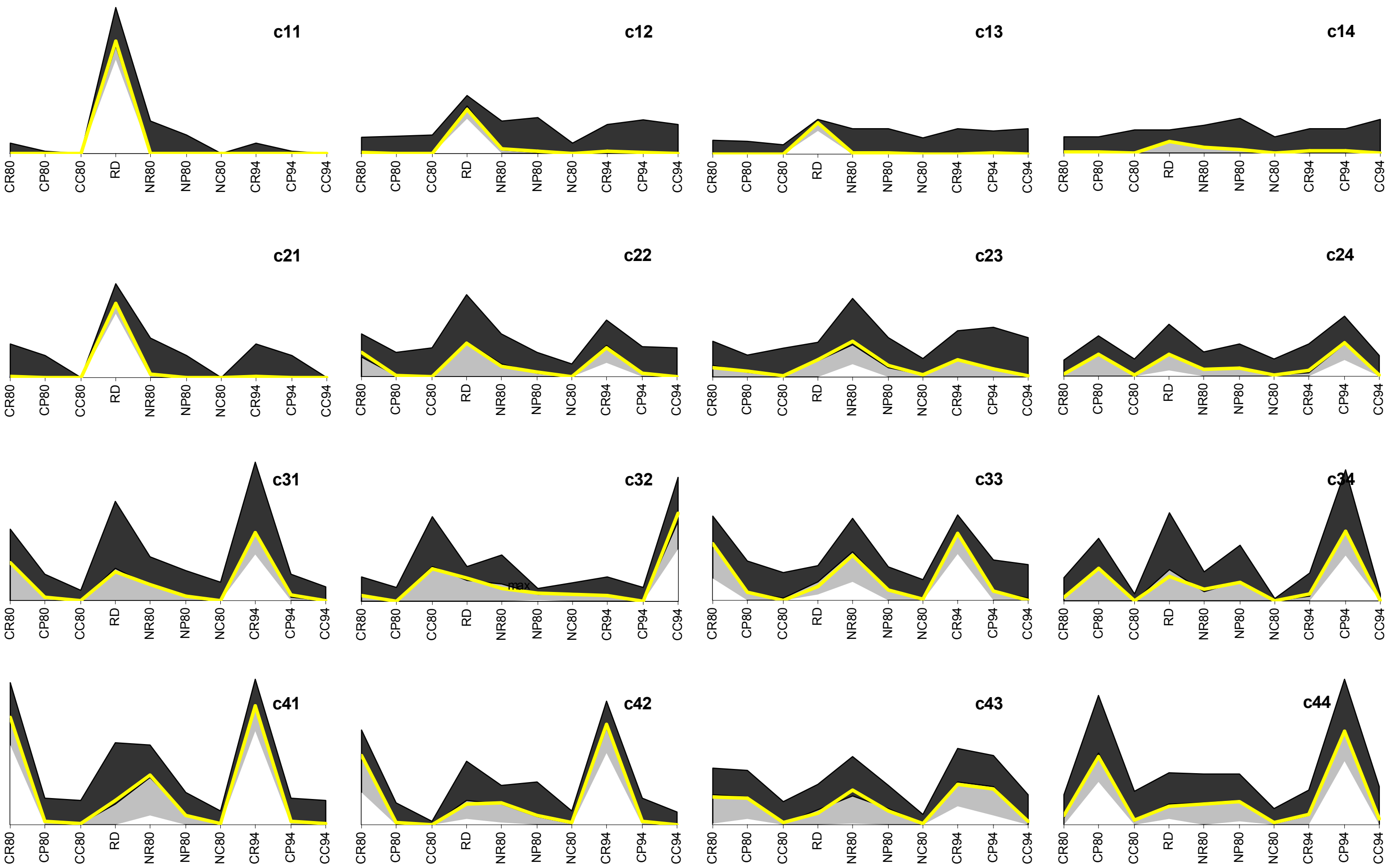


Figure 1 – *The procedure used. The SOM classes produce knowledge about the distribution of growth for each land use in each class. The probability distributions obtained feed the stochastic model which produces the forecasted growth for each cell.*



CR80	Cell, residential, 1980	**CR94**	Cell, residential, 1994
CP80	Cell, productive, 1980	**CP94**	Cell, productive, 1994
CC80	Cell, commercial, 1980	**CC94**	Cell, commercial, 1994
RD	Road distance		
NR80	Neighbourhood, residential, 1980		Codebooks range, values over the mean
NP80	Neighbourhood, productive, 1980		Codebooks range, values under the mean
NC80	Neighbourhood, commercial, 1980		Codebook

Figure 2 - *The classes of records produced by the SOM Networks. The prototypical profile (in Yellow) is the code book.*

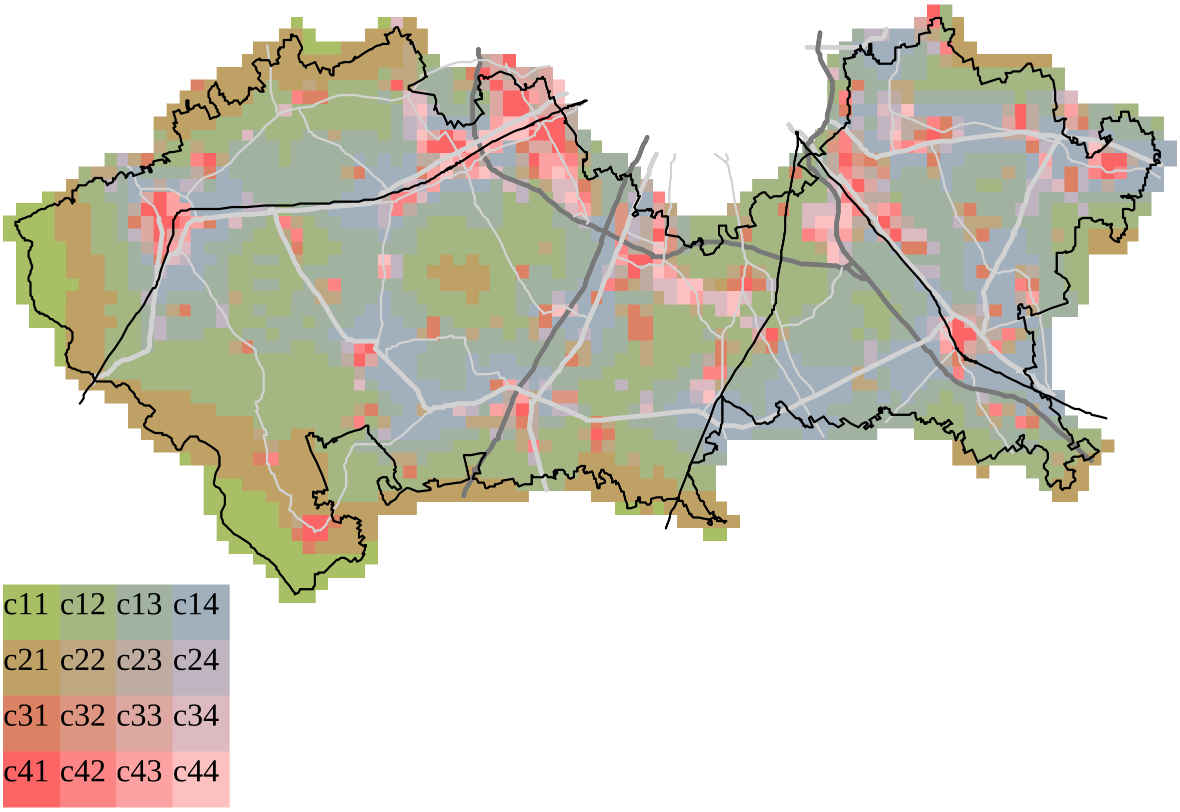


Figure 3 - *The spatial distribution of the cell classes produced by the SOM Networks.*

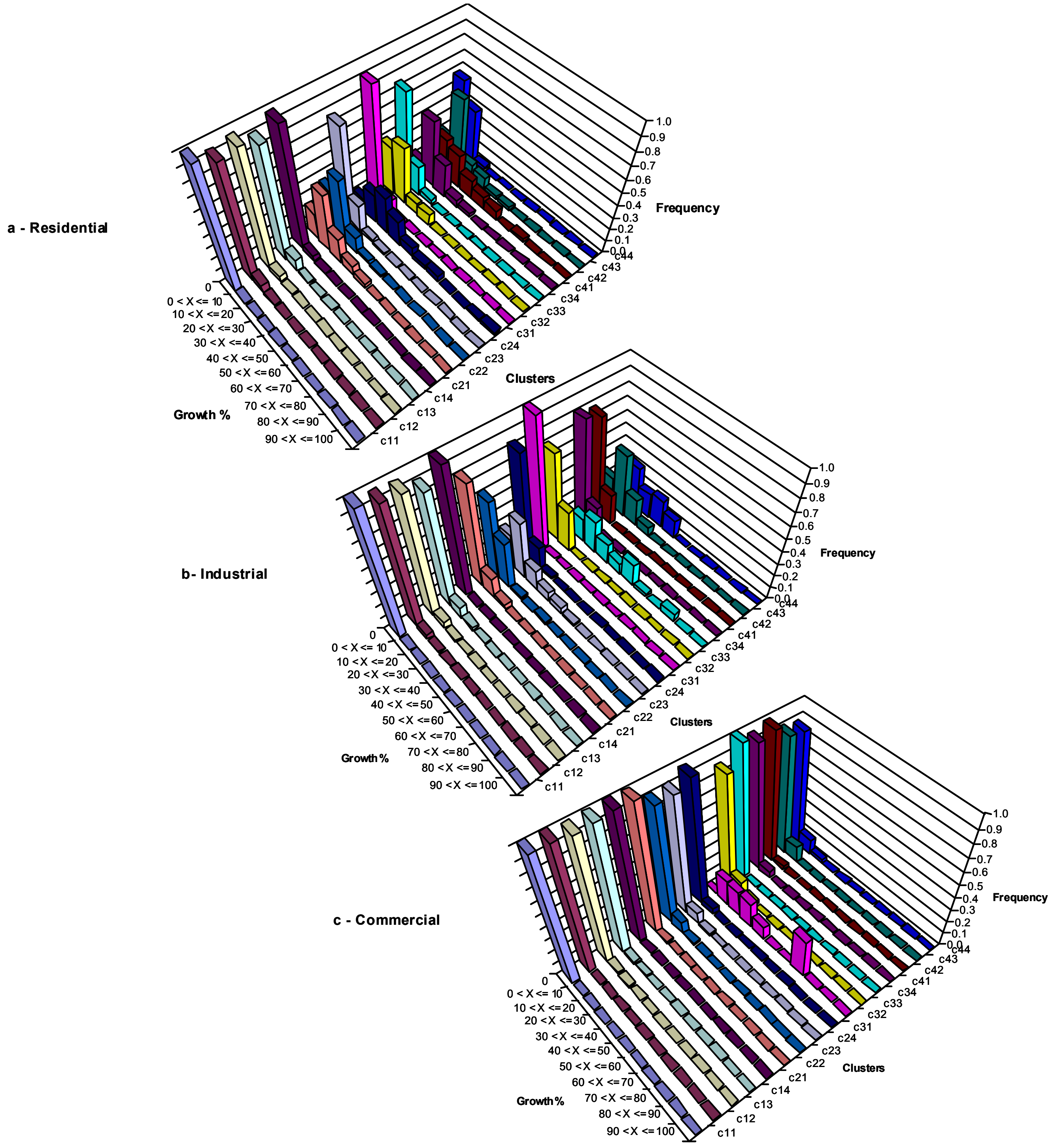


Figure 4- *The probability distributions of growth for each land use.*

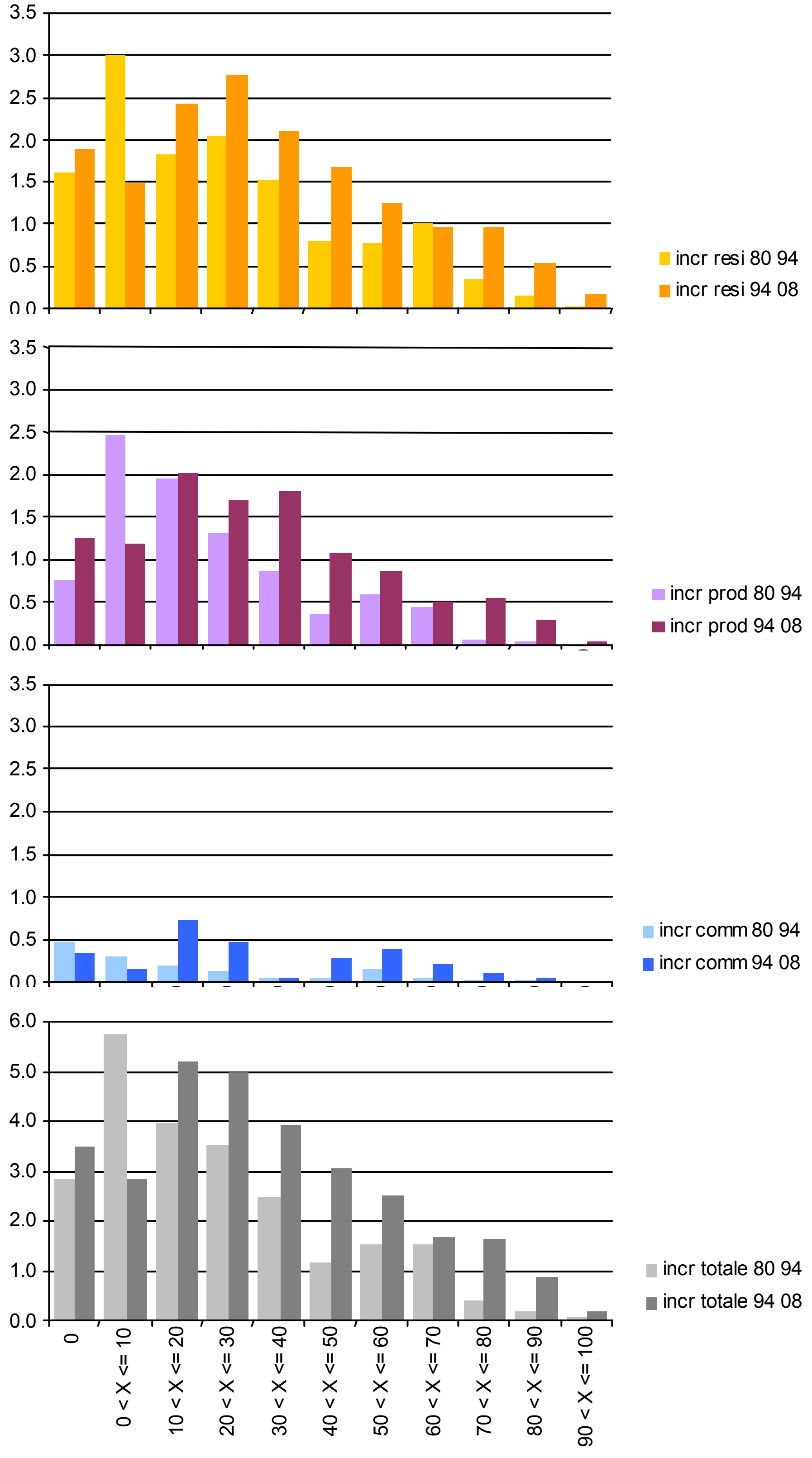


Figure 5 – *The comparison between the distributions of growth in the two periods considered.*

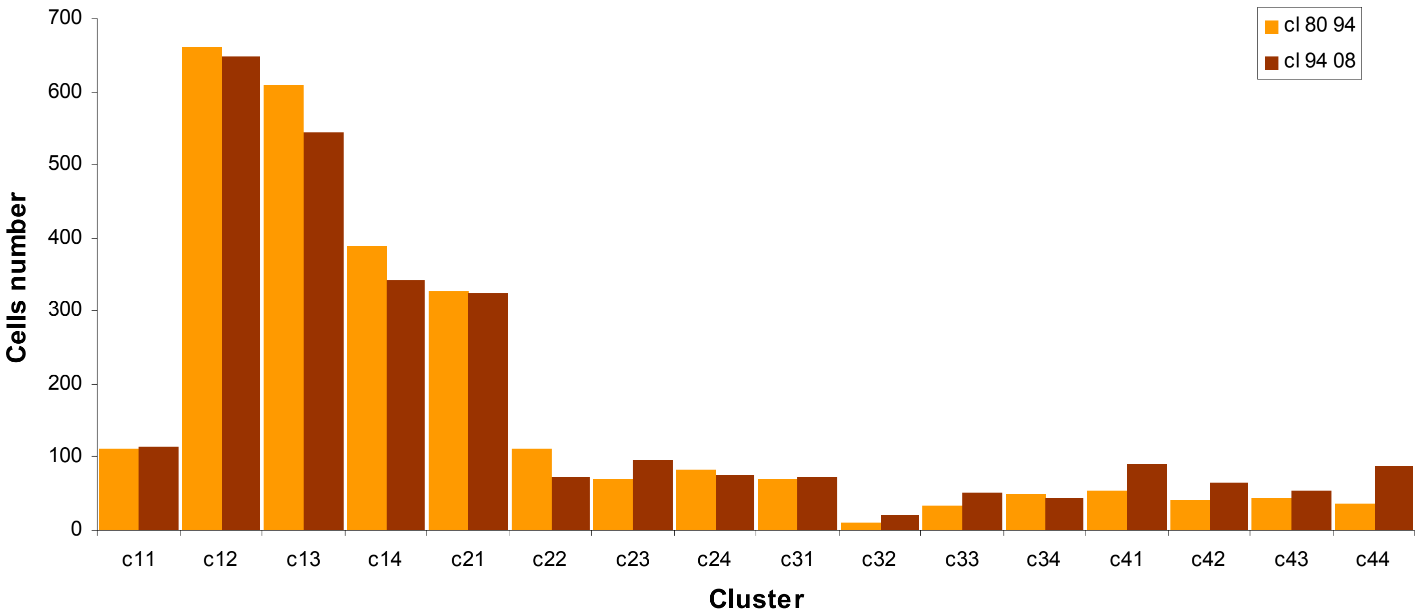
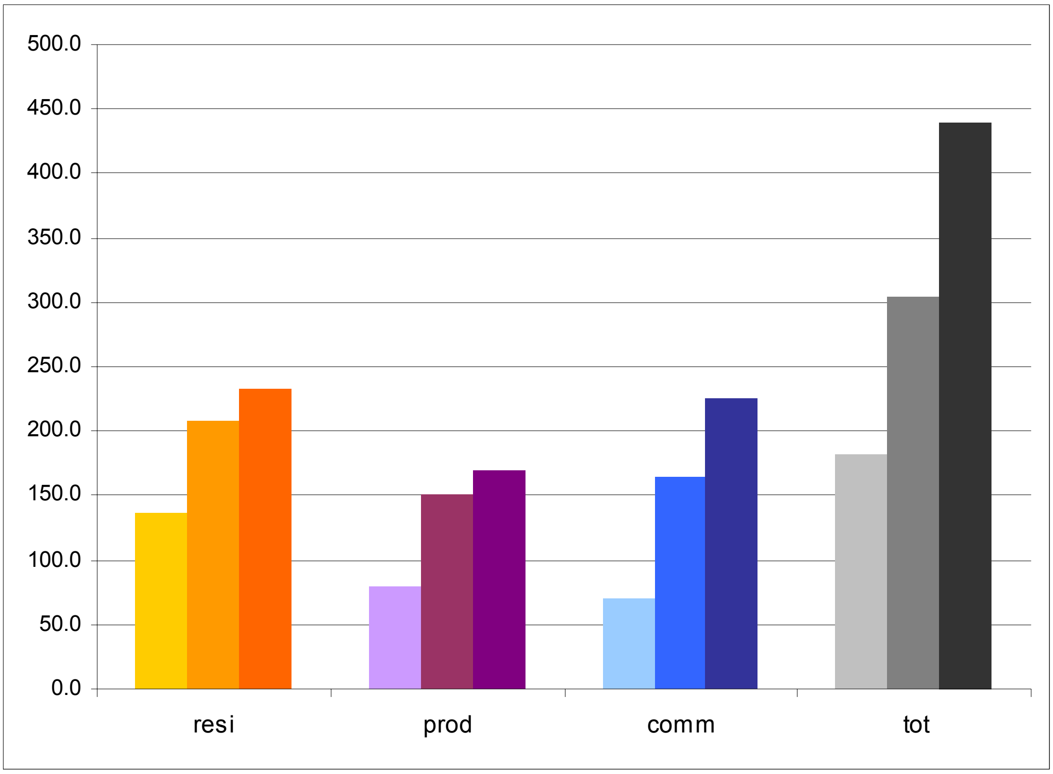


Figure 6 – *The number of cells present in each SOM class at the two temporal thresholds (1994 and 2008)*



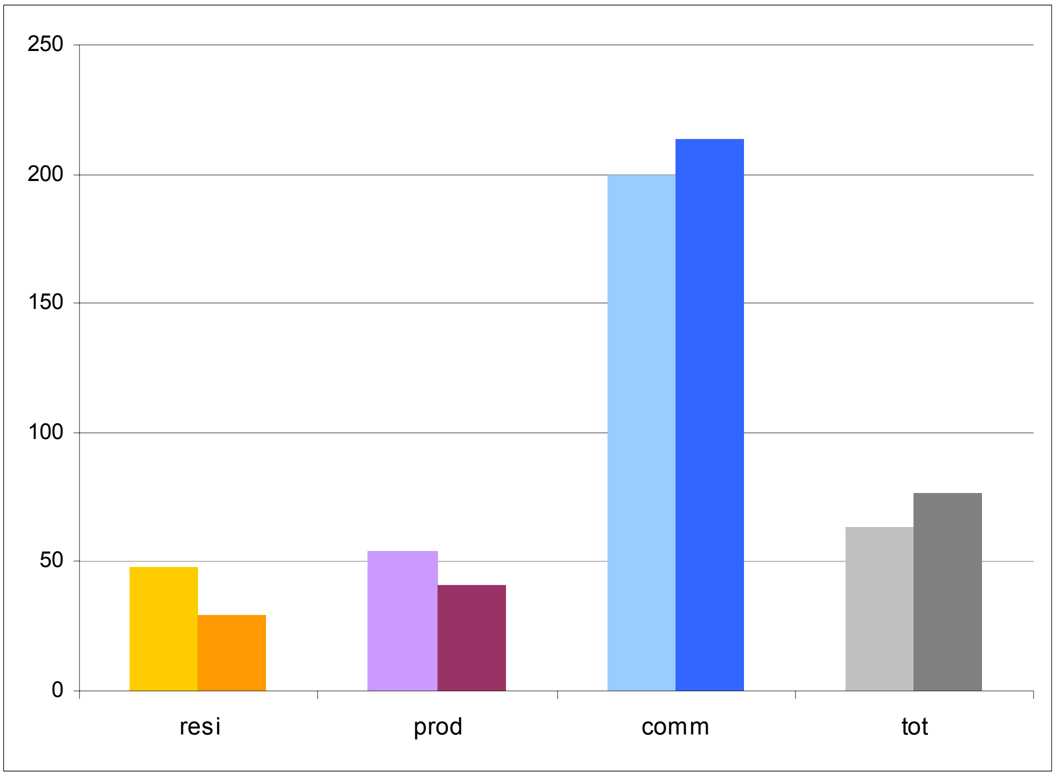


Figure 7 - *The degree of compactness first to the entire urbanised area (at the top) and then just to the new settlements (bottom).*

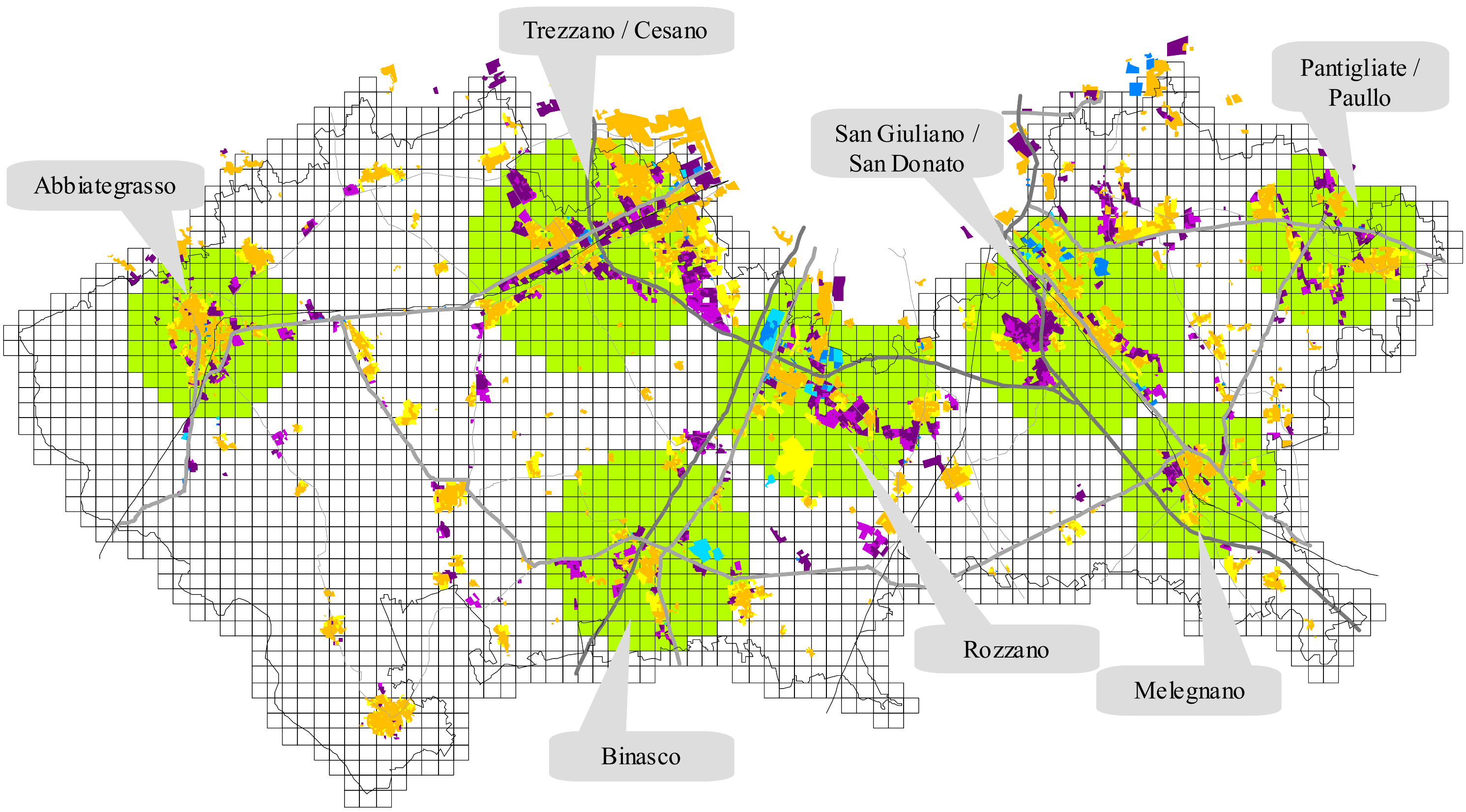


Figure 8 - *The historical centers and the circular surrounding areas which have been considered in order to measure the degree of dispersion.*

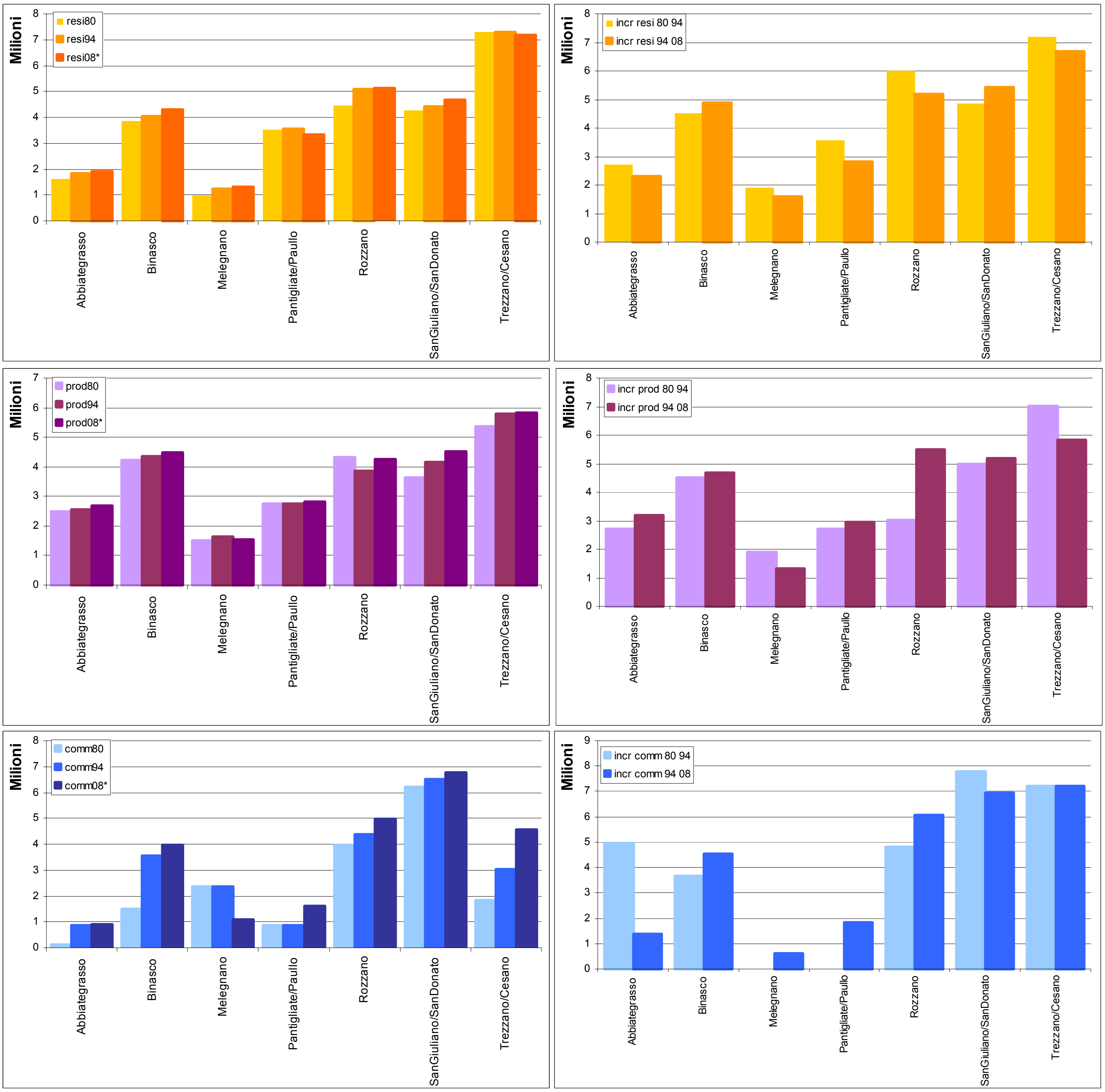


Figure 9 - *The degree of dispersion for residential (top), industrial (middle) and commercial areas (bottom)*

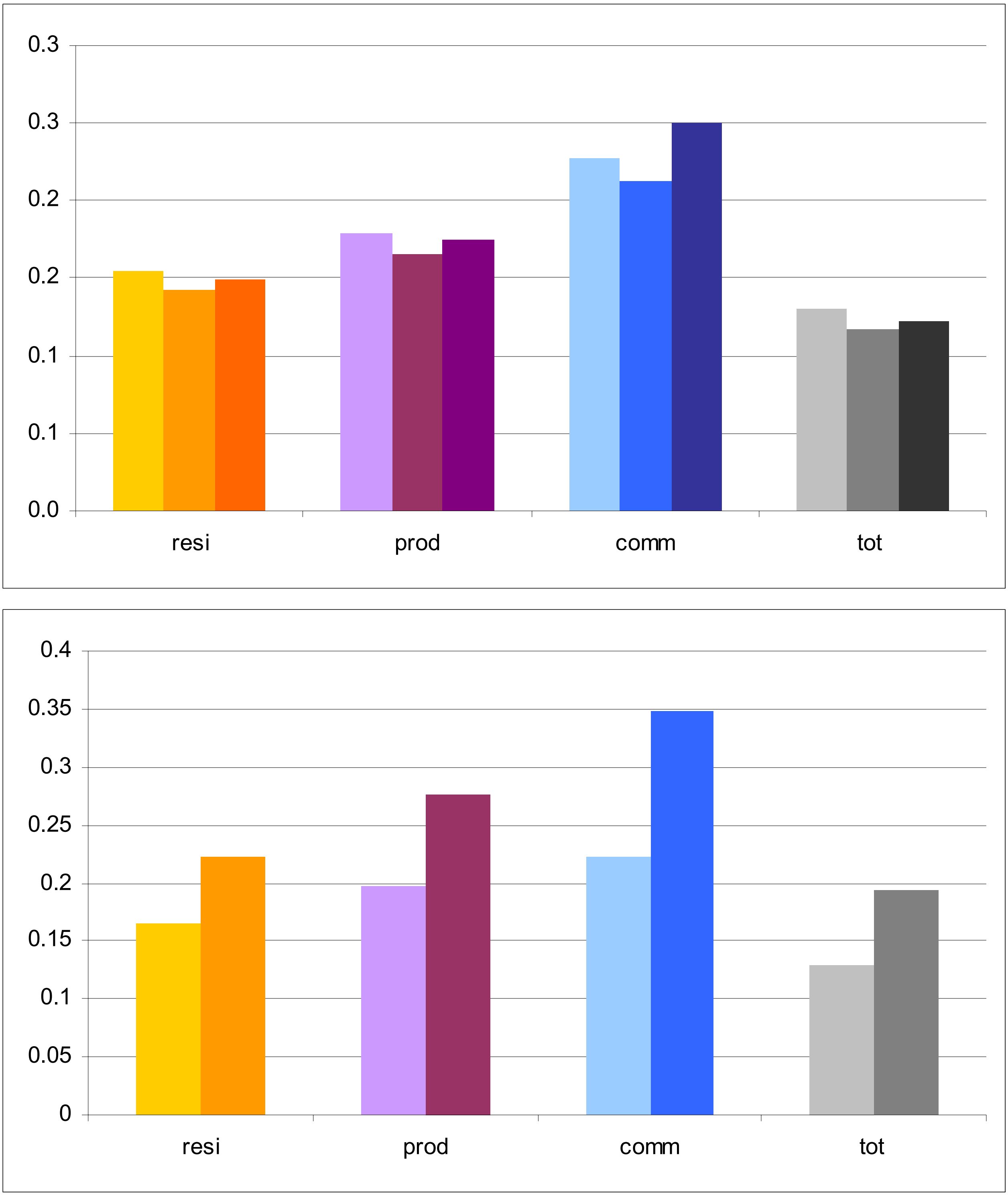


Figure 10 - *At the top the ratio perimeter/area for residential (left), industrial (center) and commercial (right) settlements at the three thresholds and , at the bottom, the index applied only to the increases.*