

# Introduction générale

Version 2019-12-04

- 24/11/2019 : début réflexion plan
- 04/12/2019 : version 0 - mise sur Google Docs pour Lena

« Modéliser, [...] c'est rechercher quelle composition de modèles rend le mieux compte d'une organisation régionale ou locale, d'une configuration de champ ou de réseau, ou d'une distribution spatiale. [...] Ces modèles ayant un sens défini et connu, représentant un processus, une réponse sociale à des problèmes spatiaux, ils éclairent ce qui est ou ce qui a été en jeu. Tel est le pari central. C'est un pari heuristique. La modélisation est une procédure de recherche. Comme telle, elle appelle des règles de rigueur et de bon usage. La modélisation en géographie commence par deux questions inévitables et associées : où sommes-nous ? qu'est-ce qui a des chances d'être en jeu ? »

Roger BRUNET (2000, p. 28)

« Si un "objet spatial", (une organisation, une entité), est observable c'est forcément qu'il est apparu et qu'il s'est maintenu suffisamment longtemps pour être considéré comme étudiable par les géographes. En conséquence, expliquer cet objet implique qu'on montre à la fois comment il est né et comment il a persisté. On ne peut donc pas se passer du temps pour faire de la géographie. Mais le temps dont les géographes doivent se servir est complexe. Ils ont à se préoccuper non du passé dans son intégralité, mais des traces de ce passé sensibles dans le monde actuel, donc de "mémoires" ».

François DURAND-DASTÈS (1999, p. 23)

« Principe 1 : Modéliser, c'est apprendre ».

Arnaud BANOS (2013, p. 77)

Le « thème » de cette thèse n'en est pas un, au sens conventionnel. Il porte sur un enjeu méthodologique : comment modéliser des dynamiques spatio-temporelles sur le temps long ? Les trois citations mises en exergue permettent de décomposer ce questionnement très générique. Roger BRUNET

aide à définir ce qu'est la modélisation en géographie. François DURAND-DASTÈS justifie la nécessité de prendre en compte la dimension temporelle, longue, d'un objet spatial contemporain pour être en mesure de l'« expliquer ». Arnaud BANOS, plus synthétiquement, nous rappelle l'un des premiers objectifs de toute procédure de modélisation en sciences humaines et sociales (SHS). Ces auteurs, figures majeures de la modélisation en géographie, nous aident à retrouver le « que », le « quoi » et le « pourquoi » de telles approches. L'ambition de ce travail de thèse est de proposer et d'explicitier un nouveau « comment ».

Cette thèse décrit, illustre et met en œuvre une approche de modélisation, vue comme « procédure de recherche » (BRUNET 2000) fondée sur la co-construction interdisciplinaire de modèles de simulation. Tout au long du manuscrit, nous proposons le recours à des méthodes accessibles aux chercheurs en SHS, basées notamment sur des formalismes partagés par ces disciplines : (1) la simulation à base d'agents, dans laquelle l'explicitation des composants et interactions d'un modèle est réalisée de manière anthropomorphique, et donc compréhensible par chacun, et (2) un usage généralisé de la représentation graphique de données, en particulier autour de la proposition d'une méthode d'évaluation des sorties de modèles, l'« évaluation visuelle ».

Le cas d'étude sur lequel la démarche proposée prend appui concerne la modélisation des transformations du système de peuplement de l'Europe du Nord-Ouest autour de l'an mil. Lors de cette période, entre 800 et 1100, l'habitat, initialement majoritairement rural et dispersé, est progressivement polarisé par les églises et châteaux, formant un système de peuplement hiérarchisé, stable, et territorialisé par le maillage paroissial. Ce cas d'étude a profondément orienté la démarche proposée dans ce travail de thèse, impliquant notamment un travail interdisciplinaire pour rendre compte de ces processus passés et qui se sont déroulés sur le temps long. Dans la suite de cette introduction, nous présenterons les principaux enjeux de la thèse à partir des contraintes que posent un tel cas d'étude.

### **Étudier les systèmes de peuplement sur le temps long**

Dans la citation mise en exergue, DURAND-DASTÈS (1999) rappelle que pour comprendre un « objet spatial », il est nécessaire d'étudier « comment il est né et comment il a persisté ». Étudier les systèmes de peuplements, c'est ainsi comprendre les raisons, générales et particulières, de leur état actuel. L'organisation structurelle des systèmes dépend en effet largement des successions de transformations qui les ont façonné tels qu'on les connaît aujourd'hui. Pour les systèmes de peuplement anciens, dont le système de villes du Nord-Ouest européen fait partie, ces successions de transformations ne peuvent être conçues que dans le temps long. C'est en effet dans cette temporalité que s'expriment les transformations sociales successives qui façonnent et transforment l'espace plus ou moins progressivement. Pour pouvoir dé-

crire et surtout comprendre un système de peuplement ancien dans son présent, il est donc nécessaire de remonter dans le passé, et de comprendre les conditions de son émergence.

DURAND-DASTÈS (1999) ajoute un élément important quant à la manière de mobiliser ce temps, long dans le cas des systèmes de peuplement : « [Les géographes] ont à se préoccuper non du passé dans son intégralité, mais des traces de ce passé sensibles dans le monde actuel, donc de “mémoires” ». Le géographe s'appuie en effet sur des « traces » persistantes pour étudier les systèmes. Pour comprendre la genèse et les transformations d'un système, on ne peut faire l'économie de mobiliser toutes les sortes de traces, c'est-à-dire de sources, permettant d'en éclairer les états passés. Les différents types de sources, dans le paysage disciplinaire français, sont justement ce qui différencie deux disciplines dont la compréhension de phénomènes passés est l'objet de recherche principal : l'histoire, basée sur l'étude des sources écrites, et l'archéologie, sur l'étude des sources matérielles.

### **Une étude interdisciplinaire**

Plus un système est ancien, plus les sources sont rares, incomplètes et lacunaires. Pour pouvoir étudier un tel système, il est alors nécessaire de compenser la rareté de chaque source par leur démultiplication, en faisant appel à une large diversité de sources et de types de sources. Cela implique *a minima* la mise en place de groupes de recherche pluridisciplinaires, composés d'archéologues, d'historiens et de géographes, chacun apportant une expertise dans un type de source ou dans la compréhension des processus qui ont mené aux transformations du système de peuplement. Cette juxtaposition d'expertises permet la multiplication des points de vue, chacun analysant avec sa culture disciplinaire les éléments directement liés du système. Au delà de l'expertise propre de chaque chercheur, c'est aussi la multiplication et la diversité des points de vue qui permet d'embrasser une vision plus globale des processus étudiés.

Cette vision globale ne pourra toutefois être détachée des *a priori* disciplinaires tant que l'analyse ne revêtira pas une dimension interdisciplinaire, et pas simplement pluridisciplinaire. C'est en effet en mettant les objets étudiés, et leur analyse, en commun, chacun s'emparant de l'ensemble des éléments, que l'on peut aboutir à une vision renouvelée, originale, de la description ou de l'explication d'un processus spatio-temporel sur le temps long. En adoptant une dimension interdisciplinaire, il est ainsi possible de dépasser les points de vue et habitudes disciplinaires classiques, en faveur d'une analyse collective, basée sur le consensus, proposant une approche plus complète et générique.

### **Modélisation**

Cette analyse collective et commune peut être formalisée, c'est-à-dire que

pour décrire les états et transformations successifs d'un système, on peut les représenter sous la forme d'un modèle. MINSKY (1965, §2) en donne cette définition très générale : « To an observer B, an object A\* is a model of an object A to the extent that B can use A\* to answer questions that interest him about A. ». La formalisation d'un modèle, quel qu'en soit le paradigme mobilisé, permet donc de répondre à des questions, c'est-à-dire, dans notre cas d'étude, de chercher les explications ou l'aspect générique de transformations dans le système de peuplement.

En matière d'interdisciplinarité, le recours à un modèle offre un autre avantage, en ce que le recours à un formalisme oblige à expliciter le sens exact, ontologique, des entités qu'on y inclue et des éventuelles relations qui les mettent en interaction. En effet, pour que plusieurs chercheurs puissent répondre à la même question sur un système (A) à partir d'un unique modèle (A\*), il est indispensable qu'ils s'accordent sur la définition précise du contenu de A\*, et donc qu'ils forment un unique « observateur » cohérent B<sup>1</sup>. Pour que les différents chercheurs, de différentes disciplines et caractérisés par différentes approches, puissent parvenir à une représentation – un modèle – commune, il est ainsi nécessaire qu'ils s'accordent sur sa définition<sup>2</sup>. Il s'agit de choisir les objets à y inclure, de circonscrire la correspondance dans le domaine empirique de ces objets, et d'harmoniser le sens précis de ce que sont les relations : « qu'est-ce qui a des chances d'être en jeu ? » dans la citation de BRUNET (2000).

L'exercice de la modélisation – la création d'un modèle – est d'autant plus intéressant qu'il favorise une forte dimension heuristique. Le premier principe de BANOS (2013), « modéliser, c'est apprendre », deviendra même le titre de la publication correspondant à cette habilitation à diriger les recherches (BANOS 2016). En effet, quelle que soit la forme choisie<sup>3</sup>, la synthèse des connaissances sous la forme d'un modèle force le chercheur à conceptualiser l'objet de sa recherche et à en repenser le fonctionnement. Pour BRUNET (2000, p.28, §F), la modélisation « construit, déconstruit et reconstruit ; elle passe par une série d'itérations entre déduction et induction ». C'est au moyen de ces itérations que le modélisateur consolide petit-à-petit ses connaissances d'un processus et devient capable de les exprimer sous la forme d'un modèle descriptif et explicatif. En contexte interdisciplinaire, l'effet heuristique de la modélisation est renforcé par la confrontation des connaissances hétérogènes et des points de vue de chacun des membres impliqués dans la modélisation. Le modèle prendra en effet appui sur les connaissances, disparates, de chacun. Dès lors, la modélisation,

---

1. Pour MINSKY (1965, §2), on ne peut en effet utiliser A\* si B n'est pas clairement identifié et compris : « Any attempt to suppress the role of the intentions of the investigator B leads to circular definitions or to ambiguities about "essential features" and the like. »

2. BRUNET (2000, p. 22) explique ainsi que « toute définition d'un objet propose un modèle ».

3. BRUNET (2000, p.27, §D) en identifie trois : « rhétorique », « mathématique » et « iconique » (qu'il préfère au terme graphique, bien plus usité).

en tant que processus de recherche, aboutit nécessairement à un enrichissement des connaissances et représentations de chacun grâce à celles des autres.

### **Simulation**

Un modèle, comme toute construction scientifique, a vocation à être éprouvé. Pour les modèles descriptifs portant sur des phénomènes sociaux ou spatiaux contemporains, une telle mise à l'épreuve peut, dans certaines mesures, être (re-)produite de manière empirique, via la mise en place d'une expérience, qu'elle soit effectuée « en laboratoire » ou par une observation extérieure. Dans notre cas, c'est-à-dire celui de la modélisation d'un processus spatial ancien et sur le long terme, il est impossible, par nature, de mener une telle expérience *in vivo*. Pour y pallier, on peut mener une expérimentation *in silico*, au moyen du « laboratoire virtuel » que constitue la simulation informatique.

En formalisant le modèle exprimé préalablement de manière computationnelle (algorithmique ou mathématique), il devient en effet possible d'en simuler le déroulement et l'aboutissement, et donc d'évaluer le modèle, c'est-à-dire de confronter ses sorties avec les connaissances empiriques du phénomène modélisé. La simulation permet donc de tester un modèle, mais aussi de passer d'une description statique à une exploration des dynamiques : contrairement à un modèle rhétorique ou « iconique », la simulation comporte une dimension temporelle et peut être observée pendant son déroulement. On peut donc évaluer un état final, mais aussi la succession d'instantanés formés par les états intermédiaires du système. Un modèle conceptuel permet de détailler des règles de fonctionnement d'un système, lesquelles peuvent mener à une transformation de ce système. Avec un modèle de simulation, on peut directement observer, et donc analyser, ces transformations. Il est alors possible d'évaluer la capacité d'un modèle à reproduire un état connu, mais aussi de vérifier la plausibilité des transformations pendant qu'elles se produisent.

### **Le choix du paradigme de la simulation multi-agents**

La conception d'un modèle rhétorique ou iconique est accessible à chacun, les formalismes employés – l'écrit et le visuel – ayant une part centrale dans nos sociétés. Au contraire, la construction d'un modèle mathématique ou informatique requiert une culture méthodologique spécifique. Cela passe notamment par l'apprentissage de différents « langages » et paradigmes, qui peuvent être mathématiques (systèmes d'équations différentielles, formalismes de la théorie des jeux, logique mathématique, etc.) ou informatiques (programmation orientée objet, programmation fonctionnelle, spécificité de chaque langage informatique, etc.). La construction de modèles de simulation demande donc une certaine maîtrise d'un langage informatique, et à ce titre, n'est pas véritablement accessible à chacun.

Pour mener un projet de simulation, il est donc nécessaire qu'au moins l'un des participants ait les compétences techniques requises, sans oublier les connaissances méthodologiques liées à la manière de mener le processus de modélisation. La maîtrise de ces éléments techniques et méthodologiques est peu fréquente chez les chercheurs en SHS, et *a fortiori* dans les humanités, parmi lesquelles l'histoire et l'archéologie. Vis-à-vis de ces disciplines, le paradigme de la modélisation à base d'agents, aussi dénommé « simulation multi-agents » (SMA) nous semble porteur et adapté. Ce courant de la simulation informatique représente les entités modélisées par des agents informatiques, qui interagissent les uns avec les autres au moyen de « règles de comportement » que l'on nomme des mécanismes. Chaque mécanisme est précisément défini par les modélisateurs, mais la somme des interactions qui en résulte ne peut être déterminée en amont.

C'est un type de simulation anthropomorphique, dans le sens où les agents et leurs mécanismes peuvent être compris comme des acteurs dotés de comportements propres. Pour des chercheurs en SHS, le parallèle avec les comportements individuels, rationnels ou non, des acteurs des sociétés passées (paysans, seigneurs, etc.), ou des agrégats d'acteurs (villages, communautés, etc.) permet d'explicitier le fonctionnement d'un modèle plus aisément que le formalisme mathématique par exemple.

### **Une modélisation en co-construction**

Si le paradigme des SMA est plus accessible aux chercheurs en SHS que certains autres, il demeure un formalisme informatique où l'on exprime les entités sous formes d'objets informatiques dotées de règles, c'est-à-dire de fonctions. Pour être certain que les chercheurs impliqués, de différentes disciplines et de niveaux de connaissances informatiques variés, aient la même compréhension du modèle – préalable nécessaire à son utilisation (voir plus haut) – il est nécessaire de mettre en place une démarche collective. Il faut ainsi veiller à ce que chaque participant au projet puisse s'approprier le modèle résultant, et pour cela, nous avons fait le choix de la « co-construction », c'est-à-dire d'une expérience collective où chacun participe à l'élaboration de l'ensemble des composants du modèle, depuis les aspects les plus conceptuels (choix des entités) jusqu'aux plus détaillés (ordre d'appel des mécanismes).

Ce faisant, chacun participe à l'ensemble des réflexions et décisions liées au modèle, dans une forme de compromis qui permet de dépasser les *a priori* de chaque discipline et position tenue dans le collectif. Cela renforce aussi la dimension heuristique du modèle, puisque chacun doit exprimer dans un formalisme commun et partagé, en explicitant de manière totale, ce qu'il souhaite ajouter au modèle : un programme informatique n'interprète pas mais exécute, et ne laisse ainsi pas de part au flou. Dans une perspective de co-construction, quand bien même chaque participant ne sera pas nécessairement amené à écrire les lignes de code du modèle, le formalisme

informatique forcera toutefois chacun à expliciter le détail de l'implémentation.

### **Des outils pour faciliter la co-construction**

Pour guider le développement et l'amélioration du modèle, il est utile de pouvoir en analyser le comportement, c'est-à-dire de procéder à son évaluation : correspond-il entièrement aux attentes des modélisateurs, ou faut-il en modifier des composantes ? L'évaluation de modèles est un domaine de recherche à part entière, souvent fondé sur des analyses formelles, statistiques ou au moins quantifiées des sorties du modèle. Évaluer un modèle selon ces méthodes requiert d'une part une certaine expertise thématique de ce qui est modélisé, mais surtout, une bonne connaissance des méthodes d'analyse quantitative des données. Comme la connaissance de la modélisation mathématique ou informatique, ces méthodes quantitatives ne font pas forcément parti de l'outillage méthodologique classique en SHS, et notamment en histoire et en archéologie.

Pour faciliter l'évaluation du modèle, nous proposons dans cette thèse une approche basée sur l'analyse graphique des sorties du modèle. Cette approche, intitulée « évaluation visuelle », doit permettre à chacun d'évaluer, de manière qualitative et experte, les sorties massives du modèle. L'évaluation visuelle consiste en l'analyse visuelle d'un ensemble d'indicateurs de sorties, représentations graphiques des différentes dynamiques du modèle de simulation, et en leur critique au regard d'une grille d'analyse explicitée *a priori*.

Une plateforme interactive a été développée afin de rendre cette évaluation plus accessible, et permet aux membres du projet de visualiser d'eux-mêmes les sorties du modèle sans requérir de compétences spécifiques en analyse et en visualisation de données. Cette plateforme donne à chacun la capacité d'explorer les sorties du modèle. L'exploration de ces sorties permet aux chercheurs impliqués de comprendre les dynamiques et structures sociales et spatiales qui émergent des interactions entre les agents du modèles, et dès lors, de tester les hypothèses thématiques introduites dans le modèle. En cela, l'évaluation visuelle – via la plateforme d'exploration – favorise une démarche abductive, guidée par les données, et augmente en cela la dimension heuristique du processus de modélisation.

### **Organisation du manuscrit**

Présenter un modèle complexe, sa démarche de construction et son évaluation, dans le cadre d'une thèse, donc de manière linéaire, revêt forcément une certaine artificialité dans l'ordre de présentation ainsi qu'une forte interdépendance entre les chapitres. Il s'agit ainsi de rendre linéaire et chronologique un processus profondément « en spirale » fait d'itérations et d'allers-retours. De même, la présentation d'un unique état de l'art initial, pour une

thèse dont les inscriptions thématiques et méthodologiques sont multiples, ne semble pas l'option la plus adaptée. Ce manuscrit prend donc une forme, sur certains aspects, peu classique, avec notamment de nombreux renvois entre chapitres (précédents et suivants), ainsi que des états de l'art propres à chaque chapitre, quand nécessaires.

Le premier chapitre (**chapitre 1**) de ce manuscrit prolonge cette introduction en introduisant et en précisant le positionnement adopté lors de ce travail, explicitant ainsi les choix qui ont amené à la modélisation sous forme de simulation multi-agents, à un usage important de la représentation graphique ainsi qu'à l'approche de co-construction. Le **chapitre 2** présente le modèle issu de cette approche, « SimFeodal », et en détaille le cadre de conception et d'application selon le protocole de présentation de modèles à base d'agents « ODD » (GRIMM et al. 2010). Le **chapitre 3** justifie et explicite le choix de l'« évaluation visuelle » comme méthode d'évaluation et d'exploration du modèle, en la replaçant dans le large cadre de l'évaluation de la modèle. Cette évaluation visuelle est ensuite mobilisée pour présenter l'évolution du modèle et le travail de « paramétrage » qui l'a guidée. Dans le **chapitre 4**, nous présentons la succession de besoins, de contraintes et de choix – thématiques, méthodologiques et techniques – qui ont menés à la création de la plateforme interactive d'exploration des sorties du modèle, « SimEDB ». Nous présentons ensuite de manière détaillée les parti-pris de sa conception en matière de stockage de données (bases de données relationnelles analytiques) et d'interactivité. En mobilisant cette plateforme, nous présentons dans le **chapitre 5** les résultats du modèle, c'est-à-dire les sorties de la version la plus récente de SimFeodal et leur réponse aux objectifs exprimés précédemment. Pour prolonger l'exploration du modèle, ce chapitre est aussi l'occasion de mener une analyse de sensibilité exploratoire du modèle, encore une fois basée sur la représentation graphique. Le **chapitre 6** introduit la conclusion générale en présentant un retour réflexif, à partir de l'expérience relatée tout au long de cette thèse, sur la manipulation de données de simulation et sur l'effectivité de la démarche de co-construction mise en œuvre.



## Références

- BANOS, Arnaud (2013). « Pour des pratiques de modélisation et de simulation libérées en Géographie et SHS ». Habilitation à Diriger des Recherches. Paris : Université Paris 1 - Panthéon-Sorbonne. 107 p. URL : <https://halshs.archives-ouvertes.fr/tel-01112668/document>.
- (2016). *Modéliser, c'est apprendre : Itinéraire d'un géographe*. Avec la coll. de Michel BUSSI. Modélisations, simulations, systèmes complexes. Paris : Matériologiques. 104 p. URL : <https://www.decitre.fr/livres/modeliser-c-est-apprendre-9782373610802.html>.
- BRUNET, Roger (2000). « Des Modèles En Géographie ? Sens d'une Recherche ». In : *Bulletin de la Société de Géographie de Liège* 2, p. 21-30.
- DURAND-DASTÈS, François (1999). « Jamais deux fois... Ou : de quelques précautions à prendre avec le temps ». In : *Travaux de l'Institut de Géographie de Reims* 26.101, p. 5-23. DOI : 10.3406/tigr.1999.1383. URL : [https://www.persee.fr/doc/tigr\\_0048-7163\\_1999\\_num\\_26\\_101\\_1383](https://www.persee.fr/doc/tigr_0048-7163_1999_num_26_101_1383).
- GRIMM, Volker et al. (2010). « The ODD Protocol : A Review and First Update ». In : *Ecological Modelling* 221.23, p. 2760-2768. DOI : 10.1016/j.ecolmodel.2010.08.019. URL : <http://www.sciencedirect.com/science/article/pii/S030438001000414X>.
- MINSKY, Marvin (1965). « Matter, Mind and Models ». In : *Proc. International Federation of Information Processing Congress* 1, p. 45-49. URL : <https://web.media.mit.edu/~minsky/papers/MatterMindModels.html>.