

## Chapitre 3

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

Rendu le 02/02/2018

Version 2018-01-31

Le modèle, tel qu'il a été 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 à une version plus adaptée aux questions des thématiciens. Le descriptif technique de cette version « finale » se trouve dans l'annexe *n* (Ref à l'annexe contenant le descriptif technique de la dernière version du modèle.) Par paramétrage, nous entendons ici le processus visant à doter le modèle de paramètres (empiriques, « commensurables » et techniques (encadré 3.2)) lui permettant de mieux répondre aux objectifs fixés en termes de comportements attendus ou d'objectifs quantitatifs. Nous nous attacherons dans un premier temps à expliciter et spécifier ce sens.

### 3.1 Paramétrer ? Quoi et quand ?

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.

#### 3.1.1 Différents points de vue sur la définition d'un paramètre

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 et REY 1993, **Paramètre**)

### 3.1.5 Comment paramétrer ?

#### Visual validation

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

#### Indicateurs

#### L'importance de la réplication

## 3.2 Paramétrage et ajustement du modèle SimFeodal

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 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 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). 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).

### 3.2.1 SimFeodal : pourquoi et quoi paramétrer ?

Description des « résultats » du modèle à l'étape v0, (c'est-à-dire reprise et completion du chap 11 de TMD).

==> Point par point, qu'est-ce qu'on observe, qu'est-ce qui va et qu'est-ce qui ne va pas.

Pour transition avec : sur quels leviers jouer pour améliorer/paramétrer le modèle ?

Le modèle SimFeodal présenté dans le chapitre 2 correspond à une « version 0 » du modèle souhaité, c'est-à-dire qu'il en constitue une base présentant l'ensemble des mécanismes voulus. Pour autant, il ne répond pas nécessairement aux attentes que l'on peut en avoir. On peut caractériser ces dernières, de manière



très générale, comme permettant de reproduire les « faits stylisés »<sup>20</sup> observés ou analysés empiriquement. La reproduction d'un fait stylisé peut s'entendre de multiples manières, tout au long du gradient qualitatif-quantitatif. Dans notre cas, en l'absence d'une connaissance quantitative totale des phénomènes mobilisés et observés, on peut s'appuyer sur deux grands ensembles d'« indices » renseignant sur la reproduction des faits stylisés par le modèle : (1) des valeurs empiriques, quantifiables et quantifiées, renseignant sur la quantité de certains éléments à obtenir en fin de simulation (nombre de châteaux, taux de foyers paysans isolés etc.), et (2) des tendances générales dans le déroulement de la simulation, notamment temporelles : allure de la courbe d'évolution du nombre d'agréats, forme de la hiérarchie rang-taille de la composition des pôles ou agréats etc.

Dans le modèle de simulation, ces valeurs empiriques (1) et tendances (2) sont symbolisées sous la forme d'indicateurs de sortie (cf. 3.1.2), lesquels sont à analyser en tenant compte de la précision du fait stylisé qu'ils représentent. Il ne faudra ainsi pas étudier la croissance du nombre d'agréats au cours de la simulation de manière fine, par exemple en étudiant le coefficient directeur de la courbe, quand l'empirie ne donne quasiment aucune information à ce sujet si ce n'est qu'il y a bien plus d'agréats (cette quantité étant elle-même analysable avec plus de précision) en fin de période qu'au début. On aurait pu vouloir quantifier la précision de ces données, par exemple à l'aide des méthodes développées dans le champ des observations floues et/ou incertaines (voir par exemple le travail de Cyril de Runz sur les données « imparfaites » (DE RUNZ 2008)). Pour autant, devant l'importante hétérogénéité dans les niveaux de connaissance des faits mobilisés, la plupart de ceux-ci demeurant à l'état d'hypothèses émises qui plus est, nous avons choisi de ne pas chercher à quantifier cette incertitude en amont.

Nous en tenons compte lors de l'analyse des sorties du modèle, sans chercher cependant à hiérarchiser l'importance des indicateurs selon cet élément. La hiérarchisation est toutefois indispensable : dans le modèle de simulation, on génère près de 50 indicateurs, et en l'absence justement d'une quantifiabilité totale de ceux-ci qui aurait pu permettre d'en extraire un indicateur composite unique d'adéquation d'une simulation aux faits stylisés modélisés, il n'est dès lors pas possible d'examiner avec la même attention chacun de ces nombreux indicateurs de sortie.

Il nous faut dès lors hiérarchiser l'observation de ces indicateurs, ou autrement dit, trouver un moyen d'ordonner leur usage. On a vu qu'il n'était pas justifié de mener cet ordonnancement à partir des propriétés intrinsèques des indicateurs du modèle. Au contraire, et cela porte bien plus de sens vis-à-vis du rôle d'un modèle, la hiérarchisation des sorties du modèle doit suivre la hiérarchie implicite qui structure les hypothèses et objectifs du modèle en lui-même. Dans le cas de SimFeodal, on cherche ainsi à reproduire des éléments extrêmement hétérogènes, qu'il nous faut donc hiérarchiser.

Dans le chapitre précédent (ref chap 2), nous présentions les dynamiques principales : polarisation, hiérarchisation et fixation des foyers paysans. À partir de ces

20. Définis ainsi par (LIVET, PHAN et SANDERS 2014) : « Un « fait stylisé » est une présentation simplifiée (i.e. taux, ratio ou écart, structure spatiale) d'une régularité empirique sur l'observation de laquelle il y a un large accord. Le terme a été popularisé en économie par Nicholas Kaldor (1961). [Les] faits stylisés peuvent être construits de la manière suivante : 1) en partant du domaine empirique, on identifie des relations saillantes ; 2) on opère quelques simplifications qui permettent d'inclure formellement ces relations dans des modèles ; 3) une fois admis que ces simplifications ne faussent pas trop les choses, on érige ces relations à la fois simplificatrices et formalisables au rang de « faits stylisés », dont les concepts théoriques doivent rendre compte. »



dynamiques, reste à trouver des indicateurs permettant de les évaluer, indicateurs jouant dès lors le rôle de « proxy ».

Si ces dynamiques s'appliquent presque toutes sur plusieurs types d'agents du modèle, on retiendra là aussi les agents les plus impactés : la polarisation, par exemple, peut être observée depuis le point de vue de ce qui polarise (les attracteurs) tout autant que de ce qui est polarisé (les foyers paysans). Ici, par soucis de hiérarchie dans l'analyse, on examinera d'abord le résultat effectif de la polarisation, c'est-à-dire la concentration des foyers paysans en agrégats, avant d'observer la répartition et les diversité des attracteurs ayant entraîné ce phénomène.

#### Encadré 3.4 : Incrémentalité des indicateurs

De la même manière que les paramètres et mécanismes d'un modèle de simulation tendent à évoluer<sup>21</sup> au cours du temps de la construction, souvent afin de raffiner un comportement observé, les indicateurs de sortie sont amenés à évoluer aussi.

Ainsi, en cas de modifications fines du modèle, il est fréquent que les indicateurs initialement choisis ne suffisent plus à départager des versions du modèle quant à un phénomène spécifique. Par exemple, quand on observe le phénomène de polarisation dans les sorties de SimFeodal, l'indicateur du nombre d'agrégats est extrêmement synthétique et informatif jusqu'à ce que l'objectif soit atteint ou que les modifications ne parviennent plus à le faire évoluer. À partir de ce moment, afin d'améliorer la vraisemblance du modèle, on peut se focaliser sur la distribution spatiale de ces agrégats, par exemple pour vérifier qu'ils sont bien répartis de manière homogène dans l'espace, et non trop concentrés.

L'observation de la répartition spatiale requiert certes de nouvelles analyses, mais surtout, par exemple, d'enregistrer les positions des agrégats au cours du temps. Si cet indicateur de sortie n'était pas utile avant cela, il n'y avait aucun intérêt à l'enregistrer. Il faut donc adapter le modèle pour enregistrer ce nouvel indicateur. Dès lors, on pourra composer un nouvel indicateur synthétique (dans cet exemple, cela pourrait être un indice de concentration spatiale).

Ce procédé incrémental dans la construction des indicateurs est très fréquent, mais pose toutefois un problème majeur : sauf à adapter chacune des anciennes version du modèle pour y ajouter l'enregistrement des nouveaux indicateurs nécessaires, on ne pourra rendre strictement comparable les sorties de tous les modèles. Et même alors, il faudrait alors ré-exécuter des répliques de chaque version du modèle à chaque ajout d'indicateur, quand bien même les indicateurs présent initialement étaient jugés suffisants. Un dernier obstacle est plus gênant : certains indicateurs sont spécifiques à des mécanismes, et en cas de changement de ces derniers, ils peuvent ne plus être calculables ou simplement comparables. Par exemple, des versions antérieures du modèle enregistraient les comportements individuels des foyers paysans quant à leur « choix » de déplacement, selon qu'ils étaient à l'origine localisés dans un agrégat ou dispersés. Une simplification du modèle a abouti à la modification des règles différenciant les possibilités de déplacement : on n'observe plus si le foyer paysan est dans un agrégat, mais plutôt s'il est dans un agrégat doté d'un pôle d'attraction. Dès lors, les analyses basées sur les choix de déplacement des foyers paysans selon leur origine ne sont plus comparables avec celles des versions antérieures au changement dans le modèle.

bon terme ?  
" de la situation  
simulée par le  
modèle

oui mais on ne  
change pas le  
modèle !

Il faut bien  
distinguer les  
2 niveaux -



Ces éléments expliquent que dans les résultats de chaque étape du paramétrage du modèle, on ne présente pas systématiquement l'ensemble des indicateurs, y compris quand ceux-ci pourraient être plus pertinents que les indicateurs présentés.

### Évaluer et paramétrer la polarisation

La polarisation des foyers paysans dans l'espace du modèle est sans doute le fait stylisé principal <sup>parmi</sup> ~~dans~~ ceux que l'on cherche à reproduire. Rappelons ici que l'objectif est de passer, au début de la simulation, d'une répartition dispersée des foyers paysans, à un état plus concentré, les foyers paysans devant, en fin de simulation, être en grande majorité rassemblés au sein d'agréats. Afin de préciser ce que l'on attend par cette concentration, on peut s'appuyer sur un certains nombres d'indicateurs simples, renseignant alors, tels des proxy, sur l'état de cette polarisation de l'habitat. } ces formules

**Taux de foyers paysans dispersés** Cet indicateur, et sa déclinaison temporelle, sont vraisemblablement les plus évidents : plus le taux de foyers paysans dispersés en fin de simulation est faible, plus le système de peuplement est polarisé. On a vu<sup>22</sup> qu'empiriquement, autour de 1160, on observe environ 20% de foyers paysans encore dispersés, alors qu'ils sont près de 95% au départ. Le modèle reproduira donc d'autant mieux l'empirie que cet indicateur s'approchera de 20% en fin de simulation. } reformuler

La « déclinaison temporelle » mentionnée juste au dessus permet de légèrement raffiner la vraisemblance de cet indicateur : il faut certes atteindre un objectif quantifié (20%), mais les hypothèses empiriques permettent aussi de penser qu'il faut que l'évolution de cet indicateur au cours du temps soit constante, diminuant ainsi plus ou moins à chaque pas de temps. Une configuration de paramètres atteignant l'objectif en fin de simulation mais fluctuant fortement ne sera ainsi pas valide, faisant apparaître un phénomène trop aléatoire dont on ne peut penser qu'il s'est déroulé historiquement. " ?  
pas d'air reformuler

Dans la version 0 de SimFeodal, on atteint en moyenne (moyenne des répliques) **57%** de foyers paysans isolés en fin de simulation. C'est un taux bien trop important, illustrant dès lors une polarisation trop faible. Qui plus est, alors que cet indicateur est en diminution constante (fig. 3.5) pendant la première moitié de la période, il stagne alors puis se met à augmenter, résultant ainsi en ce taux très éloigné de l'objectif. ) réf. à fig 3.5 effectivement mais es dire un peu +

21. De manière incrémentielle et itérative, voir dans le chapitre x? et <http://itsadeliverything.com/revisiting-the-iterative-incremental-mona-lisa>

22. Ne pas oublier de présenter les objectifs dans chapitre 2. Ou alors, on les présente ici au fur et à mesure qu'on mentionne les indicateurs



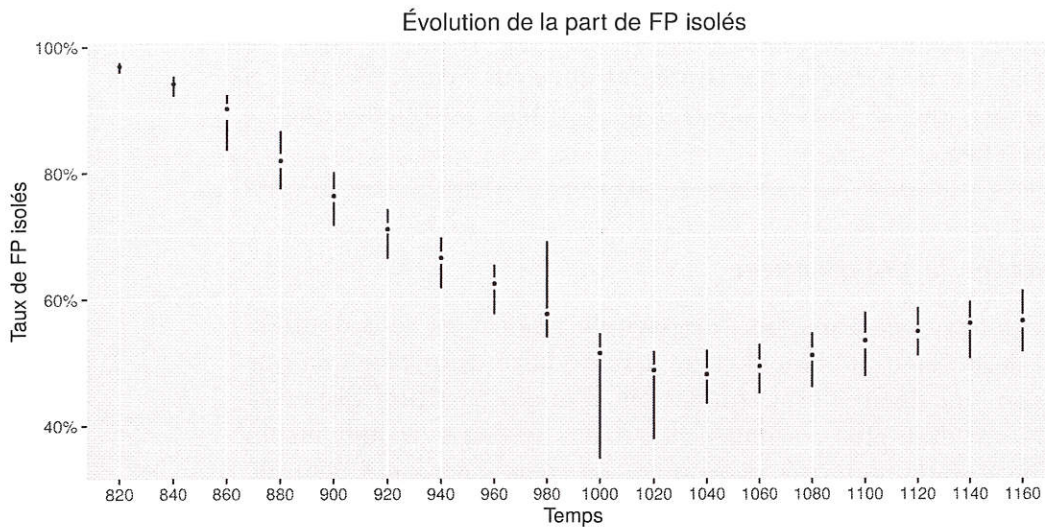


FIGURE 3.5 – Évolution de la part des foyers paysans isolés.

**Nombre d'agrégats** Les foyers paysans se concentrant au sein d'agrégats, il est aussi évident d'observer l'évolution de ceux-ci. Là aussi, (objectif à définir), on peut considérer qu'un nombre d'agrégats en fin de simulation proche de l'objectif, 200, permet de caractériser une polarisation réussie. Cet indicateur ne peut être lu seul, et c'est pour cela qu'il vient dans un second temps : en effet, un faible nombre d'agrégats peut aussi bien être révélateur d'une très faible polarisation des foyers paysans (ceux-ci restant dispersés) que d'une trop importante (un unique agrégat concentrant l'ensemble des foyers paysans par exemple). Pour autant, une fois le taux de foyers paysans dispersés connu, le nombre d'agrégats et son évolution nous renseigne sur les dynamiques de polarisation. On s'attend ainsi à ce que le nombre d'agrégats, très faible au départ (24 dans la version 0), suive trois phases : une première phase de croissance lente, le temps que les mécanismes agissent sur la polarisation, suivie d'une période de croissance plus rapide, une fois que tous les foyers paysans commenceront à être suffisamment attirés par les pôles pour y former des agrégats, et enfin, une nouvelle phase de croissance plus lente, une fois les foyers paysans répartis dans les agrégats existants et qui se déplaceront vers des agrégats plus importants, hiérarchisant le système de peuplement.

Le nombre d'agrégats est assez satisfaisant dans la version 0 du modèle. Il s'élève ainsi à **187** en moyenne, nombre qui plus est assez stable au regard des répliques. L'écart à l'objectif (200) est donc assez minime. On a toutefois vu que le taux de foyers paysans isolé était bien trop important, et dès lors, ces agrégats sont logiquement composés de trop peu de foyers paysans.

L'observation de cet indicateur au cours du temps (fig. 3.6) est quant à elle assez satisfaisante : on retrouve bien les trois phases attendues, quand bien même le début de la croissance plus importante (vers 1000 ici) est un peu trop tardive. Ce moment coïncide de plus avec la stagnation puis l'inversion de la courbe d'évolution de la part de foyers paysans isolés.

On peut quand même considérer que si la polarisation n'est pas assez importante, les structures qui en résultent semblent correspondre à l'empirie.

hyp parachute  
annoncer la 2<sup>ème</sup> ind de  
leur hiérarchie des foyers

trop souvent  
attente

reformule

hyp de  
logique  
introduit  
expérimental

attente OK  
dans la rpl ?



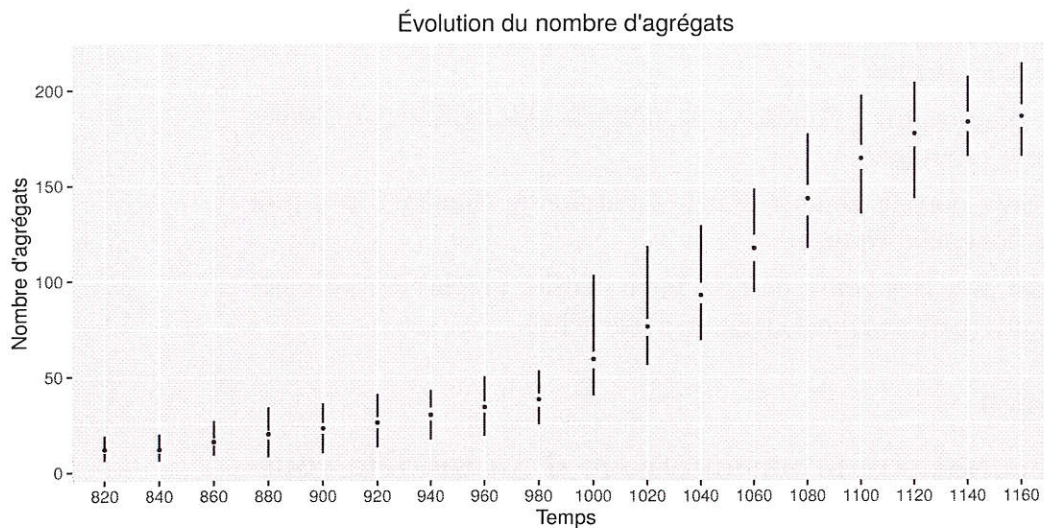


FIGURE 3.6 – Évolution du nombre d'agrégats.

Il y a un problème ici par rapport à ce qui a été mis dans le chapitre : dans le chapitre TMD, on mentionne 83 agrégats en moyenne, alors que dans l'analyse de Base, il y en a 187...

### Nombre de pôles

### Dispersion des agrégats

### Agir sur la polarisation dans le modèle.

### Évaluer et paramétrer la hiérarchisation

### Évaluer et paramétrer la fixation

On se focalise sur les agrégats :

- **Polarisation** : il faut qu'ils rassemblent les FP (taux isolés faible),
- **Hiérarchisation** : qu'ils soient hiérarchisés (composition des agrégats en FP doit tendre vers log-normale),
- **Fixation et dispersion** : et qu'ils soient dispersés dans l'espace de manière stable (fixation par paroisses).

3 leviers différents pour jouer sur ces dimensions :

- **Polarisation** : Améliorer le déplacement des FP (Agent concerné : FP) :
  - ils doivent se déplacer vers un petit agrégat/pôle et ne pas rester isolés
  - une fois dans un agrégat, ils doivent y rester ou aller vers un agrégat plus « important » : en terme de nb de FP, d'attractivité du pôle ou de présence de CA
  - Ils ne doivent pas faire des allers-retours incessants entre les agrégats.
- **Hiérarchisation** : Améliorer la hiérarchisation des attracteurs (Agent concerné : Attracteurs) :
  - La différence entre les pôles d'attraction doit être franche,
  - pour que les petits (1 église par ex.) attirent de petits agrégats,

- et que les gros (châteaux, multi-pôles etc.) polarisent des agrégats importants et stables
- **Fixation et dispersion** : Améliorer la desserte des églises paroissiales (Agent concerné : Paroisse) :
  - Les paroisses doivent couvrir tout le territoire de manière à peu près régulière (logique d'équité),
  - Mais aussi avec des zones de plus forte densité proche des gros pôles/agrégats (petites villes) (logique d'efficacité),

SimFeodal, version 0

### 3.2.2 Avant propos sur le paramétrage d'un modèle complexe

Pour introduire le schéma des étapes, rappeler ici que les paramètres et mécanismes d'un modèle complexe sont en interaction, ce qui produit des effets non linéaires.

Principe des vases communicants !

Mentionner aussi la logique des attracteurs étranges de Lorenz comme paroxysme de cette complexité : ici, moins de variation, mais des répercussions importantes tout de même qu'il faut donc « corriger » par de nombreuses itérations de modifications.

De manière générale

### 3.2.3 Les étapes du paramétrage de SimFeodal