

Modélisation et visualisation à l'interface entre les disciplines

Version 2019-12-16

- 10/10/2019 : Nouveau plan
- 22/10/2019 : fin écriture + relecture + rendu Lena
- 18/11/2019 : fin reprise + ajouts (reproductibilité + intro + ccl)
- 06/12/2019 : fin reprise + envoi relecture externe Julien
- 16/12/2019 : reprise commentaires Julien¹ + envoi Clarisse

Sommaire

Introduction	2
1.1 Parcours scientifique et professionnel : vers la modélisation en géographie	3
1.1.1 Géographie	3
1.1.2 Géomatique	4
1.1.3 Modélisation : à la confluence de la GTQ et de la géomatique	6
1.2 Modéliser l'espace sur le temps long : un travail d'interdisciplinarité	10
1.2.1 Modélisation de processus spatiaux	10
1.2.2 Des processus inscrits dans la longue durée	12
1.2.3 Un contexte fortement interdisciplinaire	14
1.3 Conditions et modalités de la modélisation collective et interdisciplinaire	19
1.3.1 Position du modélisateur : du guide au co-constructeur	19
1.3.2 Explorer plutôt que confronter des données empiriques et simulées	22
1.3.3 D'un méta-modèle à un retour sur expérience de modélisation	25
1.4 Visualiser et explorer pour co-construire un modèle : conception d'interfaces exploratoires	26
1.4.1 Favoriser une co-construction interdisciplinaire	26
1.4.2 Mettre en place des interfaces disciplinaires	30
1.4.3 Une démarche exploratoire	33
1.4.4 Une démarche reproductible	36
Conclusion	40

1. Manque : p. 2, 26, 41

Introduction

Le travail de thèse présenté dans ce manuscrit est hybride, et donc peu évident à positionner dans l'espace des disciplines académiques, tant l'objet décrit (la modélisation de dynamiques spatiales), le cas d'étude (l'évolution du système de peuplement rural en Touraine entre les IX^e et XIII^e siècles) et la méthode mise en place (la simulation multi-agents) semblent relever de disciplines différentes (géographie, sciences historiques et informatique).

Cet aspect pluridisciplinaire complexifie la structure démonstrative classique d'une thèse en géographie, et il me paraît² dès lors inadapté d'engager cette argumentation par l'habituel état de l'art, celui-ci ne pouvant être cantonné à un unique sujet dans ce travail.

Au contraire, je crois que ce premier chapitre offre plutôt l'occasion de clarifier, en l'explicitant, la position de ce travail universitaire, et plus globalement, des recherches menées dans le cadre de cette thèse. Ce manuscrit et les productions informatiques qui l'accompagnent³ ne visent pas à approfondir un questionnement thématique précis, mais s'ancrent à l'inverse dans une position assumée et engagée de démarche méthodologique. C'est dans ce sens que ce manuscrit devrait être compris, en considérant que les solutions mises en œuvre pour l'analyse d'un cas d'étude particulier ont en réalité une vocation de genericité et de reproductibilité. L'apport majeur de ce travail se veut ainsi méthodologique, et c'est à ce titre qu'il s'inscrit à l'interface entre plusieurs disciplines, mobilisant des connaissances et méthodes de chacune.

Dans ce chapitre, je tiens à expliciter le positionnement de ce travail méthodologique, à l'interface entre la géographie, l'informatique et les sciences historiques et des approches méthodologiques plurielles. Pour expliciter cette position et l'engagement collectif qui la définit, il m'apparaît nécessaire d'en décrire la genèse au travers (1) de mon parcours personnel et (2) du contexte de conception de cette thèse. La description (3) du questionnement qui en résulte et de son évolution dans le contexte de ce travail de recherche, permettront enfin (4) de justifier l'approche suivie tout au long de cette expérience, sous une forme programmatique engagée.

2. Dans ce chapitre, très personnel et consacré essentiellement à la description et justification d'un positionnement individuel, le choix de la première personne du singulier me semble tout à fait adapté. Dans le reste de ce manuscrit, qui s'inscrit dans un contexte collectif dont plusieurs dimensions sont elles aussi collectives, la première personne du pluriel sera exclusivement mobilisée.

3. La forme numérique de ces productions, organisée dans des dépôts logiciels qui sont systématiquement communiqués dans cette thèse, me paraît plus adaptée qu'une reproduction dans le manuscrit.

1.1 Parcours scientifique et professionnel : vers la modélisation en géographie

Le travail de recherche présenté dans cette thèse s'inscrit profondément à l'interface entre plusieurs courants disciplinaires liés à l'étude des phénomènes sociaux dans l'espace. Pour en comprendre aussi bien le questionnement que l'approche mobilisée et les résultats obtenus, il me paraît important de faire un rapide retour sur ma formation initiale et début de parcours dans le monde de la recherche académique, qui explique et préfigure assez largement le positionnement adopté dans ce travail. Depuis une formation classique de géographie humaine et urbaine jusqu'à l'exercice de fonctions d'ingénieur d'étude en modélisation, en passant par une spécialisation en géomatique et cartographie, chacune des étapes de ma formation permet de mieux comprendre la formulation de l'aboutissement de cette thèse fondée sur la co-construction interdisciplinaire de modèles spatiaux et sur leur exploration graphique.

1.1.1 Géographie

Ce travail de recherche s'inscrit avant tout, tant administrativement que conceptuellement, dans le champ disciplinaire de la géographie. Cette discipline, consacrée à l'étude de la dimension spatiale de phénomènes sociaux, constitue les fondements de ma formation initiale. En sortant de classes préparatoires littéraires généralistes, j'avais ainsi été frappé par l'exercice du commentaire de cartes, à visée tant verticale (cartes géologiques) qu'horizontale (cartes topographiques). On pouvait décrire et expliquer le fonctionnement économique et social d'un lieu par la seule observation de ses structures et de ses contextes spatiaux.

Géographie urbaine. Cela m'a mené vers un cursus universitaire classique de géographie, majoritairement marqué par la géographie humaine et l'étude de grandes tendances spatiales dans les interactions sociales humaines. Avec un intérêt pour l'aménagement et l'urbanisme, la géographie urbaine, dans sa dimension sociale, m'est rapidement apparue comme particulièrement stimulante dans sa capacité à décrypter, à expliquer et à comparer des processus sociaux variés à l'échelle intra-urbaine. Il ne s'agissait plus simplement de décrire un état, mais d'en expliquer les processus spatiaux et sociaux. Au regard des enseignements d'urbanisme et de politiques de la ville, ces approches permettaient ainsi de comparer le résultat de différentes politiques publiques, et de mener un début de mesure de l'écart entre leur objectif exprimé et leur action effective.

En master, j'ai voulu appliquer ces approches en initiant, sous la co-direction de Renaud Le Goix et Antonine Ribardière, un mémoire sur la comparaison de l'intégration spatiale des migrants entre les politiques francophones et anglophones. Les politiques migratoires francophones, nourries du modèle jacobin français, menaient-elles à une plus forte inclusion et mixité sociale que les modèles anglophones, fondés sur l'image d'un « *salad-bowl* » communautaire ? Le

cas d'étude choisi portait sur le Canada, pays ayant l'avantage de présenter ces deux communautés linguistiques et culturelles – francophones jacobines et anglophones communautaires –, et d'avoir une forte culture – et attractivité – migratoire, aussi bien pour les pays les plus développés que pour les Suds. Le recensement canadien, enfin, permettait les études ethniques et communautaires en offrant des statistiques ethniques rares dans les autres pays.

Géographie Théorique et Quantitative. Dans un contexte de découverte de l'analyse spatiale, de la modélisation graphique, des systèmes d'information géographiques (SIG) et d'approches plus systémiques et horizontales de description et d'explication de phénomènes socio-spatiaux, ce travail de Master s'est assez rapidement orienté vers une démarche quantitative et à visée plus généralisante. Le mémoire qui en a résulté, intitulé « Ségrégation spatiale et origines ethniques dans les métropoles canadiennes », constitue ainsi un tournant vers la géographie théorique et quantitative (GTQ). L'approche, très quantifiée, fait la part belle à la comparaison des différents indices de ségrégation caractérisant les distributions spatiales de la population urbaine canadienne. Pour être en mesure de mener cette comparaison de manière systématique, une auto-formation poussée avait été nécessaire, notamment sur les techniques d'analyse de données, de réduction de dimensionnalité (les recensements canadiens contiennent des centaines de catégories qui ne pouvaient toutes être traitées individuellement) et sur une première approche d'automatisation de traitements (via SAS). Il fallait être en mesure de tester rapidement différentes hypothèses sur les quelques milliers de « secteurs de recensement » impliqués dans l'étude.

Du point de vue méthodologique, cette première expérience d'exploration systématique d'un jeu de données m'avait montré la nécessité de parvenir à une certaine automatisation de la chaîne de traitement, depuis la sélection des données, le calcul de tel ou tel indice, jusqu'à leur représentation (carto)graphique. Les données analysées étaient en effet hétérogènes, multidimensionnelles, et je cherchais à les caractériser par le calcul d'une dizaine d'indices de ségrégation, globaux et locaux. Les quelques macros SAS mises en place ne permettaient alors pas une automatisation totale de cette démarche d'analyse, et les mois d'été passés aux traitements systématiques et répétitifs nécessaires à une approche comparative ne rendaient que plus criant le besoin d'une méthode intégrée et automatisée.

1.1.2 Géomatique

Après ce constat méthodologique, je me suis orienté, pour le master 2, vers une spécialisation en géomatique et cartographie, en intégrant le master professionnel Carthagéo. Les connaissances acquises, méthodologiques et techniques, ont toutes concouru aux démarches mises en place dans ce travail de thèse.

Programmation, automatisation et interfaces graphiques. En premier lieu, Carthagéo m'a permis de découvrir des méthodes d'automatisation de

chaînes de traitement de données spatiales. Avec l'initiation à la programmation, il devenait important d'acquérir une vision algorithmique, systématique et processuelle des analyses que je menais jusque-là de manière manuelle et répétitive. Par exemple, les successions de classifications ascendantes hiérarchiques (CAH) menées sur les données canadiennes, réalisées pour chaque métropole, pour chaque origine ethnique, et à plusieurs niveaux d'agrégation spatiale, pouvaient être automatisés une fois la démarche précisément explicitée. En préalable à l'automatisation d'une démarche, il fallait avant tout pouvoir la formaliser, sur papier d'abord, de manière à pouvoir en réaliser une implémentation informatique dans un second temps. C'est, pour moi, la découverte de la conception de modèles graphiques, non plus dédiés à la description d'un lieu comme dans la chorématique que j'avais découverte les années précédentes, mais à l'explicitation d'un processus d'analyse. Avec l'automatisation permise par l'implémentation de ces chaînes de traitement, il devenait possible de mener des études systématiques, reproductibles et paramétrables.

Dans le cadre d'un projet de programmation en systèmes d'informations géographiques (SIG), j'ai aussi dû réaliser un outil – un *plugin* SIG⁴ –, permettant une comparaison visuelle et mesurée de la qualité de différents services de géocodage. Ce projet, réalisé au sein de l'École Nationale des Sciences Géographiques, institution gérée par l'Institut Géographique National (IGN⁵), visait notamment à comparer le géocodage proposé par une société internationale – Google Maps – et par un service public national – l'IGN et la BD TOPO –. En dehors de l'aspect technique, cette première expérience de projet de programmation appliqué a surtout été l'occasion de réfléchir à des questions d'interface homme-machine. Comment rendre intuitif, pour un utilisateur n'ayant pas participé à sa conception, l'usage d'un outil interactif pensé pour vérifier la cohérence de géocodage de différentes adresses tirées de manière aléatoire dans l'espace francilien? Fallait-il privilégier la présentation de l'indicateur quantitatif – la distance entre les points issus du géocodage – ou plutôt donner une idée plus contextuelle de la localisation spécifique des résultats du géocodage? Cette sensibilité à l'« usabilité »⁶ d'un logiciel a été fondatrice pour mes travaux ultérieurs, et structure largement le présent travail de thèse (voir chapitre 5, section 5.4).

Approches géométriques. Une autre approche mobilisée dans cette thèse est la vision processuelle « géométrique » des traitements de données spatiales. Dans ce type d'approches, caractérisées par le recours aux opérateurs spatiaux, les agrégations, extractions et filtrages de données sont réalisés de manières

4. Les *plugins* sont des extensions logicielles, qui ajoutent des fonctionnalités absentes à des logiciels génériques.

5. L'IGN a depuis été renommé en « Institut National de l'information géographique et forestière ».

6. L'usabilité est définie selon le standard ISO (2018) comme la combinaison de trois propriétés que sont « l'efficacité » d'une interface (précision et qualité des résultats obtenus en l'employant), son « efficience » (coût temporel et concentration que cela demande à l'utilisateur) et la satisfaction éprouvée par l'utilisateur suite à son usage (ZHENG 2019, p. 39).

surtout spatiales plutôt qu’attributaires. Par exemple, plutôt que de définir des voisinages comme la co-appartenance à une même structure territoriale (les communes d’un département par exemple), on aura plutôt tendance à définir ces voisinages à partir de contiguïtés spatiales d’ordre n . De la même manière, pour confondre des maillages différents (bureaux de vote et IRIS par exemple), on peut procéder à une agrégation au niveau administratif supérieur (la commune), mais aussi effectuer une intersection géométrique et re-ventiler les populations dans le maillage résultant.

Il s’agit ainsi de prendre en compte le contexte spatial, topologique, pour réaliser les opérations sur les données. En somme, cela revient à mobiliser la dimension spatiale des données dans les différentes chaînes de traitement mises en œuvre, et la différencier fortement des autres dimensions, attributaires. Ces approches sont nécessaires à la réalisation d’analyses portant sur des données de différentes granularités, de différents maillages : elles permettent en effet d’homogénéiser des informations dont la dimension spatiale est primordiale. En tant que telle, cette vision « géométrique » nous a été fortement recommandée et transmise dans le cadre d’enseignements en analyse spatiale.

Dans le modèle présenté dans cette thèse (**chap2**), une large partie des mécanismes est caractérisé par des processus géométriques, qu’il s’agisse de la mise en place de zones tampons autour d’agrégats ou de pôles, de prises en compte du voisinage – définition des agrégats –, de logiques de distances euclidiennes – satisfaction de protection dépendant de la distance au château –, d’intersections et unions spatiales – héritage des agrégats –, etc. L’influence de mon parcours semble indéniable sur ces choix de modélisation où l’espace est traité de manière continue, démarche assez peu répandue dans les modèles de dynamiques spatiales.

Représentations (carto)graphiques. Un dernier point lié à mon enseignement de master a largement infusé sur les choix de ce travail de thèse. Le master Carthagéo est une formation en grande partie dédiée à la cartographie, c’est-à-dire à l’apprentissage des « règles » de représentation cartographiques et à une certaine réflexivité sur les différents messages qu’une carte peut convoier. La réflexion méthodologique sur les usages de la représentation (carto)graphique pour rendre compte d’un jeu de données est fortement présente dans l’ensemble des projets qui doivent être réalisés au cours de l’année de formation. Les choix de représentation graphiques, très présents dans ce travail de thèse, ont ainsi durablement percolé depuis cet apprentissage.

1.1.3 Modélisation : à la confluence de la GTQ et de la géomatique

Master professionnel oblige, la validation de Carthagéo impliquait la réalisation d’un stage de fin d’étude, en entreprise ou en unité de recherche. J’ai eu la chance d’entamer un stage dans l’UMR Géographie-cités, en mai 2011, sous la co-direction de Thomas Louail, Clara Schmitt et Sébastien Rey-Coyrehourcq.

Ce stage, finalement intitulé « Conception de modèles et d'outils de géosimulation » (CURA 2011), était déjà organisé autour de tâches qui résonnent fortement vis-à-vis du contenu de la thèse.

Il s'agissait, de « prendre part à toutes les étapes de la modélisation » (ibid.), par :

1. l'enrichissement d'un modèle de simulation (SimpopLocal) ;
2. la participation à la conception et à l'implémentation d'un second modèle de simulation (SimpopNet) ;
3. la création d'un outil de production de rapports de simulations (TrajPop) ;
4. la création d'un outil d'exploration cartographique des résultats de simulation (ibid., p. 12-13).

Ce stage de six mois, et les deux années de contrats d'ingénieur d'étude qui l'ont suivi pour en prolonger les recherches (au sein des projets GeoDiverCity, MIRO2 puis TransMonDyn)⁷ ont durablement marqué mon rapport à la modélisation, à l'utilité de ses méthodes et à la manière de construire un modèle de façon collective.

Découverte de la modélisation à base d'agents. Ce stage a marqué ma découverte du domaine de la modélisation, et en particulier de la modélisation à base d'agents. Par coïncidence, le premier modèle que j'ai eu à comprendre et à enrichir était un modèle de simulation de l'émergence et de la hiérarchisation d'un système de peuplement sur le temps long, au néolithique : SimpopLocal (SCHMITT 2014 ; REY-COYREHOURCQ 2015). Il s'agissait de tester des hypothèses issues de la géographie théorique et quantitative en les éprouvant, *in silico*, à l'aide d'un outil informatique permettant de simuler des dynamiques spatiales.

Pour que je parvienne à comprendre et à m'approprier le modèle, mes encadrants m'avaient demandé d'ajouter un mécanisme exogène de perturbation du système simulé, afin de prendre en compte l'effet d'incidences de catastrophes naturelles. L'idée de mon implication était de me permettre de me former, par la pratique, aux notions sous-jacentes de la modélisation de systèmes complexes : émergence, interactions entre agents, processus endogènes et exogènes, etc.

Ces éléments ont été mis en pratique dans la participation à la conception et à l'implémentation d'un second modèle de simulation, « SimpopNet-Réseaux », « modèle-jouet » servant de prototype pour le modèle SimpopNet développé plus tard (SCHMITT 2014). Il s'agissait cette fois-ci de modéliser la co-évolution entre systèmes de villes et réseaux de communication, en simulant des potentiels d'interactions entre villes par l'intermédiaire de réseaux routiers formalisés par des graphes.

7. Respectivement portés par Denise Pumain (ERC GeoDiverCity), Arnaud Banos (ANR MIRO2) et Lena Sanders (ANR TransMonDyn).

Visualisation et évaluation de modèle. Pour rendre compte de ces deux modèles – SimpopLocal et SimpopNet –, il m’avait été demandé de développer des outils de visualisation et d’exploration de leurs comportements. Les outils de visualisation intégrés à la plateforme de modélisation, NetLogo, n’étaient en effet pas suffisants pour rendre compte des différentes dynamiques produites par ces modèles. Dans un premier temps, pour étudier l’effet des perturbations sur SimpopLocal, j’ai implémenté un type de représentation utilisé pour montrer l’évolution des rangs des villes d’un système, en m’appuyant sur les « rank clocks » de BATTY (2006). Cette visualisation d’une itération à la suivante était très adaptée, mais ne permettait d’évaluer qu’une unique dimension (la stabilité des rangs) des phénomènes modélisés. De plus, il était nécessaire de régénérer manuellement ces graphiques à chaque nouvelle sortie du modèle.

C’est par le biais de la recherche d’automatisation et de proposition de plusieurs modes de représentation que j’ai été amené à découvrir le langage R et ses possibilités de création de rapports automatiquement produits à partir de jeux de données. L’outil qui en a découlé, intitulé TrajPop (analyse des trajectoires de population), a été en premier lieu mobilisé pour catégoriser et représenter les populations simulées de SimpopLocal. Cela marquait ainsi un premier pas vers une évaluation systématique des sorties de simulation, permettant en outre d’archiver de manière systématisée les résultats de simulation. Par la suite, TrajPop a été employé et amélioré pendant plusieurs années pour caractériser l’évolution des populations de systèmes de villes empiriques et non plus simulés (par exemple dans PUMAIN, SWERTS et al. (2015)).

Accompagnement à la modélisation. Un dernier aspect hérité de mes années de stagiaire/ingénieur d’étude à l’UMR Géographie-cités concerne un mode particulier de modélisation, pleinement inscrit dans le travail collectif. Il s’agit d’une approche collective, collaborative voire accompagnatrice de la modélisation. En effet, si les modèles SimpopLocal et SimpopNet étaient pilotés par des doctorants-modélisateurs, j’ai aussi participé à une expérience de modélisation commune où mon rôle, mi-modélisateur mi-accompagnateur, consistait à formaliser et implémenter des hypothèses sur la constitution de réseaux de collaboration scientifique. Marie-Noëlle Comin, géographe qui avait réalisé une thèse empirique sur le sujet (COMIN 2009), cherchait ainsi à tester différents scénarios explicatifs aux regroupements de chercheurs dans le cadre de la constitution de consortiums en vue de candidature à des financements de projets scientifiques : par affinité et historique de collaboration, par importance bibliométrique, par capacité passée à remporter des financements, etc.

Le modèle issu de cette co-construction, SearchNet, a été élaboré pendant près de deux ans, en requérant une forte perméabilité de ses concepteurs aux thématiques et usages de l’autre. Par faute de temps consacré à sa finalisation, ce modèle n’a finalement jamais été achevé et mobilisé dans une publication scientifique. Cette expérience, que l’on pourrait qualifier d’avortée, m’a pourtant permis de réaliser que pour mener à terme un projet de modélisation for-

tement collectif, il était nécessaire de disposer de beaucoup de temps, de motivation, et qu'une partie non négligeable de ces deux ressources rares devait être dédiée à l'apprentissage, des deux côtés, des thématiques et méthodologies mobilisées dans un modèle. Il me semble ainsi nécessaire que le modélisateur s'accoutume pleinement au sujet thématique traité, et que le thématicien s'intéresse aux aspects méthodologiques pour être en mesure de comprendre les implications de l'implémentation informatique du modèle.

Ma formation universitaire et mes expériences passées de modélisation à base d'agent, entre géographie et géomatique, ont considérablement influencé la manière dont le présent travail de recherche a été abordé. Ces éléments n'expliquent pas, à eux seuls, l'ensemble des choix faits dans cette thèse, mais peut-être permettent-ils de mieux les comprendre, qui plus est au regard du contexte dans lequel cette thèse a été conçue et réalisée.

1.2 Modéliser l'espace sur le temps long : un travail d'interdisciplinarité

Le sujet initial de cette thèse a été conçu alors que j'étais ingénieur d'étude en analyse de données et représentations cartographiques, sous l'encadrement de Lena Sanders, pour le projet ANR TransMonDyn⁸. Ce projet interdisciplinaire, officiellement mené de 2011 à 2014, visait à « modéliser les grandes transitions de l'évolution du peuplement dans l'Ancien et le Nouveau Monde : contraintes environnementales, interactions spatiales et innovations sociales dans la dynamique multi-échelles de systèmes complexes ».

C'est au sein de ce projet que s'est constitué le groupe de travail qui a donné lieu à SimFeodal, et c'est dans TransMonDyn que ce modèle, les hypothèses sur lesquelles il repose et les grandes lignes conceptuelles qui l'animent, ont été conçus. Il me paraît donc indispensable de préciser les cadres conceptuels et pratiques de ce projet qui ont servi de base aux grandes directions empruntées dans cette thèse.

1.2.1 Modélisation de processus spatiaux

L'entrée principale de TransMonDyn était résolument spatiale. Le projet cherchait à identifier et à modéliser de grandes transformations dans les systèmes de peuplements. Dans une première phase de travail, la quarantaine de participants initiaux ont identifié douze cas d'études (tableau 1.1) où les systèmes de peuplement avaient été profondément et durablement modifiés dans leur dimension spatiale. Ces cas d'études portaient sur des régions et échelles spatiales variées – d'une micro-région nord-américaine ou française jusqu'au monde entier –, sur des périodes et étendues temporelles variées – de milliers d'années à la préhistoire à un unique siècle de l'époque contemporaine –, mais avaient en commun, comme le descriptif du projet⁹ le mentionne, « un “avant” et un “après” radicalement différents du point de vue de l'occupation de l'espace par les sociétés ».

8. Programme « Blanc » de l'Agence Nationale de la Recherche (ANR), sous le code ANR-10-BLAN-1805, www.transmondyn.parisgeo.cnrs.fr

9. Accessible sur le site de l'ANR - <https://anr.fr/Projet-ANR-10-BLAN-1805>

N°	Titre	Période étudiée	Zone géographique	Porteur(s)
1	Sortie d'Afrique	- 70 000	Monde	J.-M. Hombert, C. Coupé
2	Néolithique Bantu	- 1000 / 1000	Afrique subsaharienne	J.-M. Hombert, C. Coupé
3	Village formation in Pueblo societies	600 / 1300	Sud-Ouest États-Unien	T. Kohler
4	Émergence des villes	- 8000 / 2000	Monde	D. Pumain
5	Concentration de l'habitat de l'Âge du Fer	- 600 / - 400	Gaule méridionale	P. Garmy, J.-L. Fiches, L. Nuninger
6	Romanisation	- 200 / 100	Gaule méridionale	M.-J. Ouriachi, F. Bertoncello
7	Antiquité tardive : une transition ?	100 / 600	Gaule méridionale	F. Favory, C. Raynaud
8	800-1100 : polarisation et territorialisation	800 / 1100	Europe du Nord-Ouest	S. Leturcq, E. Lorans, X. Rodier, E. Zadora-Rio
9	Transition urbaine : 18ème - 19ème siècles	1700 / 1900	France	A. Bretagnolle, A. Franc
10	Urbanisation de l'Afrique du Sud	? / 2000	Afrique du Sud	C. Vacchiani-Marcuzzo
11	Littoralisation des systèmes de peuplement	700 / 2010	Monde	C. Ducruet
12	Émergence de métropoles polycentriques "Mega City Regions"	1960 / 2050	Monde	F. Le Néchet

TABEAU 1.1 – Les 12 cas d'étude, ou « transitions », du projet TransMonDyn - www.transmondyn.parisgeo.cnrs.fr/transitions-etudiees/cas-empiriques

Des processus génériques : la recherche de faits stylisés. L'une des ambitions principales était de parvenir à identifier des grands types génériques de transformations spatiales (concentration, dispersion, sédentarisation, hiérarchisation...). On cherchait à caractériser ces transformations, matérialisations spatiales de processus sociaux, sous forme de « faits stylisés », c'est-à-dire de « présentation[s] simplifiée[s] (sous la forme d'une relation entre phénomènes, d'une structure temporelle ou spatiale) d'une régularité empirique sur l'observation de laquelle il y a un assez large consensus dans la communauté scientifique » (NUNINGER et al. 2017, p. 70).

Une fois les faits stylisés relatifs aux changements structurels identifiés, le projet ambitionnait de parvenir à caractériser les principaux leviers et catalyseurs de ces évolutions, et ainsi de proposer des processus explicatifs et génériques. Par exemple, quels étaient les effets spatiaux d'une période de violence sur un système de peuplement ? Ce système tendrait-il à se disperser, à se concentrer, à se hiérarchiser ?

Modéliser avec des systèmes complexes. Afin de conceptualiser et formaliser ces faits stylisés liés à des processus spatio-temporels, il avait été choisi de les modéliser en tenant compte de leur nature complexe. Pour arriver à une description systémique aussi parcimonieuse que possible de ces transformations, les membres de TransMonDyn avaient ainsi décidé de modéliser les systèmes affectés et leur évolution, en cherchant à « endogénéiser » autant que possible les éléments déclencheurs des changements. Cette endogénéisation consiste à provoquer les transformations spatiales de manière endogène, c'est-à-dire en les expliquant par les dynamiques propres d'un système plutôt que par une perturbation extérieure au système (on parle alors d'élément exogène).

Une fois conçues de manière endogènes, les dynamiques menant aux transformations peuvent être exprimées sous la forme de phénomènes émergents, les inscrivant dans le paradigme de la modélisation de systèmes complexes que Arnaud BANOS exprime ainsi :

« Selon l’acception la plus courante aujourd’hui, un système complexe est constitué d’un grand nombre d’éléments en interactions non linéaires, situés dans un environnement. Ces éléments (ou entités) actifs, dénommés agents dans la terminologie informatique usuelle, agissent dans et sur cet environnement, qui les influence en retour. Un tel système ne bénéficie pas, de plus, d’un mode de contrôle global, centralisé. Le pouvoir d’action des agents est réduit à une dimension très locale, et certaines structures globales observées sont le fait de processus d’auto-organisation. Dans une telle perspective, les multiples interactions, qui plus est localisées, entre agents peuvent conduire à l’apparition de propriétés à un autre niveau d’observation ou d’agrégation, intermédiaire ou global, non déductibles à partir des simples propriétés des agents. Ces propriétés sont dites émergentes et leur identification constitue l’un des principaux enjeux des théories de la complexité. »

BANOS 2013, p. 39-40

1.2.2 Des processus inscrits dans la longue durée

Une particularité importante du projet TransMonDyn et des systèmes complexes modélisés en son sein est leur inscription dans le temps long, et pour la plupart, dans un temps long situé dans un passé lointain relativement aux thèmes habituellement traités en géographie.

Identifier des « régimes » et des « transitions ». Ce recul historique, tant en termes de position que d’étendue temporelle, s’exprime notamment par la manière dont les transformations spatiales sont décrites. On les décrit comme des « transitions » dans le système de peuplement, selon une logique qui rappelle les transitions de phase chimiques¹⁰. Ces transitions occurred de manière brusque relativement aux « régimes » qui les précèdent (« régime 1 ») et en découlent (« régime 2 »).

Pour décrire ces termes, on peut s’appuyer sur l’une de ces transitions, originellement dénommée « Transition 8 » (de son rang dans l’ordre chronologique des cas d’étude) ou « 800-1100 : polarisation et territorialisation en Europe du Nord-Ouest » (voir le tableau 1.1), qui constitue aussi le cas d’étude de ce travail de thèse.

Dans cette transition¹¹, le régime 1, hérité de l’antiquité tardive, est caractérisé

10. Dans l’ouvrage « Peupler la Terre » (SANDERS 2017), qui constitue un bilan du projet TransMonDyn, un chapitre entier (PUMAIN, SANDERS et al. 2017) porte sur le choix de cette terminologie.

11. La description originale peut être consultée sur le site internet du projet TransMonDyn :

par un pouvoir centralisé et un habitat majoritairement dispersé et constitué de villages de faible population. La transition en tant que telle est caractérisée par un émiettement des pouvoirs seigneuriaux, l'essor urbain, une polarisation de l'habitat rural et par une stabilisation de l'habitat. Après cette transition, dans le régime 2, le système féodal a été mis en place, l'habitat est polarisé autour des églises paroissiales et socialement structuré par des communautés rurales.

Longue durée et modélisation. L'inscription conjointe dans un passé lointain et sur la longue durée pose plusieurs problèmes à la modélisation. En premier lieu, sur le plan conceptuel, l'étude d'un système sur le temps long impose d'en caractériser les composantes de manière pérenne. L'exemple des « villes » en est caractéristique : la définition de ces entités socio-spatiales varie selon les disciplines, mais aussi selon les époques considérées. Pour certains, parler de villes pour caractériser les agglomérations secondaires héritées de l'antiquité constitue ainsi un anachronisme complet. La première difficulté d'un projet tel que l'étude d'une transition est donc déjà de formaliser, de manière ontologique, des objets valides sur l'ensemble de la période étudiée, alors même que presque par définition, ces objets sont amenés à, potentiellement, se transformer d'un régime à l'autre, jusqu'à parfois même changer d'identité.

Longue durée et sources : le problème des « connaissances expertes »

Une autre difficulté majeure auquel le projet TransMonDyn devait faire face, par la volonté de modéliser des transitions sur la longue durée, est la nature forcément lacunaire des sources empiriques sur lesquelles reposent les connaissances de ces transitions. Plus l'on remonte dans le passé, moins les données empiriques sont nombreuses (voir figure 1.1-A), et plus elles sont incomplètes et incertaines. Pour augmenter la quantité de données, on peut alors procéder à de la numérisation de sources anciennes (B), ou simuler les processus anciens pour produire des données (C).

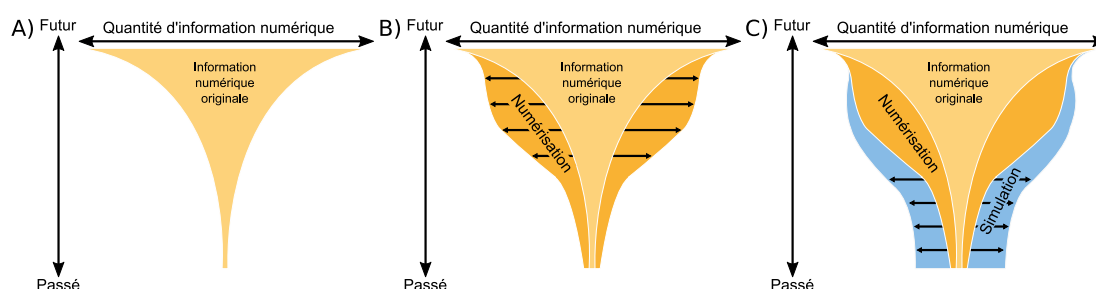


FIGURE 1.1 – Le « champignon informationnel », d'après KAPLAN (2013).

L'exercice de la modélisation demande pourtant une quantité substantielle de données, et surtout que ces données soient homogènes en termes de couverture spatio-temporelle et en termes de certitude face au risque que des « effets de source » ne biaisent les résultats. Afin d'augmenter la couverture spatio-temporelle, on a recours à des sources variées – matérielles, écrites, voire biologiques – par nature hétérogènes.

Les éléments empiriques ne peuvent dès lors reposer que sur une connaissance large et diversifiée des périodes et régions étudiées. Cette connaissance, que l'on peut qualifier d'experte, est l'apanage des « thématiciens » du projet. Le modélisateur, qui ne peut acquérir l'ensemble des connaissances expertes des thématiciens, doit accepter de leur faire entièrement confiance quant aux éléments empiriques mobilisés dans le modèle et par l'intermédiaire desquels les modèles seront ensuite évalués.

Cela mène à une double implication en matière de légitimité. Pour le modélisateur, cela implique qu'il sera toujours nécessaire de s'en remettre à la connaissance experte d'une personne (ou d'un groupe), sans possibilité d'ailleurs d'enrichir ces connaissances de son côté : mobiliser une référence scientifique dans une thématique de recherche inconnue ou distante, c'est risquer de citer des travaux non reconnus par la communauté, dépassés, ou encore anecdotiques. Dans cette thèse, les tentations de référencer certains éléments empiriques en menant des recherches bibliographiques ont été nombreuses, mais sans vision d'ensemble de l'historiographie de ces sujets, cela n'ajouterait en fait aucun gage de scientificité.

Pour le thématicien, cela implique d'être en permanence « malmené » par un modélisateur en recherche de connaissances plus précises et exhaustives. Pierre GARMY le résume ainsi à propos de son expérience en tant que thématicien dans le projet TransMonDyn :

« La collaboration interdisciplinaire entre thématiciens et modélisateurs suppose le dépassement de deux contradictions : exhaustivité tendancielle vs parcimonie recherchée d'une part et complexité vs schématisation ou stylisation d'autre part.

Il existe un véritable paradoxe entre l'incomplétude de fait des données – que les spécialistes disciplinaires cherchent à combler progressivement par l'enrichissement continu des corpus au moyen de recherches appropriées, rentabilisées par la définition de problématiques préalables aussi pointues que possible – et l'attente des interlocuteurs modélisateurs qui veulent tout savoir et se bercent souvent d'illusions sur l'état de l'art réel dans chaque champ de connaissances. »

GARMY P., « Annexe 1 - Retour sur expérience d'un "thématicien" », in OURIACHI, MATHIAN et al. (2017, p. 476)

1.2.3 Un contexte fortement interdisciplinaire

L'interdisciplinarité est le dernier élément marquant du contexte dans lequel ce travail de thèse a été initié et mené jusqu'à son terme. Celle-ci est au cœur d'un projet tel que TransMonDyn, où collaboraient des géographes, archéologues et historiens, informaticiens et géomaticiens, épistémologues, linguistes, mathématiciens... Une expérience précédente, Archaeomedes (ARCHAEOMEDES 1998), avait déjà posé les bases d'une certaine interdisciplinarité entre géo-

graphes et archéologues¹², et plusieurs de ses membres historiques se sont donc retrouvés dans TransMonDyn. Ce « noyau dur » a constitué le pôle auxquels se sont agrégés d'autres chercheurs des disciplines mentionnées, et y compris en archéologie, des chercheurs qui avaient exprimé un désaccord quant aux approches très empreintes de systémique d'Archaeomedes (voir FERDIÈRE et al. 2000, par exemple).

Dans le projet TransMonDyn, on a souvent distingué les rôles des participants selon qu'ils étaient « thématiciens », porteurs de connaissance experte vis-à-vis des processus étudiés, ou « modélisateurs », porteurs d'une connaissance experte vis-à-vis de la manière de modéliser ces processus. Au sein même de ces catégories, l'interdisciplinarité était fortement présente et a orienté les approches mises en place dans le projet.

Interdisciplinarité et thématique. En premier lieu, une forte interdisciplinarité entraîne nécessairement des points de vue disciplinaires différents sur des thématiques communes. Il me semble que ces points de vue sont fortement liés aux sources manipulées. Les données statistiques et issues d'enquêtes contemporaines de la géographie urbaine, les textes et la littérature grise historiques ou les traces et matériaux archéologiques, c'est-à-dire l'ensemble des éléments sur lesquels les disciplines se constituent, influencent fortement la manière de considérer un thème commun, comme celui de l'étude d'un système de villes par exemple. Les géographes en analysent les interactions et populations au moyen de données de recensement ou d'analyses des comportements des acteurs territoriaux. Les historiens se réfèrent aux sources historiques, le plus souvent issues des élites ou structures dominantes, qui communiquent sans doute des informations précises sur les liens politiques ou religieux entre les villes. Les archéologues cherchent dans les traces matérielles des points communs en termes de nature de construction, de répartitions spatiales à l'échelle des lieux de fouilles afin d'identifier par exemple l'existence d'échanges entre les agglomérations. Ces descriptifs, certes caricaturaux, montrent bien que pour une thématique commune, la disponibilité et la couverture spatio-temporelle des sources privilégiées par telle ou telle discipline orientent nécessairement l'approche.

Ces sources et perspectives différentes orientent aussi fortement les ontologies et lexiques utilisés. On a donné l'exemple des villes, mais chaque discipline use de son propre jargon et se dote de concepts spécifiques pour le définir. Un des défis de l'interdisciplinarité, du point de vue des thématiques, est alors de parvenir à tout expliciter, en usant d'un formalisme commun (définitions, ontologies...), de manière à ne pas laisser de place à l'implicite disciplinaire.

Une approche variée, en termes de recherche d'explication des processus observés, découle nécessairement de ces différences de vocabulaire et de sources

12. On se réfère ici aux chercheurs en géographie humaine et, ou urbaine. En géomorphologie, géophysique et géographie environnementale, les collaborations entre géographes et archéologues, sur des questions techniques notamment, sont plus fréquentes et anciennes.

considérées. Dans cette thèse, on a essayé de concilier les explications en intégrant l'ensemble des éléments constitutifs du système modélisé qui semblaient, pour l'un ou l'autre des thématiciens, pouvoir apporter une part d'explication. Il fallait ainsi concilier la vision d'une géographe modélisatrice qui cherchait à expliquer la polarisation aux moyens d'attractions différenciées, la vision d'une archéologue pour qui les effets de lignages seigneuriaux et la territorialisation due aux paroisses permettaient de comprendre cette même polarisation et fixation, ou encore un historien pour qui les communautés rurales/agraires/payannes émergentes étaient un facteur indispensable de la fixation de l'habitat rural.

Ce dernier point est aussi l'occasion d'estomper la différenciation, forte dans TransMonDyn, entre « thématiciens » et « modélisateurs ». À mon sens, l'un des accomplissements de ce projet est aussi d'avoir montré que la frontière entre ces « rôles » est fine, et extrêmement dépendante d'un positionnement particulier dans le cadre d'un projet particulier. Au-delà du poncif qui reviendrait à dire que chacun est tour à tour modélisateur ou thématicien selon son interlocuteur, il me semble ainsi pouvoir retirer de l'expérience TransMonDyn que pour modéliser en interdisciplinarité, chacun doit en même temps adopter ces deux postures, devenant un « modélicien » selon les termes d'Arnaud BANOS :

« Sous le néologisme “Modélicien”, je désignais un type d'interaction très différent, qui avait amené Robin Cura et Cécile Tannier, tous deux géographes et modélisateurs, à s'investir avec une telle intensité et une telle profondeur dans la problématique historique de leur groupe de travail qu'ils en étaient progressivement venus à s'exprimer comme s'ils avaient été eux-mêmes longuement formés à cette discipline (cf. [CURA et al. 2017]). L'évolution était tellement flagrante et systématiquement soulignée par les collègues historiens eux-mêmes que j'en suis même venu à émettre la possibilité d'un syndrome de TransMonDyn (en référence bien sûr au syndrome de Stockholm), défini de la manière suivante : “phénomène psychologique selon lequel des modélisateurs partageant longtemps la vie des thématiciens développent une empathie, voire une sympathie, ou une contagion émotionnelle avec ces derniers, devenant ainsi des modéliciens”. »

BANOS A., « Annexe 3 - Petite typologie “empirique” des modélisateurs », in OURIACHI, MATHIAN et al. (2017, p. 484)

Je crois que c'est là l'une des conditions permettant à un projet de passer de la pluridisciplinarité, « juxtaposition disciplinaire » (GRAVIER 2018, p. 14), à une interdisciplinarité définie comme une « coopération de plusieurs disciplines autour de projets communs » (ibid., p. 14, citant AERES (2014) p. 19).

Interdisciplinarité et méthodologie. Au-delà des difficultés que l'interdisciplinarité implique sur le plan thématique, on peut aussi en constater des retentissements en termes méthodologiques, très visibles dans le projet TransMonDyn. La présence de géographes-modélisateurs, de géomaticiens, d'infor-

maticiens, de mathématiciens ou encore de philosophes dans un projet génère nécessairement des approches diversifiées en matière de modélisation. Le mathématicien favorisera en effet des formalismes mathématiques tels que les systèmes dynamiques ou la théorie des jeux, les philosophes préféreront une modélisation ontologique des processus à l'œuvre, le géographe préférera une entrée spatiale, potentiellement statistique par exemple autour de modèles gravitaires, et l'informaticien usera de ses habitudes en modélisation à base d'agents, proches de la programmation à base d'objets communicants.

Le panorama des approches de modélisation mobilisées dans les transitions de TransMonDyn illustre cette diversité de pratiques : on y retrouve beaucoup de modèles à base d'agents, mais aussi des modèles basés sur la théorie des jeux, ou encore un modèle conceptuel basé sur une ontologie dynamique (FAVORY et al. 2017). En considérant l'historique des approches de modélisation du projet, on pourrait de plus y ajouter des modèles graphiques, sagittaux ou chorématiques. Ces approches ont en commun une vision systémique des processus, mais les modalités de leurs mises en place s'effectuent selon des langages (informatiques ou formels) et des paradigmes (mathématiques, statistiques ou informatiques) très différents.

Au sein même des modèles implémentés au travers de systèmes multi-agents, l'hétérogénéité est forte : entre un modèle KISS¹³, un modèle descriptif, ou encore un modèle comportemental individu-centré, les paradigmes mobilisés sont entièrement différents et surtout ont des implications diamétralement opposées en matière de conception et d'évaluation. L'interdisciplinarité, de manière générale, requiert une forte explicitation des concepts et méthodes employés. Dans le cadre de projet collectifs de modélisation, c'est vrai pour l'aspect thématique – chacun doit comprendre ce qui est modélisé afin de ne pas en mésinterpréter les résultats – aussi bien que pour l'aspect méthodologique – chacun doit comprendre la manière dont les objets sont modélisés pour les mêmes raisons. L'interdisciplinarité, sur le plan méthodologique, demande donc aussi une forte explicitation afin d'assurer la compréhension partagée des méthodes mises en places, des biais qu'elles impliquent, et des limites qu'elles comportent.

Le contexte dans lequel cette thèse a été imaginée et initiée, retracé dans cette partie, permet de mieux comprendre les enjeux – méthodologiques et thématiques – de ce travail. Les membres du groupe de travail de la « transition 8 » de TransMonDyn ont continué à œuvrer ensemble, même après les fins offi-

13. Caractérisés ainsi par AMBLARD, ROUCHIER et BOMMEL (2006, p. 110) : « Un premier courant, qui est une application directe du rasoir d'Occam, aussi appelé “principe de parcimonie”, le mouvement KISS (Keep It Simple, Stupid !) recommande de construire des modèles qui soient analysables par la suite, suffisamment simples pour être disséqués par un humain qui observe les simulations attentivement [...]. Le positionnement de ce courant peut se résumer ainsi : rien ne sert de concevoir des modèles dont on ne pourrait étudier sérieusement les propriétés et oublier ainsi la validation interne, définie comme l'existence des bonnes propriétés du modèle dans le cadre formel de ce dernier ».

cielles (2014) et effectives (2017, à la parution de l'ouvrage collectif « Peupler la Terre » dirigé par SANDERS (2017)) de ce projet de recherche interdisciplinaire. Le modèle SimFeodal, et cette thèse plus largement en constituent l'un des aboutissements.

Ma thèse, en tant que telle, n'a pourtant pas été réalisée dans le cadre officiel de l'ANR TransMonDyn. Le présent travail de thèse a ainsi bénéficié d'un financement du LabEx DynamiTe¹⁴, proposé par le groupe de travail « Systèmes de Peuplement sur le temps long »¹⁵, dont l'objectif initial était de « croiser les connaissances et savoir-faire de géographes, historiens, archéologues et mathématiciens pour décrire, conceptualiser et modéliser les dynamiques du peuplement sur le temps long dans leurs expressions spatiales et leurs rythmes temporels »¹⁶. Au regard de ces objectifs, on comprendra aisément les similarités avec TransMonDyn, que ce soit en termes d'interdisciplinarité ou de démarche.

Dans les faits, ce groupe de travail « temps long », dans lequel je me suis inscrit tout au long de mon travail de thèse, a surtout œuvré à la résolution d'une des difficultés de l'interdisciplinarité décrite plus haut : la mise en place d'un vocabulaire commun et explicite sur les concepts et notions liés à la description et à la modélisation des systèmes de peuplement sur le temps long. La réalisation d'un « lexique spatio-temporel illustré »¹⁷ qui résulte de ce projet, intitulé « Les concepts-clés des systèmes de peuplement sur le temps long » (SANDERS, BRETAGNOLLE et al. 2020), a nourri l'ensemble de cette thèse, en contribuant largement à mettre au clair les concepts employés. À ce titre, le groupe de travail du LabEx a aussi constitué un contexte fort, en mettant en place des discussions interdisciplinaires, autour des mêmes disciplines mais avec des communautés de chercheurs et d'approches distinctes de celles de TransMonDyn.

14. Laboratoire d'Excellence « Dynamiques Territoriales et Spatiales », ANR-11-LABX-0046, dans le cadre du programme « Investissements d'Avenir ».

15. Co-dirigé par Patrice BRUN, Marie-Vic OZOUF-MARIGNIER et Lena SANDERS.

16. Descriptif du groupe de travail : <http://labex-dynamite.com/fr/recherches/groupes-travail/les-systemes-de-peuplement-sur-le-temps-long/>

17. C'était le nom provisoire, modifié par la suite, de ce projet.

1.3 Du projet de thèse à une question de recherche : les conditions et modalités de la modélisation collective et interdisciplinaire

Entre mon projet de recherche initial, proposé au LabEx DynamiTe, et le présent rendu, le sujet, les questionnements et surtout les approches mises en places pour y répondre ont fortement évolué. Les grands axes du positionnement sont eux restés constants, de même que la thématique d'ensemble. Le titre initial de ce projet montre qu'il s'agit plus d'un changement de point de vue sur les approches à mobiliser que sur le thème en lui-même. Ce projet s'intitulait en effet originellement « Exploration et analyse de données spatio-temporelles : application à la modélisation des transformations des systèmes de peuplement sur le temps long ». On notera que les principales idées sont communes au titre actuel (« Accompagner la modélisation des systèmes de peuplement par l'exploration interactive de données spatio-temporelles »¹⁸), mais que l'ordre en a été bouleversé, que la logique d'accompagnement a été mise en avant (même si elle était mentionnée dans le projet de recherche, cf. *infra*) et que certaines composantes ont été spécifiées.

Le programme de recherche initial était orienté autour de trois tâches :

1. « Mise en place d'une démarche d'accompagnement des thématiciens dans la modélisation » ;
2. « développement d'une plateforme d'exploration de données spatio-temporelles : application à des transitions dans le système de peuplement » ;
3. « analyse des transitions du système de peuplement modélisées ».

Ces grandes dimensions d'analyses sont toujours présentes, largement ré-organisées et hiérarchisées différemment. Il s'agit ici de revenir sur les obstacles et solutions rencontrés durant le travail de thèse, qui ont conduit à reformuler et ré-agencer ces trois composantes initiales.

1.3.1 Position du modélisateur : du guide au con-constructeur

L'ambition d'origine était d'accompagner la modélisation des transitions en « [extrayant] les connaissances des thématiciens et [en les décryptant] afin d'en tirer des processus de causalité, de dépendance et d'inter-relations ». Il me semble maintenant que le terme de « guider » la modélisation aurait plus fidèlement représenté l'approche qui était initialement prévue. Il s'agissait ainsi plutôt d'orienter et de diriger les thématiciens vers la conception de modèles plutôt que de le faire avec eux, en m'adaptant aussi à leur point de vue.

Guider la modélisation de cas d'études suivant une approche harmonisée. Dans les faits, en matière d'accompagnement, mon projet de recherche initial visait à formaliser les discours experts des thématiciens en vue de leur modélisation. Cette formalisation devait s'appuyer sur des méthodes

18. A corriger avec titre définitif.

graphiques à l'aide de « briques de bases », modèles schématiques d'un type de processus ou d'interaction. Pour que cette approche soit généralisable à toutes les transitions identifiées dans TransMonDyn, il fallait alors que celles-ci atteignent une certaine homogénéité dans le degré de maturité de leur conceptualisation, et donc que l'ensemble des thématiciens acceptent ce type de modélisation généralisant. Il s'agissait donc de modéliser, à partir des connaissances expertes, et éventuellement des données empiriques disponibles, voire à partir de modèles déjà réalisés par les groupes de travail, chacune des transitions selon un paradigme commun et générique.

D'une part, on le constate à la lecture de l'ouvrage collectif issu du projet TransMonDyn, les modèles ont atteint des niveaux d'avancement très inégaux. D'autre part, ils ont aussi été réalisés selon des paradigmes tout aussi variés dans leur démarche (plutôt KISS ou KIDS, différents niveaux d'agentification, etc., cf. Interdisciplinarité et méthodologie, p. 16). Un type générique de modélisation, même sans aller jusqu'à l'implémentation effective, n'était donc pas applicable à l'ensemble des transitions. Une solution aurait été de sélectionner un sous-ensemble des transitions, homogène en terme d'approche et de progression vis-à-vis de l'implémentation. Là encore, chaque modélisateur ayant une approche propre (voir plus haut), une telle homogénéité était pourtant difficilement identifiable.

En allant plus loin, on peut noter que même dans le cas du projet TransMonDyn où un même modélisateur a collaboré à plusieurs transitions, il a choisi de développer des types de modèles très différents. Alain FRANC, un mathématicien qui a participé aux transitions 9 (OURIACHI, BERTONCELLO et FRANC 2017) et 12 (BRETAGNOLLE et FRANC 2017), voir tableau 1.1, a ainsi choisi d'utiliser la théorie des jeux dans le premier cas, et de réaliser un modèle descriptif statique dans le second. Cet exemple met en exergue l'obstacle qui me paraît désormais majeur dans la conception de modèles différents selon un formalisme et un paradigme harmonisé : le type de modèle est un choix collectif des membres impliqués, mais il n'est pas pour autant réalisé de manière entièrement indépendante. Le questionnement initial, les données disponibles ou encore le parcours personnel des membres impliqués jouent ainsi un rôle considérable dans les choix de modélisation.

En premier lieu, le parcours méthodologique et scientifique de chacun des participants joue un large rôle, chaque modélisateur ayant naturellement tendance à approcher un problème avec les outils qu'il a l'habitude d'utiliser. Un statisticien habitué aux systèmes dynamiques à base d'équations différentielles aura du mal à concevoir un modèle conceptuel en termes d'agents et de mécanismes, et un adepte des modèles à base d'agents très parcimonieux sera sans doute frustré par un choix collectif de modèle descriptif.

En second lieu, il me semble que la combinaison du questionnement initial et la nature des données disponibles conditionne aussi très largement le choix du type de modélisation. Selon que le questionnement repose sur l'observation de dynamiques individuelles ou au contraire agrégées, on se tournera ainsi soit

vers des méthodes où l'individu peut être suivi (modèles agents, théorie des jeux, etc.), soit vers des méthodes plus agrégées (systèmes dynamiques par exemple). Selon que le questionnement porte sur le déroulement des « transitions » ou au contraire sur la caractérisation des « régimes », on choisira aussi différentes méthodes. Au niveau des données empiriques, plus celles-ci sont nombreuses, plus le type de modélisation pourra être « *data-driven* », produisant peut-être des modèles plus parcimonieux donnant une forte importance à la validation. Au contraire, en présence de peu de données empiriques mais de nombreuses connaissances expertes, des modèles « *theory-driven* » seront sans doute privilégiés, évalués et mobilisés de manière plus qualitative. Dans l'exemple du projet TransMonDyn, le choix de la modélisation agent, basée sur des connaissances expertes et non sur les données, a été un choix collectif décidé dès le début du projet. Ce n'est donc pas la disponibilité des données qui a influencé le type de modélisation, sauf peut-être pour la modélisation de l'une des transitions (BRETAGNOLLE et FRANC 2017), où les nombreuses données disponibles et leur format ont encouragé à la création d'un modèle statistique descriptif.

De la prescription à l'accompagnement. Au-delà de la question de la faisabilité effective d'un tel projet d'accompagnement global à la modélisation d'un ensemble de transitions, rendue difficile dans le contexte très interdisciplinaire et multi-approches de TransMonDyn, la réalisation de cet « accompagnement à la modélisation » posait question en matière de méthode d'accompagnement. Dans un tel contexte, la proposition d'un cadre conceptuel commun et générique de modélisation était-elle la manière la plus adaptée de réellement « accompagner » la modélisation ? Dans un parcours, un accompagnateur peut ainsi être devant, au même niveau, ou derrière ceux qu'ils accompagnent. La position que j'avais initialement prévue me plaçait « devant », c'est-à-dire en guide chargé d'orienter la marche des suivants en proposant un seul itinéraire, c'est-à-dire une approche commune et unique de modélisation. Cela n'a pas été possible pour une raison chronologique triviale – le projet TransMonDyn avait commencé avant le début effectif de cette thèse –, mais aussi vis-à-vis des discussions collectives relatives à la position des modélisateurs, unanimement choisie comme une position d'accompagnement « au même niveau » pour reprendre la métaphore précédente. Cela interroge le positionnement du modélisateur vis-à-vis des thématiciens « avec » lesquels il cherche à concevoir un modèle, et qui voulait, dans TransMonDyn, être celui d'un « accompagnateur » et co-acteur plus que celle d'un prescripteur ou d'un assistant technique.

J'ai très vite réalisé que la proposition d'un cadre commun de modélisation ne pouvait fonctionner au regard des spécificités de chacune des approches mises en places pour décrire les transitions. Pour que chacun investisse son temps et son énergie dans une expérience de modélisation, qui peut se montrer frustrante pour les thématiciens (OURIACHI, MATHIAN et al. 2017, p. 470), il faut avant tout veiller à répondre aux attentes de ces derniers vis-à-vis de l'usage du modèle. En d'autres termes, pour que des chercheurs consacrent du temps à un travail, encore faut-il que celui-ci leur apporte quelque chose. L'identification

de qui « profite » d'un modèle¹⁹ est claire dans le cadre commun de la « modélisation d'accompagnement » (voir COMMOD 2015, par exemple), où l'on cherche à sensibiliser une population aux dangers et risques qu'elle encoure. Dans le cadre d'un projet de recherche où les objectifs des modèles sont plus théoriques, où les attentes vis-à-vis des modèles sont plus diversifiées, voire où des réticences vis-à-vis de la modélisation peuvent être présentes, il me semble qu'une conception collective du modèle et de l'approche de modélisation est la seule susceptible de rassembler les énergies nécessaires.

Dès lors, la modélisation ne pouvait se faire qu'en gardant à l'esprit, constamment et tout au long, les attentes et envies de chacune des parties prenantes. Cela n'est pas possible dans une approche prescriptive de « modélisateur éclairé » (OURIACHI, MATHIAN et al. 2017, p. 484). Un consensus permanent sur la manière de procéder est ainsi nécessaire, sous risque de voir l'expérience échouer avant son terme. Pour que le projet de modélisation interdisciplinaire aboutisse, le modélisateur ne peut être un simple « implémenteur », qui ne participerait qu'à la dernière phase – l'implémentation informatique du modèle conceptuel. Comme le relate Arnaud BANOS :

« Le travail en interdisciplinarité suppose une alchimie particulière, dont l'existence n'est jamais acquise et qui doit être entretenue, nourrie en permanence. Dans cette quête incessante et toujours renouvelée, les interactions entre modélisateurs et non modélisateurs – surtout si ces derniers s'inscrivent dans des disciplines relevant des sciences humaines et sociales – prennent une saveur tout à fait singulière [...]

L'alchimie dont je parlais au tout début de mon propos suppose une certaine forme d'intimité. Et du temps. Une denrée de plus en plus rare dans nos métiers. »

BANOS A., « Annexe 3 - Petite typologie "empirique" des modélisateurs », in OURIACHI, MATHIAN et al. (ibid., p. 483,485)

Ne disposant pas du temps nécessaire, ni pour créer cette intimité avec chaque groupe de travail, ni pour mener la modélisation de plusieurs transitions, j'ai considéré que mieux valait se concentrer sur la modélisation d'une unique transition. C'est ainsi qu'il était possible d'y investir le temps suffisant, plutôt que de chercher à multiplier les expériences de modélisation de plusieurs transitions, ce qui aurait pu être frustrant pour les thématiciens et autres modélisateurs impliqués.

1.3.2 Explorer plutôt que confronter des données empiriques et simulées

La deuxième tâche du projet initial de thèse était de construire une plate-forme interactive dédiée conjointement à l'exploration des données empiriques sur

19. « who benefits in what manner from the simulation » selon les mots de LAKE (2014, p. 260)

lesquelles sont fondées les modèles, à l'exploration des données simulées produites par ces modèles, et enfin à la confrontation de ces données empiriques et simulées. La structure des données disponibles a entraîné une reformulation de cette ambition.

D'une modélisation descriptive à une modélisation dynamique. L'idée sous-jacente était de compléter les connaissances expertes autant que possible en mobilisant les données empiriques existantes afin de faciliter/accélérer/-détailler les faits stylisés identifiés lors de la conception des modèles conceptuels. Je savais que les archéologues, en particulier, du projet TransMonDyn s'appuyaient sur de vastes corpus de données correspondant à l'enregistrement et à la formalisation informatique de matériaux issus de fouille. Les transitions portées par des géographes, elles-aussi, pouvaient compter sur des données exhaustives et complètes sur le territoire et les périodes étudiées. J'envisageais donc de créer un outil permettant à chacun, thématiciens compris, d'explorer les données compilées pour chaque transition, et ainsi de contribuer à y trouver des corrélations, des récurrences spatio-temporelles ou autres motifs dans les données qui auraient pu enrichir et faciliter la conception des modèles.

Dans les faits, et c'est aussi une limite qui n'est visible qu'une fois l'exercice entamé, l'utilisation des données archéologiques est bien plus complexe que je ne pouvais le prévoir en tant que géographe habitué aux sources de données contemporaines, institutionnelles et souvent exhaustives. De nombreuses « données » existent, mais elles ne sont par forcément harmonisées ou homogènes, et surtout, elles sont très lacunaires sur les dimensions temporelles (des périodes entières ne sont pas couvertes) et spatiales (les données ne couvrent que rarement l'ensemble du territoire étudié). À une exception près, les modèles finalement issus de TransMonDyn sont d'ailleurs tous « *theory-driven* » plutôt que « *data-driven* »²⁰ : ils reposent sur des connaissances expertes et des hypothèses thématiques plutôt que sur une information quantifiée qui serait directement injectée dans les modèles. Dans la plupart des cas²¹, les données en tant que telles ont été mobilisées comme des moyens d'évaluer les modèles plutôt que comme des bases de construction. L'usage des données empiriques, dans ma thèse, s'inscrit pleinement dans cette approche empruntée collectivement par les participants du projet TransMonDyn.

Exploration de modèles et confrontations de données. Pour « explorer les modèles », c'est-à-dire analyser l'étendue des dynamiques simulées et des sorties auxquels ils pouvaient aboutir, et mener ainsi une validation de ces modèles, je comptais mobiliser des méthodes d'exploration automatisée de mo-

20. Pour une bonne comparaison de ces approches, voir MATHIAN et SANDERS (2014, p. 120–122).

21. À l'exception de la transition 9 (cf. tableau 1.1) où les auteurs (BRETAGNOLLE et FRANC 2017) ont justement mené un travail de modélisation « statique » à partir des données disponibles, en modélisant les états successifs des potentialités d'échanges inter-urbains plutôt que de modéliser les causes de leur émergence.

dèles. Le projet ERC GeoDiverCity²², auquel je contribuais par ailleurs, comportait en effet un axe de recherche sur le développement de telles méthodes. Dans ce cadre, des outils et algorithmes d'exploration automatisée de modèles avait déjà été développés lors de la rédaction de mon projet initial de thèse. Plusieurs thèses issues de ce projet ERC avaient ainsi privilégié ce type d'approches pour explorer les modèles (SCHMITT 2014; COTTINEAU 2014; REY-COYREHOURCQ 2015).

Fondamentalement, le présupposé de telles méthodes (cf chap7) est que les modèles construits soient aussi parcimonieux que possibles, dotés d'une quantité minimale de paramètres et d'indicateurs de sortie. En automatisant les analyses, il aurait été aisé de confronter ces sorties aux données empiriques disponibles, la simulation informatique ayant l'énorme avantage de laisser au modélisateur choisir le format, le type et la quantité des données générées.

Le contexte effectif de co-construction du modèle a pourtant invalidé, ou au moins freiné, ce type d'approche. Ces méthodes d'exploration automatiques sont en effet pensées comme permettant l'exploration de modèles plutôt KISS. Pour mener une exploration complète, il faut minimiser le nombre de paramètres en entrées et le nombre d'indicateurs en sortie. Dans le cas inverse, l'« explosion combinatoire » qui résulte de leurs combinaisons empêche de mener une exploration complète.

La réalisation de modèles KISS demande cependant aux thématiciens de faire preuve d'une forte abstraction quant aux processus et entités modélisés. Dans un contexte de première expérience de modélisation pour une large part des thématiciens du projet TransMonDyn, cela rendait difficile l'adoption de ce type de modélisation qui demande « un apprentissage, une familiarisation progressive avec un mode de pensée qui n'est pas génétiquement le sien » (GARMY in OURIACHI, MATHIAN et al. 2017, p. 477). Pour des néo-modélisateurs, le choix de modèles KIDS est alors plus aisé, et le travail de « parcimonisation » qu'implique ce type de modèle dans un second temps aide justement à cette « familiarisation progressive ».

Avec des modèles plutôt descriptifs, la tâche d'exploration est plus complexe, sans parler même d'une éventuelle validation. Cette exploration des modèles, à partir des données issues de simulation, ne peut être que manuelle. À ce titre, la plate-forme interactive de confrontation des données pouvait avantageusement être repensée pour servir à l'exploration des données simulées. Au lieu de comparer des données empiriques et des données de simulation, elle servirait à explorer des données simulées, et à les confronter à d'autres données simulées afin de pouvoir comparer différents modèles ou versions de modèles par l'entremise de l'exploration interactive de données.

22. ERC Advanced Grant 269826 GeoDiverCity, dirigé par Denise PUMAIN - <http://geodivercity.parisgeo.cnrs.fr>

1.3.3 D'un méta-modèle à un retour sur expérience de modélisation

Le projet initial visait enfin à aller vers l'expression d'un modèle générique de transition dans les systèmes de peuplement – un méta-modèle –, qui serait issu d'une analyse comparée des différents modèles. La conception d'un tel méta-modèle ne pouvait se passer d'analyses poussées de plusieurs modèles de transitions, formalisés de la même manière et avec des approches homogènes. En d'autres mots, sans avoir « guidé » la modélisation de plusieurs cas d'études, ou sans intervenir complètement après l'achèvement du projet TransMonDyn, une fois que toutes les transitions auraient déjà été modélisées, il était invraisemblable d'aboutir à un tel résultat.

Cet objectif s'avérant peu réalisable au regard du contexte effectif de cette thèse, et au fur et à mesure du travail collectif de modélisation dans lequel je me suis trouvé impliqué, j'ai accordé une importance de plus en plus sensible à l'un des moyens de parvenir au méta-modèle qui avait été identifié dans le projet de thèse initial. Il s'agissait de la « la mise en place d'une démarche facilitant le dialogue et la compréhension interdisciplinaire ». À force de ré-ajustements de ce travail de thèse, c'est en définitive dans cette direction que j'ai dirigé mon travail, en cherchant à concevoir des « interfaces », méthodologiques et techniques, à la modélisation. Ces interfaces, détaillées par la suite (section 1.4.2), ambitionnent d'aider à la modélisation des transitions d'une manière qui soit satisfaisante pour chacun des acteurs impliqués, modélisateurs compris. C'est cette démarche qui est précisée dans la partie suivante et que l'ensemble de la thèse cherche à définir sur la base d'un retour sur une expérience fructueuse de modélisation interdisciplinaire. Elle permet de définir mon positionnement dans le champ de la modélisation géographique.

1.4 Visualiser et explorer pour co-construire un modèle : conception d'interfaces exploratoires

Les parties précédentes ont permis de définir et de préciser l'objet de recherche de ce travail de thèse, la modélisation de transitions dans le système de peuplement.

Le contexte dans lequel la thèse s'est déroulée (section 1.2), entre les projets TransMonDyn et le groupe de travail « temps long » du LabEx Dynamite, a amené ce thème et montré les enjeux de la modélisation des transitions comme objet de recherche interdisciplinaire, fait de l'étude de processus spatiaux complexes et de leur inscription dans le temps long.

Ma formation initiale et son extension en tant qu'ingénieur d'étude à l'UMR Géographie-cités (section 1.1) ont orienté le choix des méthodes : mobiliser le cadre méthodologique des modèles à base d'agent, en donnant une forte place aux mécanismes géographiques et géométriques, et veiller à l'importance de la représentation graphique dans leur usage.

Le questionnement initial et son évolution (section 1.3) ont permis de recentrer le sujet d'ensemble de cette recherche. Depuis la volonté initiale de créer un méta-modèle, il s'agissait désormais de documenter et promouvoir l'approche d'ensemble mise en place tout au long de ce travail collectif et interdisciplinaire.

Dans cette dernière partie, je souhaite préciser et justifier cette approche de modélisation interdisciplinaire qui est (1) résolument inscrite dans la co-construction et (2) s'appuie sur la mise en places d'interfaces disciplinaires (thématiques et méthodologiques) pour faciliter cette co-construction interdisciplinaire. Cette approche est (3) délibérément exploratoire, aussi bien pour la conception du modèle que pour son analyse, tout en visant à (4) assurer une reproductibilité des moyens, méthodologiques et techniques, mis en œuvre.

1.4.1 Favoriser une co-construction interdisciplinaire

Quand on cherche à initier une démarche de modélisation collective, qui plus est en situation d'interdisciplinarité, les approches de « modélisation d'accompagnement » (« ComMod », pour *companion modelling*) paraissent incontournables dans le champ francophone de l'étude des systèmes complexes. Il me semble dès lors important de positionner l'approche défendue dans cette thèse vis-à-vis de la modélisation d'accompagnement, et d'en identifier les points communs et de divergence.

Animer un projet interdisciplinaire. L'approche ComMod est définie par ses créateurs comme « une démarche permettant de faciliter des processus collectifs de décision, en proposant un travail d'explicitation des points de vue et des critères subjectifs auxquels se réfèrent implicitement, voire inconsciemment, les différentes parties prenantes » (C. COMMOD 2005, p. 167). Les

partisans de cette approche utilisent la modélisation de systèmes complexes comme un outil d'interface entre les différents acteurs d'une problématique souvent environnementale. Cette modélisation est pensée comme véritablement collaborative, le modèle servant de support de discussion et de négociation entre des acteurs parfois opposés. La démarche mobilisée, bien ancrée dans le paysage de la modélisation de systèmes complexes, se fonde autour de « cinq étapes incontournables » (COMMOD 2015, p. 41-132) :

1. Identifier et formuler une question clé initiale partagée ;
2. Co-construire une représentation partagée de la question (méthode « ARDI ») ;
3. Concevoir et animer un jeu de rôles ;
4. Concevoir et développer un modèle informatique ;
5. Suivre le développement et évaluer les effets de la démarche.

Cette approche émane de chercheurs impliqués dans des opérations de « recherche-action » (C. COMMOD 2005, p. 165), et donne dès lors une importance prépondérante au travail de terrain (BARRETEAU et al. 2003, p. 3.3), en particulier comme méthode d'évaluation des modèles produits. Les modélisateurs y sont d'ailleurs présentés comme des « animateurs »²³, dont l'objectif est de pousser au dialogue les « participants », décideurs politiques, acteurs de la société civile ou encore population potentiellement à risque dans les programmes de sensibilisation.

Cette dichotomie entre « animateurs » et « participants » peut rappeler celle que l'on fait entre « modélisateurs » et « thématiciens », mais elle y ajoute une hiérarchie en termes de maîtrise du processus de modélisation dans son ensemble. Dans cette thèse, je défends une approche « démocratique », où le modélisateur n'a pas pour vocation de guider les thématiciens, mais bien de construire avec eux, de manière collective, c'est-à-dire de co-construire.

L'utilité d'un « animateur », qui peut être modélisateur comme dans la démarche ComMod mais aussi thématicien, est toutefois incontestable, notamment pour initier réellement le projet collectif et, de prime abord, réussir à concilier les intérêts thématiques et méthodologiques de chacun des membres impliqués – animateur compris – autour d'un objectif consensuel commun. Dans le groupe de travail de la « transition 8 » de TransMonDyn, c'est Cécile TANNIER qui a joué ce rôle d'animatrice, en mobilisant son expérience de modélisation pour amener les thématiciens à concevoir collectivement un premier modèle conceptuel de la transition, sous la forme d'une ontologie (TANNIER et al. 2014).

Un projet collectif et intégrateur. Dans l'approche ComMod, l'implémentation informatique des modèles conceptualisés est une option, concurrente au développement d'un jeu de rôle, qui peut être mieux adaptée aux problématiques modélisées (COMMOD 2015, p. 96-97). Le modèle de simulation joue donc le rôle de compagnon dans le processus de modélisation, mais n'est

23. Dans la méthode « ARDI » (Acteurs, Ressources, Dynamiques et Interactions) (ETIENNE 2009) qui est au cœur de la démarche ComMod.

pas un objectif en soi. Le développement du modèle de simulation demande ainsi des compétences informatiques particulières. Celles-ci sont difficiles ou peu stimulantes à acquérir dans le cadre d'un projet caractérisé par le « peu de temps disponible et l'hétérogénéité des savoir-faire informatiques des participants et la diversité de leurs attentes en la matière » (COMMOD 2015, p. 29).

D'après ma compréhension de la démarche ComMod, l'implémentation est alors le plus souvent réalisée par des modélisateurs externes au projet et le modèle permet de tester les hypothèses des participants ou de les faire évoluer. Le modélisateur-informaticien est placé dans une position extérieure mais aussi surplombante. D'où un risque que le modèle implémenté soit considéré comme « l'implémentation » du modèle conceptuel plutôt que comme une de ses implémentations possibles, et constitue donc l'unique référentiel d'évaluation de la démarche. Les participants sont ainsi inscrits dans une démarche clientéliste, plus proche de la pluridisciplinarité que de l'interdisciplinarité, que le texte suivant de TUKEY me semble bien résumer, en remplaçant le « statisticien » par l'« informaticien modélisateur ».

« Like every other methodology, statistical-and-quantitative methodology has fostered a separation of roles, here into “statistician” and “client”. Today theses roles are usually embodied in separate people. Tomorrow more persons will play both roles. But these roles will not melt together, and the need to consult another person will not altogether disappear [...] »

Too often the client (whether or not a social scientist) looks to the statistician as a man who applies the final stamp of approval – perhaps by saying, “This result is significant”. Too often the statistician looks upon himself as a guardian of the proven truth – sometimes feeling it appropriate to say “That difference is not significant at 5%, *and hence* you should act as if it were *zero*”. The roles of “client seeking stamp of approval” and “statistician guarding the proven truth” are clearly rewarding to client and statistician respectively, as persons, though they are detrimental to progress and to the advance of science. »

TUKEY 1986, p. 145

Pour la conception et l'implémentation de SimFeodal, nous avons cherché à suivre une approche plus intégratrice, en séparant moins nettement les rôles de thématicien, de modélisateur, voire d'« implémenteur ». Si chacun de ces rôles correspond à des compétences particulières peu accessibles aux autres, l'idée était tout de même d'échanger au maximum sur chacun des aspects du modèle de manière à les rendre plus compréhensibles pour tous. Par exemple, on souhaitait que les raisons thématiques d'un choix de valeur de paramètre soient explicitées, justifiées et partagées par les thématiciens auprès les modélisateurs. À l'inverse, les modélisateurs devaient expliciter les partis-pris, biais ou encore alternatives des choix d'implémentation effectués, par exemple en détaillant le point de vue adopté dans l'implémentation d'un mécanisme

concernant les actions d'un type d'agent (cf. **encadré chap2**).

Sans chercher à ce que chacun des chercheurs impliqués dans la construction de SimFeodal ne devienne omniscient, notre approche visait toutefois, dans la mesure du possible, à ce que chacun soit en mesure de comprendre et donc de défendre, au moins dans les grandes lignes, les choix de modélisation et d'implémentation qui seraient réalisés.

Une position de « modélisateur-géographe ». Dans cette approche de co-construction, on essaie de faire en sorte que chacun partage un socle de connaissances commun de manière à pouvoir comprendre les implications des différentes décisions, conceptuelles, méthodologiques ou techniques. Thématiciens et modélisateurs doivent donc tendre vers un objectif idéal de transdisciplinarité, où les points de vue disciplinaires s'effacent et se confondent (GRAVIER 2018, p. 14). Pour reprendre l'expression de BANOS (in OURIACHI, MATHIAN et al. (2017)), dans un processus de co-construction absolu, chacun devrait devenir « modélicien ».

En pratique, chacun a toutefois ses compétences et connaissances spécifiques et ne cherche pas à acquérir entièrement celles des autres. La position de co-construction effective est donc bien celle de l'interdisciplinarité, complémentaire autour de sujets communs, et non celle de la transdisciplinarité.

Il me semble en effet que ces points de vue disciplinaires sont d'une part difficilement « effaçables » dans des rôles donnés, et d'autre part extrêmement enrichissants dans un projet de co-construction. Ma position, en tant que modélisateur-« implémenteur », n'est pas celle d'un informaticien distancié qui réaliserait sur commande les lignes de code correspondant au modèle conceptuel. Cette implémentation n'est qu'une partie, non négligeable certes, du travail attendu, mais surtout, elle n'est pas automatique et indiscutable, et elle est profondément influencée par mes propres expériences, envies, connaissances et surtout interactions avec les membres du projet.

Ma position dans la construction de SimFeodal n'est donc pas celle d'un pur modélisateur. C'est une position de géographe, soucieux des dynamiques spatiales observées et attaché à l'analyse de leur genericité. C'est une position de modélisateur, attentif aux choix de modélisation de ces dynamiques et curieux de l'éventail des possibles qui peut en surgir. C'est une position d'informaticien, prudent face aux biais d'implémentations du modèle et attiré par l'esthétique d'un code parcimonieux, explicite et performant. C'est une position de géomaticien, appliqué sur les détails des mécanismes géométriques mis en œuvre et enthousiaste quant à la représentation de leurs effets. C'est une position tout à la fois composée de toutes celles-ci, qui ajoute à l'interdisciplinarité de ce projet de co-construction une certaine forme d'inter-positionnement.

1.4.2 Mettre en place des interfaces disciplinaires

Dans les projets interdisciplinaires, il est utile qu'un des membres, doté de connaissances dans les disciplines impliquées, joue le rôle de médiateur par sa position d'interface entre les disciplines. Dans le groupe de travail constitué autour de SimFeodal, il n'y a pas eu de tel chercheur, et la médiation n'a pu se faire qu'au travers une vision collective et partagée d'un objet commun, le modèle.

Dans ma thèse, j'ai mis en place différentes « interfaces », entre les disciplines et entre les chercheurs, visant à faciliter la co-construction du modèle. Le terme d'interface me paraît adapté car il est polysémique, et certains de ses sens correspondent tout à fait à la démarche de ce travail. Pour les épistémologues, l'interface interdisciplinaire renvoie aux concepts et outils partagés par plusieurs disciplines, tel qu'illustré par GRAVIER (2018) dans sa thèse (voir figure 1.2), où cette chercheuse se situe elle-même à l'interface entre géographie et archéologie. En reprenant une définition grand-public, l'interface est donc ici une « limite commune à deux systèmes, permettant des échanges entre ceux-ci » (LAROUSSE 2019).

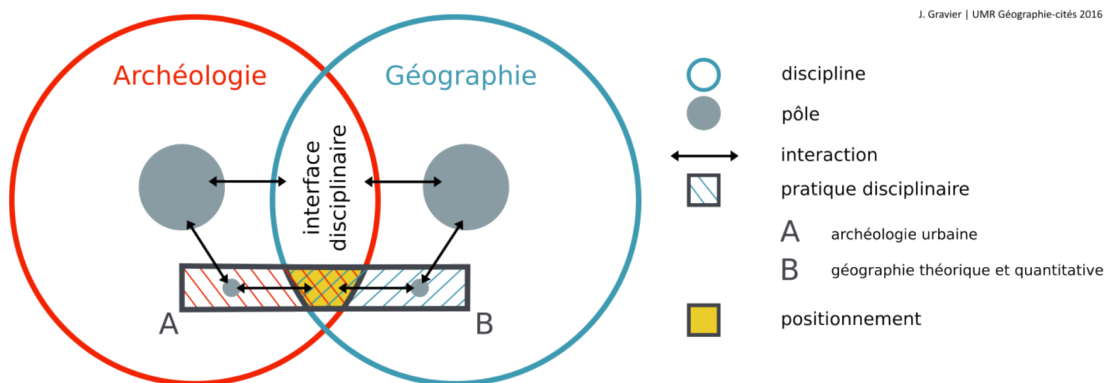


FIGURE 1.2 – Le « positionnement d'interface disciplinaire entre l'archéologie urbaine et la géographie théorique et quantitative » de GRAVIER (2018, fig. 1.3, p. 18).

Dans notre cas, c'est le modèle SimFeodal qui s'inscrit précisément dans cette même interface disciplinaire, et non l'un des chercheurs qui le co-construit. En informatique, une interface peut aussi être un « ensemble de règles, de conventions permettant un échange d'informations entre deux systèmes, deux éléments d'un système, ou entre l'utilisateur et la machine » (ACADÉMIE FRANÇAISE 2019). La dernière définition est souvent employée, sous la forme « d'interfaces utilisateur » (Interfaces Homme-Machine, IHM), qui peuvent notamment prendre une forme graphique, on parle alors d'interfaces graphiques ou, en anglais, de « *Graphical User Interface* » ou GUI.

Dans cette thèse, ces usages épistémologiques et informatiques se confondent. SimFeodal est un modèle de simulation, un outil informatique, mais sa dimension collective et heuristique constitue une interface interdisciplinaire au sens épistémologique. La démarche graphique mise en place, autour de l'évaluation visuelle, est bien une interface, c'est-à-dire une « convention qui permet un échange », pour reprendre la définition précédente (ibid.), entre les utilisateurs. Enfin, cette démarche est rendue possible par la création d'interfaces

graphiques, au sens informatique, qui facilitent la visualisation des sorties du modèle, et permettent dès lors d'appréhender le modèle de manière exploratoire.

Une interface via la modélisation. L'une des conclusions du projet Trans-MonDyn est que la modélisation joue un rôle prépondérant d'incitation au dialogue entre thématiciens et modélisateurs. La modélisation, au moins dans son aspect conceptuel ou ontologique (voir figure 1.3), est en effet reconnue comme un facteur d'explicitation et donc de mise en place de connaissances communes : « C'est l'interaction entre les thématiciens et les modélisateurs (le plus souvent un seul mais parfois deux ou trois) de chaque sous-groupe qui a été moteur dans la construction du modèle. [...] L'approche ontologique, par l'explicitation qu'elle implique, joue un rôle de médiation et facilite les échanges entre thématiciens et modélisateurs » (SANDERS, LIBOUREL et al. 2017, p. 458).

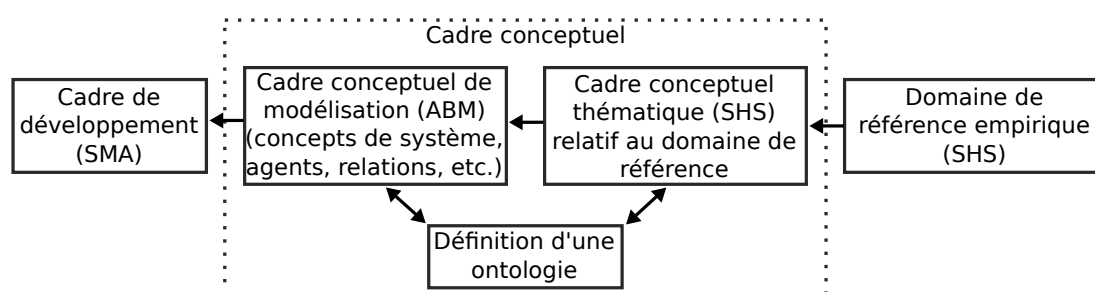


FIGURE 1.3 – « L'ontologie comme médiateur d'un dialogue » de PHAN (2014, fig. 2.7, p. 68).

Il me semble que l'on peut aller plus loin dans ce raisonnement, et le généraliser à tous les domaines du modèle, et donc aussi à l'implémentation sous forme de modèle de simulation. L'explicitation des connaissances est en effet une nécessité pour parvenir au modèle conceptuel, mais elle me semble encore plus importante lors du développement : dans un programme informatique, il ne peut y avoir aucune place au flou parce que l'ordinateur exécute sans chercher à interpréter. Les lignes de code-source sont à ce titre un formalisme non ambigu, elles expriment précisément ce que la plateforme de simulation exécutera. L'étape d'implémentation, qui aboutit sur un modèle de simulation, force donc plus encore que la modélisation conceptuelle à expliciter tous les choix effectués dans la modélisation.

Les quelque 2 500 lignes de code qui définissent le fonctionnement de Sim-Feodal ne sont pas exemptes de bugs ou d'erreurs d'implémentation, mais chacune d'elles répond à l'obligation de trancher la manière dont un mécanisme fonctionne. Chacune de ces lignes entérine l'explicitation, par la formalisation informatique, des entités et interactions du modèle. En cela, le développement d'un modèle de simulation émanant du modèle conceptuel constitue un apport qui ne se restreint pas à la possibilité de validation ou d'utilisation effective (à but prédictif par exemple) du modèle. L'implémentation permet surtout de prolonger et renforcer l'explicitation, pousse à l'échange interdisciplinaire, et à ce titre, constitue aussi une méthode qui sert d'interface interdisciplinaire.

Une interface via la visualisation. Dans cette thèse, nous mettons en place une approche basée sur la représentation graphique et l'exploration visuelle, qui est employée tout au long du cycle de construction du modèle. Le modèle est évalué en menant une analyse visuelle des sorties de simulation au travers des indicateurs de sortie majoritairement graphiques. La comparaison entre différentes versions du modèle se fait aussi par le biais d'une comparaison visuelle entre les sorties correspondantes. La sensibilité du modèle, enfin, est elle aussi analysée de manière visuelle, au moyen de représentations de la variabilité des indicateurs de sortie en fonction des répliques et valeurs de paramètres choisies.

De manière générale, je défends l'idée que les apports de la modélisation comme outil d'interdisciplinarité s'appliquent tout autant à la visualisation²⁴. Comme la modélisation à base d'agents, la visualisation est un « langage » compréhensible par le plus grand nombre, ne reposant pas sur un formalisme mathématique parfois excluant. Comme la modélisation, la réalisation d'une représentation graphique pousse à l'explicitation de ce qui est représenté, de l'origine et des particularités des données mobilisées. Comme la modélisation, la visualisation permet ainsi de créer une interface interdisciplinaire – la représentation graphique d'un ensemble de données – qui favorise alors le dialogue interdisciplinaire en le provoquant autour d'un élément accessible à tous les participants à un projet.

Pour ces raisons, les indicateurs de sortie de SimFeodal, les représentations graphiques permettant l'analyse de la sensibilité du modèle, ou encore, plus largement, SimEDB, la plateforme d'exploration des données issues de simulation, remplissent tous le rôle d'interface entre les disciplines du projet. Ces interfaces permettent aux géographes, archéologues et historiens de dialoguer autour de représentations explicites, accessibles et intelligibles.

Une interface via l'exploration. Un dernier aspect de la démarche d'ensemble choisie pour ce travail de thèse me paraît en mesure de favoriser l'interdisciplinarité, en constituant une nouvelle interface disciplinaire. Il s'agit de la nature fondamentalement exploratoire du processus (de co-construction et d'évaluation collective du modèle) mis en œuvre dans cette recherche. La sous-partie suivante du chapitre détaille cette approche, mais on peut déjà en exprimer l'intérêt dans un cadre interdisciplinaire.

L'approche exploratoire, qu'elle soit inscrite dans l'analyse de données statistiques²⁵, spatiales²⁶, ou encore spatio-temporelles²⁷ permet de favoriser un raisonnement abductif (BANOS 2005). Pour YU (1994, in BANOS (ibid., p. 2)) l'abduction désigne la « capacité de l'être humain à générer des hypothèses

24. Cette position est explicitée et détaillée dans CURA (2020, à paraître).

25. On parle alors d'*Exploratory Data Analysis* ou « EDA », d'après TUKEY (1977).

26. *Exploratory Spatial Data Analysis*, « ESDA », d'après BRUNSDON (1998) et HAINING, WISE et MA (1998).

27. On retrouve notamment cette approche, de manière très interactive, dans le champ des *geovisual analytics* (N. ANDRIENKO et G. ANDRIENKO 2006).

temporaires à partir de l'information qu'il reçoit ». En tant que position scientifique, BANOS (2005, p. 2) interprète l'abduction comme « la capacité du scientifique à se mettre en position d'étonnement, à se laisser guider par la recherche de l'inattendu ».

L'abduction est très liée à l'analyse exploratoire de données, car fortement liée à la logique exploratoire : le raisonnement abductif survient quand on accepte de raisonner sur la seule base des données disponibles, sans chercher à les interpréter en fonction d'hypothèses pré-conçues auparavant.

La modélisation peut être guidée ou influencée par un principe d'abduction (BANOS 2013, p. 77), d'autant plus quand elle s'inscrit dans une logique exploratoire, acceptant là aussi de se fier « uniquement » au modèle sans chercher forcément à interpréter ses écarts à la situation empirique comme des erreurs qui devraient être corrigées.

MORIN (1994) note qu'une approche abductive, par « l'invention d'hypothèses explicatives nouvelles », peut créer « des articulations, organisatrices ou structurales, entre des disciplines séparées et [permettre] de concevoir l'unité de ce qui était alors disjoint ».

L'approche exploratoire que nous entretenons dans ce travail collectif, en favorisant l'abduction, va aussi dans le sens d'une facilitation de l'interdisciplinarité. Cette approche de modélisation exploratoire constitue dès lors, elle aussi, une interface disciplinaire : il s'agit bien d'une limite commune à l'archéologie et à la géographie qui permet les échanges entre ces disciplines²⁸. Elle permet d'oublier, par moments, les *a priori* disciplinaires sur les causes de tel ou tel processus spatio-temporel, pour en proposer de nouvelles interprétations issues du modèle, de son exploration, et des discussions collectives entretenues autour de cette exploration.

1.4.3 Une démarche exploratoire

Dans la dernière sous-partie, je mentionnais l'intérêt d'une démarche exploratoire en ce qu'elle favorise l'abduction et donc l'interdisciplinarité. De manière plus générale, l'approche exploratoire est essentielle au raisonnement tenu tout au long de ce travail de thèse. Elle est ainsi mobilisée lors de chacune des étapes de la co-construction et de l'évaluation de SimFeodal, et constitue à chaque fois l'approche principalement suivie. Dans cette dernière partie de définition du positionnement de ce travail de thèse, il me paraît important de préciser la manière dont cette démarche exploratoire s'exprime dans cette expérience de modélisation interdisciplinaire.

Un modèle exploratoire. Les archéologues, en particulier dans le monde anglophone, font emploi de la simulation informatique depuis le début des années 1970. Parmi les chercheurs qui ont cherché à en décrire et à en catégori-

28. Pour paraphraser les termes de la définition du LAROUSSE donnée plus haut.

ser les usages, MITHEN propose une typologie des modèles de simulation selon leurs buts. Il distingue trois types de modèles de simulation, selon ²⁹ :

1. qu'ils servent à tester des hypothèses ;
2. qu'ils aient pour but d'accompagner la construction théorique ;
3. qu'ils soient conçus pour épauler le développement de nouvelles méthodes.

Ce troisième type me paraît plus distant de notre usage, l'objectif de SimFeodal n'étant pas de proposer des méthodes innovantes dans le champ de la modélisation à base d'agents. Au contraire, les objectifs de SimFeodal l'inscrivent tout à fait dans les deux premiers types, que LAKE explicite ainsi :

« When simulation models are used to test hypotheses the aim is usually to determine what actually happened in the past by comparing the output of a simulated process against the archaeological evidence. In contrast, the use of simulation models to support theory building – so-called heuristic modelling – does not necessarily, or even usually, involve detailed comparison of output against the archaeological record; in this case the purpose is not to test what happened in the past, but rather to understand how certain processes work and what sort of changes could plausibly have occurred.

[...]

[It] is important to recognise that the distinction between hypothesis-testing and theory-building simulation as conceived by Mithen is not always so clear cut in practice. »

LAKE 2014, p. 260

À mon sens, SimFeodal s'inscrit précisément dans le cas identifié où la distinction entre ces deux approches n'est pas tranchée. La conception du modèle s'appuie ainsi autant sur des théories que sur des données empiriques, en un continuum que nous appelons « connaissance experte ». Il s'agit bien de tester des hypothèses, mais sans nécessairement comparer les sorties du modèle à des données empiriques. Cette confrontation de données empiriques et simulées est d'autant moins adaptée dans notre cas que les données empiriques sont très hétérogènes et lacunaires.

L'usage principal du modèle est donc expérimental. On cherche bien à explorer les résultats d'interactions complexes entre agents, entre comportements pour accompagner la construction théorique. On cherche toutefois tout autant à expérimenter les structures et processus que le modèle fait émerger, potentiellement sous l'influence de « scénarios thématiques », ce qui nous permet de tester des hypothèses. Finalement, le modèle sert surtout de support au dialogue collectif et interdisciplinaire, mais il permet aussi, individuellement, pour chacun des participants, de repenser ou réorganiser la compréhension de la transition étudiée et des processus spatiaux modélisés. En cela, c'est l'expérience de modélisation de SimFeodal, plus encore que le modèle en lui-même,

29. MITHEN (2018, pp. 176–177 ; cité par LAKE 2014, p. 260).

qui se révèle intrinsèquement exploratoire.

Construit de manière exploratoire. La manière dont SimFeodal a été conçu, implémenté puis paramétré, pourrait sembler proche du « bricolage » à certains modélisateurs expérimentés. Par exemple, le fait que certains indicateurs de sortie de simulation – ou pire, de mécanismes – aient évolué au cours de la vie du modèle peut sembler caractéristique d’un modèle conceptuel défectueux ou incomplètement conçu.

Je souhaite au contraire défendre ce choix de modélisation, dans lequel priment les itérations entre le modèle, les résultats qu’il produit et les connaissances expertes sur lesquelles il s’appuie. Dans ce mode de développement, tout élément du modèle implémenté peut être amené à évoluer, faisant alors évoluer l’élément équivalent dans le modèle conceptuel. Si l’exploration du comportement du modèle montre qu’un type d’agent est peu utile, ou au contraire, qu’il serait judicieux de mieux distinguer des sous-types d’agents et de les doter de mécanismes propres, alors nous adapterons le modèle conceptuel. Si les thématiciens, à l’occasion de nouvelles lectures, décident que le modèle conceptuel doit être modifié, alors il sera changé et donnera lieu à une nouvelle implémentation. Ce aller-retour constant, où l’on accepte que les éléments, conceptuels ou implémentés, sont susceptibles de changer permet au fur et à mesure de consolider ou d’invalidier la représentation que l’on a du système modélisé.

Cela rend la démarche de construction de SimFeodal profondément exploratoire : dans une construction « en spirale » (MATHIAN et SANDERS 2014, p. 157), ce sont les connaissances acquises au travers de la modélisation qui guident la suite du processus de modélisation, et non directement les données empiriques ou la théorie initiale. SimFeodal est donc un modèle que l’on pourrait qualifier d’« auto-construit » – en référence à l’auto-organisation –, l’expérience exploratoire de sa construction influant sur sa forme finale de manière performative.

Évalué et analysé de manière exploratoire. De manière plus classique, l’approche exploratoire est enfin appliquée aux données de sortie du modèle, selon le cadre méthodologique déjà mentionné de l’analyse exploratoire de données (EDA). L’impermanence des indicateurs de sortie (cf. **encadré incrémentalité indicateur**) en est une conséquence directe. C’est en menant des analyses de ces indicateurs que l’on peut réaliser leurs défauts, ou leur incapacité parfois à départager. C’est aussi en menant ces analyses que l’on peut constater le besoin de créer de nouveaux indicateurs, qui donneront une vision plus précise, ou légèrement différente, d’un phénomène quelconque.

Dans l’EDA, ce sont les données – et les motifs qui en surgissent – qui guident l’analyse et orientent vers l’étude de variantes ou transformations de ces mêmes données (normalisation des données, transformation en logarithmes, etc.). Dans l’exploration de SimFeodal, ces « données », c’est-à-dire les indicateurs de sortie, guident elles-aussi l’évaluation du modèle, et leur étude entraîne de par là même des transformations de ces indicateurs (précision, distinction de catégories, changement de la méthode de calcul, etc.).

Dans l'approche classique, la visualisation n'est mobilisée qu'au début du processus d'évaluation, lors de la phase de « *face validation* ». L'approche proposée dans cette thèse, nommée « évaluation visuelle » (ref vers chapitre 3), accorde au contraire un rôle clé à la visualisation tout au long du processus d'évaluation. Les indicateurs de sortie sur lesquels on se base pour évaluer le modèle sont potentiellement classiques, mais le choix de faire reposer l'évaluation uniquement sur l'analyse visuelle de ces indicateurs me semble quant à elle assez hétérodoxe et profondément exploratoire.

De manière globale, c'est toute l'exploration du modèle qui a été menée de manière visuelle, en faisant appel à une plateforme – SimEDB – qui, comme les indicateurs et la méthode d'évaluation, a fortement évolué tout au long de cette recherche. Les « résultats » du modèle, de même que sa sensibilité, n'ont pas été analysés au crible d'objectifs pré-définis stricts et hiérarchisés. Les modalités de l'évaluation, telles que présentées dans ce manuscrit, résultent de nombreux allers-retours entre thématiciens et modélisateurs, et d'encore plus nombreux allers-retours entre le modèle et l'exploration, visuelle, de ses sorties.

1.4.4 Une démarche reproductible

La démarche d'ensemble poursuivie dans cette thèse est très exploratoire, et l'ensemble de ses composants ou étapes ont largement évolué durant les six années de cette expérience de co-construction interdisciplinaire. Pour autant, afin que cette expérience puisse avoir une portée supérieure à celle d'un projet ponctuel, nous avons collectivement porté attention à lui garantir une certaine reproductibilité. L'objectif, ce faisant, était que le modèle mis en place, et la démarche liée, puissent servir à d'autres chercheurs, notamment en histoire, en archéologie, en géographie, et de manière plus générale aux chercheurs s'intéressant à des phénomènes se déroulant sur de longues périodes temporelles. Ceci afin d'accompagner ces chercheurs, à leur tour vers la modélisation de processus spatio-temporels sur le temps long. Dans les paragraphes suivants, je détaille les méthodes mises en places dans ce projet pour appuyer la reproductibilité technique, méthodologique et conceptuelle de l'approche de co-construction de SimFeodal.

Répliquabilité technique. REY-COYREHOURCQ et al. (2017) différencient reproductibilité, répétabilité et répliquabilité en ces termes :

« Les définitions les plus courantes de ces notions sont ancrées dans le vocabulaire de la métrologie et des sciences computationnelles. Elles se rapportent donc au contrôle d'un ensemble de mesures. Dans le cas de la répétabilité, il s'agit d'effectuer des mesures répétées sur un même objet ou des objets similaires dans des conditions spécifiques. Dans un contexte de reproductibilité, ces mesures sont réalisées avec des instruments différents et dans des conditions variables (lieu, opérateur) qui doivent alors être spécifiées. La répliquabilité a été étudiée dans le domaine informatique et désigne la

capacité de reproduire à l'identique les résultats d'une analyse. »

REY-COYREHOURCQ et al. 2017, p. 411

Nous avons veillé à garantir la **réplicabilité** de tout ce qui est décrit dans cette thèse en en partageant les sources. Cela concerne bien sûr le modèle SimFeodal en lui-même³⁰, mais aussi les expériences qui ont été simulées³¹ et les outils permettant leur exploration³².

À chaque fois que possible, les différentes versions des logiciels et *packages* utilisés sont précisées dans les sources, et ce faisant, on approche ainsi d'une complétion de l'ensemble des « paliers de réplabilité » décrits dans la figure 1.4.

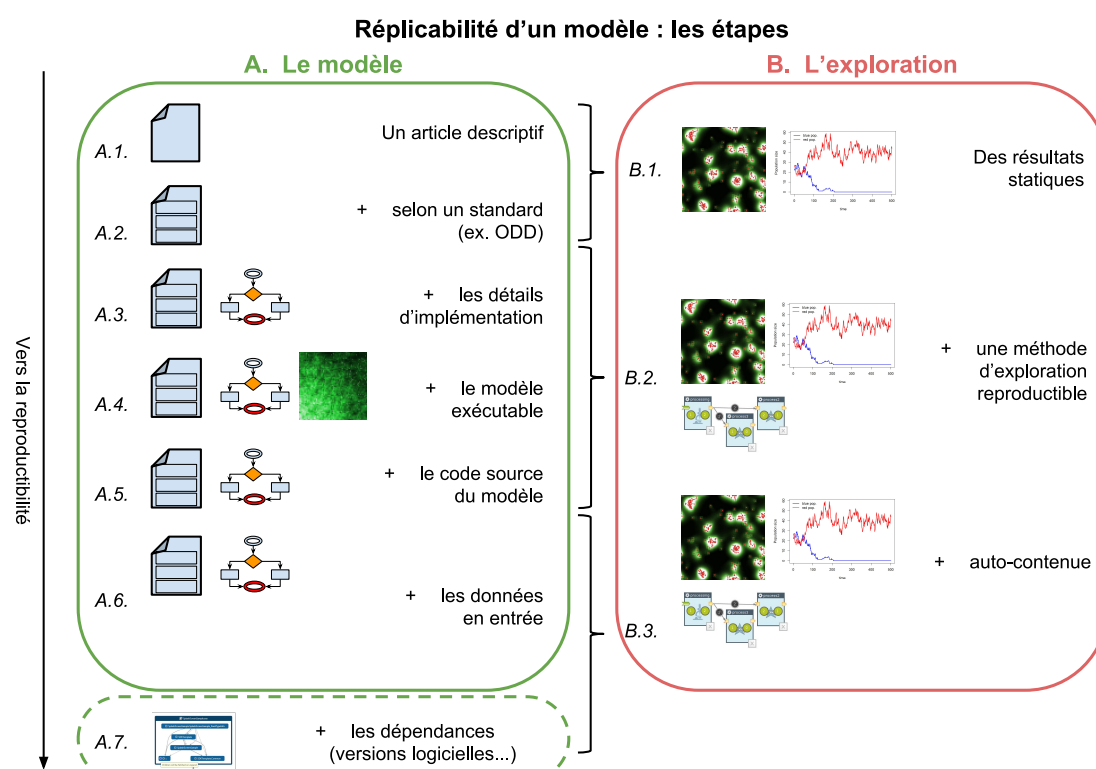


FIGURE 1.4 – « Les paliers de réplabilité d'un modèle de simulation et de son exploration » in REY-COYREHOURCQ et al. (2017, fig. 3, p. 427).

Pour renforcer cette réplabilité technique et lui donner une dimension universelle, c'est-à-dire pour que chacun puisse répliquer les expériences et résultats produits avec SimFeodal, nous avons de plus fait le choix, militant, de ne mobiliser que des outils sous licences libres. C'est le cas de la plateforme de simulation Gama (TAILLANDIER et al. 2018) avec laquelle SimFeodal a été conçue, c'est aussi le cas du langage de programmation R (R CORE TEAM 2019) utilisé pour toutes les analyses, de la base de données OmniSciDB utilisée dans la plate-forme d'exploration des données (ROOT et MOSTAK 2016), et des outils utilisés pour construire cette plateforme (CHANG et al. 2017, par exemple).

Au-delà des productions « logicielles » liées à SimFeodal et à son exploration, notons que dans le cadre de ce manuscrit de thèse, les outils mobilisés sont eux

30. <https://github.com/SimFeodal/SimFeodal>

31. Dans le dossier /experiments/ du dépôt de SimFeodal.

32. <https://github.com/RCura/SimEDB>

aussi libres : ce manuscrit est rédigé avec \LaTeX , les figures ont été produites avec R³³ ou avec le logiciel libre de dessin vectoriel Inkscape, le tout en utilisant un système d'exploitation libre pour la rédaction comme pour l'exécution des simulations.

Reproductibilité méthodologique et conceptuelle. Les solutions techniques, et la publication de leurs sources, présentées dans ce travail de thèse garantissent une répliquabilité technique. Cette dernière n'est pourtant qu'un des composants, et sans doute le plus simple, de la reproductibilité plus large que nous visons.

Nous souhaitons ainsi que l'expérience qui est la nôtre puisse servir ou inspirer d'autres projets, et avons pour cela tenté de communiquer autant que possible sur SimFeodal. Ces communications, orales ou écrites, visaient différents publics, représentatifs des concepteurs du modèle. Il s'agissait ainsi d'une part d'inspirer des archéologues et historiens à se lancer dans l'expérience de la modélisation, au moyen du paradigme des modèles à base d'agents qui nous semble porteur et accessible. Pour que ces restitutions de notre travail trouvent une résonance, et donc une audience adaptée, elles ont été effectuées devant un public d'archéologues et d'historiens sensibles aux méthodes (Journées Informatique et Archéologies de Paris, Séminaire TransMonDyn à Washington State University, organisation en cours d'un séminaire de restitution à destination d'historiens).

Il s'agissait aussi de convaincre des chercheurs pratiquant déjà la modélisation de l'intérêt de l'étude des dynamiques spatiales sur le temps long, par exemple à travers des chapitres d'ouvrages interdisciplinaires (TANNIER et al. 2014 ; CURA et al. 2017), en ciblant notamment la communauté de la géographie théorique et quantitative, que ce soit à l'occasion de colloques (ThéoQuant 2015 et 2017, ECTQG 2015, 2017 et 2019) ou de journées d'études liées à d'autres projets de modélisation.

En dehors des publications et communications qui marquent des « points d'étape » de la construction du modèle et de son exploration, nous avons fait le choix d'en « versionner » l'évolution, dans un dépôt logiciel de versionnement adapté, afin que celle-ci puisse être retracée et analysée *a posteriori*. C'est aussi le cas de la plateforme d'exploration des sorties du modèle, SimEDB, dont le dépôt retrace l'ensemble des évolutions décrites dans le **chapitre 5**.

De manière plus large, ce manuscrit s'inscrit dans la volonté de replacer le modèle SimFeodal dans le cadre plus large des modèles descriptifs, exploratoires et conçus dans un cadre interdisciplinaire. Ces dépôts logiciels concourent à l'intention de retracer le raisonnement – générique et spécifique – qui a amené au modèle SimFeodal tel que présenté dans le **chapitre 2**. En cherchant à documenter les alternatives disponibles qui ont été rencontrées tout au long de la conception du modèle et de la plateforme d'exploration de ses sorties, en

33. Les données et codes-sources en sont alors disponibles dans le dépôt logiciel de ce manuscrit.

cherchant à justifier les choix effectués, qu'ils soient conceptuels (chapitres 2 et 3), méthodologiques (chapitre 3, 5 et 6) ou parfois techniques (chapitre 5), en suivant des modes de descriptions standards (ODD pour le modèle, MCD pour la base de données), je suis convaincu d'accroître le potentiel de reproductibilité du modèle, et plus largement, de la démarche mise en place pour le construire et l'évaluer.

Cette thèse prend ainsi la forme d'un retour d'expérience, partial mais argumenté, et surtout documenté autant que possible. L'ambition est d'ouvrir la voie à d'autres expériences de ce type et de tenter de convaincre de l'intérêt, pour les projets interdisciplinaires, tout à la fois, de la modélisation à base d'agents, d'une démarche de co-construction, et de la commodité des approches visuelles pour les accompagner.

Conclusion

Dans ce chapitre, j'ai essayé de préciser les objets étudiés dans ce travail de thèse, la manière de les étudier, et les raisons, contextuelles et liées à mon parcours personnel, qui ont conduit aux approches mises en place.

De manière générale, on peut résumer ces approches au moyen des quatre termes mis en exergue dans la dernière partie de ce chapitre. Autour d'un objet de recherche constitué par la modélisation de dynamiques spatiales sur le temps long, dans un contexte interdisciplinaire, je défends ainsi une approche résolument ancrée dans la co-construction. Celle-ci va plus loin que l'accompagnement dans l'interaction entre « thématiciens » et « modélisateurs » en cherchant autant que possible à faire prendre part, à chacun des participants de projet, à l'ensemble des étapes de conception, de construction et d'évaluation du modèle. L'ambition n'est pas que chaque participant devienne expert dans chacune de ces étapes, mais que toutes les décisions adoptées dans de ces étapes soient collectives et justifiables par chacun.

La démarche de co-construction est facilitée par l'adoption d'une approche exploratoire de la modélisation. Collectivement, nous ne cherchons pas à créer un modèle qui soit validable et utilisable, par exemple à but rétro-prédictif, mais au contraire à construire un modèle exploratoire, dont la construction et l'exploration en elles-mêmes sont les vecteurs d'acquisition de connaissance par les participants au projet. En multipliant les allers-retours entre la conception, l'implémentation et l'évaluation du modèle, on adopte une position exploratoire qui s'inscrit dans une certaine visée abductive.

Pour que cette approche exploratoire puisse toutefois être mobilisée par d'autres et ne se cantonne pas à une expérimentation isolée, nous faisons le choix de documenter, pour en assurer la répliquabilité, toutes les productions informatiques mises en œuvre autour du modèle SimFeodal. Pour aller plus loin, nous documentons aussi les nombreuses évolutions du modèle, de son évaluation et de son exploration, au sein d'outils de versionnement publics et au travers de disséminations académiques diversifiées de nos travaux. On cherche ainsi à tendre vers la reproductibilité de la démarche mise en place, de manière à inscrire ce retour d'expérience dans un cadre de cumulativité des connaissances (PUMAIN (2005), citée par REY-COYREHOURCQ et al. (2017), p. 431).

Ce faisant, j'espère positionner l'ensemble des travaux décrits dans cette thèse comme des instruments d'interface disciplinaires. Les choix du paradigme de modélisation à base d'agents, d'une démarche de co-construction, de méthodes d'évaluation visuelles et interactives, sont, à mon sens, autant de solutions permettant de faciliter l'interdisciplinarité en s'inscrivant dans des domaines scientifiques accessibles et communs aux membres du projet. Le modèle SimFeodal se situe à l'interface entre géographie et sciences historiques, la plateforme d'exploration SimEDB à l'interface entre modélisation en sciences sociales et informatique liée à l'interaction homme-machine. Ces outils et les approches suivies pour leur construction participent au rôle d'interface de ce

travail de thèse, et plus largement, de ce projet collectif et interdisciplinaire de modélisation.

Ma position personnelle n'est pas celle d'une interface – je suis géographe, modélisateur, géomaticien – mais celle d'un passeur, sélectionnant des outils et méthodes éprouvés dans chacune des communautés afin de les agencer et de les mettre au service d'un projet collectif.

Références

- ACADÉMIE FRANÇAISE (2019). *Interface*. In : *Dictionnaire de l'Académie Française, 9ème Édition*. URL : <https://www.dictionnaire-academie.fr/article/A9I1677>.
- AERES (2014). *Critères d'évaluation Des Entités de Recherche : Le Référentiel de l'AERES*. Paris : AERES, p. 40. URL : <https://www.hceres.fr/fr/publications/criteres-devaluation-des-entites-de-recherche-le-referentiel>.
- AMBLARD, Frédéric, Juliette ROUCHIER et Pierre BOMMEL (2006). « Evaluation et Validation de Modèles Multi-Agents ». In : *Modélisation et simulation multi-agents. Applications pour les Sciences de l'Homme et de la Société, Hermès*, p. 103-140.
- ANDRIENKO, Natalia et Gennady ANDRIENKO (2006). *Exploratory Analysis of Spatial and Temporal Data*. Springer Berlin, Germany.
- ARCHAEOMEDES (1998). *Archaeomedes : Des Oppida Aux Métropoles*. Avec la coll. de François DURAND-DASTES et al. Paris : Anthropos.
- BANOS, Arnaud (2005). « La Voie de l'étonnement : Favoriser l'abduction Dans Les Systèmes d'Information Géographique ». In : *Apport Des SIG à La Recherche*. Colloque International Géomatique et Applications N° 1. Sous la dir. de Jean-Marie FOTSING. Orléans : Presses Universitaires d'Orléans, p. 237-254.
- (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>.
- BARRETEAU, Olivier et al. (2003). « Our Companion Modelling Approach ». In : URL : <http://jasss.soc.surrey.ac.uk/6/2/1.html>.
- BATTY, Michael (2006). « Rank Clocks ». In : *Nature* 444.7119, p. 592-596. DOI : 10/frmd5t. URL : <https://www.nature.com/articles/nature05302>.
- BRETAGNOLLE, Anne et Alain FRANC (2017). « Transition 9 : Vers Des Systèmes de Villes Intégrés (France, XVIIIe-XIXe Siècles) ». In : *Peupler La Terre : De La Préhistoire à l'ère Des Métropoles*. Sous la dir. de Lena SANDERS. Perspectives

- Villes et Territoires. Tours : Presses universitaires François-Rabelais, p. 333-359. URL : <http://books.openedition.org/pufr/10632>.
- BRUNSDON, Chris (1998). « Exploratory Spatial Data Analysis and Local Indicators of Spatial Association with XLISP-STAT ». In : *Journal of the Royal Statistical Society : Series D (The Statistician)* 47.3, p. 471-484. DOI : 10/d9fmmk. URL : <https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/1467-9884.00148>.
- CHANG, Winston et al. (2017). *Shiny : Web Application Framework for R*. URL : <https://CRAN.R-project.org/package=shiny>.
- COMIN, Marie-Noëlle (2009). « Réseaux de Villes et Réseaux d'innovation En Europe : Structuration Du Système Des Villes Par Les Réseaux de Recherche Sur Les Technologies Convergentes ». PhD Thesis. Paris 1.
- COMMOD (2015). *La modélisation d'accompagnement : partager des représentations, simuler des dynamiques*. Sous la dir. de Michel ÉTIENNE. École-chercheurs INRA 4. Bouaye : FormaSciences, FPN, INRA. 299 p. URL : <https://www.cabdirect.org/cabdirect/abstract/20163154613>.
- COMMOD, Collectif (2005). « La modélisation comme outil d'accompagnement ». In : *Natures Sciences Sociétés* 13.2, p. 165-168. DOI : 10/b3dc2p. URL : <https://www.nss-journal.org/articles/nss/abs/2005/02/nss5207/nss5207.html>.
- COTTINEAU, Clementine (2014). « L'évolution Des Villes Dans l'espace Post-Soviétique. Observation et Modélisations. » Thèse de Doctorat. Université Paris 1 Panthéon-Sorbonne.
- CURA, Robin (2011). « Conception de modèles et d'outils de géosimulation ». Mémoire - Rapport de stage. Paris : Université de Paris 1 Panthéon Sorbonne. 58 p. URL : https://documentation.ensg.eu/index.php?lvl=notice_display&id=50345.
- (2020, à paraître). « Visualisation de modèle ». In : PUMAIN, Denise. *La modélisation en géographie. Villes et territoires*. Méthodologies de modélisation en sciences sociales. Paris : ISTE.
- CURA, Robin et al. (2017). « Transition 8 : 800-1100. Fixation, Polarisation et Hiérarchisation de l'habitat Rural En Europe Du Nord-Ouest ». In : *Peupler La Terre - De La Préhistoire à l'ère Des Métropoles*. Sous la dir. de Lena SANDERS. Perspectives Villes et Territoires. Presses universitaires François Rabelais, p. 301-332. DOI : 10.4000/books.pufr.10620. URL : <https://hal.archives-ouvertes.fr/hal-01666498>.
- ETIENNE, Michel (2009). « Co-Construction d'un Modèle d'accompagnement Selon La Méthode ARDI : Guide Méthodologique ». In : *Laudun, Cardère éditeur*.
- FAVORY, François et al. (2017). « Transition 7 : Du Monde Antique Au Monde Médiéval (IVe-VIIIe Siècles) ». In : *Peupler La Terre : De La Préhistoire à l'ère Des Métropoles*. Sous la dir. de Lena SANDERS. Perspectives Villes et Territoires. Tours : Presses universitaires François-Rabelais, p. 273-299. URL : <http://books.openedition.org/pufr/10611>.

-
- FERDIÈRE, Alain et al. (2000). « La modélisation des systèmes de peuplement : débat à propos d'un ouvrage récent, Des Oppida aux métropoles ». In : URL : <https://halshs.archives-ouvertes.fr/halshs-00591267>.
- GRAVIER, Julie (2018). « Deux mille ans d'une ville en système. Proposition d'une démarche appliquée au cas de Noyon ». Thèse de Doctorat. Paris : Université Paris 1 Panthéon-Sorbonne. 404 p. URL : <http://www.theses.fr/s91330>.
- HAINING, R., S. WISE et J. MA (1998). « Exploratory Spatial Data Analysis ». In : *Journal of the Royal Statistical Society : Series D (The Statistician)* 47.3, p. 457-469. DOI : 10/drjwz2. URL : <https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/1467-9884.00147>.
- ISO, International Organization for Standardization (2018). 9241-11 :2018 - *Ergonomie de l'interaction homme-système - Utilisabilité*. URL : <https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-2:v1:fr>.
- KAPLAN, Frédéric (2013). *Lancement de la « Venice Time Machine »*. URL : <https://fkaplan.wordpress.com/2013/03/14/lancement-de-la-venice-time-machine/>.
- LAKE, Mark W. (2014). « Trends in Archaeological Simulation ». In : *Journal of Archaeological Method and Theory* 21.2, p. 258-287. DOI : 10/gd6b57. URL : <http://link.springer.com/10.1007/s10816-013-9188-1>.
- LAROUSSE (2019). *Interface*. In : *Dictionnaire de français Larousse*. URL : <https://www.larousse.fr/dictionnaires/francais/interface/43685>.
- MATHIAN, Hélène et Lena SANDERS (2014). *Objets Géographiques et Processus de Changement : Approches Spatio-Temporelles*. Systèmes d'information Géographique. ISTE Editions. 178 p.
- MITHEN, Steven (2018). « Simulating Prehistoric Hunter-Gatherer Societies ». In : *Simulating Societies*. Routledge, p. 165-193.
- MORIN, Edgar (1994). « Sur l'interdisciplinarité ». In : *Bulletin interactif du Centre international de recherches et études transdisciplinaires* 2.2. URL : <https://ciret-transdisciplinarity.org/bulletin/b2c2.php>.
- NUNINGER, Laure et al. (2017). « Un Cadre Conceptuel Générique Pour Décrire Des Transitions Dans Les Systèmes de Peuplement : Application à Un Corpus de Douze Transitions Entre 70 000 BP et 2 050 ». In : SANDERS, Lena. *Peupler La Terre : De La Préhistoire à l'ère Des Métropoles*. Perspectives Villes et Territoires. Tours : Presses universitaires François-Rabelais, p. 55-88. URL : <http://books.openedition.org/pufr/10527>.
- OURIACHI, Marie-Jeanne, Frédérique BERTONCELLO et Alain FRANC (2017). « Transition 6 Dite « Romanisation » (Ile Siècle Av. J.-C. - Ier Siècle Apr. J.-C) ». In : *Peupler La Terre : De La Préhistoire à l'ère Des Métropoles*. Sous la dir. de Lena SANDERS. Perspectives Villes et Territoires. Tours : Presses universitaires François-Rabelais, p. 243-271. URL : <http://books.openedition.org/pufr/10605>.
- OURIACHI, Marie-Jeanne, Hélène MATHIAN et al. (2017). « De l'élaboration de La Transition à Sa Modélisation : Retours Interdisciplinaires Sur l'expérience TransMonDyn ». In : *Peupler La Terre : De La Préhistoire à l'ère Des Métropoles*. Sous la dir. de Lena SANDERS. Perspectives Villes et Territoires.

-
- Tours : Presses universitaires François-Rabelais, p. 461-485. URL : <http://books.openedition.org/pufr/10671>.
- PHAN, Denis (2014). « Ontologies et modélisation par SMA en SHS ». In : *Ontologies et modélisation par SMA en SHS*. Hermes Science Publ.-Lavoisier, p. 53-94.
- PUMAIN, Denise (2005). « Cumulativité Des Connaissances ». In : *Revue européenne des sciences sociales. European Journal of Social Sciences* XLIII.131, p. 5-12. DOI : 10.4000/ress.357. URL : <http://ress.revues.org/357>.
- PUMAIN, Denise, Lena SANDERS et al. (2017). « Convergences Disciplinaires Sur Le Concept de Transition ». In : *Peupler La Terre : De La Préhistoire à l'ère Des Métropoles*. Perspectives Villes et Territoires. Tours : Presses universitaires François-Rabelais, p. 33-53. URL : <http://books.openedition.org/pufr/10515>.
- PUMAIN, Denise, Elfie SWERTS et al. (2015). « Multilevel Comparison of Large Urban Systems ». In : *Cybergeo : European Journal of Geography*. DOI : 10.4000/cybergeo.26730.
- R CORE TEAM (2019). *R : A Language and Environment for Statistical Computing*. Vienna, Austria. URL : <https://www.R-project.org/>.
- REY-COYREHOURCQ, Sébastien (2015). « Une plateforme intégrée pour la construction et l'évaluation de modèles de simulation en géographie ». Thèse de doctorat en Géographie. Paris : Université Paris I - Panthéon-Sorbonne.
- REY-COYREHOURCQ, Sébastien et al. (2017). « Vers Une Recherche Reproductible Dans Un Cadre Interdisciplinaire : Enjeux et Propositions Pour Le Transfert Du Cadre Conceptuel et La Réplication Des Modèles ». In : *Peupler La Terre. De La Préhistoire à l'ère Des Métropoles*. Sous la dir. de Sanders L. Villes et Territoires. Presses universitaires François Rabelais, p. 409-434. DOI : 10.4000/books.pufr.10647. URL : <https://hal.archives-ouvertes.fr/hal-01677950>.
- ROOT, Christopher et Todd MOSTAK (2016). « MapD : A GPU-Powered Big Data Analytics and Visualization Platform ». In : *ACM SIGGRAPH 2016 Talks*. SIGGRAPH '16. New York, NY, USA : ACM, 73 :1-73 :2. DOI : 10/gd7hg8. URL : <http://doi.acm.org/10.1145/2897839.2927468>.
- SANDERS, Lena, éd. (2017). *Peupler la terre : De la préhistoire à l'ère des métropoles*. Avec la coll. de Sander VAN DER LEEUW. Perspectives Villes et Territoires. Tours : Presses universitaires François-Rabelais. 528 p.
- SANDERS, Lena, Anne BRETAGNOLLE et al., éd. (2020). *Les concepts-clés des systèmes de peuplement sur le temps long*. Perspectives Villes et Territoires. Tours : Presses universitaires François-Rabelais.
- SANDERS, Lena, Thérèse LIBOUREL et al. (2017). « Points de Vue Ontologiques Sur Les Transitions Des Systèmes de Peuplement ». In : *Peupler La Terre : De La Préhistoire à l'ère Des Métropoles*. Perspectives Villes et Territoires. Tours : Presses universitaires François-Rabelais, p. 435-459. URL : <http://books.openedition.org/pufr/10656>.
- SCHMITT, Clara (2014). « Modélisation de la dynamique des systèmes de peuplement : de SimpopLocal à SimpopNet. » Thèse de Doctorat. paris : Univer-

-
- sité Paris I - Panthéon-Sorbonne. URL : <https://tel.archives-ouvertes.fr/tel-01077891/document>.
- TAILLANDIER, Patrick et al. (2018). « Building, Composing and Experimenting Complex Spatial Models with the GAMA Platform ». In : *GeoInformatica*. DOI : 10/gfw9nz. URL : <https://doi.org/10.1007/s10707-018-00339-6>.
- TANNIER, Cécile et al. (2014). « Une Ontologie Pour Décrire Les Transformations Du Système de Peuplement Européen Entre 800 et 1100 ». In : *Ontologies et Modélisation Par SMA En SHS*. Sous la dir. de Denis PHAN. Hermès - Lavoisier, p. 289-310. URL : <https://hal.archives-ouvertes.fr/hal-01021976>.
- TUKEY, John Wilder (1977). *Exploratory Data Analysis*. Addison-Wesley Series in Behavioral Science. Reading, Mass : Addison-Wesley Pub. Co. 688 p.
- (1986). « Statistical and Quantitative Methodology ». In : JONES, L. V. *The Collected Works of John W. Tukey : Philosophy and Principles of Data Analysis 1949-1964*. T. III. CRC Press, p. 143-183. Google Books : JuQD42kUxzIC.
- YU, Chong Ho (1994). « Abduction ? Deduction ? Induction ? Is There a Logic of Exploratory Data Analysis ? » In : Annual Meeting of American Educational Research Association. New Orleans, Lousiana, p. 28. URL : <https://eric.ed.gov/?id=ED376173>.
- ZHENG, Jiaqi (2019). « Interactive Visual Analytics for Agent-Based Simulation : Street-Crossing Behavior at Signalized Pedestrian Crossing ». Mémoire. Helsinki, Finland : Aalto University. 50 p. URL : <https://aaltodoc.aalto.fi/443/handle/123456789/40790>.