

Le statut de la modélisation dans une démarche interdisciplinaire

FRANCIS LALOË

Le domaine halieutique est sans doute l'un de ceux où le choix d'un modèle a une importance décisive, si bien que l'illustration que l'auteur en tire est particulièrement bienvenue. La démarche de modélisation et la réflexion sur la structure des objets complexes constituent les fils directeurs d'une analyse critique très rigoureuse, dans laquelle les différentes disciplines, des sciences biologiques aux sciences humaines, ont toutes leur mot à dire.

L'objet de cet article est de discuter, dans le contexte de programmes interdisciplinaires et de leur évaluation, du statut des méthodes et des disciplines relevant des sciences de la représentation des connaissances et de l'information. Les exemples donnés sont pour l'essentiel relatifs à une démarche particulière, en l'occurrence statistique. Cependant les situations décrites ne sont certainement pas spécifiques à cette seule démarche et nous semblent refléter dans une assez grande mesure celles rencontrées par les personnes relevant des sciences de l'information et de la modélisation en général.

Problématique thématique et objet commun

La nécessité de l'interdisciplinarité est largement ressentie et exposée lorsque, sur une même réalité « banalement complexe », un point de vue disciplinaire unique apparaît insuffisant pour répondre à nombre des questions « finalisées » que l'on se pose. On décide alors de confronter plusieurs points de vue, en les articulant selon une représentation qui est elle-même complexe, au sens donné par Legay (1997), relatif à « un système que la perte d'un de ses éléments fait changer de nature ».

Les représentations faites selon les divers points de vue sont nécessairement différentes et on se trouve dans une situation d'où une certaine subjectivité ne peut pas être absente. Le statut de l'objet commun, auquel il est systématiquement fait référence dans une démarche interdisciplinaire, est donc à ce titre ambigu. Il doit cependant bien exister une « définition commune de la problématique et de la stratégie de recherche » (Zanoni et al., 1998), et cette « problématique commune, strictement thématique doit être doublée par un schéma théorique et conceptuel... [qui permet de] repérer les concepts et notions clés qui vont organiser la démarche et constituer son vocabulaire commun de base ; ces notions et concepts proviennent tant des disciplines et spécialités impliquées que de la méthode utilisée pour les associer au

sein d'une même démarche. » (Chassany et Jollivet, 1997).

Une telle situation a plusieurs conséquences.

– Tout d'abord, et il s'agit là d'une autre façon de dire que la problématique commune est au départ strictement thématique, on ne peut pas construire un objet commun pouvant être reconnu comme sien par chacune des disciplines. La présence, toujours évoquée, d'un objectif finalisé et d'une confrontation avec la réalité (Legay, 1993), médiateur pour la définition d'un objet commun, est une condition nécessaire. Cet objet est délimité dans l'espace et dans le temps, il est observé selon les points de vue définis dans le programme et les résultats communs obtenus sont relatifs à cet objet.

FRANCIS LALOË

Halieutique et écosystèmes
aquatiques
IRD (anciennement Orstom)
BP 5045
34032 Montpellier cedex 1,
France
laloe@mpl.ird.fr

Abstract – The status of modelling in an interdisciplinary approach

An interdisciplinary programme always refers to a thematic problematic and applied objectives, as well as to observations about the real world. A common object is constructed, combining disciplinary points of views. The need for modelling sciences is commonly stated, providing tools that appear necessary for the construction of the common object. Using the statistical concept of sufficiency, we discuss the synthetic aspect of answers given to the questions addressed. Using modelling techniques in an interdisciplinary context may lead to enhance the need for question modification and hence rebuilding the common object. We take the example of the exploitation of a marine renewable resource, described from combinations of two points of view, one regarding the resource, the other regarding the exploitation activity. These combinations are different, according to which point of view is considered first. Based on this particular example, we indicate how 'orphan' questions may arise, which can hardly be identified by any of the individual disciplinary points of view. Such questions ought to be taken into consideration if the answers provide sound information related to the thematic domain, such as sustainable development, for example. Their 'adoption' is a step in the programme's progressive problematic. The common object makes no sense if disciplinary objects are not explicitly presented. Giving those disciplinary objects the place they should have, would make it easier to identify the role and the place of modelling sciences in interdisciplinary programmes. © 1999 Éditions scientifiques et médicales Elsevier SAS

modelling / interdisciplinarity / statistics

– Cet objet commun n'est pas au départ scientifique. Il doit être reconstruit, selon un cadre de représentation original, en associant les points de vue et les représentations des disciplines et spécialités impliquées. Comme l'indiquent Chassany et Jollivet (1997), il faut une méthode pour faire cette association, méthode qui fournit aussi des concepts et des notions.

– Cette méthode fournit aussi des outils, entre autres relatifs à la modélisation et à la collecte d'information. Les réponses aux questions posées sont des produits communs du programme et de l'équipe. Elles sont par ailleurs, à partir de l'information collectée en cohérence avec le cadre de représentation construit en commun, une synthèse elle-même réalisée en cohérence avec ce cadre. Les concepts, notions et outils relevant de la méthode devraient donc être complètement appropriés par chaque membre de l'équipe. C'est ce qu'indiquent Zanoni et al. (1998) à propos de la formulation de modèles et de celle de questionnaires d'enquêtes, « outils d'analyse construits collectivement par l'équipe ».

Des méthodes et des outils, application, développements méthodologiques et implication

On conçoit dès lors que la position des personnes dont la spécialité relève des sciences de l'information, peut être ambiguë au sein d'un programme interdisciplinaire.

En premier lieu, l'importance du domaine est bien reconnue. Pavé (1994) note ainsi que « on remarque que l'élaboration d'un point de vue commun est facilitée lorsque la modélisation et donc le (les) modélisateur(s) sont intégrés dans le projet de recherche ». Nombre d'articles publiés dans l'ouvrage *Sciences de la nature Sciences de la société* sous la direction de M. Jollivet (1992) accordent une place très importante à la modélisation au sens large. Les journées du Programme Environnement, Vie et Société consacrées aux tendances nouvelles en modélisation pour l'environnement (Blasco, 1997) ont montré la très grande diversité des démarches et des outils.

Cependant, contrairement aux autres disciplines présentes, celles qui relèvent des sciences de l'information n'ont pas de point de vue initial pour observer la « réalité ». C'est peut-être à ce titre qu'elles ne figurent généralement pas dans la liste des disciplines participant à un programme interdisciplinaire. Leur statut serait plutôt celui d'une méta-discipline pour reprendre un terme employé par Godard (1992).

Dans l'idéal, le statut de la modélisation devrait donc être purement technique, apportant l'expertise nécessaire à la construction collective d'outils d'analyse. Le savoir faire du technicien permet alors, à partir des connaissances qui lui sont fournies et des données collectées, de produire les meilleures réponses possibles aux questions identifiées par le collectif de recherche. Deux difficultés peuvent se poser, de natures et conséquences très différentes.

1. L'exercice peut être techniquement difficile au point de justifier des recherches méthodologiques pour extraire, à partir des observations et des connaissances disponibles, la meilleure réponse possible à une question donnée. Si l'intérêt de telles recherches peut être identifié dans le contexte du programme, elles peuvent être menées indépendamment de ce dernier, en ne contribuant pas à la définition des questions posées. Le terme de « recherches latérales » (Legay 1997) correspond bien à cette situation. Si un résultat théorique (pour la discipline concernée dans le domaine des sciences de l'information) est obtenu à cette occasion, il permet aussi l'amélioration de la réponse à une question extérieure à cette discipline. En référence à cette question, ce résultat peut être légitimement considéré comme « technique ». C'est de ce point de vue particulier que le terme « technique » est employé dans cette contribution.

2. L'exercice peut être techniquement impossible parce que la mise en œuvre de l'outil peut conduire à refuser de répondre à la question posée. Bien entendu ce refus ne peut être opposé qu'après avoir tenté de répondre, ce qui implique que la question a au départ été acceptée¹. Dans ce cas, soit le modélisateur ne dépasse pas son statut technique et la situation est bloquée, soit il participe à la reformulation de la question et, au delà de son application, il s'implique.

Il est important de noter que ces deux difficultés ne se situent pas dans la même dimension. Une situation de blocage peut survenir dans un contexte dans lequel le technicien ne serait intéressé (évalué) que par le seul intérêt théorique (dans sa discipline) des résultats obtenus dans une recherche latérale. Elle peut tout aussi bien survenir dans un contexte où ne serait reconnue (évaluée) que la qualité de sa contribution technique. L'opposition « appliqué-théorique » n'a donc guère de sens dans ce cas.

Synthèses

Les données utilisées pour répondre à une question font l'objet d'une synthèse. Il s'agit bien d'une synthèse parce que, à partir de la seule réponse apportée, il n'est pas possible de reconstituer exactement toutes les données utilisées. Dès lors, une qualité souhaitable de cette synthèse consiste à pouvoir affirmer que les différences entre les jeux de données qui auraient pu résulter du même protocole de collecte et dont les synthèses sont identiques, ne contiennent pas d'élément permettant de préciser la réponse à la question posée ni de répondre à une quelconque autre question. Si cette qualité est effective, toute l'information contenue dans les données est restituée dans la réponse, et les données sont bien relatives à la question posée.

Cette qualité correspond à la définition d'une information donnée par Bateson (cité par Bougnoux (1993, p.227), selon lequel une information est « une différence qui fait la différence ». Elle correspond également (Arnold, 1988) à l'un des principes fondateurs de la statistique, le principe d'exhaustivité² (en anglais *sufficiency*), énoncé par Fisher (1925) : « *a sufficient*

¹ Ainsi que l'indique Legay (1976) pour la biométrie : « Notre position est d'accepter les problèmes tels qu'ils sont posés dans la pratique des biologistes, des médecins, des agronomes... »

² Le concept d'exhaustivité est présenté ici dans son sens statistique. Il peut faire l'objet d'autres définitions (Mullon et Piron, 1998).

statistic] is equivalent for all subsequent purpose of estimation to the original data from which it was derived ». La qualité d'exhaustivité ne peut être établie qu'en référence explicite à la distribution de la variable aléatoire dont les données collectées sont issues, permettant de calculer la vraisemblance d'un jeu de données dispo-

nibles en fonction des paramètres de cette distribution (encadré 1).

Dès lors, la qualité de la réponse à une question dépend de deux aspects bien distincts (voir par exemple Matheron, 1978) : la forme de la distribution choisie pour apprécier la vraisemblance d'un jeu de

1. Exhaustivité (au sens statistique)

(Cette présentation s'inspire de celle, beaucoup plus complète, faite par Arnold, 1988).

Supposons que n observations d'une variable (une taille par exemple) soient réalisées auprès d'individus d'une population donnée, sélectionnés selon une procédure d'échantillonnage aléatoire simple. Si la distribution des tailles suit une loi Normale d'espérance m et de variance σ^2 , alors la densité (vraisemblance) de l'échantillon est égale à :

$$f(x_1, x_2, \dots, x_n) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i - m)^2}{2\sigma^2}} \quad \text{On a donc : } \text{Log}(f(x_1, x_2, \dots, x_n)) = C - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - m)^2$$

où C est une constante qui ne dépend pas des observations. En développant cette formule on obtient :

$$\text{Log}(f(x_1, x_2, \dots, x_n)) = C - \frac{1}{2\sigma^2} \left(\sum_{i=1}^n x_i^2 - 2m \sum_{i=1}^n x_i + nm^2 \right)$$

Les données n'interviennent donc dans leur vraisemblance qu'au travers de leur somme et de celle de leurs carrés. Il n'y a pas d'information dans les différences entre tous les jeux de données pouvant être issus de la mise en œuvre de ce protocole et ayant comme caractéristique d'être équivalents selon la somme de leurs valeurs et selon la somme des carrés de ces valeurs.

La statistique (synthèse) composée des deux variables :

$$\sum_{i=1}^n x_i^2 \text{ et } \sum_{i=1}^n x_i \text{ est dite exhaustive. On montre qu'elle est de plus minimale.}$$

Elle peut s'exprimer selon de multiples façons, par exemple sous la forme plus parlante :

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \text{ et } s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2, \quad \bar{x} \text{ et } s^2 \text{ étant les estimateurs classiques des deux paramètres } m \text{ et } \sigma^2.$$

Partition minimale exhaustive

Si deux jeux de données sont équivalents selon une formulation exhaustive minimale, ils le sont aussi pour toute autre. La relation selon laquelle deux jeux de données sont équivalents s'ils prennent les mêmes valeurs selon une statistique exhaustive minimale, définit une « partition exhaustive minimale ». Cette partition est unique, ce qui amène à considérer qu'elle constitue ce qui est important (Arnold, 1988). Dans l'exemple présent, deux jeux de données (x_1, x_2, \dots, x_n) et (y_1, y_2, \dots, y_n) sont dans la même classe de la partition si :

$$\frac{1}{n} \sum_{i=1}^n x_i = \frac{1}{n} \sum_{i=1}^n y_i \quad \text{et} \quad \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2$$

ou ce, qui revient au même, si :

$$\sum_{i=1}^n x_i = \sum_{i=1}^n y_i \quad \text{et} \quad \sum_{i=1}^n x_i^2 = \sum_{i=1}^n y_i^2$$

Un second exemple, relatif à une proportion, est plus naturel. Si la variable d'intérêt est une qualité quelconque dont la fréquence (inconnue) dans la population est p , on montre, toujours avec un échantillon de n individus sélectionnés selon une procédure d'échantillonnage aléatoire simple, que la statistique « nombre d'individus observés présentant la qualité parmi les n individus de l'échantillon » est exhaustive minimale. Une classe de la partition minimale exhaustive est constituée de l'ensemble des échantillons possibles parmi lesquels le même nombre d'individus présentent la qualité étudiée. Il n'y a pas d'information dans l'ordre d'apparition des observations. On peut observer que la statistique présentée dans ce petit exemple est d'un usage très courant et correspond bien à l'intuition.

Cette propriété ne peut bien entendu être montrée que sous les hypothèses qui permettent d'écrire la densité associée aux données disponibles. Si ces hypothèses ne sont pas valides, la propriété d'exhaustivité ne l'est plus non plus. Dans les exemples ci-dessus, la démonstration de la qualité d'exhaustivité repose de façon essentielle sur l'indépendance des variables, garantie par le protocole de sélection (échantillonnage aléatoire simple). Si cette hypothèse n'est pas satisfaite, par exemple si les individus ont été sélectionnés le long d'un transect, alors l'ordre d'apparition des observations contient de l'information...

Partition définie par les réponses à une question

La question posée, qui a justifié la collecte des données, définit elle aussi une partition. Chacune des classes de cette partition est définie par une réponse particulière, en regroupant tous les jeux de données dont l'analyse conduit à donner cette réponse à la question posée. Le meilleur recouvrement possible des deux partitions définies, l'une selon les réponses, l'autre selon les vraisemblances des observations, est une qualité déterminante. Si cette qualité doit être recherchée, elle ne doit pas être imposée en exigeant a priori que l'une des deux partitions soit conforme à l'autre (voir encadré 2).

données ; la procédure d'estimation utilisée (éventuellement mise au point) pour spécifier cette distribution.

Le premier aspect est relatif aux questions, le second aux réponses. Le fait même de poser une question définit une hypothèse sur la nature de la distribution des variables dont les données sont une réalisation.

La question posée peut ne pas devoir être remise en cause, parce qu'elle définit complètement le « contrat de collaboration ». Les données doivent alors être collectées selon un protocole tel que leur synthèse soit techniquement satisfaisante en permettant d'énoncer sans ambiguïté la réponse à la question posée. Si tel n'est pas le cas, on doit reconsidérer le protocole de collecte.

Dans le contexte d'une collaboration au sein d'un programme interdisciplinaire, les questions sont construites en articulant des points de vue différents, sur une réalité qui « s'impose à l'observation ». Les données doivent bien sûr ici aussi être collectées selon un protocole tel que leur synthèse devrait dans l'idéal constituer une réponse à la question posée (et acceptée). Mais cet objectif n'est plus unique, le protocole fait partie de l'objet d'étude³. L'analyse des données doit alors également porter sur la nature de la distribution dont elles sont une réalisation. De cette analyse découlent des résultats permettant d'apprécier la vraisemblance de ces données et de formuler la synthèse qu'on peut en tirer. Lorsque cette synthèse n'est pas techniquement satisfaisante⁴, parce qu'elle ne recouvre pas bien celle consistant à répondre aux questions posées (*encadré 1*), il y a un besoin de reformulation du questionnement général. Toutes les disciplines du collectif sont alors interpellées. Bien sûr ce besoin peut être aussi identifié à partir d'autres méthodes et concepts et une reformulation ne doit évidemment pas être le fait des seuls modélisateurs, encore moins des seuls statisticiens. Mais ils ne peuvent pas y être indifférents dès lors que toute reformulation du questionnement entraîne celles des synthèses et des protocoles d'observation...

Problématique progressive

Il conviendrait donc d'admettre d'une part qu'il peut exister au sein des disciplines dont sont issus les modélisateurs, des « règles » pouvant leur interdire de répondre, sous la forme demandée, à certaines sollicitations et que, d'autre part, ces refus peuvent être une des sources de progrès communs figurant parmi les résultats des programmes interdisciplinaires. Ces progrès sont identifiables au travers des « réajustements permanents [...] du dispositif et de la problématique de recherche par itérations successives », (Chassany et Jollivet, 1997), relativement auxquels ces auteurs parlent de « problématique progressive ». Bien entendu, un modélisateur participant, avec les concepts et les notions propres à sa démarche, à un programme interdisciplinaire est l'un des acteurs de ce processus.

On pourrait alors penser que l'intégration des disciplines concernées est une chose naturelle. Force est de

constater que cela n'est pas le cas. Il existe en effet une méfiance certaine à l'égard de la modélisation en général. Godard (1992) rapporte ainsi que, dans l'évaluation d'un programme à l'issue duquel une tentative d'articulation à posteriori avait été tentée au niveau empirique, cette tentative a pu être considérée comme un « coup de force tendant, sous couvert d'un empirisme humble, à conférer à l'approche technico-économique une position centrale de pouvoir à l'égard des autres disciplines ». Cet exemple est quelque peu extrême, mais il montre l'existence du risque, pour les disciplines « initiales », de se voir dépossédées de leur objet commun par des techniciens (ou par une discipline plus « technique » que les autres).

Questions à la modélisation

À l'issue de l'article abondamment cité ici de Chassany et Jollivet (1997), les auteurs posent des « questions à la modélisation ». Ils considèrent en premier lieu que la modélisation doit contenir en elle-même la progressivité de la problématique. Il nous semble ici que cette qualité découle de l'acceptation de la complexité (Legay, 1997) qui conduit à traiter de façon dialectique des deux aspects présentés plus haut, relatifs au choix d'un cadre et à la spécification du cadre. Cette progressivité ne peut guère être accompagnée par une démarche qui ne ferait pas elle-même partie intégrante du collectif de recherche.

Ils considèrent ensuite le cas des analyses des interfaces entre des systèmes sociaux et des systèmes naturels. Ils concluent à la nécessité, pour la modélisation, de « se couler dans les problématiques et les corps théoriques disciplinaires, en particulier dans ceux des sciences sociales, si elle veut être un outil d'analyse pertinent pour traiter des interfaces entre sociétés humaines et milieux biophysiques, systèmes écologiques en particulier ». La réponse à cette question est beaucoup plus délicate. Fournir un modèle compatible avec les problématiques et les corps théoriques des sciences sociales et des sciences de la nature (et, qui plus est, en cohérence avec l'objectif finalisé du programme et avec l'objet « réel » sur lequel ont été réalisées les observations) constituerait un achèvement du programme qui pourrait être vu comme une prise de pouvoir. Mais d'autre part un modèle ne présentant pas ces qualités, engendre bien le « risque du transfert analogique inconscient ou de l'artefact technique dans la compréhension des processus socio-écologiques. Faute de cette prudence, qui est aussi de la rigueur élémentaire, la modélisation, sous couvert d'outil de progrès dans l'acquisition des connaissances, ne sera qu'un outil de manipulation supplémentaire des sociétés humaines » (Chassany et Jollivet, 1997).

Cette rigueur élémentaire, qui est aussi de la prudence, est à la charge des modélisateurs et de ceux qui les écoutent.

Le modélisateur se doit d'explicitement ses modèles, avec les suppositions qui ont conduit à choisir un cadre de représentation donné et leurs conséquences. Le contexte de l'interdisciplinarité et de

³ Dans le domaine halieutique, par exemple, ce sont les pêcheurs qui décident de leurs lieux de pêche, ce qui peut avoir des conséquences importantes sur la distribution des données collectées relatives aux captures qu'ils réalisent.

⁴ Selon le sens choisi ici pour le terme « technique », dire qu'une synthèse est techniquement non satisfaisante signifie qu'elle n'est pas clairement relative à la question initialement posée. Si, selon une autre acception, « techniquement satisfaisant » signifie que la méthode ayant conduit à formuler la synthèse a été employée au mieux, la situation présentée ici doit être qualifiée de façon radicalement différente : la synthèse techniquement satisfaisante ne permet pas de répondre de façon satisfaisante à la question. C'est alors cette dernière qui est remise en cause.

Tout dépend donc du point de vue. Dans un contexte de collaboration interdisciplinaire, chacun doit avoir un point de vue et admettre qu'il y en a d'autres. Un mot tel que « technique » peut donc être utilisé, mais en ayant conscience des dangers que ses utilisations comportent. Il en est de même pour les mots qui servent à construire l'objet commun comme nous le verrons par exemple pour « ressource » ou « activité ».

la représentation selon des systèmes complexes est à ce titre exemplaire, en ce sens que les observations ne sont pas entièrement issues de plans d'expérience contrôlés. L'articulation de points de vue différents entraîne par exemple que les suppositions faites selon un des points de vue contraignent les interprétations pouvant être faites selon un autre (Laloë et al., 1998). Le modélisateur se doit de faire état du fait que son modèle n'est pas une solution unique, qu'un paramètre – l'espérance d'une distribution par exemple – peut être un objet mathématique dont l'existence est « créée » par le choix d'un cadre, et que son interprétation selon un point de vue comporte nécessairement un risque. Feller (1940) a ainsi montré, pour rendre compte de données de croissance de l'effectif d'une population, que divers modèles d'interprétations biologiques différentes s'avéraient au moins aussi satisfaisants qu'un modèle logistique (avec les célèbres r et K) qui apparaissait cependant pratiquement idéal. Sa conclusion, sur un sujet pourtant simple comparativement à ceux rencontrés dans des programmes interdisciplinaires, était la suivante : « *Thus in no special case any biological conclusion must be based solely on an agreement between the observed data and the logistic form* ». Dans le même esprit on peut se référer aux principes énoncés par Mc Cullagh and Nelder (1983) en introduction à un ouvrage « technique » sur les modèles linéaires généralisés : « [...] *all models are wrong ; some though are more useful than others and we should seek those [...] not to fall in love with one model to the exclusion of alternatives [...] some imagination and introspection is required here to determine the aspects of the model that are most important and more suspect* ». Ces quelques indications montrent que les modélisateurs peuvent être conscients des difficultés et surtout des limites de leurs « outils ».

Questions aux « autres » disciplines

Pour identifier ces difficultés, il convient de fait que les modélisateurs se coulent dans les problématiques et corps théoriques des autres disciplines, leur permettant de faire la part des choses entre ce qu'ils font et ce qu'on peut attendre d'eux. Mais la réciproque doit être également vraie. La sur-interprétation des paramètres, par exemple, peut en effet aussi venir d'un excès de confiance des autres disciplines à l'égard de modèles entièrement construits sur leurs propres questionnements. Les effets de modes liés à telle ou telle méthode sont peut-être plus causés par l'écho qu'en renvoient leurs utilisateurs potentiels que par les modélisateurs eux-mêmes. En reprenant les termes de Chassany et Jollivet (1997), « le risque du transfert analogique inconscient ou de l'artefact technique dans la compréhension des processus socio-écologiques » est sans doute d'autant plus important que le choix d'un outil, et donc d'une démarche de modélisation, est le fait des autres disciplines (ou peut-être surtout d'une des autres disciplines) du collectif. En d'autres termes, les modèles sont de toute façon

des instruments, qui doivent permettre de se rapprocher d'objectifs fixés en dehors du domaine disciplinaire dont relèvent les modélisateurs ; mais si l'évaluation du modèle (et des modélisateurs) doit bien sûr être faite en référence aux services qu'ils rendent (évaluation extrinsèque, de leur point de vue), elle doit aussi être faite d'un point de vue intrinsèque au sein des disciplines concernées⁵ (Mullon et Piron, 1998).

On pourrait penser, de ce qui précède, que la modélisation peut être questionnée de la même façon par chacune des autres disciplines. En fait les attitudes, les attentes et les demandes des disciplines vis-à-vis de la modélisation et des démarches de modélisation sont extrêmement diverses. L'idée même de système peut différer selon les disciplines (Röling, 1994) et nombre de mots en apparence évidents ont des définitions et des sens très différents. Tel est le cas, par exemple, du mot ressource dont les descripteurs sont assez stables pour les sciences de la nature (ce qui n'exclut bien sûr pas que les valeurs prises par chacun de ces descripteurs puissent montrer de grandes variabilités) et changeants, parce que sans cesse redéfinis par des acteurs, pour les sciences sociales.

On peut dès lors imaginer que représenter la ressource dans un modèle de l'exploitation d'une ressource renouvelable au sein d'un programme de recherche interdisciplinaire associant sciences de la nature et sciences de la société peut constituer une problématique à part entière pour un modélisateur impliqué dans ce programme. Ceci peut être illustré par l'exercice présenté ci-après.

Un système dynamique d'exploitation d'une ressource renouvelable, exemple d'objet interdisciplinaire

Si on s'intéresse à l'exploitation d'une ressource renouvelable, dans le domaine halieutique par exemple, deux points de vue, au moins, sont présents. L'un porte sur la dynamique de la ressource, l'autre sur celle de l'activité d'exploitation. Chacune peut être en partie expliquée par un grand nombre de déterminants. Un biologiste peut choisir de décrire l'état de la ressource en un temps t par un ensemble de descripteurs Y_t (des biomasses par espèces, et/ou des effectifs par classes d'âges et par espèces etc.). L'impact de l'activité sur cette ressource peut être pris en compte au sein d'une fonction f de cette activité (présente et éventuellement passée) et d'autres valeurs (décrivant les conditions d'environnement au sens très large de ce terme, intégrant en particulier des descriptions de l'état passé de la ressource elle-même...) :

$$Y_t = f(x_t, x_{t-1}, \dots, z_t, z_{t-1}, \dots) + e_t$$

(x_t et z_t sont des ensembles de descripteurs de l'activité et de l'environnement, e_t est un bruit de moyenne nulle).

⁵ Par qui doit être évalué un électricien qui refuse de satisfaire à une demande de branchement d'un chauffe-eau parce qu'il considère que l'installation électrique ne satisfait pas aux normes en vigueur ?

De même, un économiste pourra décrire l'activité par un ensemble de descripteurs X_t (nombres d'emplois, nombres de jours de mer, nombres d'unités de pêche de telle ou telle caractéristique etc.) sur lesquels la ressource et les conditions d'environnement exercent un impact :

$$X_t = g(y_t, y_{t-1}, \dots, z'_t, z'_{t-1}, \dots) + u_t$$

(y_t et z'_t sont des ensembles de descripteurs de la ressource et de l'environnement, u_t est un bruit de moyenne nulle, les descriptions utilisées pour l'environnement ne sont pas nécessairement les mêmes dans les fonctions f et g).

Chacune des deux composantes est donc décrite de deux points de vue différents. En adoptant un point de vue sur la ressource, on choisit pour cette dernière une description « expliquée » Y_t et une description « explicative » x_t pour l'activité. Inversement, à partir d'un point de vue sur l'activité, on choisit des descriptions « expliquée » X_t pour l'activité et « explicative » y_t pour la ressource.

Des points de vue et des représentations différentes

La description générale « expliquée » (X_t) de l'activité n'a pas de raison, bien au contraire, d'être la même que celle, « explicative » ou « instrumentalisée » (x_t), utilisée pour rendre compte de son impact sur la ressource. En effet, la façon de décrire cette activité est alors principalement déterminée par la façon de décrire la ressource. Si cette dernière est décrite par une biomasse, l'impact de l'activité est généralement décrit selon un « effort de pêche effectif », fonction monotone (souvent une simple proportion) d'une mortalité. Employer le mot « effort » pour décrire une activité constitue une faute sémantique signalée par Laurec et Le Guen (1981) qui rappellent que « l'effort est l'accumulation d'un travail et non la mesure de son résultat ». Si on distingue des classes d'âge pour décrire la ressource, on distinguera autant de mortalités pour décrire l'impact de l'activité. Si la ressource est constituée de plusieurs espèces, on s'intéressera aux mortalités subies par chacune d'elles. Dans tous ces exemples, la dimension (nombre de variables) de la description (x) de l'activité est déterminée par celle de la description (Y) de la ressource... Bien entendu l'activité peut elle-même être décrite selon bien d'autres considérations. Ce que l'on recherche alors est une synthèse de l'activité X_t selon la forme x_t la plus corrélée possible avec Y_t .

Si on considère deux exploitations A et B, les deux descriptions d'activité peuvent être réalisées pour chacune d'elles, conduisant aux résultats (X_A et x_A) et (X_B et x_B). Le fait que les deux façons de décrire l'activité ne soient pas les mêmes implique que l'une ou l'autre des contradictions suivantes peut apparaître⁶:

1. $X_A \neq X_B$ et $x_A = x_B$ (deux activités dont les descriptions montrent qu'elles sont différentes ont le même impact sur la ressource.) ; ou bien

2. $X_A = X_B$ et $x_A \neq x_B$ (deux activités dont les descriptions sont identiques ont des impacts différents sur la ressource.)

Evidemment cette propriété peut aussi être énoncée en remplaçant les « X et x » par les « Y et y ». Elle doit même l'être pour que soit respectée l'une des conditions communément admises de l'interdisciplinarité, selon laquelle chaque point de vue est instrument de chacun des autres⁷.

Dans la pratique, ces contradictions ont des conséquences très différentes sur la capacité d'articuler les deux points de vue. Ceci peut être illustré à partir d'un d'entre eux, relatif à la ressource par exemple :

La première contradiction correspond à une situation confortable. L'impact x de l'activité se déduit de sa description X globale, et il est bien naturel que certaines parties de cette information ne contribuent pas à cette synthèse. C'est la situation qui prévaut lorsqu'on compare les résultats de pêche obtenus par différentes unités de pêche pour pouvoir exprimer l'activité de chacune d'elles selon l'une d'entre elles (standardisation des efforts de pêche, voir encadré 2). Par exemple, si selon la façon de rendre compte des captures, celles d'un chalutier sont dix fois supérieures à celles d'une pirogue, alors les trois flottes constituées de 20 pirogues ou de deux chalutiers ou bien encore de 10 pirogues et un chalutier peuvent être équivalentes en termes d'impact sur la ressource ; elles ne le seront certainement pas en termes d'activité.

La seconde contradiction est moins « acceptable ». Elle correspond à une situation dans laquelle une même description globale de l'activité peut engendrer des impacts différents sur la ressource. On ne peut plus réaliser de synthèse et, surtout, l'impact de l'activité sur la ressource devient variable et présente une utilité amoindrie pour décrire cette ressource. Il est alors tout à fait légitime, la question de l'état de la ressource restant posée, de rechercher ailleurs l'information, en faisant appel à des observations directes selon les méthodes disponibles telles que les campagnes scientifiques de chalutage ou d'observation acoustique. Ceci se traduit néanmoins par une perte d'information commune relative à l'interface « activité-ressource » et une probable difficulté de communication. Les conséquences en terme de compréhension et de capacité de formalisation de « processus socio-écologiques » peuvent être très importantes, quel que soit le sens accordé à ce terme. Le recours à des observations directes ou à des méthodes ne reposant pas sur une quelconque description de l'activité peut donner les meilleures réponses possibles sur l'évolution des mortalités provoquées par la pêche, mais ce seul résultat n'est utile en termes d'aménagement ou de gestion que s'il existe un décideur susceptible de fixer (éventuellement de façon adaptative) les mortalités par pêche à venir. La mise en application de ces seules méthodes dans un contexte général halieutique engendre donc de façon implicite une hypothèse extrêmement contraignante sur la représentation de l'activité et de son déterminisme (cf. note 7)... Et le fait qu'une même activité puisse engendrer des impacts différents est une indication pouvant remettre en cause cette hypothèse...

⁶ Il peut s'agir en fait de la même exploitation en des temps différents.

⁷ Une instrumentalisation « stricte » est par ailleurs inéluctable. La liste des points de vue articulés est finie et arbitraire et les informations issues de points de vue non pris en compte dans cette liste sont alors par définition « instrumentales ». On ne peut dès lors plus être neutre ; par exemple, dans le domaine halieutique, les approches « écosystémiques » sont en développement très rapide avec un effort de recherche croissant sur l'impact de la pêche sur les écosystèmes. Il peut en découler une vision de l'aménagement dont des décisions sont directement intégrables aux modèles d'écosystèmes. C'est peut-être en partie pour cette raison que la problématique des réserves marines fait l'objet d'un intérêt croissant.

2. Standardisation des efforts de pêche, puissances de pêche

La « standardisation des efforts de pêche » consiste à rechercher une « puissance de pêche » p_i pour chaque unité de pêche i . Les résultats (capture en poids) de cette unité seront en un lieu et un moment donnés en moyenne égaux à p_i multiplié par les résultats moyens obtenus en ce lieu et à ce moment par une unité « étalon » (dont la puissance est donc égale à 1). Ce résultat doit être vrai quels que soient le lieu et le moment considérés.

L'estimation des puissances de pêche conduit donc à poser un modèle selon lequel les logarithmes des résultats de l'unité i en une combinaison spatio-temporelle j,k seront en moyenne égaux à un effet de l'unité i additionné d'un effet lié à la combinaison j,k . On impose une contrainte de nullité de l'effet de l'unité étalon. Les puissances de pêche peuvent alors être obtenues par l'exponentielle des effets individuels des unités de pêche. On pose donc un modèle linéaire à deux facteurs (unité de pêche et combinaison spatio-temporelle) sans interaction. En effet, l'existence d'une interaction entre l'effet « unité de pêche » et « combinaison spatio-temporelle » signifierait que la puissance de pêche d'une unité peut varier dans le temps et dans l'espace.

Sous les hypothèses d'absence de ces interactions, de normalité, d'indépendance et d'égalité des variances des distributions dont les logarithmes des captures sont des réalisations, les estimations par les moindres carrés (analyse de variance) des différents effets et de la variance résiduelle constituent une statistique exhaustive minimale (voir par exemple Saporta, 1990). Dans ce cas la synthèse de données correspond bien à la question posée (avec la capacité de décrire également des variations spatio-temporelles).

La structure des données habituellement disponibles (captures réalisées des diverses unités de pêche en connaissant les dates et lieux de pêche) conduit à poser un modèle incluant ces interactions. Si ces interactions sont significativement non nulles, le modèle décrivant la distribution des données doit les intégrer et les estimations des valeurs prises par ces interactions doivent figurer dans la statistique exhaustive minimale. Dans ce cas la réponse à la question posée n'est plus satisfaisante, on peut dire que l'unité de pêche i a des résultats p_i fois plus importants que l'unité étalon, mais il faut préciser où et quand. Conserver le modèle sans interaction pour répondre à la question initialement posée conduit donc ici à produire une synthèse qui n'est pas satisfaisante.

On se trouve typiquement dans une situation où la question peut devoir être reformulée. Par ailleurs l'existence d'interactions signale une variabilité des impacts de l'activité des diverses unités de pêche, variabilité sur la connaissance de laquelle elles ont éventuellement fondé leurs décisions de lieux et méthodes de pêche. On entre alors de plein pied dans l'adoption de la question orpheline présentée dans le texte. Selon que ces interactions sont prises ou non en compte dans leurs décisions par les pêcheurs, il convient de les représenter selon des facteurs de variation à effets fixes ou aléatoires, ou selon un mélange des deux. On entre alors également dans des « recherches latérales ».

D'une façon générale, ce type de difficulté se rencontre dès lors que la description « expliquée » d'un aspect n'apparaît pas suffisamment discriminante pour permettre sa description « explicative » d'un autre aspect. En privilégiant le point de vue sur l'activité, on pourrait ainsi considérer les difficultés liées aux changements de prix, par exemple causés par la dévaluation d'une monnaie, pouvant affecter dans des proportions différentes les diverses composantes de la ressource. Ces changements peuvent ne pas être restitués selon Y_t (généralement des poids ou des nombres d'individus), mais ils doivent entraîner des différences immédiates dans y_t (au travers de prix) dès lors que les pêcheurs peuvent prendre en compte ces changements dans leurs décisions. Les considérations données dans les paragraphes qui précèdent pourraient donc être reprises en partant du point de vue sur l'activité.

Des questions orphelines...

Dans la situation « moins acceptable » exposée plus haut, on peut considérer la question suivante : « quelle est la distribution (ou la variabilité) des mortalités (x) pour une activité (X) donnée ». Si l'énoncé de cette question paraît simple, son identification est difficile. En effet, du point de vue sur la ressource, l'existence de cette variabilité est un obstacle à la description de son état. Du point de vue sur l'activité, les différences entre les x n'entrent pas dans la définition de X , car dans ce cas des x différents impliqueraient des X eux-mêmes différents. On a donc affaire ici à une « question orpheline », fondée sur des différences que ne reconnaissent pas les points de vue initiaux.

Qui peuvent s'imposer...

Il convient alors de savoir si cette question mérite d'être adoptée et traitée. La réponse est positive si les différences entre les x correspondant à un même X ne laissent pas indifférents, parce qu'ils contiennent des éléments de réponses à des questions jugées importantes. Ces dernières, en étant étrangères aux points de vue initiaux, ne peuvent guère qu'être relatives à la problématique thématique du programme. C'est le cas si le problème rencontré est lié à l'existence d'une capacité d'autorégulation du système d'exploitation résultant des décisions des exploitants, par exemple lorsqu'ils choisissent entre des actions engendrant des impacts différents sur la ressource. Un programme affichant un intérêt pour le développement durable ne peut pas éluder cette question, si elle se pose.

En affichant ce type de questions, et le domaine qui peut les légitimer, on se place dans une situation paradoxale. Il s'agit d'une part de questions originales, à ce titre pouvant conduire à une activité de recherche fondamentale mais, d'autre part, en étant orphelines, leur légitimité ne peut guère être fondée dans un premier temps que sur des considérations liées à la finalité du programme, étrangères aux points de vue

disciplinaires initiaux. Elles relèvent alors d'une recherche « non fondamentale » ou en tout cas au départ difficilement reconnaissable selon un point de vue disciplinaire. Par ailleurs, les modèles tentant d'articuler les divers points de vues, au vu des données habituellement disponibles, conduisent fréquemment à des solutions cohérentes non uniques liées à des confusions d'effets, totales ou partielles, qui ne peuvent guère être résolues sans l'apport de connaissances « monodisciplinaires », renvoyant ainsi à un besoin de telles recherches... La problématique progressive est largement faite d'allers et retours entre mono- et interdisciplinarité.

Dans une redéfinition de l'objet commun

En adoptant de nouvelles questions et en cherchant à leur apporter des réponses qui seront des résultats du programme, on redéfinit l'objet commun. De l'exemple présenté plus haut, il ressort que toutes les questions ne sont pas identifiables au départ d'un projet. Leur adoption peut caractériser une étape dans la problématique progressive exposée par Chassany et Jollivet (1997). Il convient de noter que de telles étapes sont marquées par des difficultés nées de la confrontation entre points de vue. Dans l'exemple présenté ici, il aurait été dommage de commencer par s'imposer une définition commune de ce que sont l'activité et la ressource. En revanche il est nécessaire de disposer de définitions explicites et de descripteurs associés, parce que la question peut émerger de l'analyse de leurs différences. En ce sens, la confrontation des points de vue consiste pour une bonne part en l'analyse des situations que tous ne voient pas équivalentes. L'objet commun n'a de sens que s'il coexiste avec des objets propres différents, et tous doivent être explicités.

Discussion

Parmi les difficultés relatives à l'évaluation d'un programme interdisciplinaire et de ses participants, on retrouve en bonne place les critiques relatives au « flou disciplinaire » des productions individuelles. Ces critiques sont généralement associées à la mise en avant d'un objectif finalisé, conduisant à considérer que le programme relève de l'expertise et non de la recherche fondamentale. Les réponses à cette critique font état de la nécessaire compétence disciplinaire de chaque participant pour une solide construction d'un objet commun. Ce débat apparaît quelque peu sans fin avec de bons arguments de part et d'autre.

Peut-être faudrait-il que chaque participant puisse être en même temps disciplinaire et interdisciplinaire. Zani et al. (1997) évoquent plutôt à ce sujet une succession d'états, avec pour chaque chercheur, lors du dialogue interdisciplinaire, une « sorte de renoncement partiel à l'approfondissement analytique des aspects plus strictement disciplinaires de la probléma-

tique », mais avec la possibilité pour lui « à un autre moment [d']approfondir son analyse spécialisée [...] en utilisant les données relevées lors de la recherche collective au profit de travaux individuels à caractère disciplinaire ».

Si on revient au problème de l'articulation de points de vue, on constate que, s'il doit y avoir renoncement, celui-ci doit être fait en connaissance de cause. Le fait d'accepter qu'une synthèse soit opérée selon son propre point de vue, c'est-à-dire d'accepter que des choses qu'on voit différentes soient considérées équivalentes lors de la spécification d'un objet commun, implique que l'analyse disciplinaire correspondant à ce point de vue soit effectivement réalisée et disponible. Si elle est faite à un autre moment, ce doit être avant et non après, sous peine de ne produire de ce point de vue qu'une moyenne inqualifiable. De plus, il sera extrêmement difficile de concrétiser la progressivité de la problématique qui peut impliquer, à chaque itération – une reformulation – et donc une analyse – de chacun des points de vue devant être réarticulés. Ici encore on en revient au fait que l'objet commun ne peut pas être explicité sans que le soient les points de vue – et donc les objets – disciplinaires.

Les disciplines des sciences de l'information sont nécessaires, avec leurs outils, leurs démarches et leurs concepts. Si elles produisent des synthèses et des réductions, elles doivent les qualifier et pour cela se couler dans les problématiques et les corps théoriques disciplinaires, ce qui, ici encore, implique que soient explicités les objets disciplinaires.

L'objet commun, s'il reste absolument nécessaire, ne doit pas masquer les points de vues disciplinaires sans lesquels il n'existe pas. Les disciplines de modélisation ont certainement une spécificité liée à l'absence d'un

Résumé – Le statut de la modélisation dans une démarche interdisciplinaire

Les programmes interdisciplinaires font systématiquement référence à une problématique thématique et à une confrontation avec la réalité. Un objet commun est construit en confrontant et articulant des points de vue disciplinaires. L'importance de la modélisation est généralement reconnue, apportant des outils nécessaires à la construction de cet objet commun. À la lumière du concept statistique d'exhaustivité, nous discutons de la nature synthétique des réponses aux questions posées. La mise en œuvre de l'outil peut conduire, de façon légitime dans un contexte interdisciplinaire, à montrer le besoin de reformuler le questionnement général et à participer à la reformulation de l'objet commun. À partir d'un exemple, la description d'une exploitation halieutique, il s'avère que certaines questions sont « orphelines », parce que non reconnaissables à partir de chacun des points de vue disciplinaires initiaux. Ces questions peuvent s'imposer lorsqu'elles sont liées à la problématique thématique du programme. Leur adoption marque alors l'une des étapes de la problématique progressive générale du programme. L'objet commun n'a de sens que s'il coexiste avec les objets disciplinaires explicites. Donner à la présentation de ces objets l'importance qui leur revient, peut être de nature à permettre une meilleure identification du statut de la modélisation dans une démarche interdisciplinaire.
© 1999 Éditions scientifiques et médicales Elsevier SAS

modélisation / interdisciplinarité / statistique

point de vue initial sur la réalité à observer. Mais il faut préciser qu'à l'instar des autres disciplines, elles acceptent et bénéficient des autres points de vue. Comme les autres disciplines, elles acceptent la problématique thématique de départ. La mise en œuvre des outils, avec les concepts qui leurs sont liés, les conduit à participer à l'identification des besoins de reformulation de l'objet commun, et à participer à ces reformulations proprement dites. Ces actions les conduisent également à identifier des problématiques de recherche dans leurs propres domaines. Toutes ces caractéristiques sont largement communes à toutes les disciplines participant à un programme interdisciplinaire. Les disciplines de modélisation devraient donc pouvoir être considérées, simplement, comme figurant parmi celles qui collaborent au sein de tels programmes.

Remerciements

Je remercie les *referees* et Marie-Hélène Durand pour les très fructueuses remarques faites à la lecture du manuscrit

RÉFÉRENCES

- Arnold S.F. 1988. Sufficient statistics, in : Kotz S., Johnson N.L., (eds.), *Encyclopedia of statistical sciences*, Wiley Interscience, New York, Vol. 9 : 72-80.
- Blasco F., (ed.). 1997. *Tendances nouvelles en modélisation pour l'environnement*. Journées du Programme Environnement, Vie et Sociétés du CNRS. CNRS Elsevier, Paris, 445 p.
- Bougnoux D. 1993. *Sciences de l'information et de la communication*. Textes essentiels, Larousse, Paris. 809 p.
- Chassany J.P., M. Jollivet. 1997. Questions à la modélisation à partir d'une expérience d'observation continue : le cas du Causse Méjean, in : Blasco (ed.), *Tendances nouvelles en modélisation pour l'environnement*. CNRS, Elsevier, Paris : 345-349.
- Feller W. 1940. On the logistic law of growth and its empirical verifications in biology. *Acta Biotheor ser. A*, 5, 51-65.
- Fisher R., 1925. Theory of statistical estimation. *Proc. Camb. Philos. Soc.* 22, 700-725.
- Godard O. 1992. La relation interdisciplinaire : problèmes et stratégies, in : M. Jollivet (ed.) *Sciences de la nature, Sciences de la société ; les passeurs de frontières*. CNRS, Paris, 427-456.
- Jollivet M. (ed.), 1992. *Sciences de la nature, Sciences de la société ; les passeurs de frontières*. CNRS, Paris, 589 p.
- Laloë F., Pech N., Sabatier R., A. Samba. 1998. Model identification for flexible multifleet-multispecies fisheries: a simulation study. *Fisheries research*. 37, 193-202.
- Laurec A., Le Guen J.C. 1981. *Dynamique des populations marines exploitées*, Cnexo, Brest, Rapp. Sci. Techn., 45, 117 p.
- Legay J.M., 1976. Pour une biométrie. *Statistique et analyse des données* 1, 2, 5-11.
- Legay J.M., 1993. Une expérience est elle possible, in : Lebreton, Asselain (eds.) *Biométrie et environnement*. Masson, Paris 1-14.
- Legay J.M., 1997. *L'expérience et le modèle. Un discours sur la méthode*. Sciences en questions, Inra, Paris, 112 p.
- Matheron G. 1978. Estimer et choisir. Essai sur la pratique des probabilités. *Les cahiers du centre de morphologie mathématique de Fontainebleau*. 7, 175 p.
- McCullagh P., Nelder J.A. 1989. *Generalized linear models*. Second edition, Monographs on Statistics and applied Probability, 37. Chapman & Hall, London, 511 p.
- Mullon C., M. Piron. 1998. Sur la méthodologie de mise en place des observatoires socio-économiques, in : Laloë F., Perrier X. (eds.), *De l'observation à l'analyse, implication de la biométrie dans les pays en développement*. Société française de Biométrie 15, 61-77.
- Pavé A. 1994. Les termes d'une approche et d'une programmation scientifiques, in : *Plan d'action du PIR Environnement, Vie et Sociétés*. Décembre 1994, 5-17.
- Röling N. 1994. Platforms for Decision-making about ecosystems, in : L.O. Fresco, L. Stroosnijder J. Bouma, H. van Keulen (eds.), *The future of the land. Mobilising and integrating knowledge for land use options*. Wiley, New York, 385-393.
- Saporta G. 1990. *Probabilités analyse des données et statistique*. Technip, Paris 493 p.
- Zanoni M., Pivot A., Vargas M., Raynaud C., Lescure J.P., J. Quensièrre. 1998. La recherche en environnement, À propos de quelques pratiques interdisciplinaires. *Natures Sciences Sociétés* 6, 50-57.