De la géométrie des réseaux viaires à l'organisation spatiale intra-urbaine

Thomas Louail

▶ **To cite this version:**

Thomas Louail. De la géométrie des réseaux viaires à l'organisation spatiale intra-urbaine : Une approche de la comparaison des formes de villes par la simulation orientée agents. Revue Internationale de Géomatique, 2009, Simulation spatiale à base d'agents (dir. Arnaud Banos & Lael Parrott), 19 (4), pp.443-469. 10.3166/geo.19.443-469 . hal-03820428

HAL Id: hal-03820428
https://hal.science/hal-03820428

Submitted on 19 Oct 2022

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.

De la géométrie des réseaux viaires à l'organisation spatiale intra-urbaine : une approche de la comparaison des formes de villes par la simulation orientée agents

Thomas Louail [*,**]

[*] *Laboratoire IBISC (Université d'Evry Val d'Essonne - CNRS)*
Tour Evry2 – 423 Place des Terrasses, F-91000 EVRY

[**] *Laboratoire Géographie-Cités (Univ. Paris 1 - Univ. Paris 7 - CNRS)*
13 rue du four, F-75006 PARIS

thomas.louail@ibisc.fr

RÉSUMÉ. *Les villes états-uniennes et européennes présentent de nombreuses différences dans leur organisation spatiale, qui s'expriment notamment dans les formes de répartition des densités et des prix à l'intérieur des villes des deux continents. Nous faisons l'hypothèse que ces* différences de répartition *s'expliquent en partie par les* différences géométriques *historiques existant entre ces villes, ancrées dans les réseaux viaires et dans les vitesses des déplacements dans la ville. Pour tester la validité de cette hypothèse nous avons développé* SIMPOPNANO*, un modèle orienté agents simulant la localisation de fonctions urbaines dans la ville. A travers sa présentation, nous abordons plusieurs problèmes méthodologiques souvent rencontrés dans les projets de simulation en géographie.*

ABSTRACT. *European and North-American urban systems are different in many ways, and at different scales of analysis. At the intra-urban scale, they oppose by spatial repartition of socio-economical activities, and spatial configurations of density and land rent prices. We argue that these spatial repartition differences between US and European cities can be partly explained by the geometry of their streets networks. We wish to study the validity of our hypothesis by the modeling of localization mechanisms of urban functions inside the city. This environment is used to address several methodological questions often raised in geosimulation projects.*

MOTS-CLÉS : *simulation orientée agents, ville, approche comparative, modèle de localisation, couplage multi-niveaux, calibrage automatique*

KEYWORDS: *agent-based simulation, city, comparative approach, land-use modeling, multi-level coupling, automatic tuning*

Géomatique – 19/2009. Simulation spatiale à base d'agents, pages 443 à 469

1. Introduction

1.1. *Une question géographique originale…*

Les systèmes urbains européen et états-unien présentent de nombreuses différences, à plusieurs échelles d'analyse, tant au niveau des villes (Bertaud, 2004), (Mangin, 2004) qu'au niveau de leurs systèmes de villes (Cattan *et al.*, 1994), (Bretagnolle *et al.*, 2007). Au niveau intra-urbain, si l'on se restreint à l'ensemble formé par les villes dont la population est comprise entre un et trois millions d'habitants[1], plusieurs indicateurs peuvent nous éclairer sur la nature de ces différences entre villes européennes et nord-américaines.

Tout d'abord, comme l'ont mis en évidence Bretagnolle et Delisle sur l'ensemble des villes considérées, la surface d'une ville états-unienne est en moyenne deux fois plus grande que celle d'une ville européenne (Delisle, 2008). Ensuite, les formes des répartitions à l'intérieur de la ville sont différentes. En Europe la densité de population chute plus vite lorsque l'on s'éloigne du centre des villes qu'aux Etats-Unis, ce qui revient à dire que le gradient centre-périphérie est plus marqué dans les villes européennes. Cela a notamment été illustré par Bertaud et Malpezzi qui ont montré le caractère systématique de cette observation en calculant ce gradient pour une cinquantaine de métropoles mondiales (Bertaud *et al.*, 2003) (*cf.* figure 1). Enfin, plusieurs études sur les rentes foncières ont montré que celles-ci se répartissent spatialement de façon très différente : si la répartition des prix en fonction de la distance au centre suit bien une exponentielle négative dans les deux cas, tel que le décrit le modèle classique d'Alonso (1964) pour une ville monocentrique, le gradient centre-périphérie est une fois encore plus marqué en Europe (Bertaud, 2004). Ces trois types de différences observées, surface, gradient de densité et gradient des prix, seront qualifiées dans la suite de *différences dans les répartitions spatiales*.

Une autre différence majeure et historique réside dans la *géométrie* des réseaux viaires des villes de chacun des deux continents. Nous appelons *géométrie* des réseaux viaires la combinaison de leur topologie et du schéma de vitesses de déplacements qui s'y superpose. En raison des temporalités très différentes sur lesquelles ces villes se sont construites, les villes européennes diffèrent fondamentalement des villes américaines, qui sont beaucoup plus jeunes.

D'une part, les topologies des réseaux viaires de ces deux grands types de villes sont différentes. Même si à l'intérieur de chacun des continents, certaines villes sont remarquables par leur singularité et leur écart au schéma général, il reste pertinent de parler de deux topologies caractéristiques de plans de villes. La ville européenne, qui s'est construite autour d'un centre historique par le peuplement progressif de sa périphérie, a un plan généralement radio-concentrique, ses enceintes successives étant devenues des boulevard circulaires au fur et à mesure de son extension. La

1. Les données pour ces villes sont issues de la base Géopolis (Moriconi-Ebrard, 1994), dans laquelle le sens donné au mot « ville » est celui d'agglomération urbaine : les limites de la ville sont définies en termes de continuité de bâti.

ville nord-américaine présente quant à elle un plan « en damier », où la maille orthogonale élémentaire semble se reproduire à l'infini. Ces différences dans le tracé des réseaux viaires peuvent se lire comme l'opposition entre, d'un côté, une forme ayant émergé d'un développement sur le temps long et, de l'autre, une forme planifiée *a priori* et imprégnée des idées urbanistiques contemporaines de sa conception (Ghorra-Gobin, 1998).

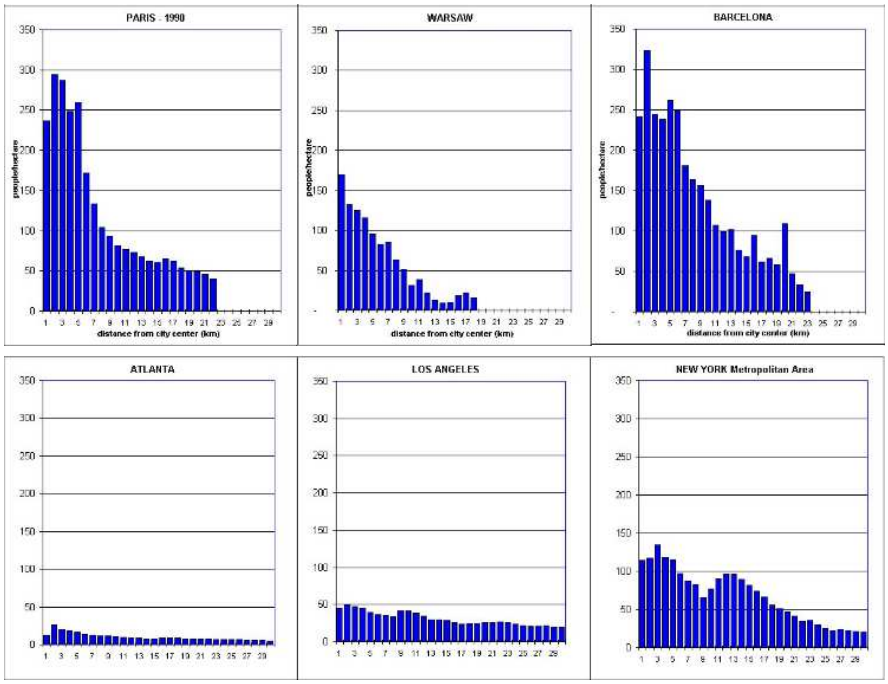


Figure 1. *Les courbes d'évolution de la densité de population en fonction de la distance au centre de trois villes européennes (Paris, Varsovie, Barcelone), et trois villes états-uniennes (Atlanta, Los Angeles, New York). Reproduites avec l'aimable autorisation de l'auteur (Bertaud* et al.*, 2003)*

D'autre part, la construction de nombreuses villes américaines (en particulier les villes de la côte Ouest les plus récentes) a été contemporaine de la construction des réseaux d'autoroutes et de l'essor de la voiture individuelle. Cela a pour conséquence que les vitesses moyennes de déplacement sont plus élevées dans les villes aux Etats-Unis que dans les villes en Europe, du fait d'infrastructures routières mieux adaptées à l'usage intensif de la voiture. On peut se faire une idée de ces différences de vitesse de circulation en comparant les résultats de deux enquêtes récentes de (Laterrasse *et al.*, 2009) et (Haas *et al.*, 2006). Cette comparaison peut se résumer ainsi : en moyenne, dans le même intervalle de temps, le navetteur états-unien peut parcourir une distance environ deux fois plus grande que son homologue européen. On peut finalement dire que ces *différences géométriques* se traduisent par des différences d'*accessibilité* entre

les quartiers de la ville dans les deux schémas, et que la géométrie nord-américaine est moins inégalitaire : elle induit moins de différences d'accessibilité entre les quartiers que ne le fait la géométrie européenne.

Dans ce contexte, la question que nous nous posons est la suivante : les *différences géométriques* existant entre ces villes suffisent-elles à expliquer les *différences dans les répartitions spatiales* ? Autrement dit, jusqu'à quel point le réseau peut-il être considéré comme responsable de la production de ces formes d'organisation, *i.e.* des variations de densité, des prix, et plus généralement des localisations des activités urbaines dans la ville ? L'une de ces deux formes géométriques est-elle, par essence, plus sujette à générer de l'étalement urbain ? L'hypothèse qu'il existe bien une relation causale entre la géométrie du réseau viaire de ces villes et les formes d'organisation qu'on y observe n'est bien sûr pas nouvelle, mais d'autres facteurs sont généralement invoqués pour expliquer ces différences : préférence de la culture américaine pour la vie à la campagne ou adaptation plus rapide des prix du marché aux Etats-Unis (Ghorra-Gobin, 1998). L'originalité de notre approche se situe dans la volonté d'étudier cette causalité par une modélisation quantitative : peut-on imaginer un modèle de développement spatio-temporel d'une ville qui soit générique, c'est-à-dire le même pour toutes les villes, et qui, en fonction de la géométrie du réseau sur laquelle il s'exécute, produise des répartitions d'activités différenciées ? A notre connaissance cela n'a jamais été fait. Notre objectif est d'aborder cette question par le biais de la simulation informatique.

1.2. *… abordée par la simulation orientée agents (SOA)*

Pour simuler des systèmes dynamiques, l'informatique met à disposition de nombreux formalismes. En choisir un entraîne une mise en balance entre d'une part la puissance expressive des formalismes disponibles et, d'autre part, la simplicité de leur mise en œuvre. La simulation orientée agent (SOA), fille des systèmes multi-agents (SMA) et de la programmation orientée objets (POO) (Drogoul *et al.*, 2002) ; (Treuil *et al.*, 2008), est une technique informatique qui permet une représentation fine et directe des objets et des processus du système modélisé. Par rapport à d'autres méthodes informatiques également dédiées à la spécification et au calcul de systèmes dynamiques discrets (réseaux de Petri, méthodes formelles, systèmes de réécriture, les différentes familles d'automates, etc.), elle représente souvent un bon compromis entre expressivité et simplicité. Le haut niveau d'abstraction et la malléabilité des concepts des SMA en font très souvent l'approche de choix en simulation sociale. Le modèle orienté agents est alors le prolongement presque « naturel » de l'énoncé thématique qui lui a donné naissance. De plus, les SMA ont depuis leurs débuts développé des métamodèles organisationnels d'inspiration sociale (agent-groupe-rôle notamment, (Gutknecht *et al.*, 1998)), dont il a été montré la capacité à aborder des questions complexes de modélisation (Abrami, 2004).

Toutefois, cette souplesse a ses contreparties : les SMA ne sont pas au sens strict un formalisme de calcul, mais plutôt une riche boîte à outils systémique et décentra-

lisée. Il n'existe pas de définition universelle de ce que sont un agent et un SMA, au contraire d'un automate cellulaire par exemple. De ce fait, le seul objet formel, calculable et déterministe de la chaîne de construction collaborative d'un modèle à base d'agents est le programme final, ce qui peut poser problème si celui-ci n'est pas vérifié par d'autres membres de l'équipe que le développeur. De plus, du fait de la complexité des règles qu'ils incorporent, une représentation analytique est très souvent hors d'atteinte, et il n'y a alors pas d'autre approche possible que simulatoire pour étudier le modèle. Enfin, l'absence de spécifications non ambigües des modèles agents limite les possibilités de reproduction des expériences. Ainsi par rapport à des approches plus parcimonieuses, on peut considérer qu'une approche simulatoire à base d'agents offre une plus grande finesse de description, mais induit aussi un glissement de la difficulté de la spécification statique du modèle à l'analyse statistique de son expression dynamique. Le modèle construit permettra de reproduire une palette plus étendue de comportements du système étudié, mais sa dynamique sera moins intelligible et il sera de fait moins contrôlable. Si la facilité apportée par la SOA pour enrichir les modèles est un avantage indéniable, ces derniers points posent d'importantes et difficiles questions relatives à la *validation* des modèles construits. Celles-ci sont au cœur des réflexions de la communauté modélisatrice en SHS (Amblard *et al.*, 2006).

Tout en cherchant à résoudre ces problèmes, la communauté modélisatrice en sciences sociales témoigne d'un grand enthousiasme pour la simulation informatique, et pour la SOA en particulier, car celle-ci permet d'attaquer des problèmes pour lesquels les approches mathématiques se révèlent rapidement limitantes. La géographe Lena Sanders a récemment analysé, à travers une présentation comparée de quelques modèles urbains, les raisons du succès de la SOA en géographie (Sanders, 2007). Elle et d'autres auteurs (Pumain, 2008), (Batty, 2008) ont expliqué les raisons de son appropriation par les géographes : approche distribuée, spatialité explicite, finesse dans l'expression des interactions spatiales, possibilité de faire correspondre à tout objet géographique un agent/objet informatique, souplesse du formalisme, puissance expressive du concept d'agent. La variété des questions pour lesquelles la SOA est mobilisée témoigne aussi de ses potentialités. Récemment, elle a notamment été utilisée pour étudier : la mobilité urbaine, en générant des trajets virtuels à partir d'une enquête ménages-déplacements (Banos *et al.*, 2008) ; les trajets d'enfants dans les rues après l'école et la croissance des boulangeries en réponse à leurs visites, à l'échelle d'un arrondissement parisien (Delâge *et al.*, 2008) ; la simulation de migrations résidentielles complexes dans la ville de Bogotà, modélisée simultanément à plusieurs niveaux d'échelle, celui des ménages et des logements, mais aussi celui de groupes statistiques de ménages et de logements (Quijano *et al.*, 2007). Des panoramas récents des utilisations de la SOA sur des questions de géographie urbaine ont été proposés par (Quijano, 2007), (Vanbergue, 2003) et (Marilleau, 2006) dans leurs thèses, mais surtout par (Benenson *et al.*, 2004) dans leur ouvrage de synthèse *Geosimulation*. Il ressort de ces états de l'art que dans une large mesure, la SOA est utilisée en géographie afin de reproduire à un niveau macro l'évolution d'un système spatial, dont la structure peut être dynamique, à partir de la description de ses composants, de leurs comportements et de leurs interactions (niveau micro). Le but est alors de mieux cer-

ner, de comprendre, voire de hiérarchiser l'importance des processus à l'œuvre dans le développement du système.

Notre question s'inscrit parfaitement dans le cadre dessiné par la définition précédente, et nous avons naturellement opté pour une approche simulatoire à base d'agents. La simulation des effets possibles de la géométrie des réseaux viaires sur les configurations urbaines implique de considérer des durées longues. Nous avons donc conçu SIMPOPNANO, un modèle générique, qui simule la localisation de fonctions urbaines dans une ville virtuelle, sur la période historique 1800-2000. Le modèle encapsule des agents, objets et processus géographiques communs aux deux grands types de villes considérés, et peut être exécuté sur des géométries-supports différentes. A terme, il doit permettre d'apporter un éclairage sur notre interrogation par l'expérimentation numérique.

Dans cet article, nous présentons la conception de ce modèle et de la plate-forme de simulation associée. Dans la section suivante, nous mettons en évidence quelques-uns des problèmes méthodologiques soulevés par cette modélisation, et les solutions originales mises en œuvre pour les résoudre. La section 3 décrit SIMPOPNANO de façon détaillée et isole les différents sous-modèles qui le composent. Enfin dans la section 4, nous discutons de la réalisation de la plate-forme de simulation qui propulse le modèle, ainsi que des outils utilisés.

2. Des questions de méthodologie

Dans cette section, nous abordons cinq des questions qui se sont posées à nous en phase de conception. Elles nous semblent intéressantes en ce qu'elles sont représentatives de « classes de problèmes » souvent rencontrés dans les projets de SOA en géographie. Pour chacune, nous discutons la solution apportée et nous la comparons avec d'autres solutions issues de la littérature.

2.1. *Comment choisir les niveaux d'échelles, acteurs et processus pertinents ?*

Dans la longue durée de l'histoire des villes choisie pour la simulation, les « événements » qui comptent, en termes de réorganisation de la morphologie urbaine, et en dehors des grandes opérations d'urbanisme « contingentes » à telle ou telle ville, sont surtout les impulsions de croissance (et donc de construction) données par les principaux cycles de l'innovation technologique et économique. La durée de ces cycles étant de l'ordre de quelques décennies, des itérations couvrant chacune un intervalle de dix ans ont été retenues pour définir la résolution temporelle du modèle.

Un modèle de ville complet et endogène sur cette période est hors d'atteinte. Au vu des ambitions affichées, un modèle de *localisation d'acteurs urbains* paraît pertinent et suffisant, car c'est bien la façon dont l'espace se « remplit » et s'organise dans le temps qu'il faut essayer de capturer dans un modèle pour reproduire ces motifs de répartition. Mais quels acteurs considérer ?

Dans leur ouvrage, Benenson et Torrens plaident pour une approche systématique des dynamiques urbaines au niveau le plus fin possible, celui des individus, des objets mobiles et des bâtiments, pour lesquels on peut recueillir des données *a priori* moins sujettes à erreurs (Benenson *et al.*, 2004). Leur argument, discuté par (Sanders, 2007), est le suivant : les objets géographiques classiques (le quartier, la ville) étant aussi des objets statistiques dont les frontières sont sujettes à discussion, leur utilisation dans les modèles introduit des erreurs de passage à l'échelle dont il est possible de s'affranchir en partant du niveau élémentaire, celui des individus et des objets « non modifiables ». Comme l'a exprimé Sanders, « il semble légitime de considérer une ville comme étant autre chose qu'une simple collection de ménages et de firmes, surtout si l'on s'intéresse aux propriétés du système de villes dont cette ville fait partie ». D'une part un agent/objet informatique peut bien sûr être associé à tout objet géographique, et d'autre part le choix des objets du modèle et du niveau de description dépend trivialement de la question que l'on se pose. Les difficultés rencontrées lorsque l'on modélise trop « bas » ont été par exemple illustrées par (Marceau *et al.*, 2008) dans le cas d'automates cellulaires d'utilisation du sol dans la région de Québec. Dans notre cas, un modèle à l'échelle des individus, des ménages et des firmes est inenvisageable pour attaquer ce problème : le nombre de niveaux géographiques intermédiaires à reconstruire serait trop grand. Pour contourner cette difficulté, nous avons choisi des objets spatiaux et des acteurs plus agrégés : les *quartiers* et les *fonctions urbaines*.

La *fonction urbaine* est un concept défini par les géographes à l'échelle inter-urbaine et porté par la théorie de la diffusion des innovations et la théorie évolutionnaire urbaine (Bretagnolle *et al.*, 2006). Elle caractérise un rôle joué par la ville qui la possède au sein du système des villes dont elle fait partie. A la différence d'une catégorie dans une nomenclature d'activités socio-économiques, notre définition des fonctions urbaines regroupe un ensemble d'activités de même génération, corrélées entre elles et qui passent par un même cycle d'apparition, de développement et d'obsolescence, qui affecte les dynamiques urbaines (Pumain, 2006a). Un aspect original de notre approche est le transfert de ce concept au niveau intra-urbain pour décrire les dynamiques de localisation dans la ville.

Le choix de ces acteurs se justifie par l'échelle temporelle considérée. En effet, durant les deux siècles que nous simulons, nous sommes passés de l'ère préindustrielle à une économie basée sur l'information. Pour résumer la variété de l'évolution des acteurs urbains pendant la période simulée, nous utilisons une palette de onze fonctions urbaines. Neuf fonctions non résidentielles ont été reprises au modèle SIMPOP2[2] : quatre fonctions de type CENTRALE, quatre de type CYCLE et une fonction de type politique. Les fonctions CENTRALE représentent des activités de service à la population de plusieurs niveaux hiérarchiques, tandis que les fonctions CYCLE représentent les grands cycles successifs d'innovation. La fonction CYCLE4 par exemple, qui représente les activités associées aux technologies de l'information, « rentre » dans la ville à partir de 1960, et son effectif croît jusqu'en 2000. A l'opposé, la fonction CYCLE1, qui correspond à un cycle d'innovation plus ancien (commerce maritime),

2. http ://www.simpop.parisgeo.cnrs.fr/theGenericModel/otherEntities.php/

est présente dès le début de la période simulée (1800). Son importance décroît au profit des fonctions CYCLE suivantes, jusqu'à ce qu'elle disparaisse en 1850. Nous avons ajouté deux fonctions résidentielles, RICH (cols blancs) et POOR (cols bleus). Nous en avons choisi deux pour aller au plus simple et pour rester à un niveau de détail, pour la modélisation du résidentiel, qui soit cohérent avec le niveau choisi pour la description des activités socio-économiques. L'ensemble de ces fonctions urbaines est présenté dans le tableau 1. Nous revenons dans la sous-section 2.3 sur les comportements de localisation de ces fonctions.

Chaque fonction urbaine est donc active pendant une période qui lui est propre, et toutes les fonctions ne sont pas présentes en même temps et en mêmes proportions dans la ville. Nous ne cherchons pas à endogénéiser dans le modèle la croissance de ces différentes fonctions urbaines, qui dépend de processus situés à un autre niveau, celui du système de villes (Paulus, 2004) ; (Cattan *et al.*, 1994). Le nombre d'actifs à ventiler dans la ville est donc fixé pour chaque pas de temps et pour chaque fonction, c'est une donnée d'entrée du modèle. Des profils d'évolution des fonctions, en effectif et en budget, ont été mis au point pour chacun des deux types de ville considérés, à partir des bases de données historiques élaborées au sein du laboratoire Géographie-Cités (Bretagnolle *et al.*, 2008), de sources de données historiques, et de l'expertise des géographes pour la simplification. Un exemple de profil d'évolution des fonctions au cours du temps est représenté par la figure 2.

Cette technique s'inscrit dans la lignée de celle mise en œuvre dans les *automates cellulaires contraints* d'utilisation du sol de White et Engelen (White *et al.*, 1993) et de toute cette famille de modèles (chapitre 4 de (Benenson *et al.*, 2004)). Comme dans ces modèles, la localisation est soumise à une contrainte de volume globale et exogène (ici le nombre d'actifs à localiser sur l'ensemble de la ville) qui empêche sa représentation par une approche purement locale. Cependant notre approche s'en distingue sur deux aspects :

1) dans ces mêmes AC, les catégories d'utilisation du sol manipulées sont souvent issues de nomenclatures d'activités détaillées qui font sens pour l'époque contemporaine, mais qui sont moins pertinentes que la fonction urbaine pour la période couverte par notre étude ;

2) ces catégories sont en général présentes toutes en même temps et pendant toute la période simulée, ce qui n'est pas notre cas avec les fonctions urbaines.

En plus de ces changements fondamentaux des activités socio-économiques, les moyens de transport urbains ont également bien changé depuis 1800 avec l'apparition du ferroviaire, puis de l'automobile. Nous allons voir dans la section suivante la façon dont cette évolution peut être prise en compte.

Nom	Description	Période d'activité dans le modèle
CHEF-LIEU	Fonction administrative de type capitale régionale	[1800-2000]
CENTRALE1	Service de base (ex. boulangerie, école, médecin généraliste)	[1800-1840]
CENTRALE2	Service de niveau moyen (ex. poissonnerie, collège, dentiste, cinéma)	[1800-2000]
CENTRALE3	Service de niveau élevé (ex. hypermarché, lycée, ophtalmologiste)	[1800-2000]
CENTRALE4	Service de niveau supérieur (ex. centre commercial, université, CHU, opéra)	[1950-2000]
CYCLE1	Grand commerce maritime	[1800-1900]
CYCLE2	1ère révolution industrielle (machine à vapeur, sidérurgie, textile)	[1800-1990]
CYCLE3	2nde révolution industrielle (pétrole, électricité, automobile)	[1900-2000]
CYCLE4	Technologies de l'information et de la communication (informatique, high-tech)	[1950-2000]
RICH	Fonction résidentielle à budget élevé	[1800-2000]
POOR	Fonction résidentielle à budget faible	[1800-2000]

Tableau 1. *Les neuf* fonctions urbaines *reprises de* SIMPOP2 *et utilisées comme des utilisations du sol dans* SIMPOPNANO*, et les deux* fonctions résidentielles

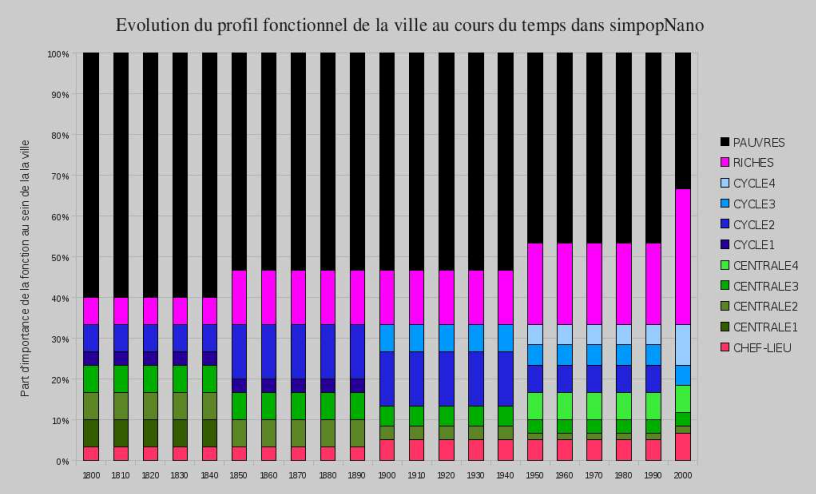


Figure 2. *Evolution de la composition fonctionnelle de la ville au cours du temps dans* SIMPOPNANO

2.2. *Quelle représentation informatique de la ville en évolution spatio-temporelle ?*

Simuler une ville en évolution spatio-temporelle impose de se choisir une représentation de l'espace qui soit cohérente et pertinente pour toute la période abordée par le modèle. La tendance est à la représentation la plus réaliste possible des espaces géographiques dans les modèles, dans la lignée des avancées des SIG et des logiciels de cartographie 3D. Ces outils permettent une représentation géométrique des objets géographiques aussi précise que souhaitée. Ils sont de plus en plus utilisés en SOA, et améliorent notamment la communication en offrant des visualisations réalistes qui facilitent le retour des experts sur le modèle. En revanche, il semble qu'à l'heure actuelle peu de modèles tirent parti de la finesse des représentations spatiales vectorielles pour décrire des processus géométriques complexes.

De notre côté, deux raisons nous incitent à préférer une représentation de l'espace au moyen d'un graphe, plutôt que de travailler avec des données vectorielles issues de SIG :

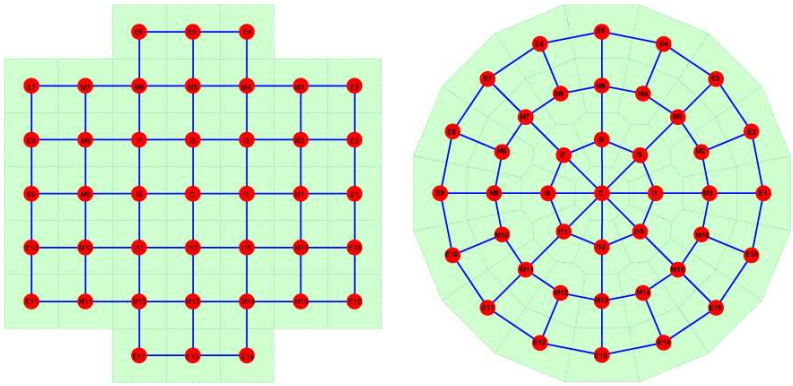
– on ne cherche pas à simuler l'évolution dans le temps de la forme géométrique de la ville et de ses quartiers, mais celle de son organisation interne ;

– l'échelle temporelle considérée rend arbitraire la définition de processus décrivant l'évolution de la forme physique de la ville.

Ainsi à l'instar des travaux présentés dans (Mathis, 2003) ou (Badariotti *et al.*, 2007), nous représentons l'espace intra-urbain par un graphe dynamique de quartiers, plutôt qu'à l'aide d'une représentation surfacique plus fine. Les nœuds du graphe correspondent aux quartiers de la ville et les arêtes aux principaux axes du réseau de transports. Nous faisons l'hypothèse forte que *les deux réseaux viaires ont une forme prédéterminée et qui reste fixe pour la période considérée*, car ils sont issus de la législation, du cadastre et de la culture pour les Etats-Unis, et de l'histoire millénaire de la ville pour l'Europe. Le « patron » du graphe est statique, c'est un paramètre du modèle. Il fixe une fois pour toute la topologie du réseau viaire de la ville simulée et les formes géométriques potentielles de son expansion. Notre modèle n'étant pas un modèle d'utilisation du sol, il ne prend pas en compte l'évolution des surfaces représentant les infrastructures de transport.

L'activation des nœuds du graphe ainsi que les poids associés aux arêtes sont dynamiques. A chaque nœud-quartier est associé une valeur *d'accessibilité* dynamique, réévaluée à chaque pas de simulation, et qui dépend à la fois des plus courts chemins dans le graphe mais également des changements dynamiques d'occupation des quartiers. Nous avons mis en place deux règles de calcul concurrentes de cette accessibilité entre lesquelles l'utilisateur peut choisir au lancement de la simulation.

Les graphes théoriques utilisés pour la ville européenne et la ville américaine ont été dessinés par Bretagnolle et Delisle (Delisle, 2008) et sont représentés en figure 3. Nous revenons de façon plus formelle sur cette représentation de l'espace dans la section 3.



a) b)

Figure 3. *Graphes théoriques des réseaux viaires de a) la ville états-unienne et b) la ville européenne (Delisle, 2008)*

2.3. *Comment représenter les processus complexes de choix de localisation des fonctions urbaines ?*

La formalisation des logiques guidant les choix de localisations des acteurs urbains fait partie des études les plus répandues dans la littérature agents en analyse spatiale. Il s'agit en effet d'une question qu'étudient les modélisateurs de l'économie géographique et ceux de la géographie économique. Ces deux courants se distinguent dans le rôle accordé à l'espace (Sanders, to be published). Celui-ci est central dans le cas de la géographie économique, et les modélisateurs de ce courant ont multiplié les efforts pour qualitativement mieux intégrer son influence, quitte à développer des modèles complexes et à délaisser la résolution analytique ; de son côté, l'économie urbaine se focalise plus sur des modèles centrés sur la maximisation d'une fonction d'utilité (par exemple (Caruso *et al.*, 2007)). Le point crucial réside alors dans la définition de cette fonction d'utilité des agents. Celle-ci doit tenir compte des aménités espérées d'une localisation particulière, mais il est parfois difficile de capturer dans un ensemble d'équations continues le jeu complexe observé entre les acteurs. La littérature économique à ce sujet est riche et abondante, mais nous essayons de faire quelque chose de différent.

Le comportement de localisation des fonctions urbaines de SIMPOPNANO est basé sur l'attribution de *scores* par les fonctions urbaines aux quartiers, ce qui permet aux fonctions d'ordonner leurs quartiers par ordre de désirabilité. L'ordre dans lequel les fonctions expriment leurs requêtes de localisation au cours d'une itération ne dépend dans notre modèle que de la propension à payer de chaque fonction. Notre hypothèse n'est pas celle d'un équilibre stable ou instantané à la Alonso, mais d'une structure évolutive dans laquelle les fonctions peuvent se substituer les unes aux autres dans l'espace urbain en fonction de la capacité dynamique qu'elles acquièrent par ailleurs

dans le système des villes. Nous nous plaçons dans le cadre d'une théorie géographique évolutive et multi-échelles des systèmes de villes (Pumain *et al.*, 2009b), et non pas dans celui d'une théorie d'économie urbaine.

Toutes les fonctions non résidentielles ont en commun de rechercher une accessibilité maximale, mais elles se différencient par le type de fonctions à côté desquelles elles cherchent à s'installer ou au contraire, dont elles cherchent à s'éloigner. Les deux fonctions résidentielles (riches et pauvres) rajoutées se distinguent uniquement par leur budget, et leurs « stratégies » de localisation sont identiques : riches comme pauvres cherchent à se localiser dans les quartiers les plus accessibles, mais leurs possibilités budgétaires pour le faire sont bien sûr différentes.

Dans les villes, les besoins en espace varient bien sûr selon les fonctions et selon les époques. Dans SIMPOPNANO, nous simplifions et l'espace nécessaire pour localiser un actif, toutes fonctions confondues, reste fixe au cours du temps et d'une fonction à l'autre. Tous les quartiers ont la même capacité d'accueil initiale, tous les emplacements au sein de ces quartiers peuvent symboliquement servir à localiser un actif et un seul.

De plus, sur la période simulée, les quartiers des villes ont été amenés à se transformer afin de se requalifier, avec l'évolution des technologies et des changements sociétaux. D'anciens quartiers manufacturiers et des friches industrielles ont été redéveloppés, des quartiers populaires gentrifiés, etc. Le modèle doit permettre de rendre compte de ces redéveloppements et de ces relocalisations d'activités. Pour cela les fonctions peuvent tout à la fois cohabiter au sein d'un quartier ; se succéder dans un quartier ; quitter un quartier pour partir s'installer dans un autre, du fait de leur incapacité nouvelle à payer le loyer (qui évolue dynamiquement, *cf.* section 3) ; pour certaines, disparaître pendant la simulation (les CYCLE d'innovation obsolescents).

Cependant, nos règles de localisation n'intègrent pas les coûts supplémentaires qui sont liés au nécessaire déplacement des personnes et des marchandises lors d'une relocalisation. Nous restons donc à un niveau agrégé, et la reproduction des phénomènes historiques mentionnés précédemment ne sera permise que « symboliquement ». L'expérimentateur pourra avoir une lecture interprétative plus fine des sorties du modèle à la lumière de sa connaissance géographique, mais sans perdre de vue le risque de sur-interprétation des résultats, l'un des six écueils identifiés par Lee (1973).

2.4. *Comment décider de la qualité des simulations en vue d'automatiser le calibrage ?*

2.4.1. *Attribution de scores aux configurations spatiales produites par le modèle*

Pour quantifier la qualité des configurations produites, et pouvoir les classifier, plusieurs scores ont été définis. De nombreux points de vue peuvent co-exister sur ce qui fait d'une simulation une « bonne » simulation. Lorsque l'objectif est de calibrer le modèle, la qualité est synonyme de capacité à générer une ville virtuelle qui puisse

être considérée comme réaliste par l'observateur, relativement au phénomène étudié. Ces scores sont alors des indicateurs de la proximité des valeurs simulées aux données dont on dispose, à plusieurs niveaux d'observation.

Ces seuls indicateurs constituent bien sûr un point de vue incomplet sur ce qui fait structure dans la ville et de la même façon les différences de répartition ne sont pas les seules qui existent entre les villes états-uniennes et européennes. Cependant ces mesures sont cohérentes avec les objectifs assignés à cette modélisation.

Notre première approche consiste donc à construire des scores basés sur la répartition des densités et des prix dans la ville. Dans la mesure où la texture spatiale de nos villes virtuelles n'est pas très détaillée, calculer une valeur de gradient n'aurait pas beaucoup de sens. Pour ce qui est de la comparaison des valeurs simulées aux données, nous pensons qu'il suffit à ce niveau de généralité de vérifier que la densité moyenne et ses variations restent dans des ordres de grandeur compatibles avec les valeurs moyennes des villes américaines ou européennes.

Nous pouvons également mesurer les écarts entre les sorties produites par le modèle selon que nous l'exécutons sur la géométrie européenne ou sur la géométrie nord-américaine. Cet écart des valeurs obtenues avec les deux géométries doit être comparable à l'écart observé entre la valeur en moyenne dans les villes européennes et celle en moyenne dans les villes états-uniennes.

Chacun de ces scores est associé à la configuration spatiale globale de la ville simulée à un instant donné. Ils peuvent être calculés à l'issue de chaque pas de temps du modèle. Dans un premier temps du calibrage, nous ne nous intéressons qu'aux valeurs de ces scores à l'instant final. Par la suite, pour vérifier la capacité du modèle à produire une évolution temporelle cohérente de ville, nous devrons nous assurer non plus uniquement de la cohérence des scores à l'instant final, mais de la cohérence de la série temporelle de leurs valeurs. Il s'agit d'une question plus difficile que nous laissons en suspens pour l'instant.

Il est possible d'imaginer d'autres scores qui seront testés dans un second temps de l'expérimentation : cœfficients de corrélation entre activités de quartiers voisins, indicateur de symétrie de la densité par rapport au centre par exemple. La définition d'un score devra être affinée pour capturer au mieux la notion de ville que nous voulons capturer avec le modèle.

L'intérêt de l'attribution systématique de scores aux configurations produites est qu'ils peuvent ensuite servir de valeurs de référence pour l'automatisation du calibrage.

2.4.2. *Automatisation du calibrage*

Comme nous l'avons dit, le calibrage constitue souvent une étape épineuse dans les projets de simulation orientée agents, du fait du grand nombre de dimensions de l'espace des paramètres. L'utilisation de techniques de calibrage automatique constitue une avancée réelle sur ces aspects. L'idée est d'aller plus loin que le simple en-

chaînement de simulations selon un plan d'expérience planifié *a priori*, en orientant le choix du prochain jeu de valeurs de paramètres à tester de façon à trouver le jeu de valeurs des paramètres qui optimise la valeur obtenue pour une fonction de référence. Dans les approches dites *génétiques* (en référence à l'inspiration évolutionniste de ces algorithmes), cette fonction est appelée *fonction de fitness* ou encore *fonction objectif*. Dans notre cas, cette fonction objectif correspondra à l'un des *scores* définis dans la section précédente. Le principe général de fonctionnement d'un algorithme génétique est représenté par la figure 4. Une *population* d'instances du modèle est instanciée avec des valeurs de paramètres différentes. A l'issue des simulations, les jeux de valeurs ayant donné les meilleurs résultats au sens de la fonction objectif (les « parents ») sont sélectionnés et sont combinés et *mutés* pour produire de nouveaux jeux de valeurs (phase de *recombinaison/mutation*), les « enfants ». On relance alors la simulation pour un jeu de modèles composé des « parents » et des « enfants », et ainsi de suite jusqu'à atteindre un critère d'arrêt de l'algorithme, qui correspond soit à l'atteinte d'une valeur seuil de la fonction objectif, ou à l'atteinte du nombre maximal de simulations défini par l'expérimentateur. En économie, ces techniques sont par exemple mises en œuvre par Deissenberg sur des modèles de marchés (Deissenberg, 2006).

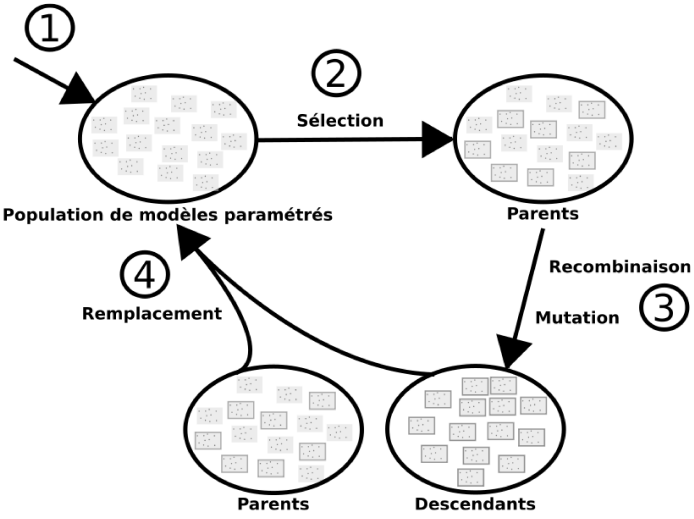


Figure 4. *Principe général du calibrage automatique de modèle par des algorithmes génétiques (d'après (Calvez, 2007))*

Le calibrage automatique de modèles orientés agents pose des questions spécifiques, puisque pour chaque paramètre associé aux agents, le nombre de configurations possibles augmente de façon exponentielle avec le nombre d'agents du modèle. Disposer d'algorithmes efficaces dans ce domaine devrait contribuer aux développements de l'expérimentation *in silico* en géomatique. Cette étape, au même titre que la définition de plans d'expériences rigoureux, l'analyse statistique en profondeur des sorties, ou l'adoption de normes et d'outils partagés par la communauté, constitue un

pas important pour l'adoption et la maîtrise des méthodes computationnelles d'expérimentation. La construction d'algorithmes d'optimisation de type génétiques, pour des modèles multi-agents, a été l'objet d'une thèse récente, avec des applications sur des modèles simples de colonies de fourmis sous Netlogo (Calvez, 2007), (Calvez *et al.*, 2005). L'interfaçage de ces algorithmes avec SIMPOPNANO est en cours de réalisation (Monzie, 2009).

2.5. *Comment coupler deux modèles indépendants et opérant à des niveaux d'échelle différents pour simuler des dynamiques multi-scalaires ?*

Les systèmes urbains sont des systèmes complexes et hiérarchiques, et il existe des interdépendances entre leurs différents niveaux d'organisation, le système des villes, les villes et les quartiers (Pumain, 2006b) ; (Pumain, 2006c). Intégrer et articuler ces différents niveaux dans un modèle doit permettre de tester des hypothèses sur la nature des processus d'inter-dépendance que les niveaux entretiennent. Pour ce faire, nous avons décidé de tenter un couplage entre SIMPOPNANO et SIMPOP2.

L'approche générique et comparative que nous avons mis en place avec SIMPOPNANO est directement inspirée de celle qui a été suivie par l'équipe P.A.R.I.S du laboratoire *Géographie-Cités* dans leurs projets de simulation orientée agents de systèmes de villes, SIMPOP2[3] et EUROSIM (Sanders *et al.*, 2005). Ces deux modèles sont les successeurs de SIMPOP, qui fut lui-même la première application des SMA en géographie (Bura *et al.*, 1996).

L'objectif de SIMPOP2 était de capturer dans un modèle générique les éléments moteurs des morphogenèses différentes ayant eu lieu dans les systèmes de villes européen, états-unien et sud-africain. Et, de façon complémentaire, de mettre en lumière les processus spécifiques à l'histoire de chacun de ces systèmes (Bretagnolle *et al.*, 2009) ; (Pumain *et al.*, 2009b). La période historique prise en compte pour la modélisation était 1650-2000.

Mettre les deux modèles SIMPOP2 et SIMPOPNANO côte à côte ne suffit pas à constituer un modèle multi-niveaux. La construction de celui-ci passe par la définition préalable d'une ontologie géographique multi-niveaux dans laquelle sont explicitées les relations qu'entretiennent les acteurs des différents niveaux (Pumain *et al.*, 2009a). Notre solution a consisté à identifier, en amont de la conception de SIMPOPNANO, une description commune aux deux niveaux, qui représente à la fois un effet de la dynamique du système des villes sur la ville, et qui puisse exprimer une dynamique interne à la ville susceptible d'avoir des effets en retour sur le système. C'est donc dans cette optique du couplage que la notion de *fonction urbaine* a été infléchie pour permettre l'articulation entre les deux modèles. Il a fallu identifier un ensemble d'attributs mesurables qui puissent représenter une sortie du niveau système de villes et une entrée du système ville, et inversement.

3. http ://www.simpop.parisgeo.cnrs.fr/

L'hypothèse non vérifiée mais plausible que nous avons utilisée pour spécifier le modèle SIMPOPNANO est que les activités qui s'implantent localement dans une ville se répartissent dans l'espace de la ville selon une logique compatible avec la typologie de fonctions adoptée dans SIMPOP2 pour décrire les spécialisations fonctionnelles des villes au sein du système des villes. Cette représentation constitue une grande simplification de la dynamique intra-urbaine des activités, mais c'est l'hypothèse qui permet sans doute le mieux de simuler les effets éventuels en retour de la dynamique intra-urbaine sur la trajectoire qu'une ville peut prendre dans le système des villes. Ainsi, dans l'optique du couplage, nous pouvons substituer aux profils d'évolution temporelle des fonctions urbaines, injectés en entrée de SIMPOPNANO, les sorties du modèle SIMPOP2, qui justement calcule l'évolution de la spécialisation fonctionnelle de chaque ville du système.

L'entité fonction urbaine sert donc de « pont » entre les deux modèles, et elle est affectée à la fois par les règles de SIMPOP2 et par celles de SIMPOPNANO. L'acquisition des fonctions par la ville, et leurs croissances différenciées, se jouent au niveau du système des villes, dans lequel la ville est inscrite dans des réseaux de collaboration et de concurrence. Ces données sur les fonctions (effectifs et budgets), calculées comme le résultat des interactions avec les autres villes, sont les sorties de SIMPOP2 et les entrées de SIMPOPNANO. Celui-ci simule alors la localisation à l'intérieur de la ville de ces effectifs. Pour « boucler la boucle », l'hypothèse que nous faisons est que la performance spatiale des fonctions au sein de la ville (*i.e.* la qualité de leurs implantations), mesurée en sortie de SIMPOPNANO, va jouer en retour sur *l'attractivité* de la ville parmi ses concurrentes, et va donc constituer un facteur supplémentaire pris en compte par les innovations futures pour choisir leurs villes d'implantation dans SIMPOP2.

3. SIMPOPNANO : un modèle de localisation de fonctions urbaines dans la ville

3.1. *Principe général*

Le principe général de notre approche est le suivant : à chaque pas de temps simulé, les fonctions urbaines possèdent toutes un effectif d'actifs qu'elles doivent localiser dans la ville, ainsi qu'un budget. La ville est constituée d'un nombre fini de quartiers qui s'inscrivent dans un réseau, représenté par un graphe. La topologie de ce graphe est immuable pendant le temps de simulation, et est donnée *a priori*. Les quartiers sont les nœuds du graphe. Ils sont décrits par des attributs dont les valeurs peuvent changer pendant la simulation. Chaque quartier de la ville possède un certain nombre de caractéristiques (accessibilité, potentiel économique, loyer). L'accessibilité et le potentiel économique des quartiers les rendent plus ou moins *attractifs* aux yeux des différentes fonctions. Certaines de ces caractéristiques (potentiel économique, prix du loyer) évoluent pendant la simulation, et sont *calculées* par le modèle. Pour l'accessibilité, l'expérimentateur peut choisir entre un modèle d'évolution où celle-ci est donnée de façon exogène (Delisle, 2008) ou bien un modèle où celle-ci est recalcu-

lée dynamiquement à chaque pas de temps, en ne tenant compte que des plus courts chemins aux quartiers déjà occupés.

Les onze fonctions se distinguent les unes des autres par leurs effectifs et leurs budgets, mais également par des « stratégies » de localisation variées. Sur la base de leurs « stratégies » , toutes les fonctions calculent à chaque pas de temps, et pour chaque quartier, un score qu'elles leurs attribuent. Sur la base de ces scores, chacune établit une liste ordonnée de ses quartiers préférés. Les fonctions sont « aveugles », elles optimisent l'endroit placé et non pas la population placée. Cela signifie qu'une fonction localise autant d'employés dans un quartier que le lui permettent son budget, le loyer du quartier, ainsi que le nombre d'emplacements libres qu'il contient. Ce mécanisme d'allocation est complètement déterministe. Des mécanismes économiquement plus sophistiqués tendraient à maximiser le nombre d'employés localisés. Ces mécanismes impliquent des hypothèses géographiques fortes sur la nature des mécanismes en jeu. Notre hypothèse est que les qualités de la localisation importent plus que l'arbitrage monétaire.

Cependant, pour nuancer ce parti pris, notre objectif informatique est de disposer d'une plate-forme qui permette de tester différentes fonctions de localisation (*cf.* sous-section 4.2). Ce n'est que dans un second temps que les tests et l'exploration permettront de sélectionner les règles d'évolution les plus adaptées pour rendre compte des différences d'organisation spatiale que l'on étudie.

La capacité d'un quartier intervient dans le modèle pour limiter la possibilité de s'y installer une fois qu'il est saturé. Cependant, le fait qu'un quartier soit « plein » n'empêche pas les fonctions de le trouver attractif, donc de le classer haut dans leur liste de quartiers préférentiels, et de lui émettre des requêtes de localisation. Celles-ci resteront insatisfaites tant que le quartier sera saturé, mais cette désirabilité du quartier fait gonfler son loyer (*cf.* équation [1] dans la sous-section 3.2.1), et dynamiquement, cette augmentation peut le rendre inabordable aux fonctions aux budgets les plus faibles. Le cas échéant, celles-ci doivent alors se relocaliser dans des quartiers moins chers, laissant la place libre aux fonctions les plus riches.

A la fin de chaque itération du modèle, les variables des fonctions (budget, effectif) sont mises à jour à partir des données d'entrée et les variables des quartiers (loyers et potentiels économiques) sont recalculées en regard des résultats de l'itération écoulée.

3.2. *Objets quartiers et agents fonctions*

3.2.1. *La ville est un réseau de quartiers...*

Un quartier q_j est un objet décrit par plusieurs observables. (X_j, Y_j) représentent les coordonnées spatiales du centre du quartier. A_j^t est la valeur de l'accessibilité de q_j à la date t. C'est un nombre qui représente le coût moyen d'un déplacement de q_j à n'importe quel autre quartier de la ville. Nous avons conçu deux modèles concurrents d'évolution de cette valeur. Ils intègrent à la fois la nodalité de q_j (le nombre de quar-

tiers auxquels il est relié par des axes du réseau), ainsi que les vitesses moyennes de circulation dans le réseau (Delisle, 2008). C_j^t représente la capacité d'accueil de q_j à la date t et correspond au nombre d'actifs que le quartier peut encore accueillir avant d'être complet. On suppose donc que les quartiers ont une capacité d'accueil limitée, et qui est la même à t_0 pour tous les quartiers du modèle. P_j^t est le loyer d'un emplacement dans le quartier q_j, et correspond à ce que doit payer une fonction pour chacun de ses employés localisés dans q_j à chaque pas de temps ($P_j^t \in \mathbb{R}^+$). $locDemands_{i,j}^t$ désigne le nombre de requêtes de localisation exprimées par la fonction f_i au quartier q_j à la date t. Cette valeur est utilisée pour calculer la valeur de τ_j^t, le taux de variation de la popularité de q_j.

Le prix P_j^t du quartier q_j à la date t nous est donné par la formule suivante :

$$P_j^t = A_j^t \cdot \tau_j^t \qquad [1]$$

avec

$$\tau_j^t = \frac{\sum_{i=0}^{M} locDemands_{i,j}^t}{\sum_{i=0}^{M} locDemands_{i,j}^{t-1}} \qquad [2]$$

L'équation [2] définit le taux de variation de la désirabilité de q_j à une date donnée comme le quotient du nombre de requêtes reçus par q_j de la part de l'ensemble des fonctions à cet instant, sur le nombre de ces requêtes à l'instant précédent. L'équation [1] exprime l'idée que le prix d'un quartier augmente quand son accessibilité augmente ; idem pour le taux de désirabilité : plus un quartier est désirable, plus il est cher.

3.2.2. … *colonisé par des fonctions urbaines*

La fonction f_i est décrite par plusieurs observables. $E_i^t (= \sum_{j=0}^{N} E_{ij}^t)$ représente le nombre d'employés de f_i à l'instant t. L_i^t représente le nombre d'employés de f_i à localiser à t (c'est-à-dire qui ne sont pas encore localisés dans un des q_j). W_i^t est le budget de f_i et $pref_i = (pref_{i,0} \dots pref_{i,k} \dots pref_{i,M})$ est le vecteur des facteurs d'attraction/répulsion de f_i. $pref_{i,k} \in [-1..1]$ est un cœfficient qui symbolise l'attraction/répulsion de f_i pour f_k lors du calcul du potentiel économique $EP_{i,j}^t$.

La valeur de $EP_{i,j}^t$ exprime l'attrait qu'exerce q_j pour f_i à la date t grâce à sa composition fonctionnelle. Ce potentiel est défini par :

$$EP_{i,j}^t = \sum_{k=0}^{M} pref_{i,k} \cdot \sum_{l=0}^{N} E_{k,l}^t / c_{j,l} \qquad [3]$$

L'équation [3] indique que la valeur de $EP_{i,j}^t$ dépend de la configuration fonctionnelle existante dans le quartier q_j considéré, mais également dans tout autre quartier q_l, en proportion inverse du coût $c_{j,l}$ du chemin qui relie à q_j à q_l dans le graphe. Il s'en suit qu'un même quartier n'a pas le même potentiel économique suivant la fonction considérée.

$S_i^t = \left(S_{i,1}^t \dots S_{i,j}^t \dots S_{i,N}^t\right)$ est le vecteur des scores attribués par f_i aux N quartiers du modèle à la date t et enfin $prefLocalisations_i^t = (q_{i_1} \dots q_{i_N})$ est la liste des quartiers triée par ordre de désirabilité au sens de f_i, avec comme critère de tri $S_{i,i_1} \geq \dots S_{i,i_k} \geq \dots \geq S_{i,i_N}$.

Le score $S_{i,j}^t$ attribué par la fonction f_i au quartier q_j à la date t nous est donné par la formule :

$$S_{i,j}^t = s_{\alpha,\beta}(A_j^t, EP_{i,j}^t) = \alpha \cdot A_j^t + \beta \cdot EP_{i,j}^t \quad [4]$$

L'équation [4] signifie que le score attribué par f_i à q_j à t dépend dans une proportion α de l'accessibilité de q_j et dans une proportion β du potentiel économique de q_j au sens de f_i. α et β sont deux paramètres globaux du modèle. Comme pour les autres sous-modèles, la fonction $s_{\alpha,\beta} : (A_j^t, EP_{i,j}^t) \rightarrow S_{i,j}^t$ n'est pas fixée *a priori*. C'est une boîte noire dans laquelle on peut inscrire n'importe quelle fonction : la plate-forme informatique développée pour l'expérimentation est neutre par rapport aux règles qu'elle incorpore.

Dans un premier temps, nous instancions s comme une combinaison linéaire de ses deux variables, mais d'autres fonctions sont à tester, *e.g.* $S_{i,j}^t = (A_j^t)^\alpha \cdot (EP_{i,j}^t)^\beta$.

3.3. *Paramètres*

Certaines valeurs manipulées dans le modèle sont constantes sur toute la durée de simulation. Ce sont des paramètres du modèle, dont l'expérimentateur fixe la valeur avant chaque simulation.

3.3.1. *Paramètres liés à l'espace et au réseau de transports urbains*

$\mathbf{A_j^t}$ est l'indicateur d'accessibilité de q_j aux dates $t = 0, \cdots, t = T$. Il représente la distance moyenne séparant q_j de n'importe quel autre quartier du modèle. $\mathbf{C_0}$ représente la capacité initiale des quartiers. C'est le nombre d'actifs que peut contenir au maximum un quartier à tout instant de la simulation. Enfin $c_{i,l}$ est le coût du plus court chemin reliant les quartiers q_i et q_l dans le graphe. Ce coût est calculé à l'initialisation du modèle.

3.3.2. *Autres paramètres*

$\mathbf{M}$ est le nombre de fonctions urbaines (11 actuellement) et $\mathbf{N}$ le nombre de quartiers (41 dans les graphes présentés). $\mathbf{\Delta E_{i,i\in 1\dots M}^{t,t\in 0\dots T}}$ est le nombre de nouveaux actifs de f_i à t, et $\mathbf{\Delta W_{i,i\in 1\dots M}^{t,t\in 0\dots T}}$ l'incrément de son budget. α et β représentent les poids donnés respectivement à *l'accessibilité* et au *potentiel économique* dans le calcul du score d'un quartier pour une fonction ($\alpha, \beta \in [0 \cdots 1]$). Enfin $\mathbf{pref_i}$ est le vecteur des facteurs d'attraction/répulsion de la fonction i envers les M fonctions du modèle. Il traduit la synergie entre f_i et les autres fonctions.

3.4. *Règles d'évolution*

Les variables des quartiers et des fonctions évoluent suivant des règles que nous présentons ci-dessous. L'ordre dans lequel nous présentons ces règles correspond à leur ordre d'occurrence pendant une itération du modèle. Nous les avons séparées en deux catégories :

– les règles *exogènes*, qui correspondent à des mises à jour de valeurs de paramètres. Ce sont des *commandes* envoyées au système depuis l'extérieur de celui-ci ;

– les règles *endogènes*, qui correspondent à des mécanismes internes au modèle, et qui aboutissent à des modifications des valeurs des variables du fait des *calculs* effectués par le modèle.

3.4.1. *Règles exogènes*

R1) *Mise à jour des paramètres : effectifs et budgets des fonctions*

$$\forall i \in \{1, \dots \mathbf{M}\}, E_i^t = E_i^t + \mathbf{\Delta E_i^t} \text{ et } L_i^t = L_i^t + \mathbf{\Delta E_i^t}$$

$$\forall i \in \{1, \dots, \mathbf{M}\}, W_i^t = W_i^t + \mathbf{\Delta W_i^t}$$

R2) *Mise à jour du paramètre : indices d'accessibilité des quartiers du réseau urbain*

$$\forall j \in \{1, \dots, \mathbf{N}\} \text{ , mettre à jour la valeur de } \mathbf{A_j^t}$$

3.4.2. *Règles endogènes*

R3) *Mise à jour des scores et de la liste des quartiers préférés de chaque fonction*

$$\forall i \in \{1, \dots, \mathbf{M}\}, \text{ calculer } S_i^t \text{ (\textit{cf.} équation [4])}, \text{et trie } prefLocalisations_i^t$$

R4) *Mise à jour des loyers des quartiers*

$$\forall j \in \{1, \dots, \mathbf{N}\}, \text{ calculer } \tau_j^t \text{ (\textit{cf.} équation [2])}, \text{et } P_j^t \text{ (\textit{cf.} équation [1])}$$

R5) *Paiement des emplacements occupés par les fonctions pour leurs employés déjà localisés* (*cf.* Annexe 7.1)

R6) *Localisation des employés qui ne sont pas localisés* (*cf.* Annexe 7.2)

R7) *Mise à jour de la matrice des potentiels économiques EP*

$$\forall i \in \{1 \dots \mathbf{M}\}, \ j \in \{1 \dots \mathbf{N}\}, \text{ calculer } EP_{i,j}^t \text{ (\textit{cf.} équation [3])}$$

4. Réalisation

4.1. *Outils utilisés*

Le modèle a été implémenté dans un logiciel écrit avec le langage orienté objets Objective-C, et en s'appuyant sur la bibliothèque logicielle de simulation orientée

agents SWARM[4](Minar *et al.*, 1996) du Santa Fe Institute. Le logiciel est fait d'environ 7500 lignes de code dans sa version actuelle, et il est peut être exécuté sur toutes les plates-formes pour lesquelles Swarm est disponible (Windows, Linux et MacOS notamment).

Notre choix s'est porté sur SWARM pour deux raisons : d'une part il s'agissait de la bibliothèque utilisée dans le modèle SIMPOP2 dont nous avons repris le code, pour faciliter les aspects techniques du couplage SIMPOPNANO-SIMPOP2 ; d'autre part SWARM offre un métamodèle structurel hiérarchique et récursif qui le rend particulièrement adapté à la conception de modèles multi-niveaux de systèmes complexes. Même si sa prise en main est reconnue comme étant moins immédiate qu'une plate-forme comme Netlogo (Castle *et al.*, 2006), elle nous semble plus adaptée pour faire de l'expérimentation en profondeur, à la fois pour des raisons de performance (temps d'exécution) et du fait que, comme il est écrit en objective-C, le code peut s'appuyer facilement sur n'importe quelle bibliothèque écrite en langage C.

Deux modes d'exécution sont possibles : un mode graphique avec une interface graphique interactive, qui permet à l'utilisateur de visualiser les sorties en temps réel (évolution des cartes et des courbes de valeurs des variables), mais aussi de réaliser dynamiquement des appels de messages ou de changer les valeurs des paramètres. Ce mode nous est utile à des fins de démonstration, et pour déboguer. La figure 5 représente une capture d'écran du modèle en train de tourner en mode graphique. L'autre mode est un mode *batch* qui permet d'enchaîner les simulations. Il est utile pour calibrer le modèle, et explorer son comportement en réponse à des variations systématiques sur un ou plusieurs paramètres, pour ensuite faire des analyses de sensibilité.

La partie graphique de la bibliothèque SWARM offre des possibilités limitées, et le parti pris privilégié pour la visualisation des sorties est de s'appuyer sur des outils existants à partir d'un format d'échange. Cela restreint les visualisations interactives mais permet des formes de visualisation très souples à un coût de développement moindre. Le modèle enregistre les données produites dans des bases au format DBF, qui sont ensuite lues sous SIG, mais aussi dans des fichiers textes, qui sont ensuite exploités par GNUPLOT[5].

4.2. *Neutralité de la plate-forme développée*

Comme nous l'avons mentionné dans les sections précédentes, notre modèle intègre plusieurs mécanismes indépendants et interchangeables de type boîte noire : fonction de calcul de l'accessibilité, fonction de localisation, fonction de prix, fonction objectif pour le calibrage, etc. Les ontologies choisies, les boîtes noires que l'on vient brancher dans la plate-forme sont suffisamment générales pour qu'elles puissent représenter toute une gamme de règles potentielles.

4. www.swarm.org/
5. http ://www.gnuplot.info/

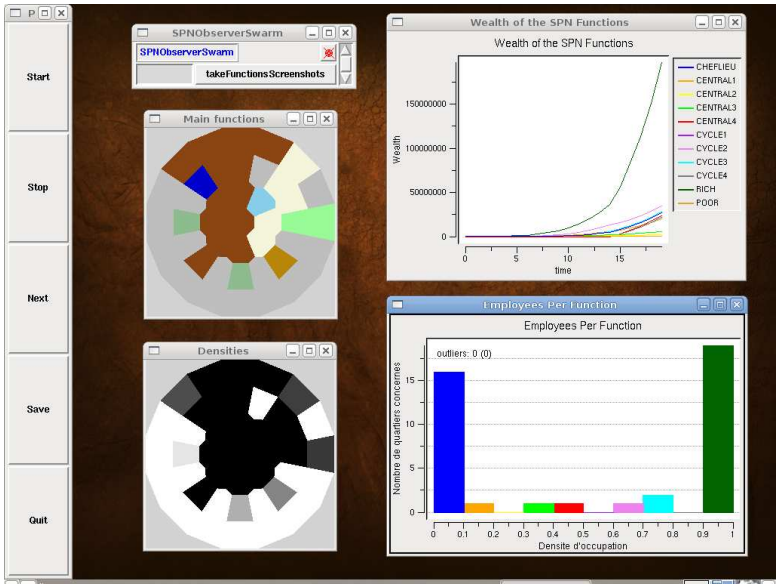


Figure 5. *Exécution en mode graphique du modèle sur une géométrie-support de type européenne. Le modèle permet déjà de reproduire des processus de croissance accrétive, au seul moyen d'une fonction d'accessibilité dynamique, et sans avoir besoin de règle exogène qui force ce phénomène.*

Dans un premier temps, nous avons choisi d'implémenter les règles des différents sous-modèles sous la forme présentée dans les équations de la section 3. Mais la plate-forme informatique développée est neutre et on peut changer, raffiner, complexifier ces différents mécanismes pour un coût en développement supplémentaire quasi nul. Chacune de ces règles est donc interchangeable avec une règle de même type mais différente. Pour certains sous-modèles (fonction d'accessibilité et fonction de scores attribués par les fonctions aux quartiers), la plate-forme permet déjà de basculer entre plusieurs instances.

L'expérimentateur peut de plus faire tourner le modèle sur n'importe quelle géométrie de ville s'il fournit les fichiers au format (DBF) défini par la plate-forme. De la même façon, il peut simuler le nombre de fonctions qu'il souhaite, chacune avec un profil (effectif et budget) de son choix, s'il fournit en entrée les fichiers au format *ad hoc*.

Enfin, comme l'expérimentateur peut choisir parmi plusieurs fonctions objectif proposées, à court terme le modèle pourra être calibré automatiquement pour optimiser sa réponse pour des fonctions objectif diverses.

5. Conclusion

A travers l'exemple de la conception d'un modèle de développement d'une ville sur le long terme, nous avons discuté de classes de problèmes fréquemment rencontrés en SOA en géographie. Nous avons proposé des solutions pour le cas particulier de SIMPOPNANO, que nous avons comparées à d'autres réalisations issues de la communauté. En particulier, nous avons montré qu'une reformulation originale de modèles classiques de localisation pour des *fonctions urbaines* et des *quartiers* permet de dessiner un cadre cohérent pour comparer villes européennes et états-uniennes sur une période historique. La phase de développement logiciel vient de se terminer, l'expérimentation débute et les résultats obtenus feront l'objet d'une publication future. Enfin, comme nous l'avons esquissé dans la section 2, la conception a été guidée par un couplage futur avec SIMPOP2, modèle de SOA réalisé pour comparer la croissance des systèmes de villes en Europe et aux Etats-Unis[6]. L'articulation de ces deux modèles pose d'ambitieuses questions relatives aux ontologies géographiques (Pumain *et al.*, 2009a). Celles-ci devraient permettre de formaliser, quantifier et tester des hypothèses sur la nature de l'articulation de niveaux d'organisation spatiale le plus souvent dissociés dans l'analyse : le quartier, la ville et le système de villes.

Remerciements

Je remercie tout particulièrement Denise Pumain et Anne Bretagnolle du laboratoire Géographie-Cités, Jean Laterrasse du LVMT, ainsi que Jean-Louis Giavitto et Guillaume Hutzler du laboratoire IBISC pour les riches discussions que j'ai pu avoir avec eux. Beaucoup d'idées qui sous-tendent ce travail sont aussi les leurs.

Un merci supplémentaire à Jean-Louis Giavitto, Guillaume Hutzler, Denise Pumain, Lena Sanders et Clara Schmitt qui m'ont accordé une aide précieuse en acceptant de faire une relecture active de cet article. Je remercie également les deux relecteurs anonymes de la *Revue de géomatique* dont les remarques ont contribué à l'amélioration du contenu, en particulier de la section 2 pour le premier relecteur et de la section 3 pour le second.
Merci enfin au Réseau de recherches sur le développement soutenable (R2DS[7]) de la région Ile-de-France, qui finance ce travail.

6. Bibliographie

Abrami G., Niveaux d'organisation dans la modélisation multi-agent pour la gestion de ressources renouvelables. Application à la mise en oeuvre de règles collectives de gestion de l'eau agricole dans la basse-vallée de la Drôme, PhD thesis, ENGREF, 2004.

Alonso W., *Location and Land Use*, Harvard University Press, 1964.

6. http ://www.simpop.parisgeo.cnrs.fr/theInstantiatedModels/introduction.php/
7. http ://www.r2ds.centre-cired.fr/

Amblard F., Rouchier J., Bommel P., « Evaluation et validation de modèles multi-agents », *in* F. Amblard, D. Phan (eds), *Modélisation et simulation multi-agents. Applications pour les Sciences de l'Homme et de la Société*, Hermès, chapter 4, p. 103-140, 2006.

Badariotti D., Banos A., Moreno D., « Conception d'un automate cellulaire non stationnaire à base de graphe pour modéliser la structure spatiale urbaine : le modèle Remus », *Cybergeo*, octobre, 2007.

Banos A., Thevenin T., « Création de champs de potentiel et simulation d'itinéraires à partir de l'enquête ménages-déplacements », *in* M. Thériault, F. Des Rosiers (eds), *Information géographique et dynamiques urbaines 1 : analyse et simulation de la mobilité des personnes*, Traité IGAT, Hermès, chapter 4, 2008.

Batty M., « Fifty Years of Urban Modeling : Macro-Statics to Micro-Dynamics », *in* S. Albeverio, D. Andrey, P. Giordano, A. Vancheri (eds), *The Dynamics of Complex Urban Systems : An Interdisciplinary Approach*, Springer, chapter 1, p. 1-21, 2008.

Benenson I., Torrens P., *Geosimulation : Automata-Based Modeling of Urban Phenomena*, Wiley, 2004.

Bertaud A., « The spatial organization of cities : deliberate outcome or unforeseen consequence ? », , available online at http ://alain-bertaud.com/, May, 2004.

Bertaud A., Malpezzi S., « The Spatial Distribution of Population in 48 World Cities : Implications for Economies in Transition », , available online at http ://alain-bertaud.com/, December, 2003.

Bretagnolle A., Daudé E., Pumain D., « From theory to modelling : urban systems as complex systems », *Cybergeo*, 2006. presented at the 13th European Colloquium on Theoretical and Quantitative Geography, Lucca (Italy), 5-9 September 2003.

Bretagnolle A., Giraud T., Mathian H., « La mesure de l'urbanisation aux Etats-Unis, des premiers comptoirs coloniaux aux Metropolitan Areas (1790-2000) », *Cybergeo*, juillet, 2008.

Bretagnolle A., Pumain D., « Simulating urban networks through multiscalar space-time dynamics (Europe and United States, 17th -20th centuries) », *Urban Studies*, 2009. to be published (accepted).

Bretagnolle A., Pumain D., Vacchiani-Marcuzzo C., « Les formes de systèmes de villes dans le monde », *in* M. Mattei, D. Pumain (eds), *Données Urbaines (vol.5)*, Economica, 2007.

Bura S., Guérin-Pace F., Mathian H., Pumain D., Sanders L., « Multi-agents system and the dynamics of a settlement system », *Geographical Analysis*, vol. 28, p. 161-178, 1996.

Calvez B., Le calibrage de modèles à base d'agents pour la simulation de systèmes complexes, PhD thesis, Université d'Evry Val d'Essonne, 2007.

Calvez B., Hutzler G., « Automatic tuning of agent-based models using genetic algorithms », *Proceedings of the 6th International Workshop on Multi-Agent Based Simulation (MABS'05)*, 2005.

Caruso G., Peeters D., Cavailhès J., Rounsevell M., « Spatial configurations in a periurban city. A cellular automata-based microeconomic model », *Regional Science and Urban Economics*, vol. 37, p. 542-567, 2007.

Castle C., Crooks A., Principles and concepts of agent-based modelling for developing geospatial simulations, Working Paper n° 106, Centre for Advanced Spatial Analysis, September, 2006.

Cattan N., Pumain D., Rozenblat C., Saint-Julien T., *Le système des villes européennes*, Economica, 1994.

Deissenberg C., « Apprentissage dans les modèles multi-agents », *in* F. Amblard, D. Phan (eds), *Modélisation et simulation multi-agents. Applications pour les Sciences de l'Homme et de la Société*, Hermès, chapter 13, p. 309-330, 2006.

Delage M., Le Néchet F., Louail T., « ACCESSIM : modélisation et simulation d'accessibilité et d'inégalités géographiques dans la ville », *Mappemonde*, décembre, 2008. second prix du concours de géovisualisation du festival de Saint-Dié.

Delisle F., *« Accessibilité et morphologie urbaine en Europe et aux Etats-Unis »*, Master's thesis, Université Paris 1, 2008.

Drogoul A., Vanbergue D., Meurisse T., « Multi-Agent Based Simulation : Where are the Agents ? », *in* J. S. Sichman, F. Bousquet, P. Davidsson (eds), *Multi-Agent-Based Simulation II. Third International Workshop, MABS 2002*, vol. 2581 of *Lecture Notes in Computer Science*, Springer, Bologna, Italy, July 15-16, 2002.

Ghorra-Gobin C., *La ville américaine*, Nathan Université, 1998.

Gutknecht O., Ferber J., « Un méta-modèle organisationnel pour l'analyse, la conception et l'exécution de systèmes multi-agents. », *JFIADSMA'98*, Hermès, p. 267-280, 1998.

Haas P., ali., Housing and Transportation Cost Trade-offs and Burdens of Working Households in 28 Metros (available online at http ://www.cnt.org), Technical report, Center for Neighborhood Technology, 2006.

Laterrasse J., Morellet O., « Mobilités et formes urbaines : difficultés d'analyse, éléments de connaissance et bonnes pratiques », *Energie, mobilité, formes urbaines et organisation de l'espace. Séminaire de l'O.I.P.R n°90.*, 2009.

Lee D., « A requiem for large scale modeling », *Journal of The American Institute of Planners*, vol. 39, n° 3, p. 163-178, 1973.

Mangin D., *La ville franchisée*, Editions de la Villette, 2004.

Marceau D., Ménard A., Moreno N., « Les automates cellulaires appliqués à la simulation des changements d'utilisation du sol : sensibilité à l'échelle spatiale et temporelle », *in* M. Thériault, F. Des Rosiers (eds), *Information géographique et dynamiques urbaines 1 : analyse et simulation de la mobilité des personnes*, Traité IGAT, Hermès, chapter 8, 2008.

Marilleau N., Méthodologie, formalismes et outils de modélisation-simulation pour l'étude des systèmes complexes : application à la mobilité géographique., PhD thesis, Université de Franche-Comté, 2006.

Mathis P. (ed.), *Graphes et réseaux. Modélisation multiniveau*, Hermès, 2003.

Minar N., Burkhart R., Langton C., Askenazi M., The Swarm simulation system : a toolkit for building multi-agent simulation, Technical report, Santa Fe Institute, 1996.

Monzie A., Interfacage algorithmes de calibrage de systèmes multi-agents - simulateur, Technical report, ENSIIE, IBISC, 2009.

Moriconi-Ebrard F., *GEOPOLIS - Pour comparer les villes du monde*, Economica, 1994.

Paulus F., Coévolution dans les systèmes de villes : croissance et spécialisation des aires urbaines françaises de 1950 à 2000, PhD thesis, Université Paris 1, 2004.

Pumain D., « Fonction urbaine (définition) », *in* D. Pumain, T. Paquot, R. Kleinschmager (eds), *Dictionnaire la ville et l'urbain*, Economica, p. 119, 2006a.

Pumain D., « Systèmes de villes et niveaux d'organisation », *in* P. Bourgine, A. Lesne (eds), *Morphogenèse. L'origine des formes.*, Belin, p. 239-263, 2006b.

Pumain D. (ed.), *Hierarchy in natural and social sciences*, vol. 3 of *Methodos Series*, Springer, 2006c.

Pumain D., « The Socio-Spatial Dynamics of Systems of Cities and Innovation Processes : a Multi-Level Model », *in* S. Albeverio, D. Andrey, P. Giordano, A. Vancheri (eds), *The Dynamics of Complex Urban Systems. An Interdisciplinary Approach*, Physica Verlag, Heidelberg, p. 373-389, 2008.

Pumain D., Louail T., « Interaction des ontologies informatique et géographique pour simuler les dynamiques multiscalaires », *Ontologie et dynamique des systèmes complexes, perspectives interdisciplinaires*, Rochebrune, Janvier, 2009a. to be published.

Pumain D., Sanders L., Bretagnolle A., Glisse B., Mathian H., « The Future of Urban Systems », *in* D. Lane, D. Pumain, S. Van der Leeuw, G. West (eds), *Complexity perspectives on innovation and social change*, Methodos Series, Springer, chapter 12, 2009b.

Quijano J. G., Modèles d'auto-organisation pour l'émergence de formes urbaines à partir de comportements individuels à Bogota, PhD thesis, Université Paris 6, 2007.

Quijano J. G., Piron M., Drogoul A., « Vers une simulation multi-agent de groupes d'individus pour modéliser les mobilités résidentielles intra-urbaines », *Revue internationale de géomatique*, vol. 17, p. 161-181, 2007.

Sanders L., « Objets géographiques et simulation agent, entre thématique et méthodologie », *Revue internationale de géomatique*, vol. 17, p. 135-160, 2007.

Sanders L., « Géographie quantitative et analyse spatiale : quelles formes de scientificités ? », *in* T. Martin (ed.), *La scientificité des Sciences Humaines et Sociales*, Philosophie des sciences, Vuibert, to be published.

Sanders L., Favaro J.-M., Mathian H., Pumain D., Glisse B., « Intelligence artificielle et agents collectifs : le modèle EUROSIM », *Cybergeo*, 2005.

Treuil J.-P., Drogoul A., Zucker J.-D., *Modélisation et simulation à base d'agents*, Dunod, 2008.

Vanbergue D., Conception de simulation multi-agents : application à la simulation des migrations intra-urbaines de la ville de Bogota, PhD thesis, Université Paris 6, 2003.

White R., Engelen G., « Cellular Automata and Fractal Urban Form : A Cellular Modeling Approach to the Evolution of Urban Land Use Patterns », *Environment and Planning A*, vol. 25, p. 1175-1199, 1993.

7. Annexes

7.1. *Algorithme du paiement par les fonctions de leurs localisations*

L'ordre de considération des fonctions est sans importance. Pour chaque fonction les quartiers sont considérés du plus attractif au moins attractif aux yeux de la fonction. Le nombre d'employés qui restera localisé dans un quartier à la prochaine itération (e_{FinLoc}) correspond au minimum entre le nombre de ceux qui y sont localisés (e_{AlrLoc}) à l'itération courante, et le nombre maximal d'employés que son budget permettrait à la fonction de localiser dans ce quartier au vu de son loyer ($e_{MaxLoc} = W_i^t / P_j^t$).

Algorithm 1 PAY_LOCALISATIONS

$toPay \leftarrow 0$
$e_{AlrLoc} \leftarrow 0$
$e_{MaxLoc} \leftarrow 0$
$e_{FinLoc} \leftarrow 0$
for $i = 0$ to M **do**
  Trier Q par ordre décroissant d'attractivité pour f_i
  for $j = 0$ to N **do**
    $e_{AlrLoc} \leftarrow E^t_{i,j}$
    $e_{MaxLoc} \leftarrow W^t_i / P^t_j$
    $e_{FinLoc} \leftarrow \min(e_{AlrLoc}, e_{MaxLoc})$
    $E^t_{i,j} = e_{FinLoc}$
    $W^t_i = W^t_i - P^t_j \cdot e_{FinLoc}$
    $L^t_i = L^t_i + (e_{AlrLoc} - e_{FinLoc})$
    $C^t_j = C^t_j + (e_{AlrLoc} - e_{FinLoc})$
  end for
end for

7.2. *Algorithme de répartition des fonctions urbaines dans les quartiers*

Algorithm 2 LOCALISE

Ordonner l'ensemble F des fonctions par ordre décroissant des budgets W.
for $i = 0$ to M **do**
  while ($L^t_i > 0$) and ($W^t_i > 0$) **do**
    $j \leftarrow prefLocalisations^t_i.next()$
    $e_{MaxLoc} \leftarrow W^t_i / P^t_j$
    $e_{FinLoc} \leftarrow \min(C^t_j, e_{MaxLoc})$
    $E^t_{i,j} = E^t_{i,j} + e_{FinLoc}$
    $W^t_i = W^t_i - P^t_j \cdot e_{FinLoc}$
    $L^t_i = L^t_i - e_{FinLoc}$
    $C^t_j = C^t_j - e_{FinLoc})$
  end while
end for

Les fonctions sont considérées par ordre décroissant de leur budget, et pour chaque fonction les quartiers sont considérés du plus attractif au moins attractif aux yeux de la fonction considérée. Les fonctions sont « aveugles », elles optimisent l'endroit placé et non pas la population placée. Cela signifie qu'une fonction localise autant d'employés dans un quartier que le lui permettent son budget W^t_i, le loyer du quartier P^t_j, ainsi que le nombre d'emplacements libres qu'il contient C^t_j.