

DU BON USAGE
D'UNE ENQUÊTE EXCEPTIONNELLE

UN STATISTICIEN, BOWLEY, ÉCRIVAIT VOICI UNE CINQUANTAINES D'ANNÉES LA PHRASE SUIVANTE DANS L'INTRODUCTION D'UN OUVRAGE SUR SA spécialité :

« Le premier pas dans l'établissement d'une statistique est de savoir qu'est-ce que l'on va compter et le dernier, quand le travail est terminé, est de se rendre compte de ce que l'on a réellement compté ».

Comment peut-on répondre à ces questions quand on considère le travail considérable accompli par le S.C.E.E.S. pour établir une image de la production fourragère en France ? Il faudra ensuite y ajouter des réflexions sur la façon d'utiliser les données recueillies et dont la validité ou le sens auront été discutés.

Rappelons que la connaissance implique que l'on soit capable d'établir une correspondance entre les constituants de deux ensembles tel qu'à tout « a » de l'ensemble A on puisse faire correspondre un élément de l'ensemble B. Or un ensemble n'a d'existence concrète que si ses constituants peuvent être définis sans ambiguïté, c'est-à-dire sans qu'il y ait hésitation entre la distinction d'un élément quelconque « a » de l'ensemble A et un élément « a' » du même ensemble.

La définition même des éléments constituants d'un ensemble implique que l'on ait défini une typologie, c'est-à-dire un ou des critères qui permettent de caractériser ses constituants. C'est là une tâche délicate qui a dû être réalisée avant même

d'entreprendre le recensement qui est à la base de l'enquête. Par exemple, il a fallu définir les prairies et les autres cultures fourragères. Cinq types ont été distingués :

- les prairies artificielles constituées exclusivement par une légumineuse ;
- les prairies temporaires : cultures de graminées pures ou associées à des légumineuses mais dont la durée prévue est limitée ;
- les prairies permanentes appelées parfois surfaces toujours en herbe parce qu'elles ne sont pas labourées ni ressemées avec une certaine périodicité ;
- les cultures fourragères principales qui comprennent les surfaces autres que l'herbe : il s'agit plus particulièrement du maïs fourrage et d'autres plantes fourragères telles que les betteraves, les choux, le ray-grass d'Italie ;
- les cultures fourragères dérobés, où l'on retrouve le ray-grass d'Italie, aux côtés d'autres plantes de plus courte durée.

Une telle typologie n'est pas neutre. N'oublions pas que l'objectif est d'établir entre ces types de production et d'autres données des correspondances qui vont préparer l'explication de leur existence. Ces autres données doivent également avoir fait l'objet d'une typologie. Citons un exemple qui montrera les difficultés de cette opération : celui des herbages pâturés ou fauchés est typique. Presque toutes les surfaces couchées en plantes fourragères pérennes font l'objet d'une double exploitation, les unes destinées à la constitution de réserves, les autres à la consommation immédiate ; il faut donc les distinguer par des fréquences ou des pourcentages d'utilisation. Mais suivant les subdivisions qui seront adoptées, on risque de faire disparaître des différences qui seraient très caractéristiques du système d'exploitation utilisé. Les typologies concernant les données caractérisant l'utilisation ou le rôle des prairies doivent donc s'inspirer d'un certain nombre d'a priori qui sont basés sur la connaissance que l'on a ou que l'on croit avoir des situations existantes : il y a donc une axiomatique derrière ces choix, il s'agit d'hypothèque sur les conclusions qui se dégageront des comparaisons. Néanmoins, celles-ci pourront faire l'objet d'une étude critique en les comparant à d'autres conclusions tirées de la confrontation de données indépendantes. Pour citer un autre

élevé, cette information peut être comparée aux productions potentielles ; si celles-ci ne semblent pas correspondre aux besoins du troupeau, alors l'exploitant doit avoir recours à des sources de fourrage extérieures à l'exploitation. En définitive, la présentation des données concernant les productions primaires et secondaires et la vérification de leur cohérence imposent que l'on s'appuie sur des modèles, dans ce cas, des bilans : ceux-ci doivent être simples mais de ce fait, s'ils sont plus robustes, ils sont nécessairement moins précis. Cette remarque implique bien entendu la confrontation de plusieurs ensembles de données : c'est ce que l'on appelle croiser les données. Cette opération peut être utilisée à deux fins : vérifier l'objectivité des données, établir des relations ayant une valeur explicative. Mais celles-ci sont indispensables, car, comme le dit le statisticien Michel VALLE dans son livre « Le métier de statisticien » (Hachette 1980 p. 207) : « La statistique n'a d'influence possible sur l'action, sur la pratique, qu'après avoir transité par une médiation théorique qui peut être parfois implicite mais est toujours présente ».

C'est justement pour assurer cette médiation que les données de la statistique doivent faire l'objet d'un double contrôle : a priori, lors du choix des axiomes ; a posteriori, lorsque les données recueillies ont été rassemblées. On retrouve bien les idées formulées par BOWLEY et rappelées au début de cet article.

Il faut maintenant reprendre la démarche suivie pour envisager le parti que l'on peut en tirer. Dès le début, on doit noter que la collecte des informations n'a pas porté sur un dénombrement exhaustif, c'est-à-dire sur toutes les exploitations existantes, mais sur un échantillon. Mais celui-ci est de bonne dimension, les enquêteurs ayant travaillé sur 20 000 exploitations réparties dans 82 départements, ceux de la PARIS et de son voisinage, et ceux situés en bordure de la Méditerranée étant exclus pour des raisons évidentes. Les enquêteurs disposaient également des données du Recensement Général de l'Agriculture, plus exhaustives mais moins précises. Il est essentiel dans une opération de ce genre que les exploitations retenues pour les enquêtes soient représentatives de l'ensemble des exploitations situées dans la même région. C'est le principe même de tous les sondages et les prévisions des résultats des élections montrent que ces opérations peuvent être très précises,

c'est-à-dire donner des résultats très voisins des données globales obtenues en fin d'opération, donc sur un tirage exhaustif. Cette représentativité de l'échantillonnage repose sur une stratification de la réalité. Cela veut dire que l'on choisit des exploitations en fonction des classes de dimension, par exemple : une fois le recensement sur échantillon terminé, les résultats peuvent être ramenés au total par une simple proportionalité, le nombre total des exploitations étant celui existant dans l'unité territoriale où s'est effectué l'échantillonnage.

Cette démarche est très importante, car si l'on fait un recensement en partant d'un découpage administratif, il existe une forte probabilité pour que les données recueillies, qu'elles soient moyennes ou extrêmes, atténuent ou exagèrent la réalité. Par exemple, si l'on considère le département de Saône-et-Loire, l'activité agricole de la zone Ouest est sensiblement différente de celle que l'on peut constater sur la partie Est. Or les comparaisons, c'est-à-dire les croisements, que l'on réalise peuvent se faire exploitation par exploitation, par exemple : surface toujours en herbe et charge de bétail, ou par groupe d'exploitations, défini sur certains critères, c'est-à-dire strate par strate. Ces choix dépendent du type de réponse que l'on désire obtenir et des moyens dont on dispose. Mais ces comparaisons se font par paire de critères, ce qui veut dire toutes choses égales par ailleurs, ce qui est rarement le cas. Si de nos jours les moyens de calcul permettent de prendre en compte de nombreuses données, il n'en reste pas moins que le volume de résultats obtenus repose aux interprétateurs des difficultés parfois insurmontables.

Dans ce cas encore il faut faire des choix, c'est-à-dire se demander quels types de réponse on veut obtenir et ces désirs sont déterminés par l'usage que l'on veut en faire.

Schématiquement, il existe deux attitudes :

- savoir ce qui est, pour le comparer soit à ce que l'on a espéré dans le passé, soit à ce qui existe ailleurs ;*
- prévoir ce qui peut arriver et, dans la mesure du possible, orienter le futur dans un sens qui paraît souhaitable.*

Les résultats de l'enquête répondent bien à certaines questions concernant les surfaces fourragères, leur nature, les conditions d'exploitation, etc. Il peut être intéressant de comparer certaines constatations aux impressions que l'on aurait pu déga-

ger en considérant les témoignages des conseillers agricoles ou des groupes de gestion. En effet, ces informations sont souvent biaisées car les échantillons sur lesquels elles portent ne sont pas représentatifs ; il s'agit d'agriculteurs dynamiques ou parfois, très rarement, de retardataires. Or le problème consiste à faire apparaître les situations correspondant aux « modes » c'est-à-dire, statistiquement parlant, aux cas les plus fréquents ou aux situations médianes partageant l'ensemble en deux fractions égales.

Ces situations peuvent alors être comparées aux situations souhaitables, celles où les données de la recherche ont été prises en compte et les techniques de pointe appliquées. Les divers articles figurant dans ce numéro évoquent plusieurs aspects de ces problèmes.

Mais en plus de la vérification de la cohérence des informations, jugement qui repose sur l'existence de correspondances rationnelles entre les diverses données recueillies, il y a la possibilité de comparer ces informations à celles existant dans les pays voisins. L'A.F.P.F. avait organisé en 1977 un Colloque pour tenter de confronter les situations et les divers points de vue des experts. Une plaquette résumant les enseignements du Colloque avait été rédigée par un petit groupe de participants français. Intellectuellement, le résultat était très satisfaisant ; pratiquement, les conclusions pouvaient paraître plus discutables parce que les données de base n'avaient pas un caractère systématiquement exhaustif. Ce n'est certes pas le cas des informations reproduites dans ce numéro 100 qui pourra ainsi servir de référence, en particulier dans les discussions concernant les décisions économiques. A une époque où l'on envisage de restreindre les productions, les mécanismes technico-politiques qui seront élaborés pour atteindre ce but devront tenir compte de réalités qui intéressent les individus, mais aussi la collectivité nationale en particulier à travers ses activités économiques. Seule une bonne statistique, c'est-à-dire dont les références géographiques et les données techniques auront été judicieusement choisies, pourra servir de base aux tentatives visant à maîtriser les productions d'une manière raisonnée.

Mais si une bonne enquête fournit une image fidèle du présent et permet d'évaluer le résultat des efforts du passé, c'est aussi un point de départ pour l'avenir. La prospective résultera de la comparaison de l'existant et du possible. La vérification

de ces interprétations et la maîtrise du futur exigent que de telles enquêtes soient effectuées périodiquement sur les mêmes bases.

En effet, tout programme de recherche et encore plus de développement doit partir d'une situation initiale et se fixer un état final à atteindre. Or, au départ, l'état pourra être défini en s'appuyant sur les informations recueillies. Compte tenu de la diversité des situations pédoclimatiques existant en France, ces opérations devront avoir un cadre régional. Ce sont évidemment les comparaisons de région à région ou, à l'intérieur d'une même zone, les différences existant entre les résultats les plus élevés, les moyens et les plus faibles qui amèneront à en rechercher les causes et les moyens de les modifier. L'existence dans cette enquête de la référence « potentialité » illustre bien l'intérêt de cette démarche.

En définitive, cette enquête a une valeur considérable en raison de l'importance de l'échantillon étudié et du fait qu'il soit représentatif. Rappelons que toute source d'information reposant sur des groupes s'étant fixé un objectif donne nécessairement des résultats biaisés. Au contraire, les données recueillies par cette enquête correspondent à une typologie clairement exposée, appliquée systématiquement à un vaste échantillon tiré au hasard. Néanmoins, ces données ainsi que leur présentation plus ou moins synthétique reposent sur des choix dont tous les utilisateurs doivent peser les conséquences.

Mais quelles que soient la qualité des données immédiates obtenues par ce travail et l'originalité des conclusions auxquelles pourront conduire leur comparaison, cette enquête n'aura sa pleine valeur que si l'on veut s'en servir car, en définitive, « la statistique n'est qu'une préparation à l'action ».

Stéphane HÉNIN.