

Évaluer et paramétrer un modèle de simulation complexe en situation d'inter-disciplinarité

Version 2019-11-11

- 31/10/2019 : Fusion anciens chapitre 3 (évaluation) et 4 (paramétrage)
- 07/11/2019 : Fin reprise (3.1 et 3.2) + création (3.3)
- 08/11/2019 : Impression Lena
- 11/11/2019 : reprises commentaires Lena du 24/11/2017 + finalisation

Sommaire

| | |
|---|-----------|
| Introduction | 2 |
| 3.1 Comment évaluer un modèle? | 3 |
| 3.1.1 Évaluation, validation, vérification... : désambiguïsation | 4 |
| 3.1.2 Les étapes de l'évaluation d'un modèle | 7 |
| 3.1.3 Une évaluation de la plausibilité d'un modèle : la « <i>face validation</i> » | 12 |
| 3.1.4 Vers une évaluation visuelle | 19 |
| 3.2 Des indicateurs pour SimFeodal | 23 |
| 3.2.1 Indices et indicateurs | 23 |
| 3.2.2 Hiérarchiser et catégoriser les indicateurs | 29 |
| 3.2.3 Les indicateurs et dimensions de SimFeodal | 33 |
| 3.3 Paramétrage du modèle SimFeodal | 42 |
| 3.3.1 Les paramètres | 43 |
| 3.3.2 Le paramétrage | 54 |
| Conclusion | 70 |

Introduction

Dans le chapitre précédent, nous avons présenté un modèle, SimFeodal, dont l'objectif est de réussir à répondre à un questionnement thématique (ref 2.1.2) portant sur la capacité des agents et mécanismes modélisés à reproduire les phénomènes de polarisation, de hiérarchisation et de fixation de l'habitat paysan entre les IX^e et XIII^e siècles. Pour juger de la capacité du modèle à reproduire ces processus empiriques, il faut être en mesure d'évaluer le modèle, c'est-à-dire de lui fixer des critères de succès et de mettre en place une méthodologie permettant de juger si oui ou non ces critères sont atteints.

Ces critères sont au cœur d'un domaine incontournable et fortement étudié de la modélisation, quelle qu'en soit le type : l'évaluation de modèle. Dans le cas de SimFeodal, modèle descriptif et complexe, l'évaluation ne peut se faire de manière formelle, c'est-à-dire analytique, tant les interactions entre agents et mécanismes sont nombreuses, non linéaires, et donc non prédictibles. L'évaluation ne peut donc être qu'expérimentale, en analysant le comportement du modèle sur la base de ses sorties.

SimFeodal pose de plus le problème d'être un modèle basé sur des connaissances expertes plutôt que sur des données directement intégrables ou confrontables. Cela en complexifie l'évaluation, car les approches classiques sont peu adaptées à ce type de modélisation interdisciplinaire basé sur des hypothèses. Il est en effet difficile de quantifier, voire parfois de spécifier de manière explicite, les attentes vis-à-vis du modèle.

Les difficultés liées à l'évaluation du modèle sont renforcées par le type de construction mis en place pour SimFeodal : en plus d'être descriptif, SimFeodal est un modèle exploratoire, dans lequel l'évaluation n'a pas vocation à valider une version définitive et statique du modèle, mais au contraire à guider son amélioration par des ajustements successifs.

Dans le chapitre précédent, on décrivait par exemple une « version » du modèle, intitulée « version 6.3 », ce qui montre qu'auparavant, il y a nécessairement eu au moins 5 versions préalables, et sans doute bien plus de sous-version. Ce chapitre vise ainsi, aussi, à présenter le processus d'amélioration du modèle, que nous nommons « paramétrage », c'est-à-dire l'approche théorique et empirique qui a guidé l'évolution de SimFeodal.

Dans ce chapitre, nous présenterons donc en premier lieu le démarche globale d'évaluation, et en particulier une proposition de méthode d'évaluation – l'évaluation visuelle – adaptée aux modèles du type de SimFeodal. Nous pourrons alors spécifier cette approche en l'appliquant au cas de SimFeodal, c'est-à-dire en présentant les critères retenus pour l'évaluation du modèle. Dans une troisième partie, il sera alors possible de décrire le paramétrage du modèle, son articulation avec l'évaluation précédemment décrite, et de mener une analyse rétrospective de l'évolution de SimFeodal.

nécessaire de spécifier les critères d'observation et les réponses attendues. Ces éléments, les critères d'évaluation, ne peuvent être formulés par n'importe qui : si le modélisateur autant qu'un expert externe peuvent les spécifier, il convient de s'assurer de l'expertise – thématique et de la connaissance du système tel que modélisé – de l'évaluateur. On obtient ainsi un système à évaluer au filtre d'une grille d'analyse qualitative et basée sur le visuel. Il devient alors possible d'apprécier l'écart le modèle et le système qu'il représente, sans chercher pour autant à quantifier ou à mesurer cet écart. Il s'agit en effet plutôt d'ordonner différentes versions ou paramétrages d'un modèle de simulation afin de juger de ceux qui semblent minimiser cet écart.

Cette méthode, contrairement à d'autres, plus quantitatives, permet donc au final de tirer avantage des méthodes qualitatives telles que la *face validation* – par exemple la capacité d'évaluer un modèle qui ne reposerait que sur peu de données empiriques ou encore sur des données incertaines –, tout en se confortant à une démarche d'évaluation rigoureuse, loin de l'estimation « au doigt mouillé » à laquelle peuvent donner lieu certaines méthodes reposant sur la plausibilité et l'estimation.

3.1.4.5 Des critères pour l'évaluation : les indicateurs

Cette petite transition a été rédigée après rendu de la version imprimée, et l'introduction de la partie 3.2 (page suivante) a été aussi reprise parce qu'elle ne collait plus du tout...

Dans la présentation de la démarche d'évaluation visuelle, nous précisons que pour que cette évaluation qualitative et experte soit rigoureuse, il était nécessaire de fixer des objectifs de manière préalable, et d'en expliciter la teneur autant que possible.

Pour l'évaluation de SimFeodal, face à la multiplicité des attentes thématiques vis-à-vis du modèle, nous avons choisis de mobiliser à cet effet des « indicateurs de sortie ». Ceux-ci relèvent du domaine de la simulation et leur évaluation doit être guidée par les connaissances empiriques, formalisées au sein d'« indices empiriques » qui correspondent à ces indicateurs.

Pour finir de décrire la démarche d'évaluation de SimFeodal, reste donc à définir plus précisément ces composantes de l'évaluation, ainsi qu'à expliciter les objectifs fixés pour chacun des indicateurs.

la base d'une seule exécution. Au contraire, seule l'exécution d'un certain nombre de **réplications** (voir **chapitre 1**) permet de s'assurer que le comportement évalué correspond bien au comportement habituel, ou tendanciel, du modèle.

3.2 Des indicateurs pour évaluer le modèle Sim-Feodal

Le modèle SimFeodal présenté dans le **chapitre 2** correspond à la « version 6.3 » du modèle souhaité, c'est-à-dire qu'il en constitue une version qui n'est ni la première, ni sans doute la dernière dans cette expérience de co-construction interdisciplinaire de modèle qui s'inscrit résolument dans le temps long. L'ensemble des mécanismes figurant dans le modèle conceptuel ont été implémentés mais l'ensemble des liens, interactions et valeurs de paramètres ne sont pas encore stabilisés. De ce fait les résultats des simulations ne répondent pas nécessairement aux attentes définies dans le **chapitre 2**.

Si l'on a déjà décrit le principal objectif du modèle dans le chapitre précédent (celui de comprendre les mécanismes sous-jacents au processus de polarisation qui s'est déroulé entre 800 et 1100), il convient ici d'explicitier comment les résultats d'un tel processus peuvent être saisis. Ceux-ci sont en effet nombreux et hétérogènes, concernant aussi bien des concentrations de foyers paysans que l'émergence de pôles. Certains sont centraux, d'autres secondaires, et le modélisateur a des attentes relativement à l'ensemble des résultats obtenus en fin de simulation. La description précise de ces attentes se révèle importante dans le cadre du paramétrage – et de l'ensemble des étapes de la vie du modèle – de SimFeodal.

Dans cette partie, on explicitera d'abord le sens que l'on prête à ces attentes, sous la forme « d'indices empiriques » et « d'indicateurs de sortie de simulation ». Ces indices et indicateurs sont nombreux, certains sont multivariés, et il s'agira donc de présenter des méthodes visant à réduire la complexité de ces indicateurs de sortie, en adoptant une démarche proche de ce qui se fait en statistiques : réduction de dimensionnalité et/ou catégorisation et hiérarchisation de ces indicateurs. L'utilisation de ces méthodes permettra, seule, de décrire et qualifier le comportement du modèle SimFeodal tel qu'il a été décrit dans le chapitre précédent, avant d'en analyser les résultats par ce biais dans le **chapitre 6**.

3.2.1 Indices et indicateurs

On attend d'un modèle, sans entrer encore dans le détail, qu'il reproduise au moins les grands traits de l'élément empirique dont il cherche à rendre compte. Ces grands traits peuvent s'entendre de multiples manières, et se formaliser avec encore plus d'approches. Ici, nous avons souhaité proposer une dichotomie simple entre le domaine de l'empirique et celui de la simulation, en systématisant l'usage d'un vocabulaire qui est souvent employé de manière plurielle. Pour être en mesure d'évaluer la vraisemblance du comportement reproduit par le modèle sur le plan empirique, il est nécessaire de mettre en correspondance des éléments empiriques et des éléments issus de la simulation. Nous caractérisons ces éléments en deux grands ensembles : (1) **les indices empiriques**, éléments quantifiables ou au moins descriptibles émanant du domaine empirique, et (2) **les indicateurs de sortie**, variables informa-

tuent le cœur.

Pour cela, on peut faire appel à une méthode d'analyse spatiale assez classique (proche de la méthode des quadrats) en carroyant l'espace et en comptant le nombre d'églises paroissiales de chacune des mailles. Un indicateur simple est alors de faire un compte des mailles contenant au moins une paroisse ce qui permettra alors d'appréhender simplement la part de l'espace couvert par des églises paroissiales.

On s'attend à ce que l'indicateur ainsi produit augmente au cours du temps, à mesure que de nouvelles églises paroissiales viennent desservir le territoire, de manière régulière (comme l'évolution du nombre d'églises paroissiales).

3.2.3.4 Évaluer SimFeodal pour en permettre le paramétrage

Ajout de cette petite conclusion intermédiaire.

Le saut de page était présent pour en « garder la place », mais sera supprimé.

La mise en place de critères d'évaluation est indispensable pour un modèle. L'évaluation peut ainsi permettre de « valider » la version définitive d'un modèle, mais surtout, elle en guide la construction et l'évolution au sein du processus que nous nommerons paramétrage. En effet, c'est en évaluant le modèle après chacune des modifications apportées que l'on peut définir si ces dernières l'améliore – approchant les sorties du modèle de l'objectif – ou le détériore – en augmentant l'écart entre sorties et attentes.

La construction de SimFeodal a été entièrement guidée par ces étapes récurrentes et systématiques de paramétrage : l'approche d'évaluation visuelle a ainsi été mise au service du processus de co-construction du modèle.

Dans la dernière partie de ce chapitre, nous montrerons donc comment le paramétrage a rythmé et déterminé la progression du modèle vers la version présentée dans le chapitre 2, en commençant par définir ce terme, peu usité, qui est construit à partir de la notion de paramètre.

3.3 Paramétrage du modèle SimFeodal

Reprise de l'intro.

Idem pour saut de page, à supprimer après inclusion dans version de Lena.

Le modèle, tel qu'il a été présenté dans le chapitre précédent, était un « état », présenté en tant que « version 6.3 », c'est-à-dire que les mécanismes, paramètres et les valeurs de ceux-ci correspondent à une étape d'un modèle amené à évoluer pour répondre aux problèmes soulevés dans la dernière partie (Ref dernière section chap 2). Cette version 6.3 implique que de nombreuses versions précédentes ont été développée, testées, puis modifiées et ajustées jusqu'à obtenir une nouvelle version plus satisfaisante du point de vue des objectifs.

L'évaluation de la « satisfaction » ressentie pour une version du modèle a été présentée dans les premières parties de ce chapitre, et nous décrivons maintenant l'étape qui suit (ou précède, dans une logique faite d'allers-retours nombreux) cette évaluation, et vise à ajuster les valeurs de paramètres du modèle et le détail des mécanismes afin d'augmenter la qualité du modèle du point de vue de l'évaluation définie.

Dans cette thèse, nous proposons l'usage du terme de « paramétrage », peu fréquent, et, à notre connaissance, forgé récemment (HIRTZEL 2015; TANNIER 2017), pour décrire ce processus d'« amélioration » itérative du modèle, répété systématiquement durant le cycle de vie du modèle SimFeodal. Nous nous attacherons ainsi à présenter le travail de paramétrage réalisé depuis la version 0 jusqu'à la version « finale » du modèle dont les résultats sont présentés dans le chap6, en passant par la version présentée dans le chapitre 2, chacune de ces itérations ayant abouti à une version plus adaptée aux questions des thématiciens.

Avant de préciser le sens du terme « paramétrage », il semble important de définir précisément ce qu'est un paramètre. C'est en particulier nécessaire en ce que ce terme recouvre de nombreux sens selon les champs disciplinaires qui l'emploient, mais aussi, au sein même de ceux-ci, par les différents chercheurs.

types de paramètres de s'exprimer en valeurs compréhensibles et exploitables. Dès lors, leurs valeurs sont propres à chaque version, sous-version ou expérience du modèle formalisé, et une comparaison de ces valeurs entre les différents modèles n'apporte pas de connaissance. Ils sont amenés à varier d'une manière uniquement guidée par l'évaluation du modèle lors du paramétrage, sans que cela n'ait le moindre ancrage ou répercussion empirique.

Exemples : distance de fusion entre les agrégats, pondération de la satisfaction matérielle des foyers paysans en fonction du nombre de droits acquittés, montants récupérés par les seigneurs selon les types de droits...

Avec cette typologie des paramètres basées non sur la nature de ceux-ci mais sur leur utilisation dans le modèle, nous nous inscrivons dans une vision fonctionnaliste et donc très subjective, rappelant la définition d'un modèle de MINSKY (cf. chapitre 1). Selon l'usage que l'on fait du modèle, un même paramètre pourra donc être vu comme un paramètre de contexte ou de mécanisme (par exemple selon l'état des connaissances empiriques liées à ce paramètre sur le cas d'étude traité).

Comment choisir les valeurs de paramètres ?

Ajout de cette petite conclusion intermédiaire.

Le saut de page était présent pour en « garder la place », mais sera supprimé.

Après avoir présenté notre définition des paramètres, nous pouvons désormais revenir sur le processus qui les mobilise et a demandé cette explicitation : le paramétrage.

Cette étape, que nous allons maintenant définir et illustrer, consiste ainsi notamment à « ajuster » les valeurs des paramètres, ou plus exactement de certains des paramètres, en se basant notamment sur la typologie mise en place dans ces pages (paramètres d'*input*, de contexte, de mécanisme, techniques).

3.3.2 Le paramétrage

Le paramétrage d'un modèle est souvent réduit à l'un de ses aspects, le « calibrage », étape finale de la construction d'un modèle qui cherchera à reproduire autant que possible des données empirique en faisant varier les valeurs des paramètres jusqu'à ce qu'une combinaison de celles-ci soit satisfaisante.

De nombreux auteurs ont montré que le paramétrage d'un modèle ne pouvait se réduire à cette étape, chacun employant des termes différents pour désigner le processus de paramétrage, processus le plus souvent inscrit comme l'une des composantes de l'évaluation des modèles (chez NGO et SEE (2012), cf. figure 3.3 par exemple).

Le plus souvent, une fois le modèle construit, le modélisateur s'attache à son « calibrage », en cherchant pour chaque paramètre la ou les valeurs qui permettront au modèle de s'approcher, au plus près, des données empiriques devant être reproduites, c'est-à-dire la conjugaison de « valeurs optimales » de paramètres minimisant l'écart entre les données simulées et les données empiriques de contrôle. Cette étape, que l'on nomme aussi souvent calibration par anglicisme, peut se faire de manière manuelle, par approximations successives – CROOKS et al. (2019, p. 253) nomme cela « calibration qualitative » –, par semi-automatisme, par exemple en effectuant des analyses de sensibilité – THIELE, KURTH et GRIMM (2014, §2.3–2.4) mobilisent ainsi l'analyse de sensibilité, via échantillonnage, pour faire de l'estimation de paramètres –, ou encore de manière entièrement automatique – par exemple, pour HEPPESTALL, EVANS et BIRKIN (2007) ou NGO et SEE (2012, p. 188), en suivant une méthode d'optimisation à base d'« algorithmes génétiques ».

Dans notre travail, nous souhaitons revenir sur cette approche de la modélisation, ancrant le paramétrage comme étape ultime de la construction d'un modèle, en particulier en ce que nous considérons que cette pratique de recherches de valeurs optimales est un exercice qui devrait s'effectuer tout au long de la construction du modèle, de manière plus itérative que conclusive.

3.3.2.1 Définition

Le terme de paramétrage recouvre deux sens différents, dont la distinction peut se faire selon qu'on l'utilise pour définir un processus ou pour caractériser une configuration. Ici, nous emploierons plutôt le premier cas, définissant dès lors le paramétrage comme le processus, manuel ou automatique, visant à constituer cette configuration de paramètres. Dans ce deuxième cas, le paramétrage désigne un ensemble de valeurs de paramètres, par exemple quand on mentionne un paramétrage par défaut, ou un paramétrage optimal. Pour ne pas risquer de contre-sens, nous préférons le terme de configuration de paramètres ou de « jeu de valeurs de paramètres ».

On tend à distinguer le paramétrage – passage obligé ne nécessitant pas d'être évoqué – du calibrage, processus systématique qui inscrirait le modèle comme un outil scientifique et incontestable. Nous choisissons ici de confondre ces approches, non pas en considérant le paramétrage comme un outil d'évalua-

tion du modèle, mais comme une composante inhérente à la construction d'un modèle, quelles que soient les formes et les temporalités que le paramétrage adopte. Le paramétrage est en effet une pratique utile dans la construction du modèle, car les résultats auxquels il aboutit, c'est-à-dire les valeurs de paramètres qui semblent mieux adaptés, renseignent aussi bien sur les biais des mécanismes adoptés que sur leur efficacité réelle. Par exemple, quand, après avoir ajouté un mécanisme, on se rend compte que des variations dans les valeurs de paramètres ne changent pas réellement les sorties du modèle, cela peut être l'occasion de repenser le mécanisme dans son ensemble, ou plus souvent, la manière dont le paramètre est mobilisé dans ce mécanisme. On retrouve cette logique dans l'exploration par Clara Schmitt du modèle SimpopLocal (SCHMITT 2014), qui a permis de réaliser que la variation de l'un des paramètres (*InnovationLife*) n'avait que peu d'impact sur les sorties du modèle, tout en rendant son calibrage plus complexe et instable :

« Au-dessous du seuil des 150 pas de simulation pour le paramétrage de *InnovationLife*, le calibrage du modèle est très difficile voire impossible. Au-dessus de ce seuil, le mécanisme associé au paramètre *InnovationLife* n'a plus d'effet sur le calibrage du modèle. Dans un souci de parcimonie du nombre et de la complexité des mécanismes simulés dans le modèle SimpopLocal, il est justifiable de retirer du modèle ce mécanisme qui n'est pas nécessaire à la simulation de la dynamique de croissance recherchée. »

SCHMITT (2014, p. 224)

Ajout des paragraphes de définition suivants, qui s'insèrent donc entre ce message et le paragraphe « Désambiguïsation », page 55 dans l'impression.

Dans ce travail de thèse, nous reprenons ainsi le sens du terme paramétrage tel qu'initialement employé par HIRTZEL (2015)³⁵ et ensuite explicité par TANNIER (2017) :

« Le paramétrage d'un modèle consiste à fixer les valeurs des variables et paramètres de mécanisme, au moyen d'analyses spatiales ou statistiques de données empiriques, de transcriptions de dires d'experts, ou de simulations avec le modèle. Le paramétrage comprend une phase d'estimation (statistique ou autre) des valeurs des paramètres et variables, et une phase de calibrage si celle-ci est nécessaire. »

TANNIER (2017, p. 52)

Nous ajoutons à cette définition une composante d'implémentation, quand celle-ci est minime, qui consiste à adapter l'implémentation des mécanismes de

35. « La notion de paramétrage d'un modèle est souvent associée à celle de calibrage, et ces deux notions, bien que différentes, sont parfois confondues dans la littérature (Richiardi et al., 2006). Le calibrage consiste à tester plusieurs jeux de paramètres possibles pour une variable et à choisir l'un d'eux pour l'exécution des simulations, selon sa capacité à atteindre les objectifs définis. Il constitue ainsi une étape du paramétrage d'un modèle. Cette étape n'est pas forcément indispensable : si les valeurs initialement affectées permettent d'atteindre les objectifs du modélisateur, celui-ci n'a pas besoin de procéder à un calibrage. » (HIRTZEL 2015, p. 136).

manière à en rendre le résultat plus satisfaisant. Cela n'implique pas de changer le modèle conceptuel, et n'est pas non plus véritablement un changement dans l'implémentation du modèle. Par exemple, changer l'ordonnancement du détail d'un mécanisme relève à notre sens du paramétrage.

On peut illustrer cela avec l'exemple du mécanisme de définition du contour spatial des agrégats (cf. **méca, 2.7.2.1, chap2**) : un *buffer* est actuellement appliqué autour de l'enveloppe convexe formée par les foyers paysans membres d'un agrégat afin de fusionner d'éventuels agrégats très proches. On peut adapter ce mécanisme en appliquant ce *buffer* à la fin du mécanisme plutôt qu'au milieu, une fois que l'héritage des agrégats précédents a été transféré par exemple.

Cela n'aurait pas un impact important sur le plan conceptuel, ni même d'ailleurs lors de l'observation agrégée de tous les agrégats. Toutefois, au niveau local, cela peut avoir un impact non négligeable sur l'historique des agrégats. À notre sens, il s'agit d'une adaptation du modèle du même ordre qu'un changement de valeur de paramètre, et nous incluons ainsi ce type de modifications dans le processus de paramétrage.

Désambiguïsation. De nombreux termes sont utilisés dans la littérature, souvent sans réelle distinction, pour désigner cette opération qui consiste à choisir un jeu de paramètres pour un modèle. Pêle-mêle, on y retrouve le paramétrage, le calibrage, l'ajustement.... Nous proposons une représentation graphique, dans la figure 3.13, de notre usage de ces termes, et en donnons des définitions dans l'encadré 3.4.

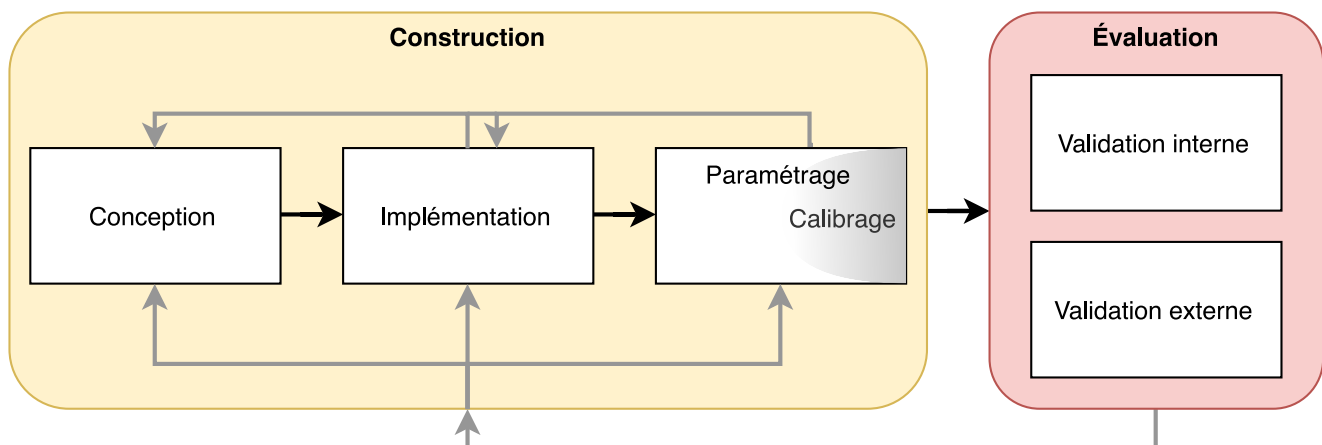


FIGURE 3.13 – Étapes du processus de modélisation.

ment que possible, mais ce paramétrage était entièrement à refaire avec la nouvelle version (modifications de la période B, figure 3.14).

Cet exemple renforce le caractère nécessaire du paramétrage, et qui plus est, montre que ce processus doit être mené de manière continue et répétée. Dans un modèle exploratoire comme SimFeodal (cf. chap1 ou chap2), on ne peut ainsi penser le paramétrage que sous la forme d'un calibrage qui serait l'étape ultime de construction du modèle avant validation. Si le paramétrage n'est pas régulier, alors chacune des versions et sous-versions successives ne pourra pas être évaluée de manière comparable, et on cours le risque de disqualifier une version du modèle qui aurait pu être plus plausible que les suivantes mais n'a pas été ajustée comme elle aurait pu l'être.

Quels paramètres ?

Ajout de cette sous-sous-sous-partie ici.

S'insère entre la fin de « Un ou des modèles à paramétrer ? » et le début de « Historique de SimFeodal. », page 62 dans l'impression.

Dans la partie précédente, nous avons distingué plusieurs types de paramètres (voir l'encadré 3.3 dans la section 3.3.1.3), selon leur usage dans le modèle et leur niveau d'inscription dans les connaissances empiriques. Cette distinction nous semble particulièrement utile pour dresser une hiérarchie des paramètres sur lesquels jouer pour adapter le modèle.

Les **paramètres d'input**, pour commencer, n'ont pas vocation à changer lors de l'évolution du modèle : ils ont été définis de manière empirique, et ne peuvent donc revêtir d'autres valeurs uniquement pour améliorer l'ajustement du modèle. Toutefois, si les connaissances expertes évoluent, les phases de paramétrage du modèle sont l'occasion de répercuter ces nouvelles connaissances. Les changements de ces paramètres peuvent résulter d'un besoin de préciser la mise en œuvre des connaissances empiriques³⁶, ou simplement de la correction d'approximations inexactes³⁷. Dans tous les cas, le paramétrage de ces paramètres n'est pas guidé directement par les résultats d'une version du modèle, mais par les connaissances qui y sont injectées. Il en va de même pour les **paramètres de contexte** : ils peuvent être précisés ou corrigés, mais leur paramétrage ne doit pas être guidé par l'évaluation du modèle.

Le cas des **paramètres de mécanisme** est plus emblématique du processus de paramétrage : ils reposent sur des ordres de grandeur plus incertains, et leur valeur peut donc assez librement être modifiée. Lors du paramétrage, il faut naturellement s'assurer de ne pas les faire varier en dehors des intervalles définis empiriquement, mais ils se prêtent très bien à des ajustements manuels. On peut ainsi sans crainte procéder à des modifications successives, par allers-retours entre l'évaluation des sorties du modèle et le changement de valeur

36. Un exemple d'un tel choix est présenté dans le chapitre 6 (partie 6.1.1.1), où l'on a choisi de faire passer les dimensions du monde simulé de 100 × 100 km à 80 × 80 km.

37. Comme c'est le cas dans l'exemple du passage de 1000 à 4000 foyers paysans présenté plus haut.

de ces paramètres. On pourra de cette manière figer, temporairement – en attendant de nouveaux ajustements dus au paramétrage d'autres paramètres –, des valeurs dont l'effet combiné sera plus satisfaisant sur le déroulement et l'aboutissement des simulations.

Les **paramètres techniques**, enfin, sont les candidats les plus évidents au paramétrage, et plus spécifiquement à sa composante de calibrage. Puisqu'ils ne reposent sur aucune connaissance empirique, ils constituent ainsi les « leviers » naturels sur lesquels on pourra mener le paramétrage, y compris de manière automatisée pour ces paramètres dont on peut faire varier les valeurs de plusieurs ordres de grandeur sans conséquence conceptuelle ou empirique sur le modèle. Comme dans l'exemple du paramétrage du modèle gravitaire (encadré 3.5), la modification de ces paramètres est extrêmement pratique pour compenser le paramétrage des autres types. Ainsi, si de nouvelles connaissances viennent pousser à modifier un paramètre de contexte par exemple (comme le paramètre de frein de la distance α), on peut espérer parvenir à compenser cette variation, sur les sorties, en modifiant un paramètre technique (k dans le modèle gravitaire) en conséquence.

Rappelons aussi, comme indiqué lors de la définition (section 3.3.2.1), qu'en dehors des valeurs des paramètres, le paramétrage est aussi l'occasion de modifier, d'une manière modérée, le détail de l'implémentation de certains mécanismes.

étant par nature d'un ordre de grandeur bien plus élevé que la médiane ⁴¹, on a normalisé (réduction sans centrage) les valeurs afin qu'elles soient comparables graphiquement. La figure 3.15 est une représentation de cette analyse ⁴², qui permet de visualiser le « rythme » des changements de SimFeodal et d'en tirer quelques remarques.

Les paragraphes suivants (+ conclusion) constituent la suite et fin du chapitre et n'étaient pas présents dans la version imprimée.

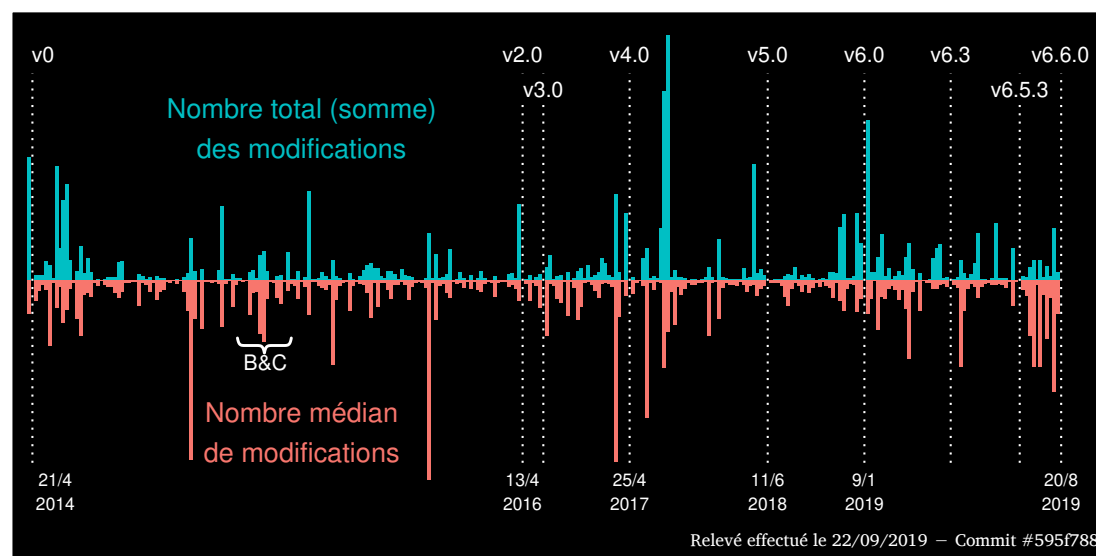


FIGURE 3.15 – Une exploration visuelle du rythme des changements de SimFeodal.

N.B. : L'axe des abscisses n'est pas linéaire du point de vue temporel : l'ordre chronologique y est respecté, mais le temps n'y est pas régulier.

On y remarque en premier lieu un certain motif, qui pourrait presque s'apparenter aux pulsations cardiaques, où l'on note des changements majeurs, espacés, entre lesquels de plus faibles phases de modifications surviennent. Ce « rythme » est avant tout une scorie de la manière dont le modèle a évolué, autour des versions représentées sur le graphique.

Les phases de plus intenses modifications, concrétisées ou précédées par des changements de version, sont ainsi intimement liées au rythme de ce travail de co-construction interdisciplinaire : les pics marquent en général la préparation, l'exécution, et les résultats immédiats des séances de travail collectives. Ces réunions étaient ainsi l'occasion d'évaluer le modèle, que ce soit en analysant ses sorties (évaluation externe) ou en ré-exposant les choix d'implémentation (évaluation interne), et en conséquence d'en mener le paramétrage, en faisant évoluer les paramètres et mécanismes. Certains pics correspondent aussi à des moments de production de « livrables », que ce soit à l'occasion de publications ou de communications, où il fallait stabiliser le modèle pour en communiquer les résultats.

41. Le changement médian maximal représente un changement de 300 lignes de codes, alors que la somme concerne, elle, plus de 1500 lignes...

42. Dans cette figure, les modifications regroupées dans l'accolade « B&C » correspondent aux phases B et C identifiées dans la figure 3.14.

À un niveau d'observation plus fin, on peut noter, aussi bien sur la somme des modifications que sur leur médiane, que les pics principaux sont le plus souvent précédés par une augmentation de la taille et de la quantité totale de modifications, puis suivis d'une diminution de ces deux indicateurs. Sur le plan « temporel », cela correspond aux phases de préparation des réunions, faite de petites explorations locales du comportement du modèle en vue de l'adapter, puis aux phases collectives résultant sur des choix de modifications plus importantes, modifications ensuite ajustées et finalement validées, jusqu'à la réunion suivante.

Sur le plan conceptuel, cela correspond à différentes phases de construction du modèle : les modifications qui précèdent les changements de version sont d'abord assez modérées, en médiane comme en somme. Cela correspond à du paramétrage assez modéré, portant sur la modification marginale de quelques valeurs de paramètres. À mesure que la taille des *commits* augmente, ces modifications sont de plus en plus importantes, et consistent ainsi à l'ajustement de plus de paramètres, à la modification plus fréquente de détails de mécanismes, etc.

Les changements de version résultent le plus souvent d'une insatisfaction vis-à-vis de l'aboutissement de ces phases de paramétrage : même en poussant le paramétrage, même en menant un calibrage approfondi on ne parvient pas à obtenir un modèle assez satisfaisant. Il est alors nécessaire de revenir sur le modèle conceptuel et son implémentation (flèches grises de la figure 3.13), et donc de mener des modifications plus conséquentes, par exemple en transformant véritablement le fonctionnement de certains mécanismes. Les mécanismes étant modifiés, avec ajout ou suppression de paramètres par exemple, la rétro-compatibilité des sorties du modèle n'est plus assurée⁴³, et l'on change alors de version.

Après ces changements de version, la taille des modifications diminue, jusqu'à atteindre des niveaux très faibles : la somme des modifications devient négligeable (à mi-chemin entre les versions 4 et 5 dans la figure 3.15 par exemple) et la médiane des *commits* est elle aussi très faible. Cela correspond exactement aux phases de paramétrages qui suivent les changements majeurs : le nombre de changements diminue, jusqu'à aboutir sur des modifications difficilement discernables qui correspondent au calibrage. Dans cette étape, seules une ou deux lignes de code changent en général, ce qui correspond aux modifications d'une valeur de paramètre.

On retrouve donc dans cette figure, et donc dans l'évolution effective de Sim-Feodal, l'ensemble des étapes décrites dans ce chapitre. Des phases de construction et d'évaluation se succèdent, avec à chaque fois un travail de paramétrage, parfois suivi de nouvelles phases d'implémentation voire de conception. On notera aussi que les versions sont assez visibles sur ce graphique, correspondant aux pics principaux, de même que les sous-versions, où les pics sont plus modé-

43. Sans même considérer d'ajout/suppression de paramètre, les modifications du sens et de l'utilisation effective de certains paramètres peuvent rendre leurs valeurs non comparables avec celles des versions précédentes.

rés : entre la version 6 et la fin du graphique (version 6.6), on arrive même sans difficulté à identifier chacune des sous-versions intermédiaires (6.1, 6.2...).

Aboutissement du paramétrage. Avec cette analyse exploratoire et l'exercice d'interprétation qui l'accompagne, nous avons voulu illustrer d'une part le processus de paramétrage mis en œuvre dans le modèle SimFeodal, mais surtout, d'autre part appuyer notre proposition de démarche de co-construction de modèles. Qui plus est dans un contexte interdisciplinaire, l'évolution du modèle ne peut ainsi passer que par des allers-retours fréquents entre son évaluation et sa modification.

Le paramétrage, et notamment sa composante de calibrage, ne peut dès lors être réduit à un exercice ultime et unique, qui viserait à « finaliser » un modèle avant son emploi. Au contraire, l'application fréquente et régulière de phases d'évaluation et de paramétrage permet de garantir que les différentes pistes de modélisation explorées sont bien comparables, car justement comparées sur des versions aussi finement ajustées que possible à chaque fois.

Conclusion

Dans ce chapitre, nous avons présenté et proposé une approche originale en matière d'évaluation de modèle de simulation à base d'agents, l'évaluation visuelle, qui nous paraît très adaptée à l'évaluation de modèles exploratoires construits dans un contexte interdisciplinaire. En proposant une évaluation basée sur la multiplication des points de vue, au moyen de nombreux indicateurs de sortie, nous pensons en effet que les différents acteurs de la modélisation peuvent chacun observer et analyser les points qui relèvent de leur expertise thématique propre. La mise en place d'une grille d'évaluation, a priori, et la hiérarchisation des critères permet de plus de garantir une vision globale de la capacité d'un modèle à produire et faire émerger les faits stylisés que l'on cherche à modéliser au sein d'une approche générative.

La forte profusion et diversité des indicateurs de sortie identifiés constitue de plus une aide précieuse au paramétrage d'un modèle : en multipliant les perspectives d'analyse des sorties d'un modèle, on facilite l'identification des mécanismes les moins satisfaisants. Avec des rapports détaillés sur le comportement d'un modèle, il est ainsi plus aisé de déterminer les composantes du modèle qui doivent être améliorés. C'est un avantage indéniable face à l'utilisation d'un indicateur quantifié unique, qui requerrait un passage systématique par une exploration globale de l'influence des paramètres et mécanismes. L'évaluation visuelle que nous proposons dans ce travail permet donc de faciliter, sur un plan conceptuel, le paramétrage, et avec lui les connaissances de chacun des membres impliqués dans la modélisation.

Nous avons toutefois aussi identifié le besoin impératif de mener cette étape de paramétrage aussi fréquemment que possible, à chaque modification du modèle. Cela implique la nécessité de mener tout aussi fréquemment une évaluation du modèle, car c'est par les allers-retours entre l'évaluation et le para-

métrage que le modèle progresse.

Alors que l'évaluation visuelle facilite l'identification des défauts du modèle, elle est en elle-même complexe à mener au regard de méthodes d'évaluation plus quantifiées et automatiques. Pour permettre une bonne progression du modèle, et donc l'augmentation de la fréquence des allers-retours entre évaluation et construction, il faut alors disposer d'outils permettant de faciliter et d'accélérer cette évaluation visuelle. La mise en place de solutions techniques et méthodologiques permettant de mener l'évaluation visuelle d'une manière collective et fluide revêt ainsi une portée cruciale dans la démarche de co-construction de modèle proposée dans ce travail de thèse.