

### 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 ?

Ajuster le chapeau introductif pour prendre en compte tout ce qui vient après, par exemple avec :

1. Indices empiriques
2. Indicateurs
3. Hiérarchie des indicateurs

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 première pré-version – implémentant l'ensemble des mécanismes décidés dans le modèle conceptuel –, mais pas encore stabilisée dans les liens, interactions et valeurs de paramètres de ceux-ci. Il ne répond donc pas nécessairement aux attentes que l'on peut en avoir.

On peut caractériser ces dernières, de manière très générale, comme permettant de reproduire les grands traits de cette transition. 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 ».

decrite à ?

indices/indicateurs

#### Ordres de grandeur

Le premier est constitué d'ordres de grandeurs empiriques estimés – avec une précision plus ou moins importante (cf. tableau 3 p. 317 du chap. TMD, à reproduire dans chap 2). Certaines valeurs sont ainsi connues, que ce soit d'après des sources primaires ou secondaires, et peuvent ainsi constituer des « indicateurs ». Par exemple, on connaît avec quasi-certitude le nombre d'églises paroissiales de la région Touraine en 1100. Certaines de ces valeurs sont toutefois plus proches de l'estimation, comme le taux de foyers paysans isolés en fin de période, que l'on ne peut qu'estimer à partir de sources secondaires et en menant des extrapolations. Ces différents indicateurs observés ou estimés peuvent cependant tous trouver une correspondance presque exacte dans le modèle de simulation, et dès lors, mener à comparaison entre données observées/estimées et données simulées. Ils peuvent ainsi jouer le rôle d'éléments d'évaluation du comportement du modèle simulé.

empiriques

#### Faits et formes stylisés

Le second type d'indice est moins précis, ne reposant pas sur une valeur estimable, mais plutôt sur la connaissance experte d'un phénomène. Il s'agit des « faits stylisés »<sup>20</sup>, plus tendanciels que les indicateurs. On fait un large usage de ces faits stylisés en économie, mais aussi en géographie, par exemple quand on estime que les systèmes de peuplement tendent à se hiérarchiser, la pente de leur rang-taille tendant vers 1 à mesure que le système perdure (Trouver ref, sans doute Pumain/Saint-Julien). De la même manière, le modèle de transition démographique d'Adolphe Landry est un fait stylisé, énoncé à partir de l'observation

quelle ?



de nombreuses récurrences de l'évolution des populations d'un pays en fonction des taux de natalité et de mortalité. Ces exemples montrent qu'au sein des faits stylisés, on retrouve aussi une diversité quant à la précision de leurs énoncés. Dans notre cas d'étude, on retiendra cette logique de faits stylisés concernant des allures de courbes (par exemple l'évolution dans le temps d'un indicateur) ou de répartition spatiale. On pourra dès lors parler de « formes stylisées », aussi bien temporelles – courbe logistique estimée pour la polarisation des foyers paysans –, spatiales – différence dans l'occupation de l'espace par les agrégats entre le début et la fin de la période –, que correspondant à une agrégation à l'échelle du système, comme dans l'observation des hiérarchies grâce aux courbes rangs-tailles.

Ah oui ?

### Indicateurs en sortie *de simul.*

Les ordres de grandeur et formes stylisées évoquées relèvent du domaine empirique. Afin de pouvoir les mobiliser, il est nécessaire de définir des **indicateurs en sortie** dans le modèle de simulation, c'est-à-dire des variables informatiques que l'on enregistrera durant l'exécution du modèle et que l'on pourra <sup>ensuite</sup> ~~alors~~ comparer aux indices empiriques définis. On pourrait définir une typologie différenciant les indicateurs en sortie, <sup>en distinguant</sup> de formes numériques simples (des nombres) <sup>et</sup> des indicateurs plus complexes, multidimensionnels, à même de permettre la confrontation du modèle de simulation avec les formes stylisées identifiées. Ici, on se contentera de regrouper ces différents types de production informatiques du modèles sous la dénomination d'« indicateurs en sortie », ou, plus simplement, d'**indicateurs**.

2 phras. - Tu tournes en rond...

si garde l'expression entière.

### Hiérarchiser les indicateurs

SimFeodal s'appuie sur une dizaine d'indicateurs numériques, ainsi que sur plus d'une trentaine d'indicateurs multidimensionnels. Tous ces indicateurs ne présentent pas le même degré de certitude, la même échelle d'observation, et surtout, le même niveau de précision sur les phénomènes modélisés. A chaque changement dans le modèle, pour une évaluation complète de la capacité de cette version à reproduire les indices empiriques, il faudrait donc observer et analyser chacun de ces nombreux et divers indicateurs. Dans le contexte du paramétrage d'un modèle s'appuyant sur une logique itérative et incrémentielle (voir encadré 3.4), on imagine bien que cela n'est pas possible : le nombre d'indicateurs est bien trop élevé pour avoir rapidement une vision globale de la qualité de représentation du modèle. Il faut dès lors, comme pour toute analyse synthétique, concevoir une hiérarchie d'observation et d'utilisation des indicateurs : il ne sera pas nécessaire d'analyser chacun des indicateurs dans la plupart des cas, seuls les indicateurs jugés plus importants pourront être analysés. Les indicateurs de moindre importance ne seront mobilisés que pour départager des situations dont la différence ne serait pas suffisamment explicitée par l'usage des indicateurs principaux.

Tu les distingues donc !

ov-

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. »



**Incertitude** Dans le modèle de simulation, les indicateurs <sup>de sorte</sup> 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'agrégats 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'agrégats (cette quantité étant elle-même analysable avec plus de précision) en fin de période qu'au début. On peut 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)). Cette quantification de l'incertitude pourrait alors servir de base à l'établissement d'une hiérarchie des indicateurs : on analyserait en premier lieu l'écart entre les ordres de grandeurs <sup>empiriques</sup> et les indicateurs <sup>qui les représentent</sup>, avant d'étudier les indicateurs plus incertains (augmentation de la charge fiscale entre 800 et 1100 par exemple) et de finir par les formes stylisées et leur représentations dans le modèle. Toutefois, SimFeodal se caractérise d'une part par une très forte hétérogénéité dans les niveaux de connaissance des ordres de grandeurs et faits stylisés modélisés, et d'autre part, se voulant un modèle théorique (A dire spécifiquement dans le chapitre 2 ; y faire une ref ici), il n'y a pas de volonté de « coller aux données » : la vraisemblance d'ensemble du modèle compte bien plus que la précision de chacune de ses composantes. <sup>ou</sup>

**Indicateur synthétique** Une autre piste, fréquemment utilisée dans les modèles théoriques, consiste à créer un <sup>h(w 711)</sup> indicateur synthétique, ou un ~~faible nombre d'entre eux~~, par exemple une moyenne de plusieurs indicateurs. Ces indicateurs synthétiques résultent d'une quantification des autres indicateurs (excluant donc les formes stylisées qui sont plus libres d'interprétation), et permettent en particulier de mener des explorations rapides et simples du comportement d'un modèle selon les valeurs de paramètres. SimFeodal n'est pas adapté à de tels indicateurs : une large partie des faits stylisés et ordres de grandeur mobilisés proviennent de connaissances expertes, et les thématiciens qui les ont consolidées rechignent à créer de tels indicateurs composites, en ce que cela demande de pondérer précisément l'importance de chacun des indicateurs par rapport aux autres. Si l'on pourrait créer quelques indicateurs synthétiques, ceux-là ne prendraient en compte qu'une faible proportion du comportement attendu du modèle, résultant en une forte perte de pouvoir explicatif ou descriptif du modèle.

**Définir des dimensions d'analyse** En présence de plus d'une quarantaine d'indicateurs, il est toutefois nécessaire, a minima, d'organiser leur analyse. 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. Ces hypothèses et objectifs sont multiples dans SimFeodal, et dès lors, une hiérarchie ne pourra être établie qu'en leur sein. Il convient donc d'identifier des ensembles d'indices permettant de caractériser les différentes dimensions à analyser, avant de définir, dans chacune de ces dimensions, une hiérarchie d'indicateurs. Dans le chapitre précédent (ref chap 2), nous présentions les principales dynamiques du modèle : (1) polarisation, (2) hiérarchisation et (3) fixation des foyers paysans. En postulant que ces dynamiques sont caractéristiques du modèle, on peut s'appuyer sur cette triade pour caractériser les sorties du modèle, c'est-à-dire mener la confrontation entre indices et indicateurs. Ces trois ensembles,



et les indicateurs qui les caractérisent dans le modèle, sont dès lors considérés comme les trois dimensions d'analyse des sorties de SimFeodal.

Hierarchiser dans les dimensions Chacune de ces dimensions s'applique à plusieurs types d'agents du modèle. Pour définir notre hiérarchie, on retiendra les agents les plus impactés par les dynamiques correspondant à ces dimensions : 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). On aura alors tendance à examiner d'abord un résultat agrégé, représentatif de la structure dans son ensemble, avant d'observer les indicateurs représentatifs des dynamiques ayant mené à cette structure. Dans cet exemple, on analysera donc 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. On peut dès lors définir des « indicateurs principaux » pour chaque dimension, représentatifs des grands traits des structures auxquelles on souhaite aboutir en sortie de simulation, et des « indicateurs secondaires », permettant d'affiner l'évaluation de chacune de ces dimensions.

Une hiérarchie mouvante Notons que l'analyse des indicateurs de sortie suit une hiérarchie parfois mouvante, et en tous les cas, assez peu quantifiable : si l'ordre d'observation est plutôt stable, l'importance que l'on portera à chacun des indicateurs peut varier. Les indicateurs principaux de chaque dynamique sont ainsi « incontournables », c'est-à-dire qu'un résultat trop loin de celui des indices empiriques est disqualifiant. Parmi les indicateurs secondaires, il n'est pas toujours possible, d'après les connaissances des experts sur le sujet, d'établir une priorité ou une pondération de chaque indicateur. L'évaluation de la polarisation par exemple (page 23), se définit principalement par rapport à un indicateur principal – le taux de foyers paysans dispersés –, mais selon les résultats des autres indicateurs de sortie, on leur attribuera une importance variable. L'étude de la dispersion des agrégats et pôles peut ainsi se révéler plus importante que celle de l'évolution du nombre d'agrégats selon les paramètres que l'on souhaite ajuster, ou se montrer tout au moins plus différenciante selon l'état du paramétrage.

#### 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 d'affiner 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 de la situation simulée par le 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

catégoriser  
ind et  
ind et

oui

un indicateur  
donc ?

secondaire ?

oui

ce n'est pas  
et la est clair,  
au-delà non