# Paramétrer un modèle dans un contexte de co-construction interdisciplinaire

Version 2017-12-11

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 (). 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 () 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 ([enc:types-parametres])) lui permettant de mieux convenir aux objectifs fixés, soient-ils en terme de comportements attendus ou d’objectifs quantitatifs.

## Paramétrer ? Quoi et quand ?

Avant de préciser le sens et le processus de 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.

Au plus général, le nouveau petit Robert définit un paramètre en ces mots :

1. math. Quantité à fixer librement, maintenue constante, dont dépend une fonction de variables indépendantes, une équation ou une expression mathématique. — Variable en fonction de laquelle on exprime chacune des variables d’une équation.  
2. fig. et didact. Élément important dont la connaissance explicite les caractéristiques essentielles de l’ensemble d’une question.  
3. par ext. Élément nécessaire pour juger, évaluer, comprendre (qqch.).  
  (Robert, Rey-Debove, and Rey 1993 **Paramètre**)

Seule la première définition correspond au sens large que l’on attend ici, mais elle est très généraliste, bien que ne correspondant pas pour autant à tous les usages du terme employés dans la littérature.

##### 

**L’acceptation mathématiques** d’un paramètre est sans doute celle qui souffre le moins d’ambiguïté : il s’agit des termes fixes d’une équation simple, par opposition aux variables qui en constituent les éléments qui seront amenés à évoluer. Par exemple, dans la formulation d’une fonction affine, , la variable résultante de dépend de la variable et des paramètres fixés et . Quelles que soient les valeurs empruntées par , ces paramètres demeurent constants. Dans un système d’équations plus complexes, par exemple le modèle de croissance logistique de population[[1]](#footnote-24), l’accroissement de population au cours du temps () dépend de la variable , population à cet instant, ainsi que de deux paramètres, , le taux de croissance, et , la capacité d’accueil, c’est-à-dire un potentiel maximum de population vers laquelle doit tendre le système modélisé. La définition mathématique est donc assez universelle et convient à la quasi-totalité des systèmes d’équations[[2]](#footnote-25).

##### 

**En statistiques**, quand bien même cette discipline fait un large usage des formalismes mathématiques, les paramètres recouvrent un ensemble assez différent : il s’agit d’une grandeur mesurable qui permet de présenter de façon plus simple, plus abrégée les caractéristiques essentielles d’un ensemble statistique. . Ainsi, pour les statisticiens, la moyenne, l’écart-type ou encore le coefficient d’asymétrie sont des paramètres. On peut toutefois y retrouver une logique commune avec les paramètres mathématiques quand on décrit une loi statistique avec ces valeurs, qui deviennent alors les éléments permettant de caractériser une variable. Ainsi, pour formaliser une distribution normale décrivant la variable , on fera appel aux paramètres de cette loi que sont la moyenne () et l’écart-type (): .

##### 

**Dans le domaine informatique**, la définition est bien plus floue et incluante que dans les éléments précédents. La faute à une hétérogénéité bien supérieure dans les pratiques et construits informatiques. Comme dans les fonctions mathématiques, les paramètres sont les arguments des fonctions informatiques. Mais ces fonctions recouvrent un ensemble extrêmement vaste, depuis une fonction mathématique basique telle que l’arrondi (round() dans sa version la plus courante) jusqu’à des fonctions plus complexes telles que la conversion d’une image d’un format à un autre[[3]](#footnote-28). Dans le cas de la fonction arrondi, le paramètre[[4]](#footnote-29) est le même qu’en mathématiques : une valeur fixée à l’exécution du programme qui renverra une variable. Dans le second cas, l’analogie est peu convaincante : les paramètres seront par exemple le chemin de stockage de l’image à convertir, ou encore le format de sortie attendu.

### Qu’est-ce qu’un paramètre ?

#### En mathématiques

#### En physique

#### En simulation informatique

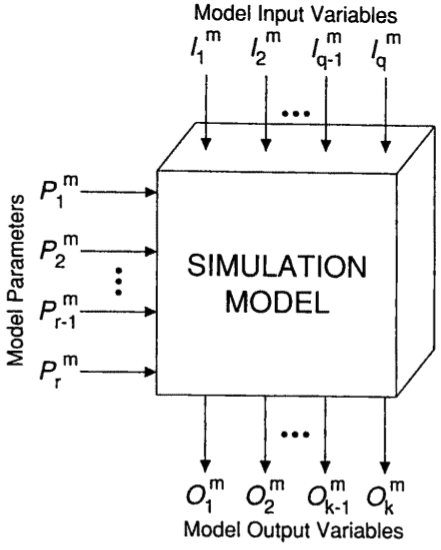
On précise ici simulation car le terme de paramètre a encore une définition usuelle différente en informatique généraliste : il renvoie ainsi aux arguments (synonyme de paramètre ici) d’une fonction informatique. Les fonctions informatiques recouvrant un ensemble extrêmement vaste, par exemple depuis une fonction mathématique telle que factorielle() () jusqu’à des fonctions plus complexes telles que la conversion d’une image d’un format à un autre.

### Les paramètres dans les modèles agents

La seconde partie de la définition renvoie à un usage plus mathématique, et nous nous inscrivons donc dans la lignée de la première partie de la définition. Celle-ci est suffisamment généraliste pour s’appliquer aussi bien au champ mathématique qu’à celui de l’informatique d’une part, et évite d’autre part les différences sémantiques que l’on retrouve souvent entre paramètre et variable dans le champs mathématique, par exemple dans l’acceptation différente qui est faite d’un paramètre selon que le modèle soit paramétrique ou non. Par ailleurs, cette définition large permet de spécifier ce que l’on entend par ces termes ici, et donc d’en faire un usage précis, là où les ouvrages de référence en modélisation agent ne définissant jamais ce terme. Un ouvrage fait exception (Treuil, Drogoul, and Zucker 2008, 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 :

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, and Zucker 2008, 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:parametres-Balci]).



Les variables d’un modèle de simulation selon Balci (Balci 1994, 122)

Dans ce schéma et l’article associé, Balci

Nous considèrerons 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.

![Schématisation des définitions de variables (V), indicateurs de sortie (I) et paramètres (P).](data:application/pdf;base64,)

Schématisation des définitions de variables (V), indicateurs de sortie (I) et paramètres (P).

#### Paramètres et variables

Variables, entrées, paramètres…termes-variables 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[[5]](#footnote-38). Dans cet ouvrage, nous distinguerons plusieurs types de paramètres ([enc:types-parametres]) 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.

#### 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 sociales[[6]](#footnote-40) distingue souvent () deux types de paramètres : les paramètres empiriques, dont la valeur a une correspondance directe dans le domaine empirique (on peut donc l’observer), et les paramètres techniques[[7]](#footnote-41), pour lesquels cette correspondance n’existe pas. Comme dans la plupart des typologies ayant trait à la catégorisation de valeurs numériques, 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[[8]](#footnote-42), 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 Minsky[[9]](#footnote-43) (Minsky 1965). Pour différents modélisateurs, un même paramètre pourra donc être considéré comme empirique ou commensurable, ou encore comme commensurable ou technique.

Trois types de paramètrestypes-parametres

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 ([desc:schelling]), le paramètre 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.

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.

Dans le modèle gravitaire ([desc:gravitaire]), est un paramètre technique : c’est un ratio qui varie afin d’ajuster l’ordre de grandeur des masses ( et ) à celui des flux échangés (). C’est une valeur d’ajustement, qui ne peut être comparée entre des modèles portant sur différents espaces et/ou thématiques.

Paramètres commensurable

Cette catégorie, forgée dans le cadre de cet ouvrage, permet de représenter les paramètres qui ne 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.

Dans le modèle gravitaire ([desc:gravitaire]), le frein de la distance (paramètre ) peut être considéré comme un paramètre commensurable. Sa valeur ne renvoit 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 comprises entre et pour exprimer la friction de la distance dans les navettes domicile-travail. Une valeur inférieure à 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 supérieur à montrera une plus forte friction de la distance, par exemple dans les choix d’une école élémentaire à proximité du foyer d’un ménage.

#### Comment choisir les valeurs de paramètres ?

### 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.

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 (Gilbert and 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*[[10]](#footnote-46), à portée plus large puisqu’il s’applique à chacune des étapes de la modélisation 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 des données empiriques devant être reproduites. Cette étape, que l’on nomme généralement calibration, peut se faire de manière manuelle, par approximations successives (), par semi-automatisme, (), ou encore de manière entièrement automatique (). L’approche souvent défendue, notamment dans les travaux les plus récents (), 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 maitrisé .

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 est un exercice qui devrait s’effectuer tout au long de la construction du modèle, de manière plus itérative que conclusive.

### Qu’est-ce que le(s) paramétrage(s) ?

#### Définition

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.

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 (Schmitt 2014), qui a permis de réaliser que la variation de l’un des paramètres () 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 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, 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.[desc:gravitaire]

##### Description

Ce modèle statistique[[11]](#footnote-52), dans la formulation qu’en a faite Stewart (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 and Saint-Julien 2001), on peut l’exprimer ainsi :

Ce modèle présente trois valeurs empiriques (, et ) et deux paramètres ( et ). est un paramètre technique ([enc:types-parametres]), puisque bien que basée sur l’empirique, sa valeur – et les fluctuations de celle-ci – ne donne aucune information sur le système étudié. est un paramètre commensurable pour les raisons explicités dans l’[enc:types-parametres].

À partir des valeurs empiriques et qui caractérisent les masses (populations, ou stocks de marchandise par exemple) des lieux et , et de la distance qui les sépare , on peut ainsi prédire , le flux entre ces lieux.

##### Objectif

Afin que ce modèle donne des ordres de grandeur réalistes quand aux quantités échangées, il faut le paramétrer en définissant une valeur de , 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 , 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 échangés.

##### Paramétrage

Pour ajuster[[12]](#footnote-55) le modèle, il faut donc en réaliser le paramétrage en s’appuyant sur les éléments connus de l’équation (les valeurs empiriques). et 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.

Le paramétrage le plus simple consiste à réaliser une régression linéaire sur les logarithmes décimaux des distances et des flux observés, prenant alors la valeur du logarithme de l’ordonnée à l’origine et celle du coefficient directeur de la courbe. Le modèle est alors calibré (voir [enc:termes-calibration]), 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 (), il faudra alors le re-soumettre à paramétrage pour adapter . De même si l’on modifie les lieux et/ou les masses sur lesquels il s’applique.

#### Illustration : le modèle de ségrégation de Schelling.[desc:schelling]

##### Description

Le modèle de Schelling[[13]](#footnote-58) 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 ségrégé ethniquement à travers une succession de comportements individuels de mobilité résidentielle. Il montre en particulier qu’on peut parvenir à un état ségrégé même quand les comportements individuels sont majoritairement tolérants.

L’espace du modèle est défini comme une grille carrée composée de [[14]](#footnote-59) 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 qui représente la part (entre et ) du nombre de cellule qui sera occupé[[15]](#footnote-60). 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 , l’agent est considéré comme insatisfait et se déplace aléatoirement dans une cellule non occupée. A terme, Schelling montre que même en considérant des valeurs de 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, , et . Les deux premier sont des paramètres techniques ([enc:types-parametres]), 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. 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 constant et à faire varier et afin d’en trouver des valeurs permettant la convergence vers une situation stable. Quand est très faible ( dans l’exemple de (Daudé and Langlois 2006)), quelles que soient les valeurs de , 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 est plus important (), l’espace disponible étant limité, toutes les valeurs de ne permettent pas la convergence du modèle. Le paramétrage du modèle aura donc pour objectif de trouver la valeur maximale possible – pour un et un donné – que peut prendre le paramètre de tolérance 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’entraîneront 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[[16]](#footnote-63) consiste à fixer un , puis à essayer d’augmenter le tant que le modèle converge. Si le modèle converge pour tout , on augmente la valeur de et on recommence à chercher la valeur maximale possible pour . 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 fixé les valeurs maximales de et de 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 et 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 en définissant les limites en matière de convergence. Une configuration de paramètres et ne sera pas meilleure qu’une autre, mais pour un donné, on saura quel est la valeur maximale de permettant cette convergence (et réciproquement).

#### Paramétrer un ou des modèles ?

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 () 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 là que pour prendre en compte la population rurale (hors Tours donc), 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 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.

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. La première version avait été paramétré aussi correctement que possible, mais ce paramétrage était entièrement à refaire avec la nouvelle version (modifications de la période B, [fig:comits-periodes]).

#### Désambiguïsation

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’[enc:termes-calibration] le sens donné à chacun de ces termes dans le cadre de ce manuscrit.

Calibration, évaluation, validation…termes-calibration

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,[[17]](#footnote-67) et la calibration aboutit donc souvent à un ensemble de configurations possibles, constituant par exemple un optimum de Pareto ().

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 a des configurations exceptionnelles. On vise ainsi à s’assurer que le modèle reproduise[[18]](#footnote-68) 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.[[19]](#footnote-69)

### Comment paramétrer ?

#### Visual validation

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

#### Indicateurs

#### L’importance de la réplication

## Paramétrage et ajustement du modèle

##### 

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 retraçant 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:etapes-construction]) 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). 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 .

### Schéma d’ensemble des étapes de paramétrage

m.05 m.08 m.12 m.65 m.10

[table:etapes-construction]  
Étape & Version & Type & Modifications & Section  
Étape & Version & Type & Modifications & Section  
0 & Base & Création & Création du modèle. & Chap 2[[20]](#footnote-77)  
1 & Base2 & paramétrage & 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 & des valeurs d’attraction des attracteurs. Facilitation de la construction de châteaux. Rationnalisation du déplacement des foyers paysans. Ajout d’un nouveau type d’attracteur, les communautés. & Chap3-x.y  
4 & Base3 2 & modélisation et paramétrage & 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 & 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

### Étape 0

##### Indicateurs généraux

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 |

### Étape 2

### Étape 3

### Étape 4

### Étape 5

### Étape 6

### Étape 7

### Étape 8

### Un bilan des changements majeurs à l’issu du paramétrage

## Comment traiter les sorties du modèle ?

### Variabilité

### Nombre de sortie

### Masse des sorties

Balci, Osman. 1994. “Validation, Verification, and Testing Techniques Throughout the Life Cycle of a Simulation Study.” *Annals of Operations Research* 53 (1): 121–73. doi:[10.1007/BF02136828](https://doi.org/10.1007/BF02136828).

Daudé, Eric, and Patrice Langlois. 2006. “Comparaison de Trois Implémentations Du Modèle de Schelling.” In *Modélisation et Simulation Multi-Agents. Applications Pour Les Sciences de L’Homme et de La Société*, edited by Hermès, 411–41. <https://halshs.archives-ouvertes.fr/halshs-01082655>.

Durand-Dastès, François. 1995. “Les Modèles En Géographie.” In *Bailly A., Ferras R., Pumain D., Encyclopédie de La Géographie*, 2ème édition, 293–307. Chapitre 16.

Gilbert, Nigel, and Klaus Troitzsch. 2005. *Simulation for the Social Scientist*. McGraw-Hill Education (UK).

ImageMagick Studio, LLC. 2008. “ImageMagick.”

Latour, Bruno. 2013. *Nous N’avons Jamais été Modernes*. La découverte.

Mathian, Hélène, and Cécile Tannier. 2015. “Formalisation des interactions spatiales dans les modèles multi-agents - Essai d’analyse comparative.” *Revue Internationale de Géomatique* 25 (3): 393–414. doi:[10.3166/RIG.25.393-414](https://doi.org/10.3166/RIG.25.393-414).

Minsky, Marvin. 1965. “Matter, Mind and Models.”

Pumain, Denise, and Thérèse Saint-Julien. 2001. “Les Interactions Spatiales.” *Armand Colin, Paris*, 19–31.

Robert, Paul, Josette Rey-Debove, and Alain Rey. 1993. *Le Nouveau Petit Robert*. le Robert.

Schelling, Thomas C. 1971. “Dynamic Models of Segregation.” *Journal of Mathematical Sociology* 1 (2): 143–86.

Schmitt, Clara. 2014. “Modélisation de La Dynamique Des Systèmes de Peuplement: De SimpopLocal à SimpopNet.” Paris 1.

Stewart, John Q. 1948. “Demographic Gravitation: Evidence and Applications.” *Sociometry* 11 (1/2): 31–58.

Tannier, Cécile, Elisabeth Zadora-Rio, Samuel Leturcq, Xavier Rodier, and Elisabeth Lorans. 2014. “Une Ontologie Pour Décrire Les Transformations Du Système de Peuplement Européen Entre 800 et 1100.” In *Ontologies et Modélisation Par SMA En SHS*, by Denis Phan, Hermès-Lavoisier, 289–310. Paris.

Treuil, Jean-Pierre, Alexis Drogoul, and Jean-Daniel Zucker. 2008. “Modélisation et Simulation à Base d’agents.” *Dunod Editions*, 1–13.

Verhulst, Pierre-François. 1838. “Notice Sur La Loi Que La Population Suit Dans Son Accroissement. Correspondance Mathématique et Physique Publiée Par a.” *Quetelet* 10: 113–21.

1. d’après (Verhulst 1838) [↑](#footnote-ref-24)
2. A l’exclusion notable des systèmes d’équation paramétriques, où, à l’inverse, le terme de paramètre désigne alors les variables indépendantes. [↑](#footnote-ref-25)
3. Par exemple la fonction convert du logiciel ImageMagick (ImageMagick Studio 2008) [↑](#footnote-ref-28)
4. Où les paramètres, cette fonction acceptant souvent un second paramètre définissant le nombre de décimales auquel arrondir la valeur d’entrée. [↑](#footnote-ref-29)
5. Pour reprendre les mots de Treuil et al. qu’ils appliquent aux entrées(Treuil, Drogoul, and Zucker 2008, 8). [↑](#footnote-ref-38)
6. 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. [↑](#footnote-ref-40)
7. Mathian et Tannier (Mathian and 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 pan 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 [enc:termes-variables]. Nous préférerons donc ce terme de technique pour éviter les confusions. [↑](#footnote-ref-41)
8. 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.(Robert, Rey-Debove, and Rey 1993).  
   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’existait pas avant leur mise au point.(Latour 2013, 153) [↑](#footnote-ref-42)
9. [↑](#footnote-ref-43)
10. *Validation, Verification and Testing* [↑](#footnote-ref-46)
11. Pour une description plus complète, se référer à l’ouvrage de Pumain et Saint-Julien (Pumain and Saint-Julien 2001). [↑](#footnote-ref-52)
12. C’est le terme employé par (Pumain and Saint-Julien 2001), que l’on peut ici retenir comme équivalent de paramétrer. François Durand-Dastès (Durand-Dastès 1995, 298) utilise pour sa part le terme de calibration à propos du même modèle. [↑](#footnote-ref-55)
13. Une description plus poussée accompagnée d’une description de l’exploration du modèle peut être lue dans (Daudé and Langlois 2006) [↑](#footnote-ref-58)
14. Nous reprenons ici la notation proposée par (Daudé and Langlois 2006, 433) en n’explicitant toutefois pas le paramètre décrivant le type de voisinage (4 ou 8) utilisé. [↑](#footnote-ref-59)
15. Il y a donc agents, et donc agents de chaque type. [↑](#footnote-ref-60)
16. C’est-à-dire manuelle, au contraire de méthodes automatiques plus rigoureuses et récentes. [↑](#footnote-ref-63)
17. Lena : il faut partir d’une def plus basique en intro [↑](#footnote-ref-67)
18. Lena : pas introduits avant [↑](#footnote-ref-68)
19. Lena : cf. JASSS : Faire un rappel du sens classique, par ex. en stats. [↑](#footnote-ref-69)
20. On s’appuie sur cette version Base dans le . [↑](#footnote-ref-77)