From the observations to the construction of a urban dynamics simulation model : an inductive approach

Des observations à la construction d'un modèle de simulation de la dynamique urbaine : une approche inductive

Cécile Tannier

cecile@lorraine.u-strasbg.fr
Laboratoire Image et Ville, Upres-a 7011 CNRS-ULP

Pierre Frankhauser

pierre.frankhauser@univ-fcomte.fr
THEMA UMR 6049 CNRS,University of Franche-Comté

Abstract : Through this paper, we chose to present a modelling concept the aim of which is the simulation of the intra-urban location dynamics. This concept includes some characteristics which relate it to the synergetic models as well as to the micro-simulation ones. In the first part of this paper, we describe and examine the modelling concept to focus in a second part on a more detailed presentation of one aspect of the model concerning the measure of the attractiveness of the neighbourhoods of a town.

Keywords : intra-urban modelling, location choices, synergetics, fuzzy sets, microsimulation

Résumé : Dans cet article, nous avons choisi de présenter un concept de modélisation dont le but est la simulation de la dynamique des localisations d'activités en milieu intra-urbain. De par ses caractéristiques, ce concept se rattache tant aux modèles synergétiques qu'à ceux de micro-simulation. Dans la première partie de cet article, nous décrivons et commentons le concept de modélisation proposé pour, dans une seconde partie, s'attacher plus particulièrement à la présentation détaillée d'un aspect du modèle, à savoir la mesure de l'attractivité des quartiers d'une ville.

Mots-clés : modélisation intra-urbaine, choix de localisation, synergétique, sous-ensembles flous, micro-simulation

Through our current research, relating to the modelling of the intra-urban location dynamics, we deal with two general objectives. The first one is to reach a better understanding concerning the urban residential, industrial and commercial location choices. More precisely, we try to extract the most information possible from different kinds of sources (quantitative or qualitative data coming from inquiries, expert statements…) ; we try to find similarities among the observations and to classify them according to their similarities. The second objective consists in improving the geographical methodological knowledge in the modelling field. We especially deal with some scale problems : relevance of the levels of analysis of the observed phenomena and connection between scales. Both objectives will be presented in this paper. However, it mainly focuses on the theoretical and methodological aspects of our researches.

The goal of the model presented here is the simulation of the urban location dynamics in order to test the sensibility of the urban system according to some given initial conditions. Such simulations could for example show that the broadening of a road would have a significant effect on the social composition of the residential area located near this road. In that way, modelling urban location dynamics appears very useful for urban planning.

The modelling concept

The model simulates the movements of agents through the neighbourhoods of a town : it deals with the intra-urban agents' movements. But, because a town is an open system, the model also takes into account the arrivals and departures of agents in the town too.
The location choices of different types of activities are considered in the model : residential, commercial and industrial ones. At each type of activities corresponds a type of agents[1]. They are :

- types of households for the residential activities,
- types of retail outlets for the commercial activities,
- types of industrial branches for the industrial activities.

Each of these three *general* types of agents is itself divided into *more detailed* types of agent (*e.g.* concerning the retail trade, we distinguish the supermarkets, the convenience stores, the DI stores…). These detailed types of agents are considered as homogeneous regarding their behaviour.
The location dynamics is a result of the variations of the number of each type of agents present in each neighbourhood.
As commonly accepted in the literature, we consider that the location process is made up of two phases : the evaluation phase, which consists of the search of a convenient site, and the phase of the choice of a site.
The evaluation phase consists of two steps. During the first step, the agents evaluate each characteristic of the neighbourhoods. The characteristics of the neighbourhoods correspond to the objective attributes of the studied space. In that sense, their evaluation can be seen as the change from the objective attributes to perceived attributes of the neighbourhoods. Only the characteristics, which play a part in the location process, are taken into account. This explains that instead of characteristics of the neighbourhoods, we may speak of location criteria. The second step of the evaluation phase consists of the evaluation of the attractiveness of each neighbourhood. For each type of agent, the perception values of the attributes are combined (aggregated) to obtain a single attractiveness measure characterising each neighbourhood. Thus, at the end of the evaluation phase, each neighbourhood is finally characterized by a set of attractiveness measures, one for each type of agent.
Following the evaluation phase, all the locations could still be chosen, even if some of them are more attractive than the others. The very decision of establishing is taken during the choice phase. The constraints, relative to the establishment of the agents in a neighbourhood (*e.g.* town planning laws), and the uncertainty relative to the representation of the location choices of a great number of agents, are introduced in the model during the choice phase.
The general architecture of the model is described on figure 1. Let us now comment on it.

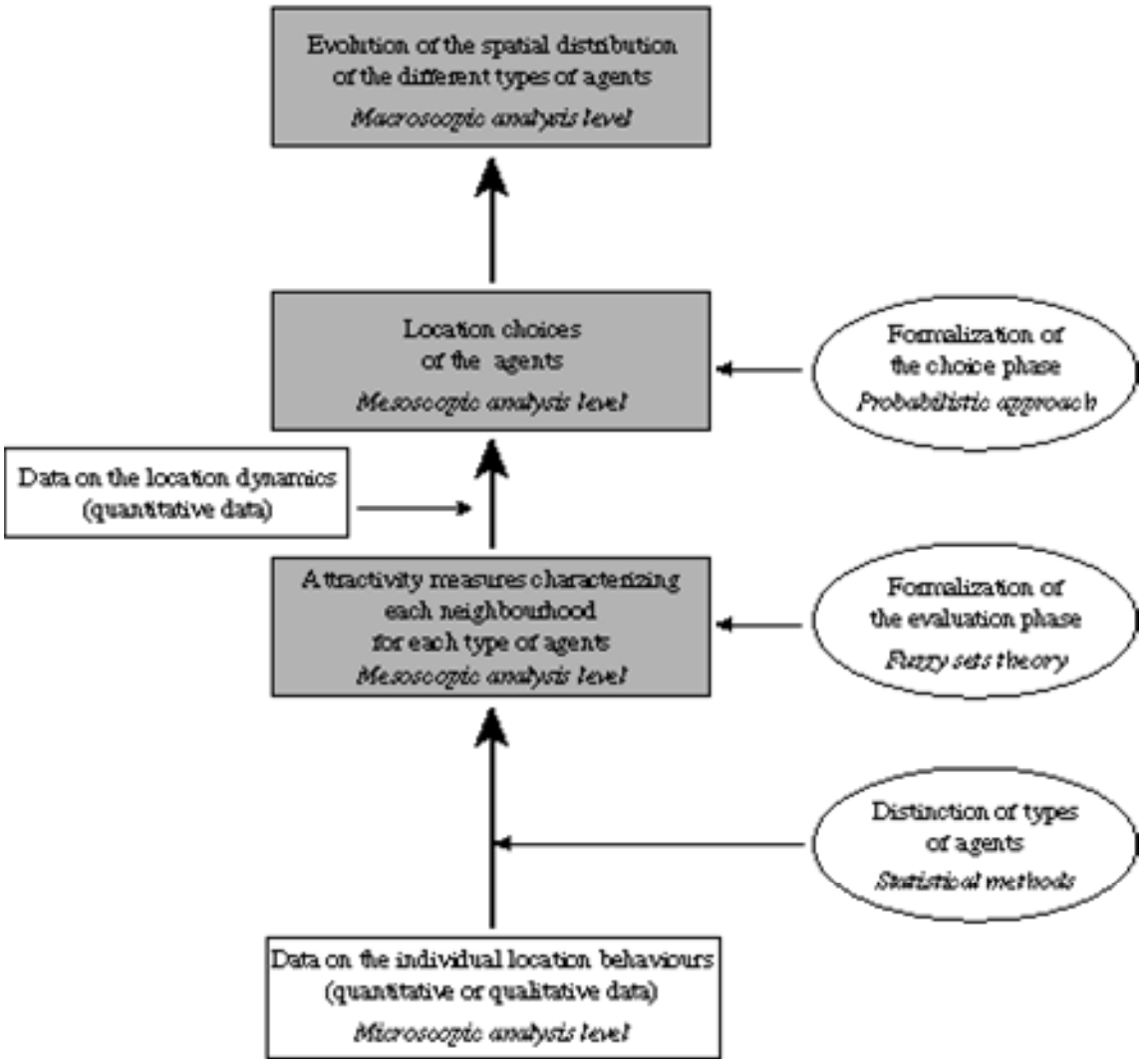

Figure 1 : General architecture of the model

Collecting data on the individual locating behaviour

Information concerning individual location behaviours may be obtained by enquiries or interviews. Because the perception of the different characteristics of space by an agent may fluctuate according to his/her state of mind, we assume that these fluctuations do not exceed the variations of the global perception of the type of agents to which he/she belongs. This corresponds to the assumption that the mean behaviour of an agent over a given time interval is equivalent to the mean behaviour of several agents belonging to the same type ("ergodicity" assumption).

The collected data should inform about the following topics :

- the location criteria that should be taken into account in the model,
- the perception of the location criteria by the agents,
- the weight of each location criterion in the evaluation of the attractiveness of the neighbourhoods.

Distinguishing types of homogeneous agents regarding their location behaviour

In order to find some types of agents without presuming the composition of these types, we chose to make use of a discriminating analysis of the empirical data. In the frame of different research projects, we are at present testing two procedures (Houot, 1999), (Pezzoli, Girerd, Frankhauser, 2000) :

- First, to identify discriminatory variables (e.g. by means of Exploratory Data Analysis). A variable is considered as discriminatory when the agents, characterized by one of the modality of the variable, show a behaviour notably different from the behaviour of agents characterized by another modality of the variable. Following the identification of discriminatory variables, types of agents can be defined, according to the modalities of the variables which characterized them.
- First, to identify the types of agents (e. g. by the way of a Hierarchical Ascending Classification ("classification ascendante hiérarchique") and secondly to find explanatory variables. The assumption is that it is possible to find common attributes such as the level of education, the type of present residential area, the age… for all agents belonging to the same type.

It will be shown later that, in the course of time, agents may change their membership to a type under certain conditions.

Formalizing the evaluation process

The use of fuzzy set theory to formalize the evaluation of the potential location sites allows us to preserve the subtleties of the agents' judgements. This supposes the introduction of a semantic graduation to represent the agents' perception of the characteristics of the neighbourhoods (*e.g.* degrees of acceptation of space characteristics, such as the distance to a service or the number of inhabitants). This procedure reminds more of the philosophy of micro-simulation. By using statistical methods such as a scaling method or the method of successive intervals (Maurin, 1984, 1986), it is possible to change the answer frequencies to each modality of a variable, obtained with inquiries, into a fuzzy graduation which represents the general perception of an attribute by a type of agents.

Through the adoption of a graduation, we implicitly assume that agents are able to organize hierarchically the different location criteria when one asks them to. But it does not mean that the agents are supposed to be rational in their location choices. On the contrary, it is well known that location choices are most often rather implicit than really founded on a comparison of the whole advantages or disadvantages of each type of urban areas. Moreover, the level of rationality of agents may vary according to their type (*e.g.* a small independent retailer is often characterized by a lower rationality level than a hypermarket when searching for a convenient site).

In order to deal with the non rationality of location choices and with the existence of different levels of rationality, we chose to base the modelling of the evaluation phase on settlement rules rather than on location strategies. It means that we consider the result of the action (*i.e.* the attractiveness of the neighbourhoods) but not the reasoning leading to this action (*i.e.* the opinion formation process).

Thus, the attractiveness measures are indicators of the regularities of the location choices which can be observed for each types of agents. More concretely, the attractiveness measures should be interpreted as follow :

Generally, an agent belonging to the type "X" will perceive the neighbourhood "Y" as very or not very attractive.

Thus, the formalization of the evaluation phase is based on the assumption that some rules exist, which govern the establishment of agents in a urban area. These rules consist of a transcription of the relation existing between the characteristics of a place and the characteristics of the agents attracted by this place.

The knowledge about the evaluation rules on the one hand, and about the attributes of the neighbourhoods, on the other hand, allows to construct synthetic indicators which describe the attractiveness of the different neighbourhoods for the different types of agents.

Formalizing the choice process

The formalization of the choice phase is based on three main assumptions.

1) The definitive decision to migrate (or respectively to open or to close a retail outlet or an industrial branch[2] in a neighbourhood) depends of the propensity of agents to realize such an intention. Thus we assume that the decision of agents depends simultaneously on two independent factors :

- the general propensity for moving or for creating/closing down a shop or an industrial branch. This translates e.g. the intention of agents to leave their actual neighbourhood since they prefer another type residential environment. We describe this phenomenon by the way of a mobility measure ;
- the attractiveness of the neighbourhoods.

2) The individual decisions are taken independently. This does not mean that the evaluation of the neighbourhoods by an agent or his/her mobility would not be influenced by the behaviour of the other agents, but we assume that this influence acts in a global way on each agent. Indeed, the attractiveness measures include the interaction between agents in the sense that every agent is also influenced when forming his opinion by the other ones. From an epistemological point of view, such a reasoning is very close to the notion of a "mean field" used in physical sciences which assumes that all influences exert by all the particles on a chosen one may be approximated by one global influence function.

3) The mobility varies with respect to the types of agents considered. For example, in the case of residential location choices, the mobility of the young singles is usually higher than the mobility of the households made up of several children. In the case of retail trade, the turn-over varies with respect to the type of retail outlet (convenience store, supermarket…).

To formalize the location choices of the agents, we accepted a probabilistic approach. Indeed, if all agents belonging to a given type would exactly have the same behaviour we could then introduce deterministic rules which govern the dynamics of the spatial system. In this case, it would be sufficient to know the mobility of each type of agents and the attractiveness of the different neighbourhoods. However, we are aware that on the one hand, the agents belonging to a same type may have a slightly different perception of the neighbourhoods, and on the other hand, that the decisions may be influenced by individual experiences (*e.g.* an agent may finally favour a particular neighbourhood since some of his/her friends or members of his/her family live there). So, it seems to us more realistic to recur to a probabilistic approach for representing the choice phase.

In order to decide if we introduce whether a time continuous or a time discrete approach, we considered the following observations.

- At a microscopic analysis level, individual decisions are taken at a particular moment and not over a time period.
- the set of individual decisions defines the dynamics on an aggregated level and thus macro-dynamics is finally the result of a sequence of individual decisions. If the number of decisions is high enough and if the decision processes are continuously distributed over a time period, it is possible to approximate this sequence by a continuous function.
- But in fact, data are available only for certain dates.
- Another more technical argument should be taken into account : computer simulations are based on sequences of discrete steps.

Thus, we decided to rely on a discrete modelling concept. Since we use a probabilistic approach, this implies that the formal description of the time evolution should be done by the way of conditional probabilities which link the state of a system at a given time t to that one at a time $t+\Delta t$.

As example, let us consider the migratory process. We introduce a conditional probability $W_i(Z_K,t+\Delta t|Z_J,t)$ for an agent belonging to a certain type i to migrate within the time range Δt from a neighbourhood characterized by an attribute vector Z_J to one characterized by an attribute vector Z_K in the following way

$$W_i(Z_K, t+\Delta t | Z_J, t) = \omega_i \cdot f(a(Z_K), a(Z_J)) \qquad (1)$$[3]

where : ω_i is the mobility of the agents belonging to type *i*
and : $f(a(Z_K),a(Z_J))$ expresses the influence of the attractiveness values $a(Z_J)$ and $a(Z_K)$ of the neighbourhoods Z_J and Z_K on the location choice.
This formalization traduces the assumptions explained previously :

- the decisions are taken independently, i.e. the decision of an agent is not influenced by the state of the other neighbourhoods (their amenities, their accessibility, etc.) or by other individuals. Thus, no reference to other agents than the considered agent or other neighbourhoods than the neighbourhoods ZJ and ZK appears in the formula.
- the decision process is not explicitly modelized ; we only assume that with respect to a certain propensity to move and to the attractiveness values of the neighbourhoods we have a certain probability that such an agent will really leave ZJ to ZK. This traduces that to represent the urban dynamics we do not find useful to consider the agents' behaviour on the very individual scale but on the aggregated level of types.
- the mobility and the attractiveness terms are linked by a product. This refers to the assumption that the evaluation of the neighbourhoods described by the term f(a(ZK),a(ZJ)) is independent of the propensity to move.

After having explained the singular process of the migration of one agent over a given time step, we will now deal with the sequence of decisions of all migrating agents on a larger time scale. The distribution of agents among the different neighbourhoods depends on the migratory decisions of all the agents within a given period and on the sequence of migratory decision taken by each agent in course of his/her existence.
Concerning the sequence of decisions of each agent in course of time, we look at it from the point of view of the Markoff assumption : agents belonging to type *i* and considering the migration from Z_K to Z_J, will always decide in the same way. Thus, the individual history of agents, *i.e.* their previous experiences do not influence their evaluation of the attractiveness of the neighbourhood and their mobility. We justify this first by the fact that we assume we have a sufficiently stable behaviour within the types introduced, secondly because we affect to another type for the next time period agents who are expected to have changed their behaviour in the previous time period. For example, by knowing the ratio of people living in Z_J, and belonging to a certain age class, we may affect them to another type when we may assume that most of them are now belonging to another age class. Concretely, we introduce a demographic model, which changes automatically the numbers of agents belonging to each types according to the previsions of this model. For the households, the model is based on their life cycle (Courgeau, 1984, 1988). For the other types of agents (retail outlets and industrial branches), the demographic model is constructed on the basis of the presumed evolution of each type of retail outlets or industrial branches over the considered time interval. Thus, the emergence of new types of agents can be simulated. Other changes should also be taken into account by a probabilistic approach. For example, we intend to introduce probabilities for certain social changes which can influence migratory behaviour like the risk to loose the job or to divorce.
Let us now come back to the fact that we consider now not only one single agent but the whole population. Since agents are supposed to be independent in their decisions, we just have to introduce in the previous formula a weighting factor representing the number of agents of types *i* living in the neighbourhood Z_J, (origin of the migration). Indeed the migration flows are just increased by the fact that now $n_i(Z_J,)$ agents are expected to migrate instead of solely one. The "loss of memory" about individual experiences of agents allows us to formalize the time evolution of our system by a Markov-chain which can now be formulated in the following way :

$$P_{t+\Delta t}(\underline{N}) = \sum_{i=1}^{n} \sum_{Z_K} (n_{i(Z_J)} + 1) \cdot W_i(Z_K | Z_J) \cdot P_t(\underline{N}^{n_{i(Z_J)}}) \qquad (2)$$

… with *L* corresponding to the number of the neighbourhoods of the town.
… with Z_J (respectively Z_K) representing the neighbourhoods considered as origin (respectively destination) of the migrations. Each neighbourhood may be at once origin and destination of the migrations.
Here we have introduced the probabilities $P_{t+\Delta t}$ to find a macroconfiguration
$\underline{N}=(n_1(Z_1),n_1(Z_1)\ldots n_2(Z_2)\ldots)$
$P_{t+\Delta t}$ is identical to P_t excepted one agent who lived in Z_J at time *t* and living in Z_K at time $t+\Delta t$.
Due to the different independence assumptions introduced above, we may consider this probability as the frequency to find simultaneously $n_1(Z_1)$ agents belonging to type *1* in the neighbourhood Z_1 and $n_1(Z_2)$ agents belonging to type *1* in the neighbourhood Z_2…
To determine the transition probability $W_i(Z_K|Z_J)$ we have to calibrate the functions ω_i and $f(a(Z_K),a(Z_J,))$ (see formula (1)).
Let us first consider the mobility ω_i. We directly go to the total number of migration processes observed for group *i* and consider that :

$$m_i = \frac{1}{L}\sum_{\alpha=1}^{L}\sum_{\beta=1}^{L} . W_i^{emp}(Z_K|Z_J) . (1-\delta_{(\alpha,\alpha)}) \qquad (3)$$

with $\delta_{(\alpha,\alpha)}=1$ if $Z_K=Z_J$ and $\delta_{(\alpha,\alpha)}=0$ if $Z_K=Z_J$

… where $W_i^{emp}(Z_K|Z_J)$ is the empirically observed number of migration between the neighbourhoods Z_J and Z_K. This number can stem from census or statistics about the creation or closing down of retail outlets or industrial branches.
Let us now consider the calibration of the attractiveness function. We make no theoretical assumption about the form of the function $f(a(Z_K),a(Z_J))$ but we compare the empirical migration frequencies with the attractiveness measures. Indeed, for one type of agents, if the attractiveness measure of the neighbourhood Z_K is superior to that of the neighbourhood Z_J, then the inflow frequencies of agents of this type in the neighbourhood Z_K should be higher than the establishment frequencies in Z_J. Then, by comparing the migration data concerning the neighbourhoods of a town with the attractiveness measures, we can find the relation $f(a(Z_K),a(Z_J))$. The advantage of using such a procedure is that if only a reduced set of migration data are available, we can even so estimate the function.
We want now to recapitulate the course of a simulation. At the beginning, the initial population distribution is considered as a frequencies distribution. On this basis, the model calculates the initial migration probabilities P_t. Then, by means of the migration frequencies $W_i(Z_K,t+\Delta t|Z_J,t)$, it calculates the results obtained for the next time steps. However, as we would obtain after some steps a great number of possible population distributions, we only keep at each step of simulation the configuration $\underline{N}=(n_1(Z_1),n_1(Z_1)\ldots n_2(Z_2)\ldots)$ characterized by the highest probabilities. Thus, the probability distribution is truncated… But bifurcations can even so be detected : indeed, such phenomena should occur when different configurations are characterized by almost the same probabilities.
After each step of simulation, the number of agents of each type located in each neighbourhood may have changed. These changes may affect the attribute vectors of the neighbourhoods and thus, their attractiveness can be modified. For example, if a supermarket establishes itself in a neighbourhood, the attractiveness of this neighbourhood for another supermarket will decrease. So, the interactions between the agents and their territory are taken into account by our modelling concept and appears with the running of the model in course of time.
A simulation corresponds to the evolution of the modelized system during about 20 years. The number of time steps is not predefined, but depends on the observed behaviour of the agents. If a great number of migrations happen during the course of a simulation, a step should correspond to a smaller time interval than if the number of migrations is small.

Some comments about the modelling concept

The model represents the behaviour of types of agents, which is a mesoscopic analysis level but (ahead of) it is fed by observations collected at a microscopic analysis level. These micro-level observations are used to represent the evaluation of the attractiveness of the neighbourhoods (mesoscopic analysis level). On the other hand, the output of the model is the settlement pattern evolution, which corresponds to a macroscopic level of observation, that is a global point of view on the location dynamics of a town. So, we see that the modelling concept tends to link different analysis levels.

The topic of the link between different analysis and representation levels of socio-economic phenomena has been recently discussed by L. Sanders (1999). In her paper, she focuses particularly on the differences and the complementary aspects between some meso-level modelling approaches, like synergetics, and models, which refer to the individual behaviour, like micro-simulation. Three months ago, N. Winder (2000) published a response to her paper and showed that microsimulation models such as ours come within the thermodynamic framework. To support his assertion, let us expose some features of synergetics.

As pointed out by L. Sanders, synergetic models, as they are generally used in social sciences, especially by W. Weidlich and G. Haag (1988), consider the urban dynamics at an aggregated level. Contrarily to micro-simulation approaches, which follow a bottom-up logic, synergetic models seem rather to correspond to a top-down logic. A typical example of this point of view is the notion of attractiveness (or "utility") as it has been used in the Weidlich-Haag model. Each spatial unit is characterized by an attractiveness measure. The attractiveness measures are estimated using migration data. Only *a posteriori* these "trend parameters" are linked to socio-economic data by means of a ranking regression analysis : the attractiveness measures are explained *a posteriori* by the socio-economic data, which are then interpreted as causes for the migration flows.

If we let aside the geographical applications of synergetics and if we consider the physical origins of the concept (Haken, 1978), we notice that the starting point of the synergetic approach was a deep quantitative knowledge about the processes peculiar to the microscopic level. Indeed, the initial goal of synergetics was to understand how the emergence of a macroscopic system, showing a high degree of order, may be explained by the microscopic behaviours.

For example, the development of the laser theory was based on a deep knowledge concerning the dynamics both of the particles and the electromagnetic field as well as the interactions between the particles and the electromagnetic field. Physicists have observed that when a light beam activates the particles, some of them may emit randomly a very coherent light (under certain conditions which refer to a critical threshold). Then, the other particles react quickly on the relatively slow changes of the amplitude of the light emitted by the first particles and emit in return a highly coherent light. This process generates the laser light, which corresponds to a highly ordered structure on a macroscopic scale. The fact that the particles act in a common way has inspired physicists to speak of a "co-operative phenomenon". This co-operation is related to an "enslaving" behaviour : the slowly varying light amplitude dictates in some sense how particles have to emit their light. The light emitted by the particles may be interpreted as the environment that they create themselves and which conditions their further behaviour.

The example of the birth of the laser theory illustrates the fact that in the original version, the concept of synergetics combines explicitly both the micro- and the meso- (or macro-) levels and that the crucial reflection focuses on the interactions between these levels. Following this point of view, the modelling concept presented in this paper is closer to the original idea of synergetics than to other geographical applications, especially those of Weidlich and Haag.

Certainly, in social sciences the interactions between agents are more complex than in the laser case. In particular, as it was presented by D. Pumain (1997), socio-economic dynamics are driven by individual decisions, which refer to personal objectives. The fact is that these processes are until now not sufficiently known to propose a detailed modelling and the partial knowledge about the individual location choices has incited Weidlich and Haag not to take into account explicitly microscopic behaviours. Their choice was confirmed by the relatively weak calculation capacities of the computers at this time and by the stability of the macro-scale. However, our point of view, shared by some other geographers and especially (Winder, 2000), is that it is possible to acquire enough information about the agents' location behaviour to describe the settlement dynamics on the aggregated level of neighbourhoods, that is to work at the mesoscopic analysis level.

Focus on the formalization of the evaluation phase

We saw previously that :

- each neighbourhood is characterized by some attributes (population, accessibility…) ;
- each attribute is perceived differently according to the type of agents ;
- the weight4 of each attribute in the final decision depends on the type of agents ;
- the evaluation of the attractiveness of a neighbourhood results of a combination of the couples (perception of an attribute ; weight of this attribute).

For the present paper, we chose to develop only the aspects relative to the formalization of the perception and of the combination of the couples (perception ; weight). How do we represent the weight of the attributes to calculate the attractiveness of the neighbourhoods is not detailed here. We just want to precise that the weight of a location criterion represents the influence of this criterion on the global attractiveness of the areas. For further explanations about this topic, you can look in (Yager, 1978), (Zimmermann, 1987) and (Tannier, 2000).

How do we represent the perception of an attribute ?

The model integrates a deterministic relation between a type of agents and its perception of the attributes of the neighbourhoods. To formalize this relation we use fuzzy variables, which are expressed as follows.

Let $\tilde{A}$ be a fuzzy set and R its reference set, the fuzzy variable *a* is defined as :

$\tilde{A} \in [0 , 1]$ and $a = \mu_{\tilde{A}}(x)$

The membership function $\mu_{\tilde{A}}$ assigns to any element of R a membership degree to the fuzzy set $\tilde{A}$, $\tilde{A} \subseteq R$. If we consider the reference set *"quality of the landscape"*, a neighbourhood *x* can be characterized by a membership degree to the fuzzy set *"good quality of landscape"*. If the quality of the landscape is neither really good nor really bad, it will be possible to write : $\mu_{\tilde{A}}(x) = 0,5$.

Thus, the determination of the membership functions can be considered as the assignation of numbers to some objects, so that the defined relations between the numbers reflect analogous relations existing between the objects. In other words, a membership function consists in a transcription of the relations between some objects in the form of same types of numerical relations (*cf.* figure 2).

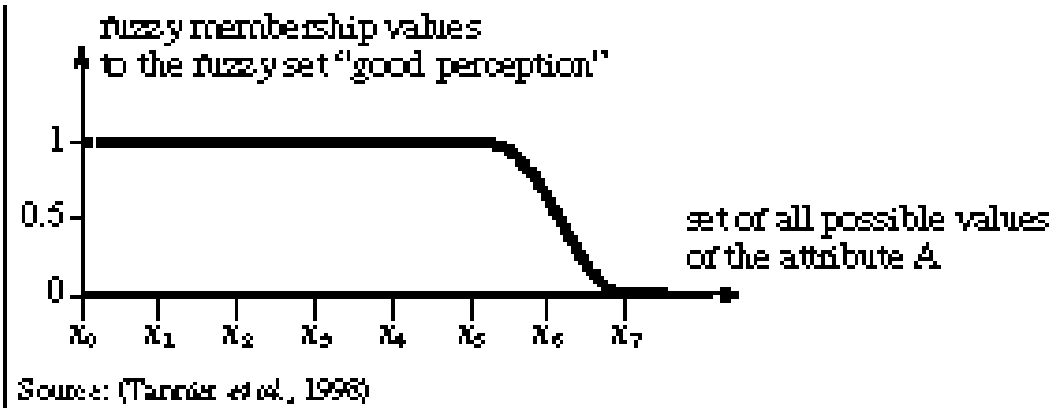

Figure 2 : A member ship function for an attribute A

In future, we intend to formalize the perception of an attribute by the way of a linguistic variable. The definition of a linguistic variable is more general than the one presented previously. It deals with one reference set R, but with several membership functions to a finite number of fuzzy sets belonging to R (*cf.* figure 3). Thus, the use of linguistic variables will permit a better taking into account of vagueness inherent in the studied phenomena.

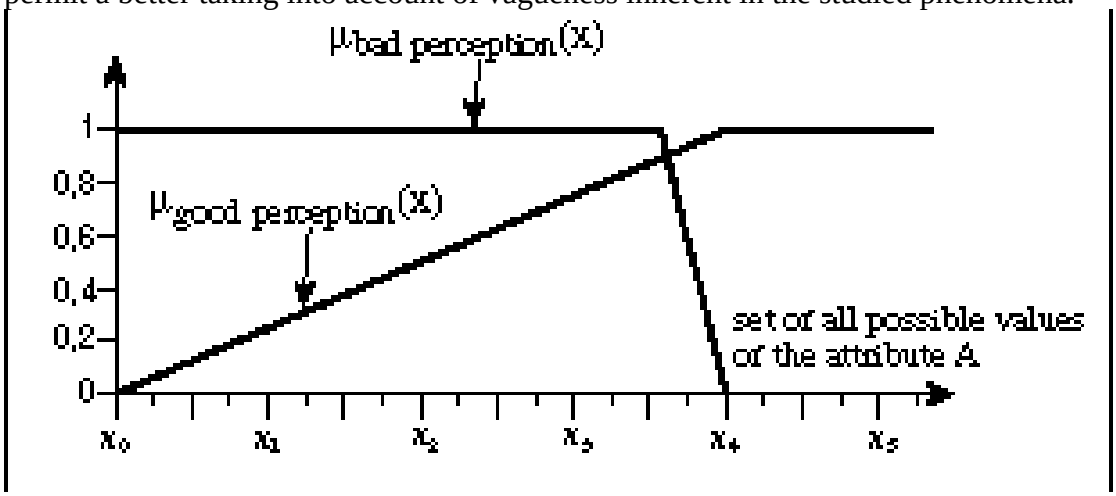

Figure 3 : Formalization of the perception of an attribute using a linguistic variable made up of two fuzzy sets

Combining both the perception of an attribute and its weight, we may qualify each attribute by a *partial attractiveness*. The agent has yet to develop a synthetic appreciation of the neighbourhoods by taking into account the partial attractiveness measures of the different attributes.

How can we obtain an attractiveness measure, coherent with the empirical observations ?

Our aim is to obtain a synthetic measure of the attractiveness of the neighbourhoods for each type of agents. So, we have to search for an aggregation function that represents in a convenient way the combination of the partial attractiveness measures for each type of agents.

When the agents evaluate the attractiveness of a neighbourhood, they arbitrate between its positive and negative aspects. For example, in the case of residential location choices, the agents will eventually be ready to accept a higher land price if the quality of the landscape and the accessibility are extremely good. So, the measure of the attractiveness of the neighbourhoods requires to deal with compensation phenomena : agents may make up for negative aspects of a neighbourhood with the positive ones[5]. The main difficulty for taking into account such phenomena lies in the fact that compensation does not play the same role for all types of decision. Some examples can illustrate this assertion.

First example : evaluation of the accessibility of a neighbourhood

If the average access time is judged worse than the mean real distance, time distance prevails over real distance : agents do not make up for time distance with distance in kilometres, or respectively for distance in kilometres with time distance. The degree of compensation is null.

Second example : evaluation of the amenities of a neighbourhood

If the distance to a school centre is evaluated by a rather low degree of satisfaction, this may be fully compensated by the existence of an effective school bus system. The degree of compensation is high.

Through these two examples, we can see that the degree of compensation corresponds to the degree of optimism (or pessimism) of the evaluation of the neighbourhoods by a type of agents. In the case of an optimistic evaluation, the bad perceived criteria are fully compensated by the good ones : the attractiveness measures are then closer to the good perception values than to the bad ones. On the contrary, if an agent is characterized by a pessimistic evaluation the attractiveness measures are closer to the bad perception values than to the good ones.

In fact, the degree of compensation (*i.e.* the degree of optimism or pessimism of the evaluation) varies not only according to the objectives of the agents : it varies also according to the characteristics of the set of all the partial attractiveness measures. More precisely, we observed that the degree of compensation vary with respect to :

- the range of the set of the partial attractiveness measures ; For example, if the range of the set of the partial attractiveness measures is small, it means that all the partial attractiveness measures of a neighbourhood are almost the same. Thus, we may expect that the agents would accept some disadvantages (translated by some low partial attractiveness measures) because these lower measures are not too much worse than the maximal attractiveness measures of the set. On the contrary, if some partial attractiveness measures are extremely high and some other extremely low, an agent may evaluate this neighbourhood in a more pessimistic way.
- the partial attractiveness measures themselves. For example, an agent may be more optimistic when the partial attractiveness measures are globally high than when they are close to the mean, or vice versa.

What can we conclude from all these remarks ? Firstly, because both the degree of compensation and the very nature of the aggregation may vary, we have to introduce in the model as many aggregation functions as the number of the observed situations. As a situation we mean a way of evaluating the attractiveness of a neighbourhood (optimistic or pessimistic evaluation, degree of optimism varying according to some features of the set of the partial attractiveness measures). We can notice here that one aggregation function can fit in with several types of agents, but different aggregation functions can also correspond with a single type of agents depending on the context of the evaluation.

Secondly, in order to choose an aggregation function that transcribes in a convenient manner the evaluation of the agents, we have to compare a set of aggregation functions on the basis of :

- the global degree of optimism (or pessimism) of their aggregation,
- the nature of their aggregation (that is with respect to a given global degree of optimism, how this degree varies according to the set of the values that will be aggregated).

In this paper, we will not linger over the question of the choice of the aggregation functions that should be introduced in the model[6]. On this subject, we only want to precise that the selection of one operator rather than another is based on the empirical observations (*i.e.* the data collected on the location behaviour of the agents, *c.f.* figure 1).

We will now focus on the step of work preceding the choice of the aggregation functions, that is the comparison of the aggregation properties of some operators.

To begin, we have to precise that the compared operators must be suitable for the fuzzy nature of the partial attractiveness measures. A great number of such operators already exist. The simplest ones were proposed by L.A. Zadeh at the time of the introduction of the concept of fuzzy sets. They are the *MIN* and the *MAX* operators.
With the *MIN* operator, the result of the aggregation is equal to the lowest of the partial attractiveness measures aggregated. Consider for example a neighbourhood characterized by a partial attractiveness equal to 0.9 for the quality of the landscape and a partial attractiveness equal to 0.4 for the land price. Using the *MIN* operator, the attractiveness of this neighbourhood is :
MIN (0.4 , 0.9) = 0.4
Through this example, we see that the quality of the landscape does absolutely not compensate the badly appreciated land price : the *MIN* operator corresponds to a totally pessimistic evaluation of the attractiveness of a neighbourhood.
Contrary to the *MIN* operator, the *MAX* is the expression of a total compensation. If nine partial attractiveness measures are equal to 0 and only one measure equal to 1, the result of the aggregation will be after all equal to 1.
MAX (0 , 1) = 1
Some previous studies have showed that when the agents evaluate a zone the compensation is most often neither null nor total (Zimmermann and Zysno, 1983). It means that the operators that would be commonly used in our model have to unable to obtain attractiveness measures, the values of which are contained between the results given by the *MIN* and the *MAX* operators. However, if a level of compensation contained between the *MIN* and the *MAX* is the more widespread, it could happen in some particular cases, that the result of the aggregation would b inferior or equal to the *MIN* or superior or equal to the *MAX*. For example, we may suppose that if all the partial attractiveness measures are very low, the result could be inferior to the *MIN*. Only the observation of the real location behaviours will show if such situations could or not exist. This is the reason for not excluding the operators giving results below the *MIN* and over the *MAX* at the time of the theoretical comparison of their aggregation properties.

To be integrate in our model, an aggregation operator must satisfy the following conditions:

- It must be characterized by a commutative nature because the order according to which the partial attractiveness values are taken into account does not play a part in the aggregation.
- It must give results contained between [0,1] because this property is a condition to remain in the application field of the fuzzy sets theory.

But, according to the type of behaviour modelized:

- the degree of pessimism or optimism (*i.e.* the level of compensation) of the operator should or not be the same,
- the operator must or not be usable for 0,
- 0 must or not represent an absorbing element for the operator,
- the operator must or not have the associativity property,
- the results of the aggregation must or not be contained between the MIN and the MAX,
- the operator must or not be symmetrical in relation to the compensation,
- the level of compensation must or not be the same whatever the values of the partial attractiveness (low or high),
- the level of compensation must or not be the same whatever the standard deviation between the values of the partial attractiveness,
- the level of compensation must or not be the same whatever the number of partial attractiveness values that are aggregated,
- the operator must or not allow a control of the level of compensation.

Taken from (Tannier, 2000)

Figure 4 : Pratical comparison schedule of aggregation operators

THE PROPERTY	...MEANS THAT...	*Example for 3 partial attractiveness measures p_A, p_B et p_C*
the operator cannot be used for 0	It is impossible to take into account partial attractiveness values equal to 0. So, with such an operator the partial attractiveness values must be contained between]0,1].	
0 represents an absorbing element for the operator	If only one partial attractiveness value is equal to 0, the result of the aggregation is also equal to 0. This property allows to introduce a condition in the aggregation.	*If p_A=0 p_B=0.6 et p_C=0.3* *The result of the aggregation will be equal to 0.*
the operator is symmetrical in relation to the compensation	The result of the aggregation is situated at the point, which minimizes the distance between this result and all the aggregated partial attractiveness values. This property is synonymous with perfect neutrality of the compensation. It leads to results situated at the equal distance between the MIN and the MAX, the pessimism and the optimism.	*If p_A=0.1 p_B=0.5 et p_C=0.3* *The result of the aggregation will be equal to 0.3.* *This result R minimizes all the distances :* *$(R-p_A)$, $(R-p_B)$ and $(R-p_C)$*
associativity	The fact that the partial attractiveness values are aggregated in groups of 3 or 4, then all together or from the very start all together has no consequence on the result of the aggregation. This property does not permit to deal with a really meaningful hierarchical aggregation.	*Let $\circ$ be an associative operator* *$\forall p_A, p_B$ and p_C* $[p_A \circ p_B] \circ p_C =$ $p_A \circ [p_B \circ p_C] =$ $p_A \circ p_B \circ p_C$

Taken from (Fournier, 2000)

Figure 5 :Practical meaning of some of the previously presented properties of the aggregation operators

Figure 4 shows a practical comparison schedule of the aggregation operators, that could be used to formalize the evaluation phase. To build this schedule, we had to identify the property of the operators that must be considered to determine their aggregation behaviour and also to understand the practical meaning of some mathematical properties of the operators (*cf.* figure 5). So, the figure 4 explains what may vary according to the type of agents when they evaluate the attractiveness of some neighbourhoods on the basis of their partial attractiveness measures, whereas figure 5 deals with the problem of the translation of the agents' behaviour into mathematical properties.

Different types of graphics permit to visualize the results of the aggregation made by an operator (*cf.* figures 6 and 7[7]). The interest of such representations is that they allow to observe and compare the aggregation properties of the operators on a global point of view, when the comparison schedule (*cf.* figure 4) consists of an analytically approach of them. With the surface graphics (*cf.* figure 6), we have a synthetic image of the aggregation made by an operator. The concave surfaces are obtained with the operators characterized by an optimistic aggregation whereas the convex surfaces correspond to a pessimistic evaluation.

The histograms give a more detailed image of an aggregation (the coordinates of the inflexion points can be read) (*cf.* figure 7).

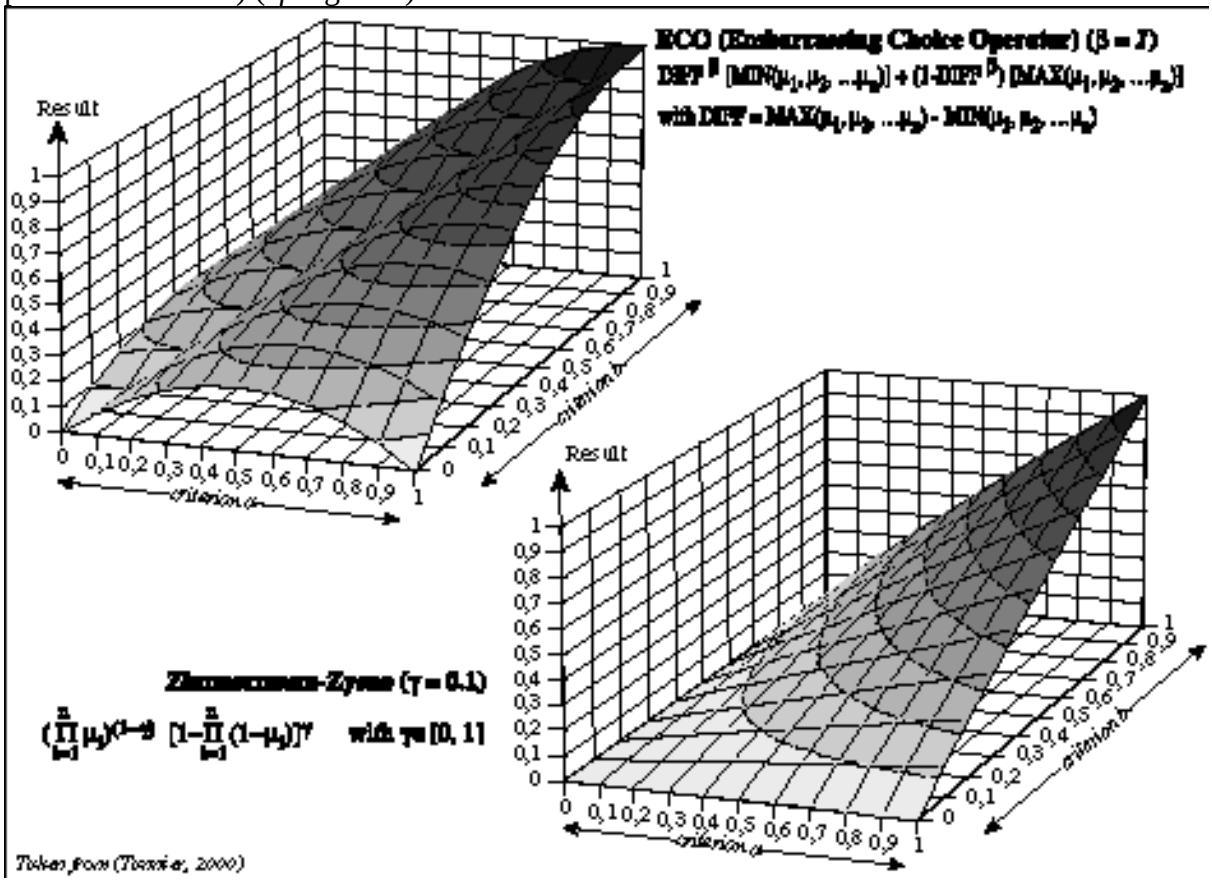

Figure 6 : Surfaces given by the operators ECO (with β=I) and Zimmermann-Zysno (with γ=0,I)

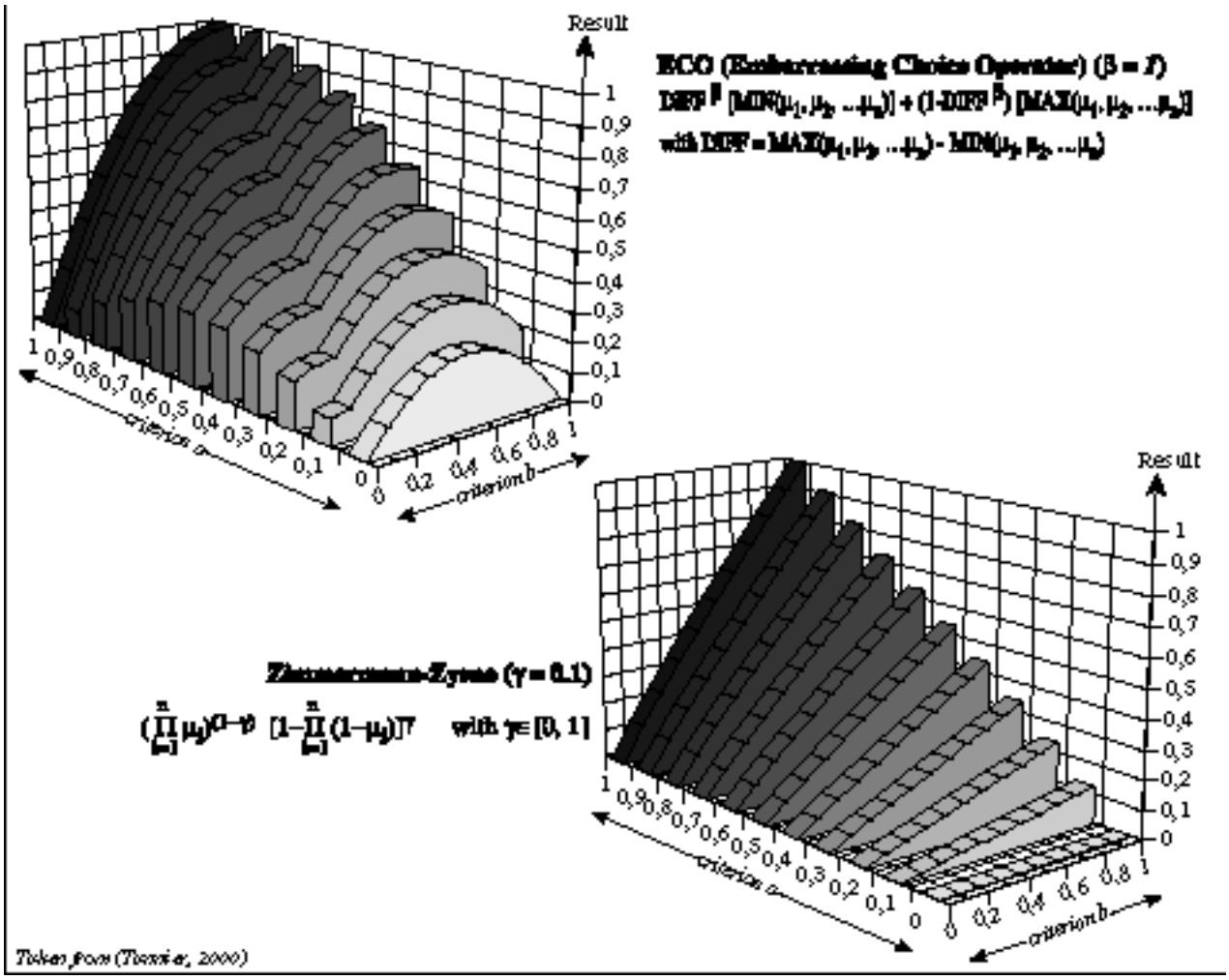

Figure 7 : Histogramms given by the operators ECO (with β=I) and Zimmermann-Zysno (with γ=0,I)

Through the graphical representations of the results of the aggregation, we can also compare the results given by some operators with the real attractiveness measures given by a set of individuals[8]. On this base, we can then choose the most convenient operator. The only constraint for doing that is to have a statistically representative sample of individuals.

However, one main problem when working with graphics is, that we deal only with two criteria (*i.e.* two partial attractiveness values). But some properties of the operators may appear with more than two criteria (*e.g.* the compensation level of the Zimmermann-Zysno operator grows with the number of criteria aggregated). So, it appears necessary to base the choice of the operators both on the comparison schedule and on the observation of the graphical results.

Before closing the description of the formalization of the evaluation phase, we would like to display quickly the main elements which determine how the agents evaluate the attractiveness of a neighbourhood, aside from the fact that an agent belongs to a type.

- Ex ante and ex post evaluation are different. The way an agent evaluates his/her own neighbourhood differs from the way he/she evaluates the other ones.
- The perception of the agents belonging to the same type could differ according to their previous experiences. For example, when an agent was living in a neighbourhood characterized by some especially big parking difficulties or a high level of noise annoyance, his/her evaluation of the neighbourhoods at the time of a new residential choice would differ from the evaluation of an agent who did not go through the same difficulties. Thus, for a same type of agents, secondary location criteria for the ones can

be determining location criteria for the others. To deal with this phenomenon two solutions can be thought of : the definition of more types of agents that would be more detailed or the weakening of the Markov assumption.

Conclusion : why do we speak of an *inductive* approach ?

The answer to this question is very simple : we tried to build a modelling concept based on knowledge taken from the observation of the reality and including the less presuppositions (apriorisms) as possible. Our aim was to remain very near to the data. We now just want to precise how this intention expresses itself through our modelling concept.

Concerning the formalization of the evaluation phase, neither the identification of types of agents nor the formalization of the perception of the attributes of the neighbourhoods by the agents include any presuppositions about the agents' behaviour. The only assumption is that types of agents can be identified and characterized by a common perception of the neighbourhoods of a town. We would like to remark here that in that sense our approach differs fundamentally from the econometric ones, where utilities functions are usually deduced with respect to some theoretical basic principles. Following the same logic, no mathematical constraints have been defined related to the aggregation operators. The choice of them are not based on a strict axiomatics and only two conditions are imposed : the operators must be characterized by the commutativity property and give results contained between zero and one (*cf.* figure 4). A particularly interesting consequence of this conception is, that the weakness of the axiomatics led us to deal with the problem of the practical meaning of some mathematical properties of the operators, question that has been not much discussed until now.

If we now look at the formalization of the choice phase, the main feature (and what distinguishes it especially from the Weidlich-Haag model) is, that the shape of the probability functions (including the two terms mobility and attractiveness) is not predefined : the probability functions are calibrated on the base of migration data. We made no assumption about the form of the function linking the mobility and the attractiveness.

The main interest of such an inductive approach is that the better we know the agents' behaviour the more we can easily improve the model. So, this approach favours the exchanges between the geographers keeper of a thematic knowledge and those who deal with more theoretical problems.

To close our purpose, we only would like to precise that our modelling concept will be applied on a concrete case in a close future. So we will next be able to expose and discuss the simulation results.

Courgeau D., 1984, "Relations entre cycles de vie et migrations", in Population, vol.3, INED, pp. 483-514.

Courgeau D., 1988, Méthodes de mesure de la mobilité spatiale, INED, PUF, 301 p.

Dubois D., Prade H., 1980, Fuzzy sets and systems - Theory and application, New-York, Academic Press, coll. Mathematics in science and engineering, 393 p.

Frankhauser P. 1996, Echelles, Hiérarchies et Fait Urbain, Habilitation à Diriger les Recherches, Université de Franche-Comté, 207 p.

Frankhauser P., 1991, Beschreibung der Evolution urbaner Prozesse mit der Mastergleichung, PHD-thesis, University of Stuttgart, 217 p.

Frankhauser P, Moine A., Bruch H., Tannier C., Josselin D., 1998 : "Simulating settlement pattern dynamics by modeling subjective attractivity evaluation of agents", paper presented for the 37th meeting on Western Regional Science Association, Monterey, California, février 1998.

Frankhauser P, Tannier C., Gillon P., Josselin D., 1997, "Simulating urban pattern dynamics by using an alternative approach of modelization", in European Workshop on Theoretical and Quantitative Geography, Rostock, août 1997.
Haken H.,1978, Synergetics : an Introduction, Springer Verlag, 355 p.
Jayet H., 1993, "Les modèles d'auto-organisation à mi-chemin ?", in Temporalités Urbaines, B. Lepetit, D.Pumain, Anthropos-Economica, pp. 69-88.
Malczewski J., 1992, «Site selection problem and a quasi-satisficing decision rule», in Geographical Analysis, vol. 24, n°4, pp. 299-316.
Maurin M., 1984, "Codage optimal et sub-optimal sur un tableau invariant", in Mathématiques et sciences humaines, 22e année, n°85, pp. 19-55.
Maurin M., 1986, "L'association du mesurage additif conjoint et des intervalles successifs", in Mathématiques et sciences humaines, 24e année, n°96, pp. 5-29.
Pumain D., 1997, "Pour une théorie évolutive des villes", in L'espace géographique, vol.2, pp. 119-134.
Pumain D., Sanders L., Saint-Julien T., 1989, Villes et auto-organisation, Paris, Economica, 191 p.
Sanders L., 1999, "Modelling within a self-organizing or a microsimulation framework : opposite or complementary approaches ?", in Cybergeo, n°90, 24-03-1999.
http://www.cybergeo.presse.fr/suger/sanders/lena.htm
Sanders L., 1992, Système de villes et synergétique, Paris, Anthropos-Economica, 274 p.
Sigg T., Koch N., Weidlich W., 1996, "Urban Evolution in Interaction with Population Dynamics", in Spatial Analysis of Biodemographic Data, Bocquet-Appelt and Pumain eds, John Libbey Eurotext, pp. 267-280.
Tannier C., 2000, Les localisations commerciales de détail en milieu urbain – Mieux connaître, par la modélisation, pour mieux aménager, Thèse de doctorat en géographie, Dirigée par M. Le Berre, Université de Franche-Comté, 323 p.
Tannier C., Josselin D., Bruch H., Frankhauser P., 1998, "Simulating retail locations using GIS and an agents' decision model for urban planning", in International conference on modelling geographical and environmental systems with GIS, Hong-Kong, juin 1998, pp. 523-531.
Tong-Tong J.-R., 1995, La logique floue, Paris, Hermès, 160 p.
Weidlich W., Haag G. eds, 1988, Interregional Migration. Dynamic Theory and Comparative Analysis ; Springer Verlag, Berlin, 387 p.
Weidlich W., Haag G., 1981, Quantitative Sociology, Springer Verlag, 217 p.
Winder N., 2000, "Modelling within a thermodynamic framework : a footnote to Sanders (2000)", in Cybergeo, n°138, 05-09-2000.
http://www.cybergeo.presse.fr/modelis/winder.htm.
Yager R. R., 1978, "Fuzzy decision making including unequal objectives", in Fuzzy sets and systems, n°1, North-Holland Publishing Company, pp. 87-95.
Zimmermann H.-J., 1987, Fuzzy sets, Decision Making and Expert Systems, Boston [u.a.], Kluwer Academic Publishers, coll. International Series in Management Science/Operations Research, 335 p.
Zimmermann H.-J., Zysno P., 1983, "Decisions and evaluations by hierarchical aggregation of information", in Fuzzy sets and systems, n°10, North-Holland, Elsevier Science Publishers B.V., pp. 243-260.

1 In future, the expression "types of agents" mentioned in this article corresponds to the detailed types.

2 From now on, all the movements observed in a town (migration of individuals or creation or closing down of shops or industrial branches) are designated by the general term "migration".

3 From now on we simplify the term Wi(ZK,t+ t|ZJ,t) and write it Wi(ZK|ZJ).

4 Instead of the weight of an attribute, we spoke in previous papers of the importance of an attribute (Tannier et al., 1998), (Frankhauser et al., 1998). But, because the term "importance" may lead to misunderstanding, we chose to use this one "weight".
5 When dealing with the problem of a possible compensation between two or more criteria, we share the economists' concern related to the problem of substitution rates.
6 Such a work has already be undertaken in the case of retail activities (Tannier, 2000).
7 The operator ECO (Embarrassing Choice Operator) has been presented before in (Frankhauser et al., 1998).
8 Such data can be obtained through enquiries.

26 avril 2001