

Chapitre 3

Paramétriser un modèle dans un contexte de co-construction interdisciplinaire

Version 2017-11-17

Le modèle présenté dans le chapitre précédent était un « état », 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). Dans ce chapitre, nous nous attacherons à présenter le travail de paramétrage réalisé à la suite, ayant abouti sur une version plus adaptée aux questions des thématiciens, et dont le descriptif technique se trouve dans l'annexe *n*(Ref à l'annexe contenant le descriptif technique de la dernière version du modèle.) Par paramétrage, et avant d'en spécifier le sens, nous entendons ici le processus visant à doter le modèle de paramètres (empiriques, « commensurables » et techniques (3.2)) lui permettant de mieux convenir aux objectifs fixés, soient-ils en terme de comportements attendus ou d'objectifs quantitatifs.

3.1 Paramétrier ? Quoi et quand ?

3.1.1 Les paramètres dans les modèles agents

Le nouveau petit Robert (Petit Robert, 2006) définit un paramètre, dans son acceptation mathématique, comme une « quantité à fixer librement, maintenue constante, dont dépend une fonction ou une expression mathématique ». Nous nous inscrivons dans la lignée de cette définition d'une part en ce qu'elle est généraliste et évite les différences sémantiques que l'on retrouve souvent entre paramètre et variable dans le champs mathématique, et d'autre part en ce que les ouvrages de référence en modélisation agent ne définissent jamais ce terme. Un ouvrage fait exception (Treuil, Drogoul et Zucker, 2008, p.8), mais nous ne pouvons souscrire à sa vision. Ainsi, pour ces auteurs, les paramètres, qui ne sont pas définis explicitement, constituent l'ensemble des « données ou éléments d'informations souvent numériques » qui viennent alimenter la structure et les mécanismes d'un modèle. Parmi ces paramètres, ils distinguent des « entrées » et des « sorties » :

à cette quand
j'aurai de fin au 2

ula aura été
dimanche ?

) év à commentez
pour corrigee.

domaine où
c'est dans porteur.

Les *entrées* d'un modèle dynamique sont des paramètres dont la valeur est définie en dehors du modèle et qui représentent ce que le simulateur peut perturber. Les *sorties* d'un modèle dynamique sont également des paramètres qui expriment ce que l'on cherche à mesurer en réponse à ces perturbations.

Treuil, Drogoul, Zucker, 2008, p.8.

Dans cet ouvrage, nous donnerons à ces termes des sens différents, voire opposés, qui nous semblent plus fréquents dans le champ de la modélisation en sciences humaines et se retrouvent partiellement dans le schéma de Balci (fig. 3.1).

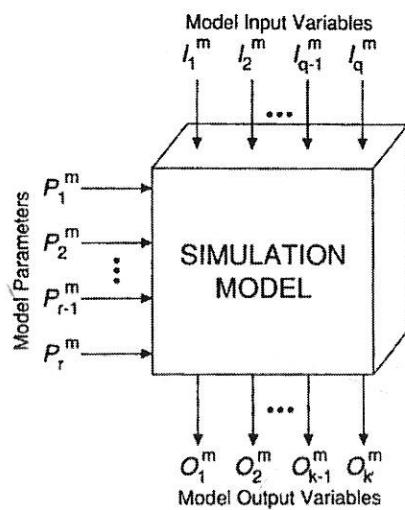


FIGURE 3.1 – Les variables d'un modèle de simulation selon Balci (Balci, 1994, p.122)

Nous considérerons donc plutôt les paramètres comme le sous-ensemble des entrées, qui peuvent s'exprimer sous forme numérique (excluant donc de fait certaines entrée telles que les configurations spatiales initiales), et que l'on fera varier, aussi bien dans la calibration du modèle que pour l'exploration de scénarios.

Paramètres et variables

Encadré 3.1 : Variables, entrées, paramètres...

Au sein des **variables**, constituées par l'ensemble des éléments d'un modèle ayant la capacité de changer, que ce soit au cours d'une simulation ou au travers de plusieurs expérimentations, nous distinguerons trois types spécifiques.

Entrées

Les **entrées** constituent l'ensemble des informations nécessaires à l'exécution d'un modèle en ce qu'elles en alimentent les mécanismes. Elles peuvent être constantes au cours d'une simulation (la date de fin d'exécution de notre modèle par exemple), ou évoluer de par l'action du modèle (le nombre d'églises paroissiale du modèle, qui est une entrée, et augmente tout au long de la simulation).

Paramètres

Certaines de ces entrées sont des **paramètres** : ce sont les entrées qui peuvent être mobilisées afin de modifier le comportement du modèle, et qui « représentent ce que le simulateur peut perturber »⁰. Dans cet ouvrage, nous distinguerons plusieurs types de paramètres (encadré 3.2) selon leur ancrage dans l'empirique. Notons toutefois dès maintenant que le nombre de paramètres d'un modèle peut être bien supérieur au nombre de paramètres que l'on fera réellement varier dans les expérimentations.

Sorties

Les sorties d'un modèle représentent ce que l'on va observer après l'exécution du modèle afin d'observer son déroulement et aboutissement. Ces sorties peuvent être très nombreuses, et concerner autant les variables individuelles (une liste des Foyers Paysans et de leur satisfaction au cours du temps par exemple) que des variables agrégées (Quartiles de satisfaction des foyers paysans par exemple), moins coûteuses en temps d'écriture ainsi qu'en espace de stockage. De la même manière que pour les paramètres, sur l'infinité potentielle de sorties d'un modèle, seules quelques-unes seront effectivement générées et enregistrées afin de minimiser les données produites à celles que le modélisateur compte explorer.

⁰. Pour reprendre les mots de Treuil et al. qu'ils appliquent aux entrées. (Treuil, Drogoul, Zucker, 2008, p.8.)

tu n'oublieras pas le terme d'indicateur ?

Différents types de paramètres

Les paramètres sont donc un sous-ensemble des entrées d'un modèle qui ont vocation à varier dans les différentes exécutions de ce modèle. Pour autant, c'est un sous-ensemble assez divers, et tous les paramètres n'apportent pas les mêmes connaissances, que ce soit par leur valeur ou par la manière dont ils varient. Il convient donc d'identifier différents types de paramètres, non selon leur caractéristique propre (c'est-à-dire leurs propriétés, qui permettent de les différencier des sorties par exemple), mais selon leur usage et donc la manière dont on pourra les mobiliser. A ce titre, quand la littérature en sciences humaines et sociales¹ distingue souvent (refs) deux types de paramètres : les paramètres empiriques (ou techniques), dont la valeur a une correspondance directe dans le domaine empirique (on peut donc l'observer), et les paramètres techniques², pour lesquels cette correspondance n'existe pas. Comme dans la plupart des typologies ayant trait à la catégorisation de valeurs numériques,

1. En physique ou en mathématiques par exemple, une telle distinction n'existe pas, la définition d'un paramètre prenant une considération plus simple d'élément rendu constant dans une exécution d'un modèle. pas clair, prendre une définition quelconque reprenant cette idée.

2. Mathian et Tannier ((Mathian & Tannier, 2015)) utilisent le terme de paramètre mécanique pour définir ce type de paramètres. Nous trouvons l'usage du mot technique plus approprié, en ce qu'il se différencie plus de l'empirique sur le plan de l'usage qui en est fait. De plus, chaque paramètre, empirique, technique ou autre, a un impact sur les mécanismes du modèle, de par la définition que nous empruntons des paramètres 3.1. Nous préférerons donc ce terme de « technique » pour éviter les confusions.

?

En y il y a aussi du paramètre technique.

de nombreux paramètres peuvent se trouver à l'interface de ces deux classes. Nous en proposons donc une troisième, que nous avons choisi de nommer « paramètres commensurables »³, qui nous semble à même caractériser des paramètres partiellement inscrits dans l'empirie, c'est-à-dire sans valeur empirique propre, mais permettant tout de même d'être mobilisés comme les paramètres empiriques, c'est-à-dire à même d'être comparés, que ce soit par les fluctuations de leurs propres valeurs ou en comparant des paramètres commensurables les uns avec les autres. Avec cette typologie des paramètres basées non sur la nature de ceux-ci mais sur leur utilisation dans un modèle, nous nous inscrivons dans une vision fonctionnaliste et donc très subjective, rappelant la définition d'un modèle de Minsky⁴ (ref, 1965). Pour différents modélisateurs, un même paramètre pourra donc être considéré comme empirique ou commensurable, ou encore comme commensurable ou technique.

phrasé

Encadré 3.2 : Trois types de paramètres

Paramètres empiriques Les paramètres empiriques sont les paramètres qui peuvent être fixés à l'aide d'une ou de valeur(s) empiriquement connue(s), ou à défaut, dont la valeur se réfère à une quantité directement identifiable par une personne ayant des connaissances thématiques sur ce qui est modélisé.

Dans le modèle de Schelling (section 3.1.3), le paramètre S est un paramètre empirique : la valeur qui lui est donnée peut être interprétée empiriquement, quand bien même il serait difficile d'obtenir, par exemple d'un enquêté, une valeur exacte à utiliser.

? le dir autrement & w compléter

h expliquer

Paramètres techniques Il s'agit des paramètres dont la valeur, ou même son /ordre de grandeur, n'apportent aucune connaissance thématique au modélisateur. Ces paramètres ont pour raison d'être de permettre à d'autres types de paramètres de s'exprimer en valeurs compréhensibles et exploitables. Dès lors, leurs valeurs sont propres à chaque modèle formalisé, et une comparaison de ces valeurs entre les différents modèles n'apporte pas de connaissance.

Sur ...

claireté fois ?

Dans le modèle gravitaire (section 3.1.3), k est un paramètre technique : c'est un ratio qui varie afin d'ajuster l'ordre de grandeur des masses (M_i et M_j) à celui des flux échangés (F_{ij}). C'est une valeur d'ajustement, qui ne peut être comparée entre des modèles portant sur différents espaces et/ou thématiques⁵.

Paramètres « commensurable » Cette catégorie, forgée dans le cadre de cet ouvrage, permet de représenter les paramètres qui ne

3. La définition du dictionnaire « Le Petit Robert » pour ce terme est celle-ci : « Se dit d'une grandeur qui a, avec une autre grandeur, une commune mesure. » Petit Robert, 2006. Cf. aussi : « Les mondes n'apparaissent commensurables ou incommensurables qu'à ceux qui s'attachent aux mesures mesurées. Or, toutes les mesures, en science dure comme en science souple, sont aussi des mesures mesurantes et celles-là construisent une commensurabilité qui n'existe pas avant leur mise au point. » p.153 in Latour B., 1997, Nous n'avons jamais été modernes : essai d'anthropologie symétrique, Paris, La Découverte, 207 p.

4. Cf. sans doute le chapitre 1, à faire.

sont pas fondamentalement issus de l'empirique mais qui peuvent toutefois être interprétés, que ce soit par rapport à une valeur canonique, définissant alors un repère, ou alors par rapport à d'autres valeurs de ce même paramètre, créant dès lors une échelle d'analyse. Outre la calibration d'un modèle, les valeurs de ces paramètres ont donc un intérêt propre et peuvent être analysées pour elles-mêmes.

rev.

Dans le modèle gravitaire (section 3.1.3), le frein de la distance (paramètre α) peut être considéré comme un paramètre commensurable. Sa valeur ne renvoie ainsi à aucune dimension empirique, mais son usage répandu permet de comparer les valeurs qu'il peut prendre à des valeurs bien connues. Par exemple, on utilise fréquemment des valeurs de α comprises entre 1 et 2 pour exprimer la friction de la distance dans les navettes domicile-travail. Une valeur inférieure à 1 peut donc être interprétée comme laissant une part plus faible à la distance, par exemple dans le cas de déplacements de loisirs. A contrario, un paramètre α supérieur à 2 montrera une plus forte friction de la distance, par exemple dans le choix d'une école élémentaire à proximité du foyer d'un ménage.

5. Dans le cas où α vaut 0, ce paramètre peut être vu comme commensurable dans la mesure où il devient alors la part de la masse qui devient un flux.

Comment choisir les valeurs de paramètres ?

Ici, faire un court paragraphe de transition pour présenter rapidement l'approche générale du paramétrage : peut-être redondant avec ce qui suit, en particulier dans les présentations des paramétrages du modèle gravitaire et de Schelling.

3.1.2 Le paramétrage, un processus d'amélioration du modèle.

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 s'approcher autant que possible des données empiriques devant être reproduites en faisant varier les valeurs des paramètres jusqu'à ce qu'une combinaison de celles-ci soit satisfaisante.

reformuler

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.

Gilbert et Troitzsch (2005) restent très flous en utilisant les termes de *checking* et de *debugging*, qu'ils inscrivent dans une démarche plus large de *verification* correspondant à la validation interne. Osman Balci (Balci, 1994) crée et emploie l'acronyme « VV&T »⁶, à portée plus large puisqu'il s'applique à chacune des étapes de la modélisation.

6. Validation, Verification and Testing

?
une ?

définir ?
5 + hank ?

idem ?

Le plus souvent (ref Seb, Clem...), une fois le modèle construit, le modéliste s'attache à son « calibrage », en cherchant pour chaque paramètre la ou les valeurs qui permettront au modèle de s'approcher des données empiriques devant être reproduites. Cette étape, que l'on nomme généralement calibration, peut se faire de manière manuelle, par approximations successives (C. Cottineau et/ou S. Rey, d'après Hermann), par semi-automatisme, exemple en effectuant des analyses de sensibilité (C. Schmitt, J. Hirtzel), ou encore de manière entièrement automatique (C. Schmitt, S. Rey, C. Cottineau avec PSE). L'approche souvent défendue, notamment dans les travaux les plus récents (Rimbault ?), voudrait que cette étape soit obligatoire pour toute modélisation, permettant par une exploration systématique de comprendre l'entièreté du comportement d'un modèle, et d'en faire dès lors un outil complètement maîtrisé rendant possible l'établissement d'une loi.

Lena : trouver d'autres exemples plus classiques et lointains

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.

Lena : expliciter

Lena : penser à bien l'introduire avant.

3.1.3 Qu'est-ce que le(s) paramétrage(s) ?

Définition

Paramétrier : « Programmer (un appareil complexe), en définissant les paramètres assurant son fonctionnement optimal » (Petit Robert, 2006).

J'aime bien cette définition, c'est la seule à avoir le sens que j'entends dans paramétrage/paramétrer. Et elle introduit une logique d'optimalité.

Le terme 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 préférerons le terme de configuration de paramètres.

+ + Amsterdam,
+ l'éco. de Marseille
en France

On tend à distinguer le paramétrage – passage obligé ne nécessitant pas d'être évoqué – de la calibration, 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'évaluation du modèle, mais comme une composante inhérente à la construction d'un modèle, quelque 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, 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 sa calibration plus complexe et instable : « Au dessous du seuil des 150 pas de simulation

précise à ?
sinon dire
simplement à quoi
cor. le paramétrage

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 soucis 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]

Afin d'illustrer ces propos, on peut s'appuyer sur des modèles bien connus de la littérature, issus de deux champs disciplinaires différents.

Illustration : le modèle gravitaire.

*un ble ?
Il ne l'est
pas vraiment
c'est la
valence qui
parle pour les ch'tis.
Mais les figures sont
pas mal non ?*

Description Ce modèle statistique⁷, dans la formulation qu'en a faite Stewart (1948), vise à prédire des flux (démographiques, marchands etc.) potentiels entre des lieux à partir d'une analogie avec la loi physique de la gravitation. Dans le formalisme le plus simple, « sans contrainte » (Pumain et Saint-Julien 2001), on peut l'exprimer ainsi :

$$F_{ij} = \frac{k \times M_i \times M_j}{d_{ij}^\alpha}$$

Ce modèle présente trois valeurs empiriques (M_i , M_j et d_{ij}) et deux paramètres (k et α). k est un paramètre technique (encadré 3.2), puisque bien que basée sur l'empirique, sa valeur – et les fluctuations de celle-ci – ne donne aucune information sur le système étudié. α est un paramètre commensurable pour les raisons explicitées dans l'encadré 3.2.

À partir des valeurs empiriques M_i et M_j qui caractérisent les masses (populations, ou stocks de marchandise par exemple) des lieux i et j , et de la distance qui les sépare d_{ij} , on peut ainsi prédire F_{ij} , le flux entre ces lieux.

pas le seul but !

Objectif Afin que ce modèle donne des ordres de grandeur réalistes quand aux quantités échangées, il faut le paramétrier en définissant une valeur de k , permettant dès lors d'obtenir un rapport entre les masses d'origine et les quantités échangées. La valeur de ce paramètre est conditionnée par celle de α , que l'on nomme fréquemment « frein de la distance », en ce qu'il permet de quantifier l'impact qu'aura un éloignement plus ou moins important sur la quantité de flux échangé.

Paramétrage Pour « ajuster »⁸ le modèle, il faut donc en réaliser le paramétrage en s'appuyant sur les éléments connus de l'équation (les valeurs empiriques). α et k étant liés dans l'équation, la valeur de chacun de ces paramètres dépend de celle choisie pour l'autre, et un changement dans l'un des paramètres entraînera la nécessité de modifier l'autre.

abus de langage

Le paramétrage le plus simple consiste à réaliser une régression linéaire sur les logarithmes décimaux des distances et des flux observés, k prenant alors la valeur du logarithme de l'ordonnée à l'origine et α celle du coefficient

7. Pour une description plus complète, se référer à l'ouvrage de Pumain et Saint-Julien, 2001.

8. C'est le terme employé par (Pumain et Saint-Julien, 2001), que l'on peut ici retenir comme équivalent de paramétrier.

La technique de(voir) et aussi de calibrer

la + simple pour paramétrer ce modèle.

a ton sens. Mais n'y a-t-il pas diff entre créer (ajouter) un env et param. et ensuite estimer leurs valeurs ?

directeur de la courbe. Le modèle est alors calibré (voir encadré 3.3), mais si l'on souhaite donner une valeur spécifique à l'un des paramètres, par exemple pour utiliser une valeur classique de 2 au frein de la distance (α), il faudra alors le re-soumettre à paramétrage pour adapter k . De même si l'on modifie les lieux et/ou les masses sur lesquels il s'applique.

Illustration : le modèle de ségrégation de Schelling.

Description Le modèle de Schelling⁹ est un modèle de simulation décrit par Thomas Schelling (Schelling, 1971) qui vise à montrer comment un espace peut passer d'un état intégré à un état ségrégué ethniquement à travers une succession de comportements individuels de mobilité résidentielle. Il montre en particulier qu'on peut parvenir à un état ségrégué même quand les comportements individuels sont majoritairement tolérants.

L'espace du modèle est défini comme une grille carrée composée de N^{10} cellules. Au début de la simulation, des agents de deux types, en proportions similaires, sont distribués aléatoirement dans l'espace du modèle. Chaque agent occupe une cellule, et le nombre total de cellules occupées (et donc d'agents) dépend d'un paramètre de densité d qui représente la part (entre 0% et 100%) du nombre de cellule qui sera occupé¹¹. Chaque agent est défini par une satisfaction. Celle-ci correspond à la part de cellules voisines occupées par des agents « étrangers », c'est-à-dire d'un autre type : plus le voisinage contient d'étrangers, moins l'agent est satisfait. Si cette satisfaction est inférieure à un paramètre de tolérance S , l'agent est considéré comme insatisfait et se déplace aléatoirement dans une cellule non occupée. À terme, Schelling montre que même en considérant des valeurs de S assez élevées, c'est-à-dire un comportement plutôt tolérant, la succession de choix individuels entraîne la mise en place d'une distribution spatiale très ségrégée. On considère que le modèle a convergé quand l'ensemble des agents sont satisfaits, ce que ne permettent pas toutes les combinaisons de valeurs de paramètres.

Dans ce modèle, on peut donc dénoter trois paramètres, N , d et S . Les deux premiers sont des paramètres techniques (encadré 3.2), en ce que, purement théoriques, ils ne se réfèrent à aucune quantité empirique, ni ne peuvent être utilisés afin de se référer à un ordre de grandeur connu qui en ferait des paramètres commensurables. S est un paramètre empirique. Bien que le modèle soit théorique, ce paramètre se réfère toutefois à une comportement ayant un sens empirique fort.

Objectif Le paramétrage de ce modèle consiste à fixer un N constant et à faire varier d et S afin d'en trouver des valeurs permettant la convergence vers une situation stable. Quand d est très faible (30% dans l'exemple de Daudé et Langlois, 2006), quelles que soient les valeurs de S , le modèle converge rapidement : quand l'espace du modèle dispose d'une faible densité d'agents, il est facile pour ceux-là de créer des agrégats homogènes distants d'agrégats de l'autre type d'agents. Quand d est plus important ($\geq 66\%$), l'espace disponible

9. Une description plus poussée accompagnée d'une description de l'exploration du modèle peut être lue dans Daudé et Langlois, 2006

10. Nous reprenons ici la notation $M(N, d, n, S)$ proposée par Daudé et Langlois, 2006, p.433 en n'explicitant toutefois pas le paramètre n décrivant le type de voisinage (4 ou 8) utilisé.

11. Il y a donc $N \times d$ agents, et donc $\frac{N \times d}{2}$ agents de chaque type.

étant limité, toutes les valeurs de S ne permettent pas la convergence du modèle. Le paramétrage du modèle aura donc pour objectif de trouver la valeur maximale possible – pour un N et un d donné – que peut prendre le paramètre de tolérance S tout en laissant le modèle converger.

Paramétrage Contrairement au modèle gravitaire, le modèle de Schelling est stochastique : les agents se déplacent aléatoirement quand non satisfaits. Dès lors, deux exécutions du modèle avec le même jeu de paramètres n'entraineront pas forcément la même configuration spatiale. Qui plus est, pour certains jeux de paramètres, seules certaines exécutions convergeront. Le paramétrage de ce modèle ressemble donc à celui qui est réalisé pour le modèle SimFeodal. La manière « traditionnelle »¹² consiste à fixer un d , puis à essayer d'augmenter le S tant que le modèle converge. Si le modèle converge pour tout S , on augmente la valeur de d et on recommence à chercher la valeur maximale possible pour S . De part la nature stochastique du modèle, chaque jeu de paramètres doit être simulé plusieurs fois, le nombre de ces réplications dépendant de la part d'aléa dans le comportement du modèle. Un modèle de Schelling calibré, c'est-à-dire dont le paramétrage est achevé, donnera pour un N fixé les valeurs maximales de d et de S atteignables.

A travers ces deux exemples ayant traits à des méthodes différentes, on retrouve deux approches de paramétrage que l'on souhaite ici voire confondues : dans le premier cas, le paramétrage du modèle gravitaire tend à sa calibration, à la recherche d'une solution optimale, c'est-à-dire meilleure que toute autre. Pour un phénomène donné (un jeu de données précis par exemple) et un formalisme donné (l'expression la plus simple du modèle, ici sans contrainte), seule un couple de paramètres α et k permet ainsi d'obtenir des flux modélisés proches des flux observés. Dans le cas du modèle de Schelling, il n'y a pas de recherche d'optimum : le paramétrage permet de comprendre le modèle en définissant les limites en matière de convergence. Une configuration de paramètres S et d ne sera pas meilleure qu'une autre, mais pour un S donné, on saura quel est la valeur maximale de d permettant cette convergence (et réciproquement).

Paramétrier un ou des modèles ?

Cottineau, Reuillon et Rey ont montré qu'on construisait plus souvent des familles de modèles que des modèles uniques. Ces auteurs plaident pour une construction modulaire de modèles, chacun des mécanismes et complexifications devant être autonomes afin de pouvoir les assembler en autant de combinaisons que nécessaire à une exploration d'ensemble de leurs interactions.

Nous reprendrons de leur discours une vision moins ambitieuse, en considérant (ref Rey) simplement qu'en fait d'un modèle, la modélisation passe par plusieurs modèles, ayant chacun différentes implémentations d'un mécanisme conceptuellement identique et chacun se devant dès lors d'être adapté par un paramétrage dédié.

Pour illustrer ce propos, on peut prendre l'exemple d'une modification qui a eu lieu sur notre modèle au cours de son développement. On considérait jusque

12. C'est-à-dire manuelle, au contraire de méthodes automatiques plus rigoureuses et récentes.

Pourquoi le seul objectif serait-il de converger ? Si oui, intér. + expliquons, et dire à quoi cela sert.

Le final ou à l'issu de l'itération.

A quoi servent alors le param. priétés ?

Objectif particulier dans la littérature ? J'en ps l'impose.

influence ?

?

Lena : trop endogame, prendre exemples modèles modulaires dans JASSS

à chercher il en parlait à sa soutenance je crois

à moins amener sinon cela parait trop basé... A priori pas d'ancien temps-lors. Et aussi le résultat du modèle qui sur avance de changement d'estimation ?

Chapitre 3

3.1. PARAMÉTRER ? QUOI ET QUAND ?

là que pour prendre en compte la population rurale (hors Tours donc), 1000 foyers paysans donnaient une bonne idée de la situation en 800 pour l'espace considéré. Lors d'une réunion, les thématiciens se sont rendus compte que ce nombre sous-estimait très largement la population réelle, et qu'il valait mieux passer ce paramètre à 4000 foyers paysans. Les autres paramètres, fixés en bonne partie empiriquement, n'avaient visiblement aucune raison d'évoluer du fait de ce changement. Les mécanismes et effets de seuil étaient en effet conçus de manière à être relatifs aux masses manipulées, et le changement attendu était une augmentation linéaire des indicateurs de sortie.

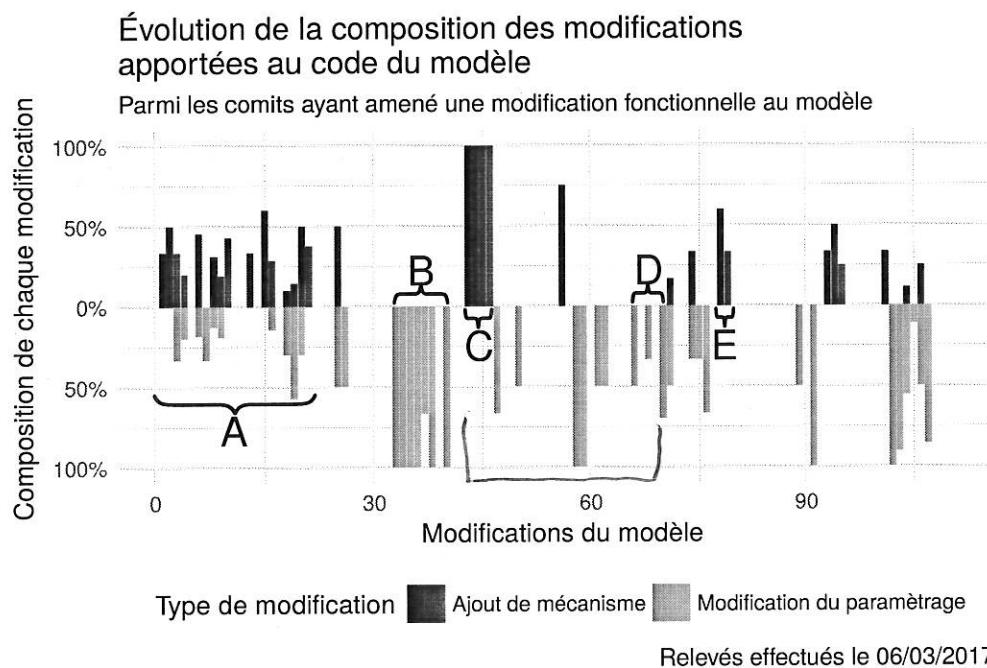


FIGURE 3.2 – Temporalité du paramétrage du modèle.

Chaque « enregistrement » (*comit*) correspond à une version modifiée et sauvegardée du modèle, contenant un ou plusieurs changements. On a ici rapporté le nombre de changements de chaque type (modification de mécanisme ou de valeur de paramètre) au nombre total de changements de chaque *comit* afin de figurer l'évolution des types de modification au cours de la construction du modèle.

Dans les faits, cette légère modification a entraîné une obligation de repenser la quasi-totalité des autres paramètres et d'ajuster une bonne part des règles. La concentration des foyers paysans était en effet bien trop rapide avec autant d'individus, et le modèle convergeait en quelques pas de temps vers une configuration presque statique et très concentrée. Tous les mécanismes de régulation de ce comportement étaient dès lors rendus ineptes, et il a fallu changer en profondeur la manière dont chaque paramètre et mécanisme interagissait avec les autres. C'est un changement majeur, que l'on peut constater dans la figure 3.2 (1ère modification de la période B), ayant donc entraîné des adaptations de tous les plans du modèle (paramètres en B, mécanismes en C). La première version avait été paramétrée aussi correctement que possible, mais ce paramétrage était entièrement à refaire avec la nouvelle version (modifications de la période B, fig. 3.2). On pourrait dès lors différencier le modèle

Lena : Là on ne suit pas, mais ce sera explicité dans le chap. 2

Lena : à discuter plus profondément

pré-existant et celui qui a suivi ce changement, et les considérer comme deux modèles puisque réagissant de manière extrêmement différente.

Cet exemple renforce le caractère nécessaire du paramétrage, et qui plus est, de la continuité de cette étape, qui ne peut être pensée que comme une calibration finale d'un modèle abouti, auquel on ne peut en fait parvenir que par des paramétrages réguliers de modèles successifs moins aboutis.

Lena : Pas clair et pas assez logique

Désambiguisation

De nombreux termes sont utilisés dans la littérature, souvent sans réelle distinction, pour désigner cette opération qui consister à choisir un jeu de paramètres pour un modèle. Pêle-mêle, on y retrouve le paramétrage, la validation, l'évaluation ou encore la calibration. Nous définissons dans l'encadré 3.3 le sens donné à chacun de ces termes dans le cadre de ce manuscrit.

Encadré 3.3 : Calibration, évaluation, validation...

Calibration

On réserve souvent, et nous nous y tiendrons, ce terme à la dernière étape dans l'aboutissement d'un modèle. Une fois les mécanismes fixés et des objectifs définis, on peut procéder à la calibration, c'est-à-dire à une exploration de l'espace des paramètres ayant pour but de stabiliser les paramètres afin de se rapprocher autant que possible de ces objectifs. Cette étape, quelques soient les moyens employés, s'approche de la résolution utilisée dans le cadre de systèmes d'équations. Le contexte des systèmes complexes, et donc d'une non-linéarité des effets des paramètres, rend toutefois difficile l'obtention d'une unique configuration de paramètres optimale¹², et la calibration aboutit donc souvent à un ensemble de configurations possibles, constituant par exemple un optimum de Pareto (Trouver ref.).

Évaluation

On emploie majoritairement ce mot pour décrire les méthodes permettant de comprendre le comportement du modèle et sa réaction aux différents paramètres. Là où la calibration cherche une configuration de paramètres optimale, l'évaluation tend surtout à caractériser la stabilité du modèle face à l'aléa ou à des configurations exceptionnelles. On vise ainsi à s'assurer que le modèle reproduise les faits stylisés¹² voulus quelque soient son réglage, ou au moins à quantifier les intervalles de paramètres qui y satisfont.

Validation

La validation (refs Seb) tire son origine de sciences plus nomothétiques, et correspond donc à la démonstration qu'un modèle reproduit correctement ce qu'il représente. Au delà de l'évaluation, le terme amène une logique de preuve formelle que le modèle réagit bien ainsi et pas autrement quelque soient les conditions d'exécution. Cette démonstration formelle peut être effectuée sur des modèles à faible nombre de paramètres et mécanismes, mais le terme n'est jamais (à vérifier) employé dès lors que les modèles se complexifient.¹²

l'en avant
de être dit
plus haut.

Pensez les étapes
et espérez où ?

défini à quel
moment ?

? exploration ?

c'est + large
que cela

) pas clair

- 12. Lena : il faut partir d'une def plus basique en intro
- 12. Lena : pas introduits avant
- 12. Lena : cf. JASSS : Faire un rappel du sens classique, par ex. en stats.

3.1.4 Comment paramétrier ?

Visual validation

(Trouver ref dans thèse Clémentine, sans doute Hermann encore)

Indicateurs

L'importance de la réplication

3.2 Premiers paramétrages

A changer : le chap 2 présentera la version 0 (Base) du modèle, et on vient donc ici montrer comment on a répondu aux limites identifiées à la fin du chapitre, cf. les conclusions/résultats du chapitre TransMonDyn.

Comme préconisé ci-dessus, le modèle SimFeodal a été paramétré tout au long de sa construction. Dans cette partie, nous reviendrons en détail sur chacune des étapes de ce paramétrage, en retracant les choix et questionnements qui ont orienté cette évolution. L'énumération est en ordre chronologique, et l'on a essayé d'organiser ces étapes en grands blocs-type de paramétrisation. Pour autant, le paramétrage et la construction d'un modèle, et du notre en particulier, sont un travail constant d'allers-retours entre l'identification et la résolution de problèmes. La structure réelle de développement, que l'on peut retrouver dans l'historique des modifications du modèle (les « *commits* »), ne correspond donc pas exactement à la chronologie organisée qui suit.

Le tableau suivant (table 3.1) se veut être un point de repère autant qu'une table des matières synthétique pour la compréhension des étapes qui suivent. Si la description de chacune de celles-ci revêt une forme assez factuelle et systématique, nous pensons toutefois que la précision de chacune permet de saisir la démarche proposée ainsi que l'intérêt de celle-ci.

Notons tout de même que l'étape 0 ici présentée correspond à une première version « exploitable » du modèle, c'est-à-dire contenant *a minima* l'ensemble des agents et mécanismes identifiés dans un premier temps par l'équipe de modélisation (Tannier et al., 2014 in Phan). Nous ne détaillerons donc pas ici les différentes phases de construction précédentes. Nous reviendrons toutefois sur plusieurs d'entre elles, afin d'illustrer nos propos, dans les chapitres 6 et 7 (Partie 3).

TABLE 3.1 – Tableau récapitulatif des étapes de construction du modèle.
Conception : C. Tannier (Tannier 2017, HDR) et R. Cura, 2017

Étape	Version	Type	Modifications	Section
0	Base	Création	Création du modèle. Cf. tableau 14 de Tannier 2017	Chap 2 ¹³
1	Base2	paramétrage	Hiérarchisation des valeurs d'attraction des attracteurs. Modification des critères de création de nouvelles paroisses. Réduction de la distance maximale de déplacement local. Assouplissement des règles de promotion d'un château en gros château.	Chap3-x.y
2	Base2 compo5 2bis	modélisation et paramétrage	Assouplissement des règles de création d'une paroisse. Réduction de la distance maximale de déplacement local. Modification des règles de calcul de la probabilité de construire un château. Modification du calcul de la satisfaction matérielle des foyers paysans.	Chap3-x.y
3	Base2 compo5 2quat	modélisation et paramétrage	Hiérarchisation des valeurs d'attraction des attracteurs. Facilitation de la construction de châteaux. Rationalisation du déplacement des foyers paysans. Ajout d'un nouveau type d'attracteur, les communautés.	Chap3-x.y
4	Base3 2	modélisation et paramétrage	Hiérarchisation des valeurs d'attraction des attracteurs. Modification des critères de création de nouvelles paroisses. Augmentation de la distance maximale de déplacement local. Modification de la procédure d'identification des agrégats et des pôles.	Chap3-x.y
5	Base4 1	modélisation et paramétrage	Modification de la répartition initiale des foyers paysans. Simplification de la procédure d'identification d'héritage des agrégats. Amélioration de la définition des pôles d'attraction et de leur enveloppe.	Chap3-x.y
6	Base4 2	modélisation et paramétrage	Hiérarchisation des valeurs d'attraction des attracteurs. Modification de l'ordonnancement des actions dans le modèle. Modification de la procédure d'identification des agrégats.	Chap3-x.y
7	Base4 3ter	modélisation et paramétrage	Dynamisation de la distance maximale de déplacement local. Modification du calcul de satisfaction des foyers paysans.	Chap3-x.y
8	Base4 4A	modélisation	Modification de l'ordonnancement des actions dans le modèle. Modification du mécanisme de déplacement local des foyers paysans.	Chap3-x.y

3.2.1 Schémas des étapes de paramétrage

3.2.2 Étape 0

Indicateurs généraux

13. On s'appuie sur cette version « Base » dans le chapitre 2 en la comparant à la version actuelle (étape 8), mais une vision plus exhaustive et linéaire en est donnée dans le chapitre d'ouvrage Cura et al., 2017, PUFR.

difficulté de conception
sous le point de vue du modèle.

TABLE 3.2 – Indicateurs synthétiques de l'étape 1

Indice	Objectif	Moyenne
Nb agrégats	200.00	145.30
Nb châteaux	50.00	69.78
Nb gros Châteaux	10.00	9.60
Nb egl. Par.	300.00	167.90
Dist entre égl.	3000.00	2944.00
FP isolés	0.20	0.49
Ratio Charge Fisc.	3.00	5.15

3.2.3 Étape 2

3.2.4 Étape 3

3.2.5 Étape 4

3.2.6 Étape 5

3.2.7 Étape 6

3.2.8 Étape 7

3.2.9 Étape 8

3.2.10 Différences finales entre version « Base » (étape 0) et version « finale » (étape 8)

3.3 Comment traiter les sorties du modèle ?

3.3.1 Variabilité

3.3.2 Nombre de sortie

3.3.3 Masse des sorties