



HAL
open science

A new Model of City Growth and its Application to a middle sized French City

Athanasios Batakis, Nga Nguyen, Michel Zinsmeister

► **To cite this version:**

Athanasios Batakis, Nga Nguyen, Michel Zinsmeister. A new Model of City Growth and its Application to a middle sized French City. 2022. hal-03551303v1

HAL Id: hal-03551303

<https://cnrs.hal.science/hal-03551303v1>

Preprint submitted on 1 Feb 2022 (v1), last revised 1 Mar 2022 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A new Model of City Growth and its Application to a middle sized French City

Athanasios Batakis Nga Nguyen Michel Zinsmeister

February 1, 2022

Abstract

In the first part of this paper we propose a new theoretical model of city growth based on percolation. The second half of the paper is devoted to a concrete application of the model, namely to the city of Montargis. It appears that the embedded algorithm is quite efficient in terms of computational time and allows to exploit big data type resources such as individual land lots.

1 Introduction

In 2022, more than half of the world population lives in cities. This is why modeling urban growth has become an important challenge for social sciences. The aim of this work is to better understand and eventually predict urban development in order to supply tools for decision makers.

In a celebrated paper of Harris and Ullman (1945), the city is considered as a multi-centered system of clusters with different hierarchical levels. In fact, for all cities, the centers, suburbs and hinterlands which form this hierarchy share the same features, both in terms of functionality and topographic patterns. Their functional scheme is repeated across different spatial scales, in a self-similar pattern: districts of different sizes at different levels in the hierarchy have a similar structure. For example, a farm-house is the point of attraction for fields and land areas of an agricultural domain, several farms together make a homogeneous pattern in a village's territory. At a higher level, a town market attracts population and activities from surrounding villages, and a regional capital is a major center of attraction for several of

those elementary farming districts. Moreover, city growth not only occurs through the addition of units of development at the most basic scale, but also through increasing specialization of key centers. This analysis strongly suggests that urban growth is in some sense a stable and intelligible process, hence accessible to various quantifications.

Sizes of the cities depend on local interaction of populations, production of goods, commercial activities, and various other factors. Nevertheless, even the largest cities grow from the tiniest seeds which, repeated at different levels of hierarchy, interact in different scales in a self-similar way (Frankhauser, 1990), making pertinent the use of fractal geometry tools.

In the middle of the 1980's, Batty and Longley (1986) have been the first to apply fractal geometry to study and model city growth. A few years later, Frankhauser (1990) proposed a more general overview of this geometry in urban geography. Fractals are used to define urbanisation rules applied at different scales of the urban areas (Frankhauser, 2008; Frankhauser and Pumain, 2010; Caruso et al., 2011; Badariotti, 2005).

The urban structure not only undergoes fractal properties but also, like all natural objects, grows and evolves with time. Any model of city growth must involve a dynamical aspect.

In fact, cities evolve through the cumulative addition and deletion of basic units, cells or particles on a determined space. Such units may be buildings, population, transportation networks,... All of them co-exist, interact in a urban system, and embody its dynamical growth.

A first attempt to describe urban expansion through mathematical models of fractal growth goes back to Batty and Longley (1994). They adapted a well known model, called diffusion limited aggregation (DLA). This model, proposed by Witten and Sander in 1981, is a process whereby particles undergoing a random walk cluster together to form aggregates. It is a simple model that was able to generate fractal structure whose self-similarity was dendritic or tree-like. In (Batty and Longley, 1994), the authors have calculated the fractal dimension of some urban structures and noticed that most of them are lying between 1.6 and 1.8, around the empirical value of dimension of DLA cluster, namely 1.71.... Even if this numerical evidence pleads in fa-

vor of their approach, this model has to be rejected since it does not fit with some other observed properties of urban systems. Most of urban structures are actually more compact near their center than DLA clusters. Moreover, DLA model does not take under account specific constraints of city's development, for example sea-sides, roads, rivers, Furthermore no-interaction between clusters occurs while a city is actually a complex system of related components.

In order to overcome some of the limitations of DLA model, Batty (1991) suggested to use the dielectric breakdown model (DBM), a modified version of DLA, in an attempt to model the urban growth of Cardiff. The DBM model, developed by Niemeyer, Pietronero, and Weismann in 1984, is a combination of DLA and previously existing electric field models. It allows to generate cities of different shapes and degrees of compactness. In addition, the physical constraints of rivers and sea were considered in the simulations.

Despite the fact that simulations of DBM are more realistic, they remain limited. A cluster generated by this model does not satisfy the complex property of an urban system: its main drawback concerns its density distribution. The density of a DBM-cluster in simulations decreases from the center as a power law

$$\rho(r) \sim r^{D-2}$$

where r is the radial distance from center of cluster, and $1 < D < 2$ is the fractal dimension of the simulated cluster. But, and this is now considered to be an universal law for cities, actual urban data have been more commonly found to fit an exponential decay (Clark, 1951)

$$\rho(r) \sim e^{-\lambda r}$$

where λ is the so called *density gradient*.

Later, Makse et al. (1998) proposed a toy model based on correlated gradient percolation to simulate Berlin with remarkable results. The model is time evolving and generates more complex urban areas which turn to be more realistic than the preceding ones . This model, compared with "real" Berlin, gives of course another "random" Berlin, but the impression of an air of family is significant, suggesting that some cities could grow according to some stochastic law.

However, this toy model faces two drawbacks. First it doesn't take in account all possible kinds of city growth. Secondly it is too rough to describe a concrete situation. For instance, in our example, the city of Montargis clearly does not grow as Berlin. Moreover, since our task was to provide a precise description of the dynamics of this city, we had to enrich the model with relevant specific parameters.

We have kept the idea of correlated gradient percolation, modified in order to overcome these weaknesses.

In Makse et al. (1998) the authors suggest that distance to the center governs city growth (in an exponentially decreasing manner $p(z) = e^{-\lambda|z|}$) and dynamics of the model is guided by the parameter λ (representing $1/\text{time}$).

In our first modified model, the role played by the distance to the center in Makse et al. (1998) is overtaken by local density of buildings at current time. In a second step, we push further this idea, replacing local density by a quantity leading to the local density in expectation at next step. This model may be run forward or backward, making possible to check the validity of its predictions.

The second model we propose is enriched by other factors that affect city growth, the most important among them being the concept of accessibility. We claim that this new parameter is better suited to towns with high dispersion which is the case of Montargis.

In both our models, the city growth is a random process governed mainly by density in a sense that may be summarized by: "density calls for more density". The simulations are carried out on land lots; we have compared simulations with real city data and obtained fairly good results.

The paper is organized as follows:

In the first section, we present some mathematical background about (un)correlated gradient percolation.

The second part is devoted to the case of a particular middle sized French city for which, after a quick geographical description, we describe two new models together with the results of actual simulations.

2 Percolation Models

Percolation is a random model that was introduced by John Hammersley in 1957 in order to describe the flow of a fluid or a gas through a porous medium with small channels which may or may not let fluid pass. During the last decades, percolation theory has brought new understanding and techniques to a broad range of topics in physics, materials science, complex networks, geography,....

Percolation theory describes the behavior of randomly formed clusters in a grid. A percolation model may be defined on any lattice in any dimension, or more generally on any graph.

Let $G = (V, E)$ be a graph, where V stands for the set of vertices and E the set of edges. There are two different kinds of percolation on such an object, bond percolation or site percolation.

In bond percolation, edges are declared open with probability p and closed with probability $1 - p$, independently of each other, and one is interested in clusters, the maximal sets of vertices such that two of them maybe connected by open edges.

In site percolation the sites are declared open or closed and the clusters are maximal sets of open vertices that are two by two graph-connected.

2.1 Site Percolation

Let G be an infinite connected graph for which we consider site percolation. There exists a well defined *critical value* p_c such that:

- if $p < p_c$ there is no infinite cluster a.s., this is called *sub-critical* regime;
- if $p > p_c$ there is a unique infinite cluster *super-critical* regime a.s.

Kesten (1982) proved that in 2-dimensions, on triangular lattice, $p_c = 1/2$. At the *critical* regime $p = p_c$, for the lattice \mathbb{Z}^d , there is no infinite cluster a.s. if $d = 2$ or $d \geq 19$ (the question remains open for other dimensions).

The value of p being given, this model provides a cluster of occupied (black) sites: in order to make it dynamical one lets p increase to 1. As p crosses the critical regime $p = p_c$, the cluster becomes connected and starts to look like a "city" if we think of occupied sites as buildings. But this model is unrealistic: these "cities" have no center and no border!

To make up for this drawback we use an alternative notion of percolation, namely gradient percolation.

2.2 Gradient Percolation

Gradient Percolation is a model of inhomogeneous percolation introduced by Sapoval, Rosso and Gouyet (1985). In this model, the occupation probability p attached at each site z is a function of z decreasing from 1 to 0 as the distance from z to the origin increases from 0 to ∞ . The existence, size and position of a cluster of occupied sites in the region depend on the set of z such that $p(z)$ is greater than the critical value p_c . The clusters have a center which is the origin and a border consisting of points where p becomes less than the critical value p_c .

As already mentioned, according to data of numerous cities, Clark (1951) claims that the population density $\rho(z)$ of the real urban systems should satisfy the relation

$$\rho(z) = \rho_0 e^{-\lambda r_z}, \quad (1)$$

with $r_z = d(z, 0)$ is the distance from z to the center ($= 0$) of the lattice-city, where λ is the so called *density gradient*.

Therefore, we assume that the occupation probability is an exponential function of r_z :

$$p(z) = \frac{\rho(z)}{\rho_0} = e^{-\lambda r_z}. \quad (2)$$

The density gradient λ influences the size of the central cluster. With small λ ($\lambda \approx 0$), the central cluster is large (corresponding to developed cities), and vice versa, the larger the value of λ ($\lambda \approx 1$) is, the smaller the central cluster is. As before, we get in this way a growth process where the role of time is played by $1/\lambda$.

This modification of the percolation process takes into account Clark's law but does not rule out the fact that "cities" obtained in this way have a rotational symmetry for example, which is rarely the case of concrete cities. Neither does this model take care of Christaller law saying that growing cities generate "satellite" smaller centers, a well-observed fact. In order to take into account these last points we need another modification explained in the next paragraph.

2.3 Correlated Gradient Percolation

Observation of real cities suggests that new constructions are not positioned independently of the existing ones. The positions/parcels seem to be occupied with probability depending on the presence of neighbours. In other

words, there exists a correlation between occupied land lots: near an occupied site, the probability of additional development is higher and decreases as we move away from it. We argue that these assumptions of site occupation affect growth of urban areas and thus introduce the following *correlated gradient percolation* model.

In order to build a bi-infinite matrix of correlated random numbers depending on the distance between two sites, we start with (X_z) , $z \in \mathbb{Z}^2$ an infinite matrix of independent variables with identical distribution $\mathcal{N}(0, 1)$, the standard normalized Gaussian distribution. From this Gaussian matrix we wish to construct another one, say (Y_z) with covariant matrix given by

$$\text{cov} \langle Y_z, Y_{z'} \rangle = C(|z - z'|), \quad z, z' \in \mathbb{Z}^2,$$

where

$$C(\ell) = \frac{1}{(1 + \ell^2)^{\alpha/2}}.$$

The parameter α has to do with the concentration, in the sense that the smaller the value of α is, the higher will be the concentration of clusters.

We look for such a matrix in the form

$$Y_z = \sum_{z' \in \mathbb{Z}^2} a(z' - z) X_{z'},$$

for which we may compute, assuming a is even,

$$a * a(z) = C(|z|),$$

where $*$ stands for convolution of functions defined on the group \mathbb{Z}^2 . This equation may be solved by taking Fourier transform of both sides,

$$\hat{a}(t, t')^2 = \hat{S}(t, t'),$$

where $S(z) = C(|z|)$, an equation which has a solution since $\hat{C} > 0$ (actually \hat{C} has been exactly computed in (Makse et al., 1996))

To finish the construction, we denote first by F the cumulative distribution function of a random variable following a $\mathcal{N}(0, 1)$ -law: the numbers $\eta(z) = F(Y_z)$ are then random numbers uniformly distributed in $[0, 1]$, and they form the desired matrix.

2.3.1 Summary of the Correlated Gradient Percolation (CGP) model

We are now able to embed these modifications into our model, namely, the use of gradient percolation instead of the standard one and the use of correlated site probabilities instead of the uncorrelated one.

Consider the following occupation probability function of site z on a square lattice of size $M \times N$.

$$p(z) = e^{-\lambda r_z} \quad (3)$$

where λ is the density gradient and $r_z = |z|$ is the distance from z to center 0 of the lattice.

We generate on the lattice a matrix $(\eta(z))$ from a correlated Gaussian matrix as in the preceding paragraph.

On the square lattice L of size $M \times N$, a site z is black and marked 1 (occupied) or white with marked 0 (vacant) with probability given by (3) in the following way:

$$L(z) = \begin{cases} 1 & \text{if } \eta(z) < p(z) \\ 0 & \text{otherwise,} \end{cases} \quad (4)$$

The algorithm of CGP model consists of the following steps:

1. Generate a correlated sequence $(Y(z))$ from an uncorrelated random sequence $(X(z))$ with normalized Gaussian distribution by using the preceding method.
2. Find the corresponding numbers $\eta(z)$.
3. Calculate the occupation probability function in (3) and simulate $L(z)$ by (4).

This model has been developed by (Makse et al., 1996) and successfully used to model the city growth of Berlin. This success is partly due to features that are particular to Berlin, which is a city with a big center with high attractivity (so that the distance to the center plays a prominent role) and rather rotationally symmetric, with no big river crossing it, for instance.

Thus, in order to deal with other kind of cities, we need to further modify the model.

3 First Modified CGP Model

Let us first introduce some notations and concepts for this new model. A town will be a finite subset of $\mathcal{S} := \{1, \dots, M\} \times \{1, \dots, N\} \subset \mathbb{Z}^2$. The set $\mathcal{M} \subset \mathcal{S}$ consists of the set of black points (corresponding to the buildings). Equivalently it is a binary $M \times N$ matrix ($\mathcal{M}(z)$). The growth model that we are going to consider consists in iterating an algorithm that we now describe. We call \mathcal{M}_t the town at time t and wish to construct $\mathcal{M}_{t+1} \supset \mathcal{M}_t$. This procedure is based on gradient percolation. More precisely we consider a function $p_t : \mathcal{S} \setminus \mathcal{M}_t \rightarrow [0, 1]$ and \mathcal{G} a correlated Gaussian $M \times N$ matrix whose entries follow a $\mathcal{N}(0, 1)$ law.

Let $\eta = \mathcal{G}(\omega)$ be a realization of the random matrix. For all $z \in \mathcal{S} \setminus \mathcal{M}_t$, we decide to include z in \mathcal{M}_{t+1} if and only if $\eta(z) < p_t(z)$ as in the above described model of correlated gradient percolation. We then consider a function $p_{t+1} : \mathcal{S} \setminus \mathcal{M}_{t+1} \rightarrow [0, 1]$ such that $p_{t+1} > p_t$ and we start over the process. The novelty in this model is that we allow p_{t+1} to depend on \mathcal{M}_t . This basic structure being described, the game is of course to choose the function p_t , as well as the parameter α .

Concerning the function p_t we already mentioned the choice made by Makse et al. (1998): $p_t = e^{-\frac{r(t)}{t}}$ where t is now a continuous parameter going from 0 to ∞ and $r(t)$ is the distance of z to the center of the city. This choice induces very "compact" towns for which the most important criterium of choice is the vicinity of the center.

Because of our main subject of study, namely the city of Montargis, we prefer to privilege more decentralized towns by taking p_t to be the concentration or density function. This is defined as

$$p_t(z) = 1_{\mathcal{M}_t} * \varphi(z) \tag{5}$$

where φ is a nonnegative function with

$$\sum_{z \in \mathcal{S}} \varphi(z) = 1,$$

and the convolution $(*)$ is the (already defined) one on \mathbb{Z}^2 after having extended the two functions by 0 outside \mathcal{S} . In order to fit with Clark's law we only consider functions φ of the form $\frac{1}{c} e^{-c|z|}$.

3.1 Yet another modified CGP-model

A slightly different model is based on the belief in harmonious growth of cities. It makes sense for a city which is already well developed and for which we want to predict the future growth, assuming its harmonious character. So let \mathcal{M}_0 be the present stage of the city and, as before, $p_0 = 1_{\mathcal{M}_0} * \varphi$, that is the density function at $t = 0$.

The idea is not to take p_0 as the percolation threshold but instead to take a function ρ to be defined in such a way that, denoting by \mathcal{M}_1 the new city built using CGP method with this threshold, we have

$$\rho = \mathbb{E}(\mathcal{M}_1) * \varphi$$

outside \mathcal{M}_0 . In other words, we want ρ to be equal to the density function of the city it builds, in expectation.

In order to compute this function we write \mathcal{M}_1 as the disjoint union of \mathcal{M}_0 and \mathcal{H}_1 . We first notice that

$$\mathbb{E}(1_{\mathcal{H}_1}) = (1 - 1_{\mathcal{M}_0})\rho,$$

so that the equation for ρ is

$$\rho = p_0 + \mathcal{T}(\rho),$$

where

$$\mathcal{T}(\rho) = [(1 - 1_{\mathcal{M}_0})\rho] * \varphi.$$

Unfortunately (but this is not much of a surprise), the only solution of this equation is the constant 1. We pass this difficulty by posing

$$\rho_t = (\mathcal{I} - t\mathcal{T})^{-1}(p_0)$$

which defines a growing one-parameter family of thresholds, and thus a growing city in time grows from 0 to 1. This model will be used later for the growth of Montargis: we call it the *forward model*.

This model may also be run backwards. In order to recover a town in the past we change the above model by using the dual $\mathcal{M}' = \mathcal{S} \setminus \mathcal{M}$, at each site z , if $1_{\mathcal{M}}(z) = 1$ then $1_{\mathcal{M}'}(z) = 0$ and vice-versa. The growth model is generated by using this matrix similarly to the forward model. But in this case the white area is growing, so the black one is decreasing, as it should be.

4 Application to a concrete case

4.1 A Brief Geographical Presentation of Montargis

The city of Montargis is the second (after Orléans) most populated town in the département du Loiret, France. The whole basin of population reaches 70000 inhabitants, which makes it, in french standards, a middle size town. It is situated 130 km south east of Paris, a proximity that, as we shall see later, plays a role in the economy of the town. Montargis is also 70 km north-east from Orléans, the local economical center. The town is crossed by the river Loing, which is an affluent of the Loire river.

In the sixties, due to the proximity of Paris, many people from the Paris area (Ile de France), had built summer houses in the Montargis area. These summer houses became principal ones as their owners retired. This fact participates to the fact that the population is rather dispersed in the vicinity of Montargis, and also rather old. The population of the center is declining, but the population of the basin is slightly increasing, a phenomenon which is shared by most middle-sized french towns, and largely due to the generalization of shopping malls that kill small shops in the centers.

The main local economical activity is low-tech industry (rubber and paper industry), a fact that may explain that as far as median income is concerned, Montargis is one of the poorest city of this size in France. In order to illustrate the long-run industrial tradition of Montargis, let us mention that in the 20's, in the framework of a program called *Work and Study*, several future leaders of the Chinese Communist Party have stayed and worked there. The most famous one is Deng Xioaping who has spent two years (1923-24) as a worker in the Hutchinson factory which still exists. For the record this fact has generated a Chinese touristic movement with "pilgrims" coming every year to visit the places (that have been carefully preserved) where Deng lived and worked.

The proximity of Paris has a second effect: more and more people, while living in Montargis, work in the Paris area, and an even more recent phenomenon (after the pandemy) is remote, th worke proximity of Paris making easy a one or two days travel to the employer in Paris.

This little presentation explains in particular why the growth of Montargis is very different from the Berlin one, and our simulations had to take this difference into account.

4.2 Simulations using the forward model

These simulations are based on real data of Montargis that we obtained from the project Transport-Réseau-Urbanisation-Croissance (TRUC) and are realized with MATLAB program. In what follows the correlation factor $\alpha = 0.001$.

In addition to the buildings, we consider the influence of the rivers (figure 1a) on the growth of the city. We assume that the effect of the rivers on the occupied probability is the same as for the buildings. The rivers are thus considered as black points like the buildings.

Firstly, the forward model is applied on Montargis 2007 (figure 1c) as the starting point ($t = 0$). The black points in the starting matrix contain the constructed buildings at 2007 and the rivers system of Montargis. The simulations are done with a time sequence of scale $\delta t = 0.01$, we get for $t = 0.01, 0.02, 0.03$ the figure 2.

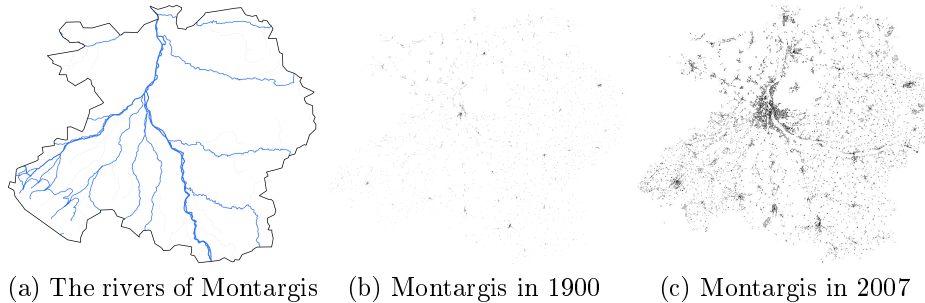


Figure 1: The river and the constructed building of Montargis at 1900 and 2007, the data are taken from the project TRUC.

Secondly, we have considered Montargis 2007 as the starting point and used the backward model to simulate in the past at $t = -0.01, -0.02$ (figure 3). We see that the number of buildings (black points) decreases very quickly at the first simulation.

Finally, we have used the simulation of backward model at $t = -0.02$ (figure 3b) as the starting point to find how it is in future by using the forward model with $t' = 0.01, 0.02$ (figure 4).

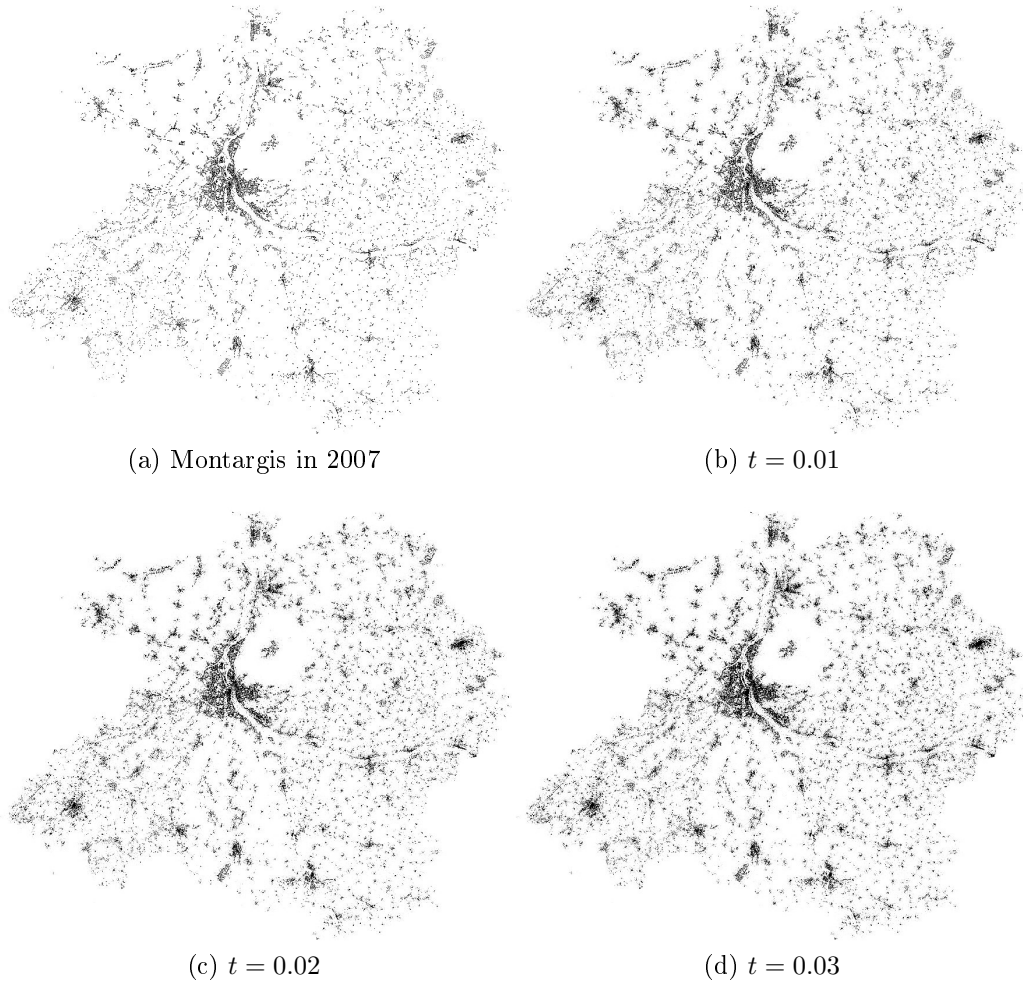


Figure 2: The simulations in future of Montargis using forward model in different “times” t .

5 A multi-parameter model

In the preceding model, we assumed that the distribution of buildings models the growth in the future or the reduction in the past on Montargis. In reality, there are many factors interacting in the urban system such as the minimum area around a building, the accessibility, the commercial areas, the schools, the forests, the risk areas We will try to include these factors in a modified

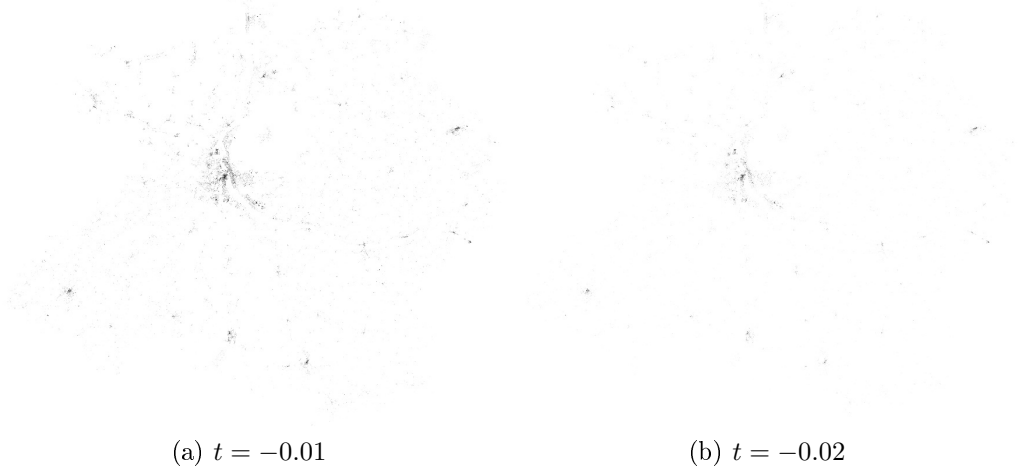


Figure 3: The simulations for the past of Montargis from 2007 using the backward model in different time t .

version of our first model in order to better simulate the growth of Montargis.

5.1 Model

A land lot (or a plot) is a land in the register of ownership (“cadaster”) with a determined area and boundary, owned by an individual or organization (see figure 5a). Hence, land lots form a partition of the landscape. Their size and forms are variable, from a few square meters to more than $10^4 m^2$.

A plot that has buildings in it will be called occupied and we will use the term empty (or vacant) if it has nothing inside.

The plots are of different size and forms (we are no longer in the square lattice paradigm). The big plots which are usually far from the center of city can be divided into the smaller ones. However, due to the cost of division, a big plot is often divided all at once to sell or construct. In fact, we can admit that a big plot will be divided when it has a high construction probability (real estate value). We will assume that an occupied plot means that it is not possible to build any more or that all its divisions were constructed. This also meets previous hypothesis that an occupied land lot will maintain its status in future (it does not become vacant or empty). These plots are the objects of our study. They replace sites (points) in the previous model.

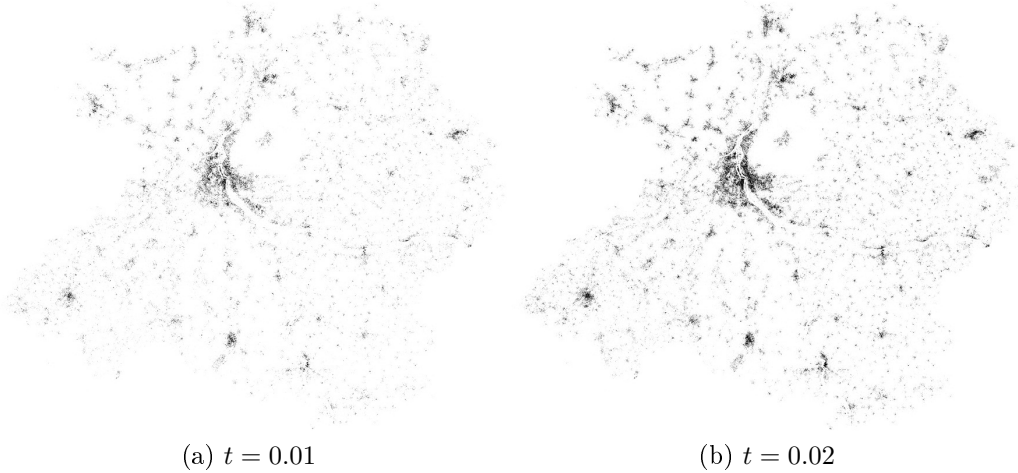


Figure 4: The simulations for the future of the simulation 3b as the starting point using the forward model with different t .

This model does not consider size, forms or type of buildings in land lots : a land lot is occupied or vacant. Since the type of the buildings (commercial, administration, schools, hospitals ...) are not differentiated in the model, their effect in city growth is assumed to be the same. It is not difficult to take under consideration different effects but this needs a precise knowledge of the use of every constructed space.

To summarize, we adapt CGP model on land lots with the following assumptions:

- The change of status of a land lot (vacant to occupied) occurs with a probability related not only to distance to the center of the city but also to local density of the urban system and accessibility,
- An occupied land lot will stay occupied, there is not changing from occupied to vacant,
- We do not distinguish the different types of buildings in land lots either from size or from purpose of use.

Let us come now to the details of the model. As before, we define the binary matrix \mathcal{M} of size $M \times N$ containing center points of land lots of

Montargis, $z \in [1, M] \times [1, N]$ (size of Montargis) as follows:

$$\mathcal{M}(z) = \begin{cases} 1 & \text{if } z \text{ is a occupied land lot,} \\ 0 & \text{if } z \text{ is empty.} \end{cases}$$

Let $\phi(t) = e^{-|t|}$, $c = \sum_{z \in \mathcal{S}} \phi(|z|)$ and $\varphi = \frac{1}{c}\phi$. We define the local density as

$$\Delta(z) = \mathcal{M} \star \varphi(z) = \frac{1}{c} \sum_{z_1 \in \mathcal{M}} e^{-|z-z_1|}. \quad (6)$$

We also assume that the probability of construction of a land lot is related to its *accessibility*, which we now define: first we define a set \mathcal{N} of “nodes”, i.e. points on the road network (mainly cross-roads). Following Gane et al. (2011), the accessibility $A(z)$ of a given node z is a (possibly ponderated) sum over all the nodes z' of the road network of the time $t(z, z')$ necessary to go to z' from z (figure 5b):

$$A(z) = \sum_{z' \in \mathcal{N}} t(z, z') \quad (7)$$

As for density, locations with high accessibility (small A) will be considered as more attractive.

Finally we make the assumption that being close to a river and/or to the center of the city does attract settlements.

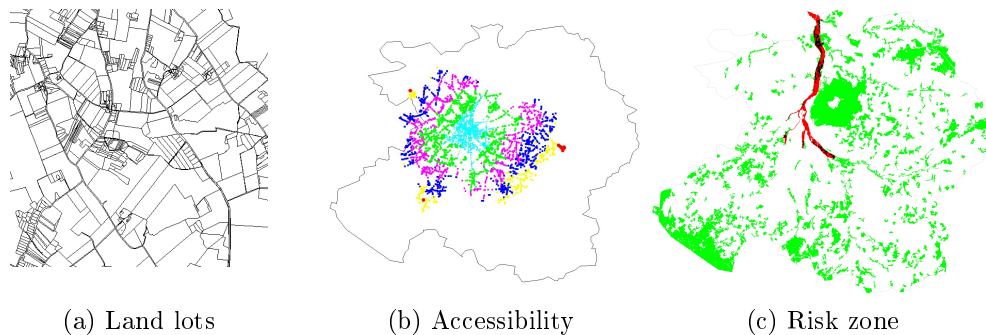


Figure 5

All these considerations lead us to modify the probability p_t defined in (5) as follows:

$$p_t(z) = \frac{1}{C} e^{-\lambda_{t,1}|z-z_0| - \lambda_{t,3}d(z,R)} \Delta(z) e^{-\lambda_{t,2}A(z)} \quad (8)$$

where

- C is a normalizing constant,
- $\lambda_{t,1}, \lambda_{t,2}, \lambda_{t,3}$ are parameters possibly depending on time t that will be fixed empirically using the historical data of Montargis,
- $z_0 = 0$ is the center of the city,
- $d(z, R)$ is the shortest distance from z to the river,

To summarize, the CGP model is used on *vacant* land lots which become *occupied* with probability given by (8). The correlation exponent in the model has been chosen to be small ($\alpha = 0.001$) implying that correlation is important. The starting point \mathcal{M}_0 is Montargis in 1900. The simulation builds \mathcal{M}_{i+1} for \mathcal{M}_i only: no external information is added.

5.2 Simulations

The model cannot predict the rhythm of the city growth, which is clearly non-linear: for instance the war period is very different from the sixties' which is itself very different from the beginning of the XXth century. Thus we change time in such a way that the total number of buildings coincides between reality and simulations at each time (figure 6).

Starting from Montargis 1900 (figure 1b) we obtain the simulations for the growth of Montargis from 1900 to 2007. In the figures 7 we present the simulations at 3 equidistributed times. All the parameters present in the formula ((8)) for p_t are empirically adjusted in order to best fit with reality. This choice is then applied once for all.

5.3 Comparisons

A comparison is performed by overlapping real Montargis 2008 with its simulation (figure 8a). There are some regions near the north boundary where the simulations do not match as well as near the center. In fact, the speed

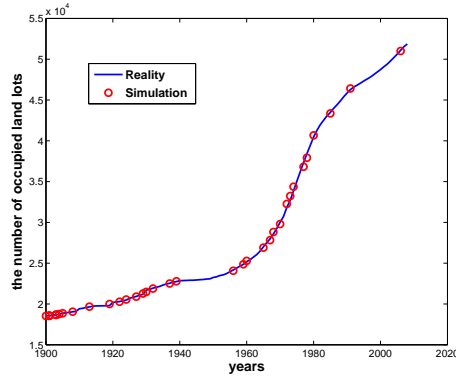


Figure 6: The evolution of the number of constructed plots of real Montargis and its simulations from 1900 to 2008.

of growth is higher there due to proximity of Paris and the emerging of peri-urban population.

Also, from place to place, there have been new subdivisions of blocks of buildings that of course could not be predicted. Besides these two exceptions, simulations fit rather well with reality.

5.4 Quantification of the error term in the simulations

In addition to the above overlap, we give some numerical comparisons between Montargis 2007 and its simulation.

Let $f_r(z)$, $f_s(z)$ be the density functions of a plot z in real Montargis \mathcal{M}_r and its simulation \mathcal{M}_s respectively; i.e.

$$\tilde{f}_j(z) = \sum_{z' \in \mathcal{M}_j} e^{-|z-z'|}, \quad j = r, s$$

Let $C = \max_z \{f_r(z); f_s(z)\}$ and take $f_j(z) = \frac{1}{C} \tilde{f}_j(z)$. We define the relative error at z as $|f_r(z) - f_s(z)|$.

The figure 8b shows the distribution of these relative errors of density.

The maximal relative error of density is

$$\epsilon_{max} = \max_z |f_s(z) - f_r(z)| = 0.25,$$

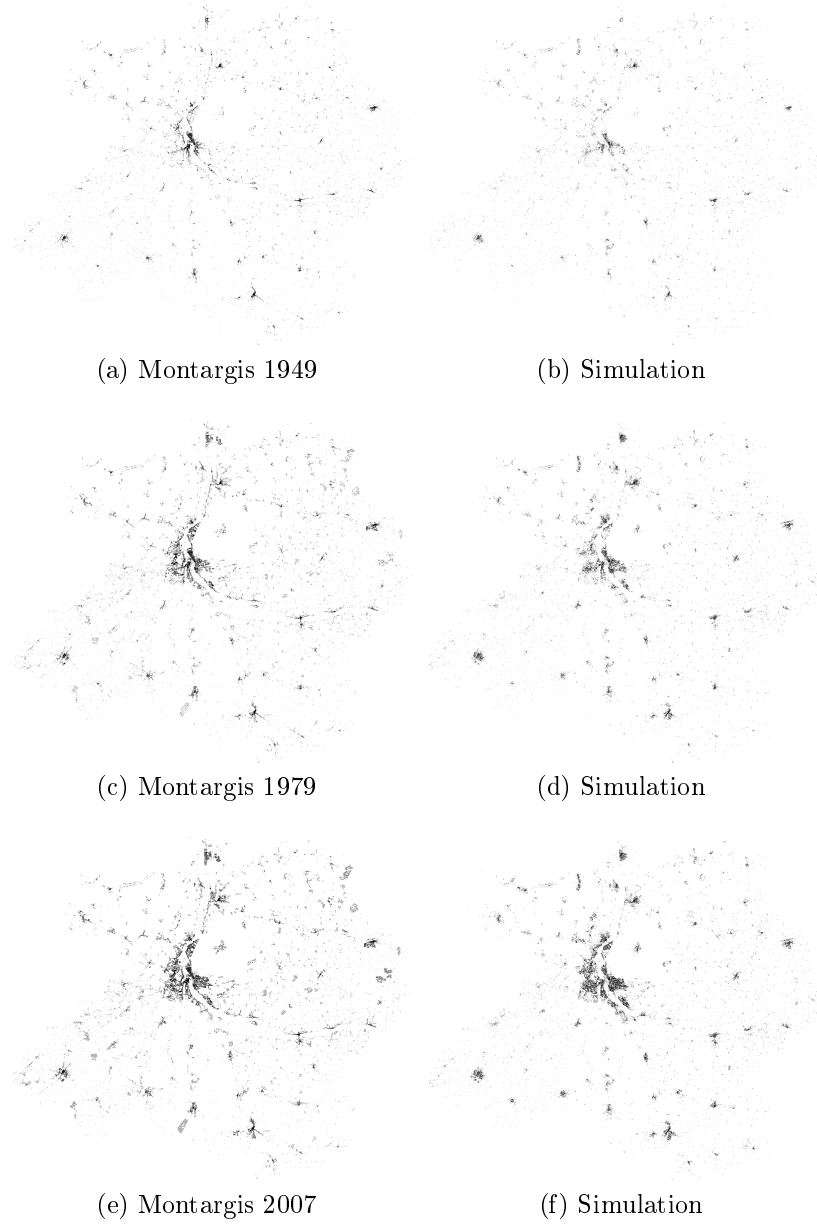
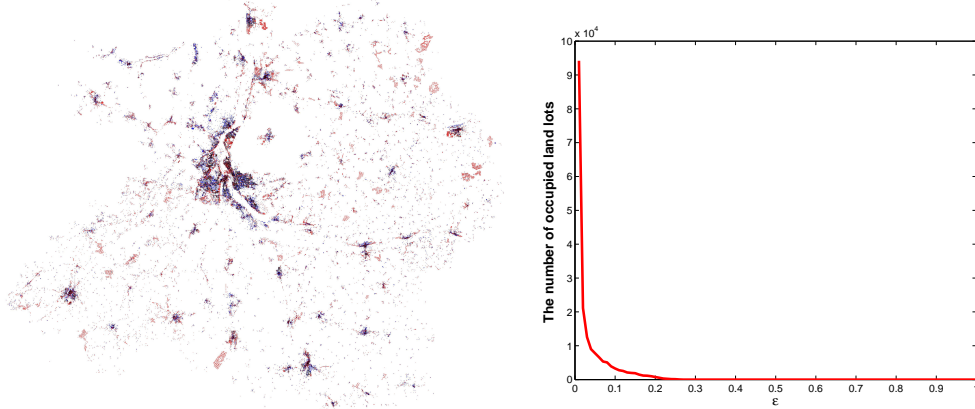


Figure 7: Montargis in real for the years 1949, 1979, 2007 (on the left) and the corresponding simulations (on the right) realized by using Correlated Gradient Percolation with occupied probability function (8).



(a) The overlap between real Montargis (red) and its 2008 simulation (blue) (b) The distribution of errors

Figure 8: Quantitative comparisons

the mean error is

$$\epsilon_{mean} = \frac{1}{n} \sum_z |f_s(z) - f_r(z)| = \mathbf{0.031}$$

and the mean squared error is

$$\epsilon_{MSE} = \frac{1}{n} \sum_z |f_s(z) - f_r(z)|^2 = \mathbf{0.003}.$$

where n is the number of terms in the sum.

6 Conclusion

Using the huge quantity of informations that has been made available to us by the local authority, we have been able to construct a powerful and realistic model for the city growth of Montargis that could serve as a useful tool for decisions makers. This multi-parametrized model is also very flexible and can thus be applied to other cities as well. Compared with the inspiring model of Makse et al. (1998), it introduces new quantified notions such as density or accessibility which have proven to be pertinent in connection

with city growth, at least for middle sized French cities. It would be interesting to investigate to what extent it may be applied to the growth of modern megapoles which seem to obey different rules whose quantification is a challenge: we intend to explore, in the near future, the case of overgrowing African megapoles that look set to overturn the actual african civilization. Another continuation of this model would be to apply the model at a different scale, such as networked cities: a first attempt has been done for the French *Région Centre*, presented at the centennial conference of the American Geophysics Union (Batakis, Nguyen and Zinsmeister (2019)). Besides its possible further uses, this model also raises many theoretical questions. As we have seen, among the main quantities that rule the models are local density and the correlation coefficient α . These two quantities are pretty much related: In the simulations it appears that the correlation factor seems to favor the emergence of "seeds" of Christaller satellites while the density factor transforms these seeds in real "sub-towns". It would be interesting to understand further this phenomenon and the mutual influence of these factors.

In a work in progress we investigate how one could statistically estimate the correlation coefficient α . It is in particular unclear if this can be done with the sole knowledge of the city map at a given time or if we really need the whole history of the town.

References

- Badariotti, Dominique. 2005. "Des fractales pour l'urbanisme? : Quelques pistes de réflexion à partir de l'exemple de Strasbourg-Kehl." *Cahiers de géographie du Québec* 49(137):133–156.
- Batakis, Athanasios, Thi-Thuy-Nga Nguyen and Michel Zinsmeister. 2019. Gradient Percolation and Growth of Networked Cities. 100th anniversary of the American Geophysics Union San Fransisco, USA: AGU.
- Batty, M and P A Longley. 1986. "The Fractal Simulation of Urban Structure." *Environment and Planning A: Economy and Space* 18(9):1143–1179. **URL:** <https://doi.org/10.1068/a181143>
- Batty, Michael. 1991. "Generating Urban Forms from Diffusive Growth." *Environment and Planning A: Economy and Space* pp. 511–544.

- Batty, Michael and Paul Longley. 1994. *Fractal Cities: A Geometry of Form and Function*. San Diego, CA, USA: Academic Press Professional, Inc.
- Caruso, Geoffrey, Gilles Vuidel, Jean Cavailhès, Pierre Frankhauser, Dominique Peeters and Isabelle Thomas. 2011. “Morphological similarities between DBM and a microeconomic model of sprawl.” *Journal of Geographical Systems* 13(1):31–48.
URL: <http://dx.doi.org/10.1007/s10109-010-0131-7>
- Clark, Colin. 1951. “Urban Population Densities.” *Journal of the Royal Statistical Society. Series A (General)* 114(4):490–496.
URL: <http://www.jstor.org/stable/2981088>
- Frankhauser, Pierre. 1990. “Aspects fractals des structures urbaines.” *L’Espace géographique* 19/20(1):45–69.
URL: <http://www.jstor.org/stable/44380008>
- Frankhauser, Pierre. 2008. *Fractal Geometry for Measuring and Modelling Urban Patterns*. Physica-Verlag HD pp. 213–243.
URL: https://doi.org/10.1007/978-3-7908-1937-3_11
- Frankhauser, Pierre and Denise Pumain. 2010. *Fractals and Geography*. ISTE pp. 281–311.
URL: <http://dx.doi.org/10.1002/9780470612255.ch10>
- Gane, Vénéétia, Malika Serrière, Kamal Serrhini and Serge Thibault. 2011. *La localisation des activités en tant que composante de l’étalement urbain : accessibilité réseau sur l’aire urbaine de Montargis’*. Thesis, École Polytechnique Universitaire, Tour.
- Harris, Chauncy D. and Edward L. Ullman. 1945. “The Nature of Cities.” *The ANNALS of the American Academy of Political and Social Science* 242(1):7–17.
URL: <https://doi.org/10.1177/000271624524200103>
- Kesten, Harry. 1982. *Percolation theory for mathematicians*. Vol. 2 of *Progress in Probability and Statistics* Mass.: Birkhäuser Boston.
URL: <http://www.math.cornell.edu/kesten/kesten-book.html>
- Makse, Hernán A., José S. Andrade, Michael Batty, Shlomo Havlin and H. Eugene Stanley. 1998. “Modeling urban growth patterns with correlated

percolation.” *Phys. Rev. E* 58:7054–7062.

URL: <https://link.aps.org/doi/10.1103/PhysRevE.58.7054>

Makse, Hernán A., Shlomo Havlin, Moshe Schwartz and H. Eugene Stanley. 1996. “Method for generating long-range correlations for large systems.” *Phys. Rev. E* 53:5445–5449.

URL: <https://link.aps.org/doi/10.1103/PhysRevE.53.5445>

Sapoval, B., M. Rosso and J.F. Gouyet. 1985. “The fractal nature of a diffusion front and the relation to percolation.” *J. Physique Lett.* 46(4):149–156.

URL: <https://doi.org/10.1051/jphyslet:01985004604014900>