

Regards

Jean-Marie Legay, entre expérience et modèle

Francis Laloë

Statisticien, IRD, UMR GRED, 34394 Montpellier cedex 5, France

À l'issue d'études universitaires hybrides, combinant statistique mathématique et génétique, je me suis retrouvé « mathématicien – informaticien en océanographie » à l'Orstom (actuel IRD, Institut de recherche pour le développement). Il m'a alors fallu du temps pour me forger une identité disciplinaire, et Jean-Marie Legay a eu une influence déterminante dans cette construction¹.

En effet, le travail au sein d'un centre de recherche océanographique au Sénégal m'a conduit à participer aux études sur la pêche artisanale sénégalaise, objet qu'aucune discipline scientifique ne peut s'approprier sans faire d'hypothèses – en général implicites, et souvent même inconscientes – dans le champ d'autres disciplines.

Le statisticien se voit demander de traiter des données pour en produire des synthèses (des statistiques) en référence à des questions posées par d'autres, et au service desquelles il accepte de se mettre. Une même synthèse peut alors contribuer aux réponses à plusieurs questions possibles, et il peut y avoir des contradictions entre les points de vue à partir desquels ces questions sont posées. Ainsi, une évolution de rendements de pêche peut s'interpréter d'au moins deux points de vue différents, comme une évolution de l'abondance de la ressource exploitée ou bien comme une évolution des choix d'espèces cibles faits par les pêcheurs. Le statisticien se trouve au centre de ces contradictions, dans une situation qui va bien au-delà de celle qu'impliquerait la simple acceptation de faire de la statistique appliquée.

Lorsque plusieurs interprétations sont possibles, le choix d'une d'entre elles revient à adopter un point de vue particulier et implique une part d'arbitraire. L'existence de ce choix est liée à la complexité de l'objet observé. Et c'est là que l'apport de Jean-Marie Legay est précieux. Il qualifie de complexe « un système que la

perte d'un de ses éléments fait changer de nature ». Cette définition figure dans l'ouvrage intitulé *L'Expérience et le Modèle*². Cela permet de comprendre que les diverses interprétations pouvant être faites d'une même observation peuvent être reliées à diverses associations d'éléments d'un système, correspondant à autant de perceptions ou constructions de ce système. L'existence de plusieurs points de vue possibles n'est pas nécessairement une faute, à condition qu'elle soit reconnue et promue : « c'est la multiplicité des points de vue et la confrontation des résultats des modèles correspondants qui conduisent à la meilleure connaissance du système étudié³ ».

Dans le titre de l'ouvrage cité, l'association des mots « expérience » et « modèle » est essentielle et originale. En effet, Jean-Marie Legay expose pour commencer les principales étapes de l'expérience, avec un premier achèvement correspondant à la planification des expériences de l'époque « fisherienne » avec la possibilité de répartir la variabilité d'observations en plusieurs parties indépendantes ; une partie contrôlée, pouvant être représentée par un nombre limité de facteurs et une partie non contrôlée dont l'existence est admise et dont les causes, connues ou non, peuvent être innombrables. Comme l'indiquait Jean-Marie Legay, « Fisher avait accepté sans le dire la complexité et ouvrait la voie à l'analyse moderne des données...⁴ ». La répartition de la variabilité en plusieurs parties indépendantes implique la mise en place de plans d'expériences contrôlées, autorisant l'analyse des données issues de l'expérience pour répondre à l'aide d'estimations et de tests d'hypothèses à la question ayant justifié la mise en place de l'expérience. Mais cela conduit en retour à s'interdire le traitement de

Auteur correspondant : francis.laloe@ird.fr

¹ Voir aussi, dans ce numéro et le suivant, les autres hommages à Jean-Marie Legay.

² Legay, J.-M., 1997. *L'Expérience et le Modèle. Un discours sur la méthode*, collection Sciences en questions, Inra Éditions, p. 18. Toutes les citations faites ici sont issues de cet ouvrage.

³ *Ibid.*, p. 58.

⁴ *Ibid.*, p. 17.

données en dehors du domaine défini par le plan d'expérience. Cela peut même conduire à abandonner le traitement de données qui ne seraient pas issues d'un plan d'expérience valide. Une telle situation peut être une erreur comme l'indique la phrase cinglante de Fisher expliquant que consulter le statisticien après l'expérience revient souvent à lui demander d'en faire l'autopsie : il pourra peut-être dire de quoi l'expérience est morte.

Il s'agissait là pour moi d'une étape qui ne pouvait guère être dépassée, et je me souviens d'une discussion avec Jean Générmont, mon directeur de thèse à qui j'annonçais mon recrutement à l'Orstom pour y travailler sur des données de pêche et qui me disait que je serais confronté aux difficultés d'avoir à traiter des données non issues de plans d'expériences... Et tel a bien été le cas.

Ainsi, face, par exemple, à des données de rendements de pêche, collectées pour disposer d'indices d'abondance, il peut apparaître évident qu'elles peuvent dans certains cas indiquer ou traduire bien d'autres choses aussi, dont les décisions des pêcheurs. On se trouve alors face à un dilemme. On peut considérer que l'expérience est morte et rechercher ailleurs l'information (auprès d'autres pêcheurs plus « fiables » ou en organisant des campagnes de pêche scientifique). Mais on peut décider de considérer que ces rendements sont un fait en eux-mêmes, qu'ils sont une expression d'un système complexe et qu'ils contiennent de l'information sur la durabilité et les possibilités de développement de ce système. L'analyse de ces rendements devient un impératif d'ordre déontologique si la caractérisation de cette durabilité et de ces possibilités de développement constitue un objectif général du programme de recherche dans le cadre duquel l'observation est réalisée.

On dépasse alors le cadre de l'expérience contrôlée et il convient de multiplier les points de vue et d'articuler les modèles correspondants pour proposer une interprétation cohérente des données disponibles. C'est ici encore de l'expérience, avec un recours aux modèles allant bien au-delà de celui du plan d'expérience avec la distribution des variables aléatoires observées selon ce plan. J'ai alors été fortement déstabilisé en constatant que Jean-Marie Legay annonçait qu'effectivement, l'étape fisherienne pouvait être dépassée et que la voie était ouverte pour « l'ère des modèles⁵ ». En effet, je ne voyais pas vraiment de continuum entre, d'un côté, le traitement de données d'enquêtes permettant d'en produire une synthèse sous forme d'estimations diverses d'objets statistiques (des espérances et des variances d'estimateurs associés à diverses sources de variation pour l'essentiel) et, de l'autre, les interprétations liées au choix d'un nom (par exemple, une abondance) pour désigner ces objets.

⁵ *Ibid.*, p. 18.

Mais cette annonce était accompagnée de précisions essentielles.

– La première consistait à dire qu'on pouvait très bien décider de ne pas procéder à l'analyse des données posant problème, et mettre en place un protocole permettant de répondre sans ambiguïté à la question posée. En d'autres termes, « la complexité est une décision⁶ ». Évidemment, cette décision a un coût, celui de la « perte des critères d'évidence⁷ » au profit de critères de cohérence. Le passer sous silence aurait signifié de la part de l'auteur une prétention démesurée. Jean-Marie Legay a été déterminant en identifiant ce coût. Il l'est encore, dans un domaine où il me semble que nombre de modélisateurs ne désespèrent pas de trouver des vérités en affichant ou restituant des évidences.

– La seconde porte sur la nature des modèles avec deux aspects. Tout d'abord, la complexité ne doit pas être associée à un type particulier de modèle ou de comportement qui pourrait être présenté comme une théorie de la complexité. Cela m'a semblé essentiel avec une phrase « je n'accepte pas de dire que les systèmes complexes sont ceux où l'on rencontre des phénomènes chaotiques⁸ » qui m'a été d'un grand secours à une époque où le chaos était effectivement devenu un synonyme de complexité... Par suite, et en combinaison avec la perte du critère d'évidence, la recherche du « bon » modèle devient une illusion. L'objectif du modèle n'est donc pas de représenter le système et sa justification doit être recherchée dans son utilité. Dès lors, il ne peut pas exister de modèle supérieur à tout autre, il y en a au moins autant qu'il peut y avoir d'objectifs.

C'est ainsi que le lien entre expérience et modèles pouvait être montré, en considérant plus précisément « qu'expérience et théorie sont devenues inséparables et que les modèles nous donnent le moyen de faire vivre cette relation⁹ ». Pour ma démarche personnelle, cela était déterminant en donnant une légitimité à un traitement statistique de données considérées comme réalisations d'un modèle cohérent, permettant d'en extraire une synthèse qui pouvait ensuite être utilisée – avec le modèle – pour tenter de répondre à des questions légitimes, parce que relatives, comme évoqué plus haut, à la durabilité du système observé. Évidemment, le modèle ne traduit qu'une cohérence et son efficacité pour représenter des données ne doit pas être présentée comme preuve d'une quelconque vérité. Ici encore, l'apport de Jean-Marie Legay est déterminant en « [donnant] au modèle dans le cadre de la recherche scientifique une acceptation instrumentale¹⁰ ».

⁶ *Loc. cit.*

⁷ *Loc. cit.*

⁸ *Ibid.*, p. 19.

⁹ *Ibid.*, p. 62.

¹⁰ *Ibid.*, p. 23.

On reste alors bien dans le contexte d'une instrumentalisation de la statistique, au service de questions posées en dehors de son domaine. Mais contrairement au cadre de l'expérience contrôlée, au service d'une question, l'observation d'un système complexe peut conduire tous

les participants à un projet à acquérir ensemble une « meilleure connaissance du système étudié » et amener ainsi chacun à contribuer à l'identification des questions nouvelles.